5 הסתברות מותנית 7-11

5.1 הסתברות מותנה ואי־תלות

נניח והטלנו קובייה וכעת חבר אומר לנו כי התוצאה איננה 1 ואיננה 6 .טרם קבלת המידע, ייחסנו הסתברות של $\frac{1}{6}$ לתוצאה 1 ולתוצאה 6, אבל לאחר קבלת המידע, ההסתברות הללו ירדו ל-0. מצד שני, תחת המידע החדש, מה הסיכוי שהקוביה נחתה על 4 ?או על 5 ?בכדי לענות על השאלה אנו נדרשים לעדכן את מרחב ההסתברות שלנו בהתאם למידע החדש. הסתברות מותנה זו הסתברות אשר עודכנה בהתאם למידע / מאורע ספציפי.

5.1 הגדרה. (הסתברות מותנית) לכל צמד מאורעות A ו־B כך ש־B בעל הסתברות חיובית, ההסתברות של מאורע A בהינתן מאורע B מוגדרת להיות

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \ . \tag{5.1}$$

5.2 דוגמא. בהטלת מטבע ניקח את המאורע $A = \{2\}$ ו $B = \overline{\{1,6\}}$. ברור ש־ $P(A) = \frac{1}{6}$. אבל, אם B מתממש ותוצאת ההטלה איננה 1 או 6 הסיכוי לקבל את התוצאה 2 משתנה. אינטואיציה שלנו מרמזת כי הסיכוי המעודכן יהיה רבע, משיקולי סימטריה בין התוצאות הנותרות. נבחן זאת לפי ההגדרה.

$$P(\{2\}|B) = \frac{P(\{2\}\cap B)}{P(B)} = \frac{P(2)}{\frac{4}{6}} = \frac{\left(\frac{1}{6}\right)}{\left(\frac{4}{6}\right)} = \frac{1}{4},$$

כפי שחזינו.

5.3 דוגמא. בדוגמא הבאה נגדיר את המאורע $A = \{1,2,5,6\}$. ברור ש $P(A) = P(B) = \frac{4}{6}$. לאחר קבלת המידע שהתוצאה איננה 1 או 6 נבחין כי

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(2,5)}{\frac{4}{6}} = \frac{\left(\frac{2}{6}\right)}{\left(\frac{4}{6}\right)} = \frac{1}{2}.$$

זאת אומרת שההסתברות של מאורע A ירדה מ$\frac{2}{3}$ ל $\frac{1}{2}$.

עדכון ההסתברות שביצענו לעיל נקרא **עדכון בייסיאני.**

5.4 דוגמא. בפונטי־פאנדי יש

- 45% גברים,
- 30% מעשנים, ו־
- 15% הם גברים מעשנים.

נבחר אדם באקראי. בהנחה ונבחר גבר,

1. מה הסיכוי שהוא מעשן?
2. מה הסיכוי שאינו מעשן?
3. כיצד התשובות תשתננה בהנחה ונבחרה אישה?

פיתרון. נסמן ב־ A את המאורע שנבחר גבר וב־ B את המאורע שהאדם שנבחר מעשן. באופן כללי, ההסתברויות של A ו־ B הן $P(A) = 0.45$ ו־ $P(B) = 0.3$.

1. נחשב את P שהאדם הנבחר מעשן, תחת ההנחה שמדובר בגבר.

$$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{0.15}{0.45} = \frac{1}{3}.$$

זאת אומרת, שיעור המעשנים באוכלוסיית הגברים גבוה מן הממוצע באוכלוסייה הכוללת.

2. בכדי לבחון את P שאינו מעשן נוכל להשתמש בכלל המשלים, (הסתברות מותנה מקיימת את כל כללי פונקצית ההסתברות). לכן,
$$P(\bar{B}|A) = 1 - P(B|A) = \frac{2}{3}.$$
כנדרש.

3. נשים לב שמדובר בהתנייה שונה ולכן מדובר במרחב שונה לחלוטין. לא ניתן להמיר את התוצאות הקודמות בגלל ששינוי ההתנייה משנה את כל הנחות הבסיס של החישוב. בהתאם לזאת, נצטרך לבצע את החישובים הקודמים בשנית, וכעת נתנה על $\bar{A}$. נשים לב כי
$$B = (B \cap A) \cup (B \cap \bar{A})$$
והאיחוד הוא איחוד זר (של מאורעות זרים). לכן נקבל
$$P(B|\bar{A}) = \frac{P(B \cap \bar{A})}{P(\bar{A})} = \frac{P(B) - P(B \cap A)}{P(\bar{A})} = \frac{0.3 - 0.15}{0.55} \approx 0.27.$$
המעשנות נמוך יותר משיעור המעשנים באוכלוסיה. כמו כן,
$$P(\bar{B}|\bar{A}) = 1 - P(B| \cap A)1 - \frac{0.3 - 0.15}{0.55} \approx 0.73.$$

■

5.5 דוגמא. בכד 10 כדורים הממספרים מ־1 עד 10 .שולפים 4 ללא החזרה.

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

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

1. נגדיר את המאורע ־ "הכדור עם המספר 7 הינו אחד מהכדורים שהוצאו" בתור A .מאחר ומדובר במרחב מדגם שווה הסתברות, אזי
$$P(A) = \frac{1.\binom{9}{3}}{\binom{10}{4}} = \frac{9!6!4!}{10!6!3!} = \frac{4}{10} = \frac{2}{5}\ .$$

2. נסמן ב־ B את המאורע המציין כי הכדור עם המספר 9 הינו אחד מהכדורים שהוצאו.
$$\begin{aligned} P(A|B) =& \frac{P(A \cap B)}{P(B)} &&= \frac{\left(\frac{1.\binom{8}{2}}{\binom{10}{4}}\right)}{\left(\frac{1.\binom{9}{3}}{\binom{10}{4}}\right)} \\ =& \frac{\binom{8}{2}}{\binom{9}{3}} &&= \frac{8!6!3!}{9!6!2!} = \frac{1}{3}. \end{aligned}$$

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

3. נגדיר מאורע C ־ כל המספרים שהוצאו גדולים או שווים ל־4 .

$$P(A|C) = \frac{P(A\cap C)}{P(C)} = \frac{\frac{\binom{6}{3}}{\binom{10}{4}}}{\frac{\binom{7}{4}}{\binom{10}{4}}} = \frac{\binom{6}{3}}{\binom{7}{4}} = \frac{6!4!3!}{7!3!3!} = \frac{4}{7} \ .$$

4. אנו נדרשים לחשב את ההסתברות $P(\bar{A}|C)$. נזכר כי הסתברות מותנה מקיימת את כל הכללים של הסתברות ולכן

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

■

5.2 כלל הכפל

הסתברות מותנה מפשט את החישוב, שכן אנו מסוגלים להתייחס לכל שלב בנפרד תחת ההנחה שהתוצאות המתבקשות בשלבים הקודמים אכן התקיימו. הגרסה הבסיסית של כלל הכפל מתקיימת עבור צמד מאורעות A ו־ B כאשר B בעל הסתברות חיובית ($P(B) > 0$). נעזר בהגדרה להסתברות מותנה [עיין משוואה (5.1)], ועל ידי העברת אגפים נקבל

$$P(A\cap B) = P(A|B)P(B) \ . \tag{5.2}$$

זאת אומרת, הסיכוי ש־ A ו־ B יתרחשו, שווה לסיכוי ש־ B יתרחש, כפול הסיכוי ש־ A יתרחש, תחת הנחה ש־ B אכן קרה. ההרחבה למספר כללי של מאורעות היא כמפורט בחוק הבא:

5.6 חוק. (כלל כפל) לכל קבוצת מאורעות $A_1, \dots, A_n$ כך ש

$$P\left(\bigcap_{i=1}^{n} A_i\right) > 0$$

מתקיים

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)\dots P\left(A_n|\bigcap_{i=1}^{n-1} A_i\right) \ . \tag{5.3}$$

קל מאוד לבדוק מדוע טענה זאת נכונה. כל מה שצריך זה לעבור על האיברים באגף ימין ולהציב את ההגדרה להסתברות מותנה. האיברים המתאימים מצטמצמים ומקבלים חזרה את הביטוי באגף שמאל.

5.7 דוגמא. בכד כדור שחור אחד וכדור לבן אחד. שולפים כדור באקראי. לאחר שמוציאים כדור באקראי מחזירים אותו יחד עם עוד כדור אחד נוסף בעל אותו הצבע. מוציאים סה"כ 5 כדורים. מה P שכולם שחורים?

פיתרון. נגדיר את המאורעות B_i $i = 1, \dots, 5$ כמאורעות בהם הכדור ה־ i הוא שחור. נשתמש בחוק הכפל [עיין משוואה (5.3)]:

$$\begin{aligned} P\left(\bigcap_{i=1}^{5} B_i\right) &= P(B_1)P(B_2|B1)P(B_3|B1\cap B2)\dots P\left(B_5|\bigcap_{i=1}^{4} B_i\right) \\ &= \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}\cdot\frac{5}{6} \\ &= \frac{1}{6} \ . \end{aligned}$$

■

5.3 נוסחת ההסתברות השלמה

איור 1: עץ הסתברות עבור דוגמה הפוליגרף

לשם המחשה ניקח סיטואציה של שימוש בפוליגרף. פוליגרף, מור לבחון האם אנשים הם דוברי אמת או שקר. הבעיה עם מכונת הפוליגרף היא שאיננה מדוייקת דיה, ולכן יש סיכוי מסויים לשקר ולא להיתפס ולהיפך. אנו יודעים כי אדם המשקר נתפס בשקר בסיכוי 80% ואדם הדובר אמת מזוהה כדובר אמת בסיכוי 90% (שימו לב שכל הנתונים הם למעשה הסתברויות מותנות!). הנחת העובדה כי 70% מן האנשים שמגיעים לבדיקת פוליגרף משקרים. מהי ההסתברות שאדם מקרי בבדיקה יימצא כדובר אמת?

נגדיר את המאורעות הבאים.

- T_1: אדם דובר אמת, L_1: אדם דובר שקר, T_2:
  אדם מוזהה כדובר אמת (על ידי הפוליגרף), L_2 אדם מזוהה כדובר שקר.

הנתונים בשאלה הם

$$P(T_2|T_1) = 0.9, \quad P(L_2|L_1) = 0.8, \quad P(L_1) = 0.7 \ .$$

אנחנו מעוניינים לחשב את ההסתברות שאדם יימצא דובר אמת, קרי $P(T_2)$. האיור לעיל מציג את הבעיה הנתונה כעץ הסתברויות. נעבור על הענפים המתאימים שנגמרים בתוצאה הרצויה בה האדם נמצא דובר אמת, T_2, ונחשב את ההסתברות של כל ענף בפני עצמו. תחילה,

$$P(T_2) = P(T_2 \cap T_1) + P(T_2 \cap L_1) \ .$$

לפי הסתברות מותנה משוואה [עיין נוסחא (5.2)]

$$P(T_1 \cap T_2) = P(T_1)P(T_1|T_1) = 0.3 \times 0.9 = 0.27 \ .$$

נבצע חישוב דומה עבור המחובר השני ונקבל

$$P(T_2 \cap L_1) = P(L_1)P(T_2|L_1) = 0.7 \times 0.2 = 0.14 \ .$$

כעת נסכום את צמד התוצאות יחד ונקבל

$$P(T_2) = P(T_2|T_1)P(T_1) + P(T_2|L_1)P(L_1) = 0.9 \times 0.3 + 0.2 \times 0.7 = 0.41 \ .$$

שימו לב כי הסיכוי שאדם יוכרז כדובר אמת הוא 41% בלבד!

5.8 חוק. (נוסחת ההסתברות השלמה) נניח כי המאורעות $B_1, \dots, B_n$ הם זרים בזוגות, בעלי הסתברות חיובית כ"א, ומחלקים את מרחב המדגם באופן מלא. ז"א

$$B_i \cap B_j = \phi$$

לכל $i \neq j$ ובנוסף

$$\bigcup_{i=1}^{n} B_i = \Omega \ ,$$

כאשר $P(B_i) > 0$. אזי לכל מאורע

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i) \ . \qquad (5.4)$$

הוכחה. במילים פשוטות, אנחנו רוצים לדעת את ההסתברות למאורע A, בעוד העולם שלנו מחולק למספר סופי של מאורעות $B_1, \dots B_n$. נתייחס לכל אחד מהמאורעות הללו בנפרד. קרי, תחת ההנחה ש־ B_i התרחש, נבדוק את הסבירות של מאורע A, ונכפול בהסתברות ש־ B_i תתרחש. במידה ונעשה זאת עבור כל מאורע B_i בנפרד, נקבל את ההסתברות של A. נפרק את A לחלקים הבאים:

$$A \cap B_1, \quad A \cap B_2, \dots, A \cap B_n \ .$$

האיחוד של כל המאורעות הוא כל מרחב המדגם, מכאן אנו יודעים שלא פספסנו אף אפשרות של A ועל כן

$$A = \bigcup_{i=1}^{n} A \cap B_i.$$

בנוסף, המאורעות הללו זרים בזוגות בגלל ש־ $B_1, \dots B_n$ זרים בזוגות. לכן

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

כנדרש, כאשר במעבר האחרון השתמשנו בכלל הכפל עבור צמד מאורעות. ∎

5.4 חוק Bayes

לפי ההגדרה של הסתברות מותנה

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

ו

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

נכפול את צמד המשוואות במכנה שלהן ונשווה בינהן בכדי להגיע לחוק בייס הבא.

5.9 חוק. (חוק בייס) עבור כל צמד מאורעות A ו B בעלי הסתברות חיובית מתקיים

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{5.5}$$

לשם המחשה, נחזור לדוגמא הקודמת של הפוליגרף.

5.10 דוגמא. כעת ננסה לחשב את ההסתברות שאדם מקרי דובר אמת, תחת ההנחה שהוא נמצא דובר אמת:

$$P(T_1|T_2).$$

הנתונים בשאלה מוצגים בתצורה של 'אם האדם עשה כך וכך, אז זה הסיכוי שהוא יוכרז כך וכך', אלו יחסי הסיבה־תוצאה בניסוי בפועל. אבל השאלה הנתונה הופכת את הסדר. השאלה מעוניינת לקבוע את הסיכוי שאדם דיבר אמת (סיבה) בהינתן הוא נמצא דובר אמת (תוצאה). במקרים הללו חוק בייס הוא יעיל במיוחד. אכן,

$$P(T_1|T_2) = \frac{P(T_2|T_1)P(T_1)}{P(T_2)} = \frac{0.8.0.3}{0.41} \approx 0.59 \ .$$

זאת אומרת, אדם שנמצא דובר אמת אכן אמר אמת בהסתברות 0.59 בלבד !

5.11 דוגמא. הרכב הסטודנטים בתואר ראשון הוא כדלקמן:

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

בנוסף, מצבם המשפחתי של הסטודנטים הנ"ל הוא:

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

ענו על השאלות הבאות.

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

פיתרון. בשאלה זו נשתמש בעיקר בנוסחת ההסתברות השלמה וחוק בייז.

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

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

2. נשתמש בחוק בייז מסיבה פשוטה ־ נתוני השאלה מציגים הסתברויות להיות נשוי בהתאם לשנת הלימוד, אבל ההתנייה ההפוכה לא נתונה. לכן,

$$P(III|M) = \frac{P(M|III)P(III)}{P(M)} = \frac{0.4.0.2}{0.27} = \frac{0.08}{0.27} \approx 0.296\ .$$

3. נחשב באופן דומה לחישוב קודם, אלא שהפעם ההתנייה היא על מאורע שונה.

$$P(III|\bar{M}) = \frac{P(\bar{M}|III)P(III)}{P(\bar{M})} = \frac{0.6.0.2}{1-0.27} = \frac{0.12}{0.73} \approx 0.164\ .$$

■

5.5 אי־תלות בין מאורעות

תלות בין מאורעות הוא מושג שמבטא את מידת הקשר ההסתברותי בין צמד המאורעות. לדוגמא, ניקח חבילת קלפים תיקנית עם 52 קלפים ונשלוף קלף באקראי. כעת עלינו לנחש את הקלף שהוצא ־ צבע וצורה. אפריורית,

$$P(\text{king}) = \frac{1}{13}, \qquad P(\text{diamond}) = \frac{13}{52},$$

כעת נניח שראינו שהקלף שהוצא הוא אדום. כיצד המידע הנוסף השפיע על ההסתברות הקודמות? ובכן,

$$P(\text{diamond}|\text{red}) = \frac{13}{26}$$

ובעוד

$$P(\text{king}|\text{red}) = \frac{2}{26} = \frac{1}{13}\ .$$

זאת אומרת, המידע הנוסף לגבי צבע הקלף השפיע על הסתברות להוציא יהלום, אך לא השפיע על ההסתברות להוציא מלך. במילים אחרות, המאורע 'יהלום' תלוי במאורע 'קלף אדום' (ולהיפך, כמובן), בעוד המאורע 'מלך' אינו תלוי במאורע 'אדום'. ננסח זאת באופן מדוייק.