8. מבוא לתורת ההסתברות הבדידה

תורת ההסתברות שימושית למטרות רבות במתמטיקה בדידה. רבים מהשימושים הללו יפים ומפתיעים. למשל, שאלות רבות בתורת הגרפים דנות בניסיון לבנות גרפים עם תכונות רצויות מסוימות. במקרים רבים איננו יודעים לבנות במפורש גרף עם התכונות האמורות. אולם, ניתן להראות שאם נבחר גרפים באקראי בצורה נאותה, אז תהיה הסתברות חיובית להעלות בחכתנו גרף כרצוי לנו. בפרק זה נניח את הבסיס הנחוץ להבנת היסודות של תורת ההסתברות והקשר בינה לבין המתמטיקה הבדידה.

פיתוח מלא של תורת ההסתברות מחייב הכנות מעמיקות בתורת המידה, דבר שמקשה לעתים על הבנת התחום החשוב הזה. למרבה המזל, אם מגבילים את הדיון למרחבי הסתברות בדידים (כפי שיוגדרו להלן), מסולקים הסיבוכים המחייבים את פיתוח תורת המידה. מרחבי הסתברות בדידים והמושגים הכרוכים בהם, הם פשוטים להבנה ויעילים באופן יוצא מן הכלל בפתרון בעיות במתמטיקה בדידה.

8.1. מרחבי הסתברות בדידים

הנושאים שיוצגו: *מרחב הסתברות בדיד, הסתברות, מרחב המדגם, מטבע מאוזן, מטבע מוטה, הסתברות אחידה, מאורע, מאורעות זרים, חסם האיחוד, עקרון ההכלה וההדחה, המאורע המשלים.*

נפתח בהגדרה הבסיסית ביותר של מרחב הסתברות.

הגדרה 8.1.1: מרחב הסתברות בדיד הוא קבוצה סופית Ω, כך שלכל אחד מאיבריה $x\in\Omega$ מיוחס משקל אי-שלילי $\Pr(x) \geq 0$ הנקרא **ההסתברות** של x, וכך שמתקיים $\sum_{x\in\Omega}\Pr(x)=1$,

כלומר סכום כל המשקלים (ההסתברויות) הוא 1. נסמן את מרחב ההסתברות וההסתברויות הנלוות לאיבריו על ידי $(\Omega, \Pr)$. הקבוצה Ω נקראת לעתים גם **מרחב המדגם**.

נתחיל בכמה דוגמאות פשוטות המוכרות לכם ודאי.

דוגמה 8.1.2: $\Omega = \{1,2,3,4,5,6\}$ היא קבוצת כל המספרים הרשומים על הדפנות של קוביית משחק. המשקלים $Pr(1) = Pr(2) = Pr(3) = Pr(4) = Pr(5) = Pr(6) = 1/6$ הם ההסתברויות (השוות) לכך שבהטלה של הקוביה יתקבל הערך המתאים.

דוגמה 8.1.3: כאשר מטילים מטבע ייתכנו שתי תוצאות "ראש" או "זנב" ("עץ" או "פלי" בישראל המנדטורית). במקרה זה אפשר להגדיר מרחב הסתברות $\Omega = \{H,T\}$, כאשר H מתאים לתוצאה "ראש" ו- T לתוצאה "זנב". אם הסיכוי שייצא ראש שווה לסיכוי שייצא זנב אז $Pr(H) = Pr(T) = ½$. במקרה זה נאמר שהמטבע **מאוזן** או **בלתי-מוטה**. אם $Pr(H) \neq Pr(T)$ אז נאמר שהמטבע **מוטה**.

בשתי הדוגמאות האחרונות היו כל ההסתברויות שוות. ההגדרה הבאה דנה בכך.

הגדרה 8.1.4: יהי (Ω, Pr) מרחב הסתברות בדיד. אם כל ההסתברויות $Pr(x)$ זהות, כלומר $Pr(x) = \frac{1}{|\Omega|}$ לכל $x \in \Omega$, אז אומרים ש- Pr היא **התפלגות אחידה** (או **הסתברות אחידה**) על המרחב Ω.

דוגמה 8.1.5: בחישוב ממוצע הציונים בבית ספר או באוניברסיטה, מקובל לשקלל את הציון בכל מקצוע בהתאם להיקף הלימודים בו. כך למשל, מייחסים לסטודנט שנה א' במדעי המחשב את מרחב ההסתברות הבדיד:

$\Omega =$ {מתמטיקה בדידה, מבוא למדעי המחשב, מבני נתונים, חדו"א, אלגברה לינארית}

עם ההסתברויות הבאות:

Pr(מתמטיקה בדידה) = 0.2, Pr(מבוא למדעי המחשב) = 0.15

Pr(מבני נתונים) = 0.15, Pr(חדו"א) = 0.25, Pr(אלגברה לינארית) = 0.25

הסתברויות אלה משקפות כאמור את היקפי השעות בכל אחד מהמקצועות. למשל, ההסתברות Pr(מתמטיקה בדידה) = 0.2 משקפת את הנתון ש- 20% מזמן הלימוד מוקדש ללימודי המתמטיקה הבדידה. לכן, אם נבחר באקראי שעת לימוד, ההסתברות לכך שתהא זו שעה שבה לומדים מתמטיקה בדידה היא 0.2. במקרה זה ההתפלגות אינה אחידה.

דוגמה 8.1.6: בהגרלה שעורכת חברת משקאות, מסומן כל פקק בבקבוקי המשקה שלה באחת מאותיות הא"ב, כלומר $\Omega =$ {א, ב, ג, ד, ..., ת}. השכיחות של האותיות משתנה מאות לאות כדלקמן Pr(א) = 0.05, Pr(ב) = 0.01, Pr(ג) = 0.05, ..., Pr(ז) = 0.0001 וכך הלאה, כאשר סכום ההסתברויות של כל האותיות הוא כמובן 1. הפרס הגדול ניתן למי שמצליח לצרף את אותיות המילה "זרזיף" מהאותיות שעל גבי הפקקים. כמו שאפשר לראות ההסתברות לקבל את האות 'ז' היא קטנה מאוד, ולכן הסיכוי לזכות בפרס הגדול אינו גדול אלא אם קונים מספר גדול מאוד של בקבוקי משקה.

לתורת ההסתברות יש מינוחים משלה. כך למשל, ההגדרה הבאה:

הגדרה 8.1.7: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד. תת-קבוצה $W \subseteq \Omega$ נקראת **מאורע**. מאורע W הכולל בדיוק איבר אחד של Ω נקרא גם **מאורע בסיסי**. למאורע W מיוחסת ההסתברות $\Pr(W) = \sum_{x \in W} \Pr(x)$ (סכום ההסתברויות של איברי הקבוצה).

דוגמה 8.1.8: נניח כעת שיש בידינו שני מטבעות שונים, המטבע הראשון כסוף והשני זהוב. הפעם ייתכנו 4 תוצאות אפשריות, ולכן מרחב ההסתברות יהיה $\Omega = \{HH, HT, TH, TT\}$. כך למשל, HT הוא המאורע הבסיסי שהמטבע הכסוף נפל על "ראש" והמטבע הזהוב על "זנב". ההסתברות של כל אחד מהמאורעות הבסיסיים היא ¼, בהנחה שהמטבעות מאוזנים.
מהי ההסתברות שלפחות אחד המטבעות נפל על ראש? במקרה זה אנו מעוניינים לדעת מהי ההסתברות של המאורע $W = \{HH, HT, TH\}$. ואכן,

$$\Pr(W) = \Pr(HH) + \Pr(HT) + \Pr(TH) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

דוגמה 8.1.9: במשחק השש-בש מטילים כידוע זוג קוביות. נניח ששתי הקוביות שונות, האחת כחולה והשניה אדומה. מרחב המדגם יהיה קבוצה Ω שכוללת את כל התוצאות האפשריות של הטלת שתי הקוביות. לכן $\Omega = \{(i,j) \mid 1 \le i, j \le 6\}$, כאשר המאורע הבסיסי (i,j) מציין את המצב שבו הקוביה הכחולה מראה את המספר i והקוביה האדומה את המספר j. במרחב הזה יש $6 \cdot 6 = 36$ מאורעות בסיסיים, ולכל אחד מהם הסתברות של 1/36.

ברצוננו לדעת כעת מהי ההסתברות ששתי הקוביות מראות אותו המספר ("דאבל" – לחובבי השש-בש). כאן מדובר במאורע $W_1 = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$. קל לוודא שההסתברות של המאורע W היא:

$$\Pr(W_1) = \sum_{x \in W_1} \Pr(x) = 6 \cdot \frac{1}{36} = \frac{1}{6}.$$

וכעת, מהי ההסתברות שסכום הערכים שמראות הקוביות הוא 7? במקרה זה מדובר במאורע $W_2 = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$ וההסתברות לכך היא $\Pr(W_2) = 6 \cdot \frac{1}{36} = \frac{1}{6}$.

מהי ההסתברות שסכום הערכים יהיה 4? כאן המאורע הוא $W_3 = \{(1,3), (2,2), (3,1)\}$ וההסתברות היא:

$$\Pr(W_3) = 3 \cdot \frac{1}{36} = \frac{1}{12}.$$

מהי ההסתברות שהקוביה הכחולה מראה מספר גדול מן הקוביה האדומה? במקרה זה:

$$W_4 = \{(2,1), (3,1), (3,2), (4,1), (4,2), (4,3), (5,1), (5,2), (5,3), (5,4), (6,1), (6,2), (6,3), (6,4), (6,5)\}$$

וההסתברות של מאורע זה היא:

$$.\ \Pr(W_4) = 15 \cdot \frac{1}{36} = \frac{5}{12}$$

כאמור מאורעות הם קבוצות חלקיות של מרחב המדגם, אולם יהיה לנו נוח לדבר על מאורעות "בשפה הסתברותית". כך, נוהגים למשל לתאר מאורע בצורה מקוצרת על פי קבוצת האיברים המאפיינים אותו. למשל, בדוגמה האחרונה, W_1 הוא המאורע "יצא דאבל", ואילו W_2 הוא המאורע "סכום הקוביות הוא 7".

תכונות של מרחבי הסתברות

רבים ממונחי היסוד שפיתחנו בעיסוקינו בקבוצות סופיות, מתאימים (בהתאמות הנחוצות) גם לתורת ההסתברות.

הגדרה 8.1.10: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד. שני מאורעות $A,B \subseteq \Omega$ נקראים **זרים** אם $A \cap B = \varnothing$ (כלומר הקבוצות A,B זרות). מובן שכל שני מאורעות בסיסיים שונים הם זרים.

הטענה הבאה נובעת ישירות מההגדרה של מאורעות זרים. זו הגרסה ההסתברותית של עקרון הסכום שראינו בסעיף 4.1.

טענה 8.1.11: לכל שני מאורעות זרים A,B מתקיים $\Pr(A \cup B) = \Pr(A) + \Pr(B)$.

באופן כללי, אם יש לנו כמה מאורעות אפשר להוכיח את המשפט הבא:

משפט 8.1.12: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהיו $A_1, A_2, \ldots, A_n \subseteq \Omega$ מאורעות זרים זה לזה. אז $\Pr\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \Pr(A_i)$.

דוגמה 8.1.13: נחזור לדוגמה של שתי הקוביות השונות. מה ההסתברות שסכומן של זוג הקוביות הוא 7 או 4? נגדיר את המאורעות הבאים. המאורע A יהיה "הסכום יצא 4" ואילו המאורע B יהיה "הסכום יצא 7". המאורעות A,B זרים כמובן, כי לא ייתכן שסכום שתי הקוביות יהיה גם 7 וגם 4, וקל לבדוק ש- $A = \{(1,3), (2,2), (3,1)\}$ ואילו $B = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$. ניתן היה לדעת זאת מראש במקרה זה. בכל מקרה נקבל:

$$\Pr(B) = 6 \cdot \frac{1}{36} = \frac{1}{6}, \qquad \Pr(A) = 3 \cdot \frac{1}{36} = \frac{1}{12}$$

מכיוון ש- $A\cap B=\varnothing$, אז:

$$\Pr(A\cup B)=\Pr(A)+\Pr(B)=\frac{1}{12}+\frac{1}{6}=\frac{1}{4}\ .$$

הוכחת הטענה הבאה ברורה מההגדרות.

טענה 8.1.14: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד. אם $A\subseteq B$ מאורעות ב- Ω, אז $\Pr(A)\le\Pr(B)$.

צירוף פשוט של האבחנות הקודמות מוליך לתוצאה הבאה, שעל אף פשטותה הרבה, היא מאוד שימושית.

משפט 8.1.15 (חסם האיחוד The Union Bound): יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד, ויהיו $A_1,\ldots,A_n$ מאורעות בו. אז $\Pr\left(\bigcup_{i=1}^{n}A_i\right)\le\sum_{i=1}^{n}\Pr(A_i)$.

הוכחה: נסמן $\bigcup_{i=1}^{n}A_i=\{x_1,\ldots,x_k\}$, כאשר $x_1,\ldots,x_k$ הם מאורעות בסיסיים. אז לפי ההגדרה $\Pr\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{j=1}^{k}\Pr(x_j)$. מאידך, נסמן ב- n_j את מספר הקבוצות A_i המכילות את x_j, כלומר:

$$n_j=\left|\{i\mid x_j\in A_i\}\right|\ .$$

כמובן n_j הם מספרים טבעיים שכולם $1\le$ מפני ש- $x_j\in\bigcup_{i=1}^{n}A_i$, ולכן יש לפחות אינדקס אחד i כך ש- $x_j\in A_i$. ההסתברות של המאורע A_i היא:

$$\Pr(A_i)=\sum_{x_j\in A_i}\Pr(x_j)\ .$$

נסכם לפי i על פני כל הקבוצות A_i ונקבל:

$$\sum_{i=1}^{n}\Pr(A_i)=\sum_{i=1}^{n}\sum_{x_j\in A_i}\Pr(x_j)\ .$$

עתה נהפוך את סדר הסכימה. כל איבר x_j ייספר פעם אחת כנגד כל מאורע A_i שמכיל אותו, ובסה"כ n_j פעמים. מכיוון ש- $n_j\ge 1$ לכל j נקבל:

$$\sum_{i=1}^{n}\Pr(A_i)=\sum_{i=1}^{n}\sum_{x_j\in A_i}\Pr(x_j)=\sum_{j=1}^{k}\Pr(x_j)\cdot n_j\ge\sum_{j=1}^{k}\Pr(x_j)=\Pr\left(\bigcup_{i=1}^{n}A_i\right)\ . \quad \square$$

דוגמה 8.1.16: שוב מטילים שתי קוביות שונות. מהי ההסתברות לכך שסכום הקוביות זוגי או גדול מ- 10? יהי A המאורע "סכום הקוביות זוגי" ו- B המאורע "סכום הקוביות גדול מ- 10". לכן:

$$A = \{(1,1), (1,3), (1,5), (2,2), (2,4), (2,6), (3,1), (3,3), (3,5), (4,2), (4,4), (4,6), (5,1), (5,3), (5,5), (6,2), (6,4), (6,6)\}$$

ואילו:

$$B = \{(5,6), (6,5), (6,6)\}$$

ולכן:

$$\Pr(B) = 3 \cdot \frac{1}{36} = \frac{1}{12}, \qquad \Pr(A) = 18 \cdot \frac{1}{36} = \frac{1}{2}$$

מכאן על פי חסם האיחוד:

$$\Pr(A \cup B) \le \Pr(A) + \Pr(B) = \frac{1}{2} + \frac{1}{12} = \frac{7}{12}$$

אפשר היה כמובן במקרה זה לחשב את ההסתברות במדויק שכן:

$$A \cup B = \{(1,1), (1,3), (1,5), (2,2), (2,4), (2,6), (3,1), (3,3), (3,5), (4,2), (4,4), (4,6), (5,1), (5,3), (5,5), (5,6), (6,2), (6,4), (6,5), (6,6)\}$$

ולכן:

$$\Pr(A \cup B) = \frac{20}{36} = \frac{5}{9}$$

ואכן $\frac{5}{9} < \frac{7}{12}$.

במקרים מסוימים קל לחשב בדיוק את $\Pr(A\cup B)$, כפי שניתן לראות בטענה הבאה (הטענה והוכחתה דומים לטענה 4.1.11).

טענה 8.1.17: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהיו A,B שני מאורעות בו. אז $\Pr(A\cup B) = \Pr(A) + \Pr(B) - \Pr(A\cap B)$.

הוכחה: נשים לב שהמאורעות $(A\setminus B)$, $(B\setminus A)$, $(A\cap B)$ זרים זה לזה, ואיחודם הוא המאורע $(A\cup B)$. כמו-כן $A = (A\setminus B)\cup(A\cap B)$ ובדומה $B = (B\setminus A)\cup(A\cap B)$. לכן, על ידי שימוש במשפט 8.1.12:

$$\Pr(A) = \Pr(A \setminus B) + \Pr(A \cap B),$$
$$\Pr(B) = \Pr(B \setminus A) + \Pr(A \cap B),$$
$$\Pr(A \cup B) = \Pr(A \setminus B) + \Pr(B \setminus A) + \Pr(A \cap B).$$

הטענה נובעת. □

דוגמה 8.1.18: עתה אפשר לחשב בדיוק את $\Pr(A\cup B)$ בדוגמה 8.1.16, כאשר A המאורע "סכום הקוביות זוגי" ו- B המאורע "סכום הקוביות גדול מ- 10". כל מה שעלינו לחשב הוא את $\Pr(A\cap B)$. אבל המאורע $A\cap B$ הוא "סכום הקוביות זוגי וגדול מ- 10", כלומר $A\cap B$ הוא

המאורע "סכום הקוביות הוא 12". לכן, $A\cap B = \{(6,6)\}$ ומכאן $\Pr(A \cap B) = \frac{1}{36}$. על פי טענה 8.1.17, נקבל:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = \frac{1}{2} + \frac{1}{12} - \frac{1}{36} = \frac{5}{9}$$

כפי שראינו בדוגמה 8.1.16.

הנוסחה המחשבת את $\Pr(A\cup B)$ אינה אלא מקרה פרטי של עקרון ההכלה וההדחה (סעיף 4.6).

משפט 8.1.19 (עקרון ההכלה וההדחה למאורעות): יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהיו $A_1,A_2,\ldots,A_n \subseteq \Omega$ מאורעות. ההסתברות של מאורע האיחוד היא:

$$\Pr\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \Pr(A_i) - \sum_{1\le i<j\le n} \Pr(A_i \cap A_j) + \ldots + (-1)^{n-1} \Pr(A_1 \cap A_2 \cap \ldots \cap A_n)$$

הוכחה: זהה להוכחה של משפט 4.6.3. □

הגדרה 8.1.20: יהי A מאורע במרחב הסתברות בדיד $(\Omega, \Pr)$. **המאורע המשלים** של A הוא המאורע $\Omega \setminus A$ שיסומן על ידי A^c או על ידי $\overline{A}$.

מכיוון שסכום ההסתברויות במרחב הסתברות הוא 1, לא קשה להוכיח את הטענה הבאה.

טענה 8.1.21: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $A\in\Omega$ מאורע. ההסתברות של המאורע המשלים היא $\Pr(A^c) = 1 - \Pr(A)$.

דוגמה 8.1.22: נשוב ונתבונן במרחב ההסתברות הכולל את כל תוצאות ההטלה של שתי קוביות שונות. יהי A המאורע "סכום שתי הקוביות גדול מ- 7". לכן, המאורע המשלים A^c הוא "סכום שתי הקוביות קטן או שווה ל- 7". מכאן:

$$A = \{(2,6), (3,5), (3,6), (4,4), (4,5), (4,6), (5,3), (5,4), (5,5), (5,6), (6,2), (6,3), (6,4), (6,5), (6,6)\}$$

ולכן:

$$\Pr(A) = \frac{15}{36} = \frac{5}{12}$$

ולכן על פי טענה 8.1.21:

$$\Pr(A^c) = 1 - \frac{5}{12} = \frac{7}{12}.$$

רבים מן הכלים שפיתחנו בעיסוקנו במניה יהיו לנו לעזר כאן, כפי שמראה הדוגמה הבאה.

דוגמה 8.1.23: בדוגמה 8.1.16 מצאנו כי ההסתברות לכך שסכום שתי קוביות הוא זוגי היא ½, וזאת על ידי מעבר שיטתי ומייגע על איברי מרחב ההסתברות המתאים. נראה זאת כעת באופן מתוחכם יותר. אם A הוא המאורע "סכום הקוביות הוא זוגי", אז A^c הוא המאורע "סכום הקוביות הוא אי-זוגי". כמובן מתקיים $Pr(A) + Pr(A^c) = 1$. אנו נראה כי $Pr(A) = Pr(A^c)$ ולכן $Pr(A) = Pr(A^c) = ½$. נראה זאת על ידי כך שנמצא פונקציה f חח"ע ועל בין איברי A לאיברי A^c. מכיוון שלכל איבר במרחב ההסתברות Ω אותה ההסתברות $1/36$, התוצאה תנבע. ואכן אם (i,j) הוא אחד מאיברי המאורע A, כלומר $1 \le i,j \le 6$ שני מספרים טבעיים כך ש- $i + j$ זוגי, אז $f((i, j)) = (7-i, j)$. קל לראות שאם $1 \le i \le 6$ אז גם $1 \le 7 - i \le 6$ מספר טבעי. כמו-כן, אם $i + j$ זוגי אז $7 - i + j$ מספר אי-זוגי (הוכיחו! בדקו מהי השארית מודולו 2). אמתו כעת שזו אכן פונקציה חח"ע ועל.

תרגילים

1. במשחק הלוטו יש לנחש 6 מספרים מתוך 45. כדי לזכות חייב המהמר לנחש מספרים זהים למספרים שנבחרים באקראי בהגרלת הלוטו.
   א. תארו במדויק את מרחב ההסתברות המתאים.
   ב. מהי הסתברות הזכייה בלוטו?
   ג. מהי ההסתברות לנחש נכונה לפחות 5 מתוך 6 המספרים?
   ד. מהמר בלוטו מקשיב לתוצאות ההגרלה. לשמחתו, שלושת המספרים הראשונים שעלו בגורל מתאימים לניחוש שלו. אולם בשלב זה נפסק השידור. מה ההסתברות שהמהמר ניחש נכונה את כל 6 המספרים?

2. מטילים שלוש קוביות, כחולה, אדומה ולבנה, כאשר לכל קוביה הסתברות שווה ליפול על כל אחד מ- 6 המספרים. מה ההסתברות שסכום הקוביות הוא 11? מה ההסתברות שסכומן 12?

3. בכד יש שבעה כדורים, שלושה מתוכם לבנים וארבעה אדומים. מוציאים מהכד באופן מקרי שלושה כדורים.
   א. תארו במדויק את מרחב ההסתברות המתאים.
   ב. מה ההסתברות שיצאו שני כדורים אדומים ואחד לבן?

4. מחלקים חפיסה של 52 קלפים לשני חלקים באופן מקרי.
   א. מהו מרחב ההסתברות המתאים?
   ב. מה ההסתברות שיש אותו מספר של קלפים שחורים בשני החלקים (בחבילת קלפים יש 26 קלפים שחורים ו- 26 אדומים)?

5. במשפחה יש 10 ילדים. נניח שההסתברות שנולד בן שווה להסתברות שנולדה בת.
   א. מה ההסתברות שיש במשפחה 5 בנים ו- 5 בנות?
   ב. מה ההסתברות שמספר הבנים הוא בין 3 ל- 8?

6. נתונה קוביה מוטה שבה ההסתברות לכך שהקוביה תראה מספר מסוים היא יחסית למספר הזה. כך למשל, ההסתברות שהקוביה תראה 6 גבוהה פי 3 מאשר ההסתברות שהקוביה תראה 2.
   א. מה ההסתברות של כל אחד מן המספרים?
   ב. מה ההסתברות שהקוביה תראה מספר זוגי?

7. יהיו A,B שני מאורעות כך ש- $\Pr[A\cap B] = 1/4$, $\Pr[A^c] = 1/3$, $\Pr[B] = 1/2$. למה שווה $\Pr[A\cup B]$? ומהי $\Pr[A^c\cap B^c]$?

8. א. מטילים 3 מטבעות: של 10 אגורות, של חצי שקל ושל שקל, ומניחים שכל המטבעות מאוזנים. מה ההסתברות לכך שמספר המופעים של "ראש" הוא זוגי?
ב. עתה מטילים n מטבעות מאוזנים שונים זה מזה. מה ההסתברות לקבל מספר זוגי של "ראש"?

8.2. אי-תלות והסתברות מותנה

הנושאים שיוצגו: *מאורעות בלתי-תלויים, הסתברות מותנה, צמצום של מרחב המדגם.*

הקוראים יכולים לתמוה בשלב זה מה יתרון יש בהכנסת מושגי ההסתברות. לכאורה כל מה שעשינו כאן אינו אלא גרסה משוקללת של עובדות ומושגים שכבר ראינו בתחום המניה הקומבינטורית בפרק 4. המושג הראשון שנציג המושתת על אינטואיציה הסתברותית ממש, הוא מושג האי-תלות.

אי-תלות של מאורעות

הגדרה 8.2.1: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד. שני מאורעות $A,B \subseteq \Omega$ נקראים **בלתי תלויים** אם $\Pr(A\cap B) = \Pr(A)\cdot\Pr(B)$.

דוגמה 8.2.2: נטיל שוב שתי קוביות, אדומה וכחולה. נסמן ב- (i,j) את המאורע הבסיסי שבו i הוא ערכה של הקוביה הכחולה ו- j ערכה של הקוביה האדומה. יהי A המאורע "הקוביה הכחולה מראה מספר זוגי", ויהי B המאורע "הקוביה האדומה מראה מספר גדול מ- 4". אנו טוענים כי אלה מאורעות בלתי תלויים. נבדוק זאת תחילה לאור ההגדרה. ואכן,

$$A = \{(2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$$

ולכן, $\Pr(A) = \frac{18}{36} = \frac{1}{2}$.

בדומה אפשר לבדוק ולמצוא כי $\Pr(B) = \frac{1}{3}$ (בדקו). ואילו:

$$A\cap B = \{(2,5), (2,6), (4,5), (4,6), (6,5), (6,6)\}$$

ואכן:

$$\Pr(A\cap B) = \frac{6}{36} = \frac{1}{6} = \frac{1}{2}\cdot\frac{1}{3} = \Pr(A)\cdot\Pr(B)$$

כאן כדאי לעבור ולדון במשמעות האינטואיטיבית של אי-תלות, ומדוע צפוי היה שהמאורעות A,B יהיו בלתי-תלויים. במקרה זה העניין פשוט: המאורע A מוגדר רק על פי התוצאה שמראה

הקוביה הכחולה, בעוד ש- B מוגדר רק על פי התוצאה של הקוביה האדומה. לא קשה לשער ש"אין תלות" (סיבתית, סטטיסטית) בין המאורעות האלה. אנו רואים שאכן ההגדרה של אי-תלות הולמת את האינטואיציה במקרה זה.
ראוי להזהיר ולומר שיש מקרים שבהם מאורעות A,B הם בלתי-תלויים אף כי קשה למצוא הסבר אינטואיטיבי לדבר. ולהיפך, יש מצבים שבהם נראה אולי כאילו A,B מסוימים הם בלתי תלויים, אף כי הדבר איננו כך. אחת המטרות שלנו בפרק זה היא לפתח אצל הקוראים אינטואיציה הסתברותית טובה, העוזרת בין היתר בזיהוי נכון של תלות של מאורעות ואי-תלות של מאורעות.

דוגמה 8.2.3: נחזור לדוגמה של שתי הקוביות, ונביט הפעם במאורעות הבאים. המאורע A הוא "תוצאת הקוביה הכחולה היא 4", ואילו המאורע B הוא "$i - j \geq 3$", כאשר i הוא תוצאת הקוביה הכחולה ו- j תוצאת הקוביה האדומה. לכן,

$$A = \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$$
$$B = \{(6,1), (6,2), (6,3), (5,1), (5,2), (4,1)\}$$

ומכאן:

$$\Pr(B) = \frac{6}{36} = \frac{1}{6} \qquad \Pr(A) = \frac{6}{36} = \frac{1}{6}$$

מצד שני, $A \cap B = \{(4,1)\}$ ולכן:

$$\Pr(A \cap B) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \Pr(A) \cdot \Pr(B)$$

ולכן המאורעות A,B בלתי-תלויים. די קשה לספק במקרה זה הסבר אינטואיטיבי לאי-תלות בין המאורעות A,B. על מנת להמחיש את הנקודה, נתבונן במאורע A' שהוא "הקוביה הכחולה מראה 6". במקרה זה:

$$A' \cap B = \{(6,1), (6,2), (6,3)\}$$

ולכן:

$$\Pr(A' \cap B) = \frac{3}{36} = \frac{1}{12} \neq \frac{1}{6} \cdot \frac{1}{6} = \Pr(A') \cdot \Pr(B).$$

אפשר לומר שבמקרה הראשון המאורעות A,B בלתי-תלויים "במקרה". זאת אומרת הם בלתי-תלויים אף כי אין בידינו הסבר אינטואיטיבי פשוט לכך.

הסתברות מותנה

נעבור עתה לדון במושג של הסתברות מותנה. נפתח תחילה בדוגמה שתבהיר את האינטואיציה של מושג זה.

דוגמה 8.2.4: משפחות אדמוני ולבני גרות בבניין דו-משפחתי ("דופלקס") מפואר. בבניין חיות בסה"כ 20 נפשות, כאשר מפקד האוכלוסין המדויק הוא כדלקמן.

	משפחת אדמוני	**משפחת לבני**
נשים	3	2
גברים	7	3
כלבים	5	0
סה"כ	15	5

השכנים הרכלנים שממול (משפחת סקרני) מנהלים מעקב אחרי בואם וצאתם של דיירי הבניין. נגדיר את מרחב המדגם Ω כאוסף כל המאורעות הבסיסיים "x יצא מהבניין", ונייחס למרחב הזה הסתברות אחידה. כך למשל , אם A_1 הוא המאורע "מר יעקב אדמוני יוצא כרגע מהבניין", ו- A_2 הוא המאורע "יוצא מהבניין איזשהו כלב", אז ההסתברות למאורעות אלה מחושבת כפי שראינו עד כה והיא:

$$\Pr(A_1)=\frac{1}{20} \qquad \Pr(A_2)=\frac{5}{20}=\frac{1}{4}.$$

נסמן כעת ב- X את המאורע "יוצאת מהבניין דמות ששיערה אדום" וב- Y את המאורע "יוצאת מהבניין אישה". ההסתברויות הרגילות של מאורעות אלה במרחב המדגם Ω הן:

$$\Pr(Y)=\frac{2+3}{20}=\frac{1}{4} \qquad \Pr(X)=\frac{15}{20}=\frac{3}{4}$$

אנו מניחים כאן ולהלן שלכל בני משפחת אדמוני שיער אדום (כולל הכלבים...), ולבני משפחת לבני שיער לבן. נניח שהשכנים הצופים מבחינים כבר שעומדת לצאת מהבניין דמות ששיערה אדום. מהי ההסתברות שהדמות היא אישה **בהינתן** המידע הזה ששיער הדמות אדום? ברצוננו לדעת מהי ההסתברות של המאורע Y בהינתן שהמאורע X מתקיים. לא קשה לראות שהתשובה היא $\frac{3}{15}=\frac{1}{5}$, שכן משפחת אדמוני כוללת בסה"כ 15 נפשות ומתוכן יש 3 נשים. נוהגים לסמן זאת כך:

$$\Pr(Y\mid X)=\frac{1}{5}$$

ולומר "ההסתברות של המאורע Y **בהינתן** (או **מותנה** בקיום) המאורע X היא $\frac{1}{5}$".

ברוח זו נתבונן גם במקרה הבא. יהי Z המאורע "יוצאת מהבניין אישה ממשפחת לבני". הסתברותו של המאורע Z היא כמובן

$$\Pr(Z)=\frac{2}{20}=\frac{1}{10}.$$

נניח שמשפחת סקרני מצליחה להבחין כי הדמות היוצאת מהבניין היא אישה, כלומר המאורע Y שהגדרנו קודם מתקיים. מהי כעת ההסתברות שהאישה היוצאת מהבניין היא ממשפחת לבני,

או במילים אחרות מהי ההסתברות של המאורע Z בהינתן שהמאורע Y מתקיים? לא קשה לראות שהתשובה היא $\frac{2}{5}$, וזאת מכיוון שיש בסה"כ 5 נשים בבניין ו- 2 מתוכן שייכות למשפחת לבני. הראינו אם כן כי:

$$\Pr(Z|Y)=\frac{2}{5}\,.$$

מה המשותף לדוגמאות שראינו זה עתה? בהיעדר מידע מקדים כגון: "יוצא מישהו ששיערו אדום", "יוצאת אישה" וכו', החישוב רגיל. אולם כיצד נטפל בחישובים בהינתן מידע נוסף כזה? בשפה ההסתברותית מדובר **בחישוב ההסתברות המותנה**. למשל, כאשר ידוע לנו שיוצא מישהו ממשפחת אדמוני, איך חישבנו את ההסתברות לכך שיוצאת אישה בהינתן התנאי הזה? למעשה במקרה זה התעלמנו לחלוטין ממשפחת לבני, **וצמצמנו** את תחום העניין שלנו למרחב הסתברות עם הסתברות אחידה הכולל בדיוק את חברי משפחת אדמוני. כלומר הצטמצמנו למרחב המדגם הבא הכולל 15 דיירים:

	משפחת אדמוני
נשים	3
גברים	7
כלבים	5
סה"כ	15

ובמרחב הזה ההסתברות שתצא אישה היא כמובן $\frac{3}{15}=\frac{1}{5}$.

או בדוגמה השניה שבה חישבנו את ההסתברות שיוצאת מישהי ממשפחת לבני, בהינתן התנאי שיוצאת מהבניין אישה, הצטמצמנו למרחב המדגם הבא:

	משפחת אדמוני	משפחת לבני
נשים	3	2

מהו התיאור הכללי של התהליך שביצענו כאן, ומהו הקשר בין ההסתברויות המקוריות להסתברויות המותנות? נשוב ונסמן ב- X את המאורע "יוצאת כרגע מהבניין דמות ששיערה אדום", וב- Y את המאורע "יוצאת מהבניין אישה", ונבדוק שנית מהי ההסתברות של המאורע Y מותנה בקיום המאורע X. כזכור, במרחב המדגם Ω מיוחסת לכל אחת מ- 5 = 2+3 הנשים הסתברות $\frac{1}{20}$, ובסה"כ לקבוצת הנשים במרחב הדיירים Ω יש הסתברות $\Pr(Y)=\frac{5}{20}=\frac{1}{4}$.

לעומת זאת, בבעיה המותנית, אנו מייחסים לכל אישה הסתברות $\frac{1}{15}$. מהו הקשר? המאורע $X\cap Y$ הוא "יוצאת מהבניין אישה ששיערה אדום (כלומר, ממשפחת אדמוני)", והסתברותו

במרחב Ω היא $\Pr(X \cap Y) = \frac{3}{20}$. לעומת זאת אם אומרים לנו שיוצאת מהבניין דמות ששיערה אדום, כלומר שהמאורע X מתקיים, אז אנו דנים במרחב המדגם המצומצם X. ראינו שבמקרה זה ההסתברות שהדמות היוצאת היא אישה היא:

$$\frac{\Pr(X \cap Y)}{\Pr(X)} = \frac{\frac{3}{20}}{\frac{15}{20}} = \frac{3}{15} = \frac{1}{5}.$$

באופן כללי יותר, יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד, ונניח ש- $X \subseteq \Omega$ מאורע בעל הסתברות חיובית. האם ניתן לראות גם את X כמרחב הסתברות? כמעט: X היא קבוצה סופית שלכל איבריה $x \in X$ מיוחס מספר אי-שלילי $\Pr(x)$. הקושי היחיד הוא שכרגיל $\sum_{x \in X} \Pr(x) < 1$, בעוד שבמרחב הסתברות סכום ההסתברויות צריך להיות 1. אולם אין קושי לתקן זאת אם נגדיר לכל $x \in X$ הסתברות חדשה:

$$Q(x) = \frac{\Pr(x)}{\sum_{y \in X} \Pr(y)} = \frac{\Pr(x)}{\Pr(X)}.$$

קל לראות כי הזוג (X,Q) הוא מרחב הסתברות בדיד מכיוון ש- $\sum_{x \in X} Q(x) = 1$. מרחב זה נקרא **הצמצום** של מרחב ההסתברות $(\Omega, \Pr)$ ל- X.

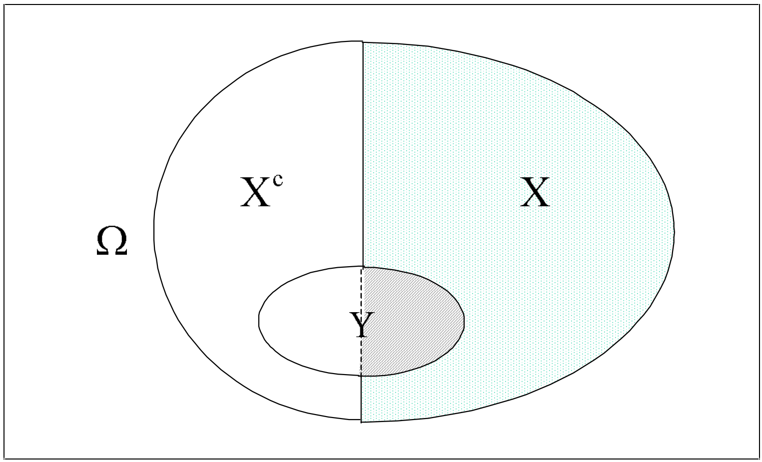
כעת אנו יכולים לשאול מהי ההסתברות של מאורע $Y \subseteq \Omega$ בהינתן המידע שהתרחש X, או במילים אחרות: מהי ההסתברות של Y מותנה בקיום המאורע X?

הגדרה 8.2.5: יהיו X,Y שני מאורעות במרחב הסתברות בדיד $(\Omega, \Pr)$, כאשר למאורע X יש הסתברות חיובית. **ההסתברות המותנה** של המאורע Y בהינתן המאורע X מסומנת על ידי $\Pr(Y|X)$ ומוגדרת על ידי:

$$\Pr(Y \mid X) = \frac{\Pr(X \cap Y)}{\Pr(X)}.$$

הערה: במקרה ש- $\Pr(X) = 0$ ההסתברות המותנה הנ"ל איננה מוגדרת כלל.

תרשים 8.2.1 מבהיר את ההגדרה. אנו מעוניינים בהסתברות של האזור הכהה. אולם אנו מעוניינים להעריך זאת לא במונחי מרחב ההסתברות $(\Omega, \Pr)$, אלא רק במונחים של חלקו הימני המנוקד X.



תרשים 8.2.1: האזור הכהה הוא $X\cap Y$, והסתברותו בהינתן ש- X מתקיים היא $Pr(Y|X)$.

דוגמה 8.2.6: נחזור לדוגמה של הטלת שתי הקוביות, הכחולה והאדומה, כאשר מרחב ההסתברות הוא $\Omega = \{(i,j) \mid 1 \le i, j \le 6\}$. כזכור, אם Y הוא המאורע "סכום הקוביות הוא זוגי" אז $Pr(Y) = ½$. נניח שנודע לנו בנפרד כי תוצאת הקוביה הכחולה גדולה מתוצאת הקוביה האדומה. כיצד משפיע המידע הנוסף הזה על ההסתברות שהסכום הוא זוגי? המידע הזה אומר שבעצם אנו מתעניינים רק במאורע: $X = \{(i,j) \mid 1 \le i,j \le 6,\ \ i > j\}$, המכיל 15 איברים (בדקו). הצמצום של (Ω, Pr) ל- X הוא לפי הגדרתנו מרחב ההסתברות X עם ההסתברות האחידה $Q(x) = \frac{1}{15}$ לכל $x\in X$. מהי ההסתברות לסכום זוגי בעוברנו למרחב המצומצם (X,Q)? קבוצת האיברים ב- X שבהם הסכום זוגי היא: $Y\cap X = \{(6,2), (6,4), (5,1), (5,3), (4,2)\}$.
מאורע זה כולל 5 איברים, ולכן הסתברותו במרחב המצומצם (X,Q) היא $5 \cdot \frac{1}{15} = \frac{1}{3}$, ובסימונים של הגדרה 8.2.5:

$$Pr(Y \mid X) = \frac{Pr(Y \cap X)}{Pr(X)} = \frac{5/36}{15/36} = \frac{5}{15} = \frac{1}{3}.$$

מושג ההסתברות המותנה מאיר באור חדש גם את מושג האי-תלות. מה ניתן לומר על ההסתברות המותנה $Pr(Y|X)$ במקרה ש- X,Y מאורעות בלתי תלויים (בעלי הסתברות חיובית)?

טענה 8.2.7: יהי (Ω, Pr) מרחב הסתברות בדיד. אם X,Y מאורעות בלתי-תלויים ול- X יש הסתברות חיובית אז $Pr(Y|X) = Pr(Y)$.

הוכחה: כזכור על פי ההגדרה

$$\Pr(Y \mid X) = \frac{\Pr(Y \cap X)}{\Pr(X)}.$$

אבל היות ש- X,Y בלתי-תלויים מתקיים $\Pr(X \cap Y) = \Pr(X) \cdot \Pr(Y)$, ולכן השבר מצטמצם. □

בעצם לא קשה להבחין שהגרירה הלוגית בטענה האחרונה עובדת גם בכיוון ההפוך, ולמעשה הוכחנו את המשפט הבא.

משפט 8.2.8: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהיו $X,Y \subseteq \Omega$ מאורעות בעלי הסתברות חיובית. אז התנאים הבאים שקולים:

1. X,Y בלתי-תלויים.
2. $\Pr(Y|X) = \Pr(Y)$.
3. $\Pr(X|Y) = \Pr(X)$.

הוכחת המשפט קלה כפי שראינו וחשיבותו היא בעיקר בבירור מושג האי-תלות. שלושת התנאים השקולים שבמשפט אומרים, בשפת בני אדם, שאם X,Y בלתי-תלויים אז המידע בדבר קיומו או אי-קיומו של המאורע X, אינו משפיע על ההסתברות שהמאורע Y יתקיים. זו באמת התכונה שהיינו מצפים באופן אינטואיטיבי שמושג האי-תלות ייצג. שימו לב, גם שהיחס "המאורעות X,Y בלתי-תלויים" הוא יחס סימטרי.

חכמים היזהרו בהסתברות...

אחת הבעיות העיקריות בלימוד יסודות ההסתברות היא העובדה שגם קודם לימוד התורה יש לנו תחושה אינטואיטיבית מהי הסתברות. אינטואיציה נאיבית כזו היא לרוב מדריך גרוע להבנה אמיתית של תורת ההסתברות. במילים אחרות, עם החשיפה השיטתית הראשונה לתורת ההסתברות, מומלץ לקוראים לנטוש כל אינטואיציה קודמת, ולסמוך רק על ההגדרות והמשפטים שיוצגו כאן. אחד המציינים המובהקים למבוכה הנובעת מידע קודם ולקוי, היא ריבוי הפרדוקסים שבעזרתם אפשר להביך את מי שמסתמך על "ידע עממי" של הסתברות. נעיר שהפרדוקסים מושתתים לרוב על ערפול מכוון של מרחב ההסתברות. בכל מקרה כזה מומלץ לקורא לברר מהו בדיוק מרחב ההסתברות Ω, מהן ההסתברויות $\Pr(x)$ הנלוות, וכן מהו המאורע הנדון. דוגמה ל"פרדוקס" המושתת על כך שמרחב ההסתברות לא הוגדר היטב היא החידה הבאה.

חידה: למשה יש שני ילדים. נתון שלפחות אחד מהם הוא בן. מה ההסתברות שלמשה יש שני בנים?

הפתרון האינטואיטיבי (והמוטעה) יאמר: "המידע שניתן לנו על כך שלמשה יש בן אינו משפיע על התשובה. לכל יילוד יש סיכוי $1/2$ להיוולד בן ו- $1/2$ להיוולד בת. לכן, התשובה היא $1/2$". הטענה הזו אינה נכונה ולאמיתו של דבר התשובה היא $1/3$, כפי שנראה מיד.
הבה נגדיר את הבעיה במדויק. מרחב ההסתברות שלנו $\Omega = \{BB, BG, GB, GG\}$ כולל את כל האפשרויות השונות לשני ילדיו של משה, כאשר B מציין בן ו- G מציין בת. כך למשל, המאורע

BG מתאר את המקרה שהילד הראשון הוא בן והילד השני הוא בת. לכל אחד מ- 4 האיברים של Ω יש הסתברות אחידה של $\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$. כעת נתון לנו המידע שאחד הילדים הוא בן. יהי X המאורע "לפחות אחד הילדים הוא בן" ויהי Y המאורע "למשה שני בנים". החידה שואלת אם כן במונחים מדויקים מהי ההסתברות של המאורע $\Pr(Y|X)$. קל לראות שהמאורע X כולל את המאורעות הבסיסיים $X = \{BB, BG, GB\}$ ולכן, $\Pr(X) = 3/4$. כמו-כן $X \cap Y = \{BB\}$, ולכן $\Pr(X \cap Y) = 1/4$. לכן ההסתברות של Y מותנה בקיום X תהיה:

$$\Pr(Y \mid X) = \frac{\Pr(X \cap Y)}{\Pr(X)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

תרגילים

1. אם $\Pr[A^c] = ¼$ ו- $\Pr[B|A] = ½$, למה שווה $\Pr[A \cap B]$?

2. מטילים קוביה פעמיים. מה ההסתברות שסכום המספרים גדול מ- 10 כאשר:
   א. לפחות באחת מהזריקות יצא 6?
   ב. בדיוק באחת מהזריקות יצא 6?
   ג. בזריקה הראשונה יצא 6?

3. נתונה קוביית משחק עם 4 פיאות (פירמידה משוכללת), אשר פיאה אחת שלה צבועה באדום, אחת בשחור, אחת בכחול והפיאה הרביעית רב-גונית וכוללת את כל שלושה הצבעים. ההסתברות ליפול על כל אחד מהצדדים היא ¼.
   א. יהי A המאורע שהפיאה שעליה נחתה הקוביה כוללת את הצבע האדום. מה ההסתברות של A?
   ב. יהי B המאורע שהפיאה שעליה נחתה הקוביה כוללת את הצבע הכחול. מה ההסתברות של B?
   ג. מה ההסתברות של המאורע $A \cap B$? הגדירו במילים מהו המאורע הזה.
   ד. האם A ו- B מאורעות בלתי-תלויים?

8.3. משתנים מקריים ותוחלת

הנושאים שיוצגו: *משתנה מקרי, תוחלת, הליניאריות של התוחלת, משתנה מקרי מציין, המספר הממוצע של נקודות שבת של תמורה.*

נעבור כעת לדון בכמה מושגי יסוד בתורת ההסתברות: משתנה מקרי, תוחלת ושונות. נפתח לשם כך בכמה דוגמאות.

דוגמה 8.3.1: נחזור לדוגמה של חישוב ממוצע הציונים של סטודנט שנה א' במדעי המחשב. כזכור מרחב ההסתברות הבדיד הוא:

Ω = {מתמטיקה בדידה, מבוא למדעי המחשב, מבני נתונים, חדו"א, אלגברה ליניארית}

וההסתברויות המתאימות לאיברי המרחב הן:

Pr(מבוא למדעי המחשב) =0.15, Pr(מתמטיקה בדידה) = 0.2,

Pr(אלגברה ליניארית) = 0.25, Pr(חדו"א) = 0.25, Pr(מבני נתונים) = 0.15.

כידוע, בחישוב ממוצע ציונים, נהוג לחשב ממוצע משוקלל הלוקח בחשבון את משקלם של המקצועות השונים. נניח שתלמיד כלשהו השיג את הציונים שלהלן:

מבוא למדעי המחשב 60, מתמטיקה בדידה 85,

אלגברה ליניארית 92, חדו"א 84, מבני נתונים 80.

אם כן, הממוצע המשוקלל של ציוני התלמיד יהיה:

$85\cdot0.2 + 60\cdot0.15 + 80\cdot0.15 + 84\cdot0.25 + 92\cdot0.25 = 82$.

דוגמה 8.3.2: משחקים משחק גורל עם שתי קוביות משחק שונות, קוביה כחולה וקוביה אדומה. נניח שתוצאת "דאבל" (אותו ערך בשתי הקוביות) מזכה את השחקן ברווח של 21 ש"ח, בעוד שכל תוצאה אחרת תגרום לו הפסד של 3 ש"ח. מהו הרווח/ההפסד הצפוי לשחקן?
מרחב המדגם כאן הוא תוצאות ההטלה של שתי הקוביות, כלומר $\Omega = \{(i,j) \mid 1 \le i,j \le 6\}$. ניתן לוודא שיש 30 צירופים שבהם הקוביות שונות, בעוד שיש 6 צירופים של "דאבל" ולכל אחד מהם הסתברות 1/36. לכן הרווח הצפוי לשחקן הוא:

$$21\cdot\frac{6}{36} - 3\cdot\frac{30}{36} = 1$$

כלומר שקל אחד.

דוגמאות אלה כללו מספר מושגים שלא הוגדרו, ולכן נפנה מיד לדיון שיטתי בשתי הדוגמאות האחרונות ובמשותף להן. לכל אחד מאיברי מרחב ההסתברות הבדיד Ω הותאם מספר ממשי. בדוגמת ממוצע הקורסים מותאם לכל קורס הציון שקיבל התלמיד בקורס. בדוגמת המשחק בקוביות מותאם לכל צירוף של הקוביות הרווח או ההפסד של השחקן במקרה הנדון. נגדיר זאת במדויק.

הגדרה 8.3.3: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד. פונקציה ממשית $f:\Omega \to \mathbb{R}$ המוגדרת על איברי Ω, נקראת **משתנה מקרי ממשי** על מרחב ההסתברות Ω.

הגדרה 8.3.4: התוחלת של משתנה מקרי ממשי f מסומנת על ידי $E[f]$, והיא הממוצע המשוקלל $E[f] = \sum_{x\in\Omega} \Pr(x)\cdot f(x)$. לעתים מסמנים את $E[f]$ גם על ידי μ_f או פשוט על ידי μ.

בדוגמה של הממוצע בקורסים, הפונקציה f מתאימה למקצוע $x\in\Omega$ את הציון $f(x)$, והתוחלת $E[f]$ היא הממוצע המשוקלל של התלמיד. בדוגמה של המשחק בקוביות, מתאימה הפונקציה f לתוצאה האפשרית של הקוביות את הרווח/ההפסד של השחקן במקרה המתאים. כאן $E[f]$ הוא הרווח/ההפסד הצפוי לשחקן.

הערה: הערה טכנית שכדאי להעיר היא שמטעמים היסטוריים מקובל יותר לסמן משתנים מקריים באותיות X,Y,Z ולא בסימון f המקובל יותר לפונקציות. אנו מעדיפים את הסימון הנוכחי מכיוון שהוא מדגיש את העיקר, כלומר שמשתנה מקרי אינו אלא פונקציה.

דוגמה 8.3.5: נניח שוב שאנו מטילים שני מטבעות שונים, המטבע הראשון כסוף והשני זהוב. כל מטבע יכול ליפול על "ראש" או "זנב". לכן מרחב ההסתברות יהיה תוצאת ההטלה של שני המטבעות, כלומר:

$$\Omega = \{HH, HT, TH, TT\},$$

כאשר ההסתברות של כל אחד מהמאורעות הבסיסיים היא 1/4.

ברצוננו לדעת כמה פעמים בממוצע יוצא "ראש" בניסוי שבו מטילים שני מטבעות שונים. נתבונן במשתנה המקרי f המתאים לכל איבר ב- Ω את מספר הפעמים שיצא "ראש". כך:

$$f(HH) = 2,\ \ f(HT) = 1,\ \ f(TH) = 1,\ \ F(TT) = 0.$$

נחשב כעת את התוחלת של f:

$$E[f] = \sum_{x\in\Omega} \Pr(x)\cdot f(x) = \frac{2}{4}+\frac{1}{4}+\frac{1}{4}+\frac{0}{4} = \frac{4}{4} = 1,$$

כלומר בממוצע מטבע אחד נופל על "ראש".

דוגמה 8.3.6: מהו הערך הממוצע של סכום שתי קוביות שונות? נתבונן שוב במרחב ההסתברות $\Omega = \{(i,j) \mid 1 \le i,j \le 6\}$ המתאר את תוצאות ההטלה של שתי קוביות שונות. נגדיר את המשתנה המקרי f המתאים לכל הטלה את סכום הקוביות, כלומר: $f((i,j)) = i + j$. מכיוון שההסתברות לכל מאורע היא 1/36, אז התוחלת של f היא:

$$E[f] = \sum_{x\in\Omega} \Pr(x)\cdot f(x) = \sum_{x\in\Omega} \frac{1}{36}\cdot f(x) = \frac{1}{36}\sum_{x\in\Omega} f(x) = \frac{1}{36}\sum_{1\le i,j\le 6} (i+j).$$

נחשב את הסכום האחרון:

$$\sum_{1\le i,j\le 6} (i+j) = \sum_{i=1}^{6}\sum_{j=1}^{6} (i+j) = \sum_{i=1}^{6}\left(6i + \sum_{j=1}^{6} j\right) = \sum_{i=1}^{6}\left(6i+21\right) = \sum_{i=1}^{6} 6i + \sum_{i=1}^{6} 21 = 252$$

ולכן התוחלת של f היא $\frac{252}{36} = 7$, כלומר הסכום הממוצע של הקוביות הוא 7.

דוגמה 8.3.7: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד, ויהי $A \subseteq \Omega$ מאורע כלשהו. הפונקציה המציינת של A היא כזכור הפונקציה $f : \Omega \to \mathbb{R}$ המוגדרת על ידי

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

שימו לב ש- f היא משתנה מקרי על Ω. לא קשה לראות שהתוחלת של f היא בדיוק ההסתברות של המאורע A מפני ש:

$$E[f]=\sum_{x\in\Omega}\Pr(x)\cdot f(x)=\sum_{x\in A}\Pr(x)\cdot 1+\sum_{x\notin A}\Pr(x)\cdot 0=\Pr(A)$$

הגדרה 8.3.8: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד. משתנה מקרי ממשי $f:\Omega\to\{0,1\}$ שהטווח שלו כולל רק את הערכים 0 ו- 1 נקרא **משתנה מקרי מציין.**

למרות השם המיוחד שניתן להם כאן, משתנים מקריים אינם אלא פונקציות. וכך, בהינתן שני משתנים מקריים f,g, אפשר להגדיר משתנים מקריים חדשים כמו $f + g$, או $f\cdot g$ וכך הלאה. למשל המשתנה המקרי $h_1 = f + g$ מוגדר על ידי $h_1(x) = f(x) + g(x)$. בדומה המשתנה המקרי $h_2 = f\cdot g$ מוגדר על ידי $h_2(x) = f(x)\cdot g(x)$. את המשתנה המקרי $h_3 = f\cdot f$ נוהגים לסמן בקיצור גם על ידי f^2 (כמקובל ברישום פונקציות ממשיות). כיצד נחשב את התוחלת של משתנים מקריים חדשים אלה? מתברר שלתוחלת יש כמה תכונות פשוטות אך מועילות ביותר כפי שנראה בהמשך.

טענה 8.3.9: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד, ויהיו $f,g:\Omega\to\mathbb{R}$ שני משתנים מקריים ממשיים. אז, הפונקציה $h = f + g$ היא משתנה מקרי ממשי על Ω ומתקיים $E[h] = E[f] + E[g]$.

הוכחה: לפי הגדרת התוחלת:

$$\begin{aligned}E[h] &= \sum_{x\in\Omega}\Pr(x)\cdot h(x)=\sum_{x\in\Omega}\Pr(x)\cdot(f+g)(x)\\ &= \sum_{x\in\Omega}\Pr(x)\cdot(f(x)+g(x))\\ &= \sum_{x\in\Omega}\Pr(x)\cdot f(x)+\sum_{x\in\Omega}\Pr(x)\cdot g(x)=E[f]+E[g]\end{aligned}$$

ובזאת מסתיימת ההוכחה. □

ניתן להכליל תוצאה פשוטה זו למספר כלשהו של משתנים מקריים.

משפט 8.3.10 (הליניאריות של התוחלת): יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהיו $f_1,\ldots,f_n:\Omega\to\mathbb{R}$ משתנים מקריים ממשיים. גם הפונקציה $f=\sum_{i=1}^{n} f_i$ היא משתנה מקרי ממשי על Ω, והתוחלת של f היא $E[f]=\sum_{i=1}^{n}E[f_i]$.

דוגמה 8.3.11: בסיאטל ירדו בשנת 1947 בממוצע מדי יום 5.3 מילימטר גשם, 1.2 מילימטר טל ו- 2.0 מילימטר שלג. לכן ממוצע המשקעים היומי בסיאטל באותה השנה היה:

$5.3 + 1.2 + 2.0 = 8.5$ מילימטר.

כאן מרחב ההסתברות היומי שלנו הוא 365 ימי השנה בשנת 1947. המשתנים המקריים f_1,f_2,f_3 מתאימים לכמויות הגשם, הטל והשלג בהתאמה שירדו כל יום באותה השנה. כך למשל,

f_1(5 בינואר 1947) = 32.0

מבטא את העובדה שבתאריך הנ"ל ירדו 32.0 מילימטר גשם.
התוחלת $E[f_3] = 2.0$ מבטאת את הערך הממוצע היומי של ירידת שלגים בשנת 1947 וכדומה.
וכך,

$$E[f_1+f_2+f_3] = E[f_1] + E[f_2] + E[f_3] = 5.3 + 1.2 + 2.0 = 8.5$$

היא כפי שראינו ממוצע המשקעים היומי בשנת 1947.

דוגמה 8.3.12: המשכורת הממוצעת ברוטו בהולנד היא 8400 גילדן בחודש. אזרח הולנדי משלם מס ממוצע של 2200 גילדן בחודש. לכן המשכורת הממוצעת נטו בהולנד היא $8400 - 2200 = 6200$ גילדן בחודש.

זו כמובן דוגמה אינטואיטיבית ביותר, אך מהו בדיוק הביסוס המתמטי שלה? נשים לב לטענה הפשוטה הבאה.

טענה 8.3.13: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהיו $f,g:\Omega\to\mathbb{R}$ שני משתנים מקריים ממשיים. אז: $E[f-g] = E[f] - E[g]$.
הוכחה: ניתן להוכיח את הטענה ישירות מהגדרת התוחלת, או לחילופין להשתמש בטענה 8.3.9 על המשתנים המקריים g, $f-g$. השלימו את הפרטים. □

דוגמה 8.3.14: נחזור לדוגמה 8.3.12. מרחב ההסתברות Ω כולל את כל אזרחי הולנד בהתפלגות אחידה. המשתנה המקרי f מתאים לכל אזרח הולנדי את משכורת הברוטו שלו לחודש, והמשתנה המקרי g מתאים לכל אזרח את המס שהוא משלם לחודש. המשתנה המקרי $f-g$ הוא לכן משכורת הנטו של האזרח לחודש. אם נפעיל את הטענה האחרונה נקבל:

$$E[f-g] = E[f] - E[g] = 8400 - 2200 = 6200$$

כנדרש.

טענה 8.3.15: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ותהי $f:\Omega\to\mathbb{R}$ פונקציה קבועה, כלומר $f(x) = c$ לכל $x \in \Omega$. אז $E[f] = c$. לפעמים נסמן זאת בפשטות $E[c] = c$, כאשר $c\in\mathbb{R}$ קבוע כלשהו.

טענה 8.3.15 היא למעשה מקרה פרטי של המשפט הבא.

משפט 8.3.16: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד, יהיו $f,g:\Omega\to\mathbb{R}$ שני משתנים מקריים ממשיים, ויהיו $a,b,c\in\mathbb{R}$ קבועים ממשיים. אז:

1. הפונקציה $a\cdot f$ היא משתנה מקרי ממשי על Ω והתוחלת שלו היא: $E[a\cdot f] = a\cdot E[f]$.
2. הפונקציה $a\cdot f + b\cdot g$ היא משתנה מקרי ממשי על Ω והתוחלת שלו היא: $E[a\cdot f + b\cdot g] = a\cdot E[f] + b\cdot E[g]$.
3. הפונקציה $a\cdot f + b\cdot g + c$ היא משתנה מקרי ממשי על Ω והתוחלת שלו היא: $E[a\cdot f + b\cdot g + c] = a\cdot E[f] + b\cdot E[g] + c$.

הערה: למשפט האחרון יש כמובן הרחבה גם לצירוף ליניארי של משתנים מקריים רבים.

כפי שרואים כאן, תכונת הליניאריות של התוחלת היא פשוטה ביותר. אף על פי כן היא מועילה מאוד כפי שמדגימה הדוגמה הבאה.

דוגמה 8.3.17: כזכור, בסעיף 4.4 דנו במספר התמורות ללא נקודות שבת של איברי הקבוצה $\{1,\ldots,n\}$. כעת נדון בשאלה מהו המספר **הממוצע** של נקודות שבת של תמורה. איך ננסח את השאלה הזו במדויק? מרחב המדגם Ω כולל את כל $n!$ התמורות האפשריות של הקבוצה $\{1,\ldots,n\}$, בהתפלגות אחידה – כלומר, לכל תמורה ב- Ω יש הסתברות שווה של $\frac{1}{n!}$. נתבונן במשתנה המקרי $f:\Omega\to\mathbb{R}$ המתאים לכל תמורה את מספר נקודות השבת שלה. מטרתנו היא לחשב את התוחלת של f. הדרך לחישוב התוחלת הזאת אופיינית לפתרונן של שאלות רבות.

ננסח תחילה במילים מהו בעצם המשתנה המקרי f: אם $\pi\in\Omega$ תמורה כלשהי, אז $f(\pi)$ הוא מספר האינדקסים $1\le i\le n$ כך ש- $\pi(i)=i$ (כלומר i נקודת שבת).

כדאי לכן להגדיר n משתנים מקריים מציינים $f_1,\ldots,f_n\colon \Omega\to\{0,1\}$ כדלקמן: תהי π תמורה כלשהי. אז,

$$f_i(\pi)=\begin{cases}1, & \pi(i)=i\\ 0, & \pi(i)\neq i\end{cases}$$

כלומר, $f_i(\pi)=1$ אם המספר i הוא נקודת שבת של התמורה π, ואחרת, $f_i(\pi)=0$. האבחנה העיקרית שאנו זקוקים לה היא ש:

$$f=f_1+\ldots+f_n.$$

ואכן, $f(\pi)$ הוא בדיוק מספר האינדקסים i כך ש- $f_i(\pi)=1$, זאת אומרת מספר האינדקסים i כך ש- $\pi(i)=i$, וזהו מספר נקודות השבת של π. בגלל הליניאריות של התוחלת:

$$E[f]=\sum_{i=1}^{n}E[f_i].$$

נותר לנו לחשב את התוחלת של כל אחד מהמשתנים המקריים f_i. לפי הגדרת התוחלת:

$$E[f_i]=\sum_{\pi\in\Omega}\Pr(\pi)\cdot f_i(\pi)$$

כזכור לכל תמורה $\pi\in\Omega$ מתקיים במקרה שלנו $\Pr(\pi)=\frac{1}{n!}$. לכן,

$$E[f_i]=\sum_{\pi\in\Omega}\frac{1}{n!}\cdot f_i(\pi)=\frac{1}{n!}\sum_{\pi\in\Omega}f_i(\pi).$$

ואולם הערך של $f_i(\pi)$ הוא 0 או 1, ולכן $E[f_i]$ איננו אלא מספר התמורות שבהן i נקודת שבת כשמחלקים מספר זה ב- $n!$. נברר אם כן מהו מספר התמורות π ש- i נקודת שבת שלהן, כלומר $\pi(i)=i$. לא קשה לראות שמספרן בדיוק $(n-1)!$ (מדוע?). לכן, התוחלת של f_i היא:

$$E[f_i]=\frac{(n-1)!}{n!}=\frac{1}{n}.$$

בסה"כ קיבלנו כי:

$$E[f]=\sum_{i=1}^{n}E[f_i]=n\cdot\frac{1}{n}=1$$

כלומר, המספר הממוצע של נקודות שבת בתמורה מקרית הוא בדיוק 1.

תרגילים

1. נהפוך את האוסף של $2^5 = 32$ הסדרות באורך 5 הבנויות מ- $\{0,1\}$ למרחב הסתברות עם התפלגות אחידה. נתאים לכל סדרה כזו את מספר ה- 1 שבה. מה התוחלת של המשתנה המקרי הזה?

2. נהפוך את אוסף הסדרות באורך 5 הבנויות מאותיות הא"ב $\{1,2,3\}$ למרחב הסתברות עם התפלגות אחידה.
   א. מהי התוחלת של מספר ה- 1 בסדרה כזו?
   ב. מהי התוחלת של מספר הרצפים 11?

3. א. במשחק המזל "קטינא" מטילים קוביה בלתי מוטה. הפרס הניתן לשחקן הוא ביחס הפוך למספר שמראה הקוביה. כך למשל, אם יצא המספר 4 אז השחקן יזכה ברבע שקל, ואם יצא המספר 1 הוא יזכה בשקל אחד. מהי תוחלת הזכייה בהטלה אחת של הקוביה?
   ב. את המשחק "קטינא" אפשר לשחק גם בשתי קוביות. אם הסכום של שתי הקוביות הוא k אז השחקן זוכה ב- 1/k שקלים. מהי תוחלת סכום הזכייה במקרה זה?

4. אומרים שהתמורה π של $\{1,2,\ldots,n\}$ **מחליפה** בין i ל- j אם $\pi(i)=j$ וכן $\pi(j)=i$. למשל, אם n = 7 אז התמורה $\pi=(5,3,2,6,1,7,4)$ מחליפה בין 1 ל- 5 כי $\pi(1)=5$ ו- $\pi(5)=1$. כמו-כן יש כאן גם חילוף בין 2 ל- 3. מהי תוחלת מספר החילופים בתמורה π של $\{1,2,\ldots,n\}$?
   הדרכה: הגדירו $\binom{n}{2}$ משתנים מקריים $f_{i,j}$ על אוסף התמורות של $\{1,2,\ldots,n\}$ כך ש- $f_{i,j}(\pi)=1$ אם יש חילוף בין i ל- j, ו- $f_{i,j}(\pi)=0$ אחרת.

8.4. ההתפלגות של משתנה מקרי

הנושאים שיוצגו: *התפלגות הערכים של משתנה מקרי וחישוב התוחלת בעזרת התפלגות זו.*

נפתח בשאלה האינטואיטיבית הבאה: "מהו הסיכוי שאם נטיל זוג קוביות ייצא סכום ≤ 8"? או: "מה הסיכוי שייצא הסכום 7"? וכדומה. כבר ידוע לנו שהסתברויות מיוחסות למאורעות, בעוד שכאן מנוסחת השאלה במונחים של משתנים מקריים (המשתנה המקרי המתאים להטלת זוג קוביות את סכומן). איך מיישבים את הדברים?

למעשה כבר ראינו קשר הדוק בין מאורעות למשתנים מקריים כשמדובר במשתנים מקריים מציינים (דוגמה 8.3.7). אם $f:\Omega\to\{0,1\}$ הוא משתנה מקרי מציין אז מתאים לו המאורע

$$A_f=\{x\in\Omega\,|\,f(x)=1\}$$

הכולל את כל המאורעות הבסיסיים $x\in\Omega$ שעבורם $f(x) = 1$. מהי התוחלת של המשתנה המקרי f? נזכיר את הטענה הבאה שהוכחה קודם לכן בדוגמה 8.3.7:

טענה 8.4.1: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד, יהי $f:\Omega\to\{0,1\}$ משתנה מקרי מציין ויהי $A_f=\{x\in\Omega\,|\,f(x)=1\}$. אז: $E[f]=\Pr(A_f)$.

באופן כללי יותר כל משתנה מקרי ממשי $f:\Omega\to\mathbb{R}$ משרה מאורעות שונים במרחב המדגם Ω. כך למשל, אם a הוא מספר ממשי אז המאורע

$$T=\{x\in\Omega\,|\,f(x)=a\}$$

כולל בדיוק את איברי המרחב שעליהם מקבל המשתנה המקרי f את הערך a. את המאורע הזה מקובל לסמן באופן הבא (הסימון הזה עלול לבלבל בקלות, ולכן מומלץ להתעכב כאן):

הגדרה 8.4.2: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $f:\Omega\to\mathbb{R}$ משתנה מקרי ממשי. נסמן על ידי $\mathbf{(f = a)}$ את המאורע $\{x \in \Omega \mid f(x) = a\}$. באופן דומה המאורע $\{x \in \Omega \mid f(x) \geq a\}$ יסומן על ידי $\mathbf{(f \geq a)}$.

על השאלה האינטואיטיבית "מה הסיכוי ש- $f = a$" נוכל לענות כעת על ידי:

$$\Pr(f=a)=\sum_{\substack{x\in\Omega,\\ f(x)=a}}\Pr(x)$$

כי הרי ההסתברות של מאורע היא סכום ההסתברויות של איברי המאורע, ו- $(f = a)$ הוא מאורע.

דוגמה 8.4.3: נחזור לדוגמה 8.1.8, שבה הטלנו שני מטבעות שונים, אחד כסוף והשני זהוב. מרחב המדגם היה כזכור:

$$\Omega = \{HH, HT, TH, TT\},$$

וההסתברות של כל מאורע בסיסי היא 1/4.
יהי f המשתנה המקרי המתאים לכל איבר ב- Ω את מספר הפעמים שיצא "ראש". ברצוננו לדעת מהי ההסתברות שיצא לפחות "ראש" אחד, כלומר מהי ההסתברות ש- $f \geq 1$. המאורע $(f \geq 1)$ הוא הקבוצה $\{HH, HT, TH\}$ וההסתברות של מאורע זה היא כמובן:

$$\Pr(f\geq 1)=\Pr(HH)+\Pr(HT)+\Pr(TH)=\frac{3}{4}$$

דוגמה 8.4.4: מהי ההסתברות שבהטלת זוג קוביות שונות הסכום יצא 7? כאן כמקודם מרחב המדגם הוא $\Omega = \{(i,j) \mid 1 \leq i,j \leq 6\}$, והמשתנה המקרי $f:\Omega\to\mathbb{R}$ מוגדר על ידי $f((i,j)) = i + j$, כלומר סכום הקוביות. המאורע "הסכום הוא 7" כולל את האיברים הבאים:

$$(f=7)=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

ולכן:

$$\Pr(f=7)=\frac{6}{36}=\frac{1}{6}$$

כי במאורע $(f=7)$ יש 6 מאורעות בסיסיים ולכל אחד הסתברות של $\frac{1}{36}$.

מהי ההסתברות שסכום הקוביות ≥ 9? כאן מדובר במאורע:

$$(f\geq 9)=\{(3,6),(4,5),(4,6),(5,4),(5,5),(5,6),(6,3),(6,4),(6,5),(6,6)\}.$$

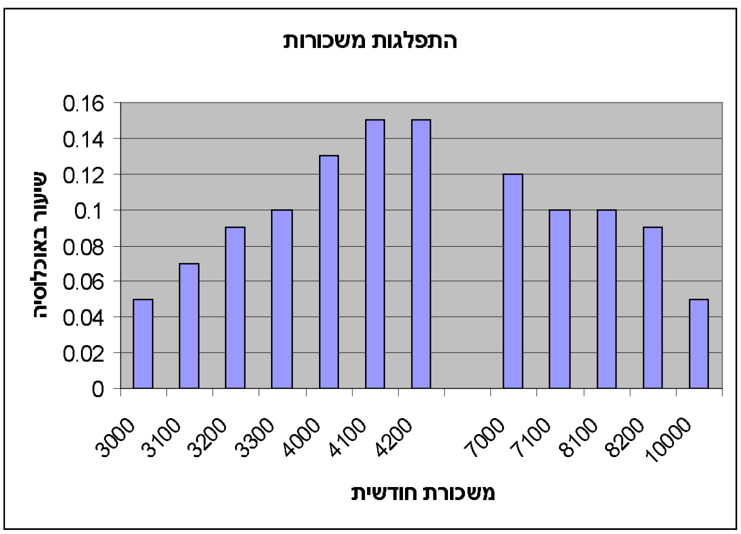
לכן,

$$\Pr(f\geq 9)=\frac{10}{36}=\frac{5}{18}.$$

מהי לעומת זאת ההסתברות שסכום הקוביות קטן מ- 9? כלומר מהי ההסתברות של המאורע $(f<9)$? כאן אין כבר צורך בחישוב מפורש. הרי המאורעות $(f\geq 9)$ ו- $(f<9)$ הם מאורעות משלימים. לכן:

$$\Pr(f<9)=1-\Pr(f\geq 9)=1-\frac{5}{18}=\frac{13}{18}.$$

בשימושים רבים של ההסתברות כאשר דנים במשתנה מקרי ממשי $f:\Omega\rightarrow\mathbb{R}$, איננו מעוניינים בידיעה מלאה של הפונקציה f, אלא רק **בהתפלגות** הערכים שלה. דהיינו, בהסתברויות השונות $\Pr(f=a)$ על פני כל ערכי a הממשיים.



תרשים 8.4.1: חלק מתוך הגרף של התפלגות המשכורות באוכלוסייה.

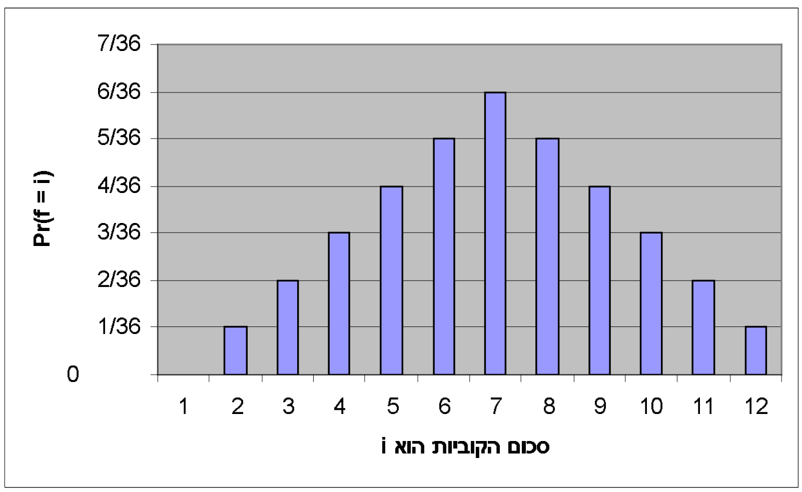
דוגמה 8.4.5: ניקח לדוגמה את מרחב המדגם Ω הכולל את כל בתי האב במדינה בהסתברות אחידה, כאשר $f(x)$ היא ההכנסה החודשית של משפחת x (מעוגלת לכפולה הקרובה של 100 ש״ח). הלשכה המרכזית לסטטיסטיקה מתעניינת למשל בשאלות מהסוג: מה שיעור המשפחות שהכנסתן החודשית היא בין 7000-8000 ש״ח. איננו מתעניינים מיהם אותם אנשים, אלא רק מהו שיעורם באוכלוסייה.

הגרף בתרשים 8.4.1, מתאר דוגמה של התפלגות ההכנסות של בתי האב במדינה. ציר ה- x מתאר את המשכורת החודשית, ואילו ציר ה- y את שיעור בתי האב באוכלוסייה. כך למשל, אפשר לראות כי $\Pr(7000 \le f < 7100) = 0.12$.

דוגמה 8.4.6: בדיוק באותו אופן ניתן לתאר גם את התפלגות הערכים של המשתנה המקרי f המוגדר על ידי ״הסכום של שתי קוביות שונות״. בתרשים 8.4.2, ציר ה- x מתאר את הסכומים האפשריים של סכומן של שתי קוביות. ואילו, ציר ה- y מתאר את ההסתברות שהמשתנה המקרי f שווה לסכום כלשהו. כך למשל, אנו רואים ש-

$$\Pr(f = 7) = \frac{6}{36} = \frac{1}{6}$$

כפי שראינו קודם (דוגמה 8.1.9). בדקו שסכום הגבהים של העמודות בגרף הוא 1. עובדה זאת מבטאת את העובדה שהמאורעות $(f = 2), (f = 3), \dots, (f = 12)$ מהווים חלוקה של מרחב ההסתברות $\Omega = \{(i,j) \mid 1 \le i,j \le 6\}$.



תרשים 8.4.2: גרף התפלגות הערכים של סכומן של שתי קוביות שונות.

אנחנו רואים אם כן את המשתנה המקרי כמוכר או ידוע למעשה, כאשר ידועה לנו במלואה ההתפלגות שלו. בשימושים רבים קשה לרוב להשיג ידיעה כה מלאה על המשתנה המקרי, ואנו מסתפקים לעתים קרובות במידע חלקי יותר. להלן נראה שמתוך ידיעת ההתפלגות של משתנה מקרי, ניתן להסיק למשל מהי התוחלת שלו.

משפט 8.4.7: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $f:\Omega \to \mathbb{R}$ משתנה מקרי ממשי. אז $E[f]=\sum_{a\in\mathbb{R}} a\cdot \Pr(f=a)$.

הערה: שימו לב, הסכום הנ״ל הוא על פני כל המספרים הממשיים $a\in\mathbb{R}$, אף כי בדיוננו מרחבי ההסתברות סופיים, ולכן יהיה רק מספר סופי של מספרים ממשיים a שבשבילם $\Pr(f = a)$ חיובי.

הוכחה: המשפט נובע ישירות מהגדרת התוחלת והמאורעות $f = a$, וזאת מכיוון ש-

$$E[f]=\sum_{x\in\Omega} f(x)\cdot \Pr(x)=\sum_{a\in\mathbb{R}}\left(a\cdot \sum_{\substack{x\in\Omega,\\ f(x)=a}} \Pr(x)\right)=\sum_{a\in\mathbb{R}} a\cdot \Pr(f=a)$$

שימו לב שהצעד העיקרי בהוכחה איננו אלא החלפה של סדר הסכימה. □

הנוסחה הזאת מבטאת באופן נוסף את האינטואיציה שתוחלת של משתנה המקרי איננה אלא ממוצע משוקלל של ערכי פונקציה. כזכור כבר ראינו מקרה פרטי של נוסחה זו עבור משתנה מקרי מציין (ראו דוגמה 8.3.7).

דוגמה 8.4.8: נחזור שוב לדוגמה של שתי הקוביות השונות, ונתבונן במשתנה המקרי f המוגדר על ידי $f((i,j)) = i + j$. טווח הערכים של המשתנה המקרי f הוא $\{2,\ldots,12\}$. לכן, התוחלת של f היא:

$$E[f]=\sum_{k=2}^{12} k\cdot \Pr(f=k)=2\cdot\frac{1}{36}+3\cdot\frac{2}{36}+4\cdot\frac{3}{36}+\ldots+12\cdot\frac{1}{36}=7$$

כפי שכבר ראינו קודם (ראו דוגמה 8.3.6).

כפי שכבר הערנו, אי-תלות היא בין המושגים המרכזיים ביותר בתורת ההסתברות. ראינו כבר מתי שני מאורעות נקראים בלתי-תלויים. המושג הזה מוכלל באופן טבעי גם להגדרה של אי-תלות בין משתנים מקריים.

סימון: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד, יהיו $f,g:\Omega\to\mathbb{R}$ שני משתנים מקריים ממשיים, ויהיו $a,b\in\mathbb{R}$ שני קבועים. המאורע ״$(f = a)$ וגם $(g = b)$״ יסומן על ידי $(\mathbf{f = a, g = b})$.

הגדרה 8.4.9: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהיו $f,g:\Omega\to\mathbb{R}$ שני משתנים מקריים ממשיים. נאמר ש- f,g **בלתי-תלויים** אם המאורעות $(f = a), (g = b)$ בלתי תלויים לכל $a,b\in\mathbb{R}$. דהיינו, לכל $a,b\in\mathbb{R}$ מתקיים:

$$\Pr(f=a,g=b)=\Pr(f=a)\cdot\Pr(g=b).$$

דוגמה 8.4.10: נתבונן במרחב המדגם $\Omega = \{(i,j) \mid 1 \le i,j \le 6\}$, המתאר את תוצאות ההטלה של שתי קוביות שונות, האחת כחולה והשניה אדומה. כזכור ההסתברות לכל מאורע בסיסי ב- Ω היא $\frac{1}{36}$. נגדיר שני משתנים מקריים f,g על מרחב המדגם Ω באופן הבא: המשתנה המקרי f יתאר את תוצאת ההטלה של הקוביה הכחולה, ואילו המשתנה המקרי g יתאר את תוצאת

ההטלה של הקוביה האדומה. כך למשל, $f((2,3)) = 2$ ואילו $g((2,3)) = 3$. קל לראות שההסתברות שהמשתנה המקרי f מקבל ערך כלשהו $1 \le i \le 6$ היא:

$$\Pr(f = i) = \frac{6}{36} = \frac{1}{6},$$

וזאת מכיוון שהמאורע $(f = i)$ כולל את האיברים הבאים $\{(i,1), (i,2), (i,3), (i,4), (i,5), (i,6)\}$.
בדומה לכל $1 \le j \le 6$:

$$\Pr(g = j) = \frac{6}{36} = \frac{1}{6}.$$

לעומת זאת, המאורע $(f = i, g = j)$ מתאר את המאורע שתוצאת הקוביה הכחולה היא i ותוצאת הקוביה האדומה היא j, כלומר את המאורע (i,j). ההסתברות של מאורע זה היא כמובן $\frac{1}{36}$.
ראינו שלכל $1 \le i,j \le 6$:

$$\Pr\left(f = i, g = j\right) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \Pr(f = i) \cdot \Pr(g = j)$$

ולכן המשתנים המקריים f,g הם בלתי-תלויים.
הדוגמה הזאת ממחישה את האינטואיציה שאכן התוצאה בקוביה הכחולה והתוצאה בקוביה האדומה אינן תלויות זו בזו (במובן של היעדר קשר סיבתי בין שתי התוצאות).

כפי שראינו קל לחשב את התוחלת כשמדובר בסכום של משתנים מקריים, אולם חישוב התוחלת של מכפלת משתנים מקריים הוא קשה. יוצא מכלל זה המצב כשהמשתנים המקריים הם בלתי תלויים.

טענה 8.4.11: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהיו $f,g : \Omega \to \mathbb{R}$ שני משתנים מקריים ממשיים בלתי-תלויים. אז $E[f \cdot g] = E[f] \cdot E[g]$.
הוכחה: לפי ההגדרה של התוחלת מתקיים

$$E\left[f \cdot g\right] = \sum_{x \in \Omega} f(x) \cdot g(x) \cdot \Pr(x).$$

עתה נפצל את Ω למאורעות על פי זוג הערכים $(f(x), g(x))$. כלומר לכל מספר a בטווח של f ומספר b בטווח של g, נביט במאורע $(f = a, g = b)$ הכולל את כל האיברים $x \in \Omega$ כך ש- $f(x) = a$ ו- $g(x) = b$. את הסכום שלפנינו נפצל בהתאם למאורעות האלה. בכל מאורע כזה, $(f = a, g = b)$ הביטוי $f(x) \cdot g(x)$ הוא קבוע ושווה ל- $a \cdot b$. לכן נקבל:

$$\sum_{x \in \Omega} f(x) \cdot g(x) \cdot \Pr(x) = \sum_{a,b \in \mathbb{R}} a \cdot b \cdot \Pr(f = a, g = b).$$

מכיוון ש- f,g משתנים מקריים בלתי תלויים אז $\Pr(f = a, g = b) = \Pr(f = a) \cdot \Pr(g = b)$. לכן:

$$\sum_{a,b \in \mathbb{R}} a \cdot b \cdot \Pr(f = a, g = b) = \sum_{a,b \in \mathbb{R}} a \cdot b \cdot \Pr(f = a) \cdot \Pr(g = b).$$

את הסכום הכפול הזה ניתן לבטא כמכפלה של שני סכומים:

$$\sum_{a,b\in\mathbb{R}} a\cdot b\cdot \Pr(f=a)\cdot\Pr(g=b)=\left(\sum_{a\in\mathbb{R}} a\cdot\Pr(f=a)\right)\cdot\left(\sum_{b\in\mathbb{R}} b\cdot\Pr(g=b)\right)=E[f]\cdot E[g]$$

השוויון האחרון נובע שוב מהגדרת התוחלת. □

דוגמה 8.4.12: בקזינו "מועדון המכפלה" לא מסכמים את ערכי הקוביות, אלא מכפילים אותם זה בזה. נשוב ונתבונן במרחב המדגם $\Omega=\{(i,j)\mid 1\le i,j\le 6\}$, המתאר את תוצאות ההטלה של שתי קוביות שונות, האחת כחולה והשניה אדומה. נתבונן במשתנה המקרי $f:\Omega\to\mathbb{R}$ המוגדר על ידי $f((i,j))=i\cdot j$. מהי התוחלת של f?

כאן, $f=f_1\cdot f_2$, כאשר f_1 הוא ערכה של הקוביה הכחולה ו- f_2 ערכה של הקוביה האדומה. פורמלית נגדיר שני משתנים מקריים $f_1,f_2:\Omega\to\mathbb{R}$ באופן הבא:

$$f_1((i,j))=i \quad , \quad f_2((i,j))=j \, .$$

כפי שראינו המשתנים המקריים f_1,f_2 הם בלתי תלויים. לכן,

$$E[f_1\cdot f_2]=E[f_1]\cdot E[f_2] \, .$$

נחשב את התוחלת של f_1.

$$E[f_1]=\sum_{i=1}^{6} i\cdot\Pr(f_1=i)=1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{1}{6}+6\cdot\frac{1}{6}=\frac{7}{2} \, .$$

מובן גם ש- $E[f_2]=\frac{7}{2}$.

לכן,

$$E[f]=E[f_1]\cdot E[f_2]=\frac{7}{2}\cdot\frac{7}{2}=\frac{49}{4}=12\frac{1}{4} \, .$$

8.5. השונות של משתנה מקרי

הנושאים שיוצגו: *השונות, סטייה מהתוחלת, תכונות של השונות.*

כפי שכבר הערנו, רק לעתים נדירות אנו יודעים ידיעה מלאה את התפלגות הערכים של משתנה מקרי. לרוב יש לנו רק מידע חלקי אודות ההתפלגות. כרגיל אנו יכולים לחשב את התוחלת אשר מבטאת מידע חשוב – הערך הממוצע המשוקלל של הפונקציה המתאימה. אולם בשימושים רבים חשוב גם לדעת עד כמה מבטא הערך הממוצע את ההתנהגות הטיפוסית של הפונקציה. כך למשל, מפורסם הסיפור על האיש שטבע באגם שעומקו הממוצע 10 ס"מ. עומק המים באגם הזה משתנה מן הסתם באופן קיצוני.

מדד חשוב אחר, **השונות**, מבטא באופן מוצלח את התנודות והשינויים בערכו של המשתנה המקרי. כיצד נמדוד זאת? יהי f משתנה מקרי שתוחלתו $\mu=E[f]$. אנחנו רוצים לדעת מהם הערכים האופייניים של $f(x)-\mu$, או במילים אחרות, מהו המרחק בין $f(x)$ לערכו הממוצע של

המשתנה המקרי (התוחלת). נתבונן בפונקציה $g = f - \mu$ המוגדרת על ידי $g(x) = f(x) - \mu$.
נשים לב כי לפי משפט 8.3.16:

$$E[g] = E[f - \mu] = E[f] - \mu = 0.$$

שימו לב, המשתנה המקרי $g = f - \mu$ מודד את **הסטייה** (המרחק) של המשתנה המקרי f מהתוחלת שלו. אולם הממוצע (התוחלת) של המשתנה המקרי g הוא 0. כלומר מיצוע רגיל של ערכי הסטיות לא ילמד אותנו דבר. הסיבה היא כמובן שהפונקציה g מקבלת הן ערכים חיוביים והן שליליים. נשים לב שאם ברצוננו לדעת מהו המרחק של המשתנה המקרי f מהתוחלת שלו, העובדה כי $g(x) = 3$ או $g(x) = -3$ היא שקולה למעשה. שני המקרים מצביעים על סטייה של המשתנה המקרי f ב- 3 מהתוחלת שלו (כלפי מעלה או כלפי מטה). מה שדרוש, הוא לכן, אמצעי שיטפל בערכים חיוביים ושליליים של g באותו אופן. מועמד טבעי לשם כך הוא המשתנה המקרי g^2. ואכן, $E\left[g^2\right]$ מוגדרת כשונות של f. נסכם זאת בהגדרה הבאה.

הגדרה 8.5.1: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $f : \Omega \to \mathbb{R}$ משתנה מקרי ממשי עם תוחלת $\mu = E[f]$. **השונות** של f מסומנת על ידי $Var[f]$ ומוגדרת על ידי

$$Var[f] = E\left[(f - \mu)^2\right].$$

הערה: השונות של משתנה מקרי f היא תמיד אי-שלילית מכיוון שהמשתנה המקרי $(f - \mu)^2$ מקבל רק ערכים אי-שליליים, ולפיכך גם תוחלתו אי-שלילית.

הערה נוספת: נבהיר שנית, $(f - \mu)^2$ הוא משתנה מקרי שערכו בכל $x \in \Omega$ הוא $(f(x) - \mu)^2$.

דוגמה 8.5.2: נתבונן במרחב המדגם $\Omega = \{H, T\}$ המתאר את תוצאות ההטלה של מטבע, עם התפלגות אחידה, כלומר $\Pr(H) = \Pr(T) = \frac{1}{2}$. נגדיר משתנה מקרי $f : \Omega \to \mathbb{R}$ על ידי $f(H) = 1$, $f(T) = 0$. קל לחשב את התוחלת של המשתנה המקרי f על פי ההגדרה של התוחלת:

$$E[f] = \Pr(H) \cdot f(H) + \Pr(T) \cdot f(T) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 = \frac{1}{2}.$$

נחשב כעת את השונות של f:

$$\begin{aligned} Var[f] &= E\left[\left(f - \frac{1}{2}\right)^2\right] \\ &= \Pr(H) \cdot \left(f(H) - \frac{1}{2}\right)^2 + \Pr(T) \cdot \left(f(T) - \frac{1}{2}\right)^2 \\ &= \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)^2 + \frac{1}{2} \cdot \left(0 - \frac{1}{2}\right)^2 = \frac{1}{4} \end{aligned}$$

דוגמה 8.5.3: נתבונן שוב במרחב המדגם $\Omega=\{H,T\}$ המתאר את תוצאות ההטלה של מטבע, אולם הפעם נניח שמדובר במטבע מוטה. כלומר ההסתברות לראש ולזנב אינה שווה. נניח כי $\Pr(H)=p$ ואילו $\Pr(T)=q$, כאשר כמובן $p+q=1$ (סכום ההסתברויות של איברי המדגם הוא כרגיל 1). נגדיר שוב משתנה מקרי $f:\Omega\to\mathbb{R}$ על ידי $f(H)=1,\ f(T)=0$. התוחלת של f במקרה זה תהיה:

$$E[f]=\Pr(H)\cdot f(H)+\Pr(T)\cdot f(T)=p\cdot 1+q\cdot 0=p\,.$$

מה לגבי השונות? לא קשה לחשב כי:

$$\begin{aligned}\mathrm{Var}[f]&=E\left[(f-p)^2\right]\\&=\Pr(H)\cdot(f(H)-p)^2+\Pr(T)\cdot(f(T)-p)^2\\&=p\cdot(1-p)^2+q\cdot(0-p)^2\end{aligned}$$

אולם $p+q=1$, ולכן אפשר לפשט את הביטוי האחרון ולקבל:

$$\begin{aligned}\mathrm{Var}[f]&=p\cdot(1-p)^2+q\cdot(0-p)^2\\&=p\cdot q^2+q\cdot p^2=p\cdot q\cdot(q+p)=p\cdot q\end{aligned}$$

במקרה ש- $p=q=½$ ראינו כבר שהשונות היא אכן $p\cdot q=¼$.

דוגמה 8.5.4: מהי השונות של הטלת קוביה אחת? נתבונן במרחב המדגם $\Omega=\{1,2,3,4,5,6\}$ המתאר את תוצאת ההטלה של קוביה אחת, עם הסתברות אחידה של $\frac{1}{6}$ לכל תוצאה. יהי $f:\Omega\to\mathbb{R}$ משתנה מקרי המוגדר על ידי $f(i)=i$, כלומר תוצאת ההטלה. התוחלת של f היא:

$$E[f]=\sum_{i\in\Omega}\Pr(i)\cdot f(i)=\frac{1}{6}\sum_{i=1}^{6}i=\frac{7}{2}\,.$$

ואילו השונות של f היא:

$$\mathrm{Var}[f]=E\left[\left(f-\frac{7}{2}\right)^2\right]=\sum_{i\in\Omega}\Pr(i)\cdot\left(f(i)-\frac{7}{2}\right)^2=\frac{1}{6}\sum_{i=1}^{6}\left(i-\frac{7}{2}\right)^2=\frac{35}{12}\,.$$

הטענה הבאה מפשטת לעתים את חישובי השונות.

טענה 8.5.5: יהי f משתנה מקרי ממשי. אז $\mathrm{Var}[f]=E\left[f^2\right]-(E[f])^2$.

הוכחה: כמקודם נסמן $\mu=E[f]$ ונשתמש בהגדרת השונות. אז,

$$\mathrm{Var}[f]=E\left[(f-\mu)^2\right]=E\left[f^2-2\mu f+\mu^2\right]\,.$$

בשוויון האחרון פתחנו סוגריים כרגיל בביטוי של הפונקציה $(f-\mu)^2$. עתה נשתמש בליניאריות של התוחלת ונרשום:

$$E\left[f^2-2\mu f+\mu^2\right]=E\left[f^2\right]-E\left[2\mu f\right]+E\left[\mu^2\right]=E\left[f^2\right]-2\mu E\left[f\right]+\mu^2$$

שימו לב שהמספר μ הוא קבוע. לכן, בגורם האמצעי $E\left[2\mu f\right]$, המספר 2μ הוא גורם קבוע הכופל את f, ולכן $E\left[2\mu f\right]=2\mu E\left[f\right]$. ואילו בגורם האחרון חושבים על μ^2 כעל פונקציה קבועה שזהו ערכה. ואולם $\mu=E\left[f\right]$ ולכן נקבל,

$$\begin{aligned}\mathrm{Var}\left[f\right] &= E\left[f^2\right]-2\mu E\left[f\right]+\mu^2\\ &= E\left[f^2\right]-2\mu^2+\mu^2\\ &= E\left[f^2\right]-\mu^2\\ &= E\left[f^2\right]-\left(E\left[f\right]\right)^2\end{aligned}$$

למען הסר ספק: הביטוי $E\left[f^2\right]$ הוא התוחלת של המשתנה המקרי f^2, ואילו $\left(E\left[f\right]\right)^2$ הוא הריבוע של המספר $E\left[f\right]$. □

דוגמה 8.5.6: נתבונן שוב במרחב $(\Omega, \Pr)$ המתאר את תוצאות ההטלה של שתי קוביות שונות, ויהי f המשתנה המקרי המתאים לכל הטלה את סכום הקוביות. ראינו כבר שהתוחלת של f היא $E[f] = 7$. נחשב כעת את השונות. המשתנה המקרי f^2 מוגדר על ידי $f^2((i,j))=(i+j)^2$. התוחלת של f^2 היא:

$$E\left[f^2\right]=\sum_{x\in\Omega}\Pr(x)\cdot f^2(x)=\frac{1}{36}\sum_{x\in\Omega}f^2(x)=\frac{1}{36}\sum_{1\le i,j\le 6}(i+j)^2=\frac{329}{6}.$$

את הסכום האחרון חישבנו בעזרת טענה 8.5.7 שהוכחתה היא תרגיל לקוראים. יוצא שהשונות היא:

$$\mathrm{Var}[f]=E\left[f^2\right]-\left(E[f]\right)^2=\frac{329}{6}-7^2=\frac{35}{6}.$$

טענה 8.5.7: יהי n מספר טבעי. אז $\sum_{i=1}^{n}\sum_{j=1}^{n}(i+j)^2=\frac{n^2(n+1)(7n+5)}{6}$.

לאור מה שלמדנו על התוחלת, מתבקש לשאול אם גם לשונות יש תכונת ליניאריות. כלומר, עלינו לברר מה קורה כשמכפילים משתנה מקרי בקבוע וכאשר מסכמים שני משתנים מקריים. לשאלה הראשונה הדנה בכפל בקבוע, יש תשובה פשוטה, אך בנוגע לסכום אין ביטוי פשוט.

טענה 8.5.8: יהי f משתנה מקרי ממשי, ויהי a מספר ממשי כלשהו. אז $Var[a\cdot f]= a^2\cdot Var[f]$.
הוכחה: נשתמש בטענה 8.5.5 ובליניאריות של התוחלת ונקבל,

$$\begin{aligned}
Var[a\cdot f] &= E\left[a^2\cdot f^2\right]-\left(E[a\cdot f]\right)^2\\
&= a^2\cdot E\left[f^2\right]-\left(a\cdot E\left[f\right]\right)^2\\
&= a^2\cdot E\left[f^2\right]-a^2\cdot\left(E[f]\right)^2\\
&= a^2\cdot\left(E\left[f^2\right]-\left(E[f]\right)^2\right)\\
&= a^2\cdot Var[f]
\end{aligned}$$

ובזאת מסתיימת ההוכחה. □

מה ניתן לומר על השונות כאשר מסכמים שני משתנים מקריים? באופן כללי אם f,g הם שני משתנים מקריים, אין זה נכון ש- $Var[f + g] = Var[f] + Var[g]$. כך למשל, על ידי שימוש בטענה 8.5.8, נקבל ש-

$$Var\left[f+f\right]=Var\left[2\cdot f\right]=4Var\left[f\right]\neq Var\left[f\right]+Var\left[f\right]$$

דוגמה נוספת לכך אפשר לראות בדוגמה הבאה.

דוגמה 8.5.9: נתבונן במרחב המדגם $\Omega=\{H,T\}$ המתאר את תוצאות ההטלה של מטבע, עם ההסתברויות $Pr(H)=Pr(T)=\frac{1}{2}$. נגדיר שני משתנים מקריים $f,g:\Omega\to\mathbb{R}$ על ידי

$$\begin{aligned}
f(H)=1,\quad & f(T)=-1\\
g(H)=-1,\quad & g(T)=1
\end{aligned}$$

אפשר לחשוב על המשתנים המקריים האלה כך: שניהם מתארים ניסוי בהטלת מטבע. במקרה של f אנחנו מרוויחים שקל אחד כשהמטבע מראה "ראש" ומפסידים שקל אחד ב"זנב". המשתנה המקרי g מוגדר ההיפך: רווח של שקל ב"זנב" והפסד של שקל ב"ראש".
קל לוודא כי:

$$\begin{aligned}
&E\left[f\right]=E\left[g\right]=0\\
&Var\left[f\right]=E\left[f^2\right]-\left(E\left[f\right]\right)^2=1
\end{aligned}$$

ובדומה גם $Var\left[g\right]=1$. ואולם המשתנה המקרי f + g הוא פונקצית האפס, ולכן $Var\left[f+g\right]=0$ (בדקו). כלומר גם במקרה זה $Var\left[f+g\right]\neq Var\left[f\right]+Var\left[g\right]$.

אולם יש מקרה מיוחד שבו גם השונות מתנהגת היטב, וזאת כאשר מסכמים שני משתנים מקריים בלתי-תלויים.

טענה 8.5.10: יהיו f,g שני משתנים מקריים בלתי-תלויים. אז $Var[f + g] = Var[f] + Var[g]$.

הוכחה: המשתנים המקריים f,g בלתי תלויים ולכן כפי שראינו (טענה 8.4.11):
$E[f\cdot g]=E[f]\cdot E[g]$.

מכאן, על ידי שימוש חוזר בליניאריות של התוחלת נקבל:

$$\begin{aligned}
\mathrm{Var}[f+g] &= E\left[(f+g)^2\right]-\left(E[f+g]\right)^2\\
&= E\left[f^2+2f\cdot g+g^2\right]-\left(E[f]+E[g]\right)^2\\
&= E\left[f^2\right]+2E[f\cdot g]+E\left[g^2\right]-E[f]\cdot E[f]-2E[f]\cdot E[g]-E[g]\cdot E[g]\\
&= E\left[f^2\right]-\left(E[f]\right)^2+E\left[g^2\right]-\left(E[g]\right)^2+2E[f\cdot g]-2E[f]\cdot E[g]
\end{aligned}$$

אולם, לפי ההגדרה $E\left[f^2\right]-\left(E[f]\right)^2=\mathrm{Var}[f]$ וגם $E\left[g^2\right]-\left(E[g]\right)^2=\mathrm{Var}[g]$. כמו-כן, $2E[f\cdot g]-2E[f]\cdot E[g]=0$ כי $E[f\cdot g]=E[f]\cdot E[g]$ למשתנים מקריים בלתי-תלויים. לכן, נקבל:

$$\mathrm{Var}[f+g]=\mathrm{Var}[f]+\mathrm{Var}[g]$$

כנדרש. □

תרגילים

1. נחזור לתרגיל 6 מסעיף 8.1. חשבו את התוחלת של הערך שמראה הקוביה. מהי השונות של המשתנה המקרי הזה?
2. בדוגמה 8.3.17, בסעיף 8.3, דנו במשתנה המקרי המונה את מספר נקודות השבת של תמורה π של $\{1,\ldots,n\}$, וראינו שהתוחלת שלו היא 1. מהי השונות שלו?
3. נתבונן במרחב הסדרות באורך n הבנויות מ- $\{0,1\}$, כאשר ההתפלגות היא אחידה, ונגדיר משתנה מקרי f המונה את מספר הרצפים 11 בסדרה (ייתכנו חפיפות ברצפים). כך למשל, $f(011101011)=3$. מהי התוחלת של המשתנה המקרי הזה? מהי השונות שלו?
4. נביט במרחב ההסתברות של הטלת זוג קוביות שונות, ונגדיר משתנה מקרי על ידי $f((i,j))=i\cdot j$ (כמו בדוגמה 8.4.12). מהי השונות של f?
5. נביט במרחב כל התמורות של $\{1,\ldots,n\}$ ובמשתנה המקרי המונה את מספר הזוגות המוחלפים (ראו תרגיל 4, סעיף 8.3). מהי השונות של המשתנה המקרי הזה?

8.6. מרחבי מכפלה

הנושאים שיוצגו: *חוק המספרים הגדולים, מרחב מכפלה, התוחלת והשונות של מספר תוצאות "ראש" בהטלת n מטבעות (מוטים), ההתפלגות הבינומית.*

נניח שעומד לרשותנו מטבע מאוזן (כלומר ההסתברות שיצא "ראש" שווה להסתברות שיצא "זנב"), ואנו מטילים אותו מספר גדול של פעמים – n. האינטואיציה הפשוטה אומרת שצפוי כי מספר המופעים של "ראש" ומספר המופעים של "זנב" יהיו "די קרובים" ל- $n/2$. זהו אכן מקרה פרטי של משפט יסודי בהסתברות שנקרא **חוק המספרים הגדולים**. בהמשך נראה גרסה מסוימת של עקרון חשוב זה (ראו משפט 8.7.10 בסעיף 8.7). כרגיל ראשית עלינו להגדיר באופן מסודר את מרחב ההסתברות שבו נשתמש. נגדיר תחילה את המושג של מרחב מכפלה.

הגדרה 8.6.1: יהיו $(\Omega_1, P_1), (\Omega_2, P_2)$ שני מרחבי הסתברות בדידים. **מרחב המכפלה** $(\Omega_1 \times \Omega_2, Pr)$ הוא מרחב הסתברות שקבוצת האיברים שלו היא המכפלה הקרטזית $\Omega_1 \times \Omega_2$, כלומר אוסף כל הזוגות (x,y) כאשר $x\in\Omega_1,\ y\in\Omega_2$. פונקצית ההסתברות המוגדרת על מרחב זה היא $Pr(x,y) = P_1(x)\cdot P_2(y)$.

נעיר שבעצם ראינו כבר בעבר מרחבי מכפלה. כך למשל, כאשר אנו דנים בהטלה של זוג קוביות שונות, מרחב המדגם הוא $\Omega = \{(i,j) \mid 1 \le i, j \le 6\}$ עם התפלגות אחידה. במקרה זה Ω הוא פשוט מכפלה קרטזית של שני מרחבי מדגם מהצורה $\{i \mid 1 \le i \le 6\}$ עם התפלגות אחידה, המתארים את תוצאת ההטלה של קוביה אחת.

כמובן, ניתן להרחיב את הגדרת המכפלה ולדון במכפלה של מספר כלשהו של מרחבי הסתברות.

הגדרה 8.6.2: יהיו $(\Omega_1, P_1),\ldots,(\Omega_k, P_k)$ מרחבי הסתברות בדידים. **מרחב המכפלה** $(\Omega_1 \times\ldots\times \Omega_k, Pr)$ הוא מרחב הסתברות שקבוצת האיברים שלו היא המכפלה הקרטזית $\Omega_1 \times\ldots\times \Omega_k$, כלומר אוסף כל ה- k-יות $(x_1,\ldots,x_k)$ כאשר $x_1\in\Omega_1,\ldots,x_k\in\Omega_k$. פונקצית ההסתברות במרחב זה היא $Pr(x_1,\ldots,x_k) = \prod_{i=1}^{k} P_i(x_i)$.

אם כל מרחבי ההסתברות במכפלה זהים, משתמשים בסימון פשוט יותר.

הגדרה 8.6.3: יהי (Ω, P) מרחב הסתברות בדיד. **מרחב המכפלה** (Ω^n, Pr) הוא מרחב הסתברות שקבוצת האיברים שלו היא המכפלה הקרטזית $\Omega^n = \underbrace{\Omega\times\ldots\times\Omega}_{n}$, כלומר אוסף כל ה- n-יות $(x_1,\ldots,x_n)$ כאשר $x_1,\ldots,x_n\in\Omega$. פונקצית ההסתברות במרחב זה היא $Pr(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(x_i)$.

דוגמה 8.6.4: נתבונן במרחב המדגם $\Omega = \{H,T\}$ המתאר את תוצאות ההטלה של מטבע מאוזן, כלומר $\Pr(H) = \Pr(T) = ½$. אז Ω^n הוא המרחב הכולל את כל 2^n הסדרות של n הטלות מטבע, כאשר לכל סדרה כזו יש הסתברות $\left(\frac{1}{2}\right)^n$.

כך למשל, אם $n = 3$, אז $\Omega^3 = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$, וההסתברות של כל סדרה באורך 3 היא $\left(\frac{1}{2}\right)^3 = \frac{1}{8}$.

על מנת לברר את השאלות שהעלינו בתחילת הסעיף, נגדיר את המשתנה המקרי $f : \Omega^n \to \mathbb{R}$ המייחס לסדרה כלשהי של n הטלות מטבע את מספר הפעמים שמופיע "ראש" בסדרה. כך למשל, במקרה ש- $n = 3$, $f(HHT) = 2$.

נחשב תחילה מהי ההסברות שהמשתנה המקרי f מקבל ערך כלשהו k. מספר הסדרות באורך n שיש בהן k "ראשים" ו- $(n-k)$ "זנבות" הוא $\binom{n}{k}$, וזאת מכיוון שיש לבחור k מתוך n המקומות בסדרה שבהם נרשום "ראש" (טענה 4.2.15). לכל סדרה כזו יש הסתברות $\left(\frac{1}{2}\right)^n$. לכן,

$$\Pr(f = k) = \binom{n}{k} \cdot \left(\frac{1}{2}\right)^n .$$

מהי התוחלת של f? נראה בשתי דרכים כי $E[f] = \frac{n}{2}$, ישירות מההגדרה ובעזרת הליניאריות של התוחלת.

הוכחה א': על פי משפט 8.4.7, ניתן לחשב את התוחלת ישירות כך:

$$E[f] = \sum_{k=0}^{n} k \cdot \Pr(f = k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} \cdot \left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^n \sum_{k=0}^{n} k \cdot \binom{n}{k} = \frac{n}{2} .$$

השוויון האחרון נובע ממשפט 4.3.4.

הוכחה ב': נגדיר את המשתנה המקרי המציין $f_i : \Omega^n \to \{0,1\}$ על ידי

$$f_i(x_1,\dots,x_n) = \begin{cases} 1, & x_i = H \\ 0, & x_i = T \end{cases} ,$$

כלומר ערכו של המשתנה המקרי f_i הוא 1 אם המטבע ה- i בסדרה נופל על "ראש", ואחרת ערכו 0. לכן, בדוגמה של $n = 3$, נקבל למשל $f_2(HHT) = 1$ ואילו $f_3(HHT) = 0$. כמובן:

$$f = \sum_{i=1}^{n} f_i .$$

לכן בגלל הליניאריות של התוחלת נקבל:

$$E[f]=\sum_{i=1}^{n}E[f_i] .$$

מהי התוחלת של f_i? לא קשה לראות כי:

$$E[f_i]=1\cdot\frac{1}{2}+0\cdot\frac{1}{2}=\frac{1}{2} .$$

ולכן,

$$E[f]=\sum_{i=1}^{n}E[f_i]=\frac{n}{2} .$$

הראינו שמספר המופעים הממוצע של "ראש" בסדרה של n הטלות מטבע הוא n/2. תוצאה זו די צפויה. מה שיהיה פחות מובן מאליו ורב חשיבות הוא הדיון בשאלה מהי הסטייה הצפויה מהתוחלת של n/2. לשם כך עלינו לחשב כמובן את השונות של f. מכיוון ש- $f=\sum_{i=1}^{n}f_i$ והמשתנים המקריים f_i בלתי-תלויים (בדקו!), אז לפי טענה **8.5.10**:

$$Var[f]=\sum_{i=1}^{n}Var[f_i] .$$

נחשב תחילה את התוחלת של המשתנים המקריים f_i^2. מכיוון שהמשתנה המקרי f_i מקבל רק את הערכים 0 ו- 1, אז כך גם המשתנה המקרי f_i^2. לכן:

$$E\left[f_i^2\right]=1\cdot\frac{1}{2}+0\cdot\frac{1}{2}=\frac{1}{2} .$$

לכן השונות של f_i היא:

$$Var[f_i]=E\left[f_i^2\right]-\left(E[f_i]\right)^2=\frac{1}{2}-\frac{1}{4}=\frac{1}{4} .$$

ולכן,

$$Var[f]=\sum_{i=1}^{n}Var[f_i]=\frac{n}{4} .$$

דוגמה 8.6.5: מעניין להרחיב את הדיון לסדרת הטלות של מטבע מוטה. נתבונן שוב במרחב המדגם $\Omega=\{H,T\}$ המתאר את תוצאות ההטלה של מטבע, כשהפעם ההסתברויות הן $Pr(H)=p$, $Pr(T)=q$, כאשר $0\le p,q\le 1$ וכמובן $p+q=1$.

נעבור עתה לניתוח המרחב Ω^n. נגדיר כמקודם את המשתנה המקרי f המונה את מספר מופעי "ראש" בסדרה של n הטלות מטבע. גם הפעם נחשב את התוחלת של f בשתי דרכים.

דרך א׳: אנו נחשב לכל k את ההסתברות שהמשתנה המקרי f מקבל ערך k. הפעם לסדרה באורך n, עם k ״ראש״ ו- $(n-k)$ ״זנב״, יש הסתברות $p^k q^{n-k}$. מספר הסדרות האלה הוא $\binom{n}{k}$ ולכן,

$$\Pr(f=k)=\binom{n}{k}p^k q^{n-k}\ .$$

נחשב כעת את התוחלת של f:

$$\begin{aligned}
E[f] &= \sum_{k=0}^{n} k\cdot \Pr(f=k) = \\
&= \sum_{k=0}^{n} k\cdot\binom{n}{k}\cdot p^k\cdot q^{n-k} = \\
&= \sum_{k=0}^{n-1} n\cdot\binom{n-1}{k-1}\cdot p^k\cdot q^{n-k} \\
&= n\cdot p\cdot\sum_{k=0}^{n-1}\binom{n-1}{k-1}\cdot p^{k-1}\cdot q^{n-k} \\
&= n\cdot p\cdot(p+q)^{n-1} = n\cdot p
\end{aligned}$$

בשוויון הלפני אחרון השתמשנו בנוסחת הבינום של ניוטון (ראו משפט 4.3.1).

אגב, להתפלגות של המשתנה המקרי f שראינו זה עתה יש חשיבות גדולה במגוון רחב של בעיות. כפי שהראינו ההתפלגות של f היא:

$$\Pr(f=k)=\binom{n}{k}p^k\cdot q^{n-k}$$

שמה המקובל הוא **ההתפלגות הבינומית** עם פרמטרים p,n וסימונה $B(n, p)$.

דרך ב׳: נגדיר כמקודם את המשתנים המקריים f_i, כש- $1 \le i \le n$. כלומר, f_i הוא משתנה מקרי מציין למאורע ״המטבע ה- i בסדרה נפל על ראש״. בדומה לדוגמה הקודמת, נטען כי $f=\sum_{i=1}^{n} f_i$ ובגלל הליניאריות של התוחלת $E[f]=\sum_{i=1}^{n}E[f_i]$. ואולם כאן $E[f_i]=p$ לכל i, ולכן:

$$E[f]=\sum_{i=1}^{n}E[f_i]=n\cdot p\ .$$

הראינו בשתי דרכים כי ממוצע מספר הראשים בסדרה באורך n הוא $n\cdot p$. זהו מקרה פרטי של חוק המספרים הגדולים בתורת ההסתברות. חוק זה אומר במילים כך: יהיה f משתנה מקרי ממשי. נדגום את f מספר גדול של פעמים n ונמצע את הערכים המתקבלים. וקצת ביתר פירוט:

נגריל n איברים $x_1,\dots,x_n$ מתוך מרחב המדגם Ω, כשההסתברות להגריל את האיבר x_i היא $\Pr(x_i)$ וההגרלות נעשות באופן בלתי תלוי. נתבונן בממוצע $\frac{1}{n}\sum_{i=1}^{n} f\left(x_i\right)$. אז כאשר n שואף לאינסוף הממוצע שואף לגבול, וגבול זה הוא E[f]. הניסוח המלא של עקרון זה מופיע כאמור במשפט 8.7.10.

נחשב לסיום את השונות של המשתנה המקרי f. גם כאן אין קושי להרחיב את החישוב מהמקרה של המטבע המאוזן ($p = ½$) למקרה כללי. גם הפעם המשתנים המקריים f_i בלתי-תלויים ולכן,

$$.\ \mathrm{Var}\left[f\right] = \sum_{i=1}^{n} \mathrm{Var}\left[f_i\right]$$

השונות של f_i היא:

$$.\ \mathrm{Var}\left[f_i\right] = E\left[f_i^2\right] - \left(E\left[f_i\right]\right)^2 = p - p^2 = p(1-p) = p \cdot q$$

ולכן:

$$.\ \mathrm{Var}\left[f\right] = \sum_{i=1}^{n} \mathrm{Var}\left[f_i\right] = n \cdot p \cdot q$$

יש עדיין בעיה מציקה שלא באה על פתרונה. הוכחנו למשל שב- n הטלות של מטבע לא מוטה, התוחלת של מספר תוצאות ה"ראש" היא n/2. הדעת נותנת שלא צפויה סטייה גדולה מדי ממספר זה. כך למשל, אילו דווח לנו שבניסוי שבו הוטל מטבע מספר גדול של פעמים, יצא "ראש" למעלה מ- 60% מהפעמים, היינו סבורים, ובצדק, שהמטבע מוטה. זוהי הנחה נכונה שאותה נוכיח בסעיף הבא.

תרגילים

1. מטילים קוביה, כאשר ההסתברות שיצא כל אחד מ- 6 המספרים שעל הקוביה היא 1/6. כעת מטילים מטבע מאוזן מספר פעמים על פי המספר שמראה הקוביה. תארו במפורט את מרחב ההסתברות וחשבו את ההסתברות שבכל אחת מן ההטלות של המטבע יצא "ראש"?

2. א. מטילים מטבע מאוזן מספר פעמים ועוצרים אם שתי הטלות עוקבות יוצאות זהות, או כאשר הגענו להטלה ה- n. מה ההסתברות שאכן נגיע להטלה ה- n?
   ב. כנ"ל אבל עוצרים כאשר יוצאת אותה התוצאה שלוש פעמים ברציפות.
   ג. כנ"ל אבל עוצרים כאשר יוצאת אותה התוצאה k פעמים ברציפות.

3. באיזה מבין שני המשחקים הבאים יש הסתברות גבוהה יותר לנצח:
   א. מטילים 4 קוביות ומנצחים אם לפחות אחת מהן מראה את המספר 1.
   ב. מטילים 24 פעמים 2 קוביות ומנצחים אם לפחות באחת ההטלות שתי הקוביות הראו את המספר 1.
   הדרכה: חשבו בכל אחד מהמקרים מהי ההסתברות להפסיד.

4. מטילים קוביה 10 פעמים. מהי ההסתברות שהמספר 5 הופיע בדיוק 3 פעמים? שערו מהו המספר הטבעי k כך שההסתברות ש- 5 הופיע בדיוק k פעמים היא מירבית. אמתו או הפריכו את השערתכם.

8.7. אי-שוויונות יסודיים בהסתברות

הנושאים שיוצגו: *אי-שוויון מרקוב, אי-שוויון צ'בישב, הערכות זנב להתפלגות הבינומית, החוק החלש של המספרים הגדולים, אי-שוויון צ'רנוף, התפלגות נורמלית, משפט הגבול המרכזי.*

כפי שציינו, לרוב איננו יכולים לחשב באופן מלא את ההתפלגות של המשתנים המקריים הנדונים, אולם די במידע חלקי עליהם כדי לענות על רבות מן השאלות שברצוננו לפתור. בסעיף זה נציג דוגמאות של אי-שוויונות הסתברותיים המאפשרים לנו לקבל מידע על תכונות יסודיות של התפלגויות, תוך שימוש בתוחלת ובשונות בלבד. זהו קצה הקרחון של תחום רחב ופעיל של מחקר הדן בשאלה הבאה: נניח שלפנינו משתנה מקרי f וידועה לנו התוחלת שלו $\mu = \mathrm{E}[\mathrm{f}]$. עד כמה שכיח המצב שערכו של f קרוב לתוחלת שלו μ? בעיות אלה והדיון בהן נקראות ניתוח של ריכוז מידה או הערכות זנב. כאמור, אנו נטעם משטח זה רק על קצה המזלג.

נניח ש- f משתנה מקרי ותוחלתו μ ידועה לנו. האם אפשרי ש- f נוטה לקבל ערכים גדולים, נניח למעלה מ- 10μ? באופן כללי הדבר אפשרי בהחלט.

דוגמה 8.7.1: נניח ש- f משתנה מקרי המקבל רק את הערכים 100 ו- 3200- בהסתברויות הבאות:

$$\Pr(\mathrm{f} = 100) = \frac{97}{100}, \quad \Pr(\mathrm{f} = -3200) = \frac{3}{100}.$$

התוחלת של f היא כמובן:

$$\mathrm{E}[\mathrm{f}] = 100 \cdot \frac{97}{100} + (-3200) \cdot \frac{3}{100} = 1.$$

אולם בהסתברות גדולה מאוד של 0.97 הערך של f גדול בהרבה מהתוחלת והוא 100. מה קורה בדוגמה הזו? הערך השלילי מאוד (3200- במקרה זה) מקטין במידה ניכרת את התוחלת, על אף הסתברותו הנמוכה $\frac{3}{100}$.

אם f משתנה מקרי שאינו מקבל ערכים שליליים, תופעה כזאת איננה אפשרית, כפי שמראה המשפט הבא.

משפט 8.7.2 (אי-שוויון מרקוב): יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $\mathrm{f}: \Omega \to \mathbb{R}$ משתנה מקרי ממשי אי-שלילי. אז לכל מספר ממשי $\lambda > 0$,

$$\Pr\left(\mathrm{f} \geq \lambda \cdot \mathrm{E}[\mathrm{f}]\right) \leq \frac{1}{\lambda}.$$

הוכחה: מכיוון שהמשתנה המקרי f הוא אי-שלילי אז:

$$\begin{aligned} E[f] &= \sum_{x\in\Omega} \Pr(x)\cdot f(x) \\ &= \sum_{y\geq 0} y\cdot \Pr(f=y) \end{aligned}$$

עתה נתבונן רק בחלק מהסכום, וניקח בחשבון רק ערכים y בתחום $y \geq \lambda\cdot E[f]$. מכיוון ש- $f \geq 0$, כל המחוברים הם אי-שליליים ולכן:

$$\begin{aligned} E[f] &= \sum_{y\geq 0} y\cdot \Pr(f=y) \\ &\geq \sum_{y\geq \lambda\cdot E[f]} y\cdot \Pr(f=y) \end{aligned}$$

כעת נחליף כל y בסכום האחרון בחסם התחתון ונקבל:

$$\begin{aligned} E[f] &\geq \sum_{y\geq \lambda\cdot E[f]} y\cdot \Pr(f=y) \\ &\geq \lambda\cdot E[f]\cdot \sum_{y\geq \lambda\cdot E[f]} \Pr(f=y) \\ &= \lambda\cdot E[f]\cdot \Pr\left(f\geq \lambda\cdot E[f]\right) \end{aligned}$$

נחלק כעת את שני צדי האי-שוויון שקיבלנו ב- $\lambda\cdot E[f]$, ונקבל את המשפט. □

דוגמה 8.7.3: אם ממוצע הציונים שלכם הוא חלילה 20 מתוך 100, אז השגתם ציון 80 ומעלה לכל היותר ברבע מהמבחנים שנבחנתם.

נניח שוב ש- f משתנה מקרי אי-שלילי. האם סביר שהערכים ש- f מקבל יהיו קטנים בהרבה מן התוחלת שלו? בפרט, האם ייתכן ש- $f = 0$ בהסתברות גבוהה? ללא מגבלות הדבר אפשרי בהחלט. למשל, אם f מקבל רק את הערכים 10,000 ו- 0 בהסתברויות:

$$\Pr(f = 10{,}000) = 0.01 \qquad ,\Pr(f = 0) = 0.99$$

אז $E[f] = 100$ (בדקו!). ואולם בהסתברות גבוהה של 0.99 המשתנה המקרי f מתאפס. הסיבה לתופעה זו היא ש- f מקבל ערכים גבוהים מאוד (במקרה זה 10,000). ואכן ניתן להוכיח:

טענה 8.7.4: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $f:\Omega\to\mathbb{R}$ משתנה מקרי ממשי אי-שלילי שכל ערכיו הם בתחום $[0,a]$. אז: $\Pr(f=0)\leq 1-\frac{E[f]}{a}$.

הוכחה: נחשב את התוחלת של f:

$$E[f] = \sum_{x\in\Omega} x\cdot \Pr(f=x) = \sum_{\substack{x\in\Omega,\\ x\neq 0}} x\cdot \Pr(f=x) \leq a\cdot \sum_{\substack{x\in\Omega,\\ 0<x\leq a}} \Pr(f=x) = a\left(1-\Pr(f=0)\right).$$

הטענה נובעת על ידי העברת אגפים. □

נשים לב שבאי-שוויון מרקוב מעורבת רק התוחלת, אך יש צורך בהנחה ש- f משתנה מקרי שמקבל ערכים אי-שליליים. כפי שראינו השונות היא מדד למידת הפיזור או הריכוז של ערכי

המשתנה המקרי f. אין פלא לכן, שבעזרת התוחלת והשונות ניתן לקבל הערכות טובות יותר בדבר הריכוז של ערכי f. זהו האי-שוויון הקלאסי של צ'בישב. הניסוח הישיר ביותר לכך שהמשתנה המקרי f מרוכז סביב התוחלת $\mu = E\left[f\right]$ הוא: בהסתברות גבוהה $\left|f-\mu\right|$ קטן. ובמילים: אם נגריל איבר x מהמרחב, אז קרוב לוודאי ש- f(x) יהיה קרוב לערך הממוצע μ.

משפט 8.7.5 (אי-שוויון צ'בישב): יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $f:\Omega \to \mathbb{R}$ משתנה מקרי ממשי. אז לכל מספר ממשי $C > 0$,

$$\Pr\left[\left|f - E[f]\right| \geq C\right] \leq \frac{\mathrm{Var}[f]}{C^2}.$$

הוכחה: לא נוח לעבוד עם ערכים מוחלטים. נשים לב שהתנאי $|a| > b > 0$ שקול לתנאי $a^2 > b^2$.
לכן, נגדיר משתנה מקרי אי-שלילי $g = (f - E[f])^2$. טענת המשפט היא אם כן:

$$\Pr\left[g \geq C^2\right] \leq \frac{\mathrm{Var}[f]}{C^2}.$$

יתרון נוסף של הגדרת g הוא שעל פי הגדרת השונות $E[g] = \mathrm{Var}[f]$. נפעיל את אי-שוויון מרקוב על המשתנה המקרי האי-שלילי g עם הקבוע החיובי $\lambda > 0$ המוגדר כך:

$$\lambda = \frac{C^2}{\mathrm{Var}[f]} = \frac{C^2}{E[g]}.$$

לכן,

$$\Pr\left[g \geq \lambda \cdot E[g]\right] \leq \frac{1}{\lambda}$$

כלומר:

$$\Pr\left[g \geq C^2\right] \leq \frac{\mathrm{Var}[f]}{C^2}$$

ומכאן:

$$\Pr\left[(f - E[f])^2 \geq C^2\right] \leq \frac{\mathrm{Var}[f]}{C^2}$$

כנדרש. □

הערה: האינטואיציה הבסיסית אכן מתממשת באי-שוויון צ'בישב. ככל שהשונות קטנה יותר, כן מרוכזים יותר ערכי f סביב התוחלת שלו. וההסתברות לכך שהערכים של f יסטו כדי C מן התוחלת אינה עולה על $\frac{\mathrm{Var}[f]}{C^2}$.

דוגמה 8.7.6: נחזור לדוגמאות הכרוכות בסדרה של n הטלות של מטבע מאוזן, כאשר המשתנה המקרי f מציין את מספר הפעמים שיצא "ראש" בסדרה. אמנם בסדרה של n הטלות מטבע, הערך בעל ההסתברות הגבוהה ביותר של f הוא n/2, כאשר ההסתברות לערך זה היא:

$$. \Pr\left(f = \frac{n}{2}\right) = \binom{n}{n/2} \cdot \frac{1}{2^n} = \Theta\left(\frac{1}{\sqrt{n}}\right)$$

(השתמשנו במסקנה 7.4.1 לקירוב המקדם הבינומי האמצעי.) זהו אמנם הערך המסתבר ביותר, אך ההסתברות שלו קטנה (שואפת ל- 0 כש- n גדול). במונחים אינטואיטיביים סביר כי תהיה גם סטייה מסוימת מן הערך $f = n/2$, אך מהי? נפעיל את אי-שוויון צ'בישב. כזכור:

$$. E[f] = \frac{n}{2}, \quad Var[f] = \frac{n}{4}$$

לכן,

$$. \Pr\left(\left|f - \frac{n}{2}\right| \geq C\right) \leq \frac{n}{4C^2}$$

אם נציב למשל $C = 5\sqrt{n}$ נקבל שבהסתברות של לפחות 0.99 הערך של f יהיה בין:

$$. \frac{n}{2} - 5\sqrt{n} \leq f \leq \frac{n}{2} + 5\sqrt{n}$$

בניתוח מדויק יותר שלא יוצג כאן, ניתן להוכיח כי ההערכה של חסם צ'בישב קרובה להדוקה. כלומר, ב- n הטלות מטבע כמעט וודאי שנקבל $\frac{n}{2} \pm \Theta\left(\sqrt{n}\right)$ מופעים של "ראש".

דוגמה 8.7.7 (הערכות זנב להתפלגות הבינומית $B(n,p)$): נחזור שוב לסדרה של n הטלות של מטבע מאוזן. מהי ההסתברות שמספר המופעים של "ראש" יהיה קטן או שווה ל- $n/10$? זהו בוודאי מאורע מאוד לא סביר. ואכן, אי-שוויון צ'בישב מראה כי:

$$. \Pr\left(f \leq \frac{n}{10}\right) \leq \Pr\left(\left|f - \frac{n}{2}\right| \geq \frac{4}{10}n\right) \leq \frac{1}{0.64n} < \frac{2}{n}$$

לאמיתו של דבר, ההסתברות הזו נמוכה בהרבה, והיא:

$$\Pr\left(f \leq \frac{n}{10}\right) = \sum_{k=0}^{n/10} k \cdot \Pr(f = k)$$
$$= \sum_{k=0}^{n/10} k \cdot \binom{n}{k} \cdot \frac{1}{2^n}$$

בסכום האחרון יש $n/10 +1$ מחוברים והגדול מביניהם הוא $\frac{n}{10}\binom{n}{n/10}$, ולכן אפשר לחסום את הסכום שקיבלנו כך:

$$\begin{aligned}\Pr\left(f \le \frac{n}{10}\right) &= \sum_{k=0}^{n/10} k \cdot \binom{n}{k} \cdot \frac{1}{2^n} \\ &\le \left(1+\frac{n}{10}\right)\frac{n}{10}\binom{n}{n/10}\frac{1}{2^n} \\ &= 2^{-n(1-H(1/10)+o(1))}\end{aligned}$$

השוויון האחרון נובע מהקירובים של המקדמים הבינומיים בעזרת פונקצית האנטרופיה (ראו מסקנה 7.4.3). כלומר, החסם העליון על ההסתברות למאורע זה קטן **מעריכית** מהחסם המתקבל מאי-שוויון צ'בישב. בהמשך נצטט ללא הוכחה את אי-שוויון צ'רנוף המספק הערכות מדויקות בהרבה מאי-שוויון צ'בישב במצבים כגון זה.

משפט 8.7.8: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $f:\Omega \to \mathbb{N}$ משתנה מקרי שהטווח שלו הוא המספרים הטבעיים. אז: $\Pr(f=0) \ge 1-E[f]$.

הוכחה: נחשב את התוחלת של f. מכיוון ש- f מקבל רק ערכים שלמים אי-שליליים אז:

$$E[f] = \sum_{i=0}^{\infty} i \cdot \Pr(f=i).$$

המחובר הראשון של $i = 0$ אינו תורם דבר לסכום, ולכן:

$$E[f] = \sum_{i=1}^{\infty} i \cdot \Pr(f=i).$$

נחליף כעת את i ב- **1** בכל המחוברים. הסכום רק יקטן, ולכן נקבל:

$$E[f] \ge \sum_{i=1}^{\infty} \Pr(f=i) = 1-\Pr(f=0).$$

נעביר אגפים ונקבל את הנדרש. □

מסקנה שימושית פשוטה מהמשפט האחרון היא שאם $f:\Omega \to \mathbb{N}$ משתנה מקרי שהטווח שלו הוא המספרים הטבעיים, ואם $E[f]<1$, אז יש הסתברות חיובית לכך שהמשתנה המקרי f מתאפס, ובפרט יש מאורע $x \in \Omega$ כך ש- $f(x) = 0$. אנחנו נראה שימושים למסקנה זו בסעיף 8.8.

הטענה הפשוטה הבאה שימושית מאוד, כפי שנראה בהמשך.

טענה 8.7.9: יהי $(\Omega, \Pr)$ מרחב הסתברות בדיד ויהי $f:\Omega \to \mathbb{R}$ משתנה מקרי ממשי עם תוחלת $\mu = E[f]$. אז f מקבל ערך גדול או שווה ל- μ וגם f מקבל ערך קטן או שווה ל- μ.

נצטט עוד כמה מהמשפטים המרכזיים החשובים בהסתברות.

משפט 8.7.10 (החוק החלש של המספרים הגדולים): יהיו $f_1,\ldots,f_n$ משתנים מקריים בלתי-תלויים שלכולם אותה התוחלת $\mu = E[f_i]$ לכל $1 \le i \le n$, וכן אותה השונות $\alpha = Var[f_i]$ לכל $1 \le i \le n$. יהי $S_n = f_1+\ldots+f_n$ ויהי $\varepsilon > 0$. אז: $\lim_{n\to\infty} Pr\left(\left|\frac{S_n}{n} - \mu\right| \ge \varepsilon\right) = 0$.

הוכחה: נשתמש בליניאריות של התוחלת ובעובדה שלמשתנים המקריים f_i יש תוחלת משותפת μ ונקבל:

$$E\left[\frac{S_n}{n}\right] = \frac{1}{n}\sum_{i=1}^{n} E[f_i] = \frac{n\mu}{n} = \mu .$$

כמו-כן, מכיוון שהמשתנים המקריים בלתי-תלויים ויש להם שונות משותפת α, אז לפי טענה 8.5.10:

$$Var\left[\frac{S_n}{n}\right] = \frac{1}{n^2}\sum_{i=1}^{n} Var[f_i] = \frac{n\alpha}{n^2} = \frac{\alpha}{n} .$$

לכן, בעזרת אי-שוויון צ'בישב נקבל:

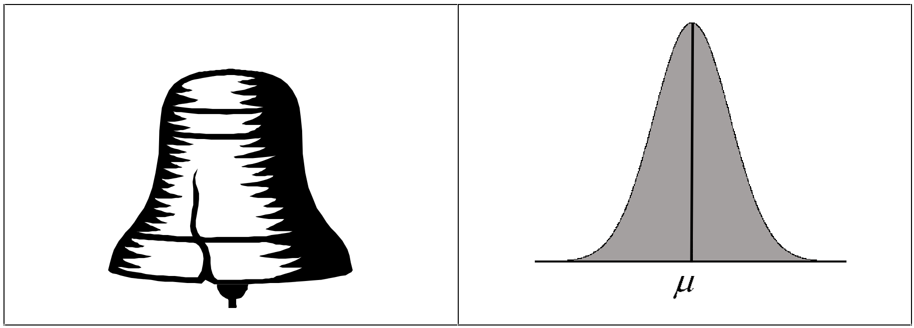
$$Pr\left(\left|\frac{S_n}{n} - \mu\right| \ge \varepsilon\right) \le \frac{\alpha}{n\varepsilon^2} .$$

ולכן כאשר $n \to \infty$ החסם הזה שואף ל- 0 כפי שטענו. □

משפט 8.7.11 (אי-שוויון צ'רנוף): יהיו $f_1,\ldots,f_n$ משתנים מקריים מציינים בלתי-תלויים כך ש- $Pr[f_i = 1] = p_i$ לכל $1 \le i \le n$, כאשר $0 < p_i < 1$. יהי $f = f_1+\ldots+f_n$ ותהי $\mu = E[f]$ התוחלת של f. אז,

1. לכל $0 < \delta \le 1$ מתקיים $Pr\left(f < (1-\delta)\cdot\mu\right) < e^{-\mu\cdot\delta^2/2}$.
2. לכל $0 < \delta \le 2e-1$ מתקיים $Pr\left(f > (1+\delta)\cdot\mu\right) < e^{-\mu\cdot\delta^2/4}$.

נזכיר בקצרה גם את **משפט הגבול המרכזי**. זהו אחד המשפטים המרשימים ביותר בהסתברות. נתקלתם ודאי במקומות שונים במושג **עקומת פעמון** או מה שמכונה בשפה המקצועית **התפלגות נורמלית** או **התפלגות גאוסיאנית**. משפט הגבול המרכזי אומר שאם f משתנה מקרי שהוא סכום של משתנים מקריים רבים $f = f_1 + \ldots + f_n$, כאשר f_i שווי התפלגות ובלתי תלויים, אז ההתפלגות של f קרובה להתפלגות נורמלית. הנקודה המפתיעה היא שאין זה חשוב מהי ההתפלגות של כל אחד מהמשתנים המקריים f_i. בכל מקרה לסכומם f תהיה התפלגות נורמלית בקירוב.



תרשים 8.7.1: גרף הפעמון של ההתפלגות הנורמלית ולידו פעמון שהתפלג קצת...

תרגילים

1. הראו שהחסם בטענה 8.7.4 הדוק. מהי התפלגות הערכים של f אם יש שוויון בטענה הנ"ל?
2. האם החסם באי-שוויון צ'בישב הדוק? אם כן, מה תוכלו לומר על התפלגות הערכים של משתנה מקרי f שבשבילו מתקיים אי-שוויון צ'בישב כשוויון?

8.8. שימושים של הסתברות בקומבינטוריקה

הנושאים שיוצגו: *מאורע שמתקיים כמעט בוודאות, פרדוקס יום ההולדת, אספן קופונים, השיטה ההסתברותית, חסם תחתון למספרי רמזי, 2-צביעה של היפרגרפים.*

כמה מהעובדות הבסיסיות והמועילות ביותר בתורה של מרחבי הסתברות בדידים מתקבלות מעיון בתהליך הפשוט הבא: יש n תאים ואנחנו משליכים לתוכם באקראי כדורים. נתעניין בעיקר בשתי שאלות:

1. מהו מספר הכדורים שעלינו להשליך לתאים עד שלראשונה יכיל כל תא לפחות כדור אחד? התשובה: בערך $n \log n$ (בעיית אספן הקופונים).
2. מתי לראשונה יהיה תא המכיל יותר מכדור אחד? התשובה: בערך $\sqrt{n}$ (פרדוקס יום ההולדת).

בשאלות אלו ואחרות נעסוק בסעיף זה. אנו נתעניין אמנם בשאלות האלה גם לערכים ספציפיים של n, אולם מוקד הדיון שלנו הוא בנקודת המבט האסימפטוטית. כך למשל, אנו נראה בהמשך שאם משליכים הרבה יותר מ- $\sqrt{n}$ כדורים באקראי לתוך n תאים, אז **כמעט בוודאות** יהיה תא עם יותר מכדור אחד. משמעות המושג הזה היא שההסתברות לכך שיהיה תא עם יותר מכדור אחד היא $1-o(1)$, כלומר הסתברות זו שואפת ל- 1 כאשר $n \to \infty$.

פרדוקס יום ההולדת

נציג כעת את פרדוקס יום ההולדת שהוא בין השימושים הבסיסיים של תורת ההסתברות בקומבינטוריקה. נפתח בתיאור החידה הקלאסית: מה ההסתברות שבכיתה שלכם (הכוללת נאמר **60** תלמידים) יש שניים בעלי אותו יום הולדת? הכוונה לאותו תאריך בשנה, אך לאו דווקא באותה שנה. שיקולים נאיביים יכולים להוביל לניחושים שונים, אך התשובה הנכונה מפתיעה למדי. ההסתברות שכל ימי ההולדת של **60** התלמידים הם שונים היא רק:

$$1\cdot\frac{364}{365}\cdot\frac{363}{365}\cdots\frac{305}{365}=0.0049112 .$$

כלומר ההסתברות לכך היא פחות מחצי אחוז (ודאו שאתם יכולים להוכיח זאת).

להלן נדון בשאלה בצורה כללית. נתון מרחב מדגם $\Omega=\{1,2,...,n\}$. אנו חוזרים k פעמים על ניסוי שבו אנו בוחרים בכל צעד באופן בלתי-תלוי איבר מתוך הקבוצה Ω. מהי ההסתברות לכך שכל האיברים הנבחרים שונים זה מזה ואף אחד מהם לא נבחר פעמיים או יותר? קל לראות שבחידת ימי ההולדת, המרחב Ω מתאר את $n = 365$ ימי השנה, ובחירת $k = 60$ איברים ממנו מתארת בחירת יום הולדת לתלמידים השונים.
אגב, התיאור הנ"ל זהה לבעיה שבה פתחנו את הסעיף שדנה בהשלכת k כדורים באקראי ובאופן בלתי-תלוי ל- n תאים. אנו שואלים כאן מהי ההסתברות שבאף תא אין יותר מכדור אחד.

נשוב לניתוח הבעיה. נתבונן במרחב ההסתברות Ω^k, הכולל את כל הסדרות באורך k מתוך הקבוצה $\Omega=\{1,2,...,n\}$. המרחב כולל $\left|\Omega^k\right|=n^k$ איברים, עם הסתברות אחידה של n^{-k} לכל סדרה $(x_1,...,x_k)\in\Omega^k$. המאורע המעניין אותנו הוא המאורע הבא:

$$A = \{ (x_1,\ldots,x_k) \in \Omega^k \mid \text{כל האיברים } x_1,\ldots,x_k \text{ שונים זה מזה} \}.$$

לפי משפט 4.2.8:

$$|A|=n(n-1)\cdots(n-k+1) .$$

לכן, ההסתברות של המאורע A היא:

$$\begin{aligned}\Pr(A)&=\frac{n(n-1)\cdots(n-k+1)}{n^k}\\&=1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\\&=\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right)\end{aligned}$$

נחזור כעת לשאלת יום ההולדת, שבה $n = 365$, ונבדוק מהו ה- k שעבורו $\Pr(A) > ½$, ומתי $\Pr(A) < ½$. נקבל שאם $k \geq 23$ אז ההסתברות של- 23 אנשים מקריים יש ימי הולדת שונים קטנה מ- ½. לעומת זאת, אם $k \leq 22$ אז ההסתברות של- 22 אנשים מקריים יש ימי הולדת

שונים גדולה מ- ½. לכן, אם תבקרו במשחק כדורגל, אז יש הסתברות קטנה מ- ½ שלשחקני שתי הקבוצות ולשופט, המהווים יחד קבוצה של 11 + 11 + 1 = 23 אנשים, יהיו ימי הולדת שונים.

חישוב המכפלה האחרונה לערכים ספציפיים של n,k הוא מייגע, ולכן נוח יהיה לקבל הערכה של ערכה האסימפטוטי של המכפלה. אנו נוכיח את המשפט הבא:

משפט 8.8.1: $\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right)=e^{-\Theta\left(k^2/n\right)}$, כאשר $k < n/3$.

(הערה: ההנחה כי $k < n/3$ אינה הכרחית ומשמשת כאן רק כדי לפשט את הדיון. ראו תרגיל 4.)

הוכחה: נמצא חסם עליון ותחתון למכפלה הזאת.

חסם תחתון: נשתמש באי-שוויון $1-x > e^{-2x}$ המתקיים לכל $x < 1/3$ (ראו משפט 7.3.3). לכן,

$$\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right) \ge \prod_{i=0}^{k-1} e^{-2i/n} = e^{-\sum_{i=0}^{k-1} 2i/n} = e^{-k(k-1)/n}.$$

כאמור, כדי להשתמש באי-שוויון $1-x > e^{-2x}$ עלינו להניח כי $x < 1/3$. במקרה שלנו ההנחה מתקיימת שהרי $\frac{i}{n} < \frac{1}{3}$ לכל $0 \le i \le k-1$, שהרי הנחנו ש- $k < n/3$.

חסם עליון: מצד שני, נשתמש באי-שוויון $1-x \le e^{-x}$ המתקיים לכל x (מסקנה 7.3.2). נקבל:

$$\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right) \le \prod_{i=0}^{k-1} e^{-i/n} = e^{-\sum_{i=0}^{k-1} i/n} = e^{-k(k-1)/(2n)}.$$

נשלב את שני החסמים ונקבל כי $\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right)=e^{-\Theta\left(k^2/n\right)}$ כפי שטענו. □

בשאלות רבות בהסתברות אנו מסתפקים בידיעה האם מאורע מסוים הוא כמעט וודאי (הסתברות קרובה מאוד ל- 1) או כמעט בלתי אפשרי (הסתברות קרובה מאוד ל- 0). מהו המצב כאן? נשים לב כי $e^{-\Theta\left(k^2/n\right)}$ קרוב מאוד ל- 1 אם k קטן יחסית ל- $\sqrt{n}$, ומאידך הביטוי קרוב מאוד ל- 0 אם k גדול יחסית ל- $\sqrt{n}$.

נסכם: אם מתוך עולם של n איברים בוחרים עם חזרות הרבה פחות מ- $\sqrt{n}$ איברים, אז כמעט בוודאות לא ייבחר אותו איבר פעמיים. מאידך, אם בוחרים הרבה יותר מ- $\sqrt{n}$ איברים, אז כמעט בוודאות יהיה איבר שייבחר לפחות פעמיים.

אספן קופונים

בעיה קלאסית נוספת שנדון בה היא בעיית אספן הקופונים. ניתן לנסחה בכמה אופנים שקולים. למשל, חברת המשקאות "מי אפסיים" מקדמת בקיץ את מכירותיה במבצע נושא פרסים. בכל

פקק של בקבוק מתוצרתה מוטבעת אחת מאותיות הא"ב. כל אחת מאותיות הא"ב מופיעה בשכיחות שווה של 1/22. הפרס ניתן למי שמצליח לאסוף את כל אותיות הא"ב. מה מספר הבקבוקים שעליכם לקנות על מנת שיהיה לכם סיכוי טוב לזכות בפרס?

באופן כללי יותר יהיה לנו א"ב של n אותיות, כלומר מרחב המדגם הוא $\Omega = \{1, 2, ..., n\}$. כרגיל נגדיר את Ω^k כאוסף הסדרות מאורך k מעל Ω, כאשר יש הסתברות אחידה של n^{-k} לכל סדרה ב- Ω^k. המאורע שבו נתעניין הוא:

$$A = \{(x_1,\dots,x_k) \in \Omega^k \mid \text{כל } n \text{ האותיות מופיעות בין } x_1,\dots,x_k\}.$$

נרצה להעריך את ההסתברות $\Pr(A)$ כפונקציה של n ו- k. מטרתנו להוכיח שחל מעבר בסביבות $k = n{\cdot}\ln(n)$. כלומר, אם k קטן ממש מ- $n{\cdot}\ln(n)$ אז כמעט וודאי שסדרה מאורך k אינה כוללת את כל אותיות הא"ב, ולמאורע A יש הסתברות קטנה. לעומת זאת אם k ממש גדול מ- $n{\cdot}\ln(n)$ אז כמעט וודאי שסדרה מאורך k כוללת את כל אותיות הא"ב, והסתברות המאורע A קרובה ל- 1. נוכיח כעת את המשפט הבא.

משפט 8.8.2: ההסתברות שסדרה מקרית באורך k מתוך א"ב בגודל n, כוללת את כל אותיות הא"ב היא לפחות $1 - n \cdot e^{-k/n}$.

הוכחה: נגדיר משתנה מקרי $f : \Omega^k \to \mathbb{R}$ המונה בכל סדרה $(x_1,\dots,x_k)$ את מספר האותיות השונות **המופיעות** בה. לצורך הניתוח נוח יותר להתבונן במשתנה המקרי $g = n - f$. משתנה מקרי זה מונה את מספר האותיות **החסרות** בסדרה. ניתן לבטא את המשתנה המקרי g כך. נגדיר n משתנים מקריים $g_1,\dots,g_n$ על ידי:

$$g_i(x_1,...,x_k) = \begin{cases} 1, & i \notin \{x_1,...,x_k\} \\ 0, & i \in \{x_1,...,x_k\} \end{cases},$$

כלומר המשתנה המקרי g_i מקבל את הערך 1 על סדרה $(x_1,\dots,x_k)$ אם האות i אינה מופיעה בסדרה. אחרת, ערכו של g_i הוא 0. כמובן מתקיים:

$$g = \sum_{i=1}^{n} g_i,$$

ולכן:

$$E[g] = \sum_{i=1}^{n} E[g_i].$$

איך נחשב את התוחלת של g_i? מכיוון ש- g_i משתנה מקרי מציין, כפי שראינו התוחלת שלו איננה אלא ההסתברות ש- $g_i = 1$, כלומר שהאות i אינה מופיעה בסדרה $(x_1,\dots,x_k)$. לכן,

$$\begin{aligned} E[g_i] &= \Pr(x_1 \neq i, x_2 \neq i, \dots, x_k \neq i) \\ &= \Pr(x_1 \neq i) \cdot \Pr(x_2 \neq i) \cdots \Pr(x_k \neq i) \\ &= \left(1 - \frac{1}{n}\right)^k \end{aligned}$$

וזאת כי האות i אינה מופיעה בסדרה $(x_1,\dots,x_k)$ אם ורק אם $x_1 \neq i, x_2 \neq i,\dots,x_k \neq i$. אלה k מאורעות בלתי תלויים ולכל אחד מהם הסתברות $1-\frac{1}{n}$.

נשוב ונחשב כעת את התוחלת של g. על פי הליניאריות של התוחלת:

$$E[g]=\sum_{i=0}^{n}E[g_i]=n\left(1-\frac{1}{n}\right)^k \le n\cdot e^{-k/n}.$$

באי-שוויון האחרון השתמשנו בעובדה כי $(1-x) < e^{-x}$.

ברצוננו לדעת מהי ההסתברות שבסדרה $(x_1,\dots,x_k)$ מופיעות כל אותיות הא״ב $\{1,2,\dots,n\}$. כזכור, המשתנה המקרי g מונה את מספר האותיות הנעדרות. לכן, אם בסדרה $(x_1,\dots,x_k)$ מופיעות כל אותיות הא״ב אז $g(x_1,\dots,x_k) = 0$. אנו רוצים להעריך את ההסתברות $\Pr(g = 0)$. נפעיל את משפט 8.7.8 (אפשר להפעילו כי g משתנה מקרי המקבל ערכים טבעיים) ונקבל:

$$\Pr(g=0)\ge 1-E[g]\ge 1-n\cdot e^{-k/n}$$

ובכך הוכח המשפט. □

כעת נשתמש במשפט כדי לבדוק מהו הערך של k שיבטיח כמעט בוודאות שסדרה באורך k תכיל את כל אותיות הא״ב. ואכן, אם $k \ge a\cdot n\cdot \ln(n)$ כאשר $a > 1$ קבוע, אז ההסתברות שסדרה מקרית באורך k כוללת את כל אותיות הא״ב היא לפחות:

$$1-n\cdot e^{-k/n}\ge 1-n\cdot e^{-a\ln n}=1-n^{-(a-1)}.$$

אולם $n^{-(a-1)}$ שואף לאפס כאשר n שואף לאינסוף, ולכן ההסתברות שהסדרה כוללת את כל אותיות הא״ב קרובה מאוד ל- 1. כלומר, בסדרה מקרית מאורך $a\cdot n\cdot \ln(n)$, כמעט וודאי שכל אותיות הא״ב תופענה.

ניתן גם להוכיח שאם $k = a\cdot n\cdot \ln(n)$ ו- $a < 1$, אז בסדרה מקרית מאורך k כמעט בוודאות לא כל האותיות תופענה. את העובדה הזאת מוכיחים באמצעות הפעלת אי-שוויון צ׳בישב על המשתנה המקרי g שהגדרנו בהוכחת משפט 8.8.2. סוג כזה של שיקול נקרא **שיטת המומנט השני**, אך לא נעסוק בכך בספר זה.

חסם תחתון למספרי רמזי

כזכור, בסעיף 5.7 הוגדר מספר רמזי $R(s,t)$ בתור המספר הטבעי n המזערי כך שבכל צביעה של צלעות הגרף השלם K_n בכחול ואדום יש בהכרח תת-גרף שלם K_s שצבוע בכחול או יש תת-גרף שלם K_t שצבוע באדום. ראינו גם חסם עליון (משפט 5.7.2):

$$R(s,t)\le\binom{s+t-2}{s-1}$$

ובפרט כאשר $s = t$ נקבל:

$$R(t,t) \le \binom{2t-2}{t-1} = O\left(\frac{4^t}{\sqrt{t}}\right).$$

החסם העליון מסתמך על הערכותינו למקדמים הבינומיים (ראו מסקנה 7.4.1). ובפרט אם $n > 4^t$ אז בכל צביעה של צלעות K_n בכחול ואדום נקבל תת-גרף שלם בגודל t שצלעותיו חד-גוניות. כאן ברצוננו למצוא חסם תחתון מעריכי ל- $R(t,t)$. אנו נוכיח כי

$$R(t,t) = \Omega\left(2^{t/2}\right).$$

טענה זו פירושה שניתן לצבוע את צלעות K_n בכחול ואדום כאשר n שווה בערך ל- $2^{t/2}$, כך שלא יתקבל תת-גרף שלם בגודל t שכל צלעותיו כחולות או כל צלעותיו אדומות. קוראים שטרם נחשפו **לשיטה ההסתברותית** יניחו, מן הסתם, שמשפט כזה יוכח על ידי תיאור מפורש של הצביעה. כלומר, יש לומר לכל זוג אינדקסים $1 \le i < j \le n$, האם הצלע $\{i,j\}$ צבועה בכחול או באדום. אולם, עד היום איננו יודעים לתת הוכחה כזו, הקרויה **בנייה מפורשת** של הצביעה, וזאת למרות שהושקע מאמץ רב בחיפוש אחריה. כוחה של השיטה ההסתברותית יודגם כאן. השיטה הזאת מאפשרת לנו להוכיח שצביעה כזו **קיימת** בלי להצביע במפורש על צביעה כנ"ל. ניגש כעת להוכחתו של המשפט.

משפט 8.8.3: $R(t,t) = \Omega\left(2^{t/2}\right)$.

הערה: שימו לב שאם $n = 2^{t/2}$ אז $t = 2\log_2 n$, ולכן ניסוח אחר למשפט הוא שיש צביעה של צלעות הגרף השלם K_n בכחול ובאדום כך שאין תת-גרף שלם חדגוני מגודל $t = 2\log_2 n + 2$.

הוכחה: אנו נראה את אוסף כל הצביעות של צלעות K_n בכחול ואדום כמרחב הסתברות Ω עם התפלגות אחידה. יש במרחב הזה $2^{\binom{n}{2}}$ איברים (צביעות) ולכן ההסתברות של כל צביעה היא $2^{-\binom{n}{2}}$. אנו רואים איבר כללי ω במרחב Ω כפונקציה

$$\omega : \{\{i,j\} \mid 1 \le i < j \le n\} \to \{b,r\}$$

המתאימה לכל צלע $\{i,j\}$ של K_n צבע כחול b או אדום r.

נאמר שקבוצת קדקודים S היא **חדגונית** בצביעה מסוימת ω אם לכל זוג קדקודים $i,j \in S$ מתקיים $\omega(i,j) = b$, או שלכל $i,j \in S$ מתקיים $\omega(i,j) = r$. כלומר, צלעות התת-גרף השלם הכולל בדיוק את קדקודי הקבוצה S צבועות רק בצבע אחד.

ההוכחה מסתמכת על דיון במשתנה המקרי f המונה את מספר התת-קבוצות S מגודל t שהן חדגוניות. כלומר, אם $\omega \in \Omega$ צביעה של צלעות K_n, אז $f(\omega)$ הוא מספר התת-קבוצות $S \subseteq \{1,\dots,n\}$ מעוצמה $|S| = t$, כך שהצביעה ω מייחסת אותו צבע לכל צלע $\{i,j\}$ כאשר $i,j \in S$.

רעיון ההוכחה הוא זה: אנו נראה שמהנחת המשפט נובע כי התוחלת של המשתנה המקרי f מקיימת $E[f] < 1$. לפי משפט 8.7.8 יוצא ש-

$$\Pr(f = 0) \ge 1 - E[f] > 0.$$

כלומר המאורע $\{\omega \mid f(\omega)=0\}$ אינו ריק. פירוש הדבר הוא שיש צביעה ω ללא תת-גרף שלם חדגוני מגודל t. שימו לב, כפי שציינו מקודם, שיטת ההוכחה הזאת אינה מצביעה על דרך לחשב את הצביעה ω או למצוא אותה, אלא רק מוכיחה את קיומה.

ניגש לחישוב התוחלת של f. כמו בדוגמאות קודמות, נוח יהיה לבטא את f כסכום של משתנים מקריים מציינים. תהי $S \subseteq \{1,\dots,n\}$ תת-קבוצה מעוצמה $|S| = t$. המשתנה המקרי f_S יציין אם הצלעות בין קדקודי הקבוצה S צבועות באותו הצבע בצביעה ω. כלומר, $f_S(\omega)=1$ אם הקבוצה S חדגונית בצביעה ω, ואחרת $f_S(\omega)=0$. קל לראות כי:

$$f=\sum_{|S|=t} f_S ,$$

ולכן,

$$E[f]=\sum_{|S|=t} E[f_S] .$$

נחשב כעת את $E[f_S]$. היות ש- f_S משתנה מקרי מציין, התוחלת שלו $E[f_S]$ שווה להסתברות שצביעה מקרית ω היא חדגונית על הצלעות שבין קדקודי S. נראה כי ההסתברות ש- ω חדגונית על תת-קבוצה S היא $\left(\frac{1}{2}\right)^{\binom{t}{2}-1}$. ניתן לראות את המצב כאילו אנו עובדים במרחב מכפלה:

$$\underbrace{\{b,r\}\times\{b,r\}\times\ldots\times\{b,r\}}_{\binom{t}{2}}$$

עם גורם אחד כנגד כל צלע בין קדקודי S. זהו מרחב הסתברות עם $2^{\binom{t}{2}}$ איברים והתפלגות אחידה. המאורע "הצביעה ω חדגונית על צלעות S" כוללת שני מאורעות בסיסיים $(b,b,\dots,b)$ ו- $(r,r,\dots,r)$, כלומר כל הצלעות כחולות או כל הצלעות אדומות. לכן, ההסתברות ש- ω חדגונית על S היא:

$$2\cdot 2^{-\binom{t}{2}}=\left(\frac{1}{2}\right)^{\binom{t}{2}-1}$$

נחזור כעת לחישוב התוחלת על פי הביטוי

$$E[f]=\sum_{|S|=t} E[f_S] .$$

בסכום הזה יש $\binom{n}{t}$ מחוברים, שכן זהו מספר התת-קבוצות מגודל t מתוך איברי הקבוצה $\{1,2,\dots,n\}$. כל מחובר שווה ל- $\left(\frac{1}{2}\right)^{\binom{t}{2}-1}$ כפי שהראינו. קיבלנו לכן,

$$.\ E[f]=\binom{n}{t}\left(\frac{1}{2}\right)^{\binom{t}{2}-1}$$

כזכור, עלינו להראות עתה שאם $t > 2\log n + 1$ אז $E[f]<1$. נשתמש בהערכה הפשוטה למקדמים הבינומיים (ראו תרגיל 2 בסעיף 7.4):

$$.\ \binom{n}{t}<n^t$$

ובסה"כ נקבל:

$$.\ E[f]<2n^t\left(\frac{1}{2}\right)^{t(t-1)/2}=2\left(\frac{n}{2^{(t-1)/2}}\right)^t$$

ואכן, אם $t > 2\log_2 n + 1$ נקבל כי $E[f]<1$ כנדרש. ובזאת נשלמת ההוכחה. □

והנה עוד דוגמה שבה משתמשים בשיטה ההסתברותית.

בעיית הלבוש הססגוני או 2-צביעה של היפרגרפים

כזכור, בסעיף 5.5 עסקנו בצביעה של גרפים. גרף הוגדר כ- 2-צביע אם אפשר לצבוע את קדקודי הגרף בשני צבעים, כך שאף צלע לא תהיה חדגונית. כלומר, כל שני קדקודים שכנים בגרף חייבים להיות צבועים בצבעים שונים. **היפרגרף** הוא הרחבה של מושג הגרף המאפשרת לצלע לכלול יותר משני קדקודים. במקרה זה הצלעות הן קבוצות חלקיות של קדקודים. גם במקרה זה אפשר לדבר על 2-צביעה של ההיפרגרף. שוב נצבע את קדקודי הגרף בשני צבעים ונדרוש שאף צלע לא תהיה חדגונית. האם תמיד הדבר אפשרי? אנו נעסוק בשאלה זאת כעת, אולם תחילה נציג אותה באופן ציורי יותר.

נתון אוסף של $k = 500$ קבוצות $A_1,\ldots,A_k$ כשבכל קבוצה יש 10 אנשים, כלומר $|A_i| = 10$ לכל $1 \le i \le k$. כמו-כן ייתכן שאדם מסוים שייך ליותר מקבוצה אחת, ולכן החיתוך של הקבוצות אינו בהכרח ריק. יהי $V=\bigcup_{i=1}^{k}A_i$ אוסף כל האנשים המשתייכים ל- k הקבוצות הנ"ל, ונניח ש- $|V| = n$. ברצוננו להלביש כל אחד מהאנשים בחולצה אדומה או צהובה, כך שאף אחת מ- k הקבוצות לא תהיה חדגונית. כלומר, בכל אחת מהקבוצות יהיה אדם אחד לפחות הלבוש בצהוב ואדם אחד לפחות הלבוש באדום. אנו לא נראה במפורש כיצד לעשות זאת, אולם נוכיח שהדבר אפשרי.

אם נחזור לרגע לבעיית ההיפרגרפים, קבוצת הקדקודים של ההיפרגרף היא הקבוצה V של n האנשים, וצלעות ההיפרגרף הן הקבוצות החלקיות $A_1,\ldots,A_k$. במקרה זה כל צלע מכילה 10 קדקודים. ברצוננו לצבוע את קדקודי הגרף בשני צבעים, אדום וצהוב, כך שלא תהיה צלע חדגונית.

נשוב אל בעיית החולצות ונניח שבחירת החולצה נעשית באקראי. מרחב ההסתברות שלנו הוא $\Omega^n=\{Y,R\}^n$, כאשר Y מייצג חולצה צהובה, ו- R מייצג חולצה אדומה. מאורע אופייני $\omega\in\Omega^n$ הוא סדרה באורך n המתארת כיצד לבוש כל אדם בקבוצה V. נגדיר k משתנים מקריים $f_1,\ldots,f_k$ כדלקמן:

$f_i(\omega)=1$ אם הקבוצה A_i חדגונית במאורע הבסיסי ω, ו- $f_i(\omega)=0$ אחרת.

כלומר $f_i(\omega)=1$ אם כל חברי הקבוצה A_i לבושים באופן זהה. נגדיר גם משתנה מקרי $f=\sum_{i=1}^{k} f_i$. שימו לב ש- $f(\omega)$ הוא מספר הקבוצות החדגוניות המתקבלות במאורע הבסיסי ω, ומטרתנו להראות שיש $\omega\in\Omega^n$ כך ש- $f(\omega)=0$, כלומר אף קבוצה אינה חדגונית. כדי להראות זאת נוכיח ש- $E[f]<1$. מכיוון שהמשתנה המקרי f מקבל ערכים שלמים אי-שליליים, אז יש בהכרח $\omega\in\Omega^n$ כך ש- $f(\omega)=0$, כדרוש (ראו משפט 8.7.8 וההערה שאחריו).

נראה שאכן $E[f]<1$ כנדרש. בגלל הליניאריות של התוחלת מתקיים:

$$E[f]=\sum_{i=1}^{k}E\left[f_i\right].$$

מכיוון ש- f_i משתנה מקרי מציין, אז $E\left[f_i\right]$ שווה להסתברות ש- A_i חדגונית. מכלל 2^{10} הבחירות האפשריות של בגדים לחברי הקבוצה A_i, יש בדיוק שתיים חדגוניות – הבחירה שבה כולם לובשים אדום והבחירה שבה כולם לובשים צהוב. לכן,

$$E\left[f_i\right]=2\left(\frac{1}{2}\right)^{10}=2^{-9}=\frac{1}{512}.$$

לכן,

$$E[f]=\sum_{i=1}^{k}E\left[f_i\right]=k\cdot\frac{1}{512}=500\cdot\frac{1}{512}<1.$$

ולכן יש $\omega\in\Omega^n$ כך ש- $f(\omega)=0$, כלומר יש בחירה של הביגוד שבה אף אחת מ- k הקבוצות אינה חדגונית.

בדוגמה הנ"ל היו $k = 500$ קבוצות, וכל קבוצה A_i כללה $t = 10$ אנשים. המשפט הבא מגדיר את היחס בין k ל- t המבטיח שתהיה בחירה של החולצות כך שאף קבוצה לא תהיה חדגונית. הוכחת המשפט זהה לניתוח שסיימנו זה עתה.

משפט 8.8.4: נתבונן בהיפרגרף שבו יש k צלעות שכל אחת מהן מכילה בדיוק t קדקודים. אם $k<2^{t-1}$ אז ההיפרגרף 2-צביע.

תרגילים

1. הוכיחו שאם בוחרים k בחירות חוזרות מתוך $\{1,...,n\}$, כאשר k גדול ממש מ- $n^{2/3}$, אז יש הסתברות גבוהה לכך שאותו איבר ייבחר לפחות 3 פעמים. מאידך גיסא אם k קטן ממש מ- $n^{2/3}$, אז כמעט בוודאות אף איבר לא ייבחר יותר מפעמיים.

2. מהו המספר המזערי של אנשים שמבטיח כי בהסתברות גדולה מ- ½ לשלושה מהם אותו יום הולדת?

3. בוחרים k פעמים באופן בלתי-תלוי איבר מתוך הקבוצה $\{1,2,...,n\}$. במשפט **8.8.1** מצאנו חסם עליון וחסם תחתון למכפלה $\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right)$, שהיא ההסתברות שכל k האיברים שבחרנו שונים זה מזה. השתמשו בחסמים אלה כדי למצוא:
   א. חסם עליון לערך של k שיבטיח שבהסתברות גדולה מ- ½ כל האיברים שבחרנו שונים זה מזה.
   ב. חסם תחתון לערך של k המבטיח שבהסתברות גדולה מ- ½ לא כל האיברים שבחרנו שונים זה מזה.
   ג. השוו את החסמים שקיבלתם לערכים המדויקים של $k = 23$ ו- $k = 22$ המתאימים לבעיית היום הולדת, שקיבלנו על ידי שימוש ישיר במכפלה.

4. הוכיחו שמשפט **8.8.1** נכון לכל $k \le n$.
הדרכה: השתמשו בנוסחת סטירלינג כדי להוכיח שהמשפט נכון ל- $k = n$. כעת הסיקו את המשפט לכל $n/3 < k < n$.

5. זורקים $n\log^2 n$ כדורים באקראי ל- n תאים. נניח שהתא הכי מלא מכיל a כדורים והתא הכי ריק מכיל b כדורים. הוכיחו שכמעט בוודאות מתקיים $a/b = O(1)$.
הדרכה: הוכיחו שכמעט בוודאות גם $b = \Theta(\log^2 n)$ וגם $a = \Theta(\log^2 n)$.

6. קיים היפרגרף הכולל 7 קדקודים ו- 7 צלעות, שבכל אחת מצלעותיו יש 3 קדקודים והוא אינו 2-צביע. האם תוכלו למצאו?
זו אינה שאלה הסתברותית, אלא דרושה בנייה מפורשת של ההיפרגרף.

הערות היסטוריות

אנדרי אנדרייביץ' מרקוב Andrei Andreyevich Markov (רוסיה 1856-1922). בתחילת דרכו כמתמטיקאי עסק מרקוב בתורת המספרים ובאנליזה מתמטית, בנושאים כמו גבולות של אינטגרלים, תורת הקירובים והתכנסות של סדרות. הוא השתמש בשיטה של שברים משולבים, שפותחה על ידי מורהו צ'בישב, להוכחת תוצאות בתורת ההסתברות. שמו נקשר בעיקר במושג של שרשראות מרקוב: המודל הבסיסי בתהליכים סטוכסטיים.

פפנוטי לבוביץ' צ'בישב Pafnuty Lvovich Chebyshev (רוסיה 1821-1894). תרם תרומות חשובות לתורת המספרים. הוא גילה שמספר המספרים הראשוניים בין 1 ל- n הוא $\Theta\left(\frac{n}{\log n}\right)$.

התוצאה המדויקת של $\left(1+o(1)\right)\frac{n}{\ln(n)}$ הוכחה רק שנתיים מאוחר יותר על ידי ז'אק אדמר Jacque Hadamard ועל ידי שרל דה לה ואלה פוסין Charles de La Vallée Poussin. צ'בישב כתב גם עבודות חשובות בתחומי האנליזה המתמטית (תורת הקירובים), במכניקה ובהסתברות. ב- 1850 הוא הוכיח את ההשערה של ברטרנד מ- 1845, האומרת שיש לפחות מספר ראשוני אחד בין n ל- 2n לכל n טבעי.