יסודות תורת ההסתברות

אוהד נ. פלדהיים
אורי גוראל גורביץ'

מכון איינשטיין למתמטיקה
האוניברסיטה העברית בירושלים

3 בנובמבר 2018

מטרת הפרק הזה היא לנסות לענות על שאלות הסתברותיות
מעניינות ממרחב ההסתברות בלי שיהיה צורך לבנות מחדש כל פעם מרחב לשאלה
כזו או אחרת

פרק 4

משתנים מקריים בדידים

בזמן שכתבתי את הספר [תהליכים סטוכסטים 1953] היה לי ויכוח עם פֶלֶר. הוא אמר שכולם אומרים **משתנה מקרי** (random variable) ואילו אני אמרתי שכולם אומרים **משתנה אקראי** (chance variable). ברור היה שמוטב שנשתמש באותו שם בשני הספרים שלנו. החלטנו להכריע באופן מקרי. הטלנו מטבע והוא זכה.

– ג'וֹזף דֻּוּבּ בשיחה עם לָאוּרִי סְנֶל, 1997.

בפרקים הקודמים הנחנו את יסודותיה של תורת ההסתברות על ידי כך שהגדרנו את מרחב ההסתברות וניסחנו את הכללים לחישוב הסתברותם של המאורעות שבו. אומנם כלים אלו אפשרו לנו להשיב על שאלות קלאסיות רבות ולאפיין תופעות הסתברותיות, אולם הפתרונות היו לעיתים כרוכים בסרבול רב. אחד מהגורמים לסרבול זה היה הצורך להכריע בין הגדרת מרחב הסתברות מורכב מאוד שיאפשר לנו להתמודד עם מגוון רחב של שאלות לבין הגדרת מרחב הסתברות מצומצם ופשוט יותר, המתאים רק לשאלה מסויימת. בפרק זה נציג מונח חדש - **משתנה מקרי**, אשר יאפשר לנו הן להתמודד באופן טוב יותר עם שאלות הסתברותיות כמותיות והן לחקור מגוון שאלות על אותו מרחב הסתברות. חידוש זה, הפעוט כביכול, יסלול את הדרך למונחים יסודיים המופיעים לעיתים גם בשימושים יום-יומיים בסטטיסטיקה כמו **תוחלת**, **שונות** ו**סטיית תקן**.

אנלוג לאיקין

4.1 משתנה מקרי בדיד

הגדרה 4.1. יהי Ω מרחב מדגם בדיד. **משתנה מקרי (מ"מ) (discrete random variable)** על מרחב זה הוא פונקציה $X : \Omega \to \mathbb{R}$. נאמר כי X מוגדר על כל מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ שזהו מרחב המדגם שלו.

דוגמא 4.2 (דוגמא למשתנים מקריים על מרחבי הסתברות מוכרים).

תיאור המ"מ	מרחב מדגם	המ"מ
סכום שתי קוביות	$[6]^2$	$X((a_1, a_2)) = a_1 + a_2$
מספר העצים בהטלת שלושה מטבעות	$[H, T]^3$	$X((a_1, a_2, a_3)) = \|\{i \in [3] \ : \ a_i = H\}\|$
מספר נקודות השבת בתמורה מקרית	S_n	$X(\sigma) = \|\{i \in [N] \ : \ \sigma(i) = i\}\|$

אנלוגי לפונק' הסתברות נק'

הגדרה 4.3 (פונקציית התפלגות נקודתית של משתנה מקרי בדיד)**.** נגדיר את **פונקציית ההתפלגות הנקודתית (probability mass function)**, של משתנה מקרי X על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ להיות הפונקציה $p_X : \mathbb{R} \to [0,1]$ הנתונה על ידי $p_X(s) = \mathbb{P}\left(X^{-1}(s)\right)$. כאשר $X^{-1}(s) := \{\omega \in \Omega \ : \ X(\omega) = s\}$.
נכנה את הקבוצה $\mathrm{Supp}(X) = \{x \in \mathbb{R} \ : \ p_X(x) > 0\}$ בשם **התומך** של X.

[illegible] עמוד הבא

כמו במקרה של מרחבי הסתברות - נרחיב את ההגדרה הנקודתית להגדרה עבור מאורעות.

הגדרה 4.4 (התפלגות מ"מ)**.** יהי X מ"מ על מרחב הסתברות בדיד $(\Omega, \mathcal{F}, \mathbb{P})$. תהי $A \subset \mathbb{R}$. נגדיר את $\mathbb{P}_X$, **התפלגותו של** X, להיות פונקציה $\mathbb{P}_X : 2^{\mathbb{R}} \to [0,1]$ הנתונה על ידי,

אנלוגי לפונק' הסתברות

$$\mathbb{P}_X(A) := \mathbb{P}(X^{-1}(A)) = \sum_{\omega : X(\omega) \in A} \mathbb{P}(\{\omega\}).$$

כמו כן נסמן $\mathbb{P}(X \in A) = \mathbb{P}_X(A)$ ונאמר כי זו היא ה**הסתברות ש-**X **נמצא ב-**A.

כך נפרש עבור $A, B \subset \mathbb{R}$ את הכתיב $\mathbb{P}(X \in A \mid X \in B)$, לאור הגדרה 4.24, כ-$\mathbb{P}(X \in A \cap B)/\mathbb{P}(X \in B)$. כאשר נעסוק במאורעות הנוגעים למשתנים מקריים, נשמיט לעיתים קרובות את סימון הקבוצה. כך נכתוב למשל $\mathbb{P}(X = 5) := \mathbb{P}_X(\{5\}) = \mathbb{P}(X \in \{5\})$. לרוב נמנע משימוש בפונקציית ההתפלגות הנקודתית תוך שימוש בכך ש-$\mathbb{P}(X = x) = p_X(x)$.

באופן זה משתנה מקרי הוא **דחיפה קדימה** (forward push) של פונקציית ההסתברות שבחרנו אל $\mathbb{R}$. אלא שלמרות שמרחב ההסתברות ממנו התחלנו עשוי להיות בדיד, הרי ש-$\mathbb{R}$ אינו בדיד. לפיכך נאמר כי "X משתנה מקרי בדיד", כאשר מדובר במשתנה מקרי שניתן להגדירו על מרחב הסתברות בדיד כלשהו. בכדי לדבר על המושג המקביל לפונקציית הסתברות ב-$\mathbb{R}$ - נגדיר את מונח ה**התפלגות**.

הגדרה 4.5 (התפלגות בדידה)**.** פונקציה $\mathbb{P} : \mathbb{R} \to [0,1]$ נקראת **התפלגות בדידה** אם

(א) קיימת קבוצה בת-מניה $A \subset \mathbb{R}$ כך ש-$\mathbb{P}(A) = 1$.

(ב) לכל אוסף קבוצות זרות $\{A_i\}_{i \in \mathbb{N}}$ של מספרים ממשיים מתקיים $\mathbb{P}(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} \mathbb{P}(A_i)$

כמו כן נגדיר את **התומך** של $\mathbb{P}$ להיות: $\mathrm{Supp}(\mathbb{P}) := \{x \in \mathbb{R} \ : \ \mathbb{P}(x) > 0\}$. משתנה מקרי בעל התפלגות בדידה נקרא **בדיד**.

✎ **בעיה 4.1.** נתונה התפלגות בדידה $\mathbb{P}$. על איזה מרחב הסתברות נוכל להגדיר מ"מ X כך שיתקיים $\mathbb{P}_X = \mathbb{P}$?

אם לשני משתנים מקריים שונים X ו-Y (שעשויים להיותר מוגדרים בשני מרחבי הסתברות שונים) יש את אותה התפלגות, נאמר כי הם **שווי התפלגות** ונכתוב $Y \overset{\mathrm{d}}{=} X$. בכדי לסמן את העובדה ש**למשתנה** X **התפלגות** $\mathcal{D}$ נסמן $X \sim \mathcal{D}$.

[illegible] עמוד הבא

דוגמה 4.6 (סכום הטלת שתי קוביות כמשתנה מקרי)**.**
שחקן מטיל שתי קוביות הוגנות. נסמן ב-X את המ"מ המתאר את סכום התוצאות. בנה מרחב הסתברות מתאים, הגדר את X וחשב את פונקציית ההתפלגות הנקודתית שלו, ואת ההסתברות ש-X בין 4 ל-7.

תשובה: מרחב ההסתברות המתאים לשאלה הוא המרחב האחיד על $[6]^2$. נגדיר $X((a,b)) = a + b$.

אז מה ההבדל בין פונק' התפלגות נק' לבין פונק' ההתפלגות:

תחילה סימונים פונק' התפלגות נק' תסומן (p קטן)
פונק' התפלגות תסומן $\mathbb{P}$.

נתחיל מפונק' ההתפלגות נק': שמוגדרת כך: $P_X:\mathbb{R}\longrightarrow[0,1]$ ומוגדרת $P_X(X^{-1}(s))$ כאשר
$X^{-1}(s)=\{w\in\Omega \mid X(w)=s\}$ באופן אינטואיטיבי נוכל לרשום $P_X(s)=P(\{w\in\Omega \mid X(w)=s\})$

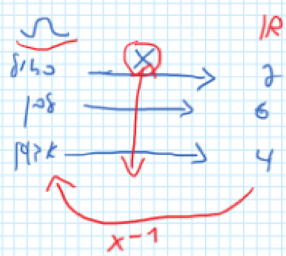
אז מה כתוב כאן: לכל יחידון ששייך ל X יש מספר שמתאים לו בין 0 ל-1. לכן נבין בכל היחידונים שהסיכוי שהם יקרו הוא s ולכל יחידון כזה ניתן ערך בין 0 ל-1

פונק' התפלגות: פונקציה זו מתפלגת כך נתבונן ב $X\subset A\subset\mathbb{R}$. כלומר נביט בקבוצה A שבתוכה X מוכל בה
כעת $\mathbb{P}_X: 2^{\mathbb{R}}\longrightarrow[0,1]$ כלומר פונק' ההתפלגות אוכלת קבוצה מ$2^{\mathbb{R}}$ ואוכלת בהכרח קבוצה שמקיימת

$$\mathbb{P}_X(A)=P(X^{-1}(A))=\sum_{w\in X(w)\in A}\overbrace{\mathbb{P}(\{w\})}^{P_X(w)}$$

פונק' התפלגות נק'

כלומר הקבוצה A זה מאורע שזו קבוצה שגודלה ≤ 1. וקבוצה זו מקבלת ערך בין 0 ל 1. כעת נבין איך
$X^{-1}(A)$ פועלת במקרה הבדידותי בייחוד לפי הגדרה ברורה כאן ניקח את האיברים של A ולכל איבר נפעיל
את $X^{-1}(a)$ דוגמא: אם $X=\{2,6,4\}$ ומתקיים



אז $X^{-1}(A)$ תחזיר לי כחול ואדום שהם יחידונים ב Ω כעת לכל יחידון את מסתכל על הנוסחה
הבאה $\sum_{w\in X(w)\in A}\mathbb{P}(\{w\})$ אני עובר על כל מה ש $X^{-1}(A)$ מחזיר לי ואני סוכם את מה שהפונק'
ההתפלגות הנק' החזירה לכל יחידון וכאן בא לידי ביטוי הקשר בין פונק'
ההתפלגות הנק' לבין פונק' ההתפלגות

$P_X(w)$

נשים לב שאנו דורשים בנוסף מ $\mathbb{P}$ 2 התכונות הבאות: (1) $\mathbb{P}(A)=1$

(2) $\mathbb{P}\left(\bigcup_{i\in N}A_i\right)=\sum_{i\in N}\mathbb{P}(A_i)$

שימו לב לכל ההבדלים הקטנים! $X = y : R \longrightarrow \Omega$

1) אם $X = y$ זה אומר שהפונק' שוות ממש כלומר איבר במקור יפלטו את אותו איבר בטווח

2) אם $X \overset{d}{=} y$ זה אומר ש $P_y = P_x$ [ל"מאורעות" עם אותה הסתברות] כלומר באופן נורמלי x שווה להתפלגות של y

בונוס להמחשת ההבדלים: בין $X = y$ לבין $X \overset{d}{=} y$

ברור כי $X \neq y$ כהגדרה הפונק' מ $\Omega \to R$ אבל הן פועלות בדיוק את אותן הסתברויות למאורעות ב Ω

Ω	P	$X(w)$	$y(w)$
0	$1/3$	w	$w \bmod 3$
1	$1/3$		
2	$1/3$		
3	0		

3) מהזוג' הנ"ל נוצר הגדרה 3: $X \overset{a.s}{=} y$ פונק' הסתברות [התפלגות] כמעט שוות למעט המקומות בהם יש מאורע/משתנה מקרי בעל הסתברות -0!

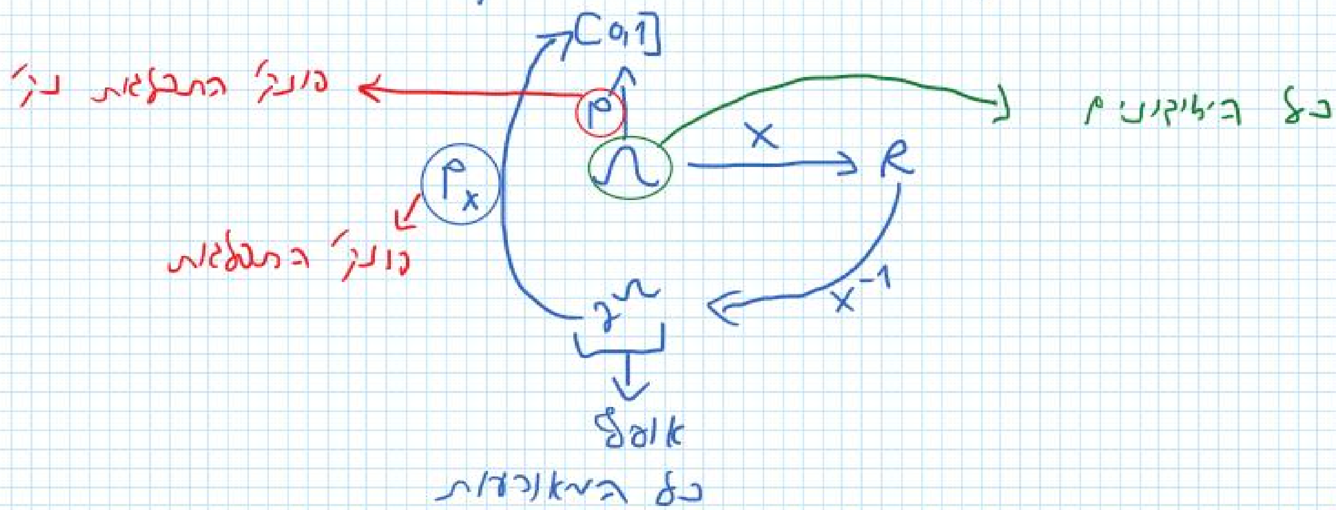
מהגדרה זו נסיק כי x, y יהיו רגישים לשינויי הסתברות משתנים עצמם ע"פ ההגדרות באמצע

נסמן באופן פורמלי

$$X \overset{a.s}{=} y = P(X = y) = 1 = P\left(\{w : X(w) = y(w)\}\right) = 1 \leftarrow$$ נכון בכל המאורעות למעט

$$P\left(w : X(w) \neq X(w)\right) = 0 \longleftarrow$$ מקבלים ערכים שונים במאורעות בעלי הסתברות 0!

יצרנו כאן מערכת שמתנהגת כך:



נדגים את חישוב

$$p_x(4) = \mathbb{P}(\{\omega \in \Omega \ : \ X(\omega) = 4\}) = \mathbb{P}(\{(1,3),(2,2),(3,1)\}) = \frac{3}{36}$$

באופן דומה נקבל

x	2	3	4	5	6	7	8	9	10	11	12
$p_X(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

כעת נחשב $\mathbb{P}(X \in \{4,5,6,7\}) = \sum_{n=4}^{7} p_X(n) = \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} = \frac{1}{2}$.

טענה 4.7 (התפלגות נקודתית קובעת את פונקציית ההתפלגות). יהי X מ"מ בדיד, אזי לכל $A \in \mathbb{R}$,

$$\mathbb{P}(X \in A) = \sum_{x \in \mathrm{Supp}(X) \cap A} \mathbb{P}(X = x).$$

דווקא שאני סוכמים רק
בת מנייה

טענה 4.8 (שוויון התפלגויות נשמר תחת הפעלת פונקציה). יהיו X, Y משתנים מקריים שווי התפלגות (לאו דווקא על אותו מרחב הסתברות), ותהי $f : \mathbb{R} \to \mathbb{R}$ אזי $f(X) \overset{\mathrm{d}}{=} f(Y)$.

הוכחה. נבטא את פונקציות ההתפלגות הנקודתיות $p_{f(X)}, p_{f(Y)}$ מבלי להזדקק למרחב ההסתברות.

$$p_{f(X)}(x) = \mathbb{P}(X \in f^{-1}(\{x\})) = \mathbb{P}(Y \in f^{-1}(\{x\})) = p_{f(Y)}(x),$$

כאשר השוויון האמצעי נובע מכך ש-$Y \overset{\mathrm{d}}{=} X$. ■

טענה 4.8 מהווה צעד ראשון בתהליך של השתחררות ממרחבי הסתברות ומעבר לתיאור בשפה של התפלגויות. לראשונה ראינו ששאלות מסויימות שאנו מתעניינים בהן אינן תלויות כלל במרחב ההסתברות ורק התפלגותו של המשתנה המקרי משפיעה עליהן. עד עתה כאשר רצינו לתאר ניסוי מסויים תארנו אותו באמצעות מרחב הסתברות. כך עשינו למשל עבור ניסוי ברנולי בדוגמא 2.7. בעוד שהגדרה זו הייתה נוחה לצורך תיאור של ניסוי הטלת מטבע יחיד, הרי שלעיתים קרובות נרצה לתאר ניסוי מורכב יותר, ולטעון שמשתנה מקרי מסויים שמוגדר עליו למעשה מתנהג כמו הטלת מטבע. לצורך כך נגדיר התפלגויות שתתארנה משתנים מקריים המתאימים לתוצאות ניסויים נפוצים שיעניינו אותנו. נפתח בפשוטה ביותר בין אלה – **התפלגות ברנולי**.

הגדרה 4.9 (התפלגות ברנולי).נאמר שמ"מ X מתפלג לפי **התפלגות ברנולי עם סיכוי הצלחה** p (בקיצור - ברנולי p) ונכתוב $X \sim \mathrm{Ber}(p)$ אם $p_X(1) = p$ ו-$p_X(0) = 1 - p$.

לעיתים נרצה לקשור בין משתנה ברנולי למאורע במרחב ההסתברות. לשם כך נציג את ההגדרה הבאה.

הגדרה 4.10 (משתנה מציין).יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהי מאורע $A \in \mathcal{F}$. נכנה בשם **משתנה מציין (אינדיקטור) של** A את המשתנה המקרי $X = \mathbb{1}(A)$.

נשים לב שלמשתנה מציין יש התפלגות ברנולי, וכל משתנה ברנולי X הוא משתנה מציין של הקבוצה $X^{-1}(1)$.

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

דוגמא 4.11 (מציאת משתנה בעל התפלגות מסויימת במרחב הסתברות)**.**
עליסה והמלכה משחקות במשחק הבא. בכד אטום שני כדורים לבנים ושלושה כדורים אדומים. ארבעה כדורים נשלפים בזה אחר זה. אם יצאו יותר כדורים אדומים - המלכה מנצחת ואחרת עליסה מנצחת. נגדיר משתנה מציין למאורע "המלכה ניצחה". הראה שזהו משתנה ברנולי $2/5$.

תשובה: נשתמש במרחב ההסתברות על תמורות בגודל 5 כאשר נתייחס למספרים 1 ו-2 כמייצגים כדורים לבנים וליתר ככדורים אדומים. נשים לב שהמאורע המבוקש ניתן לתיאור באופן הבא:

$$A = \{S \in \Omega \ : \ S_5 \in [2]\}$$

ולפי הגדרה

$$\mathbb{P}(X) = \mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{2 \cdot 4!}{5!} = \frac{2}{5}.$$

כפי שתיארנו באמצעות התפלגות ברנולי את מספר העצים בניסוי הטלת המטבע, כך נגדיר **התפלגות אחידה**, אשר תתאים לניסויים מספריים המתוארים על ידי מרחב הסתברות אחיד, כמו למשל תוצאת הטלת קוביה.

הגדרה 4.12 (התפלגות אחידה)**.** נאמר שמ"מ X מתפלג לפי **התפלגות אחידה על קבוצה סופית** $A \subset \mathbb{R}$ ונכתוב $X \sim \mathrm{Unif}(A)$ אם $\forall i \in A : p_X(i) = \frac{1}{|A|}$.

על פי רוב נתעניין במשתנים בדידים בעלי התפלגות אחידה על $[N]$ ולכן נקצר $\mathrm{Unif}(N) := \mathrm{Unif}([N])$.

✎ **בעיה 4.2** (קוביות סיקרמן Sicherman dice)**.** נתונות זוג קוביות הוגנות בעלות שש פאות. על פאות קוביה אחת כתובים המספרים $\{1, 2, 2, 3, 3, 4\}$ ועל פאות השניה $\{1, 3, 4, 5, 6, 8\}$. יש להראות שהתפלגות סכום שתי הקוביות שווה להתפלגות סכומן של שתי קוביות משחק רגילות.

4.2 יחסים בין מספר משתנים מקריים

עד כה דנו במשתנה מקרי יחיד המוגדר על מרחב הסתברות, אך לעיתים קרובות נתעניין דווקא ביחסים בין מספר משתנים מקריים המוגדרים על אותו מרחב. נפתח בדוגמא.

דוגמא 4.13. מוטלות שתי קוביות משחק הוגנות. יש לתאר באמצעות משתנים מקריים את תוצאת כל אחת מהקוביות ואת סכומן. מה ההסתברות שסכום הקוביות יצא בין שש לשמונה?

תשובה: מרחב ההסתברות הוא המרחב האחיד על $[6]^2$. נגדיר משתנים מקריים X, תוצאת הקוביה הראשונה, Y, תוצאת הקוביה השניה ו-Z סכום הקוביות. כלומר

$$X((a,b)) = a \qquad Y((a,b)) = b \qquad Z = X + Y.$$

נשים לב כי ל-X ול-Y התפלגות אחידה. כעת נחשב לפי נוסחת ההסתברות השלמה

$$\mathbb{P}(Z \in \{6,7,8\}) = \sum_{i=1}^{6} \mathbb{P}(X = i)\mathbb{P}(Z \in \{6,7,8\} \,|\, X = i) = \frac{1}{6}\left(\frac{2}{6} + \frac{3}{6} + \frac{3}{6} + \frac{3}{6} + \frac{3}{6} + \frac{2}{6}\right) = \frac{4}{9}.$$

בדוגמא 4.13 נתקלנו לראשונה בשתי תופעות מעניינות: מצב שבו שני משתנים מקריים מוגדרים על אותו מרחב הסתברות, והפעלה של אופרטור החיבור על משתנים מקריים. המידע הנוגע ליחס בין שני משתנים מקריים על אותו מרחב הסתברות אינו מצוי בהתפלגותם. כיוון שאנחנו מעוניינים להימנע מתיאור מפורש של מרחבי הסתברות, נייצג מידע זה באמצעות מונח חדש.

הגדרה 4.14 (התפלגות משותפת)**.** יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות בדיד ויהיו $\{X_i\}_{i\in[N]}$ משתנים מקריים עליו. **ההתפלגות המשותפת** של $\{X_i\}_{i\in[N]}$ היא פונקציה $\mathbb{P}_{X_1,\dots,X_N}$ מתת-קבוצות של $\mathbb{R}^N$ ל-$[0,1]$ המוגדרת לכל A על ידי

$$\mathbb{P}_{X_1,\dots,X_N}(A) := \mathbb{P}\Big((X_1,\dots,X_N)\in A\Big) = \mathbb{P}\left(\{\omega\in\Omega : (X_1(\omega),\dots,X_n(\omega))\in A\}\right)$$

בכדי לתאר התפלגות משותפת של מספר משתנים מקריים בדידים די לתאר את פונקציית ההתפלגות הנקודתית שלהם המוגדרת כדלהלן.

הגדרה 4.15 (פונקציית התפלגות משותפת נקודתית)**.** יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות בדיד ויהיו $\{X_i\}_{i\in[N]}$ משתנים מקריים עליו. ***פונקציית ההתפלגות המשותפת הנקודתית*** של $\{X_i\}_{i\in[N]}$ מוגדרת להיות הפונקציה $p_X : \mathbb{R}^N \to [0,1]$ שנתונה על ידי

$$p_{X_1,\dots X_N}(x_1,\dots,x_N) := \mathbb{P}(X_1 = x_1,\dots,X_N = x_N) = \mathbb{P}\left(\bigcap_{i=1}^{N} X_i^{-1}(x_i)\right).$$

ואכן, את פונקציית ההתפלגות המשותפת אפשר לבטא בתור

$$\mathbb{P}_{X_1,\dots,X_N}(A) = \mathbb{P}(X_1,\dots,X_N\in A) = \sum_{\substack{\forall i\in[N]: x_i\in \mathrm{Supp}(X_i) \\ (x_1,\dots,x_n)\in A}} p_{X_1,\dots,X_N}((x_1,\dots,x_N)).$$

למעשה ניתן לראות את ההתפלגות המשותפת כהכללה של מושג המשתנה המקרי למרחב $\mathbb{R}^d$. כאשר נתונה התפלגות משותפת נכנה בשם **התפלגות שולית** את התפלגותם בנפרד של כל אחד מהמשתנים. מעתה בכל פעם שנדבר על מספר משתנים מקריים נניח כי כולם מוגדרים על אותו מרחב הסתברות.

דוגמא 4.16. מוטלות שתי קוביות משחק סטנדרטיות. נסמן ב-X_1, X_2 את תוצאת כל אחת מהקוביות. כמו כן נסמן ב-Y את התוצאה הנמוכה מבין השתיים וב-Z את התוצאה הגבוהה מביניהן. יש לחשב את התפלגותם המשותפת של Y ו-Z ואת ההתפלגויות השוליות.

תשובה: מרחב ההסתברות המתאים לבעיה הוא המרחב האחיד על $\Omega = [6]^2$. נחשב את התפלגותם של X_1 ו-X_2 על ידי מעבר על כל האפשרויות. בכל משבצת יצויין "ערך גבוה-ערך נמוך".

Min - Max

X_2 \ X_1	1	2	3	4	5	6
1	1-1	1-2	1-3	1-4	1-5	1-6
2	1-2	2-2	2-3	2-4	2-5	2-6
3	1-3	2-3	3-3	3-4	3-5	3-6
4	1-4	2-4	3-4	4-4	4-5	4-6
5	1-5	2-5	3-5	4-5	5-5	5-6
6	1-6	2-6	3-6	4-6	5-6	6-6

לכל אחד מהתאים הסתברות של $1/36$. נבחן ונגלה שהתפלגותם המשותפת של Y ו-Z הנה כמתואר להלן

מקרה יחיד
max = min = 1
לכן
$\frac{1}{36}$

(1-6)
(6-1)
אלו המקרים
היחידים לכן
$\frac{2}{36}$
מסומן בצהוב
כלפי למעלה

Min Y \ Max Z	1	2	3	4	5	6	Y שולי
1	1/36	2/36	2/36	2/36	2/36	2/36	11/36
2	0	1/36	2/36	2/36	2/36	2/36	9/36
3	0	0	1/36	2/36	2/36	2/36	7/36
4	0	0	0	1/36	2/36	2/36	5/36
5	0	0	0	0	1/36	2/36	3/36
6	0	0	0	0	0	1/36	1/36
Z שולי	1/36	3/36	5/36	7/36	9/36	11/36	1

אין סיכוי
שmin>max
לכן
הסתברות 0

שולי כי הוא
לא תלוי
בהתפלגות
של Z

בטבלה מתוארות גם התפלגויותיהם השוליות של Y ו-Z. את התפלגותו השולית של Y למשל, אנו מחשבים לפי נוסחאת ההסתברות השלמה תוך סכימת הסתברותן של כל הקונפיגורציות המתאימות לערך Y מסויים (כלומר סכימת השורה המתאימה בטבלה).

נכליל כעת את טענה 4.8.

אבחנה 4.17 (שוויון התפלגויות משותפות נשמר תחת הפעלת פונקציה)**.** יהיו $X_1, \ldots, X_N$ ו-$Y_1, \ldots, Y_N$ שני אוספי משתנים מקריים בעלי אותה התפלגות משותפת ותהיינה $f_1, \ldots, f_M : \mathbb{R}^N \to \mathbb{R}$ פונקציות כלשהן. אזי $(f_1(X_1, \ldots, X_N), \ldots, f_M(X_1, \ldots, X_N))$ ו-$(f_1(Y_1, \ldots, Y_N), \ldots, f_M(Y_1, \ldots, Y_N))$ בעלות אותה התפלגות משותפת ובפרט $f_1(X_1, \ldots, X_N) \overset{\mathrm{d}}{=} f_1(Y_1, \ldots, Y_N)$.

✎ **בעיה 4.3.** להוכיח את אבחנה 4.17 בדומה לטענה 4.8.

מאבחנה 4.17 אנו למדים שהתפלגותו של סכום משתנים מקריים תלויה רק בהתפלגותם המשותפת.

4.2.1 אי-תלות בין משתנים מקריים

עד עתה שימש אותנו מושג האי-תלות בשתי צורות. האחת - בצורת אי-תלות בין מאורעות המתארת את העובדה שהידיעה האם התרחש אחד המאורעות אינה משליכה על ההסתברות שהמאורע השני התרחש. השניה - בצורת מרחבי מכפלה - מרחבים בהם יש חוסר תלות בין כל מאורע. כעת נכיר צורה שלישית - **אי-תלות בין משתנים מקריים**. בדומה לאי-התלות שבמרחבי מכפלה, אי-תלות זו תבטא מצב שבו לתוצאה של ניסוי אחד אין כל השפעה על התפלגות תוצאתו של ניסוי אחר והיא תתגלם בקשר בין התפלגותם של משתנים מקריים להתפלגותם המשותפת.

הגדרה 4.18 (אי-תלות של שני מ"מ בדידים)**.** נאמר שמ"מ X ו-Y בדידים (על אותו מרחב הסתברות) הנם **בלתי תלויים (ב"ת)** ונסמן $X \perp Y$, אם לכל שתי קבוצות $A, B \subset \mathbb{R}$ מתקיים שהמאורעות $\{X \in A\}$ ו-$\{Y \in B\}$ בלתי תלויים, כלומר

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

באופן דומה נדבר על חוסר תלות של קבוצה של מ"מ מקריים.

הגדרה 4.19 (אי-תלות של מ"מ בדידים). יהי $\mathcal{X}$ אוסף של מ"מ בדידים (על אותו מרחב הסתברות). נאמר שהמ"מ באוסף בלתי-תלויים אם לכל $\{X_i\}_{i\in[N]}$, תת קבוצה סופית של משתנים מקריים ב-$\mathcal{X}$ ולכל $\{A_i\}_{i\in[N]}$, אוסף קבוצות ב-$\mathbb{R}$, מתקיים:

$$\mathbb{P}(\forall i \in [N]\ X_i \in A_i) = \prod_{i\in[N]} \mathbb{P}(X_i \in A_i).$$

✎ **בעיה 4.4.** הוכח כי משתנה מקרי הוא ב"ת בעצמו אם ורק אם הוא קבוע (כלומר מקבל ערך מסויים בהסתברות 1).

אבחנה 4.20 (אי-תלות של משתנים מקריים בדידים היא תכונה של פונקציות התפלגות נקודתיות). יהיו $X_1,\dots,X_n$ מ"מ ב"ת במרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. המ"מ ב"ת אם ורק אם לכל $x_1,\dots,x_n\in\mathbb{R}$ מתקיים

$$\mathbb{P}(X_1=x_1,\dots,X_n=x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i=x_i).$$

✎ **בעיה 4.5.** להוכיח את אבחנה 4.20.

✎ **בעיה 4.6.** יהיו X ו-Y מ"מ בלתי-תלויים. יש להוכיח כי $|X|,|Y|$ בלתי תלויים.

✎ **בעיה 4.7.** להראות כי Y ו-Z מדוגמא 4.16 הנם תלויים אך $|Y-3.5|$ ו-$|Z-3.5|$ הנם בלתי-תלויים.

✎ **בעיה 4.8.** להראות כי אם $X\sim \mathrm{Ber}(p)$ ו-$Y\sim\mathrm{Ber}(q)$ הנם ב"ת אז XY מתפלג ברנולי ולחשב את סיכוי ההצלחה.

דוגמא 4.21 (בעיית האספן [coupon collector's problem]).
בכל ביצת הפתעה ניתן למצוא אחת מבין n סוגי בובות אהובות. הבובות מוגרלות באופן אקראי ובלתי-תלוי בכל ביצה. כיצד נחסום מלמעלה את ההסתברות שיהיה צורך ברכישת יותר מ-k ביצים בכדי לאסוף את כל סוגי הבובות.

תשובה: מטרתנו היא לחסום את הסתברותו של המאורע שבין k הביצים הראשונות שקנינו לא הופיעו כל סוגי הבובות. בלי הגבלת הכלליות נסמן את סוגי הבובות האפשריים במספרים מ-1 ועד n. יהי X_i סוג הבובה בביצה ה-i. נסמן ב-A_j את המאורע שבין k הביצים שרכשנו אף אחת לא הכילה בובה מסוג j. המאורע שאנו מעוניינים לחשב את הסתברותו הוא $E=\bigcup_{j\in[n]}A_j$.

ההסתברות שהביצה ה-i אינה מכילה בובה מסוג j הנה $\mathbb{P}(X_i\neq j)=\frac{n-1}{n}$ ולכן, הואיל וסוגי הבובות בביצים שונות בלתי תלויים, הסתברותו של A_j היא $\mathbb{P}(A_j)=\prod_{i\in[k]}\mathbb{P}(X_i\neq j)=\left(\frac{n-1}{n}\right)^k$. נשתמש בחסם האיחוד (משפט 2.17) ונקבל

$$\mathbb{P}(E)\le\sum_{j\in[n]}\mathbb{P}(A_j)=\sum_{j\in[n]}\left(\frac{n-1}{n}\right)^k=n\left(1-\frac{1}{n}\right)^k\le ne^{-k/n}.$$

חסם ברנולי

לפיכך אם נקנה $k=c\cdot n\log n$ ביצי הפתעה ההסתברות של E חסומה על ידי n^{1-c}. אם ניקח $c>1$ כלשהו הרי שההסתברות שנקבל את כל סוגי הבובות תשאף ל-1 כאשר נשאיף את n לאינסוף.

הערה: לעת עתה הצגנו רק חסם עליון למספר ביצי ההפתעה הדרושות באופן טיפוסי בכדי לזכות בכל סוגי הבובות ואילו כחסם תחתון נאלץ להסתפק בחסם המובן מאליו הקובע שדרושות לפחות n ביצים. בפרקים הבאים, נתוודע לכלים חדשים שיאפשרו לנו להראות שהחסם שהצגנו הדוק באופן אסימפטוטי.

בבקשנו להכליל את מושג האי-תלות של זוג משתנים מקריים לאי-תלות קבוצתית – יסתבר לנו שההגדרות וההוכחות מסתבכות והולכות. יתר ההגדרות וההוכחות בפרק זה מציגות את הגישה הישירה להכללה זו. בפרק 4.5, מוצגת גישה חילופית, עמוקה יותר, אשר מפשטת הגדרות אלו במחיר של הפשטה והצגת מושגים חדשים. היכרות עם גישה זו היא הצעד הראשון בדרך לבניין מרחבי הסתברות שאינם בדידים.

הגדרה 4.22 (אי-תלות של שתי קבוצות מ"מ בדידים). יהיו $\mathcal{X}, \mathcal{Y}$ שני אוספים של מ"מ בדידים (על אותו מרחב הסתברות). נאמר ש**האוספים בלתי-תלויים** ונסמן $\mathcal{X} \perp\!\!\!\perp \mathcal{Y}$, אם לכל $\{X_i\}_{i\in[N]}$, תת קבוצה סופית של משתנים מקריים ב-$\mathcal{X}$ ו-$\{Y_i\}_{i\in[M]}$, תת קבוצה סופית של משתנים מקריים ב-$\mathcal{Y}$ ולכל $\{A_i\}_{i\in[N]}, \{B_i\}_{i\in[N]}$ אוספי קבוצות ב-$\mathbb{R}$, מתקיים:

$$\mathbb{P}(\forall i \in [N]\ X_i \in A_i, \forall i \in [M]\ Y_i \in B_i) = \mathbb{P}(\forall i \in [N]\ X_i \in A_i)\mathbb{P}(\forall i \in [M]\ Y_i \in B_i).$$

אי-תלות בין קבוצות משתנים מקריים עוברת בירושה לפונקציות על אוספים אלה.

טענה 4.23 (שימור אי-תלות תחת הפעלת פונקציות). יהיו $X_1,\ldots,X_n$ ו-$Y_1,\ldots,Y_m$ שתי קבוצות בלתי תלויות של מ"מ במרחב הסתברות. תהיינה $f:\mathbb{R}^n\to\mathbb{R}$, $g:\mathbb{R}^m\to\mathbb{R}$ פונקציות ממשיות. אזי המ"מ $f(X_1,\ldots,X_n)$ ו-$g(Y_1,\ldots,Y_m)$ ב"ת.

הוכחה. נסמן $X = f(X_1,\ldots,X_n), Y = f(Y_1,\ldots,Y_m)$. לאור אבחנה 4.20 די להראות כי לכל שני ממשיים $s \in \mathrm{Supp}(X)$ ו-$t \in \mathrm{Supp}(Y)$ מתקיים כי $\mathbb{P}(X=s, Y=t) = \mathbb{P}(X=s)\mathbb{P}(Y=t)$. נחשב,

$$\begin{aligned}
\mathbb{P}(X=s,Y=t) &= \sum_{\substack{\{x_i\}_{i\in[n]} \\ x_i\in\mathrm{Supp}(X_i) \\ f(x_1,\ldots,x_n)=s}} \sum_{\substack{\{y_j\}_{j\in[m]} \\ y_j\in\mathrm{Supp}(Y_j) \\ g(y_1,\ldots,y_m)=t}} \mathbb{P}\Big(X_1=x_1,\ldots,X_n=x_n,Y_1=y_1,\ldots,Y_m=y_m\Big) \\
&\overset{\text{אי-תלות}}{=} \sum_{\substack{\{x_i\}_{i\in[n]} \\ x_i\in\mathrm{Supp}(X_i) \\ f(x_1,\ldots,x_n)=s}} \sum_{\substack{\{y_j\}_{j\in[m]} \\ y_j\in\mathrm{Supp}(Y_i) \\ g(y_1,\ldots,y_m)=t}} \mathbb{P}(\forall i\in[n]:X_i=x_i)\mathbb{P}(\forall i\in[m]:Y_i=y_i) \\
&= \left(\sum_{\substack{\{x_i\}_{i\in[n]} \\ x_i\in\mathrm{Supp}(X_i) \\ f(x_1,\ldots,x_n)=s}} \mathbb{P}(\forall i\in[n]:X_i=x_i)\right)\left(\sum_{\substack{\{y_j\}_{j\in[m]} \\ y_j\in\mathrm{Supp}(Y_j) \\ g(y_1,\ldots,y_m)=t}} \mathbb{P}(\forall i\in[m]:Y_i=y_i)\right) \\
&= \mathbb{P}(X=s)\mathbb{P}(Y=t)
\end{aligned}$$

לוגית על מינוך
התפלגות משותפת

מינוך התפלגות על ?

■

✎ **בעיה 4.9** (קוביות אפרון [Efron's dice]). להלן המספרים על צדדיהן של ארבע קוביות הוגנות, A, B, C, D

כמה הערות חשובות:

כאשר מסתכלים על אוספים של מ"מ זרי[ם]

$x_1, — x_n$ מ"מ ב"ת

$y_1, — y_m$ מ"מ ב"ת

אז שימו לב x בינם לבין עצמם ב"ת

ואותו דבר לגבי y הם ב"ת בינם לבין עצמם כלומר אם נגדיר $z = x_i \cdot x_j$ אני לא יודע האם

z תלוי ב x_i או ב y_j ועוד נג' משובב

נשים לב שלהסתכל על האוספים של הנ"ל ולהגדיר שהם ב"ת על ידי הסתכלות על תתי קב' ושקיימות

$$P(\forall i \in [n]\ x_i \in A_i, \forall i \in [m]\ y_i \in B_i) = P(\forall i \in [n]\ x_i \in A_i)\, P(\forall i \in [m]\ y_i \in B_i)$$

נזכור שהיחידונים של x_i, y_j קרי:

$$P(x_1 = x_1, — x_n = x_n, y_1 = y_1 — y_m = y_m) = \underbrace{P(\forall i \in [n]\ x_i = x_i)}_{\text{פונק' התפלגות נק' עבור יחידון}} \cdot \underline{P(\forall i \in [m]\ y_i = y_i)}$$

זה בעצם אומר את אותו הדבר ועל זה הולכה מתבססת?

הדיון בעניין שימור אי תלות עבור פונק':

אם $x_1, — x_n$ מ"מ ב"ת לא משנה אילו פעולות נבצע אליהם הם ישארו ב"ת מכיון שהיחידונים שלהם ב"ת?

באופן דומה עבור $y_1 — y_m$ אבל שימו לב שאם אקח מ"מ x_i, y_j ואבצע עליהם פעולות אני לא

יודע מה יקרה וזה נקודה קריטית שחשוב להבין

פונק' התפלגות נק'?

המוטלות באופן בלתי-תלוי.

$$A = \{4,4,4,4,0,0\} \qquad B = \{3,3,3,3,3,3\}$$
$$C = \{6,6,2,2,2,2\} \qquad D = \{5,5,5,1,1,1\}$$

יש להשתמש במשתנים מקריים מתאימים ולחשב את ההסתברויות ש:

(א) תוצאת ההטלה של A גדולה מזו של B,
(ב) תוצאת ההטלה של B גדולה מזו של C,
(ג) תוצאת ההטלה של C גדולה מזו של D,
(ד) תוצאת ההטלה של D גדולה מזו של A,

4.3 התפלגות מותנית במאורע

בפרק 3.1 התוודענו למושג ההסתברות המותנית אשר תאר את השפעתו של מידע על הסתברותו של מאורע. כעת, נכליל מונח זה לתאר כיצד מידע מחולל שינוי בהתפלגותו של משתנה מקרי.

הגדרה 4.24 (התפלגות מותנית). יהי X משתנה מקרי על המרחב הבדיד $(\Omega, \mathcal{F}, \mathbb{P})$ ויהי $A \in \mathcal{F}$ מאורע המקיים $\mathbb{P}(A) > 0$. נסמן ב-$(X \mid A)$ את המשתנה המקרי X על מרחב ההסתברות $(\Omega, \mathcal{F}, \mathbb{P}_A)$ כאשר $\mathbb{P}_A$ היא פונקצית ההתפלגות המתקבלת כאשר מתנים ב-A (ראה טענה 3.3).

הערות

הסימון בהגדרה 4.24 מטעה במידה מסויימת, משום ש-$(X \mid A)$ ו-X הם סימונים שונים לאותה פונקציה, והם נבדלים רק כאשר מחושבות הסתברויות עליהם, מפני שהללו תלויות במרחב ההסתברות שנבחר.

דוגמא 4.25. נסמן ב-X וב-Y את תוצאות ההטלה של שתי קוביות משחק הוגנות וב-$Z = X + Y$ את סכומן. מהי התפלגות של Z בהינתן שהוא זוגי? בהינתן ש-Z זוגי איזה מאורע סביר יותר - $\{X = 4\}$ או $\{Z = 4\}$?
תשובה: נסמן ב-A את המאורע ש-Z זוגי, כלומר, $A = \{Z \in \{2,4,6,8,10,12\}\}$. לפי נוסחת ההסתברות השלמה מתקיים:

$$\mathbb{P}(A) = \sum_{i \in [6]} \mathbb{P}(A \mid X = i)\mathbb{P}(X = i) = \sum_{i \in [6]} \frac{1}{2}\mathbb{P}(X = i) = \frac{1}{2}.$$

משיקולי ספירת תצריפים נקבל כי

$$\mathbb{P}(Z = 2) = \mathbb{P}(Z = 12) = \frac{1}{36} \quad \mathbb{P}(Z = 4) = \mathbb{P}(Z = 10) = \frac{3}{36} \quad \mathbb{P}(Z = 6) = \mathbb{P}(Z = 8) = \frac{5}{36}$$

ונחשב את ההתפלגות המותנית של Z בהנתן A לפי הנוסחא

$$\mathbb{P}(Z = i \mid A) = \frac{\mathbb{P}(Z = i \cap A)}{\mathbb{P}(A)} = \frac{P(Z = i)}{\mathbb{P}(A)}$$

לכל i זוגי. נקבל

i	2	4	6	8	10	12
$\mathbb{P}(Z = i \mid A)$	1/18	1/6	5/18	5/18	1/6	1/18

כמה דגשים:

נשים לב כי $P_{X|A}: \mathcal{B} \to \mathbb{R}$ כך ש: $P_{X|A}(B) = P_A(X \in B) = \frac{P(X\in B, A)}{P(A)}$

נשים לב כי $P_{X|A}$, P_X הן אותה פונק' **אבל** עבור מרחבי הסתברות שונים

A – המאורע קרה

Ω – מרחב מדגם כולו

לכן כמובן אין משמעות לשאלה מהי ההתפלגות המשותפת של $P_{X|A,\, X}$

לגבי השאלה שמופיע בהמשך

II \ I	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

סכום זוגי:

$$P(\text{סכום זוגי}) = \frac{18}{36} = \frac{1}{2}$$

וחישבנו את התפלגות $Z \mid A$. כעת נשים לב כי

$$\mathbb{P}(X = i \mid A) = \frac{P(X = i)\mathbb{P}(A \mid X = i)}{\mathbb{P}(A)} = \frac{(1/6) \cdot (1/2)}{1/2} = \frac{1}{6}.$$

כלומר X בלתי תלוי במאורע A. קיבלנו כי בהנתן ש-Z זוגי, ההסתברויות ש-$X = 4$ וש-$Z = 4$ שתיהן שוות ל-$1/6$.

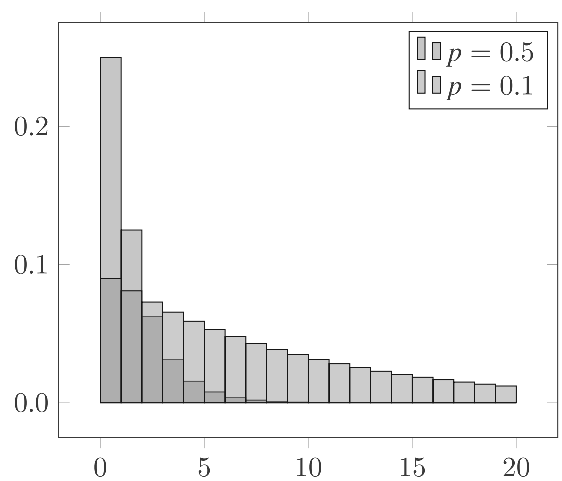
4.4 התפלגויות נפוצות

למרות שהניסויים התיאורתיים שאנו יכולים להעלות במחשבתנו מגוונים, מסתבר שרבים מהם צופנים בחובם משתנים מקריים דומים זה לזה. חמש התפלגויות כאלה תקבלנה את מרבית אור הזרקורים בפרקים הבאים. שתי הראשונות, התפלגות ברנולי, והתפלגות אחידה כבר הוצגו בהגדרות 9.9-14.9 בהתאמה, וכעת בשלו התנאים לערוך היכרות עם שלוש התפלגויות נוספות.

4.4.1 התפלגות גיאומטרית

נסתכל על ניסוי חוזר בעל הסתברות הצלחה p שאינה תלויה בתוצאות הניסויים הקודמים. אנו מבקשים לדעת מה הוא מספר הניסויים עד להצלחה ראשונה. התפלגות משתנה מקרי זה מכונה **התפלגות גיאומטרית** והיא נתונה להלן.

הגדרה 4.26 (התפלגות גיאומטרית).נאמר שמ"מ X מתפלג לפי **התפלגות גיאומטרית עם סיכוי הצלחה p** ונכתוב $X \sim \mathrm{Geo}(p)$ אם לכל $n \in \mathbb{N}$ מתקיים $p_X(n) = (1-p)^{n-1}p$.



התפלגות גיאומטרית עבור ערכים עד 20 לשני ערכי p.

תכונה מעניינת של התפלגות גיאומטרית היא **פונקציית ההתפלגות השיורית** שלה (מושג שנתוודע אליו באופן כללי יותר בפרק הפניה). נגדיר לכל $r \in \mathbb{N}_0$,

$$\bar{F}_X(r) = \mathbb{P}(X > r) = \sum_{n=r+1}^{\infty} (1-p)^{n-1}p = (1-p)^r \sum_{n=1}^{\infty} (1-p)^{n-1}p = (1-p)^r.$$

בפרק הפניה, נראה הצדקה לכך שהתפלגות גיאומטרית אומנם מתארת את התפלגות מספר ההטלות עד להצלחה בהטלות חוזרות של מטבע מוטה.

סגולתה המיוחדת של התפלגות גיאומטרית היא תכונת חוסר הזיכרון לכשלונות כפי שעולה מהטענה הבאה.

טענה 4.27 (אפיון התפלגות גיאומטרית). יהי X משתנה מקרי הנתמך על $\mathbb{N}$. שלושת הבאים שקולים:

(א) $\mathcal{D}$ הנה התפלגות גיאומטרית,

(ב) X ו-$(X-1 \mid X>1)$ שווי התפלגות,

(ג) X ו-$(X-s \mid X>s)$ שווי התפלגות לכל $s \in \mathbb{N}$.

הוכחה. **(א) גורר את (ב).** לכל $n \in \mathbb{N}$ מתקיים

$$\mathbb{P}(X-1=n \mid X>1) = \frac{\mathbb{P}(X=n+1)}{\mathbb{P}(X>1)} = \frac{(1-p)^n p}{1-p} = (1-p)^{n-1}p = \mathbb{P}(X=n).$$

(ב) גורר את (ג). לפי (ב) $(X-1 \mid X>1)$ ו- X שווי התפלגות. נגדיר $Y := (X-(s-1))$, נניח באינדוקציה כי המשתנים $X \stackrel{\mathrm{d}}{=} (Y \mid Y>0)$ שווי התפלגות. נגדיר $Y := (X-(s-1))$ ונחשב:

$$(Y \mid Y>0)) \stackrel{\mathrm{d}}{=} X = (X-1 \mid X>1) \stackrel{\mathrm{d}}{=} (Y-1 \mid Y>0, Y>1) = (Y-1 \mid Y>1) = (X-s \mid X>s).$$

כאשר השוויון השמאלי ביותר נובע מהנחת האינדוקציה ומהגדרת Y, השני מתכונה (ב), השלישי – שוב משימוש בהנחת האינדוקציה, הרביעי מאבחנה 3.4א והאחרון מהגדרת Y.

(ג) גורר את (א). יהי X מ"מ בעל ערכים ב-$\mathbb{N}$ המקיים את תכונה (ג). נרשום לפי כלל השרשרת (טענה 3.4ב) לכל $s \in \mathbb{N}_0$,

$$\mathbb{P}(X>s) = \mathbb{P}(X>s \mid X>s-1)\mathbb{P}(X>s-1 \mid X>s-2)\cdots\mathbb{P}(X>1) \stackrel{(ג)}{=} \mathbb{P}(X>1)^s,$$

נסמן $1-p = \mathbb{P}(X>1)$ וקיבלנו את פונקציית ההתפלגות השיורית של מ"מ גיאומטרי עם פרמטר p. כעת נוכל להשתמש בכך ש $\{X=s\} \uplus \{X>s\} = \{X>s-1\}$ ולחשב

$$\mathbb{P}(X=s) = \mathbb{P}(X>s-1) - \mathbb{P}(X>s) = (1-p)^{s-1} - (1-p)^s = p(1-p)^s,$$

בעקיפין למדנו כי פונקציית ההתפלגות השיורית של התפלגות על השלמים קובעת אותה. ■

הערה: גם את ההוכחה ש-(א) גורר את (ב) בטענה לעיל יכולים היינו לבצע באמצעות פונקציות ההתפלגות השיוריות. לכל $t \in \mathbb{R}$ נרשום $s = \lfloor t+1 \rfloor$:

$$\bar{F}_{X-1 \mid X>1}(t) = \frac{\bar{F}_X(t+1)}{\bar{F}_X(1)} = \frac{(1-p)^{s+1}}{1-p} = (1-p)^s = \bar{F}_X(t),$$

והלא ראינו שפונקציית ההתפלגות השיורית קובעת את ההתפלגות על ידי $\mathbb{P}(X=i) = \bar{F}(i-1) - \bar{F}(i)$.

✎ **בעיה 4.10.** יהיו $X_1 \sim \mathrm{Geo}(p)$ ו-$X_2 \sim \mathrm{Geo}(q)$ ב"ת. יש להראות כי

$$\min(X_1, X_2) \sim \mathrm{Geo}(1-(1-p)(1-q)).$$

מי הראשון שהצליח מסתכלים על 2 הניסויים במקביל
ורואים מי הצליח קודם תרגיל 4 בשיעורי בית

✎ **בעיה 4.11.** יהיו $X_1, X_2 \sim \mathrm{Geo}(p)$ בלתי תלויים. יש להראות כי בהינתן המאורע $\{X_1 + X_2 = n\}$ מתקיים כי X_1 מתפלג אחיד על $[n-1]$.

✎ **בעיה 4.12** (התפלגות בינומית שלילית). יהיו $X_1, X_2, \ldots, X_r \sim \mathrm{Geo}(p)$ בלתי תלויים. יש להראות

(א) כי $\mathbb{P}(X_1 + X_2 = 2 + k) = \binom{2+k}{r}(1-p)^2 p^k$,

(ב) $\mathbb{P}(\sum_{i=1}^{r} X_i = r + k) = \binom{k+r}{r}(1-p)^2 p^k$.

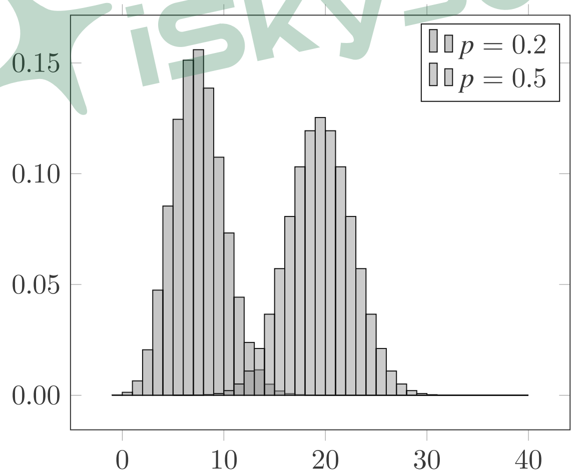
הערה: התפלגותו של הסכום $r - (\sum_{i=1}^{r} X_i) = X$ בבעיה 4.12 מכונה **התפלגות בינומית שלילית** או **התפלגות פסקל** עם r הצלחות ו-p סיכויי הצלחה, ומסמנים $X \sim \mathrm{NB}(r, p)$. התפלגות זו מתאימה למספר הכשלונות עד ל-r הצלחות בסדרה של ניסויים בלתי-תלויים.

4.4.2 התפלגות בינומית

שוב נסתכל על ניסוי חוזר בעל הסתברות הצלחה p, ב"ת בתוצאות הניסויים הקודמים. הפעם נבקש לדעת כמה ניסויים הצליחו מתוך N הניסויים הראשונים. התפלגות משתנה מקרי זה מכונה **התפלגות בינומית**.

הגדרה 4.28 (התפלגות בינומית). נאמר שמ"מ X מתפלג לפי **התפלגות בינומית על N ניסיונות עם סיכוי הצלחה** p ונכתוב $X \sim \mathrm{Bin}(N, p)$ אם לכל $n \in \{0, \ldots, N\}$ מתקיים $p_X(n) = \binom{N}{n} p^n (1-p)^{N-n}$.

פונק' התפלגות נק'



התפלגות בינומית עבור $N = 40$ לשני ערכי p.

בְּלֵאִיז פָּסְקַל (1623-1662), מתמטיקאי צרפתי ותיאולוג קתולי. בגיל 16 פיתח את הגיאומטריה הפרוייקטיבית ולאורך חייו תרם רבות לפורמליזם המתמטי, ולחשבון הדיפרנציאלי והאינטגרלי. בשנות העשרים לחייו פיתח בהתכתבויות עם פייר פרמה ועם חברו אַנְטוּאַן גוֹמְבּוֹ, שנודע כשבלייר דה מארה, את יסודות תורת ההסתברות ובמיוחד נודע בשל פתרון שלל בעיות בתחום ההימורים. הופעתה של ההתפלגות הבינומית השלילית בפתרונות אלו הביא לכך שנקראה על שמו. בגיל 31 חווה חוויה רוחנית עזה (יתכן שבעקבות תאונת מרכבה בה היה מעורב) והפנה חלק ניכר ממרצו לעיסוקים תיאולוגיים.

נשים לב שהתפלגות בינומית עם $N = 1$ היא לא אחרת מאשר התפלגות ברנולי עם סיכוי הצלחה p.

תוך כדי כך שנוודא כי התפלגות בינומית היא אומנם התפלגות נשים לב לקשר מעניין בינה ובין הבינום של ניוטון. נרשום

$$\mathbb{P}(X \in \{0, \ldots, N\}) = \sum_{n=0}^{N} \binom{N}{n} p^n (1-p)^{N-n} = \big(p + (1-p)\big)^N = 1$$

כאשר השוויון המרכזי הוא בדיוק נוסחת הבינום.

$P_x(0) = 1-p$

$P_x(1) = p$

טענה 4.29 (אפיון התפלגות בינומית)**.** יהיו $\{X_i\}_{i \in [N]}$ משתני ברנולי p בלתי-תלויים, אזי

$$\sum_{i \in [N]} X_i \sim \mathrm{Bin}(N, p).$$

$P_x(0) = 1-p$

$P_x(1) = p$

הוכחה. יהי $n \in \{0, \ldots, N\}$ ונסמן $Y = \sum_{i \in [N]} X_i$ ונחשב

$$\mathbb{P}(Y = n) = \sum_{\substack{(x_1, \ldots, x_N) \in \{0,1\}^N \\ \sum_{i \in [N]} x_i = n}} \mathbb{P}(\forall i \in [N],\ X_i = x_i) \overset{\text{אי-תלות}}{=} \sum_{\substack{(x_1, \ldots, x_N) \in \{0,1\}^N \\ \sum_{i \in [N]} x_i = n}} \left(\prod_{i \in [N]} \mathbb{P}(X_i = x_i) \right) = \binom{N}{n} p^n (1-p)^{N-n}$$

■

מסקנה 4.30 (חיבור התפלגויות בינומיות)**.** אם $X \sim \mathrm{Bin}(N, p)$, $Y \sim \mathrm{Bin}(M, p)$ ב"ת, אז $X + Y \sim \mathrm{Bin}(N + M, p)$.

הוכחה. יהיו $X_1, \ldots, X_{N+M}$ מ"מ ברנולי ב"ת עם סיכוי הצלחה p. לפי טענה 4.29 מתקיים

$$\sum_{n=1}^{N} X_n \sim \mathrm{Bin}(N, p), \qquad \sum_{n=N+1}^{N+M} X_n \sim \mathrm{Bin}(M, p).$$

כיוון שאלו פונקציות של קבוצות משתנים שונות באוסף ב"ת, הרי ששני סכומים אלו הנם מ"מ ב"ת ולכן לפי טענה 4.23 התפלגותם המשותפת זהה לזו של X ו-Y. לכן, כיוון שהתפלגות הסכום תלויה רק בהתפלגויות המשותפת (אבחנה 4.17), הרי שמתקיים:

$$\sum_{n=1}^{N+M} X_i \overset{\mathrm{d}}{=} X + Y.$$

לפי טענה 4.29 התפלגותו של $\sum_{n=1}^{N+M} X_i$ היא בינומית עם $N + M$ נסיונות והסתברות הצלחה p. ■

✎ **בעיה 4.13.** מטילים מטבע הוגן 10 פעמים. לאחר מכן מטילים שוב את כל ההטלות שתוצאתן הייתה עץ. יש להראות שמספר תוצאות הפלי לאחר ההטלות החוזרות מתפלג בינומית ולמצוא את סיכוי ההצלחה.

✎ **בעיה 4.14.** בתהליך הייצור של מסך דק כל פִּיקְסֶל עלול להיות פגום בהסתברות של אחד לשני מיליון. אם במסך יש פיקסל פגום הוא נחשב סוג ב', ואם יש בו שני פיקסלים פגומים הוא מושמד. חשב (הנך רשאי להשתמש במכונת חישוב) מה ההסתברות שהמסך יהיה סוג ב' ומה ההסתברות להשמדת המסך ברזולוציות הבאות:

א. VGA (640×480 פיקס') ב. 720p (1280×720 פיקס') ג. 1080p (1920×1080 פיקס')

ד. WQHD (2560×1440 פיקס') ה. UHD-1 (3840×2160 פיקס') ו. 4K (4096×2160 פיקס')

(עד לסדרה של התקדמויות טכנולוגיות סביב שנת 2010, הציבו פיקסלים פגומים מגבלה משמעותית על רזולוציית מסכים.)

✎ **בעיה 4.15.** נניח כי הסתברותו של נדאל לנצח את פדרר במערכה של משחקי טניס היא שני שליש, באופן בלתי תלוי בתוצאת המערכות הקודמות. השניים משחקים עד אשר אחד מהם זוכה בשלוש מערכות בסך הכל. מה ההסתברות שנדאל יזכה בטורניר?

דוגמא 4.31 (חישוב ההסתברות לשוויון במשחק הוגן).
מוטלים $2N$ מטבעות הוגנים. הערך את ההסתברות שבדיוק N מטבעות יפלו על עץ.

תשובה: לפי טענה 4.29 מספר המטבעות שתוצאת הטלתם היא עץ מתפלג בינומית עם מספר ניסויים $2N$, וסיכוי הצלחה $1/2$. על כן, ההסתברות לשוויון היא בדיוק $\binom{2N}{N}2^{-2N}$, נשתמש בנוסחת סטרלינג האסימפטוטית ונקבל

$$\binom{2N}{N}2^{-2N} = \frac{2^{-2N}\sqrt{4\pi N}(2N/e)^{2N}(1+R(2N))}{\left(\sqrt{2\pi N}(N/e)^N(1+R(N))\right)^2}.$$

כאשר $\lim_{n\to\infty} R(n) = 0$. ומכאן שמתקיים: $\lim_{N\to\infty} \sqrt{\pi N}\binom{2N}{N}2^{-2N} = 1$,

כלומר ההסתברות שתוצאת מחצית ההטלות תהיה עץ היא בקירוב $1/\sqrt{\pi N}$. אם ברצוננו בחסמים הנכונים לכל $N \in \mathbb{N}$ נוכל להשתמש בגרסא המתאימה של נוסחת סטרלינג ולקבל את ההערכה

$$\frac{0.47}{\sqrt{N}} \le \frac{2\sqrt{\pi}}{e^2\sqrt{N}} = \frac{\sqrt{2\pi}(2N)^{2N+1/2}e^{-2N}}{\left(eN^{N+1/2}e^{-N}\right)^2 2^{2N}} \le \binom{2N}{N}2^{-2N} \le \frac{e(2N)^{2N+1/2}e^{-2N}}{\left(\sqrt{2\pi}N^{N+1/2}e^{-N}\right)^2 2^{2N}} = \frac{e\sqrt{2}}{2\pi\sqrt{N}} \le \frac{0.62}{\sqrt{N}}.$$

✎ **בעיה 4.16.** להראות כי אם $X \sim \mathrm{Bin}(N,p)$ אז $(N-X) \sim \mathrm{Bin}(N,1-p)$.

4.4.3 התפלגות פואסון

ההתפלגות האחרונה שנציג בפרק זה שונה מקודמותיה בכך שהיא אינה מתאימה באופן ישיר למדד כלשהו המתייחס לניסוי חוזר, אלא למעין מקביל רציף שלו. היא גם הייתה האחרונה להתגלות בין ההתפלגויות המתוארות כאן, שכן פורסמה ברבים לראשונה על ידי סימון דניס פואסון (Siméon Denis Poisson) ב-1837 (אף על פי שהופיעה בחיבור של אברהם דה-מואבר ב-1711 וכנראה שהיה ראוי שתיקרא על שמו). בכדי להבין את התועלת שבהתפלגות זו נתאר מחקר שביצע הכלכלן לדיסלאוץ וון בורטקייביץ' (Ladislaus Von Bortkwicz) ב-1898 שבמסגרתו הוא מקדם את השימוש בהתפלגות זו.

הברון **סִימוֹן דֶנִיס פּוּאָסוֹן** (1781-1840), מתמטיקאי צרפתי פורה אשר פרסם למעלה משלוש-מאות מאמרים ועבודות, חבר האקדמיה הצרפתית למדעים, זכה להוקרה במשך כל חייו על כישוריו כחוקר וכמרצה. פואסון שהיה תלמידו של לפלאס אשר נהג בו קרבה יתירה, ראה עצמו כממשיך דרכו. תרם תרומה כבדת משקל לאנליזה, למתמטיקה שימושית ולמתמטיקה פיסיקלית (מדינמיקה של נוזלים, דרך תנועת גרמי השמיים ועד תיאור האור כגל). דווקא תרומתו להסתברות הייתה צנועה יחסית, ופרסומה של ההתפלגות הקרויה על שמו, אך אשר נתגלתה על ידי אברהם-דה-מואבר, הופיע בפרסום מתחום המתמטיקה השימושית שעסק במודל הסתברותי להרשעות שווא במערכת המשפט.

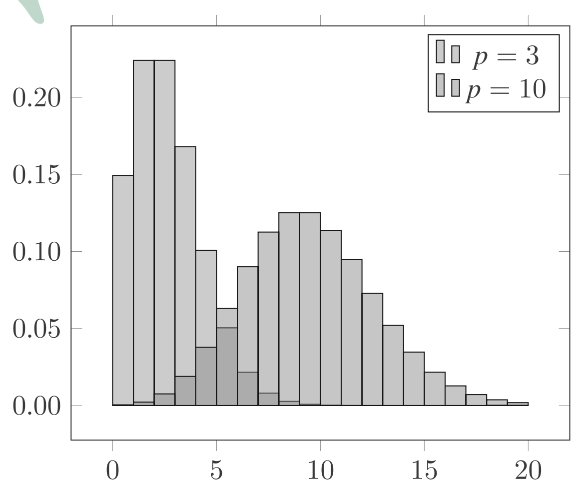
בספרו ''חוק המספרים הקטנים'', מתאר וון בורטקייביץ' מחקר שערך על מספר החיילים בצבא הפרוסי שנהרגים מבעיטת סוס ידידותי. לכאורה נצפה כי העובדה שחייל מסויים נבעט למוות על ידי סוס היא כמעט ב''ת בגורלם של חיילים אחרים. לכן, טבעי שנשתמש במודל של התפלגות בינומית. ואולם מספר החיילים הוא גדול מאוד והסיכוי של כל אחד מהם להיבעט למוות קטן יחסית (ולראיה רק מספר מצומצם של חיילים אומנם נבעטים למוות – בכל שנה 0.61 חיילים בממוצע בשנה בחטיבה המונה כאלף חיילים). התפלגות בינומית היא מורכבת לניתוח מדוייק. כדי להתגבר על בעיה זו נוכל להשאיף את מספר החיילים בחטיבה לאינסוף ולהשאיף את ההסתברות של כל אחד מהם להיבעט לאפס תוך שימור התכונה שבאופן טיפוסי יבעטו 0.61 חיילים בחטיבה בשנה. ההתפלגות הגבולית במקרה זה היא התפלגות פואסון עם פרמטר 0.61, וון בורטקייביץ' אומנם ממחיש במחקרו כי התפלגות פואסון אומנם מתארת בדיוק מפליא את כמות החיילים שנבעטו למוות.

התפלגות פואסון משמשת אפוא לקירוב מספר המאורעות המתרחשים בפועל מתוך מספר רב של מאורעות שעשויים להתרחש באופן בלתי תלוי ובהסתברות נמוכה כל אחד. מלבד הדוגמא הקלאסית שלעיל, היא מתארת היטב את מספר האטומים המתפרקים בפרק זמן נתון בחומר רדיואקטיבי ואת מספר שיחות הטלפון המתבצעות בנקודת זמן מסוימת.

הגדרה 4.32 (התפלגות פואסון)**.** נאמר שמ''מ X מתפלג לפי **התפלגות פואסון (או פואסונית) עם שכיחות** λ ונכתוב $X \sim \mathrm{Po}(\lambda)$ אם לכל $n \in \mathbb{N}_0$ מתקיים

$$p_X(n) = \frac{\lambda^n e^{-\lambda}}{n!}.$$

נשים לב ש- $\sum_{n\in\mathbb{N}_0} p_X(n) = \sum_{n\in\mathbb{N}_0} \frac{\lambda^n e^{-\lambda}}{n!} = 1$ שכן זהו טור טיילור של הפונקציה e^x בנקודה λ מוכפל ב $e^{-\lambda}$.



התפלגות פואסון עבור ערכים עד 20 לשני ערכי λ.

בפרק הפניה, נמחיש כיצד ניתן לקבל את התפלגות פואסון כגבול של סדרת התפלגויות בינומיות.

טענה 4.33 (תכונות התפלגות פואסונית)**.** יהי $X \sim \mathrm{Po}(\lambda)$.

דוג' נורמלית עבור התפלגות פואסון

ההתפלגות הפואסונית משמשת קירוב טוב להתפלגות הבינומית כאשר מספר הניסוים המבוצעים (n) הוא גדול מאוד, וההסתברות ל"הצלחה" בכל ניסוי p , קרובה ל-0.

אולם, להתפלגות זו חשיבות לכשלעצמה. כללית היא מתארת את ההסתברות כי אירוע מסויים יקרה k פעמים בזמן נתון t כאשר האירוע קורה באופן אקראי לחלוטין, וקיימת אי-תלות בין ההופעות השונות של האירוע.

לדוגמא:

1. מה ההסתברות כי בקטע כביש מסויים יעברו 10 מכוניות (k=10) בזמן של 5 שניות (t=5).
2. מה ההסתברות כי למרכזיית טלפונים יגיעו 25 שיחות (k=25) במשך 3 דקות (t=3).
3. מה ההסתברות כי למחסן הטכני במפעל יגיעו בפרק זמן של שעה (t=1) שבעה עובדים (k=7) להוצאת ציוד.

יהי x משתנה מקרי בדיד היכול לקבל את הערכים ...x=0,1,2

תהיה λ קצב המופע של האירועים (דהיינו: ממוצע האירועים ביחידת זמן).

x הינו משתנה מקרי פואסוני , אזי פונקצית ההסתברות שלו היא:

$$P(\text{יתרחשו } k \text{ אירועים בזמן } t) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

כאשר:

(הבסיס של הלוגריתם הטבעי) e = 2.718...

דוג' פרקטית:

שאלה L5.10

אם ידוע שמספר פניות בדקה למודיעין של שירותי טלפון מתפלג פואסונית עם ממוצע של 5 פניות **בדקה אחת**, מהי ההסתברות

א. שבין השעה 10:00 ל—10:01 לא תתקבל אף פנייה?
ב. שבדקה הזאת יתקבלו לכל היותר 3 פניות?
ג. שבמשך **2 דקות** לא תכנס אף שיחה?
ד. שב**שעה** הראשונה יכנסו 333 שיחות?

פתרון.

יהיה משתנה מקרי X_1 = {**מספר פניות בדקה אחת**}. על פי נתוני השאלה, $X_1 \sim \mathrm{P}(\lambda = 5)$.

א. $P(X_1 = 0) = \frac{5^0}{0!}e^{-5} \approx 0.0067$.

ב.

$$P(X_1 \le 3) = P(X_1 = 0) + P(X_1 = 1) + P(X_1 = 2) + P(X_1 = 3)$$
$$= e^{-5} + \frac{5^1}{1!}e^{-5} + \frac{5^2}{2!}e^{-5} + \frac{5^3}{3!}e^{-5} \approx 0.265$$

ג. יהיה משתנה מקרי X_2 = {מספר פניות במשך **שתי דקות**}. על פי נתוני השאלה, $X_2 \sim \mathrm{P}(\lambda = 10)$.

$$P(X_2 = 0) = \frac{10^0}{0!}e^{-10} \approx 0.000045$$

ד. יהיה משתנה מקרי X_3 = {מספר פניות במשך **שעה אחת**}. על פי נתוני השאלה, $X_3 \sim \mathrm{P}(\lambda = 300)$.

$$P(X_3 = 4) = \frac{300^{333}}{333!}e^{-300} \approx 0.0038$$

(א) אם $Y \sim \mathrm{Po}(\eta)$ ב"ת ב-X אז $X+Y \sim \mathrm{Po}(\lambda+\eta)$

(ב) אם $Y \mid X \sim \mathrm{Bin}(X,p)$ אז $Y \sim \mathrm{Po}(\lambda p)$.

הוכחה. עבור סעיף (א) נחשב לפי נוסחת ההסתברות השלמה עבור $n \in \mathbb{N}_0$,

$$\mathbb{P}(X+Y=n) = \sum_{i\in\mathbb{N}_0} \mathbb{P}(X=i)\mathbb{P}(X+Y=n \mid X=i) \overset{\text{המחוברים } i=n+1 \text{ ואילך הם אפס}}{=} \sum_{i=0}^{n} \frac{\lambda^i e^{-\lambda}}{i!}\frac{\eta^{n-i}e^{-\eta}}{(n-i)!}$$

$$= \frac{e^{-\lambda-\eta}}{n!}\sum_{i=0}^{n} \binom{n}{i}\lambda^i\eta^{n-i} \overset{\text{נוסחת הבינום}}{=} \frac{(\lambda+\eta)^n e^{-\lambda-\eta}}{n!}$$

עבור סעיף (ב) נחשב לכל $n \in \mathbb{N}_0$,

$$\mathbb{P}(Y=n) = \sum_{i=0}^{\infty} \mathbb{P}(X=i)\mathbb{P}(Y=n \mid X=i) \overset{n \text{ המחוברים הראשונים הם אפס}}{=} \sum_{i=n}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!}\binom{i}{n}p^n(1-p)^{i-n}$$

$$\overset{\text{שינוי אינדקס}}{=} (p\lambda)^n \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{(n+i)!}\binom{n+i}{n}(1-p)^i = \frac{(p\lambda)^n e^{-\lambda}}{n!}\sum_{i=0}^{\infty}\frac{(\lambda-\lambda p)^i}{i!} \overset{\text{טור טיילור}}{=} \frac{(p\lambda)^n e^{-\lambda p}}{n!}$$

■

לשני הסעיפים בטענה 4.33 ישנו גם הסבר אינטואיטיבי. אשר לסעיף (א), אם נשוב לפרשים הפרוסים הנבעטים למוות על ידי סוסיהם. מובן שנצפה שכאשר נחבר את הנספים בשתי חטיבות, קצב ההיבעטות המשותף יהיה בדיוק סכום קצבי שתי היחידות. באשר לסעיף (ב), נניח שנסתכל על ניסוי לפיו מקרה של בעיטה למוות מדווח באופן בלתי תלוי בכל מקרה אחר בסיכוי חצי. כיוון שהיות שהמאורע שפרש ייבעט למוות ומותו ידווח עודנה בלתי תלויה בקירוב עבור פרשים שונים - הרי שהתפלגות מקרי המוות מבעיטת סוסים שדווחו היא בקירוב פואסונית. היות שאנו מתוודעים למחצית המקרים בקירוב נצפה שקצב המאורעות המדווחים יהיה מחצית מקצב המאורעות בכלל.

✎ **בעיה 4.17.** להראות כי אם $X \sim \mathrm{Po}(\lambda)$, $Y \sim \mathrm{Po}(\eta)$ ב"ת, אז לכל $N \in \mathbb{N}$ המשתנה המקרי המותנה $(X \mid X+Y=N)$ מתפלג בינומית עם N נסיונות וסיכוי הצלחה $\lambda/(\lambda+\eta)$.

*4.5 סיגמא-אלגברה: המידע המקודד במשתנה מקרי

כפי שנוכחנו, ההגדרה והטיפול במערכת תלויות מורכבת בין משתנים מקריים (לדוגמא שני אוספים בלתי-תלויים של משתנים מקריים), היא מסובכת ומסורבלת בהשוואה לטיפול באי-תלות של מערכת מקבילה עבור מאורעות במרחב הסתברות. בפרק זה נכיר את מושג ה-σ-אלגברה (קרי: סִיגְמָא-אַלְגֶּבְּרָה) שיאפשר לנו להתמודד ביתר קלות עם מורכבות זו. תחום זה דורש בשלות מתמטית גבוהה יותר מיתר הספר ומכיוון שיתר הספר אינו מסתמך על ידיעתו, לקורא שאינו מנוסה בעבודה עם ביטויים מורכבים מוטב לדלג עליו בקריאה ראשונה.

נפתח בהגדרת ה-σ-אלגברה כתכונה של אוסף של קבוצות.

הגדרה 4.34 (סיגמא-אלגברה)**.** נאמר כי אוסף קבוצות $\mathcal{A} \subset 2^{\Omega}$ המכיל את $\emptyset$ הוא σ-אלגברה מעל Ω, אם מתקיים לכל $\{A_i\}_{i\in\mathbb{N}}$ קבוצות ב-$\mathcal{A}$.

(א) **סגירות להשלמה.** $\Omega \setminus A_1 \in \mathcal{A}$.

(ב) **סגירות לחיתוך סופי.** $A_1 \cap A_2 \in \mathcal{A}$.

(ג) **סגירות לאיחוד בן מניה.** $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}$.

אנו נתעניין במשפחה מצומצמת של סיגמא-אלגבראות, המכונה סיגמא-אלגבראות חלוקתיות.

הגדרה 4.35 (סיגמא-אלגברה חלוקתית)**.** תהי $\{A_i\}_{i\in\Lambda}$ חלוקה של קבוצה Ω (כמתואר בהגדרה 3.6). הקבוצה $\mathcal{A} := \{\bigcup_{i\in\Lambda'} A_i \;:\; \Lambda' \subset \Lambda\}$ מכונה **ה-σ-אלגברה החלוקתית הנוצרת ע"י** $\{A_i\}_{i\in\Lambda}$.

טענה 4.36. σ-אלגברה חלוקתית היא σ-אלגברה.

הוכחה. תהי $\{A_i\}_{i\in\Lambda}$ חלוקה של Ω ותהי $\mathcal{A} := \{\bigcup_{i\in\Lambda'} A_i \;:\; \Lambda' \subset \Lambda\}$. אומנם, $\emptyset = \bigcup_{i\in\emptyset} A_i \in \mathcal{A}$. עלינו לוודא אפוא, כי $\mathcal{A}$ מקיימת את שלוש התכונות הנדרשות לפי הגדרה . **תכונה (א).** היות ש-A_i היא חלוקה, הרי שאם $A = \bigcup_{i\in\Lambda'} A_i$ נמצאת ב-$\mathcal{A}$, אז גם $A^c = \bigcup_{i\in\Lambda\setminus\Lambda'} A_i \in \mathcal{A}$. **תכונה (ב).** אם $A = \bigcup_{i\in\Lambda'} A_i$ ו-$B = \bigcup_{i\in\Lambda''} A_i$ אז $A \cap B = \bigcup_{i\in\Lambda'\cap\Lambda''} A_i$ ולכן $\mathcal{A}$ סגורה לחיתוכים. **תכונה (ג).** אם $\{B_j\}_{j\in\mathbb{N}}$ מקיימות $B_j = \bigcup_{i\in\Lambda_j} A_i$ אז $\bigcup_{j\in\mathbb{N}} B_j = \bigcup_{i\in\bigcup_{j\in\mathbb{N}}\Lambda_j} A_i$ ולכן $\mathcal{A}$ סגורה לאיחוד בן-מניה. קיבלנו אפוא, ש-$\mathcal{A}$ היא σ-אלגברה. ∎

יהי X משתנה מקרי על מרחב הסתברות בדיד $(\Omega, \mathcal{F}, \mathbb{P})$. נזכר כי הביטוי $X \in A$ הוא תיאור אחר למאורע ב-$\mathcal{F}$ הנתון על ידי $X^{-1}(A)$. הבה נעיין באוסף כלל המאורעות שניתן לגשת אליו באופן זה ונווכח כי הוא σ-אלגברה חלוקתית.

טענה 4.37 (סיגמא-אלגברה נוצרת על ידי משתנה מקרי)**.** יהי X משתנה מקרי על מרחב הסתברות בדיד $(\Omega, \mathcal{F}, \mathbb{P})$. הקבוצה

$$\sigma(X) = \{X^{-1}(A) \;:\; A \subset \mathbb{R}\}$$

של מאורעות ב-$\mathcal{F}$ היא σ-אלגברה, המכונה **ה-σ-אלגברה הנוצרת מ-X.** זוהי ה-σ-אלגברה החלוקתית הנוצרת מ- $\{X^{-1}(X(\omega)) \;:\; \omega \in \Omega\}$.

הוכחה. די שנראה כי $\sigma(X)$ היא ה-σ-אלגברה החלוקתית הנוצרת מ- $\{X^{-1}(X(\omega)) \;:\; \omega \in \Omega\}$, שכן מטענה 4.36 ינבע כי זו σ-אלגברה. לשם כך נסמן ב- $\mathcal{A}$ את ה-σ-אלגברה החלוקתית הנתונה בשאלה. נבחין כי לכל $R \subset \mathbb{R}$, מתקיים $X^{-1}(R) = X^{-1}(X(X^{-1}(R))) = \bigcup_{\omega: X(\omega)\in R} X^{-1}(X(\omega))$, ולכן $\sigma(X) \subset \mathcal{A}$. מצד שני, עבור $A \in \mathcal{A}$ קיימת $\Lambda \subset \Omega$ כך ש- $A = \bigcup_{\omega\in\Lambda} X^{-1}(X(\omega)) = X^{-1}(\bigcup_{\omega\in\Lambda} X^{-1}(X(\omega)))$ ולכן $\mathcal{A} \subset \sigma(X)$. קיבלנו כי $\mathcal{A} = \sigma(X)$ כנדרש. ∎

נשים לב שלפי הגדרה 4.18, אי-תלות בין משתנים מקריים X_1, X_2 היא למעשה אי תלות בין אוספי המאורעות $\sigma(X_1)$ ו-$\sigma(X_2)$ (ר' הגדרה 3.24). נסכם זאת באבחנה הבאה.

אבחנה 4.38 (אי תלות בין-משתנים בשפה של σ-אלגבראות). לכל זוג משתנים מקריים X_1, X_2 על אותו מרחב הסתברות, מתקיים כי $X_1 \perp\!\!\!\perp X_2$ אם ורק אם $\sigma(X_1) \perp\!\!\!\perp \sigma(X_2)$.

כמו כן, ניזכר כי אי-תלות של אוספי קבוצות היא תכונה תורשתית.

אבחנה 4.39 (תורשתיות של אי-תלות בין σ-אלגבראות). אם $\mathcal{A}, \mathcal{A}', \mathcal{B}, \mathcal{B}'$ σ-אלגבראות של מאורעות מקיימות $\mathcal{A}' \subset \mathcal{A}$, $\mathcal{B}' \subset \mathcal{B}$ ו-$\mathcal{A} \perp\!\!\!\perp \mathcal{B}$, אז מתקיים $\mathcal{A}' \perp\!\!\!\perp \mathcal{B}'$.

דוגמא 4.40. עבור מרחב הסתברות על $\{0,1\}^2$ יהי $X((a,b)) = a + b \pmod 2$ מ"מ. מה היא $\sigma(X)$?

תשובה: $\sigma(X) = \{X^{-1}(A) \ : \ A \subset \mathbb{R}\}$ המקור לפי X של שתי קבוצות ב-$\mathbb{R}$ שונה רק אם הן נבדלות לפחות באיבר אחד של Ω, לכן אם נסמן $R = \cup_{\omega\in\Omega} X(\omega)$ נקבל כי $\sigma(X) = \{X^{-1}(A) \ : \ A \subset \mathbb{R}\}$. במקרה שלנו $R = \{0,1\}$ ולכן

$$\sigma(X) = \{X^{-1}(\emptyset), X^{-1}(\{0\}), X^{-1}(\{1\}), X^{-1}(\{0,1\})\} = \{\emptyset, \{(0,0),(1,1)\}, \{(0,1),(1,0)\}\Omega\}.$$

✎ **בעיה 4.18.** יהי $X(a,b) = a + b$ עבור $\Omega = [3]^2$. מה היא ה-σ-אלגברה הנוצרת ע"י X?

באמצעות מושג זה נוכל להראות כי הפעלה של פונקציה על משתנה מקרי מצמצמת את ה-σ-אלגברה.

הגדרה 4.41 (תת σ-אלגברה). אם $\mathcal{A}$ ו-$\mathcal{B}$ הן σ-אלגבראות כך ש- $\mathcal{A} \subset \mathcal{B}$ אז $\mathcal{A}$ נקראת **תת σ-אלגברה** של $\mathcal{B}$.

כיוון שאנו עוסקים ב-σ-אלגבראות חלוקתיות, נציג יחסי עידון וגסות בין חלוקות שיוכלו לשמש אותנו לנתח מצבים שבהם מגלמת אלגברה אחת בתוכה את כל המידע המצוי באלגברה אחרת.

הגדרה 4.42 (עידון וגסות של חלוקות). תהיינה $\{A_i\}_{i\in\Lambda_A}$ ו-$\{B_i\}_{i\in\Lambda_B}$ חלוקות של קבוצה Ω. נאמר כי $\{A_i\}_{i\in\Lambda_A}$ **מעדנת את** $\{B_i\}_{i\in\Lambda_B}$ או כי $\{B_i\}_{i\in\Lambda_B}$ **גסה יותר מ-** $\{A_i\}_{i\in\Lambda_A}$ אם לכל $j \in \Lambda_B$ קיים $I_j \subset \Lambda_A$ כך ש- $\bigcup_{i\in I_j} A_i = B_j$.

עידון של חלוקה הוא למעשה שבירה של חלק ממחלקותיה לתת-מחלקות קטנות יותר.

טענה 4.43 (קשר בין עידון לתת-σ-אלגבראות). תהיינה $\mathcal{A}$ ו-$\mathcal{B}$ שתי σ-אלגבראות חלוקתיות על מרחב Ω המוגדרות על ידי $\{A_i\}_{i\in\Lambda_A}$ ו-$\{B_i\}_{i\in\Lambda_B}$ בהתאמה. אזי $\mathcal{B} \subset \mathcal{A}$ אם ורק אם $\{A_i\}_{i\in\Lambda_A}$ היא עידון של $\{B_i\}_{i\in\Lambda_B}$.

הוכחה. **בכיוון אחד,** נניח כי $\mathcal{B} \subset \mathcal{A}$ ונקבל כי לכל $j \in \Lambda_B$ קיים Λ'_A כך ש-$B_j = \bigcup_{i\in\Lambda'_A} A_i$, כך **בכיוון השני,** נניח כי $\{A_i\}_{i\in\Lambda_A}$ היא עידון של $\{B_i\}_{i\in\Lambda_B}$ ולכל $\Lambda'_B \subset \Lambda_B$ נגדיר $\Lambda'_A = \bigcup_{j\in\Lambda'_B} A_{I_j}$ ומיד ניווכח כי $\bigcup_{i\in\Lambda'_A} A_i = \bigcup_{i\in\Lambda'_B} B_i$, כך ש-$\mathcal{B} \subset \mathcal{A}$. ■

טענה 4.44 (פונקציות על מ"מ מצמצמות את ה-σ-אלגברה). יהי X משתנה מקרי על מרחב מדגם Ω ותהי פונקציה $f : \mathbb{R} \to \mathbb{R}$. אז $\sigma(f(X))$ היא תת-σ-אלגברה של $\sigma(X)$.

הוכחה. נוכח טענה 4.43, די שנראה כי החלוקה שמשרה $Y = f(X)$ גסה יותר מהחלוקה שמשרה X. ואכן:

$$Y^{-1}(Y(\omega)) = X^{-1} \circ f^{-1}(f(X(\omega)) = \bigcup_{\omega' \in f^{-1}(f(X(\omega)))} X^{-1}(\omega').$$

■

4.5.1 סיגמא-אלגברה נוצרת ממאורעות

כעת נבקש להכליל את השימוש ב-σ-אלגבראות למספר רב של משתנים. ראשית נשים לב כי חיתוך של σ-אלגבראות – גם הוא σ-אלגברה.

טענה 4.45 (חיתוך σ-אלגבראות)**.** תהיינה $\{\mathcal{A}_i\}_{i\in\Lambda}$ σ-אלגבראות מעל קבוצה Ω. אזי $\bigcap_{i\in\Lambda} \mathcal{A}_i$ היא σ-אלגברה מעל Ω.

הוכחה. הקבוצה הריקה מוכלת בכל אחת מהקבוצות $\mathcal{A}_i$, ולכן גם בחיתוך. נוודא כי $\mathcal{A} = \bigcap_{i\in\Lambda} \mathcal{A}_i$ מקיימת את שלוש התכונות הנדרשות לפי הגדרה . **תכונה (א) - סגירות להשלמה.** אם $A \in \mathcal{A}$, אז $A \in \mathcal{A}_i$ לכל $i \in \Lambda$ ומכיון שאלה σ-אלגבראות, הרי ש$A^c \in \mathcal{A}_i$ לכל $i \in \Lambda$ ולכן $A^c \in \mathcal{A}$. **תכונה (ב) - סגירות לחיתוך.** אם $A, B \in \mathcal{A}$, אז $A, B \in \mathcal{A}_i$ לכל $i \in \Lambda$ ומכיון שאלה σ-אלגבראות, הרי ש$A \cap B \in \mathcal{A}_i$ לכל $i \in \Lambda$ ולכן $A \cap B \in \mathcal{A}$. **תכונה (ג) - סכימות בת-מניה.** אם $\{A_j : j \in \mathbb{N}\} \subset \mathcal{A}$, אז $\{A_j : j \in \mathbb{N}\} \subset \mathcal{A}_i$ לכל $i \in \Lambda$ ומכיון שאלה σ-אלגבראות, הרי ש$\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}_i$ לכל $i \in \Lambda$ ולכן $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}$. ■

העובדה שקבוצת ה-σ-אלגבראות מעל Ω סגורה לחיתוך, מאפשרת לנו להגדיר את **ה-σ-אלגברה המינימלית שמקיימת תכונה מסויימת**. זאת על ידי כך שנסתכל על כלל ה-σ-אלגבראות המקיימות תכונה זו ונחתוך אותן – האלגברה שתתקבל חייבת להיות מוכלת בכל σ-אלגברה המקיימת את התכונה. כעת נוכל להגדיר את ה-σ-אלגברה שנוצרת מקבוצת מאורעות.

הגדרה 4.46 (σ-אלגברה נוצרת מאוסף מאורעות)**.** יהי $\mathcal{S}$ אוסף תת-קבוצות של קבוצה Ω. ה-σ-אלגברה המינימלית המכילה את $\mathcal{S}$ היא **ה-σ-אלגברה הנוצרת מ-$\mathcal{S}$**. σ-אלגברה זו תסומן ב-$\sigma(\mathcal{S})$.

באמצעות מושג הσ-אלגברה הנוצרת מאוסף מאורעות נוכל להכליל את טענה 3.20.

טענה 4.47. יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות בדיד ויהיו $\mathcal{A} \perp\!\!\!\perp \mathcal{B}$, שני אוספי מאורעות סגורים לחיתוכים סופיים ובלתי-תלויים. אז מתקיים $\sigma(\mathcal{A}) \perp\!\!\!\perp \mathcal{B}$

הוכחה. נגדיר חלוקה $\{A_i\}_{i\in\mathbb{N}}$ של Ω על ידי יחס השקילות $\omega_1 \overset{\mathcal{A}}{\sim} \omega_2$ אם ורק אם לכל $A \in \mathcal{A}$ כך ש-$\omega_1 \in A$ מתקיים $\omega_2 \in A$. נשים לב כי מספר המחלקות של יחס שקילות זה אינו יכול לעלות על מספר אברי Ω ולכן זוהי אומנם חלוקה בת-מניה לכל היותר. נראה כי $\mathcal{A}'$, ה-σ-אלגברה החלוקתית המתאימה ל-$\{A_i\}_{i\in\mathbb{N}}$ בלתי תלויה ב-$\mathcal{B}$. כיוון שכל קבוצה היא איחוד זר ובן מניה של מחלקות, די אם נראה כי כל מחלקה בלתי-תלויה ב-$\mathcal{B}$. תהי A מחלקה כזו המתאימה לאיבר ω' ונגדיר $A = \bigcap_{\omega\in A^c} A_{n_\omega}$ כאשר A_{n_ω} היא קבוצה המכילה את ω' אך לא את

ω (קבוצה כזו חייבת להיות קיימת, אחרת ω היה ב-A). תהי $B \in \mathcal{B}$. היות ש-Ω בן-מניה נגדיר מניה של איבריו על ידי $\Omega = \{\omega_n\}_{n\in\mathbb{N}}$ ונחשב

$$\mathbb{P}(A\cap B) = \mathbb{P}\left(\bigcap_{\omega\in A^c} A_{n_\omega} \cap B\right) = \lim_{N\to\infty} \mathbb{P}\left(\bigcap_{\substack{i\in[N]\\ \omega_i\in A^c}} A_{n_{\omega_i}} \cap B\right)$$

$$= \lim_{N\to\infty} \mathbb{P}\left(\bigcap_{\substack{i\in[N]\\ \omega_i\in A^c}} A_{n_{\omega_i}}\right)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B),$$

כאשר השוויון הראשון נובע מהגדרת A, השני מרציפות פונקציית ההסתברות על מאורעות יורדים (מסקנה 2.21), השלישי מהנחת האי-תלות ומכך ש$\mathcal{A}$ סגורה לחיתוכים סופיים, והאחרון - שוב מרציפות פונקציית ההסתברות על מאורעות יורדים (מסקנה 2.21). מכאן נקבל כי $\mathcal{A}'$ היא σ-אלגברה המכילה את $\mathcal{A}$ ולכן היא מכילה את $\sigma(\mathcal{A})$. מפני ש- $\mathcal{A}'$ בלתי-תלויה ב-$\mathcal{B}$, נסיק כי $\sigma(\mathcal{A})$, המוכלת בה, אף היא בלתי-תלויה בה, כנדרש. ■

הגדרה 4.48 (σ-אלגברה נוצרת מאוסף מ"מ). יהי $\mathfrak{X}$ אוסף משתנים מקריים על מרחב מדגם Ω. ה-σ-אלגברה המינימלית המכילה את $\sigma(X)$ לכל $X \in \mathfrak{X}$ היא **ה-σ-אלגברה הנוצרת מ-**$\mathfrak{X}$ שתסומן ב-$\sigma(\mathfrak{X})$. כאשר $\mathfrak{X} = \{X_1,\ldots,X_n\}$ נסמן σ-אלגברה זו גם ב-$\sigma(X_1,\ldots,X_n)$.

טענה 4.49 (יצירת σ-אלגבראות במושגי עידון). יהיו $\{X_j\}_{j\in[N]}$ משתנים מקריים כך ש-$\sigma(X_j)$ מתאימה לחלוקה $\{A^j_k\}_{k\in\Lambda_j}$. אזי $\sigma(\{X_j\}_{j\in[N]})$ היא ה-σ-אלגברה החלוקתית המתאימה לחלוקה $\{A_{k_1,\ldots,k_N}\}_{k_j\in\Lambda_j}$ עבור הקבוצות $A_{k_1,\ldots,k_N} = \{\cap_{j\in[N]} A^j_{k_j}\}$.

במילים אחרות, החלוקה המתאימה ל-σ-אלגברה הנוצרת מאוסף של משתנים בדידים $\{X_n\}_{n\in[N]}$ היא החלוקה הנוצרת מכל האפשרויות לבחור מחלקה אחת מהחלוקה אשר מתאימה לכל אחד מהמשתנים ולחתוך את כל המחלקות הנבחרות.

הוכחה. ראשית נוודא כי $\{A_{i_1,\ldots,i_N}\}_{i_k\in\Lambda_k}$ היא אומנם חלוקה. לשם כך נשים לב כי

$$\bigcup_{i_1\in\Lambda_1}\cdots\bigcup_{i_k\in\Lambda_k} A_{i_1,\ldots,i_k} = \Omega$$

וכי אם $(k_1,\ldots,k_N) \neq (k'_1,\ldots,k'_N)$ אז $A_{k_1,\ldots,k_N}\cap A_{k'_1,\ldots,k'_N} = \emptyset$. נסמן את ה-$\sigma$-אלגברה המתאימה לחלוקה דנן ב-$\mathcal{A}$. ראשית **נראה כי** $\sigma(\{X_n\}_{n\in[N]}) \subset \mathcal{A}$. לשם כך עלינו להראות, לפי טענה 4.43, כי $A_{k_1,\ldots,k_N}$ היא עידון של כל חלוקה המתאימה ל-X_j, ואכן לכל A^j_k עבור $k\in\Lambda_j$ מתקיים

$$A^j_k = \bigcup_{k_1\in\Lambda_1}\cdots\bigcup_{k_{j-1}\in\Lambda_{j-1}}\bigcup_{k_{j+1}\in\Lambda_{j+1}}\cdots\bigcup_{k_k\in\Lambda_k} A_{k_1,\ldots,k_N}$$

כעת, **נראה כי** $\mathcal{A}\subset\sigma(\{X_n\}_{n\in[N]})$. לשם כך נשים לב כי לכל $j\in[N], k\in\Lambda_j$ מתקיים $A^j_k\in\sigma(\{X_n\}_{n\in[N]})$. מכאן, לפי סגירות לחיתוכים סופיים של $\sigma(\{X_n\}_{n\in[N]})$, הרי שכל אברי החלוקה $\{A_{i_1,\ldots,i_N}\}_{i_k\Lambda_k}$ מקיימים כי $A_{i_1,\ldots,i_N}\in\sigma(\{X_n\}_{n\in[N]})$. כיוון ש-$\sigma(\{X_n\}_{n\in[N]})$ סגורה לאיחודים בני-מניה, נסיק שכל אברי $\mathcal{A}$ נמצאים ב-$\sigma(\{X_n\}_{n\in[N]})$, כנדרש. ■

הטענה הבאה מאפשרת לנו להגדיר אי-תלות בין קבוצות של משתנים מקריים באופן טבעי יותר, באמצעות σ-אלגבראות.

טענה 4.50 (אי-תלות בין אוספי משתנים מקריים)**.** יהיו $\mathcal{X}$ ו-$\mathcal{Y}$ שני אוספי משתנים מקריים בדידים. אזי $\mathcal{X} \perp\!\!\!\perp \mathcal{Y}$ אם ורק אם $\sigma(\mathcal{X}) \perp\!\!\!\perp \sigma(\mathcal{Y})$.

הוכחה. ראשית **נניח כי** $\mathcal{X} \perp\!\!\!\perp \mathcal{Y}$ **ונראה כי** $\sigma(\mathcal{X}) \perp\!\!\!\perp \sigma(\mathcal{Y})$. יהיו $n, m \in \mathbb{N}$ שני מספרים טבעיים ויהיו $X_1, \dots, X_n \in \mathcal{X}$ ו-$Y_1, \dots, Y_m \in \mathcal{Y}$, $A_1 \in \sigma(X_1), \dots, A_n \in \sigma(X_n)$ ו-$B_1 \in \sigma(Y_1), \dots, B_m \in \sigma(Y_m)$. לפי הגדרה 4.22 מתקיים

$$\bigcap_{i\in[n]} A_i \perp\!\!\!\perp \bigcap_{i\in[m]} B_i$$

לכן אם נרשום $\mathcal{A}$ עבור אוסף כל המאורעות המופיעים באגף שמאל ו-$\mathcal{B}$ עבור אוסף כל המאורעות המופיעים באגף ימין, הרי שאלה אוספים בלתי-תלויים של מאורעות סגורים לחיתוכים סופיים. מטענה 4.47 נובע כי $\mathcal{B} \perp\!\!\!\perp \sigma(\mathcal{A})$. מהפעלה נוספת של אותה טענה נסיק כי $\sigma(\mathcal{B}) \perp\!\!\!\perp \sigma(\mathcal{A})$. נשים לב שלכל $X \in \mathcal{X}$ מתקיים $\sigma(X) \subset \mathcal{A}$ ולכל $Y \in \mathcal{Y}$ מתקיים $\sigma(Y) \subset \mathcal{B}$. נסיק כי, לפי הגדרה, $\sigma(\mathcal{X}) \subset \sigma(\mathcal{A})$ וכן $\sigma(\mathcal{Y}) \subset \sigma(\mathcal{B})$ ולכן, לפי אבחנה 4.39, מתקיים $\sigma(\mathcal{Y}) \perp\!\!\!\perp \sigma(\mathcal{X})$, כנדרש.

כעת **נניח כי** $\sigma(\mathcal{Y}) \perp\!\!\!\perp \sigma(\mathcal{X})$ **ונראה כי** $\mathcal{Y} \perp\!\!\!\perp \mathcal{X}$. יהיו $X_1, \dots, X_N$ ו-$Y_1, \dots, Y_M$ משתנים מקריים מהאוספים $\mathcal{X}$ ו-$\mathcal{Y}$ בהתאמה. נסמן $\mathcal{A} = \sigma(X_1, \dots, X_n)$, $\mathcal{B} = \sigma(Y_1, \dots, Y_n)$. נשים לב שלפי טענה 4.37, לכל $n \in [N]$ מתקיים כי $\sigma(X_n)$ הנה חלוקתית ומחלקותיה הן הקבוצות מהצורה $\{\omega : X_n(\omega) = x_n\} = X_n^{-1}(x_n)$ שאינן ריקות, עבור $x_n \in \mathbb{R}$. מטענה 4.49 נובע כי גם $\mathcal{A}$ הנה חלוקתית וכי מחלקותיה הן קבוצות מהצורה $\{\omega : X_1(\omega) = x_1 \dots X_n(\omega) = x_n\}$ שאינן ריקות עבור $x_1, \dots, x_N \in \mathbb{R}$. באופן דומה נקבל שהקבוצות הלא-ריקות מהצורה $\{\omega : Y_1(\omega) = Y_1 \dots Y_m = y_m\}$ עבור $y_1, \dots, y_M \in \mathbb{R}$ מהוות את מחלקותיה של $\mathcal{B}$. מכיוון ש-$\mathcal{A} \subset \sigma(\mathcal{X})$ ו-$\mathcal{B} \subset \sigma(\mathcal{Y})$ וכן $\sigma(\mathcal{Y}) \perp\!\!\!\perp \sigma(\mathcal{X})$, הרי שלפי אבחנה 4.39 מתקיים $\mathcal{B} \perp\!\!\!\perp \mathcal{A}$, כלומר לכל $x_1, \dots, x_N, y_1, \dots, y_M \in \mathbb{R}$ מתקיים:

$$\mathbb{P}(\{X_n = x_n, Y_m = y_m \ : \ n \in [N], m \in [M]\}) = \mathbb{P}(\{X_n = x_n \ : \ n \in [N]\})\mathbb{P}(\{Y_m = y_m \ : \ m \in [M]\}).$$

לאור הגדרה 4.18 נסיק כי $\mathcal{Y} \perp\!\!\!\perp \mathcal{X}$, כנדרש. ■

נשאיר לקורא להוכיח את פישוט הגדרתו של אוסף בלתי-תלוי של משתנים מקריים.

טענה 4.51 (אי-תלות בין מספר רב של משתנים מקריים)**.** יהי $\mathcal{X}$ אוסף משתנים מקריים בדידים. אזי המשתנים באוסף בלתי-תלויים אם ורק אם לכל $X \in \mathcal{X}$ מתקיים $\sigma(X) \perp\!\!\!\perp \sigma(\mathcal{X} \setminus \{X\})$.

✎ **בעיה 4.19.** להוכיח את טענה 4.51.

✎ **בעיה 4.20.** להוכיח את טענה 4.23 על סמך טענות פרק 4.5.

✎ **בעיה 4.21.** למצא σ-אלגברה מעל $\mathbb{N}$ שאינה חלוקתית. מדוע אין הדבר בסתירה לטענה 4.37?

4.5.2 סיגמא-אלגברה ומרחבי הסתברות

בפרק 2 הצגנו את אוסף המאורעות $\mathcal{F}$ במרחב הסתברות בדיד $(\Omega, \mathcal{F}, \mathbb{P})$ בתור $\mathcal{F} = 2^{\Omega}$. כך יכולנו להבטיח שכל אוסף תוצאות במרחב ההסתברות שלנו יכול להוות מאורע בעל הסתברות. מסתבר שאת מרבית תכונות מרחב ההסתברות יכולים היינו לשמר גם אילו בחרנו את $\mathcal{F}$ להיות σ-אלגברה כלשהי מעל Ω. זאת משום שתכונותיה של σ-אלגברה מאפשרות לנו לבצע על מאורעות את מרבית הפעולות שאנו זקוקים להן לצורך מתן מענה לשאלות הסתברותיות והגדרת משתנים מקריים. אי לכך, בתורת ההסתברות המודרנית מוגדר בדרך כלל מרחב הסתברות כשלשה $(\Omega, \mathcal{F}, \mathbb{P})$ המורכבת ממרחב מדגם, σ-אלגברה עליו ופונקציית הסתברות ממנה ל-$[0, 1]$. הדבר יתגלה כחשוב במיוחד אם נבקש להכליל את מרחבי ההסתברות למרחבי מדגם שאינם בני-מניה.

בעיות הרחבה והעשרה

✎ **בעיה 4.22.** עליסה נמצאת בחדר המבואה עם דלת כניסה ושלוש דלתות יציאה. דלת אחת מוליכה לארץ הפלאות, דלת שניה מוליכה לאנגליה והדלת השלישית מוליכה בחזרה לחדר המבואה. עליסה בכל ביקור בחדר המבואה עליסה בוחרת בדלת המוליכה לאנגליה בסיכוי p_1 ובדלת המוליכה לארץ הפלאות בסיכוי p_2 עבור $p_1 + p_2 < 1$ נתונים.

(א) מה הסיכוי שעליסה תגיע בסופו של דבר לארץ הפלאות?

(ב) כיצד מתפלג מספר הביקורים של עליסה בחדר המבואה בטרם תגיע לארץ הפלאות?

✎ **בעיה 4.23** (התפלגות פואסון מורכבת). מספר הנכנסים לחנות בגדים בשעה מתפלג פואסונית עם פרמטר 3. כל מבקר רוכש בהסתברות 0.5 פריט אחד ובהסתברות 0.5 שני פריטים. רשום נוסחא מפורשת להתפלגות מספר הפריטים שנרכשו במשך שעתיים.

✎ **בעיה 4.24** (דילול מ"מ גיאומטרי). יהי X מ"מ המקיים $(X+1) \sim \mathrm{Geo}(p)$ ויהי Y מ"מ המקיים $(Y \mid X) \sim \mathrm{Bin}(X, q)$. יש להראות כי $(Y+1) \sim \mathrm{Geo}\left(p/(p+q-pq)\right)$. כיצד נסביר תופעה זו?

✎ **בעיה 4.25** (יחס סדר). יהיו X, Y ו-Z שלושה מ"מ שווי התפלגות ובלתי תלויים. נסמן ב-A את המאורע שאין אף זוג משתנים שערכו זהה וב-B את המאורע $X < Y < Z$ יש לחשב את $\mathbb{P}(B \mid A)$ תוך שימוש קפדני בנימוקים פורמליים.

✎ **בעיה 4.26 (*).** [פואסון כקירוב לבינומי] מכונת צילום שכפלה מאתיים בחינות בהסתברות. בכל בחינה עלולה ליפול טעות שכפול בהסתברות 0.01 באופן בלתי-תלוי. הסתכל על מ"מ X_i המהווה מציין לטעות בבחינה ה-i. ונסמן ב- $[200] \in \{Y_i\}_i$ סדרה של משתני פואסון עם פרמטר $\ln 0.99$. כמו כן נרשום $X = \sum_{i=0}^{200} X_i$ ו-$Y = \sum_{i=0}^{200} Y_i$

(א) יש להראות כי $Y_1|(Y_1 \in \{0,1\})$ ו-X_1 שווי התפלגות.

(ב) מה ההסתברות ש-$Y_i \in \{0,1\}$?

(ג) כיצד נחסום משני הצדדים באמצעות התפלגות פואסון את ההסתברות שתפולנה טעויות שכפול בדיוק ב-k בחינות.

(רמז: ניתן להראות כי $\mathbb{P}(Y=k) - \mathbb{P}(Y \neq X) < \mathbb{P}(X=k) < \mathbb{P}(Y=k) + \mathbb{P}(Y \neq X)$)

✎ **בעיה 4.27 (**).** [פרדוקס הבחירה] יהיו X ו-Y שני משתנים מקריים שווי התפלגות ובלתי תלויים על $\mathbb{N}$ שהתפלגותם אינה קבועה. בשתי מעטפות נפרדות מונחים X שקלים ו-Y שקלים. יש למצא אסטרטגיה (שיכולה להיות הסתברותית ותלויה באקראיות נוספת) לבחור את המעטפה שבה סכום גבוה יותר בהסתברות **גדולה ממש** מחצי בהינתן ש-$Y \neq X$. לאסטרטגיה אסור להתחשב בהתפלגות של X ו-Y (היא אינה ידועה).