מבוא להסתברות

נתלי תמם ודודו לגזיאל, בית הספר למתמטיקה, אוניברסיטת ת"א

סמסטר א', תשע"ה

תקציר

חוברת זו נכתבה כחוברת מתרגלים בלבד ולכן ייתכנו (וככל הנראה ישנם..) טעויות ואי דיוקים. האחריות על השימוש בהן היא על המשתמש בלבד! החוברת הוכנה בהתבסס על תרגולי העבר המעולים של ד"ר רועי טפר, ד"ר יובל הלר וד"ר אסף כהן ועל רשימות הקורס של פרופ' אילון סולן.

תוכן עניינים

1 יסודות: מרחבי הסתברות ומאורעות.

הבהרה: בקורס אנחנו נסתמך בצורה די רצינית על דרישות הקדם, ובעיקר על הקורסים מבוא לתורת הקבוצות ומבוא לקומבינטוריקה ותורת הגרפים. אנחנו לא נחזור על החומר של הקורסים הנ"ל ולכן כל מי שיש לו פער מבחינת הקורסים הנדרשים, חייב להשלים את החומר בהקדם. בנוסף, נשתמש בכלים שראיתם בקורסי החדו"א והליניארית הראשוניים וכן בהמשך נשתמש באלמנטים אשר נלמדים במקביל בקורסי החדו"א והליניארית המתקדמים.

1.1 ניסויים ומרחבי הסתברות כלליים

1.1.1 מרחבי מדגם ומאורעות

לרוב אנחנו נתחיל בניסוי כלשהו, בין אם מדובר בהטלת מטבע, הטלת קובייה, הוצאת כדורים מכד וכו', כמעט תמיד יהיה לנו ניסוי אשר יהווה את הבסיס לשאלות ולבעיות אותן נפגוש.

הגדרה 1.1 מרחב מדגם זו קבוצה, המסומנת לרוב באות Ω, ומכילה את כל התוצאות האפשריות של הניסוי.

דוגמה 1.2 נסתכל על מספר ניסויים אשר יופיעו רבות בקורס. בכל אחת מן הדוגמאות הבאות, תארו את מרחב המדגם:

1. הטלה מטבע הוגן כאשר על צד אחד שלו רשומה הספרה $'1'$ ועל הצד השני מופיעה הספרה $'0'$ - $\Omega = \{0,1\}$.
2. הטלת אותו המטבע ההוגן שצויין לעיל מספר סופי של פעמים $n \in \mathbb{N}$ - $\Omega = \{0,1\}^n = \{(x_1,\ldots,x_n):\ x_i \in \{0,1\},\ 1 \le i \le n\}$.
3. הטלת קוביה עם 6 פאות - $\Omega = \{1,2,3,4,5,6\}$.
4. הוצאה עם החזרה של 2 כדורים מתוך כד בו יש 2 כדורים לבנים ו־1 שחור - $\Omega = \{(W,W),(W,B),(B,W),(B,B)\}$.
5. הוצאה ללא החזרה של 2 כדורים מתוך כד בו יש 2 כדורים לבנים ו־1 שחור - $\Omega = \{(W,W),(W,B),(B,W)\}$.

הגדרה 1.3 לכל תת קבוצה $A \subseteq \Omega$ של מרחב המדגם נקרא מאורע. איברים במרחב המדגם יסומנו ב־ $\omega \in \Omega$.

מאורע A^c ייקרא המאורע המשלים של מאורע A ויוגדר ביחס למרחב המדגם ע"י $A^c = \Omega \setminus A$. בקורס אנחנו נעסוק במרחבי מדגם סופיים או לכל היותר בני מניה ולכן מעתה נניח כי כל מרחבי המדגם שלנו הם לכל היותר בני־מניה. כמובן שבאופן כללי מרחב מדגם יכול להיות גם מעוצמת הרצף, דוגמת כל הסדרות הבינאריות ועוד.

תרגיל 1.4 בכל אחד מארבעת הימים הבאים מבצעים ניסוי. כל אחד מהניסויים יכול להצליח או להיכשל. נגדיר את המאורע A_i כמאורע בו הניסויי ביום ה־i הצליח.

1. האם המאורעות זרים בזוגות? האם המאורעות זרים?

2. תארו את המאורעות הבאים בעזרת חיתוכים, איחודים ומשלימים של המאורעות שהוגדרו בשאלה בלבד:

   (א) כל הניסויים הצליחו.
   (ב) לפחות אחד מהניסויים הצליח.
   (ג) כל הניסויים נכשלו.
   (ד) הניסוי הראשון שהצליח היה ביום השלישי.
   (ה) הניסוי השני שהצליח היה ביום השלישי.

3. הוכיחו כי המאורע "כל הניסויים נכשלו" הוא המאורע המשלים למאורע "לפחות אחד מהניסויים הצליח".

פתרון: נענה על כל סעיף בנפרד.

1. בשאלה זו למעשה שואלים אותנו האם

$$A_i \cap A_j \quad = \quad \phi, \quad \forall i \neq j \tag{1}$$

והאם

$$.\bigcap_{i=1}^{4} A_i \quad = \quad \phi \tag{2}$$

הביטוי הראשון מתייחסים למצב בו המאורעות זרים בזוגות והשני למצב בו מדובר במאורעות זרים. נגדיר תחילה את מרחב המדגם בבעיה.

$$.\Omega \quad = \quad \{(b_1, \ldots, b_4): \ b_i \in \{0,1\}, \ i = 1,2,3,4\}$$

זהו תיאור מתמטי מדויק של מרחב המדגם כאשר קוארדינטה i בוקטור שמרכיב את מרחב המדגם מקבלת את הערך '1' כאשר הייתה הצלחה בניסוי ביום ה־i ו־'0' אחרת. כעת נוכל להגדיר במפורש את מאורע A_i. נניח כי $i = 1$, אזי

$$A_1 = \{(1, b_2, b_3, b_4): \ b_i \in \{0,1\}, \ i = 2,3,4\}$$

ובאופן דומה

$$A_2 = \{(b_1, 1, b_3, b_4): \ b_i \in \{0,1\}, \ i = 1,3,4\}$$

וכן הלאה. נשים לב כי

$$A_1 \cap A_2 = \{(1, 1, b_3, b_4): \ b_i \in \{0,1\}, \ i = 3,4\} \neq \phi$$

ולכן קיימים i, j כך ש־ $A_i \cap A_j \neq \phi$. על כן המאורעות אינם זרים בזוגות לפי ההגדרה המובעת במשוואה (1). באופן דומה נשים לב שהמאורעות אינם זרים מאחר וקל לראות כי

$$.\bigcap_{i=1}^{4} A_j \quad = \quad \{(1,1,1,1)\}$$

2. בסעיף הנוכחי אנחנו יכולים להיעזר אך ורק במאורעות $A_i,\ i=1,\dots,4$ אשר הוגדרו בשאלה בלבד.

(א) כל הניסויים הצליחו: $\bigcap_{i=1}^{4} A_i$.

(ב) לפחות אחד מהניסויים הצליח: $\bigcup_{i=1}^{4} A_i$.

(ג) כל הניסויים נכשלו: $\bigcap_{i=1}^{4} A_i^c$.

(ד) הניסוי הראשון שהצליח היה ביום השלישי: $A_1^c \cap A_2^c \cap A_3$.

(ה) הניסוי השני שהצליח היה ביום השלישי: $(A_1 \cap A_3 \cap A_2^c) \cup (A_2 \cap A_3 \cap A_1^c)$.

3. נתחיל בתזכורת קטנה בנוגע לחוקי דה־מורגן. יהיו מאורעות כלשהם $A_1,\dots,A_n \subseteq \Omega$. אזי

$$\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} A_i^c$$
$$\left(\bigcap_{i=1}^{n} A_i\right)^c = \bigcup_{i=1}^{n} A_i^c \quad . \tag{3}$$

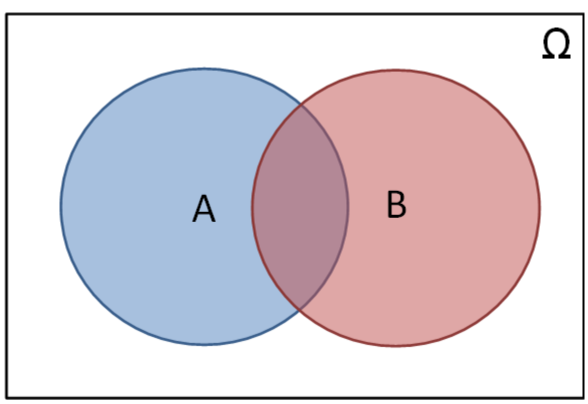
כעת נוכל לראות כי הנדרש בסעיף הוא למעשה תוצאה ישירה של חוקי דה־מורגן,

$$\bigcap_{i=1}^{4} A_i^c = \left(\bigcup_{i=1}^{4} A_i\right)^c$$

כאשר אגף שמאל זה המאורע בו כל הניסויים נכשלו (סעיף 2.א) ובאגף ימין רשום המאורע המשלים למאורע בו לפחות אחד מהניסויים הצליח (סעיף 2.ב).

1.1.2 דיאגרמת ואן וכללי דה־מורגן

דרך נוחה להמחיש את מרחב המדגם והמאורעות היא באמצעות דיאגרמת ואן. מדובר בדיאגרמה אבסטרקטית המתארת את כל מרחב המדגם והמאורעות הרלוונטיים מתוארים בעזרת מעגלים בתוכה.



איור 1: דיאגרמת ואן

נוכל לראות בעזרת הדיאגרמה מדוע, לדוגמא, הכלל הראשון של דה־מורגן המופיע במשוואה (3) נכון.

טענה 1.5 יהי Ω מרחב מדגם. עבור כל צמד מאורעות $A,B\subseteq\Omega$ מתקיים

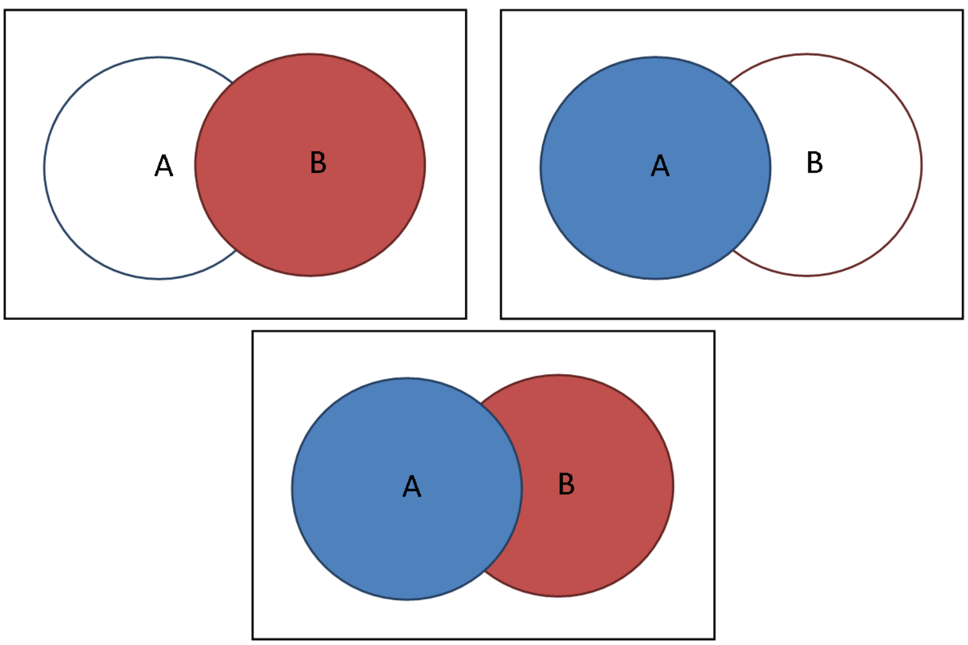
$$(A\cup B)^c \quad = \quad A^c\cap B^c \; .$$

אנו נוכיח טענה זו ובנוסף נשתמש בדיאגרמת ואן בכדי להצדיק אותה.

הוכחה: נוכיח בעזרת הכלה דו־כיוונית. נראה תחילה ש־ $(A\cup B)^c \subseteq (A^c\cap B^c)$. יהי $\omega\in(A\cup B)^c$ (במידה ולא קיים כזה ω אז ההכלה מתקיימת באופן טריוויאלי). ז"א ש־ $\omega\notin A\cup B$. מכאן נובע ש־ $\omega\notin A$ וגם $\omega\notin B$. לכן $\omega\in A^c\wedge\omega\in B^c$ וקיבלנו ש־ $(A\cup B)^c\subseteq(A^c\cap B^c)$.

נוכיח את ההכלה בכיוון השני, $(A^c\cap B^c)\subseteq(A\cup B)^c$. יהי $\omega\in(A^c\cap B^c)$, על כן $\omega\in A^c\wedge\omega\in B^c$. לכן נובע ש־ $\omega\notin A,B$ ובפרט $\omega\notin A\cup B$. לכן קיבלנו ש־ $\omega\in(A\cup B)^c$ והטענה נובעת. ■

נראה זאת בעזרת דיאגרמת ואן. ניקח את המקרה הכללי ביותר המופיע באיור 1 ונסמן את האזור אליו מתייחס כל אגף בטענה. אגף שמאל של המשוואה, $(A\cup B)^c$, מתאר את המאורע המשלים של איחוד המאורעות, שזה האזור הלבן באיור 1. לחילופין, באגף ימין נתון חיתוך של המאורעות המשלימים כפי שהם מוצגים באיור 2. ניתן לראות כי האזורים הלבנים באיור 1 ובשרטוט השלישי באיור 2 הם למעשה אותם האזורים.



איור 2: מרשם סכמטי של אגף ימין של המשוואה $(A\cup B)^c = A^c\cap B^c$. בצמד השרטוטים הראשונים, האזור הצבוע בלבן מתאר את המאורע המשלים של כל אחד מן המאורעות והשרטוט האחרון מציג את החיתוך של המשלימים.

1.1.3 מרחבי הסתברות וחסם האיחוד

הגדרה 1.6 פונקצית הסתברות מסומנת באות $\Pr:\Omega\to[0,1]$ ומוגדרת ממרחב המדגם לקטע $[0,1]$.

לפונקצית ההסתברות ישנן מספר תכונות בסיסיות:

1. נרמול: $\sum_{\omega\in\Omega}\Pr(\omega)=1$.
2. אדיטיביות:
$$\Pr(A)=\sum_{\omega\in A}\Pr(\omega)\,. \tag{4}$$
זוהי למעשה הרחבה של הגדרת פונקצית ההסתברות כך שהיא תוגדר על תתי קבוצות ולא רק על איברים. שימו לב שניתן באותה מידה להגדיר את הפונקציה מראש על תתי קבוצות, ז"א מ־2^{Ω} לקטע $[0,1]$.
3. ההסתברות של הקבוצה הריקה היא אפס: $\Pr(\phi)=0$. תכונה זו נובעת ישירות מן התכונה הקודמת בעקבות כך שאנחנו סוכמים, למעשה, על 0 איברים.
4. ההסתברות של כל מרחב המדגם היא אחת: $\Pr(\Omega)=1$ (כתוצאה של שילוב תכונות הנרמול והאדיטיביות).

בעזרת פונקצית ההסתברות ומרחב המדגם, אנחנו יכולים להגדיר מרחב הסתברות באופן הבא.

הגדרה 1.7 מרחב הסתברות זה זוג $(\Omega,\Pr)$ כאשר Ω הוא מרחב מדגם ו־$\Pr$ היא פונקצית הסתברות המוגדרת על מרחב המדגם Ω.

טענה 1.8 יהיה $A\subseteq\Omega$ מאורע במרחב הסתברות $(\Omega,\Pr)$. אזי
$$\Pr(A)=1-\Pr(A^c)\,. \tag{5}$$

הוכחה: נזכור כי $A^c=\Omega\setminus A$ ולכן,
$$\begin{aligned}\Pr(A)+\Pr(A^c) &= \sum_{\omega\in A}\Pr(\omega)+\sum_{\omega\in\Omega\setminus A}\Pr(\omega)=\\ &= \sum_{\omega\in\Omega}\Pr(\omega)=\\ &= \Pr(\Omega)=1\end{aligned}$$

והתוצאה נובעת מהעברת $\Pr(A^c)$ לאגף השני. ■

תרגיל 1.9 יהיה מרחב מדגם $\Omega=\mathbb{N}\setminus\{0\}$ ונגדיר פונקצית ההסתברות עליו $\Pr(i)=\frac{k}{3^i}$ כאשר k קבוע כלשהו. נגריל מספר $i\in\Omega$ על פי ההתפלגות $\Pr$.

1. מצאו את k.
2. חשבו את הסיכוי ש־ i זוגי?
3. חשבו את הסיכוי ש־i אי־זוגי?

פתרון: נשים לב תחילה כי מדובר במרחב מדגם לא־אחיד (לא סימטרי).

1. בכדי לפתור סעיפים מסוג זה נשתמש בתנאי הנרמול של פונקצית ההסתברות

$$\sum_{i=1}^{\infty} \Pr(i) = \sum_{i=1}^{\infty} \frac{k}{3^i} = 1.$$

נשים לב כי יש לנו טור של סדרה הנדסית אותו נוכל לחשב באופן ישיר. נגדיר

$$S_n = \sum_{i=1}^{n} \frac{k}{3^i},$$

ולכן מתקיים ש־ $S_n = S_{n-1} \cdot \frac{1}{3} + \frac{k}{3}$ ומצד שני $S_n = S_{n-1} + \frac{k}{3^n}$. נשווה את צמד הביטויים ונקבל

$$S_{n-1} + \frac{k}{3^n} = S_{n-1} \cdot \frac{1}{3} + \frac{k}{3}$$
$$S_{n-1} = \frac{\frac{k}{3} - \frac{k}{3^n}}{1 - \frac{1}{3}} = k \cdot \frac{\frac{1}{3} - \left(\frac{1}{3}\right)^n}{1 - \frac{1}{3}}.$$

מאחר ו־ $S_\infty = \lim_{n\to\infty} S_n = \sum_{i=1}^{\infty} \frac{k}{3^i} = 1$ נקבל ש־ $k \cdot \frac{1-0}{2} = 1 \Leftarrow k = 2$.
חשוב להכיר את הנוסחה לסכום סדרה הנדסית אחר ונפגוש אותה רבות במהלך הקורס

$$\sum_{i=1}^{n} a_1 q^i = a_1 \cdot \frac{q - q^n}{1 - q}, \quad \forall\ 0 \le q < 1, \tag{6}$$

וכמובן הסכום לסדרה אינסופית הינו

$$\sum_{i=1}^{\infty} a_1 q^i = a_1 \cdot \frac{q}{1 - q}, \quad \forall\ 0 \le q < 1. \tag{7}$$

הערה חשובה: שימו לב כי האיבר הראשון בסדרה בה דנו הוא $a_1 \cdot q^1$. במקרים רבים האיבר הראשון יהיה דווקא $a_1 q^0 = a_1$ ולכן נצטרך לחלק את הנוסחאות שקיבלנו ב־q. קרי,

$$\sum_{i=0}^{\infty} a_1 q^i = a_1 \cdot \frac{1}{1 - q}, \quad \forall\ 0 \le q < 1. \tag{8}$$

בכדי לפשט את הבעיה, תמיד תשתמשו בנוסחה הבאה:

$$\sum_{i=0}^{\infty} q^i = 1 + q + q^2 + \cdots = \frac{1}{1 - q}$$

ותכפילו את צמד האגפים באיבר הראשון לקבלת משוואה (8).

2. נגדיר מאורע $A_{even}=\{\omega\in\Omega:\ \omega\ is\ even\}$ ונחשב את $\Pr\left(A_{even}\right)$.

$$\begin{aligned}\Pr\left(A_{even}\right) &= \sum_{\omega\in A_{even}}\Pr\left(\omega\right)=\sum_{i=1}^{\infty}\Pr\left(2i\right)=\\ &= \sum_{i=1}^{\infty}\frac{2}{3^{2i}}=\sum_{i=1}^{\infty}\frac{2}{9^{i}}=2\cdot\frac{\frac{1}{9}}{1-\frac{1}{9}}=\frac{1}{4}\end{aligned}$$

כאשר השתמשנו בנוסחה (6) בכדי לחשב את הסכום.

3. נגדיר מאורע $A_{odd}=\{\omega\in\Omega:\ \omega\ is\ odd\}$ ונחשב את $\Pr\left(A_{odd}\right)$. נוכל לפתור באופן דומה לאופן בו פתרנו את סעיף ב', אך דרך יותר מהירה תהיה בעזרת משוואה (5). נשים לב ש־ $A_{odd}=A_{even}^{c}$ ולכן נובע ש־

$$\Pr\left(A_{odd}\right)=\Pr\left(A_{even}^{c}\right)=1-\Pr\left(A_{even}\right)=\frac{3}{4}\,.$$

טענה 1.10 יהיו $A,B\subseteq\Omega$ מאורעות כך ש־ $A\subseteq B$. אזי $\Pr\left(B\setminus A\right)=\Pr\left(B\right)-\Pr\left(A\right)$.

הוכחה: נשים לב כי $B=A\cup\left(B\setminus A\right)$ ושני המאורעות הללו הם מאורעות זרים. מכאן נובע ש־

$$\Pr\left(B\right)=\sum_{\omega\in A\cup\left(B\setminus A\right)}\Pr\left(\omega\right)=\sum_{\omega\in A}\Pr\left(\omega\right)+\sum_{\omega\in B\setminus A}\Pr\left(\omega\right)=\Pr\left(A\right)+\Pr\left(B\setminus A\right)\,,$$

כנדרש כאשר המעבר השני נובע מכך שהקבוצות זרות זו לזו. ■

מסקנה 1.11 נשים לב שכפועל יוצא מטענה 1.10 נובע שלכל $A,B\subseteq\Omega$ מאורעות כך ש־ $A\subseteq B$ מתקיים $\Pr\left(B\right)\geq\Pr\left(A\right)$.

משפט 1.12 יהיו $A_1,\ldots,A_n$ מספר סופי של מאורעות במרחב הסתברות $(\Omega,\Pr)$, אזי

$$\Pr\left(\bigcup_{i=1}^{n}A_i\right)\leq\sum_{i=1}^{n}\Pr\left(A_i\right)\,. \tag{9}$$

הוכחה: נשתמש בתכונת האדיטיביות ונקבל

$$\begin{aligned}\Pr\left(\bigcup_{i=1}^{n}A_i\right) &= \sum_{\omega\in\bigcup_{i=1}^{n}A_i}\Pr\left(\omega\right)\leq\\ &\leq\sum_{i=1}^{n}\sum_{\omega\in A_i}\Pr\left(\omega\right)=\\ &=\sum_{i=1}^{n}\Pr\left(A_i\right)\end{aligned}$$

כאשר המעבר השני נכון מאחר ואנחנו רק מוסיפים איברים בכך שאנחנו סוכמים על כל מאורע בנפרד. אם $\omega \in \bigcup_{i=1}^{n} A_i$ אזי קיים A_i כך ש־ $\omega \in A_i$ ולכן בפרט $\Pr(\omega)$ נמצא בסכום $\sum_{\omega \in A_i} \Pr(\omega)$. ■

הוכחה נוספת ויותר מדויקת למשפט זה נראה בהמשך כאשר נסקור את נוסחת ההכלה וההפרדה הכללית.

תרגיל 1.13 יהיו מאורעות $A_1, \ldots, A_n, B_1, \ldots, B_n \subseteq \Omega$ כך שלכל i מתקיים $A_i \subseteq B_i$. הראו כי

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

פתרון: נתחיל בלהוכיח ש־

$$\left(\bigcup_{i=1}^{n} B_i\right) \setminus \left(\bigcup_{i=1}^{n} A_i\right) \subseteq \bigcup_{i=1}^{n} (B_i \setminus A_i) .$$

אם $\omega \in \left(\bigcup_{i=1}^{n} B_i\right) \setminus \left(\bigcup_{i=1}^{n} A_i\right)$ אז $\omega \in \bigcup_{i=1}^{n} B_i \wedge \omega \notin \bigcup_{i=1}^{n} A_i$. לכן קיים B_i כך ש־ $\omega \in B_i$ ובנוסף לכל i מתקיים ש־ $\omega \notin A_i$. לכן $\omega \in B_i \setminus A_i$ ובפרט $\omega \in \bigcup_{i=1}^{n} (B_i \setminus A_i)$. בנוסף, נשים לב ש־$\left(\bigcup_{i=1}^{n} A_i\right) \subseteq \left(\bigcup_{i=1}^{n} B_i\right)$. לכן נקבל ש־

$$\begin{aligned}
\Pr\left(\bigcup_{i=1}^{n} B_i\right) - \Pr\left(\bigcup_{i=1}^{n} A_i\right) &= \Pr\left(\left(\bigcup_{i=1}^{n} B_i\right) \setminus \left(\bigcup_{i=1}^{n} A_i\right)\right) \leq \\
&\leq \Pr\left(\bigcup_{i=1}^{n} (B_i \setminus A_i)\right) \leq \\
&\leq \sum_{i=1}^{n} \Pr(B_i \setminus A_i) = \\
&= \sum_{i=1}^{n} \left(\Pr(B_i) - \Pr(A_i)\right) = \\
&= \sum_{i=1}^{n} \Pr(B_i) - \sum_{i=1}^{n} \Pr(A_i)
\end{aligned}$$

השיוויון הראשון נובע מטענה 1.10, האי־שוויון הראשון נובע מההכלה שהוכחנו בתחילת ההוכחה, האי־שוויון השני נובע מחסם האיחוד והשיוויון השני נובע ממסקנה 1.11.

תרגיל 1.14 תהיי $A_1, A_2, \ldots$ סדרה אינסופית של מאורעות. הוכיחו כי

$$\Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{N \to \infty} \Pr\left(\bigcup_{n=1}^{N} A_n\right) .$$

פתרון: נגדיר סדרת מאורעות חדשה

$$B_1 = A_1$$

$$.B_k = \bigcup_{n=1}^{k} A_n \setminus \bigcup_{n=1}^{k-1} A_n$$

ונשים לב ש־

$$\bigcup_{n=1}^{k} B_n = \bigcup_{n=1}^{k} A_n$$

ובנוסף שלכל $1 \leq i \neq j \in \mathbb{N}$ מתקיים $B_i \cap B_j = \phi$ (זה אוסף מאורעות זר). לכן

$$\begin{aligned}
\Pr\left(\bigcup_{n=1}^{\infty} A_n\right) &= \Pr\left(\bigcup_{n=1}^{\infty} B_n\right) = \\
&= \sum_{n=1}^{\infty} \Pr\left(B_n\right) = \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \Pr\left(B_n\right) = \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \Pr\left(\bigcup_{i=1}^{n} A_i \setminus \bigcup_{i=1}^{n-1} A_i\right) = \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \Pr\left(\bigcup_{i=1}^{n} A_i\right) - \Pr\left(\bigcup_{i=1}^{n-1} A_i\right) = \\
&= \lim_{N\to\infty} \Pr\left(\bigcup_{i=1}^{N} A_i\right)
\end{aligned}$$

כאשר השיוויון השני מתקיים מאחר ומדובר באיחוד זר, השיוויון החמישי נובע מהעובדה ש־ $\bigcup_{i=1}^{n-1} A_i \subseteq \bigcup_{i=1}^{n} A_i$ וטענה 1.10 והשיוויון האחרון נובע מכך שהטור הינו טור טלסקופי.

1.1.4 הבינום של ניוטון

נבצע רענון קצר בכל הקשור לקומבינטוריקה. נניח שיש לנו קבוצה A סופית בעלת n איברים, $A = \{a_1, \ldots, a_n\}$.

- מספר הסידורים האפשריים של הקבוצה: $n!$.
- מספר הדרכים לבחור k עצמים מתוך הקבוצה (ללא החזרה וללא חשיבות לסדר):

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

טענה 1.15 (הבינום של ניוטון) לכל $a, b \in \mathbb{R}$ ו־ $n \in \mathbb{N}$ מתקיים ש־

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

תרגיל 1.16 ענו על הסעיפים הבאים.

1. חשבו את הטור

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

2. הוכיחו כי

$$\sum_{\substack{k=1 \\ k \text{ is odd}}}^{n} \binom{n}{k} = \sum_{\substack{k=0 \\ k \text{ is even}}}^{n} \binom{n}{k} .$$

3. הוכיחו (ללא חישובים אלגבריים!) את הזהות

$$\sum_{k=1}^{n} \left(k \cdot \binom{n}{k} \right) = n \cdot 2^{n-1} .$$

פתרון: נענה על כל סעיף בנפרד.

1. נחשב בעזרת הבינום של ניוטון.

$$\sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} \cdot 1^k \cdot 1^{n-k} = (1+1)^n = 2^n .$$

2. נשים לב כי

$$(-1)^k = \begin{cases} 1 & k \text{ is even} \\ -1 & k \text{ is odd} \end{cases} .$$

ולכן

$$
\begin{aligned}
0 = (-1+1)^n &= \sum_{k=0}^{n} \binom{n}{k} \cdot (-1)^k \cdot 1^{n-k} = \\
&= \sum_{\substack{k=1 \\ k \text{ is odd}}}^{n} (-1)^k \binom{n}{k} + \sum_{\substack{k=0 \\ k \text{ is even}}}^{n} (-1)^k \binom{n}{k} = \\
&= (-1) \sum_{\substack{k=1 \\ k \text{ is odd}}}^{n} \binom{n}{k} + \sum_{\substack{k=0 \\ k \text{ is even}}}^{n} \binom{n}{k} \\
\Rightarrow 0 &= \sum_{\substack{k=0 \\ k \text{ is even}}}^{n} \binom{n}{k} - \sum_{\substack{k=1 \\ k \text{ is odd}}}^{n} \binom{n}{k}
\end{aligned}
$$

כנדרש.

3. נוכיח סעיף זה ללא חישובים אלגבריים כלל, אלא באמצעות סיפור! נניח שיש לנו n סטודנטים ואנו מעוניינים לבחור יו"ר לועד הסטודנטים ובנוסף לבחור ועדה מייעצת (בגודל כלשהו, ייתכן ותהיה ריקה). מה מספר הדרכים לבצע בחירה זו? מצד אחד, ינתן לבחור תחילה את היו"ר, n אפשרויות, ולאחר מכן לבחור ועדה מייעצת בגודל כלשהו. אם ניתן לבחור ועדה בכל גודל אפשרי מתוך $n-1$ הסטודנטים הנותרים יש לנו 2^{n-1} אפשרויות לכך (כמספרי תתי הקבוצות של קבוצה בגודל $n-1$). לכן ישנן $n \cdot 2^{n-1}$ דרכים לבצע את הבחירה. מצד שני, ניתן לבצע את הספירה באופן הבא. תחילה לבחור תת קבוצה בגודל כלשהו מ־1 עד n ולאחר מכן לבחור מתוכה את היו"ר,

$$\sum_{k=1}^{n} \binom{n}{k} \cdot k = n \cdot 2^{n-1}$$

הסכימה מתבצעת מאחר ובחירה של ועד בגודל מסויים לעומת בחירה של ועד בגודל אחר הן שתי בחירות שונות זו מזו ולכן אנחנו סוכמים על כל אפשרויות הבחירה.

תרגיל 1.17 הוכיחו כי

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

פתרון: נוכיח זאת ללא חישובים אלגבריים, אלא שוב באמצעות ספירת מצבים בסיטואציה ספציפית. מה מספר הדרכים לבחור ועד בגודל n מתוך קבוצה של n גברים ו־n נשים? מצד אחד, ניתן לספור בצורה ישירה ולקבל שמספר הדרכים הן $\binom{2n}{n}$. מצד שני, ניתן לבחון את

כמות הנשים שיכולות להיבחר במדגם. ייתכן וישנן n נשים בועד ואז מספר הדרכים לבחור הן $\binom{n}{n}\binom{n}{0}$, ייתכן וישנן רק $n-1$ נשים במדגם וגבר יחיד ואז מספר הדרכים לבצע בחירה זו הן $\binom{n}{n-1}\binom{n}{1}$ וכן הלאה. לכן אם נסכום על כל המצבים הזרים האפשריים שציינו נקבל שמספר הדרכים לבחור ועד שכזה הן

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

כאשר המעבר הראשון נובע מכך ש־ $\binom{n}{k}=\binom{n}{n-k}$ (בדקו זאת!).

1.2 מרחבי הסתברות אחידים

החלוקה הראשונה שנבצע עבור מרחבי הסתברות היא ־ מרחבי הסתברות אחידים (סימטריים) ומרחבי הסתברות לא אחידים (א־סימטריים).

הגדרה 1.18 מרחב הסתברות $(\Omega,\Pr)$ נקרא אחיד אם לכל $\omega\in\Omega$,

$$\Pr(\omega)=\frac{1}{|\Omega|}.\qquad(10)$$

ז"א, לכל איבר במרחב המדגם הסתברות שווה.

באופן משלים, מרחב הסתברות לא אחיד הוא מרחב הסתברות בו ישנו לפחות איבר אחד במרחב המדגם עם הסתברות שונה מהאיברים האחרים.

טענה 1.19 יהי $(\Omega,\Pr)$ מרחב הסתברות אחיד ו־$A\subseteq\Omega$ מאורע. אזי

$$\Pr(A)=\frac{|A|}{|\Omega|}.\qquad(11)$$

הוכחה: במרחב הסתברות אחיד לכל מאורע $A\subseteq\Omega$ מתקיים

$$\Pr(A)=\sum_{\omega\in A}\Pr(\omega)=\sum_{\omega\in A}\frac{1}{|\Omega|}=\frac{|A|}{|\Omega|},$$

כאשר המעבר הראשון נובע ממשוואה (4) והמעבר השני נובע ממשוואה (10). ■

1.3 פרמוטציות

תהי A קבוצה סופית כלשהי.

הגדרה 1.20 פרמוטציה (תמורה) σ היא העתקה חד־חד ערכית ועל מ־A לעצמה.

נניח כי $A=\{a_1,a_2,a_3\}$. דרך נוחה להציג את σ, פרמוטציה כלשהי של A היא

$$\sigma=\left(\begin{array}{ccc} a_1 & a_2 & a_3 \\ a_2 & a_1 & a_3 \end{array}\right).$$

זוהי העתקה אשר מעבירה את a_1 ל־ a_2, את a_2 ל־ a_1 ואת a_3 לעצמו. לעיתים נהוג לתת לכל אחד מאיברי הקבוצה מספר מ־1 עד $|A|$ ואז ניתן לרשום את הפרמוטציה באופן הבא:

$$\sigma=\left(\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 1 & 3 \end{array}\right),$$

כאשר האיבר a_i משוייך במקור למספר ה־i.

נניח כי A קבוצה סופית ולכן מספר הפרמוטציות עבור קבוצה עם $|A|$ איברים היא $|A|!$ (כמספר הדרכים לסדר $|A|$ עצמים שונים בשורה).

הערה 1.21 במידה ולא יצויין אחרת, אנו נניח מעתה כי A היא קבוצה סופית בעלת $n\in\mathbb{N}$ איברים ו־ σ היא פרמוטציה על קבוצה זאת.

הגדרה 1.22 בפרמוטציה σ קיים מעגל (מחזור) באורך $1\leq k\leq n$, $k\in\mathbb{N}$ אם קיימים k איברים שונים $a_{i_1},\ldots,a_{i_k}\in A$ כך ש־

$$\sigma\left(a_{i_j}\right)=a_{i_{j+1}}$$

לכל $1\leq j\leq k-1$ ו־ $\sigma\left(a_{i_k}\right)=a_{i_1}$.

זאת אומרת, שהאיבר a_{i_1} מועבר לאיבר $\sigma\left(a_{i_1}\right)$, האיבר $\sigma\left(a_{i_1}\right)$ מועבר לאיבר $\sigma^2\left(a_{i_1}\right)$ וכן הלאה, עד שלבסוף האיבר $\sigma^{k-1}\left(a_{i_1}\right)$ מועבר חזרה לאיבר a_{i_1} ובכך יצרנו מעגל / מחזור באורך k.

דוגמה 1.23 נראה מספר דוגמאות למעגלים בפרמוטציות שונות.

1. בפרמוטציה $\sigma=\left(\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 1 & 3 \end{array}\right)$ קיים מעגל באורך 2 משום ש־ $1\underset{\sigma}{\mapsto}2\underset{\sigma}{\mapsto}1$.
2. בפרמוטציה $\sigma=\left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{array}\right)$ ישנו מעגל באורך 4 מאחר ו־

$$1\underset{\sigma}{\mapsto}3\underset{\sigma}{\mapsto}2\underset{\sigma}{\mapsto}4\underset{\sigma}{\mapsto}1.$$

הגדרה 1.24 נקודת שבת בפרמוטציה σ היא מעגל באורך 1. קרי, ל־σ קיימת נקודת שבת אם קיים $a\in A$ כך ש־ $\sigma\left(a\right)=a$.

תרגיל 1.25 מה ההסתברות שבפרמוטציה σ שנבחרה באופן אחיד יש מעגל באורך n?

פתרון: מרחב המדגם שלנו בבעיה הוא קבוצת כל התמורות של $A=\{1,\ldots,n\}$,

$$\Omega=\{\sigma \text{ such that } \sigma : A\to A, \text{ is a permutation of } A\},$$

ובנוסף $|\Omega|=n!$. המאורע המבוקש, B, מוגדר באופן הבא:

$$B=\left\{\sigma :\ \forall i\ \forall j<n,\ \sigma^{j}\left(i\right)\neq i\right\}.$$

מצב זה הוא מאוד מיוחד מאחר ומעגל באורך n משמע שבכל נקודה בה נתחיל תמיד נחזור לאותה נקודה רק לאחר הפעלה של הפרמוטציה n פעמים. ננסה לרשום את B בצורה של קבוצת סדרות בכדי למצוא את $|B|$. נשים לב שכל פרמוטציה σ ב־ B נוכל לתאר ע"י סדרה באורך n, $(a_1,\ldots,a_n)$ כאשר $a_1=\sigma\left(1\right)$ ובנוסף $a_{i+1}=\sigma\left(a_i\right)$ לכל $2\leq i\leq n-1$. לכל $\sigma\in B$ קיימת הצגה אחת ויחידה כסדרה $(a_1,\ldots,a_n)$ וגם לחילופין לכל סדרה שכזאת קיימת פרמוטציה יחידה המתאימה לה ב־B ולכן מדובר בהעתקה חד־חד ערכית ועל. על כן, נוכל להשתמש בכך בכדי להציג את B באופן הבא:

$$B\ \ =\ \ \{(a_1,\ldots,a_n):\ a_i\in A\ \forall 1\leq i\leq n,\ a_n=1\},$$

כעת, נוכל לראות ש־$|B|=(n-1)!$ כמספר כל הסדרות של איברי A באורך nשנגמרות בספרה 1, לכן $\Pr\left(B\right)=\frac{(n-1)!}{n!}=\frac{1}{n}$.

תרגיל 1.26 מה ההסתברות שבפרמוטציה σ באורך $2n$ שנבחרה באופן אחיד יש בדיוק 2 מעגלים באורך n כל אחד?

פתרון: נסמן את המאורע המבוקש ב־ B. נשים לב שכל פרמוטציה σ ב־ B נוכל לתאר ע"י סדרה באורך $2n$, $(a_1,\ldots,a_{2n})$ כאשר $(a_1,\ldots,a_n)$ מתאר את המעגל הראשון ו־ $(a_{n+1},\ldots,a_{2n})$ מתאר את המעגל השני. כעת, נוכל לראות שגודל המאורע B הוא -

$$|B|=\frac{1}{2}\binom{2n}{n}\left(n-1\right)!\left(n-1\right)!$$

כאשר $\frac{1}{2}\binom{2n}{n}$ זה מספר הדרכים לחלק את הקבוצה המקורית ל־2 קבוצות שוות גודל (בוחרים n איברים לקבוצה הראשונה והיתר נכנסים לקבוצה השנייה; החלוקה ב־2 היא מאחר וסדר הקבוצות לא רלוונטי) ובנוסף מספר הדרכים ליצור מעגל באורך n הוא $n-1!$ לכל קבוצת איברים, לכן

$$\Pr\left(B\right)=\frac{\frac{1}{2}\binom{2n}{n}\left(n-1\right)!\left(n-1\right)!}{2n!}=\frac{2n!\left(n-1\right)!\left(n-1\right)!}{2\left(n!n!2n!\right)}=\frac{1}{2n^2}.$$

1.3.1 הכלה והפרדה

חסם האיחוד אשר למדנו קודם לכן נתן לנו חסם על ההסתברות של איחוד מאורעות כלשהו. לעיתים, החסם הזה יכול להיות טריוויאלי ולא רלוונטי, ולכן נרצה למצוא דרך לחשב את ההסתברות לאיחוד של מאורעות באופן מדוייק. הנוסחה המשמשת אותנו לטובת עניין זה נקראת נוסחת הכלה והפרדה. נתחיל במקרה הפשוט המתייחס לאיחוד של 2 מאורעות.

טענה 1.27 (נוסחת הכלה והפרדה עבור 2 מאורעות) יהי מרחב הסתברות $(\Omega, \Pr)$ ו־ 2 מאורעות $A, B \subseteq \Omega$ כלשהם. אזי,

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

הוכחה: נשים לב כי הקבוצות $B, A \setminus B$ הן זרות ובנוסף ש־ $(A \setminus B) \cup B = A \cup B$ ולכן

$$.\Pr(A \cup B) = \sum_{\omega \in (A \setminus B) \cup B} \Pr(\omega) = \sum_{\omega \in A \setminus B} \Pr(\omega) + \sum_{\omega \in B} \Pr(\omega) = \Pr(A \setminus B) + \Pr(B)$$

לכן נקבל ש־

$$.\Pr(A \setminus B) = \Pr(A \cup B) - \Pr(B) \tag{12}$$

באופן דומה נקבל ש־ $A \cap B, A \setminus B$ הן זרות והאיחוד שלהן מקיים $(A \cap B) \cup (A \setminus B) = A$ לכן

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \setminus B) \\ &\Rightarrow P(A \setminus B) = P(A) - P(A \cap B) \end{aligned} \tag{13}$$

מהשוואת צמד המשוואות (13) ו־ (12) נקבל את הנוסחה הנתונה. ■

כעת נוכל לעבור למקרה הכללי אשר יוצג ללא הוכחה (ההוכחה מתבצעת באינדוקציה על מספר המאורעות תוך שימוש במקרה הפרטי של $n = 2$ אשר הוכחנו בטענה 1.27).

משפט 1.28 (נוסחת הכלה והפרדה עבור n מאורעות) יהי מרחב הסתברות (Ω, P) ו־$n \in \mathbb{N}$ מאורעות $A_1, \ldots, A_n \subseteq \Omega$ כלשהם. אזי,

$$.\Pr\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{\phi \neq I \subseteq \{1, \ldots, n\}} (-1)^{|I|+1} \Pr\left(\bigcap_{i \in I} A_i\right)$$

המקרים בהם נשתמש לרוב הן כאשר $n = 2$ או לחילופין כאשר $n = 3$:

$$\begin{aligned} .\Pr\left(\bigcup_{i=1}^{3} A_i\right) &= \Pr(A_1) + \Pr(A_2) + \Pr(A_3) - \Pr(A_1 \cap A_2) \\ &- \Pr(A_1 \cap A_3) - \Pr(A_2 \cap A_3) + \Pr(A_1 \cap A_2 \cap A_3) \end{aligned}$$

תרגיל 1.29 10 זוגות מתיישבים באקראי סביב שולחן עגול. מה ההסתברות שאף זוג לא יושב יחד (כיסאות סמוכים)?

פתרון: נגדיר את המאורע B־ אף זוג לא יושב יחד ומאורעות שמרכיבים את מרחב המדגם $A_i,\ i=1,\ldots,10$ ־ הזוג ה־i יושב יחדיו. מאחר ומדובר בסידור אנשים במעגל, גודל מרחב המדגם שלנו הוא $|\Omega|=19!$ (מיקום האדם הראשון איננו רלוונטי מאחר וישנה סימטרית סיבוב). אם נגדיר, לדוגמא, את a_i להיות השם של האדם ה־i מתוך 20 האנשים המרכיבים את הזוגות אזי

$$.\Omega=\{(a_1,\ldots,a_{19}):\ a_i=person\ in\ seat\ i\ w.r.t\ seat\ of\ a_{20}\}$$

אנחנו מעוניינים לחשב את $\Pr(B)$ ולפי כללי דה מורגן נובע ש־

$$.\Pr(B)=\Pr\left(\bigcap_{i=1}^{10}A_i^c\right)=1-\Pr\left(\bigcup_{i=1}^{10}A_i\right)$$

נשים לב שיש סימטריה בין הזוגות, בכך ש־$\Pr(A_i)=\Pr(A_j)$ לכל $1\le i\le j\le 10$ ובאופן כללי

$$.\Pr\left(A_{i_1}\cap\cdots\cap A_{i_k}\right)=\Pr\left(A_{j_1}\cap\cdots\cap A_{j_k}\right)\ \forall\, k=1,\ldots,10,\ i_1<\cdots<i_k,j_1<\cdots<j_k$$

נשתמש שוב בנוסחת הכלה והפרדה ל־10 מאורעות ונקבל,

$$\begin{aligned}\Pr(B) &= 1-\Pr\left(\bigcup_{i=1}^{10}A_i\right)=\\ &= 1-\sum_{\phi\neq I\subseteq\{1,\ldots,n\}}(-1)^{|I|+1}\Pr\left(\bigcap_{i\in I}A_i\right)=\\ &= 1-\sum_{k=1}^{10}(-1)^{k+1}\binom{10}{k}\Pr\left(A_1\cap\cdots\cap A_k\right)\end{aligned}$$

כאשר המעבר בין השורה השנייה לשלישית נובע מכך שיש סימטריה בין המאורעות ולכן ניתן להסתכל על k המאורעות הראשונים בלבד ($1\le k\le 10$) ולהכפיל במספר הדרכים לבחור k מאורעות מתוך 10. נותר לנו לחשב את $\Pr(A_1\cap\cdots\cap A_k)$, הסתברות המאורע בו k הזוגות הראשונים יושבים יחדיו. לשם כך, נתייחס לכל זוג מהזוגות המדוברים כיחידה אחת בסידור הכללי ונסדר את כלל הגורמים, כולל סידורים פנימיים:

$$\Pr\left(A_1\cap\cdots\cap A_k\right)=\frac{|A_1\cap\cdots\cap A_k|}{|\Omega|}=\frac{2^k\cdot(19-k)!}{19!}$$

כאשר המעבר האחרון בוצע בכך שישנן $20-k$ גופים לסדר במעגל (סה"כ האנשים פחות מספר הזוגות שאוחדו ליחידה אחת) והכפלה במספר הסידורים הפנימיים של כל זוג בפני עצמו.

$$\begin{aligned}P(B) &= 1-\sum_{k=1}^{10}(-1)^{k+1}\binom{10}{k}P\left(A_1\cap\cdots\cap A_k\right)=\\ &= 1+\sum_{k=1}^{10}(-1)^{k}\binom{10}{k}\frac{2^k\cdot(19-k)!}{19!}\end{aligned}$$

תרגיל 1.30 (בעיית המעטפות והמכתבים) ישנן n מעטפות ו־ n מכתבים, כאשר כל מכתב מתאים למעטפה ספציפית. המזכיר לא ידע כי לכל מכתב קיימת מעטפה ספציפית ולכן הכניס מכתבים למעטפות באקראי. מה הסיכוי שלא יהיה אפילו מכתב אחד שיגיע ליעדו?

הערה 1.31 שימו לב שאם אנו מניחים שהמעטפות ממוספרות מ־ 1 עד n וכן המכתבים ממוספרים באותו אופן (לכל מכתב ומעטפה מתאימים ישנו אותו המספר), אזי השאלה בעצם היא מה הסיכוי שלא תהיה נקודת שבת בפרמוטציה האקראית σ שבוחר במזכיר.

פתרון: נתחיל בהגדרת המאורע B־ אף מכתב לא הגיע ליעדו ומאורעות שמרכיבים את מרחב המדגם $A_i,\ i=1,\ldots,n$ ־ המכתב ה־i הגיע ליעדו. מרחב המדגם שלנו מורכב מכל הפרמוטציות σ על הקבוצה $\{1,\ldots,n\}$ כאשר כל מכתב מציין מספר והמעטפות מסודרות לפי הסדר המתאים מ־1 עד n, לכן $A_i=\{\sigma:\ \sigma(i)=i\}$. נשים לב ש־ $B=\bigcap_{i=1}^{n}A_i^c$ ולכן

$$\Pr(B)=\Pr\left(\bigcap_{i=1}^{n}A_i^c\right)=1-\Pr\left(\left(\bigcap_{i=1}^{n}A_i^c\right)^c\right)=1-\Pr\left(\bigcup_{i=1}^{n}A_i\right),$$

כאשר המאורע $\bigcup_{i=1}^{n}A_i$ הוא המאורע בו הגיע לפחות מכתב 1 ליעדו והוא גם המאורע המשלים ל־B (חשוב להבין זאת גם באופן אינטואיטיבי, ולא רק דרך כללי דה־מורגן). מאחר וישנה סימטריה בין המכתבים (הוכנסו באקראי למעטפות), אנו יודעים כי $\Pr(A_i)=\Pr(A_j)$ לכל $1\le i\le j\le n$. נשתמש בנוסחת הכלה והפרדה ונקבל,

$$\begin{aligned}
\Pr(B) &= 1-\Pr\left(\bigcup_{i=1}^{n}A_i\right)= \\
&= 1-\sum_{\phi\neq I\subseteq\{1,\ldots,n\}}(-1)^{|I|+1}\Pr\left(\bigcap_{i\in I}A_i\right)= \\
&= 1-\sum_{k=1}^{n}(-1)^{k+1}\binom{n}{k}\Pr(A_1\cap\cdots\cap A_k)= \\
&= 1-\sum_{k=1}^{n}(-1)^{k+1}\binom{n}{k}\frac{1\cdot(n-k)!}{n!}= \\
&= 1+\sum_{k=1}^{n}\frac{(-1)^k}{k!}= \\
&= \sum_{k=0}^{n}\frac{(-1)^k}{k!}
\end{aligned}$$

כאשר המעבר בין השורה השנייה לשלישית נובע מכך שיש סימטריה בין המאורעות ולכן ניתן להסתכל על k המאורעות הראשונים בלבד ($1\le k\le n$) ולהכפיל במספר הדרכים לבחור k מאורעות מתוך n, והמעבר שמגיע לאחר מכן הוא חישוב קומבינטורי ישיר של מספר הדרכים ש־k המעטפות הראשונות יוכנסו למעטפות שלהן תוך חלוקה בגודל מרחב המדגם $n!$.

תרגיל 1.32 תהיי קבוצה $A=\{1,\dots,n\}$ ונבחר פרמוטציה σ של A באקראי. מה הסיכוי שיש בדיוק נקודת שבת אחת ב־σ?

פתרון: מרחב המדגם שלנו בבעיה הוא קבוצת כל התמורות של A,

$$\Omega=\{\sigma:\ \sigma:A\to A,\ permutation\ of\ A\},$$

ובנוסף $|\Omega|=n!$. נגדיר את המאורע A_i להיות המאורע בו ישנה נקודת שבת בערך ה־i, $A_i=\{\sigma:\ \sigma(i)=i\}$. אנחנו מעוניינים לחשב את המאורע B ־ יש בדיוק נקודת שבת אחת ב־σ,

$$B=\{\sigma:\ \exists 1\le i\le n:\ \sigma(i)=i\ \wedge\ \sigma(j)\neq j\ \forall j\neq i\}.$$

נשים לב ש־

$$B=\bigcup_{i=1}^{n}\left(A_i\cap\left(\bigcap_{j\neq i}A_j^c\right)\right)$$

ומאחר והאיחוד הוא איחוד זר ובנוסף קיימת סימטריה מלאה בין המאורעות, נובע ש־

$$\begin{aligned}\Pr(B) &= \sum_{i=1}^{n}\Pr\left(A_i\cap\left(\bigcap_{j\neq i}A_j^c\right)\right)=\\ &= \binom{n}{1}\Pr\left(A_1\cap A_2^c\cap\cdots\cap A_n^c\right)=\\ &= \binom{n}{1}\frac{1\cdot|A_2^c\cap\cdots\cap A_n^c|}{n!}\end{aligned}$$

המאורע $A_1\cap A_2^c\cap\cdots\cap A_n^c$ שקיבלנו זה בדיוק המאורע בו אין נקודות שבת בפרמוטציה של הקבוצה $A_2^c\cap\cdots\cap A_n^c$ והנקודה 1 מועברת לעצמה. נזכור כי פתרנו כבר את המקרה בו אין כלל נקודות שבת בבעיית המכתבים והמעטפות ואנו יודעים כי עבור פרמוטציה של $n-1$ אובייקטים ההסתברות שלא תהיה נקודות שבת היא $\sum_{k=0}^{n-1}\frac{(-1)^k}{k!}$ לכן

$$|A_2^c\cap\cdots\cap A_n^c|=(n-1)!\left(\sum_{k=0}^{n-1}\frac{(-1)^k}{k!}\right),$$

ולסיכום

$$\begin{aligned}\Pr(B) &= \binom{n}{1}\frac{1\cdot(n-1)!}{n!}\left(\sum_{k=0}^{n-1}\frac{(-1)^k}{k!}\right)=\\ &= \sum_{k=0}^{n-1}\frac{(-1)^k}{k!}\end{aligned}$$

תרגיל 1.33 תהיי קבוצה $A=\{1,\ldots,2n\}$ ונבחר פרמוטציה σ של A באקראי. מה הסיכוי שיש ב־σ מעגל באורך n בדיוק?

פתרון: מרחב המדגם שלנו בבעיה הוא קבוצת כל התמורות של A,

$$\Omega=\{\sigma:\ \sigma:A\to A,\ permutation\ of\ A\},$$

ובנוסף $|\Omega|=2n!$. מספר הדרכים לבחור n איברים שונים שירכיבו מעגל באורך n הוא $\binom{2n}{n}$. נגדיר את הקבוצה $\mathcal{I}=\{I\subseteq\{1,\ldots,2n\}:\ |I|=n\}$. לכל אפשרות $I\in\mathcal{I}$ נגדיר את המאורע A_I להיות המאורע בו ישנו מעגל שכולל את האיברים שנבחרו בתת קבוצה I. אנחנו מעוניינים לחשב את המאורע B -

$$B=\bigcup_{I\in\mathcal{I}}A_I.$$

נניח בלי הגבלת הכלליות שלכל $I\in\mathcal{I}$ מתקיים ש־ A_{I^c} הוא המאורע בו יש מעגל באורך n על קבוצת הקודקודים המשלימה של I. נשים לב ש־ $\Pr(A_I\cap A_J)=\Pr(\phi)=0$, אלא אם כן $I=J$ או לחילופין $J=I^c$. לכן

$$\begin{aligned}
\Pr(B) &= \Pr\left(\bigcup_{I\in\mathcal{I}}A_I\right)= \\
&= \sum_{\phi\neq L\subseteq\mathcal{I}}(-1)^{|L|+1}\Pr\left(\bigcap_{I\in L}A_I\right)= \\
&= \sum_{I\in\mathcal{I}}(-1)^{1+1}\Pr(A_I)+\sum_{I\in\mathcal{I}}(-1)^{2+1}\frac{1}{2}\Pr(A_I\cap A_{I^c})= \\
&= \sum_{I\in\mathcal{I}}\Pr(A_i)-\frac{1}{2}\sum_{I\in\mathcal{I}}\Pr(A_I\cap A_{I^c})= \\
&= \frac{\binom{2n}{n}(n-1)!n!}{2n!}-\frac{\frac{1}{2}\binom{2n}{n}(n-1)!(n-1)!}{2n!}= \\
&= \frac{1}{n}-\frac{1}{2n^2}=\frac{1}{n}\left(1-\frac{1}{2n}\right)
\end{aligned}$$

כאשר השיוויון השלישי נובע מכך שהחיתוך של המאורעות הוא קבוצה ריקה למעט המקרים שצויינו קודם לכן, כמו כן בסכום השני יש פקטור $\frac{1}{2}$ מאחר ולכל קבוצה A_I צריך להתאים רק את המאורע המשלים לה ובמקרה של הסכום שמציין לעיל, אנחנו בחרנו בנפרד בכל פעם גם את הקבוצה וגם את המשלים שלה כשתי אופציות שונות.

1.4 הילוך מקרי

נניח כי אנו מגרילים באופן אחיד סדרה $(S_0,\ldots,S_n)$ כאשר $S_0=0$, $S_i\in\mathbb{Z}$ לכל איבר S_i, $0\leq i\leq n$ בסדרה ובנוסף מתקיים $S_i-S_{i-1}\in\{-1,1\}$ לכל $1\leq i\leq n$. ז"א, אנחנו, למעשה, מגרילים סדרה של ערכים שההפרשים בין כל שני איברים עוקבים בסדרה הוא 1 או -1 כאשר האיבר הראשון הינו תמיד 0. סדרה שכזאת נקראת הילוך מקרי. בהמשך

הקורס אנחנו נראה דרכים נוספות להגדיר הילוך מקרי, אשר מבוססות בפועל על הגרלה של סדרות ערכים של -1 או 1 (כמובן שמדובר במצב שקול למצב המתואר לעיל).

כמה סדרות כאלה ישנן? ובכן, מאחר ולמעשה אנחנו בכל שלב בוחרים את התנועה של ההילוך כלפי מעלה או מטה במרווח של 1 בדיוק, אזי ישנם 2^n אפשרויות להגריל את n הצעדים הללו. ההילוך מקרי, או במקרה הגבולי הנקרא תנועה בראונית בו $n\to\infty$ וצפיפות הצעדים שואפת לאפס, הם אלמנטים שימושיים ביותר מאחר ונעזרים בהם רבות במודלים שונים בכלכלה ובכלל (תמחור אופציות וכו').

תרגיל 1.34 יהי הילוך מקרי פשוט $(S_0,\dots,S_n)$ שנבחר באקראי. מה ההסתברות ש־ $S_n=k$ עבור $k\in\mathbb{Z}$ כלשהו?

פתרון: נתחיל בפיצול למקרים. המקרה הפשוט הוא המקרה בו $|k|>n$, קרי $k>n$ או $k<-n$. מצב זה כמובן אינו אפשרי ולכן

$$.\Pr\left(S_n=k\right)=\begin{cases}0 & k>n\ \vee\ k<-n\\ ? & -n\le k\le n\end{cases}$$

נניח מעתה והלאה כי $k\ge 0$ ומשיקולי סימטריה עבור כל $k<0$ נקבל ש־

$$.\Pr\left(S_n=k\right)=\Pr\left(S_n=-k\right)$$

נפצל למקרים שוב - $n-k$ מספר זוגי או $n-k$ מספר אי־זוגי.

- אם $n-k$ הוא מספר זוגי אזי ההילוך היה צריך לבצע a צעדים לכיוון מטה ו־ $a+k$ לכיוון מעלה. לכן נקבל ש - $n=2a+k$ ואז $a=\frac{n-k}{2}$. נסיק כי

$$\Pr\left(S_n=k\right)=\frac{\binom{n}{\frac{n-k}{2}}\binom{\frac{n+k}{2}}{\frac{n+k}{2}}}{2^n}=\frac{\binom{n}{\frac{n-k}{2}}}{2^n}$$

כאשר במונה יש את מספר הדרכים לבצע $a=\frac{n-k}{2}$ צעדים לכיוון מטה והיתר כמובן יהיו לכיוון מעלה.

- אם $n-k$ הוא מספר אי־זוגי אזי נתקל בבעיה מאחר ו־ $n=2a+k \Leftarrow n-k=2a$ וזה אינו אפשרי. הסיבה לכך היא פשוטה, אם נניח ויש בהילוך המקרי 10 צעדים, $n=10$, אזי המיקום של הצעד האחרון S_n חייב להיות זוגי גם כן. באותה מידה, אם מתבצעים מספר אי־זוגי של צעדים אזי המיקום הסופי חייב להיות אי־זוגי (כל זאת תחנת ההנחה שיוצאים מנקודת האפס). לכן לסיכום:

$$.\Pr\left(S_n=k\right)=\begin{cases}\frac{\binom{n}{\frac{n-|k|}{2}}}{2^n} & n-|k|\ \ is\ even,\ |k|\le n\\ 0 & otherwise\end{cases}$$

תרגיל 1.35 בכד יש 8 כדורים: 3 כחולים, 2 אדומים, 2 ירוקים ו־1 צהוב. בוחרים n כדורים עם החזרה. מה הסיכוי שכל הצבעים מופיעים במדגם? אנו מניחים ש־ $n\ge 4$.

פתרון: נשים לב שמרחב המדגם הוא $|\Omega| = 8^n$ (ניתן לחשוב על הכדורים כאילו הם ממסופרים מ־1 עד 8 ואז מרחב המדגם יותר קל להבנה כסדרות של מספרים). נגדיר מאורעות A_1 ־ הצבע הכחול אינו מופיע, A_2 ־ הצבע האדום אינו מופיע, A_3 ־ הצבע הירוק אינו מופיע ולבסוף A_4 ־ הצבע הצהוב אינו מופיע. המאורע המבוקש B מקיים ש־

$$\begin{aligned}
B &= \bigcap_{n=1}^{4} A_i^c \\
\Pr(B) &= \Pr\left(\bigcap_{n=1}^{4} A_i^c\right) = \\
&= 1 - \Pr\left(\left(\bigcap_{n=1}^{4} A_i^c\right)^c\right) = \\
&= 1 - \Pr\left(\bigcup_{n=1}^{4} A_i\right) = \\
&= \sum_{\phi \neq I \subseteq \{1,2,3,4\}} (-1)^{|I|+1} \Pr\left(\bigcap_{i \in I} A_i\right) = \\
&= \sum_{i=1}^{4} \Pr(A_i) - \sum_{i<j} \Pr(A_i \cap A_j) + \\
&\quad + \sum_{i<j<k} \Pr(A_i \cap A_j \cap A_k) - 0
\end{aligned}$$

נחשב כל איבר בנפרד.

$$\begin{aligned}
\Pr(A_1) &= \frac{5^n}{8^n} \\
\Pr(A_2) = \Pr(A_3) &= \frac{6^n}{8^n} \\
\Pr(A_4) &= \frac{7^n}{8^n} \\
\Pr(A_1 \cap A_2) = \Pr(A_1 \cap A_3) &= \frac{3^n}{8^n}, \\
\Pr(A_1 \cap A_4) &= \frac{4^n}{8^n} \\
\Pr(A_2 \cap A_4) = \Pr(A_3 \cap A_4) &= \frac{5^n}{8^n} \\
\Pr(A_2 \cap A_3) &= \frac{4^n}{8^n} \\
\Pr(A_1 \cap A_2 \cap A_4) = \Pr(A_1 \cap A_3 \cap A_4) &= \frac{2^n}{8^n} \\
\Pr(A_2 \cap A_3 \cap A_1) &= \frac{1^n}{8^n} \\
\Pr(A_2 \cap A_3 \cap A_4) &= \frac{3^n}{8^n}
\end{aligned}$$

לכן לסיכום,

$$\begin{aligned}
\Pr\left(B\right) &= \sum_{i=1}^{4}\Pr\left(A_i\right)-\sum_{i<j}\Pr\left(A_i\cap A_j\right)+\sum_{i<j<k}\Pr\left(A_i\cap A_j\cap A_k\right)-0= \\
&= \frac{5^n+2\cdot 6^n+7^n-2\cdot 3^n-2\cdot 4^n-2\cdot 5^n+2\cdot 2^n+3^n+1}{8^n}= \\
&= \frac{7^n+2\cdot 6^n-5^n-2\cdot 4^n-3^n+2\cdot 2^n+1}{8^n}
\end{aligned}$$

2 הסתברות מותנה ואי־תלות

ישנן סיטואציות בהן אנחנו מקבלים מידע נוסף ולכן נדרשים לעדכן את מרחב ההסתברות שלנו בהתאם למידע זה. הסתברות מותנה זו הסתברות אשר עודכנה בהתאם למידע / מאורע ספציפי.

הגדרה 2.1 יהי מרחב הסתברות $(\Omega, \Pr)$ ומאורע בעל הסתברות חיובית $B \subseteq \Omega$, $\Pr(B) > 0$. מרחב ההסתברות המותנה $(\Omega, \Pr(\cdot|B))$ מוגדר לכל $\omega \in \Omega$ ע"י

$$\Pr(\omega|B) \begin{cases} \frac{\Pr(\omega)}{\Pr(B)} & \omega \in B \\ 0 & \omega \notin B \end{cases}.$$

מסקנה ישירה מהגדרה 2.1 היא:

מסקנה 2.2 לכל מאורע $A \subseteq \Omega$ מתקיים ש־

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

בנוסף, פונקצית ההסתברות המותנה $\Pr(\cdot|B)$ היא בעצמה פונקצית הסתברות, קרי היא מקיימת את כל התכונות ההכרחיות של פונקצית הסתברות:

1. **אי־שלילית**: $1 \leq \Pr(\omega|B) \leq 0$ לכל $\omega \in \Omega$. תכונה זו נובעת ישירות מכך שההסתברות היא פונקציה אי־שלילית ולכן $\Pr(A|B) = \frac{\Pr(A\cap B)}{\Pr(B)} \geq 0$ כמנה של ערכים אי־שליליים ומכנה חיובי.
2. **נרמול**: נשים לב כי $\Pr(\Omega|B) = \frac{\Pr(\Omega\cap B)}{\Pr(B)} = \frac{\Pr(B)}{\Pr(B)} = 1$.

על כן פונקצית ההסתברות המותנה $\Pr(\cdot|B)$, לכל מאורע בעל הסתברות חיובית $\Pr(B) > 0$, $B \subseteq \Omega$, היא בעצמה פונקצית הסתברות ולכן כל מה שהוכחנו עד כה עבור פונקצית הסתברות תקף באותה המידה להסתברות מותנה.

הערה 2.3 מעתה בכל פעם שנתייחס להסתברות מותנה, נניח כי ההתנייה היא על מאורע כלשהו בעל הסתברות חיובית.

תרגיל 2.4 בכד 10 כדורים הממוספרים מ־1 עד 10. שולפים 4 ללא החזרה.

1. מה הסיכוי שהכדור עם המספר 7 בחוץ?
2. ידוע ש־9 בחוץ, מה הסיכוי שגם 7 בחוץ?
3. ידוע שאין מספר קטן מ־4 בחוץ, מה הסיכוי ש־7 בחוץ?
4. מה הסיכוי ש־7 לא בחוץ אם ידוע שאין מספר קטן מ־4 בחוץ?

פתרון: נתחיל בהגדרת מרחב המדגם בבעיה - ישנם 10 כדורים ומוציאים 4 ללא החזרה ולכן מרחב המדגם מורכב מכל האפשרויות לבחור 4 כדורים מתוך 10, סה"כ $|\Omega| = \binom{10}{4}$.

1. נגדיר את המאורע - "הכדור עם המספר 7 הינו אחד מהכדורים שהוצאו" בתור A. מאחר ומדובר במרחב מדגם שווה הסתברות, אזי

$$\Pr(A) = \frac{1 \cdot \binom{9}{3}}{\binom{10}{4}} = \frac{9!6!4!}{10!6!3!} = \frac{4}{10} = \frac{2}{5}.$$

דרך נוספת להסתכל על זה היא פשוט לחשוב על הוצאות מכד כמו סידורים בשורה. עבור כל הוצאה מכד יש סידור בשורה ולהיפך. לכן ישנם $10!$ סידורים בשורה ואותנו מעניין רק מי נמצא ב־4 המקומות הראשונים ומי לא. משיקולי סימטריה הסיכוי ש־7 יהיה במקום הראשון בסידור בשורה הוא 0.1, ואותו הדבר גם עבור המקום השני והשלישי וכן הלאה ולכן סה"כ ההסתברות שהכדור הזה יוצא היא $4 \cdot \frac{1}{10} = 0.4$.

2. נסמן ב־B את המאורע המציין כי הכדור עם המספר 9 הינו אחד מהכדורים שהוצאו.

$$\begin{aligned}\Pr(A|B) &= \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\frac{1\cdot\binom{8}{2}}{\binom{10}{4}}}{\frac{1\cdot\binom{9}{3}}{\binom{10}{4}}} \\ &= \frac{\binom{8}{2}}{\binom{9}{3}} = \frac{8!6!3!}{9!6!2!} = \frac{1}{3}\end{aligned}$$

האם היינו יכולים לדעת זאת מראש? כן! האינטואיציה היא מאוד פשוטה. נתון שהכדור עם המספר 9 יצא כבר ולכן נותרו 3 כדורים שצריך להוציא מתוך 9 נותרים, הסיכוי ש־7 הוא אחד מהם הוא בדיוק $\frac{1}{3}$ (ע"ב אותו רעיון של סידור בשורה שראינו קודם לכן).

3. נגדיר מאורע C - כל המספרים שהוצאו גדולים או שווים ל־4. במידה והיינו צריכים לענות אינטואיטיבית אז היינו יכולים להניח שכעת הכדורים נבחרים בהסתברות שווה מתוך הערכים $4, \ldots, 10$ ולכן הסיכוי להוציא את הכדור עם המספר 7 הוא $\frac{4}{7}$ (הסיכוי ש־7 נמצא בארבעת המקומות הראשונים בסידור בשורה של המספרים $4, \ldots, 10$). נחשב זאת מפורשות:

$$\begin{aligned}\Pr(A|C) = \frac{\Pr(A \cap C)}{\Pr(C)} = \frac{\frac{1\cdot\binom{6}{3}}{\binom{10}{4}}}{\frac{1\cdot\binom{7}{4}}{\binom{10}{4}}} = \\ = \frac{\binom{6}{3}}{\binom{7}{4}} = \frac{6!4!3!}{7!3!3!} = \frac{4}{7}\end{aligned}$$

4. אנו נדרשים לחשב כעת את ההסתברות $\Pr(A^c|C)$. נזכור כי הסתברות מותנה מקיימת את כל הכללים של הסתברות ולכן

$$\Pr(A^c|C) = 1 - \Pr(A|C) = 1 - \frac{4}{7} = \frac{3}{7}.$$

טענה 2.5 (כלל הכפל) יהיו מאורעות $A_1,\ldots,A_n$ במרחב הסתברות $(\Omega,\Pr)$ ונניח כי $\Pr\left(\bigcap_{i=1}^{n}A_i\right)>0$. על כן

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

שימו לב שההוכחה עבור המקרה בו $n=2$ היא מיידית מההגדרה של הסתברות מותנה וניתן להכליל תוצאה זו עבור n מאורעות באינדוקציה.

משפט 2.6 (נוסחת ההסתברות השלמה) תהיי $B_1,\ldots,B_n$ חלוקה של מרחב המדגם למאורעות זרים בעלי הסתברות חיובית כ"א. ז"א $B_i\cap B_j=\phi$ לכל $1\le i<j\le n$, $\bigcup_{i=1}^{n}B_i=\Omega$ ובנוסף לכל $i=1,\ldots,n$ מתקיים ש־ $\Pr\left(B_i\right)>0$. אזי לכל מאורע $A\subseteq\Omega$ מתקיים

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

האינטואיציה מאחורי נוסחת ההסתברות השלמה היא די פשוטה. אנחנו רוצים לדעת את ההסתברות למאורע A כלשהו ויכולים להתרחש מספר מצבים בעולם שלנו $B_1,\ldots,B_n$.B_i התרחש, נבדוק את הסבירות של מאורע A תחת הנחה זו ונכפיל באפשרות שאכן המצב הוא B_i. במידה ונעשה זאת עבור כל מצב שונה בנפרד נקבל בסה"כ את ההסתברות של A. אפשר לחשוב על זה כמו עץ, תחילה לכל ענף / מצב ישנה הסתברות מסויימת ואז ישנה הסתברות למאורע להתרחש תחת אותו מצב / ענף בו נמצאים.

טענה 2.7 יהיו צמד מאורעות C,B במרחב הסתברות כך ש־ $\Pr\left(B\cap C\right)>0$. מרחב ההסתברות המותנה בו מתנים תחילה על B ולאחר מכן על C הוא $(\Omega,\Pr\left(\cdot|B\cap C\right))$.

הוכחה: יהי מאורע $A\subseteq\Omega$. כפי שכבר הוכחנו, $\Pr\left(\cdot|C\right)$ זו פונקצית התפלגות כלשהי, לכן נסמן אותה ב־$Q\left(\cdot\right)=\Pr\left(\cdot|C\right)$. מאחר ו־$Q$ היא פונקצית התפלגות נוכל להשתמש בהגדרת מרחב הסתברות מותנה עבורה ונקבל כי

$$Q\left(A|B\right)=\frac{Q\left(A\cap B\right)}{Q\left(B\right)}.$$

נשתמש בהגדרה של Q ונקבל

$$Q\left(A|B\right)=\frac{Q\left(A\cap B\right)}{Q\left(B\right)}=\frac{\Pr\left(A\cap B|C\right)}{\Pr\left(B|C\right)}=\frac{\frac{\Pr(A\cap B\cap C)}{\Pr(C)}}{\frac{\Pr(B\cap C)}{\Pr(C)}}=\frac{\Pr\left(A\cap B\cap C\right)}{\Pr\left(B\cap C\right)},$$

כאשר השיוויון השלישי נובע מהגדרת הסתברות מותנה גם כן. לסיכום קיבלנו ש־

$$\Pr\left(\left(A|B\right)|C\right)=Q\left(A|B\right)=\frac{\Pr\left(A\cap B\cap C\right)}{\Pr\left(B\cap C\right)}=\Pr\left(A|B\cap C\right).$$

■

2.1 חוק Bayes

חוק בייז הוא חוק פשוט למידי שנובע ישירות מההגדרה של הסתברות המותנה, אבל יחד עם זאת הוא מאוד יעיל בחישובי הסתברות.

משפט 2.8 (חוק בייז) יהי צמד מאורעות A,B במרחב הסתברות (Ω,P) בעלי הסתברות חיובית. אזי

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

חוק בייז מאוד נוח לשימוש כאשר ישנה הסתברות מותנה ודווקא ההסתברות ההפוכה היא זו שידועה לנו.

הוכחה: ההוכחה נובעת ישירות מההגדרה.

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

מטעמי סימטריה נובע גם ש־ $\Pr\left(B\cap A\right)=\Pr\left(B|A\right)\Pr\left(A\right)$ וכעת נשווה את צמד המשוואות זו לזו ונחלק ב־ $\Pr\left(B\right)$. ■

תרגיל 2.9 הרכב הסטודנטים בתואר ראשון הוא כדלקמן:

- 50% מהסטודנטים הם בשנה א' לתואר.
- 30% מהסטודנטים הם בשנה ב' לתואר.
- 20% מהסטודנטים הם בשנה ג' לתואר.

בנוסף, מצבם המשפחתי של הסטודנטים הנ"ל הוא:

- 20% מהסטודנטים שנמצאים בשנה א' ־ נשואים.
- 30% מהסטודנטים שנמצאים בשנה ב' ־ נשואים.
- 40% מהסטודנטים שנמצאים בשנה ג' ־ נשואים.

1. בהנחה שפגשנו באקראי את אלון, סטודנט מתואר ראשון, מה הסיכוי שהוא נשוי?
2. מסתבר שאלון אכן נשוי, מה הסיכוי שהוא משנה ג'?
3. פגשנו סטודנט אחר, בן, גם כן מתואר ראשון ומסתבר שהוא איננו נשוי, מה הסיכוי שהוא משנה ג'?

פתרון: בשאלה זו נשתמש בעיקר בנוסחת ההסתברות השלמה וחוק בייז. שימו לב שנוסחת ההסתברות השלמה היא נוסחה מאוד בסיסית ושימושית ביותר ולכן נשתמש בה לכל אורך הקורס במספר רב של מקרים.

1. נסמן ב־M את המאורע בו הסטודנט נשוי וב־I, II, III את המאורעות בהם הסטודנט הוא בשנה א',ב' ו־ג' לתואר בהתאמה.

$$\begin{aligned}\Pr(M) &= \Pr(M|I)\Pr(I) + \Pr(M|II)\Pr(II) + \Pr(M|III)\Pr(III) = \\ &= \frac{2}{10}\cdot\frac{5}{10} + \frac{3}{10}\cdot\frac{3}{10} + \frac{4}{10}\cdot\frac{2}{10} = \frac{27}{100} = 0.27\end{aligned}$$

2. נשתמש בחוק בייז מהסיבה הפשוטה שנתונות לנו ההסתברויות להיות נשוי בהתאם לשנת הלימוד אבל ההתנייה ההפוכה לא נתונה ולכן

$$.\Pr(III|M) = \frac{\Pr(M|III)\Pr(III)}{\Pr(M)} = \frac{0.4\cdot 0.2}{0.27} = \frac{0.08}{0.27} \cong 0.296$$

3. נחשב באופן דומה אלא שהפעם ההתנייה היא על מאורע חדש.

$$.\Pr(III|M^c) = \frac{\Pr(M^c|III)\Pr(III)}{\Pr(M^c)} = \frac{0.6\cdot 0.2}{1-0.27} = \frac{0.12}{0.73} \cong 0.164$$

סטודנטים רבים מתבלבלים וחושבים שהתוצאה של סעיף ג' היא התוצאה המשלימה למה שקיבלנו בסעיף ב'. זה כמובן לא המצב. הסתברות מותנה היא הסתברות לכל הדבר, אבל אין זה אומר שהתנייה על מאורע אחד והתנייה על המאורע המשלים לו ייסכמו ל־1.

תרגיל 2.10 בן הוא כורה פחם שהחליט לתת לבתו, אלונה, מתנה לחתונה. הוא נותן לה לבחור בין 2 קופסאות הזהות חיצונית ־ קופסא א' המכילה 2 יהלומים ופחם וקופסא ב' המכילה 2 פחמים ויהלום. מאחר ובן הוא אב מתחשב, הוא החליט לאפשר לה להוציא באקראי אבן מקופסא אחת ורק לאחר מכן להחליט איזו קופסא לקחת. מה האסטרטגיה האופטימלית מבחינת אלונה?

פתרון: זו שאלה המתקשרת לתורת המשחקים ולכן צריך להבין תחילה מה זאת בכלל אסטרטגיה בסיטואציה הנוכחית. אסטרטגיה, באופן כללי, זו פונקציה מהאינפורמציה של אלונה לפעולות שלה. במקרה הנוכחי אלונה רואה את האבן שהיא הוציאה ולכן המידע שזמין עבורה הוא "יהלום" או "פחם" והיא צריכה להחליט מה לעשות בכל סיטואציה.

- אם אלונה מחליטה להישאר עם הקופסא שבחרה בכל מקרה (בין אם הוציאה פחם ובין אם הוציאה יהלום) אז ההסתברות שהיא תקבל לבסוף את קופסא א' היא $\frac{1}{2}$. הסיבה היא שההסתברות לבחור את קופסא א' תחילה היא חצי, ואלונה לא מחליפה קופסא, אזי ההסתברות נשמרת.
- באופן סימטרי ־ במידה ואלונה מחליטה להחליף קופסא בכל מקרה, אז ההסתברות שתקבל את קופסא א' בסופו של דבר היא גם כן חצי (שכנעו את עצמכם למה זה נכון, בדומה לסעיף הקודם עם שיקולי סימטריה).
- אלונה מחליטה להחליף את הקופסא במידה ותוציא פחם ולהישאר עם הקופסא במידה ותוציא יהלום. נחשב את ההסתברות לקבל לבסוף את קופסא א'. נסמן ב־A_1, B_1

את המאורעות בהן אלונה בחרה תחילה בקופסא א' ו־ב' בהתאמה וב־A_2, B_2 את המאורעות בהן קיבלה לבסוף את קופסאות א' ו־ב' בהתאמה.

$$\begin{aligned}\Pr(A_2) &= \Pr(A_2|A_1)\Pr(A_1)+\Pr(A_2|B_1)\Pr(B_1)= \\ &= \frac{2}{3}\cdot\frac{1}{2}+\frac{2}{3}\cdot\frac{1}{2}=\frac{2}{3}\end{aligned}$$

ואנו רואים שאסטרטגיה זו כבר יותר טובה מצמד הקודמות.

- אלונה מחליטה להחליף את הקופסא במידה ותוציא יהלום ולהישאר עם הקופסא במידה ותוציא פחם.

$$\begin{aligned}\Pr(A_2) &= \Pr(A_2|A_1)\Pr(A_1)+\Pr(A_2|B_1)\Pr(B_1)= \\ &= \frac{1}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{2}=\frac{1}{3}\end{aligned}$$

ולכן האסטרטגיה האופטימאלית היא האסטרטגיה השלישית.

באותה מידה ניתן לפתור את התרגיל הקודם בעזרת עץ הסתברויות.

תרגיל 2.11 (הכד של פוליה) בכד 5 כדורים שחורים ו־3 לבנים. בכל הוצאה של כדור מהכד, מחזירים אותו חזרה יחד עם 4 נוספים מאותו הצבע.

1. מה ההסתברות שהכדור הראשון שנוציא יהיה שחור?
2. מה ההסתברות שהכדור השני שנוציא יהיה שחור?
3. מה ההסתברות שהכדור ה־100 שנוציא יהיה שחור?

פתרון: נפתור את צמד הסעיפים הראשונים בעזרת חישובים קומבינטוריים ונוסחת ההסתברות השלמה.

1. הפתרון לסעיף הראשון הוא מיידי והתשובה היא $\frac{5}{8}$ (חלוקה של גודל המאורע במרחב המדגם).
2. נחשב בעזרת נוסחת ההסתברות השלמה. נסמן ב־ W_i, B_i , $i=1,2$ את המאורעות בהם הכדור ה־i הוא שחור או לבן בהתאמה.

$$\begin{aligned}\Pr(B_2) &= \Pr(B_2|B_1)\Pr(B_1)+\Pr(B_2|W_1)\Pr(W_1)= \\ &= \frac{9}{12}\cdot\frac{5}{8}+\frac{5}{12}\cdot\frac{3}{8}=\frac{45+15}{96}=\frac{5}{8}\end{aligned}$$

באופן מפתיע קיבלנו שההסתברות להוציא שחור בשלב הראשון ובשלב השני היא זהה.

3. לפני שנענה על הסעיף הנוכחי, ננסה למצוא הוכחה יותר פשוטה שלא תצריך התניות רבות. נתחיל במצב יחסית פשוט: מה היה קורה לו היו לנו רק 2 כדורים בכד בהתחלה ־ 1 שחור ו־1 לבן? אזי משיקולי סימטריה ההסתברות שהכדור ה־100 יהיה שחור היא $\frac{1}{2}$ (מאחר והכדור ה־100 הוא שחור או לבן, וישנה סימטריה מלאה בין השחורים

ללבנים, אזי ההסתברות שנוציא כל אחד מהם בפעם ה־100 היא שווה). מה לגבי מצב בו ישנם 8 כדורים בכד ללא צבעים אשר ממוספרים מ־1 עד 8 (לאחר הוצאה מחזירים את הכדורים עם אותו המספר של הכדור שהוצא)? במצב זה, גם כן יש סימטריה בין שמונת הכדורים ולכן ההסתברות להוציא כדור עם ערך $i = 1, \ldots, 8$ היא $\frac{1}{8}$. כעת נוכל להניח שיש מספרים על הכדורים בבעיה המקורית, הסיכוי שנוציא את כדורים 1 עד 5 (אלו הכדורים הצבועים בשחור) הוא $\frac{5}{8}$ כסכום ההסתברויות להוציא כל מספר מ־1 עד 5 בנפרד. ולסיכום, ההסתברות שנוציא כדור שחור בפעם ה־100 היא $\frac{5}{8}$ וזה נכון עבור כל שלב שנבחר.

טענה 2.12 יהיו צמד מאורעות בלתי־תלויים A, B במרחב הסתברות $(\Omega, \Pr)$ ונניח כי למאורע B הסתברות חיובית. אזי

$$. \Pr(A|B) = \Pr(A)$$

הוכחה: ישירות מההגדרה.

$$. \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)\Pr(B)}{\Pr(B)} = \Pr(A)$$

■

2.2 אי תלות בין מאורעות

הגדרה 2.13 יהיו צמד מאורעות A, B במרחב הסתברות $(\Omega, \Pr)$. המאורעות הם בלתי־תלויים אם

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

תלות בין מאורעות בהסתברות מבטא למעשה את מידת הקשר ההסתברותי בין צמד המאורעות הללו. מן ההגדרה נובע שכאשר המאורעות בלתי־תלויים אזי האחד לא משפיע על הסבירות של המאורע האחר להתרחש. נאמר על המאורעות A, B במרחב הסתברות $(\Omega, \Pr)$ שהם תלויים אם הם לא בלתי־תלויים, ז"א

$$. \Pr(A \cap B) \neq \Pr(A)\Pr(B)$$

תרגיל 2.14 (בעיית מונטי־הול) נניח ואתם משתתפים בתוכנית טלוויזיה. בתוכנית מוצגים לכם שלושה וילונות כאשר ידוע שמאחורי אחד מהם ישנה מכונית ומאחורי השניים האחרים ־ דחליל. הפרסים מחולקים באקראי בהסתברות שווה מאחורי הוילונות. אתם נדרשים לבחור וילון. לאחר שבחרתם, פותח המנחה של התוכנית וילון אחד באקראי בהסתברות שווה מתוך הוילונות שמצד אחד, אין בהם פרס, ומצד שני, לא בחרתם בהם. לאחר כל זה, יש באפשרותכם לשנות את החלטתכם. תחת ההנחה שאתם שואפים למקסמם את הסתברויות הזכייה שלכם ובהנחה שבחרתם בוילון 1 תחילה והמנחה פתח את וילון 3, מה האסטרטגיה האופטימאלית מבחינתכם?

הבעיה הנ"ל היא בעיה ידועה בשם ־ בעיית מונטי הול. הבעיה נקראת על שם מנחה התוכנית, "Let's make a deal", ששודרה במשך מספר עשורים בארה"ב. הסיפור מאחורי הבעיה הזאת הוא די מעניין. בחורה מאוד חכמה (נטען שהיא מחזיקה בשיא גינס עבור מנת משכל) בשם מרלין ווס סאוונט הייתה בעלת טור בעיתון ובו הייתה עונה על שאלות וחידות, לרוב מתמטיות, שהיו שולחים אליה האנשים מחבי ארה"ב. בספטמבר 1990, בחור ממרילנד שאל אותה את השאלה הנ"ל והיא ענתה לו שהאסטרטגיה האופטימאלית היא החלפה. התשובה הזאת גררה אלפי ואולי אף יותר תגובות (ובינהם גם אנשים שטענו שהם דוקטורנטים למתמטיקה) שטענו שהיא טועה. כעת, אנחנו נוכיח שהיא דווקא צדקה.

פתרון: אחד הדברים החשובים בבעיה הנוכחית היא להגדיר היטב את מרחב המדגם. נניח שהדלתות ממוספרות מ־1 עד 3. נגדיר את מרחב המדגם כ־

$$\Omega = \{(i,j,k):\ i,j,k \in \{1,2,3\}\},$$

כאשר הקוארדינטה הראשונה מתייחסת לבחירה הראשונית של המתמודד, הקוארדינטה השנייה מתייחסת למיקום המכונית והשלישית מתייחסת לוילון שהמנחה בחר לפתות. פונקצית ההסתברות על מרחב המדגם היא:

$$\Pr(i,j,k) = \begin{cases} 0 & k=i \ or\ k=j \\ \frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{2} & i=j, k\neq i,j \\ \frac{1}{3}\cdot\frac{1}{3}\cdot 1 & i\neq j, k\neq i,j \end{cases},$$

כאשר השורה הראשונה נובעת מכך שהמנחה לא יכול לפתוח את הוילון בו יש מכונית או את הוילון בו בחרתם; השורה השנייה נובעת מכך שמיקום המכונית והבחירה הראשונית של המתמודדים הן ב"ת וכן המנחה בוחר את אחד מהוילונות האחרים בהסתברות שווה; השורה השלישית נובעת מכך שברגע ששני וילונות "תפוסים" (אחד ע"י הבחירה הראשונית של המתמודד והאחר ע"י מיקום המכונית), למנחה יש רק וילון אחד אפשרי לפתוח.

נגדיר כעת צמד אסטרטגיות: הראשונה תהיה "החלפה" והיא אומרת למתמודד להחליף את בחירתו המקורית לאחר פתיחת הדלת. נסמן את המאורע בו אסטרטגיה זו מנצחת ב־$SWITCH$. האסטרטגיה השנייה תהיה "הישארות" והיא ואומרת למתמודד להישאר עם בחירתו המקורית לאחר פתיחת הדלת. נסמן את המאורע בו אסטרטגיה זו מנצחת ב־$STAY$.

$$\begin{aligned} \Pr\left(SWITCH | \bigcup_{j=1}^{3}(1,j,3)\right) &= \frac{\Pr\left(SWITCH \cap \left(\bigcup_{j=1}^{3}(1,j,3)\right)\right)}{\Pr\left(\bigcup_{j=1}^{3}(1,j,3)\right)} = \\ &= \frac{\Pr(1,2,3)}{\Pr\left(\bigcup_{j=1}^{3}(1,j,3)\right)} = \\ &= \frac{\Pr(1,2,3)}{\Pr(1,1,3)+\Pr(1,2,3)+\Pr(1,3,3)} = \\ &= \frac{\frac{1}{3}\cdot\frac{1}{3}\cdot 1}{\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{3}\cdot 1+0} = \frac{1}{\frac{1}{2}+1} = \frac{2}{3}. \end{aligned}$$

לעומת זאת,

$$
\begin{aligned}
\Pr\left(STAY|\bigcup_{j=1}^{3}(1,j,3)\right) &= \frac{\Pr\left(STAY\cap\left(\bigcup_{j=1}^{3}(1,j,3)\right)\right)}{\Pr\left(\bigcup_{j=1}^{3}(1,j,3)\right)} = \\
&= \frac{\Pr(1,1,3)}{\Pr\left(\bigcup_{j=1}^{3}(1,j,3)\right)} = \\
&= \frac{\Pr(1,2,3)}{\Pr(1,1,3)+\Pr(1,2,3)+\Pr(1,3,3)} = \\
&= \frac{\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{2}}{\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{3}\cdot 1+0} = \frac{\frac{1}{2}}{\frac{1}{2}+1} = \frac{1}{3},
\end{aligned}
$$

ומכאן נובעת התשובה הרצויה.

תרגיל 2.15 בספרייה ישנם 10 ספרי הסתברות. 5 ספרים הכוללים פתרונות ו־5 ללא פתרונות. הספרנית איבדה ספר אחד מתוך ה־10 (כלשהו). אלון מגיע לקחת ספר, מקבל ספר מתוך התשעה הנותרים באקראי ומחזיר לאחר יומיים. לאחר שבועיים מגיע בן ולוקח גם כן ספר 1 מתוך ה־9 ומחזיר לאחר יומיים. נגדיר:

- A ־ הספר של אלון כלל פתרונות.
- B ־ הספר של בן כלל פתרונות.

האם A,B תלויים?

פתרון: לכאורה, אין סיבה שתהיה תלות. מצד אחד, אלון לקח ספר והחזיר לאחר יומיים. שבועיים לאחר מכן, אחרי שאלון החזיר כבר את הספר, מגיע בן ולוקח ספר כלשהו, לכן למה שתהיה תלות?! אבל במקרה זה אכן המאורעות תלויים. נראה זאת בעזרת חישוב ישיר. נגדיר C ־ הספר שאבד כלל פתרונות. משיקולי סימטריה, $\Pr(B)=\Pr(A)$ ולכן נחשב רק אחד מהם.

$$
\begin{aligned}
\Pr(A) &= \Pr(A|C)\Pr(C)+\Pr(A|C^c)\Pr(C^c) = \\
&= \frac{4}{9}\cdot\frac{1}{2}+\frac{5}{9}\cdot\frac{1}{2} = \frac{1}{2}
\end{aligned}
$$

על כן נקבל ש־ $\Pr(A)\cdot\Pr(B)=\frac{1}{4}$. מצד שני,

$$
\begin{aligned}
\Pr(A\cap B) &= \Pr(A\cap B|C)\Pr(C)+\Pr(A\cap B|C^c)\Pr(C^c) = \\
&= \left(\frac{4}{9}\right)^2\cdot\frac{1}{2}+\left(\frac{5}{9}\right)^2\cdot\frac{1}{2} \cong 0.253 \neq \Pr(A)\cdot\Pr(B)
\end{aligned}
$$

לכן המאורעות תלויים. ההסבר הוא פשוט יחסית, ניקח את הסיטואציה למצב קיצון ונניח שבין אלון ובן היו מגיעים עוד 1000 איש ומבצעים את הניסוי, והיינו שמים לב שרק כ־$\frac{4}{9}$

מתוכם קיבלו ספר עם פתרונות. במצב כזה היינו חושבים שככל הנראה אבד ספר עם פתרונות בהתחלה. שימו לב שלעומת זאת,

$$\Pr(A\cap B|C) = \Pr(A|C)\Pr(B|C)$$

זאת אי־תלות מותנה ונראה אותה בהמשך.

הגדרה 2.16 תהיי סדרת מאורעות $A_1,\ldots,A_n$ במרחב הסתברות $(\Omega,\Pr)$. המאורעות הם בלתי־תלויים אם

$$\begin{aligned}\Pr\left(\bigcap_{i=1}^{n}A_i\right) &= \prod_{i=1}^{n}\Pr(A_i)\\ \Pr\left(\left(\bigcap_{i=1}^{n-1}A_i\right)\cap A_n^c\right) &= \left(\prod_{i=1}^{n-1}\Pr(A_i)\right)\Pr(A_n^c)\\ &\vdots\\ \Pr\left(\bigcap_{i=1}^{n}A_i^c\right) &= \prod_{i=1}^{n}\Pr(A_i^c)\end{aligned}$$

סה"כ ישנם 2^n משוואות אשר אמורות להתקיים בכדי שהמאורעות יהיו ב"ת.

הערה 2.17 המקרה בו $n=2$ הוא מקרה פרטי בו השיוויון $\Pr(A\cap B)=\Pr(A)\Pr(B)$ גורר את השיוויונות

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

לדוגמא:

$$\begin{aligned}\Pr(A\cap B^c) &= \Pr(A)-\Pr(A\cap B)=\\ &= \Pr(A)-\Pr(A)\Pr(B)=\\ &= \Pr(A)(1-\Pr(B))=\\ &= \Pr(A)\Pr(B^c)\end{aligned}$$

באופן האופן ניתן להכליל את הגדרת האי־תלות לסדרה של מאורעות באופן הבא:

טענה 2.18 המאורעות $A_1,\ldots,A_n$ הם ב"ת אם ורק אם לכל תת קבוצה $A_{i_1},\ldots,A_{i_k}$ עבור $2\le k\le n$ מתקיים

$$\Pr\left(\bigcap_{j=1}^{k}A_{i_j}\right)=\prod_{j=1}^{k}\Pr\left(A_{i_j}\right).$$

שימו לב שההוכחה לטענה זאת מתבצעת בדיוק באותו האופן בו הוכחנו את ההערה הקודמת.

תרגיל 2.19 בקופסא ישנם 3 כדורים שחורים, 4 כדורים אדומים ו־93 כדורים לבנים. מציאים כדור בכל שלב עם החזרה ועוצרים בפעם הראשונה בה מתקבל כדור שאינו לבן.

1. מה הסיכוי שהתהליך יימשך לנצח?
2. מה ההסתברות שהתהליך יסתיים בהוצאת כדור אדום?

פתרון: נגדיר את המאורעות W_k, R_k כמאורע בו הוצאנו בשלב ה־k כדור אדום או לבן בהתאמה עבור כל $k \geq 1$ טבעי.

1. אנחנו מעוניינים לחשב את ההסתברות $\Pr\left(\bigcap_{k=1}^{\infty} W_k\right)$. נשים לב שמדובר במאורעות ב"ת מאחר וההוצאה מתבצעת עם החזרה ולכן
$$\Pr\left(\bigcap_{k=1}^{N} W_k\right) = \prod_{k=1}^{N} P\left(W_k\right) = \left(\frac{93}{100}\right)^N.$$
לכן נקבל ש־
$$\begin{aligned} 0 \leq \Pr\left(\bigcap_{k=1}^{\infty} W_k\right) &= \lim_{N\to\infty} \Pr\left(\bigcap_{k=1}^{N} W_k\right) = \\ &= \lim_{N\to\infty} \left(\frac{93}{100}\right)^N \to 0 \end{aligned}$$
לפי כלל הסנדביץ' נובע שההסתברות הינה 0.
2. נגדיר את המאורע R כמאורע בו התהליך הסתיים עם הוצאה של כדור אדום.

$$\begin{aligned} \Pr(R) &= \Pr(R_1) + \sum_{N=1}^{\infty} \Pr\left(\bigcap_{k=1}^{N} W_k\right) \Pr\left(R_{N+1} \middle| \bigcap_{k=1}^{N} W_k\right) = \\ &= \Pr(R_1) + \sum_{N=1}^{\infty} \Pr\left(\bigcap_{k=1}^{N} W_k\right) \Pr(R_1) = \\ &= \Pr(R_1) + \Pr(R_1) \sum_{N=1}^{\infty} \left(\frac{93}{100}\right)^N = \\ &= \Pr(R_1) \sum_{N=0}^{\infty} \left(\frac{93}{100}\right)^N = \\ &= \frac{4}{100} \cdot \frac{1}{1-\frac{93}{100}} = \frac{4}{7} \end{aligned}$$

והאינטואיציה מאחורי התוצאה הזאת היא מאוד פשוטה. אם התהליך נגמר רק כאשר יוצא כדור אדום או שחור ואנחנו יודעים שהתהליך ייגמר בודאות בסופו של דבר, משמעות הדבר היא שהכדור האחרון שיוצא הוא אדום או שחור. לכן ההסתברות שהוצא כדור אדום בהינתן שהתהליך נגמר חייב להיות $\frac{4}{7}$, כמספר הכדורים האדומים חלקי מספר כל הכדורים האפשריים בשלב האחרון.

תרגיל 2.20 (הכד של פוליה 2) בכד כדור שחור וכדור לבן, לאחר שמוציאים כדור באקראי מחזירים אותו יחד עם עוד כדור 1 מאותו הצבע. מוציאים סה"כ 5 כדורים.

1. מה הסיכוי שכולם שחורים?
2. מה הסיכוי שהוצאנו בדיוק לבן 1?

פתרון: נגדיר את המאורעות W_i, B_i, $i = 1, \ldots, 5$ כמאורעות בהם הכדור ה־i הוא שחור או לבן בהתאמה.

1. נשתמש בחוק הכפל לטובת החישוב.

$$\begin{aligned}\Pr\left(\bigcap_{i=1}^{5} B_i\right) &= \Pr(B_1)\Pr(B_2|B_1)\Pr(B_3|B_1\cap B_2)\cdots\Pr\left(B_5|\bigcap_{i=1}^{4} B_1\right) = \\ &= \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}\cdot\frac{5}{6} = \frac{1}{6}\end{aligned}$$

2. נשים לב שמדובר באיחוד של מאורעות זרים.

$$\Pr\left(\bigcup_{i=1}^{5}\left(W_i \cap \left(\bigcap_{j\neq i} B_j\right)\right)\right) = \sum_{i=1}^{5}\Pr\left(W_i \cap \left(\bigcap_{j\neq i} B_j\right)\right)$$

ננסה לפתור את המקרה הכללי עבור $1 \le i \le 5$ כלשהו ולהציב חזרה בסכום.

$$\begin{aligned}\Pr\left(W_i \cap \left(\bigcap_{j\neq i} B_j\right)\right) &= \Pr(B_1)\Pr(B_2|B_1)\cdots\Pr\left(B_{i-1}|\bigcap_{j=1}^{i-2} B_j\right)\cdot \\ &\cdot\Pr\left(W_i|\bigcap_{j=1}^{i-1} B_j\right)\Pr\left(B_{i+1}|\left(\bigcap_{j=1}^{i-1} B_j\right)\cap W_i\right)\cdots \\ &\cdots\Pr\left(B_5|\left(\bigcap_{j\neq i,5} B_j\right)\cap W_i\right) = \\ &= \frac{1}{2}\cdot\frac{2}{3}\cdots\frac{i-1}{i}\cdot\frac{1}{i+1}\cdot\frac{i}{i+2}\cdots\frac{4}{6}\end{aligned}$$

נשים לב שהמכנה תמיד גדל ב־1 כי מרחב המדגם גדל ב־1 בכל שלב בניסוי. יחד עם זאת בשלב ה־i ישנה רק אפשרות אחת להוציא לבן ומיד לאחר מכן מספר האפשרויות להוציא שחור הוא i (כמספר הכדורים השחורים בכד באותו השלב). ההמשך זהה לתרגיל הקודם. לכן קיבלנו כי

$$\begin{aligned}\Pr\left(W_i \cap \left(\bigcap_{j\neq i} B_j\right)\right) &= \frac{1}{2}\cdot\frac{2}{3}\cdots\frac{i-1}{i}\cdot\frac{1}{i+1}\cdot\frac{i}{i+2}\cdots\frac{4}{6} = \\ &= \frac{1\cdot 4!}{6!} = \frac{1}{30}\end{aligned}$$

ולסיכום,

$$\begin{aligned}\Pr\left(\bigcup_{i=1}^{5}\left(W_i\cap\left(\bigcap_{j\neq i}B_j\right)\right)\right) &= \sum_{i=1}^{5}\Pr\left(W_i\cap\left(\bigcap_{j\neq i}B_j\right)\right)=\\ &= \sum_{i=1}^{5}\frac{1}{30}=\frac{1}{30}\cdot 5=\frac{1}{6}.\end{aligned}$$

תרגיל 2.21 (סעיף משאלה של מועד ב', סמסטר ב' תשע"ד) לבן יש מטבע כחול ולאלון מטבע אדום. למטבע הכחול הסתברות p לנחות על 'עץ' (H) ולמטבע האדום הסתברות q. אלון ובן משחקים את המשחק הבא:

- בכל תור, מטילים שניהם את המטבע שלהם באופן ב"ת באחר או בהטלות הקודמות.
- אם התקבל פעמיים H, אז הם מחליפים ביניהם את המטבעות.
- אחרת, כל אחד נשאר עם המטבע שהתחיל איתו את הסיבוב.
- התהליך חוזר על עצמו שוב ושוב.

לכל $n\in\mathbb{N}$, יהי A_n המאורע שבתום n סיבובים אלון מחזיק את מהטבע האדום.

1. חשבו את $\Pr\left(A_2\right)$.
2. עבור אילו ערכי p ו־q המאורעות A_2 ו־A_3 ב"ת?

פתרון:

1. ננסה לחשוב מתי A_2 אפשרי. זהו המאורע בו אחרי 2 סיבובים אלון מחזיק את המטבע האדום איתו התחיל. דבר זה יכול להתרחש אם ורק אם לא הוחלפו מטבעות כלל, או לחילופין, התבצעו שתי החלפות. ההסתברות להחלפה של מטבע בסיבוב נתון היא pq. לכן נקבל ש־

$$\Pr\left(A_2\right)=\left(1-pq\right)^2+\left(pq\right)^2 .$$

2. דרך אחת לבחון אי־תלות זה בעזרת ההגדרה שאומרת שאם $\Pr\left(A_2\cap A_3\right)=\Pr\left(A_2\right)\Pr\left(A_3\right)$ אז המאורעות ב"ת. ישנה עוד דרך להראות אי־תלות וזאת על ידי השיוויון

$$\Pr\left(A_3|A_2\right)=P\left(A_3\right) .$$

מדוע זה אפשרי? ובכן, לפי הגדרה, נובע ש־

$$\Pr\left(A_3|A_2\right)=\frac{\Pr\left(A_3\cap A_2\right)}{\Pr\left(A_2\right)} ,$$

ואם בנוסף אנו יודעים ש־ $\Pr(A_3|A_2) = P(A_3)$ אזי נובע ש־

$$\frac{\Pr(A_3 \cap A_2)}{\Pr(A_2)} = \Pr(A_3),$$

כנדרש. נחשב מפורשות את $\Pr(A_3|A_2)$ ואת $P(A_3)$. נסמן $s := pq$ ונקבל

$$\begin{aligned} P(A_3) &= (1-s)^3 + 3(1-s)s^2 \\ \Pr(A_3|A_2) &= 1-s \end{aligned}$$

ולכן נבדוק מתי מתקיים

$$1 - s = (1-s)^3 + 3(1-s)s^2$$

פתרון ראשון הוא $s_1 = 1$. אחרת

$$\begin{aligned} &1 = (1-s)^2 + 3s^2 \\ &0 = -2s + 4s^2 \\ &0 = s(2s-1) \end{aligned}$$

ולכן נקבל את הפתרונות $s_1 = 1,\ s_2 = 0,\ s_3 = \frac{1}{2}$. נתרגם זאת למונחי p, q.

$$\begin{aligned} s = 1 &\quad\Leftrightarrow\quad p = q = 1 \\ s = 0 &\quad\Leftrightarrow\quad p = 0\ \vee q = 0 \\ s = \frac{1}{2} &\quad\Leftrightarrow\quad p = \frac{1}{2q}. \end{aligned}$$

ואילו המקרים בהם ישנה אי־תלות.

התוצאה שקיבלנו היא די הגיונית ואפילו אינטואיטיבית. כאשר $pq = 0$ או $pq = 1$ אז מקבלים שהמאורעות A_2 ו־A_3 מתרחשים בהסתברות 0 או 1 ולכן ראינו כבר שהם ב"ת בכל מאורע אחר. בנוסף, כאשר $pq = \frac{1}{2}$, אז בכל שלב יש סיכוי שווה להיות בכל מצב ללא תלות בעבר. למעשה, ישנה סימטריה בבעיה מבחינת מיקום המטבעות אצל אלון ובן ולכן מתקיימת אי־תלות.

2.2.1 אי־תלות מותנה

כעת נוכל לשלב את צמד הרעיונות שהצגנו עד כה בפרק: הסתברות מותנה ואי־תלות.

יהי מרחב הסתברות $(\Omega, \Pr)$ ונניח כי למאורע C הסתברות חיובית. אזי, נוכל להגדיר את מרחב ההסתברות המותנה $(\Omega, \Pr(\cdot|C))$. נבחר צמד מאורעות $A, B \subseteq \Omega$.

הגדרה 2.22 מאורעות A, B הם בלתי־תלויים בהינתן C אם

$$\Pr(A \cap B|C) = \Pr(A|C)\Pr(B|C).$$

שימו לב! ההערה הבאה היא מאוד חשובה וסטודנטים רבים נוטים להתבלבל איתה.

הערה 2.23 אם המאורעות A,B בלתי־תלויים בהינתן C , אין זה אומר שהם בלתי־תלויים באופן כללי. כמו כן, גם הכיוון השני איננו נכון. מרחב ההסתברות המותנה הוא מרחב הסתברות חדש ולכן קיום של תלות או אי־תלות בו איננה מחייבת דבר על מרחבי הסתברות אחרים (למרות שכמובן יש קשר בין מרחב ההסתברות המקורי למרחב ההסתברות המותנה).

תרגיל 2.24 על השולחן מונחים 2 כדים: בכד הראשון (כד א') שלושה כדורים לבנים וחמישה כדורים אדומים; בכד השני (כד ב') ישנם שישה כדורים לבנים ושלושה כדורים אדומים. אלון מטיל מטבע הוגן. אם תוצאת ההטלה היא "עץ", הוא יוציא מכד א' 2 כדורים עם החזרה. אחרת, הוא יוצא מכד ב' 2 כדורים עם החזרה. נסמן ב־A את המאורע שהכדור הראשון לבן, ב־B את המאורע שהכדור השני לבן וב־C את המאורע בו נבחר כד א'.

1. האם המאורעות A,B תלויים?
2. האם המאורעות A,B תלויים בהינתן C ?

פתרון: נתחיל בבניית מרחב המדגם. נסמן את הכדורים בכד א' במספרים $1,\ldots,8$ כאשר שלושת הראשונים יהיו הכדורים הלבנים והיתר אדומים. באופן דומה נסמן את הכדורים בכד ב' במספרים $1,\ldots,9$ כאשר ששת הראשונים לבנים והיתר אדומים. מרחב המדגם מורכב משלשות כאשר הקוארדינטה הראשונה היא זהות הכד שנבחר והקוארדינטות האחרונות הן ערכי הכדורים שהוצאו מהכד שנבחר. נשים לב שמרחב המדגם מכיל

$$|\Omega| = 8^2 + 9^2 = 145$$

אפשרויות שונות. בנוסף נשים לב שמערכת המדגם אינו מרחב מדגם שווה הסתברות. לדוגמא: ההסתברות להוציא צמד כדורים כלשהם מכד א' היא $\frac{1}{2}\cdot\frac{1}{64}$ בעוד ההסתברות להוציא צמד כדורים מכד ב' היא $\frac{1}{2}\cdot\frac{1}{81}$.

1. נחשב את ההסתברות של כל מאורע בנפרד.

$$\Pr(A) = \frac{1}{2}\cdot\frac{3\cdot 8}{64} + \frac{1}{2}\cdot\frac{6\cdot 9}{81} = \frac{3}{16} + \frac{3}{9} = \frac{75}{144}$$

כאשר ביצענו התנייה על הכד שנבחר, הסתברות $\frac{1}{2}$ לכל כד, ולאחר מכן ספרנו את מספר המאורעות בכל חלק ממרחב המדגם. באופן זהה לחישוב הקודם, נקבל ש־

$$\Pr(B) = \frac{1}{2}\cdot\frac{3\cdot 8}{64} + \frac{1}{2}\cdot\frac{6\cdot 9}{81} = \frac{3}{16} + \frac{3}{9} = \frac{75}{144},$$

וכמובן, $\Pr(C) = \frac{1}{2}$. נחשב כעת את ההסתברות למאורע $A\cap B$.

$$\Pr(A\cap B) = \frac{1}{2}\cdot\frac{3\cdot 3}{64} + \frac{1}{2}\cdot\frac{6\cdot 6}{81} = \frac{9}{128} + \frac{2}{9} = \frac{337}{1152}.$$

נשים לב ש־ $\Pr(A)\Pr(B) \neq \Pr(A\cap B)$ ולכן המאורעות תלויים.

2. נחשב את ההסתברות של כל מאורע מותנה בנפרד את ההסתברות של חיתוך המאורעות תחת ההתנייה.

$$\Pr(A|C) = \frac{3 \cdot 8}{64} = \frac{3}{8} = \Pr(B|C)$$
$$\Pr(A \cap B|C) = \frac{3 \cdot 3}{64} = \frac{9}{64} = \Pr(A|C)\Pr(B|C)$$

לכן המאורעות הללו בלתי־תלויים בהינתן C.

משפט 2.25 יהי מרחב הסתברות $(\Omega, \Pr)$ ו־3 מאורעות A, B, C בלתי־תלויים כך ש־ $\Pr(C) > 0$. אזי A, B בלתי־תלויים בהינתן C.

הוכחה: נשתמש בהגדרה לאי־תלות מותנה של צמד מאורעות ונראה ש־ $\Pr(A \cap B|C) = \Pr(A|C)\Pr(B|C)$.

$$\begin{aligned}
\Pr(A \cap B|C) &= \frac{\Pr(A \cap B \cap C)}{\Pr(C)} = \\
&= \frac{\Pr(A)\Pr(B)\Pr(C)}{\Pr(C)} = \\
&= \frac{\Pr(A)\Pr(C)}{\Pr(C)} \cdot \frac{\Pr(B)\Pr(C)}{\Pr(C)} = \\
&= \frac{P(A \cap C)}{\Pr(C)} \cdot \frac{\Pr(B \cap C)}{\Pr(C)} = \\
&= \Pr(A|C)\Pr(B|C)
\end{aligned}$$

כנדרש. ■

2.3 גרפים מקריים

יהי גרף $G = (V, E)$. בחירת תת גרף מקרי של G או בשמו הנוסף, פרקולציה, הוא תהליך בו לוקחים גרף כלשהו ומתחילים למחוק קשתות באופן מקרי וב"ת. בצורה כזאת ניתן לגבל מספר רב של תתי־גרפים של הגרף המקורי, כל אחד בהסתברות המתאימה לו. לשם הגדרת המודל, נתחיל בהגדרה בסיסית של גרף.

הגדרה 2.26 גרף הוא זוג $G = (V, E)$ כאשר V היא קבוצה סופית של קודקודים ו־E היא קבוצה סופית של קשתות המחברות בין קודקודים שונים. קרי $E \subseteq V \times V$ כך ש־ $(v, v) \notin E$ לכל קודקוד $v \in V$ ובנוסף, אם $(u, v) \in E \Leftarrow$ $(v, u) \in E$.

התנאי האחרון נקרא "גרף לא מכוון" והוא אומר בעצם שאין משמעות לסדר בו נכתבים הקודקודים בתיאור של כל קשת (אין נקודת התחלה וסוף מוגדרות בקשתות) והתנאי הקודם אומר שאין קשתות עצמיות לאף קודקוד (אין קשת שמחברת קודקוד לעצמו).

ישנן 2 דרכים להגדיר **גרף מקרי**:

1. מתחילים עם קבוצה V של קודקודים ובכל שלב בוחרים צמד קודקודים שלא נבחר עד כה ומבצעים הגרלה כך שבהסתברות p תמצא קשת בין צמד הקודקודים (באופן ב"ת בקשתות אחרות).

2. מתחילים עם גרף שלם G ובכל שלב בוחרים קשת שלא נבחרה עד כה ובהסתברות $1-p$ מחליטים למחוק אותה (באופן ב"ת בקשתות אחרות).

חשוב להבדיל בין המונחים: גרף מקרי הוא תת־גרף של הגרף השלם שנוצר ע"ב התהליך שהוגדר לעיל, לבין פרקולציה / תת גרף מקרי שזה תהליך בו לוקחים תת־גרף מקרי של גרף כלשהו G, שאינו בהכרח הגרף השלם.

תרגיל 2.27 ישנן 2 כיתות בנות n תלמידים כל אחת. הכיתות נפגשות וכל זוג תלמידים, אחד מכל כיתה, נהיים חברים בסיכוי p באופן ב"ת ביתר זוגות התלמידים. יהי A המאורע שקיים תלמיד מהכיתה הראשונה שאין לו אף חבר מהכיתה השניה. יהי $0<\epsilon\ll 1$

1. הוכיחו שאם $p=(1+\epsilon)\frac{\log(n)}{n}$, אזי ההסתברות של מאורע A שואפת ל־0 כאשר $n\to\infty$.
2. הוכיחו שאם $p=(1-\epsilon)\frac{\log(n)}{n}$, אזי ההסתברות של מאורע A שואפת ל־1 כאשר $n\to\infty$.

פתרון: תחילה נבין את הקשר לגרפים מקריים. אפשר לחשוב על כל כיתה כקבוצה של קודקודים. כל תלמיד מייצג קודקוד. כל קודקוד יכול להיות מחובר אך ורק לקודקוד מהכיתה השנייה. למבנה הזה קוראים "גרף דו־צדדי מקרי". $V=V_1\cup V_2$ כאשר V_i - זו קבוצת התלמידים של כיתה i וקבוצת הקשתות E היא תת־קבוצה של הקבוצה הבאה

$$E\subseteq\{(v_1,v_2):\ v_i\in V_i\ i=1,2\}.$$

1. נחשב מפורשות את ההסתברות של מאורע A. $\Pr(A)=\binom{n}{1}(1-p)^n$ מאחר והקשתות הן בלתי־תלויות אז ההסתברות של כל תלמיד מכיתה 1 להיות ללא חברים מכיתה 2 היא שווה וב"ת בתלמידים האחרים. נשתמש באי־שוויון $1-p\le e^{-p}$ לכל $p\in\mathbb{R}$ (במידה ורוצים להוכיח זאת עבור כל $p\in[0,1]$ נשים לב שהפונקציה $f(p)=e^{-p}+p-1$ היא אי־שלילית ב־$p=0$ ומונוטונית עולה לכל $p>0$).

$$\begin{aligned}\Pr(A) &\le \binom{n}{1}(1-p)^n\le\\ &\le ne^{-np}=\\ &= ne^{-n(1+\epsilon)\frac{\log(n)}{n}}=\\ &= n\cdot n^{-1-\epsilon}=\\ &= n^{-\epsilon}\to 0\end{aligned}$$

כאשר $n\to\infty$.

2. לשם ההוכחה נצטרך אי־שוויון נוסף והוא : $e^{-p-2p^2}\le 1-p$ לכל $p\in\left[0,\frac{1}{2}\right]$. נשים לב תחילה כי הפונקציה $f(x)=1-(1+4x)e^{-x-2x^2}$ היא פונקציה רציפה וגזירה ובנוסף, יש לה נקודת מינימום בקטע $\left[0,\frac{1}{2}\right]$.

$$\begin{aligned}f(0) &= 0\\ f\left(\frac{1}{2}\right) &< 0\end{aligned}$$

ולכן הפונקציה היא מתאפסת או שלילית בקטע $\left[0,\frac{1}{2}\right]$, זאת אומרת

$$-\left(1+4x\right)e^{-x-2x^2}\leq -1\ .$$

נשתמש במשפט לגראנג' עבור הפונקציה $g\left(x\right)=e^{-x-2x^2}$ בקטע $[0,p]$ ונקבל

$$\begin{aligned}\frac{e^{-p-2p^2}-e^{-0-2\cdot 0^2}}{p-0} &= g'\left(c\right)= \\ &= \left(-1-4c\right)e^{-c-2c^2}\leq \\ &\leq -1\end{aligned}$$

נעביר אגפים ונקבל

$$e^{-p-2p^2}-1\quad\leq\quad -p$$

לכל $p\in\left[0,\frac{1}{2}\right]$ כנדרש. נעבור כעת להוכיח את הטענה ההסתברותית. נגדיר את המאורעות A_i לכל $i=1,\ldots,n$ כמאורע בו הילד ה־i הוא ללא חברים. לפי חוקי דה־מורגן נקבל ש־

$$A^c=\left(\bigcup_{i=1}^{n}A_i\right)^c=\bigcap_{i=1}^{n}A_i^c$$

ולכן

$$\begin{aligned}\Pr\left(A^c\right) &= \Pr\left(\bigcap_{i=1}^{n}A_i^c\right)= \\ &= \prod_{i=1}^{n}\Pr\left(A_i^c\right)= \\ &= \prod_{i=1}^{n}\left(1-\Pr\left(A_i\right)\right)= \\ &= \left(1-\Pr\left(A_1\right)\right)^n\end{aligned}$$

כאשר במעבר השני השתמשנו באי־תלות, במעבר השלישי השתמשנו בהגדרת המשלים ובשלב האחרון השתמשנו בסימטריה בין המאורעות. נזכור כי $\Pr\left(A_1\right)=\left(1-p\right)^n$, נציב זאת ונראה ש־ $\Pr\left(A^c\right)\to 0$ כאשר $n\to\infty$. לפי אי־השוויון שהוכחנו (תקף במקרה זה מאחר ו־$p\to 0$) נקבל

$$\begin{aligned}\Pr\left(A_1\right) &= \left(1-p\right)^n\geq \\ &\geq e^{-np-2np^2}= \\ &= e^{-n(1-\epsilon)\frac{\log(n)}{n}}\cdot e^{-2n(1-\epsilon)^2\frac{\log^2(n)}{n^2}}= \\ &= n^{-1+\epsilon}\cdot e^{-2(1-\epsilon)^2\frac{\log^2(n)}{n}}\geq \\ &\geq \frac{1}{2n^{1-\epsilon}}\end{aligned}$$

ולכן $1-\Pr\left(A_1\right)\leq 1-\frac{1}{2n^{1-\epsilon}}$. נציב חזרה ב־ $\Pr\left(A^c\right)$,

$$\begin{aligned}\Pr\left(A^c\right) &= \left(1-P\left(A_1\right)\right)^n \leq \\ &\leq \left(1-\frac{1}{2n^{1-\epsilon}}\right)^n = \\ &= \left(1-\frac{n^\epsilon}{2n}\right)^{\frac{n}{n^\epsilon}n^\epsilon} \to 0\end{aligned}$$

לסיכום, $P\left(A^c\right)\to 0$ כאשר $n\to\infty$ ולכן התוצאה נובעת.

תרגיל 2.28 מבצעים סדרה של הטלות ב"ת של קוביה הוגנת והתוצאה הגבוהה ביותר עד השלב ה־n נקראת "מקסימום זמני בשלב ה־n".

1. מה הסיכוי ש־3 יהיה מקסימום זמני לפחות פעם אחת?
2. מה הסיכוי שכל המספרים מ־1 עד 6 יהיו מקסימום זמני לפחות פעם אחת?

פתרון: מאחר וישנן הרבה אפשרויות נשים לב שהמצב בו 3 הוא מקסימום זמני מתרחש אם ורק אם הוא יוגרל לפני שיוגרלו המספרים $\{4,5,6\}$. הטלות הקובייה יכולות ליפול רק מספר סופי של פעמים על הערכים $1,2$ ולכן אנו למעשה מחפשים את המאורע בו בהטלת קובייה בה לא יצא 1 או 2 התקבלה הספרה 3.

$$\begin{aligned}\Pr\left(\{3\}\,|\,\{1,2\}^c\right) &= \Pr\left(\{3\}\,|\,\{3,4,5,6\}\right) = \\ &= \frac{\Pr\left(\{3\}\right)}{\Pr\left(\{3,4,5,6\}\right)} = \\ &= \frac{\frac{1}{6}}{\frac{4}{6}} = \frac{1}{4}\end{aligned}$$

בכדי למצוא את הפתרון לסעיף ב', אנחנו צריכים להפעיל את אותו הרעיון מסעיף א' אבל הפעם עבור כל ערך. נשים לב ש־

$$\begin{aligned}\Pr\left(\{1\}\,|\,\{1,2,3,4,5,6\}\right) &= \frac{1}{6}; \\ \Pr\left(\{2\}\,|\,\{2,3,4,5,6\}\right) &= \frac{1}{5}; \\ &\vdots \\ \Pr\left(\{5\}\,|\,\{5,6\}\right) &= \frac{1}{2}; \\ \Pr\left(\{6\}\,|\,\{6\}\right) &= 1;\end{aligned}$$

ולכן ההסתברות של המאורע המבוקש היא $\frac{1}{6!}$.

תרגיל 2.29 (סעיף משאלה של מועד א', סמסטר ב' תשע"ד) ישנו מטבע מוטה אשר נופל על 'עץ' (H) בהסתברות $\frac{1}{3}$. מטילים את המטבע שוב ושוב n פעמים באופן ב"ת. לכל $\epsilon>0$ נגדיר את המאורע $A_{n,\epsilon}$ להיות המאורע בו הרצף הארוך ביותר של עצים גדול יותר מ־$(1+\epsilon)\log_3 n$. הוכיחו כי לכל $\epsilon>0$ מתקיים $\Pr\left(A_{n,\epsilon}\right)\to 0$ כאשר $n\to\infty$.

פתרון: יהי $\epsilon > 0$. נגדיר את המספר $\underline{A}_{n,\epsilon} = \lfloor (1+\epsilon)\log_3 n + 1 \rfloor$. נשים לב שהמאורע $A_{n,\epsilon}$ מוכל במאורע $B_{n,\epsilon}$ שמוגדר כך ש־"קיים רצף באורך $\underline{A}_{n,\epsilon}$". זה מתקיים מהסיבה הפשוטה שאם הרצף הארוך ביותר של עצים גדול יותר מ־$(1+\epsilon)\log_3 n$, אזי ישנו רצף באורך $\lfloor (1+\epsilon)\log_3 n + 1 \rfloor$. כעת נחשב את ההסתברות של המאורע $B_{n,\epsilon}$. נשים לב שניתן לפרק את $B_{n,\epsilon}$ לאיחוד של מאורעות לא זרים $B^i_{n,\epsilon}$ כך שכל מאורע הוא המאורע בו ישנו רצף באורך $A_{n,\epsilon}$ שמתחיל במיקום i.

$$\begin{aligned}
\Pr\left(B_{n,\epsilon}\right) &= \Pr\left(\bigcup_{i=1}^{n-A_{n.\epsilon}+1} B^i_{n,\epsilon}\right) \le \\
&\le \sum_{i=1}^{n-A_{n.\epsilon}+1} \Pr\left(B^i_{n,\epsilon}\right) = \\
&= \left(n - A_{n.\epsilon}+1\right)\Pr\left(B^1_{n,\epsilon}\right) = \\
&= \left(n - A_{n.\epsilon}+1\right)3^{-\underline{A}_{n,\epsilon}} \le \\
&\le n3^{-\underline{A}_{n,\epsilon}}.
\end{aligned}$$

נשים לב ש־

$$\underline{A}_{n,\epsilon} > (1+\epsilon)\log_3 n$$
$$3^{\underline{A}_{n,\epsilon}} > 3^{(1+\epsilon)\log_3 n}$$
$$\frac{1}{n^{1+\epsilon}} = 3^{-(1+\epsilon)\log_3 n} > 3^{-\underline{A}_{n,\epsilon}}.$$

ולכן,

$$,\Pr\left(B_{n,\epsilon}\right) < \frac{n}{n^{1+\epsilon}} \to 0$$

כאשר $n \to \infty$ כנדרש.

תרגיל 2.30 (סעיפים משאלה של מועד א', סמסטר ב' תשע"ד) יהי $n \ge 10$. מסדרים את המספרים מ־1 עד n בסדר מקרי הנבחר באופן אחיד מכל הסידורים האפשריים. עבור $2 \le i \le n$, יהי A_i המאורע בו המספר ה־i בשורה גדול יותר מכל קודמיו.

1. מהי $\Pr\left(A_i\right)$?
   (א) $\frac{1}{i!}$.
   (ב) $\frac{(i-1)!(n-i+1)!}{n!}$.
   (ג) $\frac{1}{i}$.
   (ד) אף אחת מהנ"ל.
2. המאורעות A_2, A_3 ו־A_7 הם:
   (א) ב"ת.

(ב) תלויים אך ב"ת בזוגות.

(ג) A_2, A_3 הם ב"ת, אך A_3, A_7 תלויים.

(ד) אף אחת מהנ"ל.

פתרון:

1. נרצה לחשב בעזרת התנייה שתפשט לנו את הפתרון. נשים לב שלכל קבוצה של i מספרים, ישנו רק איבר מקסימאלי יחיד. מאחר והמספרים מסודרים באופן אחיד, המיקום של האיבר המקסימאלי גם כן נקבע באופן אחיד. לכן הסיכוי שהאיבר המקסימאלי יהיה במקום ה־i הוא $\frac{1}{i}$. נגדיר מאורע B_Iכמאורע בו I זו קבוצת הערכים שמופיעים ב־i האיברים הראשונים. על כן,

$$\begin{aligned}
\Pr\left(A_i\right) &= \sum_{I\subseteq\{1,\ldots,n\},|I|=i} \Pr\left(A_i|B_I\right)\Pr\left(B_I\right) = \\
&= \sum_{I\subseteq\{1,\ldots,n\},|I|=i} \frac{1}{i}\Pr\left(B_I\right) = \\
&= \frac{1}{i}\sum_{I\subseteq\{1,\ldots,n\},|I|=i} \Pr\left(B_I\right) = \frac{1}{i}.
\end{aligned}$$

2. נבצע חישוב מפורש עבור A_2 ו־A_3. נשים לב שלכל שלשת מספרים שנבחר (x,y,z), ישנו רק סידור יחיד בו גם A_2 מתרחש וגם A_3 מתרחש מתוך סה"כ $3!$ סידורים אפשריים. לכן, אם נתנה על שלושת הערכים הראשונים בשורה נקבל ש־

$$\begin{aligned}
\Pr\left(A_2\cap A_3\right) &= \sum_{I\subseteq\{1,\ldots,n\},|I|=3} \Pr\left(A_2\cap A_3|B_I\right)\Pr\left(B_I\right) = \\
&= \sum_{I\subseteq\{1,\ldots,n\},|I|=3} \frac{1}{3!}\Pr\left(B_I\right) = \\
&= \frac{1}{6} = \Pr\left(A_2\right)\Pr\left(A_3\right).
\end{aligned}$$

באופן כללי, נשים לב שלכל סידור של $i-1$ איברים ראשונים בשורה, הסיכוי שהמקסימום יהיה במקום ה־i נשאר $\frac{1}{i}$ לכן גם במקרה הכללי אנחנו צפויים לקבל אי־תלות.

$$\begin{aligned}
\Pr\left(A_7\cap A_3\right) &= \sum_{I\subseteq\{1,\ldots,n\},|I|=7} \Pr\left(A_7\cap A_3|B_I\right)\Pr\left(B_I\right) = \\
&= \sum_{I\subseteq\{1,\ldots,n\},|I|=3} \frac{\binom{6}{3}\cdot 1\cdot 2!3!\cdot 1}{7!}\Pr\left(B_I\right) = \\
&= \frac{1}{21} = \Pr\left(A_7\right)\Pr\left(A_3\right),
\end{aligned}$$

כאשר הספירה מתבצעת באופן הבא: ספירה של מספר השלשות הראשונות האפשריות מתוך שבעת המספרים הראשונים כאשר מורידים מהספרות את המספר המקסימאלי

שיוצב בקצה השורה, מתוך השלשה שנבחרה מיקום המקסימאלי במקום השלישי (אפשרות אחת), סידור צמד המספרים הראשונים (2 אפשרויות), סידור המספרים שבין המיקום ה־4 ל־6 (6 אפשרויות) ומיקום המקסימאלי בקצה השורה.

תרגיל 2.31 (סעיפים משאלה של מועד ב', סמסטר א' תשע"ה) יהיו $1 \leq k \leq n$ מספרים שלמים. בספארי נולדו לקרנף n גורים אחד אחרי השני כאשר כל גור יכל להיוולד לכל מין בהסתברות שווה ובאופן ב"ת במין הגורים האחרים. דוליטול מסתובב בגן החיות ופוגש גור שנבחר באקראי באופן אחיד מבין הגורים. נסמן ב־A את המאורע בו k הגורים הראשונים שנולדו הם ממין זכר. נסמן ב־B את המאורע שנולדו לפחות k גורים ממין זכר. נסמן ב־C את המאורע בו דוליטול פוגש גור ממין זכר.

1. עבור $k=1, n=2$, מהי $\Pr(C|B)$?
   (א) $\frac{1}{2}$.
   (ב) $\frac{2}{3}$.
   (ג) $\frac{3}{4}$.
   (ד) אף אחת מהנ"ל.
2. עבור $k=4, n=12$, מהי $\Pr(A|C)$?
   (א) $\frac{1}{12}$.
   (ב) $\frac{1}{16}$.
   (ג) $\frac{1}{24}$.
   (ד) אף אחת מהנ"ל.

פתרון:

1. נסמן ב־D_i את המאורע בו נולדו i גורים ממין זכר. נעזר בפירוק של מרחב המדגם על פי המאורעות הללו.

$$\begin{aligned}
\Pr(C|B) &= \frac{\Pr(C\cap B)}{\Pr(B)} = \\
&= \frac{\sum_{i=0}^{2}\Pr(C\cap B\cap D_i)}{\Pr(B)} = \\
&= \frac{\Pr(C\cap B\cap D_1)+\Pr(C\cap B\cap D_2)}{\Pr(B)} = \\
&= \frac{\Pr(C\cap D_1)+\Pr(C\cap D_2)}{1-\Pr(B^c)} = \\
&= \frac{\Pr(C|D_1)\Pr(D_1)+\Pr(C|D_2)\Pr(D_2)}{1-\Pr(B^c)} = \\
&= \frac{\frac{1}{2}\cdot 2\cdot\left(\frac{1}{2}\right)^2+1\cdot\left(\frac{1}{2}\right)^2}{1-\left(\frac{1}{2}\right)^2} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3}.
\end{aligned}$$

2. נשתמש בחוק בייס לפתרון התרגיל. נסמן ב־E את המאורע בו דוליטול פוגש אחד מארבעת הגורים הראשונים שנולדו.

$$\begin{aligned}
\Pr\left(A|C\right) &= \frac{\Pr\left(C|A\right)\Pr\left(A\right)}{\Pr\left(C\right)} = \\
&= \frac{\left[\Pr\left(C|A\cap E\right)\Pr\left(E|A\right)+\Pr\left(C|A\cap E^{c}\right)\Pr\left(E^{c}|A\right)\right]\cdot\left(\frac{1}{2}\right)^{4}}{\frac{1}{2}} = \\
&= \frac{\left[1\cdot\frac{4}{12}+\frac{1}{2}\cdot\frac{8}{12}\right]\cdot\left(\frac{1}{2}\right)^{4}}{\frac{1}{2}} = \frac{1}{12}.
\end{aligned}$$

3 משתנים מקריים

הגדרה 3.1 משתנה מקרי (מ"מ) היא התאמה של ערך מספרי לארועים האפשריים במרחב הסתברות, $X:\Omega\rightarrow\mathbb{R}$.

דוגמה 3.2 (בעיית המעטפות והמכתבים) תזכורת־ ישנם 5 מכתבים, לכל מכתב ישנה מעטפה שמתאימה לו. המזכירה מפזרת את 5 המכתבים במעטפות באקראי. נוכל להגדיר פונקציה כלשהי, $X:\Omega\rightarrow\mathbb{R}$ המתאימה לכל מאורע את מספר המכתבים בו שנכנסו למעטפה הנכונה. אזי, X משתנה מקרי.

הגדרה 3.3 (התפלגות) התפלגות זו פונקציה $\Pr:\mathbb{R}\rightarrow[0,1]$ כך שקבוצת המספרים A עליה הפונקציה שונה מאפס היא סופית או בת־מנייה ובנוסף

$$\sum_{n\in A}\Pr(n)=1\,.$$

לקבוצת המספרים A שכזאת אנחנו נקרא התומך ($support$) של ההתפלגות P. ההסתברויות של כל הערכים שמשתנה מקרי ממשי X מקבל, מגדירות את **ההתפלגות של** X. התפלגות זאת קובעת את ההסתברות שהמ"מ X יהיה שווה לערך k כלשהו.

כלומר, אם למשל $X:\Omega\rightarrow\{1,\ldots,n\}$, ההתפלגות של X היא רשימת ההסתברויות:

$$\Pr(X=1)\,,\ \Pr(X=2)\,,\ldots,\ \Pr(X=n)\,.$$

הערה 3.4 בדרך כלל אנו נשתמש בכתיב המקוצר $\{X=k\}$ בכדי לציין את המאורע בו המשתנה המקרי X מקבל את הערך k. יחד עם זאת, חשוב להבין את המשמעות האמיתית שמאחורי סימון זה. מדובר פה בכל המצבים האפשריים במרחב המדגם $\{\omega_i\}_{i\in I}\subset\Omega$ כך ש־ $X(\omega_i)=k$ או במדויק

$$\{X=k\}=\{\omega\in\Omega|X(\omega)=k\}\,.$$

אך בכדי לקצר את הסימון, אנו נשתמש לרוב בסימון המקוצר.

תרגיל 3.5 נטיל 2 קוביות הוגנות (עם חשיבות לסדר). נגדיר משתנה מקרי X להיות התוצאה המקסימלית (בפרט, $X\in\{1,\ldots,6\}$). מהי התפלגות X?

פתרון: נרצה למצוא את $\Pr(X=k)$ עבור $k\in\{1,\ldots,6\}$. עבור $k=1$, המאורע היחידי שייתן את הנדרש הוא המאורע בו שתי הקוביות נותנות 1, לכן

$$\Pr(X=1)=\frac{1}{36}\,.$$

עבור $k=2$, המאורעות המתאימים הם אלו בהם נקבל

$$(1,2)\,,(2,1)\ \text{ or }\ (2,2)\,,$$

לכן

$$\Pr\left(X=2\right)=\frac{3}{36}\,.$$

באופן כללי, המאורעות שייתנו תוצאה מקסימלית k הם

$$\left(j,k\right),\left(k,j\right),\left(k,k\right)\qquad j<k\,.$$

לכן, ההסתברות היא

$$\Pr\left(X=k\right)=\frac{2\left(k-1\right)+1}{36}=\frac{2k-1}{36}\,.$$

דרך פתרון נוספת ־ מאחר ו־X הוא מקסימום, קל יותר לחשב את ההסתברות שהוא קטן או שווה לערך כלשהו מאשר רק שווה לערך כלשהו. כלומר,

$$\begin{aligned}\Pr\left(X=k\right)=\Pr\left(X\leq k\right)-\Pr\left(X\leq k-1\right)&=\\ =\frac{k^2}{6^2}-\frac{\left(k-1\right)^2}{6^2}&=\\ =\frac{k^2-k^2+2k-1}{36}&=\frac{2k-1}{36}.\end{aligned}$$

כעת נראה קבוצה של התפלגויות שכיחות בעולם ההסתברות.

3.1 משתנה גאומטרי

מבצעים סדרה של ניסויים בלתי תלויים כך שכל ניסוי מצליח בסיכוי p, ונכשל בסיכוי $1-p$. נגדיר, X סך כל הניסויים עד (כולל) ההצלחה הראשונה. אזי $X:\Omega\to\mathbb{N}\setminus\{0\}$, ולכל $k\in\mathbb{N}\setminus\{0\}$,

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

נסמן $X\sim G\left(p\right)$.

תרגיל 3.6 נזכר בדוגמא נוספת שראינו כבר־ קופסא בה ישנם 3 כדורים שחורים, 4 כדורים אדומים ו־93 כדורים לבנים. מציאים כדור בכל שלב עם החזרה ועוצרים בפעם הראשונה בה מתקבל כדור שאינו לבן. הפעם נסמן ב־ Y את מספר ההוצאות עד לעצירה (כולל האחרונה). מאחר והוכחנו $P\left(Y=\infty\right)=0$, Y מ"מ. מהי התפלגותו?

פתרון: מתקיים $\Pr:\Omega\to\mathbb{N}\setminus\{0\}$. עבור $k\in\mathbb{N}\setminus\{0\}$ ההסתברות שהוצאנו בדיוק k כדורים עד לעצירה היא

$$\Pr\left(Y=k\right)=\left(\frac{93}{100}\right)^{k-1}\cdot\frac{7}{100}\,,$$

הוצאת $k-1$ כדורים לבנים, ואז כדור שאינו לבן. כלומר, זהו משתנה מקרי גאומטרי עם $p=\frac{7}{100}$, $Y\sim G\left(\frac{7}{100}\right)$.

הערה 3.7 עבור $Z:\Omega\to\mathbb{R}$ מ"מ בהכרח מתקיים השיוויון

$$\sum_{k\in\mathbb{R}}\Pr\left(Z=k\right)=1.$$

האם גם עבור $Z\sim G\left(p\right)$, $0<p<1$, יתקיים שיוויון (בפרט, האם לכל p כנ"ל נקבל מ"מ חוקי)? כן! נראה זאת:

$$\begin{aligned}\sum_{k=1}^{\infty}\Pr\left(X=k\right) & =\sum_{k=1}^{\infty}p\left(1-p\right)^{k-1}=p\sum_{k=1}^{\infty}\left(1-p\right)^{k-1}=\\ & =p\sum_{l=0}^{\infty}\left(1-p\right)^{l}=p\cdot\frac{1}{1-\left(1-p\right)}=\\ & =p\cdot\frac{1}{p}=1\end{aligned}$$

כאשר המעבר השלישי הוא בעזרת ההתאמה $l=k-1$, והרביעי מתקבל מסכום סדרה הנדסית.

3.2 משתנה אחיד

משתנה מקרי X אשר מקבל ערכים ב־ $\{1,\ldots,n\}$ כך ש־ $P\left(X=k\right)=\frac{1}{n}$ לכל k נקרא מ"מ אחיד. נסמן $Y\sim U\left(n\right)$ (לפעמים גם $U\left(\{1,\ldots,n\}\right)$ ובאופן כללי $U\left(I\right)$ לכל קבוצה סופית I שכל איבר בה מקבל את אותה הסתברות).

תרגיל 3.8 בכד 9 כדורים לבנים ו־ 1 שחור. מושכים באקראי בזה אחר זה כדור מהכד. נסמן ב־ X את מספר המשיכות עד לקבלת שחור וב־ Y את מספר המשיכות עד לקבלת 3 שחורים.

1. מצאו את התפלגות X כאשר המשיכה היא עם החזרה.
2. מצאו את התפלגות X כאשר המשיכה היא ללא החזרה.
3. מצאו את התפלגות Y כאשר המשיכה היא עם החזרה.

פתרון: נענה על כל סעיף בנפרד.

1. X מקבל ערכים ב־ $\mathbb{N}\setminus\{0\}$. בכל הוצאה ישנה הסתברות של $\frac{9}{10}$ לכדור לבן (כישלון) והסתברות של $\frac{1}{10}$ לכדור לבן (הצלחה). כלומר, קיבלנו מ"מ גאומטרי, $X\sim G\left(\frac{1}{10}\right)$.
2. X מקבל ערכים ב־ $\{1,\ldots,10\}$, כאשר מטעמי סימטריה לכל $1\le k\le10$

$$\Pr\left(X=k\right)=\frac{1}{10}$$

(ההסתברות של כל כדור לצאת במשיכה ה־ k הוא $\frac{1}{10}$, שימו לב־ אנו לא יודעים איזה כדור יצא במשיכות $\{1,\ldots,k-1\}$).

3. Y מקבל ערכים ב־ $\{3,4,5,\dots\}$. יהי $3 \le k$, נחשב

$$\Pr(Y=k) = \binom{k-1}{2}\left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right)^{k-3}\frac{1}{10} = \binom{k-1}{2}\left(\frac{1}{10}\right)^3\left(\frac{9}{10}\right)^{k-3}$$

כאשר $\binom{k-1}{2}$ הוא מספר הבחירות ל־ 2 המקומות בהם יצאו כדורים שחורים ב־ $k-1$ ההוצאות הראשונות, $\left(\frac{1}{10}\right)^2$ היא ההסתברות ל־ 2 שחורים, $\left(\frac{1}{10}\right)^{k-3}$ ההסתברות ל־ $k-3$ לבנים (בשאר ההוצאות), ו־ $\frac{1}{10}$ היא ההסתברות לכדור שחור בהוצאה האחרונה.

ישנה דרך נוספת להסתכל על שאלה זו ־ אם עבור $i \in \{1,2,3\}$ נסמן ב־ Y_i את מספר ההוצאות עד לקבלת שחור בפעם ה־ i לאחר השחור ה־ $i-1$ (כאשר Y_1 הוא בדיוק X), מתקיים

$$.Y_i \sim G\left(\frac{1}{10}\right) \quad \text{and} \quad Y = Y_1+Y_2+Y_3$$

האם יש לנו דרך לדעת מהי ההתפלגות של סכום מ"מ?

3.3 משתנה בינומי שלילי

מבצעים סדרה של ניסויים בלתי תלויים. כל ניסוי מצליח בסיכוי p ונכשל בסיכוי $1-p$. X מספר הניסויים עד ל־ r, $r \in \mathbb{N}\setminus\{0\}$, הצלחות. אזי ערכי X הינם $k \in \{r, r+1, \dots\}$ ומתקיים

$$.\Pr(X=k) = \binom{k-1}{r-1}p^r(1-p)^{k-r}$$

ל־ Y קוראים מ"מ בינומי שלילי ומסמנים $X \sim NB(r,p)$. בדוגמא שלנו־

$$.Y \sim NB\left(3, \frac{1}{10}\right)$$

טענה 3.9 אם X מתפלג בינומי שלילי אזי ניתן להציג אותו כסכום של משתנים גאומטריים עם פרמטר זהה.

הוכחה: עבור אותם ניסויים מהגדרת X, נגדיר משתנים מקריים:

$X_1 =$ מספר הניסויים עד להצלחה הראשונה.

$X_i =$ מספר הניסויים לאחר ההצלחה ה־ $i-1$ ועד להצלחה ה־ i.

אזי ברור כי מתקיים

$$.X_i \sim G(p) \quad \text{and} \quad X = \sum_{i=1}^{r} X_i$$

■

הערה 3.10 גם ההפך נכון. אם $X_i \sim G(p)$ (עם אותו p) הנקבעים על בסיס מאורעות בלתי תלויים, $1 \le i \le r$, אזי מתקיים

$$\sum_{i=1}^{r} X_i \sim NB(r,p).$$

3.4 משתנה בינומי

מבצעים סדרה של n ניסויים בלתי תלויים כך שכל ניסוי מצליח בסיכוי p, ונכשל בסיכוי $1-p$. נגדיר X מספר ההצלחות. אזי $X:\Omega \to \{0,\ldots,n\}$, ולכל $k \in \{1,\ldots,n\}$,

$$\Pr(X=k) = \binom{n}{k} p^k \cdot (1-p)^{n-k}.$$

נסמן $X \sim Bin(n,p)$ (יכול להופיע גם $B(n,p)$).

תרגיל 3.11 (הילוך שיכור) שיכור הולך בכל צעד ימינה מטר אחד בהסתברות p ושמאלה בהסתברות $1-p$, במשך N צעדים. נגדיר:

$X =$ מיקום השיכור אחרי N צעדים,

$Z =$ מספר הקפיצות ימינה.

מצאו את התפלגות Z ו־ X.

פתרון: ראשית, מתקיים

$$Z \sim Bin(N,p).$$

בנוסף, נשים לב כי $N-Z$ הוא מספר הצעדים שמאלה. לכן

$$X = Z-(n-Z) = 2Z-n$$

(מספר הצעדים ימינה פחות מספר הצעדים שמאלה). כעת, נוכל לחשב את התפלגות X. X מקבל ערכים $\{-N,\ldots,0,\ldots,N\}\in, k$ ו־

$$\Pr(X=k) = \Pr(2Z-N=k) = \Pr\left(Z=\frac{N+k}{2}\right)$$

נשים לב שעבור k שאינו מאותה זוגיות כמו N קיבלנו $\Pr(X=k)=0$. עבור k מאותה זוגיות של N מהתפלגות בינומית נקבל

$$\Pr(X=k) = \binom{N}{\frac{N+k}{2}} p^{\frac{N+k}{2}} (1-p)^{\frac{N-k}{2}}.$$

כלומר, סה"כ

$$\Pr(X=k) = \begin{cases} \binom{N}{\frac{N+k}{2}} p^{\frac{N+k}{2}} (1-p)^{\frac{N-k}{2}} & \frac{N-k}{2} \text{ is even}, |k|\le N \\ 0 & \text{else} \end{cases}.$$

תרגיל 3.12 בבריכת דגים 5 דגי זהב 4 כרישים ו־ 11 קרפיונים. דייג דג 6 מהם (ללא החזרה). X מספר דגי הזהב שדג בחכתו. מהי התפלגות X?

פתרון: X מקבל ערכים ב־ $k \in \{0,\dots,5\}$. נחשב

$$\Pr\left(X=k\right)=\frac{\binom{5}{k}\binom{15}{6-k}}{\binom{20}{6}},$$

כיוון שסך הבחירות האפשריות $\binom{20}{6}$, וכדי שיהיו בדיוק k דגי זהב נבחר $\binom{5}{k}$ דגי זהב ו־ $\binom{15}{6-k}$ משאר הדגים.

התפלגות זו נקראת התפלגות היפר־גאומטרית.

3.5 משתנה היפר־גאומטרי

תהי קבוצת עצמים בגודל N כאשר $D\le N$ מתוכם הם מיוחדים. דוגמים ללא החזרה n $(n\le N)$ עצמים מתוך הקבוצה. H מספר העצמים המיוחדים במדגם. מתקיים

$$\Pr\left(H=k\right)=\frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}$$

(באותו אופן כמו בתרגיל האחרון). H נקרא מ"מ היפר־גאומטרי. נסמן $H\sim H\left(N,D,n\right)$.

הערה 3.13 בתרגיל האחרון מתקיים $X\sim H\left(20,5,6\right)$.

תרגיל 3.14 לדני יש 5 מטבעות הוגנים ובלתי תלויים. הוא מטיל את המטבעות. אם ישנם מטבעות שהוא קיבל בהם "פלי" הוא מסלק אותם ומטיל שוב את שאר המטבעות. כך הלאה בכל שלב. מה הסיכוי שבסבב ההטלות השני הוא יקבל בדיוק 3 "עץ".

פתרון: נסמן ב־ Y_i, $i=1,2$, את מספר ה"עץ"ים בסבב ה־ i.
נוכל לפתור בעזרת נוסחאת ההסתברות השלמה

$$\Pr\left(Y_2=3\right)=\sum_{k=0}^{5}\Pr\left(Y_2=3\mid Y_1=k\right)\Pr\left(Y_1=k\right).$$

לכל k מתקיים

$$\Pr\left(Y_1=k\right)=\binom{5}{k}\left(\frac{1}{2}\right)^{k}\left(\frac{1}{2}\right)^{5-k}=\binom{5}{k}\left(\frac{1}{2}\right)^{5}$$

לבחור את k המטבעות שיצאו "עץ" כפול ההסתברויות שכל אחד מהמטבעות יצא כפי שבחרנו (בפרט, $Y_1\sim Bin\left(5,\frac{1}{2}\right)$). נשים לב שעבור $k<3$ מתקיים

$$\Pr\left(Y_2=3\mid Y_1=k\right)=0,$$

כיוון שאחרי הסבב הראשון נשאר עם מעט מידי מטבעות. עבור $k\in\{3,4,5\}$ מתקיים

$$\Pr\left(Y_2=3\mid Y_1=k\right)=\binom{k}{3}\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{k-3}=\binom{k}{3}\left(\frac{1}{2}\right)^{k},$$

באותו אופן כמו מקודם. נשים לב שקיבלנו, $Y_2 \mid Y_1 = k \sim Bin\left(k, \frac{1}{2}\right)$. לכן סה"כ

$$\begin{aligned}\Pr\left(Y_2 = 3\right) &= \sum_{k=3}^{5} \binom{5}{k} \left(\frac{1}{2}\right)^5 \binom{k}{3} \left(\frac{1}{2}\right)^k = \\ &= \binom{5}{3} \left(\frac{1}{2}\right)^8 + \binom{5}{4} \left(\frac{1}{2}\right)^5 \binom{4}{3} \left(\frac{1}{2}\right)^4 + \binom{5}{5} \left(\frac{1}{2}\right)^5 \binom{5}{3} \left(\frac{1}{2}\right)^5 = \\ &= \frac{45}{512}\end{aligned}$$

דרך II לפתרון: נתבונן במטבע בודד כלשהו. אזי ההסתברות שהוא הוטל "עץ" בסבב השני היא $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$. מאחר והמטבעות בלתי תלויים מתקיים

$$Y_2 \sim Bin\left(5, \frac{1}{4}\right),$$

ולכן

$$\Pr\left(Y_2 = 3\right) = \binom{5}{3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2 = \frac{45}{512}.$$

עד כה ראינו מספר משתנים מקריים ולכל משנה מקרי ניתן להגדיר התפלגות.

דוגמה 3.15 יהי $X \sim Bin\left(n, p\right)$ ונגדיר התפלגות בינומית ע"י $\Pr(k) = \binom{n}{k} p^k \left(1-p\right)^{n-k}$ לכל $k \in A = \{0, 1, \ldots, n\}$. באופן דומה נוכל להגדיר התפלגות גיאומטרית עם פרמטר p על ידי $\Pr(k) = \left(1-p\right)^{k-1}$ לכל $k \in \mathbb{N} \setminus \{0\}$ וכן הלאה.

כאשר נגיד שיש משתנה מקרי X שמתפלג על פי התפלגות בינומית / גיאומטרית / בינומית־שלילית וכן הלאה, הכוונה היא שקבוצת הערכים עליה מוגדר X היא אותה קבוצת ערכים עליה מוגדרת ההתפלגות וכן שההסתברות של X לקבל כל ערך נתונה על ידי ההתפלגות. בצורה כזאת לא נצטרך להגדיר בכל פעם מחדש מרחב מדגם, תיאור ניסוי וכן הלאה ונוכל לדון במקרים כלליים יותר.

3.6 משתנים מקריים דו־מימדיים והתפלגות משותפת

הגדרה 3.16 עבור X ו־ Y משתנים מקריים, נגדיר משתנה מקרי חדש, דו־מימדי

$$(X, Y) : \Omega_1 \times \Omega_2 \to \mathbb{R}^2$$

כאשר פונקציית ההסתברות (המשותפת של המשתנים המקריים X ו־Y) מוגדרת לכל $x, y \in \mathbb{R}$ ע"י

$$\Pr\left(X = x, Y = y\right) = \Pr\left(\{X = x\} \cap \{Y = y\}\right)$$

והתנאי

$$\sum_{x,y} \Pr\left(X = x, Y = y\right) = 1$$

מתקיים.

באופן דומה לאיך שהגדרנו התפלגות במקרה החד־מימדי נוכל גם להגדיר התפלגות דו־מימדית.

הגדרה 3.17 (התפלגות דו־מימדית) התפלגות דו־מימדית זו פונקציה $\Pr:\mathbb{R}^2\to[0,1]$ כך שקבוצת הוקטורים $A\subsetneq\mathbb{R}^2$ עליה הפונקציה שונה מאפס היא סופית או בת־מנייה ובנוסף

$$\sum_{(m,n)\in A}\Pr(m,n)=1\,.$$

לכל משתנה מקרי דו־מימדי ישנה התפלגות דו־מימדית ולהיפך. ההתפלגות מתארת את הערכים שהמשתנה המקרי מקבל בהסתברות חיובית (שימו לב שהערכים הם וקטורים) ואת ההסתברות לקבל כל ערך.

תרגיל 3.18 מטילים שתי קוביות הוגנות, לבנה ושחורה. X התוצאה של הקוביה הלבנה, Y התוצאה המקסימאלית בין השתיים. מצאו את ההתפלגות המשותפת של (X,Y) ואת ההתפלגויות השוליות.

פתרון: צריך למצוא לכל $k,l\in\{1,\ldots,6\}$ את

$$\Pr(X=k,Y=l)=\Pr(\{X=k\}\cap\{Y=l\})\,.$$

נחשב בטבלה (כאשר כל תא בטבלה מתאר את ההסתברות שמתאימה לשורה והעמודה הרלוונטיים, קרי $\Pr(\{X=k\}\cap\{Y=l\})$):

$Y\setminus X$	1	2	3	4	5	6	$Y\sim$
1	$\frac{1}{36}$	0	0	0	0	0	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{2}{36}$	0	0	0	0	$\frac{3}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{3}{36}$	0	0	0	$\frac{5}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{4}{36}$	0	0	$\frac{7}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{5}{36}$	0	$\frac{9}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$	$\frac{11}{36}$
$X\sim$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Table 1: טבלת התפלגות משותפת של משתנים מקריים X ו־ Y.

הסבר:

- עבור $k>l$ (כלומר $X>Y$) לא ייתכן שבקוביה הלבנה נקבל תוצאה הגבוהה יותר ממקסימום הקוביות, לכן $\Pr(X=k,Y=l)=0$.
- עבור $k=l$ (כלומר $X=Y$) בקוביה הלבנה צריך לצאת k, ובשחורה יכול לצאת כל מספר הקטן או שווה ל־ k, לכן

$$\Pr(X=k,Y=l)=\frac{1}{6}\cdot\frac{k}{6}=\frac{k}{36}\,.$$

- עבור $k<l$ (כלומר $X<Y$) בקוביה הלבנה צריך לצאת k ובשחורה l, לכן

$$.\Pr\left(X=k,Y=l\right)=\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$$

הערה 3.19 כפי שרואים בטבלה, ניתן למצוא את התפלגויות X או Y מסכימת עמודות או שורות הטבלה בהתאמה. בנוסף, אנחנו יכולים לוודא אם התשובה שלנו נכונה על ידי סכימה ובדיקה שסכום ההסתברויות שווה ל־1.

שימו לב שלרוב אנו כותבים את הכתיב המקוצר $\Pr\left(X=k\right)$ ללא ציון מרחב המדגם והתייחסות אליו כלל. בצורה המלאה צריך לרשום את הביטוי כ־

$$.\Pr\left(X=k\right)=\Pr\left(\left\{\omega\in\Omega|X\left(\omega\right)=k\right\}\right)$$

3.6.1 הילוך מקרי ושיטת השיקוף

בתרגולים קודמים דנו בהילוכים מקריים ובהתפלגות על נקודות הקצה שהילוך באורך n מייצר. כעת אנו מעוניינים להסתכל על נקודות הסיום של ההילוך ובמקביל על הערך המקסימאלי אליו יגיע ההילוך.

באופן פורמאלי, יהי הילוך מקרי פשוט באורך n. נסמן ב־S_n את המיקום הסופי של ההילוך לאחר n צעדים וב־M_n את המיקום המקסימאלי אליו הגיע ההילוך.

תרגיל 3.20 מצאו את ההתפלגות המשותפת של (S_n,M_n).

פתרון: נתחיל עם המצבים הקלים עבור $\Pr\left(S_n=s,\ M_n=m\right)$. ההסתברות תהיה אפס בהינתן אחד המקרים הבאים:

1. $s>m$.
2. $s<-n$.
3. $m<0$.
4. $m>n$.
5. $m,s\notin\mathbb{Z}$.
6. ל־s ול־n אין אותה הזוגיות (האחד זוגי והאחר אי־זוגי).

בכל אחד מהמקרים הנ"ל נקבל ש־ $P\left(S_n=s,\ M_n=m\right)=0$ ולכן נניח שהם אינם מתקיימים. קרי, ל־s ול־n יש אותה הזוגיות, $-n\le s\le m\le n$, $0\le m$, $s,m\in\mathbb{Z}$.

נגדיר צמד מאורעות: $A=\{M_n\geq m,\ S_n=s\}$ ו־$B=\{S_n=2m-s\}$, ונראה שההסתברות של מאורעות אלו שווה. כיצד זה יעזור לנו? אנחנו יודעים לחשב את $\Pr(B)$ באופן הבא:

$$\Pr(B)=\Pr(S_n=2m-s)=\frac{|\{S_n=2m-s\}|}{2^n}=$$
$$=\frac{\binom{n}{\frac{n}{2}+\frac{2m-s}{2}}\binom{\frac{n}{2}-\frac{2m-s}{2}}{\frac{n}{2}-\frac{2m-s}{2}}}{2^n}=\frac{\binom{n}{\frac{n+2m-s}{2}}}{2^n}$$

שימו לב ש־$\frac{n+2m-s}{2}\in\mathbb{Z}$, $0\leq\frac{n+2m-s}{2}\leq n$. כעת נוכל להשתמש במעבר הבא ולקבל את התוצאה המבוקשת.

$$\begin{aligned}\Pr(S_n=s,\ M_n=m) &= (S_n=s,\ M_n\geq m)-(S_n=s,\ M_n\geq m+1)=\\ &= \frac{\binom{n}{\frac{n+2m-s}{2}}}{2^n}-\frac{\binom{n}{\frac{n+2(m+1)-s}{2}}}{2^n}=\\ &= \frac{\binom{n}{\frac{n+2m-s}{2}}-\binom{n}{\frac{n+2m-s}{2}+1}}{2^n}\end{aligned}$$

בכדי להראות ש־ $\Pr(A)=\Pr(B)$ נשתמש בטכניקה שנקראת "עיקרון השיקוף". אנחנו נבצע שיקוף של הילוכים מקריים במאורע אחד למאורע אחר ובאופן חד־חד ערכי ועל ובכך נוכיח את הטענה. לשם כך נגדיר משתנה מקרי חדש $T:=\min\{i:\ S_i=m\}$ זה משתנה מקרי חדש שמציין את השלב הראשון בו ההילוך הגיע לגובה m. שימו לב שייתכן ואין כזה שלב ולכן במקרים כאלה נגדיר $T:=n$. יחד עם זאת, לכל מאורע ב־A קיים שלב כזה.

נגדיר את המיפוי הבא:

$$f:\Omega\to\Omega,$$
$$f(\omega)=f(\omega_1,\ldots,\omega_n)=\left(\omega_1,\ldots,\omega_{T(\omega)},-\omega_{T(\omega)+1},\ldots,-\omega_n\right).$$

מה שהפונקציה עושה זה בעצם לשקף את ההילוך מהרגע הראשון שהוא מגיע לגובה m, ואם הוא לא מגיע אז הפונקציה לא עושה דבר.

טענה 3.21 f מעתיקה את A ל־B באופן חח"ע ועל.

הוכחה: שימו לב ש־ $f(f(\omega))=\omega$ וזה נובע מאחר ו־ $T(f(\omega))=T(\omega)$ (הזמן הראשון בו מגיע ההילוך לגובה m לא משתנה תחת הפונקציה f) ולכן מאותו שלב והלאה הפונקציה, כאשר היא מופעלת על עצמה, מבטלת את השינויים הראשונים שביצעה ב־ω. מאחר ו־ $f(f(\omega))=\omega$ נובע שהפונקציה חח"ע. יהי $f(\omega_1)=f(\omega_2)$ אזי נפעיל את הפונקציה על צמד האגפים ונקבל ש־

$$\omega_1=f(f(\omega_1))=f(f(\omega_2))=\omega_2.$$

בכדי להראות שכל $\omega\in A$ מועתקת ל־$f(\omega)\in B$ ניקח $\omega\in A$. לכן,

$$S_{T(\omega)}=m,\ S_n(\omega)=s.$$

מכאן נסיק כי $S_n(f(\omega))=2m-s$ (לאחר שלב $T(\omega)$ ההילוך עתיד לבצע $m-s$ יותר צעדים כלפי מטה מאשר כלפי מעלה, לאחר השיקוף $f(\omega)$ יבצע את אותה כמות צעדים

עודפים כלפי מעלה ולכן יסיים במיקום $2m - s$). בכדי לראות שהפוקנציה הינה על ניקח $\omega' \in B$, נגדיר $\omega = f(\omega')$ ונראה ש־ $\omega \in A$. בגלל התכונה $f(f(\omega)) = \omega$ נובע ש־

$$f(\omega) = f(f(\omega')) = \omega' \in B.$$

אם $\omega' \in B$, אזי

$$M_n(\omega') \geq S_n(\omega') = 2m - s \geq m,$$

לכן

$$\begin{aligned} M_n(\omega) &= M_n(f(\omega')) \geq m, \\ S_n(\omega) &= S_n(f(\omega')) = s \end{aligned}$$

כנדרש ולכן הטענה נובעת. ■

3.7 התפלגות מותנה

ראינו כבר את ההגדרה להסתברות מותנה. בחלק זה נדבר על ההתפלגות המותנה, כלומר התפלגות $X \mid Y = y$ עבור y כלשהו.

תרגיל 3.22 בן נוסע לקזינו בלאס וגאס להמר על קוביות. הוא מהמר על קובייה הוגנת המוטלת שוב ושוב. נגדיר את X להיות מספר ההטלות עד שיוצא לראשונה המספר 6, ו־Y מספר ההטלות עד שיוצא לראשונה מספר זוגי קטן מ־6, ז"א 2 או 4.

1. מצאו את ההתפלגות המשותפת של (X, Y).
2. מצאו את ההתפלגויות השוליות של X ו־ Y.
3. מצאו את ההתפלגות של $W = \min\{X, Y\}$.
4. מצאו כיצד מתפלג $Y | X = 2$.

פתרון: נענה על כל סעיף בנפרד.

1. מאחר ומדובר במספר לא מוגבל של הטלות, המשתנים המקריים יכולים לקבל ערכים הולכים וגדלים ולכן לא נוכל לרשום טבלה שתתאר את ההתפלגות שלהם באופן מוחלט. לכן נבחר להציג את ההתפלגות המשותפת בעזרת נוסחה.

$$\Pr(X = k, Y = l) = \begin{cases} \left(\frac{3}{6}\right)^{l-1}\left(\frac{2}{6}\right)^{1}\left(\frac{5}{6}\right)^{k-(l+1)}\left(\frac{1}{6}\right)^{1} & k > l \\ & k, l \in \mathbb{N} \setminus \{0\} \\ \left(\frac{3}{6}\right)^{k-1}\left(\frac{1}{6}\right)^{1}\left(\frac{4}{6}\right)^{l-(k+1)}\left(\frac{2}{6}\right)^{1} & l > k \end{cases}.$$

הסבר: כאשר $k > l$ אז בסדרת ההטלות יצא תחילה 2 או 4 (בהטלה ה־l) ורק לאחר מכן יצא 6 (בהטלה ה־k). לכן ב־$l - 1$ ההטלות הראשונות קיבלנו רק $1, 3, 5$ (ההסתברות לכך היא $\left(\frac{3}{6}\right)^{l-1}$), בהטלה ה־$l$ קיבלנו 2 או 4 (ההסתברות לכך היא $\frac{2}{6}$), ומאותה הטלה והלאה עד ההטלה ה־k יכלו לצאת כל המספרים למעט 6 (בהסתברות $\left(\frac{5}{6}\right)^{k-l-1}$). בהטלה ה־$k$ בהכרח יצא 6 (בהסתברות $\frac{1}{6}$). החישוב השני מתבצע באופן דומה.

2. בכדי לענות על הסעיף השני, אין צורך להתייחס לסעיף הראשון. למעשה מדובר בסדרה של הטלות בלתי־תלויות וצמד המשתנים המקריים סופרים את מספר ההטלות עד לפעם הראשונה שמתרחש תנאי מסויים, בפרט עד שיוצא 2 או 4 או עד שיוצא 6. לכן אלו משתנים מקריים גיאומטריים על פי הגדרתם, $X \sim G\left(\frac{1}{6}\right),\ Y \sim G\left(\frac{2}{6}\right)$.

3. ישנן 2 דרכים לגשת לסעיף השלישי ־ בחישוב ישיר או על ידי הבנת מהו המשתנה המקרי W. החישוב הישיר הוא ארוך ומסובך ולכן עדיף באופן כללי להימנע ממנו. נתחיל בהבנת המשמעות של פונקצית המינימום הנ"ל. W הוא הערך המתנימאלי המתקבל בצמד הפונקציות בכל מאורע, ובפרט אם בהטלות הקוביה יצא 2 או 4 לפני שיצא 6 אזי W סופר את מספר ההטלות עד שיצא 2 או 4, ואם יצא 6 לפני 2 או 4 אז W סופר את מספר ההטלות עד שיצא 6. זו המשמעות של פונקצית המינימום. באופן כללי, W סופר את מספר ההטלות עד שיוצא מספר זוגי כלשהו ולכן גם הוא משתנה מקרי גיאומטרי על פי הגדרתו, $W \sim G\left(\frac{3}{6}\right)$.

4. נמצא את הערכים שמקבל המשתנה המקרי $Y|X=2$. קבוצת הערכים האפשריים היא $\mathbb{N} \setminus \{0,2\}$. נשתמש בהגדרת להסתברות מותנה בכדי לחשב את התפלגות המשתנה המקרי החדש.

$$\Pr\left(Y=l|X=2\right) = \frac{\Pr\left(Y=l, X=2\right)}{\Pr\left(X=2\right)} = $$
$$= \begin{cases} \frac{\frac{2}{6}\cdot\frac{1}{6}}{\frac{5}{6}\cdot\frac{1}{6}} = \frac{2}{5} & l=1 \\ \frac{\frac{3}{6}\cdot\frac{1}{6}\left(\frac{4}{6}\right)^{l-3}\cdot\frac{2}{6}}{\frac{5}{6}\cdot\frac{1}{6}} = \frac{1}{5}\left(\frac{2}{3}\right)^{l-3} & l\in\mathbb{N}, l\geq 3 \end{cases}.$$

3.8 אי תלות של משתנים מקריים

באופן דומה לאי־תלות של צמד מאורעות, ניתן להגדיר אי־תלות של משתנים מקריים.

הגדרה 3.23 (אי־תלות של צמד משתנים מקריים) מ"מ X ו־ Y הם ב"ת אם ורק אם לכל $k,l\in\mathbb{R}$ מתקיים

$$\Pr\left(X=k, Y=l\right) = \Pr\left(X=k\right)\cdot\Pr\left(Y=l\right).$$

למעשה, משתנים מקריים הם בלתי־תלויים אם ההסתברות שאחד יקבל ערך מסויים לא מושפעת כלל על ידי הערכים שהאחר מקבל. דרך אלטרנטיבית להגדיר אי־תלות בין משתנים מקריים היא X ו־ Y הם ב"ת אם לכל $k,l\in\mathbb{R}$ מתקיים

$$\Pr\left(X=k|Y=l\right) = \Pr\left(X=k\right).$$

ההגדרות הנ"ל תקפות עבור צמד משתנים מקריים. ההכללה עבור מספר כלשהו של משתנים מקריים היא ישירה.

הגדרה 3.24 (אי תלות של משתנים מקריים) משתנים מקריים $X_1,\ldots,X_n$ הם ב"ת אם לכל $k_1,\ldots,k_n\in\mathbb{R}$ מתקיים

$$\Pr\left(X_1=k_1,\ldots,X_n=k_n\right)=\prod_{k=1}^{n}\Pr P\left(X_i=k_i\right).$$

ישנם משפטים רבים אשר ניתן להוכיח מתוך ההגדרות הנ"ל. אנחנו לא נספק את הדוגמאות כעת, אך בהחלט נציין את המשפטים הללו מאחר והם יהיו שימושיים ביותר בהמשך.

משפט 3.25 כל תת קבוצה של משתנים מקריים ב"ת, מהווה קבוצה של משתנים מקריים ב"ת.

לכן כאשר נתונה לנו קבוצה של מ"מ ב"ת, נוכל תמיד להסתכל על תת קבוצה ולהשתמש באי־תלות של תת הקבוצה שאנו מעוניינים בה.

משפט 3.26 נתונות $\mu_1,\ldots,\mu_n$ התפלגויות כלשהן. אזי קיים מרחב הסתברות $(\Omega,\Pr)$ ומשתנים מקריים $X_1,\ldots,X_n$ כך שלכל $i=1,\ldots,n$ ההתפלגות של X_i היא μ_i והמשתנים הם ב"ת.

המשפט הבא מתאר אי־תלות של קבוצות של משתנים מקריים ב"ת.

משפט 3.27 יהיו $X_1,\ldots,X_n$ מ"מ ב"ת ותהי פונקציה $f:\mathbb{R}^k\to\mathbb{R}$. אזי המשתנים המקריים $Y,X_{k+1},\ldots,X_n$ הם ב"ת כאשר $Y=f\left(X_1,\ldots,X_k\right)$.

שימו לב, ניתן באותה מידה לקחת פונקציות (יותר מאחת) של אוספים זרים של משתנים מקריים ב"ת ועדיין נקבל שהפונקציות הנ"ל הן אוסף של משתנים מקריים ב"ת. מסקנה ישירה ממשפט זה היא -

מסקנה 3.28 אם $X_1,\ldots,X_n$ מ"מ ב"ת ו־$E_1,\ldots,E_n\subseteq\mathbb{R}$ תת־קבוצות ב־$\mathbb{R}$, אז

$$\Pr\left(X_1\in E_1,\ldots,X_n\in E_n\right)=\prod_{k=1}^{n}\Pr P\left(X_i\in E_i\right).$$

תרגיל 3.29 מספר ביצי החופש שמטילה תרנגולת מתפלג אחיד בין 1 ל־4. מכל ביצה עתיד לבקוע אפרוח מאושר בהסתברות $\frac{1}{3}$ באופן ב"ת בביצים האחרות ובמספר הביצים שהוטלו. יהי X - מספר הביצים שהוטלו ו־Y מספר האפרוחים שבקעו.

1. מצאו את ההתפלגות המשותפת של (X,Y).
2. האם X,Y תלויים?
3. מצאו כיצד מתפלג $X|Y=3$.

פתרון: אנו מניחים בבעיה שהביצים שמוטלות הן שונות זו מזו.

1. ישנה דרך קלה לחשב את ההתפלגות המשותפת בעזרת התפלגות מותנה. $X\sim U[1,4]$ ובהינתן $X=n$ נתון ש־Y הוא משתנה מקרי בינומי עם פרמטרים n ו־ $\frac{1}{3}$. לכן

$$\begin{aligned}\Pr(X=n,Y=k) &= \Pr(Y=k|X=n)\Pr(X=n) = \\ &= \binom{n}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{n-k}\cdot\frac{1}{4}, \quad \left\{\begin{matrix},k=0,\ldots,4\\ ,n=1,\ldots,4\\ k\le n\end{matrix}\right\}\end{aligned}$$

2. דרך קלה לראות שהמשתנים המקריים הללו תלויים היא בעזרת בחירת צמד ערכים למ"מ כך שההגדרה לאי־תלות לא מתקיימת. לרוב נבחר ערכים כך שההתפלגות המשותפת הא 0 והשוליות לא. במקרה הנוכחי קל לראות ש־

$$0=\Pr(X=1,Y=4)\neq\Pr(X=1)\Pr(Y=4)$$

מאחר ו־ $\Pr(X=1)=\frac{1}{4}>0$ ו־

$$\Pr(Y=4)=\Pr(X=4,Y=4)=\frac{1}{4}\cdot\left(\frac{1}{3}\right)^4>0\,.$$

על כן, המ"מ הללו תלויים.

3. הערכים שמקבל המשתנה המקרי $X|Y=3$ הם $3,4$ בלבד וזאת בגלל שלא ייתכן שיהיו יותר אפרוחים מאשר ביצים שהוטלו. נשתמש בחוק בייס בכדי למצוא את ההסתברויות הרלוונטיות.

$$\begin{aligned}\Pr(X=3|Y=3) &= \frac{\Pr(Y=3|X=3)\Pr(X=3)}{\Pr(Y=3)} = \\ &= \frac{\binom{3}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{3-3}\cdot\frac{1}{4}}{\sum_{n=0}^{4}\Pr(Y=3|X=n)\Pr(X=n)} = \\ &= \frac{\frac{1}{27}\cdot\frac{1}{4}}{\sum_{n=3}^{4}\binom{n}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{n-3}\cdot\frac{1}{4}} = \\ &= \frac{\frac{1}{27}\cdot\frac{1}{4}}{\binom{3}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{3-3}\cdot\frac{1}{4}+\binom{4}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{4-3}\cdot\frac{1}{4}} = \\ &= \frac{\frac{1}{27}\cdot\frac{1}{4}}{\frac{1}{27}\cdot\frac{1}{4}+\frac{1}{27}\cdot\frac{2}{3}} = \frac{\frac{1}{4}}{\frac{1}{4}+\frac{2}{3}} = \frac{3}{11}\end{aligned}$$

מאחר ו־$X|Y=3$ הוא משתנה מקרי, בשל תנאי נרמול נקבל ש־ $\Pr(X=4|Y=3)=\frac{8}{11}$.

דוגמה 3.30 אלון מטיל שתי קוביות הוגנות, פעם אחת כל אחת. X ־ התוצאה של הקוביה הראשונה, Y ־ התוצאה המקסימאלית. נמצא את ההתפלגויות

1. $Y\mid X=2$.

2. $X \mid Y = 3$.
3. האם X, Y תלויים?

פתרון: נבצע את החישוב לכל סעיף בנפרד.

1. נחשב ישירות, בעזרת התוצאות שכבר מצאנו בתרגיל קודם.

$$\Pr\left(Y=1 \mid X=2\right) = \frac{\Pr\left(X=2, Y=1\right)}{\Pr\left(X=2\right)} = \frac{0}{1/6} = 0$$
$$\Pr\left(Y=2 \mid X=2\right) = \frac{\Pr\left(X=2, Y=2\right)}{\Pr\left(X=2\right)} = \frac{2/36}{1/6} = \frac{1}{3}$$
$$.\Pr\left(Y=k \mid X=2\right) = \frac{\Pr\left(X=2, Y=1\right)}{\Pr\left(X=2\right)} = \frac{1/36}{1/6} = \frac{1}{6}, \quad k \in \{3,4,5,6\}$$

2. שוב, נעזר בטבלה לחישוב

$$\Pr\left(X=1 \mid Y=3\right) = \frac{\Pr\left(X=1, Y=3\right)}{\Pr\left(Y=3\right)} = \frac{1/36}{5/36} = \frac{1}{5}$$
$$\Pr\left(X=2 \mid Y=3\right) = \frac{\Pr\left(X=2, Y=3\right)}{\Pr\left(Y=3\right)} = \frac{1/36}{5/36} = \frac{1}{5}$$
$$\Pr\left(X=3 \mid Y=3\right) = \frac{\Pr\left(X=3, Y=3\right)}{\Pr\left(Y=3\right)} = \frac{3/36}{5/36} = \frac{3}{5}$$
$$.\Pr\left(X=k \mid Y=3\right) = \frac{\Pr\left(X=k, Y=3\right)}{\Pr\left(Y=3\right)} = \frac{0}{5/36} = 0, \quad k \in \{4,5,6\}$$

3. כמובן שהמ"מ תלויים. נראה שההגדרה לאי־תלות אינה מתקיימת בעזרת שני ערכים ממשיים,

$$.\Pr\left(X=1 | Y=3\right) = \frac{1}{5} \neq \frac{1}{6} = \Pr\left(X=1\right)$$

במקרה זה השתמשנו בדרך השנייה להראות תלות / אי־תלות.

הערה 3.31 ניתן לחשוב על בעיה זו באופן נוסף ־ כנירמול העמודה $X=2$ או השורה $Y=3$, בהתאמה, בטבלה כדי לקבל הסתברות (הרי שאחד התנאים להיסתברות הוא שסכום ההסתברויות שווה 1).

תרגיל 3.32 (סמסטר א', תשע"ב, מועד ב') גודלו של מרחב ההסתברות הקטן ביותר (מבחינת מספר איברים) שבו ניתן להגדיר לושה משתנים מקריים X, Y, Z המקבלים את הערך 0 בהסתברות $\frac{2}{3}$ ואת הערך 1 בהסתברות $\frac{1}{3}$ ומקיימים ש־X, Y ב"ת הוא:

1. 2.

2. 4.

3. 9.

4. אף תשובה אינה נכונה.

פתרון: נרצה לנסות לבנות מרחב הסתברות עליו מוגדרים המשתנים הללו וכן נרצה לגרום ל־X,Y להיות ב"ת. מאחר ואין כל מגבלה על Z, נוכל להגדיר את המרחב כך ש־$Z=Y$ ולהתרכז בהגדרת המרחב עבור X ו־Y. נרצה שהמרחב יכלול איברים עבור המצבים הבאים: $\{X=0,Y=0\}$, $\{X=0,Y=1\}$, $\{X=1,Y=0\}$ ו־ $\{X=1,Y=1\}$. לכן נגדיר את המרחב באופן הבא: $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$ כך ש־

$$X(\omega_1)=X(\omega_2)=0\ ;\ X(\omega_3)=X(\omega_4)=1$$
$$Y(\omega_1)=Y(\omega_3)=0\ ;\ Y(\omega_2)=Y(\omega_4)=1.$$

כעת נותר לבדוק האם ניתן לקבוע את פונקצית ההסתברות על המרחב ולקבל אי־תלות בין המשתנים. נגדיר

$Y\setminus X$	0	1	$\Pr(Y=k)$
0	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{2}{3}$
1	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{1}{3}$
$\Pr(X=k)$	$\frac{2}{3}$	$\frac{1}{3}$	

Table 2: טבלת התפלגות משותפת של משתנים מקריים X ו־ Y.

אפשר כעת לבדוק שאכן יש אי־תלות בין המשתנים. למעשה הדרך הקלה ביותר היא פשוט להגדיר מרחב הסתברות עבור כל משתנה בפני עצמו ואז להגדיר מרחב מכפלה של המרחבים הנתונים. מאחר ואין לנו כל מגבלה על Z, ניתן להגדירו ע"ב אחד המרחבים האחרים ובגלל שכל משתנה מקבל רק צמד ערכים, אין ביה להגדיר כל מרחב רק עם 2 איברים.

תרגיל 3.33 (סמסטר ב', תשע"ב, מועד א' ־ בערך...) תהיי פונקציה $B:\mathbb{N}\cup\{0\}\to(0,1)$ כך ש־

$$B(k)=ce^{-k},\quad \forall k\in\mathbb{N}\cup\{0\}$$

ו־c הוא קבוע נרמול כלשהו. נתון כי הפונקציה מגדירה התפלגות בשם "בן־בן". יהיו שני משתנים מקריים ב"ת X,Y בעלי התפלגות בן־בן, זאת אומרת, לכל $k\in\mathbb{N}\cup\{0\}$ מתקיים

$$\Pr(X=k)=\Pr(Y=k)=ce^{-k}.$$

אילו מהטענות הבאות נכונות:

1. המשתנה המקרי $2X$ הוא בעל התפלגות בן־בן.

2. $X+Y$ הוא בעל התפלגות בן־בן.

3. $X+Y$ ו־ $X-Y$ הם ב"ת.

4. אף הטענה אינה נכונה.

פתרון: נבחן כל טענה בנפרד. הטענה הראשונה איננה נכונה מאחר והמשתנה המקרי $2X$ מקבל רק ערכים זוגיים ולכן לא יכול להתפלג ע"פ ההתפלגות הנתונה. נעבור לטענה השנייה. נחשב את ההסתברות ש־$X+Y=0$,

$$\begin{aligned}\Pr(X+Y=0) &= \Pr(X=0,Y=0) = \\ &= \Pr(X=0)\Pr(Y=0) = \\ &= c^2 \neq c\end{aligned}$$

מאחר ו־$c\neq 0,1$ ולכן גם הטענה השנייה איננה נכונה. הטענה השלישית מעט יותר מסובכת. נשים לב שהמשתנים המקריים הללו תלויים, לדוגמא:

$$\Pr(X+Y=0,X-Y=1)=0\neq\Pr(X+Y=0)\Pr(X-Y=1).$$

הסיבה שההסתברות שרשומה מימין שווה ל־0 היא שהמצב בו $X+Y=0$ מחייב ש־ $X=Y=0$ ולכן לא ניתן לקבל ש־ $X-Y=1$.

3.9 צימוד

תכונה חשובה הנוגעת לצמדים של משתנים מקריים היא צימוד. נניח שנתונים צמד משתנים מקריים $X_1:\Omega_1\to\mathbb{R}$ ו־ $X_2:\Omega_2\to\mathbb{R}$ ואנחנו מעוניינים להציג אותם על מרחב מדגם יחיד. מדוע שנרצה לעשות זאת? ישנם מקרים שבהם ישנם צמד משתנים מקריים כנ"ל עם קשר כלשהו בינהם, אבל הם מוגדרים על מרחבי הסתברות שונים. בכדי להשוות בין המשתנים המקריים, אנחנו צריכים מרחב חדש ויחיד ששניהם מוגדרים עליו ומשמרים את תכונותיהן המקוריות. במידה וישנו כזה מרחב, יהיה הרבה יותר קל להשוות בין המשתנים הללו, מאחר ומסתכלים על מאורעות במרחב ספציפי ולא על כל מרחב בנפרד.

נתחיל בהגדרת צימוד עבור צמד התפלגויות μ_1,μ_2.

הגדרה 3.34 יהיו צמד התפלגויות μ_1,μ_2. צימוד של ההתפלגויות היא התפלגות דו־מימדית μ כך שההתפלגויות השוליות שלה הן μ_1,μ_2.

אם נרצה לבחון במה מדובר בפועל בהיבט הטכני גרידא. בעצם מדובר בדרך כלשהי למלא את הטבלה של ההתפלגות המשותפת באופן שהסכומים בכל שורה ועמודה ייתנו את ההתפלגויות הנתונות. נרחיב את הסוגייה עבור משתנים מקריים.

הגדרה 3.35 צימוד בין ההתפלגויות של משתנים מקריים X ו־Y משמעו שקיים מרחב הסתברות Ω עליו מוגדרים X ו־Y (או זוג מ"מ בעלי אותה התפלגות כמו המ"מ X ו־Y) כך שההתפלגות המשותפת שלהם היא הצימוד בין ההתפלגויות השוליות שלהם.

הצימודים הנ"ל יהיו מאוד הכרחיים כאשר נדון במשתנים מקריים פואסוניים וכיצד ניתן לקבל אותם בעזרת גבול של סכום של משתנים מקריים בינומיים / ברונלי.

3.10 תכונות של משתנים מקריים ספציפיים

3.10.1 משתנה מקרי פואסוני

משתנה מקרי פואסוני X זה משתנה מקרי X שמקבל ערכים ב־ $\mathbb{N}$, וההתפלגות שלו נתונה על ידי

$$\Pr\left(X=k\right)=e^{-\lambda}\frac{\lambda^{k}}{k!},\quad\forall k\in\mathbb{N}\,.$$

נסמן משתנה מקרי כזה ב־ $X\sim P\left(\lambda\right)$, משתנה מקרי פואסוני עם פרמטר λ.

התפלגות פואסונית היא התפלגות מאוד נפוצה בחיים היומיומיים שלנו. היא מתארת את הסבירות לאירועים שמתרחשים בממוצע בקצב קבוע. לדוגמא:

- מספר רעידות האדמה באזור מסויים מתפלג פואסונית.
- מספר שגיאות הכתיב בעמוד עיתון מתפלג פואסונית.
- מספר השיחות שנכנסות למוקד שירות ביחידת זמן נתונה מתפלג פואסונית.

ועוד. תשימו לב שבכל אחת מהדוגמאות ישנה התייחסות למימד הזמן / השטח (עמוד / אזור תחום / יחידת זמן), והדבר הזה לא קורה במקרה. הפרמטר λ בו אנו דנים הוא בעצם ממוצע האירועים ליחידת זמן / שטח. איך מגיעים למסקנה הזאת? מדוע דווקא ההתפלגות הזאת מתארת את מה שאנו טוענים שהיא מתארת? לשם כך צריכים לבצע מספר חישובי גבול ולהגיע למסקנה שהתפלגות פואסונית מתקבלת מהתפלגות בינומית.

משפט 3.36 (התכנסות סדרת מ"מ בינומיים למ"מ פואסוני) יהיו סדרת מ"מ $X_1,X_2,\ldots$ כך ש־ $X_n\sim Bin\left(n,p_n\right)$. נתון כי $n\cdot P_n\to\lambda$ קבוע כלשהו. אזי לכל קבוע k מתקיים ש־

$$\Pr\left(X_n=k\right)\underset{n\to\infty}{\longrightarrow}e^{-\lambda}\frac{\lambda^{k}}{k!}\,.$$

למשפט זה קיימת גם ההכללה כך שאין צורך להסתכל בכל שלב על מ"מ בינומי X_n שמורכב מ־n ניסויים ב"ת ש"ה. ניתן לחילופין להסתכל על סכום של n משתנים מקריים ברנולי ב"ת כך שכל משתנה מקבל את הערך 1 בהסתברות $p_{n,k},\ k=1,\ldots,n$ ובנוסף מתקיים $\sum_{k=1}^{n}P_{n,k}\to\lambda$.

תרגיל 3.37 אלון עובד במרכז שירות אליו מתקשרים בממוצע $\lambda=10$ אנשים בדקה. מה ההסתברות שבדקה מסויימת יתקשרו 10 אנשים למרכז? ו־30 איש? הניחו כי האוכלוסיה הכללית גדולה מאוד.

פתרון: לשם החישוב נצטרך לבצע מספר הנחות. נניח שיש סה"כ N אנשים ולאחר מכן ניקח את הגבול כאשר $N\to\infty$. לכל אדם ישנו סיכוי של $\frac{\lambda}{N}$ להתקשר למוקד בדקה נתונה (בצורה זו סה"כ יתקשרו למוקד בממוצע λ אנשים). לכן, אם X הוא מ"מ שסופר את מספר האנשים שהתקשרו למוקד בדקה אחת, נקבל ש־

$$\begin{aligned}\Pr\left(X=10\right) &= \binom{N}{10}\left(\frac{\lambda}{N}\right)^{10}\left(1-\frac{\lambda}{N}\right)^{N-10}=\\ &= \frac{\binom{N}{10}}{N^{10}}\lambda^{10}\frac{\left(1-\frac{\lambda}{N}\right)^{N}}{\left(1-\frac{\lambda}{N}\right)^{10}}\end{aligned}$$

כעת נחשב את הגבול של כל שבר בנפרד.

$$\begin{aligned}
\lim_{N\to\infty}\frac{\binom{N}{10}}{N^{10}} &= \lim_{N\to\infty}\frac{N!}{N^{10}\,(10)!\,(N-10)!}\\
&= \lim_{N\to\infty}\frac{N\cdot(N-1)\cdots(N-9)}{N^{10}10!} =\\
&= \frac{1}{10!}\\
\lim_{N\to\infty}\frac{\left(1-\frac{\lambda}{N}\right)^{N}}{\left(1-\frac{\lambda}{N}\right)^{10}} &= \frac{\lim_{N\to\infty}\left(1-\frac{\lambda}{N}\right)^{N}}{\lim_{N\to\infty}\left(1-\frac{\lambda}{N}\right)^{10}} =\\
&= \frac{\lim_{N\to\infty}\left(1-\frac{\lambda}{N}\right)^{\frac{N}{-\lambda}\cdot(-\lambda)}}{1} =\\
&= e^{-\lambda}
\end{aligned}$$

ולכן ההסתברות היא,

$$\lim_{N\to\infty}\Pr\left(X=10\right) = e^{-\lambda}\frac{\lambda^{10}}{10!}.$$

שזאת למעשה התפלגות פואסונית. באופן דומה היינו רואים שההסתברות ל־30 איש היא

$$\lim_{N\to\infty}\Pr\left(X=30\right) = e^{-\lambda}\frac{\lambda^{30}}{30!}.$$

3.10.2 פיצול פואסונים

למשתנים מקריים פואסונים יש תכונה מאוד מעניינת ־ מכל משתנה מקרי פואסוני יתן לייצר צמד משתנים מקריים פואסונים הקשורים למשתנה המקורי, להתליך הזה קוראים: פיצול פואסונים. אנחנו נדגים את התהליך באמצעות תרגיל.

תרגיל 3.38 מספר המכוניות X הנכנסות לצומת ארלוזרוב־נמיר בשעה מסויימת מתפלג פואסונית עם פרמטר λ. כל מכונית הנכנסת לצומת פונה ימינה בהסתברות p ופונה שמאלה בהסתברות $1-p$ באופן ב"ת במכוניות האחרות (אבל תוך זהירות ותשומת לב מלאה לתנאי הדרך והסובבים!). נסמן ב־X_R את מספר המכוניות הפונות ימינה בשעה מסויימת.

1. מצאו את התפלגות X_R.
2. נניח כי X_L הוא מספר המכוניות שנכנסות לצומת ופונות שמאלה. מצאו את התפלגות X_L.
3. האם X_L, X_R תלויים?
4. האם X_L, X_R תלויים בהינתן $X=n$?

פתרון: משיקולי סימטריה - יספיק לנו למצוא את הפתרון לסעיף א' ובאופן מיידי נקבל את הפתרון לסעיף ב':

1. נשים לב תחילה שמספר המכוניות הנכנסות לצומת הוא משתנה מקרי פואסוני המקבל ערכים ב-$\mathbb{N}$. כעת ננתח את ההתפלגות המותנה של $X_R|X=n$ עבור $n \in \mathbb{N}$. אם נתון שנכנסו לצומת n מכוניות, אזי בהסתברות p כל מכונית תפנה ימינה ובהסתברות המשלימה, תפנה שמאלה. נוכל להתייחס לכל מכונית בתור ניסוי ואז נקבל ש-$X_R|X=n \sim Bin\left(n,p\right)$, כי ישנה סדרה של n ניסויים ב"ת ו-"הצלחה בניסוי" זו פנייה ימינה.

בנוסף, נשים לב שהערכים האפשריים ל-X_L הם $\mathbb{N}$ בלבד. נחשב את ההתפלגות של המשתנה המקרי מפורשות בעזרת נוסחת ההסתברות השלמה והתנייה על X באופן הבא:

$$\begin{aligned}
\Pr\left(X_R=k\right) &= \sum_{n=0}^{\infty}\Pr\left(X_R=k|X=n\right)\Pr\left(X=n\right)= \\
&= \sum_{n=k}^{\infty}\Pr\left(X_R=k|X=n\right)\Pr\left(X=n\right)
\end{aligned}$$

שימו לב שבמעבר מהשורה הראשונה לשנייה, הסכום השתנה מאחר ולא ייתכן שהיו יותר מכוניות שפנו ימינה מאשר מספר המכוניות שנכנסו לצומת,

$$\Pr\left(X_R=k|X=n\right)=0\ \forall n<k\,.$$

נמשיך בפיתוח ונשתמש בנוסחאות להתפלגות פואסונית ובינומית.

$$\begin{aligned}
\Pr\left(X_R=k\right) &= \sum_{n=k}^{\infty}\binom{n}{k}\left(1-p\right)^{n-k}p^k e^{-\lambda}\frac{\lambda^n}{n!}= \\
&= e^{-\lambda}\sum_{n=k}^{\infty}\frac{n!}{k!\left(n-k\right)!}\left(1-p\right)^{n-k}p^k\frac{\lambda^n}{n!}= \\
&= e^{-\lambda}\frac{p^k}{k!}\sum_{n=k}^{\infty}\frac{\left(1-p\right)^{n-k}\lambda^n}{\left(n-k\right)!}= \\
&= e^{-\lambda}\frac{p^k}{k!}\sum_{m=0}^{\infty}\frac{\left(1-p\right)^{m}\lambda^{m+k}}{m!}= \\
&= e^{-\lambda}\frac{\left(\lambda p\right)^k}{k!}\sum_{m=0}^{\infty}\frac{\left(\left(1-p\right)\lambda\right)^{m}}{m!}= \\
&= e^{-\lambda}\frac{\left(\lambda p\right)^k}{k!}e^{\left(1-p\right)\lambda}=e^{-\lambda p}\frac{\left(\lambda p\right)^k}{k!}
\end{aligned}$$

וקיבלנו ש-$X_R \sim P\left(\lambda p\right)$, למעשה מספר המכוניות הפונות ימינה בצומת בשעה מסויימת הוא משתנה מקרי פואסני עם פרמטר λp ובאופן סימטרי מספר המכוניות הפונות שמאלה יהיה מ"מ פואסוני עם פרמטר $\lambda\left(1-p\right)$. תהליך זה נקרא "פיצול פואסוני" מסיבה די ברורה - משתנה פואסוני אחד פוצל לשני משתנים מקריים פואסוניים שסכומם נותן את המשתנה המקורי.

2. משיקולי סימטריה, $X_L \sim P(\lambda(1-p))$.

3. באופן מפתיע אנחנו נקבל שהמשתנים המקריים X_L, X_R הם ב"ת. יהיו $l, r \in \mathbb{R}$ כלשהם. אם $l, r \notin \mathbb{N}$ אזי

$$\Pr(X_L = l, X_R = r) = 0 = \Pr(X_L = l)\Pr(X_R = r)$$

והאי־תלות מתקיימת. אחרת,

$$\begin{aligned}
\Pr(X_L = l, X_R = r) &= \Pr(X_L = l, X_R = r | X = l + r)\Pr(X = l + r) = \\
&= \binom{l+r}{r} p^r (1-p)^l e^{-\lambda} \frac{\lambda^{r+l}}{(r+l)!} = \\
&= \lambda^r \frac{p^r}{r!} \left(e^{-\lambda p} e^{-\lambda(1-p)}\right) \lambda^l \frac{(1-p)^l}{l!} = \\
&= e^{-\lambda p} \frac{(\lambda p)^r}{r!} e^{-\lambda(1-p)} \frac{(\lambda(1-p))^l}{l!} = \\
&= \Pr(X_R = r)\Pr(X_L = l)
\end{aligned}$$

ולכן המ"מ ב"ת.

4. במקרה הנוכחי ישנה תלות והיא די פשוטה להצגה מאחר והיא ליניארית - $X_L = n - X_R$. לדוגמא, נשים לב שעבור $n = 1$ נקבל

$$\Pr(X_L = 0, X_R = 0 | X = 1) = 0 \neq \Pr(X_L = 0)\Pr(X_R = 0) > 0\,.$$

3.10.3 חיבוריות של מ"מ ב"ת

ישנה תכונה מאוד נחמדה אחרת של משתנים מקריים פואסוניים ב"ת והיא חיבוריות.

למה 3.39 (חיבור של מ"מ פואסוניים) יהי צמד מ"מ פואסוניים ב"ת $X_1 \sim P(\lambda_1)\,,\ X_2 \sim P(\lambda_2)$, אזי $X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$.

הדרישה של האי־תלות היא קריטית. אם המ"מ יהיו תלויים אזי הנ"ל לא בהכרח יתקיים. לדוגמא: נניח שמדובר באותו משתנה מקרי $X_1 = X_2 \sim P(\lambda)$ אזי הסכום שלהם מקבל רק ערכים זוגיים ולכן הוא לא יכול להיות משתנה מקרי פואסוני רק מהסתכלות על התומך של ההתפלגות של $X_1 + X_2$. תכונה זו נקראת תכונת החבורה למחצה ($semigroup$) של מ"מ פואסוניים.

זו איננה ההתפלגות היחידה שמקיימת תכונת חיבוריות שכזאת. גם עבור מקרה ספציפי של משתנים מקריים בינומיים נקבל סגירות מסויימת לחיבור.

למה 3.40 (חיבור של מ"מ בינומיים) יהי צמד מ"מ בינומיים ב"ת $X_1 \sim Bin(n,p)\,,\ X_2 \sim Bin(m,p)$, אזי $X_1 + X_2 \sim Bin(n+m,p)$.

תרגיל 3.41 הוכיחו את הלמות הנ"ל.

הוכחה: (עבור חיבור של מ"מ פואסוניים) נחשב את ההסתברות $\Pr(X_1+X_2=n)$ עבור $n\geq 0$ שלם ונראה שמתקבלת תוצאה המתאימה להתפלגות של מ"מ פואסוני כפי שנדרש.

$$\begin{aligned}
\Pr(X_1+X_2=n) &= \sum_{k=0}^{n}\Pr(X_1+X_2=n|X_1=k)\Pr(X_1=k)= \\
&= \sum_{k=0}^{n}\Pr(X_2=n-k)\Pr(X_1=k)= \\
&= \sum_{k=0}^{n}e^{-\lambda_2}\frac{\lambda_2^{n-k}}{n-k!}e^{-\lambda_1}\frac{\lambda_1^k}{k!}= \\
&= e^{-(\lambda_1+\lambda_2)}\sum_{k=0}^{n}\binom{n}{k}\frac{\lambda_2^{n-k}\lambda_1^k}{n!}= \\
&= e^{-(\lambda_1+\lambda_2)}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}\lambda_2^{n-k}\lambda_1^k= \\
&= e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_2+\lambda_1)^n}{n!},
\end{aligned}$$

כאשר השיוויון הראשון נובע מנוסחת ההסתברות השלמה, השיוויון השני נובע מאי־תלות של X_1 ו־X_2 והמעבר האחרון נובע מהבינום של ניוטון. ■

האינטואיציה מאחורי למה 3.40 היא די פשוטה. X_1 סופר את מספר ההצלחות מתוך n ניסויים ב"ת עם הסתברות p להצלחה. גם X_2 מבצע את אותו הדבר עבור m ניסויים ב"ת. מתון שהמתשנים הללו ב"ת ולכן נוכל להסיק שישנם $n+m$ ניסויים ב"ת עם הסתברות p להצלחה בכל אחד מהם. על כן X_1+X_2 סופר את מספר ההצלחות בניסויים הללו וזו היא ממש ההגדרה למשתנה מקרי בינומי עם פרמטרים $n+m$ ו־ p.

3.10.4 חוסר זיכרון של מ"מ והתפלגות גיאומטריים

תכונת חוסר זיכרון היא תכונה נפוצה במשתנים מקריים כמו אצל אנשים. נראה תופעה זאת במשתנה מקרי גיאומטרי.

תרגיל 3.42 אלון הוא בעליו של מטבע מיוחד עם הסתברות p ליפול על "עץ" ו־$1-p$ לנחות על "פלי". אלון מחליט להטיל אותו עד לפעם הראשונה שהוא נוחת על "עץ". נסמן ב־X את מספר ההטלות.

1. מצאו את התפלגות X.
2. בהינתן שאלון ביצע כבר n הטלות שיצאו "פלי", מה ההסתברות שהוא יבצע סה"כ $n+1$ הטלות?

פתרון: נוכל לזהות בקלות את ההתפלגות של המשתנה המקרי הנ"ל. נתייחס לכל הטלה כניסוי ולנחיתה של המטבע על "עץ" כהצלחה. אזי, X הוא משתנה מקרי המתפלג לפי התפלגות גיאומטרית עם פרמטר p (מאחר והוא סופר מספר ניסויים עד ההצלחה הראשונה).

נשים לב ששואלים אותנו בסעיף השני לגבי הסתברות מותנה. נגדיר את המאורע הנתון על ידי

$$A_{n+1} = \{X \geq n+1\}.$$

זה המאורע בו אלון ביצע n הטלות שיצאו "פלי" ולכן תוצאות הניסוי הן בקבוצה A_{n+1}. אנו נדרשים לחשב את ההסתברות הבאה:

$$\begin{aligned}
\Pr\left(X = n+1 | X \geq n+1\right) &= \frac{\Pr\left(\{X = n+1\} \cap \{X \geq n+1\}\right)}{\Pr\left(X \geq n+1\right)} = \\
&= \frac{\Pr\left(X = n+1\right)}{\Pr\left(X \geq n+1\right)} = \\
&= \frac{\left(1-p\right)^n p}{\left(1-p\right)^n} = p
\end{aligned}$$

כאשר במעבר האחרון השתמשנו בעובדה שהמשתנה המקרי יקבל ערכים שגדולים או שווים ל־$n+1$ אם ורק אם יתרחשו n כשלונות בכלל הניסויים הראשונים. לחילופין, ניתן פשוט לחשב את הטור ההנדסי

$$\Pr\left(X \geq n+1\right) = \sum_{k=n+1}^{\infty} \Pr\left(X = k\right) = \sum_{k=n+1}^{\infty} \left(1-p\right)^{k-1} p.$$

ההסתברות שקיבלנו היא בדיוק p, הסיכוי להצלחה בניסוי בודד. המשמעות של תוצאה זו היא שבהינתן והמשתנה המקרי הגיאומטרי הוא לפחות ערך כלשהו, אז ההסתברות שהוא יגיע לערך הבא היא בדיוק p ולא צריך להתחשב בהסתברות של כל הערכים המוקדמים.

4 מדדי מרכז ומדדי פיזור של התפלגויות ומשתנים מקריים

4.1 תוחלת

בפרק הקודם עסקנו במשתנים מקריים והתפלגויות. ראינו כי ישנן התפלגויות רבות שניתן לקטלג לפי משפחות וכעת נעבור לדון במאפיינים נוספים (מדדי מרכז ופיזור) שיש למשתנים מקריים והתפלגויות ואחד מהם הוא התוחלת. התוחלת היא למעשה ממוצע משוקלל של ערכי המשתנה המקרי. מה המשמעות של "ממוצע משוקלל"? למעשה, לוקחים את כל הערכים שמקבל המ"מ, מכפילים כל ערך בהסתברות לקבל את אותו הערך וסוכמים. לצערנו, למרות שההגדרה היא לכאורה מאוד פשוטה ואינטואיטיבית ישנם הרבה בעיות שיכולות לעלות כפי שנראה בהמשך.

הגדרה 4.1 (תוחלת) יהי משתנה מקרי X המוגדר על מרחב הסתברות $(\Omega,\Pr)$. התוחלת של X, $E\left[X\right]$ היא

$$E\left[X\right]=\sum_{\omega\in\Omega}X\left(\omega\right)\Pr\left(\omega\right). \tag{14}$$

כאשר מרחב ההסתברות סופי אזי הסכום הוא סופי ומוגדר היטב. הבעיות מתחילות כאשר מרחב ההסתברות הוא בן־מניה. נניח כעת כי מרחב המדגם הוא בן־מניה. התוחלת של X היא סופית אם

$$\sum_{\omega\in\Omega}\left|X\left(\omega\right)\right|\Pr\left(\omega\right)<\infty.$$

זאת אומרת, כאשר הטור מתכנס בהחלט, נגיד שהתוחלת של X היא סופית ואז גם אין בעיה של הגדרת התוחלת לפי משוואה 14.

ישנם שני תתי־מצבים בהם נאמר התוחלת היא אינסופית:

1. אם $\sum_{\{\omega\in\Omega:X(\omega)>0\}}X\left(\omega\right)\Pr\left(\omega\right)=\infty$ ו־ $\sum_{\{\omega\in\Omega:X(\omega)<0\}}X\left(\omega\right)\Pr\left(\omega\right)>-\infty$ אז נגיד ש־ $E\left[X\right]=\infty$.
2. אם $\sum_{\{\omega\in\Omega:X(\omega)>0\}}X\left(\omega\right)\Pr\left(\omega\right)<\infty$ ו־ $\sum_{\{\omega\in\Omega:X(\omega)<0\}}X\left(\omega\right)\Pr\left(\omega\right)=-\infty$ אז נגיד ש־ $E\left[X\right]=-\infty$.

בכל מצב אחר, התוחלת תהיה לא מוגדרת. שימו לב שתוחלת של מ"מ אי־חיובי או אי־שלילי הן תמיד מוגדרות, אבל ייתכן והן אינסופיות.

משפט 4.2 יהי מ"מ X בעל תוחלת. אזי

$$E\left[X\right]=\sum_{k}\Pr\left(X=k\right)k.$$

המסקנה המתבקשת מהמשפט האחרון הוא שהדבר היחיד שמשפיע על התוחלת של מ"מ X זו ההתפלגות שלו.

דוגמה 4.3 תוחלות של מ"מ בעלי התפלגויות מוכרות.

• מ"מ ברנולי ־ $X \sim B(p)$.

$$E[X] = 1 \cdot p + 0 \cdot (1-p) = p.$$

• מ"מ גיאומטרי ־ $X \sim G(p)$, $0 < p < 1$.

$$\begin{aligned} E[X] &= \sum_{k=1}^{\infty} \Pr(X=k) k = \sum_{k=1}^{\infty} (1-p)^{k-1} pk = \\ &= p \sum_{k=1}^{\infty} \frac{d}{dp} \left[(1-p)^k\right](-1) = -p \frac{d}{dp} \left[\sum_{k=1}^{\infty} (1-p)^k\right] = \\ &= -p \frac{d}{dp} \left[\frac{1-p}{p}\right] = -p \frac{d}{dp} \left[\frac{1}{p} - 1\right] = \frac{1}{p} \end{aligned}$$

• מ"מ בינומי ־ $X \sim Bin(n,p)$, $0 < p < 1$.

$$\begin{aligned} E[X] &= \sum_{k=0}^{n} \Pr(X=k) k = \sum_{k=1}^{n} \binom{n}{k} (1-p)^{n-k} p^k k = \\ &= \sum_{k=1}^{n} \frac{n!}{k-1!n-k!} (1-p)^{n-k} p^k = np \sum_{k=1}^{n} \binom{n-1}{k-1} (1-p)^{n-k} p^{k-1} = \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} (1-p)^{n-k-1} p^k = np(1-p+p)^{n-1} = np \end{aligned}$$

4.1.1 ליניאריות התוחלת

לתוחלת ישנן מספר תכונות מאוד חשובות שמסייעות לנו בחישובה במגוון רחב של מקרים. התכונה העיקרית היא ־ ליניאריות.

משפט 4.4 (ליניאריות התוחלת) יהי מ"מ X, Y המוגדרים על אותו מרחב הסתברות ובעלי תוחלת (סופית או אינסופית) וקבוע $c \in \mathbb{R}$. אזי התוחלת של cX מוגדרת ו־$E[cX] = cE[X]$ (במידה ו־$c = 0$, אזי $E[cX] = 0$). בנוסף, התוחלת של $X + Y$ מוגדרת ו־

$$E[X+Y] = E[X] + E[Y]$$

בכל עת בו אגף ימין מוגדר.

4.1.2 אינדיקטורים ותוחלת

אינדיקטור הוא מ"מ המוגדר על מרחב הסתברות כלשהו $(\Omega, \Pr)$ המסומן לרוב על ידי $\mathbf{1_A}$ כאשר A הוא מאורע ב־Ω. $\mathbf{1_A}$ מקבל את הערך "1" כאשר A מתרחש ואחרת הוא מקבל

את הערך "0". לעיתים הוא נקרא גם מ"מ מציין. בפועל מדובר במשתנה מקרי המקבל את הערכים $0,1$ כאשר $\Pr(\mathbf{1_A}=1)=\Pr(A)$.

הסיבה שזה מ"מ כל כך חיוני בהקשר של תוחלת הוא על ידי היכולת לפרק משתנים מקריים לסכום של אנדיקטורים (תלויים או ב"ת) ועל ידי חישוב התוחלת של כל אינדיקטור ושימוש בליניאריות התוחלת, לקבל את התוחלת של המ"מ המקורי. אנו נפתור בהמשך מספר תרגילים בטכניקה זו.

4.2 חציון ושכיח

ישנם שני מדדי מרכז נוספים שנעשה בהם שימוש נרחב והם ־ החציון והשכיח. השכיח, כמו כן הוא, זה כל ערך בעל הסתברות מקסימאלית להתקבל. זאת אומרת שייתכן מצב בו ישנם מספר ערכים שכיחים. לדוגמא: בהתפלגות אחידה ניתן להתייחס לכל ערך שמתקבל בהסתברות חיובית כשכיח.

הגדרה 4.5 (חציון) יהי X מ"מ. החציון מוגדר להיות כל ערך m כך ש־

$$\Pr(X\geq m)\geq\frac{1}{2}\quad,\quad\Pr(X\leq m)\geq\frac{1}{2}\,.$$

נשים לב שאם לדוגמא ישנו מ"מ שמקבל ארבע ערכים בהסתברות שווה אזי ישנם מספר חציונים, אבל ייתכן מצב בו ישנו חציון יחיד.

תרגיל 4.6 יהי $X\sim Bin(n,p)$, נגדיר $Y=2^x$. חשבו את $E[Y]$ ואת $E\left[Y^2\right]$.

פתרון: נשתמש בהגדרת התוחלת עבור החישובים הנ"ל.

$$\begin{aligned}
E[Y] &= \sum_{k=2^0}^{2^n} k\cdot\Pr(Y=k)= \\
&\overset{m=\log_2(k)}{=} \sum_{m=0}^{n} 2^m\cdot\Pr\left(2^X=2^m\right)= \\
&= \sum_{m=0}^{n} 2^m\cdot\Pr(X=m)= \\
&= \sum_{m=0}^{n} 2^m\binom{n}{m}p^m(1-p)^{n-m}= \\
&= \sum_{m=0}^{n}\binom{n}{m}(2p)^m(1-p)^{n-m}= \\
&= (2p+1-p)^n=(1+p)^n
\end{aligned}$$

כאשר השתמשנו בבינום של ניוטון לחישוב הסכום.

$$\begin{aligned}
E\left[Y^2\right] &= \sum_{k=2^{2\cdot 0}}^{2^{2\cdot n}} k\cdot\Pr\left(Y^2=k\right) = \\
&= \sum_{k=4^0}^{4^n} k\cdot\Pr\left(Y^2=k\right) \\
&\overset{m=\log_4(k)}{=} \sum_{m=0}^{n} 4^m\cdot\Pr\left(2^{2X}=4^m\right) = \\
&= \sum_{m=0}^{n} 4^m\cdot\Pr\left(X=m\right) = \\
&= \sum_{m=0}^{n} 4^m\binom{n}{m}p^m\left(1-p\right)^{n-m} = \\
&= \sum_{m=0}^{n} \binom{n}{m}\left(4p\right)^m\left(1-p\right)^{n-m} = \\
&= \left(4p+1-p\right)^n = \left(1+3p\right)^n
\end{aligned}$$

תרגיל 4.7 בוחרים קוד באורך 10 תווים, כאשר כל תו נבחר באקראי מתוך הספרות $0-9$ ללא קשר לתווים האחרים שנבחרו. נסמן ב־X מ"מ הסופרים את מספר הרצפים 11 בקוד. חשבו את $E\left[X\right]$.

פתרון: במידה ואנו מעוניינים למצוא את התוחלת של X על פי ההתפלגות שלו, נצטרך תחילה למצוא את ההתפלגות הזאת. הדבר אומנם אפשרי, אך יחד עם זאת מאוד מסובך. לכן בד"כ ננסה למצוא דרך יותר קלה לחישוב התוחלת כמו, לדוגמא, פירוק לאינדיקטורים. נגדיר 9 מ"מ $X_i,\ i=1,\ldots 9$ כאשר X_i מקבל את הערך "1" אם ורק אם במקום ה־i מתחיל רצף של 11. נשים לב כי השיוויון הפונקציונאלי הבא מתקיים

$$.X=\sum_{i=1}^{9}X_i$$

מדובר בשיוויון בין פונקציות וחשוב להבין זאת. לכן,

$$\begin{aligned}
E\left[X\right] = E\left[\sum_{i=1}^{9}X_i\right] &= \sum_{i=1}^{9}E\left[X_i\right] = \\
&= \sum_{i=1}^{9}\left[1\cdot\Pr\left(X_i=1\right)+0\cdot\Pr\left(X_i=0\right)\right] = \\
&= \sum_{i=1}^{9}\Pr\left(X_i=1\right) = \sum_{i=1}^{9}\frac{1}{100} = \frac{9}{100}
\end{aligned}$$

תרגיל 4.8 יהיו שני מ"מ X,Y כך ש־ $X\sim Bin\left(n,p\right),\ Y\sim HG\left(N,D,n\right)$. מצאו את התוחלת של צמד המ"מ הללו.

פתרון: נשתמש בפירוק של המשתנים המקריים לאינדיקטורים בכדי לקבל את התוחלת בצורה פשוטה. עבור X נגדיר n מ"מ ברנולי X_i שמקבלים את הערך 1 אם הייתה הצלחה בניסוי ה־i ו־0 אחרת. לכן $X=\sum_{i=1}^{n}X_i$. שימו לב שמדובר בשיוויון פונקציונאלי (של צמד פונקציות).

$$\begin{aligned} E\left[X\right] &= E\left[\sum_{i=1}^{n}X_i\right]=\sum_{i=1}^{n}E\left[X_i\right]= \\ &= \sum_{i=1}^{n}\left[1\cdot\Pr\left(X_i=1\right)+0\cdot\Pr\left(X_i=0\right)\right]= \\ &= \sum_{i=1}^{n}\Pr\left(X_i=1\right)=\sum_{i=1}^{n}p=np \end{aligned}$$

באופן דומה נגדיר n מ"מ ברנולי Y_i שמקבלים את הערך 1 אם האובייקט ה־i שנשלף הוא מיוחד ו־0 אחרת. לכן $Y=\sum_{i=1}^{n}Y_i$ ובחישוב התוחלת נקבל ש־

$$\begin{aligned} E\left[Y\right] &= E\left[\sum_{i=1}^{n}Y_i\right]=\sum_{i=1}^{n}E\left[Y_i\right]= \\ &= \sum_{i=1}^{n}\Pr\left(Y_i=1\right)= \\ &= \sum_{i=1}^{n}\frac{D}{N}=n\frac{D}{N} \end{aligned}$$

שימו לב שישנו דימיון מסויים בין התוחלת של צמד המ"מ. בשני המקרים התוחלת היא מספר "הניסויים" כפול ההסתברות "להצלחה" בכל ניסוי. באופן כללי ניתן להראות שישנו גם דימיון רב בין ההתפלגויות כאשר דנים באוכלוסיות גדולות. זה לא אמור להפתיע אותנו מאחר וכאשר לוקחים מדגם קטן מאוכלוסיה גדולה, אז הוצאה של מספר מצומצם של גורמים מהאוכלוסיה לא משפיעה יותר מידי על ההסתברות להוציא אובייקט מיוחד נוסף ולכן בגבול כאשר האוכלוסיה היא אינסופית והמדגם סופי אנחנו אכן מקבלים אי־תלות כמו במקרה של מ"מ בינומי.

תרגיל 4.9 בחנות ירקות קטנה ישנם 7 תפוחים ו־8 אגסים. אלון קונה 5 פירות באקראי. בן המוכר מרוויח על כל תפוח 2 ש"ח ועל כל אגס 3 ש"ח. מצאו את תוחלת הרווח של בן.

פתרון: נגדיר מ"מ Q עבור הרווח של בן ויהי מ"מ X־ מספר התפוחים שאלון קנה. על כן,

$$Q=2X+3\left(5-X\right)=2X+15-3X=15-X.$$

התוחלת של X יחסית קלה לחישוב וזאת מאחר והוא בעל התפלגות היפרגאומטרית, $X\sim HG\left(15,7,5\right)$. על כן, $E\left[X\right]=5\cdot\frac{7}{15}=\frac{7}{3}$. נשתמש בתכונת הליניאריות לחישוב התוחלת של Q.

$$E\left[Q\right]=15-E\left[X\right]=15-\frac{7}{3}=\frac{38}{3},$$

המעבר הראשון נכון משום שניתן להתייחס למספר 15 כמ"מ קבוע שמקבל את הערך 15 ואנו יודעים שלמ"מ X קיימת תוחלת סופית ולכן גם ל־$-X$.

תרגיל 4.10 מטילים n מטבעות הוגנים כאשר על כל מטבע רשומים הספרות $0,1$, אחת מכל צד. נגדיר מ"מ X_i המציינים את תוצאות במקומות $i=1,\ldots,n$ בהתאמה. מערבבים באקראי את המטבעות בלי להפוך אותם ומסדרים מחדש במקומות $1,\ldots,n$. נגדיר מ"מ Y_i המציינים את תוצאות במקומות $i=1,\ldots,n$ לאחר הערבוב. נגדיר $X=\sum_{i=1}^{n}X_i$, $Z_i=\mathbf{1}_{\{X_i=Y_i\}}$ ו־$Z=\sum_{i=1}^{n}Z_i$ (שימו לב כי Z_i מקבל את הערך "1" כאשר $X_i=Y_i$ ו־0 אחרת). ענו על הסעיפים הבאים:

1. $\Pr\left(Y_i=1\right)=?$.
2. $\Pr\left(Y_i=1|X_i=1\right)=?$.
3. $\Pr\left(Z_i=1\right)=?$.
4. $E\left[Z\right]=?$.
5. $\Pr\left(Z=n|X=k\right)=?$.
6. $\Pr\left(Z=n\right)=?$.
7. עבור $n=15$, $\Pr\left(Z=0\right)=?$.
8. כאשר $n\to\infty$, מה הגבול של סדרת ההסתברויות $\Pr\left(Z=0\right)$?

פתרון: נתייחס לכל סעיף בנפרד.

1. $\Pr\left(Y_i=1\right)=?$: לכל אחד מהמטבעות הסתברות שווה להגיע למקום ה־i לאחר הערבוב ($\frac{1}{n}$). בנוסף, לכל מטבע ישנה הסתברות של $\frac{1}{2}$ ליפול על הצד עם הספרה "1" כלפי מעלה, ולכן בעזרת נוסחת ההסתברות השלמה נקבל ש־

$$\Pr\left(Y_i=1\right)=\sum_{i=1}^{n}\frac{1}{2}\cdot\frac{1}{n}=\frac{1}{2}\sum_{i=1}^{n}\frac{1}{n}=\frac{1}{2}\ .$$

2. $\Pr\left(Y_i=1|X_i=1\right)=?$: נוכל לפתור סעיף זה בעזרת התנייה על המטבע שמגיע למקום i לאחר הערבוב. נסמן ב־$A_{j\to i}$ את המאורע בו מטבע j עובר למקום i לאחר

הערבוב.

$$\begin{aligned}
\Pr\left(Y_i=1|X_i=1\right) &= \sum_{j=1}^{n}\Pr\left(Y_i=1|X_i=1,A_{j\to i}\right)\Pr\left(A_{j\to i}\right)= \\
&= \left[\sum_{j\neq i}\Pr\left(Y_i=1|X_i=1,A_{j\to i}\right)\Pr\left(A_{j\to i}\right)\right]+\Pr\left(Y_i=1|X_i=1,A_{i\to i}\right)\cdot\frac{1}{n}= \\
&= \left[\sum_{j\neq i}\Pr\left(Y_i=1|A_{j\to i}\right)\frac{1}{n}\right]+1\cdot\frac{1}{n}= \\
&= \left[\sum_{j\neq i}\frac{1}{2}\cdot\frac{1}{n}\right]+\frac{1}{n}=\frac{n-1}{2n}+\frac{1}{n}=\frac{n+1}{2n}
\end{aligned}$$

3. $\Pr\left(Z_i=1\right)=?$: נשתמש שוב בנוסחת ההסתברות השלמה ונקבל

$$\begin{aligned}
\Pr\left(Z_i=1\right) &= \sum_{j\neq i}\Pr\left(Z_i=1|A_{j\to i}\right)\frac{1}{n}+\Pr\left(Z_i=1|A_{i\to i}\right)\frac{1}{n}= \\
&= \frac{1}{n}\sum_{j\neq i}\Pr\left(Z_i=1|A_{j\to i}\right)+1\cdot\frac{1}{n}= \\
&= \frac{1}{n}\sum_{j\neq i}\left[\Pr\left(X_i=X_j|A_{j\to i}\right)\right]+\frac{1}{n}= \\
&= \frac{1}{n}\sum_{j\neq i}\frac{1}{2}+\frac{1}{n}=\frac{n-1}{2n}+\frac{1}{2}=\frac{n+1}{2n}
\end{aligned}$$

4. $E\left[Z\right]=?$: נשתמש בליניאריות התוחלת ובתוצאות מסעיפים קודמים.

$$\begin{aligned}
E\left[Z\right] &= E\left[\sum_{i=1}^{n}Z_i\right]= \\
&= \sum_{i=1}^{n}E\left[Z_i\right]= \\
&= \sum_{i=1}^{n}\Pr\left(Z_i=1\right)= \\
&= \sum_{i=1}^{n}\frac{n+1}{2n}= \\
&= \frac{n+1}{2}
\end{aligned}$$

5. $\Pr\left(Z=n|X=k\right)=?$: נצטרך תחילה להבין את הנתון. נתון כי ישנם k מטבעות שנחתו על 1 והיתר על 0. כמה דרכים יש לסדר את המטבעות? $\binom{n}{k}$, כמספר הדרכים לבחור k מקומות מתוך n. בשביל שהמאורע $Z=n$ יתרחש, אנו צריכים שבכל מקום הערך לא ישתנה. על כן ישנה רק אפשרות אחת לסידור שכזה ולכן

$$\Pr\left(Z=n|X=k\right)=\frac{1}{\binom{n}{k}}\,.$$

שימו לב שהנחנו שהמטבעות זהים. לו היינו מניחים שהמטבעות שונים זה מזה, אזי ישנן $n!$ דרכים לסדר אותם בשורה ובנוסף מספר האפשרויות ה־"טובות" (שהערכים לא משתנים לאחר הערבוב) הוא $k!\left(n-k\right)!$ וההסתברות הייתה יוצאת באופן זהה

$$\Pr\left(Z=n|X=k\right)=\frac{k!\left(n-k\right)!}{n!}\,.$$

6. $\Pr\left(Z=n\right)=?$: נשתמש בנוסחת ההסתברות השלמה והתוצאות מסעיפים קודמים.

$$\begin{aligned}\Pr\left(Z=n\right)&=\sum_{k=0}^{n}\Pr\left(Z=n|X=k\right)\Pr\left(X=k\right)=\\&=\sum_{k=0}^{n}\frac{k!\left(n-k\right)!}{n!}\cdot\binom{n}{k}\left(\frac{1}{2}\right)^{k}\left(\frac{1}{2}\right)^{n-k}=\\&=\sum_{k=0}^{n}\left(\frac{1}{2}\right)^{n}=\frac{1}{2^{n}}\sum_{k=0}^{n}1=\frac{n+1}{2^{n}}\end{aligned}$$

7. עבור $n=15$, $\Pr\left(Z=0\right)=?$: התשובה היא מיידית 0. הסיבה היא יחסית פשוטה. מאחר ומספר המטבעות אי־זוגי יש יותר מטבעות שנפלו על צד אחד מאשר מטבעות שנפלו על הצד האחר. לכן, אפילו אם נרצה לייצר מקרה בו כל הערכים מתחלפים לאחר הערבוב, לא נוכל לעשות זאת משום שתמיד יישארו מטבעות עודפים (מן הצד הרב יותר) שנאלץ לשים במקום בו הערך שלהם הופיע טרם הערבוב.

8. כאשר $n\to\infty$, מה הגבול של סדרת ההסתברויות $\Pr\left(Z=0\right)$: עבור ערכי n אי־זוגיים נקבל תמיד הסתברות 0. נבדוק מה קורה עבור ערכים זוגיים. יהי n זוגי.

$$\begin{aligned}\Pr\left(Z=0\right)&=\Pr\left(Z=0\cap X=\frac{n}{2}\right)=\\&=\Pr\left(Z=0|X=\frac{n}{2}\right)\Pr\left(X=\frac{n}{2}\right)=\\&=\frac{1}{\binom{n}{\frac{n}{2}}}\binom{n}{\frac{n}{2}}\left(\frac{1}{2}\right)^{n/2}\left(\frac{1}{2}\right)^{n-n/2}=\\&=\frac{1}{2^{n}}\to0\;\;as\;\;n\to\infty\end{aligned}$$

ולכן ההסתברות שואפת ל־0.

תרגיל 4.11 קודקוד v בגרף $G=(V,E)$ נקרא קודקוד מבודד אם הוא אינו נמצא בשום קשת, ז"א הוא אינו מחובר בקשת לאף קודקוד אחר. יהיו $2\leq n\in\mathbb{N}$ ו־$p\in(0,1)$ ונגדיר גרף מקרי עם n קודקודים כאשר כל קשת מתקבלת בהסתברות p באופן ב"ת בקשתות האחרות. יהי X מ"מ הסופר את מספר הקודקודים המבודדים בגרף המקרי. חשבו את $E[X]$.

פתרון: נבצע פירוק של X לאינדיקטורים. לכל קודקוד $v\in V$ נגדיר X_v מ"מ מציין שמקבל את הערך 1 אם ורק אם הקודקוד v מבודד ואת הערך 0 אחרת. לפי הגדרת המ"מ מתקיים השיוויון הפונקציונאלי

$$X = \sum_{v\in V} X_v$$

ולכן

$$\begin{aligned} E[X] &= E\left[\sum_{v\in V} X_v\right] = \sum_{v\in V} E[X_v] = \\ &= \sum_{v\in V}\Pr(X_v=1) = \\ &= \sum_{v\in V}(1-p)^{n-1} = \\ &= n(1-p)^{n-1} \end{aligned}$$

נשים לב שככל שמספר הקודקודים עולה, תוחלת מספר הקודקודים המבודדים שואפת ל־0.

4.3 תוחלת של מכפלה ופונקציה של מ"מ

כאשר נתון מ"מ בעל תוחלת ונתון מ"מ חדש המוגדר כפונקציה של המ"מ הראשון, אזי ישנה דרך יחסית קלה לחישוב התוחלת של המ"מ החדש.

משפט 4.12 יהי X מ"מ בעל תוחלת ותהי פונקציה $g:\mathbb{R}\to\mathbb{R}$. אזי

$$E(g(X)) = \sum_{\omega\in\Omega} g(X(\omega))\Pr(\omega) = \sum_k g(k)\Pr(X=k).$$

כמו שניתן לראות, התוחלת של $g(x)$ ניתנת לחישוב ע"ב ההתפלגות של X בלבד. יחד עם זאת, נשים לב שהשיוויון $E(g(X))=g(E(x))$ לרוב לא יהיה נכון.

לרוב נוטים להתייחס ל־ $E(X^n)$ כאשר $n\in\mathbb{N}$ בתור המומנט ה־n־י של X. המשפט הבא מראה שאם מומנט מסדר מסויים קיים וסופי עבור מ"מ X כלשהו, אזי גם כל מומנט קטן ממנו קיים וסופי.

משפט 4.13 יהי X מ"מ ו־n מספר טבעי כך ש־ $E\left(\left|X^{n+1}\right|\right)<\infty$, אזי $E(|X^n|)<\infty$.

פתרון: לשם ההוכחה נשתמש באי־שוויון פשוט להוכחה והוא שלכל $k \in \mathbb{R}$,

$$|k^n| \leq \left|k^{n+1}\right| + 1 \,.$$

נתון ש־ $E\left(\left|X^{n+1}\right|\right) < \infty$ ונוכל להשתמש במשפט שאומר שאם $X(\omega) \leq Y(\omega)$, אזי $E(X) \leq E(Y)$ (כאשר התוחלות קיימות). על כן

$$E(|X^n|) \quad \leq \quad E\left(\left|X^{n+1}\right| + 1\right) = E\left(\left|X^{n+1}\right|\right) + 1 < \infty \,,$$

כאשר השתמשנו בליניאריות התוחלת בשיוויון האחרון.

4.3.1 משתנים מקריים ב"ת מתואמים

נניח כי ישנם צמד מ"מ בעלי תוחלת X, Y. שאלה אחת שיכולה להישאל היא האם $E(XY) = E(X)E(Y)$ ובאילו תנאים?

הגדרה 4.14 יהיו צמד מ"מ בעלי תוחלת X, Y ונניח שגם התוחלת של המכפלה XY מוגדרת היטב. אם

$$E(XY) = E(X)E(Y)$$

אז המשתנים המקריים בלתי־מתואמים.

משפט 4.15 נתונים X, Y מ"מ בעלי תוחלת סופית ובלתי־תלויים. אזי X, Y בלתי־מתואמים.

כל זוג משתנים מקריים בלתי־תלויים עם תוחלת סופית הם מ"מ בלתי־מתואמים. שימו לב שהכיוון השני אינו נכון. העובדה שצמד מ"מ הם בלתי־מתאמים, לא אומרת שהם בלתי־תלויים.

תרגיל 4.16 יהי משתנה מקרי $X \sim B\left(\frac{1}{2}\right)$ ונגדיר $Z = X^2$. האם X, Z הם בלתי־תלויים? בלתי־מתואמים?

פתרון: המשתנים המקריים הנ"ל אינם בלתי־תלויים. קל לראות ש־

$$\Pr(X = 1, Z = 0) = 0 \neq \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \Pr(X = 1)\Pr(Z = 0) \,.$$

נשים לב ש־ $Z = X$ וזאת משום ש־X הוא בעל התפלגות ברנולי ולכן העלאה שלו בכל חזקה לא משנה את הערכים שהוא מקבל או את ההסתברות לקבל כל ערך. ולכן

$$E(XZ) = E(X \cdot X) = E(X) = \frac{1}{2} \neq \frac{1}{4} = (E(X))^2 = E(X)E(Z) \,.$$

תרגיל 4.17 יהי $(\Omega, \Pr) = \left(\{\omega_1, \dots, \omega_4\}, \Pr(\omega_i) = \frac{1}{4}\ \forall i = 1, \dots, 4\right)$. נגדיר X מ"מ באופן הבא:

$$X(\omega) = \begin{cases} -2 & \omega = \omega_1 \\ -1 & \omega = \omega_2 \\ 1 & \omega = \omega_3 \\ 2 & \omega = \omega_4 \end{cases},$$

ונגדיר $Y = X^2$. האם X, Y הם בלתי־תלויים? בלתי־מתואמים?

פתרון: קל לראות שהמשתנים אינם בלתי־תלויים. ניקח לדוגמא את הערכים $k = 1, l = 4$.

$$\Pr(X = 1, Y = 4) = \Pr(\{\omega_3\} \cap \{\omega_1, \omega_4\}) = \Pr(\phi) = 0.$$

לעומת זאת,

$$\Pr(X = 1)\Pr(Y = 4) = \Pr(\{\omega_3\})\Pr(\{\omega_1, \omega_4\}) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}.$$

נעבור לחישוב התוחלת של כל אחד מהמשתנים המקריים.

$$E(X) = \frac{1}{4}(-2 - 1 + 1 + 2) = 0;$$
$$E(Y) = \frac{1}{4}\left((-2)^2 + (-1)^2 + 1^2 + 2^2\right) = \frac{1}{4}(4 + 1 + 1 + 4) = \frac{5}{2}$$

ולכן $E(X)E(Y) = 0$. נחשב את $E(XY)$.

$$\begin{aligned} E(XY) &= \sum_{\omega_i \in \Omega} X(\omega_i) Y(\omega_i) \Pr(\omega_i) = \\ &= \sum_{\omega_i \in \Omega} X(\omega_i) X^2(\omega_i) \frac{1}{4} = \\ &= \frac{1}{4} \sum_{\omega_i \in \Omega} X^3(\omega_i) = \\ &= \frac{1}{4}(-8 - 1 + 1 + 8) = 0 \end{aligned}$$

לכן המשתנים המקריים בלתי־מתואמים.

שימו לב שהתרגיל האחרון מדגים למה לא כל זוג מ"מ בלתי־מתואמים הם בלתי־תלויים.

4.4 שונות

תוחלת, חציון ושכיח הם מדדי אמצע. באופן אינטואיטיבי, ניתן לקבוע שהם מנסים להעריך את הערך הסביר ביותר של כל משתנה מקרי. בנוסף, ישנו סוג אחר של מדדים למשתנים מקריים ואילו הם מדדי פיזור. מדדי פיזור באים לבחון כמה כל משתנה מקרי מקבל ערכים שונים זה מזה, או לחילופין, כמה המשתנה המקרי מפוזר או מרוכז סביב ערכים ספציפיים.

הגדרה 4.18 יהי X משתנה מקרי כך ש־ $E\left(X^2\right) < \infty$. השונות של X, המסומנת ב־ $V(x)$ (או לפעמים על ידי $Var(x)$) מוגדרת על ידי

$$V(X) = E\left[(X - E(X))^2\right].$$

התוחלת של X מוגדרת היטב וסופית בגלל הנתון ש־ $E\left(X^2\right) < \infty$ יחד עם משפט 4.13. כאשר המומנט השני של מ"מ X מוגדר היטב וסופי, ניתן לחשב את השונות בדרך יחסית קלה בעזרת הליניאריות של התוחלת.

$$\begin{aligned} V(X) &= E\left[(X - E(X))^2\right] = \\ &= E\left[X^2 - 2XE(X) + (E(X))^2\right] = \\ &= E\left[X^2\right] - 2E\left[XE(X)\right] + E\left[(E(X))^2\right] = \\ &= E\left[X^2\right] - 2E(X)E[X] + (E(X))^2 = \\ &= E\left[X^2\right] - (E[X])^2 \end{aligned}$$

אנו יודעים כי הביטויים באגף ימין מוגדרים היטב וסופיים מאחר והמומנט השני של X הוא סופי. למרות שהתוחלת היא ליניארית, הדבר אינו נכון עבור השונות. יחד עם זאת, ישנן מספר תכונות מאוד חשובות שיש לשונות כפי שהמשפטים הבאים מדגימים.

משפט 4.19 (שונות של קבוע) נניח כי X הוא מ"מ על מרחב הסתברות $(\Omega, \Pr)$ כך ש־ $X(\omega) = c$ לכל $\omega \in \Omega$, אזי $V(X) = 0$.

ניתן להוכיח תכונה זאת ישירות מן ההגדרה.

משפט 4.20 (הומוגוניות־חלקית מדרגה 2) יהי X מ"מ בעל מומנט שני סופי ו־ $a, b \in \mathbb{R}$. אזי

$$V(aX + b) = a^2 V(X).$$

ההוכחה היא יחסית פשוטה ונובעת ישירות מההגדרה,

$$\begin{aligned} V(aX + b) &= E\left[(aX + b - E(aX + b))^2\right] = \\ &= E\left[(aX + b - aE(X) - b)^2\right] = \\ &= E\left[a^2 (X - E(X))^2\right] \\ &= a^2 V(X) \end{aligned}$$

הגדרה 4.21 (סטיית התקן של X) סטיית התקן של משתנה מקרי X מוגדרת להיות שורש השונות ומסומנת ב־

$$.\sigma_X = \sqrt{V(X)}$$

4.4.1 שונות משותפת

ישנו סוג נוסף של שונות בה נדון והיא השונות המשותפת. השונות המשותפת היא מדד עבור צמד משתנים מקריים והיא בוחנת את מידת התיאום בינהם (עד כמה מדובר במ"מ מתואמים).

הגדרה 4.22 יהיו צמד מ"מ X,Y כך שהתוחלת שלהם ושל מכפלתם מוגדרת היטב וסופית. אזי, השונות המשותפת של X,Y היא

$$.Cov(X,Y) = E\left((X-E(X))(Y-E(Y))\right)$$

בעזרת מספר פעולות אלגבריות פשוטות ניתן להראות ש־

$$\begin{aligned} .Cov(X,Y) &= E\left((X-E(X))(Y-E(Y))\right) = \\ &= E\left(XY - XE(Y) - YE(X) + E(X)E(Y)\right) = \\ &= E(XY) - E(XE(Y)) - E(YE(X)) + E(E(X)E(Y)) = \\ &= E(XY) - 2E(X)E(Y) + E(X)E(Y) = \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

גם לשונות המשותפת יש מספר תכונות חשובות וקלות להוכחה.

משפט 4.23 יהיו צמד מ"מ X,Y כך שהתוחלת שלהם ושל מכפלתם מוגדרת היטב וסופית, אזי

- השונות המשותפת שלהם היא סימטרית: $.Cov(X,Y) = Cov(Y,X)$
- אם $Z = aX + b$ עבור קבועים $a,b \in \mathbb{R}$ כלשהם, אזי $.Cov(Z,Y) = aCov(X,Y)$
- אם Z הוא מ"מ בעל תוחלת כך ש־ $E(YZ)$ מוגדרת היטב וסופית, אזי ,

$$.Cov(X+Z,Y) = Cov(X,Y) + Cov(Z,Y)$$

הנקודה האחרונה ניתנת להכללה עבור $n \in \mathbb{N}$ משתנים מקריים כלשהם $(X_i)_{i=1}^n$ ומ"מ Y על ידי

$$Cov\left(\sum_{i=1}^{n} X_i, Y\right) = \sum_{i=1}^{n} Cov\left(X_i, Y\right),$$

תחת ההנחה שצמד האגפים מוגדרים היטב.

הערה 4.24 אם X הוא מ"מ בעל מומנט שני, אזי נשים לב ש־

$$Cov\left(X, X\right) = E\left(\left(X - E\left(X\right)\right)\left(X - E\left(X\right)\right)\right) = V\left(X\right),$$

והשונות המשותפת של X עם עצמו היא השונות של X.

בעזרת השונות המשותפת אנו מגדירים פונקציה חדשה שנקראת מקדם המתאם ותסומן ב־ρ. לכל צמד משתנים מקריים X, Y לא קבועים בעלי שונות סופית נגדיר

$$\rho\left(X, Y\right) = \frac{Cov\left(X, Y\right)}{\sigma_X \sigma_Y}.$$

טענה 4.25 יהיו צמד משתנים מקריים X, Y לא קבועים בעלי שונות סופית ונתון מקדם המתאם בינהם $\rho\left(X, Y\right)$. אזי,

1. $\left|\rho\left(X, Y\right)\right| \leq 1$.
2. $\rho\left(X, Y\right) = 1$ אם ורק אם $\Pr\left(Y = aX + b\right) = 1$ כאשר a, b קבועים ו־$a > 0$.
3. $\rho\left(X, Y\right) = -1$ אם ורק אם $\Pr\left(Y = aX + b\right) = 1$ כאשר a, b קבועים ו־$a < 0$.

הוכחה: מספיק להוכיח את המשפט עבור

$$X' = \frac{X - E\left(X\right)}{\sigma_X} \quad , \quad Y' = \frac{Y - E\left(Y\right)}{\sigma_Y}$$

מאחר ו־

$$\begin{aligned}
\rho\left(X', Y'\right) &= \frac{Cov\left(\frac{X - E(X)}{\sigma_X}, \frac{Y - E(Y)}{\sigma_Y}\right)}{\sqrt{V\left(\frac{Y - E(Y)}{\sigma_Y}\right) V\left(\frac{X - E(X)}{\sigma_X}\right)}} = \\
&= \frac{\frac{1}{\sigma_X \sigma_Y} Cov\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right)}{\sqrt{\left(\frac{1}{\sigma_X \sigma_Y}\right)^2 V\left(Y\right) V\left(X\right)}} = \\
&= \frac{Cov\left(X, Y\right)}{\sigma_X \sigma_Y} = \rho\left(X, Y\right),
\end{aligned}$$

כאשר השתמשנו בליניאריות של השונות המשותפת ובתכונות ידועות של שונות. וכן נשים לב ש־

$$\Pr\left(Y'=aX'+b\right)=1\Leftrightarrow\Pr\left(\frac{Y-E\left(Y\right)}{\sigma_Y}=a\frac{X-E\left(X\right)}{\sigma_X}+b\right)=1$$
$$\Leftrightarrow\Pr\left(Y=\frac{\sigma_Y a}{\sigma_X}X-\frac{\sigma_Y}{\sigma_X}aE\left(X\right)+\sigma_Y b+E\left(Y\right)\right)=1$$

כאשר $a>0$ אם ורק אם $\frac{\sigma_Y a}{\sigma_X}>0$.

נשים לב ש־ $E\left(X'\right)=E\left(Y'\right)=0$ ובנוסף $V\left(X'\right)=V(Y')=1$. מאחר ולצמד המשתנים המקריים קיימת שונות סופית נובע שהאי־שיוויונים הבאים מוגדרים היטב ומתקיימים

$$E\left[\left(X'\pm Y'\right)^2\right]\geq 0.$$

בעזרת פתיחת סוגריים והעברת אגפים נקבל ש־

$$\begin{aligned}E\left(X'^2\right)+E\left(Y'^2\right)&\geq\pm2E\left(X'Y'\right)\\V\left(X'\right)+V\left(Y'\right)&\geq\pm2\left(E\left(X'Y'\right)-E\left(X'\right)E\left(Y'\right)\right)\\2&\geq\pm2\rho\left(X',Y'\right)\\1\geq\rho\left(X,Y\right)&\geq-1.\end{aligned}$$

נשים לב ששיוויון מתקיים אם ורק אם $E\left[\left(X'\pm Y'\right)^2\right]=0$ וזאת אומרת ש־ $V\left(X'\pm Y'\right)=0$, או במילים אחרות $X'\pm Y'=const$, מה שמוכיח את הטענה. ■

4.4.2 שונות של סכום מ"מ

ראינו שתוחלת היא ליניארית במ"מ בלי קשר לתלות שלהם. הדבר אינו נכון עבור השונות באופן כללי אלא בתנאים מסויימים. במקרה של שונות של סכום מ"מ, השונות תהיה לינארית אם ורק אם המשתנים המקריים בלתי־מתואמים כפי שהמשפט הבא מדגים.

משפט 4.26 (שונות של סכום מ"מ) יהיו $n\in\mathbb{N}$ משתנים מקריים $\left(X_i\right)_{i=1}^n$ בעלי שונות ונניח שהשונות המשותפת של כל צמד מוגדרת היטב. אזי

$$\begin{aligned}V\left(\sum_{i=1}^n X_i\right)&=\sum_{i=1}^n V\left(X_i\right)+\sum_{i\neq j}Cov\left(X_i,X_j\right)=\\&=\sum_{i=1}^n V\left(X_i\right)+2\sum_{1\leq i<j\leq n}Cov\left(X_i,X_j\right)\end{aligned}$$

תרגיל 4.27 במבחן בקורס ישנן 20 שאלות אמריקאיות כך שלכל שאלה 4 תשובות אפשריות שרק 1 נכונה. בכל שאלה ניתן לסמן רק תשובה אחת. תשובה נכונה מזכה ב־6 נקודות ותשובה שגויה גורעת 2 נקודות. ישנן 8 שאלות קלות שעליהן יודע אלון לענות נכונה ו־12 שאלות קשות בהן הוא מנחש תשובה באקראי. יהי Z הציון של אלון. מצאו את התוחלת והשונות של Z.

פתרון: נסמן ב־X את מספר התשובות הנכונות של אלון מבין השאלות הקשות. השאלות הן בלתי־תלויות ולכן $X \sim Bin\left(12, \frac{1}{4}\right)$. נוכל לרשום מפורשות את היחס בין Z ו־X על ידי

$$\begin{aligned} Z &= 6X + (12 - X)(-2) + 8 \cdot 6 = \\ &= 8X + 24. \end{aligned}$$

נזכור כי התוחלת והשונות של מ"מ בינומי $X \sim Bin\left(n, p\right)$ הן

$$\begin{aligned} &E\left(X\right) = np \\ &V\left(X\right) = np\left(1 - p\right). \end{aligned}$$

נחשב בעזרת נוסחאות אלו את התוחלת והשונות של Z,

$$\begin{aligned} &E\left(Z\right) = 8E\left(X\right) + 24 = 8 \cdot \left(\frac{12}{4}\right) + 24 = 48 \\ &V\left(Z\right) = V\left(8X + 24\right) = 64V\left(X\right) = 64 \cdot 12 \cdot \frac{1}{4} \cdot \frac{3}{4} = 144. \end{aligned}$$

תרגיל 4.28 לבן יש סכום כסף אותו הוא רוצה להשקיע. הוא צריך לבחור בין אפיק השקעות א' שמפיק רווח הנתון על ידי מ"מ X לבין אפיק ב' שמפיק רווח הנתון על ידי מ"מ Y (בשני המקרים מדובר ברווח עבור השקעת כל הכסף). בנוסף, הוא יכול לפצל את ההשקעה כאשר $\alpha \in [0,1]$ מהכסף מושקע באפיק הראשון ו־$1 - \alpha$ לשני, במקרה כזה הרווח יהיה $Z\left(\alpha\right) = \alpha X + \left(1 - \alpha\right) Y$. הניחו כי

- $0 < E\left(X\right) = E\left(Y\right) < \infty$.
- $0 < V\left(X\right) < V\left(Y\right) < \infty$.

מצאו כיצד בן צריך להשקיע את כספו אם הוא מעוניין למקסם את תוחלת הרווח שלו ובו־זמנית למזער את הסיכון? (רמת הסיכון נמדדת על ידי השונות).

פתרון: ליניאריות התוחלת קובעת שלכל $\alpha \in [0,1]$ נקבל ש־

$$\begin{aligned} E\left(Z\left(\alpha\right)\right) &= E\left(\alpha X + \left(1 - \alpha\right) Y\right) = \\ &= \alpha E\left(X\right) + \left(1 - \alpha\right) E\left(Y\right) = \\ &= \alpha E\left(X\right) + \left(1 - \alpha\right) E\left(X\right) = \\ &= E\left(X\right) = E\left(Y\right) \end{aligned}$$

לכן תוחלת הרווח לא תשתנה בכל אופן בו יחליט להשקיע. נעבור לבחון את השונות של ההשקעה.

$$\begin{aligned} V\left(Z\left(\alpha\right)\right) &= V\left(\alpha X + \left(1 - \alpha\right) Y\right) = \\ &= V\left(\alpha X\right) + V\left(\left(1 - \alpha\right) Y\right) + 2Cov\left(\alpha X, \left(1 - \alpha\right) Y\right) = \\ &= \alpha^2 V\left(X\right) + \left(1 - \alpha\right)^2 V\left(Y\right) + 2\alpha\left(1 - \alpha\right) Cov\left(X, Y\right) = \\ &= \alpha^2\left(V\left(X\right) + V\left(Y\right) - 2Cov\left(X, Y\right)\right) + 2\alpha\left(Cov\left(X, Y\right) - V\left(Y\right)\right) + V\left(Y\right) = \\ &= \alpha^2 V\left(X - Y\right) + 2\alpha\left(Cov\left(X, Y\right) - V\left(Y\right)\right) + V\left(Y\right) \end{aligned}$$

וקיבלנו פולינום ב־α. נשים לב ש־$X \neq Y$ ולכן $V(X-Y) > 0$. זאת אומרת, יש לנו פרבולה עם מקדם חיובי, ולכן יש לה נקודת מינימום. נגזור ונשווה לאפס ונקבל

$$\alpha_{min} = \frac{V(Y) - Cov(X,Y)}{V(X-Y)}. \tag{15}$$

נשים לב ש־$\alpha_{min} \in [0,1]$ אם ורק אם

1. $V(Y) \geq Cov(X,Y)$.
2. $V(Y) - Cov(X,Y) \leq V(X-Y)$. זאת אומרת,

$$V(Y) - Cov(X,Y) \leq V(X) + V(Y) - 2Cov(X,Y)$$
$$Cov(X,Y) \leq V(X).$$

שקלול צמד התנאים נותן לנו את התנאי $Cov(X,Y) \leq V(X)$ (התנאי הראשון מתקיים כתוצאה מהשני כי נתון $V(X) < V(Y)$). לסיכום, ממשוואה (15) נקבל שני תנאים:

- אם $V(X) < Cov(X,Y)$, אז $\alpha_{min} \notin [0,1]$ ובפרט $\alpha_{min} > 1$. מאחר ומדובר בפרבולה, המינימום האפשרי יתקבל כאשר $\alpha = 1$, על כן צריך להשקיע את כל הכסף באפיק א'.
- אם $V(X) \geq Cov(X,Y)$, אז יש לפזר את ההשקעה על פי משוואה (15).

תרגיל 4.29 אלון ובן ארגנו מסיבה עם n זוגות וסידרו את $2n$ האנשים בשורה באקראי. יהי X מספר הזוגות שבהם בני הזוג צמודים. מצאו את התוחלת והשונות של X.

פתרון: נשתמש בפירוק לאינדיקטורים בכדי לפתור את הבעיה. נגדיר עבור כל זוג $1 \leq i \leq n$ אינדיקטור

$$X_i = \begin{cases} 1 & the\ i^{th}\ couple\ is\ adjacent \\ 0 & otherwise \end{cases}.$$

נמצא כיצד מתפלג כל אינדיקטור. עבור $i = 1, \ldots, n$ נקבל

$$\Pr(X_i = 1) = \frac{(2n-1)!2!}{(2n)!} = \frac{1}{n}.$$

על כן נקבל ש־

$$\begin{aligned} E(X) &= E\left(\sum_{i=1}^{n} X_i\right) = \\ &= \sum_{i=1}^{n} E(X_i) = \\ &= \sum_{i=1}^{n} \frac{1}{n} = 1. \end{aligned}$$

נעבור כעת לחישוב השונות גם בעזרת אינדיקטורים.

$$\begin{aligned} V(X) &= V\left(\sum_{i=1}^{n} X_i\right) = \\ &= \sum_{i=1}^{n} V(X_i) + \sum_{i\neq j} Cov(X_i, X_j) = \\ &= \sum_{i=1}^{n} \frac{1}{n}\left(1-\frac{1}{n}\right) + \sum_{i\neq j} Cov(X_i, X_j). \end{aligned}$$

המעבר האחרון בוצע בעזרת הנוסחה לשונות של אינדיקטור - אם $Y \sim B(p)$ אזי

$$\begin{aligned} V(Y) &= E\left(Y^2\right) - (E(Y))^2 = \\ &= E(Y) - (p)^2 = \\ &= p - p^2 = p(1-p). \end{aligned}$$

נקבע זוג אינדקסים $1 \leq i \neq j \leq n$ ונחשב את $Cov(X_i, X_j)$. לשם חישוב השונות המשותפת נצטרך למצוא את ההתפלגות של X_iX_j. מאחר ומדובר בצמד אינדיקטורים, אזי גם המכפלה שלהם היא אינדיקטור ובנוסף

$$X_iX_j = 1 \ \Leftrightarrow\ X_i = 1 \ and\ X_j = 1.$$

לכן

$$\begin{aligned} \Pr(X_iX_j = 1) &= \Pr(X_i = 1, X_j = 1) = \\ &= \frac{(2n-2)!\,(2!)^2}{(2n)!} = \\ &= \frac{2}{(2n-1)\,n}. \end{aligned}$$

זאת אומרת שלכל זוג אינדקסים $1 \leq i \neq j \leq n$

$$\begin{aligned} Cov(X_i, X_j) &= E(X_iX_j) - E(X_i)\,E(X_j) = \\ &= \frac{2}{(2n-1)\,n} - \frac{1}{n}\cdot\frac{1}{n} = \\ &= \frac{1}{n}\left[\frac{1}{(2n-1)\,n}\right] = \\ &= \frac{1}{(2n-1)\,n^2} \end{aligned}$$

נציב את התוצאה שקיבלנו בשונות של X,

$$\begin{aligned} V(X) &= \sum_{i=1}^{n} \frac{n-1}{n^2} + \sum_{i\neq j} \frac{1}{(2n-1)\,n^2} = \\ &= \frac{n-1}{n} + 2\binom{n}{2} \frac{1}{(2n-1)\,n^2} = \\ &= \frac{n-1}{n} + \frac{n(n-1)}{(2n-1)\,n^2} = \\ &= \frac{n-1}{n} + \frac{(n-1)}{(2n-1)\,n} = \\ &= \frac{(2n-1)(n-1)+(n-1)}{n(2n-1)} = \\ &= \frac{2n-2}{2n-1} \end{aligned}$$

תרגיל 4.30 מטילים קובייה 100 פעמים. יהי X מספר ההטלות בהן קיבלנו 2 או 3 ו־Y מספר ההטלות בהן קיבלנו 1 או 2. חשבו את $Cov(X,Y)$.

פתרון: נשים לב ש־ $X,Y \sim Bin\left(100, \frac{1}{3}\right)$ ולכן

$$.E(X) = E(Y) = \frac{100}{3}$$

הבעיה מתחילה כאשר רוצים למצוא את התפלגות XY. זאת התפלגות די מורכבת ולכן נצטרך לחשב את הנדרש בדרך אחרת. נפרק כל מ"מ לאינדיקטורים כאשר כל הטלת מטבע תוגדר כניסוי. נגדיר $X = \sum_{i=1}^{100} X_i,\ Y = \sum_{i=1}^{100} Y_i$ ו־ $X_i = 1$ (לחילופין $Y_i = 1$) אם ורק אם בהטלה ה־i יצא 2 או 3 (1 או 2 בהתאמה). נמצא כעת את השונות המשותפת של כל זוג אינדיקטורים X_i, Y_j.

$$\begin{aligned} Cov(X_i, Y_j) &= E(X_iY_j) - E(X_i)\,E(Y_j) = \\ &= \begin{cases} \frac{1}{6} - \frac{1}{3}\cdot\frac{1}{3} & i=j \\ \frac{1}{3}\cdot\frac{1}{3} - \frac{1}{3}\cdot\frac{1}{3} & i\neq j \end{cases} \\ &= \begin{cases} \frac{1}{18} & i=j \\ 0 & i\neq j \end{cases} \end{aligned}$$

שימו לב שאם $i \neq j$ אזי X_i ו־ Y_j הם בלתי־תלויים ובפרט בלתי־מתואמים, על כן $Cov\left(X_i, Y_j\right) = 0$. לסיכום נקבל ש־

$$\begin{aligned} Cov\left(X, Y\right) = Cov\left(\sum_{i=1}^{100} X_i, \sum_{j=1}^{100} Y_j\right) = \\ = \sum_{i=1}^{100}\sum_{j=1}^{100} Cov\left(X_i, Y_j\right) = \\ = \sum_{i=1}^{100} Cov\left(X_i, Y_i\right) = \\ = 100 \cdot \frac{1}{18} = \frac{50}{9} \end{aligned}$$

4.4.3 שונויות של משפחות התפלגות מוכרות

ראינו כי השונות של מ"מ ברנולי עם פרמטר p היא $p\left(1-p\right)$ והחישוב בוצע ישירות מההגדרה. באופן דומה, אם $X \sim Bin\left(n, p\right)$ אז נוכל לפרק אותו ל־n אינדיקטורים ב"ת $\left(X_i\right)_{i=1}^{n}$ שלכל אחד מהם התפלגות ברנולי $B\left(p\right)$, ובגלל שהם בלתי־תלויים, השונות המשותפת של כל זוג תהיה 0. לכן

$$V\left(X\right) = \sum_{i=1}^{n} V\left(X_i\right) = np\left(1-p\right).$$

בכדי לחשב את השונות של משתנה מקרי היפרגאומטרי, גם נצטרך לבצע פירוק כזה, אלא שהפעם המשתנים המקריים יהיו תלויים ולכן נצטרך גם לחשב את השונות המשותפת של כל צמד אינדיקטורים.

משפט 4.31 יהי $X \sim G\left(p\right)$ כאשר $0 < p < 1$. אזי $V\left(X\right) = \frac{1-p}{p^2}$.

הוכחה: נשתמש בחישוב טור מוכר של התוחלת של X, $E\left(X\right)$.

$$\begin{aligned} \sum_{k=1}^{\infty} k\left(1-p\right)^{k-1} p &= \frac{1}{p} \\ \sum_{k=1}^{\infty} k\left(1-p\right)^{k} &= \frac{1-p}{p^2}. \end{aligned}$$

נגזור את צמד האגפים במקביל לפי p ונקבל

$$\begin{aligned}
\frac{d}{dp}\left[\sum_{k=1}^{\infty} k\left(1-p\right)^{k}\right] &= \frac{d}{dp}\left[\frac{1-p}{p^{2}}\right]\\
-\sum_{k=1}^{\infty} k^{2}\left(1-p\right)^{k-1} &= \frac{-1\cdot p^{2}-\left(1-p\right)2p}{p^{4}}\\
\sum_{k=1}^{\infty} k^{2}\left(1-p\right)^{k-1}p &= \frac{p^{2}+\left(1-p\right)2p}{p^{3}}\\
.\sum_{k=1}^{\infty} k^{2}\left(1-p\right)^{k-1}p &= \frac{2-p}{p^{2}}
\end{aligned}$$

בכך חישבנו את המומנט השני של X. נשתמש בתוצאה זאת עבור חישוב השונות של X.

$$\begin{aligned}
V\left(X\right) &= E\left(X^{2}\right)-\left(E\left(X\right)\right)^{2}=\\
&= \left[\sum_{k=1}^{\infty} k^{2}\left(1-p\right)^{k-1}p\right]-\left(\frac{1}{p}\right)^{2}=\\
&= \frac{2-p}{p^{2}}-\frac{1}{p^{2}}=\\
&= \frac{2-p-1}{p^{2}}=\\
&= \frac{1-p}{p^{2}}
\end{aligned}$$

■

שימו לב שנוכל להשתמש בתוצאה האחרונה לחישוב תוחלת של מ"מ בינומי־שלילי. ראינו שניתן לפרק כל מ"מ בינומי־שלילי לסכום של מ"מ גיאומטריים ב"ת. לכן עבור מ"מ בינומי־שלילי כלשהו $X\sim NB\left(n,p\right)$ נקבל

$$.V\left(X\right)=n\left(\frac{1-p}{p^{2}}\right)$$

מושג נוסף הקשור לשונות הוא סטיית התקן. סטיית התקן המסומנת לרוב ב־σ מוגדרת להיות שורש השונות של מ"מ כלשהו,

$$.\sigma_{X}=\sqrt{V\left(X\right)}$$

4.4.4 מקדם מתאם

לאחר שהגדרנו שונות משותפת, נוכל להגדיר מקדם מתאם של צמד מ"מ X,Y.

הגדרה 4.32 (מקדם מתאם) יהיו צמד משתנים מקריים X, Y בעלי מומנט שני סופי. אזי מקדם המתאם של X, Y הוא

$$.\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}$$

מקדם המתאם הוא פונקציה מנורמלת שמקבלת ערכים בין -1 ל־1. ניתן להוכיח כי:

- $\rho(X,Y)=1$ אם ורק אם $Y=aX+b$ כאשר $a>0$. זאת אומרת, כאשר קיים קשר ליניארי חיובי בין צמד המשתנים המקריים.
- $\rho(X,Y)=-1$ אם ורק אם $Y=aX+b$ כאשר $a<0$. זאת אומרת, כאשר קיים קשר לינארי שלילי בין צמד המשתנים המקריים.

תרגיל 4.33 בקערה יש 10 אגסים, 10 תפוחים ו־10 בננות. מוצאים באקראי וללא החזרה 10 פירות. יהיו Y, X מספר התפוחים והבננות שהוצאו מהכד בהתאמה. חשבו את $\rho(X,Y)$.

פתרון: נגדיר את Z להיות מספר האגסים שהוצאו ונשים לב ש־

$$,X+Y=10-Z$$

לכן, $V(X+Y)=V(10-Z)=V(Z)$. בנוסף, משיקולי סימטריה נובע ש־

$$.V(X)=V(Y)=V(Z)$$

נשתמש בנוסחה לשונות של סכום מ"מ,

$$\begin{aligned}
V(Z) &= V(X+Y)= \\
&= V(X)+V(Y)+2Cov(X,Y)= \\
&= 2V(X)+2Cov(X,Y) \\
Cov(X,Y) &= \frac{V(Z)-2V(X)}{2}= \\
&= \frac{V(X)-2V(X)}{2}=-\frac{V(X)}{2}
\end{aligned}$$

נחשב את מקדם המתאם לפי הגדרתו.

$$\begin{aligned}
\rho(X,Y) &= \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}= \\
&= \frac{-\frac{V(X)}{2}}{\sqrt{V(X)^2}}=-\frac{1}{2}
\end{aligned}$$

תרגיל 4.34 ישנם 100 כדורים בכד ועליהם המספרים 1 עד 100 כאשר על כל כדור מספר שונה. מוציאים בזה אחר זה ללא החזרה 10 כדורים. יהי X_i הערך המופיע על הכדור ה־i שהוצא. חשבו את $Cov(X_i, X_j)$.

פתרון: נניח שממשיכים להוציא את כלל הכדורים ונגדיר משתנים מקריים $X_i,\ i=1,\dots,100$ כאשר X_i הוא הערך המופיע על הכדור ה־i שהוצא. אנחנו יודעים כי הסכום של כלל המ"מ הללו הוא קבוע ולכן השונות של הסכום שווה ל־0,

$$V\left(\sum_{i=1}^{100}X_i\right)=V\left(Const\right)=0.$$

בנוסף, נזכור שניתן לבצע הקבלה בין הוצאות מכד ללא החזרה עם סידור עצמים בשורה, כפי שראינו בעבר בבעיות קומבינטוריות ולכן למעשה ישנה סימטריה מלאה בין המ"מ הללו. זאת אומרת, כלל המשתנים בעלי אותה ההתפלגות ובנוסף

$$Cov\left(X_i,X_j\right)=Cov\left(X_k,X_l\right)$$

לכל $1\le i,j,k,l\le 100$ השונים זה מזה. אם $i=j$, אזי $Cov\left(X_i,X_j\right)=V\left(X_i\right)$. מאחר ו־ $X_1\sim X_i\sim U\left[1,100\right]$ נוכל להיעזר בנוסחה לשונות של מ"מ בעל התפלגות אחידה (אם $X\sim U\left[a,b\right]$ אז $V\left(X\right)=\frac{(b-a+1)^2-1}{12}$) ו־

$$\begin{aligned}V\left(X_i\right)&=\frac{\left(b-a+1\right)^2-1}{12}=\\&=\frac{100^2-1}{12}=\frac{9999}{12}\end{aligned}$$

בנוסף,

$$\begin{aligned}0&=V\left(\sum_{i=1}^{100}X_i\right)=\\&=\sum_{i=1}^{100}V\left(X_i\right)+2\sum_{i=1}^{100}\sum_{1\le j\ne i}^{100}Cov\left(X_i,X_j\right)=\\&=100V\left(X_1\right)+100\cdot 99Cov\left(X_1,X_2\right)\\Cov\left(X_i,X_j\right)&=-\frac{100V\left(X_1\right)}{100\cdot 99}=\\&=-\frac{\frac{9999}{12}}{99}=-\frac{101}{12}\end{aligned}$$

לכל $1\le i\ne j\le 100$ שלמים.

4.5 תוחלת מותנה ושונות מותנה

עד כה עסקנו בתוחלת ושונות של מ"מ מעל מרחב הסתברות נתון $(\Omega,\Pr)$. כעת נרחיב את הדיון גם למרחבי הסתברות מותנה $(\Omega,\Pr(\cdot|A))$ כאשר $A\subseteq\Omega$ ו־ $\Pr(A)>0$.

יהי מרחב הסתברות מותנה $(\Omega,\Pr(\cdot|A))$ כאמור לעיל ונניח כי $X:\Omega\to\mathbb{R}$ הוא מ"מ. אזי, X הוא משתנה מקרי המוגדר, הן על מרחב הסתברות $(\Omega,\Pr)$ והן על מרחב ההסתברות

המותנה $(\Omega, \Pr(\cdot|A))$. לכן נוכל להגדיר את התוחלת המותנה של X באופן הבא:

$$\begin{aligned} E[X|A] &= \sum_{\omega\in\Omega} X(\omega)\Pr(\omega|A) = \\ &= \sum_k k\Pr(X=k|A) \end{aligned}$$

וכן את השונות המותנה על ידי

$$\begin{aligned} V[X|A] &= E\left[(X-E[X|A])^2|A\right] = \\ &= E\left[X^2|A\right] - (E[X|A])^2 \end{aligned}$$

טענה 4.35 אם X הוא מ"מ בעל תוחלת (שונות) סופית במרחב $(\Omega, \Pr)$, אזי הוא בעל תוחלת (שונות) מותנה סופית במרחב $(\Omega, \Pr(\cdot|A))$ כאשר $A \subseteq \Omega$ ו־ $\Pr(A) > 0$.

4.5.1 התנייה במשתנה מקרי

יהיו צמד משתנים מקריים X, Y המוגדרים על מרחב הסתברות $(\Omega, \Pr)$. נגדיר את $E[X|Y]$ להיות משתנה מקרי חדש על המרחב הנתון כך ש־

$$E[X|Y](\omega) = E[X|Y=Y(\omega)].$$

כאשר אם $\Pr(Y=Y(\omega))=0$ נקבע שרירותית כי $E[X|Y](\omega)=0$.
לחילופין, עבור כל k שנמצא בתומך של Y ($\Pr(Y=k)>0$) נגדיר

$$f(k) = E[X|Y=k]$$

(שימו לב ש־ $\{Y=y\}$ הוא מאורע בעל הסתברות חיובית ולכן הביטוי האחרון מוגדר היטב) ואז

$$E[X|Y] = f(Y).$$

משפט 4.36 (משפט התוחלת השלמה) יהיו X, Y מ"מ המוגדרים על אותו מרחב הסתברות $(\Omega, \Pr)$. אם $E[X]<\infty$ אזי גם $E[E[X|Y]]<\infty$ ומתקיים

$$E[E[X|Y]] = E[X].$$

הערה 4.37 המשפט הנ"ל תקף גם כאשר X הוא מ"מ אי־שלילי כלשהו. זאת אומרת, במידה ו־X הוא בעל תוחלת סופית, אזי המשפט הנ"ל תקף, אחרת צמד האגפים הם אינסופיים ובפרט שווים זה לזה.

משפט 4.38 (משפט השונות השלמה) יהיו X, Y מ"מ המוגדרים על אותו מרחב הסתברות $(\Omega, \Pr)$. אם $V[X] < \infty$ אזי $V[E[X|Y]] < \infty$, $E[V[X|Y]] < \infty$ ומתקיים

$$.V[X] = E[V[X|Y]] + V[E[X|Y]]$$

הערה 4.39 גם משפט השונות השלמה תקף כאשר X הוא מ"מ כלשהו בעל תוחלת סופית. זאת אומרת, במידה ו־$E[X] < \infty$, $V(X) = \infty$ אז צמד האגפים הם בעלי ערכים אינסופיים ובפרט שווים זה לזה.

טענה 4.40 יהיו X, Y מ"מ המוגדרים על אותו מרחב הסתברות $(\Omega, \Pr)$ ו־$g : \mathbb{R} \to \mathbb{R}$ פונקציה ממשית. אזי

$$.E[Xg(Y)|Y] = g(Y)E[X|Y]$$

תרגיל 4.41 אלון כורה הפחם נמצא במערה עם 3 פתחים כאשר רק פתח אחד מוביל ליציאה מהמערה. אם הוא ייבחר בפתח הנכון הוא ייצא מהמערה תוך 7 שעות, אחרת הוא יילך 3 שעות או 5 שעות (בהתאם לפתח שבחר) ויחזור שוב למערה. בכל פעם, הוא בוחר דרך בהסתברות שווה מבין ה־3 ובאופן ב"ת בבחירות קודמות. יהי X משך הזמן עד שייצא מהמערה. מצאו את $E[X]$.

פתרון: נגדיר מ"מ חדש Y כך ש־

$$Y = \begin{cases} 1 & \text{Alon first chooses path A (7 hours)} \\ 2 & \text{Alon first chooses path B (5 hours)} \\ 3 & \text{Alon first chooses path C (3 hours)} \end{cases}$$

ונשתמש במשפט התוחלת השלמה בכדי לפתור את הבעיה.

$$\begin{aligned} E[X] &= E[E(X|Y)] = \\ &= \sum_{k=1}^{3} E(X|Y=k)P(Y=k) = \\ &= \frac{1}{3}[E(X|Y=1) + E(X|Y=2) + E(X|Y=3)] = \\ &= \frac{1}{3}[7 + E(X) + 5 + E(X) + 3] = \\ E[X] &= \frac{1}{3}[15 + 2E[X]] \\ 3E[X] &= 15 + 2E[X] \\ E[X] &= 15 \end{aligned}$$

טענה 4.42 יהי N מ"מ שמקבל ערכים טבעיים ויהיו $\{X_i\}_{i=1}^{N}$ מ"מ ב"ת (גם ב־N) ושווי התפלגות, אזי

$$.E\left[\sum_{i=1}^{N} X_i\right] = E[N]\,E[X_1]$$

הוכחה: ההוכחה מבוססת על נוסחת התוחלת השלמה.

$$\begin{aligned} E\left[\sum_{i=1}^{N} X_i\right] &= E\left[E\left[\sum_{i=1}^{N} X_i | N\right]\right] = \\ &= \sum_{n=1}^{\infty} E\left[\sum_{i=1}^{N} X_i | N=n\right] \Pr(N=n) = \\ &= \sum_{n=1}^{\infty} E\left[\sum_{i=1}^{n} X_i | N=n\right] \Pr(N=n) = \\ &= \sum_{n=1}^{\infty} E\left[\sum_{i=1}^{n} X_i\right] \Pr(N=n) \end{aligned}$$

המעבר האחרון נכון מאחר ואין תלות בין בין N ל־ $\{X_i\}_{i=1}^{n}$.

$$\begin{aligned} E\left[\sum_{i=1}^{N} X_i\right] &= \sum_{n=1}^{\infty}\left[\sum_{i=1}^{n} E[X_i]\right] \Pr(N=n) = \\ &= \sum_{n=1}^{\infty} [nE[X_1]] \Pr(N=n) = \\ &= E[X_1]\sum_{n=1}^{\infty} n \Pr(N=n) = E[X_1]\,E[N] \end{aligned}$$

כנדרש. ■

תרגיל 4.43 מטילים מטבע הוגן עד קבלת $'H'$ בפעם הראשונה. יהי N מספר ההטלות. מכינים מטבע עם סיכוי $\frac{1}{N}$ ל־$'H'$ ומטילים אותו N פעמים. יהי R מספר הפעמים שיצא $'H'$ במטבע החדש.

1. האם N ו־ R תלויים?
2. חשבו את תוחלת R.

פתרון: קל לראות שהמשתנים הם תלויים, מאחר ו־

$$\begin{aligned} ,R|\{N=1\} &= 1 \\ .R|\{N=2\} &\sim Bin\left(2,\frac{1}{2}\right) \end{aligned}$$

לכן ישנה תלות בין המשתנים. נמצא את התוחלת של R בעזרת נוסחת התוחלת השלמה. תחילה נבחין כי

$$R|N=n\sim Bin\left(n,\frac{1}{n}\right)$$

על כן $E\left[R|N=n\right]=1$. זאת אומרת ש־ $E\left[R|N\right]$ הוא מ"מ קבוע שמקבל תמיד את הערך 1 ולכן

$$E\left[R\right]=E\left[E\left[R|N\right]\right]=E\left[1\right]=1.$$

תרגיל 4.44 נתון מטבע עם סיכוי $0<p<1$ ליפול על H''. מטילים אותו שוב ושוב עד שמתקבלת אותה תוצאה כמו ההטלה הראשונה. יהי X מספר ההטלות, כולל הראשונה. מצאו את התוחלת והשונות של X.

פתרון: נגדיר מ"מ אינדיקטור על ידי $\mathbf{1_A}$ כאשר A הוא מאורע בו ההטלה הראשונה יצאה H''. נשים לב ש־

$$\begin{aligned}X-1|1_A=1 &\sim G\left(p\right),\\ X-1|1_A=0 &\sim G\left(1-p\right).\end{aligned}$$

לכן

$$\begin{aligned}E\left[X-1\right] &= E\left[X-1|1_A=1\right]\Pr\left(1_A=1\right)+\\ &+ E\left[X-1|1_A=0\right]\Pr\left(1_A=0\right)=\\ &= \frac{1}{p}\cdot p+\frac{1}{1-p}\cdot\left(1-p\right)=2\\ E\left[X\right] &= 3\end{aligned}$$

בכדי לחשב את השונות $V\left(X\right)$, נחשב את השונות של $V\left(X-1\right)$.

$$\begin{aligned}V\left(X\right) &= V\left(X-1\right)=\\ &= E\left(\left(X-1\right)^2\right)-\left(E\left(X-1\right)\right)^2=\\ &= E\left(\left(X-1\right)^2\right)-4\end{aligned}$$

נותר לנו למצוא את $E\left(\left(X-1\right)^2\right)$. נשתמש בנוסחה לשונות של מ"מ בעל התפלגות גיאומטרית בכדי למצוא את המומנט השני. יהי $Z\sim G\left(p\right)$ אזי

$$\begin{aligned}V\left(Z\right) &= E\left(Z^2\right)-\left(E\left(Z\right)\right)^2\\ E\left(Z^2\right) &= V\left(Z\right)+\left(E\left(Z\right)\right)^2=\\ &= \frac{1-p}{p^2}+\frac{1}{p^2}=\frac{2-p}{p^2}\end{aligned}$$

לכן

$$\begin{aligned}
E\left((X-1)^2\right) &= E\left[(X-1)^2|1_A=1\right]\Pr(1_A=1)+\\
&+ E\left[(X-1)^2|1_A=0\right]\Pr(1_A=0)=\\
&= \frac{2-p}{p^2}\cdot p+\frac{2-(1-p)}{(1-p)^2}\cdot(1-p)=\\
&= \frac{2-p}{p}+\frac{2-(1-p)}{(1-p)}=\\
&= \frac{(2-p)(1-p)+p(1+p)}{p(1-p)}=\\
&= 2\frac{p^2-p+1}{p(1-p)}
\end{aligned}$$

נשקלל את התוצאות יחדיו ונקבל

$$\begin{aligned}
V(X) &= E\left((X-1)^2\right)-4=\\
&= \frac{2p^2-2p+2-4p+4p^2}{p(1-p)}=\\
&= \frac{6p^2-6p+2}{p(1-p)}
\end{aligned}$$

תרגיל 4.45 יהי $n\geq 1$ טבעי. נגדיר מ"מ T_n להיות מספר ההטלות של מטבע הוגן עד לקבלת רצף של בדיוק n עצים.

1. $E[T_n|T_{n-1}]=?$.
2. $E[T_n]=?$.

פתרון: בכדי לפתור את הבעיה חשוב שנבין את המשמעות של ההתנייה. ישנו מרחב הסתברות הנבנה על בסיס הטלות של מטבע הוגן. מטילים את המטבע עד שמקבלים רצף של בדיוק n עצים. כאשר נתון $T_{n-1}=k$ זה אומר שרצף של $n-1$ עצים רצפוים התקבל לאחר k הטלות. משלב זה והלאה יכולים להתרחש שני דברים:

1. בהסתברות 0.5 התקבל שוב עץ ולכן מספר ההטלות הכולל עד שהתקבלו n עצים הוא $k+1$.
2. בהסתברות 0.5 התקבל פלי בהטלה הבאה ולכן מספר ההטלות הכולל עד שהתקבלו n עצים הוא $k+1+T_n$.

לכן

$$\begin{aligned}
E[T_n|T_{n-1}] &= \frac{1}{2}[T_{n-1}+1]+\frac{1}{2}[T_{n-1}+1+E[T_n]]=\\
&= T_{n-1}+1+\frac{1}{2}E[T_n]
\end{aligned}$$

נשתמש במשפט התוחלת השלמה לחישוב $E\left[T_n\right]$.

$$\begin{aligned} E\left[T_n\right] &= E\left[E\left[T_n|T_{n-1}\right]\right] = \\ &= E\left[T_{n-1}+1+\frac{1}{2}E\left[T_n\right]\right] = \\ &= E\left[T_{n-1}\right]+1+\frac{1}{2}E\left[T_n\right] \end{aligned}$$

נעביר אגפים ונקבל

$$E\left[T_n\right] = 2\left[E\left[T_{n-1}\right]+1\right].$$

זוהי נוסחה רקורסיבית לחישוב התוחלת.

$$\begin{aligned} E\left[T_n\right] &= 2\left[E\left[T_{n-1}\right]+1\right] = \\ &= 2\left[2\left[E\left[T_{n-2}\right]+1\right]+1\right] = \\ &= 2^2E\left[T_{n-2}\right]+2^2+2 = \\ &= 2^3E\left[T_{n-3}\right]+2^3+2^2+2 = \\ &= 2^{n-1}E\left[T_1\right]+\sum_{i=1}^{n-1}2^i \end{aligned}$$

נחשב בנפרד את $E\left[T_1\right]$. מדובר במשתנה מקרי גיאומטרי עם פרמטר $\frac{1}{2}$ לכן התוחלת שלו היא 2. נציב זאת ונסכום את האיברים בטור.

$$\begin{aligned} E\left[T_n\right] &= 2^{n-1}E\left[T_1\right]+\sum_{i=1}^{n-1}2^i = \\ &= 2^n+\sum_{i=1}^{n-1}2^i = \\ &= \sum_{i=1}^{n}2^i = 2\sum_{i=0}^{n-1}2^i = \\ &= 2\frac{2^n-1}{2-1} = 2^{n+1}-2 \end{aligned}$$

5 חסמים על הסתברויות

ישנם מקרים בהם קשה או לא אפשרי לחשב הסתברות בצורה מדוייקת. למשל אם עבור $X \sim \mathrm{Bin}\left(100, \frac{1}{4}\right)$ ננסה למצוא את $P\left(X \geq 49\right)$,

$$\Pr\left(X \geq 49\right) \sum_{k=49}^{100} \Pr\left(X = k\right) = \sum_{k=49}^{100} \binom{100}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{100-k} =?$$

או אם נתונה רק התוחלת ולא ההתפלגות המלאה. במקרים כאלו ואחרים נוכל למצוא חסמים עבור ההסתברויות.

5.1 אי שיוויון מרקוב

משפט 5.1 (א"ש מרקוב) יהי X מ"מ אי שלילי (כלומר $\Pr\left(X \geq 0\right) = 1$), אזי עבור כל $\alpha > 0$ מתקיים

$$\Pr\left(X \geq \alpha\right) \leq \frac{E\left[X\right]}{\alpha}.$$

הערה 5.2 עפ"י משפט זה, נקבל שעבור $X \sim \mathrm{Bin}\left(100, \frac{1}{4}\right)$ מתקיים

$$\Pr\left(X \geq 49\right) \leq \frac{100 \cdot \frac{1}{4}}{49} = \frac{25}{147}.$$

תרגיל 5.3 יהי X מ"מ אי שלילי, $E\left[X\right] = 10$.

1. מצא חסם על $\Pr\left(X \geq 12\right)$.
2. נתון בנוסף כי $P\left(X \geq 2\right) = 1$. מצא חסם טוב יותר ל- $\Pr\left(X \geq 12\right)$.

פתרון:

1. מאי שיוויון מרקוב נקבל

$$\Pr\left(X \geq 12\right) \leq \frac{10}{12} = \frac{5}{6}.$$

2. במקרה זה, נוכל להסתכל על המ"מ $X - 2$ (כיוון שאי שלילי). מלינאריות התוחלת, $E\left[X - 2\right] = 8$, לכן נקבל

$$\Pr\left(X \geq 12\right) = \Pr\left(X - 2 \geq 10\right) \leq \frac{8}{10} = \frac{4}{5},$$

וזהו אכן חסם טוב יותר.

תרגיל 5.4 (ממבחן) משתנה מקרי X מקיים $\Pr\left(0\le X\le 1\right)=1$ ו־ $E\left[X\right]=0.5$. נסמן $p=\Pr\left(X>0.25\right)$. בתנאים אלו, מהו הערך המקסימלי האפשרי עבור p ומהו הערך המינימלי האפשרי?

1. הערך המקסימלי האפשרי הוא 1 והמינימלי האפשרי 0.
2. p בהכרח שווה ל־ 0.75.
3. הערך המקסימלי האפשרי הוא 1 והמינימלי האפשרי $1/3$.
4. אף אחת מהנ"ל.

פתרון: ראשית נראה שהערך המקסימלי האפשרי הוא 1. זאת מאחר ונוכל להגדיר X המקיים את התנאים ע"י

$$\Pr\left(X=0.5\right)=1\,.$$

כדי למצוא את הערך המינימלי נגדיר $Y=1-X$ ונעזר בא"ש מרקוב. מתקיים $E\left[Y\right]=1-E\left[X\right]=0.5$, לכן

$$1-p=\Pr\left(X\le 0.25\right)=\Pr\left(Y\ge 0.75\right)\le\frac{E\left(Y\right)}{0.75}=\frac{2}{3}\,.$$

כלומר, $p\ge\frac{1}{3}$. אם נמצא דוגמא שבה $p=\frac{1}{3}$ סיימנו. האפשרות המקסימלי היא

$$\Pr\left(X=1\right)=\frac{1}{3},\quad \Pr\left(X=0.25\right)=\frac{2}{3}$$

ואכן נקבל $E\left[X\right]=0.5$, לכן התשובה הנכונה היא 3.

5.2 אי שיוויון צ'בישב

א"ש מרקוב מצא לנו חסם כאשר הנתון העיקרי הוא התוחלת. כאשר נתונה לנו גם השונות נוכל לשפר חסם זה.

משפט 5.5 (א"ש צ'בישב) יהי X מ"מ ו־ $\alpha>0$ אזי

$$\Pr\left(\left|X-E\left[X\right]\right|\ge\alpha\right)\le\frac{V\left[X\right]}{\alpha^{2}}\,.$$

הערה 5.6 ניזכר שוב בדוגמא מתחילת השיעור. כאשר $X\sim\mathrm{Bin}\left(100,\frac{1}{4}\right)$ מאחר ו־

$$E\left[X\right]=100\cdot\frac{1}{4}=25,\quad V\left[X\right]=100\cdot\frac{1}{4}\cdot\frac{3}{4}=\frac{75}{4}\,,$$

מא"ש צ'בישב נקבל

$$\begin{aligned}\Pr(X \geq 49) &= \Pr(X-25 =\geq 24) \leq \\ &\leq \Pr(|X-25| \geq 24) \leq \\ &\leq \frac{75}{24^2 \cdot 4} = \\ &= \frac{75}{2304} < \frac{25}{147}\end{aligned} .$$

נשים לב שמאחר שמ"מ בינומי הינו סימטרי סביב התוחלת, נוכל לשפר את החסם

$$\Pr(X \geq 49) = \frac{1}{2}\Pr(|X-25| \geq 24) \leq \frac{75}{4608} .$$

תרגיל 5.7 (שאלה ממבחן) מסדרים 10 כדורים, שלושה לבנים והיתר שחורים בשורה באופן אקראי. חוזרים על הניסוי עד שכל הכדורים הלבנים נמצאים במקומות זוגיים. יהי X מספר החזרות על הניסוי. בעזרת א"ש צ'בישב ניתן לומר כי הסתברות

1. הסיכוי שערכו של X הוא 0 או גדול מ 24, קטן או שווה $^{11}/_{12}$.
2. הסיכוי שערכו של X גדול מ 12, קטן או שווה $^{10}/_{12}$.
3. הסיכוי שערכו של X רחוק מתוחלתו ביותר משתיים, קטן או שווה $^{11}/_{12}$.
4. אף אחד מהנ"ל.

פתרון: ההסתברות לסידור בו כל הכדורים הלבנים נמצאים במקומות זוגיים הוא

$$\frac{\binom{5}{3}}{\binom{10}{3}} = \frac{1}{12}$$

(לבחור 3 מקומות עבור הכדורים הלבנים מתוך 5 המקומות הזוגיים, חלקי האפשרויות לבחור 3 מקומות אקראים ללבנים מתוך 10 האפשריים). לכן $X \sim G\left(\frac{1}{12}\right)$. בפרט,

$$E[X] = \frac{1}{^1/_{12}} = 12, \quad V[X] = \frac{1-\frac{1}{12}}{\left(\frac{1}{12}\right)^2} = 132 .$$

מאי שיוויון צ'בישב נקבל

$$\Pr(|X-12| \geq \alpha) \leq \frac{132}{\alpha^2} .$$

אם נציב $\alpha = 12$, נקבל

$$\Pr(|X-12| \geq 12) = \Pr(\{X=0\} \cup \{X \geq 24\}) \leq \frac{11}{12}$$

(השיוויון נובע מכך ש־ X אי שלילי).

תרגיל 5.8 זמן הנסיעה ממחלף קק"ל עד למחלף השלום הוא בעל התפלגות לא ידועה עם תוחלת μ. ברצוננו לחשב תוחלת זו ולשם כן אנחנו מתכננים לבצע n נסיעות (ב"ת) ולחשב את ממוצע הזמן שלקחו. נסמן ב־ $X_1,\ldots,X_n$ את זמני הנסיעות שביצענו ונסמן את הממוצע של התוצאות ב־ $\bar{X}$. עפ"י א"ש צ'בישב, מה צריך להיות n כדי שהסיכוי שהממוצע שחישבנו $\bar{X}$ רחוק מהתוחלת ביותר מ־ 10 דקות הוא לכל היותר 1% (כלומר, הערך של n עבורו $\Pr\left(\left|\bar{X}-\mu\right|\geq 10\right)\leq 0.01$)? נניח כי שונות זמן הנסיעה ידועה והיא 100.

פתרון: לפי אי שיויון צ'בישב מתקיים

$$\Pr\left(\left|\bar{X}-\mu\right|\geq 10\right)\leq\frac{V\left[\bar{X}\right]}{10^2}=0.01V\left[\bar{X}\right].$$

כלומר, נרצה לבדוק עבור איזה n השונות של $\bar{X}$ קטנה שווה ל־1. מתכונות השונות מתקיים

$$V\left[\bar{X}\right]=V\left[\frac{1}{n}\sum_{k=1}^{n}X_i\right]=\frac{1}{n^2}\left[\sum_{k=1}^{n}V\left(X_i\right)+\sum_{i\neq j}Cov\left(X_i,X_j\right)\right]=$$

מאחר ו־ X_i ב"ת נקבל

$$=\frac{1}{n^2}\sum_{k=1}^{n}V\left(X_i\right)=\frac{1}{n^2}\cdot n\cdot 100=\frac{100}{n}$$

כלומר n צריך להיות 100.

5.3 אי שיוויון צ'בישב החד צדדי

משפט 5.9 (א"ש צ'בישב החד צדדי) יהי X מ"מ ו־ $0\leq\alpha$ אזי

$$\Pr\left(X-E\left[X\right]\geq\alpha\right)\leq\frac{V\left[X\right]}{V\left[X\right]+\alpha^2},$$

ו־

$$\Pr\left(X-E\left[X\right]\leq-\alpha\right)\leq\frac{V\left[X\right]}{V\left[X\right]+\alpha^2}.$$

תרגיל 5.10 (שאלה ממבחן) יהי S_k , $1\leq k\leq n$ הילוך מקרי פשוט. כלומר, $S_0=0$, ולכל $1\leq i\leq n$ מתקיים $S_i-S_{i-1}\in\{-1,1\}$ בהסתברות שווה ובאופן ב"ת.

1. חשבו את $E\left[S_n^2\right]$.
2. הוכיחו כי $E\left[S_n^4\right]=n\left(3n-2\right)$.
3. השתמשו באי שיוויון צ'בישב החד צדדי כדי להוכיח כי $\Pr\left(S_n^2\geq\frac{n}{2}\right)\geq\frac{1}{9}$ לכל $n\geq 1$.

פתרון: ראשית נסמן $X_i = S_i - S_{i-1}$, $i \geq 1$, אזי

$$S_n = \sum_{i=1}^{n} X_i, \quad \Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}.$$

1. נחשב

$$\begin{aligned} E\left[S_n^2\right] &= E\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = \\ &= E\left[\sum_{i=1}^{n} X_i^2 + \sum_{i \neq j} X_i X_j\right] = \\ &= \sum_{i=1}^{n} E\left[X_i^2\right] + \sum_{i \neq j} E\left[X_i X_j\right] \end{aligned}$$

X_i בלתי תלויים ומקיימים

$$E\left[X_i\right] = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0,$$

ו־

$$E\left[X_i^2\right] = 1^2 \cdot \frac{1}{2} + (-1)^2 \cdot \frac{1}{2} = 1,$$

לכן (מאי תלות נקבל $\sum_{i \neq j} E\left[X_i X_j\right] = \sum_{i \neq j} E\left[X_i\right] E\left[X_j\right] = \sum_{i \neq j} E\left[X_i\right]^2$)

$$\begin{aligned} E\left[S_n^2\right] &= \sum_{i=1}^{n} E\left[X_i^2\right] + \sum_{i \neq j} E\left[X_i\right]^2 = \\ &= n \cdot 1 + n(n-1) \cdot 0 = n \end{aligned}$$

2. נראה באינדוקציה. עבור $n = 0$,

$$E\left[S_0\right] = E[0] = 0 = 0 \cdot (3 \cdot 0 - 2).$$

נניח כי מתקיים

$$E\left[S_{n-1}^4\right] = (n-1)\left[3(n-1) - 2\right].$$

מתקיים $S_n = S_{n-1} + X_n$ לכן

$$\begin{aligned} E\left[S_n^4\right] &= E\left[\left(S_{n-1} + X_n\right)^4\right] = E\left[\sum_{k=0}^{4} \binom{4}{k} X_n^k \cdot S_{n-1}^{4-k}\right] = \\ &= \sum_{k=0}^{4} \binom{4}{k} E\left[X_n^k \cdot S_{n-1}^{4-k}\right] = \sum_{k=0}^{4} \binom{4}{k} E\left[X_n^k\right] \cdot E\left[S_{n-1}^{4-k}\right] \end{aligned}$$

כאשר השיוויון השלישי נובע מלינאריות והרביעי מאי תלות. כעת, נחשב את התוחלות של X_n^k

$$.E\left[X_n^k\right]=\begin{cases}1^k\cdot\frac{1}{2}+(-1)^k\cdot\frac{1}{2}=1 & k\text{ even}\\ 1^k\cdot\frac{1}{2}+(-1)^k\cdot\frac{1}{2}=0 & k\text{ odd}\end{cases}$$

לכן נקבל

$$\begin{aligned}E\left[S_n^4\right]&=\sum_{k=0}^{4}\binom{4}{k}E\left[X_n^k\right]\cdot E\left[S_{n-1}^{4-k}\right]=\\&=\binom{4}{0}\cdot E\left[S_{n-1}^4\right]+\binom{4}{2}\cdot E\left[S_{n-1}^2\right]+\binom{4}{4}\cdot E\left[S_{n-1}^0\right]\end{aligned}$$

מהנחת האינדוקציה וסעיף קודם נקבל

$$\begin{aligned}&=1\cdot(n-1)\left[3(n-1)-2\right]+6\cdot(n-1)+1\cdot 1=\\&=n(3n-2)\end{aligned}$$

כנדרש.

3. ראשית, נחשב את התוחלת והשונות של S_n^2. את התוחלת חישבנו בסעיף 1, $E\left[S_n^2\right]=n$, ועפ"י סעיף קודם השונות היא

$$\begin{aligned}V\left[S_n^2\right]&=E\left[S_n^4\right]-E\left[S_n^2\right]^2=\\&=n(3n-2)-n^2=\\&=2n(n-1)\end{aligned}$$

לפי אי שיוויון צבישב החד צדדי מתקיים

$$.\Pr\left(S_n^2-E\left[S_n^2\right]\leq-\frac{n}{2}\right)\leq\frac{V\left[S_n^2\right]}{V\left[S_n^2\right]+\left(\frac{n}{2}\right)^2}$$

נציב,

$$\begin{aligned}.\Pr\left(S_n^2\leq\frac{n}{2}\right)=\Pr\left(S_n^2-n\geq-\frac{n}{2}\right)&\leq\frac{2n(n-1)}{2n(n-1)+\left(-\frac{n}{2}\right)^2}\\&=\frac{2n(n-1)}{n\left(\frac{9}{4}n-2\right)}=\frac{8(n-1)}{9n-8}<\frac{8}{9}\end{aligned}$$

כעת, נוכל לקבל

$$.\Pr\left(S_n^2\geq\frac{n}{2}\right)=1-\Pr\left(S_n^2<\frac{n}{2}\right)\geq 1-\frac{8}{9}=\frac{1}{9}$$

הוכחה: (לאי שיוויון צ'בישב החד צדדי) בה"כ נניח ש־ $E[X]=0$, אזי צריך להראות כי

$$\Pr(X\geq\alpha)\leq\frac{V[X]}{V[X]+\alpha^2}=\frac{E\left[X^2\right]}{E[X^2]+\alpha^2}$$

(נוכל להניח זאת כיוון שאז עבור X כללי נוכל להסתכל על $Y=X-E[X]$ ולקבל את הטענה). לכל $-\alpha<c\in\mathbb{R}$ מתקיים

$$\begin{aligned}
\Pr(X\geq\alpha)&=\Pr(X+c\geq\alpha+c)\leq,\\
&\leq\Pr\left((X+c)^2\geq(\alpha+c)^2\right)\leq\\
&\leq\frac{E\left[(X+c)^2\right]}{(\alpha+c)^2}=
\end{aligned}$$

כאשר אי השיוויון האחרון מתקיים לפי מרקוב. מאחר ו־$E[X]=0$ נקבל

$$\begin{aligned}
&=\frac{E\left[X^2+2cX+c^2\right]}{(\alpha+c)^2}=\\
&=\frac{E\left[X^2\right]+c^2}{(\alpha+c)^2}
\end{aligned}$$

אי שיוויון זה מתקיים לכל $c>-\alpha$, בפרט עבור $c=\frac{E[X^2]}{\alpha}$ (ניתן לגזור ולראות שזה המינימום של הפונקציה). נציב ונקבל

$$\begin{aligned}
\Pr(X\geq\alpha)&\leq\frac{E\left[X^2\right]+\left(\frac{E[X^2]}{\alpha}\right)^2}{\left(\alpha+\frac{E[X^2]}{\alpha}\right)^2}=\\
&=\frac{E\left[X^2\right]\left(1+\frac{E[X^2]}{\alpha^2}\right)}{(\alpha^2+E[X^2])\left(1+\frac{E[X^2]}{\alpha^2}\right)}=\\
&=\frac{E\left[X^2\right]}{E[X^2]+\alpha^2}
\end{aligned}$$

כנדרש. ■

6 גבולות של סכומי מ"מ

6.1 החוק החלש של המספרים הגדולים

משפט 6.1 (החוק החלש של המספרים הגדולים) יהיו $X_1, \ldots, X_n$ מ"מ בלתי מתאמים, ש"ה עם תוחלת μ ושונות σ^2 (בפרט, σ סטיית התקן). נסמן $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, אזי לכל $c > 0$,

$$\Pr\left(\left|\bar{X} - \mu\right| \geq c\right) \leq \frac{\sigma^2}{nc^2}.$$

הערה 6.2 שימו לב שלכל $c > 0$, כאשר n שואף לאינסוף מתקיים

$$\frac{\sigma^2}{nc^2} \underset{n\to\infty}{\longrightarrow} 0.$$

תרגיל 6.3 מכשיר פיננסי מסויים מתנהג באופן הבא: בכל יום המכשיר מכפיל את ערך ההשקעה פי 5 בסיכוי 0.4 או פי $1/5$ בסיכוי 0.6 (הימים ב"ת).

א. מהי תוחלת הרווח של המכשיר לאחר n ימים (כאשר ביום ה־0 יש שקל אחד)?

ב. מה הסיכוי שלאחר n ימים יהיה לי יותר מאגורה?

פתרון:

א. נגדיר X_n ־ תוחלת הרווח של המכשיר לאחר n ימים. נגדיר $\mathbf{1}_i$ ־ בכמה הכפיל המכשיר את ההשקעה ביום ה־ i. אזי

$$X_n = \prod_{i=1}^{n} \mathbf{1}_i, \quad \text{and} \quad \mathbf{1}_i = \begin{cases} 5 & \text{with prob. } 0.4 \\ 1/5 & \text{with prob. } 0.6 \end{cases}.$$

בנוסף, נוכל לחשב

$$E\left[\mathbf{1}_i\right] = 5 \cdot 0.4 + 1/5 \cdot 0.6 = 2.12.$$

מאי התלות בין הימים נקבל

$$\begin{aligned} E\left[X_n\right] = E\left[\prod_{i=1}^{n} \mathbf{1}_i\right] &= \\ = \prod_{i=1}^{n} E\left[\mathbf{1}_i\right] &= \\ = 2.12^n & \end{aligned}$$

נשים לב, שבפרט קיבלנו שתוחלת הרווח שואפת לאינסוף,

$$E\left[X_n\right] = 2.12^n \underset{n\to\infty}{\longrightarrow} \infty.$$

ב. שימו לב, בעוד הסעיף הקודם שאל מהי תוחלת הרווח, כאן השאלה היא מה הסיכוי שבאמת הצלחתי לשמור על הכסף, או לחילופין, כמה גבוהים הסיכונים בשיטה זו. נרצה להשתמש בחוק המספרים הגדולים, אך X_n מתואר כמכפלה ולא כסכום. מתקיים

$$\begin{aligned} & X_n = \mathbf{1}_i \cdot \dots \cdot \mathbf{1}_n = e^{\ln(\mathbf{1}_i \cdot \dots \cdot \mathbf{1}_n)} = e^{\sum_{i=1}^{n} \ln(\mathbf{1}_i)} \\ & \Rightarrow \ln(X_n) = \sum_{i=1}^{n} \ln(\mathbf{1}_i) \end{aligned}$$

בפרט, נפעיל את החוק החלש של המספרים הגדולים על $\frac{1}{n}\ln X_n = \frac{1}{n}\sum_{i=1}^{n}\ln(\mathbf{1}_i)$. לכל $1 \le i \le n$,

$$.\ln(\mathbf{1}_i) = \begin{cases} \ln(5) & \text{with prob. } 0.4 \\ \ln(1/5) & \text{with prob. } 0.6 \end{cases}$$

נוכל לחשב בצורה ישירה ולקבל

$$.E[\ln(\mathbf{1}_i)] \sim -0.32, \quad V[\ln(\mathbf{1}_i)] \sim 2.91$$

עפ"י החוק החלש של המספרים הגדולים נקבל

$$.\Pr\left(\left|\frac{1}{n}\ln(X_n) + 0.32\right| \ge c\right) = \Pr\left(\left|\frac{\sum_{i=1}^{n}\ln(\mathbf{1}_i)}{n} + 0.32\right| \ge c\right) \le \frac{\sigma^2}{nc^2}$$

אנו רוצים למצוא את ההסתברות,

$$\begin{aligned} .\Pr(X_n > 0.01) &= \Pr\left(\frac{1}{n}\ln(X_n) + 0.32 > \frac{1}{n}\ln(0.01) + 0.32\right) \le \\ &\le \Pr\left(\left|\frac{1}{n}\ln(X_n) + 0.32\right| > \left|\frac{1}{n}\ln(0.01) + 0.32\right|\right) \end{aligned}$$

כאשר המעבר האחרון תקף לכל $n \ge 15$, ($\frac{1}{n}\ln(0.01) + 0.32 > 0.0129$) בשל הכלה של מאורעות. לכן יחד נקבל

$$\begin{aligned} .\Pr(X > 0.01) &\le \frac{2.91^2}{n\left(\frac{1}{n}\ln(0.01) + 0.32\right)^2} \le \\ &\le \frac{2.91^2}{n \cdot 0.000169} \le \frac{50107.2}{n} \xrightarrow[n\to\infty]{} 0 \end{aligned}$$

כלומר, הסיכוי שנישאר עם אגורה אחרי "מספיק" זמן שואף לאפס.

6.2 משפט הגבול המרכזי (CLT)

משפט 6.4 (משפט הגבול המרכזי) יהיו $X_1, \dots, X_n$ מ"מ ב"ת ש"ה עם תוחלת μ ושונות σ^2. נסמן $S_n = \sum_{i=1}^{n} X_i$. אזי

$$,\Pr\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \le b\right) \sim \phi(b)$$

בפרט
$$\Pr\left(a\le\frac{S_n-n\mu}{\sqrt{n\sigma^2}}\le b\right)\sim\phi\left(b\right)-\phi\left(a\right)$$
כאשר
$$.\phi\left(z\right)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)dt$$

הערה 6.5 תכונות ϕ:

1. נרמול $1 \underset{z\to\infty}{\longrightarrow} \phi\left(z\right)$.
2. $0 \underset{z\to-\infty}{\longrightarrow} \phi\left(z\right)$.
3. סימטריה סביב 0, $\phi\left(z\right)=1-\phi\left(-z\right)$.

למעשה, ϕ היא פונקצית התפלגות מצטברת של מ"מ בעל התפלגות נורמאלית. זו התפלגות רציפה (מהצורות שלא ראינו עד כה) אבל יחד עם זאת היא מאוד שימושית כאשר מסתכלים על ניסויים שמבוצעים שוב ושוב.

תרגיל 6.6 יוסי ודנה מטילים כל אחד מטבע הוגן n פעמים באופן ב"ת אחד בשני. יהי $S-$ מספר העצים של יוסי, $T-$ מספר העצים של דנה. חשבו בקירוב את
$$.\Pr\left(S-T\ge\sqrt{n/2}\right)$$

פתרון: נגדיר S_i- האינדיקטור שמציין האם בהטלה ה־ i של יוסי יצא עץ, T_i- האינדיקטור שמציין האם בהטלה ה־ i של דנה יצא עץ. על כן נקבל,
$$T=\sum_{i=1}^{n}T_i,\quad S=\sum_{i=1}^{n}S_i$$
$$\Rightarrow S-T=\sum_{i=1}^{n}\left(S_i-T_i\right)$$
כאשר S_i-T_i ב"ת בעלי תוחלת
$$E\left[S_i-T_i\right]=E\left[S_i\right]-E\left[T_i\right]=0$$
(S_i,T_i ש"ה, בפרט בעלי תוחלת זהה), ושונות
$$\begin{aligned}V\left[S_i-T_i\right]&=V\left[S_i\right]+V\left[-T_i\right]+2Cov\left[S_i,-T_i\right]=\\&=V\left[S_i\right]+V\left[T_i\right]=\\&=2\cdot\frac{1}{2}\cdot\left(1-\frac{1}{2}\right)=\\&=\frac{1}{2}\end{aligned}$$

כאשר השיוויון השני נובע מכך ש־ S_i, T_i ב"ת והשלישי מהנוסחא לשונות של מ"מ ברנולי. כעת, נעביר את הביטוי לצורה של משפט הגבול המרכזי

$$\Pr\left(S-T\geq\sqrt{n/2}\right)=\Pr\left(\frac{(S-T)-n\cdot 0}{\sqrt{n\cdot\frac{1}{2}}}\geq 1\right)=1-\Pr\left(\frac{(S-T)-n\cdot 0}{\sqrt{n/2}}<1\right),$$

לפי המשפט, מתקיים

$$\Pr\left(\frac{(S-T)-n\cdot 0}{\sqrt{n/2}}<1\right)\sim\phi(1)=0.8413,$$

לכן סה"כ נקבל

$$\Pr\left(S-T\geq\sqrt{n/2}\right)\sim 0.1587.$$

תרגיל 6.7 בכיתה 100 תלמידים. כל תלמיד מטיל קובייה עד שלראשונה מקבל 6. יהי X_i- מספר ההטלות של התלמיד ה־ i.

1. מה הסיכוי שאלון יטיל בין 5 ל 7? כלומר, מהו $\Pr(5\leq X_1\leq 7)$?
2. מה בקירוב הסיכוי, ע"פ א"ש צ'בישב ומשפט הגבול המרכזי, שממוצע ההטלות של 100 התלמידים יהיה בין 5 ל־ 7?
3. מהו מספר הסטודנטים המינימאלי, ע"פ א"ש צ'בישב ומשפט הגבול המרכזי, שיבטיח שממוצע התוצאות יהיה בין 5 ל־ 7 בסיכוי של לפחות 0.99?

פתרון: ראשית נשים לב שלכל $1\leq i\leq 100$

$$X_i\sim G\left(\frac{1}{6}\right).$$

1. מכך נקבל

$$\begin{aligned}\Pr(5\leq X_1\leq 7)&=\Pr(X_1\geq 5)-\Pr(X_1\geq 8)=\\&=\left(\frac{5}{6}\right)^4-\left(\frac{5}{6}\right)^7\cong\\&\cong 0.203\end{aligned}$$

2. נפתור בשתי דרכים: בדרך הראשונה נמצא חסם על ההסתברות ובדרך השנייה נמצא קירוב של ההסתברות.

פתרון 1־ א"ש צ'בישב. ראשית, נסמן $X = \frac{1}{100}\sum_{i=1}^{100} X_i$. אזי מחפשים מהי $P\left(5 \le X \le 7\right)$. נחשב את התוחלת והשונות של X,

$$E\left[X\right] = E\left[\frac{1}{100}\sum_{i=1}^{100} X_i\right] = E\left[X_1\right] = \frac{1}{1/6} = 6 \ ,$$

$$V\left[X\right] = V\left[\frac{1}{100}\sum_{i=1}^{100} X_i\right] = \frac{1}{100^2}\left(\sum_{n=1}^{100} V\left[X_i\right] + \sum_{i \ne j} Cov\left(X_i, X_j\right)\right) =$$
$$= \frac{1}{100} \cdot \frac{1-\frac{1}{6}}{\left(1/6\right)^2} = \frac{3}{10}$$

כאשר השיוויון השלישי בחישוב השונות נובע מכך שה־ X_i ב"ת ומנוסחת השונות של מ"מ גאומטרי. נעביר את ההסתברות המבוקשת לצורה המתאימה בא"ש צ'בישב

$$\Pr\left(5 \le X \le 7\right) = \Pr\left(\left|X-6\right| \le 1\right) = 1 - \Pr\left(\left|X-6\right| > 1\right) .$$

כעת, עפ"י המשפט נקבל

$$\Pr\left(\left|X-6\right| > 1\right) \le \frac{3/10}{1^2} = 0.3 .$$

לכן

$$\Pr\left(\left|X-6\right| \le 1\right) \ge 0.7 .$$

פתרון 2־ משפט הגבול המרכזי. כאן נסתכל על $S = \sum_{i=1}^{100} X_i$, ואז מחפשים $\Pr\left(500 \le S \le 700\right)$. כפי שחישבנו בחלק קודם,

$$E\left[X_i\right] = 6$$
$$V\left[X_i\right] = 30 .$$

שוב, נעביר את ההסתברות המבוקשת לצורה המתאימה

$$\Pr\left(500 \le S \le 700\right) = \Pr\left(\frac{500 - 100\cdot 6}{\sqrt{100\cdot 30}} \le \frac{S - 100\cdot 6}{\sqrt{100\cdot 30}} \le \frac{700 - 100\cdot 6}{\sqrt{100\cdot 30}}\right) = .$$
$$= \phi\left(\sqrt{\frac{10}{3}}\right) - \phi\left(-\sqrt{\frac{10}{3}}\right) =$$
$$= 2\phi\left(\sqrt{\frac{10}{3}}\right) - 1$$

כאשר המעבר האחרון נובע מכך ש־ $\phi\left(z\right) = 1 - \phi\left(-z\right)$. בעזרת הטבלה נקבל

$$\phi\left(\sqrt{\frac{10}{3}}\right) \sim \phi\left(1.82\right) = 0.9656 \ ,$$
$$\Rightarrow \Pr\left(500 \le S \le 700\right) = 0.9312$$

שימו לב־ הפתרון המדויק (בעזרת התפלגות בינומית־שלילית) 0.933. כלומר, משפט הגבול המרכזי נתן הערכה יותר מדויקת בצורה משמעותית.

3. נפתור בשתי דרכים:

פתרון 1־ א"ש צ'בישב. ראינו בסעיף קודם שמתקיים

$$\Pr\left(5 \le X \le 7\right) = 1 - \Pr\left(|X-6| > 1\right) \ge 1 - \frac{\frac{1}{n}\cdot 30}{1^2} = \frac{n-30}{n}.$$

אנו רוצים שהסתברות זו תהיה לפחות 0.99, כלומר ש־

$$\frac{n-30}{n} \ge 0.99 \quad\iff\quad n \ge 3000,$$

והמספר המינימלי הוא 3000.

פתרון 2־ משפט הגבול המרכזי. כאן, ראינו שמתקיים

$$\Pr\left(5\cdot n \le S \le 7\cdot n\right) = \Pr\left(-\sqrt{\frac{n}{30}} \le \frac{S-n\cdot 6}{\sqrt{n\cdot 30}} \le \sqrt{\frac{n}{30}}\right) - 1 \sim 2\phi\left(\sqrt{\frac{n}{30}}\right) - 1$$

כלומר נרצה ש־

$$2\phi\left(\sqrt{\frac{n}{30}}\right) - 1 \ge 0.99 \quad\iff\quad \phi\left(\sqrt{\frac{n}{30}}\right) \ge 0.995$$
$$\iff\quad \sqrt{\frac{n}{30}} \ge 2.58 \quad\iff\quad n \ge 199.692$$

והמספר המינימלי הוא 200.

גם כאן קיבלנו שהפתרון ההדוק יותר מתקבל בעזרת משפט הגבול המרכזי.

תרגיל 6.8 (ממבחן) מסדרים 10 כדורים, שלושה לבנים והיתר שחורים בשורה באופן אקראי. חוזרים על ניסוי זה 100 פעמים. נגדיר X_i- מספר הכדורים הלבנים במקומות הזוגיים בחזרה ה־ i של הניסוי. נגדיר $\bar{X} = \frac{\sum_{i=1}^{100} X_i}{100}$. העריכו לפי משפט הגבול המרכזי את ההסתברות

$$p = \Pr\left(1.45 \le \bar{X} \le 1.51\right).$$

1. $0 \le p < 0.25$.
2. $0.25 \le p < 0.5$.
3. $0.5 \le p < 0.75$.
4. $0.75 \le p < 1$.

פתרון: מתקיים

$$,X_i \sim HG\left(10,5,3\right)$$

לכן, לכל $1 \le i \le 100$

$$E\left[X_i\right] = \frac{3 \cdot 5}{10} = 1.5$$
$$.V\left[X_i\right] = \frac{3 \cdot \frac{5}{10} \cdot \left(1 - \frac{5}{10}\right) \cdot \left(10-3\right)}{10-1} = \frac{7}{12}$$

כעת, לפי משפט הגבול המרכזי

$$\begin{aligned}
\Pr\left(1.45 \le \bar{X} \le 1.51\right) &= \Pr\left(145 \le \sum_{i=1}^{100} X_i \le 151\right) = \\
&= \Pr\left(\frac{145 - 100 \cdot 1.5}{\sqrt{100 \cdot \frac{7}{12}}} \le \frac{\sum_{i=1}^{100} X_i - 100 \cdot 1.5}{\sqrt{100 \cdot \frac{7}{12}}} \le \frac{151 - 100 \cdot 1.5}{\sqrt{100 \cdot \frac{7}{12}}}\right) \sim \\
&\sim \phi\left(\frac{151-150}{10\sqrt{\frac{7}{12}}}\right) - \phi\left(\frac{145-150}{10\sqrt{\frac{7}{12}}}\right) = \\
&= \phi\left(\frac{1}{5}\sqrt{\frac{3}{7}}\right) - \phi\left(-\sqrt{\frac{3}{7}}\right) = \\
&= \phi\left(0.13\right) - \left(1 - \phi\left(0.65\right)\right) = \\
&= 0.5517 - 1 + 0.7422 = 0.2939
\end{aligned}$$

כלומר, תשובה 2 היא הנכונה.

7 שרשראות מרקוב

7.1 הגדרות

בכדי להגדיר שרשרת מרקוב אנו צריכים להגדיר תחילה שלושה אובייקטים מתמטיים:

1. קבוצת מצבים סופית S־ זו קבוצה כלשהי שמכילה את כל המצבים האפשריים אליהם ניתן להגיע בתהליך.
2. מטריצה P של הסתברויות מעבר בין המצבים ב־S ־ לכל $x,y\in S$ תהי $P\left(x,y\right)=\Pr\left(x\to y\right)$ ההסתברות לעבור ל־y ממצב x בצעד / בשלב 1. התנאי על מטריצת המעבר הוא
$$.\sum_{y\in S}P\left(x,y\right)=1\quad\forall x\in S$$
3. התפלגות התחלתית μ על S ־ על כן $\mu:S\to\left[0,1\right]$ היא פונקציה על S כך ש־
$$.\sum_{x\in S}\mu\left(x\right)=1$$

הגדרה 7.1 שרשרת מרקוב עד זמן n היא אוסף של המשתנים המקריים $X_0,X_1,\ldots,X_n$ המקבלים ערכים ב־S כך ש־

$$,\Pr\left(X_0=x\right)=\mu\left(x\right)\quad\forall x\in S$$

ולכל אוסף מצבים $x_0,x_1,\ldots,x_n\in S$ מתקיים

$$.\Pr\left(X_0=x_0,\ldots,X_n=x_n\right)=\mu\left(x_0\right)P\left(x_0,x_1\right)\cdots P\left(x_{n-1},x_n\right)$$

הערה 7.2 מבחינת סימונים אנו נשתמש בסימון $\Pr_\mu$ בשביל ההסתברות כאשר ההתפלגות ההתחלתית היא μ בכדי להדגיש התפלגות זאת. במידה ו־$\mu\left(x\right)=1$ עבור מצב כלשהו $x\in S$, נסמן את ההסתברות ב־ $\Pr_x$.

7.2 תכונות בסיסיות

הגדרה 7.3 מאורע A תלוי רק ב־ $X_0,\ldots,X_t$ אם קיימת פונקציה $f:S^t\to\{0,1\}$ כך ש־

$$.\mathbf{1_A}=f\left(X_0,\ldots,X_t\right)$$

למה 7.4 (תכונת מרקוב) לכל $t \geq 0$ ומאורעות A, B כך ש־ A נקבע ע"פ $X_0, \ldots, X_{t-1}$ (עבר), $A \subseteq \{X_t = x\}$ (הווה) ו־B נקבע ע"פ $X_{t+1}, \ldots, X_n$ (עתיד) אז $\Pr_\mu(B|A) = \Pr_x(B')$, כאשר אם $\mathbf{1_B} = f(X_{t+1}, \ldots, X_n)$ אז $\mathbf{1_{B'}} = f(X_1, \ldots, X_{n-t})$.

תכונת מרקוב למעשה אומרת שמה שייקרה בעתיד, לאחר זמן t, תלוי אך ורק בהווה ובכך שלאחר ש־A מתרחש, הגענו למצב x בשלב t. מאותו רגע t והלאה, מה שייקבע יהיה תלוי רק ב־x ולא בעבר או בזמן עצמו.

מטריצת המעבר P בין מצבים היא מאוד נוחה מהבחינה הזאת שניתן לתאר את ההתפלגות של מצב t בצורה יחסית פשוטה על ידי מכפלה של וקטור במטריצה. נזכור כי עמודה i במטריצה מייצגת את ההסתברות לעבור מהמצבים השונים למצב i, ולכן כאשר נכפול את ההתפלגות ההתחלתית μ ב־P, $\mu \cdot P$, ככפל של וקטור שורה במטריצה, נקבל וקטור שורה חדש שמתאר את ההתפלגות של X_1.

למה 7.5 (קידום התפלגויות) נסמן ב־ μ_t את ההתפלגות של X_t, קרי $\mu_t(x) = \Pr_\mu(X_t = x)$ לכל $x \in S$. אזי

$$\mu_t = \mu \cdot P^t.$$

הגדרה 7.6 (התפלגות סטציונרית) התפלגות π על S נקראת סטציונרית אם $\pi = \pi \cdot P$.

זאת אומרת, התפלגות היא סטציונרית אם היא וקטור עצמי של מטריצת המעברים. מדוע התפלגויות כאלה כל כך חשובות? ובכן, במקרים רבים התפלגות המצבים מתכנסת בסופו של דבר להתפלגות כלשהי, במובן הבא:

$$\mu_t(x) \underset{t \to \infty}{\to} \pi(x),$$

ומאחר ו־ $\mu_{t+1} = \mu_t \cdot P$ נובע ש־ $\pi = \pi \cdot P$. עבור כל שרשרת מרקוב אנחנו יודעים שקיימת התפלגות סטציונרית כלשהי כפי שהמשפט הבא קובע.

משפט 7.7 (קיום התפלגות סטציונרית) לכל שרשרת מרקוב קיימת התפלגות סטציונרית. כלומר, קיימת π על S כך ש־

$$\pi = \pi \cdot P.$$

תרגיל 7.8 נניח שיש משחק מחשב עם 5 שלבים שונים. בן משחק במשחק ובכל שלב הוא יכול לעלות שלב אחד בהסתברות α או לרדת שלב אחד בהסתברות $1 - \alpha$ באופן ב"ת במיקומו (כאשר השלבים מסודרים על מעגל ולכן משלב 1 ניתן לעלות לשלב 2 או לרדת לשלב 5 ומשלב 5 ניתן לעלות לשלב 1 או לרדת לשלב 4). תארו את הבעיה הנתונה כשרשרת מרקוב ומצאו את ההתפלגות הסטציונרית שלה.

פתרון: נשים לב שקבוצת המצבים היא $S=\{1,2,3,4,5\}$ ואת מטריצת המעבר ניתן להציג באופן הבא:

$$P = \begin{bmatrix} 0 & \alpha & 0 & 0 & 1-\alpha \\ 1-\alpha & 0 & \alpha & 0 & 0 \\ 0 & 1-\alpha & 0 & \alpha & 0 \\ 0 & 0 & 1-\alpha & 0 & \alpha \\ \alpha & 0 & 0 & 1-\alpha & 0 \end{bmatrix}.$$

ההתפלגות הסטציונרית מקיימת את המשוואה הוקטורית $\pi \cdot P=\pi$ ובצורה מפורשת

$$(\pi_1,\pi_2,\pi_3,\pi_4,\pi_5) \begin{bmatrix} 0 & \alpha & 0 & 0 & 1-\alpha \\ 1-\alpha & 0 & \alpha & 0 & 0 \\ 0 & 1-\alpha & 0 & \alpha & 0 \\ 0 & 0 & 1-\alpha & 0 & \alpha \\ \alpha & 0 & 0 & 1-\alpha & 0 \end{bmatrix} = (\pi_1,\pi_2,\pi_3,\pi_4,\pi_5)$$

או לחילופין,

$$\begin{cases} \pi_1 & =(1-\alpha)\,\pi_2+\alpha\pi_5 \\ \pi_2 & =(1-\alpha)\,\pi_3+\alpha\pi_1 \\ \pi_3 & =(1-\alpha)\,\pi_4+\alpha\pi_2 \\ \pi_4 & =(1-\alpha)\,\pi_5+\alpha\pi_3 \\ \pi_5 & =(1-\alpha)\,\pi_1+\alpha\pi_4 \end{cases}$$

אנו אמורים למצוא פתרון למערכת המשוואות הנ"ל. נשים לב שמאחר ומדובר במעגל אז משיקולי סימטריה בין המצבים אפשר לטעון שההתפלגות הסטציונרית צריכה להיות אחידה. כמו כן, אנו יכולים לראות ש־$\pi=\left(\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5}\right)$ מקיים את מערכת המשוואות הנ"ל וזאת מאחר וכל איבר הוא ממוצע משוקלל של צמד איברים אחרים ולכן ברגע שלכולם אותו הערך, נקבל שהתוצאות עקביות בין המשוואות. בנוסף, אפשר להראות שזו התפלגות סטציונרית יחידה. נניח שהיא איננה יחידה ולכן קיים i כך ש־ $\pi_i>\pi_{i+1}$ או $\pi_i<\pi_{i+1}$. נניח בה"כ שהא"ש הראשון מתרחש עבור $i=1$. ולכן $\pi_1>\pi_2$. מהמשוואה השנייה נקבל ש־ $\pi_3<\pi_2$. מהשלישית נקבל ש־ $\pi_4<\pi_3$. מהמשוואה הבאה אחריה נקבל ש־ $\pi_5<\pi_4$. באופן דומה נקבל מהמשוואה האחרונה ש־ $\pi_1<\pi_5$ ולכן המשוואה הראשונה גוררת ש־ $\pi_2<\pi_1$. סתירה.

תרגיל 7.9 (משחק הסטודנטים המעתיקים) אינסוף סטודנטים הנבחנים בכיתה מנסים להעביר בינהם תשובה לשאלת כן/לא. האדם הראשון הוא היחיד שיודע את הפתרון. הוא אומר את תשובה כלשהי לאדם היושב לידו. האדם הזה מעביר אותה לבא אחריו וכן הלאה. אם התשובה היא "כן" אז האנשים מעבירים אותה בהסתברות 1 מאחד לשני. אם התשובה היא "לא" אז בהסתברות 0.5 התשובה שתועבר היא "לא" ובהסתברות 0.5 התשובה שתועבר היא "כן". מצאו כיצד תראה ההתפלגות הגבולית של התשובות. קרי, בהנחה שבכל שלב n לאדם ה־n יש סיכוי מסויים לשמוע פתרון "כן" או "לא", מצאו מה הגבול של התפלגות זאת.

פתרון: תחילה נחליט על המצבים $S=\{Y,N\}$ בתור התשובה המועברת בין האנשים. מטריצת המעברים נתונה על ידי

$$P=\begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

נניח שההתפלגות ההתחלתית היא μ ונמצא את ההתפלגות הגבולית. נחשב את P^3, P^2 וכן הלאה.

$$P^2 = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$$
$$P^3 = \begin{bmatrix} 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{7}{8} & \frac{1}{8} \end{bmatrix}$$
$$P^4 = \begin{bmatrix} 1 & 0 \\ \frac{7}{8} & \frac{1}{8} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{15}{16} & \frac{1}{16} \end{bmatrix}$$

ובאופן כללי, נוכיח באינדוקציה כי $P^n = \begin{bmatrix} 1 & 0 \\ 1-\left(\frac{1}{2}\right)^n & \left(\frac{1}{2}\right)^n \end{bmatrix}$. נניח שהטענה נכונה עבור n ונוכחי עבור $n+1$:

$$P^{n+1} = \begin{bmatrix} 1 & 0 \\ 1-\left(\frac{1}{2}\right)^n & \left(\frac{1}{2}\right)^n \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1-\left(\frac{1}{2}\right)^{n+1} & \left(\frac{1}{2}\right)^{n+1} \end{bmatrix}.$$

לסיכום, נשים לב ש־

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

ולכן בלי קשר לתשובה ההתחלתית של האדם שיודע את התשובה, בגבול נקבל שהתשובה היחידה שתשרוד היא "כן".

תרגיל 7.10 (הטובים לטייס) בעבר החשוך הקשר בין התפקיד של אדם מסויים בצבא היה קשור באופן הדוק לתפקיד של אביו בצבא. נניח כי בצבא יש שלושה תפקידים אפשריים: טייס (T), חיל־רגלים (R) ושריונר (S). כמו כן, נניח שהסיכויים להתקבל לתפקידים נתונים באופן הבא:

- 80% מהילדים של טייסים נהפכו בסופו של דבר לטייסים. היתר הפכו להיות חי"רניקים.
- 30% מהילדים של חי"רניקים נהפכו לטייסים, 40% נשארו בחי"ר כמו הוריהם ו־30% סיימו בשריון.
- 20% מהילדים של שריונרים הפכו להיות טייסים, 10% עבור לחי"ר והיתר נותרו בשריון.

מצאו מה הסיכוי של ילד להיות טייס בהנחה שסבו היה טייס? מה לגבי מצב בו הסבא היה שריונר?

פתרון: נגדיר את קבוצת המצבים להיות $\{T, R, S\}$ כתפקיד של האדם הרלוונטי בשושלת (זאת אומרת, כל ילד יכול להיות טייס, שריונר או בחיל רגלים). נרשום את מטריצת המעבר על פי הנתונים בשאלה:

$$P = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.1 & 0.7 \end{bmatrix}.$$

בכדי למצוא את ההתפלגויות השונות, נחשב את P^2.

$$P^2 = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.1 & 0.7 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.1 & 0.7 \end{bmatrix} =$$
$$= \begin{bmatrix} 0.7 & 0.24 & 0.06 \\ 0.42 & 0.25 & 0.33 \\ 0.33 & 0.15 & 0.52 \end{bmatrix}.$$

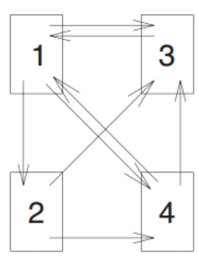
לכן, אם הסבא היה טייס אז בסיכוי 0.7 גם הנכד טייס, ואם הסבא היה שרוינר אז בסיכוי 0.33הנכד יהיה טייס.

7.2.1 הוקטור הסטציונרי ששווה מאות מיליארדי דולרים

<u>אלגוריתם גוגל למדרוג חשיבות אתרים</u>

יתרונו העיקרי של גוגל כאתר חיפוש על פני אתרים אחרים שצצו באותו זמן היה שהצליח לתת את ה"תוצאות הטובות" ראשונות. אבל איך מכמתים מוצלחות אתר? ואיך הם מצליחים לפגוע נכון בעזרת אלגוריתם מכני? כלומר: "איך נותנים ציונים לכל האתרים (בעלום) שיעידו על "חשיבותם"? מסתבר ששאלה זו פותרים בעזרת שרשראות מרקוב.

כשלב ראשון ממספרים את כל האתרים (כן, זה יצא דיי הרבה), ואפילו עמודים מסוימים בתוכם, ומסמנים את כל הלינקים בינהם. לדוגמא, ראו איור (3).



איור 3: רשת בעלת ארבעה אתרים.

כעת, נרצה לתת לכל אתר k ציון x_k המעיד על "חשיבותו". הדרך הפשוטה ביותר היא עפ"י מספר האתרים שמכוונים אליו. למשל, אצלנו נקבל $x_1 = 2, x_2 = 1, x_3 = 3, x_4 = 2$.

אך בדרך זו אנו לא נותנים משקל לאתר המכוון. כלומר, האתרים שלנו יקבלו את אותה תוספת משקל עבור לינק מיאהו (yahoo) ומאתר לא מוכר. כדי לפתור בעיה זו נרצה לשכלל את הרעיון המקורי ולתת ל־ x_k את סכום המשקולות המכוונות אליו. אצלנו, למשל, נקבל $x_1 = x_3 + x_4, x_4 = x_2 + x_1$. מאחר ול־ x_3, x_4 מקשרים יותר אתרים מ־ x_2, x_1, נקבל בדרך זו שאתר 1 "חשוב" יותר מאתר 4.

לרעיון זה חסרון ברור־ x_k מוגדר ע"י עצמו וכך הוא יכול לשפר את מצבו. אם הוא יוסיף קישורים לכל האתרים שמקשרים אליו, הוא ישפר את הניקוד שלהם וכך גם את שלו. כדי

לטפל בבעיה זו ננרמל את תוספת ה"חשיבות" שכל אתר יכול לתת. לכל אתר k נסמן ב־ n_k את מספר הלינקים שמופיעים ב־ x_k (לינקים מהאתר לעצמו לא נספרים). בנוסף, לכל k נסמן ב־ L_k את קבוצת האתרים שמקשרים ל־ k. כעת, נוכל להגדיר

$$x_k = \sum_{j \in L_k} \frac{x_j}{n_j}. \tag{16}$$

נשים לב שלכל k, x_k חיובי כלשהו (ייתכן שגדול מ־ 1).

כבר מן ההגדרה הנ"ל קיבלנו נוסחה מעגלית ־ החשיבות של אתר k תלויה בחשיבות של האתרים באחרים שהם בעצמם תלויים באתר k. כיצד מציגים את הבעיה הזאת בצורה יותר נוחה? בעזרת כתיב מטריציוני כמובן. ניקח לדוגמא את המצב המוצג באיור (3). נגדיר וקטור דירוגים $x = (x_1, x_2, x_3, x_4)$ ואת המטריצה

$$A = \begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix},$$

ונקבל שאת מערכת המשוואות (16) ניתן להציג על ידי

$$Ax = x.$$

או, באופן מפורש,

$$\begin{aligned} \begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \\ x_1 &= 1 \cdot x_3 + \frac{1}{2} x_4 \\ x_2 &= \frac{1}{2} x_1 \\ x_3 &= \frac{1}{3} x_1 + \frac{1}{2} x_2 + \frac{1}{2} x_4 \\ x_4 &= \frac{1}{3} x_1 + \frac{1}{2} x_2. \end{aligned}$$

קיבלנו את מערכת המשוואת (16). אם נמצא וקטור x כך ש־ $Ax = x$ פתרנו את הבעיה. כלומר, נרצה ו"ע עבור ע"ע 1 ואנו יודעים שאם יש כזה נוכל לבחור אחד שסכום האיברים בו הוא 1 (ו"ע יוצרים תת מרחב). אמנם המטריצה שקיבלנו איננה שרשרת מרקוב, אך נשים לב שהצמודה שלה, A^T, כן. זאת מאחר ובעמודה ה־ k יופיע בדיוק n_k פעמים $\frac{1}{n_k}$. בנוסף מתקיים

$$x^T A^T = (Ax)^T = x^T,$$

כלומר הוקטור שאנו מחפשים הוא בדיוק וקטור ההתפלגות הסטציונרית של מטריצת המעבר A^T!

האם ישנה סיבה מיוחדת לכך או שזה סתם יצא כך במקרה? ובכן, אם נסתכל על המטריצה A^T במקרה הנתון נקבל

$$A^T = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

וזו המטריצה המתקבלת אם מתייחסים לרשת המתוארת לעיל כמו גרף מקרי והשרשרת מרקוב עליו מוגדרת על ידי הילוך מקרי. הדירוג של גוגל למעשה מתבסס על הילוך מקרי על גרף (שמבוסס על הקישורים בין האתרים) ומוצג בעזרת שרשרת מרקוב מתאימה.

בעיות אפשריות־ עמודת/שורת אפסים (אתרים מבודדים) ושרשרת פירקה (אנו רוצים שהפתרון יהיה יחיד). בעיות אלו נפתרות בעזרת אלגברה לינארית־ מציאת ו"ע עבור הע"ע הגדול ביותר (ראינו בשיעור קודם שכל הע"ע קטנים או שווים ל־ 1) ומעבר למערכת

$$M = (1-m)\,A + mS$$

כאשר S מטריצת $n \times n$ (תחת ההנחה ש־A הינה $n \times n$) שכל רכיביה הינם $\frac{1}{n}$ ו־ $0 < m < 1$. הערך המקורי הינו $m = 0.15$, אך ניתן לקחת ערכים אחרים. פעולה זו מבטיחה שיתקבל וקטור סטציונרי יחיד (להמשך קריאה ־ Kurt Bryan, Tanya Leise, The 25,000,000,000 eigenvector. The linear algebra behind Google. SIAM Review, 48(3), 569-81.2006).

7.3 פריקות ומחזוריות

הגדרה 7.11 (שרשרת אי־פריקה) שרשרת מרקוב נקראת אי־פריקה אם לכל צמד מצבים $x, y \in S$ קיים שלב t כך ש־ $P^t(x,y) > 0$ או, במילים אחרות, אם $\Pr_x(X_t = y) > 0$.

שרשרת שאיננה אי־פריקה, תיקרא פריקה. המשמעות של אי־פריקות היא יחסית פשוטה. שרשרת אי־פריקה אומרת שתמיד נוכל להגיע מכל מצב x לכל מצב y באיזשהו שלב t מרגע ההגעה למצב x.

מושג נוסף וחשוב בשרשראות מרקוב הוא מחזור של מצב $x \in S$. לכל $x \in S$ נגדיר את הקבוצה

$$T(x) = \{t \geq 1 : \; P_x(X_t = x) > 0\}.$$

קבוצה זו מכילה את כל האורכים האפשריים למעגלים שכוללים את x. נקבע את המחזור של x להיות ה־ $gcd(T(x))$, ז"א המחלק הכי גדול של הקבוצה $T(x)$.

למה 7.12 אם השרשרת היא אי־פריקה, אז המחזור של כל המצבים הוא זהה ונקרא המחזור של השרשרת.

הערה 7.13

1. אם לכל $x \in S$ המחזור הוא 1 אז השרשרת תיקרא חסרת מחזור, אחרת השרשרת היא שרשרת מחזורית.

2. העובדה שקיימת התפלגות סטציונרית למטריצת מעבר לא אומרת שההתפלגות של שלב t, μ_t, מתכנסת ל־π. לדוגמא: אם השרשרת היא אי־פריקה ומחזורית, אזי μ_t לא מתכנסת בהכרח (עבור כל התפלגות התחלתית).

משפט 7.14 נתונה שרשרת מרקוב עם התפלגות התחלתית μ ויהי זמן $t \in \mathbb{N}$. נגדיר פונקציה π_t מ־S באופן הבא

$$\pi_t\left(x\right) = \frac{1}{t} E_\mu \left[\# \text{ of visitis to } x \text{ by stage t} - 1\right].$$

אזי

$$\pi_t \underset{t \to \infty}{\to} \pi$$

ו־ π זאת ההתפלגות הסטציונרית של השרשרת מרקוב.

המשפט הנ"ל למעשה מוכיח שלכל שרשרת מרקוב יש התפלגות סטציונרית. יחד עם זאת, אין זה אומר שלכל התפלגות התחלתית, תמיד נתכנס להתפלגות זאת. במקרה שיש אי־פריקות ומחזוריות אז לא תתקיים התכנסות (עבור התפלגויות התחלתיות מסויימות) וזאת מאחר וכל מספר שלבים נחזור חזרה להתפלגות ממנה התחלנו. לעומת זאת, במקרה של אי־פריקות וחוסר מחזוריות אין הדבר כך.

משפט 7.15 (המשפט הארגודי של שרשראות מרקוב) נתונה שרשרת מרקוב אי־פריקה וחסרת מחזור בעלת התפלגות סטציונרית π. קיימים קבועים $C > 0$ ו־ $0 < \alpha < 1$ כך שלכל התפלגות התחלתית μ ולכן מצב $x \in S$ מתקיים

$$\left|P_\mu\left(X_t = x\right) - \pi\left(x\right)\right| < C\alpha^t .$$

תרגיל 7.16 תהי קבוצת מצבים $S = \{0, 1, 2\}$ ומטריצת מעבר

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

מצאו התפלגות סטציונרית π עבור P וקבעו האם $\mu_t \underset{t \to \infty}{\to} \pi$ לכל התפלגות התחלתית μ.

פתרון: נמצא את ההתפלגות הסטציונרית π בעזרת חישוב ישיר. נניח כי $\pi = (a, b, 1-b-a)$ כאשר $0 \leq a, b, 1-b-a \leq 1$ ונקבל את מערכת המשוואות הבאות:

$$(a, b, 1-b-a) \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = (a, b, 1-b-a)$$

או בצורה מפורשת,

$$\begin{aligned} 1-b-a &= a \\ a &= b \\ b &= 1-b-a. \end{aligned}$$

נציב ונעביר אגפים ונקבל ש־ $a = 1 - 2a$ ולכן $a = b = \frac{1}{3}$, לכן ההתפלגות הסטציונרית היחידה היא $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. האם $\mu_t \underset{t\to\infty}{\rightarrow} \pi$? אם $\mu = \pi$ אז $\mu_t = \mu \cdot P^t = \pi \cdot P^t = \pi$ ולכן כמובן שההתפלגות נשמרת ואנו מתכנסים אליה באופן טריוויאלי. לעומת זאת, בכדי ש־ $\mu_t \underset{t\to\infty}{\rightarrow} \pi$ בלי קשר להתפלגות ההתחלתית μ צריך להתקיים ש־ $P^t \underset{t\to\infty}{\rightarrow} M$ כאשר

$$M = \begin{bmatrix} \pi(0) & \pi(1) & \pi(2) \\ \pi(0) & \pi(1) & \pi(2) \\ \pi(0) & \pi(1) & \pi(2) \end{bmatrix}.$$

זאת משום שאם $\mu = (\alpha, \beta, \gamma)$ אזי השיוויון הבא

$$(\alpha, \beta, \gamma) \cdot M = \pi$$

יתקיים לכל α, β, γ אם ורק אם המטריצה M היא מהצורה הנ"ל. אבל

$$\begin{aligned} P^2 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ P^3 &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ P^4 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \end{aligned}$$

וכמו שאנו רואים לא מתקיים ש־

$$P^t \underset{t\to\infty}{\rightarrow} M = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

לכן לא מתקיים ש־ $\mu_t \underset{t\to\infty}{\rightarrow} \pi$ לכל התפלגות התחלתית μ.

תרגיל 7.17 (שאלה מבחינה) לאלון יש מטבע אדום ולבן יש מטבע כחול. המטבע האדום נופל על "עץ" בסיכוי p והכחול נופל על "עץ" בסיכוי q. אלון ובן משחקים את המשחק הבא: בכל שלב הם מטילים את המטבעות, כל אחד את המטבע שנמצא אצלו. הם בשניהם יצא "עץ" אז הם מחליפים מטבעות, אחרת כל אחד נשאר עם המטבע שכרגע אצלו. יהי A_n המאורע בו בתום n תורות בן מחזיק את המטבע הכחול.

1. חשבו את $P\left(A_2\right)$ ו־$P\left(A_3\right)$.
2. קבעו לאילו ערכי $p,q\in[0,1]$ המאורעות A_2 ו־A_3 ב"ת.
3. קבעו לאילו ערכי $p,q\in[0,1]$ הגבול $\lim_{n\to\infty}\Pr\left(A_n\right)$ קיים וחשבו אותו, אם הוא קיים.

פתרון:

1. נפתור בעזרת שרשראות מרקוב. נגדיר את המצבים שלנו בתור $S=\{1,2\}$ כאשר מצב 1 הוא המצב שאלון עם המטבע האדום ו־2 זה המצב בו לאלון יש את המטבע הכחול. נסמן $a=pq$. מטריצת המעברים שלנו היא

$$P=\begin{bmatrix}1-a & a\\ a & 1-a\end{bmatrix}.$$

ההתפלגות ההתחלתית שלנו היא $\mu=(1,0)$. נשים לב ש־

$$P^2=\begin{bmatrix}1-a & a\\ a & 1-a\end{bmatrix}\begin{bmatrix}1-a & a\\ a & 1-a\end{bmatrix}=\begin{bmatrix}(1-a)^2+a^2 & 2a(1-a)\\ 2a(1-a) & (1-a)^2+a^2\end{bmatrix}$$

$$\begin{aligned}P^3&=\begin{bmatrix}(1-a)^2+a^2 & 2a(1-a)\\ 2a(1-a) & (1-a)^2+a^2\end{bmatrix}\begin{bmatrix}1-a & a\\ a & 1-a\end{bmatrix}=\\&=\begin{bmatrix}(1-a)^3+(1-a)a^2+2a^2(1-a) & a^3+(1-a)^2a+2a(1-a)^2\\ a^3+(1-a)^2a+2a(1-a)^2 & (1-a)^3+(1-a)a^2+2a^2(1-a)\end{bmatrix}=\\&=\begin{bmatrix}(1-a)^3+3a^2(1-a) & a^3+3a(1-a)^2\\ a^3+3a(1-a)^2 & (1-a)^3+3a^2(1-a)\end{bmatrix}.\end{aligned}$$

לכן ההסתברויות הרצויות נתונות על ידי

$$P^2\mu=\begin{bmatrix}(1-a)^2+a^2 & 2a(1-a)\\ 2a(1-a) & (1-a)^2+a^2\end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}(1-a)^2+a^2\\ 2a(1-a)\end{pmatrix}$$

שזו ההסתברות להיות במצב ההתחלתי אחרי 2 שלבים היא $(1-a)^2+a^2$ ו־$\Pr\left(A_2\right)$ ובאופן דומה $\Pr\left(A_3\right)=(1-a)^3+3a^2(1-a)$. כמובן שיכולנו לחשב זאת גם אחרת. בכדי שאחרי 2 סיבובים נחזור למצב המקורי יש 2 אפשרויות - או שלא יהיו חילופים כללי או שיהיו 2 חילופים. הסיכוי לכך הוא $a^2+(1-a)^2$. באופן דומה נוכל לחשב את $\Pr\left(A_3\right)$ מאחר ובמקרה זה צריך שיהיו או 0 חילופים או 2 חילופים והסיכוי לכך הוא $a^3+\binom{3}{2}a^2(1-a)$.

2. בכדי לבחון אי־תלות נבדוק את הערך של $\Pr\left(A_3|A_2\right)$. אם בשלב השני לאלון יש את המטבע האדום, הסיכוי שזה יישמר גם לאחר השלב השלישי הוא $(1-a)$. נדרוש ש־ $\Pr\left(A_3|A_2\right)=\Pr\left(A_3\right)$ ונקבל

$$(1-a)^3+3a^2(1-a)=1-a\,.$$

לכן הפתרון הראשון הוא $a_1 = pq = 1$.

$$\begin{aligned}(1-a)^2 + 3a^2 &= 1\\ 1 - 2a + a^2 + 3a^2 &= 1\\ 4a^2 &= 2a\\ a_2 &= 0;\\ a_3 &= \frac{1}{2}.\end{aligned}$$

לכן בכדי שתהיה אי־תלות צריכים ש־ $p = 0$ או $q = 0$ או $p = q = 1$ או $pq = \frac{1}{2}$.

3. נשים לב שאם $a = 0$ אז תמיד נשארים במצב ההתחלתי ולכן הגבול קיים והוא $\lim_{n\to\infty} \Pr(A_n) = 1$. לעומת זאת, אם $a = 1$ אז בכל שלב יש החלפה והגבול לא קיים. בנוסף, אם $0 < a < 1$ אז השרשרת הנתונה היא שרשרת מחזורית (בכל שלב יש הסתברות חיובית שנשאר במצב עצמו, לכן המחזור של כל המצבים הוא 1) ואי־פריקה (ניתן לעבור מכל מצב למצב האחר בשלב 1 בהסתברות חיובית) ולכן יש התכנסות של השרשרת לכל התפלגות התחלתית μ. ההתכנסות היא לוקטור העצמי π של המטריצה עם ערך עצמי 1 ומתקיים ש־ $\lim_{n\to\infty} P^n\mu = \pi$. נמצא את הוקטור העצמי של המטריצה. נוכל לבצע חישוב מפורש ולמצוא את הוקטור העצמי, מצד שני ניתן לשים לב שמדובר במטריצה דו־סטוכסטית ולכן $\pi = \left(\frac{1}{2}, \frac{1}{2}\right)$ זה וקטור עצמי של P עם ערך עצמי 1.

תרגיל 7.18 יהי גרף $G = (V, E)$ ונניח כי מבצעים על הגרף הילוך מקרי פשוט ־ אם נמצאים בשלב t בקודקוד $x \in V$ אז עוברים בהסתברות $\frac{1}{deg(x)}$ לכל קודקוד y כך ש־ $(x, y) \in E$.

1. הגדירו שרשרת מרקוב ע"ב ההילוך המקרי הנ"ל.
2. הוכיחו כי השרשרת אי־פריקה אם ורק אם הגרף קשיר.
3. הניחו כי G הוא גרף קשיר. הוכיחו כי השרשרת היא בעלת מחזור 2 אם ורק אם הגרף דו־צדדי. אחרת השרשרת בעלת מחזור 1.
4. הניחו כי G הוא גרף קשיר. הוכיחו כי ההתפלגות הסטציונרית בקודקוד $x \in V$ שווה ל־ $\frac{deg(x)}{\sum_{y\in V} deg(y)}$.

פתרון:

1. הילוך מקרי על גרף פשוט נותן לנו באופן מייד שרשרת מרקוב. קבוצת המצבים שלנו היא קבוצת הקודקודים, $S = V$. מטריצת המעבר P מוגדרת באופן הבא

$$P(x, y) = \begin{cases} \frac{1}{deg(x)} & (x, y) \in E \\ 0 & (x, y) \notin E \end{cases}.$$

נשים לב שהגדרנו אכן מטריצת מעבר מאחר ו־

$$\sum_{y\in S} P(x, y) = \sum_{y\in S:\ (x,y)\in E} \frac{1}{deg(x)} = 1,$$

ובנוסף $P(x, y) \in [0, 1]$ לכל $x, y \in S$.

2. יהיו $x,y\in S$. אנו נוכיח טענה כללית שאומרת כי קיים מסלול באורך t מ־x_0 ל־ x_t,$(x_0,x_1,\ldots,x_{t-1},x_t)$, אם ורק אם קיים t כך ש־ $P^t\left(x_0,x_t\right)>0$. נוכיח טענה זאת באינדוקציה על $t\leq 1$. בסיס האינדוקציה עבור $t=1$ הוא מיידי כי אם $\left(x_0,x_1\right)\in E$ אז לפי הגדרת מטריצת המעבר נובע ש־ $P^1\left(x_0,x_1\right)>0$ ולהיפך. נניח כי הטענה נכונה עבור t ונוכיח עבור $t+1$.

<u>כיוון ראשון:</u> נניח שקיים מסלול באורך $t+1$ מ־x_0 ל־ x_{t+1},$(x_0,x_1,\ldots,,x_{t+1})$, ונוכיח כי $P^{t+1}\left(x_0,x_{t+1}\right)>0$. אם $P\left(x_t,x_{t+1}\right)>0$ (כי $\left(x_t,x_{t+1}\right)\in E$) מתקיים

$$\begin{aligned}P^{t+1}\left(x_0,x_{t+1}\right) &= \left(P^tP\right)\left(x_0,x_{t+1}\right)=\\ &= \sum_{x\in S}P^t\left(x_0,x\right)P\left(x,x_{t+1}\right)\geq\\ &\geq P^t\left(x_0,x_t\right)P\left(x_t,x_{t+1}\right)>0\end{aligned}$$

כאשר $P^t\left(x_0,x_t\right)>0$ מהנחת האינדוקציה ובגלל שקיים מסלול באורך t מ־x_0 ל־ x_t. לכן הכיוון הראשון נובע באינדוקציה.

<u>כיוון שני:</u> נניח ש־ $P^{t+1}\left(x_0,x_{t+1}\right)>0$ ונוכיח שקיים מסלול באורך $t+1$ מ־x_0 ל־ x_{t+1},$(x_0,x_1,\ldots,,x_{t+1})$. נשים לב כי

$$\begin{aligned}0<P^{t+1}\left(x_0,x_{t+1}\right) &= \left(P^tP\right)\left(x_0,x_{t+1}\right)=\\ &= \sum_{x\in S}P^t\left(x_0,x\right)P\left(x,x_{t+1}\right)\end{aligned}$$

ואנו יודעים שישנו $x\in S$ כך ש־ $P^t\left(x_0,x\right)P\left(x,x_{t+1}\right)>0$. זאת אומרת ש־ $P^t\left(x_0,x\right)>0,\ P\left(x,x_{t+1}\right)>0$ ולפי הנחת האינדוקציה נובע שקיים מסלול באורך t מ־x_0 ל־x ומאחר ו־ $P\left(x,x_{t+1}\right)>0$ קיים גם מסלול באורך $t+1$ מ־x_0 ל־x_{t+1}.

3. אם הגרף קשיר ודו־צדדי, משמע שניתן להגיע מכל קודקוד, חזרה לעצמו ב־2 צעדים. בנוסף, ידוע שגרף הוא דו־צדדי אם ורק אם אין בו מעגל מאורך אי־זוגי ולכן כל מעגל הוא מאורך זוגי ובפרט המחזור של השרשרת הוא 2. מצד שני, אם נניח שהגרף קשיר והשרשרת היא בעלת מחזור 2 אזי נובע שכל מעגל בגרף הוא באורך זוגי ובפרט לא קיימים בגרף מעגלים באורך אי־זוגי לכן, ע"פ אותו המשפט, הגרף הוא דו־צדדי.
אז מדוע במקרה בו הגרף איננו דו־צדדי המחזור של השרשרת הוא 1? אם המעגל איננו דו־צדדי משמע שיש בו מעגל באורך אי־זוגי. אם יש בו מעגל באורך אי־זוגי ובנוסף יש בו מעגל באורך 2 בכל קודקוד, נובע שהמחלק הגדול ביותר האפשרי הוא 1. שימו לב שבמקרה זה צריך להשתמש בכך שגרף קשיר גורר אי־פריקות של השרשרת מרקוב ולכן לכל המצבים יש את אותו המחזור.

4. נוכיח זאת על ידי בדיקה ישירה ובשילוב המשפט שאומר שאם שרשרת מרקוב היא אי־פריקה אז ההתפלגות הסטציונרית שלה היא יחידה. מאחר ו־π היא התפלגות

סטציונרית נובע ש־

$$\begin{aligned}\pi(x) &= \sum_{y\in S}\pi(y)P(y,x) = \\ &= \sum_{y\in S:\ (x,y)\in E}\pi(y)\frac{1}{deg(y)} = \\ &= \sum_{y\in S:\ (x,y)\in E}\left(\frac{deg(y)}{\sum_{z\in S}deg(z)}\right)\frac{1}{deg(y)} = \\ &= \sum_{y\in S:\ (x,y)\in E}\frac{1}{\sum_{z\in S}deg(z)} = \\ &= \frac{deg(x)}{\sum_{z\in S}deg(z)} = \frac{deg(x)}{\sum_{y\in V}deg(y)}\end{aligned}$$

כנדרש.

תרגיל 7.19 גאריק ובנץ אחים החולקים שלושה כובעים. שלושת הכובעים בצבעים לבן, שחור ואדום. בתחילת השנה לבש אריק את הכובע הלבן ובנץ את הכובע השחור בעוד הכובע האדום נשאר בצד. מאז, בכל יום הם בוחרים את אחד האחים באופן אחיד ואח זה מחליף בין הכובע שלבש אתמול לכובע שבצד בעוד האח השני נשאר עם הכובע שלבש אתמול.

1. נסמן ב־X_n את כמות הפעמים שלבש אריק את הכובע השחור ב־n הימים הראשונים של השנה. חשבו את הערך $\frac{E[X_n]}{n}$ כאשר n שואף לאינסוף?
2. חשבו את ההסתברות שאריק לובש את הכובע האדום ובנץ את הכובע השחור ביום ה־n־י של השנה, כאשר n שואף לאינסוף?
3. בשנת 2015 יצא הכובע האדום מעט מהאופנה ועל כן החליטו אריק ובנץ על השיטה הבאה: בתחילת השנה לובש אריק את הכובע הלבן ובנץ את הכובע השחור. בכל יום לאחר מכן, אם הכובע הלבן או השחור מונח בצד, יפעלו כמו בעבר ע"י בחירת אח באופן אחיד והחלפת הכובע שלו לכובע שמונח בצד. לעומת זאת, אם הכובע האדום מונח בצד אז בסיכוי $1/3$ יתחלפו אריק ובנץ בכובעיהם מהיום הקודם, ובסיכוי $2/3$ יפעלו כמו בעבר, כלומר יבחרו אח באופן אחיד ויחליפו בין הכובע שלו מאתמול לכובע האדום. השאלות הבאות מתייחסות לשיטה זו. חשבו את ההסתברות שבנץ לובש את הכובע השחור ביום ה־n־י של השנה, כאשר n שואף לאינסוף?
4. חשבו את ההסתברות שבנץ לובש את הכובע האדום ביום ה־n־י של השנה, כאשר n שואף לאינסוף?

פתרון:

1. נתחיל בבניית השרשרת. ישנם 6 מצבים אפשריים $\{BW, WB, BR, RB, WR, RW\}$ומטריצת

המעבר היא

$$.P=\begin{bmatrix}0&0&0.5&0&0&0.5\\0&0&0&0.5&0.5&0\\0.5&0&0&0&0.5&0\\0&0.5&0&0&0&0.5\\0&0.5&0.5&0&0&0\\0.5&0&0&0.5&0&0\end{bmatrix}$$

זו שרשרת אי־פריקה דו סטוכסטית שבה לכל מצב יש התפלגות סטציונרית של $\frac{1}{6}$. למעשה מדובר בהילוך מקרי על מעגל באורך 6. לכן הפתרון הגבולי הוא $\frac{1}{3}$.

2. השרשרת המתוארת היא מחזורית (מחזור 2 כמו כל הילוך מקרי על מעגל באורך זוגי) ולכן אין למצב הסתברות גבולית.

3. נרשום את מטריצת המעבר מחדש.

$$.P=\begin{bmatrix}0&\frac{1}{3}&\frac{1}{3}&0&0&\frac{1}{3}\\\frac{1}{3}&0&0&\frac{1}{3}&\frac{1}{3}&0\\0.5&0&0&0&0.5&0\\0&0.5&0&0&0&0.5\\0&0.5&0.5&0&0&0\\0.5&0&0&0.5&0&0\end{bmatrix}$$

קיבלנו הילוך מקרי פשוט על מעגל באורך 6 כאשר חיברנו זוג קודקודים מסויים. ראינו כי בשרשרת של הילוך מקרי פשוט על גרף, ההתפלגות הסטציונרית פרופורציונאלית לדרגות. לכן ההתפלגות הסטציונרית היא

$$.\pi=\left(\frac{3}{14},\frac{3}{14},\frac{2}{14},\frac{2}{14},\frac{2}{14},\frac{2}{14}\right)$$

דרך אחרת לחישוב ההתפלגוץ הסטציונרית היא על ידי ההבנה שיש 2 מצבים (1 ו־2) שבהם ההתפלגות היא שווה, ונסמנה ב־x וביתר המצבים, משיקולי סימטריה, ההתפלגות שווה גם כן, נסמנה ב־y. נרשום את המשוואות להתפלגות ונקבל:

$$x=\frac{1}{3}x+\frac{1}{2}y+\frac{1}{2}y$$
$$2x+4y=1$$

כעת כל שנותר זה לפתור מערכת של 2 משוואות ב־2 נעלמים.
נשים לב שממצב 1 ניתן לעבור למצבים 2,3 ו־6 וממצבים 2,3 ו־6 ניתן לעבור למצב 1,4,5. אותו הדבר נכון למצבים 4,5 וכן הלאה. למעשה בזמנים האי־זוגיים, בהנחה שאנו נמצאים בשלב הראשון במצב 1, אז תמיד נהיה במצבים 1,4,5 ובזמנים הזוגיים תמיד נהיה במצבים 2,3,6. במידה ורושמים את השרשראות לימים הזוגיים והאי־זוגיים אז מקבלים צמד שרשראות לא־מחזוריות (כי ניתן להישאר במצב נתון) ולא פריקות ולכן מתכנסות להתפלגות הסטציונרית שלהן. המאורע המבוקש הוא בעל הסתברות סטציונרית $\frac{3}{7}=\frac{\frac{3}{14}}{\frac{3}{14}+\frac{2}{14}+\frac{2}{14}}$ בימים האי־זוגיים, ולעומת זאת בזמנים הזוגיים הסתברות הגבולית היא $\frac{2}{7}=\frac{\frac{2}{14}}{\frac{3}{14}+\frac{2}{14}+\frac{2}{14}}$ ולכן אין התכנסות.

4. אומנם מדובר בשרשרת שלא מתכנסת, אבל באופן דומה לסעיף הקודם, המאורע המבוקש מתייחס לשני מצבים אפשריים ־ אחד בזמנים הזוגיים (מצב 6) ואחד בזמנים האי־זוגיים (מצב 4). בשני המקרים ההתפלגות הסטציונרית היא

$$\frac{\frac{2}{14}}{\frac{3}{14}+\frac{2}{14}+\frac{2}{14}}=\frac{2}{7}$$

ולכן יש התכנסות וזו ההסתברות הגבולית.

תרגיל 7.20 הוכיחו כי כל הערכים העצמיים של מטריצת מעבר P קטנים או שווים ל־1 בערכם המוחלט. בנוסף, הראו ש־$\lambda=1$ הוא ערך עצמי של P (עם וקטור עצמי ימני) ומצאו וקטור עצמי $v\neq\overrightarrow{0}$ (וקטור עמודה) שמתאים לע"ע זה.

פתרון: נוכיח תחילה את הטענה כי כאשר לוקחים מטריצת מעבר אחת P ומכפילים אותה במטריצת מעבר שנייה Q מקבלים גם מטריצת מעבר PQ. נסתכל על $(PQ)(i,j)$. ידוע כי

$$(PQ)(i,j)=\sum_{x\in S}P(i,x)\,Q(x,j)\leq\sum_{x\in S}P(i,x)\cdot1=1\,,$$
$$(PQ)(i,j)=\sum_{x\in S}P(i,x)\,Q(x,j)\geq\sum_{x\in S}P(i,x)\cdot0=0\,.$$

בנוסף נשים לב כי

$$\begin{aligned}\sum_{j\in S}(PQ)(i,j)&=\sum_{j\in S}\sum_{x\in S}P(i,x)\,Q(x,j)=\\&=\sum_{x\in S}P(i,x)\sum_{j\in S}Q(x,j)=\\&=\sum_{x\in S}P(i,x)\cdot1=1\end{aligned}$$

ולכן קיבלנו מטריצת מעבר חדשה בה כל האיברים הם אי־שליליים וסכום כל שורה הוא 1. נניח בשלילה שקיים ע"ע λ כך ש־ $|\lambda|>1$ ולכן קיים גם וקטור עצמי v כך ש־

$$vP=\lambda v$$

ואם נכפיל את צמד האגפים במטריצת מעבר מימין נקבל ש־

$$vP^2=\lambda vP=\lambda^2v.$$

כאשר נבצע את אותה הפעולה $n\gg1$ פעמים נקבל באגף ימין וקטור עם אורך שגדל אקספוננציאלית

$$vP^n=\lambda^nv.$$

חשוב לזכור כי v הוא וקטור עצמי כלשהו לא טריוויאלי ולכן בעל אורך חיובי וכאשר מוכפל ב־ λ^n, אורכו שואף לאינסוף. על כן, חייב להיות תא במטריצה P^n עם ערך גדול מ־1.

אבל הטענה מוכיחה כי מדובר במטריצת מעבר ולכן כל התאים במטריצה נעים בין 0 ל־1. סתירה. הראנו שכל הערכים העצמיים בעלי ערך מוחלט קטן או שווה ל־1. נוכיח כעת כי $\lambda = 1$ הוא ע"ע ונמצא ו"ע שמתאים לו. מאחר וסכום הערכים בכל שורה הוא 1, אזי כל וקטור עמודה קבוע v^T (שכל הקוארדינטות שלו שוות, נניח ל־c) שנכפול במטריצה ייתן,

$$\begin{aligned} \left(Pv^T\right)_i &= \sum_{x\in S} P(i,x)\, c = c\sum_{x\in S} P(i,x) = c \\ \Rightarrow Pv^T &= v^T = 1v^T. \end{aligned}$$

זאת אומרת, כל וקטור קבוע הוא וקטור עצמי עם ע"ע 1.

תרגיל 7.21 הוכיחו כי ההתפלגות הסטציונרית π של שרשרת אי־פריקה נותנת הסתברות חיובית לכל מצב $x \in S$.

פתרון: נניח בשלילה שקיים מצב x כך ש־$\pi(x) = 0$. מאחר וזו התפלגות קיים לפחות מצב אחד נוסף $y \neq x$ כך ש־ $\pi(y) > 0$. השרשרת היא אי־פריקה ומכאן נובע שקיים t כך ש־ $P^t(y,x) > 0$. נשים לב כי

$$.\pi P^t = \pi P P^{t-1} = \pi P^{t-1} = \cdots = \pi$$

על כן נסיק כי

$$\begin{aligned} \pi(x) = \left(\pi P^t\right)(x) &= \\ &= \sum_{z\in S} \pi(z)\, P^t(z,x) \geq \\ &\geq \pi(y)\, P^t(y,x) > 0. \end{aligned}$$

סתירה.

תרגיל 7.22 הוכיחו כי אם שרשרת אי־פריקה היא בעלת מחזור p אז ניתן לחלק את מרחב המצבים ל־p קבוצות זרות $S_1, \ldots, S_p$ כך שאם $x \in S_i$ ו־$P(x,y) > 0$ אז $y \in S_{i+1}$ (אם $i = p$ אז $y \in S_1$).

פתרון: יהי $x_0 \in S$. לכל $k = 1, \ldots, p$ נגדיר

$$.S_k = \left\{x \in S :\ P^{mp+k}(x_0, x) > 0 \text{ for some m}\right\}$$

מן הנתון שהשרשרת היא אי־פריקה נובע בהכרח שכל מצב $x \in S$ שייך לקבוצה כלשהי. נוכיח כי כל $x \in S$ שייך לקבוצה S_k אחת בלבד. נניח כי $P^{mp+k}(x_0,x) > 0$ ו־ $P^{mp+j}(x_0,x) > 0$ כאשר $j \leq k$ בה"כ. מאחר והשרשרת היא אי־פריקה, קיים t כך ש־ $P^t(x, x_0) > 0$. לכן $mp + k + t \in T(x_0)$. בגלל שהשרשרת היא בעלת מחזור p נובע ש־ p מחלק את $k + t$. באופן דומה נקבל ש־p מחלק את $j + t$, על כן p מחלק את $k - j = k + t - (j + t)$. ידוע כי $j \leq k \leq p$ ולכן בהכרח נובע ש־ $j = k$. לסיכום, הוכחנו כי $S_1, \ldots, S_p$ היא חלוקה של S לקבוצות זרות.

נניח כי $x \in S_i$ ו־$P(x,y) > 0$. לפי הגדרה קיים m כך ש־ $P^{mp+i}(x_0,x) > 0$ ולכן

$$, P^{mp+i+1}(x_0,y) \geq P^{mp+i}(x_0,x)\, P(x,y) > 0$$

מכאן נובע ש־ $y \in S_{i+1}$ (ואם $i = p$ אז $y \in S_1$).

תרגיל 7.23 הוכח שאם $z \neq 1$ בעל $|z| = 1$ הוא ערך עצמי של מטריצת המעבר של שרשרת אי־פריקה, אז קיים p כך ש־$z = exp(2\pi i/p)$.

פתרון: נסתכל על המטריצה $Q = P^T$. אנו יודעים שיש לה אותם ע"ע כמו ל־P. ניקח וקטור עצמי v של Q עם הע"ע z וננרמל אותו כך שבמצב נתון x_0, $v(x_0) = 1$ ויתר הקואורדינטות הן קטנות או שוות ל־1 בערכן המוחלט. נשים לב כי

$$\begin{aligned} z &= zv(x_0) = \\ &= (vQ)(x_0) = \\ &= \sum_{y \in S} v(y) Q(y, x_0) = \\ &= \sum_{y \in S} v(y) P^T(y, x_0) \end{aligned}$$

נשים לב שקיבלנו ממוצע משוקלל של ערכי $v(y)$ שכולם קטנים או שווים ל־1 בערכם המוחלט ויחד עם זאת הממוצע שלהם שווה ל־z. על כן, בכדי שערכו המוחלט של הערך הממוצע יהיה 1 והם יהיו שווים ל־z, כלל הערכים $v(y)$ בסכום כך ש־ $Q(y, x_0) > 0$ חייבים להיות שווים ל־z. נסמן קבוצה זאת ב־S_1. ניקח איבר כלשהו $x_1 \in S_1$. נבצע את אותו החישוב הנ"ל עבור x_1.

$$\begin{aligned} z^2 &= zv(x_1) = \\ &= (vQ)(x_1) = \\ &= \sum_{y \in S} v(y) Q(y, x_1) = \\ &= \sum_{y \in S} v(y) P^T(y, x_1) \end{aligned}$$

ונקבל קבוצה חדשה של איברים, S_2, כך ש־ $v(y) = z^2$ לכל $y \in S_2$. נמשיך באותו האופן בצורה אינדוקטיבית. מאחר וקבוצת המצבים סופית, נקבל בשלב מסויים קבוצה המכילה איברים שהופיעו בקבוצה עם אינדקס קטן יותר ולכן נובע שקיימת חזקה p כך ש־ $z^p = 1$ כנדרש.

7.4 דוגמאות נוספות

תרגיל 7.24 לדן יש סכום התחלתי של 3 שקלים ולערן יש סכום התחלתי של 2 שקלים. בכל שלב, כל עוד אף אחד מהם לא התרושש, כל אחד מהם שם שקל אחד על השולחן. מתקיימת הגרלה בה דן זוכה בסכום שעל השולחן בסיכוי p וערן זוכה בו בסיכוי $q = 1 - p$, $0 < p < 1$. התהליך מסתיים כאשר אחד מהם מתרושש. מה הן ההתפלגויות הסטציונריות?

פתרון: נסמן כמרחב המצבים S את הסכום שיש לדן. אזי, יש כאן שרשרת בת 6 מצבים,

כאשר מצב i מייצג את המצב שבידי דן יש i שקלים. מטריצת המעבר נראית כך

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & 0 \\ 0 & 0 & q & 0 & p & 0 \\ 0 & 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

שימו לב שממצבים 0 ו־ 5 לא ניתן לצאת־ אחד מהם התרושש (נקראים גם מצבים סופגים). בפרט, קיבלנו שהשרשרת פריקה. נחשב התפלגות סטציונרית

$$\pi P = \pi, \quad \pi = (\pi_1, \ldots, \pi_6)$$

נחשב,

$$(\pi_1 + q\pi_2, q\pi_3, p\pi_2 + q\pi_4, p\pi_3 + q\pi_5, p\pi_4, p\pi_5 + \pi_6) = (\pi_1, \ldots, \pi_6)$$

$$\Rightarrow \begin{cases} \pi_1 + q\pi_2 = \pi_1 & \Rightarrow \pi_2 = 0 \\ q\pi_3 = \pi_2 & \Rightarrow \pi_3 = 0 \\ p\pi_2 + q\pi_4 = \pi_3 & \Rightarrow \pi_4 = 0 \\ p\pi_3 + q\pi_5 = \pi_4 & \Rightarrow \pi_5 = 0 \\ p\pi_4 = \pi_5 & \\ p\pi_5 + \pi_6 = \pi_6 & \end{cases}$$

כלומר קיבלנו ש־ $\pi_2 = \pi_3 = \pi_4 = \pi_5 = 0$, אך אין הגבלות על π_1, π_6. קיבלנו שקבוצת הוקטורים הסטציונרים הם

$$\{(\alpha, 0, 0, 0, 0, 1 - \alpha) \ : \ 1 \geq \alpha \geq 0\}.$$

הגדרה 7.25 שרשרת מרקוב נקראת דו־סטוכסטית אם גם כל עמודה בה נסכמת ל־ 1 (כלומר, $\sum_i P(i,j) = 1$ לכל $j \in S$).

תרגיל 7.26 מצאו את ההתפלגות הסטציונרית של שרשרת דו־ סטוכסטית אי־פריקה בעלת מרחב מצבים $|S| = n$.

פתרון: נראה ש־ $\pi = ({}^1\!/\!{}_n, \ldots, {}^1\!/\!{}_n)$ היא התפלגות סטציונרית. נחשב את האיבר ה־ i בוקטור πP

$$(\pi P)_i = \sum_{j \in S} {}^1\!/\!{}_n \cdot P(j,i) = {}^1\!/\!{}_n \cdot \sum_{j \in S} P(j,i) = {}^1\!/\!{}_n \cdot 1 = {}^1\!/\!{}_n$$

וקיבלנו, $\pi P = \pi$ כנדרש. כעת, עפ"י משפט הוקטור הסטציונרי של שרשרת מרקוב אי־פריקה הינו יחיד.

תרגיל 7.27 (עירבוב קלפים) רוצים לערבב חפיסת קלפים בגודל n. בכל שלב לוקחים את הקלף העליון ושמים אותו באקראי באחד מ־ n המקומות האפשריים. הציגו את הבעיה כשרשרת מרקוב ומצאו לה התפלגות סטציונרית.

פתרון: ראשית, מרחב המצבים S הוא כל הפרמוטציות האפשריות על הקלפים, בפרט $|S| = n!$. עבור $x, y \in S$ נסמן $x \sim y$ אם ניתן לקבל את y מ־ x בצורה זו (שימו לב, זהו אינו יחס סימטרי). מטריצת המעבר היא

$$P(x,y) = \begin{cases} \frac{1}{n} & x \sim y \\ 0 & x \nsim y \end{cases},$$

מאחר וישנם בדיוק n סידורים שניתן לעבור אליהם בצורה כזו (לכל בחירה של מיקום חדש לקלף העליון ב־ x). נשים לב ששרשרת זו היא אי־פריקה (ע"י מספר סופי של החלפות נוכל להגיע מכל סידור x לכל סידור y) וחסרת מחזור (נוכל להשאיר את הקלף הראשון במקום הראשון). בנוסף, שרשרת זו היא דו־סטוכסטית. זאת מאחר ולכל y יש גם בדיוק n סידורים שונים שניתן להגיע מהם ל־ y (נעשה את הצעד ההפוך־ נקח קלף מתוך החפיסה ונשים אותו עליון), לכן

$$\sum_{y \in S} P(x,y) = n \cdot \frac{1}{n} = 1.$$

כעת, לפי תרגיל קודם, ההתפלגות הסטציונרית היא $\pi = \left(\frac{1}{n!}, \ldots, \frac{1}{n!}\right)$.

8 תרגילים ממבחנים ישנים

תרגיל 8.1 (סוגיה) בקופסא יש X כדורים אדומים ו־ $(100 - X)$ כדורים שחורים. כאשר X הינו מ"מ עם תמונה ב־ $\{0, \ldots, 100\}$ כך ש־ $E[X] = 25$. דוגמים שני כדורים עם החזרה.

1. הסיכוי שהכדור השני הוא אדום הוא:

   (א) $\frac{1}{4}$.

   (ב) $\frac{1}{2}$.

   (ג) $\frac{24}{100}$.

   (ד) לא ניתן לקבוע.

   פתרון: נסמן A_2 המאורע שהכדור השני אדום, ונחשב

$$\begin{aligned}\Pr(A_2) &= \sum_{k=0}^{100} \Pr(A_2 \mid X = k) \Pr(X = k) = \\ &= \sum_{k=0}^{100} \frac{k}{100} \Pr(X = k) = \\ &= \frac{1}{100} \sum_{k=0}^{100} k \Pr(X = k) =\end{aligned}$$

   מתקיים $\sum_{k=0}^{100} k \Pr(X = k) = E[x] = 25$, ולכן

$$= \frac{1}{100} \cdot 25 = \frac{1}{4}$$

   כלומר, התשובה הנכונה היא (א).

2. הצבעים של שני הכדורים הינם ב"ת.

   (א) תמיד.

   (ב) לעולם לא.

   (ג) רק כאשר X מנוון (קבוע).

   (ד) התשובה תלויה בהתפלגות הלא מנוונת של X.

   פתרון: התשובה הנכונה היא (ג). כיוון שכפי שראינו בדוגמאות במהלך הסמסטר, כאשר X אינו מנוון נוכל ללמוד מהתוצאה הראשונה על התוצאה השניה ולהפך. לעומת זאת, כאשר X מנוון, אז הוא בהכרח שווה זהותית ל־25 ואז הניסויים הופכים להיות ב"ת (הסיכוי לכל כדור אדום הוא $1/4$ באופן ב"ת בכדורים האחרים).

3. כאשר X לא מנוון, ההסתברות לכך ששני הכדורים אדומים היא

(א) $\frac{1}{16}$.

(ב) גדולה מ־ $\frac{1}{16}$.

(ג) קטנה מ־ $\frac{1}{16}$.

(ד) תלויה בהתפלגות X.

פתרון: נסמן A_1 המאורע שהראשון אדום. אזי מחפשים את $A_1 \cap A_2$.

דרך I: באותו אופן כמו בסעיף 1 מתקיים $\Pr(A_1) = \frac{1}{4}$. בנוסף, באופן דומה

$$\begin{aligned}
\Pr(A_2 \mid A_1) &= \sum_{k=0}^{100} \Pr(A_2 \mid A_1 \cap X = k) \Pr(X = k \mid A_1) = \\
&= \sum_{k=0}^{100} \Pr(A_2 \mid X = k) \Pr(X = k \mid A_1) = \\
&= \sum_{k=0}^{100} \frac{k}{100} \Pr(X = k \mid A_1) = \\
&= \frac{E[X \mid A_1]}{100} > \frac{1}{4}
\end{aligned}$$

כאשר השיוויון הראשון נובע מכך שבהינן ערך X, A_1 ו־ A_2 בלתי תלויים, ואי השיוויון האחרון נובע מכך שבהינתן שהכדור הראשון היה אדום ישנו סיכוי גבוה יותר ש־ X גדול מאשר ללא ההתניה (כאשר X אינו מנוון). נציב ונקבל

$$\Pr(A_1 \cap A_2) = \Pr(A_1) \Pr(A_2 \mid A_1) > \frac{1}{16}.$$

דרך II:

$$\begin{aligned}
\Pr(A_1 \cap A_2) &= \sum_{k=0}^{100} \Pr(A_2 \cap A_1 \mid X = k) \Pr(X = k) = \\
&= \sum_{k=0}^{100} \left(\frac{k}{100}\right)^2 \Pr(X = k) = \\
&= \left(\frac{1}{100}\right)^2 E\left[X^2\right] = \\
&= \frac{E[X]^2 + V(X)}{100^2} = \\
&= \frac{25^2 + V(X)}{100^2} > \\
&> \frac{1}{16}
\end{aligned}$$

כאשר אי השיוויון נובע מכך ש־ $V(X) > 0$ (תמיד אי שלילי, ואינו אפס מאחר שנתון כי X אינו מנוון). בשני המקרים קיבלנו את תשובה (ב).

4. נניח כעת ש־ $\sigma(X)=25$. מה הסיכוי ששניהם אדומים?

(א) $\frac{1}{16}$.
(ב) $\frac{1}{8}$.
(ג) $\frac{1}{32}$.
(ד) תלויה בהתפלגות X.

פתרון: נוכל להשתמש במה שקיבלנו בסעיף קודם (X אינו מנוון)

$$\Pr(A_1\cap A_2)=\frac{25^2+V(X)}{100^2}=\frac{25^2+25^2}{100}=\frac{1}{8}.$$

קיבלנו (ב). שימו לב שהיינו יכולים לפסול את סעיפים (א) ו־ (ג) עפ"י פתרון שאלה קודמת.

תרגיל 8.2 (סוגיה) מפזרים 20 כדורים (עם שמות מ־1 עד 20) בשלושה כדים, כך שכל 3^{20} תוצאות הפיזור שוות סיכוי. נסמן ב־ X_i את מספר הכדורים בעלי שם זוגי בכד ה־i ($i=1,2,3$), וב־ Y_i את מספר הכדורים בעלי שם אי־זוגי בכד ה־ i. (באם כל התשובות נראות לך לא נכונות סמן (ד))

.1

(א) X_1 ו־ X_2 תלויים, אך X_1 ו־ Y_1 בלתי תלויים.
(ב) X_1 ו־ X_2 בלתי תלויים, אך X_1 ו־ Y_1 תלויים.
(ג) X_1 ו־ X_2 תלויים, וגם X_1 ו־ Y_1 תלויים.

פתרון: ראשית ברור כי X_1 ו־ X_2 תלויים, זאת מאחר ו־ $X_1+X_2+X_3=10$, בפרט

$$\Pr(X_1=10)\Pr(X_2=10)\neq 0=\Pr(X_1=X_2=10).$$

נבדוק האם X_1 ו־ Y_1 תלויים. מתקיים

$$X_i,Y_i\sim\text{Bin}\left(10,\frac{1}{3}\right).$$

כעת, לכל $0\le l,k\le 10$

$$\Pr(X_1=l,Y_1=k)=\binom{10}{l}\left(\frac{1}{3}\right)^l\left(\frac{2}{3}\right)^{10-l}\cdot\binom{10}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{10-k}$$

כלומר ב"ת. זה מתקבל מכך שהכדורים הזוגיים והאי זוגיים, אומנם נראים תלויים אבל בעצם לא משפיעים בעת הבחירה.

.2

(א) מ"מ X_1+X_2 בעל התפלגות לא בינומית וגם מ"מ X_1+Y_1 בעל התפלגות לא בינומית.

(ב) מ"מ X_1+X_2 בעל התפלגות בינומית, וגם מ"מ X_1+Y_1 בעל התפלגות בינומית.

(ג) מ"מ X_1+X_2 בעל התפלגות בינומית, אך מ"מ X_1+Y_1 בעל התפלגות לא בינומית.

פתרון: ראינו ש- $X_1, Y_1 \sim \text{Bin}\left(10, \frac{1}{3}\right)$ ובלתי תלויים, לכן גם סכומם מתפלג בינומי (בפרט $X_1+X_2 \sim \text{Bin}\left(20, \frac{1}{3}\right)$). X_1+X_2 הוא משתנה מקרי הסופר את הכדורים הזוגיים בשני הכדים הראשונים, בפרט

$$.X_1+X_2 \sim \text{Bin}\left(10, \frac{2}{3}\right)$$

על כן (ב) היא התשובה הנכונה.

3. ההתפלגות המותנית של X_1 בהנתן $X_1+Y_1=9$ היא

(א) אחידה.

(ב) היפרגיאומטרית

(ג) בינומית.

פתרון: תחת התניה זו, אנו יודעים את מספר סך הכדורים בכד 1, ונרצה לדעת כמה מהם זוגיים, או "מיוחדים". כלומר

$$X \mid X_1+Y_1=9 \sim \text{HG}\left(20, 10, 9\right)$$

(20 כדורים, 10 מיוחדים, 9 הוצאות). התשובה הנכונה היא (ב).

4. ההתפלגות המותנית של X_1 בהנתן $X_1+X_2=9$ היא

(א) אחידה.

(ב) בינומית.

(ג) היפרגיאומטרית

פתרון: כאן יש לנו בחירה של 9 מתוך סך הכדורים ואנו רוצים לדעת כמה מהם בכד הראשון, כלומר

$$X_1 \mid X_1+X_2=9 \sim \text{Bin}\left(9, \frac{1}{2}\right)$$

(ההסתברות להיות בראשון או בשני שווה) - התשובה הנכונה היא (ב).

5. ההתפלגות המותנית של X_1 בהנתן $Y_2=5$ היא

(א) בינומית.

(ב) היפרגיאומטרית.

(ג) אחידה.

פתרון: ראינו כבר ש־ X_1 ב"ת ב־ Y_1, מאותה סיבה הוא ב"ת ב־ Y_2, לכן (א) היא התשובה הנכונה.

תרגיל 8.3 יהי מ"מ $Y \sim G(p)$ כאשר $0 < p < 1$. נגדיר סדרת משתנים מקריים $(X_i)_{i\geq 1}$ ב"ת בינם לבין עצמם וב־Y ובעלי אותה התפלגות כמו Y. מצאו כיצד מתפלג $S = \sum_{i=1}^{Y} X_i$.

פתרון: נשים לב שהתומך של S הוא כל המספרים הטבעיים ונפתור בעזרת התנייה על Y. עבור $n \in \mathbb{N}$,

$$\begin{aligned}
\Pr(S=n) &= \sum_{k=1}^{n} \Pr(S=n|Y=k)\Pr(Y=k) = \\
&= \sum_{k=1}^{n} \Pr\left(\sum_{i=1}^{k} X_i = n|Y=k\right)(1-p)^{k-1}p = \\
&= \sum_{k=1}^{n} \Pr\left(\sum_{i=1}^{k} X_i = n\right)(1-p)^{k-1}p
\end{aligned}$$

כאשר המעבר האחרון נובע מאי־תלות של $(X_i)_{i\geq 1}$ ב־Y. אנו יודעים שסכום של מ"מ גיאומטריים ב"ת בעלי אותו פרמטר מתפלגים כמו משתנה מקרי בינומי־שלילי ולכן

$$\begin{aligned}
\Pr(S=n) &= \sum_{k=1}^{n}\left[\left(\binom{n-1}{k-1}(1-p)^{n-k}p^k\right)(1-p)^{k-1}p\right] = \\
&= \sum_{k=1}^{n}\binom{n-1}{k-1}p^{k+1}(1-p)^{n-1} = \\
&= p^2(1-p)^{n-1}\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1} = \\
&= p^2(1-p)^{n-1}\sum_{k=0}^{n-1}\binom{n-1}{k}p^{k} = \\
&= p^2(1-p)^{n-1}\sum_{k=0}^{n-1}\binom{n-1}{k}p^{k}1^{n-1-k} = \\
&= p^2(1-p)^{n-1}(1+p)^{n-1} = \\
&= p^2((1-p)(1+p))^{n-1} = \\
&= p^2\left(1-p^2\right)^{n-1}.
\end{aligned}$$

לכן קיבלנו ש־S מתפלג גיאומטרית עם פרמטר p^2.