88-165 – מבוא להסתברות וסטטיסטיקה
מערכי תרגול

תיכוניסטים, סמסטר ב', תשפ"ב

תוכן העניינים

1 תרגול ראשון

1.1 חזרה מתורת הקבוצות

הגדרה 1.1. יהיו A, B קבוצות המוכלות בקבוצה אוניברסלית Ω.

א. **האיחוד** שלהן, המסומן $A \cup B$, הוא קבוצת האיברים הנמצאים ב-A או ב-B. כלומר

$$A \cup B = \{x \mid x \in A \vee x \in B\}.$$

ב. **החיתוך** שלהן, המסומן $A \cap B$, הוא קבוצת האיברים הנמצאים ב-A וגם ב-B. כלומר

$$A \cap B = \{x \mid x \in A \wedge x \in B\}.$$

ג. **ההפרש** שלהן, המסומן $A \setminus B$, הוא קבוצת האיברים הנמצאים ב-A ואינם נמצאים ב-B. כלומר

$$A \setminus B = \{x \mid x \in A, x \notin B\} = \{x \in A \mid x \notin B\}$$

ד. **המשלים** של A ב-Ω, המסומן A^c, הוא $\Omega \setminus A$.

ה. **ההפרש הסימטרי** של A ו-B, המסומן $A \triangle B$, הוא קבוצת האיברים הנמצאים בדיוק באחת הקבוצות A ו-B. כלומר

$$A \triangle B = \{x \mid (x \in A \wedge x \notin B) \vee (x \notin A \wedge x \in B)\}.$$

ניתן להכליל את הגדרות האיחוד והחיתוך גם למשפחות של תת־קבוצות: יהיו $\{A_i\}_{i \in I}$ תת־קבוצות של Ω עם $I \neq \varnothing$. אז

$$\bigcup_{i \in I} A_i = \{x \mid \exists i \in I : x \in A_i\},$$
$$\bigcap_{i \in I} A_i = \{x \mid \forall i \in I : x \in A_i\}.$$

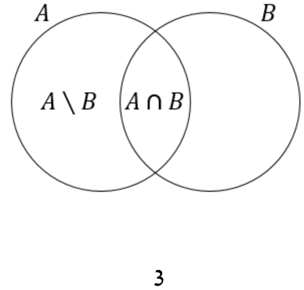
הערה. $A \setminus B = A \cap B^c$.

טענה 1.2 (דה-מורגן). לכל אוסף קבוצות $\{A_i\}_{i \in I}$,

$$\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c, \qquad \left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c.$$

תרגיל 1.3. הוכיחו כי לכל שתי קבוצות A, B,

$$A \setminus (A \setminus B) = A \cap B.$$



הוכחה. אפשר באמצעות הכלה דו-כיוונית, ואפשר גם באמצעות דה-מורגן:

$A\setminus(A\setminus B) = A\cap(A\cap B^c)^c = A\cap(A^c\cup B) = (A\cap A^c)\cup(A\cap B) = \varnothing\cup(A\cap B) = A\cap B$.

□

תרגיל 1.4. הוכיחו שלכל שתי קבוצות A, B,

$$(A\cup B)\cap(A\cap B)^c = (A\setminus B)\cup(B\setminus A)\,.$$

הוכחה. נשים לב כי

$$\begin{aligned}(A\cup B)\cap(A\cap B)^c &= (A\cup B)\cap(A^c\cup B^c) = ((A\cup B)\cap A^c)\cup((A\cup B)\cap B^c) = \\ &= (A\cap A^c)\cup(B\cap A^c)\cup(A\cap B^c)\cup(B\cap B^c) = \\ &= \varnothing\cup(B\setminus A)\cup(A\setminus B)\cup\varnothing = (A\setminus B)\cup(B\setminus A)\,.\end{aligned}$$

□

הגדרה 1.5.

א. אומרים ש-A ו-B **זרות**, אם $A\cap B=\varnothing$.

ב. אומרים ש-$\{A_i\}_{i\in I}$ **זרות בזוגות**, אם לכל $i,j\in I$ כך ש-$j\neq i$ מתקיים $A_i\cap A_j=\varnothing$. במקרה זה מסמנים את האיחוד שלהן $\dot\bigcup_{i\in I} A_i$.

דוגמה 1.6. ההבדל בין קבוצות זרות בזוגות לקבוצות זרות: הקבוצות $\{1,2\}$, $\{1,3\}$ ו-$\{2,3\}$ הן זרות, אך לא זרות בזוגות.

תרגיל 1.7. נניח ש-A ו-B קבוצות זרות. האם A^c ו-B^c זרות גם כן?

פתרון. לא! למשל, ניקח $\Omega=\{1,2,3\}$, $A=\{1\}$ ו-$B=\{2\}$. אז $A\cap B=\varnothing$, אבל $A^c\cap B^c=\{2,3\}\cap\{1,3\}=\{3\}$.

תרגיל 1.8 (תרגיל שימושי)**.** הוכיחו כי לכל סדרת קבוצות $A_1, A_2, \ldots$ מתקיים

$$\bigcup_{n=1}^{\infty} A_n = \dot\bigcup_{n=1}^{\infty}\left(A_n\setminus(A_1\cup\cdots\cup A_{n-1})\right).$$

(שימו לב: צריך להוכיח גם שהאיחוד באגף ימין הוא איחוד זר.)

הוכחה. נסמן $B_n = A_n\setminus(A_1\cup\cdots\cup A_{n-1})$. נוכיח שהאיחוד באגף ימין הוא איחוד זר. אכן, אם $m<n$ אז $B_m = A_m\setminus(A_1\cup\cdots\cup A_{m-1})\subseteq A_m$, ואילו

$$B_n = A_n\setminus(A_1\cup\cdots\cup A_{n-1})\subseteq A_n\setminus A_m = A_n\cap A_m^c\subseteq A_m^c.$$

לכן $B_m\cap B_n=\varnothing$, והאיחוד הוא אכן איחוד זר.

כעת נראה שוויון בין שני האגפים. לכל m מתקיים $B_m\subseteq A_m\subseteq\bigcup_{n=1}^{\infty}A_n$, ולכן $\bigcup_{n=1}^{\infty}A_n\supseteq\dot\bigcup_{m=1}^{\infty}B_m$. להכלה $\subseteq$, יהי $x\in\bigcup_{n=1}^{\infty}A_n$. ניקח את ה-$n$ המינימלי שעבורו $x\in A_n$. לכן $x\notin A_1\cup\cdots\cup A_{n-1}$, כלומר $x\in A_n\setminus(A_1\cup\cdots\cup A_{n-1})=B_n$. מכאן ש-$x\in\dot\bigcup_{n=1}^{\infty}B_n$, כנדרש. □

הגדרה 1.9. **קבוצת החזקה של** A היא קבוצת כל תת־הקבוצות של A. בשפה מתמטית,

$$P(A) = 2^A = \{B \mid B\subseteq A\}.$$

טענה 1.10. אם $|A|=n$, אז $|P(A)|=2^n$.

1.2 מבוא לקומבינטוריקה

המטרה הבסיסית של קומבינטוריקה היא ספירת איברים בקבוצות סופיות, כלומר לכל קבוצה סופית A לחשב את $|A|$. בחלק הראשון של הקורס נציג מספר שיטות שעוזרות לחישוב גדלים כאלו.

טענה 1.11 (עיקרון הסכום).

א. לכל שתי קבוצות A ו-B, $|A \cup B| = |A| + |B| - |A \cap B|$.

ב. בפרט: אם A ו-B זרות, נקבל $|A \uplus B| = |A| + |B|$.

ג. ל-n קבוצות: אם $A_1, \ldots, A_n$ קבוצות זרות בזוגות, אז $|\biguplus_{i=1}^n A_i| = \sum_{i=1}^n |A_i|$.

<u>משמעות קומבינטורית:</u> אם A, B זרות, אז מספר האפשרויות שיתרחש $A \cup B$ הוא מספר האפשרויות שיתרחש A ועוד מספר האפשרויות שיתרחש B.

דוגמה 1.12. מספר האפשרויות לבחור תלמיד או תלמידה מתוך כיתה של 15 בנים ו-12 בנות הוא 27.

טענה 1.13 (עיקרון המכפלה).

א. לכל שתי קבוצות A ו-B, $|A \times B| = |A| \times |B|$.

ב. ל-n קבוצות: אם $A_1, \ldots, A_n$ קבוצות, אז $|A_1 \times \cdots \times A_n| = |A_1| \cdot \cdots \cdot |A_n|$.

<u>משמעות קומבינטורית:</u> מספר תוצאותיו של ניסוי בן r שלבים, כשלשלב ה-i יש n_i תוצאות אפשריות, הוא $n_1 \cdot \cdots \cdot n_r$.

דוגמה 1.14. מספר האפשרויות לבחור תלמיד ותלמידה מתוך כיתה של 15 בנים ו-12 בנות הוא $15 \cdot 12 = 180$.

תרגיל 1.15. מזהה רישיון רכב מורכב משתי אותיות בעברית ולאחר מכן שלוש ספרות. כמה מזהי רישיונות רכב ניתן ליצור?

פתרון. $22^2 \cdot 10^3 = 484,000$.

תרגיל 1.16.

א. בכמה דרכים ניתן לסדר כדורים אדומים, ירוקים וכחולים בשורה של n כדורים?

ב. בכמה סידורים כנ"ל אין שני כדורים סמוכים עם אותו הצבע?

פתרון.

א. לכל כדור בשורה צריך לבחור את הצבע שלו. זו בחירה של n מתוך 3, עם החזרה (כי יכולים להיות כמה כדורים באותו צבע) ועם חשיבות לסדר (כי הסדר של הצבעים בשורה משנה), ולכן כמות הסידורים היא 3^n.

ב. נפעל לפי עיקרון המכפלה. לכדור הימני ביותר יש 3 אפשרויות. לאחר מכן, לכל כדור יש 2 אפשרויות – הרי הוא לא יכול להיות זהה לכדור שהיה מימינו. לכן כמות האפשרויות הכוללת היא $3 \cdot 2^{n-1}$.

הגדרה 1.17. **תמורה על קבוצה** X היא פונקציה חח"ע ועל $\sigma : X \to X$.

תרגיל 1.18. כמה תמורות שונות יש על קבוצה X בגודל n? (באופן שקול: בכמה דרכים ניתן לסדר n אנשים בשורה?)

פתרון. נסמן את הקבוצה $X=\{x_1,\ldots,x_n\}$. עבור $\sigma(x_1)$ יש n אפשרויות; הוא יכול להיות כל איבר של X. עבור $\sigma(x_2)$ יש $n-1$ אפשרויות, שכן הוא לא יכול להיות שווה ל-$\sigma(x_1)$; ניתן להמשיך כך ולקבל של-$\sigma(x_k)$ יש $n-(k-1)$ אפשרויות. בסך הכל, כמות התמורות על קבוצה בגודל n היא $n!=n\cdot(n-1)\cdot\cdots\cdot 1$.

תרגיל 1.19. בכמה דרכים ניתן לסדר n אנשים בשורה, אם יש זוג אנשים שרוצים להיות אחד ליד השני?

פתרון. נחשוב על שני האנשים כעל "בלוק" אחד. נסדר את הבלוק ואת שאר האנשים בשורה, ואז נבחר את הסידור בין שני האנשים בזוג. בשלב הראשון צריך לסדר $n-1$ עצמים בשורה (שהם $n-2$ אנשים ובלוק אחד), ולכך יש $(n-1)!$ אפשרויות; בשלב השני צריך לסדר שני אנשים, ולכך יש $2!=2$ אפשרויות. בסך הכל כמות הדרכים היא $2(n-1)!$.

תרגיל 1.20. בכמה דרכים ניתן לסדר n אנשים במעגל?

פתרון. בתרגיל הקודם בעצם דיברנו על סידור של n אנשים בשורה, וראינו שמדובר ב-$n!$ אפשרויות. כל סידור בשורה אפשר כמובן לתרגם לסידור במעגל, אם נמספר את המקומות של המעגל $1,\ldots,n$. אבל כיוון שלא אכפת לנו מאיפה מתחילים, בעצם ספרנו כל סידור n פעמים, כתלות בנקודת ההתחלה. לכן כמות האפשרויות הכוללת היא $\frac{1}{n}\cdot n!=(n-1)!$.

1.3 בחירה של k מתוך n

	עם החזרה	ללא החזרה
עם חשיבות לסדר	n^k	$P(n,k)={}^nP_k={}_nP_k=\frac{n!}{(n-k)!}$
ללא חשיבות לסדר	$\binom{n+k-1}{n-1}=\binom{n+k-1}{k}$	$C(n,k)={}^nC_k={}_nC_k=\binom{n}{k}=\frac{n!}{k!(n-k)!}$

עם החזרה – האם מותר לבחור את אותו האיבר פעמיים?

עם חשיבות לסדר – האם אנחנו רוצים לספור סידורים של אותם איברים אך בסדר שונה כסידורים נפרדים, או לא?

דוגמה 1.21. נניח שמסדרים k כדורים ב-n תאים. זו בחירה של k מתוך n: לכל כדור אנו בוחרים את התא שבו הוא יונח. נתרגם מה אומר כל אחד מהמונחים שאנו מכירים.

- ללא החזרה – אסור לבחור לשני כדורים את אותו התא; כלומר, בכל תא יכול להיות לכל היותר כדור אחד.
- עם החזרה – מותר לבחור לשני כדורים את אותו התא; כלומר, ייתכנו מספר כדורים באותו התא.
- ללא חשיבות לסדר – אין חשיבות לסדר שבו אנו בוחרים את התאים. זה אומר שאם בחרנו לכדור הראשון את תא 1 ולכדור השני את תא 2, לא נראה הבדל לעומת אם בחרנו לכדור הראשון את תא 2 ולכדור השני את תא 1. זה קורה כאשר הכדורים זהים זה לזה.
- עם חשיבות לסדר – יש חשיבות לסדר שבו אנו בוחרים את התאים. כלומר, נראה הבדל אם נשים קודם את הכדור הראשון בתא 1 ואז את הכדור השני בתא 2, לעומת האפשרות ההפוכה. זה קורה כאשר הכדורים שונים זה מזה.

כעת נעבור למספר תרגילים.

תרגיל 1.22. כמה פונקציות יש מהקבוצה $\{1,2,3\}$ לקבוצה $\{1,2,3,4,5,6,7\}$?

פתרון. רוצים לבחור את $f(i)$ לכל $i\in\{1,2,3\}$. זו בחירה של 3 מתוך 7, עם חשיבות לסדר ועם החזרה. לכן התשובה היא 7^3.

תרגיל 1.23. מהו מספר האפשרויות לבחור 5 ילדים ל-5 תפקידים שונים בהצגה, מתוך קבוצה של 12 ילדים?

פתרון. רוצים לבחור 5 תלמידים מתוך 12; יש חשיבות לסדר (כי התפקידים שונים), ורוצים בחירה ללא החזרה (כי כל תפקיד צריך שחקן אחר). אז בסך הכל יש $\frac{12!}{7!}=95,040$ דרכים לעשות זאת.

תרגיל 1.24. בכמה דרכים ניתן לבחור ארבעה מתנדבים להדגמת קסם במופע של 100 אנשים בקהל?

פתרון. זו בחירה ללא החזרה וללא חשיבות לסדר, ולכן התשובה היא $\binom{100}{4}$.

תרגיל 1.25. כידוע, כדי לזכות במשחק לוטו טיפוסי בישראל צריך לנחש נכונה את ששת המספרים העולים בגורל (לא ייצא אותו מספר פעמיים, ואין חשיבות לסדר שבו הם ייצאו), וכן לנחש את "המספר החזק" שהוא מספר בין 1 ל-7. בהנחה שאני ממלא טפסי לוטו בצורה אקראית לחלוטין, מה הסיכוי שלי לזכות בלוטו?

פתרון. בעצם השאלה היא כמה טפסי לוטו אפשריים יש. עבור המספרים הרגילים, צריך לבחור 6 מתוך 37, בלי החזרה וללא חשיבות לסדר, ולכן מדובר על $\binom{37}{6}$ אפשרויות; לכל אחת מהן יש את המספר החזק, שלו 7 אפשרויות. לכן כמות הלוחות האפשריים היא

$$\binom{37}{6}\cdot 7=16273488,$$

והסיכוי לזכות בלוטו הוא $\frac{1}{16273488}$.

תרגיל 1.26. מגרילים מספר בעל 5 ספרות מהספרות $1,\dots,9$. מה הסיכוי שמופיעה בו הספרה 3?

פתרון. ראשית, נשים לב שכמות המספרים הכוללת בעלי 5 ספרות מהספרות האלו היא 9^5. כדי לחשב את הסיכוי שמופיעה במספר הספרה 3, נחשב את כמות המספרים הנ"ל.

<u>פתרון שגוי:</u> יש 5 מקומות שבהם יכולה להופיע הספרה 3, ואחרי שבחרנו את המקום שלה יש 9^4 אפשרויות למלא את שאר הספרות. בסך הכל $5\cdot 9^4$ אפשרויות. (הבעיה: אם יש במספר יותר מפעם אחת את הספרה 3, ספרנו אותו מספר פעמים).

<u>פתרון נכון:</u> נתבונן במשלים. כמות המספרים בעלי 5 הספרות מהספרות $1,\dots,9$ היא 9^5; כמות המספרים האלו שבהם לא מופיעה הספרה 3 היא 8^5. בסך הכל כמות המספרים שבהם כן מופיעה הספרה 3 היא $9^5-8^5=26281$, והסיכוי הוא $\frac{26281}{9^5}\approx 0.445$.

תרגיל 1.27. שולפים 5 קלפים מתוך חבילה של 30 קלפים – 10 אדומים, 10 כחולים ו-10 ירוקים, כששני קלפים בעלי אותו צבע נחשבים זהים. חמשת הקלפים נשלפים בבת אחת, ולכן אין חשיבות לסדר. כמה אפשרויות יש לחמישיית הקלפים הנשלפת?

פתרון. זו בחירה של 5 מתוך 3 עם החזרה (כי מותר לשלוף את אותו הצבע מספר פעמים) ובלי חשיבות לסדר (כי הקלפים נשלפים ביחד). לכן בסך הכל יש $\binom{5+3-1}{3-1}=\binom{7}{2}=21$ דרכים.

תרגיל 1.28. בבוקר יום בהיר עלו על השאטל של אוניברסיטת בר-אילן 6 סטודנטים. השאטל עוצר ב-16 תחנות, ונהג השאטל התבקש לתעד כמה נוסעים ירדו בכל תחנה. הנוסעים אינם תלויים זה בזה. כמה תיעודים שונים יכול נהג השאטל להציג בסוף הנסיעה?

פתרון. זו בחירה של 6 מתוך 16, עם החזרה (כמה סטודנטים יכולים לרדת באותה תחנה) וללא חשיבות לסדר (הסדר שבו הם יורדים לא משנה, אלא רק הכמות שיורדת בכל תחנה). לכן כמות התיעודים האפשריים היא $\binom{6+16-1}{16-1} = \binom{21}{15} = 54,264$.

טענה 1.29 (אם יש זמן). כמות הפתרונות של המשוואה $x_1 + \cdots + x_n = k$, כאשר כל $x_i \geq 0$ הוא מספר שלם, היא $\binom{n+k-1}{n-1}$.

הוכחה. ניתן לתרגם את המשוואה לחלוקה של k כדורים זהים ל-n תאים. זו בחירה עם החזרה (כי מותר לשים שני כדורים באותו התא) וללא חשיבות לסדר (כיוון שהכדורים זהים), ולכן בסך הכל יש $\binom{n+k-1}{n-1}$ דרכים לעשות זאת. $\square$

תרגיל 1.30 (אם יש זמן). כמה פתרונות שלמים יש למשוואה $x_1 + \cdots + x_4 = 12$ כאשר:

א. $x_i \geq 0$ לכל i?

ב. $x_i \geq 0$ לכל i, וידוע שלפחות אחד מהמשתנים הוא 0?

ג. $x_i \geq 2$ לכל i?

פתרון.

א. לפי המסקנה הקודמת, כמות הפתרונות למשוואה היא

$$\binom{12+4-1}{4-1} = \binom{15}{3} = \frac{15 \cdot 14 \cdot 13}{6} = 5 \cdot 7 \cdot 13 = 455.$$

ב. נחשב את מספר הפתרונות כאשר כל המשתנים חיוביים ממש. אם נגדיר $y_i = x_i - 1$, אז כל $y_i \geq 0$, וגם $y_1 + \cdots + y_4 = 8$. לכן כמות הפתרונות היא $\binom{8+4-1}{4-1} = \binom{11}{3} = \frac{11 \cdot 10 \cdot 9}{6} = 11 \cdot 5 \cdot 3 = 165$. והכמות של הפתרונות של המשוואה המקורית שבהם לפחות אחד המשתנים הוא 0 היא $455 - 165 = 290$.

ג. נגדיר $y_i = x_i - 2$. כעת רוצים לפתור את המשוואה $y_1 + \cdots + y_4 = 4$, כאשר $y_i \geq 0$. כמות הפתרונות למשוואה הזו היא $\binom{4+4-1}{4-1} = \binom{7}{3} = 35$.

2 תרגול שני

2.1 זהויות על מקדמים בינומיים

מקדמים בינומיים הם מושג חשוב בהרבה מקומות. יש כל מיני קשרים בין מקדמים בינומיים שונים, שניתן להוכיח בשלוש דרכים מרכזיות:

- דרך קומבינטורית – מראים ששני הצדדים של המשוואה סופרים את אותה הכמות.
- דרך אלגברית – באמצעות פתיחת הביטויים מראים ששני הצדדים שווים.
- אינדוקציה.

$$
\begin{array}{ccccccccccccccccccccc}
&&&&&&&&&&1&&&&&&&&&&\\
&&&&&&&&&1&&1&&&&&&&&&\\
&&&&&&&&1&&2&&1&&&&&&&&\\
&&&&&&&1&&3&&3&&1&&&&&&&\\
&&&&&&1&&4&&6&&4&&1&&&&&&\\
&&&&&1&&5&&10&&10&&5&&1&&&&&\\
&&&&1&&6&&15&&20&&15&&6&&1&&&&\\
&&&1&&7&&21&&35&&35&&21&&7&&1&&&\\
&&1&&8&&28&&56&&70&&56&&28&&8&&1&&\\
&1&&9&&36&&84&&126&&126&&84&&36&&9&&1&\\
1&&10&&45&&120&&210&&252&&210&&120&&45&&10&&1
\end{array}
$$

איור 1: משולש פסקל

משולש פסקל. המקדמים הבינומיים $\binom{n}{k}$ מוגדרים לכל $n \in \mathbb{N} \cup \{0\}$ ולכל $0 \le k \le n$. אם נסדר אותם במשולש, נקבל את משולש הפסקל:

אפשר לראות שלמשולש הזה יש הרבה סימטריות. למשל:

דוגמה 2.1 (בהרצאה)**.** לכל $0 \le k \le n$,

$$\binom{n}{k} = \binom{n}{n-k}.$$

דוגמה 2.2 (בהרצאה)**.** לכל $k < n$,

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

תרגיל 2.3. הוכיחו שלכל $n \ge 0$,

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

הוכחה קומבינטורית. באגף ימין כתובה כמות תת־הקבוצות בגודל n של $\{1, \dots, 2n\}$. נראה שגם אגף ימין סופר את אותה הכמות. אפשר לספור את תת־הקבוצות האלו באופן הבא: נקבע כמה איברים k מבין $\{1, \dots, n\}$ ישתתפו בתת־הקבוצה; נבחר את k האיברים האלו מ-$\{1, \dots, n\}$; ואז נבחר את $n-k$ האיברים הנותרים מ-$\{n+1, \dots, 2n\}$. אחרי שבחרנו את k, כמות הדרכים לבחור תת־קבוצה בגודל k של $\{1, \dots, n\}$ היא $\binom{n}{k}$, וכמות הדרכים לבחור תת־קבוצה בגודל $n-k$ של $\{n+1, \dots, 2n\}$ היא $\binom{n}{n-k} = \binom{n}{k}$. בסך הכל קיבלנו את אגף שמאל.

בסך הכל הראינו ששני האגפים סופרים את אותה הכמות, ולכן שווים. $\square$

משפט 2.4 (הבינום של ניוטון)**.** לכל $a, b \in \mathbb{R}$,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

דוגמה 2.5. המקדם של x^2y^3 בביטוי $(x+y)^5$ הוא $\binom{5}{2} = 10$.

דוגמה 2.6. אם מציבים $a=b=1$, מקבלים את הזהות $\sum_{k=0}^{n}\binom{n}{k}=2^n$. אפשר להוכיח אותה גם קומבינטורית: שני האגפים סופרים את כמות תת־הקבוצות של $[n]=\{1,\ldots,n\}$.

תרגיל 2.7. הראו כי לכל $n\in\mathbb{N}$,

$$\sum_{\substack{k=0\\ \text{זוגי } k}}^{n}\binom{n}{k}=\sum_{\substack{k=0\\ \text{אי־זוגי } k}}^{n}\binom{n}{k}=2^{n-1}.$$

הוכחה. נשתמש בבינום של ניוטון עם $a=-1$ ו-$b=1$:

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

מצד שני,

$$\sum_{k=0}^{n}\binom{n}{k}(-1)^k=\sum_{\substack{k=0\\ \text{זוגי } k}}^{n}\binom{n}{k}-\sum_{\substack{k=0\\ \text{אי־זוגי } k}}^{n}\binom{n}{k}.$$

לכן שני הסכומים האלו שווים. אך סכומם הוא 2^n לפי הדוגמה הקודמת, ולכן כל אחד מהם שווה 2^{n-1}. □

2.2 המקדם המולטינומי

הגדרה 2.8. יהי $n\in\mathbb{N}$ ו-$m_1,\ldots,m_k\in\mathbb{N}$ כך ש-$m_1+\cdots+m_k=n$. **המקדם המולטינומי של n מעל $m_1,\ldots,m_k$** מוגדר להיות

$$\binom{n}{m_1,\ldots,m_k}=\frac{n!}{m_1!\cdots m_k!}.$$

הוא סופר את כמות הדרכים לחלק n כדורים לקבוצות בגודל $m_1,\ldots,m_k$.

הערה 2.9. שימו לב כי

$$\binom{n}{m_1,\ldots,m_k}=\binom{n}{m_1}\binom{n-m_1}{m_2}\binom{n-(m_1+m_2)}{m_3}\cdots\binom{n-(m_1+\cdots+m_{k-1})}{m_k}.$$

גם את הזהות הזו ניתן להוכיח באופן אלגברי או קומבינטורי.

הוכחה קומבינטורית. רוצים לחלק את $\{1,\ldots,n\}$ לקבוצות בגודל $m_1,\ldots,m_k$. נבחר ראשית את הקבוצה הראשונה, בת m_1 איברים; לבחירה הזו יש $\binom{n}{m_1}$ אפשרויות. נמשיך לבחירת הקבוצה השנייה; כעת יש $\binom{n-m_1}{m_2}$ אפשרויות. ניתן להמשיך כך ולקבל את הביטוי באגף ימין. □

הערה 2.10. הכללה של הבינום של ניוטון:

$$(a_1+\cdots+a_k)^n=\sum_{m_1+\cdots+m_k=n}\binom{n}{m_1,\ldots,m_k}a_1^{m_1}\cdots a_k^{m_k}.$$

תרגיל 2.11. מה מספר האפשרויות לחלק 7 תלמידים ל-3 קבוצות, אחת עם 3 תלמידים והשתיים האחרות עם 2 תלמידים?

פתרון. באמצעות המקדם המולטינומי,

$$\binom{7}{3,2,2} = \frac{7!}{3!\cdot 2!\cdot 2!}.$$

(כמו בהערה, אפשר גם לפתור עם מקדמים בינומיים: לבחירת הקבוצה הראשונה יש $\binom{7}{3}$ אפשרויות, לשנייה יש $\binom{4}{2}$ ולשלישית יש $1 = \binom{2}{2}$, ובסך הכל $\binom{7}{3}\cdot\binom{4}{2}\cdot\binom{2}{2} = \frac{7!}{3!\cdot 2!\cdot 2!}$).

2.3 מרחבי הסתברות בדידים

הגדרה 2.12.

א. **מרחב מדגם** יהיה עבורנו כל התוצאות האפשריות בניסוי כלשהו. לרוב נסמן אותו על ידי Ω. מרחב המדגם יכול להיות סופי או אינסופי.

ב. **מאורע פשוט** הינו תוצאה אפשרית אחת של הניסוי.

ג. **מאורע** הינו אוסף של מספר תוצאות אפשריות של הניסוי.

הגדרה 2.13. יהי Ω מרחב מדגם (בינתיים סופי או בן מנייה). **פונקציית הסתברות** היא פונקציה $P : 2^{\Omega} \to \mathbb{R}$ המקיימת את שלושת התנאים הבאים:

א. לכל $A \subseteq \Omega$, $0 \leq P(A) \leq 1$.

ב. לכל אוסף סופי או בן־מנייה של מאורעות זרים בזוגות $A_1, A_2, \ldots \subseteq \Omega$ מתקיים $P\left(\dot\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$.

ג. $P(\Omega) = 1$.

לזוג (Ω, P) נקרא **מרחב הסתברות** (**בדיד**).

הערה 2.14. בהמשך נדבר על מרחבי הסתברות רציפים, שבהם מרחב המדגם יכול להיות גם לא בן מנייה.

דוגמה 2.15. נסתכל על המרחב של הטלות מטבע $\Omega = \{H, T\}$. במקרה זה ניתן להגדיר פונקציית הסתברות $P(\{H\}) = P(\{T\}) = \frac{1}{2}$ שמתאימה להטלה של מטבע הוגן. במקרה שבו המטבע לא הוגן, אפשר להגדיר פונקציית הסתברות אחרת: למשל $P(\{H\}) = \frac{2}{3}$, $P(\{T\}) = \frac{1}{3}$.

טענה 2.16. יהי (Ω, P) מרחב הסתברות.

א. לכל $A \subseteq \Omega$, $P(A^c) = 1 - P(A)$.

ב. אם $A \subseteq B$, אז $P(B \setminus A) = P(B) - P(A)$.

ג. אם $A \subseteq B$, אז $P(A) \leq P(B)$.

ד. לכל $A, B \subseteq \Omega$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

תרגיל 2.17. לקוח מגיע לבית קפה. ההסתברות שהלקוח יזמין קפה היא 0.7, ההסתברות שיזמין מאפה היא 0.6, וההסתברות שיזמין קפה או מאפה היא 0.8.

א. תארו את מרחב המדגם.

ב. מהי ההסתברות שיזמין קפה ומאפה?

ג. מהי ההסתברות שיזמין פריט אחד?

פתרון.

א. מרחב המדגם Ω מכיל את כל האפשרויות להזמנה של הלקוח. ניעזר בסימוני עזר: אם A הוא המאורע שהלקוח יזמין קפה, ו-B הוא המאורע שהלקוח יזמין מאפה, אז מרחב המדגם הוא $\Omega = \{A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c\}$.

ב. נתרגם את הנתונים: $P(A) = 0.7$, $P(B) = 0.6$ ו-$P(A \cup B) = 0.8$. לפי התרגיל הקודם,

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.6 + 0.7 - 0.8 = 0.5.$$

ג. כעת, רוצים לחשב את

$$P((A \cup B) \setminus (A \cap B)) = P(A \cup B) - P(A \cap B) = 0.8 - 0.5 = 0.3$$

כשבמעבר הראשון השתמשנו בכך ש-$A \cap B \subseteq A \cup B$.

2.4 מרחבי הסתברות אחידים

הגדרה 2.18. מרחב הסתברות אחיד הוא מרחב הסתברות סופי (Ω, P) כך שלכל $A \subseteq \Omega$ מתקיים $P(A) = \frac{|A|}{|\Omega|}$. כלומר, כל תוצאה בניסוי מתקבלת בסיכוי זהה, וכל מה שמעניין אותנו הוא כמה תוצאות אפשריות יש במאורע A וכמה יש במרחב המדגם Ω.

לכן כל מה שלמדנו בקומבינטוריקה לא היה לשווא; בכל שאלה על מרחב הסתברות אחיד, עובדים לפי המודל הבא:

- מגדירים את מרחב המדגם Ω;
- סופרים כמה איברים יש ב-Ω;
- מגדירים את המאורע A כתת־קבוצה של Ω;
- סופרים כמה איברים יש ב-A;
- מחלקים את הכמויות שקיבלנו.

תרגיל 2.19. רותם וירון זוג מבולגן וחובב גרביים. במגירת הגרביים שלהם מפוזרים באקראי 7 גרביים שחורים, 9 כחולים ו-8 עם ציורים של דינוזאורים. ירון קם לעבודתו בבית הספר ליידי דייויס ושולף מן המגירה 2 גרביים באקראי. חשבו את ההסתברויות הבאות:

א. ירון בחר 2 גרביים מאותו הסוג.

ב. ירון בחר לפחות גרב אחד שחור.

ג. ירון בחר גרב אחד שחור וגרב אחד עם דינוזאורים.

פתרון. מרחב המדגם שלנו Ω הוא כל הזוגות של גרביים במגירה. נשים לב שאין חשיבות לסדר כי הגרביים נשלפים ביחד, ולא ניתן לשלוף את אותו הגרב פעמיים. לכן $|\Omega| = \binom{7+8+9}{2} = \binom{24}{2} = 276$.

א. נסמן את המאורע שלנו A. נפרק למקרים לפי עיקרון הסכום: המאורע שלנו A מורכב מאיחוד זר של שלושה מאורעות: שולפים שני גרביים שחורים ($\binom{7}{2}$ אפשרויות), שולפים שני גרביים כחולים ($\binom{8}{2}$ אפשרויות), ושולפים שני גרביים ירוקים ($\binom{9}{2}$ אפשרויות). בסך הכל

$$P\left(A\right) = \frac{|A|}{|\Omega|} = \frac{\binom{7}{2}+\binom{8}{2}+\binom{9}{2}}{276} = \frac{85}{276}.$$

ב. נסמן את המאורע על ידי B. נחשב את המאורע המשלים, שבו לא בחרנו אף גרב שחור. במקרה הזה מותר לנו רק לבחור מתוך 17 גרביים, לכן

$$P\left(B\right) = 1 - P\left(B^c\right) = 1 - \frac{|B^c|}{|\Omega|} = 1 - \frac{\binom{17}{2}}{276} = 1 - \frac{136}{276} = \frac{140}{276} = \frac{35}{69}.$$

ג. נסמן את המאורע על ידי C. יש לנו 7 אפשרויות לבחור גרב שחור ו-8 אפשרויות לבחור גרב כחול, לכן

$$P\left(C\right) = \frac{8 \cdot 7}{276} = \frac{56}{276} = \frac{14}{69}.$$

תרגיל 2.20. 8 אנשים עומדים וביניהם אליס, בוב, גל ודור מסתדרים בשורה באקראי. חשבו את ההסתברויות הבאות:

א. אליס וגל עומדות זו ליד זו.

ב. בוב או דור נמצאים בקצה.

ג. אליס וגל עומדות זו ליד זו, אבל בוב ודור לא עומדים אחד ליד השני.

פתרון. פה מרחב המדגם Ω הוא אוסף כל הסידורים האפשריים של שמונה אנשים בשורה; לכן $|\Omega| = 8!$.

א. נסמן את המאורע על ידי A. במקרה הזה נחשוב על אליס ועל גל כעל "בלוק" אחד. כעת יש לסדר 7 אובייקטים בשורה, כשאחד מהם הוא ה"בלוק" והיתר הם ששת האנשים האחרים. סך האפשרויות לסידור הוא $7!$. נכפול בכמות האפשרויות לסידור הפנימי של הבלוק, שהיא $2 = 2!$, ונקבל ש-$|A| = 7! \cdot 2!$. לכן

$$P\left(A\right) = \frac{|A|}{|\Omega|} = \frac{7! \cdot 2!}{8!} = \frac{2}{8} = \frac{1}{4}.$$

ב. נסמן את המאורע על ידי B. ניעזר במשלים B^c, שבו בוב ודור לא עומדים בקצה. נסדר את שמונת האנשים כך: ראשית נמצא מקום לבוב, לזה יש 6 אפשרויות; כעת נמצא מקום לדור מבין 5 האפשרויות שאינן בקצה שנותרו; ונסדר את שאר 6 האנשים בשורה ($6!$ אפשרויות). בסך הכל $|B^c| = 6 \cdot 5 \cdot 6!$. לכן

$$P\left(B\right) = 1 - P\left(B^c\right) = 1 - \frac{6 \cdot 5 \cdot 6!}{8!} = 1 - \frac{6 \cdot 5}{8 \cdot 7} = 1 - \frac{15}{28} = \frac{13}{28}.$$

ג. נסמן על ידי C את המאורע שבו אליס וגל עומדות זו ליד זו וגם בוב ודור עומדים זה ליד זו, ועל ידי D את המאורע שבו אליס וגל עומדות זו ליד זו אבל בוב ודור לא עומדים זה ליד זה. נשים לב כי $D = A \setminus C$ וכי $C \subseteq A$; לכן $P(D) = P(A) - P(C)$. נחשב את ההסתברות של C: בדומה לחישוב A, הפעם נייצר שני בלוקים, נסדר את 6 האובייקטים שנשארו, ואז נסדר כל אחד מהבלוקים באחד מ-$2 = 2!$ הסידורים הפנימיים שלו. בסך הכל

$$P(D) = P(A) - P(C) = \frac{1}{4} - \frac{6! \cdot 2! \cdot 2!}{8!} = \frac{1}{4} - \frac{4}{56} = \frac{1}{4} - \frac{1}{14} = \frac{5}{28}.$$

3 תרגול שלישי

3.1 רציפות ההסתברות

טענה 3.1. יהי (Ω, P) מרחב הסתברות. אז:

א. אם $A_1 \subseteq A_2 \subseteq \cdots$ סדרה עולה, אז $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$.

ב. אם $B_1 \supseteq B_2 \supseteq \cdots$ סדרה יורדת, אז $P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} P(B_n)$.

תרגיל 3.2. מטילים מטבע הוגן אינסוף פעמים. הוכיחו שההסתברות לקבל רק מספר סופי של פעמים "עץ" היא 0.

הוכחה. יהי A המאורע שבו מתקבל רק מספר סופי של פעמים "עץ". נגדיר B_n להיות המאורע שהחל מההטלה ה-n-ית כל ההטלות הן "פלי". לכן $B_1 \subseteq B_2 \subseteq \cdots$, ומתקיים $A = \bigcup_{n=1}^{\infty} B_n$. לכן נרצה לחשב את $P(B_n)$.

כדי לחשב את B_n, לכל $k \geq n$ נסמן $C_{n,k}$ להיות המאורע שכל ההטלות מההטלה ה-n עד להטלה ה-k הן "פלי". אז $C_{n,n} \supseteq C_{n,n+1} \supseteq \cdots$, ומתקיים $B_n = \bigcap_{k=n}^{\infty} C_{n,k}$. נשים לב כי $P(C_{n,k}) = \left(\frac{1}{2}\right)^{k-n+1}$, ולכן $P(B_n) = \lim_{k\to\infty} P(C_{n,k}) = \lim_{k\to\infty} \left(\frac{1}{2}\right)^{k-n+1} = 0$. מכאן נקבל $P(A) = \lim_{n\to\infty} P(B_n) = 0$. $\square$

3.2 חסם האיחוד

טענה 3.3 (חסם האיחוד). לכל אוסף מאורעות $\{A_i\}_{i\in I}$ סופי או בן־מנייה, מתקיים

$$P\left(\bigcup_{i\in I} A_i\right) \leq \sum_{i\in I} P(A_i).$$

השוו עם האקסיומה שהגדירה פונקציית הסתברות. שם דרשנו שיהיה שוויון אם המאורעות $\{A_i\}$ זרים בזוגות; פה, כיוון שיכולים להיות חיתוכים בין הקבוצות, מקבלים רק אי־שוויון.

דוגמה 3.4. במעבדה מבצעים בדיקות קורונה. במקום לבדוק כל דגימה בנפרד, מחלקים אותן לקבוצות של n דגימות, מערבבים חלק מכל דגימה, ואז בודקים האם התערובת החדשה מכילה נגיפי קורונה (כלומר, אם לפחות אחת מהדגימות חיובית). אם היא יצאה חיובית, בודקים כל אדם בנפרד.

לפי חסם איחוד, אם הסיכוי שהבדיקה תצא חיובית היא p, הסיכוי שהבדיקה המעורבת תצא חיובית הוא לכל היותר np.

תרגיל 3.5. בכפר יש n אנשים. כל שניים מהם מכירים אחד את השני בהסתברות p באופן בלתי־תלוי בשאר האנשים. תנו חסם להסתברות שיש אדם שאינו מכיר אף אחד אחר בכפר.

בשפה של גרפים: אפשר לחשוב על האנשים כעל n קודקודים, ומוסיפים קשת בין שני קודקודים בסיכוי p באופן בלתי־תלוי בקשתות האחרות. כך מקבלים גרף מקרי על n קודקודים, שנקרא גם **מודל ארדש-רייני** (ומסומן $G(n,p)$).

פתרון. יהי A_i המאורע שבו האדם ה-i לא מכיר אף אחד. אז $\bigcup_{i=1}^{n} A_i$ הוא בדיוק המאורע שבו יש אדם שאינו מכיר אף אחד אחר בכפר. לפי חסם איחוד,

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

מצד שני, ההסתברות שאדם מסוים אינו מכיר אף אחד בכפר היא $(1-p)^{n-1}$ (כי יש $n-1$ אנשים אחרים, ואת כל אחד מהם הוא לא מכיר בהסתברות $1-p$). בסך הכל מחסם איחוד נקבל

$$.P\left(\bigcup_{i=1}^{n} A_i\right) \leq n(1-p)^{n-1}$$

3.3 עיקרון ההכלה וההפרדה

תזכורת. עבור שתי קבוצות A, B, מתקיים

$$.P\left(A \cup B\right) = P\left(A\right) + P\left(B\right) - P\left(A \cap B\right)$$

מה לגבי שלוש קבוצות?

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

היינו רוצים להבין איך לחשב באופן דומה את ההסתברות של איחוד סופי של מאורעות.

טענה 3.6 (עיקרון ההכלה וההפרדה). לכל אוסף מאורעות $A_1, \ldots, A_n \subseteq \Omega$,

$$.P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} P\left(A_{i_1} \cap \cdots \cap A_{i_k}\right)$$

תרגיל 3.7. מגרילים מספר באופן אחיד בין 1 ל-1000. מה הסיכוי שהוא יתחלק ב-2, ב-3 או ב-5?

פתרון. נגדיר את המאורעות A_2, A_3 ו-A_5 כאוסף כל המספרים בין 1 ל-1000 המתחלקים ב-2, ב-3 וב-5 בהתאמה. אנחנו רוצים לחשב את $|A_2 \cup A_3 \cup A_5|$. נחשב את גדלי הקבוצות:

$$.\,|A_2| = \left\lfloor \frac{1000}{2} \right\rfloor = 500; \qquad |A_3| = \left\lfloor \frac{1000}{3} \right\rfloor = 333; \qquad |A_5| = \left\lfloor \frac{1000}{5} \right\rfloor = 200$$

כעת, החיתוך $A_2 \cap A_3$ מכיל את כל המספרים המתחלקים ב-2 וב-3, ולכן מכיל בדיוק את המספרים המתחלקים ב-6. משיקול דומה עבור $A_2 \cap A_5$ ו-$A_3 \cap A_5$,

$$.\,|A_2 \cap A_3| = \left\lfloor \frac{1000}{6} \right\rfloor = 166; \quad |A_2 \cap A_5| = \left\lfloor \frac{1000}{10} \right\rfloor = 100; \quad |A_3 \cap A_5| = \left\lfloor \frac{1000}{15} \right\rfloor = 66$$

המספרים הנמצאים ב-$A_2\cap A_3\cap A_5$ הם בדיוק המספרים בין 1 ל-1000 המתחלקים ב-30, ולכן

$$|A_2\cap A_3\cap A_5| = \left\lfloor\frac{1000}{30}\right\rfloor = 33.$$

לפי עיקרון ההכלה וההפרדה,

$$P\left(A_2\cup A_3\cup A_5\right) = \frac{1}{|\Omega|}\left(|A_2|+|A_3|+|A_5|-|A_2\cap A_3|-|A_2\cap A_5|-|A_3\cap A_5|+|A_2\cap A_3\cap A_5|\right) =$$
$$= \frac{1}{1000}\left(500+333+200-166-100-66+33\right) = 0.734.$$

מסקנה 3.8. נניח שהקבוצות "סימטריות", כלומר החיתוכים בגודל k תמיד שווים זה לזה בגודלם. אז אפשר לקבל ניסוח פשוט יותר:

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}P\left(A_1\cap\cdots\cap A_k\right).$$

תרגיל 3.9. בכמה דרכים ניתן לחלק 200 כדורים זהים ל-10 תאים, כך שלא יהיה תא עם 30 כדורים או יותר?

פתרון. מרחב המדגם Ω הוא כל הדרכים לחלק 200 כדורים זהים ל-10 תאים. זו בחירה של 200 מתוך 10, עם החזרה וללא חשיבות לסדר (כי הכדורים זהים), ולכן $|\Omega| = \binom{10+200-1}{10-1} = \binom{209}{9}$.

נסמן על ידי A_i את המאורע שבו בתא ה-i יש 30 כדורים או יותר. הגודל שאנחנו רוצים לחשב הוא $\left|\left(\bigcup_{i=1}^{10} A_i\right)^c\right| = |\Omega| - \left|\bigcup_{i=1}^{10} A_i\right|$. לפי נוסחת ההכלה וההפרדה,

$$\left|\bigcup_{i=1}^{10} A_i\right| = \sum_{k=1}^{10}(-1)^{k-1}\sum_{1\le i_1<\cdots<i_k\le 10}|A_{i_1}\cap\cdots\cap A_{i_k}|.$$

נשים לב שכיוון שהחלוקה נעשית באקראי, החיתוכים $A_{i_1}\cap\cdots\cap A_{i_k}$ סימטריים (כמות החלוקות שבהן בתאים $i_1,\ldots,i_k$ יש לפחות 30 כדורים היא זהה, לא משנה באילו תאים מדובר), ולכן אפשר להשתמש בניסוח הסימטרי ולקבל

$$\left|\bigcup_{i=1}^{10} A_i\right| = \sum_{k=1}^{10}(-1)^{k-1}\binom{10}{k}|A_1\cap\cdots\cap A_k|.$$

נחשב את $|A_1\cap\cdots\cap A_k|$. החלוקות ב-$A_1\cap\cdots\cap A_k$ הן כל החלוקות של 200 הכדורים ל-10 התאים, כך שב-k התאים הראשונים יש לפחות 30 כדורים. נשים לב שבהכרח $k\le 6$, אחרת חילקנו יותר מדי כדורים. אפשר לספור זאת כך: בכל אחד מ-k התאים הראשונים נשים ראשית 30 כדורים (ואין חשיבות לאילו כדורים, כי הם זהים), ולאחר מכן את $200-30k$ הכדורים הנותרים נחלק בין כל 10 התאים. זו בחירה של $200-30k$ מ-10, עם החזרה וללא חשיבות לסדר, ולכן כמות הדרכים לחלק כדורים כך היא $|A_1\cap\cdots\cap A_k| = \binom{10+200-30k-1}{10-1} = \binom{209-30k}{9}$.

אם נציב חזרה בנוסחת ההכלה וההפרדה, נקבל

$$\left|\left(\bigcup_{i=1}^{10} A_i\right)^c\right| = |\Omega| - \left|\bigcup_{i=1}^{10} A_i\right| = \binom{209}{9} - \sum_{k=1}^{6}(-1)^{k-1}\binom{10}{k}\binom{209-30k}{9}.$$

תרגיל 3.10. $2n$ אנשים שמחולקים ל-n זוגות מתיישבים באקראי סביב שולחן מעגלי. חשבו את ההסתברויות של המאורעות הבאים:

א. רק $n-1$ זוגות יושבים ביחד?

ב. אף זוג אינו יושב ביחד?

פתרון. מרחב המדגם שלנו Ω מכיל את כל הסידורים האפשריים של $2n$ אנשים במעגל. כפי שחישבנו בתרגול הראשון, $|\Omega| = (2n-1)!$.

א. נסמן את המאורע על ידי A. נחלק את הבחירה של סידור למספר שלבים. בשלב הראשון צריך לבחור איזה אחד מבין n הזוגות יהיה הזוג שמופרד (n אפשרויות). בשלב השני נסדר את $n-1$ הזוגות האחרים כך שיישארו יחדיו ($2^{n-1}\cdot(n-2)!$ אפשרויות). כעת נמקם את הראשון מהזוג שבחרנו באחד המקומות בין $n-1$ הבלוקים שסידרנו ($n-1$ אפשרויות), ואת השני מהזוג באחד מבין $n-2$ המקומות שנותרו ($n-2$ אפשרויות). בסך הכל

$$P\left(A\right)=\frac{|A|}{|\Omega|}=\frac{n\cdot 2^{n-1}\cdot(n-2)!\cdot(n-1)\cdot(n-2)}{(2n-1)!}=\frac{2^{n-1}\cdot(n-2)\cdot n!}{(2n-1)!}.$$

ב. נסמן על ידי B_i את המאורע שבו הזוג ה-i יושב זה ליד זה. אנו רוצים לחשב את $P\left(\left(\bigcup_{i=1}^{n}B_i\right)^c\right)$. גם פה הקבוצות סימטריות, ולכן

$$P\left(\left(\bigcup_{i=1}^{n}B_i\right)^c\right)=1-\sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}P\left(B_1\cap\cdots\cap B_k\right).$$

נחשב את ההסתברות של החיתוכים. בחיתוך $B_1\cap\cdots\cap B_k$ מופיעים כל הסידורים שבהם k זוגות נמצאים ביחד. שוב על ידי שיטת הבלוקים, נקבל שיש $2n-2k$ אנשים ו-k בלוקים, ולכן יש $(2n-k-1)!$ דרכים לסדר אותם במעגל ו-2^k סידורים פנימיים אפשריים. לכן

$$P\left(\left(\bigcup_{i=1}^{n}B_i\right)^c\right)=1+\sum_{k=1}^{n}(-1)^{k}\binom{n}{k}\frac{2^k\cdot(2n-1-k)!}{(2n-1)!}.$$

תרגיל 3.11 (אם יש זמן, הלמה של Bonferroni). הוכיחו כי לכל אוסף מאורעות $A_1,\ldots,A_n$,

$$P\left(\bigcup_{i=1}^{n}A_i\right)\geq\sum_{i=1}^{n}P\left(A_i\right)-\sum_{1\leq i<j\leq n}P\left(A_i\cap A_j\right).$$

הוכחה. נוכיח באינדוקציה על n. עבור $n=1$, הטענה ברורה ואין מה להוכיח. עבור $n=2$, הטענה טוענת כי

$$P\left(A_1\cup A_2\right)\geq P\left(A_1\right)+P\left(A_2\right)-P\left(A_1\cap A_2\right)$$

ולמעשה יש שוויון.

נניח שהטענה נכונה לכל n מאורעות, ונוכיח עבור $n+1$. אז נשים לב כי

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\left(\bigcup_{i=1}^{n} A_i\right) \cup A_{n+1}\right) \geq \\
\underset{n=2}{\text{המקרה}} &\geq P\left(\bigcup_{i=1}^{n} A_i\right) + P\left(A_{n+1}\right) - P\left(\left(\bigcup_{i=1}^{n} A_i\right) \cap A_{n+1}\right) \geq \\
\underset{\text{האינדוקציה}}{\text{הנחת}} &\geq \sum_{i=1}^{n} P\left(A_i\right) - \sum_{1\leq i<j\leq n} P\left(A_i \cap A_j\right) + P\left(A_{n+1}\right) - P\left(\bigcup_{i=1}^{n}\left(A_i \cap A_{n+1}\right)\right) \geq \\
\underset{\text{איחוד}}{\text{חסם}} &\geq \sum_{i=1}^{n} P\left(A_i\right) + P\left(A_{n+1}\right) - \sum_{1\leq i<j\leq n} P\left(A_i \cap A_j\right) - \sum_{i=1}^{n} P\left(A_i \cap A_{n+1}\right)
\end{aligned}$$

כנדרש. $\square$

הערה 3.12. חסם איחוד הוא מקרה אחד של הלמה של Bonferroni, והתרגיל האחרון הוא מקרה נוסף. הלמה הכללית נותנת חסמים להסתברות אם מחשבים את הסכום של עיקרון ההכלה וההפרדה, אבל עוצרים אחרי כמה צעדים. אם עוצרים אחרי כמות צעדים אי־זוגית, מקבלים חסם עליון; אם עוצרים אחרי כמות צעדים זוגית, מקבלים חסם תחתון.

4 תרגול רביעי

4.1 הסתברות מותנית

הגדרה 4.1. יהיו $A, B \subseteq \Omega$ מאורעות כלשהם כך ש-$P(B) > 0$. **ההסתברות של** A **בהינתן ש-**B **התרחש** היא ההסתברות

$$P\left(A \mid B\right) = \frac{P\left(A \cap B\right)}{P\left(B\right)}.$$

דוגמה 4.2. בכד יש 10 כדורים שחורים, 10 כדורים לבנים ו-5 כדורים צהובים. אם הוצאנו כדור שאינו שחור, ההסתברות שהכדור שהוצאנו צהוב היא $\frac{1}{3} = \frac{5}{15} = \frac{\frac{5}{25}}{\frac{15}{25}}$.

הערה. הפונקציה $P\left(\cdot \mid B\right)$ מגדירה בפני עצמה פונקציית הסתברות. במקרה זה, המשלים של $A \mid B$ הוא $A^c \mid B$.

הערה. על פי ההגדרה, $P\left(B \mid A\right) = \frac{P(A\cap B)}{P(A)}$, ומכאן ניתן לגזור את חוק בייס:

$$P\left(B \mid A\right) = \frac{P\left(A \mid B\right) \cdot P\left(B\right)}{P\left(A\right)}.$$

נציג עוד שתי טענות חשובות בהקשר הסתברות מותנית:

משפט 4.3 (נוסחת ההסתברות השלמה)**.** תהי $\left\{A_i\right\}_{i\in I}$ חלוקה של מרחב המדגם (כלומר הקבוצות A_i זרות בזוגות וכן $\dot\bigcup_{i\in I} A_i = \Omega$) למספר סופי או בן־מנייה של קבוצות. אז לכל $B \subseteq \Omega$,

$$P\left(B\right) = \sum_{i\in I} P\left(B \mid A_i\right) \cdot P\left(A_i\right).$$

משפט 4.4 (עיקרון הכפל). יהיו $A_1,\dots,A_n$ מאורעות כלשהם. אזי

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P\left(A_1\right)\cdot P\left(A_2 \mid A_1\right)\cdots P\left(A_n \mid A_1\cap\cdots\cap A_{n-1}\right).$$

תרגיל 4.5. בכד יש 5 כדורים, מתוכם 2 לבנים ו-3 שחורים. נוציא באקראי 2 כדורים: אם הראשון לבן – נשאיר אותו בחוץ ואז נוציא את השני, ואם הראשון שחור – נחזיר אותו פנימה ונוציא את הכדור השני.

א. מה ההסתברות שהכדור השני שחור?

ב. אם הכדור השני שהוצאנו שחור, מה ההסתברות שהכדור הראשון היה לבן?

פתרון. אפשר לצייר עם עץ ולחשב את ההסתברויות, ובאופן שקול להשתמש בנוסחת ההסתברות השלמה:

$$P\left(\text{הכדור השני שחור}\right) = \frac{3}{5}\cdot\frac{3}{5}+\frac{2}{5}\cdot\frac{3}{4} = \frac{9}{25}+\frac{3}{10} = \frac{33}{50}.$$

כעת, לפי חוק בייס,

$$P\left(\text{הכדור הראשון לבן} \mid \text{הכדור השני שחור}\right) = \frac{P\left(\text{הכדור השני שחור} \mid \text{הכדור הראשון לבן}\right)\cdot P\left(\text{הכדור הראשון לבן}\right)}{P\left(\text{הכדור השני שחור}\right)} =$$

$$= \frac{\frac{3}{4}\cdot\frac{2}{5}}{\frac{33}{50}} = \frac{\frac{3}{10}}{\frac{33}{50}} = \frac{5}{11}.$$

תרגיל 4.6. בחברת "בדיקורונה" פיתחו בדיקות קורונה. החברה בדקה את יעילות הבדיקה שלה על אוכלוסייה ש-10% ממנה חולים, וקיבלה שבסך הכל 12% מהבדיקות חזרו חיוביות. כמו כן, ה-false negative של הבדיקה הוא 10%, כלומר אם אדם חולה, הסיכוי שהבדיקה שלו תצא שלילית הוא 10%.

א. אם התוצאה של אדם חזרה חיובית, מה הסיכוי שהוא בריא?

ב. אם אדם בריא, מה הסיכוי שהתוצאה של הבדיקה תחזור חיובית?

פתרון. נסמן על ידי A את המאורע שהאדם חולה ועל ידי B את המאורע שהבדיקה חזרה חיובית. הנתונים אומרים כי $P(A)=0.1$, $P(B)=0.12$ ו-$P(B^c\mid A)=0.1$. לכן $P(A\cap B^c)=P(B^c\mid A)\cdot P(A)=0.01$. נכתוב את הנתונים בטבלה ונמלא אותה:

	בדיקה חיובית B	בדיקה שלילית B^c	
האדם חולה A	0.09	0.01	0.1
האדם בריא A^c	0.03	0.87	0.9
	0.12	0.88	

(מה שצבוע בכחול הוא מידע שהשלמנו). נחשב את ההסתברויות הדרושות:

$$P\left(A^c\mid B\right) = \frac{P\left(A^c\cap B\right)}{P\left(B\right)} = \frac{0.03}{0.12} = 0.25,$$

ולפי חוק בייס,

$$P\left(B\mid A^c\right) = \frac{P\left(A^c\mid B\right)\cdot P\left(B\right)}{P\left(A^c\right)} = \frac{\frac{1}{4}\cdot\frac{3}{25}}{\frac{9}{10}} = \frac{1}{30}\approx 0.033.$$

תרגיל 4.7. תהי S קבוצה בת n איברים. מבין 2^n תת־הקבוצות של S נבחר באופן אחיד ועם החזרה שתי קבוצות A ו-B.

א. חשבו את ההסתברות ש-A בת k איברים.

ב. היעזרו בסעיף א' כדי לחשב את ההסתברות ש-$A \subseteq B$.

פתרון.

א. יש $\binom{n}{k}$ תת־קבוצות של S מגודל k, ולכן $P\left(\{|A| = k\}\right) = \frac{\binom{n}{k}}{2^n}$.

ב. נשתמש בנוסחת ההסתברות השלמה ביחס למאורעות $\{|A| = k\}$:

$$P\left(\{A \subseteq B\}\right) = \sum_{k=0}^{n} P\left(\{A \subseteq B\} \mid \{|A| = k\}\right) \cdot P\left(\{|A| = k\}\right) = \sum_{k=0}^{n} \left(\frac{1}{2}\right)^k \binom{n}{k} \left(\frac{1}{2}\right)^n =$$
$$= \left(\frac{1}{2}\right)^n \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{2}\right)^k 1^{n-k} = \left(\frac{1}{2}\right)^n \cdot \left(\frac{3}{2}\right)^n = \frac{3^n}{4^n}.$$

תרגיל 4.8. מטילים על השולחן מטבעות הוגנים עד שיוצא בפעם הראשונה "עץ", ואז מגרילים באקראי מטבע מבין המטבעות שעל השולחן. מהי ההסתברות שהמטבע שנבחר ייצא "עץ"?

פתרון. נסמן על ידי B את המאורע שהמטבע שנרים מן השולחן יהיה "עץ". ניעזר בנוסחת ההסתברות השלמה על החלוקה הבאה של המרחב: A_n = כל ההטלות שבהן יצא "עץ" בפעם הראשונה בהטלה ה-n. (זו אמנם לא חלוקה בדיוק של כל המרחב, כי נשאר מאורע פשוט אחד שלא נכנס לחלוקה – כל ההטלות הן "פלי", אך זהו מאורע מהסתברות 0 ולכן הוא לא משפיע):

$$P\left(B\right) = \sum_{n=1}^{\infty} P\left(B \mid A_n\right) \cdot P\left(A_n\right) = \sum_{n=1}^{\infty} \frac{1}{n} \cdot \left(\frac{1}{2}\right)^n.$$

ניזכר מאינפי 2 כי טור טיילור של $\ln\left(1+x\right)$ הינו $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}x^n}{n}$; לכן טור טיילור של $\ln\left(1-x\right)$ הוא $-\sum_{n=1}^{\infty} \frac{x^n}{n}$, ונקבל

$$P\left(B\right) = -\ln\left(1 - \frac{1}{2}\right) = -\ln\frac{1}{2} = \ln 2.$$

תרגיל 4.9. יהי (Ω, P) מרחב הסתברות, ויהי $B \subseteq \Omega$ מאורע בעל הסתברות חיובית. נתבונן במרחב ההסתברות (Ω, Q) כאשר פונקציית ההסתברות Q מוגדרת באופן הבא: $Q\left(A\right) = P\left(A \mid B\right)$.

א. יהי C מאורע עם הסתברות חיובית תחת Q (כלומר $Q\left(C\right) > 0$). הוכיחו כי $B \cap C$ הוא מאורע בעל הסתברות חיובית תחת P.

ב. הוכיחו כי $Q\left(A \mid C\right) = P\left(A \mid B \cap C\right)$ לכל $A \subseteq \Omega$.

ג. השתמשו בתוצאה מסעיף ב' כדי לבטא את נוסחת ההסתברות השלמה במרחב ההסתברות (Ω, Q) במונחים של P בלבד.

הוכחה.

א. לפי ההגדרה $0 < P(B \cap C) \cdot P(B) = P(C \mid B) = Q(C)$. לכן בהכרח $P(B \cap C) > 0$.

ב. מחישוב ישיר,

$$Q(A \mid C) = \frac{Q(A \cap C)}{Q(C)} = \frac{P(A \cap C \mid B)}{P(C \mid B)} = \frac{\frac{P(A \cap B \cap C)}{P(B)}}{\frac{P(B \cap C)}{P(B)}} = \frac{P(A \cap B \cap C)}{P(B \cap C)} = P(A \mid B \cap C).$$

ג. תהי $C_1, \ldots, C_n$ חלוקה של Ω. אזי

$$Q(A) = \sum_{i=1}^{n} Q(A \mid C_i) \cdot Q(C_i)$$

ובמונחי P נקבל

$$P(A \mid B) = \sum_{i=1}^{n} P(A \mid B \cap C_i) \cdot P(C_i \mid B).$$

$\square$

4.2 תלות ואי-תלות

הגדרה 4.10. נאמר כי מאורעות A ו-B הם **בלתי־תלויים**, אם $P(A \cap B) = P(A) \cdot P(B)$. אחרת, נאמר שהם **תלויים**.

הערה. A ו-B בלתי־תלויים אם ורק אם $P(A \mid B) = P(A)$, אם ורק אם $P(B \mid A) = P(B)$. כלומר, התרחשותו של אחד המאורעות לא משפיעה על התרחשות המאורע השני.

הערה. מאורעות זרים שהסתברותם אינה שווה 0 או 1 הם בהכרח תלויים! כי אם אנחנו יודעים שאחד התרחש, אנחנו יודעים אוטומטית שהשני לא התרחש.

תרגיל 4.11. נטיל שתי קוביות. יהי A המאורע "סכום הקוביות שהוטלו זוגי", B המאורע "סכום הקוביות שהוטלו מתחלק ב-3", ו-C המאורע "סכום הקוביות שהוטלו מתחלק ב-5".

א. האם A ו-B תלויים?

ב. האם A ו-C תלויים?

פתרון. נחשב את ההסתברויות. כדי שסכום הקוביות יהיה זוגי, לכל הטלה של אחת הקוביות יש 3 הטלות אפשריות של הקובייה השנייה כך שהסכום זוגי. לכן

$$P(A) = \frac{6 \cdot 3}{36} = \frac{1}{2}.$$

בדומה

$$P(B) = \frac{6 \cdot 2}{36} = \frac{1}{3}.$$

לגבי החיתוך, רוצים שסכום הקוביות יתחלק ב-6, ופה לכל הטלה של קובייה אחת יש בדיוק תוצאה אחת של הקובייה השנייה כך שסכומן יתחלק ב-6. לכן

$$P(A \cap B) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = P(A) \cdot P(B).$$

מכאן ש-A ו-B בלתי־תלויים.

עבור A ו-C, נחשב את $P\left(C\right)$: לכל הטלה למעט 4 של הקובייה הראשונה, יש הטלה אחת בלבד של הקובייה השנייה שסכומן יתחלק ב-5. לעומת זאת, להטלה 4 יש שתי אפשרויות כאלו. לכן

$$P\left(C\right)=\frac{7}{36}.$$

כעת $A\cap C$ הוא המאורע שסכום הקוביות מתחלק ב-10, כלומר הוא חייב להיות 10. לכן

$$P\left(A\cap C\right)=\frac{3}{36}=\frac{1}{12}\neq\frac{1}{2}\cdot\frac{7}{36}=P\left(A\right)\cdot P\left(C\right).$$

4.3 תרגילים נוספים

תרגיל 4.12. בקבוצה של קייטנת היצורונים יש 80 טרולים וכמות לא ידועה של דרדסים. 75% מהטרולים הם בריונים, ו-25% מהדרדסים הם בריונים אף הם. בוחרים באקראי משתתף בקבוצה, ואם מדובר בדרדס בריון – הוא נשלח חזרה לכפר. ידוע כי ההסתברות לתרחיש הזה בבחירת המשתתף האקראי הראשון היא 0.05.

א. האם בכיתת הקייטנה יש תלות בין סוג המשתתף לבין האם הוא בריון או לא?

ב. כמה דרדסים יש בקבוצה?

ג. מה ההסתברות שמשתתף הנבחר באקראי אינו בריון?

ד. בקייטנה כולה יש 1000 דרדסים, ובכל בוקר כל אחד מהם בוחר באופן אקראי האם להיות בריון (בסיכוי p) או לא, באופן בלתי־תלוי באחרים. לדרדסבא יש שני דרדסונים בקייטנה, והוא קיבל טלפון להגיע לקייטנה כיוון שלדרדסים אסור להיות בריונים. מה ההסתברות ששני הדרדסונים של דרדסבא הם בריונים ביום הזה?

פתרון.

א. יש תלות. אילו לא הייתה תלות, ההסתברות להיות בריון לא הייתה משתנה אם מדובר בטרול או בדרדס, אך לפי הנתונים אנו רואים שהיא כן שונה.

ב. נסמן את כמות הדרדסים בקבוצה על ידי n. נסמן A = המשתתף שבוחרים הוא טרול, B = המשתתף שנבחר הוא בריון. לפי הנתונים, נקבל $P\left(B\mid A\right)=0.75$, $P\left(B\mid A^{c}\right)=0.25$, $P\left(A\right)=\frac{80}{n+80}$, $P\left(A^{c}\right)=\frac{n}{n+80}$, $P\left(A^{c}\cap B\right)=0.05$. לפי ההגדרה, $P\left(B\mid A^{c}\right)=\frac{P\left(A^{c}\cap B\right)}{P\left(A^{c}\right)}$, ולכן נקבל את המשוואה $0.25=\frac{0.05}{\frac{n}{n+80}}$. נפתור:

$$0.25n=0.05\left(n+80\right)=0.05n+4\Longrightarrow0.2n=4\Longrightarrow n=20.$$

ג. רוצים לחשב את $P\left(B^{c}\right)$. ניעזר בנוסחת ההסתברות השלמה:

$$\begin{aligned}P\left(B^{c}\right)=1-P\left(B\right)=1-P\left(B\mid A\right)\cdot P\left(A\right)-P\left(B\mid A^{c}\right)\cdot P\left(A^{c}\right)=\\=1-0.75\cdot0.8-0.25\cdot0.2=0.35.\end{aligned}$$

ד. כעת ידוע לנו שלפחות אחד משני הדרדסונים הוא בריון, והשאלה היא מה ההסתברות ששניהם בריונים. מרחב המדגם שלנו הוא

$$\Omega=\left\{(\text{בריון},\text{בריון}),(\text{בריון},\text{לא בריון}),(\text{לא בריון},\text{בריון}),(\text{לא בריון},\text{לא בריון})\right\},$$

וההסתברות היא

$$P\left(\text{לפחות אחד בריון} \mid \text{שניהם בריונים}\right) = \frac{P\left(\left\{\left(\text{בריון},\text{בריון}\right)\right\}\right)}{P\left(\left\{\left(\text{בריון},\text{בריון}\right),\left(\text{בריון},\text{לא בריון}\right),\left(\text{לא בריון},\text{בריון}\right)\right\}\right)} =$$
$$= \frac{p^2}{p^2+2p\left(1-p\right)} = \frac{p}{p+2\left(1-p\right)} = \frac{p}{2-p}.$$

תרגיל 4.13. התדירות של נשאות HIV באוכלוסיית ארה"ב הינה 0.0013. בגיוס לצבא האמריקני עושים שלושה סבבים של בדיקה מסוימת לנשאות. ה-TPR (true positive rate) של הבדיקה הינו 92%, וה-TNR (true negative rate) של הבדיקה הינו 99%, והדיוק בכל בדיקה בלתי־תלוי בדיוק של הבדיקות האחרות.

א. מה ההסתברות שאדם שקיבל תשובה חיובית בבדיקה הראשונה הינו נשא של HIV?

ב. מה ההסתברות שאדם שיצא חיובי גם בבדיקה השנייה הינו נשא של HIV?

ג. מה ההסתברות שאדם שאיננו נשא יקבל שלוש תשובות חיוביות?

פתרון. נסמן על ידי A את המאורע שהאדם הוא נשא, ועל ידי B_1,B_2,B_3 את המאורעות שיקבל תשובה חיובית בבדיקות המתאימות. לפי הנתון, $P\left(A\right)=0.0013$, $P\left(B_i\mid A\right)=0.92$, $P\left(B_i^c\mid A^c\right)=0.99$.

א. נחשב:

$$P\left(A\mid B_1\right)=\frac{P\left(A\cap B_1\right)}{P\left(B_1\right)}=\frac{P\left(B_1\mid A\right)\cdot P\left(A\right)}{P\left(B_1\mid A\right)\cdot P\left(A\right)+P\left(B_1\mid A^c\right)\cdot P\left(A^c\right)}=$$
$$=\frac{\text{TPR}\cdot f}{\text{TPR}\cdot f+\text{TNR}\cdot\left(1-f\right)}=$$
$$=\frac{0.92\cdot 0.0013}{0.92\cdot 0.0013+0.01\cdot 0.9987}=0.10695$$

ב. בדומה,

$$P\left(A\mid B_1\cap B_2\right)=\frac{P\left(A\cap B_1\cap B_2\right)}{P\left(B_1\cap B_2\right)}=\frac{P\left(B_1\cap B_2\mid A\right)\cdot P\left(A\right)}{P\left(B_1\cap B_2\mid A\right)\cdot P\left(A\right)+P\left(B_1\cap B_2\mid A^c\right)\cdot P\left(A^c\right)}=$$
$$=\frac{0.92^2\cdot 0.0013}{0.92^2\cdot 0.0013+0.01^2\cdot 0.9987}=0.91679$$

ג. כעת $P\left(B_1\cap B_2\cap B_3\mid A^c\right)=\left(1-\text{TNR}\right)^3=0.01^3=10^{-6}$.

5 תרגול חמישי

5.1 משתנים מקריים

המטרה. להיות מסוגלים "לכמת" את אוסף כל התוצאות של ניסוי מסוים. במקום להסתכל כל פעם על כל תוצאה בנפרד כמאורע ("האם יצא 1 בהטלת קוביה", "האם יצא 2 בהטלת קוביה", ...), נרצה להסתכל על אובייקט אחד שירכז לנו את כל המידע שמעניין אותנו מהניסוי.

הגדרה 5.1. יהי $(\Omega,\mathcal{F},P)$ מרחב הסתברות. **משתנה מקרי על** Ω הוא פונקציה $X:\Omega\to\mathbb{R}$.
נסמן $\{X\in S\}=X^{-1}(S)$; למשל, $\{X\le 1\}=\{\omega\in\Omega\,|\,X(\omega)\le 1\}$.

דוגמה 5.2.

א. הטלת מטבע: $\Omega=\{\text{פלי,עץ}\}$ אפשר לייצג על ידי משתנה מקרי $X(\text{עץ})=0$ ו-$X(\text{פלי})=1$.

ב. הטלת קוביה: $\Omega=\{1,\ldots,6\}$ אפשר לייצג על ידי $X(\omega)=\omega$.

ג. הטלת שני מטבעות: $\Omega=\{(0,0),(0,1),(1,0),(1,1)\}$. פה השאלה תהיה מה אנחנו רוצים למדוד. אפשר למשל להגדיר את המשתנה המקרי $X(a,b)=a+b$ שיספור כמה פעמים יצא "פלי" בהטלה; אפשר מצד שני להגדיר את המשתנה המקרי $X(a,b)=\begin{cases}1, & a=b\\ 0, & a\neq b\end{cases}$ שימדוד האם שתי ההטלות זהות זו לזו.

5.2 פונקציית ההתפלגות של משתנה מקרי

הגדרה 5.3. יהי $X:\Omega\to\mathbb{R}$ משתנה מקרי. **פונקציית ההתפלגות של** X היא הפונקציה $P_X(a)=P(X=a)$. שימו לב שהסימון $X=a$ הוא מבחינתנו קיצור למאורע

$$\{X=a\}=\{\omega\in\Omega\mid X(\omega)=a\}\,.$$

דוגמה 5.4.

א. פונקציית ההתפלגות של הטלת קוביה: $P_X(k)=\frac{1}{6}$ לכל $1\le k\le 6$.

ב. פונקציית ההתפלגות של סכום הטלות שתי קוביות: $P_X(2)=\frac{1}{36}$, $P_X(3)=\frac{2}{36}$ וכן הלאה. באופן כללי, לכל $2\le k\le 12$ מתקיים

$$P_X(k)=\begin{cases}\frac{k-1}{36}, & k\le 7\\ \frac{13-k}{36}, & k>7\end{cases}.$$

תרגיל 5.5. נסתכל על הטלת שלושה מטבעות $\Omega=\{(0,0,0),(0,0,1),(0,1,0),\ldots\}$. $|\Omega|=8$. נגדיר את המשתנה המקרי X הסופר כמה פעמים יצא 1, ו-Y כמה פעמים יצא 0.

א. מהי $P(X=2)$?

ב. מהי $P(X<2)$?

ג. האם X ו-Y שווים?

ד. האם X ו-Y שווי התפלגות?

פתרון.

א. $P(X=2)=P(\{(1,1,0),(1,0,1),(0,1,1)\})=\frac{3}{8}$.

ב. $P(X<2)=P(\{(0,0,0),(1,0,0),(0,1,0),(0,0,1)\})=\frac{1}{2}$.

ג. המשתנים אינם שווים כי אפשר למצוא מאורעות שהם ייתנו ערכים שונים בהם. למשל, ניקח את התוצאה $\omega = (1,1,1)$. אז $X(\omega) = 3$ אבל $Y(\omega) = 0$.

ד. כן. פונקציית ההתפלגות שלהם היא

$$P_X(a) = P_Y(a) = \begin{cases} \frac{1}{8}, & a = 0,3 \\ \frac{3}{8}, & a = 1,2 \end{cases}.$$

המסר של התרגיל האחרון: יכולים להיות משתנים מקריים על אותו מרחב מדגם, או על מרחבי מדגם שונים, שיהיו שונים זה מזה אבל בעלי אותה התפלגות. כאשר אנחנו באים ללמוד את המשתנה המקרי, בהרבה מקרים כל מה שיעניין אותנו הוא ההתפלגות שלו – אפשר יהיה לשכוח ממרחב המדגם, ורק לשאול את עצמנו מה ההתפלגות של המשתנה המקרי.

תרגיל 5.6. משתנה מקרי X מקבל את הערכים $0,1,2,3,\ldots$ בהסתברות $P_X(i) = \frac{c}{3^i}$.

א. מצאו את c.

ב. מצאו את $P(X > 5)$.

ג. מצאו את $P(\text{זוגי } X)$.

פתרון.

א. נשים לב כי $\sum_{i=0}^{\infty} P_X(i) = 1$. מפה נקבל את המשוואה

$$1 = \sum_{i=0}^{\infty} \frac{c}{3^i} = c \cdot \frac{1}{1-\frac{1}{3}} = \frac{3c}{2}$$

ולכן $c = \frac{2}{3}$.

ב. מחישוב ישיר,

$$P(X > 5) = \sum_{i=6}^{\infty} \frac{2}{3} \cdot \frac{1}{3^i} = \frac{2}{3} \cdot \frac{\frac{1}{3^6}}{1-\frac{1}{3}} = \frac{1}{3^6}.$$

ג. נחשב:

$$P(\text{זוגי } X) = \sum_{i=0}^{\infty} P_X(2i) = \sum_{i=0}^{\infty} \frac{2}{3} \cdot \left(\frac{1}{3}\right)^{2i} = \frac{2}{3} \cdot \frac{1}{1-\frac{1}{9}} = \frac{2}{3} \cdot \frac{9}{8} = \frac{3}{4}.$$

תרגיל 5.7. מתוך קבוצה של 100 נבדקים שמתוכם 12 חולים בוחרים באקראי 10 אנשים ובודקים אותם. יהי X המשתנה המקרי המתאר את מספר החולים מתוך ה-10 שנבדקו.

א. תארו את מרחב המדגם ואת ההתפלגות של X.

ב. מהו הערך השכיח ביותר ש-X מקבל?

פתרון.

א. מרחב המדגם Ω הוא אוסף כל הקבוצות מגודל 10 של קבוצת הנבדקים. לכן $|\Omega| = \binom{100}{10}$. פונקציית ההתפלגות של X היא

$$.P\left(X=k\right)=\frac{\binom{12}{k}\cdot\binom{88}{10-k}}{\binom{100}{10}}$$

ב. כדי לחשב את הערך השכיח ביותר, נחשב את המנות $q_k=\frac{P(X=k+1)}{P(X=k)}$:

$$q_k=\frac{\frac{\binom{12}{k+1}\cdot\binom{88}{10-k-1}}{\binom{100}{10}}}{\frac{\binom{12}{k}\cdot\binom{88}{10-k}}{\binom{100}{10}}}=\frac{\binom{12}{k+1}\cdot\binom{88}{10-k-1}}{\binom{12}{k}\cdot\binom{88}{10-k}}=\frac{\frac{12!}{(k+1)!(11-k)!}\cdot\frac{88!}{(9-k)!(79+k)!}}{\frac{12!}{k!(12-k)!}\cdot\frac{88!}{(10-k)!(78+k)!}}=$$
$$=\frac{k!}{(k+1)!}\cdot\frac{(12-k)!}{(11-k)!}\cdot\frac{(10-k)!}{(9-k)!}\cdot\frac{(78+k)!}{(79+k)!}=$$
$$=\frac{\left(12-k\right)\left(10-k\right)}{\left(k+1\right)\left(79+k\right)}=\frac{120-22k+k^2}{79+80k+k^2}$$

לכן $q_k>1$ אם ורק אם $120-22k+k^2>79+80k+k^2$, אם ורק אם $102k<41$, שקורה אם ורק אם $k=0$. לכן הערך השכיח ביותר הוא 1.

5.3 התפלגות משותפת

הגדרה 5.8. יהיו X ו-Y משתנים מקריים. נגדיר את **פונקציית ההתפלגות המשותפת** שלהם להיות

$$.P_{X,Y}\left(a,b\right)=P\left(X=a,Y=b\right)$$

ההתפלגויות של X ושל Y נקראות **ההתפלגויות השוליות**.

דוגמה 5.9. מטילים שתי קוביות סטנדרטיות. נסמן על ידי X ו-Y את המשתנה המקרי המתאים לכל קוביה. ההתפלגות המשותפת היא

$$P_{X,Y}\left(a,b\right)=\frac{1}{36}$$

לכל $1\leq a,b\leq 6$ טבעיים. ההתפלגות השולית של X היא $P_X\left(a\right)=\frac{1}{6}$, וכנ"ל ההתפלגות של Y.

הגדרה 5.10. אומרים ש-X ו-Y הם **בלתי־תלויים**, אם לכל a ו-b מתקיים

$$.P_{X,Y}\left(a,b\right)=P_X\left(a\right)\cdot P_Y\left(b\right)$$

כלומר, לכל a,b, המאורעות $\{X=a\}$ ו-$\{Y=b\}$ הם בלתי־תלויים.

תרגיל 5.11. בכד יש 5 כדורים לבנים ו-10 כדורים אדומים. מטילים מטבע פעמיים, ולפי כמות הפעמים שיצא "עץ" שולפים את אותה כמות כדורים. נסמן על ידי X את מספר העצים שיצאו, ועל ידי Y את מספר הכדורים הלבנים שנשלפו.

א. מהי ההתפלגות המשותפת של X ו-Y?

ב. מהן ההתפלגויות השוליות של X ושל Y?

ג. האם X ו-Y תלויים?

ד. חשבו את $P\left(1\leq X\leq 2, 1\leq Y\leq 2\right)$.

ה. מהי $P\left(X=1\mid Y=1\right)$?

ו. מהי $P\left(X>Y\right)$?

פתרון. הכי קל בשאלות כאלו הוא להשתמש בטבלה.

$Y\backslash X$	0	1	2	$P_Y\left(b\right)$
0	$\frac{1}{4}$	$\frac{1}{2}\cdot\frac{2}{3}=\frac{1}{3}$	$\frac{1}{4}\cdot\frac{2}{3}\cdot\frac{9}{14}=\frac{3}{28}$	$\frac{29}{42}$
1	0	$\frac{1}{2}\cdot\frac{1}{3}=\frac{1}{6}$	$\frac{1}{4}\cdot\left(\frac{1}{3}\cdot\frac{10}{14}+\frac{2}{3}\cdot\frac{5}{14}\right)=\frac{5}{42}$	$\frac{12}{42}=\frac{2}{7}$
2	0	0	$\frac{1}{4}\cdot\frac{1}{3}\cdot\frac{4}{14}=\frac{1}{42}$	$\frac{1}{42}$
$P_X\left(a\right)$	$\frac{1}{4}$	$\frac{1}{3}+\frac{1}{6}=\frac{1}{2}$	$\frac{1}{4}$	1

לכן:

א. בטבלה.

ב. בטבלה.

ג. כן! שהרי $P_{X,Y}\left(0,0\right)=\frac{1}{4}\neq\frac{29}{42}\cdot\frac{1}{4}=P_X\left(0\right)\cdot P_Y\left(0\right)$.

ד. $P\left(1\leq X\leq 2, 1\leq Y\leq 2\right)=\sum_{a=1}^{2}\sum_{b=1}^{2}P_{X,Y}\left(a,b\right)=\frac{13}{42}$.

ה. $P\left(X=1\mid Y=1\right)=\frac{P(X=1,Y=1)}{P(Y=1)}=\frac{\frac{1}{6}}{\frac{2}{7}}=\frac{7}{12}$

ו. נחשב:

$$\begin{aligned}P\left(X>Y\right) &= \sum_{a=1}^{2}\sum_{b=0}^{a-1}P_{X,Y}\left(a,b\right)= \\ &= P_{X,Y}\left(1,0\right)+P_{X,Y}\left(2,0\right)+P_{X,Y}\left(2,1\right)=\frac{47}{84}\end{aligned}$$

תרגיל 5.12. מגרילים n ביטים באקראי באופן בלתי־תלוי. יהי X המשתנה המקרי הסופר את מספר המופעים של 01 במחרוזת שהתקבלה, ו-Y מספר המופעים של 10 במחרוזת שהתקבלה. בדקו האם X,Y תלויים או בלתי־תלויים עבור $n=2,3,4$.

פתרון. עבור $n=2$, המשתנים X ו-Y יכולים לקבל את הערכים $0,1$.

$$P_{X,Y}\left(0,0\right)=P\left(\{00,11\}\right)=\frac{1}{2}$$
$$P_X\left(0\right)\cdot P_Y\left(0\right)=\frac{3}{4}\cdot\frac{3}{4}=\frac{9}{16}$$

ולכן יש תלות.

עבור $n=3$, נרשום את כל המחרוזות בטבלה ונראה מהם הערכים של X ושל Y:

	000	001	010	011	100	101	110	111
X	0	1	1	1	0	1	0	0
Y	0	0	1	0	1	1	1	0

לכן אפשר לכתוב את טבלת ההתפלגות המשותפת:

$Y\backslash X$	0	1	$P_Y(b)$
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$P_X(a)$	$\frac{1}{2}$	$\frac{1}{2}$	1

מפה אפשר לראות שהם בלתי־תלויים.

עבור $n=4$, שוב יש תלות. אכן, $P_{X,Y}(2,2)=0\neq P_X(2)\cdot P_Y(2)$.

תרגיל 5.13 (אם יש זמן)**.** בדוגמה הקודמת, לאילו ערכי n המשתנים X ו-Y בלתי־תלויים?

פתרון. נניח כי X ו-Y בלתי־תלויים. בפרט

$$P_{X,Y}(0,0)=P_X(0)\cdot P_Y(0).$$

נחשב כל אחת מההסתברויות. עבור $P_X(0)$, צריך שלא יופיע הרצף 01; המחרוזות היחידות האפשריות באורך n שלא מופיע בהן 01 הן מחרוזות קבועות או מחרוזות שמתחילות ב-1, מתישהו מתחלפות ל-0 ולא מתחלפות חזרה ל-1. לכן $P_X(0)=\frac{2+n-1}{2^n}=\frac{n+1}{2^n}$. באופן סימטרי, $P_Y(0)=\frac{n+1}{2^n}$.

לעומת זאת, אם רוצים לחשב את $P_{X,Y}(0,0)$, רוצים לספור בכמה מחרוזות באורך n לא מופיעים הרצפים 01 ו-10. יש בדיוק שתי מחרוזות כאלו – המחרוזות הקבועות. לכן $P_{X,Y}(0,0)=\frac{2}{2^n}=\frac{1}{2^{n-1}}$.

כדי ש-X ו-Y יהיו בלתי־תלויים, נהיה חייבים שיתקיים

$$\frac{2}{2^n}=\left(\frac{n+1}{2^n}\right)^2,$$

כלומר $2^{n+1}=(n+1)^2$.

אפשר לוודא שלמשוואה $2^x=x^2$ יש שני פתרונות יחידים שהם $x=2,4$, ולכן המשתנים יכולים להיות בלתי־תלויים רק אם $n=1,3$. ואכן לשני הערכים האלו המשתנים בלתי־תלויים (עבור $n=1$ שניהם שווים תמיד 0, ועבור $n=3$ הראינו בתרגיל הקודם שהם בלתי־תלויים).

6 תרגול שישי

6.1 התפלגויות נפוצות

הגדרה 6.1. התפלגות אחידה בדידה – משתנה מקרי X המקבל ערכים מקבוצה סופית S, כאשר לכל ערך יש סיכוי זהה להתקבל. לרוב הקבוצה S היא קבוצת מספרים שלמים עוקבים $\{a,a+1,\dots,b\}$, ואז מסמנים $X\sim U[a,b]$. פונקציית ההתפלגות של X היא

$$P_X(k)=\frac{1}{b-a+1}$$

לכל $k\in\{a,a+1,\dots,b\}$.

תרגיל 6.2. יהי X משתנה מקרי אחיד על הקבוצה $\{4, 5, \dots, 49\}$.

א. מה הסיכוי ש-X יקבל את הערך 17?

ב. מה הסיכוי שהערך של X יתחלק ב-5?

פתרון.

א. בקבוצה הזו יש $49 - 4 + 1 = 46$ איברים, ולכן ההסתברות היא $\frac{1}{46}$.

ב. הערכים בקבוצה שמתחלקים ב-5 הם $\{5, 10, \dots, 45\}$. בסך הכל יש 9 ערכים כאלה, ולכן הסיכוי ש-X יתחלק ב-5 הוא $\frac{9}{46}$.

הגדרה 6.3. יהי A מאורע. **משתנה האינדיקטור של** A, מסומן $X = 1_A$, הוא המשתנה הבא:

$$1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}.$$

זאת אומרת שהוא בדיוק מייצג את המאורע A – שואל האם התוצאה שלנו ממרחב המדגם נמצאת בתוך המאורע או לו.

פונקציית ההתפלגות של משתנה אינדיקטור היא

$$P(1_A = a) = \begin{cases} P(A^c), & a = 0 \\ P(A), & a = 1 \end{cases}.$$

שימו לב שכל משתני האינדיקטור שעבורם $P(A) = p$ לאיזשהו p קבוע הם שווי התפלגות, וזו נקראת התפלגות ברנולי:

הגדרה 6.4. התפלגות ברנולי – משתנה מקרי X המקבל שני ערכים: 0 או 1 (כישלון או הצלחה). מסמנים $X \sim \mathrm{Ber}(p)$ או $X \sim \mathrm{b}(p)$, כאשר p היא ההסתברות לקבל את הערך 1 (הסתברות ההצלחה בניסוי).

תרגיל 6.5. נתון מרחב מדגם ובו שני מאורעות A ו-B. נסתכל על המשתנים המקריים 1_A ו-1_B. הוכיחו ש-$1_A 1_B \sim \mathrm{Ber}(p)$ כאשר $p = P(A)P(B)$ אם ורק אם A ו-B בלתי־תלויים.

הוכחה. נשים לב כי המשתנה $1_A 1_B$ מקבל גם הוא רק שתי אפשרויות לערכים – 0 ו-1. לכן הוא צריך להיות משתנה ברנולי. של איזה מאורע?

$$(1_A 1_B)(\omega) = 1 \iff 1_A(\omega) 1_B(\omega) = 1 \iff 1_A(\omega) = 1_B(\omega) = 1 \iff \omega \in A \cap B.$$

לכן $1_A 1_B = 1_{A\cap B}$. מכאן שההתפלגות שלו היא $1_A 1_B \sim \mathrm{Ber}(P(A \cap B))$, וההסתברות הזו שווה ל-$P(A)P(B)$ אם ורק אם A ו-B בלתי־תלויים. $\square$

הגדרה 6.6. התפלגות בינומית – מבצעים סדרה של n ניסויים זהים ובלתי־תלויים עם הסתברות הצלחה p, וסופרים את כמות ההצלחות שלהם. מסמנים $X \sim \mathrm{Bin}(n,p)$ אז פונקציית ההתפלגות היא

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

לכל $k \in \{0, \dots, n\}$.

באופן שקול: יש לנו n משתני ברנולי בלתי־תלויים עם סיכוי הצלחה p, ומסתכלים על ההתפלגות של הסכום שלהם.

תרגיל 6.7. מתוך סקר שנערך בין תושבי עיר מסוימת התברר כי 45% מהתושבים היו בעד פתיחת הקניון בעיר, 30% נגד ול-25% אין דעה בעניין. בוחרים באקראי ארבעה אנשים מתושבי העיר (הניחו שהדעה של כל אחד נקבעת באופן אקראי לפי ההסתברויות הנ"ל באופן בלתי־תלוי באחרים). חשבו את ההסתברויות הבאות:

א. ארבעתם בעד פתיחת הקניון.

ב. שלפחות אחד יהיה נגד הפתיחה.

ג. שלכל אחד מהם תהיה דעה.

פתרון.

א. תהי X כמות האנשים שבעד. אז $X \sim \operatorname{Bin}(4, 0.45)$, ולכן

$$P(X=4) = \binom{4}{4} 0.45^4 = 0.041.$$

ב. תהי Y כמות האנשים שנגד. אז $Y \sim \operatorname{Bin}(4, 0.3)$, ולכן

$$P(Y \geq 1) = 1 - P(Y=0) = 1 - \binom{4}{0} 0.7^4 = 0.7599.$$

ג. תהי Z כמות האנשים עם דעה. אז $Z \sim \operatorname{Bin}(4, 0.75)$, ולכן

$$P(Z=4) = \binom{4}{4} 0.75^4 = 0.316.$$

תרגיל 6.8. משחקים במשחק הבא: מטילים n מטבעות. את אלו שנפלו על "עץ" שומרים, ואת היתר מחזירים לקופה. מטילים את אלה ששמרנו שוב, והפעם שומרים רק את אלו שנפלו על "פלי" (ואת היתר מחזירים). כיצד מתפלג מספר המטבעות שבידינו בסוף המשחק?

פתרון. אילו עצרנו אחרי השלב הראשון, המספר הזה היה מתפלג $\operatorname{Bin}\left(n, \frac{1}{2}\right)$. נרצה לומר שאחרי שני השלבים המספר מתפלג $\operatorname{Bin}\left(n, \frac{1}{4}\right)$. אכן, נדמיין שאנחנו זורקים כל מטבע פעמיים, גם את אלו שנפלו בהתחלה על "פלי", ושומרים רק את אלו שיצאו בהתחלה "עץ" ואחר כך "פלי". ברור שמספר המטבעות שבידינו זהה לניסוי הזה. כעת אנחנו מבצעים ניסוי n פעמים (הטלת מטבע) שסיכוי הצלחתו הוא $\frac{1}{4}$, ולכן התשובה היא אכן $\operatorname{Bin}\left(n, \frac{1}{4}\right)$.

תרגיל 6.9. כיתה מחולקת לשלוש קבוצות, ובכל קבוצה שלושה תלמידים. כל תלמיד בוחר אם הוא מעדיף ללמוד מתמטיקה (בהסתברות p) או פיזיקה (בהסתברות $1-p$). כל קבוצה בוחרת את נושא הלימוד המועדף עליה על פי רוב, והכיתה בוחרת את נושא הלימוד לפי מה שרוב קבוצות הלמידה החליטו.

מה ההסתברות שהכיתה תלמד מתמטיקה? העריכו את ההסתברות כפונקציה של $p = \frac{1+\varepsilon}{2}$ כאשר $\varepsilon > 0$ קטן מאוד.

פתרון. ההסתברות של קבוצה לבחור במתמטיקה היא

$$P(\operatorname{Bin}(3, p) \geq 2) = 3p^2(1-p) + p^3 = 3p^2 - 2p^3 = p^2(3-2p).$$

לכן ההסתברות של הכיתה לבחור במתמטיקה היא

$$.P\left(\mathrm{Bin}\left(3,p^2\left(3-2p\right)\right)\geq 2\right)=\left(p^2\left(3-2p\right)\right)^2\left(3-2p^2\left(3-2p\right)\right)$$

כאשר $p=\frac{1+\varepsilon}{2}$, אפשר לקרב $p^2\approx\frac{1+2\varepsilon}{4}$ (מזניחים חזקות גבוהות של ε), ואז

$$p^2\left(3-2p\right)\approx\frac{1+2\varepsilon}{4}\left(2-\varepsilon\right)\approx\frac{2+3\varepsilon}{4}=\frac{1+\frac{3}{2}\varepsilon}{2}$$

ובהצבה פעמיים מתקבל $\frac{1+\frac{9}{4}\varepsilon}{2}\approx$.

לדוגמה, אם כל תלמיד בוחר מתמטיקה בסיכוי 0.51, הכיתה כולה תלמד מתמטיקה בסיכוי $0.5225\approx$ (קרוב לתשובה האמיתית עד כדי הספרה החמישית אחרי הנקודה).

הגדרה 6.10. התפלגות גיאומטרית – מבצעים ניסוי עם הסתברות הצלחה p, וסופרים את כמות הפעמים עד להצלחה הראשונה. מסמנים $X\sim\mathrm{Geo}\left(p\right)$ אז פונקציית ההתפלגות היא

$$P\left(X=k\right)=\left(1-p\right)^{k-1}\cdot p$$

לכל $k\in\{1,2,\dots\}$.

טענה 6.11 (תכונות של התפלגות גיאומטרית). אם $X\sim\mathrm{Geo}(p)$, אז מתקיים

א. $P\left(X>k\right)=\left(1-p\right)^k$.

ב. X מקיים את **תכונת חוסר הזיכרון**, כלומר $P\left(X=m+n\,|\,X>n\right)=P\left(X=m\right)$.

תרגיל 6.12. תייר מחכה לאוטובוס שייקח אותו למלון. הסיכוי שבאוטובוס יהיה מקום פנוי הוא $\frac{2}{3}$, ללא תלות באוטובוסים האחרים.

א. מה הסיכוי שהתייר לא יעלה על ארבעת האוטובוסים הראשונים?

ב. ידוע שהתייר לא עלה על ארבעת האוטובוסים הראשונים. מה הסיכוי שהוא יעלה על האוטובוס השישי?

פתרון. נסמן על ידי X את כמות האוטובוסים עד שמגיע אוטובוס עם מקום פנוי (כולל). אז $X\sim\mathrm{Geo}\left(\frac{2}{3}\right)$.

א. $P\left(X>4\right)=\left(1-\frac{2}{3}\right)^4=\frac{1}{3^4}=\frac{1}{81}$.

ב. לפי תכונת חוסר הזיכרון, $P\left(X=6\,|\,X>4\right)=P\left(X=2\right)=\frac{2}{3}\cdot\left(\frac{1}{3}\right)^{2-1}=\frac{2}{9}$.

תרגיל 6.13. לפניכם 10 קוביות. 9 מהן הוגנות, ואחת מהן מוטה, והסיכוי שלה ליפול על 6 הוא $\frac{1}{20}$ (ושאר התוצאות בסיכוי $\frac{19}{100}$). החלטתם לבצע את המבחן הבא: תיקחו קוביה, תטילו אותה עד שתקבלו 6, ואם לוקח לכם יותר מ-12 הטלות עד אז – תקבעו שהקוביה מזויפת.

א. אם נבחר קוביה אקראית, מה הסיכוי שייקח לה יותר מ-12 הטלות עד שנקבל 6?

ב. אם לקח לקוביה יותר מ-12 הטלות עד שנקבל 6, מה הסיכוי שהיא המזויפת?

פתרון. נסמן על ידי A את המאורע שבחרנו קוביה מזויפת, ועל ידי X את כמות ההטלות עד שנקבל 6.

א. לפי נוסחת ההסתברות השלמה,

$$.P\left(X>12\right)=P\left(X>12\,|\,A\right)\cdot P\left(A\right)+P\left(X>12\,|\,A^{c}\right)\cdot P\left(A^{c}\right)$$

נשים לב כי $P\left(A\right)=\frac{1}{10}$. אם מניחים שהקוביה מזויפת, אז $X\sim\text{Geo}\left(\frac{1}{20}\right)$, ולכן $P\left(X>12\,|\,A\right)=\left(1-\frac{1}{20}\right)^{12}=\left(\frac{19}{20}\right)^{12}$. מצד שני, אם הקוביה הוגנת, אז $X\sim\text{Geo}\left(\frac{1}{6}\right)$, ולכן $P\left(X>12\,|\,A^{c}\right)=\left(1-\frac{1}{6}\right)^{12}=\left(\frac{5}{6}\right)^{12}$. בסך הכל

$$.P\left(X>12\right)=\left(\frac{19}{20}\right)^{12}\cdot\frac{1}{10}+\left(\frac{5}{6}\right)^{12}\cdot\frac{9}{10}=0.155$$

ב. כעת רוצים לחשב את $P\left(A\,|\,X>12\right)$. לפי חוק בייס,

$$.P\left(A\,|\,X>12\right)=\frac{P\left(X>12\,|\,A\right)\cdot P\left(A\right)}{P\left(X>12\right)}=\frac{\left(\frac{19}{20}\right)^{12}\cdot\frac{1}{10}}{0.155}=0.349$$

תרגיל 6.14. בר ואילן החליטו להתחרות ביניהם: שניהם מנסים לקלוע לסל, ומנצח מי שקולע ראשון. שניהם לא שיחקו כדורסל כבר כמה שנים, אז הסיכוי של כל אחד מהם לקלוע בניסיון בודד הוא p, ואינו תלוי בניסיונות האחרים או זה בזה. מה הסיכוי שיהיה תיקו (כלומר ששניהם יקלעו בפעם הראשונה ביחד)?

פתרון. נסמן על ידי X את כמות הניסיונות עד שבר מצליח לקלוע לראשונה, ועל ידי Y את כמות הניסיונות עד שאילן מצליח לקלוע לראשונה. לפי הנתונים, $X\sim\text{Geo}(p)$ ו-$Y\sim\text{Geo}(p)$ הם בלתי־תלויים. לפי נוסחת ההסתברות השלמה,

$$\begin{aligned}P\left(X=Y\right)&=\sum_{n=1}^{\infty}P\left(X=Y,Y=n\right)=\sum_{n=1}^{\infty}P\left(X=n,Y=n\right)=\\&=\sum_{n=1}^{\infty}P\left(X=n\right)P\left(Y=n\right)=\sum_{n=1}^{\infty}\left(p\left(1-p\right)^{n-1}\right)^{2}=\\&=p^{2}\sum_{n=1}^{\infty}\left(\left(1-p\right)^{2}\right)^{n-1}=\frac{p^{2}}{1-\left(1-p\right)^{2}}=\frac{p^{2}}{2p-p^{2}}=\frac{p}{2-p}\end{aligned}$$

7 תרגול שביעי

הגדרה 7.1. התפלגות פואסון – מסמנים $X\sim\text{Poi}\left(\lambda\right)$. אז פונקציית ההתפלגות היא

$$P\left(X=k\right)=e^{-\lambda}\cdot\frac{\lambda^{k}}{k!}$$

לכל $k\in\left\{0,1,2,\dots\right\}$.

דוגמה 7.2. מה ההתפלגות מתארת?

- מספר המופעים של תופעה מסוימת המתרחשים בפרק זמן נתון, כאשר מניחים שהזמן בין מופע למופע הוא "חסר זיכרון":

– מספר האנשים המתווספים לתור בשעה;
– מספר שיחות הטלפון למרכזייה בדקה;
– כמות שגיאות ההקלדה בעמוד מודפס בעיתון;
– מספר החיילים בגדוד של חיל הפרשים הפרוסי הנהרגים בשנה מבעיטת סוס.

- דרך אחרת לחשוב עליה: קירוב להתפלגות בינומית עם פרופורציה קבועה: אם λ קבוע, אז $\mathrm{Bin}\left(n, \frac{\lambda}{n}\right) \to \mathrm{Poi}\left(\lambda\right)$.

תרגיל 7.3. במאפייה מערבבים פצפוצי שוקולד עם $1,000$ ק"ג בצק עוגיות כך שנוצרת עיסה במשקל טון להכנת עוגיות שוקולד צ'יפס (משקל הפצפוצים זניח). כמות הפצפוצים המעורבבת מתפלגת פואסונית, כשבממוצע נכנסים $10,000$ פצפוצים על כל טון בצק. העיסה עורבבה היטב כך שהפצפוצים מפוזרים באופן אחיד (עד כדי סטייה הסתברותית) בנפח הבצק. כעת מכינים מהבצק 100 חבילות עוגיות במשקל 1 ק"ג כל אחת.

א. מהי התפלגות כמות הפצפוצים בכל חבילה?

ב. עורך הדין בייקר הגיש תביעה בגין פרסום שגוי לאחר שקנה 100 חבילות וגילה שבאחת מהן לא היו כלל פצפוצי שוקולד. הטבח, מר קוק, טוען שזה מקרי, אך אינו מעיד על רמאות מצד החברה. מהי ההסתברות לתקרית כמו זו שקרתה לעורך הדין אם אכן החברה איננה מרמה?

ג. מחקר שפורסם בכתב העת "פצפוצים פליליים" העלה שאחת מכל 10^3 מאפיות אכן מרמה, ובכל פעם שלקוח קונה 100 חבילות ומעלה היא משחילה בזדון אחת ללא פצפוצים (ודואגת שבשאר יהיו פצפוצים כדי שלא יהיו חשדות). מה ההסתברות שהמאפייה של מר קוק אכן רמאית, כפי שטוען עורך הדין?

פתרון.

א. התפלגות הפצפוצים הכוללת היא $\mathrm{Poi}(10000)$. לכן בכל חבילה יש בממוצע 10 פצפוצים, ונקבל שכמות הפצפוצים בכל חבילה מתפלגת $\mathrm{Poi}(10)$.
הסבר נוסף: המשמעות של הפרמטר λ הוא מספר המאורעות (הפצפוצים) הממוצע שמתרחש לכל דגימה. מאחר שהן מפוזרות באופן אחיד, נוכל לומר ש-$\lambda = 10$.

ב. הסיכוי שבחבילה אחת לא יהיו פצפוצים כלל הוא $P\left(\mathrm{Poi}\left(10\right) = 0\right) = e^{-10}$. נסמן על ידי X את כמות החבילות מתוך ה-100 שאין בהן פצפוצים. אם החברה לא מרמה, $X \sim \mathrm{Bin}\left(100, e^{-10}\right)$. לכן

$$P\left(X = 1 \mid \text{החברה לא מרמה}\right) = P\left(\mathrm{Bin}\left(100, e^{-10}\right) = 1\right) = \binom{100}{1} \cdot e^{-10} \cdot \left(1 - e^{-10}\right)^{99} = 0.0045.$$

ג. נסמן על ידי C את המאורע שהחברה מרמה. נרצה להשתמש בחוק בייס:

$$P\left(C \mid X = 1\right) = \frac{P\left(X = 1 \mid C\right) \cdot P\left(C\right)}{P\left(X = 1\right)}.$$

אנחנו יודעים ש-$P\left(X = 1 \mid C\right) = 1$ ו-$P\left(C\right) = 10^{-3}$. מה לגבי $P\left(X = 1\right)$? ניעזר בנוסחת ההסתברות השלמה:

$$\begin{aligned} P\left(X = 1\right) &= P\left(X = 1 \mid C\right) \cdot P\left(C\right) + P\left(X = 1 \mid C^c\right) \cdot P\left(C^c\right) = \\ &= 1 \cdot 10^{-3} + 0.0045 \cdot \left(1 - 10^{-3}\right) = 0.0055 \end{aligned}$$

ובסך הכל

$$.P\left(C \mid X=1\right)=\frac{1\cdot 10^{-3}}{0.0055}=0.18$$

תרגיל 7.4. נניח ש-$X\sim \mathrm{Poi}\left(\lambda\right)$ ו-$Y\sim \mathrm{Poi}\left(\mu\right)$ בלתי־תלויים. מה ההתפלגות של $X+Y$?

פתרון. אינטואיטיבית, אם X מייצג את כמות הפעמים שקורה מאורע עם קצב λ ביחידת זמן, ו-Y מייצג מאורע עם קצב μ, אז $X+Y$ אמור גם הוא להתפלג פואסונית – ויספור את כמות הפעמים שקורה המאורע הראשון או השני.

נחשב ישירות כדי לוודא את האינטואיציה. התומך של $X+Y$ הוא גם $\{0,1,2,\dots\}$.

$$\begin{aligned}P\left(X+Y=k\right)&=\sum_{i=0}^{k}P\left(X=i,Y=k-i\right)=\sum_{i=0}^{k}P\left(X=i\right)P\left(Y=k-i\right)=\\&=\sum_{i=0}^{k}e^{-\lambda}\frac{\lambda^{i}}{i!}\cdot e^{-\mu}\frac{\mu^{k-i}}{(k-i)!}=e^{-(\lambda+\mu)}\sum_{i=0}^{k}\frac{\lambda^{i}\mu^{k-i}}{i!\cdot(k-i)!}=\\&=e^{-(\lambda+\mu)}\cdot\frac{1}{k!}\sum_{i=0}^{k}\binom{k}{i}\lambda^{i}\mu^{k-i}=e^{-(\lambda+\mu)}\frac{\left(\lambda+\mu\right)^{k}}{k!}\end{aligned}$$

לכן $X+Y\sim \mathrm{Poi}\left(\lambda+\mu\right)$.

תרגיל 7.5. לחברת דיור יש $10,000$ לקוחות. אם ידוע שבכל חודש, כל לקוח מבקש לעזוב את השירות בהסתברות 0.002 (בלי תלות באחרים), מה ההסתברות שבחודש אחד יעזבו 30 אנשים?

פתרון. נגדיר את X להיות מספר האנשים העוזבים מדי חודש. ההתפלגות המדויקת של X היא $X\sim \mathrm{Bin}\left(10000,0.002\right)$. כלומר, רוצים לחשב את

$$.P\left(X=30\right)=\binom{10,000}{30}\cdot 0.002^{30}\cdot 0.998^{9970}$$

שמישהו יוציא מחשבון.

אלטרנטיבית, ניעזר בקירוב פואסון. אנחנו יודעים שאם הקצב λ קבוע, אז $\mathrm{Bin}\left(n,\frac{\lambda}{n}\right)\approx \mathrm{Poi}\left(\lambda\right)$. במקרה שלנו, $\mathrm{Bin}\left(10000,0.002\right)\approx \mathrm{Poi}\left(20\right)$, ולכן

$$P\left(X=30\right)\approx P\left(\mathrm{Poi}\left(20\right)=30\right)=e^{-20}\cdot\frac{20^{30}}{30!}=0.00834\dots$$

(התשובה האמיתית היא $0.00831\dots$).

תרגיל 7.6 (התפלגות בינומית שלילית)**.** נניח שבמקום לספור את כמות הניסיונות עד להצלחה הראשונה, נספור את כמות הניסיונות עד שיש r הצלחות, כשבכל ניסיון יש הסתברות הצלחה p שאינה תלויה בניסיונות האחרים. יהי X_i המשתנה המקרי הסופר את כמות הניסיונות מההצלחה ה-$(i-1)$ (לא כולל) להצלחה ה-i (כולל), ויהי $X=\sum_{i=1}^{r}X_i$.

א. מה ההתפלגות של $X_1,\dots,X_r$? האם הם תלויים?

ב. מצאו את ההתפלגות של X.

פתרון.

א. לפי ההגדרה, כל $X_i \sim \mathrm{Geo}(p)$, והם בלתי־תלויים, כי ההצלחות בניסויים השונים הן בלתי־תלויות זו בזו.

ב. רוצים לכל $k \in \mathbb{N}$ לחשב את $P\left(X=k\right)$. כמובן, בהכרח $k \geq r$ (כי יש r הצלחות בין כל הניסיונות), כלומר התומך של ההתפלגות הוא $\{r, r+1, \ldots\}$. כדי שיהיו k ניסיונות עד ההצלחה ה-r, צריך שיהיו בתוך הסדרה הזו $k-r$ כישלונות, ושהניסיון האחרון יהיה הצלחה. בין שאר הניסיונות אין לנו אילוצים.
לכל בחירה של סידור של $k-r$ כישלונות ו-$r-1$ הצלחות, ההסתברות לקבל את הסידור הזה היא $p^{r-1}\left(1-p\right)^{k-r}$, וההסתברות להצלחה האחרונה היא p. כמות הסידורים האלו שקולה לבחירה של $r-1$ הצלחות מתוך $k-1$ מקומות, ללא החזרה וללא חשיבות לסדר, ולכן יש $\binom{k-1}{r-1}$ אפשרויות לבחירה הזו. בסך הכל, נקבל $P\left(X=k\right)=\binom{k-1}{r-1}p^r\left(1-p\right)^{k-r}$.

הגדרה 7.7. להתפלגות הזו קוראים **התפלגות בינומית שלילית**; מסמנים $X \sim \mathrm{NB}(r,p)$.

תרגיל 7.8 (התפלגות היפרגיאומטרית)**.** בכד יש N עצמים שמתוכם יש D מיוחדים. שולפים n בלי החזרה ($n \leq N$), ויהי X כמות העצמים המיוחדים שנשלפו. מצאו את ההתפלגות של X.

פתרון. ראשית, נבין מהו התומך של X. כמובן $0 \leq X \leq \min\left\{n, D\right\}$. מצד שני, $D-X \leq N-n$ (כי $D-X$ מייצג את מספר העצמים המיוחדים שלא נשלפו, והוא צריך להיות קטן מאשר מספר העצמים שלא נשלפו), ולכן $X \geq N+D-n$. בסך הכל התומך הוא המספרים השלמים בתחום $\max\left\{0, N+D-n\right\} \leq X \leq \min\left\{n, D\right\}$.

נעבור לחשב את ההתפלגות. מרחב המדגם Ω מכיל את כל התוצאות של שליפות של n כדורים מתוך ה-N. כיוון שאנחנו רק מתעניינים במספר העצמים שנשלפו, ולא בסדר שלהם, זו בחירה של n מתוך N בלי החזרה ובלי חשיבות לסדר, כלומר $\left|\Omega\right|=\binom{N}{n}$. לכל k בתומך, כמות הדרכים שבהן נשלפים k עצמים מיוחדים היא $\binom{D}{k}\binom{N-D}{n-k}$ (כי צריך לבחור k עצמים מיוחדים מתוך ה-D ו-$(n-k)$ עצמים לא מיוחדים מתוך ה-$N-D$). בסך הכל,

$$.P\left(X=k\right)=\frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}$$

הגדרה 7.9. להתפלגות הזו קוראים **התפלגות היפרגיאומטרית**, ומסמנים $X \sim \mathrm{HG}(N,D,n)$.

תרגיל 7.10 (אם יש זמן)**.** הראו כי במקרה הבדיד חוסר זיכרון שקול לגיאומטריות. באופן פורמלי: נניח כי X משתנה מקרי עם תומך $\{1,2,\ldots\}$. הראו כי X מתפלג גיאומטרית אם ורק אם לכל $m \geq 1$, $P\left(X=m+1 \,|\, X>1\right)=P\left(X=m\right)$.

פתרון. בהרצאה ראיתם את הגרירה $\boxed{\Leftarrow}$. נוכיח את $\boxed{\Rightarrow}$. נניח כי לכל $m \geq 1$ מתקיים $P\left(X=m+1 \,|\, X>1\right)=P\left(X=m\right)$. נראה כעת כי X מתפלג גיאומטרית. אכן, לפי ההגדרה,

$$.P\left(X=m+1 \,|\, X>1\right)=\frac{P\left(X=m+1, X>1\right)}{P\left(X>1\right)}=\frac{P\left(X=m+1\right)}{P\left(X>1\right)}$$

לפי הנתון, נקבל שלכל m מתקיים $P\left(X=m+1\right)=P\left(X=m\right)\cdot P\left(X>1\right)$. אם נסמן $p=P\left(X>1\right)$, נקבל שהסדרה $P\left(X=m\right)$ היא סדרה הנדסית שמנתה p, ואיברה הראשון הוא $P\left(X=1\right)=1-P\left(X>1\right)=1-p$ (כי X נתמך על המספרים הטבעיים). לכן $P\left(X=m\right)=p\left(1-p\right)^{m-1}$, כנדרש.

7.1 התוחלת של משתנה מקרי

המטרה. למצוא את ה"ממוצע" של ההתפלגות. היינו לוקחים את הממוצע של הערכים של המשתנה המקרי, אבל מן הסתם נרצה להתחשב גם בהסתברויות שלו – אם משתנה מקרי מקבל את הערך 0 בהסתברות 0.99 ואת הערך 10000 בהסתברות 0.01, לא היינו רוצים שהממוצע יהיה 5000, שהרי הרבה יותר סביר לקבל 0.

הגדרה 7.11. יהי X משתנה מקרי. **התוחלת** של X הינה

$$\mathbb{E}\left[X\right]=\sum_{a}a\cdot P\left(X=a\right).$$

דוגמה 7.12.

א. נניח ש-X הוא הטלת קוביה, כלומר X מתפלג אחיד על $\{1,\ldots,6\}$. אז $\mathbb{E}\left[X\right]=1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+\cdots+6\cdot\frac{1}{6}=3.5$.

ב. באופן כללי, אם X מתפלג אחיד על פני המספרים $a,a+1,\ldots,b$, אז התוחלת היא הממוצע של המספרים – $\mathbb{E}\left[X\right]=\frac{a+b}{2}$.

ג. נניח ש-$X\sim\mathrm{Ber}\left(p\right)$ הוא משתנה ברנולי. מה התוחלת של X?

$$\mathbb{E}\left[X\right]=1\cdot p+0\cdot\left(1-p\right)=p.$$

תרגיל 7.13. בחברת "מעסיקים בזול מוכרים ביוקר" ישנם 100 עובדים. מתוכם המנכ"ל מרוויח $1,000,000$ ש"ח בשנה, 9 המנהלים הבכירים מרוויחים $10,000$ ש"ח בשנה, והעובדים הזוטרים מרוויחים 10 ש"ח בשנה (בושה וחרפה).

א. מהו השכר הממוצע בחברה?

ב. מהו השכר החציוני בחברה?

פתרון. יהי X המשתנה המקרי של הרווח של עובד אקראי בחברה.

א. $P\left(X=1,000,000\right)=\frac{1}{100}$, $P\left(X=10,000\right)=\frac{9}{100}$, $P\left(X=10\right)=\frac{90}{100}=\frac{9}{10}$. לכן

$$\mathbb{E}\left[X\right]=\frac{1}{100}\cdot1,000,000+\frac{9}{100}\cdot10,000+\frac{9}{10}\cdot10=10909$$

ב. חציון של קבוצת ערכים $a_1,\ldots,a_n$ מוגדר בתור המספר האמצעי בסדרה אם היינו מסדרים אותה בסדר עולה, כולל חזרות (אם היא באורך זוגי – הממוצע בין השניים שבאמצע). במקרה הזה כמעט כל הערכים הם 10, ולכן גם השכר החציוני הוא 10 ש"ח בשנה.

טענה 7.14. יהיו X,Y משתנים מקריים.

א. אדיטיביות התוחלת: $\mathbb{E}\left[X+Y\right]=\mathbb{E}\left[X\right]+\mathbb{E}\left[Y\right]$, גם כאשר הם תלויים!

ב. אם a סקלר, אז $\mathbb{E}\left[aX\right]=a\cdot\mathbb{E}\left[X\right]$.

תרגיל 7.15. שי, רון ואיתי משחקים במשחק הבא: בכל תור,

- שי מטיל מטבע מוטה שהסיכוי שלו לצאת "פלי" הוא 0.6. אם אכן יוצא "פלי" הוא מקבל שלושה שקלים.
- רון מקבל שני שקלים באופן קבוע.
- איתי מקבל ארבעה שקלים; הוא מטיל מטבע דומה לזה של שי, ואם יוצא "פלי" הוא נאלץ להחזיר שלושה שקלים.

בכל תור, המנצח הוא מי שהרוויח הכי הרבה (באותו תור).
חשבו את הסיכוי שכל אחד מנצח את השני. בהנחה שמשחקים הרבה תורות, מי ירוויח הכי הרבה?

פתרון. נכתוב את התפלגויות הרווח של כל שחקן:

$$S \sim 3 \cdot \mathrm{Ber}(0.6), \qquad R \sim 2, \qquad I \sim 4 - 3 \cdot \mathrm{Ber}(0.6).$$

נשווה את זוגות השחקנים:

$$P(S > R) = P(3 \cdot \mathrm{Ber}(0.6) > 2) = P(\mathrm{Ber}(0.6) = 1) = 0.6$$

$$P(R > I) = P(2 > 4 - 3 \cdot \mathrm{Ber}(0.6)) = P\left(\mathrm{Ber}(0.6) > \frac{2}{3}\right) = P(\mathrm{Ber}(0.6) = 1) = 0.6$$

לכן במרבית התורות שי מנצח את רון, ורון מנצח את איתי. מה לגבי שי ואיתי? נכתוב את כל המאורעות בטבלה:

המטבע של שי / המטבע של איתי	עץ	פלי
עץ	$S = 0, I = 4$ – איתי מנצח	$S = 0, I = 1$ – איתי מנצח
פלי	$S = 3, I = 4$ – איתי מנצח	$S = 3, I = 1$ – שי מנצח

לכן $P(S > I) = P(\text{פלי, פלי}) = 0.6^2 = 0.36$. כלומר, במרבית התורות שי מנצח את רון, במרבית התורות רון מנצח את איתי, ובמרבית התורות איתי מנצח את שי!
כדי להבין מי השחקן הרווחי ביותר לאורך מספר תורות רב, נחשב את התוחלת:

$$\mathbb{E}[S] = 3 \cdot 0.6 = 1.8, \qquad \mathbb{E}[R] = 2, \qquad \mathbb{E}[I] = 4 - 3 \cdot 0.6 = 2.2.$$

כלומר איתי הוא השחקן הרווחי ביותר (וזאת על אף שבמרבית התורות רון מנצח אותו).

תרגיל 7.16. חשבו את התוחלות של ההתפלגויות הבאות:

א. $\mathrm{Bin}(n, p)$.

ב. $\mathrm{Geo}(p)$.

פתרון. התוחלת של משתנה מקרי תלויה רק בהתפלגות שלו, ולכן מספיק תמיד לבחור משתנה אחד עם ההתפלגות המתאימה ולחשב את התוחלת שלו.

א. יהיו $X_1, \ldots, X_n$ משתנים מקריים בלתי־תלויים בעלי התפלגות $\mathrm{Ber}(p)$. אז נקבל $X = \sum_{i=1}^n X_i \sim \mathrm{Bin}(n, p)$. מלינאריות התוחלת, $\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] = np$.

ב. נחשב ישירות: אם $X \sim \mathrm{Geo}(p)$, נסמן $q = 1 - p$. אז

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \cdot P(X = k) = \sum_{k=0}^{\infty} kpq^{k-1} = p\sum_{k=0}^{\infty} kq^{k-1} = p\sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}q}\left(q^k\right) =$$
$$= p \cdot \frac{\mathrm{d}}{\mathrm{d}q}\left(\sum_{k=0}^{\infty} q^k\right) = p \cdot \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{1}{1-q}\right) = p \cdot \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

8 תרגול שמיני

תרגיל 8.1. בעיר מסוימת ב-80% מבתי האב יש 2 ילדים, ואילו ב-20% מבתי האב יש 8 ילדים.

א. מה תוחלת מספר הילדים בבתי אב בעיר?

ב. בוחרים ילד באקראי ושואלים אותו כמה אחים יש לו. מה תוחלת מספר האחים עבור ילדי העיר?

פתרון.

א. נסמן על ידי X את מספר הילדים בבית אב. אז $P(X=2) = 0.8$ ו-$P(X=8) = 0.2$, ולכן
$$.\mathbb{E}[X] = 2 \cdot 0.8 + 8 \cdot 0.2 = 3.2$$

ב. רציתם להגיד 2.2? וואלה באסה.
נסמן על ידי Y את כמות האחים של ילד אקראי בעיר. ברור שהערכים של Y הם 1 או 7. אבל ההסתברויות הולכות להשתנות, כי עכשיו הסיכוי לבחור ילד מתוך בית אב עם 8 ילדים גדול יותר מאשר הסיכוי לבחור רק את בית האב הזה. נניח שיש n בתי אב בסך הכל בעיר. אז יש $8 \cdot 0.2n = 1.6n$ ילדים מבית אב עם 8 ילדים, ו-$2 \cdot 0.8n = 1.6n$ ילדים מבית אב עם 2 ילדים. כלומר:

$$P\left(\text{ילד אקראי הוא מבית אב עם 2 ילדים}\right) = P\left(\text{ילד אקראי הוא מבית אב עם 8 ילדים}\right) = \frac{1}{2}$$

ולכן
$$.P(Y=1) = P(Y=7) = \frac{1}{2}$$
כלומר $\mathbb{E}[Y] = 1 \cdot \frac{1}{2} + 7 \cdot \frac{1}{2} = 4$.

טענה 8.2 (תכונות התוחלת). בנוסף ללינאריות שהזכרנו בתרגול הקודם, נזכיר עוד שלוש תכונות של התוחלת:

א. מונוטוניות: אם $X \geq 0$, אז $\mathbb{E}[X] \geq 0$.
באופן שקול: אם $X \geq Y$ אז $\mathbb{E}[X] \geq \mathbb{E}[Y]$.

ב. טרנספורמציה: אם $f\colon \mathbb{R} \to \mathbb{R}$ פונקציה, אז $\mathbb{E}[f(X)] = \sum_a f(a)\, P(X=a)$.

ג. חישוב על ידי הזנב: אם X מקבל רק ערכים טבעיים, אז $\mathbb{E}[X] = \sum_{n=1}^{\infty} P(X \geq n)$.

תרגיל 8.3. יהי $X \sim U[-3,3]$ משתנה מקרי אחיד. נגדיר $Y = X^2$.

א. חשבו את $\mathbb{E}[Y]$ על ידי חישוב ההתפלגות של Y ישירות.

ב. חשבו את $\mathbb{E}[Y]$ על ידי שימוש בנוסחה הנ"ל.

פתרון.

א. הערכים ש-Y מקבל הם $\{0,1,4,9\}$, בהסתברויות $P(Y=0)=\frac{1}{7}$ ו-$P(Y=a)=\frac{2}{7}$ לכל $a\in\{1,4,9\}$. לכן

$$\mathbb{E}[Y]=0\cdot\frac{1}{7}+1\cdot\frac{2}{7}+4\cdot\frac{2}{7}+9\cdot\frac{2}{7}=14\cdot\frac{2}{7}=4.$$

ב. על פי הנוסחה,

$$\begin{aligned}\mathbb{E}[Y]=\mathbb{E}\left[X^2\right]&=\sum_{a=-3}^{3}a^2\cdot P(X=a)=\frac{1}{7}\sum_{a=-3}^{3}a^2=\\&=\frac{9+4+1+0+1+4+9}{7}=\frac{28}{7}=4\end{aligned}$$

תרגיל 8.4 (פרדוקס סנט פטרבורג)**.** מנהלים הגרלה באופן הבא. בקופה יש שקל אחד. השחקן מקבל הזדמנות להטיל מטבע הוגן עד לקבלת ה"עץ" הראשון, ואז לקבל את כל המטבעות שבקופה, כשבכל תור מכפילים את כמות המטבעות שבקופה. עבור איזה מחיר תסכימו להיכנס למשחק?

(הפרדוקס קרוי כך על שם מאמרו של דניאל ברנולי שהציג אותו ב-1738 בכתב העת "כתבי האקדמיה המלכותית למדעים של סנט פטרבורג", למרות שמקורו הוא במכתב שכתב ניקולאוס ברנולי, בן דודו של דניאל, ב-1713)

פתרון. יהי X המשתנה המקרי המייצג את הפעם הראשונה שבה יצא עץ. הרווח שלנו במשחק הוא 2^{X-1}. אם נחשב את התוחלת של 2^{X-1}, נדע מתי כדאי להיכנס (יהיה כדאי להיכנס לכל ערך שנמצא מתחת לתוחלת, כי אז תוחלת הרווח שלנו חיובית). אנחנו יודעים ש-$X\sim\mathrm{Geo}\left(\frac{1}{2}\right)$, ולכן

$$\mathbb{E}\left[2^{X-1}\right]=\sum_{k=1}^{\infty}2^{k-1}\cdot P(X=k)=\sum_{k=1}^{\infty}2^{k-1}\cdot\left(\frac{1}{2}\right)^{k-1}\cdot\frac{1}{2}=\frac{1}{2}\sum_{k=1}^{\infty}1=\infty.$$

כלומר כדאי להיכנס בכל מחיר למשחק הזה.

הערה 8.5. באופן כללי, נגיד שלמשתנה מקרי X קיימת תוחלת סופית אם הטור $\mathbb{E}[X]=\sum_a a\cdot P(X=a)$ מתכנס בהחלט. בפרט, סדר הסכימה לא משנה את התוחלת. בתרגיל הקודם ראינו דוגמה למשתנה מקרי שהתוחלת שלו אינסופית. נראה דוגמאות שבהן אין תוחלת בכלל:

דוגמה 8.6. נניח ש-$X\sim\mathrm{Geo}(\frac{1}{2})$.

א. נגדיר $Y=(-2)^X$. זהו משתנה מקרי, והתוחלת שלו היא

$$\mathbb{E}[Y]=\sum_{k=1}^{\infty}(-2)^k\cdot P(X=k)=\sum_{k=1}^{\infty}(-2)^k\cdot\left(\frac{1}{2}\right)^k=\sum_{k=1}^{\infty}(-1)^k$$

שאינו מתכנס. לכן אין ל-Y תוחלת.

ב. נגדיר $Z = \frac{(-2)^X}{X}$. זהו משתנה מקרי, והתוחלת שלו היא

$$\mathbb{E}\left[Z\right] = \sum_{k=1}^{\infty} \frac{(-2)^k}{k} \cdot P\left(X = k\right) = \sum_{k=1}^{\infty} \frac{(-2)^k}{k} \cdot \left(\frac{1}{2}\right)^k = \sum_{k=1}^{\infty} \frac{(-1)^k}{k}.$$

הטור הזה מתכנס בתנאי (ל-$\ln 2$), אבל כיוון שהוא לא מתכנס בהחלט – נגיד של-Z אין תוחלת סופית.

אם רוצים לחשב תוחלת של כמות עצמים שמקיימת תכונה מסוימת, אחת השיטות השימושיות היא להציג את הכמות הזו כסכום של אינדיקטורים, לחשב את התוחלת של כל אחד מהם, ואז לסכום את התוצאות.

דוגמה 8.7. כשחישבנו את התוחלת של התפלגות בינומית $\mathrm{Bin}\left(n,p\right)$, חשבנו על ההתפלגות הזו כעל כמות ההצלחות בסדרה של n ניסויי ברנולי בלתי־תלויים בעלי סיכוי הצלחה p. הצגנו את המשתנה כסכום של n אינדיקטורים, שכל אחד מהם בודק האם הניסוי ה-i היה הצלחה; לכל אחד מהם יש סיכוי (ולכן תוחלת) p, ובסך הכל התוחלת היא np.

תרגיל 8.8. חשבו את התוחלת של $\mathrm{HG}(N,D,n)$.

פתרון. אפשר לחשב את התוחלת ישירות מהנוסחה. לעומת זאת, נפעל באופן הבא: נניח שיש לנו כד עם N כדורים צהובים וכחולים, שמתוכם יש D כחולים. שולפים מתוכם n כדורים, ו-X סופר כמה כדורים כחולים שלפנו. אז $X \sim \mathrm{HG}(N,D,n)$. נכתוב $X = \sum_{j=1}^{n} I_j$, כאשר I_j הוא משתנה האינדיקטור הבודק האם הכדור ה-j שנשלף הוא כחול.

לכל $1 \leq j \leq n$, נטען כי $\mathbb{E}\left[I_j\right] = \frac{D}{N}$. אכן, צריך לחשב את הסיכוי שהכדור ה-j שנשלף הוא כחול. אפשר לדמיין כאילו היינו ממשיכים לשלוף את הכדורים באופן אקראי מן הכד, ומסדרים אותם בשורה על סמך הסדר שבו הם יצאו. באופן זה אנחנו מקבלים סידור אקראי של כל הכדורים מהכד, והכדור ה-j שנשלף הוא הכדור במקום ה-j בשורה. כעת ברור שהסיכוי שהכדור הזה הוא כחול הוא $\frac{D}{N}$, והוכחנו את הטענה.

מאדיטיביות התוחלת, נקבל $\mathbb{E}\left[X\right] = \sum_{i=1}^{n} \frac{D}{N} = \frac{nD}{N}$.

תרגיל 8.9. מספר הדגים באגם אינו ידוע. דגו 1000 דגים וסימנו להם את הסנפיר. אחרי שבוע שוב דגו 1000 דגים ומצאו כי 17 דגים מתוכם מסומנים. העריכו את כמות הדגים באגם על סמך התוחלת.

פתרון. נסמן ב-N את כמות הדגים באגם וב-D את כמות הדגים המסומנים. אם X היא כמות הדגים המסומנים שדגו לאחר שבוע, נשים לב ש-X הוא משתנה מקרי היפרגיאומטרי:

$$X \sim \mathrm{HG}\left(N, D, n = 1000\right).$$

מכיוון שכך, $\mathbb{E}\left[X\right] = \frac{nD}{N}$. כדי להעריך מהו N נניח ש-17 הדגים שהתקבלו הם בערך התוחלת של X, כלומר $\frac{nD}{N} \approx 17$, ולכן $N \approx \frac{nD}{17} = \frac{1000^2}{17} = 58823.529$.

תרגיל 8.10. במפעל עוגיות שוקולד צ'יפס מייצרים עוגיות באופן הבא: ראשית, לוקחים גוש שוקולד, ומחלקים אותו ל-n חלקים שווים. כעת מגיעה פיית הבצק, בוחרת באקראי ועם החזרות n חלקים, וכל אחד מהם היא הופכת לבצק. מה תוחלת אחוז השוקולד שנותר? נתחו את התשובה כאשר $n \to \infty$.

פתרון. לכל $1 \leq i \leq n$, נסמן על ידי A_i את המאורע שהחלק ה-i נשאר שוקולד. אז כמות החלקים שנותרו שוקולד בסוף התהליך היא $X = \sum_{i=1}^{n} 1_{A_i}$. לכל $1 \leq i \leq n$, ההסתברות שהפייה לא תבחר את החלק ה-i ב-n הבחירות שלה היא $\left(1-\frac{1}{n}\right)^n$ (כי היא צריכה לבחור את אחד מ-$n-1$ החלקים האחרים), ולכן $\mathbb{E}\left[1_{A_i}\right] = P\left(A_i\right) = \left(1-\frac{1}{n}\right)^n$. מאדיטיביות התוחלת, $\mathbb{E}\left[X\right] = \sum_{i=1}^{n} \mathbb{E}\left[1_{A_i}\right] = n \cdot \left(1-\frac{1}{n}\right)^n$, כלומר אחוז השוקולד שנותר הוא בתוחלת $\mathbb{E}\left[\frac{1}{n}X\right] = \left(1-\frac{1}{n}\right)^n$. כאשר $n \to \infty$, הכמות הזו שואפת ל-$\frac{1}{e}$.

איור 2: עוגיות עם $37\% \approx \frac{1}{e}$ שוקולד

טענה 8.11 (תוחלת של מכפלה). אם X, Y בלתי־תלויים, אז $\mathbb{E}\left[XY\right] = \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$.

דוגמה 8.12. נניח ש-X, Y מייצגים תוצאות הטלה של קוביות הוגנות שונות. אז מתקיים $X, Y \sim U\left[1,6\right]$, ולכן $\mathbb{E}\left[X\right] = \mathbb{E}\left[Y\right] = \frac{7}{2}$, כלומר $\mathbb{E}\left[XY\right] = \left(\frac{7}{2}\right)^2 = \frac{49}{4}$.
מצד שני, $\mathbb{E}\left[X^2\right] = \frac{1}{6}\sum_{a=1}^{6} a^2 = \frac{1+4+9+16+25+36}{6} = \frac{91}{6} \neq \mathbb{E}\left[X\right]^2$.

תרגיל 8.13. תנו דוגמה למשתנים מקריים תלויים X, Y שעבורם $\mathbb{E}\left[XY\right] = \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$.

פתרון. נניח ש-Z_1 ו-Z_2 הם תוצאות הטלות של שתי קוביות שונות, וניקח $X = Z_1 + Z_2$ ו-$Y = Z_1 - Z_2$. מתקיים

$$\mathbb{E}\left[Y\right] = \mathbb{E}\left[Z_1 - Z_2\right] = \mathbb{E}\left[Z_1\right] - \mathbb{E}\left[Z_2\right] = 0,$$

ולכן $\mathbb{E}\left[X\right]\mathbb{E}\left[Y\right] = 0$. מצד שני,

$$\mathbb{E}\left[XY\right] = \mathbb{E}\left[\left(Z_1+Z_2\right)\left(Z_1-Z_2\right)\right] = \mathbb{E}\left[Z_1^2 - Z_2^2\right] = \mathbb{E}\left[Z_1^2\right] - \mathbb{E}\left[Z_2^2\right] = 0,$$

ולכן יש שוויון.
מדוע X ו-Y תלויים? אם למשל ידוע $Y = 0$, אז בהכרח $Z_1 = Z_2$, כלומר X הוא זוגי. אז למשל $P\left(X=3, Y=0\right) = 0 \neq \frac{2}{36} \cdot \frac{1}{6} P\left(X=3\right) \cdot P\left(Y=0\right)$.

9 תרגול תשיעי

9.1 השונות של משתנה מקרי

התוחלת נותנת לנו מידע על ה"מרכז" של ההתפלגות שלנו. לעומת זאת, אם נרצה להבין גם כמה ההתפלגות שלנו "מפוזרת", נרצה מדד נוסף.

הגדרה 9.1. יהי X משתנה מקרי. **השונות** של X הינה

$$\mathrm{Var}\left(X\right)=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)^{2}\right]=\mathbb{E}\left[X^{2}\right]-\mathbb{E}\left[X\right]^{2}.$$

סטיית התקן של X הינה

$$\mathrm{std}\left(X\right)=\sqrt{\mathrm{Var}\left(X\right)}.$$

איך מחשבים את השונות? נזכיר שלכל פונקציה $f:\mathbb{R}\to\mathbb{R}$ מתקיים

$$\mathbb{E}\left[f\left(X\right)\right]=\sum_{a}f\left(a\right)P\left(X=a\right).$$

בפרט אפשר לקחת $f\left(x\right)=x^{2}$ ולקבל

$$\mathbb{E}\left[X^{2}\right]=\sum_{a}a^{2}\cdot P\left(X=a\right).$$

אז מחשבים את $\mathbb{E}\left[X\right]$ ואת $\mathbb{E}\left[X^{2}\right]$ לפי הנוסחאות האלו, ואז מחשבים את השונות מההגדרה.

טענה 9.2. יהיו X,Y משתנים מקריים.

א. $\mathrm{Var}\left(X\right)\geq 0$. יותר מזה, $\mathrm{Var}\left(X\right)=0$ אם ורק אם X קבוע. כלומר אם המשתנה שלנו לא קבוע, השונות (ולכן גם סטיית התקן) חיוביים!

ב. $\mathrm{Var}\left(aX+b\right)=a^{2}\mathrm{Var}\left(X\right)$.

ג. אם X,Y בלתי־תלויים, $\mathrm{Var}\left(X+Y\right)=\mathrm{Var}\left(X\right)+\mathrm{Var}\left(Y\right)$.

תרגיל 9.3. יהי X משתנה מקרי המקיים

$$P\left(X=0\right)=0.4,\qquad P\left(X=1\right)=0.3,\qquad P\left(X=2\right)=0.2,\qquad P\left(X=3\right)=0.1.$$

א. חשבו את התוחלת ואת השונות של X.

ב. נגדיר $Y=5X-13$ ו־$Z=\left(X-2\right)^{2}$. חשבו את התוחלת ואת השונות של Y ו־Z.

ג. חשבו את התוחלת ואת השונות של $L=\frac{3}{X-3}$.

פתרון.

א. נחשב ישירות:

$$\mathbb{E}\left[X\right]=0\cdot 0.4+1\cdot 0.3+2\cdot 0.2+3\cdot 0.1=1$$
$$\mathbb{E}\left[X^{2}\right]=0^{2}\cdot 0.4+1^{2}\cdot 0.3+2^{2}\cdot 0.2+3^{2}\cdot 0.1=2$$

לכן $\mathrm{Var}\left(X\right)=1$.

ב. עבור Y זה קל:

$$\mathbb{E}\left[Y\right]=\mathbb{E}\left[5X-13\right]=5\mathbb{E}\left[X\right]-13=5-13=-8$$
$$\mathrm{Var}\left(Y\right)=\mathrm{Var}\left(5X-13\right)=5^{2}\cdot\mathrm{Var}\left(X\right)=25$$

עבור Z עדיף לחשב ישירות: ההתפלגות שלו היא

$$\begin{aligned} P(Z=0) &= P(X=2) = 0.2 \\ P(Z=1) &= P(X \in \{1,3\}) = 0.4 \\ P(Z=4) &= P(X=0) = 0.4 \end{aligned}$$

לכן

$$\begin{aligned} \mathbb{E}[Z] &= 0 \cdot 0.2 + 1 \cdot 0.4 + 4 \cdot 0.4 = 2 \\ \mathbb{E}\left[Z^2\right] &= 0^2 \cdot 0.2 + 1^2 \cdot 0.4 + 4^2 \cdot 0.4 = 6.8 \end{aligned}$$

ומכאן $\mathrm{Var}(Z) = 4.8$.

ג. L אינו משתנה מקרי, כי הוא אינו מוגדר לכל המרחב.

תרגיל 9.4. חשבו את השונות של ההתפלגויות הבאות:

א. $U[a,b]$.

ב. $\mathrm{Ber}(p)$.

ג. $\mathrm{Geo}(p)$.

ד. $\mathrm{Poi}(\lambda)$.

פתרון.

א. יהי $X \sim U[a,b]$. אז $X-a+1 \sim U[1, b-a+1]$, ו-$\mathrm{Var}(X-a+1) = \mathrm{Var}(X)$. לכן מספיק לחשב את השונות של $Y = X - a + 1$.
לצורך הנוחות, נסמן $c = b - a + 1$; לכן $Y \sim U[1,c]$. לחישוב $\mathbb{E}\left[Y^2\right]$, נזכיר כי $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$. לכן

$$\mathbb{E}\left[Y^2\right] = \sum_{k=1}^{c} \frac{1}{c} k^2 = \frac{(c+1)(2c+1)}{6}$$

ולכן

$$\begin{aligned} \mathrm{Var}(X) = \mathrm{Var}(Y) = \mathbb{E}\left[Y^2\right] - \mathbb{E}[Y]^2 &= \frac{(c+1)(2c+1)}{6} - \left(\frac{c+1}{2}\right)^2 \\ &= \frac{(c+1)(2c+1)}{6} - \frac{(c+1)^2}{4} = \\ &= \frac{c+1}{12} \cdot (2(2c+1) - 3(c+1)) = \frac{c+1}{12}\left(4c^2 - c - 3\right) = \\ &= \frac{(c+1)(c-1)}{12} = \frac{c^2-1}{12} = \frac{(b-a+1)^2 - 1}{12} \end{aligned}$$

ב. יהי $X \sim \mathrm{Ber}(p)$. אז $X^2 = X$, לכן $\mathbb{E}\left[X^2\right] = \mathbb{E}[X] = p$, ונקבל $\mathrm{Var}(X) = p - p^2 = p(1-p)$.

ג. יהי $X \sim \mathrm{Geo}(p)$. אפשר לחשב את $\mathbb{E}\left[X^2\right]$ בדומה לחישוב שעשינו עבור התוחלת, על ידי גזירה פעמיים. לעומת זאת, נראה שיטה שונה לחישוב:

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \sum_{k=1}^{\infty} k^2 p(1-p)^{k-1} = \sum_{k=1}^{\infty} (k-1+1)^2 p(1-p)^{k-1} = \\
&= \sum_{k=1}^{\infty} (k-1)^2 p(1-p)^{k-1} + 2\sum_{k=1}^{\infty} (k-1) p(1-p)^{k-1} + \sum_{k=1}^{\infty} p(1-p)^{k-1} = \\
&= \sum_{j=0}^{\infty} j^2 p(1-p)^{j} + 2\sum_{j=0}^{\infty} jp(1-p)^{j} + 1 = \\
&= (1-p)\sum_{j=1}^{\infty} j^2 p(1-p)^{j-1} + 2(1-p)\sum_{j=1}^{\infty} jp(1-p)^{j-1} + 1 = \\
&= (1-p)\mathbb{E}\left[X^2\right] + 2(1-p)\mathbb{E}\left[X\right] + 1
\end{aligned}$$

ובסך הכל

$$\mathbb{E}\left[X^2\right] = \frac{2(1-p)\mathbb{E}\left[X\right]+1}{p} = \frac{\frac{2(1-p)}{p}+1}{p} = \frac{2-p}{p^2}.$$

לכן $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$.

ד. יהי $X \sim \mathrm{Poi}(\lambda)$. בתרגיל הבית הוכחתם כי $\mathbb{E}\left[X(X-1)\right] = \lambda^2$, ומכאן נסיק כי $\mathbb{E}\left[X^2\right] = \mathbb{E}\left[X\right] + \lambda^2 = \lambda + \lambda^2$. בסך הכל $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2 = \lambda$.

תרגיל 9.5. חברת ממתקים מחביאה כרטיס בצבע מסוים (מבין שלושה צבעים) בכל חפיסת שוקולד; שיעור צבעי הכרטיסים שווה בין כל החפיסות. מה תוחלת מספר החפיסות שיש לקנות עד לקבלת כרטיסים בכל שלושת הצבעים? ומהי שונות כמות החפיסות הזו?

פתרון. יהי X המשתנה המקרי המייצג את מספר החפיסות שצריך לקנות עד לצבירת כרטיסים מכל שלושת הצבעים. בחפיסה הראשונה הקונים יוצא צבע כלשהו; נקרא לו 'הצבע הראשון'. נכתוב: $X = 1 + Y + Z$ כאשר Y הוא המשתנה המקרי המייצג את מספר החפיסות שיש לקנות לאחר החפיסה הראשונה ועד לקבלת כרטיס מצבע אחר; ו-Z הוא מספר החפיסות שיש לקנות לאחר שקיבלנו כרטיסים משני צבעים ועד לקבלת הצבע השלישי. אנחנו יודעים כי התוחלת אדיטיבית ולכן $\mathbb{E}\left[X\right] = 1 + \mathbb{E}\left[Y\right] + \mathbb{E}\left[Z\right]$.

כיצד מתפלג Y? אנחנו מבצעים ניסוי עד לקבלת הצלחה ראשונה, כאשר הסתברות ההצלחה היא $\frac{2}{3}$. לכן $Y \sim \mathrm{Geo}\left(\frac{2}{3}\right)$, ומכאן $\mathbb{E}\left[Y\right] = \frac{3}{2}$. בדומה, $Z \sim \mathrm{Geo}\left(\frac{1}{3}\right)$, ולכן $\mathbb{E}\left[Z\right] = 3$. בסך הכל $\mathbb{E}\left[X\right] = 1 + \frac{3}{2} + 3 = 5.5$.

לגבי השונות, נשים לב שפה Y ו-Z הם בלתי־תלויים. אכן, אם אני יודע שלקחו לי 17 חפיסות עד שקיבלתי את הצבע השני, אין שום השפעה על כמה חפיסות אצטרך לקנות עד לקבלת הצבע השלישי. לכן

$$\mathrm{Var}\left(X\right) = \mathrm{Var}\left(1+Y+Z\right) = \mathrm{Var}\left(1\right) + \mathrm{Var}\left(Y\right) + \mathrm{Var}\left(Z\right) = 0 + \frac{\frac{1}{3}}{\left(\frac{2}{3}\right)^2} + \frac{\frac{2}{3}}{\left(\frac{1}{3}\right)^2} = \frac{3}{4} + 6 = 6.75.$$

9.2 השונות המשותפת ומקדם המתאם של פירסון

כעת, אם יש לנו שני משתנים מקריים, המטרה שלנו תהיה ללמוד את הקשר ביניהם.

הגדרה 9.6. יהיו X ו-Y משתנים מקריים. **השונות המשותפת** של X ו-Y היא

$$\mathrm{Cov}\left(X,Y\right)=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)\left(Y-\mathbb{E}\left[Y\right]\right)\right]=\mathbb{E}\left[XY\right]-\mathbb{E}\left[X\right]\mathbb{E}\left[Y\right].$$

אם $\mathrm{Cov}\left(X,Y\right)=0$, אומרים ש-$X$ ו-Y **בלתי־מתואמים**.

טענה 9.7. יהיו X,Y,Z משתנים מקריים.

א. $\mathrm{Cov}\left(X,X\right)=\mathrm{Var}\left(X\right)$.

ב. $\mathrm{Cov}\left(aX+bY,Z\right)=a\mathrm{Cov}\left(X,Z\right)+b\mathrm{Cov}\left(Y,Z\right)$. בדומה ברכיב השני, בגלל:

ג. $\mathrm{Cov}\left(X,Y\right)=\mathrm{Cov}\left(Y,X\right)$.

ד. אם X ו-Y בלתי־תלויים, אז הם בלתי־מתואמים.

ה. $\mathrm{Var}\left(X+Y\right)=\mathrm{Var}\left(X\right)+\mathrm{Var}\left(Y\right)+2\mathrm{Cov}\left(X,Y\right)$. בפרט, אם הם בלתי־מתואמים, השונות אדיטיבית.
הנוסחה לסכום כללי: אם $X_1,\ldots,X_n$ משתנים מקריים, אז

$$\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathrm{Var}\left(X_i\right)+\sum_{i\neq j}\mathrm{Cov}\left(X_i,X_j\right)=\sum_{i=1}^{n}\mathrm{Var}\left(X_i\right)+2\sum_{i<j}\mathrm{Cov}\left(X_i,X_j\right).$$

.

דוגמה 9.8. ראינו בתרגול הקודם (בתרגיל 8.13) דוגמה לזוג משתנים מקריים תלויים אך בלתי־מתואמים.

תרגיל 9.9. n זוגות נעמדים בשורה באופן אקראי. יהי X המשתנה המקרי הסופר את כמות הזוגות שנעמדו זה לצד זה. חשבו את $\mathbb{E}\left[X\right]$ ואת $\mathrm{Var}(X)$.

פתרון. נכתוב $X=\sum_{j=1}^{n}I_j$, כאשר I_j הוא משתנה האינדיקטור של האם הזוג ה-j עומד זה לצד זה. על מנת לחשב את התוחלת ואת השונות של X, נחשב את ההתפלגות של כל I_j.
המשתנים $I_1,\ldots,I_n$ הם שווי התפלגות, וכולם מתפלגים ברנולי. כדי לחשב את $P\left(I_j=1\right)$, צריך לחשב את הסיכוי שהזוג ה-j עומד זה לצד זה. יש $(2n)!$ סידורים אפשריים לכל הזוגות. מתוכם, אם רוצים שהזוג ה-j יישב ביחד, נתייחס אליהם כבלוק, נסדר את $2n-2$ האנשים והבלוק בשורה ($(2n-1)!$ אפשרויות), ונבחר את הסידור הפנימי של הזוג (2 אפשרויות).
בסך הכל

$$P\left(I_j=1\right)=\frac{(2n-1)!\cdot 2}{(2n)!}=\frac{1}{n},$$

ולכן $I_j\sim\mathrm{Ber}(\frac{1}{n})$.
כעת קל לחשב את התוחלת של X, על סמך לינאריות התוחלת:

$$\mathbb{E}\left[X\right]=\mathbb{E}\left[\sum_{j=1}^{n}I_j\right]=\sum_{j=1}^{n}\mathbb{E}\left[I_j\right]=\sum_{j=1}^{n}\frac{1}{n}=1,$$

כלומר בממוצע יהיה זוג אחד שיעמוד זה לצד זה.

לחישוב השונות, נצטרך להתאמץ קצת יותר, שהרי המשתנים $I_1, \ldots, I_n$ ודאי תלויים. לכן

$$\operatorname{Var}(X) = \sum_{j=1}^{n} \operatorname{Var}(I_j) + \sum_{j \neq k} \operatorname{Cov}(I_j, I_k).$$

לכל $1 \leq j \leq n$, $\operatorname{Var}(I_j) = \frac{1}{n}\left(1 - \frac{1}{n}\right)$. לחישוב השונות המשותפת, אם $j \neq k$, נזכור כי

$$\operatorname{Cov}(I_j, I_k) = \mathbb{E}\left[I_j I_k\right] - \mathbb{E}\left[I_j\right]\mathbb{E}\left[I_k\right].$$

אנחנו יודעים כבר ש-$\mathbb{E}\left[I_j\right] = \mathbb{E}\left[I_k\right] = \frac{1}{n}$. ניזכר שבתרגיל 6.5 ראינו כי $I_j I_k$ הוא האינדיקטור של חיתוך המאורעות שהם מייצגים, כלומר זהו משתנה האינדיקטור של האם הזוג ה-j עומד זה לצד זה, וגם הזוג ה-k עומד זה לצד זה. כמו קודם, יש $(2n)!$ אפשרויות לסידור הכללי, והפעם משיטת הבלוקים נקבל שצריך לסדר $2n-4$ אנשים ו-2 בלוקים בשורה, ו-2^2 סידורים פנימיים. בסך הכל

$$P\left(I_j I_k = 1\right) = \frac{(2n-2)! \cdot 4}{(2n)!} = \frac{2}{n(2n-1)}$$

ולכן $\mathbb{E}\left[I_j I_k\right] = \frac{2}{2n-1}$ לכל $j \neq k$, כלומר $\operatorname{Cov}(I_j, I_k) = \frac{2}{n(2n-1)} - \frac{1}{n^2} = \frac{2n-2n+1}{n^2(2n-1)} = \frac{1}{n^2(2n-1)}$. בסך הכל

$$\begin{aligned}\operatorname{Var}(X) &= n \cdot \frac{1}{n}\left(1 - \frac{1}{n}\right) + n(n-1) \cdot \frac{1}{n^2(2n-1)} = \frac{n-1}{n} + \frac{n-1}{n(2n-1)} = \\ &= \frac{(n-1)(2n-1+1)}{n(2n-1)} = \frac{2n-2}{2n-1}\end{aligned}$$

(בפרט, אפשר לראות שכאשר n גדול מאוד, השונות שואפת ל-1. למרות שיכולים להיות הרבה זוגות שעומדים זה לצד זה, הסיכוי לכך קטן מספיק מהר, ונצפה שכמות הזוגות שעומדים זה לצד זה תהיה קטנה גם אם n גדול מאוד.)

הגדרה 9.10. יהיו X ו-Y משתנים מקריים. **מקדם המתאם של פירסון** של X ו-Y מוגדר להיות

$$\rho\left(X, Y\right) = \frac{\operatorname{Cov}\left(X, Y\right)}{\sqrt{\operatorname{Var}\left(X\right)} \cdot \sqrt{\operatorname{Var}\left(Y\right)}}.$$

טענה 9.11. יהיו X ו-Y משתנים מקריים.

א. $-1 \leq \rho\left(X, Y\right) \leq 1$.

ב. $\rho\left(X, Y\right) = 0$ אם ורק אם X ו-Y בלתי־מתואמים.

ג. $\rho\left(X, Y\right) = 1$ אם ורק אם $Y = aX + b$ עבור $a > 0$.

ד. $\rho\left(X, Y\right) = -1$ אם ורק אם $Y = aX + b$ עבור $a < 0$.

כפי שרואים מהטענה, מקדם המתאם עוזר לנו למדוד קשרים לינאריים בין משתנים מקריים. העובדה שיש משתנים מקריים בלתי־מתואמים שהם לא בלתי־תלויים מראה לנו שמקדם המתאם יודע לראות קשרים לינאריים, אבל עשוי לפספס קשרים סטטיסטיים (כמו תלות יותר מורכבת). ככל שמקדם המתאם קרוב לאפס, כך גם הקשר הלינארי בין המשתנים חלש יותר.

10 תרגול עשירי

10.1 ריכוז מידה

המטרה. בהינתן משתנה מקרי X, להבין מה הסיכוי שלו לקבל ערך "קיצוני" אם אנחנו יודעים רק את התוחלת והשונות של X.

טענה 10.1 (אי-שוויון מרקוב). לכל $X \geq 0$ משתנה מקרי אי-שלילי, אז לכל $a > 0$ מתקיים

$$.P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

תרגיל 10.2. השכר הממוצע במשק הוא $11,000$ ש"ח. הוכיחו שבעשירון העליון יש מישהו שמרוויח לכל היותר $110,000$ ש"ח.

פתרון. נסמן על ידי X את המשתנה המקרי של השכר של כל אדם. לפי הנתון $\mathbb{E}\left[X\right] = 11000$. אם השכר המינימלי של האנשים בעשירון העליון הוא a, אז $P(X \geq a) = 0.1$. לפי אי-שוויון מרקוב,

$$0.1 = P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

ולכן $a \leq \frac{\mathbb{E}[X]}{0.1} = 110,000$, כנדרש.

תרגיל 10.3. תנו חסם לא טריוויאלי לסיכוי שבכיתה של 30 תלמידים יהיו לפחות חמישה זוגות של תלמידים שנולדו באותו היום.

פתרון. נסמן על ידי $I_{i,j}$ את משתנה האינדיקטור של המאורע "התלמיד ה-i וה-j נולדו באותו היום", ונסמן $X = \sum_{1 \leq i < j \leq 30} I_{i,j}$ כמספר הזוגות של תלמידים שנולדו באותו היום. לפי אדיטיביות התוחלת,

$$.\mathbb{E}[X] = \sum_{1 \leq i < j \leq 30} \mathbb{E}[I_{i,j}] = \binom{30}{2} \cdot \frac{1}{365} \approx 1.19$$

מאי-שוויון מרקוב,

$$.P(X \geq 5) \leq \frac{1.19}{5} \approx 0.24$$

טענה 10.4 (אי-שוויון צ'בישב). לכל משתנה מקרי X בעל שונות ולכל $a > 0$,

$$.P(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}$$

תרגיל 10.5. שפרו את החסם מתרגיל 10.3.

פתרון. נרצה לשפר את החסם שנתנו קודם על ידי שימוש באי-שוויון צ'בישב. לצורך כך נרצה לחשב את השונות של X. הבעיה היא שהמשתנים $I_{i,j}$ הם כן תלויים (למשל בשלשות). לכן

$$.\mathrm{Var}(X) = \mathrm{Var}(\sum_{i,j} I_{i,j}) = \sum_{(i,j)} \mathrm{Var}(I_{i,j}) + \sum_{(i,j) \neq (k,l)} \mathrm{Cov}(I_{i,j}, I_{k,l})$$

נחשב כל גורם:

- השונות של כל $I_{i,j}$ היא $\mathrm{Var}(I_{i,j}) = \frac{1}{365} \cdot \frac{364}{365}$ (כי זו השונות של משתנה ברנולי).
- אם הזוגות (i,j) ו-(k,l) זרים, אז המשתנים $I_{i,j}$ ו-$I_{k,l}$ בלתי־תלויים, והשונות המשותפת שלהם תהיה 0.
- אם יש איבר משותף בין הזוגות, נניח שהם (i,j) ו-(i,k), אז

$$\mathbb{E}[I_{i,j}I_{i,k}] = P(\text{לשלושתם יש אותו יום הולדת}) = \frac{1}{365^2}$$

ולכן

$$\mathrm{Cov}(I_{i,j}, I_{i,k}) = \frac{1}{365^2} - \left(\frac{1}{365}\right)^2 = 0$$

כלומר הם עדיין בלתי־מתואמים (ולמעשה, בלתי־תלויים).

לכן

$$.\mathrm{Var}(X) = \sum_{(i,j)} \mathrm{Var}(I_{i,j}) = \binom{30}{2} \cdot \frac{364}{365^2} \approx 1.1885$$

נזכור כי $\mathbb{E}[X] \approx 1.19$, ולכן

$$P(X \geq 5) = P(X - 1.19 \geq 5 - 1.19) \leq P(|X - \mathbb{E}[X]| \geq 3.81) \leq \frac{1.1885}{3.81^2} \approx 0.08$$

שזה שיפור משמעותי ביחס לחסם מרקוב שהשגנו.

10.2 חוק המספרים הגדולים

בהרבה מקרים האינטואיציה שלנו להסתברות נובעת מאמירות כמו "ההסתברות שהטלת מטבע תהיה עץ היא $\frac{1}{2}$, כי אם נטיל מטבע הרבה פעמים – בערך בחצי מהפעמים נקבל עץ". חוק המספרים הגדולים נותן ביסוס פורמלי לאינטואיציה שלנו.

משפט 10.6 (חוק המספרים הגדולים בהנחת שונות)**.** יהיו $X_1, X_2, \ldots$ משתנים מקריים בלתי־תלויים ושווי התפלגות, בעלי תוחלת μ ובעלי שונות. אזי לכל $\varepsilon > 0$ מתקיים

$$.\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) = 0$$

דוגמה 10.7. אם $X_1, X_2, \cdots \sim \mathrm{Ber}(\frac{1}{2})$ מציינים האם בהטלת מטבע קיבלנו עץ, אז $\frac{1}{n}\sum_{i=1}^{n} X_i$ מייצג את אחוז ההטלות מבין n ההטלות הראשונות שבהן קיבלנו עץ, והמשפט אומר שהגודל הזה מתקרב ל-$\frac{1}{2}$ כאשר n גדול.

תרגיל 10.8. יהיו $X_1, X_2, \ldots$ משתנים מקריים בלתי־תלויים המתפלגים $\mathrm{Geo}(\frac{1}{2})$. לכל n, יהי $Y_n = \sum_{i=1}^{n} X_i \sim \mathrm{NB}(n, \frac{1}{2})$. חשבו את הגבולות הבאים:

א. $\lim_{n\to\infty} P\left(Y_n \geq 3n\right)$.

ב. $\lim_{n\to\infty} P\left(Y_n \geq n+1\right)$.

פתרון. נזכור כי $\mathbb{E}\left[X_i\right] = 2$, והם בעלי שונות. לכן אפשר להשתמש בחוק המספרים הגדולים.

א. נכתוב את ההסתברות בצורה הבאה:

$$.P\left(Y_n \geq 3n\right) = P\left(\frac{1}{n}\sum_{i=1}^{n} X_i - 2 \geq 1\right) \leq P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - 2\right| \geq 1\right)$$

על פי חוק המספרים הגדולים, ההסתברות שבאגף ימין שואפת ל-0, ולכן גם ההסתברות שבאגף שמאל שואפת ל-0, וזו התשובה.

ב. כיוון ש-$Y_n \sim \mathrm{NB}(n, \frac{1}{2})$, בהכרח $Y_n \geq n$. לכן

$$P\left(Y_n \geq n+1\right) = 1 - P\left(Y_n < n+1\right) = 1 - P\left(Y_n = n\right) = 1 - p^n$$

והביטוי הזה שואף ל-1 כאשר $n \to \infty$.

10.3 תוחלת מותנית

הגדרה 10.9. יהי X משתנה מקרי, ויהי A מאורע מהסתברות חיובית. מגדירים את $\mathbb{E}\left[X \mid A\right] = \sum_a a \cdot P\left(X = a \mid A\right)$. זו בעצם התוחלת של X ביחס לפונקציית ההסתברות $P\left(\cdot \mid A\right)$.

טענה 10.10 (נוסחת התוחלת השלמה). אם $\Omega = \bigcup_i A_i$ פירוק של המרחב לקבוצות זרות, אז

$$.\mathbb{E}\left[X\right] = \sum_i \mathbb{E}\left[X \mid A_i\right] \cdot P\left(A_i\right)$$

תרגיל 10.11. מטילים קוביה. אם התוצאה היא מספר זוגי, עליי לשלם את המספר שיצא בשקלים. אם התוצאה היא מספר אי-זוגי, אני מרוויח פי שניים מהמספר הכתוב. מה תוחלת הרווח שלי?

פתרון. נסמן על ידי X את הרווח שלי. אז

$$\mathbb{E}\left[X \mid \text{יצא זוגי}\right] = \frac{1}{3} \cdot (-2) + \frac{1}{3} \cdot (-4) + \frac{1}{3} \cdot (-6) = -4$$
$$\mathbb{E}\left[X \mid \text{יצא אי-זוגי}\right] = \frac{1}{3} \cdot 2 + \frac{1}{3} \cdot 6 + \frac{1}{3} \cdot 10 = 6$$

לכן

$$\begin{aligned}\mathbb{E}\left[X\right] &= \mathbb{E}\left[X \mid \text{יצא זוגי}\right] \cdot P\left(\text{יצא זוגי}\right) + \mathbb{E}\left[X \mid \text{יצא אי-זוגי}\right] \cdot P\left(\text{יצא אי-זוגי}\right) = \\ &= (-4) \cdot \frac{1}{2} + 6 \cdot \frac{1}{2} = 1\end{aligned}$$

הגדרה 10.12. יהיו X ו-Y משתנים מקריים. לכל תוצאה b של Y אפשר להסתכל על המשתנה המקרי $X \mid Y = b$. נגדיר את **התוחלת המותנית של X ב-Y**, שנסמן $\mathbb{E}\left[X \mid Y\right]$, להיות המשתנה המקרי המוגדר לפי

$$.\mathbb{E}\left[X \mid Y\right](\omega) = \mathbb{E}\left[X \mid Y = Y(\omega)\right]$$

אינטואיטיבית: ההתנייה של X ב-Y אומרת שאנחנו רוצים ללמוד את ההתפלגות של X אם אנחנו יודעים מה ההתפלגות של Y. המידע הכי טוב שאנחנו יכולים לספק הוא התוחלת של X.

דוגמה 10.13. מטילים מטבע פעמיים. X יסמן את תוצאת ההטלה הראשונה, Y יסמן את כמות הפעמים שקיבלנו 1. המשתנה $X \mid Y=0$ הוא משתנה שקבוע ל-0, לכן $\mathbb{E}\left[X \mid Y=0\right]=0$. המשתנה $X \mid Y=1$ הוא משתנה אחיד על $0,1$, לכן $\mathbb{E}\left[X \mid Y=1\right]=\frac{1}{2}$. המשתנה $X \mid Y=2$ הוא משתנה שקבוע ל-1, לכן $\mathbb{E}\left[X \mid Y=2\right]=1$.

טענה 10.14 (חוק התוחלת החוזרת). $\mathbb{E}\left[\mathbb{E}\left[X \mid Y\right]\right]=\mathbb{E}\left[X\right]$.

תרגיל 10.15. מטילים מטבעות הוגנים על שולחן עד לקבלת "עץ". עכשיו לוקחים את אחד המטבעות האלו (באופן אקראי); אם הוא ה"עץ", אני מרוויח את כל המטבעות, ואחרת לא מקבל כלום. מה תוחלת הרווח שלי?

פתרון. נסמן על ידי X את כמות המטבעות עד לקבלת "עץ", ונסמן על ידי Y את הרווח שלי בסוף המשחק. מה ההתפלגות של Y בהינתן X?

$$P\left(Y=a \mid X=n\right)=\begin{cases} n, & \frac{1}{n} \\ 0, & 1-\frac{1}{n} \end{cases}.$$

לכן

$$\mathbb{E}\left[Y \mid X=n\right]=n\cdot\frac{1}{n}+0\cdot\frac{n-1}{n}=1$$

ולפי חוק התוחלת החוזרת,

$$\mathbb{E}\left[X\right]=\mathbb{E}\left[\mathbb{E}\left[X \mid Y\right]\right]=\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot 1=1.$$

תרגיל 10.16. וולדמורט מענה את סיריוס במשרד הקסמים. הוא מטיל קוביה. אם התוצאה היא 1 וולדמורט מתעצבן ורוצח את סיריוס, ואם התוצאה היא $n>1$, הוא מענה את סיריוס למשך $n-1$ דקות, ואז חוזר על התהליך. כמה זמן סיריוס צפוי לסבול?

פתרון. נסמן על ידי X את הזמן הצפוי עד שסיריוס יוצא להורג, ועל ידי Y את תוצאת הקוביה. אז

$$\mathbb{E}\left[X \mid Y=k\right]=\begin{cases} 0, & k=1 \\ \mathbb{E}\left[X\right]+k-1, & k>1 \end{cases}.$$

לכן

$$\mathbb{E}\left[X\right]=\sum_{k=1}^{6}\mathbb{E}\left[X \mid Y=k\right]\cdot P\left(Y=k\right)=\frac{1}{6}\cdot 0+\frac{1}{6}\sum_{k=2}^{6}\left(\mathbb{E}\left[X\right]+k-1\right)=\frac{5}{6}\cdot\mathbb{E}\left[X\right]+\frac{1}{6}\cdot 15$$

ומכאן נקבל $\mathbb{E}\left[X\right]=15$.

טענה 10.17 (חוק פירוק השונות). אם X ו-Y משתנים מקריים, אז

$$\mathrm{Var}\left(X\right)=\mathbb{E}\left[\mathrm{Var}\left(X \mid Y\right)\right]+\mathrm{Var}\left(\mathbb{E}\left[X \mid Y\right]\right)$$

תרגיל 10.18. מספר שיחות הטלפון המתקבל במרכזייה מתפלג $\mathrm{Poi}\left(\lambda\right)$. שיחה היא חשובה בהסתברות p. חשבו את מקדם המתאם בין מספר השיחות בשעה למספר השיחות החשובות בשעה.

פתרון. נסמן את מספר השיחות ב-N ואת מספר השיחות החשובות ב-X. מהנתונים, $N \sim \mathrm{Poi}\left(\lambda\right)$ ו-$X \sim \mathrm{Bin}\left(N,p\right)$. כדי לחשב את השונות המשותפת, נחשב:

$$\begin{aligned}\mathbb{E}\left[N\right] &= \lambda \\ \mathbb{E}\left[X\right] &= \mathbb{E}\left[\mathbb{E}\left[X \mid N\right]\right] = \mathbb{E}\left[Np\right] = p\lambda \\ \mathbb{E}\left[XN\right] &= \mathbb{E}\left[\mathbb{E}\left[XN \mid N\right]\right] = \mathbb{E}\left[N \cdot \mathbb{E}\left[X \mid N\right]\right] = \mathbb{E}\left[N \cdot Np\right] = p \cdot \mathbb{E}\left[N^2\right].\end{aligned}$$

נזכור כי $\mathrm{Var}\left(N\right) = \mathbb{E}\left[N^2\right] - \mathbb{E}\left[N\right]^2$, ולכן $\mathbb{E}\left[N^2\right] = \lambda^2 + \lambda$. כלומר

$$\mathbb{E}\left[XN\right] = p\left(\lambda^2 + \lambda\right).$$

בסך הכל

$$\mathrm{Cov}\left(X,N\right) = p\left(\lambda^2 + \lambda\right) - p\lambda^2 = p\lambda.$$

נחשב את השונויות:

$$\mathrm{Var}\left(N\right) = \lambda$$

ועבור X ניעזר בפירוק השונות:

$$\begin{aligned}\mathrm{Var}\left(X\right) &= \mathbb{E}\left[\mathrm{Var}\left(X \mid N\right)\right] + \mathrm{Var}\left(\mathbb{E}\left[X \mid N\right]\right) = \\ &= \mathbb{E}\left[Np\left(1-p\right)\right] + \mathrm{Var}\left(Np\right) = \\ &= p\left(1-p\right)\lambda + p^2\lambda = p\lambda\end{aligned}$$

בסך הכל

$$\rho\left(X,N\right) = \frac{p\lambda}{\sqrt{\lambda} \cdot \sqrt{\lambda p}} = \sqrt{p}.$$

11 תרגול אחד עשר

11.1 מרחבי הסתברות כלליים

הגדרה 11.1. ***מרחב הסתברות*** הוא שלשה $\left(\Omega,\mathcal{F},P\right)$, כאשר:

א. Ω הוא מרחב המדגם שלנו.

ב. $\mathcal{F}$ היא ***σ-אלגברה***, כלומר משפחה של תת־קבוצות של Ω המקיימת:

(i) $\Omega \in \mathcal{F}$.

(ii) אם $A \in \mathcal{F}$ אז $A^c \in \mathcal{F}$.

(iii) אם $A_1, A_2, \dots \in \mathcal{F}$ אז $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

ג. P היא ***פונקציית הסתברות***, כלומר פונקציה $P\colon \mathcal{F} \to [0,1]$ המקיימת:

(i) $P\left(\Omega\right) = 1$.

(ii) σ-אדיטיביות: אם $A_1, A_2, \dots \in \mathcal{F}$ זרות בזוגות, אז מתקיים $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P\left(A_n\right)$.

11.2 משתנים מקריים רציפים

הגדרה 11.2. יהי $(\Omega, \mathcal{F}, P)$ מרחב הסתברות. **משתנה מקרי** על Ω הוא פונקציה $X : \Omega \to \mathbb{R}$ כך שלכל $a < b$, $X^{-1}((a,b)) \in \mathcal{F}$.

דוגמה 11.3.

א. לכל מרחב הסתברות, $X = 165$ הוא משתנה מקרי.

ב. אם $\Omega = \mathbb{R}$ ו-$\mathcal{F} = \mathbb{B}(\mathbb{R})$, אז $X = 1_{[a,b]}$ הוא משתנה מקרי.

ההתפלגות של משתנה מקרי. נרצה כלים לתאר את ההתפלגות של משתנה מקרי. הבעיה היא שפונקציית ההתפלגות שהכרנו עד כה, $a \mapsto P(X = a)$, היא לא טובה. למשל, אם אנחנו מגרילים מספר באופן אחיד בין 0 ל-1, הסיכוי לקבל כל נקודה בודדת הוא 0. אז הפונקציה הזו לא תיתן לנו אינפורמציה בכלל.

הגדרה 11.4. פונקציית ההתפלגות המצטברת (cummulative distribution function) של משתנה מקרי X היא

$$F_X(a) = P(X \leq a).$$

טענה 11.5. F_X מקיימת את התכונות הבאות:

א. F_X מונוטונית לא יורדת.

ב. $\lim_{a\to\infty} F_X(a) = 1$, $\lim_{a\to-\infty} F_X(a) = 0$.

ג. F_X רציפה מימין. כלומר, $\lim_{h\to 0^+} F_X(a+h) = F_X(a)$.

ד. F_X רציפה בנקודה a אם ורק אם $P(X = a) = 0$.

הגדרה 11.6. אומרים ש-X **משתנה מקרי רציף בהחלט** אם קיימת פונקציה $f_X : \mathbb{R} \to \mathbb{R}$ המקיימת

$$P(a < X < b) = \int_a^b f_X(t)\, \mathrm{d}t$$

לכל $a < b$ ממשיים. לפונקציה f_X קוראים **פונקציית הצפיפות** (probability density function) של X, ומתקיים $f_X = F_X'$.

טענה 11.7. f_X מקיימת את התכונות הבאות:

א. $\int_{-\infty}^{\infty} f_X(x)\, \mathrm{d}x = 1$.

ב. לכל $a < b$, $\int_a^b f_X(x)\, \mathrm{d}x = P(a \leq X \leq b)$.

דוגמה 11.8. איך נגדיר הגרלה אחידה בקטע $[0,1]$? נרצה להגיד שהסיכוי ליפול בקטע מסוים הוא פרופורציוני לאורך שלו בתוך $[0,1]$. כלומר לכל $0 \leq a < b \leq 1$ נרצה

$$P(a \leq X \leq b) = b - a.$$

אז אפשר להגדיר

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x \geq 1 \end{cases}$$

ופונקציית הצפיפות היא

$$.f_X(x)=\begin{cases}0, & x>1 \text{ או } x<0\\ 1, & 0<x<1\end{cases}$$

איך צריך לדמיין את f_X? ככל שהשטח מתחת לגרף שלה גדול יותר, כך ההסתברות ליפול באותו תחום גדולה.

תרגיל 11.9. יהי X משתנה מקרי בעל פונקציית צפיפות

$$.f(x)=\begin{cases}cx^n, & 0\le x\le 1\\ 0, & \text{אחרת}\end{cases}$$

א. מצאו את c.

ב. מצאו את פונקציית ההתפלגות המצטברת של X.

ג. חשבו את $P\left(X\le\frac{1}{2}\right)$.

ד. חשבו את החציון של X, כלומר a שעבורו $P(X\le a)=\frac{1}{2}$.

פתרון.

א. נחשב את c לפי תכונת האינטגרל:

$$\int_{-\infty}^{\infty}f(x)\,\mathrm{d}x=1$$

אצלנו התומך הוא $[0,1]$, ולכן

$$.\int_{-\infty}^{\infty}f(x)\,\mathrm{d}x=\int_0^1 cx^n\,\mathrm{d}x=\left(\frac{cx^{n+1}}{n+1}\right)_0^1=\frac{c}{n+1}=1$$

כלומר $c=n+1$.

ב. לפי ההגדרה,

$$.F_X(x)=\int_{-\infty}^{x}f_X(t)\,\mathrm{d}x$$

אם $x\le 0$, אז $F_X(x)=0$. אם $x\ge 1$, אז $F_X(x)=1$. אם $0<x<1$,

$$.F_X(x)=\int_0^x(n+1)\,t^n\,\mathrm{d}t=x^{n+1}$$

ג. $P\left(X\le\frac{1}{2}\right)=F_X\left(\frac{1}{2}\right)=\frac{1}{2^{n+1}}$.

ד. מן הסתם $0\le a\le 1$. לכן רוצים $a^{n+1}=F_X(a)=\frac{1}{2}$, כלומר $a=\frac{1}{\sqrt[n+1]{2}}$.

הגדרה 11.10. יהי X משתנה מקרי רציף. נגדיר את **התוחלת** של X להיות

$$\mathbb{E}\left[X\right]=\int_{-\infty}^{\infty}xf_X\left(x\right)\,\mathrm{d}x$$

ואת **השונות** שלו להיות

$$\mathrm{Var}\left(X\right)=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)^2\right]=\mathbb{E}\left[X^2\right]-\mathbb{E}\left[X\right]^2.$$

מתקיימות אותן התכונות כמו שהכרנו אצל משתנים מקריים בדידים.

תרגיל 11.11. משתנה מקרי X מוגדר על ידי פונקציית הצפיפות

$$f_X\left(t\right)=\begin{cases}at^2+bt, & 0\leq t\leq 1\\ 0, & \text{אחרת}\end{cases}$$

כאשר $a,b\in\mathbb{R}$. נתון כי $\mathbb{E}\left[X\right]=0.6$. חשבו את $P\left(X\leq\frac{1}{2}\right)$ ואת $\mathrm{Var}\left(X\right)$.

פתרון. ראשית, נשתמש בתכונה של פונקציית הצפיפות:

$$1=\int_{-\infty}^{\infty}f_X\left(x\right)\,\mathrm{d}x=\int_0^1\left(at^2+bt\right)\,\mathrm{d}t=\frac{a}{3}+\frac{b}{2}.$$

מצד שני, לפי הנתון על התוחלת,

$$0.6=\mathbb{E}\left[X\right]=\int_{-\infty}^{\infty}xf_X\left(x\right)\,\mathrm{d}x=\int_0^1\left(at^3+bt^2\right)\,\mathrm{d}t=\frac{a}{4}+\frac{b}{3}.$$

קיבלנו מערכת משוואות לינאריות

$$\frac{a}{3}+\frac{b}{2}=1,\qquad \frac{a}{4}+\frac{b}{3}=0.6.$$

אם נפתור אותה נקבל $a=-2.4$ ו-$b=3.6$.
ההסתברות המבוקשת:

$$P\left(X\leq\frac{1}{2}\right)=\int_0^{\frac{1}{2}}\left(-2.4t^2+3.6t\right)\,\mathrm{d}t=\left(-0.8t^3+1.8t^2\right)_0^{\frac{1}{2}}=0.35.$$

לחישוב השונות,

$$\mathbb{E}\left[X^2\right]=\int_{-\infty}^{\infty}t^2f_X\left(t\right)\,\mathrm{d}t=\int_0^1\left(-2.4t^4+3.6t^3\right)\,\mathrm{d}t=-0.48\cdot1^5+0.9\cdot1^4=0.42$$

ונקבל

$$\mathrm{Var}\left(X\right)=\mathbb{E}\left[X^2\right]-\mathbb{E}\left[X\right]^2=0.42-0.6^2=0.42-0.36=0.06.$$

תרגיל 11.12 ("המקרה המעורב"). זמן החיים בחודשים של שפופרת אלקטרונית הוא משתנה מקרי X המוגדר באמצעות פונקציית הצפיפות

$$f_X\left(t\right)=\begin{cases}te^{-t}, & t\geq 0\\ 0, & \text{אחרת}\end{cases}.$$

א. חשבו את תוחלת זמן החיים של שפופרת.

ב. במפעל שבו מותקנות 10 שפופרות בלתי־תלויות, מה הסיכוי שלפחות שתיים מהן תחזקנה מעמד לפחות שלושה חודשים?

פתרון.

א. נחשב על פי הנוסחה:

$$\mathbb{E}\left[X\right] = \int_0^\infty t f_X\left(t\right)\,\mathrm{d}t = \int_0^\infty t^2 e^{-t}\,\mathrm{d}t = \left(-t^2e^{-t}\right)_0^\infty - \int_0^\infty 2t\left(-e^{-t}\right)\,\mathrm{d}t = \\ = 2\int_0^\infty te^{-t}\,\mathrm{d}t = 2$$

ב. הסיכוי של שפופרת מסוימת להחזיק מעמד יותר משלושה חודשים הוא

$$P\left(X \geq 3\right) = \int_3^\infty f_X\left(t\right)\,\mathrm{d}t = \int_3^\infty te^{-t}\,\mathrm{d}t = \left(-te^{-t}\right)_3^\infty - \int_3^\infty \left(-e^{-t}\right)\,\mathrm{d}t = \\ = 3e^{-3} + \int_3^\infty e^{-t}\,\mathrm{d}t = 3e^{-3} + \left(-e^{-t}\right)_3^\infty = 4e^{-3}$$

כעת, ההסתברות המבוקשת היא ההסתברות ללפחות 2 הצלחות מתוך 10 ניסויים בלתי־תלויים, כלומר התפלגות בינומית:

$$P\left(\mathrm{Bin}\left(10, 4e^{-3}\right) \geq 2\right) = 1 - P\left(\mathrm{Bin}\left(10, 4e^{-3}\right) < 2\right) = \\ = 1 - \left(1-4e^{-3}\right)^{10} - 10\cdot 4e^{-3}\cdot\left(1-4e^{-3}\right)^9 \approx 0.622$$

11.3 טרנספורמציות של משתנים מקריים

בהרבה מקרים יהיה נתון לנו משתנה מקרי X שאנחנו מבינים, ונרצה ללמוד באמצעותו התפלגות של טרנספורמציה עליו, $h\left(X\right)$.

תרגיל 11.13. נתון משתנה מקרי X בעל צפיפות $f_X\left(t\right) = 2te^{-t^2}, t > 0$. מצאו את פונקציית הצפיפות של $Y = X^2$.

פתרון. נחשב את פונקציית ההתפלגות המצטברת של Y, ואז נסיק את הצפיפות כנגזרת. לכל $y \geq 0$,

$$F_Y\left(y\right) = P\left(Y \leq y\right) = P\left(X^2 \leq y\right) = P\left(-\sqrt{y} \leq X \leq \sqrt{y}\right) = F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right) = F_X\left(\sqrt{y}\right)$$

נגזור ונקבל

$$f_Y\left(y\right) = \frac{\mathrm{d}}{\mathrm{d}y}F_Y\left(y\right) = \frac{\mathrm{d}}{\mathrm{d}y}\left(F_X\left(\sqrt{y}\right)\right) = \frac{1}{2\sqrt{y}}f_X\left(\sqrt{y}\right) = \frac{1}{2\sqrt{y}}\cdot 2\sqrt{y}e^{-y} = e^{-y}.$$

זו **התפלגות מעריכית** שנפגוש בהמשך.

תרגיל 11.14. יהי X משתנה מקרי אחיד על $[0,1]$, כלומר

$$f_X\left(x\right) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{אחרת} \end{cases}.$$

מהי ההתפלגות של $Y = -\log X$?

פתרון. נחשב:

$$F_Y(y)=P(Y\le y)=P(-\log X\le y)=P\left(X\ge e^{-y}\right)=1-F_X\left(e^{-y}\right).$$

קל לוודא שפונקציית ההתפלגות המצטברת של X היא

$$F_X(x)=\begin{cases}0, & x<0\\ x, & 0\le x\le 1\\ 1, & x>1\end{cases}$$

ולכן

$$F_Y(y)=\begin{cases}0, & y<0\\ 1-e^{-y}, & y\ge 0\end{cases}.$$

אם רוצים את הצפיפות נקבל

$$f_Y(y)=\begin{cases}0, & y<0\\ e^{-y}, & y\ge 0\end{cases}$$

תרגיל 11.15 (אם יש זמן)**.** נתון $X\sim U\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$. מהי הצפיפות של $Y=\tan X$?

פתרון. נשים לב כי

$$F_X(x)=\begin{cases}0, & x<-\frac{\pi}{2}\\ \frac{x+\frac{\pi}{2}}{\pi}, & -\frac{\pi}{2}\le x\le\frac{\pi}{2}\\ 1, & x>\frac{\pi}{2}\end{cases}$$

ולכן

$$P(Y\le y)=P(\tan X\le y)=P(X\le\arctan y)$$

שימו לב שתמיד $\arctan X\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$, ולכן

$$P(Y\le y)=\frac{\arctan y+\frac{\pi}{2}}{\pi},$$

כלומר פונקציית הצפיפות היא

$$f_Y(y)=\frac{1}{\pi\left(1+y^2\right)}.$$

להתפלגות הזו קוראים **התפלגות קושי סטנדרטית**.

שימו לב שלהתפלגות הזו אין תוחלת! זה נכון כי

$$\mathbb{E}[Y]=\int_{-\infty}^{\infty}\frac{y}{\pi\left(1+y^2\right)}\,\mathrm{d}y$$

לא מתכנס.

12 תרגול שנים עשר

12.1 זוגות של משתנים מקריים רציפים

הגדרה 12.1. יהיו X,Y משתנים מקריים רציפים. **פונקציית הצפיפות המשותפת שלהם** היא פונקציה $f_{X,Y}:\mathbb{R}^2\to\mathbb{R}$ המקיימת

$$.P\left(a\le X\le b,c\le Y\le d\right)=\int_a^b\int_c^d f_{X,Y}\left(x,y\right)\,\mathrm{d}x\,\mathrm{d}y$$

אומרים ש-X,Y הם **בלתי־תלויים** אם $.f_{X,Y}\left(x,y\right)=f_X\left(x\right)\cdot f_Y\left(y\right)$

במקרה הזה, כדי לחשב הסתברויות על זוג המשתנים המקריים, נבצע אינטגרל כפול על השטח הרצוי.

תרגיל 12.2. נתונים שני משתנים מקריים X ו-Y בעלי צפיפות משותפת

$$.f_{X,Y}\left(x,y\right)=\begin{cases}c\left(x^2+\frac{xy}{2}\right), & 0<x<1,\ 0<y<2\\ 0, & \text{אחרת}\end{cases}$$

א. מצאו את ערכו של c.

ב. חשבו את פונקציית הצפיפות השולית של X.

ג. חשבו את ההסתברות ש-$X>Y$.

פתרון.

א. נבצע אינטגרל כפול ונשווה ל-1:

$$\begin{aligned}1&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)\,\mathrm{d}y\,\mathrm{d}x=\int_0^1\int_0^2 c\left(x^2+\frac{xy}{2}\right)\,\mathrm{d}y\,\mathrm{d}x=\\&=c\int_0^1\left(x^2y+\frac{xy^2}{4}\right)_0^2\mathrm{d}x=c\int_0^1\left(2x^2+x\right)\,\mathrm{d}x=c\left(\frac{2}{3}x^3+\frac{1}{2}x^2\right)_0^1=\\&=c\cdot\left(\frac{2}{3}+\frac{1}{2}\right)=c\cdot\frac{7}{6}\end{aligned}$$

לכן $.c=\frac{6}{7}$

ב. נחשב את פונקציית הצפיפות השולית:

$$\begin{aligned}f_X\left(x\right)&=\int_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)\,\mathrm{d}y=\int_0^2\frac{6}{7}\left(x^2+\frac{xy}{2}\right)\,\mathrm{d}y=\\&=\frac{6}{7}\left(x^2y+\frac{xy^2}{4}\right)_0^2=\frac{6}{7}\left(2x^2+x\right)\end{aligned}$$

ג. כדי לחשב את ההסתברות הזו, צריך לבצע אינטגרל כפול על השטח המתאים למאורע $X > Y$. במקרה שלנו,

$$P(X > Y) = \int_0^1 \int_0^x \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \mathrm{d}y\,\mathrm{d}x = \frac{6}{7}\int_0^1 \left(x^2 y + \frac{xy^2}{4}\right)_0^x \mathrm{d}x =$$
$$= \frac{6}{7}\int_0^1 \left(x^3 + \frac{x^3}{4}\right) \mathrm{d}x = \frac{15}{14}\left(\frac{x^4}{4}\right)_0^1 = \frac{15}{56}$$

אופציה אחרת: להשתמש בנוסחת ההסתברות השלמה הרציפה. אז

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{x^2 + \frac{xy}{2}}{2x^2 + x} = \frac{2x + y}{4x + 2}$$

ולכן

$$P(X > Y) = \int_0^1 F_{Y|X=x}(x) f_X(x)\,\mathrm{d}x$$

ואז צריך לחשב גם את פונקציית ההצטברות של $Y \mid X = x$ (תרגיל לבית).

12.2 התפלגויות רציפות מיוחדות

הגדרה 12.3. התפלגות אחידה רציפה – משתנה מקרי X המקבל ערכים בקטע סופי $[a,b]$. פונקציית ההתפלגות המצטברת היא

$$F_X(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & x \ge b \end{cases}.$$

פונקציית הצפיפות היא

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{אחרת} \end{cases}.$$

מסמנים $X \sim U[a,b]$. התוחלת שלו היא $\frac{a+b}{2}$, והשונות $\frac{(b-a)^2}{12}$.

תרגיל 12.4. נניח כי $X \sim U[0,5]$ משתנה מקרי רציף. מהי ההסתברות ששורשי המשוואה

$$4t^2 + 4Xt + X + 2 = 0$$

הם ממשיים?

פתרון. שורשי המשוואה $at^2 + bt + c = 0$ הם ממשיים אם ורק אם $b^2 - 4ac \ge 0$. עבור המשוואה שלנו רוצים

$$16(X-2)(X+1) = 16\left(X^2 - X - 2\right) = 16X^2 - 16X - 32 = (4X)^2 - 4\cdot 4\cdot (X+2) \ge 0,$$

כלומר $X \ge 2$ או $X \le -1$. אבל $X \sim U[0,5]$, לכן

$$P(X \le -1 \text{ או } X \ge 2) = P(X \ge 2) = \frac{3}{5}.$$

הגדרה 12.5. התפלגות מעריכית – מתארת את הזמן עד שמתרחש אירוע שהסיכוי שלו להתרחש קבוע בזמן. מסמנים $X \sim \mathrm{Exp}(\lambda)$. פונקציית ההתפלגות המצטברת היא

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

ופונקציית הצפיפות היא

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

אם $X \sim \mathrm{Exp}(\lambda)$, אז $\mathbb{E}[X] = \frac{1}{\lambda}$ והשונות היא $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.

טענה 12.6. התפלגות מעריכים מקיימת את תכונת חוסר הזיכרון, כלומר

$$P(X > a + b \mid X > a) = P(X > b).$$

בנוסף, זו ההתפלגות הרציפה היחידה עם תומך $[0, \infty)$ המקיימת את תכונת חוסר הזיכרון.

הערה 12.7. אם מאורעות מתרחשים בקצב פואסוני עם ממוצע λ, אז משך הזמן עד שקורה מאורע מתפלג $\mathrm{Exp}(\lambda)$.

תרגיל 12.8. כמות האוטובוסים המגיעה לתחנה בשעה מתפלג פואסונית עם פרמטר $\lambda = 6$. מה הסיכוי שאצטרך לחכות לפחות 20 דקות עד לאוטובוס הבא? ואם כבר חיכיתי 20 דקות, מה הסיכוי שאצטרך לחכות עוד 20 דקות?

פתרון. יהי X כמות הזמן שצריך לחכות בדקות. לפי הערה 12.7, $X \sim \mathrm{Exp}\left(\frac{1}{10}\right)$. אז

$$P(X \geq 20) = 1 - F_X(20) = 1 - \left(1 - e^{-\frac{1}{10} \cdot 20}\right) = e^{-2} = 0.1353.$$

זו גם התשובה לשאלה השנייה בגלל חוסר הזיכרון.

תרגיל 12.9. הגובה של הרי געש מתפלג $X \sim \mathrm{Exp}(\lambda)$.

א. מה ההסתברות שמבין n הרי געש שנבחרו באקראי, יהיו בדיוק m שגובהם מעל λ?

ב. בוחרים שני הרי געש אקראיים בלתי־תלויים. מה ההסתברות שהראשון גבוה לפחות פי 2 מהשני?

ג. מטפסת הרים מטפסת מדי יום על הר אקראי. עלות הטיפוס על כל הר היא S, וכאשר היא מוצאת הר שגובהו יותר מ-λ היא זוכה בפרס בגובה R. מה צריך להיות ערך הפרס על מנת שתוחלת הרווח של המטפסת תהיה חיובית?

פתרון.

א. הסיכוי של הר געש מסוים להיות עם גובה גדול מ-λ הוא $P(X > \lambda) = e^{-\lambda^2}$. לפיכך אנחנו מתעניינים בסיכוי

$$P\left(\mathrm{Bin}\left(n, e^{-\lambda^2}\right) = m\right) = \binom{n}{m} e^{-m\lambda^2} \left(1 - e^{-\lambda^2}\right)^{n-m}.$$

ב. נסמן את הגבהים שלהם X,Y. אפשר לפתור על ידי אינטגרציה כפולה:

$$\begin{aligned}
P\left(X>2Y\right) &= \int_0^\infty \int_{2y}^\infty f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\infty \int_{2y}^\infty f_X(x) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y = \\
&= \int_0^\infty f_Y(y)\left(\int_{2y}^\infty f_X(x)\,\mathrm{d}x\right)\,\mathrm{d}y = \int_0^\infty f_Y(y)\cdot P\left(X>2y\right)\,\mathrm{d}y = \\
&= \int_0^\infty \lambda e^{-\lambda y}\cdot e^{-2\lambda y}\,\mathrm{d}y = \left(-\frac{1}{3}e^{-3\lambda y}\right)_0^\infty = \frac{1}{3}
\end{aligned}$$

ג. אם היא מטפסת על T הרים, אז $T\sim\mathrm{Geo}\left(e^{-\lambda^2}\right)$. לכן תוחלת הרווח שלה היא

$$\mathbb{E}\left[R-T\cdot S\right] = R-\mathbb{E}\left[T\right]\cdot S = R-e^{\lambda^2}S$$

ונקבל שכדי שהתוחלת תהיה חיובית צריך $R>e^{\lambda^2}S$.

הגדרה 12.10. התפלגות נורמלית – התפלגות נורמלית עם תוחלת μ ושונות σ^2 היא ההתפלגות $X\sim\mathrm{N}\left(\mu,\sigma^2\right)$ בעלת פונקציית הצפיפות

$$f_X\left(x\right)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

אם $X\sim\mathrm{N}\left(\mu,\sigma^2\right)$ אז אפשר לנרמל ולקבל $\frac{X-\mu}{\sigma}\sim\mathrm{N}\left(0,1\right)$.

כדי לחשב את ההסתברויות של התפלגות נורמלית, אפשר להיעזר בטבלת התפלגות נורמלית.

תרגיל 12.11. ציון ה-IQ באוכלוסיה מתפלג $X\sim\mathrm{N}\left(100,15^2\right)$.

א. מהי ההסתברות שלאדם יש IQ פחות מ-70?

ב. מהו ה-IQ המינימלי שצריך על מנת להתקבל למועדון "המאיון העליון" של מנת המשכל?

פתרון. הנרמול הוא $Z=\frac{X-100}{15}\sim\mathrm{N}\left(0,1\right)$.

א. ההסתברות היא

$$P\left(X<70\right)=P\left(\frac{X-100}{15}<\frac{70-100}{15}\right)=P\left(Z<-2\right)=0.02275.$$

ב. אנחנו רוצים לפתור $P\left(X>a\right)=0.01$, ובאופן שקול $P\left(Z>\frac{a-100}{15}\right)=0.01$. אפשר לבדוק בטבלה ולראות שצריך $\frac{a-100}{15}\approx 2.325$, ולכן $a\approx 134.9$.

תרגיל 12.12. מגרילים נקודות על ציר ה-x וציר ה-y לפי $P,Q\sim\mathrm{N}\left(0,1\right)$ באופן בלתי־תלוי. חשבו את תוחלת שטח המשולש שהנקודות האלו יוצרות עם הצירים.

פתרון. שטח המשולש הוא $S = \frac{|P|\cdot|Q|}{2}$. לכן

$$\mathbb{E}\left[S\right] = \mathbb{E}\left[\frac{|P|\cdot|Q|}{2}\right] = \frac{1}{2}\mathbb{E}\left[|P|\right]\mathbb{E}\left[|Q|\right].$$

כעת, נחשב שאם $Z \sim \mathrm{N}\left(0,1\right)$, אז

$$\mathbb{E}\left[|Z|\right] = \int_{-\infty}^{\infty}|t|\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,\mathrm{d}t = \frac{2}{\sqrt{2\pi}}\int_0^{\infty}te^{-\frac{t^2}{2}}\,\mathrm{d}t = \sqrt{\frac{2}{\pi}}\left(-e^{-\frac{t^2}{2}}\right)_0^{\infty} = \sqrt{\frac{2}{\pi}}.$$

לכן

$$\mathbb{E}\left[S\right] = \frac{1}{\pi}.$$

תרגיל 12.13. נתון ש-X הוא משתנה מקרי המתפלג נורמלית. ידוע כי $P\left(X<0.9\right)=0.9$ ו-$P\left(X<0.7\right)=0.7$. מצאו את $P\left(X<0.5\right)$.

פתרון. התפלגות נורמלית נקבעת על פי שני פרמטרים: התוחלת μ וסטיית התקן σ. נזכור את הנרמול $Z = \frac{X-\mu}{\sigma}$. לפי הנתונים,

$$P\left(X<0.9\right) = P\left(\frac{Z-\mu}{\sigma} < \frac{0.9-\mu}{\sigma}\right) = 0.9 \Longrightarrow \frac{0.9-\mu}{\sigma} \approx 1.285$$
$$P\left(X<0.7\right) = P\left(\frac{Z-\mu}{\sigma} < \frac{0.7-\mu}{\sigma}\right) = 0.7 \Longrightarrow \frac{0.7-\mu}{\sigma} \approx 0.525$$

משתי המשוואות מקבלים $\mu \approx 0.562$ ו-$\sigma \approx 0.263$, כלומר

$$P\left(X<0.5\right) = P\left(Z < \frac{0.5-0.562}{0.263}\right) = P\left(Z<-0.236\right) = 0.408$$

הערה. סכום של משתנים מקריים נורמליים ב"ת הוא גם נורמלי: כלומר,

$$\mathrm{N}(\mu_1,\sigma_1^2) + \mathrm{N}(\mu_2,\sigma_2^2) = \mathrm{N}(\mu_1^2+\mu_2^2,\sigma_1^2+\sigma_2^2).$$

12.3 פונקציה יוצרת מומנטים

הגדרה 12.14. יהי X משתנה מקרי.

א. **המומנט מסדר** k **של** X הוא $\mathbb{E}[X^k]$.

ב. **הפונקציה יוצרת המומנטים של** X היא $M_X(t) = \mathbb{E}[e^{tX}]$.

טענה 12.15. אם $k<l$ והמומנט ה-l של X קיים, אז גם המומנט ה-k של X קיים.

טענה 12.16 (תכונות הפונקציה יוצרת המומנטים). יהיו X,Y משתנים מקריים, ויהיו $a,b\in\mathbb{R}$.

א. $M_X(t)$ מוגדרת עבור $t=0$, ומתכנסת בקטע סביב 0 (אולי אינסופי, לא בהכרח סימטרי).

ב. אם X ו-Y בלתי־תלויים, $M_{X+Y}(t) = M_X(t)\cdot M_Y(t)$.

ג. $M_{aX+b}(t) = e^{tb} \cdot M_X(at)$.

ד. אם $M_X(t)$ קיימת וסופית בקטע פתוח סביב 0, אז כל המומנטים של X קיימים, ויש לה פיתוח טיילור $M_X(t) = \sum_{n=0}^{\infty} \frac{\mathbb{E}[X^n]}{n!} t^n$. בפרט, $M_X^{(n)}(0) = \mathbb{E}[X^n]$.

ה. אם $M_X(t)$ קיימת לכל t, אז היא קובעת את ההתפלגות של X (ללא הוכחה).

תרגיל 12.17.

א. חשבו את הפונקציה יוצרת המומנטים של $X \sim \mathrm{Poi}(\lambda)$.

ב. הסיקו כי אם $X \sim \mathrm{Poi}(\lambda)$ ו-$Y \sim \mathrm{Poi}(\mu)$ בלתי־תלויים, אז $X+Y \sim \mathrm{Poi}(\lambda+\mu)$.

פתרון.

א. נחשב על פי ההגדרה:

$$M_X(t) = \mathbb{E}[e^{tX}] = \sum_{n=0}^{\infty} e^{tn} P(X=n) = \sum_{n=0}^{\infty} e^{tn} e^{-\lambda} \frac{\lambda^n}{n!} =$$
$$= e^{-\lambda} \cdot \sum_{n=0}^{\infty} \frac{(\lambda e^t)^n}{n!} = e^{-\lambda} \cdot e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

הפונקציה מוגדרת בכל $\mathbb{R}$.

ב. כיוון ש-X ו-Y בלתי־תלויים,

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t) = e^{\lambda(e^t-1)} \cdot e^{\mu(e^t-1)} = e^{(\lambda+\mu)(e^t-1)}.$$

אבל זו הפונקציה יוצרת המומנטים של $\mathrm{Poi}(\lambda+\mu)$, והפונקציה יוצרת המומנטים קובעת את ההתפלגות, לכן $X+Y \sim \mathrm{Poi}(\lambda+\mu)$.

תרגיל 12.18. נניח כי $X \sim \mathrm{Exp}(1)$. חשבו את $\mathbb{E}[X^{17}]$.

פתרון. אפשר לנסות לחשב ישירות, זה ידרוש הרבה אינטגרציה בחלקים והרבה סבלנות. או לחשב את הפונקציה יוצרת המומנטים של $Y \sim \mathrm{Exp}(\lambda)$:

$$M_Y(t) = \mathbb{E}[e^{tY}] = \int_0^{\infty} e^{ty} \cdot \lambda e^{-\lambda y}\, \mathrm{d}y = \lambda \int_0^{\infty} e^{(t-\lambda)y}\, \mathrm{d}y = \lambda \cdot \left(\frac{e^{(t-\lambda)y}}{t-\lambda} \right)_{y=0}^{y=\infty} = \frac{\lambda}{\lambda - t}$$

אבל הפעם זה מוגדר רק כאשר $t < \lambda$ (אחרת האינטגרל לא מתכנס). במקרה שלנו $\lambda = 1$, לכן

$$M_X(t) = \frac{1}{1-t} = \sum_{n=0}^{\infty} t^n.$$

מפיתוח טיילור של $M_X(t)$, נקבל $\frac{\mathbb{E}[X^{17}]}{17!} = 1$, לכן $\mathbb{E}[X^{17}] = 17!$.

13 תרגול שלושה עשר

13.1 אי-שוויוני צ'רנוף והופדינג

משפט 13.1 (אי-שוויון צ'רנוף)**.** יהי X משתנה מקרי, ויהי $t>0$ שעבורו $M_X(t)$ מוגדרת וסופית. אז לכל $a>0$, $P\left(X\geq a\right)\leq M_X\left(t\right)e^{-ta}$.

דוגמה 13.2. יהי $X\sim\mathrm{Poi}(\lambda)$. ניתן חסם טוב ל-$P\left(X\geq a\right)$. ראינו בתרגיל 12.17 כי הפונקציה יוצרת המומנטים של X היא $M_X(t)=e^{\lambda(e^t-1)}$, ובפרט מוגדרת לכל $t\in\mathbb{R}$. לכן, לכל $t>0$,

$$.P\left(X\geq a\right)\leq e^{\lambda(e^t-1)}e^{-ta}=e^{\lambda e^t-ta-\lambda}$$

כיוון שזה נכון לכל $t>0$, נרצה למצוא את ה-t שעבורו הביטוי באגף ימין מינימלי. נסמן את המעריך $f(t)=\lambda e^t-ta-\lambda$. נחפש מינימום: הנגזרת היא $f'(t)=\lambda e^t-a$, לכן יש קיצון כאשר $t=\log\frac{a}{\lambda}$. קל לוודא שזוהי נקודת מינימום; לכן החסם שנקבל הוא

$$P\left(X\geq a\right)\leq e^{\lambda e^{\log\frac{a}{\lambda}}-a\log\frac{a}{\lambda}-\lambda}=e^{a-\lambda-a\log a-a\log\lambda}=e^{-\lambda}\cdot\left(\frac{e}{\lambda a}\right)^a$$

(תרגיל – השוו לחסם המתקבל מאי-שוויון מרקוב ומאי-שוויון צ'בישב.)

משפט 13.3 (אי-שוויון הופדינג)**.** יהיו $X_1,\ldots,X_n$ משתנים מקריים בלתי־תלויים עם $\mathbb{E}\left[X_k\right]=0$ ו-$\left|X_k\right|\leq 1$. אזי לכל $a>0$,

$$.P\left(\sum_{i=1}^{n}X_k\geq a\right)\leq\exp\left(-\frac{a^2}{2n}\right)$$

תרגיל 13.4. בבחירות מתמודדות שתי מפלגות. מתוך ייאוש, ברירת המחדל של $4,000,001$ המצביעים היא להטיל מטבע הוגן ולהצביע בהתאם אליו לאחת משתי המפלגות. $3,000$ תושבי שכונת "לב הפארק" ברעננה החליטו להתאגד ולהצביע כולם למפלגה ב'. מה הסיכוי שהם ישפיעו על תוצאות הבחירות?

פתרון. נסמן על ידי X את ההפרש בין כמות הקולות של מפלגה א' לכמות הקולות הקולות של מפלגה ב', ונסמן $m=4,000,001-3,000$. אז אפשר לכתוב $X=\sum_{i=1}^{m}X_i$ כשכל $X_i=\begin{cases}1, & \frac{1}{2}\\ -1, & \frac{1}{2}\end{cases}$. המצביעים ישפיעו אם ורק אם $0<X<3000$. נשים לב כי

$$\begin{aligned}P(0<X<3000)&=P(X<3000)-P(X\leq 0)\leq P(X<3000)-\frac{1}{2}=\\&=1-P(X\geq 3000)-\frac{1}{2}=\frac{1}{2}-P(X\geq 3000)\end{aligned}$$

לפי אי-שוויון הופדינג,

$$P(X\geq 3000)\leq e^{-\frac{3000^2}{2\cdot m}}\approx 0.3244$$

ולכן

$$.P(\text{ישפיעו})\geq\frac{1}{2}-0.3244=0.1756$$

13.2 משפט הגבול המרכזי

הגדרה 13.5. תהי $X_1, X_2, \ldots$ סדרת משתנים מקריים, ויהי X משתנה מקרי. אומרים ש-$X_n \xrightarrow[n\to\infty]{\text{בהתפלגות}} X$, אם לכל נקודת רציפות t של F_X מתקיים

$$F_{X_n}(t) = P\left(X_n \le t\right) \xrightarrow[n\to\infty]{} F_X(t) = P\left(X \le t\right).$$

תרגיל 13.6. נניח $X_n \sim U\left[1, n\right]$ בדיד, ויהיו $Y_n = \frac{1}{n} X_n$. יהי $Y \sim U\left[0,1\right]$ רציף. הראו כי $Y_n \xrightarrow[n\to\infty]{\text{בהתפלגות}} Y$.

פתרון. נזכיר כי

$$F_Y(t) = \begin{cases} 0, & t < 0 \\ t, & 0 \le t \le 1 \\ 1, & t > 1 \end{cases}.$$

עבור $t < 0$ מתקיים כי $F_{Y_n}(t) = P\left(Y_n \le t\right) = 0 \xrightarrow[n\to\infty]{} 0 = F_Y(t)$; בדומה יש התכנסות עבור $t > 1$. לכן נניח $0 \le t \le 1$, ונראה $F_{Y_n}(t) \xrightarrow[n\to\infty]{} F_Y(t)$. אכן, על פי הגדרת Y_n,

$$F_{Y_n}(t) = P\left(Y_n \le t\right) = P\left(X_n \le nt\right) = \frac{\lfloor nt \rfloor}{n} \xrightarrow[n\to\infty]{} t = F_Y(t)$$

כנדרש.

משפט 13.7 (משפט הגבול המרכזי). תהי $X_1, X_2, \ldots$ סדרת משתנים מקריים בלתי־תלויים ושווי התפלגות עם תוחלת μ ושונות σ^2. אז

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow[n\to\infty]{\text{בהתפלגות}} \mathrm{N}(0,1).$$

מפה אנחנו מקבלים דרך להעריך את הסכום או הממוצע של דגימות. נניח שנתונים לנו n משתנים מקריים בלתי־תלויים ושווי התפלגות $X_1, \ldots, X_n$, ונניח ש-n מספיק גדול (לצורך חישובים סטטיסטיים למשל, "מספיק גדול" לרוב נחשב $n \ge 25$). אז

$$\sum_{i=1}^{n} X_i \approx \mathrm{N}\left(\mu n, \sigma^2 n\right)$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i \approx \mathrm{N}\left(\mu, (\frac{\sigma}{\sqrt{n}})^2\right)$$

דוגמה 13.8. נניח שכל $X_i \sim U\left[-1, 1\right]$ רציף. נעריך את $P\left(\sum_{i=1}^{100} X_i \ge 10\right)$. נזכור כי $\mathbb{E}\left[X_i\right] = 0$ ו-$\mathrm{Var}\left(X_i\right) = \frac{(1+1)^2}{12} = \frac{1}{3}$. לפי משפט הגבול המרכזי,

$$P\left(\sum_{i=1}^{100} X_i \ge 10\right) = P\left(\frac{\sum_{i=1}^{100} X_i - 0}{\sqrt{\frac{1}{3}} \cdot \sqrt{100}} \ge \frac{10 - 0}{\sqrt{\frac{1}{3}} \cdot \sqrt{100}}\right) \approx P\left(\mathrm{N}(0,1) \ge \sqrt{3}\right) =$$
$$= 1 - P\left(\mathrm{N}(0,1) \le 1.73\right) = 1 - 0.95818 = 0.04182$$

הערה 13.9. כאשר מקרבים משתנה בדיד על ידי משתנה רציף, נהוג לבצע **תיקון רציפות**. התיקון נעשה באופן הבא: נניח שיש משתנה מקרי בדיד X, שהתומך שלו מוכל ב-$\mathbb{N}$, ומקרבים אותו על ידי משתנה רציף Y. אז נהוג לקרב את המאורע $\{X = a\}$ על ידי המאורע $\left\{a - \frac{1}{2} \le Y \le a + \frac{1}{2}\right\}$. לכן את המאורע $\{X \le a\}$ מקרבים על ידי $\left\{Y \le a + \frac{1}{2}\right\}$, ואילו את $\{X < a\}$ מקרבים על ידי $\left\{Y \le a - \frac{1}{2}\right\}$.

דוגמה 13.10. הרבה התפלגויות אפשר לקרב על ידי התפלגות נורמלית. למשל:

$$\mathrm{Bin}(n,p) \approx \mathrm{N}(np, np(1-p))$$
$$\mathrm{Poi}(\lambda) \approx \mathrm{N}(\lambda, \lambda)$$

כאשר מקרבים משתנה מקרי בדיד על ידי נורמלי, צריך להשתמש בתיקון רציפות.

תרגיל 13.11. מטילים קוביה הוגנת 720 פעמים.

א. העריכו את הסיכוי לקבל בדיוק 120 פעמים 1.

ב. העריכו את הסיכוי לקבל בין 90 ל-100 פעמים 1 (כולל).

פתרון. המשתנה שסופר את כמות ה-1-ים הוא $X \sim \mathrm{Bin}(720, \frac{1}{6})$.

א. נעריך על ידי התפלגות נורמלית: X מתפלג בקירוב $\mathrm{N}(720 \cdot \frac{1}{6}, 720 \cdot \frac{1}{6} \cdot \frac{5}{6}) = \mathrm{N}(120, 100)$. לכן

$$P(X = 120) \approx P(119.5 \le \mathrm{N}(120,100) \le 120.5) = P\left(\frac{119.5 - 120}{10} \le \mathrm{N}(0,1) \le \frac{120.5 - 120}{10}\right) =$$
$$= P(-0.05 \le \mathrm{N}(0,1) \le 0.05) = \Phi(0.05) - \Phi(-0.05) \approx 0.51994 - 0.48006 = 0.03988$$

(התשובה האמיתית: ...0.03987; התשובה על ידי קירוב פואסון: ...0.03639)

ב. נעריך שוב על ידי התפלגות נורמלית:

$$P(90 \le X \le 100) \approx P(89.5 \le \mathrm{N}(120,100) \le 100.5) = P\left(\frac{89.5 - 120}{10} \le \mathrm{N}(0,1) \le \frac{100.5 - 120}{10}\right) =$$
$$= P(-3.05 \le \mathrm{N}(0,1) \le -1.95) = \Phi(-1.95) - \Phi(-3.05) \approx 0.02559 - 0.00114 = 0.02445$$

(התשובה האמיתית: $0.02282\ldots$).

תרגיל 13.12. ב-120 חממות בצפון הארץ מגדלים עגבניות. ידוע שתוחלת יבול העגבניות מחממה אחת היא 1.2 טון, וסטיית התקן היא 0.4 טון.

א. מה ההסתברות שיבול העגבניות הכולל מכל החממות יהיה גדול מ-140 טון?

ב. ברמת הגולן יש 30 חממות. מדריך הטיולים הפופולרי באזור העריך שהיבול הממוצע מכל חממה באזור הוא לפחות 1.5 טון. בהסתמך על נתוני השאלה, מה ההסתברות שהמדריך צודק?

ג. בעקבות נזקי הקורונה הוחלט במשרד הבריאות לסגור את כל החממות שהיבול שלהן הוא פחות מטון. בהנחה שהתפלגות היבול מחממה היא נורמלית עם הנתונים הנ"ל, מה ההסתברות שייסגרו לפחות 30 חממות?

פתרון.

א. נסמן על ידי $X_1, \dots, X_{120}$ את כמות היבול מכל חממה, ונסמן $Y_1 = \sum_{i=1}^{120} X_i$.
לפי משפט הגבול המרכזי, ההתפלגות של Y_1 היא בקירוב $\mathrm{N}(1.2 \cdot 120, 0.4^2 \cdot 120) = \mathrm{N}(144, 19.2)$. לכן

$$P(Y_1 > 140) = P(\frac{Y_1 - 144}{\sqrt{19.2}} > \frac{140 - 144}{\sqrt{19.2}}) = P(Z > -0.913) =$$
$$= 1 - P(Z \le -0.913) \approx 1 - 0.1814 = 0.8186$$

ב. כעת יש לנו 30 חממות $X_1, \dots, X_{30}$, והמשתנה הוא $Y_2 = \frac{1}{30}\sum_{i=1}^{30} X_i$. לפי משפט הגבול המרכזי, ההתפלגות של Y_2 היא בקירוב $\mathrm{N}(1.2, \frac{0.4^2}{30})$. לכן

$$P(Y_2 \ge 1.5) = P(\frac{Y_2 - 1.2}{0.4/\sqrt{30}} \ge \frac{1.5 - 1.2}{0.4/\sqrt{30}}) = P(Z \ge 4.1) =$$
$$= 1 - P(Z < 4.1) \approx 0$$

ג. נחשב את הסיכוי של חממה כלשהי להיסגר:

$$P(X_i < 1) = P(\frac{X_i - 1.2}{0.4} < \frac{1 - 1.2}{0.4}) = P(Z < -0.5) = 0.3085.$$

כמות החממות שייסגרו היא משתנה בינומי $W \sim \mathrm{Bin}(120, 0.3085)$. רוצים לחשב את $P(W \ge 30)$. נקרב את W על ידי משתנה נורמלי (כדי לוודא שהקירוב טוב, כלל אצבע הוא $\mathbb{E}[W] \ge 5$, וזה אכן המצב פה). אז בקירוב

$$W \sim \mathrm{N}(120 \cdot 0.3085, 120 \cdot 0.3085 \cdot 0.6915) = \mathrm{N}(37.02, 25.6).$$

לכן

$$P(W \ge 30) \approx P(\mathrm{N}(37.02, 25.6) \ge 29.5) = P(Z \ge \frac{30 - 37.02}{\sqrt{25.6}}) = P(Z \ge -1.48) =$$
$$1 - P(Z \le -1.48) \approx 1 - 0.06944 = 0.93056$$

תרגיל 13.13. חשבו את הגבול $\lim_{n\to\infty} e^{-n}\sum_{k=1}^{n} \frac{n^k}{k!}$.

פתרון. נשים לב כי $e^{-n}\sum_{k=1}^{n} \frac{n^k}{k!} = P\left(\mathrm{Poi}(n) \le n\right)$. יהיו $X_1, X_2, \dots$ בלתי־תלויים המתפלגים $\mathrm{Poi}(1)$. אז $\sum_{k=1}^{n} X_k \sim \mathrm{Poi}(n)$, ולכן

$$\lim_{n\to\infty} e^{-n}\sum_{k=1}^{n} \frac{n^k}{k!} = \lim_{n\to\infty} P\left(\sum_{k=1}^{n} X_k \le n\right) = \lim_{n\to\infty} P\left(\sum_{k=1}^{n} X_k - n \le 0\right) =$$
$$= \lim_{n\to\infty} P\left(\frac{\sum_{k=1}^{n} X_k - n}{\sqrt{n}} \le 0\right) \overset{(\star)}{=} P\left(\mathrm{N}(0,1) \le 0\right) = \frac{1}{2}$$

כשהמעבר $(\star)$ מתבסס על משפט הגבול המרכזי.

תרגיל 13.14 (ממבחן). נניח ש-$X_i \sim U\left[0,e\right]$ משתנים מקריים רציפים, בלתי־תלויים, ובעלי התפלגות אחידה בקטע שצוין.

א. תארו את ההתפלגות של $Y_i = \log X_i$; רצוי על ידי פונקציית הצפיפות f_{Y_i}.

ב. חשבו את התוחלת ואת השונות $\mathbb{E}[Y_i]$, $\mathrm{Var}(Y_i)$.

ג. מצאו מספר a כך ש-$P(\prod_{i=1}^{100} X_i > a) = 0.95$.

פתרון.

א. אפשר להיעזר בטענה על הצפיפות, או לחשב ישירות. נשים לב שהתומך של Y_i הוא $(-\infty, 1)$. לכל t בטווח הזה,

$$F_{Y_i}(t) = P(Y_i \leq t) = P(\log X_i \leq t) = P(X_i \leq e^t) = e^{t-1}.$$

נגזור ונקבל $f_{Y_i}(t) = e^{t-1}$. אפשר לשים לב ש-$1 - Y_i \sim \mathrm{Exp}(1)$. אכן,

$$P\left(1 - Y_i \leq t\right) = P\left(Y_i \geq 1-t\right) = 1 - F_{Y_i}(1-t) = 1 - e^{-t}.$$

ב. מהתובנה האחרונה נובע ישירות $\mathbb{E}[Y_i] = 1 - \mathbb{E}[\mathrm{Exp}(1)] = 1 - 1 = 0$ ו-$\mathrm{Var}(Y_i) = \mathrm{Var}(\mathrm{Exp}(1)) = 1$.

ג. כדי למצוא את a נבחין ראשית כי $P(\prod_{i=1}^{100} X_i > a) = P(\sum_{i=1}^{100} Y_i > \log a)$. נסמן $S = \sum_{i=1}^{100} Y_i$; לפי משפט הגבול המרכזי, $\frac{S - 100\cdot 0}{\sqrt{100\cdot 1}} = \frac{S}{10} \approx \mathrm{N}(0,1)$. לכן עלינו למצוא a שעבורו $P(S > \log(a)) = P(\frac{S}{10} > \frac{\log a}{10}) = 0.95$. אם נלך לטבלת ההתפלגות הנורמלית נקבל $\frac{\log a}{10} = 1.645$, כלומר $a = e^{16.45}$.

14 תרגול ארבעה עשר – לקריאה עצמית

תרגול זה הוא בנושא סטטיסטיקה. לעומת הסתברות, שבה ניסינו לחזות תכונות של דגימה מקרית על סמך התפלגות תיאורטית, בסטטיסטיקה מטרתנו היא להיעזר במדגם נתון שהתקבל מניסוי על מנת להעריך / לאמוד את ההתפלגות שייצרה אותו.

14.1 הפרדת השערות פשוטות

בפרק זה מטרתנו היא לדעת להפריד בין שתי השערות: H_0, שנקראת **השערת האפס**, ו-H_1, שהיא ההשערה האלטרנטיבית. השערת האפס לרוב מייצגת את המצב הקיים, או את ברירת המחדל, שלפיה למשל אין הבדל בין שתי דגימות. המטרה שלנו היא לדעת האם לקבל את H_0 או לדחות אותה ביחס ל-H_1.

הגדרה 14.1. **השערה פשוטה** היא השערה הקובעת את התפלגותו של המשתנה המקרי. **השערה מורכבת** היא השערה שלפיה התפלגות המשתנה המקרי שייכת למשפחה של התפלגויות.

דוגמה 14.2. ההשערה "$X \sim \mathrm{Geo}(\frac{1}{2})$" היא השערה פשוטה; ההשערה "$X \sim \mathrm{Geo}(p)$ לאיזשהו p" היא השערה מורכבת.

הגדרה 14.3. **מבחן** הוא הגדרה של מאורע $A \subseteq \Omega$ כך שאם $\omega \in A$, מקבלים את H_0, ואחרת דוחים את H_0 (ומקבלים את H_1).

דוגמה 14.4. ברצוננו לקבוע האם מטבע נתון הוא הוגן, או שהסיכוי שלו ליפול על עץ הוא 0.6. מבחן לדוגמה: נטיל את המטבע 1000 פעמים, ואם נקבל פחות מ-550 פעמים עץ, נקבע שהוא הוגן. אחרת, נקבע שהוא מוטה.

בדוגמה הקודמת, המספר 550 שנבחר הוא שרירותי. כל בחירה שונה הייתה נותנת לנו מבחן שונה. ברור שאם נבחר את החסם להיות, למשל, 800, המבחן לא יהיה יעיל במיוחד. על כן ברצוננו לפתח דרך למדוד עד כמה מבחן הוא טוב.

הגדרה 14.5.

א. **טעות מסוג ראשון** היא דחייה של H_0 בהינתן ש-H_0 היא נכונה. את ההסתברות לטעות כזו מסמנים $\alpha = P_{H_0}(\neg H_0)$, והיא נקראת **רמת המובהקות** של המבחן.

ב. **טעות מסוג שני** היא קבלה של H_0 בהינתן ש-H_0 אינה נכונה. את ההסתברות לטעות כזו מסמנים $\beta = P_{H_1}(H_0)$. ההסתברות המשלימה, $1-\beta$, נקראת **העוצמה** של המבחן.

בסימוני הגדרה 14.3, $\alpha = P_{H_0}(A^c)$ ו-$\beta = P_{H_1}(A)$.

דוגמה 14.6. נחשב את רמת המובהקות ואת העוצמה של המבחן שהוצע בדוגמה 14.4. יהי X המשתנה המקרי הסופר את כמות העצים שנקבל. ההשערות הן: H_0 – המטבע הוגן; H_1 – המטבע מוטה. המבחן שלנו מתבסס על המאורע $\{X < 550\}$. ניעזר במשפט הגבול המרכזי על מנת לקרב את ההסתברויות:

- חישוב α: נשים לב כי $X|H_0 \sim \mathrm{Bin}(1000, 0.5)$. על פי משפט הגבול המרכזי, נקבל שבקירוב $X|H_0 \sim \mathrm{N}(500, 250)$. לכן

$$\alpha = P_{H_0}(\neg H_0) = P_{H_0}\left(X \geq 550\right) \approx P\left(\mathrm{N}(500,250) \geq 549.5\right) =$$
$$= P\left(\mathrm{N}(0,1) \geq \frac{549.5-500}{\sqrt{250}}\right) = 1 - \Phi\left(3.13\right) = 0.00087$$

- חישוב β: נשים לב כי $X|H_1 \sim \mathrm{Bin}(1000, 0.6)$. על פי משפט הגבול המרכזי, נקבל שבקירוב $X|H_1 \sim \mathrm{N}(600, 240)$. לכן

$$\beta = P_{H_1}(H_0) = P_{H_1}\left(X < 550\right) \approx P\left(\mathrm{N}(600,240) < 549.5\right) =$$
$$= P\left(\mathrm{N}(0,1) \leq \frac{549.5-600}{\sqrt{240}}\right) = \Phi\left(-3.26\right) = 0.00056$$

הגדרה 14.7. נניח שאנו מפרידים בין שתי השערות פשוטות: $H_0: P = P_0$ ו-$H_1: P = P_1$. **יחס הנראות** מוגדר על ידי

$$\lambda\left(x_1, \ldots, x_n\right) = \frac{P_1\left(x_1, \ldots, x_n\right)}{P_0\left(x_1, \ldots, x_n\right)}$$

כאשר $x_1, \ldots, x_n$ הן תוצאות הניסוי. אם מדובר על התפלגויות בדידות, נכתוב את הסיכוי לקבלת התוצאות המדויקות; אם מדובר בהתפלגויות רציפות, נשתמש בפונקציית הצפיפות (המשותפת). ככל שיחס הנראות גדול יותר, כך הסיכוי שהתוצאה התקבלה מ-H_1 גדול יותר מאשר מ-H_0.

הגדרה 14.8. מבחןניימן-פירסון הוא מבחן מהצורה $\lambda\left(X_1,\ldots,X_n\right)\leq K$ עבור K קבוע. כלומר, משווים את יחס הנראות לקבוע K; אם הוא נמוך יותר, מקבלים את H_0, ואחרת דוחים את H_0. על ידי שינוי הערך של K ניתן לקבל רמות מובהקות שונות.

משפט 14.9 (הלמה של ניימן-פירסון)**.** בהפרדה בין השערות פשוטות עם רמת מובהקות נתונה α, מבחן ניימן-פירסון הוא המבחן בעל העוצמה המקסימלית.

איך נתכנן ניסוי המבוסס על מבחן ניימן-פירסון?

- נבחר את רמת המובהקות שבה אנו רוצים לעבוד;
- נחשב את הסף K_α שעבורו מבחן ניימן-פירסון המתאים יהיה בעל רמת מובהקות α באופן תיאורטי (על סמך H_0);
- נחשב את יחס הנראות של הדגימות הנתונות, ונשווה אותו ל-K_α.

תרגיל 14.10. במפעל לייצור נורות טוענים שמשך החיים של נורות שהם מייצרים היא מעריכית עם תוחלת 200 שעות. עיתונאי חוקר שמע מהמקורות שלו שתוחלת החיים היא דווקא 140 שעות. הוא נזכר בקורס בסטטיסטיקה שלמד, והחליט לקנות 100 נורות (בלתי־תלויות זו בזו) ולמדוד את משך החיים שלהן. הוא קיבל שבממוצע הן שרדו 160 שעות.

א. כתבו מבחן ניימן-פירסון שיבדוק האם טענת החברה נכונה. האם ניתן לדחות את טענתה ברמת מובהקות 5%?

ב. מהי העוצמה של המבחן שכתבתם?

פתרון.

א. אם נסמן על ידי $X_1,\ldots,X_{100}$ את תוחלת החיים של הנורות שהעיתונאי רכש, השערת האפס שלנו היא $H_0\colon X_i\sim\mathrm{Exp}(\frac{1}{200})$, ואילו ההשערה האלטרנטיבית היא $H_1\colon X_i\sim\mathrm{Exp}(\frac{1}{120})$. יחס הנראות הוא

$$\begin{aligned}\lambda\left(x_1,\ldots,x_{100}\right)&=\frac{f_{X_1,\ldots,X_{100}|H_1}\left(x_1,\ldots,x_{100}\right)}{f_{X_1,\ldots,X_{100}|H_0}\left(x_1,\ldots,x_{100}\right)}=\frac{\prod_{i=1}^{100}f_{X_i|H_1}\left(x_i\right)}{\prod_{i=1}^{100}f_{X_i|H_0}\left(x_i\right)}=\\&=\prod_{i=1}^{100}\frac{f_{X_i|H_1}\left(x_i\right)}{f_{X_i|H_0}\left(x_i\right)}=\prod_{i=1}^{100}\frac{\frac{1}{140}e^{-\frac{x_i}{140}}}{\frac{1}{200}e^{-\frac{x_i}{200}}}=\left(\frac{10}{7}\right)^{100}\cdot\prod_{i=1}^{100}e^{\frac{x_i}{200}-\frac{x_i}{140}}=\\&=\left(\frac{10}{7}\right)^{100}e^{-\frac{3}{1400}\sum_{i=1}^{100}x_i}\end{aligned}$$

לכן מבחן ניימן-פירסון המתאים יהיה מהצורה $\left(\frac{10}{7}\right)^{100}e^{-\frac{3}{1400}\sum_{i=1}^{100}x_i}\leq C$, ששקול לכך ש-$\frac{1}{100}\sum_{i=1}^{100}x_i\geq C$ לקבוע C כלשהו.

נרצה למצוא את C המתאים שעבורו $\alpha=P_{H_0}\left(\frac{1}{100}\sum_{i=1}^{100}X_i<C\right)=0.05$. לשם כך ניעזר במשפט הגבול המרכזי: תחת השערת האפס, $\frac{1}{100}\sum_{i=1}^{100}X_i$ מתפלג בקירוב $\mathrm{N}(200,\frac{200^2}{100})=\mathrm{N}(200,400)$. לכן

$$P_{H_0}\left(\frac{1}{100}\sum_{i=1}^{100}X_i<C\right)\approx P_{H_0}\left(\mathrm{N}(200,400)<C\right)=\Phi\left(\frac{C-200}{20}\right)=0.05,$$

נקבל $\frac{C-200}{20} = -1.645$, ולכן $C = 200 - 20 \cdot 1.645 = 167.1$. לכן נוכל לדחות את טענת החברה ברמת מובהקות 5%.

ב. העוצמה של המבחן היא $1 - \beta = P_{H_1}(\neg H_0) = P_{H_1}\left(\frac{1}{100}\sum_{i=1}^{100} X_i \leq 167.1\right)$. תחת ההשערה האלטרנטיבית, $X_i \sim \mathrm{Exp}(\frac{1}{140})$, ולכן לפי משפט הגבול המרכזי $\frac{1}{100}\sum_{i=1}^{100} X_i$ מתפלג בקירוב $\mathrm{N}(140, \frac{140^2}{100}) = \mathrm{N}(140, 196)$. לכן

$$1 - \beta \approx P\left(\mathrm{N}(140, 196) \leq 167.1\right) = \Phi\left(\frac{167.1 - 140}{14}\right) = 0.9735.$$

14.2 אמידה

בחלק זה נניח שנתון מדגם שיוצר מתוך התפלגות ידועה, שהפרמטרים שלה אינם ידועים. מטרתנו תהיה להעריך את הפרמטרים על סמך המדגם.

הגדרה 14.11.

א. **סטטיסטי** הוא פונקציה של המדגם שאינה תלויה בפרמטר.

ב. **אומד** הוא סטטיסטי שערכו אמור להיות קרוב לפרמטר המבוקש.

ג. **אומדן** הוא ערך של האומד (לאחר הצבת ערכי המדגם).

דוגמה 14.12. בהינתן דגימות $X_1, \ldots, X_n \sim \mathrm{Ber}(p)$, X_1 הוא אומד ל-p. גם $\frac{1}{n}\sum_{i=1}^{n} X_i$ הוא אומד ל-p.

הגדרה 14.13. נניח ש-$X_1, \ldots, X_n \sim F_\theta$, כאשר θ הוא פרמטר הקובע את ההתפלגות. אומרים שאומד $T = t(X_1, \ldots, X_n)$ הוא **אומד חסר הטייה** של θ, אם לכל θ מתקיים $\mathbb{E}\left[T \mid X_1, \ldots, X_n \sim F_\theta\right] = \theta$.

דוגמה 14.14.

א. בכל התפלגות, X_1 הוא אומד חסר הטייה לתוחלת.

ב. גם $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ הוא אומד חסר הטייה לתוחלת.

מה לגבי השונות? בדומה לתוחלת, היינו מצפים ש-$T = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ יהיה אומד חסר הטייה לשונות של ההתפלגות. מסתבר שזה לא המצב:

תרגיל 14.15. יהיו $X_1, \ldots, X_n$ משתנים מקריים בלתי־תלויים שווי התפלגות, בעלי תוחלת μ ושונות σ^2.

א. הראו כי T שהגדרנו אינו אומד חסר הטייה לשונות, כלומר $\mathbb{E}[T] \neq \sigma^2$.

ב. כיצד תוכלו לתקן את האומד T כך שיהיה חסר הטייה?

פתרון.

א. ראשית, נשים לב כי

$$\begin{aligned}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2&=\sum_{i=1}^{n}\left((X_i-\mu)-\left(\overline{X}-\mu\right)\right)^2=\\&=\sum_{i=1}^{n}\left((X_i-\mu)^2-2\left(X_i-\mu\right)\left(\overline{X}-\mu\right)+\left(\overline{X}-\mu\right)^2\right)=\\&=\sum_{i=1}^{n}\left(X_i-\mu\right)^2-2\left(\overline{X}-\mu\right)\sum_{i=1}^{n}\left(X_i-\mu\right)+n\left(\overline{X}-\mu\right)^2=\\&=\sum_{i=1}^{n}\left(X_i-\mu\right)^2-2n\left(\overline{X}-\mu\right)^2+n\left(\overline{X}-\mu\right)^2=\\&=\sum_{i=1}^{n}\left(X_i-\mu\right)^2-n\left(\overline{X}-\mu\right)^2\end{aligned}$$

לכן

$$\begin{aligned}\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]&=\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left(X_i-\mu\right)^2\right]-\mathbb{E}\left[\left(\overline{X}-\mu\right)^2\right]=\\&=\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}\left(X_i\right)-\mathrm{Var}\left(\overline{X}\right)=\sigma^2-\frac{\sigma^2}{n}=\frac{n-1}{n}\sigma^2\end{aligned}$$

ולא σ^2.

ב. אם נחשב דווקא את $S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$ במקום T, נקבל $\mathbb{E}\left[S^2\right]=\sigma^2$.

תרגיל 14.16. יהיו $X_1,\ldots,X_n\sim U\left[0,\theta\right]$ משתנים מקריים בלתי־תלויים מהתפלגות אחידה רציפה על $[0,\theta]$.

א. הוכיחו כי $\frac{2}{n}\sum_{i=1}^{n}X_i$ הוא אומד חסר הטייה ל-θ.

ב. מצאו קבוע a שעבורו $a\cdot\max\left\{X_1,\ldots,X_n\right\}$ הוא אומד חסר הטייה ל-θ.

(זו גרסה רציפה של בעיית הטנק הגרמני; ניתן לקרוא עליה בקישור הזה).

פתרון.

א. את זה קל לחשב, שהרי

$$\mathbb{E}\left[\frac{2}{n}\sum_{i=1}^{n}X_i\right]=\frac{2}{n}\sum_{i=1}^{n}\mathbb{E}\left[X_i\right]=\frac{2}{n}\sum_{i=1}^{n}\frac{\theta}{2}=\theta.$$

ב. ראשית יש לחשב את ההתפלגות של $M=\max\left\{X_1,\ldots,X_n\right\}$. לכל $0\le m\le\theta$ מתקיים

$$P\left(M\le m\right)=P\left(X_1,\ldots,X_n\le m\right)=\prod_{i=1}^{n}P\left(X_i\le m\right)=\left(\frac{m}{\theta}\right)^n$$

כשהמעבר האחרון מסתמך על האי־תלות. לכן ל-M יש פונקציית צפיפות

$$.f_M(m) = \begin{cases} \frac{n\cdot m^{n-1}}{\theta^n}, & 0 < m < \theta \\ 0, & \text{אחרת} \end{cases}$$

על כן, התוחלת של M היא

$$\mathbb{E}\left[M\right] = \int_0^\theta m \cdot f_M(m)\,\mathrm{d}m = \int_0^\theta m \cdot \frac{n\cdot m^{n-1}}{\theta^n}\,\mathrm{d}m = \frac{n}{\theta^n}\int_0^\theta m^n\,\mathrm{d}m =$$
$$= \frac{n}{\theta^n}\cdot\left(\frac{m^{n+1}}{n+1}\right)_0^\theta = \frac{n}{\theta^n}\cdot\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\cdot\theta$$

לכן אם ניקח $a = \frac{n+1}{n}$ נקבל ש-$\frac{n+1}{n}M$ הוא אומד חסר הטייה ל-θ.