אוניברסיטת תל אביב

הסתברות למהנדסים

יבגני צודיקוביץ'

12 בפברואר 2019

אין לשכפל, להעתיק, לצלם, לתרגם, לאחסן במאגר מידע, לשדר או להקליט בכל דרך או בכל אמצעי אלקטרוני, אופטי, מכני, או אחר כל חלק מהחומר בספר זה. שימוש מסחרי מכל סוג בחומר הכלול בספר זה אסור בהחלט, אלא ברשות מפורשת בכתב מהמחבר.

גירסא 1 - בהרצה

הקדמה

חוברת זו החלה את דרכה כרשימות מרצה עבור הקורס "הסתברות להנדסת ביו-רפואה" אך ככל שהתקדמה הכתיבה טפחה החוברת והפכה לסיכום של כל ההרצאות, התרגולים, התרגילים והמבחנים אותם העברתי במהלך השנים למסלולי ההנדסה השונים. לפיכך, לא רק היקף החומר השתנה במהלך הכתיבה אלא גם אופי הכתיבה וכעת החוברת אינה כתובה כרשימות לצורך הוראה בכיתה אלא כספר המאפשר לימוד עצמי וכולל, בנוסף להסברים, דוגמאות ותרגילים - גם פתרונות שגויים נפוצים של סטודנטים אותם אספתי לאורך השנים. כך, במידה וגם אתם חזרתם על אותה טעות תוכלו להבין מיד מה לא היה בסדר בה ובכל מקרה תוכלו ללמוד מטעויות של אחרים.

אחד המפתחות החשובים להצלחה בקורס הוא תרגול ושוב תרגול. אין זה מהקורסים אותם ניתן לעבור עם ידע מעורפל בחומר ודף נוסחאות טוב. חשוב להקדיש זמן רב לפתרון התרגילים המופיעים בחוברת באופן עצמאי, מבלי להציץ בפתרונות או "להבין את הפתרונות" כתחליף לפתרון העצמאי. רק באופן זה יכולה להיווצר ההבנה הדרושה ולהתפתח האינטואיציה ההסתברותית הנכונה. לפיכך – בסופו של כל פרק תמצאו מקבץ גדול של שאלות לתרגול. פיתרו אותן! בנוסף, שימו לב שהפרק הראשון, פרק 0, אינו חלק רשמי מהקורס וכולל סיכום קצר של טכניקות מתמטיות אותם למדתם בקורסים קודמים וקורס זה מתבסס עליהן, דוגמת אינטגרציה וחישוב טורים. החומר ההסתברותי מתחיל בפרק 1, העוסק בלוגיקה ותורת הקבוצות. ההנחה היא שלוגיקה אתם יודעים (מהחיים...) ותורת הקבוצות כבר למדתם, אך היות וחומר זה הינו בסיסי בתורת ההסתברות וכן היות והמונחים בהם אנו נשתמש שונים מהמונחים המוכרים מתורת הקבוצות, בחרתי להציג את הנושא במלואו.

מאחר והמטרה העיקרית בקורס היא להקנות ידע מספיק בהסתברות בכדי לבצע חישובים הסתברותיים ככלי עזר בקורסים אחרים דוגמת סטטיסטיקה או אותות אקראיים ורעש, לעיתים קרובות ויתרתי על הדיוק המתמטי לטובת הקריאות והפישוט של החומר. ייתכן ותיתקלו בגירסאות אחרות של תוצאות המופיעות כאן במקומות אחרים תחת הנחות אחרות או עם סייגים לגבי תקפות התוצאה. מאחר ובפרקטיקה סייגים אלו נדירים להחריד ויתרתי על ציונם המפורש על חשבון הכשרות של חלק מהמעברים והחישובים. עמכם הסליחה.

לסיום, חוברת זו מבוססת על המרצים והמתרגלים איתם עבדתי או שלימדו באוניברסיטת תל אביב לפני תקופתי ועל רשימותיהם התבססתי - ד"ר דוד לגזיאל, ד"ר גלית גולן-אשכנזי, ד"ר אילה יעקובי-משיח, ד"ר אופיר הררי פרופ' יובל הלר, פרופ' אילון סולן וד"ר שלומי רובינשטיין. תודה לכם!

בהצלחה בלימוד,

יבגני צודיקוביץ'

תוכן העניינים

0 רקע מתמטי

פרק זה כולל את הרקע הבסיסי באנליזה מתמטית הנדרש לצורך לימוד תורת ההסתברות, דוגמת אינטגרציה וחישוב סכומים של טורים. חומר זה נלמד לרוב בקורסים המהווים קורסי קדם לקורס "תורת ההסתברות" ולכן פרק זה אינו מיועד ללמד את החומר אלא רק להזכיר את המושגים הבסיסיים והטכניקות הבסיסיות ולייצר נקודת פתיחה אחידה לתחילת הלימוד. מומלץ לרפרף על ראשי הפרקים של פרק זה, בכדי לוודא שכל המושגים נהירים ובמידה ולא – לרענן את הידע בנושאים החסרים.

היות והמטרה המרכזית של ספר זה היא לימוד החשיבה ההסתברותית וההבנה של תורת ההסתברות, אנו נימנע מפתרון תרגילים טכניים מדי הדורשים חישובים מסובכים מאוד. תרגילים כאלו לרוב לא תורמים להבנה של החומר ומבזבזים את זמנו של הסטודנט על טכניקה לא רלוונטית וחישובים ארוכים שלא תורמים להבנה ההסתברותית. לכן, התרגילים שיופיעו בפרק זה ובפרקים הבאים שיתבססו על חומר זה יהיו יחסית פשוטים מבחינה טכנית. כך למשל, מרבית האינטגרלים שנפגוש יהיו אינטגרלים פשוטים של פולינומים או פונקציות אלמנטריות אחרות במקום פונקציות מסובכות יותר העשויות להופיע בחיים האמיתיים.

בנוסף, על מנת לא לקטוע את רצף הקריאה ואת ההסברים ההסתברותיים בפרקים הבאים, חלק מהחישובים הטכניים הועברו לפרק זה. בכך, פרק זה יינקז לתוכו חישובים רבים ומייגעים שאינם רלוונטיים והפרקים האחרים יוכלו להתמקד יותר בהסתברות ופחות בטכניקה מתמטית.

0.1 עצרת, מקדמים בינומיים ובינום ניוטון

הגדרה 0.1 (עצרת)**.** לכל n טבעי נגדיר בתור $n!$ (קרי: n עצרת) את מכפלת המספרים הטבעיים מ-1 ועד n:

$$n! = \prod_{k=1}^{n} k = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1 \tag{0.1}$$

בנוסף, מגדירים $0! = 1$.
ניתן להגדיר עצרת גם באופן רקורסיבי, כלומר להגדיר $0! = 1$ ולכל $n > 0$: $n! = n \cdot (n-1)!$.

עבור ערכים גדולים של n ניתן להשתמש בקירוב סטרלינג:

$$n! \simeq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \tag{0.2}$$

בכדי לקבל נוסחא לפונקציית העצרת שהיא יותר נוחה לעבודה.

הגדרה 0.2 (מקדמים בינומיים)**.** לכל מספר טבעי n ולכל $0 \le k \le n$ טבעי מגדירים את המקדם הבינומי $\binom{n}{k}$ על ידי

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{0.3}$$

המקדם הבינומי $\binom{n}{k}$ נקרא גם "n מעל k" או "n בחר k" (ובאנגלית n choose k), כאשר הקשר בינו לבין "בחירה" יובהר בפרק 3.1.3.

לרוב נשאיר את מלאכת החישוב של המקדמים הבינומיים למחשבון, אך כדאי לשים לב לערכים של מקרי הקצה:

$$\binom{n}{0} = \binom{n}{n} = 1 \tag{0.4}$$

וכן

$$\binom{n}{1} = \binom{n}{n-1} = n \tag{0.5}$$

בנוסף, ניתן לראות על ידי הצבה ישירה שהזהות

$$\binom{n}{k} = \binom{n}{n-k} \tag{0.6}$$

מתקיימת תמיד.

בנוסף, כמוסכמה, נגדיר את המקדם הבינומי להיות אפס עבור כל צירוף אחר, כלומר $\binom{n}{k} = 0$ כאשר $k \notin \{0, \ldots, n\}$.

תרגיל 0.1. הוכיחו את הטענות הבאות:

1. לכל $n \in \mathbb{N}$ ולכל $n \geq k \in \mathbb{N}$ מתקיים

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \tag{0.7}$$

2. לכל $1 \leq k \leq \frac{n}{2}$ שלם מתקיים

$$\binom{n}{k} > \binom{n}{k-1} \tag{0.8}$$

3. לכל $0 \leq i \leq n$ מתקיים

$$\sum_{k=0}^{n} \binom{k}{i} = \binom{n+1}{i+1} \tag{0.9}$$

רמז: אינדוקציה.

***משפט* 0.1** (בינום ניוטון)**.** *לכל x,y ממשיים ולכל n טבעי מתקיים:*

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^ky^{n-k} \tag{0.10}$$

עבור $n=2$ נוסחא זו היא נוסחת הכפל המקוצר המוכרת מהתיכון: $(x+y)^2=x^2+2xy+y^2$.

ניתן להוכיח נוסחא זו על ידי אינדוקציה תוך שימוש בזהויות מתרגיל 0.1. בנוסף, קיימת הוכחה לנוסחת הבינום המתבססת על קומבינטוריקה אותה נראה בתרגיל 3.3.

תרגיל 0.2. הוכיחו את הזהויות הבאות:

1. לכל n טבעי מתקיים

$$2^n=\sum_{k=0}^{n}\binom{n}{k} \tag{0.11}$$

2. לכל n טבעי מתקיים

$$\binom{n}{0}+\binom{n}{2}+\ldots=\binom{n}{1}+\binom{n}{3}+\ldots \tag{0.12}$$

הערה: לא מדובר על סכימה אינסופית אלא על סכימה עד המספר הזוגי (באגף שמאל) או האי-זוגי (באגף ימין) הגדול ביותר הקטן או שווה ל-n, למרות שממילא עבור ערכים החורגים מ-n המקדם הבינומי מתאפס.

3. לכל n טבעי ולכל $0\le p\le 1$ מתקיים

$$np=\sum_{k=0}^{n}k\binom{n}{k}p^k\left(1-p\right)^{n-k} \tag{0.13}$$

תרגיל 0.3. יהיו $N,D,n\in\mathbb{N}$ כך ש-$N>D,n$. הוכיחו

$$\binom{N}{n}=\sum_{k=0}^{n}\binom{D}{k}\cdot\binom{N-D}{n-k}$$

רמז: הכפילו שני בינומי ניוטון עם אותם x,y כך שסכום החזקות יהיה N.

0.2 סכימה וטורים

***משפט* 0.2** (טור גאומטרי)**.** *עבור $0<p<1$ מתקיים*

$$\sum_{k=0}^{n}p^k=\frac{1-p^{n+1}}{1-p} \tag{0.14}$$

ובמקרה שמספר המחוברים הוא אינסופי, נקבל

$$\sum_{k=0}^{\infty} p^k = \frac{1}{1-p} \tag{0.15}$$

הטורים הגאומטריים שאנו נפגוש יהיו בדרך כלל מהצורה $p\,(1-p)^{k-1}$ *ועל ידי הצבה בנוסחא הקודמת נקבל*

$$\sum_{k=1}^{\infty} (1-p)^{k-1} p = 1 \tag{0.16}$$

הערה: *נוסחא 0.14 נכונה לכל* $p \neq 1$.

תרגיל 0.4. הוכיחו כי מתקיים

$$\sum_{k=1}^{\infty} k \cdot p\,(1-p)^{k-1} = \frac{1}{p} \tag{0.17}$$

רמז: $\frac{d}{dp}\left((1-p)^k\right) = -k\,(1-p)^{k-1}$

משפט* 0.3** (טור טיילור של הפונקציה e^x)**.** ***עבור כל $x > 0$ *מתקיים:*

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{0.18}$$

ולכן

$$\sum_{n=0}^{\infty} e^{-x} \frac{x^n}{n!} = 1 \tag{0.19}$$

טורים נוספים שכדאי להכיר (ללא הוכחה):

- סכום המספרים הטבעיים:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \tag{0.20}$$

- סכום החזקות הריבועיות:

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \tag{0.21}$$

- סכום החזקות השלישיות:

$$\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4} \tag{0.22}$$

- סכום הופכיי החזקות הריבועיות:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} \tag{0.23}$$

- הטור ההרמוני מתבדר:

$$\sum_{k=1}^{\infty} \frac{1}{k} = \infty \tag{0.24}$$

0.3 אינטגרציה

0.3.1 אינטגרלים בסיסיים

לפניכם רשימת האינטגרלים הלא-מסויימים של הפונקציות האלמנטריות העיקריות (הנוסחאות מגיעות מכך שאנחנו יודעים לגזור את אגף ימין ומתקבל האינטגרנד[1] באגף שמאל):

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} \tag{0.25}$$

$$\int \frac{dx}{x} = \ln x \tag{0.26}$$

$$\int e^x\,dx = e^x \tag{0.27}$$

עבור $a > 0$ מתקיים:

$$\int a^x\,dx = \frac{1}{\ln a} a^x \tag{0.28}$$

[1] האינטגרנד היא הפונקציה עליה מבצעים את האינטגרל.

$$\int \sin x \, dx = -\cos x \tag{0.29}$$

$$\int \cos x \, dx = \sin x \tag{0.30}$$

במידה והאינטגרנד מורכב יותר, ניתן להיעזר בתכונות הבאות בכדי לפשט אותו:

- ליניאריות:

$$\int \alpha \cdot f(x) + \beta \cdot g(x) \, dx = \alpha \int f(x) \, dx + \beta \int g(x) \, dx \tag{0.31}$$

- נגזרת של פונקציה:

$$\int f(x) \cdot f'(x) \, dx = \frac{1}{2} f^2(x) \tag{0.32}$$

- כלל ההצבה: אם $x = g(y)$ אז $dx = g'(y)dy$ ועל ידי ההצבה של ביטויים אלו באינטגרל ניתן לעיתים להגיע לאינטגרל פשוט יותר. חשוב לשים לב שגם גבולות האינטגרציה משתנים לאחר ההצבה.
- בקורסים העוסקים בחשבון דיפרנציאלי ואינטגרלי לומדים שיטות אינטגרציה נוספות ומתקדמות אשר לא נזדקק להן שכן הן מוסיפות סיבוכיות טכנית לשאלות אך לא מוסיפות תובנות הסתברותיות.

בעיה 0.1. יהיו $\lambda, t > 0$ קבועים. חשבו את האינטגרל $\int_0^t \lambda e^{-\lambda x} \, dx$.

הוכחה. נציב $x = -\frac{y}{\lambda}$. לפיכך $dx = -\frac{dy}{\lambda}$ וגבולות האינטגרציה של המשתנה החדש יהיו מ-$y(0) = 0$ ועד $y(t) = -\lambda t$:

$$\int_0^t \lambda e^{-\lambda x} \, dx = \int_0^{-\lambda t} -e^y \, dy = -e^y \big|_0^{-\lambda t} = 1 - e^{-\lambda t} \tag{0.33}$$

ולכן כאשר הגבול העליון הוא ∞ נקבל

$$\int_0^\infty \lambda e^{-\lambda x} \, dx = 1 - \lim_{t \to \infty} e^{-\lambda t} = 1 \tag{0.34}$$

□

משפט 0.4* (אינטגרציה בחלקים). *אם $F'(x) = f(x)$ ***וגם*** $G'(x) = g(x)$ ***אז***

$$\int_a^b f(x)G(x) \, dx = F(x)G(x)\big|_a^b - \int_a^b F(x)g(x) \, dx \tag{0.35}$$

ניתן להשתמש בנוסחת האינטגרציה בחלקים כאשר האינטגרל $\int_a^b F(x)g(x)\,dx$ יותר פשוט לחישוב מאשר $\int_a^b f(x)G(x)\,dx$.

תרגיל 0.5. חשבו באמצעות אינטגרציה בחלקים את הפונקציה הקדומה של $\ln x$.

תרגיל 0.6. הראו כי לכל $n \geq 0$ שלם מתקיים

$$\int_0^\infty x^n e^{-x}\,dx = n! \tag{0.36}$$

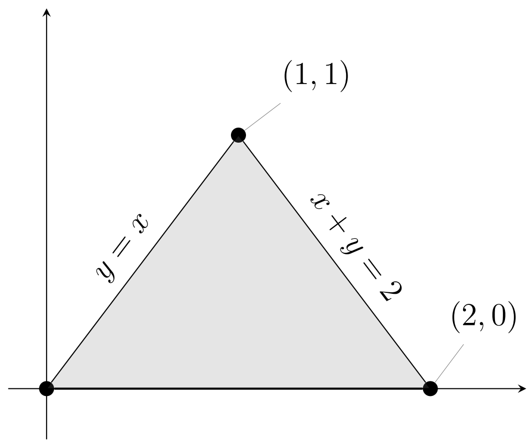
רמז: הגדירו $I_n = \int_0^\infty x^n e^{-x}\,dx$ ומצאו נוסחת נסיגה ותנאי התחלה עבור הסדרה I_n באמצעות אינטגרציה בחלקים.

0.3.2 אינטגרלים דו-ממדיים

באינטגרל דו-ממדי האינטגרנד הוא פונקציה של שני משתנים ותחום האינטגרציה הוא תחום כלשהו במישור XY. במקרה כזה מבצעים שני אינטגרלים, אחד לכל משתנה, כאשר גבולות האינטגרציה של משתנה אחד יכולים להיות תלויים באחר. גבולות האינטגרציה של האינטגרל החיצוני (המבוצע שני) הם קבועים הנקבעים לפי הערך המינימלי והמקסימלי שמשתנה האינטגרציה יכול לקבל ואילו גבולות האינטגרציה של האינטגרל הפנימי (המבוצע ראשון) יכולים להיות תלויים במשתנה השני בהתאם לצורת התחום. בכדי לקבוע בצורה נכונה את גבולות האינטגרציה, מומלץ ראשית לשרטט את תחום האינטגרציה ולקבוע על סמך השרטוט מהן הגבולות של האינטגרל הפנימי וכן על איזה משתנה לבצע את האינטגרל הראשון, כפי שיומחש להלן.

בעיה 0.2. חשבו את האינטגרל $\iint_D xy\,dxdy$ כאשר D הוא המשולש שקודקודיו ב- $(0,0),(2,0),(1,1)$.

הוכחה. באופן כללי, המשתנה x יכול לקבל את הערכים 0 עד 2 ואילו המשתנה y מקבל את הערכים 0 עד 1, אך יש תלות בין המשתנים, כמוראה בתרשים:



כך למשל, לכל ערך של $0 \le y \le 1$, מתקיים בהכרח $y \le x \le 2-y$. לפיכך נבצע את האינטגרל החיצוני לפי y בגבולות 0 עד 1 ואילו את האינטגרל הפנימי לפי x כאשר הגבולות תלויים בערך של y: הגבול התחתון של x הינו y והגבול העליון הינו $2-y$:

$$\begin{aligned}\iint_D xy\,dxdy &= \int_0^1 \int_y^{2-y} xy\,dxdy \\ &= \int_0^1 y \cdot \left.\frac{x^2}{2}\right|_y^{2-y} dy = \frac{1}{2}\int_0^1 y\left((2-y)^2 - y^2\right)\,dy \\ &= \int_0^1 2y - 2y^2\,dy = 2\left(\frac{y^2}{2} - \frac{y^3}{3}\right)_0^1 = \frac{1}{3}\end{aligned}$$

בדומה, ניתן לבצע קודם את האינטגרל על y ורק לאחר מכן את האינטגרל על x מ-0 ועד 2. במקרה כזה, הגבול התחתון של המשתנה y הוא תמיד 0 אך הגבול העליון תלוי בערך של x: כאשר $0 \le x \le 1$, הגבול העליון הינו x ואילו כאשר $1 \le x \le 2$ הגבול העליון הינו $2-x$. לפיכך, יש לפצל את האינטגרל בהתאם לתחומים אלו:

$$\begin{aligned}\iint_D xy\,dydx &= \int_0^1 \int_0^x xy\,dydx + \int_1^2 \int_0^{2-x} xy\,dydx \\ &= \int_0^1 x\left.\frac{y^2}{2}\right|_0^x dx + \int_1^2 x\left.\frac{y^2}{2}\right|_0^{2-x} dx \\ &= \int_0^1 \frac{x^3}{2}\,dx + \int_1^2 \frac{x^3 - 4x^2 + 4x}{2}\,dx \\ &= \left(\frac{x^4}{8}\right)_0^1 + \left(\frac{x^4}{8} - \frac{2x^3}{3} + x^2\right)_1^2 = \frac{1}{3}\end{aligned}$$

□

באופן טבעי, סדר האינטגרציה לא השפיע על התשובה הסופית, אך הפתרון הראשון פשוט יותר שכן גבולות האינטגרציה במקרה זה נקבעים פעם אחת ואין צורך לפצל למקרים ולבצע חישובים עודפים. לכן מומלץ תמיד להתחיל משרטוט תחום האינטגרציה ובהתאם לשרטוט לקבוע הן את גבולות האינטגרציה והן את סדר האינטגרציה כך שהאינטגרלים יהיו פשוטים והחישוב יהיה נוח ככל הניתן.

בנוסף, כאשר לאינטגרל יש סימטריה מעגלית, נוח לעבוד בקואורדינטות קוטביות (פולריות) בכדי לפשט את האינטגרל (לפחות בחלק מהמקרים). המעבר לקואורדינטות קוטביות מתבצע על ידי ההצבה:

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \\ dxdy = r\,drd\theta \end{cases} \tag{0.37}$$

וכמובן שיש לשנות את גבולות האינטגרציה בהתאם.

תרגיל 0.7. חשבו את האינטגרלים הבאים:

1. $\iint_{D_1} \lambda_1\lambda_2 e^{-\lambda_1 x-\lambda_2 y}\, dxdy$
כאשר D_1 הינו התחום בו $0 \le x \le y$ ו- λ_1, λ_2 קבועים חיוביים.
2. $\iint_{D_2} c\, dxdy$
כאשר D_2 הינו התחום ברביע הראשון המקיים $x^2+y^2 \le 1$ ו-c קבוע חיובי.
3. $\iint_{D_3} e^{-x}\, dxdy$
כאשר D_3 הינו התחום בו מתקיים $1 \le y \le 2$ וגם $x \ge 0$.

בהמשך לסעיף 2 של תרגיל 0.7, אינטגרל על קבוע שווה לשטח עליו מבוצע האינטגרל. נסכם תוצאה זו במשפט הבא:

משפט 0.5. *אינטגרל על קבוע של פונקציה על פני קטע שווה לאורך הקטע כפול הקבוע: לכל $a, b \in \mathbb{R}$ כך ש-$a < b$ ולכל $c \in \mathbb{R}$ מתקיים*

$$\int_a^b c\, dx = (b-a)\, c \tag{0.38}$$

באופן דומה, אינטגרל על קבוע של פונקציה על פני תחום דו-ממדי שווה לשטח התחום כפול הקבוע: לכל תחום דו-ממדי חסום $D \subset \mathbb{R}^2$ ולכל $c \in \mathbb{R}$ מתקיים

$$\iint_D c\, dxdy = S_D$$

כאשר S_D הוא שטח התחום D.

0.4 פתרונות לשאלות הפרק

תרגיל 0.1. 1. נחשב מפורשות את אגף ימין:

$$\begin{aligned}\binom{n-1}{k-1}+\binom{n-1}{k} &= \frac{(n-1)!}{(k-1)!(n-k)!}+\frac{(n-1)!}{k!(n-1-k)!}\\ &= \frac{k\cdot(n-1)!}{k!(n-k)!}+\frac{(n-k)\cdot(n-1)!}{k!(n-k)!}\\ &= \frac{(k+n-k)\cdot(n-1)!}{k!(n-k)!}=\frac{n!}{k!(n-k)!}=\binom{n}{k}\end{aligned}$$

כנדרש.

2. ההוכחה באינדוקציה על n. ראשית, הטענה נכונה באופן מיידי עבור $n=0,1,2$. אחרת, עבור $n\geq 2, k>1$, ניתן לרשום לפי הסעיף הקודם:

$$\begin{aligned}\binom{n}{k} &= \binom{n-1}{k-1}+\binom{n-1}{k}\\ \binom{n}{k-1} &= \binom{n-1}{k-2}+\binom{n-1}{k-1}\end{aligned}$$

לפי הנחת האינדוקציה, כל אחד מהמחוברים באגף ימין של המשוואה העליונה גדול מהמחובר המתאים לו במשוואה התחתונה ולכן גם $\binom{n}{k}>\binom{n}{k-1}$. בנוסף, באופן מיידי הטענה מתקיימת גם עבור $k=1$ ובכך ההוכחה הושלמה.

3. נוכיח באינדוקציה. הטענה נכונה עבור $n=0,1$ באופן מיידי. נניח שעבור n טבעי כלשהו מתקיים: $\sum_{k=0}^{n}\binom{k}{i}=\binom{n+1}{i+1}$ ולכן עבור $n+1$ נקבל

$$\begin{aligned}\sum_{k=0}^{n+1}\binom{k}{i} &= \binom{n+1}{i}+\sum_{k=0}^{n}\binom{k}{i}\\ \text{(*)} &= \binom{n+1}{i}+\binom{n+1}{i+1}\overset{eq.\,0.7}{=}\binom{n+2}{i+1}\end{aligned}$$

ובכך הושלמה ההוכחה, כאשר המעבר * מתקבל מהנחת האינדוקציה.

□

תרגיל 0.2. 1. ההוכחה מיידית ומתקבלת על ידי ההצבה $x=y=1$ בבינום ניוטון.

2. נציב $x=-1, y=1$ בנוסחת הבינום ונקבל

$$0=\sum_{k=0}^{n}\binom{n}{k}(-1)^k$$

לפיכך, מקדמים בינומיים עם k זוגי יופיעו עם סימן פלוס ומקדמים בינומיים עם k אי זוגי עם סימן מינוס. נעביר את החזקות האי זוגיות אגף ונקבל את השוויון המבוקש.

3. נחשב מפורשות את אגף ימין:

$$\begin{aligned}\sum_{k=0}^{n} k\binom{n}{k}p^{k}\left(1-p\right)^{n-k} &= np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}\left(1-p\right)^{n-k}\\ &= np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}\left(1-p\right)^{n-k}\\ (k\text{`}=k-1) &= np\sum_{k\text{`}=0}^{n-1}\binom{n-1}{k\text{`}}p^{k\text{`}}\left(1-p\right)^{n-1-k\text{`}}\\ (eq.\,0.10) &= np\left(p+1-p\right)^{n-1}=np\end{aligned}$$

$\square$

תרגיל 0.3. יהיו $x,y\in\mathbb{R}$ ונרשום את $(x+y)^{N}$ ככפל של שני בינומי ניוטון:

$$\left(x+y\right)^{N}=\left(x+y\right)^{D}\left(x+y\right)^{N-D}=\left(\sum_{k=0}^{D}\binom{D}{k}x^{k}y^{D-k}\right)\cdot\left(\sum_{j=0}^{N-D}\binom{N-D}{j}x^{j}y^{N-D-j}\right)$$
$$=\sum_{k=0}^{D}\sum_{j=0}^{N-D}\binom{D}{k}\cdot\binom{N-D}{j}x^{k+j}y^{N-k-j}$$

נסמן $n-k=j$ ונשנה את סדר הסכימה כך שנבצע את הסכימה החיצונית על n ואת הסכימה הפנימית על k. בנוסף, נשים לב כי ניתן להרחיב את גבולות האינדקסים כרצוננו, שכן המקדמים הבינומיים מתאפסים עבור כל הערכים שאינם בטווח. נקבל את הביטוי:

$$=\sum_{n=0}^{N}\left[\sum_{k=0}^{D}\binom{D}{k}\cdot\binom{N-D}{n-k}\right]x^{n}y^{N-n}$$

סכום זה הוא בדיוק בינום ניוטון שהיה מתקבל לו פתחנו את $(x+y)^{N}$ישירות ולכן הסוגריים המרובעות חייבות להיות המקדם הבינומי המתאים, כלומר

$$\sum_{k=0}^{D}\binom{D}{k}\cdot\binom{N-D}{n-k}=\binom{N}{n}$$

כנדרש. $\square$

תרגיל 0.4. נשתמש ברמז ונחליף את הסדר של הגזירה והסכימה:

$$\begin{aligned}\sum_{k=1}^{\infty} kp\,(1-p)^{k-1} &= p\sum_{k=1}^{\infty} -\frac{d}{dp}\left((1-p)^k\right) \\ &= p\frac{d}{dp}\left(\sum_{k=1}^{\infty} -(1-p)^k\right) \\ &= p\frac{d}{dp}\left(-\frac{1}{p}\right) = p\cdot\frac{1}{p^2} = \frac{1}{p}\end{aligned}$$

$\square$

תרגיל 0.5. נגדיר $f(x)=1$ וכן $G(x)=\ln x$. מתקיים $F(x)=x$ וכן $g(x)=G'(x)=\frac{1}{x}$ ולכן

$$\begin{aligned}\int \ln x\,dx &= \int f(x)G(x)\,dx = F(x)G(x) - \int F(x)g(x)\,dx \\ &= x\cdot\ln(x) - \int x\cdot\frac{1}{x}\,dx = x\ln(x) - x\end{aligned}$$

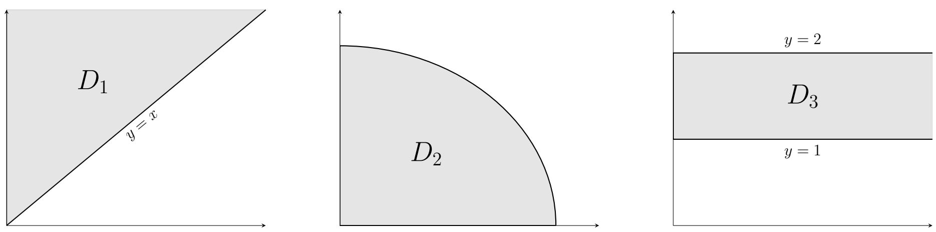
$\square$

תרגיל 0.6. נסמן $I_n=\int\limits_0^\infty x^n e^{-x}\,dx$ ונחשב את האינטגרל עבור $n>0$ באמצעות אינטגרציה בחלקים:

$$I_n = \int\limits_0^\infty x^n e^{-x}\,dx = \cancelto{0}{x^n e^{-x}\Big|_0^\infty} + n\int\limits_0^\infty x^{n-1}e^{-x}\,dx = nI_{n-1}$$

עבור $n=0$ זהו פשוט אינטגרל של אקספוננט ומתקיים $I_0=1$. לפיכך הסדרה I_n מקיימת $I_0=1$ ולכל $n>0$ $I_n=nI_{n-1}$ וזו בדיוק נוסחת הנסיגה של פונקציית העצרת, כלומר $I_n=n!$. $\square$

תרגיל 0.7. עבור כל אחד מהאינטגרלים נשרטט ראשית את תחום האינטגרציה ובהתאם לכך נקבע את סדר האינטגרציה וגבולות האינטגרציה:



1. נבצע ראשית את האינטגרל על y בין x לאינסוף ולאחר מכן את האינטגרל על x בין 0 לאינסוף:

$$\begin{aligned}\iint_{D_1} \lambda_1\lambda_2 e^{-\lambda_1 x-\lambda_2 y}\,dxdy &= \int_0^\infty\int_x^\infty \lambda_1\lambda_2 e^{-\lambda_1 x-\lambda_2 y}\,dydx \\ &= \int_0^\infty \lambda_1 e^{-\lambda_1 x}\left(-e^{-\lambda_2 y}\right)_x^\infty\,dx = \int_0^\infty \lambda_1 e^{-\lambda_1 x}e^{-\lambda_2 x}\,dx \\ &= -\frac{\lambda_1}{\lambda_1+\lambda_2}e^{-(\lambda_1+\lambda_2)x}\Big|_0^\infty = \frac{\lambda_1}{\lambda_1+\lambda_2}\end{aligned}$$

2. נשים לב שהאינטגרנד הינו קבוע ולכן תוצאת חישוב האינטגרל תהיה הקבוע כפול שטח תחום האינטגרציה. במקרה זה תחום האינטגרציה הינו רבע מעגל ברדיוס 1 ולכן:

$$\iint_{D_2} c\,dxdy = c \cdot S_{D_2} = c\frac{\pi}{4}$$

ניתן לנסות לפתור אינטגרל זה גם על ידי חישוב רגיל:

$$\iint_{D_2} c\,dxdy \;=\; \int_0^1 \int_0^{\sqrt{1-x^2}} c\,dydx = \int_0^1 c\sqrt{1-x^2}\,dx$$

האינטגרל האחרון אפשרי לחישוב (למשל, באמצעות אינטגרציה בחלקים) אך זהו חישוב מסובך מאוד וטכני. לפיכך, נוותר על התענוג המפופק ונשים לב שבבעיה יש סימטריה מעגלית וניתן לפשט את האינטגרל הכפול על ידי מעבר לקואורדינטות קוטביות. היות וזהו אינטגרל על רבע מעגל ברדיוס 1, גבולות האינטגרציה של הזווית יהיו בין 0 ל-$\frac{\pi}{2}$ ושל הרדיוס יהיו בין 0 ל-1:

$$\begin{aligned}\iint_{D_2} c\,dxdy &= \int_0^1 \int_0^{\pi/2} cr\,d\theta dr \\ &= \int_0^1 \frac{\pi cr}{2}\,dr = \frac{\pi c}{2}\cdot \left.\frac{r^2}{2}\right|_0^1 = \frac{\pi c}{4}\end{aligned}$$

3. במקרה הזה תחום האינטגרציה של שני המשתנים קבוע ואינו תלוי זה בזה ולכן אין זה משנה על איזה מהמשתנים נבצע את האינטגרל הראשון:

$$\begin{aligned}\iint_{D_3} e^{-x}\,dxdy &= \int_0^\infty \int_1^2 e^{-x}\,dydx = \int_0^\infty e^{-x}dx \\ &= \left.-e^{-x}\right|_0^\infty = 1\end{aligned}$$

□

0.4.1 עובדות שכדאי לדעת

1. למטבע יש שני צדדים – עץ ופלי.
2. לסביבון יש 4 צדדים – נ' ג' ה' פ'.
3. לקוביה יש 6 פאות הממוספרות מ-1 עד 6.
4. בחפיסת קלפים יש 4 צורות (לב, תלתן, יהלום ועלה). בכל צורה יש 13 קלפים: קלפי מספר אס, המספרים 2-10, נסיך, מלכה ומלך.
5. בלוח שחמט 8 שורות ו-8 עמודות. לכל צד יש 16 כלים בתחילת המשחק: 8 רגלים, 2 צריחים, 2 רצים, 2 פרשים, מלכה ומלך.
6. במשחק הדומינו כל אבן מחולקת ל-2, ועל כל חלק שלה מצוייר מספר נקודות בין 0 ל-6. אין אבני דומינו זהות, אבל ישנן אבני דומינו עם אותו מספר על שני החלקים (למשל 5-5). בסה"כ יש 28 אבנים.

1 תורת הקבוצות ומבוא ללוגיקה

תורת הקבוצות היא הבסיס הפורמלי עליו מתבססת תורת ההסתברות. נשתמש בקבוצות על מנת לתאר מאורעות שאת הסתברותם נרצה לחשב. בפרק זה נסקור בקצרה את ההגדרות והתוצאות החשובות לצרכינו של תורת הקבוצות תוך שימוש בטרמינולוגיה המתאימה לתורת ההסתברות. את חישובי ההסתברות עצמם נשאיר לפרקים הבאים.

הגדרה 1.1. מאורע (קבוצה) היא אוסף של תוצאות אפשריות של ניסוי מסויים (איברי הקבוצה).

לדוגמא, אם הניסוי הוא הטלת קוביה אזי $A = \{2, 4, 6\}$ הוא המאורע שבהטלת קוביה התקבלה תוצאה זוגית ואילו $B = \{2, 5\}$ הוא המאורע שבהטלת קוביה התקבל 2 או 5. סדר רישום התוצאות או חזרה על תוצאה מספר פעמים אינם משנים את המאורע: $B = \{5, 2\} = \{2, 5, 2\}$. כאשר תוצאה מסוימת שייכת למאורע נסמן זאת על ידי הסימן $\in$ (למשל, $2 \in A$) וכאשר תוצאה אינה חלק ממאורע נסמן זאת על ידי $\notin$ ($1 \notin A$). בנוסף, נגדיר את הקבוצות הבאות שהינן קבוצות בסיסיות במתמטיקה וישמשו אותנו בהמשך:

- $\{\} = \emptyset$ הקבוצה הריקה, שאינה מכילה אף איבר (המאורע הריק).
- $\mathbb{N} = \{1, 2, ...\}$ קבוצת המספרים הטבעיים.
- $\mathbb{R}$- קבוצת המספרים הממשיים.

הגדרה 1.2. מרחב המדגם הוא אוסף כל התוצאות האפשריות בניסוי (איברים). מרחב המדגם יסומן באות Ω (אומגה גדולה).

למשל, מרחב המדגם כאשר מטילים קוביה הינו $\Omega = \{1, 2, 3, 4, 5, 6\}$ ואילו מרחב המדגם המתאר את שתי התוצאות האפשריות בהטלת מטבע הינו $\Omega = \{H, T\}$[1]. מרחב המדגם לא חייב להיות קבוצה סופית. מרחב המדגם יכול להיות קבוצה בת מנייה, למשל, מספר הזריקות לסל עד הקליעה הראשונה של שחקן כדורסל יכול להיות כל מספר טבעי ($\Omega = \mathbb{N}$) או אפילו קבוצות מעוצמת הרצף, למשל כשהניסוי הוא ירי לעבר מטרה ומדידת המרחק ממרכזה, מרחב המדגם יהיה $\Omega = [0, \infty)$. ניתן לתאר את מרחב המדגם האחרון על ידי $\Omega = \{x \in \mathbb{R} | x \geq 0\}$, קרי Ω היא קבוצת המספרים הממשיים x המקיימים $x \geq 0$. באמצעות צורת רישום זו ניתן לתאר קבוצות מורכבות יותר שתנאי הכניסה אליהן ניתן לתיאור באמצעות משוואות או אי-שוויונים.

כאשר תוצאת הניסוי שהתקבלה בפועל שייכת למאורע, נאמר שהמאורע התרחש. למשל, אם הקוביה נפלה על המספר 4 נאמר שהמאורע A התרחש ואילו המאורע B לא התרחש.

תרגיל 1.1. רשמו את מרחב המדגם ואת המאורע המתוארים להלן במונחי תורת הקבוצות:

1. מטילים שני מטבעות. A- התקבלה תוצאה שונה בשניהם.
2. בוחרים סטודנט באקראי ומודדים את גובהו. B- התלמיד גבוה מ-2 מטרים.
3. מטילים שתי קוביות. C- סכומן יצא 5.
4. מתוך קופסא עם 10 מפתחות מוציאים מפתח באקראי ומנסים לפתוח איתו את הדלת עד שמצליחים. מפתח שלא פותח את הדלת מוחזר לקופסא.D – המפתח ה-15 שנוסה פתח את הדלת.

---

[1] התוצאות הן עץ (H) או פלי (T).

5. יותם קורא ספר בן 60 עמודים עד אשר נרדם באקראי תוך קריאת אחד העמודים. E הוא המאורע שיותם לא נרדם בעמוד ה-5.

הגדרה 1.3. מאורע A חלקי למאורע B אם כל תוצאה הכלולה ב-A כלולה גם ב-B. במקרה כזה נסמן $A \subseteq B$. מאורע A חלקי-ממש ל-B אם A חלקי ל-B אך הם לא שווים. במקרה כזה נסמן $A \subset B$. בפרט, כל מאורע מוכל במרחב המדגם, קרי $A \subseteq \Omega$ וכל מאורע מכיל את המאורע הריק $\emptyset \subseteq A$.

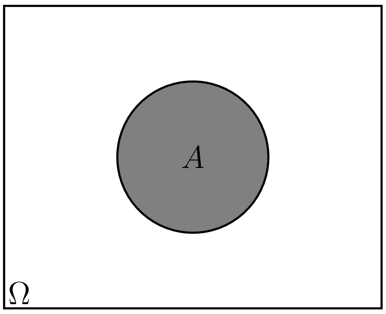
לדוגמא, כאשר הניסוי הוא הטלת קוביה פעם אחת, המאורע שיצא 1 חלקי-ממש למאורע שיצאה תוצאה אי זוגית שכן $\{1\} \subset \{1,3,5\}$ אך לא חלקי למאורע שיצאה תוצאה זוגית, שכן $\{1\} \not\subseteq \{2,4,6\}$.

משפט 1.1. *שני מאורעות שווים אם ורק אם הם מכילים בדיוק את אותן תוצאות, כלומר אם כל אחד חלקי לשני: $A = B$ אם"ם $A \subseteq B$ וגם $B \subseteq A$.*

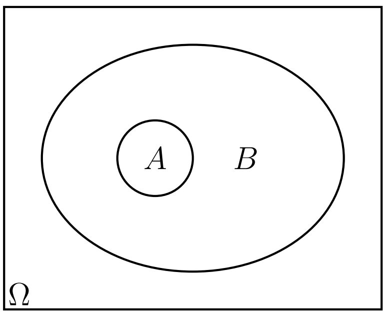
משפט זה מהווה את הדרך המקובלת להוכיח ששני מאורעות שווים ונשתמש בה בהמשך בהוכחת חוקי הפעולות על קבוצות.

1.0.1 דיאגרמת ון

דיאגרמת ון (Venn) היא שיטה גרפית נוחה לתאר קבוצות ואת היחסים בין הקבוצות. מרחב המדגם יתואר באמצעות מלבן הכולל בתוכו הכל ואילו מאורעות יוצגו באמצעות שטחים חלקיים לו, לרוב מעגלים או אליפסות. הצורות עצמן, גודלן ושטחן לא משנים אלא רק היחסים ההדדיים ביניהן - האם הן נחתכות, האם הן מוכלות אחת בשניה וכדומה. תוצאות אפשריות בניסוי יהיו נקודות הכלולות במאורעות המכילים אותם. לרוב לא נציג את הנקודות עצמן ונתעניין רק במאורעות וביחסי הגומלין שלהם. למשל, דיאגרמה המכילה מאורע אחד A תשורטט באופן הבא:



ואילו שני מאורעות המקיימים $A \subseteq B$ ישורטטו באופן הבא:



1.1 פעולות על קבוצות

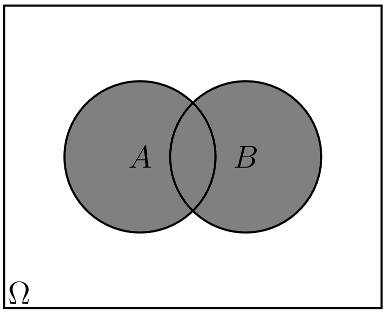
1.1.1 איחוד

הגדרה 1.4. האיחוד של שני מאורעות A ו-B מכיל את כל התוצאות הנמצאות ב-A או ב-B או בשניהם:

$$A \cup B = \{\omega \in \Omega | \omega \in A \text{ or } \omega \in B\} \tag{1.1}$$

מבחינה מילולית, המשמעות של איחוד מאורעות זהה למשמעות של המילה "או" בשפת הדיבור. $A \cup B$ משמעו A או B (או שניהם[2]). למשל, אם הניסוי הוא הטלה של קוביה, $A = \{2, 4, 6\}$ הוא המאורע שהתוצאה זוגית ו-$B = \{1, 2\}$ הוא המאורע שהתוצאה קטנה מ-3 אז $A \cup B = \{1, 2, 4, 6\}$, כלומר התוצאה היא זוגית או קטנה מ-3.

בדיאגרמת ון, האיחוד של המאורעות A, B הוא השטח השייך לפחות לאחד מהמאורעות:



באופן טבעי, איחוד הינה פעולה חילופית - המאורע "A או B" זהה למאורע "B או A":

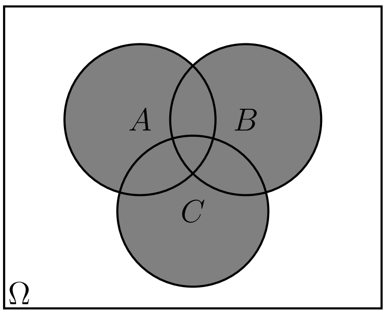
$$A \cup B = B \cup A \tag{1.2}$$

כמו כן, ניתן לראות שכל אחד מהמאורעות מוכל באיחוד של שניהם: $A \subseteq A \cup B, B \subseteq A \cup B$.

עבור שלושה מאורעות ומעלה, מתקיימת תכונת הקיבוציות הקובעת שהסדר בו מבוצע האיחוד אינו משנה:

$$(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C \tag{1.3}$$

והתחום המתאים בדיאגרמת ון הינו:



[2] באופן כללי במתמטיקה, משמעות המילה "או" היא "ו/או", כלומר או האחד, או השני, או שניהם גם יחד.

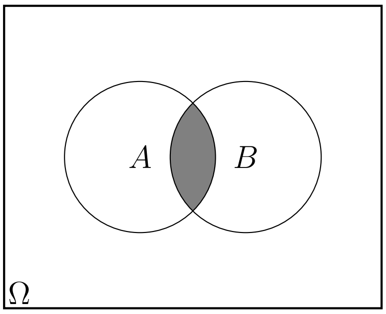
1.1.2 חיתוך

הגדרה 1.5. החיתוך של שני מאורעות A ו-B מכיל את כל התוצאות הנמצאות גם ב-A וגם ב-B:

$$A \cap B = \{\omega \in \Omega | \omega \in A \text{ and } \omega \in B\} \tag{1.4}$$

מבחינה מילולית, המשמעות של חיתוך מאורעות זהה למשמעות של המילה "וגם" בשפת הדיבור. $A \cap B$ משמעו A וגם B. למשל, אם הניסוי הוא הטלה של קוביה, $A = \{2, 4, 6\}$ הוא המאורע שהתוצאה זוגית ו-$B = \{1, 2\}$ הוא המאורע שהתוצאה קטנה מ-3 אז $A \cap B = \{2\}$, שהרי 2 היא התוצאה היחידה שהינה גם זוגית וגם קטנה מ-3.

בדיאגרמת ון, החיתוך של המאורעות A, B הוא השטח השייך לשני המאורעות:



באופן טבעי, חיתוך הינה פעולה חילופית - המאורע "A וגם B" זהה למאורע "B וגם A":

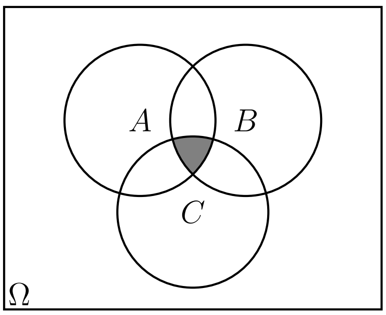
$$A \cap B = B \cap A \tag{1.5}$$

כמו כן, ניתן לראות שכל אחד מהמאורעות מכיל את החיתוך של שניהם: $A \cap B \subseteq A, A \cap B \subseteq B$.

עבור שלושה מאורעות ומעלה, מתקיימת תכונת הקיבוציות, שהרי הסדר בו מבוצע החיתוך אינו משנה:

$$(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C \tag{1.6}$$

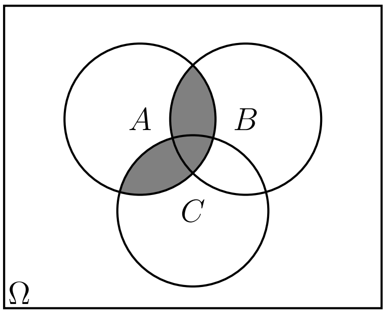
והתחום המתאים בדיאגרמת ון הינו:



מאורע הכולל איחודים וחיתוכים ניתן לעיתים לפישוט על ידי חוקי הפילוג, המזכירים פתיחת סוגריים:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \qquad \textbf{(1.7)}$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \qquad \textbf{(1.8)}$$

ניתן להוכיח את שני החוקים בצורה רגורוזית באמצעות הכלה דו-צדדית, בהתאם למשפט **1.1**, אך יותר פשוט להתרשם מנכונותם בעזרת דיאגרמת ון. כך למשל, אגף שמאל של חוק הקיבוץ הראשון הוא החלק המשותף של $B \cup C$ עם A, כלומר:

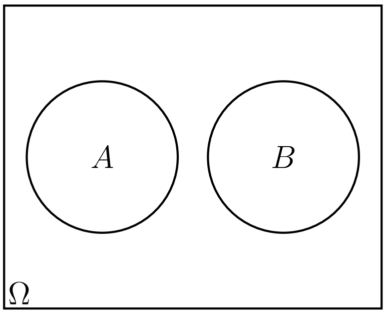


וזהו בדיוק אגף ימין, המורכב מצביעת שני תחומים: $A \cap B$ ו-$A \cap C$. ניתן כמובן להכליל חוקים אלו גם למספר רב יותר של מאורעות.

תרגיל 1.2. הוכיחו את חוק הפילוג השני (משוואה **1.8**) באמצעות הכלה דו-צדדית.

במקרים רבים יעניינו אותנו מאורעות שאינם חולקים תוצאות משותפות. במקרה כזה החיתוך של המאורעות לא יכיל אף תוצאה של הניסוי אלא יהיה המאורע הריק, $\emptyset$. היות ואין אף תוצאה המשותפת לשני המאורעות, נוכל לקבוע בוודאות שאם אחד המאורעות התרחש אזי המאורע השני לא התרחש. למקרה זה מתייחסת ההגדרה הבאה:

הגדרה 1.6. מאורעות A ו-B שעבורם מתקיים $A \cap B = \emptyset$ נקראים *מאורעות זרים*. בדיאגרמת ון נדגיש את העובדה שמאורעות זרים על ידי כך שנשרטט אותם ללא תחום משותף:



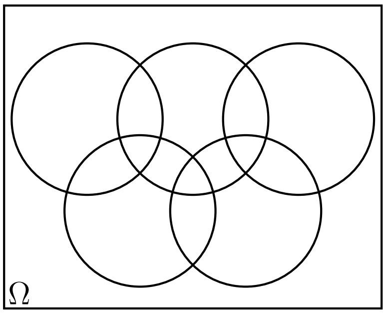
תרגיל 1.3. נתונים שני מאורעות זרים A, B ונתון שבניסוי מסויים המאורע A לא התרחש. האם מכאן ניתן להסיק שהמאורע B כן התרחש?

תרגיל 1.4. נתונים שני מאורעות זרים A, B. האם גם $\overline{A}, \overline{B}$ זרים?

ההגדרה נותרת בעינה גם כאשר מדובר על יותר מאורעות. למשל, ארבעת המאורעות A, B, C, D יקראו זרים אם מתקיים $A \cap B \cap C \cap D = \emptyset$. במקרה כזה לא ייתכן שארבעת המאורעות ייתרחשו בו זמנית שכן אין אף תוצאה אפשרית של הניסוי המוכלת בכל המאורעות. יחד עם זאת, בהחלט ייתכן שחלק מהמאורעות הללו מתרחשים בו זמנית. למשל, אם הניסוי הוא הטלת קוביה והמאורעות הינם $A = \{1\}, B = \{2\}, C = \{1,2\}, D = \{3\}$ הרי שארבעתם מאורעות זרים אך שני המאורעות A ו-C מתרחשים בו זמנית כאשר תוצאת ההטלה היא 1. ניתן להכליל את ההגדרה של מאורעות זרים ליותר משני מאורעות תוך שימור המשמעות שרק אחד מהמאורעות יכול להתרחש בכל פעם:

הגדרה 1.7. מאורעות $A_1, A_2, \ldots$ (מספר סופי או אינסופי של מאורעות) נקראים זרים בזוגות אם לכל $i \neq j$ מתקיים $A_i \cap A_j = \emptyset$.

מאורעות זרים בזוגות הם בוודאי גם זרים, שהרי אם חיתוך של שניים מהם הוא ריק אזי גם חיתוך של כולם יחד יהיה ריק, אך ההפך אינו בהכרח נכון כפי שהודגם. בדיאגרמת ון נשרטט מאורעות זרים בזוגות כסדרה של עיגולים שאף זוג מתוכם לא נחתך ואילו מאורעות זרים יכולים להיחתך, אך אסור שיהיה תחום המשותף לכל המאורעות. למשל, המאורעות הבאים זרים אך לא זרים בזוגות:

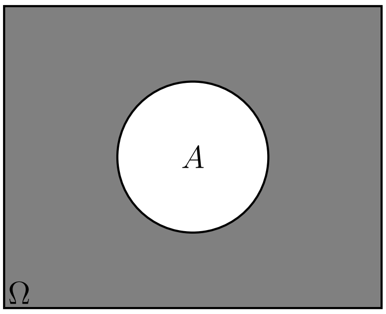


1.1.3 משלים

הגדרה 1.8. המשלים[3] של המאורע A מכיל את כל התוצאות הנמצאות במרחב המדגם שאינן נכללות ב-A:

$$\overline{A} = \{\omega \in \Omega | \omega \notin A\} \tag{1.9}$$

מבחינה מילולית, המשמעות של משלים זהה למשמעות של המילה "לא" בשפת הדיבור. $\overline{A}$ משמעו כל מה שלא A. למשל, אם הניסוי הוא הטלת קוביה ו-$A = \{2,4,6\}$ הוא המאורע שהתוצאה זוגית, $\overline{A} = \{1,3,5\}$ הוא המאורע שהתוצאה אי זוגית. בנוסף, $\overline{\emptyset} = \Omega$ וכן $\overline{\Omega} = \emptyset$. בדיאגרמת ון, המשלים של מאורע A הוא כל שטח מרחב המדגם הנמצא מחוץ ל-A:



[3] בחלק מהספרים משתמשים בסימון A^c.

באופן טבעי, שלילה כפולה ("לא לא A") מבטלת זו את זו ותכלול את כל התוצאות שכן נמצאות ב- A:

$$\overline{(\overline{A})} = A \tag{1.10}$$

תרגיל 1.5. מטילים מטבע 10 פעמים. תארו מילולית את המשלים של כל אחד מהמאורעות הבאים:

1. התקבל עץ לפחות 3 פעמים.
2. לכל היותר 4 פעמים התקבל עץ.
3. לא התקבל עץ כלל.
4. מספר העצים גדול ממספר הפלים.
5. בשתי ההטלות האחרונות התקבל עץ.

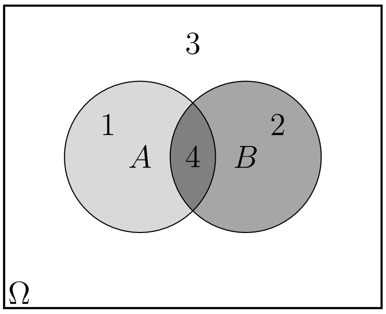
כאשר מבצעים משלים על איחוד או חיתוך של מאורעות, ניתן "לפתוח סוגריים" באמצעות חוקי דה-מורגן, המוצגים כאן עבור שני מאורעות אך נכונים גם למספר רב של מאורעות:

$$\overline{A \cup B} \;=\; \overline{A} \cap \overline{B} \tag{1.11}$$
$$\overline{A \cap B} \;=\; \overline{A} \cup \overline{B} \tag{1.12}$$

ניתן להוכיח את שני החוקים בצורה רגורוזית באמצעות הכלה דו-צדדית, בהתאם למשפט **1.1**, אך יותר פשוט להתרשם מנכונותם בעזרת דיאגרמת ון או תרגום הביטוי המתמטי לשפה המילולית. כך למשל, בחוק דה-מורגן הראשון (משוואה **1.11**) אנו מחשבים את $\overline{A \cup B}$ שמשמעו "לא (A או B)". כל תוצאה שתתאים לנו אסור לה להיות ב-A ואסור לה להיות ב-B (שהרי אז היא תהיה ב-$A \cup B$), ולכן התוצאות המבוקשות מקיימות "לא ב-A וגם לא ב-B" וזהו בדיוק אגף ימין של המשוואה: $\overline{A} \cap \overline{B}$.

תרגיל 1.6. הוכיחו את חוק דה-מורגן השני (משוואה **1.12**) באמצעות הכלה דו-צדדית.

תרגיל 1.7. רשום את כל אחד מהתחומים הבאים באמצעות פעולות החיתוך והמשלים:



לאור תרגיל **1.7** ניתן להציג כל מאורע המורכב מהמאורעות A ו-B באמצעות פעולות האיחוד, החיתוך והמשלים בלבד ולכן שלושת הפעולות הללו הן הפעולות הבסיסיות של תורת הקבוצות שישמשו אותנו בחישובי הסתברות. במקרים מסויימים הצגה באמצעות איחוד, חיתוך ומשלים בלבד אינה הנוחה ביותר להבנה או הפשוטה ביותר לרישום ולכן לעיתים נשתמש בפעולת ההפרש בין הקבוצות לצורך תיאור תחומים דוגמת 1 ו-2, תחומים השייכים לקבוצה אחת אך לא לאחרת:

$$A \backslash B = A \cap \overline{B} = \{x \in A | x \notin B\} \quad (1.13)$$

לפיכך, ניתן להציג כל מאורע A בתור איחוד של שני מאורעות זרים: המאורע "A וגם B" והמאורע "A אבל לא B", ונקבל את הנוסחא השימושית הבאה:

$$A = (A \setminus B) \cup (A \cap B) = \left(A \cap \overline{B}\right) \cup (A \cap B) \quad (1.14)$$

תרגיל 1.8. מטילים קוביה הוגנת פעם אחת. נסמן ב-A את המאורע שיצאה תוצאה זוגית, ב-B את המאורע שיצאה הספרה 5 וב-C את המאורע שיצאה תוצאה המתחלקת ב-3.

1. רשמו מפורשות את המאורעות A, B, C ואת מרחב המדגם Ω.
2. רשמו מפורשות את המאורעות $A \cup B$, $B \cap \overline{C}$, $(A \cup B) \cap C$ ו- $A \cup (B \cap C)$.
3. האם המאורעות A, B, C זרים? האם הם זרים בזוגות?

תרגיל 1.9. יותם, סופי ובן מסתדרים בשורה.

1. רשמו מפורשות את מרחב המדגם, קבוצת כל הסידורים האפשריים.
2. נגדיר את המאורעות A - יותם עומד בקצה הימני, B- בן עומד בקצה הימני, C - סופי עומדת בין שני הבנים. כתבו מפורשות את המאורעות הבאים ונסחו אותם במילים:
$A \cup B, A \cap B \cap C, (A \cap C) \cup B, \overline{A} \cap \overline{B}$

1.2 פתרונות לשאלות הפרק

תרגיל 1.1. 1. $\Omega=\{(H,H),(H,T),(T,H),(T,T)\}$. $A=\{(H,T),(T,H)\}$.

2. $\Omega=[0,3]$. $B=(2,3]$. ברמת העקרון ניתן להסתפק במרחב מדגם קטן יותר (למשל, $[0.5,3]$) אולם איננו יודעים את הטווח האמיתי ועדיף להגדיר מרחב מדגם גדול יותר הכולל תוצאות שלא יתקבלו לעולם מאשר להגדיר מרחב מדגם קטן מדי ולפסול תוצאות אפשריות בניסוי.

3. $\Omega=\{(1,1),(1,2),...,(5,6),(6,6)\}$ - בסה"כ 36 אפשרויות. $C=\{(1,4),(2,3),(3,2),(4,1)\}$. עקרונית אפשר גם לבחור במרחב מדגם המכיל את הסכומים האפשריים ובמקרה כזה $\Omega=\{2,3,4,...,11,12\}$ וכן $C=\{5\}$. בהמשך נראה איזה ממרחבי המדגם הללו עדיף לצורך חישובי הסתברות.

4. ברמת העיקרון כל מספר טבעי של ניסיונות אפשרי (שכן בכל שלב מחזירים את המפתחות שלא פתחו את הדלת לתיבה) ולכן התוצאות האפשריות בניסוי הן מהצורה "המפתח ה-n פתח את הדלת (והקודמים לא)". לכן $\Omega=\mathbb{N}$ ו-$D=\{15\}$.

5. מרחב המדגם הינו $\Omega=\{1,2,...,60\}$ ואילו $E=\{1,2,3,4,6,7,...,60\}$. טעות נפוצה היא להגדיר את המאורע E בתור $\{6,7,...,60\}$ מתוך מחשבה שכדי לא להירדם בעמוד החמישי צריך קודם כל להגיע אליו, אך לא כך הדבר שכן אם הסטודנט נרדם בעמוד השלישי, למשל, הוא בהכרח לא נרדם באף אחד מהעמודים האחרים ובפרט לא נרדם בעמוד החמישי.

□

תרגיל 1.2. נוכיח את חוק הפילוג באמצעות הכלה דו-צדדית:
יהי $x\in A\cup(B\cap C)$. לכן $x\in A$ או $x\in B\cap C$. אם $x\in A$ אז מתקיים $x\in A\cup B$ וגם $x\in A\cup C$ ולכן $x\in(A\cup B)\cap(A\cup C)$. אם $x\in B\cap C$ אז $x\in B$ וגם $x\in C$ ולכן מתקיים $x\in A\cup B$ וגם $x\in A\cup C$ ומכאן $x\in(A\cup B)\cap(A\cup C)$. לפיכך הראנו הכלה חד-צדדית: $A\cup(B\cap C)\subseteq(A\cup B)\cap(A\cup C)$.
הוכחת ההכלה ההפוכה דומה. □

תרגיל 1.3. הטענה אינה נכונה. למשל, אם הניסוי הוא הטלת קוביה, המאורעות $A=\{1\}$ ו-$B=\{2\}$ זרים, אך אם A לא התרחש לא נוכל לקבוע בוודאות ש-B כן התרחש, כי ייתכן שתוצאת ההטלה היא 5 ובמקרה כזה שני המאורעות לא התרחשו. □

תרגיל 1.4. הטענה אינה נכונה. למשל, אם הניסוי הוא הטלת קוביה, המאורעות $A=\{1\}$ ו-$B=\{2\}$ זרים אך $\overline{A}=\{2,3,4,5,6\}$ ו-$\overline{B}=\{1,3,4,5,6\}$ אינם זרים כי חיתוכם כולל את כל התוצאות הגדולות או שוות ל-3. □

תרגיל 1.5. 1. לא התקבל עץ לפחות 3 פעמים, כלומר התקבל עץ לכל היותר פעמיים.

2. לפחות 5 פעמים התקבל עץ.
3. התקבל עץ לפחות פעם אחת.
4. מספר העצים קטן או שווה למספר הפלים.
5. לפחות באחת משתי ההטלות האחרונות לא התקבל עץ.

□

תרגיל 1.6. נוכיח את חוק דה-מורגן באמצעות הכלה דו צדדית:
יהי $x\in\overline{A\cap B}$. לכן $x\notin A\cap B$ כלומר x לא שייך לאחת מהקבוצות (או לשתיהן). אם $x\notin A$ אז בהכרח $x\in\overline{A}$ ולכן $x\in\overline{A}\cup\overline{B}$. באותה צורה אם $x\notin B$ אז $x\in\overline{A}\cup\overline{B}$ ולסיכום $\overline{A\cap B}\subseteq\overline{A}\cup\overline{B}$.
הוכחת ההכלה ההפוכה דומה. □

תרגיל 1.7. **1.** $A\cap\overline{B}$

2. $\overline{A}\cap B$

3. $\overline{A}\cap\overline{B}$

4. $A\cap B$

□

תרגיל 1.8. **1.** מרחב המדגם הינו $\Omega=\{1,2,3,4,5,6\}$ ואילו המאורעות הינם $A=\{2,4,6\}, B=\{5\}$ ו-$C=\{3,6\}$.

2. לפי הסעיף הקודם, המאורעות המבוקשים הינם:

$$\begin{aligned}
A\cup B &= \{2,4,5,6\}\\
B\cap\overline{C} &= \{5\}\cap\{1,2,4,5\}=\{5\}\\
A\cup(B\cap C) &= \{2,4,6\}\cup\emptyset=\{2,4,6\}\\
(A\cup B)\cap C &= \{2,4,5,6\}\cap\{3,6\}=\{6\}
\end{aligned}$$

בשתי הנוסחאות האחרונות אנו רואים שמתקיים $A\cup(B\cap C)\neq(A\cup B)\cap C$, כלומר תכונת הקיבוציות לא מתקיימת כאשר מעורבות בחישוב גם פעולות איחוד וגם פעולות חיתוך וחשוב להקפיד על רישום סוגריים.

3. המאורעות B ו-C זרים ולכן $A\cap B\cap C=\emptyset$, כלומר שלושת המאורעות זרים אך אינם זרים בזוגות שכן $A\cap C=\{6\}\neq\emptyset$.

□

תרגיל 1.9. **1.** ישנם שישה סידורים אפשריים:

א) יותם, סופי, בן

ב) יותם, בן, סופי

ג) סופי, יותם, בן

ד) סופי, בן, יותם

ה) בן, סופי, יותם

ו) בן, יותם, סופי

2. $A\cup B$ הוא המאורע שבן או יותם עומדים בקצה הימני ויש 4 אפשרויות כאלו כפי שרשמנו בסעיף הקודם.
$A\cap B\cap C$ הוא המאורע שגם בן וגם יותם עומדים בקצה הימני וגם סופי עומדת ביניהם. היות שניהם לא יכולים לעמוד בקצה הימני הרי שמאורע זה הינו המאורע הריק.
$(A\cap C)\cup B$ הוא המאורע שבן עומד בקצה הימני או ש(יותם עומד בקצה הימני וסופי עומדת ביניהם). ברשימה של הסעיף הקודם, יש רק סידור אחד בו $A\cap C$ מתקיים (יותם, סופי, בן) וישנם עוד שני סידורים בהם בן עומד בקצה הימני ולכן בסך הכל מאורע זה כולל בתוכו שלושה סידורים אפשריים.
$\overline{A}\cap\overline{B}$ הוא המאורע שבן לא עומד בקצה הימני וגם יותם לא עומד בקצה הימני. לכן בהכרח סופי עומדת בקצה הימני ויש שני סידורים כאלו.

□

2 מבוא לתורת ההסתברות

נהוג לייחס את ראשיתה של תורת ההסתברות המודרנית למאה ה-17, כאשר האציל הצרפתי וחובב ההימורים שבלייר דה-מרה (Chevalier de-Mere) פנה למתמטיקאי בלייז פסקל (Blaise Pascal) וביקש עזרה בחישובי הכדאיות של משחקי המזל אותם נהג לשחק (ראו תרגיל 3.9). פניה זו הובילה לחליפת מכתבים בין פסקל לבין מתמטיקאי מפורסם אחר בן זמנו, פייר דה-פרמה (Pierre de-Fermat), במהלכה פותחה תורת ההסתברות הקלאסית תוך התמקדות בבעיות העוסקות במשחקי מזל עם תוצאות סימטריות. בבעיות אלו, אם במשחק מסוים יש n תוצאות אפשריות וכולן סימטריות אזי לכולן אותו סיכוי להתקבל. לפיכך הסיכוי שתתקבל כל אחת מהתוצאות הוא $\frac{1}{n}$, הסיכוי שתתקבל אחת משתי תוצאות מסוימות הוא $\frac{2}{n}$ וכן הלאה. כך למשל, בהטלה של קוביה הסיכוי שנקבל את הספרה 2 הוא $\frac{1}{6}$ (יש בסך הכל 6 תוצאות אפשריות) והסיכוי שנקבל תוצאה זוגית הוא $\frac{3}{6}$ כי יש 3 תוצאות זוגיות מתוך 6 התוצאות האפשריות לניסוי. תאוריה זו מאפשרת חישובי הסתברויות רבים במסגרת נושא הקומבינטוריקה היעמוד בליבו של הפרק הבא, פרק 3.

למרות שהגישה של פסקל ופרמה טובה לפתרון מגוון רחב של בעיות, היא אינה מתאימה תמיד שכן ייתכן והתוצאות האפשריות אינן סימטריות ושוות הסתברות. כך למשל, נניח שאנחנו מעוניינים לחשב את ההסתברות ששחקן כדורסל הקולע לסל בהסתברות של 0.9 יקלע 2 סלים ב-2 זריקות. לניסוי זה יש 4 תוצאות אפשריות[1] אך ברור שתוצאות אלו אינן סימטריות והסיכוי לשני סלים צריך להיות גבוה יחסית לאפשרויות האחרות שכן זהו שחקן טוב הקולע בהסתברות גבוהה. כחלופה לגישה הקלאסית פותחה במאה ה-19 על ידי ריכרד פון-מיזס (Richard von-Mises) גישה שכיחותית הגורסת שהסתברות של מאורע היא השכיחות שבו המאורע יקרה אם נחזור על ניסוי הרבה פעמים. לכן, כדי לחשב את ההסתברות שהכדורסלן יקלע לסל פעמיים, עלינו לבקש מהכדורסלן לחזור על הניסוי בו הוא זורק לסל פעמיים n פעמים ולספור את S_n – מספר הפעמים בהן הוא קלע פעמיים. אנחנו מצפים שעבור n גדול מספיק, המנה $\frac{S_n}{n}$ תתאר בקירוב את ההסתברות לכך שמאורע זה יקרה. באופן כללי, אם נחזור על הניסוי n פעמים ונגדיר בתור $S_n(A)$ בתור מספר הפעמים שמאורע A התרחש אזי ההסתברות של המאורע תחושב באמצעות:

$$Pr(A) = \lim_{n\to\infty} \frac{S_n(A)}{n} \tag{2.1}$$

באופן טבעי, גישה זו אינה נוחה לעבודה שכן לא תמיד ניתן לחזור על ניסוי מספר רב של פעמים בכדי למדוד הסתברות ובכל מקרה מדובר על גישה אמפירית-ניסיונית ואילו אנחנו מחפשים גישה תאורטית שתאפשר חישובי הסתברויות מתוך נוסחאות ולא מתוך חזרה פיזית על הניסוי. יחד עם זאת, גישה זו מאוד חשובה כי היא תופסת את המשמעות האינטואיטיבית של מושג ההסתברות כשכיחות של מאורע. כאשר אנחנו אומרים שההסתברות של הספרה 2 בהטלת קוביה היא $\frac{1}{6}$, אנחנו מתכוונים שבערך בשישית מהטלות הקוביה נקבל את הספרה 2. לכן, אם נטיל את הקוביה 60 פעמים אנחנו מצפים לקבל בערך 10 פעמים את הספרה 2 ואם נטיל את הקוביה 600 פעמים אנחנו מצפים שבערך[2] 100 מההטלות יפלו על 2. לכן, במקרים רבים נחזור אל הגישה השכיחותית בכדי לנמק תוצאות, משפטים ונוסחאות ולהראות שאכן הם מתיישבים עם האינטואיציה ועם האופן בו אנחנו רוצים להגדיר הסתברות.

[1] שני סלים, שתי החטאות, קליעה והחטאה והחטאה וקליעה.

[2] בהמשך נלמד לכמת את ה"בערך".

הגישה עליה נתבסס לצורך פיתוח תורת ההסתברות היא הגישה האקסיומטית שנוקטתה בתחילת המאה ה-20 על ידי המתמטיקאי הרוסי אנדריי קולמוגורוב (Andrey Kolmogorov). באותה תקופה תחומים מתמטיים רבים זכו למתיחת פנים וסידור מחדש והובאו לצורתם הנוכחית והמודרנית המופיעה בספרי הלימוד הסטנדרטיים תוך נקיטה בגישה האקסיומטית. הדוגמא הבולטת של הגישה האקסיומטית היא הגאומטריה האוקלידית המוכרת מהתיכון. כזכור, גאומטריה אוקלידית מתבססת על חמש אקסיומות (למשל "בין כל שתי נקודות ניתן להעביר קו ישר") המהוות תכונות יסודיות אשר אינן דורשות הוכחה ועל סמך תכונות אלו וכללי היסק הלוגיים מפתחים נוסחאות ומוכיחים משפטים. באופן דומה, קולמוגורוב הגדיר שלוש אקסיומות בסיסיות ועל סמך אקסיומות אלו פיתח נוסחאות ומשפטים רבים לחישוב הסתברויות. אנו ננקוט בגישה זו להגדרת הסתברויות כאשר נשתמש בגרסאות פשוטות יותר של אקסיומות קולמגורוב תוך ויתור מסוים על הדיוק המתמטי לטובת הפשטות וההתמקדות בהסתברויות "מעשיות" המעניינות אותנו.

נסיים את ההקדמה בפירוש נוסף הקיים להסתברות - שכיחות של תופעות באוכלוסיה. לפי פירוש זה, השכיחות של תכונה מסויימת באוכלוסיה היא גם ההסתברות שאם נבחר פריט אחד באוכלוסיה, הוא יקיים תכונה זו. למשל, את הנתון ש- 75% מתושבי ישראל הם יהודים ניתן לפרש גם בתור תוצאה של ניסוי: כאשר בוחרים באקראי ישראלי, הסיכוי שייבחר יהודי הוא 0.75. באופן זה, ניתן להשתמש בנוסחאות ההסתברותיות שנפתח בפרק זה לצורך חישובי שכיחויות של תופעות עליהן ידועים נתונים אמפיריים.

2.1 אקסיומות ההסתברות ונוסחאות בסיסיות

נסתכל על ניסוי כלשהו ונגדיר את מרחב המדגם, Ω, בתור קבוצת כל התוצאות האפשריות בניסוי. אנו מעוניינים להתאים לכל מאורע אפשרי $A \subseteq \Omega$ מספר ממשי, היסומן ב-$Pr(A)$, ויתאר את ההסתברות שמאורע A יתרחש. לצורך ניסוח אקסיומות ההסתברות עלינו להבהיר לעצמנו מהן הציפיות שפונקציית ההסתברות, המתאימה לכל מאורע את ההסתברות להתרחשותו, צריכה לקיים. ראשית, כפי שהגישה השכיחותית מרמזת, ההסתברות של כל מאורע מתקבלת על ידי ספירה של מספר הניסויים בהן המאורע התרחש מתוך כלל הניסויים. מאחר ומספר הפעמים שמאורע מתרחש חייב להיות אי-שלילי, הרי שגם ההסתברות להתרחשותו צריכה להיות אי-שלילית.

***עובדה* 2.1** (אקסיומת קולמוגורוב הראשונה)**.** *הסתברות של מאורע היא מספר אי-שלילי, כלומר לכל* $A \subseteq \Omega$ *מתקיים* $Pr(A) \geq 0$.

האקסיומה השניה עוסקת במרחב המדגם כולו. ההסתברות שמרחב המדגם יתרחש, $Pr(\Omega)$, היא בהכרח **1** שכן מרחב המדגם מוגדר כאוסף כל התוצאות האפשריות בניסוי ולכן בכל חזרה בניסוי, אחת מתוצאות אלו תקרה. מבחינה שכיחותית, בכל פעם שנחזור על הניסוי בהכרח נקבל תוצאה מתוך מרחב המדגם ולכן, בסימוני משוואה **2.1**, תמיד יתקיים $S_n(\Omega) = n$ כך שההסתברות לקבלת מרחב המדגם תהיה **1**.

***עובדה* 2.2** (אקסיומת קולמוגורוב השנייה)**.** *ההסתברות של מרחב המדגם כולו היא* $Pr(\Omega) = 1$.

האקסיומה השלישית והאחרונה עוסקת בהסתברות של איחוד מאורעות זרים. נניח שנתונים מאורעות זרים $A, B \subseteq \Omega$ שהסתברותם ידועה ואנחנו מעוניינים לחשב את ההסתברות של האיחוד, $Pr(A \cup B)$. בכל חזרה על הניסוי תיתכן אחת בדיוק משלוש האפשרויות הבאות: **(1)** מאורע A התרחש, או; **(2)** מאורע B התרחש, או; **(3)** אף אחד מהם לא התרחש, כלומר $\overline{A} \cap \overline{B}$ התרחש. מספר הפעמים בהן התרחש המאורע $A \cup B$ יהיה שווה ל-$S_n(A) + S_n(B)$ ולכן גם $Pr(A \cup B) = Pr(A) + Pr(B)$. לדוגמא, ההסתברות שבהטלת קוביה נקבל את המספר **1** או **2** היא $\frac{2}{6} = \frac{1}{6} + \frac{1}{6}$ שכן המאורע "בהטלת קוביה התקבל המספר **1**" זר למאורע "בהטלת קוביה התקבל המספר **2**". העובדה שהמאורעות זרים חיונית בכדי שחישוב זה יהיה נכון, שכן אם המאורעות אינם זרים ישנן תוצאות שמתקבלות בניסוי ונספרות גם כחלק ממאורע A וגם כחלק ממאורע B. האקסיומה השלישית מכלילה אינטואיציה זו למספר סופי או בן מנייה של מאורעות, כשהתנאי הנדרש הוא שלא תהיה אף תוצאה של הניסוי המשותפת לאף זוג של מאורעות, כלומר שהמאורעות יהיו זרים בזוגות.

***עובדה* 2.3** (אקסיומת קולמוגורוב השלישית). *יהיו* $A_1, A_2, ...A_n \subseteq \Omega$ *מאורעות זרים בזוגות (מספר סופי* $n \in \mathbb{N}$ *או אינסופי בן מנייה* $n = \infty$*). ההסתברות של האיחוד היא*

$$Pr\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} Pr\left(A_i\right) \tag{2.2}$$

חמושים בשלושת האקסיומות שפונקציית ההסתברות מקיימת אנחנו מוכנים לפתח נוסחאות נוספות לחישובי הסתברויות. עד כה אנחנו יודעים שההסתברות היא מספר חיובי ושההסתברות של מרחב המדגם כולו היא 1. מבחינה אינטואיטיבית, אנחנו מבינים את המשמעות של הסתברות קטנה מ-1 אך לא של הסתברויות גדולות מ-1. למשל, ההסתברות לקבל 2 בקוביה היא $\frac{1}{6}$ ולכן בערך בשישית מההטלות הקוביה תיפול על 2, אבל מה המשמעות של מאורע שההסתברות שלו היא 1.5? אם נעשה את הניסוי 100 פעמים אז המאורע אמור להתרחש 150 פעמים? תוצאה זו לא ברורה ולכן הגיוני שהסתברות תהיה מספר קטן מ-1, כפי שנקבל מהמשפט הבא:

***משפט* 2.1.** *לכל מאורע* $\Omega \supseteq A$ *מתקיים* $Pr(A) \le 1$*, כלומר הסתברות של כל מאורע היא מספר חיובי הקטן או שווה ל-1.*

הוכחה. מתקיים $\Omega = A \cup \overline{A}$ וכן המאורעות A ו-$\overline{A}$ זרים ולכן לפי אקסיומה 2.3 נקבל

$$Pr(\Omega) = Pr\left(A \cup \overline{A}\right) = Pr\left(A\right) + Pr\left(\overline{A}\right) \tag{2.3}$$

לפי אקסיומה 2.1 ההסתברות של המשלים היא חיובית, $Pr\left(\overline{A}\right) \ge 0$, ואילו לפי אקסיומה 2.2 סכום ההסתברויות הינו $Pr(\Omega) = 1$, לכן בהכרח $Pr(A) \le 1$, כדרוש. □

נשים לב שבמהלך ההוכחה קיבלנו משוואה מאוד שימושית - סכום ההסתברויות של מאורע ושל המשלים שלו שווים ל-1. לפיכך, די אם נדע לחשב את ההסתברות של אחד מהמאורעות כדי שנדע את ההסתברות של שניהם. בעתיד נראה שישנם מקרים בהם חישוב ישיר של הסתברות של מאורע יכול להיות מסובך וארוך ואילו חישוב ההסתברות של המשלים שלו יהיה פשוט בהרבה ולכן נעדיף לטפל במשלים ולהשתמש בתכונה זו (ראו למשל את תרגיל 2.1). מאחר ותכונה זו חשובה מאוד, ננסח אותה בתור משפט בפני עצמו:

***משפט* 2.2.** *לכל מאורע* $A \subseteq \Omega$ *מתקיים* $Pr(\overline{A}) = 1 - Pr(A)$.

הוכחה. ההוכחה מתקבלת מיידית על סמך משוואה 2.3. □

באמצעות המשפט האחרון אפשר לחשב את ההסתברות של המאורע הריק, המתאר את המצב בו לא יצאה אף אחת מהתוצאות של מרחב המדגם בניסוי. היות ואנחנו מניחים שכל התוצאות האפשריות מוכלות במרחב המדגם, אחת מהן חייבת לצאת בניסוי והסיכוי שאף אחת מהן לא תקרה הינו 0. ניתן להוכיח זאת גם בצורה ישירה, תוך הסתמכות על כך שהמאורע הריק הינו המשלים של מרחב המדגם:

***משפט* 2.3.** *ההסתברות של המאורע הריק הינה* $Pr(\emptyset) = 0$.

הוכחה. נובע ישירות מכך ש-$\overline{\emptyset} = \Omega$ בשילוב עם אקסיומה 2.2:

$$Pr(\emptyset) = 1 - Pr\left(\overline{\emptyset}\right) = 1 - Pr(\Omega) = 0$$

□

לסיום סעיף זה נדון בקשר שבין שני מאורעות כך שאחד מוכל בשני. אם $A \subseteq B$ אז בכל פעם שהמאורע A מתרחש בהכרח גם המאורע B מתרחש, אך ייתכנו מקרים בהם יתקבלו תוצאות שנמצאות רק ב-B ולא ב-A. לכן, מספר הפעמים שבהם המאורע A מתרחש צריך להיות קטן יותר או שווה למספר הפעמים שמאורע B מתרחש. אינטואיציה זו מסוכמת במשפט הבא:

***משפט* 2.4.** *עבור כל שני מאורעות המקיימים $A \subseteq B$ מתקיים* $Pr(A) \leq Pr(B)$.

הוכחה. ניתן לרשום $B = A \cup (B\backslash A)$ כאשר המאורעות A ו-$B\backslash A$ זרים (לפי נוסחא **1.14**) ולכן

$$Pr(B) = Pr\left(A \cup (B\backslash A)\right) \overset{eq.2.2}{=} Pr(A) + Pr(B\backslash A) \geq Pr(A)$$

כאשר האי-שוויון האחרון מתקיים מאקסיומה **2.1** שכן $Pr(B\backslash A) \geq 0$. □

בהוכחת המשפט האחרון השתמשנו באסטרטגיה אשר תסייע לנו רבות לחישוב הסתברות של מאורע מורכב – הצגתו כאיחוד של מאורעות זרים בזוגות אשר כל אחד מהם ניתן לחישוב בצורה פשוטה יותר. במהלך הוכחת המשפט פרקנו באופן זה את B ל-A ו-$B \setminus A$ אך פירוק זה אפשרי גם למספר רב יותר של מאורעות, כפי שתראו בתרגילים הבאים. בפרט, ניתן לחשב את ההסתברות של כל מאורע על ידי סכימה על כל תוצאות הניסוי הכלולות במאורע, שהרי אם $A = \{\omega_1, \omega_2 ..., \omega_n\}$ אזי $A = \{\omega_1\} \cup \{\omega_2\} \cup ... \cup \{\omega_n\}$ ולכן

$$Pr(A) = \sum_{i=1}^{n} Pr\left(\{\omega_i\}\right) \tag{2.4}$$

בסעיף הבא נראה כיצד ניתן לחשב הסתברות של מאורע שמוצג כאיחוד של מאורעות שאינם זרים.

תרגיל 2.1. נתון מרחב המדגם $\Omega = \{1, 2, ..., 10\}$ ונניח שההסתברות שתצא התוצאה $i \in \Omega$ היא $c \cdot i^2$ כאשר $c > 0$ קבוע.

1. מצאו את ערכו של c.
2. מה ההסתברות שתוצאת הניסוי תתחלק ב-3?
3. מה ההסתברות שתוצאת הניסוי תהיה גדולה מ-2?

תרגיל 2.2. נתון מרחב המדגם $\Omega = \mathbb{N}$ ונניח שההסתברות שתצא התוצאה $i \in \Omega$ היא $\frac{c}{i^2}$ כאשר $c > 0$ קבוע.

1. מצאו את ערכו של c (רמז: נוסחא **0.23**).
2. מה הסיכוי שתוצאת הניסוי תהיה זוגית?
3. מה הסיכוי שתוצאת הניסוי תהיה זוגית אך לא תתחלק ב-3?

אקסיומה חשובה נוספת שאינה חלק מאקסיומות קולמוגורוב אבל הכרחית לביצוע חישובים הסתברותיים עוסקת בתוצאות סימטריות: אם ישנן שתי תוצאות סימטריות במרחב המדגם, ההסתברות שלהן צריכה להיות שווה. למשל, לקוביה הוגנת יש 6 פאות וכל הפאות סימטריות – הקוביה לא נוטה ליפול יותר על "4" מאשר על "1" או להפך. לכן ההסתברות של כל אחד ממספרי הקוביה צריכה להיות שווה ובשילוב עם העובדה שסכום ההסתברויות הוא 1, ההסתברות של כל פאה היא $\frac{1}{6}$. באותה צורה, הסיכוי לקבל עץ בהטלת מטבע הוא $\frac{1}{2}$ והסיכוי להוציא כדור לבן מתוך כד בו 3 כדורים לבנים ו-7 כדורים שחורים הוא $\frac{3}{10}$.

***עובדה* 2.4** (עקרון הסימטריה). *לתוצאות סימטריות הסתברות שווה להתקבל.*

בפרק זה נשתמש בעקרון הסימטריה לחישוב הסתברויות בסיסיות, דוגמת ההסתברות של התוצאות האפשריות בהטלת קוביה, מטבע וכד'. בפרק הבא, פרק 3, נתבסס על עקרון זה בכדי לחשב הסתברויות של מאורעות מסובכים יותר בניסויים מורכבים יותר מאשר הוצאת כדור בודד מכד. למשל, נלמד כיצד לחשב את ההסתברות שבניסוי בו מוצאים בדיוק 8 כדורים מכד המכיל 10 כדורים, כדור לבן אחד נשאר בכד ושני הכדורים הלבנים האחרים הוצאו בזה אחר זה (ראו תרגיל 3.14).

2.1.1 נוסחת ההכלה וההפרדה

אנו מעוניינים לחשב הסתברות של איחוד מאורעות, אך אקסיומה 2.3 מאפשרת חישוב הסתברות של איחוד של מאורעות רק בתנאי שמאורעות אלו זרים בזוגות. אם המאורעות אינם זרים בזוגות לא ניתן לחשב את הסתברות האיחוד על ידי סכום פשוט שכן סכימה זו כוללת בתוכה ספירה כפולה של מאורעות הנמצאים בחיתוך. נתבונן למשל בהטלת קוביה ונגדיר A – המאורע שהתוצאה זוגית ו-B – המאורע שהתוצאה היא לכל היותר 3. בכל אחד מהמאורעות יש 3 תוצאות אפשריות ולכן ההסתברות של כל אחד משני מאורעות אלו היא $\frac{3}{6}$. יחד עם זאת, התוצאה 2 מופיעה בשני המאורעות ולכן כאשר נבדוק את מאורע האיחוד נגלה שכלולות בו רק 5 התוצאות האפשריות $\{1,2,3,4,6\}$ ולכן הסתברות האיחוד היא $\frac{5}{6}$. סכום ההסתברויות מניב הסתברות גבוהה מדי ובכדי לתקן אותה יש להפחית את הסתברות החיתוך, שכן כל אחת מהתוצאות בחיתוך נספרה פעמיים.

תרגיל 2.3 (אי-שוויון בול (Boole)). הוכיחו את המקרה הכללי של התופעה שתוארה בפסקה האחרונה: אם $A_1, A_2, .., A_n$ מאורעות כלשהם אזי

$$Pr\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} Pr\left(A_i\right)$$

תרגיל 2.4 (אי-שוויון פרשט (Frechet)). הוכיחו: אם $A_1, A_2, .., A_n$ מאורעות כלשהם אזי

$$Pr\left(\bigcap_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} Pr\left(A_i\right) - n + 1$$

מבלי לדעת פרטים נוספים על ההסתברויות ניתן להשתמש באי-שיוויונות בול ופרשט (תרגילים 2.3 ו-2.4) בכדי לחשב חסמים להסתברות של איחוד וחיתוך מאורעות. אי-שוויון פרשט שימושי במיוחד שכן באמצעותו ניתן לחשב חסם מלרע לשכיחות תופעות באוכלוסיה שאין בידינו מידע נוסף עליהן. כך למשל, אם בכיתה מסוימת 70% מהילדים טובים בכדורגל ו-40% מהילדים טובים בכדורסל, מובטח לנו שלפחות $0.4+0.7-2+1 = 10\%$ מהילדים טובים בשני ענפי ספורט אלו (ולכל היותר?).

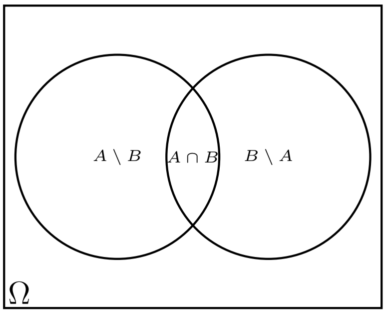
כאמור, סכום ההסתברויות של שני מאורעות יהיה גדול או שווה להסתברות של האיחוד, שכן תוצאות אפשריות של הניסוי המופיעות בחיתוך של המאורעות ייספרו פעמיים. על מנת לקבל את ההסתברות האמיתית, עלינו להפחית את ההסתברות של החיתוך מסכום ההסתברויות. נוסחא זו, המחשבת הסתברות של איחוד מאורעות תוך התחשבות בהסתברויות של החיתוכים נקראת נוסחת ההכלה וההפרדה[3] (Inclusion-Exculsion Principle).

***משפט 2.5** (נוסחת ההכלה וההפרדה לשני מאורעות). עבור כל שני מאורעות A, B מתקיים*

$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B) \qquad (2.5)$$

[3] או "נוסחת ההלכה וההדחה"

הוכחה. המאורע $A \cup B$ ניתן לרישום כאיחוד של שלושה מאורעות זרים בזוגות: "רק A התרחש" ($A \setminus B$), "רק B התרחש" ($B \setminus A$) ו"שניהם התרחשו ($A \cap B$) כמוראה באיור:



לפי האקסיומה השלישית הסתברות האיחוד תהיה:

$$Pr(A \cup B) = Pr(A \setminus B) + Pr(B \setminus A) + Pr(A \cap B) \quad \text{(2.6)}$$

לפי אותו השיקול (ראו נוסחא 1.14), גם את המאורע A ניתן לרשום כאיחוד של שני המאורעות הזרים $A \setminus B$ ו-$A \cap B$ ואת המאורע B כאיחוד של $B \setminus A$ ו-$A \cap B$, ולכן הסתברותם תהיה:

$$\begin{aligned} Pr(A) &= Pr(A \setminus B) + Pr(A \cap B) \\ Pr(B) &= Pr(B \setminus A) + Pr(A \cap B) \end{aligned}$$

לצורך קבלת נוסחא 2.5 נותר להציב את $Pr(A \setminus B)$ ואת $Pr(B \setminus A)$ משתי הנוסחאות האחרונות בנוסחא 2.6. $\square$

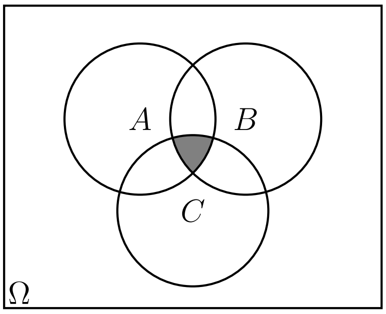
תרגיל 2.5. נתונים שני מאורעות המקיימים $Pr\left(A\right) = 0.7, Pr\left(B\right) = 0.8$. האם ייתכן כי

1. $Pr\left(A \cap B\right) = 0.4$?
2. $Pr\left(\overline{A} \cap B\right) = 0.3$?

ננסה להכליל את הנוסחא עבור שלושה מאורעות. בדומה לנוסחת ההכלה וההפרדה לשני מאורעות, הסכום של ההסתברויות $Pr(A) + Pr(B) + Pr(C)$ יהיה גדול מהסתברות האיחוד $Pr(A \cup B \cup C)$ שכן מאורעות הנמצאים בחיתוכים נספרים פעמיים. בהתאם לשיקול שהנחה אותנו מקודם, נפחית את הסתברויות של כל זוג חיתוכים ונקבל את הנוסחא הבאה:

$$\begin{aligned} Pr(A \cup B \cup C) = {} & Pr(A) + Pr(B) + Pr(C) \\ & - Pr(A \cap B) - Pr(A \cap C) - Pr(B \cap C) \end{aligned}$$

האם שוויון זה אכן מתקיים? בכדי שהשוויון יתקיים עלינו לוודא שכל תוצאה אפשרית בניסוי אכן נלקחה בחשבון פעם אחת בלבד ולא כך הדבר עבור התוצאות הנמצאות בחיתוך כל המאורעות $A \cap B \cap C$:



תוצאות אלו נסכמו שלוש פעמים בשורה הראשונה של המשוואה, כחלק מכל אחד מהמאורעות A, B, C אך גם הופחתו שלוש פעמים, כחלק מכל אחד מהחיתוכים של זוג מאורעות בשורה השנייה. לכן, על מנת להגיע להסתברות הנכונה עלינו להוסיף לסכום גם את ההסתברות של חיתוך שלושת המאורעות:

$$\begin{aligned} Pr(A \cup B \cup C) &= Pr(A) + Pr(B) + Pr(C) \\ &\quad -Pr(A \cap B) - Pr(A \cap C) - Pr(B \cap C) \\ &\quad +Pr(A \cap B \cap C) \end{aligned} \tag{2.7}$$

תרגיל 2.6. הוכיחו את נוסחת ההכלה וההפרדה ל-3 מאורעות.

באופן דומה ניתן לקבל נוסחא כללית עבור הכלה והפרדה של n מאורעות. בשלב הראשון נסכום את ההסתברויות של המאורעות. בשלב השני נפחית את ההסתברויות של חיתוכים של זוגות של מאורעות, אותם סכמנו פעמיים בשלב הראשון. בשלב השלישי, נוסיף בחזרה את ההסתברויות של חיתוכים של שלושה מאורעות, אותם הורדנו בשלב השני. בשלב הרביעי נחסיר חיתוכים של ארבע מאורעות, אותם ספרנו פעמיים בשלב הקודם וכן הלאה. הלוגיקה בה נבנית נוסחא זו מבהירה את שמה האנגלי, שכן בכל שלב אנחנו כוללים חיתוכים שלא כללנו בשלב הקודם או מחסירים ולא כוללים חיתוכים שספרנו יותר מדי פעמים בשלב הקודם. לסיכום, עבור n מאורעות $A_1, A_2, \ldots, A_n$ נקבל שהסתברות האיחוד הינה:

$$\begin{aligned} Pr\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{i=1}^{n} Pr(A_i) - \sum_{1 \le i < j \le n} Pr(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} Pr(A_i \cap A_j \cap A_k) \\ &\quad \ldots \quad +(-1)^{n-1} Pr\left(\bigcap_{i=1}^{n} A_i\right) \end{aligned} \tag{2.8}$$

לא נוכיח נוסחא זו וכמעט ולא נשתמש בה שכן שימושים בנוסחת ההכלה וההפרדה ליותר מ-3 מאורעות נוטים להיות ארוכים וטכניים מפאת מספר ההסתברויות הרב שיש לחשב. חמושים בנוסחאת ההכלה וההפרדה והמשפטים שהוכחו מאקסיומות ההסתברות, אנחנו יכולים עתה לחשב את ההסתברות של מאורעות מורכבים הנגזרים ממאורעות אחרים גם אם הם לא זרים כפי שביצענו בתרגילים 2.1 ו-2.2. נמחיש את השיטות השונות לפתרון בעיות באמצעות הבעיה הבאה:

בעיה 2.1. בכיתה מסוימת 0.7 מהילדים טובים בכדורגל, 0.4 מהילדים טובים בכדורסל וחצי מהילדים שטובים בכדורגל טובים גם בכדורסל. מה ההסתברות שילד שנבחר באקראי יהיה טוב רק בענף ספורט אחד?

השלב הראשון, בכל שיטה שבה נפתור את התרגיל, הוא להגדיר מאורעות ולהמיר את הסיפור המילולי לנוסחאות מתמטיות המתארות את הנתונים. נגדיר אם כך את A להיות המאורע שהילד טוב בכדורגל ו-B שהילד טוב בכדורסל. לפי הנתונים, $Pr(A) = 0.7$ וכן $Pr(B) = 0.4$. הנתון השלישי מתייחס לקבוצת הילדים שטובים בשני ענפי הספורט, כלומר לחיתוך בין המאורעות, ולכן $Pr\left(A\cap B\right) = \frac{0.7}{2} = 0.35$. המאורע שאת הסתברותו יש לחשב הוא המאורע שהילד טוב בענף ספורט אחד, קרי טוב בכדורגל אך לא בכדורסל או טוב בכדורסל אך לא בכדורגל, ולכן עלינו למצוא את $Pr\left(\left(A\cap\overline{B}\right)\cup\left(\overline{A}\cap B\right)\right)$.

השיטה האלגברית. השיטה הראשונה היא שיטה אלגברית. בשיטה זו ננסה לרשום את המאורע המבוקש כאיחוד, חיתוך ומשלים של מאורעות אחרים שאת הסתברותם אנו יודעים, וחישוב ההסתברות של המאורע המבוקש באמצעות הנוסחאות שפיתחנו. למשל, המאורע $A\cup B$ מתאר מצב שהילד טוב בלפחות בענף ספורט אחד וניתן לרשום אותו כאיחוד של שני מאורעות זרים - המאורע שהילד טוב רק בענף ספורט אחד, $\left(A\cap\overline{B}\right)\cup\left(\overline{A}\cap B\right)$, והמאורע שהילד טוב בשני ענפי ספורט, $A\cap B$. את הסתברות החיתוך אנחנו יודעים מהנתונים ואילו את הסתברות האיחוד ניתן למצוא באמצעות נוסחת ההכלה וההפרדה:

$$Pr(A\cup B) = Pr(A) + Pr(B) - Pr(A\cap B) = 0.4 + 0.7 - 0.35 = 0.75$$

לפיכך ההסתברות המבוקשת תהיה:

$$Pr\left(\left(A\cap\overline{B}\right)\cup\left(\overline{A}\cap B\right)\right) = Pr(A\cup B) - Pr(A\cap B) = 0.75 - 0.35 = 0.4$$

ניתן היה כמובן לחשב את ההסתברות המובקשת בדרכים אחרות. למשל, תוך שימוש בנוסחא **1.14** נקבל

$$Pr\left(A\cap\overline{B}\right) = Pr(A) - Pr(A\cap B) = 0.7 - 0.35 = 0.35$$

ובאופן דומה

$$Pr\left(\overline{A}\cap B\right) = Pr(B) - Pr(A\cap B) = 0.4 - 0.35 = 0.05$$

ומכאן

$$Pr\left(\left(A\cap\overline{B}\right)\cup\left(\overline{A}\cap B\right)\right) = Pr\left(A\cap\overline{B}\right) + Pr\left(\overline{A}\cap B\right) = 0.4$$

□

שיטת הטבלה. נבנה את הטבלה הבאה, המתארת את כל החיתוכים האפשריים של שני המאורעות A ו-B:

$\cap$	A	$\overline{A}$	
B	$Pr\left(A\cap B\right)$	$Pr\left(\overline{A}\cap B\right)$	$Pr(B)$
$\overline{B}$	$Pr\left(A\cap\overline{B}\right)$	$Pr\left(\overline{A}\cap\overline{B}\right)$	$Pr\left(\overline{B}\right)$
	$Pr(A)$	$Pr\left(\overline{A}\right)$	1

ראשית, נרשום את ההסתברויות הידועות לנו מתוך נתוני השאלה. לאחר מכן, נמלא את התאים האחרים תוך שימוש בעובדה שסכום שני התאים הראשונים בכל שורה צריך להיות שווה לתא האחרון בשורה וסכום שני התאים הראשונים בכל עמודה צריך להיות שווה לתא האחרון (למשל - $Pr(B) = Pr\left(A\cap B\right) + Pr\left(\overline{A}\cap B\right)$. ודאו שאכן שאר הנוסחאות מתקיימות). הטבלה ההתחלתית בשאלה זו תהיה:

	$\overline{A}$	A	$\cap$
0.4		0.35	B
			$\overline{B}$
1		0.7	

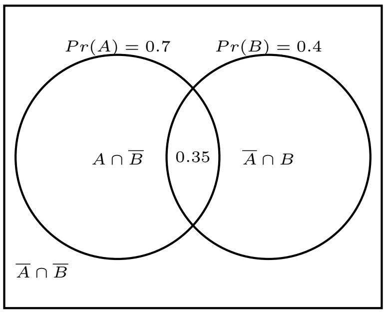
ולאחר מילוי שאר התאים נקבל את הטבלה הבאה:

	$\overline{A}$	A	$\cap$
0.4	0.05	0.35	B
0.6	0.25	0.35	$\overline{B}$
1	0.3	0.7	

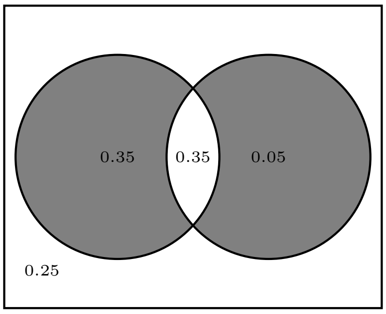
לאחר שכל התאים מלאים, כל שנותר לעשות הוא לזהות אילו מהתאים מתארים את המאורע המבוקש. במקרה זה המאורע הוא איחוד של $\overline{A} \cap B$ ושל $A \cap \overline{B}$ והיות והמאורעות זרים ההסתברות המבוקשת תהיה סכום ההסתברויות של מאורעות אלו כפי שמופיעות בטבלה, קרי $0.35 + 0.05 = 0.4$.

לרוב שיטה זו מהירה בהרבה מהשיטה האלגברית שכן לאחר ארגון הנתונים בטבלה, החישובים הבאים שיש לבצע הם מידיים ומהירים, וכל שנותר לעשות לאחר מילוי הטבלה הוא לזהות את המאורע המבוקש בשאלה. בנוסף, באמצעות שיטה זו אנחנו מחשבים בבת אחת את כל ההסתברויות שניתן לחשב בתרגיל, ולכן אם ישנם סעיפים נוספים בשאלה ניתן יהיה לענות עליהם בלי לפתח נוסחאות נוספות או לבצע חישובים נוספים. למשל, אם אנחנו מעוניינים לדעת מה ההסתברות שהילד הנבחר לא טוב באף אחד מענפי הספורט, כל שעלינו לעשות הוא להבין שמדובר על המאורע $\overline{A} \cap \overline{B}$ ולכן ההסתברות היא 0.25. בשיטה האלגברית הינו נאלצים לנסות להביע מאורע זה באמצעות מאורעות אחרים ולמעשה לפתור תרגיל נוסף וחדש. החיסרון העיקרי של שיטה זו הוא יכולתה לטפל רק שאלות עם 2 מאורעות, שכן עבור 3 מאורעות הטבלה צריכה להיות תלת מימדית ולהכיל את כל החיתוכים האפשריים של שלושת המאורעות. $\square$

שיטת דיאגרמת ון. נשרטט את דיאגרמת ון המתאימה לבעיה זו, ועל גבי הדיאגרמה נרשום את ההסתברויות הידועות לנו מהשאלה:



באופן דומה לשיטת הטבלה, נשתמש בנתונים על מנת לחשב את ההסתברויות של כל האזורים המופיעים בדיאגרמה. למשל, אנו יודעים שההסתברות של B כולו היא 0.4 ואילו ההסתברות של $A \cap B$ היא 0.35 ולכן ההסתברות של החלק שנותר מ-B צריכה להיות ההפרש, 0.05. באותה צורה ניתן לחשב את ההסתברות של החלק שנמצא ב-A ולא ב-$A \cap B$. בנוסף, שטח כל 4 התחומים הבסיסיים בדיאגרמה מסתכם ל-1 ולכן השטח של התחום שמחוץ ל-$A \cup B$ יהיה המשלים ל-1, כלומר 0.25. לאחר רישום ההסתברויות שחישבנו על הדיאגרמה, נקבל את השרטוט הבא:



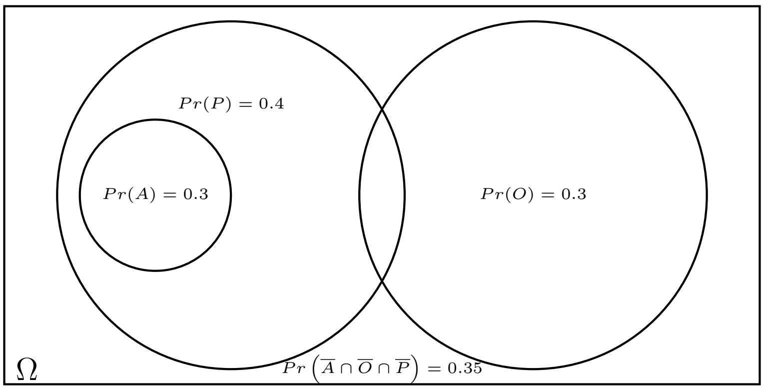
האזור הצבוע הוא בדיוק האזור המתאים למאורע המבוקש, המאורע שהילד הנבחר באקראי טוב רק בענף אחד מבין השניים, ולכן הסתברותו היא $0.4 = 0.05 + 0.35$. גם כאן, מרגע שחישבנו את ההסתברויות של כל התחומים, ניתן לענות על סעיפים נוספים בשאלה מבלי לבצע חישובים מורכבים נוספים, על ידי איתור התחומים המתארים את המאורע בלבד וסכימת ההסתברויות. למשל, אם נידרש למצוא את ההסתברות שהילד טוב רק בכדורסל, הרי שמדובר במאורע $\overline{A} \cap B$ והסתברותו 0.05. $\square$

העבודה עם דיאגרמת ון נוחה כמו העבודה עם הטבלה, אך בניגוד לטבלה, ניתן להשתמש בדיאגרמות ון גם לצורך פתרון תרגילים בהם מופיעים יותר מ-2 מאורעות. לפיכך, זו תהיה השיטה העיקרית והעדיפה מבחינתנו לפתרון תרגילים. נמחיש את נוחותה על ידי פתרון תרגיל נוסף הכולל שלושה מאורעות.

בעיה 2.2. בסקר שנערך בקרב סטודנטים באוניברסיטה התגלה ש-30% מהם אוהבים פיצה עם אננס, 40% אוהבים פיצה עם פפרוני ו-30% אוהבים פיצה עם בצל. בנוסף, התגלה שכל הסטודנטים שאוהבים פיצה עם אננס אוהבים גם פיצה עם פפרוני, שאין סטודנטים שאוהבים גם אננס וגם בצל וכן ש-35%מהסטודנטים כלל לא אוהבים פיצה. בוחרים באקראי את אחד הסטודנטים.

1. מה ההסתברות שהוא אוהב פיצה עם פפרוני בלבד?
2. מה ההסתברות שהוא אוהב בדיוק $k \in \{0, 1, 2, 3\}$ סוגי תוספות?

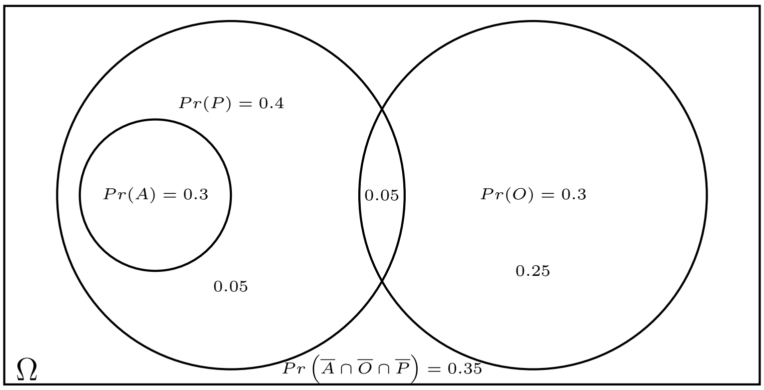
הוכחה. נסמן ב-A את המאורע שהסטודנט אוהב פיצה עם אננס, ב-P את המאורע שהסטודנט אוהב פיצה עם פפרוני, ב-O את המאורע שהסטודנט אוהב פיצה עם בצל ונשרטט את דיאגמת ון המאתימה. מאחר וכל הסטודנטים האוהבים אננס אוהבים גם פפרוני מתקיים $A \subseteq P$. בנוסף, אין סטודנטים האוהבים גם אננס וגם בצל ולכן $A \cap O = \emptyset$. לכן בשרטוט העיגול המתאים ל-A יהיה כולו בתוך P ולא יהיה חיתוך בינו לבין O:



אם נתעלם מהמאורע A, נקבל בעיה עם שני מאורעות P, O המקיימים $Pr\left(\overline{P} \cap \overline{O}\right) = Pr\left(\overline{P \cup O}\right) = 0.35$ ולכן $Pr\left(P \cup O\right) = 1 - Pr\left(\overline{P \cup O}\right) = 0.65$. נשתמש בנוסחת ההכלה וההפרדה ונקבל

$$Pr\left(P\cap O\right)=Pr\left(O\right)+Pr\left(P\right)-Pr\left(P\cup O\right)=0.4+0.3-0.65=0.05$$

לפיכך ניתן להשלים את השרטוט עם התוצאות המיידיות $Pr\left(O\setminus P\right)=0.25$ וכן $Pr\left(P\setminus\left(A\cup O\right)\right)=0.05$:



לאחר שחושבו ההסתברויות של כל המאורעות הבסיסיים, ניתן לענות על השאלה.

1. ההסתברות שהסטודנט אוהב פיצה עם פפרוני בלבד הוא התחום הנמצא ב-P ולא באף מאורע אחר וההסתברות המתאימה הינה 0.05.
2. לפי הנתונים, הסיכוי שהוא אוהב בדיוק תוספות הוא **0** (אין חיתוך בין כל שלושת המאורעות) והסיכוי שהוא אוהב בדיוק **0** תוספות הוא 0.35 (זה התחום שמחוץ לכל המאורעות). הוא אוהב בדיוק סוג אחד של תוספת כאשר הוא אוהב רק פפרוני (הסתברות של 0.05) או כאשר הוא אוהב רק בצל (הסתברות של 0.25). מאורעות אלו זרים ולכן ההסתברות שהוא אוהב רק סוג אחד של תוספת הינו $0.05+0.25=0.3$. המאורע שהוא אוהב בדיוק שני סוגי תוספות מתרחש כאשר הוא אוהב אננס ופפרוני או פפרוני ובצל (אף אחד לא אוהב אננס ובצל) והסיכוי לכך הינו $0.3+0.05=0.35$.
   מאחר ואלו בדיוק כל האפשרויות (k חייב להיות **0** או **1** או **2** או **3** ורק אחד מהם) ולכן סכום ההסתברויות צריך להיות **1**, כפי שאכן מתקיים: $0+0.35+0.3+0.35=1$.

□

תרגיל 2.7. סקר שנערך באוניברסיטה גילה ש-30% מהסטודנטים לא צורכים ריטלין, 50% מהסטודנטים לא שותים קפה ואילו 35% מהסטודנטים שותים קפה וצורכים ריטלין. כמה סטודנטים לא שותים קפה ולא צורכים ריטלין?

תרגיל 2.8. לסנטה קלאוס שלושה אילים שמושכים את מזחלתו. שני אילים רגילים (ויקסן וקומט) ואיל עם אף אדום (רודולף). בכדי לעוף ברחבי העולם ולחלק מתנות לילדים, יש צורך לפחות בשני אילים בריאים. לאורך השנים סנטה קלאוס שם לב שכל אחד מהאילים הרגילים בריא בהסתברות 0.55 ושניהם חולים בו-זמנית בהסתברות 0.05. בנוסף, הסיכוי שלפחות אחד משלושת האילים בריא הינו 0.95, הסיכוי של רודולף להיות בריא הינו 0.25 והסיכוי של כל שני איילים להיות בריאים זהה לכל הזוגות. בסך הכל סנטה קלאוס יכול לעוף ברחבי העולם ב-35% מהזמן.

1. מה ההסתברות שרק רודולף בריא?
2. מה ההסתברות שכל האילים בריאים?

תרגיל 2.9. במבחן הסיום בקורס יש 3 שאלות. כל סטודנט ענה נכון על שאלה אחת לפחות לפי הטבלה הבאה:

שאלות	1	2	3	$1+2$	$1+3$	$2+3$	$1+2+3$
מספר העונים נכון	76	72	98	25	41	27	10

כמה סטודנטים נבחנו במבחן?

2.2 הסתברות מותנית

לעיתים בעת ביצוע ניסוי אנחנו עשויים לגלות מידע חלקי אודות תוצאת הניסוי או אופן התפתחותו. מידע זה עשוי לשנות את ההסתברויות של המאורעות השונים, שכן מרחב המדגם השתנה. נדגים זאת על ידי שימוש בנתוני הקליעות של משחק מגמר ה-NBA בין גולדן-סטייט ווריורס לבין קליבלנד קאבאלירס. במשחק זה גולדן-סטייט ניצחה וזכתה באליפות לשנת 2017 ובמהלכו נזרקו 51 זריקות עונשין לפי הפירוט הבא:

קבוצה	זריקות	קליעות	אחוזי קליעה
גולדן-סטייט	28	23	82.1%
קליבלנד	23	15	65.2%
סה"כ	51	38	74.5%

הניסוי שאנחנו מבצעים הינו בחירה אקראית של אחת מזריקות העונשין שנזרקה במשחק והמאורע המעניין אותנו הוא המאורע A, המאורע שנקלע סל באותה זריקה. בהתאם לטבלה, במשחק נזרקו 51 זריקות ונקלעו 38 סלים, לכן ההסתברות שהזריקה שנבחרה באקראי תוביל לסל היא $\frac{38}{51}$. נניח שאנחנו מגלים שהזריקה שנבחרה נזרקה על ידי שחקן של גולדן-סטייט. במקרה כזה, כבר לא כל 51 הזריקות אפשריות אלא רק 28 הזריקות של גולדן-סטייט. 23 מתוך הזריקות הללו נכנסו לסל לכן ההסתברות המעודכנת לכך שהזריקה נכנסה לסל בהינתן שנזרקה על ידי שחקן של גולדן-סטייט הינה $\frac{23}{28}$. באותו אופן אם אנחנו מגלים שהזריקה נזרקה על ידי שחקן של קליבלנד ההסתברות לסל תשתנה ל-$\frac{15}{23}$. תוספת המידע אודות זהות הקבוצה הזורקת שינתה את של המאורע מאחר ולקבוצות שונות סיכוי קליעה שונים.

נכליל את הדיון ונשתמש בגישה השכיחותית בכדי לפתח את הנוסחא הכללית לכך שמאורע A מתרחש כאשר ידוע ש-B התרחש. נניח שחוזרים על הניסוי n פעמים. ב-$S_n\left(B\right)$ מהחזרות המאורע B התרחש. אם ידוע שאנחנו באחת מהחזרות הללו בהן המאורע B התרחש, הדרך היחידה בה המאורע A יכול להתרחש היא אם המאורע $A\cap B$ התרחש, וזה קורה $S_n\left(A\cap B\right)$ פעמים. לכן, מספר הפעמים שבהן המאורע A מתרחש מתוך הפעמים בהן B מתרחש הינו

$$\lim_{n\to\infty}\frac{S_n\left(A\cap B\right)}{S_n\left(B\right)}=\lim_{n\to\infty}\frac{S_n\left(A\cap B\right)/n}{S_n\left(B\right)/n}=\frac{\lim_{n\to\infty}\left(S_n\left(A\cap B\right)/n\right)}{\lim_{n\to\infty}\left(S_n\left(B\right)/n\right)}=\frac{Pr\left(A\cap B\right)}{Pr(B)}$$

נשים לב שזהו למעשה החישוב שביצענו בדוגמת הכדורסל. הסיכוי לבחור זריקה של גולדן-סטייט הינו $\frac{28}{51}$ ואילו הסיכוי לבחור זריקה של גולדן-סטייט שנגמרה בסל (זהו מאורע החיתוך) הינו $\frac{23}{51}$ ולכן ההסתברות שנכנס סל בהינתן שבחרנו בזריקה של גולדן-סטייט היא $\frac{23}{28}=\frac{23/51}{28/51}$, כמקודם.

הגדרה 2.1. יהיו $A,B\subseteq\Omega$ שני מאורעות כך ש-$Pr\left(B\right)\neq0$. ההסתברות המותנית של המאורע A בהינתן המאורע B, קרי הסיכוי ש-A התקבל בניסוי בו ידוע שהמאורע B התרחש, מסומנת על ידי $Pr\left(A|B\right)$ ומחושבת על ידי

$$Pr\left(A|B\right)=\frac{Pr\left(A\cap B\right)}{Pr\left(B\right)} \qquad (2.9)$$

הסתברות זו נקראת בפשטות "ההסתברות של A בהינתן B".

ההגבלה למאורעות $Pr\left(B\right)\neq 0$ אינה רק הגבלה טכנית המונעת מאיתנו לחלק באפס אלא נובעת מכך שמאורעות עם הסתברות 0 לא יכולים להתרחש לעולם ולכן השאלה חסרת משמעות. למשל, הסיכוי שתיבחר זריקה של שחקן שיקאגו בולס במשחק בין גולדן-סטייט לקליבלנד הוא 0 ולכן אין משמעות לשאלה "מה הסיכוי שנקלע סל בהינתן שנבחרה זריקה של שחקן שיקאגו". לא ייתכן שנבחרה זריקה של שחקן שיקאגו ואין זריקות של שחקן כזה שנכנסו לסל! לכן, בכל פעם שנחשב הסתברויות מותנות או נדבר על מאורעות בהקשר זה, נניח שכל המאורעות עליהם מתנים הם מאורעות עם הסתברות חיובית.

עבור כל מאורע $B\subseteq\Omega$ שהסתברותו חיובית פונקציית ההסתברות המותנית $Pr\left(\cdot|B\right)$ המתאימה לכל $A\subseteq\Omega$ את $Pr\left(A|B\right)$ היא פונקציית הסתברות, כלומר מקיימת את שלושת אקסיומות ההסתברות של קולמוגורוב (ודאו זאת!):

1. הסתברות מותנית היא חיובית, כלומר לכל מאורע $A\subseteq\Omega$ מתקיים $Pr\left(A|B\right)\geq 0$.
2. ההסתברות המותנית של מרחב המדגם כולו היא $Pr\left(\Omega|B\right)=1$.
3. עבור מאורעות זרים בזוגות $A_1,A_2,...A_n\subseteq\Omega$ ההסתברות המותנית של האיחוד מקיימת

$$Pr\left(\bigcup_{i=1}^{n}A_i\middle|B\right)=\sum_{i=1}^{n}Pr\left(A_i|B\right) \tag{2.10}$$

לאור זאת, כל המשפטים שהוכחנו בפרק 2.1 תקפים גם עבור פונקציית ההסתברות המותנית על B. למשל, מסקנה 2.2 תלבש את הצורה המותנית הבאה:

$$Pr\left(\overline{A}|B\right)=1-Pr\left(A|B\right) \tag{2.11}$$

ובאופן זהה ישתנו שאר הנוסחאות וניתן באמצעותן לבצע חישובי הסתברויות כפי שעשינו בתרגילים הקודמים, כשההבדל היחיד הוא שכל ההסתברויות מותנות על אותו המאורע B, שלמעשה הפך למרחב המדגם החדש.

תרגיל 2.10. האם הנוסחאות הבאות נכונות? אם כן - הוכיחו אותן. אם לא - תנו דוגמא לניסוי בו הן אינן מתקיימות.

1. אם $Pr(A)=Pr(B)\neq 0$ אז $Pr(A|B)=Pr(B|A)$.
2. אם $Pr(A|B)=Pr(B|A)$ אז $Pr(A)=Pr(B)$.
3. $Pr(A|B)+Pr(A|\overline{B})=1$.
4. $Pr(A|B)+Pr(\overline{A}|B)=1$.
5. אם $Pr\left(A\right)<Pr(B)$ ואם $Pr(C)\neq 0$ אז גם $Pr\left(A|C\right)<Pr\left(B|C\right)$

לצורך המחשת העבודה עם ההסתברות המותנית, נחזור למרחב המדגם הנתון בתרגיל 2.1, נגדיר את המאורעות A- המספר הנבחר מתחלק ב-3 ו-B - המספר הנבחר גדול מ-2, ונחשב את ההסתברות המותנית $Pr\left(A|B\right)$, קרי ההסתברות שהמספר הנבחר מתחלק ב-3 כאשר ידוע שהוא גדול מ-2. במהלך פתרון התרגיל חישבנו את ההסתברויות של שני המאורעות ומצאנו כי $Pr(A)=\frac{126}{385}$ וכן $Pr(B)=\frac{380}{385}$. כמו כן, היות ואין במרחב המדגם מספרים קטנים מ-2 המתחלקים ב-3, הרי שהחיתוך בין "להיות גדול מ-2" לבין "להתחלק ב-3" הוא "להתחלק ב-3", כלומר $A\cap B=A$ וההסתברות המותנית הינה $Pr\left(A|B\right)=\frac{126/385}{380/385}=\frac{126}{380}$.

שימו לב כי $Pr\left(A|B\right) > Pr(A)$, כלומר העבודה שאנחנו יודעים שהתוצאה גדולה מ-2 הגדילה את ההסתברות שהתוצאה מתחלקת ב-3. קשר זה אינו מפתיע שהרי מבין 10 התוצאות האפשריות של הניסוי, כל התוצאות המתחלקות ב-3 כלולות ב-B. לכן, אם ידוע שהמאורע B התרחש, סימן שהתוצאות 1 ו-2 לא נבחרו ולכן הסיכוי שנבחר כל אחד מהמספרים האחרים, ובפרט הסיכוי שנבחר אחד המספרים המתחלקים ב-3 גדל. בדומה, ההסתברות שהמספר גדול מ-2 בהינתן שהוא מתחלק ב-3 היא $Pr\left(B|A\right) = 1$, שהרי אם ידוע שהמספר הנבחר מתחלק ב-3 אזי בוודאות הוא גם גדול מ-2.

בדוגמא האחרונה ראינו שמידע נוסף על תוצאת הניסוי יכול להגדיל את ההסתברויות הלא מותנות של מאורעות. באופן דומה, מידע נוסף יכול גם להקטין את ההסתברויות הללו. למשל, ההסתברות שהמספר הנבחר לא מתחלק ב-3 בהינתן שהמספר הנבחר גדול מ-2 תהיה $Pr\left(\overline{A}|B\right) = 1 - Pr\left(A|B\right) = \frac{254}{380}$ והסתברות זו קטנה מההסתברות שהמספר הנבחר לא מתחלק ב-3, שהיא $Pr\left(\overline{A}\right) = \frac{259}{385}$. בנוסף, $Pr\left(\overline{B}|A\right) = 0$ שהרי אם המספר הנבחר מתחלק ב-3, אזי לא ייתכן שהוא קטן מ-2, והסתברות זו קטנה מההסתברות שהמספר הנבחר קטן מ-2, $Pr\left(\overline{B}\right) = \frac{5}{385}$.

מתוך הגדרת ההסתברות המותנית ניתן לקבל נוסחא לחישוב הסתברות של חיתוך מאורעות.

***משפט* 2.6** (חוק הכפל). *עבור שני מאורעות עם הסתברות חיובית מתקיים:*

$$Pr\left(A \cap B\right) = Pr\left(A|B\right) \cdot Pr(B) = Pr\left(B|A\right) \cdot Pr(A) \tag{2.12}$$

נוסחא זו ניתנת להכללה לחיתוך של מספר רב של מאורעות. עבור מאורעות $A_1, A_2, ..., A_n$ שאינם זרים מתקיים

$$Pr\left(\bigcap_{i=1}^{n} A_i\right) = Pr\left(A_1\right) \cdot Pr\left(A_2|A_1\right) \cdot \ldots \cdot Pr\left(A_n|\bigcap_{i=1}^{n-1} A_i\right) \tag{2.13}$$

הוכחה. עבור שני מאורעות התוצאה טריוויאלית ומתקבלת ישירות מההגדרה. עבור שלושה מאורעות A, B, C נקבל:

$$Pr\left(A \cap B \cap C\right) = Pr\left(\left(A \cap B\right) \cap C\right) = Pr\left(A \cap B\right) \cdot Pr\left(C|A \cap B\right) \tag{2.14}$$

כאשר במעבר האחרון השתמשנו בנוסחא 2.12 עבור שני המאורעות C ו-$A \cap B$. נשתמש בנוסחא זו שנית לחישוב $Pr\left(A \cap B\right) = Pr\left(B|A\right) \cdot Pr(A)$ ולפיכך

$$Pr\left(A \cap B \cap C\right) = Pr(A) \cdot Pr\left(B|A\right) \cdot Pr\left(C|A \cap B\right) \tag{2.15}$$

כנדרש. ההוכחה עבור מספר כללי של מאורעות תבוצע באינדוקציה באופן דומה ולכן נשמיטה. □

מבחינה אינטואיטיבית, אפשר לחשוב על חוק הכפל כעל הכנסה של סדר כרונולוגי להתרחשות של אירועים. כדי ששלושת המאורעות A, B ו-C יתרחשו גם יחד, אפשר קודם כל לדרוש שמאורע A יתרחש, לאחר מכן לדרוש שמאורע B יתרחש כאשר A כבר התרחש ולסיום לדרוש ש-C יתרחש בהינתן ששני המאורעות הקודמים התרחשו. כמובן שסדר כרונולוגי זה לא בהכרח מתאר "זמן" או "שלבים" שבאמת קיימים בניסוי ובמציאות,

אבל במקרים רבים נוח לחשוב על הניסוי כניסוי המתקדם בשלבים כשבכל שלב אנחנו מחשבים את ההסתברות שהמאורע בשלב זה יתרחש, בהינתן שהמאורעות של השלבים הקודמים אכן התרחשו.

חוק הכפל מציג גישה נוספת לשימוש בהסתברות מותנית. בגישה זו, במקום להתנות על מאורעות שאנחנו יודעים בוודאות שהתרחשו, אנחנו מתנים על מאורע שלאו דווקא התרחש, אך ההתניה עליו מפשטת את החישוב או נחוצה לחישוב. במקרה זה, אנחנו מעוניינים לחשב את ההסתברות שגם מאורע A וגם מאורע B יתרחשו בו זמנית, מבלי שאנחנו יודעים דבר על תוצאת הניסוי. במידה ואחת ההסתברויות המותנות ניתנת לחישוב, למשל A בהינתן B, אזי ניתן יהיה לחשב את ההסתברות של החיתוך באמצעות חוק הכפל על ידי חישוב ההסתברות $Pr\left(A|B\right)$ והכפלתה ב-$Pr(B)$. ננקוט בגישה זו כאשר מידע נוסף מסייע בחישוב ההסתברות – נניח שהוא מצוי ברשותנו, נחשב את ההסתברות המותנית ובאמצעותה נחשב את ההסתברות של המאורע הדרוש. שימוש חשוב לשיטה זו מוצג בסעיף 2.2.1.

תרגיל 2.11. מתוך חפיסת קלפים שולפים 3 קלפים בזה אחר זה. מה הסיכוי שהקלפים שנבחרו הם נסיך, מלכה ומלך בדיוק בסדר הזה?

תרגיל 2.12. בכד 3 כדורים שחורים ו-5 לבנים. בכל שלב מוציאים כדור אחד ומחזירים אותו בתוספת 5 כדורים נוספים מאותו הצבע. מוציאים מהכד 4 כדורים בסך הכל, מה ההסתברות שכולם שחורים?

2.2.1 נוסחת ההסתברות השלמה

נפתח בדוגמא שתמחיש את נוסחת ההסתברות השלמה, הצורך בה ואופן השימוש בה.

בעיה 2.3. כדורסלן יכול לזרוק לסל משני מקומות על המגרש. כאשר השחקן עומד על קו העונשין, הסיכוי שלו לקלוע הינו 0.9. כאשר הכדורסלן עומד על קשת השלוש, הסיכוי שלו לקלוע הוא 0.4. בכדי להחליט מאין לזרוק, הכדורסלן מטיל קוביה: אם יצא 3 הוא זורק מקשת השלוש ואחרת – מקו העונשין. מה הסיכוי שיקלע לסל?

בניגוד לתרגילים שפתרנו עד כה, בבעיה זו הניסוי הוא ניסוי דו-שלבי והתוצאות של השלב השני (הזריקה לסל) תלויות בתוצאות של השלב הראשון (הטלת הקוביה). אם היינו יודעים מהי תוצאת הטלת הקוביה בסיבוב הראשון, הפתרון היה מיידי אבל ידע זה אינו מצוי בידינו ולכן צריך למצוא דרך לשלב אותו בפתרון ולהתייחס לכל התוצאות האפשריות בהטלת הקוביה.

הוכחה. נגדיר את המאורעות: B – המאורע שהכדורסלן קלע סל ו-T– המאורע שבקוביה יצא 3. המאורע B יכול להתרחש בשתי דרכים שונות: סל וגם בקוביה יצא 3 $(B\cap T)$ או סל וגם בקוביה לא יצא 3 $(B\cap\overline{T})$. מאורעות אלו זרים ואיחודם הינו B ולכן הסיכוי לסל הינו:

$$Pr(B)=Pr\left(B\cap T\right)+Pr\left(B\cap\overline{T}\right) \qquad (2.16)$$

בכדי לחשב כל אחד מהחיתוכים נשתמש בחוק הכפל (משפט 2.6), שכן אנו יודעים מהנתונים את שתי ההסתברויות המותנות $Pr\left(B|T\right)=0.4$ ו-$Pr\left(B|\overline{T}\right)=0.9$. נציב כל זאת בנוסחא הקודמת ונקבל:

$$\begin{aligned}Pr(B) &= Pr\left(B|T\right)\cdot Pr(T)+Pr\left(B|\overline{T}\right)\cdot Pr\left(\overline{T}\right)\\ &= 0.4\cdot\frac{1}{6}+0.9\cdot\frac{5}{6}=\frac{49}{60}\end{aligned}$$

$\square$

בחישוב זה ביצענו, למעשה, התניה על השלב הראשון של הניסוי. לו הינו יודעים את תוצאת הקוביה בשלב הראשון אזי היינו יודעים את ההסתברות לסל בשלב השני. היות ואנחנו לא יודעים את תוצאת הקוביה, אנחנו מתנים על כל האפשרויות שיכולות להיות ומבצעים סכימה משוקללת של הסתברויות אלו, כאשר המשקולות הן ההסתברויות לכל תוצאה אפשרית של השלב הראשון. באופן זה, אנחנו עוברים על כל התרחישים בניסוי שיכולים להיגמר בסל וסוכמים את ההסתברויות של כל תרחיש כזה. הנוסחא האחרונה שקיבלנו נקראת נוסחת ההסתברות השלמה והוכחתה לכל מספר של אירועים נתונה במשפט הבא:

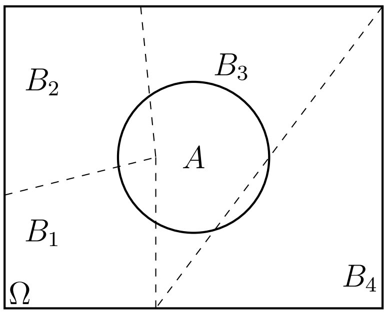
***משפט* 2.7** (נוסחת ההסתברות השלמה)**.** *יהיו $B_1, \ldots, B_n$ (ייתכן גם $n = \infty$) מאורעות זרים בזוגות שאיחודם הוא מרחב המדגם כולו. אזי ההסתברות של מאורע A ניתנת לחישוב על ידי*

$$Pr(A) = \sum_{i=1}^{n} Pr\left(A|B_i\right) \cdot Pr\left(B_i\right) \tag{2.17}$$

כאשר $n = 2$ למעשה מתנים על מאורע בודד ומשלימו:

$$Pr(A) = Pr\left(A|B\right) \cdot Pr(B) + Pr\left(A|\overline{B}\right) \cdot Pr\left(\overline{B}\right) \tag{2.18}$$

הרעיון מאחורי הנוסחא וההוכחה מומחש בשרטוט הבא:



המאורעות B_1, B_2, B_3, B_4 זרים בזוגות ומחלקים את מרחב המדגם ל-4 חלקים. כתוצאה מכך הם מחלקים גם את המאורע A, כך שהחלק המשותף למאורע A ולמאורע B_i הינו $A \cap B_i$. חיתוכים אלו זרים בזוגות גם כן ואיחודם הוא כל A ולכן אם נחשב את ההסתברות שלהם נקבל את ההסתברות של A על ידי סכימה. ההסתברות של כל אחד מהחיתוכים ניתנת לחישוב על ידי התנייה על המאורע B_i באמצעות חוק הכפל (נוסחא 2.12) ומכאן מתקבלת נוסחת ההסתברות השלמה. נרשום את ההוכחה הכללית בצורה מפורשת:

הוכחה. לכל i נגדיר $C_i = A \cap B_i$. מאחר והמאורעות B_i זרים בזוגות כך גם המאורעות C_i וכן מתקיים $C_i \subseteq B_i$. נחשב את איחוד כל המאורעות:

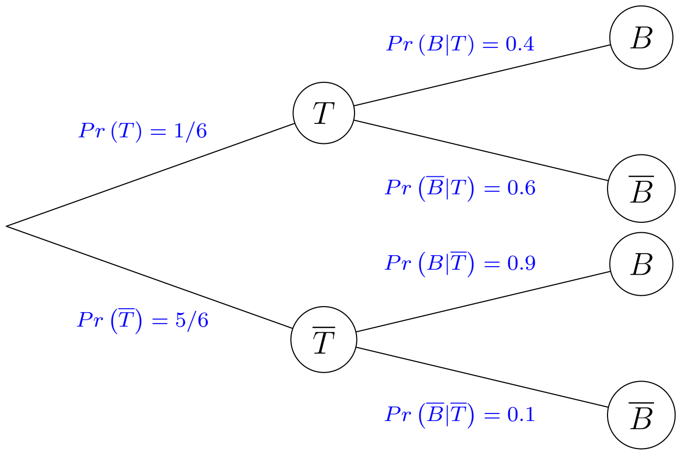
$$\begin{aligned} \bigcup_{i=1}^{n} C_i &= \bigcup_{i=1}^{n} \left(A \cap B_i\right) \overset{eq.1.7}{=} A \cap \left(\bigcup_{i=1}^{n} B_i\right) \\ &= A \cap \Omega = A \end{aligned}$$

לכן ההסתברות של המאורע A תתקבל על ידי

$$\begin{aligned}Pr(A) &= Pr\left(\bigcup_{i=1}^{n} C_i\right) = \sum_{i=1}^{n} Pr\left(C_i\right)\\ &= \sum_{i=1}^{n} Pr\left(A\cap B_i\right) \overset{eq.2.12}{=} \sum_{i=1}^{n} Pr\left(A|B_i\right)\cdot Pr\left(B_i\right)\end{aligned}$$

כנדרש. □

דרך נוחה להציג את ההשתלשלות של ניסוי רב-שלבי ולחשב את ההסתברויות של המאורעות השונים היא באמצעות שרטוט עץ. בשיטה זו, קודקודים מתארים תוצאות אפשריות של שלבי הניסוי והקשתות מתארות את המעברים בין השלבים. שורש העץ הוא התחלת הניסוי וממנו יוצאות קשתות המחברות אותו לתוצאות האפשריות של השלב הראשון. מכל תוצאה אפשרית של השלב הראשון יוצאות קשתות אל תוצאות אפשריות של השלב השני הנובעות ממנו וכן הלאה. העלים מתארים את התוצאות הסופיות של הניסוי. על כל קשת נרשום את ההסתברות המותנית לעבור לתוצאה של השלב הבא בהינתן השלב בו אנחנו נמצאים כעת. השרטוט הבא מציג את העץ המתאים לדוגמת הכדורסל:



בהתאם לחוק הכפל, ההסתברות להגיע לכל עלה היא מכפלת ההסתברויות המותנות המובילות לעלה זה. למשל, ההסתברות להגיע לעלה B העליון (וזהו למעשה $B\cap T$) היא $\frac{1}{6}\cdot 0.4$ וההסתברות להגיע לעלה B התחתון (שהוא $B\cap\overline{T}$) היא $\frac{5}{6}\cdot 0.9$. בכדי לחשב את ההסתברות של המאורע B יש למצוא את כל המסלולים המובילים לעלים עליהם רשום B ולסכום את ההסתברויות הללו – זו למעשה נוסחת ההסתברות השלמה.

נשתמש לרוב בנוסחת ההסתברות השלמה כאשר אנחנו מעוניינים לחשב הסתברות של מאורע שאיננו יודעים לחשב ישירות שכן "חסר לנו נתון" – מידע על מאורע אחר. לו היינו יודעים אם מאורע אחר התרחש או לא, החישוב של ההסתברות המותנית היה אפשרי. במקרה כזה, ניתן להשתמש בנוסחא ולהוסיף את הנתון החסר – פעם אחת נתנה על כך שהמאורע קרה ובפעם הנוספת נתנה על כך שהמאורע לא קרה. בדוגמת הכדורסלן, כדי לדעת את הסתברות הקליעה אנחנו צריכים נתון נוסף – מאין הוא זורק (או ליתר דיוק, את תוצאת ההטלה של הקוביה). לכן אנחנו מוסיפים נתון זה על ידי כך שאנחנו מתנים פעם אחת על כך שהקוביה נפלה על 3 ופעם אחת על כך שלא נפלה על 3.

תרגיל 2.13. במחקר (אמיתי) שנערך בכדי לבדוק את תופעת "היד החמה" בכדורסל, התגלה כי ישנו קשר בין זריקות עוקבות של השחקן לארי בירד. לאחר סל מוצלח, הסיכוי שלארי בירד יקלע גם את הזריקה הבאה הוא 0.49 ואילו לאחר החטאה, הסיכוי שיקלע הינו 0.38.

1. השתמשו בנוסחת ההסתברות השלמה על מנת לחשב את הסיכוי שזריקה כלשהי של לארי בירד תיכנס לסל.
2. במחקר התגלה כי הסיכוי לסל של לארי בירד שונה מהמספר אותו חישבתם בסעיף הקודם. שערו - מדוע יכולים להיות הבדלים?
3. לארי בירד קולע את הזריקה הראשונה שלו במשחק. מה הסיכוי שיקלע גם את הזריקה השלישית?

תרגיל 2.14. בשק יש n פתקים הממוספרים מ-1 עד n. מוציאים פתק באקראי, רושמים את מספרו, מחזירים לשק ושולפים פתק נוסף.

1. מה ההסתברות שהפתק השני יהיה גדול יותר מהפתק הראשון? בדקו את התשובה עבור $n = 2$ על ידי רישום מפורש של כל האפשרויות.
2. חזרו על הסעיף הקודם כאשר הבחירה נעשית ללא החזרה.

תרגיל 2.15. מטבע לא הוגן הנותן עץ בהסתברות p מוטל שוב ושוב עד אשר מתקבל עץ לראשונה. מה ההסתברות שהניסוי נעצר לאחר מספר זוגי של הטלות?
רמז: השתמשו בנוסחת ההסתברות השלמה והתנו על תוצאת ההטלה הראשונה.

בתרגיל 2.15 ראינו טכניקה נוספת לשימוש בנוסחת ההסתברות השלמה. בתרגיל זה, הניסוי המתואר הוא מחזורי וחזרתי, שכן הניסוי החל מההטלה השניה דומה מאוד לניסוי החל מההטלה הראשונה, עם הבדל קל בהגדרת המאורע. במצב זה, ניתן להשתמש בנוסחת ההסתברות השלמה כדי להתנות על ההטלה של הניסוי ולהגיע למשוואה המתארת את ההסתברות של המאורע המבוקש. בהמשך נראה דרך נוספת לטפל בשאלות כאלו תוך חישוב ההסתברות של העץ להופיע לראשונה בכל אחת מההטלות וסכימה של כל האפשרויות המתאימות לתוצאה זוגית, כפי שמופיע בתרגיל 6.31.

תרגיל 2.16. חוזרים על ניסוי בו יכולים להתקבל בין היתר המאורעות הזרים A ו-B שוב ושוב, כאשר ההסתברות לכל מאורע נתונה וקבועה. מה ההסתברות שהפעם הראשונה שהמאורע A יקרה תהיה לפני הפעם הראשונה בה המאורע B יקרה?

2.2.2 תלות וחוסר תלות בין מאורעות

עד כה חישבנו הסתברויות של מאורעות והסתברויות מותנות של מאורעות וראינו שמידע נוסף על תוצאות הניסוי יכול להשפיע על ההסתברות. למשל, בהטלת קוביה הסיכוי לקבל 1 הוא $\frac{1}{6}$. אם ידוע שהתוצאה שהתקבלה אי-זוגית, סיכוי זה משתנה והופך ל-$\frac{1}{3}$ ואם ידוע שהתקבלה תוצאה זוגית הסיכוי יהיה כעת 0.[4] בין המאורעות "לקבל 1 בקוביה", ו"לקבל תוצאה אי-זוגית" ישנו קשר הסתברותי, המתבטא בכך שההסתברות של כל מאורע שונה מההסתברות המותנית של המאורע בהינתן המאורע השני. קשר זה מכונה בתורת ההסתברות "תלות בין מאורעות". קיימת גם האפשרות שהמידע הנוסף לא משנה את ההסתברויות של המאורעות האחרים. במקרה כזה נאמר שהמאורעות בלתי תלויים.

הגדרה 2.2. המאורעות A, B נקראים בלתי תלויים אם מידע על התרחשותו של אחד המאורעות לא משנה את ההסתברות שהמאורע השני יתרחש, כלומר אם מתקיימים שני התנאים הבאים:

[4] למעשה, מאורעות זרים תמיד יהיו תלויים, שהרי אינם יכולים להתרחש בו זמנית ולכן אם ידוע שאחד מהם התרחש, הסיכוי שהשני התרחש יהיה אפס. ראו תרגיל 2.20

1. $Pr\left(A|B\right)=Pr\left(A\right)$.
2. $Pr\left(B|A\right)=Pr\left(B\right)$.

אחרת, המאורעות יקראו תלויים.

בעיה 2.4. בוחרים באקראי קלף אחד מתוך חפיסת קלפים תקנית. נגדיר את המאורע A– נשלף קלף מסוג "לב" ו-B – נשלף קלף "מלך". האם מאורעות אלו תלויים?

הוכחה. בחפיסת קלפים 52 קלפים, מתוכם 4 קלפי מלך ו-13 קלפי לב. לכן הסיכוי לבחור קלף לב הינו $Pr\left(A\right)=\frac{1}{4}$. בהינתן שנבחר קלף "מלך", כל אחד מארבעת המלכים יכל היה להיבחר בהסתברות שווה והסיכוי שזה מלך מסוג לב הינו $Pr\left(A|B\right)=\frac{1}{4}$. לאור זאת ולאור תרגיל 2.17, המאורעות בלתי תלויים. □

בהרבה מקרים המאורעות יהיו בלתי תלויים מאופן עריכת הניסוי. למשל, הטלות של קוביה הוגנת הן בלתי תלויות שכן הסיכוי לכל פאה קבוע ולא תלוי בתוצאה של ההטלה הקודמת או כמה פעמים הקוביה הוטלה כבר בעבר.

תרגיל 2.17. הוכיחו שהתנאים בהגדרה 2.2 שקולים: $Pr\left(A|B\right)=Pr\left(A\right)$ אם ורק אם $Pr\left(B|A\right)=Pr\left(B\right)$.

תרגיל 2.18. נניח שהמאורעות A ו-B בלתי תלויים. האם המאורעות $\overline{A}$ ו-$\overline{B}$ בלתי תלויים גם כן? הוכיחו!

בתרגילים 2.17 ו-2.18 הוכחנו שתי תוצאות בסיסיות ואינטואיטיביות של תלות בין מאורעות. ראשית, יחס התלות הוא יחס סימטרי, כלומר אם A תלוי ב-B אזי בוודאי גם B תלוי ב-A. אם התרחשותו של מאורע A משנה את הסיכוי של מאורע B, אזי מהתרחשותו של מאורע B גם תשתנה ההסתברות של המאורע A. בנוסף, אם הסיכוי של B משתנה כאשר אנחנו יודעים ש-A התרחש, אז הסיכוי צריך להשתנות גם כאשר אנחנו יודעים שהמאורע A לא התרחש ולכן גם B ו-$\overline{A}$ תלויים. יתרה מזאת, כחלק מההוכחה של תרגיל 2.18 הראנו שאם A ו-B בלתי תלויים אזי גם כל הזוגות $(\overline{A},B)$, $(\overline{B},A)$ ו-$(\overline{A},\overline{B})$ בלתי תלויים.

ניתן לפשט את נוסחת הכפל (משפט 2.6) עבור מאורעות בלתי תלויים ולקבל שהסתברות החיתוך היא מכפלת ההסתברויות של המאורעות. מסתבר שזהו גם תנאי מספיק לכך שהמאורעות יהיו בלתי תלויים - אם ההסתברות של חיתוך מאורעות שווה למכפלת ההסתברויות שלהם, אזי המאורעות בלתי תלויים. חישוב זה יכול לשמש במקרים רבים לבדיקה האם מאורעות תלויים או בלתי תלויים.

משפט 2.8. *המאורעות A ו-B הם בלתי תלויים אם ורק אם מתקיים*

$$Pr\left(A\cap B\right)=Pr(A)\cdot Pr(B) \tag{2.19}$$

הוכחה. המאורעות A ו-B בלתי תלויים אם ורק אם $Pr\left(A|B\right)=Pr(A)$ ועל ידי הצבה בנוסחא 2.12 נקבל $Pr\left(A\cap B\right)=Pr(A)\cdot Pr(B)$, כנדרש. □

ייתכן ונתקלתם בנוסחא זו בקורסים קודמים בהסתברות (למשל, כחלק מהחומר לבגרות ולפסיכומטרי) שם מלמדים שההסתברות של "A וגם B" היא מכפלת ההסתברויות של שני המאורעות, מבלי להדגיש מספיק את העובדה שהמאורעות צריכים להיות בלתי תלויים ולכן נדמה שהנוסחא נכונה תמיד. כאמור, הנוסחא איננה נכונה תמיד! נוסחא זו מתבססת על כך שהמאורעות בלתי תלויים ולפני שמבצעים את המעבר מהסתברות החיתוך למכפלת ההסתברויות חשוב לוודא ולנמק היטב שאכן מדובר על מאורעות בלתי תלויים והשימוש בנוסחא מוצדק.

תרגיל 2.19. יהי $A\subseteq\Omega$ מאורע הבלתי תלוי בכל מאורע אחר. מהם הערכים האפשריים עבור $Pr(A)$? הסבירו את המשמעות של התוצאה שקיבלתם.

תרגיל 2.20. יהיו מאורעות A,B זרים עם הסתברות חיובית. האם מאורעות אלו תלויים?

תרגיל 2.21. בספארי הפילים ישנים 25% מהזמן, ההיפופוטמים 20% מהזמן והאריות 80% מהזמן (כל החיות בלתי תלויות זו בזו). מה הסיכוי שבביקור בספארי תראו לפחות חיה אחת ערה?

1. חשבו באמצעות נוסחת ההכלה וההפרדה.
2. חשבו באמצעות כללי דה מורגן (נוסחא 1.11).

נחזור לדוגמת הכדורסלן (בעיה 2.3) ונניח שלאחר שהטיל קוביה ובחר מאין לזרוק, הוא זורק מהמיקום הנבחר פעמיים לסל. נסמן ב-F את המאורע שהוא קולע לסל בזריקה הראשונה וב-S את המאורע שהוא קולע לסל בזריקה השנייה. האם מאורעות אלו תלויים?

תרגיל. לפני שתמשיכו לקרוא הלאה, ענו אינטואיטיבית - האם המאורעות תלויים או בלתי תלויים?

תשובה אינטואיטיבית אפשרית היא שהמאורעות בלתי תלויים שהרי, לפי הנתונים, ההסתברות לקלוע מכל טווח היא קבועה. למשל, הסיכוי לקלוע זריקת עונשין הוא 0.9 וזה לא משנה אם זו הזריקה הראשונה שהכדורסלן זורק או החמישית. לכן, תוצאת הזריקה הראשונה לא צריכה להשפיע על הזריקה השניה והסיכוי לסל בה זהה בין אם ידועה תוצאת הזריקה ובין אם לאו.

נבחן אינטואיציה זו ונענה מפורשות על השאלה על ידי בדיקה האם $Pr\left(F\cap S\right)=Pr\left(F\right)\cdot Pr(S)$. את ההסתברות לסל בזריקה הראשונה חישבנו כבר וקיבלנו כי $Pr(F)=\frac{49}{60}$. באופן דומה, על ידי נוסחת ההסתברות השלמה והתניה על המאורע שהקוביה נפלה על 3 (המאורע T) נקבל כי הסיכוי לסל בזריקה השניה הינו $Pr(S)=\frac{49}{60}$ גם כן.

נותר לחשב את ההסתברות של המאורע שהכדורסלן קלע סל בשתי הזריקות, $F\cap S$. לו היינו יודעים מהי תוצאת הקוביה, היינו יודעים מאין נזרקו הזריקות וההסתברות לסל הייתה נתונה. למשל, אם הקוביה נפלה על 3 הזריקות בוצעו מטווח השלשה בה הסיכוי לסל הוא 0.4 ולכן הסיכוי לשני סלים הינו 0.4^{2}. היות ואנחנו לא יודעים את תוצאת הקוביה, נתנה על האפשרויות השונות באמצעות נוסחת ההסתברות השלמה באופן דומה לזה בו חושבה ההסתברות של מאורע F:

$$\begin{aligned}Pr\left(F\cap S\right) & = Pr\left(F\cap S|T\right)Pr(T)+Pr\left(F\cap S|\overline{T}\right)Pr\left(\overline{T}\right)\\ & = 0.4^{2}\cdot\frac{1}{6}+0.9^{2}\cdot\frac{5}{6}=\frac{421}{600}\end{aligned} \tag{2.20}$$

כלומר, בניגוד לאינטואיציה הראשונית קיבלנו $Pr\left(F\cap S\right)\approx0.7\neq0.66=Pr(F)\cdot Pr(S)$ והמאורעות תלויים! כיצד זה ייתכן, שהרי סל בזריקה הראשונה לא גורם לשחקן הכדורסל להיות יותר טוב ולא אמור להגדיל או להקטין את ההסתברות של סל בזריקה השניה?

הסיבה לכך שהמאורעות תלויים היא שהמיקום ממנו השחקן זורק לסל אינו ידוע, אך כל תוצאת זריקה לסל מלמדת אותנו משהו על המקום ממנו הוא זורק ומכאן משתנה ההסתברות לסל בזריקה הבאה. בדוגמא זו, השחקן קולע לסל בסיכוי גבוה מקו העונשין ובסיכוי נמוך מהשלוש, לכן אם הוא קלע סל בזריקה הראשונה, טבעי להניח שהזריקה נעשתה מקו העונשין ולכן הסיכוי שהוא יקלע את הזריקה השנייה גדול יותר מאשר $\frac{49}{60}$. לעומת זאת, אם הוא החטיא את הזריקה הראשונה, ככל הנראה הוא זורק מקשת השלוש ולכן גם את הזריקה השניה הוא צפוי להחטיא.

נחזק אינטואיציה זו על ידי כך שנלך לשני תרחישי קיצון. בתרחיש הראשון, נניח שהכדורסלן תמיד מצליח לקלוע לסל מהעונשין ותמיד מחטיא משלוש. אם הוא קלע את הראשונה, הוא בוודאות זורק מהעונשין ולכן בוודאות

גם יקלע את השנייה. אם הוא החטיא את הראשונה, הוא בוודאות זורק מהשלוש ובוודאות יחטיא את השניה - באופן ברור שתי הזריקות תלויות. בתרחיש השני, נניח שהכדורסלן קולע לסל בהסתברויות הנתונות בשאלה אך במקום לזרוק פעמיים הוא זורק אלף ואחת פעמים לסל. הזריקה ה-1,001 תלויה בכל אלף הזריקות הקודמות. אם אנחנו יודעים שרוב הזריקות הקודמות נכנסו לסל, סביר מאוד שהוא זורק מהעונשין ולכן בסיכוי גדול יותר גם הזריקה הבאה תיכנס לסל. לעומת זאת, אם אנחנו יודעים שרוב הזריקות הקודמות לא נכנסו לסל סביר להניח שהוא זורק משלוש ולכן כנראה שיחטיא גם את הזריקה הבאה.

לאור הדוגמא הקודמת חשוב להבדיל בין תלות הסתברותית לסיבתיות. המאורעות בשאלה תלויים אך אין בהם קשר סיבתי – סל בזריקה הראשונה לא "גורם"[5] לכדור להיכנס לסל בסיכוי גבוה יותר בפעם הבאה. הקשר בין המאורעות הוא קשר עקיף, הנובע מכך שהטלת הקוביה משפיעה על ההסתברות של שתי הזריקות. אם הזריקה הראשונה נכנסה לסל, זה מלמד משהו על תוצאת הטלת הקוביה ומכאן גם על הסיכוי שייקלע סל בזריקה השניה (ראו תרגיל 2.24). באופן דומה, הציונים של סטודנט מסוים בקורסים "אלגברה לינארית" ו"מבוא להסתברות" תלויים, למרות שמדובר על קורסים שונים, חומר שונה, מרצים שונים ומבחנים שונים: ציון גבוה באחד הקורסים מצביע על כך שמדובר על סטודנט טוב וסביר להניח שיקבל ציון גבוה גם בקורס השני. לעומת זאת, כאשר יש תלות סיבתית בין המאורעות בוודאי גם יש תלות הסתברותית ביניהם. אם המאורע A גורם למאורע B לקרות, אזי $Pr\left(B|A\right)=1$. כאשר המאורע B לא קורה תמיד (כלומר $Pr\left(B\right)\neq 1$) נקבל שהמאורעות אכן תלויים.

תרגיל 2.22. יהיו A,B מאורעות בעלי הסתברות חיובית הקטנה מ-1 המקיימים $A\subseteq B$. האם מאורעות אלו תלויים?

תרגיל 2.23. יהיו A,B,C שלושה מאורעות כך שכל זוג מאורעות הוא בלתי תלוי. האם גם השלישייה כולה בלתי תלויה, כלומר האם מתקיים $Pr\left(A\cap B\cap C\right)=Pr\left(A\right)\cdot Pr\left(B\right)\cdot Pr\left(C\right)$?

תרגיל 2.23 מלמד אותנו שעל מנת להכליל את מושג חוסר התלות לקבוצות גדולות של מאורעות, לא מספיק לדרוש שהמאורעות יהיו בלתי תלויים בזוגות (כפי שנעשה כאשר הגדרנו מאורעות זרים בזוגות) אלא יש לדרוש שכל מספר של מאורעות יהיו זרים.

כדאי לשים לב שתלות וחוסר תלות בין מאורעות יכולה להשתנות בהתאם לתוצאות הניסוי או כתוצאה מהתניה על מאורעות אחרים. למשל, כאשר מטילים שתי קוביות המאורעות A – "יצא 5 בקוביה הראשונה" ו-B– "יצא 5 בקוביה השניה" הם בלתי תלויים. יחד עם זאת, אם ידוע שסכום הקוביות הוא 10, מאורעות אלו תלויים, שהרי בהינתן שהסכום הוא 10 ובהינתן שמאורע A התרחש, ההסתברות של מאורע B היא כעת 1. זוהי טעות נפוצה להתייחס למאורעות שנתון שהם בלתי תלויים כמאורעות בלתי תלויים תמיד, מבלי לשים לב שנתונים נוספים או התניות נוספות יכולות להפוך את המאורעות לתלויים.

2.2.3 נוסחת בייס

לפי חוק הכפל (נוסחא (2.12)) את הסתברות החיתוך של שני מאורעות ניתן לחשב על ידי התניה על כל אחד מהם. אם נחלק את הנוסחא ב-$Pr(B)$ (למשל) נקבל את המשפט השימושי הבא:

משפט 2.9 (נוסחת בייס). *עבור מאורעות A,B בעלי הסתברות חיובית מתקיים*

$$Pr\left(A|B\right)=\frac{Pr\left(B|A\right)\cdot Pr(A)}{Pr(B)} \qquad (2.21)$$

[5] במציאות תופעה זו יכולה להתרחש. למשל, שחקן יכול לאבד ביטחון אחרי ההחטאה הראשונה ולכן להחטיא בסיכוי גדול יותר בזריקה השניה ולהפך (ראו תרגיל 2.13)

נוסחת בייס מאפשרת את החלפת כיוון ההתניה מ- $A|B$ ל-$B|A$. ניתן להשתמש בנוסחא זו כאשר אנחנו מעוניינים לחשב את אחת ההתניות, אבל דווקא השניה ניתנת לחישוב פשוט יותר כי היא נתונה או כי היא יותר טבעית. למשל, כאשר ניסוי מתקיים בשני שלבים, נוח מאוד להתנות את תוצאת השלב השני על תוצאת השלב הראשון (ככל הנראה יש ביניהם קשר ברור) אך מאוד מוזר ולא טבעי לחשוב על תוצאת השלב הראשון בהינתן השלב השני. במקרה כזה נרצה להחליף את כיוון ההתניה כך שנחשב את ההסתברות המותנית הטבעית יותר.

בעיה 2.5. מתוך ערימה של מטבעות שחציים הוגנים וחציים מוטים (מראים עץ בהסתברות $\frac{3}{4}$) בוחרים מטבע באקראי ומטילים אותו פעם אחת. מה ההסתברות שנבחר מטבע הוגן (המאורע F) אם המטבע נפל על עץ (המאורע H)?

הוכחה. נתחיל מדיון אינטואיטיבי בנוגע לערך הצפוי. לפני שמטילים את המטבע, הסיכוי לבחור מטבע הוגן ולא הוגן שווה, כלומר $Pr(F) = 0.5$. הסתברות זו מכונה הפריור (Prior) ומתארת את ההסתברות שאנחנו מייחסים לתוצאה כלשהי בטרם ביצוע הניסוי או קבלת מידע נוסף. הטלת המטבע מהווה ניסוי המאפשר להסיק משהו על זהות המטבעות, שכן המטבעות נופלים בהסתברויות שונות על עץ ופלי. אם המטבע נפל על עץ, סביר יותר להעריך שמדובר על המטבע הלא הוגן הנופל על עץ לעיתים קרובות יותר ביחס למטבע ההוגן, ולכן אנחנו מצפים שיתקיים $Pr\left(F|H\right) < 0.5$. לעומת זאת, המטבע הלא הוגן נופל על פלי לעיתים נדירות ולכן אם יצא פלי נעריך שזהו המטבע הוגן, כלומר $Pr\left(F\middle|\overline{H}\right) > 0.5$. ההסתברות המותנית המעודכנת הלוקחת בחשבון את המידע הנוסף שברשותנו (במקרה זה – תוצאת הטלת המטבע) מכונה הפוסטיריור (Posterior) וזה מה שעלינו לחשב בדוגמא זו.

נפנה עתה לחישוב כמותי של ההסתברות $Pr\left(F|H\right)$. מאורע זה מתאר התניה לא הגיונית ולא טבעית לעבודה – התניה של סוג המטבע בהינתן תוצאת הטלתו. לעומת זאת, ההתנייה ההפוכה, של תוצאת ההטלה בהינתן סוג המטבע היא נוחה וטריוויאלית: אם המטבע הוגן ההסתברות לעץ היא 0.5 ואם המטבע אינו הוגן הסיכוי לעץ הינו 0.75. נשתמש אם כך בנוסחת בייס בכדי להחליף את כיוון ההתניה ונקבל:

$$Pr\left(F|H\right) = \frac{Pr\left(H|F\right)Pr(F)}{Pr(H)} = \frac{0.5 \cdot 0.5}{Pr(H)}$$

נותר לחשב את ההסתברות לקבל עץ בניסוי זה. בהמשך לדיון הקודם, לו היינו יודעים איזה מטבע נבחר חישוב ההסתברות לעץ היה קל, אך אין בידינו מידע זה ואנחנו צריכים לחשב את ההסתברות הכוללת לעץ בניסוי. נשתמש אם כך בנוסחת ההסתברות השלמה ונתנה על סוג המטבע שנבחר:

$$Pr(H) = Pr\left(H|F\right)Pr(F) + Pr\left(H\middle|\overline{F}\right)Pr\left(\overline{F}\right) = 0.5 \cdot 0.5 + 0.75 \cdot 0.5 = 0.625$$

לפיכך $Pr\left(F|H\right) = \frac{0.25}{0.625} = 0.4$. כצפוי, הסתברות זו נמוכה מההסתברות הלא מותנית שהמטבע הוגן. $\square$

מבחינה טכנית, ישנה מסקנה חשובה שאפשר ללמוד מהדוגמא האחרונה לגבי אופן חישוב ההסתברויות באגף ימין של נוסחא **2.21**. המוטיבציה שלנו להשתמש בנוסחת בייס היא שההתפלגות המותנית $Pr\left(B|A\right)$ פשוטה יותר לחישוב מאחר $Pr\left(A|B\right)$ ולרוב גם $Pr(A)$ נתון או קל לחשב. לכן הבעיה העיקרית בשימוש בנוסחת בייס היא חישוב המכנה, $Pr(B)$. כפי שנעשה בדוגמא האחרונה, ניתן לנסות לחשב הסתברות זו באמצעות נוסחת ההסתברות השלמה כאשר מתנים על המאורע A (ועל תוצאות נוספות שיכלו להתרחש בשלב הראשון של הניסוי אם ישנן כאלו). בדוגמא האחרונה, התנו על סוג המטבע שנבחר. לו היו שלושה סוגי מטבעות, אופן חישוב המונה כלל לא היה משתנה (רק המספרים היו משתנים) ובמכנה היינו עדיין מחשבים את ההסתברות לקבל עץ באמצעות נוסחת ההסתברות השלמה, כאשר בכל פעם היינו מתנים על סוג אחר של מטבע.

תרגיל 2.24. בהמשך לבעיית הכדורסלן (בעיה **2.3** בעמוד **44**) חשבו את ההסתברות שהכדורסלן זורק מקו העונשין במקרים הבאים:

1. הוא זרק זריקה אחת וקלע אותה.
2. הוא זרק שתי זריקות וקלע את שתיהן.
3. הוא זרק שתי זריקות, קלע את הראשונה והחטיא את השנייה.

כאמור, אחד השימושים לנוסחת בייס הוא בחישוב ההסתברות המותנית של שלב ביניים בניסוי (למשל, בחירת המטבע) בהינתן התוצאה הסופית של הניסוי (יצא עץ). תיארנו התנייה זו כהתנייה "לא טבעית" שכן במסגרתה אנחנו מתנים בניגוד לכיוון ההגיוני והקל לחישוב של ההתנייה, אך במציאות הסקה זו נפוצה מאוד הן בחיים והן בניסוי מעבדה בכל תחומי המדעים. למשל, אנחנו נכנסים למעבדה בכדי לאושש תאוריה פיזיקלית מסויימת. אם התאוריה נכונה, אנחנו מצפים לתוצאות מסוימות ואם התאוריה שגויה אנחנו מצפים לתוצאות אחרות, אבל עקב מגבלות של דיוק, רעש וכד' ייתכן ונקבל את התוצאות הלא נכונות. לאחר ביצוע הניסוי כל שאנחנו יודעים הוא התוצאה הסופית ועל סמך תוצאה זו מנסים להסיק מסקנות לגבי האם התאוריה נכונה או לא. ניתן לתאר ניסוי זה כניסוי דו שלבי. בשלב הראשון הטבע מגריל מהי התאוריה הנכונה ובהתאם לכך מוגרלת תוצאה לניסוי. אנחנו רואים רק את השלב השני, התוצאה של הניסוי, ולפיה מנסים להבין את התוצאה הראשונה. זהו למעשה התהליך אשר עושים פעמים רבות בתחום הסטטיסטיקה – חישוב ההסתברויות שתאוריות מתחרות נכונות בהינתן התוצאות שנערכו. דוגמא נוספת מגיעה מתחום התקשורת. הצד הקולט יודע רק את התוצאה הסופית הכוללת גם את המסר שנשלח וגם רעש שנוסף במהלך השידור, ועל סמך מידע זה צריך לשערך את המסר.

בעיה 2.6. למחלה מסויימת שכיחות של $\frac{1}{10,000}$, כלומר זהו הסיכוי של אדם מקרי לחלות במחלה. בכדי לבדוק האם אדם חולה הוא נשלח לבדיקת דם הנותנת את התשובה הנכונה בהסתברות 0.95. האדם מקבל תשובה חיובית, קרי לפי בדיקת הדם הוא חולה. מה ההסתברות שהוא אכן חולה?

תרגיל. לפני שתמשיכו לקרוא הלאה, ענו אינטואיטיבית - מהו תחום הערכים שאתם מצפים לקבל עבור הסתברות זו (קרוב ל-0.95? גדול יותר? קטן יותר? בכמה קטן יותר?). רמז: קראו את התשובה של תרגיל **2.24**.

הוכחה. תשובה אינטואיטיבית היא שההסתברות צריכה להיות באזור ה-0.95, שכן זהו הדיוק של בדיקת הדם. תשובה זו לא לוקחת בחשבון את ההשפעה של הפריור על הפוסטיריור, וכפי שראינו בנוסחת בייס ובפתרונו של תרגיל **2.24**, יש לפריור השפעה. אם מראש ההסתברות הפריורית למאורע מסויים היא נמוכה, הרי שגם אם תוצאת הניסוי מתאימה למאורע זה, הפריור הנמוך יגרום לכך שהפוסטיריור לא יעלה משמעותית שכן מראש מדובר על מאורע נדיר. לשם המחשה, נניח ש-$10,000$ אנשים ניגשים לבדיקה. שכיחות המחלה היא $\frac{1}{10,000}$ לכן בקירוב רק אדם אחד חולה והיתר בריאים. הבדיקה נותנת את התשובה הנכונה בהסתברות 0.95 לכן בסיכוי גבוה אותו אדם חולה אכן יקבל תשובה חיובית. אבל, גם $9,999$ הבריאים ניגשים לבדיקה ובהסתברות 0.05 יקבלו את התשובה הלא נכונה מבחינתם, התשובה החיובית. לכן אנחנו נצפה לקבל $500 \simeq 9,999 \times 0.05$ תשובות חיובית (ושגויות) לאנשים הבריאים. בסך הכל מתוך $10,000$ האנשים שנבדקו 501 קיבלו תשובה חיובית. לכל אדם שקיבל תשובה חיובית יש סיכוי של $\frac{1}{501} \simeq 0.001996$ להיות האדם החולה באמת, הרבה פחות מ-0.95. □

תרגיל 2.25. חשבו בצורה מדוייקת את ההסתברות שאדם המקבל תשובה חיובית אכן חולה.

אמנם תוצאה זו נראית מוזרה ולא אינטואיטיבית במבט ראשון, אבל אם תשימו לב תראו שאנשים מסביבכם ואף אתם פועלים לפיה באופן לא מודע. למשל, כאשר אדם עובר בגלאי המתכות בכניסה לרכבת וגלאי המתכות מצפצף, אף אחד מהמאבטחים לא מתרגש יתר על המידה. אמנם גלאי המתכות מזהה מטען חבלה בהסתברות מאוד גבוהה, אבל גלאי המתכות מזהה דברים נוספים דוגמת טלפונים בכיסים. זאת ועוד, ההסתברות הפריורית שמחבל עם מטען יגיע לתחנה הוא מאוד נמוך ולכן גם כאשר גלאי המתכות מצפצף, ההסתברות שהצפצוף מתריע על מטען הוא נמוך מאוד וסביר להניח שמדובר על התראת שווא נוספת.

דוגמא נוספת לכך שהסתברות מותנית יכולה להיות מאוד לא אינטואיטיבית (וכנראה הדוגמא המפורסמת ביותר) נקראת פרדוקס מונטי הול, על שמו של מנחה הטלוויזיה מונטי הול (במקור, הלפרין). מונטי הול הגיש תוכנית טלוויזיה בשם "Let's make a deal" (בגירסתה הישראלית נקראה "עשינו עסק" ושודרה בהנחיית צביקה הדר ואברי גלעד). באחד משלבי התוכנית היו מציגים למתמודד שלוש דלתות סגורות כאשר מאחורי אחת מהן נמצא פרס גדול ויקר (נניח, מכונית) ומאחורי שתי הדלתות האחרות היו פרסים פעוטי ערך או כלום (נניח, כפי שמקובל לספר את הסיפור, שיש עז מאחורי כל אחת משתי הדלתות האחרות). המתמודד התבקש לבחור באחת הדלתות ומונטי הול, שידע היכן מצויה המכונית, היה פותח דלת אחרת מזו שהמתמודד בחר ומציג את העז העומדת מאחוריה. לאחר מכן, היה מציע מונטי הול למתמודד להחליף את הבחירה שלו ולבחור את הדלת הסגורה השניה.

תרגיל. ענו אינטואיטיבית – מה אתם הייתם עושים, מחליפים או נשארים בבחירה המקורית? כיצד המידע הנוסף שחשף בפניכם המנחה משפיע על החלטתכם?

בעיה זו הפכה למפורסמת ומעניינת במיוחד בשנת 1990 בזכות מרלין ווס-סאוואנט (Marilyn Vos Savant). ווס-סאוואנט החזיקה בשיא גינס למנת המשכל הגבוהה בהיסטוריה (IQ מעל 220!) והייתה כותבת טור קבוע במגזין Parade בו הייתה עונה על שאלות מהקוראים. במסגרת הטור היא נשאלה לגבי המשחק ומה כדאי למתמודד לעשות ותשובתה יצרה מהומה רבתי. מעל לאלף מכתבים נשלחו למערכת ממתמטיקאים רבים המסבירים לה שהיא טעתה. היא כמובן צדקה ורבים מהמתמטיקאים התקשו להשתכנע בכך גם אחרי הוכחות רבות וסימולציות מחשב. בתרגיל 2.26 תחשבו את שתי האפשרויות ותקבעו – מה על המתמודד לעשות.

תרגיל 2.26. מה כדאי למתמודד לעשות בבעיית מונטי הול? מה הסיכוי לזכיה אם יחליף דלת ומה הסיכוי לזכיה אם לא יחליף דלת?

תשובה אפשרית (ושגויה) היא ששתי הדלתות שהמנחה לא פתח הן סימטריות ולכן הסיכוי לזכות בכל אחת מהן הוא $\frac{1}{2}$. תשובה זו לא נכונה שכן הדלתות לא סימטריות – הדלת שהמתמודד בחר שונה מהדלת שאינה בחר שכן ייתכן מאוד שדלת זו היא הדלת שהמנחה היה חייב לא לפתוח כדי לא לחשוף את המכונית. למשל, אם המתמודד בחר בדלת מספר 1 והמכונית נמצאת מאחורי דלת מספר 2, המנחה חייב לפתוח את דלת מספר 3 כדי לא לחשוף את המכונית והעובדה שהדלת נותרה נעולה מגדיל את הסיכוי שמאחוריה יש מכונית.

ניתן לחשוב על תוצאה זו גם בדרך אחרת – המתמודד יכול לבחור או בדלת אחת או בשתי הדלתות האחרות. הסיכוי לזכות כאשר הוא בוחר בשתי דלתות הוא כמובן $\frac{2}{3}$. העובדה שהמנחה פותח דלת ומראה שמאחורי אחת משתי הדלתות יש עז לא משנה דבר, שכן ממילא ידוע שלפחות מאחורי אחת משתי הדלתות יש עז. אין בפתיחת הדלת מידע חדש ובהינתן הדלת שנפתחה, הסיכוי לזכות בבחירה של שתי הדלתות כפול מהסיכוי לזכות בבחירה של דלת אחת. כדי להקצין אינטואיציה זו, נניח שהיו 1000 דלתות, המתמודד בחר אחת מהן ולאחר מכן נפתחו 998 דלתות ונחשפו 998 עזים. נותרו שתי דלתות, אבל הדלת הבודדת שלא נפתחה על ידי המנחה כנראה לא נפתחה מסיבה טובה ובבירור הדלתות לא סימטריות.

2.3 שאלות מסכמות

תרגיל 2.27. נתונים שלושה מאורעות המקיימים $Pr(A) = 0.5, Pr(B) = 0.3, Pr(C) = 0.2$. מה ההסתברות שלפחות אחד מהם יקרה כאשר

1. המאורעות זרים בזוגות.
2. המאורעות בלתי תלויים.
3. מתקיים $C \subseteq B \subseteq A$.

תרגיל 2.28. כדורסלן זורק לסל פעמיים, פעם אחת מקו העונשין ופעם אחת מקשת השלוש. הסיכוי שיקלע מקו העונשין הינו 0.8 והסיכוי שיקלע מקשת השלוש הינו 0.7. ההסתברות שיקלע לפחות סל אחד הינו 0.95.

1. מה ההסתברות שיצליח לקלוע את שני הסלים?
2. מה ההסתברות שיקלע בדיוק סל אחד?
3. מה הסיכוי שיקלע בדיוק סל אחד בהינתן שהוא קלע מהעונשין?

תרגיל 2.29. כדורסלן זורק לסל משתי עמדות אפשריות – קו העונשין וקשת השלוש. הסיכוי שיקלע מקו העונשין הינו 0.8 והסיכוי שיקלע מקשת השלוש הינו 0.7. הכדורסלן בוחר באקראי את המיקום ממנו יזרוק את הזריקה הראשונה וזורק לסל. אם קלע – הוא זורק ממקום זה פעם נוספת ואם החטיא – זורק מהעמדה השניה פעם אחת. מה ההסתברות שהכדורסלן

1. יקלע פעמיים?
2. יחטיא פעמיים?
3. יקלע בדיוק פעם אחת?

תרגיל 2.30. בקרקס יש שלושה לוליינים המבצעים תרגילים אקרובטיים על טרפז: אביטל, דניאל ואלירן. אביטל נעדרת מ-40% מהופעות, דניאל מ-50% מההופעות ואלירן מ-40% מההופעות. כל זוג לוליינים נעדר יחד מ-20% מההופעות וכולם ביחד נעדרים מ-5% מההופעות. לכל אחד מהלוליינים יש מחליף אשר יודע את התפקיד שלו ויכול להחליף אותו כאשר הוא נעדר.

1. מה ההסתברות שבמופע מסויים לא יהיו מחליפים כלל?
2. חשבו את ההסתברות שיש בדיוק $k \in \{0, 1, 2, 3\}$ מחליפים.

תרגיל 2.31. במשפחת כהן שני ילדים אשר כל אחד מהם יכול להיות בן או בת בהסתברות שווה ללא תלות בילד השני. אתה דופק בדלת ובן פותח את הדלת. מה ההסתברות שבמשפחה שני בנים כאשר:

1. במשפחת כהן נהוג שהבכור פותח את הדלת תמיד.
2. במשפחת כהן נהוג שהילד הפותח את הדלת נבחר באקראי בכל פעם.

תרגיל 2.32. לכל אדם שני גנים האחראים על צבע השיער ונניח לשם הפשטות שהצבעים האפשריים הם שחור וג'ינג'י. הגן הג'ינג'י הוא גן רציסיבי, לכן רק אם לאדם שני גנים ג'ינג'יים הוא יהיה ג'ינג'י. אחרת, צבע השיער ייקבע לפי הגן הדומיננטי ויהיה שחור. כל ילד יורש באקראי גן אחד מכל הורה ולפי שני הגנים שירש – נקבע צבע שערו. לאם ג'ינג'ית ולאב שחור שיער שני ילדים.

1. מהן האפשרויות השונות עבור הגנים של האב?
2. נניח שכל אחת מהאפשרויות עבור הגנים של האב מתקבלת בהסתברות שווה. מה הסיכוי שהבכור יהיה ג'ינג'י? מה הסיכוי שהצעיר יהיה ג'ינג'י?
3. ענו מבלי לחשב האם צבע השיער של הבכור וצבע השיער של הצעיר תלויים או בלתי תלויים.

4. בהינתן שהבכור שחור שיער, מה הסיכוי שהצעיר יהיה ג'ינג'י?

תרגיל 2.33. מערכת מורכבת מ-n רכיבים בלתי תלויים המחוברים במקביל (די בכך שאחד יפעל על מנת שהמערכת כולה תפעל). כל רכיב פועל בהסתברות 0.5.

1. מה ההסתברות שהמערכת תפעל? מה קורה כשמספר הרכיבים שואף לאינסוף?
2. מה ההסתברות שהמערכת תפעל בהינתן שרכיב מספר 1 פועל?
3. מה ההסתברות שרכיב מספר 1 פועל בהינתן שהמערכת פועלת? מה קורה כאשר מספר הרכיבים שואף לאינסוף?
4. האם המאורע "המערכת פועלת" והמאורע "רכיב 1 פועל" תלויים או בלתי תלויים? האם המאורעות תלויים או בלתי תלויות אסימפטוטית (בגבול $n \to \infty$)? הסבירו את התוצאה אינטואיטיבית.

תרגיל 2.34. מזג האוויר בעיר מסויימת יכול להיות שמשי או גשום בהסתברויות שוות. בבוקר מסויים, לפני נסיעתכם לעיר, אתם מתקשרים ל-3 מחבריכם הגרים בעיר ומבררים את מזג האוויר. כל אחד מחבריכם דובר אמת בהסתברות $\frac{2}{3}$ (ומשקר בהסתברות המשלימה) ללא תלות בחברים האחרים.

1. שלושת החברים אמרו שיורד גשם. מה ההסתברות שאכן יורד גשם?
2. ענו ללא חישוב נוסף - כיצד תשתנה תשובתכם אם הסיכוי לגשם בעיר זו הוא 0.1 במקום 0.5.
3. האם המאורעות A - "חבר א' אמר שיורד גשם" ו-B - "חבר ב' אמר שיורד גשם" תלויים או בלתי תלויים? הוכיחו!

תרגיל 2.35. נתונים שלושה מעגלים חשמליים, אשר כל אחד כולל נורה. הסיכוי שהנורה שבמעגל הראשון תדלוק הינו 0.3, הסיכוי שהנורה שבמעגל השני תדלוק הינו 0.5 ואילו הנורה שבמעגל השלישי דולקת תמיד.

1. בוחרים באקראי מעגל חשמלי ובודקים אם הנורה דולקת. לאחר זמן מה, בודקים שנית האם אותה הנורה דולקת. מה הסיכוי שבפעם השניה הנורה דולקת אם היא דלקה בפעם הראשונה?
2. בוחרים באקראי מעגל חשמלי ובודקים האם הנורה דולקת. לאחר מכן, בוחרים באקראי שוב מעגל חשמלי ובודקים האם הנורה דולקת. מה הסיכוי שבפעם הראשונה הנורה דלקה אם ידוע שבפעם השניה היא דלקה?

2.4 פתרונות לשאלות הפרק

תרגיל 2.1. נגדיר את המאורעות $A_i = \{i\}$ לכל $i \in \Omega$.

1. המאורעות A_i זרים בזוגות ומקיימים $Pr(A_i) = c \cdot i^2$ ולכן

$$Pr\left(\bigcup_{i=1}^{10} A_i\right) = \sum_{i=1}^{10} Pr(A_i) = \sum_{i=1}^{10} c \cdot i^2 = 385c$$

כאשר במעבר האחרון השתמשנו בנוסחא הידועה של סכום ריבועים $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.
יחד עם זאת, איחוד כל המאורעות הללו הוא בדיוק מרחב המדגם כך שהסתברות האיחוד צריכה להיות 1 ולכן $c = \frac{1}{385}$.

2. המאורע שתוצאת הניסוי תתחלק ב-3 הוא איחוד של שלושת המאורעות A_3, A_6, A_9 ולכן ההסתברות המבוקשת הינה

$$Pr(A_3) + Pr(A_6) + Pr(A_9) = \frac{3^2}{385} + \frac{6^2}{385} + \frac{9^2}{385} = \frac{126}{385}$$

3. נסמן ב-B את המאורע שתוצאת הניסוי גדולה מ-2, כלומר, $B = A_3 \cup A_4 \cup \ldots \cup A_{10}$. לצורך חישוב ההסתברות של B יש לסכום את 8 ההסתברויות המתאימות למאורעות הללו הזרות בזוגות. אמנם חישוב זה אפשרי, אך ניתן לפשטו בצורה ניכרת על ידי התבוננות במאורע המשלים, שתוצאת הניסוי קטנה או שווה ל-2:

$$Pr\left(\overline{B}\right) = Pr(A_1 \cup A_2) = Pr(A_1) + Pr(A_2) = \frac{5}{385}$$

ולכן $Pr(B) = 1 - Pr\left(\overline{B}\right) = \frac{380}{385} = \frac{76}{77}$.

□

תרגיל 2.2. נגדיר את המאורעות $A_i = \{i\}$ לכל $i \in \Omega$.

1. המאורעות A_i זרים בזוגות ומקיימים $Pr(A_i) = \frac{c}{i^2}$ ולכן

$$Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} Pr(A_i) = \sum_{i=1}^{\infty} \frac{c}{i^2} = \frac{c\pi^2}{6}$$

כאשר במעבר האחרון השתמשנו בסכום הידוע מנוסחא 0.23.
יחד עם זאת, איחוד כל המאורעות הללו הוא בדיוק מרחב המדגם כך שהסתברות האיחוד צריכה להיות 1 ולכן $c = \frac{6}{\pi^2}$.

2. נסמן ב-E את המאורע שתוצאת הניסוי זוגית. מתקיים $E = A_2 \cup A_4 \cup \ldots$ ולכן

$$Pr(E) = Pr(A_2 \cup A_4 \cup \ldots) = \sum_{i=1}^{\infty} Pr(A_{2i}) = \sum_{i=1}^{\infty} \frac{6}{(2i)^2\pi^2} = \frac{1}{4}\sum_{i=1}^{\infty} \frac{6}{\pi^2 i^2} = \frac{1}{4}$$

כאשר הסכום במעבר האחרון הוא בדיוק סכום כל ההסתברויות של כל ה-A_i ולכן שווה ל-1. באופן דומה ניתן לחשב לכל $n \in \mathbb{N}$ את ההסתברות שתוצאת הניסוי מתחלקת ב-n השווה ל- $\frac{1}{n^2}$.

3. נסמן ב-T את המאורע שתוצאת הניסוי מתחלקת ב-3. בסעיף זה אנחנו מתבקשים לחשב את ההסתברות שתוצאת הניסוי מתחלקת ב-2 אך לא ב-3, כלומר את ההסתברות של המאורע $E\cap\overline{T}$. אם נרצה לפתור תרגיל זה בדומה לסעיף הקודם, ניתקל בבעיה שכן יהא עלינו למצוא נוסחא המתארת את כל המספרים הטבעיים הזוגיים שאינם מתחלקים ב-3 ומציאת נוסחא כזו היא משימה מורכבת. במקום זאת נרצה לעבור למאורע המשלים ("כן מתחלק ב-3") אך המעבר כאן מורכב יותר שכן לפי חוקי דה-מורגן, המשלים של המאורע המבוקש הינו "תוצאת הניסוי לא מתחלקת ב-2 או כן מתחלקת ב-3" $\left(\left(\overline{E\cap\overline{T}}\right)=\overline{E}\cup T\right)$ ומאורע זה מסובך גם כן. היות והחלק הבעייתי הוא "לא מתחלק ב-3" נרצה להפעיל רק עליו את המשלים, והדרך לעשות זאת היא באמצעות רישום המאורע E כאיחוד של שני מאורעות זרים: $E=(E\cap T)\cup\left(E\cap\overline{T}\right)$ (ראו נוסחא 1.14).
המאורע $E\cap T$ מתאר תוצאות המתחלקת ב-2 וגם ב-3, כלומר ב-6, וההסתברות לכך היא $\frac{1}{36}$. לפיכך ההסתברות המבוקשת הינה:

$$Pr\left(E\cap\overline{T}\right)=Pr(E)-Pr(E\cap T)=\frac{1}{4}-\frac{1}{36}=\frac{8}{36}=\frac{2}{9}$$

□

תרגיל 2.3. נגדיר מאורעות חדשים וזרים בזוגות שאיחודם שווה ל-$\bigcup_{i=1}^{n}A_i$ על מנת שנוכל להשתמש באקסיומת ההסתברות השלישית לצורך חישוב הסתברות האיחוד. נעשה זאת על ידי כך שבמאורע ה-i נכלול רק תוצאות מתוך A_i שלא נכללו במאורעות הקודמים.
נסמן $B_1=A_1$ ונגדיר $B_2=A_2\backslash A_1$. בבירור המאורעות B_1,B_2 זרים בזוגות ומקיימים $B_1\cup B_2=A_1\cup A_2$.
באותו אופן, נניח כי הגדרנו כבר $k<n$ מאורעות $B_1,B_2,...,B_k$. נגדיר את המאורע ה-$k+1$ על ידי:

$$B_{k+1}=A_{k+1}\setminus(A_1\cup...\cup A_k)$$

כלומר המאורע B_{k+1} כולל את התוצאות הנמצאות רק ב-A_{k+1} ולא במאורעות $A_1,...,A_k$. מאורע זה זר לכל המאורעות הקודמים שהוגדרו ומתקיים $B_1\cup...\cup B_{k+1}=A_1\cup...\cup A_{k+1}$. כמו-כן, בבירור $B_{k+1}\subseteq A_{k+1}$ ולכן $Pr\left(B_{k+1}\right)\leq Pr\left(A_{k+1}\right)$.
לאחר שהגדרנו את כל n המאורעות B, ניתן להשתמש באקסיומת ההסתברות השלישית לחישוב הסתברות האיחוד:

$$Pr\left(\bigcup_{i=1}^{n}A_i\right)=Pr\left(\bigcup_{i=1}^{n}B_i\right)=\sum_{i=1}^{n}Pr\left(B_i\right)\leq\sum_{i=1}^{n}Pr\left(A_i\right)$$

□

תרגיל 2.4. נחשב את ההסתברות של המאורע המשלים ונשתמש בחוקי דה-מורגן על מנת לעבור מההסתברות של החיתוך להסתברות של איחוד, בכדי שנוכל להשתמש באי-שוויון בול (תרגיל 2.3).

$$Pr\left(\overline{\bigcap_{i=1}^{n}A_i}\right)=Pr\left(\bigcup_{i=1}^{n}\overline{A_i}\right)\leq\sum_{i=1}^{n}Pr\left(\overline{A_i}\right)=\sum_{i=1}^{n}\left(1-Pr\left(A_i\right)\right)=n-\sum_{i=1}^{n}Pr\left(A_i\right)$$

לכן ההסתברות של המאורע המבוקש הינה:

$$Pr\left(\bigcap_{i=1}^{n}A_i\right)=1-Pr\left(\overline{\bigcap_{i=1}^{n}A_i}\right)\geq 1-\left(n-\sum_{i=1}^{n}Pr\left(A_i\right)\right)=\sum_{i=1}^{n}Pr\left(A_i\right)-n+1$$

כנדרש. $\square$

תרגיל 2.5. **1.** במקרה כזה, לפי נוסחת ההכלה וההפרדה יתקיים

$$Pr\left(A\cup B\right)=Pr\left(A\right)+Pr\left(B\right)-Pr\left(A\cap B\right)=0.7+0.8-0.4=1.1>1$$

לכן מצב זה לא ייתכן.

2. מצב זה אפשרי, למשל כאשר $B\subseteq\overline{A}$ ואז $\overline{A}\cap B=\overline{A}$ והסתברות החיתוך תהיה $Pr\left(\overline{A}\right)=1-Pr\left(A\right)=0.3$.

$\square$

תרגיל 2.6. נפתח את נוסחת ההכלה וההפרדה ל-3 מאורעות על ידי שימוש חוזר בנוסחת ההכלה וההפרדה ל-2 מאורעות:

$$\begin{aligned}Pr(A\cup B\cup C) &= Pr\left(A\cup(B\cup C)\right)=Pr(A)+Pr(B\cup C)-Pr\left(A\cap(B\cup C)\right)\\ &= Pr(A)+Pr(B)+Pr(C)-Pr(B\cap C)-Pr\left(A\cap(B\cup C)\right)\end{aligned}$$

לפי חוק הפילוג (נוסחא 1.7) מתקיים $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$ ולכן

$$Pr\left(A\cap(B\cup C)\right) = Pr\left((A\cap B)\cup(A\cap C)\right)=Pr(A\cap B)+Pr(A\cap C)-Pr(A\cap B\cap C)$$

נציב זאת בנוסחא הקודמת ונקבל את נוסחת ההכלה וההפרדה ל-3 מאורעות:

$$\begin{aligned}Pr(A\cup B\cup C) &= Pr(A)+Pr(B)+Pr(C)\\ &\quad -Pr(A\cap B)-Pr(A\cap C)-Pr(B\cap C)\\ &\quad +Pr(A\cap B\cap C)\end{aligned}$$

את המקרה הכללי (נוסחא 2.8) ניתן להוכיח באמצעות טכניקה דומה, למשל על ידי אינדוקציה. $\square$

תרגיל 2.7. נסמן ב-R את המאורע שהסטודנט צורך ריטלין וב-C את המאורע שהסטודנט שותה קפה. הנתונים מספקים את ההסתברויות של המאורעות המשלימים, $Pr\left(\overline{R}\right)=0.3, Pr\left(\overline{C}\right)=0.5$ ולכן ההסתברויות של המאורעות עצמם הינם $Pr\left(R\right)=0.7, Pr\left(C\right)=0.5$. בנוסף, נתון $Pr\left(R\cap C\right)=0.35$ ולפי נוסחת ההכלה וההפרדה נקבל

$$Pr\left(R\cup C\right)=Pr\left(R\right)+Pr\left(C\right)-Pr\left(R\cap C\right)=0.7+0.5-0.35=0.85$$

המאורע המבוקש, סטודנטים שלא שותים קפה ולא צורכים ריטלין הינו $\overline{R} \cap \overline{C} = \overline{R \cup C}$ ולכן הסתברותו היא המשלימה של הסתברות האיחוד:

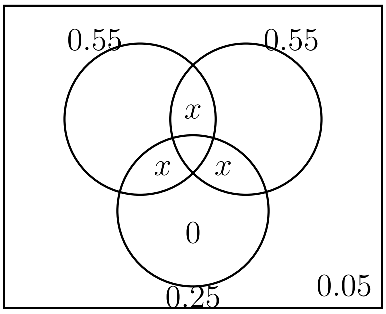
$$Pr\left(\overline{R \cup C}\right) = 1 - Pr\left(R \cup C\right) = 0.15$$

$\square$

תרגיל 2.8. נסמן ב-R את המאורע שרודולף בריא, C את המאורע שקומט בריא ו-V את המאורע שויקסן בריא. לפי הנתונים, $Pr\left(V\right) = Pr\left(C\right) = 0.55$ וכן $Pr\left(\overline{V} \cap \overline{C}\right) = 0.05$. לכן

$$Pr\left(V \cup C\right) = 1 - Pr\left(\overline{V \cup C}\right) = 1 - Pr\left(\overline{V} \cap \overline{C}\right) = 0.95$$

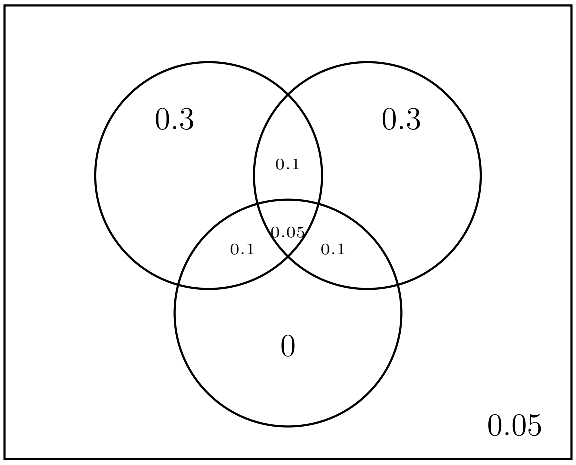
יחד עם זאת, לפי הנתונים מתקיים גם $Pr\left(R \cup C \cup V\right) = 0.95$. לכן בהכרח $R \subseteq V \cup C$ שכן אם הייתה אפשרות לרודולף להיות בריא לבדו, ההסתברות שלפחות איל אחד היה בריא הייתה צריכה להיות גבוהה מההסתברות שלפחות אחד משני האילים הרגילים בריא. לכן $Pr\left(R \setminus \left(V \cup C\right)\right) = 0$. נסמן ב-$x$ את הסיכוי שבדיוק שני אילים בריאים. לפי הנתון, סיכוי זה שווה לכל הזוגות ולכן נקבל את דיאגרמת ון הבאה:



סנטה קלאוס יכול לעוף ברחבי העולם בהסתברות 0.35 וזה קורה כאשר לפחות שני אילים בריאים. מצב זה יכול לקרות כאשר רודולף בריא (בהכרח עוד איל בריא) או כאשר רודולף חולה אבל ויקסן וקומט בריאים. לכן נקבל את המשוואה הבאה:

$$Pr\left(R\right) + x = 0.35$$

ופתרונה $x = 0.1$. בנוסף, המאורע שרודולף בריא ניתן לרישום כאיחוד של שלושה מאורעות זרים: רק רודולף וויקסן בריאים, רק רודולף וקומט בריאים וכולם ביחד בריאים. לפי הנתון, ההסתברות של שני המאורעות הראשונים הינה $x = 0.1$ ולכן ההסתברות שכולם בריאים תהיה $Pr\left(R\right) - 2x = 0.05$. נוסיף נתונים אלו לדיאגרמה ונקבל את ההסתברויות של כל המאורעות הבסיסיים:



כעת אנחנו מוכנים לענות על השאלות: ההסתברות שרק רודולף בריא היא 0 וההסתברות שכל האילים בריאים היא 0.05. □

תרגיל 2.9. נסמן ב-A_i את המאורע שסטודנט ענה על שאלה i במבחן. הסיכוי לכל מאורע כזה הינו $Pr\left(A_i\right) = \frac{|A_i|}{x}$ כאשר $|A_i|$ הינו מספר הסטודנטים הכולל שענה על השאלה ו-x הינו מספר הסטודנטים הכולל שנבחן במבחן. נחשב את ההסתברות שסטודנט מקרי ענה לפחות על שאלה אחת לפי נוסחת ההכלה וההפרדה:

$$\begin{aligned} Pr\left(A_1 \cup A_2 \cup A_3\right) = & Pr\left(A_1\right) + Pr\left(A_2\right) + Pr\left(A_3\right) - Pr\left(A_1 \cap A_2\right) \\ & - Pr\left(A_1 \cap A_3\right) - Pr\left(A_2 \cap A_3\right) + Pr\left(A_1 \cap A_2 \cap A_3\right) \end{aligned}$$

נציב את המספרים המופיעים בטבלה כאשר ההסתברויות לחיתוכים השונים מחושבות באופן דומה להסתברות לענות על כל שאלה:

$$Pr\left(A_1 \cup A_2 \cup A_3\right) = \frac{76}{x} + \frac{72}{x} + \frac{98}{x} - \frac{25}{x} - \frac{41}{x} - \frac{27}{x} + \frac{10}{x} = \frac{163}{x}$$

לפי הנתון, כל סטודנט ענה לפחות על שאלה אחת לכן $Pr\left(A_1 \cup A_2 \cup A_3\right) = 1$ ולכן $x = 163$ סטודנטים נבחנו במבחן.
ניתן היה לפתור שאלה זו גם באמצעות דיאגרמת ון אך הפעם הרבה יותר נוח להשתמש בנוסחא פעם אחת מבלי לחשב את ההסתברות של כל מאורע אפשרי במרחב המדגם. □

תרגיל 2.10. 1. הטענה נכונה: $Pr(A|B) = \frac{Pr(A \cap B)}{PrB)} = \frac{Pr(A \cap B)}{Pr(A)} = Pr(B|A)$.

2. הטענה לא נכונה. נסמן A- בהטלת קוביה יצא 1, B- בהטלת קוביה יצאה תוצאה זוגית. המאורעות זרים ולכן שתי ההסתברויות המותנות שוות ל-0 אבל ברור כי ההסתברות לקבל 1 שונה מההסתברות לקבל תוצאה זוגית.
3. הדוגמא הנגדית מהסעיף הקודם מתאימה גם לסעיף זה (ודאו זאת!)
4. הטענה נכונה:

$$\begin{aligned}
Pr\left(A|B\right)+Pr\left(\overline{A}|B\right) &= \frac{Pr\left(A\cap B\right)}{Pr(B)}+\frac{Pr\left(B\cap \bar{A}\right)}{Pr(B)}\\
&= \frac{Pr\left(A\cap B\right)+Pr\left(B\cap \bar{A}\right)}{Pr(B)}\\
(*) &= \frac{Pr\left(\left(A\cap B\right)\cup\left(B\cap \bar{A}\right)\right)}{Pr(B)}\\
(**) &= \frac{Pr\left(B\cup\left(A\cap \overline{A}\right)\right)}{Pr(B)}=\frac{Pr(B)}{Pr(B)}=1
\end{aligned}$$

כאשר (*) מתקיים כי המאורעות זרים ו-(**) מתקיים מחוק הפילוג.

5. הטענה לא נכונה. הדוגמא של סעיף 2 מתאימה גם כאן, כאשר נגדיר $C=A$. מתקיים $\frac{1}{6}=Pr\left(A\right)<Pr(B)=\frac{1}{2}$ אבל $1=Pr\left(A|C\right)>Pr\left(B|C\right)=0$.

□

תרגיל 2.11. נסמן ב-J את המאורע שיצא נסיך בהוצאה הראשונה, ב-Q את המאורע שיצאה מלכה בהוצאה השניה ו-K את המאורע שיצא מלך בהוצאה השלישית. באמצעות חוק הכפל נקבל:

$$\begin{aligned}
Pr\left(J\cap Q\cap K\right) &= Pr\left(J\right)\cdot Pr\left(Q|J\right)\cdot Pr\left(K|Q\cap J\right)\\
&= \frac{4}{52}\cdot\frac{4}{51}\cdot\frac{4}{50}=0.00048
\end{aligned}$$

שכן בחפיסת קלפים מלאה יש 52 קלפים ו-4 נסיכים (ראו פרק 0.4.1) ולכן הסיכוי לשלוף נסיך הוא $\frac{4}{52}$. אם נשלף נסיך בהוצאה הראשונה, נשארנו עם חפיסה בה 51 קלפים ו-4 מלכות ולכן הסיכוי לשלוף מלכה הוא $\frac{4}{51}$ ולאחר שנשלפה גם המלכה, נשארים עם חפיסה בה 50 קלפים ו-4 מלכים והסיכוי למלך הינו $\frac{4}{50}$. □

תרגיל 2.12. נסמן ב-B_i את המאורע שבהוצאה ה-$i=1,2,3,4$ יצא כדור שחור. ניתן לחשב את המאורע ש-4 הכדורים שחורים על ידי התניה על המאורע ש-3 הכדורים הראשונים שחורים:

$$Pr\left(B_4\right)=Pr\left(B_4|B_3\right)\cdot Pr\left(B_3\right)=\frac{18}{23}Pr\left(B_3\right)$$

שכן בהינתן שבשלושת ההוצאות הראשונות יצאו כדורים שחורים, יש לפני ההוצאה הרביעית בכד 5 כדורים לבנים ו-$18=3+3\cdot 5$ כדורים שחורים. באופן דומה ניתן לחשב את המאורע ששלושת הכדורים הראשונים שחורים על ידי התניה על כך ששני הראשונים שחורים ואת המאורע ששני הראשונים שחורים על ידי התניה על כך שהכדור הראשון שחור. בסיכומו של דבר לפי נוסחת הכפל נקבל

$$Pr\left(B_4\right)=Pr\left(B_4|B_3\right)\cdot Pr\left(B_3|B_2\right)\cdot Pr\left(B_2|B_1\right)\cdot Pr\left(B_1\right)=\frac{18}{23}\cdot\frac{13}{18}\cdot\frac{8}{13}\cdot\frac{3}{8}=\frac{3}{23}$$

□

תרגיל 2.13. 1. נתבונן בזריקה כלשהי לסל של לארי בירד ונסמן ב-A את המאורע שנקלע סל בזריקה זו. מאחר ויש קשר בין זריקות עוקבות, נסמן ב-B את המאורע שלארי בירד קלע את הזריקה הקודמת. נחשב את ההסתברות למאורע A לפי נוסחת ההסתברות השלמה:

$$Pr(A) = Pr(A|B)\cdot Pr(B) + Pr\left(A|\overline{B}\right)\cdot Pr\left(\overline{B}\right) = 0.49Pr(B) + 0.38\left(1 - Pr(B)\right)$$

בהיעדר מידע על הזריקה הקודמת, הסיכוי לסל בכל אחת מהזריקות הוא קבוע וזהה (זה מה שעלינו לחשב) ולכן $Pr(A) = Pr(B)$. לכן הנוסחא הקודמת היא בעצם משוואה ועל ידי פתירתה נקבל $Pr(A) = 0.43$.

2. במחקר התגלה שלארי בירד קולע 44% מזריקותיו ולא 43% כפי שחישבנו בסעיף הקודם. הסיבה לכך היא שבמציאות גם זריקות מוקדמות יותר משפיעות על הזריקה הנוכחית. למשל, במחקר התגלה כי הסיכוי לסל אחרי שתי זריקות מוצלחות הינו 51% והסיכוי לסל אחרי שלוש זריקות מוצלחת הינו 47% ואילו בסעיף הראשון הנחנו שרק הזריקה הקודמת משפיעה על הזריקה הנוכחית. לכן הנחנו שגם אחרי שתי זריקות מוצלחות הסיכוי לסל צריך להיות 49%. הסיבה שההסתברות לסל משתנה היא שבפועל במשחק כדורסל הזריקות תלויות. למשל, ככל ששחקן קולע יותר ההגנה תתמקד בו יותר ותפריע לו יותר.

3. נחשב באמצעות נוסחת ההסתברות השלמה. נסמן ב-A את המאורע שנקלע סל בזריקה השלישית וב-B את המאורע שנקלע סל בזריקה השנייה. היות ונקלע סל בזריקה הראשונה, $Pr(B) = 0.49$ ולכן

$$Pr(A) = Pr(A|B)\cdot Pr(B) + Pr\left(A|\overline{B}\right)\cdot Pr\left(\overline{B}\right) = 0.49\cdot 0.49 + 0.38\cdot 0.51 = 0.433$$

ניתן לרשום את נוסחת ההסתברות השלמה גם במונחים של כפל מטריצות ובכך לקבל נוסחא כללית הקושרת בין תוצאת הזריקה הראשונה לתוצאת הזריקה ה-n, אך טכניקה זו חורגת ממסגרת הדיון הנוכחי.

$\square$

תרגיל 2.14. 1. נשתמש בנוסחת ההסתברות השלמה ונתנה על הפתק הראשון שיצא. אם הפתק הראשון שיצא הוא m (נסמן מאורע זה ב- A_m) אז עבור הפתק השני יש רק $n-m$ אפשרויות שגדולות ממש מ-m. לפיכך, ההסתברות של המאורע B, שהפתק השני גדול מהראשון היא:

$$\begin{aligned} Pr\left(B\right) &= \sum_{m=1}^{n} Pr\left(B|A_m\right)\cdot Pr\left(A_m\right) \\ &= \sum_{m=1}^{n}\frac{n-m}{n}\cdot\frac{1}{n} \\ &= 1 - \frac{1}{n^2}\sum_{m=1}^{n} m = \frac{n-1}{2n} = \frac{1}{2} - \frac{1}{2n} \end{aligned}$$

עבור $n = 2$ אנו מקבלים $Pr(B) = \frac{1}{4}$ וקל לראות שזו אכן התוצאה הנכונה. מתוך 4 התוצאות האפשריות: $\{(1,1),(1,2),(2,1),(2,2)\}$ רק אחת עונה על הקריטריון הנדרש.
דרך נוספת לפתור את השאלה היא באמצעות שיקולי סימטריה. נסמן ב-p את ההסתברות שהפתק השני יהיה גדול ממש מהפתק הראשון. מסימטריה, זו גם ההסתברות שהפתק הראשון יהיה גדול מהפתק השני. האפשרות השלישית היא ששני הפתקים יהיו שווים. ניתן לחשב זאת באופן דומה לאופן בו חושב המאורע

B על ידי נוסחת ההסתברות השלמה או פשוט על ידי כך שנזכור שההוצאה של הפתק הראשון לא מוגבלת ולא משפיעה על המאורע אבל אחרי שערכו של הפתק הראשון ידוע - הסיכוי שהפתק השני יהיה זהה לו מתוך n הפתקים הוא בדיוק $\frac{1}{n}$. שלושת אפשריות אלו זרות בזוגות ובדיוק אחת מהן תיתכן בכל חזרה על הניסוי, לכן מתקיימת המשוואה

$$p+p+\frac{1}{n}=1$$

שפתרונה $p=\frac{1}{2}-\frac{1}{2n}$ כמקודם.

2. מאחר וכל הפתקים זהים וכל המספרים סימטריים, הסיכוי שבהוצאה הראשונה יצא i ובהוצאה השניה יצא j שווה לסיכוי שבו j יוצא בהוצאה הראשונה ו-i בשנייה. לכן חצי מהאפשרויות מתאימות למאורע בו הפתק הראשון גדול מהשני וחצי מהאפשרויות מתאימות למאורע בו הפתק השני גדול מהראשון וההסתברות המבוקשת הינה $\frac{1}{2}$. כאשר יש בדיוק 2 פתקים, מרחב המדגם כולל שתי אפשרויות בלבד: $\{(1,2),(2,1)\}$ ואכן הסיכוי שהפתק השני גדול מהראשון הוא חצי.

$\square$

תרגיל 2.15. נסמן ב-E את המאורע שהעץ הראשון התקבל בהטלה זוגית ונתנה על תוצאת ההטלה הראשונה. אם ההטלה הראשונה היא עץ (המאורע H)- המאורע לא קרה. אם ההטלה הראשונה היא פלי $\left(\overline{H}\right)$, החל מההטלה השניה אנחנו למעשה מתחילים את הניסוי מחדש, רק שהפעם דורשים שהעץ הראשון בניסוי החדש יגיע בהטלה אי זוגית (שכן זו הטלה זוגית, כשסופרים את ההטלות של הניסוי המקורי):

$$\begin{aligned} Pr\left(E\right) &= Pr\left(E|H\right)\cdot Pr\left(H\right)+Pr\left(E|\overline{H}\right)\cdot Pr\left(\overline{H}\right) \\ &= 0+(1-Pr(E))(1-p) \end{aligned}$$

מפתרון המשוואה מקבלים $Pr\left(E\right)=\frac{1-p}{2-p}$.
שימו לב לתוצאה מעניינת - כצפוי, כאשר $p=1$ תמיד יתקבל עץ בהטלה הראשונה ולא בהטלה זוגית ואכן $Pr(E)=0$. לעומת זאת, כאשר $p=0$ יתקבל $Pr(E)=\frac{1}{2}$, אך במקרה כזה אף פעם לא מתקבל עץ, אז איך ההסתברות שיתקבל בעץ בהטלה זוגית הוא חצי ולא 0? הבעיה בהצבה זו היא שהנוסחא פותחה תוך שימוש בנוסחת ההסתברות השלמה והתנייה על המאורע שהתקבל עץ בהטלה הראשונה. כאשר $p=0$ התננו על מאורע עם הסתברות אפס וזה לא חוקי.
מדוע בכל זאת כאשר $p\to 0$ נקבל $Pr(E)\to\frac{1}{2}$? הסיבה היא שהסיכוי לעץ במקרה זה הינו מאוד נמוך ולכן אין כבר יתרון להטלות האי זוגיות. באופן כללי, ההסתברות לעץ בהטלה הראשונה הוא p ואילו ההסתברות לעץ לראשונה בהטלה השניה הוא $p<p(1-p)$. לכן, הסיכוי לעץ בהטלה אי זוגית תמיד גדול יותר מהסיכוי לעץ בהטלה הזוגית העוקבת, כך שיש סיכוי גדול מחצי שהעץ הראשון יפול בהטלה אי זוגית (שימו לב ש $Pr(E)<\frac{1}{2}$). כאשר $p\to 0$ היתרון של ההטלות האי זוגיות הולך וקטן והסיכוי לעץ בכל אחת משתי הטלות עוקבות הופך לזהה. $\square$

תרגיל 2.16. נחשב את ההסתברות של המאורע $A\,\text{before}\,B$ על ידי התניה על תוצאת הניסוי הראשון. מאחר והמאורעות A ו-B זרים ייתכנו שלוש אפשרויות: מאורע A קרה, מאורע B קרה או אף אחד משניהם לא קרה, כלומר $\overline{A}\cap\overline{B}$ קרה. לפיכך:

$$\begin{aligned} Pr\left(A \text{ before } B\right) &= Pr\left(A \text{ before } B | \text{first } A\right) Pr\left(A\right) \\ &+ Pr\left(A \text{ before } B | \text{first } B\right) Pr\left(B\right) \\ &+ Pr\left(A \text{ before } B | \text{first } \overline{A} \cap \overline{B}\right) Pr\left(\overline{A} \cap \overline{B}\right) \end{aligned}$$

אם בחזרה הראשונה על הניסוי המאורע A קרה אז בהכרח הוא קרה לפני B כלומר $Pr\left(A \text{ before } B | \text{first } A\right) = 1$. מאותה סיבה, $Pr\left(A \text{ before } B | \text{first } B\right) = 0$. בנוסף, אם בחזרה הראשונה אף אחד משני המאורעות לא קרה, אנחנו חוזרים על הניסוי ולפי הנתון ההסתברויות לקבל את שני המאורעות לא משתנות. כלומר, החל מהחזרה השניה אנחנו מבצעים את בדיוק את אותו הניסוי כמו החל מהחזרה הראשונה, והסיכוי ש-A יקרה לפני B זהה בניסוי המתחיל בחזרה השנייה לזה המתחיל בחזרה הראשונה. לכן $Pr\left(A \text{ before } B | \text{first } \overline{A} \cap \overline{B}\right) = Pr\left(A \text{ before } B\right)$. לסיכום, קיבלנו את המשוואה:

$$Pr\left(A \text{ before } B\right) = Pr(A) + Pr\left(A \text{ before } B\right)\left(1 - Pr(A) - Pr(B)\right)$$

ופתרונה $Pr\left(A \text{ before } B\right) = \frac{Pr(A)}{Pr(A)+Pr(B)}$. □

תרגיל 2.17. נניח שמתקיים $Pr\left(A|B\right) = Pr(A)$. נציב את נוסחא **2.9** במקום ההסתברות המותנית ונקבל $Pr(A) = \frac{Pr(A\cap B)}{Pr(B)}$, כלומר $Pr(B) = \frac{Pr(A\cap B)}{Pr(A)}$. אגף ימין של המשוואה האחרונה הוא בדיוק ההסתברות המותנית $Pr\left(B|A\right)$ וקיבלנו $Pr\left(B|A\right) = Pr(B)$, כנדרש. הוכחת הכיוון השני זהה. □

תרגיל 2.18. נניח כי המאורעות A ו-B בלתי תלויים. לכן

$$Pr\left(\overline{A}|B\right) \overset{\text{eq. } \mathbf{2.11}}{=} 1 - Pr\left(A|B\right) = 1 - Pr(A) = Pr\left(\overline{A}\right)$$

כלומר גם המאורעות $\overline{A}$ ו-B בלתי תלויים. באופן זהה, גם המאורעות A ו-$\overline{B}$ בלתי תלויים ולכן גם המאורעות $\overline{A}$ ו-$\overline{B}$ בלתי תלויים. □

תרגיל 2.19. נניח שהמאורע A בלתי תלוי בכל מאורע אחר. אנחנו יודעים שכל מאורע תלוי במאורעות הזרים לו (ראו תרגיל **2.20**) ולכן נתבונן במאורע כזה, $\overline{A}$. לפי הנתון, A ו-$\overline{A}$ בלתי תלויים ולכן מתקיים $Pr\left(A \cap \overline{A}\right) = Pr(A) \cdot Pr\left(\overline{A}\right)$.

יחד עם זאת, $A \cap \overline{A} = \emptyset$ ולכן באגף שמאל רשומה ההסתברות של המאורע הריק, **0**. לכן בהכרח $Pr(A) = 0$ או $Pr\left(\overline{A}\right) = 0$, כלומר $Pr(A) = 1$. לפיכך מאורע יכול להיות בלתי תלוי בכל מאורע אחר רק אם הסתברותו היא **0** (ואז הוא לעולם לא מתרחש, ולא משנה איזה מאורע כן התרחש בניסוי) או **1** (ואז הוא תמיד מתרחש, ולא משנה איזה מאורע כן התרחש בניסוי). □

תרגיל 2.20. מאורעות זרים בהכרח יהיו תלויים, שהרי אם משמעות הדבר שהמאורעות זרים היא שהם אינם יכולים להתרחש בו זמנית. לכן, אם אחד המאורעות התרחש, בוודאי השני לא התרחש וההסתברות המותנית שאחד התרחש בהינתן השני תהיה אפס. ניתן להראות זאת בחישוב ישיר: המאורעות A, B זרים ולכן מתקיים $Pr\left(A \cap B\right) = 0$. מצד שני, המאורעות בעלי הסתברות חיובית ולכן $Pr(A) \cdot Pr(B) \neq 0$ ולכן המאורעות תלויים. תרגיל זה ממחיש שמאורעות זרים בהכרח תלויים, אך מאורעות שאינם זרים יכולים להיות תלויים או לא, כתלות בהסתברויות שלהם. □

תרגיל 2.21. נסמן ב-E את המאורע שהפיל ער ($Pr(E) = 0.75$), ב-H את המאורע שההיפופוטם ער ($Pr(H) = 0.8$) וב-L את המאורע שהאריה ער ($Pr(L) = 0.2$). נחשב את ההסתברות של המאורע $E \cup H \cup L$ בשתי דרכים שונות:

1. לפי נוסחת ההכלה וההפרדה:

$$\begin{aligned} Pr\left(E \cup H \cup L\right) &= Pr(E) + Pr(H) + Pr(L) \\ &- Pr\left(E \cap H\right) - Pr\left(E \cap L\right) - Pr\left(H \cap L\right) \\ &+ Pr\left(E \cap H \cap L\right) = 0.75 + 0.8 + 0.2 \\ &- 0.75 \cdot 0.8 - 0.75 \cdot 0.2 - 0.8 \cdot 0.2 + 0.75 \cdot 0.8 \cdot 0.2 \\ &= 0.96 \end{aligned}$$

2. לפי חוקי דה-מורגן:

$$\begin{aligned} Pr\left(E \cup H \cup L\right) &= 1 - Pr\left(\overline{E \cup H \cup L}\right) \\ &= 1 - Pr\left(\overline{E} \cap \overline{H} \cap \overline{L}\right) = 1 - 0.25 \cdot 0.2 \cdot 0.8 \\ &= 0.96 \end{aligned}$$

□

תרגיל 2.22. מאורעות המוכלים אחד בשני בהכרח תלויים, שהרי אם $A \subseteq B$ וידוע ש-A קרה, בהכרח גם B קרה: $Pr\left(B|A\right) = \frac{Pr(A \cap B)}{Pr(A)} = \frac{Pr(A)}{Pr(A)} = 1 > Pr(B)$. □

תרגיל 2.23. הטענה אינה נכונה. למשל נתבונן בניסוי בו מטילים פעמיים מטבע הוגן שעל צד אחד שלו רשום 0 ועל הצד השני 1. נגדיר A - בהטלה הראשונה יצא 0, B - בהטלה השניה יצא 1 ו-C - סכום שתי ההטלות הוא מספר אי זוגי. בבירור $Pr\left(A\right) = Pr\left(B\right) = \frac{1}{2}$. בנוסף, $Pr\left(C\right) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$ שכן סכום אי זוגי יכול להופיע כאשר בהטלה הראשונה יצא 0 ובשנייה 1 או להפך. יחד עם זאת, שלושת המאורעות זרים ולכן

$$Pr\left(A \cap B \cap C\right) = 0 \neq \frac{1}{8} = Pr\left(A\right) \cdot Pr\left(B\right) \cdot Pr\left(C\right)$$

□

תרגיל 2.24. נשתמש בסימוני השאלה: T הוא המאורע שהכדורסלן זורק מ-3 ו-B_i הוא המאורע שהוא קלע את הזריקה ה-$i = 1, 2$. לו היינו יודעים את המקום ממנו נזרקת הזריקה חישוב הסתברות הקליעה הייתה מיידית אך כאן נשאלנו על ההתנייה ההפוכה ולכן נשתמש בנוסחת בייס:

1. $Pr\left(\overline{T}|B_1\right) = \frac{Pr\left(B_1|\overline{T}\right) \cdot Pr\left(\overline{T}\right)}{Pr(B_1)} = \frac{0.9 \cdot \frac{5}{6}}{\frac{49}{60}} = \frac{45}{49} \simeq 0.92$.

מראש, הסיכוי שהכדורסלן זורק מהעונשין הוא מאוד גבוה $\left(\frac{5}{6} \simeq 0.83\right)$. בנוסף, מהעונשין השחקן קולע בהסתברות מאוד גבוהה ואילו מקשת השלוש הוא קולע בהסתברות נמוכה, לכן סל מחזק את ההסתברות שהוא נמצא על קו העונשין והפוסטיריור כצפוי גדל. באותו אופן ניתן לחשב ולקבל $Pr\left(T|\overline{B}_1\right) = \frac{6}{11} \simeq 0.55$. כלומר אחרי החטאה יש כמעט אותו סיכוי שהוא זרק מהעונשין ומשלוש, למרות שהחטאות בדרך כלל קורות בזריקות משלוש ולא מהעונשין. הסיבה לכך היא שההסתברות הנמוכה להחטיא מהעונשין מאוזנת על ידי הפריור - ישנה הסתברות גבוהה לזרוק מהעונשין ביחס להסתברות הנמוכה לזרוק משלוש.

2. $Pr\left(\overline{T}\middle|B_1\cap B_2\right)=\frac{Pr\left(B_1\cap B_2|\overline{T}\right)\cdot Pr\left(\overline{T}\right)}{Pr(B_1\cap B_2)}=\frac{0.9^2\cdot\frac{5}{6}}{\frac{421}{600}}=\frac{405}{421}\simeq 0.96$.
ההסתברות לקלוע שני סלים (במכנה) חושבה בנוסחא 2.20 שם הראנו שהמאורע "סל בזריקה הראשונה" ו-"סל בזריקה השנייה" תלויים. בסעיף הקודם ראינו זאת כמותית - סל בזריקה הראשונה הגדיל את הסיכוי שהזריקה בוצעה מקו העונשין ולכן גם הגדיל את הסיכוי לסל בזריקה הבאה. כצפוי, העובדה שנקלעו שני סלים הגדילה עוד יותר את ההסתברות שהכדורסלן זורק מהעונשין.

3. $Pr\left(\overline{T}\middle|B_1\cap \overline{B}_2\right)=\frac{Pr\left(B_1\cap\overline{B}_2|\overline{T}\right)\cdot Pr\left(\overline{T}\right)}{Pr\left(B_1\cap\overline{B}_2\right)}=\frac{0.9\cdot 0.1\cdot\frac{5}{6}}{0.9\cdot 0.1\cdot\frac{5}{6}+0.4\cdot 0.6\cdot\frac{1}{6}}=\frac{45}{69}\simeq 0.65$.
החטאה מקו העונשין היא מאורע נדיר (הסתברות של 0.1 בלבד) ולכן ההחטאה מורידה משמעותית את הסיכוי שהוא אכן זורק משם.

□

תרגיל 2.25. נסמן ב-S את המאורע שהאדם חולה וב-$\oplus$ את המאורע שהתקבלה תשובה חיובית. לפי נוסחת בייס:

$$\begin{aligned}Pr\left(S|\oplus\right)&=\frac{Pr\left(\oplus|S\right)\cdot Pr(S)}{Pr\left(\oplus\right)}=\frac{Pr\left(\oplus|S\right)\cdot Pr(S)}{Pr\left(\oplus|S\right)\cdot Pr(S)+Pr\left(\oplus|\overline{S}\right)\cdot Pr\left(\overline{S}\right)}\\&=\frac{0.95\cdot\frac{1}{10{,}000}}{0.95\cdot\frac{1}{10{,}000}+0.05\cdot\frac{9{,}999}{10{,}000}}=0.001897\end{aligned}$$

ותוצאה זו קרובה מאוד לחישוב האינטואיטיבי $\frac{1}{501}$. □

תרגיל 2.26. נניח לשם הפשטות שהמתמודד תמיד בוחר את דלת מספר 1 ונסמן ב-A_i את המאורע שהמכונית מאחורי דלת מספר i. המשחק הוגן ולכן $Pr\left(A_i\right)=\frac{1}{3}$.
אם המתמודד לא מחליף את בחירתו, הסיכוי שלו לזכות הוא $Pr\left(A_1\right)=\frac{1}{3}$ וכלל לא משנה מה המנחה עושה. לכן, הסיכוי של המתמודד לזכות אם כן יחליף את בחירתו היא המאורע המשלים, $Pr\left(\overline{A_1}\right)=\frac{2}{3}$. החלפה, כך מסתבר, מכפילה את הסיכוי לזכות! □

תרגיל 2.27. בשלושת הסעיפים אנחנו נדרשים לחשב את ההסתברות שלפחות מאורע אחד מתרחש, כלומר את ההסתברות של המאורע $A\cup B\cup C$.

1. המאורעות זרים בזוגות ולכן לפי אקסיומה 2.3 מתקיים:

$$Pr\left(A\cup B\cup C\right)=Pr\left(A\right)+Pr\left(B\right)+Pr\left(C\right)=1$$

2. במקום לחשב את ההסתברות שלפחות מאורע אחד מתרחש נחשב את ההסתברות של המשלים, שאף אחד מהמאורעות לא מתרחש ולצורך כך נשתמש בחוקי דה-מורגן (משוואה 1.11):

$$\begin{aligned}Pr\left(A\cup B\cup C\right)&=1-Pr\left(\overline{A\cup B\cup C}\right)=1-Pr\left(\overline{A}\cap\overline{B}\cap\overline{C}\right)\\&\overset{*}{=}1-Pr\left(\overline{A}\right)\cdot Pr\left(\overline{B}\right)\cdot Pr\left(\overline{C}\right)=1-0.28=0.72\end{aligned}$$

כאשר במעבר $*$ השתמשנו בעובדה שאם מאורעות בלתי תלויים אז גם משלימיהם בלתי תלויים (תרגיל 2.18).
הערה: ניתן להשתמש גם בנוסחת ההכלה וההפרדה לשלושה מאורעות, אך דרך זו מצריכה יותר חישובים וארוכה יותר.

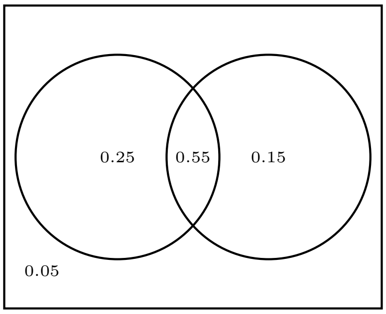
3. המאורעות מוכלים זה בזה ולכן $A \cup B \cup C = A$, כך שההסתברות המבוקשת הינה $Pr(A) = 0.5$.

□

תרגיל 2.28. נסמן ב-A את המאורע שהשחקן קולע מקו העונשין וב-B את המאורע שהוא קולע מקשת השלוש. לפי הנתונים, $Pr(A) = 0.8, Pr(B) = 0.7$. ההסתברות לפחות לסל אחד הינו $Pr(A \cup B) = 0.95$. לפי נוסחת ההכלה וההפרדה ההסתברות שיקלע שני סלים הינה

$$Pr(A \cap B) = Pr(A) + Pr(B) - Pr(A \cup B) = 0.8 + 0.7 - 0.95 = 0.55$$

לכן הסיכוי שיקלע רק מהעונשין הינו $0.8 - 0.55 = 0.25$ והסיכוי שיקלע רק מקשת השלוש הינו $0.7 - 0.55 = 0.15$. נציב את כל הערכים שמצאנו עד כה בדיאגרמת ון:



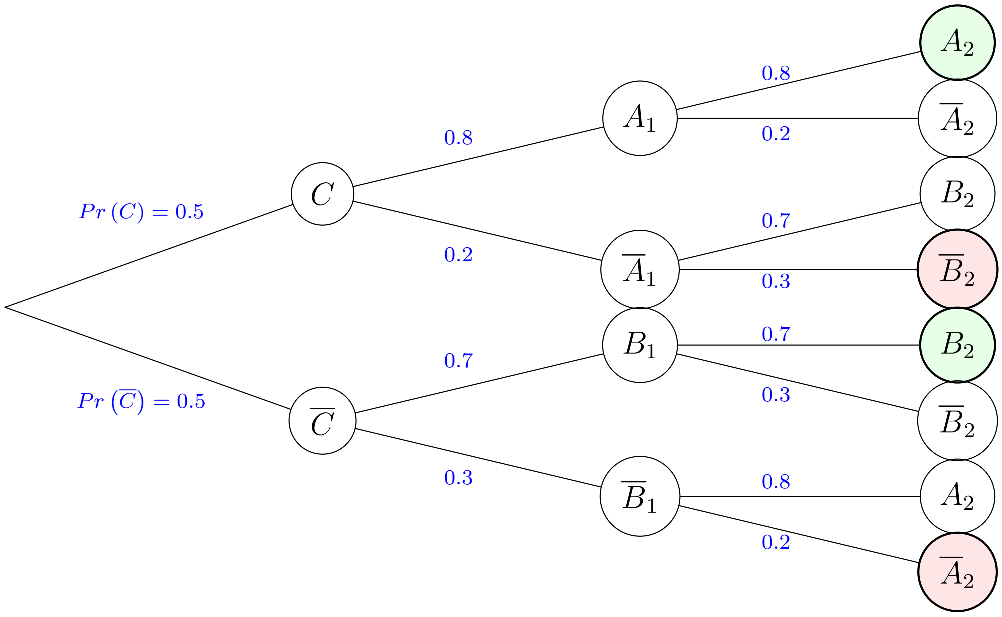
כעת ניתן לענות על השאלות:

1. הכדורסלן קולע פעמיים בהסתברות $Pr(A \cap B) = 0.55$.
2. הכדורסלן קולע פעם אחת בדיוק בהסתברות $Pr(A \setminus B) + Pr(B \setminus A) = 0.25 + 0.15 = 0.4$.
3. כדי שהכדורסלן יקלע בדיוק פעם אחת אם הוא קלע מהעונשין, עליו להחטיא מקשת השלוש. לכן ההסתברות המבוקשת היא

$$Pr\left(\overline{B}|A\right) = \frac{Pr\left(A \cap \overline{B}\right)}{Pr(A)} = \frac{0.25}{0.8} = 0.3125$$

□

תרגיל 2.29. נסמן ב-A_i את המאורע שהכדורסלן קלע מקו העונשין בזריקה ה-$i \in \{1,2\}$ (אם נזרקו) וב-B_i את המאורע שהכדורסלן קלע מקשת השלוש בזריקה ה-$i \in \{1,2\}$ (אם נזרקו). בנוסף, נסמן ב-C את המאורע שהכדורסלן בחר לזרוק בפעם הראשונה מקו העונשין. נשרטט את עץ ההסתברויות המתאים:



לצורך הפשטות, סומנו עלי העץ אשר התקבלו לאחר שני סלים בירוק ואלו אשר התקבלו לאחר שתי החטאות באדום. שאר העלים התקבלו לאחר בדיוק סל אחד. ההסתברות של כל עלה מתקבל לפי מכפלת ההסתברויות בדרך המובילה אליו וההסתברות של כל מאורע היא סכום ההסתברויות של העלים המתאימים למאורע זה:

1. הסיכוי לשני סלים הינו

$$\frac{1}{2} \cdot 0.8 \cdot 0.8 + \frac{1}{2} \cdot 0.7 \cdot 0.7 = 0.565$$

2. הסיכוי לשתי החטאות הינו

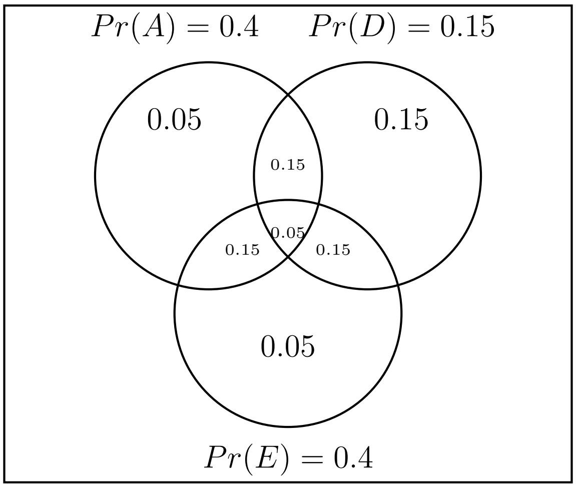
$$\frac{1}{2} \cdot 0.2 \cdot 0.3 + \frac{1}{2} \cdot 0.3 \cdot 0.2 = 0.06$$

3. את ההסתברות לסל אחד בדיוק ניתן לחשב או על ידי המשלים של שתי התוצאות הקודמות או באופן ישיר על ידי העץ:

$$\frac{1}{2} \cdot 0.8 \cdot 0.2 + \frac{1}{2} \cdot 0.2 \cdot 0.7 + \frac{1}{2} \cdot 0.7 \cdot 0.3 + \frac{1}{2} \cdot 0.3 \cdot 0.8 = 0.375$$

□

תרגיל 2.30. נסמן ב-A את המאורע שאביטל נעדרת, ב-D את המאורע שדניאל נעדר וב-E את המאורע שאלירן נעדר. לפי הנתון, כל זוג נעדר מ-20% מההופעות אבל השלישייה כולה נעדרת מ-5%מההופעות, לכן הסיכוי שזוג מסויים ייעדר בעוד שהשלישי מגיע להופעה הינו 15%. לאור מסקנה זו, נשלים את הנתונים החסרים בדיאגרמת הון המתאימה:



כעת ניתן לענות בקלות על השאלות:

1. במופע אין מחליפים כלל אם כל האקרובטים מגיעים. הסיכוי שלפחות אחד מהם נעדר הינו

$$Pr\left(A\cup D\cup E\right) = 0.05+0.15+0.05+3\cdot 0.15+0.05 = 0.75$$

ולכן הסיכוי שלא יהיו מחליפים הוא המשלים, 0.25 (ומספר זה צריך להיות מחוץ לכל העיגולים בדיאגרמת ון).

2. בסעיף הקודם חישבנו את הסיכוי שאין מחליפים כלל, כלומר $k=0$. עבור $k=1$ יש בדיוק מחליף אחד, כלומר בדיוק אקרובט אחד לא הגיע והמאורעות המתאימים הם אלו שנמצאים רק בתוך מאורע אחד, כך שההסתברות הינה $0.25 = 0.05+0.15+0.05$. המאורע $k=2$ מתרחש כאשר יש בדיוק 2 מחליפים והמאורעות המתאימים הם אלו שבחיתוך של שני מאורעות אך לא בחיתוך של כולם ביחד, וההסתברות המתאימה היא $0.45 = 0.15+0.15+0.15$. לסיום, הסיכוי שכל האקרובטים נעדרים הוא 0.05 וזו ההסתברות ש-$k=3$.

□

תרגיל 2.31. 1. מאחר שהילדים בלתי תלויים, הצעיר יכול להיות בן או בת בהסתברות שווה ללא תלות בכך שהבכור, שפתח את הדלת, הוא בן. לכן ההסתברות שבמשפחה שני בנים היא $\frac{1}{2}$.

2. נסמן ב-A את המאורע שבמשפחה שני בנים וב-B את המאורע שבן פתח את הדלת. נשתמש בחוק בייס בכדי לחשב את ההסתברות המתאימה:

$$Pr\left(A|B\right) = \frac{Pr\left(B|A\right)Pr\left(A\right)}{Pr\left(B\right)}$$

כל ילד יכול להיות בן או בת בהסתברות שווה, לכן הסיכוי שבמשפחה שני בנים הינו $Pr\left(A\right) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$. כמו כן, אם במשפחה שני בנים, בוודאי שבן יפתח את הדלת כך ש-$Pr\left(B|A\right) = 1$. נחשב את הסיכוי שבן יפתח את הדלת באמצעות נוסחת ההסתברות השלמה, כאשר המאורעות עליהם אנחנו מתנים זה צורת המשפחה: שני בנים בהסתברות רבע, שתי בנות בהסתברות רבע (נסמן ב- GG) או בן ובת בהסתברות חצי (נסמן ב- BG):

$$\begin{aligned}Pr\left(B\right) &= Pr\left(B|A\right)Pr\left(A\right)+Pr\left(B|BG\right)Pr\left(BG\right)+Pr\left(V|GG\right)Pr\left(GG\right)\\ &= 1\cdot\frac{1}{4}+\frac{1}{2}\cdot\frac{1}{2}+0\cdot\frac{1}{4}=\frac{1}{2}\end{aligned}$$

לכן, כמו בסעיף הקודם, ההסתברות המותנית הינה

$$Pr\left(A|B\right)=\frac{\frac{1}{4}}{\frac{1}{2}}=\frac{1}{2}$$

תשובה מוטעית אפשרית היא שהסיכוי הוא $\frac{1}{3}$: במקור, כל אחת מ-4 האפשרויות להרכב המשפחה (שני בנים, שתי בנות, בן בכור ובת או בת בכורה ובן) הן שוות הסתברות. בהינתן שבן פתח את הדלת, האפשרות ששתי בנות אינה אפשרית ולכן הסיכוי לשני בנים אמור להיות $\frac{1}{3}$ כי זו אפשרות אחת מתוך שלוש. אולם, אפשרות זו אינה סימטרית לשאר האפשרויות שכן היא כוללת שתי תתי-אפשרויות לגבי זהות הבן שפתח את הדלת, בן א' או בן ב'. באופן כללי, מאורעות שהיו סימטריים במרחב מדגם אחד לאו דווקא נשארים סימטריים בהינתן מידע נוסף והמידע הנוסף כאן הוא שבן פתח את הדלת. □

תרגיל 2.32. 1. מאחר וצבע שערו של האב שחור, בהכרח יש לו לפחות גן אחד של צבע שחור. הגן השני יכול להיות שחור גם כן או ג'ינג'י. נסמן את המאורע ששני הגנים של האב שחורים ב-B.

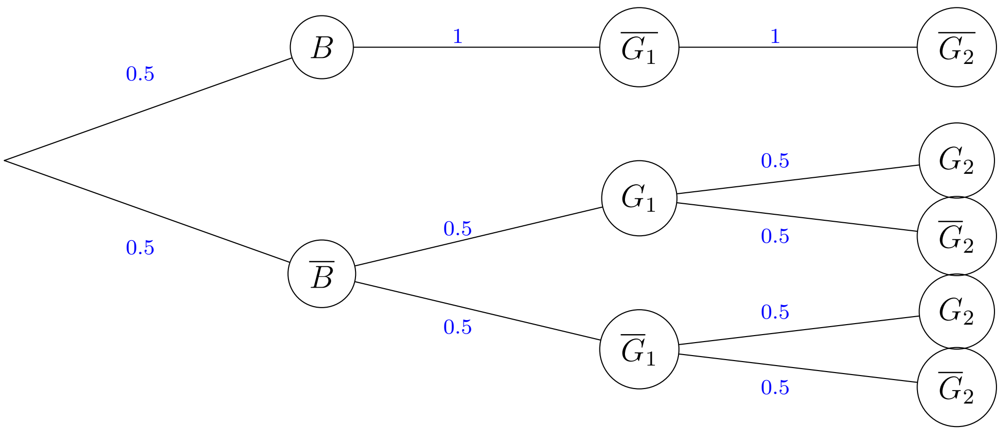
2. מאחר והאמא ג'ינג'ית, היא בהכרח תורמת גן ג'ינג'י לכל אחד מבניה ולכן צבע שיערם נקבע לפי האב – אם הם מקבלים ממנו גם גן ג'ינג'י יהיו ג'ינג'ים ואחרת שחורי שיער. לו היינו יודעים מהם הגנים של האב ניתן היה לחשב את ההסתברות שהם ג'ינג'ים. בהעדר מידע זה, נתנה עליו לפי נוסחת ההסתברות השלמה. נסמן את המאורע שהבן ה-$i\in\{1,2\}$ ג'ינג'י ב-G_i ונקבל:

$$Pr\left(G_i\right)=Pr\left(G_i|B\right)Pr\left(B\right)+Pr\left(G_i|\overline{B}\right)Pr\left(\overline{B}\right)=0\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$$

שכן בהנתן שלאב גן שחור וגן ג'ינג'י, הילד יהיה ג'ינג'י בהסתברות חצי – רק אם יירש את הגן הג'ינג'י.

3. צבע השיער של שני הילדים תלוי שכן הוא מרמז על הגנים של אביהם. אם למשל הבכור ייצא ג'ינג'י, זה אומר שלאב יש גן ג'ינג'י מה שמגדיל את הסיכוי של הבן הצעיר להיות ג'ינג'י גם כן. לעומת זאת, אם הבכור שחור שיער, זה מקטין את הסיכוי שהצעיר יהיה ג'ינג'י כי זה מרמז על כך שלאב כנראה אין גן ג'ינג'י.

4. נשרטט את העץ המתאים לבעיה:



ההסתברות המבוקשת היא ההסתברות המותנית $Pr\left(G_2|\overline{G}_1\right) = \frac{Pr\left(G_2\cap\overline{G}_1\right)}{Pr\left(\overline{G}_1\right)}$. המכנה הינו פשוט המשלים של החישוב שביצענו כבר ולכן $Pr\left(\overline{G}_1\right) = \frac{3}{4}$. את המונה נחשב באופן דומה – נתנה על הגנים של האב ונחשב את ההסתברות של החיתוך בהתאם להסתברות השלמה כפי שמוצג בעץ:

$$Pr\left(\overline{G}_1\cap G_2\right) = Pr\left(\overline{G}_1\cap G_2|B\right)Pr\left(B\right)+Pr\left(\overline{G}_1\cap G_2|\overline{B}\right)Pr\left(\overline{B}\right) = 1\cdot 0\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{8}$$

לכן ההסתברות המותנית שהילד הצעיר ג'ינג'י בהינתן שהבכור שחור שיער הינה $\frac{\frac{1}{8}}{\frac{3}{4}} = \frac{1}{6}$.

כפי שחזינו בסעיף הקודם, $Pr\left(G_2\right) \neq Pr\left(G_2|\overline{G}_1\right)$ והמידע אודות האח הבכור שינה את ההסתברות שהאח הצעיר יהיה ג'ינג'י – המאורעות תלויים.

$\square$

תרגיל 2.33. נסמן ב-A את המאורע שרכיב מספר 1 פועל וב-B את המאורע שהמערכת פועלת.

1. כדי שהמערכת תפעל צריך שלפחות רכיב אחד יפעל ולכן יש הרבה אפשרויות (רכיב אחד בדיוק פועל, שני רכיבים פועלים וכן הלאה). נחשב במקום זאת את המשלים: המערכת כבויה כאשר כל הרכיבים לא פועלים, וההסתברות לכך היא $Pr\left(\overline{B}\right) = \left(\frac{1}{2}\right)^n$, כלומר $Pr\left(B\right) = 1-\left(\frac{1}{2}\right)^n$. כאשר מספר הרכיבים שואף לאינסוף הסתברות זו שואפת ל-1 כי יש הרבה מאוד רכיבים ובסיכוי גבוה לפחות אחד מהם יפעל וידליק את המערכת כולה.
2. באופן טריויאלי הסתברות זו היא $Pr\left(B|A\right) = 1$.
3. נשתמש בנוסחת בייס: $Pr\left(A|B\right) = \frac{Pr(B|A)\cdot Pr(A)}{Pr(B)} = \frac{\frac{1}{2}}{1-\left(\frac{1}{2}\right)^n}$. בגבול, כאשר מספר הרכיבים שואף לאינסוף נקבל $Pr\left(A|B\right) = \frac{1}{2}$.
4. בבירור, פעולת המערכת ופעולת הרכיב תלויות שכן אם הרכיב פועל אז גם המערכת פועלת ויש ביניהם קשר סיבתי. לעומת זאת, בגבול בו יש אינסוף רכיבים, מתקיים $Pr\left(A|B\right) = Pr\left(A\right)$ כלומר הרכיב והמערכת בלתי תלויים. הסיבה לכך היא שדי ברכיב פועל אחד כדי שהמערכת תפעל וכאשר יש הרבה מאוד רכיבים, כמעט תמיד יהיה רכיב חוץ מהרכיב הראשון שיפעל וידליק את המערכת. לכן, העובדה שרכיב מספר אחד פועל לא מוסיפה לנו מידע לגבי ההסתברות לפעולת המערכת, שהיא 1. ניתן גם לבצע את השיקול ההפוך - אם אנחנו יודעים שהמערכת פועלת, בסיכוי מאוד גבוה היא פועלת בגלל אחד הרכיבים האחרים שכן ההסתברות שדווקא רכיב 1 הוא זה שעושה את ההבדל בין מערכת כבויה לנמוכה הוא מאוד נמוך. לכן, בהינתן שהמערכת פועלת, רכיב 1, שהוא כמעט חסר השפעה, יפעל או יהיה כבוי לפי ההסתברות הלא מותנית.

□

תרגיל 2.34. נסמן ב-R_i את המאורע שהחבר ה-$i = 1, 2, 3$ אמר שיורד גשם וב-R את המאורע שיורד גשם.

1. לפי הנתונים ההסתברות שכל חבר יגיד שיורד גשם כאשר באמת יורד גשם היא $Pr(R_i|R) = \frac{2}{3}$, אבל בסעיף זה אנחנו נדרשים לחשב את ההסתברות שיורד גשם בהינתן שהחברים אומרים שיורד גשם. נשתמש בנוסחת בייס בכדי להפוך את כיוון ההתנייה:

$$Pr(R|R_1 \cap R_2 \cap R_3) = \frac{Pr(R_1 \cap R_2 \cap R_3|R)\, Pr(R)}{Pr(R_1 \cap R_2 \cap R_3)}$$
$$= \frac{\left(\frac{2}{3}\right)^3 \cdot \frac{1}{2}}{\left(\frac{2}{3}\right)^3 \cdot \frac{1}{2} + \left(\frac{1}{3}\right)^3 \cdot \frac{1}{2}} = \frac{8}{9}$$

כאשר המכנה חושב לפי נוסחת ההסתברות השלמה, כשבכל פעם מתנים על מזג האוויר בעיר.

2. ההסתברות שאכן יורד גשם תקטן. ביחס לסעיף הקודם, ההסתברות שיורד גשם בעיר קטנה יותר ולכן ההסתברות שהחברים משקרים ואומרים שיורד גשם ביום שמשי תגדל ביחס להסתברות שהחברים דוברים אמת ואומרים שיורד גשם ביום גשום. כדי להבין זאת באופן אינטואיטיבי, אפשר לחשוב על מצב בו הסיכוי לגשם הוא אפסי. במקרה כזה, כמעט תמיד כאשר החברים יגידו שיורד גשם הם בעצם משקרים והסיכוי שיורד גשם גם כשהם אומרים שיורד גשם יהיה עדיין קטן. חישוב ישיר בדומה לסעיף א' יניב את התוצאה $\frac{8}{17} < \frac{8}{9}$.

3. המאורעות תלויים. היות וחבר א' אומר אמת בהסתברות גבוהה יותר, כאשר הוא אומר אמת סימן שהסיכוי לגשם גדול יותר מהפריור. לכן, בסיכוי גבוה יותר אכן יורד גשם ולכן גם החבר השני יגיד שיורד גשם. נחשב זאת מפורשות. לפי נוסחת ההסתברות השלמה:

$$Pr(A) = Pr(B) = \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{2}$$

וכן

$$Pr(A \cap B) = \left(\frac{2}{3}\right)^2 \cdot \frac{1}{2} + \left(\frac{1}{3}\right)^2 \cdot \frac{1}{2} = \frac{5}{18}$$

הסתברות החיתוך שונה ממכפלת הסתברויות המאורעות והמאורעות תלויים.
ניתן עתה לחשב מפורשות את ההסתברות שחבר ב' אומר שיורד גשם בהינתן שזה מה שחבר א' אמר ולקבל:

$$Pr(B|A) = \frac{Pr(A \cap B)}{Pr(A)} = \frac{\frac{5}{18}}{\frac{1}{2}} = \frac{5}{9} > \frac{1}{2}$$

כאמור בהסבר האינטואיטיבי, אם חבר א' אמר שיורד גשם, חבר ב' יגיד שיורד גשם בהסתברות גבוהה יותר.

□

תרגיל 2.35. **1.** נסמן ב-C_i את המאורע שנבחר המעגל ה-i וב-L_j את המאורע שהנורה דולקת בבדיקה ה-j. לפי נוסחת ההסתברות השלמה:

$$Pr\left(L_1\right) = \sum_{i=1}^{3} Pr\left(L_1|C_i\right) Pr\left(C_i\right) = 0.3 \cdot \frac{1}{3} + 0.5 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0.6$$

כאשר מבצעים שתי בדיקות, הנורה דולקת בכל אחת מהבדיקות באופן בלתי תלוי ולכן ההסתברות שתדלוק בשתיהן היא ריבוע ההסתברות לדלוק פעם אחת:

$$Pr\left(L_1 \cap L_2\right) = \sum_{i=1}^{3} Pr\left(L_1 \cap L_2|C_i\right) Pr\left(C_i\right) = 0.3^2 \cdot \frac{1}{3} + 0.5^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{1}{3} = 0.447$$

ולכן

$$Pr\left(L_2|L_1\right) = \frac{Pr\left(L_1 \cap L_2\right)}{Pr\left(L_1\right)} = \frac{0.447}{0.6} = 0.744$$

ההסתברות גדלה. הסיבה לכך היא שאם הנורה דולקת בבדיקה הראשונה, כנראה שבחרנו מעגל שבו הנורה דולקת בהסתברות גבוהה יותר ולכן גם בבדיקה השניה היא תדלוק.

2. המאורעות בלתי תלויים, שכן אנחנו בוחרים את המעגל החשמלי בכל פעם מחדש. לכן בכל בדיקה הסיכוי שהנורה דולקת הוא 0.6 והמידע הנוסף על כך שבבדיקה השניה הנורה דולקת לא משפיע על ההסתברות שהנורה דולקת בבדיקה הראשונה.

□

3 קומבינטוריקה

בפרק הקודם חישבנו הסתברויות של מאורעות מורכבים מתוך הסתברויות של מאורעות פשוטים יותר. למשל, באמצעות נוסחת ההכלה וההפרדה יכולנו לחשב הסתברות של איחוד מאורעות שהסתברותם נתונה. הסייג היחיד לכך היה עקרון הסימטריה (אקסיומה 2.4) באמצעותו חישבנו את ההסתברות של מאורעות סימטריים, כמו תוצאות ההטלה של קוביה, הטלה של מטבע או שליפה של פתק מכובע (ראו תרגילים 2.12 ו-2.14). בפרק זה נרחיב את השימוש בעקרון הסימטריה על מנת לחשב הסתברויות של מאורעות במרחבי מדגם מורכבים יותר, הכוללים מספר רב של תוצאות ואפשרויות. הכלל החשוב היחיד שמרחב המדגם צריך לקיים הוא היותו סימטרי, כלומר שלכל התוצאות יש בדיוק את אותה הסתברות להתקבל. מרחבי מדגם סימטריים הם למעשה מרחבי ההסתברות היחידים בהם ניתן לחשב את ההסתברויות של מאורעות בצורה תאורטית מבלי להתבסס על תוצאות של ניסויים או הסתברויות אחרות נתונות.

הגדרה 3.1. מרחב מדגם סימטרי הוא מרחב מדגם (סופי) בו כל התוצאות סימטריות כך שההסתברות לקבל כל אחת מהן שווה. במרחב מדגם סימטרי ההסתברות של כל מאורע $A\subseteq\Omega$ הינה:

$$Pr\left(A\right)=\frac{|A|}{|\Omega|} \tag{3.1}$$

כאשר $|A|$ הינו מספר התוצאות הנמצאות במאורע A ו-$|\Omega|$ הוא מספר התוצאות הנמצאות במרחב המדגם כולו.

תרגיל 3.1. האם מרחבי המדגם הבאים סימטריים? אם לא – איזה מרחב מדגם סימטרי תציעו במקום?

1. מטילים זוג קוביות. מרחב המדגם הינו הסכום: $\Omega=\{2,...,12\}$.
2. מטילים קוביה בה הפאות 1-4 צבועות בלבן והפאות 5-6 בשחור. מרחב המדגם הינו $\Omega=\{\text{white,black}\}$.
3. מטילים מטבע פעמיים. מרחב המדגם הוא מספר העצים שהתקבלו: $\Omega=\{0,1,2\}$.

אחד הכלים החשובים בבניית מרחב מדגם סימטרי הוא להתייחס אל כל הפריטים כאל שונים, גם אם השוני ביניהם לא מעניין אותנו או שעקרונית לא ניתן להבדיל ביניהם. במקרה כזה, נתייחס אל הפריטים כאל שונים וכשנספור את המצבים נזכור שהשוני בין הפריטים לא מעניין. שתי סיבות מובילות אותנו להתייחס תמיד לפריטים כאל שונים. ראשית, תמיד אפשר לפני תחילת הניסוי לצבוע את הפריטים בצבעים שונים ואז לבצע את הניסוי. הצביעה מאפשרת לנו להבדיל בין הפריטים אבל לא צריכה להשפיע על ההסתברות. לכן, תמיד ניתן להתייחס אל הפריטים כאל שונים. שנית, מרחב מדגם המורכב מפריטים זהים שלא ניתן להבדיל ביניהם אינו מרחב הסתברות סימטרי. נדגים זאת עם בעיה פשוטה: מחלקים 2 כדורים זהים באקראי ל-2 תאים, מה ההסתברות שכל התאים מלאים? כאשר אנחנו מסתכלים על הפריטים כזהים, יכולות להיות שלוש אפשרויות: 2 כדורים בתא הראשון, 2 כדורים בתא השני וכדור אחד בכל תא. לו מרחב המדגם הזה היה סימטרי, ההסתברות שכל התאים מלאים הייתה $\frac{1}{3}$. נחזור על הניסוי עם כדור לבן וכדור שחור. אנחנו עדיין מחלקים 2 כדורים ל-2 תאים והניסוי זהה לחלוטין, לכן ההסתברות לא צריכה להשתנות. יחד עם זאת, הפעם יש 4 תוצאות אפשריות בניסוי: 2 כדורים בתא הראשון, 2 כדורים בתא השני, כדור לבן בתא הראשון וכדור שחור בתא השני וההפך – כדור לבן בתא

השני וכדור שחור בתא הראשון. לכן, המצב בו שני התאים מלאים יכול להתקבל בשתי דרכים שונות וההסתברות (האמיתית) לכך שכל התאים יהיו מלאים הינה $\frac{2}{4}$.

הכלי המרכזי לחישובי הסתברויות במרחבי מדגם סימטריים הוא מדע הקומבינטוריקה, העוסק בספירת אפשרויות. באמצעותו נחשב את מספר האפשרויות שיש בניסוי כולו ללא הגבלה (זהו מרחב המדגם), את מספר האפשרויות המתאימות למאורע ונחלק אותם לקבלת ההסתברות. הרעיון המרכזי בקומבינטוריקה הוא להמציא ניסוי המורכב משלבים פשוטים, במסגרתו נבנות כל האפשרויות הכלולות במרחב המדגם או המאורע. השאיפה תהיה שכל אחד מהשלבים יהיה פשוט מאוד עם מספר אפשרויות קבוע וכך ניתן יהיה לחשב את מספר האפשרויות הכולל בניסוי כולו על ידי עקרון הכפל:

***עובדה* 3.1** (עקרון הכפל). *יהי ניסוי המורכב מ-n שלבים כך שמספר האפשרויות בשלב הראשון הינו m_1, מספר האפשרויות בשלב השני הינו m_2... ומספר האפשרויות בשלב ה-n הינו m_n. בהנחה שמספרים אלו קבועים ולא תלויים באפשרויות שנבחרו בפועל, מספר האפשרויות הכולל לביצוע הניסוי הינו $m_1 \cdot \ldots \cdot m_n$.*

לשם המחשה, נספור כמה מספרים ארבע ספרתיים קיימים. נבצע זאת על ידי ניסוי בן 4 שלבים: בחירת ספרת האחדות, בחירת ספרת העשרות, בחירת ספרת המאות ובחירת ספרת האלפים. בכל אחד משלושת השלבים הראשונים יש 10 אפשרויות (הספרות 0-9) ובבחירת ספרת האלפים יש 9 אפשרויות (הספרות 1-9[1]). לפי עקרון הכפל, מספר האפשרויות הכולל בניסוי, קרי מספר המספרים הארבע ספרתיים הקיימים, הינו $9 \cdot 10 \cdot 10 \cdot 10 = 9000$. אמנם התשובה המספרית יחידה, אבל הדרך להגיע אליה אינה יחידה וניתן להמציא הרבה מאוד תהליכים שיבנו את כל המספרים הארבע ספרתיים הקיימים. למשל, ניתן לבחור קודם את ספרת האלפים ואחרי זה את ספרות אחרות וכדומה. למעשה, ברוב התרגילים בקומבינטוריקה יהיו מספר דרכים לתכנן את הניסוי ולהגיע אל התשובה הסופית.

בנוסף לעקרון הכפל, העוסק במספר האפשרויות בניסוי המורכב ממספר שלבים נשתמש בעקרון נוסף, עקרון החיבור, העוסק במספר האפשרויות בניסוי שלם אותו ניתן לבצע במספר דרכים שונות.

***עובדה* 3.2** (עקרון החיבור). *יהי ניסוי הניתן לביצוע ב-n דרכים אפשריות, כך שמספר האפשרויות לעריכת הניסוי בדרך הראשונה הינו m_1, מספר האפשרויות לעריכת הניסוי בדרך השניה הינו m_2... ומספר האפשרויות לעריכת הניסוי בדרך ה-n הינו m_n. אם בכל אחת מהדרכים מקבלים תוצאות שונות, מספר האפשרויות הכולל לניסוי כולו הינו $m_1 + \ldots + m_n$.*

לשם המחשה, נספור בכמה דרכים ניתן לבחור זוג אנשים מתוך קבוצה הכוללת 3 גברים ו-2 נשים, כך שבזוג לא יהיו שני גברים. בכדי שיתקיים המבוקש, הזוג יכול לכלול גבר ואישה או שתי נשים. נספור את מספר הזוגות הקיימים בכל אחת מאפשרויות אלו ובהתאם לעקרון החיבור, נסכום אותן. היות ויש 2 נשים, יש רק זוג אחד המורכב רק מנשים וזה הזוג המתקבל כאשר שתיהן נבחרות. נספור את מספר הזוגות המעורבים. כדי ליצור זוג מעורב ניתן לחשוב על הניסוי הדו-שלבי הבא: בשלב הראשון בוחרים גבר (יש 3 אפשרויות) ובשלב השני בוחרים אישה (יש 2 אפשרויות). לפי עקרון הכפל יש $3 \cdot 2 = 6$ אפשרויות לבחירת זוג מעורב ובסך הכל $6 + 1 = 7$ זוגות אפשריים שאינם כוללים רק גברים.

נדגיש את ההבדל בין שני העקרונות. בעקרון הכפל משתמשים בשביל לחשב את מספר האפשרויות בניסוי כולו, כאשר המכפלה נעשית בין מספרי האפשרויות בכל אחד מהשלבים השונים של הניסוי. בדוגמא לעיל, השלבים היו בחירת הגבר ובחירת האישה והמכפלה נתנה את מספר הזוגות הכולל. בעקרון החיבור משתמשים כדי לספור את מספר האפשרויות בניסוי, כאשר המחוברים מתארים אפשרויות שונות לביצוע הניסוי **כולו**. בדוגמא לעיל, חישבנו כמה זוגות של נשים יש, חישבנו כמה זוגות מעורבים יש וסכמנו מספרים אלו כדי לדעת כמה זוגות משני הזוגים יש (השוו לאקסיומה 2.3 ולעקרון ההכלה וההפרדה).

[1] אסור לבחור את הספרה 0 שכן אחרת המספר לא יהיה ארבע ספרתי.

3.1 נוסחאות קומבינטוריות בסיסיות

נציג שלוש נוסחאות בסיסיות המתארות את שתי אבני הבניין המרכזיות של הקומבינטוריקה: סידור פריטים בשורה ובחירה של מספר פריטים מתוך קבוצה של פריטים. אמנם קיימות נוסחאות נוספות פרט לנוסחאות אלו, אך ריבוי נוסחאות בהקשר זה יגרום ליותר נזק מתועלת. עיקר העבודה בפתרון בעיות קומבינטוריות תהיה לבנות שיטה למניית כל האפשרויות בניסוי על ידי תיאורו כניסוי בשלבים המתבססים על נוסחאות בסיסיות אלו ועל עקרונות הכפל והחיבור. אין טעם להוסיף ולהזכיר נוסחאות רבות שהרי ממילא הניסוי לרוב לא יתאים בדיוק לאחת מהן. לאור זאת, פרק זה לא יכלול הרבה תאוריה ורוב הפרק יתמקד בפתרון תרגילים והמחשה של שיטות אפשריות ליצירת סידורים מסויימים ופתרון תרגילים.

3.1.1 סידור פריטים בשורה

משפט 3.1. *מספר הסידורים האפשריים של n פריטים בשורה הינו $n!$.*

הוכחה. נבצע את הניסוי ב-n שלבים, כאשר בכל פעם נבחר איזה פריט למקם במקום הבא בשורה. יש n אפשרויות לבחור את הפריט שיעמוד במקום מספר 1. לאחר שפריט זה נבחר, נותרו $n-1$ פריטים ולכן בשלב השני יש $n-1$ אפשרויות לבחור את הפריט שיעמוד במקום השני. באופן דומה, יש $n-2$ אפשרויות לבחירת הפריט שיעמוד במקום השלישי, $n-3$ אפשרויות לבחירת הפריט במקום הרביעי וכן הלאה. לבחירת הפריט שיעמוד במקום הלפני אחרון יש 2 אפשרויות ועבור המקום האחרון יש רק אפשרות אחת – הפריט שטרם הוצב. בסך הכל, לפי עקרון הכפל, ישנן $n\cdot(n-1)\cdot\ldots\cdot 2\cdot 1=n!$ אפשרויות. □

יש לציין כי הסידור שתיארנו כאן אינו הסידור האפשרי היחיד. אמנם טבעי למקם פריטים בשורה החל מהמקום הראשון, אבל נקבל את כל הסידורים האפשריים גם נתחיל בבחירת הפריט שיעמוד במקום ה-5, לאחר מכן הפריט שיעמוד במקום ה-2 וכן הלאה. זה חלק מהחופש הנתון לנו בפתרון בעיות קומבינטוריות, לפרק את הניסוי לשלבים כרצוננו, כל עוד בסופו של דבר נמנה את כל האפשרויות שיש בניסוי.

תרגיל 3.2. בכמה דרכים ניתן לסדר n פריטים במעגל?

3.1.2 בחירה עם סדר ועם החזרה

משפט 3.2. *מספר האפשרויות לבחור k פריטים מתוך n אפשרויות כאשר ניתן לבחור כל פריט מספר לא מוגבל של פעמים ("עם החזרה") וכאשר הסדר בו נבחרו הפריטים משנה ("עם סדר") הינו n^k.*

הוכחה. נבחר את הפריטים לפי הסדר. מספר האפשרויות לבחירת הפריט הראשון הינו n. מאחר והפריט הנבחר יכול להיבחר שוב, מספר האפשרויות לבחירת הפריט השני הינו גם כן n. זהו גם מספר האפשרויות לפריט השלישי, הרביעי וכן הלאה, כך שמספר האפשרויות הכולל לפי עקרון הכפל יהיה $n\cdot n\cdot\ldots\cdot n=n^k$. □

במקרים רבים זהות הבוחר לא תהיה חד משמעית ונתונה אלא אנחנו נצטרך לקבוע כחלק מהפתרון איך יתנהל הניסוי ו"מי בוחר את מי". למשל – בכמה אפשרויות ניתן לחלק 5 כדורים ל-4 תאים? האם במקרה זה התאים הם אלו שבוחרים את הכדורים שנכנסים לתוכם או שמא הכדורים בוחרים את התאים אליהם יכנסו? המפתח להחלטה הנכונה הוא לדרוש שהבוחר יבחר מתוך מספר קבוע של אפשרויות פעם אחת ולא יבחר פעם נוספת. נניח שאנחנו מחליטים שהתאים יבחרו את הכדורים. כמה אפשרויות יש לתא הראשון? הרי אין הגבלה על כמות הכדורים שנמצאים בכל תא, לכן התא הראשון יכול לבחור כל אחד מהכדורים, אבל הוא גם יכול לבחור לא לקבל אף כדור, או לבחור כל מספר אחר של כדורים. כבר בחישוב מספר האפשרויות לתא הראשון אנחנו נתקלים בבעיה, לא כל שכן עבור התא השני (שמספר האפשרויות שלו תלוי גם במספר הכדורים שהתא הראשון בחר)

ובוודאי גם עבור התא השלישי. לעומת זאת, כאשר הכדורים בוחרים את התאים התהליך מסודר יותר ומספר האפשרויות בכל שלב ברור. לכדור הראשון יש בדיוק 4 אפשרויות בחירה. לכדור השני יש גם 4 אפשרויות בחירה, ללא תלות בבחירה של הכדור הראשון. כך גם לכדור השלישי, הרביעי והחמישי ולכן מספר האפשרויות לחלוקת 5 כדורים ל-4 תאים הינו $4^5 = 1024$.

3.1.3 בחירה ללא סדר וללא החזרה

משפט 3.3. *מספר האפשרויות לבחור k פריטים מתוך n אפשרויות כאשר כל פריט נבחר לכל היותר פעם אחת ("ללא החזרה") והסדר בו נבחרו הפריטים איננו משנה ("ללא סדר") הינו $\binom{n}{k}$.*

הוכחה. נדגים כיצד ניתן על ידי בחירה של k פריטים מתוך n ללא סדר וללא החזרה ליצור בשלבים את כל הסידורים בשורה של n פריטים ועל ידי כך נמנה את מספר האפשרויות הדרוש. נסדר n פריטים בשורה בשלושה שלבים. בשלב הראשון נבחר k פריטים מתוך n ללא סדר וללא החזרה. בשלב השני את הפריטים הללו נסדר במקומות 1 עד k שבשורה. בשלב השלישי ניקח את $n-k$ הפריטים שלא נבחרו בשלב הראשון ונסדר אותם במקומות הנותרים. באופן זה אנחנו מקבלים את כל הסידורים של n פריטים בשורה. לכן סך האפשרויות הכולל של הניסוי שתואר כאן הוא $n!$. בשלב השני סידרנו k פריטים בשורה ומספר האפשרויות הוא $k!$. בשלב השלישי סידרנו $n-k$ פריטים בשורה ומספר האפשרויות הוא $(n-k)!$. לכן מספר האפשרויות בשלב הראשון הינו $\frac{n!}{k!(n-k)!} = \binom{n}{k}$. □

תרגיל 3.3. הוכיחו את נוסחת הבינום של ניוטון (נוסחא 0.10) באמצעות קומבינטוריקה.

תרגיל 3.4. בכמה דרכים ניתן לבחור k פריטים מתוך n ללא החזרה ועם סדר?

3.1.4 דוגמאות

נציג בסעיף זה מספר דוגמאות הממחישות את האופן בו מחושבות הסתברויות באמצעות קומבינטוריקה. בסעיף הבא נדון בקשר שבין קומבינטוריקה להסתברות מותנית ונראה כיצד ידיעה שמאורע מסויים קרה או לא קרה משנה את מרחב המדגם ואת אופן החישוב של ההסתברות. לאחר מכן נעבור על מספר רב של דוגמאות הכוללות טכניקות ושיטות שונות לסידורים שיכולים להיות שימושיים כחלק מפתרון שאלות בקומבינטוריקה.

בעיה 3.1 (פרדוקס ימי ההולדת). באוטובוס 30 נוסעים. מה הסיכוי שלפחות שניים מהם חולקים את אותו תאריך יום הולדת (יום וחודש, לא שנה)?

הוכחה. המאורע המבוקש קשה לחישוב מאחר והוא כולל אפשרויות רבות: בדיוק שניים חולקים את אותו יום הולדת, בדיוק 10 חולקים את אותו יום הולדת, ישנה שלישיה החולקת יום הולדת אחד וזוג החולק יום הולדת אחר ועוד. לעומת זאת, המשלים הרבה יותר פשוט וכולל בדיוק אפשרות אחת – אין אף שני אנשים החולקים את אותו יום הולדת. לפיכך נחשב את הסתברות המאורע המשלים ונשתמש במשפט 2.2 לחישוב ההסתברות הרצויה.

בהנחה שתאריכי יום ההולדת של כל האנשים בלתי תלויים (אין תאומים, למשל) ושכל 365 הימים בשנה סימטריים (מתעלמים מה-29.2), מרחב המדגם יהיה בחירה של 30 ימים מתוך 365 עם החזרה (כי אותו יום יכול להיבחר כמה פעמים) ועם סדר (הסדר הוא הזהות השונה של האנשים. המצב בו סופי נולדה ב-14.9 ובן נולד ב-4.3 שונה מהמצב בו סופי נולדה ב-4.3 ובן נולד ב-14.9, למרות שאותם תאריכי הלידה נבחרו). לכן מרחב המדגם הינו 365^{30}. המאורע יהיה בחירה של 30 ימים עם סדר (מאותה סיבה) אך הפעם בלי החזרה, שכן אנחנו דורשים שאותו יום לא ייבחר פעמיים: $\frac{365!}{(365-30)!}$ (תרגיל 3.4). נציב את כל המספרים ונקבל שההסתברות שבקבוצה של

30 אנשים לפחות שניים נולדו באותו יום היא:

$$1 - \frac{\frac{365!}{335!}}{365^{30}} = 0.7$$

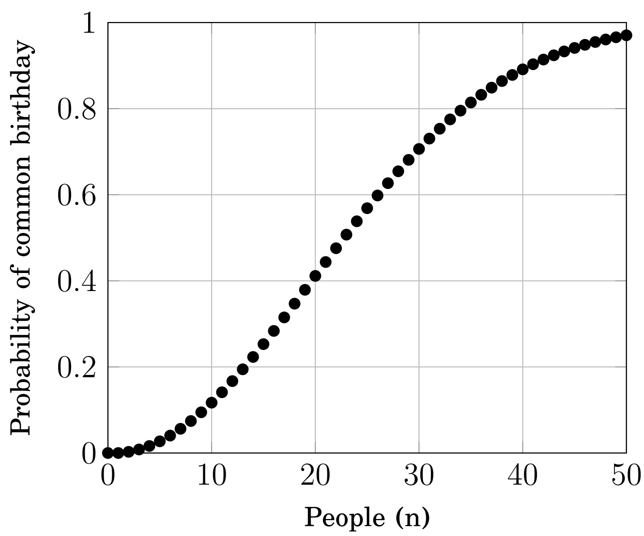
□

אמנם בעיה זו מכונה *"פרדוקס ימי ההולדת"*, אין כאן באמת פרדוקס אלא פשוט תוצאה לא אינטואיטיבית. ראשית, בהרבה מקרים אנשים נוטים לייחס למאורע זה הסתברות נמוכה מאוד בגלל שהם בטעות חושבים על המאורע "אף אחד לא חולק יום הולדת איתי" במקום על המאורע "אף זוג לא חולק יום הולדת", ומניסיונם אכן לרוב גם כשיש 30 אנשים בחדר, אף אחד מהם לא חולק יום הולדת איתם (מה ההסתברות של מאורע זה?). שנית, מרבית האנשים חושבים שההסתברות צריכה להיות הרבה יותר קטנה מאחר ונדמה שכאשר בוחרים רק 30 ימים מתוך 365, יש הרבה מאוד אפשרויות ו"מרווח" לבחור ימים כך שאף יום לא ייבחר פעמיים. אבל, למעשה, הבחירה לא כל כך מרווחת כמו שנדמה. אם ננסה לבחור את ימי ההולדת לאנשים לפי הסדר, הסיכוי שהאדם השני לא יחלוק יום הולדת עם הראשון היא הסתברות מאוד גבוהה, $\frac{364}{365}$. ההסתברות שהאדם השלישי יוולד בתאריך בו לא נולד אף אחד משני הראשונים (בהנחה שהם לא נולדו באותו תאריך) היא $\frac{363}{365}$. ככל שנחלק עוד ועוד ימי הולדת (ועדיין אין התנגשויות), המונה יילך ויקטן וההסתברות שלא נקבל יום הולדת משותף הולכת וקטנה. מאחר והסתברות המאורע מחושבת על ידי הכפלה של הרבה הסתברויות הולכות וקטנות שכולן קטנות מאחד, הרי שמהר מאוד גם המכפלה עצמה תהיה מאוד קטנה.

הנוסחא הכללית, עבור קבוצה של n אנשים ניתנת לחישוב באותה הדרך:

$$1 - \frac{\binom{365}{n} \cdot n!}{365^n} \tag{3.2}$$

מסתבר שביטוי זה גדל מאוד מהר ל-1 כאשר n גדל. כבר ב-$n = 23$ הסיכוי שיהיה זוג שחולק יום הולדת חוצה את ה-50% (כלומר, בערך בחצי מהמקרים ששיתם בחדר עם עוד 22 אנשים, לפחות זוג אחד חלק יום הולדת!) וב-$n = 40$ מגיע כבר ל-90%, כמוראה בתרשים:



מסקנה: בכל כיתה באוניברסיטה ובכל אוטובוס מלא כמעט בוודאות יש לפחות זוג אחד של אנשים החולק תאריך לידה.

בעיה 3.2. "יד" של פוקר כוללת 5 קלפים הנבחרים באקראי מתוך ערימת קלפים תקנית (ראו פרק 0.4.1). חשבו את ההסתברויות של המאורעות הבאים:

1. יותם קיבל 5 קלפים עם ערכים שונים.
2. דריה קיבלה "פול האוס (Full House)", כלומר ביד שלה יש 3 קלפים עם אותו ערך ו-2 קלפים עם ערך אחר.
3. בן קיבל "סטרייט פלאש (Straight Flash)", כלומר יד עם 5 קלפים בערכים עוקבים שכולם מאותה הצורה (אס יכול להיחשב בתור הקלף הבא אחרי מלך או בתור 1).
4. סופי קיבלה "סטרייט (Straight)", כלומר יד עם 5 קלפים בערכים עוקבים אך לא כולם מאותה הצורה (אס יכול להיחשב בתור הקלף הבא אחרי מלך או בתור 1).

הוכחה. נתחיל במרחב המדגם המשותף לכל הסעיפים – יד של פוקר כוללת 5 קלפים הנבחרים מתוך 52 קלפים ללא החזרה וללא סדר, שכן כל מה שמשנה הוא הקלפים שנבחרו ולא באיזה סדר הם נשלפו מהחפיסה והגיעו ליד. לכן מרחב המדגם בכל הסעיפים יהיה $\binom{52}{5}$.

1. על מנת שיותם יקבל 5 קלפים עם ערכים שונים עלינו לבחור את הערכים שהוא יקבל, כלומר לבחור 5 ערכים מתוך 13 הערכים האפשריים ללא סדר וללא החזרה (כי כאמור, אסור לבחור באותו הערך פעמיים). לצורך כך יש $\binom{13}{5}$ אפשרויות. אבל בחירה זו קובעת רק את הערכים שייבחרו ולא את הצורות שיבחרו, שכן כל ערך (למשל, מלך) יכול להתקבל ב-4 דרכים שונות (לב, תלתן, יהלום ועלה). לכן עבור כל אחד מחמשת הקלפים יש לבחור צורה מתוך 4 הצורות האפשריות עם החזרה (צורות יכולות להיבחר כמה פעמים. למעשה חייבים שלפחות צורה אחת תחזור על עצמה כי יש 4 צורות אבל בוחרים 5 קלפים) ועם סדר (הסדר מבטא את הקלפים השונים. מלך לב ומלכה יהלום שונים ממלך יהלום ומלכה לב). לכן יש 4^5 אפשרויות לשלב זה ובסה"כ ההסתברות המבוקשת הינה:

$$\frac{\binom{13}{5}\cdot 4^5}{\binom{52}{5}} = 0.5$$

2. על מנת לקבל "פול האוס" צריך לבחור איזה ערך ייבחר 3 פעמים (13 אפשרויות), אילו 3 צורות ייבחרו עבור ערך זה מתוך 4 הצורות ($\binom{4}{3}$ אפשרויות), איזה ערך ייבחר פעמיים (נותרו רק 12 אפשרויות) ואילו שתי צורות ייבחרו עבור ערך זה מתוך 4 הצורות ($\binom{4}{2}$ אפשרויות):

$$\frac{13\cdot\binom{4}{3}\cdot 12\cdot\binom{4}{2}}{\binom{52}{5}} = 0.00144$$

3. על מנת לקבל סטרייט פלאש יש לבחור את הצורה המשותפת שתהיה לכל הקלפים (4 אפשרויות) ולאחר מכן לבחור את הקלף עם הערך הכי נמוך. ברגע שנקבע הקלף הזה, נקבעו כל שאר הקלפים שבן מחזיק ביד. למשל, אם הקלף הכי נמוך הינו 3, הרי שהרצף חייב לכלול גם את 4,5,6,7. המספר הכי קטן שניתן לבחור לתחילת הרצף הינו אס (ואז הרצף הינו אס עד 4) והמספר הכי גדול שניתן לבחור לתחילת הרצף הינו 10 (ואז הרצף הינו 10 עד אס). לכן יש 10 אפשרויות בחירה למספר הכי קטן ברצף וההסתברות הכוללת הינה

$$\frac{4\cdot 10}{\binom{52}{5}} = 1.54\cdot 10^{-5}$$

4. נבנה את הסידורים המתאימים לסטרייט בדומה לסטרייט פלאש. כאמור, יש 10 אפשרויות לבחור את המספר הקטן ביותר ברצף וברגע שהוא נקבע, נקבעים גם שאר הערכים באופן יחיד. נותר רק לקבוע את הצורות של הקלפים. לכל אחד מחמשת הקלפים יש 4 אפשרויות ולכן מספר האפשרויות הכולל הינו 4^5.

אבל, חישוב זה כלל גם את האפשרות שכל הקלפים בחרו את אותה הצורה ומצב זה אסור (זהו סטרייט פלאש ולא "סתם" סטרייט). לכן נפחית 4 אפשרויות אלו (כולם לב, כולם יהלום, כולם תלתן וכולם עלה) מסך האפשרויות ונקבל:

$$\frac{10\cdot(4^5-4)}{\binom{52}{5}}=0.0039$$

טעות נפוצה להתייחס לדרישה שהקלפים מצורות שונות באופן הבא: יש 4 אפשרויות לצורה של הקלף הראשון, 3 אפשרויות לצורה של הקלף השני (כדי שיהיו שונים) ואז 4 אפשרויות לכל אחד מהקלפים האחרים, כי הדרישה שהקלפים מצורות שונות כבר מולאה. ההסתברות שנקבל בדרך זו תהיה קטנה מההסתברות האמיתית מאחר ובמאורע זה ספרנו רק את המצבים שבהם הצורות של הקלף הראשון והשני שונות. ייתכנו "ידיים" אחרות בהן הקלף הראשון והשני מאותה צורה, אך הקלף השלישי מצורה שונה וכדומה. באופן כללי, הרעיון של "קודם נדאג שהמאורע יתקיים ואז הסידור של הפריטים האחרונים לא מוגבל" הוא לרוב לא נכון ויוביל בדרך כלל להסתברות קטנה מההסתברות האמיתית.

□

במשחק הפוקר "סטרייט פלאש" היא כמעט היד החזקה ביותר האפשרית והיא מנצחת את שתי הידיים האחרות שהצגנו כאן. "פול האוס" היא היד הכי חלשה מבין השלוש. לאור החישובים שעשינו, הסדר בין הידיים ברור – ככל שיד יותר נדירה והסיכוי לקבל אותה יותר קטן, כך היא נחשבת לחזקה יותר. חובבי המשחק מוזמנים להמשיך ולחשב את ההסתברויות של כל הידיים האפשריות כהכנה (חשובה!) לערב הפוקר הבא שלהם.

בעיה 3.3 (בעיית הלוטו הבולגרי). ב-6.9.2009 עלו בהגרלת הלוטו בבולגריה המספרים 4, 15, 23, 24, 35 ו-42. כעבור 4 ימים, בהגרלת הלוטו הבאה שהתקיימה ב-10.9.2009 עלו בגורל בדיוק אותם מספרים. תוצאת ההגרלה השנייה עוררו סערה תקשורתית והאשמות רבות על כך שההגרלה לא הייתה כשרה. זאת ועוד, שר הספורט של בולגריה אף הורה על פתיחת חקירה שלא הניבה שום ראיה לכך שבוצע זיוף וקבעה שמדובר על צירוף מקרים נדיר. האם אכן מדובר על צירוף מקרים נדיר במיוחד?

הוכחה. בטרם נחשב את ההסתברות של המאורע המבוקש ונבחן עד כמה תופעה זו נדירה, נציין שלפי חוקי הלוטו הבולגרי, בכל הגרלה נבחרים באקראי, ללא החזרה וללא סדר, 6 מספרים בין 1 ל-49. בכדי לזכות בפרס הראשון יש לנחש את כולם. מרחב המדגם בהגרלת לוטו בודדת הוא $\binom{49}{6}$ וההסתברות שבהגרלה הראשונה יצאו בדיוק המספרים שצוינו לעיל הינה $\frac{1}{\binom{49}{6}}=7.15\cdot 10^{-8}$. זוהי גם ההסתברות שמספרים אלו יעלו גם בהגרלה השניה ובהנחה שההגרלות בלתי תלויות, ההסתברות שבשתי ההגרלות הללו יצאו המספרים שצוינו לעיל הינה

$$\left(\frac{1}{\binom{49}{6}}\right)^2=5.11\cdot 10^{-15}$$

אכן, למאורע זה הסתברות זעומה ביותר המעוררת חשד. אבל, האמנם זהו המאורע שאת הסתברותו היינו צריכים לחשב? הרי אין משהו מיוחד במספרים אלו וגם אם היו מתקבלים מספרים אחרים בשתי ההגרלות שהתקיימו בספטמבר הדבר היה מעורר חשדות ותמיהה. נבדוק אם כך מה ההסתברות שבשתי ההגרלות הללו יצאו אותם המספרים, אבל לאו דווקא המספרים שבאמת התקבלו בבולגריה. בהגרלה הראשונה יכולה לצאת כל שישייה אפשרית, זה לא ישפיע על המאורע. בהגרלה השנייה, מרחב המדגם הוא שוב $\binom{49}{6}$ וההסתברות לקבל בדיוק את השישייה שיצאה בהגרלה הראשונה היא

$$\frac{1}{\binom{49}{6}}=7.15\cdot 10^{-8}$$

עדיין מאורע נדיר, אבל פחות. האם יש משהו מיוחד בספטמבר 2009? הרי היינו מתפלאים לא פחות אם בדיוק אותם המספרים היו יוצאים בשתי הגרלות רצופות בינואר 2009 או בדצמבר 2009. לכן, נחשב את ההסתברות שמתישהו במהלך שנת 2009 לפחות בשתי הגרלות רצופות של הלוטו הבולגרי התקבלו אותן התוצאות. בהנחה שיש 53 שבועות בשנה ו-2 הגרלות בשבוע, בסה"כ התקיימו בשנה זו 106 הגרלות. מאורע זה מסובך לחישוב שכן הוא כולל הרבה אפשרויות (דוגמת בדיוק ב-7 הגרלות יצאו אותן התוצאות, ב-3 הגרלות יצאו אותן התוצאות וב-5 הגרלות אחרות יצאו תוצאות זהות וכו') ולכן נפנה לחישוב המשלים – ההסתברות שבהגרלות רצופות לא יצאו אותן התוצאות. בהגרלה הראשונה של השנה יכלה לצאת כל תוצאה, כי אין הגרלה לפניה. הסיכוי שבהגרלה השניה יצאה תוצאה שונה, בהמשך לחישוב הקודם, הינה $1-\frac{1}{\binom{49}{6}}$. ההסתברות שבהגרלה השלישית יצאה תוצאה שונה מההגרלה השניה היא גם כן $1-\frac{1}{\binom{49}{6}}$ ולכן ההסתברות שב-3 ההגרלות הראשונות של השנה באף הגרלה לא יצאה התוצאה של ההגרלה הקודמת היא $\left(1-\frac{1}{\binom{49}{6}}\right)^2$. באותו אופן, ההסתברות שבאף אחד מ-106 ההגרלות של השנה לא יצאה אותה תוצאה כמו בהגרלה שקדמה לה היא

$$\left(1-\frac{1}{\binom{49}{6}}\right)^{105}=1-7.51\cdot10^{-6}$$

ולכן ההסתברות שלפחות פעם אחת כן יצאה אותה תוצאה כמו בהגרלה הקודמת היא המשלים, $7.51\cdot10^{-6}$. אם נמשיך באותו הכיוון, 2009 היא לא שנה מיוחדת והיינו מתפלאים לא פחות גם אם הדבר היה קורה בשנת 1980 או בשנת 2011. לכן, אם נחזור על אותו חישוב, ההסתברות שלאורך 50 השנים האחרונות (בהן התקיימו $50\cdot106=5300$ הגרלות) לפחות בהגרלה אחת של הלוטו הבולגרי יצאה אותה תוצאה כמו בהגרלה שקדמה לה היא

$$1-\left(1-\frac{1}{\binom{49}{6}}\right)^{5299}=0.000379$$

הגורם האחרון בו נשאר לטפל הוא במדינה. הרי היינו מתפלאים לא פחות אם זה היה קורה בבלגיה ולא בבולגריה. נניח שב-100 מדינות ברחבי העולם נערכות הגרלות בחמישים השנים האחרונות וההגרלות כולן לפי אותם חוקים של הלוטו הבולגרי. ההסתברות שלפחות פעם אחת מתישהו, במהלך חמישים השנים האחרונות, במדינה כלשהי בעולם התקבלו בשתי הגרלות עוקבות אותן התוצאות היא

$$1-\left(\left(1-\frac{1}{\binom{49}{6}}\right)^{5299}\right)^{100}=0.037$$

הסתברות זו עדיין נמוכה, אבל נמצאת בתחום המספרים הסבירים וה"לא מפתיעים". אף אחד לא מרים גבה כשמאורע שההסתברות שלו היא 3.7% קורה. לשם השוואה, הסיכוי ששני המספרים 23 ו-24 יצאו בהגרלת לוטו מסויימת קטן יותר ושווה ל- $\frac{\binom{47}{4}}{\binom{49}{6}}=0.013$ (במונה – חייבים לבחור את המספרים 23 ו-24 ולכן נותר רק לבחור את 4 המספרים הנוספים שיצאו מתוך 47 המספרים האחרים). אם ניקח בחשבון שבעולם מתקיימות יותר הגרלות לוטו ממה שהנחנו ושבחלק מההגרלות מספר האפשרויות קטן יותר, הרי שזה כלל לא מפתיע שמתישהו, איפשהו בעולם, יצאה אותה תוצאה פעמיים ברציפות. רצה הגורל וזה קרה בבולגריה בספטמבר 2009, אבל

כאמור, זה יכל היה לקרות בכל מקום אחר בעולם בכל זמן אחר (ולמעשה, קרה פעמים נוספות, למשל בהגרלה מקומית שנערכה בצפון קרוליינה ב-2007). □

פרט לאופן בו מחושבת ההסתברות, הדוגמא האחרונה ממחישה כמה תופעות מעניינות בתורת ההסתברות ובאופן בו אנשים תופסים הסתברויות של מאורעות. אנשים נוטים לייחס משמעות רבה מדי לפרטים מסויימים של אירועים ולחשוב שמדובר על אירועים נדירים ומפתיעים בגינם, אך גם אם הפרטים היו אחרים (בדוגמת הלוטו – המספרים, התאריכים או המדינה) הם היו מתפלאים באותה מידה. החופש הנוסף מכך שניתן לבחור את הפרטים הללו מגדיל משמעותית את ההסתברות לכך שהמאורע ייתרחש. בנוסף, אנשים נוטים לחשוב על דברים אקראיים בתור דברים מפוזרים מאוד ולא מסודרים. למשל, המספרים שעלו בגורל בלוטו הבולגרי נראים כמו מספרים "אקראיים", אבל אם המספרים שהיו עולים בגורל היו 1, 2, 3, 4, 5 ו-6, זה היה מאוד חשוד, למרות שלשתי השישיות הללו יש בדיוק אותו סיכוי להתקבל. באותה מידה, אם תבקשו ממישהו לדמיין רצף של 100 הטלות מטבע אקראיות ולרשום אותן על דף, תגלו שברשימת ההטלות כמעט ולא מופיעים רצפים ארוכים של אותה התוצאה (כי הם נראים "מסודרים מדי" ולא אקראיים) למרות שהסיכוי לרצפים כאלו הוא לא נמוך ולמעשה שווה לכל רצף אחר אפשרי באותו אורך.

מסקנה נוספת וחשובה היא שכל דבר שיש לו הסתברות חיובית בסופו של דבר ייתרחש. בבעיית הלוטו הבולגרי, הסיכוי לחזור על אותה תוצאה היה מאוד נמוך, אבל כאשר לקחנו בחשבון הרבה מאוד הגרלות (50 שנות לוטו ב-100 מדינות!) הסיכוי עלה משמעותית ואם היינו מוסיפים עוד ניסויים, ההסתברות הייתה הולכת וגדלה (עבור 500 שנות לוטו ב-200 מדינות, ההסתברות כבר עוברת את ה-50%!). בתורת ההסתברות משפט זה נקרא הלמה של בורל-קנטלי והגירסא הבאה היא גירסא פשוטה ושימושית לצרכינו של הלמה:

***משפט* 3.4** (משפט הקוף המקליד). *יהי A מאורע עם הסתברות $Pr(A) > 0$ בניסוי כלשהו. אם חוזרים על הניסוי מספר רב של פעמים והחזרות בלתי תלויות, הסיכוי שבאחת החזרות המאורע A יתרחש שואפת ל-1 כאשר מספר החזרות שואף לאינסוף. בניסוח מרושל: כאשר חוזרים על ניסוי אינסוף פעמים, כל מה שיכול להתרחש – ייתרחש מתישהו.*

הוכחה. ההסתברות שהמאורע A לא ייתרחש ב-n החזרות הראשונות הינה $(1 - Pr(A))^n$. מתקיים $Pr(A) > 0$ כך שבסיס החזקה קטן מ-1 ולכן $\lim_{n\to\infty}(1 - Pr(A))^n = 0$, כלומר ההסתברות שהמאורע לעולם לא יתרחש היא 0. לכן כאשר חוזרים על הניסוי מספר רב של חזרות, המאורע A מתרחש בהסתברות 1 באחת החזרות. □

שמו של המשפט לקוח מתוך מאמרו המקורי של בורל ב-1913 שם נטען שאם יתנו לקוף להקליד באקראי על מכונת כתיבה אינסוף זמן, בהסתברות אחת מתישהו יקליד הקוף את כל כתבי ויליאם שייקספיר (ולמעשה, כל טקסט אפשרי אחר).

תרגיל 3.5. חוקי הלוטו הישראלי עברו מספר שינויים לאורך השנים. עד שנת 2000, היה צריך לנחש 6 מספרים מתוך 49. בין השנים 2004-2009 היה צריך לנחש 6 מספרים מתוך 34 וכן לנחש "מספר חזק" אחד מתוך 10. החל משנת 2011 יש לנחש 6 מספרים מתוך 37 וכן "מספר חזק" אחד מתוך 7. באיזו משלוש הגירסאות הללו הסיכוי לזכות בלוטו הוא הכי גדול?

תרגיל 3.6. בהגרלת משחק ה"צ'אנס" נבחרים 4 קלפים מתוך חפיסה המכילה 32 קלפים (4 סדרות של קלפים ללא הקלפים 2-6), כאשר מכל סדרה נבחר קלף אחד בדיוק. מהי ההסתברות לנחש את ארבעת הקלפים?

תרגיל 3.7. משחק הדומינו מורכב מ"אבנים" המחולקות לשניים. על כל חצי אבן רשום מספר בין 0 ל-6. ייתכנו אבנים שאותו מספר רשום על שני החלקים, אך אין שתי אבנים זהות. כמה אבני דומינו קיימות בסך הכל?

תרגיל 3.8. בכיתה 30 תלמידים. 10 מתוכם טובים בכדורסל והיתר טובים בכדורגל.

1. בוחרים באקראי חמישה תלמידים לנבחרת הכדורסל הכיתתית. מהו מספר הנבחרות המכילות לפחות תלמיד אחד שטוב בכדורסל?
2. בוחרים באקראי 11 תלמידים לנבחרת הכדורגל הכיתתית (שוער ועשרה שחקנים). מהו מספר הנבחרות המכילות בדיוק 4 תלמידים שטובים בכדורגל, אך לא בתפקיד השוער?

תרגיל 3.9. שבי משחק בקזינו במשחק הבא: הוא מטיל 4 קוביות ומנצח רק אם לפחות באחת ההטלות התקבל 6. בעל הקזינו מציע לו את המשחק הבא: להטיל בכל סיבוב זוג קוביות במקום קוביה אחת כאשר ניצחון יתקבל רק אם לפחות פעם אחת ייצא שש-שש. שבי טוען שהסיכוי לקבל שש-שש קטן פי שש מהסיכוי לקבל שש בודד, ולכן כפיצוי עליו להטיל את זוג הקוביות פי שש פעמים, כלומר עליו להטיל 24 פעמים זוג קוביות. חשבו את הסיכוי לנצחון בכל אחד מהמשחקים וקבעו האם שבי צודק.

תרגיל 3.10. מחלקים באקראי 5 כדורים זהים ל-4 תאים. מה הסיכוי שכל התאים מלאים?

3.1.5 קומבינטוריקה והתנייה

בפרק 2 ראינו שידיעה על מאורעות שקרו בניסוי יכולה להשפיע על ההסתברות שמאורעות אחרים יקרו בניסוי ולמדנו כיצד לחשב הסתברויות מותנות אלו. בהקשר הקומבינטורי, הידיעה על כך שמאורע מסויים קרה בניסוי יכולה לשנות את מרחב המדגם של הניסוי למרחב מדגם הלוקח בחשבון רק תוצאות המתאימות למאורע שקרה. באופן זה, ניתן לחשב את ההסתברות המותנית ישירות מבלי לנסות לחשב את שתי ההסתברויות של שני המאורעות המופיעים בנוסחא 2.9.

בעיה 3.4. מטילים קוביה 5 פעמים. בהינתן ש-6 יצא לראשונה בהטלה החמישית, מה הסיכוי שבשאר ההטלות יצאו רק מספרים זוגיים?

הוכחה. אמנם בניסוי מטילים קוביה 5 פעמים ולכן טבעי להסתכל על מרחב המדגם הכולל 5 הטלות קוביה (6^5), אבל אנחנו יודעים משהו על התוצאות שהתקבלו בקוביה ולכן ניתן לבנות את מרחב המדגם כך שיתחשב במידע זה. לאור זאת, נבנה את מרחב המדגם בהתחשב בעובדה ש-6 יצא לראשונה בהטלה החמישית. אנחנו מסתכלים על 4 הטלות ובכל הטלה יכל לצאת כל אחד מהמספרים 1-5 לכן מרחב המדגם הוא בחירה של 4 מספרים מבין 5 האפשרויות הללו, עם החזרה ועם סדר: 5^4. בכדי שהמאורע יקרה, בכל אחת מ-4 ההטלות צריך להיבחר אחד משני המספרים הזוגיים האפשריים (כי 6 לא יצא בהטלות אלו) ולכן המאורע הוא בחירה של 4 מספרים מתוך 2 האפשרויות, עם החזרה ועם סדר: 2^4. ההסתברות היא לכן: $\frac{2^4}{5^4}$. □

תרגיל 3.11. מטילים קוביה 5 פעמים. בהינתן ש-6 יצא לראשונה בהטלה החמישית, מה הסיכוי שבחמש ההטלות יצאו כל המספרים הזוגיים ורק הם?

תרגיל 3.12. מטילים זוג קוביות עד אשר לראשונה סכומן מתחלק ב-5. מה ההסתברות שבהטלה בה הסתיים הניסוי סכום הקוביות מראה על 5?

שימוש נוסף בהתנייה בהקשר הקומבינטורי הוא בנוסחת ההסתברות השלמה (פרק 2.2.1). כאשר נשתמש בנוסחת ההסתברות השלמה נשתדל להתנות על מאורעות אשר בונים מרחב מדגם, כך שאם אנחנו יודעים שהם התרחשו אנחנו יודעים איך נראה מרחב המדגם, מהו גודלו ואיך לייצר את המאורע במרחב מדגם זה.

בעיה 3.5. בכד יש 5 כדורים לבנים. מטילים קוביה הוגנת ומכניסים לכד כדורים שחורים בהתאם למספר הרשום על הקוביה. לאחר מכן, מוציאים כדור אחד. מה הסיכוי שהכדור שנשלף לבן?

הוכחה. נגדיר את המאורע שנשלף כדור לבן ב-A. לא ניתן לחשב את ההסתברות של A מפורשות שכן מרחב המדגם לא ידוע – אנחנו יודעים שיש 5 כדורים לבנים אך לא יודעים מהו המספר הכולל של הכדורים בכד. לו היינו יודעים את תוצאת הטלה היינו יודעים גם את מרחב המדגם, לכן נשתמש בנוסחת ההסתברות השלמה בכדי להתנות על תוצאת הטלת הקוביה. נגדיר את המאורע שהקוביה נפלה על i ב-B_i. אם המאורע B_i קרה יש בכד $5+i$ כדורים ולכן ההסתברות להוציא כדור לבן היא $\frac{5}{5+i}$. לפי נוסחת ההסתברות השלמה הסיכוי להוציא כדור לבן בניסוי כולו הינו:

$$Pr\left(A\right)=\sum_{i=1}^{6}Pr\left(A|B_i\right)\cdot Pr\left(B_i\right)=\sum_{i=1}^{6}\frac{5}{5+i}\cdot\frac{1}{6}=0.614$$

□

תרגיל 3.13. יש שלוש שידות, בכל אחת מהן שתי מגירות וכולן זהות כלפי חוץ. בשידה הראשונה יש מטבע זהב בכל אחת מצמד המגירות, בשידה השנייה יש מטבע כסף בכל אחת מהמגירות ובשלישית מטבע זהב במגירה אחת ומטבע כסף במגירה השניה. בוחרים שידה באקראי ובוחרים מגירה באקראי ומוצאים מטבע זהב. מה הסיכוי שבמגירה השניה יש גם מטבע זהב?

3.2 שיטות לביצוע חישובים קומבינטוריים

הנוסחאות הבסיסיות לסידור בשורה ובחירה של פריטים מתוך אוכלוסיה מספיקות בשביל לפתור את השאלות בקומבינטוריקה כפי שראינו בשאלות שבסעיף הקודם. המפתח לפתרון שאלה בקומבינטוריקה הוא לתאר את הניסוי כתהליך עם שלבים הכוללים נוסחאות אלו וכך לחשב את המאורע ואת מרחב המדגם. בזאת מסתיימת התאוריה וכל הדרוש להצלחה הינו תרגול רב ככל הניתן ופתרון של שאלות רבות בקומבינטוריקה. בכדי להקל על המשימה, בסעיף זה מוצגות בעיות קומבינטוריות נפוצות (למשל – סידור פריטים כך שמספר פריטים עומדים זה לצד זה, בדיקה איזה פריט נבחר בהוצאה השמינית וכד') והדרכים לפתור אותן. מומלץ לעבור על השאלות ולראות כיצד ניתן לבצע את הסידורים הקומבינטוריים המבוקשים כהכנה לסעיף האחרון בו מבחר של שאלות מסכמות לתרגול.

3.2.1 האנלוגיה בין סידורים בשורה להוצאת פריטים ללא החזרה

כאשר בוחרים פריטים באקראי ללא החזרה מייצרים עבורם סדר ולכן תהליך זה שקול לסידור פריטים בשורה. הפריט הראשון שנבחר הוא הפריט הראשון בשורה, הפריט השני שנבחר הוא הפריט השני בשורה וכך הלאה. לכן, כל סידור פריטים בשורה מתאר תהליך כלשהו בו מוצאים כל הפריטים ללא החזרה ובעיות אלו שקולות. למעשה, גם אם לא מוציאים את כל הפריטים (אלא רק 10 מתוך 20, למשל) הבעיות שקולות, שכן ניתן להחליט שבכל זאת מוציאים את כל הפריטים ומתייחסים לעשרת הראשונים בתור הפריטים ש"נבחרו" ולשאר הפריטים בשורה כאל פריטים ש"לא נבחרו".

בעיה 3.6. מתוך חפיסה אשר אחד מקלפיה אבד שולפים קלף באקראי. מה הסיכוי שקלף זה הינו 2 - לב?

פתרון לפי הוצאה ללא החזרה. התהליך המתואר כאן הוא הוצאה ללא החזרה ועם סדר של שני קלפים, הקלף הראשון הוא הקלף שאבד והקלף השני הוא הקלף שעליו אנחנו מסתכלים. לכן מרחב המדגם הינו $52\cdot 51$. בכדי לחשב את המאורע צריך שהקלף השני יהיה 2 - לב (אפשרות אחת) ושהקלף הראשון יהיה כל אחד מ-51 הקלפים האפשריים האחרים, ולכן ההסתברות המבוקשת הינה $\frac{51\cdot 1}{52\cdot 51}=\frac{1}{52}$. □

פתרון לפי סידורים בשורה. נשתמש באנלוגיה שבין סידורים בשורה להוצאה ללא החזרה. כאשר מסדרים בשורה את הקלפים לכל הקלפים יש את אותו סיכוי להיות במקום השני בשורה ובפרט הסיכוי של 2 - לב להיות במקום השני הינו $\frac{1}{52}$. □

תרגיל 3.14. מתוך כד בו 7 כדורים שחורים ו-3 לבנים מוציאים באקראי 8 כדורים. מה הסיכוי ששני כדורים לבנים הוצאו מהכד בזה אחר זה והכדור השלישי נשאר בתוך הכד?

3.2.2 פריטים צמודים ולא צמודים

כאשר מעוניינים שפריטים מסויימים יסודרו ביחד ויעמדו זה לצד זה בשורה או יוכנסו לאותו התא, ניתן לאחד אותם לפריט אחד ולסדר אותם ביחד. דבר זה נעשה, למשל, בתרגיל 3.10 כאשר רצינו ששני כדורים יחלקו תא ולכן התייחסנו אליהם כאל פריט אחד וסידרנו אותם ביחד.

בעיה 3.7. עשרה זוגות מתיישבים באקראי בשורה אחת באולם קולנוע.

1. מה הסיכוי שאדון וגברת סמית' יושבים זה לצד זו?
2. מה הסיכוי שכל הזוגות יושבים במושבים צמודים?
3. מה הסיכוי שכל הזוגות יושבים במושבים צמודים וכן שכל גבר לא יושב ליד נשים אחרות פרט לאישתו ואף אישה לא יושבת ליד גברים אחרים פרט לבעלה?

הוכחה. בכל אחד מהסעיפים אנחנו מסדרים 20 אנשים בשורה ולכן בכולם מרחב המדגם יהיה $20!$.

1. על מנת שהזוג סמית' ישבו במושבים צמודים, נתייחס אליהם כאל פריט אחד במהלך הסידורים בשורה. לכן אנחנו מסדרים בשורה 19 פריטים: 18 אנשים והזוג סמית' ולכך יש $19!$ אפשרויות. כל אחת מהאפשרויות הללו מפרטת איפה יושבים כל האנשים ביחס לזוג סמית' ואיפה יושב הזוג סמית' כפריט אחד, אך לא מתייחס אל האופן בו ניתן לסדר את בני הזוג בינם לבין עצמם, שהרי בכל סידור כזה מר סמית' יכול לשבת מימין לאשתו או משמאלה. לכן, יש לכפול את האפשרויות הללו במספר הסידורים הפנימיים של הזוג סמית', השווה לסידור בשורה של שני אנשים – $2!$. בסך הכל ההסתברות המבוקשת הינה $\frac{1}{10} = \frac{19!\cdot 2!}{20!}$.
2. נפתור בדרך דומה לסעיף הקודם. הפעם כל אחד מהזוגות צריך לשבת יחד ולכן נתייחס לכל זוג כאל פריט אחד ובסך הכל – 10 פריטים. לאחר שבחרנו את מקומות הישיבה של כל הזוגות, עלינו לקבוע עבור כל זוג מי מבני הזוג יושב במושב הימני ומי במושב השמאלי. יש $2!$ אפשרויות לכל זוג וישנם 10 זוגות לכן יש $(2!)^{10}$ סידורים פנימיים. בסך הכל ההסתברות הינה $\frac{10!\cdot(2!)^{10}}{20!}$.
3. כמקודם, נתייחס אל כל הזוגות כאל פריטים ולכן אנחנו מסדרים 10 פריטים ולכך יש $10!$ אפשרויות. נתבונן בזוג היושב בקצה הימני של השורה. אם הבעל יושב בכיסא הימני (מושב מספר 1) אזי אישתו יושבת במושב מספר 2 וצמודה לאחד מבני הזוג היושבים במקומות 3-4. היות ואסור שתשב ליד גבר שאינו בעלה, במקום מספר 3 חייבת לשבת האישה של הזוג ובמקום מספר 4 גבר. לפי אותו שיקול, במקום מספר 5 צריך לשבת גבר ובמקום מספר 6 אישתו וכן הלאה. לכן, לאחר שקבענו איזה זוג יושב באילו מקומות וקבענו את מין האדם היושב בכיסא הימני ביותר (יש 2 אפשרויות), נקבעים כל הסידורים הפנימיים של כל הזוגות האחרים. ההסתברות הכוללת היא אם כך $\frac{10!\cdot 2}{20!}$.

□

כאשר מעוניינים שפריטים מסויימים לא יהיו צמודים, ניתן להשתמש בטכניקה שתוארה לעיל על מנת לחשב את המשלים, המצב בו הם כן צמודים. דרך נוספת היא לבצע את הניסוי בשלבים. בשלב הראשון מסדרים את הפריטים האחרים שלא מעניינים אותנו. בשלב השני, מוסיפים לשורה את הפריטים שאסור שיהיו צמודים, על ידי זה שמשתמשים בפריטים שסודרו בשלב הראשון כמחיצות ביניהם.

בעיה 3.8. מסדרים באקראי את האותיות של המילה "הסתברותוסטטיסטיקה".

1. כמה "מילים" שונות ניתן ליצור בצורה זו?
2. בכמה מתוך ה"מילים" אף אחת מהאותיות "ט" לא עומדת ליד "ט" אחרת?

הוכחה. נשים לב שבמילה "הסתברותוסטטיסטיקה" יש 17 אותיות: 3 פעמים "ס" ו-"ט", פעמיים "ו", "ת", "ה" ו-"י" ופעם אחת "ב", "ר" ו-"ק".

1. אנחנו מסדרים 17 פריטים בשורה (17 אותיות שונות) ולכן יש $17!$ סידורים שונים, אך לא כל הסידורים השונים מתארים מילים שונות. למשל, אם נחליף את המקום של שתי האותיות "ת" בינן לבין עצמן נקבל סידור שונה של האותיות בשורה למרות שבפועל זו אותה מילה שכן שתי ה"ת" זהות. לכן סידורים פנימיים של האותיות "ת" לא יוצרים מילים חדשות וצריך לחלק בסידורים הפנימיים של ה"ת"-ים (ספרנו כל מילה פי $2!$ פעמים מהרצוי). באותו אופן, גם כל סידור פנימי של שאר האותיות החוזרות על עצמן לא משנים ובסה"כ מספר המילים השונות יהיה

$$\frac{17!}{(2!)^4 \cdot (3!)^2}$$

2. ראשית נסדר את כל האותיות פרט לשלושת אותיות ה-"ט". בהמשך לסעיף הקודם, מספר המילים המתקבלות באופן זה הינו $\frac{14!}{(2!)^4 \cdot 3!}$. כעת נותר למקם את אותיות ה"ט" כך שלא יהיו אחת ליד השניה. נעשה זאת על ידי כך שנמקם את ה"ט"-ים ברווחים שבין אותיות סמוכות (יש 13 רווחים כאלו), או בתחילת המילה או בסוף המילה (בסך הכל – 15 רווחים אפשריים). אנחנו צריכים לבחור 3 רווחים כאלו, כאשר הבחירה נעשית ללא החזרה (אחרת שני "ט"-ים לא יהיו מופרדים על ידי אות אחרת) וללא סדר (כי ה"ט"-ים זהים ולכן לא משנה איזו מהן תמוקם באיזה מהרווחים שיבחרו) ולכן יש $\binom{15}{3}$ אפשרויות. מספר המילים הכולל הוא לפיכך

$$\frac{14!}{(2!)^4 \cdot 3!} \cdot \binom{15}{3}$$

ניתן לנסות לפתור סעיף זה על ידי המשלים אבל הפתרון יהיה מסובך יותר. במקרה זה, המשלים הוא שלפחות שתיים מה"ט"-ים עומדות אחת ליד השניה ויש לכך שתי אפשרויות. אפשרות אחת היא שכל 3 ה"ט"-ים עומדות ברצף ואפשרות שניה היא ש-2 מהאותיות סמוכות זו לזו והאות השלישית לא עומדת לידן. בנוסף, צריך גם להיזהר לא לייצר בטעות שלשה כאשר מחשבים את האפשרות השניה (למשל, על ידי הכלה והפרדה). במקרה שהיו יותר מ-3 אותיות "ט" הפתרון היה הופך למסובך עוד יותר.

□

תרגיל 3.15. ארבעה זוגות מסתדרים באקראי בשורה.

1. מה ההסתברות שאף זוג לא יושב ביחד?
2. דני מנסה לפתור את הסעיף הקודם באופן הבא:

ראשית, הוא מסדר בשורה את כל הגברים. לאחר מכן הוא מוסיף את הנשים אחת אחת ברווחים שבין העומדים בשורה (או בקצוות השורה) כך שאף אישה לא תעמוד ליד בעלה. הסידור הפנימי של הגברים בטרם הוכנסו הנשים לא משנה, כי ממילא ללא הנשים הסידור שלהם לא משפיע על המאורע, לכן נתחיל מסידור כלשהו של הגברים ונצרף לשורה את הנשים אחת אחת. לגברת א' יש 5 אפשרויות (2 מקומות בקצוות השורה ו-3 מקומות בין כל זוג גברים), מתוכם 2 מקומות הם מצידי בעלה, לכן ההסתברות שלא תעמוד לידו היא $\frac{3}{5}$. כעת עומדים בשורה 5 אנשים. לגברת ב' יש 6 אפשרויות (2 מקומות בקצוות השורה ו-4 מקומות בין העומדים בשורה, הכוללת כעת גם את גברת א'). כמקודם, 2 מהמקומות הללו צמודים

לבעלה ולכן הסיכוי שלא תעמוד לידו הוא $\frac{4}{6}$. באותו אופן, הסיכוי שגברת ג' לא תעמוד ליד בעלה הוא $\frac{5}{7}$ והסיכוי שגברת ד' לא תעמוד ליד בעלה הוא $\frac{6}{8}$. ההסתברות המבוקשת היא מכפלת ההסתברויות בכל השלבים של הניסוי, כלומר

$$\frac{3}{5}\cdot\frac{4}{6}\cdot\frac{5}{7}\cdot\frac{6}{8}$$

ענו, מבלי לחשב הסתברות זו האם דני צודק. אם כן, מדוע? אם לא --מהי הטעות והאם ההסתברות שתחושב באופן זה תהיה קטנה או גדולה מההסתברות האמיתית?

3. מה ההסתברות שמר סמית' יושב ליד גברת סמית', בהינתן שהוא יושב בקצה השורה.
4. מה ההסתברות שמר סמית' יושב בקצה השורה, בהינתן שהוא יושב ליד גברת סמית'?
5. מה ההסתברות שמר סמית' יושב ליד גברת סמית' וכן שמר כהן לא יושב ליד גברת כהן?

3.2.3 צמצום מרחב המדגם

כאשר המאורע מתייחס רק לפריטים מסויימים ודן רק בסידורים שלהם או ביחסים ביניהם (למשל, מי מהם נבחר קודם) ניתן לצמצם את מרחב המדגם ולהתייחס רק לפריטים המעניינים אותנו ולא לכל שאר הפריטים. באופן זה, אנחנו מפשטים משמעותית את החישוב שכן אנחנו מורידים הרבה דרגות חופש מהבעיה וצריכים להתעסק בהרבה פחות פריטים.

בעיה 3.9. מתוך חפיסה תקנית שולפים קלפים בזה אחר זה. מה ההסתברות שארבעת המלכות נשלפו מהחפיסה אחרי שהמלך הראשון נשלף ממנה ולפני המלך השלישי?

ניסיון לפתרון ללא צמצום מרחב מדגם. נדגים את הקושי הרב בפתרון השאלה כמו שהיא, תוך התייחסות לכל החפיסה וללא צמצום מרחב המדגם. באמצעות שימוש באנלוגיה שבין שליפה ללא החזרה לבין סידורים בשורה, ניתן להתייחס אל מרחב המדגם כאל סידור בשורה של כל הקלפים ולכן גודלו $52!$. נפנה לחישוב המאורע. בכדי שהמאורע יתרחש צריך לבחור מקום למלך הראשון (45 אפשרויות, כי אחריו יש להשאיר מקום לעוד 7 קלפים: 4 מלכות ו-3 מלכים) ולבחור איזה מהמלכים הוא המלך שיצא (4 אפשרויות). כעת צריך לבחור מי מבין הקלפים הללו ייצא אחריו (המלך השני או המלכה הראשונה), לבחור איזה קלף זה ולבחור עבורו מקום. אבל מספר המקומות האפשריים לבחירה תלוי בבחירה שלנו עבור המלך הראשון, וככל שנמשיך לסדר קלפים כך התלות בכל הבחירות הקודמות תגדל ותסתבך (ועוד לא סידרנו את שאר הקלפים!). לכן, לא ניתן לחשב את ההסתברות של מאורע זו על ידי סידור כל הקלפים בשורה באופן זה ועלינו לצמצם את מרחב המדגם. □

צמצום מרחב המדגם ל-8 הקלפים המעניינים. כמקודם, נשתמש באנלוגיה בין הוצאה ללא החזרה לבין סידורים בשורה, אך הפעם נתייחס רק אל שמונת הקלפים המעניינים אותנו: 4 המלכים ו-4 המלכות. ברגע שהסדר היחסי בין המלכים והמלכות נקבע, אנחנו יודעים האם המאורע הרצוי התרחש או לא ולא משנה מתי נשלפו שאר הקלפים של החפיסה ובאיזה סדר. לכן, ניתן להתעלם מהם ולהסתכל על מרחב מדגם הכולל סידור של 8 הקלפים המעניינים בשורה וגודלו $8!$. בכדי שהמאורע הנדרש יתרחש, הקלף הראשון מבין השמונה שצריך להישלף הוא קלף של מלך (אחרת מלכה נבחרה לפני המלך הראשון) ולכך יש 4 אפשרויות. בנוסף, הקלף העומד במקום האחרון צריך להיות גם קלף של מלך, כי המלך הרביעי נשלף אחרי כל המלכות, ועבורו יש 3 אפשרויות. מאותה סיבה, הקלף במקום השביעי צריך להיות גם קלף של מלך, כי זהו המלך השלישי שנשלף ולכן הוא חייב להיות אחרי כל המלכות. נותרו 2 מלכים בחפיסה ולכן 2 אפשרויות עבור קלף זה. נותר למקם את חמשת הקלפים הנותרים (מלך ו-4 מלכות) במקומות 2-5. לצורך כך, ראשית נבחר מקום עבור המלך (יש 5 מקומות אפשריים) ולאחר מכן ניוותר עם 4 מלכות ו-4 מקומות אפשריים. אין שום הגבלה על הסידור של המלכות במקומות הללו,

ולכן מספר הסידורים האפשרי הינו $4!$. לסיכום, הסתברות המאורע הינה

$$\frac{4\cdot 3\cdot 2\cdot 5\cdot 4!}{8!}=0.071$$

$\square$

צמצום מרחב המדגם לבחירת מקומות למלכים. ניתן לפשט את מרחב המדגם עוד יותר, על ידי זה שמתעלמים מהשוני בין הצורות. בפתרון הקודם, כל שהיה חשוב הוא שמבין 8 הקלפים, הקלף הראשון יהיה קלף של מלך ולא היה אכפת לנו אם זה מלך לב או מלך תלתן. לכן, במקום להסתכל על סידור של 8 קלפים שונים בשורה, ניתן להסתכל על מרחב מדגם שקול הכולל בחירה של 4 המקומות מתוך השמונה בו יוצבו המלכים. לא משנה איזה מלך בדיוק יוצב באיזה מקום, שכן ברגע שנבחרו 4 המקומות מתוך ה-8, נדע אם המאורע התרחש או לא. למשל, אם נבחרו המקומות 1, 2, 4, 6, המאורע לא התרחש, כי יש מלכות במקומות 7,8 שיצאו אחרי כל המלכים. לעומת זאת, אם נבחרו המקומות 1, 5, 7, 8, אז המאורע התרחש, שכן המלך הראשון יצא ראשון מבין הקלפים, אחרי כן יצאו 3 מלכות, ואז המלך השלישי ואז מלכה (כל המלכות יצאו בשלב הזה) ולאחר מכן המלך השלישי והרביעי. לכן מרחב המדגם יהיה בחירה של 4 המקומות בהן יצאו המלכים מתוך 8 המקומות – $\binom{8}{4}$. המאורע כולל את כל הסידורים המתאימים לכך שכל המלכות יצאו בין המלך הראשון לשלישי. מהנימוקים שפורטו בפתרון הקודם ובדוגמא המספרית, חייב לצאת מלך במקום הראשון, השביעי והשמיני, לכן הבחירה היחידה היא מיקום אחד מבין 5 המקומות האפשריים עבור המלך השני. לכן יש 5 אפשריות עבור המאורע וההסתברות המבוקשת היא

$$\frac{5}{\binom{8}{4}}=\frac{5}{\frac{8!}{4!4!}}=\frac{4!\cdot 5\cdot 4!}{8!}=0.071$$

ובאופן לא מפתיע, קיבלנו בדיוק את אותה התוצאה כמו בפתרון הקודם. $\square$

משתי דרכי הפתרון הללו עולה אזהרה חשובה – המאורע חייב להתאים למרחב המדגם. אם, כמו באפשרות הראשונה, מרחב המדגם הוא סידור בשורה של שמונה קלפים שונים אזי גם המאורע צריך לספור את כל הסידורים האפשריים של 8 קלפים שונים בהם המאורע מתקיים. לעומת זאת, בפתרון השני מרחב המדגם הוא בחירה של 4 המקומות בהם עומדים המלכים, מבלי להחליט איזה מלך יעמוד איפה ואיזו מלכה תעמוד איפה ולכן גם המאורע צריך לספור רק בחירה של מקומות המלכים, בלי לסדר את המלכים עצמם במקומות אלו ובלי לסדר את המלכות במקומות האחרים. טעות נפוצה היא לנסות לסדר את המלכים והמלכות במאורע של הפתרון השני, על ידי חישוב מהצורה

$$\frac{5\cdot 4!\cdot 4!}{\binom{8}{4}}$$

כלומר, בחירת המקומות עבור המלכים והמלכות כמו בפתרון הנכון ואז סידור 4 המלכים במקומות המיועדים להם וסידור 4 המלכות במקומות המיועדים להן. במרחב המדגם קבענו שסופרים רק בחירה של 4 מקומות מתוך 8, לכן גם במאורע, המהווה תת קבוצה של מרחב המדגם, מותר לנו לבחור רק 4 מקומות מתוך 8 (המתאימים לדרישה של המאורע). חשוב לפיכך בשלב הראשון של הפתרון לקבוע את מרחב המדגם ולהבין מהם הסידורים שנספרים (האם מסדרים פריטים? האם בוחרים מקומות? האם יש סדר? האם אין סדר? וכו') ולשמור על אחידות בין מרחב המדגם למאורע (פרט לשינויים הנדרשים על ידי המאורע עצמו).

נותר להראות שצמצום מרחב המדגם רק לשמונת הקלפים הללו אכן מוצדק. כאמור, כאשר מסדרים את כל הקלפים בשורה, ניתן ראשית לקבוע את הסדר היחסי של שמונת הקלפים המעניינים אותנו ואחרי כן להציב בשורה את שאר הקלפים, ביניהם, לפניהם או אחריהם ודבר לא ישתנה. לכן, רק הסדר היחסי ביניהם משנה.

טיעון אינטואיטיבי זה ניתן להציג גם בצורה מתמטית, תוך סידור כל הקלפים בשורה אבל בדרך מעט שונה, המדגישה את הסדר היחסי בין המלכים והמלכות:

הצדקה מתמטית לצמצום מרחב המדגם. נשתמש באנלוגיה שבין הוצאת קלפים ללא החזרה לסידור קלפים בשורה ונסדר את כל 52 הקלפים בשורה. במקום לבצע את הסידור בבת אחת ולומר שיש $52!$ אפשרויות, נבצע זאת במספר שלבים: ראשית, נבחר באילו שמונה מקומות נציב את המלכים והמלכות. לצורך כך נבחר 8 מקומות מתוך השורה, ללא סדר וללא החזרה ($\binom{52}{8}$). שנית, נסדר את המלכים והמלכות במקומות שנבחרו. יש 8 מקומות אפשריים ו-8 קלפים לסדר, לכן מספר הסידורים הינו $8!$. לסיום, נסדר את שאר הקלפים שאינם מלכים ומלכות בשאר המקומות. גם כאן, ישנם 44 קלפים ו-44 מקומות ולכן $44!$ אפשרויות. סך כל האפשרויות הינו כצפוי $\binom{52}{8} \cdot 8! \cdot 44! = 52!$.

נחזור על שלבים אלו בבואנו לספור את מספר האפשרויות של המאורע. כמקודם, נבחר בשורה את 8 המקומות של המלכים והמלכות ולאחר מכן נסדר את המלכים והמלכות כך שהמאורע יתקיים. השיקולים בסידור רק המלכים והמלכות זהים לאלו שהיו בפתרון הראשון ולכן מספר האפשרויות הוא שוב, $4 \cdot 3 \cdot 2 \cdot 5 \cdot 4!$. נותר לסדר את שאר הקלפים וכאמור, לצורך כך יש $44!$ מקומות. לכן ההסתברות שכל המלכות יצאו אחרי המלך הראשון ולפני השלישי הינה:

$$\frac{\cancel{\binom{52}{8}} \cdot 4 \cdot 3 \cdot 2 \cdot 5 \cdot 4! \cdot \cancel{44!}}{\cancel{\binom{52}{8}} \cdot 8! \cdot \cancel{44!}}$$

בחירת המקומות ל-8 הקלפים הצטמצמה כמו גם סידור כל שאר הקלפים בשורה וכך נותר הביטוי המתאר רק את סידור המלכים והמלכות בלבד. □

כאשר מצמצמים את מרחב המדגם, חשוב לוודא שאנחנו אכן מתייחסים לכל הפריטים הרלוונטיים למאורע. למשל, כאשר שאלה זו הופיעה במבחן, טעות נפוצה של הרבה סטודנטים הייתה לסדר שבעה קלפים: 4 מלכות ו-3 מלכים (המלך הראשון, השני והשלישי) ולדרוש שהמלך הראשון יהיה ראשון מבין השבעה והמלך השלישי יהיה אחרון. אבל, הקלף "המלך הראשון" לא מוגדר היטב, שכן כדי שמלך יהיה "ראשון" חייבים לדעת מה הסדר היחסי של ארבעת המלכים. כאשר מסדרים רק שלושה מלכים ולא נקבע מיקומו של המלך הנוסף, לא בטוח שהמלך שקבענו שהוא ראשון הוא אכן ראשון והמלך שקבענו שהוא שלישי הוא אכן שלישי מבין המלכים. זאת ועוד, צמצום מרחב המדגם במקרה זה אינו מוגדר היטב, שכן לא ברור מיהם אותם המלכים שמסודרים - האם שלושת המלכים הם לב יהלום ותלתן או לב עלה ויהלום או כל אחת מאפשרויות האחרות? לעומת זאת, אם בשאלה היינו נשאלים מה ההסתברות שארבע המלכות יצאו בין מלך לב למלך יהלום, אז מספיק היה לסדר רק את ששת הקלפים הרלוונטיים: ארבע המלכות ושני המלכים הללו.

תרגיל 3.16. בכד 10 כדורים בחמישה צבעים, שניים מכל צבע (ממספרים ב-1 ו-2). מוציאים את כל הכדורים מהכד בזה אחר זה.

1. מה ההסתברות שכדור מספר 1 מכל צבע יצא מהכד לפני כדור מספר 2?
2. האם המאורעות A– "כדור מספר 1 שחור יצא לפני כדור מספר 2 שחור" ו-B – "כדור מספר 1 שחור יצא לפני כדור מספר 1 אדום" תלויים או בלתי תלויים? חשבו ונמקו!
3. מה הסיכוי ששני הכדורים השחורים יצאו מהכד לפני ששני הכדורים האדומים יצאו מהכד?
4. בהנתן שהכדור הרביעי שיצא מהכד היה כדור שחור מספר 2, מה הסיכוי שכדור שחור מספר 1 היה הכדור השני שיצא מהכד?

3.2.4 בחירת פריטים שונים כאשר הבחירה עם החזרה

כאשר הבחירה נעשית עם החזרה אבל יש לבחור פריטים שונים, ניתן לעשות את הבחירה ללא החזרה ואז, במידה והסדר משנה, להכפיל בכמות הסידורים הרלוונטיים.

בעיה 3.10. מטילים קוביה 3 פעמים. מה ההסתברות שהתקבלו ערכים בסדר עולה? מה ההסתברות שתוצאת ההטלה השלישית גדולה משתי ההטלות האחרות?

הוכחה. מטילים קוביה 3 פעמים, לכל הטלה יש 6 אפשרויות ולכן מרחב המדגם הוא בחירה של 3 מספרים מתוך ה-6 האפשריים עם החזרה ועם סדר – 6^3. בכדי שהערכים יהיו בסדר עולה ממש, יש לבחור את שלושת המספרים השונים שהתקבלו בהטלות, והבחירה נעשית ללא החזרה (כי התוצאות שונות) וללא סדר (בינתיים). יש לכך $\binom{6}{3}$ אפשרויות. לאחר שבחרנו את המספרים שהתקבלו, נשאר להחליט איזה מספר התקבל באיזו הטלה. למילוי הדרישה שהסדרה תעלה ממש יש רק אפשרות אחת – המספר הכי קטן יהיה ראשון, הכי גדול שלישי והאמצעי – שני. ההסתברות הכוללת הינה

$$\frac{\binom{6}{3}}{6^3}$$

עבור המאורע שההטלה השלישית היא הכי גדולה ישנן שתי אפשרויות – או שבהטלה הראשונה והשנייה יצאה תוצאה זהה, או שיצאו תוצאות שונות. במקרה הראשון עלינו לבחור את שני המספרים השונים שהתקבלו בניסוי, כאשר הגדול ביניהם התקבל בהטלה השלישית והקטן התקבל בראשונה ובשניה. במקרה השני, עלינו לבחור את שלושת התוצאות השונות שהתקבלו בניסוי, כאשר הגדול מביניהם חייב להיות המספר השלישי ואילו שני האחרים יכולים להיות הראשון והשני או השני והראשון. לסיכום, ההסתברות הכוללת הינה

$$\frac{\binom{6}{2} + \binom{6}{3} \cdot 2}{6^3}$$

□

תרגיל 3.17. מהו ערכו של הביטוי

$$\sum_{i=1}^{25} \sum_{j=1}^{i-1} c$$

כאשר $c \in \mathbb{R}$ קבוע?

3.3 שאלות מסכמות

תרגיל 3.18. כמה פתרונות יש למשוואה $ab(c+d+e+f)=33$ כאשר כל אות מייצגת מספר שונה בין 0 ל-9?

תרגיל 3.19. תהי A קבוצה בת n איברים ונסמן ב-$\mathbf{P}(A)$ את קבוצת החזקה שלה, קרי קבוצת כל הקבוצות החלקיות לה. לדוגמא: $\mathbf{P}(\phi)=\{\phi\}, \mathbf{P}(\{1\})=\{\phi,\{1\}\}, \mathbf{P}(\{1,2\})=\{\phi,\{1\},\{2\},\{1,2\}\}$.
הוכיחו בשתי דרכים קומבינטוריות שונות כי ישנם 2^n איברים בקבוצת החזקה.
רמז לאחת הדרכים: תרגיל 0.2.

תרגיל 3.20. כמה מספרים בני 6 ספרות המורכבים רק מהספרות 2,3 ו-9:

1. מתחלקים ב-2?
2. מתחלקים ב-3?
3. מתחלקים ב-6?

תרגיל 3.21. נתונות שתי הקבוצות $A=\{1,2,3,4,5\}$ ו-$B=\{6,7,8\}$. כמה פונקציות $f:A\rightarrow B$ קיימות כך ש:

1. אין הגבלה נוספת?
2. כך שהפונקציה חד-חד-ערכית?
תזכורת: פונקציה היא חד-חד-ערכית אם $x\neq y$ גורר $f(x)\neq f(y)$.
3. כך שהפונקציה על?
תזכורת: פונקציה היא על אם לכל $y\in B$ קיים $x\in A$ כך ש-$f(x)=y$.

תרגיל 3.22. נתונים עשרה מאורעות $A_1,\ldots,A_{10}$ סימטריים לחלוטין, כלומר ההסתברות של חיתוך של k מאורעות מתוכם שווה ל-p_k ללא תלות במיהם המאורעות המשתתפים בחיתוך. למשל, $Pr\left(A_1\cap A_3\right)=Pr\left(A_7\cap A_8\right)=p_2$. חשבו באמצעות נוסחת ההכלה וההפרדה (נוסחא 2.8) את ההסתברות שלפחות אחד המאורעות יתרחש.

תרגיל 3.23. משבצים באקראי 5 רופאים ל-10 תורנויות, בלי הגבלה על מספר התורנויות שכל רופא יכול לקבל.

1. מהו מספר החלוקות האפשרי?
2. מה ההסתברות שכל רופא ישובץ בדיוק לשתי תורנויות?

תרגיל 3.24. בכד 20 כדורים: 8 שחורים, 7 אדומים ו-5 לבנים. מוציאים באופן מקרי 8 כדורים. חשבו את ההסתברויות של המאורעות הבאים:

1. הוצאו 6 כדורים שחורים ו-2 לבנים.
2. הוצאו 3 כדורים שחורים, 2 לבנים ו-3 אדומים
3. הוצא לפחות כדור לבן אחד.

תרגיל 3.25. מחלקים באקראי m כדורים זהים ל-n תאים. חשבו את ההסתברות שכל התאים מלאים כאשר

1. $m=n$
2. $m=n+1$

תרגיל 3.26. מסדרים באקראי את הספרות של המספר 20042004. מה ההסתברות שהרצף 2004 יופיע במספר החדש?

תרגיל 3.27. עשרה סטודנטים נוסעים באוטובוס לאוניברסיטה. לאוטובוס יש 5 תחנות הסמוכות לקמפוס וכל סטודנט יורד באקראי באחת מהתחנות.

1. מה ההסתברות שבכל תחנה ירדו בדיוק שני סטודנטים?
2. מה ההסתברות שבתחנה הסופית ירדו בדיוק 5 סטודנטים?
3. בתחנה הראשונה ירדו מהאוטובוס שלושה סטודנטים. מה ההסתברות שהאוטובוס יגיע ריק לתחנה האחרונה?

תרגיל 3.28. בקופסא יש 100 סוללות, מתוכן 10 מלאות והשאר ריקות. בוחרים באקראי 5 סוללות ומנסים להפעיל איתן שלט הדורש שתי סוללות. מה ההסתברות שהשלט יעבוד, כלומר שמתוך 5 הסוללות שנבחרו לפחות 2 מלאות?

תרגיל 3.29. מטילים קוביה הוגנת 10 פעמים.

1. מה הסיכוי שהמספר המינימלי מבין 10 התוצאות הינו 4?
2. מה הסיכוי שהמספר 4 התקבל בדיוק 3 פעמים בהינתן שזהו המספר המינימלי?
3. מה הסיכוי שהמספר 5 התקבל בדיוק 3 פעמים בהינתן שהמספר המינימלי הינו 4?

תרגיל 3.30. נתונים $n+1$ כדים $(n>3)$ כל אחד מכיל בדיוק n כדורים. כד מספר $i\in\{0,...,n\}$ מכיל i כדורים לבנים והשאר שחורים. בוחרים באקראי כד ומוציאים ללא החזרה כדורים מתוכו.

1. בהנתן שבהוצאה הראשונה יצא כדור לבן, מהי ההסתברות שנבחר הכד ה-i?
2. בהנתן שבהוצאה הראשונה יצא כדור לבן, מהי ההסתברות שבהוצאה השניה גם כן יצא כדור לבן?
3. בהנתן שבשתי ההוצאות הראשונות יצא כדור לבן, מהי ההסתברות שגם בהוצאה השלישית ייצא כדור לבן?

תרגיל 3.31. אבי מטיל קוביה שוב ושוב עד שיוצאת לראשונה הספרה 6. טל מטילה מטבע שוב ושוב עד שיוצא לראשונה עץ. מה ההסתברות שמספר ההטלות של אבי שווה למספר ההטלות של טל?

3.4 פתרונות לשאלות הפרק

תרגיל 3.1. אף אחד ממרחבי המדגם המתוארים בשאלה אינם סימטריים:

1. הסכומים אינם סימטרים. הסכום 2 מתקבל רק כאשר שתי הקוביות נופלות על 1 ואילו הסכום 7 מתקבל בהרבה יותר אפשרויות. מרחב מדגם סימטרי יהיה הזוגות הסדורים של התוצאות שהתקבלו בקוביה: $\Omega = \{(1,1),(1,2),...,(6,5),(6,6)\}$ והוא כולל 36 תוצאות.
2. מרחב המדגם לא סימטרי: לבן מתקבל יותר פעמים משחור. מרחב מדגם סימטרי יהיה מרחב המדגם הרגיל $\Omega = \{1,2,3,4,5,6\}$.
3. עץ אחד יכול להתקבל בשתי אפשרויות (עץ ואז פלי או פלי ואז עץ) ואילו התוצאה "שני עצים" מתקבלת בדרך אחת בלבד. לכן מרחב המדגם לא סימטרי. מרחב מדגם סימטרי יכלול את כל הזוגות הסדורים האפשריים של ההטלות: $\Omega = \{HH, HT, TH, TT\}$.

□

תרגיל 3.2. נניח שיש α אפשרויות לסדר n פריטים במעגל. נסתכל על סידור כזה ונהפוך את המעגל לשורה על ידי כך שנבחר את אחד הפריטים (יש n אפשרויות בחירה), נמקם אותו במקום הראשון בשורה ואחריו נציב לפי הסדר את כל הפריטים שנמצאו במעגל עם כיוון השעון מהפריט הנבחר. למשל, אם אנחנו מסתכלים על השעות בשעון אנלוגי והתחלנו מהשעה 4, במקום השני נציב את 5, במקום השלישי את 6 ובמקום ה-12 את 3. באופן זה אנחנו יכולים לייצר את כל הסידורים של n פריטים בשורה ולכן מתקיים $n! = \alpha \cdot n$, כלומר מספר האפשרויות לסדר n פריטים במעגל הינו $(n-1)!$. □

תרגיל 3.3. נרשום את החזקה בצורה מפורשת בתור מכפלה: $(x+y)^n = (x+y)\cdot\ldots\cdot(x+y)$. כאשר נפתח את הסוגריים באגף ימין, כל גורם במכפלה יתרום x או y ולכן נקבל סכום של מכפלות הכוללות n נכפלים, חלקם x-ים והשאר y-ים, כלומר ביטויים מהצורה $x^k y^{n-k}$ כאשר $k = 0,1,\ldots,n$. נותר לקבוע מהו המקדם של $x^k y^{n-k}$, קרי מספר הפעמים שגורם זה מופיע בסכימה. בכדי ליצור את הביטוי הזה עלינו לבחור מתוך n הסוגריים את ה-k שיתרמו את ה-x והשאר יתרמו את ה-y. הבחירה נעשית ללא החזרה (אי אפשר לבחור את אותם סוגריים פעמיים) וללא סדר, שכן רק הסוגריים שנבחרו משנים ולא באיזה סדר הם נבחרו. לכן יש $\binom{n}{k}$ אפשרויות בחירה וזהו מספר הפעמים בו יופיע הגורם $x^k y^{n-k}$ בפתיחת הסוגריים, כדרוש לפי נוסחת הבינום. □

תרגיל 3.4. ניתן לבחור k פריטים מתוך n עם סדר וללא החזרה על ידי הניסוי הדו-שלבי הבא: בשלב הראשון, נבחר את k הפריטים ללא סדר וללא החזרה, ולאחר מכן נסדר את הפריטים שנבחרו. בשלב הראשון יש $\binom{n}{k}$ אפשרויות, בשלב השני יש $k!$ אפשרויות ובסה"כ יש $\binom{n}{k}\cdot k! = \frac{n!}{(n-k)!}$ אפשרויות.

דרך אחרת לספור אפשרויות אלו היא ישירות על ידי בחירת הפריטים. יש n אפשרויות לבחירת הפריט הראשון, $n-1$ אפשרויות לפריט השני וכן הלאה. בסך הכל נבחרים k פריטים לכן לפריט האחרון יש $n-(k-1)$ אפשרויות (כי נבחרו עד כה $k-1$ פריטים ועכשיו עומד להיבחר הפריט ה-k). מספר האפשרויות הכולל הינו $n\cdot(n-1)\cdot\ldots\cdot(n-k+1) = \frac{n!}{(n-k)!}$. □

תרגיל 3.5. בגירסא הראשונה הסיכוי זהה לסיכוי בלוטו הבולגרי שהופיע בדוגמא, $\frac{1}{\binom{49}{6}} = 7.15\cdot 10^{-8}$. בשנים 2004-2009 מרחב המדגם כלל 6 מספרים מתוך המספרים 1-34 וכן מספר חזק אחד מתוך 10 אפשרויות, לכן, לפי עקרון הכפל, מרחב המדגם הינו $\binom{34}{6}\cdot 10$ והסתברות הזכייה היא $\frac{1}{\binom{34}{6}\cdot 10} = 7.43\cdot 10^{-8}$. כיום יש לבחור מתוך 37 אפשרויות אבל יש רק 7 מספרים חזקים ולכן ההסתברות לזכייה היא $\frac{1}{\binom{37}{6}\cdot 7} = 6.14\cdot 10^{-8}$. סיכוי הזכיה היום הוא הכי נמוך מבין האפשרויות הללו והסיכוי הכי גבוה היה בשנים 2004-2009. □

תרגיל 3.6. יש 4 סדרות של קלפים ובכל סדרה יש 8 אפשרויות (13 קלפים ללא הקלפים 2-6) ולכן מרחב המדגם כולל בחירה של 4 קלפים מתוך 8 קלפים עם החזרה (כי אותו מספר יכול להיבחר כמה פעמים על ידי כמה סדרות) ועם סדר (הסדר הוא הסדרות השונות. לבחור קלף מלך-לב וקלף מלכה-יהלום זו בחירה שונה מקלף מלך-יהלום ומלכה-לב). לכן הסיכוי לנחש נכון את ארבעת הקלפים הינו $\frac{1}{4096} = \frac{1}{8^4}$. □

תרגיל 3.7. כל אבן דומינו מורכבת או משני מספרים שונים או מאותו המספר. כדי ליצור אבן המורכבת משני מספרים שונים מתוך 7 האפשרויות (0 עד 6) יש לבחור 2 מספרים מתוך 7 ללא החזרה וללא סדר ולכן יש $\binom{7}{2}$ אפשרויות. בנוסף, יש 7 אבני דומינו שרשום עליהן אותו המספר (צריך לבחור את המספר ויש 7 אפשרויות) ולכן בסה"כ יש $\binom{7}{2} + 7 = 28$ אבני דומינו שונות. □

תרגיל 3.8. 1. נפתור את התרגיל בשתי דרכים שונות. בדרך הראשונה נספור בכל פעם את מספר הנבחרות בהן יש שחקן טוב אחד (בדיוק) בכדורסל, שני שחקנים טובים, שלושה שחקנים טובים, ארבעה שחקנים טובים וחמישה שחקנים טובים ונחבר את התוצאות לפי עקרון החיבור. כדי לבחור נבחרת עם 3 שחקנים טובים בכדורסל (לדוגמא) צריך לבחור את 3 השחקנים הטובים מתוך ה-10 ולבחור את 2 השחקנים הנותרים מתוך ה-20 שאינם טובים בכדורסל. כל אחת מהבחירות נעשית ללא החזרה וללא סדר (כי מה שחשוב הוא זהות השחקנים שנבחרו ולא הסדר בו הם נבחרו לנבחרת) ולכן מספר הנבחרות הכולל הינו:

$$\binom{10}{1} \cdot \binom{20}{4} + \binom{10}{2} \cdot \binom{20}{3} + \binom{10}{3} \cdot \binom{20}{2} + \binom{10}{4} \cdot \binom{20}{1} + \binom{10}{5}$$

דרך זו אפשרית אבל מצריכה חישובים רבים במיוחד אם המספרים היו גדולים יותר. הדרך השניה היא להשתמש במשלים - נספור את מספר הנבחרות הכולל בלי אף הגבלה, נספור את מספר הנבחרות בהן אין אף שחקן טוב בכדורסל וההפרש בין המספרים הללו הוא מספר הנבחרות בהן יש לפחות שחקן אחד טוב (השוו למשפט 2.2):

$$\binom{30}{5} - \binom{20}{5}$$

2. בניגוד לקבוצת כדורסל, בקבוצת כדורגל לא כל השחקנים זהים כי יש שוער ויש שחקני שדה. לכן, שתי נבחרות המורכבות בדיוק מאותם 11 שחקנים אך עם שחקנים אחרים בתפקיד השוער צריכות להיספר כנבחרות שונות. נייצר את ההבדלה הזו בין נבחרות על ידי יצירת סדר בתהליך הבחירה: נבחר בשלב הראשון את השוער ובשלב השני נבחר את השחקנים האחרים (כל השאר נבחרים ללא סדר, כי אין תפקידים נוספים). באופן זה יהיה הבדל בין השחקן שנבחר ראשון לשחקנים האחרים.

$$\binom{10}{1} \cdot \binom{20}{4} \cdot \binom{9}{6}$$

□

תרגיל 3.9. בשני הניסויים אנחנו צריכים לחשב את ההסתברות שבסדרה של ניסויים קרה משהו לפחות פעם אחת (בניסוי הראשון 6, בניסוי השני שש-שש) והדרך הנוחה יותר לחשב את זה היא באמצעות המשלים - הסיכוי שהמאורע לא קרה בכלל בכל ההטלות. לכן הסיכוי לנצח במשחק הראשון הינו:

$$1 - \left(\frac{5}{6}\right)^4 = 0.5177$$

במשחק השני מטילים בכל סיבוב 2 קוביות ולכן מרחב המדגם הינו בחירה של 2 מספרים מתוך המספרים 1-6 עם החזרה ועם סדר (הסדר בא לידי ביטוי בכך שיש קוביות שונות ואפשר, למשל, להניח שמטילים אותן לפי הסדר). המאורע כולל את כל ההטלות פרט לאפשרות הבודדת שש-שש ולכן יש 35 אפשרויות שאינן שש-שש. לכן הסיכוי שב-24 הטלות לא יצא שש-שש הינו:

$$1-\left(\frac{35}{36}\right)^{24}=0.491$$

שבי, גיבור השאלה, הוא לא אחר מאשר האציל הצרפתי שבלייר דה-מרה (ראו את פסקת הפתיחה של פרק 2). דה-מרה נהג לשחק במשחק הראשון ושם לב שלאורך זמן הוא קצת יותר מרוויח מאשר מפסיד. כאשר הציעו לו לשחק במשחק השני הוא חשב מהשיקולים שפורטו בשאלה, שמדובר על משחקים שקולים והסכים. לאורך זמן דה-מרה שם לב שהוא מפסיד במשחק זה יותר מאשר מנצח. בניסיון להבין למה המשחקים לא שקולים ולמה הוא עבר להפסיד הוא פנה לפסקל והיתר היסטוריה. □

תרגיל 3.10. מרחב המדגם הוא חלוקה של 5 כדורים זהים ל-4 תאים. נמספר את הכדורים ונבצע את החלוקה על ידי כך שהכדורים יבחרו לפי הסדר את התא אליו הם נכנסים. הבחירה של התאים נעשית עם סדר (זהו סדר הכדורים) ועם החזרה ולכן מרחב המדגם הינו 4^5.
כדי לייצר את המאורע אנחנו צריכים לחלק את 5 הכדורים ל-4 תאים כך שכל התאים יהיו מלאים. לכן בהכרח יהיה תא אחד עם שני כדורים וביתר התאים יהיה כדור בודד. לכן עלינו לבחור את זוג הכדורים שילך ביחד ($\binom{5}{2}$ אפשרויות) ולאחר מכן נישאר עם 4 פריטים (זוג ושלושה כדורים בודדים) לחלק ל-4 תאים. חלוקה זו היא למעשה סידור בשורה (בחירת הפריט לתא הראשון, בחירת הפריט לתא השני וכו') ולכן בשלב הזה יש $4!$ אפשרויות. לסיכום, ההסתברות המבוקשת הינה:

$$\frac{\binom{5}{2}\cdot 4!}{4^5}=0.234$$

טעות נפוצה היא לפתור את התרגיל באופן הבא: נכניס את הכדורים אחד אחד לתאים ונדאג למלא את כל התאים: יש לנו 4 אפשרויות לכדור הראשון, 3 אפשרויות לכדור השני, 2 אפשרויות לכדור השלישי ואפשרות אחת לכדור הרביעי. לאחר שסידרנו את כל 4 הכדורים הראשונים מילאנו כבר את כל התאים ודאגנו שהמאורע מתקיים, לכן הכדור החמישי יכול להיכנס לכל אחד מהתאים בלי הגבלה, כלומר יש לו 4 אפשרויות. בסה"כ: $\frac{4!\cdot 4}{4^5}=0.09$. ההסתברות שקיבלנו בדרך זו קטנה מההסתברות האמיתית והסיבה לכך היא שבמאורע ספרנו רק חלק מהסידורים האפשריים. בכל אחד מהסידורים במאורע תמיד גרמנו לכך שכדור מספר חמש יהיה זה שנמצא בתא עם כדור אחד נוסף, אך זה לא מצב הכרחי. ייתכן בהחלט שכדורים 1 ו-2 הם אלו שחולקים תא וכדור 5 שוכן לבדו בתא, אפשרות שלא נספרה בדרך זו. □

תרגיל 3.11. תרגיל זה דומה מאוד לבעיה 3.4. גם כאן 6 יצא לראשונה בהטלה האחרונה ואנחנו מחפשים את הסיכוי שבשאר ההטלות יצאו רק זוגיים, אך הפעם אנחנו דורשים שכל המספרים הזוגיים יתקבלו. לכן מרחב המדגם לא משתנה ואילו המאורע, בו יצאו רק 2 ו-4 כולל שתי תוצאות מיותרות: המצב בו יצא רק 2 בכל ההטלות והמצב בו יצא רק 4 בכל ההטלות. לכן בתרגיל זה מספר האפשרויות עבור המאורע הינו 2^4-2 וההסתברות היא $\frac{2^4-2}{5^4}$. □

תרגיל 3.12. בכל אחת מההטלות מרחב המדגם כולל את כל אחת מ-6^2 הזוגות האפשריים בהטלת שתי קוביות (ולא הסכומים, שכן אינם שווי הסתברות). אנחנו לא יודעים כמה הטלות נדרשו עד אשר יצא סכום המתחלק ב-5, אבל אנחנו יודעים שזה קרה בסופו של דבר (לפי משפט הקוף המקליד, משפט 3.4) ולכן ניתן להתנות על ההטלה

זו. בהינתן שאנחנו בהטלה האחרונה של הניסוי, בה יצא מספר המתחלק ב-5, ייתכנו רק שני סכומים אפשריים: 10 או 5. הטלות הקוביה המתאימות לסכומים אלו הן $\{(1,4),(2,3),(3,2),(4,1),(4,6),(5,5),(6,4)\}$. יש 7 אפשרויות שוות הסתברות, מתוכן 4 מתאימות לסכום 5 ו-3 מתאימות לסכום 10, ולכן הסיכוי שבהטלה האחרונה סכום הקוביות הינו 5 הוא $\frac{4}{7}$.

הערה: ניתן לפתור את השאלה גם על ידי שימוש בתרגיל 2.16. □

תרגיל 3.13. נגדיר את המאורעות A– נבחרה השידה עם 2 מטבעות זהב, B– נבחרה השידה עם מטבע זהב ומטבע כסף, C– נבחרה השידה עם שני מטבעות כסף ו-G– נבחר מטבע זהב. אנו מעוניינים לחשב את ההסתברות $Pr(A|G)$ (במגירה השניה יש מטבע זהב רק אם נבחרה השידה A). אנחנו לא יודעים לחשב הסתברות זו ישירות, אבל אנחנו יודעים לחשב את ההסתברות שנבחר מטבע זהב בהינתן השידה שנבחרה (באופן טריוויאלי, $Pr(G|A)=1$) ולכן נהפוך את כיוון ההתנייה על ידי נוסחת בייס:

$$Pr(A|G)=\frac{Pr(G|A)\cdot Pr(A)}{Pr(G)}$$

בבירור, $Pr(A)=\frac{1}{3}$. על מנת לחשב את ההסתברות שנבחר מטבע זהב המופיעה במכנה עלינו לדעת מתוך איזו שידה הבחירה נעשית. נשתמש בנוסחת ההסתברות השלמה ונתנה על זהות השידה שנבחרה:

$$\begin{aligned} Pr(G) &= Pr(G|A)\cdot P(A)+Pr(G|B)\cdot P(B)+Pr(G|C)\cdot P(C) \\ &= 1\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}=\frac{1}{2} \end{aligned}$$

למעשה יכולנו לדעת שההסתברות היא חצי מאחר ויש סימטריה מוחלטת בבעיה בין מטבעות הזהב והכסף ולכן הסיכוי להוציא מטבע זהב שווה לסיכוי להוציא מטבע כסף. נציב את כל המספרים שקיבלנו בנוסחת בייס ונקבל $Pr(A|G)=\frac{2}{3}$. □

תרגיל 3.14. נשתמש באנלוגיה לסידורים בשורה. מרחב המדגם הוא סידור של 10 כדורים בשורה וגודלו $10!$. נבנה את המאורע בשלבים: נבחר את הכדור הלבן שנשאר בתוך הכד (3), נבחר את המקום שלו בשורה (2 אפשרויות, מקומות 9 או 10 בשורה הם המקומות שנמצאים בתוך הכד), נבחר את הכדור שיצא קודם מבין שני הלבנים (2 אפשרויות) ונבחר את המקום שבו הוא נמצא בשורה (7 אפשרויות, שכן צריך להשאיר מקום גם למקום הלבן האחרון שיצא מיד אחריו). הכדור הלבן הנותר נמצא במקום הבא בתור ויש לכך אפשרות אחת. סיימנו לסדר את הכדורים הלבנים. נשארנו עם 7 כדורים שחורים ו-7 מקומות אפשריים, אין הגבלה על סידור המקומות השחורים ולכן יש $7!$ אפשרויות. ההסתברות המבוקשת היא:

$$\frac{3\cdot 2\cdot 2\cdot 7\cdot 7!}{10!}=0.12$$

□

תרגיל 3.15. 1. נסמן את המאורע שהזוג ה-i יושב ביחד ב-A_i. ההסתברות שאף זוג לא יושב ביחד היא המשלים של המאורע לפחות זוג אחד היושב ביחד, $A_1\cup\ldots\cup A_4$, אותו ניתן לחשב באמצעות נוסחת ההכלה וההפרדה ל-4 מאורעות (ראו נוסחא 2.8):

$$Pr\left(A_1\cup A_2\cup A_3\cup A_4\right)=\sum_{i=1}^{4}Pr\left(A_i\right)-\sum_{i\neq j}Pr\left(A_i\cap A_j\right)$$
$$+\sum_{i\neq j\neq k}Pr\left(A_i\cap A_j\cap A_k\right)-Pr\left(A_1\cap\ldots\cap A_4\right)$$

בכדי לחשב את הסיכוי שזוג אחד מסויים יושב ביחד, נתייחס אליו כאל פריט אחד ונקבל $Pr\left(A_i\right)=\frac{2!\cdot 7!}{8!}=\frac{1}{4}$. באותה צורה, בכדי ששני זוגות מסויימים יישבו ביחד, יש להתייחס לכל אחד מהזוגות כאל פריט אחד ונקבל $Pr\left(A_i\cap A_j\right)=\frac{(2!)^2\cdot 6!}{8!}=\frac{1}{14}$, ההסתברות ששלושה זוגות יושבים ביחד היא $\frac{(2!)^3\cdot 5!}{8!}$ וההסתברות שכל הזוגות יושבים ביחד היא $\frac{(2!)^4\cdot 4!}{8!}$. בנוסף יש $\binom{4}{2}$ זוגות אפשריים ו-$\binom{4}{3}$ שלשות אפשריות, לכן ההסתברות שלפחות זוג אחד יושב ביחד תהיה:

$$4\cdot\frac{2!\cdot 7!}{8!}-\binom{4}{2}\cdot\frac{(2!)^2\cdot 6!}{8!}+\binom{4}{3}\cdot\frac{(2!)^3\cdot 5!}{8!}-\frac{(2!)^4\cdot 4!}{8!}=0.657 \qquad (3.3)$$

וההסתברות שאף זוג לא יושב ביחד היא $1-0.657=0.343$.

2. חישובו של דני אינו נכון, שכן הוא אינו סופר את כל האפשרויות. למשל, כאשר הוא מציב את גברת א', הוא מכריח אותה לעמוד כך שבינה לבין מר א' יהיה לפחות גבר אחד. אבל, כאשר מסתכלים על כל הסידורים האפשריים, גברת א' יכולה לא לעמוד ליד מר א' כאשר ביניהם יש רק נשים ולא גברים אחרים. באותה צורה, בין גברת ב' ומר ב' בהכרח מפרידים גברים או גברת א', אבל ניתן לחשוב גם על סידורים בהם בין מר ב' לגברת ב' עומדת רק גברת ג'. דני לא סופר סידורים כאלו, לכן ההסתברות שלו תהיה קטנה מההסתברותה אמיתית, כפי שאכן ניתן לראות על ידי חישוב המכפלה.

3. בהינתן שמר סמית' יושב בקצה השורה, לגברת סמית' יש 7 מקומות אפשריים אבל רק אחד מהם הוא ליד בעלה ולכן ההסתברות היא $\frac{1}{7}$.

4. ניתן לפתור את התרגיל הזה באמצעות קומבינטוריקה, אבל חישבנו כבר את כל ההסתברויות בסעיפים הקודמים ולכן יותר נוח לפתור את התרגיל על ידי נוסחת בייס. נגדיר A – מר וגברת סמית' יושבים יחד ו-B– מר סמית' יושב בקצה. מתקיים:

$$Pr\left(B|A\right)=\frac{Pr\left(A|B\right)\cdot Pr\left(B\right)}{Pr\left(A\right)}=\frac{\left(\frac{1}{7}\right)\cdot\left(\frac{2}{8}\right)}{\frac{1}{4}}=\frac{1}{7}$$

5. נסמן ב-S את המאורע שמר וגברת סמית' יושבים יחד וב-C את המאורע שמר וגברת כהן יושבים יחד. חישבנו כבר את $Pr\left(S\right)=\frac{1}{4}$ ואת $Pr\left(S\cap C\right)=\frac{1}{14}$ ונשאלנו על $Pr\left(S\cap\overline{C}\right)$. מאחר ומתקיים $S=\left(S\cap C\right)\cup\left(S\cap\overline{C}\right)$ ומאורעות אלו זרים בזוגות, מתקיים $Pr\left(S\right)=Pr\left(S\cap C\right)+Pr\left(S\cap\overline{C}\right)$, כלומר $Pr\left(S\cap\overline{C}\right)=\frac{1}{4}-\frac{1}{14}=\frac{5}{28}$.

□

תרגיל 3.16 1. מסימטריה, בכל צבע יש הסתברות של $\frac{1}{2}$ שכדור מספר 1 ייצא לפני כדור מספר 2. הסדר היחסי של שני הכדורים השחורים לא משפיע על הסדר היחסי של הכדורים האדומים ולכן כל הצבעים הם בלתי תלויים וההסתברות המבוקשת היא $\left(\frac{1}{2}\right)^5$.

2. המאורעות תלויים. נתחיל בהסבר אינטואיטיבי ולאחר מכן נחשב זאת מפורשות. אם ידוע שכדור מספר 1 שחור יצא לפני כדור מספר 2 שחור, זה סימן שכדור 1 שחור יצא יחסית בהתחלה, מה שמגדיל את ההסתברות שיצא גם לפני כדור מספר 1 אדום. לעומת זאת, אם כדור מספר 1 שחור יצא אחרי כדור מספר 2 שחור, הרי שהוא יצא יחסית מאוחר וכנראה שיצא גם אחרי 1 אדום.
נסמן A – כדור 1 שחור יצא לפני כדור 2 שחור ו-B – כדור 1 שחור יצא לפני כדור מספר 1 אדום. בדומה לסעיף הקודם, $Pr\left(A\right)=Pr\left(B\right)=\frac{1}{2}$. נחשב את ההסתברות של החיתוך, $A\cap B$, כלומר שכדור מספר 1 שחור יצא לפני 2 שחור וגם כדור מספר 1 שחור יצא לפני כדור מספר 1 אדום. המאורע עוסק בסידורים הפנימיים של שלושת כדורים אלו בלבד ולכן ניתן לצמצם את מרחב המדגם ולהסתכל רק עליהם. המאורע מתרחש כאשר כדור 1 שחור הוא הראשון מבין השלושה והסיכוי לכך הוא $\frac{1}{3}$. כלומר

$$Pr\left(A\cap B\right)=\frac{1}{3}\neq\frac{1}{2}\cdot\frac{1}{2}=Pr\left(A\right)\cdot Pr\left(B\right)$$

והמאורעות תלויים.

3. מסימטריה בין הצבעים, הסיכוי ששני השחורים יצאו לפני שני האדומים זהה לסיכוי ששני האדומים יצאו לפני שני השחורים, כלומר ההסתברות היא חצי. אפשר לראות זאת גם על ידי צמצום מרחב המדגם לסידורים של ארבעת הכדורים השחורים והאדומים. שני השחורים יצאו לפני שני האדומים אם במקום הרביעי ביניהם עומד כדור אדום, וההסתברות לכך היא $\frac{2}{4}$.

4. בהינתן שהכדור הרביעי שיצא מהכד היה כדור שחור מספר 2, לכדור השחור מספר 1 יש אותו סיכוי להיות בכל אחד מ-9 המקומות האחרים ולכן הסיכוי שהוא הכדור השני שיצא מהכד יהיה $\frac{1}{9}$.

□

תרגיל 3.17. הביטוי $\sum_{i=1}^{25}\sum_{j=1}^{i-1}c$ הוא למעשה $c+c+\ldots$ כאשר השאלה היחידה היא כמה פעמים סוכמים את c. הסכימה על i נעשית בין 1 ל-25 והסכימה על j נעשית על כל המספרים בין 1 ל-$i-1$, כלומר האינדקסים i,j הם שני מספרים שונים בתחום 1-25 כך שהקטן הוא j. יש $\binom{25}{2}$ אפשרויות בחירה של שני מספרים שונים בתחום הזה ולאחר שהם נבחרו, יש רק סידור אחד אפשרי בו הקטן הוא j. לכן בסכום הזה המספר c מופיע $\binom{25}{2}$ פעמים ולכן $\sum_{i=1}^{25}\sum_{j=1}^{i-1}c=\binom{25}{2}c$. □

תרגיל 3.18. נשים לב שיש רק דרך אחת לרשום את 33 כמכפלה של 3 מספרים: $33=1\cdot3\cdot11$. לכן בהכרח $a=1,b=3$ או להפך. סכום 4 הספרות הקטנות ביותר שנותרו נותן בדיוק $11=0+2+4+5$ ולכן אלו הן בדיוק האופציות עבור שאר המשתנים. נסכם – יש 2 אופציות עבור a,b ו-$4!$ אופציות לשאר המשתנים ולכן בסה"כ מספר הפתרונות של המשוואה הינו $2\cdot4!=48$. □

תרגיל 3.19. תהי A קבוצה בת n איברים ונסמן ב-$\mathbf{P}(A)$ את קבוצת החזקה שלה. נוכיח שיש ב-$\mathbf{P}(A)$ 2^n איברים.

1. <u>דרך א:</u> כדי לבנות תת קבוצה של A, כל איבר מאיבריה צריך ל"בחור" האם הוא חלק מהתת-קבוצה או לא. לכל אחד מ-n-האיברים יש 2 אפשרויות בחירה עם החזרה (להיות חלק מהקבוצה או לא להיות חלק מהקבוצה) והבחירה נעשית עם סדר (הסדר מייצג את העובדה שהאיברים שונים). לכן מספר האפשרויות הוא 2^n, כלומר ישנן 2^n תתי קבוצות אפשריות.

2. <u>דרך ב:</u> כדי לייצר תת קבוצה בת k איברים צריך לבחור k איברים מתוך n, ללא חשיבות לסדר (כך מוגדרת קבוצה) וללא חזרה. לכן, מספר תתי הקבוצות בגודל k הוא $\binom{n}{k}$ ומספר תתי הקבוצות מכל הגדלים האפשריים הוא $\sum_{k=0}^{n}\binom{n}{k}=2^n$ (ראו תרגיל 0.2).

□

תרגיל 3.20. 1. כדי שמספר יתחלק ב-2 הוא צריך להסתיים ב-2 (מבין האפשרויות הנתונות בשאלה). לכן יש אפשרות אחת בלבד עבור ספרת האחדות ו-3 אפשרויות לכל אחת מהספרות האחרות. בסה"כ יש $1 \cdot 3^5 = 243$ אפשרויות.

2. כדי שמספר יתחלק ב-3 סכום הספרות שלו צריך להתחלק ב-3. מאחר וכל הופעה של 3 או 9 בסכום הספרות לא משנה את השארית בחלוקה ל-3, הרי שמה שיקבע האם הסכום מתחלק ב-3 הוא מספר ההופעות של 2 במספר. ישנן 3 אפשרויות:

   א) 0 הופעות של 2. לכן המספר מורכב רק מ-3 ומ-9 ויש לכך 2^6אפשרויות.

   ב) 3 הופעות של 2. יש $\binom{6}{3}$ לבחור את המקומות בהן ה-2 יופיעו ו-2^3 אפשרויות להציב 3 או 9 בשאר שלושת המקומות. בסה"כ באפשרות זו יש $\binom{6}{3} \cdot 2^3$ אפשרויות.

   ג) 6 הופעות של 2. במקרה כזה יש רק אפשרות אחת, המספר הוא $222,222$.

   לסיכום, יש $2^6 + \binom{6}{3} \cdot 2^3 + 1 = 225$

3. כדי שמספר יתחלק ב-6 הוא צריך להסתיים ב-2 וגם להתחלק ב-3. מתוך האפשרויות שמנינו בסעיף הקודם, ייתכנו רק האפשרויות הבאות:

   א) 6 הופעות של 2.

   ב) 3 הופעות של 2, כאשר אחת מהן היא בספרת האחדות. במקרה כזה יש $\binom{5}{2}$ אפשרויות לבחור את שני המקומות הנוספים עבור ה-2 ו2^3 אפשרויות עבור שאר המקומות.

   לסיכום, נקבל $1 + \binom{5}{2} \cdot 2^3 = 81$.

□

תרגיל 3.21. 1. כל פונקציה מהווה בחירה של ערך מהקבוצה B עבור כל אחד מהאיברים ב-A. ישנם ב-A 5 איברים ולכן אנחנו צריכים לבחור מתוך שלושת איברי B חמישה איברים עם החזרה ועם סדר, כך שמספר הפונקציות הכולל הינו 3^5.

2. לא ניתן להגדיר פונקציה חד-חד-ערכית מקבוצה בת 5 איברים לקבוצה בת 3 איברים, שכן בוודאות לפחות אחד מאיברי הטווח ייבחר לפחות פעמיים.

3. כאמור, יש בסך הכל 3^5 פונקציות מ-A ל-B. נבדוק אילו מהן אינן על. כל הפונקציות מ-A ל-$\{6,7\}$ אינן על ויש 2^5 פונקציות כאלו. ישנן 3 תתי קבוצות בגודל 2 ולכן בסך הכל יש $3 \cdot 2^5$ פונקציות לא מתאימות. נשים לב שכל פונקציה קבועה נספרה פעמיים (הפונקציה הקבועה $f(x) = 6$ נספרה גם במסגרת הפונקציות לקבוצה $\{6,7\}$ וגם במסגרת הפונקציות לקבוצה $\{6,8\}$) לכן נוריד אותן. יש 3 פונקציות קבועות ולכן בסך הכל יש $3 \cdot 2^5 - 3 = 93$ פונקציות שאינן על B או $3^5 - 93 = 150$ פונקציות על.

□

תרגיל 3.22. בנוסחת ההכלה וההפרדה (נוסחא 2.8) מופיעות ההסתברויות של כל המאורעות, של כל החיתוכים של שני מאורעות, של כל החיתוכים של שלושה מאורעות וכן הלאה. מספר החיתוכים האפשריים של k מאורעות הינו $\binom{10}{k}$, כמספר האפשרויות לבחור את k המאורעות שישתתפו בחיתוך. לכן ההסתברות של חיתוך k מאורעות, p_k, נסכמת בנוסחא $\binom{10}{k}$ פעמים וניתן להחליפה סכימה זו בכפל ב-$\binom{10}{k}$. נציב זאת בנוסחת ההכלה וההפרדה ונקבל $\sum_{k=1}^{10} (-1)^{k+1} \binom{10}{k} p_k$. באופן דומה ההסתברות שלפחות מאורע אחד מתוך n מאורעות סימטריים לחלוטין יתרחש הינה

$$\sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} p_k \qquad (3.4)$$

השתמשנו במשוואה זו עבור $n = 4$ כחלק מפתרון תרגיל 3.15 (נוסחא 3.3). □

תרגיל 3.23. 1. מספר התורנויות שכל רופא יכול לבחור לא נתון ולא ידוע מראש, אך מספר הרופאים שכל תורנות צריכה לבחור הוא בדיוק 1, לכן בבעיה זו התורנויות הן אלו שיבחרו את הרופאים ומרחב המדגם כולל בחירה עם החזרה של 5 הרופאים עם סדר (הסדר מבטא את התורנויות השונות) וכולל 5^{10} אפשרויות.

2. כדי שכל רופא יקבל שתי תורנויות, כל רופא צריך לבחור בדיוק 2 תורנויות מתוך אלו שעדיין לא נבחרו. הבחירה של כל אחד מהרופאים לא החזרה וללא סדר (כי לא משנה באיזה סדר הרופא בחר את התורנויות שלו, אלא רק את התורנויות שהוא בחר). לכן לרופא הראשון יש $\binom{10}{2}$ אפשרויות, לרופא השני יש $\binom{8}{2}$ וכן הלאה. ההסתברות של המאורע המבוקש היא אם כך:

$$\frac{\binom{10}{2} \cdot \binom{8}{2} \cdot \binom{6}{2} \cdot \binom{4}{2} \cdot \binom{2}{2}}{5^{10}} = 0.012$$

□

תרגיל 3.24. נתייחס אל כל הכדורים כאל כדורים שונים, גם אם הם מאותו הצבע. מרחב המדגם הוא מספר הדרכים השונות לבחור 8 כדורים מתוך 20 ללא החזרה וללא סדר: $\binom{20}{8} = 125970$.

1. כדי שמאורע זה יתקיים צריך לבחור את 6 הכדורים שיבחרו מתוך 8 השחורים ואת 2 הכדורים שייבחרו מתוך 5 הלבנים, לכן מספר האפשרויות הוא $\binom{8}{6} \cdot \binom{5}{2} = 280$ וההסתברות היא $\frac{280}{125970} = 0.0022$.
2. כדי שהמאורע יתקיים צריך לבחור 3 כדורים מתוך 8 השחורים, 2 כדורים מתוך 5 הלבנים ו-3 כדורים מתוך 7 האדומים, בדומה לסעיף הקודם. לכן ההסתברות היא $\frac{\binom{8}{3}\cdot\binom{5}{2}\cdot\binom{7}{3}}{\binom{20}{8}} = 0.1556$.
3. מאחר ויש הרבה אפשרויות להוציא לפחות כדור לבן אחד, נחשב באמצעות המשלים – אפס כדורים לבנים. במקרה כזה המאורע מוגבל לבחירה רק של כדורים מתוך 15 הכדורים שאינם לבנים, כשאין חשיבות חלוקה הפנימית בין הצבעים שאינם לבן. לכן ההסתברות המבוקשת היא $1 - \frac{\binom{15}{8}}{\binom{20}{8}} = 0.9489$.

□

תרגיל 3.25. כדי לקבוע את מרחב המדגם של הבעיה (הזהה לכל הסעיפים) עלינו להחליט מי בוחר את מי – התאים את הכדורים או להפך. מאחר וכל כדור בוחר בדיוק תא יחיד להיכנס אליו ואילו כל תא יכול לבחור מספר של כדורים, נקבע שהכדורים הם אלו שבוחרים את התאים. לכן התהליך כולל בחירה של m תאים (אחד לכל כדור) מתוך n אפשרויות עם החזרה. אמנם לצורך הבעיה הכדורים זהים ולא חשוב איזה בדיוק כדור נמצא בכל תא אלא רק שכל התאים מלאים, אך כדי שמרחב המדגם יהיה סימטרי עלינו להניח שהכדורים אינם זהים. נניח אם כך שהכדורים ממוספרים מ-1 עד m והסדר בו הכדורים בוחרים את התאים אליהם הם נכנסים הוא לפי מספרם. לסיכום, מרחב המדגם הינו בחירה של m תאים מתוך n אפשריים עם סדר ועם החזרה – n^m.

1. ישנם n כדורים ו-n תאים ויש להתאים לכל כדור תא. לכן לכדור הראשון יש n אפשרויות, לכדור השני יש $n-1$ אפשרויות וכן הלאה. בסך הכל מספר הסידורים האפשרי הינו $n!$ וההסתברות היא $\frac{n!}{n^n}$.
2. מקרה זה הוא הכללה של תרגיל 3.10 ולכן ההסתברות היא $\frac{\binom{n+1}{2}\cdot n!}{n^{n+1}}$.

□

תרגיל 3.26. נניח שכל הספרות שונות זו מזו, גם אלו הדומות (למשל, נניח שכל אחת מהספרות החוזרות על עצמן צבועה בצבע אחר). לכן מרחב המדגם הוא סידור של שמונה ספרות שונות בשורה וגודלו $8!$. בכדי שהרצף 2004 יופיע במספר החדש נהפוך את 4 הספרות הללו לפריט אחד ונסדר אותו כמו שהוא עם 4 הספרות האחרות.

בכדי לייצר את הרצף צריך לבחור איזו מהספרות 2 תופיע ברצף (יש 2 אפשרויות), איזו מהספרות 4 תופיע ברצף (2 אפשרויות) ואיזו מהספרות 0 תופיע במספר (יש 4אפשרויות לאפס הראשון ו-3 לשני). בסך הכל יש $2\cdot2\cdot4\cdot3=48$ אפשרויות להרכיב את הרצף. לאחר מכן ישנם 5 פריטים: 4 הספרות שלא נבחרו והרצף ויש $5!$ אפשרויות לסדר את הספרות הללו.
חשוב לשים לב כי הרצף 20042004 נספר פעמיים: פעם אחת כאשר ארבעת הספרות הראשונות הן הרצף שבחרנו, הצבנו אותו במקום הראשון ובמקרה שאר הספרות הסתדרו ברצף זה ופעם נוספת כאשר הצבנו את הרצף במקום החמישי ובמקרה 4 הספרות האחרות הסתדרו במקומות 1-4 בסדר זה. לכן נפחית מ-$5!$ הסידורים האפשריים את הסידור הכפול שהתקבל והתוצאה הסופית תהיה:

$$\frac{48\cdot(5!-1)}{8!}=0.142$$

□

תרגיל 3.27. כל סטודנט בוחר תחנה בה הוא יכול לרדת מתוך 5 התחנות האפשריות, הבחירה נעשית עם החזרה ועם סדר (הסדר מייצג את הסטודנטים השונים) ולכן מרחב המדגם בשני הסעיפים הראשונים הינו 5^{10}.

1. בכל תחנה נבחר זוג סטודנטים שיורדים בתחנה זו. הבחירה נעשית ללא החזרה וללא סדר, שכן משנה רק זהות הסטודנטים שנבחרו לרדת מהאוטובוס ולא באיזה סדר הם נבחרו לרדת ממנו. בתחנה הראשונה יש $\binom{10}{2}$ אפשרויות לבחור את הסטודנטים שירדו. בתחנה השניה יש על האוטובוס 8 סטודנטים ולכן יש $\binom{8}{2}$ אפשרויות לבחור את הזוג שירד בה וכך הלאה. ההסתברות הכוללת הינה

$$\frac{\binom{10}{2}\cdot\binom{8}{2}\cdot\binom{6}{2}\cdot\binom{4}{2}\cdot\binom{2}{2}}{5^{10}}=0.012$$

2. כדי שבתחנה הסופית ירדו בדיוק 5 סטודנטים, יש לבחור את חמשת הסטודנטים שירדו בתחנה זו. שאר הסטודנטים רשאיים לבחור כל תחנה מבין ארבע התחנות האחרות (עם החזרה) ולכן ההסתברות הינה

$$\frac{\binom{10}{5}\cdot4^5}{5^{10}}=0.026$$

3. בהינתן שבתחנה הראשונה ירדו מהאוטובוס שלושה סטודנטים, החל מהתחנה השניה יש 7 סטודנטים על האוטובוס ו-4 תחנות, לכן מרחב המדגם החדש הינו 4^7. האוטובוס מגיע לתחנה האחרונה ריק כאשר כל אחד מהסטודנטים בוחר אחת מתוך שלושת תחנות האחרות ולכן ההסתברות היא $\frac{3^7}{4^7}=0.133$.

□

תרגיל 3.28. השלט יעבוד כאשר יש לפחות 2 סוללות תקינות. נחשב את המאורע המשלים הכולל פחות אפשרויות, כלומר שיש 0 סוללות מלאות או אחת מלאה. מרחב המדגם הוא בחירה של 5 סוללות מתוך ה-100, ללא החזרה וללא סדר. כדי לבחור 0 סוללות מלאות, יש לבחור את כל ה-5 מתוך 90 הריקות. כדי לבחור סוללה מלאה אחת, יש לבחור אותה מתוך 10 המלאות ולבחור עוד 4 מתוך 90 הריקות. ההסתברות המבוקשת היא אם כך:

$$1-\frac{\binom{90}{5}}{\binom{100}{5}}-\frac{\binom{10}{1}\cdot\binom{90}{4}}{\binom{100}{5}}=0.077$$

□

תרגיל 3.29. מטילים קוביה הוגנת 10 פעמים, לכן מרחב המדגם הוא בחירה של 6 אפשרויות על ידי כל הטלה עם החזרה ועם סדר (ההטלות שונות) ולכן מרחב המדגם הינו 6^{10}.

1. כדי שהמינימום של ההטלות יהיה 4 צריך שבכל ההטלות יתקבלו רק מספרים מתוך 4, 5, ו-6, כשבהכרח 4 חייב להתקבל. לכן לכל הטלה יש רק שלוש אפשרויות ומספר האפשרויות הכולל הוא 3^{10}. יחד עם זאת, בספירה זו אנחנו גם כוללים את האפשרויות שרק 5 התקבל או, בהכללה, שרק המספרים 5 ו-6 התקבלו. יש 2^{10} אפשרויות כאלו ולכן מספר האפשרויות שבהטלות יצאו המספרים 4, 5 ו-6 ובהכרח 4 יתקבל הוא $3^{10} - 2^{10}$. לכן ההסתברות שהמספר המינימלי הוא

$$\frac{3^{10} - 2^{10}}{6^{10}} = 0.00096$$

2. בהינתן שהמינימום הוא 4, מרחב המדגם החדש כולל את כל תוצאות ההטלה בהן המינימום הוא 4 ומספר האפשרויות הוא $3^{10} - 2^{10}$. נעבור לחישוב המאורע. כדי ש-4 יתקבל בדיוק 3 פעמים יש לבחור את מספרי ההטלות בהן זה יקרה (זו בחירה של 3 ההטלות בלי החזרה ובלי סדר) ולבחור תוצאות לכל אחת מ-7 ההטלות האחרות, כאשר יש לכל הטלה 2 אפשרויות. לכן ההסתברות ש-4 יתקבל בדיוק 3 פעמים בהנתן שהתוצאה המינימלית היא 4 היא

$$\frac{\binom{10}{3} \cdot 2^7}{3^{10} - 2^{10}} = 0.265$$

3. באופן דומה לסעיף הקודם, מרחב המדגם כולל $3^{10} - 2^{10}$ ולחישוב המאורע יש לבחור את שלושת ההטלות בהן יצא המספר 5. עבור כל אחת מההטלות האחרות יש כמקודם 2 אפשרויות (4 או 6) אבל לא ייתכן שבכל ההטלות נבחרה הסיפרה 6 (כי אז המינימום אינו 4) לכן צריך להוריד אפשרות זו. לסיכום, ההסתברות המבוקשת הינה

$$\frac{\binom{10}{3} \cdot (2^7 - 1)}{3^{10} - 2^{10}} = 0.263$$

□

תרגיל 3.30. נסמן את המאורעות A_i – נבחר הכד ה-i, W_j – יצא כדור לבן בהוצאה ה-j. נשים לב שלכל כד יש אותו סיכוי להיבחר $Pr\left(A_i\right) = \frac{1}{n+1}$ ומסימטריה בין הכדורים הלבנים לשחורים (יש אותו מספר כדורים לבנים ושחורים, ועל כל כד שבו יש i לבנים יש כד מקביל עם i שחורים) מתקיים $Pr\left(W_j\right) = \frac{1}{2}$.

1. נשתמש בנוסחת בייז:

$$Pr\left(A_i|W_1\right) = \frac{Pr\left(W_1|A_i\right) Pr\left(A_i\right)}{Pr\left(W_1\right)} = \frac{\frac{i}{n} \cdot \frac{1}{n+1}}{\frac{1}{2}} = \frac{2i}{n(n+1)}$$

2. נחשב באופן ישיר לפי הגדרת ההסתברות המותנית:

$$\begin{aligned} Pr\left(W_2|W_1\right) = \frac{Pr\left(W_2 \cap W_1\right)}{Pr\left(W_1\right)} &= \frac{\sum_{j=1}^{n} Pr\left(W_1 \cap W_2|A_j\right) Pr\left(A_j\right)}{Pr\left(W_1\right)} \\ = \frac{\sum_{j=1}^{n} \frac{1}{n+1} \cdot \frac{j}{n} \cdot \frac{j-1}{n-1}}{\frac{1}{2}} &= 2\frac{\sum_{j=1}^{n} j^2 - \sum_{j=1}^{n} j}{n(n+1)(n-1)} = \frac{2}{3} \end{aligned}$$

כאשר המונה חושב באמצעות נוסחת ההסתברות השלמה ושני הטורים חושבו על ידי הנוסחאות הידועות של סכום סדרת הטבעיים (נוסחא 0.20) וסכום סדרת הריבועים (נוסחא 0.21).

3. לפי הסעיף הקודם, $P(W_1 \cap W_2) = \frac{1}{3}$. נשתמש באותה השיטה ונקבל:

$$Pr\left(W_3|W_1 \cap W_2\right) = \frac{Pr\left(W_1 \cap W_2 \cap W_3\right)}{Pr\left(W_1 \cap W_2\right)} = \frac{\sum_{j=1}^{n} Pr\left(W_1 \cap W_2 \cap W_3|A_j\right) Pr\left(A_j\right)}{\frac{1}{3}}$$
$$= 3\sum_{j=1}^{n} \frac{1}{n+1}\frac{j(j-1)(j-2)}{n(n-1)(n-2)} = \frac{3}{4}$$

כאשר הפעם בחישוב משתמשים גם בסכום הידוע של סדרת החזקות השלישיות (נוסחא 0.22).

□

תרגיל 3.31. נסמן ב-A_k את המאורע שאבי הטיל את הקוביה k פעמים עד שיצא עץ לראשונה. נסמן ב-B_k את המאורע שטל הטילה את המטבע k פעמים. המאורע שמספר ההטלות של שניהם זהה הוא איחוד של המאורעות הזרים בזוגות "שניהם הטילו פעם אחת", "שניהם הטילו פעמיים" וכו'. לכן ההסתברות המבוקשת הינה:

$$Pr\left(\bigcup_{k=1}^{\infty} A_k \cap B_k\right) = \sum_{k=1}^{\infty} Pr\left(A_k \cap B_k\right) = \sum_{k=1}^{\infty} Pr\left(A_k\right) \cdot Pr\left(B_k\right)$$

כאשר במעבר האחרון הסתמכנו על העובדה שהקוביה של אבי בלתי תלויה במטבע של טל. ההסתברות שאבי יטיל את הקוביה בדיוק k פעמים הינה $\left(\frac{5}{6}\right)^{k-1}\frac{1}{6}$, שכן מאורע זה מורכב מ-$k-1$ הטלות שלא יצאו 6 וההטלה האחרונה כן יצאה 6. באותה צורה, ההסתברות שטל הטילה את הקוביה בדיוק k פעמים הינה $\left(\frac{1}{2}\right)^{k-1}\frac{1}{2}$. נציב מספרים אלו:

$$\sum_{k=1}^{\infty}\left(\frac{5}{6}\right)^{k-1}\frac{1}{6}\cdot\left(\frac{1}{2}\right)^{k-1}\frac{1}{2} = \frac{1}{12}\sum_{k=1}^{\infty}\left(\frac{5}{12}\right)^{k-1} = \frac{1}{12}\sum_{k=0}^{\infty}\left(\frac{5}{12}\right)^{k}$$

זהו טור גאומטרי מתכנס (נוסחא 0.15) וערכו $\frac{12}{7} = \frac{1}{1-\frac{5}{12}}$. לכן ההסתברות המבוקשת הינה $\frac{1}{7} = \frac{12}{7} \cdot \frac{1}{12}$. □

4 משתנים מקריים

הדיון במאורעות שירת אותנו נאמנה עד כה, אך יש בו מין הסירבול שכן אנחנו עוסקים במאורעות ופעולות בין מאורעות גם כאשר מה שמעניין אותנו הוא ערך מספרי מסויים שיכול להתקבל בניסוי. למשל, כאשר מתקשרים למוקד תמיכה של חברה מסויימת, רוצים לדעת הסתברויות הקשורות לזמן ההמתנה – מה הסיכוי שנמתין פחות מ-3 דקות, מה הסיכוי שנמתין יותר מ-10 דקות, מה הסיכוי שזמן ההמתנה יהיה ארוך מהממוצע וכד׳. אין זה נוח להגדיר כל אחד מהמאורעות הללו בדרך הרגילה (בתור קבוצה) שכן אנחנו מתעניינים בערך מספרי מסויים, זמן ההמתנה, ובשוויונות ואי-שוויונות שהוא יכול לקיים. במקום זאת נגדיר משתנה המתאר את זמן ההמתנה ונדון בהסתברויות הקשורות למשתנה זה. מאחר והמשתנה אינו דטרמיניסטי אלא מקבל ערכים באקראי לפי איזושהי התפלגות. משתנה כזה נקרא משתנה מקרי.

מבחינה רעיונית, לפחות בהתחלה, לא יהיה הבדל גדול ביחס לפרקים הקודמים והשינוי יהיה סמנטי בלבד. במקום להסתכל על מאורעות כעל קבוצות עם שמות סתמיים, נגדיר משתנה מקרי X ונתבונן במאורעות מהצורה $\{X=5\},\{X\leq 0\},\{X>3\}$ וכדומה. לפיכך, התרגילים והדוגמאות בהם נפתח את הפרק לא יהיו שונים מהפרק הקודם. ההבדל יהיה טמון באופן בו אנחנו יוצג התרגיל והטרמינולוגיה. בהמשך נראה שבאמצעות מעבר למשתנים מקרים אנחנו יכולים לבצע אנליזה מורכבת יותר של תוצאות של ניסויים אקראיים ולדבר על ערך ממוצע שמתקבל בניסוי, פיזור סביב הממוצע ועוד. מבחינה פורמלית, אנחנו נגדיר משתנה מקרי בתור פונקציה הממפה תוצאות אפשריות של הניסוי למספרים ממשיים.

הגדרה 4.1 (משתנה מקרי). יהי Ω מרחב מדגם כלשהו. משתנה מקרי הוא פונקציה המתאימה לכל תוצאה אפשרית במרחב המדגם ערך ממשי: $X:\Omega\to\mathbb{R}$.

חשוב לשים לב שההגדרה לא דורשת שהפונקציה תהיה חד-חד ערכית, כלומר תוצאות שונות של מרחב המדגם יכולות להיות משוייכות לערך בודד. למשל, אם הניסוי הוא הטלה של קוביה, ניתן להגדיר את המשתנה המקרי X להיות 1 כאשר התוצאה זוגית ו-0 אחרת. במקרה זה, אם תתקבל אחת מ-3 התוצאות $\{1,3,5\}$ הערך של X יהיה 0. כל שחשוב הוא שלכל תוצאה אפשרית בניסוי ייקבע ערכו של X.

הגדרה 4.2 (תומך). יהי Ω מרחב מדגם כלשהו ו-X משתנה מקרי המוגדר עליו. התומך של X, המסומן על ידי $supp(X)$, הינו קבוצת כל הערכים ש-X יכול לקבל בניסוי, כלומר $X(\Omega)$.

בדוגמת הקוביה, התומך של X הינו הקבוצה $\{0,1\}$. כאשר הניסוי הוא המתנה למענה של נציג במוקד תמיכה והמשתנה המקרי הוא זמן ההמתנה, קבוצת הערכים יכולה להיות כל מספר גדול מ-0 ולכן התומך יהיה $[0,\infty)$. מבחינה רעיונית והסתברותית אין הבדל בין דוגמאות אלו, אך מבחינה מתמטית ופרקטית יש. הראשון הוא משתנה מקרי המוגדר על קבוצה סופית והשני הוא משתנה מקרי המוגדר על קבוצה אינסופית מעוצמת הרצף, לכן הכלים המתמטיים בהם נשתמש בחישובים של כל אחד מהם יהיו כלים שונים. משתנים מקריים שהתומך שלהם הוא קבוצה סופית או קבוצה בת-מנייה (בעיקר $\mathbb{N}$) נקראים משתנים מקריים בדידים ומטופלים בפרק 4.1. משתנים מקריים שהתומך שלהם הוא קטע, קרן או כל המספרים הממשיים $\mathbb{R}$ נקראים משתנים מקריים רציפים ומטופלים בפרק 4.2.

כאשר מבצעים את הניסוי מתקבלת תוצאה כלשהי במרחב המדגם ובעקבות זאת נבחר ערך עבור המשתנה המקרי X מתוך התומך שלו. לכן, ניתן לראות בתומך כמעין מרחב מדגם חדש מתוכו נבחר ערך מספרי לפי איזושהי

הסתברות בהתאם למרחב המדגם המקורי. כך למשל עשינו בתרגילים 2.1 ו-2.2, שם הסתכלנו על הערכים שמתקבלים בניסוי וההסתברויות לקבלם, מבלי לציין מפורשות מהו הניסוי המבוצע. יכולת הפשטה זו הינה יתרון נוסף של משתנים מקריים – ניתן לקשור בין בעיות שונות על ידי המרתן למשתנים מקריים דומים. למשל, נניח והניסוי הוא הטלת מטבע. אם נגדיר משתנה מקרי Y המקבל את הערכים $Y = 0$ כאשר התקבל עץ ו-$Y = 1$ כאשר התקבל פלי, נקבל משתנה מקרי זהה ל-X מדוגמת הקוביה: שני המשתנים המקריים מקבלים בדיוק את אותם ערכים, והסיכוי לכל ערך הוא אותו סיכוי – $\frac{1}{2}$. לכן, לא מדובר על בעיות שונות ואם נבצע את החישובים הדרושים לנו עבור X, התוצאות יהיו תקפות גם עבור Y. בהמשך נראה שישנן משפחות של משתנים מקריים המתארות בעיות נפוצות בחיים, עבורן בוצעו כבר כל החישובים ולנו יוותר רק להתאים בין הסיפור לבין המשתנה המקרי המתאים לו, מבלי שהמתמטיקה עליה מבוססת הבעיה תשתנה.

תרגיל 4.1. מהם התומכים של המשתנים המקריים הבאים? האם זהו משתנה מקרי רציף או בדיד?

1. מטילים קוביה שוב ושוב. X – מספר ההטלות עד שהתקבל 4 לראשונה.
2. בוחרים באקראי אדם באוכלוסיה. X - גובהו.
3. יורים לעבר מטרה. X – מרחב הפגיעה ממרכז המטרה.
4. מתוך קבוצה של 10 בנים ו-4 בנות בוחרים באקראי 5 ילדים (ללא החזרה). X – מספר הבנות שנבחרו.
5. מטילים מטבע שוב ושוב. X – מספר ההטלות עד שהתקבלו שתי התוצאות האפשריות.
6. שיכור עומד בנקודה בנקודה 0 על ציר ה-x וצועד 10 צעדים בגודל מטר ימינה (כיוון חיובי) ושמאלה (באקראי). X – מיקום השיכור ביחס לנקודת המוצא לאחר 10 הצעדים.

4.1 משתנים מקריים בדידים

כאמור, משתנים מקריים בדידים הם אלו שהתומך שלהם הוא סופי או בן מנייה. כל אחד מהערכים בתומך יכול להתקבל בהסתברות כלשהי ואנחנו מעוניינים לדעת את ההסתברויות הללו. הפונקציה המתאימה לכל ערך את הסיכוי לקבלו בניסוי נקראת פונקציית ההסתברות של המשתנה המקרי.

הגדרה 4.3 (פונקציית הסתברות)**.** יהי X משתנה מקרי בדיד. פונקציית ההסתברות של X היא פונקציה המתאימה לכל ערך בתומך את הסיכוי לקבל ערך כזה, כלומר

$$P_X(k) = Pr\left(X = k\right) \tag{4.1}$$

כאשר $k \in supp(X)$ ו-$P_X(k) = 0$ כאשר $k \notin supp(X)$.

ברגע שנקבעה פונקציית ההסתברות של המשתנה המקרי, אנחנו יודעים את כל הערכים שהמשתנה המקרי מקבל, את ההסתברויות בהם הוא מקבל אותם ולמעשה אנחנו יודעים עליו הכל ויכולים לחשב את הסיכוי של כל מאורע אפשרי בניסוי. ידע זה מכונה ההתפלגות של X וזו תהיה המטרה המרכזית שלנו בסעיף זה – לחשב את ההתפלגויות של משתנים מקריים. כאשר מדובר על משתנים מקריים בדידים, חישוב ההתפלגות יהיה למעשה אוסף של בעיות קומבינטוריות (אחת לכל ערך בתומך) ולכן השלבים בהם ננקוט יהיו:

1. הבנת הגדרת המשתנה המקרי וזיהוי התומך.
2. זיהוי מרחב המדגם הקומבינטורי המתאים לבעיה. לרוב נשתמש באותו מרחב מדגם לכל ערכי התומך אך אין זה הכרחי.

3. חישוב ההסתברות של המאורע $\{X=k\}$ לכל ערך $k \in supp(X)$. זהו ערך פונקציית ההסתברות בנקודה זו.

בעיה 4.1. 3 גברים ו-3 נשים מתיישבים באקראי על 6 כסאות הממוספרים מ-1 עד 6. נסמן ב-X את מספר הכסא הקטן ביותר עליו יושבת אישה. כיצד X מתפלג?

הוכחה. נתחיל בחישוב התומך. כאשר יושבת אישה בכיסא מספר 1 אז $X=1$ וזה הערך המינימלי של המשתנה המקרי. כאשר הנשים יושבות בכסאות 4,5,6 אז $X=4$ וזהו גם הערך המרבי ש-X יכול לקבל. לכן $supp(X)=\{1,2,3,4\}$.

מאחר ולכל הערכים בתומך הבעיה הקומבינטורית היא בעיה זהה (סידור 6 אנשים ב-6 כסאות) ורק המאורע הוא המשתנה, ניתן לבחור מרחב מדגם אחד שיתאים לכל הערכים בתומך. מאחר וכל מה שמעניין הוא מספר הכיסא שעליו יושבת אישה ולא זהות האישה שיושבת עליו, ניתן לצמצם את מרחב המדגם ולהתייחס אליו בתור בחירה של 3 מקומות עבור הנשים מתוך 6 המקומות האפשריים במקום, למשל, סידור 6 אנשים בשורה. הבחירה נעשית ללא החזרה וללא סדר (כי כאמור, ברגע שאנחנו יודעים את מספרי הכסאות שנבחרו אנחנו יודעים את ערכו של X, אין צורך לבחור איזו אישה יושבת באיזה כיסא) ולכן גודלו של מרחב המדגם הינו $20=\binom{6}{3}$.

- $X=1$

כדי שהמאורע $X=1$ יקרה, חייבים לבחור את כיסא מספר 1 עבור אישה. מאחר ומרחב המדגם כולל בחירה של 3 כסאות בסך הכל, עלינו לבחור עוד 2 כסאות מתוך 5 הכסאות שנותרו ולכן ההסתברות תהיה:

$$Pr\left(X=1\right)=\frac{\binom{5}{2}}{20}=\frac{1}{2}$$

הערה: ניתן היה לחשב הסתברות זו גם משיקולי סימטריה – הסיכוי שבמקום מספר 1 יושבת אישה הוא חצי כי יש מספר שווה של גברים ונשים. זו דוגמא טובה לכך שלא תמיד חייבים להשתמש באותו מרחב המדגם שהגדרנו בתחילת השאלה בכדי לחשב את כל ערכי פונקציית ההסתברות.

- $X=2$

כדי שהמאורע $X=2$ יקרה, חייבים לבחור את כיסא מספר 2 עבור אישה ואסור לנו לבחור את כיסא מספר 1 (כי אם הוא היה נבחר, אז $X=1$ ולא $X=2$). לכן ל-2 הנשים הנותרות יש לבחור 2 כיסאות מתוך 4 הכיסאות המותרים וההסתברות לכך היא:

$$Pr\left(X=2\right)=\frac{\binom{4}{2}}{20}=\frac{3}{10}$$

- $X=3$

כדי שהמאורע $X=3$ יקרה, חייבים לבחור את כיסא מספר 3 עבור אישה ואסור לבחור את הכיסאות 1 ו-2. לכן ל-2 הנשים הנותרות יש לבחור 2 כיסאות מתוך 3 הכיסאות המותרים וההסתברות לכך היא:

$$Pr\left(X=3\right)=\frac{\binom{3}{2}}{20}=\frac{3}{20}$$

- $X=4$

כדי שהמאורע $X=4$ יקרה, חייבים לבחור את כיסא מספר 4 עבור אישה ולשאר הנשים אסור לבחור את כיסאות 1,2,3. לכן אין לנו למעשה אפשרויות בחירה נוספות, הנשים חייבות לשבת בכיסאות 4,5,6 ויש רק אפשרות בחירה אחת. ההסתברות המבוקשת היא אם כך:

$$Pr\left(X=4\right)=\frac{1}{20}$$

נסכם. פונקציית ההסתברות של המשתנה המקרי X, מספר הכיסא הקטן ביותר עליו יושבת אישה, הינה:

$$P_X(k) = \begin{cases} 0.5 & k=1 \\ 0.3 & k=2 \\ 0.15 & k=3 \\ 0.05 & k=4 \end{cases}$$

וכמובן 0 אחרת, אבל אין צורך לציין זאת מפורשות שכן זו תכונה כללית של פונקציית ההסתברות.
ניתן להציג את פונקציית ההסתברות בשתי דרכים נוספות: טבלה או נוסחה. טבלה היא דרך הצגה נוחה כאשר מדובר על תומך עם מעט ערכים:

k	$Pr\left(X=k\right)$
1	0.5
2	0.3
3	0.15
4	0.05

כאשר התומך גדול וכולל הרבה אפשרויות, לשרטט טבלה הכוללת את כל האופציות היא לא אפשרות פרקטית וצריך לנסות למצוא נוסחא כללית, המתארת את ההסתברות כפונקציה של k. במקרה דנן, שימו לב שלכל ערך של k, חייבים לבחור את כיסא מספר k ועבור שתי הנשים האחרות אסור לבחור את הכיסאות $1,2,\ldots,k-1$ ואי אפשר לבחור שוב את כיסא k, לכן עבורן יש לבחור 2 כיסאות מתוך $6-k$ הכסאות המותרים:

$$P_X(k) = Pr\left(X=k\right) = \frac{\binom{6-k}{2}}{20}$$

□

אם נסכם את כל הערכים בטבלה נקבל $P_X(1)+P_X(2)+P_X(3)+P_X(4)=1$. תוצאה זו אינה מפתיעה, שהרי כבר ציינו שבמובן מסויים, התומך של כל משתנה מקרי הוא מרחב המדגם החדש. בכל חזרה על הניסוי נבחר בדיוק ערך אחד עבור המשתנה המקרי מתוך התומך ולכן סכום הסיכויים לבחור ערך כלשהו מתוך התומך צריך להיות 1.

משפט 4.1. *יהי X משתנה מקרי בדיד. סכום ההסתברויות של כל הערכים בתומך הוא 1:*

$$\sum_{k\in supp(X)} P_X(k) = 1 \qquad (4.2)$$

תרגיל 4.2. הוכיחו את משפט 4.1.

למשפט 4.1 שני שימושים סותרים. השימוש העיקרי של המשפט הוא בבדיקת החישובים. לאחר שחושבה פונקציית ההסתברות לכל הערכים בתומך, ניתן לסכום את התוצאות ולבדוק האם אכן מתקבל 1. אם לא מתקבל 1, בוודאות משהו בחישובים לא בסדר והסכום שמתקבל מרמז גם על אופי השגיאה. סכומים גדולים מ-1 מרמזים על כך שביצענו ספירת יתר של חלק מהמצבים ואילו סכומים קטנים מ-1 מרמזים על כך שלא ספרנו את כל המצבים האפשריים או שטעינו בהגדרת התומך ולא כללנו בו ערכים אפשריים. חשוב להדגיש שגם אם נקבל 1 זה עדיין לא סימן לכך שהפתרון נכון, אך רמת הביטחון בכך בוודאי תעלה.

שימוש אפשרי נוסף במשפט 4.1 הוא לצורך חישוב הסתברויות של מאורעות שקשה לחשב ישירות. למשל, אם ישנו ערך בתומך שמתקבל בהרבה צורות או שהחישוב הקומבינטורי שלו מסובך, כל עוד יש רק ערך אחד כזה, ניתן לוותר על חישוב ישיר שלו ולחשב את ההסתברות שלו על ידי המשלים. למשל, ניתן היה לוותר על החישוב הקומבינטורי האחרון ולקבל את $P_X(4)$ על ידי

$$P_X(4) = 1 - Pr\left(X \neq 4\right) = 1 - \left(P_X(1) + P_X(2) + P_X(3)\right)$$

באופן זה אנחנו מפשטים את החישובים אבל מאבדים את יכולת הבדיקה. לכן כדאי להשתמש באפשרות זו רק אם חישוב ישיר של ההסתברות הוא אכן מסובך מדי (או שהזמן במבחן עומד להיגמר...).

בעיה 4.2. מטילים קוביה שוב ושוב. כיצד מתפלג מספר ההטלות עד הפעם הראשונה בה התקבל 6 אם ידוע שבהטלה ה-5 התקבל 2 לראשונה?

הוכחה. המשתנה המקרי אינו מוגדר באופן מפורש בשאלה ולכן נגדירו כחלק מהפתרון: X – מספר ההטלות עד הפעם הראשונה בה התקבל 6 אם ידוע שבהטלה ה-5 התקבל 2 לראשונה. כמקודם, נתחיל מחישוב התומך. ברמת העקרון, מספר ההטלות עד ההופעה הראשונה של 6 יכול להיות כל מספר טבעי, שכן בהחלט ייתכן שנטיל קוביה מיליוני פעמים מבלי לקבל את הספרה 6 (אם כי הסיכוי לכך כמובן נמוך). יחד עם זאת, אנחנו יודעים בוודאות ש-6 לא יצא בהטלה ה-5 שכן בהטלה זו התקבל 2, לכן מרחב המדגם הינו $supp(X) = \mathbb{N} \setminus \{5\}$. היות והתומך אינסופי, אין אפשרות לעבור אפשרות-אפשרות ולחשב את ההסתברויות באופן זה או להציג את התוצאות בטבלה, ויהא עלינו לפתח נוסחא כללית. נתחיל בחישוב עבור מספר ערכים בודדים בכדי לראות את הכיוון הכללי והחוקיות. נבצע זאת עבור שתי תוצאות: $X = 2$ ו-$X = 6$ ולאחר מכן נכליל.

- $X = 2$

  בכדי שמאורע זה יקרה, צריך שבהטלה הראשונה לא יצא 6 ובהטלה השנייה כן ייצא 6. אנחנו יודעים שבשתי הטלות אלו בוודאות לא יצא 2 (הוא התקבל לראשונה רק בהטלה ה-5) ולכן מספר האפשרויות שיכולות לצאת הוא 5 (הכל פרט ל-2). לכן הסיכוי שבהטלה הראשונה לא יצא 6 הוא $\frac{4}{5}$ והסיכוי שבהטלה השנייה כן יצא 6 הוא $\frac{1}{5}$. ההטלות בלתי תלויות ולכן

$$Pr\left(X = 2\right) = \frac{4}{5} \cdot \frac{1}{5} = \frac{4}{25}$$

- $X = 6$

  בכדי שמאורע זה יקרה, צריך שב-4 ההטלות הראשונות לא יצא 6 ובהטלה השישית ייצא 6. כמקודם, ב-4 ההטלות הראשונות יכולות לצאת רק 5 תוצאות אפשריות ולכן הסיכוי שבכולם לא יצא 6 הוא $\left(\frac{4}{5}\right)^4$. בהטלה החמישית בוודאות מתקבל 2 (זה נתון) ובהטלה השישית כבר יכולות לצאת כל התוצאות כולל 2, לכן הסיכוי שיצא 6 הוא $\frac{1}{6}$. בסך הכל ההסתברות המבוקשת הינה

$$Pr\left(X = 6\right) = \left(\frac{4}{5}\right)^4 \cdot \frac{1}{6} = \frac{128}{1875}$$

נכליל. אנחנו רואים שיש שוני בין ההטלות לפני ההטלה החמישית להטלות אחריה. השוני הוא שב-4 ההטלות הראשונות יש רק 5 תוצאות אפשריות בכל הטלה ואילו החל מההטלה השישית יש כבר 6 תוצאות אפשריות. לכן עבור $k \leq 4$ ההסתברות היא:

$$Pr\left(X = k\right) = \left(\frac{4}{5}\right)^{k-1} \cdot \frac{1}{5}$$

ועבור $k \geq 6$ ההסתברות היא

$$Pr\left(X=k\right)=\left(\frac{4}{5}\right)^{4}\cdot\left(\frac{5}{6}\right)^{k-6}\cdot\frac{1}{6}$$

לפיכך פונקציית ההסתברות הינה

$$P_X(k)=\begin{cases}\left(\frac{4}{5}\right)^{k-1}\cdot\frac{1}{5} & k=1,2,3,4\\ \left(\frac{4}{5}\right)^{4}\cdot\left(\frac{5}{6}\right)^{k-6}\cdot\frac{1}{6} & k=6,7,8,\ldots\end{cases} \tag{4.3}$$

□

תרגיל 4.3. ודאו שסכום ההסתברויות של ההתפלגות הנתונה בנוסחא 4.3 הוא אכן 1.

תרגיל 4.4. מטילים 2 קוביות. נסמן ב-X את סכום התוצאות שהתקבלו וב-Y את הפרשן בערכו המוחלט. מצאו את ההתפלגות של X ו-Y.

תרגיל 4.5. שחקן המשחק במשחק מזל מחליט להפסיק כשיפסיד לראשונה או לאחר 10 סיבובים. הסיבובים בלתי תלויים וההסתברות לנצח בסיבוב בודד היא p. יהי X מספר הסיבובים הכולל.

1. כיצד מתפלג X?
2. עבור $p=\frac{1}{2}$ חשבו את ההסתברות שיפסיק לשחק לפני הסיבוב העשירי.

פונקציית ההסתברות של משתנה מקרי מחשבת את ההסתברויות של מאורעות המוגדרים דרך שוויון, כלומר מאורעות מהצורה $\{X=k\}$. לצורך חישוב הסתברויות של אי-שוויונים נגדיר פונקציה נוספת, פונקציית ההתפלגות המצטברת, שערכיה יתאימו להסתברויות של המאורעות $\{X\leq k\}$.

הגדרה 4.4 (פונקציית התפלגות מצטברת (פה"מ)). פונקציית ההתפלגות המצטברת של המשתנה המקרי X היא פונקציה $F_X(t):\mathbb{R}\to\mathbb{R}$ המוגדרת לכל t ממשי על ידי:

$$F_X(t)=Pr\left(X\leq t\right) \tag{4.4}$$

שימו לב שבניגוד לפונקציית הסתברות, פונקציית ההתפלגות המצטברת מוגדרת לכל מספר ממשי ולא רק לערכים בתומך של המשתנה המקרי. עבור משתנים מקריים בדידים נשתמש בפונקציית ההתפלגות המצטברת כשלב ביניים בדרך לחישוב פונקציית ההסתברות, במקרים בהם חישוב ההסתברות של אי-השוויון פשוטה יותר מאשר חישוב ההסתברות של השיוויון. פרט לכך, צורתה של פונקציית ההתפלגות המצטברת היא מסורבלת מדי עבור משתנים מקריים בדידים (כפי שיודגם בהמשך) ועיקר השימוש בה יהיה עבור משתנים מקריים רציפים. יחד עם זאת, כבר עכשיו ניתן להוכיח מספר תכונות של פונקציית ההתפלגות המצטברת הנובעות מעצם הגדרתה וכלל לא תלויות במשתנה המקרי ובהיותו בדיד או רציף.

משפט 4.2. *יהי X משתנה מקרי. פונקציית ההתפלגות המצטברת $F_X(t)$ מקיימת את התכונות הבאות:*

1. *הפונקציה עולה חלש, כלומר אם $a<b$ אז $F_X(a)\leq F_X(b)$.*
2. *הגבול באינסוף הוא 1: $\lim_{t\to\infty}F_X(t)=1$.*
3. *הגבול במינוס אינסוף הוא 0: $\lim_{t\to-\infty}F_X(t)=0$.*

תרגיל 4.6. הוכיחו את משפט 4.2.

בעיה 4.3. מטילים קוביה 3 פעמים. כיצד מתפלג X, המספר המקסימלי שהתקבלת בשלושת ההטלות?

מציאת פונקציית ההסתברות. המספר המקסימלי הוא בוודאות מספר הרשום על הקוביות ולכן התומך הוא $supp(X) = \{1, 2, \ldots, 6\}$. מאחר ומרחב המדגם כולל בחירה של מספר אחד מבין 6 המספרים בקוביה עם החזרה ועם סדר (ההטלה הראשונה, ההטלה השנייה וההטלה השלישית) הרי שמרחב המדגם הינו 6^3.
ננסה לחשב מפורשות פונקציית ההסתברות עבור אחד הערכים בתומך, למשל את $X = 3$. בכדי שמאורע זה יתרחש המספר 3 צריך להתקבל לפחות פעם אחת ובנוסף המספרים האחרים חייבים להיות 1 ו-2. לכן ישנן 3 אפשרויות:

- התוצאה 3 יצאה 3 פעמים ויש אפשרות אחת כזו.
- התוצאה 3 יצאה פעמיים ובהטלה הנוספת יצא 1 או 2. יש 2 אפשרויות לתוצאת ההטלה הנוספת ויש $\binom{3}{2}$ אפשרויות לבחור אילו מ-3 ההטלות נפלו על 3.
- התוצאה 3 יצאה פעם אחת ובהטלות הנוספות יצאו 1 או 2. יש 2 אפשרויות לכל הטלה נוספת ויש $\binom{3}{1}$ אפשרויות לבחור איזו מ-3 ההטלות היא ההטלה שבה יצא 3.

לכן ההסתברות הכוללת היא:

$$Pr\left(X=3\right) = \frac{1+\binom{3}{2}\cdot 2+\binom{3}{2}\cdot 2^2}{6^3} = \frac{19}{216}$$

ניתן לחזור על חישוב זה גם עבור ערכים אחרים אך החישוב הוא מסורבל מאוד, כולל אפשרויות רבות ולא ברור כיצד להכליל אותו למקרה הכללי ולקבל נוסחא. המוקש בחישוב טמון בכך שכאשר מחשבים את הסתברות המאורע $\{X=3\}$, מוכרחים ש-3 יצא לפחות באחת ההטלות ולכן צריך לספור כמה פעמים הוא יצא, באילו מההטלות הוא התקבל ומה התקבל בשאר ההטלות. במקום זאת נחשב את הפה"מ שכן הפה"מ מחושב דרך המאורע $\{X \leq 3\}$ הכולל פחות אילוצים ולכן פשוט יותר. על מנת שהמקסימום יהיה קטן משלוש, כל שנדרש הוא שבכל הטלה יתקבל מספר קטן מ-3. אין דרישה ש-3 אכן יתקבל או כל דרישה אחרת, רק שהבחירה בכל אחת מההטלות של הקוביה תהיה מהמספרים 1,2,3. לכן ההסתברות שהמקסימום יהיה קטן מ-3 הינה $\frac{3^3}{6^3}$ובאופן כללי, ההסתברות שהמקסימום יהיה קטן מ-k עבור $k \in supp(X)$ היא

$$Pr\left(X \leq k\right) = \frac{k^3}{6^3} \tag{4.5}$$

שכן במונה מתואר המאורע בו בכל הטלה בחרנו מספרים מ1- עד k. נוסחא 4.5 היא כבר צעד משמעותי בכיוון הנכון שכן זוהי נוסחא כללית ופשוטה יחסית, אך זו אינה פונקציית ההסתברות המוגדרת עבור המאורע $\{X=k\}$. לשם חישובו נשים לב כי את אי השוויון $\{X \leq k\}$ ניתן לרשום כאיחוד של שני מאורעות זרים:

$$\{X \leq k\} = \{X = k\} \cup \{X < k\}$$

ולכן

$$Pr\left(X \leq k\right) = Pr\left(X = k\right) + Pr\left(X < k\right)$$

בנוסף, זהו משתנה מקרי בדיד עם תומך הגדל בקפיצות של 1, לכן הוא לא יכול לקבל אף ערך בתחום $(k-1,k)$. כך שהמאורעות $\{X<k\}$ ו-$\{X\leq k-1\}$ זהים. למשל, אם המקסימום קטן ממש מ-4 אזי הוא 1, 2 או 3, כלומר המקסימום קטן או שווה ל-3. נציב זאת בנוסחא האחרונה ויחד עם נוסחא 4.5 נקבל שפונקציית ההסתברות היא

$$P_X(k)=Pr\left(X\leq k\right)-Pr\left(X\leq k-1\right)=\frac{k^3-(k-1)^3}{6^3}$$

□

ראוי לשים לב לנקודה חשובה. נוסחא 4.5 איננה פונקציית ההתפלגות המצטברת של המשתנה המקרי X מאחר והיא מוגדרת רק על ערכי התומך ולא עבור כל ערך ממשי.

מציאת פונקציית ההתפלגות המצטברת. נחשב את פונקציית ההתפלגות המצטברת לכל ערך t החל מערכים קטנים מאוד וכלה בערכים גדולים. למשל, $F_X(-4)=Pr\left(X\leq -4\right)=0$ שכן בוודאות X גדול או שווה ל-1 ולא ייתכן שהוא יהיה קטן או שווה ל--4. באותו אופן, $F_X(t)=0$ לכל $t<1$.

בנקודה $t=1$ ערך הפונקציה משתנה, לאור החישוב לעיל: $F_X(1)=Pr\left(X\leq 1\right)=\frac{1^3}{6^3}$. נמשיך לערכים גדולים יותר. עבור $t=1.5$ ניתן לפרק את המאורע $\{X\leq 1.5\}$ לאיחוד של שני מאורעות זרים $\{X\leq 1.5\}=\{X\leq 1\}\cup\{1<X\leq 1.5\}$. הואיל ו-X מקבל רק ערכים שלמים, לא ייתכן שהוא גדול ממש מ-1 אבל קטן מ-1.5 וההסתברות של מאורע זה היא אפס. לכן פונקציית ההתפלגות המצטברת בנקודה $t=1.5$ היא

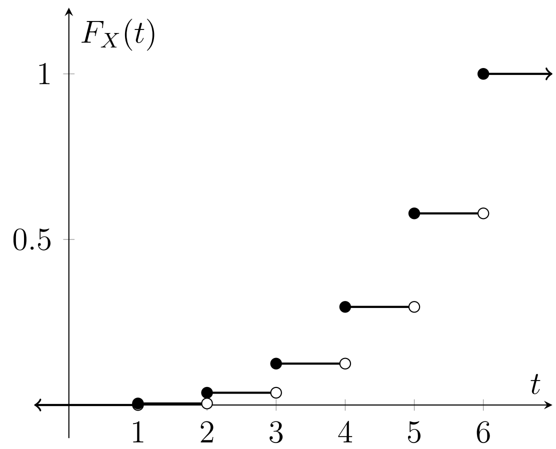
$$F_X(1.5)=Pr\left(X\leq 1.5\right)=Pr\left(X\leq 1\right)+Pr\left(1<X\leq 1.5\right)=Pr\left(X\leq 1\right)=\frac{1^3}{6^3}$$

הדבר נכון לכל $t\in[1,2)$ ולכן הפה"מ קבוע בתחום זה וערכו $\frac{1^3}{6^3}$. עבור $t=2$ מתרחשת קפיצה בפונקציה: $F_X(2)=Pr\left(X\leq 2\right)=\frac{2^3}{6^3}$. גודלה של הקפיצה היא בדיוק ההסתברות של המאורע שהוספנו לחישוב: $\{X=2\}$. משיקולים דומים, הפונקציה תהיה קבועה גם בתחום $[2,3)$ עם קפיצה עבור $t=3$ וכן הלאה – בין כל שני ערכים של התומך הפונקציה קבועה.

עבור $t\geq 6$, המאורע $\{X\leq t\}$ מתקיים תמיד, שכן X תמיד קטן או שווה ל-6 ולכן בוודאי קטן או שווה ממספר הגדול מ-6. לכן החל מ-6 ועד לאינסוף הפונקציה קבועה על הערך 1 (כפי שאכן משפט 4.2 מנבא).

לסיכום, פונקציית ההתפלגות המצטברת של X היא פונקציה קבועה למקוטעין המתוארת על ידי הנוסחא הבאה:

$$F_X(t)=\begin{cases}0 & t<1\\ 1^3/6^3 & 1\leq t<2\\ 2^3/6^3 & 2\leq t<3\\ 3^3/6^3 & 3\leq t<4\\ 4^3/6^3 & 4\leq t<5\\ 5^3/6^3 & 5\leq t<6\\ 1 & 6\leq t\end{cases}\qquad (4.6)$$



הקפיצות מתקבלות בדיוק בערכי התומך וגודלן הוא בדיוק פונקציית ההסתברות של X בנקודה זו.
מתוך גרף זה ניתן להבין את המקור לשמה של פונקציית ההתפלגות המצטברת: הפונקציה הולכת מ-$-\infty$ אל $+\infty$ ובדרך "צוברת" את כל ההסתברויות של הערכים בהם היא עוברה עד כה. □

דוגמא זו ממחישה את חוסר שימושיותה של פונקציית ההתפלגות המצטברת עבור משתנים מקריים בדידים. פונקציה זו תמיד תהיה פונקציה קבועה למקוטעין, עם קפיצות בערכי התומך ונוסחא מסורבלת הכוללת הרבה תחומים. אנחנו נראה בהמשך שעבור משתנים מקריים רציפים פונקציית ההתפלגות המצטברת היא פונקציה רציפה וגזירה ולכן גם שימושית.

תרגיל 4.7. בכד יש n כדורים הממוספרים מ-1 ועד n. מוציאים מהכד m כדורים. יהי X המספר המקסימלי שהתקבל. חשבו את התפלגותו של X אם:

1. ההוצאה היא ללא החזרה.
2. ההוצאה היא עם החזרה.

תרגיל 4.8. נתון לוח משובץ 3×3. צובעים באקראי 6 משבצות בצבע שחור. נסמן על ידי X את מספר השורות בהן יש בדיוק שתי משבצות שאינן צבועות. כיצד X מתפלג?

תרגיל 4.9. מחלקים באקראי 4 כדורים ל-5 תאים. כיצד מתפלג X, מספר התאים בהם יש בדיוק 2 כדורים?

4.2 משתנים מקריים רציפים

נעבור לדון במשתנים מקריים היכולים לקבל כל ערך בקטע מסויים או אפילו כל ערך ממשי, דוגמת זמן, משקל, גובה, מרחק מנקודה מסויימת וכד'. היינו שמחים להגדיר פונקציית הסתברות גם עבור משתנים מקריים אלו, אך ישנן שתי בעיות. ראשית, משתנים מקריים אלו יכולים לקבל "הרבה יותר"[1] ערכים מאשר משתנים מקריים בדידים (כולל אלו בעלי התומך האינסופי) ולכן לא ברור כיצד ניתן להגדיר את ההסתברות לכל נקודה $Pr\left(X=x\right)$ כך שסכום ההסתברויות יהיה 1 (משפט 4.1) ולמעשה כלל לא ברור כיצד מבצעים סכימה על קבוצה שאינה בת מנייה. שנית, ניתן להוכיח באמצעות כלים של חשבון אינפיניטסימלי שפונקציה המקיימת את תנאי משפט 4.2 חייבת להיות רציפה פרט, אולי, למספר בן מנייה של נקודות. היות וכל נקודת אי רציפות היא "קפיצה" בפונקציה המתאימה לערך המתקבל בהסתברות חיובית ויש רק מספר בן מנייה של קפיצות, נקבל שלא ניתן להגדיר פונקציית צפיפות כך ש-$P_X(x) \neq 0$ לכל x בקטע מסויים. לאור הדיון לעיל כלל לא ברור שהמאורע $\{X=x\}$ מוגדר

[1] במובן של עוצמה גדולה יותר

(וגם אם כן, מה המשמעות של הסתברות 0 למאורע שיכול לקרות?). לפיכך במקום פונקציית הסתברות נגדיר משתנה מקרי רציף באמצעות פונקציית ההתפלגות המצטברת.

הגדרה 4.5 (משתנה מקרי רציף)**.** משתנה מקרי X אשר פונקציית ההתפלגות המצטברת שלו, $F_X(t) = Pr\left(X \leq t\right)$, היא פונקציה רציפה בכל הישר הממשי $\mathbb{R}$ נקרא משתנה מקרי רציף. במקרה כזה, אין "קפיצות" בפונקציית ההתפלגות המצטברת ולכן לכל $x \in \mathbb{R}$ מתקיים $Pr\left(X = x\right) = 0$.

כאשר עוסקים במשתנים מקריים רציפים, אין משמעות לנקודות בודדות או אפילו לקבוצה בת מנייה של נקודות. הסיכוי לקבל נקודה בודדת הוא תמיד 0. המאורעות שיעניינו אותנו אם כך יהיו מאורעות המוגדרים על ידי אי שוויון, למשל $\{X > 3\}$, $\{X \leq 5\}$ או $\{2 \leq X \leq 4\}$. היות ותופעות טבע הן לרוב בדידות לחלוטין (כמו הטלה של קוביה) או רציפות לחלוטין (כמו קליעה למטרה ומדידת המרחק ממרכזה) גם מרבית המשתנים המקריים יהיו בדידים או רציפים. בפרק 4.3 נדון בסוג השלישי של משתנים מקריים, משתנים מקריים מעורבים, הכוללים גם מרכיב בדיד וגם מרכיב רציף, אך אלו פחות שימושיים.

תרגיל 4.10. יהי X משתנה מקרי רציף בעל פונקציית התפלגות מצטברת

$$F_X\left(t\right) = \begin{cases} 0 & t < 0 \\ at + b & 0 \leq t \leq 2 \\ ct^2 & 2 < t \leq 5 \\ d & t > 5 \end{cases}$$

כאשר a, b, c, d קבועים.

1. העזרו במשפט 4.2 ובתכונת הרציפות של הפה"מ וחשבו את a, b, c, d.
2. מהי ההסתברות של שלושת המאורעות $\{X > 3\}$, $\{X \leq 5\}$ ו-$\{2 \leq X \leq 4\}$.

הטכניקה שהודגמה בתרגיל הקודם לצורך חישוב הסתברויות של אי-שוויונים חשובה מספיק על מנת שנכליל אותה ונסכם אותה במשפט.

***משפט* 4.3** (חישובי הסתברויות באמצעות פה"מ)**.** *יהי X משתנה מקרי רציף בעל פונקציית התפלגות מצטברת $F_X\left(t\right)$. אזי לכל $a, b \in \mathbb{R}$ מתקיים:*

$$Pr\left(X \leq a\right) = Pr\left(X < a\right) = F_X(a) \tag{4.7}$$

$$Pr\left(X \geq a\right) = Pr\left(X > a\right) = 1 - F_X(a) \tag{4.8}$$

$$Pr\left(a \leq X \leq b\right) = F_X(b) - F_X(a) \tag{4.9}$$

כאשר המשוואה האחרונה תקפה גם כאשר אחד או שני אי-השוויונות הוא אי-שוויון חזק.

הוכחה. ההוכחה זהה לאופן בו בוצעו החישובים בתרגיל 4.10. □

הקצב בו פונקציית ההתפלגות המצטברת משתנה מתאר את שכיות התוצאות. תחומים בהם הפה"מ עולה מהר הם תחומים בעלי סיכוי גבוה יותר להתקבל מאשר תחומים בהם הפה"מ עולה לאט יותר. במשתנים מקריים בדידים אינטואיציה זו ניתנת להסבר אינטואיטיבי פשוט יותר - ככל שהעליה גדולה יותר, קרי ככל שהקפיצה בפה"מ גדולה יותר, כך הסיכוי לקבל את הערך הזה בניסוי גדול יותר. בפה"מ של משתנה מקרי רציף אין קפיצות וקצב השינוי של הפה"מ נקבע על ידי הנגזרת. לפיכך, נגדיר פונקציה חדשה, פונקציית צפיפות ההסתברות, שתהיה הנגזרת של פונקציית ההתפלגות המצטברת ותתאר, במובן מסויים, את השכיחות של כל נקודה.

הגדרה 4.6 (פונקציית צפיפות הסתברות)**.** יהי X משתנה מקרי רציף בעל פונקציית התפלגות מצטברת $F_X(t)$. פונקציית צפיפות ההסתברות של X מוגדרת על ידי

$$f_X(t) = \frac{dF_X(t)}{dt} \tag{4.10}$$

כלומר הצפיפות היא הנגזרת[2] של הפה"מ והפה"מ הינה הפונקציה הקדומה של הצפיפות:

$$F_X(t) = \int_{-\infty}^{t} f_X(u)\, du \tag{4.11}$$

תרגיל 4.11. חשבו את פונקציית הצפיפות של המשתנה המקרי X המוגדר בתרגיל **4.10**.

ייתכן ונתקלתם במושג הצפיפות בקורסים אחרים, דוגמת "צפיפות מסה" במכניקה או "צפיפות מטען" בחשמל. הרעיון העומד מאחורי כל הצפיפויות הוא דומה ונמחיש את האינטואיציה העומדת מאחורי צפיפות הסתברות על ידי שימוש בצפיפות מסה. נסתכל על מקל באורך 4 מטרים במסה של M ק"ג. המסה מפולגת על פני המקל לפי איזושהי פונקציית צפיפות: $\rho\,(x)$ הנתונה על ידי יחידות של גרם למטר $\left[\frac{g}{m}\right]$. ככל שהצפיפות גבוהה יותר בנקודה, כך יש יותר מסה באזור של הנקודה וככל שהצפיפות נמוכה יותר, כך אזור זה של המקל הוא פחות דחוס ויותר קל. היות ומסה היא גודל חיובי, פונקציית הצפיפות חייבת להיות חיובית $\rho\,(x) \geq 0$. אם ניקח חתיכת מקל באורך Δx קבוע מאזורים שונים, חתיכת מקל שנלקחה מאזור עם צפיפות גבוהה תשקול יותר מאשר חתיכת מקל שנלקחה מאזור עם צפיפות נמוכה. כדי לדעת כמה מסה יש בין שתי נקודות, כלומר מה יהיה משקל המקל שיתקבל על ידי חיתוך המקל המקורי בנקודות $0 < a < b < 4$ עלינו לחשב את האינטגרל $\int_a^b \rho\,(x)\ dx$. סך כל המסה, כאמור היא $\int_0^4 \rho\,(x)\ dx = M$.

נחזור להסתברות. נסתכל על משתנה מקרי X המוגדר על כל המספרים הממשיים. ההסתברות מפולגת על פני הממשיים לפי איזושהי פונקציית צפיפות: $f_X(x)$ הנתונה על ידי יחידות של הסתברות ליחידת אורך. ככל שהצפיפות ההסתברות גבוהה יותר בנקודה, כך יש יותר סיכוי שבניסוי מקרי תצא תוצאה בסביבת הנקודה וככל שהצפיפות נמוכה יותר, כך אזור זה של הממשיים מתקבל בהסתברות נמוכה יותר. היות והסתברות היא גודל חיובי, פונקציית הצפיפות חייבת להיות חיובית $f_X\,(x) \geq 0$. אם נסתכל על תחום באורך Δx קבוע מאזורים שונים של הציר הממשי, תחום שבו הצפיפות גבוהה יתקבל בהסתברות גבוהה יותר מאשר תחום שבו הצפיפות נמוכה. כדי לדעת כמה הסתברות יש בין שתי נקודות, כלומר מה הסיכוי שבניסוי יתקבל איזשהו ערך בין הנקודות $a < b$ עלינו לחשב את האינטגרל $\int_a^b f_X\,(x)\ dx$. סך כל ההסתברות, כאמור היא $\int_{-\infty}^{\infty} f_X\,(x)\ dx = 1$.
לאור האמור בפסקה זו נקבל את שתי התכונות של פונקציית צפיפות ההסתברות:

***משפט* 4.4.** *יהי X משתנה מקרי רציף ותהי $f_X(t)$ פונקציית צפיפות ההסתברות שלו. אזי $f_X(t)$ מקיימת את שתי התכונות הבאות:*

[2] לפי משפט לבג לפונקציות מונוטוניות, הפה"מ, בהיותה פונקציה עולה, היא גזירה, פרט, אולי, למספר בן מנייה של נקודות.

1. נרמול:

$$\int_{-\infty}^{\infty} f_X(t)\, dt = 1 \qquad (4.12)$$

2. חיוביות: לכל $t \in \mathbb{R}$ מתקיים $f_X(t) \geq 0$.

כל פונקציה המקיימת שתי תכונות אלו יכולה לשמש כפונקציית צפיפות הסתברות של משתנה מקרי רציף.

תרגיל 4.12. האם הפונקציות הבאות יכולות לשמש כפונקציות התפלגות מצטברת של משתנה מקרי רציף, כפונקציות צפיפות הסתברות או לא זה ולא זה?

1. $f(x) = \begin{cases} 4e^{-4(x-5)} & x \geq 5 \\ 0 & x < 5 \end{cases}$
2. $g(x) = \frac{1}{x^2}$
3. $h(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(x)$
4. $k(x) = x^3$
5. $l(x) = \begin{cases} \frac{x^n}{n!} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$
6. $m(x) = \begin{cases} \frac{15}{16} x^2 (2-x)^2 & x \in [0,2] \\ 0 & x \notin [0,2] \end{cases}$

היות ופונקציית צפיפות ההסתברות היא נגזרת של פונקציית ההתפלגות המצטברת, ניתן לנסח משפט דומה למשפט 4.3 עבור פונקציית הצפיפות, תוך שימוש באינטגרלים. חישובים אלו פחות נוחים, שכן כל אחד מהם מצריך חישוב של אינטגרל ולרוב נעדיף לחשב מתוך צפיפות ההסתברות את ההסתברות המצטברת פעם אחת ולאחר מכן להציב בה ערכים עבור חישובי ההסתברויות במקום לחשב מחדש את האינטגרל מחדש לכל חישוב הסתברות.

***משפט 4.5** (חישובי הסתברות באמצעות צפיפות). יהי X משתנה מקרי רציף בעל פונקציית צפיפות הסתברות $f_X(t)$. אזי לכל $a, b \in \mathbb{R}$ מתקיים:*

$$Pr\left(X \leq a\right) = Pr\left(X < a\right) = F_X(a) = \int_{-\infty}^{a} f_X(t)\, dt \qquad (4.13)$$

$$Pr\left(X \geq a\right) = Pr\left(X > a\right) = 1 - F_X(a) = \int_{a}^{\infty} f_X(t)\, dt \qquad (4.14)$$

$$Pr\left(a \leq X \leq b\right) = F_X(b) - F_X(a) = \int_{a}^{b} f_X(t)\, dt \qquad (4.15)$$

כאשר המשוואה האחרונה תקפה גם כאשר אחד או שני אי-השוויונות הוא אי-שוויון חזק.

לאור משפט 4.5 כאשר עוסקים במשתנה מקרי רציף אין הבדל בין אי-שוויון חזק לאי-שוויון חלש ובאופן יותר כללי – קבוצה סופית או בת מנייה של נקודות אינה משפיעה על ההתפלגות ועל המשתנה המקרי. לכן, כל עוד מדובר על משתנים מקריים רציפים, לא נקפיד על ההבדל בין אי-שוויון חזק או חלש ולא תמיד נטרח להגדיר את כל הפונקציות בכל מקום. למשל, פונקציית ההתפלגות המצטברת המוגדרת בתרגיל 4.10 אינה גזירה בנקודות החיבור בין התחומים $t = 0, 2, 5$ אך מאחר ומדובר על 3 נקודות בסך הכל, ניתן להגדיר את הצפיפות בהן כרצוננו או אפילו להתעלם מהן ולא להגדיר שם את הצפיפות כלל.

מאחר ומשתנה מקרי רציף לא ניתן להגדרה דרך תהליכים "סופיים" כמו חלוקת כדורים לתאים או שליפת קלפים מחפיסה, אין מרכיבים קומבינטוריים בחישובי ההתפלגות הרציפה. הקושי העיקרי במשתנים מקריים רציפים הוא הקושי הטכני של חישובי האינטגרלים[3] ולכן, לרוב, שאלות על משתנים מקריים רציפים יהיו קלות יותר מאשר שאלות על משתנים מקריים בדידים. נדגים את החישובים השונים אותם ניתן לבצע ואת השימוש במשפטים 4.3 ו-4.5 באמצעות הבעיה הבאה.

בעיה 4.4. הזמן העובר (בשעות) מרגע שמזמינים פיצה ועד שהמשלוח מגיע מתפלג לפי

$$f_X(x) = \begin{cases} cx^2(2-x)^2 & x \in [0,2] \\ 0 & x \notin [0,2] \end{cases}$$

1. חשבו את המקדם c.
2. מהי פונקציית ההתפלגות המצטברת של X?
3. אם המשלוח מגיע אחרי יותר מ-45 דקות, הפיצה מגיעה קרה. מה הסיכוי לקבל פיצה קרה?
4. עברה חצי שעה והפיצה עדיין לא הגיעה. מה הסיכוי שבכל זאת תגיע פיצה חמה?
5. כיצד מתפלג הזמן בו יוני קיבל את הפיצה, בהינתן שהוא יושב ואוכל פיצה קרה?

הוכחה. 1. הפונקציה $f_X(x)$ אי-שלילית בכל תחום ההגדרה ולכן ערכו של הקבוע c ייקבע לפי דרישת הנרמול במשפט 4.4:

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^2 cx^2(2-x)^2\,dx \\ &= c\left(\frac{x^5}{5} - x^4 + \frac{4x^3}{3}\right)_0^2 = \frac{16}{15}c \end{aligned}$$

כלומר $c = \frac{15}{16}$.

2. לפי ההגדרה: $F_X(t) = \int_{-\infty}^{t} f_X(x)\,dx$. מאחר והאינטגרנד משתנה בתחומים שונים, גם החישוב יבוצע בתחומים שונים.
עבור $t < 0$:

$$F_X(t) = \int_{-\infty}^{t} 0\,dx = 0$$

[3] לא נפרט באריכות את חישובי האינטגרלים. לפרטים נוספים ראו פרק 0.3.

עבור $t \in [0,2]$:

$$\begin{aligned}F_X(t) &= \int_{-\infty}^{t} f_X(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{t} \frac{15}{16}x^2(2-x)^2\,dx = \\ &= \frac{15}{16}\left(\frac{x^5}{5} - x^4 + \frac{4x^3}{3}\right)_0^t = \frac{3t^5}{16} - \frac{15t^4}{16} + \frac{5t^3}{4}\end{aligned}$$

ועבור $t > 2$:

$$F_X(t) = \int_{-\infty}^{t} f_X(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{2} \frac{15}{16}x^2(2-x)^2\,dx + \int_{2}^{t} 0\,dx = 1$$

נסכם:

$$F_X(t) = \begin{cases} 0 & t < 0 \\ \frac{3t^5}{16} - \frac{15t^4}{16} + \frac{5t^3}{4} & t \in [0,2] \\ 1 & t > 2 \end{cases}$$

3. הפיצה מגיעה קרה אם זמן המשלוח גדול מ-45 דקות, כלומר אם מתרחש המאורע $\{X > 0.75\}$. ניתן לחשב את המאורע על ידי שימוש בצפיפות או בהתפלגות המצטברת. מאחר וכבר חישבנו את ההתפלגות המצטברת, יותר פשוט להשתמש בה מאשר לחשב את האינטגרל של הצפיפות פעם נוספת.

$$\begin{aligned}Pr(X > 0.75) &= 1 - Pr(X \le 0.75) = 1 - F_X(0.75) = \\ &= 1 - 0.275 = 0.725\end{aligned}$$

4. הפיצה תגיע חמה אם $X < 0.75$ אבל כבר עברה חצי שעה, לכן ידוע ש-$X > 0.5$. כלומר המאורע המבוקש הינו $X < 0.75 | X > 0.5$. לפי נוסחא 2.9 נקבל:

$$\begin{aligned}Pr\left(X < 0.75 | X > 0.5\right) &= \frac{Pr\left(X < 0.75 \cap X > 0.5\right)}{Pr\left(X > 0.5\right)} = \frac{Pr\left(0.5 < X < 0.75\right)}{Pr\left(X > 0.5\right)} = \\ &= \frac{F_X(0.75) - F_X(0.5)}{1 - F_X(0.5)} = \frac{0.172}{0.896} = 0.192\end{aligned}$$

5. נסמן את הזמן בו הגיעה הפיצה הקרה ב-Y. המשתנה המקרי Y רציף לכן התפלגותו נקבעת לפי הצפיפות או פונקציית ההתפלגות המצטברת. נעדיף לחשב את ההתפלגות המצטברת, שכן היא מוגדרת דרך מאורע ועם מאורעות אנחנו יכולים להתמודד.

$$F_Y(t) = Pr\left(Y \le t\right) = Pr\left(X \le t | X > 0.75\right) = \frac{Pr\left(X \le t \cap X > 0.75\right)}{Pr\left(X > 0.75\right)}$$

אם $t < 0.75$ החיתוך במונה ריק וההסתברות היא אפס, כצפוי, שכן פיצות קרות מגיעות רק אחרי יותר

מ-45 דקות ולכן הסיכוי שהיא הגיעה מוקדם יותר, בהינתן שהיא קרה הוא אפס. אחרת, נמשיך כמו בסעיף הקודם עבור $t \in [0.75, 2]$:

$$\begin{aligned} F_Y(t) &= \frac{F_X(t) - F_X(0.75)}{1 - F_X(0.75)} = \frac{\frac{3t^5}{16} - \frac{15t^4}{16} + \frac{5t^3}{4} - 0.275}{1 - 0.275} \\ &= 0.259t^5 - 1.293t^4 + 1.724t^3 - 0.379 \end{aligned}$$

ועבור $t > 2$ פונקציית ההתפלגות המצטברת היא 1. נסכם:

$$F_Y(t) = \begin{cases} 0 & t < 0.75 \\ 0.259t^5 - 1.293t^4 + 1.724t^3 - 0.379 & t \in [0.75, 2] \\ 1 & t > 2 \end{cases}$$

□

תרגיל 4.13. זמן החימום (בשעות) של פיצה מוקפאת בתנור ביתי, Z, מתפלג לפי פונקציית ההתפלגות המצטברת

$$F_Z(t) = \begin{cases} 0 & t < 0 \\ \frac{ct}{t+1} & t \geq 0 \end{cases}$$

1. מהו הקבוע c?
2. מצאו את פונקציית הצפיפות של Z.
3. בכל פעם שיוני מעוניין לאכול פיצה, הוא מטיל מטבע: אם יצא עץ הוא מזמין פיצה (המגיעה בהתאם למשתנה המקרי המוגדר בבעיה 4.4) ואם יצא פלי, הוא מחמם פיצה מהמקפיא בתנור. מה ההסתברות שהוא יאכל פיצה תוך פחות מחצי שעה לאחר שחשקה נפשו בפיצה?
4. תוך פחות מחצי שעה מהרגע שחשקה נפשו של יוני בפיצה הוא החל לאכול אותה. מה ההסתברות שהוא הזמין את הפיצה אותה הוא אוכל מפיצריה?

תרגיל 4.14. הציון בבחינה הוא משתנה מקרי רציף המתפלג לפי פונקציית הצפיפות

$$f_X(t) = \begin{cases} ct(100 - t) & t \in [c_1, c_2] \\ 0 & t \notin [c_1, c_2] \end{cases}$$

כאשר c_1, c_2 קבועים ו-$c > 0$ קבוע חיובי.

1. מהם הערכים האפשריים עבור c_1, c_2? האם ייתכן כי $c_1 = c_2$?
עבור שאר הסעיפים נקבע $c_1 = 0, c_2 = 100$.
2. מהו ערכו של הקבוע c?
3. מהי פונקציית ההתפלגות המצטברת של X?
4. על מנת לעבור את הבחינה יש לקבל יותר מ-60. על מנת להצטיין בקורס, יש לקבל יותר מ-90. מהי ההסתברות להיכשל במבחן? מהי ההסתברות להצטיין בקורס?
5. סופי עברה את המבחן. מה הסיכוי שהיא הצטיינה בקורס?

6. הסטודנטים התלוננו וטענו שיותר מחצי כיתה נכשלה במבחן. האם הם צודקים? בכמה נקודות יש להעלות את הציונים של כל הסטודנטים ("פקטור") על מנת שאכן רק חצי כיתה תיכשל במבחן?

פונקציית הצפיפות והכלים המתמטיים המתאימים מאפשרים טיפול במרחבי מדגם שאינם בדידים. ההכללה השימושית ביותר למקרה הרציף הינה ההכללה של נוסחת ההסתברות השלמה (פרק 2.2.1). בנוסחת ההסתברות השלמה (נוסחא 2.17) ההנחה הייתה שמספר המאורעות הוא סופי או אינסופי אך בן-מנייה, כך שהסכימה מוגדרת. במקרים רבים מאורעות תלויים בגדלים שאינם בני מנייה. למשל, זמן היציאה מהבית לעבודה בבוקר הוא משתנה מקרי רציף ואילו הסיכוי לאחר לעבודה תלוי בזמן היציאה. אם נרצה לחשב את הסיכוי הכולל לכך שנאחר לעבודה, עלינו לקחת בחשבון את כל זמני היציאה האפשריים. באופן כללי, כאשר נדרש לחלק את מרחב המדגם למספר שאינו בן-מנייה של מאורעות, הביצוע יהיה דרך פונקציית צפיפות ואינטגרציה, בדומה לנוסחת ההסתברות השלמה. באופן פורמלי, אם קיים משתנה מקרי רציף X כך שבהינתן ערכו חישוב ההסתברות של המאורע A פשוט יותר, אזי ניתן לחשב את ההסתברות של המאורע A על ידי

$$Pr\left(A\right)=\int_{-\infty}^{\infty}Pr\left(A|X=x\right)\cdot f_{X}\left(x\right)\,dx \tag{4.16}$$

תרגיל 4.15. במפעל מסויים מייצרים מטבעות שהסיכוי שלהם ליפול על עץ הינו משתנה מקרי X עם פונקציית הצפיפות:

$$f_{X}\left(t\right)=\begin{cases}12t^{2} & t\in\left[0,\frac{1}{2}\right]\\ 12\left(t-1\right)^{2} & t\in\left[\frac{1}{2},1\right]\\ 0 & t\notin\left[0,1\right]\end{cases}$$

הארווי מטיל מטבע שיוצר במפעל פעמיים. האם המאורעות A - המטבע נפל על עץ בהטלה הראשונה ו-B- המטבע נפל על עץ בהטלה השנייה תלוים או בלתי תלויים? ענו (מבלי לחשב) ואז ודאו את תשובתכם על ידי חישוב.

חסרונה העיקרי של פונקציית הצפיפות ביחס לפונקציית ההתפלגות המצטברת היא שהאחרונה מוגדרת דרך הסתברות של מאורע ואנחנו יודעים לטפל בהסתברויות של מאורעות (פרק 2). בנוסף, המאורע דרכו פונקציית ההתפלגות המצטברת מוגדרת הוא אי-שוויון, ולכן ניתן להציב בו קשרים שיש בין המשתנה המקרי למשתנים מקריים אחרים בכדי לחשב את ההתפלגות שלהם. דוגמא כזו ראינו באחד הסעיפים של בעיה 4.4, אך מפאת חשיבותה של טכניקה זו נדגים אותה שנית כבעיה בפני עצמה.

בעיה 4.5. נתון מקל באורך 5 מטרים. המקל נשבר לשניים בנקודה המרוחקת X מקצהו השמאלי. פונקציית הצפיפות של נקודת השבר נתונה על ידי

$$f_X(x)=\begin{cases}\frac{6}{125}\cdot x(5-x) & x\in[0,5]\\ 0 & x\notin[0,5]\end{cases}$$

מהי הצפיפות של Y, שבר המקל הארוך מבין השניים?

הוכחה. אמנם השאלה נתונה במונחי צפיפות והתשובה הסופית הנדרשת היא צפיפות, אבל בהמשך לדוגמאות הקודמות, נחשב את פונקציית ההתפלגות המצטברת של X ונשתמש בה לחישוב פונקציית ההתפלגות המצטברת

של Y. לאחר מכן, נגזור אותה לקבלת הצפיפות.

$$F_X(t) = \int_{-\infty}^{t} f_X(x)\,dx = \begin{cases} 0 & t < 0 \\ \frac{1}{125}t^2(15-2t) & t \in [0,5] \\ 1 & t > 5 \end{cases}$$

על מנת להשתמש ב-X למציאת Y עלינו למצוא את הקשר ביניהם. כאשר המקל נשבר לשניים נוצרים שני מקלות (כמוראה בתרשים), אחד באורך X והשני באורך $5 - X$. שבר המקל הארוך יותר הוא

$$Y(X) = \max\{X, 5 - X\}$$

לפיכך ברור כי אורכו של שבר המקל הארוך יותר צריך להיות בין 2.5 מטרים ל-5 מטרים, ומספיק לחשב את הפה"מ של Y רק בתחום זה. יהי $t \in [2.5, 5]$ ונחשב את הפה"מ של Y ב-t:

$$\begin{aligned} F_Y(t) &= Pr\left(Y \le t\right) = Pr\left(\max\{X, 5-X\} \le t\right) \\ (*) &= Pr\left(\{X \le t\} \cap \{5 - X \le t\}\right) = Pr\left(5 - t \le X \le t\right) \end{aligned}$$

כאשר המעבר$*$ נובע מכך שאם המקסימום של שני ביטויים קטן מ-t אזי כל אחד מהם גם צריך להיות קטן מ-t.

$$\begin{aligned} Pr\left(5 - t \le X \le t\right) &= F_X(t) - F_X(5-t) \\ &= \frac{1}{125}t^2(15-2t) - \frac{1}{125}(5-t)^2\left(15 - 2\cdot(5-t)\right) \\ &= \frac{-4t^3 + 30t^2 - 125}{125} \end{aligned}$$

הפה"מ המלא של Y הינו

$$F_Y(t) = \begin{cases} 0 & t < 2.5 \\ \frac{-4t^3+30t^2-125}{125} & t \in [2.5, 5] \\ 1 & t > 5 \end{cases}$$

והצפיפות מתקבלת מגזירה:

$$f_Y(t) = \frac{dF_Y(t)}{dt} = \begin{cases} \frac{-12t^2+60t}{125} & t \in [2.5, 5] \\ 0 & t \notin [2.5, 5] \end{cases}$$

$\square$

תרגיל 4.16. על מנת למנוע מצב שסטודנט ייכשל על חודה של נקודה, הוחלט להוסיף 10 נקודות לכל הסטודנטים מתרגיל 4.14 שקיבלו בין 50 ל-60 במבחן. כיצד מתפלג ציון המבחן לאחר התוספת?

תרגיל 4.17. עשרה סטודנטים ניגשים למבחן המתואר בתרגיל 4.14. כל אחד מהסטודנטים מקבל ציון בהתאם לפונקציית הצפיפות המתוארת בשאלה באופן בלתי תלוי בסטודנטים האחרים. כיצד מתפלג הציון הכי גבוה שהתקבל במבחן? כיצד מתפלג הציון הכי נמוך שהתקבל במבחן?

משפט 4.6 (התפלגות של מקסימום ומינימום). *יהי X משתנה מקרי עם פונקציית התפלגות מצטברת $F_X(t)$. מגרילים באופן בלתי תלוי n תוצאות שונות לפי ההתפלגות של X ונסמן ב-M את המקסימום של התוצאות שהתקבלו וב-m את המינימום. אזי התפלגותיהם נתונות לפי:*

$$F_M(t) = (F_X(t))^n \tag{4.17}$$
$$F_m(t) = 1 - (1 - F_X(t))^n \tag{4.18}$$

הוכחה. ההוכחה זהה לפתרון של תרגיל 4.17. □

תרגיל 4.18. יהי X משתנה מקרי עם תומך סופי $[a, b]$, כלומר $f_X(t) = 0$ כאשר $t \notin [a, b]$ ו-$f_X(t) > 0$ כאשר $t \in [a, b]$. כיצד מתפלג המקסימום והמינימום של n הגרלות בלתי תלויות של המשתנה המקרי כאשר $n \to \infty$?

רמז: מה יהיה המקסימום והמינימום של הטלת קוביה? (ראו בעיה 4.3 ותרגיל 4.22).

בעיה 4.6. יורים למטרה עגולה ברדיוס 10 ס"מ. מיקום הפגיעה מתפלג באקראי על פני שטח המטרה. כיצד מתפלג המרחק מנקודת הפגיעה?

בטרם נפתור את הבעיה חשוב להבהיר את משמעות המושג "באקראי" - כאשר נתון שמשהו קורה באקראי, בהיעדר נתונים אחרים, הכוונה היא שלכל התוצאות הסתברות זהה להתקבל. למשל, כאשר מסדרים אנשים באקראי, הכוונה היא שכל סידור אפשרי וכל הסידורים מתקבלים באותה ההסתברות. במקרה דנן, כל נקודת פגיעה על פני שטח המטרה אפשרית וכולן סימטריות. הבנה זו היא המפתח לחישוב ההתפלגות הרצויה.

הוכחה. נגדיר ב-R את המרחק של נקודת הפגיעה ממרכז המטרה. בבירור, המרחק בהכרח יהיה בין 0 ל-10, לכן נותר לנו לחשב את פונקציית ההתפלגות המצטברת $F_R(t)$ רק עבור ערכי $t \in [0, 10]$. מאחר ונקודת הפגיעה אקראית על פני המטרה ואין נקודות עדיפות יותר או פחות, הסיכוי לפגוע בריבוע בשטח 1 סמ"ר זהה וקבוע, ללא תלות במיקום הריבוע על פני המטרה. אם נכליל היגיון זה, הרי שהסיכוי לפגוע בכל תחום יהיה שווה לשטח התחום ביחס לשטח הכולל של המטרה. לכן, המאורע $R \leq t$ מתאר פגיעה בעיגול ברדיוס t וההסתברות לכך הינה היחס בין שטח העיגול לשטח המטרה כולה:

$$F_R(t) = Pr\left(R \leq t\right) = \frac{\pi t^2}{\pi 10^2} = \frac{t^2}{100}$$

ובסך הכל פונקציית ההתפלגות המצטברת הינה

$$F_R(t) = \begin{cases} 0 & t < 0 \\ \frac{t^2}{100} & t \in [0, 10] \\ 1 & t > 100 \end{cases}$$

נגזרתה היא פונקציית הצפיפות:

$$f_R(t) = \begin{cases} \frac{t}{50} & t \in [0,10] \\ 0 & t \notin [0,10] \end{cases}$$

שימו לב שהצפיפות עולה, כלומר מרחקים גדולים מנקודת המרכז מתקבלים בסיכוי גדול יותר מאשר מרחקים קטנים. הסיבה לכך היא שבטבעת שבין $R=9$ ל-$R=10$ יש הרבה יותר שטח מאשר בטבעת שבין $R=2$ ל-$R=3$, למרות שההבדל בין הרדיוסים זהה. לכן, יש סיכוי גדול יותר לפגוע בטבעת החיצונית מאשר בטבעת הפנימית. $\square$

תרגיל 4.19. נתון משולש שגובהו h. נקודה נבחרת באקראי בתוך המשולש. נסמן ב-X את מרחקה האנכי של הנקודה מבסיס המשולש. מצאו את פונקציית ההתפלגות המצטברת ואת פונקצית הצפיפות של X.

4.3 משתנה מקרי מעורב

הסוג השלישי של המשתנים המקריים הם המשתנים המקריים המעורבים, הכוללים מרכיבים בדידים ומרכיבים רציפים. משתנים מקריים אלו כוללים ערכים שיכולים להתקבל בהסתברות חיובית, בדומה למשתנה מקרי בדיד (ערכים אלו מכונים "סינגלטונים" של ההתפלגות) ותחומי ערכים אחרים בהם פונקציית ההתפלגות המצטברת עולה ורציפה, בדומה למשתנה מקרי רציף. נטפל במשתנים מקריים מעורבים רק דרך פונקציית ההתפלגות המצטברת שכן לא הצפיפות ולא פונקציית ההסתברות מתאימות לתיאור שני החלקים בו זמנית.

היות ומשתנים מקריים מעורבים כוללים גם חלקים בדידים, יש לשוב ולהקפיד על אי-השוויון הנכון ($\leq$ או $<$) ועל תחומי ההגדרה הנכונים. מאחר ופונקציית ההתפלגות המצטברת מוגדרת דרך אי-שוויון חלש ($X \leq t$), ערכה בכל נקודה יכלול גם את ההסתברות עצמה, לכן אמנם הפונקציה לא בכל מקום תהיה רציפה אך היא תמיד תהיה רציפה מימין, גם למשתנים מקריים בדידים וגם למעורבים, כפי שניתן לראות בשרטוטים המופיעים בבעיות 4.3 ו-4.7.

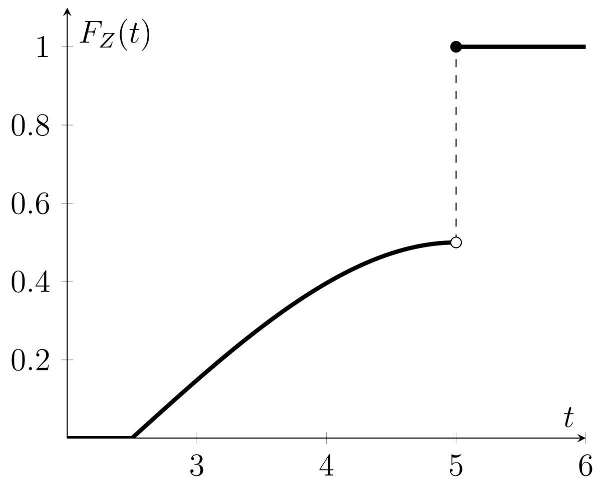
בעיה 4.7. נתון מקל עץ באורך 5 מטרים. מטילים מטבע ובהתאם לתוצאה מחליטים האם לשבור או לא לשבור את המקל. אם יוצא עץ – לא שוברים את המקל ואם יוצא פלי – שוברים אותו לשניים בהתאם לפונקציית הצפיפות המתוארת בבעיה 4.5. כיצד מתפלג Z, שבר המקל הארוך יותר ($Z=5$ אם המקל לא נשבר).

הוכחה. בדומה לבעיה 4.5, שבר המקל הארוך יותר הוא לפחות 2.5 ולכל היותר 5. על מנת לחשב את ההתפלגות של Z, נחשב את פונקציית ההתפלגות המצטברת שלו בתחום $t \in (2.5,5)$:

$$F_Z(t) = Pr\left(Z \leq t\right) = \frac{1}{2} \cdot Pr\left(Y \leq t\right) = \frac{1}{2} \cdot \frac{-4t^3 + 30t^2 - 125}{125}$$

שכן המקל נשבר לפי Y הנתון בבעיה 4.5 רק אם במטבע יצא פלי, כלומר בהסתברות חצי. לכן, הפה"מ בכל תחום ההגדרה יהיה

$$F_Z(t) = \begin{cases} 0 & t \leq 2.5 \\ \frac{-4t^3+30t^2-125}{250} & t \in [2.5,5) \\ 1 & t \geq 5 \end{cases}$$



כצפוי, ישנה קפיצה עבור $Z=5$ שכן זהו סינגלטון של ההתפלגות ויכול להתקבל בהסתברות חצי (כאשר המטבע נופל על עץ). $\square$

תרגיל 4.20. על מנת לוודא שסטודנטים, שציוניהם מתפלגים לפי פונקציית הצפיפות המופיעה בתרגיל **4.14**, לא יכשלו על חודה של נקודה, הוחלט להעלות את ציוניהם של כל הסטודנטים בתחום $[55,60]$ ל-60. כיצד מתפלגים ציוני הסטודנטים כעת?

תרגיל 4.21. קלעי יורה לעבר מטרה עגולה ברדיוס 10 סנטימטרים. המרחק של נקודת הפגיעה (בסנטימטרים) ממרכז המטרה מתפלג לפי פונקציית הצפיפות

$$f_R(t)=\begin{cases}\frac{1}{20} & t\in[0,20]\\ 0 & t\notin[0,20]\end{cases}$$

אם הוא פוגע במרחק הקטן מסנטימטר ממרכז המטרה, הוא זוכה ב-10 נקודות. אם הוא מפספס את המטרה לחלוטין, הוא זוכה ב-0 נקודות. אחרת, אם הוא פוגע במרחק R ממרכז המטרה, הוא מקבל $\frac{10}{R}$ נקודות. כיצד מתפלג הניקוד שהוא יקבל עבור יריה בודדת?

4.4 שאלות מסכמות

4.4.1 משתנים מקריים בדידים

תרגיל 4.22. מטילים קוביה 3 פעמים. נסמן ב-Y את התוצאה המינימלית שהתקבלה. כיצד מתפלג Y?

תרגיל 4.23. לשיעור ריקודים מגיעים 3 זוגות, 2 מהם זוגות של רקדנים חובבים וזוג אחד של רקדנים מקצועיים. מחלקים מחדש את הזוגות באקראי כך שלכל גבר יש סיכוי זהה לרקוד עם כל אחת מהנשים. נגדיר X – מספר הזוגות (המקוריים) שרוקדים זה עם זו, Y – מספר הזוגות (המקוריים) שרוקדים זה עם זו מתוך הזוגות החובבים. כיצד משתנים מקריים אלו מתפלגים?

תרגיל 4.24. במסיבת ריקודים משתתפים 50 זוגות. בוחרים באקראי 10 אנשים מבין מאה המשתתפים. מה הסיכוי שהזוג סמית' נבחר? כיצד מתפלג מספר הזוגות שנבחרו (בשלומתם, כלומר שני בני הזוג נבחרו)?

תרגיל 4.25. לתחרות שחמט מגיעים 16 משתתפים, 6 מתוכם נשים והיתר גברים. המשתתפים מחולקים באקראי לזוגות לצורך המשחק. נגדיר X – מספר המשחקים המתנהלים בין שתי נשים, Y – מספר המשחקים המתנהלים בין שני גברים ו-Z – מספר המשחקים המתנהלים בין גבר לאישה. כיצד כל אחד ממשתנים מקריים אלו מתפלג?

תרגיל 4.26. הסיכוי להיכשל במועד א' של קורס מסויים הינו 0.5 והסיכוי להיכשל במועד ב' הינו 0.4. למועד א' ניגשים כל 100 הסטודנטים בקורס ולמועד ב' ניגשים אך ורק הסטודנטים שנכשלו במועד א'. סטודנטים שנכשלו בשני המועדים יכולים לגשת גם למועד ג'. כיצד מתפלג X, מספר מועדי הבחינה שיתקיימו בקורס בהנחה שהסטודנטים לא מעתיקים זה מזה ולכן ציוניהם בלתי תלויים?
לדוגמא: אם כולם עברו במועד א', לא יתקיימו מועדים נוספים ולכן $X = 1$.

4.4.2 משתנים מקריים רציפים

תרגיל 4.27. נתונה פונקציית צפיפות ההתפלגות הבאה:

$$f_X(t) = \begin{cases} -t & t \in [-1,0] \\ ct^2 & t \in (0,1] \\ 0 & t \notin [-1,1] \end{cases}$$

1. מצאו את ערכו של c.
2. חשבו את פונקציית ההתפלגות המצטברת של X.
3. חשבו את ההסתברויות:
   א) $Pr(X \leq -0.5)$
   ב) $Pr(X < -0.5)$
   ג) $Pr(X \leq 0.5)$
   ד) $Pr(-0.2 \leq X \leq 0.3)$

תרגיל 4.28. אספקת הדלק למאגר של תחנת דלק מתבצעת אחת לשבוע. הביקוש לדלק (באלפי ליטרים) בתחנה זו בכל שבוע הוא משתנה מקרי המתפלג בהתאם לפונקציית הצפיפות:

$$f_X(t) = \begin{cases} c(1-t)^4 & t \in [0,1] \\ 0 & t \notin [0,1] \end{cases}$$

1. חשבו את c.
2. כאשר הביקוש גדול מקיבולת המאגר, הוא מתרוקן. מה צריכה להיות קיבולת המאגר כך שהוא יתרוקן בהסתברות קטנה מ-0.05?

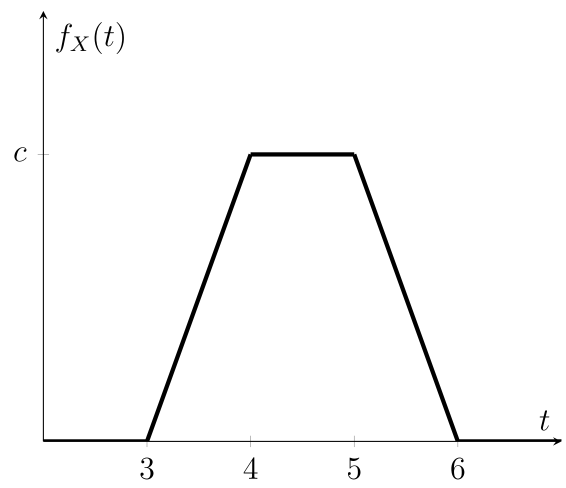
תרגיל 4.29. פונקציית ההתפלגות של מיקום אלקטרון בתוך בור קוונטי באורך L נתונה על ידי:

$$f_X(t) = \begin{cases} A \cdot \cos^2\left(2\pi n \frac{t}{L}\right) & t \in \left[-\frac{L}{2}, \frac{L}{2}\right] \\ 0 & t \notin \left[-\frac{L}{2}, \frac{L}{2}\right] \end{cases}$$

כאשר n טבעי.

1. חשבו את A.
2. מה ההסתברות למצוא את האלקטרון בדיוק במרכז הבור?
3. מה ההסתברות שהאלקטרון נמצא במרחק קטן מ-$0.1L$ ממרכז הבור?

תרגיל 4.30. מפעל מייצר ברגים עם ראש עגול על ידי מכונה לא מדוייקת. הקוטר (במילימטרים) של ראשי הברגים מתפלג לפי פונקציית צפיפות ההסתברות הבאה:



1. רשמו, במונחים של c, את הנוסחא המתארת את פונקציית צפיפות ההסתברות.
2. מהו c?
3. מהי פונקציית ההתפלגות המצטברת של X?
4. כיצד מתפלג Y, השטח של ראש הבורג?

תרגיל 4.31. יהי X משתנה מקרי רציף ונגדיר משתנה מקרי חדש על ידי $Y = aX + b$ כאשר $a, b \in \mathbb{R}$. מצאו את פונקציית צפיפות ההסתברות של Y במונחי פונקציית צפיפות ההסתברות של X.

4.5 פתרונות לשאלות הפרק

תרגיל 4.1. 1. מספר ההטלות הוא משתנה מקרי בדיד והתומך הינו $\mathbb{N}$, שכן מספר ההטלות יכול להיות רק מספר שלם.

2. הגובה הוא משתנה מקרי רציף (ישנו רצף של גבהים אפשריים, למשל 1.54, 1.545, 1.76443 וכד'). היות ואנחנו לא יודעים מהם ערכי הקיצון האפשריים, נבחר תומך שבוודאות כולל את כל האפשרויות, למשל את הקטע $[0,3]$, למרות שטרם נתקלנו באדם בגובה 3 מטרים. באופן כללי, רצוי שהתומך יכלול בדיוק את האפשרויות המתקבלות בניסוי (במיוחד אם זהו משתנה מקרי בדיד), אבל אם לא בטוחים בערכים מסויימים, עדיף לכלול אותם בתומך ולגלות בהמשך הפתרון שההסתברות לקבלם היא 0 מאשר לא לכלול אותם ולטעות.

3. מרחק הפגיעה יכול להיות כל מספר ממשי חיובי, ולכן התומך הינו $[0,\infty)$ והמשתנה המקרי רציף.

4. מספר הבנות שנבחרות הוא משתנה מקרי בדיד ויכול לקבל את הערכים $\{0,1,2,3,4\}$.

5. מספר ההטלות הוא משתנה מקרי בדיד. כדי שיצאו שתי התוצאות האפשריות צריך להטיל את המטבע לפחות פעמיים, אבל ברמת העיקרון המטבע יכול גם ליפול אלפי פעמים על עץ לפני שיופיע הפלי הראשון. לכן התומך הינו $\mathbb{N}\setminus\{1\}=\{2,3,\ldots\}$.

6. מיקום השיכור לאחר 10 צעדים הוא משתנה מקרי בדיד. אם הוא צעד Y צעדים ימינה אז הוא בהכרח צעד גם $10-Y$ שמאלה ולכן מיקומו יהיה $X=Y-(10-Y)=2Y-10$. מאחר ש-Y מקבל כל שלם בין 0 ל-10, X יקבל כל ערך זוגי בין -10 ל-10 ותומכו יהא $\{-10,-8,\ldots,8,10\}$.

□

תרגיל 4.2. לכל ערך $k\in supp(X)$ נגדיר את המאורע $A_k=\{X=k\}$. המשתנה המקרי X הוא פונקציה ממרחב המדגם לממשיים ולכן כל תוצאה במרחב המדגם כלולה בדיוק במאורע אחד: התוצאה $\omega\in\Omega$ כלולה רק במאורע $A_{X(\omega)}$. לכן המאורעות A_k זרים בזוגות. בנוסף, איחודם מהווה את כל מרחב המדגם (מאותה הסיבה) ולכן לפי אקסיומה 2.3 נקבל:

$$\sum_{k\in supp(X)} P_X(k)=\sum_{k\in supp(X)} Pr\left(A_k\right)=Pr\left(\bigcup_{k\in supp(X)} A_k\right)=Pr\left(\Omega\right)=1$$

□

תרגיל 4.3. נחשב את סכום ההסתברויות של כל הערכים בתומך: באמצעות הנוסחאות של טור גאומטרי (נוסחאות 0.14 ו-0.14):

$$\begin{aligned}\sum_{k\in supp(X)} P_X(k) &= \sum_{k=1}^{4} P_X(k)+\sum_{k=6}^{\infty} P_X(k)\\ &= \sum_{k=1}^{4}\left(\frac{4}{5}\right)^{k-1}\cdot\frac{1}{5}+\sum_{k=6}^{\infty}\left(\frac{4}{5}\right)^{4}\cdot\left(\frac{5}{6}\right)^{k-6}\cdot\frac{1}{6}\end{aligned}$$

הסכום הראשון הוא טור גאומטרי סופי (נוסחא 0.14) והטור השני הוא טור גאומטרי אינסופי (נוסחא 0.14). נשנה את הסכימה כך שתתחיל מ-0 ותתאים לנוסחא ונקבל:

$$\sum_{k\in supp(X)} P_X(k) = \frac{1}{5}\cdot\sum_{k=0}^{3}\left(\frac{4}{5}\right)^k + \left(\frac{4}{5}\right)^4\cdot\frac{1}{6}\cdot\sum_{k=0}^{\infty}\cdot\left(\frac{5}{6}\right)^k$$
$$= \frac{1}{5}\cdot\frac{1-\left(\frac{4}{5}\right)^4}{1-\frac{4}{5}} + \left(\frac{4}{5}\right)^4\cdot\frac{1}{6}\cdot\frac{1}{1-\frac{5}{6}} = 1-\left(\frac{4}{5}\right)^4 + \left(\frac{4}{5}\right)^4 = 1$$

כנדרש. □

תרגיל 4.4. נתחיל מהתומך של X– התוצאות האפשרויות בסכום של 2 קוביות הן המספרים הטבעיים מ-2 עד 12. בכדי לחשב את ההסתברות לכל ערך נשתמש במרחב מדגם סימטרי הכולל את כל הזוגות האפשריים של הטלות. גודל מרחב המדגם הוא $6^2 = 36$, כמספר האפשרויות לבחור תוצאה לקוביה הראשונה מתוך 6 ותוצאה לקוביה השניה מתוך 6.

כעת ניתן לחשב את ההסתברות של כל תוצאה:

$$\begin{aligned} Pr\left(X=2\right) &= \frac{\left|\{(1,1)\}\right|}{6^2} = \frac{1}{36} \\ Pr\left(X=3\right) &= \frac{\left|\{(1,2),(2,1)\}\right|}{6^2} = \frac{2}{36} \\ Pr\left(X=4\right) &= \frac{\left|\{(1,3),(2,2),(3,1)\}\right|}{6^2} = \frac{3}{36} \end{aligned}$$

ובאופן דומה ממשיכים לחשב את האפשרויות לכל הערכים האפשריים בתומך. לאחר שמסיימים את החישוב ניתן לראות שעד 7 המונה גדל ב-1 והחל מ-7 המונה קטן ב-1 ולכן הנוסחא הכללית היא $P_X\left(k\right) = \frac{6-|7-k|}{36}$.

עבור Y, ההפרשים נעים בין 0 ל-5 ולכן התומך הינו $\{0,\ldots,5\}$. נשתמש באותו מרחב מדגם המתאים להטלה של שתי קוביות. ראשית, $P_Y(0) = Pr\left(Y=0\right) = \frac{6}{36} = \frac{1}{6}$ שכן ישנן 6 אפשרויות בהן תוצאת שתי הקוביות זהה ולכן הפרשן בערכו המוחלט הוא 0. עבור $k \in \{1,\ldots,5\}$ ערכו המוחלט של ההפרש יכול להיות k בשתי צורות: כאשר הקוביה הראשונה קטנה ב-k מהשניה או להפך. עבור האפשרות הראשונה, עלינו לבחור ערך לקוביה הראשונה l ובהכרח הערך של הקוביה השניה יהיה $l+k$. מאחר וערך זה צריך להיות ערך חוקי בקוביה, $l+k \le 6$ ולכן יש $6-k$ אפשרות לבחור את התוצאה של הקוביה הראשונה: $\{1,2,\ldots,6-k\}$. מסימטריה, זהו גם מספר האפשרויות עבור האפשרות השניה ולכן

$$P_Y(k) = Pr\left(Y=k\right) = \frac{2\cdot(6-k)}{36} = \frac{6-k}{18}$$

לסיכום:

$$P_Y(k) = \begin{cases} \frac{1}{6} & k=0 \\ \frac{6-k}{18} & k=1,\ldots,5 \end{cases}$$

□

תרגיל 4.5. 1. התומך הוא מספר הסיבובים במשחק, כלומר מספר טבעי בין 1 ל-10. המאורע $\{X=k\}$ עבור $1 \le k \le 9$ מתקבל כאשר השחקן ניצח בכל אחד מ-$k-1$ הסיבובים הקודמים והפסיד בסיבוב ה-k,

לכן ההסתברות היא $Pr\left(X=k\right)=p^{k-1}\left(1-p\right)$. המאורע $\{X=10\}$ מתקבל בשני מצבים - או שהשחקן ניצח בכל 10 הסיבובים (ואז הוא פורש) או שהוא ניצח ב-9 סיבובים והפסיד בסיבוב העשירי. לכן, המאורע הזה מתרחש תמיד אחרי שהשחקן ניצח ב-9 סיבובים, ללא תלות בתוצאת הסיבוב העשירי. ההסתברות המתאימה היא $Pr\left(X=10\right)=p^9$ ופונקציית ההסתברות הינה:

$$P_X(k)=\begin{cases} p^{k-1}\left(1-p\right) & k=1,\ldots,9 \\ p^9 & k=10 \end{cases}$$

2. המאורע שהוא מפסיק לשחק לפני הסיבוב העשירי הינו $\{X<10\}$ ועבור התומך של המשתנה המקרי מאורע זה זהה ל-$\{X\neq 10\}$. לכן:

$$Pr\left(X<10\right)=Pr\left(X\neq 10\right)=1-Pr\left(X=10\right)=1-\left(\frac{1}{2}\right)^9=0.998$$

□

תרגיל 4.6. יהיו $a<b$ ממשיים. המאורע $\{X\leq a\}$ מוכל במאורע $\{X\leq b\}$ שכן כל תוצאה של הניסוי הקטנה מ-a בוודאי גם קטנה מ-b. לפיכך, לפי משפט 4.3 מתקיים

$$F_X(a)=Pr\left(X\leq a\right)\leq Pr\left(X\leq b\right)=F_X(b)$$

והפונקציה עולה חלש. בנוסף, מדובר על פונקציה חסומה שכן לכל t, $F_X(t)$ הינה ההסתברות של המאורע $\{X\leq t\}$ ולכן $0\leq F_X(t)\leq 1$. זוהי פונקציה עולה וחסומה ולכן קיימים הגבולות ב-∞ וב-$-\infty$. מאחר ו-X משתנה מקרי המקבל בוודאות ערך ממשי כלשהו, כלומר מתקיים $Pr\left(X<\infty\right)=1$ ולכן הגבול באינסוף הוא 1. מאותה סיבה, $Pr\left(X>-\infty\right)=1$ ולכן $Pr\left(X\leq-\infty\right)=0$ וזהו גם הגבול ב-$-\infty$. □

תרגיל 4.7. 1. מאחר ונבחרים m כדורים ללא החזרה, המקסימום הקטן ביותר האפשרי הוא m (כאשר נבחרו הכדורים $1,2,\ldots,m$) והמרבי הוא n (כאשר נבחר הכדור n). לכן התומך הינו $\{m,m+1,\ldots,n\}$. לכל l בתומך, הערך של X יהיה l רק כאשר יצא הכדור שמספרו l ובשאר $m-1$ ההוצאות נבחרו רק כדורים מתוך הכדורים שמספריהם בין 1 ל-$l-1$. ההוצאה מתבצעת ללא החזרה וללא סדר (מסתכלים רק על המספרים שיצאו ולא משנה לנו הסדר בו המספרים נבחרו) ולכן:

$$P_X(l)=Pr\left(X=l\right)=\frac{\binom{l-1}{m-1}}{\binom{n}{m}}$$

מעניין לציין שבאמצעות תכונות פונקציית ההתפלגות ניתן להוכיח זהויות אלגבריות. למשל, מכך שסכום ההסתברויות הוא 1 נקבל

$$1=\sum_{l=m}^{n}P_X\left(l\right)=\sum_{l=m}^{n}\frac{\binom{l-1}{m-1}}{\binom{n}{m}}=\frac{1}{\binom{n}{m}}\sum_{l=m}^{n}\binom{l-1}{m-1}$$

כלומר

$$\sum_{l=m}^{n}\binom{l-1}{m-1}=\binom{n}{m} \qquad (4.19)$$

מקרה פרטי של טענה זו הוכחנו בתרגיל 0.1 באמצעות אינדוקציה.

2. כאשר מתבצעת החזרה, השאלה היא הכללה של החישוב עבור הקוביה (בעיה 4.3). הפעם התומך יהיה $\{1,\ldots,n\}$ומרחב המדגם יהיה n^m (כי יש n אפשרויות ולא 6) לכן פונקציית ההסתברות תהיה:

$$P_X(l)=Pr\left(X\le l\right)-Pr\left(X\le l-1\right)=\frac{l^m-(l-1)^m}{n^m}$$

□

תרגיל 4.8. נתחיל מחישוב התומך. מאחר ונצבעות 6 משבצות, רק 3 משבצות נותרות בצבען המקורי. לכן ייתכן מצב שבו אין אף שורה בה יש שתי משבצות לא צבועות וייתכן מצב בו יש שורה אחת בו יש שתי משבצות לא צבועות אך לא ייתכנו מצבים אחרים שכן מספר גדול יותר של שורות עם שתי משבצות לא צבועות מצריך יותר מ-3 משבצות לא צבועות. אם כך, $supp(X)=\{0,1\}$. מרחב המדגם הוא בחירה של 3 המשבצות מתוך 9 המשבצות שלא ייצבעו, בחירה זו נעשית ללא החזרה וללא סדר, כי הדבר היחיד שחשוב הוא זהות המשבצות שנבחרו ולא הסדר בו הן נבחרו. לכן גודלו של מרחב המדגם הינו $\binom{9}{3}$.

- $X=0$
תוצאה זו יכולה להתרחש כאשר יש שורה שכל משבצותיה לא נצבעו או כאשר בכל שורה לא נצבעה בדיוק משבצת אחת. במקרה הראשון יש בדיוק 3 אפשרויות, כי ברגע שנבחרה השורה מתוך ה-3, בחירת המשבצות נקבעת באופן יחיד. במקרה השני, צריך לבחור עבור כל שורה מי מ-3 המשבצות תהיה לא צבועה ולכן בכל שורה יש 3 אפשרויות בחירה ובסך הכל יש 3^3 אפשרויות בחירה. לפיכך:

$$Pr\left(X=0\right)=\frac{3+3^3}{\binom{9}{3}}=\frac{5}{14}$$

- $X=1$
תוצאה זו יכולה להתרחש רק אם באחת השורות שתי משבצות לא נצבעו ובאחת משתי השורות האחרות משבצת אחת לא נצבעה. במקרה כזה אנחנו צריכים לבחור את השורה בה שתי משבצות לא ייצבעו (יש 3 אפשרויות), לבחור את שתי המשבצות מתוך ה-3 של השורה (יש $\binom{3}{2}$ אפשרויות) ולאחר מכן יש לבחור אחת מ-6 המשבצות הנותרות בשתי השורות האחרונות ולכך יש 6 אפשרויות. לפיכך:

$$Pr\left(X=1\right)=\frac{3\cdot\binom{3}{2}\cdot 6}{\binom{9}{3}}=\frac{9}{14}$$

בנוסף, קל לראות שאכן מתקיים $P_X(0)+P_X(1)=1$ כפי שאכן צריך להתקיים ולכן לא פספסנו אף אפשרות ולא ספרנו אף אפשרות פעמיים. □

תרגיל 4.9. כאשר מחלקים 4 כדורים לתאים וסופרים את מספר התאים בהם יש זוג כדורים, יכולים להיות לכל היותר שני תאים כאלו, ולכן התומך הינו $supp(X)=\{0,1,2\}$. בדומה לתרגיל 3.25, נניח שהכדורים ממוספרים ומרחב המדגם יהיה בחירה עם סדר ועם החזרה של התאים על ידי הכדורים: 5^4. נחשב את ההסתברות המתאימה לכל ערך בנפרד.

- $X=0$
מאורע זה יכול להתרחש באחת משלוש צורות: כל כדור בתא משלו, יש תא עם שלושה כדורים ותא נוסף עם כדור אחד או שיש תא אחד ובו כל הכדורים. בכל המצבים הללו אין אף תא עם בדיוק שני כדורים ולכן $X=0$. נחשב את מספר האפשרויות לכל מצב:
  - כל כדור בתא משלו: לו היו בדיוק 4 כדורים ו-4 תאים, חלוקת הכדורים לתאים הייתה שקולה לסידור הכדורים בשורה. מאחר ויש תא מיותר, בתור שלב ראשון נבחר מתוך התאים את 4 התאים אליהם יכנסו כדורים ולאחר מכן נסדר את 4 הכדורים לתוך 4 התאים שנבחרו: $\binom{5}{4}\cdot 4!=120$.

- תא עם שלושה כדורים ותא עם כדור בודד: ראשית נבחר מתוך 4 הכדורים שלישייה שתהיה באותו התא. לאחר מכן, נבחר את התא המתאים לה ולבסוף נבחר את התא הנוסף עבור הכדור האחרון שנשאר מתוך 4 התאים שלא נבחרו בשלב הראשון: $\binom{4}{3}\cdot 5\cdot 4=80$.
- כל הכדורים בתא אחד: יש לבחור את התא ומרגע שהוא נבחר אין יותר אפשרויות בחירה – כל הכדורים חייבים להיכנס אליו: 5.

$$Pr\left(X=0\right)=\frac{120+80+5}{625}=\frac{41}{125}$$

- $X=1$

מאורע זה מתרחש רק כאשר יש תא אחד עם שני כדורים. לכן שני הכדורים האחרים צריכים להיות כל אחד בתא משלו ואין אפשרויות נוספות. נבחר אם כך את 2 הכדורים שיהיו זוג, נבחר תא עבור הזוג ולאחר מכן נמשיך את תהליך הבחירה של הכדורים הבודדים: לכדור עם המספר הקטן יותר מבין אלו שנותרו יש 4 אפשרויות ולכדור הנוסף נשארו רק 3.

$$Pr\left(X=1\right)=\frac{\binom{4}{2}\cdot 5\cdot 4\cdot 3}{625}=\frac{72}{125}$$

- $X=2$

מאורע זה יכול להתרחש רק באופן יחיד – שני תאים מלאים, בכל אחד בדיוק שני כדורים. לכן נבחר את זוג הכדורים הראשון, נבחר תא עבורם ואז נבחר עוד תא אליו נכניס את שני הכדורים שלא נבחרו.

$$Pr\left(X=2\right)=\frac{\binom{4}{2}\cdot 5\cdot 4}{625}=\frac{24}{125}$$

נבדוק ונקבל $P_X(0)+P_X(1)+P_X(2)=\frac{137}{125}>1$, כלומר יש טעות באחד החישובים וכנראה בוצעה ספירה כפולה של תוצאות באחד המאורעות. האם ניתן להגיע לאותה תוצאה סופית (אותו סידור כדורים בתאים) במספר דרכים שונות? למשל, נסתכל על הסידור שבו כדורים 1,2 מוכנסים לתא ג' וכדורים 3,4 מוכנים לתא ה', המתאים ל-$\{X=2\}$. בהתאם לתיאור דלעיל, סידור כזה יכול להתקבל כאשר כדורים 1,2 נבחרים בשלב הראשון, תא ג' נבחר עבורם ולאחר מכן תא ה' נבחר עבור שני הכדורים הנותרים. אפשרות נוספת היא שבשלב הראשון של הניסוי נבחר את כדורים 3,4, תא ה' נבחר עבורם ולאחר מכן תא ג' נבחר עבור שני הכדורים הנותרים. בשתי האפשרויות הסידור הסופי שהתקבל זהה לחלוטין אך אפשרויות אלו נספרו כאפשרויות שונות בחישוב של $\{X=2\}$ ולכן יש בחישוב המאורע ספירה מיותרת. ניתן לתקן זאת על ידי חלוקה ב-2 כדי לצמצם את הספירות הכפולות, אבל עדיף למצוא דרך לתאר את הניסוי כך שכל תוצאה תיבחר רק פעם אחת. למשל: נבחר קודם את שני התאים (בלי סדר ובלי החזרה) ולאחר מכן נבחר זוג כדורים שיכנס לתא עם האות הנמוכה יותר ואת זוג הכדורים שלא נבחרו נכניס לתא השני. נקבל:

$$Pr\left(X=2\right)=\frac{\binom{5}{2}\cdot\binom{4}{2}}{625}=\frac{12}{125}$$

ואכן מתקיים $P_X(0)+P_X(1)+P_X(2)=1$.

הבעיה בפתרון הראשון שהמונחים בהם השתמשנו "הזוג הראשון" ו"הזוג השני" אינם מוגדרים היטב שכן כאשר מסתכלים רק על התוצאה הסופית של סידור הכדורים בתאים, לא ניתן לשחזר מי מהם היה הראשון ומי מהם היה השני. שיטה טובה כדי למצוא טעויות מסוג זה היא לדמיין את הסידור הסופי של הכדורים ולבדוק האם ממנו ניתן לשחזר מה נבחר בכל שלב, כאשר ידוע תהליך הבחירה אך לא התוצאות. למשל, אם ניכנס לחדר בו יש תא אחד עם שני כדורים ושני תאים עם כדור אחד, נוכל לשחזר בדיוק מה אירע לפי פרוטוקול הניסוי – אילו כדורים נבחרו להיות הזוג, איזה כדור היה הבא בתור לבחור תא ואיזה כדור הושם בתא אחרון. לעומת זאת, בחדר בו

יש שני זוגות כדורים לא נדע בשיטת הפתרון הראשונה מי מהזוגות הוא "הזוג הראשון" וישנן שתי אפשרויות קבילות. בשיטת הסידור השניה, לעומת זאת, ברור שהזוג שנמצא בתא עם האות הנמוכה יותר הוא הזוג שנבחר ראשון ושוב אין ספירה כפולה. □

תרגיל 4.10. **1.** לפי משפט 4.2 מתקיים $\lim_{t\to\infty} F_X(t) = 1$ ולכן $d = 1$. בנוסף, פונקציית ההתפלגות המצטברת היא פונקציה רציפה. בתוך כל תחום היא מוגדרת כפונקציה רציפה ולכן כדי שתהיה רציפה תמיד צריך לוודא שהגבולות שלה בכל נקודת תפר שווים. למשל, מהדרישה $\lim_{t\to 0^-} F_X(t) = \lim_{t\to 0^+} F_X(t)$ נבנה את המשוואה $a \cdot 0 + b = 0$, כלומר $b = 0$. על מנת שהפונקציה תהיה רציפה גם ב-$t = 2$, הגבולות החד צדדים צריכים להיות שווים כלומר $2a = c \cdot 2^2$ וכדי שהפונקציה תהיה רציפה ב-5 חייב להתקיים גם $c \cdot 5^2 = 1$. לכן $c = \frac{1}{25}$ ו-$a = \frac{2}{25}$.
קל לראות כי יתר התכונות הנדרשות מתקיימות גם כן - הפונקציה עולה, רציפה, והגבולות הם 0 ו-1.

2. לפי ההגדרה, $Pr\left(X \le 5\right) = F_X(5) = 1$.
כדי לחשב את ההסתברות ש-X גדול מ-3, נשתמש במשלים ונהפוך את המאורע ל-X קטן מ-3 המתאים לפונקציית ההתפלגות המצטברת:

$$Pr\left(X > 3\right) = 1 - Pr\left(X \le 3\right) = 1 - F_X(3) = 1 - \frac{1}{25} \cdot 3^2 = \frac{16}{25}$$

כדי לחשב את המאורע $\{2 \le X \le 4\}$ נרשום אותו כחלק מאיחוד של מאורעות שאנחנו יודעים לחשב באמצעות הפה"מ:

$$\{X \le 4\} = \{X < 2\} \cup \{2 \le X \le 4\}$$

ולכן

$$Pr\left(X \le 4\right) = Pr\left(X < 2\right) + Pr\left(2 \le X \le 4\right)$$

למאורעות $\{X < 2\}$ ו-$\{X \le 2\}$ יש אותה ההסתברות (כי ההבדל ביניהם הוא רק המאורע $\{X = 2\}$ בעל ההסתברות 0) ולכן ניתן לחשב גם אי שוויון חזק באמצעות הפה"מ ונקבל:

$$Pr\left(2 \le X \le 4\right) = F_X(4) - F_X\left(2\right) = \frac{16}{25} - \frac{2}{25} \cdot 2 = \frac{12}{25}$$

□

תרגיל 4.11. בתרגיל 4.10 חישבנו את הפה"מ וקיבלנו, לאחר מציאת הקבועים, את הנוסחא הבאה:

$$F_X\left(t\right) = \begin{cases} 0 & t < 0 \\ \frac{2}{25} \cdot t & 0 \le t \le 2 \\ \frac{1}{25} \cdot t^2 & 2 < t \le 5 \\ 1 & t > 5 \end{cases}$$

נגזור ונקבל:

$$f_X(t) = \begin{cases} \frac{2}{25} & 0 < t < 2 \\ \frac{2t}{25} & 2 < t < 5 \\ 0 & t \notin [0, 5] \end{cases}$$

$\square$

תרגיל 4.12. עבור כל פונקציה נבדוק האם היא יכולה לשמש כפונקציית התפלגות מצטברת על ידי כך שנבדוק את גבולותיה ב-$\pm\infty$ והאם היא עולה והאם היא יכולה לשמש כפונקציית צפיפות הסתברות על ידי כך שנבדוק האם היא חיובית ומנורמלת.

1. פונקציה זו היא פונקציית צפיפות. קל לראות שהיא חיובית בכל $\mathbb{R}$ ובנוסף

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\,dx &= \int_{5}^{\infty} 4e^{-4(x-5)}\,dx\\ (u=x-5) &= \int_{0}^{\infty} 4e^{-4u}\,du\\ (eq.\,0.34) &= 1\end{aligned}$$

2. פונקציה זו לא יכולה להיות פונקציית התפלגות מצטברת (הגבול ב-∞ הוא 0) ולא יכולה להיות פונקציית צפיפות כי האינטגרל שלה אינו מתכנס.

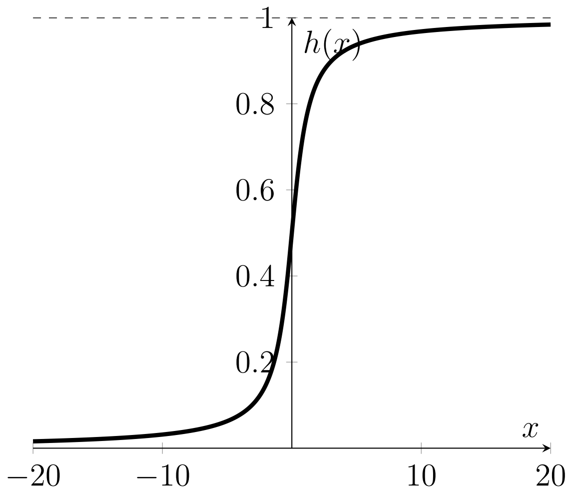
3. פונקציה זו היא פונקציית התפלגות מצטברת. מתקיים $\lim_{x\to\pm\infty}\arctan(x) = \pm\frac{\pi}{2}$, ולכן

$$\lim_{x\to-\infty} h(x) = 0 \qquad \lim_{x\to\infty} h(x) = 1$$

בנוסף,

$$\frac{dh(x)}{dx} = \frac{1}{\pi}\frac{d\,\mathbf{arctan}(x)}{dx} = \frac{1}{\pi}\cdot\frac{1}{1+x^2} \geq 0$$

כלומר הפונקציה עולה ממש, כך שכל הדרישות מפה"מ מתקיימות.



4. הפונקציה לא יכולה לשמש כפונקציית התפלגות מצטברת ולא כפונקציית צפיפות שכן אינה מקיימת כמעט אף תכונה: גבולותיה ב-$\pm\infty$ הם $\pm\infty$ (לכן גם האינטגרל עליה מתבדר) וכן הפונקציה שלילית כאשר $x<0$.

5. פונקציה זו יכולה לשמש כפונקציית צפיפות הסתברות. קל לראות שהפונקציה אי-שלילית ולפי נוסחא 0.36 האינטגרל עליה שווה ל-1, כנדרש.

6. פונקציה זו יכולה לשמש כפונקציית צפיפות הסתברות. קל לראות שהפונקציה היא אי-שלילית ומתקיים:

$$\int_{-\infty}^{\infty} m(x)\,dx = \int_0^2 \frac{15}{16}x^2(2-x)^2\,dx = \ldots = 1$$

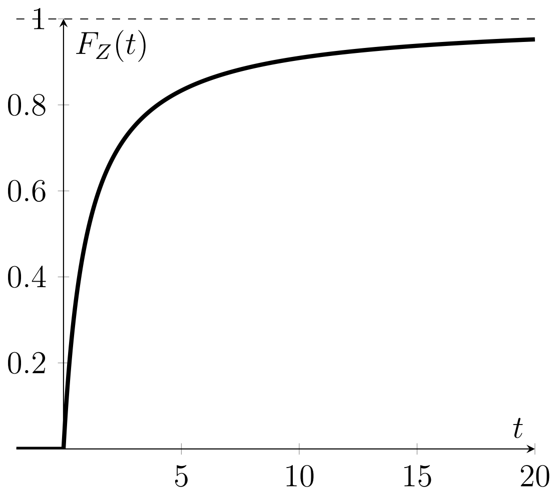
את פרטי החישוב של האינטגרל עצמו ניתן לראות בפתרון של בעיה **4.4**.

$\square$

תרגיל 4.13. 1. על מנת שפונקציה זו תהיה פונקציית התפלגות מצטברת האינטגרל ב-∞ צריך להיות 1:

$$\begin{aligned}\lim_{t\to\infty} F_Z(t) &= \lim_{t\to\infty}\frac{ct}{t+1} = c\cdot\lim_{t\to\infty}\frac{1}{1+\frac{1}{t}}\\ &= c\cdot\frac{1}{1+0} = c\end{aligned}$$

לכן $c = 1$. בנוסף, הגבול ב-$-\infty$ הוא אכן 0 ועל מנת לראות שהפונקציה אכן לא יורדת בכל תחום הגדרתה נמצא בסעיף הבא את הנגזרת ונראה שהיא חיובית תמיד.



2. נגזור עבור $t > 0$:

$$\begin{aligned}f_Z(t) &= \frac{dF_Z(t)}{dt} = \frac{d}{dt}\left(\frac{t}{t+1}\right) = \frac{(t+1)-t}{(t+1)^2}\\ &= \frac{1}{(t+1)^2}\end{aligned}$$

כלומר

$$f_Z(t) = \begin{cases}\frac{1}{(t+1)^2} & t \geq 0\\ 0 & t < 0\end{cases}$$

3. נסמן ב-A את המאורע שיוני אוכל פיצה תוך פחות מחצי שעה מרגע ההחלטה וב-B את המאורע שיצא עץ. נחשב את ההסתברות של A לפי נוסחת ההסתברות השלמה, בהתאם לתוצאת המטבע. בהינתן שיצא עץ, מאורע זה הינו $X \leq 0.5$ ובהינתן שיצא פלי מאורע זה הינו $Z \leq 0.5$. לכן:

$$\begin{aligned} Pr\left(A\right) &= Pr\left(A|B\right) Pr\left(B\right) + Pr\left(A|\overline{B}\right) Pr\left(\overline{B}\right) \\ &= Pr\left(X \leq 0.5\right) \cdot \frac{1}{2} + Pr\left(Z \leq 0.5\right) \cdot \frac{1}{2} \\ &= \frac{1}{2} \cdot F_X(0.5) + \frac{1}{2} \cdot F_Z(0.5) = 0.218 \end{aligned}$$

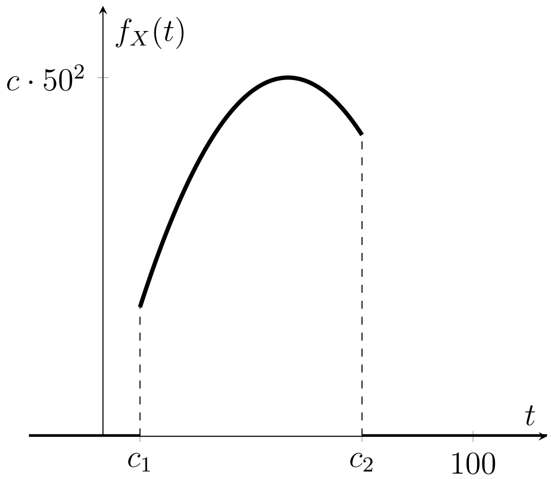
4. יוני הזמין פיצה רק אם במטבע יצא עץ, כלומר השאלה שקולה לשאלה מה ההסתברות שיצא עץ בהינתן שמאורע A התרחש. בסעיף הקודם חישבנו את ההסתברות של המאורע A וכן את ההסתברות של A בהינתן B, לכן ניתן לחשב את ההסתברות המבוקשת באמצעות כלל בייס (נוסחא **2.21**):

$$Pr\left(B|A\right) = \frac{Pr\left(A|B\right) Pr\left(B\right)}{Pr\left(A\right)} = \frac{F_X(0.5) \cdot \frac{1}{2}}{0.218} = 0.237$$

שאלות מהסוג הזה נפגוש הרבה בחיים האמיתיים: אנחנו מודדים תופעת טבע מסויימת שאיננו בטוחים כיצד היא מתפלגת ועל סמך תצפיות ספורות צריכים להעריך מהי ההתפלגות ממנה היא נלקחה.

□

תרגיל 4.14. **1.** הפונקציה $t(100-t)$ חיובית עבור $t \in [0,100]$ ושלילית בשאר הישר הממשי. על מנת ש-$f_X(t)$ תהיה פונקציית צפיפות היא חייבת להיות חיובית בכל תחום הגדרתה, לכן בהכרח $0 \leq c_1 \leq c_2 \leq 100$.
לא ייתכן כי $c_1 = c_2$ שכן אז הפונקציה היא תמיד 0 פרט לנקודה אחת, הנקודה $t = c_1 = c_2$ ואינטגרל על פונקציה כזו יהיה תמיד אפס (נקודה לא משפיעה על אינטגרל). לכן בהכרח $c_1 \neq c_2$, כך שלמעשה הדרישה היא $0 \leq c_1 < c_2 \leq 100$. שרטוט לדוגמא עבור $c_1 = 10,\ c_2 = 70$:



2. עם הקביעה $c_1 = 0, c_2 = 100$ דרישת החיוביות מולאה וכל שנדרש הוא לוודא שהצפיפות מנורמלת:

$$1 = \int_{-\infty}^{\infty} f_X(t)\,dt = \int_0^{100} ct(100-t)\,dt$$
$$= c\left(100\cdot\frac{t^2}{2} - \frac{t^3}{3}\right)_0^{100} = \frac{500000}{3}c$$

כלומר $c = \frac{3}{500000} = 6\cdot 10^{-6}$.

3. בסעיף הקודם כבר חישבנו את הפונקציה הקדומה של הצפיפות בתחום $t \in [0,100]$ לכן פונקציית ההתפלגות המצטברת תהא:

$$F_X(t) = \begin{cases} 0 & t<0 \\ 6\cdot 10^{-6}\left(50t^2 - \frac{t^3}{3}\right) & t\in[0,100] \\ 1 & t>100 \end{cases}$$

4. ההסתברות להיכשל במבחן הינה $Pr\left(X<60\right) = F_X(60) = 0.648$ ואילו הסיכוי להצטיין בקורס הינו $Pr\left(X>90\right) = 1-F_X(90) = 0.028$

5. סופי עברה את המבחן, כלומר ידוע שהמאורע $X>60$ התרחש ולכן ההסתברות המבוקשת תהיה:

$$Pr\left(X>90|X>60\right) = \frac{Pr\left(X>90\cap X>60\right)}{Pr\left(X>60\right)} = \frac{Pr\left(X>90\right)}{Pr\left(X>60\right)}$$
$$= \frac{1-F_X(90)}{1-F_X(60)} = \frac{0.028}{0.352} = 0.08$$

6. התלמידים צודקים. ניתן לראות זאת בשתי דרכים. ראשית, קל לראות שההתפלגות סימטרית סביב 50 ולכן החציון הוא 50, כך שיותר מחצי כיתה קיבלה מתחת ל-60. דרך אפשרית נוספת, $F_X(60) = 0.648$ כלומר 64.8% מהכיתה נכשלו וכאמור, יותר מחצי.

□

תרגיל 4.15. המאורעות יהיו תלויים. אם המטבע נפל בהטלה הראשונה על עץ, סביר להניח שהסיכוי ליפול על עץ גבוה יותר (אם היה סיכוי גבוה מאוד ליפול על פלי, הוא היה נופל על פלי) כך שהסיכוי ליפול על עץ גם בהטלה השנייה גדול יותר. נחשב זאת מפורשות באמצעות נוסחת ההסתברות השלמה (נוסחא 4.16):

$$Pr\left(A\right) = \int_0^1 Pr\left(A|X=x\right)f_X\left(x\right)\,dx = \int_0^{0.5} x\cdot 12x^2\,dx + \int_{0.5}^1 x\cdot 12(x-1)^2\,dx = \frac{1}{2}$$

ניתן היה להגיע לתשובה זו גם מטעמי סימטריה, שכן פונקציית הצפיפות סימטרית סביב $x=\frac{1}{2}$. באותה צורה, $Pr\left(B\right)=\frac{1}{2}$. נחשב את הסתברות החיתוך. בהינתן הערך של $X=x$, הסיכוי שהמטבע יפול פעמיים על עץ הינו x^2 ולכן:

$$Pr\left(A\cap B\right) = \int_0^1 Pr\left(A\cap B|X=x\right)f_X\left(x\right)dx = \int_0^{0.5} x^2\cdot 12x^2\,dx + \int_{0.5}^1 x^2\cdot 12(x-1)^2 dx = 0.275$$

ואכן $Pr\left(B\right) \cdot Pr\left(A\right) = 0.25 \neq 0.275 = Pr\left(A\cap B\right)$ כך שהמאורעות תלויים. בהינתן שההטלה הראשונה הייתה עץ, ניתן לפסול ערכי X מאוד נמוכים מה שמעלה את ההסתברות שההטלה השנייה יהיה עץ, והסתברות זו הינה

$$Pr\left(B|A\right) = \frac{Pr\left(A\cap B\right)}{Pr\left(A\right)} = \frac{0.275}{0.5} = 0.55$$

□

תרגיל 4.16. נסמן ב-Y את הציון אחרי הפקטור. הציון של סטודנטים שקיבלו עד 50 לא השתנה ולכן לכל $t < 50$ מתקיים $F_Y(t) = F_X(t)$. באופן דומה, לכל $t > 70$ הסיכוי לקבל פחות מ-t לפני ואחרי הפקטור נותר כשהיה, שכן הפקטור משפיע רק על החלוקה הפנימית של הצפיפות שבין 50 ל-70 ולא על סך הצפיפות שבתחום זה. לכן גם לכל $t > 70$ מתקיים $F_Y(t) = F_X(t)$.
נשאר לבדוק מה השתנה. לאחר הפקטור, לא ניתן לקבל ציונים בתחום $60 - 50$ לכן לכל $t \in [50, 60]$ מתקיים

$$F_Y(t) = Pr\left(Y \leq t\right) = Pr\left(Y \leq 50\right) + \cancelto{0}{Pr\left(50 < Y \leq t\right)} = F_X(50)$$

ולכל $t \in [60, 70]$ ניתן לקבל ציון קטן מ-t בשתי אפשרויות: או שהציון המקורי קטן מ-$t - 10$ או שהציון המקורי היה בין 60 ל-t. למעשה, מבין כל הציונים המקוריים הקטנים מ-t, רק ציונים שהיו בין $t - 10$ לבין 60 אחרי הפקטור חרגו מהתחום והם גדולים מ-t. לכן:

$$\begin{aligned} F_Y(t) &= Pr\left(Y \leq t\right) = Pr\left(X \leq t - 10 \cup 60 < X \leq t\right) \\ &= F_X(t-10) + F_X(t) - F_X(60) \end{aligned}$$

נסכם הכל בנוסחא אחת ונציב:

$$\begin{aligned} F_Y(t) &= \begin{cases} F_X(t) & t < 50 \\ F_X(50) & t \in [50, 60] \\ F_X(t-10) + F_X(t) - F_X(60) & t \in [60, 70] \\ F_X(t) & t > 70 \end{cases} \\ &= \begin{cases} 0 & t < 0 \\ 6 \cdot 10^{-6}\left(50t^2 - \frac{t^3}{3}\right) & t \in [0, 50] \cup [70, 100] \\ 0.5 & t \in [50, 60] \\ -4 \cdot 10^{-6}t^3 + 6.6 \cdot 10^{-4}t^2 - 6.6 \cdot 10^{-3}t - 0.616 & t \in [60, 70] \\ 1 & t > 100 \end{cases} \end{aligned}$$

□

תרגיל 4.17. נסמן ב-M את הציון הכי גבוה מבין הסטודנטים, ב-m את הציון הכי נמוך וב-X_i את הציון של הסטודנט ה-$i \in \{1, \ldots, 10\}$. נחשב את הפה"מ של M:

$$F_M(t) = Pr\left(M \leq t\right) = Pr\left(\max_i \{X_i\} \leq t\right)$$
$$(*) = Pr\left(X_1 \leq t \cap \ldots \cap X_{10} \leq t\right) \overset{(**)}{=} Pr\left(X_1 \leq t\right) \cdot \ldots \cdot Pr\left(X_{10} \leq t\right)$$
$$(***) = F_X(t) \cdot \ldots \cdot F_X(t) = \left(F_X(t)\right)^{10} = \begin{cases} 0 & t < 0 \\ 6 \cdot 10^{-5}\left(50t^2 - \frac{t^3}{3}\right)^{10} & t \in [0, 100] \\ 1 & t > 100 \end{cases}$$

כאשר במעבר $(*)$ השתמשנו בכך שאם המקסימום של מספר ערכים קטן מ-t, כל אחד מהם צריך להיות קטן מ-t, במעבר $(**)$ השתמשנו בחוסר התלות בין הסטודנטים וב-$(***)$ השתמשנו בכך שכולם מקבלים ציונים לפי ההתפלגות של X מתרגיל **4.14**. באותו אופן, נחשב את הפה"מ של המינימום רק שהפעם כדי להשתמש במעבר דוגמת $(*)$ עלינו לחסום את המינימום מלרע ולא מלעיל:

$$F_m(t) = Pr\left(m \leq t\right) = 1 - Pr\left(m > t\right) = 1 - Pr\left(\min_i \{X_i\} > t\right)$$
$$= 1 - Pr\left(X_1 > t \cap \ldots \cap X_{10} > t\right) = 1 - Pr\left(X_1 > t\right) \cdot \ldots \cdot Pr\left(X_{10} > t\right)$$
$$= 1 - \left(1 - F_X(t)\right)^{10}$$

ומפאת הסרבול הכרוך בכך, נמנענו מהצבת הנוסחא המפורשת של הפה"מ של X בנוסחא. $\square$

תרגיל 4.18. נתחיל בהסבר האינטואיטיבי אודות התפלגות המקסימום והמינימום של אינסוף הטלות קוביה. בהתאם למשפט הקוף המקליד (משפט **3.4**), ככל שנחזור על הטלת הקוביה יותר פעמים הסיכוי שלפחות פעם אחת ייצא 6 יגדל וישאף ל-1. היות ו-6 הוא הערך המרבי שיכול לצאת בקוביה, מרגע שהוא יצא המקסימום הוא 6 ללא תלות בשאר ההטלות. לכן, לאחר אינסוף הטלות קוביה בהסתברות 1 המקסימום יהיה 6. מאותו נימוק, המינימום יהיה 1. ניתן להראות זאת מתמטית גם על ידי שימוש בנוסחאות שפותחו בבעיה **4.3** ובתרגיל **4.22**.

נעבור לחישוב המתמטי עבור הנתונים שבשאלה. היות ו-$f_X(t) > 0$ בתחום $t \in [a, b]$ הרי שפונקציית ההתפלגות המצטברת מקיימת לכל $t \in (a, b)$:

$$F_X(t) = \int_a^t f_X(x)\, dx < \int_a^b f_X(x)\, dx = 1$$

כלומר עבור כל $t < b$ מתקיים $F_X(t) < 1$ ולכן $F_M(t) = \lim_{n \to \infty}\left(F_X(t)\right)^n = 0$ ועבור $t \geq b$ מתקיים $F_X(t) = 1$ ולכן גם $F_M(t) = 1$. לפיכך, לפונקציית ההתפלגות המצטברת של המקסימום יש קפיצה בגודל 1 בנקודה b והיא קבועה על שאר הישר הממשי, כך ש-M הוא משתנה מקרי בדיד המקבל את הערך b בהסתברות 1. לפי אותו החישוב, המשתנה המקרי m הוא משתנה מקרי בדיד המקבל את הערך a בהסתברות 1.

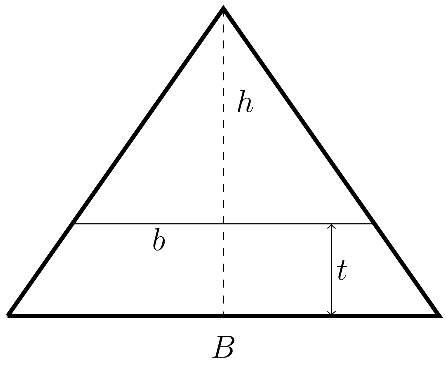
הרעיון העומד מאחורי תוצאה זו זהה לרעיון העומד מאחורי הטלת הקוביה. ככל שמגרילים יותר ערכים מההתפלגות של X, כך גובר הסיכוי שנקבל ערך קרוב מאוד מאוד ל-b. לכן המקסימום של אינסוף הטלות לא יכול להיות אף ערך קטן ממש מ-b והערך b יהיה המקסימום. $\square$

תרגיל 4.19. נחשב את פונקציית ההתפלגות המצטברת, $F_X(t) = P(X \leq t)$. בבירור, המרחק מהבסיס יכול להיות בין 0 לבין גובה המשולש, h, לכן נתמקד בתחום זה. עבור כל $t \in [0, h]$, נחלק את המשולש למשולש עליון ולטרפז תחתון על ידי קו המקביל לבסיס שמרחקו מהבסיס הוא t. המאורע שנבחרה נקודה שמרחקה קטן

t-מ הוא המאורע שנבחרה נקודה בתוך הטרפז. מאחר והנקודה נבחרת באקראי במשולש, ההסתברות שתיבחר נקודה בתוך הטרפז מתוך כל המשולש הוא היחס בין שטחיהם,

$$Pr(X \leq t) = \frac{\frac{t(B+b)}{2}}{\frac{Bh}{2}} = \frac{t\left(1+\frac{b}{B}\right)}{h}$$

כאשר B הוא בסיס המשולש המקורי ו-b הוא הבסיס של המשולש הקטן (הישר המקביל שהוספנו – הצלע העליונה של הטרפז).



זו עדיין לא התשובה הסופית שכן היא כוללת פרמטרים שהוספנו ולא הופיעו בשאלה המקורית. נותר להביע את היחס $\frac{b}{B}$ באמצעות נתוני השאלה. מדמיון משולשים נקבל $\frac{h-t}{b} = \frac{h}{B}$ ולכן $\frac{h-t}{h} = \frac{b}{B}$. נציב זאת בנוסחא להתפלגות המצטברת ונקבל:

$$F_X(t) = \begin{cases} 0 & t < h \\ 1 - \frac{(h-t)^2}{h^2} & t \in [0,h] \\ 1 & t > h \end{cases}$$

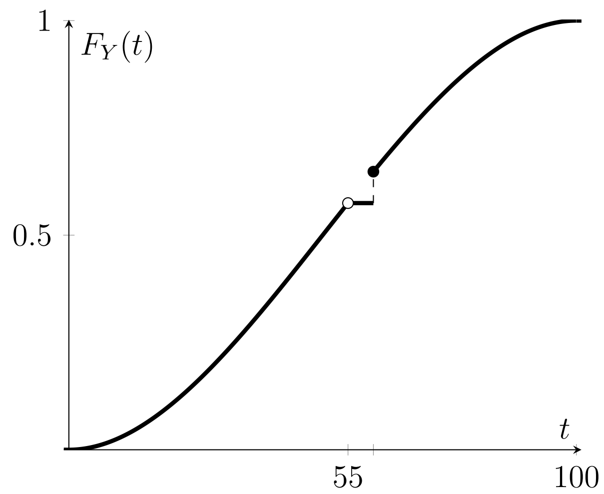
פונקציית הצפיפות תתקבל על ידי גזירה:

$$f_X(t) = \begin{cases} \frac{2(h-t)}{h^2} & t \in [0,h] \\ 0 & t \notin [0,h] \end{cases}$$

□

תרגיל 4.20. כמו בתרגיל **4.16**, הפה"מ ישתנה רק בתחום $[55,60]$, שכן הסיכוי לקבל ציונים מתחת ל-55 לא משתנה והסיכוי לקבל ציונים מעל 60 לא משתנה. ההבדל היחיד הוא שכל ההסתברות שהייתה קודם לקבל ציון בין 55 ל-60 מרוכזת כעת ב-60. לכן, בתחום זה פונקציית ההתפלגות המצטברת תהיה קבועה ושווה ל-$F_X(55)$:

$$F_Y(t) = \begin{cases} 0 & t < 0 \\ 6 \cdot 10^{-6}\left(50t^2 - \frac{t^3}{3}\right) & t \in [0, 55) \\ F_X(55) = 0.575 & t \in [55, 60) \\ 6 \cdot 10^{-6}\left(50t^2 - \frac{t^3}{3}\right) & t \in [60, 100) \\ 1 & t \geq 100 \end{cases}$$



□

תרגיל 4.21. נסמן ב-X את הניקוד ביריה בודדת. בהתאם לניקוד המתואר בשאלה, לא ניתן לקבל ניקוד שלילי אלא רק ניקוד בתחום $\{0\} \cup [1, 10]$. מקבלים 0 נקודות כאשר פוגעים מחוץ למטרה, כלומר

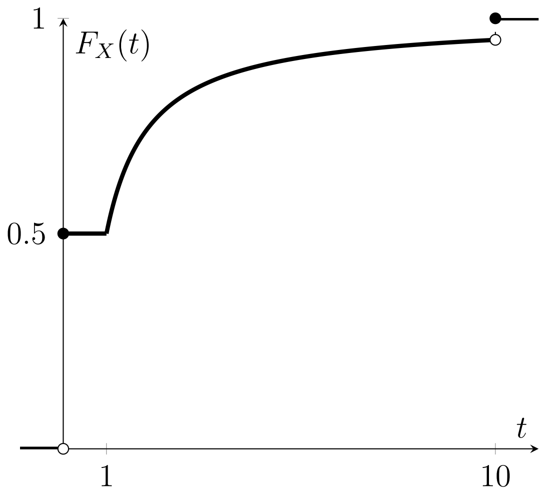
$$Pr\left(X = 0\right) = Pr\left(R > 10\right) = \int_{10}^{20} \frac{1}{20}\, dt = \frac{1}{2}$$

לכל ניקוד בתחום $t \in [1, 10)$ פונקציית ההתפלגות המצטברת תהיה:

$$\begin{aligned} F_X(t) &= Pr\left(X \leq t\right) = Pr\left(\frac{10}{R} \leq t\right) \\ &= Pr\left(\frac{10}{t} \leq R\right) = \int_{\frac{10}{t}}^{20} \frac{1}{20}\, du = \frac{1}{20}\left(20 - \frac{10}{t}\right) \\ &= 1 - \frac{1}{2t} \end{aligned}$$

ולסיכום:

$$F_X(t)=\begin{cases}0 & t<0\\ \frac{1}{2} & t\in[0,1)\\ 1-\frac{1}{2t} & t\in[1,10)\\ 1 & t\geq 10\end{cases}$$



□

תרגיל 4.22. השאלה דומה לבעיה 4.3 ובהתאם לכך גם התומך של Y הינו $supp(Y)=\{1,2,\ldots,6\}$ ומרחב המדגם יהיה הטלה של קוביה 3 פעמים: 6^3. גם הפעם, חישוב ישיר של פונקציית ההסתברות הוא בעייתי, שכן אם המינימום הוא 3 אנחנו יודעים ש-3 צריך לצאת לפחות פעם אחת, אך לא ידוע כמה פעמים ובאילו הטלות ופרט לכך יש לדאוג לכך שבשאר ההטלות יצאו מספרים גדולים מ-3. לכן נשתמש באותה שיטה כמו בבעיה 4.3 ונחשב את הפה"מ.

המאורע $\{Y\leq k\}$ מתאר מצב בו לפחות אחת הקוביות נפלה על תוצאה קטנה מ-k ולכך יש אפשרויות רבות. לעומת זאת, המשלים, $\{Y>k\}$, הוא פשוט יותר ומשמעו שכל ההטלות גדולות ממש מ-k (שכן אחרת, המינימום היה גם קטן מ-k). לאור זאת, נחשב את הפה"מ באמצעות המשלים:

$$\begin{aligned}F_Y(k)&=Pr\left(Y\leq k\right)=1-Pr\left(Y>k\right)\\&=1-\left(\frac{6-k}{6}\right)^3\end{aligned}$$

שכן יש $6-k$ תוצאות הגדולות ממש מ-k. לכן פונקציית ההסתברות תהיה

$$\begin{aligned}P_Y(k)&=F_Y(k)-F_Y\left(k-1\right)\\&=1-\left(\frac{6-k}{6}\right)^3-\left(1-\left(\frac{6-(k-1)}{6}\right)^3\right)\\&=\frac{(7-k)^3-(6-k)^3}{6^3}\end{aligned}$$

כדאי לשים לב להבדל בין בעיה 4.3 לתרגיל זה. בבעיה 4.3 חישבנו את המקסימום ולצורך כך חסמנו אותו מלעיל על ידי חישוב המאורע $\{X \leq k\}$. חישוב זה היה פשוט שכן אם המקסימום קטן מ-k אז בהכרח כל המספרים קטנים מ-k. בתרגיל זה חישבנו את המינימום לצורך כך השתמשנו היה נוח לחסום את המינימום מלרע על ידי חישוב המאורע $\{Y \geq k\}$ (למעשה $\{Y \geq k+1\} = \{Y > k\}$ כך שזה אותו הדבר). חישוב זה היה פשוט שכן אם המינימום גדול מ-k אז בהכרח כל המספרים גדולים מ-k. מסיבה זו נסיק שבאופן כללי תמיד כדאי לחסום את המקסימום מלמעלה ואת המינימום מלמטה. דוגמא נוספת כזו ראינו בהתפלגות המקסימום והמינימום (משפט 4.6). □

תרגיל 4.23. מחלקים מחדש את הזוגות ולכן מספר הזוגות המקוריים שרוקדים אחד עם השני יכול להיות בין 0 ל-3 ומספר הזוגות של הרקדנים החובבים הרוקדים זה עם זה יכול להיות בין 0 ל-2. יחד עם זאת, לא יתכן ש-2 זוגות בדיוק רוקדים זה עם זה, שכן עבור שני האנשים האחרים (שהיו זוג במקור) נותרת רק אפשרות אחת – לרקוד זה עם זו ולהיות זוג נוסף. לכן $supp(X) = \{0,1,3\}$ ו-$supp(Y) = \{0,1,2\}$. לצורך ביצוע הניסוי נניח שהגברים עומדים בשורה ועלינו לשדך אליהם את הנשים, כלומר מרחב המדגם הוא סידור בשורה של 3 הנשים למול הגברים איתם ירקדו – $3!$ אפשרויות. לשם הפשטות נסמן את הזוגות ב-א׳,ב׳ (החובבים) ו-ג׳ (המקצועיים).

- $X = 0$
במקרה זה אף זוג לא רוקד יחד. לכן לאישה א׳ יש 2 אפשרויות שאינן בן זוגה המקורי. נניח שנבחר גבר ב׳. לאישה ב׳ יש שתי אפשרויות, אבל אם תיבחר בגבר א׳ אז לאישה ג׳ תהיה אפשרות אחת – גבר ג׳ ויהיה זוג שירקוד יחד. לכן אישה ב׳ חייבת לרקוד עם גבר ג׳ ואישה ג׳ חייבת לרקוד עם גבר א׳. כלומר, מהרגע שנבחר בן הזוג של אישה א׳, נבחרו כל בני הזוג באופן יחיד ולסיכום:

$$Pr\left(X=0\right) = \frac{2}{3!} = \frac{1}{3}$$

- $X = 1$
במקרה זה ישנו זוג אחד שירקוד ביחד. יש 3 אפשרויות לבחור את הזוג שרוקד ביחד. לאחר מכן, נותרו שני זוגות שצריך לשדך מחדש ויש רק דרך אחת לעשות זאת. לכן:

$$Pr\left(X=1\right) = \frac{3}{3!} = \frac{1}{2}$$

- $X = 3$
במקרה זה כל הזוגות המקוריים רוקדים זה עם זו ויש רק אפשרות אחת כזו:

$$Pr\left(X=3\right) = \frac{1}{3!} = \frac{1}{6}$$

- $Y = 0$
אף אחד מהרקדנים החובבים לא רוקד זה עם זו. לכן לאישה א׳ יש 2 אפשרויות. אם נבחר עבורה גבר ב׳, אז בוודאות אף זוג של רקדנים חובבים לא רוקד ביחד, ולכן כל אחד מ-$2!$ הסידורים האפשריים של שתי הנשים האחרות מתאים למאורע זה. לעומת זאת, אם נבחר גבר ג׳, אז לאישה ב׳ חייבת לרקוד עם גבר א׳ כדי שלא יהיו זוגות של רקדנים חובבים. שתי אפשרויות אלו מתארות מאורעות זרים ולכן ההסתברות במקרה זה תהיה:

$$Pr\left(Y=0\right) = \frac{2!+1}{3!} = \frac{1}{2}$$

- $Y=1$
במקרה זה יש זוג אחד של רקדנים חובבים שרוקדים זה עם זו. לצורך כך נבחר את הזוג (יש שתי אפשרויות), ואת הרקדנית החובבת השנייה חייבים לשדך לרקדן המקצועי. לכן, מרגע שנבחר איזה זוג רקדנים חובבים רוקד יחד, שאר החלוקה נקבעת באופן יחיד וההסתברות תהיה:

$$Pr\left(Y=1\right)=\frac{2}{3!}=\frac{1}{3}$$

- $Y=2$
מקרה זה מתאים לכך ששני זוגות הרקדנים החובבים רוקדים יחד ולזוג המקצועי אין ברירה אלא לרקוד זה עם זו, כלומר ישנה רק אפשרות אחת וההסתברות היא:

$$Pr\left(Y=2\right)=\frac{1}{3!}=\frac{1}{6}$$

לסיכום:

$$P_X(k)=\begin{cases}\frac{1}{3} & k=0\\ \frac{1}{2} & k=1\\ \frac{1}{6} & k=3\end{cases}\quad P_Y(k)=\begin{cases}\frac{1}{2} & k=0\\ \frac{1}{3} & k=1\\ \frac{1}{6} & k=2\end{cases}$$

קל לראות שאכן שתי פונקציות ההסתברות מסתכמות לאחד. □

תרגיל 4.24. במסיבה יש 50 זוגות המהווים 100 אנשים ובוחרים באקראי 10 מתוכם. לכן מרחב המדגם כולל בחירה של 10 אנשים מתוך 100 ללא החזרה וללא סדר: $\binom{100}{10}$. על מנת לבחור את הזוג סמית' חייבים לבחור את שניהם ובנוסף אליהם יש לבחור עוד 8 אנשים מתוך 98 האנשים האחרים, לכן ההסתברות שהזוג סמית' נבחר הינה

$$\frac{\binom{98}{8}}{\binom{100}{10}}=\frac{1}{110}$$

נסמן ב-X את מספר הזוגות שנבחרו בשלמותם ונחשב את התפלגותו. כאשר בוחרים 10 אנשים ניתן לבחור בין 0 ל-5 זוגות, לכן $supp(X)=\{0,\ldots,5\}$. כמקודם, מרחב המדגם הינו $\binom{100}{10}$. המאורע $\{X=k\}$ מתרחש כאשר בדיוק k זוגות נבחרים. לצורך כך יש לבחור את kהזוגות מתוך 50 הזוגות ולבחור את שני בני הזוג. נותר לבחור עוד $10-2k$ אנשים, מבלי לבחור אף זוג נוסף בשלמותו. ניתן לוודא זאת על ידי כך שנבחר מתוך $50-k$ הזוגות שנותרו $10-2k$ זוגות ומכל אחד מהם נבחר כנציג רק את אחד משני בני הזוג. באופן זה, בוודאות לא נבחר בטעות שני בני זוג. לסיכום:

$$P_X(k)=Pr\left(X=k\right)=\frac{\binom{50}{k}\cdot\binom{50-k}{10-2k}\cdot 2^{10-2k}}{\binom{100}{10}}$$

□

תרגיל 4.25. נתחיל במציאת התומכים. יש 6 נשים ולכן יהיו לכל היותר 3 משחקים בין נשים ולכן $supp(X)=\{0,1,2,3\}$. יש 10 גברים לכן יהיו לכל היותר 5 משחקים בין גברים. מצד שני, מאחר ויש רק 6 נשים, גם אם כל אישה תשחק עם גבר, עדיין יישארו 4 גברים שיצטרכו לשחק זה עם זה, לכן לכל הפחות יהיו 2 משחקים בין גברים: $supp(Y)=\{2,3,4,5\}$. לאור כל אלו, מספר המשחקים בין גברים לנשים יהיה לכל היותר 6 ולכל

הפחות 0. יחד עם זאת, מראש יש 6 נשים וכל זוג נשים שמשחק עם עצמו מותיר את מספר הנשים שלא משחקות עם אישה על מספר זוגי, לכן מספר המשחקים בין נשים לגברים צריך להיות זוגי: $supp(Z) = \{0, 2, 4, 6\}$.

מרחב המדגם הינו חלוקה של 16 האנשים לזוגות. בחירה של זוג היא בחירה ללא החזרה וללא סדר (חשובים שני האנשים שנבחרו, לא הסדר בו הם נבחרו), לכן יש $\binom{16}{2}$ אפשרויות לבחירת הזוג הראשון, $\binom{14}{2}$ אפשרויות לזוג השני, $\binom{12}{2}$ לשלישי וכן הלאה. באופן זה חולקו 16 השחקנים לזוגות, אך תהליך זה בוצע עם סדר בין הזוגות. למשל, אם הזוג א-ב נבחר בשלב הראשון והזוג ג-ד נבחר בשלב השני, זה זהה לחלוטין לסידור בו הזוג ג-ד נבחר בשלב הראשון והזוג א-ב בשלב השני. כדי לנטרל את הסדר המלאכותי של הזוגות שהכנסנו, נחלק בסידורים האפשריים של 8 הזוגות שהתקבלו ונקבל שגודלו של מרחב המדגם הינו:

$$\frac{\binom{16}{2}\binom{14}{2}\cdot\ldots\cdot\binom{2}{2}}{8!}$$

דרך שקולה ולעיתים יותר נוחה לייצר את מרחב המדגם היא על ידי סידור האנשים בשורה. נקבע שנסדר את האנשים בשורה וכל שחקן שעומד במקום אי זוגי ישחק עם השחקן שעומד במקום הזוגי הבא בתור (1 עם 2, 3 עם 4 וכן הלאה). יש $16!$ אפשרויות לסדר את האנשים בשורה. יחד עם זאת, סידור פנימי של כל זוג לא משנה (1 ישחק עם 2 גם אם שני האנשים העומדים שם יתחלפו) ולכן עבור כל זוג נחלק בסידורים הפנימיים ובסך הכל נחלק ב-$(2!)^8$. בנוסף, בדומה לאפשרות הראשונה נחלק גם בסידורים הפנימיים של הזוגות ונקבל שמרחב המדגם הינו

$$\frac{16!}{2^8\cdot 8!}$$

נחשב את ההתפלגות של כל אחד מהמשתנים המקריים. המאורע $\{X = k\}$ מתקיים כאשר $2k$ נשים משחקות זו עם זו והיתר משחקות עם גברים (ושאר הגברים משחקים עם גברים אחרים). נחשב כל אחת מהקבוצות בנפרד. עבור הנשים, עלינו לבחור את $2k$ הנשים שישחקו זו עם זו ואז לחלקן לזוגות באופן דומה בו חושב מרחב המדגם (אלו הסוגריים הראשונים בנוסחא הבאה). נותרו $6-2k$ נשים ועלינו לבחור להן יריבים מתוך 10הגברים. נסדר את הנשים שנותרו בסדר כלשהו (למשל, אלפבתי) ונבחר $6-2k$ גברים שישחקו מולן. חלוקת היריבים תתבצע על ידי סידור הגברים בשורה. נותרנו עם $10-(6-2k) = 2k+4$ גברים שצריכים לשחק זה עם זה ונחלקם באופן דומה לזה שנעשה במרחב המדגם (אלו הסוגריים האחרונים בנוסחא):

$$\begin{aligned} Pr\left(X=k\right) &= \frac{\left(\binom{6}{2k}\cdot\frac{(2k)!}{2^k\cdot k!}\right)\cdot\left(\binom{10}{6-2k}\cdot(6-2k)!\right)\cdot\left(\frac{(2k+4)!}{2^{k+2}\cdot(k+2)!}\right)}{\frac{16!}{2^8\cdot 8!}} \\ &= \frac{6!\cdot 10!}{2^{2k+2}\cdot k!\cdot(k+2)!\cdot(6-2k)!} \end{aligned}$$

מאחר ומרחב המדגם קטן יחסית ניתן היה לוותר על הפישוט ובמקום זאת להציב את הערכים האפשריים לקבלת פונקציית ההסתברות:

$$P_X(k) = \begin{cases} 0.224 & k=0 \\ 0.559 & k=1 \\ 0.210 & k=2 \\ 0.007 & k=3 \end{cases}$$

נמשיך לחשב את ההתפלגויות של Y, Z. ניתן לבצע את החישוב בדומה לחישוב עבור X, אך מדובר על משתנים מקריים הקשורים לאותו הניסוי ולכן ניתן למצוא קשר ביניהם ולפשט את העבודה תוך ניצול העבודה שאת ההתפלגות של X כבר מצאנו. אם מתנהלים X משחקים בין נשים, אזי יש $6-2X$ נשים שחייבות לשחק עם גברים ולכן מספר המשחקים בין נשים לגברים הינו $Z=6-2X$. גברים אלו לא יכולים לשחק עם גברים אחרים ולכן הגברים שנותרו משחקים בינם לבין עצמם ומספר המשחקים בין גברים הינו $Y=\frac{10-(6-2X)}{2}=X+2$. מכאן שפונקציות ההסתברות המתאימות תהיינה:

$$
\begin{aligned}
P_Y(k) &= Pr\left(Y=k\right)=Pr\left(X+2=k\right)\\
&= Pr\left(X=k-2\right)=\begin{cases}0.224 & k=2\\ 0.559 & k=3\\ 0.210 & k=4\\ 0.007 & k=5\end{cases}
\end{aligned}
$$

וכן

$$
\begin{aligned}
P_Z(k) &= Pr\left(Z=k\right)=Pr\left(6-2X=k\right)\\
&= Pr\left(X=\frac{6-k}{2}\right)=\begin{cases}0.224 & k=6\\ 0.559 & k=4\\ 0.210 & k=3\\ 0.007 & k=2\end{cases}
\end{aligned}
$$

□

תרגיל 4.26. ייתכנו לכל היותר 3 מועדים לכן $supp(X)=\{1,2,3\}$. כאמור, $X=1$ מתרחש כאשר כולם עוברים במועד א', כלומר

$$Pr\left(X=1\right)=0.5^{100}$$

המאורע $X=2$ מתרחש כאשר לפחות סטודנט אחד נכשל במועד א', אבל כל אלו שנכשלו במועד א' עוברים במועד ב'. מאחר ומספר הנכשלים במועד א' יכול להיות כל מספר, נפרק מאורע זה לאיחוד של המאורעות הזרים בזוגות A_k– בדיוק k סטודנטים נכשלו במועד א' וכולם עברו במועד ב'. ההסתברות של כל מאורע כזה היא

$$Pr\left(A_k\right)=\binom{100}{k}\cdot 0.5^k\cdot 0.5^{100-k}\cdot 0.6^k$$

שכן יש לבחור את k הסטודנטים שיכשלו, לדרוש שהם אכן יכשלו (כל אחד בסיכוי 0.5), לדרוש שכל האחרים יעברו, וכן לדרוש שאלו שנכשלו יעברו במועד ב' (כל אחד בסיכוי 0.6) לפיכך

$$\begin{aligned}
Pr\left(X=2\right) &= Pr\left(\bigcup_{k=1}^{100}A_k\right)=\sum_{k=1}^{100}Pr\left(A_k\right)\\
&=\sum_{k=1}^{100}\binom{100}{k}\cdot 0.5^{100}\cdot 0.6^k=0.5^{100}\cdot\sum_{k=1}^{100}\binom{100}{k}\cdot 0.6^k\cdot 1^{100-k}\\
(*) &= 0.5^{100}\cdot\left(\sum_{k=0}^{100}\binom{100}{k}\cdot 0.6^k\cdot 1^{100-k}-\binom{100}{0}\cdot 0.6^0\cdot 1^{100}\right)\\
&=0.5^{100}\left(1.6^{100}-1\right)=0.8^{100}-0.5^{100}
\end{aligned}$$

כאשר במעבר $*$ הוספנו והחסרנו את הערך המתאים ל-$k=0$ כדי שנוכל להשתמש בבינום ניוטון (**נוסחא 0.10**) לחישוב הסכום.

המאורע $X=3$ מתאים למצב בו לפחות סטודנט אחד נכשל בשני המועדים. נחשב אותו על ידי המשלים – הסיכוי שאף סטודנט לא נכשל בשני המועדים. הסיכוי של סטודנט מסויים להיכשל בשני המועדים הינו $0.5\cdot 0.4=0.2$ ולכן הסיכוי שהסטודנט לא נכשל בשני המועדים הינו 0.8. הסיכוי שכל הסטודנטים לא נכשלו בשני המועדים הינו 0.8^{100} ולכן

$$Pr\left(X=3\right)=1-0.8^{100}$$

מעניין לציין שכצפוי קיבלנו $P_X(3)\simeq 1$ מאחר ויש סיכוי לא זניח להיכשל בשני המועדים (0.2) ולכן מתוך 100 ניסיונות, כמעט בוודאות המאורע הזה יקרה ויגרום לכך שיהיו שלושה מועדים. □

תרגיל 4.27. 1. מתכונת הנרמול נקבל

$$1=\int_{-\infty}^{\infty}f_X(t)\,dt=\int_{-1}^{0}-t\,dt+\int_0^1 ct^2\,dt=\frac{1}{2}+c\left(\frac{1}{3}-0\right)$$

ולכן $c=1.5$.

2. נחשב את האינטגרל של פונקציית הצפיפות ונקבל:

$$F_X(t)=\int_{-\infty}^{t}f_X(x)\,dx=\begin{cases}0 & t<-1\\ \frac{1}{2}-\frac{t^2}{2} & -1\le t<0\\ \frac{1}{2}+\frac{t^3}{2} & 0\le t\le 1\\ 1 & t>1\end{cases}$$

טעות נפוצה היא לחשב אינטגרל לא מסויים בתחום $t\in[0,1]$ ולקבל רק $\frac{t^3}{2}$ ללא החצי. זו טעות מאחר ופונקציית ההתפלגות המצטברת מוגדרת להיות האינטגרל החל מ-$-\infty$ ולכן לא מספיק לחשב אינטגרל בתחום מסויים אלא יש לכלול את תוצאת האינטגרל בתחומים הקודמים. במקרה זה:

$$\int_{-\infty}^{t}f_X(x)\,dx=\int_{-\infty}^{-1}0\,dx+\int_{-1}^{0}-x\,dx+\int_0^t\frac{3}{2}x^2\,dx=0+\frac{1}{2}+\frac{t^3}{2}$$

3. נחשב את ההסתברויות באמצעות הצבה בפונקציית ההתפלגות המצטברת:

א) $P(X \leq -0.5) = F_X(-0.5) = 0.375$

ב) $P(X < -0.5) = P(X \leq -0.5) = 0.375$ שכן כזכור, נקודה אחת אינה משפיעה על תוצאת אינטגרל וההסתברות להיות שווה בדיוק ל-0.5 היא אפס.

ג) $P(X \leq 0.5) = F_X(0.5) = 0.5625$

ד) $P(-0.2 \leq X \leq 0.3) = F_X(0.3) - F_X(-0.2) = 0.0335$

□

תרגיל 4.28. 1. מתכונת הנרמול נקבל

$$1 = \int_{-\infty}^{\infty} f_X(t)\,dt = c\int_0^1 (1-t)^4\,dt = -\frac{c}{5}(1-t)^5|_{t=0}^{t=1} = \frac{c}{5}$$

כלומר $c = 5$.

2. נסמן את קיבולת המאגר ב-M. אנו דורשים שההסתברות שצריכת הדלק תהיה גדולה מהקיבולת תהיה קטנה מ-0.05, כלומר $Pr(X > M) \leq 0.05$. נחשב את ההסתברות המבוקשת:

$$P(X > M) = \int_M^{\infty} f_X(t)\,dt = \int_M^1 f_X(t)\,dt = (1-M)^5 \leq 0.05$$

מפתרון המשוואה נקבל $M \geq 0.4507$, כלומר כאשר קיבולת המאגר גדולה מ-450 ליטרים, הביקוש יעלה על ההיצע בהסתברות הקטנה מ-0.05 והמאגר יתרוקן לחלוטין רק ב-5% מהשבועות.

□

תרגיל 4.29. 1. נחשב את הקבוע A לפי תכונת הנרמול:

$$1 = \int_{-\infty}^{\infty} f_X(t)\,dt = \int_{-\frac{L}{2}}^{\frac{L}{2}} A\cdot\cos^2\left(2\pi n\frac{t}{L}\right)\,dt = A\left(\frac{t}{2} + \frac{L\sin(4\pi n\frac{x}{L})}{8\pi n}\right)\Bigg|_{-\frac{L}{2}}^{\frac{L}{2}} = A\frac{L}{2}$$

ולכן $A = \frac{2}{L}$.

2. ההסתברות למצוא את האלקטרון בדיוק במרכז הבור הוא אפס, שכן זו התפלגות רציפה.

3. האלקטרון נמצא במרחק של פחות מ-$0.1L$ ממרכז הבור כאשר $|X| \leq 0.1L$ וההסתברות לכך הינה

$$\begin{aligned} Pr\left(-0.1L \leq X \leq 0.1L\right) &= \int_{-0.1L}^{0.1L} f_X(t)\,dt = \int_{-0.1L}^{0.1L} \frac{2}{L}\cdot\cos^2\left(2\pi n\frac{t}{L}\right)\,dt \\ &= \frac{2}{L}\left(\frac{t}{2} + \frac{L\sin(4\pi n\frac{t}{L})}{8\pi n}\right)\Bigg|_{-0.1L}^{0.1L} = 0.2 + \frac{\sin(0.4\pi n)}{2\pi n} \end{aligned}$$

□

תרגיל 4.30. **1.** בין 3ל-4 הפונקציה לינארית עולה, בין 5 ל-6 לינארית יורדת ובשאר הישר הממשי היא קבועה. לכן פונקציית צפיפות ההסתברות הינה:

$$f_X(t)=\begin{cases} c\,(t-3) & t\in[3,4] \\ c & t\in[4,5] \\ c\,(6-t) & t\in[5,6] \\ 0 & t\notin[3,6] \end{cases}$$

2. נדרוש נרמול:

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f_X(t)\,dt = \int_3^4 c(t-3)\,dt + \int_4^5 c\,dt + \int_5^6 c(6-t)\,dt \\ &= \frac{c}{2}+c+\frac{c}{2}=2c \end{aligned}$$

לכן $c=\frac{1}{2}$.

3. בהמשך לסעיף הקודם, פונקציית ההתפלגות המצטברת תתקבל על ידי אינטגרציה:

$$F_X(t)=\begin{cases} 0 & t<3 \\ \frac{1}{4}\,(t-3)^2 & t\in[3,4] \\ \frac{1}{2}t-\frac{7}{4} & t\in[4,5] \\ 1-\frac{1}{4}(t-6)^2 & t\in[5,6] \\ 1 & t>6 \end{cases}$$

4. על מנת למצוא את ההתפלגות של Y, ראשית יש למצוא את הקשר בינו לבין X. מאחר ו-X הוא הקוטר ו-Y הוא השטח, מתקיים $Y=\pi\left(\frac{X}{2}\right)^2$. לכן:

$$F_Y(t)=Pr\left(Y\le t\right)=Pr\left(\frac{\pi X^2}{4}\le t\right)=Pr\left(X^2\le\frac{4t}{\pi}\right)$$

מאחר ו-Y הינו שטח, אזי עבור $t<0$ הפה"מ יהיה 0. אחרת, נוציא שורש ונקבל:

$$F_Y(t)=Pr\left(X\le\sqrt{\frac{4t}{\pi}}\right)=F_X\left(\sqrt{\frac{4t}{\pi}}\right)$$

נותר להציב את הביטוי האחרון בפונקציית ההתפלגות המצטברת של X ולסדר את הביטויים:

$$F_Y(t) = F_X\left(\sqrt{\frac{4t}{\pi}}\right) == \begin{cases} 0 & \sqrt{\frac{4t}{\pi}} < 3 \\ \frac{1}{4}\left(\sqrt{\frac{4t}{\pi}} - 3\right)^2 & \sqrt{\frac{4t}{\pi}} \in [3,4] \\ \frac{1}{2}\sqrt{\frac{4t}{\pi}} - \frac{7}{4} & \sqrt{\frac{4t}{\pi}} \in [4,5] \\ 1 - \frac{1}{4}\left(\sqrt{\frac{4t}{\pi}} - 6\right)^2 & \sqrt{\frac{4t}{\pi}} \in [5,6] \\ 1 & \sqrt{\frac{4t}{\pi}} > 6 \end{cases}$$

$$== \begin{cases} 0 & t < \frac{9\pi}{4} \\ \frac{1}{4}\left(\sqrt{\frac{4t}{\pi}} - 3\right)^2 & t \in \left[\frac{9\pi}{4}, 4\pi\right] \\ \sqrt{\frac{t}{\pi}} - \frac{7}{4} & t \in \left[4\pi, \frac{25\pi}{4}\right] \\ 1 - \left(\sqrt{\frac{t}{\pi}} - 3\right)^2 & t \in \left[\frac{25\pi}{4}, 9\pi\right] \\ 1 & t > 9\pi \end{cases}$$

□

*תרגיל **4.31**.* נסמן את פונקציית הצפיפות של X ב-f_X ונחשב את פונקציית הצפיפות של Y. לשם כך, כרגיל, נתחיל בחישוב פונקציית ההתפלגות המצטברת:

$$F_Y(t) = Pr(Y \le t) = Pr(aX + b \le t) = Pr\left(X \le \frac{t-b}{a}\right) = F_X\left(\frac{t-b}{a}\right)$$

על מנת לחשב את פונקציית הצפיפות נגזור ונקבל:

$$f_Y(t) = \frac{dF_Y(t)}{dt} = \frac{dF_X\left(\frac{t-b}{a}\right)}{dt} = f_X\left(\frac{t-b}{a}\right) \cdot \frac{1}{a}$$

המשימה טרם הושלמה, שכן בחישוב דלעיל ביצענו הנחה סמויה: $0 < a$. דבר זה איפשר לחלק ב-a מבלי שסימן האי-שוויון יתהפך. נחשב את פונקציית הצפיפות עבור $a < 0$:

$$F_Y(t) = Pr(Y \le t) = Pr(aX + b \le t) = Pr\left(X \ge \frac{t-b}{a}\right) = 1 - F_X\left(\frac{t-b}{a}\right)$$

ולאחר גזירה נקבל:

$$f_Y(t) = \frac{dF_Y(t)}{dt} = \frac{d\left(1 - F_X\left(\frac{t-b}{a}\right)\right)}{dt} = -f_X\left(\frac{t-b}{a}\right) \cdot \frac{1}{a} = \frac{1}{|a|} f_X\left(\frac{t-b}{a}\right)$$

ולמעשה נוסחא זו מתאימה גם למקרה ש-$a > 0$ שכן במקרה כזה $a = |a|$.

המקרה האחרון אותו יש לבדוק הוא המקרה בו $a = 0$. במקרה כזה ללא תלות בערך של X מתקיים $Y = b$ כלומר Y משתנה מקרי בדיד המקבל את הערך b בהסתברות 1. □

5 תוחלת, שונות ומדדים מרכזיים

"מי שיגיע עד הסוף
ירצה לבכות מרוב תוחלת"
ערב עירוני/יוסי בנאי

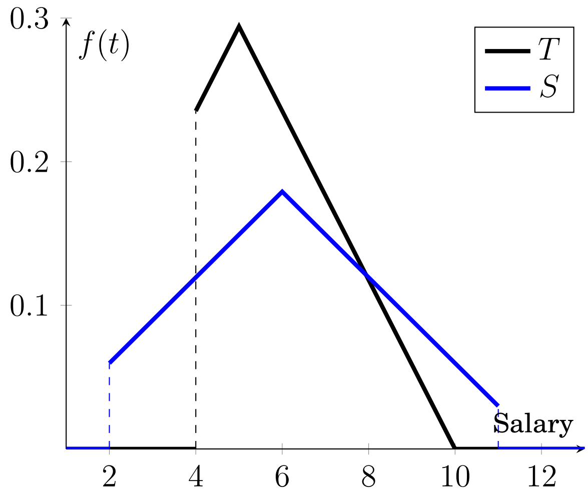
ההתפלגות מספקת את כל המידע הקיים עבור משתנה מקרי מסויים ומתוך ידיעת ההתפלגות ניתן לענות על כל שאלה הסתברותית הנוגעת לאותו המשתנה המקרי. במובן מסויים, ההתפלגות מכילה יותר מדי מידע אשר קשה להתמודד עימו בו זמנית ובחלקו לא תמיד נחוץ או לא תמיד רלוונטי. למשל, נניח שמהמר צריך להחליט האם לשחק בלוטו או ב"צ'אנס". הסתברויות הזכייה של משחקים אלו מחושבות בתרגילים 3.5 ו-3.6. במשחק הלוטו הפרס הראשון הינו 5 מיליון ש"ח ומחיר שליחת הניחוש הינו 3 ש"ח. לעומת זאת, במשחק ה"צ'אנס" הפרס הראשון הינו 20,000 ש"ח ועלות ההשתתפות הינה 10 ש"ח. איזה מהמשחקים עדיף? כיצד נשקלל את הרווח היותר גדול במשחק הלוטו ומחיר ההשתתפות הנמוך יותר עם כך שהסתברות הזכייה נמוכה יותר?

בפרק הנוכחי נפתח מדדים שונים המאפשרים להשוות בין התפלגויות שונות. כל מדד כזה הינו מספר בודד המחושב עבור ההתפלגות ותופס פן מסויים של ההתפלגות. ערכו של המדד מאפשר השוואה בין ההתפלגויות השונות ביחס למידע הטמון במדד זה, תוך התעלמות מאספקטים אחרים של ההתפלגות שפחות רלוונטיים לשאלה בה עוסקים. שני המדדים העיקריים בהם עוסקים בתורת ההסתברות הם התוחלת והשונות, אך בטרם נגדירם נציג מספר מדדים נוספים ומעניינים, חלקם מופיעים תדיר בתחומי דעת נוספים וכן בעיתונות. לאורך כל הדיון תלווה אותנו השאלה הבאה מתחום הכלכלה.

בעיה 5.1. רעות מתלבטת האם ללמוד אילוף אריות (T) או אשורולוגיה (S). מאחר ושני מקצועות אלו מעניינים אותה במידה שווה, הבחירה ביניהם תיעשה לפי השכר המוצע בכל מקצוע. מעלעול בנתוני לשכת הסטטיסטיקה הממשלתית היא מגלה שהתפלגות השכר בשני מקצועות אלו נתונה לפי פונקציות הצפיפות הבאות (באלפי ש"ח):

$$f_T(t)=\begin{cases}\frac{1}{17}t & t\in[4,5]\\ \frac{1}{17}(10-t) & t\in[5,10]\\ 0 & t\notin[4,10]\end{cases}\qquad f_S(t)=\begin{cases}\frac{2}{67}t & t\in[2,6]\\ \frac{2}{67}(12-t) & t\in[6,11]\\ 0 & t\notin[2,11]\end{cases}$$

שתי פונקציות הצפיפות משרוטטות באיור הבא:



כמו בדוגמת הלוטו וה"צ'אנס", קשה להחליט איזו מההתפלגויות עדיפה. באשורולוגיה, למשל, השכר המרבי גדול יותר מאשר באילוף אריות, אך מצד שני גם השכר המינימלי קטן יותר. אם כך, איזה מהמקצועות כדאי לבחור ואיך ניתן לבצע את הבחירה?

5.0.1 שכיח

דרך אחת להכריע בין שני המקצועות היא לבדוק מהו השכר שרוב המועסקים במקצוע מרוויחים. דרך פשוטה לעשות זאת היא לעבור בין כל העובדים, לשאול אותם לשכרם ולבדוק מהי התשובה שתתקבל הכי הרבה פעמים. ערך זה, המכונה "השכיח", הוא גם ניבוי טוב לשכר של רעות בכל אחד מהמקצועות. אם רוב המועסקים באילוף אריות מרוויחים כ-5,000 ש"ח, הגיוני שגם רעות תרוויח סכום דומה כי זו התוצאה הכי נפוצה. באופן דומה, הסיכוי להרוויח בהגרלת לוטו הוא מאוד קטן והתוצאה השכיחה היא להפסיד. זו גם התוצאה שלה נצפה (אך לא נייחל לה) כשנשתתף בלוטו.

הגדרה 5.1 (שכיח)**.** השכיח הוא הערך בעל ההסתברות הגבוהה ביותר (עבור משתנים מקריים בדידים) או בעל הצפיפות הגבוהה ביותר (עבור משתנים מקריים רציפים). כאשר ישנם מספר מקסימה מקומיים, מקובל לקחת את כולם בתור שכיחים, גם אם המקסימום המקומי קטן מהמקסימום המוחלט של פונקציית הצפיפות.

אילוף אריות או אשורולוגיה - שכיח. שתי ההתפלגויות לינאריות מקוטעין ולכן המקסימום מתקבל באחד מקצוות תחומי ההגדרה. קל לראות (דרך האיור או על ידי הצבת הערכים) שהערך השכיח עבור אילוף אריות הינו 5 ואילו הערך השכיח עבור אשורולוגיה הינו 6, כלומר לפי הערך השכיח, אשורולוגיה עדיפה על פני אילוף אריות. □

השכיח, בהיותו מדד פשוט יחסית סובל מבעיות רבות. ראשית, כלל לא בטוח שיש שכיח יחיד. למשל, כאשר מטילים קוביה, לכל התוצאות יש את אותה ההסתברות להתקבל ולכן כולן יהיו השכיחים של ההתפלגות. אמנם רוב ההתפלגויות שנתקלים בהן בחיים כוללות שכיח בודד אך במקרה ואין זה כך שוב לא ניתן להשתמש בשכיח כדי להשוות בין התפלגויות שכן לא מדובר על מספר בודד כפי שרצינו.

שנית, השכיח לא מתחשב כלל בערכים אחרים של ההתפלגות, במקסימום, במינימום ובפיזור של הערכים סביבו. למשל, עבור שני המשתנים המקריים הבאים יש את אותו השכיח, למרות שבבירור היינו מעדיפים להשתתף בהגרלה שזכייתה נקבעת לפי ההתפלגות של X מאשר לפי ההתפלגות של Y:

$$X = \begin{cases} 5 & 0.51 \\ 100 & 0.49 \end{cases} \quad Y = \begin{cases} 5 & 0.51 \\ 4 & 0.49 \end{cases}$$

שלישית, השכיח לא לוקח בחשבון מידע נוסף שיכול להיות רלוונטי ולשנות את ההתפלגות הרלוונטית. למשל, אם רעות יודעת שהיא מוכשרת בכל מעשי ידיה, היא תצפה למשכורת שמקבלים אנשים מוכשרים ולאו דווקא את המשכורת השכיחה. בדומה, השם השכיח בעולם הוא "מוחמד" בהטיותיו השונות, אך אם תסתובבו בבני ברק תתקשו למצוא אנשים העונים לשם זה.

תרגיל 5.1. יהי X משתנה מקרי רציף עם פונקציית צפיפות גזירה $f_X(t)$. נגדיר משתנה מקרי חדש $Y = aX+b$ כאשר $a,b \in \mathbb{R}$ קבועים. הוכיחו שאם x_m הוא השכיח של ההתפלגות X אזי $y_m = ax_m + b$ הוא השכיח של התפלגות Y.
הערה: הטענה נכונה גם למשתנים מקריים בדידים.

תרגיל 5.2. מצאו דוגמא למשתנה מקרי X ולפונקציה g כך שאם נגדיר משתנה מקרי $Y = g(X)$ לא יתקיים $y_m = g\left(x_m\right)$ כאשר x_m ו-y_m הם השכיחים של X ו-Y, בהתאמה. לשם הפשטות, כדאי לבחור משתנה מקרי בדיד.

5.0.2 אחוזונים

לאורך חייה עד כה, רעות הצטיינה בכל מעשי ידיה. היא סיימה את התיכון בהצטיינות וקיבלה ציון גבוה בפסיכומטרי. היא מאמינה שהיא תמשיך להצטיין גם בעבודתה החדשה ולכן שכר המינימום או השכר שרוב מאלפי האריות או האשורולוגים מרוויחים לא מעניין אותה. היא חושבת שהיא תהיה בין הטובים בתחום ולכן רוצה לדעת מהו השכר שהטובים מקבלים. נניח שבהיותה מצטיינת היא תמיד הייתה יותר טובה מאשר 90% מהאחרים. מעניין אותה אם כך השכר שמקבלים בשני המקצועות אלו הטובים מ-90% העובדים האחרים, כלומר השכר שהסיכוי לקבל פחות ממנו שווה ל-0.9. מספר זה נקרא האחוזון ה-0.9 או האחוזון ה-90 [1] של ההתפלגות.

הגדרה 5.2. האחוזון ה-$\alpha \in (0,1)$ של המשתנה המקרי X, המסומן ב-x_α, הוא המספר הממשי המקיים

$$x_\alpha = \mathbf{sup}\left\{y | Pr\left(X \le y\right) \le \alpha\right\} = \mathbf{sup}\left\{y | F_X(y) \le \alpha\right\}$$

כלומר זהו המספר המקסימלי אשר ההסתברות לקבל ערכים קטנים ממנו היא α לכל היותר.

אילוף אריות או אשורולוגיה - האחוזון ה-90. נבדוק מהו האחוזון ה-90 של כל התפלגות. מאחר ולא חישבנו את פונקציית ההתפלגות המצטברת של כל משתנה מקרי, נחשב את ההסתברויות המבוקשות באמצעות חישוב האינטגרלים המתאימים.
עבור אילוף אריות אנחנו מחפשים את המספר $t_{0.9}$ המקיים $Pr\left(T \le t_{0.9}\right) = 0.9$. נשים לב כי

$$Pr\left(T \le 5\right) = \int_4^5 \frac{1}{17} t\, dt = \frac{9}{34} < 0.9$$

כלומר האחוזון ה-90 יתקבל בהכרח בתחום $[5,10]$ ולכן המשוואה שיש לפתור הינה

$$\begin{aligned} 0.9 = Pr\left(T \le t_{0.9}\right) &= \int_4^5 \frac{1}{17} t\, dt + \int_5^{t_{0.9}} \frac{1}{17}\left(10-t\right)\, dt \\ &= \frac{9}{34} + \frac{1}{34}\left(-t_{0.9}^2 + 20t_{0.9} - 75\right) \end{aligned}$$

[1] כשהכוונה, כמובן, היא ל-90%.

לאחר סידור המשוואה נקבל את המשוואה הריבועית

$$t_{0.9}^2 - 20t_{0.9} + 96.6 = 0$$

שפתרונה בתחום הנכון הינו בקירוב $t_{0.9} = 8.156$.

באותה צורה עבור אשורולוגיה נקבל את החישוב

$$\begin{aligned} 0.9 = Pr\left(S \le s_{0.9}\right) &= \int_2^6 \frac{2}{67} t\, dt + \int_6^{s_{0.9}} \frac{2}{67}\left(12 - t\right)\, dt \\ &= \frac{32}{67} + \frac{1}{67}\left(-s_{0.9}^2 + 24 s_{0.9} - 108\right) \end{aligned}$$

המוביל למשוואה

$$s_{0.9}^2 - 24s_{0.9} + 136.3 = 0$$

שפתרונה $s_{0.9} = 9.225$.

לאור חישובים אלו, האחוזון ה-90 באשורולוגיה גבוה יותר מאשר באילוף אריות, לכן אם אכן רעות תהיה בין 10% המצטיינים במקצועה, כדאי לה לבחור באשורולוגיה. □

הרעיון העומד מאחורי האחוזונים הוא חיפוש ספים בהתפלגות אשר עד אליהם יש אחוז מסויים של האוכלוסיה המעניין אותנו. בתחומי מחקר שונים מעניינים אחוזונים שונים. למשל, בסטטיסטיקה ובמחקר ניסיוני מעניינות תוצאות יוצאות דופן המוגדרות בדרך כלל ככאלה הגדולות מהאחוזון ה-95 (כלומר, הסיכוי לקבל אותן הוא רק 0.05). בכלכלה, מקובל לדבר על עשירונים. למשל, העשירון העליון הוא ה-10% באוכלוסיה המרוויחה את המשכורת הגבוהה ביותר מבין כל השכירים במשק. כדי להיכלל בעשירון העליון צריך להרוויח שכר הגבוה משכרם של 90% מהאוכלוסיה, כלומר להרוויח משכורת הגבוהה מהאחוזון ה-90 (זהו החישוב שביצענו עבור רעות). באופן דומה, העשירון התחתון הוא ה-10% מהאוכלוסיה המרוויחה את המשכורות הנמוכות ביותר, ולכן כדי להיכלל בעשירון זה צריך להרוויח שכר הקטן מהאחוזון ה-10 של התפלגות השכר. באופן דומה מוגדר המאיון העליון והאלפיון העליון.

לפי נתוני הלשכה המרכזית לסטטיסטיקה, בשנת **2017** עשירוני ההכנסה (ברוטו ליחיד) היו:

עשירון 9	עשירון 8	עשירון 7	עשירון 6	עשירון 5	עשירון 4	עשירון 3	עשירון 2	עשירון 1
12,926	10,107	8,451	7,209	6,208	5,314	4,045	3,435	2,396

על סמך טבלה זו ניתן לקבל תמונת מצב של רמות ההכנסה במשק הישראלי, מבלי להיכנס לפרטים או להיזדקק לכל המידע הקיים, ולהסיק את המסקנות הרצויות לו. למשל, על סמך הטבלה, שכר המינימום (**4,300** ש"ח) נמצא בעשירון הרביעי כלומר בין 30% ל-40% מהשכירים במשק משתכרים פחות משכר המינימום.

אחוזון נוסף הנמצא בשימוש היומיומי הוא האחוזון ה-50, המכונה החציון. החציון הוא הערך שהסיכוי להיות מתחתיו הוא 0.5, כלומר הוא חוצה את ציר המספרים לשני חלקים והסיכוי להיות בכל אחד מהחלקים שווה. בכלכלה, למשל, מסתכלים על השינוי במשכורת החציונית כמדד לאופן בו השכר משתנה במשק לאורך השנים. בנוסף, מעמד הביניים מוגדר באמצעות המשכורת החציונית בתור חלק האוכלוסיה המרוויח בין 125% − 75% מהמשכורת החציונית.

תרגיל 5.3. מהו החציון ומהו תחום השכר אותו מרוויחים מאלפי האריות במעמד הביניים ומהו החציון ותחום השכר אותו מרוויחים האשורולוגים במעמד הביניים? האם גודלו של מעמד הביניים שווה בשני מקצועות אלו?

תרגיל 5.4. מהו החציון של התופעות הבאות:

1. התפלגות המקסימום בהטלת 3 קוביות (בעיה 4.3)?
2. זמן חימום פיצה בתנור ביתי (תרגיל 4.13)?
3. ציון בבחינה המתפלג לפי תרגיל 4.14? כיצד תשתנה תשובתכם אם ינתנו הפקטורים המוצעים בתרגילים 4.16 ו-4.20?

תרגיל 5.5. יהי X משתנה מקרי רציף ויהיו $a, b \in \mathbb{R}$ קבועים כאשר $a > 0$. הוכיחו כי האחוזון ה-α של $Y = aX + b$ הינו $y_\alpha = ax_\alpha + b$. מה יהיה האחוזון ה-α של Y כאשר $a < 0$?

5.1 תוחלת

השכיח וחישוב אחוזונים מסויימים איפשר לנו לקבל מסקנות לגבי ההבדל בין לימודי אשורולוגיה לבין לימודי אילוף אריות, אך מדדים אלו לא פתרו עדיין את כל הבעיות. למשל, הסוגיה האם להמר בלוטו או בצ'אנס טרם יושבה, שכן בשני המקרים השכיח הוא "לא לזכות"" ובגלל ההסתברות הנמוכה מאוד לזכות, הרוב המוחלט של האחוזונים שווים לעלות ההשתתפות. נפתח מדד המבדיל בין הימורים אלו על ידי חזרה לגישה השכיחותית ונבדוק כמה נרוויח בטווח הארוך מכל אחד מההימורים.

נניח שמהמר משחק בלוטו n פעמים. נסמן ב-A את המאורע שזוכים בלוטו. מתוך n הניסויים ב- $S_n(A)$ המהמר יזכה בניסוי ויקבל $5,000,000$ ש"ח וב-$S_n(\overline{A})$ האחרים יפסיד וישלם 3 ש"ח. בסך הכל, אחרי n ההימורים, הכסף שיהיה למהמר הינו $5,000,000 \cdot S_n(A) - 3 \cdot S_n(\overline{A})$. סך הכסף שהמהמר מרוויח או מפסיד בכל הימור, בממוצע, יהיה

$$\begin{aligned}
\lim_{n\to\infty} \frac{5,000,000 \cdot S_n(A) - 3 \cdot S_n(\overline{A})}{n} &= 5,000,000 \cdot \lim_{n\to\infty} \frac{S_n(A)}{n} - 3 \cdot \lim_{n\to\infty} \frac{S_n(\overline{A})}{n} \\
&= 5,000,000 \cdot Pr(A) - 3 \cdot Pr(\overline{A}) \\
(exer.\, 3.5) &= 5,000,000 \cdot 6.14 \cdot 10^{-8} - 3 \cdot \left(1 - 6.14 \cdot 10^{-8}\right) \\
&= -2.693
\end{aligned}$$

כלומר, המהמר בממוצע יפסיד כסף בהימורים אלו. אמנם אם הוא ישחק הרבה מאוד פעמים מדי פעם הוא צפוי לזכות, אך הכסף שהוא ירוויח בכל זכייה כזו לא יכסה את סך עלויות ההימורים הכושלים שלו. תוצאה זו אינה מפתיעה שכן הלוטו לא פושט רגל...

החישוב שהתקבל הוא למעשה הממוצע המשוקלל של הערכים האפשריים בלוטו – רווח של 5 מיליון ש"ח או הפסד של 3 ש"ח, כאשר המשקולת הן ההסתברויות של כל תוצאה. ממוצע משוקלל זה של כל התוצאות האפשריות בתומך של המשתנה המקרי מכונה התוחלת של המשתנה המקרי ומהווה מדד לערך הממוצע שיתקבל בכל חזרה של הניסוי.

הגדרה 5.3 (תוחלת של משתנה מקרי בדיד)**.** יהי X משתנה מקרי בדיד. התוחלת שלו מסומנת על ידי $\mathbb{E}(X)$ ומוגדרת על ידי

$$\mathbb{E}(X) = \sum_{k\in supp(X)} k \cdot P_X(k) = \sum_{k\in supp(X)} k \cdot Pr(X = k) \qquad (5.1)$$

נחשב בהתאם להגדרה את התוחלת של משחק הצ'אנס על מנת שנוכל להשוות ביניהם. נסמן ב-X את הרווח ממשחק הצ'אנס. כאמור לעיל, הפרס במשחק הצ'אנס הינו $20,000$ ש"ח וסיכוי הזכיה לפי תרגיל **3.6** הינו $\frac{1}{4096}$, כלומר $Pr\left(X=20,000\right)=\frac{1}{4096}$. האפשרות השניה היא הפסד ובמקרה זה משלמים רק את עלות ההימור, **10** ש"ח. לכן $Pr\left(X=-10\right)=1-\frac{1}{4096}$. תוחלת ההימור במשחק הצ'אנס היא אם כך

$$\mathbb{E}\left(X\right)=20,000\cdot\frac{1}{4096}-10\cdot\left(1-\frac{1}{4096}\right)=-5.11$$

כצפוי, גם במקרה זה קיבלנו תוחלת שלילית שכן מפעל הפיס לא מתרושש. בהשוואה בין הלוטו לצ'אנס, התוחלת (בממוצע) בהימור של במשחק הצ'אנס נמוכה יותר מהתוחלת בהימור בלוטו, לכן בכל הימור במשחק הצ'אנס המהמר מפסיד יותר כסף ממה שהוא מפסיד במשחק הלוטו. התוחלת, אם כך, היא מדד לתוצאה הממוצעת שצפויה להתקבל כאשר משחקים במשחק מספר פעמים רב מאוד. בחלק מהחזרות על ההימור המהמר יקבל תוצאה גבוהה יותר, בחלק יקבל תוצאה נמוכה יותר, והתוחלת היא הממוצע ארוך הטווח של סך כל התוצאות. למשל, אם המהמר ישחק מיליון משחקי צ'אנס, הוא יגלה שמדי פעם הוא הרוויח, לרוב הוא הפסיד, ובממוצע הוא הפסיד יותר מ-**5** ש"ח בכל משחק.

היות והתוחלת היא רק ממוצע, היא לא חייבת להיות תוצאה אפשרית בניסוי. באף אחת מהתוצאות האפשריות במשחק הצ'אנס המהמר לא יפסיד בדיוק 5.11 ש"ח (שזו התוחלת). יחד עם זאת, ממוצע זה לוקח בחשבון את כל התוצאות האפשריות ואת ההסתברויות לקבלן ולכן במקרים רבים הוא המדד המעניין ואחד המדדים המרכזיים בתורת ההסתברות להערכת התוצאה של ניסוי מקרי.

בעיה 5.2. מטילים קוביה הוגנת. נגדיר X – תוצאת ההטלה. מהי התוחלת של X?

הוכחה. התומך של X הינו $\{1,\ldots,6\}$ וכל התוצאות מתקבלות בהסתברות שווה, כלומר בהסתברות $\frac{1}{6}$. לכן התוחלת של הקוביה הינה

$$\mathbb{E}\left(X\right)=\sum_{n=1}^{6}n\cdot Pr\left(X=n\right)=\sum_{n=1}^{6}\frac{n}{6}=3.5$$

כמקודם, קיבלנו תוצאה בלתי אפשרית. באף הטלה לא נקבל אף פעם 3.5, אבל אם נטיל קוביה מיליון פעמים ונסכום את כל התוצאות, נקבל בערך $3,500,000$ והממוצע בכל הטלה יהיה קרוב מאוד לתוחלת – 3.5. □

תרגיל 5.6. מטילים קוביה **3** פעמים ונגדיר X – המקסימום שהתקבל בשלושת ההטלות (ראו בעיה **4.3**). מהי התוחלת של X?

תרגיל 5.7. יורים באקראי על מטרה עגולה שרדיוסה **10** ס"מ, כאשר נקודת הפגיעה מתפלגת אחיד על פני המטרה. פגיעה במרחק של עד סנטימטר ממרכז המטרה מזכה ב-**10** נקודות, פגיעה במרחק של בין **1** ס"מ ל-**4** ס"מ ממרכז המטרה מזכה ב-**7** נקודות, פגיעה במרחק של בין **4** ס"מ ל-**8** ס"מ מזכה ב-**3** נקודות ופגיעה בשאר המטרה מזכה ב-**0** נקודות. מהי תוחלת הניקוד ביריה בודדת?
הערה: כדאי להיעזר בבעיה **4.6**.

התוחלות שחישבנו עד כה היו תוחלות של משתנים מקריים עם תומך סופי. במקרה כזה נוסחא **5.1** היא פשוט סכימה על כל האפשרויות. ראינו כבר שהתומך של משתנה מקרי בדיד יכול להיות קבוצה בת מנייה, ובמקרה כזה הנוסחא זהה והשוני הוא שהסכימה הופכת מסכום סופי לחישוב סכום של טור, כפי שממחישה הבעיה הבאה.

בעיה 5.3. בכד שני כדורים – כדור לבן וכדור שחור. מוציאים מהכד כדור, בודקים את צבעו ומחזירים אותו פנימה בצירוף כדור נוסף מאותו הצבע. התהליך מסתיים כאשר נבחר לראשונה כדור שחור. מהי תוחלת מספר הפעמים שהוצא כדור מהכד?

הוכחה. נסמן את מספר הפעמים שהוצא כדור מהכד ב-X. על מנת לחשב את התוחלת של X ראשית עלינו לחשב את ההתפלגות שלו. הכדור השחור יכול לצאת אחרי כל מספר של הוצאות, לכן התומך הוא אינסופי ושווה לכל הטבעיים $supp\left(X\right)=\mathbb{N}$. המאורע $X=k$ מתקיים כאשר הכדור הראשון היה לבן, השני לבן, השלישי לבן וכן הלאה, ורק הכדור האחרון שהוצא בהוצאה ה-k היה שחור. בכל הוצאה מספר הכדורים הלבנים גדל ב-1 מההוצאה הקודמת ומספר הכדורים הכולל גדל ב-1, לכן ההסתברות תהיה:

$$Pr\left(X=k\right)=\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{k-1}{k}\cdot\frac{1}{k+1}=\frac{1}{k\left(k+1\right)}$$

נחשב את התוחלת:

$$\mathbb{E}\left(X\right)=\sum_{k=1}^{\infty}k\cdot Pr\left(X=k\right)=\sum_{k=1}^{\infty}k\cdot\frac{1}{k\left(k+1\right)}=\sum_{k=1}^{\infty}\frac{1}{k+1}$$

הטור הזה הוא הטור ההרמוני (נוסחא **0.24**) ולכן מתבדר. במקרה כזה נאמר שהתוחלת היא אינסופית. $\square$

נעמוד על המשמעות של תוחלת אינסופית. כאשר נחזור על הניסוי שוב ושוב, לרוב נקבל תוצאות עם הסתברות גבוהה כמו $X=1$ ו-$X=2$. מדי פעם נקבל תוצאה מאוד גבוהה אשר תגדיל את הממוצע. מאחר וההסתברויות לקבל תוצאות מאוד גבוהות לא קטנות מספיק, מספרים גדולים יופיעו בקצב גבוה וככל שנחזור על הניסוי עוד ועוד פעמים, כך הממוצע ילך ויגדל. יחד עם זאת, התכנסות של הטור עדיין אינה ערובה לכך שהתוחלת תהיה קיימת. כדי שהתוחלת תהיה קיימת הטור חייב לא רק להתכנס אלא להתכנס בהחלט, כלומר שהערך המוחלט של ערכי הטור יתכנס.

הגדרה 5.4 (תוחלת אינסופית)**.** יהי X משתנה מקרי בדיד עם תומך אינסופי. נאמר שהתוחלת של X קיימת רק כאשר הטור המתאר אותה מנוסחא **5.1** מתכנס בהחלט, כלומר כאשר הטור

$$\sum_{n\in supp(X)}\left|n\right|\cdot Pr\left(X=n\right)$$

מתכנס. כאשר הטור מתבדר או מתכנס בתנאי (דהיינו, מתכנס ללא הערך המוחלט אבל הטור עם הערכים המוחלטים מתבדר), נאמר שלמשתנה המקרי "אין תוחלת" או "תוחלת אינסופית".

הסיבה שהתכנסות הטור אינה מספיקה אלא שהדרישה היא להתכנסות בהחלט היא שבמקרה שהטור מתכנס בתנאי הערך אליו הטור מתכנס תלוי בסדר האיברים בטור. לכן, אם נבצע את הניסוי אינסוף פעמים פעמיים, נקבל בשני הניסויים את התוצאות בסדר שונה והממוצע של התוצאות יתכנס לערך שונה למרות שהשכיחות של כל תוצאה זהה. למעשה, כאשר הטור מתכנס בתנאי, ניתן להתכנס לכל מספר ממשי על ידי החלפת סדר המחוברים בטור. תופעה זו מאוד לא רצויה לתוחלת, שכן משמעותה היא שאם נבצע את הניסוי אינסוף פעמים פעמיים, ונחשב את הממוצע בכל אחד משני הניסויים, נקבל ממוצעים שונים, מה שמעקר מתוכן את הרעיון שמאחורי התוחלת.

נעבור לדון במשתנים מקריים רציפים. גם כאן נרצה לבצע ממוצע משוקלל על הערכים האפשריים. בהיעדר פונקציית הסתברות, נשתמש בפונקציית הצפיפות כדי לשקלל את הערכים האפשריים של המשתנה המקרי.

הגדרה 5.5 (תוחלת של משתנה מקרי רציף)**.** יהי X משתנה מקרי רציף. התוחלת שלו מוגדרת על ידי

$$\mathbb{E}\left(X\right)=\int_{-\infty}^{\infty}x\cdot f_{X}\left(x\right)\,dx \tag{5.2}$$

הערה: כמו בהגדרה 5.4 ומאותה סיבה, נדרוש שהאינטגרל יתכנס בהחלט. אם הוא מתבדר או מתכנס בתנאי, נאמר שלמשתנה המקרי הרציף X "אין תוחלת" או "תוחלת אינסופית".

אילוף אריות או אשורולוגיה – תוחלת. נחשב את התוחלות של שני משתנים מקריים אלו, דהיינו את השכר הממוצע, על מנת להשוות בין שני המשתנים המקריים. עבור אילוף אריות נקבל

$$\mathbb{E}\left(T\right)=\int_{-\infty}^{\infty}t\cdot f_T(t)\,dt=\int_4^5\frac{t^2}{17}\,dt+\int_5^{10}\frac{t}{17}(10-t)\,dt=6.098$$

ואילו עבור אשורולוגיה נקבל

$$\mathbb{E}\left(S\right)=\int_{-\infty}^{\infty}t\cdot f_S(t)\,dt=\int_2^6\frac{2t^2}{67}\,dt+\int_6^{11}\frac{2t}{67}(12-t)\,dt=6.199$$

כלומר גם מבחינת השכר הממוצע, אשורולוגיה עדיפה על אילוף אריות. □

אמנם קיבלנו שבכל המדדים שנבדקו אשורולוגיה עדיפה על אילוף אריות, אך אין הדבר תמיד נכון. זהו מקרה שנובע מהצורה הספציפית של פונקציות הצפיפות של המשתנים המקריים והתפלגויות אחרות יכולות להוביל למצבים אחרים בהם משתנה מקרי אחד עדיף על פני האחר בחציון אך נופל ממנו בתוחלת ולהפך. בנוסף, גם כאשר בוחנים את המדדים הללו עבור משתנה מקרי בודד, ניתן לקבל כל סדר ביניהם. כך למשל, יהיו מקרים שבהם התוחלת, החציון והשכיח הם אותו הערך, אך גם מקרים בהם התוחלת גדולה מהחציון ומקרים בהם התוחלת קטנה ממנו.

ההבדל בין התוחלות של השכר באשורולוגיה ואילוף אריות הוא מאוד קטן, בטח ביחס להבדל בין החציונים של שני משתנים מקריים אלו. הסיבה לכך היא שהתוחלת לוקחת בחשבון את הערכים הקיצוניים לטוב ולרע. למשל באשורולוגיה האחוזונים הגבוהים מרוויחים משכורות גבוהות יותר מאשר מקביליהם באילוף אריות אך גם ההפך נכון – האשורולוגים שמרוויחים משכורות נמוכות מאוד מרוויחים פחות מאשר מאלפי האריות בעשירונים הנמוכים כך שהממוצע דומה. רגישות זו של התוחלת לתוצאות קיצוניות הופכת אותה במקרים רבים למדד נחות ביחס לחציון. למשל, אם כל הבכירים במשק יחליטו להגדיל את שכרם, המשכורת הממוצעת תעלה למרות ששכרם של כמעט כל השכירים נותר כשהיה. לעומת זאת, המשכורת החציונית לא תשתנה ובכך היא מהווה מדד יותר טוב ל"כמה מרוויחים בשוק".

לעומת זאת, במדעים מדוייקים ותחומי ההנדסה השונים, הממוצע הוא מדד טוב לגדלים נמדדים. הסיבה לכך היא שהרעש המתווסף במערכת המדידה הוא לרוב עם תוחלת 0, כך שכאשר מודדים גודל מסויים מספר פעמים ומחשבים את הממוצע, ממוצע הרעש ישאף לתוחלת (ראו משפט הגבול המרכזי, משפט 6.16, וחוק המספרים הגדולים, משפט 5.10) ויתאפס. הממוצע שיתקבל יהיה בקירוב טוב הגודל האמיתי אותו באמת מנסים למדוד ובו מעוניינים. זו הסיבה שבמקרים רבים פרקטיקה נפוצה בעבודת מעבדה היא למדוד את אותו הגודל מספר רב של פעמים וחישוב הממוצע.

פרט לכך, ניתן לייחס לתוחלת גם משמעות פיזיקלית. כזכור, ניתן להקביל את צפיפות ההסתברות לצפיפות מסה ולחשוב על הסתברות כגודל פיזי המפוזר על פני מוט אינסופי (ציר המספרים). במקרה כזה, לפי נוסחא 5.2 נסיק שהתוחלת היא בדיוק מרכז המסה (מרכז הכובד) של המוט.

תרגיל 5.8. נתון משתנה מקרי רציף שצפיפותו

$$f_X(t)=\begin{cases}c & t\in[-1,0]\\ t^2 & t\in(0,1]\\ 0 & t\notin[-1,1]\end{cases}$$

כאשר $c>0$ קבוע. נגדיר משתנה מקרי חדש באופן הבא:

$$Y=\begin{cases}2 & X>\frac{1}{2}\\ 1 & X\leq\frac{1}{2}\end{cases}$$

מהי $\mathbb{E}(Y)$?

תרגיל 5.9. יהי X משתנה מקרי עם פונקציית צפיפות

$$f_X(t)=\begin{cases}\frac{t^n e^{-t}}{n!} & t\geq 0\\ 0 & t<0\end{cases}$$

מהי התוחלת $\mathbb{E}\left(X\right)$?
רמז: נוסחא **0.36**

תרגיל 5.10. רובה מונח בנקודה $(0,1)$ ויורה בזווית θ ביחס לציר y לכיוון ראשית הצירים (כאשר $\theta=0$ הרובה מכוון בדיוק לראשית הצירים, כאשר $\theta=\frac{\pi}{2}$ הרובה מקביל לציר ה-x). זווית הירי מתפלגת לפי

$$f_\theta(t)=\begin{cases}\frac{1}{\pi} & t\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]\\ 0 & t\notin\left[-\frac{\pi}{2},\frac{\pi}{2}\right]\end{cases}$$

נסמן ב-X את נקודת הפגיעה של הקליע בציר ה-x. חשבו את הצפיפות ואת התוחלת של X.

5.1.1 תכונות של תוחלת

כאשר ההתפלגות של משתנה מקרי נתונה, חישוב התוחלת לפי הנוסחאות הופך לתהליך טכני של חישוב סכום, טור או אינטגרל והוא עשוי להיות מורכב מבחינה חישובית אך פשוט מבחינה רעיונית. הקושי הגדול, במקרים רבים, הוא לא בחישוב התוחלת עצמה אלא, למשל, בחישוב ההתפלגות של המשתנה המקרי שאת תוחלתו מעוניינים לחשב. בסעיף זה נוכיח מספר תכונות של התוחלת אשר יכולות לייתר את הצורך בחישוב ההתפלגות. כך ייצא שבמקרים מסויימים ניתן לדעת את התוחלת של המשתנה המקרי ללא ידע על התפלגותו. למניעת התארכות פרק זה שלא לצורך, נוכיח בכל פעם את המשפטים רק עבור אחד המקרים – הבדיד או הרציף, בהתאם לנוחות. הוכחת האפשרות השנייה דומה מאוד ונשארת לקורא כתרגיל. כמו כן, בכל המשפטים נניח מראש שמדובר על משתנים מקריים בעלי תוחלת. מבחינה אינטואיטיבית, רבות מהתכונות שיוזכרו להלן מוכרות כתכונות של ממוצע והרי התוחלת היא לא יותר מאשר ממוצע.

***משפט* 5.1** (ליניאריות התוחלת). *לכל משתנה מקרי X ולכל $a,b\in\mathbb{R}$ מתקיים*

$$\mathbb{E}\left(aX+b\right)=a\mathbb{E}\left(X\right)+b \tag{5.3}$$

הוכחה. נוכיח עבור משתנה מקרי בדיד X. ההוכחה למקרה הרציף דומה ומסתמכת על תרגיל **4.31**. נגדיר $Y=aX+b$ ונחשב את התוחלת ישירות על ידי המעבר בין ההתפלגות של Y לזו של X:

$$\begin{aligned}\mathbb{E}\left(Y\right)&=\sum_{n\in supp(Y)}n\cdot Pr\left(Y=n\right)=\sum_{n\in supp(Y)}n\cdot Pr\left(aX+b=n\right)\\&=\sum_{n\in supp(Y)}n\cdot Pr\left(X=\frac{n-b}{a}\right)\end{aligned}$$

נבצע את החלפת המשתנים $k=\frac{n-b}{a}$ ונקבל

$$\mathbb{E}\left(Y\right)=\sum_{k\in supp(X)}\left(a\cdot k+b\right)Pr\left(X=k\right)=a\cdot\sum_{k\in supp(X)}k\cdot Pr\left(X=k\right)+b\cdot\sum_{k\in supp(X)}Pr\left(X=k\right)$$

הטור הראשון הוא בדיוק התוחלת של X והטור השני הוא הסכום כל הערכים של פונקציית ההסתברות של X השווה ל-1. לכן $\mathbb{E}\left(Y\right)=a\mathbb{E}\left(X\right)+b$, כנדרש. $\square$

משמעותה של לינאריות התוחלת היא שאם מכפילים את כל הערכים האפשריים בקבוע – גם הממוצע מוכפל באותו קבוע ואם מוסיפים לכל הערכים האפשריים קבוע – גם לממוצע מתווסף קבוע זה. למשל, נניח שגובה של גברים הוא משתנה מקרי עם תוחלת 1.75 מטרים כאשר הם יחפים. נמדוד את הגברים עם נעליים שמוסיפות עוד סנטימטר לגובה ונבצע את החישוב בסנטימטרים במקום במטרים. המשתנה המקרי החדש, גובה הגברים בסנטימטרים כולל הנעל, הוא טרנספורמציה לינארית של המשתנה המקרי הקודם. כדי לדעת מה יהיה גובהו של גבר מסויים בשיטת המדידה החדשה, יש לכפול את גובהו בשיטת המדידה הקודמת ב-100 ולהוסיף לתוצאה 1. פונקציה זו יש להפעיל גם על התוחלת ולכן תוחלת גובה הגברים בשיטת המדידה החדשה תהיה $100\cdot1.75+1=176$ סנטימטרים.

משפט זה ניתן להכללה לא רק לטרנספורמציות לינאריות של המשתנה המקרי אלא גם (כמעט) לכל פונקציה שלו. נוותר הפעם על ההוכחה ולא ניכנס לפרטים בנוגע למה צריכה הפונקציה של המשתנה המקרי לקיים. נסתפק בכך שכל פונקציה שנפגוש בחיים האמיתיים תקיים את המשפט הבא:

***משפט* 5.2** (תוחלת של פונקציה של משתנה מקרי). *יהי X משתנה מקרי ו-$g:\mathbb{R}\to\mathbb{R}$ פונקציה כלשהי. נגדיר משתנה מקרי חדש $Y=g\left(X\right)$. אזי ניתן לחשב את התוחלת של Y ללא מציאת התפלגותו על ידי*

$$\mathbb{E}\left(Y\right)=\sum_{k\in supp(X)}g\left(k\right)\cdot Pr\left(X=k\right)\qquad(5.4)$$

$$\mathbb{E}\left(Y\right)=\int_{-\infty}^{\infty}g(x)\cdot f_X(x)\,dx\qquad(5.5)$$

כאשר נוסחא 5.4 מתאימה למצב בו X משתנה מקרי בדיד ונוסחא 5.5 מתאימה למצב בו X משתנה מקרי רציף.

בעיה 5.4. נתון מקל באורך 5 מטרים. המקל נשבר לשניים בנקודה המרוחקת X מקצהו השמאלי. פונקציית הצפיפות של נקודת השבר נתונה על ידי

$$f_X(x)=\begin{cases}\frac{6}{125}\cdot x(5-x) & x\in[0,5]\\ 0 & x\notin[0,5]\end{cases}$$

מהי התוחלת של Y, שבר המקל הארוך מבין השניים?

הוכחה. בבעיה 4.5 ראינו שהקשר בין X ל-Y הינו $Y=\max\left\{X,5-X\right\}$ ועל סמך קשר זה מצאנו את הצפיפות של Y. אמנם אפשרי להשתמש בצפיפות של Y לצורך חישוב תוחלתו, אך בזכות נוסחא 5.5 ניתן לוותר על כך ולחשב את התוחלת של Y רק על סמך ההתפלגות של X. לצורך כך נרשום את הקשר ביניהם בצורה יותר פשוטה (ראו האיור המופיע בבעיה 4.5):

$$Y(X)=\begin{cases}X & X\geq 2.5\\ 5-X & X\leq 2.5\end{cases}$$

ולכן התוחלת תחושב לפי

$$\begin{aligned}\mathbb{E}(Y)=\mathbb{E}(Y(X))=\int_{-\infty}^{\infty}Y(x)f_X(x)\ dx&=\int_0^{2.5}(5-x)f_X(x)\ dx+\int_{2.5}^{5}xf_X(x)\ dx\\ =\int_0^{2.5}(5-x)\cdot\frac{6}{125}\cdot x(5-x)\,dx+\int_{2.5}^{5}x\cdot\frac{6}{125}\cdot x(5-x)\,dx&=3.438\end{aligned}$$

$\square$

***משפט* 5.3** (סימטריית התוחלת). *יהי X משתנה מקרי בעל תוחלת. אם קיים מספר $\mu\in\mathbb{R}$ כך שלכל $t\in\mathbb{R}$ מתקיים אחד מהתנאים הבאים:*

1. $Pr(X\leq\mu-t)=Pr(X\geq\mu+t)$.
2. *עבור משתנה מקרי בדיד* $P_X(\mu-t)=P_X(\mu+t)$.
3. *עבור משתנה מקרי רציף* $f_X(\mu-t)=f_X(\mu+t)$.

אזי $\mathbb{E}(X)=\mu$.

הוכחה. יהי X משתנה מקרי בעל תוחלת ונוכיח שאם קיים מספר $\mu\in\mathbb{R}$ כך שלכל $t\in\mathbb{R}$ מתקיים $Pr(X\leq\mu-t)=Pr(X\geq\mu+t)$ אזי $\mathbb{E}(X)=\mu$. שאר התנאים בשאלה גוררים תנאי זה מיידית ולכן ניתן להסתפק רק בוהכחה שלו. נגדיר $Y=2\mu-X$ ונחשב את פונקציית ההתפלגות המצטברת של Y:

$$\begin{aligned}F_Y(t)&=Pr(Y\leq t)=Pr(2\mu-X\leq t)=Pr(X\geq 2\mu-t)\\&=Pr(X\geq\mu+(\mu-t))=Pr(X\leq\mu-(\mu-t))=Pr(X\leq t)=F_X(t)\end{aligned}$$

כלומר, פונקציית ההתפלגות המצטברת של X זהה לפונקציית ההתפלגות המצטברת של Y ולכן משתנים מקריים אלו מתפלגים אותו הדבר. לפיכך, יש להם את אותה התוחלת, $\mathbb{E}(X)=\mathbb{E}(Y)$.
יחד עם זאת, מלינאריות התוחלת, $\mathbb{E}(Y)=2\mu-\mathbb{E}(X)$ ולאחר הצבת $\mathbb{E}(Y)=\mathbb{E}(X)$ נקבל $\mathbb{E}(X)=\mu$, כנדרש. $\square$

לאור תכונת הסימטריה, בכל פעם שנמצא סימטריה בבעיה סביב ערך מסויים, ערך זה יהפוך מיד לתוחלת של המשתנה המקרי, גם אם חישוב ישיר של האינטגרל או הטור המגדירים את התוחלת הוא מסובך. יחד עם זאת, חשוב להדגיש שיש לבדוק שהאינטגרל או הטור אכן מתכנסים בהחלט, שכן תנאי מקדים לקיומו של המשפט הוא שהתוחלת קיימת. בתרגיל 5.10 ראינו שדרישה זו חשובה, שכן בתרגיל מופיע משתנה מקרי שאמנם הוא סימטרי סביב 0 אבל לא קיימת לו תוחלת ולכן הוכחה זו לא תקפה עבורו.

תרגיל 5.11. חשבו את התוחלת של המשתנה המקרי הרציף X בעל פונקציית הצפיפות:

$$f_X(t)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)$$

***משפט* 5.4** (חיבוריות התוחלת). *לכל שני משתנים מקריים X,Y מתקיים*

$$\mathbb{E}\left(X+Y\right)=\mathbb{E}\left(X\right)+\mathbb{E}\left(Y\right) \tag{5.6}$$

ובהכללה, לכל n סופי או אינסופי של משתנים מקריים מתקיים:

$$\mathbb{E}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathbb{E}\left(X_i\right) \tag{5.7}$$

הוכחה. נניח ש-X,Y משתנים מקריים בדידים בעלי תוחלת. נזכור כי משתנה מקרי הוא פונקציה ממרחב מדגם Ω ל-$\mathbb{R}$ (הגדרה 4.1), לכן המאורע $\{X=k\}$ הוא למעשה איחוד של מאורעות זרים בזוגות מהצורה $\{\omega_k\}$, כאשר המשותף לכולם הוא ש-$\omega_k\in\Omega$ היא תוצאה במרחב המדגם עבורה $X=k$, כלומר $X\left(\omega_k\right)=k$, וכן $Pr\left(X=k\right)=\sum_{\omega_k}Pr\left(\{\omega_k\}\right)$. נחליף בנוסחא 5.1 את $Pr\left(X=k\right)$ בסכום $\sum_{\omega_k}Pr\left(\{\omega_k\}\right)$. כאשר נבצע את הסכימה על כל התוצאות האפשריות במרחב המדגם, כך שנוסחא 5.1 תשתנה ותהא

$$\mathbb{E}\left(X\right)=\sum_{k\in supp(X)}k\cdot Pr\left(X=k\right)=\sum_{\omega\in\Omega}X\left(\omega\right)\cdot Pr\left(\{\omega\}\right) \tag{5.8}$$

נוסחא חדשה זו לרוב פחות שימושית, שכן עברנו לדון במשתנים מקריים כדי לפשט את הדיון ולהימנע משימוש במרחב המדגם ובכל התוצאות שלו, אך במקרה זה היא תפשט משמעותית את חישוב התוחלת של סכום המשתנים המקריים:

$$\begin{aligned}\mathbb{E}\left(X+Y\right)&=\sum_{\omega\in\Omega}\left(X+Y\right)\left(\omega\right)\cdot Pr\left(\omega\right)=\sum_{\omega\in\Omega}\left(X(\omega)+Y(\omega)\right)\cdot Pr\left(\omega\right)\\&=\sum_{\omega\in\Omega}X(\omega)\cdot Pr\left(\omega\right)+\sum_{\omega\in\Omega}Y(\omega)\cdot Pr\left(\omega\right)=\mathbb{E}\left(X\right)+\mathbb{E}\left(Y\right)\end{aligned}$$

□

בזכות חיבוריות התוחלת ניתן לחשב תוחלות של משתנים מקריים גם אם לא ידועה את התפלגותם, כל עוד יש אפשרות לרשום אותם כסכום של משתנים מקריים אחרים שהתפלגותם או תוחלתם ידועה. כך למשל, אם אנחנו רוצים לדעת את ההכנסה הממוצעת לזוג, איננו צריכים לדעת כמה כל זוג מרוויח אלא די לנו אם נדע את ההכנסה הממוצעת של גברים, ההכנסה הממוצעת של נשים וסכום ממוצעים אלו יהיו ההכנסה הממוצעת של זוגות.

בעיה 5.5. בהגרלת לוטו מוציאים שישה כדורים ללא החזרה מתוך כד הכולל 37 כדורים ממוספרים מ-1 ועד 37. מהי תוחלת סכום המספרים שנבחרו?

הוכחה. נסמן ב-X את סכום ששת המספרים שיצאו בהגרלת הלוטו. לצורך חישוב תוחלתו באמצעות נוסחא 5.1 יש לחשב את התפלגותו אך מדובר על חישוב מורכב מאוד שכן הסכומים השונים יכולים להתקבל בהרבה אופנים שונים וקשה מאוד למנותם לצורך חישוב ההסתברות ולפתח נוסחא כללית. במקום זאת, נשתמש בתכונת החיבוריות של התוחלת ונרשום את X כסכום של משתנים מקריים פשוטים יותר. היות ו-X מוגדר מראש כסכום של שישה כדורים, טבעי להגדיר משתנה מקרי לכל כדור, כלומר נגדיר ב-X_i את הכדור ה-$i\in\{1,\dots,6\}$

שנבחר בהגרלה.
מטעמי סימטריה, הכדור ה-i יכול להיות כל אחד מ-37 הכדורים ולכן $Pr\left(X_i=k\right)=\frac{1}{37}$ לכל $k\in\{1,\ldots,37\}$. לפיכך

$$\mathbb{E}\left(X_i\right)=\sum_{k=1}^{37}k\cdot Pr\left(X=k\right)=\sum_{k=1}^{37}\frac{k}{37}\overset{eq.\,0.20}{=}19$$

כך שתוחלת הסכום של ששת הכדורים תהיה

$$\begin{aligned}\mathbb{E}\left(X\right)=\mathbb{E}\left(\sum_{i=1}^{6}X_i\right)&\overset{eq.\,5.7}{=}\sum_{i=1}^{6}\mathbb{E}\left(X_i\right)\\&=\sum_{i=1}^{6}19=6\cdot19=114\end{aligned}$$

□

חשוב להבהיר שחיבוריות התוחלת נכונה תמיד, ללא תלות בקשר אפשרי כלשהו בין המשתנים המקריים. אמנם בבעיה זו יש קשר בין המשתנים המקריים שכן אם, למשל, המספר 3 נבחר בהוצאה הראשונה, בבירור הוא לא ייבחר בהוצאות הבאות ולכן התוחלת של הכדורים הבאים כבר לא צריכה לכלול אותו. אבל, כאשר אנחנו מחשבים את ההתפלגות והתוחלת של כל משתנה מקרי בנפרד, אנחנו מתעלמים מהכדורים האחרים כי ערכם לא ידוע. אחרי הכל, עד כה בכל פעם שחישבנו התפלגות של משתנה מקרי כלשהו, היינו יכולים עקרונית להגדיר עוד משתנים מקריים בבעיה, אבל עצם ההגדרה של משתנה מקרי נוסף בבעיה לא יכולה להשפיע על ההתפלגות של משתנה מקרי אחר. בהעדר מידע נוסף על הערכים של המשתנים המקריים האחרים, ההתפלגות של כל אחד מהמשתנים המקריים נותרת כשהייתה. דוגמא לכך ראינו בבעיה 3.6, שם לא ידענו מה הקלף שאבד ולכן הסיכוי לכל אחד מהקלפים בחפיסה להיות הקלף הראשון שיישלף זהה.

תרגיל 5.12. חיזרו על בעיה 5.5 כאשר הפעם הכדורים נבחרים עם החזרה.

חשוב לשים לב שהן בבעיה 5.5 והן בתרגיל 5.12 הקשר בין המשתנים המקריים שהגדרנו הוא שהם מתפלגים אותו הדבר ולכן התוחלת שלהם שווה. יחד עם זאת, משתנים מקריים אלו אינם שווים שכן בכל ניסוי הם יכולים (וכאשר הבחירה נעשית ללא החזרה, מוכרחים) לקבל ערכים שונים. לו היינו רושמים בטעות $X=6X_1$ ומשתמשים בלינאריות התוחלת, היינו מקבלים את אותה התשובה המספרית אך חישוב זה מוטעה שכן לרשום $X=6X_1$ בעצם כופה על כל המשתנים המקריים להיות שווים זה לזה או לחילופין מגדיר את המשתנה המקרי X להיות הערך שיצא בהוצאה הראשונה כפול 6, ולא כך הדבר.

תרגיל 5.13 (מונוטוניות התוחלת)**.** יהיו X,Y משתנים מקריים.

1. נניח ש-X משתנה מקרי עם תומך המוכל בקטע $[a,b]$, כלומר בכל תוצאה של הניסוי מתקבל ערך $a\leq X\leq b$. הוכיחו כי גם $a\leq\mathbb{E}\left(X\right)\leq b$.
2. נניח שבכל תוצאה אפשרית של הניסוי $\omega\in\Omega$ מתקיים $X\left(\omega\right)\leq Y\left(\omega\right)$ (ובקיצור: $X\leq Y$). הוכיחו כי גם $\mathbb{E}\left(X\right)\leq\mathbb{E}\left(Y\right)$.

כאשר המשתנה המקרי סופר מספר התרחשויות של מאורעות מוגדרים מראש, ניתן להיעזר בסוג מיוחד של משתנים מקריים המכונים משתנים מציינים (אינדיקטורים) בכדי לפרק את המשתנה המקרי לסכום של משתניים

מקריים פשוטים יותר. הסיבה שאינדיקטורים פשוטים יותר הוא שכל אחד מהם מקבל שני ערכים בלבד: 1 כאשר המאורע המתאים לו קרה ו-0 כאשר הוא לא קרה. אם נסכום על כל האינדיקטורים, אחד לכל מאורע, נדע בדיוק כמה מאורעות קרו.

הגדרה 5.6 (משתנה מקרי מציין (אינדיקטור)). **יהי A מאורע כלשהו. אינדיקטור למאורע A הוא משתנה מקרי היסומן על ידי $\mathbf{1}_A$ ומקבל שני ערכים: 1 כאשר A קרה ו-0 כאשר המאורע A לא קרה. לכן ההתפלגות של האינדיקטור היא:**

$$Pr\left(\mathbf{1}_A = k\right) = \begin{cases} Pr\left(A\right) & k = 1 \\ 1 - Pr\left(A\right) & k = 0 \end{cases} \tag{5.9}$$

והתוחלת שלו היא

$$\mathbb{E}\left(\mathbf{1}_A\right) = 1 \cdot Pr\left(A\right) + 0 \cdot \left(1 - Pr\left(A\right)\right) = Pr\left(A\right) \tag{5.10}$$

בעיה 5.6. באוטובוס 30 נוסעים. נסמן ב-X את מספר תאריכי הלידה השונים שיש לנוסעי האוטובוס. למשל, אם $X = 1$ אזי כולם נולדו באותו היום ואם $X = 30$ אזי כולם נולדו בימים שונים. מהי תוחלתו של X?

הוכחה. בפרדוקס ימי ההולדת (בעיה 3.1) ראינו כיצד לחשב את ההסתברות שלפחות שניים חולקים יום הולדת באותו יום ועשינו זאת באמצעות חישוב ההסתברות שלכולם ימי הולדת בימים שונים, כלומר את הסתברות המאורע $X = 30$. אם נרצה להמשיך ולחשב את שאר ההסתברויות נגלה שזו משימה מסובכת. למשל, המאורע $X = 10$ משמעו שיש 10 תאריכי יום הולדת שונים ובכדי לחשב את הסתברותו צריך לבחור את ימי ההולדת שנבחרו, לבחור איזה מהנוסעים בחר כל יום הולדת ולוודא שאכן 10 הימים נבחרו לפחות פעם אחת.

במקום זאת, נחשב את התוחלת על ידי פירוק לאינדיקטורים. המשתנה המקרי X סופר את מספר הימים בשנה שנבחרו, לכן נגדיר לכל יום בשנה אינדיקטור, $\mathbf{1}_i$, השווה ל-1 אם היום ה-$i \in \{1, \ldots, 365\}$ נבחר ו-0 אם לא. מספר הימים שנבחרו, X, הוא בדיוק מספר האינדיקטורים ששווים ל-1 ולכן

$$\mathbb{E}\left(X\right) = \mathbb{E}\left(\sum_{i=1}^{365} \mathbf{1}_i\right) = \sum_{i=1}^{365} \mathbb{E}\left(\mathbf{1}_i\right)$$

נחשב את ההתפלגות של כל אינדיקטור. המאורע $\{\mathbf{1}_i = 1\}$ מתרחש כאשר היום ה-i נבחר לפחות על ידי אדם אחד ולכן יהיה יותר נוח לחשב דרך המשלים – המאורע שאף אחד לא בחר ביום זה:

$$\mathbb{E}\left(\mathbf{1}_i\right) = Pr\left(\mathbf{1}_i\right) = 1 - \left(\frac{364}{365}\right)^{30} = 0.08$$

נציב בנוסחא לתוחלת של X ונקבל $\mathbb{E}\left(X\right) = 365 \cdot 0.08 = 28.8$.

למרות ש-$1 \leq X \leq 30$, הערך הממוצע יצא מאוד קרוב ל-30. הסיבה לכך היא שכאשר בוחרים באקראי 30 תאריכים מתוך 365, הסיכוי לבחור מספר מועט של תאריכים (למשל, שכולם יבחרו בדיוק את אותו היום) הוא מאורע עם הסתברות מאוד נמוכה ולכן ערכים נמוכים של X יוכפלו בהסתברויות מאוד נמוכות ביחס לערכים יותר גדולים (כזכור מבעיה 3.1, הסיכוי לקבל $X = 30$ הוא 0.3!). הממוצע המשוקלל של הערכים ש-X יכול לקבל יהיה אם כך קרוב לערכים המתקבלים בהסתברויות גבוהות. $\square$

תרגיל 5.14. 50 זוגות מסתדרים באקראי בשורה. מהי התוחלת של מספר האנשים העומדים בין אדון וגברת סמית'?

תרגיל 5.15. פתרו את בעיה 5.5 על ידי פירוק X לסכום של משתנים מקריים אחרים הבודקים האם מספר מסויים נבחר או לא.

עד כה ראינו כיצד מתוך ההתפלגות ניתן לחשב את התוחלת וכיצד ניתן למצוא את התוחלת גם כאשר ההתפלגות לא ידועה. את התהליך ההפוך, חישוב ההתפלגות מתוך התוחלת, לא ניתן לעשות שכן התוחלת היא בסך הכל ממוצע של תוצאות אפשריות בניסוי וניתן לייצר את אותו הממוצע בהרבה דרכים שונות. יחד עם זאת, ניתן לפתח חסמים ואי-שוויונים עבור הסתברויות של מאורעות הקשורים ל-X המתבססים על התוחלת. החסם הראשון שנפגוש הוא חסם מרקוב התקף למשתנים מקריים חיוביים.

***משפט* 5.5** (חסם (אי-שוויון) מרקוב)**.** *אם X משתנה מקרי חיובי, כלומר* $Pr\left(X\geq 0\right)=1$*, אזי לכל* $0<a\in\mathbb{R}$ *מתקיים:*

$$Pr\left(X\geq a\right)\leq\frac{\mathbb{E}\left(X\right)}{a} \tag{5.11}$$

הוכחה. יהי $0<a\in\mathbb{R}$ ונגדיר משתנה מקרי חדש

$$X_a=\begin{cases}a & X\geq a\\ 0 & X<a\end{cases}$$

בבירור מתקיים $X\geq X_a$ וממונוטוניות התוחלת (תרגיל 5.13) גם $\mathbb{E}\left(X\right)\geq\mathbb{E}\left(X_a\right)$. אבל $\mathbb{E}\left(X_a\right)=a\cdot Pr\left(X\geq a\right)$ ולכן

$$\mathbb{E}\left(X\right)\geq a\cdot Pr\left(X\geq a\right)$$

כנדרש. □

הרעיון מאחורי חסם מרקוב הוא שהתוחלת היא ממוצע ההתפלגות ואם ערכים רחוקים מדי מהתוחלת יתקבלו בהסתברות גבוהה, הם ימשכו את הממוצע אליהם. הדרישה לחיוביות המשתנה המקרי היא הכרחית, שכן אם המשתנה המקרי יכול להיות גם שלילי, ניתן לאזן ערכים חיוביים גדולים עם ערכים שליליים גדולים. למשל, לשני המשתנים המקריים הבאים אותה תוחלת למרות שהערכים שהם מקבלים שונים מאוד ובכל מקרה חסם מרקוב לא רלוונטי עבורם:

$$X=\begin{cases}1 & \frac{1}{2}\\ -1 & \frac{1}{2}\end{cases}\quad Y=\begin{cases}10^9 & \frac{1}{2}\\ -10^9 & \frac{1}{2}\end{cases} \tag{5.12}$$

5.2 שונות

בנוסחא 5.12 מוצגים שני הימורים אפשריים. בהימור X ניתן לזכות או להפסיד שקל אחד בהסתברות שווה ואילו בהימור Y ניתן לזכות או להפסיד מיליארד שקלים בהסתברות שווה. בשני ההימורים התוחלת שווה ל-0 אך בכל זאת יש הבדל מהותי בין ההימורים. בהימור X התוצאות האפשריות קרובות יחסית לתוחלת ואילו

בהימור Y התוצאות האפשריות מאוד רחוקות מהתוחלת. בכדי לכמת הבדל זה בין ההימורים נוסיף מדד חדש שתפקידו למדוד את הפיזור של ההתפלגות ביחס לתוחלת, כלומר את המרחק (בריבוע, כדי להימנע מערך מוחלט) הממוצע של הערכים האפשריים בניסוי ביחס לתוחלת. מדד זה נקרא השונות (ו' ראשונה מנוקדת בחולם, ו' שנייה בשורוק) של המשתנה המקרי ומתאר עד כמה התוצאות האפשריות בניסוי קרובות לתוחלת.

הגדרה 5.7 (שונות)**.** יהי X משתנה מקרי בעל תוחלת. השונות של X, המסומנת על ידי $\mathbb{V}\left(X\right)$[2], היא מדד לפיזור של המשתנה המקרי סביב התוחלת שלו ושווה לממוצע המרחק הריבועי של המשתנה המקרי מהתוחלת:

$$\mathbb{V}\left(X\right)=\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^{2}\right) \tag{5.13}$$

תוך שימוש בתכונות של תוחלת ניתן לפשט את נוסחא **5.13** לנוסחא פשוטה הרבה יותר. נפתח סוגריים לפי כפל מקוצר וחיבוריות התוחלת ונקבל:

$$\begin{aligned}\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^{2}\right)&=\mathbb{E}\left(X^{2}-2X\mathbb{E}\left(X\right)+\left(\mathbb{E}\left(X\right)\right)^{2}\right)\\&=\mathbb{E}\left(X^{2}\right)+\mathbb{E}\left(-2X\mathbb{E}\left(X\right)\right)+\mathbb{E}\left(\left(\mathbb{E}\left(X\right)\right)^{2}\right)\end{aligned}$$

למניעת סרבול נסמן $\left(\mathbb{E}\left(X\right)\right)^{2}=\mathbb{E}^{2}\left(X\right)$, כלומר $\mathbb{E}^{2}\left(X\right)$ הוא ריבוע התוחלת של המשתנה המקרי X. בנוסף, נזכור כי $\mathbb{E}\left(X\right)$ הינו מספר ממשי קבוע ולכן ניתן להוציא אותו מחוץ לתוחלת לפי לינאריות התוחלת:

$$\mathbb{E}\left(X^{2}\right)-2\mathbb{E}\left(X\right)\cdot\mathbb{E}\left(X\right)+\mathbb{E}^{2}\left(X\right)=\mathbb{E}\left(X^{2}\right)-\mathbb{E}^{2}\left(X\right)$$

לסיכום, נקבל את הנוסחא השימושית הבאה לחישוב התוחלת:

$$\mathbb{V}\left(X\right)=\mathbb{E}\left(X^{2}\right)-\mathbb{E}^{2}\left(X\right) \tag{5.14}$$

כאשר חישוב התוחלת של X^2 נעשה על ידי הנוסחא לחישוב תוחלת של פונקציה (נוסחאות **5.4** ו-**5.5**) עם הפונקציה $g(x)=x^2$.

אילוף אריות מול אשורולוגיה - שונות. בחישוב התוחלות גילינו כי השכר הממוצע בשני המקצועות מאוד קרוב – $\mathbb{E}\left(T\right)=6.098$ למול $\mathbb{E}\left(S\right)=6.199$. נחשב עבור שתי ההתפלגויות את השונות על מנת לבחון עד כמה ההתפלגות מרוחקת מהממוצע. היות וכבר חישבנו את התוחלת, נותר רק לחשב את התוחלת של המשתנה המקרי בריבוע:

$$\begin{aligned}\mathbb{E}\left(T^{2}\right)&=\int_{-\infty}^{\infty}t^{2}f_{T}\left(t\right)\,dt=\int_{4}^{5}\frac{t^{3}}{17}\,dt+\int_{5}^{10}\frac{t^{2}}{17}(10-t)\,dt=39.128\\\mathbb{E}\left(S^{2}\right)&=\int_{-\infty}^{\infty}t^{2}f_{S}\left(t\right)\,dt=\int_{2}^{6}\frac{2t^{3}}{67}\,dt+\int_{6}^{11}\frac{2t^{2}}{67}(12-t)\,dt=43.097\end{aligned}$$

---

[2] בספרים רבים נפוץ גם הסימון $Var\left(X\right)$

ולכן השונויות הינן

$$\mathbb{V}(T)=\mathbb{E}\left(T^2\right)-\mathbb{E}^2(T)=39.128-6.098^2=2.307$$
$$\mathbb{V}(S)=\mathbb{E}\left(S^2\right)-\mathbb{E}^2(S)=43.097-6.199^2=4.669$$

כצפוי, השונות של האשורולוגיה גבוהה יותר (ראו שרטוט בתחילת הפרק). לפיכך, למרות שממוצע השכר בשני המקצועות דומה, באשורולוגיה ישנם הרבה יותר אנשים ששכרם רחוק מהממוצע מאשר באילוף האריות. ניתן להתייחס לשונות כאל מדד לאי-השוויון בהתפלגות השכר, כך שאילוף האריות הוא מקצוע שיוויוני יותר בו כולם מרוויחים שכר דומה ואילו באשורולוגיה אי-השוויון גדול יותר וישנם הבדלי שכר גדולים יותר בין אלו שמרוויחים שכר גבוה לאלו המרוויחים שכר נמוך. □

בעוד שלשאר המדדים המרכזיים (שכיח, אחוזונים ותוחלת) יש "יחידות" כמו למשתנה המקרי, השונות היא ממוצע משוקלל של המרחק בריבוע ולכן היחידות שלו הן כמו של המשתנה המקרי בריבוע. למשל, עבור מאלפי האריות גילינו שהשכר הממוצע הוא 6.098 אלפי ש"ח ואילו השונות היא 2.307 (אלפי ש"ח)^2. כדי שניתן יהיה להשתמש בערך של השונות כמדד לפיזור, יש לתקן את היחידות שלה כך שיהיו שוות ליחידות של המשתנה המקרי על ידי הוצאת שורש. מספר זה מכונה סטיית התקן. כדאי לציין שהוצאת השורש תמיד אפשרית, שכן לפי נוסחא **5.13** השונות היא תוחלת של משתנה מקרי חיובי ולכן תמיד $\mathbb{V}(X)\geq 0$.

הגדרה 5.8 (סטיית תקן). יהי X משתנה מקרי בעל שונות $\mathbb{V}(X)$. נגדיר את סטיית התקן של המשתנה המקרי בתור

$$\sigma_X=\sqrt{\mathbb{V}(X)}$$

בניגוד לשונות, לסטיית התקן אותן "יחידות" כמו ל-X ולכן היא מהווה מדד עם משמעות לפיזור של הערכים האפשריים של X סביב התוחלת שלו.

עבור אילוף אריות, למשל, סטיית התקן הינה $\sigma_T=\sqrt{2.307}=1.519$. משמעות הדבר היא שחלק ניכר ממאלפי האריות ירוויחו שכר מרוחק סטיית תקן אחת מהתוחלת, כלומר בתחום $[4.579,7.617]$. יחד עם זאת, לא ניתן לקבוע איזה חלק בדיוק ממאלפי האריות נמצאים בטווח זה רק על סמך ערך סטיית התקן ולצורך חישוב זה יש צורך בידיעת ההתפלגות עצמה.

תרגיל 5.16. חשבו את השונות של המשתנה המקרי המופיע בתרגיל **5.9**.

תרגיל 5.17. חשבו את התוחלת והשונות של משתנה מקרי שפונקציית הצפיפות שלו הינה

$$f_X(t)=\begin{cases}\frac{3}{t^4} & t\geq 1\\ 0 & t<1\end{cases}$$

כפי שלא תמיד מוגדרת התוחלת, לא תמיד מוגדרת השונות. על מנת שהשונות תהיה מוגדרת הטור או האינטגרל של המשתנה המקרי בריבוע צריך להתכנס, כלומר אנו דורשים את התכנסות הטור $\sum_{k\in supp(X)}k^2P_X(k)$ עבור משתנה מקרי בדיד ואת ההתכנסות של האינטגרל $\int_{-\infty}^{\infty}t^2f_X(t)\,dt$ עבור משתנה מקרי רציף. כאשר תנאי זה לא מתקיים נאמר שאין למשתנה המקרי שונות. במקרה כזה למשתנה המקרי יש סיכוי גבוה לקבל ערכים רחוקים מאוד מהתוחלת, כך שהמרחק הממוצע שלהם מהתוחלת הוא אינסופי.

תרגיל 5.18. בכל אחד מהסעיפים הבאים מצאו משתנה מקרי רציף המקיים את התנאי המבוקש או הוכיחו שלא קיים משתנה מקרי כזה:

1. משתנה מקרי רציף בעל תוחלת ובעל שונות.
2. משתנה מקרי רציף בעל תוחלת וללא שונות.
3. משתנה מקרי רציף ללא תוחלת ובעל שונות.
4. משתנה מקרי רציף ללא תוחלת וללא שונות.

5.2.1 תכונות של שונות

בדומה לתוחלת, גם עבור השונות ישנן מספר תכונות שימושיות אשר מפשטות את חישוב השונות גם במקרים בהם ההתפלגות מסובכת. חשוב לזכור שהשונות מחושבת באמצעות התוחלת ולכן הרבה מהתכונות של השונות דומות לאלו של התוחלת בשינויים הנדרשים, הנובעים מכך שהתוחלת היא אופרטור לינארי ואילו השונות היא תוחלת של ריבוע המשתנה המקרי ולא המשתנה המקרי עצמו. נתחיל מהמקבילה של השונות לתכונת הלינאריות של התוחלת.

***משפט* 5.6** ("לינאריות" השונות)**.** *יהי X משתנה מקרי בעל שונות. אזי לכל $a, b \in \mathbb{R}$ מתקיים*

$$\mathbb{V}\left(aX+b\right) = a^2\mathbb{V}\left(X\right) \qquad (5.15)$$

הוכחה. נחשב את השונות ישירות מתוך ההגדרה תוך שימוש בלינאריות התוחלת:

$$\begin{aligned}\mathbb{V}\left(aX+b\right) &= \mathbb{E}\left(\left(aX+b\right)^2\right) - \mathbb{E}^2\left(aX+b\right) = \mathbb{E}\left(a^2X^2+2abX+b^2\right) - \left(a\mathbb{E}\left(X\right)+b\right)^2 \\ &= a^2\mathbb{E}\left(X^2\right) + 2ab\mathbb{E}\left(X\right) + b^2 - \left(a\mathbb{E}\left(X\right)\right)^2 - 2ab\mathbb{E}\left(X\right) - b^2 \\ &= a^2\left(\mathbb{E}\left(X^2\right) - \mathbb{E}^2\left(X\right)\right) = a^2\mathbb{V}\left(X\right)\end{aligned}$$

□

לאור נוסחא 5.15, הוספה של קבוע כלל לא משפיע על הפיזור. תוצאה זו הגיונית, שכן אם נוסיף לכל הסטודנטים בקורס מסויים פקטור של 5 נקודות, דבר זה יגדיל את הציון של כולם ב-5 נקודות וכך גם את הממוצע אך לא ישנה את הפיזור של הסטודנטים סביב הממוצע החדש. לעומת זאת, אם נכפיל את השכר של כל מאלפי האריות פי 10, מה שבעצם ביצענו זה מתיחה של ציר ה-X בהתפלגות פי 10. מתיחה זו תגדיל פי 10 את הפיזור סביב הממוצע (וכמובן גם תשנה את הממוצע), ולכן סטיית התקן תגדל פי 10 והשונות, שהיא ריבוע סטיית התקן, תגדל פי 100.

כמו תכונת הלינאריות כך גם תכונת החיבוריות לא מתקיימת במלואה עבור שונות. הבעיה נעוצה בכך שייתכן ויש קשר בין המשתנים המקריים שיגדיל או יקטין את השונות. למשל, אם $Y = 2X$, אזי השונות של $X+Y = 3X$ לפי ההגדרה צריכה להיות $\mathbb{V}\left(3X\right) = 9\mathbb{V}\left(X\right)$ אך אם נחבר את השונויות של X ו-Y כמו שהן נקבל רק $\mathbb{V}\left(X\right) + \mathbb{V}\left(2X\right) = 5\mathbb{V}\left(X\right)$. בהמשך (פרק 7.4) נגדיר מפורשות מהו טיב הקשר שצריך להיות בין המשתנים כדי שהשונות לא תהיה חיבורית ונפתח את הנוסחא הכללית המתארת שונות של סכום משתנים מקריים תוך לקיחת הקשר בין המשתנים המקריים בחשבון. בפרק זה נסתפק באמירה הכללית שכאשר המשתנים המקריים בלתי תלויים, כלומר כאשר ההתפלגות של משתנה מקרי אחד לא מושפעת מהמשתנה המקרי האחר, השונות היא חיבורית. גם אל מושג התלות בין משתנים מקריים נחזור בהמשך (פרק 7) ונגדירו פורמלית.

משפט 5.7 ("חיבוריות" השונות)***. יהיו X,Y משתנים מקריים בלתי תלויים בעלי שונות. אזי מתקיים***

$$\mathbb{V}\left(X+Y\right)=\mathbb{V}\left(X\right)+\mathbb{V}\left(Y\right) \tag{5.16}$$

זהו מקרה פרטי של משפט 7.8 העוסק במקרה הכללי של כל שני משתנים מקריים, לאו דווקא בלתי תלויים.

נשמיט את הוכחת המשפט שכן מדובר על מקרה פרטי של משפט 7.8. מסיבה זו גם נמעיט בחישובי השונות המתבססים על חיבוריות כל עוד לא נראה את הנוסחא הכללית והמלאה.

תרגיל 5.19. יהי X משתנה מקרי עם תוחלת μ ושונות σ^2. מגרילים את הערכים של X שוב ושוב ובאופן בלתי תלוי n פעמים. נסמן ב-X_i את תוצאת ההגרלה ה-$i\in\{1,\ldots,n\}$ וב-$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$ את ממוצע התוצאות שהתקבלו. מהי התוחלת ומהי השונות של הממוצע?

ניתן לנצל את השונות בכדי לפתח חסם נוסף להתפלגות. בחסם מרקוב (נוסחא **5.11**) התבססנו על זה שמשתנה מקרי חיובי לא יכול לקבל ערכים גדולים מדי שכן הם "יסחבו" אחריהם את התוחלת. באופן דומה, משתנה מקרי לא יכול לקבל ערכים רחוקים מדי מהתוחלת שלו, שכן אז הם "יסחבו" אחריהם את השונות. לכן, באמצעות חישוב השונות, ניתן לחסום את הסיכוי שמשתנה מקרי כלשהו רחוק מדי מהתוחלת שלו. חסם זה נקרא חסם (אי-שוויון) צ'בישב.

משפט 5.8 (חסם (אי-שוויון) צ'בישב)***. יהי X משתנה מקרי בעל תוחלת ושונות. אזי לכל $a>0$ מתקיים***

$$Pr\left(\left|X-\mathbb{E}\left(X\right)\right|\geq a\right)\leq\frac{\mathbb{V}\left(X\right)}{a^2}$$

הוכחה. ברור כי מתקיים $Pr\left(\left|X-\mathbb{E}\left(X\right)\right|\geq a\right)=Pr\left(\left(X-\mathbb{E}\left(X\right)\right)^2\geq a^2\right)$ ולכן אם נסמן $Y=\left(X-\mathbb{E}\left(X\right)\right)^2$ נקבל משתנה מקרי חיובי עליו ניתן להפעיל את אי שוויון מרקוב:

$$Pr\left(Y\geq a^2\right)\leq\frac{\mathbb{E}\left(Y\right)}{a^2}=\frac{\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^2\right)}{a^2}\overset{eq.\,5.13}{=}\frac{\mathbb{V}\left(X\right)}{a^2}$$

כנדרש. $\square$

באמצעות אי-שוויון מרקוב ניתן לקבל חסם להסתברות להתרחק מהתוחלת תוך ניצול העובדה שהשונות היא מדד לריכוזיות ההתפלגות. כך בין היתר, ניתן להעריך את ההסתברות להיות במרחק קטן ממספר סטיות תקן מהתוחלת. למשל, אם נבחר $a=3\sigma_X$ נקבל מאי-שוויון צ'בישב כי $Pr\left(\left|X-\mathbb{E}\left(X\right)\right|\geq 3\sigma_X\right)\leq\frac{\mathbb{V}(X)}{9\sigma_X^2}=\frac{1}{9}$. כלומר בתוך טווח של $\pm3\sigma_X$ מסביב לתוחלת צריכה להיות לפחות $\frac{8}{9}$ מההתפלגות.

בעיה 5.7. חוקר מעוניין לגלות כמה מתוך הסטודנטים באוניברסיטת תל אביב הם סטודנטים להנדסה. לצורך כך הוא בוחר באקראי n סטודנטים ושואל אותם האם הם סטודנטים להנדסה או לא. לאחר מכן הוא מחשב את Y – אחוז הסטודנטים מתוך כלל הנשאלים שאמרו לו שהם לומדים הנדסה. כמה סטודנטים עליו לשאול על מנת שיוכל להיות בטוח בהסתברות של לפחות 0.8 שהערך שהוא קיבל בניסוי לא סוטה מהערך האמיתי ביותר מ-5%? הניחו לשם הפשטות שהבחירה של הסטודנטים נעשית עם החזרה ושמתוך כלל הסטודנטים רק $p\in[0.1,0.3]$ הם סטודנטים להנדסה (זהו הגודל הלא ידוע אותו החוקר מנסה לגלות).

הוכחה. נפתח בהמרת השאלה ממילים לנוסחא מתמטית. בשאלה מבקשים למצוא את n עבורו הסטייה מערך האמיתי, p, קטנה מ-5% בהסתברות של לפחות 0.8, כלומר

$$Pr\left(|Y-p|\leq 0.05\right)\geq 0.8$$

נוסחאו זו מזכירה את חסם צ'בישב לכן נחשב את התוחלת והשונות של Y (התלויים כנראה ב-n) וננסה לחסום את ההסתברות המבוקשת. נחשב את התוחלת והשונות באמצעות פירוק לאינדיקטורים.

נסמן ב-$\mathbf{1}_i$ את האינדיקטור לתשובה של הסטודנט ה-i במדגם (**1** אם הוא סטודנט להנדסה, **0** אם לאו). במדגם כולו נבחרו $\sum_{i=1}^{n}\mathbf{1}_i$ סטודנטים להנדסה ואחוז הסטודנטים להנדסה מתוך כלל הסטודנטים, אשר הוגדר בשאלה בתור Y, הינו

$$Y=\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_i$$

המשתנים המקריים $\mathbf{1}_i$ הם שווי התפלגות ומקבלים את הערך **1** בהסתברות p כך שלפי נוסחא **5.10** מתקיים $\mathbb{E}\left(\mathbf{1}_i\right)=p$. בנוסף, כאשר $\mathbf{1}_i=1$ גם $\mathbf{1}_i^2=1$ וכאשר $\mathbf{1}_i=0$ גם $\mathbf{1}_i^2=0$, כלומר המשתנים המקריים $\mathbf{1}_i$ ו-$\mathbf{1}_i^2$ מקבלים בדיוק את אותם ערכים באותן ההסתברויות ולכן שווי התפלגות ושווי תוחלת: $\mathbb{E}\left(\mathbf{1}_i^2\right)=\mathbb{E}\left(\mathbf{1}_i\right)=p$.
לאור זאת, השונות של כל אינדיקטור תהא

$$\mathbb{V}\left(\mathbf{1}_i\right)=\mathbb{E}\left(\mathbf{1}_i^2\right)-\mathbb{E}^2\left(\mathbf{1}_i\right)=p-p^2=p(1-p) \qquad (5.17)$$

הסטודנטים הנבחרים הם בלתי תלויים והבחירה נעשית עם החזרה, לכן האינדיקטורים הם בלתי תלויים והשונות של Y הינה

$$\mathbb{V}\left(Y\right)=\mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n}Y_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}\left(Y_i\right)=\frac{n\left(p-p^2\right)}{n^2}=\frac{p(1-p)}{n}$$

לפי אי-שוויון צ'בישב מתקיים

$$Pr\left(|Y-p|\geq 0.05\right)\leq\frac{\mathbb{V}\left(Y\right)}{0.05^2}$$

ולכן ניתן לחסום את ההסתברות המבוקשת דרך המשלים:

$$Pr\left(|Y-p|\leq 0.05\right)=1-Pr\left(|Y-p|\geq 0.05\right)\geq 1-\frac{\mathbb{V}\left(Y\right)}{0.05^2}=1-400\cdot\frac{p(1-p)}{n}$$

הדרישה היא שהסתברות זו תהיה לפחות 0.8 כלומר

$$1-400\cdot\frac{p(1-p)}{n}\geq 0.8$$

ועל ידי חילוץ n נקבל:

$$n\geq\frac{400}{0.2}\cdot p(1-p)=2000\cdot p(1-p)$$

הנחנו כי $p \in [0.1, 0.3]$ והמקסימום של הפונקציה $p(1-p)$ בקטע מתקבל בנקודה $p = 0.3$. נציב אותה ונקבל $n \geq 420$. כלומר, אם מסתמכים על חסם צ'בישב, יש לשאול לפחות 420 סטודנטים בכדי להיות קרובים להסתברות האמיתית עד כדי 5% בהסתברות של לפחות 0.8. בהמשך נראה דרך אחרת באמצעותה ניתן לחשב הסתברות זו בקירוב טוב יותר ולקבל חסם הדוק יותר על התוצאה באמצעות משפט הגבול המרכזי (ראו בעיה 6.8). $\square$

תרגיל 5.20. שיכור יוצא מראשית הצירים וצועד באקראי על ציר המספרים. השיכור צועד ימינה בהסתברות 0.6 ושמאלה בהסתברות המשלימה. מאחר וההסתברות לצעוד ימינה גבוהה יותר, אנחנו מצפים שככל שהוא יצעד יותר, כך הוא יצעד יותר לכיוון ימין ויימצא בעיקר בחצי הציר הימני. הוכיחו זאת מפורשות – הראו שההסתברות להימצא בחצי הציר השלילי שואפת ל-0 כאשר מספר הצעדים שואף לאינסוף.

תרגיל 5.21. מטילים קוביה הוגנת 10 פעמים. מצאו חסם הדוק ככל הניתן להסתברות שסכום ההטלות יהיה גדול או שווה ל-50.

הלוגיקה שהובילה לפיתוח חסם צ'בישב ניתן לסיבוב גם לכיוון השני לפיתוח חסם על השונות עצמה. אם השונות נתונה, המשתנה המקרי לא יכול לקבל ערכים רחוקים מהתוחלת בהסתברות גבוהה מדי, שכן אחרת השונות הייתה צריכה להיות גדולה יותר. ההפך גם נכון – אם טווח הערכים האפשרי של המשתנה המקרי חסום, גם השונות לא יכולה להיות גבוהה מאוד כי יש גבול עד כמה הערכים יכולים להיות רחוקים מהתוחלת.

***משפט* 5.9.** *יהי X משתנה מקרי כך שקיימים מספרים $a, b \in \mathbb{R}$ עבורם תמיד $a \leq X \leq b$. אזי השונות של X חסומה ומתקיים*

$$\mathbb{V}(X) \leq \frac{(b-a)^2}{4} \tag{5.18}$$

הוכחה. נוכיח למקרה בו $a = 0, b = 1$. במקרה כזה $0 \leq X \leq 1$ ולכן $X^2 \leq X$, כך שממונוטוניות התוחלת גם $\mathbb{E}\left(X^2\right) \leq \mathbb{E}(X)$. מכאן השונות חסומה על ידי

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X) \leq \mathbb{E}(X) - \mathbb{E}^2(X)$$

נתבונן בפונציה $f(x) = x - x^2$ זוהי פרבולה המקבלת מקסימום בנקודה $x = \frac{1}{2}$ וערך המקסימום הוא $f\left(\frac{1}{2}\right) = \frac{1}{4}$. בפרט, אם נציב $x = \mathbb{E}(X)$ נקבל $\mathbb{E}(X) - \mathbb{E}^2(X) = f(\mathbb{E}(X)) \leq f\left(\frac{1}{2}\right) = \frac{1}{4}$, כלומר $\mathbb{V}(X) \leq \frac{1}{4}$, כנדרש. $\square$

תרגיל 5.22. השלימו את ההוכחה למקרה הכללי והראו שלא ניתן למצוא חסם טוב יותר, כלומר שאכן קיים משתנה מקרי ש-$\frac{(b-a)^2}{4}$ היא שונותו.

בדיון שסבב את הגדרת התוחלת ציינו שכאשר חוזרים על ניסוי שוב ושוב, ממוצע התוצאות שיתקבלו הוא התוחלת. תוצאה זו מכונה החוק החלש של המספרים הגדולים וניתן להוכיחה באמצעות אי-שוויון צ'בישב. מבחינה פורמלית, החוק טוען שאם נחזור על ניסוי שוב ושוב, ההסתברות שהתוצאה הממוצעת תסטה מהתוחלת היא אפס כאשר מספר החזרות שואף לאינסוף. הסיבה שתוצאה זו נקראית חוק המספרים הגדולים היא שתופעת ההתכנסות לממוצע היא תופעה שדורשת הרבה הטלות. במספרים קטנים, קל לקבל תוצאות נדירות. ככל שנחזור על הניסוי יותר פעמים, התוצאות הנדירות יתמצעו החוצה.

לדוגמא, נניח שישנם שני בתי חולים, באחד יש בכל יום אלף לידות ובשני, הקטן יותר, יש בכל יום עשר לידות. בנים ובנות נולדים בהסתברות שווה, לכן נצפה שבשני בתי החולים, בממוצע, ב-50% מהלידות יוולדו בנים וב-50% מהלידות יוולדו בנות. לפי חוק המספרים הגדולים, הממוצע בבית החולים הגדול יהיה קרוב יותר לתוחלת

והסטייה מ-50% תקרה בהסתברות קטנה יותר, לעומת בית החולים הקטן. כך למשל, הסיכוי לקבל ביום מסויים לידות רק ממין אחד בבית החולים הקטן הוא בסך הכל $\left(\frac{1}{2}\right)^9 = 0.00195$ ולכן בערך פעם בשנתיים צפוי יום כזה. מצד שני, בבית החולים הגדול, הסיכוי שביום מסויים כל הלידות הוא רק ממין אחד הוא $\left(\frac{1}{2}\right)^{999} \to 0$ וגם אם נמתין עד קץ כל הימים מאורע זה ככל הנראה לא יתרחש.

***משפט* 5.10** (החוק החלש של המספרים הגדולים)**.** *יהי X משתנה מקרי עם תוחלת μ ושונות σ^2 ונסמן ב-$X_1, \ldots, X_n$ הגרלות חוזרות ובלתי תלויות של ערכים לפי ההתפלגות של X. אזי לכל $\epsilon > 0$ מתקיים*

$$\lim_{n\to\infty} Pr\left(\left|\overline{X}_n - \mu\right| > \epsilon\right) = 0 \tag{5.19}$$

כאשר $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ הוא ממוצע n ההגרלות הראשונות.

הוכחה. לפי תרגיל 5.19 מתקיים $\mathbb{E}\left(\overline{X}_n\right) = \mu$ וכן $\mathbb{V}\left(\overline{X}_n\right) = \frac{\sigma^2}{n}$. לכן לכל $\epsilon > 0$ מתקיים לפי אי-שוויון צ'בישב:

$$Pr\left(\left|\overline{X}_n - \mu\right| > \epsilon\right) \le \frac{\sigma^2}{n\epsilon} \xrightarrow[n\to\infty]{} 0$$

ולכן גם

$$\lim_{n\to\infty} Pr\left(\left|\overline{X}_n - \mu\right| > 0\right) = 0$$

□

קיים גם חוק חזק של מספרים גדולים אשר משמעותו דומה לחוק החלש וההבדל ביניהם עדין מכדי להיכנס אליו במסגרת זו. שני החוקים יחד מכונים לרוב "חוקי המספרים הגדולים" וברוב המוחלט של המקרים שניהם תקפים ובעלי משמעות זהה. לעומתם, העובדה שבחזרות מועטות קל מאוד לקבל תוצאות נדירות מכונה לעיתים בבדיחות "חוק המספרים הקטנים".

5.3 פונקציה יוצרת מומנטים

פונקציות יוצרות הוא רעיון שאינו ייחודי רק להסתברות ומופיע בתחומים נוספים ורבים. הרעיון הכללי הוא איגוד מספרים רבים המתייחסים לתופעה מסויימת לכדי פונקציה אחת (או באופן כללי, לכדי אובייקט מתמטי אחד), כך שתכונות מסויימות וחשובות של המספרים ייגזרו מתוך האובייקט המתמטי. התופעות הנחקרות בתורת ההסתברות הן משתנים מקריים והמספרים המעניינים אותנו הן המומנטים של המשתנים המקריים, קרי התוחלות של חזקות שונות של משתנה מקרי. שני מומנטים פגשנו כבר בפרק זה, המומנט הראשון, הלוא הוא התוחלת, והמומנט השני, התוחלת של X^2 המהווה את המרכיב המרכזי בחישוב השונות. המומנטים מסדרים גבוהים יותר מתארים גם כן את הפילוג של ההתפלגות. למשל המומנט מסדר שלישי מתאר את הסימטריה של ההתפלגות ביחס לתוחלת (מכונה צידוד), המומנט הרביעי מתאר את עוצמתן של תוצאות נדירות בהתפלגות (מכונה גבנוניות) וכן הלאה[3].

[3]מומנטים דומים מופיעים בתחומי מדע אחרים העוסקים בהתפלגויות של צפיפויות. למשל, במכניקה של גוף קשיח, המומנט הראשון הינו מרכז המסה (מחושב בדרך זהה לזו של התוחלת, כאשר צפיפות ההסתברות מוחלפת בצפיפות המסה) והמומנט השני הוא מומנט ההתמד של הגוף כאשר מסובבים אותו ביחס למרכז המסה.

הגדרה 5.9 (מומנטים של משתנה מקרי)**.** יהי X משתנה מקרי. המומנט מסדר $n \in \mathbb{N}$ של המשתנה המקרי הינו התוחלת של X^n, כלומר $\mathbb{E}\left(X^n\right)$, והיא קיימת כאשר האינטגרל (למשתנה מקרי רציף) או הטור (למשתנה מקרי בדיד) המתארים את התוחלת מתכנסים בהחלט.

הדרך לאסוף את כל המומנטים לכדי פונקציה אחת היא על ידי טור חזקות של פונקציה מעריכית (נוסחא 0.18). לכל $t \in \mathbb{R}$ נתבונן בטור $\sum_{n=0}^{\infty} \frac{t^n \mathbb{E}(X^n)}{n!}$ ונשתמש בלינאריות התוחלת בכדי להפוך את הטור לפונקציה בודדת:

$$\sum_{n=0}^{\infty} \frac{t^n \mathbb{E}\left(X^n\right)}{n!} = \sum_{n=0}^{\infty} \mathbb{E}\left(\frac{t^n X^n}{n!}\right) = \mathbb{E}\left(\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right) \overset{eq. 0.18}{=} \mathbb{E}\left(e^{tX}\right) \tag{5.20}$$

הפונקציה $M_X\left(t\right) = \mathbb{E}\left(e^{tX}\right)$ נקראת הפונקציה יוצרת מומנטים של המשתנה המקרי X ולאור פיתוח זה הפונקציה מכילה בתוכה את כל המידע על כל המומנטים של המשתנה המקרי X.

הגדרה 5.10 (פונקציה יוצרת מומנטים)**.** יהי X משתנה מקרי. הפונקציה $M_X\left(t\right) : \mathbb{R} \to \mathbb{R}$ המוגדרת על ידי

$$M_X\left(t\right) = \mathbb{E}\left(e^{tX}\right) \tag{5.21}$$

היא הפונקציה יוצרת מומנטים של X, כאשר חישוב התוחלת נעשה על ידי נוסחאות 5.4 ו-5.5 כאשר t הינו פרמטר קבוע.
חשוב לציין שפונקציה זו לא חייבת להיות מוגדרת לכל t (כי לא בהכרח הטור או האינטגרל יתכנסו) או בכלל. לצורך הדיון מספיק שהפונקציה תהיה מוגדרת בסביבה כלשהי של $t = 0$.

כאמור, הפונקציה יוצרת המומנטים מכילה בתוכה את כל המידע על כל המומנטים של המשתנה המקרי. הדרך לחלץ את המומנטים היא באמצעות גזירה:

***משפט* 5.11.** *יהי X משתנה מקרי ותהי $M_X\left(t\right)$ פונקציה יוצרת מומנטים שלו. אזי $M_X\left(0\right) = 1$ ולכל $n > 0$ המומנט ה-n של X מתקבל על ידי גזירה (אם הפונקציה גזירה ב-0):*

$$\mathbb{E}\left(X^n\right) = \left.\frac{d^n}{dt^n} M_X\left(t\right)\right|_{t=0} \tag{5.22}$$

הוכחה. נשתמש בנוסחא 5.20 על מנת לגזור $n > 0$ פעמים את $M_X\left(t\right)$:

$$\frac{d^n}{dt^n} M_X\left(t\right) = \frac{d^n}{dt^n} \sum_{k=0}^{\infty} \frac{t^k \mathbb{E}\left(X^k\right)}{k!} = \sum_{k=0}^{\infty} \frac{d^n}{dt^n} \frac{t^k \mathbb{E}\left(X^k\right)}{k!} = \sum_{k=n}^{\infty} n \cdot \ldots \cdot (n-k+1) \frac{t^{k-n} \mathbb{E}\left(X^k\right)}{k!}$$

שכן כאשר גוזרים את t^k יותר מאשר k פעמים מקבלים זהותית 0. נציב $t = 0$ בביטוי האחרון. אם $k > n$ אז $t^{k-n} = 0$ ולכן הביטוי היחיד שלא מתאפס בטור הינו הביטוי המתאים ל-$k = n$ וערכו:

$$\frac{d^n}{dt^n} M_X\left(t\right) = n! \cdot \frac{E\left(X^n\right)}{n!} = \mathbb{E}\left(X^n\right)$$

עבור $n = 0$ לא גוזרים אלא מציבים מפורשות בפונקציה היוצרת מומנטים $t = 0$ ונקבל:

$$M_X\left(0\right) = \mathbb{E}\left(e^{0 \cdot X}\right) = \mathbb{E}\left(1\right) = 1$$

□

תרגיל 5.23. חשבו את הפונקציה יוצרת מומנטים של המשתנים המקריים הבאים:

1. משתנה מקרי בדיד עם פונקציית ההתפלגות

$$Pr\left(X=k\right)=\binom{n}{k}p^{k}(1-p)^{n-k} \qquad (5.23)$$

כאשר $n\in\mathbb{N}$ ו-$p\in[0,1]$ קבועים ו-$supp\left(X\right)=\{0,\ldots,n\}$.

2. משתנה מקרי רציף עם פונקציית הצפיפות

$$f_{X}\left(t\right)=\begin{cases}\lambda e^{-\lambda t} & t\geq0\\ 0 & t<0\end{cases}$$

כאשר $\lambda>0$ קבוע.

תרגיל 5.24. הראו כי אם X מ"מ בעל פונקציה יוצרת מומנטים $M_X(t)$ ואם $a,b\in\mathbb{R}$ אז $M_{aX+b}(t)=e^{bt}M_X(at)$.

בשלב זה ההגדרה של פונקציה יוצרת מומנטים לא נראית שימושית במיוחד. ניתן למצוא כל אחד מהמומנטים בצורה יחסית פשוטה על ידי חישוב האינטגרל או הטור המתאים ישירות. חישוב המומנט דרך פונקציה יוצרת מומנטים מהווה עבודה רבה מדי – יש לחשב אינטגרל או טור מסובכים יותר (עם אקספוננט במקום עם חזקה) ולאחר מכן יש לגזור את הפונקציה המתקבלת. אכן, השימוש בפונקציה יוצרת מומנטים לחישוב המומנטים אינו השימוש העיקרי בה, פרט למקרים פשוטים, בהם ידוע הפירוק של הפונקציה המתקבלת לטור חזקות. תחת זאת, לפונקציה היוצרת מומנטים שלושה שימושים עיקריים בהם נפגוש.

ראשית, הפונקציה היוצרת מומנטים עשויה לעזור בחישוב אינטגרלים או טורים מסובכים. בפרק 6 נכיר משפחות של משתנים מקריים נפוצים ולכל אחד מהם נחשב את הפונקציה היוצרת מומנטים. באמצעותה, ניתן יהיה לחסוך זמן בחישוב אינטגרלים וטורים שצורתם מזכירה את הצורה של הפונקציה היוצרת מומנטים. נניח למשל שבמהלך פתרון של בעיה הגענו לביטוי $\sum_{k=0}^{10}\binom{10}{k}e^{-3k}\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{10-k}$ אותו יש לחשב. במקום להתאמץ ולנסות לחשב את הטור הזה מפורשות, קל לשים לב שמדובר למעשה על $M_X\left(-3\right)$ של המשתנה המקרי המופיע בנוסחא 5.23 עם $n=10$ ו-$p=\frac{1}{4}$. לכן ניתן להציב -3 בפונקציה היוצרת מומנטים שחושבה בתרגיל 5.23 ולקבל את ערך הטור במהירות.

שנית, הפונקציה היוצרת מומנטים שימושית מאוד בהוכחות של משפטים העוסקים בסכום של משתנים מקריים בלתי תלויים. בהמשך נראה שחישוב ההתפלגות של סכום של שני משתנים מקריים היא עבודה מאתגרת והדבר נכון שבעתיים לסכום של שלושה, ארבעה ו-n משתנים מקריים. אך כפי שתוכיחו בתרגיל 7.38, הפונקציה היוצרת של הסכום היא מכפלת הפונקציות היוצרות של המשתנים המקרים וניתן להשתמש בעובדה זו על מנת לקבל באמצעים פשוטים יחסית את הפונקציה היוצרת מומנטים של סכום המשתנים המקריים ולהסיק על הסכום מסקנות. למשל, אם תתקבל פונקציה יוצרת מומנטים של משתנה מקרי ידוע ומוכר – הבעיה נפתרה. בנוסף, עובדה זו תהיה מרכזית בהוכחת משפט הגבול המרכזי (משפט 6.16) ומשפטים אחרים העוסקים בסכום של משתנים מקריים.

שלישית, ניתן להשתמש בפונקצה יוצרת מומנטים על מנת לבנות חסם נוסף להסתברות ש-$X\geq a$, המכונה חסם צ'רנוף. מאחר וחסם צ'רנוף משתמש בבת אחת בכל המומנטים האפשריים, בניגוד לחסמים הקודמים אשר כל אחד מהם הסתמך על מומנט אחד בלבד, בהרבה מקרים חסם זה יהיה יותר טוב מאשר מרקוב וצ'בישב.

***משפט* 5.12** **(חסם (אי-שוויון) צ'רנוף).** *יהי X משתנה מקרי עם פונקציה יוצרת מומנטים $M_X(t)$ ויהי $a \in \mathbb{R}$.*
אזי לכל $t > 0$ מתקיים

$$Pr(X \geq a) \leq M_X(t)\, e^{-ta} \tag{5.24}$$

אי-השוויון מתקיים לכל $t > 0$ ולכן כדאי לחקור את הפונקציה $M_X(t)\, e^{-ta}$ למציאת המינימום המוחלט שלה בתחום $(0, \infty)$ ולהשתמש בערך המינימלי בתור חסם.

לרוב מציאת חסם לפי צ'רנוף תהיה מדויקת יותר אך גם מסובכת יותר, שכן היא דורשת חקירת פונקציה ומציאת מינימום ולכן כמעט ולא נעסוק בחסם זה בהמשך.

5.4 שאלות מסכמות

תרגיל 5.25. הכדורסלן מתרגיל 2.29 מקבל נקודה אחת על קליעה מוצלחת מהעונשין ושלוש נקודות על קליעה מוצלחת מקשת השלוש. מהי תוחלת הניקוד שיקבל הכדורסלן אחרי שתי הזריקות?

תרגיל 5.26. מהי התוחלת והשונות של X, מספר השחקנים המחליפים שמופיעים במופע המתואר בשאלה 2.30?

תרגיל 5.27. נתון משתנה מקרי X המקבל רק שני ערכים בהסתברות חיובית, אחד מהם הינו -3. ידוע כי $\mathbb{E}(X) = 0, \mathbb{V}(X) = 4.5$. מהו הערך השני ש-$X$ מקבל?

תרגיל 5.28. מטילים קוביה הוגנת פעמיים. מהי התוחלת והשונות של סכום התוצאות? מהי התוחלת והשונות של ההפרש בין תוצאת ההטלה הראשונה לתוצאת ההטלה השניה?

תרגיל 5.29. יהי X משתנה מקרי רציף המתפלג לפי פונקציית הצפיפות

$$f_X\left(t\right) = \begin{cases} 2t & t \in [0,1] \\ 0 & t \notin [0,1] \end{cases}$$

1. ודאו שזו אכן פונקציית צפיפות של משתנה מקרי.
2. מהי התוחלת והשונות של X?
3. מהי התוחלת והשונות של $Y = \sqrt{X}$?
4. מהי התוחלת והשונות של $Z = \frac{1}{X}$?

תרגיל 5.30. בבחינה אמריקאית n שאלות עם 4 תשובות. סטודנט שנמנם רק בחלק מההרצאות יודע בהסתברות p את התשובה הנכונה לשאלה. אם אינו יודע את התשובה הנכונה – הוא מנחש באקראי את אחת התשובות. על כל תשובה נכונה הסטודנט יקבל $\frac{100}{n}$ נקודות ועל כל תשובה שגוייה ייקנס ב-$\frac{40}{n}$ נקודות.

1. מהי תוחלת ציון הבחינה של הסטודנט?
2. כמה נקודות צריך המרצה להוריד לכל תשובה שגויה בכדי שסטודנט שלא יודע כלום יקבל בתוחלת 0 בבחינה?

תרגיל 5.31. בכד n כדורים ממוספרים מ-1 ועד n. מוציאים באקראי וללא החזרה m כדורים. מהי התוחלת של X, מספר הכדור המקסימלי שהוצא.
היעזרו בתרגיל 4.7.

תרגיל 5.32. יהי X משתנה מקרי אשר לא מקבל את הערך 0. הוכיחו או הפריכו:

$$\mathbb{E}\left(\frac{1}{X}\right) = \frac{1}{\mathbb{E}\left(X\right)}$$

תרגיל 5.33. יהי X משתנה מקרי רציף בעל פונקציית צפיפות

$$f_X\left(t\right) = \begin{cases} \frac{2}{t^3} & t \geq 1 \\ 0 & t < 1 \end{cases}$$

מצאו חסם הדוק ככל הניתן (מבין חסם מרקוב וחסם צ'בישב) להסתברות $Pr\left(X \geq 100\right)$ וכן חשבו הסתברות זו מפורשות. מהו טיב החסם?

תרגיל 5.34. אליס בוחרת מספר באקראי לפי איזושהי התפלגות. נסמן את המספר הנבחר ב-X. בוב, היודע את ההתפלגות אך לא את המספר הנבחר, מנסה לנחש אותו. אם בוב ניחש את המספר y והמספר שאליס הגרילה הוא X, התשלום שבוב משלם לאליס יהיה $(y-X)^2$. איזה מספר כדאי לבוב לנחש כדי שתוחלת התשלום שהוא משלם יהיה הנמוך ביותר? מה המשמעות של התוצאה שקיבלתם?

תרגיל 5.35. יהי X משתנה מקרי אי-שלילי. הוכיחו כי אם X בדיד עם תומך $\mathbb{N}$ אזי

$$\mathbb{E}\left(X\right)=\sum_{k=0}^{\infty}Pr\left(X>k\right)$$

ואם X רציף אזי

$$\mathbb{E}\left(X\right)=\int_{0}^{\infty}Pr\left(X>t\right)\,dt=\int_{0}^{\infty}\left(1-F_{X}\left(t\right)\right)\,dt$$

5.5 פתרונות לשאלות הפרק

תרגיל 5.1. הטענה נכונה עבור $a = 0$ באופן טריוויאלי. אחרת, לפי תרגיל 4.31 פונקציית הצפיפות של Y תהיה $\frac{1}{|a|} f_X\left(\frac{t-b}{a}\right)$ ולמציאת השכיח נגזור אותה ונשווה לאפס:

$$0 = \frac{df_Y(t)}{dt} = \frac{d\left(\frac{1}{|a|} f_X\left(\frac{t-b}{a}\right)\right)}{dt} = \frac{1}{|a|} \cdot \frac{1}{a} \cdot f_X'\left(\frac{t-b}{a}\right)$$

הערך אפס מתקבל רק כאשר הנגזרת של f_X מתאפסת, כלומר בנקודה x_m. לכן השכיח של Y מקיים $x_m = \frac{y_m-b}{a}$ או $y_m = ax_m + b$ כנדרש. גזירה נוספת תראה שזהו אכן מקסימום שכן סימן הנגזרת השנייה של f_Y נקבע רק לפי הסימן של f_X''. □

תרגיל 5.2. לשם הדוגמא נבנה פונקציה שאינה חד-חד ערכית כך ששני ערכים שונים של X יותאמו לאותו ערך של Y. ההסתברות של ערך זה תגדל והוא יהפוך לשכיח.

נגדיר משתנה מקרי בדיד X על ידי $X = \begin{cases} -1 & \frac{4}{12} \\ 0 & \frac{5}{12} \\ 1 & \frac{3}{12} \end{cases}$ ונתבונן בפונקציה $g(x) = x^2$. לכן המשתנה המקרי Y מוגדר על ידי $Y = g(X) = \begin{cases} 1 & \frac{7}{12} \\ 0 & \frac{5}{12} \end{cases}$. קל לראות שהערך השכיח של X הינו 0 אבל הערך השכיח של Y הינו 1 ולא מתקיים $g(0) = 1$. □

תרגיל 5.3. באופן דומה לחישוב האחוזון ה-90 המשוואה שתתקבל עבור אילוף אריות תהיה

$$t_{0.5}^2 - 20t_{0.5} + 83 = 0$$

ופתרונה $t_{0.5} = 5.877$. לפיכך מעמד הביניים מורכב מכל מאלפי האריות המרוויחים בין $0.75 \cdot 5.877 = 4.408$ אלפי ש"ח לבין $1.25 \cdot 5.877 = 7.346$ אלפי ש"ח. לכן הסיכוי להיות במעמד הביניים של תחום אילוף האריות הינו

$$Pr\left(4.408 \le T \le 7.346\right) = \int_{4.408}^{5} \frac{1}{17} t\, dt + \int_{5}^{7.346} \frac{1}{17}(10-t)\, dt = 0.692$$

עבור אשורולוגיה המשוואה תהיה

$$s_{0.5}^2 - 24s_{0.5} + 109.5 = 0$$

ופתרונה $s_{0.5} = 6.126$. לכן מעמד הביניים נמצא בתחום $[0.75s_{0.5}, 1.25s_{0.5}] = [4.595, 7.658]$ והסיכוי להיות במעמד הביניים הינו

$$Pr\left(4.595 \le S \le 7.658\right) = \int_{4.595}^{6} \frac{2}{67} t\, dt + \int_{6}^{7.658} \frac{2}{67}(12-t)\, dt = 0.478$$

ניתן לראות שמעמד הביניים באשורולוגיה קטן יותר מאשר באילוף אריות. ניתן היה לצפות לתשובה זו גם על סמך גרפי הצפיפות, שכן התפלגות ההכנסות בקרב מאלפי האריות מרוכזת יחסית בתחום קטן ואילו התפלגות ההכנסות של האשורולוגים רחבה בהרבה. לאור זאת נסיק שמעמד הביניים אינו פלח קבוע של האוכלוסיה אלא משתנה עם שינוי התפלגות ההכנסות. □

תרגיל 5.4. **1.** פונקציית ההתפלגות המצטברת של המשתנה המקרי המתאר את המקסימום בהטלת קוביה 3 פעמים מחושב בנוסחא **4.6**. לפי הנוסחא, לכל $t < 5$ מתקיים

$$F_X(t) \leq F_X(4) = \frac{4^3}{6^3} = \frac{8}{27} < \frac{1}{2}$$

בעוד ש $F_X(5) = \frac{5^3}{6^3} > \frac{1}{2}$, כלומר החציון הוא $x_{0.5} = 5$.

2. פונקציית ההתפלגות המצטברת של זמן חימום הפיצה הינה פונקציה עולה ממש בתחום $t > 0$, ולכן תהיה לה רק נקודת חיתוך אחת עם 0.5. נפתור את המשוואה $F_X(x_{0.5}) = 0.5$ למציאת החציון:

$$\frac{x_{0.5}}{1 + x_{0.5}} = \frac{1}{2}$$

והפתרון הינו $x_{0.5} = 1$.

3. אמנם פונקציית ההתפלגות המצטברת היא פונקציה עולה, אבל היות והיא פולינום ממעלה שלישית, המשוואה שתתקבל תהיה משוואה ממעלה שלישית שפתרונה האנליטי מסובך. במקום זאת, נשים לב שפונקציית הצפיפות סימטרית ביחס ל-50, שכן לכל $t > 0$ מתקיים $f_X(50+t) = f_X(50-t)$. לכן, על ידי חישוב האינטגרל נקבל $Pr(X > t) = Pr(X < t)$. לכן $Pr(X > 50) = Pr(X < 50) = \frac{1}{2}$ כך שהחציון הוא 50.
בתרגיל **4.16** ניתנת תוספת של **10** נקודות לתחום $[50, 60]$. לכן, גם $F_X(60) = 0.5$ וזהו הערך המקסימלי עבורו ההתפלגות המצטברת שווה ל-**50**, לכן החציון יהיה 60.
בתרגיל **4.20** משתנה רק ההתפלגות בתחום $[55, 60]$ והשינוי **לא** משפיע על החציון שנותר **50**.

$\square$

תרגיל 5.5. יהי X משתנה מקרי ויהיו $a, b \in \mathbb{R}$ כך ש-$a > 0$. נחשב את האחוזון ה-α של Y:

$$\alpha \geq Pr(Y \leq y_\alpha) = Pr(aX + b \leq y_\alpha) = Pr\left(X \leq \frac{y_\alpha - b}{a}\right) = F_X\left(\frac{y_\alpha - b}{a}\right)$$

לפי ההגדרה, הערך המקסימלי המקיים $F_X(x) \leq \alpha$ הוא x_α ולכן $x_\alpha = \frac{y_\alpha - b}{a}$ כך ש-$y_\alpha = ax_\alpha + b$.
כאשר $a < 0$ החילוק ב-a הופך את סימן אי-השוויון:

$$\alpha \geq Pr\left(X \geq \frac{y_\alpha - b}{a}\right) = 1 - F_X\left(\frac{y_\alpha - b}{a}\right)$$

לכן למעשה אנחנו מחפשים את האחוזון ה-$1 - \alpha$ ומתקיים $y_\alpha = ax_{1-\alpha} + b$. $\square$

תרגיל 5.6. בבעיה **4.3** חישבנו ומצאנו כי פונקציית הצפיפות של התוצאה המקסימלית בשלוש הטלות קוביה הינה

$$P_X(k) = \frac{k^3 - (k-1)^3}{6^3} = \frac{3k^2 - 3k + 1}{6^3}$$

לכן התוחלת תתקבל על ידי חישוב ישיר:

$$\mathbb{E}(X)=\sum_{k=1}^{6}k\cdot P_X(k)=\sum_{k=1}^{6}\frac{3k^3-3k^2+k}{6^3}=\frac{1071}{216}=4.958$$

תוצאה זו הגיונית שכן לערכים נמוכים יש הסתברות נמוכה כך שהממוצע המשוקלל צריך להיות קרוב לערכים הגבוהים בתומך. $\square$

תרגיל 5.7. בבעיה **4.6** ראינו שהמרחק ממרכז המטרה מתפלג לפי פונקציית ההתפלגות המצטברת

$$F_R(r)=\begin{cases}0 & r<0\\ \frac{r^2}{10^2} & r\in[0,10]\\ 1 & r>10\end{cases}$$

נחשב באמצעותה את ההתפלגות של המשתנה המקרי המתאר את הניקוד המתקבל ביריה בודדת, X. המאורע $X=10$ מתקבל כאשר מרחק הפגיעה קטן מסנטימטר, כלומר

$$Pr(X=10)=Pr(R<1)=F_R(1)=\frac{1}{100}$$

באותו האופן ההסתברויות של התוצאות האחרות הינן:

$$Pr(X=7)=Pr(1<R<4)=F_R(4)-F_R(1)=\frac{15}{100}$$
$$Pr(X=3)=Pr(4<R<8)=F_R(8)-F_R(4)=\frac{48}{100}$$

אין צורך לחשב את ההסתברות של המאורע $X=0$ שכן הסתברות זו מוכפלת באפס ולא משפיעה ישירות על חישוב התוחלת. לכן תוחלת הניקוד הינה

$$\begin{aligned}\mathbb{E}(X)&=10\cdot Pr(X=10)+7\cdot Pr(X=7)+3\cdot Pr(X=3)\\&=10\cdot\frac{1}{100}+7\cdot\frac{15}{100}+3\cdot\frac{48}{100}=2.59\end{aligned}$$

$\square$

תרגיל 5.8. מאחר ו-Y משתנה מקרי בדיד המקבל רק שני ערכים, מספיק שנחשב את ההסתברות של אחד מהם על מנת לחשב את התוחלת. נעדיף לחשב את ההסתברות של המאורע $\{Y=2\}$שכן מאורע זה מתייחס רק לתחום ה-X-ים עבורם הצפיפות ידועה ואין צורך לחשב את הקבוע c(אם כי הדבר אפשרי). מתקיים

$$Pr(X>0.5)=\int_{0.5}^{1}t^2\,dt=\left.\frac{t^3}{3}\right|_{x=0.5}^{1}=0.292$$

כלומר $Pr(Y=2)=0.292$ ולכן $Pr(Y=1)=1-Pr(Y=2)$ כך שהתוחלת הינה

$$\mathbb{E}(Y)=2\cdot0.292+1\cdot(1-0.292)=1.292$$

הערה: ניתן לחשב תוחלת זו גם על ידי הנוסחא לתוחלת של פונקציה של משתנה מקרי. □

תרגיל 5.9. נחשב את התוחלת ישירות לפי ההגדרה:

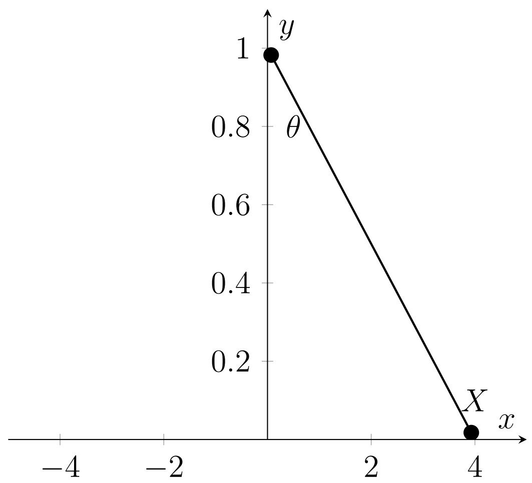
$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t \cdot f_X(t)\ dt = \int_0^{\infty} \frac{t^{n+1}e^{-t}}{n!}\ dt \overset{eq.0.36}{=} \frac{(n+1)!}{n!} = n+1$$

□

תרגיל 5.10. ראשית נחשב את פונקציית ההתפלגות המצטברת של θ. הצפיפות שונה מאפס רק בתחום $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ ולכן למעשה רק שם האינטגרל על הצפיפות יהיה לא טריוויאלי ונקבל:

$$F_\theta(t) = \int_{-\infty}^{t} f_\theta(u)\ du = \begin{cases} 0 & t < -\frac{\pi}{2} \\ \frac{t}{\pi} + \frac{1}{2} & t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ 1 & t > \frac{\pi}{2} \end{cases}$$

שנית, נחשב את פונקציית ההתפלגות המצטברת של X. נקודת הפגיעה מקיימת $X = \tan(\theta)$ ולכן נחשב את ההתפלגות דרך חישוב ההתפלגות המצטברת:



$$F_X(t) = Pr(X \le t) = Pr(\tan(\theta) \le t) = Pr(\theta \le \arctan(t)) = F_\theta(\arctan(t))$$

מאחר ולכל $t \in \mathbb{R}$ מתקיים $\arctan(t) \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, נציב רק בחלק זה של פונקציית הצפיפות ונקבל

$$F_X(t) = F_\theta(\arctan(t)) = \frac{\arctan(t)}{\pi} + \frac{1}{2}$$

לקבלת הצפיפות נגזור את פונקציית ההתפלגות המצטברת:

$$f_X(t) = \frac{dF_X(t)}{dt} = \frac{1}{\pi(1+t^2)}$$

מפתה לחשב את התוחלת על ידי ביצוע האינטגרל:

$$\mathbb{E}\left(X\right)=\int_{-\infty}^{\infty}t\cdot f_X\left(t\right)\,dt=\int_{-\infty}^{\infty}\frac{t}{\pi\left(1+t^2\right)}\,dt=\lim_{s\to\infty}\int_{-s}^{s}\frac{t}{\pi\left(1+t^2\right)}\,dt=0$$

כאשר תוצאת האינטגרל מתקבלת מכך שזהו אינטגרל על פונקציה אי-זוגית בקטע סימטרי סביב 0. תוצאה זו גם הגיונית, שכן נקודת הפגיעה סימטרית סביב $X=0$ ולכן הגיוני שבממוצע נקבל 0, אבל חישוב זה אינו נכון! מדובר על משתנה מקרי שיכול לקבל ערכים שליליים ולכן על מנת שהתוחלת תהיה מוגדרת יש קודם לוודא שהאינטגרל מתכנס בהחלט:

$$\int_{-\infty}^{\infty}|t|\cdot f_X(t)\,dt=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{|t|}{1+t^2}\,dt=\frac{1}{\pi}\lim_{t\to\infty}\ln(1+t^2)=\infty$$

ומכאן שהאינטגרל לא מתכנס בהחלט ו-X אינו בעל תוחלת.

בנוסף, החישוב שהוצע לעיל כלל לא נכון מתמטית - אינטגרל מ-$-\infty$ ל-$+\infty$ שקול לגבול

$$\lim_{t\to\infty}\lim_{s\to-\infty}\int_{s}^{t}f(x)\,dx$$

כלומר גבול כפול בו הגבול העליון והגבול התחתון של האינטגרל לא חייבים לשאוף לאינסוף באותו הקצב. □

תרגיל 5.11. לכל $t\in\mathbb{R}$ מתקיים

$$f_X\left(-t\right)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\left(-t\right)^2}{2}\right)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)=f_X\left(t\right)$$

כלומר פונקציית הצפיפות של X סימטרית סביב $\mu=0$. בנוסף, למשתנה המקרי X קיימת תוחלת שכן האינטגרל $\int_{-\infty}^{\infty}|t|\exp\left(-t^2\right)\,dt$ מתכנס (למשל, על ידי משפטי השוואה סטנדרטיים מחדו"א) ולכן לפי משפט 5.3 $\mathbb{E}\left(X\right)=0$.

ניתן כמובן גם לחשב את האינטגרל ישירות על ידי זיהוי הנגזרת $\frac{d}{dt}\left(\exp\left(-\frac{t^2}{2}\right)\right)=-t\exp\left(-\frac{t^2}{2}\right)$:

$$\mathbb{E}\left(X\right)=\int_{-\infty}^{\infty}\frac{t}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)\,dt=\frac{-1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)\Big|_{-\infty}^{\infty}=0$$

אך זו דרך ארוכה יותר ומורכבת יותר. □

תרגיל 5.12. העובדה שהניסוי נעשה עם החזרה לא משנה דבר בחישוב. ההגדרה של כל המשתנים המקריים נותרת זהה, התוחלות זהות ולכן גם התוחלת של הסכום זהה. ההבדל היחיד הוא שהפעם גם אם היינו יודעים את הערך של X_1 או היינו יודעים שכדור 3 כבר נבחר, זה לא היה משנה את ההתפלגות של ההוצאות הבאות ולא את התוחלת שלהן, שכן הניסוי נעשה עם החזרה. □

תרגיל 5.13. 1. יהי X משתנה מקרי בדיד עם תומך המוכל בקטע $[a,b]$. לכן מתקיים:

$$\mathbb{E}\left(X\right)=\sum_{k\in supp(X)}k\cdot Pr\left(X=k\right)\geq\sum_{k\in supp(X)}a\cdot Pr\left(X=k\right)=a\cdot\sum_{k\in supp(X)}Pr\left(X\right)$$

הסכום האחרון הוא סכום כל ההסתברויות ולכן שווה ל-1, כלומר $\mathbb{E}\left(X\right)\geq a$. באופן דומה, $\mathbb{E}\left(X\right)\leq b$ וההוכחה דומה למקרה הרציף.

2. נגדיר $Z = Y - X$. לכל תוצאה אפשרית של הניסוי מתקיים $X \leq Y$ ולכן $Z \geq 0$, כך שלפי הסעיף הקודם גם $\mathbb{E}(Z) \geq 0$. נשתמש בלינאריות התוחלת:

$$\mathbb{E}(Z) = \mathbb{E}(Y - X) = \mathbb{E}(Y) - \mathbb{E}(X) \geq 0$$

ולכן גם $\mathbb{E}(Y) \geq \mathbb{E}(X)$.

$\square$

תרגיל 5.14. נגדיר X – מספר האנשים בין אדון וגברת סמית' ונשתמש בפירוק לאינדיקטורים בכדי לחשב את התוחלת. לכל אדם $i = 1, 2..., 98$ נסמן ב- $\mathbf{1}_i$ את האינדיקטור לכך שהאדם ה-i עומד בין מר וגברת סמית'. מתקיים $X = \sum_{i=1}^{98} \mathbf{1}_i$ ולכן

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{98} \mathbf{1}_i\right) = \sum_{i=1}^{98} \mathbb{E}(\mathbf{1}_i)$$

אמנם השאלה עוסקת בסידורים בשורה של 100 אנשים, אבל כאשר רוצים לבדוק האם אדם i נמצא בין אדון וגברת סמית', אפשר לצמצם את מרחב המדגם רק לשלושת אנשים אלו ונקבל שהסיכוי שהוא אכן עומד ביניהם הינו $Pr(\mathbf{1}_i = 1) = \frac{1}{3}$ (מבין 3 המקומות ל-3 האנשים, כדי לעמוד ביניהם הוא חייב לבחור במקום האמצעי). זוהי גם התוחלת של האינדיקטור ולכן $\mathbb{E}(X) = \sum_{i=1}^{98} \frac{1}{3} = \frac{98}{3}$.

$\square$

תרגיל 5.15. לכל אחד מהכדורים $i \in \{1, \ldots, 37\}$ נגדיר משתנה מקרי באופן הבא: $X_i = i$ כאשר מספר זה עלה בגורל בהגרלת הלוטו ו-0 אחרת. הגדרה זו אינה אינדיקטור שכן הערכים ש-X_i מקבל אינם 0 ו-1, אבל עדיין התומך של כל משתנה מקרי כזה בן 2 איברים ולכן קל לחשב את התוחלת שלו. היות ו-X הוא סכום הכדורים שנבחרו, בדיוק שישה מתוך 37 המשתנים המקריים שהגדרנו יהיו שונים מאפס והסכום שלהם הוא סכום הכדורים שנבחרו, כך שנקבל $X = \sum_{i=1}^{37} X_i$. נותר לחשב את ההסתברות שכל מספר נבחר. הסיכוי שכדור i נבחר הינו

$$Pr(X_i = i) = \frac{6}{37}$$

שכן הוצאה ללא החזרה אנלוגית לסידורים בשורה של כל הכדורים וכדי שכדור i יבחר הוא צריך להיות באחד המקומות 1-6. לכן התוחלת של X תהיה:

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}\left(\sum_{i=1}^{37} X_i\right) = \sum_{i=1}^{37} \mathbb{E}(X_i) \\ &= \sum_{i=1}^{37} \left(i \cdot Pr(X_i = i) + 0 \cdot (1 - Pr(X_i = i))\right) = \sum_{i=1}^{37} \frac{6}{37} \cdot i = 114\end{aligned}$$

וכמובן שקיבלנו את אותה התוצאה כמקודם.

$\square$

תרגיל 5.16. פונקציית הצפיפות הינה

$$f_X(t) = \begin{cases} \frac{t^n e^{-t}}{n!} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

וראינו בתרגיל 5.9 כי $\mathbb{E}\left(X\right)=n+1$. בנוסף, לפי נוסחא 0.36 מתקיים

$$\begin{aligned}\mathbb{E}\left(X^2\right) &= \int_{-\infty}^{\infty} t^2\cdot f_X\left(t\right)\,dt = \int_0^{\infty}\frac{t^{n+2}e^{-t}}{n!}\,dt\\ &\overset{eq.0.36}{=} \frac{(n+2)!}{n!} = (n+2)(n+1)\end{aligned}$$

לכן השונות הינה

$$\begin{aligned}\mathbb{V}\left(X\right) &= \mathbb{E}\left(X^2\right)-\mathbb{E}^2\left(X\right) = (n+2)(n+1)-(n+1)^2\\ &= (n+1)(n+2-n-1) = n+1\end{aligned}$$

□

תרגיל 5.17. נחשב את התוחלת לפי ההגדרה:

$$\mathbb{E}\left(X\right)=\int_1^{\infty} t\cdot\frac{3}{t^4}\,dt=\int_1^{\infty}\frac{3}{t^3}\,dt=-\frac{3}{2t^2}\Big|_1^{\infty}=\frac{3}{2}$$

לצורך חישוב השונות נחשב את התוחלת של X^2:

$$\mathbb{E}\left(X^2\right)=\int_1^{\infty} t^2\cdot\frac{3}{t^4}\,dt=\int_1^{\infty}\frac{3}{t^2}\,dt=-\frac{3}{t}\Big|_1^{\infty}=3$$

ולכן השונות הינה

$$\mathbb{V}\left(X\right)=\mathbb{E}\left(X^2\right)-\mathbb{E}^2\left(X\right)=3-\left(\frac{3}{2}\right)^2=\frac{3}{4}$$

□

תרגיל 5.18.

1. כל משתנה מקרי המוגדר על קטע סופי הוא משתנה מקרי בעל תוחלת ובעל שונות, למשל המשתנים המקריים המופיעים בסיפור על רעות.

2. קל לראות כי הפונקציה $f_X(t)=\begin{cases}0 & t<1\\ 2t^{-3} & t\geq 1\end{cases}$ היא פונקציית צפיפות של משתנה מקרי רציף שכן היא חיובית בכל התחום ומנורמלת. מתקיים

$$\mathbb{E}\left(X\right)=\int_{-\infty}^{\infty} t\cdot f_X\left(t\right)\,dt=\int_1^{\infty}\frac{2}{t^2}\,dt=2$$

לכן X בעל תוחלת. לעומת זאת, האינטרל של $\mathbb{E}\left(X^2\right)$ הינו

$$\mathbb{E}\left(X^2\right)=\int_{-\infty}^{\infty} t^2\cdot f_X\left(t\right)\,dt=\int_1^{\infty}\frac{2}{t}\,dt$$

וזהו אינטגרל מתבדר, לכן אין למשתנה מקרי זה שונות.

3. משתנה מקרי כזה לא קיים, שהרי השונות מודדת פיזור סביב התוחלת ואם אין תוחלת, כיצד וסביב מה נמדוד את הפיזור?

ניתן להוכיח זאת גם בצורה מתמטית: נניח בשלילה כי קיים משתנה מקרי רציף ללא תוחלת ובעל שונות שפונקציית הצפיפות שלו הינה $f_X(t)$. לכן האינטגרל $\int_{-\infty}^{\infty}|t|f_X(t)\,dt$ מתבדר בעוד שהאינטגרל $\int_{-\infty}^{\infty}t^2f_X(t)\,dt$ מתכנס. נחסום את האינטגרל עבור התוחלת ובכך נפריך את הטענה:

$$\begin{aligned}\int_{-\infty}^{\infty}|t|f_X(t)\,dt &= \int_{-\infty}^{-1}|t|f_X(t)\,dt+\int_{-1}^{1}|t|f_X(t)\,dt+\int_{1}^{\infty}|t|f_X(t)\,dt\\ &\le \int_{-\infty}^{-1}t^2f_X(t)\,dt+\int_{-1}^{1}f_X(t)\,dt+\int_{1}^{\infty}t^2f_X(t)\,dt\\ &\le \int_{-\infty}^{\infty}t^2f_X(t)\,dt-\int_{-1}^{1}t^2f_X(t)\,dt+1\end{aligned}$$

כאשר באי השוויון הראשון אנו מסתמכים על כך ש $t^2\ge|t|$ עבור $|t|>1$ ובאי השוויון השני חסמנו את $\int_{-1}^{1}f_X(t)\,dt$ ע"י 1 (תכונת הנרמול). כמו כן השתמשנו בכך שכאשר $|t|<1$ גם $t^2<1$ ולכן $\int_{-1}^{1}t^2f_X(t)\,dt\le\int_{-1}^{1}f_X(t)\,dt\le 1$. הסכום אליו הגענו הוא סכום של שני אינטגרלים מתכנסים. לפיכך, $\int_{-\infty}^{\infty}|t|f_X(t)\,dt$ הינו אינטגרל של פונקציה חיובית החסומה מלעיל על ידי אינטגרל מתכנס ולכן לא יכול להתבדר וזו סתירה. לסיכום, לא יתכן משתנה מקרי רציף בעל שונות וללא תוחלת (הטענה נכונה באופן כללי לכל משתנה מקרי, לאו דווקא רציף).

4. קל לראות כי הפונקציה $f_X(t)=\begin{cases}0 & t<1\\ t^{-2} & t\ge 1\end{cases}$ היא פונקציית צפיפות של משתנה מקרי רציף שכן היא חיובית בכל התחום ומנורמלת אבל האינטגרלים $\int_{-\infty}^{\infty}|t|f_X(t)\,dt$ ו- $\int_{-\infty}^{\infty}t^2f_X(t)\,dt$ אינם מתכנסים ולכן לפונקציה זו אין תוחלת ואין שונות.

□

תרגיל 5.19. מאחר וכל ה-X_i-ים מתפלגים לפי ההתפלגות של X, לכולם אותה התוחלת ואותה השונות. לכן, ניתן לחשב את התוחלת של הממוצע על ידי תכונות הלינאריות והחיבוריות:

$$\mathbb{E}\left(\overline{X}_n\right)=\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i\right)=\frac{1}{n}\sum_{i=1}^{n}\mu=\frac{n\mu}{n}=\mu$$

כצפוי, התוחלת של הממוצע היא התוחלת של המשתנה המקרי, דבר התואם את ההגדרה והאינטואיציה הראשונית, שהתוחלת היא הממוצע של התוצאות כאשר נחזור על הניסוי פעמים רבות.

לצורך חישוב השונות נשתמש גם כן בלינאריות ובחיבוריות, כפי שהן מוגדרות עבור השונות:

$$\mathbb{V}\left(\overline{X}_n\right)=\mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}\left(X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$$

כלומר, ככל שחוזרים על הניסוי יותר פעמים, המרחק של הממוצע מהתוחלת ילך ויקטן. ניתן לתאר תופעה זו כהתכנסות של הממוצע אל התוחלת, כפי שנראה במשפט הגבול המרכזי (משפט 6.16). □

תרגיל 5.20. נסמן ב-X את מיקום השיכור לאחר n צעדים. לצורך חישוב התוחלת והשונות של X, נרשום אותו כסכום הצעדים שלו. נסמן את הצעד ה-i של השיכור ב-$\mathbf{1}_i$. לאור נתוני השאלה, ההתפלגות של כל צעד הינה $\mathbf{1}_i = \begin{cases} 1 & 0.6 \\ -1 & 0.4 \end{cases}$. זהו לא בדיוק אינדיקטור, שכן הוא מקבל ערך אחר במקום 0, אך באופן דומה לחישוב שביצענו עבור אינדיקטורים נקבל:

$$\begin{aligned}
\mathbb{E}\left(\mathbf{1}_i\right) &= 1 \cdot 0.6 - 1 \cdot 0.4 = 0.2 \\
\mathbb{E}\left(\mathbf{1}_i^2\right) &= 1^2 \cdot 0.6 + (-1)^2 \cdot 0.4 = 1 \\
\mathbb{V}\left(\mathbf{1}_i\right) &= \mathbb{E}\left(\mathbf{1}_i^2\right) - \mathbb{E}^2\left(\mathbf{1}_i\right) = 1 - 0.2^2 = 0.96
\end{aligned}$$

ומיקום השיכור לאחר n צעדים הינו $X = \sum_{i=1}^n \mathbf{1}_i$. מאחר והצעדים בלתי תלויים, ניתן לחשב את התוחלת והשונות של מיקומו על ידי סכימת התוחלות והשונויות של כל אחד מהצעדים:

$$\begin{aligned}
\mathbb{E}\left(X\right) &= \mathbb{E}\left(\sum_{i=1}^n \mathbf{1}_i\right) = \sum_{i=1}^n \mathbb{E}\left(\mathbf{1}_i\right) = 0.2n \\
\mathbb{V}\left(X\right) &= \mathbb{V}\left(\sum_{i=1}^n \mathbf{1}_i\right) = \sum_{i=1}^n \mathbb{V}\left(\mathbf{1}_i\right) = 0.96n
\end{aligned}$$

נחסום באמצעות אי-שוויון צ'בישב את הסטייה של השיכור מהתוחלת ועל ידי כך נוכיח שהוא תמיד יהיה בחצי הציר החיובי. נבחר למשל $a = 0.1n$. מתקיים

$$Pr\left(|X - \mathbb{E}\left(X\right)| \geq 0.1n\right) \leq \frac{\mathbb{V}\left(X\right)}{\left(0.1n\right)^2} = \frac{0.96n}{0.1^2 n^2} \xrightarrow[n\to\infty]{} 0$$

לכן המשלים שואף ל-1:

$$Pr\left(|X - \mathbb{E}\left(X\right)| \leq 0.1n\right) = Pr\left(0.1n \leq X \leq 0.3n\right) \xrightarrow[n\to\infty]{} 1$$

המאורע המבוקש, $\{X \geq 0\}$, מכיל את המאורע $\{0.1n \leq X \leq 0.3n\}$ ולכן גם הסתברותו שואפת ל-1 כאשר מספר הצעדים שואף לאינסוף וההסתברות למצוא את השיכור בחצי המישור השלילי שואף ל-0. □

תרגיל 5.21. נסמן את תוצאת ההטלה ה-$i \in \{1, \ldots, 10\}$ ב-X_i ואת סכום עשרת ההטלות ב-$X = \sum_{i=1}^{10} X_i$. לצורך חישוב התוחלת והשונות של הסכום ראשית נחשב את התוחלת והשונות של כל הטלה. בכל הטלה מתקבל ערך מבין הערכים 1 עד 6 בהסתברות שווה, לכן

$$\mathbb{E}\left(X_i\right) = \sum_{k=1}^{6} \frac{k}{6} = 3.5$$

בנוסף,

$$\mathbb{E}\left(X_i^2\right)=\sum_{k=1}^{6}\frac{k^2}{6}=\frac{91}{6}$$
$$\mathbb{V}\left(X_i\right)=\mathbb{E}\left(X_i^2\right)-\mathbb{E}^2\left(X_i\right)=\frac{35}{12}$$

מאחר וההטלות בלתי תלויות, התוחלת והשונות של סכום ההטלות תהיה הסכום של התוחלות והשונויות (בהתאמה) של כל אחת מההטלות:

$$\mathbb{E}\left(X\right)=\sum_{i=1}^{10}\mathbb{E}\left(X_i\right)=35$$
$$\mathbb{V}\left(X\right)=\sum_{i=1}^{10}\mathbb{V}\left(X_i\right)=\frac{175}{6}$$

נחסום את ההסתברות המבוקשת על ידי שני החסמים שראינו ונבחר את החסום ההדוק יותר, כלומר הקטן יותר מביניהם. היות ומדובר על סכום הטלות של קוביה והסכום תמיד יהיה חיובי, ניתן להשתמש בחסם מרקוב:

$$Pr\left(X\geq 50\right)\leq\frac{\mathbb{E}\left(X\right)}{50}=0.7$$

בכדי להשתמש בחסם צ'בישב, יש ראשית להביא את האי-שוויון המבוקש לצורה של חסם צ'בישב:

$$Pr\left(X\geq 50\right)=Pr\left(X-35\geq 50-35\right)=Pr\left(X-35\geq 15\right)$$

בחסם צ'בישב הסטייה מהתוחלת מופיעה בערכה המוחלט, לכן בכדי שניתן יהיה להשתמש בו יש לעבור מהמאורע $\{X-35\geq 15\}$ למאורע $\{|X-35|\geq 15\}$. הבעיה היא שהמאורע עם הערך המוחלט הוא איחוד של שני המאורעות הזרים $\{X-35\geq 15\}$ ו-$\{X-35\leq -15\}=\{X\leq 20\}$, לכן במעבר מהמאורע בלי ערך מוחלט למאורע עם ערך מוחלט, ההסתברות גדלה:

$$Pr\left(X-35\geq 15\right)\leq Pr\left(|X-35|\geq 15\right)\leq\frac{\mathbb{V}\left(X\right)}{15^2}=0.13$$

חסם צ'בישב קטן בהרבה מחסם מרקוב ולכן הוא חסם הדוק יותר להסתברות המבוקשת.

ניתן לשפר את התוצאה אף יותר אם נחשב עד כמה הגדלנו את ההסתברות כאשר חישבנו אותה עם חסם צ'בישב. נשים לב ש-X סימטרי ביחס לתוחלת: למשל, הסכום 60 מתקבל רק אם בכל הקוביות יצא 6 ולעומת זאת הסכום 10 מתקבל רק אם בכל הקוביות יצא 1. באופן כללי, כל תוצאה שרשומה על הקוביות ומובילה לתוצאה גדולה מ-50 מתקבלת בהסתברות שווה לתוצאה אחרת, בה כל אחת מהקוביות מוחלפת במשלימתה ל-7 (מחליפים בין "1" ל-"6", בין "2" ל-"5" ובין "3" ל-"4") ובה התוצאה קטנה מ-20. לכן במעבר מ-$\{X-35\geq 15\}$ ל-$\{|X-35|\geq 15\}$ הוספנו מאורע שהסתברותו שווה להסתברות של המאורע $\{X-35\geq 15\}$. לכן, ניתן להפוך את אי-השוויון לשוויון על ידי חלוקה ב-2 ונקבל:

$$Pr\left(X-35\geq 15\right)=\frac{1}{2}\cdot Pr\left(|X-35|\geq 15\right)\leq\frac{1}{2}\cdot\frac{\mathbb{V}\left(X\right)}{15^2}=0.065$$

וזהו החסם הטוב ביותר אותו ניתן להשיג מבין חסמי מרקוב וצ'בישב. $\square$

תרגיל 5.22. נגדיר משתנה מקרי חדש, $Y = \frac{X-a}{b-a}$. היות ו-X מקיים $a \le X \le b$ הרי ש-Y מקיים $0 \le Y \le 1$ ולאור ההוכחה הקודמת, $\mathbb{V}(Y) \le \frac{1}{4}$. יחד עם זאת, לפי לינאריות השונות מתקיים $\mathbb{V}(Y) = \mathbb{V}\left(\frac{X-a}{b-a}\right) = \frac{1}{(b-a)^2}\mathbb{V}(X)$ כלומר $\frac{1}{(b-a)^2}\mathbb{V}(X) \le \frac{1}{4}$ או $\mathbb{V}(X) \le \frac{(b-a)^2}{4}$, כנדרש.

זהו החסם הטוב ביותר שכן קיים משתנה מקרי שזוהי אכן השונות שלו. המשתנה המקרי המבוקש הוא המשתנה המקרי בו הפיזור הגדול ביותר, לכן הערכים היחידים שהוא יקבל יהיו ערכי הקצוות:

$$X = \begin{cases} a & \frac{1}{2} \\ b & \frac{1}{2} \end{cases}$$

מתקיים $\mathbb{E}(X) = \frac{a+b}{2}$ וכן $\mathbb{E}(X^2) = \frac{a^2+b^2}{2}$ ולכן

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X) = \frac{a^2+b^2}{2} - \left(\frac{a+b}{2}\right)^2 = \frac{a^2+b^2}{2} - \frac{a^2+2ab+b^2}{4}$$
$$= \frac{a^2-2ab+b^2}{4} = \frac{(b-a)^2}{4}$$

כנדרש. $\square$

תרגיל 5.23. 1. נחשב לפי ההגדרה:

$$M_X(t) = \mathbb{E}\left(e^{tX}\right) = \sum_{k \in supp(X)} e^{kt} Pr(X=k)$$
$$= \sum_{k=0}^{n} e^{kt} \cdot \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} \left(e^t p\right)^k (1-p)^{n-k}$$

את הסכום האחרון ניתן לחשב על ידי בינום ניוטון (נוסחא 0.10) ונקבל

$$M_X(t) = \left(e^t p + 1 - p\right)^n$$

2. נחשב לפי ההגדרה:

$$M_X(t) = \mathbb{E}\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tu} \cdot f_X(u)\, du$$
$$= \int_0^{\infty} e^{tu} \lambda e^{-\lambda u}\, du = \int_0^{\infty} \lambda e^{(t-\lambda)u}\, du$$

אינטגרל זה מתכנס רק אם $t - \lambda < 0$ לכן הפונקציה היוצרת מומנטים מוגדרת רק בתחום $(-\infty, \lambda)$. תחום זה כולל את הנקודה $t = 0$ ולכן אין בהגבלה זו בעיה ונוכל לחשב באמצעות הנוסחא שתתקבל

את המומנטים של משתנה מקרי זה. נמשיך את האינטגרל תחת ההנחה ש-$t < \lambda$ ונקבל אינטגרל של אסקפוננט, כלומר:

$$M_X\left(t\right) = \lambda \cdot \left.\frac{e^{(t-\lambda)u}}{t-\lambda}\right|_0^{\infty} = \frac{\lambda}{\lambda - t}$$

$\square$

תרגיל 5.24. יהי X משתנה מקרי עם פונקציה יוצרת מומנטים $M_X\left(t\right)$, יהיו $a, b \in \mathbb{R}$ ונגדיר משתנה מקרי $Y = aX + b$. נחשב את הפונקציה היוצרת מומנטים של Y ישירות מההגדרה:

$$M_Y\left(t\right) = \mathbb{E}\left(e^{tY}\right) = \mathbb{E}\left(e^{t(aX+b)}\right) = \mathbb{E}\left(e^{taX} \cdot e^{tb}\right)$$

התוחלת מחושבת על X ולכן הגודל e^{tb} הינו קבוע כפלי:

$$\mathbb{E}\left(e^{taX} \cdot e^{tb}\right) = e^{tb}\mathbb{E}\left(e^{(ta)X}\right) = e^{tb}M_X\left(at\right)$$

כנדרש. $\square$

תרגיל 5.25. בתרגיל 2.29 חושבו כל ההסתברויות הרלוונטיות ונותר רק לחשב את ההתפלגות של ניקוד הכדורסלן. למשל, $X = 0$ מתקיים כאשר הוא מחטיא פעמיים ולכן $Pr\left(X = 0\right) = 0.06$. באופן דומה,

$$Pr\left(X = k\right) = \begin{cases} 0.06 & k = 0 \\ \frac{1}{2} \cdot 0.8 \cdot 0.2 + \frac{1}{2} \cdot 0.3 \cdot 0.8 = 0.2 & k = 1 \\ \frac{1}{2} \cdot 0.8 \cdot 0.8 = 0.32 & k = 2 \\ \frac{1}{2} \cdot 0.7 \cdot 0.3 + \frac{1}{2} \cdot 0.2 \cdot 0.7 = 0.175 & k = 3 \\ \frac{1}{2} \cdot 0.7 \cdot 0.7 = 0.245 & k = 6 \end{cases}$$

לכן תוחלת הניקוד שלו תהא

$$\mathbb{E}\left(X\right) = 0 \cdot 0.06 + 1 \cdot 0.2 + \ldots + 6 \cdot 0.245 = 2.835$$

$\square$

תרגיל 5.26. לפי הפתרון של תרגיל 2.30 מספר המחליפים במופע, X, מתפלג עם פונקציית ההסתברות

$$P_X\left(k\right) = \begin{cases} 0.25 & k = 0 \\ 0.25 & k = 1 \\ 0.45 & k = 2 \\ 0.05 & k = 3 \end{cases}$$

תוחלת מספר המחליפים היא לפיכך

$$\mathbb{E}\left(X\right) = 0 \cdot 0.25 + 1 \cdot 0.25 + 2 \cdot 0.45 + 3 \cdot 0.05 = 1.3$$

תוחלת ריבוע מספר המחליפים הינה

$$\mathbb{E}\left(X^2\right) = 0^2 \cdot 0.25 + 1^2 \cdot 0.25 + 2^2 \cdot 0.45 + 3^2 \cdot 0.05 = 2.5$$

ולכן שונות מספר המחליפים הינה

$$\mathbb{V}\left(X\right) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2\left(X\right) = 2.5 - 1.3^2 = 0.81$$

□

תרגיל 5.27. נסמן את הערך השני ש-X מקבל ב-a ואת ההסתברות ב-$Pr\left(X = a\right) = p_a$. מתקיים:

$$0 = \mathbb{E}(X) = -3 \cdot (1 - p_a) + a \cdot p_a$$

וכן

$$4.5 = \mathbb{V}(X) = \mathbb{E}\left(X^2\right) = 9 \cdot (1 - p_a) + a^2 p_a$$

זו מערכת של שתי משוואות בשני נעלמים שפתרונה $a = 1.5, p_a = \frac{2}{3}$, כלומר המשתנה המקרי X הינו:

$$X = \begin{cases} \frac{3}{2} & \frac{2}{3} \\ -3 & \frac{1}{3} \end{cases}$$

□

תרגיל 5.28. ניתן לחשב את התוחלות והשונויות המבוקשות בשתי דרכים. בדרך הראשונה והמעט יותר ארוכה נשתמש בתרגיל **4.4** בו חושבו ההתפלגויות בצורה מפורשת ומתוך התפלגויות אלו נחשב את התוחלות והשונויות ישירות לפי ההגדרה. הדרך השניה היא שימוש בתכונות של התוחלת והשונות. נסמן ב-X_1 את תוצאת הקוביה הראשונה וב-X_2 את תוצאת הקוביה השנייה. סכום הקוביות הינו $X = X_1 + X_2$ והפרש הקוביות הינו $Y = X_1 - X_2$. לכן

$$\begin{aligned} \mathbb{E}\left(X\right) &= \mathbb{E}\left(X_1 + X_2\right) = \mathbb{E}\left(X_1\right) + \mathbb{E}\left(X_2\right) \overset{Prob\ 5.2}{=} 3.5 + 3.5 = 7 \\ \mathbb{E}\left(Y\right) &= \mathbb{E}\left(X_1 - X_2\right) = \mathbb{E}\left(X_1\right) - \mathbb{E}\left(X_2\right) = 0 \end{aligned}$$

לצורך חישוב השונות נחשב תחילה את השונות של הטלה בודדת. מתקיים:

$$\mathbb{E}\left(X_1^2\right) = \sum_{n=1}^{6} \frac{n^2}{6} = \frac{91}{6}$$

לכן

$$\mathbb{V}\left(X_1\right) = \mathbb{E}\left(X_1^2\right) - \mathbb{E}^2\left(X_1\right) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

מסימטריה בין ההטלות, זוהי גם השונות של X_2. ההטלות בלתי תלויות ולכן

$$\mathbb{V}(X) = \mathbb{V}(X_1 + X_2) = \mathbb{V}(X_1) + \mathbb{V}(X_2) = \frac{35}{6}$$
$$\mathbb{V}(Y) = \mathbb{V}(X_1 - X_2) = \mathbb{V}(X_1) + (-1)^2\,\mathbb{V}(X_2) = \mathbb{V}(X)$$

כך שלשני המשתנים המקריים אותה השונות. □

תרגיל 5.29. 1. בבירור פונקציית הצפיפות אי-שלילית. כמו כן, הפונקציה מנורמלת שכן

$$\int_0^1 2t\,dt = t^2\Big|_0^1 = 1$$

ולכן מתארת משתנה מקרי רציף.

2. לצורך חישוב התוחלת והשונות נחשב ישירות את האינטגרלים המתאימים:

$$\mathbb{E}(X) = \int_0^1 2t^2\,dt = 2\cdot\frac{t^3}{3}\Big|_0^1 = \frac{2}{3}$$
$$\mathbb{E}\left(X^2\right) = \int_0^1 2t^3\,dt = 2\cdot\frac{t^4}{4}\Big|_0^1 = \frac{1}{2}$$

ולכן

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X) = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

3. נשתמש בנוסחא לתוחלת של פונקציה של משתנה מקרי בשביל לפשט את החישוב ולחשב את התוחלת והשונות מבלי לחשב את ההתפלגות עצמה. לכן

$$\mathbb{E}(Y) = \mathbb{E}\left(\sqrt{X}\right) = \int_0^1 \sqrt{t}\cdot 2\cdot t\,dt = \int_0^1 2t^{1.5}\,dt = 2\cdot\frac{t^{2.5}}{2.5}\Big|_0^1 = \frac{4}{5}$$

וכן

$$\mathbb{E}\left(Y^2\right) = \mathbb{E}\left(\sqrt{X}^2\right) = \mathbb{E}(X) = \frac{2}{3}$$

כך שהשונות תהא

$$\mathbb{V}(Y) = \frac{2}{3} - \left(\frac{4}{5}\right)^2 = \frac{2}{75}$$

4. באותה צורה נחשב את התוחלת והשונות של Z על ידי שימוש בנוסחא לתוחלת של פונקציה:

$$\mathbb{E}(Z) = \mathbb{E}\left(\frac{1}{X}\right) = \int_0^1 \frac{1}{t} \cdot 2 \cdot t\, dt = 2$$

$$\mathbb{E}\left(Z^2\right) = \mathbb{E}\left(\frac{1}{X^2}\right) = \int_0^1 \frac{1}{t^2} \cdot 2 \cdot t\, dt = \int_0^1 \frac{2}{t}\, dt$$

אינטגרל זה מתבדר שכן הפונקציה הקדומה היא $\ln(t)$ ששואפת ל-$-\infty$ ב-$t = 0$ ולכן ל-$Z = \frac{1}{X}$ יש תוחלת אבל אין שונות.

□

תרגיל 5.30. נסמן ב-X את ציון הבחינה וב-X_i את ציון השאלה ה-i. מתקיים $X = \sum_{i=1}^{n} X_i$ ולכן $\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i)$ ונותר לחשב את ההתפלגות של X_i.

1. לפי נוסחת ההסתברות השלמה, הסיכוי שהסטודנט יענה את התשובה הנכונה ויקבל את מלוא הניקוד הינו

$$Pr\left(X_i = \frac{100}{n}\right) = p + (1-p) \cdot \frac{1}{4} = \frac{1}{4} + \frac{3}{4} \cdot p$$

לכן ההסתברות לתשובה שגויה הינה הסתברות המשלימה $Pr\left(X_i = -\frac{40}{n}\right) = \frac{3}{4} - \frac{3}{4} \cdot p$. נחשב את תוחלת הניקוד על השאלה ה-i:

$$\mathbb{E}(X_i) = \frac{100}{n} \cdot \left(\frac{1+3p}{4}\right) - \frac{40}{n}\left(\frac{3-3p}{4}\right) = \frac{-5+105p}{n}$$

כך שתוחלת הציון הכולל תהא $\mathbb{E}(X) = 105p - 5$.

2. לפי הקנס מהסעיף הקודם, כאשר הסטודנט לא יודע כלום ($p = 0$) הוא יקבל בתוחלת -5 נקודות, כלומר הקנס גדול מדי. בכדי שהציון יהיה 0 נקנוס את הסטודנט רק ב-$\frac{33.3}{n}$ נקודות על כל תשובה שגויה. ניתן לודא זאת על ידי שחזור הפתרון מהסעיף הקודם.

□

תרגיל 5.31. לפי תרגיל **4.7** המשתנה המקרי X מתפלג לפי $P_X(l) = \frac{\binom{l-1}{m-1}}{\binom{n}{m}}$. נחשב את התוחלת ישירות לפי ההגדרה:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{l=m}^{n} l \cdot P_X(l) = \sum_{l=m}^{n} l \cdot \frac{\binom{l-1}{m-1}}{\binom{n}{m}} = \frac{\sum_{l=m}^{n} \frac{l \cdot (l-1)!}{(l-m)!(m-1)!}}{\binom{n}{m}} \\
&= \frac{1}{\binom{n}{m}} \sum_{l=m}^{n} \frac{l!}{(l-m)!m!} \cdot m = \frac{m}{\binom{n}{m}} \sum_{l=m}^{n} \binom{l}{m} \overset{eq.\,4.19}{=} \frac{m}{\binom{n}{m}} \cdot \binom{n+1}{m+1} \\
&= m \cdot \frac{n+1}{m+1}
\end{aligned}$$

□

תרגיל 5.32. הטענה הזו כמעט אף פעם לא נכונה שכן התוחלת היא אופרטור לינארי (נסו למצוא משתנה מקרי עבורו הטענה נכונה!) ולכן ישנן דוגמאות נגדיות רבות. למשל, המשתנה המקרי X המוגדר על ידי

$$X = \begin{cases} 1 & \frac{1}{2} \\ 3 & \frac{1}{2} \end{cases}$$

מקיים $\mathbb{E}(X) = 2$ אולם ההופכי שלו

$$\frac{1}{X} = \begin{cases} 1 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} \end{cases}$$

מקיים $\mathbb{E}\left(\frac{1}{X}\right) = \frac{1+\frac{1}{3}}{2} = \frac{2}{3} \neq \frac{1}{2} = \frac{1}{\mathbb{E}(X)}$. □

תרגיל 5.33. מאחר ופונקציית הצפיפות נתונה והיא פשוטה, ניתן לחשב את ההסתברות המבוקשת ישירות:

$$Pr(X \geq 100) = \int_{100}^{\infty} \frac{2}{t^3}\, dt = \frac{1}{10,000}$$

נבנה את שני החסמים המוכרים ונבדוק איזה מהם הכי טוב ומהו טיב החסם ביחס לערך האמיתי. לצורך חישוב חסם מרקוב עלינו לחשב את התוחלת:

$$\mathbb{E}(X) = \int_{1}^{\infty} t \cdot \frac{2}{t^3}\, dt = 2$$

ולכן לפי חסם מרקוב (נוסחא 5.11) נקבל

$$Pr(X \geq 100) \leq \frac{\mathbb{E}(X)}{100} = \frac{1}{50}$$

לצורך חסם צ'בישב נחשב גם את השונות:

$$\mathbb{E}\left(X^2\right) = \int_{1}^{\infty} t^2 \cdot \frac{2}{t^3}\, dt = \int_{1}^{\infty} \frac{2}{t}\, dt$$

אינטגרל זה מתבדר ולכן למשתנה המקרי אין שונות ולא ניתן לחשב עבורו חסם צ'בישב. החסם הטוב ביותר אם כך הוא חסם מרקוב שערכו $\frac{1}{50}$ רחוק יחסית מההסתברות האמיתית. □

תרגיל 5.34. נגדיר $f(y) = \mathbb{E}\left((y-X)^2\right)$, תוחלת התשלום שבוב משלם כאשר הוא מנחש y. נשתמש בתכונות התוחלת בשביל לפשט את הפונקציה:

$$f(y) = \mathbb{E}\left(y^2 - 2yX + X^2\right) = y^2 - 2y\mathbb{E}(X) + \mathbb{E}\left(X^2\right)$$

כאשר y הוא קבוע לא אקראי הנקבע על ידי בוב ולכן יוצא מהתוחלות כקבוע. הנוסחא שהתקבלה היא פרבולה והמינימום שלה מתקבל בנקודה $y_{\max} = \frac{2\mathbb{E}(X)}{2} = \mathbb{E}(X)$ וערכו $= f(\mathbb{E}(X)) = \mathbb{E}\left((\mathbb{E}(X) - X)^2\right)$ $\mathbb{V}(X)$.

המסקנה היא שהתוחלת היא הגודל שמקטין למינימום את ההפסד הריבועי מניחוש שגוי. במקרה כזה, השונות היא בדיוק תוחלת ההפסד הטובה ביותר. כל ניחוש אחר יוביל לתוחלת הפסד יותר גדולה. □

תרגיל 5.35. יהי Xמשתנה מקרי אי-שלילי ונניח כי X בדיד. לכל $k \in \mathbb{N}$ מתקיים $Pr(X > k) = \sum_{n=k+1}^{\infty} Pr(X = n)$ ולכן

$$\sum_{k=0}^{\infty} Pr(X > k) = \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} Pr(X = n)$$

יהי $i \in supp(X)$. הביטוי $Pr(X = i)$ מופיע בסכימה רק כאשר $0 \leq k \leq i - 1$ שכן רק אז $n = k + 1 \leq i$. לפיכך, הסתברות זו מופיעה בסכום בדיוק i פעמים. לכן טור זה שווה לטור

$$\sum_{i=0}^{\infty} i \cdot Pr(X = i) = \mathbb{E}(X)$$

כנדרש.

נניח כי X רציף עם פונקציית צפיפות $f_X(t)$ ופונקציית התפלגות מצטברת $F_X(t)$. נשתמש בכך שלכל $t > 0$ מתקיים $t = \int_0^t 1\, dx$ ונחשב את האינטגרל המתאר את התוחלת:

$$\mathbb{E}(X) = \int_0^{\infty} t f_X(t)\, dt = \int_0^{\infty} f_X(t) \left(\int_0^t 1\, dx \right) dt = \int_0^{\infty} \int_0^t f_X(t)\, dx dt$$

לכל $t > 0$ נגדיר את הפונקציה $g_t(x) = \begin{cases} 1 & x > t \\ 0 & x < t \end{cases}$ ונציב באינטגרל האחרון:

$$\mathbb{E}(X) = \int_0^{\infty} \int_0^t f_X(t)\, dx dt = \int_0^{\infty} \int_0^{\infty} g_t(x) f_X(t)\, dx dt$$

נחליף את סדר האינטגרציה ונבצע את האינטגרל לפי t קודם. עבור x קבוע, $g_t(x) f_X(t) = \begin{cases} f_X(t) & t < x \\ 0 & t > x \end{cases}$ ולכן:

$$\mathbb{E}(X) = \int_0^{\infty} \int_0^{\infty} g_t(x) f_X(t)\, dt dx = \int_0^{\infty} \int_x^{\infty} f_X(t)\, dt dx = \int_0^{\infty} Pr(X > x)\, dx$$

כנדרש. $\square$

6 משפחות של משתנים מקריים

דריה מבקרת בפרדס. אחרי שקטפה 10 פירות הדר, ביניהם 3 תפוזים, דריה רוצה לאכול תפוז והיא מוציאה באקראי פירות מהסל עם החזרה, עד אשר תוציא תפוז. יותם, אחיה, נשאר באותו הזמן בבית ומשחק בקוביה. הוא מטיל את הקוביה שוב ושוב וסופר את מספר ההטלות עד הפעם הראשונה שהקוביה נופלת על 4. על פניו, שני הילדים מבצעים פעולות שונות אבל ניתן למצוא ביניהן קווי דימיון רבים. שניהם מבצעים ניסוי שוב ושוב. בכל חזרה על הניסוי, לכל אחד מהם יש תוצאה רצויה (תפוז או 4) עם סיכוי קבוע בכל חזרה ללא תלות בחזרות הקודמות. שניהם סופרים את מספר החזרות עד אשר התוצאה הרצויה מתקבלת. לכן, למעשה, שניהם מבצעים את אותו הניסוי עם הבדל מספרי – הסיכוי שדריה תצליח להוציא תפוז בכל חזרה על הניסוי הינו $\frac{3}{10}$ והסיכוי שיותם יקבל את התוצאה 4 הוא $\frac{1}{6}$. לכן, מספר הניסויים עד התוצאה המבוקשת בשני המקרים הוא אותו משתנה מקרי מבחינת הנוסחאות המתארות אותו, כאשר ההבדל היחיד הוא רק ההבדל המספרי.

באופן כללי, בעיות רבות המופיעות בחיים (ובמבחן) דומות זו לזו. במקום לפתור כל בעיה בנפרד, נכליל את הבעיות הנפוצות ונפתור אותן פרמטרית כך שכל שיוותר זה לזהות את הבעיה המתאימה ולהציב את הערכים הנכונים. מרגע שהבעיה זוהתה, ניתן יהיה להשתמש בכל הנוסחאות עבור פונקציית ההסתברות, התוחלת, השונות ושאר המדדים והתכונות שיפותחו להלן. ניתן, כמובן, לדלג על כל החומר המופיע בפרק זה ולפתור כל בעיה בנפרד, אך יהיה בכך בזבוז גדול מאוד של זמן ומאמץ.

נסיים בהוספת סימון חדש שמשמעותו "מתפלג כמו". נניח שמטילים קוביה פעמיים ו- X הוא תוצאת ההטלה הראשונה ו-Y הוא תוצאת ההטלה השניה. בשתי ההטלות התוצאות יכולות להיות שונות לכן לא נכון להגיד ש-$X = Y$. יחד עם זאת, לשתי ההטלות אותה ההתפלגות, שכן כל ערך אפשרי מתקבל בהטלה הראשונה ובהטלה השניה בהסתברויות שוות. לכן, נאמר ש"X מתפלג כמו Y" ונסמן $X \sim Y$. באופן טבעי, אם X מתפלג כמו Y אז שני המשתנים המקריים מקבלים בדיוק את אותם ערכים באותן הסתברויות ולכן כל החישובים שביצענו בפרק 5 יהיו זהים לשניהם, כלומר $\mathbb{E}(X) = \mathbb{E}(Y)$, $\mathbb{V}(X) = \mathbb{V}(Y)$ וכן הלאה.

6.1 משפחות של משתנים מקריים בדידים

בסעיף זה נתאר את המשפחות הנפוצות של המשתנים המקריים הבדידים. מאחורי כל משתנה מקרי כזה עומד "סיפור". כאשר פותרים שאלות, מרגע שזוהתה התאמה בין הסיפור בשאלה לסיפור המתאר את המשתנה המקרי, ניתן להשתמש בנוסחאות המתאימות עבור ההתפלגות, התוחלת, השונות ועוד ולמעשה השאלה כמעט נפתרה לחלוטין ברגע הזיהוי. לצורך זיהוי המשתנה המקרי כדאי להתחיל מחישוב התומך והשוואתו לתומכים של המשפחות המוכרות – בדיקה כזו יכולה לפסול חלק מהמשפחות ולהשאיר מספר מועט של משפחות של משתנים מקריים להשוואה ובדיקה. כמו כן, כדאי לזכור שאמנם מדובר על משפחות שמופיעות במקרים רבים, לא תמיד המשתנה המקרי מתאים לאחת המשפחות ולעיתים לא תהיה ברירה אלא לחשב את פונקציית ההסתברות ישירות כפי שביצענו בפרק 4.1.

6.1.1 משתנה מקרי אחיד בדיד

ההתפלגות הפשוטה ביותר מתקבלת כאשר התוצאות נבחרות באקראי בהסתברות שווה, כמו בקוביה. משתנה מקרי שמקבל את כל הערכים האפשריים בהסתברות שווה מכונה משתנה מקרי אחיד. אנחנו נתמקד במשתנים מקריים אחידים מסוג מיוחד, בהם התומך כולל קבוצה של ערכים שלמים עוקבים בתחום מסויים. אחרת, אם הקבוצה לא כוללת חוקיות כלשהי, לא ניתן לחשב באופן כללי את התוחלת ואת שאר המאפיינים של ההתפלגות.

הגדרה 6.1 (משתנה מקרי אחיד). משתנה מקרי המקבל את כל הערכים בתחום $\{a, a+1, \ldots, b\}$ כאשר $a, b \in \mathbb{Z}$ מכונה משתנה מקרי אחיד ומסומן $U\{a, \ldots, b\}$. אם $X \sim U\{a, \ldots, b\}$ אזי כל אחד מ-$b-a+1$ ערכים מתקבל בהסתברות שווה ולכן לכל $k \in \{a, \ldots, b\}$ מתקיים

$$P_X(k) = \frac{1}{b-a+1} \tag{6.1}$$

והתומך הוא כמובן $supp(X) = \{a, \ldots, b\}$.

תרגיל 6.1. יהי $X \sim U\{a, \ldots, b\}$. הוכיחו את הנוסחאות הבאות:

$$\mathbb{E}(X) = \frac{a+b}{2} \tag{6.2}$$

$$\mathbb{V}(X) = \frac{(b-a+1)^2-1}{12} \tag{6.3}$$

$$M_X(t) = \frac{e^{at}-e^{(b+1)t}}{(b-a+1)(1-e^t)} \tag{6.4}$$

רמז: כדאי להתחיל מ-$Y \sim U\{0, \ldots, N\}$.

6.1.2 משתנה מקרי בינומי

נסתכל על ניסוי בו ישנן שתי תוצאות אפשריות ולמען הפשטות נכנה אחת מהן בתור "הצלחה" ואת השניה בתור "כשלון". נניח כי ההסתברות להצלחה היא מספר קבוע $p \in [0,1]$ ואינה תלויה בתוצאות של חזרות קודמות על הניסוי, אם היו. ניסוי כזה מכונה ניסוי ברנולי. למשל – מטילים קוביה ומחשיבים את התוצאה "5" בתור הצלחה או זורקים לסל (עם הסתברות קבועה לקלוע) ומחשיבים קליעה בתור הצלחה. כאשר חוזרים על הניסוי $n \in \mathbb{N}$ פעמים וסופרים בכמה מתוך חזרות אלו הניסוי הצליח, מקבלים משתנה מקרי המכונה משתנה מקרי בינומי.

הגדרה 6.2 (משתנה מקרי בינומי). חוזרים על ניסוי $n \in \mathbb{N}$ פעמים כאשר בכל חזרה יש הסתברות קבועה $p \in [0,1]$ להצלחה. נגדיר X– מספר ההצלחות שהתקבלו. אזי X מתפלג בינומית עם הפרמטרים n ו-p ונסמן

$$X \sim Bin(n,p)$$

נחשב את פונקציית ההסתברות של המשתנה המקרי הבינומי. נתחיל מחישוב התומך. המשתנה המקרי סופר הצלחות לכן מספר ההצלחות חייב להיות מספר טבעי, כאשר המספר המינימלי הוא 0 הצלחות והמספר המרבי הוא n הצלחות, לכן $supp(X) = \{0, \ldots, n\}$.

יהי $k\in\{0,\ldots,n\}$. המאורע $\{X=k\}$ מתרחש כאשר יש בדיוק k הצלחות ולכן בהכרח $n-k$ כשלונות. הסיכוי שיהיו בדיוק k הצלחות ומיד אחריהם בדיוק $n-k$ כשלונות הינו $p^k\left(1-p\right)^{n-k}$. אולם מאורע זה יכול להתרחש גם בסידורים אחרים, למשל $n-k$ כשלונות ואז k הצלחות או הצלחה אחת, כשלון אחד, ואז $k-1$ הצלחות ואז יתר הכשלונות וכן הלאה. מספר האפשרויות הכולל הינו כמספר האפשרויות לבחור את k הניסויים המוצלחים מתוך n הניסויים שהתקיימו, כאשר הבחירה נעשית ללא החזרה (ניסוי לא יכול להיבחר פעמיים) וללא סדר (משנים רק מספרי הניסויים המוצלחים, לא באיזה סדר הם נבחרו), כלומר יש $\binom{n}{k}$ אפשרויות ולסיכום:

$$P_X\left(k\right)=\binom{n}{k}\cdot p^k\left(1-p\right)^{n-k} \tag{6.5}$$

ניתן לראות שפונקציה זו היא אכן פונקציית הסתברות באמצעות בינום ניוטון (נוסחא 0.10) המקנה למשתנה המקרי את שמו:

$$\sum_{k=0}^{n}Pr\left(X=k\right)=\sum_{k=0}^{n}\binom{n}{k}p^k\left(1-p\right)^{n-k}=\left(p+\left(1-p\right)\right)^n=1$$

שלושה מרכיבים מרכזיים מתארים את הסיפור המתאים למשתנה מקרי בינומי:

1. הסתברות קבועה להצלחה.
2. מספר חזרות קבוע.
3. המשתנה המקרי סופר את מספר ההצלחות.

למשל - מטילים קוביה 10 פעמים וסופרים את מספר הפעמים שהתקבלה התוצאה 5 (מתפלג $Bin\left(10,\frac{1}{6}\right)$), שולפים 12 קלפים עם החזרה וסופרים כמה פעמים התקבל מלך ($Bin\left(12,\frac{1}{13}\right)$) או שחקן כדורסל הקולע בסיכוי 0.3 זורק לסל פעמיים וסופר כמה פעמים הוא קלע ($Bin\left(2,0.3\right)$). ברגע שזוהה המשתנה המקרי ניתן מיידית להשתמש בנוסחא 6.5 בכדי לחשב את ההסתברות של כל תוצאה אפשרית. בנוסף, ניתן לחשב את התוחלת והשונות של מספר ההצלחות הצפוי.

***משפט* 6.1.** *יהי $X\sim Bin\left(n,p\right)$. אזי*

$$\mathbb{E}\left(X\right)=np \tag{6.6}$$
$$\mathbb{V}\left(X\right)=np\left(1-p\right) \tag{6.7}$$

הוכחה. ניתן לחשב את התוחלת והשונות ישירות לפי הנוסחא תוך שימוש בזהויות של מקדמים בינומיים או על ידי פירוק לאינדיקטורים. נגוון ונחשב את התוחלת על ידי שימוש בנוסחא ואת השונות על ידי פירוק לאינדיקטורים:

$$\mathbb{E}\left(X\right)=\sum_{k=0}^{n}k\cdot\binom{n}{k}p^k\left(1-p\right)^{n-k}\overset{eq.0.13}{=}np$$

לצורך חישוב השונות נסמן ב-$\mathbf{1}_i$ את האינדיקטור המקבל את הערך 1 כאשר הייתה הצלחה בניסוי ה-i. כאשר $X=k$ ישנם בדיוק k ניסויים מוצלחים לכן גם k אינדיקטורים עם הערך 1 ויתר האינדיקטורים עם הערך 0, כך שמתקבל $X=\sum_{i=1}^{n}\mathbf{1}_i$. בנוסף, לפי הגדרת המשתנה המקרי הניסויים בלתי תלויים כך שגם האינדיקטורים בלתי תלויים וניתן להשתמש בחיבוריות השונות (משפט 5.7):

$$\mathbb{V}(X) = \mathbb{V}\left(\sum_{i=1}^{n} \mathbf{1}_i\right) = \sum_{i=1}^{n} \mathbb{V}(\mathbf{1}_i)$$
$$(eq.\,5.17) = \sum_{i=1}^{n} p(1-p) = np(1-p)$$

□

נעיר שאת הפונקציה היוצרת מומנטים של משתנה מקרי בינומי כבר מצאנו בתרגיל 5.23:

$$M_X(t) = \left(e^t p + 1 - p\right)^n \tag{6.8}$$

בעיה 6.1. מטילים מטבע הוגן n פעמים. נסמן ב-X את מספר הפעמים בהן תוצאת ההטלה השתנתה ביחס לתוצאת ההטלה הקודמת וב-Y את מספר הפעמים בהן תוצאת ההטלה השתנתה מעץ לפלי. מהן התוחלות של X ו-Y?

הוכחה. נזהה את המשתנה המקרי X. בכל הטלה, הסיכוי שהמטבע יפול על הצד ההפוך מהצד של ההטלה הקודמת הוא חצי והוא לא תלוי במספר ההטלה או בתוצאות הקודמות. בנוסף, יש בסך הכל $n-1$ הטלות בהן התוצאה יכולה להתחלף (היא לא יכולה להתחלף בהטלה הראשונה כי אין הטלה קודמת) והמשתנה המקרי X סופר את מספר ההחלפות, לכן זהו בדיוק הסיפור של משתנה מקרי בינומי: $X \sim Bin\left(n-1, \frac{1}{2}\right)$. התוחלת תתקבל לפי נוסחא 6.6:

$$\mathbb{E}(X) = \frac{n-1}{2}$$

נזהה את המשתנה המקרי Y– גם כאן יש $n-1$ הטלות בהן התוצאה יכולה להתחלף והמשתנה המקרי סופר את מספר ההחלפות. הפעם ההסתברות להחלפה מעץ לפלי הינה $p = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$ שכן על מנת שתהיה בהטלה השניה החלפה מעץ לפלי, יש לדרוש שבהטלה הראשונה המטבע יפול על עץ ובהטלה השניה יפול על פלי. יחד עם זאת, הסתברות זו אינה קבועה בין הסיבובים ותלויה בתוצאות של הסיבוב הקודם. למשל, אם בהטלה השניה הייתה החלפה מעץ לפלי, בהטלה השלישית לא יכולה להיות החלפה נוספת כזו כי זה ידרוש מההטלה השניה להיות עץ. מתקבל ש-Y אינו משתנה מקרי בינומי! במקרה כזה ניתן לנסות לחשב את ההתפלגות של Y מפורשות אך היות וכל מה שמעניין אותנו הוא התוחלת, נשתמש בתכונות התוחלת בכדי לחשב את התוחלת של Y תוך היעזרות בקשר בין Y ל-X.

נסמן ב-Z את מספר ההחלפות בין פלי לעץ. בבירור, כל החלפה היא החלפה מעץ לפלי או מפלי לעץ, ולכן $X = Y + Z$. בנוסף, המשתנים המקריים Y ו-Z שווי התפלגות ולכן $\mathbb{E}(Y) = \mathbb{E}(Z)$. נשתמש בחיבוריות התוחלת (נוסחא 5.6):

$$\mathbb{E}(X) = \mathbb{E}(Y+Z) = \mathbb{E}(Y) + \mathbb{E}(Z) = 2\mathbb{E}(Y)$$

ולכן $\mathbb{E}(Y) = \frac{n-1}{4}$. □

לפעמים המשתנה המקרי הבינומי לא מוגדר מפורשות בבעיה או לחילופין מוגדר משתנה מקרי אחר, אבל כל עוד הסיפור המתאים לבעיה מזכיר משתנה מקרי בינומי – מספר חזרות קבוע וסיכוי קבוע להצלחה בכל חזרה, ניתן להגדיר משתנה מקרי בינומי הסופר הצלחות ולהשתמש בו לצורך חישוב ההתפלגות המבוקשת.

בעיה 6.2. שיכור מתחיל לצעוד מנקודה 0, על ציר המספרים הממשיים. השיכור צועד ימינה (לכיוון החיובי של הציר) ושמאלה (לכיוון השלילי של הציר) בהסתברות שווה וצועד 100 צעדים בגודל 1.

1. כיצד מתפלג מיקום השיכור?
2. מה ההסתברות שבסופו של דבר ימצא את עצמו השיכור בנקודה $X = 10$? ובנקודה $X = 11$?
3. מצאו חסם הדוק ככל הניתן להסתברות שהשיכור יסיים את צעידתו במרחק של יותר מ-90 צעדים מנקודת ההתחלה.

הוכחה. 1. נסמן ב-X את מיקום השיכור ונתחיל מחישוב התומך. ברמה העקרונית, במצב הקיצוני ביותר השיכור יצעד 100 צעדים לכיוון אחד, לכן הוא יכול להיות בכל מקום שלם בין $X = -100$ ל-$X = 100$ כך שהתומך אינו מתאים לתומך של משתנה מקרי בינומי. יחד עם זאת, יש כאן את מרבית הסיפור הבינומי: מספר הניסויים (צעדים) קבוע מראש ויש הסתברות קבועה להצלחה (נניח, צעד ימינה). הבעיה היא שהמשתנה המקרי שהגדרנו לא סופר את ההצלחות אלא את מיקום השיכור, הקשור איכשהו למספר ההצלחות. נגדיר אם כך משתנה חדש שיספור הצלחות, כלומר נסמן ב-Y את מספר הצעדים ימינה. לאור הדיון הקודם, $Y \sim Bin\left(100, \frac{1}{2}\right)$.
נותר למצוא את הקשר בין מספר הצעדים ימינה למיקום השיכור. אם השיכור צעד Y צעדים ימינה אזי בהכרח הוא צעד $100 - Y$ צעדים שמאלה ולכן מיקומו הסופי יהיה

$$X = Y - (100 - Y) = 2Y - 100$$

נשתמש בקשר זה בכדי למצוא את ההתפלגות של X:

$$\begin{aligned} Pr\left(X = k\right) &= Pr\left(2Y - 100 = k\right) = Pr\left(Y = \frac{100 + k}{2}\right) \\ &= \binom{100}{\frac{k+100}{2}} \cdot \left(\frac{1}{2}\right)^{\left(\frac{k+100}{2}\right)} \cdot \left(\frac{1}{2}\right)^{100 - \frac{k+100}{2}} = \binom{100}{\frac{k+100}{2}} \cdot \left(\frac{1}{2}\right)^{100} \end{aligned}$$

נשים לב שהתומך של X איננו $\{-100, -99, \ldots, 100\}$ שכן על מנת ש-k יהיה בתומך על $\frac{k+100}{2}$ להיות בתומך של Y, כלומר עליו להיות מספר שלם. לכן $k + 100$ חייב להיות מספר זוגי ולכן התומך האמיתי של X הינו $\{-100, -98, \ldots, 98, 100\}$.

2. ניתן להציב בנוסחא שמצאנו ולקבל $Pr\left(X = 10\right) = Pr(Y = 55) = \binom{100}{55} \cdot \left(\frac{1}{2}\right)^{100}$. באופן כללי, השיכור יכול להגיע רק למקומות זוגיים שכן מספר הצעדים ימינה ועוד מספר הצעדים שמאלה צריך להיות 100 ולכן או ששניהם זוגיים או ששניהם אי-זוגיים ובכל מקרה – הפרשם, מיקומו הסופי, זוגי. לאור זאת (ולאור הצבה ישירה בנוסחא, אם טרם השתכנעתם) $Pr(X = 11) = 0$.

3. המאורע המבוקש הינו $\{X > 90\} \cup \{X < -90\}$ ומאחר ומיקומים אי זוגיים בלתי אפשריים, המאורע שקול ל-$\{X \geq 92\} \cup \{X \leq -92\}$. מאחר והמשתנה המקרי X אינו ממשפחה ידועה וכן אינו משתנה מקרי חיובי, יותר נוח יהיה לעבוד עם Y. לאור הקשר $Y = \frac{X+100}{2}$ נקבל

$$Pr\left(X \geq 92 \cup X \leq -92\right) = Pr\left(Y \geq 96 \cup Y \leq 4\right)$$

נחסום את Y על ידי כל אחד מהחסמים המוכרים ונמצא את החסם הטוב ביותר.

- מרקוב:

המאורעות $Y \geq 96$ ו-$Y \leq 4$ שווי הסתברות מאחר ושניהם מתארים לכל היותר **4** צעדים מסוג מסויים (שמאלה וימינה, בהתאמה) והסיכוי לצעד ימינה וצעד שמאלה זהה. לכן:

$$Pr\left(Y \geq 96 \cup Y \leq 4\right) = 2Pr\left(Y \geq 96\right) \leq 2 \cdot \frac{\mathbb{E}\left(Y\right)}{96} = 2 \cdot \frac{100 \cdot \frac{1}{2}}{96} = 1.04$$

כלומר קיבלנו לפי מרקוב חסם חסר תועלת, כי ממילא ידענו שההסתברות הזו קטנה מ-**1**.

- צ'בישב:

חסם צ'בישב דן בסיכוי להתרחק מהתוחלת, לכן נפחית את התוחלת $\mathbb{E}\left(Y\right) = 100 \cdot \frac{1}{2} = 50$ משני האגפים:

$$\begin{aligned} Pr\left(Y \geq 96 \cup Y \leq 4\right) &= Pr\left(\left\{Y - \mathbb{E}\left(Y\right) \geq 46\right\} \cup \left\{Y - \mathbb{E}\left(Y\right) \leq -46\right\}\right) \\ &= Pr\left(\left|Y - \mathbb{E}\left(Y\right)\right| \geq 46\right) \leq \frac{\mathbb{V}\left(Y\right)}{46^2} = 0.012 \end{aligned}$$

וחסם זה כבר יותר מוצלח בייחוד לאור העובדה שאנחנו מצפים להסתברות נמוכה למאורע זה, הדורש סטייה כל כך רחוק מהממוצע וצעידה של כמעט כל הצעדים בכיוון אחד.

- צ'רנוף:

נשתמש בחסם צ'רנוף ונציב את הפונקציה היוצרת מומנטים של Y:

$$\begin{aligned} Pr\left(Y \geq 96\right) &\leq M_Y\left(t\right)e^{-96t} = \left(e^t \cdot \frac{1}{2} + 1 - \frac{1}{2}\right)^{100} e^{-96t} \\ &= \frac{1}{2^{100}}\left(e^t + 1\right)^{100} e^{-96t} \end{aligned}$$

נחקור את הפונקציה $g\left(t\right) = \left(e^t + 1\right)^{100} e^{-96t}$ למציאת מינימום בתחום $t \in [0, \infty)$. בנקודת המינימום הנגזרת מתאפסת:

$$0 = g'\left(t\right) = 100e^{-95t}\left(e^t + 1\right)^{99} - 96e^{-96t}\left(e^t + 1\right)^{100}$$

המשוואה האחרונה שקולה לכך ש-$100e^t = 96\left(e^t + 1\right)$ ופתרונה $t = \ln 24$ היא נקודת המינימום. נציב ערך זה לקבלת החסם:

$$Pr\left(Y \geq 96 \cup Y \leq 4\right) = 2Pr\left(Y \geq 96\right) \leq 2 \cdot \frac{1}{2^{100}} g\left(\ln 24\right) = 3.1 \cdot 10^{-23}$$

כך שההסתברות האמיתית מאוד מאוד קטנה ולכל צורך פרקטי ניתן להגיד שההסתברות הזו היא אפס – המאורע בו השיכור יתרחק יותר מ-**90** צעדים מנקודת המוצא לא מתרחש. מאחר ויש מעט מאוד ערכים המתאימים לתוצאה $Y \geq 96$, ניתן להציב את כל הערכים האפשריים בנוסחא

להתפלגות בינומית ולקבל את ההסתברות האמיתית של המאורע:

$$\begin{aligned} Pr\left(Y \geq 96 \cup Y \leq 4\right) &= 2 \cdot Pr\left(Y \geq 96\right) \\ &= 2 \cdot \sum_{k=96}^{100} \binom{100}{k} \cdot \left(\frac{1}{2}\right)^{100} \\ &= 6.4 \cdot 10^{-24} \end{aligned}$$

ואכן אי-שוויון צ'רנוף סיפק חסם הדוק מאוד להסתברות האמיתית, בוודאי ביחס לחסמים שהושגו באמצעות מרקוב וצ'בישב.

□

תרגיל 6.2. בהמשך לבעיה 6.2:

1. מה ההסתברות שהצעד האחרון הוא צעד ימינה?
2. לאחר 100 צעדים השיכור הגיע לנקודה $X = 10$. מה ההסתברות שהצעד האחרון שלו היה צעד ימינה?
3. חזרו על שני הסעיפים בהנחה שהשיכור נע ימינה בהסתברות p כלשהי.

תרגיל 6.3. מחשב משתמש במילים באורך 8 ביטים. כאשר משדרים ביטים בקו התקשורת יש סיכוי של 0.1 לשגיאה בביט בודד שמשודר, כך ש-0 ששודר יפוענח בתור 1 או להפך. בכדי לשפר את איכות הקו, משדרים כל ביט 3 פעמים ברצף והפענוח של כל שלשת ביטים מתבצע על ידי החלטת הרוב. למשל: אם רוצים לשדר 0, משדרים בקו "000" והמקלט מפענח את התשדורת כ-0 אם השלשה שהגיעה אליו מורכבת מלפחות שני אפסים. כפי שנראה בתרגיל זה, טכניקה של שידור שלושה ביטים במקום אחד עוזרת להקטין את ההסתברות לשגיאה אך המחיר מגיע בכך שכל ביט משודר 3 פעמים (ולכן מילה כעת מורכבת מ-24 ביטים) וקצב התקשורת יורד.

1. חשבו את ההסתברות שמילת מחשב תפוענח נכון.
2. מצאו את התפלגות מספר המילים שפוענחו באופן שגוי בקובץ המכיל 1000 מילים.

תרגיל 6.4. לפניכם חמישה מטבעות הוגנים. מטילים את כולם ומוציאים מהמשחק את המטבעות שנפלו על עץ. מטילים שנית את המטבעות שלא הוצאו מהמשחק ומוציאים שוב את המטבעות שנפלו על עץ. ממשיכים שוב ושוב עד שכל המטבעות הוצאו מהמשחק.

1. מה ההסתברות שלאחר הסיבוב השני נותרו 2 מטבעות?
2. מהי ההסתברות שלאחר הסיבוב החמישי נותרו 3 מטבעות?
3. מהי ההסתברות שלאחר 4 סיבובים לא נותרו מטבעות?

תרגיל 6.5. בצרפת ישנם שני בתי קזינו בהם משחקים ברולטה. בקזינו של ואל-ז'אן ישנה ההסתברות זהה להרוויח ולהפסיד בכל סיבוב ברולטה ואילו בקזינו של טנרדייה ההסתברות להפסד היא 0.6 בכל סיבוב. מהמר נכנס באקראי לאחד מבתי הקזינו ומהמר על פרנק אחד בכל סיבוב (אם זוכה – מקבל פרנק נוסף ואם מפסיד – מאבד את הפרנק עליו הימר).

1. כיצד מתפלג מספר הזכיות של המהמר בכל אחד מבתי הקזינו לאחר 10 סיבובי משחק?
2. מהי תוחלת הרווח של המהמר בכל אחד מבתי הקזינו אחרי 10 סיבובים?
3. לאחר 10 סיבובים המהמר הרוויח 8 פרנקים. מה ההסתברות שהוא בקזינו של ואל-ז'אן?

6.1.3 משתנה מקרי גאומטרי

שאלה נוספת אותה ניתן לשאול בהקשר של חזרה שיטתית על ניסוי ברנולי בלתי תלויים היא מתי אירעה ההצלחה הראשונה. למשל - כדורסלן קולע לסל בהסתברות קבועה 0.3, כיצד מתפלג מספר הזריקות עד הסל הראשון? מה הסיכוי שקוביה שמוטלת שוב ושוב תוטל 10 פעמים לפחות לפני שיצא 6 לראשונה? שיכור צועד באקראי ימינה ושמאלה, מה הסיכוי שיבצע לכל היותר 4 צעדים שמאלה לפני הצעד הראשון ימינה? המשותף לכל השאלות הללו הוא שישנו ניסוי ברנולי עם סיכוי קבוע להצלחה, הניסוי מבוצע שוב ושוב ואנחנו מתעניינים בהופעה הראשונה של התוצאה המכונה "הצלחה" בניסוי. המשתנה המקרי הסופר את מספר החזרות על הניסוי עד אשר הופיעה ההצלחה הראשונה מכונה משתנה מקרי גאומטרי.

הגדרה 6.3 (משתנה מקרי גאומטרי)**.** חוזרים על ניסוי שוב ושוב כאשר בכל חזרה יש הסתברות קבועה $p \in [0,1]$ להצלחה. נגדיר X- מספר החזרות על הניסוי עד אשר התקבלה ההצלחה הראשונה, כולל הניסוי בו היא התקבלה. אזי X מתפלג גאומטרית עם הפרמטר p ונסמן

$$X \sim Geom\left(p\right)$$

נחשב את פונקציית ההסתברות של המשתנה המקרי הגאומטרי. נתחיל מחישוב התומך. המשתנה המקרי סופר מספר חזרות ולכן חייב להיות מספר טבעי, כאשר המספר המינימלי הוא 1 (מי שלא מנסה לא מצליח). הפעם, אין מספר מקסימלי בתומך שכן ברמת העקרון ניתן להיכשל ברצף גם מיליון ומיליארד פעמים לפני ההצלחה הראשונה. לכן $supp\left(X\right) = \mathbb{N}$.

יהי $k \in \mathbb{N}$. המאורע $\{X = k\}$ מתרחש כאשר $k-1$ הניסויים הראשונים הסתיימו בכשלון והניסוי האחרון הסתיים בהצלחה. כל הניסויים בלתי תלויים ולכן ההסתברות של המאורע המבוקש הינה:

$$P_X\left(k\right) = \left(1-p\right)^{k-1} p \qquad (6.9)$$

ניתן לראות שפונקציה זו היא אכן פונקציית הסתברות באמצעות חישוב הטור הגאומטרי המתאים (נוסחא 0.15) המקנה למשתנה המקרי את שמו:

$$\sum_{k=1}^{\infty} Pr\left(X=k\right) = \sum_{k=1}^{\infty} \left(1-p\right)^{k-1} p \overset{eq.\,0.16}{=} 1$$

שלושה מרכיבים מרכזיים מתארים את הסיפור המתאים למשתנה מקרי גאומטרי:

1. הסתברות קבועה להצלחה.
2. מספר חזרות לא ידוע מראש (אינסופי).
3. המשתנה המקרי סופר את מספר החזרות עד ההצלחה הראשונה (כולל).

בעיה 6.3. במשחק הרולטה ישנם 37 מספרים (המספרים 0-36) כאשר המספר 0 צבוע בירוק, המספרים הזוגיים האחרים צבועים בשחור והמספרים האי-זוגיים באדום. מהמר משחק ברולטה ובכל סיבוב מהמר שהרולטה תעצור על מספר אדום. אם אכן כך יקרה - יכפיל את סכום ההימור ואם לא - יפסיד אותו.

1. כיצד מתפלג מספר הסיבובים עד הזכייה הראשונה של המהמר?
2. מה הסיכוי שהזכייה הראשונה תתרחש אחרי יותר מ-5 סיבובים?
3. המהמר הפסיד חמישה סיבובים רצופים. מה הסיכוי שיזכה בסיבוב הבא?

4. המהמר מחליט לנקוט באסטרטגיה הבאה: בסיבוב הראשון יהמר על דולר אחד ויפרוש מיד אם הוא יזכה. אם יפסיד, בסיבוב השני הוא יהמר על שני דולרים ויפרוש מיד אם יזכה. אם יפסיד, בסיבוב השלישי יהמר על ארבעה דולרים ויפרוש מיד אם יזכה וכן הלאה. מהי תוחלת הרווח של המהמר הנוקט באסטרטגיה זו?

הוכחה. ישנם 37 מספרים מתוכם 18 מספרים אדומים, לכן הסיכוי לזכות בכל סיבוב הינו $p = \frac{18}{37} = 0.486$.

1. כאמור, הסיכוי לזכות בכל סיבוב ברולטה הינו $p = \frac{18}{37}$ וסיכוי זה קבוע ולא תלוי במספר הסיבובים או בתוצאות הקודמות. לכן מספר החזרות עד הזכייה הראשונה מתפלג לפי $X \sim Geom\left(\frac{18}{37}\right)$.
2. בסימוני הסעיף הקודם, ההסתברות המבוקשת הינה ההסתברות של המאורע $\{X > 5\}$. ניתן לחשב הסתברות זו בשלוש דרכים. דרך אחת היא על ידי המעבר למשלים וסכימת כל האפשרויות המתאימות:

$$\begin{aligned}Pr\left(X>5\right) &= 1 - Pr\left(X \le 5\right) = 1 - \sum_{k=1}^{5} Pr\left(X=k\right)\\ &= 1 - \left(\left(1-\frac{18}{37}\right)^0 \cdot \frac{18}{37} + \ldots + \left(1-\frac{18}{37}\right)^5 \cdot \frac{18}{37}\right) =\\ (eq.\,0.14) &= 1 - \frac{18}{37} \cdot \frac{1-\left(1-\frac{18}{37}\right)^{4+1}}{1-\left(1-\frac{18}{37}\right)} = \left(1-\frac{18}{37}\right)^5 = \left(\frac{19}{37}\right)^5 = 0.036\end{aligned}$$

הדרך השניה היא להשתמש בנוסחא של טור גאומטרי וחישוב כל האפשרויות הגדולות מ-5:

$$\begin{aligned}Pr\left(X>5\right) &= \sum_{k=6}^{\infty} Pr\left(X=k\right) = \sum_{k=6}^{\infty}\left(1-\frac{18}{37}\right)^{k-1} \cdot \frac{18}{37}\\ &= \frac{18}{37} \cdot \sum_{k=0}^{\infty}\left(1-\frac{18}{37}\right)^5\left(1-\frac{18}{37}\right)^k \overset{eq.\,0.15}{=} \frac{18}{37} \cdot \left(1-\frac{18}{37}\right)^5 \cdot \frac{1}{1-\left(1-\frac{18}{37}\right)}\\ &= \left(1-\frac{18}{37}\right)^5 = 0.036\end{aligned}$$

הדרך השלישית מתבססת על הבנת המאורע המבוקש. משמעות המאורע $\{X > 5\}$ הוא שההצלחה הראשונה הגיעה לאחר יותר מ-5 ניסיונות וזה קורה רק כאשר חמשת הניסיונות הראשונים נגמרו בהפסד, מאורע שהסתברותו $\left(1-\frac{18}{37}\right)^5$.

3. המהמר הפסיד חמישה סיבובים לכן ידוע שהמאורע $\{X > 5\}$ התרחש. הסיבוב הבא הוא אם כך הסיבוב השישי ולכן המאורע המבוקש הוא המאורע המותנה $X = 6|X > 5$. נחשב הסתברות זו על ידי הנוסחא להסתברות מותנית (נוסחא 2.9):

$$Pr\left(X=6|X>5\right) = \frac{Pr\left(X=6 \cap X>5\right)}{Pr\left(X>5\right)}$$

החיתוך של $X = 6$ עם $X > 5$ הינו $X = 6$ (שכן המאורע הראשון מוכל באחרון). נציב את התוצאה

מהסעיף הקודם ונקבל:

$$Pr\left(X=6|X>5\right)=\frac{Pr\left(X=6\right)}{Pr\left(X>5\right)}=\frac{\left(1-\frac{18}{37}\right)^{5}\cdot\frac{18}{37}}{\left(1-\frac{18}{37}\right)^{5}}=\frac{18}{37}$$

תוצאה זו מאוד הגיונית, שהרי הסיכוי לזכיה בסיבוב בודד הינה $\frac{18}{37}$ ובפרט הסיכוי לזכות בסיבוב הבא, יהא מספרו אשר יהיה, הוא $\frac{18}{37}$. למרות שתוצאה זו הגיונית, היא מאוד לא אינטואיטיבית. מרבית האנשים מצפים שאחרי רצף ארוך של מספרים אי-זוגיים הסיכוי למספר זוגי יעלה שכן התוצאות צריכות להיות "מאוזנות" וצריכה להיות "התאמה" בין ההסתברויות של התוצאות האפשריות לבין התוצאות בפועל. הם מצפים ש"הסטטיסטיקה תתיישר" כך שאחרי רצף של אי-זוגיים תגדל השכיחות של הזוגיים ויגיעו רצפים של מספרים זוגיים כדי "לפצות" על הרצף ולתקן את השכיחות של הזוגיים לשכיחות התאורטית - $\frac{18}{37}$. כשל זה של האינטואיציה מכונה "כשל המהמר" ונובע מאי הבנה נכונה של הסתברויות וייחוס מוטעה של תלות לאירועים בלתי תלויים.

4. נסמן ב-Y את הרווח של המהמר. לצורך חישוב התוחלת של Y נחשב ראשית את הקשר בין X, מספר הסיבוב בו זכה המהמר לראשונה, לבין הרווח שלו Y. אם המהמר זכה בסיבוב ה-X לראשונה, אזי הוא הרוויח 2^{X-1} דולרים מהימור זה אבל בכל סיבוב אחר $j<X$ הוא הפסיד 2^{j-1} ולכן הרווח שלו יהיה:

$$\begin{aligned}Y\left(X\right)&=2^{X-1}-\sum_{j=1}^{X-1}2^{j-1}=2^{X}-\sum_{j=0}^{X-2}2^{j}\\(eq.\,0.14)&=2^{X-1}-\frac{1-2^{X-1}}{1-2}=1\end{aligned}$$

כלומר, ללא תלות בסיבוב בו המהמר מנצח לראשונה ופורש, הרווח שלו יהיה 1, כך שגם תוחלת הרווח תהיה 1. אסטרטגיה זו מבטיחה שהמהמר תמיד יצא מהקזינו ברווח של דולר אחד, גם אם סיכוי הזכיה לרעתו (כמו במקרה זה, $p<\frac{1}{2}$) והיא מתוארת לעיתים כשיטה המבטיחה לנצח את הקזינו. טענה זו נכונה תאורטית אך לא ישימה פרקטית שכן במרבית בתי הקזינו יש סכום הימור מינימלי ומקסימלי ואילו סכום ההימור הדרוש לנקיטת שיטה זו גדל מעריכית ובמהרה חורג מגבולות המותר או מתקציב המהמר. לשם הדגמה, נניח כי המהמר מגיע לקזינו עם תקציב של $2^{n}-1$ דולרים. לאור החישוב לעיל, תקציב זה מספיק ל-n הימורים עוקבים. אם המהמר ינצח באחד מ-n ההימורים הראשונים, הרווח שלו יהא 1. אחרת, הוא יתרושש וההפסד הכולל שלו יהיה כל כספו. תוחלת הרווח של המהמר יהא:

$$\begin{aligned}1\cdot Pr\left(X\leq n\right)-\left(2^{n}-1\right)\cdot Pr\left(X>n\right)&=1-\left(1-p\right)^{n}-\left(2^{n}-1\right)\cdot\left(1-p\right)^{n}\\&=1-2^{n}\left(1-p\right)^{n}\end{aligned}$$

וכל עוד סיכוי הזכיה לרעת המהמר, מספר זה שלילי לכל n.

□

בבעיה 6.3 הוצגו שני מקרים פרטיים של תכונות הנכונות בהתפלגות גאומטרית: נוסחת הזנב וחוסר הזיכרון. בסעיף 2 חושבה ההסתברות של המאורע $\{X>5\}$. חישוב הסתברויות מהצורה הזו עבור משתנים מקריים בדידים היא לרוב משימה קשה שכן החישוב מחייב סכימה של מספר רב של הסתברויות המהוות טור שלא תמיד

יש עבורו נוסחא סגורה או פשוטה. בהתפלגות גאומטרית, לעומת זאת, הסכום המתואר הוא סכום המתאים לטור גאומטרי ולכן החישוב אפשרי.

***משפט* 6.2** (נוסחת הזנב של התפלגות גאומטרית)*. יהי $X \sim Geom\left(p\right)$ ויהי $n \in \mathbb{N}$. אזי*

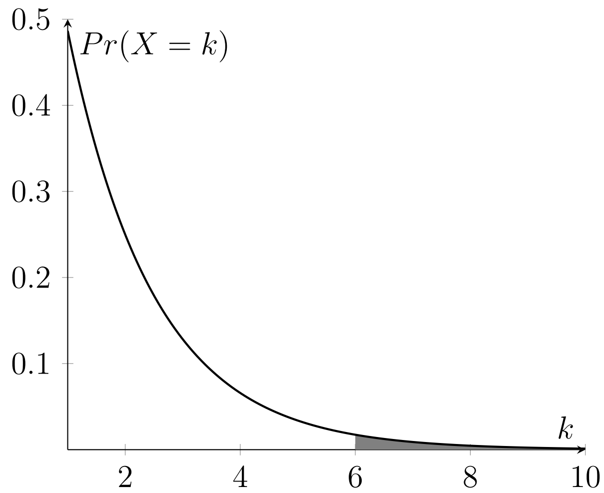
$$Pr\left(X > n\right) = \left(1 - p\right)^{n} \tag{6.10}$$

ולכן

$$Pr\left(X < n\right) = 1 - Pr\left(X \geq n\right) = 1 - Pr\left(X > n - 1\right) = 1 - \left(1 - p\right)^{n-1} \tag{6.11}$$

הוכחה. ניתן לחזור על כל אחת מההוכחות המוצגות בסעיף 2 של בעיה 6.3 כהוכחה כללית, תוך החלפת הערכים המספריים בפרמטרים. $\square$

נוסחה זו מכונה נוסחת הזנב עבור התפלגות גאומטרית שכן היא מתארת את ההסתברות של זנב ההתפלגות – התחום המתאים לערכים גדולים מערך מסויים. השרטוט הבא, בו צויירה פונקציית הסתברות של משתנה מקרי גאומטרי (לכל $k \in \mathbb{R}$ ולא רק לטבעיים, על מנת שהגרף יהיה רציף ו"יפה") ממחיש את משמעות השם.



בסעיף 3 של בעיה 6.3 ראינו כי בהינתן שהמהמר לא זכה בחמשת הסיבובים הראשונים, הסיכוי שלו לזכות כעת, בסיבוב השישי, זהה לסיכוי הזכייה של הסיבוב הראשון. באופן כללי, התפלגות מספר הסיבובים עד הזכייה הראשונה בהינתן שהוא כבר הפסיד חמש פעמים היא שוב התפלגות גאומטרית, כאילו הוא כלל לא שיחק. תכונה זו מכונה "חוסר זיכרון" ומתארת את העובדה שהעבר לא משפיע על העתיד ולרולטה אין "זיכרון" של אירועי העבר שישפיע על התפלגות הזכיות בהמשך. לכן, הסיכוי לזכות בסיבוב השמיני אם כבר שיחקנו חמישה סיבובים (כלומר, בעוד 3 סיבובים) זהה לסיכוי לזכות בסיבוב השלישי כאילו רק עתה התחלנו את המשחק.

***משפט* 6.3** (חוסר זיכרון של משתנה מקרי גאומטרי)*. יהי $X \sim Geom\left(p\right)$. המשתנה X חסר זיכרון, כלומר בהינתן שההצלחה הראשונה לא הגיעה אחרי a סיבובים, הסיכוי להצלחה אחרי b סיבובים נוספים שווה לסיכוי*

להצלחה אחרי b סיבובים מנקודת ההתחלה:

$$Pr\left(X=a+b|X>a\right)=Pr\left(X=b\right) \tag{6.12}$$

ובניסוח שקול:

$$Pr\left(X>a+b|X>a\right)=Pr\left(X>b\right) \tag{6.13}$$

הוכחה. יהי $X\sim Geom\left(p\right)$ ויהיו $a,b\in\mathbb{N}$. בדומה לפתרון סעיף 3 של בעיה 6.3 :

$$\begin{aligned}Pr\left(X=a+b|X>a\right)&=\frac{Pr\left(X=a+b\cap X>a\right)}{Pr\left(X>a\right)}=\frac{Pr\left(X=a+b\right)}{Pr\left(X>a\right)}\\(eq.\,6.10)&=\frac{\left(1-p\right)^{a+b-1}\cdot p}{\left(1-p\right)^{a}}=\left(1-p\right)^{b-1}\cdot p=Pr\left(X=b\right)\end{aligned}$$

ובכך נוסחא 6.12 הוכחה. נוסחא 6.13 מוכחת בצורה דומה מאוד:

$$\begin{aligned}Pr\left(X>a+b|X>a\right)&=\frac{Pr\left(X>a+b\cap X>a\right)}{Pr\left(X>a\right)}=\frac{Pr\left(X>a+b\right)}{Pr\left(X>a\right)}\\(eq.\,6.10)&=\frac{\left(1-p\right)^{a+b}}{\left(1-p\right)^{a}}=\left(1-p\right)^{b}=Pr\left(X>b\right)\end{aligned}$$

□

משתנה מקרי גאומטרי הוא המשתנה המקרי היחיד הבדיד הניחן בתכונת חוסר הזיכרון. למשל, אם מטילים מטבע 10 פעמים וסופרים את מספר העצים שהתקבלו, הסיכוי שיהיו 5 עצים אם ידוע שיש לפחות 4 עצים שונה מהסיכוי שיהיה עץ אחד (ודאו זאת!). לעומת זאת, אם הטלתם את הקוביה עשרים פעמים ועדיין לא יצא 6, נחתם לרגע והמשכתם את מלאכת ההטלות, התפלגות מספר ההטלות הנוספות עד שייצא 6 לראשונה היא גאומטרית וזהה להתפלגות מספר ההטלות החל מההטלה הראשונה.

תרגיל 6.6. במבחן אמריקאי 5 שאלות כאשר בכל שאלה יש 4 תשובות אפשריות אך רק אחת נכונה. יש לענות נכון על לפחות 3 שאלות בכדי לעבור את המבחן.

1. בכיתה 50 סטודנטים שלא למדו למבחן והם מנחשים באקראי ובאופן בלתי תלוי את התשובות לכל השאלות. כיצד מתפלג מספר הסטודנטים שעוברים את המבחן?
2. סטודנט ניגש למבחן מבלי ללמוד ומנחש באקראי את כל התשובות. כיצד מתפלג מספר הפעמים בהם עליו להיבחן עד אשר יעבור את המבחן?

תרגיל 6.7. מטילים מטבע הוגן שוב ושוב. כיצד מתפלג מספר ההטלות עד אשר יופיעו שני הצדדים של המטבע לפחות פעם אחת?

תרגיל 6.8. מטילים בכל פעם שני מטבעות לא הוגנים עד אשר אחד המטבעות נופל על עץ בזמן שהשני נופל על פלי. הסיכוי של המטבע הראשון ליפול על עץ הינו p והסיכוי של המטבע השני ליפול על עץ הינו q.

1. כיצד מתפלג מספר ההטלות עד סוף הניסוי?
2. מה הסיכוי שבסוף הניסוי המטבע הראשון הוא זה שנפל על עץ?

משפט 6.4. *יהי* $X\sim Geom\left(p\right)$. *אזי*

$$\mathbb{E}\left(X\right)=\frac{1}{p} \tag{6.14}$$

$$\mathbb{V}\left(X\right)=\frac{1-p}{p^{2}} \tag{6.15}$$

הוכחה. נחשב את התוחלת מפורשות לפי ההגדרה:

$$\mathbb{E}\left(X\right)=\sum_{k=1}^{\infty}k\cdot Pr\left(X=k\right)=\sum_{k=1}^{\infty}k\cdot\left(1-p\right)^{k-1}p\overset{eq.\,0.17}{=}\frac{1}{p}$$

חישוב השונות אפשרי גם בדרך זו אך הטור המתקבל מסובך יותר וה"טריק" בו יש לנקוט מסובך יותר. במקום זאת, נחשב את השונות דרך הפונקציה היוצרת מומנטים המחושבת בתרגיל 6.9. לפי נוסחא 5.22 המומנט השני הינו

$$\begin{aligned}\mathbb{E}\left(X^{2}\right) &= \frac{d^{2}}{dt^{2}}M_{X}\left(t\right)\Big|_{t=0}=\frac{d^{2}}{dt^{2}}\left(\frac{p\cdot e^{t}}{1-\left(1-p\right)\cdot e^{t}}\right)\Big|_{t=0}\\ &= \frac{d}{dt}\left(\frac{p\cdot e^{t}}{\left(1-\left(1-p\right)\cdot e^{t}\right)^{2}}\right)\Big|_{t=0}=\frac{pe^{t}\left(\left(1-(1-p)e^{t}\right)^{2}+2\left(1-(1-p)e^{t}\right)(1-p)e^{t}\right)}{\left(1-\left(1-p\right)e^{t}\right)^{4}}\Bigg|_{t=0}\\ &= \frac{2-p}{p^{2}}\end{aligned}$$

כך שהשונות תהא

$$\mathbb{V}\left(X\right)=\mathbb{E}\left(X^{2}\right)-\mathbb{E}^{2}\left(X\right)=\frac{2-p}{p^{2}}-\frac{1}{p^{2}}=\frac{1-p}{p^{2}}$$

□

במקרה של המהמר ברולטה (בעיה 6.3) התוחלת תהא $\mathbb{E}\left(X\right)=\frac{37}{18}=2.05$, כך שבממוצע המהמר ינצח ברולטה בכל סיבוב שני מה שיעודד אותו להמשיך ולשחק עוד ועוד במשחק המוטה במקצת לטובת הקזינו, שירוויח בטווח הארוך בעוד המהמר יתרושש.

תרגיל 6.9. הוכיחו שהפונקציה היוצרת מומנטים של משתנה מקרי גאומטרי $X\sim Geom\left(p\right)$ הינה

$$M_{X}\left(t\right)=\frac{p\cdot e^{t}}{1-\left(1-p\right)\cdot e^{t}} \tag{6.16}$$

תרגיל 6.10. מטילים קוביה שוב ושוב ונגדיר X – מספר ההטלות עד אשר כל הפאות התקבלו פעם אחת. מהי התוחלת והשונות של X?

6.1.4 משתנה מקרי בינומי שלילי

בפרקטיקה מספר החזרות עד ההצלחה הראשונה הוא המשתנה המקרי הנפוץ יותר, אבל אין שום דבר מיוחד דווקא בהצלחה הראשונה. לעיתים מה שמשנה הוא מספר החזרות עד ההצלחה השנייה או העשירית. המשתנה המקרי הסופר את מספר החזרות על ניסויי ברנולי בלתי תלויים עד ההצלחה ה-n נקרא משתנה מקרי בינומי שלילי. משתנה מקרי זה הינו הכללה של משתנה מקרי גאומטרי וניתן לקבל את כל הנוסחאות וכל התוצאות של משתנה מקרי גאומטרי על ידי הצבת $n=1$.

הגדרה 6.4 (משתנה מקרי בינומי שלילי). חוזרים על ניסוי שוב ושוב כאשר בכל חזרה יש הסתברות קבועה $p\in[0,1]$ להצלחה. נגדיר X– מספר החזרות על הניסוי עד אשר התקבלה ההצלחה ה-$n\in\mathbb{N}$, כולל הניסוי בו היא התקבלה. אזי X מתפלג בינומית שלילית עם פרמטרים n ו-p ונסמן

$$X\sim NB\left(n,p\right)$$

נחשב את פונקציית ההסתברות של המשתנה המקרי הבינומי השלילי. נתחיל מחישוב התומך. המשתנה המקרי סופר מספר חזרות ולכן חייב להיות מספר טבעי, כאשר המספר המינימלי הוא n כי חייבים לנסות לפחות n פעמים כדי להצליח n פעמים. כמו במשתנה מקרי גאומטרי, אין חסם עליון למספר הניסיונות האפשרי ולכן $supp\left(X\right)=\left\{n,n+1,\ldots\right\}$.

יהי $k\in supp\left(X\right)$. המאורע $\left\{X=k\right\}$ מתרחש כאשר מתוך k הניסויים היו $k-n$ ניסויים כושלים ו-n מוצלחים, כשההגבלה היחידה היא שהניסוי האחרון, ניסוי מספר k, הוא הניסוי בו אירעה ההצלחה ה-n ולכן התרחשה בו הצלחה. כמו בהתפלגות בינומית, עלינו לבחור מבין $k-1$ הניסויים הראשונים אילו ניסויים יהיו הניסויים המוצלחים וצריכים להיות $n-1$ כאלו. מספר אפשרויות הבחירה הינו $\binom{k-1}{n-1}$ וההסתברות הכוללת הינה

$$P_X\left(k\right)=\binom{k-1}{n-1}\left(1-p\right)^{k-n}p^n \qquad (6.17)$$

ניתן לראות שפונקציה זו היא אכן פונקציית הסתברות באמצעות חישוב הטור המתאים:

$$\begin{aligned}\sum_{k=n}^{\infty}Pr\left(X=k\right)&=\sum_{k=n}^{\infty}\binom{k-1}{n-1}\left(1-p\right)^{k-n}p^n\overset{eq.\,0.6}{=}\sum_{k=n}^{\infty}\binom{k-1}{k-n}\left(1-p\right)^{k-n}p^n\\ \left(i=k-n\right)&=\sum_{i=0}^{\infty}\binom{i+n-1}{i}\left(1-p\right)^{i}p^n\end{aligned}$$

נחשב את המקדם הבינומי המופיע בנוסחא האחרונה:

$$\binom{i+n-1}{i}=\frac{\left(i+n-1\right)!}{i!\left(n-1\right)!}=\frac{\left(i+n-1\right)\cdot\left(i+n-2\right)\cdot\ldots\cdot n}{i!}$$

נוציא $\left(-1\right)$ מכל אחד מהגורמים המופיעים במונה. יש i כאלו ולכן בסך הכל הגורם שיצא החוצה יהיה $\left(-1\right)^i$:

$$=\left(-1\right)^i\frac{\left(-n-\left(i-1\right)\right)\cdot\left(-n-\left(i-2\right)\right)\cdot\ldots\cdot\left(-n\right)}{i!}=\left(-1\right)^i\frac{\left(-n\right)\cdot\left(-n-1\right)\cdot\ldots\cdot\left(-n-\left(i-1\right)\right)}{i!}$$

אמנם הגדרנו מקדמים בינומים רק למספרים חיוביים, אבל מבחינה פורמלית הביטוי שהתקבל זהה לביטוי המופיע בנוסחא **0.3** ולכן ניתן לרשום אותה בתור $\left(-1\right)^i\binom{-n}{i}$. רישום זה הוא שמקנה להתפלגות את השם "בינומי

שלילי". נציב בטור המתאר את ההסתברויות ונשלים את החישוב:

$$\sum_{i=0}^{\infty}\binom{i+n-1}{i}(1-p)^i p^n=\sum_{i=0}^{\infty}(-1)^i\binom{-n}{i}(1-p)^i p^n=p^n\sum_{i=0}^{\infty}\binom{-n}{i}(p-1)^i$$

הטור האחרון מזכיר מאוד את נוסחת הבינום של ניוטון 0.10. תוך סטייה מאוד גדולה מהחומר ובאמצעות שימוש בטורי חזקות ניתן להוכיח שנוסחת הבינום של ניוטון נכונה גם במקרה הזה ומתקיים

$$p^n\sum_{i=0}^{\infty}\binom{-n}{i}(p-1)^i\cdot 1^{-n-i}=p^n(p-1+1)^{-n}=1$$

כנדרש.

ארבעה מרכיבים מרכזיים מתארים את הסיפור המתאים למשתנה מקרי בינומי שלילי:

1. הסתברות קבועה להצלחה.
2. מספר חזרות לא ידוע מראש (אינסופי).
3. מספר הצלחות נדרש וקבוע מראש, n.
4. המשתנה המקרי סופר את מספר החזרות עד ההצלחה ה-n (כולל).

חישוב סכום ההסתברויות היה מסובך מספיק בכדי להבין שלא כדאי לחשב את התוחלת, השונות והפונקציה היוצרת מומנטים של המשתנה המקרי הבינומי שלילי בשיטה הישירה. תחת זאת, נשים לב שניתן לתאר את המשתנה המקרי הבינומי השלילי כסכום של משתנים מקריים גאומטריים, שהרי ספירה עד ההצלחה ה-n ניתנת לחלוקה למספר ספירות עוקבות: ספירה עד ההצלחה הראשונה, ספירה מההצלחה הראשונה לשניה, ספירה מההצלחה השניה לשלישית וכן הלאה. הפירוק הפורמלי מופיע בהוכחת המשפט הבא.

***משפט* 6.5.** *יהי $X\sim NB(n,p)$. אזי*

$$\mathbb{E}(X)=\frac{n}{p} \tag{6.18}$$

$$\mathbb{V}(X)=n\cdot\frac{1-p}{p^2} \tag{6.19}$$

$$M_X(t)=\left(\frac{p\cdot e^t}{1-(1-p)\cdot e^t}\right)^n \tag{6.20}$$

הוכחה. יהי $X\sim NB(n,p)$ משתנה מקרי הסופר את מספר החזרות עד ההצלחה ה-n. ניתן לרשום משתנה מקרי זה כסכום של משתנים מקריים הסופרים את מספר החזרות בין כל שתי הצלחות סמוכות. ביתר פירוט, נגדיר X_1 - מספר החזרות מתחילת הניסוי ועד ההצלחה הראשונה, X_2- מספר החזרות מההצלחה הראשונה ועד ההצלחה השניה, X_3- מספר החזרות מההצלחה השניה ועד השלישית וכן הלאה עד ל-X_n - מספר החזרות מההצלחה ה-$n-1$ ועד ההצלחה ה-n. כל אחד ממשתנים אלו סופר את מספר החזרות עד ההצלחה "הראשונה" הבאה בתור, ולכן לכל $i\in\{1,\ldots,n\}$ מתקיים $X_i\sim Geom(p)$ וכן $X=\sum_{i=1}^{n}X_i$.

נשתמש בחיבוריות התוחלת (נוסחא 5.7) ונקבל:

$$\mathbb{E}(X)=\mathbb{E}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathbb{E}(X_i)\overset{eq.\,6.14}{=}\sum_{i=1}^{n}\frac{1}{p}=\frac{n}{p}$$

מספר החזרות בין הצלחה להצלחה בלתי תלוי במספר החזרות בין הצלחות אחרות ואינו תלוי במספר הסידורי של ההטלה (כל זאת מתכונת חוסר הזיכרון של המשתנה המקרי הגאומטרי) ולכן ניתן להשתמש גם בחיבוריות השונות (נוסחא 5.16):

$$\mathbb{V}\left(X\right)=\mathbb{V}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathbb{V}\left(X_i\right)\overset{eq.6.15}{=}\sum_{i=1}^{n}\frac{1-p}{p^2}=n\cdot\frac{1-p}{p^2}$$

לצורך חישוב הפונקציה היוצרת מומנטים נשתמש בשאלה 7.38 לפיה הפונקציה היוצרת מומנטים של סכום משתנים מקריים שווה למכפלת הפונקציות היוצרות מומנטים. לכן:

$$M_X\left(t\right)=\prod_{i=1}^{n}M_{X_i}\left(t\right)=\prod_{i=1}^{n}\frac{p\cdot e^t}{1-(1-p)\cdot e^t}=\left(\frac{p\cdot e^t}{1-(1-p)\cdot e^t}\right)^n$$

כנדרש. $\square$

תרגיל 6.11. מטילים מטבע שוב ושוב עד הפעם הרביעית בה יצא עץ. כיצד מתפלג מספר ההטלות אם ידוע שהעץ הראשון התקבל בהטלה החמישית?

6.1.5 משתנה מקרי היפרגאומטרי

תהליך חשוב החוזר על עצמו בתחומים רבים ובעיקר בהסקה סטטיסטית הוא תהליך הדגימה. מתוך אוכלוסיה קבועה נבחרת קבוצה (ללא החזרה וללא חשיבות לסדר) ומעוניינים ללמוד על האוכלוסיה הכללית מתוך הקבוצה הנבחרת. למשל, מתוך אוכלוסיית ישראל נבחרים באקראי 500 אנשים ונשאלים לגבי הצבעתם בבחירות הקרובות על מנת לקבל תחזית לגבי התוצאה הצפויה הבחירות. באופן דומה, מכל 100 מוצרים שמפעל מייצר נבחרים 2 באקראי ונבדקים בדיקה מעמיקה כדי לבדוק שפס היצור פועל כפי שצריך ואין תקלות. תחום ההסקה הסטטיסטית עוסק בשאלות מהסוג הזה, המבקש ללמוד על האוכלוסיה הכללית מתוך מדגם שלה. אנחנו נתמקד בתהליך ההסתברותי - מתוך אוכלוסיה ידועה עם מספר מסויים של פריטים "מיוחדים" (מצביעי ליכוד, למשל, בדוגמא הראשונה, או פריטים תקולים בדוגמא השניה) דוגמים באקראי מספר פריטים ובוחנים כמה מהם "מיוחדים". המשתנה המקרי הסופר את כמות המיוחדים במדגם נקרא משתנה מקרי היפרגאומטרי.

הגדרה 6.5 (משתנה מקרי היפרגאומטרי). נתונה אוכלוסיה בת N פריטים מתוכם D פריטים "מיוחדים". בוחרים באקראי וללא החזרה n פריטים מתוך האוכלוסיה. נגדיר X - מספר הפריטים המיוחדים שנבחרו. אזי X מתפלג היפרגאומטרית עם הפרמטרים N,D,n ונסמן

$$X\sim HG\left(N,D,n\right)$$

נחשב את פונקציית ההסתברות של המשתנה המקרי ההיפרגאומטרי. נתחיל מחישוב התומך. המשתנה המקרי סופר את מספר הפריטים המיוחדים שנבחרו ולכן חייב להיות מספר טבעי, כאשר ברמת העקרון המספר המינימלי הוא 0 פריטים והמספר המקסימלי הוא D פריטים. יחד עם זאת, אם $n<D$ לא ניתן יהיה לבחור יותר מ-n פריטים מיוחדים ולכן המקסימום של התומך יהיה $\min\left\{n,D\right\}$. באופן דומה, אם מספר הפריטים המיוחדים גדול ומספר הפריטים הנבחרים גדול, הרי שאין ברירה אלא לבחור לפחות חלק מהפריטים המיוחדים. כך למשל, אם יש בקערה 8 תפוחים ו-2 אגסים ובוחרים באקראי 5 פירות, חייבים לבחור לפחות 3 תפוחים. באופן כללי, המספר המקסימלי של "לא מיוחדים" שניתן לבחור הינו $N-D$ ולכן אם $n>N-D$ אזי חייבים לבחור לפחות $n-(N-D)$ פריטים מיוחדים. לכן המינימום של התומך יהיה $\max\left\{0,n-(N-D)\right\}$. לסיכום, התומך של משתנה מקרי היפרגאומטרי הינו $supp\left(X\right)=\left\{\max\left\{0,n-(N-D)\right\},\ldots,\min\left\{n,D\right\}\right\}$.

לעיתים, לשם הפשטות, נתייחס לתומך בתור $\{0,\ldots,n\}$ שכן הצבה של ערכים בלתי אפשריים בפונקציית ההסתברות (נוסחא 6.21) תניב את הערך 0.

יהי $k\in supp\left(X\right)$. מרחב המדגם כולל את כל הבחירות של n פריטים מתוך האוכלוסיה הכוללת, כאשר אין החזרה ואין חשיבות לסדר (כל שחשוב הוא מספר המיוחדים ולא הסדר בו יצאו). לכן מרחב המדגם הינו $\binom{N}{n}$. המאורע $\{X=k\}$ מתרחש כאשר מתוך D המיוחדים נבחרו k פריטים ומתוך $N-D$ הלא מיוחדים נבחרו $n-k$ פריטים. ההסתברות של מאורע זה הינה

$$P_X\left(k\right)=\frac{\binom{D}{k}\cdot\binom{N-D}{n-k}}{\binom{N}{n}} \qquad (6.21)$$

ניתן לראות שפונקציה זו אכן מתארת פונקציית הסתברות כפי שמוראה בתרגיל 0.3.

בעיה 6.4. בארגז תפוחים 100 תפוחים, 20 מתוכם רקובים. בוחרים באקראי 10 תפוחים. מה ההסתברות שנבחרו 4 תפוחים רקובים כאשר

1. הבחירה נעשית עם החזרה.
2. הבחירה נעשית ללא החזרה.

הוכחה. נשים לב שלא מוגדר משתנה מקרי בשאלה ואכן, ברמת העקרון, אין צורך במשתנה מקרי לפתרון השאלה. ניתן לפתור את שני הסעיפים בכלים רגילים של קומבינטוריקה. אולם, אם נגדיר נכון את המשתנה המקרי המסתתר בבעיה, נקבל מיד ובלי מאמץ נוסף את פונקציית ההסתברות שלו וכל שיוותר הוא להציב בנוסחא ולקבל ערך. הגודל המעניין בשני הסעיפים הוא מספר התפוחים הרקובים שנבחרו ולכן נגדיר X- מספר התפוחים הרקובים שנבחרו. ההסתברות המבוקשת בשני הסעיפים הינה $Pr\left(X=4\right)$.

1. הבחירה נעשית עם החזרה ולכן ניתן לתאר את הניסוי בשלבים. בשלב הראשון, הסיכוי לבחור תפוח רקוב הוא $\frac{20}{100}=\frac{1}{5}$. בשלב השני, ארגז התפוחים מכיל בדיוק את אותם התפוחים כמו בסיבוב הראשון ולכן הסיכוי לקחת תפוח רקוב הוא שוב $\frac{1}{5}$. דבר זה נכון גם לכל שאר בחירות התפוחים. לפיכך, הניסוי המתואר הוא ניסוי חזרתי, בו נבחרים 10 תפוחים עם סיכוי $\frac{1}{5}$ לבחור תפוח רקוב בכל סיבוב והמשתנה המקרי סופר את מספר התפוחים הרקובים. זהו בדיוק הסיפור המתאר התפלגות בינומית כך ש-$X\sim Bin\left(10,\frac{1}{5}\right)$ ולכן

$$Pr\left(X=4\right)=\binom{10}{4}\left(\frac{1}{5}\right)^4\left(\frac{4}{5}\right)^6=0.088$$

2. הבחירה נעשית ללא החזרה והתהליך המתואר הוא דגימה - מתוך 100 פריטים נדגמים 10 ונבדק כמה מתוכם הם רקובים, כלומר $X\sim HG\left(100,20,10\right)$ ולכן

$$Pr\left(X=4\right)=\frac{\binom{20}{4}\binom{80}{6}}{\binom{100}{10}}=0.084$$

□

בבעיה 6.4 קיבלנו הסתברויות דומות מאוד לבחור 4 תפוחים רקובים, בין אם הבחירה נעשית עם החזרה ובין אם הבחירה נעשית ללא החזרה. הסיבה לכך היא שהאוכלוסיה גדולה ביחס לגודל המדגם - נבחרים רק 10 תפוחים מתוך 100. במצב כזה, אין כמעט הבדל בין "עם החזרה" לבין "בלי החזרה", שכן הסיכוי לבחור את אותו הפריט פעמיים הוא נמוך. תופעה דומה ראינו בפרדוקס ימי ההולדת (בעיה 3.1), כאשר מספר האנשים

קטן ביחס ל-365, הסיכוי ששניים חוגגים יום הולדת באותו יום הוא אפסי. זוהי דוגמא פרטית של קשר מיוחד הקיים בין התפלגות בינומית להיפרגאומטרית. ההתפלגות הבינומית היא קירוב טוב (ובהרבה מקרים, נוח מאוד) להתפלגות היפרגאומטרית כאשר הסיכוי לדגימה כפולה נמוך, כלומר כאשר האוכלוסיה גדולה בהרבה מהמדגם.

משפט 6.6 (הקירוב הבינומי למשתנה מקרי היפרגאומטרי). *יהי* $X \sim HG(N, D, n)$. *כאשר* $N \gg n$ *המשתנה המקרי* X *מתפלג בקירוב בינומית עם הפרמטרים* $Bin\left(n, \frac{D}{N}\right)$. *ככלל אצבע, ניתן לבצע קירוב זה כאשר* $10n < N$.

הוכחה. נגדיר $p = \frac{D}{N}$. נתחיל מחישוב ההסתברות של המאורע $\{X = 0\}$:

$$Pr\left(X=0\right) = \frac{\binom{N-D}{n}}{\binom{N}{n}} = \frac{n!\left(N-n\right)!\left(N-D\right)!}{n!\left(N-D-n\right)!N!} = \frac{\left(N-n\right)!}{N!} \cdot \frac{\left(N-D\right)!}{\left(N-D-n\right)!}$$

כאשר $N \gg n$, בקירוב טוב $N = N-1 = \ldots = N-n+1$ ולכן

$$Pr\left(X=0\right) \simeq \frac{1}{N^n} \cdot \left(N-D\right) \cdot \ldots \cdot \left(N-D-n+1\right) = \left(1 - \frac{D}{N}\right) \cdot \ldots \cdot \left(1 - \frac{D+n-1}{N}\right)$$

כמוקדם, n זניח ביחס ל-N ולכן כל אחד מהסוגריים שווה בקירוב ל-$\left(1 - \frac{D}{N}\right) = 1-p$ וההסתברות הכוללת היא

$$Pr\left(X=0\right) = \left(1-p\right)^n$$

תוצאה זו מתאימה להתפלגות בינומית. נמשיך בחישוב היחס בין ההסתברויות של כל שני ערכים עוקבים בתומך:

$$\frac{Pr\left(X=k+1\right)}{Pr\left(X=k\right)} = \frac{\binom{D}{k+1} \cdot \binom{N-D}{n-k-1}}{\binom{D}{k} \cdot \binom{N-D}{n-k}} = \frac{\frac{D!}{(k+1)!(D-k-1)!} \cdot \frac{(N-D)!}{(n-k-1)!(N-D-n+k+1)!}}{\frac{D!}{k!(D-k)!} \cdot \frac{(N-D)!}{(n-k)!(N-D-n+k)!}} =$$
$$= \frac{D-k}{k+1} \cdot \frac{n-k}{N-D-n+k+1} = \frac{n-k}{k+1} \cdot \frac{\frac{D-k}{N}}{1 - \frac{D+n-k-1}{N}}$$

בגבול שבו N גדול המונה הוא p והמכנה הוא $1-p$. לפיכך, היחס בין הסתברויות של כל שני ערכים עוקבים בתומך הינו

$$\frac{Pr\left(X=k+1\right)}{Pr\left(X=k\right)} = \frac{n-k}{k+1} \cdot \frac{p}{1-p}$$

לסיכום, בגבול בו $N \gg n$ ההסתברות של כל ערך בתומך הינה

$$Pr\left(X=k\right) = \frac{n-k+1}{k} \cdot \frac{p}{1-p} \cdot Pr\left(X=k-1\right) = \ldots =$$
$$= \left(1-p\right)^n \cdot \prod_{i=1}^{k} \left(\frac{n-i+1}{i} \cdot \frac{p}{1-p}\right) = p^k \left(1-p\right)^{n-k} \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{1 \cdot \ldots \cdot k}$$
$$= \binom{n}{k} p^k \left(1-p\right)^{n-k}$$

וזו בדיוק ההתפלגות הבינומית. □

שלושה מרכיבים מרכזיים מתארים את הסיפור המתאים למשתנה מקרי היפרגאומטרי:

1. אוכלוסיה בגודל נתון, מתוכה מספר ידוע מראש של פריטים מיוחדים.
2. בחירה של מספר נתון של פריטים מתוך האוכלוסיה ללא החזרה.
3. המשתנה המקרי סופר את מספר הפריטים המיוחדים שנבחרו.

כמו במשתנים המקריים האחרים, ניתן לחשב גם עבור משתנה מקרי היפרגאומטרי תוחלת, שונות ופונקציה יוצרת מומנטים. הבעיה שנוסחא **6.21** והטורים המתקבלים ממנה מסובכים לחישוב ישיר. במקום זאת נחשב את התוחלת והשונות על ידי פירוק לאינדיקטורים. חישוב הפונקציה היוצרת מומנטים מורכב והתוצאה הסופית מסובכת וכוללת את הפונקציה ההיפרגאומטרית ${}_2F_1$ המקנה למשתנה המקרי את שמו.

משפט 6.7. *יהי* $X \sim HG\left(N,D,n\right)$. *אזי*

$$\mathbb{E}\left(X\right)=n\frac{D}{N} \quad (6.22)$$

$$\mathbb{V}\left(X\right)=n\frac{D}{N}\left(1-\frac{D}{N}\right)\left(\frac{N-n}{N-1}\right) \quad (6.23)$$

הוכחה. נניח שמוציאים את n הפריטים עם סדר. נגדיר ב-$\mathbf{1}_i$ את האינדיקטור השווה ל-1 כאשר הפריט ה-i מיוחד כאשר $i\in\left\{1,\ldots,n\right\}$. מספר הפריטים המיוחדים הוא בדיוק מספר האינדיקטורים השווה ל-1, לכן

$$X=\sum_{i=1}^{n}\mathbf{1}_i$$

ההסתברות שהפריט ה-i יהיה מיוחד הוא $Pr\left(\mathbf{1}_i=1\right)=\frac{D}{N}$ שכן יש N אפשרויות עבור הפריט ה-i מתוכן D אפשרויות מיוחדות. התוחלת של X מתקבלת מתכונת החיבוריות (נוסחא **5.7**):

$$\mathbb{E}\left(X\right)=\mathbb{E}\left(\sum_{i=1}^{n}\mathbf{1}_i\right)=\sum_{i=1}^{n}\mathbb{E}\left(\mathbf{1}_i\right)=\sum_{i=1}^{n}\frac{D}{N}=n\frac{D}{N}$$

חישוב השונות אינו אפשרי כרגע שכן האינדיקטורים תלויים. אם הפריט הראשון מיוחד זה מקטין את הסיכוי שהפריט השני מיוחד ולהפך. את השונות נחשב בתרגיל **7.21**. □

תרגיל 6.12. מטילים קוביה 10 פעמים. כיצד מתפלג מספר הפעמים שיצא 6 ב-5 ההטלות הראשונות בהינתן ש-6 יצא בסך הכל 4 פעמים?

תרגיל 6.13. בבריכה 20 דגים. 5 דגי זהב, 4 כרישים ו-11 קרפיונים. דייג דג 6 דגים. כיצד מתפלגים X– מספר דגי הזהב שהדייג דג ו-Y– מספר דגי הזהב שנשארו בבריכה.

6.1.6 משתנה מקרי פואסוני

בקניון מסויים מבקרים 10,000 אנשים ביום. לכל מבקר יש הסתברות של $p=0.0001$ לקנות ספר של עמוס עוז בחנות הספרים באופן בלתי תלוי במבקרים האחרים. מאחר וכל לקוח הוא ניסוי בלתי תלוי באחרים עם הסתברות קבועה לקנות ספר, מספר הספרים הכולל שיימכר הוא משתנה מקרי בינומי עם הפרמטרים

$Bin(10,000, 0.0001)$. מכאן והלאה אין בעיה לחשב את כל ההסתברויות של כל המאורעות שיכולים לעניין אותנו. למשל, הסיכוי שיימכרו בדיוק 2 ספרים הינו:

$$\binom{10,000}{2}0.0001^2 0.9999^{9998} = 0.184$$

חישובים אלו מסובכים יחסית שכן מלכתחילה הנוסחא המתארת התפלגות בינומית לא נוחה לעבודה ואילו בחישוב זה ישנן חזקות מאוד גבוהות של מספרים מאוד קרובים ל-1 וכן עצרות של מספרים מאוד גדולים. גורמים אלו עשויים לגרום לחישוב להיות לא נוח ולעיתים גם פחות מדויק. נרצה לקרב את ההתפלגות על ידי התפלגות פשוטה יותר עם נוסחא נוחה יותר להצבה – התפלגות פואסונית.

הגדרה 6.6 (משתנה מקרי פואסוני)**.** יהי $0 < \lambda \in \mathbb{R}$. משתנה מקרי פואסוני עם פרמטר λ הוא משתנה מקרי עם פונקציית ההתפלגות

$$P_X(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \tag{6.24}$$

כאשר $k \in supp(X) = \{0, 1, ...\}$. במקרה כזה נסמן

$$X \sim Pois(\lambda)$$

לפי תרגיל **0.19**, זו אכן פונקציית הסתברות.

***משפט* 6.8** (הקירוב הפואסוני להתפלגות בינומית)**.** *יהי* $X \sim Bin(n,p)$ *כאשר* $n \gg 1$, $p \ll 1$ *והגודל* $\lambda = np$ *קבוע. אזי בקירוב* $X \sim Pois(\lambda)$. *כלל האצבע הוא שקירוב זה טוב כאשר* $n > 100$, $p < 0.01$ *ומכפלתם היא מסדר גודל של 1:* $\lambda \in [0.05, 20]$.

הוכחה. יהי $k \in \{0, 1, \ldots\}$ ונחשב את ההסתברות של המאורע $\{X = k\}$ כאשר $n \gg 1, p \ll 1$ אבל מכפלתם $\lambda = np$ קבועה. כלומר, נציב $p = \frac{\lambda}{n}$ ונפשט את הביטוי:

$$\begin{aligned} Pr(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{n!}{k!(n-k)!} \lambda^k \frac{1}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

נשאיף את n לאינסוף. המכפלה הזו מורכבת ממכפלה של שלושה גבולות:

$$\lim_{n\to\infty} \frac{n!}{(n-k!)\, n^k} = 1$$
$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = 1$$

וכן

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n}=e^{-\lambda} \tag{6.25}$$

ובסך הכל נקבל שההסתברות שואפת עבור $n\to\infty$ ל-$e^{-\lambda}\frac{\lambda^k}{k!}$, כנדרש. $\square$

תרגיל 6.14. מפעל מייצר מנועים חשמליים. ההסתברות שמנוע ייצא פגום מקו היצור היא 0.01. מהי ההסתברות שפס ייצור של 300 מנועים יכלול יותר מ-2 מנועים פגומים? תנו תשובה מדויקת ותשובה מקורבת. מהו טיב הקירוב?

משתנה מקרי פואסוני מתאר הרבה תופעות טבע הדומות לקניון וספריו של עמוס עוז, בהן מספר הניסויים גדול מאוד, הסיכוי להצלחה קטן מאוד אך מכפלתם קבועה וסבירה. למשל, מספר ההתפרקויות של חומר רדיאקטיבי (מספר אטומים גדול, סיכוי קטן שכל אחד מהם יתפרק), מספר התקרים במהלך נסיעה (מספר קילומטרים גדול, סיכוי קטן לתקר בכל קילומטר), מספר טילי ה-V2 שפגעו בנהר התמזה של לונדון במהלך מלחמת העולם השניה (מספר טילים גדול, סיכוי קטן לפגיעה בנהר) ועוד. כאשר מכניסים למשתנה המקרי מימד נוסף של זמן או דמוי זמן (התפרקויות ליחידת זמן, תקרים ליחידת מרחק, פגיעות ליחידת זמן) מקבלים תהליך פואסון בו נדון בפרק 8. תהליכי פואסון מרכזיים בתחומים רבים של שכן ביכולתם לתאר תופעות טבע אקראיות רבות.

חשוב לציין שבניגוד למשתנים המקריים הבדידים האחרים, מאחורי משתנה מקרי פואסוני לא עומד סיפור קומבינטורי. לא ניתן להגדיר משתנה מקרי השייך לתהליך של הוצאת כדורים מכד או סידור פריטים בשורה כך שהמשתנה המקרי יתפלג פואסונית. כל שניתן הוא לבצע קירוב בינומי לפואסוני כפי שתואר לעיל או, לחילופין, התרגיל עצמו יגדיר שתופעת טבע מסויימת מתרחשת לפי תהליך פואסון. כאמור, נראה דוגמאות רבות מאוד להתפלגות פואסונית ולשימושיה בפרק 8.

משפט 6.9. *יהי* $X\sim Pois\left(\lambda\right)$. *אזי*

$$\mathbb{E}\left(X\right)=\mathbb{V}\left(X\right)=\lambda \tag{6.26}$$

הוכחה. נחשב לפי ההגדרה:

$$\begin{aligned}\mathbb{E}\left(X\right)&=\sum_{k=0}^{\infty}k\cdot e^{-\lambda}\frac{\lambda^{k}}{k!}=\sum_{k=1}^{\infty}e^{-\lambda}\frac{\lambda^{k}}{(k-1)!}\\&=\lambda\cdot\sum_{k=1}^{\infty}e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!}=\lambda\cdot\sum_{i=0}^{\infty}e^{-\lambda}\frac{\lambda^{i}}{i!}\overset{eq.\,0.19}{=}\lambda\end{aligned}$$

בדרך דומה,

$$\begin{aligned}\mathbb{E}\left(X^{2}\right)&=\sum_{k=0}^{\infty}k^{2}\cdot e^{-\lambda}\frac{\lambda^{k}}{k!}=\sum_{k=1}^{\infty}k\cdot e^{-\lambda}\frac{\lambda^{k}}{(k-1)!}\\(i=k-1)&=\sum_{k=1}^{\infty}k\cdot e^{-\lambda}\frac{\lambda^{k}}{(k-1)!}=\sum_{i=0}^{\infty}\left(i+1\right)e^{-\lambda}\frac{\lambda^{i+1}}{i!}\\&=\lambda\left(\sum_{i=0}^{\infty}ie^{-\lambda}\frac{\lambda^{i}}{i!}+\sum_{i=0}^{\infty}e^{-\lambda}\frac{\lambda^{i}}{i!}\right)\end{aligned}$$

הטור הראשון הוא הטור המתאר את התוחלת ואילו הטור השני הוא סכום כל ההסתברויות. לכן

$$\mathbb{E}\left(X^2\right)=\lambda\left(\lambda+1\right)$$

והשונות לפיכך תהיה

$$\mathbb{V}\left(X\right)=\mathbb{E}\left(X^2\right)-\mathbb{E}^2\left(X\right)=\lambda^2+\lambda-\lambda^2=\lambda$$

כנדרש. $\square$

תרגיל 6.15. יהי $X\sim Pois\left(\lambda\right)$. חשבו את הפונקציה היוצרת מומנטים של X.

תרגיל 6.16. במדינת ישראל יש כ-10^5 עורכי דין ושופטים. ועדה לאיתור שופטים מחפשת מועמדים לצורך משרת שופט שהתפנתה בבית המשפט העליון. להערכתם, הסיכוי של כל מועמד לענות על כל הדרישות (השכלה, ניסיון, ותק ועוד) הוא 10^{-5}. בטרם הועדה מתחילה לעבור על כל עורכי הדין והשופטים ולבחון את התאמתם לתפקיד, מודיע שר המשפטים כי הוא מכיר אישית שני מועמדים שיכולים להתאים. מה ההסתברות שהועדה לא תמצא מועמדים נוספים פרט לאלו שמכיר השר?

6.2 משפחות של משתנים מקריים רציפים

בסעיף זה נתאר את המשפחות העיקריות של משתנים מקריים רציפים. כפי שראינו כבר, בניגוד למשתנים מקריים בדידים, אין סיפור קומבינטורי המתאר את המשתנים המקריים הרציפים (שהרי התומך אינו בן מנייה). עיקר העבודה הינה עבודה טכנית עם התפלגות נתונה ולא זיהוי המשתנה המקרי המתאים לבעיה. המייחד משתנים מקריים אלו ביחס לסתם פונקציות צפיפות אפשריות הוא שהם מופיעים בתופעות טבע רבות ולכן שימושיים בחיים האמיתיים.

כפי שראינו בפרק 4, כל משתנה מקרי רציף ניתן לתיאור על ידי פונקציית ההתפלגות המצטברת או נגזרתה – פונקציית הצפיפות. למען השלמות, נציג לכל משתנה מקרי את שתי הפונקציות וכן נשרטט את פונקציית הצפיפות בכדי להבין כיצד ההתפלגות נראית ובאיזה טווח ערכים המשתנה המקרי צפוי להימצא. לא נטרח להוכיח שאכן פונקציית ההתפלגות המצטברת היא הקדומה של פונקציית הצפיפות והקורא החרוץ יוכל להשלים את הבדיקה לבד בקלות על ידי גזירה.

6.2.1 משתנה מקרי אחיד רציף

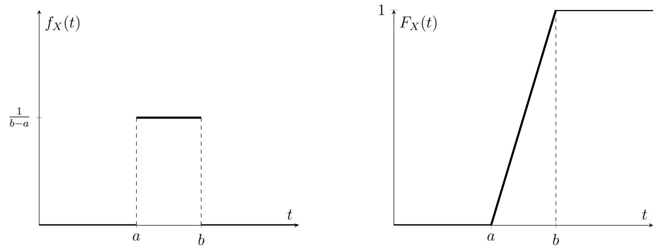
המשתנה המקרי הרציף הפשוט ביותר מתקבל כאשר הצפיפות היא פונקציה קבועה על פני קטע מסויים. במקרה כזה, כל ערך בקטע זה יכול להתקבל בהסתברות שווה ואין תחום ערכים עם הסתברות גבוהה יותר על פני תחומים אחרים. משתנה מקרי שצפיפותו קבועה על פני תחום מסויים מכונה משתנה מקרי אחיד רציף.

הגדרה 6.7 (משתנה מקרי אחיד רציף). משתנה מקרי רציף אשר צפיפותו קבועה בתחום $[a,b]$ נקרא משתנה מקרי אחיד רציף ומסומן $X\sim U\left(a,b\right)$. במקרה כזה פונקציית הצפיפות הינה

$$f_X\left(t\right)=\begin{cases}\frac{1}{b-a} & t\in[a,b]\\ 0 & t\notin[a,b]\end{cases} \tag{6.27}$$

ופונקציית ההתפלגות המצטברת הינה

$$F_X\left(t\right)=\begin{cases}0 & t<a\\ \frac{t-a}{b-a} & t\in\left[a,b\right]\\ 1 & t>b\end{cases} \tag{6.28}$$



שימו לב להבדל בין אחיד בדיד לאחיד רציף. אמנם שני המשתנים המקריים מסומנים באמצעות האות U (Uniform) אך משתנה מקרי אחיד בדיד היא התפלגות על קבוצה סופית ולכן גם הסימן הוא של קבוצה עם צומדיים $\{1,\ldots,4\}$ ואילו משתנה מקרי אחיד רציף מתפלג על קטע ולכן הסימון הוא הסימון המתאים לקטע $(1,4)$. כמובן שנקודות הקצה יכולות להיות חלק מהקטע אך היות ומדובר על משתנה מקרי רציף ואין משמעות לנקודה בודדת, הרי שאין הבדל בין התפלגות על הקטע הפתוח $(1,4)$או הקטע הסגור $[1,4]$.

משפט 6.10. *יהי $X\sim U\left(a,b\right)$. אזי*

$$\mathbb{E}\left(X\right)=\frac{a+b}{2} \tag{6.29}$$

$$\mathbb{V}\left(X\right)=\frac{\left(b-a\right)^2}{12} \tag{6.30}$$

$$M_X\left(t\right)=\frac{e^{tb}-e^{ta}}{t\left(b-a\right)} \tag{6.31}$$

תרגיל 6.17. הוכיחו את משפט 6.10

תרגיל 6.18. יהי $X\sim U\left(-1,1\right)$. כיצד מתפלג $Y=X^2$?

תרגיל 6.19. השעה שבה סטודנט מתעורר בבוקר ויוצא לאוניברסיטה, X, מתפלגת אחיד בין 8 ל-12 בבוקר. ככל שהסטודנט מתעורר מאוחר יותר, כך זמן הנסיעה קצר יותר לפי הנוסחא $Y=1+\frac{1}{X}$. כיצד מתפלג זמן הנסיעה של הסטודנט?

בשני התרגילים 6.18 ו-6.19 ראינו שפונקציה של משתנה מקרי אחיד רציף בדרך כלל אינה מתפלגת באופן אחיד.

6.2.2 משתנה מקרי מעריכי

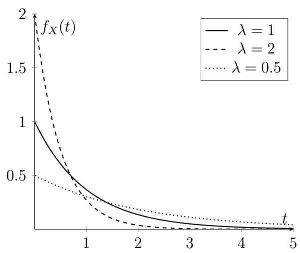
משתנה מקרי מעריכי מתאר במקרים רבים זמן בין תקלות או זמן חיים של מוצרים כתוצאה מאירוע "טראומטי" ולא בלייה טבעית. למשל, הזמן עד שנורה תישרף, הזמן עד שיחת הטלפון הבאה שאקבל ומרחק הנסועה עד לתקר הבא – כולם מתפלגים מעריכית. נעמוד עוד על משמעותו של משתנה מקרי זה בפרק 8, שם הוא יופיע כזמן עד האירוע הראשון והזמן בין אירועים סמוכים של תהליך פואסון. הסיבה שמשתנה מקרי זה נקרא "מעריכי" הוא שפונקציית הצפיפות שלו מתוארת על ידי הפונקציה המעריכית $\exp\left(-x\right)$.

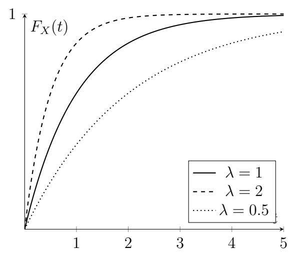
הגדרה 6.8 (משתנה מקרי מעריכי)**.** יהי $0<\lambda\in\mathbb{R}$. משתנה מקרי עם פרמטר λ יסומן על ידי $X\sim Exp\left(\lambda\right)$ והוא בעל פונקציית צפיפות

$$f_X\left(t\right)=\begin{cases}\lambda e^{-\lambda t} & t\geq0\\ 0 & t<0\end{cases}\tag{6.32}$$

ופונקציית ההתפלגות המצטברת

$$F_X\left(t\right)=\begin{cases}1-e^{-\lambda t} & t\geq0\\ 0 & t<0\end{cases}\tag{6.33}$$





מבחינת יחידות, היחידות של λהופכיות ליחידות של X. כך למשל, אם X מתאר את המרחק (בקילומטרים) עד התקר הבא הרי שהיחידות של λ יהיו $\left[\frac{1}{km}\right]$.

בבעיה 0.1 מוכיחים שאכן פונקציית ההתפלגות המצטברת של משתנה מקרי מעריכי היא הפונקציה הקדומה של פונקציית הצפיפות וכן שפונקציית הצפיפות מנורמלת.

תרגיל 6.20. הזמן (בדקות) שלוקח לבריסטה להכין כוס קפה, X, מתפלג מעריכית עם פרמטר 1. בוצע סידור מחדש של עמדת העבודה של הבריסטה וכעת פרק הזמן שלוקח לו להכין כוס קפה מקיים $Y=\frac{X}{2}+5$. כיצד מתפלג Y?

תרגיל 6.21. יהי $X\sim Exp\left(1\right)$ ונגדיר $Y=\lfloor X\rfloor+1$ ו-$Z=X-\lfloor X\rfloor$, כאשר $\lfloor a\rfloor$ הוא החלק השלם של a.

1. חשבו את ההתפלגות של Y.
2. חשבו את התוחלת והשונות של Y.

3. חשבו את ההתפלגות של Z.

בעיה 6.5. הזמן (בשעות) עד שריפת נורה חדשה, X, מתפלג מעריכית עם פרמטר $\frac{1}{1000}$.

1. מה הסיכוי שנורה חדשה תדלוק מעל ל-2000 שעות?
2. מה הסיכוי שנורה שכבר דלקה מעל ל-2000 שעות תדלק לפחות עוד 2000 שעות?

הוכחה. לפי הנתון, $X \sim Exp\left(\frac{1}{1000}\right)$.

1. נורה חדשה תדלוק מעל ל-2000 שעות אם $X > 2000$ ולכן

$$\begin{aligned} Pr\left(X > 2000\right) &= 1 - F_X\left(2000\right) = 1 - \left(1 - e^{-\frac{1}{1000}\cdot 2000}\right) \\ &= e^{-2} = 0.135 \end{aligned}$$

2. ידוע שהנורה דלקה כבר 2000שעות, לכן המאורע $X > 2000$ התקיים. הנורה תדלוק לפחות עוד 2000 שעות אם בסך הכל $X > 4000$. לכן המאורע המבוקש הוא המאורע המותנה $X > 4000|X > 2000$ שהסתברותו

$$\begin{aligned} Pr\left(X > 4000|X > 2000\right) &= \frac{Pr\left(X > 4000 \cap X > 2000\right)}{Pr\left(X > 2000\right)} = \frac{Pr\left(X > 4000\right)}{Pr\left(X > 2000\right)} \\ &= \frac{1 - F_X\left(4000\right)}{1 - F_X\left(2000\right)} = \frac{e^{-4}}{e^{-2}} = e^{-2} \end{aligned}$$

□

ההסתברות שנורה שדלקה 2000 שעות תדלוק עוד 2000 שעות זהה להסתברות שנורה חדשה תדלוק 2000 שעות. העובדה שהנורה כבר דלקה הרבה זמן לא הוסיפה לה בלאי או הגדילה את הסיכוי שהיא תישרף בקרוב, וההתפלגות של זמן שריפת הנורה לאחר 2000 שעות זהה להתפלגות זמן שריפת נורה חדשה. בתכונה זו נפגשנו כבר בהקשר של משתנה מקרי גאומטרי – תכונת חוסר הזיכרון (נוסחא 6.13). הנורה "לא זוכרת" שהיא דלקה כבר זמן רב וזמן עבודה זה לא משפיע על התפלגות הזמן שנותר עד שהיא תישרף. כשם שמשתנה מקרי גאומטרי הוא המשתנה המקרי הבדיד היחיד הניחן בתכונת חוסר הזיכרון, כך גם המשתנה המקרי המעריכי הוא היחיד הניחן בתכונת חוסר הזיכרון.

***משפט* 6.11** (תכונת חוסר הזיכרון של משתנה מקרי מעריכי). *יהי* $X \sim Exp\left(\lambda\right)$. *המשתנה המקרי* X *חסר זיכרון, כלומר לכל* $0 < t, s \in \mathbb{R}$ *מתקיים*

$$Pr\left(X > t + s|X > t\right) = Pr\left(X > s\right) \qquad (6.34)$$

הוכחה. נחשב ישירות כפי שביצענו בדוגמא:

$$\begin{aligned} Pr\left(X > s + t|X > t\right) &= \frac{Pr\left(X > s + t \cap X > t\right)}{Pr\left(X > t\right)} = \frac{Pr\left(X > s + t\right)}{Pr\left(X > t\right)} \\ &= \frac{1 - F_X\left(s + t\right)}{1 - F_X\left(t\right)} = \frac{1 - \left(1 - e^{-\lambda(s+t)}\right)}{1 - \left(1 - e^{-\lambda t}\right)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} \\ &= e^{-\lambda s} = 1 - \left(1 - e^{-\lambda s}\right) = 1 - F_X\left(s\right) = Pr\left(X > s\right) \end{aligned}$$

□

תכונת חוסר הזיכרון הופכת את המשתנה המקרי המעריכי לנוח מאוד לעבודה, שכן כאשר זמן של מאורע כלשהו מתפלג לפי התפלגות מעריכית, אין צורך לדעת את העבר או כמה זמן עבר מאז האירוע הקודם כדי לדעת מתי יקרה האירוע הבא. למשל, אם מרחק הנסועה עד התקר הבא בגלגל אכן מתפלג מעריכית, הרי זה לא משנה כמה כבר הגלגל עבר כדי לקבוע מה הסיכוי שהוא ישלים את הנסיעה הקרובה. באופן דומה, לו זמן הגעת האוטובוס לתחנה היה מתפלג מעריכית, לא היה טעם לשאול את הממתינים בתחנה האם הקו שלכם עבר וכמה זמן הם ממתינים שם. מהרגע שהגעתם, זמן ההגעה של האוטובוס הוא מעריכי ללא תלות בזמן ההמתנה. לעומת זאת, אם זמן הגעת האוטובוס לתחנה לא היה מתפלג מעריכית, הרי שמידע זה על הזמן בו האוטובוס לא הגיע היה מאוד חשוב ומשנה את התפלגות המותנית של זמן ההגעה של האוטובוס הקרוב.

תרגיל 6.22. דני ויונתן הולכים לשחק טניס. בהסתברות $\frac{1}{4}$ יש להם כוח ואז זמן המשחק (בשעות) מתפלג מעריכית עם פרמטר $\frac{1}{2}$. אחרת, אין להם כוח לשחק וזמן המשחק (בשעות) מתפלג מעריכית עם פרמטר 1.

1. מה הסיכוי שהמשחק יארך בין שעה לשעתיים?
2. כיצד מתפלג זמן המשחק?
3. השחקנים משחקים כבר שעה. מה הסיכוי שהמשחק יימשך יותר משעתיים בסך הכל?

משפט 6.12. *יהי* $X \sim Exp\left(\lambda\right)$*. אזי*

$$\mathbb{E}\left(X\right) = \frac{1}{\lambda} \tag{6.35}$$

$$\mathbb{V}\left(X\right) = \frac{1}{\lambda^2} \tag{6.36}$$

$$M_X\left(t\right) = \frac{\lambda}{\lambda - t} \tag{6.37}$$

הוכחה. נתחיל מחישוב הפונקציה היוצרת מומנטים ונראה כיצד באמצעותה ניתן לחשב את כל המומנטים בצורה פשוטה יחסית.

$$\begin{aligned} M_X\left(t\right) = \mathbb{E}\left(e^{tX}\right) &= \int_0^\infty e^{tu}\lambda e^{-\lambda u}\,du = \lambda\int_0^\infty e^{(t-\lambda)u}\,du \\ (eq.\,0.33) &= \frac{\lambda}{t-\lambda}\left. e^{(t-\lambda)u}\right|_0^\infty = \frac{\lambda}{\lambda - t} \end{aligned}$$

כאשר נוסחה זו נכונה כמובן רק כאשר $t < \lambda$, אחרת האינטגרל מתבדר. אין עם הגבלה זו בעיה שכן העיקר שהפונקציה היוצרת מומנטים מוגדרת בסביבה של $t = 0$. ניתן לרשום את הפונקציה היוצרת מומנטים גם כסכום של סדרה גאומטרית (נוסחא 0.15) באופן הבא:

$$M_X\left(t\right) = \frac{\lambda}{\lambda - t} = \frac{1}{1 - \frac{t}{\lambda}} = \sum_{k=0}^{\infty}\left(\frac{t}{\lambda}\right)^k$$

נחשב את המומנט ה-n על ידי נוסחא 5.22. עבור $k < n$, הנגזרת ה-n של t^n מתאפסת ולכן נשארות החזקות הגדולות מ-n:

$$\mathbb{E}\left(X^n\right) = \left.\frac{d^n}{dt^n}M_X\left(t\right)\right|_{t=0} = \left.\sum_{k=n}^{\infty}\frac{k\cdot(k-1)\cdot\ldots\cdot(k-n+1)}{\lambda^k}t^{k-n}\right|_{t=0}$$

כאשר $k > n$ הביטוי t^{k-n} מתאפס כאשר מציבים $t = 0$. הגורם היחיד שאינו מתאפס בטור הוא זה עבורו $k = n$ וערכו

$$\mathbb{E}\left(X^n\right) = \frac{n!}{\lambda^n}$$

בפרט,

$$\mathbb{E}\left(X\right) = \frac{1}{\lambda}$$
$$\mathbb{E}\left(X^2\right) = \frac{2}{\lambda^2}$$

ולכן

$$\mathbb{V}\left(X\right) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2\left(X\right) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

כנדרש. חישוב זה מדגים חלק מהעוצמה של הפונקציה היוצרת מומנטים, המאפשרת לחשב את כל המומנטים של המשתנה המקרי על ידי ביצוע אינטגרל בודד ופירוק התוצאה לטור חזקות. $\square$

6.2.3 משתנה מקרי נורמלי

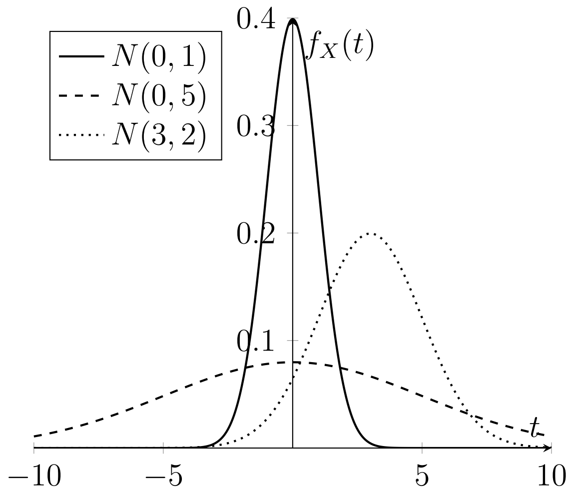
משתנה מקרי רציף נוסף המתאר מספר רב של תופעות טבע הוא משתנה מקרי נורמלי המוכר גם בתור התפלגות גאוסית או פעמון גאוס. בהתפלגות זו, רוב ההסתברות מרוכזת סביב התוחלת וככל שמתרחקים ממנה הצפיפות יורדת באופן סימטרי ומעריכי. משתנה מקרי נורמלי מתאר למשל רעש במערכות מדידה, התפלגות של תכונות דוגמת משקל, גובה, ו-IQ באוכלוסיה וכן ציונים בבחינה. בנוסף, לפי משפט הגבול המרכזי (משפט **6.16**) ממוצע של הרבה משתנים מקריים בלתי תלויים ושווי התפלגות מתפלג נורמלית גם כן. היות ובמקרים רבים אנו מודדים ממוצעים (למשל במעבדה, חוזרים על הניסוי מספר פעמים ומחשבים את ממוצע המדידות), הרי שהתפלגות זו חשובה במיוחד.

הגדרה 6.9 (משתנה מקרי נורמלי). יהיו $\mu, \sigma \in \mathbb{R}$. משתנה מקרי נורמלי הוא משתנה מקרי בעל פונקציית הצפיפות

$$f_X\left(t\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \qquad (6.38)$$

ונסמן $X \sim N\left(\mu, \sigma^2\right)$. כאשר מתקיים $\mu = 0, \sigma = 1$ המשתנה המקרי נקרא נורמלי סטנדרטי ומסומן לרוב על ידי האות $Z \sim N\left(0, 1\right)$. פונקציית הצפיפות המתאימה הינה

$$f_Z\left(t\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \qquad (6.39)$$



בהתאם לאיור, הפרמטר μ של ההתפלגות קובע את מרכזה ואילו הפרמטר σ^2 קובע את רוחבה. ככל ש-σ^2 קטן יותר כך ההתפלגות מרוכזת יותר קרוב ל-μ וככל שהוא גדול יותר כך ההתפלגות רחבה יותר והסיכוי לקבל ערכים רחוקים מ-μ גדל. אם תיאור זה מזכיר את התוחלת והשונות הרי שזה לא מפתיע שכן כפי שיוכח במשפט **6.15**, פרמטרים אלו הם בדיוק התוחלת והשונות של ההתפלגות.

הבעיה המרכזית של המשתנה המקרי הנורמלי היא שפונקציית הצפיפות היא מהצורה e^{-x^2} ולפונקציה זו אין פונקציה קדומה אנליטית. אמנם ניתן לחשב את האינטגרל על כל הישר בשיטות שונות החורגות ממסגרת הדיון ולקבל שאכן הצפיפות מנורמלת, כלומר אכן מתקיים

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx=1 \qquad \textbf{(6.40)}$$

אבל אין דרך לרשום אינטגרלים מהצורה $\int_{-\infty}^{t}e^{-x^2}\,dx$ כנוסחא סגורה הכוללת פונקציות אלמנטריות. בלית ברירה, יש לחשב את הפונקציה הקדומה באמצעים נומריים. בעמוד **224** מוצג חישוב זה עבור משתנה מקרי נורמלי סטנדרטי והפונקציה המתקבלת מסומנת על ידי

$$\Phi\left(z\right)=Pr\left(Z\le z\right)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,dt \qquad \textbf{(6.41)}$$

כל שורה בטבלה מייצגת את ספרת האחדות והספרה העשרונית הראשונה של המספר z עבורו מנסים לחשב את $\Phi\left(z\right)$ וכל עמודה מייצגת את הספרה העשרונית השנייה. המספר הרשום בשורה ובעמודה המתאימים למספר z הוא $\Phi\left(z\right)$. למשל, אם אנחנו מעוניינים לחשב את $\Phi\left(2.34\right)$עלינו לבחור בשורה בה רשום 2.3 ובעמודה בה רשום 0.04. המספר הרשום בהצטלבות השורה והעמודה הללו, 0.99036, הוא בדיוק $\Phi\left(2.34\right)$.

הערך הכי קטן שניתן לחשב ישירות באמצעות הטבלה הוא $\Phi\left(0\right)=0.5$ שכן בשורות לא מופיעים מספרים שליליים. הסיבה לכך היא שההתפלגות סימטרית סביב אפס:

$$f_Z\left(-t\right)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(-t)^2}{2}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}=f_Z\left(t\right)$$

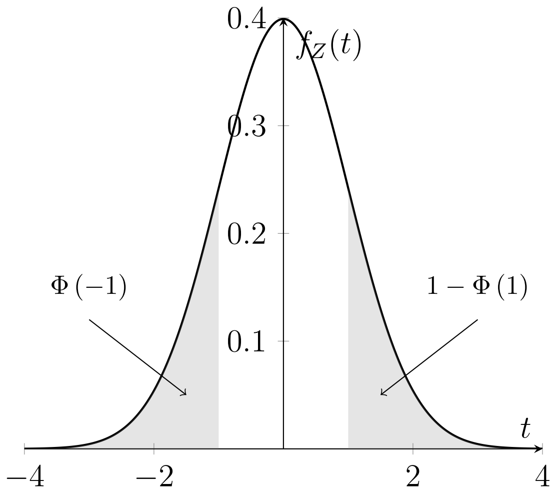
ולכן לכל $z<0$ מתקיים

$$Pr\left(Z\leq z\right)=\int_{-\infty}^{z}f_{Z}\left(t\right)\ dt=\int_{|z|}^{\infty}f_{Z}\left(t\right)\ dt=1-\Phi\left(-z\right)$$

כלומר, מתקיים הקשר החשוב

$$\Phi\left(z\right)=1-\Phi\left(-z\right) \tag{6.42}$$

באמצעותו ניתן לחשב את פונקציית ההתפלגות המצטברת גם של מספרים שליליים בעזרת הטבלה.



כאשר הערך המבוקש חורג בערכו המוחלט מטווח הערכים המכוסה על ידי הטבלה, ניתן לעגל את התוצאה ל-1 כאשר החריגה היא כלפי מעלה או ל-0 כאשר החריגה היא כלפי מטה. כך למשל, $\Phi\left(20\right)=1$ וכן $\Phi\left(-10\right)=0$. תוצאות אלו כמובן הינן רק קירוב אך למרבית הצרכים הפרקטיים קירוב זה מספיק. בנוסף, כאשר הערך המבוקש לא מופיע בטבלה שכן הוא כולל מספר רב יותר של ספרות משמעותיות (למשל, $\Phi\left(2.32334\right)$) ניתן לעגל למספר הקרוב ביותר הכולל רק שתי ספרות לאחר הנקודה העשרונית ולחשב בהתאם. ישנן שיטות נוספות כיצד להתגבר על בעיה זו (למשל, אינטרפולציה לינארית בין נקודות סמוכות המופיעות בטבלה), אך עיגול בהחלט מספיק שכן ממילא הטבלה הינה כלי ארכאי המופיע בעיקר בקורסים בהסתברות לצורך חישובי הסתברויות של התפלגות נורמלית סטנדרטית במהלך תרגילים והמבחן, אך לא בחיים האמיתיים שם ניתן להיעזר במחשב ולחשב את פונקציית ההתפלגות המצטברת של כל ערך רצוי בדיוק טוב כרצוננו. לא נתעכב על השימוש בטבלה מעבר לתרגיל הבא.

תרגיל 6.23. רעש מדידת המתח במעגל חשמלי מתפלג נורמלית סטנדרטית.

1. מהי ההסתברות שהרעש יהיה גדול מ-1?
2. מהי ההסתברות שהרעש יהיה גדול מ-1 בערכו המוחלט?
3. תוצאה חריגה היא תוצאה הגדולה מ-95% מהתוצאות האחרות. מהו המתח אשר מעבר לו התוצאה תיחשב לחריגה?

כאמור, הטבלה מאפשרת לחשב את פונקציית ההתפלגות המצטברת של משתנה מקרי נורמלי סטנדרטי. כאשר המשתנה המקרי מתפלג נורמלית עם פרמטרים כלשהם, יש צורך בטבלה חדשה ומתאימה. לשמחתנו, פונקציה לינארית של משתנה מקרי נורמלי היא גם כן משתנה מקרי נורמלי ולכן כל משתנה מקרי נורמלי ניתן להפוך למשתנה מקרי נורמלי סטנדרטי על ידי בחירת המקדמים המתאימים. לאור זאת, אין צורך בחישוב טבלה חדשה לכל הפרמטרים האפשריים אלא די בטבלה אחת עבור המשתנה המקרי הנורמלי הסטנדרטי ובתהליך ההפיכה של כל משתנה מקרי נורמלי למשתנה מקרי נורמלי סטנדרטי, תהליך המכונה "תיקנון".

משפט 6.13. *יהי $X \sim N(\mu, \sigma^2)$ ויהיו $a, b \in \mathbb{R}$ כאשר $a > 0$. נגדיר משתנה מקרי חדש על ידי $Y = aX + b$. אזי המשתנה המקרי Y מתפלג נורמלית גם כן ומתקיים $Y \sim N(a\mu + b, a^2\sigma^2)$.*
בפרט, $\frac{X-\mu}{\sigma} = Z \sim N(0,1)$ ולכן ניתן לחשב את פונקציית ההתפלגות המצטברת של X על ידי תהליך תיקנון - החסרת μ וחלוקה ב-σ:

$$F_X(t) = \Phi\left(\frac{t-\mu}{\sigma}\right) \tag{6.43}$$

הוכחה. נחשב את פונקציית ההתפלגות המצטברת של Y בנקודה $t \in \mathbb{R}$ ונראה שהיא מתאימה לפונקציית התפלגות מצטברת של משתנה מקרי נורמלי.

$$F_Y(t) = Pr(Y \le t) = Pr(aX + b \le t) = Pr\left(X \le \frac{t-b}{a}\right) = \int_{-\infty}^{\frac{t-b}{a}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

נבצע את החלפת המשתנים $y = ax + b$. לכן $dy = a \cdot dx$ והאינטגרל האחרון הינו

$$F_Y(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}} \frac{dy}{a} = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi(a\sigma)^2}} e^{-\frac{(y-a\mu-b)^2}{2a^2\sigma^2}}\, dy$$

האינטגרנד האחרון הוא בדיוק פונקציית הצפיפות של משתנה מקרי נורמלי $N(a\mu + b, a^2\sigma^2)$ ולכן $Y \sim N(a\mu + b, a^2\sigma^2)$.
בפרט, כאשר מבצעים תיקנון ומגדירים $Y = \frac{X-\mu}{\sigma}$ מקבלים $Y \sim N(0,1)$ שכן זה שקול להצבת $a = \frac{1}{\sigma}$ ו-$b = -\frac{\mu}{\sigma}$ בנוסחא האחרונה. □

בעיה 6.6. ציון הבחינה הפסיכומטרית מתפלג נורמלית עם $\mu = 500$ ו-$\sigma = 100$.

1. מה הסיכוי שנבחן אקראי יקבל ציון גבוה מ-700?
2. מה הסיכוי שנבחן אקראי יקבל ציון בין 400 ל-600?
3. סף הקבלה לפקולטה להנדסה הוא האחוזון ה-90. מהו הציון המינימלי הדרוש בכדי להתקבל לפקולטה?

הוכחה. נסמן ב-$X \sim N(500, 100^2)$ את ציון הפסיכומטרי. בכדי לחשב את ההסתברויות נבצע בכל אחד מהסעיפים תיקנון על ידי $Z = \frac{X-500}{100}$ כך שנוכל להשתמש בטבלה של משתנה מקרי נורמלי סטנדרטי.

.1

$$Pr(X > 700) = Pr\left(\frac{X-500}{100} > \frac{700-500}{100}\right) = Pr(Z > 2) = 1 - \Phi(2) = 1 - 0.977 = 0.023$$

.2

$$\begin{aligned} Pr\left(400 < X < 600\right) &= Pr\left(\frac{400-500}{100} < \frac{X-500}{100} < \frac{600-500}{100}\right) = Pr\left(-1 < Z < 1\right) \\ &= \Phi\left(1\right) - \Phi(-1) = \Phi\left(1\right) - \left(1-\Phi\left(1\right)\right) = 2\Phi\left(1\right) - 1 \\ &= 2\cdot 0.841 - 1 = 0.682 \end{aligned}$$

3. הפעם נדרש האחוזון ה-90 של ההתפלגות, $x_{0.9}$, המקיים $Pr\left(X \le x_{0.9}\right) = 0.9$. נתקנן את המשוואה האחרונה ונקבל

$$0.9 = Pr\left(X \le x_{0.9}\right) = Pr\left(\frac{X-500}{100} \le \frac{x_{0.9}-500}{100}\right) = \Phi\left(\frac{x_{0.9}-500}{100}\right)$$

האחוזון ה-90 של התפלגות נורמלית סטנדרטית הינו $z_{0.9} = 1.285$ שכן בקירוב $\Phi\left(1.285\right) = 0.9$ ולכן $\frac{x_{0.9}-500}{100} = 1.285$ כלומר $x_{0.9} = 628.5$.

□

לאור התוצאה של הסעיף האחרון נקבל את הקשר בין האחוזונים של התפלגות נורמלית סטנדרטית להתפלגות נורמלית כלשהי.

משפט 6.14. *יהי $X \sim N\left(\mu,\sigma^2\right)$. האחוזון ה-$\alpha$ של המשתנה המקרי, x_α, מקיים את המשוואה*

$$x_\alpha = \sigma z_\alpha + \mu \tag{6.44}$$

כאשר z_α הוא האחוזון ה-α של ההתפלגות הנורמלית הסטנדרטית.

הוכחה. האחוזון x_α מקיים $Pr\left(X \le x_\alpha\right) = \alpha$. נתקנן:

$$\alpha = Pr\left(X \le x_\alpha\right) = Pr\left(\frac{X-\mu}{\sigma} \le \frac{x_\alpha-\mu}{\sigma}\right) = \Phi\left(\frac{x_\alpha-\mu}{\sigma}\right)$$

לפי ההגדרה, $\alpha = \Phi\left(z_\alpha\right)$ ולכן $z_\alpha = \frac{x_\alpha-\mu}{\sigma}$ כנדרש. □

כאמור לעיל, הפרמטרים המאפיינים את ההתפלגות הנורמלית הם התוחלת והשונות שלה. נוכיח כעת עובדה זו.

משפט 6.15. *יהי $X \sim N\left(\mu,\sigma^2\right)$. אזי*

$$\mathbb{E}\left(X\right) = \mu \tag{6.45}$$

$$\mathbb{V}\left(X\right) = \sigma^2 \tag{6.46}$$

הוכחה. יהי $X \sim N\left(\mu,\sigma^2\right)$. ניתן לחשב מפורשות את התוחלת על ידי אינטגרציה, אבל יותר פשוט להשתמש במשפט 6.13 ולינאריות התוחלת (נוסחא 5.3):

$$\mathbb{E}\left(X\right) = \mathbb{E}\left(\sigma Z + \mu\right) = \sigma\mathbb{E}\left(Z\right) + \mu$$

טבלה 6.1: $\Phi(z)$

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,5279	0,53188	0,53586
0,1	0,53983	0,5438	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,6293	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,6591	0,66276	0,6664	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,7054	0,70884	0,71226	0,71566	0,71904	0,7224
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,7549
0,7	0,75804	0,76115	0,76424	0,7673	0,77035	0,77337	0,77637	0,77935	0,7823	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,8665	0,86864	0,87076	0,87286	0,87493	0,87698	0,879	0,881	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,9032	0,9049	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,9222	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,9452	0,9463	0,94738	0,94845	0,9495	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,9608	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,9732	0,97381	0,97441	0,975	0,97558	0,97615	0,9767
2	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,9803	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,983	0,98341	0,98382	0,98422	0,98461	0,985	0,98537	0,98574
2,2	0,9861	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,9884	0,9887	0,98899
2,3	0,98928	0,98956	0,98983	0,9901	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,9918	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,9943	0,99446	0,99461	0,99477	0,99492	0,99506	0,9952
2,6	0,99534	0,99547	0,9956	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,9972	0,99728	0,99736
2,8	0,99744	0,99752	0,9976	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,999
3,1	0,99903	0,99906	0,9991	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,9994	0,99942	0,99944	0,99946	0,99948	0,9995
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,9996	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,9997	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,9998	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,9999	0,9999	0,9999	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998

את התוחלת של משתנה מקרי נורמלי סטנדרטי Z חישבנו בתרגיל **5.11** באמצעות סימטריה וראינו שערכה $\mathbb{E}(Z) = 0$. לכן $\mathbb{E}(X) = \mu$.

לצורך חישוב השונות נשתמש באותו תהליך אך הפעם באמצעות לינאריות השונות (נוסחא **5.15**)

$$\mathbb{V}(X) = \mathbb{V}(\sigma Z + \mu) = \sigma^2 \mathbb{V}(Z)$$

נותר לחשב את השונות של משתנה מקרי נורמלי סטנדרטי. מאחר והתוחלת היא 0, השונות שווה לתוחלת של Z^2:

$$\mathbb{V}(Z) = \mathbb{E}\left(Z^2\right) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} t^2 e^{-\frac{t^2}{2}}\, dt = 2 \cdot \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} t^2 e^{-\frac{t^2}{2}}\, dt$$

כאשר הכפל ב-2 נובע מכך שזהו אינטגרל על פונקציה זוגית בתחום סימטרי סביב ה-0. נבצע אינטגרציה בחלקים כאשר $f(t) = 2t \cdot e^{-\frac{t^2}{2}}$ ו-$G(t) = t$. לכן $F(t) = -e^{-\frac{t^2}{2}}$ ו-$g(t) = 1$:

$$2 \cdot \int_0^{\infty} \frac{1}{\sqrt{2\pi}} t^2 e^{-\frac{t^2}{2}}\, dt = -\frac{1}{\sqrt{2\pi}} t e^{-\frac{t^2}{2}} \bigg|_0^{\infty} - \int_0^{\infty} -\frac{2}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

הביטוי הראשון מתאפס בשני הגבולות. האינטגרל השני הינו אינטגרל על פונקציית הצפיפות של משתנה מקרי נורמלי סטנדרטי בין 0 ל-∞ ולכן שווה ל-$Pr(Z > 0) = \frac{1}{2}$. לסיכום

$$\mathbb{V}(Z) = 0 + 2 \cdot \frac{1}{2} = 1$$

כנדרש. לכן לכל $X \sim N(\mu, \sigma^2)$ מתקיים $\mathbb{V}(X) = \sigma^2 \mathbb{V}(Z) = \sigma^2$. □

תרגיל 6.24. חשבו את הפונקציה היוצרת מומנטים של משתנה מקרי $X \sim N(\mu, \sigma^2)$.
רמז: כדאי להתחיל מחישוב הפונקציה היוצרת מומנטים של משתנה מקרי נורמלי סטנדרטי ואז להשתמש במשפט **6.13** ותרגיל **5.24**.

תרגיל 6.25. הגובה של גברים באוכלוסיה מתפלג נורמלית עם תוחלת 175 ס"מ וסטיית תקן 10ס"מ. למיונים לקבוצת כדורסל מזמינים גברים שגובהם 2 מטרים ומעלה.

1. מהו אחוז הגברים המוזמן למיונים לקבוצת הכדורסל?
2. באוניברסיטה לומדים 10,000 גברים. מהי תוחלת מספר הגברים שיוזמנו למיונים?
3. כל גבר המוזמן למיונים נבחן בסדרה של מבחני מיומנות ובסיכוי $\frac{1}{3}$ מתקבל לקבוצת הכדורסל. המאמן מחליט לקבל את 12 הגברים הראשונים שיצליחו במבחני המיומנות. כיצד מתפלג מספר הגברים שנבחנו במבחני המיומנות עד שהנבחרת הושלמה?

הסיבה שהתפלגות נורמלית חשובה מאוד ואחת המרכזיות בהסתברות וסטטיסטיקה היא משפט הגבול המרכזי, לפיו ממוצעים של משתנים מקריים מתפלג נורמלית גם אם המשתנים המקריים לא מתפלגים נורמלית. חישוב ממוצע של משתנים מקריים היא פרקטיקה ידועה בתחומים מדעיים רבים ומשפט הגבול המרכזי מאפשר ביצוע חישובים הסתברותיים על ממוצע זה ללא העבודה הקשה של חישוב ההתפלגות של הגודל הנמדד. כך למשל, בניסויים רבים בהם רעש המדידה גדול יחסית, מקובל לחזור על הניסוי מספר רב של פעמים ולחשב את הממוצע. באופן זה, הרעש הגדול קטן משמעותית ובמקום זאת נקבל ערך קרוב יותר לערך האמיתי. מאחר ולפי משפט הגבול המרכזי ההתפלגות של הערך הממוצע מתפלגת נורמלית, אין צורך בניתוח מעמיק של מערך הניסוי וחישוב ההתפלגות האמיתית של הרעש.

***משפט* 6.16** (משפט הגבול המרכזי)**.** *תהי $X_1, X_2, \ldots$ סדרה של משתנים מקריים בלתי תלויים ושווי התפלגות בעלי תוחלת μ ושונות σ^2. נגדיר את ממוצע n המשתנים המקריים הראשונים בתור*

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

כאשר מספר המחוברים n שואף לאינסוף, הממוצע מתפלג בקירוב נורמלית עם תוחלת μ וסטיית תקן $\frac{\sigma}{\sqrt{n}}$: $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

לכן, לפי משפט 6.13, הסכום של n המחוברים הראשונים מתפלג גם כן נורמלית $S_n = n\cdot\overline{X}_n \sim N\left(n\mu, n\sigma^2\right)$.

הוכחה. נניח לשם הפשטות כי $\mu = 0, \sigma = 1$. נחשב את הפונקציה היוצרת מומנטים של $\frac{S_n}{\sqrt{n}}$:

$$M_{\frac{S_n}{\sqrt{n}}}(t) = \mathbb{E}\left(e^{\frac{S_n t}{\sqrt{n}}}\right) = M_{S_n}\left(\frac{t}{\sqrt{n}}\right) = M_{X_1+\ldots+X_n}\left(\frac{t}{\sqrt{n}}\right) = \left(M_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

כאשר המעבר האחרון בוצע לפי תרגיל 7.38 בו מוכיחים שהפונקצה היוצרת מומנטים של סכום של משתנים מקריים בלתי תלויים היא מכפלת הפונקציות היוצרות מומנטים של המשתנים המקריים. במקרה זה, כל המשתנים המקריים שווי התפלגות ולכן יש להם אותה פונקציה יוצרת מומנטים. נקרב את הפונקציה היוצרת מומנטים באמצעות טור טיילור לפי 5.20:

$$M_{X_1}(t) = \sum_{i=0}^{\infty}\frac{\mathbb{E}\left(X_1^i\right)t^i}{i!} = 1 + \mathbb{E}\left(X_1\right)t + \mathbb{E}\left(X_1^2\right)\frac{t^2}{2} + \ldots = 1 + \frac{t^2}{2} + \ldots$$

נציב בנוסחא $\frac{t}{\sqrt{n}}$ ונקבל את הקירוב

$$M_{X_1}\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{t^2}{2n}$$

כלומר

$$M_{\frac{S_n}{\sqrt{n}}}(t) = \left(1 + \frac{t^2}{2n}\right)^n \xrightarrow[n\to\infty]{} e^{\frac{t^2}{2}}$$

זוהי הפונקציה היוצרת מומנטים של משתנה מקרי נורמלי סטנדרטי, לכן $\frac{S_n}{\sqrt{n}} \sim N(0,1)$. נכפול ב-$\sqrt{n}$ כדי לקבל $S_n \sim N(0,n)$. לחילופין, ניתן לחלק ב-$\sqrt{n}$ כדי לקבל $\overline{X}_n \sim N\left(0, \frac{1}{n}\right)$.

באופן כללי, כאשר המשתנים המקריים $X_1, X_2, \ldots$ מתפלגים עם תוחלת μ ושונות σ^2כלשהן, נגדיר $Y_i = \frac{X_i - \mu}{\sigma}$. לפי לינאריות התוחלת והשונות מתקיים

$$\mathbb{E}\left(Y_i\right) = \mathbb{E}\left(\frac{X_i-\mu}{\sigma}\right) = \frac{\mathbb{E}\left(X_i\right)-\mu}{\sigma} = 0$$

וכן

$$\mathbb{V}\left(Y_i\right) = \mathbb{V}\left(\frac{X_i-\mu}{\sigma}\right) = \mathbb{V}\left(\frac{X_i}{\sigma}\right) = \frac{1}{\sigma^2}\mathbb{V}\left(X_i\right) = 1$$

לפי ההוכחה הקודמת, $\overline{Y}_n \sim N\left(0, \frac{1}{n}\right)$ והקשר ביניהם הינו $\overline{X}_n = \sigma \overline{Y}_n + \mu$, כלומר $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.
□

טיב הקירוב של משפט הגבול המרכזי תלוי בהתפלגות של המשתנים המקריים $X_1, \ldots$. כאשר ההתפלגות מראש היא התפלגות "יפה" ודומה להתפלגות נורמלית (למשל, סימטרית), כך הקירוב מתקבל מהר יותר וניתן להשתמש במשפט הגבול המרכזי החל מ $n \geq 20$. עבור התפלגויות שאינן יפות הקירוב טוב החל מ $n \geq 100$.

בעיה 6.7. בדירה של דוד ויונתן מדליקים דוד חשמל בכל יום בזמן ששניהם מתקלחים. זמן המקלחת של דוד (בדקות) מתפלג מעריכית עם פרמטר $\frac{1}{10}$ וזמן המקלחת של יונתן מתפלג מעריכית עם פרמטר $\frac{1}{5}$, ללא תלות בזמן המקלחת של דוד או בימים האחרים. עלות הפעלת הדוד למשך שעה שלמה היא חצי ש"ח. מה ההסתברות שבמשך שנה שלמה שילמו השותפים יותר מ-50 ש"ח בעבור הפעלת הדוד?

הוכחה. זמן המקלחת הכולל הוא סכום של 365 משתנים מקריים המתארים את זמן המקלחת של דוד ועוד 365 משתנים מקריים המתארים את זמן המקלחת של יונתן. לא ניתן למצוא את ההתפלגות הישירה של זמן המקלחת לאורך השנה כולה ולכן נשתמש במשפט הגבול המרכזי בכדי לקרב זמן זה להתפלגות נורמלית. נסמן את זמן המקלחת של דוד ביום ה-$i \in \{1, \ldots, 365\}$ ב-$X_i \sim Exp\left(\frac{1}{10}\right)$, את זמן המקלחת של יונתן ב-$Y_i \sim Exp\left(\frac{1}{5}\right)$ ואת זמן הפעלת הדוד הכולל ביום זה ב-$W_i = X_i + Y_i$. לפי חיבוריות התוחלת והשונות (המשתנים המקריים בלתי תלויים) נקבל שזמן פעולת דוד החשמל ביום ה-i בעל תוחלת

$$\mathbb{E}(W_i) = \mathbb{E}(X_i + Y_i) = \mathbb{E}(X_i) + \mathbb{E}(Y_i) \overset{eq.\,6.35}{=} 10 + 5 = 15$$

ושונות

$$\mathbb{V}(W_i) = \mathbb{V}(X_i + Y_i) = \mathbb{V}(X_i) + \mathbb{V}(Y_i) \overset{eq.\,6.36}{=} 100 + 25 = 125$$

נסמן ב-$W = \sum_{i=1}^{365} W_i$ את זמן פעולת הדוד לאורך שנה שלמה (בדקות) וב-$T = 0.5 \cdot \frac{W}{60} = \frac{W}{120}$ את המחיר שהשותפים שילמו עבור הפעלת הדוד לאורך השנה כולה. ההסתברות שהשותפים שילמו לאורך השנה יותר מ-50 ש"ח הינה

$$Pr(T \geq 50) = Pr\left(\frac{W}{120} \geq 50\right) = Pr(W \geq 6000)$$

המשתנה המקרי W הוא סכום של 365 משתנים מקריים בלתי תלויים ולכן ניתן לחשב את ההסתברות האחרונה באמצעות קירוב למשפט הגבול המרכזי, שכן בקירוב $W \sim N(365 \cdot 15, 365 \cdot 125) = N(5475, 45625)$. נתקנן בכדי לעבור למשתנה מקרי נורמלי סטנדרטי ונקבל:

$$Pr(W \geq 6000) = Pr\left(\frac{W - 5475}{\sqrt{45625}} \geq \frac{6000 - 5475}{\sqrt{45625}}\right) = 1 - \Phi(2.46) = 0.007$$

□

בדוגמא האחרונה ראינו כיצד משתמשים במשפט הגבול המרכזי – במקום לחשב את ההתפלגות של המשתנה המקרי הרצוי, משתמשים בעובדה שניתן לרשום אותו כסכום של משתנים מקריים אשר את התוחלת והשונות שלהם אנחנו יודעים ומחשבים את ההסתברות של המאורע הדרוש על ידי התפלגות נורמלית. מה היה קורה בדוגמא זו אם זמן המקלחת בכל יום של שניהם ביחד היה מתפלג אחיד בדיד $W_i \sim U\{1, \ldots, 10\}$? במקרה כזה, גם W, סכום זמני המקלחת לאורך השנה, היה בדיד והמאורעות $\{W \geq 6000\}$ ו-$\{W > 5999\}$ היו זהים ובעלי אותה ההסתברות. אולם, אם נפעיל את משפט הגבול המרכזי על שני המאורעות נקבל תוצאה שונה שכן

התפלגות נורמלית היא התפלגות רציפה וישנה הסתברות חיובית להיות בתחום $[5999, 6000]$. בכדי לפתור בעיה זו, בה אותו המאורע מקבל שתי הסתברויות שונות כתלות באופן הצגת המאורע, הונהג תיקון רציפות. תיקון זה, המבוצע כאשר מקרבים משתנה מקרי בדיד באמצעות משפט הגבול המרכזי פותר את הבעיה ומוסיף דיוק מסויים לחישוב.

הגדרה 6.10 (תיקון רציפות). יהי X משתנה מקרי בדיד המקבל ערכים עוקבים ומקורב באמצעות משפט הגבול המרכזי. לכל $k \in \mathbb{Z}$ מתקיים $\{X > k-1\} = \{X \geq k\}$ אולם בעת הפעלת משפט הגבול המרכזי נקבל מאורעות שונים עם הסתברות שונה. לכן, תמיד נפעיל את משפט הגבול המרכזי על הערך הנמצא בחצי הדרך בין אי-השוויון החלש לאי-השוויון החזק. במקרה זה, כאשר נשתמש במשפט הגבול המרכזי נחשב את ההסתברות של המאורע $\{X > k-0.5\}$. באופן כללי עבור אי-שוויון מסוג זה נקבל:

$$Pr(X > k) = Pr(X \geq k+1) = Pr(X > k+0.5)$$

ועבור אי-שוויון הפוך נקבל

$$Pr(X < k) = Pr(X \leq k-1) = Pr(X < k-0.5)$$

עבור אי-שוויון דו-צדדי נקבל

$$Pr(a \leq X \leq b) = Pr(a-0.5 < X < b+0.5)$$

ובאופן זה ניתן גם לחשב הסתברויות מקורבות של שיוויונים:

$$Pr(X = k) = Pr(k \leq X \leq k) = Pr(k-0.5 < X < k+0.5)$$

כאשר הביטויים אותם נציב במשפט הגבול המרכזי הם הביטויים באגף ימין, עם גורם התיקון, 0.5.

אחד המשתנים המקריים הבדידים המרכזיים הוא משתנה מקרי בינומי וכחלק ממשפט **6.1** ראינו כי ניתן להציג משתנה מקרי בינומי כסכום של אינדיקטורים המתארים הצלחה בכל אחד מהסיבובים. לכן, ניתן לקרב משתנה מקרי בינומי על ידי התפלגות נורמלית. קירוב זה מאוד שימושי מאחר ולא ניתן לחשב הסתברויות של אי-שוויונים בצורה נוחה בהתפלגות בינומית. הסיבה לכך היא שאי-שוויון מצריך סכימה של הרבה אפשרויות מתוך התומך המתאימות לאי-השוויון ואין ברשותנו נוסחא פשוטה המתארת סכומים אלו.

***משפט* 6.17** (הקירוב הנורמלי להתפלגות בינומית). *יהי $X \sim Bin(n,p)$ כאשר $np \geq 5$ וגם $n(1-p) \geq 5$. אזי לפי משפט הגבול המרכזי ניתן לקרב את X לפי התפלגות נורמלית: $X \sim N(np, np(1-p))$. תוך שימוש בתיקון רציפות נקבל:*

$$Pr(X \leq k) = Pr(X < k+0.5) = \Phi\left(\frac{k+0.5-np}{\sqrt{np(1-p)}}\right) \tag{6.47}$$

$$Pr(X \geq k) = Pr(X > k-0.5) = 1-\Phi\left(\frac{k-0.5-np}{\sqrt{np(1-p)}}\right) \tag{6.48}$$

הוכחה. המשפט הוא תוצאה מיידית של משפט הגבול המרכזי שכן ניתן לרשום משתנה מקרי בינומי בתור הסכום $X = \sum_{i=1}^{n} \mathbf{1}_i$ כאשר $\mathbf{1}_i$ הוא משתנה מקרי מציין השווה ל-**1** כאשר הייתה הצלחה בניסוי ה-i ו-**0** אחרת. □

במקרים רבים, למשל כאשר n גדול או בשאלות שיש למצוא את n, ניתן להזניח את קירוב הרציפות בכדי לקבל משוואה פשוטה יותר לפתרון מבלי לאבד דיוק משמעותי.

בעיה 6.8 (המשך בעיה 5.7). נחזור לבעיה 5.7, שם חישבנו את מספר הסטודנטים שיש לדגום על מנת שבהסתברות של 0.8 לפחות, אחוז הסטודנטים להנדסה במדגם לא יסטה מהערך האמיתי ביותר מ-5% באמצעות חסם מרקוב וצ'בישב. נחשב ערך זה הפעם באמצעות משפט הגבול המרכזי.

הוכחה. מספר הסטודנטים להנדסה הנדגמים בדרך זו מתפלג $X \sim Bin\left(n,p\right)$ ולכן אחוז הסטודנטים להנדסה, כפי שמוגדר בשאלה, הינו $Y = \frac{X}{n}$. המאורע המבוקש בשאלה הינו $\{|Y-p| \le 0.05\}$ ועל ידי כפל ב-n נקבל את המאורע השקול $\{np-0.05n \le X \le np+0.05n\} = \{|X-np| \le 0.05n\}$. נציב זאת בקירוב הנורמלי להתפלגות בינומית ונקבל:

$$\begin{aligned} Pr\left(np-0.05n \le X \le np+0.05n\right) &= \Phi\left(\frac{np+0.05n+0.5-np}{\sqrt{np\left(1-p\right)}}\right) - \Phi\left(\frac{np-0.05n-0.5-np}{\sqrt{np\left(1-p\right)}}\right) \\ &= \Phi\left(\frac{0.05n+0.5}{\sqrt{np\left(1-p\right)}}\right) - \Phi\left(-\frac{0.05n+0.5}{\sqrt{np\left(1-p\right)}}\right) \\ &= 2\Phi\left(\frac{0.05n+0.5}{\sqrt{np\left(1-p\right)}}\right) - 1 \end{aligned}$$

הדרישה היא שהסתברות זו תהיה לפחות 0.8 כלומר $2\Phi\left(\frac{0.05n+0.5}{\sqrt{np(1-p)}}\right) - 1 \ge 0.8$ השקול ל- $\Phi\left(\frac{0.05n+0.5}{\sqrt{np(1-p)}}\right) \ge 0.9$. לפי הטבלה של התפלגות נורמלית סטנדרטית $\Phi\left(1.285\right) = 0.9$ ולכן נדרוש $\frac{0.05n+0.5}{\sqrt{np(1-p)}} \ge 1.285$. לשם הפשטות נזניח את תיקון הרציפות. זה לא ישנה בהרבה את התשובה הסופית אך ישפר משמעותית את יכולת הפתרון של משוואה זו. נקבל:

$$\sqrt{n} \ge \frac{1.285}{0.05}\sqrt{p\left(1-p\right)} = 25.7\sqrt{p\left(1-p\right)}$$

בבעיה נתון היה כי $p \in [0.1, 0.3]$ ובתחום זה המקסימום של הפונקציה $p\left(1-p\right)$ מתקבל ב-$p = 0.3$. נציב ערך זה ונקבל

$$n \ge \left(25.7 \cdot \sqrt{0.3 \cdot 0.7}\right)^2 = 138.7$$

נבחר את הערך השלם הבא בתור ונקבל שיש לדגום לניסוי זה לפחות $n = 139$ סטודנטים בכדי שאחוז הסטודנטים להנדסה במדגם לא יסטה מהערך האמיתי ביותר מ-0.05 בהסתברות גבוהה מ-0.8. $\square$

תרגיל 6.26. כדי לבחון האם מטבע הוא הוגן מטילים אותו 100 פעמים וסופרים את מספר העצים שהתקבלו. אם מספר העצים סוטה מהתוחלת הצפויה ביותר מ-10, המטבע מוכרז כלא הוגן. מה ההסתברות שמטבע הוגן יוכרז כלא הוגן?

תרגיל 6.27. הפידיון היומי של דוכן קפה הינו משתנה מקרי עם תוחלת 2000 ש"ח ושונות 1000. בעל הדוכן סופר את ההכנסות מיום פתיחת הדוכן ומחשב את ממוצע ההכנסות היומי מיום זה.

1. מה הסיכוי שלאחר 100 ימים הממוצע היומי קטן מ-2010?

2. האם קיים יום אשר לאחריו בהסתברות של 0.99 לפחות ממוצע ההכנסות היומי מעל 1999? אם כן – מהו יום זה?

בתרגיל 6.27 בשני הסעיפים התקבלו הסתברויות גבוהות מ-0.99, אך בסעיף הראשון תוצאה זו התקבלה מאוד מהר, אחרי 100 ימים, ואילו בסעיף השני זה קרה רק אחרי הרבה מאוד זמן (כמעט 15 שנים...). ההבדל בין התוצאות נעוץ בכך שדרשנו מרחקים שונים מהתוחלת. בסעיף הראשון הסתפקנו בקירבה של 10 לתוחלת ואילו בסעיף השני הדרישה היא קירבה של 1 בלבד. ככל שמבקשים לקבל תוצאה קרובה יותר לתוחלת, כך יש לבצע יותר מדידות שכן סטיית התקן של הממוצע ביחס לתוחלת הולכת וקטנה כמו $\frac{1}{\sqrt{n}}$. זו דעיכה יחסית איטית – אם רוצים להקטין את השגיאה פי 2, יש לבצע פי 4 יותר מדידות ואם, כמו בדוגמא זו, מבקשים להקטין את השגיאה פי 10, מספר המדידות גדל בערך פי 100. יש לקחת בחשבון תוצאה זו כאשר מתכננים ניסוי שכן הדיוק שניתן להשיג בדרך זו מוגבל על ידי מספר המדידות אותן ניתן לבצע כחלק מהממוצע.

6.3 שאלות מסכמות

6.3.1 משתנים מקריים בדידים

שלבים שכדאי לבצע בעת פתרון שאלה על משתנים מקרים בדידים:

1. חישוב התומך באופן מלא או באופן חלקי (האם התומך סופי או אינסופי? האם הוא כולל ערכים עוקבים או לא וכד'). בהתאם לתומך ניתן לפסול משתנים מקריים בעלי תומך לא מתאים (למשל, כאשר התומך סופי, בוודאות המשתנה המקרי אינו גאומטרי).
2. לבדוק האם הסיפור המתאר את המשתנה המקרי מתאים לאחד המשתנים המקריים הנותרים באופן מלא או שניתן להגדיר משתנה עזר שיתאים (כפי שנעשה למשל בבעיה 6.2).
3. האם ההתפלגות בכלל נדרשת? למשל, אם השאלה עוסקת רק בתוחלת ושונות, אולי ניתן לחשב אותם ללא חישוב ההתפלגות או זיהוי המשתנה המקרי, למשל באמצעות חיבוריות. אם השאלה כוללת סכימה או ממוצע, ייתכן ומספיק להשתמש במשפט הגבול המרכזי בכדי לקבל קירוב טוב להסתברות מבלי חישוב ההסתברות עצמה.

תרגיל 6.28. שחקן מהמר באקראי על אחד מהמספרים $1-6$ ואז מטיל 3 קוביות. הוא מקבל $i\in\{1,2,3\}$ שקלים אם המספר עליו הימר הופיע i פעמים ומשלם שקל בודד אם המספר לא הופיע כלל.

1. מהי התפלגות הרווח של השחקן?
2. מהי תוחלת הפרס?

תרגיל 6.29. שיכור מגיע לביתו הנעול. השיכור בוחר באקראי מפתחות מתוך צרור הכולל 10 מפתחות עד אשר הוא מצליח לפתוח את הדלת. כיצד מתפלג מספר הניסיונות כאשר

1. השיכור מספיק פיקח כדי לזכור אילו מפתחות הוא כבר ניסה ולא לנסות אותם שנית.
2. השיכור לא זוכר אילו מפתחות כבר ניסה ובכל פעם בוחר באקראי מפתח מאלו שברשותו.

תרגיל 6.30. מטילים מטבע הוגן n פעמים. מה ההסתברות ש"עץ" הופיע מספר זוגי של פעמים בסדרת הטלות זו?

תרגיל 6.31. מטבע לא הוגן הנותן עץ בהסתברות p מוטל שוב ושוב עד אשר מתקבל עץ לראשונה. מה ההסתברות שהניסוי נעצר לאחר מספר זוגי של הטלות?
הערה: שאלה זו זהה לתרגיל 2.15. פיתרו אותו הפעם על ידי זיהוי המשתנה המקרי המופיע בשאלה.

תרגיל 6.32. נתון לוח משבצות 3×3. צובעים באקראי כל משבצת בשחור או לבן באופן בלתי תלוי במשבצות האחרות.

1. איך מתפלג מספר המשבצות הלבנות על הלוח?
2. איך מתפלג מספר השורות בהן יש בדיוק שתי משבצות לבנות?
3. בהינתן שיש משבצת אחת לבנה בכל הלוח, כיצד מתפלג מספר השורה בו משבצת זו נמצאת?
4. בהינתן שיש 3 משבצות לבנות בכל הלוח. כיצד מתפלג מספר המשבצות הלבנות בשורה העליונה?

תרגיל 6.33. בהגרלת הלוטו הישראלי (ראו תרגיל 3.5) משתתפים מיליון מתושבי ישראל. מה ההסתברות שלא יזכה אף אחד? ידוע שיש לפחות זוכה אחד, מה הסיכוי שהוא הזוכה היחיד?
הניחו שכל אדם שולח ניחוש אחד באופן בלתי תלוי באנשים האחרים.

תרגיל 6.34. יהי $X\sim Bin\left(n,p\right)$ ונסמן $\lambda=np$. חשבו את $\lim_{n\to\infty}M_X\left(t\right)$ כאשר λ קבוע. מה קיבלתם? מדוע?

תרגיל 6.35. בתרגיל הבא נבחן כיצד ניתן לקבל משתנים מקריים באמצעות אנלוגיות לסידורי כדורים. בכל אחד מהסעיפים המציאו תהליך אקראי והגדירו משתנה מקרי מתאים בהתאם לדרישה:

1. הראו כיצד ניתן ליצור משתנה מקרי המתפלג אחיד $U\{1, 2, ..., n\}$ על ידי סידור אקראי של כדורים בשורה.
2. הראו כיצד ניתן ליצור משתנה מקרי המתפלג $Bin(n, p)$ לכל $p \in \mathbb{Q}$ (רציונלי) על ידי חלוקה אקראית של כדורים לתאים.
3. הראו כיצד ניתן ליצור משתנה מקרי המתפלג $Geom(p)$ לכל $p \in \mathbb{Q}$ (רציונלי) על ידי הוצאת כדורים מכד.

6.3.2 משתנים מקריים רציפים

תרגיל 6.36. בבנק אוצר הסטודנט בוחנים את חשבונות הלקוחות ואת מדיניות הריבית. הם מעריכים כי יתרת החשבון של לקוח מתפלגת נורמלית עם $\mu = 2200$ ו-$\sigma = 600$ (ש"ח).

1. לאיזה אחוז מהלקוחות יש יתרה של למעלה מ-3000 ש"ח? לאיזה אחוז יתרה של פחות מ-1400 ש"ח? איזה אחוז מהלקוחות במינוס?
2. הבנק מעוניין להציע ריבית גבוהה יותר ל-5% מהלקוחות - בעלי היתרה הגבוהה ביותר. מהי היתרה המינימלית הנדרשת כדי לקבל את הריבית המיוחדת?
3. בבוקר יום א' נמצאים בבנק 30 לקוחות. מהי תוחלת מספר הלקוחות בעלי הריבית המיוחדת מהסעיף הקודם?

תרגיל 6.37. יהי X משתנה מקרי רציף המתפלג $U(1, 3)$. נגדיר $Y = \ln X$.

1. כיצד מתפלג Y?
2. מהי התוחלת והשונות של Y?
3. המ"מ $X_1, X_2, ..., X_{20}$ מוגרלים באקראי ובאופן בלתי תלוי לפי ההתפלגות של X. חשבו בקירוב את ההסתברות ש- $X_1 \cdot X_2 \cdot \ldots \cdot X_{20} > 1500000$?

תרגיל 6.38. אילון מתלבט האם לצאת לחופשה באיטליה או בצרפת. בכדי לבחור הוא מבצע הגרלה אך מאחר והוא מעדיף מעט יותר את צרפת על פני איטליה, הוא מעוניין שהסיכוי לבחור את איטליה בהגרלה יהיה $\frac{1}{\pi}$. ברשותו של אילון מצוי רק מטבע הוגן ובכדי לממש את ההגרלה הזו הוא מטיל את המטבע 1000 פעמים וסופר את מספר הפעמים שהמטבע נפל על עץ. כיצד על סמך תוצאות של ניסוי זה יכול אילון להחליט בין איטליה וצרפת כך שהסיכוי לבחור באיטליה יהיה אכן $\frac{1}{\pi}$?

תרגיל 6.39. תוכנות מחשב רבות יודעות להגריל מספרים אקראיים המתפלגים אחיד רציף או נורמלי. בשאלה זו נראה כיצד ניתן באמצעות הגרלות אלו להגריל משתנים מקריים אחרים. נדגים את הטכניקה באמצעות משתנה מקרי מעריכי.

1. יהי X משתנה מקרי רציף עם פונקציית התפלגות מצטברת **הפיכה** F_X ו-U משתנה מקרי אחיד רציף בקטע $(0, 1)$. הוכיחו שהמ"מ המקרי $Y = F^{-1}(U)$ מתפלג כמו X, כאשר F^{-1} היא ההופכית של F_X.

   הדרכה: הראו שפונקציית ההתפלגות המצטברת של Y היא גם F_X.
2. הראו כי הפונקציות $f(x) = -\frac{1}{\lambda}\ln(1-x)$ ו-$g(x) = 1-e^{-\lambda x}$ הן הופכיות זו לזו, קרי $f(g(x)) = x$ לכל $x \in \mathbb{R}$.
3. על סמך הסעיפים הקודמים, הסיקו שהמשתנה המקרי $X = f(U)$ הוא משתנה מקרי מעריכי עם פרמטר λ.

4. **סכמו: אם ברשותכם מדגם $x_1, x_2, ..., x_n$ שנלקח מהתפלגות אחידה בקטע $(0,1)$, כיצד ניתן ליצור ממנו מדגם מהתפלגות מעריכית עם פרמטר λ?**

תרגיל 6.40. יהי X משתנה מקרי בעל צפיפות רציפה הנתמך בקטע סופי $[a,b]$ (קרי, הצפיפות היא 0 מחוץ לקטע). הוכיחו כי אם קיימת פונקציה g כך שלכל $a \leq c \leq d \leq b$ מתקיים $Pr\left(c \leq X \leq d\right) = g(d-c)$ אז X הוא משתנה מקרי אחיד בקטע $[a,b]$.

6.4 פתרונות לשאלות הפרק

תרגיל 6.1 . לשם הפשטות נסתכל על $Y \sim U\left\{0, \ldots, N\right\}$.

$$\mathbb{E}\left(Y\right) = \sum_{k=0}^{N} k \cdot P_Y\left(k\right) = \sum_{k=0}^{N} \frac{k}{N+1} \overset{eq.\ 0.20}{=} \frac{N\left(N+1\right)}{2\left(N+1\right)} = \frac{N}{2}$$

לצורך חישוב השונות נחשב את התוחלת של Y^2:

$$\begin{aligned}\mathbb{E}\left(Y^2\right) &= \sum_{k=0}^{N} k^2 \cdot P_Y\left(k\right) = \sum_{k=0}^{N} \frac{k^2}{N+1} \overset{eq.\ 0.21}{=} \frac{N(N+1)(2N+1)}{6\left(N+1\right)} \\ &= \frac{N(2N+1)}{6}\end{aligned}$$

ולכן השונות הינה

$$\begin{aligned}\mathbb{V}\left(Y\right) &= \mathbb{E}\left(Y^2\right) - \mathbb{E}^2\left(Y\right) = \frac{N(2N+1)}{6} - \frac{N^2}{4} = \frac{2N(2N+1) - 3N^2}{12} \\ &= \frac{4N^2 + 2N - 3N^2}{12} = \frac{N^2 + 2N}{12}\end{aligned}$$

לסיום, הפונקציה יוצרת המומנטים הינה

$$\begin{aligned}M_X\left(t\right) &= \mathbb{E}\left(e^{tX}\right) = \sum_{k=0}^{N} \frac{e^{tk}}{N+1} = \frac{1}{N+1} \overset{eq.\ 0.14}{=} \frac{1}{N+1} \cdot \frac{1 - \left(e^t\right)^{N+1}}{1 - e^t} \\ &= \frac{1}{N+1} \cdot \frac{1 - e^{(N+1)t}}{1 - e^t}\end{aligned}$$

אם $X \sim U\{a, \ldots, b\}$ אזי $X - a \sim U\left\{0, \ldots, b-a\right\}$. נסמן $Y = X - a$ ו-$N = b - a$ ונשתמש בנוסחאות שפיתחנו לעיל. לפי לינאריות התוחלת נקבל:

$$\begin{aligned}\mathbb{E}\left(X\right) &= \mathbb{E}\left(Y + a\right) = \mathbb{E}\left(Y\right) + a \\ &= \frac{N}{2} + a = \frac{b-a}{2} + a = \frac{b+a}{2}\end{aligned}$$

ולפי לינאריות השונות:

$$\begin{aligned}\mathbb{V}\left(X\right) &= \mathbb{V}\left(Y + a\right) = \mathbb{V}\left(Y\right) = \frac{N^2 + 2N}{12} \\ &= \frac{\left(N+1\right)^2 - 1}{12} = \frac{\left(b - a + 1\right)^2 - 1}{12}\end{aligned}$$

לפי תרגיל 5.24:

$$M_X(t) = M_{Y+a}(t) = e^{at}M_X(t) = e^{at} \cdot \frac{1}{N+1} \cdot \frac{1-e^{(N+1)t}}{1-e^t}$$
$$= \frac{e^{at}}{b-a+1} \cdot \frac{1-e^{(b-a+1)t}}{1-e^t} = \frac{e^{at}-e^{(b+1)t}}{(b-a+1)(1-e^t)}$$

□

תרגיל 6.2. **1.** לפי הנתון, ההסתברות של כל צעד להיות צעד ימינה היא $p=\frac{1}{2}$ ובפרט ההסתברות שהצעד האחרון הוא צעד ימינה היא גם $p=\frac{1}{2}$.

2. לאחר 100 צעדים השיכור הגיע לנקודה $X=10$. משמעות הדבר היא שהוא בהכרח צעד $Y=55$ צעדים ימינה ואת היתר שמאלה. יש $\binom{100}{55}$ דרכים שונות בהן מאורע זה יכל היה להתרחש מבחינת סדר הצעדים ימינה ושמאלה אותם ביצע השיכור, ולכל סידור כזה יש את אותה הסתברות להתקבל. לכן, הבעיה שקולה לסידור 100 צעדים בשורה באקראי, מתוכם 55 צעדים מסוג "ימין" ו-45 צעדים מסוג "שמאל". הסיכוי שבמקום האחרון (ולמעשה בכל מקום) עומד צעד ימינה הוא $\frac{55}{100}$ וזו ההסתברות המבוקשת.
ניתן לחשב הסתברות זו גם בצורה ישירה תוך שימוש בנוסחאות של הסתברות מותנית והתפלגות בינומית.
נסמן ב-A את המאורע שהצעד האחרון היה ימינה. לכן ההסתברות המבוקשת הינה

$$Pr(A|X=10) = \frac{Pr(A\cap\{X=10\})}{Pr(X=10)}$$

המאורעות A ו-$\{X=10\}$ תלויים ולכן כדי לטפל במונה נשנה את המאורע הדרוש לחיתוך של מאורעות בלתי תלויים. כדי שהשיכור יגיע ל-$X=10$ וגם יצעד את הצעד האחרון שמאלה, בהכרח אחרי 99 צעדים הוא הגיע ל-9. לכן נסמן ב-Z את מספר הצעדים ימינה שהשיכור ביצע ב-99 הצעדים הראשונים. על מנת שאחרי 99 צעדים הוא יהיה בנקודה 9, עליו לבצע 54 צעדים ימינה ו-45 צעדים שמאלה (בדומה לקשר בין X ל-Y) ולכן $A\cap\{X=10\} = A\cap\{Z=54\}$. הצעד ה-100 בלתי תלוי בכל הצעדים הקודמים ולכן מאורעות אלו בלתי תלויים

$$\frac{Pr(A\cap\{Z=54\})}{Pr(X=10)} = \frac{Pr(A)\cdot Pr(Z=54)}{Pr(X=10)} = \frac{\frac{1}{2}\cdot\binom{99}{54}\cdot\left(\frac{1}{2}\right)^{54}\left(\frac{1}{2}\right)^{45}}{\binom{100}{55}\cdot\left(\frac{1}{2}\right)^{55}\left(\frac{1}{2}\right)^{45}}$$
$$= \frac{\binom{99}{54}}{\binom{100}{55}} = \frac{\frac{99!}{54!45!}}{\frac{100!}{55!45!}} = \frac{55}{100}$$

כפי שקיבלנו קודם.

3. במקרה שהשיכור מבצע את הצעדים בסיכוי p כלשהו ימינה ולאו דווקא חצי, הסיכוי שהצעד האחרון ימינה יהיה p וזו התשובה לסעיף 1. לגבי סעיף 2 – התוצאה שקיבלנו כלל לא תלויה בהסתברות לצעוד ימינה ולכן גם אם הסיכוי לצעוד ימינה הוא p כלשהו, בהינתן שהוא הגיע ל-$X=10$ אחרי 100 צעדים, הסיכוי שהצעד האחרון היה צעד ימינה הינו $\frac{55}{100}$. ההסתברות הפריורית לצעד ימינה לא משפיעה על ההסתברות הפוסטיריות לצעוד צעד ימינה, בהינתן $X=10$. הסיבה היא שהמאורע $X=10$ כבר קרה ולא משנה מה הייתה הסתברותו הראשונית. ניתן להבהיר נקודה זו עם דוגמא פשוטה: מתוך כד שבו מיליון כדורים לבנים, כדור אדום וכדור שחור מוציאים כד באקראי. בהינתן שיצא כדור שאינו לבן, מה

הסיכוי שהוא שחור? ובכן, מראש הסיכוי להוציא כדור שאינו לבן הוא מאוד נמוך, אבל בהינתן שהכדור לא לבן הוא יכול להיות רק אחד משני הכדורים האחרים בהסתברויות שוות ולכן הסיכוי הוא $\frac{1}{2}$. תשובה זו תישאר בעינה גם אם בכד יש מיליארד כדורים לבנים או כדור לבן בודד – מרחב המדגם הצטמצם לכדי מה שקרה בפועל ולא מה שיכל היה לקרות מראש בניסוי.

□

תרגיל 6.3. נשים לב שכמות הטעויות בשליחה של ביט בודד מתפלגת $X \sim Bin(3, 0.1)$ ולכן ההסתברות שתהיה שגיאה בפענוח ביט בודד היא $Pr(X \geq 2) = \binom{3}{2}0.1^2 0.9 + 0.1^3 = 0.028 < 0.1$. השגיאה לטעות בביט בודד ירדה פי למעלה מ-3!

1. מספר השגיאות במילה באורך 8 ביטים מתפלגת $Y \sim Bin(8, 0.028)$ ולכן המילה תפוענח נכון רק כאשר לא יהיו שגיאות באף אחד מהביטים, כלומר בהסתברות $Pr(Y = 0) = (1-0.028)^8 = 0.796$.
2. לאור הסעיף הקודם, ההסתברות לשגיאה בפענוח מילה בודדת היא $1 - 0.796 = 0.203$ ולכן מספר המילים שפוענחו באופן שגוי בקובץ המכיל **1000** מילים מתפלג $Bin(1000, 0.203)$.

□

תרגיל 6.4. אם ננסה להסתכל על הניסוי באופן בו הוא מתואר ונסתכל על התקדמות הניסוי בפועל – נסתבך. כדי לתאר את המאורע שאחרי 2 הטלות נותרו 2 מטבעות, נצטרך להתחשב בכל האפשרויות בהן מאורע זה קרה: בסיבוב הראשון נפסלו 3 מטבעות ו-2 הנותרים נשארו בסיבוב השלישי, בסיבוב הראשון נפסלו 2 מטבעות ומטבע נוסף נפסל בסיבוב השני וכן הלאה. במקום זאת, נסתכל על כל מטבע כעל ניסוי באופן הבא: ההסתברות שמטבע מסוים ישרוד n סיבובים היא $\left(\frac{1}{2}\right)^n$. כל מטבע שורד סיבובים אלו באופן בלתי תלוי במטבעות האחרים ולכן מספר המטבעות ששרדו n סיבובים מתפלגת $X_n \sim Bin\left(5, \left(\frac{1}{2}\right)^n\right)$. לפיכך:

1. $Pr\left(X_2 = 2\right) = \binom{5}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3 = 0.2636$.
2. $Pr\left(X_5 = 3\right) = \binom{5}{3}\left(\frac{1}{2^5}\right)^3\left(1 - \frac{1}{2^5}\right)^2 = 0.000286$.
3. $Pr\left(X_4 = 0\right) = \binom{5}{0}\left(\frac{1}{2^4}\right)^0\left(1 - \frac{1}{2^4}\right)^5 = 0.724$.

□

תרגיל 6.5. **1.** מספר הזכיות בכל אחד מבתי הקזינו מתפלג בינומית כאשר השוני הוא רק בהסתברות לזכייה. אצל ואל-ז'אן מספר הזכיות מתפלג $Bin\left(10, \frac{1}{2}\right)$ ואילו אצל טנרדייה $Bin\left(10, \frac{4}{10}\right)$.

2. נגדיר X_i – הרווח בסיבוב ה-i. משתנה מקרי זה מתפלג לפי

$$X_i = \begin{cases} 1 & p \\ -1 & 1-p \end{cases}$$

כאשר ההסתברות לזכייה אצל ואל-ז'אן היא $p = \frac{1}{2}$ וההסתברות לזכייה אצל טנרדייה היא $p = \frac{4}{10}$. לכן תוחלת הרווח בכל סיבוב הינה

$$\mathbb{E}\left(X_i\right) = 1 \cdot p + (-1) \cdot (1-p) = 2p - 1 = \begin{cases} 0 & p = \frac{1}{2} \\ -\frac{1}{5} & p = \frac{4}{10} \end{cases}$$

סך הזכיה לאחר **10** סיבובים היא סכום הזכיות בכל אחד מהסיבובים ומחיבוריות התוחלת נקבל שאצל ואל-זאן תוחלת הרווח אחרי **10** סיבובים היא **0** פרנקים ואילו אצל טנרדייה תוחלת הרווח אחרי **10** סיבובים היא -2 פרנקים, כך שאצל טנרדייה המהמר בממוצע מפסיד.

3. נסמן ב-V את המאורע שהמהמר נכנס לקזינו של ואל-ז'אן וב-X את הרווח שלו אחרי 10 סיבובים. נשתמש בנוסחת בייס ונקבל:

$$Pr\left(V|X=8\right)=\frac{Pr\left(X=8|V\right)Pr\left(V\right)}{Pr\left(X=8\right)}$$

לצורך קבלת רווח של 8 פרנקים יש צורך ב-9 זכיות והפסד אחד. בהינתן שהמהמר נמצא אצל ואל-ז'אן, מספר הזכיות מתפלג $Bin\left(10,0.5\right)$ ולכן 9 זכיות יתקבלו בהסתברות

$$Pr\left(X=8|V\right)=\binom{10}{9}\left(\frac{1}{2}\right)^{9}\left(\frac{1}{2}\right)=0.00978$$

את המכנה נחשב באמצעות נוסחת ההסתברות השלמה כאשר מתנים על הקזינו אליו נכנס – אל ואן-ז'אל או טנרדייה:

$$\begin{aligned}Pr\left(X=8\right)&=Pr\left(X=8|V\right)Pr\left(V\right)+Pr\left(X=8|\overline{V}\right)Pr\left(\overline{V}\right)\\&=\frac{1}{2}\cdot0.00978+\frac{1}{2}\cdot\binom{10}{9}\left(\frac{4}{10}\right)^{9}\left(\frac{6}{10}\right)=0.00567\end{aligned}$$

נציב ערכים אלו בחזרה בנוסחת בייס ונקבל

$$Pr\left(V|X=8\right)=\frac{0.00978\cdot0.5}{0.00567}=0.861$$

כצפוי, קיבלנו הסתברות גבוהה. הסיבה לכך היא שבקזינו של ואל-ז'אן המהמר בממוצע לא מרוויח ולא מפסיד ובקזינו של טנדרייה הוא בעיקר מפסיד. לכן, אם המהמר הרוויח, סביר להניח שהוא בקזינו ההוגן ולא בקזינו הפועל לרעתו, שם הסיכוי לרווח קטן יותר.

□

תרגיל 6.6. עבור שני הסעיפים נחשב ראשית את ההסתברות לעבור את המבחן. ישנן 5 שאלות, הסיכוי לענות על כל שאלה הוא $\frac{1}{4}$ ללא תלות בשאלות האחרות ולכן מספר התשובות הנכונות מתפלג $X\sim Bin\left(5,\frac{1}{4}\right)$. ציון עובר מתקבל כאשר עונים על לפחות 3 שאלות לכן הסיכוי לציון עובר הינו

$$p=Pr\left(X\geq3\right)=\sum_{k=3}^{5}\binom{5}{k}\cdot\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{5-k}=0.1$$

1. בכיתה 50 תלמידים בלתי תלויים, לכל אחד סיכוי של $p=0.1$ לעבור את הבחינה ולכן מספר הסטודנטים שעוברים את הבחינה מתפלג בינומית $Bin\left(50,0.1\right)$.
2. מספר המבחנים שיש להיבחן עד הפעם הראשונה בה עוברים את המבחן, בהנחה שסיכוי המעבר בכל מבחן הינו $p=0.1$, מתפלג גאומטרית $Geom\left(0.1\right)$.

□

תרגיל 6.7. נניח שבהטלה הראשונה יצא צד "א", יהא צד זה אשר יהא. מעכשיו והלאה, כל הטלה בה יוצא צד "א" תחשב לכישלון ואילו הפעם הראשונה בה תתרחש הצלחה, כלומר המטבע יפול על הצד השני, היא הפעם הראשונה בה ראינו אני שני הצדדים של המטבע ומספר ההטלות עד אליה הוא מספר ההטלות המבוקש. לכן,

מספר ההטלות בין ההטלה הראשונה לזו בה יצא הצד השני של המטבע מתפלג $X\sim Geom\left(\frac{1}{2}\right)$ ואילו מספר ההטלות המבוקש הינו $Y=X+1$. מאחר ו-Y לא מתפלג לפי אחת המשפחות הבדידות המוכרות, נרשום במפורש את פונקציית ההסתברות שלו:

$$\begin{aligned}Pr\left(Y=k\right)&=Pr\left(X+1=k\right)=Pr\left(X=k-1\right)\\&=\left(1-\frac{1}{2}\right)^{k-1-1}\frac{1}{2}=\left(\frac{1}{2}\right)^{k-1}\end{aligned}$$

כאשר התומך הינו $k\in\{2,3,\ldots\}$, שכן נחוצות לפחות שתי הטלות בכדי לקבל את שני הצדדים של המטבע.

□

תרגיל 6.8. עבור שני הסעיפים, נחשב את ההסתברות שאחד המטבעות יפול על עץ והשני על פלי. מצב זה אפשרי בשתי דרכים זרות: הראשון על עץ והשני על פלי או להפך. המטבעות בלתי תלויים לכן ההסתברות שהראשון יפול על עץ והשני על פלי היא מכפלת ההסתברויות והסיכוי הכולל שמטבע אחד בדיוק יפול על עץ והשני על פלי יהא:

$$r=p\left(1-q\right)+q\left(1-p\right)$$

1. בכל הטלה יש סיכוי של r שאחד המטבעות יפול על עץ והמטבע השני על פלי, וסיכוי זה לא תלוי בהטלות הקודמות. לכן, אנחנו מבצעים ניסוי חוזר עם סיכוי קבוע להצלחה ומספר החזרות עד ההצלחה הראשונה בה הניסוי מסתיים מתפלג $Geom\left(r\right)$.
2. נסמן ב-A את המאורע שאחד המטבעות נפל על עץ והשני על פלי וב-B את המאורע שהמטבע הראשון הוא זה שנפל על עץ והשני נפל על פלי. בבירור, $B\subseteq A$. לפיכך, ההסתברות המבוקשת הינה

$$\begin{aligned}Pr\left(B|A\right)&\overset{eq.\,2.9}{=}\frac{Pr\left(B\cap A\right)}{Pr\left(A\right)}\overset{(*)}{=}\frac{Pr\left(B\right)}{Pr\left(A\right)}\\&=\frac{p\left(1-q\right)}{p\left(1-q\right)+q\left(1-p\right)}\end{aligned}$$

כאשר מעבר * נובע מכך ש-B מוכל ב-A ולכן $A\cap B=B$.

□

תרגיל 6.9. נחשב את הפונקציה היוצרת מומנטים לפי ההגדרה:

$$\begin{aligned}M_X\left(t\right)&=\mathbb{E}\left(e^{tX}\right)=\sum_{k=1}^{\infty}e^{tk}\cdot\left(1-p\right)^{k-1}\cdot p\\&=pe^t\cdot\sum_{k=1}^{\infty}e^{t(k-1)}\left(1-p\right)^{k-1}=pe^t\cdot\sum_{k=0}^{\infty}\left(e^t\left(1-p\right)\right)^k\\(eq.\,0.15)&=\frac{p\cdot e^t}{1-\left(1-p\right)\cdot e^t}\end{aligned}$$

כנדרש. □

תרגיל 6.10. נגדיר X – מספר ההטלות עד אשר כל הפאות התקבלו פעם אחת. בשלב הראשון נבדוק האם X מתפלג לפי התפלגות מוכרת כלשהי. התומך של X הינו $\{6,7,\ldots\}$ שכן בשביל לקבל את כל הפאות שעל הקוביה יש להטילה לפחות 6 פעמים אך ברמת העקרון מספר ההטלות יכול להיות כל מספר טבעי שכן יתכנו רצפים ארוכים מאוד בהם אחת הפאות לא יוצאת. תומך זה לא מתאים לאף אחת מההתפלגויות המופיעות בפרק ובנוסף הסיפור אינו מתאים לסיפורים המתאימים. אמנם יש כאן מרכיב גאומטרי, שכן חוזרים על ניסוי שוב ושוב עד שמשהו יקרה, אבל אותו משהו לא יכול להיות מוגדר כ"הצלחה" שכן הוא לא יכול להתקבל בכל סיבוב וההסתברות לקבל אותו משתנה מהטלה להטלה. לכן, משתנה מקרי זה אינו חלק ממשפחה מוכרת. ניתן לנסות לחשב את ההתפלגות של X מפורשות, אך היות וכל שנדרשנו לחשב הוא התוחלת והשונות שלו, נחשב אותם מבלי לחשב מפורשות את ההתפלגות, על ידי רישום X כסכום של משתנים מקריים אחרים ופשוטים יותר.

נגדיר X_1 בתור מספר ההטלות מתחילת הניסוי ועד שיצא מספר שלא יצא קודם ("מספר חדש"), X_2 בתור מספר ההטלות מהמספר החדש הראשון ועד המספר החדש השני, X_3 מספר ההטלות מהמספר החדש השני ועד המספר החדש השלישי וכן הלאה, עד ל-X_6 המוגדר להיות מספר ההטלות בין המספר החדש החמישי לבין המספר החדש השישי. כאשר יוצא המספר החדש השישי למעשה התקבלו כבר כל המספרים המופיעים על הקוביה, לכן מספר ההטלות הכולל הינו $X=\sum_{i=1}^{6}X_i$ ולכן מחיבוריות התוחלת (נוסחא 5.7) נקבל $\mathbb{E}\left(X\right)=\sum_{i=1}^{6}\mathbb{E}\left(X_i\right)$.

נותר לחשב את ההתפלגות של כל אחד מה-X_i-ים. תמיד ההטלה הראשונה היא מספר שלא היה אף פעם בעבר, לכן תמיד $X_1=1$. נניח שבהטלה הראשונה התקבלה התוצאה "5". כל אחת מהתוצאות 1-4,6 תהיה תוצאה חדשה כך שמספר ההטלות עד שיצא המספר החדש השני הוא גאומטרי $X_2\sim Geom\left(\frac{5}{6}\right)$. נניח שהמספר החדש הבא הוא "6". המספר החדש הבא יכול להיות רק אחד מארבעת המספרים 1-4 והסיכוי בכל הטלה לקבל את המספר החדש השלישי הינו $\frac{4}{6}$. לכן $X_3\sim Geom\left(\frac{4}{6}\right)$. באופן דומה, $X_4\sim Geom\left(\frac{3}{6}\right)$, $X_5\sim Geom\left(\frac{2}{6}\right)$ ו-$X_6\sim Geom\left(\frac{1}{6}\right)$. נציב את התוחלת של כל אחד מהם לפי נוסחא 6.14 והתוחלת הכוללת תהא:

$$\mathbb{E}\left(X\right)=\sum_{i=1}^{6}\frac{6}{i}=14.7$$

משתנים מקריים אלו בלתי תלויים שכן מספר ההטלות עד שהתקבל המספר החדש השלישי, למשל, לא משפיע על הסיכוי לקבל אחריו מספר חדש רביעי. ניתן אם כך לחשב את השונות על ידי שימוש בתכונת החיבוריות (נוסחא 6.15):

$$\mathbb{V}\left(X\right)=\mathbb{V}\left(\sum_{i=1}^{6}X_i\right)=\sum_{i=1}^{6}\mathbb{V}\left(X_i\right)=\sum_{i=1}^{6}\frac{1-\frac{i}{6}}{\left(\frac{i}{6}\right)^2}=38.99$$

אמנם מפתה להגדיר את המשתנים המקריים באופן הבא: X_i – מספר ההטלות מתחילת הניסוי ועד שהופיע המספר i (ואז כל אחד מהם גאומטרי עם פרמטר $\frac{1}{6}$) אך זה יהיה שגוי שכן במקרה כזה $X\neq\sum_{i=1}^{6}X_i$. הסיבה לכך היא שההטלה הראשונה, למשל, תיספר בתוך כל אחד מהמשתנים המקריים ולכן תיספר 6 פעמים ולא פעם אחת כדרוש. אם יוגדרו המשתנים המקריים באופן זה, הרי שיתקיים $X=\max\left\{X_i\right\}$ ולא ניתן יהיה להשתמש בחיבוריות של התוחלת והשונות לצורך חישוב מספר ההטלות הכולל. □

תרגיל 6.11. נסמן ב-X את המשתנה המקרי הסופר את מספר ההטלות עד הפעם הרביעית בה יופיע עץ בהינתן שההופעה הראשונה הייתה בהטלה החמישית. הסיכוי לעץ בכל הטלה הוא חצי ולכן מספר ההטלות בין העץ הראשון לבין העץ הרביעי מתפלג לפי התפלגות $Y\sim NB\left(3,\frac{1}{2}\right)$. המאורע $\{X=k\}$ מתרחש כאשר בין העץ הראשון לעץ הרביעי היו בדיוק $k-5$ הטלות, כלומר

$$Pr\left(X=k\right)=Pr\left(Y=k-5\right)=\binom{k-5}{2}\cdot\left(\frac{1}{2}\right)^{3}\cdot\left(\frac{1}{2}\right)^{k-5-3}=\binom{k-5}{2}\left(\frac{1}{2}\right)^{k-5}$$

כאשר התומך הינו $supp\left(X\right)=\{8,9,\ldots\}$, שכן אם העץ הראשון היה בהטלה החמישית, העץ הרביעי יכול להיות רק החל מההטלה השמינית (במקרה בו גם ההטלה השישית, השביעית והשמינית היו עץ). □

תרגיל 6.12. בהינתן ש-6 יצא 4 פעמים, יש לנו 10 פריטים: 4 הטלות "6" ו-6 הטלות "לא 6". לכל אחת מהתוצאות יש סיכוי שווה להתקבל בכל אחת מההטלות ולכן השאלה היא בעצם בחירה של 5 תוצאות מתוך 10 וספירה כמה מתוכן הן "6". זו בדיוק ההגדרה של התפלגות היפרגאומטרית ולכן $HG\left(10,4,5\right)$. □

תרגיל 6.13. הדייג בוחר למעשה 6 דגים מתוך אוכלוסיה בגודל 20 ובודק כמה מתוכם הם דגי זהב. יש 5 דגי זהב ולכן ההתפלגות של דגי הזהב שנדוגו הינה $X\sim HG\left(20,5,6\right)$. באופן דומה, במקום בחירת הדגים שהדייג דג, ניתן לבחור את הדגים שהדייג משאיר בבריכה. לכן הדגימה היא של 14 פריטים (הדגים שנשארים בבריכה) ומתקיים $Y\sim HG\left(20,5,14\right)$. □

תרגיל 6.14. בהתאם לנתוני השאלה, מספר המנועים הפגומים מתפלג $X\sim Bin(300,0.01)$ ולכן

$$\begin{aligned}Pr\left(X>2\right)&=1-Pr\left(X=0\right)-Pr\left(X=1\right)-Pr\left(X=2\right)\\&=1-0.01^{300}-\binom{300}{1}0.01^{299}\cdot0.99-\binom{300}{2}0.01^{298}\cdot0.99^{2}\\&=0.577936\end{aligned}$$

בהמשך לנתונים נקבל $\lambda=np=3$ וניתן לבצע קירוב להתפלגות פואסונית. נשתמש בנוסחא הקודמת עבור המשלים אך הפעם נציב את ההסתברויות המתאימות מהתפלגות פואסון:

$$\begin{aligned}Pr\left(X>2\right)&=1-Pr\left(X=0\right)-Pr\left(X=1\right)-Pr\left(X=2\right)\\&=1-e^{-3}\cdot\frac{3^{0}}{0!}-e^{-3}\cdot\frac{3^{1}}{1!}-e^{-3}\cdot\frac{3^{2}}{2!}=0.57681\end{aligned}$$

הקירוב הפואסוני במקרה זה מהווה קירוב מעולה עם שגיאה של כ-0.2%. □

תרגיל 6.15. יהי $X\sim Pois\left(\lambda\right)$. נחשב את הפונקציה היוצרת מומנטים ישירות לפי ההגדרה:

$$\begin{aligned}M_{X}\left(t\right)&=\mathbb{E}\left(e^{tX}\right)=\sum_{k=0}^{\infty}e^{tk}\cdot e^{-\lambda}\frac{\lambda^{k}}{k!}=\sum_{k=0}^{\infty}e^{-\lambda}\frac{\left(\lambda e^{t}\right)^{k}}{k!}\\(eq.\,0.18)&=e^{-\lambda}\cdot e^{\left(\lambda e^{t}\right)}=e^{-\lambda\left(1-e^{t}\right)}\end{aligned}$$

□

תרגיל 6.16. המועמדים בלתי תלויים זה בזה, לכן מספר המועמדים המתאימים מתפלג לפי $X \sim Bin\left(10^5, 10^{-5}\right)$. זהו בדיוק המצב הקלאסי לשימוש בקירוב הפואסוני - מספר ניסויים גדול וסיכוי קטן להצלחה. לכן נגדיר $\lambda = 10^5 \cdot 10^{-5} = 1$ ובקירוב $X \sim Pois\left(1\right)$. השר הודיע שהוא מכיר 2 מועמדים מתאימים, לכן ידוע כי $X \geq 2$ ואין מועמדים נוספים מתאימים אם בדיוק $X = 2$. לכן ההסתברות המבוקשת היא

$$\begin{aligned} Pr\left(X = 2 | X \geq 2\right) &= \frac{Pr\left(X = 2 \cap X \geq 2\right)}{Pr\left(X \geq 2\right)} = \frac{Pr\left(X = 2\right)}{1 - Pr\left(X = 0\right) - Pr\left(X = 1\right)} \\ &= \frac{e^{-1} \cdot \frac{1^2}{2!}}{1 - e^{-1} \cdot \frac{1^0}{0!} - e^{-1} \cdot \frac{1^1}{1!}} = 0.696 \end{aligned}$$

הסתברות גבוהה זו מתאימה לאינטואיציה, שהרי התוחלת והשונות של X שוות ל-1, לכן הציפיה היא שיהיה בערך מועמד אחד מתאים ובוודאי לא הרבה יותר מ-2. □

תרגיל 6.17. פונקציית הצפיפות של המשתנה המקרי האחיד הרציף היא מאוד פשוטה (קבוע) ולכן האינטגלים המתארים את התוחלת, השונות והפונקציה היוצרת מומנטים פשוטים מאוד. עבור התוחלת נקבל את האינטגרל:

$$\begin{aligned} \mathbb{E}\left(X\right) &= \int_a^b \frac{t}{b-a}\, dt = \frac{1}{b-a} \cdot \left.\frac{t^2}{2}\right|_a^b = \frac{b^2 - a^2}{2\left(b-a\right)} \\ &= \frac{\left(b-a\right) \cdot \left(b+a\right)}{2\left(b-a\right)} = \frac{a+b}{2} \end{aligned}$$

כצפוי התוחלת, המייצגת את הערך הממוצע של ההתפלגות, היא אמצע הקטע שכן כל התוצאות בעלות אותה ההסתברות. לצורך החישוב של השונות נחשב ראשית את $\mathbb{E}\left(X^2\right)$:

$$\begin{aligned} \mathbb{E}\left(X^2\right) &= \int_a^b \frac{t^2}{b-a}\, dt = \frac{1}{b-a} \cdot \left.\frac{t^3}{3}\right|_a^b = \frac{b^3 - a^3}{3\left(b-a\right)} \\ &= \frac{\left(b-a\right) \cdot \left(b^2 + ab + a^2\right)}{3\left(b-a\right)} = \frac{b^2 + ab + a^2}{3} \end{aligned}$$

ולכן השונות הינה

$$\begin{aligned} \mathbb{V}\left(X\right) &= \mathbb{E}\left(X^2\right) - \mathbb{E}^2\left(X\right) = \frac{b^2 + ab + a^2}{3} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{4b^2 + 4ab + 4a^2 - 3a^2 - 6ab - 3b^2}{12} = \frac{b^2 - 2ab + a^2}{12} \\ &= \frac{\left(b-a\right)^2}{12} \end{aligned}$$

כדאי לשים לב שבניגוד לתוחלת, עבור השונות קיבלנו נוסחא שונה מאשר במקרה הבדיד (נוסחא 6.3). לסיום, נחשב את הפונקציה היוצרת מומנטים:

$$M_X\left(t\right) = \mathbb{E}\left(e^{tX}\right) = \int_a^b \frac{e^{ty}}{b-a}\, dy = \frac{1}{b-a} \cdot \left.\frac{e^{ty}}{t}\right|_a^b = \frac{e^{tb} - e^{ta}}{t\left(b-a\right)}$$

כנדרש. □

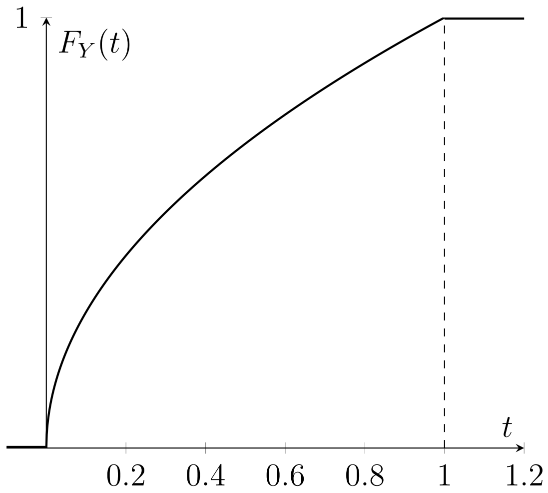
תרגיל 6.18. שאלה זו זהה לחלוטין לשאלות דומות שפתרנו עבור משתנים מקריים רציפים בפרק 4, למשל בעיה 4.4. כפי שמוסבר שם, נעדיף לחשב את פונקציית ההתפלגות המצטברת של Y כאשר ההבדל היחידי הוא שההתפלגות של X נתונה בשמה ולא בפונקציית הצפיפות המפורשת.

מאחר ו-$-1 \le X \le 1$ הרי ש-$0 \le X^2 \le 1$ ולכן הערכים המעניינים עבור פונקציית ההתפלגות המצטברת של Y הינם בתחום $t \in [0,1]$:

$$F_Y(t) = Pr(Y \le t) = Pr\left(X^2 \le t\right) = Pr\left(-\sqrt{t} \le X \le \sqrt{t}\right)$$
$$= F_X\left(\sqrt{t}\right) - F_X\left(-\sqrt{t}\right) \overset{*}{=} \frac{\sqrt{t}+1}{2} - \frac{-\sqrt{t}+1}{2} = \sqrt{t}$$

כאשר במעבר $*$ הצבנו את $\sqrt{t} \in [0,1]$ בפונקציית ההתפלגות המצטברת של משתנה מקרי אחיד רציף $U(0,1)$.
לסיכום, פונקציית ההתפלגות המצטברת תהיה

$$F_Y(t) = \begin{cases} 0 & t < 0 \\ \sqrt{t} & t \in [0,1] \\ 1 & t > 1 \end{cases}$$



□

תרגיל 6.19. לאור הנתונים, זמן היציאה של הסטודנט מביתו מתפלג לפי $X \sim U(8,12)$ ואילו זמן הנסיעה הינו $Y = 1 + \frac{1}{X}$. נמצא את ההתפלגות של Y על ידי חישוב פונקציית ההתפלגות המצטברת שלו. ראשית נשים לב לתחום: $8 \le X \le 12$ לכן $\frac{1}{12} \le \frac{1}{X} \le \frac{1}{8}$ כלומר $\frac{13}{12} \le Y \le \frac{9}{8}$ וזהו טווח הערכים המעניין עבורו יש לחשב את פונקציית ההתפלגות המצטברת. יהי $t \in \left[\frac{13}{12}, \frac{9}{8}\right]$:

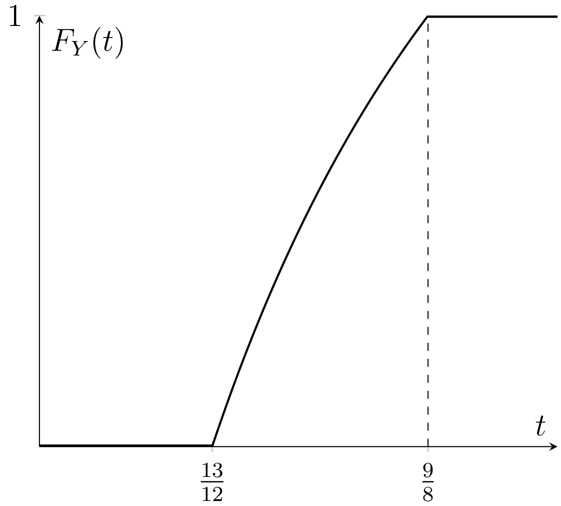
$$F_Y(t) = Pr(Y \le t) = Pr\left(1 + \frac{1}{X} \le t\right) = Pr\left(\frac{1}{X} \le t-1\right) = Pr\left(X \ge \frac{1}{t-1}\right)$$

כאשר המעבר האחרון אפשרי בזכות העובדה ש-$1 < t$ ואיננו מחלקים במספר שלילי, פעולה שהייתה הופכת את כיוון אי-השוויון. נציב את הביטוי שקיבלנו בפונקציית ההתפלגות המצטברת של משתנה מקרי אחיד $X \sim U(8,12)$:

$$F_Y(t) = 1 - Pr\left(X \leq \frac{1}{t-1}\right) = 1 - F_X\left(\frac{1}{t-1}\right) = 1 - \frac{\frac{1}{t-1} - 8}{4} = \frac{12t-13}{4(t-1)}$$

לסיכום, פונקציית ההתפלגות המצטברת הינה

$$F_Y(t) = \begin{cases} 0 & t < \frac{13}{12} \\ \frac{12t-13}{4(t-1)} & t \in \left[\frac{13}{12}, \frac{9}{8}\right] \\ 1 & t > \frac{9}{8} \end{cases}$$



□

תרגיל 6.20. יהי $X \sim Exp(1)$ונחשב את ההתפלגות של $Y = \frac{X}{2} + 5$ על ידי חישוב פונקציית ההתפלגות המצטברת:

$$\begin{aligned} F_Y(t) &= Pr(Y \leq t) = Pr\left(\frac{X}{2} + 5 \leq t\right) = Pr(X \leq 2(t-5)) \\ &= F_X(2(t-5)) = \begin{cases} 0 & 2(t-5) < 0 \\ 1 - e^{-2(t-5)} & 2(t-5) \geq 0 \end{cases} = \begin{cases} 0 & t < 5 \\ 1 - e^{-2(t-5)} & t \geq 5 \end{cases} \end{aligned}$$

□

תרגיל 6.21. נשים לב ש-Y מתקבל על ידי פונקציית הערך השלם, לכן הוא בהכרח שלם והתומך הינו $\mathbb{N}$, כלומר זהו משתנה מקרי בדיד. לעומת זאת, Z הוא החלק "הלא שלם" ולכן הוא יכול לקבל כל ערך בתחום $[0,1]$ וזהו משתנה מקרי רציף.

1. יהי $k \in \mathbb{N}$. ההתפלגות של Y תתקבל על ידי חישוב המאורע $\{Y = k\}$:

$$P_Y(k) = Pr(Y = k) = Pr(\lfloor X \rfloor + 1 = k) = Pr(\lfloor X \rfloor = k - 1)$$

כדי שהערך השלם של X יהיה שווה ל-$k-1$ על X עצמו להיות בין $k-1$ ל-k:

$$\begin{aligned} P_Y(k) &= Pr(k-1 \le X < k) = F_X(k) - F_X(k-1) \\ &= \left(1 - e^{-k}\right) - \left(1 - e^{-(k-1)}\right) = \left(\frac{1}{e}\right)^{k-1}\left(1 - \frac{1}{e}\right) \end{aligned}$$

נסמן $p = 1 - e^{-1}$. הנוסחא האחרונה היא מהצורה $p\,(1-p)^{k-1}$ המאפיינת התפלגות גאומטרית, לכן $Y \sim Geom\left(1 - e^{-1}\right)$.

2. לפי תכונות המשתנה המקרי הגאומטרי (נוסחאות 6.14 ו- 6.15) התוחלת הינה $\mathbb{E}(X) = \frac{1}{p} = 1.58$ והשונות הינה $\mathbb{V}(X) = \frac{1-p}{p^2} = 0.92$. אמנם ניתן היה להשאיר את התשובה לסעיף הקודם בתור הנוסחא $\left(\frac{1}{e}\right)^{k-1}\left(1 - \frac{1}{e}\right)$, אך ללא זיהוי התפלגות זו כהתפלגות גאומטרית, חישוב התוחלת והשונות היה מצריך עבודה רבה נוספת.

3. נחשב את פונקציית ההתפלגות המצטברת של Z. כאמור, התחום המעניין הוא $[0,1]$ ולכן יהי $t \in [0,1]$. כדי שהערך הלא שלם של X יהיה קטן מ-t, הערך שלו צריך להיות בין n ל-$n+t$ עבור n שלם כלשהו. לכן המאורע $Z \le t$ הוא איחוד של מאורעות זרים מהצורה $\{n \le X \le n+t\}$:

$$\begin{aligned} F_Z(t) &= Pr(Z \le t) = \sum_{n=0}^{\infty} Pr(n \le X \le n+t) = \sum_{n=0}^{\infty} F_X(n+t) - F_X(n) \\ &= \sum_{n=0}^{\infty}\left(e^{-(n+t)} - e^{-n}\right) = \left(1 - e^{-t}\right)\sum_{n=0}^{\infty} e^{-n} \\ (eq.\, 0.15) &= \frac{1 - e^{-t}}{1 - e^{-1}} \end{aligned}$$

ולכן

$$F_Z(t) = \begin{cases} 0 & t < 0 \\ \frac{1-e^{-t}}{1-e^{-1}} & t \in [0,1] \\ 1 & t > 1 \end{cases}$$

□

תרגיל 6.22. נסמן את זמן המשחק של דני ויונתן כשיש להם כוח ב-$X \sim Exp\left(\frac{1}{2}\right)$, כשאין להם כוח ב-$Y \sim Exp(1)$ ואת זמן המשחק ב-Z. כמו כן, נסמן ב-A את המאורע שלשחקנים יש כוח לשחק היום.

1. המאורע המבוקש הוא $\{1 \le Z \le 2\}$. לו היינו יודעים אם יש להם כוח לשחק או אין להם, ההתפלגות הייתה ידועה וניתן היה לחשב את ההסתברות המבוקשת. היות ופרט זה לא ידוע, נחשב את ההסתברות המבוקשת על ידי התנייה באמצעות נוסחת ההסתברות השלמה (נוסחא 2.18):

$$\begin{aligned} Pr\left(1 \le Z \le 2\right) &= Pr\left(1 \le Z \le 2|A\right) Pr\left(A\right) + Pr\left(1 \le Z \le 2|\overline{A}\right) Pr\left(\overline{A}\right) \\ &= Pr\left(1 \le X \le 2\right) \cdot \frac{1}{4} + Pr\left(1 \le Y \le 2\right) \cdot \frac{3}{4} = \frac{1}{4}\left(F_X\left(2\right) - F_X(1)\right) + \frac{3}{4}\left(F_Y\left(2\right) - F_Y\left(1\right)\right) \\ &= \frac{1}{4}\left(e^{-\frac{1}{2}\cdot 1} - e^{-\frac{1}{2}\cdot 2}\right) + \frac{3}{4}\left(e^{-1} - e^{-2}\right) = 0.234 \end{aligned}$$

2. בדומה לסעיף הקודם, נחשב את פונקציית ההתפלגות המצטברת של Z. ראשית, הן X והן Y חיוביים, לכן זמן המשחק בהכרח חיובי ומתקיים $Pr\left(Z < t\right) = 0$ לכל $t < 0$. שנית, עבור $t \ge 0$ נחשב את ההסתברות של המאורע $\{Z \le t\}$ על ידי נוסחת ההסתברות השלמה על המאורע A:

$$\begin{aligned} F_Z\left(t\right) &= Pr\left(Z \le t\right) = Pr\left(Z \le t|A\right) Pr\left(A\right) + Pr\left(Z \le t|\overline{A}\right) Pr\left(\overline{A}\right) = \frac{1}{4}F_X\left(t\right) + \frac{3}{4}F_Y\left(t\right) \\ &= \frac{1}{4}\left(1 - e^{-\frac{t}{2}}\right) + \frac{3}{4}\left(1 - e^{-t}\right) = 1 - \frac{1}{4}e^{-\frac{t}{2}} - \frac{3}{4}e^{-t} \end{aligned}$$

3. ידוע שזמן המשחק ארוך משעה, כלומר ידוע שהמאורע $\{Z > 1\}$ התרחש ועלינו לחשב את ההסתברות של המאורע $\{Z > 2\}$:

$$\begin{aligned} Pr\left(Z > 2|Z > 1\right) &= \frac{Pr\left(Z > 2 \cap Z > 1\right)}{Pr\left(Z > 1\right)} = \frac{Pr\left(Z > 2\right)}{Pr\left(Z > 1\right)} = \frac{1 - F_Z\left(2\right)}{1 - F_Z\left(1\right)} \\ &= \frac{\frac{1}{4}e^{-1} + \frac{3}{4}e^{-2}}{\frac{1}{4}e^{-\frac{1}{2}} + \frac{3}{4}e^{-1}} = 0.453 \end{aligned}$$

שימו לב שההתפלגות של Z אינה מעריכית ולכן תכונת חוסר הזיכרון לא מתקיימת:

$$Pr\left(Z > 1\right) = \frac{1}{4}e^{-\frac{1}{2}} + \frac{3}{4}e^{-1} = 0.427$$

העובדה שהמשחק נמשך כבר יותר משעה מלמדת על זה שכנראה לשחקנים יש כוח היום ולכן מגדילה את הסיכוי שהמשחק יימשך יותר זמן.

$\square$

תרגיל 6.23. יהי $Z \sim N\left(0, 1\right)$. נחשב את ההסתברויות המבוקשות על ידי הצבה בטבלה:

1. המאורע המבוקש הינו $\{Z > 1\}$ ולכן

$$Pr\left(Z > 1\right) = 1 - Pr\left(Z \le 1\right) = 1 - \Phi\left(1\right) = 1 - 0.841 = 0.159$$

2. הפעם המאורע המבוקש הינו $\{|Z|>1\}$ ולכן ניתן לרשום אותו כאיחוד של שני המאורעות הזרים $\{Z>1\}\cup\{Z<-1\}$:

$$\begin{aligned}Pr\left(|Z|>1\right)&=Pr\left(Z>1\right)+Pr\left(Z<-1\right)=1-\Phi\left(1\right)+\Phi\left(-1\right)\\&=1-\Phi\left(1\right)+1-\Phi\left(1\right)=0.318\end{aligned}$$

ניתן היה לשים מראש שמאחר וההתפלגות סימטרית, הסיכוי להיות גדול מ-1 שווה לסיכוי להיות קטן מ--1 ולכן ההסתברות של סעיף זה כפולה מההסתברות של הסעיף הקודם.

3. מחפשים את הערך שהסיכוי להיות מתחתיו הוא 95%, כלומר את האחוזון ה-0.95 של ההתפלגות. לצורך כך נפעל הפוך - נאתר בטבלה את הערך 0.95 ונמצא לאיזו שורה ולאיזו עמודה הוא מתאים. ערך זה לא מופיע בטבלה אבל מתקיים $\Phi\left(1.64\right)<0.95$ בעוד ש-$0.95<\Phi\left(1.65\right)$ לכן ניתן לבחור בקירוב את הערך $z_{0.95}=1.645$.

□

תרגיל 6.24. נחשב ראשית את הפונקציה היוצרת מומנטים של משתנה מקרי נורמלי סטנדרטי. יהי $Z\sim N\left(0,1\right)$. מתקיים

$$M_Z\left(t\right)=\mathbb{E}\left(e^{Zt}\right)=\int_{-\infty}^{\infty}e^{ut}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}\,du=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2-2ut}{2}}\,du$$

נשלים את המעריך לריבוע: $u^2-2ut=\left(u-t\right)^2-t^2$ ונקבל:

$$M_Z\left(t\right)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{\left(u-t\right)^2}{2}}e^{\frac{t^2}{2}}\,du=e^{\frac{t^2}{2}}$$

שכן האינטגרל על u הוא אינטגרל על כל הישר של משתנה מקרי המתפלג $N\left(t,1\right)$ ולכן שווה ל-1.
עבור $X\sim N\left(\mu,\sigma^2\right)$ מתקיים $X=\sigma Z+\mu$ ולכן לפי תרגיל **5.24**

$$M_X\left(t\right)=M_{\sigma Z+\mu}\left(t\right)=e^{\mu t}M_Z\left(\sigma t\right)=e^{\mu t}e^{\frac{\left(\sigma t\right)^2}{2}}=e^{\frac{\sigma^2t^2+2\mu t}{2}}$$

□

תרגיל 6.25. נסמן את גובה בגברים ב-$X\sim N\left(175,10^2\right)$.

1. גבר אקראי יוזמן למיונים אם גובהו גבוה מ-200 ס"מ, תופעה המתרחשת בסיכוי

$$Pr\left(X>200\right)=1-\Phi\left(\frac{200-175}{10}\right)=1-\Phi\left(2.5\right)=1-0.99379=0.00621$$

כלומר רק 0.621% מהגברים גבוהים משני מטרים ויכולים להיבחן לקבוצת הכדורסל.

2. לכל גבר יש סיכוי 0.00621 להיות גבוה מ-2 מטרים ומספר הגברים הכולל הוא **10,000**. היות והגברים בלתי תלויים, מספר הגברים הגבוהים מ-2 מטרים מתפלג $Bin\left(10,000,0.00621\right)$ ותוחלת מספר הגברים המתאימים הינה (נוסחא **6.6**) $10000\cdot0.00621=62.1$.

3. לכל נבחן יש סיכוי של $\frac{1}{3}$ להיות בנבחרת ובוחנים שחקנים עד שמוצאים 12 מתאימים, לכן מספר הנבחנים מתפלג בינומית שלילית $NB\left(12, \frac{1}{3}\right)$.

□

תרגיל 6.26. נסמן את מספר העצים שהתקבל ב-X. בכל הטלה יש סיכוי של $\frac{1}{2}$ לקבל עץ ו-X סופר את מספר העצים שהתקבלו, לכן $X \sim Bin\left(100, \frac{1}{2}\right)$. תוחלת מספר המופעים של עץ אם כך הינו (נוסחא 6.6) $\mathbb{E}\left(X\right) = 100 \cdot \frac{1}{2} = 50$ ולכן מטבע יוכרז כלא הוגן כאשר $|X - 50| > 10$ כלומר $\{X > 60\} \cup \{X < 40\}$. מסימטריה, לשני מאורעות אלו אותה ההסתברות ולכן נחשב את ההסתברות של אחד מהם ונכפול ב-2. מאחר ולא ניתן לחשב מפורשות את ההסתברות של המאורע $\{X > 60\}$עבור התפלגות בינומית נקרב על ידי משתנה מקרי נורמלי לפי נוסחא 6.48:

$$Pr\left(X > 60\right) = Pr\left(X > 59.5\right) = 1 - \Phi\left(\frac{59.5 - 50}{\sqrt{50 \cdot \frac{1}{2} \cdot \frac{1}{2}}}\right) = 1 - \Phi\left(2.69\right) = 1 - 0.9964 = 0.0036$$

לכן ההסתברות שמטבע הוגן יוכרז כמטבע לא הוגן הינה $2 \cdot 0.0036 = 0.0072$. הסתברות מכונה בסטטיסטיקה "ההסתברות לשגיאה מסוג ראשון" ומתארת מצב בו אנחנו עורכים ניסוי ובטעות מקבלים שההנחה הראשונית שלנו, שהמטבע הוגן, לא נכונה למרות שהיא כן נכונה. במקרה הזה, ההסתברות הזו מאוד נמוכה ולכן כמעט ולא נטעה ולא נחשוב שמטבעות הוגנים הם בעצם לא הוגנים. □

תרגיל 6.27. נסמן את ההכנסות ביום ה-i החל מפתיחת הדוכן ב-X_i. לפי הנתונים, $\mathbb{E}\left(X_i\right) = 2000$ ו-$\mathbb{V}\left(X_i\right) = 1000$. נסמן ב-$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ את ממוצע ההכנסות של הדוכן מיום הקמתו ועד היום ה-n.

1. עבור 100 ימים נציב $n = 100$ ונחשב את ההסתברות של המאורע $\left\{\overline{X}_{100} < 2010\right\}$. לפי משפט הגבול המרכזי $\overline{X}_{100} \sim N\left(2000, \frac{1000}{100}\right)$ ולכן:

$$Pr\left(\overline{X}_{100} < 2010\right) = Pr\left(\frac{\overline{X}_{100} - 2000}{\sqrt{\frac{1000}{100}}} < \frac{2010 - 2000}{\sqrt{\frac{1000}{100}}}\right) = \Phi\left(\frac{10}{\sqrt{10}}\right) = 3.16 = 0.99921$$

2. בטרם נפנה לחישוב, אנחנו יודעים בוודאות שקיים יום כזה שכן לפי משפט הגבול המרכזי, ככל שנחכה יותר ימים כך הממוצע היומי ישאף לתוחלת וילך ויתקרב ל-2000. ככל שנחכה יותר, כך הממוצע יהיה קרוב יותר לערך זה וההסתברות שהוא יסטה מהתוחלת ב-1 ילך ויקטן. נחשב זאת מפורשות. לפי משפט הגבול המרכזי, $\overline{X}_n \sim N\left(2000, \frac{1000}{n}\right)$ ולכן

$$Pr\left(\overline{X}_n > 1999\right) = Pr\left(\frac{\overline{X}_n - 2000}{\sqrt{\frac{1000}{n}}} > \frac{1999 - 2000}{\sqrt{\frac{1000}{n}}}\right) = 1 - \Phi\left(-\sqrt{\frac{n}{1000}}\right) = \Phi\left(\sqrt{\frac{n}{1000}}\right)$$

נדרוש שהסתברות זו תהיה גבוהה מ-0.99ולכן הארגומנט של Φצריך להיות גדול מ-2.33 שכן $\Phi\left(2.33\right) = 0.99$. לכן $\sqrt{\frac{n}{1000}} \geq 2.33$ או $n \geq 1000 \cdot 2.33^2 = 5429$.

□

תרגיל 6.28. 1. הסיכוי לקבל 6 בכל אחת מהקוביות הוא $\frac{1}{6}$ לכן מספר המופעים של 6 הינו $X \sim Bin\left(3, \frac{1}{6}\right)$. יחד עם זאת, זהו לא הרווח של השחקן שכן כאשר $X = 0$ השחקן משלם שקל.

לכן, ההתפלגות של הרווח, Y, היא כמעט בינומית ונציגה מפורשות באמצעות פונקציית ההסתברות:

$$Pr\left(Y=k\right)=\begin{cases}Pr\left(X=0\right) & k=-1\\ Pr\left(X=k\right) & k\in\left\{1,2,3\right\}\end{cases}=\begin{cases}\left(\frac{5}{6}\right)^{3} & k=-1\\ \binom{3}{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{3-k} & k\in\left\{1,2,3\right\}\end{cases}$$

2. ניתן לחשב את התוחלת על ידי חישוב ישיר וסכימת כל הערכים, אך יותר פשוט להציג את Y במונחים של X ולהשתמש בתוחלת הידועה של משתנה מקרי בינומי. כאשר $X\neq 0$ מתקיים $Y=X$ וכאשר $X=0$ מתקיים $Y=-1$ לכן ניתן לרשום $Y=X-\mathbf{1}_{X=0}$. לפיכך, תוך שימוש בחיבוריות התוחלת (נוסחא 5.6) נקבל:

$$\begin{aligned}\mathbb{E}\left(Y\right)&=\mathbb{E}\left(X-\mathbf{1}_{X=0}\right)=\mathbb{E}\left(X\right)-\mathbb{E}\left(\mathbf{1}_{X=0}\right)\\&=3\cdot\frac{1}{6}-\left(\frac{5}{6}\right)^{3}=-0.0787\end{aligned}$$

□

תרגיל 6.29. 1. כאשר אף מפתח לא מנוסה פעמיים ההוצאה מתבצעת למעשה ללא החזרה ולכן הדבר שקול לסידור כל המפתחות בשורה (פרק 3.2.1) ובדיקתם אחד-אחד. הסיכוי של המפתח הנכון להיות בכל אחד מהמקומות הוא $\frac{1}{10}$ (מסימטריה) ולכן מספר הניסיונות מתפלג $U\left\{1,\ldots,10\right\}$.
ניתן לחשב זאת גם בצורה מפורשת על ידי חישוב ההסתברויות של כל הערכים האפשריים. למשל, כדי שמספר הניסיונות יהיה 3, צריך כישלון בניסיון הראשון (הסתברות $\frac{9}{10}$), כישלון בניסיון השני (הסתברות $\frac{8}{9}$ כי מפתח אחד כבר לא בצרור) והצלחה בניסיון השלישי (הסתברות $\frac{1}{8}$) ובסך הכל:

$$Pr\left(X=3\right)=\frac{9}{10}\cdot\frac{8}{9}\cdot\frac{1}{8}=\frac{1}{10}$$

ובאופן דומה החישוב יניב $\frac{1}{10}$ לכל $k\in\left\{1,\ldots,10\right\}$.

2. בכל ניסיון לשיכור יש סיכוי של $\frac{1}{10}$ להוציא את המפתח הנכון באופן בלתי תלוי בכשלונות הקודמים, שכן המפתחות שלא הצליחו חוזרים לצרור וניתן לבחור בהם שוב. מספר הניסיונות אם כך מתפלג $Geom\left(\frac{1}{10}\right)$.

□

תרגיל 6.30. יהי $n\in\mathbb{N}$. כאשר מטילים מטבע הוגן n פעמים, מספר המופעים של "עץ" מתפלג לפי התפלגות בינומית $X\sim Bin\left(n,\frac{1}{2}\right)$. נסכום את ההסתברויות של כל הערכים הזוגיים בתומך:

$$\begin{aligned}Pr\left(X\text{ is even}\right)&=\sum_{\substack{k=0\\(even)}}^{n}\binom{n}{k}\cdot\left(\frac{1}{2}\right)^{k}\left(\frac{1}{2}\right)^{n-k}=\sum_{\substack{k=0\\(even)}}^{n}\binom{n}{k}\cdot\left(\frac{1}{2}\right)^{n}\\&=\left(\frac{1}{2}\right)^{n}\left(\binom{n}{0}+\binom{n}{2}+\ldots\right)\overset{eq.\,0.12}{=}\left(\frac{1}{2}\right)^{n}\left(\binom{n}{1}+\binom{n}{3}+\ldots\right)\\&=\sum_{\substack{k=0\\(odd)}}^{n}\binom{n}{k}\cdot\left(\frac{1}{2}\right)^{n}=\sum_{\substack{k=0\\(odd)}}^{n}\binom{n}{k}\cdot\left(\frac{1}{2}\right)^{k}\left(\frac{1}{2}\right)^{n-k}=Pr\left(X\text{ is odd}\right)\end{aligned}$$

כלומר הסיכוי שמספר העצים זוגי שווה לסיכוי שמספר העצים אי-זוגים. מאחר ואלו שתי האפשריות היחידות והן זרות, ההסתברות של כל אחת מהן היא חצי (ללא תלות בזוגיות של n).

ניתן לקבל תוצאה זו ללא ביצוע אף חישוב. נניח שאחרי $n-1$ הטלות מספר העצים הוא זוגי. הסיכוי שמספר העצים יהיה זוגי גם אחרי ההטלה הבאה הוא $\frac{1}{2}$ (בהטלה הבאה צריך לצאת פלי) והסיכוי שמספר העצים יהיה אי-זוגי הוא $\frac{1}{2}$ גם כן. ניתן לבצע שיקול זה גם כאשר אחרי $n-1$ הטלות מספר העצים הוא אי-זוגי. לכן, לא משנה מהי זוגיות מספר העצים אחרי $n-1$ הטלות, הסיכוי שמספר העצים אחרי n הטלות (ובניסוי כולו) יהיה זוגי הוא $\frac{1}{2}$, כמקודם. $\square$

תרגיל 6.31. נגדיר X - מספר ההטלות עד אשר יצא עץ לראשונה. הסיכוי לעץ בכל הטלה הינו p באופן בלתי תלוי בהטלות האחרות ולכן $X\sim Geom\left(p\right)$. המאורע שעץ יוצא לראשונה בהטלה זוגית הינו איחוד של מאורעות זרים מהצורה $\left\{ X=2k\right\}$ כאשר $k\in\mathbb{N}$ ולכן הסתברותו:

$$Pr\left(\bigcup_{k=1}^{\infty}\left\{ X=2k\right\} \right)=\sum_{k=1}^{\infty}Pr\left(X=2k\right)=\sum_{k=1}^{\infty}\left(1-p\right)^{2k-1}p=\frac{p}{1-p}\sum_{k=1}^{\infty}\left(1-p\right)^{2k}$$

נציב $q=\left(1-p\right)^{2}$ ונקבל את הטור הגאומטרי המוכר (נוסחא **0.15**) אך ללא האיבר $k=0$. נוסיף ונחסר אותו נקבל:

$$\frac{p}{1-p}\sum_{k=1}^{\infty}q^{k}=\frac{p}{1-p}\left(\frac{1}{1-q}-1\right)=\frac{p}{1-p}\left(\frac{1}{1-\left(1-p\right)^{2}}-1\right)=\frac{1-p}{2-p}$$

כצפוי, קיבלנו את אותה התשובה כמו בתרגיל **2.15**. $\square$

תרגיל 6.32.

1. לכל משבצת יש הסתברות חצי להיות לבנה ויש 9 משבצות בלתי תלויות, לכן מספר המשבצות הלבנות מתפלג $Bin\left(9,\frac{1}{2}\right)$.
2. ישנן שלוש שורות בלתי תלויות ובכל שורה יש סיכוי מסויים לכך שיהיו בה שתי משבצות לבנות, לכן מספר השורות בהן יש בדיוק שתי משבצות לבנות מתפלג בינומית. נותר רק לחשב את ההסתברות שבשורה מסויימת יהיו שתי משבצות לבנות. בכל שורה ישנן 3 משבצות וכל משבצת יכולה להיות לבנה בסיכוי $\frac{1}{2}$, לכן מספר המשבצות הלבנות בשורה מתפלג $Bin\left(3,\frac{1}{2}\right)$. לפיכך, הסיכוי שבשורה יהיו בדיוק 2 משבצות לבנות הינו

$$\binom{3}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{1}=\frac{3}{8}$$

ולסיכום, מספר השורות בהן יש בדיוק שתי משבצות לבנות מתפלג $Bin\left(3,\frac{3}{8}\right)$.
3. מסימטריה, המשבצת הלבנה יכולה להיות כל אחת מ-9 המשבצות שעל הלוח ויכולה להיות בכל אחת מהשורות בהסתברות שווה. לכן, מספר השורה בה המשבצת נמצאת מתפלג $U\left\{ 1,2,3\right\}$.
4. ניתן להסתכל על הבעיה השקולה הבאה: יש 3 משבצות לבנות ו-6 משבצות שחורות והשורה הראשונה צריכה לבחור 3 משבצות שירכיבו אותה. המשתנה המקרי סופר את מספר המשבצות הלבנות שנבחרו וההתפלגות המתאימה היא התפלגות היפרגאומטרית כאשר האוכלוסיה היא 9 משבצות, 3 מתוכן מיוחדות (לבנות), נבחרות 3 והמשתנה המקרי סופר כמה מיוחדות נבחרו: $HG\left(9,3,3\right)$.

$\square$

תרגיל 6.33. לפי תרגיל 3.5 הסיכוי לזכות בהגרלת הלוטו הינו $6.14 \cdot 10^{-8}$, לכן מספר הזוכים בהגרלה הכוללת מיליון אנשים מתפלג לפי $X \sim Bin\left(10^6, 6.14 \cdot 10^{-8}\right)$. זהו מצב קלאסי בו מספר הניסויים גדול, הסיכוי להצלחה קטן ואילו מכפלתם $\lambda = 10^6 \cdot 6.14 \cdot 10^{-8} = 0.0614$ היא מספר סביר. לכן בקירוב, $X \sim Pois\left(0.0614\right)$ ונשתמש בקירוב זה לצורך חישוב ההסתברויות.

אף אחד לא זוכה כאשר $X = 0$ ולכן ההסתברות הינה

$$Pr\left(X = 0\right) = e^{-0.0614} \cdot \frac{0.0614^0}{0!} = 0.94$$

ערך זה גבוה ומתאים לכך שבפועל, בהרבה מן ההגרלות, לא נמצאים זוכים. יחד עם זאת, חלק ניכר מההנחות שהנחנו בפתרון השאלה לא תקפות במציאות. למשל, האנשים הם לא בלתי תלויים שכן ישנם מספרים יותר פופולריים (מספרי מזל, תאריכי לידה וכו') אותם רבים מנחשים ובנוסף – אנשים נוטים לשלוח מספר ניחושים (וניחושים של אותו האדם, כמובן, תלויים).

אם יש זוכה, הרי ש-$X \geq 1$. בהינתן מאורע זה, הסיכוי שיש זוכה בודד הינו

$$Pr\left(X = 1 | X \geq 1\right) = \frac{Pr\left(X = 1 \cap X \geq 1\right)}{Pr\left(X \geq 1\right)} = \frac{e^{-0.0614} \cdot \frac{0.0614^1}{1!}}{1 - Pr\left(X = 0\right)} = 0.97$$

תוצאה זו צפויה גם כן שהרי הסיכוי שיהיו זוכים הוא נמוך מאוד ולכן אם כבר מישהו זכה, הסיכוי שיש הרבה זוכים זעום. □

תרגיל 6.34. בתרגיל 6.8 ראינו שהפונקציה היוצרת מומנטים של משתנה מקרי בינומי הינה

$$M_X\left(t\right) = \left(e^t p + 1 - p\right)^n = \left(1 + \frac{\lambda e^t}{n} - \frac{\lambda}{n}\right)^n$$

כאשר n שואף לאינסוף מקבלים גבול מהצורה של משוואה 6.25 ולכן

$$\lim_{n \to \infty} M_X\left(t\right) = \lim_{n \to \infty} \left(1 + \frac{\lambda e^t}{n} - \frac{\lambda}{n}\right)^n = e^{-\lambda\left(1 - e^t\right)}$$

כשם שהתפלגות בינומית היא בקירוב פואסונית כאשר $n \to \infty$, כך גם הפונקציה היוצרת מומנטים של משתנה מקרי בינומי שואף לפונקציה היוצרת מומנטים של משתנה מקרי פואסוני. □

תרגיל 6.35.

1. ניקח n כדורים שאחד מהם שונה מהאחרים ונסדרם באופן אקראי בשורה. נסמן ב-X את מיקום הכדור השונה. מטעמי סימטריה, ההסתברות שהכדור המיוחד יהיה במקום ה-m הינה $\frac{1}{n}$ לכל $1 \leq m \leq n$, כלומר X מתפלג אחיד.
2. יהי $p = \frac{k}{m}$ מספר רציונלי, כאשר $0 < k < m$ טבעיים. נפזר בין m תאים באופן אקראי n כדורים ונסמן ב-Y את מספר הכדורים בתאים $1, 2, ...k$ (בסך הכל). ההסתברות של כל כדור להיות בתאים אלו הוא בדיוק p ולכן $Y \sim Bin(n, p)$.
3. יהי $p = \frac{k}{m}$ מספר רציונלי, כאשר $0 < k < m$ טבעיים. נכניס לכד k כדורים לבנים ו-$m - k$ כדורים שחורים. נוציא מהכד כדורים עם החזרה עד אשר לראשונה נוציא כדור לבן. נסמן ב-Z את הסיבוב בו הוצאנו כדור לבן לראשונה. $Z \sim Geom(p)$.

□

תרגיל 6.36. נסמן את היתרה של לקוח אקראי ב- $X \sim N\left(2200, 600^2\right)$ וכתמיד $Z \sim N(0, 1)$.

1. הסיכוי שללקוח יש יתרה הגבוהה מ-3000 ש"ח הינה:

$$Pr\left(X \geq 3000\right) = Pr\left(\frac{X-2200}{600} \geq \frac{800}{600}\right) = Pr\left(Z \geq \frac{4}{3}\right) = 1-\Phi\left(\frac{4}{3}\right) = 1-0.9066 = 0.0934$$

באותו אופן

$$Pr\left(X \leq 1400\right) = Pr\left(Z \leq \frac{1400-2200}{600}\right) = \Phi\left(-\frac{4}{3}\right) = 1-\Phi\left(\frac{4}{3}\right) = 0.0934$$

קיבלנו את אותה התוצאה שכן שני הערכים נמצאים במרחק זהה מהתוחלת ומשתנה מקרי נורמלי הוא סימטרי.
לסיום

$$Pr\left(X \leq 0\right) = Pr\left(Z \leq \frac{-2200}{600}\right) = \Phi\left(-3.66\right) = 1-\Phi\left(3.66\right) \sim 0.0001$$

2. אנחנו מחפשים את האחוזון ה-95 של ההתפלגות, כלומר את x_{95} כך ש-$Pr\left(X \leq x_{95}\right) = 0.95$. נתקנן לקבלת משתנה מקרי נורמלי סטנדרטי ונקבל $Pr\left(Z \leq \frac{X_{95}-2200}{600}\right) = 0.95$. מהטבלה ניתן לראות שהאחוזון ה-95 של Z הינו 1.65 ולכן $\frac{x_{95}-2200}{600} = 1.65$ כך שסכום הסף הנדרש על מנת להינות מהריבית המיוחדת הינו $x_{95} = 1.65 \cdot 600 + 2200 = 3190$.
3. ההסתברות שללקוח מקרי יש את הריבית שחושבה בסעיף הקודם היא 0.05 באופן בלתי תלוי בלקוחות האחרים. לכן מספר הלקוחות עם הריבית המיוחדת מתפלג $Bin(30, 0.05)$ והתוחלת היא $30 \cdot 0.05 =$ 1.5.

□

תרגיל 6.37. המשתנה המקרי X מתפלג אחיד ולכן פונקציית ההתפלגות המצטברת שלו הינה

$$F_X\left(u\right) = \begin{cases} 0 & u < 1 \\ \frac{u-1}{2} & 1 \leq u \leq 3 \\ 1 & u > 3 \end{cases}$$

1. נחשב את פונקציית ההתפלגות המצטברת של $Y = \ln X$:

$$\begin{aligned} F_Y(t) &= Pr\left(Y \leq t\right) = Pr\left(\ln X \leq t\right) \\ &= Pr\left(X \leq e^t\right) = F_X\left(e^t\right) = \begin{cases} 0 & e^t < 1 \\ \frac{e^t-1}{2} & 1 \leq e^t \leq 3 \\ 1 & e^t > 3 \end{cases} \end{aligned}$$

נשאר לסדר את הפונקציה שקיבלנו ונקבל לסיכום:

$$F_Y\left(t\right) = \begin{cases} 0 & t < 0 \\ \frac{e^t-1}{2} & 0 \leq t \leq \ln 3 \\ 1 & t > \ln 3 \end{cases}$$

2. לצורך חישוב התוחלת והשונות יש למצוא את פונקציית הצפיפות ולכן נגזור את פונקציית ההתפלגות המצטברת:

$$f_Y(t)=\frac{dF_Y(t)}{dt}=\begin{cases}\frac{e^t}{2} & t\in[0,\ln 3]\\ 0 & t\notin[0,\ln 3]\end{cases}$$

לכן ניתן לחשב את התוחלת והשונות לפי ההגדרה:

$$\mathbb{E}\left(Y\right)=\int_0^{\ln 3} t\cdot\frac{e^t}{2}\,dt=0.648$$
$$\mathbb{E}\left(Y^2\right)=\int_0^{\ln 3} t^2\cdot\frac{e^t}{2}\,dt=0.515$$

ולכן

$$\mathbb{V}\left(Y\right)=\mathbb{E}\left(Y^2\right)-\mathbb{E}^2\left(Y\right)=0.515-0.648^2=0.095$$

הערה: ניתן היה לחשב את התוחלת והשונות תוך שימוש בנוסחא לתוחלת של פונקציה של משתנה מקרי (נוסחא 5.5), גם בלי חישוב ההתפלגות של Y.

3. נשתמש במשפט הגבול המרכזי ולצורך כך נוציא ln משני האגפים כדי להפוך את המכפלה לסכום:

$$\begin{aligned}Pr\left(X_1\cdot\ldots\cdot X_{20}>1500000\right)&=Pr\left(\ln X_1+\ldots+\ln X_{20}>\ln 1500000\right)\\&=Pr\left(Y_1+\ldots+Y_{20}>\ln 1500000\right)\end{aligned}$$

כאשר לכל $i\in\{1,\ldots,20\}$ המשתנה המקרי $Y_i=\ln X_i$ מתפלג לפי ההתפלגות של Y ולכן ניתן להשתמש במשפט הגבול המרכזי:

$$\begin{aligned}Pr\left(X_1\cdot\ldots\cdot X_{20}>1500000\right)&=Pr\left(\overline{Y}_{20}>\frac{\ln 1500000}{20}\right)\\&=Pr\left(\frac{\overline{Y}_{20}-0.648}{\sqrt{0.095/20}}>\frac{\frac{\ln 1500000}{20}-0.648}{\sqrt{0.095/20}}\right)=1-\Phi(0.9148)\\&=1-0.8186=0.1814\end{aligned}$$

□

תרגיל 6.38. מספר הפעמים שהמטבע נפל על עץ מתפלג לפי $X\sim Bin\left(1000,\frac{1}{2}\right)$. אילון מעוניין על סמך תוצאות הניסוי לבחור בין איטליה לצרפת, כך שההסתברות לבחור את איטליה (כלומר ההסתברות לקבל תוצאות המתאימות לבחירה באיטליה) תהיה $\frac{1}{\pi}$. דרך נוחה לבצע מיפוי זה ולקבוע כלל החלטה הוא על ידי בחירת ערך מסויים c כך שאם $X\le c$ אילון יסע לאיטליה ואחרת יסע לצרפת. לכן ערך זה צריך לקיים $Pr\left(X\le c\right)=\frac{1}{\pi}$.

נשתמש בקירוב הנורמלי להתפלגות בינומית ונקבל:

$$\frac{1}{\pi}=Pr\left(X\leq c\right)=\Phi\left(\frac{c+0.5-500}{\sqrt{1000\cdot0.5\cdot0.5}}\right)=\Phi\left(\frac{c-499.5}{15.811}\right)$$

לפי הטבלה הערך המתאים להסתברות של $0.318=\frac{1}{\pi}$ הוא -0.475 כי $\Phi\left(-0.475\right)=1-\Phi\left(0.475\right)=$ $0.317=1-0.683$ וזה הערך הקרוב ביותר ל-$\frac{1}{\pi}$ הנמצא בטבלה. לכן המשוואה שעל c לקיים הינה $-0.475=\frac{c-499.5}{15.811}$ ופתרונה $c=491.98$. מאחר ו-c חייב להיות מספר שלם (מייצג את מספר העצים שהתקבלו) נבחר $c=492$ ובקירוב טוב בהסתברות $\frac{1}{\pi}$ אכן תיבחר איטליה. באופן דומה, על ידי הטלת מטבע מספר רב של פעמים וספירת מספר העצים שהתקבלו, ניתן לממש כל משתנה מקרי בדיד. □

תרגיל 6.39. 1. נחשב את פונקציית ההתפלגות המצטברת של Y:

$$F_{Y}(t)=Pr\left(Y\leq t\right)=Pr\left(F^{-1}(U)\leq t\right)$$

נשים לב שמתכונות פונקציית ההתפלגות המצטברת אנו יודעים שהיא עולה חלש. מאחר והיא הפיכה הרי שהיא לא קבועה באף קטע ולכן היא עולה חזק תמיד. לכן ניתן להפעילה על שני אגפי אי-השוויון $F^{-1}(U)\leq t$ ונקבל $U=F_{X}\left(F^{-1}(U)\right)\leq F_{X}(t)$. נציב זאת בנוסחא שקיבלנו קודם ונקבל $F_{Y}(t)=Pr\left(U\leq F_{X}(t)\right)=F_{X}(t)$ שכן ההסתברות ש-U קטן מ-$F_{X}(t)\in[0,1]$ היא בדיוק $F_{X}\left(t\right)$ (התפלגות אחידה). לכן לכל t מתקיים $F_{Y}(t)=F_{X}(t)$ ולמשתנים המקריים X ו-Y יש אותה פונקציית התפלגות מצטברת כלומר הם שווי התפלגות.

2. לכל $x\in\mathbb{R}$ מתקיים $x=-\frac{1}{\lambda}\ln\left(1-\left(1-e^{-\lambda x}\right)\right)=f\left(1-e^{-\lambda x}\right)=f\left(g(x)\right)$ ולכן הפונקציות הופכיות זו לזו.

3. נשים לב ש-g היא פונקציית ההתפלגות המצטברת של משתנה מקרי מעריכי עם פרמטר λ ולפי הסעיף הקודם זוהי פונקציה הפיכה, לכן לפי הסעיף הראשון $X=f(U)$ מתפלג גם כן מעריכית עם פרמטר λ.

4. ראינו שאם $x_{1},x_{2},...,x_{n}$ נלקחו מהתפלגות אחידה אז $f\left(x_{1}\right),f\left(x_{2}\right),...,f\left(x_{n}\right)$ יהיו דגימות מהתפלגות מעריכית עם פרמטר λ. לכן, אם אנחנו משתמשים בתוכנה שאינה יודעת להגריל מספרים לפי התפלגות מעריכית אך מעוניינים בכל זאת לעשות זאת, ניתן להגריל מספרים שמתפלגים אחיד בין 0 ל-1 ולהפעיל עליהם את הפונקציה f בכדי לקבל מספרים המתפלגים מעריכית.

□

תרגיל 6.40. יהי $a<c<b$. לכל $t>0$ עבורו $c+t<b$ יתקיים לפי הנתון

$$F_{X}\left(c+t\right)-F_{X}\left(c\right)=Pr\left(c\leq X\leq c+t\right)=g\left(t\right)$$

נגזור את שני האגפים לפי c ונקבל $f_{X}(c+t)-f_{X}(c)=0$ שכן אגף ימין כלל לא תלוי ב-c. לכן לכל $t\in(0,b-c)$ מתקיים $f_{X}(c+t)=f_{X}(c)$ כלומר בקטע $[c,b]$ הצפיפות קבועה. הוכחה זו נכונה לכל $c\in(a,b)$ לכן הצפיפות קבועה גם בכל הקטע $[a,b]$ וזו בדיוק ההגדרה של משתנה מקרי אחיד רציף. □

7 משתנים מקריים דו-ממדיים

בפרק 4 הוחלפה הטרמינולוגיה ממאורעות למשתנים מקריים, בטענה שבניסויים רבים ישנו גודל אחד מעניין ולכן טבעי לדון רק בגודל זה ובערכים אותו הוא מקבל. בפועל, לרוב ישנם מספר גדלים אשר מעניינים אותנו בו זמנית. כך היה למשל בתרגיל 4.4 שם הוגדרו שני משתנים מקריים עבור אותו הניסוי של הטלת שתי קוביות. בתרגיל זה חושבו ההתפלגויות של כל אחד מהמשתנים המקריים בנפרד אך ברור שישנו קשר בין המשתנים המקריים (הם מוגדרים על אותן שתי הטלות קוביה) ומעניין לחקור את הקשר ביניהם. דוגמא נוספת היא בחירה מקרית של אדם מהאוכלוסיה ומדידת גובהו ומשקלו. שני משתנים מקריים אלו ניתנים לחישוב בנפרד אך יש עניין רב במציאת קשרים ביניהם. למשל, כיצד מתפלג גובהו של האדם הנבחר אם ידוע משקלו? מהו המשקל הממוצע של אנשים הגבוהים מ-2 מטרים? האם אנשים גבוהים מהממוצע שוקלים גם יותר מהממוצע או להפך? בפרק זה נדון בקשרי הגומלין בין שני משתנים מקריים המוגדרים עבור אותו ניסוי ונענה על שאלות מסוג זה ודומות להן.

ההרחבה הזו ממשתנה מקרי אחד לשני משתנים מקריים, המכונים משתנים מקריים דו-ממדיים, היא טבעית ומזכירה מאוד את הנעשה בפרק 2. שם דנו בזוגות של מאורעות, חישבנו הסתברויות של חיתוכים, הסתברויות של התניות וכדומה. גם כאן נעסוק בהסתברויות של חיתוכים של מאורעות (מה הסיכוי ש-X מקבל ערך מסויים וגם Y מקבל ערך מסויים), התניות של ערכי משתנה מקרי אחד על ערכי השני וכדומה. כמקודם, נחלק את הדיון למשתנים מקריים דו-ממדיים בדידים ורציפים מאחר והטכניקות המתמטיות יהיו שונות. לסיום נציג שיטה נוספת לחישוב תוחלת ונשלים חוב ישן בכך שנציג את הנוסחא המלאה לשונות של סכום משתנים מקריים.

7.1 משתנים מקריים דו-ממדיים בדידים

נסתכל על ניסוי מסויים ונגדיר עבורו שני משתנים מקריים. כך עשינו למשל בתרגיל 4.4, שם הטלנו קוביה פעמיים והגדרנו את X להיות סכום ההטלות ואת Y להיות הפרשן בערכו המוחלט. בתרגיל זה מצאנו את ההתפלגות של כל משתנה מקרי תוך כדי התעלמות מהמשתנה המקרי השני המוגדר בתרגיל. בפרק זה נתעניין בהתפלגות המושתפת שלהם ונדון במאורעות מהצורה $\{X=k,Y=l\}$. התפלגות זו, העוסקת בשני משתנים מקריים בבת אחת, מכונה ההתפלגות המשותפת של X,Y.

הגדרה 7.1 (התפלגות משותפת). יהי Ω מרחב מדגם כלשהו ונגדיר עבורו שני משתנים מקריים בדידים X,Y. ההתפלגות המשותפת של X,Y היא פונקציה המתאימה לכל זוג של ערכים אפשריים של X ושל Y את ההסתברות לקבלם בו זמנית:

$$P_{X,Y}\left(k,l\right)=Pr\left(X=k,Y=l\right) \tag{7.1}$$

כאשר $k\in supp\left(X\right),l\in supp\left(Y\right)$. כאשר אחד הארגומנטים אינו בתומך, ההסתברות הינה כמובן 0.

הדרך הבסיסית לחישוב ההתפלגות המשותפת היא על ידי ניתוח המאורע $\{X=k,Y=l\}$ וחישוב הסתברותו במרחב המדגם Ω, למשל על ידי קומבינטוריקה.

בעיה 7.1. בקלמר יש 2 עטים אדומים, אחד ירוק ואחד כחול. מוציאים מהקלמר את העטים בזה אחר זה עד שיוצא העט הירוק. נסמן ב-X את מספר העטים שהוצאו בניסוי (כולל הירוק) וב-Y את מספר העטים האדומים שהוצאו בניסוי. מהי ההתפלגות המשותפת של X,Y?

הוכחה. נתחיל בחישוב התומכים. מספר העטים הכולל שיצאו בניסוי הוא לפחות 1 ולכל היותר 4, לכן $supp\left(X\right) = \{1,2,3,4\}$. מספר העטים האדומים שנבחרו בניסוי הינו לכל היותר 2 ולפחות 0, לכן $supp\left(Y\right) = \{0,1,2\}$. לאור זאת, יש בסך הכל 12 זוגות אפשריים של k,l עבורם יש לחשב את ההסתברות. מאחר ומספר האפשרויות מאוד קטן, ניתן לחשב את האפשרויות אחת-אחת ואין צורך בפיתוח נוסחא כללית.

נבחן מספר מאורעות לדוגמא. המאורע $\{X=1,Y=0\}$ מתקיים כאשר יצא בדיוק עט אחד בסך הכל ו-0 עטים אדומים וזה ייתכן רק כאשר העט הראשון שיצא היה הירוק. לפיכך, $P_{X,Y}(1,0)=\frac{1}{4}$. באופן דומה, לכל $l>0$ מתקיים $P_{X,Y}\left(1,l\right)=0$ שכן לא ייתכן שגם יצא בדיוק עט אחד, העט הירוק, וגם יצאו עטים אדומים. באופן דומה, המאורע $\{X=2,Y=1\}$ מתקיים כאשר יצאו בדיוק שני עטים, מתוכם אחד אדום ואחד ירוק, שבהכרח יצא שני ואחרון. ההסתברות שהעט הראשון אדום הינה $\frac{1}{2}$ וההסתברות שהעט השני ירוק (לאחר שיצא כבר עט אדום) הינה $\frac{1}{3}$ לכן $P_{X,Y}\left(2,1\right)=\frac{1}{2}\cdot\frac{1}{3}=\frac{1}{6}$. באופן דומה ניתן להמשיך ולתאר כל מאורע בתור סידור כלשהו של עטים שיצאו כך שהעט האחרון חייב להיות ירוק ולהציג את ההתפלגות המשותפת בטבלה:

$Y\backslash X$	**1**	**2**	**3**	**4**
0	$\frac{1}{4}$	$\frac{1}{12}$	**0**	**0**
1	**0**	$\frac{1}{6}$	$\frac{1}{6}$	**0**
2	**0**	**0**	$\frac{1}{12}$	$\frac{1}{4}$

□

כשם שהתייחסנו אל התומך של משתנה מקרי כאל מרחב מדגם חדש, ניתן לחשוב גם על הטבלה כמרחב המדגם החדש של הניסוי. בכל פעם שנערוך את הניסוי ייקבע ערך עבור X וערך עבור Y, כך שתיבחר אחת מהמשבצות בטבלה לפי ההסתברויות הרשומות עליהן. היות ובכל ניסוי נבחרת בדיוק משבצת אחת, סכום ההסתברויות שרשומות במשבצות צריך להיות 1 (ודאו זאת).

משפט 7.1. *עבור משתנה מקרי דו-ממדי בדיד X,Y מתקיים*

$$\sum_{k\in supp(X)}\sum_{l\in supp(Y)}P_{X,Y}\left(k,l\right)=1 \tag{7.2}$$

הוכחה. ההוכחה מבססת על כך שכל המאורעות מהצורה $\{X=k,Y=l\}$ זרים בזוגות (לא ייתכן שבו זמנית $\{X=3,Y=1\}$ וגם $\{X=3,Y=2\}$, למשל) ומרכיבים את מרחב המדגם כולו (בכל ניסוי מתקבל ערך כלשהו עבור X וערך כלשהו עבור Y). לשם הקיצור, נשמיט את הרישום הפורמלי של ההוכחה. □

כאשר התומכים של המשתנים המקריים גדולים יותר (וכמובן, כאשר הם אינסופיים) אין טעם לנסות לחשב את כל המקרים אחד-אחד ולא ניתן להציג את ההתפלגות באמצעות טבלה אלא יש למצוא נוסחא כללית המתארת את ההתפלגות הדו-ממדית.

בעיה 7.2. מטילים קוביה שוב ושוב. נסמן ב-X את מספר ההטלות עד ההופעה הראשונה של המספר 5 וב-Y את מספר ההטלות עד ההופעה הראשונה של מספר אי-זוגי (כולל 5). מהי ההתפלגות המשותפת של X,Y?

הוכחה. מספר ההטלות עד ההצלחה הראשונה (אי-זוגי או 5, בהתאם למשתנה המקרי) יכול להיות כל מספר טבעי ולכן במקרה הזה התומך של כל משתנה מקרי הוא $\mathbb{N}$ ובוודאי שאין אפשרות לערוך טבלה של כל האפשרויות. יהיו $k,l\in\mathbb{N}$ ונחשב את ההסתברות של המאורע $\{X=k,Y=l\}$. נשים לב כי לא ייתכן המצב $k<l$ שכן

5 הוא מספר אי-זוגי לכן אם הוא יצא לראשונה בהטלה ה-k אזי המספר האי-זוגי הראשון היה חייב לצאת קודם לכן או, לכל המאוחר, הוא יצא בהטלה זו. לכן $l \leq k$. נבחן את שתי האפשרויות בנפרד:

כאשר $k = l$ משמעות המאורע $\{X = k, Y = l\} = \{X = Y = k\}$ **היא ש-5 הראשון יצא בהטלה ה-k** וזוהי גם ההטלה הראשונה בה יצא מספר אי-זוגי. לכן ב-$k - 1$ **ההטלות הראשונות יצאו מספרים זוגיים.** ההסתברות לכל מספר זוגי היא $\frac{1}{2}$ ולכן ההסתברות הכוללת של מאורע זה הינה

$$P_{X,Y}(k,k) = \left(\frac{1}{2}\right)^{k-1} \cdot \frac{1}{6}$$

כאשר $\frac{1}{6}$ היא ההסתברות שבהטלה מספר k יצא 5.

כאשר $k < l$ משמעות המאורע $\{X = k, Y = l\}$ הוא שב-$l - 1$ **ההטלות הראשונות יצאו מספרים זוגיים.** לאחר מכן, בהטלה ה-l יצא מספר אי-זוגי שאינו 5. בהטלות הבאות, יצאו מספרים זוגיים או אי-זוגיים פרט ל-5, שכן 5 יצא לראשונה רק בהטלה ה-k. בין ההטלה ה-l להטלה ה-k יש בסך הכל $k - l - 1$ הטלות ולכן הסתברות של מאורע זה הינה

$$P_{X,Y}(k,l) = \left(\frac{1}{2}\right)^{l-1} \cdot \frac{2}{6} \cdot \left(\frac{5}{6}\right)^{k-l-1} \cdot \frac{1}{6}$$

לסיכום, פונקציית ההתפלגות המשותפת לכל $k, l \in \mathbb{N}$ תהא:

$$P_{X,Y}(k,l) = \begin{cases} \left(\frac{6}{10}\right)^{l} \cdot \left(\frac{5}{6}\right)^{k} \cdot \frac{20}{6} & l < k \\ \left(\frac{1}{2}\right)^{k} \cdot \frac{1}{3} & l = k \\ 0 & l > k \end{cases}$$

גם כאן, סכימה כפולה על כל הערכים האפשריים תסתכם ל-1. $\square$

תרגיל 7.1. חשבו את ההתפלגות המשותפת של X, Y המוגדרים בתרגיל 4.4.

תרגיל 7.2. מטילים קוביה שוב ושוב. נסמן ב-X את מספר ההטלות עד ההופעה הראשונה של 6 וב-Y את מספר ההטלות עד ההופעה הראשונה של מספר אי-זוגי. מהי ההתפלגות המשותפת של X, Y?

תרגיל 7.3. תייר מגיע לעיר בה פועלות שלוש מסעדות. בכל ערב התייר בוחר באחת המסעדות באקראי וסועד בה. נסמן ב-$\mathbf{1}_A$ את האינדיקטור לכך שהתייר סעד במסעדה A וב-$\mathbf{1}_B$ את האינדיקטור לכך שהתייר סעד במסעדה B. מהי ההתפלגות המשותפת של $\mathbf{1}_A, \mathbf{1}_B$ אם התייר בסך הכל נשאר בעיר 5 ערבים?

נחזור לבעיה 7.1 ונניח שאנחנו מתעניינים רק ב-Y, מספר העטים האדומים שיצאו מהקלמר. כלומר, כל שמעניין אותנו הוא ההתפלגות של המשתנה המקרי Y, תוך התעלמות מהמשתנה המקרי X והקשר האפשרי ביניהם. ניתן למצוא התפלגות זו בשתי גישות שונות. בגישה הראשונה נחזור לפרק 4 וננתח את ההתפלגות של Y בתור משתנה מקרי חד-ממדי רגיל, תוך התעלמות מוחלטת מכך שבכלל הוגדר משתנה מקרי נוסף בשאלה (כפי שנעשה בכל השאלות של פרק 4 בהן הוגדרו שני משתנים מקריים וחושבה התפלגות של כל אחד מהם בנפרד, דוגמת תרגיל 4.4).

חישוב ההתפלגות של Y תוך התעלמות מ-X. הניסוי הוא שליפה של עטים מקלמר ללא החזרה. כפי שראינו בפרק 3.2.1, ניתן להתייחס לניסוי זה כסידור פריטים בשורה כאשר כל העטים שלפני העט הירוק "יצאו" וכל העטים שאחריו "לא יצאו". מעניין אותנו כמה עטים אדומים יצאו ולכן ניתן לצמצם את מרחב המדגם ולהסתכל רק על שלושת הפריטים: עט ירוק ושני עטים אדומים. ההסתברות שהעט הירוק נמצא בכל אחד מהמקומות האפשריים זהה ולכן ההסתברות של כל ערך זהה, כך ש-$Y \sim U\{1,2,3\}$. $\square$

בגישה השנייה נשתמש בהתפלגות המשותפת של המשתנים המקריים בכדי לחשב את ההתפלגות של Y. לאור הטבלה, הערך $Y=1$, למשל, יכול להתקבל באחת משתי דרכים: $\{X=2,Y=1\}$ או $\{X=3,Y=1\}$. מאורעות אלו זרים ואיחודם הוא $\{Y=1\}$ לכן

$$P_Y\left(1\right)=P_{X,Y}\left(2,1\right)+P_{X,Y}\left(3,1\right)=\frac{1}{6}+\frac{1}{6}=\frac{1}{3}$$

באופן דומה ניתן לחשב את ההסתברות של כל ערך אחר בתומך של Y ולקבל את ההתפלגות שלו. דרך נוחה מאוד להציג את החישוב היא על ידי הוספת עמודה בטבלה בה רשומים כל הסכומים של הערכים באותה השורה. באופן דומה, הוספת שורה בה רשומים הסכומים של כל ההסתברויות באותה העמודה תיתן את ההתפלגות של X תוך התעלמות מכך שהמשתנה המקרי Y מוגדר. ההתפלגות המתקבלת באופן זה מכונה ההתפלגות השולית של Y (ובהתאמה, ההתפלגות השולית של X):

$Y\backslash X$	**1**	**2**	**3**	**4**	$P_Y\left(l\right)$
0	$\frac{1}{4}$	$\frac{1}{12}$	**0**	**0**	$\frac{1}{3}$
1	**0**	$\frac{1}{6}$	$\frac{1}{6}$	**0**	$\frac{1}{3}$
2	**0**	**0**	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$
$P_X\left(k\right)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	**1**

לאור הטבלה, לא רק ש-Y מתפלג אחיד כפי שחישבנו מקודם אלא גם X מתפלג אחיד (כיצד ניתן לנמק זאת מבלי לחשב במפורש את ההתפלגות?).

הגדרה 7.2 (התפלגות שולית). יהי X,Y משתנה מקרי דו-ממדי בדיד עם פונקציית התפלגות משותפת $P_{X,Y}$. ההתפלגות השולית של X מוגדרת בתור ההתפלגות של X כמשתנה מקרי בדיד וניתנת לחישוב לכל $k\in supp\left(X\right)$ על ידי

$$P_X\left(k\right)=\sum_{l\in supp\left(Y\right)}P_{X,Y}\left(k,l\right) \tag{7.3}$$

באופן דומה מוגדרת ההתפלגות השולית של Y.

במעבר מהתפלגות משותפת להתפלגות שולית, ההופעות של המשתנה המקרי השני נעלמות (שכן נסכמים כל הערכים האפשריים שלו) והתוצאה היא התפלגות חד-ממדית של משתנה מקרי בדיד חד-ממדי בדומה לפרק 4. כפי שהדגמנו לעיל, את ההתפלגות השולית ניתן לחשב או על ידי הסכימה המופיעה בנוסחא 7.3 או על ידי מציאת התפלגותו כאילו הוא המשתנה המקרי היחיד המוגדר בבעיה תוך התעלמות מהמשתנה המקרי השני. כך למשל, בבעיה 7.2 קל לראות שכל אחד מהמשתנים המקריים מוגדר כמספר החזרות על ניסוי עם הסתברות קבועה להצלחה עד ההצלחה הראשונה, ולכן כל אחד מהם הוא משתנה מקרי גאומטרי אך עם פרמטר שונה: $X\sim Geom\left(\frac{1}{6}\right)$ ו-$Y\sim Geom\left(\frac{1}{2}\right)$.

הצעד הטבעי הבא הוא להגדיר התניה של משתנה מקרי אחד על אחר, בדומה לאופן בו הוגדרה התניה של מאורעות בפרק 2. התניה זו יכולה להתרחש במציאות כאשר ידוע ערכו של אחד המשתנים המקריים אך לא של האחר ומעוניינים לדעת מהי ההתפלגות של המשתנה המקרי האחר בהינתן מידע זה. למשל, נניח שמבצעים את הניסוי המתואר בבעיה 7.1 ומגלים שבסוף הניסוי הוצאו שני העטים האדומים מהקלמר, כלומר $Y=2$. בהתאם לטבלה, בהכרח $X\geq 3$ שכן על מנת להוציא שני עטים אדומים ואת העט הירוק, יש להוציא לפחות 3 עטים. בהינתן מידע זה על Y ההסתברויות של הערכים השונים בתומך משתנות. למשל, ההסתברות של המאורע $\{X=3\}$ כאשר המאורע $\{Y=2\}$ התרחש תהא

$$Pr\left(\{X=3\}|\{Y=2\}\right)\overset{eq.\,2.9}{=}\frac{Pr\left(\{X=3\}\cap\{Y=2\}\right)}{Pr\left(\{Y=2\}\right)}=\frac{\frac{1}{12}}{\frac{1}{3}}=\frac{1}{4}$$

אם נבצע חישוב זה לכל ערך אפשרי של X נקבל את ההתפלגות שלו בהינתן $Y = 2$. התפלגות זו של משתנה מקרי כאשר ערכו של משתנה מקרי אחר ידוע מכונה התפלגות מותנית.

הגדרה 7.3 (התפלגות מותנית)**.** יהי X, Y משתנה מקרי דו-ממדי בדיד עם פונקציית התפלגות משותפת $P_{X,Y}$. לכל $l \in supp\,(Y)$ נגדיר את ההתפלגות המותנית של X בהינתן $Y = l$ בתור

$$P_{X|Y=l}\,(k) = \frac{P_{XY}\,(k,l)}{P_Y\,(l)} \tag{7.4}$$

כאשר P_Y היא ההתפלגות השולית של Y ו-$k \in supp\,(X)$. נסמן את המשתנה המקרי המתקבל על ידי $X|Y = l$ כאשר $l \in \mathbb{R}$ הוא מספר שנקבע בבעיה ו-$X|Y$ לציון התפלגות המותנית באופן כללי כאשר l הוא רק פרמטר.

נוסחא 7.4 נובעת ישירות מההגדרה של הסתברות מותנית (נוסחא 2.9) שכן המונה הוא בדיוק החיתוך בין המאורעות $\{X = k\}$ ו-$\{Y = l\}$ ואילו המכנה הוא המאורע עליו מתנים. חשוב מאוד לשים לב שבאגף ימין של הנוסחא מופיעים שני המשתנים המקריים אך תפקידיהם אינם זהים. בנוסחא זו X הינו המשתנה המקרי הלא ידוע שאת הסתברויותיו מחשבים ואילו Y ידוע ולכן ערכו הוא פרמטר של הפונקציה. לכן, נוסחא זו מתארת למעשה משתנה מקרי חד-ממדי בדיד: $X|Y = l$. ההבדל בין משתנה מקרי זה לבין X (שאת התפלגותו חישבנו בנוסחא 7.3) הוא שההתפלגות השולית מניחה שכלום לא ידוע על תוצאות הניסוי ואילו ההתפלגות המותנית מתבססת על כך שהערך של Y ידוע. זהו גם ההבדל בין חישוב ההסתברות של המאורע A לבין חישוב ההסתברות של מאורע מהצורה $A|B$.

התפלגות מותנית בבעיה 7.1. נחשב את ההתפלגות המותנית של X בהינתן כל ערך אפשרי של Y לפי נוסחא 7.4. למשל אם $Y = 0$ הרי שהערכים האפשריים עבור X הינם 1 או 2 וההתפלגות המותנית הינה

$$P_{X|Y=0}\,(k) = \begin{cases} \frac{\frac{1}{12}}{\frac{1}{3}} = \frac{1}{4} & k = 1 \\ \frac{\frac{1}{4}}{\frac{1}{3}} = \frac{3}{4} & k = 2 \end{cases}$$

באופן דומה, אם $Y = 1$ הרי שההתפלגות המותנית תהא

$$P_{X|Y=1}\,(k) = \begin{cases} \frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2} & k = 2 \\ \frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2} & k = 3 \end{cases}$$

אם $Y = 2$ אזי

$$P_{X|Y=2}\,(k) = \begin{cases} \frac{\frac{1}{4}}{\frac{1}{3}} = \frac{3}{4} & k = 3 \\ \frac{\frac{1}{12}}{\frac{1}{3}} = \frac{1}{4} & k = 4 \end{cases}$$

כפי שקורה רבות במשתנים מקריים בדידים, ניתן לחשב את ההסתברויות גם על ידי חישוב קומבינטורי. למשל, בהינתן $Y = 2$, העט הירוק חייב להיות במקום השלישי או הרביעי ואילו שני עטים אדומים חייבים להיות לפניו. יש רק סידור אחד אפשרי בו העט הירוק במקום השלישי ויש 3 סידורים אפשריים בהם העט הירוק במקום הרביעי (כחול במקום 1, כחול במקום 2 וכחול במקום 3) לכן מרחב המדגם כולל 4 סידורים בשורה שווי הסתברות והסיכוי שהעט הירוק במקום השלישי הוא בדיוק רבע. $\square$

לאור חישובים אלו, בכל בעיה דו-ממדית אפשר לשאול עבור כל משתנה מקרי שתי שאלות: כיצד אותו משתנה מקרי מתפלג בניסוי (כלומר, מהי ההתפלגות השולית) וכיצד משתנה מקרי זה מתפלג בניסוי כאשר ערכו של המשתנה השני האחר ידוע (ההתפלגות המותנית). מצב מיוחד מתרחש כאשר התפלגויות אלו זהות. במקרה כזה, העובדה שידוע ערכו של משתנה מקרי אחד לא משנה את ההתפלגות של המשתנה המקרי השני, כלומר כל אחת מההסתברויות המותנות שווה להסתברות הלא מותנה. מאורעות המקיימים תכונה זו, קרי שההסתברות של מאורע מסויים שווה להסתברות של מאורע זה בהינתן המאורע האחר, נקראו מאורעות בלתי תלויים. משתנים מקריים המקיימים תכונה זו יכונו לפיכך משתנים מקריים בלתי תלויים.

הגדרה 7.4. משתנים מקריים דו-ממדיים בדידים X,Y נקראים בלתי תלויים אם

$$P_{XY}(k,l)=P_X(k)\,P_Y(l) \tag{7.5}$$

לכל $k\in supp(X)$ ו-$l\in supp(Y)$. אחרת, המאורעות נקראים תלויים. נוסחא 7.5 זהה להגדרת חוסר התלות המופיעה בנוסחא 2.19: ההסתברות של החיתוך שווה למכפלת ההסתברויות.
הערה: לתיאור תכונה זו בחלק מהספרים והקורסים משתמשים בטרמינולוגיה "בלתי תלויים סטטיסטית" או "בלתי תלויים סטוכסטית".

המשתנים המקריים המופיעים בבעיות 7.1 ו-7.2 תלויים. קל לוודא זאת על ידי חיפוש אפסים של פונקציית ההתפלגות המצטברת עבור ערכים המופיעים בתומכים של שני המשתנים המקריים. אפסים אלו מלמדים על מצבים שאפשריים לכל משתנה מקרי בנפרד אך לא ביחד. למשל, בבעיה 7.1 מתקיים $P_{X,Y}(1,1)=0$ שכן לא ייתכן שהעט הירוק היה הראשון שיצא ובו זמנית הוצא גם עט אדום אחד. מצד שני, כל אחד מהמרכיבים של המאורע אפשריים. ייתכן שהעט הירוק יהיה הראשון שיצא (בהסתברות $P_X(1)=\frac{1}{4}$) וייתכן שבתום תהליך ההוצאה יהיה עט אדום אחד מחוץ לקלמר (בהסתברות $P_Y(1)=\frac{1}{3}$). מתקיים

$$P_X(1)\cdot P_Y(1)=\frac{1}{12}\neq 0=P_{X,Y}(1,1)$$

ולכן המשתנים המקריים תלויים. באופן דומה, בבעיה 7.2 מתקיים $P_{X,Y}(3,7)=0$ למרות שההסתברויות של שני המאורעות $X=3$ ו-$Y=7$ שונות מאפס.

קיום אפסים בפונקציית ההתפלגות המשותפת הוא תנאי מספיק לכך שהמשתנים המקריים יהיו תלויים, אך אינו הכרחי. ייתכן מצב שבפונקציית ההתפלגות המשותפת של משתנים מקריים תלויים לא יהיו אפסים כלל. במקרה כזה אין מנוס אלא לבדוק את כל הערכים בתומך ולוודא שהם מקיימים את נוסחא 7.5. אם הנוסחא מתקיימת תמיד, אזי המשתנים המקריים בלתי תלויים. מספיק שעבור זוג ערכים בודד הנוסחא לא תתקיים כדי שהם יהיו תלויים. ניתן לפעמים לקצר בדיקה זו על ידי חישוב ההתפלגויות השוליות ולהשוות את המכפלה שלהן להתפלגות המצטברת כנוסחא ולא כערכים בודדים.

תרגיל 7.4. יהיו X,Y משתנה מקרי בדיד דו-ממדי עם פונקציית התפלגות משותפת נתונה על ידי

$$P_{X,Y}(x,y)=\begin{cases}\frac{xy}{27} & x=1,2;y=2,3,4\\ 0 & otherwise\end{cases}$$

האם הם בלתי תלויים?

תרגיל 7.5. נגדיר מרחב מדגם Ω באופן הבא: המרחב כולל את כל התמורות האפשריות של הקבוצה $\{1,2,3\}$ ובנוסף את השלשות $(1,1,1),(2,2,2),(3,3,3)$ ונניח כי זהו מרחב הסתברות סימטרי. נסמן X_i - הערך של הקוארדינטה ה-$i=1,2,3$ בוקטור שנבחר.

1. רשמו מפורשות את מרחב המדגם.
2. עבור כל $i \neq j$, האם X_i, X_j בלתי תלויים?
3. האם X_1, X_2, X_3 בלתי תלויים?

הערה: כדי ששלושה משתנים מקריים יהיו בלתי תלויים כל זוג צריך להיות בלתי תלוי ובנוסף צריך להתקיים

$$Pr\left(X=x,Y=y,Z=z\right)=P_X\left(x\right)P_Y\left(y\right)P_Z\left(z\right)$$

לכל $x \in supp\left(X\right)$, $y \in supp\left(Y\right)$ ו-$z \in supp\left(Z\right)$.

ידיעת התלות בין המשתנים המקריים יכולה לסייע בחישוב ההתפלגות המשותפת שלהם. למשל, כאשר נתון שהמשתנים המקריים בלתי תלויים ניתן לקבל את ההתפלגות המשותפת באמצעות מכפלת ההתפלגויות השוליות. אפשרות אחרת היא לנצל את התלות ביניהם כדי לבנות את ההתפלגות המשותפת באמצעות נוסחא 7.4. במקרה כזה, ההתפלגות המותנית היא הנתונה בבעיה (או יותר נוחה לחישוב) ומתוכה ניתן לחלץ את ההתפלגות המשותפת.

בעיה 7.3. נהג נוסע מתל אביב לחיפה במהירות מופרזת. במהלך הנסיעה חולף הנהג על פני שמונה מכמונות מהירות, כאשר כל מכמונת יכולה להיות אמיתית או דמה בהסתברות שווה וללא תלות במכמונות האחרות. כל מכמונת אמיתית יעילה ב-25%, כלומר זוהי ההסתברות שהיא תתפוס את הנהג ותשלח לו דו"ח. נסמן ב-X את מספר מכמונות המהירות האמיתיות על פניהן חולף הנהג וב-Y את מספר הקנסות שיקבל בתום הנסיעה.

1. האם משתנים מקריים אלו תלויים או בלתי תלויים? (ענו בלי לחשב).
2. מהי ההתפלגות השולית של X? מהי ההתפלגות המותנית $Y|X$?
3. מהי ההתפלגות המשותפת של X, Y?
4. מהי תוחלת עלות הקנסות הכוללת שעל הנהג לשלם אם כל קנס הוא בסך 250 ש"ח?

הוכחה. 1. משתנים מקריים אלו בוודאי תלויים. למשל, אם $X=0$ (כלומר אין מכמונות אמיתיות בכלל) הרי שבהכרח $Y=0$ (אין גם קנסות). לעומת זאת, אם יש מכמונות מהירות הרי שיש הסתברות כלשהי לקבל קנסות. הידיעה אודות הערך של X משנה את ההתפלגות של Y. בסעיפים הבאים נחשב את ההתפלגות המשותפת ונוכל לענות על שאלה זו גם מספרית.

2. לפי הנתונים, יש שמונה מכמונות מהירות וכל אחת מהן יכולה להיות אמיתית בהסתברות $\frac{1}{2}$ ללא תלות במכמונות האחרות, לכן $X \sim Bin\left(8,\frac{1}{2}\right)$. נניח שיש בדיוק X מכמונות מהירות אמיתיות. כל אחת תופסת את הנהג בהסתברות $\frac{1}{4}$ ולכן מספר הקנסות שהנהג יקבל בהינתן שיש X מכמונות אמיתיות מתפלג $Y|X \sim Bin\left(X,\frac{1}{4}\right)$.

3. לצורך חישוב ההתפלגות המשותפת ניעזר בנוסחא 7.4. נחלץ מנוסחא זו את ההתפלגות המשותפת ונקבל

$$P_{X,Y}\left(x,y\right)=P_{Y|X=x}\left(y\right)P_X\left(x\right)=\left(\binom{x}{y}\left(\frac{1}{4}\right)^y\left(\frac{3}{4}\right)^{x-y}\right)\left(\binom{8}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{8-x}\right)$$

הסוגריים השמאליים מתקבלים על ידי הצבת הערך y בהתפלגות המותנית שהינה התפלגות בינומית. הסוגריים הימניים מתארים את הצבת הערך x בהתפלגות הבינומית המתארת את מספר המכמונות האמיתיות. נשים לב כי מספר זה שונה מאפס רק כאשר $y \leq x$, שהרי לא ניתן לקבל יותר קנסות מאשר מספר המכמונות האמיתיות. נסדר את הנוסחא ונרשום תנאי זה מפורשות לקבלת ההתפלגות

המשותפת עבור $x, y \in \{0, \ldots, 8\}$:

$$P_{X,Y}(x,y) = \begin{cases} \frac{\binom{x}{y}\binom{8}{x}}{2^8}\left(\frac{1}{3}\right)^y\left(\frac{3}{4}\right)^x & y \leq x \\ 0 & y > x \end{cases}$$

4. היות שכל קנס שווה ל-250 ש"ח הרי שעלות הקנסות הכוללת הינה $250Y$ ובשאלה מבקשים לחשב את התוחלת $\mathbb{E}(250Y) = 250\mathbb{E}(Y)$. ניתן להמשיך חישוב זה בשלוש דרכים. דרך אחת היא לחשב את ההתפלגות השולית מפורשות, על ידי סכימה של כל ערכי ה-X המתאימים. כלומר, חישוב לכל $y \in \{0, ..., 8\}$ את ההסתברות של

$$P_Y(y) = \sum_{x=0}^{y} P_{X,Y}(x,y)$$

וחישוב התוחלת על ידי נוסחא 5.1. דרך זו אפשרית אבל מצריכה חישובים רבים ואינה נוחה במיוחד. דרך שניה היא על ידי פירוק לאינדיקטורים. נסמן ב-$\mathbf{1}_i$ את האינדיקטור לכך שמכמונות המהירות ה-$i \in \{1, \ldots, 8\}$ שלחה דו"ח לנהג. בכדי שזה יקרה נדרש שהיא תהיה אמיתית (בהסתברות $\frac{1}{2}$) ושתתפוס את הנהג (הסתברות של $\frac{1}{4}$ בהינתן שהיא אמיתית). לכן הסיכוי הכולל הינו

$$Pr(\mathbf{1}_i = 1) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

מספר הדו"חות הכולל הוא כמספר המכמונות ששלחו דו"ח, כלומר $Y = \sum_{i=1}^{8} \mathbf{1}_i$ ולכן

$$\mathbb{E}(Y) = \mathbb{E}\left(\sum_{i=1}^{8} \mathbf{1}_i\right) \overset{eq.\,5.7}{=} \sum_{i=1}^{8} \mathbb{E}(\mathbf{1}_i) \overset{eq.\,5.10}{=} \sum_{i=1}^{8} \frac{1}{8} = 1$$

כך שתוחלת הקנס הינה 250 ש"ח.

דרך שלישית, המתבססת על התלות בין Y ל-X מופיעה בתרגיל 7.14 בהמשך.

$\square$

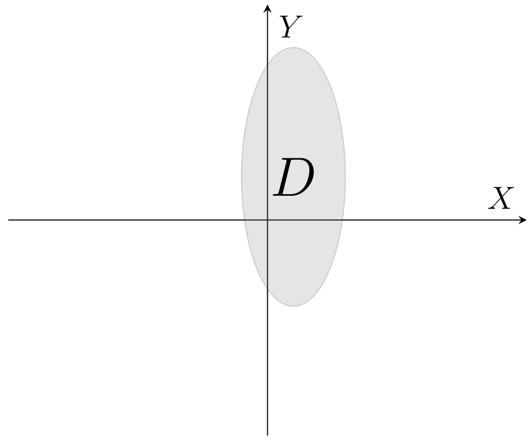
תרגיל 7.6. נתונה חפיסת קלפים הכוללת חמישה קלפים הממוספרים 1-5 ומסודרים באופן מקרי. נסמן ב-X את ערכו של הקלף העליון בחפיסה וב-Y את ערכו של הקלף התחתון. חשבו את ההתפלגות המשותפת של X, Y ואת ההתפלגויות השוליות של כל אחד ממשתנים מקריים אלו.

7.2 משתנים מקריים דו-ממדיים רציפים

באופן דומה נרחיב גם משתנים מקריים רציפים חד-ממדיים למקרה הדו-ממדי. במקרה זה אין פונקציית הסתברות (כזכור, ההסתברות לכל נקודה בודדת היא 0) ובמקומה הפונקציה שתהפוך לדו-מקומית היא פונקציית צפיפות ההסתברות. מסיבות שיובהרו בהמשך, פונקציית ההתפלגות המצטברת פחות שימושית במקרה הדו-ממדי. האינטואיציה שמאחורי פונקציית צפיפות ההסתברות הדו-ממדית זהה למקרה החד-ממדי. הפעם הצפיפות מפוזרת על פני מישור (מישור XY) ולא על פני הציר הממשי והערך של פונקציית צפיפות ההסתברות בכל נקודה מתארת עד כמה ההסתברות היחסית גבוהה בנקודה זו. גם כאן, לנקודות בודדות אין הסתברות והפעם גם לא לקטעים או ישרים אלא רק למקביליהם הדו-ממדיים: תחומים בעלי שטח. אלו יהיו המאורעות שאת הסתברותם נחשב והחישוב יבוצע כמו במקרה החד-ממדי – באמצעות אינטגרל. בסעיף זה נעקוב אחרי הסדר הכללי של הסעיף הקודם אך במקרה הרציף.

הגדרה 7.5 **(צפיפות משותפת).** עבור שני משתנים מקריים רציפים X, Y, הצפיפות המשותפת היא פונקציה המתאימה לכל זוג של ערכים אפשריים של X ו-Y את צפיפות ההסתברות בנקודה זו. פונקציה זו מסומנת על ידי $f_{XY}(x,y)$. במקרה הדו-ממדי מאורעות הם תחומים דו-ממדיים במישור וההסתברות לקבלם נתונה על ידי

$$Pr(D) = \iint_D f_{XY}(x,y)\ dxdy \tag{7.6}$$



נדגיש שנית שההסתברות לקבל נקודה או אפילו קו ישר היא אפס, ומאורעות בעלי הסתברות שונה מאפס חייבים להיות דו-ממדיים (כלומר צורות בעלות שטח).

לצפיפות המשותפת אותן תכונות של פונקציית הצפיפות החד-ממדית (משפט 4.4): הצפיפות צריכה להיות חיובית ומנורמלת. ההבדל הוא שהנרמול מתבצע על כל המישור, כלומר הדרישה שפונקציית צפיפות צריכה לקיים היא

$$\iint_{\mathbb{R}^2} f_{XY}(x,y)\ dxdy = 1 \tag{7.7}$$

כאשר נתונה פונקציית הצפיפות המשותפת, ניתן להפיק ממנה את הצפיפות החד-ממדית של כל משתנה מקרי על ידי אינטגרציה על המשתנה המקרי השני (במקרה הבדיד: סכימה). באופן זה תתקבל צפיפות חד-ממדית רגילה של המשתנה המקרי הרצוי ובאמצעותה ניתן לחשב את תוחלתו, שונותו ושאר החישובים הרצויים לגביו בלבד, כאילו המשתנה המקרי השני כלל לא היה מוגדר בשאלה. צפיפות זו מכונה צפיפות שולית.

הגדרה 7.6 **(צפיפות שולית).** יהי X, Y משתנה מקרי דו-ממדי רציף עם פונקציית צפיפות משותפת f_{XY}. הצפיפות השולית של X מוגדרת בתור ההתפלגות של X כמשתנה מקרי חד-ממדי רציף וניתנת לחישוב לכל $x \in \mathbb{R}$ על ידי

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\ dy \tag{7.8}$$

באופן דומה מוגדרת הצפיפות השולית של Y.

כאשר מחשבים את האינטגרל לצורך חישוב הצפיפות השולית, המשתנה x הוא קבוע ויש להתייחס אליו כאל פרמטר, למשל כאשר מחשבים את גבולות האינטגרל שיכולים להיות תלויים בו. האינטגרל מבוצע לפי y ולכן התשובה הסופית לא יכולה להיות תלויה בו ומתקבלת פונקציה של x בלבד, בדיוק כפי שראינו בפרק **4**.

בעיה 7.4. נתון משתנה מקרי דו-ממדי רציף X,Y בעל פונקציית צפיפות משותפת

$$f_{XY}\left(x,y\right)=\begin{cases} c\left(1+x^2-2xy^2\right) & 0\leq x,y\leq 1\\ 0 & \text{otherwise}\end{cases}$$

1. חשבו את ערכו של הקבוע c.
2. מצאו את הצפיפות השולית של X ואת הצפיפות השולית של Y.
3. מהי ההסתברות ש-$X+Y\leq 1$?

הוכחה. 1. נדרוש נרמול של פונקציית הצפיפות:

$$1=\iint_{\mathbb{R}^2} f_{XY}\left(x,y\right)\,dxdy=\int_0^1\int_0^1 c\left(1+x^2-2xy^2\right)\,dxdy=c$$

לכן בהכרח $c=1$.

2. נחשב לפי ההגדרה. ראשית, כאשר $x\notin[0,1]$ הצפיפות השולית היא 0 לפי תחום ההגדרה של הצפיפות המשותפת. אחרת, כאשר $x\in[0,1]$ נקבל

$$f_X\left(x\right)=\int_0^1 1+x^2-2xy^2\,dy=1+x^2-2x\left.\frac{y^3}{3}\right|_0^1=x^2-\frac{2x}{3}+1$$

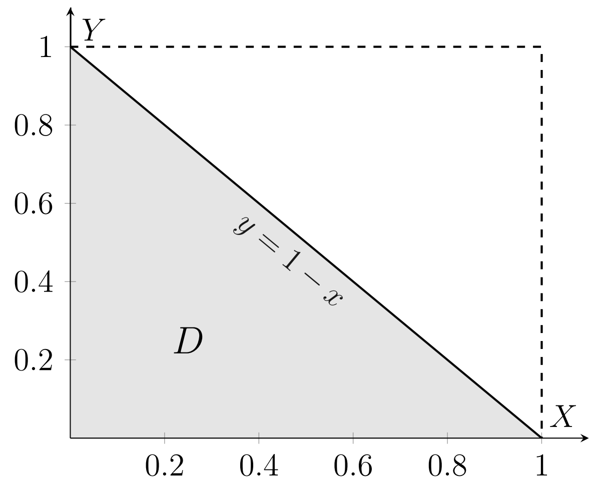
באופן דומה, כאשר $y\notin[0,1]$ הצפיפות השולית היא 0 ואילו כאשר $y\in[0,1]$ הצפיפות השולית תהא

$$f_Y\left(y\right)=\int_0^1 1+x^2-2xy^2\,dx=1+\left.\frac{x^3}{3}\right|_0^1-2y^2\left.\frac{x^3}{3}\right|_0^1=-y^2+\frac{4}{3}$$

לסיכום, הצפיפויות השוליות הינן:

$$f_X\left(x\right)=\begin{cases} x^2-\frac{2x}{3}+1 & x\in[0,1]\\ 0 & x\notin[0,1]\end{cases}\qquad f_Y\left(y\right)=\begin{cases} -y^2+\frac{4}{3} & y\in[0,1]\\ 0 & y\notin[0,1]\end{cases}$$

3. ההסתברות של המאורע $D=\{X+Y\leq 1\}$ מתקבלת על ידי חישוב האינטגרל $\iint_D f_{XY}\left(x,y\right)\,dxdy$. לצורך מציאת גבולות האינטגרציה מומלץ מאוד לשרטט על פני מישור XY את תחום ההגדרה של הצפיפות המשותפת ואת התחום D. החיתוך ביניהם הוא התחום עליו יש לבצע את האינטגרל:



כלומר גבולות האינטגרל החיצוני יהיו $x = 0$ ועד 1 ואילו גבולות האינטגרל הפנימי יהיו $y = 0$ ועד $1 - x$:

$$\begin{aligned} Pr\left(D\right) &= \int_{x=0}^{1}\int_{y=0}^{1-x} 1 + x^2 - 2xy^2 \, dydx = \int_0^1 \left(1 + x^2\right)(1-x) - 2x\left.\frac{y^3}{3}\right|_0^{1-x} dx \\ &= \int_0^1 1 + x^2 - x - x^3 - \frac{2}{3}x(1-x)^3 \, dx = \frac{11}{20} \end{aligned}$$

□

דוגמא זו ממחישה מדוע פונקציית התפלגות מצטברת דו-ממדית אינה שימושית. פונקציה כזו, בדומה לפונקציית ההתפלגות המצטברת החד-ממדית, מוגדרת על ידי

$$F_{XY}\left(x, y\right) = Pr\left(X \leq x, Y \leq y\right)$$

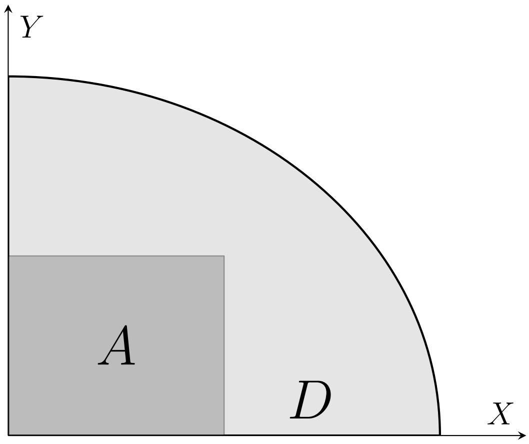
ומתייחסת למאורעות שהם רבע מישור. הסתברויות של מאורעות דו-ממדים כלליים (למשל, בצורת משולש) לא ניתנות לתיאור בפשטות על ידי חיסור רבעי מישורים. זאת בניגוד למקרה החד-ממדי שם כל קטע ניתן היה לתיאור בתור הפרש של שתי קרנות.

מקרה מיוחד ופשוט יחסית הוא כאשר ההתפלגות היא אחידה, כלומר כאשר הצפיפות המשותפת קבועה. מאחר ואינטגרל (דו-ממדי) על קבוע הוא השטח בו מבוצע האינטגרל, הרי שחישובי מאורעות הופכים לחישובי שטחים גאומטריים פשוטים יותר. בנוסף, מאחר והאינטגרל על הצפיפות על פני כל תחום ההגדרה צריך להיות **1**, הצפיפות הקבועה תהיה ההופכי של שטח הצורה עליה היא מוגדרת. למשל, נניח ש-X, Y מתפלגים אחיד ברבע המעגל ברדיוס **1** הנמצא ברביע הראשון, כלומר בתחום $D = \left\{(x, y) \in \mathbb{R}_+^2 \middle| x^2 + y^2 \leq 1\right\}$. שטח רבע המעגל הינו $S_D = \frac{\pi}{4}$ (תרגיל **0.7**) ולכן הצפיפות תהיה

$$f_{XY}\left(x, y\right) = \begin{cases} \frac{1}{S_D} = \frac{4}{\pi} & (x, y) \in D \\ 0 & (x, y) \notin D \end{cases}$$

נחשב את ההסתברות של המאורע $A = \left\{(x, y) \in \mathbb{R}_+^2 \middle| x \leq 0.5, y \leq 0.5\right\}$. מאורע זה מתאר ריבוע הנמצא כולו בתוך רבע המעגל ושטחו $S_A = 0.5^2 = 0.25$. לכן ההסתברות המבוקשת היא שטח הריבוע ביחס לשטח רבע המעגל, קרי

$$Pr\left(A\right)=\frac{S_A}{S_D}=\frac{0.25}{\pi/4}=\frac{1}{\pi}$$



מעניין לשים לב שגם כאשר המשתנה המקרי מתפלג באופן אחיד, זה לא מבטיח שכל אחד מהמשתנים מתפלג שולית בצורה אחידה. למשל קל לראות שבדוגמא זו יש יותר סיכוי לקבל ערכים קטנים עבור X שכן שטח התחום $\{x\in[0,0.1]|(x,y)\in D\}$ גדול בהרבה משטח התחום $\{x\in[0.9,1]|(x,y)\in D\}$. חישוב ישיר יראה שפונקציית הצפיפות של X בתחום $[0,1]$ הינה:

$$f_X\left(x\right)=\int_0^{\sqrt{1-x^2}}\frac{4}{\pi}\,dy=\frac{4\sqrt{1-x^2}}{\pi}$$

וזו כמובן אינה פונקציה קבועה כך ש-$X\not\sim U\left(0,1\right)$.

גם במקרה הרציף ניתן להתנות על אחד המשתנים המקריים ולנתח את ההתפלגות של המשתנה המקרי האחר בהינתן הערך הידוע של המשתנה המקרי השני. הנוסחא זהה למקרה הבדיד כאשר הצפיפות מחליפה את פונקציית ההסתברות.

הגדרה 7.7 (צפיפות מותנית). יהי X,Y משתנה מקרי דו-ממדי רציף עם פונקציית צפיפות משותפת $f_{XY}\left(x,y\right)$. לכל $y\in\mathbb{R}$ עבורו $f_Y\left(y\right)\neq0$ נגדיר את הצפיפות המותנית של X בהינתן $Y=y$ בתור

$$f_{X|Y=y}\left(x\right)=\frac{f_{XY}\left(x,y\right)}{f_Y\left(y\right)}\qquad(7.9)$$

כאשר f_Y היא הצפיפות השולית של Y. נסמן את המשתנה המקרי המתקבל על ידי $X|Y=y$. כאשר y הוא מספר או $X|Y$ באופן כללי.

גם כאן הצפיפות המותנית המתקבלת מתארת את ההתפלגות של משתנה מקרי חד ממדי רציף: $X|Y$. ההבדל בין הצפיפות המותנית לצפיפות השולית היא שהצפיפות השולית לא מניחה דבר על תוצאות הניסוי ואילו הצפיפות המותנית מניחה שערכו של Y ידוע ולכן ההתפלגות של X בהינתן מידע זה תהיה שונה.

למרות שמדובר על משתנה מקרי חד-ממדי, בנוסחא 7.9 מופיעים גם x וגם y, אך בתפקידים שונים. y הוא פרמטר, הערך ש-Y קיבל בניסוי ובהינתן ערך זה מחשבים את כל ההסתברויות. לעומתו, x הוא הארגומנט של

הפונקציה המתאים לערך של המשתנה המקרי X. תחום ההגדרה של הפונקציה ושאר המאפיינים שלה, אם כך, יכולים להיות תלויים גם ב-y.

כאשר הצפיפות המותנית שווה לצפיפות השולית נאמר שהמשתנים המקריים בלתי תלויים. דרך נוחה יותר להגדיר חוסר תלות היא על ידי פירוק הצפיפות המשותפת למכפלת הצפיפויות השוליות:

הגדרה 7.8. משתנים מקריים דו-ממדיים רציפים X,Y נקראים בלתי תלויים אם

$$f_{XY}(x,y)=f_X(x)\,f_Y(y) \tag{7.10}$$

לכל $x,y\in\mathbb{R}$.

בעיה (המשך בעיה 7.4). 1. מהי הצפיפות המותנית של X בהינתן $Y=\frac{1}{2}$?

2. מהי הצפיפות המותנית $X|Y=y$ לכל $y\in[0,1]$?

3. האם המשתנים המקריים הללו תלויים או בלתי תלויים? הוכיחו!

הוכחה. 1. לפי ההגדרה הצפיפות המותנית הינה

$$f_{X|Y=\frac{1}{2}}(x)=\frac{f_{XY}\left(x,\frac{1}{2}\right)}{f_Y\left(\frac{1}{2}\right)}=\frac{1+x^2-2x\left(\frac{1}{2}\right)^2}{-\left(\frac{1}{2}\right)^2+\frac{4}{3}}=\frac{12x^2-6x+12}{13}$$

בתחום $x\in[0,1]$ ו-0 אחרת. קל לראות שהפונקציה שהתקבלה היא אכן פונקציית צפיפות: חיובית ומנורמלת.

2. באופן כללי, הצפיפות המותנית עבור כל $y\in[0,1]$ הינה

$$f_{X|Y=y}(x)=\frac{f_{XY}(x,y)}{f_Y(y)}=\frac{1+x^2-2xy^2}{-y^2+\frac{4}{3}}$$

כאשר $x\in[0,1]$ ו-0 אחרת.

3. המשתנים המקריים תלויים. קל לראות זאת למשל על ידי השוואה בין השולית של X לבין $X|Y=\frac{1}{2}$. אם המשתנים המקרים היו בלתי תלויים שתי הצפיפויות היו זהות אך אין כך הדבר. כמו כן, ניתן לראות ישירות לפי ההגדרה. בתחום $x,y\in[0,1]$ מתקיים:

$$f_X(x)\cdot f_Y(y)=\left(x^2-\frac{2x}{3}+1\right)\cdot\left(-y^2+\frac{4}{3}\right)\neq 1+x^2-2xy^2=f_{XY}(x,y)$$

□

תרגיל 7.7. הצפיפות המשותפת של זוג משתנים מקריים רציפים ניתנת על ידי:

$$f_{X,Y}(x,y)=\begin{cases}cx(x-y) & (x,y)\in D\\ 0 & otherwise\end{cases}$$

כאשר $D=\{(x,y)|0\le x\le 2,-x\le y\le x\}$.

1. מצאו את c.
2. חשבו את הצפיפויות השוליות ואת התוחלות של כל אחד מהמשתנים המקריים.

תרגיל 7.8. המשתנים המקריים X,Y מתפלגים לפי $f_{XY}(x,y)=c(5-x-y)$ כאשר $x\in[1,2]$, $y\in[0,3]$ ו-0 אחרת.

1. חשבו את c.
2. מצאו את הצפיפות השוליות של המשתנים המקריים.
3. חשבו את ההסתברויות: $Pr\left(X-Y<0.5\right),Pr\left(XY<3\right)$.

7.2.1 התפלגות של פונקציה של משתנה מקרי דו-ממדי

סוג מיוחד של בעיות הוא חישוב התפלגות של משתנה מקרי חד-ממדי רציף המוגדר כפונקציה של משתנה מקרי דו-ממדי. סוג זה מזכיר בעיות מפרק 4, דוגמת בעיה 4.5, שם חישבנו כיצד מתפלג משתנה מקרי אחד הנתון כפונקציה של משתנה מקרי אחר. הגישה לפתרון זהה. גם כאן נחשב את ההתפלגות על ידי מציאת פונקציית ההתפלגות המצטברת, שהרי פונקציה זו מתארת מאורע אשר את הסתברותו ניתן לחשב. ההבדל הוא שהמאורע הוא מאורע הקשור למשתנים מקריים דו-ממדיים ולכן מתאר תחום מסויים במישור. חישוב הסתברותו יהיה על ידי חישוב אינטגרל כפול של פונקציית הצפיפות המשותפת בתחום זה, כפי שביצענו לאורך הפרק. נמחיש את השיטה הכללית באמצעות הבעיה הבאה:

בעיה 7.5. אלכס משרטט מלבנים על דף. אורך צלעו האחת של המלבן אקראית בין 0 ל-1 סנטימטרים ואורך צלעו השניה אקראית בין 0 ל-2 סנטימטרים, ללא תלות באורך הצלע הראשונה. כיצד מתפלג היקף המלבן שאלכס שרטט?

הוכחה. נפתח בהגדרת המשתנים המקריים המופיעים בבעיה. נסמן את אורך הצלע הראשונה ב- X ואת אורך הצלע השניה ב-Y. לפי הנתון מתקיים $X\sim U\left(0,1\right)$ ו-$Y\sim U\left(0,2\right)$. כל צלע משורטטת פעמיים ולכן היקף המלבן הינו $Z=2X+2Y$ ואת התפלגותו של משתנה מקרי זה אנחנו מחפשים. נחשב את פונקציית ההתפלגות המצטברת של Z. לכל $t\in\mathbb{R}$ פונקציית ההתפלגות המצטברת מוגדרת על ידי:

$$F_Z\left(t\right)=Pr\left(Z\le t\right)=Pr\left(2X+2Y\le t\right)\overset{eq.7.6}{=}\iint_{\{2X+2Y\le t\}}f_{XY}\left(x,y\right)\,dxdy$$

נותר לקבוע שני דברים: מהי פונקציית הצפיפות המשותפת וכיצד נראה התחום במישור בו יש לבצע את האינטגרל. נתונות שתי הצפיפויות השוליות וכן נאמר שהמשתנים המקריים בלתי תלויים, לכן הצפיפות המשותפת היא מכפלתן של הצפיפויות השוליות. הצפיפויות השוליות (נוסחא 6.27) הן

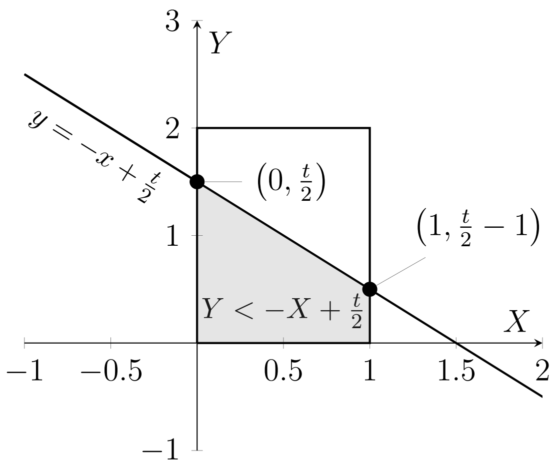
$$f_X\left(x\right)=\begin{cases}1 & x\in[0,1]\\ 0 & x\notin[0,1]\end{cases}\qquad f_Y\left(y\right)=\begin{cases}0.5 & y\in[0,2]\\ 0 & y\notin[0,2]\end{cases}$$

כך שהצפיפות המשותפת הינה

$$f_{XY}\left(x,y\right)=f_X\left(x\right)f_Y\left(y\right)=\begin{cases}0.5 & 0\le x\le 1,0\le y\le 2\\ 0 & \text{otherwise}\end{cases}$$

כלומר הצפיפות היא קבועה בכל המלבן $[0,1]\times[0,2]$ ולכן ניתן להחליף את חישובי האינטגרלים בחישובי שטחים וחלוקתם בשטח המלבן – 2. נרשום בכל זאת את סימני האינטגרל בכל המקומות על מנת שהדוגמא תישאר כללית.

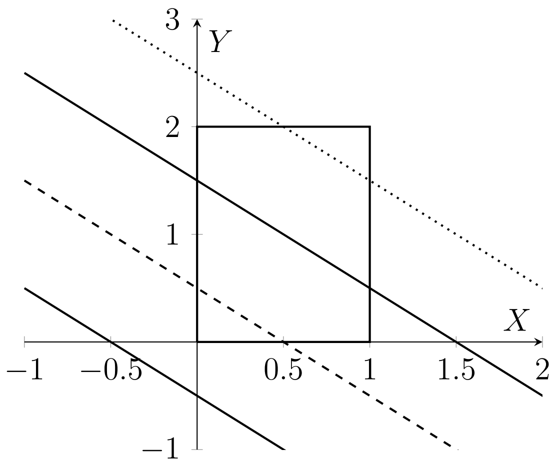
תחום האינטגרציה הינו $2X+2Y\leq t$ השקול ל- $Y\leq\frac{t}{2}-X$. כלומר תחום האינטגרציה הוא חצי המישור הנמצא מתחת לישר $y=-x+\frac{t}{2}$. נשרטט את תחום ההגדרה של הצפיפות המשותפת וכן ישר מסוג זה עבור ערך t כללי בכדי לקבוע את גבולות האינטגרציה:



כלומר תחום האינטגרציה הינו טרפז. נחשב את ההסתברות המבוקשת (ניתן היה לחשב הסתברות זו על ידי חישוב שטח הטרפז והכפלתו ב-0.5 מאחר והצפיפות אחידה):

$$Pr\left(X+Y\leq\frac{t}{2}\right)=\int_{x=0}^{1}\int_{y=0}^{\frac{t}{2}-x}\frac{1}{2}\,dydx=\frac{t}{4}-\frac{1}{4}$$

עבור ערכי t אחרים שיפוע הישר לא ישתנה (כי אינו תלוי ב-t) ונקבל ישרים המקבילים לישר ששרטטנו, מעליו או מתחתיו. עבור ערכי t אחרים נקבל תחומי אינטגרציה שונים מהטרפז. כך למשל, כאשר t מאוד קטן, הישר עובר מתחת למלבן, החיתוך ביניהם ריק וההסתברות הינה 0. תוצאה זו אינטואיטיבית מאוד - צלעות המלבן חיוביות והסיכוי שההיקף יהיה שלילי הוא 0. באופן דומה, כאשר t גדול מאוד, הישר עובר מעל המלבן, החיתוך ביניהם יהיה המלבן כולו וההסתברות להיות בו היא 1. גם תוצאה זו אינטואיטיבית - אורך הצלע הקטנה הוא לכל היותר 1 ואורך הצלע הגדולה הוא לכל היותר 2, לכן ההיקף הוא לכל היותר $6=(2+1)\,2$, כך שהסיכוי שההיקף קטן מ-20 הוא **1**. עבור ערכי t אחרים החיתוך בין הישר למלבן יכול להיות, בנוסף לטרפז, גם משולש (מקווקו) או מחומש (נקודות), כמוראה בתרשים:



המעבר בין חיתוך ריק למשולש מתבצע כאשר הישר חותך את המלבן בדיוק בנקודה אחת, $(0,0)$, כלומר כאשר $t=0$. כאשר החיתוך הינו משולש ההסתברות תהא

$$Pr\left(X+Y\le\frac{t}{2}\right)=\int_{x=0}^{\frac{t}{2}}\int_{y=0}^{\frac{t}{2}-x}\frac{1}{2}\,dydx=\frac{t^2}{16}$$

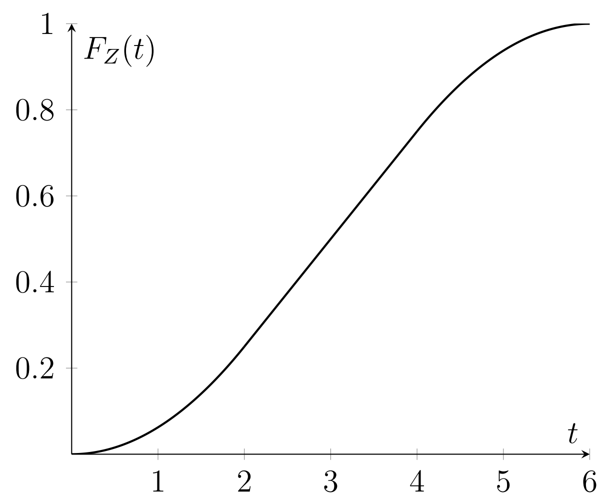
המעבר בין משולש לטרפז מתבצע כאשר הישר עובר דרך $(1,0)$ כלומר כאשר $t=2$. המעבר מטרפז למחומש מתבצע כאשר הישר עובר דרך $(0,2)$ כלומר עבור $t=4$. גבולות האינטגרל במקרה זה מעט יותר מסובכים ויותר נוח לחשב את ההסתברות על ידי המשלים:

$$Pr\left(X+Y\le\frac{t}{2}\right)=1-\int_{x=\frac{t}{2}-2}^{1}\int_{y=\frac{t}{2}-x}^{1}\frac{1}{2}\,dydx=1-\frac{(6-t)^2}{16}$$

החיתוך הופך להיות המלבן כולו כאשר הישר עובר דרך הנקודה $(1,2)$ כלומר כאשר $t=6$. לסיכום, פונקציית ההתפלגות המצטברת הינה

$$F_Z\left(t\right)=\begin{cases}0 & t<0\\ \frac{t^2}{16} & t\in[0,2]\\ \frac{t}{4}-\frac{1}{4} & t\in[2,4]\\ 1-\frac{(6-t)^2}{16} & t\in[4,6]\\ 1 & t>6\end{cases}$$

דרך נוחה לוודא שלא פוספס אף תחום של ערכי t ושהפונקציה שהתקבלה היא אכן פונקציית התפלגות מצטברת היה על ידי וידוא תכונותיה לפי משפט 4.2 בשילוב עם רציפות: הגבול ב-$-\infty$ הוא 0, הפונקציה עולה, רציפה (הבדיקה המשמעותית היא במעבר בין התחומים) והגבול ב-$+\infty$ הוא 1:



פונקציית ההתפלגות המצטברת מספיקה בכדי לתאר את ההתפלגות. כרגיל, במידה ומעוניינים דווקא בפונקציית הצפיפות ניתן לקבלה על ידי גזירה:

$$f_Z(t) = \frac{dF_Z(t)}{dt} = \begin{cases} 0 & t \notin [0,6] \\ \frac{t}{8} & t \in [0,2] \\ \frac{1}{4} & t \in [2,4] \\ \frac{6-t}{8} & t \in [4,6] \end{cases}$$

הסדר המומלץ לפתרון תרגילים מסוג זה הוא הסדר ההפוך מזה שבו נקטנו: קודם שרטוט ומציאת תחומי צורות האינטגרציה השונות כתלות ב-t ורק לאחר מכן חישוב כל אחד מהאינטגרלים. $\square$

תרגיל 7.9. בהמשך לבעיה 7.5, כיצד מתפלג שטח המלבן שאלכס משרטט?

תרגיל 7.10. שתי נקודות נבחרות באקראי ובאופן בלתי תלוי בקטע $(0,1)$. חשבו את התפלגות המרחק ביניהן, $Z = |X - Y|$ ואת $\mathbb{E}(Z)$.

תרגיל 7.11. יורים למטרה מישורית אינסופית. נסמן את קוארדינטות נקודות הפגיעה ב-(X,Y) ונניח שכל אחת מהן מתפלגת $N(0,1)$ באופן בלתי תלוי באחרת. נגדיר R – המרחק של נקודת הפגיעה מ-$(0,0)$. חשבו את הצפיפות ופונקציית ההתפלגות המצטברת של R.

מקרה פרטי ומעניין במיוחד הוא סכום של משתנים מקריים בלתי תלויים, למשל כפי שראינו בדיון על משפט הגבול המרכזי (משפט 6.16). ניתן לחשב סכומים אלו באמצעות הטכניקה שהוצגה לעיל אך עבור סכומים ישנה עוד אפשרות. כפי שתוכיחו בתרגיל 7.38, הפונקציה היוצרת של סכום משתנים מקריים בלתי תלויים שווה למכפלת הפונקציות יוצרות המומנטים שלהם. לכן, אם ידועות הפונקציות היוצרות מומנטים של המשתנים המקריים ניתן לחשב את הפונקציה היוצרת של הסכום וממנה להסיק את התפלגות המשתנה המקרי. במקרים רבים המעבר ממשתנים מקריים לפונקציה יוצרת מומנטים ובחזרה אינו פשוט מספיק בכדי להצדיק דרך זו, אך עבור ההתפלגות הנורמלית זוהי דרך מאוד נוחה לקבל תוצאה חשובה – סכום של משתנים מקריים נורמליים מתפלג אף הוא נורמלית.

משפט 7.2. *יהיו $X \sim N(\mu_X, \sigma_X^2)$ ו-$Y \sim N(\mu_Y, \sigma_Y^2)$ משתנים מקריים בלתי תלויים. אזי סכומם מתפלג אף הוא נורמלית:*

$$X + Y \sim N\left(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2\right)$$

הוכחה. נשתמש בפונקציות יוצרות מומנטים. לפי תרגיל 6.24 הפונקציות היוצרות מומנטים של המשתנים המקריים הינן

$$M_X(t) = \exp\left(\frac{\sigma_X^2 t^2 + 2\mu_X t}{2}\right) \qquad M_Y(t) = \exp\left(\frac{\sigma_Y^2 t^2 + 2\mu_Y t}{2}\right)$$

ולפי תרגיל 7.38 הפונקציה היוצרת מומנטים של הסכום הינה

$$\begin{aligned} M_{X+Y}(t) = M_X(t) M_Y(t) &= \exp\left(\frac{\sigma_X^2 t^2 + 2\mu_X t}{2}\right) \cdot \exp\left(\frac{\sigma_Y^2 t^2 + 2\mu_Y t}{2}\right) \\ &= \exp\left(\frac{(\sigma_X^2 + \sigma_Y^2) t^2 + 2(\mu_X + \mu_Y) t}{2}\right) \end{aligned}$$

וזו בדיוק הפונקצה היוצרת מומנטים של משתנה מקרי נורמלי עם תוחלת $\mu_X + \mu_Y$ ושונות $\sigma_X^2 + \sigma_Y^2$. לכן $X + Y$ מתפלג נורמלית עם פרמטרים אלו. $\square$

על פניו, תוצאה זו מאוד אינטואיטיבית. כאשר מחברים שני משתנים מקריים בלתי תלויים, תוחלת הסכום היא סכום התוחלות ושונות הסכום היא סכום השונויות, בדיוק כפי שקיבלנו זה עתה. היות ושני גדלים אלו הם בדיוק הגדלים הקובעים את מאפייניה של ההתפלגות הנורמלית, ניתן היה לנחש שאם הסכום הוא נורמלי הוא בהכרח נורמלי עם מאפיינים אלו. העובדה היחידה אותה לא ניתן היה לנחש מראש היא שאכן הסכום מתפלג נורמלית, שהרי לא מובטח שסכום של שני משתנים מקריים מאותה משפחה יהיה גם כן באותה המשפחה (בבעיה 7.5 ראינו שסכום אחידים אינו אחיד).

תרגיל 7.12. גובהם של גברים באוכלוסיה מתפלג נורמלית עם תוחלת 175 ס"מ וסטיית תקן 10 ס"מ ואילו גובהן של נשים מתפלג נורמלית עם תוחלת 165 ס"מ וסטיית תקן 9 ס"מ. נניח שגובהו של כל ילד הוא ממוצע משוקלל של הגבהים של הוריו, כאשר המשקל של גובה האב הינו $\frac{3}{4}$ אם הוא בן והמשקל של גובה האב הינו $\frac{1}{4}$ אם היא בת (משקל גובה האם הוא $\frac{1}{4}$ ו-$\frac{3}{4}$ בהתאמה). גבר ואישה נפגשים באקראי ומולידים ילד.

1. כיצד מתפלג גובהו של הילד בהינתן שהוא בן?
2. כיצד מתפלג גובהו של הילד מבלי שידוע מינו?

במקרה הבדיד הרעיון הכללי זהה אבל התהליך הטכני מורכב יותר. כזכור, במקרה הבדיד פונקציית ההתפלגות המצטברת לא נוחה לעבודה ובמקומה ניתן לחשב את ההסתברויות המבוקשות ישירות באמצעות פונקציית ההסתברות. הבעיה היא שלרוב הביטויים שיתקבלו הם סכומים מורכבים ויש הרבה פחות זהויות ונוסחאות סגורות עבור סכומים מאשר אינטגרלים. למשל, נתבונן במשתנים המקריים הבלתי תלויים $X\sim Bin\left(10,\frac{1}{2}\right)$ ו-$Y\sim Bin\left(10,\frac{1}{4}\right)$ ונגדיר $Z=X+Y$. המאורע $\{Z=9\}$ מתקיים כאשר אחד המשתנים הוא k והשני הוא $9-k$ ולכן הסתברותו:

$$\begin{aligned}
Pr\left(Z=9\right) &= \sum_{k=0}^{9} Pr\left(X=k,Y=9-k\right)=\sum_{k=0}^{9} Pr\left(X=k\right)Pr\left(Y=9-k\right)\\
&=\sum_{k=0}^{9}\binom{10}{k}\binom{10}{9-k}\left(\frac{1}{2}\right)^{k}\left(\frac{1}{2}\right)^{10-k}\left(\frac{1}{4}\right)^{9-k}\left(\frac{3}{4}\right)^{10-(9-k)}\\
&=\frac{1}{2^{10}4^{10}}\sum_{k=0}^{9}\binom{10}{k}\binom{10}{9-k}3^{k+1}
\end{aligned}$$

כלל לא ברור כיצד להמשיך הלאה (אם בכלל ניתן) ולקבל תוצאה סופית עבור $Z=9$ ובטח שלא עבור המקרה הכללי. יחד עם זאת, עבור מספר מקרים פשוטים ניתן לקבל תוצאה כללית בזכות התוצאה של תרגיל 7.38 או באמצעות הבנה מה הסכום מייצג (שהרי משתנים מקריים בדידים מתארים לרוב "סיפור"):

משפט 7.3. *יהיו X,Y משתנים מקריים בלתי תלויים ונגדיר $Z=X+Y$. אזי:*

1. *אם $X\sim Bin\left(n,p\right)$ ו-$Y\sim Bin\left(m,p\right)$ אז $Z\sim Bin\left(m+n,p\right)$.*
2. *אם $X\sim NB\left(n,p\right)$ ו-$Y\sim NB\left(m,p\right)$ אז $Z\sim NB\left(m+n,p\right)$. בפרט, אם X,Y גאומטריים ($n=m=1$) עם פרמטר p אזי סכומם $NB\left(2,p\right)$.*
3. *אם $X\sim Pois\left(\lambda_X\right)$ ו-$Y\sim Pois\left(\lambda_Y\right)$ אז $Z\sim Pois\left(\lambda_X+\lambda_Y\right)$.*

הוכחה. 1. כל אחד מהמשתנים המקריים סופר הצלחות כאשר חוזרים על ניסוי מספר פעמים והסיכוי להצלחה בכל סיבוב הוא p. לכן $X+Y$ סופר את מספר ההצלחות כאשר מבצעים $m+n$ חזרות ולכן מתפלג $Bin\left(m+n,p\right)$.

2. כל אחד מהמשתנים המקריים סופר את מספר החזרות עד ההצלחה ה-m או ה-n. לכן סכומם סופר את מספר ההצלחות עד ההצלחה ה-$n+m$ ולכן מתפלג בינומית שלילית. בפרט, כאשר $m=n=1$ אלו משתנים מקריים גאומטריים ולכן סכום של משתנים מקריים גאומטריים הוא משתנה מקרי בינומי שלילי $NB\left(2,p\right)$.
חשוב להדגיש שבעבור שני סעיפים אלו הסיכוי להצלחה חייב להיות זהה בשני המשתנים המקריים. אחרת לא ניתן לחשוב על משתנה הסכום כעל ספירה של הצלחות או חזרות של אותו הניסוי ולכן הוא יתפלג אחרת.
3. נשתמש בפונקציה יוצרת מומנטים ותרגיל 7.38. מתקיים:

$$\begin{aligned}M_{X+Y}\left(t\right)&=M_X\left(t\right)M_Y\left(t\right)\overset{exer.\,6.15}{=}e^{-\lambda_X\left(1-e^t\right)}e^{-\lambda_Y\left(1-e^t\right)}\\&=e^{-\left(\lambda_X+\lambda_Y\right)\left(1-e^t\right)}\end{aligned}$$

הפונקציה היוצרת מומנטים שהתקבלה מתאימה למשתנה מקרי פואסוני עם פרמטר $\lambda_X+\lambda_Y$, כדרוש.
□

7.3 תוחלת שלמה ושונות שלמה

נסתכל על משתנה מקרי דו-ממדי X,Y. ניתן להתייחס לשתי התפלגויות חד-ממדיות המאפיינות את המשתנה המקרי X. התפלגות אחת היא ההתפלגות השולית, המתארת את ההתפלגות משתנה מקרי זה כאשר מתעלמים מקיומו של Y ומתעניינים רק בערכים ש-X יכול לקבל מבלי לדעת דבר על תוצאת הניסוי. בהתפלגויות מסוג זה דנו כבר בפרקים הקודמים. ההתפלגות השניה היא ההתפלגות המותנית של X בהינתן Y. התפלגות זו מניחה שידוע הערך של Y שהתקבל בניסוי ובהינתן מידע זה מעוניינים לדעת את ההתפלגות של המשתנה המקרי X. זוהי התפלגות חד-ממדית רגילה (עם פרמטר - הערך של Y) ולכן ניתן לחשב את מאפייניה השונים דוגמת התוחלת והשונות.

הגדרה 7.9 (תוחלת מותנית). יהי X,Y משתנה מקרי דו-ממדי. התוחלת המותנית של X בהינתן $Y=y$ מוגדרת על ידי

$$\mathbb{E}\left(X|Y=y\right)=\int_{-\infty}^{\infty}xf_{X|Y=y}\left(x\right)\,dx \tag{7.11}$$

כאשר המשתנים רציפים ו-

$$\mathbb{E}\left(X|Y=y\right)=\sum_{k\in supp(X)}k\cdot Pr\left(X=k|Y=y\right) \tag{7.12}$$

כאשר הם בדידים. בשני המקרים נרשום בקיצור $\mathbb{E}\left(X|Y\right)$ במקרה הכללי (ללא ערך ספציפי של Y).

התוחלת המותנית המוגדרת באופן זה מקיימת את כל התכונות הרגילות של תוחלת מפרק 5 ובייחוד היא חיבורית (נוסחא 5.6) ולינארית (נוסחא 5.3). התוספת היא שבחישוב התוחלת המותנית מתייחסים אל הערך של Y כאל גודל ידוע ולכן גם כל הגדלים הקשורים ב-Y ידועים, קבועים ויוצאים מהתוחלת לפי הלינאריות.

משפט 7.4 (לינאריות התוחלת המותנית). *יהי X,Y משתנה מקרי דו-ממדי ותהיינה $h,g_1,g_2:\mathbb{R}\to\mathbb{R}$ פונקציות כלשהן. אזי*

$$\mathbb{E}\left(g_1\left(Y\right)h(X)+g_2\left(Y\right)|Y\right)=g_1\left(Y\right)\mathbb{E}\left(h\left(X\right)|Y\right)+g_2\left(Y\right) \tag{7.13}$$

ההבדל העיקרי בין תוחלת "רגילה" לבין תוחלת מותנית היא שתוחלת מותנית היא פונקציה של Y ובהיותה פונקציה של משתנה מקרי, תוחלת מותנית היא בעצמה משתנה מקרי! הרעיון הוא שהערך של Y נקבע באקראי ובהתאם לערך זה – נקבע ערך התוחלת המותנית. לאור זאת, ניתן לחשב גם את ההתפלגות של התוחלת המותנית ואפילו את התוחלת של התוחלת המותנית. במקרה כזה, נקבל את התוחלת "הרגילה" של X.

משפט 7.5 (משפט התוחלת השלמה). *יהי X,Y משתנה מקרי דו-ממדי. אזי*

$$\mathbb{E}\left(X\right)=\mathbb{E}\left(\mathbb{E}\left(X|Y\right)\right) \tag{7.14}$$

הוכחה. נוכיח עבור המקרה הרציף כאשר ההוכחה למקרה הבדיד דומה. התוחלת המותנית של X בהינתן Y היא פונקציה של משתנה מקרי Y ולכן נחשב את תוחלתה בהתאם לתוחלת של פונקציה של משתנה מקרי Y (נוסחא 5.5):

$$\begin{aligned}\mathbb{E}\left(\mathbb{E}\left(X|Y\right)\right)&=\int_{-\infty}^{\infty}\mathbb{E}\left(X|Y=y\right)f_Y\left(y\right)\,dy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{X|Y=y}\left(x\right)f_Y\left(y\right)\,dxdy\\(eq.\,7.9)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{XY}\left(x,y\right)\,dxdy=\int_{-\infty}^{\infty}x\left(\int_{-\infty}^{\infty}f_{XY}\left(x,y\right)\,dy\right)dx\\(eq.\,7.8)&=\int_{-\infty}^{\infty}xf_X\left(x\right)\,dx=\mathbb{E}\left(X\right)\end{aligned}$$

□

האינטואיציה למשפט התוחלת השלמה היא שהתוחלת של X מהווה ממוצע ולכן ניתן לחשבה גם על ידי חלוקת הערכים האפשריים לקבוצות, חישוב ממוצע לכל קבוצה ולאחר מכן ממוצע של ממוצעים. נניח למשל שמחשבים את ממוצע הציונים בכיתה מסויימת. דרך אחת היא לחשב את הממוצע כרגיל: סכום כל הציונים חלקי מספר התלמידים. חישוב כזה יהיה התוחלת "הרגילה" של X. אפשרות נוספת היא לחלק את הכיתה לשתי קבוצות, נאמר הבנים והבנות, ולחשב את הממוצע בכל קבוצה. הממוצע בקבוצת הבנים הוא למעשה התוחלת של הציון בהינתן שהנבחן הוא בן והממוצע בקבוצת הבנות הוא התוחלת של הציון בהינתן שהנבחנת היא בת. בכדי לחשב את הציון הממוצע בכיתה כולה יש לחשב את הממוצע המשוקלל של שני הממוצעים (המותנים) הללו, כאשר המשקולות הן היחס בין מספר הבנים למספר הבנות בכיתה.

משפט התוחלת השלמה מהווה דרך נוספת לחישוב התוחלת של X כאשר ההתפלגות שלו לא ידועה, בדומה לפירוק משתנה מקרי לסכום של משתנים מקריים אחרים שתוחלתם ידועה. לצורך שימוש במשפט זה נדרשת ידיעה על התפלגות המותנית של X בהינתן משתנה מקרי אחר שהתפלגותו ידועה.

בעיה 7.6. זמן ההגעה של אוטובוס לתחנה הוא משתנה מקרי אחיד בדיד $T\sim U\left\{1,\ldots,5\right\}$ (בדקות). בתחנה עומדים 50 אנשים. בכל דקה עד הגעת האוטובוס מטיל כל אחד מהאנשים מטבע ואם יוצא עץ – מתעייף מההמתנה והולך הבייתה. מתי תוחלת מספר הנוסעים שנמצאים בתחנה כאשר האוטובוס מגיע?

הוכחה. נסמן את מספר הנוסעים הנמצאים בתחנה בזמן הגעת האוטובוס ב-X. חישוב ההתפלגות השולית של X מסובכת אבל לצורך חישוב התוחלת די אם נמצא את הקשר בין X ל-T. לצורך כך נניח ש-T ידוע ונחשב את ההתפלגות של $X|T$. כל נוסע מטיל מטבע בכל דקה ונשאר בתחנה רק אם יצא פלי. לכן, אם האוטובוס הגיע בזמן T הרי שהוא הטיל את המטבע T פעמים והסיכוי שהוא נשאר בתחנה הינו $\left(\frac{1}{2}\right)^T$. ישנם **50** נוסעים בתחילת התהליך ולכן מספר הנוסעים שנשאר בתחנה כאשר האוטובוס הגיע הינו $X|T \sim Bin\left(50, \frac{1}{2^T}\right)$. לכן התוחלת המותנית מתקבלת לפי הנוסחא של תוחלת של משתנה מקרי בינומי: $\mathbb{E}\left(X|T\right) = \frac{50}{2^T}$. נוסחא זו ממחישה את העובדה שתוחלת מותנית היא משתנה מקרי שכן התוחלת המותנית יכולה לקבל כל אחד מהערכים $\left\{\frac{50}{2^1}, \frac{50}{2^2}, \frac{50}{2^3}, \frac{50}{2^4}, \frac{50}{2^5}\right\}$ בהסתברות שווה. התוחלת של X, לפי משפט התוחלת השלמה, הינה

$$\begin{aligned}\mathbb{E}\left(X\right) &= \mathbb{E}\left(\mathbb{E}\left(X|T\right)\right) = \mathbb{E}\left(\frac{50}{2^T}\right) = 50\mathbb{E}\left(\frac{1}{2^T}\right)\\ &= 50\left(\frac{1}{2^1}\cdot\frac{1}{5} + \frac{1}{2^2}\cdot\frac{1}{5} + \frac{1}{2^3}\cdot\frac{1}{5} + \frac{1}{2^4}\cdot\frac{1}{5} + \frac{1}{2^5}\cdot\frac{1}{5}\right) = 9.7\end{aligned}$$

□

תרגיל 7.13. כל אחד מ-n חבריך הוא אוהד של אחת מבין m הנבחרות המשתתפות במונדיאל, כאשר לכל נבחרת הסתברות זהה להיות אהודה על ידי כל אחד. נסמן ב-R_i את מספר האוהדים של הנבחרת ה-i. אתם מתכנסים לצפות בטורניר ובתחילת הערב כל אחד מניח בתיבה כובע של נבחרתו האהודה. בסוף הערב כולם שולפים באקראי כובעים מהתיבה. נסמן ב-X_i את מספר אוהדי הנבחרת ה-i שישלפו בסוף הערב באקראי את הכובע של הנבחרת אותה הם אוהדים (לאו דווקא הכובע המקורי שלהם) וב-X את סך האנשים ששלפו כובע של נבחרתם.

1. כיצד מתפלג $X_i|R_i$? מהי $\mathbb{E}\left(X_i|R_i\right)$?
2. כיצד מתפלג R_i? מהי $\mathbb{E}\left(R_i^2\right)$?
3. מהי $\mathbb{E}\left(X_i\right)$? מהי $\mathbb{E}\left(X\right)$?

תרגיל 7.14. חשבו את תוחלת מספר הקנסות שהנהג מבעיה 7.3 מקבל באמצעות משפט התוחלת השלמה.

תרגיל 7.15. המשתנים המקריים X, Y מתפלגים אחיד בטרפז שקודקודיו $(0,1), (0,2), (1,0), (1,3)$. חשבו את התוחלת של Y באמצעות משפט התוחלת השלמה.

באופן דומה ניתן לחשב גם את השונות של משתנה מקרי אחד על ידי התנייה על המשתנה המקרי השני. במקרה כזה השונות של Y היא סכום של התוחלת של השונות המותנית עם השונות של התוחלת המותנית ונקראת משפט השונות השלמה או משפט פירוק השונויות.

***משפט* 7.6** (משפט השונות השלמה)**.** *יהי X, Y משתנה מקרי דו-ממדי. אזי*

$$\mathbb{V}\left(X\right) = \mathbb{V}\left(\mathbb{E}\left(X|Y\right)\right) + \mathbb{E}\left(\mathbb{V}\left(X|Y\right)\right) \qquad (7.15)$$

כאשר השונות המותנית הינה

$$\mathbb{V}\left(X|Y\right) = \mathbb{E}\left(X^2|Y\right) - \mathbb{E}^2\left(X|Y\right) \qquad (7.16)$$

הוכחה. נשתמש בהגדרה של שונות $\mathbb{V}(X)=\mathbb{E}(X^2)-\mathbb{E}^2(X)$ ונחשב כל אחת מהתוחלות באמצעות משפט התוחלת השלמה:

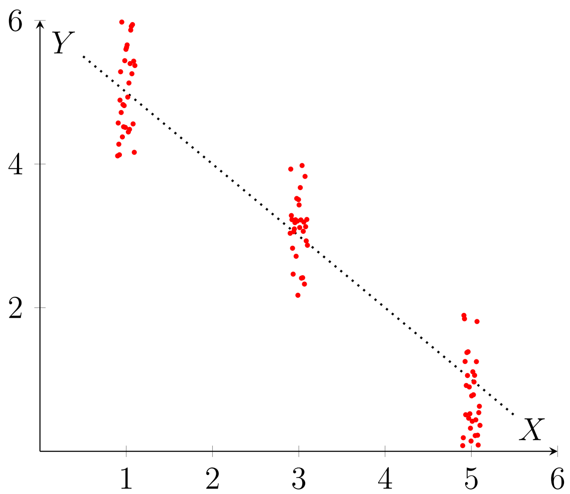
$$\mathbb{E}\left(X^2\right)=\mathbb{E}\left(\mathbb{E}\left(X^2|Y\right)\right)=\mathbb{E}\left(\mathbb{V}\left(X|Y\right)+\mathbb{E}^2\left(X|Y\right)\right)=\mathbb{E}\left(\mathbb{V}\left(X|Y\right)\right)+\mathbb{E}\left(\mathbb{E}^2\left(X|Y\right)\right)$$

לכן

$$\begin{aligned}\mathbb{V}(X)&=\mathbb{E}\left(\mathbb{V}\left(X|Y\right)\right)+\mathbb{E}\left(\mathbb{E}^2\left(X|Y\right)\right)-\mathbb{E}^2(X)\\&=\mathbb{E}\left(\mathbb{V}\left(X|Y\right)\right)+\mathbb{E}\left(\mathbb{E}^2\left(X|Y\right)\right)-\mathbb{E}^2\left(\mathbb{E}\left(X|Y\right)\right)\\&=\mathbb{E}\left(\mathbb{V}\left(X|Y\right)\right)+\mathbb{V}\left(\mathbb{E}\left(X|Y\right)\right)\end{aligned}$$

□

משפט השונות השלמה הינו הבסיס למשפחה של מבחנים סטטיסטיים הנקראים ניתוח שונויות (ANOVA-ANalysis Of VAriation). במבחנים אלו בודקים האם יש קשר בין שני משתנים מקריים על ידי ניתוח הפיזור של המשתנה Y ובדיקת איזה חלק ממנו ניתן להסביר על ידי המשתנה המקרי X. נניח שלמחלה מסויימת ניתן לתת תרופה בשלושה מינונים שונים $(X=1,3,5)$ ומעוניינים לבדוק עד כמה זמן ההחלמה בימים (מסומן ב-Y) מושפע מהמינון. באופן טבעי, זמן ההחלמה מושפע לא רק מהמינון ולכן גם עבור מינון זהה הוא יכול להיות שונה לחולים שונים. תוצאות הניסוי משורטטות בגרף הבא:



נסתכל על ערך בודד של X, למשל $X=3$. הערכים בגרף המתאימים לערך זה הם בדיוק ההתפלגות המותנית $Y|X=3$. התפלגות זו מורכבת משני חלקים: הממוצע (כלומר, התוחלת המותנית) והפיזור ביחס לממוצע (כלומר, השונות המותנית). בהנחה שיש קשר בין Y ל-X ושינוי המינון משפיע על מספר ימי ההחלמה, נצפה לראות שוני בתוחלת המותנית. שוני זה מתואר בשרטוט על ידי קו ישר מקווקו המחבר את התוחלות המותנות.[1] פיזור ממוצעים אלו הוא הפיזור אותו אנחנו יודעים להסביר על ידי ההשערה שיש קשר בין X ל-Y וזוהי בדיוק השונות המוסברת, $\mathbb{V}\left(\mathbb{E}\left(Y|X\right)\right)$. ככל שהקשר בין המשתנים המקריים הללו משמעותי יותר, הרי שההבדל בין התוחלות המותנות יהיה גדול יותר ולכן גם שונות זו. מצד שני, ככל שהקשר בין המשתנים הללו פחות משמעותי,

[1] באופן כללי הקשר לא חייב להיות לינארי

התוחלות המותנות יהיו קרובות זו לזו ולכן השונות תהיה קטנה. במקרה הקיצון, כאשר המשתנים המקרים בלתי תלויים, היא תהיה 0.

המרכיב השני של השונות של Y הוא השונות שלא מוסברת על ידי X, $\mathbb{E}\left(\mathbb{V}\left(Y|X\right)\right)$. שונות זו מתקבלת כאשר מסתכלים על כל ערך של X ובודקים את הפיזור של ערכי Y ביחס לממוצע (אך לא את השינוי בממוצע עצמו!). עבור כל ערך של X, השונות המותנית מתארת את הפיזור של Y הנובע מכל הגורמים האחרים האפשריים פרט ל-X. עבור ערכי X שונים אנו מצפים לאותו ערך של שונות וממוצע השונויות הוא השונות הלא מוסברת על ידי הבדלים ב-X. כאשר מנתחים את השונות הכוללת של Y, הרי שהיא מורכבת משני חלקים: השונות המוסברת הנובעת מהקשר בין המשתנים המקריים והשונות הלא מוסברת הנובעת מהאקראיות של Y שלא תלות ב-X. ככל שהשונות המוסברת מהווה מרכיב גדול יותר בשונות הכוללת, ניתן לומר שיש קשר יותר גדול בין המשתנים המקריים שכן חלק גדול יותר של השוני בערכי ה-Y נובע מההבדל בערכי ה-X.

זהו השימוש במשפט פירוק השונויות על קצה המזלג. הרחבות נוספות ניתן למצוא בקורסי המשך בסטטיסטיקה. כאן נשתמש במשפט זה בעיקר בכדי לחשב את השונות של Y באמצעות ההתפלגות המותנית, בדומה למשפט התוחלת השלמה.

7.4 שונות משותפת ומתאם

נסיים את הדיון במשתנים מקריים דו-ממדיים בהרחבת השונות למקרה הדו-ממדי ובכך גם נוכל להחזיר חוב ישן מפרק 5 ולהציג נוסחא לחישוב שונות של סכום משתנים מקריים. כשם שהשונות הוגדרה בתור הפיזור הממוצע של משתנה מקרי מסביב לתוחלת שלו, כך נגדיר גם את השונות המשותפת בתור הפיזור הממוצע של מכפלת המשתנים המקריים ביחס לממוצעיהם.

הגדרה 7.10 (שונות משותפת). יהיו X,Y משתנים מקריים דו-ממדיים. השונות המשותפת שלהם (Covariance) תוגדר על ידי

$$Cov\left(X,Y\right)=\mathbb{E}\left[\left(X-\mathbb{E}\left(X\right)\right)\left(Y-\mathbb{E}\left(Y\right)\right)\right] \qquad (7.17)$$

על ידי פתיחת סוגריים וחישוב התוחלות ניתן להגיע לנוסחא יותר נוחה:

$$Cov\left(X,Y\right)=\mathbb{E}\left(XY\right)-\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right) \qquad (7.18)$$

השונות המשותפת מרחיבה את מושג השונות הרגילה. אם נציב $X=Y$ נקבל

$$Cov\left(X,X\right)=\mathbb{E}\left(X^2\right)-\mathbb{E}^2\left(X\right)=\mathbb{V}\left(X\right) \qquad (7.19)$$

בניגוד לשונות הרגילה, השונות המשותפת יכולה להיות גם שלילית. ניתן לראות זאת לפי נוסחא 7.17 – השונות המשותפת היא ממוצע של ביטויים חיוביים כאשר $X>\mathbb{E}\left(X\right)$,$Y>\mathbb{E}\left(Y\right)$ או כאשר $X<$,$Y<\mathbb{E}\left(Y\right)$ $\mathbb{E}\left(X\right)$ וביטויים שליליים במקרים האחרים. לכן, ממוצע זה יכול להיות שלילי במקרה שהביטויים השליליים גוברים על הביטויים החיוביים. אינטואיציה זו מלמדת אותנו על המשמעות של השונות המשותפת. כאשר המשתנים המקריים גדלים ביחד, כלומר כאשר X מקבל ערך גבוה (נמוך) לרוב גם Y מקבל ערך גבוה (נמוך), השונות המשותפת תהיה חיובית שכן במצב זה מרבית הגורמים בממוצע יהיו חיוביים. לעומת זאת, כאשר המשתנים

המקריים גדלים בכיוונים מנוגדים, כלומר ערכים גבוהים (נמוכים) של X מלווים בערכים נמוכים (גבוהים) של Y, בממוצע יופיעו בעיקר ביטויים שליליים והשונות המשותפת תהיה שלילית (ראו שרטוט בהמשך).

לדוגמא, נניח שהניסוי הוא בחירה של אדם באקראי ומדידת גובהו (X) ומשקלו (Y). היות ואנשים גבוהים מהממוצע גם נוטים לשקול יותר מהממוצע ואילו אנשים נמוכים מהממוצע גם נוטים לשקול פחות מהממוצע, מקדם המתאם במקרה זה יהיה חיובי. אין זה אומר שלא ייתכנו גם המצבים האחרים, כלומר אנשים גבוהים מהממוצע ששוקלים פחות מהממוצע או להפך, אך מספרם של אנשים אלו קטן יחסית (ולכן גם ההסתברות לבחור בהם קטנה) והם לא מספיקים בשביל להפוך את הממוצע לשלילי.

בעיה 7.7. מהי השונות המשותפת של המשתנים המקריים X,Y שצפיפותם המשותפת הינה

$$f_{XY}\left(x,y\right)=\begin{cases}\frac{6}{7}\left(x+y\right)^{2} & 0\leq x,y\leq 1\\ 0 & \text{otherwise}\end{cases}$$

הוכחה. נשתמש בנוסחא 7.18. תוחלת המכפלה הינה

$$\mathbb{E}\left(XY\right)=\iint xyf_{XY}\left(x,y\right)\,dydx=\int_{x=0}^{1}\int_{y=0}^{1}\frac{6}{7}xy\left(x+y\right)^{2}\,dydx=\frac{17}{42}$$

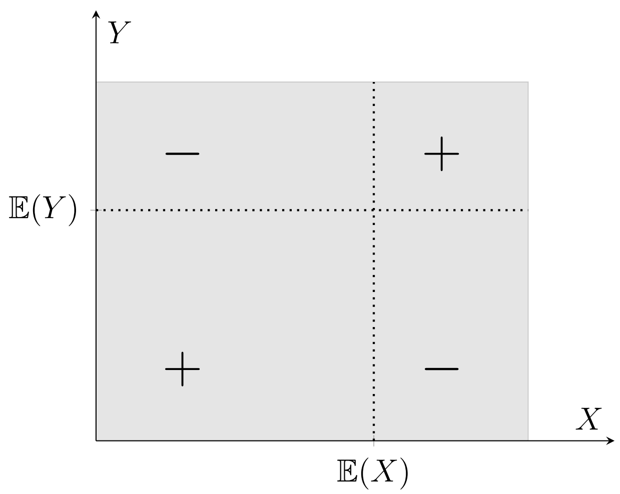
הצפיפות השולית של X בתחום $x\in[0,1]$ הינה

$$f_{X}\left(x\right)=\int_{y=0}^{1}\frac{6}{7}\left(x+y\right)^{2}\,dy=\frac{2}{7}\left(3x^{2}+3x+1\right)$$

ו-0 אחרת. לכן התוחלת של X הינה

$$\mathbb{E}\left(X\right)=\int_{x=0}^{1}x\cdot\frac{2}{7}\left(3x^{2}+3x+1\right)\,dx=\frac{9}{14}$$

מסימטריה בין המשתנים המקריים זוהי גם התוחלת של Y. בשרטוט להלן מופיע תחום ההגדרה של המשתנים המקריים וכן התחומים במישור XY בהם התרומה לשונות המשותפת חיובית ושלילית (הקווים המקווקווים הם התוחלות של כל אחד מהם):



לסיום, השונות המשותפת הינה

$$Cov\left(X,Y\right)=\mathbb{E}\left(XY\right)-\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)=\frac{17}{42}-\left(\frac{9}{14}\right)^{2}=-\frac{5}{588}=-0.0085$$

כלומר השונות המשותפת שלילית וערכים גבוהים של X נוטים להגיע עם ערכים נמוכים יחסית של Y. אינטואיטיבית, ניתן לקבל רמז לכך מהשרטוט. כאשר X מעל תוחלתו, רוב הערכים של Y תורמים גודל שלילי לממוצע ולכן הממוצע יהיה שלילי. יחד עם זאת, חשוב להיזהר עם אינטואיציה שכזו שכן הצפיפות המשותפת, שלא נראית בשרטוט, קובעת את המשקלות של ועבור פונקציות צפיפות אחרות הנותנות משקל רב הרבה יותר לאזור בו $X,Y>\frac{9}{14}$ ניתן לקבל גם שונות משותפת חיובית. □

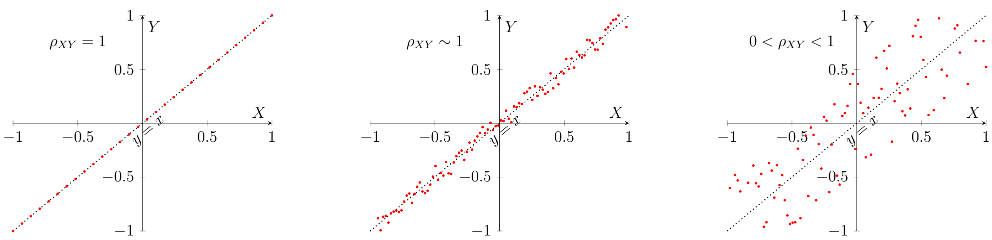
בבעיה **7.7** קיבלנו שהשונות המשותפת הינה שלילית וערכה -0.0085. האם זהו גודל גדול או קטן? כמו לשונות, גם לערך של השונות המשותפת קשה לשייך משמעות שכן היחידות של השונות המשותפת הן לא טבעיות לבעיה. במקרה של השונות, פתרנו בעיה זו על ידי הוצאת שורש והפיכת השונות לסטיית תקן בעלת יחידות דומות ליחידות של המשתנה המקרי עצמו. במקרה של שונות משותפת הוצאת שורש לאו דווקא תעזור, שכן כלל לא בטוח שיש לה יחידות שהן ריבוע של משהו. למשל, כאשר הניסוי הוא בחירה של אדם ומדידת גובהו ומשקלו, לשונות המשותפת יחידות של משקל*גובה והוצאת שורש לא תוביל לגודל עם משמעות. במקום זאת נחלק את השונות המשותפת בסטיית התקן של כל אחד מהמשתנים המקריים לקבלת גודל ללא יחידות כלל. גודל זה מכונה מקדם המתאם.

הגדרה 7.11 (מקדם מתאם)**.** יהי X,Y משתנה מקרי דו-ממדי. מקדם המתאם בין משתנים מקריים אלו מוגדר על ידי

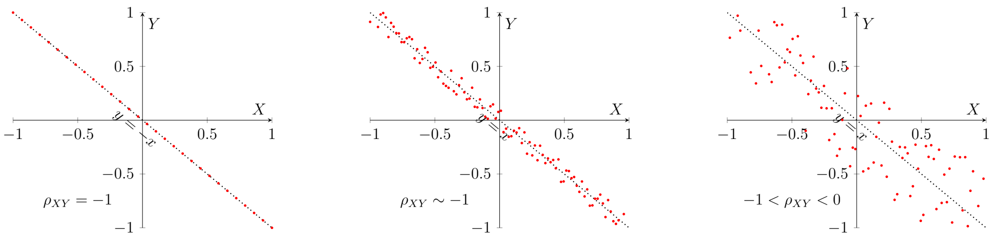
$$\rho\left(X,Y\right)=\frac{Cov\left(X,Y\right)}{\sqrt{\mathbb{V}\left(X\right)\mathbb{V}\left(Y\right)}}$$

לרוב נסמן את מקדם המתאם ב-$\rho_{X,Y}$. מקדם המתאם מודד את היחס בין השונות המשותפת לבין השונויות של כל אחד מהמשתנים המקריים וככל שהוא יותר גדול (בערכו המוחלט) כך הקשר חזק יותר.

לפי אי-שוויון קושי-שוורץ', מתקיים תמיד $\left|Cov\left(X,Y\right)\right|^{2}\leq\mathbb{V}\left(X\right)\mathbb{V}\left(Y\right)$ ולכן תמיד $\left|\rho_{XY}\right|\leq 1$. לפיכך, מקדם המתאם הוא מדד לעוצמת הקשר הלינארי בין המשתנים המקריים וככל שהוא גדול יותר בערכו המוחלט הקשר הלינארי גדול יותר. ככל שמקדם המתאם קרוב יותר ל-1, כך הגדילה המשותפת של המשתנים המקריים חזקה יותר. במקרה הקצה, כאשר $\rho_{XY}=1$, הקשר בין המשתנים המקריים הוא לינארי דטרמיניסטי, כלומר קיימים $a>0,b\in\mathbb{R}$ כך ש-$Y=aX+b$ (ראו תרגיל **7.20**).



באופן דומה, ככל שהמתאם שלילי יותר כך הקשר ההפוך בין המשתנים המקריים גדול יותר. במקרה הקצה, כאשר $\rho_{XY}=-1$, הקשר בין המשתנים המקריים הוא לינארי דטרמיניסטי, כלומר קיימים $a<0,b\in\mathbb{R}$ כך ש-$Y=aX+b$ (ראו תרגיל **7.20**).



אם נמשיך את בעיה 7.7 נקבל שהשונות של כל אחד מהמשתנים המקריים הינה

$$\mathbb{V}(Y) = \mathbb{V}(X) = \int_0^1 x^2 f_X(x)\, dx - \left(\frac{9}{14}\right)^2 = 0.067$$

ולכן מקדם המתאם הינו

$$\rho_{XY} = \frac{-0.0085}{\sqrt{0.067 \cdot 0.067}} = -0.126$$

כך שהקשר היורד בין המשתנים המקרים הוא לא חזק במיוחד. אנחנו נתמקד בדיונים על סימנו של מקדם המתאם ועל ערכי הקצה האפשריים ולא נרחיב את הדיון על עוצמת הקשר ועל ההבדלים בין מתאם של 0.5 ל-0.9. דיון שכזה מתאים יותר לקורסים בסטטיסטיקה כחלק ממדידת עוצמת הקשר הלינארי בין המשתנים המקריים. במקום זאת נציג את מקרה הקצה השלישי האפשרי, המתאר העדר קשר לינארי בין המשתנים המקריים: $\rho_{XY} = 0$.

הגדרה 7.12 (משתנים מקריים בלתי מתואמים). יהיו X, Y משתנים מקריים דו-ממדיים. נאמר שמשתנים מקריים אלו בלתי מתואמים כאשר $\rho_{XY} = 0$ השקול לכך ש-$Cov(X,Y) = 0$.

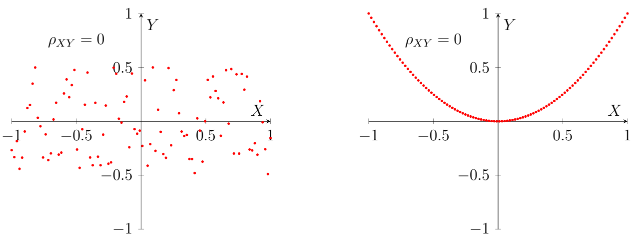
משתנים מקריים בלתי תלויים יהיו בהכרח בלתי מתואמים, שהרי אם X, Y בלתי תלויים אזי ההתפלגות המותנית $X|Y$ שווה להתפלגות השולית X ולפי משפט התוחלת השלמה נקבל

$$\mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(XY|Y)) \overset{eq.7.13}{=} \mathbb{E}(Y\mathbb{E}(X|Y)) = \mathbb{E}(X)\mathbb{E}(Y)$$

כלומר $Cov(X,Y) = 0$. תוצאה זו טריוויאלית - כאשר המשתנים המקריים בלתי תלויים אין ביניהם קשר ובפרט אין ביניהם קשר לינארי או קשר של גדילה משותפת. כאשר אחד גדל השני לא מושפע מכך כלל. הכיוון השני אינו נכון, משתנים מקריים יכולים להיות בלתי מתואמים ובו זמנית תלויים. למשל, $X \sim N(0,1)$ ו-$Y = X^2$. משתנים מקריים אלו בהכרח תלויים (Y הוא פונקציה דטרמיניסטית של X) אך $\mathbb{E}(X) = 0$ וכן

$$\mathbb{E}(XY) = \mathbb{E}\left(X \cdot X^2\right) = \int_{-\infty}^{\infty} x^3 e^{-\frac{x^2}{2}}\, dx = 0$$

ולכן גם $Cov(X,Y) = 0$. במקרה זה המשתנים המקריים בלתי מתואמים מאחר ואין לתלות ביניהם שום מרכיב לינארי. אמנם כאשר X גדול יחסית גם Y גדול יחסית, אך ההפך גם נכון - כאשר X קטן יחסית גם Y גדול יחסית.



תרגיל 7.16. חשבו את המתאם בין המשתנים המקריים המופיעים בתרגיל 7.15.

תרגיל 7.17. נתונה פונקציית צפיפות משותפת

$$f_{XY}(x,y)=\begin{cases}6(1-x-y) & 0\le x,y\cap x+y\le 1\\ 0 & \text{otherwise}\end{cases}$$

ענו, מבלי לחשב, מהו סימנו של מקדם המתאם בין X ל-Y? לאחר מכן, חשבו אותו וודאו שצדקתם.

בדומה לתוחלת ולשונות, גם לשונות המשותפת תכונות המזכירות לינאריות וחיבוריות:

משפט 7.7. *יהיו X,Y,Z משתנים מקריים ויהיו $a,b\in\mathbb{R}$. השונות המשותפת היא:*

1. *חיבורית:*

$$Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z) \tag{7.20}$$

2. *"לינארית":*

$$Cov(aX+b,Y)=aCov(X,Y) \tag{7.21}$$

שילוב של שתי תכונות אלו מכונה בי-לינאריות, המזכיר פתיחת סוגריים:

$$Cov\left(\sum_{i=1}^{n}a_iX_i+b_i,\sum_{j=1}^{m}c_jY_j+d_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{m}a_ic_jCov(X_i,Y_j) \tag{7.22}$$

הוכחה. נוכיח את החיבוריות והלינאריות. הבי-לינאריות מתקבלת מהפעלה חוזרת של תכונות אלו ולכן הוכחתה טכנית בלבד.

$$\begin{aligned}Cov(X+Y,Z)&=\mathbb{E}((X+Y)Z)-\mathbb{E}(X+Y)\mathbb{E}(Z)\\&=\mathbb{E}(XZ)+\mathbb{E}(YZ)-\mathbb{E}(X)\mathbb{E}(Z)-\mathbb{E}(Y)\mathbb{E}(Z)\\&=Cov(X,Z)+Cov(Y,Z)\end{aligned}$$

וכן

$$\begin{aligned}Cov(aX+b,Y)&=\mathbb{E}((aX+b)Y)-\mathbb{E}(aX+b)\mathbb{E}(Y)\\&=a\mathbb{E}(XY)+b\mathbb{E}(Y)-a\mathbb{E}(X)\mathbb{E}(Y)-b\mathbb{E}(Y)\\&=a(\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y))=aCov(X,Y)\end{aligned}$$

□

בזכות תכונות אלו ניתן לפתח נוסחא כללית עבור חישוב שונות של סכום.

משפט 7.8. *לכל שני משתנים מקריים X, Y מתקיים*

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2Cov(X,Y) \quad (7.23)$$

ובאופן כללי עבור n משתנים מקריים $X_1, \ldots, X_n$ מתקיים

$$\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i) + \sum_{i\neq j} Cov(X_i, X_j) = \sum_{i=1}^{n} \mathbb{V}(X_i) + 2\sum_{i<j} Cov(X_i, X_j) \quad (7.24)$$

הוכחה. נוכיח את הנוסחא עבור סכום של שני משתנים מקריים, כאשר התוצאה הכללית מתקבלת בדרך דומה. נשתמש בכך ששונות של משתנה מקרי שווה לשונות המשותפת שלו עם עצמו, ולכן

$$\begin{aligned}\mathbb{V}(X+Y) &= Cov(X+Y, X+Y) = Cov(X,X) + Cov(X,Y) + Cov(Y,X) + Cov(Y,Y)\\ &= \mathbb{V}(X) + \mathbb{V}(Y) + 2Cov(X,Y)\end{aligned}$$

□

כאשר המשתנים המקריים בלתי מתואמים (ובפרט, בלתי תלויים) שונות הסכום שווה לסכום השונויות, כפי שראינו בפרק 5. החידוש הוא שמשפט זה מאפשר את חישוב השונות של הסכום גם במקרה הכללי, כאשר המשתנים המקריים תלויים ומתואמים.

במקרים רבים ניעזר במשפט זה לחישוב השונות המשותפת ומקדם המתאם מתוך השונות. הסיבה לכך היא שחישוב השונות המשותפת ישירות מההגדרה עבור משתנים מקריים בדידים הוא תהליך מסובך ודורש סכימה כפולה על כל המכפלות האפשריות של איברים בתומכים של המשתנים המקריים. סכימה זו לא תמיד נוחה לביצוע ולא תמיד יש נוסחא כללית המתארת את הסכום. תחת זאת, במקרה הבדיד נוכל לזהות את המשתנה המקרי $X+Y$ מתוך הסיפור וההבנה מה משתנה מקרי זה מייצג ואז להשתמש בנוסחא 7.23 לחישוב השונות המשותפת. לעומת זאת, במקרה הרציף נוסחא 7.18 נוחה לחישוב כפי שראינו בדוגמאות הקודמות.

בעיה 7.8. בקערה יש 10 תפוחים, 10 אגסים ו-10 בננות. מוציאים מהקערה בלי החזרה 10 פירות. נסמן ב-X את מספר התפוחים שנבחרו וב-Y את מספר האגסים שנבחרו. מהו מקדם המתאם ביניהם?

הוכחה. ראשית נעיר שאנחנו מצפים לתשובה שלילית. ככל שהוצאו יותר תפוחים כך יש פחות הוצאות בהן ניתן להוציא אגסים, לכן ערכים גבוהים של X יגיעו עם ערכים נמוכים של Y ולהפך. נחשב את המתאם מפורשות ולשם כך נסמן ב-Z את מספר הבננות שנבחרו.

בסך הכל נבחרו 10 פריטים לכן $X+Y+Z=10$ תמיד. כמו כן, לכל אחד מהמשתנים המקריים הללו אותה ההתפלגות השולית (ההבדל היחיד הוא בשם הפרי) ולכן $\mathbb{V}(X)=\mathbb{V}(Y)=\mathbb{V}(Z)$. זאת ועוד, $\mathbb{V}(X+Y)=\mathbb{V}(10-Z)=\mathbb{V}(Z)$ ולכן השונות המשותפת הינה (נוסחא 7.23):

$$Cov(X,Y) = \frac{\mathbb{V}(X+Y) - \mathbb{V}(X) - \mathbb{V}(Y)}{2} = -\frac{\mathbb{V}(X)}{2}$$

מקדם המתאם מתקבל על ידי חלוקה במכפלת סטיות התקן וערכו

$$\rho_{XY} = \frac{-\mathbb{V}(X)/2}{\sqrt{\mathbb{V}^2(X)}} = -\frac{1}{2}$$

הצלחנו לחשב את מקדם המתאם של המשתנים המקריים מבלי לחשב את ההתפלגות שלהם או אף ערך אחר דוגמת התוחלת או השונות. ניתן היה לזהות גם שכל אחד מהמשתנים המקריים X, Y, Z מתפלג $HG(30, 10, 10)$ (דוגמים 10 פריטים מקערה עם 30 ושואלים כמה מהם הם הפרי הספציפי המעניין את המשתנה המקרי) וכן $X+Y \sim HG(30, 20, 10)$ (הפעם יש 20 פריטים מיוחדים: גם התפוחים וגם האגסים) ולהציב את השוניות המתאימות בנוסחא. התוצאה, כמובן, תהיה זהה. לעומת זאת, חישוב תוחלת המכפלה לא פרקטית שכן ישנן כ-50 מכפלות אפשריות של ערכי X ב-Y. $\square$

תרגיל 7.18. בהמשך לתרגיל 7.6:

1. חשבו את מקדם המתאם בין Xל-Y.
2. נסמן ב-Z את סכום שני הקלפים העליונים. חשבו את מקדם המתאם בין X ל-Z.

תרגיל 7.19. בכד כדורים אדומים שחורים ולבנים (4 מכל צבע). מוציאים מהכד 4 כדורים עם החזרה. יהיו X,Y,Z מספר הכדורים האדומים, השחורים והלבנים שהוצאו בהתאמה.

1. כיצד משתנים מקריים אלו מתפלגים?
2. כיצד מתפלג $X+Y$? מהי שונותו?
3. מהו מקדם המתאם בין X ל-Y?
4. מהו מקדם המתאם בין $X+Y$ לבין Z? (לחשוב לפני שמחשבים!)

תרגיל 7.20. הוכיחו שכאשר מקדם המתאם הוא 1 הקשר בין המשתנים המקריים לינארי. כלומר, הוכיחו שכאשר $\rho_{XY} = \pm 1$ קיימים $a, b \in \mathbb{R}$ כך ש-$Y = aX + b$, כאשר סימנו של a זהה לסימנו של ρ_{XY}.
רמז: מצאו את המינימום של הפונקציה $g(a) = \mathbb{V}(Y - aX)$.

בזכות השונות המשותפת ונוסחא 7.24 ניתן לחשב שונות של סכום משתנים מקריים באופן דומה לחישוב תוחלת של סכום אותו הצגנו בפרק 5 .

בעיה 7.9 (המזכירה המבולבלת)**.** למזכירה יש n מכתבים אישיים ו-n מעטפות עם כתובות. המזכירה מכניסה את המכתבים באקראי למעטפות ושולחת אותן בדואר. נסמן ב-X את מספר הנמענים שקיבלו מכתב שיועד עבורם. מהי התוחלת והשונות של X?

הוכחה. חישוב ההתפלגות של X מסובכת (ראו למשל את תרגיל 4.23 המתאים למקרה $n=3$). במקום זאת נשים לב ש-X סופר את מספר הנמענים שקיבלו מכתב המיועד להם ולכן ניתן להציג אותו כסכום של אינדיקטורים הבודקים האם כל נמען קיבל את המכתב שלו או לא. נגדיר ב-$\mathbf{1}_i$ אינדיקטור לכך שנמען $i = 1, \ldots, n$ קיבל את המכתב שמיועד לו. ההסתברות של נמען כלשהו לקבל את המכתב שלו הינה $\frac{1}{n}$ ולכן לפי תוחלת ושונות של אינדיקטור (נוסחאות 5.10, 5.17)

$$\mathbb{E}(\mathbf{1}_i) = \frac{1}{n} \qquad \mathbb{V}(\mathbf{1}_i) = \frac{1}{n}\left(1 - \frac{1}{n}\right)$$

התוחלת של X הינה

$$\mathbb{E}\left(X\right)=\mathbb{E}\left(\sum_{i=1}^{n}\mathbf{1}_{i}\right)=\sum_{i=1}^{n}\mathbb{E}\left(\mathbf{1}_{i}\right)=\sum_{i=1}^{n}\frac{1}{n}=1$$

והשונות הינה

$$\mathbb{V}\left(X\right)=\mathbb{V}\left(\sum_{i=1}^{n}\mathbf{1}_{i}\right)=\sum_{i=1}^{n}\mathbb{V}\left(\mathbf{1}_{i}\right)+2\sum_{i<j}Cov\left(\mathbf{1}_{i},\mathbf{1}_{j}\right)=1-\frac{1}{n}+2\sum_{i<j}Cov\left(\mathbf{1}_{i},\mathbf{1}_{j}\right)$$

נשאר לחשב את השונות המשותפת של כל זוג אינדיקטורים. לשם כך נחשב את התוחלת של המכפלה $\mathbf{1}_{i}\mathbf{1}_{j}$. מכפלה זו היא מכפלה של שני משתנים מקריים שכל אחד מהם מקבל את הערכים 0 או 1, ולכן גם המכפלה יכולה להיות או 0 או 1, והיא שווה ל-1 רק כאשר שני האנשים, גם i וגם j, קיבלו את המכתב שלהם. ההסתברות שזה יקרה הינה

$$Pr\left(\mathbf{1}_{i}\mathbf{1}_{j}=1\right)=\frac{1}{n}\cdot\frac{1}{n-1}$$

ולכן

$$\begin{aligned}Cov\left(\mathbf{1}_{i},\mathbf{1}_{j}\right)&=\mathbb{E}\left(\mathbf{1}_{i}\mathbf{1}_{j}\right)-\mathbb{E}\left(\mathbf{1}_{i}\right)\mathbb{E}\left(\mathbf{1}_{j}\right)=\frac{1}{n}\cdot\frac{1}{n-1}-\left(\frac{1}{n}\right)^{2}\\&=\frac{1}{n}\left(\frac{1}{n-1}-\frac{1}{n}\right)=\frac{1}{n^{2}\left(n-1\right)}\end{aligned}$$

מספר הזוגות השונים של אינדיקטורים $i<j$ שקיימים הוא כמספר המספרים השונים שניתן לבחור מהקבוצה $\{1,\ldots,n\}$ כך שהקטן יהיה i והגדול יהיה j. מספר זוגות זה הינו $\binom{n}{2}$ שכן זהו מספר האפשרויות לבחור שני מספרים ללא סדר וללא החזרה ולאחר שמספרים אלו נבחרו יש רק דרך אחת להקצות אותם לאינדקסים i,j כך שהקטן יהיה i. לכן

$$\begin{aligned}\mathbb{V}\left(X\right)&=1-\frac{1}{n}+2\binom{n}{2}\frac{1}{n^{2}\left(n-1\right)}=1-\frac{1}{n}+2\cdot\frac{n\left(n-1\right)}{2}\cdot\frac{1}{n^{2}\left(n-1\right)}\\&=1-\frac{1}{n}+\frac{1}{n}=1\end{aligned}$$

□

תרגיל 7.21. יהי $X\sim HG\left(N,D,n\right)$. חשבו את $\mathbb{V}\left(X\right)$ באמצעות פירוק לאינדיקטורים (ראו את הגדרת האינדיקטורים במשפט 6.7).

תרגיל 7.22. מטילים קוביה הוגנת n פעמים. נסמן X – מספר הפעמים שהתקבלה התוצאה 1, Y – מספר הפעמים שהתקבלה התוצאה 6.

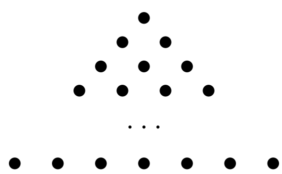
1. כיצד מתפלגים X,Y? מהן שונותיהם?

2. **נגדיר אינדיקטורים $\mathbf{1}_i^X$ השווים ל-1 אם יצא 1 בהטלה ה-i ו-$\mathbf{1}_i^Y$ השווים ל-1 אם יצא 6 בהטלה ה-i. בטאו את X,Y באמצעות אינדיקטורים אלו.**
3. **הסבירו מדוע $Cov\left(\mathbf{1}_i^X,\mathbf{1}_j^Y\right)=0$ עבור $i\neq j$ ומדוע $\mathbb{E}\left(\mathbf{1}_i^X\mathbf{1}_i^Y\right)=0$.**
4. **חשבו את השונות המשותפת $Cov\left(\mathbf{1}_i^X,\mathbf{1}_i^Y\right)$.**
5. **חשבו את מקדם המתאם ρ_{XY}. חוו דעתכם על סימנו.**

7.5 שאלות מסכמות

תרגיל 7.23. על מגש נמצאים 3 נקניקים, 2 פרוסות פסטרמה ו-4 סטייקים. הכלבה קרמבו מתגנבת לקערה ואוכלת באקראי שני מוצרים לפני ששמים לב אליה ועוצרים אותה. נסמן ב-X את מספר הנקניקים שקרמבו אכלה וב-Y את מספר פרוסות הפסטרמה אותן קרמבו אכלה. מהי ההתפלגות המשותפת של X,Y? מהן ההתפלגויות השוליות של כל אחד מהמשתנים המקריים?

תרגיל 7.24. נתונים כדורים המסודרים במבנה משולש, כך שבשורה ה-$i \in \{1,\dots,n\}$ יש i כדורים כמוראה בתרשים:



בוחרים באקראי את אחד מהכדורים במשולש. נסמן את השורה של הכדור הנבחר ב-X ואת מיקומו בשורה ב-Y (כאשר המקום הכי שמאלי הוא 1).

1. מה הסיכוי שהכדור נבחר מאחת משלושת השורות הראשונות אם ידוע שהוא נבחר מאחד משלושת המיקומים הראשונים בשורה ($n \geq 3$)?
2. מהי ההתפלגות המשותפת של X,Y?
3. מהן ההתפלגויות השוליות של כל אחד מהמשתנים המקריים?
4. האם משתנים מקריים אלו תלויים או בלתי תלויים?

תרגיל 7.25. מטילים מטבע הוגן פעמיים. נסמן ב-$\mathbf{1}_i$ את האינדיקטור לכך שיצא עץ בהטלה ה-$i \in \{1,2\}$. נגדיר $X = \mathbf{1}_1 + \mathbf{1}_2$ ו-$Y = |\mathbf{1}_1 - \mathbf{1}_2|$. האם X,Y תלויים? האם X,Y מתואמים?

תרגיל 7.26. יהיו X,Y שני משתנים מקריים רציפים. הוכיחו או הפריכו: גם סכומם, $X+Y$, הוא משתנה מקרי רציף.

תרגיל 7.27. נוסע מגיע לתחנה ריקה. הזמן עד הגעת האוטובוס הראשון לתחנה מתפלג מעריכית עם פרמטר λ_1 ואילו הזמן עד הגעת המונית הראשונה מתפלג מעריכית עם פרמטר λ_2. בהנחה שזמני ההגעה הללו בלתי תלויים, מהי ההסתברות שהרכב הראשון שיגיע לתחנה יהיה אוטובוס?

תרגיל 7.28 (המחט של בופון)**.** על הרצפה משורטטים אינסוף קווים ישרים מקבילים במרחק d זה מזה. זורקים על הרצפה מחט באורך $l \leq d$ באקראי, כלומר הזווית בו נוחתת המחט ביחס לקווים היא אחידה וכך גם המרחק בין מרכז המחט לישר הקרוב ביותר. מה הסיכוי שהמחט חוצה את אחד הקווים?

תרגיל 7.29. נתונים שני משתנים מקריים בלתי תלויים $X \sim N\left(0,\sigma_X^2\right), Y \sim N\left(0,\sigma_Y^2\right)$. חשבו:

1. $P(X>Y)$
2. $P(X>Y>0)$

תרגיל 7.30. יהיו $X \sim Exp\left(1\right)$ ו-$Y \sim Exp\left(1\right)$ שני משתנים מקריים בלתי תלויים. כיצד מתפלג $Z = \frac{X}{X+Y}$?

תרגיל 7.31. יהי X,Y משתנה מקרי דו ממדי עם פונקציית הצפיפות המשותפת

$$f_{XY}\left(x,y\right)=\begin{cases} x+y & x,y\in[0,1] \\ 0 & \text{otherwise}\end{cases}$$

מהי הצפיפות המותנית $Y|X=x$ ומהי התוחלת המותנית $\mathbb{E}\left(Y|X=x\right)$?

תרגיל 7.32. חרק מטיל $k\in\mathbb{N}$ ביצים בהסתברות $p(1-p)^{k-1}$ עבור $p=0.1$. כל ביצה בוקעת בהסתברות 0.5 ללא תלות בביצים האחרות. נסמן ב- Y את מספר הביצים שבקעו.

1. כיצד מתפלג Y אם הוטלו 7 ביצים?
2. חשבו את התוחלת והשונות של Y.
3. מהי ההסתברות $Pr(Y=0)$?

תרגיל 7.33. 50 זוגות מסתדרים באקראי בשורה. מהי השונות של מספר האנשים העומדים בין אדון וגברת סמית'?
רמז: תרגיל **5.14**.

תרגיל 7.34. מתוך 50 זוגות נשואים נבחרים באקראי 30 אנשים. מהי התוחלת והשונות של מספר הזוגות שנבחרו בשלמותם (כלומר, שנבחרו שני בני הזוג)?

תרגיל 7.35. מהמר מגיע לקזינו ומטיל קוביה. אם יוצא 1 – הוא זוכה ב-10 דולרים. אם יוצא 2,3,4 או 5 – הוא מפסיד דולר אחד. אם יוצא 6 – הוא מפסיד 5 דולרים. המהמר משחק במשחק שוב ושוב עד הפעם הראשונה שהוא זוכה. מהי תוחלת הרווח של המהמר בסיום המשחק?

תרגיל 7.36. באוניברסיטה בה ישנו בית קפה יחיד לומדים 10,000 סטודנטים. בכל יום, כל אחד מהסטודנטים קונה כוס קפה בהסתברות 0.35 ללא תלות בסטודנטים האחרים. כוס הקפה יכולה להיות קטנה (5 ש"ח) או גדולה (8 ש"ח) בהסתברות שווה. חשבו את התוחלת והשונות של הרווח היומי של בית הקפה.

תרגיל 7.37. אל תחנת אוטובוס אחת מגיעים בכל בוקר 100 אנשים ושני אוטובוסים. לכל נוסע יש סיכוי של 0.75 להתעורר מוקדם ולהגיע לאוטובוס הראשון. אביטל, אחת הנוסעות, טוענת שהאוטובוסים צפופים מדי. על מנת להוכיח זאת היא בוחרת באקראי אחד מ-100 הנוסעים ושואלת אותו כמה נוסעים נסעו איתו באוטובוס באותו היום. נסמן את תשובתו ב-X. מהי תוחלת התשובה שתקבל אביטל, כלומר מהי $\mathbb{E}\left(X\right)$? האם היא תואמת את טענת משרד התחבורה, שמספר הנוסעים הממוצע הינו $\frac{100}{2}=50$? מדוע?
הערה: הניחו שלא עולים על האוטובוס נוסעים נוספים פרט ל-100 הנוסעים המקוריים.

תרגיל 7.38. יהיו X,Y שני משתנים מקריים בלתי תלויים ונגדיר $Z=X+Y$. הוכיחו כי הפונקציה היוצרת של Z היא המכפלה של הפונקציות היוצרות של X ו-Y, כלומר

$$M_{X+Y}\left(t\right)=M_X\left(t\right)\cdot M_Y\left(t\right) \tag{7.25}$$

תרגיל 7.39. נתונים שני מטבעות, מטבע הוגן ומטבע שהסיכוי לקבל בו עץ הינו 0.25. בוחרים באקראי את אחד המטבעות ומטילים אותו שוב ושוב. נסמן ב-$\mathbf{1}_i$ את האינדיקטור למאורע שהמטבע נפל על עץ בהטלה ה-i. האם הסדרה $\mathbf{1}_1,\mathbf{1}_2,\ldots$ מקיימת את החוק החלש של המספרים הגדולים (משפט **5.10**)?

7.6 פתרונות לשאלות הפרק

תרגיל 7.1. ראשית נשים לב שלסכום מספרים ולהפרש יש בדיוק את אותה זוגיות (כלומר הם שקולים מודולו 2). כמו כן, לכל זוג אפשרי של תוצאות (פרט למקרה בו $Y = 0$) מתאימות רק שני צירופים מתוך ה-36. למשל, המאורע $X = 4, Y = 2$ מתקבל רק כאשר אחת הקוביות היא 3 והשניה היא 1 והסתברותו היא $\frac{2}{36}$. המקרה בו $Y = 0$ אפשרי רק כששתי ההטלות זהות וערכי ה-X המתאימים הם $2, 4, 6, 8, 10, 12$. ההסתברות לכל תוצאה כזו היא $\frac{1}{36}$. לסיכום:

$$P_{X,Y}(k,l) = \begin{cases} \frac{1}{36} & l = 0; k = 2, 4, \dots, 12 \\ \frac{2}{36} & l = 1, \dots, 5;\ k \equiv_2 l \\ 0 & \text{otherwise} \end{cases}$$

□

תרגיל 7.2. המאורע $\{X = k, Y = l\}$ עבור $k < l$ מתקבל כאשר ב-$k-1$ ההטלות הראשונות יצאו רק 2,4, בהטלה ה-k יצא 6, בהטלות הבאות עד ההטלה ה-l יצאו רק זוגיים ובהטלה ה-l יצא מספר אי זוגי. הסתברותו של מאורע זה הינה

$$Pr(X = k, Y = l) = \left(\frac{2}{6}\right)^{k-1} \frac{1}{6} \left(\frac{1}{2}\right)^{l-k-1} \frac{1}{2}$$

באופן אנלוגי כאשר $k > l$ ההסתברות הינה

$$Pr(X = k, Y = l) = \left(\frac{2}{6}\right)^{l-1} \frac{1}{2} \left(\frac{5}{6}\right)^{k-l-1} \frac{1}{6}$$

וכמובן המצב בו $l = k$ לא ייתכן כי במקרה כזה באותה הטלה יצא גם 6 וגם מספר אי זוגי. לסיכום, ההתפלגות המשותפת עבור כל $k, l \in \mathbb{N}$ הינה

$$P_{X,Y}(k,l) = \begin{cases} \left(\frac{4}{6}\right)^k \left(\frac{1}{2}\right)^{l+1} & k < l \\ \frac{3}{10} \left(\frac{2}{5}\right)^l \left(\frac{5}{6}\right)^k & k > l \\ 0 & k = l \end{cases}$$

□

תרגיל 7.3. ישנן בסך הכל 4 אפשרויות ולכן נחשב אותן אחת-אחת. ראשית,

$$Pr(\mathbf{1}_A = 0, \mathbf{1}_B = 0) = \left(\frac{1}{3}\right)^5 = \frac{1}{243}$$

שכן במצב זה התייר סעד את כל ארוחותיו במסעדה השלישית. לפיכך, נראה כי הרבה יותר נוח לחשב מאורעות בהן התייר לא סועד במסעדה מסויימת מאשר כאשר הוא כן סועד בה, שכן אז אין צורך לבחור את מספר הסעודות

והערבים בהם סעודות אלו התקיימו. נשתמש בתובנה זו ונחשב את המאורע שהוא סעד בשתי המסעדות:

$$\begin{aligned}Pr\left(\mathbf{1}_A=1,\mathbf{1}_B=1\right) &= 1-Pr\left(\overline{\mathbf{1}_A=1\cap\mathbf{1}_B=1}\right)=1-Pr\left(\overline{\mathbf{1}_A=1}\cup\overline{\mathbf{1}_B=1}\right)\\ &= 1-Pr\left(\mathbf{1}_A=0\cup\mathbf{1}_B=0\right)\\ (eq.\,2.5) &= 1-\left(Pr\left(\mathbf{1}_A=0\right)+Pr\left(\mathbf{1}_B=0\right)-Pr\left(\mathbf{1}_A=0\cap\mathbf{1}_B=0\right)\right)\end{aligned}$$

ההסתברות שהוא לא אכל במסעדה מסויימת הינה

$$Pr\left(\mathbf{1}_A=0\right)=Pr\left(\mathbf{1}_B=0\right)=\left(\frac{2}{3}\right)^5$$

ואת ההסתברות שהוא לא סעד בשתיהן כבר חישבנו. לפיכך

$$Pr\left(\mathbf{1}_A=1,\mathbf{1}_B=1\right)=1-2\left(\frac{2}{3}\right)^5+\left(\frac{1}{3}\right)^5=\frac{180}{243}$$

את שתי ההסתברויות החסרות נחשב באמצעות הסתברות מותנית:

$$Pr\left(\mathbf{1}_A=1,\mathbf{1}_B=0\right)=Pr\left(\mathbf{1}_A=1|\mathbf{1}_B=0\right)Pr\left(\mathbf{1}_B=0\right)$$

בהינתן שהוא לא סעד במסעדה B, הסיכוי שהוא כן סעד במסעדה A הינו $1-\left(\frac{1}{2}\right)^5$ (המשלים של המאורע שהוא לא סעד בה). לכן

$$Pr\left(\mathbf{1}_A=1,\mathbf{1}_B=0\right)=\left(1-\left(\frac{1}{2}\right)^5\right)\left(\frac{2}{3}\right)^5=\frac{31}{243}$$

וזו כמובן גם ההסתברות של המאורע $\{\mathbf{1}_A=0,\mathbf{1}_B=1\}$. לסיכום:

$$P_{\mathbf{1}_A,\mathbf{1}_B}\left(a,b\right)=\begin{cases}\frac{1}{243} & a=b=0\\ \frac{31}{243} & (a,b)\in\left\{\left(1,0\right),\left(0,1\right)\right\}\\ \frac{180}{243} & a=b=1\end{cases}$$

$\square$

תרגיל 7.4. ניתן לפתור שאלה זו בשתי דרכים. מאחר ומספר האפשרויות קטן, ניתן לבדוק את כולן, למשל על ידי הצבת ערכים ושרטוט ההתפלגות המשותפת בטבלה:

$Y\backslash X$	**1**	**2**	P_Y
2	$\frac{2}{27}$	$\frac{4}{27}$	$\frac{2}{9}$
3	$\frac{3}{27}$	$\frac{6}{27}$	$\frac{1}{3}$
4	$\frac{4}{27}$	$\frac{8}{27}$	$\frac{4}{9}$
P_X	$\frac{1}{3}$	$\frac{2}{3}$	**1**

קל לראות שכל משבצת בהתפלגות המשותפת שווה למכפלת הערכים המתאימים בהתפלגויות השוליות, ולכן המשתנים המקריים הללו בלתי תלויים.

אפשרות נוספת היא חישוב ההתפלגות השולית של כל משתנה מקרי ובדיקת התנאי המופיע בנוסחא 7.5. לכל $x \in supp\left(X\right)$ מתקיים

$$P_X\left(x\right) = \sum_{y=2}^{4} \frac{xy}{27} = \frac{x}{27}\left(2+3+4\right) = \frac{9x}{27}$$

ולכל $y \in supp\left(Y\right)$ מתקיים

$$P_Y\left(y\right) = \sum_{x=1}^{2} \frac{xy}{27} = \frac{y}{27}(1+2) = \frac{3y}{27}$$

ואכן מתקיים

$$P_X\left(x\right) P_Y\left(y\right) = \frac{9x}{27} \cdot \frac{3y}{27} = \frac{xy}{27} = P_{XY}\left(x,y\right)$$

כך שהמשתנים המקריים בלתי תלויים. □

תרגיל 7.5. 1. מרחב המדגם כולל 9 אפשרויות:

$$\Omega = \left\{\left(1,2,3\right), \left(1,3,2\right), (2,1,3), (2,3,1), (3,1,2), (3,2,1), (1,1,1), (2,2,2), (3,3,3)\right\}$$

2. לפי מרחב המדגם, כל קואורדינטה מקבלת כל ערך 3 פעמים ולכן עבור כל קואורדינטה k וכל ערך אפשרי $l \in \{1,2,3\}$ נקבל $Pr\left(X_k = l\right) = \frac{1}{3}$.
כל צירוף של שני ערכים מתאים לשלישיה יחידה מתוך 9 השלישיות האפשריות. למשל אם הקואורדינטה הראשונה היא 1 והקואורדינטה השניה היא 3, הרי שבהכרח מדובר על הוקטור $(1,3,2)$. לכן $Pr\left(X_i = k, X_j = l\right) = \frac{1}{9}$ וזו בדיוק מכפלת ההתפלגויות השוליות. לכן לכל $i \neq j$ המשתנים בלתי תלויים.
3. שלושת המשתנים תלויים. למשל, אם ידוע שהקואורדינטה הראשונה היא 2 והקואורדינטה השניה היא 2, מדובר בהכרח על השלישיה $(2,2,2)$. לפיכך,

$$Pr\left(X_3 = 2 | X_1 = 2, X_2 = 2\right) = 1 \neq Pr\left(X_3 = 2\right) = \frac{1}{3}$$

דרך נוספת להשתכנע בכך היא על ידי חיפוש אפסים. למשל, אין שלישיה מהצורה $(1,1,2)$ ולכן

$$Pr\left(X_1 = 1, X_2 = 1, X_3 = 2\right) = 0 \neq Pr(X_1 = 1)Pr(X_2 = 1)Pr(X_3 = 2)$$

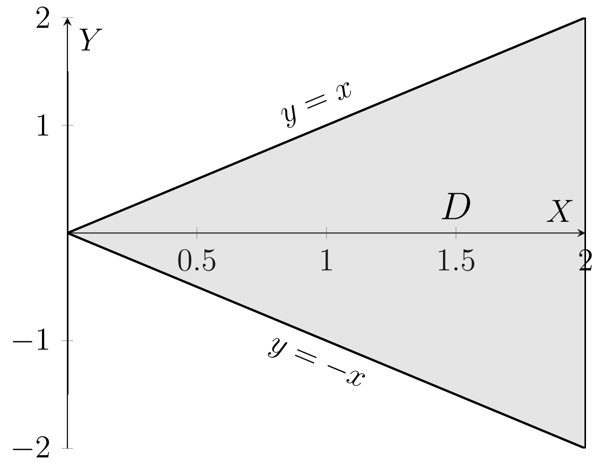
□

תרגיל 7.6. מרחב המדגם כולל בחירה של 2 מספרים מתוך 5 עם סדר (הראשון הוא הקלף העליון, השני הוא התחתון) וללא החזרה ולכן גודלו $\binom{5}{2}2! = 20$. לכל זוג ערכים יש הסתברות שווה להתקבל ולכן ההתפלגות המשותפת הינה

$$P_{X,Y}\left(k,l\right) = \begin{cases} \frac{1}{20} & k \neq l; k,l \in \{1,2,3,4,5\} \\ 0 & k = l \end{cases}$$

הקלף העליון יכול להיות בהסתברות שווה כל אחד מהמספרים ולכן $X \sim U\left\{1,\ldots,5\right\}$ ובאופן זהה $Y \sim U\left\{1,\ldots,5\right\}$. ניתן כמובן לחשב זאת גם על ידי סכימה של כל האפשרויות ישירות לפי נוסחא 7.3. □

תרגיל 7.7. **ניתן לראות כי התחום D מתאר משולש שקודקודיו $(0,0),(2,2),(2,-2)$.**



1. **מדרישת הנרמול נקבל**

$$1=\iint_D f_{XY}(x,y)\ dxdy=\int_{x=0}^{2}\int_{y=-x}^{x}cx(x-y)\,dydx=2c\int_0^2 x^3\,dx=8c$$

ולכן $c=\frac{1}{8}$.

2. **נקבל את הצפיפויות השוליות על ידי אינטגרציה. הצפיפות היא אפס כאשר** $x\notin[0,2]$ **וכאשר** $x\in[0,2]$ **מתקיים:**

$$f_X(x)=\int_{-\infty}^{\infty}f_{XY}(x,y)\ dy=\int_{-x}^{x}\frac{1}{8}x(x-y)\,dy=\frac{x^3}{4}$$

ולכן

$$\mathbb{E}(X)=\int_0^2 xf_X(x)\,dx=\frac{1}{4}\int_0^2 x^4\,dx=1.6$$

באופן דומה:

$$f_Y(y)=\int_{-\infty}^{\infty}f_{XY}(x,y)\,dx=\begin{cases}\frac{1}{8}\int_{-y}^{2}x(x-y)\,dx & y\in[-2,0]\\ \frac{1}{8}\int_{y}^{2}x(x-y)\,dx & y\in[0,2]\\ 0 & y\notin[-2,2]\end{cases}=\begin{cases}\frac{5y^3}{48}-\frac{y}{4}+\frac{1}{3} & y\in[-2,0]\\ \frac{y^3}{48}-\frac{y}{4}+\frac{1}{3} & y\in[0,2]\\ 0 & y\notin[-2,2]\end{cases}$$

ולכן

$$\mathbb{E}(Y)=\int_{-2}^{0}\frac{5y^3}{48}-\frac{y}{4}+\frac{1}{3}\,dy+\int_0^2\frac{y^3}{48}-\frac{y}{4}+\frac{1}{3}\,dy=-\frac{8}{15}$$

□

תרגיל 7.8. 1. מתכונת הנרמול

$$1 = \int_{x=1}^{2} \int_{y=0}^{3} c(5 - x - y)\, dxdy = 6c$$

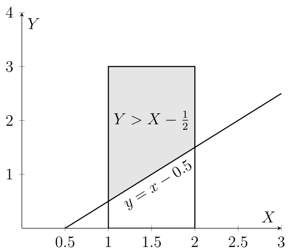
ולכן $c = \frac{1}{6}$.

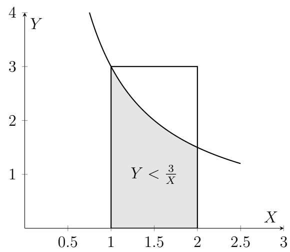
2. נמצא את הצפיפויות השוליות לפי ההגדרה:

$$f_X\left(x\right) = \int_{y=0}^{3} f_{XY}\left(x, y\right)\, dy = \frac{7}{4} - \frac{x}{2}$$

בתחום $x \in [1, 2]$ ו-0 אחרת. ובאותו אופן $f_Y(y) = \frac{7}{12} - \frac{y}{6}$ כאשר $y \in [0, 3]$ ו-0 אחרת.

3. נשרטט את שני המאורעות המבוקשים בשאלה:





לפיכך:

$$Pr\left(XY < 3\right) = Pr\left(Y < \frac{3}{X}\right) = \int_{x=1}^{2} \int_{y=0}^{\frac{3}{x}} f_{XY}\left(x, y\right)\, dxdy = 0.857$$

באופן דומה:

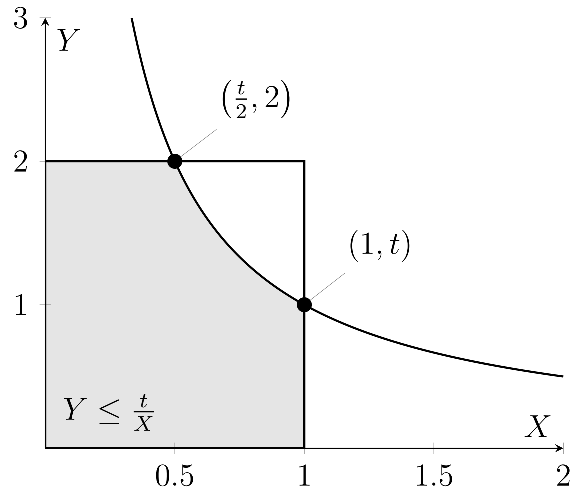
$$Pr\left(Y > X - \frac{1}{2}\right) = \int_{x=1}^{2} \int_{y=x-0.5}^{3} f_{XY}\left(x, y\right)\, dxdy = 0.52$$

□

תרגיל 7.9. נסמן ב-$S = XY$ את שטח המלבן. מאחר ו-$X \in [0, 1]$ ו-$Y \in [0, 2]$ הרי שבהכרח $S \in [0, 2]$. נחשב את פונקציית ההתפלגות המצטברת של S עבור $t \in [0, 2]$:

$$F_S\left(t\right) = Pr\left(S \leq t\right) = Pr\left(Y \leq \frac{t}{X}\right)$$

והתחום המתאים הינו



על מנת לבצע אינטגרל אחד במקום שניים נשתמש במשלים ונחשב את התחום הנותר מהמלבן:

$$\begin{aligned}F_S(t) &= 1-\int_{x=\frac{t}{2}}^{1}\int_{y=\frac{t}{x}}^{2}\frac{1}{2}\,dxdy = 1-\frac{1}{2}\int_{x=\frac{t}{2}}^{1}2-\frac{t}{x}\,dx\\ &= 1-\frac{1}{2}\,2x-t\ln x\big|_{x=\frac{t}{2}}^{1} = \frac{t}{2}\left(1-\ln\frac{t}{2}\right)\end{aligned}$$

לכן פונקציית ההתפלגות המצטברת (ודאו שהפונקציה אכן מתאימה - עולה ורציפה) הינה

$$F_S(t)=\begin{cases}0 & t<0\\ \frac{t}{2}\left(1-\ln\frac{t}{2}\right) & t\in[0,2]\\ 1 & t>2\end{cases}$$

נעיר שאמנם ייתכן מצב בו $X=0$ אך מדובר על משתנה מקרי רציף ומאורע זה עם הסתברות אפס, כך שאין בעיה לחלק ב-X. □

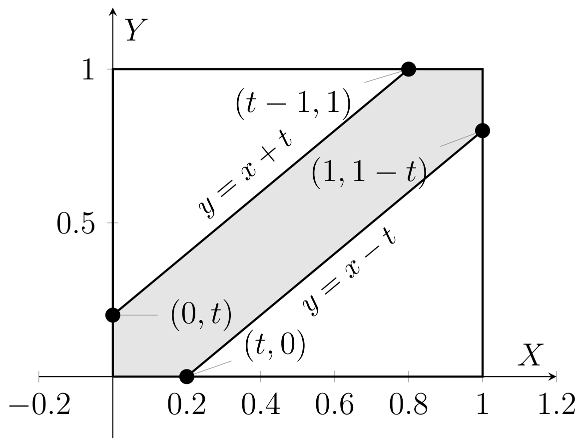
תרגיל 7.10. בהתאם למוגדר בשאלה, $X,Y\sim U(0,1)$ ו-$Z=|X-Y|$. המשתנים המקריים בלתי תלויים ולכן פונקציית הצפיפות המשותפת שווה למכפלת הצפיפויות השוליות:

$$f_{XY}(x,y)=\begin{cases}1 & 0\le x,y\le 1\\ 0 & \text{otherwise}\end{cases}$$

בבירור, $Z\ge 0$ ולכן $F_Z(t)=0$ לכל $t<0$. כמו כן, המרחק המירבי בין X ל-Y הוא 1 לכן $F_Z(t)=1$ כאשר $t>1$. נותר לחשב את פונקציית ההתפלגות המצטברת בתחום $t\in[0,1]$:

$$F_Z(t)=Pr(Z\le t)=Pr(|X-Y|\le t)=Pr(-t\le X-Y\le t)$$

תחום האינטגרציה הוא התחום $\{-t\le X-Y\le t\}$. נעביר אגף לקבלת צורה נוחה יותר (Y כתלות ב-X) ונקבל שהתחום הינו $\{X-t\le Y\le X+t\}$, כלומר זהו התחום הנמצא בין שני הישרים $y=x+t$ ו-$y=x-t$:



תחום האינטגרציה שהתקבל מורכב יחסית ויותר פשוט לבצע את החישוב באמצעות המשלים – שני המשולשים. מאחר והצפיפות אחידה אין צורך בחישוב האינטגרל וניתן להסתפק בחישוב השטח, כך שהסתברות תהא:

$$Pr\left(-t \leq X-Y \leq t\right) = 1 - 2 \cdot \frac{(1-t)^2}{2} = 1-(1-t)^2$$

לסיכום, פונקציית ההתפלגות המצטברת הינה:

$$F_Z(t) = \begin{cases} 0 & t<0 \\ 1-(1-t)^2 & t\in[0,1] \\ 1 & t>1 \end{cases}$$

לצורך חישוב התוחלת נחשב ראשית את הצפיפות באמצעות גזירה:

$$f_Z(t) = \frac{dF_Z\left(t\right)}{dt} = \begin{cases} 2(1-t) & t\in[0,1] \\ 0 & t\notin[0,1] \end{cases}$$

ולכן התוחלת הינה

$$\mathbb{E}(Z) = \int_0^1 t\cdot 2(1-t)\,dt = \frac{1}{3}$$

□

תרגיל 7.11. בהתאם למוגדר בשאלה, $X,Y \sim N\left(0,1\right)$ ובלתי תלויים, לכן

$$f_{XY}\left(x,y\right) = f_X\left(x\right) f_Y\left(y\right) = \frac{1}{2\pi}\,\mathbf{exp}\left(-\frac{x^2+y^2}{2}\right)$$

נחשב את פונקציית ההתפלגות המצטברת של המרחק מנקודת הפגיעה, $R=\sqrt{X^2+Y^2}$. בבירור, $F_R\left(t\right)=0$ לכל $t<0$ (המרחק לא יכול להיות שלילי) ולכן נתמקד רק בתחום החיובי:

$$F_R\left(t\right) = Pr\left(R\leq t\right) = Pr\left(\sqrt{X^2+Y^2}\leq t\right) = Pr\left(X^2+Y^2\leq t^2\right)$$

המאורע $\{X^2+Y^2\le t^2\}$ הינו מעגל במישור XY ברדיוס t לכן נוח לעבור לקואורדינטות קוטביות לצורך ביצוע האינטגרל. נעבור לקואורדינטות קוטביות (נוסחא **0.37**) ונקבל:

$$F_R(t)=\int_{r=0}^{t}\int_{\theta=0}^{2\pi}\frac{1}{2\pi}\exp\left(-\frac{r^2}{2}\right)rdrd\theta=\int_0^t r\exp\left(-\frac{r^2}{2}\right)dr=1-e^{-\frac{r^2}{2}}$$

לסיכום, פונקציית ההתפלגות המצטברת של המרחק מנקודת הפגיעה הינה:

$$F_R(t)=\begin{cases}1-e^{-\frac{r^2}{2}} & r\ge 0\\ 0 & r<0\end{cases}$$

והצפיפות היא נגזרתה:

$$f_R(t)=\begin{cases}re^{-\frac{r^2}{2}} & r\ge 0\\ 0 & r<0\end{cases}$$

$\square$

תרגיל 7.12. נסמן את גובה הגבר ב-$X\sim N(175,10^2)$ ואת גובה האישה ב-$Y\sim N(165,9^2)$.

1. גובהו של בן מתפלג לפי $B=\frac{3}{4}X+\frac{1}{4}Y$. לפי משפט **6.13**, $\frac{3}{4}X\sim N\left(\frac{3}{4}\cdot 175,\left(\frac{3}{4}\right)^2\cdot 10^2\right)$ וכן $\frac{1}{4}Y\sim N\left(\frac{1}{4}\cdot 165,\left(\frac{1}{4}\right)^2\cdot 9^2\right)$. סכום משתנים מקריים נורמליים הוא גם כן נורמלי ולכן ההתפלגות של B הינה

$$B\sim N\left(\frac{3}{4}\cdot 175+\frac{1}{4}\cdot 165,\left(\frac{1}{4}\right)^2\cdot 9^2+\left(\frac{3}{4}\right)^2\cdot 10^2\right)=N\left(172.5,7.83^2\right)$$

שימו לב שקיבלנו סטיית תקן קטנה מזו של כל אחת מההתפלגויות. הסיבה היא שיש סיכוי נמוך לקבל תוצאה חריגה לאותו הכיוון של שני ההורים (שניהם גבוהים מאוד או שניהם נמוכים מאוד) ולכן גם אם אחד ההורים חריג, ההורה השני יהיה לרוב בגובה ממוצע וגובה הבן (הממוצע המשוקלל) יהיה קרוב לתוחלת. אנו נראה זאת גם במשפט הגבול המרכזי - כאשר עושים ממוצע של משתנים מקריים, הפיזור שלהם מסביב לתוחלת קטן.

2. הפעם לא ידוע אם הילד הוא בן או בת. אם הוא בת, בדומה לסעיף הקודם נקבל $G=\frac{1}{4}X+\frac{3}{4}Y$ ולכן

$$G\sim N\left(\frac{1}{4}\cdot 175+\frac{3}{4}\cdot 165,\left(\frac{3}{4}\right)^2\cdot 9^2+\left(\frac{1}{4}\right)^2\cdot 10^2\right)=N\left(167.5,7.198^2\right)$$

נסמן את הגובה של הילד ב-T. שימו לב שלא נכון להגיד $T=\frac{1}{2}B+\frac{1}{2}G$ כי T איננו ממוצע של שני המשתנים המקריים אלא הוא או האחד או השני בסיכויים שווים. לכן נחשב את ההתפלגות שלו על ידי חישוב פונקציית ההתפלגות המצטברת לפי נוסחת ההסתברות השלמה. לכל $t\in\mathbb{R}$ מתקיים:

$$\begin{aligned}F_T(t) &= \frac{1}{2}F_B(t)+\frac{1}{2}F_G(t)=\frac{1}{2}Pr(B\le t)+\frac{1}{2}Pr(G\le t)\\ &=\frac{1}{2}\left(Pr\left(\frac{B-172.5}{7.83}\le\frac{t-172.5}{7.83}\right)+Pr\left(\frac{G-167.5}{7.198}\le\frac{t-167.5}{7.198}\right)\right)\\ &=\frac{1}{2}\left(\Phi\left(\frac{t-172.5}{7.83}\right)+\Phi\left(\frac{t-167.5}{7.198}\right)\right)\end{aligned}$$

□

תרגיל 7.13. **1.** נניח שהערך של R_i ידוע. בהינתן ערך זה, בתיבה יש n כובעים, מתוכם R_i של הנבחרת ה-i. בנוסף, ישנם R_i אוהדים לכן נבחרים מהתיבה על ידי אוהדי הקבוצה ה-i בדיוק R_i פריטים. תיאור זה מתאים להתפלגות היפרגאומטרית - אוכלוסיית כובעים בגודל n, מתוכה R_i כובעים מיוחדים ו-R_i קובעים שנשלפים. לכן מספר הכובעים של הנבחרת ה-i שנשלפו על ידי אוהדיה מתפלג לפי $X_i|R_i\sim HG(n,R_i,R_i)$. נציב בנוסחא לתוחלת של התפלגות היפרגאומטרית (נוסחא 6.22) ונקבל $\mathbb{E}(X_i|R_i)=\frac{R_i^2}{n}$.

2. לכל אוהד הסתברות זהה השווה ל-$\frac{1}{m}$ לאהוד נבחרת מסוימת ולכן מספר אוהדיה של כל נבחרת מתפלג $R_i\sim Bin\left(n,\frac{1}{m}\right)$. נשתמש בנוסחא לתוחלת ושונות של משתנה מקרי בינומי ונקבל

$$\mathbb{E}\left(R_i^2\right)=\mathbb{V}(R_i)+\mathbb{E}^2(R_i)=\frac{n(n+m-1)}{m^2}$$

3. לפי משפט התוחלת השלמה:

$$\mathbb{E}(X_i)=\mathbb{E}(\mathbb{E}(X_i|R_i))=\mathbb{E}\left(\frac{R_i^2}{n}\right)=\frac{n+m-1}{m^2}$$

ולכן לפי לינאריות התוחלת:

$$\mathbb{E}(X)=\mathbb{E}\left(\sum_{i=1}^{m}X_i\right)=\sum_{i=1}^{m}\mathbb{E}(X_i)=\frac{n+m-1}{m}$$

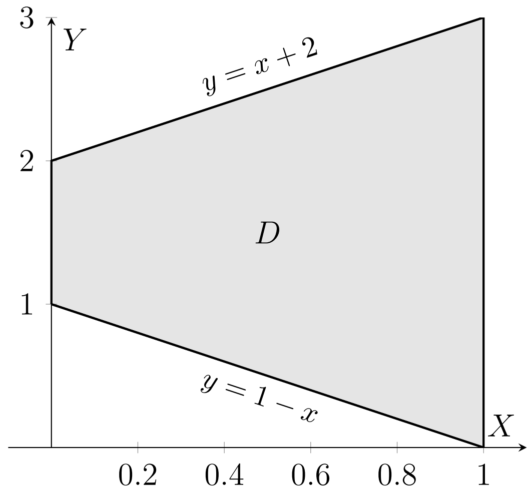
□

תרגיל 7.14. **בבעיה 7.3 מצאנו כי** $X\sim Bin\left(8,\frac{1}{2}\right)$ **ו-**$Y|X\sim Bin\left(X,\frac{1}{4}\right)$. **נחשב את התוחלת של Y באמצעות משפט התוחלת השלמה תוך שימוש בתוחלת של משתנה מקרי בינומי (נוסחא 6.6):**

$$\begin{aligned}\mathbb{E}(Y)&=\mathbb{E}(\mathbb{E}(Y|X))=\mathbb{E}\left(\frac{1}{4}X\right)=\frac{1}{4}\mathbb{E}(X)\\ &=\frac{1}{4}\cdot 8\cdot\frac{1}{2}=1\end{aligned}$$

וזו בדיוק אותה התוצאה אותה קיבלנו בבעיה 7.3 אך בצורה הרבה יותר פשוטה. □

תרגיל 7.15. נתחיל משרטוט תחום ההגדרה של המשתנים המקריים:



שטח הטרפז הינו $S_D = \frac{1+3}{2} = 2$ ולכן הצפיפות המשותפת הינה

$$f_{XY}(x,y) = \begin{cases} \frac{1}{2} & (x,y) \in D \\ 0 & (x,y) \notin D \end{cases}$$

ניתן לחשב את התוחלת של Y באופן ישיר על ידי חישוב הצפיפות השולית של Y ולאחר מכן חישוב האינטגרל המתאים, אך במקרה זה חישוב דרך התוחלת השלמה יהיה פשוט יותר. הסיבה היא שהצפיפות השולית של Y מורכבת וכוללת שלושה אינטגרלים לחישוב: כאשר $y \in [0,1]$ תחום הערכים המתאים ל-x הינו $[1-y,1]$ ואלו יהיו גבולות האינטגרל בחישוב הצפיפות השולית. בדומה, כאשר $y \in [1,2]$ גבולות האינטגרל יהיו $[0,1]$ וכאשר $y \in [2,3]$ גבולות האינטגרל יהיו $[y-2,1]$. לפיכך, הצפיפות השולית תכלול שלוש צורות פונקציונליות שונות ואת האינטגרל המתאר את התוחלת נצטרך לחשב בשלושה תחומים שונים. התשובה הסופית תהיה זהה, אך פתרון דרך משפט התוחלת השלמה יהיה פשוט יותר. לשם כך עלינו למצוא את הצפיפות השולית של X ואת הצפיפות המותנית $Y|X$.

יהי $x \in [0,1]$. הצפיפות השולית של X בנקודה זו תהיה

$$f_X(x) = \int_{y=1-x}^{x+2} 0.5\, dy = \frac{(x+2)-(1-x)}{2} = x + \frac{1}{2}$$

ואפס אחרת. לכן הצפיפות המשותפת כאשר $x \in [0,1]$ הינה

$$f_{Y|X=x}(y) = \frac{f_{XY}(x,y)}{f_X(x)} = \begin{cases} \frac{\frac{1}{2}}{x+\frac{1}{2}} & y \in [1-x, x+2] \\ 0 & y \notin [1-x, x+2] \end{cases}$$

הצפיפות המותנית שהתקבלה היא פונקציה קבועה של y כך שזהו משתנה מקרי אחיד: $Y|X \sim U(1-X, X+2)$. לכן הצפיפות המותנית הינה

$$\mathbb{E}(Y|X) = \frac{X+2+1-X}{2} = \frac{3}{2}$$

התוחלת המותנית קבועה ולכן זוהי גם התוחלת הלא מותנית:

$$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}\left(\frac{3}{2}\right) = \frac{3}{2}$$

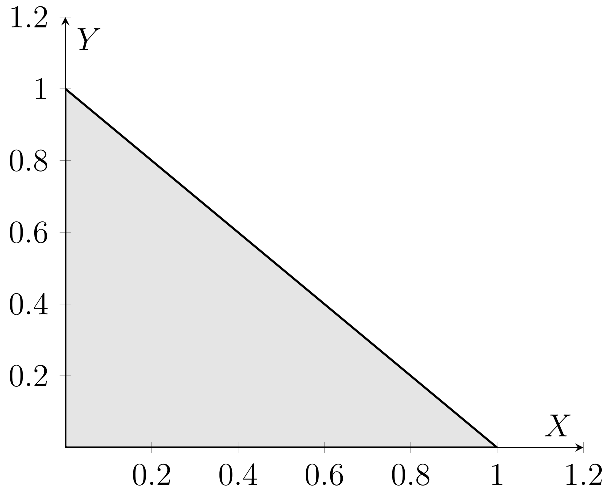
שימו לב שיכולנו לצפות לתוצאה זו. היות והטרפז סימטרי סביב $Y = \frac{3}{2}$ והצפיפות אחידה, לכל נקודה $\left(x, \frac{3}{2} + \Delta y\right) \in D$ יש נקודה מקבילה לה $\left(x, \frac{3}{2} - \Delta y\right) \in D$ והממוצע של כל ערכי ה-Y יהיה בדיוק $\frac{3}{2}$. □

תרגיל 7.16. בתרגיל 7.15 ראינו שנוח לטפל ב-Y באמצעות התניה על X מאחר והתפלגות זו היא התפלגות אחידה. לכן נחשב גם את תוחלת המכפלה באמצעות משפט התוחלת השלמה והתניה על X:

$$\begin{aligned}\mathbb{E}(XY) &= \mathbb{E}(\mathbb{E}(XY|X)) \overset{eq.7.13}{=} \mathbb{E}(X\mathbb{E}(Y|X)) = \mathbb{E}\left(X \cdot \frac{3}{2}\right) \\ &= \frac{3}{2}\mathbb{E}(X) = \mathbb{E}(X)\,\mathbb{E}(Y)\end{aligned}$$

שכן ראינו בתרגיל 7.15 ש-$\mathbb{E}(Y) = \mathbb{E}(Y|X) = \frac{3}{2}$. כלומר תוחלת המכפלה שווה למכפלת התוחלות ולכן המשתנים המקריים בלתי מתואמים. תוצאה זו הגיונית לאור העובדה שהתוחלת המותנית של Y קבועה. כאשר X מקבל ערכים גבוהים מהממוצע שלו, Y יהיה בממוצע $\frac{3}{2}$ וכך גם כאשר X מקבל ערכים קטנים מהממוצע שלו. כלומר, הממוצע של Y נשאר קבוע ללא תלות בערך של X וזו בדיוק המשמעות של חוסר תיאום. □

תרגיל 7.17. נשרטט את תחום ההגדרה:



מבלי לבצע חישובים נוספים ניתן להסיק שמקדם המתאם יהיה שלילי. ככל ש-X מקבל ערכים גבוהים יותר, כך תחום הערכים ש-Y יכול לקבל קטן יותר ומרוכז יותר בערכים הנמוכים שלו, כך שהקשר ביניהם הוא הפוך – כשאחד גדל מהממוצע השני בהכרח יהיה קטן מהממוצע. נוודא זאת על ידי חישוב מפורש.

הצפיפות השולית של X בתחום $x \in [0, 1]$ הינה

$$f_X(x) = \int_{y=0}^{1-x} 6(1-x-y)\,dy = 3(x-1)^2$$

ו-0 אחרת. לכן התוחלת הינה

$$\mathbb{E}(X) = \int_0^1 3x(x-1)^2\,dx = \frac{1}{4}$$

והתוחלת של X^2 הינה

$$\mathbb{E}\left(X^2\right) = \int_0^1 3x^2(x-1)^2\,dx = \frac{1}{10}$$

כך שהשונות הינה

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X) = \frac{1}{10} - \left(\frac{1}{4}\right)^2 = \frac{3}{80}$$

נשים לב שהמשתנים המקריים סימטריים ולכן אלו גם התוחלת והשונות של Y. לצורך חישוב מקדם המתאם נותר לחשב את התוחלת של מכפלתם:

$$\mathbb{E}(XY) = \int_{x=0}^1 \int_{y=0}^{1-x} 6xy(1-x-y)\,dydx = \frac{1}{20}$$

ומקדם המתאם הינו

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}} = \frac{\frac{1}{20} - \left(\frac{1}{4}\right)^2}{\sqrt{\left(\frac{3}{80}\right)^2}} = -\frac{1}{3} < 0$$

כצפוי. □

תרגיל 7.18. ראינו בתרגיל 7.6 כי ההתפלגות השולית של X,Y היא אחידה ולכן $\mathbb{E}(X) = \mathbb{E}(Y) = 3$ וכן $\mathbb{V}(X) = \mathbb{V}(Y) = 2$.

1. נחשב את תוחלת המכפלה בצורה ישירה. לשם כך נרשום את לוח הכפל בין X ל-Y ונראה מהן כל התוצאות האפשריות:

X/Y	1	2	3	4	5
1		2	3	4	5
2	2		6	8	10
3	3	6		12	15
4	4	8	12		20
5	5	10	15	20	

התאים על האלכסון ריקים מאחר והם מתארים מאורעות בלתי אפשריים (לא ייתכן ששני הקלפים עם אותו הערך). כל תוצאת מכפלה אפשרית מופיעה בשני תאים בדיוק ומאחר והסיכוי לכל קומבינציה של X ו-Y הוא $\frac{1}{20}$ הרי שהסיכוי לכל אחת מהתוצאות האפשריות הוא $\frac{1}{10}$ והתוחלת של המכפלה היא הממוצע שלהן:

$$\mathbb{E}(XY) = \frac{2+3+4+5+6+8+10+12+15+20}{10} = 8.5$$

לפיכך השונות המשותפת הינה $Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = -0.5$. כצפוי, המתאם

הוא שלילי – כאשר הקלף העליון מקבל ערך גבוה יחסית, הקלף התחתון לא יכול לקבל ערך זה ולכן בהכרח יקבל ערכים יחסית נמוכים (ולהפך). מספרית, מקדם המתאם הינו

$$\rho_{XY}=\frac{Cov\left(X,Y\right)}{\sqrt{\mathbb{V}\left(X\right)\mathbb{V}\left(Y\right)}}=\frac{-\frac{1}{2}}{2}=-\frac{1}{4}$$

2. ההתפלגות של הקלף השני זהה להתפלגות הקלף התחתון (אין ביניהם הבדל) וכך גם ההתפלגות המשותפת של הקלף העליון והתחתון והקלף העליון והשני. לכן מקדם המתאם בין X ל-Z שווה למקדם המתאם בין X לבין $X+Y$. ניעזר בתוצאות של הסעיף הקודם בכדי לחשב אותו:

$$Cov\left(X,Z\right)=Cov\left(X,X+Y\right)=Cov\left(X,X\right)+Cov\left(X,Y\right)=2-0.5=1.5$$

השונות של סכום הקלפים הינה

$$\mathbb{V}\left(Z\right)=\mathbb{V}\left(X+Y\right)=\mathbb{V}\left(X\right)+\mathbb{V}\left(Y\right)+2Cov\left(X,Y\right)=4-1=3$$

ומכאן שמקדם המתאם הינו

$$\rho_{X,Z}=\rho_{X,X+Y}=\frac{1.5}{3}=\frac{1}{2}$$

כצפוי מקדם המתאם חיובי. אם X קיבל ערך גבוה סביר שהסכום שלו עם עוד קלף יהיה גם גבוה ואם X קיבל ערך נמוך סביר שהסכום שלו עם עוד קלף יהיה גם כן נמוך, כך שהמשתנים המקריים גדלים וקטנים ביחד.

□

תרגיל 7.19

1. מסימטריה, שלושת המשתנים המקריים הם שווי התפלגות. עבור צבע מסויים, ההסתברות שצבע זה ייצא בהוצאה מסויימת הינה $\frac{1}{3}$ ולכן כל אחד מתפלג $Bin\left(4,\frac{1}{3}\right)$. לפי שונות של משתנה מקרי בינומי, $\mathbb{V}\left(X\right)=\mathbb{V}\left(Y\right)=\mathbb{V}\left(Z\right)=4\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{8}{9}$.
2. המשתנה המקרי המתואר על ידי הסכום $X+Y$ הוא מספר האדומים והשחורים שהוצאו. מאחר וההסתברות להוציא אדום או שחור היא $\frac{2}{3}$ הרי ש - $X+Y\sim Bin\left(4,\frac{2}{3}\right)$. לפיכך $\mathbb{V}\left(X+Y\right)=4\cdot\frac{2}{3}\cdot\frac{1}{3}=\frac{8}{9}$.
3. נשתמש בנוסחה 7.23 ונקבל

$$Cov\left(X,Y\right)=\frac{\mathbb{V}\left(X+Y\right)-\mathbb{V}\left(X\right)-\mathbb{V}\left(Y\right)}{2}=-\frac{4}{9}$$

לפיכך: $\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}=\frac{-4/9}{8/9}=-\frac{1}{2}$ וזהו ערך שלילי, כצפוי. ככל שיצאו יותר כדורים מצבע מסויים ייצאו פחות כדורים מצבעים אחרים.

4. מתקיים $\left(X+Y\right)+Z=4$ לכן בין $X+Y$ לבין Z יש קשר לינארי עם שיפוע שלילי ובהתאם $\rho\left(X+Y,Z\right)=-1$.

□

תרגיל 7.20. נניח כי $\rho = \pm 1$ ונוכיח כי יש קשר לינארי בין המשתנים המקריים. לצורך כך מספיק להוכיח שקיים a עבורו $Y - aX$ קבוע. היות והשונות של קבוע היא 0 הרי שנחפש a כך ש-$\mathbb{V}(Y - aX) = 0$. לצורך כך נגדיר את הפונקציה

$$g(a) = \mathbb{V}(Y - aX)$$

מאחר והשונות היא אי שלילית הרי שהנקודה בה הפונקציה מתאפסת היא מינימום שלה. נפתח את השונות לפי נוסחא 7.23 ונמצא את נקודת המינימום.

$$g(a) = \mathbb{V}(Y) + a^2\mathbb{V}(X) - 2aCov(X,Y)$$

הפונקציה $g(a)$ היא פרבולה והמינימום שלה מתקבל בנקודה $a' = \frac{Cov(X,Y)}{\mathbb{V}(X)} = \rho_{XY}\sqrt{\frac{\mathbb{V}(Y)}{\mathbb{V}(X)}}$. נציב לחישוב ערך הפונקציה (שימו לב ש-$\rho_{XY}^2 = 1$):

$$g\left(\rho_{XY}\sqrt{\frac{\mathbb{V}(Y)}{\mathbb{V}(X)}}\right) = \mathbb{V}(Y) + \frac{\mathbb{V}(Y)}{\mathbb{V}(X)}\cdot\mathbb{V}(X) - 2\rho_{XY}\sqrt{\frac{\mathbb{V}(Y)}{\mathbb{V}(X)}}\cdot\rho_{XY}\cdot\sqrt{\mathbb{V}(X)\mathbb{V}(Y)} = 0$$

כלומר קיים מספר ממשי a' עבורו השונות היא אפס כלומר ההפרש $Y - a'X$ קבוע ויש בין המשתנים המקריים קשר לינארי. בנוסף, לפי הנוסחא שקיבלנו עבור a' מתקיים שסימנו שווה לסימן של ρ_{XY}. □

תרגיל 7.21. נמשיך לפי הרעיון שהוצג בחישוב התוחלת במשפט 6.7. האינדיקטור $\mathbf{1}_i$ מסמן את העובדה שהפריט ה-i שיצא מיוחד. כפי שראינו שם $X = \sum_{i=1}^{n}\mathbf{1}_i$ וכן $Pr(\mathbf{1}_i = 1) = \frac{D}{N}$. לכן $\mathbb{E}(\mathbf{1}_i) = \frac{D}{N}$ ו-$\mathbb{V}(\mathbf{1}_i) = \frac{D}{N}\left(1 - \frac{D}{N}\right)$ (נוסחא 5.17). לצורך חישוב השונות של X באמצעות נוסחא 7.24 נותר לחשב את $Cov(\mathbf{1}_i, \mathbf{1}_j)$ עבור $i < j$. לצורך כך נחשב את התוחלת של המכפלה $\mathbf{1}_i\mathbf{1}_j$. זוהי מכפלה של שני אינדיקטורים השווה ל-1 או ל-0. ערכה 1 רק כאשר שני הפריטים מיוחדים וההסתברות לכך הינה

$$\mathbb{E}(\mathbf{1}_i\mathbf{1}_j) = Pr(\mathbf{1}_i\mathbf{1}_j = 1) = \frac{D}{N}\cdot\frac{D-1}{N-1}$$

לכן השונות המשותפת היא

$$\begin{aligned} Cov(\mathbf{1}_i, \mathbf{1}_j) &= \mathbb{E}(\mathbf{1}_i\mathbf{1}_j) - \mathbb{E}(\mathbf{1}_i)\mathbb{E}(\mathbf{1}_j) = \frac{D}{N}\cdot\frac{D-1}{N-1} - \left(\frac{D}{N}\right)^2 \\ &= \frac{D}{N}\left(\frac{D-1}{N-1} - \frac{D}{N}\right) = -\frac{D}{N}\left(1 - \frac{D}{N}\right)\frac{1}{N-1} \end{aligned}$$

כצפוי, השונות המשותפת שלילית. כאשר אחד האינדיקטורים שווה ל-1 משמעות הדבר שנבחר פריט מיוחד אחד ולכן הסיכוי שאינדיקטור אחר יהיה 1 קטן יותר. ישנם $\binom{n}{2}$ זוגות אפשריים של אינדקסים שונים (בכל זוג כזה נבחר את i להיות הקטן מבין המספרים שנבחרו) ולכן השונות הכוללת הינה

$$\begin{aligned}
\mathbb{V}\left(X\right) = \mathbb{V}\left(\sum_{i=1}^{n}\mathbf{1}_i\right) &= \sum_{i=1}^{n}\mathbb{V}\left(\mathbf{1}_i\right) + 2\sum_{i<j}Cov\left(\mathbf{1}_i,\mathbf{1}_j\right)\\
&= n\cdot\frac{D}{N}\left(1-\frac{D}{N}\right) - 2\binom{n}{2}\frac{D}{N}\left(1-\frac{D}{N}\right)\frac{1}{N-1}\\
&= \frac{nD}{N}\left(1-\frac{D}{N}\right)\left(\frac{N-n}{N-1}\right)
\end{aligned}$$

□

תרגיל 7.22. **1.** כל אחד מהמשתנים המקריים מתבונן בסדרה של n ניסויים בלתי תלויים וסופר את מספר ההצלחות (1 או 6, בהתאם למשתנה המקרי). לכן לשניהם התפלגות $Bin\left(n,\frac{1}{6}\right)$ ומתקיים

$$\mathbb{V}\left(X\right) = \mathbb{V}\left(Y\right) = n\frac{1}{6}\left(1-\frac{1}{6}\right) = \frac{5n}{36}$$

2. ניתן לתאר כל אחד מהמשתנים המקריים באמצעות סכום האינדיקטורים המתאימים:

$$XY = \sum_{i=1}^{n}\mathbf{1}_i^X \qquad = \sum_{i=1}^{n}\mathbf{1}_i^Y$$

3. הטלות שונות הן בלתי תלויות ולכן האינדיקטורים המתאימים בלתי מתואמים והשונות המשותפת שלהם שווה ל-0. כמו כן, שני האינדיקטורים $\mathbf{1}_i^X,\mathbf{1}_i^Y$ מביטים באותה ההטלה. בוודאות אחד מהם יהיה שווה ל-0, שכן לא ייתכן שבאותה ההטלה ייצא גם 1 וגם 6, כך שגם מכפלתם קבועה ושווה ל-0 וזוהי גם התוחלת $\mathbb{E}\left(\mathbf{1}_i^X\mathbf{1}_i^Y\right) = 0$.

4. לפי ההגדרה, $Cov\left(\mathbf{1}_i^X,\mathbf{1}_i^Y\right) = \mathbb{E}\left(\mathbf{1}_i^X\mathbf{1}_i^Y\right) - \mathbb{E}\left(\mathbf{1}_i^X\right)\mathbb{E}\left(\mathbf{1}_i^Y\right) = -\frac{1}{36}$.

5. נחשב את השונות המשותפת באמצעות תכונת הבי-לינאריות:

$$\begin{aligned}
Cov\left(X,Y\right) = Cov\left(\sum_{i=1}^{n}\mathbf{1}_i^X,\sum_{i=1}^{n}\mathbf{1}_i^Y\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n}Cov\left(\mathbf{1}_i^X,\mathbf{1}_j^Y\right)\\
&= \sum_{i=1}^{n}Cov\left(\mathbf{1}_i^X,\mathbf{1}_i^Y\right) = -\frac{n}{36}
\end{aligned}$$

ומקדם המתאם הינו

$$\rho_{XY} = \frac{Cov\left(X,Y\right)}{\sqrt{\mathbb{V}\left(X\right)\mathbb{V}\left(Y\right)}} = \frac{-n/36}{5n/36} = -\frac{1}{5}$$

כצפוי, מקדם המתאם שלילי שהרי ככל שהיו יותר הטלות השוות ל-1 יש בהכרח פחות הטלות השוות ל-6.

□

תרגיל 7.23. הניסוי המתואר הוא בחירה של שני פריטים ללא החזרה מתוך תשעה פריטים. כל אחד מהמשתנים המקריים סופר כמה מתוך הפריטים שנבחרו הוא פריט "מיוחד" ולכן ההתפלגות השולית של כל משתנה מקרי היא התפלגות היפרגאומטרית: $X \sim HG\left(9,3,2\right)$ ו-$Y \sim HG\left(9,2,2\right)$. המאורע $\{X=k,Y=l\}$ מתרחש כאשר נבחרים k נקניקים מתוך ה-3, l פרוסות פסטרמה מתוך ה-2. להשלמת המאורע יש לבחור גם $2-k-l$ סטייקים מתוך ה-4. ההסתברות של מאורע זה ולמעשה ההתפלגות המשותפת היא

$$P_{X,Y}\left(k,l\right)=\frac{\binom{3}{k}\binom{2}{l}\binom{4}{2-k-l}}{\binom{9}{2}}$$

כאשר $k,l,k+l\in\{0,1,2\}$ ו-0 אחרת. □

תרגיל 7.24. במבנה כולו יש $1+2+\ldots+n\overset{eq.0.20}{=}\frac{n(n+1)}{2}$ כדורים.

1. בכל שורה $i\geq 3$ יש 3 כדורים בלבד בשלושת המקומות הראשונים ולכן בסך הכל יש $1+2+3+\ldots+3=3\left(n-1\right)$ כדורים במקומות אלו. מתוכם 6 כדורים נמצאים בשלושת השורות הראשונות ולכל הכדורים אותו סיכוי להיבחר, לכן הסיכוי שהכדור הנבחר הוא משלושת השורות הראשונות בהינתן שהוא נבחר מאחד משלושת המקומות הראשונים הינו $\frac{6}{3(n-1)}=\frac{2}{n-1}$.
2. לכל כדור סיכוי שווה להיבחר ולכן ההתפלגות המשותפת הינה

$$P_{X,Y}\left(k,l\right)=\begin{cases}\frac{2}{n(n+1)} & 1\leq l\leq k\leq n\\ 0 & \text{otherwise}\end{cases}$$

שכן בהכרח המיקום בשורה בה עומד הכדור צריך להיות קטן או שווה למספר השורה.

3. בשורה ה-k יש k כדורים ולכן ההתפלגות השולית של X הינה

$$P_X\left(k\right)=\frac{2k}{n\left(n+1\right)}$$

כאשר $k\in supp\left(X\right)=\{1,\ldots,n\}$. יש n כדורים שעומדים במקום הראשון בשורה (אחד בכל שורה), $n-1$ כדורים שעומדים במקום השני בשורה ובאופן כללי, $n-k+1$ כדורים במקום ה-k בכל שורה. לכן

$$P_Y\left(k\right)=\frac{2\left(n-k+1\right)}{n\left(n+1\right)}$$

כאשר $k\in supp\left(Y\right)=\{1,\ldots,n\}$.

4. המשתנים המקריים בבירור תלויים. ניתן לראות זאת על ידי הכפלת התפלגויות השוליות או על ידי חישוב ההתפלגות המותנית - אם ידועה השורה ממנה נבחרו הכדורים, אז לכל כדור בשורה יש סיכוי שווה להיבחר אבל אם לא ידועה השורה יש סיכוי גדול יותר לבחור את הכדור שעומד במקום הראשון בשורה.

□

תרגיל 7.25. נשים לב ש-X סופר את מספר העצים שהתקבלו בניסוי ולכן $X\sim Bin\left(2,\frac{1}{2}\right)$. נבדוק האם המשתנים המקריים תלויים על ידי בחינת מקרה בו ההתפלגות השולית שונה מההתפלגות המותנית. אם $Y=0$ בהכרח בשתי ההטלות יצאה אותה התוצאה ולכן $X=0$ או $X=2$, כלומר $Pr\left(X=1|Y=0\right)=0$. יחד עם זאת, $Pr\left(X=1\right)=\frac{1}{2}$ ולכן הידיעה על הערך של Y משפיעה על ההתפלגות של X והם תלויים.

נשים לב שבהינתן X, הערך של Y קבוע:

$$Y\left(X\right)=\begin{cases}0 & X=0,2\\ 1 & X=1\end{cases}$$

לכן $Pr\left(Y=0\right)=Pr\left(Y=1\right)=\frac{1}{2}$, כלומר $\mathbb{E}\left(Y\right)=\frac{1}{2}$. לצורך חישוב השונות המשותפת נחשב את התוחלת של XY באמצעות תוחלת של פונקציה:

$$\mathbb{E}\left(YX\right)=\mathbb{E}\left(Y\left(X\right)X\right)=0\cdot Y\left(0\right)\cdot Pr\left(X=0\right)+1\cdot Y\left(1\right)\cdot Pr\left(X=1\right)+2\cdot Y\left(2\right)\cdot Pr\left(X=2\right)=\frac{1}{2}$$

נשתמש בתוחלת של משתנה מקרי בינומי ונקבל שמקדם המתאם הינו

$$Cov\left(X,Y\right)=\mathbb{E}\left(XY\right)-\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)=\frac{1}{2}-1\cdot\frac{1}{2}=0$$

כך שמשתנים מקריים אלו בלתי מתואמים. ניתן לראות זאת כבר מהתלות בין Y ל-X: לא משנה אם X גדל או קטן ביחס ל-1, הערך של Y הוא 2, כלומר הם לא גדלים ביחד ולא גדלים בכיוונים הפוכים. □

תרגיל 7.26. הטענה לא נכונה כאשר המשתנים המקריים תלויים. למשל, $X\sim U\left(-1,1\right)$ ו-$Y=-X$ הם משתנים מקריים רציפים ושווי התפלגות אך $X+Y=0$ ולכן סכומם אינו משתנה מקרי רציף. באופן דומה ניתן לבנות גם משתנים מקריים בדידים. למשל אם נגדיר

$$Y=\begin{cases}-X+3 & X>0\\ -X-2 & X<0\end{cases}$$

נקבל משתנה מקרי רציף גם כן והפעם סכומם הוא משתנה מקרי בדיד המקבל את הערכים $\{-2,3\}$ בהסתברות שווה. □

תרגיל 7.27. נסמן ב-$X\sim Exp\left(\lambda_1\right)$ את הזמן עד הגעת האוטובוס וב-$Y\sim Exp\left(\lambda_2\right)$ את הזמן עד הגעת המונית. הצפיפות המשותפת הינה מכפלת הצפיפויות השוליות מאחר ומשתנים מקריים אלו בלתי תלויים:

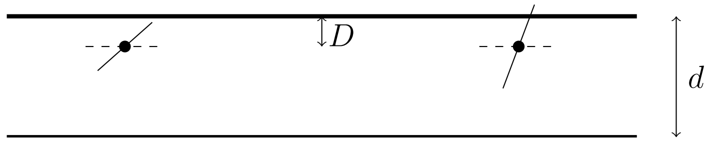
$$f_{XY}\left(x,y\right)=f_X\left(x\right)f_Y\left(y\right)=\begin{cases}\lambda_1\lambda_2e^{-\lambda_1x}e^{-\lambda_2y} & x,y>0\\ 0 & \text{otherwise}\end{cases}$$

הרכב הראשון הוא אוטובוס כאשר $X<Y$ כך שההסתברות המבוקשת הינה

$$Pr\left(X<Y\right)=\int_{x=0}^{\infty}\int_{y=x}^{\infty}\lambda_1\lambda_2e^{-\lambda_1x}e^{-\lambda_2y}=\frac{\lambda_1}{\lambda_1+\lambda_2}$$

כאשר האינטגרל חושב בתרגיל 0.7. □

תרגיל 7.28. נשרטט את הבעיה. המחט השמאלית לא חותכת את הקו ואילו הזווית של המחט הימנית מספיק גדולה כך שהיא כן חותכת את הקו:



נסמן ב- $D \sim U\left(0, \frac{d}{2}\right)$ את המרחק בין מרכז המחט לישר הקרוב ביותר וב- $\Theta \sim U\left(0, \frac{\pi}{2}\right)$ את הזווית של המחט ביחס לקו המקווקו שבשרטוט. מאחר ומשתנים מקריים אלו בלתי תלויים הרי שפונקציית הצפיפות המשותפת הינה

$$f_{\Theta,D}(x,y) = \begin{cases} \frac{4}{\pi d} & x \in \left[0, \frac{\pi}{2}\right], y \in \left[0, \frac{d}{2}\right] \\ 0 & \text{otherwise} \end{cases}$$

המחט חותכת את הישר כאשר ההיטל שלה במאונך לקווים גדול מהמרחק אל הקו, כלומר כאשר $D \geq \frac{l}{2}\sin(\Theta)$. בהתאם לכך ההסתברות הינה

$$Pr\left(\frac{l}{2}\sin(\Theta) \geq D\right) = \int_{x=0}^{\frac{\pi}{2}} \int_{y=0}^{\frac{l\sin x}{2}} \frac{4}{\pi d}\, dydx = \frac{2l}{\pi d}$$

ניסוי המחט של בופון מאפשרת את חישוב הערך של π באמצעות ניסוי פשוט: הטלת מקלות באקראי בין שני קווים וספירת מספר הפעמים שהמקל נוגע בקו ביחס למספר הזריקות. $\square$

תרגיל 7.29. 1. המשתנים המקריים נורמליים ובלתי תלויים לכן הפרשם גם כן נורמלי: $X - Y \sim N\left(0, \sigma_X^2 + \sigma_Y^2\right)$. משתנה מקרי נורמלי סימטרי סביב ה-0 ולכן

$$Pr(X > Y) = Pr(X - Y > 0) = \frac{1}{2}$$

2. נחשב ישירות. ראשית, נפשט על ידי תיקנון: $X_0 = \frac{X}{\sigma_X} \sim N(0,1)$, $Y_0 = \frac{Y}{\sigma_Y} \sim N(0,1)$. במונחים של המשתנים המקריים החדשים, ההסתברות המבוקשת היא $Pr(X_0 > sY_0 > 0)$ כאשר $s = \frac{\sigma_Y}{\sigma_X}$. הצפיפות המשותפת היא מכפלת הצפיפויות השוליות (חוסר תלות) ולכן:

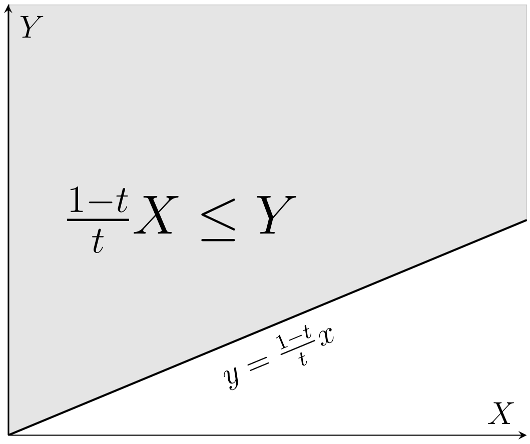
$$\begin{aligned} Pr(X_0 > sY_0 > 0) &= \int_{y=0}^{\infty} \int_{x=sy}^{\infty} \frac{1}{2\pi} \exp\left(-\frac{x^2+y^2}{2}\right) dxdy \\ (*) &= \int_{r=0}^{\infty} \int_{\theta=0}^{\arctan s^{-1}} \frac{1}{2\pi} r \exp\left(-\frac{r^2}{2}\right) drd\theta \\ &= \frac{\arctan\left(\frac{\sigma_X}{\sigma_Y}\right)}{2\pi} \end{aligned}$$

כאשר ב-(*) עברנו לקואורדינטות קוטביות (נוסחא 0.37). את האינטגרל לפי r ניתן לחשב בצורה פשוטה אם שמים לב שהפונקציה $r\exp\left(-\frac{r^2}{2}\right)$ היא הצפיפות שהתקבלה בתרגיל 7.11 ולכן מנרמול האינטגרל הוא 1.

$\square$

תרגיל 7.30. נחשב את פונקציית ההתפלגות המצטברת של Z. נשים לב ששני המשתנים המקריים X,Y חיוביים ולכן גם Z חיובי. בנוסף, $X \leq X+Y$ ולכן בהכרח $Z \leq 1$. לכן התחום המעניין בו יש לחשב את $F_Z(t)$ הוא התחום $t \in [0,1]$:

$$
\begin{aligned}
F_Z(t) &= Pr(Z \leq t) = Pr\left(\frac{X}{X+Y} \leq t\right) = Pr(X \leq t(X+Y)) \\
&= Pr\left(\frac{1-t}{t}X \leq Y\right) = \iint_{\frac{1-t}{t}X \leq Y} f_{XY}(x,y)\, dxdy
\end{aligned}
$$



הצפיפות המשותפת היא מכפלת הצפיפויות השוליות מאחר והמשתנים המקריים בלתי תלויים, כלומר $f_{XY}(x,y) = \begin{cases} e^{-x-y} & x,y \geq 0 \\ 0 & \text{otherwise} \end{cases}$. תחום האינטגרציה הוא החלק מהרביע הראשון הנמצא מעל הישר $Y = \frac{1-t}{t}X$ ולכן

$$
F_Z(t) = \int_{x=0}^{\infty}\int_{y=\frac{1-t}{t}x}^{\infty} e^{-x-y}\, dydx = t
$$

כלומר

$$
F_Z(t) = \begin{cases} 0 & t<0 \\ t & t \in [0,1] \\ 1 & t>1 \end{cases}
$$

פונקציית ההתפלגות המצטברת שהתקבלה היא פונקציית התפלגות של משתנה מקרי אחיד ולכן $Z \sim U(0,1)$.
כדאי לשים לב שאם לא היינו קובעים מראש $t \in [0,1]$ היה עלינו להיזהר בחלוקה ב-t (שהרי פעולה זו הופכת את סימן אי-השוויון) ובשרטוט תחום האינטגרציה (השיפוע יכל היה להיות שלילי). לכן מומלץ מאוד וחשוב לוודא מראש מהו תחום הערכים המתאים ולבצע את החישוב בהתאם. $\square$

תרגיל 7.31. לצורך התניה על X עלינו לחשב את הצפיפות השולית של X. לכל $x \in [0,1]$ הצפיפות הינה

$$
f_X(x) = \int_0^1 x+y\, dy = x + \frac{1}{2}
$$

ו-0 אחרת. לכן הצפיפות המותנית $Y|X=x$ כאשר $x \in [0,1]$ הינה

$$f_{Y|X=x}(y)=\frac{f_{XY}(x,y)}{f_X(x)}=\begin{cases}\frac{x+y}{x+0.5} & y\in[0,1]\\ 0 & y\notin[0,1]\end{cases}$$

והתוחלת המותנית הינה

$$\mathbb{E}(Y|X=x)=\int_0^1 y\cdot f_{Y|X=x}(y)\,dy=\int_0^1\frac{xy+y^2}{x+0.5}\,dy=$$
$$=\frac{0.5x+\frac{1}{3}}{x+0.5}=\frac{x+\frac{2}{3}}{2x+1}$$

□

תרגיל 7.32. נשים לב שלפי נוסחת ההסתברות הנתונה, מספר הביצים שהחרק מטיל מתפלג $X\sim Geom(p)$.

1. אם הוטלו 7 ביצים, מספר הביצים שבקעו מתפלג $Bin(7,0.5)$ שכן כל ביצה היא ניסוי בלתי תלוי עם סיכוי 0.5 לבקוע.
2. בדומה לסעיף הקודם, אם הוטלו X ביצים מתקיים $Y|X\sim Bin(X,0.5)$. לכן ניתן למצוא את התוחלת לפי משפט התוחלת השלמה:

$$\mathbb{E}(Y)=\mathbb{E}(\mathbb{E}(Y|X))=\mathbb{E}(0.5X)=0.5\cdot\frac{1}{p}=5$$

בדומה, נמצא את התוחלת של Y^2 לפי משפט התוחלת השלמה (ניתן גם לחשב את השונות באמצעות משפט השונות השלמה):

$$\begin{aligned}\mathbb{E}\left(Y^2\right)&=\mathbb{E}\left(\mathbb{E}\left(Y^2|X\right)\right)=\mathbb{E}\left(\mathbb{V}(Y|X)+\mathbb{E}^2(Y|X)\right)\\&=\mathbb{E}(0.5(1-0.5)X)+\mathbb{E}\left(0.5^2X^2\right)=0.25\cdot\frac{1}{p}+0.25\left(\mathbb{V}(X)+\mathbb{E}^2(X)\right)\\&=50\end{aligned}$$

ולכן $\mathbb{V}(Y)=50-5^2=25$

3. לו היינו יודעים את הערך של X, ניתן היה לחשב את ההסתברות של המאורע $Y=0$ בקלות שהרי בהינתן מספר הביצים שהוטלו, מאורע זה מתקיים רק אם כולן לא בקעו וההסתברות לכך היא 0.5^X. מאחר והערך לא ידוע, נשתמש בנוסחת ההסתברות השלמה (נוסחא **2.17**) ונתנה על כל האפשרויות עבור X:

$$\begin{aligned}Pr(Y=0)&=\sum_{k=1}^{\infty}Pr(Y=0|X=k)\,Pr(X=k)=\sum_{k=1}^{\infty}0.5^kp(1-p)^{k-1}\\(eq.\,0.15)&=\frac{0.05}{1-0.45}=0.0909\end{aligned}$$

□

תרגיל 7.33. בהמשך לתרגיל 5.14, נסמן ב- $\mathbf{1}_i$ את האינדיקטור לכך שהאדם ה-$i \in \{1,\dots,98\}$ עומד בין אדון וגברת סמית'. מתקיים $\mathbb{E}(\mathbf{1}_i) = Pr(\mathbf{1}_i = 1) = \frac{1}{3}$ ולכן $\mathbb{V}(\mathbf{1}_i) = \frac{1}{3}\cdot\left(1-\frac{1}{3}\right) = \frac{2}{9}$. מאחר ומתקיים $X = \sum_{i=1}^{98}\mathbf{1}_i$ הרי שלפי נוסחא 7.24 נקבל

$$\mathbb{V}(X) = \sum_{i=1}^{98}\mathbb{V}(\mathbf{1}_i) + 2\sum_{i<j} Cov(\mathbf{1}_i,\mathbf{1}_j) = 98\cdot\frac{2}{9} + 2\sum_{i<j}\left(\mathbb{E}(\mathbf{1}_i\mathbf{1}_j) - \mathbb{E}(\mathbf{1}_i)\,\mathbb{E}(\mathbf{1}_j)\right)$$

המכפלה של המשתנים המקריים $\mathbf{1}_i\mathbf{1}_j$ שווה ל-1 רק אם שני האנשים i,j עומדים בין מר וגברת סמית' ו-0 אחרת. נחשב את ההסתברות על ידי צמצום מרחב המדגם לארבעת האנשים הרלוונטיים. מרחב המדגם הוא בחירה של 2 מקומות מתוך ה-4 עבור מר וגברת סמית' והמאורע כולל רק אפשרות אחת: שנבחרו שני המקומות הקיצונים. לכן $Pr(\mathbf{1}_i\mathbf{1}_j = 1) = \frac{1}{\binom{4}{2}} = \frac{1}{6}$. יש בסך הכל $\binom{98}{2}$ זוגות אפשריים ולכן

$$\mathbb{V}(X) = \frac{196}{9} + 2\cdot\binom{98}{2}\cdot\left(\frac{1}{6} - \left(\frac{1}{3}\right)^2\right) = \frac{4949}{9}$$

□

תרגיל 7.34. בתרגיל 4.24 חושבה התפלגות של משתנה מקרי הדומה למשתנה מקרי זה, אך החישוב של התוחלת והשונות על סמך התפלגות זו מסובך מאוד. במקום זאת נשתמש בפירוק לאינדיקטורים. נסמן ב-$\mathbf{1}_i$ אינדיקטור לכך שהזוג ה-$i \in \{1,\dots,50\}$ נבחר. מרחב המדגם הינו בחירה של 30 אנשים מתוך ה-100 בלי סדר ובלי החזרה. המאורע הינו בחירה של 28 אנשים מתוך 98 האנשים שנותרו (בהכרח נבחרו שני בני הזוג ה-i). לפי נוסחאות 5.10 ו-5.17 מתקיים

$$\mathbb{E}(\mathbf{1}_i) = Pr(\mathbf{1}_i = 1) = \frac{\binom{98}{28}}{\binom{100}{30}} \qquad \mathbb{V}(\mathbf{1}_i) = \frac{\binom{98}{28}}{\binom{100}{30}}\left(1 - \frac{\binom{98}{28}}{\binom{100}{30}}\right)$$

מאחר ומספר הזוגות הכולל שנבחר, X, מקיים $X = \sum_{i=1}^{50}\mathbf{1}_i$ הרי ש-$\mathbb{E}(X) = 50\cdot\frac{\binom{98}{28}}{\binom{100}{30}} = 4.39$. לצורך חישוב השונות עלינו לחשב את המתאם בין כל זוג אינדיקטורים, כלומר את התוחלת של $\mathbf{1}_i\mathbf{1}_j$ כאשר $i<j$. מכפלה זו שווה ל-1 רק כאשר נבחרו שני הזוגות. לשם כך יש לבחור את שני הזוגות (4 אנשים) ולבחור עוד 26 אנשים מתוך 96 הנותרים. אחרת, המכפלה שווה ל-0 ולכן

$$\mathbb{E}(\mathbf{1}_i\mathbf{1}_j) = Pr(\mathbf{1}_i\mathbf{1}_j = 1) = \frac{\binom{96}{26}}{\binom{100}{30}}$$

שונות מספר הזוגות שייבחרו הינה

$$\mathbb{V}(X) = \sum_{i=1}^{50}\mathbb{V}(\mathbf{1}_i) + 2\sum_{i<j} Cov(\mathbf{1}_i,\mathbf{1}_j) = 50\cdot 0.08 + 2\binom{50}{2}\left(0.00699 - 0.087^2\right) = 2.21$$

□

תרגיל 7.35. לבעיה זו תכונת חוסר זיכרון, במובן שאם המהמר לא זכה בסיבוב הראשון, החל מהסיבוב השני הבעיה חוזרת על עצמה בניכוי מספר דולרים שכבר לא מצויים בכיסו. לכן, באמצעות התנייה על תוצאת הקוביה

בסיבוב הראשון ומשפט התוחלת השלמה ניתן לבנות משוואה המערבת את תוחלת הרווח של המהמר. נסמן ב-X את הרווח של המהמר בסיום המשחק וב-Y את הרווח שלו בסיבוב הראשון. נחשב את ההתפלגות המותנית של X בהינתן Y:

$$X|Y = \begin{cases} 10 & Y = 10 \\ Z-1 & Y=-1 \\ Z-5 & Y=-5 \end{cases}$$

כאשר Z הוא משתנה מקרי המתאר את הרווח של המהמר מהסיבוב השני ועד סיום המשחק. לאור החזרתיות של המשחק, ל-X ול-Z אותה התפלגות ולכן אותה תוחלת. לפיכך

$$\mathbb{E}\left(X|Y\right) = \begin{cases} 10 & Y = 10 \\ \mathbb{E}\left(X\right)-1 & Y=-1 \\ \mathbb{E}\left(X\right)-5 & Y=-5 \end{cases}$$

ולפי משפט התוחלת השלמה נקבל

$$\begin{aligned}\mathbb{E}\left(X\right) &= \mathbb{E}\left(\mathbb{E}\left(X|Y\right)\right) = 10\cdot Pr\left(Y=10\right) + \left(\mathbb{E}\left(X\right)-1\right)Pr\left(Y=-1\right) + \left(\mathbb{E}\left(X\right)-5\right)Pr\left(Y=-5\right) \\ &= 10\cdot\frac{1}{6} + \left(\mathbb{E}\left(X\right)-1\right)\frac{4}{6} + \left(\mathbb{E}\left(X\right)-5\right)\frac{1}{6} = \frac{5}{6}\mathbb{E}\left(X\right) + \frac{1}{6}\end{aligned}$$

פתרון המשוואה הינו $\mathbb{E}\left(X\right)=1$.

דרך נוספת: נסמן ב-U את מספר הסיבובים עד שהמהמר זוכה, $U\sim Geom\left(\frac{1}{6}\right)$. בהינתן U, מספר הסיבובים בהם הוא הפסיד דולר אחד הינו $V|U\sim Bin\left(U-1,\frac{4}{5}\right)$. בכל סיבוב כזה הוא הפסיד דולר. בשאר $U-1-V$ הסיבובים הוא הפסיד 5 דולרים ולכן הרווח הכולל שלו עד הזכייה בהינתן U הינו $X|U = 10-V-5\left(U-1-V\right) = 15+4V-5U$. תוחלת הרווח המותנה בהינתן U הינה $\mathbb{E}\left(X|U\right) = 15+4\cdot\frac{4}{5}\cdot\left(U-1\right)-5U = 11.8-1.8U$. לפי נוסחת התוחלת השלמה נקבל

$$\mathbb{E}\left(X\right) = \mathbb{E}\left(\mathbb{E}\left(X|U\right)\right) = \mathbb{E}\left(11.8-1.8U\right) = 11.8-1.8\cdot 6 = 1$$

$\square$

תרגיל 7.36. לו היינו יודעים כמה בדיוק סטודנטים קנו קפה, ניתן היה לחשב את התוחלת על ידי פירוק לסכום. היות ואנחנו לא יודעים, נגדיר מספר זה בתור משתנה מקרי ונחשב את התוחלת המותנית. נסמן ב-N את מספר הסטודנטים שקונים כוס קפה וב-X_i את המחיר אותו הסטודנט ה-$i\in\{1,\ldots,N\}$ משלם. הרווח הכולל של בית הקפה הינו $X=\sum_{i=1}^{N}X_i$ ולפי משפט התוחלת השלמה:

$$\mathbb{E}\left(X\right) = \mathbb{E}\left(\mathbb{E}\left(X|N\right)\right) = \mathbb{E}\left(\mathbb{E}\left(\sum_{i=1}^{N}X_i\middle|N\right)\right) = \mathbb{E}\left(\sum_{i=1}^{N}\mathbb{E}\left(X_i|N\right)\right)$$

כל כוס יכולה להיות גדולה או קטנה בהסתברות שווה כך ש-$\mathbb{E}\left(X_i|N\right) = \frac{5+8}{2} = 6.5$. יש N מחוברים בסכום ולכן $\mathbb{E}\left(X\right) = \mathbb{E}\left(6.5N\right)$. מאחר ו-$N\sim Bin\left(10,000,0.35\right)$ הרי שתוחלת הרווח הכוללת הינה $\mathbb{E}\left(X\right) = 6.5\cdot 10000\cdot 0.35 = 22750$.

דרך נוספת היא על ידי הגדרת Y_i בתור הכסף ששילם הסטודנט ה-$i \in \{1, \ldots, 10,000\}$. Y_i יכול להיות 0 (אם הוא לא קנה כלום), 5 אם הוא קנה כוס קטנה או 8 אם הוא קנה כוס גדולה, לכן ההתפלגות של Y_i הינה

$$Y_i = \begin{cases} 0 & 0.65 \\ 5 & 0.35 \cdot 0.5 \\ 8 & 0.35 \cdot 0.5 \end{cases}$$

לכן התוחלת של הכסף שמשלם כל סטודנט הינה $\mathbb{E}(Y_i) = 2.275$ ותוחלת הרווח הכולל הינה $\mathbb{E}(Y) = \mathbb{E}\left(\sum_{i=1}^{10,000} Y_i\right) = 22750$.

ניתן לחשב את השונות בשתי הדרכים שצוינו לעיל, באמצעות משפט השונות השלמה (משפט 7.6) או באמצעות שונות של סכום תוך שימוש בכך שהסטודנטים בלתי תלויים:

$$\mathbb{V}(Y) = \mathbb{V}\left(\sum_{i=1}^{10,000} Y_i\right) = \sum_{i=1}^{10,000} \mathbb{V}(Y_i) = 10,000 \cdot \left(8^2 \cdot 0.175 + 5^2 \cdot 0.175 - 2.275^2\right) = 103,994$$

□

תרגיל 7.37. נסמן ב-Y את מספר הנוסעים שנסעו באוטובוס הראשון. לכל אדם יש הסתברות זהה וקבועה לנסוע באוטובוס זה, לכן $Y \sim Bin(100, 0.75)$. בהינתן הערך של Y, הסיכוי שאביטל שאלה מישהו מהאוטובוס הראשון הוא $\frac{Y}{100}$ והתשובה שתקבל תהיה Y. לכן, התשובה המותנית שאביטל תקבל בהינתן Y הינה

$$X|Y = \begin{cases} Y & \frac{Y}{100} \\ 100 - Y & 1 - \frac{Y}{100} \end{cases}$$

ולכן התוחלת המותנית הינה

$$\mathbb{E}(X|Y) = \frac{Y^2}{100} + \frac{(100-Y)^2}{100} = \frac{2Y^2 - 200Y + 100^2}{100}$$

ולפי נוסחת התוחלת השלמה התשובה הממוצעת שאביטל תקבל הינה

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}\left(\frac{2Y^2 - 200Y + 100^2}{100}\right) = \frac{1}{50}\mathbb{E}\left(Y^2\right) - 2\mathbb{E}(Y) + 100 \\ &= \frac{1}{50}\left(\mathbb{V}(Y) + \mathbb{E}^2(Y)\right) - 2\mathbb{E}(Y) + 100\end{aligned}$$

נציב את התוחלת והשונות של משתנה מקרי בינומי ונקבל

$$\mathbb{E}(X) = \frac{100 \cdot 0.75 \cdot 0.25 + (100 \cdot 0.75)^2}{50} - 2 \cdot 100 \cdot 0.75 + 100 = 62.875$$

התשובה שתקבל אביטל גדולה מהממוצע עליו מדווח משרד התחבורה, 50. הסיבה היא שהאוטובוסים לא סימטריים ויש אוטובוס צפוף ואוטובס מרווח. בממוצע, באוטובוס הראשון נוסעים הרבה יותר נוסעים מאשר באוטובוס השני, לכן לאביטל יש סיכוי גבוה יותר לשאול מישהו מהאוטובוס הצפוף וכל אדם כזה בילה את הנסיעה עם יותר אנשים מהממוצע. באופן כללי, מספר גדול יותר של אנשים נוסעים באוטובוס צפופים ולכן יותר

אנשים יתלוננו על צפיפות באוטובוסים. באופן דומה, כאשר אוטובוס מאחר, יותר אנשים חווים את האיחור (כי הם הולכים ומגיעים) ולכן שוב – יותר אנשים יתלוננו על איחורים ויהיה רושם שהאיחור הממוצע גבוה ממה שהוא באמת. $\square$

תרגיל 7.38. המשתנים המקריים X, Y בלתי תלויים ולכן גם פונקציות שלהם בלתי תלויות, כלומר e^{tX}, e^{tY} בלתי תלויים לכל t. בפרט פונקציות אלו גם בלתי מתואמות ולכן תוחלת המכפלה שווה למכפלת התוחלות:

$$M_{X+Y}(t) = \mathbb{E}\left(e^{t(X+Y)}\right) = \mathbb{E}\left(e^{tX}\right)\mathbb{E}\left(e^{tY}\right) = M_X(t)\, M_Y(t)$$

$\square$

תרגיל 7.39. נסמן ב-A את המאורע שנבחר מטבע הוגן. הסיכוי ההטלה ה-i היא עץ תחושב לפי נוסחת ההסתברות השלמה:

$$Pr(\mathbf{1}_i = 1) = Pr(\mathbf{1}_i = 1|A)\,Pr(A) + Pr\left(\mathbf{1}_i = 1|\overline{A}\right)Pr\left(\overline{A}\right) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{2} = \frac{3}{8}$$

ולכן גם $\mathbb{E}(\mathbf{1}_i) = \frac{3}{8}$. נבחר $\epsilon = \frac{1}{16}$ ונבדוק האם החוק החלש של המספרים הגדולים מתקיים. לכל n מתקיים:

$$Pr\left(|\overline{\mathbf{1}}_n - \frac{3}{8}| > \frac{1}{16}\right) = Pr\left(|\overline{\mathbf{1}}_n - \frac{3}{8}| > \frac{1}{16}\middle|A\right)Pr(A) + Pr\left(|\overline{\mathbf{1}}_n - \frac{3}{8}| > \frac{1}{16}\middle|\overline{A}\right)Pr\left(\overline{A}\right)$$

בהינתן A, המטבע הוגן ולכן ההטלות בלתי תלויות בעלות תוחלת מותנית $\mathbb{E}(\mathbf{1}_i|A) = \frac{1}{2}$ ולפי החוק החלש מתקיים $Pr\left(|\overline{\mathbf{1}}_n - \frac{1}{2}| > \frac{1}{16}|A\right) \xrightarrow[n\to\infty]{} 0$ כך ש-$Pr\left(|\overline{\mathbf{1}}_n - \frac{3}{8}| > \frac{1}{16}|A\right) \xrightarrow[n\to\infty]{} 1$. באותו אופן, בהינתן $\overline{A}$, המטבע לא הוגן ולכן ההטלות בלתי תלויות, בעלות תוחלת מותנית $\mathbb{E}\left(\mathbf{1}_i|\overline{A}\right) = \frac{1}{4}$ ולפי החוק החלש מתקיים $Pr\left(|\overline{\mathbf{1}}_n - \frac{1}{4}| > \frac{1}{16}|\overline{A}\right) \xrightarrow[n\to\infty]{} 0$ כך ש-$Pr\left(|\overline{\mathbf{1}}_n - \frac{3}{8}| > \frac{1}{16}|\overline{A}\right) \xrightarrow[n\to\infty]{} 1$. לסיכום, בשני המקרים, בגבול $n \to \infty$ הגבול הוא **1** ולכן

$$\lim_{n\to\infty} Pr\left(|\overline{\mathbf{1}}_n - \frac{3}{8}| > \frac{1}{16}\right) = 1$$

בסתירה לחוק החלש של המספרים הגדולים. הסיבה שהסדרה $\mathbf{1}_1, \mathbf{1}_2, \ldots$ לא מקיימת את החוק החלש של המספרים הגדולים היא שהמשתנים המקריים תלויים. אם יצא פלי באחת ההטלות, זה מגדיל את הסיכוי שנבחר המטבע הלא הוגן ולכן מגדיל את הסיכוי שבהטלות אחרות יצא גם פלי. לכן כאשר נבצע את הניסוי בפועל, הממוצע של האינדיקטורים יתקרב ל-$\frac{1}{2}$ אם נבחר מטבע הוגן או ל-$\frac{1}{4}$ אם המטבע הוא לא הוגן ובכל מקרה לא יתקרב ל-$\frac{3}{8}$. מסיבה זו גם משפט הגבול המרכזי (משפט **6.16**) לא מתקיים עבור משתנים מקריים אלו. $\square$

8 תהליך פואסון

בפרק זה נשתמש בכלים שרכשנו בפרקים הקודמים על מנת לנתח תופעה שכיחה מאוד המכונה תהליך פואסון. תהליך זה מאפשר לתאר מתמטית בעיות רבות מהחיים, דוגמת התפרקויות של חלקיקים רדיואקטיביים, שערים במשחק כדורגל, תקלות בקווי תקשורת, הגעת קונים לחנות ספרים ועוד. בכל הבעיות הללו נספור את מספר האירועים (התפרקויות, שערים, תקלות, קונים) שמתרחשים בפרקי זמן שונים וננתח את התפלגות הזמן בין האירועים. מימד הזמן הוא שהופך את הבעיה ל"תהליך", בניגוד למשתנים מקריים רגילים המתארים ניסוי בודד ללא מרכיב של "זמן".[1]

נפתח בדוגמא. בקניון מסויים מבקרים **10,000** אנשים ביום. לכל מבקר יש הסתברות של $p = 0.0001$ לקנות ספר של עמוס עוז בחנות הספרים באופן בלתי תלוי במבקרים האחרים. בהנחה וכל לקוח הוא ניסוי בלתי תלוי באחרים עם הסתברות קבועה לקנות ספר, מספר הספרים הכולל שיימכר הוא משתנה מקרי בינומי עם הפרמטרים $Bin(10,000, 0.0001)$. בפרק **6.1.6** ראינו שניתן לקרב התפלגות זו באמצעות משתנה מקרי פואסוני עם פרמטר $\lambda = 10,000 \cdot 0.0001 = 1$. בעל החנות מעוניין להעריך כמה ספרים יימכרו במשך שבוע. בהנחה שבקניון מבקרים **10,000** אנשים בכל יום ללא תלות בימים האחרים, הרי שמספר המבקרים בשבוע יהיה **70,000** ולכן, בקירוב, מספר הספרים שיימכרו יתפלג לפי התפלגות פואסון עם פרמטר $\lambda = 70,000 \cdot 0.0001 = 7$. באופן דומה ובהנחה שהמבקרים מגיעים בצורה אחידה לאורך היום, מספר המבקרים בפרק זמן של Δt יהיה $10,000 \cdot \Delta t$ ולכן מספר הספרים שיימכרו יתפלג לפי התפלגות פואסון עם פרמטר $\lambda = 10,000 \cdot \Delta t \cdot 0.0001 = \Delta t$. תהליך כזה, בו מספר האירועים (ספרים שנמכרים) בפרק זמן כלשהו מתפלג לפי התפלגות פואסון עם פרמטר הלינארי במשכו של פרק הזמן, מכונה תהליך פואסון.

הגדרה 8.1 (תהליך פואסון). תהליך אקראי $N(t)$ הסופר את מספר האירועים מזמן 0 עד לזמן $t \geq 0$ נקרא תהליך פואסון כאשר התנאים הבאים מתקיימים:

1. $N(0) = 0$.
2. קיים מספר קבוע $\lambda > 0$ כך שמספר האירועים בין כל שני זמנים $0 \leq t < s$ מתפלג לפי $Pois(\lambda(s-t))$.
3. פרקי זמן זרים הם בלתי תלויים.

הקבוע λ מכונה "הקצב של התהליך" ויחידותיות הן "אירועים ליחידת זמן".

בדוגמת הקניון, הקצב של התהליך הוא מכירה של ספר אחד ליום, $\lambda = 1$. ההנחה הראשונה קובעת שבתחילת התהליך לא נמכר אף ספר ומשמשת כתנאי שפה לחישובים האחרים. ההנחה השנייה קובעת, כפי שראינו, שמספר הספרים שנמכרו בפרק זמן Δt מתפלג לפי התפלגות פואסון עם פרמטר Δt. ההנחה השלישית קובעת שפרקי זמן זרים הם בלתי תלויים. הנחנו זאת כאשר קבענו שמספר המבקרים בכל יום הוא קבוע, שהקונים מגיעים בצורה אחידה לאורך היום ושלכל אחד יש סיכוי שווה לקנות ספר, ללא תלות בקונים האחרים או ביום.

באמצעות התפלגות פואסון ניתן לענות על שאלות רבות לגבי פרק זמן נתון, דוגמת מה הסיכוי שיימכרו ביום מסויים 2 ספרים, מהי תוחלת מספר הספרים שיימכרו במשך שבוע, מהי שונות מספר הספרים שיימכרו במשך חודש וכן הלאה. מאחר וישנו מימד נוסף בבעיה, והוא מימד הזמן, ניתן לשאול גם שאלות מעניינות אחרות

[1] לרוב אכן נעסוק בתהליכי פואסון הקורים לאורך ציר הזמן (כמו בדוגמאות לעיל) אך באותה מידה "זמן" יכול לתאר כל ציר רציף שתהליכים יכולים לקרות לאורכו. למשל – מספר התקרים שקורים בגלגל ליחידת מרחק.

העוסקות בזמן העובר בין מכירות ספרים, דוגמת כיצד מתפלג הזמן מפתיחת החנות ועד הספר הראשון שנמכר, מהי תוחלת הזמן עד המכירה של הספר הרביעי, מהי שונות הזמן שעבר בין הספר השלישי שנמכר ועד הספר העשירי וכן הלאה. מתברר, שפרק הזמן עד האירוע הפואסוני הראשון (ובאופן כללי, בין כל אירוע לאירוע העוקב) מתפלג לפי התפלגות מעריכית (פרק 6.1.6).

***משפט* 8.1** (התפלגות הזמן בין האירועים בתהליך פואסון). *הזמן בין תחילת התהליך לבין האירוע הראשון בתהליך פואסון עם קצב λ, כמו גם הזמן בין האירוע ה-n ועד האירוע ה-$n+1$ מתפלג מעריכית עם פרמטר λ.*

הוכחה. נחשב את פונקציית ההתפלגות המצטברת של הזמן עד האירוע הראשון, T. בבירור, לכל $t<0$ מתקיים $F_T\left(t\right)=0$. עבור $t>0$ המאורע $T\leq t$ מתקיים אם ורק אם $N\left(t\right)>0$. מאורע זה יותר נוח לחשב על ידי המשלים והסתברותו:

$$F_T\left(t\right)=Pr\left(N\left(t\right)>0\right)=1-Pr\left(N\left(t\right)=0\right)=1-e^{-\lambda t}$$

זוהי פונקציית ההתפלגות המצטברת של התפלגות מעריכית (נוסחא 6.33) ולכן $T\sim Exp\left(\lambda\right)$. באופן דומה ובשילוב תכונת חוסר הזיכרון של התפלגות מעריכית, נקבל שהזמן בין כל שני אירועים סמוכים, דהיינו הזמן בין האירוע ה-n ועד האירוע הבא אחריו, מתפלג $Exp\left(\lambda\right)$ גם כן. $\square$

תרגיל 8.1. כיצד מתפלג הזמן עד האירוע השני בתהליך פואסון עם קצב λ?

מאחר ומשתנה מקרי מעריכי הוא חסר זיכרון, הרי שניתן להתחיל לעקוב אחרי התהליך בכל נקודת זמן מבלי לאבד מידע. לא משנה באיזו שעה אחרי הפתיחה יגיע מנהל החנות לחנות, הזמן שיצטרך לחכות עד הקונה הבא מתפלג מעריכית עם אותו הפרמטר, ללא תלות במספר הקונים הקודמים או השעה שבה החנות נפתחה. מסיבה זו תהליכי פואסון הם נוחים מאוד – לא צריך לדעת כלום על העבר של התהליך בכדי לחשב את כל הדרוש לנו לגבי העתיד. למשל, אם אוטובוסים מגיעים לתחנת אוטובוס לפי תהליך פואסון עם קצב של 6 אוטובוסים בשעה, הרי שהזמן עד האוטובוס הבא **מרגע הגעתנו לתחנה**, מתפלג מעריכית עם פרמטר 6, ללא תלות בזמן שבו עבר האוטובוס הקודם. אם אכן אוטובוסים היו מגיעים לפי תהליך פואסון, לא היה טעם לשאול את הנוסעים האחרים בתחנה אם הקו שאנחנו מחכים לו עבר או לא. ללא תלות בתשובה שלהם, הזמן שניאלץ לחכות מתפלג מעריכית. במציאות אוטובוסים לא מגיעים לפי תהליך פואסון ולכן מקובל לשאול את הנוסעים אם הקו עבר ומתי. ככל שחלף יותר זמן מאז האוטובוס האחרון, כך גדל הסיכוי שהאוטובוס הבא יגיע בקרוב. בדומה, כאשר מבצעים ניסויים עם חומרים רדיואקטיביים, ניתן להתחיל את הניסוי בכל רגע שכן לא משנה מתי יחל הניסוי – החלקיקים ייפלטו מהחומר לפי תהליך פואסון. אחרת, היינו צריכים לדעת את ההיסטוריה של החומר ולהתחשב בכל העבר שלו כהכנה לניסוי.

בעיה 8.1. אוטובוסים מגיעים תחנה לפי תהליך פואסון עם קצב של 6 לשעה החל מ-5 בבוקר.

1. מה ההסתברות שבשעה הראשונה יגיעו 5 אוטובוסים ובעשר הדקות הבאות לא יגיע אף אוטובוס?
2. מה ההסתברות שבשעה הראשונה יגיעו 5 אוטובוסים ובין 5:00 ל-5:30 יגיעו 4 אוטובוסים?
3. בהינתן שבשעה הראשונה הגיעו 5 אוטובוסים, מה הסיכוי שבשעה השניה יגיעו 7 אוטובוסים?
4. דני מגיע לתחנה בשעה 6:50. מהי תוחלת הזמן שעליו להמתין עד האוטובוס הבא?

הוכחה. נסמן ב-$N\left(t\right)$ את מספר האוטובוסים המגיעים לתחנה במהלך t השעות הראשונות לאחר 5 בבוקר. לפי הנתונים, מדובר על תהליך פואסון עם קצב של $\lambda=6$ אוטובוסים בשעה.

1. המאורע המבוקש הינו חיתוך של שני מאורעות: בשעה הראשונה מגיעים 5 אוטובוסים, כלומר $N\left(1\right)=5$, ובעשר הדקות הבאות לא מגיע אף אוטובוס, כלומר $N\left(1\frac{1}{6}\right)-N\left(1\right)=0$. מאחר ומדובר על

פרקי זמן זרים[2] ולכן הם בלתי תלויים וניתן לחשב את ההסתברות של החיתוך על ידי מכפלת ההסתברויות:

$$Pr\left(N(1)=5\cap N\left(1\frac{1}{6}\right)-N(1)=0\right)=Pr\left(N(1)=5\right)\cdot Pr\left(N\left(1\frac{1}{6}\right)-N(1)=0\right)$$

לפי ההגדרה של תהליך פואסון, מספר האירועים בשעה מתפלג פואסונית עם פרמטר $\lambda\cdot 1=6$. מספר האירועים בעשר הדקות (שישית שעה) מתפלג פואסונית עם פרמטר $\lambda\cdot\frac{1}{6}=1$. לכן כל אחת מההסתברויות ניתן לחשב על ידי הצבה בנוסחא של משתנה מקרי פואסוני:

$$e^{-6}\frac{6^5}{5!}\cdot e^{-1}\frac{1^0}{0!}=0.059$$

2. בדומה לסעיף הקודם, המאורע המבוקש הינו $N(1)=5\cap N\left(\frac{1}{2}\right)=4$. ההבדל הוא שהפעם לא מדובר על פרקי זמן זרים ולכן לא ניתן לפרק את ההסתברות של החיתוך למכפלת ההסתברויות בצורה ישירה. לשם כך נתאר את המאורע המבוקש באמצעות מאורע שקול על פרקי זמן זרים. במאורע המבוקש מגיעים בחצי השעה הראשונה 4 אוטובוסים ובשעה הראשונה בסך הכל חמישה, וזה שקול לכך שבחצי השעה הראשונה מגיעים 4 אוטובוסים ובחצי השעה הבאה מגיע אוטובוס אחד. לפיכך, $\left\{N(1)=5\cap N\left(\frac{1}{2}\right)=4\right\}=\left\{N\left(\frac{1}{2}\right)=4\cap N(1)-N\left(\frac{1}{2}\right)=1\right\}$ וההסתברות המבוקשת הינה

$$\begin{aligned}Pr\left(N(1)=5\cap N\left(\frac{1}{2}\right)=4\right)&=Pr\left(N\left(\frac{1}{2}\right)=4\cap N(1)-N\left(\frac{1}{2}\right)=1\right)\\&=Pr\left(N\left(\frac{1}{2}\right)=4\right)\cdot Pr\left(N(1)-N\left(\frac{1}{2}\right)=1\right)\\&=e^{-3}\frac{3^4}{4!}\cdot e^{-3}\frac{3^1}{1!}=0.025\end{aligned}$$

3. לפי התכונות של תהליך פואסון, פרקי זמן זרים הם בלתי תלויים ולכן השעה הראשונה לא משפיעה על השעה השניה. מספר האוטובוסים בשעה השניה מתפלג פואסונית עם פרמטר $\lambda=6$ ולכן הסיכוי שיגיעו בה 7 אוטובוסים הינו

$$Pr\left(N(2)-N(1)=7|N(1)=5\right)=Pr\left(N(2)-N(1)=7\right)=e^{-6}\frac{6^7}{7!}=0.138$$

4. הזמן בין האוטובוס שדני פספס לאוטובוס הבא מתפלג מעריכית עם פרמטר $\lambda=6$. מתכונת חוסר הזיכרון, לא משנה כמה זמן אחרי האוטובוס הקודם דני הגיע, זמן ההמתנה עד האוטובוס הבא יתפלג בכל זאת מעריכית עם פרמטר $\lambda=6$. לכן תוחלת זמן ההמתנה עד האוטובוס הראשון מתקבלת לפי תוחלת של משתנה מקרי מעריכי (נוסחא 6.35): $\frac{1}{\lambda}=\frac{1}{6}$ שעה, כלומר 10 דקות.

□

תרגיל 8.2. שיחות טלפון מגיעות למרכזיה בין השעות 8:00 ל-16:00 בתהליך פואסון עם קצב ממוצע של 20 שיחות בשעה.

[2] החיתוך בין שני פרקי הזמן הוא נקודה בודדת, $t=1$, אבל הסיכוי שאירוע יקרה בנקודה בודדת בזמן הוא 0 (משתנה מקרי רציף) ולכן פרקי זמן אלו אכן זרים.

1. מה ההסתברות שב-10 הדקות הראשונות התקבלו 2 שיחות? מה ההסתברות שב-5 הדקות האחרונות התקבלה שיחה אחת?
2. מה ההסתברות שהשיחה הראשונה תגיע למרכזיה בין דקה אחת ל-3 דקות לאחר הפתיחה?
3. אף שיחה לא הגיעה ב-3 הדקות הראשונות. מה ההסתברות שנמתין לפחות דקה נוספת עד שתגיע השיחה הראשונה?

תרגיל 8.3. חלקיקים נפלטים מחומר רדיואקטיבי בתהליך פואסון עם קצב של 0.5 חלקיק לשנייה. חשבו את ההסתברויות של המאורעות הבאים:

1. ייפלט לפחות חלקיק אחד בשנייה נתונה.
2. ייפלטו יותר מ-3 חלקיקים ב-5 שניות.
3. בחמש השניות הראשונות נפלטו שלושה חלקיקים. מהי התפלגות מספר החלקיקים שנפלטו בשנייה הראשונה?

נדגיש את ההבדל בין שני המשתנים המקריים המופיעים בתהליכי פואסון: משתנה מקרי פואסוני ומשתנה מקרי מעריכי. המשתנה המקרי הפואסוני הינו משתנה מקרי **בדיד** הסופר את **מספר האירועים** שהתרחשו בפרק זמן נתון. לעומתו, המשתנה המקרי המעריכי הינו משתנה מקרי **רציף** הסופר את **הזמן** עד האירוע הבא. לרוב נשתמש במשתנה המקרי הפואסוני בכדי לענות על שאלות העוסקות במספר אירועים ובמשתנה מקרי מעריכי בכדי לענות על שאלות העוסקות בזמן בין אירועים. יחד עם זאת, חשוב לזכור ששני המשתנים המקריים מתייחסים לאותו תהליך ובמקרים מסויימים ניתן לענות על אותה השאלה בעזרת שניהם. למשל, את המאורע "לא אירע אירוע פואסוני במשך שעה אחת" ניתן לתאר על ידי התייחסות למספר האירועים שאירעו במשך שעה ($N(1) = 0$) או על ידי דרישה שהאירוע הראשון היה לאחר יותר משעה ($T > 1$). שתי האפשרויות יניבו כמובן את אותה התשובה.

שני המשפטים הבאים מלמדים שכאשר ידוע מספר האירועים הפואסוניים שהתרחשו בפרק זמן מסויים, הם כבר לא מתפלגים לפי התפלגות פואסון. במקום זאת, זמן התרחשות האירוע (במסגרת פרק הזמן בו ידוע שהוא התרחש) מתפלג התפלגות אחידה ומספר האירועים בפרק זמן מסויים מתפלג לפי התפלגות בינומית.

משפט 8.2. *יהי תהליך פואסון עם קצב λ. אם ידוע שאירע אירוע פואסוני מתישהו בין זמן $t = 0$ לזמן $t = T$ אזי זמן האירוע מתפלג אחיד $U(0, T)$.*

הוכחה. נסמן ב-$N(t)$ את מספר האירועים מזמן 0 ועד זמן t ונסמן ב-X את זמן האירוע. נחשב את פונקציית ההתפלגות המצטברת של X עבור $t \in [0, T]$

$$F_X(t) = Pr(X \le t) = Pr(N(t) = 1 | N(T) = 1)$$

שכן הזמן עד האירוע קטן מ-t אם בזמן t כבר היה אירוע (וכל זאת מותנה בכך שהיה אירוע עד זמן T). נחשב את ההסתברות המותנית על ידי נוסחא 2.9 ופירוק החיתוך של תחומי הזמן הלא זרים לחיתוך של פרקי זמן זרים:

$$\begin{aligned} F_X(t) &= \frac{Pr(N(t) = 1 \cap N(T) = 1)}{Pr(N(T) = 1)} = \frac{Pr(N(t) = 1 \cap N(T) - N(t) = 0)}{Pr(N(T) = 1)} \\ &= \frac{e^{-\lambda t}\frac{(\lambda t)^1}{1!} \cdot e^{-\lambda(T-t)}\frac{(\lambda(T-t))^0}{0!}}{e^{-\lambda T}\frac{(\lambda T)^1}{1!}} = \frac{\lambda t}{\lambda T} = \frac{t}{T} \end{aligned}$$

זוהי פונקציית התפלגות מצטברת של משתנה מקרי אחיד רציף (נוסחא 6.28) כך ש-$X \sim U(0, T)$, כדרוש. □

תרגיל 8.4. כיצד מתפלג הזמן עד האירוע הראשון בהינתן שהאירוע הראשון היה מתישהו בין זמן $t=0$ לזמן $t=T$? מדוע יש הבדל בין תרגיל זה לבין משפט 8.2?

משפט 8.3. *יהי תהליך פואסון עם קצב λ. אם ידוע שאירעו n אירועים פואסוניים בין זמן $t=0$ לבין זמן $t=T$, אזי מספר האירועים הפואסוניים בפרק הזמן באורך ΔT המוכל כולו בקטע $[0,T]$מתפלג $Bin\left(n,\frac{\Delta T}{T}\right)$. הדבר נכון גם אם ΔT מורכב ממספר פרקי זמן זרים שסכום זמניהם הוא ΔT.*

הוכחה. נוכיח עבור פרק זמן יחיד ΔT המתחיל בזמן t_0 (כך ש-$[t_0,t_0+\Delta T]\subseteq[0,T]$) . כאשר מדובר על מספר פרקי זמן זרים שסכום זמניהם הוא ΔT, ההוכחה דומה.

נסמן ב-$N\left(t\right)$ את מספר האירועים הפואסוניים בין זמן 0 לזמן t וב-X את מספר האירועים שאירעו בקטע $[t_0,t_0+\Delta T]$ בהינתן שבקטע $[0,T]$ אירעו n אירועים. מספר האירועים שיכול להיות בקטע $[t_0,t_0+\Delta T]$ הוא לפחות 0 ולכל היותר n (כולם) ולכן $supp\left(X\right)=\{0,\ldots,n\}$. יהי $k\in supp\left(X\right)$ ונחשב את פונקציית ההסתברות של X עבור ערך זה:

$$\begin{aligned}P_X\left(k\right)&=Pr\left(N\left(t_0+\Delta T\right)-N\left(t_0\right)=k|N\left(T\right)=n\right)\\&=\frac{Pr\left(N\left(t_0+\Delta T\right)-N\left(t_0\right)=k\cap N\left(T\right)=n\right)}{Pr\left(N\left(T\right)=n\right)}\end{aligned}$$

נחלק את המונה לשני פרקי זמן זרים. פרק הזמן הראשון הינו $[t_0,t_0+\Delta T]$ בו הדרישה היא שיהיו k אירועים ופרק הזמן השני הינו $[0,t_0]\cup[t_0+\Delta T,T]$ בו בהכרח היו שאר $n-k$ האירועים. אמנם פרק הזמן השני מורכב משני פרקי זמן לא רציפים, אך מתכונות תהליך פואסון זה לא משנה ומספר האירועים בו מתפלג פואסונית עם פרמטר λ כפול משך הזמן הכולל: $T-\Delta T$. נציב זאת בנוסחא האחרונה ונקבל:

$$\begin{aligned}P_X\left(k\right)&=\frac{e^{-\lambda\Delta T}\frac{\left(\lambda\Delta T\right)^k}{k!}\cdot e^{-\lambda\left(T-\Delta T\right)}\frac{\left(\lambda\left(T-\Delta T\right)\right)^{n-k}}{(n-k)!}}{e^{-\lambda\Delta T}\frac{\left(\lambda T\right)^n}{n!}}=\frac{n!}{k!\left(n-k\right)!}\frac{\Delta T^k\left(T-\Delta T\right)^{n-k}}{T^n}\\&=\binom{n}{k}\left(\frac{\Delta T}{T}\right)^k\left(1-\frac{\Delta T}{T}\right)^{n-k}\end{aligned}$$

הנוסחא שהתקבלה מתארת התפלגות בינומית ולכן $X\sim Bin\left(n,\frac{\Delta T}{T}\right)$. $\square$

שימו לב שהן במשפט 8.2 והן במשפט 8.3, ההתפלגות לא תלויה כלל בקצב האירועים של התהליך, λ. הסיבה לכך היא ששתי ההתפלגויות מחושבות עבור מספר אירועים ידוע. ברגע שנודע שהיה אירוע פואסוני במהלך זמן מסויים, לא משנה כלל אם מראש האירועים היו בקצב מהיר או איטי, אם אירוע אחד בזמן זה הוא מאורע שכיח או נדיר - אירוע זה קרה כבר ועכשיו רק נותר לתהות מתי. תופעה זו מזכירה את תרגיל 6.2, שם ראינו שבהינתן ששיכור צעד 100 צעדים והגיע לנקודה מסוימת, הסיכוי שלו לצעוד את הצעד האחרון ימינה קבוע וכלל לא מושפע מהסיכוי שהיה לו מראש לצעוד ימינה. גם אם מראש היה לו סיכוי מאוד גבוה לצעוד ימינה, ברגע שאנחנו יודעים שהוא צעד רק 55 צעדים מתוך ה-100ימינה, הסיכוי לצעד האחרון להיות ימינה הוא $\frac{55}{100}$. באופן דומה, כאשר מטילים קוביה 100 פעמים, הסיכוי לקבל 100 פעמים 1 או 100 פעמים 5 מאוד נמוך. אבל בהינתן שאחד משני מאורעות אלו התרחש, הסיכוי של כל אחד מהם הוא חצי.

תרגיל 8.5. תקלות בקו תקשורת מתרחשות לפי תהליך פואסון עם קצב של $\lambda=0.25$ תקלות בשעה.

1. מהי ההסתברות שב-4 השעות הראשונות התרחשה לכל היותר תקלה אחת וב-8 השעות הבאות התרחשו לכל הפחות 2 תקלות?
2. מאז התקלה האחרונה עברו 4 שעות. מה ההסתברות שנשלים 10 שעות רצופות ללא תקלה?
3. ב-12 השעות הראשונות התרחשה תקלה אחת בדיוק. מה הסיכוי שהיא התרחשה בשעה הראשונה או בשעה האחרונה של פרק זמן זה?
4. ב-10 השעות הראשונות התרחשו 20 תקלות. מה הסיכוי שבכל שעה התרחשו בדיוק 2 תקלות?

תרגיל 8.6. אל מסעדה מגיעים לקוחות בתהליך פואסון עם קצב של 4 לקוחות לשעה. המסעדה נפתחת ב-8 בבוקר ומגישה ארוחות בוקר עד השעה 12:30. לאחר מכן מוגשות ארוחות צהריים עד שהמסעדה נסגרת ב-17:00.

1. מהי תוחלת מספר הלקוחות שהגיעו למסעדה במשך יום שלם?
2. לקוח המגיע לארוחת בוקר משאיר טיפ המתפלג לפי $U\left(20,50\right)$. לקוח המגיע לארוחת צהריים משאיר טיפ המתפלג לפי $U\left(40,80\right)$. מהי תוחלת סך הטיפים שהושאר במסעדה במשך יום שלם?
3. כיצד מתפלג מספר הלקוחות שהגיעו לאכול בין השעות 12:00-13:00 בהינתן שלארוחת הבוקר הגיעו 2 לקוחות?

תופעה מעניינת מתרחשת כאשר מספר תהליכים פואסוניים בלתי תלויים מתאחדים לתהליך אחד בו נספרים כלל האירועים, ללא תלות בתהליך הפואסוני בו הם התרחשו. למשל, נניח שאל שדה תעופה מגיעים מטוסים מהמזרח בתהליך פואסון עם פרמטר 10 מטוסים לשעה והמערב בתהליך פואסון עם פרמטר 11 מטוסים לשעה, ללא תלות בין התהליכים. פקח הטיסה שבשדה מתעניין רק בתהליך הגעת המטוסים וכלל לא אכפת לו האם הגיעו מהמזרח או מהמערב. מאחר וסכום של משתנים מקריים פואסונים בלתי תלויים הוא גם כן פואסוני (משפט 7.3) הרי שגם פקח הטיסה חווה תהליך פואסון, אך עם קצב של 21 מטוסים לשעה. באופן כללי, סכום של תהליכי פואסון בלתי תלויים הוא גם תהליך פואסון, אשר הקצב שלו שווה לסכום הקצבים של התהליכים המרכיבים אותו.

***משפט* 8.4.** *יהיו* $N_1\left(t\right), N_2\left(t\right)$ *תהליכי פואסון פואסון בלתי תלויים עם קצבים* λ_1, λ_2 *(בהתאמה). התהליך* $N\left(t\right) = N_1\left(t\right) + N_2\left(t\right)$ *הינו תהליך פואסון גם כן עם קצב* $\lambda = \lambda_1 + \lambda_2$.

הוכחה. עבור כל אחד מהתהליכים פרקי זמן זרים הם בלתי תלויים לכן גם עבור הסכום פרקי זמן זרים הם בלתי תלויים. בנוסף, לפי משפט 7.3 מספר האירועים בפרק זמן כלשהו מתפלג פואסונית עם פרמטר $\lambda = \lambda_1 + \lambda_2$ ולכן $N\left(t\right)$ תהליך פואסון עם קצב λ. $\square$

תרגיל 8.7. מטוסים מהמזרח מגיעים לשדה התעופה לפי תהליך פואסון עם קצב של 10 מטוסים לשעה. מטוסים מהמערב מגיעים לשדה התעופה לפי תהליך פואסון עם קצב של 11 מטוסים לשעה, ללא תלות במטוסים מהמזרח. פקח הטיסה מגיע למשרד בשעה 8:00. מה הסיכוי שהמטוס הראשון שיגיע לשדה יהיה מטוס מהמזרח? מה הסיכוי שהמטוס השמיני שיגיע לשדה יהיה מטוס מהמזרח?

נתבונן במצב ההפוך, בו תהליך פואסון בודד מתפצל לשני תהליכים. למשל, מטוסים ממריאים לפי תהליך פואסון עם קצב של λ מטוסים בשעה וכל מטוס טס מזרחה בהסתברות $p \in [0,1]$ ללא תלות במטוסים האחרים. נסמן ב-$N\left(t\right)$ את תהליך המראת המטוסים וב-$N_E\left(t\right)$ את תהליך טיסת המטוסים מזרחה. נחשב את ההסתברות שמזמן 0 ועד t המריאו $k \in \mathbb{N}$ מטוסים מזרחה. לו היינו יודעים כמה מטוסים המריאו בסך הכל, הרי שמספר המטוסים שטס מזרחה היה מתפלג בינומית (כי המטוסים בלתי תלויים וטסים מזרחה בהסתברות קבועה). היות ואנחנו לא יודעים כמה מטוסים המריאו בסך הכל, נתנה על כל האפשרויות באמצעות משפט ההסתברות השלמה, כאשר n הוא מספר המטוסים שטסים מערבה:

$$\begin{aligned}Pr\left(N_E\left(t\right)=k\right)&=\sum_{n=0}^{\infty}Pr\left(N_E\left(t\right)=k|N\left(t\right)=n+k\right)\cdot Pr\left(N\left(t\right)=n+k\right)\\&=\sum_{n=0}^{\infty}\binom{n+k}{k}p^k\left(1-p\right)^n\cdot e^{-\lambda t}\frac{\lambda t^{n+k}}{(n+k)!}\\&=e^{-\lambda t}\sum_{n=0}^{\infty}\frac{(n+k)!}{n!k!}p^k\left(1-p\right)^n\cdot\frac{\left(\lambda t\right)^n\left(\lambda t\right)^k}{(n+k)!}\\&=e^{-\lambda t}\frac{\left(\lambda tp\right)^k}{k!}\sum_{n=0}^{\infty}\frac{\left(\lambda t\left(1-p\right)\right)^n}{n!}\end{aligned}$$

הטור שנותר הוא טור טיילור של אקספוננט (נוסחא **0.18**) ושווה ל-$\exp\left(\lambda t\left(1-p\right)\right)$. לכן ההסתברות שהמריאו k מטוסים מזרחה הינה

$$Pr\left(N_E\left(t\right)=k\right)=e^{-\lambda pt}\frac{\left(\lambda pt\right)^k}{k!} \qquad \textbf{(8.1)}$$

וזוהי בדיוק ההסתברות המתאימה למשתנה מקרי פואסוני עם פרמטר λpt. לכן התהליך המתאר את טיסת המטוסים מזרחה הוא תהליך פואסון עם קצב λp. באופן דומה, התהליך המתאר את טיסת המטוסים מערבה הוא תהליך פואסוני עם קצב $\lambda\left(1-p\right)$ ותהליכים אלו, למרות שמקורם בתהליך אחד, הם בלתי תלויים.

***משפט* 8.5** (פיצול פואסון). *יהי $N\left(t\right)$ תהליך פואסון כך שלכל אירוע יש סיכוי של p להיות מסוג מסויים ("מזרח") וסיכוי $1-p$ להיות מהסוג השני ("מערב"). אזי תהליך הופעת האירועים המזרחיים הוא תהליך פואסון עם קצב λp, תהליך הופעת האירועים המערביים הוא תהליך פואסון עם קצב $\lambda\left(1-p\right)$ ותהליכים אלו בלתי תלויים.*

הוכחה. העובדה שכל אחד מהתהליכים הוא תהליך פואסוני עם הקצב המתאים הוכחה לעיל. נותר רק להראות שהתהליכים בלתי תלויים. יהיו $m_1,m_2\in\mathbb{N}$. לפי נוסחא **8.1** ההסתברות שהיו m_1 אירועים פואסוניים מזרחיים בפרק זמן t הינה $e^{-\lambda pt}\frac{(\lambda pt)^{m_1}}{m_1!}$ וההסתברות שהיו m_2 אירועים פואסוניים מערביים הינה $e^{-\lambda t(1-p)}\frac{(\lambda t(1-p))^{m_2}}{m_2!}$. נחשב את ההסתברות של החיתוך ונראה שהיא שווה למכפלת הסתברויות אלו. לשם כך נפשט את מאורע החיתוך ובמקום לדון בחיתוך בין שני התהליכים נדון בחיתוך בין התהליך המזרחי לתהליך הכולל ממנו התהליכים התפצלו:

$$Pr\left(N_E\left(t\right)=m_1\cap N_W\left(t\right)=m_2\right)=Pr\left(N_E\left(t\right)=m_1\cap N\left(t\right)=m_1+m_2\right)$$

הסיבה שמאורע זה נוח יותר הוא שניתן לחשב את החיתוך על ידי חוק הכפל והתנייה על המספר הכולל של האירועים בתהליך המקורי: כאשר ידוע מספר האירועים בתהליך המקורי, מספר האירועים המזרחיים הינו בינומי:

$$\begin{aligned}Pr\left(N_E\left(t\right)=m_1\cap N\left(t\right)=m_1+m_2\right)&=Pr\left(N_E\left(t\right)=m_1|N\left(t\right)=m_1+m_2\right)Pr\left(N\left(t\right)=m_1+m_2\right)\\&=\binom{m_1+m_2}{m_1}p^{m_1}\left(1-p\right)^{m_2}e^{-\lambda t}\frac{\left(\lambda t\right)^{m_1+m_2}}{(m_1+m_2)!}\\&=\frac{\left(\lambda pt\right)^{m_1}}{m_1!}\frac{\left(\lambda t\left(1-p\right)\right)^{m_2}}{m_2!}e^{-\lambda t}\end{aligned}$$

כל שנותר הוא לפרק את האקספוננט לפי $e^{-\lambda t} = e^{-\lambda tp}e^{-\lambda t(1-p)}$ ולקבל שאכן הסתברות החיתוך שווה למכפלת ההסתברויות ולכן המאורעות בלתי תלויים והתהליכים בלתי תלויים. $\square$

תרגיל 8.8. אוטובוסים מגיעים לתחנה בתהליך פואסון עם קצב של 6 אוטובוסים לשעה. כל אוטובוס יכול להיות אדום או ירוק בהסתברות שווה ללא תלות באוטובוסים האחרים.

1. מה הסיכוי שהאוטובוס הראשון הוא אוטובוס ירוק?
2. האם המאורעות A– "הגיעו 4 אוטובוסים אדומים בין 12 ל-13" ו-B– "הגיעו 2 אוטובוסים ירוקים בין 12 ל-13" תלויים? הוכיחו!
3. בסך הכל הגיעו 6 אוטובוסים לתחנה בשעה מסויימת. מהו מקדם המתאם בין מספר האוטובוסים האדומים למספר האוטובוסים הירוקים שהגיעו בשעה זו?
4. מונית מגיעה לתחנה בזמן כלשהו המתפלג אחיד בין 12 ל-13 באופן בלתי תלוי בתנועת האוטובוסים. נוסע מגיע לתחנה בשעה 12 ועולה על כלי הרכב הראשון שמגיע. מה ההסתברות שהנוסע יעלה על המונית?

8.1 שאלות מסכמות

תרגיל 8.9. אוטובוסים מגיעים לתחנה לפי תהליך פואסון עם קצב של 2 אוטובוסים לשעה. נוסעים מגיעים לתחנה לפי תהליך פואסון בקצב של 4 נוסעים לשעה. שני התהליכים בלתי תלויים. הניחו כי זמן ההתחלה של התהליכים הינו 12:00 בצהריים.

1. נוסע הגיע לתחנה בשעה 13:00. מה הסיכוי כי יאלץ להמתין לאוטובוס לפחות חצי שעה?
2. מבקר בודק את זמני הגעת האוטובוסים החל מהשעה 12:00. מה הסיכוי שהאוטובוס השלישי יגיע לתחנה אחרי השעה 14:00?
3. נתון שמהשעה 12:00 ועד 13:00 הגיעו 4 אוטובוסים ונוסעים לתחנה (לא ידוע לפי איזו חלוקה). מה הסיכוי כי 3 מהם הגיעו בין 12:30 ל-13:00?
4. נוסע המגיע לתחנה ממתין לאוטובוס הראשון שיגיע לתחנה. מה תוחלת זמן ההמתנה של הנוסע שהגיע ראשון לתחנה?

תרגיל 8.10. בעיר מסוימת מספר התינוקות שנולדים בחודש מתפלג פואסונית עם תוחלת של 20 תינוקות בחודש (ללא תלות במספר הימים בחודש).

1. מה הסיכוי שבשלושת החודשים הראשונים של השנה ייולדו 55 תינוקות?
2. אם בשלושת החודשים הראשונים של השנה נולדו 55 תינוקות, מה הסיכוי ש-20 מתוכם נולדו במרץ?
3. מה הסיכוי שהתינוק הראשון נולד במחצית השנייה של ינואר?
4. בנים ובנות נולדים בהסתברות חצי בכל לידה באופן בלתי תלוי בלידות האחרות. האם המאורעות A– "בחודש ינואר נולדו 10 בנים" ו-B – "בחודש ינואר נולדו 10 בנות" תלויים?
5. ענו על הסעיף הקודם בהינתן שנולדו 20 ילדים בחודש ינואר.

תרגיל 8.11. בגלידריה שני מוכרים. הזמן הלוקח לכל מוכר לטפל בלקוח מתפלג לפי התפלגות מעריכית עם תוחלת של 3 דקות.

1. לקוח מגיע לגלידריה ומגלה ששני המוכרים תפוסים. כיצד מתפלג זמן ההמתנה שלו עד שהראשון יתפנה ויוכל לטפל בו?
2. מהי תוחלת מספר הלקוחות בהם מטפלים המוכרים במשך שעה, בהנחה שבכל פעם שמוכר מסיים לטפל בלקוח, מיד מגיע לקוח חדש?

תרגיל 8.12. סטודנט מקבל מיילים לפי תהליך פואסון עם קצב של 4 מיילים ביום. הסטודנט יוצא לטייל למשך שלושה ימים ולא בודק את המיילים שלו.

1. מה הסיכוי שיקבל 7 מיילים בפרק זמן זה?
2. מה הסיכוי שבכל יום הגיעו לפחות 2 מיילים?
3. מה הסיכוי ש-4 מיילים הגיעו ביום הראשון של הטיול אם בסך הכל הגיעו 4 מיילים?
4. מהי ההתפלגות של מספר הימים בהם הסטודנט קיבל לפחות מייל אחד (מתוך שלושת ימי הטיול)?

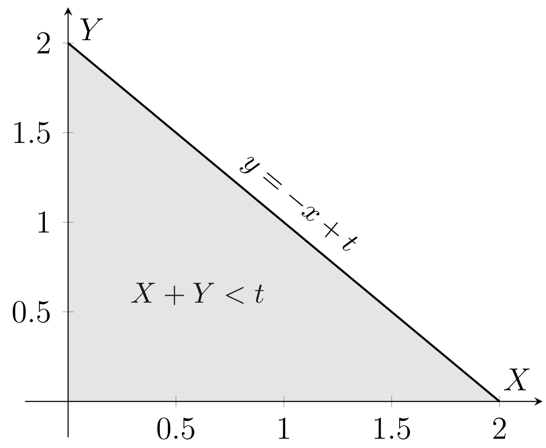
8.2 פתרונות לשאלות הפרק

תרגיל 8.1. נסמן ב-X את הזמן מתחילת התהליך ועד האירוע הראשון וב-Y את הזמן בין האירוע הראשון ועד האירוע השני. לפי התכונות של תהליך פואסון, שני משתנים מקריים אלו הם מעריכיים עם פרמטר λ ובלתי תלויים. נחשב את התפלגות הזמן עד האירוע הפואסוני השני מתחילת התהליך, כלומר את ההתפלגות של $Z = X+Y$. אופן החישוב זהה לשיטה שתוארה בסעיף 7.2.1. הצפיפות המשותפת של המשתנים המקריים הינה

$$f_{XY}(x,y) = \begin{cases} \lambda^2 e^{-\lambda x} e^{-\lambda y} & x,y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

ולכן פונקציית ההתפלגות המצטברת של Z בתחום $t \geq 0$ הינה

$$\begin{aligned} F_Z(t) &= Pr(Z \leq t) = Pr(X+Y \leq t) \\ &= \int_{x=0}^{t} \int_{y=0}^{t-x} \lambda^2 e^{-\lambda x} e^{-\lambda y} \, dy dx \\ &= 1 - e^{-\lambda t} - \lambda t e^{-\lambda t} \end{aligned}$$



וכמובן 0 אחרת. לסיכום, פונקציית ההתפלגות המצטברת של זמן המאורע השני הינה

$$F_Z(t) = \begin{cases} 1 - e^{-\lambda t} - \lambda t e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

התפלגות זו היא מקרה פרטי של התפלגות הנקראית התפלגות גאמא (Gamma Distribution). □

תרגיל 8.2. נסמן ב-X_t את המשתנה המקרי המייצג את מספר השיחות שהגיעו מרגע הפתיחה ועד זמן t (בשעות). זהו משתנה מקרי פואסוני המתפלג $X_t \sim Pois(20t)$.

1. מספר השיחות ב-10 הדקות הראשונות, $X_{1/6}$, מתפלג פואסוני עם פרמטר $20\cdot\frac{1}{6}=3\frac{1}{3}$ ולכן $Pr\left(X_{1/6}=2\right)=e^{-3\frac{1}{3}}\frac{\left(3\frac{1}{3}\right)^2}{2!}=0.198$.
מספר השיחות ב-5 הדקות האחרונות נקבע רק לפי משך פרק הזמן ולא לפי שעת ההתחלה שלו, ולכן מתפלג פואסוני עם פרמטר $\frac{20}{12}$. ההסתברות שהתקבלה שיחה אחת בלבד היא $e^{-\frac{20}{12}}\frac{\left(\frac{20}{12}\right)^1}{1!}=0.3148$.
2. לשם הפשטות נעבור ליחידות זמן של שיחות לדקה. ביחידות אלו, מדובר על תהליך פואסון עם פרמטר $\frac{1}{3}$ שיחות לדקה. פרק הזמן בין אירועים פואסוניים, T, מתפלג מעריכית עם פרמטר $\frac{1}{3}$ ולכן

$$Pr(1\le T\le 3)=F_T(3)-F_T(1)=1-e^{-1}-\left(1-e^{-1/3}\right)=0.349$$

3. נשתמש בתכונת חוסר הזיכרון של משתנה מקרי מעריכי ונקבל:

$$Pr\left(T\ge 3+1|T\ge 3\right)=Pr\left(T\ge 1\right)=1-F_T(1)=0.717$$

□

תרגיל 8.3. נסמן ב-X_t את המשתנה המקרי המייצג את מספר החלקיקים שנפלטו עד זמן t (בשניות).

1. מספר החלקיקים בשניה הראשונה מתפלג $X_1\sim Pois\left(\lambda\right)$. המאורע $X_1\ge 1$ כולל הרבה אפשרויות ולכן נעבור למאורע המשלים הכולל רק אפשרות אחת: $X_1=0$. ההסתברות המבוקשת הינה

$$Pr\left(X_1\ge 1\right)=1-Pr\left(X_1=0\right)=1-e^{-0.5}\frac{0.5^0}{0!}=0.393$$

2. מספר החלקיקים בחמש שניות, X_5, מתפלג $Pois(0.5\cdot 5)=Pois\left(2.5\right)$. נשתמש במשלים בכדי לחשב את ההסתברות המבוקשת:

$$Pr\left(X_5>3\right)=1-Pr\left(X_5=0\right)-Pr\left(X_5=1\right)-Pr\left(X_5=2\right)-Pr\left(X_5=3\right)=0.242$$

3. נסמן ב-X_1 את כמות החלקיקים שנפלטו בשניה הראשונה, X_5 את כמות החלקיקים שנפלטו ב-5 השניות הראשונות. נחשב את ההסתברות שמתוך 3 החלקיקים שידוע שנפלטו בחמש השניות הראשונות, בדיוק $k\in\{0,1,2,3\}$ נפלטו בשניה הראשונה:

$$\begin{aligned}Pr\left(X_1=k|X_5=3\right)&=\frac{Pr\left(X_1=k,X_5=3\right)}{Pr\left(X_5=3\right)}\overset{*}{=}\frac{Pr\left(X_1=k,X_5-X_1=3-k\right)}{Pr\left(X_5=3\right)}\\(**)&=\frac{Pr\left(X_1=k\right)Pr\left(X_5-X_1=3-k\right)}{Pr\left(X_5=3\right)}\end{aligned}$$

כאשר ב-* הפרדנו את תחום הזמן לשני פרקי זמן בלתי תלויים: השנייה הראשונה ו-4 השניות שלאחר מכן וב-** ניצלנו את חוסר התלות בין פרקי זמן אלו. נציב את שלושת ההסתברויות המבוקשות במשתנה המקרי הפואסוני:

$$\frac{e^{-0.5}\frac{0.5^k}{k!}\cdot e^{-4\cdot 0.5}\frac{(4\cdot 0.5)^{3-k}}{(3-k)!}}{e^{-5\cdot 0.5}\frac{(5\cdot 0.5)^3}{3!}}=\binom{3}{k}\frac{4^{3-k}}{5^3}=\binom{3}{k}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{3-k}$$

כלומר, בהינתן שנפלטו שלושה חלקיקים ב-5 השניות הראשונות, מספר החלקיקים בשניה הראשונה מתפלג $Bin\left(3,\frac{1}{5}\right)$. זוהי תוצאה פרטית של משפט 8.3.

□

תרגיל 8.4. נסמן ב-Y את הזמן עד האירוע הראשון וב-X את הזמן עד האירוע הראשון בהינתן שאירע לפני זמן T. נחשב את פונקציית ההתפלגות המצטברת של X עבור $t \in [0, T]$:

$$F_X(t) = Pr(X \leq t) = Pr(Y \leq t | Y \leq T) = \frac{Pr(Y \leq t \cap Y \leq T)}{Pr(Y \leq T)}$$

המאורע $Y \leq t$ מוכל במאורע $Y \leq T$ ולכן ההסתברות המבוקשת הינה

$$F_X(t) = \frac{Pr(Y \leq t)}{Pr(Y \leq T)} = \frac{1 - e^{-\lambda t}}{1 - e^{-\lambda T}}$$

ההתפלגות שהתקבלה אינה התפלגות אחידה כמו במשפט 8.2. הסיבה היא שבמשפט 8.2 מחשבים את ההתפלגות של הזמן של האירוע הפואסוני בהינתן שהיה אירוע אחד ואילו כאן ההתניה על כך שהאירוע הראשון היה לפני זמן T. במונחי הוכחת המשפט, ההתניה כאן היא על המאורע $N(T) \geq 1$ ולא על $N(T) = 1$, כמקודם. כלומר, אם ידוע שהיה אירוע אחד, אירוע זה יכול להיות בכל זמן בהסתברות שווה. לעומת זאת, אם ידוע שהיה לפחות אירוע אחד, האירוע הראשון יתפלג באופן לא שווה, אלא יהיה יותר סיכוי למצוא אותו בתחילת הקטע (כדי "להשאיר מקום" לשאר האירועים שהתרחשו). □

תרגיל 8.5. נסמן את מספר התקלות בקו מזמן 0 ועד t ב-$N(t)$.

1. המאורע המבוקש הינו $N(4) \leq 1$ וגם $N(12) - N(4) \geq 2$. מדובר על שני פרקי זמן זרים ולכן הם בלתי תלויים וניתן לחשב כל אחת מההסתברויות בנפרד ולכפול לקבלת החיתוך. מספר התקלות ב-4 השעות הראשונות מתפלג פואסונית עם פרמטר $4 \cdot 0.25 = 1$ ולכן

$$\begin{aligned} Pr(N(4) \leq 1) &= Pr(N(4) = 0) + Pr(N(4) = 1) \\ &= e^{-1} + e^{-1}\frac{1^1}{1!} = 0.736 \end{aligned}$$

בשמונה השעות הבאות מספר התקלות מתפלג פואסונית עם פרמטר $8 \cdot 0.25 = 2$ ולכן

$$\begin{aligned} Pr(N(12) - N(4) \geq 2) &= 1 - Pr(N(12) - N(4) < 2) \\ &= 1 - e^{-2} - e^{-2}\frac{2^1}{1!} = 0.594 \end{aligned}$$

ההסתברות של החיתוך הינה המכפלה, $0.736 \cdot 0.594 = 0.437$.

2. נסמן את הזמן עד התקלה הראשונה ב-T. ידוע ש-$T \geq 4$ ועל מנת שנשלים 10 שעות רצופות ללא תקלה נדרוש שיתקיים גם $T \geq 10$. מאחר ו-T מתפלג מעריכית ניתן להפעיל את תכונת חוסר הזיכרון:

$$\begin{aligned} Pr(T \geq 10 | T \geq 4) &= Pr(T \geq 6) = 1 - F_T(6) \\ &= 1 - \left(1 - e^{-6 \cdot 0.25}\right) = e^{-1.5} \end{aligned}$$

ניתן היה לפתור תרגיל זה גם על ידי חישוב ההסתברות של מאורע מהצורה $N(10) = 0 | N(4) = 0$.

3. ב-12 השעות הראשונות התרחשה תקלה אחת בדיוק. לפי משפט 8.2 זמן התרחשות האירוע מתפלג אחיד על פני 12 השעות ולכן הסיכוי שיתרחש בפרק זמן של שעתיים (השעה הראשונה והשעה האחרונה) הינו $\frac{2}{12}$.

4. בהמשך לסעיף הקודם, לכל אירוע יש סיכוי שווה להיות בכל אחת מהשעות. לכן ניתן לחשוב על הבעיה כעל בעיה של סידור כדורים בתאים: מעוניינים לסדר 20 כדורים ב-10 תאים, כך שבכל תא יהיו בדיוק 2 כדורים. מרחב המדגם הינו סידור הכדורים לתאים עם סדר ועם החזרה כך שגודלו 10^{20}. המאורע הוא סידור הכדורים לתאים כך שבכל תא יש שני כדורים. לכן עבור התא הראשון יש $\binom{20}{2}$ אפשרויות בחירה. עבור התא השני יש $\binom{18}{2}$ אפשרויות בחירה וכן הלאה. ההסתברות המבוקשת הינה

$$\frac{\binom{20}{2}\cdot\binom{18}{2}\cdot\ldots\cdot\binom{4}{2}\cdot\binom{2}{2}}{10^{20}}$$

□

תרגיל 8.6. נסמן ב-$N(t)$ את מספר הלקוחות במסעדה מהפתיחה ועד t שעות לאחריה.

1. ביום שלם המסעדה פתוחה 9 שעות. מאחר ומדובר על תהליך פואסון, מספר הלקוחות במהלך היום מתפלג לפי התפלגות פואסון עם פרמטר $9\cdot 4 = 36$ ולכן זוהי גם תוחלת מספר הלקוחות ביום שלם (נוסחא 6.26).

2. נסמן ב-X את מספר הלקוחות שהגיעו לארוחת בוקר. מדובר על תהליך פואסון ולכן $X\sim Pois\left(4\cdot 4.5 = 18\right)$. בנוסף, נסמן ב-$Z_i\sim U\left(20,50\right)$ את הטיפ של הלקוח ה-i בארוחת הבוקר. סך הטיפ שהושאר בארוחת הבוקר הינו

$$Z = Z_1+\ldots+Z_X$$

לו היינו יודעים את מספר הלקוחות, X, ניתן היה לחשב את סך הטיפים שהושארו בארוחת הבוקר על ידי תוחלת של סכום. היות ומספר הלקוחות לא ידוע, נתנה עליו ונשתמש במשפט התוחלת השלמה (נוסחא 7.14):

$$\begin{aligned}\mathbb{E}\left(Z\right) &= \mathbb{E}\left(\mathbb{E}\left(Z|X\right)\right) = \mathbb{E}\left(\sum_{i=1}^{X}\mathbb{E}\left(Z_i|X\right)\right)\\ &= \mathbb{E}\left(\sum_{i=1}^{X}35\right) = 35\mathbb{E}\left(X\right) = 35\cdot 18 = 630\end{aligned}$$

באותו אופן נחשב את תוחלת סך הטיפים של משמרת הצהריים: $60\cdot 18 = 1080$ ולכן תוחלת הטיפים לאורך כל היום הינה 1710.

3. נסמן ב-Y את מספר הלקוחות שהגיעו לאכול בין השעות 12:00-13:00. פרק זמן זה מורכב משני פרקי זמן אשר עליהם יש מידע שונה. פרק הזמן הראשון, 12:00-12:30, חלקי לזמן ארוחת הבוקר בו ידוע שהגיעו 2 לקוחות. לכל אחד מהם סיכוי של $\frac{0.5}{4.5}=\frac{1}{9}$ להגיע בין 12:00 ל-12:30 ולכן מספר הלקוחות מארוחת הבוקר שאכלו בפרק הזמן הראשון מתפלג $Bin\left(2,\frac{1}{9}\right)$.
פרק הזמן השני, 12:30-13:00, הינו פרק זמן של תהליך פואסון בו לא ידוע מספר הלקוחות ולכן מספר הלקוחות בפרק זמן זה מתפלג פואסונית עם פרמטר $0.5\cdot 4 = 2$. נחשב את פונקציית ההסתברות של Y עבור המקרים הראשונים ולאחר מכן נכליל.

המאורע $Y=0$ מתרחש כאשר אף לקוח לא הגיע בפרק זמן זה, כלומר שני הלקוחות שבאו לארוחת הבוקר הגיעו בזמן אחר ואף לקוח לא הגיע בחצי השעה הנוספת:

$$P_Y(0)=\left(\frac{8}{9}\right)^2\cdot e^{-2}$$

המאורע $Y=1$ יכול להתרחש באחת משתי צורות זרות: או שאחד מהלקוחות של ארוחת הבוקר הגיע בפרק זמן זה ואף אחד לא הגיע בחצי השעה הנוספת או ששני הלקוחות של ארוחת הבוקר לא הגיעו בפרק זמן זה ולקוח אחד הגיע בחצי השעה הנוספת. לכן

$$P_Y(1)=\binom{2}{1}\frac{1}{9}\cdot\frac{8}{9}\cdot e^{-2}+\left(\frac{8}{9}\right)^2e^{-2}\frac{2^1}{1!}$$

עבור $k\geq 2$ ייתכנו **3** אפשרויות: $n\in\{0,1,2\}$ מתוך הלקוחות של ארוחת הבוקר הגיעו בפרק זמן זה ו-$k-n$ הנותרים הגיעו בין **12:30** ל-**13:00**. לכן ההסתברות המבוקשת הינה

$$P_Y(k)=\sum_{n=0}^{2}\binom{2}{n}\left(\frac{1}{9}\right)^n\left(\frac{8}{9}\right)^{2-n}e^{-2}\frac{2^{k-n}}{(k-n)!}$$

□

תרגיל 8.7. הזמן עד הופעת המטוס הראשון מהמזרח מתפלג מעריכית עם פרמטר **10** והזמן עד הופעת המטוס הראשון מהמערב מתפלג מעריכית עם פרמטר **11**. זהו מקרה פרטי של תחרות בין מעריכיים וכפי שמצאנו בתרגיל **7.27**, הסיכוי שהמטוס הראשון מגיע מהמזרח הינו $\frac{10}{21}=\frac{10}{10+11}$.

באופן זהה, מתכונת חוסר הזיכרון, הזמן בין הגעת המטוס השביעי (בסך הכל) ועד הגעת המטוס הבא מהמזרח מתפלג מעריכית עם פרמטר **10** והזמן בין הגעת המטוס השביעי (בסך הכל) ועד הגעת המטוס הבא מהמערב מתפלג מעריכית עם פרמטר **11**. לכן גם כאן, הסיכוי שהמטוס הבא יגיע מהמזרח הינו $\frac{10}{10+11}$.

ננסח מחדש את טענת התרגיל בצורה כללית: כאשר שני תהליכי פואסון בלתי תלויים עם פרמטרים λ_1,λ_2 מתאחדים, הסיכוי שהאירוע הראשון יהיה אירוע מהסוג של התהליך הראשון הינו $\frac{\lambda_1}{\lambda_1+\lambda_2}$. יתרה מכך, זהו גם הסיכוי של כל אחד מהאירועים בתהליך המאוחד להיות אירוע מהסוג של התהליך הראשון. □

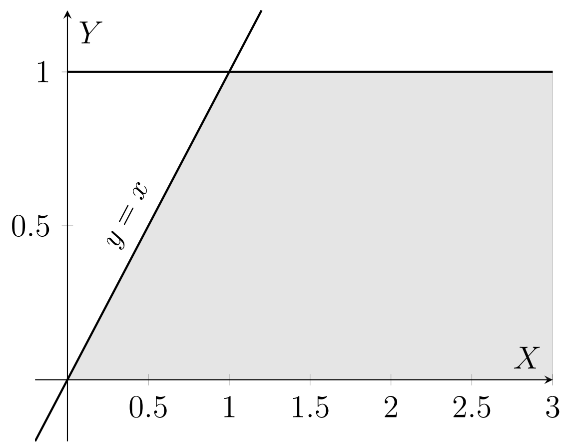
תרגיל 8.8. **1.** לכל אוטובוס סיכוי שווה להיות אדום או ירוק, בפרט לאוטובוס הראשון. לכן הסיכוי שהאוטובוס הראשון הוא ירוק הינו $\frac{1}{2}$.

2. התהליך המתואר הוא פיצול פואסון בו תהליך הגעת האוטובוסים מתפצל לשני תהליכים: הגעת אוטובוסים אדומים והגעת אוטובוסים ירוקים. המאורע A מתייחס לתהליך האדומים והמאורע B לתהליך האוטובוסים הירוקים. שני התהליכים בלתי תלויים ולכן גם המאורעות.

3. אם X הוא מספר האוטובוסים האדומים ו-Y הוא מספר האוטובוסים הירוקים, הרי שמתקיים $X+Y=6$, כלומר בין המשתנים המקריים יש קשר לינארי מושלם יורד ומקדם המתאם הינו $\rho_{XY}=-1$.

4. נסמן ב-X את הזמן עד הגעת האוטובוס הראשון וב-Y את הזמן עד הגעת המונית. לפי הנתונים, $X\sim Exp(6)$ ו-$Y\sim U(0,1)$ (הכל ביחידות של שעה). הנוסע יעלה על המונית אם מתקיים $Y<X$. נחשב את ההסתברות של מאורע זה, תוך שימוש בכך שמשתנים מקריים אלו בלתי תלויים ולכן הצפיפות המשותפת היא מכפלת הצפיפויות השוליות:

$$Pr(Y<X)=\int_{y=0}^{1}\int_{x=y}^{\infty}6e^{-6x}\,dxdy=0.166$$



□

תרגיל 8.9. **1.** נסמן את זמן ההמתנה לאוטובוס ב-T. זמן ההמתנה לאוטובוס המגיע לפי תהליך פואסוני עם מקדם 2 הוא מעריכי עם מקדם 2 ולכן

$$Pr(T>0.5)=1-Pr(T\le 0.5)=1-F_T(0.5)=1-\left(1-e^{-0.5\cdot 2}\right)=e^{-1}$$

2. נסמן ב-X_2 את מספר האוטובוסים שמגיעים בין השעות 12 ל-14. מדובר על פרק זמן של שעתיים בתהליך פואסון עם קצב של 2 ולכן $X_2\sim Pois(4)$. האוטובוס השלישי יגיע אחרי 14:00 אם ורק אם בשעתיים הראשונות הגיעו פחות מ-3 אוטובוסים, כלומר $X_2<3$. לפיכך, ההסתברות המבוקשת הינה

$$\begin{aligned}Pr(X_2<3)&=Pr(X_2=0)+Pr(X_2=1)+Pr(X_2=2)\\&=e^{-4}\frac{4^0}{0!}+e^{-4}\frac{4^1}{1!}+e^{-4}\frac{4^2}{2!}\end{aligned}$$

3. בהינתן שאירע מופע פואסוני, הזמן המדוייק בו הוא התרחש מתפלג אחיד (משפט 8.2). מאחר והמופעים בלתי תלויים, כל אחד יהיה בהסתברות שווה בין 12:30 ל-13:00 ומספר המופעים בפרק זמן זה מתפלג $Bin(4,0.5)$ (ראו משפט 8.9). ההסתברות המבוקשת היא $\binom{4}{3}\frac{1}{2^3}\cdot\frac{1}{2}=0.25$.

4. מתכונת חוסר הזיכרון, מרגע התחלת ההמתנה הזמן מתפלג מעריכית עם פרמטר 2 ולכן תוחלת זמן ההמתנה היא $\frac{1}{2}$שעה.

□

תרגיל 8.10. נסמן את תהליך הפואסון של לידת התינוקת ב-X_t. זהו תהליך עם קצב של 20 תינוקות בחודש.

1. נשים לב כי X_3 מתפלג פואסונית עם פרמטר 60, לכן מחישוב ישיר

$$Pr(X_3=55)\quad=\quad e^{-60}\frac{60^{55}}{55!}$$

2. לפי משפט 8.9 מספר הלידות בחודש מרץ מתפלג $Bin\left(55,\frac{1}{3}\right)$. מכאן נובע שהסיכוי ל-20 לידות בחודש מרץ הוא $\binom{55}{20}\left(\frac{1}{3}\right)^{20}\left(\frac{2}{3}\right)^{35}$.

3. נסמן את הזמן עד ללידה הראשונה ב-Y. ידוע כי $Y \sim Exp\left(20\right)$. לכן

$$\begin{aligned} Pr\left(0.5<Y<1\right) &= F_Y\left(1\right)-F_Y\left(\frac{1}{2}\right)=1-e^{-20}-\left(1-e^{-\frac{1}{2}\cdot 20}\right) \\ &= e^{-10}-e^{-20} \end{aligned}$$

4. מדובר בפיצול פואסונים ולכן המאורעות בלתי תלויים. נחשב את כל הגדלים בהתנייה על מספר הלידות הכללי בחודש ינואר ונראה זאת ישירות:

$$\begin{aligned} Pr\left(A\right) &= \sum_{k=10}^{\infty} Pr\left(A|X_1=k\right)Pr\left(X_1=k\right) \\ &= \sum_{k=10}^{\infty}\left[\binom{k}{10}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^{k-10}e^{-20}\frac{20^k}{k!}\right] \\ &= e^{-20}\sum_{k=10}^{\infty}\left[\frac{k!}{10!\left(k-10\right)!}\left(\frac{1}{2}\right)^{k}\frac{20^k}{k!}\right] \\ &= \frac{e^{-20}}{10!}\sum_{k=10}^{\infty}\frac{10^k}{\left(k-10\right)!} \\ &= \frac{e^{-20}10^{10}}{10!}\sum_{m=0}^{\infty}\frac{10^m}{m!}=\frac{e^{-20}10^{10}}{10!}e^{10}=e^{-10}\frac{10^{10}}{10!} \end{aligned}$$

משיקולי סימטריה בין בנים ובנות, נקבל תוצאה דומה עבור מאורע B. נותר לחשב את הסיכוי של מאורע $A\cap B$:

$$\begin{aligned} Pr\left(A\cap B\right) &= Pr\left(A\cap B|X_1=20\right)Pr\left(X_1=20\right) \\ &= \binom{20}{10}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^{20-10}e^{-20}\frac{20^{20}}{20!} \\ &= \frac{e^{-20}20^{20}}{10!10!2^{20}}=\frac{e^{-20}10^{20}}{10!10!} \\ &= \left[e^{-10}\frac{10^{10}}{10!}\right]^2=Pr\left(A\right)Pr\left(B\right) \end{aligned}$$

ולכן הם בלתי תלויים.

5. בהינתן שנולדו בדיוק 20 תינוקות בחודש ינואר, יש בין מספר הבנים ומספר הבנות קשר לינארי דטרמיניסטי: אם נולדו בחודש ינואר 10 בנים בהכרח נולדו גם 10 בנות ואם לא נולדו 10 בנים בהכרח גם לא נולדו 10 בנות. המאורעות תלויים.

□

תרגיל 8.11. נסמן ב-X את הזמן עד שהמוכר הראשון מתפנה וב-Y את הזמן עד שהמוכר השני מתפנה. שניהם משתנים מקריים מעריכיים עם פרמטר $\frac{1}{3}$.

1. ניתן לפתור את השאלה בשתי דרכים. דרך אחת היא לסמן $Z = \min\{X,Y\}$ ולחשב את ההתפלגות של Z בדומה לנעשה בסעיף 7.2.1. זו הייתה הדרך היחידה לפעול לו ההתפלגויות של X ו-Y היו אחרות, אך היות ומדובר על משתנים מקרים מעריכיים, ניתן לפשט משמעותית את החישוב באמצעות התכונות של תהליך פואסון.
היות והזמן עד שהמוכר מתפנה מתפלג מעריכית, הרי שהאירועים כאן הם "ההתפנויות" של המוכר, המתרחשות לפי תהליך פואסון עם קצב $\frac{1}{3}$ (התפנויות לדקה), בהנחה שכל מוכר שמתפנה מיד מתחיל לטפל בלקוח נוסף. מאחר ומדובר על שני מוכרים, מדובר על סכום של תהליכי פואסון בלתי תלויים ולכן ההתפנויות מתרחשות עתה בקצב של $\frac{2}{3}$ לדקה, כלומר הזמן עד ההתפנות הראשונה בתהליך הסכום מתפלגת מעריכית עם פרמטר $\frac{2}{3}$.
2. כאמור בסעיף הקודם, התהליך המתרחש בגלידריה הוא תהליך פואסון עם קצב של $\frac{2}{3}$ לדקה, כאשר כל "התפנות" מסמנת סיום טיפול בלקוח. מספר ההתפנויות בשעה מתפלג פואסונית עם פרמטר $\frac{2}{3}\cdot 60 = 40$ ולכן תוחלת מספר הלקוחות שטופלו בגלידריה בשעה הוא 40.
ניתן לקבל תוצאה זו גם לפי תוחלת של סכום. כל מוכר מטפל בממוצע בלקוח ב-3 דקות, לכן מטפל בשעה בממוצע ב-20 לקוחות ולכן שניהם מטפלים יחד בממוצע ב-40 לקוחות.

□

תרגיל 8.12. נסמן ב-$N(t)$ את מספר המיילים שהסטודנט מקבל במהלך t הימים הראשונים להיעדרו. זהו תהליך פואסון עם קצב 4.

1. מספר המיילים בשלושה ימים מתפלג פואסונית עם פרמטר $3\cdot 4 = 12$ ולכן הסיכוי שיקבל בפרק זמן זה 7 מיילים הינו
$$e^{-12}\frac{12^7}{7!}$$
2. הסיכוי שביום מסויים יגיעו לפחות 2 מיילים הינו המשלים של המאורע שהגיע מייל אחד או לא הגיעו מיילים כלל, כלומר $1 - e^{-4} - e^{-4}\frac{4^1}{1!} = 0.908$. הימים בלתי תלויים שכן אלו פרקי זמן זרים ולכן הסיכוי שבכל יום הגיעו לפחות 2 מיילים הנו $0.908^3 = 0.749$.
3. בהינתן שהגיעו 4 מיילים, מספר המיילים שהגיעו ביום הראשון של הטיול מפלג בינומית $Bin\left(4,\frac{1}{3}\right)$. לכן ההסתברות שכל המיילים הגיעו ביום הראשון הינה $\left(\frac{1}{3}\right)^4$.
4. הסיכוי לקבל מייל אחד לפחות הינו $1 - e^{-4} = 0.981$ והימים בלתי תלויים, לכן התפלגות מספר הימים בהם הגיע לפחות מייל אחד הינה $Bin(3, 0.981)$.

□

9 שאלות חזרה וסיכום

9.1 פתרונות לשאלות הפרק

תרגיל . □