יסודות תורת ההסתברות

אוהד נ. פלדהיים
אורי גוראל גורביץ'

מכון איינשטיין למתמטיקה
האוניברסיטה העברית בירושלים

19 באוקטובר 2018

פרק 3

הסתברות מותנית ואי-תלות

"...התורה של ניסויים בלתי-תלויים היא בעת ובעונה אחת החלק הפשוט ביותר מבחינה אנליטית והמתקדם ביותר של תורת ההסתברות."

– ויליאם פלר, מבוא לתורת ההסתברות ושימושיה.

בפרק זה נערוך היכרות עם שני רעיונות מרכזיים בתורת ההסתברות. הראשון - הסתברות מותנית, מעניק לנו כלי פורמלי ראשון לתאר את השינויים שמתרחשים בהערכותינו באשר להסתברותן של התרחשויות לאור מידע חדש שנחשפנו אליו. רעיון זהו הוא מקור לא-אכזב לפרדוקסים ומסקנות לא אינטואיטיביות אך נכונות אשר כמה מן היפות שבהן ננסה לסקור בפרק זה. השני - מושג האי-תלות, המבטא את הרעיון לפיו ישנן עובדות שידיעתן הנה חסרת כל השפעה על הערכת ההסתברות לתוצאתו של ניסוי מסויים. מושג זה יכליל את מושג מרחב המכפלה וימשיך לשחק תפקיד מרכזי בפרקים הבאים.

3.1 הסתברות מותנית

הסתברות משמשת להערכת שכיחותן של תוצאות של ניסוי שטרם נערך. משעה שנערך הניסוי והתוצאה ידועה - הרי שמנקודת מבט הסתברותית הסתברות התוצאה הזו היא אחת והסתברות כל תוצאה אחרת – אפס. הניסוי הופך דטרמיניסטי. אלא שמעבר זה ממצב של חוסר וודאות למצב של ודאות מוחלטת אינו תמיד כה חד. נסתכל למשל על הדוגמא הבאה.

מר ישראלי מרים את שפופרת הטלפון ומחייג לבתאל, בתו. בעודו מרים את השפופרת, הוא משער בנפשו כמה סביר שקולה של בת-שיחו אומנם ישמע מעברו השני של הקו. משמשע את צליל החיוג הראשון מתחזק בטחונו שבת-שיחו תענה. הרי הצליל מהווה עדות בטוחה לכך שהתקשורת הצליחה והמכשיר מן העבר השני מצלצל. שניה נוקפת וצלצול שני נשמע. כעת בטחונו בכך שבתו תענה פוחת קמעה, אך אין זה אירוע נדיר שהמענה אינו מגיע בצלצול השני. נוקפים צלצול שלישי ורביעי ובטחונו פוחת והולך ורגע לפני שהחליט לנתק, פתאום נשמעת הקריאה מן העבר השני "האלו?". ומר ישראלי משיב, "הו, כבר סבור הייתי שלא תעני...".

בדוגמא זו ביצע מר ישראלי הערכה ראשונית להסתברותו של מאורע, במקרה זה ההסתברות לשמוע את קולו של בן-שיחו. כעבור רגעים מספר נוסף לדובר מידע - השמיע הטלפון צליל חיוג ראשון. בעקבות המידע ביצע הדובר שקלול מחודש של ההסתברות. הוא אמד את **ההסתברות המותנית** של מענה אנושי **בהנתן** צליל חיוג ראשון. כלומר בהנתן מידע רלוונטי לניסוי שאינו מכריע אותו באופן חד-משמעי. אותה התופעה חוזרה על עצמה עם כל צליל חיוג כך שלמרות שבתחילה סבור היה שיענה בן שיחו, כשהדובר אומנם קרה לבסוף היה הדובר מופתע.

הבה נבחן דוגמא פשוטה יותר לניתוח פורמלי הניתנת לתיאור באמצעות המשוגים שכבר למדנו.

דוגמא 3.1 (סכום קוביות מותנה).

שתי קוביות הוגנות מוטלות. מה היא ההסתברות שסכומן הוא 8?

אם נודע לנו שתוצאת הקוביה הראשונה היא 3, מה כעת הסיכוי שהסכום יהיה 8? (I)

ואם תחת זאת נודע לנו שתוצאת לפחות אחת הקוביות היא 3, מה הסיכוי שהסכום יהיה 8? (II)

תשובה: לטובת הסעיף הראשון נתאר את מרחב ההסתברות של הניסוי באמצעות חזקת מרחב ההסתברות של הטלת קוביה בודדת שתואר בדוגמא (2.8). זהו מרחב הסתברות אחיד עבור $\Omega = [6]^2$. את המאורע "סכום הקוביות יהיה 8" נסמן

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

לפיכך $\mathbb{P}(A) = |A|/|\Omega| = 5/36$. לצורך החלק השני נסתכל על מרחב הסתברות של הטלת קוביה בודדת ונבקש שתוצאתה תהיה 5 (שהלא תוצאת הקוביה השניה ידוע לנו וערכה 3). זו אפשרות אחת מתוך שש במרחב אחיד ולכן ההסתברות היא $\frac{1}{6}$. אך ישנה גם דרך אחרת להסתכל על בעיה זו. נוכל להסתכל על Ω, מרחב ההסתברות שהגדרנו קודם, ולצמצם את פונקציית ההסתברות שלו למאורעות שעדיין עשויים להתרחש. נסמן $S = \{(3,i) \ : \ i \in [6]\}$ כעת ננסה להגדיר פונקציית הסתברות מתאימה $\mathbb{P}^*$ שתייצג את הסתברותם של המאורעות השונים לאחר שנודע לנו שתוצאת הקוביה הראשונה היא 3. היות שהמידע שברשותנו לא משפיע על יחס ההסתברויות של מאורעות המוכלים ב-S_1, והיות ששומא עלינו להגדיר $\mathbb{P}^*(S_1) = 1$, הרי לכל $\omega \in S_1$ יהא עלינו להגדיר $\mathbb{P}^*(\{\omega\}) = \mathbb{P}(\{\omega\})/\mathbb{P}(S_1)$. ואילו לכל $\omega \notin S_1$ עלינו להגדיר $\mathbb{P}^*(\{\omega\}) = 0$ (הלא כבר נודע לנו שמאורעות אלו לא התרחשו). כעת נחשב את הסתברות המאורע A לפי $\mathbb{P}^*$ במרחב ההסתברות $(\Omega, 2^\Omega, \mathbb{P}^*)$ והיות שגודלו של המרחב הוא 6 אומנם נקבל $\frac{1}{6}$. אומנם זו דרך סבוכה יותר מזו שהוצגה קודם, אך היא תאפשר לנו להתמודד גם עם הסעיף השלישי. לשם כך נגדיר

$$S_2 = \{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),(1,3),(2,3),(4,3),(5,3),(6,3)\}\,.$$

שוב נגדיר פונקציית הסתברות אחידה על S_2 המעניקה ערך 0 לכל איבר ב-$\Omega \setminus S_2$. אנו מעוניינים במאורע $\{(3,5),(5,3)\}$ והיות ש-$|S_2| = 11$, נסיק שההסתברות המבוקשת הנה 2/11.

את הדוגמא נכליל באמצעות ההגדרה הבאה.

נשתמש בטבלה שמתארת את מרחב המדגם!

קוביה II \ קוביה 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

בשאלה I רק העמודה 3 מעניינת לכן $\frac{1}{6}$

בשאלה II העמודה והשורה מעניינות אותי לכן $\frac{2}{11}$

הגדרה 3.2. יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהי $B \in \mathcal{F}$ מאורע בעל הסתברות חיובית. לכל מאורע $A \in \mathcal{F}$, נגדיר את **ההסתברות המותנית של** A **בהנתן** B על ידי

$$\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

הערה: את מושג ההסתברות המותנית ניתן לפרש בצורות שונות, כתלות בנקודת ההשקפה על תוקפה של תורת-ההסתברות. מנקודת מבט שכיחותנית ההסתברות המותנית מתארת את מספר המקרים שבהם ארע מאורע A מתוך כלל המקרים שבהם ארע מאורע B. מנקודת מבט הכרתית היא מייצגת את מידת הוודאות של התרחשות A לאחר שנודע לנו ש-B התרחש. מנקודת השקפה מהותית הרי היא מיצגת את אחוז העתידים האפשריים שבהם A התרחשה מאותה נקודה בזמן שבה B כבר ארע. כמובן שהטיפול הפורמלי במונח, אינו מושפע מההשקפה אותה נבחר לאמץ.

ראשית נראה כי התניה תמיד יוצרת פונקציית הסתברות מוגדרת היטב על Ω, מרחב המדגם של הניסוי.

טענה 3.3 (הסתברות מותנית היא הסתברות). יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהי $B \in \mathcal{F}$ מאורע בעל הסתברות חיובית. נגדיר $\mathbb{P}_B(A) = \mathbb{P}(A \mid B)$. אזי $\mathbb{P}_B$ היא פונקציית הסתברות על Ω.

הוכחה. עלינו לוודא כי $\mathbb{P}_B$ מוגדרת היטב ומקיימת את תכונות פונקציית ההסתברות, הגדרה 2.3. מתכונת המונוטוניות (2.4ג) נקבל כי $\mathbb{P}(A \cap B) \le \mathbb{P}(B)$ כך ש $\mathbb{P}_B$ היא אכן פונקציה מ-$\mathcal{F}$ לקטע $[0,1]$. תכונה (2.3א) נובעת מיידית מכך ש

$$\mathbb{P}_B(\Omega) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = 1.$$

נותר אפוא להראות סכימות בת מניה (הגדרה 2.3ב). תהי $(A_n)_{n \in \mathbb{N}}$ סדרת מאורעות זרים ב-$\mathcal{F}$. קל לראות שהמאורעות $(A_n \cap B)$ אף הם זרים ולכן, לפי תכונה (2.3ב) של $\mathbb{P}$, נקבל

$$\mathbb{P}_B\left(\dot{\bigcup_n} A_n\right) = \frac{\mathbb{P}((\dot{\bigcup}_n A_n) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\dot{\bigcup}_n (A_n \cap B))}{\mathbb{P}(B)}$$
$$= \frac{1}{\mathbb{P}(B)} \sum_n \mathbb{P}(A_n \cap B) = \sum_n \mathbb{P}_B(A_n).$$

■

כאשר נתנה ביותר ממאורע אחד, נשתמש לעיתים קרובות בסימון $\mathbb{P}(D \mid A, B) = \mathbb{P}(D \mid A \cap B)$.

אבחנה 3.4 (תכונות בסיסיות של הסתברות מותנית). יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהיו $A, B, D \in \mathcal{F}$ מאורעות בעלי הסתברות חיובית. אזי

(א) אם $A \subset B$ אזי $\mathbb{P}(D \mid A) = \mathbb{P}(D \mid A, B)$.

(ב) **כלל השרשרת (chain-rule):** $\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A \mid B)$

(ג) **התניה חוזרת:** $\mathbb{P}_D(A \mid B) = \mathbb{P}(A \mid B, D) = \frac{\mathbb{P}(A \cap B \cap D)}{\mathbb{P}(B \cap D)}$

הוכחות בעמוד הבא

(א) ממש בס'ע' כי

$$P(D\mid A,B)=\frac{P(A\cap B\cap D)}{P(D)}\overset{A\subset B}{=}\frac{P(A\cap D)}{P(D)}=P(D\mid A)$$

(ב) כלל השרשרת: נובע מהגדרה:

$$P(A\mid B)=\frac{P(A\cap B)}{P(B)}\iff P(B)\cdot P(A\mid B)=P(A\cap B)$$

(ג) נובע מהגדרה

דוגמא 3.5 (הסתברות מותנית).

(א) בכד 5 כדורים צהובים, 10 לבנים ו-10 שחורים. כדור נשלף באקראי, מה ההסתברות שהוא צהוב? מה ההסתברות שהוא צהוב בהינתן שאינו שחור?

תשובה: נשתמש במרחב ההסתברות האחיד על $\Omega = [25]$, כאשר הכדורים הצהובים מיוצגים על ידי המספרים 1-5, הלבנים על ידי 6-15, והשחורים על ידי 16-25. המאורע המבוקש הוא $A = [5]$ ומתוקף אחידות המרחב הסתברותו $5/25 = 1/5$. כעת נתנה במאורע $B = [15]$ ונקבל

$$\mathbb{P}(A \mid B) = \frac{|A \cap B|}{|B|} = \frac{|A|}{|B|} = \frac{5}{15} = \frac{1}{3}.$$

הסבר ממוקד הבא

(ב) גיל ידוע כקלע מוכשר הפוגע במטרה בהסתברות 70% בכל ירייה, לעומתו דוד אלרואי קולע רק בהסתברות 10%. בתחרות קליעה בין השניים נבחר הקלע הראשון באקראי באמצעות הטלת מטבע הוגן וירה קליע. בהינתן שהקליע פגע במטרה מה ההסתברות שהוא נורה על ידי גיל?

תשובה: נשתמש במרחב ההסתברות $\Omega = \{\text{גיל}, \text{דוד}\} \times \{\text{החטיא}, \text{פגע}\}$. המתאר מי ניסה לקלוע ומה הייתה התוצאה. את ההסתברויות המתאימות קיבלנו במונחים של הסתברות מותנית -

$$\mathbb{P}\Big(\{(\text{דוד}, \cdot)\}\Big) = 0.5 \qquad \mathbb{P}\Big(\{(\text{גיל}, \cdot)\}\Big) = 0.5$$

$$\mathbb{P}\Big(\{(\cdot, \text{פגע})\} \,\Big|\, \{(\text{דוד}, \cdot)\}\Big) = 0.1 \qquad \mathbb{P}\Big(\{(\cdot, \text{פגע})\} \,\Big|\, \{(\text{גיל}, \cdot)\}\Big) = 0.7$$

ולפי כלל השרשרת, נקבל

$$\mathbb{P}\Big(\{(\text{גיל}, \text{פגע})\}\Big) = \mathbb{P}\Big(\{(\text{גיל}, \cdot)\}\Big)\mathbb{P}\Big(\{(\cdot, \text{פגע})\} \,\Big|\, \{(\text{גיל}, \cdot)\}\Big) = 0.5 \cdot 0.7 = 0.35$$

באופן דומה נחשב.

$$\mathbb{P}\Big(\{(\text{גיל}, \text{החטיא})\}\Big) = 0.15, \quad \mathbb{P}\Big(\{(\text{דוד}, \text{החטיא})\}\Big) = 0.45, \quad \mathbb{P}\Big(\{(\text{דוד}, \text{פגע})\}\Big) = 0.05.$$

כעת נשתמש בנוסחא להסתברות מותנית ונקבל

$$\mathbb{P}\Big(\{(\text{גיל}, \cdot)\} \,\Big|\, \{(\cdot, \text{פגע})\}\Big) = \frac{\mathbb{P}\Big(\{(\text{גיל}, \text{פגע})\}\Big)}{\mathbb{P}\Big(\{(\cdot, \text{פגע})\}\Big)} = \frac{0.35}{0.35 + 0.5} = 7/8$$

כלומר בהסתברות של שמונים ושבעה אחוזים וחצי – היורה היה גיל.

(ג) נשוב לדוגמא המילולית שפתחה את הפרק. נניח כי מר ישראלי מתקשר לבתו בתאל. ההסתברות שתקלה תמנע תקשורת היא 0.1. אחרת ההסתברות שבת-אל תענה מיד לאחר הצלצול ה-n נתונה על ידי הטבלה הבאה:

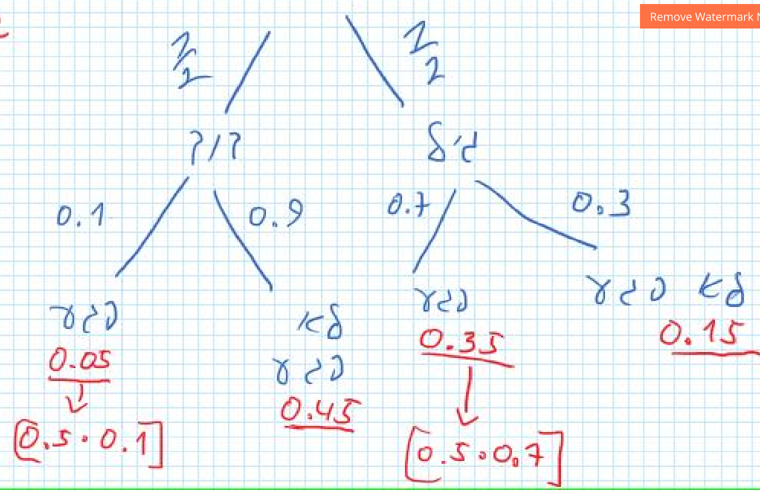
צלצול:	1	2	3	4	לא תענה
הסתברות:	0.05	0.3	0.3	0.05	0.2

חשב את ההסתברות שבתאל תענה בהינתן שנשמעו n צלצולים עבור $n = 0, 1, 2, 3, 4$.

תשובה: נגדיר מרחב הסתברות בן חמישה מצבים שיסומנו על ידי

$$\Omega = \{(F, 0), (F, 4), (T, 1), (T, 2), (T, 3), (T, 4)\}$$

שימו לב כי לא עם טבלה
של מספרי מקרים בודדים
למרות שיש כאן כלל בין ההסתברויות



$$P(\text{גיל} \mid \text{פגע}) = \frac{P(\text{גיל} \cap \text{פגע})}{P(\text{פגע})} = \frac{0.35}{0.05 + 0.35} = \frac{7}{8}$$

האות הראשונה בכל צמד תייצג את התשובה לשאלה האם בת-אל ענתה או לא, ואילו המספר ייצג את מספר הצלצולים שנשמעו. פונקציית ההסתברות המתאימה תהיה זו המקיימת

$$\mathbb{P}(\{(F,0)\}) = 0.1 \quad \mathbb{P}(\{(T,1)\}) = 0.05 \quad \mathbb{P}(\{(T,2)\}) = 0.3$$
$$\mathbb{P}(\{(F,4)\}) = 0.2 \quad \mathbb{P}(\{(T,3)\}) = 0.3 \quad \mathbb{P}(\{(T,4)\}) = 0.05$$

נסמן $A = \{(a,b) \in \Omega \ : \ a = T\}$ וכן $B_n = \{(a,b) \in \Omega \ : \ b \geq n\}$. נציב בנוסחא להסתברות מותנית ונקבל

מאורע	$A\cap B_0$	$A\cap B_1$	$A\cap B_2$	$A\cap B_3$	$A\cap B_4$	B_0	B_1	B_2	B_3	B_4
הסתברות:	0.7	0.7	0.65	0.35	0.05	1	0.9	0.85	0.55	0.25

נחלק ונקבל

מאורע	$A\|B_0$	$A\|B_1$	$A\|B_2$	$A\|B_3$	$A\|B_4$
הסתברות:	7/10	7/9	13/17	7/11	1/5

קל לראות כיצד טבלה זו מסבירה את שינויי הערכת ההסתברות לקבלת תשובה לאורך המתנתו של מר ישראלי למענה.

כפי שראינו בדוגמא 3.5ב, הסתברות מותנית משמשת אותנו לא רק לניתוח ניסויים אלא גם לתיאורם. לעיתים קרובות אנו מיטיבים להעריך את הסתברותו של מאורע מותנה בהשוואה להסתברותו המוחלטת. הדבר נכון במיוחד בניסויים רב-שלביים.

✎ **בעיה 3.1.** יהיה $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהיו $A, B \in \mathcal{F}$ זוג מאורעות בעלי הסתברות חיובית. יש להראות כי אם $\mathbb{P}(A) < \mathbb{P}(A \mid B)$ אז $\mathbb{P}(A) > \mathbb{P}(A \mid B^c)$. מה משמעותה של קביעה זו?

✎ **בעיה 3.2.** יהיו A ו-B שני מאורעות בעלי הסתברות חיובית במרחב הסתברות כלשהו. נאמר כי A **מאשש** את B אם $\mathbb{P}(B \mid A) > P(B)$. האם לכל A, B, C, מאורעות בעלי הסתברות חיובית. יש להראות כי אישוש אינה תכונה עוברת [טרנזיטיבית]. כלומר כי קיימים מאורעות A, B ו-C במרחב הסתברות כך ש-A מאשש את B ו-B מאשש את C אך A אינו מאשש את C.

✎ **בעיה 3.3** (כלל השרשרת - הכללה). השוויון הבא מתקיים לכל שלושה מאורעות A, B ו-C המקיימים $\mathbb{P}(B) > 0$ ו-$\mathbb{P}(B \cap C) > 0$.

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \mid B, C)\,\mathbb{P}(B \mid C)\,P(C).$$

כיצד נוכיח את נכונותו ונכלילו למספר רב יותר של איברים?

3.2 נוסחת ההסתברות השלמה וכלל בייס

נשווה בנפשנו את המצב הבא:

בשניית הסיום של משחק כדורסל מכריע, כאשר התוצאה 55:56 לטובת רוסיה, ביצע שחקן נבחרת רוסיה עבירה על נשרקה, שחקן נבחרת ישראל . נשרקה עולה לקו העונשים לשתי קליעות עונשין. אם יקלע את שתיהן - תנצח ישראל, אם יקלע אחת מהן - יגרר המשחק להערכה שבה סיכוי ישראל לנצח הם 20%. אם יחטיא את שתיהן תנצח רוסיה. לגרשון 75% לקלוע מקו העונשין בכל קליעה באופן בלתי-תלוי.

מה סיכוייה של ישראל לנצח? לשם כך נרצה לבטא את סיכויי המאורע "ישראל ניצחה", באמצעות סיכויי הקליעה של גרשון וסיכויי הנבחרת לנצח בהארכה. ביטוי זה יגולם ב**נוסחת ההסתברות השלמה**.

הגדרה 3.6 (חלוקה). יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. סדרה בת מניה של מאורעות $(A_n)_{n\in\mathbb{N}}$ תיקרא **חלוקה** של Ω אם $A_i \cap A_j = \emptyset$ לכל $i, j \in \mathbb{N}$. כלומר $(A_n)_{n\in\mathbb{N}}$ היא סדרה של קבוצות זרות ב-$\mathcal{F}$ שאיחודן הוא כל Ω (במילים אחרות - כל איבר באומגה נמצא בקבוצה אחת ויחידה בחלוקה).

טענה 3.7 (נוסחת ההסתברות השלמה). תהי $(A_n)_{n\in\mathbb{N}}$ חלוקה של מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. לכל מאורע B מתקיים

$$\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{\infty} \mathbb{P}(B \mid A_i)\mathbb{P}(A_i).$$

כאשר אנו מפרשים איברים באגף ימין כשווים ל-0 אם $\mathbb{P}(A_i) = 0$.

הוכחה. כל מאורע B נוכל לפרק לפי

$$B = \dot{\bigcup_{i=1}^{\infty}} (B \cap A_i).$$

ולכן, בזכות תכונת הסכימות בת-המניה (ב2.3) נקבל

$$\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{\infty} \mathbb{P}(B \mid A_i)\mathbb{P}(A_i).$$

■

נשים לב שנוסחת ההסתברות השלמה פועלת גם בחלוקות סופיות. כלי זה מאפשר לנו לבצע למידת מקרים פרטיים בהקשר הסתברותי, כלומר לפרק את מרחב ההסתברות למקרים שבהם התרחש כל אחד מקבוצת מאורעות משלימים.

נשוב לדוגמא מתחילת הפרק.

דוגמא 3.8 (נשרקה - נוסחת ההסתברות השלמה).

נגדיר מרחב הסתברות - $\Omega = \{W, L\} \times [2]$ המתאר האם ישראל ניצחה או הפסידה במשחק הכדורסל, וכמה קליעות עונשין קלע נשרקה. ידוע כי פונקציית ההסתברות על מרחב זה מקיימת:

$$\mathbb{P}(\{(\cdot, 0)\}) = \left(\frac{1}{4}\right)^2 \qquad \mathbb{P}(\{(\cdot, 1)\}) = 2\left(\frac{3}{4}\right)\left(\frac{1}{4}\right) \qquad \mathbb{P}(\{(\cdot, 2)\}) = \left(\frac{3}{4}\right)^2$$

$$\mathbb{P}(\{(W, \cdot)\} \mid \{(\cdot, 0)\}) = 0 \qquad \mathbb{P}(\{(W, \cdot)\} \mid \{(\cdot, 1)\}) = \frac{1}{5} \qquad \mathbb{P}(\{(W, \cdot)\} \mid \{(\cdot, 2)\}) = 1$$

כיצד נחשב את $\mathbb{P}(\{(W, \cdot)\}$.

תשובה: נציב בנוסחת ההסתברות השלמה ונקבל

$$\begin{aligned}\mathbb{P}(\{(W,\cdot)\} &= \sum_{k\in\{0,1,2\}} \mathbb{P}(\{(W,\cdot)\} \mid \{(\cdot,k)\})\mathbb{P}(\{(\cdot,k)\}) \\ &= \left(\frac{1}{4}\right)^2 \cdot 0 + 2\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)\cdot\left(\frac{1}{5}\right) + \left(\frac{3}{4}\right)^2 \cdot 1 = \frac{51}{80}.\end{aligned}$$

בהמשך לדוגמא זו נשווה בנפשנו תסריט נוסף.

בדיוק ברגע המכריע התקלקל מקלט הטלוויזיה והצופה האומלל לא זכה לראות את סופו של המשחק. למחרת מספר לו חברו שישראל אומנם ניצחה.

כעת ירצה הצופה לדעת מה הסיכוי שנשרקה קלע פעם אחת בדיוק. בכדי לענות על שאלה זו - עומד לרשותנו **כלל בייס** המכונה גם **נוסחת ההיפוך**, והוא קושר בין ההסתברות $\mathbb{P}(A \mid B)$ להסתברות $\mathbb{P}(B \mid A)$.

טענה 3.9 (כלל בייס [נוסחת ההיפוך]). יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. ויהיו $A, B \in \mathcal{F}$ שני מאורעות. אזי

$$\mathbb{P}(B)\mathbb{P}(A \mid B) = \mathbb{P}(B \mid A)\mathbb{P}(A).$$

כאשר אנו מפרשים את אגף ימין כשווה ל-0 אם $\mathbb{P}(B) = 0$ ואת אגף שמאל כשווה ל-0 אם $\mathbb{P}(A) = 0$.

הוכחה. הוכחת המשפט מיידית מהגדרת ההסתברות המותנית:

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

■

כאשר מדובר במאורעות בעלי הסתברות חיובית, נוח לרשום את כלל בייס באופן הבא:

$$\mathbb{P}(A \mid B) = \mathbb{P}(B \mid A)\frac{\mathbb{P}(A)}{\mathbb{P}(B)}.$$

תוֹמָאס בֵּייס (1701-1761), מתמטיקאי אנגלי וכומר פרסבטריאני אשר למד את תורת ההסתברות על סמך ספרו של אברהם דה-מואבר "תורת המקריות" שיצא ב-1718. הערותיו לספר שנתפרסמו לאחר מותו תחת הכותרת "מאמר החותר לפתרון בעיה בתורת הסיכויים" (An Essay towards solving a Problem in the Doctrine of Chances). פורשו על ידי פייר-סימון לפלס ככלי מתמטי לעדכון הסתברותה של תיאוריה נוכח עדויות אמפיריות והוא החל מבסס תורת הסתברות סובייקטיבית. בשנות החמישים של המאה העשרים היוותה הנוסחא כלי מרכזי בפיתוח תורת ההסקה הבייסיאנית המאפשרות להכריע בין השערות על סמך עדויות. כך למשל אם אנו סוברים שאם השערה H נכונה אזי מאורע B מתרחש בהסתברות מסויימת, ואילו אם השערה H' נכונה הרי הוא מתרחש בהסתברות שונה, הרי שאם נראה את B מתרחש נוכל לאמוד את יחס הסבירויות של שתי ההשערות. ספק רב אם בייס עצמו היה שותף לפרשנות זו, שכן אין לה כל ראיות בכתביו.

נשים לב שמהנוסחא נובע גם כי לפי אבחנה 3.4ג מתקיים (בכל מקרה בו ההתניות מוגדרות היטב)

$$\mathbb{P}(A \mid B, D) = \mathbb{P}_D(A \mid B) = \mathbb{P}_D(B \mid A)\frac{\mathbb{P}_D(A)}{\mathbb{P}_D(B)} = \mathbb{P}(B \mid A, D)\frac{\mathbb{P}(A \mid D)}{\mathbb{P}(B \mid D)},$$

נייש כלי זה במצב המתואר בדוגמא 3.8.

דוגמא 3.10 (נשרקה - נוסחת בייס).

הפעם נחשב $\mathbb{P}(\{(\cdot,1)\} \mid \{(W,\cdot)\})$, כלומר את ההסתברות שנשרקה קלע בדיוק פעם אחת בהינתן שישראל ניצחה. נציב בנוסחת בייס ונקבל

$$\mathbb{P}(\{(\cdot,1)\} \mid \{(W,\cdot)\}) = \mathbb{P}(\{(W,\cdot)\} \mid \{(\cdot,1)\})\frac{\mathbb{P}(\{(\cdot,1)\})}{\mathbb{P}(\{(W,\cdot)\})} = \frac{1}{5}\cdot\frac{2\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)}{\frac{51}{80}} = \frac{2}{17}$$

נמחיש כעת את הגישה ההכרתית לכלל בייס בדוגמא ממחזהו של תום סטופארד "רוזנקרנץ וגילדרשטרן מתים". נקודת מבט זו היא בסיס הגישה הבייסיאנית להסקה סטטיסטית שתידון בפרק הפניה.

דוגמא 3.11 (רוזנקרנץ וגילדרשטרן).

כאשר רוזנקרנץ מוציא מטבע מכיסו, גילדרשטרן מאמין שמדובר במטבע הוגן, אך מוכן לקבל כי בהסתברות של אחד למיליון מדובר במטבע מכושף המוציא תמיד את התוצאה עץ. לאחר מכן רוזנקרנץ מתחיל להטיל את המטבע ופעם אחר פעם התוצאה המתקבלת היא עץ. לאחר שחזה גילדרשטרן בשלושים הטלות מטבע, כיצד עליו לעדכן את ההסתברות שהוא מייחס למאורע שמדובר במטבע מכושף?

תשובה: נשתמש במרחב ההסתברות $\Omega = \{M,F\}\times\{H,T\}^{30}$ כאשר הסמל M מיצג את העובדה שהמטבע מכושף ואילו F את העובדה שהוא הוגן. כמו כן נסמן $A = \{M\}\times\{H,T\}^{30}$ ו-$B = \{M,F\}\times\{H\}^{30}$. חיצוני של גולדשטין

להשקפתו של גילדרשטרן $\mathbb{P}(A) = 1/10^6$, $\mathbb{P}(B \mid A) = 1$ ולפי העובדה שבשלושים הטלות מטבע הוגן יש 2^{30} סדרות תוצאות אפשריות שוות הסתברות $\mathbb{P}(B \mid A^c) = 2^{-30}$. נחשב לפי נוסחת ההיפוך ונקבל שההסתברות המבוקשת היא

(יצא עץ שתיין 30 הטלות) $= B$ $\quad P(B|A) = 1$
זה נקרא לכן 1
כי במטבע יקוץ
שיצא עץ

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B \mid A)\mathbb{P}(A) + \mathbb{P}(B \mid A^c)\mathbb{P}(A^c)} = \frac{\frac{1}{10^6}}{\frac{1}{10^6} + \frac{(10^6-1)2^{-30}}{10^6}} \approx 1 - \frac{1}{1073}$$

נוסחת הסתברות השלמה

עתה נמחיש בדוגמאות מספר עד כמה נוח לשגות בשימוש בהסתברות מותנית, למרות פשוטתה הפורמלית, ועד כמה שימוש שגוי זה מוביל לשגיאות ולכשלים לוגיים.

דוגמא 3.12 (בדיקות סטטיסטיות לתופעה נדירה).

אדם מקרי ניגש לבדוק האם הוא נשא של וירוס מסוכן. ידוע שסיכויו של נבדק מקרי הוא כ-1% לשאת את המחלה. בדיקה חדשה שפותחה מזהה טועה בזיהוי אדם שאינו נשא בסיכוי α וטועה בזיהוי נשא בסיכוי β. אדם נבדק ונמצא נשא.

מה הסיכוי שהוא באמת נושא את הווירוס? חשב בפרט לערכים $\alpha,\beta \in \{0, 0.05\}$.

תשובה: ראשית נגדיר את מרחב המדגם $\Omega = \{\text{לא נשא, נשא}\}\times\{+,-\}$ ונסמן $A = \{\text{נשא}\}\times\{+,-\}$ ו-$B = \{\text{לא נשא, נשא}\}\times\{+\}$.

נתון לנו כי $\mathbb{P}(B^c \mid A) = \alpha$, $\mathbb{P}(B \mid A^c) = \beta$ ו- $\mathbb{P}(A) = 0.01$. נחשב מנתונים אלו את $\mathbb{P}(B)$ ומתוכו את $\mathbb{P}(A \mid B)$ – הנתון המבוקש. לפי נוסחת ההסתברות השלמה

$P(B^c|A) + P(B|A) = 1$
$P(B|A^c) + P(B|A) = 1$

$$\mathbb{P}(B) = \mathbb{P}(B \mid A)\mathbb{P}(A) + \mathbb{P}(B \mid A^c)\mathbb{P}(B) = 0.01(1-\alpha) + 0.99\beta$$

ומכאן, לפי כלל בייס, נקבל

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{(1-\alpha) \cdot 0.01}{0.01(1-\alpha) + 0.99\beta}$$

כמובן שאם נציב $\alpha = \beta = 0$ נקבל $\mathbb{P}(A \mid B) = 1$, כלומר תוצאה אידיאלית. לעומת זאת, אם נציב $\alpha = \beta = 0.05$ נקבל

$$\mathbb{P}(A \mid B) = 0.0095/0.059 \approx 1/6,$$

כלומר רק שישית מהנבדקים שנמצאו נשאים באמת נושאים את הוירוס!

תופעה דומה מתרחשת עבור $\alpha = 0$ ו-$\beta = 0.05$. ובמקרה זה נקבל $\mathbb{P}(A \mid B) = 0.01/0.0495 \approx 1/5$. לעומת זאת כאשר $\alpha = 0$ נקבל $\mathbb{P}(A \mid B) = 0.01/0.01 = 1$.

בכדי להיטיב להבין את הקשר בין שכיחות תופעה וחשיבותן של השגיאות השונות הקורא החרוץ מוזמן לחשב את השפעתן של α ו-β על ההסתברות שאדם נשא של הווירוס בהינתן שתוצאת הבדיקה הייתה דווקא שלילית.

דוגמא 3.13 (משפטו של או. ג'יי. סימפסון).
נוסחאת בייס רלוונטית לעיתים קרובות בעולם המשפט, במיוחד בהקשר של כשל התובע (prosecutor's fallacy). נסתכל למשל על הדוגמא המפורסמת הבאה ממשפטו של או. ג'יי. סימפסון ב-1994 בגין רציחתה של אשתו, ניקול בראון. במשפט טענה התביעה שהחשוד המיידי הוא הבעל, מפני שנודע במהלך המשפט שהוא נהג להכותה, והסתברותו של אדם מכה להרוג את אשתו גבוהה יותר. להגנתו טען עורך הדין אלן דרשוביץ כי העובדה הזו כמעט חסרת חשיבות מפני שרק 0.4% מהבעלים המכים רוצחים את נשותיהם.

הבה ננתח שאלה הסתברותית בעלת זיקה למקרה. נסתכל על אדם נשוי מקרי ב-1994. נסמן ב-M את המאורע שאשתו נרצחה, ב-G את המאורע שהוא רצח את אשתו וב-H את המאורע שהוא נהג להכותה. נשים לב כי $G \subset M$. נתונים סטטיסטים מאותה שנה קובעים בקירוב כי $\mathbb{P}(M \mid H) = 1/2250$, ואילו, כפי שנטען, $\mathbb{P}(G \mid H) = 1/2500$.

השאלה שעניינה את המשתתפים במשפט היא הערכת ההסתברות שהאיש אשם בהינתן נתוני הרצח, כלומר $\mathbb{P}(G \mid H, M)$. נחשב:

$$\mathbb{P}(G \mid H, M) \overset{\text{Bayes}}{=} \frac{\mathbb{P}(M \mid G, H) \cdot \mathbb{P}(G \mid H)}{\mathbb{P}(M \mid H)} \overset{G \subset M}{=} \frac{\mathbb{P}(G \mid H)}{\mathbb{P}(M \mid H)} = \frac{1/2500}{1/2250} = 0.9$$

במילים אחרות - אילו היה מדובר באדם מקרי באותה שנה שכל מה שאנו יודעים עליו הוא שהוא הכה את אשתו והיא נרצחה, אזי ישנם 90 אחוזים שהוא הרוצח. גם התביעה וגם ההגנה במקרה זה התעלמו מהנתון הידוע לנו והוא שהאישה אומנם נרצחה.

דוגמא 3.14 (בעיית שני הילדים).
הבעיה הבאה הוצגה על ידי מרטין גארדנר (Martin Gardner) ב-1959 כפרדוקס הסתברותי:

- למר ג'ונס שני ילדים, אשר לפחות אחד מהם הוא בן, מה ההסתברות שגם השני הוא בן?
- למר ג'ונס שני ילדים, הבכור הוא בן, מה ההסתברות שהצעיר אף הוא בן?

בהנחה שמר ג'ונס הוא אדם מקרי בעל שני ילדים, הרי שמרחב ההסתברות הוא המרחב האחיד על ^{2}{בן,בת}. תחת פרשנות זו המאורע הראשון יכול להתפרש כ- $A=\{(b,b),(g,b),(b,g)\}$ ואילו השני כ- $B=\{(b,b),(b,g)\}$ נחשב ונקבל:

$$\mathbb{P}\Big(\{(b,b)\}\,\Big|\,A\Big)=1/3, \qquad \mathbb{P}\Big(\{(b,b)\}\,\Big|\,B\Big)=1/2.$$

אלא שנדמיין את המצב הבא בעולם האמיתי. אנו ניגשים לדלתו של בית ומקישים עליה. פתוח לנו את הדלת נער ואנו שואלים אותו כמה אחים הם. נדמיין שלוש תשובות אפשרויות: "שניים", "שניים, אני הבכור", "שניים, אני הצעיר". לכאורה, לפי הבעיה שפתרנו זה עתה, נוכח התשובה הראשונה עלינו להסיק שבסיכוי שליש הוא אח של בן זכר, ואילו נוכח שתי האחרות - בסיכוי חצי. כעת לכאורה נקבל סתירה לנוסחת ההסתברות השלמה. נסמן $E=$ {הילד שלפנינו בן}, $F=$ {הילד שלפנינו בכור} ו- $M=$ {הילד האחר הוא בן} ונקבל

$$\frac{1}{3}\overset{?}{=}\mathbb{P}(M\mid E)=\mathbb{P}(M\mid F,E)P(F)+\mathbb{P}(M\mid E,F^c)(1-\mathbb{P}(F))\overset{?}{=}\frac{1}{2},$$

היכן הטעות?

תשובה: נוסחת ההסתברות השלמה נכונה, כמובן, אלא שלא התייחסנו כאן לדרך שבה נבחר הילד המשיב לדלת ויצרנו עמימות. אם בהינתן לפחות בן אחד במשפחה ישיב לדלת בן מקרי, באמת נקבל מקרה התואם את המאורע הראשון של מרטין. במקרה זה $\mathbb{P}\Big(M\,\Big|\,E\Big)=\mathbb{P}\Big(\{(b,b)\}\,\Big|\,A\Big)$ ונחשב:

$$\begin{aligned}\frac{1}{3}=\mathbb{P}\Big(\{(b,b)\}\,\Big|\,A\Big)&=\mathbb{P}\Big(\{(b,b)\}\,\Big|\,A,F\Big)(1/2)+\mathbb{P}\Big(\{(b,b)\}\,\Big|\,A,F^c\Big)(1/2)\\&\overset{\text{Bayes}}{=}\mathbb{P}\Big(F\,\Big|\,\{(b,b)\},A\Big)\frac{\mathbb{P}\Big(\{(b,b)\}\,\Big|\,A\Big)}{\mathbb{P}\Big(F\,\Big|\,A\Big)}+\mathbb{P}\Big(F^c\,\Big|\,\{(b,b)\},A\Big)\frac{\mathbb{P}\Big(\{(b,b)\}\,\Big|\,A\Big)}{\mathbb{P}\Big(F^c\,\Big|\,A\Big)}\\&=(1/2)\frac{1/3}{1/2}(1/2)+(1/2)\frac{1/3}{1/2}(1/2)=\frac{1}{3}.\end{aligned}$$

מצד שני אם הילד המשיב נבחר באקראי נקבל מרחב הסתברות אחיד על $\Omega=\{b,g\}^2\times\{1,2\}$ כאשר הקואורדינטה הנוספת מייצגת מי מהילדים השיב לדלת. במקרה כזה: $\mathbb{P}\Big(M\,\Big|\,E\Big)=\mathbb{P}\Big(\{(b,b,\cdot)\}\,\Big|\,\{(x_1,x_2,x_3)\ :\ x_{x_3}=b\}\Big)$ ונקבל

$$\begin{aligned}\frac{1}{2}&=\mathbb{P}\Big(\{(b,b,\cdot)\}\,\Big|\,\{(x_1,x_2,x_3)\ :\ x_{x_3}=b\}\Big)\\&=\mathbb{P}\Big(\{(b,b,1)\}\mid\{(b,\cdot,1)\}\Big)(1/2)+\mathbb{P}\Big(\{(b,b,2)\}\mid\{(\cdot,b,2)\}\Big)(1/2)=\frac{1}{2}.\end{aligned}$$

העמימות לגבי הדרך בה מוכרע איזה מן הילדים פותח את הדלת היא שייצרה את הפרדוקס, כיוון שלא היה נהיר לנו מה המאורע בו אנו מתנים.

דוגמא 3.15 (עשינו עסק [Monty Hall]).

בשעשועון הטלוויזיה "עשינו עסק" ששודר בערוץ השני בשנות התשעים היה נהוג התהליך המקרי הבא. מאחורי וילון שנבחר באופן אחיד מבין שלושה וילונות הייתה מוסתרת מכונית ומאחורי שני האחרים - דחליל.

המשתתף התבקש לבחור וילון. לאחר מכן היה מנחה השעשועון חושף באקראי את אחד הוילונות שלא בחר בהם המשתתף מבין אלו שמאחוריהם נמצא בדחליל. אז ניתנה למשתתף אפשרות לבחור אם ברצונו להחליף וילון. האם מבחינה הסתברותית כדאי למשתתף להחליף את הוילון אם ברצונו לגלות מאחוריו מכונית?

תשובה: בלי הגבלת הכלליות, מטעמי סימטריה, נוכל להניח שהמשתתף בוחר תמיד את וילון מספר אחד. נשתמש במרחב ההסתברות האחיד על $\Omega = [3]$. קל לראות שזכיה ללא החלפה מתוארת על ידי המאורע $A = \{1\}$ ולכן הסתברותה שליש. מה באשר לזכיה לאחר ביצוע החלפה? נשנה את מרחב ההסתברות להיות $\Omega = \{(1,2),(1,3),(2,3),(3,2)\}$ כאשר האיבר הראשון בזוג יתאר את הוילון מאחוריו המכונית והשני - את הוילון שנפתח. הסתברותם של המאורעות כפי שהיא מתוארת היא

$$\mathbb{P}\Big(\{(1,2)\}\Big) = \mathbb{P}\Big(\{(1,3)\}\Big) = \mathbb{P}\Big(\{(1,\cdot)\}\Big)\mathbb{P}\Big(\{(1,3)\} \mid \{(1,\cdot)\}\Big) = \frac{1}{3}\frac{1}{2} = \frac{1}{6}$$
$$\mathbb{P}\Big(\{(2,3)\}\Big) = \mathbb{P}\Big(\{(3,2)\}\Big) = \mathbb{P}\Big(\{(3,\cdot)\}\Big)\mathbb{P}\Big(\{(3,2)\} \mid \{(3,\cdot)\}\Big) = \frac{1}{3}$$

הגיוני

$P(\{(1,\cdot)\}) = 1/3$

$P(\{1,3\} \mid \{1,\cdot\}) =$

אני יודע שבחרתי 1

צב את שירה בסוף

$\frac{\{1,3\}}{|\{1,\cdot\}|} = \frac{1}{2}$

כמו כן נשים לב שהחלפה פירושה בחירה בוילון המכונית במאורע $\{(2,3),(3,2)\}$ ולכן הסתברות הזכיה במקרה של החלפה היא $(2/3)$. ואמונם - החלפה נכשלת אך ורק במקרה שבו בחרנו בניחוש ההתחלתי את וילון המכונית. קורא שחושיו מוחים כנגד הפתרון, עשוי להסתייע בכך שידמיין מצב שבו מכונית מסתתרת מאחורי אחד ממאה וילונות, ולאחר שנבחר אחד מהם, המנחה פותח תשעים ושמונה שאינם מכילים מכונית. במקרה כזה קל לראות שקלושים הסיכויים שהוילון שנבחר בתחילה הוא הנכון, ואפוא למנחה אין ברירה אלא להותיר רק את הוילון מאחוריו מסתתרת המכונית.

3.3 אי-תלות

בפרק הקודם עסקנו בהשפעת ההדדית של מאורעות אלה על אלה. ראינו שלעיתים בהינתן מאורע מסויים, הסתברותו של מאורע אחר משתנה והיא עשויה לקטון או לגדול, ואולם לא תמיד זהו המצב. נסתכל על השאלה הבאה:

"גד ויוקי מטילים כל אחד מטבע. בהינתן שליוקי יצא עץ, מה ההסתברות שגם לגד יצא עץ?"

ההגיון הפיסיקלי גורס שכיוון שלא התרחשו כל יחסי-גומלין בין המטבעות, הרי שאין כל חשיבות לתוצאת ההטלה של יוקי ולכן אם נסמן את המאורע - גד הטיל עץ ב-A ואת המאורע יוקי הטילה עץ ב-B, נקבל

$$\mathbb{P}(A \mid B) = \mathbb{P}(A).$$

במקרה כזה נאמר שהמאורעות **בלתי-תלויים**. אולם, הגדרה זו אינה נראית סימטרית, ואילו חוסר יחסי-הגומלין – דווקא כן. נְפַתֵּחַ את ההגדרה ונקבל $\mathbb{P}(A \cap B)/\mathbb{P}(B) = \mathbb{P}(A)$. נכפיל ב- $\mathbb{P}(B)$ ונקבל את ההגדרה הבאה.

הגדרה 3.16 (אי-תלות)**.** יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. שני מאורעות $A, B \in \mathcal{F}$ יקראו **בלתי-תלויים (ב"ת).** אם

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

אחרת המאורעות יקראו **תלויים**.

נשים לב שלפי טענה 2.15, מאורעות שוליים במרחב מכפלה הנם בלתי תלויים. במובן זה מרחב מכפלה הוא מקרה קיצוני של חוסר תלות. נשוב ונפתח רעיון זה בפרק הפניה.

אבחנה 3.17 (תכונות בסיסיות של אי-תלות). יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהיו $A, B \in \mathcal{F}$ מאורעות ב"ת בעלי הסתברות חיובית, אזי

(א) A בלתי-תלוי ב- Ω וב-$\emptyset$.

(ב) $\mathbb{P}(A \mid B) = \mathbb{P}(A), \mathbb{P}(B \mid A) = \mathbb{P}(B)$.

(ג) $\mathbb{P}(A^c \cap B) = \mathbb{P}(A^c)\mathbb{P}(B)$. כלומר גם A^c ו-B ב"ת.

לצורך הוכחת סעיף ג' נשים לב כי $\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B) = (1 - \mathbb{P}(A))\mathbb{P}(B) = \mathbb{P}(A^c)\mathbb{P}(B)$.

דוגמא 3.18 (ההבדל בין היעדר יחסי-גומלין בין ניסויים לאי-תלות בין מאורעות).

פאוסט בוחר באקראי ובאופן אחיד מספר בין 1 ל-10. נסמן ב-A את המאורע שהמספר קטן מחמש וב-E את המאורע שהוא זוגי. האם A ו-E ב"ת?

תשובה: למרות ששני המאורעות נוגעים לאותו אובייקט, אין הם בהכרח תלויים. נחשב אפוא. מדובר במרחב האחיד על $\Omega = [10]$.

$$\mathbb{P}(A) = \frac{|[4]|}{|[10]|} = \frac{4}{10} \qquad \mathbb{P}(E) = \frac{|\{2,4,6,8,10\}|}{|[10]|} = \frac{1}{2} \qquad \mathbb{P}(A \cap E) = \frac{|\{2,4\}|}{|[10]|} = \frac{2}{10}$$

ומסתבר שהמאורעות אומנם ב"ת.

את מונח האי-תלות ניתן להכליל גם לקבוצה של מאורעות.

הגדרה 3.19 (אי-תלות של אוסף מאורעות). יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. אוסף מאורעות $\mathcal{A} \subset \mathcal{F}$ יקראו **בלתי-תלויים** אם לכל תת-קבוצה סופית שלהם $\{A_n\}_{n\in[N]}$ מתקיים

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \prod_{n\in\mathbb{N}} \mathbb{P}(A_n).$$

✎ **בעיה 3.4.** הוכח כי אם אוסף מאורעות $\mathcal{A} \cup \{B\} \subset \mathcal{F}$ הנם בלתי-תלויים, אזי $\mathcal{A} \cup \{B^c\} \subset \mathcal{F}$

✎ **בעיה 3.5.** הוכח כי אם מאורע A ב"ת בעצמו אזי הסתברותו היא 0 או 1.

✎ **בעיה 3.6.** בדוגמא 2.10 נגדיר, לכל k, את המאורע $A_k = k\mathbb{N} = \{k, 2k, 3k, \ldots\}$.

האם המאורעות $\{A_k\}_{k\in\mathbb{N}}$ הם בלתי תלויים?

האם המאורעות $\{A_k\}_{k\in P}$, כאשר P היא קבוצת כל הראשוניים, הם בלתי תלויים?

טענה 3.20 (מאפייני אי-תלות של אוסף מאורעות). יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהיו $\{A_n\}_{n\in[N]}, B, C$ מאורעות ב"ת. אזי $(B \cap C)$ ו-$\{A_n\}_{n\in[N]}$ ב"ת וכן $(B \cup C)$ ו-$\{A_n\}_{n\in[N]}$ ב"ת.

הוכחה. בכדי לראות כי $(B\cap C)$ ו-$\{A_n\}_{n\in[N]}$ בלתי-תלויים נשים לב כי לפי חוסר התלות של $B, C, \{A_n\}_{n\in[N]}$ מתקיים שלכל תת-קבוצה סופית $\{B_n\}_{n\in[N]}$ של $B, C, \{A_n\}_{n\in[N]}$ מתקיים

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \prod_{n\in\mathbb{N}} \mathbb{P}(B_n)$$

בפרט לכל קבוצה המכילה הן את B והן את C. בכדי לראות כי $(B\cup C)$ ו-$\{A_n\}_{n\in[N]}$ בלתי-תלויים נשתמש בבעיה 3.4 מספר פעמים לקבל כי $B^c, C, \{A_n\}_{n\in[N]}$ בלתי-תלויים, וכן $B^c, C^c, \{A_n\}_{n\in[N]}$ בלתי-תלויים. לפי חלקה הראשון של הטענה נקבל כי $B^c\cap C^c, \{A_n\}_{n\in[N]}$ בלתי-תלויים ולפי החלק הראשון שוב $(B^c\cap C^c)^c, \{A_n\}_{n\in[N]}$ בלתי-תלויים והלא $(B^c\cap C^c)^c = (B\cup C)$. ■

באופן דומה ניתן להראות כי עבור אוסף מאורעות $\mathcal{A}$ בלתי-תלויים, ועבור תת-קבוצה שלו $\mathcal{B}\subset\mathcal{A}$, מתקיים שכל סדרת פעולות של חיתוך ואיחוד של מאורעות ב-$\mathcal{B}$ תיצור מאורע שיהיה בלתי תלוי בכל המאורעות ב-$\mathcal{A}\setminus\mathcal{B}$. מעתה נרשה לעצמנו להשתמש בעובדה זו כמובנת מאליה. בפרק הבא נוכיח טענה דומה בנודע למשתנים מקריים.

טבעי לשאול האם לצורך הוכחת אי-תלות של אוסף מאורעות אנו אומנם נדרשים לבדוק את הסתברותו אל של כל מקבץ סופי של מאורעות. לשם כך נבחן את ההגדרה הבאה.

הגדרה 3.21 (אי-תלות בזוגות). יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות. אוסף מאורעות $\mathcal{A}\subset\mathcal{F}$ יקראו **בלתי-תלויים בזוגות** אם כל זוג מאורעות $A, B\in\mathcal{A}$ הנם ב"ת.

בדוגמא הבאה נראה כי אי-תלות בזוגות **אינה** מספיקה בכדי להוכיח אי-תלות.

דוגמא 3.22 (ההבדל בין אי-תלות לאי-תלות בזוגות).

פלורין מטיל שני מטבעות הוגנים, אשר תוצאות הטלותיהם בלתי-בלויות. נסמן ב-A את המאורע שיצאה בשניהם אותה התוצאה, ב-B_1 את המאורע שהראשון יצא עץ וב-B_2 את ההמאורע שהשני יצא פלי. האם המאורעות האלו בלתי-תלויים? האם הם בלתי-תלויים בזוגות?

תשובה: זהו מרחב אחיד $\Omega = [H,T]^2$.

$$\mathbb{P}(A) = \frac{|\{(H,H),(T,T)\}|}{|\Omega|} = \frac{1}{2} \quad \mathbb{P}(B_1) = \frac{|\{(H,H),(H,T)\}|}{|\Omega|} = \frac{1}{2} \quad \mathbb{P}(B_2) = \frac{|\{(H,T),(T,T)\}|}{|\Omega|} = \frac{1}{2}$$

$$\mathbb{P}(A\cap B_1) = \frac{|\{(H,H)\}|}{|\Omega|} = \frac{1}{4} \qquad \mathbb{P}(A\cap B_2) = \frac{|\{(T,T)\}|}{|\Omega|} = \frac{1}{4} \qquad \mathbb{P}(B_1\cap B_2) = \frac{|\{(H,T)\}|}{|\Omega|} = \frac{1}{4}$$

ומסתבר שהמאורעות אומנם ב"ת בזוגות. ואולם $\mathbb{P}(A\cap B_1\cap B_2) = \mathbb{P}(\emptyset) = 0$, ולכן הם אינם ב"ת.

✎ **בעיה 3.7.** בדוגמא 2.18 האם המאורעות B_{ij} בלתי תלויים בזוגות? האם הם בלתי תלויים?

טענה 3.23 (אי-תלות של מאורעות מכפלה במרחב מכפלה). יהיו $(\Omega_1,\mathcal{F}_1,\mathbb{P}_1)$ ו-$(\Omega_2,\mathcal{F}_2,\mathbb{P}_2)$ מרחבי הסתברות בדידים ויהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב המכפלה שלהם. יהיו $A\in\mathcal{F}_1$ ו-$B\in\mathcal{F}_2$. אזי המאורעות $A\times\Omega_2$ ו-$\Omega_1\times B$ הנם בלתי-תלויים ב-$(\Omega,\mathcal{F},\mathbb{P})$.

אוסף מאורעות

הבדל בין אי תלות כללית לאי תלות בזוגות

נבין בזוגי:

$$\Omega = \begin{Bmatrix} (0,0,0) \\ (0,1,1) \\ (1,0,1) \\ (1,1,0) \end{Bmatrix} \qquad A = \{(1,\cdot,\cdot)\} \qquad B = \{(\cdot,1,\cdot)\}$$

$$C = \{(\cdot,\cdot,1)\}$$

$$P(A\cap B) = 1/4 = P(B\cap C) = P(A\cap C)$$

$$P(A) = P(B) = P(C) = 1/2$$

אבל !! $P(A\cap B\cap C) = 0$ יש כאן אי תלות בזוגות

אין אי תלות כללית

הוכחה. נחשב לפי תכונת מרחב המכפלה (טענה 2.15).

$$\mathbb{P}(A \times \Omega_2) = \mathbb{P}_1(A) \qquad \mathbb{P}(\Omega_1 \times B) = \mathbb{P}_1(B) \qquad \mathbb{P}(A \times B) = \mathbb{P}_1(A)\mathbb{P}_2(B)$$

כנדרש. ■

בעיות הרחבה והעשרה

✎ **בעיה 3.8** (על פי לואיס קרול). בכד אטום נמצא כדור אשר בהסתברות חצי הנו שחור ובהסתברות חצי לבן. הכניסו לכד כדור לבן ולאחר מכן הוציאו ממנו כדור מקרי. מה ההסתברות שהכדור שנותר בפנים לבן?

✎ **בעיה 3.9.** שלושה שופטים מכריעים את גורלו של נאשם על פי דעת רוב. שניים מהשופטים מנוסים ומזהים נכונה את אשמתו של הנאשם בסיכוי 90%. האחרון אינו מנוסה ומזהה נכונה את אשמתו של הנאשם בסיכוי 51%. השווה בין ההסתברויות לזיהוי נכון במקרים הבאים:

(א) החלטת כל שופט בלתי-תלויה

(ב) השופטים המנוסים מחליטים באופן בלתי-תלוי והשופט שאינו מנוסה בוחר אחד מהם באקראי ומצטרף לדעתו

✎ **בעיה 3.10.** על שולחן שלושה כדים אטומים אשר בכל אחד כדור אדום ושני כדורים שחורים. מעבירים כדור מקרי מכד א' לכד ב'. לאחר מכן מעבירים כדור מקרי מכד ב' לכד ג' ולאחר מכן מעבירים כדור מקרי מכד ג' לכד א'. לבסוף מוציאים כדור מכד א'. מה ההסתברות שהכדור שהוצא אדום?

✎ **בעיה 3.11** (פרדוקס האסירים). שלושה אסירים יושבים בבית האסורים. אחד הסוהר מבשר להם כי למחרת היום אחד מהם יוצא להורג והיתר יוצאו לחופשי. באישון לילה אחד האסירים מבקש מהסוהר שיגלה לו את שמו של אחד האסירים האחרים שיצא לחופשי, לטענתו אין בכך מידע לגביו משום שכבר ידוע שלפחות אחד משני האסירים האחרים ישתחרר. הסוהר מסרב מפני שלטענתו לאחר שיגלה את שמו של אסיר כזה - סיכוייו של כל אחד מהנותרים להיהרג יעלה לחצי. בנה מודל הסתברותי והכרע מי צודק.

✎ **בעיה 3.12** (פרדוקס סימפסון). הסיכוי שגלגמש יצוד חיה (צבי או יעל) שהוא מתחקה אחר עקבותיה הוא 80% ואילו סיכוייו של נמרוד לעשות כן הם 70%. זאת למרות שאם נמרוד מתחקה אחר צבי הוא לוכד אותו בהסתברות 1 ואילו גלגמש לוכד צבאים שהתחקה אחר עכבותיהם בהסתברות 7/8, ואילו אם נמרוד מתחקה אחר יעל ההסתברות שילכדה היא 5/8 ואילו הסתברותו של גלגמש לעשות כן היא 1/2. כיצד תסביר (בכלים הסתברותיים) את העובדה שבכל אחד מסוגי הציד בנפרד מצטיין נמרוד ובכל זאת הסתברותו של גלגמש לצוד חיה כלשהי גבוהה יותר?