

הסתברות וסטטיסטיקה לכלכלנים

ד"ר דוד לגזיאל

המחלקה לכלכלה, אוניברסיטת בן-גוריון בנגב

תקציר

חוברת זו נכתבה **ברשימות מרצה**, בהתבסס על חומרי העזר המעולים של ד"ר שירלי שדות, פרופ' אורי חיימנקו, ד"ר רועי טפר, פרופ' יובל הלר, ד"ר אסף כהן, מר ייבגני צ'ודיקוביץ', ופרופ' אילון סולן. ייתכנו אי דיוקים ושגיאות דפוס. אחריות השימוש ברשימות היא על המשתמשים בלבד.



אין לשכפל, להעתיק, לצלם, לתרגם, לאחסן במאגר מידע, לשדר או להקליט בכל דרך או בכל אמצעי אלקטרוני, אופטי, מכני, או אחר כל חלק מהחומר בספר זה. שימוש מסחרי מכל סוג בחומר הכלול בספר זה אסור בהחלט, אלא ברשות מפורשת בכתב מהמחבר.

תוכן עניינים

4	I חלק א' - הסתברות
5	1 יסודות: מרחבי הסתברות
5	1.1 ניסוי, מרחב מדגם, ומאורעות
6	1.2 מבוא לתורת הקבוצות ולוגיקה
7	1.2.1 תורת הקבוצות: פעולות בסיסיות בין מאורעות
10	1.2.2 חוקי דה־מורגן
11	1.2.3 מאורעות זרים
13	1.3 מרחבי הסתברות כלליים
14	1.3.1 חוקי הסתברות בסיסיים
16	1.4 מרחבי הסתברות אחידים
17	1.5 קומבינטוריקה
17	1.5.1 עקרון הכפל
18	1.5.2 מדגמים עם/בלי החזרה, סדורים ולא סדורים
22	2 הסתברות מותנה ואי־תלות
24	2.1 כלל הכפל
25	2.2 נוסחת ההסתברות השלמה
26	2.3 חוק Bayes
29	2.4 אי תלות בין מאורעות
33	3 משתנים מקריים בדידים
33	3.1 תומך והתפלגות
35	3.1.1 פונקציית התפלגות מצטברת
36	3.2 תוחלת
38	3.2.1 לינאריות התוחלת
38	3.2.2 תוחלת פונקציה של מ"מ
40	3.3 שונות
41	3.3.1 תכונות השונות
42	3.4 משפחות התפלגות
43	3.4.1 ברנולי
43	3.4.2 אחידה
44	3.4.3 בינומית
48	3.4.4 גיאומטרית
50	3.4.5 פואסונית
53	4 משתנים מקריים דו־מימדיים
53	4.1 תומך והתפלגות משותפת
54	4.1.1 התפלגויות שוליות
55	4.2 התפלגות מותנה
56	4.3 תוחלת מותנה ונוסחת התוחלת השלמה
58	4.4 אי תלות בין משתנים מקריים

60	שונות משותפת ומקדם מתאם	4.5
61	שונות משותפת	4.5.1
63	מקדם מתאם	4.5.2
65	דוגמאות גרפיות וקו רגרסיה	4.5.3
66	שונות סכום משתנים מקריים	4.5.4
71	משתנים מקריים רציפים	5
71	פונקצית התפלגות מצטברת ופונקצית הצפיפות	5.1
71	פונקצית ההתפלגות המצטברת	5.1.1
72	פונקצית הצפיפות	5.1.2
73	הקשר בין הפה"מ לפונקצית הצפיפות	5.1.3
80	תוחלת ושונות	5.2
81	משפחות התפלגויות רציפות	5.3
81	התפלגות אחידה רציפה	5.3.1
82	התפלגות נורמאלית	5.3.2
89	התפלגות מעריכית	5.3.3

92 II חלק ב' - סטטיסטיקה

93	סטטיסטיקה תיאורית	6
93	הסטטיסטיקה בעולם האמיתי	6.1
94	עיבוד נתונים	6.2
94	סכימה	6.2.1
96	טרנספורמציה לינארית	6.2.2
97	טבלת שכיחויות	6.2.3
97	הצגת הנתונים	6.2.4
99	מדדי מרכז ופיזור	6.3
99	מדדי מרכז	6.3.1
100	מדדי פיזור	6.3.2
102	מדד מיקום יחסי	6.3.3

חלק I

חלק א' - הסתברות

1 יסודות: מרחבי הסתברות

הבסיס לתורת ההסתברות הוא הניסוי המקרי. הניסוי המקרי הוא ניסוי שעושה האדם או הטבע וישנה קבוצה מוגדרת וידועה מראש של תוצאות אפשריות, אבל רק אחת מהן יכולה להתממש בפועל בניסוי יחיד. לדוגמא: הטלת מטבע. אדם המטיל מטבע יודע מראש כי ישנן רק שתי תוצאות אפשריות - "עץ" או "פלי". יחד עם זאת, אותו אדם אינו מסוגל לחזות מבעוד מועד אם התוצאה תהיה "עץ" או "פלי" בפועל. בנוסף, אם יטיל שוב את המטבע, התוצאה בניסוי החדש יכולה לצאת זהה לניסוי הקודם או שונה ממנו, שכן תוצאות עלולות להשתנות בין ניסויים מקריים שונים.

בעולם הכלכלי אנחנו מבצעים ניסויים מקריים רבים, אפילו במצבים פשוטים לכאורה כגון מסחר או משא ומתן. המקריות יכולה להתבטא בפרמטר הזמן - מכירה של נכס נדל"ן או נייר ערך היום או מחר; ויכולה להתבטא בפרמטר המחיר - בקשה של מחיר מסויים או בקשה של מספר אחוזים נוספים. כמובן שישנם עוד אינספור מצבים בהם מקריות באה לידי ביטוי. לדוגמא, בעולם הפיזיקאלי ישנה מקריות בתנועת חלקיקים.

המקור של תורת ההסתברות מגיע מחיי היומיום היותר צבעוניים - משחקי המזל. במאה ה-16, החוקר האיטלקי ג'וראמו קורדנו (Georlamo Cordano, 1501-1576), עסק בנושא משחקי המזל וההימורים (בין יתר מחקריו בענפי המתמטיקה, פיזיקה, ביולוגיה, כימיה, אסטרונומיה וכן הלאה). הבעיות הבסיסיות שעלו מענייני ההימורים, בדגש על הטלות מטבע וקוביה, היוו את הבסיס לתורה ההסתברותית כשם שאנו מכירים אותה כיום. לאחר קורדנו, הלפיד עבר לשני מתמטיקאים צרפתיים ידועים בשם פרמה (Pierre De Fermat, 1607-1665) ופסקל (Blaise Pascal, 1623-1662). צמד המתמטיקאים ביצע חליפת מכתבים ענפה במהלכה התמקדו בבעיות העוסקות במשחקי מזל עם תוצאות סימטריות. התורה הושלמה על ידי החוקר הצרפתי לפלס (Pierre-Simon, marquis de Laplace, 1749-1827) במאה ה-19, ואת גרסתה הנוכחית והגישה האקסיומטית שבבסיסה ניסח אנדריי קולמוגורוב (Andrey Nikolaevich Kolmogorov, 1903-1987) בתחילת המאה ה-20. קולמוגורוב מיסד את תורת ההסתברות בעזרת כללים ברורים שאותם אנו לומדים עד היום.

1.1 ניסוי, מרחב מדגם, ומאורעות

הניתוח הראשוני של כל בעיה הסתברותית מתרכז בתוצאות האפשריות של הניסוי אותו מבצעים. התוצאות הללו מוגדרות תחת קבוצה כללית הנקראת "מרחב המדגם".

הגדרה 1.1 (מרחב מדגם) מרחב מדגם זו קבוצה, המסומנת לרוב באות היוונית Ω (אומגה), המכילה את כל התוצאות האפשריות של הניסוי.¹

נסתכל על מספר ניסויים אשר נזכיר רבות בקורס. בעת הטלת מטבע הוגן כאשר על צד אחד שלו רשומה הספרה '1' ועל הצד השני מופיעה הספרה '0', מרחב המדגם יהיה $\Omega = \{0, 1\}$. מצד שני, אם על המטבע נמצאים הסימונים המוכרים של T ו- H ² אזי מרחב המדגם יהיה $\Omega = \{H, T\}$. בהטלת קוביה סטנדרטית, מרחב המדגם יהיה $\Omega = \{1, 2, 3, 4, 5, 6\}$, ואם נטיל את הקוביה פעמיים, אז מרחב המדגם יהיה מורכב מצמדי מספרים שלמים מ-1 עד 6, $\Omega = \{(i, j) : i, j = 1, 2, 3, 4, 5, 6\}$.

בקורס אנחנו נעסוק במרחבי מדגם סופיים או לכל היותר בני מניה (קרי, מרחבים שניתן לספור כמה איברים יש בהם) ולכן מעתה נניח כי כל מרחבי המדגם שלנו הם לכל היותר בני-מניה.³ לדוגמא, נבצע ניסוי של קליעה לסל עד לפעם הראשונה שנקלע. במקרה כזה מרחב המדגם איננו חסום שכן מספר הזריקות יכול להיות גדול כרצוננו. יחד עם זאת, המרחב הוא בן מניה וניתן להציגו בעזרת כל המספרים הטבעיים $\Omega = \{1, 2, 3, \dots\} = \mathbb{N}$.

הצגה של כל התוצאות האפשריות בניסוי היא החלק הראשון בתיאור הניסוי, אבל ברוב המקרים נתעניין דווקא בקבוצה ספציפית של תוצאות ולא בכולן. לדוגמא, בהטלה של צמד קוביות, ייתכן ונתעניין דווקא בכל המקרים בהם סכום ההטלות הוא 7 או 12 (כפי שקורה בחלק ממשחקי המזל). התיאור של מצבים אלו מבוסס על קבוצה חלקית למרחב המדגם הנקראת **מאורע**.

הגדרה 1.2 (מאורע) לכל אוסף של תוצאות אפשריות בניסוי נקרא מאורע.⁴

¹אנו לרוב מסמנים קבוצה בסוגריים מסולסלות. לדוגמא, הקבוצה שמכילה את הערכים 1, 2, 3 נרשמת כך $\{1, 2, 3\}$.
² H^2 מציין את התוצאה "עץ", בעוד T מתייחס לתוצאה "פלי".
³כמובן שבאופן כללי מרחב מדגם יכול להיות גם מעוצמת הרצף, דוגמת כל הסדרות הבינאריות ועוד, אבל זו כבר סוגיה מתמטית פחות רלוונטית עבורנו.
⁴תת קבוצה היא קבוצה שמכילה חלק מאיברי הקבוצה המקורית או את כל איברי הקבוצה המקורית.

לדוגמא, בהטלת מטבע עם ערכים 0 ו-1 על כל פאה (ערך נפרד בכל פאה), מרחב המדגם הוא $\Omega = \{0, 1\}$, והמאורע $\{0, 1\}$ מתאר את התוצאות בהן המטבע נוחת על צד כלשהו. מנגד, המאורע $\{0\}$ מתאר את המצב בו המטבע נוחת על הצד עליו הספרה 0. במקרה כזה, נקבע כי מאורע $\{0\}$ **התרחש**. בנוסף, אם המטבע נחת על הצד עם הספרה 0 אז גם המאורע $\{0, 1\}$ התרחש, שכן אחת האפשרויות שבו התממשה. אם נחזור לדוגמא הקודמת של הטלת צמד קוביות, אז המאורע בו סכום ההטלות הוא 4 נתון על ידי $\{(1, 3), (2, 2), (3, 1)\}$, בעוד המאורע בו סכום ההטלות שווה ל-12 הוא $\{(6, 6)\}$.

תרגיל 1.3 רשמו את מרחב המדגם והמאורעות המצויינים בהמשך במונחי תורת הקבוצות.

1. הוצאה של כדור מתוך כד בו יש 2 כדורים לבנים וכדור שחור. A - הוצא כדור שחור.
2. הוצאה של כדור מתוך כד בו יש 2 כדורים לבנים הממוספרים ב-1,2 וכדור שחור. A - הוצא כדור לבן.
3. הוצאה עם החזרה של 2 כדורים מתוך כד בו יש 2 כדורים לבנים ו-1 שחור. A - הכדור השני שהוצא איננו לבן.
4. הוצאה ללא החזרה של 2 כדורים מתוך כד בו יש 2 כדורים לבנים ו-1 שחור. A - הכדור השני שהוצא הוא לבן.

פתרון: נענה על כל סעיף בנפרד.

1. מרחב המדגם בהוצאה של כדור מתוך כד בו יש 2 כדורים לבנים וכדור שחור הוא $\Omega = \{W, B\}$. המאורע נתון על ידי $A = \{B\}$.
2. מרחב המדגם בהוצאה של כדור מתוך כד בו יש 2 כדורים לבנים הממוספרים ב-1,2 וכדור שחור הוא $\Omega = \{W_1, W_2, B\}$. המאורע נתון על ידי $A = \{W_1, W_2\}$.
3. מרחב המדגם בהוצאה עם החזרה של 2 כדורים מתוך כד בו יש 2 כדורים לבנים ו-1 שחור הוא $\Omega = \{(W, W), (W, B), (B, W), (B, B)\}$.

$$A = \{(B, B), (W, B)\}$$

4. מרחב המדגם בהוצאה ללא החזרה של 2 כדורים מתוך כד בו יש 2 כדורים לבנים ו-1 שחור הוא $\Omega = \{(W, W), (W, B), (B, W)\}$. המאורע נתון על ידי $A = \{(W, W), (B, W)\}$.

1.2 מבוא לתורת הקבוצות ולוגיקה

הגדרת מרחב המדגם והגדרת המאורע דומות זו לזו - בשני המקרים מדובר בקבוצות המכילות אוסף של תוצאות אפשריות בניסוי. ההבדל נובע מכך שמרחב המדגם מכיל את כלל התוצאות האפשריות, בעוד מאורע יכול להכיל רק חלק מן התוצאות. בהסתכלות רחבה, מרחב המדגם הוא דוגמא פרטית למאורע שמכיל את כל תוצאות הניסוי. לכן נרצה להבין את היחס בין הקבוצות הללו ולשם כך נגדיר את מושג השייכות, עם סימן $\in$, ואת ההכלה עם הסימונים $\subset$ ו- $\subseteq$.

נתחיל במושג הבסיסי ביותר - **השייכות**, ונחשוב על הדוגמא הבאה. ניקח את האיבר 2 ואת המאורע $A = \{1, 2, 3\}$. נבחין כי האיבר 2 הוא איבר במאורע A , על כן הוא שייך למאורע A , ונסמן זאת ב- $2 \in A$. מצד שני, האיבר 4 אינו נמצא ב- A ולכן אינו שייך למאורע. נסמן זאת $4 \notin A$. באופן כללי, איבר ω שייך למאורע A (או נמצא במאורע A) אם הוא אחד מן האפשרויות באותו המאורע. כאמור, סימון השייכות הוא $\omega \in A$. כמוסכמה, איברים במרחב המדגם יסומנו ב- $\omega \in \Omega$.

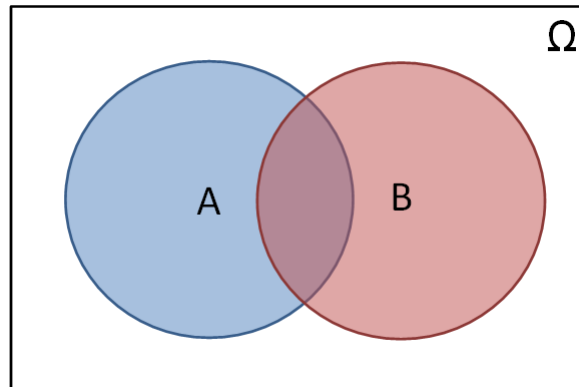
בעזרת מושג השייכות נגדיר את מושג **ההכלה**. מאורע A מוכל במאורע B אם כל האיברים של A שייכים גם ל- B . במילים פשוטות, אם כל האיברים של A נמצאים ב- B אז המאורע A מוכל במאורע B ונסמן זאת ב- $A \subseteq B$. ברגע שיחס ההכלה הנ"ל מתקיים נאמר כי מאורע A הוא **תת קבוצה** של מאורע B . המשמעות מבחינת הניסוי היא מיידית - אם A התרחש אז גם B התרחש. לדוגמא, אם הניסוי הוא הטלת קוביה ומאורע $A = \{6\}$ מציין את התוצאה 6 בעוד מאורע $B = \{4, 5, 6\}$ מציין שהתוצאה גדולה או שווה ל-4, אז ברגע ש- A מתרחש גם

B מתרחש ו- $A \subseteq B$. נשים לב כי $A = B$ זה מקרי פרטי של הכלה שכן כל איברי A נמצאים ב- B ולהיפך. בנוסף, באופן טריוויאלי כל מאורע (מהגדרתו!) הוא תת קבוצה של מרחב המדגם. זאת אומרת, מאורע הוא קבוצה המקיימת $A \subseteq \Omega$. לעיתים נוטים להפריד בין הסימון $\subseteq$ (במילים, "מוכל או שווה"), לבין הסימון $\subset$ (במילים, "מוכל ממש" או "מוכל ולא שווה") בכדי להפריד בין מצבים בהם ייתכן שיוויון בין A ל- B , לעומת המקרים בהם אין שיוויון כזה.

בעוד מרחב המדגם הוא מאורע אשר מכיל את כל התוצאות האפשריות בניסוי, עולה השאלה כיצד נראה מאורע קיצוני אחר אשר לא מכיל אף תוצאה? עבור מצבים כאלה אנו מגדירים את **הקבוצה הריקה** (או, לחילופין, המאורע הריק), והיא מסומנת ב- ϕ . הקבוצה הריקה היא קבוצה שאין בה כלל איברים, $\phi = \{\}$, ולכן זה מצב בו לא התקבלה אף תוצאה בניסוי. בניסוי ממשי המצב הזה איננו אפשרי, ונדון בכך בהמשך.

1.2.1 תורת הקבוצות: פעולות בסיסיות בין מאורעות

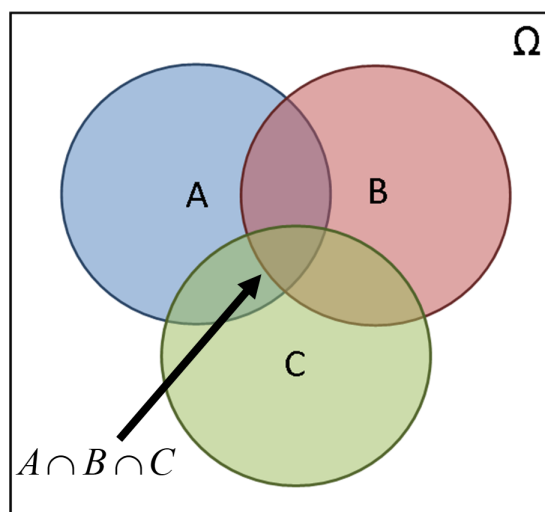
ישנן מספר פעולות בסיסיות בין מאורעות. פעולות אלו יחזרו על עצמן במספר קורסים מתמטיים ולכן כדאי להבין אותן לעומק. נדון תחילה בפעולה הבסיסית ביותר בין מאורעות והיא פעולת **האיחוד**, המסומנת ב- $\cup$. פעולת האיחוד היא פעולה שלוקחת את האיברים של צמד מאורעות ומאגדת אותם לקבוצה חדשה שהיא האיחוד של הקבוצות המקוריות. על מנת להמחיש זאת בצורה גרפית, נשתמש בדיאגרמת ואן. דיאגרמת ואן היא תיאור אבסטרקטי של מרחב המדגם כאשר המאורעות מתוארים בעזרת מעגלים בתוכה. באיור 1 נוכל לראות דיאגרמה של מרחב המדגם עם צמד מאורעות A, B . האיחוד של מאורעות אלו, מסומן ב- $A \cup B$, הוא אוסף כל התוצאות השייכות לפחות לאחד המאורעות A או B . באיור 1 מדובר הן בשטח האדום, הן בשטח הכחול, והן בשטח האדום-כחול.



איור 1: דיאגרמת ואן כללית לצמד מאורעות

מבחינה אלגברית, נתבונן במרחב המדגם של הטלת קובייה $\Omega = \{1, 2, 3, 4, 5, 6\}$. נגדיר $A = \{1, 2\}$ ו- $B = \{1, 3, 6\}$, אזי האיחוד של המאורעות הללו הוא $A \cup B = \{1, 2, 3, 6\}$. אין צורך לרשום את האיבר 1 פעמיים, שכן בכל מקרה הטלה אשר תניב את הערך 1 נכללת במאורע המדובר.

הפעולה הבאה אותה ננתח היא **חיתוך** בין מאורעות. החיתוך $A \cap B$ של צמד המאורעות זו קבוצה שמכילה את כל האיברים שנמצאים הן ב- A והן ב- B . במונחי ניסויים, המשמעות היא שגם מאורע A התרחש וגם מאורע B התרחש. באיור 1 מדובר אך ורק בשטח האדום-כחול. איור 2 מציג דוגמא נוספת עם 3 מאורעות A, B, C . החיתוך $A \cap B \cap C$ הוא השטח המשותף למעגל הכחול, למעגל האדום, ולמעגל הירוק (מסומן בחץ מתאים). זאת אומרת, מדובר בכל האפשרויות שנמצאות בו זמנית במאורע A , במאורע B , ובמאורע C . אלגברית, נמשיך עם מאורעות A ו- B שהגדרנו קודם לכן. החיתוך $A \cap B = \{1\}$ כי האיבר 1 הוא האיבר היחיד המשותף לצמד המאורעות. בעת שימוש בפעולות האיחוד והחיתוך על מספר גדול של מאורעות ניתן (ולפעמים אף חובה) להשתמש בסוגריים בכדי לייצר סדר בין הפעולות, בדומה לפעולות אריתמטיות של חיבור, חיסור, כפל, וחילוק. נעבור על שלושה חוקים עיקריים הנוגעים לסדר הפעולות בין קבוצות. נתחיל עם **חוק החילוף**. חוק החילוף קובע כי סדר כתיבת המאורעות באיחוד או בחיתוך אינו משפיע על התוצאה הסופית. בצורה מדויקת,



איור 2: דיאגרמת ואן כללית ל-3 מאורעות. החיתוך של שלושת המאורעות הוא האזור שצבוא הן באדום, הן בכחול, והן בירוק, מסומן בחץ.

עבור כל צמד מאורעות A ו- B מתקיים

$$A \cap B = B \cap A,$$

$$A \cup B = B \cup A.$$

במילים, האיחוד והחיתוך של צמד מאורעות אינו מושפע מסדר כתיבת המאורעות. תכונה זאת נובעת בראש ובראשונה מאופן הגדרת פעולות החיתוך והאיחוד. תכונה דומה מיוצגת ע"י **חוק הקיבוץ**. חוק הקיבוץ קובע כי לכל שלושה מאורעות A, B, C מתקיים

$$(A \cup B) \cup C = A \cup (B \cup C),$$

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

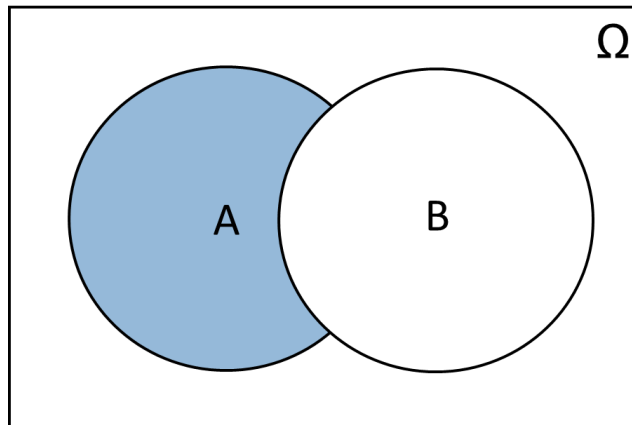
במילים פשוטות, סדר הפעולות הכולל איחודים בלבד או חיתוכים בלבד אינו משנה כל עוד מדובר באותה הפעולה שמבוצעת שוב ושוב. לעומת זאת, כאשר מבוצעת יותר מפעולה אחת מסוגים שונים, סדר הפעולות חשוב. לטובת זאת נגדיר את **חוק הפילוג**. חוק הפילוג קובע כי לכל שלושה מאורעות A, B, C מתקיים

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

באופן כללי, אין צורך לזכור מפורשות את כל הכללים הללו. הדרך הקלה להבין ולשחזר אותם היא באמצעות דיאגרמת ואן. נסו להשתמש באיור 2 בכדי להדגים את נכונות החוקים הנ"ל.

הפעולה האחרונה שננתח כעת היא פעולת **ההפרש בין מאורעות** אשר לוקחת מאורע אחד ומורידה ממנו את איברי מאורע אחר (במידה ויש איברים משותפים לצמד המאורעות). פורמאלית, $A \setminus B$ מייצג את המאורע בו לוקחים את כל איברי מאורע A ומורידים מהם את האיברים המשותפים ל- A ו- B . במידה ואין ל- A איברים משותפים, אז $A \setminus B = A$ ונקבל חזרה את המאורע A ללא כל שינוי. ראו איור 3. בכתוב מתמטי, אם $A = \{1, 2\}$ ו- $B = \{1, 3, 6\}$, אזי $A \setminus B = \{2\}$ (השמטנו את האיבר 1 ונותרנו רק עם האיבר 2 אשר נמצא ב- A ולא נמצא ב- B).



איור 3: דיאגרמת ואן כללית לצמד מאורעות. החלק הצבוע בכחול מייצג את $A \setminus B$.

דרך נוספת להגדיר מאורע היא באמצעות בחירה של כל האיברים שלא נמצאים במאורע אחר. זה הרעיון בבסיס **המאורע המשלים**. מאורע $\bar{A}$ ייקרא המאורע המשלים של מאורע A ויוגדר ביחס למרחב המדגם Ω כל האיברים שנמצאים ב- Ω אך אינם נמצאים ב- A . הדרך המתמטית לסמן זאת היא באמצעות הפרש בין מאורעות, $\bar{A} = \Omega \setminus A$.⁵ ניקח לדוגמא ניסוי הטלת מטבע. המאורע המשלים למאורע בו יצא "עץ" הוא המאורע בו יצא "פלי". לחילופין בהטלת קוביה, המאורע בו התוצאה היא זוגית הוא מאורע משלים למאורע שהתוצאה היא אי-זוגית, וכן הלאה. נעבור כעת לתרגול קצר של הפעולות אותן סקרנו.

תרגיל 1.4 בן מטיל קוביה הוגנת.

- רשמו את מרחב המדגם.
- רשמו את המאורעות הבאים: A - התוצאה קטנה מ-4, B - התוצאה גדולה או שווה ל-3, C - התוצאה זוגית, D - התוצאה אי זוגית.
- האם $4 \in A$? האם $3 \in B$?
- רשמו מפורשות את המאורעות הבאים: $A \cap B$, $A \cup B$, $C \cap B$, $(A \cap B) \cap C$, $A \setminus D$, $\bar{D}$, $((A \setminus B) \cap C)$.

פתרון:

- מרחב המדגם מוגדר על ידי הקבוצה $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- $A = \{3, 2, 1\}$, $B = \{3, 4, 5, 6\}$, $C = \{2, 4, 6\}$, $D = \{1, 3, 5\}$.
- נשים לב כי $A = \{1, 2, 3\}$, $4 \notin A$, בעוד $3 \in B = \{3, 4, 5, 6\}$.

⁵כדאי לדעת שמאורע משלים לעיתים מסומן ב- A^c .

$$\begin{aligned}
A \cap B &= \{3\}; \\
A \cup B &= \{1, 2, 3, 4, 5, 6\} = \Omega; \\
C \cap B &= \{4, 6\}; \\
(A \cap B) \cap C &= \{3\} \cap \{2, 4, 6\} = \{\} = \phi; \\
A \setminus D &= \{1, 2, 3\} \setminus \{1, 3, 5\} = \{2\}; \\
\bar{D} &= \Omega \setminus \{1, 3, 5\} = \{2, 4, 6\}; \\
((A \setminus B) \cap C) &= \{1, 2\} \cap \{2, 4, 6\} = \{2\}.
\end{aligned}$$

תרגיל 1.5 ניקח את קבוצות הסטודנטים הנמצאים בכיתה ונגדיר את המאורעות הבאים: C - הסטודנטים שאוהבים חתולים; D - הסטודנטים שאוהבים כלבים; F - הסטודנטים שאוהבים דגים. רשמו את המאורעות הבאים:

- A_1 - הסטודנטים שאוהבים לפחות חיה אחת.
- A_2 - הסטודנטים שלא אוהבים אף חיה.
- A_3 - הסטודנטים שאוהבים רק חתולים.
- A_4 - הסטודנטים שאוהבים את כל החיות.
- A_5 - הסטודנטים שאוהבים בעל חיים אחד בלבד.
- A_6 - הסטודנטים שאוהבים לפחות 2 בעלי חיים.

פתרון:

$$\begin{aligned}
A_1 &= C \cup D \cup F; \\
A_2 &= \bar{A}_1 = \bar{C} \cap \bar{D} \cap \bar{F}; \\
A_3 &= C \cap \bar{D} \cap \bar{F} = (C \setminus D) \setminus F; \\
A_4 &= C \cap D \cap F; \\
A_5 &= (C \cap \bar{D} \cap \bar{F}) \cup (D \cap \bar{C} \cap \bar{F}) \cup (F \cap \bar{D} \cap \bar{C}); \\
A_6 &= (C \cap D) \cup (D \cap F) \cup (F \cap C).
\end{aligned}$$

1.2.2 חוקי דה־מורגן

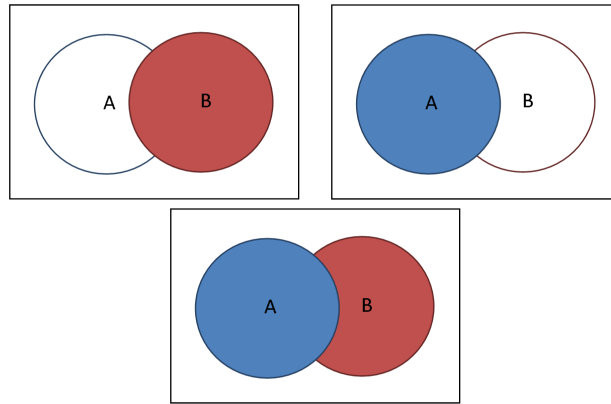
בתרגיל האחרון ראינו שילוב של מאורעות משלימים ופעולות איחוד וחיתוך רבות. באופן טבעי עולה השאלה - האם יש חוקים מתאימים לשילוב של פעולות המשלים עם הכללים שראינו קודם לכן? התשובה היא כן ולשם כך נוסחו חוקי דה־מורגן. חוקי דה־מורגן הם חוקים המשלבים את חוקי הקיבוץ יחד עם רעיון המאורע המשלים. חוקי דה מורגן קובעים כי עבור כל צמד מאורעות A, B מתקיים

$$\begin{aligned}
\overline{(A \cup B)} &= \bar{A} \cap \bar{B}, \\
\overline{(A \cap B)} &= \bar{A} \cup \bar{B}.
\end{aligned}$$

בנוסף, עבור מספר כללי של n מאורעות נקבל⁶

$$\begin{aligned}\overline{\left(\bigcup_{i=1}^n A_i\right)} &= \bigcap_{i=1}^n \overline{A_i}, \\ \overline{\left(\bigcap_{i=1}^n A_i\right)} &= \bigcup_{i=1}^n \overline{A_i}.\end{aligned}\quad (1)$$

ניקח לדוגמא את כלל דה־מורגן הראשון ונשתמש בדיאגרמת ואן להדגים אותו. נסמן את האזור אליו מתייחס כל אגף בטענה. אגף שמאל של המשוואה, $\overline{(A \cup B)}$, מתאר את המאורע המשלים של $A \cup B$, הנתון על ידי האזור הלבן באיור 1. מצד שני, באגף ימין של המשוואה ישנו חיתוך $\overline{A} \cap \overline{B}$ של המאורעות המשלימים כפי שהם מוצגים באיור 4. ניתן לראות כי חיתוך האזורים הלבנים בדיאגרמות העליונות באיור 4 מניב את החלק הלבן בשרטוט התחתון באותו האיור, ולכן מדגים את נכונות הטענה.



איור 4: מרשם סכמטי של אגף ימין של המשוואה $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$. בצמד הדיאגרמות העליונות, האזור הצבוע בלבן מתאר את המאורעות המשלימים של A ו- B , בעוד השטח הלבן בדיאגרמה התחתונה מציג את החיתוך $\overline{A} \cap \overline{B}$.

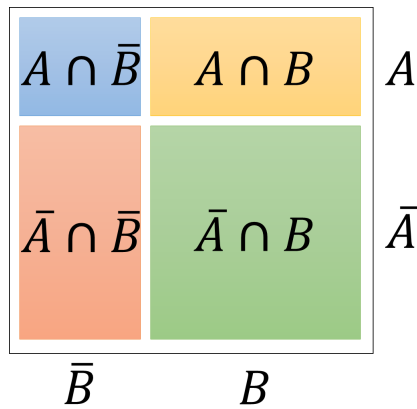
דיאגרמה נוספת שיכולה לסייע בהבנה של כללי דה־מורגן היא דיאגרמת 2×2 המוצגת באיור 5. כדאי לציין כי דיאגרמת 2×2 רלוונטית בעיקר עבור ניתוח של צמד מאורעות, בעוד ניתוח של 3 מאורעות או יותר מצריך שימוש בדיאגרמת ואן כללית. לדוגמא, באיור 6 מוצגת דיאגרמת ואן המסבירה את יחס ההכלה $C \subseteq B \subseteq A$ אותו פגשנו קודם לכן.

1.2.3 מאורעות זרים

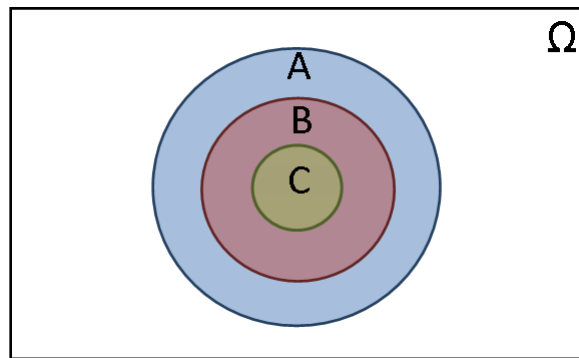
נסתכל שוב על הסעיף האחרון בתרגיל 1.4. ראינו כי $(A \cap B) \cap C = \phi$, כלומר, אין איברים משותפים למאורע C ולמאורע $A \cap B$. צמד מאורעות כאלה נקראים **זרים**.

הגדרה 1.6 (מאורעות זרים) מאורעות A ו- B נקראים זרים זה לזה אם אין להם איברים משותפים, ולכן $A \cap B = \phi$. כאשר ישנה קבוצה כללית של מאורעות, לדוגמא A, B, C , נוכל לשאול אם כל צמד מהם זר אחד לשני, ואם זה אכן המצב אז המאורעות יקראו **זרים בזוגות**. לעומת זאת, אם החיתוך של כלל המאורעות הוא קבוצה ריקה, אז המאורעות ייקראו רק זרים. פורמאלית, $A \cap B \cap C = \phi$ משמע שהמאורעות זרים, בעוד הגדרת זרים בזוגות מחייבת ש- $A \cap B = \phi$, $A \cap C = \phi$, $B \cap C = \phi$. התנאי עבור זרים בזוגות הוא חזק יותר מאשר התנאי עבור מאורעות זרים, שכן אם המאורעות זרים בזוגות הם גם בהכרח זרים (נסו להבין למה באמצעות דוגמא!).

⁶הסימון $\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup \dots \cup A_n$ מתאר איחוד של $A_1, A_2, \dots, A_n$. באופן דומה, $\bigcap_{i=1}^n A_i$ מתאר חיתוך של אותם המאורעות.



איור 5: דיאגרמת 2 על 2



איור 6: דיאגרמת ואן לתיאור יחס ההכלה $C \subseteq B \subseteq A$.

תרגיל 1.7 בכל אחד מארבעת הימים הבאים מבצעים ניסוי. כל אחד מהניסויים יכול להצליח או להיכשל. נגדיר את המאורע A_i כמאורע בו הניסוי ביום ה- i הצליח.

1. האם המאורעות זרים אחד לשני? האם המאורעות זרים בזוגות?

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

(א) כל הניסויים הצליחו.

(ב) לפחות ניסוי אחד הצליח.

(ג) כל הניסויים נכשלו.

(ד) הניסוי הראשון שהצליח היה ביום השלישי.

(ה) הניסוי השני שהצליח היה ביום השלישי.

פתרון: נענה על כל סעיף בנפרד.

1. בשאלה זו למעשה שואלים אותנו האם קיימים זוגות $i \neq j$ כך ש- $A_i \cap A_j = \emptyset$ והתשובה היא לא. לדוגמא, $A_1 \cap A_2$ הוא המאורע בו הייתה הצלחה הן ביום הראשון והן ביום השני. לכן עבור כל זוג i, j מתקיים $A_i \cap A_j \neq \emptyset$. באופן דומה נקבל כי $A_1 \cap A_2 \cap A_3 \cap A_4 \neq \emptyset$ שכן המאורע המדובר מכיל את האפשרות שבכל יום ויום הייתה הצלחה בניסוי.

2. בסעיף הנוכחי אנחנו יכולים להיעזר אך ורק במאורעות A_i , $i = 1, 2, 3, 4$ אשר הוגדרו בשאלה.

(א) כל הניסויים הצליחו: $\bigcap_{i=1}^4 A_i = A_1 \cap A_2 \cap A_3 \cap A_4$.

(ב) לפחות אחד מהניסויים הצליח: $\bigcup_{i=1}^4 A_i = A_1 \cup A_2 \cup A_3 \cup A_4$.

(ג) כל הניסויים נכשלו: $\bigcap_{i=1}^4 \bar{A}_i$.

(ד) הניסוי הראשון שהצליח היה ביום השלישי: $\bar{A}_1 \cap \bar{A}_2 \cap A_3$.

(ה) הניסוי השני שהצליח היה ביום השלישי: $(A_1 \cap A_3 \cap \bar{A}_2) \cup (A_2 \cap A_3 \cap \bar{A}_1)$.

1.3 מרחבי הסתברות כלליים

עד כה דנו אך ורק בתוצאות שיכולות להתקבל בניסוי ולא התייחסנו כלל לסבירות של אותן התוצאות. יחד עם זאת, עצם העובדה שאנו מבינים אלו תוצאות יכולות להתקבל בניסוי, מרמזת שיש לנו הבנה בסיסית של הסתברות (אחרת לא היינו מסוגלים לקבוע מה כן מתרחש ומה לא מתרחש). הדיון ההסתברותי נפתח בהגדרת פונקציית ההסתברות. פונקציית ההסתברות היא פונקציה המשייכת לכל מאורע מספר כלשהו בין 0 ל-1.

הגדרה 1.8 (פונקציית הסתברות) פונקציית הסתברות, המסומנת באותיות $\Pr$, זו פונקציה אשר משייכת לכל מאורע A מספר כלשהו, המסומן ב- $\Pr(A)$.

באופן אינטואיטיבי, נסתכל על הטלת קוביה הוגנת. אנו מבינים, מטעמי סימטריה, כי הסיכוי לקבל תוצאה זוגית הוא $\frac{1}{2}$, ואותה הטענה נכונה עבור תוצאה אי זוגית. מה לגבי תוצאה הגדולה ממש מ-4? ובכן, במקרה הזה ההסתברות של המאורע $\{5, 6\}$ היא שליש. מדוע? כי יש סיכוי (הסתברות) שווה לכל פאה ואנחנו מעוניינים ב-2 מתוך 6 התוצאות האפשריות.

פונקציית ההסתברות בעלת מספר תכונות בסיסיות. התכונות הללו הן למעשה אקסיומות (עובדות או הנחות עבודה) שנוסחו בתחילת המאה ה-20 על ידי קולמוגורוב, על כן הן נקראות על שמו. האקסיומה הראשונה והבסיסית ביותר היא **אי־שליליות**. תכונת האי־שליליות קובעת כי הסתברות לעולם איננה שלילית $\Pr(A) \geq 0$. האקסיומה השנייה היא **הנרמול** אשר קובעת כי $\Pr(\Omega) = 1$. זאת אומרת, הסיכוי שמשוהו יתרחש בניסוי הוא 1.⁷ זו תכונה די נוחה מבחינה טכנית שכן היא קובעת לנו אמת מידה לסבירות של מאורעות שונים. האקסיומה השלישית והאחרונה היא **החיבוריות (האדיטיביות)** והיא קובעת ש- $\Pr(A) = \sum_{\omega \in A} \Pr(\omega)$. במילים פשוטות, ההסתברות של מאורע נתונה על ידי סכום ההסתברויות של האפשרויות הנמצאות באותו המאורע. הסימן $\sum$ הוא סימן סכימה ובמקרה הנ"ל אנחנו סוכמים את ההסתברות של כל האפשרויות הנמצאות בתוך מאורע A . מתכונת החיבוריות אנחנו מסיקים כי ההסתברות של הקבוצה הריקה היא אפס, $\Pr(\emptyset) = 0$, שכן אנו סוכמים על 0 איברים.

לעיתים אקסיומת החיבוריות מנוסחת בצורה שונה. מבחינה מהותית, החיבוריות מאפשרת לנו לסכום הסתברויות של אפשרויות שונות. לכן ניתן לבצע את פעולת החיבור גם עבור מאורעות זרים ולא רק עבור אפשרויות שונות. כלומר, תכונת החיבוריות למעשה קובעת כי עבור כל צמד מאורעות זרים A ו- B מתקיים $\Pr(A \cup B) = \Pr(A) + \Pr(B)$. באופן דומה, עבור n מאורעות $A_1, A_2, \dots, A_n$ זרים בזוגות מתקיים

$$\Pr\left(\bigcup_{i=1}^n A_i\right) = \Pr(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n \Pr(A_i).$$

זאת הצגה אחרת של תכונת החיבוריות, והיא שקולה לה לחלוטין.

⁷לעיתים מייחסים להסתברות אחוזים מתאימים בין 0% ל-100%.

הגדרה 1.9 (מרחב הסתברות) מרחב הסתברות זה זוג $(\Omega, \Pr)$ כאשר Ω הוא מרחב מדגם ו- $\Pr$ היא פונקצית הסתברות המוגדרת על מרחב המדגם Ω .

במילים פשוטות, תיאור של מרחב הסתברות מצריך תיאור של מרחב המדגם ושל פונקצית ההסתברות. לדוגמא, להלן מספר מרחבי הסתברות: בהטלת מטבע הוגן מרחב ההסתברות נתון על ידי $\Omega = \{0, 1\}$, $\Pr(0) = \Pr(1) = \frac{1}{2}$; בהטלת מטבע מוטה עם סיכוי $\frac{1}{3}$ לנחות על 1 נקבל מרחב הסתברות שונה הנתון על ידי $\Omega = \{0, 1\}$, $\Pr(0) = \frac{2}{3}$, $\Pr(1) = \frac{1}{3}$; בהטלת קוביה הוגנת נקבל מרחב הסתברות הנתון על ידי $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\Pr(\omega) = \frac{1}{6}$, $\forall \omega \in \Omega$. שימו לב שתכונת האדיטיביות מאפשרת לנו להגדיר את פונקצית ההסתברות עבור התוצאות הבודדות בלבד, ואין צורך להגדירה עבור כל מאורע.

תרגיל 1.10 נתון מרחב המדגם הבא: $\Omega = \{1, 2, \dots, 10\}$, $\Pr(i) = c \cdot i^2$, $\forall i \in \Omega$, כאשר c הוא קבוע כלשהו. מצאו את ערכו של c , וחשבו את ההסתברות שתוצאות הניסוי תחלק ב-3.

פתרון: נוכל למצוא את הקבוע c מתנאי הנרמול. בפרט אנחנו יודעים שסכום ההסתברויות חייב להיות 1, לכן

$$1 = \sum_{i=1}^{10} \Pr(i) = \sum_{i=1}^{10} c \cdot i^2 = 385c.$$

מכאן נובע ש- $c = \frac{1}{385}$. המאורע המבוקש בשאלה (תוצאת הניסוי מתחלקת ב-3) הוא $A = \{3, 6, 9\}$. מאדיטיביות נקבל

$$\Pr(A) = \Pr(3) + \Pr(6) + \Pr(9) = \frac{3^2 + 6^2 + 9^2}{385} = \frac{126}{385}.$$

1.3.1 חוקי הסתברות בסיסיים

לאחר שהגדרנו את מרחב ההסתברות, נתפנה להוכיח מספר תכונות בסיסיות שלו. התכונות הללו יסייעו לנו בחישוב הסתברות למאורעות המורכבים מאיחודים, חיתוכים, ומשלימים של מאורעות אחרים. נפתח ב**נוסחת ההכלה וההפרדה**. הנוסחה קובעת כי לכל צמד מאורעות A ו- B מתקיים

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

הדרך הקלה להמחיש את נוסחת ההכלה וההפרדה היא בעזרת דיאגרמת ון. כאשר ישנם איברים משותפים לצמד המאורעות, אז סכימה בנפרד של ההסתברויות יכולה לייצר כפילויות. הכפילויות נובעות מכך שאנחנו עוברים פעמיים על אותם האיברים המשותפים לצמד המאורעות כאשר אנו סוכמים $\Pr(A) + \Pr(B)$. לכן, אנו מחסרים פעם אחת את האיברים שסכמנו פעמיים ומקבלים בדיוק את הסכום המבוקש. לדוגמא, בניסוי הטלת קוביה הוגנת נגדיר $A = \{1, 2, 3\}$ ו- $B = \{2, 3, 4\}$. אז,

$$\Pr(A \cap B) = \Pr(\{2, 3\}) = \frac{2}{6},$$

$$\Pr(A) = \Pr(B) = \frac{3}{6},$$

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

$$\Pr(A) + \Pr(B) - \Pr(A \cap B) = \frac{3}{6} + \frac{3}{6} - \frac{2}{6} = \frac{4}{6}.$$

אם A ו- B זרים, אז $\Pr(A \cap B) = \Pr(\emptyset) = 0$. לכן נוסחת ההכלה וההפרדה היא הרחבה ישירה של תכונת האדיטיביות. במילים אחרות, תכונת האדיטיביות נובעת ישירות מנוסחת ההכלה וההפרדה והיא מקרה פרטי המתרחש כאשר המאורעות זרים. המסקנה הבאה מכלילה את נוסחת ההכלה וההפרדה עבור צמד מאורעות לשלושה מאורעות.

מסקנה 1.11 עבור שלושה מאורעות A, B, C נוסחת ההכלה וההפרדה קובעת כי

$$\begin{aligned}\Pr(A \cup B \cup C) &= \Pr(A) + \Pr(B) + \Pr(C) \\ &+ \Pr(A \cap B) + \Pr(A \cap C) + \Pr(B \cap C) \\ &- \Pr(A \cap B \cap C).\end{aligned}$$

נשים לב שניתן בקלות לקבל את הנוסחה עבור 2 מאורעות מהמסקנה האחרונה על ידי הצבת $c = \phi$. נעבור לדון בהסתברות של המאורע המשלים $\bar{A}$ בעזרת הכללים שראינו לעיל. נזכור כי $\bar{A} \cup A = \Omega$ וצמד המאורעות זרים זה לזה. על כן, מתכונת האדיטיביות אנו למדים כי $\Pr(A) + \Pr(\bar{A}) = \Pr(A \cup \bar{A}) = \Pr(\Omega) = 1$, כאשר במעבר האחרון השתמשנו באקסיומת הנרמול. נסכם את התוצאה במסקנה הבאה.

מסקנה 1.12 (הסתברות מאורע משלים) עבור כל מאורע A , ההסתברות של המאורע המשלים לו היא

$$\Pr(\bar{A}) = 1 - \Pr(A).$$

תרגיל 1.13 בהצגה לכבוד פורים מככבים אסתר, מרדכי והמן הרשע. לכל אחד מהכוכבים ההצגה יש שחקן מחליף למקרה של היעדרות. אסתר נעדרת ב-40% מההופעות, מרדכי ב-50% מההופעות והמן ב-35% מההופעות. ידוע שכל שניים מהכוכבים נעדרים יחדיו ב-15% מההופעות וכולם ביחד נעדרים מ-5% מההופעות. חשבו את ההסתברות שבהצגה אליה קניתם כרטיסים לא יהיה אף שחקן מחליף.

פתרון: נסמן A אסתר נעדרת, B מרדכי נעדר, C המן נעדר. מהנתונים ידוע ש $\Pr(A) = 0.4$, $\Pr(B) = 0.5$, $\Pr(C) = 0.35$, $\Pr(A \cap B) = \Pr(A \cap C) = \Pr(B \cap C) = 0.15$ וכן $\Pr(A \cap B \cap C) = 0.05$. לפי דה-מורגן מתקיים $\Pr(\bar{A} \cap \bar{B} \cap \bar{C}) = 1 - \Pr(A \cup B \cup C)$ ולפי כלל ההכלה וההפרדה נקבל

$$\begin{aligned}\Pr(A \cup B \cup C) &= \Pr(A) + \Pr(B) + \Pr(C) - \Pr(A \cap B) - \Pr(A \cap C) \\ &- \Pr(B \cap C) + \Pr(A \cap B \cap C) = 0.85\end{aligned}$$

$$\Pr(\bar{A} \cap \bar{B} \cap \bar{C}) = 0.15 \text{ מכאן נובע ש-}$$

נמשיך למקרים בהם יש יחס של הכלה בין מאורעות A ו- B . במקרים הללו נוכל לחשב ישירות את ההסתברות של ההפרש בין המאורעות. נניח לצורך העניין כי $A, B \subseteq \Omega$ הם מאורעות כך ש- $A \subseteq B$. נשים לב ש- $B = A \cup (B \setminus A)$ ושני המאורעות הללו הם מאורעות זרים. מכאן נובע ש- $\Pr(B) = \Pr(A \cup (B \setminus A)) = \Pr(A) + \Pr(B \setminus A)$, ולכן $\Pr(B \setminus A) = \Pr(B) - \Pr(A)$. נסכם זאת בטענה הבאה.

טענה 1.14 לכל צמד מאורעות A ו- B כך ש- $A \subseteq B$, מתקיים $\Pr(B \setminus A) = \Pr(B) - \Pr(A)$ ובנוסף $\Pr(B) \geq \Pr(A)$.

טענה 1.14 נכונה אך ורק כאשר יש יחס הכלה בין המאורעות. אם אין יחס הכלה בין המאורעות, אז $\Pr(B) - \Pr(A)$ עלול להיות שלילי והטענה לעיל לא יכולה להיות נכונה (כי הסתברות לעולם איננה שלילית!). מצד שני, תמיד נוכל לטעון כי $B = (B \cap \bar{A}) \cup (B \cap A)$, ובמילים - המאורע B מורכב מהחלק שאינו נחתך עם המאורע A , הרי הוא $B \cap \bar{A}$, ומן החלק המשותף עם מאורע A , והוא $B \cap A$. שימו לב כי שני החלקים הללו זרים אחד לשני, ועל כן נקבל את המסקנה הכללית הבאה.

מסקנה 1.15 עבור כל צמד מאורעות A, B מתקיים $\Pr(B) = \Pr(B \cap A) + \Pr(B \cap \bar{A})$.

תרגיל 1.16 בסקר שנערך בעיר מסוימת נמצא ש-60% מהתושבים מגדלים כלב. בנוסף, 30% מהתושבים מגדלים חתולים ו-15% מגדלים גם כלב וגם חתול. ענו על הסעיפים הבאים:

1. מה ההסתברות שתושב מקרי מגדל לפחות בעל חיים אחד?

2. מה ההסתברות שתושב מקרי לא מגדל 2 בעלי חיים?

3. מה ההסתברות שתושב מקרי מגדל כלב, אך לא חתול?

4. מה ההסתברות שתושב מקרי מגדל חתול, אך לא כלב?

פתרון: נסמן ב- D את בעלי הכלבים וב- C את בעלי החתולים. נענה על כל סעיף בנפרד.

1. המאורע המבוקש הוא $C \cup D$. נשתמש בכללי הכלה והפרדה ונקבל $\Pr(C \cup D) = \Pr(C) + \Pr(D) - \Pr(C \cap D) = 0.6 + 0.3 - 0.15 = 0.75$.

2. המאורע המבוקש הוא $\overline{C} \cup \overline{D}$, או במילים, התושב לא מגדל כלב או לא מגדל חתול. נשתמש בכללי דה־מורגן ובכלל המשלים, ונקבל $\Pr(\overline{C} \cup \overline{D}) = \Pr(\overline{C \cap D}) = 1 - \Pr(C \cap D) = 1 - 0.15 = 0.85$.

3. המאורע המבוקש הוא $D \cap \overline{C}$. נחשב בעזרת מסקנה קודמת $\Pr(D \cap \overline{C}) = \Pr(D) - \Pr(C \cap D) = 0.6 - 0.15 = 0.45$.

4. המאורע המבוקש הוא $C \cap \overline{D}$. נחשב בעזרת מסקנה קודמת $\Pr(C \cap \overline{D}) = \Pr(C) - \Pr(C \cap D) = 0.3 - 0.15 = 0.15$.

1.4 מרחבי הסתברות אחידים

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

הגדרה 1.17 (מרחב הסתברות אחיד) מרחב הסתברות $(\Omega, \Pr)$ נקרא אחיד אם לכל איבר $\omega \in \Omega$ במרחב מתקיים,⁸

$$\Pr(\omega) = \frac{1}{|\Omega|}. \quad (2)$$

ז"א, לכל תוצאה במרחב המדגם ישנה הסתברות שווה, בניגוד למרחב הסתברות אסימטרי בו ישנה לפחות תוצאה אחת המתקבלת בהסתברות שונה מהתוצאות האחרות. מרחבי הסתברות אחידים קלים לתיאור בגלל הסימטריה בין התוצאות: בהטלת קוביה הוגנת, התוצאות 1 עד 6 המתקבלות כל אחת בהסתברות $\frac{1}{6}$; בהטלת מטבע הוגן, מרחב המדגם הוא $\Omega = \{H, T\}$ וכל תוצאה מתקבלת בהסתברות חצי $\Pr(H) = \Pr(T) = \frac{1}{2}$; בניסוי בו מוציאים כדור מכד המכיל כדור לבן, כדור שחור, וכדור אדום מרנב המדגם הוא $\Omega = \{B, W, R\}$ וכל תוצאה מתקבלת בהסתברות שליש; בעת בחירת סטודנט באקראי, מרחב המדגם הוא $\Omega = \{Students\}$ וההסתברות לכל תוצאה היא

$$\Pr(student\ i) = \frac{1}{\text{number of students}}.$$

דוגמא פשוטה למרחב לא סימטרי נקבל בניסוי במהלכו זורקים כדור לסל ועוצרים בקליעה הראשונה. המרחב מורכב מאינסוף איברים, ולא ייתכן כי כולם בעלי אותה הסתברות שכן סכום ההסתברויות יהיה אינסופי.

במרחב אחיד ההסתברות של מאורע תלויה אך ורק במספר התוצאות המוכלות באותו המאורע ובמספר האפשרויות הכולל. נסמן את מספר התוצאות במאורע A ב- $|A|$ (לעיתים מסומן גם ב- $\#A$) ונחשב ישירות

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

כאשר המעבר הראשון נובע מתכונת האדיטיביות והמעבר השני נובע ממשוואה (2). נסכם את התוצאה בטענה הבאה.

⁸הסימון $|\Omega|$ מציין את מספר האיברים שנמצאים ב- Ω .

טענה 1.18 (הסתברות מאורע במרחב אחיד) לכל מאורע $A \subseteq \Omega$ במרחב הסתברות אחיד מתקיים $\Pr(A) = \frac{|A|}{|\Omega|}$.

טענה 1.18 לא אמורה להפתיע אותנו. זאת תוצאה אשר אנחנו מבינים בצורה די אינטואיטיבית. לדוגמא, נסו לחשוב מה ההסתברות לקבל תוצאה הקטנה מ-3 בהטלת קוביה הוגנת? ההסתברות היא $\frac{2}{6}$ וזאת בגלל שיש 2 תוצאות אפשריות $\{1, 2\}$ מתוך 6 סה"כ. באופן דומה, מה הסיכוי שבהטלה כפולה של מטבע הוגן נקבל HH? הסיכוי במקרה הזה הוא רבע. כי ישנן 4 תוצאות במרחב המדגם ואנחנו מעוניינים בהסתברות של אחת בלבד.

1.5 קומבינטוריקה

בשל העובדה שהסתברות במרחבים אחידים נקבעת בראש ובראשונה על ידי מספר האפשרויות במאורע, ישנה חשיבות רבה ליכולת לספור את כמות האפשרויות בכל מאורע. זאת בדיוק הרעיון מאחורי התחום המתמטי הנקרא **קומבינטוריקה**. קומבינטוריקה הוא תחום יחסית חמקמק. מצד אחד, טכניקות הספירה יכולות להיות די מסובכות ולא קלות להבנה. מצד שני, התיאוריה כולה מתנקזת ליכולת הבסיסית לספור אפשרויות. זאת אומרת, הרקע התיאורטי הנדרש הוא אך ורק היכולת לספור וזאת נלמדת בבי"ס היסודי. על כן, חלק נכבד מן החישובים תלוי בלוגיקה וביצירתיות שלנו. למרות כל זאת, ישנו עקרון בסיסי אחד שאנחנו צריכים להבין וליזכור - **עקרון הכפל**.

1.5.1 עקרון הכפל

עקרון הכפל מסייע לנו לנתח את מספר האפשרויות כאשר ישנו תהליך המתבצע במספר שלבים. בכל שלב נוספות אפשרויות, בלי קשר לשלבים המוקדמים, ולכן בעזרת כפל של מספר האפשרויות בכל השלבים נקבל את סך האפשרויות הכולל. לצורך העניין, נניח ויש בפנינו ניסוי המורכב מ- k שלבים. בכל שלב i ישנן n_i תוצאות אפשרויות. אנו מעוניינים לחשב את מספר האפשרויות הכולל בניסוי ה- k שלבי. הדרך הקלה ביותר לענות על השאלה הנ"ל, היא לתאר עץ או סכימה שפורסת באופן מופשט את תוצאות הניסוי. כל תוצאה אפשרית בשלב הראשון מתפצלת ל- n_2 תוצאות בשלב השני, המתפצלות כ"א ל- n_3 תוצאות בשלב השלישי וכן הלאה. לכן סה"כ נקבל $\prod_{i=1}^k n_i = n_1 \cdot n_2 \cdot n_3 \cdots n_k$ תוצאות אפשריות סה"כ.⁹ בכדי להבין את עיקרון הכפל כראוי, נעבור על כמה דוגמאות.

- מספר התוצאות האפשריות בהטלת מטבע פעמיים הוא 4 $\{HH, HT, TH, TT\}$. 2 אפשרויות בשלב הראשון, ועבור כל אפשרות בשלב הראשון יש 2 אפשרויות בשלב השני.
- מספר התוצאות האפשריות בהטלת מטבע 3 פעמים הוא $2^3 = 8$.
- מספר התוצאות האפשריות בהטלת קוביה פעמיים הוא $6 \cdot 6 = 36$ (6 תוצאות אפשריות בהטלה הראשונה ו-6 תוצאות בהטלה השנייה). באופן כללי, ישנן 36 זוגות המורכבים מהמספרים השלמים מ-1 עד 6.
- מספר התוצאות האפשריות בהטלת קוביה ולאחר מכן מטבע הוא $6 \cdot 2 = 12$.
- מספר המילים השונות בנות 5 אותיות שניתן ליצור מהא"ב העברי הוא 22^5 .

באופן כללי, אם בכל שלב בניסוי ה- k שלבי יש n אפשרויות, אזי ישנן n^k תוצאות אפשריות סה"כ.

תרגיל 1.19 בקורס סטטיסטיקה לכלכלנים מקצים באקראי 4 מרצים ל-8 קבוצות, בלי הגבלה על מספר הקבוצות שכל מרצה ילמד. מה מספר החלוקות האפשרי?

פתרון: כל קורס צריך לבחור את המרצה שילמד אותו. הקורסים בוחרים את המרצים, ולכן לקורס הראשון יש 4 אפשרויות, לקורס השני יש 4 אפשרויות וכן הלאה. סך הכל יש 4^8 חלוקות אפשריות.

⁹הסימון $\prod$ מסמל כפל של מספר גורמים; במקרה הנ"ל מדובר בכפל של אפשרויות בהתאם לשלבי הניסוי.

תרגיל 1.20 אלון בוחר סיסמא למחשב באורך 5 תווים. שני התווים הראשונים חייבים להיות אותיות לטיניות ($A, B, C, \dots$), שלושת התווים הבאים נקבעים בעזרת הטלת קובייה. סיסמא נחשבת לחוקית אם היא לא מכילה 2 אותיות לטיניות זהות. מה ההסתברות שהסיסמא שנבחרה היא חוקית?

פתרון: תחילה נבנה את מרחב המדגם. עבור התו הראשון יש 26 אפשרויות ואותו הדבר נכון עבור התו השני. עבור כל אחד מהתווים הבאים יש 6 אפשרויות סה"כ. לכן נעזר בעיקרון הכפל ונקבל

$$|\Omega| = 26 \cdot 26 \cdot 6 \cdot 6 \cdot 6 = 26^2 \cdot 6^3.$$

כעת נעבור לחישוב ההסתברות. נסמן את המאורע בו יש 2 אותיות זהות בסיסמא ב- A . עבור התו הראשון יש לנו 26 אפשרויות, אבל ברגע שנבחר התו הראשון, התו השני באופן מיידי חייב להיות זהה לראשון בכדי שהם יהיו זהים. על כן, מספר האפשרויות לבחור את צמד התווים הראשונים הוא $26 \cdot 1$. מכאן נובע כי $\Pr(A) = \frac{|A|}{|\Omega|} = \frac{26 \cdot 1 \cdot 6 \cdot 6 \cdot 6}{26^2 \cdot 6^3} = \frac{1}{26}$. ולכן לפי כלל ההסתברות למאורע משלים $\Pr(\bar{A}) = 1 - \Pr(A) = \frac{25}{26}$.

1.5.2 מדגמים עם/בלי החזרה, סדורים ולא סדורים

את החישובים הקומבינטוריים ניתן לחלק בכלליות לפי שתי קטגוריות: **מדגמים סדורים/לא סדורים, ודגימה עם/בלי החזרה**. החלוקה לעיל איננה פותרת את בעיות החישוב אותן נפגוש למכביר, אבל הן מסייעות ברמה מסויימת בהבנתן. נתחיל עם החלוקה הפשוטה למדגם סדור ולא סדור.

נניח ויש ברשותנו סלסלה עם 3 פירות: בננה, תפוח, ואגס. אנחנו רוצים לחלק אותם באקראי לאלון, בן, ובר. כמה אפשרויות חלוקה יש בפנינו? הדרך הקלה לדמיין את הבעיה היא להניח שאלון, בן, ובר מסודרים בשורה ואנחנו עוברים ומחלקים להם את הפירות. עבור אלון יש לנו 3 אפשרויות, עבור בן יוותרו לנו 2 אפשרויות, ועבור בר נשארה אפשרות אחת בלבד, לכן מספר האפשרויות הכללי, לפי עקרון הכפל, הוא $3! = 1 \cdot 2 \cdot 3 = 6$. כפי שניתן להבין, למעשה סידרנו את הפירות בשורה. הגרלנו את הפרי הראשון שקיבל אלון, לאחר מכן הגרלנו את הפרי השני שעבר לבן, ולבסוף הגרלנו את הפרי האחרון שבר קיבלה. בעזרת הדוגמא הזאת נוכל לנסח את המסקנה הכללית לגבי סידור איברים שונים בשורה.

מסקנה 1.21 (מדגם סדור) מספר הדרכים לסדר n איברים שונים בשורה הוא $n! = 1 \cdot 2 \cdot 3 \cdots n$.

כעת נחזור לסלסלת הפירות ונניח שאנו רוצים לשלוח 2 פירות מתוך ה-3, ולחלק אותם לאלון ובן. מספר האפשרויות לביצוע החלוקה דומה למקרה הקודם. ישנן 3 אפשרויות לבחירת הפרי עבור אלון ו-2 אפשרויות לבחירה עבור בן. נשים לב שעברנו על כל האפשרויות שכן כל אחד מהילדים יכול לקבל כל פרי מן הסלסלה. ננסח את המסקנה הכללית.

מסקנה 1.22 (מדגם סדור ללא החזרה) מספר הדרכים לדגום k איברים מתוך n איברים שונים, עם חשיבות לסדר וללא החזרה, הוא

$$n \cdot (n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!}$$

ההסבר למסקנה הוא מיידי. עבור האיבר הראשון ישנן n אפשרויות, עבור האיבר השני ישנן $n-1$ אפשרויות, וכן הלאה, עד האיבר ה- k . מדגם שכזה ייקרא **מדגם סדור ללא החזרה**.

נמשיך למדגם הבא - מדגם **לא סדור ללא החזרה**. נניח כי אנו מעוניינים להוציא 2 פירות מתוך ה-3 שבסלסלה ללא חשיבות לסדר. כמה אפשרויות עומדות בפנינו? עבור הפרי הראשון יש לנו 3 אפשרויות ועבור הפרי השני יוותרו לנו 2 אפשרויות בלבד. לכן, לפי עקרון הכפל, יש לכאורה $3 \cdot 2 = 6$ אפשרויות סה"כ. אלא מה? הניתוח הזה הוא בעייתי כי הוא לא לוקח בחשבון שסדר בחירת הפירות איננו רלוונטי עבורנו. אין לנו כל עניין להפריד בין המקרים בהם בהתחלה בחרנו בתפוח ורק לאחר מכן בחרנו באגס לבין המקרים בהם תחילה הוצאנו אגס ורק לאחר מכן את התפוח. לכן צריך למצוא דרך לבטל את הכפילויות שספרנו. במקרה הנ"ל, כל מה שצריך זה לחלק במספר

הסידורים האפשריים של צמד הפירות, שהרי אם ספרנו את כל הסידורים בנפרד, אז חלוקה במספר הסידורים תבטל את הספירה הכפולה. לכן, מספר האפשרויות הוא בעצם $\frac{6}{2!} = 3$ (המכנה מייצג את מספר הדרכים לסדר 2 אובייקטים בשורה כפי שניסחנו במסקנה 1.21). דרך אחרת לחשוב על אותה הבעיה היא: כמה אפשרויות יש לנו לבחור פרי בודד שלא יישלף החוצה? ישנם רק 3 פירות בסלסלה, ולכן מספר האפשרויות לבחור את הפרי שלא יישלף הוא 3. שימו לב שבחירה של הפירות שלא נשלפים שקולה לחלוטין לבחירה של הפירות שכן נשלפים, והתוצאה זהה לחישוב הקודם. נעבור לניסוח המסקנה הכללית.

מסקנה 1.23 (מדגם לא סדור ללא החזרה) מספר הדרכים לדגום k איברים מתוך n איברים שונים, ללא חשיבות לסדר וללא החזרה, הוא

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!}$$

ההסבר לנוסחה הכללית דומה לדוגמא המקורית ולמסקנות קודמות. אם אנו מעוניינים לדגום k עצמים מתוך n שונים ללא החזרה וללא חשיבות לסדר, אז ישנן n אפשרויות בחירה לראשון, $n-1$ אפשרויות בחירה לשני, וכן הלאה עד לאיבר ה- k עבורו יש $n-k+1$ אפשרויות בחירה. מעקרון הכפל נקבל $n \cdot (n-1) \cdot \dots \cdot (n-k+1)$. כעת צריך לחלק בכל הסידורים האפשריים כי אין לנו כל עניין בסידור ספציפי שלהם (אותו כבר לקחנו בחשבון, כמו במדגם סדור ללא החזרה).

האובייקט המתמטי $\binom{n}{k}$ נקרא **מקדם בינומיאלי**. למקדמים הללו ישנם מספר תכונות קומבינטוריות מעניינות שאת חלקן הזכרנו. לדוגמא, $\binom{n}{k} = \binom{n}{n-k}$ (בדקו מפורשות לפי הנוסחה!). הסיבה לשיוויון הזה היא שבחירה של k איברים מתוך n שונים, היא שקולה לבחירה של יתר האיברים, $n-k$ במספר, מתוך ה- n . בדומה לדוגמא עם הפירות בסלסלה, הוצאה של 2 פירות מתוך 3 היא שקולה לאי הוצאה של פרי בודד מתוך 3.

דוגמה 1.24 ישנה כיתה עם 30 סטודנטים וצריכים לבחור ועד של 3 סטודנטים. יתר התלמידים שלא נבחרו לוועד יעסקו בפעילות אקדמית אחרת. כמה אפשרויות שונות יש לבחירת ועד? כמה אפשרויות שונות יש לבחירת התלמידים שיעסקו בפעילות אקדמית אחרת? ובכן התשובות, כפי שצינו כבר, הן זהות. נשתמש בנוסחה שראינו לדגימה ללא החזרה וללא חשיבות לסדר ונקבל

$$\binom{30}{3} = \frac{30!}{3!(30-3)!} = \frac{30!}{3!27!}$$

$$\binom{30}{27} = \frac{30!}{27!(30-27)!} = \frac{30!}{3!27!}.$$

המדגם האחרון ששנתח הוא **מדגם סדור עם החזרה**. לשם כך ניקח את הדוגמא של הטלת קובייה פעמיים. כמה תוצאות אפשריות קיימות בניסוי המדובר? ראינו כבר כי ישנן 36 אפשרויות סה"כ - 6 אפשרויות בהטלה הראשונה כפול 6 של ההטלה השנייה. למעשה, הניתוח של מדגם סדור עם החזרה זהה לעיקרון הכפל כאשר בכל שלב יש n אפשרויות, לכן בתהליכי באורך k נקבל סה"כ n^k אפשרויות.

מסקנה 1.25 (מדגם סדור עם החזרה) מספר הדרכים לדגום k איברים מתוך n שונים, עם חשיבות לסדר ועם החזרה, הוא n^k .

חשוב להקדיש מספר מילים לגבי מדגם לא סדור עם החזרה. נחשוב שוב על הדוגמא של הטלת קובייה פעמיים. במדגם הסדור ישנן סה"כ 36 תוצאות אפשריות לניסוי, ובפרט אנחנו מבדילים בין התוצאה $(1, 2)$ לתוצאה $(2, 1)$. אך כיצד נהפוך את המדגם ללא סדור? ובכן, במקרה הזה מדובר בבעיה רצינית. החלוקה בסידורים הפעם איננה רלוונטית שכן יש זוגות דוגמת $(6, 6)$ שלא ביצענו עבורם ספירה כפולה. לכן נצטרך לספור באופן מפורש את כל הזוגות האפשריים כעת. ספירה תניב את מספר האפשרויות 21 המתקבל מ- $1+2+3+4+5+6=21$ (נסו לחשוב למה!). כמו כן, אם מתייחסים לניסוי המקורי, אז הורדת הסדר הופכת את מרחב המדגם ללא סימטרי. לדוגמא, התוצאה $(6, 6)$ בניסוי המקורי מתקבלת בהסתברות של $\frac{1}{36}$, בעוד התוצאה $(1, 2)$ כאשר מסתכלים על זוגות ללא חשיבות לסדר, מתקבלת בהסתברות $\frac{2}{36}$ (כי התוצאות $(1, 2)$ ו- $(2, 1)$ הן זהות כאשר לא מייחסים חשיבות לסדר).

תרגיל 1.26 בבית ספר ובו 7 מורים מלמדים המורים 5 ימים בשבוע. כל מורה בחר מקרית יום בו לא יעבוד (א"ו). חשבו את ההסתברויות הבאות: המורה לאנגלית והמורה למתמטיקה לא יהיו חופשיים באותו יום; כל המורים בחרו ביום ו' כיום החופשי שלהם; אף מורה לא בחר את יום ו' כיום החופשי שלו; כל יום בשבוע נבחר כיום חופשי של לפחות אחד המורים.

פתרון: נתחיל ממרחב המדגם. לכל מורה יש 6 אפשרויות לבחירת היום החופשי שלו לכן כמות הצירופים הכוללת היא 6^7 .

1. למורה לאנגלית יש 6 אפשרויות לבחור את היום החופשי שלו. לאחר מכן, למורה למתמטיקה יש 5 אפשרויות עבור היום החופשי שלו ובסה"כ יש $30 = 6 \cdot 5$ אפשרויות. לכן יש להוסיף את 5 המורים האחרים, קרי לכפול ב- 6^5 . לכן ההסתברות המבוקשת היא $\frac{30 \cdot 6^5}{6^7} = 0.83$. הערה: ניתן גם לפתור את השאלה על ידי התעלמות מהמורים האחרים והתמקדות בשני המורים הנתונים. כמקודם, יש 30 אפשרויות להקצות שני ימי חופש שונים לשני המורים, אך הפעם מרחב המדגם יכול לכלול רק את אפשרויות הבחירה שלהם ויכלול רק 6^2 אפשרויות.

2. יש רק דרך אחת לייצר את המאורע הדרוש ולכן ההסתברות המבוקשת היא $\frac{1}{6^7}$.

3. במקרה הזה לכל מורה יש רק 5 אפשרויות בחירה ולכן כמות הצירופים האפשריים היא 5^7 . לכן ההסתברות המבוקשת היא $\frac{5^7}{6^7} = 0.279$.

4. במאורע המבוקש בהכרח צריכים להיות 2 מורים שיבחרו את אותו היום ורק הם. יש $\binom{7}{2}$ אפשרויות לבחור את זוג המורים שינחו באותו היום ולאחר שבחרנו אותם יש לנו 6 פריטים (5 מורים ו"זוג" מורים) שצריך לסדר בשורה (כל מקום הוא היום החופשי) ולכן יש 6! אפשרויות. לכן ההסתברות תהיה $\frac{\binom{7}{2} 6!}{6^7} = 0.054$.

תרגיל 1.27 מרכיבים באקראי מילה בת 4 תווים מן האותיות א' עד ו'. חשבו את ההסתברויות של המאורעות הבאים: A - האות א' מופיעה פעם אחת לפחות; B - האות א' מופיעה בדיוק פעם אחת; C - אין אות שחוזרת על עצמה בסיסמא.

פתרון: נתחיל ממרחב המדגם. לכל תו יש 6 אפשרויות ולכן $|\Omega| = 6^4$. נעבור לחישוב ההסתברויות. נעזר במאורע המשלים $\bar{A}$ המציין את האפשרויות שהאות א' לא נבחרה כלל. זאת אומרת, כל האפשרויות בהן הסיסמא מורכבת אך ורק מן התווים ב' עד ו'. משמע, $\Pr(A) = 1 - \Pr(\bar{A}) = 1 - \frac{5^4}{6^4}$. עבור מאורע B נשתמש בטכניקה אחרת של פירוק לתתי מאורעות. נגדיר את B_i להיות המאורע בו התו ה- i הוא היחיד שמכיל את האות א'. המאורע הזה קל לחישוב שכן הוא מקבע את האות א' בתו ה- i , בעוד כל יתר התווים נבחרים באקראי מתוך האותיות ב' עד ו'. לכן $B = \bigcup_{i=1}^4 B_i$ והאיחוד הוא על פני מאורעות זרים. על כן,

$$\begin{aligned} \Pr(B) &= \Pr\left(\bigcup_{i=1}^4 B_i\right) = \sum_{i=1}^4 \Pr(B_i) \\ &= \sum_{i=1}^4 \left(\frac{5^3}{6^4}\right) = \frac{4 \cdot 5^3}{6^4} = \frac{5^4}{6^4} \approx 0.384. \end{aligned}$$

החישוב האחרון הוא מיידי. לתו הראשון יש 6 אפשרויות, לתו השני יש 5 אפשרויות, לתו השלישי 4 אפשרויות וכן הלאה. לכן $\Pr(C) = \frac{6 \cdot 5 \cdot 4 \cdot 3}{6^4} = \frac{6!}{2! \cdot 6^4} \approx 0.28$.

שימו לב למספר נקודות חשובות שעלו עד כה. תחילה, שימוש במשלים. ישנם מקרים בהם החישוב של המאורע המשלים הוא יותר פשוט מחישוב המאורע המקורי. נצלו זאת! שנית, פרקו מאורעות מורכבים לתתי מאורעות פשוטים. ברגע שנפרק מאורע מורכב למאורעות זרים ופשוטים, החישוב יהיה מיידי.

תרגיל 1.28 בכיתה n סטודנטים. מהי ההסתברות שיש לפחות זוג אחד של סטודנטים עם אותו תאריך יום הולדת (מבלי להתייחס לשנת הלידה)?

פתרון: גודל מרחב המדגם שלנו הוא 365^n משום שלכל סטודנט יש 365 ימים אפשריים ליום הולדת וישנם n סטודנטים. נסמן ב- A את המאורע בו לפחות זוג אחד של סטודנטים נולד באותו היום, ונבחן את המאורע המשלים $\bar{A}$, המאורע בו לכל הסטודנטים ימי הולדת שונים. מספר האפשרויות במאורע $\bar{A}$ הוא $\binom{365}{n} \cdot n!$ שכן צריך לבחור n תאריכי ימי הולדת שונים (בחירה ללא החזרה) ולאחר הבחירה צריך לחלק את התאריכים בין הסטודנטים, לכן נכפיל בכל הסידורים האפשריים, $n!$. מכאן נקבל ש-

$$\begin{aligned}\Pr(\bar{A}) &= 1 - \Pr(A) = 1 - \frac{\binom{365}{n} \cdot n!}{365^n} \\ &= 1 - \frac{365!}{365^n (365 - n)!}.\end{aligned}$$

התוצאה המפתיעה היא שעבור $n = 23$ ההסתברות גדולה מ-50% (0.507 בקירוב) ועבור 60 סטודנטים ההסתברות היא 0.994.

תרגיל 1.29 בספרייה 10 ספרים בכלכלה: 3 במיקרו, 2 במאקרו, ו-5 בתורת המשחקים. מסדרים את הספרים באקראי על המדף. מה הסיכוי שכל הספרים בכל תחום יעמדו בסמוך אחד לשני על המדף?

פתרון: נסמן את המאורע המבוקש ב- A , ונחשב את מרחב המדגם, קרי מספר הדרכים לסדר 9 ספרים שונים על המדף, $|\Omega| = 9!$. נתייחס לכל תחום כ-"בלוק" ונסדר תחילה את הבלוקים בינם לבין עצמם. לאחר מכן, נכפול במספר הסידורים הפנימיים של כל בלוק.

$$\Pr(A) = \frac{3! \cdot 4! \cdot 3! \cdot 2!}{9!} = \frac{1}{210}.$$

תרגיל 1.30 במחלקה לכלכלה 25 חברי סגל - 9 דוקטורים ו-16 פרופסורים. בוחרים באקראי ועדת הוראה בת 4 חברי סגל. מצאו את ההסתברות של המאורעות הבאים: A - בועדה 2 דוקטורים ו-2 פרופסורים; B - בועדה לפחות 3 פרופסורים; C - לפחות דוקטור אחד ולפחות פרופסור אחד בועדה.

פתרון: נחשב כל מאורע בנפרד. מרחב המדגם הוא $|\Omega| = \binom{25}{4}$.

$$\begin{aligned}\Pr(A) &= \frac{\binom{9}{2} \binom{16}{2}}{\binom{25}{4}}, \\ \Pr(B) &= \frac{\binom{9}{1} \binom{16}{3} + \binom{9}{0} \binom{16}{4}}{\binom{25}{4}}, \\ \Pr(C) &= 1 - \Pr(\bar{C}) = 1 - \frac{\binom{9}{4}}{\binom{25}{4}} - \frac{\binom{16}{4}}{\binom{25}{4}}.\end{aligned}$$

2 הסתברות מותנה ואי-תלות

ישנן סיטואציות בהן, לאחר ביצוע הניסוי, אנחנו מקבלים מידע נוסף על תוצאת הניסוי. המידע הנוסף איננו מספק לנו את כל המידע לגבי הניסוי, אבל הוא כן עוזר לנו להעריך מחדש את האפשרויות העומדות בפנינו. לדוגמא, 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, ההסתברויות הללו יורדת 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. מבחינת מינוח, ההסתברות טרם קבלת המידע החדש נקראת הסתברות **אפריורית** (prior probability), בעוד ההסתברות לאחר קבלת המידע נקראת **פוסטריורית** (posterior probability).

הגדרה 2.1 לכל צמד מאורעות A ו- B כך ש- B בעל הסתברות חיובית, ההסתברות של מאורע A בהינתן מאורע B מוגדרת להיות

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

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

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

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

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

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

עדכון ההסתברות שביצענו לעיל נקרא **עדכון בייסיאני** והוא מאוד נפוץ בכלכלה. בעולם הכלכלי, סוכנים מקבלים אותות מידע בכל נקודת זמן. הם מפרשים את האותות הללו באופן הנ"ל (לפחות בתיאוריה...) בכדי להסיק מסקנות על העולם. לדוגמא, אדם אשר מעוניין לקנות דירה מחזיק במספר אמונות לגבי שוק הדירור. לאחר קריאת העיתון והסתכלות באתר הלמ"ס הוא מקבל מידע נוסף שעוזר לו לעדכן את האמונות שלו לגבי השוק. אותן האמונות אולי ינועו לכיוון הרכישה או ההמתנה, הכל בהתאם לאינפורמציה שהוא מקבל ולאמונות המוקדמות שלו.

משפט 2.2 פונקציית ההסתברות המותנה היא פונקציית הסתברות.

המשפט האחרון לכאורה נראה טריוויאלי, אבל הוא טומן בתוכו משקל רב. המשפט קובע כי כל הכללים שראינו עד כה עבור פונקציית הסתברות (כלל המשלים, הכלה והפרדה וכן הלאה) תקפים גם כאשר מדובר בהסתברות מותנה. בנוסף, כל אקסיומות ההסתברות תקפות גם עבור הסתברות מותנה. למעשה, **הסתברות מותנה היא פונקציית הסתברות** על כל המשתמע מכך, ובפרט כל מה שהוכחנו עד כה עבור פונקציית הסתברות תקף עבור הסתברות מותנה. לדוגמא, הפונקציה היא אי-שלילית, $0 \leq \Pr(A|B) \leq 1$ לכל מאורע A וכן מנורמלת $\Pr(\Omega|B) = \frac{\Pr(\Omega \cap B)}{\Pr(B)} = \frac{\Pr(B)}{\Pr(B)} = 1$.

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

דוגמה 2.4 בפונטי-פאנדי יש 45% גברים, 30% מעשנים, ו-15% הם גברים מעשנים. נבחר אדם באקראי. בהנחה ונבחר גבר, מה הסיכוי שהוא מעשן? מה הסיכוי שאינו מעשן? כיצד התשובות תשתנה בהנחה ונבחרה אישה? נסמן ב- A את המאורע שנבחר גבר וב- B את המאורע שהאדם שנבחר מעשן. באופן כללי, ההסתברויות של A ו- B הן $\Pr(A) = 0.45$ ו- $\Pr(B) = 0.3$. נחשב את הסיכוי שהאדם הנבחר מעשן, תחת ההנחה שמדובר בגבר. $\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{0.15}{0.45} = \frac{1}{3}$. זאת אומרת, שיעור המעשנים באוכלוסיית הגברים גבוה מן הממוצע באוכלוסייה הכוללת. בכדי לבחון את הסבירות שאינו מעשן נוכל להשתמש בכלל המשלים, שהרי הסתברות מותנה מקיימת את כל כללי פונקצית ההסתברות. לכן, $\Pr(\bar{B}|A) = 1 - \Pr(B|A) = \frac{2}{3}$, כעת נעבור לחלק האחרון בשאלה - ההנחה ונבחרה אישה. 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 $\Pr(B|\bar{A}) = \frac{\Pr(B \cap \bar{A})}{\Pr(\bar{A})} = \frac{\Pr(B) - \Pr(B \cap A)}{\Pr(\bar{A})} = \frac{0.3 - 0.15}{0.55} \approx 0.27$. ושיעור הנשים המעשנות נמוך יותר משיעור המעשנים באוכלוסייה. כמו כן, $\Pr(\bar{B}|\bar{A}) = 1 - \Pr(B|\bar{A}) = 1 - \frac{0.3 - 0.15}{0.55} \approx 0.73$.

תרגיל 2.5 בכד 10 כדורים הממסופרים מ-1 עד 10. שולפים 4 ללא החזרה.

1. מה הסיכוי שהכדור עם המספר 7 בחוץ?

2. ידוע ש-9 בחוץ, מה הסיכוי שגם 7 בחוץ?

3. ידוע שאין מספר קטן מ-4 בחוץ, מה הסיכוי ש-7 בחוץ?

4. מה הסיכוי ש-7 לא בחוץ אם ידוע שאין מספר קטן מ-4 בחוץ?

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

1. נגדיר את המאורע - "הכדור עם המספר 7 הינו אחד מהכדורים שהוצאו" בתור A . מאחר ומדובר במרחב מדגם שווה הסתברות, אזי $\Pr(A) = \frac{1 \cdot \binom{9}{3}}{\binom{10}{4}} = \frac{9!6!4!}{10!6!3!} = \frac{4}{10} = \frac{2}{5}$. דרך נוספת להסתכל על זה היא פשוט לחשוב על הוצאות מכד כמו סידורים בשורה. עבור כל הוצאה מכד יש סידור בשורה ולהיפך. לכן ישנם $10!$ סידורים בשורה ואותו מעניין רק מי נמצא ב-4 המקומות הראשונים ומי לא. משיקולי סימטריה הסיכוי ש-7 יהיה במקום הראשון בסידור בשורה הוא 0.1, ואותו הדבר גם עבור המקום השני והשלישי וכן הלאה ולכן $\Pr(A) = 0.1 \cdot 4 = 0.4$.

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

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

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

3. נגדיר מאורע C - כל המספרים שהוצאו גדולים או שווים ל-4. במידה והיינו צריכים לענות אינטואיטיבית אז היינו יכולים להניח שכעת הכדורים נבחרים בהסתברות שווה מתוך הערכים 10, ..., 4, ולכן הסיכוי להוציא את הכדור עם המספר 7 הוא $\frac{4}{7}$ (הסיכוי ש-7 נמצא

בארבעת המקומות הראשונים בסידור בשורה של המספרים $10, \dots, 4$. נחשב זאת מפורשות:

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

4. אנו נדרשים לחשב את ההסתברות $\Pr(A^c|C)$. נזכור כי הסתברות מותנה מקיימת את כל הכללים של הסתברות ולכן $\Pr(A^c|C) = 1 - \Pr(A|C) = 1 - \frac{4}{7} = \frac{3}{7}$.

2.1 כלל הכפל

כלל הכפל בהסתברות מסייע לנו לבצע חישוב הסתברויות של מאורע בניסויים הנערכים בשלבים. השימוש בהסתברות 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 האחרון בעל הסתברות חיובית. נעזר בהגדרה להסתברות מותנה, ועל ידי העברת אגפים נקבל $\Pr(A \cap B) = \Pr(A|B) \Pr(B)$. זאת אומרת, הסיכוי A -ש- B יתרחש, שווה לסיכוי B -ש- A יתרחש, כפול הסיכוי A -ש- B יתרחש, תחת הנחה B -אכן קרה. ההרחבה למספר כללי של מאורעות היא מיידית.

טענה 2.6 (כלל הכפל) לכל קבוצת מאורעות $A_1, \dots, A_n$ כך ש- $\Pr(\bigcap_{i=1}^n A_i) > 0$ מתקיים

$$\Pr\left(\bigcap_{i=1}^n A_i\right) = \Pr(A_1) \cdot \Pr(A_2|A_1) \cdot \Pr(A_3|A_1 \cap A_2) \cdots \Pr\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right).$$

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.

תרגיל 2.7 (הכד של פוליה) 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 הסיכוי שכולם שחורים?

פתרון: נגדיר את המאורעות $B_i, i = 1, \dots, 5$, כמאורעות בהם הכדור ה- i הוא שחור. נשתמש בחוק הכפל לטובת החישוב.

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

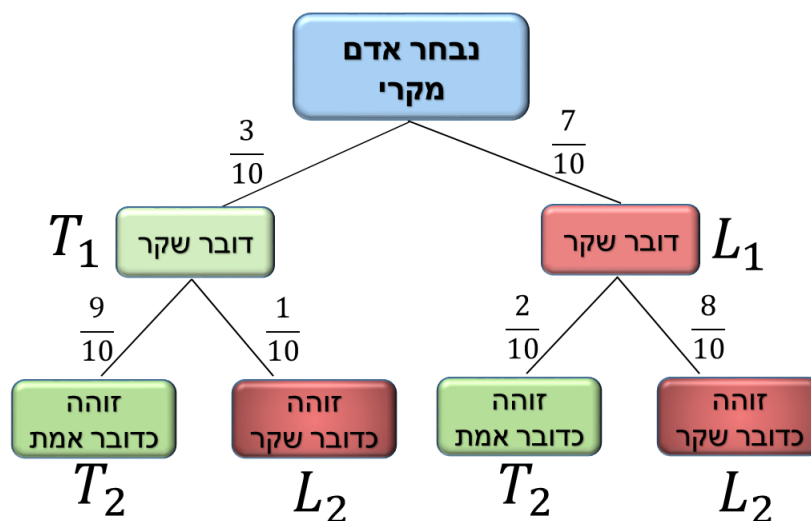
2.2 נוסחת ההסתברות השלמה

נוסחת ההסתברות השלמה היא כנראה הנוסחה החשובה ביותר בהסתברות בסיסית, ונשתמש בה (בין אם אנו שמים לב לכך או לא) ברוב המוחלט של הבעיות אותן נפתור. בהתאמה לחיוניות שבה, היא גם פשוטה. הרעיון הבסיסי מאחורי נוסחת ההסתברות השלמה לא שונה מההבנה של תרשים זרימה או עץ הסתברויות. בדרך כלל, בכדי לחשב את ההסתברות של מאורע כלשהו, אנחנו צריכים לחלק את הניסוי כולו לתתי מקרים. עבור כל תת מקרה, נחשב את ההסתברות שאכן המאורע המבוקש התרחש ואז נשקלל יחד את התוצאות לקבל את הסיכוי הכולל שהמאורע אכן התרחש.

לשם המחשה ניקח סיטואציה קצת קיצונית של שימוש בפוליגרף. פוליגרף, או בשמו העברי "מכונת אמת", אמור לבחון האם אנשים הם דוברי אמת או שקר. הפוליגרף נפוץ הן בחקירות משטרה, הן באישורים ביטחוניים, ואפילו בתעשייה וקבלה לעבודה. הבעיה עם מכונת הפוליגרף היא שאיננה מדויקת 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 העבודה כי 70% מן האנשים שמגיעים לבדיקת פוליגרף משקרים. כעת ננסה לחשב את ההסתברות שאדם מקרי בבדיקה ימצא כדובר אמת.

בכדי לענות על השאלה הזאת נגדיר את המאורעות הבאים. T_1 - אדם דובר אמת, L_1 - אדם דובר שקר, T_2 - אדם מזוהה כדובר אמת (על ידי הפוליגרף), L_2 - אדם מזוהה כדובר שקר. הנתונים בשאלה הם $\Pr(L_1) = 0.7$, $\Pr(L_2|L_1) = 0.8$, $\Pr(T_2|T_1) = 0.9$. אנחנו מעוניינים לחשב את ההסתברות שאדם ימצא דובר אמת, $\Pr(T_2)$. מאחר ואנחנו לא יודעים אם אותו אדם משקר או לא, נצטרך לחלק את המרחב שלנו לאפשרויות - אדם דובר אמת ואדם דובר שקר. כל אחד מהמקרים הללו ננתח בנפרד ואז נשקלל אותם יחד לכדי תוצאה סופית. זה בדיוק הרעיון מאחורי נוסחת ההסתברות השלמה.

נבחן את איור 7 המציג את הבעיה הנתונה כעץ הסתברויות. נעבור על הענפים המתאימים שנגמרים בתוצאה הרצויה בה האדם



איור 7: עץ הסתברויות עבור דוגמת הפוליגרף

נמצא דובר אמת, T_2 , ונחשב את ההסתברות של כל ענף בפני עצמו. בפרט אנו מבצעים את החישוב לפי השלבים הבאים. תחילה, אנחנו מפרקים את המרחב לפי הפעולה המקורית של האדם הנבחר - האם הוא שיקר או לא. מבחינה הסתברותית זה נראה כך $\Pr(T_2) = \Pr(T_2 \cap T_1) + \Pr(T_2 \cap L_1)$. לאחר מכן, אנחנו מנתחים כל חלק בנפרד. תחילה מנתחים את האיבר הראשון בסכום, ולטובת החישוב אנחנו משתמשים בהסתברות מותנה. בפרט, $\Pr(T_1 \cap T_2) = \Pr(T_1) \Pr(T_2|T_1) = 0.3 \cdot 0.9 = 0.27$. נבצע חישוב דומה עבור המחורב השני ונקבל $\Pr(T_2 \cap L_1) = \Pr(L_1) \Pr(T_2|L_1) = 0.7 \cdot 0.2 = 0.14$. כעת נסכום את צמד התוצאות יחד ונקבל $\Pr(T_2) = \Pr(T_2|T_1) \Pr(T_1) + \Pr(T_2|L_1) \Pr(L_1) = 0.9 \cdot 0.3 + 0.2 \cdot 0.7 = 0.41$. שימו לב כי הסיכוי שאדם יוכרז כדובר אמת הוא 41% בלבד! נעבור לניסוח הטענה הכללית.

משפט 2.8 (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 $B_i \cap B_j = \emptyset$ לכל $i \neq j$, ובנוסף $\bigcup_{i=1}^n B_i = \Omega$, כאשר $\Pr(B_i) > 0$. אזי לכל מאורע $A \subseteq \Omega$ מתקיים

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

האינטואיציה מאחורי נוסחת ההסתברות השלמה איננה מסובכת. 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 התרחש. במידה ונעשה זאת עבור כל מאורע B_i בנפרד, נקבל את ההסתברות של A . דרך טובה להמחיש את הנוסחה היא באמצעות עץ הסתברויות כפי שראינו בדוגמה הקודמת, שכן ההוכחה הפורמאלית היא מיידית. נפרק את A לחלקים הבאים: $A \cap B_1, A \cap B_2, \dots, A \cap B_n$. האיחוד של כל המאורעות B_i הוא כל מרחב המדגם, מכאן אנו יודעים שלא פספסנו אף אפשרות של A ו- $A = \bigcup_{i=1}^n A \cap B_i$. בנוסף, המאורעות הללו זרים בזוגות בגלל ש- $B_1, \dots, B_n$ זרים בזוגות. לכן, $\Pr(A) = \sum_{i=1}^n \Pr(A \cap B_i) = \sum_{i=1}^n \Pr(A|B_i) \Pr(B_i)$ כנדרש, כאשר במעבר האחרון השתמשנו בכלל הכפל עבור צמד מאורעות.

2.3 חוק Bayes

חוק בייס הוא חוק פשוט למידי שנובע ישירות מההגדרה של הסתברות מותנה. יחד עם זאת, הוא מאוד יעיל בחישובי הסתברות. בפרט, הוא מסייע לנו להפוך את החישוב כאשר ישנם יחסי סיבה ותוצאה בין המאורעות השונים. באופן כללי, אין באמת צורך לזכור אותו שכן ניתן לפתח אותו ישירות מן ההגדרה של הסתברות מותנה. נראה כיצד זה עובד. לפי ההגדרה של הסתברות מותנה $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$ ו- $\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}$. נכפול את צמד המשוואות במכנה שלהן ונשווה ביניהן בכדי להגיע לחוק בייס הבא.

משפט 2.9 (חוק בייס) עבור כל צמד מאורעות A ו- B בעלי הסתברות חיובית מתקיים

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

חוק בייס מאוד נוח לשימוש כאשר ישנה הסתברות מותנה ודווקא ההסתברות ההפוכה היא זו שידועה לנו. לשם המחשה, נחזור לדוגמה הקודמת של הפוליגרף. כעת ננסה לחשב את ההסתברות שאדם מקרי דובר אמת, תחת ההנחה שהוא נמצא דובר אמת. באופן מפורש, אנו מעוניינים לחשב את ההסתברות $\Pr(T_1|T_2)$, ולשם כך נשתמש בחוק בייס. הסיבה שאנו משתמשים בחוק בייס היא יחסי סיבה-תוצאה בתהליך המדובר ונתוני השאלה. 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 שהוא נמצא דובר אמת (תוצאה). במקרים הללו חוק בייס הוא יעיל במיוחד. אכן, $\Pr(T_1|T_2) = \frac{\Pr(T_2|T_1) \Pr(T_1)}{\Pr(T_2)} = \frac{0.8 \cdot 0.3}{0.41} \approx 0.59$. זאת אומרת, אדם שנמצא דובר אמת אכן אמר אמת בהסתברות 0.59 בלבד!

תרגיל 2.10 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'?

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

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

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

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

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

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

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

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

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

1. מה ההסתברות שהכדור הראשון שנוציא יהיה שחור?

2. מה ההסתברות שהכדור השני שנוציא יהיה שחור?

3. מה ההסתברות שהכדור ה-100 שנוציא יהיה שחור?

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

1. הפתרון לסעיף הראשון הוא מיידי והתשובה היא $\frac{5}{8}$ (חלוקה של גודל המאורע במרחב המדגם).

2. נחשב בעזרת נוסחת ההסתברות השלמה. נסמן ב- W_i, B_i , $i = 1, 2$ את המאורעות בהם הכדור ה- i הוא שחור או לבן, בהתאמה.

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

באופן מפתיע קיבלנו שההסתברות להוציא שחור בשלב הראשון ובשלב השני שוות.

3. לפני שנענה על הסעיף הנוכחי, ננסה למצוא הוכחה יותר פשוטה שלא תצריך התניות רבות. נתחיל במצב יחסית פשוט: מה היה קורה לו היו לנו רק 2 כדורים בכד בהתחלה - 1 שחור ו-1 לבן? אזי משיקולי סימטריה ההסתברות שהכדור ה-100 יהיה שחור היא $\frac{1}{2}$ (מאחר והכדור ה-100 הוא שחור או לבן, וישנה סימטריה מלאה בין השחורים ללבנים, אזי ההסתברות שנוציא כל אחד מהם בפעם ה-100 היא שווה). מה לגבי מצב בו ישנם 8 כדורים בכד ללא צבעים אשר ממוספרים מ-1 עד 8 (לאחר הוצאה מחזירים את הכדורים עם אותו המספר של הכדור שהוצא)? במצב זה, גם כן יש סימטריה בין שמונת הכדורים ולכן ההסתברות להוציא כדור עם ערך $i = 1, \dots, 8$ היא $\frac{1}{8}$. כעת נוכל להניח שיש מספרים על הכדורים בבעיה המקורית, הסיכוי שנוציא את כדורים 1 עד 5 (אלו הכדורים הצבועים בשחור) הוא $\frac{5}{8}$ כסכום ההסתברויות להוציא כל מספר מ-1 עד 5 בנפרד. ולסיכום, ההסתברות שנוציא כדור שחור בפעם ה-100 היא $\frac{5}{8}$ וזה נכון עבור כל שלב שנבחר.

תרגיל 2.12 (בעיית מונטי-הול) נניח ואתם משתתפים בתוכנית טלוויזיה. בתוכנית מוצגים לכם שלושה וילונות כאשר ידוע שמאחורי אחד מהם ישנה מכונית ומאחורי השניים האחרים - דחליל. הפרסים מחולקים באקראי בהסתברות שווה מאחורי הוילונות. אתם נדרשים לבחור וילון. לאחר שבחרתם באקראי את וילון מספר 1, המנחה של התוכנית פותח וילון אחד באקראי (בהסתברות שווה) מתוך כל הוילונות שאין בהם פרס מצד אחד, ולא בחרתם בהם, מצד שני,¹⁰ ומראה לכם שאין כל פרס מאחורי וילון זה. כעת המנחה פונה אליכם המנחה ומציע לכם להחליף לוילון האחר שנותר סגור. האם כדאי לכם לשמוע בעצתו? זאת אומרת, בהנחה ואתם מעוניינים לזכור במכונית, האם כדאי לכם להחליף את הבחירה המקורית שלכם?

הבעיה הנ"ל היא בעיה ידועה בשם - בעיית מונטי הול. הבעיה נקראת על שם מנחה התוכנית, "Let's make a deal", ששודרה במשך מספר עשורים בארה"ב. הסיפור מאחורי הבעיה הזאת הוא די מעניין. בחורה מאוד חכמה (נטען שהיא מחזיקה בשיא גינס עבור מנת משכל) בשם מרלין ווס סאוונט הייתה בעלת טור בעיתון ובו ענתה על שאלות וחדות, לרוב מתמטיות, שהיו שולחים אנשים מרחבי ארה"ב. בספטמבר 1990, בחור ממרילנד שאל אותה את השאלה הנ"ל והיא ענתה לו שהאסטרטגיה האופטימאלית היא החלפה. התשובה הזאת גררה אלפי תגובות (אם לא יותר, וביניהם גם מתמטיקאים) אשר טענו כי היא טועה. כעת, אנחנו נוכיח שהיא דווקא צדקה.

פתרון: אחד הדברים החשובים בבעיה הנוכחית היא להגדיר היטב את המאורעות. נניח שהוילונות ממוספרים מ-1 עד 3. נגדיר D_i - המכונית מאחורי וילון i ו- O_i - הוילון שפתח המנחה הוא וילון i , עבור $i = 1, 2, 3$. נשים לב כי המכונית יכולה להיות בהסתברות שווה מאחורי כל אחד מהוילונות, לכן $\Pr(D_i) = \frac{1}{3}$ לכל i . ההנחה הבסיסית היא שבחרנו בוילון מספר 1. לכן ישנן שתי אפשרויות - המנחה יכול לפתוח את וילון מספר 2 או את וילון מספר 3. בגלל שאין הבדל מהותי בין הוילונות, אנחנו יודעים ש- $\Pr(O_2) = \Pr(O_3) = \frac{1}{2}$, בעוד $\Pr(O_1) = 0$, כי בחרנו בוילון מספר 1.

נתחיל בניתוח המקרה שהמנחה פותח את וילון מספר 2, ונחשב את ההסתברות $\Pr(D_1|O_2)$.

$$\begin{aligned}\Pr(D_1|O_2) &= \frac{\Pr(O_2|D_1)\Pr(D_1)}{\Pr(O_2)} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\Pr(O_2|D_1)\Pr(D_1) + \Pr(O_2|D_2)\Pr(D_2) + \Pr(O_2|D_3)\Pr(D_3)} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{1}{3}.\end{aligned}$$

נסביר את שלבי החישוב. השיוויון הראשון נובע מחוק בייס והשיוויון השני נובע מנוסחת ההסתברות השלמה עבור המכנה תוך התנייה על מיקום הפרס. בשיוויון השלישי הצבנו את הערכים המתאימים, תחת ההנחה שבחרנו בדלת מספר 1. זאת אומרת, $\Pr(O_2|D_1) = \frac{1}{2}$ (אם בחרנו בוילון מספר 1 והפרס אכן שם, אז המנחה מגריל באקראי וילון מבין 2 ו-3), $\Pr(O_2|D_2) = 0$ (אם הפרס מאחורי וילון מספר 2, אז וילון זה לא יכול להיפתח), $\Pr(O_2|D_3) = 1$ (אם הפרס מאחורי וילון 3 ובחרנו בוילון 1, אז הוילון היחיד שנותר לפתוח הוא וילון מספר 2). שימו לב לנקודה החשובה - הסיכוי לזכות בפרס ללא החלפה הוא $\frac{1}{3}$, על כן החלפה לוילון מספר 3 שנותר סגור תעניק לנו את הפרס בסיכוי $\Pr(D_3|O_2) = \frac{2}{3}$!

¹⁰שימו לב לתנאי הכפול. המנחה לא יכול לפתוח את הוילון שבחרתם ולא יכול לפתוח וילון שמאחוריו פרס. מצד שני, עבור כל הוילונות שעומדים בתנאים הללו, הוא רשאי לבחור באחד באקראי.

אם נניח כי המנחה פתח את וילון מספר 3 ונבצע חישוב דומה, נקבל את אותה התוצאה (שכן, אין אפריורית הבדל בין פתיחה של וילון מספר 2 לפתיחה של וילון מספר 3). לסיכום, $\Pr(D_1|O_2) = \Pr(D_1|O_3) = \frac{1}{3}$. על כן, לפי נוסחת ההסתברות השלמה נקבל $\Pr(D_1) = \Pr(D_1|O_2) \cdot \Pr(O_2) + \Pr(D_1|O_3) \cdot \Pr(O_3) = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{3}$. לכן, הסיכוי שהפרס נמצא מאחורי הוילון שבחרנו בהתחלה הוא $\frac{1}{3}$, ומכאן נובע שהחלפה תעניק לנו את הפרס בסיכוי $\frac{2}{3}$!

תרגיל 2.13 בן הוא כורה פחם שהחליט לתת לבתו, אלונה, מתנה לחתונה. הוא נותן לה לבחור בין 2 קופסאות הזהות חיצונית - קופסא א' המכילה 2 יהלומים ופחם וקופסא ב' המכילה 2 פחמים ויהלום. מאחר ובן הוא אב מתחשב, הוא החליט לאפשר לה להוציא באקראי אבן מקופסא אחת ורק לאחר מכן להחליט איזו קופסא לקחת. מה האסטרטגיה האופטימלית מבחינת אלונה?

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

- אם אלונה מחליטה להישאר עם הקופסא שבחרה בכל מקרה (בין אם הוציאה פחם ובין אם הוציאה יהלום) אז ההסתברות שהיא תקבל לבסוף את קופסא א' היא $\frac{1}{2}$. הסיבה היא שההסתברות לבחור את קופסא א' תחילה היא חצי, ואלונה לא מחליפה קופסא, אזי ההסתברות נשמרת.
- באופן סימטרי - במידה ואלונה מחליטה להחליף קופסא בכל מקרה, אז ההסתברות שתקבל את קופסא א' בסופו של דבר היא גם כן חצי (שכנעו את עצמכם למה זה נכון, בדומה לסעיף הקודם עם שיקולי סימטריה).
- אלונה מחליטה להחליף את הקופסא במידה ותוציא פחם ולהישאר עם הקופסא במידה ותוציא יהלום. נחשב את ההסתברות לקבל לבסוף את קופסא א'. נסמן ב- B_1, A_1 את המאורעות בהן אלונה בחרה תחילה בקופסא א' ו-ב' בהתאמה, ונסמן ב- B_2, A_2 את המאורעות בהן קיבלה לבסוף את קופסאות א' ו-ב' בהתאמה.

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

ואנו רואים שאסטרטגיה זו כבר יותר טובה מצמד הקודמות.

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

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

ולכן האסטרטגיה האופטימלית היא האסטרטגיה השלישית. באותה מידה ניתן לפתור את התרגיל בעזרת עץ הסתברויות.

2.4 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 ונשלף קלף באקראי. כעת עלינו לנחש את הקלף שהוצא - צבע וצורה. אפריורית, $\Pr(\text{king}) = \frac{1}{13}$ ו- $\Pr(\text{diamond}) = \frac{13}{52}$. כעת נניח שראינו שהקלף שהוצא הוא אדום. כיצד המידע הנוסף השפיע על ההסתברות הקודמות? ובכן, $\Pr(\text{diamond}|\text{red}) = \frac{13}{26}$ בעוד $\Pr(\text{king}|\text{red}) = \frac{2}{26} = \frac{1}{13}$. 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, כמוכן), בעוד המאורע 'מלך' אינו תלוי במאורע 'אדום'. ננסח זאת באופן מדויק.

הגדרה 2.14 (תלות בין מאורעות) צמד מאורעות A ו- B הם בלתי-תלויים אם $\Pr(A \cap B) = \Pr(A) \Pr(B)$. מנגד, המאורעות A ו- B הם תלויים אם הם לא בלתי-תלויים, ז"א אם $\Pr(A \cap B) \neq \Pr(A) \Pr(B)$.

ישירות מן ההגדרה נסיק כי אם המאורעות בלתי-תלויים, אזי קיום של אחד לא משפיע על ההסתברות של המאורע האחר. בכדי להוכיח זאת נבחן את ההסתברות המותנה $\Pr(A|B)$. אם המאורעות הם בלתי תלויים, אזי $\Pr(A|B) = \Pr(A)$. מסקנה - אם המאורעות A ו- B בלתי תלויים, אזי $\Pr(A|B) = \Pr(A)$. לחילופין, במילים פשוטות, ההסתברות של מאורע A איננה מושפעת מקיומו או אי-קיומו של B .

תרגיל 2.15 בספרייה ישנם 10 ספרי הסתברות. 5 ספרים הכוללים פתרונות ו-5 ללא פתרונות. הספרנית איבדה ספר אחד מתוך ה-10 (כלשהו). אלון מגיע לקחת ספר, מקבל ספר מתוך התשעה הנותרים באקראי ומחזיר לאחר יומיים. לאחר שבועיים מגיע בן ולוקח גם כן ספר 1 מתוך ה-9 ומחזיר לאחר יומיים. נגדיר את המאורעות A - הספר של אלון כלל פתרונות, ו- B - הספר של בן כלל פתרונות. האם A ו- B תלויים?

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

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

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

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

לכן המאורעות תלויים. ההסבר הוא פשוט יחסית, ניקח את הסיטואציה למצב קיצון ונניח שבין אלון ובן היו מגיעים עוד 1000 איש ומבצעים את הניסוי, והיינו שמים לב שרק כ- $\frac{4}{9}$ מתוכם קיבלו ספר עם פתרונות. במצב כזה היינו חושבים שכלל הנראה אבד ספר עם פתרונות בהתחלה. כדאי לשים לב כי $\Pr(A \cap B|C) = \Pr(A|C) \Pr(B|C)$. קרי, קיימת אי תלות תחת הידיעה לגבי זהות הספר האבוד. אי-תלות זאת נקראת אי-תלות מותנה ונראה אותה בהמשך.

הגדרה 2.16 קבוצה של מאורעות $A_1, \dots, A_n$ הם בלתי-תלויים אם $\Pr(\bigcap_{i=1}^n B_i) = \prod_{i=1}^n \Pr(B_i)$ לכל $B_i \in \{A_i, \bar{A}_i\}$. זאת אומרת, ישנן 2^n משוואות אשר אמורות להתקיים בכדי שהמאורעות יהיו ב"ת.

תרגיל 2.17 לבן יש מטבע כחול ולאלון מטבע אדום. למטבע הכחול הסתברות p לנחות על 'עץ' (H) ולמטבע האדום הסתברות q . אלון ובן משחקים את המשחק הבא: בכל תור, מטילים שניהם את המטבע שלהם באופן ב"ת באחר או בהטלות הקודמות. אם התקבל פעמיים H , אז הם מחליפים ביניהם את המטבעות. אחרת, כל אחד נשאר עם המטבע שהתחיל איתו את הסיבוב. התהליך חוזר על עצמו שוב ושוב. לכל $1 \leq n$, נגדיר את המאורע A_n כמאורע שבתום n סיבובים אלון מחזיק את המטבע האדום. חשבו את $\Pr(A_2)$, ומצאו עבור אילו ערכי p ו- q המאורעות A_2 ו- A_3 ב"ת.

פתרון: ננסה לחשוב מתי A_2 אפשרי. זהו המאורע בו אחרי 2 סיבובים אלון מחזיק את המטבע האדום איתו התחיל. דבר זה יכול להתרחש אם ורק אם לא הוחלפו מטבעות כלל, או לחילופין, התבצעו שתי החלפות. ההסתברות להחלפה של מטבע בסיבוב נתון היא pq . לכן נקבל ש- $\Pr(A_2) = (1 - pq)^2 + (pq)^2$. דרך אחת לבחון אי-תלות היא בעזרת ההגדרה $\Pr(A_2 \cap A_3) = \Pr(A_2) \Pr(A_3)$. מנגד, ישנה עוד דרך להראות אי-תלות וזאת על ידי השוואת $\Pr(A_3|A_2) = \Pr(A_3)$. מדוע זה אפשרי? ובכן, לפי ההגדרה, נובע ש-

$\Pr(A_3|A_2) = \frac{\Pr(A_3 \cap A_2)}{\Pr(A_2)}$, ואם בנוסף אנו יודעים ש- $\Pr(A_3|A_2) = P(A_3)$, אזי נובע ש- $\frac{\Pr(A_3 \cap A_2)}{\Pr(A_2)} = \Pr(A_3)$ כנדרש. נחשב מפורשות את $\Pr(A_3|A_2)$ ואת $P(A_3)$. נסמן $s = pq$ ונקבל $\Pr(A_3|A_2) = 1 - s$, $P(A_3) = (1 - s)^3 + 3(1 - s)s^2$. לכן, נבדוק מתי מתקיים $1 - s = (1 - s)^3 + 3(1 - s)s^2$. פתרון ראשון הוא $s_1 = 1$. אחרת,

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

$$0 = -2s + 4s^2$$

$$0 = s(2s - 1)$$

ולכן נקבל את הפתרונות $s_1 = 1$, $s_2 = 0$, $s_3 = \frac{1}{2}$. נתרגם זאת למונחי p ו- q באופן הבא: $s = \frac{1}{2} \Leftrightarrow p = \frac{1}{2q}$, ואלו המקרים בהם ישנה אי-תלות. התוצאה שקיבלנו היא די הגיונית ואפילו אינטואיטיבית. כאשר $pq = 0$ או $pq = 1$ אז מקבלים שהמאורעות A_2 ו- A_3 מתרחשים בהסתברות 0 או 1 ולכן הם ב"ת בכל מאורע אחר. בנוסף, כאשר $pq = \frac{1}{2}$, אז בכל שלב יש סיכוי שווה להיות בכל מצב ללא תלות בעבר. למעשה, ישנה סימטריה בבעיה מבחינת מיקום המטבעות אצל אלון ובן ולכן מתקיימת אי-תלות.

תרגיל 2.18 על השולחן מונחים 2 כדים: בכד הראשון (כד א') שלושה כדורים לבנים וחמישה כדורים אדומים; בכד השני (כד ב') ישנם שישה כדורים לבנים ושלושה כדורים אדומים. אלון מטיל מטבע הוגן. אם תוצאת ההטלה היא "עץ", הוא יוציא מכד א' 2 כדורים עם החזרה. אחרת, הוא יוצא מכד ב' 2 כדורים עם החזרה. נסמן ב- A את המאורע שהכדור הראשון לבן, ב- B את המאורע שהכדור השני לבן וב- C את המאורע בו נבחר כד א'. האם המאורעות A, B תלויים? האם המאורעות A, B תלויים בהינתן C ?

פתרון: נתחיל בבניית מרחב המדגם. נסמן את הכדורים בכד א' במספרים $1, \dots, 8$ כאשר שלושת הראשונים יהיו הכדורים הלבנים והיתר אדומים. באופן דומה נסמן את הכדורים בכד ב' במספרים $1, \dots, 9$ כאשר ששת הראשונים לבנים והיתר אדומים. מרחב המדגם מורכב משלוש מאורעות: $\Omega = 8^2 + 9^2 = 145$ אפשרויות שונות. בנוסף נשים לב שמערכת המדגם אינו מרחב מדגם שווה הסתברות. לדוגמא:

$$\frac{1}{2} \cdot \frac{1}{81} \text{ היא } \frac{1}{2} \cdot \frac{1}{64} \text{ בעוד ההסתברות להוציא צמד כדורים מכד ב' היא } \frac{1}{2} \cdot \frac{1}{81}.$$

נחשב את ההסתברות של כל מאורע בנפרד. $\Pr(A) = \frac{1}{2} \cdot \frac{3 \cdot 8}{64} + \frac{1}{2} \cdot \frac{6 \cdot 9}{81} = \frac{3}{16} + \frac{3}{9} = \frac{75}{144}$. נשכח, הסתברות $\frac{1}{2}$ לכל כד, ולאחר מכן ספרנו את מספר האפשרויות בכל חלק ממרחב המדגם. באופן זה לחישוב הקודם, נקבל ש- $\Pr(A \cap B) = \frac{1}{2} \cdot \frac{3 \cdot 3}{64} + \frac{1}{2} \cdot \frac{6 \cdot 6}{81} = \frac{9}{128} + \frac{2}{9} = \frac{337}{1152}$. נשים לב ש- $\Pr(A) \Pr(B) \neq \Pr(A \cap B)$ ולכן המאורעות תלויים.

נעבור לחישוב השני בשאלה. נחשב את ההסתברות של כל מאורע מותנה בנפרד את ההסתברות של חיתוך המאורעות תחת ההתניה $\Pr(A|C) = \frac{3 \cdot 8}{64} = \frac{3}{8} = \Pr(B|C)$ ובנוסף, $\Pr(A \cap B|C) = \frac{3 \cdot 3}{64} = \frac{9}{64} = \Pr(A|C) \Pr(B|C)$, לכן המאורעות הללו בלתי-תלויים בהינתן C .

תרגיל 2.19 נקבע $1 \leq k \leq n$ מספרים שלמים. בספארי נולדו לקרנף n גורים אחד אחרי השני כאשר כל גור יכל להיוולד לכל מין בהסתברות שווה ובאופן ב"ת במין הגורים האחרים. דוליטול מסתובב בגן החיות ופוגש גור שנבחר באקראי באופן אחיד מבין הגורים. נסמן ב- A את המאורע בו k הגורים הראשונים שנולדו הם ממין זכר. נסמן ב- B את המאורע שנולדו לפחות k גורים ממין זכר. נסמן ב- C את המאורע בו דוליטול פוגש גור ממין זכר. עבור $k = 1, n = 2$, חשבו את $\Pr(C|B)$. עבור $k = 4, n = 12$, חשבו את $\Pr(A|C)$.

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

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

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

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

3 משתנים מקריים בדידים

החישובים שלנו עד כה עסקו במרחבי הסתברות ובמאורעות. זאת הטרימינולוגיה היחידה שנדרשנו לה. בפרק הנוכחי אנחנו מרחיבים את ההגדרות ועוברים לדון במבנים יותר מורכבים על מרחבי ההסתברות שאינם רק מאורעות. למעשה, אנחנו נדון בפונקציות על מרחבי ההסתברות. למה הכוונה? במציאות היומיומית מתקבל מידע רב מכל ניסוי נתון. לדוגמא, אם נבחן את מאגר המידע של הלמ"ס אשר נוצר מסקר אוכלוסיה כלשהו, כנראה נזהה פרמטרים רבים כגון- הכנסות משקי הבית, הוצאות משקי הבית, פריון, מספר חדרים למשק בית, וכן הלאה. על כן, עבור כל ניסוי נתון נוכל להגדיר פונקציה אשר מחזירה את המידע שמעניין אותנו. אותה הפונקציה, הנקראת משתנה מקרי ותסומן לרוב באות לטינית גדולה (לרוב X , אבל לא רק), תשתנה בהתאם לתוצאות הניסוי.

לדוגמא, בהטלת צמד קוביות נגדיר את X להיות סכום התוצאות. כעת נוכל לשאול לא מעט שאלות בנוגע למשתנה המקרי X . אילו ערכים הוא מקבל? (קרי, לאילו ערכים הוא יוכל להיות שווה?) מה הסיכוי לקבל ערך מקסימאלי? או ערך מינימאלי? ועוד. נשים לב ש- X הוא פונקציה ממרחב המדגם Ω ל- $\mathbb{R}$. זאת אומרת, $X(\omega) \in \mathbb{R}$ עבור כל תוצאה במרחב המדגם ω . במקרה לעיל המשתנה המקרי X מקבל זוגות (i, j) , כאשר i, j הם מספרים שלמים בין 1 ל-6, ומחזיר מספר ממשי שהוא סכום הערכים שקיבל.

השם **משתנה מקרי** לעיתים מטעה. צריך לזכור שמדובר בפונקציה לכל דבר ועניין. מצד אחד, הפונקציה משתנה בהתאם לתוצאות הניסוי, ומצד שני תוצאת הניסוי היא סטוכסטית (אקראית) ולכן השינוי הוא מקרי. בשלב הראשון אנחנו נעסוק רק במשתנים מקריים בדידים וחד-מימדים. זאת אומרת, משתנים מקריים שמקבלים (קרי, שווים ל-) ערכים בדידים (בניגוד למשתנים מקריים רציפים שיכולים לקבל רצף של ערכים), ובנוסף כל משתנה מחזיר מספר בודד, $X(\omega) \in \mathbb{R}$, ולא וקטור כלשהו.

הגדרה 3.1 (משתנה מקרי בדיד) משתנה מקרי בדיד (מ"מ) הוא פונקציה המתאימה ערך מספרי יחיד לתוצאות במרחב הסתברות, $X : \Omega \rightarrow \mathbb{R}$.

3.1 תומך והתפלגות

לאחר שהגדרנו את המשתנה המקרי, נעבור לדון בתכונות שלו. השאלה הראשונה שתעסיק אותנו היא הערכים שהמשתנה יכול לקבל. לדוגמא, בהטלת צמד קוביות ו- X אשר מוגדר להיות סכום הטלות, הערכים שיכולים להתקבל הם 2, 3, ..., 11, 12. הערכים הללו נקראים **התומך (Support) של X** . כדוגמא נוספת ניקח את המ"מ Y להיות הערך המקסימאלי מבין צמד ההטלות. התומך של Y יהיה 1, 2, 3, 4, 5, 6 שכן כל קוביה מקבלת ערכים רק בין 1 ל-6.

הערה 3.2 (תומך של מ"מ בדיד) התומך של משתנה מקרי בדיד X זו קבוצת המספרים ש- X יכול להיות שווה להם בהסתברות חיובית, קרי $\text{Supp}(X) = \{k \in \mathbb{R} : \Pr(X = k) > 0\}$.

לאחר הגדרת התומך, נתעניין בסיכוי לקבל כל ערך בתומך. לשם כך מגדירים פונקציה חדשה שנקראת - **ההתפלגות (Distribution) של X** . ההתפלגות P_X של משתנה מקרי X זו פונקציה ממשיית (ז"א, היא מקבל מספר ומחזירה מספר), ומוגדרת עבור כל משתנה מקרי X באופן הבא: $P_X(k) = \Pr(X = k)$. זאת אומרת, ההתפלגות של X בנקודה k מציינת את ההסתברות של המאורע $\{X = k\}$. למעשה, פונקציית התפלגות ופונקציית הסתברות הן די דומות מבחינת התכונות, אבל כדאי לא להתבלבל ביניהן. פונקציית הסתברות מוגדרת על מרחבי מדגם ולכן הקלט של פונקציית ההסתברות יכול להיות מאוד מורכב. מנגד, התפלגויות מקבלות מספרים ומחזירות מספרים. אלו פונקציות ממשיות כמו כל הפונקציות שראיתם בלימודים הטרנס אקדמיים.

משתנים מקריים שונים על מרחבים שונים יכולים להיות בעלי אותה התפלגות. לדוגמא, ניקח מטבע הוגן, שיכול ליפול בסיכוי שווה על 0 או על 1, ונגדיר את X להיות תוצאת ההטלה. מנגד, נטיל קוביה ונגדיר את המשתנה Y להיות שווה ל-1 אם התוצאה היא זוגית, ו-0 אחרת. אומנם צמד המשתנים מוגדרים על מרחבים שונים (הטלה מטבע מול הטלת קוביה), אבל יש להם את אותה ההתפלגות, שכן $P_X(1) = P_Y(1) = \frac{1}{2}$ ו- $P_X(0) = P_Y(0) = \frac{1}{2}$.

התכונה הבסיסית ביותר של פונקציית ההתפלגות היא אי-שליליות. תכונה זאת היא מיידית מכך שהסתברות לעולם איננה שלילית. בנוסף, ההתפלגות היא מנורמלת בכך ש- $\sum_{k \in \text{Supp}(X)} P_X(k) = 1$. במילים, סכימה של ההתפלגות על כל ערכי התומך תניב 1, כשם

שסכום ההסתברויות תמיד יהיה 1 כאשר סוכמים על כל האפשרויות במרחב. נשים לב שהתפלגויות ניתן להציג כנוסחה או בטבלה. לדוגמא, אם $X(\omega) \in \{1, \dots, n\}$, אז ההתפלגות של X נתונה על ידי רשימת ההסתברויות $\Pr(X=1), \Pr(X=2), \dots, \Pr(X=n)$.

הערה 3.3 בדרך כלל אנו נשתמש בכתוב המקוצר $\{X=k\}$ בכדי לציין את המאורע בו המשתנה המקרי X מקבל את הערך k . יחד עם זאת, חשוב להבין את המשמעות האמיתית שמאחורי סימון זה. מדובר פה בכל המצבים האפשריים $\{\omega_i\}$ במרחב המדגם כך ש- $X(\omega_i) = k$ או במדויק $\{X=k\} = \{\omega \in \Omega | X(\omega) = k\}$. אך בכדי לקצר את הרישום, אנו נשתמש לרוב בסימון המקוצר.

דוגמה 3.4 נפתח בדוגמא פשוטה. נטיל קוביה הוגנת ונניח כי X מחזיר את תוצאת ההטלה. אזי $Supp(X) = \{1, 2, 3, 4, 5, 6\}$ וההתפלגות של X היא $P_X(k) = \frac{1}{6}$ לכל k בתומך, ו-0 אחרת. חשוב לזכור שהתפלגות היא פונקציה ולכן אנחנו צריכים לתאר אותה בכל נקודה. משמע, זה לא מספיק לציין את ערכי הפונקציה רק על התומך, אלא בכל נקודה $k \in \mathbb{R}$.

נמשיך לדוגמא יותר מורכבת. אלון הוא סוחר ממולח בשוק ההון. הוא משקיע סכום של \$1000 במכשיר פיננסי אשר בכל שנה מכפיל את הכסף בסיכוי $\frac{2}{3}$, או מפסיד חצי מההשקעה בסיכוי $\frac{1}{3}$. השנים הן בלתי תלויות. נגדיר את X להיות שווי ההשקעה של אלון לאחר שנתיים. התומך של X כעת הוא $\{250, 100, 4000\}$ שכן ייתכנו שנתיים רצופות של הפסדים, שנתיים רצופות של רווח, ושנתיים בהן היה רווח והפסד לסירוגין. על כן, ההתפלגות של X היא $P_X(250) = \frac{1}{9}, P_X(4000) = \frac{4}{9}, P_X(1000) = \frac{4}{9}$, ו-0 אחרת.

תרגיל 3.5 מטילים 3 קוביות הוגנות. נגדיר משתנה מקרי X להיות התוצאה המקסימלית מבין ההטלות. מהי התפלגות X ?

פתרון: תחילה נציין כי התומך הוא $Supp(X) = \{1, 2, 3, 4, 5, 6\}$, לכן נרצה למצוא את $\Pr(X=k)$ עבור $k \in \{1, \dots, 6\}$. עבור $k=1$, המאורע היחידי שייתן את הנדרש הוא המאורע בו כל ההטלות הניבו 1, לכן $\Pr(X=1) = \frac{1}{216}$. עבור $k=2$, המאורעות המתאימים הם אלו בהם נקבל לפחות פעם אחת 2, אבל לא ערך גבוה יותר. מספירה מהירה של האפשרויות מגיעים ל- $\Pr(X=2) = \frac{7}{216}$. ברגע שנרצה לעבור לערך הבא נגלה שמספר האפשרויות הוא מסובך לחישוב. לכן ננסה למצוא טכניקה יותר פשוטה. מאחר ו- X הוא מקסימום על שלושת ההטלות, קל יותר לחשב את ההסתברות שהוא קטן או שווה לערך כלשהו מאשר רק שווה לערך כלשהו. למה הכוונה? המאורע $\{X \leq k\}$ שקול לכך שכל הטלה בנפרד קטנה או שווה ל- k . זאת אומרת, המאורע $\{X \leq k\}$ מתקיים אם ורק אם כל ההטלות קטנות או שוות ל- k . במילים אחרות, לכל הטלה יש k אפשרויות, ולכן $\Pr(X \leq k) = \frac{k^3}{216}$. כתוצאה מכך

$$\begin{aligned} \Pr(X=k) &= \Pr(X \leq k) - \Pr(X \leq k-1) = \\ &= \frac{k^3}{6^3} - \frac{(k-1)^3}{6^3} = \\ &= \frac{k^3 - (k^3 - 3k^2 + 3k - 1)}{216} = \frac{3k^2 - 3k + 1}{216}. \end{aligned}$$

המעבר הראשון מדויק כי ההפרש מקז את כל המקרים הלא רלוונטיים עבורנו, בהם $X \leq k-1$, ואנו נותרים רק עם המקרים בהם $X=k$.

תרגיל 3.6 מטילים 2 קוביות. נסמן ב- X את סכום התוצאות שהתקבלו. מצאו את ההתפלגות של X .

פתרון: נתחיל מהתומך - התוצאות האפשרויות בסכום של 2 קוביות הן המספרים הטבעיים מ-2 עד 12. בכדי לחשב את ההסתברות לכל ערך נשתמש במרחב מדגם סימטרי: גודל מרחב המדגם הוא $6^2 = 36$, ככמות האפשרויות לבחור תוצאה לקוביה הראשונה מתוך 6 ותוצאה לקוביה השנייה מתוך 6. נחשב את ההסתברות של כל תוצאה בנפרד:

$$\begin{aligned}\Pr(X=2) &= \frac{|\{(1,1)\}|}{6^2} = \frac{1}{36} \\ \Pr(X=3) &= \frac{|\{(1,2), (2,1)\}|}{6^2} = \frac{2}{36} \\ \Pr(X=4) &= \frac{|\{(1,3), (2,2), (3,1)\}|}{6^2} = \frac{3}{36}\end{aligned}$$

ובאופן דומה ממשיכים לחשב את האפשרויות לכל הערכים האפשריים בתומך. באופן כללי קל לראות שהנוסחה היא $\Pr(X=k) = \frac{6-|7-k|}{36}$.

3.1.1 פונקצית התפלגות מצטברת

בתרגיל 3.5 השתמשנו בביטוי $\Pr(X \leq k)$ בכדי לחשב את ההתפלגות של משתנה מקרי כלשהו. זה לא היה במקרה, שכן הביטוי $\Pr(X \leq k)$ הוא פונקציה חשובה בהקשר של משתנים מקריים. עבור כל משתנה מקרי X אנו מגדירים פונקציה ממשית, בנוסף להתפלגות, שנקראת פונקצית התפלגות מצטברת, ומסומנת ב- F_X . **פונקצית ההתפלגות המצטברת (פה"מ)** מוגדרת ע"י $F_X(k) = \Pr(X \leq k)$ בכל נקודה על הציר הממשי, ושמה מרמז על התכונה המרכזית שלה - היא צוברת את כל ההתפלגות מן הערך המינימאלי עד לנקודה k . לדוגמא, נניח כי X הוא משתנה מקרי בדיד הנקבע על ידי הטלת קוביה. מה הערך של הפה"מ בנקודה $k = 2.3$? בכדי לענות על השאלה, נחזור להגדרה. $F_X(2.3) = \Pr(X \leq 2.3) = \frac{2}{6}$ והערך של הפה"מ הוא $\frac{2}{6}$ כי ההתפלגות הנצברת עד לנקודה 2.3 היא $\frac{2}{6}$ (בפרט, ההתפלגות היא $\frac{1}{6}$ בנקודה 1 ו- $\frac{1}{6}$ בנקודה 2). הדוגמא הזאת מרמזת לנו על תכונה חשובה של פונקצית ההתפלגות המצטברת - **מונוטוניות לא יורדת**. הפה"מ היא פונקציה לא יורדת, ובמקרה הבדיד נקבל פונקצית מדרגות, כפי שנראה בהמשך. בנוסף, הפה"מ היא **אי-שלילית**, וכן $\lim_{k \rightarrow -\infty} F_X(k) = 0$, $\lim_{k \rightarrow \infty} F_X(k) = 1$, כתוצאה ישירה מן ההגדרה. לסיכום, התכונות העיקריות של פונקצית התפלגות מצטברת הן: (1) פונקציה מונוטונית ולא יורדת; (2) עולה מ-0 עד ל-1; (3) רציפה מימין.

תרגיל 3.7 רשמו את פונקצית ההתפלגות המצטברת של המשתנה המקרי המוגדר בתרגיל 3.5.

פתרון: נרשום את פונקצית ההתפלגות המצטברת בצורה מפורשת בהתאם למה שחישבנו קודם לכן.

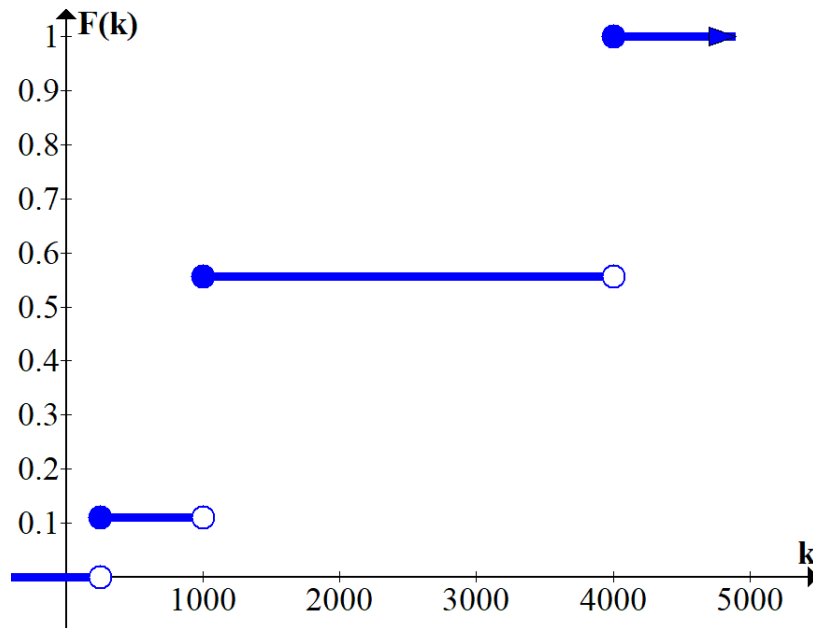
$$F_X(k) = \begin{cases} 0, & k < 1, \\ \frac{|k|^3}{216}, & 1 \leq k < 6, \\ 1, & 6 \leq k. \end{cases}$$

הסימון $[k]$ מציין עיגול של המספר כלפי מטה, למספר השלם הקרוב ביותר. מכאן נדגיש שוב שפונקצית ההתפלגות המצטברת מוגדרת בכל נקודה על הישר ולא רק על התומך של X !

תרגיל 3.8 אלון הוא סוחר ממולח בשוק ההון. הוא משקיע סכום של \$1000 במכשיר פיננסי אשר בכל שנה מכפיל את הכסף בסיכוי $\frac{2}{3}$, או מפסיד חצי מההשקעה בסיכוי $\frac{1}{3}$. השנים הן בלתי תלויות. נגדיר את X להיות שווי ההשקעה של אלון לאחר שנתיים. מצאו את פונקצית ההתפלגות המצטברת של X וציירו אותה.

פתרון: כפי שציינו קודם לכן, התומך של X כעת הוא $\{250, 100, 4000\}$ שכן ייתכנו שנתיים רצופות של הפסדים, שנתיים רצופות של רווח, ושנתיים בהן היה רווח והפסד לסירוגין. על כן, פונקצית ההתפלגות המצטברת של X היא

$$F_X(k) = \Pr(X \leq k) = \begin{cases} 0, & k < 250, \\ \frac{1}{9}, & 250 \leq k < 1000, \\ \frac{5}{9}, & 1000 \leq k < 4000, \\ 1, & 4000 \leq k. \end{cases}$$



איור 8: שרטוט פונקציית התפלגות מצטברת עבור תרגיל 3.8

3.2 תוחלת

בפרק הקודם עסקנו במשתנים מקריים והתפלגויות. ראינו כי לכל משתנה מקרי יש התפלגות המפרטת את ערכיו וההסתברויות המתאימות, על כן ידיעה של ההתפלגות מספקת את כל המידע הדרוש עבור משתנה מקרי יחיד כלשהו. יחד עם זאת, כאשר בוחנים מספר משתנים מקריים בו זמנית, ההתייחסות לכל אחת מההתפלגויות ביחס לאחרות יכולה להיות בעייתית ומורכבת. בנוסף, בעולם האמיתי איננו יודעים במדויק את ההתפלגויות של המשתנים הסובבים אותנו, לכן ישנן דרכים פשוטות יותר לקטלג משתנים מקריים - באמצעות מדדים.

באופן כללי, מדדים באשר הם מספקים לנו מידע חשוב (אך מוגבל) אודות הגורמים אותם הם מודדים. לדוגמא, מדד המחירים לצרכן משמש למדידת קצב האינפלציה, קרי קצב עליית המחירים, במדינת ישראל. המדד כמובן אינו מדויק באופן מוחלט, ולכן המידע שהוא מספק לנו הוא חלקי מצד אחד, אך מייצג, מצד שני. אם מדד המחירים לצרכן מצביע על עלייה של 1% ברמות המחירים, אין הדבר אומר שכל המוצרים התייקרו באחוז בודד. חלק מן המוצרים דווקא הוזלו וחלק התייקרו בשיעורים שונים, אך בממוצע רמות המחירים, עבור סל צריכה נתון, התייקרו באחוז אחד. למה הכוונה ב"סל צריכה נתון"? ובכן זה החלק החשוב לענייננו.

נניח ובמדינה א' צורכים אך ורק תפוחים. סל הצריכה של המדינה הזו אינו מושפע כלל ממחירי הדירות, מחירי המכוניות, או מחירי הטלפונים הסלולריים, שכן אזרחי מדינה א' לא צורכים מוצרים אלו, הם צורכים רק תפוחים. משמע, עלייה של חצי אחוז במחירי התפוחים תתבטא בעלייה של חצי אחוז במדד המחירים ובהוצאה של כל נפש, בניגוד לעלייה של חצי אחוז במחירי העגבניות שלא תשפיע בכלל

על מדד המחירים במדינה א'.¹¹ לעומת זאת, במדינה ב' הצריכה של כל משק בית מבוססת על 80% תפוחים ו-20% בננות. במקרה שכזה, עלייה במחירי מכוניות ונדל"ן עדיין לא תשפיע על מדד המחירים לצרכן שכן אזרחי מדינה ב' לא קונים מוצרים כאלה. מצד שני, עלייה של 5% במחירי הבננות תתבטא בעלייה של אחוז אחד במדד המחירים לצרכן, כי 20% מהסל התייקר בחמישה אחוזים ונקבל $101 = 1.05 \cdot \frac{20}{100} + 1 \cdot \frac{80}{100}$. זאת אומרת שעלייה של חלק קטן מהסל בשיעור גבוה, העלתה את המחיר הכולל של הסל בשיעור נמוך יחסית. לעומת זאת, עלייה של מחירי התפוחים ב-5% הייתה מתבטאת בעלייה של הסל כולו ב-4%, $104 = 1.05 \cdot \frac{80}{100} + 1 \cdot \frac{20}{100}$. זו תוצאה צפויה - מחירי התפוחים משפיעים בצורה יותר משמעותית על מדד המחירים, כי מרבית הצריכה מבוססת תפוחים.

החישובים שהצגנו לעיל הם ממוצעים משוקללים של מחירי מוצרים, וזה גם הרעיון שעומד בבסיס המדד שנקרא "תוחלת". התוחלת היא מדד אמצעי, לכן היא מעניקה לנו אינדיקציה מאוד בסיסית בנוגע לערך הממוצע של משתנה מקרי נתון. התוחלת לוקחת את הערכים השונים שמקבל המשתנה המקרי ומשקללת אותם, בהתאם לסבירות של כל ערך, לכדי מספר בודד שמציין את הערך הממוצע של אותו המשתנה. קרי, בדומה למדד המחירים לצרכן אשר משקלל את עלות סל צריכה בהתאם לצריכה, התוחלת היא ממוצע משוקלל של ערכי המשתנה המקרי בהתאם לסבירות של הערכים. בפועל, החישוב של התוחלת הוא מיידי: לוקחים את כל הערכים שמקבל המ"מ X , מכפילים כל ערך בהסתברותו וסוכמים.

הגדרה 3.9 (תוחלת) התוחלת של משתנה מקרי X היא $E[X] = \sum_{\omega \in \Omega} X(\omega) \Pr(\omega)$. נהוג לסמן את התוחלת באות היוונית μ (מיו, mu).

ההגדרה הפורמאלית לתוחלת היא מדויקת, אבל לא ממש נוחה לשימוש, מכיוון שהיא מצריכה לסכום על פני מרחב המדגם. לכן ישנה גרסה יותר פשוטה ונוחה לחישוב תוחלת המתבססת אך ורק על ההתפלגות של המשתנה הנדון.¹² באמצעות שינוי סדר הסכימה בהגדרה, נוכל לרשום את התוחלת כ-

$$E[X] = \sum_{k \in \text{Supp}(X)} k \Pr(X = k) = \sum_k k P_X(k),$$

ונקבל כי התוחלת התלויה אך ורק בהתפלגות של המשתנה המקרי. נציין כי האות E נובעת משמה הלועזי של התוחלת - expectation. נתחיל בחישוב תוחלת עבור דוגמאות פשוטות. התוחלת של הטלת קוביה היא $E[X] = \sum_{k=1}^6 k \cdot \frac{1}{6} = 3.5$, כאשר נעזרנו רק בהתפלגות של המשתנה ללא צורך בתיאור מרחב המדגם כולו. החישוב האחרון הוא טריוויאלי, אך מעלה נקודה חשובה - התוחלת היא ערך מספרי שאינו בהכרח נמצא בתומך של המשתנה X . הטלת קוביה לא יכולה להסתיים ב-3.5, ויחד עם זאת התוחלת של הטלת קובייה היא 3.5. אם כן מה המשמעות של המספר 3.5? אם נטיל את הקובייה שוב ושוב מספיק פעמים, אז נקבל כל ערך שלם מ-1 עד 6, בערך שישית מהפעמים. לכן הערך הממוצע שנקבל יהיה 3.5. זאת הסיבה לכך שמתייחסים לתוחלת בתור "הניבוי הטוב ביותר עבור המשתנה המקרי".

נמשיך עם דוגמא נוספת. בן הוא סוחר מתוחכם ולכן הוא מציע לאלון את ההצעה הבאה: הפקד ברשותי 100 שקלים בתחילת השנה. בסוף השנה, בסיכוי $\frac{2}{5}$ הסכום יוכפל, ובסיכוי $\frac{3}{5}$ הסכום יצטמצם בחצי. במילים אחרות, בן מציע לאלון הגרלה המניבה את הפרס 100 בסיכוי $\frac{2}{5}$ ואת "הפרס" -50 בסיכוי $\frac{3}{5}$. מה תהיה תוחלת הכסף של אלון לאחר קבלת ההצעה? $E[X] = 200 \cdot \frac{2}{5} + 50 \cdot \frac{3}{5} = 110$ זאת אומרת שהצעתו של בן מגלמת תוחלת רווח של 10%! (כמובן שהסיכון בהתאם...) בהזדמנות הזאת כדאי לציין שהתוחלת של מספר קבוע a היא a , מאחר ומספר קבוע a הוא גם משתנה מקרי שמקבל את הערך a בהסתברות 1.

תרגיל 3.10 חברה קידוחים מפתחת 2 שדות גז, בתקווה לאתר נפט. מלכתחילה, כל שדה מכיל נפט, בנוסף לגז הקיים, בסיכוי $\frac{1}{5}$. עלות הפיתוח של כל שדה היא 2.5 מיליארד ש"ח, הרווח מהפקת הגז הוא 2 מיליארד ש"ח בכל שדה, הרווח מהפקת הנפט (אם יימצא כזה) הוא 3 מיליארד ש"ח נוספים לשדה. נניח כי X הוא הרווח הנקי של החברה. חשבו את $E[X]$.

פתרון: בסיכוי $\frac{16}{25}$ לא יימצא כלל נפט, והחברה תרשום הפסד של מיליארד ש"ח, שכן עלויות הפיתוח הן 5 מיליארד והרווח מהפקת הגז הוא 4 מיליארד ש"ח. בסיכוי $\frac{1}{25}$ יימצא נפט בשתי השדות, ואז הרווח הנקי של החברה יהיה $5 = 5 - 2 + (3 + 2)$ מיליארד ש"ח. בסיכוי $\frac{8}{25}$ יימצא נפט רק בשדה אחד והרווח הנקי של החברה יהיה $2 = 5 - 2 + 5$ מיליארד ש"ח. לכן, $E[X] = \frac{16}{25} \cdot (-1) + \frac{8}{25} \cdot 2 + \frac{1}{25} \cdot 5 = 0.2$. קרי, תוחלת הרווח הנקי של החברה הוא 200 מיליון ש"ח (עם סיכון לא קטן).

¹¹לטובת הדוגמא אנו מתעלמים ממספר בעיות כלכליות ברורות, כגון: כיצד נקבעים המחירים במדינה שהמוצר לא נצרך בה, וכן הלאה.
¹²ישנה הנחה סמויה באמור לעיל והיא שהתוחלת מוגדרת. ישנם מצבים מתמטיים בהם התוחלת לא תהיה מוגדרת, כאשר עוסקים בסכומים אינסופיים (חיובים ושיליים יחדיו). בקורס הנוכחי אנחנו לא נפרט את המצבים הללו, אך כדאי לזכור את האפשרות הזאת להמשך.

לאור התרגיל האחרון כדאי לשים לב להבדל הבסיסי בין ממוצע לתוחלת. ממוצע פשוט מחושב על ידי סכימת כל הערכים וחלוקה במספר הערכים, ללא התחשבות בסיכוי של כל ערך. לכן, הממוצע הפשוט של X יהיה $\frac{5+2-1}{3} = 2$. מצד שני, התוחלת לוקחת בחשבון את הסיכוי של כל אחד מהערכים להתקבל, ולכן מתקבלת התוצאה של 0.2 מיליארד (ולא 2 מיליארד).

3.2.1 לינאריות התוחלת

תוחלת ניתן לחשב במגוון דרכים, ורובן מתבססות על אותה תכונה חיונית לינאריות. עבור כל משתנה מקרי X וקבוע $a \in \mathbb{R}$, התוחלת של aX היא $E[aX] = aE[X]$. בנוסף, עבור כל צמד משתנים מקריים X ו- Y , התוחלת של $X + Y$ היא $E[X + Y] = E[X] + E[Y]$. כמוכן, שהתכונה הנ"ל ניתנת להכללה גם עבור מספר כללי של משתנים מקריים, שכן $E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i]$.

תרגיל 3.11 בחנות ירקות קטנה ישנם 3 תפוחים ו-4 אגסים. אלון קונה 2 פירות באקראי. בן המוכר מרוויח על כל תפוח 2 ש"ח ועל כל אגס 4 ש"ח. מצאו את תוחלת הרווח של בן.

פתרון: נגדיר מ"מ Q עבור הרווח של בן משתנה מקרי X כמספר התפוחים שאלון קנה. על כן, $Q = 2X + 4(2 - X) = 8 - 2X$. התוחלת של X יחסית קלה לחישוב שכן $\Pr(X=0) = \frac{4}{7} \cdot \frac{3}{6} = \frac{2}{7}$, $\Pr(X=1) = \frac{4}{7} \cdot \frac{2}{6} = \frac{1}{7}$, $\Pr(X=2) = \frac{3}{7} \cdot \frac{2}{6} = \frac{1}{7}$. מכאן נובע ש- $E[X] = 0 \cdot \frac{2}{7} + 1 \cdot \frac{4}{7} + 2 \cdot \frac{1}{7} = \frac{6}{7}$. נשתמש בתכונת הלינאריות לחישוב התוחלת של Q . $E[Q] = 8 - 2E[X] = 8 - \frac{12}{7} = \frac{44}{7}$.

תרגיל 3.12 בוחרים קוד באורך 10 תווים, כאשר כל תו נבחר באקראי מתוך הספרות 0–9 ללא קשר לתווים האחרים שנבחרו. נסמן ב- X מ"מ הסופרים את מספר הרצפים 11 בקוד. חשבו את $E[X]$.

פתרון: במידה ואנו מעוניינים למצוא את התוחלת של X על פי ההתפלגות שלו, נצטרך תחילה למצוא את ההתפלגות הזאת. הדבר אומנם אפשרי, אך יחד עם זאת מאוד מסובך. לכן בד"כ ננסה למצוא דרך יותר קלה לחישוב התוחלת, כגון פירוק המשתנה לסכום משתנים פשוטים יותר. נגדיר 9 מ"מ X_i , $i = 1, \dots, 9$ כאשר X_i מקבל את הערך "1" אם ורק אם במקום ה- i מתחיל רצף של 11. לכן $X = \sum_{i=1}^9 X_i$.

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

3.2.2 תוחלת פונקציה של מ"מ

נחזור חזרה לסוחר המתוחכם בן ולהצעה שהוא מציע לאלון (הגרלה אשר מניבה רווח של 100 בסיכוי $\frac{2}{5}$ או "רווח" של -50 בסיכוי $\frac{3}{5}$) וננסה לחשב את $E[X^2]$. שימו לב, לא מדובר ב- $(E[X])^2$, שכן את התוחלת כבר חישובנו וקיבלנו תוצאה של 110, לכן העלאת התוחלת בריבוע איננה מסובכת. אנחנו מעוניינים לחשב את התוחלת של משתנה חדש, הנתון על ידי X^2 . נוכל למצוא את התפלגותו ולקבל כי $\Pr(X^2 = 10000) = \Pr(X = 100) = \frac{2}{5}$ בעוד $\Pr(X^2 = 2500) = \Pr(X = -50) = \frac{3}{5}$. על כן, לפי הגדרת התוחלת נקבל $E[X^2] = 2500 \cdot \frac{3}{5} + 10000 \cdot \frac{2}{5} = 5500 \neq 110^2$.

החישוב הנ"ל הוא נקודתי ושאלת השאלה - כיצד ניתן לייצר נוסחה כללית למקרים בהם איננו מעוניינים לחשב תוחלת של משתנה מקרי נתון, אלא מעוניינים לחשב תוחלת של פונקציה של אותו המשתנה המקרי? (במקרה לעיל הפונקציה היא העלאה בחזקה שנייה). מסתבר שהתשובה לכך היא יחסית פשוטה. ניתן להשתמש בהתפלגות של המשתנה המקורי ולהגיע לתשובה בקלות יחסית. באופן פורמאלי,

נניח כי X הוא משתנה מקרי כלשהו וניקח פונקציה ממשית $g: \mathbb{R} \rightarrow \mathbb{R}$. אזי $E(g(X)) = \sum_{k \in \text{Supp}(X)} g(k) \Pr(X = k)$. שימו לב שאנו נעזרים בהתפלגות ובתומך של X בכדי לחשב את התוחלת של $g(X)$, אשר ייתכן והיא משתנה מקרי אחר לחלוטין. כאמור, נשים לב שהשיויון $E(g(X)) = g(E(x))$ לרוב לא יהיה נכון, כפי שראינו בדוגמא הקודמת.

תרגיל 3.13 ניקח משתנה מקרי X בעל התפלגות

$$\Pr(X = k) = \begin{cases} \frac{1}{3}, & k = -1, \\ \frac{1}{6}, & k = 0, \\ \frac{1}{2}, & k = 1. \end{cases}$$

חשבו את $E[X]$ ואת $E[X^4]$.

פתרון: נשתמש בהגדרת התוחלת ונקבל

$$E[X] = (-1) \cdot \frac{1}{3} + 0 \cdot \frac{1}{6} + 1 \cdot \frac{1}{2} = \frac{1}{6},$$

$$E[X^2] = (-1)^4 \cdot \frac{1}{3} + 0^4 \cdot \frac{1}{6} + 1^4 \cdot \frac{1}{2} = \frac{5}{6}.$$

תרגיל 3.14 בהגרלת לוטו עולים בגורל 6 מספרים מתוך 37. הפרס הראשון, למי שמנחש את כל ה-6 הינו 5 מיליון ש"ח. מחיר שליחת ניחוש הינו 3 ש"ח. לעומת זאת, בהגרלת צ'אנס נבחרים 4 קלפים מתוך חפיסה המכילה 32 קלפים (4 סדרות של קלפים ללא הקלפים 2-6), כאשר מכל סדרה נבחר קלף אחד בדיוק. עלות ההשתתפות היא 10 ש"ח ובמידה וניחשתם את כל 4 הקלפים תזכו ב-20,000 ש"ח. לשם הפשטות, נזניח כרגע את כל שאר האפשרויות במשחק הלוטו ובמשחק הצ'אנס.

1. חשבו את תוחלת הרווח במשחק הלוטו.

2. חשבו את תוחלת הרווח במשחק הצ'אנס.

3. במה תבחרו לשחק?

פתרון: נסמן את הרווח במשחק הלוטו ב- L ואת הרווח במשחק הצ'אנס ב- C .

1. נחשב את ההתפלגות של L : $\Pr(L = 5,000,000 - 3) = \frac{1}{\binom{37}{6}} = 4.3 \cdot 10^{-7}$ ולכן $\Pr(L = -3) = 1 - 4.3 \cdot 10^{-7}$. לכן $E(L) = 5000000 \cdot 4.3 \cdot 10^{-7} - 3 \cdot 1 = -0.85$.

2. באופן דומה לסעיף הקודם, $\Pr(C = 20,000 - 10) = \frac{1}{8^4} = 0.000244$ ולכן $\Pr(C = -10) = 0.999756$. לכן התוחלת תהא $E(C) = 20000 \cdot 0.000244 - 10 \cdot 1 = -5.11$.

3. על פניו, תוחלת הרווח בלוטו גבוהה יותר ולכן עדיף לשחק בו אולם בשני המקרים תוחלת הרווח היא שלילית - בממוצע נפסיד מהשתתפות בכל אחד מהמשחקים ולכן לא כדאי להשתתף באף אחד מהם.

תרגיל 3.15 בבחינה אמריקאית n שאלות עם 4 תשובות. סטודנט שנמנס רק בחלק מההרצאות יודע בהסתברות p את התשובה הנכונה לשאלה. אם אינו יודע את התשובה הנכונה - הוא מנחש. על כל תשובה נכונה הסטודנט יקבל $\frac{100}{n}$ נקודות ועל כל תשובה שגויה ייקנס ב- $\frac{40}{n}$ נקודות.

1. מהי תוחלת ציון הבחינה של הסטודנט?

2. כמה נקודות צריך המרצה להוריד לכל תשובה שגויה בכדי שסטודנט שלא יודע כלום יקבל בתוחלת 0 בבחינה?

פתרון: נסמן ב- X את ציון הבחינה וב- X_i את ציון השאלה ה- i . מתקיים $X = \sum_{i=1}^n X_i$ ולכן $E(X) = \sum_{i=1}^n E(X_i)$ ונותר לחשב את ההתפלגות של X_i .

1. לפי נוסחת ההסתברות השלמה: $\Pr(X_i = \frac{100}{n}) = p + (1-p)\frac{1}{4} = \frac{1}{4} + \frac{3}{4}p$ לכן $\Pr(X_i = -\frac{40}{n}) = \frac{3}{4} - \frac{3}{4}p$ ולכן $E(X_i) = \frac{100}{4n}(1+3p) - \frac{40}{4n}(3-3p) = \frac{-5+105p}{n}$.
 $E(X) = 105p - 5$ תהא $E(X)$ כך שתוחלת הציון תהא $E(X) = 105p - 5$.

2. לפי הקנס מהסעיף הקודם, כאשר הסטודנט לא יודע כלום ($p = 0$) הוא יקבל בתוחלת -5 נקודות, לכן הקנס גדול מדי. בכדי שהציון יהיה 0 נקנס את הסטודנט רק ב- $\frac{33.3}{n}$ נקודות על כל תשובה שגויה. ניתן לודא זאת על ידי שחזור הפתרון מהסעיף הקודם.

3.3 שונות

מדוע היה חשוב לנו כל כך לדון בערך של $E[X^2]$? ובכן, התשובה לכך נעוצה בממד אחר, חשוב מאוד מבחינה כלכלית, והוא השונות. השונות (Variance) של משתנה מקרי X , מסומנת ב- $Var(X)$ או ב- $V(X)$, היא מדד פיזור. זאת אומרת, היא מספקת לנו אינפורמציה בנוגע למידה שבה המשתנה המקרי מקבל ערכים שונים זה מזה. בניגוד למדדי אמצע כמו תוחלת, חציון, או שכיח אשר מנסים לאמוד את ערכו של המשתנה, מדדי פיזור בוחנים את המידה בה המשתנה המקרי מפוזר או מרוכז סביב ערכים ספציפיים. במקרה של שונות, הפיזור נבחן ביחס לתוחלת. במילים פשוטות, השונות בוחנת את המידה בה המשתנה המקרי רחוק מן התוחלת שלו.

לשם המחשה נתחיל בדוגמא פשוטה. ניקח את המשתנה המקרי שמקבל את הערכים ± 1 בהסתברות שווה. התוחלת שלו היא כמובן 0. כעת נחשב את השונות שלו שמוגדרת על ידי $V(X) = E[(X - E(X))^2] = E[X^2] = 1$. מחישוב ישיר נקבל $V(X) = E[(X - 0)^2] = E[X^2] = 1$. והשונות של המשתנה היא 1. נבחן את ההגדרה לעומק. תחילה לוקחים את המשתנה ומוודדים את המרחק הריבועי שלו מן התוחלת שלו (לשם הנוחות נעבור לסמן את התוחלת ב- $\mu_X = E[X]$), קרי $(X - \mu_X)^2$. זה המרחק הריבועי בין המשתנה לתוחלת שלו. יחד עם זאת, עדיין מדובר במשתנה מקרי שהוא למעשה פונקציה של המשתנה המקרי המקורי X . לכן ניתן לחשב לו את התוחלת ולקבל $V(X) = E[(X - \mu_X)^2]$. במילים, זהו המרחק הריבועי המשוקלל של המשתנה X מהתוחלת שלו μ_X . בנקודה זו חשוב להדגיש כי μ_X הוא מספר קבוע(!) אשר מייצג את התוחלת של X , בעוד X הוא משתנה מקרי (פונקציה על מרחב המדגם). השונות מודדת את המרחק הריבועי ולכן היחידות הן יחידות ריבועיות. בכדי לחזור ליחידות המקוריות צריך לבצע פעולה נוספת של הוצאת שורש, ובזאת מוגדרת סטיית התקן.

הגדרה 3.16 (סטיית התקן של X) סטיית התקן של משתנה מקרי X מוגדרת להיות שורש השונות ומסומנת ב- $\sigma_X = \sqrt{V(X)}$.

סטיית התקן מוגדרת להיות שורש השונות והיא מהווה מדד די שימושי למרחק של המשתנה מן התוחלת שלו. בנוסף, חישוב השונות באמצעות ההגדרה המקורית יכול להיות די מסורבל, לכן נשתמש בתכונות הליניאריות של התוחלת בכדי לפשט את החישוב באופן הבא:

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

כעת קיבלנו ביטוי יחסית נוח לחישוב אשר מסביר את ההתעניינות המוקדמת שלנו בתוחלת $E[X^2]$.

קל לראות כי משתנים מקריים שונים יכולים להיות בעלי אותה התוחלת אבל עם שונות שונה ולהיפך. לדוגמא, אם ניקח את המשתנה המקרי הקבוע 0 (זאת אומרת, המספר 0 שהוא גם משתנה מקרי שמקבל את הערך 0 בהסתברות 1), אז התוחלת שלו תהיה 0 וגם השונות שלו תהיה 0 (חשבו וראו, זה מייד). בניגוד לחישוב הקודם שביצענו, במהלכו ראינו שהתוחלת של המשתנה X (המקבל את הערכים ± 1 בהסתברות שווה) היא 0 בעוד השונות היא 1.

תרגיל 3.17 בר היא בנקאית השקעות מתוחכמת המציעה ללקוחותיה את ההגרלה הבאה: בסיכוי $\frac{2}{5}$ תקבלו רווח של 10, בסיכוי $\frac{1}{5}$ לא תרוויחו ולא תפסידו דבר, ובסיכוי $\frac{3}{5}$ תפסידו 13.5 .¹³ חשבו את התוחלת וסטיית התקן של X המייצג את הרווח מהשקעה אצל בר.

פתרון: החישוב הוא מייד. מצד אחד, תוחלת הרווח היא חיובית שכן

$$\begin{aligned}\mu_X &= 10 \cdot \frac{2}{5} + 0 \cdot \frac{1}{5} - 5 \cdot \frac{2}{5} = 2, \\ \mu_{X^2} &= 100 \cdot \frac{2}{5} + 0^2 \cdot \frac{1}{5} + 25 \cdot \frac{2}{5} = 50, \\ V(X) &= \mu_{X^2} - (\mu_X)^2 = 50 - 4 = 46 \\ &\Rightarrow \mu_X = 2, \sigma_X = \sqrt{46} \approx 6.78.\end{aligned}$$

זאת אומרת, התוחלת אומנם 2, אבל יש פיזור ותנודתיות לא מבוטלת סביב ההשקעה הנ"ל - ניתן להרוויח היטב, אבל גם ניתן להפסיד.

3.3.1 תכונות השונות

למרות שהתוחלת היא ליניארית, הדבר אינו בהכרח נכון עבור השונות. יחד עם זאת, ישנן מספר תכונות מאוד חשובות שיש לשונות כפי שנראה כעת. נתחיל עם החלק הפשוט והוא חישוב **שונות של מספר קבוע**. ניקח את המספר הקבוע b ונחשב את התוחלת והשונות שלו על פי ההגדרה שראינו. התוחלת היא b , בעוד השונות היא $V(b) = E[b - \mu_b] = E[b - b] = 0$. זאת אומרת, השונות של מספר קבוע היא אפס. תמיד. מצד שני, השונות של כל משתנה אחר היא בהכרח גדולה מ-0 כי שונות מודדת מרחק ריבועי והוא לא יוכל להיות שלילי. לכן, השונות היא **אי-שלילית**, ושווה ל-0 אם ורק אם מדובר בקבוע.

כעת ננתח מצב יותר כללי, עבור משתנה X כלשהו וקבועים a, b . נחשב את השונות של המשתנה החדש $aX + b$, ולשם החישוב נעזר בתכונות הליניאריות של התוחלת.

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

כאשר השתמשנו שוב ושוב בתכונת הליניאריות, ובכך הוכחנו את הטענה הבאה.

טענה 3.18 (שונות של טרנספורמציה ליניארית של משתנה מקרי) עבור X מ"מ וקבועים $a, b \in \mathbb{R}$, נקבל $V(aX + b) = a^2 V(X)$.

¹³קל לראות שכן, בר, ואלון מנהלים קרטל השקעות ערמומי, אבל על כך נרחיב בהזדמנות אחרת...

זאת אומרת, הזזה בקבוע (הוספה של קבוע חיובי או שלילי) למשתנה המקרי איננה משפיעה על השונות שלו. זה הגיוני מאחר ושונות מודדת את הפיזור של ערכי המשתנה ביחס לתוחלת, בעוד הוספה או חיסור של מספר איננה משפיעה על הפיזור שכן כל הערכים זזים באופן דומה, והמרחק היחסי ביניהם לא משתנה. לעומת זאת, כפל בקבוע כן משפיע על הפיזור. גם זה הגיוני שכן אם נכפיל את כל ערכי המשתנה ב-2, אז המרחק היחסי ביניהם יגדל פי 2, והשונות (אשר מייצגת את המרחק הריבועי) תגדל פי 4, כפי שקיבלנו בנוסחה.

התכונה האחרונה שדורשת התייחסות (בשלב הנוכחי בצורה די מוגבלת) נוגעת לשונות של סכום משתנים מקריים. עבור תוחלת קיימת לינאריות ביחס למשתנים, קרי $E[X + Y] = E[X] + E[Y]$, אבל התכונה הזאת איננה נכונה באופן כללי עבור שונות. ישנם מקרים בהם השונות של סכום המשתנים תהיה שווה לסכום השונות, אבל הטענה הזאת איננה נכונה באופן כללי. נרחיב שוב על הנקודה הזאת כאשר נדון בהרחבה במשתנים מקריים דו-מימדיים והיחסים ביניהם.

תרגיל 3.19 חברת חיפוש גז קיבלה רישיון קידוח לשלושה אתרים: 1, 2, ו-3. ידוע שרק באתר אחד בדיוק יש גז, אבל זהות האתר לא ידועה. בכל שנה החברה מגרילה באקראי אתר ובודקת אם יש בו גז או לא. במידה ולא נמצא גז, החברה מגרילה אתר נוסף מהנותרים. במידה ונמצא גז, החברה מפסיקה את הקידוחים. זאת אומרת, תקופת הקידוח יכולה להגיע ל-3 שנים לכל היותר. נסמן ב- X_i את המשתנה המקרי שמקבל את הערך 1 אם נמצא גז לראשונה בשנה ה- i , ו-0 אחרת. נסמן $X = X_1 + X_2 + X_3$ ו- $Y = 2X_1 - 1$. מצאו את תוחלת ושונות X, X_i, Y .

פתרון: נשים לב שההתפלגויות של המשתנים הנתונים הן קלות לחישוב. תחילה, כל המשתנים X_i בעלי אותה התפלגות משיקולי סימטריה. קרי, אין הבדל בין השנה הראשונה לשנה השנייה או לשנה השלישית כי הסיכויים למצוא גז (בהנחה ורק שדה אחד מכיל גז) שווה בכל אחת מן השנים. סדר בחירת השדות הוא אקראי ולכן הסיכוי לקבל 1 בכל אחד מן המשתנים הוא $\frac{1}{3}$, $\frac{1}{3}$. $\Pr(X_i = 1) = \frac{1}{3}$. חישובי תוחלת ושונות יניבו

$$\begin{aligned}\mu_{X_i} &= 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{3}, \\ \mu_{X_i^2} &= 1^2 \cdot \frac{1}{3} + 0^2 \cdot \frac{2}{3} = \frac{1}{3}, \\ V(X_i) &= \frac{1}{3} - \left(\frac{1}{3}\right)^2 = \frac{2}{9}.\end{aligned}$$

כעת נעבור למשתנה Y . הדרך הטובה לחשוב על Y היא בתור הרווח של החברה לאחר השנה הראשונה. נניח כי גילוי שדה גז שווה לחברה 2 מיליארד דולר, בעוד הקידוחים עצמם עולים מיליארד דולר. לכן לאחר שנה, הרווח יכול להיות 1 אם נמצא גז, או -1, אחרת. לפי תכונת הלינאריות שראינו, $E[Y] = 2E[X_1] - 1 = -\frac{1}{3}$, ו- $V(Y) = 4V(X_1) = \frac{8}{9}$. נקבל את אותן התוצאות אם נחשב את התוחלת והשונות ישירות מהתפלגות Y אשר שווה ל-1 בסיכוי $\frac{1}{3}$ ו-1 בסיכוי $\frac{2}{3}$ (בדקו זאת!).

נעבור לניתוח המשתנה X . נשים לב שסכומים המשתנים הוא בהכרח 1 כי יש שדה גז אחד והוא ימצא באחת מן השנים. לכן המשתנה X הוא למעשה המספר הקבוע 1, והתוחלת שלו היא 1 והשונות היא 0. זאת דוגמה למצב בו $V\left(\sum_{i=1}^3 X_i\right) = 0 \neq \sum_{i=1}^3 V(X_i)$.

3.4 משפחות התפלגות

בחלק הקודם סקרנו בהרחבה את המשתנים המקריים באשר הם, יחד עם התוחלת והשונות. כעת אנחנו נתמקד בקבוצת המשתנים המקריים וננתח משתנים ספציפיים המופיעים לרוב בניתוחים סטטיסטיים. ישנן מספר סוגים של משתנים כאלה אשר מחולקים למשפחות ובכל משפחה יש פרמטר אחד או יותר המאפיינים אותה.

המשתנה הראשון והבסיסי ביותר שנפגוש הוא משתנה ברנולי המבוסס על ניסוי ברנולי. ניסוי ברנולי הוא ניסוי עם תוצאה בינארית - כן או לא. 0 או 1. הצלחה או כישלון. בניסויים רפואיים, בסקרים כלכליים, ובכלל בעולם המדע אלו משתנים מאוד נפוצים כי הם פשוט מציניים קיום של תופעה או היעדר. ניתן לחשוב על ניסוי ברנולי בתור הטלת מטבע, לא בהכרח הוגן. המטבע נוחת על עץ או פלי, כאשר אחד מהם מגדיר הצלחה = 1, והאחר מגדיר כישלון = 0. הסיכוי להצלחה נתון על ידי הפרמטר $0 \leq p \leq 1$, וברוב המקרים לא מתייחסים לערכי הקצה שכן מדובר בניסוי טריויאלי (תמיד תהיה הצלחה או, לחילופין, תמיד נקבל כישלון).

3.4.1 ברנולי

משתנה ברנולי הוא משתנה המקבל את הערכים 0 ו-1 בהתבסס על ניסוי ברנולי עם סיכוי p להצלחה. המשתנה מסומן ב- $X \sim \text{Ber}(p)$ (במילים, X מתפלג לפי התפלגות ברנולי עם פרמטר p). פורמאלית,

$$\Pr(X = k) = \begin{cases} p, & k = 1, \\ 1 - p, & k = 0. \end{cases}$$

התוחלת והשונות של משתנה ברנולי קלות לחישוב והן $E[X] = 1 \cdot p + 0 \cdot (1 - p) = p$ ו- $E[X^2] = E[X] = p$ כאשר השיויון נובע מכך שהעלאה של 0 או של 1 בריבוע איננה משנה את הערכים ולכן, באופן כללי, $X^2 = X = X^n$ לכל חזקה טבעית n . לסיכום, $V(X) = p - p^2 = p(1 - p)$.

הערה 3.20 שיויון התפלגויות איננו שיויון בין משתנים! משפחות התפלגות מציינות אך ורק את ההתפלגות של המשתנים, ולא את מרחב ההסתברות עליהם המשתנים מוגדרים. יכולים להיות משתנים שונים זה מזה, בעלי אותה ההתפלגות. זאת אומרת, ייתכן מצב בו X ו- Y הם שני משתנים שונים, אבל יש להם את התפלגות זהה. לדוגמא, נחשוב על צמד מטבעות הוגנים. שני המטבעות בעלי התפלגות ברנולי עם פרמטר חצי ולכן יש להם את אותה ההתפלגות, אבל אין זה אומר שהם שווים אחד לשני. כאשר מטבע אחד נוחת על 0, זה לא אומר שהשני גם יינחת על 0. אותו הדבר לגבי צמד קוביות הוגנות. יש להן את אותה ההתפלגות, אבל הטלה אחת יכולה להראות 2 והאחרת יכולה להניב 5. על כן המשתנים אינם שווים זה לזה, אלא רק בעלי אותה התפלגות.

3.4.2 אחידה

התפלגות אחידה היא הדוגמא הבסיסית הבאה למשפחות התפלגות אשר מרחיבה במידה מסויימת את התפלגות ברנולי. הדוגמא המוכרת היא הטלת קוביה הוגנת. בהטלת קוביה הוגנת מקבלים ערכים מ-1 עד 6, וכל ערך מתקבל בהסתברות שווה. זה הרעיון שעומד בבסיס ההתפלגות האחידה. ניקח שני מספרים שלמים $a < b$ ונגיד שמשתנה מקרי X מתפלג אחיד על $[a, b]$, בסימונים $X \sim [a, b]$, אם $\Pr(X = k) = \frac{1}{b-a+1}$ לכל $a \leq k \leq b$. זאת אומרת, התומך של משתנה אחיד שכזה הוא $\text{Supp}(X) = \{a, a+1, a+2, \dots, b\}$ וכל הערכים מתקבלים בהסתברות שווה. כמובן שנוסחה זאת עיקבית עם הדוגמא הבסיסית של קוביה הוגנת בעלת התפלגות אחידה מ-1 עד 6, עבורה כל ערך מתקבל בסיכוי $\frac{1}{6-1+1} = \frac{1}{6}$. חישוב התוחלת והשונות של משתנה בעלת התפלגות אחידה מבוסס על נוסחאות של סכום סדרה חשבונית וסכום ריבועים של מספרים שלמים עוקבים.

$$\begin{aligned} E[X] &= \sum_{k=a}^b \left(k \cdot \frac{1}{b-a+1} \right) \\ &= \frac{1}{b-a+1} \sum_{k=a}^b k \\ &= \frac{1}{b-a+1} \cdot \frac{(b+a)(b-a+1)}{2} \\ &= \frac{b+a}{2}, \end{aligned}$$

כאשר המעבר השלישי מבוסס על סכום סדרה חשבונית. חישוב יותר מתקדם עבור השונות יניב $V(X) = \frac{(b-a+1)^2-1}{12}$.

תרגיל 3.21 ריבית הלייבור בפונטי-פנדי נקבעת באקראי על ידי דגימה של אחד מעשרת הבנקים המסחריים. הבנקים ממסופרים מ-1 עד 10 כאשר בנק i מעניק ריבית בשיעור i ללקוחותיו. לאחר זמן מה נסגרו בנקים 1 ו-2 (בשל הענקת ריבית לא כדאית...). חשבו את הסיכוי שריבית הלייבור נמוכה מ-6 וכן את תוחלת ושונות הריבית. בצעו את החישובים עבור התקופה שקדמה לסגירת הבנקים ולאחר הסגירה.

פתרון: נגדיר צמד משתנים מקריים. X עבור הריבית טרם הסגירה ו- Y עבור הריבית לאחר שנסגרו הבנקים המדוברים. לכן $X \sim U[1, 10]$ ו- $Y \sim U[3, 10]$, $E[X] = \frac{1+10}{2} = 5.5$, $E[Y] = \frac{3+10}{2} = 6.5$, $V(X) = \frac{(10-1+1)^2-1}{12} = \frac{99}{12}$, $V(Y) = \frac{(10-3+1)^2-1}{12} = \frac{63}{12}$, $\Pr(Y \leq 5) = \frac{3}{8}$, $\Pr(X \leq 5) = \frac{5}{10} = 0.5$.

3.4.3 בינומית

ההתפלגות הבינומית מכלילה את התפלגות ברנולי למספר ניסויי ברנולי בלתי תלויים. נדמיין ניסוי ברנולי, כגון הטלת מטבע, אשר מבוצע שוב ושוב ובסה"כ n פעמים. השאלה הטבעית שעולה היא מה יהיו מספר ההצלחות בכלל הניסויים. לדוגמא, אם בוחנים השפעות של תרופה על n נבדקים, נרצה לדעת בכמה מקרים התרופה אכן הייתה אפקטיבית. המשתנה אשר מציין את מספר ההצלחות ב- n ניסויי ברנולי ב"ת עם פרמטר p הוא משתנה בינומי, בעל התפלגות בינומית. נסמן זאת ב- $X \sim \text{Bin}(n, p)$. התומך של משתנה מקרי בינומי הוא $\text{Supp}(X) = \{0, 1, 2, 3, \dots, n\}$: יכולות להתקבל 0 הצלחות, הצלחה אחת וכן הלאה, עד למקסימום של n הצלחות. ההתפלגות עצמה נתונה על ידי הנוסחה $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ עבור כל k בתומך. נסביר את הנוסחה. המאורע $X = k$ מציין k הצלחות. הניסויים עצמם בלתי תלויים ולכן הסיכוי שיהיה בדיוק k הצלחות ספציפיות ב- n ניסויים הוא $p^k (1-p)^{n-k}$ (הסתברות לקבלת k הצלחות ו- $n-k$ כשלונות מסויימים). בנוסף, צריך לקחת בחשבון את כלל הקומבינציות לקבלת אותן ההצלחות - ניתן לבחור באילו ניסויים, מתוך סה"כ הניסויים, יהיו הצלחות ומספר האפשרויות הללו הוא בחירה של k מתוך n , $\binom{n}{k}$ כידוע. חישובי התוחלת והשונות של משתנה בינומי אינם טריוויאליים. יחד עם זאת, בחישוב עצמו ישנם אלמנטים חשובים ללמידה אשר ישמשו אותנו רבות בהמשך הקורס ובכלל בפתרון בעיות הסתברותיות וסטטיסטיות. לכן נחשב את התוחלת והשונות של משתנה בינומי כללי כתרגיל.

תרגיל 3.22 חשבו את $E[X]$ ו- $V(X)$ עבור משתנה בינומי $X \sim \text{Bin}(n, p)$.

פתרון: החישוב מתחיל ברעיון יחסית פשוט ומאוד יעיל של פירוק X למשתנים מקריים פשוטים. נזכור כי X מציין את מספר ההצלחות ב- n ניסויי ברנולי בלתי תלויים. לכן, עבור כל ניסוי ברנולי i , נגדיר משתנה ברנולי X_i אשר מקבל את הערך 1 (הצלחה) בסיכוי p אם ורק אם הייתה הצלחה בניסוי המתאים. הרעיון המיוחד בהגדרת n המשתנים הללו מופיע בשיויון הבא: $X = \sum_{i=1}^n X_i$. במילים, אגף שמאל הוא משתנה בינומי המודד את מספר ההצלחות ב- n ניסויי ברנולי ב"ת, בעוד אגף ימין הוא סכום של משתני ברנולי כאשר כל אחד מהם מציין הצלחה בניסוי אחר מבין הניסויים שהוזכרו לעיל. הסכום באגף ימין שווה זהותית לאגף שמאל משום שהוא סופר את מספר ההצלחות באותם הניסויים בדיוק! שימו לב, זהו שיויון בין שני משתנים מקריים! קרי, כאשר אגף שמאל שווה ל-1, גם אגף ימין שווה ל-1 וכן הלאה. כעת נשתמש בתכונת הלינאריות של התוחלת. $E[X_i] = p$ כפי שחישבנו קודם לכן, ומכאן נובע ש-

$$E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n p = np.$$

התוחלת של הסכום שווה לסכום התוחלת ולכן קיבלנו את התוצאה $E[X] = np$. עבור השונות החישוב יהיה מעט יותר מורכב. נחשב את

$$E[X^2]$$

$$\begin{aligned} E[X^2] &= E\left[\left(\sum_{i=1}^n X_i\right)^2\right] \\ &= E\left[\left(\sum_{i=1}^n X_i\right)\left(\sum_{j=1}^n X_j\right)\right] \\ &= E\left[\sum_{i=1}^n X_i^2 + 2\sum_{i \neq j} X_i X_j\right] \\ &= E\left[\sum_{i=1}^n X_i^2\right] + 2E\left[\sum_{i \neq j} X_i X_j\right] \end{aligned}$$

נזכור כי $X_i^2 = X_i$ עבור משתני ברנולי, ובנוסף ההתפלגות של $X_i X_j$ היא די פשוטה. התומך של המכפלה הוא 0 או 1, כי צמד המשתנים מקבלים רק ערכים אלה ומכפלתם תניב את אותם הערכים. בנוסף, המכפלה תהיה שווה ל-1 אם ורק אם צמד המשתנים שווים ל-1. זאת אומרת, $\Pr(X_i X_j = 1) = \Pr(X_i = X_j = 1)$ ואם הניסויים הם בלתי תלויים, אז המאורע המדובר מתקבל בהסתברות p^2 (הסיכוי לשתי הצלחות ספציפיות ב-2 ניסויי ברנולי בלתי תלויים). לכן $X_i X_j \sim \text{Ber}(p^2)$ משתנה המקבל את הערך 1 בסיכוי p^2 ו-0 אחרת, והתוחלת שלו היא p^2 כפי שכבר הוכחנו. לסיכום,

$$\begin{aligned} E[X^2] &= E\left[\sum_{i=1}^n X_i^2\right] + 2E\left[\sum_{i \neq j} X_i X_j\right] \\ &= \sum_{i=1}^n E[X_i] + 2\sum_{i \neq j} E[X_i X_j] \\ &= \sum_{i=1}^n p + 2\sum_{i \neq j} p^2 \\ &= np + 2\binom{n}{2}p^2 \\ &= np + n(n-1)p^2. \end{aligned}$$

בשורה השלישית רשומים צמד סכומים כאשר האיברים בכל סכום, בפני עצמו, הם זהים. כל האיברים בסכום הראשון שווים ל- p (ויש n איברים כאלה בסכום), ולכן הוא מסתכם ל- np . בסכום השני כל האיברים שווים ל- p^2 וישנם $\binom{n}{2}$ איברים כאלה (צריך לבדוק צמד אינדקסים שונים i ו- j מתוך n אינדקסים סה"כ), לכן הסכום מניב את הערך $\binom{n}{2}p^2$. מנקודה זו חישוב השונות הוא מיידי,

$$\begin{aligned} V(X) &= E[X^2] - (\mu_x)^2 \\ &= np + n(n-1)p^2 - (np)^2 \\ &= np - np^2 \\ &= np(1-p) = \sum_{i=1}^n V(X_i). \end{aligned}$$

קיבלנו שהשונות של הסכום שווה לסכום השונויות וזה לא במקרה. המשתנים בהם דנו הם בלתי תלויים אחד בשני, ז"א הם מבוססים על ניסויים בלתי תלויים. בהמשך נראה טענה כללית יותר הקובעת כי שונות של סכום של משתנים בלתי תלויים שווה לסכום השונויות (כאמור, לינאריות).

תרגיל 3.23 מחשב משתמש במילים באורך 8 ביטים. כאשר משדרים ביטים בקו התקשורת יש סיכוי של 0.1 לשגיאה בביט בודד שמשודר. בכדי לשפר את איכות הקו, משודר כל ביט 3 פעמים ברצף והפענוח של כל שלשת ביטים מתבצע על ידי החלטת הרוב. למשל: אם רוצים לשדר 0, משדרים בקו "000" והמקלט מפענח את התשדורת כ-0 אם השלשה שהגיעה אליו מורכבת מלפחות שני אפסים.

1. חשבו את ההסתברות שמילת מחשב תפוענח נכון.

2. מצאו את התפלגות מספר המילים שפוענחו באופן שגוי בקובץ המכיל 1000 מילים.

פתרון: נשים לב שכמות הטעויות בשליחה של ביט בודד מתפלגת $X \sim \text{Bin}(3, 0.1)$ ולכן ההסתברות שתהיה שגיאה בפענוח ביט בודד היא $\Pr(X \geq 2) = \binom{3}{2} 0.1^2 0.9 + 0.1^3 = 0.028$.

1. כמות השגיאות במילה באורך 8 ביטים מתפלגת $Y \sim \text{Bin}(8, 0.028)$, ולכן המילה תפוענח נכון בהסתברות $\Pr(Y = 0) = (1 - 0.028)^8 = 0.796$.

2. לאור הסעיף הקודם, ההסתברות לשגיאה בפענוח מילה בודדת היא $1 - 0.796 = 0.203$, לכן מספר המילים שפוענחו באופן שגוי בקובץ המכיל 1000 מילים מתפלג $\text{Bin}(1000, 0.203)$.

תרגיל 3.24 בלאס וגאס שני בתי קזינו. בקזינו "הצדיקים" ההסתברות להרוויח ולהפסיד היא שווה ואילו בקזינו "הרשעים" ההסתברות להפסיד היא 0.6 בכל סיבוב. מהמר נכנס באקראי לאחד מבתי הקזינו ומהמר על דולר אחד בכל סיבוב.

1. כיצד מתפלג מספר הזכיות של המהמר בכל אחד מבתי הקזינו לאחר 10 סיבובי משחק?

2. מהי תוחלת הרווח של המהמר בכל אחד מבתי הקזינו אחרי 10 סיבובים?

3. לאחר 10 סיבובים המהמר הרוויח \$8. מה ההסתברות שהוא בקזינו "הצדיקים"?

פתרון:

1. מספר הזכיות בכל אחד מבתי הקזינו מתפלג בינומית, כאשר השוני הוא רק בהסתברות לזכייה. אצל "הצדיקים" מספר הזכיות מתפלג $\text{Bin}(10, \frac{1}{2})$ ואילו אצל "הרשעים" $\text{Bin}(10, 0.4)$.

2. נגדיר X_i - הרווח בסיבוב i -י. באופן בלתי תלוי בסיבובים האחרים, $E(X_i) = p - (1 - p) = 2p - 1$, כאשר $p = \frac{1}{2}$ אצל "הצדיקים" (ולכן $E(X_i) = 0$) ו- $p = 0.4$ אצל "הרשעים" (ולכן $E(X_i) = -0.2$). הזכייה לאחר 10 סיבובים היא סכום הזכיות בכל אחד מ-10 הסיבובים ומחיבוריות התוחלת נקבל שאצל "הצדיקים" תוחלת הרווח אחרי 10 סיבובים (ולמעשה, אחרי כל מספר של סיבובים) היא 0 ואילו אצל הרשעים אחרי 10 סיבובים המהמר יהיה בתוחלת עם \$-2.

3. נסמן ב- G את המאורע שהמהמר נכנס לקזינו "הצדיקים" ו- X את הרווח שלו אחרי 10 סיבובים. נשתמש בנוסחת בייס ונקבל $\Pr(G|X=8) = \frac{\Pr(X=8|G)\Pr(G)}{\Pr(X=8)}$. עבור רווח של \$8 ב-10 סיבובים אנחנו זקוקים ל-9 זכיות בדיוק ולכן במונה נקבל $\Pr(X=8) = \binom{10}{9} (\frac{1}{2})^9 (\frac{1}{2}) = 0.00978$. את המכנה יש לחשב באמצעות הסתברות שלמה:

$$\Pr(X=8) = \Pr(X=8|G)\Pr(G) + \Pr(X=8|\bar{G})\Pr(\bar{G}) = \frac{1}{2} 0.00978 + \frac{1}{2} \binom{10}{9} 0.4^9 0.6 = 0.00567.$$

$$\text{לסיכום } \Pr(G|X=8) = \frac{0.00978 \cdot 0.5}{0.00567} = 0.861$$

תרגיל 3.25 במבחן בקורס ישנן 20 שאלות אמריקאיות כך שלכל שאלה 4 תשובות אפשריות שרק 1 נכונה. בכל שאלה ניתן לסמן רק תשובה אחת. תשובה נכונה מזכה ב-6 נקודות ותשובה שגויה גורעת 2 נקודות. ישנן 8 שאלות קלות שעליהן יודע אלון לענות נכונה ו-12 שאלות קשות בהן הוא מנחש תשובה באקראי. יהי Z הציון של אלון. מצאו את התוחלת והשונות של Z .

פתרון: נסמן ב- X את מספר התשובות הנכונות של אלון מבין השאלות הקשות. השאלות הן בלתי-תלויות ולכן $X \sim \text{Bin}(12, \frac{1}{4})$. נוכל לרשום מפורשות את היחס בין Z ו- X על ידי $Z = 6X + (12 - X)(-2) + 8 \cdot 6 = 8X + 24$. נזכור כי התוחלת והשונות של מ"מ בינומי $X \sim \text{Bin}(n, p)$ הן $E(X) = np$, $V(X) = np(1-p)$. נחשב בעזרת נוסחאות אלו את התוחלת והשונות של Z ,

$$E(Z) = 8E(X) + 24 = 8 \cdot \left(\frac{12}{4}\right) + 24 = 48,$$

$$V(Z) = V(8X + 24) = 64V(X) = 64 \cdot 12 \cdot \frac{1}{4} \cdot \frac{3}{4} = 144.$$

תרגיל 3.26 מניה מתחיל לנוע מגובה $x = 0$, על פני זמן בדיד. בכל שלב המניה עולה או יורדת, ביחידה אחת, בהסתברות שווה במשך 100 יחידות זמן.

1. כיצד מתפלג מיקום המניה בסוף התהליך? חשבו את תוחלת ושונות המיקום.
2. מה ההסתברות שבסופו של דבר המניה תמצא בנקודה בנקודה $x = 10$ ובנקודה $x = 11$?
3. לאחר 100 יחידות זמן המניה הגיע לנקודה $x = 10$. מה ההסתברות שהתנועה האחרונה שלה הייתה לכיוון מעלה?
4. חזרו על סעיף 4 בהנחה שהמניה נעה כלפי מעלה בהסתברות p כלשהי.

פתרון: נסמן ב- X את המיקום האנכי של המניה וב- Y את כמות התנועות מעלה. קל לראות ש $Y \sim \text{Bin}(100, \frac{1}{2})$ וכן $X = Y - 100$. המיקום שווה למספר התנועות כלפי מעלה בחיסור מספר התנועות כלפי מטה.

1. לאור התוצאות הקודמות,

$$\Pr(X = k) = \Pr(2Y - 100 = k) = \Pr(Y = \frac{k+100}{2}) = \binom{100}{\frac{k+100}{2}} \left(\frac{1}{2}\right)^{100},$$

כאשר k זוגי ו-0 אחרת. בנוסף, $E[Y] = 100 \cdot \frac{1}{2} = 50$ ו- $V(Y) = 100 \cdot \frac{1}{2} \cdot \frac{1}{2} = 25$, לכן

$$E[X] = E[2Y - 100] = 2 \cdot E[Y] - 100 = 0,$$

$$V(X) = 4V(Y) = 100.$$

2. ניתן להציב בנוסחה שמצאנו ולקבל $\Pr(X = 11) = 0$, $\Pr(X = 10) = \Pr(Y = 55) = \binom{100}{55} 2^{-100}$.

3. בהינתן שהמנייה הגיעה לגובה 10, היא ביצעה בדיוק 55 תנועות לכיוון מעלה ו-45 לכיוון מטה. ההסתברות שהתנועה האחרונה תהיה כלפי מעלה היא $\frac{55}{100}$. ההסבר הוא קומבינטורי. ישנם 55 צעדים כלפי מעלה ו-45 כלפי מטה. צריכים לסדר אותם בשורה. הסיכוי שהצעד האחרון יהיה כלפי מעלה הוא בחירה של צעד מעלה, 55 אפשריים, מתוך 100 סה"כ.

4. התשובה לסעיף הקודם לא תלויה בהסתברות p ונקבל את התוצאה 0.55 בשנית.

תרגיל 3.27 לפניכם חמישה מטבעות הוגנים. מטילים את כולם ומוציאים מהמשחק את המטבעות שנפלו על עץ. מטילים שנית את המטבעות שלא הוצאו מהמשחק ומוציאים שוב את המטבעות שנפלו על עץ. ממשיכים שוב ושוב עד שכל המטבעות הוצאו מהמשחק.

1. מה ההסתברות שלאחר הסיבוב השני נותרו 2 מטבעות?

2. מהי ההסתברות שלאחר הסיבוב החמישי נותרו 3 מטבעות?

3. מהי ההסתברות שלאחר 4 סיבובים לא נותרו מטבעות?

פתרון: במקום להסתכל על כל הטלה כעל ניסוי, נסתכל על כל מטבע כעל ניסוי. ההסתברות שמטבע מסוים ישרוד n סיבובים היא $(0.5)^n$. לכן כמות המטבעות ששרדו n סיבובים מתפלגת $X_n \sim \text{Bin}(5, 0.5^n)$. לפיכך:

$$1. \Pr(X_2 = 2) = \binom{5}{2} 0.25^2 0.75^3 = 0.2636$$

$$2. \Pr(X_5 = 3) = \binom{5}{3} 0.031^3 0.968^2 = 0.000286$$

$$3. \Pr(X_4 = 0) = \binom{5}{0} 0.0625^0 (1 - 0.625)^5 = 0.724$$

3.4.4 גיאומטרית

ההתפלגות הגיאומטרית מבוססת גם כן על ניסויי ברנולי, בדומה להתפלגות הבינומית, אלא שכעת המדידה מתייחסת לגורם שונה. נפתח בשאלה לדוגמא: נניח כי אנו מטילים צמד קוביות שוב ושוב עד שמתקבלת לראשונה התוצאה $(6, 6)$. מה מספר ההטלות שנדרש לבצע? ובכן, התשובה לשאלה הזו היא משתנה מקרי כלשהו. ייתכן וכבר בסיבוב הראשון נקבל את התוצאה הרצויה, ייתכן ותתקבל בסיבוב השני, וגם קיימת האפשרות כי נאלץ לחכות מאה סיבובים (ואף יותר מכך) עד שנקבל את התוצאה המבוקשת. מספר הסיבובים הנדרש אינו חסום ויוכל להיות גדול כרצוננו. המשתנה המקרי המתאר את מספר הסיבובים נקרא משתנה מקרי גיאומטרי, והתפלגותו נקראת התפלגות גיאומטרית.

פורמאלית, משתנה מקרי גיאומטרי הוא משתנה מקרי המציין את מספר ניסויי הברנולי הבלתי תלויים שמתקיימים עד לקבלת הצלחה ראשונה. התפלגות גיאומטרית נתונה על ידי פרמטר יחיד p והוא הסיכוי לקבלת הצלחה בניסוי ברנולי בודד. נסמן את המשתנה ב- $X \sim G(p)$ והתפלגותו היא $\Pr(X = k) = (1 - p)^{k-1} p$ עבור כל $k \geq 1$ טבעי. ההסבר להתפלגות הוא מידי. בכדי להגיע להצלחה הראשונה בניסוי ה- k , צריכים להתרחש $k - 1$ כשלונות רצופים ומיד לאחר מכן הצלחה. זאת בדיוק הנוסחה ההסתברותית שכתבנו. התומך של ההתפלגות הוא כל המספרים הטבעיים, כאשר הערך המינימאלי הוא 1, כי צריך להתקיים לפחות ניסוי אחד בכדי להגיע להצלחה אחת. חישובי תוחלת ושונות עבור התפלגות גיאומטרית מתבססים על סכומים גיאומטריים, סכום סדרה הנדסית. החישוב עצמו די עדין ומצריך טכניקות שלא למדנו עד כה, לכן נרשום את התוצאה הסופית בלבד,

$$E[X] = \sum_{k=1}^{\infty} k \Pr(X = k) = \sum_{k=1}^{\infty} k (1 - p)^{k-1} p = \frac{1}{p},$$

$$V(X) = \frac{1 - p}{p^2}.$$

יש היגיון די ברור מאחורי הנוסחה לתוחלת של משתנה גיאומטרי. ניקח לדוגמא מטבע הוגן. כמה הטלות בממוצע נצטרך לבצע עד שנקבל לראשונה 'עץ'? בהנחה והמטבע הוא הוגן, אז 'עץ' מתקבל בהסתברות חצי, לכן זה רק הגיוני שבממוצע נבצע 2 הטלות עד שנקבל 'עץ'. כמובן, אין הדבר אומר שבהכרח התוצאה תתקבל אחרי 2 הטלות. התהליך יכול להימשך גם עשרות שלבים ומעבר לכך. התפלגות גיאומטרית בעלת צמד תכונות חשובות: נוסחת זנב וחוסר זיכרון. נדון תחילה בנוסחת הזנב. נסתכל על המאורע $\{X > k\}$ וננסה לחשב את הסתברותו. המאורע מתאר מצב בו מספר הניסויים עד להצלחה הראשונה גדול מ- k . למה הדבר שקול? בכדי שההצלחה הראשונה תתקבל לאחר מעל k ניסויים, אזי צריכים להתקבל לפחות k כשלונות רצופים מן השלב הראשון. קרי, אם יתקבלו k כשלונות רצופים, אזי המשתנה הגיאומטרי יהיה גדול ממש מ- k . נשים לב ששני הדברים שקולים, ז"א ישנם k כשלונות רצופים אם ורק אם המשתנה גודל ממש מ- k . לכן הסתברות המאורע הנ"ל שווה להסתברות לקבל k כשלונות ברצף, והיא $(1 - p)^k$. זאת אומרת, $\Pr(X > k) = (1 - p)^k$. זאת נוסחת הזנב. באותה המידה היינו יכולים לחשב את ההסתברות למאורע $\{X \geq k\}$ אשר שקול למאורע $\{X \geq k - 1\}$ ולקבל $\Pr(X \geq k) = (1 - p)^{k-1}$. נעבור לדון בתכונת חוסר הזיכרון.

אנו מעוניינים לחשב את ההסתברות $\Pr(X > k + n | X > n)$. במילים, ההסתברות שמספר הניסויים עד ההטלה הראשונה יהיה גדול ממש $n + k$ בהנחה וכבר עברנו n ניסויים ללא הצלחה. האינטואיציה מרמזת על כך שכל מה שנדרש הוא לעבור עוד k ניסויים ללא הצלחה (עברנו כבר n כשלונות רצופים ואנו מעוניינים להגיע ל- $n + k$ כשלונות רצופים), ואכן זה מה שנקבל מחישוב ישיר לפי הגדרה של ההסתברות מותנה ונוסחת הזנב.

$$\begin{aligned}\Pr(X > k + n | X > n) &= \frac{\Pr(X > k + n, X > n)}{\Pr(X > n)} = \frac{\Pr(X > k + n)}{\Pr(X > n)} \\ &= \frac{(1-p)^{k+n}}{(1-p)^n} = (1-p)^k = \Pr(X > k),\end{aligned}$$

לכן ההסתברות המותנה איתה התחלנו את החישוב שווה להסתברות לקבלת k כשלונות רצופים, כפי שחזינו. הנוסחה שקיבלנו היא ההגדרה לחוסר זיכרון של משתנה מקרי - $\Pr(X > k + n | X > n) = \Pr(X > k)$. במילים פשוטות, ניתן להתעלם מאשר התרחש בעבר, ולהתמקד רק בשארית שנותרה בכדי להגיע למצב הרצוי, "חוסר זיכרון".

תרגיל 3.28 בכד 9 כדורים לבנים ו-1 שחור. מושכים באקראי בזה אחר זה כדור מהכד. נסמן ב- X את מספר המשיכות עד לקבלת שחור וב- Y את מספר המשיכות עד לקבלת 3 שחורים.

1. מצאו את התפלגות X כאשר המשיכה היא עם החזרה.

2. מצאו את התפלגות X כאשר המשיכה היא ללא החזרה.

3. מצאו את התפלגות Y כאשר המשיכה היא עם החזרה.

פתרון: נענה על כל סעיף בנפרד.

1. X מקבל ערכים ב- $\mathbb{N} \setminus \{0\}$. בכל הוצאה ישנה הסתברות של $\frac{9}{10}$ לכדור לבן (כישלון) והסתברות של $\frac{1}{10}$ לכדור לבן (הצלחה). כלומר, קיבלנו מ"מ גאומטרי, $X \sim G\left(\frac{1}{10}\right)$.

2. X מקבל ערכים ב- $\{1, \dots, 10\}$, כאשר מטעמי סימטריה לכל $1 \leq k \leq 10$ מתקיים $\Pr(X = k) = \frac{1}{10}$, וההסתברות של כל כדור לצאת במשיכה ה- k הוא $\frac{1}{10}$. שימו לב- אנו לא יודעים איזה כדור יצא במשיכות $\{1, \dots, k-1\}$.

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

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

כאשר $\binom{k-1}{2}$ הוא מספר הבחירות ל-2 המקומות בהם יצאו כדורים שחורים ב- $k-1$ ההוצאות הראשונות, $\left(\frac{1}{10}\right)^2$ היא ההסתברות להוצאת 2 שחורים, $\left(\frac{9}{10}\right)^{k-3}$ ההסתברות להוצאת $k-3$ לבנים, ו- $\frac{1}{10}$ היא ההסתברות לכדור שחור בהוצאה האחרונה.

תרגיל 3.29 שחקן המשחק מזל מחליט להפסיק כשיפסיד לראשונה או לאחר 10 סיבובים. הסיבובים בלתי תלויים וההסתברות לנצח בסיבוב בודד היא p . יהי X מספר הסיבובים הכולל. כיצד מתפלג X ? עבור $p = \frac{1}{2}$ חשבו את ההסתברות שיפסיק לשחק לפני הסיבוב העשירי.

פתרון: ההתפלגות המתוארת היא התפלגות גאומטרית קטומה - השחקן משחק עד ההפסד הראשון או עד הסיבוב העשירי, מה שיבוא קודם. לכן עבור כל $1 \leq k \leq 9$ נקבל $\Pr(X = k) = p^{k-1}(1-p)$. השחקן ישחק בדיוק 10 סיבובים אם הוא ניצח ב-9 הסיבובים הראשונים, לכן $\Pr(X = 10) = p^9$. לפיכך $\Pr(X < 10) = 1 - \Pr(X = 10) = 1 - p^9 = 0.998$.

משפחת ההתפלגויות הבדידה האחרונה עליה נדון היא ההתפלגות הפואסונית. התפלגות פואסונית מופיעה רבות בחיי היומיום, באינספור סיטואציות, והמקור שלה הוא דווקא ההתפלגות הבינומית. נחשוב על הסיטואציה הבאה. מוקד שירות של חברת סלולר מגיב לקריאות של לקוחות החברה. בכל שעה נתונה כל לקוח מתקשר בהסתברות קטנה, בלי תלות ביתר הלקוחות. מספר הלקוחות הוא יחסית גדול ולכן נוצר איזון בין המספר הגדול של הלקוחות והסיכוי של כל לקוח להתקשר באותה השעה. במונחים של התפלגות בינומית, הפרמטר n מייצג את מספר הלקוחות והוא יחסית גדול, בעוד הפרמטר p מייצג את הסיכוי של כל לקוח להתקשר והוא יחסית קטן. זאת אומרת ש- np הוא (בקירוב) גודל קבוע כלשהו (מכפלה של ערך גדול עם ערך קטן). הגודל np מייצג את מספר הלקוחות הממוצע שיתקשרו בשעה כלשהי. ומה מספר הלקוחות הממוצע שיתקשרו בשעתיים? מתכונות הלינאריות אנחנו יודעים שזה $2np$. מה לגבי מספר הלקוחות הממוצע אשר יתקשרו בחצי שעה כלשהי? שוב, מתכונות הלינאריות, אנו יודעים כי מדובר בכמות של $\frac{1}{2}np$, וכן הלאה עבור כל יחידת זמן שנבחר. נפתח את הדוגמא הנ"ל למאפייני ההתפלגות הפואסונית, ולמספר תכונות חשובות שלה:

- מספר האירועים הממוצע ליחידת זמן כלשהי (או ליחידת שטח כלשהי) הוא ערך קבוע $\lambda > 0$ (האות היוונית למדא). זהו הפרמטר המרכזי בהתפלגות פואסון. באנלוגיה להתפלגות הבינומית, הפרמטר λ מייצג את המכפלה np .
- מספר האירועים בקטעי זמן (או שטח) זרים הם בלתי תלויים זה בזה. במילים אחרות, מספר השיחות שנקבל בשעה הראשונה הוא בלתי תלוי במספר השיחות שנקבל בשעה השנייה.
- הסיכוי לאירוע בקטע זמן (שטח) מסויים תלוי אך ורק באורך קטע הזמן (השטח). במילים אחרות, הסיכוי שתתקבל שיחה בשעה הראשונה שווה לסיכוי שתתקבל שיחה בשעה השנייה ושניהם שווים לסיכוי שתתקבל שיחה בשעה השלישית וכן הלאה. באופן דומה, הסיכוי לקבל שיחה בחצי השעה הראשונה שווה לסיכוי לקבל שיחה בחצי השעה השנייה וכו'. אורך קטע הזמן הוא המשפיע על ההסתברות, ולא המיקום של הקטע על ציר הזמן.

כאשר התכונות הנ"ל מתקיימות, נדע שמדובר במשתנה מקרי פואסוני, בעל התפלגות פואסון, ונסמנו ב- $X \sim P(\lambda)$. התהליך כולו של האירועים המתרחשים בהתאם לכללים לעיל נקרא **תהליך פואסון** או **זרם אירועים פואסוני**. נעבור על מספר דוגמאות מציאותיות של אירועים המתרחשים לפי תהליך פואסון. נתחיל באירועים המתרחשים על פני יחידות זמן: זרם הלקוחות שנכנסים לבית עסק במשך יום עבודה או במהלך שעת עבודה, מספר תאונות הדרכים בקטע דרך מסויים בחודש, זמן החיים של מכשיר חשמלי כלשהו או של נורת חשמל, ביקורת פתע של רמה ממונה ועוד, כל אלו הם תהליכי פואסון. מצד שני, אפשר גם להציג תהליכים המתקיימים על פני אישור או שטח מסויים, לדוגמא: פיזור הכובים בחלל, פיזור טיפות הגשם של בקילומטר רבוע, מספר שגיאות הכתיב בעמוד עיתון ועוד.

לאחר הבנת התהליך והתכונות שלו, נדון בהתפלגות עצמה. עבור $X \sim P(\lambda)$ התומוך הוא כל מספר שלם $k \geq 0$. מספר האירועים אינו שלילי, ויכול להתקבל כל מספר אירועים, החל מ-0 ומעלה. ההתפלגות עצמה נתונה על ידי הנוסחה $\Pr(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ עבור כל ערך בתומוך. כיצד הגענו לנוסחה הזאת? התשובה נעוצה בסיפור ממנו נפתח הדיון. לוקחים התפלגות בינומית ומניחים כי $np = \lambda$. כעת לוקחים את הגבול $n \rightarrow \infty$ ומחשבים את הערך $\binom{n}{k} p^k (1-p)^{n-k}$. זאת ההסתברות של משתנה בינומי לקבל את הערך k . הגבול המתקבל תחת ההנחה ש- $np = \lambda$ זה בדיוק הביטוי שכתבנו $\Pr(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$. לא נוכל כעת לבצע את החישוב בשל מורכבותו, אבל כל מי שמתעניין יכול בקלות למצוא את ההוכחה הרלוונטית. לסיכום, משתנה מקרי פואסוני מבוסס על תהליך פואסון, וסופר את מספר האירועים ליחידת זמן (או שטח) כאשר מספר האירועים הממוצע ליחידת זמן (או שטח) הוא קבוע ושווה ל- λ .

התוחלת והשונות של התפלגות פואסון הן די קלות לזיכרון - $E[X] = V(X) = \lambda$.

$$\begin{aligned} E[X] &= \sum_{k=0}^{\infty} k \Pr(X = k) \\ &= \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} = \lambda \cdot 1 = \lambda, \end{aligned}$$

כאשר בשורה השלישית השתמשנו בעובדה שהאיבר הראשון, $k=0$, הוא אפס ורשמנו את הסכום רק מהאיבר $k=1$; בשורה הבאה צמצנו את k והוצאנו λ מחוץ לסכום; בשורה האחרונה רשמנו מחדש את הסכימה על ידי החלפת $k-1$ ב- n (רשמנו את שני הסכומים מפורשות ותראו כי מדובר בסכומים זהים!); והשיויון בשורה האחרונה נובע מכך שהסכום על כל הערכים של התפלגות כלשהי שווה ל-1. חישוב דומה ניתן לבצע עבור $E[X^2]$ ולקבל את התוצאה שציינו קודם לכן.

תרגיל 3.30 שיחות טלפון מגיעות למרכזיה בין השעות 8:00 ל-16:00 בתהליך פואסון עם קצב ממוצע של 20 שיחות בשעה.

1. מה ההסתברות שב-10 הדקות הראשונות התקבלו 2 שיחות? מה ההסתברות שב-5 הדקות האחרונות התקבלה שיחה אחת?

2. מה ההסתברות שהשיחה הראשונה תגיע למרכזיה בין דקה ל-3 דקות לאחר הפתיחה?

3. אף שיחה לא הגיעה ב-3 הדקות הראשונות. מה ההסתברות שנמתין דקה נוספת לפחות עד שתגיע השיחה הראשונה?

פתרון: נסמן ב- X_t את המ"מ המייצג את כמות השיחות שהגיעו מרגע הפתיחה ועד זמן t [בשעות]. זהו מ"מ פואסוני המתפלג $X_t \sim P(20t)$.

1. כמות השיחות ב-10 הדקות הראשונות, $X_{1/6}$, מתפלג פואסוני עם פרמטר $3\frac{1}{3}$ ולכן $\Pr(X_{1/6} = 2) = \frac{(3\frac{1}{3})^2}{2!} e^{-3\frac{1}{3}} = 0.198$. מתכונת חוסר הזיכרון, מספר השיחות ב-5 הדקות האחרונות נקבע רק לפי משך פרק הזמן ולא לפי שעת ההתחלה שלו, ולכן מתפלג פואסוני עם פרמטר $\frac{20}{12}$. לכן ההסתברות שהתקבלה שיחה אחת בלבד היא $\frac{(20/12)^1}{1!} \exp(-20/12) = 0.3148$.

2. לשם הפשטות נעבור ליחידות זמן של שיחות לדקה ולכן הפרמטר הינו $\frac{1}{3}$ שיחות לדקה. פרק הזמן בין אירועים פואסוניים מתפלג מעריכית עם פרמטר $\frac{1}{3}$ ולכן $\Pr(1 \leq T \leq 3) = F_T(3) - F_T(1) = 1 - e^{-1} - (1 - e^{-1/3}) = 0.349$.

3. נשתמש בתכונת חוסר הזיכרון של מ"מ מעריכי ונקבל: $\Pr(T \geq 3 + 1 | T \geq 3) = \Pr(T \geq 1) = 1 - F_T(1) = 0.717$.

תרגיל 3.31 חלקיקים נפלטים מחומר רדיואקטיבי בתהליך פואסון עם קצב של 0.5 חלקיקים לשנייה. חשבו את ההסתברויות הבאות:

1. ייפלט לפחות חלקיק אחד בשנייה נתונה.

2. ייפלטו יותר מ-3 חלקיקים ב-5 שניות.

3. בחמש השניות הראשונות נפלטו שלושה חלקיקים. מהי התפלגות מספר החלקיקים שנפלטו בשנייה הראשונה?

פתרון:

$$1. \Pr(X \geq 1) = 1 - \Pr(X = 0) = 1 - e^{-0.5} \frac{0.5^0}{0!} = 0.393$$

$$2. \Pr(X_5 > 3) = 1 - \Pr(X_5 = 0) - \Pr(X_5 = 1) - \Pr(X_5 = 2) - \Pr(X_5 = 3) = 0.242$$

3. נסמן ב- X_1 את כמות החלקיקים שנפלטו בשניה הראשונה, X_5 את כמות החלקיקים שנפלטו ב-5 השניות הראשונות.

$$\Pr(X_1 = k | X_5 = 3) = \frac{\Pr(X_1 = k, X_5 = 3)}{\Pr(X_5 = 3)} \stackrel{*}{=} \frac{\Pr(X_1 = k, X_5 - X_1 = 3 - k)}{\Pr(X_5 = 3)} \stackrel{**}{=} \frac{\Pr(X_1 = k) \Pr(X_5 - X_1 = 3 - k)}{\Pr(X_5 = 3)}$$

כאשר ב-* הפרדנו את הזמנים לשני פרקי זמן בלתי תלויים: השנייה הראשונה ו-4 השניות שלאחר מכן וב-** ניצלנו את חוסר התלות בין פרקי זמן אלו. אם נציב את המספרים בנוסחא לתהליך פואסון נקבל $\binom{3}{k} (\frac{1}{5})^k (\frac{4}{5})^{3-k}$. כלומר, בהינתן שב-5 שניות נפלטו 3 חלקיקים, כמות החלקיקים בשנייה הראשונה מתפלג בינומית $(3, \frac{1}{5})$. כלומר, ניתן להתייחס לכל חלקיק שנפלט כ-"ניסוי" והסתברות שהחלקיק נפלט דווקא בפרק זמן של שנייה אחת ספיציפית מתוך 5 השניות הוא חמישית.

4 משתנים מקריים דו-מימדיים

עד כה עסקנו במשתנים מקריים חד-מימדיים בדידים. במסגרת ההרחבה הזאת יכולנו לנתח לא מעט גורמים חדשים כגון ממוצעים ופיזור, ועדיין הדיון חסר מעט ביחס לנדרש בעולם האמיתי. ניקח לדוגמא סקרי אוכלוסין שמבצעת הלשכה המרכזית לסטטיסטיקה (הלמ"ס). הלמ"ס בודקת את הכנסות משקי הבית, הוצאות משקי הבית, תנאי המחיה והדיוור, מספר נפשות, רכבים, ועוד. ברוב המקרים אנחנו נרצה לדעת מה הקשר בין הגורמים השונים - כיצד ההכנסה של משקי הבית תלויה באזור המגורים ומספר הנפשות? כיצד מתפלג מספר הרכבים בהתאם למספר הנפשות ולהכנסה, וכן ביחס למקום המגורים? כיצד משתנה ההכנסה לנפש כפונקציה של מרחק האוכלוסיה ממרכז הארץ? ועוד. כל השאלות הללו הן קריטיות לשיקולי מדיניות. הקשר בין המשתנים יירמז לנו על דרכי פעולה אפשריות לשיפור הרווחה החברתית (שיפור תשתיות, תחבורה ציבורית, סבסוד צהרונים, העלאת או קיצוץ קצבאות וכן הלאה).

בכדי להתמודד עם הסוגיות הללו, בפרק הנוכחי נעסוק במשתנים מקריים דו-מימדיים. נתחיל בניתוח הבסיסי של התפלגות משותפת עבור צמד משתנים מקריים דו-מימדיים בדידים, אשר מרכיבים משתנה מקרי דו-מימדי בדיד. נתחיל לדון בהתפלגויות המשותפות ובהתפלגויות השוליות, ובהמשך נעבור לדון במתאם (קורלציה) בין המשתנים השונים.

נפתח בדוגמא. אלו, בר ובן זורקים 2 קוביות הוגנות סה"כ. זהו הניסוי אשר עומד בבסיס הדוגמא שלנו. כעת נרצה לשאול מספר שאלות שונות על הניסוי. הראשונה תהיה מה סכום ההטלות? סכום ההטלות הוא אקראי ונסמן אותו במשתנה המקרי X . השאלה השנייה שתעסיק אותנו היא מה הערך המקסימאלי שהתקבל? גם ערך זה הוא אקראי ונסמנו במשתנה המקרי Y . נשים לב שצמד המשתנים מוגדרים על אותו מרחב מדגם אחד המורכב מזוגות (i, j) כאשר $1 \leq i, j \leq 6$ שלמים. זאת אומרת, אלו שתי פונקציות שונות המוגדרות על אותו המרחב ולכן נוכל לשאול כיצד הערך של משתנה אחד מתייחס לערך של משתנה אחר. לדוגמא, אם $Y = 4$, אזי בהכרח $X > 4$ שכן X הוא סכום התוצאות, ומנגד אם $X = 2$, אזי בהכרח $Y = 1$, כי צמד ההטלות הניבו את הערך 1.

אם כן, מה חידשנו עד כה? לא הרבה, למען האמת. באופן כללי, משתנים מקריים דו-מימדיים הם בסה"כ צירוף של צמד משתנים מקריים חד-מימדיים והצגתם כוקטור (X, Y) . באופן דומה, משתנים מקריים תלת-מימדיים הם וקטור של 3 משתנים מקריים חד מימדיים, וכן הלאה למימדים גבוהים יותר. נעבור להגדרות פורמאליות.

4.1 תומך והתפלגות משותפת

הגדרה 4.1 (משתנה מקרי דו-מימדי) עבור X ו- Y משתנים מקריים חד-מימדיים המוגדרים על מרחב $(\Omega, \Pr)$, נגדיר משתנה מקרי חדש, דו-מימדי על ידי $(X, Y) : \Omega \times \Omega \rightarrow \mathbb{R}^2$.

נשים לב שפונקציית ההסתברות $\Pr$ לוקחת ערכים ממרחב המדגם מחזירה מספר. זאת אומרת, מרחב ההסתברות עצמו (הניסוי בבסיס השאלה!) לא השתנה כלל. זה אותו מבנה של מרחב הסתברות הנלמד מתחילת הקורס. ההבדל היחיד הוא ההסתכלות שלנו על צמד משתנים ומאורעות מתאימים להם באופן הבא. עבור $k, l \in R$ מספרים ממשיים, המאורע $\{X = k, Y = l\}$ מוגדר על ידי $\{\omega \in \Omega \mid X(\omega) = k, Y(\omega) = l\}$. ההסתברות מוגדרת בהתאם לסכום ההסתברויות של כל איברי ω המרכיבים את המאורע, ועל פי פונקציית ההסתברות $\Pr$. נשתמש במונחים הללו בכדי להגדיר התפלגות דו-מימדית.

הגדרה 4.2 (התפלגות דו-מימדית) לכל משתנה דו מימדי (X, Y) ישנה התפלגות דו-מימדית $P_{X,Y} : \mathbb{R}^2 \rightarrow [0, 1]$ כך ש-

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

התפלגות דו מימדית היא אי-שלילית, והתומך של (X, Y) מוגדר כאוסף הנקודות בהן התפלגות היא חיובית ממש - $Supp(X, Y) = \{k, l \in \mathbb{R} \mid P_{X,Y}(k, l) > 0\}$. באנלוגיה למקרה החד-מימדי, התפלגות דו-מימדית היא פונקציה מנורמלת כך שסכימה על כלל התומך של המשתנה הדו-מימדי תניב $\sum_{k,l \in Supp(X,Y)} P_{X,Y}(k, l) = 1$.

הערה 4.3 (הצגת התפלגות דו-מימדית) ישנם שני אופנים להציג התפלגות דו-מימדית. הראשונה היא באמצעות טבלה דו-מימדית, אשר סוקרת את ערכי X, Y והשנייה היא באמצעות נוסחה. אלו גם שתי הדרכים אשר בחנו במקרה החד-מימדי. השימוש בטבלה או בנוסחה הוא

תלוי מצב. עבור משתנים מקריים לא חסומים, אנחנו נאלץ להשתמש בנוסחה (אלא אם כן תרצו "לשרוף זמן" על שרטוט טבלה אינסופית). מנגד, במקרים סופיים שימוש בטבלה יכול להיות די נוח, במיוחד לחישוב התפלגויות שוליות, כפי שנלמד מיד.

4.1.1 התפלגויות שוליות

התפלגות שולית של X (או של Y) במשתנה הדו-מימדי (X, Y) היא ההתפלגות של משתנה X . המילה "שולית" מתווספת בכדי להבדיל בין ההתפלגות המקורית הדו-מימדית להתפלגויות החד-מימדיות. חישוב התפלגות שולית הופך פשוט יותר כאשר נתונה לנו ההתפלגות המשותפת. אם נתונה ההתפלגות $P_{X,Y}$, אז נוכל למצוא את ההתפלגות השולית באמצעות סכימה על ערכי המשתנה האחר. לדוגמא,

$$\begin{aligned} P_X(k) &= \Pr(X = k) \\ &= \sum_{l \in \text{Supp}(Y)} \Pr(X = k, Y = l) \\ &= \sum_{l \in \text{Supp}(Y)} P_{X,Y}(k, l), \end{aligned}$$

כאשר השורה השנייה היא פירוק של הביטוי בשורה הראשונה לחלקים (בדומה לנוסחת ההסתברות השלמה). כאשר ההתפלגות המשותפת מוצגת בטבלה, אז הסכימה היא מיידית (לאורך עמודות או שורות, בהתאם למשתנה). בנוסף, בטבלה די קל לזהות טעויות באמצעות סכימת כלל הערכים ויודא כי הסכום מנורמל ל-1 כנדרש. נעזר בדוגמא הבאה להמחשה.

תרגיל 4.4 מטילים שתי קוביות הוגנות, לבנה ושחורה. X התוצאה של הקוביה הלבנה, Y התוצאה המקסימלית בין השתיים. מצאו את ההתפלגות המשותפת של (X, Y) ואת ההתפלגויות השוליות.

פתרון: צריך למצוא לכל $k, l \in \{1, \dots, 6\}$ את $P_{X,Y}(k, l)$. נשתמש בטבלה כאשר כל תא בטבלה מתאר את ההסתברות שמתאימה לשורה והעמודה הרלוונטיים, קרי $\Pr(X = k, Y = l)$:

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

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

הסבר:

- עבור $k > l$ (כלומר $X > Y$) לא ייתכן שבקוביה הלבנה נקבל תוצאה הגבוהה יותר ממקסימום הקוביות, לכן $\Pr(X = k, Y = l) = 0$.
- עבור $k = l$ (כלומר $X = Y$) בקוביה הלבנה צריך לצאת k , ובשחורה יכול לצאת כל מספר הקטן או שווה ל- k , לכן $\Pr(X = k, Y = l) = \frac{1}{6} \cdot \frac{k}{6} = \frac{k}{36}$.
- עבור $k < l$ (כלומר $X < Y$) בקוביה הלבנה צריך לצאת k ובשחורה l , לכן $\Pr(X = k, Y = l) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$.

כפי שרואים בטבלה, ניתן למצוא את ההתפלגויות השוליות של X או Y מסכימת עמודות או שורות הטבלה בהתאמה. בנוסף, אנחנו יכולים לוודא אם התשובה שלנו נכונה על ידי סכימה ובדיקה שסכום ההסתברויות שווה ל-1. ראוי לציין שההסבר הוא, בפני עצמו, הצגה של ההתפלגות באמצעות נוסחה.

4.2 התפלגות מותנה

בהמשך לקו עימו התחלנו את הפרק, ברגע שיש בידינו התפלגות משותפת של (X, Y) , נרצה לדעת כיצד משתנה אחד מתנהג כאשר השני נותר קבוע. לדוגמה, אם X מייצג את ההכנסה לנפש ו- Y מייצג את מקום המגורים בכפונקציה של המרחק ממרכז הארץ, אז נוכל לקבע את המרחק ממרכז הארץ, $Y = l$, ולבחון את התפלגות ההכנסה לנפש במרחק הנתון. בכך מתקבל משתנה מקרי חדש, המסומן על ידי $X|Y = l$, ובמילים - X בהינתן $Y = l$ (ההכנסה לנפש בהינתן מרחק l ממרכז הארץ). ההגדרה של **התפלגות מותנה** מבוססת על הסתברות מותנה אשר ראינו בחלק הראשון של הקורס, בדיון על הסתברות בסיסית. ההתפלגות של $X|Y = l$ היא פונקציה ממשית המוגדרת על ידי

$$P_{X|Y=l}(k) = \Pr(X = k|Y = l) = \frac{\Pr(X = k, Y = l)}{\Pr(Y = l)}.$$

קרי, הסיכוי ש- $X = k$, תחת ההנחה ש- $Y = l$. נשתמש בדוגמה הבאה בכדי לחשב התפלגות מותנה. החלק החשוב להבנה הוא ש- $X|Y = l$ זה משתנה מקרי חדש - הוא תלוי במשתנים X ו- Y , אבל הוא משתנה בפני עצמו וניתן להסתכל על ההתפלגות שלו, התוחלת, והשונות שלו וכן הלאה. זכרו זאת כאשר נדון בסעיף הבא בתוחלת מותנה.

תרגיל 4.5 בקלמר יש 2 עטים אדומים, אחד ירוק ואחד כחול. מוציאים מהקלמר את העטים בזה אחר זה עד שיוצא הירוק. נסמן ב- X את מספר העטים (כולל הירוק) שמוציאים וב- Y את מספר האדומים שיצאו. רשמו את ההתפלגות המשותפת של X, Y . כיצד מתפלג X בהינתן $Y = 1$?

פתרון: נבחן את המאורעות האפשריים. למשל $P(X = 1, Y = 0) = \frac{1}{4}$ כי זו ההסתברות שהוצאנו בדיוק עט אחד וזהו עט ירוק. בנוסף, $P(X = 1, Y > 0) = 0$ כי מצב כזה לא ייתכן. באופן דומה, $P(X = 2, Y = 1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$ כי זו ההסתברות שנוציא עט אדום אחד בהתחלה ומיד אחריו עט ירוק. ניתן להמשיך את החישובים הנ"ל ונקבל את הטבלה הבאה:

$Y \backslash X$	1	2	3	4	P_Y
0	$\frac{1}{4}$	$\frac{1}{12}$	0	0	$\frac{1}{3}$
1	0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$
2	0	0	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$
P_X	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

טבלה 2: התפלגות משותפת של (X, Y)

נשים לב שקיבלנו התפלגויות שוליות אחידות: $X \sim U\{1, 2, 3, 4\}$ וכן $Y \sim U\{0, 1, 2\}$. נעבור לסעיף הבא. בהינתן $Y = 1$ ניתן לראות מהטבלה ש- X יכול להיות 2 או 3 בלבד, וחישוב ההסתברויות מראה שההסתברות לכל תוצאה היא בדיוק $0.5 = \frac{1}{2}$. לכן $X|Y = 1 \sim U\{2, 3\}$, כלומר בהינתן $Y = 1$ המ"מ X הוא בדיד אחיד על הקבוצה $\{2, 3\}$.

תרגיל 4.6 בן נוסע לקניון בלאס וגאס להמר על קוביות. הוא מהמר על קובייה הוגנת המוטלת שוב ושוב. נגדיר את X להיות מספר ההטלות עד שיוצא לראשונה המספר 6, ו- Y מספר ההטלות עד שיוצא לראשונה מספר זוגי קטן מ-6, ז"א 2 או 4.

1. מצאו את ההתפלגות המשותפת של (X, Y) , ואת ההתפלגויות השוליות.

2. מצאו את התפלגות $Y|X=2$.

3. מצאו את התפלגות $W = \min\{X, Y\}$.

פתרון: נענה על כל סעיף בנפרד.

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

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

הסבר: כאשר $k > l$, בסדרת ההטלות יצא תחילה 2 או 4 (בהטלה ה- l), ורק לאחר מכן יצא 6 (בהטלה ה- k). לכן ב- $l-1$ ההטלות הראשונות קיבלנו רק 1, 3, 5 (ההסתברות לכך היא $\left(\frac{3}{6}\right)^{l-1}$), בהטלה ה- l קיבלנו 2 או 4 (ההסתברות לכך היא $\frac{2}{6}$), ומאותה הטלה והלאה עד ההטלה ה- k יכלו לצאת כל המספרים למעט 6 (בהסתברות $\left(\frac{5}{6}\right)^{k-l-1}$). בהטלה ה- k בהכרח יצא 6 (בהסתברות $\frac{1}{6}$). החישוב השני מתבצע באופן דומה. בכדי למצוא את ההתפלגויות השוליות, אין צורך להתרכז בהתפלגות המשותפת. למעשה מדובר בסדרה של הטלות בלתי-תלויות וצמד המשתנים המקריים סופרים את מספר ההטלות עד לפעם הראשונה שמתרחש תנאי מסויים, בפרט עד שיוצא 2 או 4 או עד שיוצא 6. לכן אלו משתנים מקריים גיאומטריים על פי הגדרתם, $X \sim G\left(\frac{1}{6}\right)$, $Y \sim G\left(\frac{2}{6}\right)$.

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

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

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

4.3 תוחלת מותנה ונוסחת התוחלת השלמה

בסעיף הקודם ציינו כי משתנה מותנה הוא משתנה מקרי חדש העומד בזכות עצמו ובעל התפלגות משל עצמו, לכן ניתן לחשב את התוחלת והשוונות של המשתנה המותנה. התוחלת של המשתנה המותנה נקראית **תוחלת מותנה**, מסומנת על ידי $E[X|Y=l]$ ומחושבת בעזרת ההתפלגות המותנה כ- $E[X|Y=l] = \sum_{k \in \text{supp}(X)} k \Pr(X=k|Y=l)$. הגדרה זאת איננה שונה מהותית מהגדרת התוחלת שראינו קודם לכן. משקללים את ערכי המשתנה $X|Y=l$ בעזרת ההסתברות לקבלת כל ערך. לכן, אין כל חידוש בהגדרה האחרונה של תחלת מותנה.

נשתמש בתרגיל קודם בכדי להדגים את הרעיון בבסיס תוחלת מותנה. מטילים שתי קוביות הוגנות, לבנה ושחורה. X מייצג את התוצאה של הקוביה הלבנה, ו- Y מייצג את התוצאה המקסימאלית בין השתיים. נבחן את התוחלת של כל משתנה בנפרד. ראינו כי $X \sim U[1, 6]$ לכן $E[X] = 3.5$ ו- $E[Y] = \sum_{l=1}^6 l \cdot \frac{2l-1}{36} = \frac{161}{36} \approx 4.47$. מה ייקרה לתוחלת של X אם נקבע את Y לערך כלשהו? לדוגמא, $X|Y=1$ זה משתנה מקרי המקבל את הערך 1 בהסתברות 1 ולכן $E[X|Y=1] = 1 < E[X]$. מנגד, $X|Y=6$ מקבל ערכים מ-1 עד 6 והתפלגותו היא

$$\begin{aligned} \Pr(X = k|Y = 6) &= \frac{\Pr(X = k, Y = 6)}{\Pr(Y = 6)} \\ &= \begin{cases} \frac{1}{11}, & k = 1, 2, 3, 4, 5, \\ \frac{6}{11}, & k = 6, \end{cases} \end{aligned}$$

לכן, $E[X|Y=6] = \frac{1}{11} \cdot (1 + 2 + 3 + 4 + 5) + \frac{6}{11} \cdot 6 \approx 4.64 > E[X]$. בחישוב אחד קיבלנו ערך הגדול מהתוחלת של X ובחישוב אחר קיבלנו ערך קטן יותר. זה די הגיוני שכן כל הסטואציות הללו אמורות להשתקלל יחד ולהניב חזרה את התוחלת של X . טענה זאת מסוכמת היטב בנוסחת התוחלת השלמה הבאה.

משפט 4.7 (נוסחת התוחלת השלמה) עבור כל צמד משתנים מקריים מתקיים $E[X] = \sum_{l \in \text{supp}(Y)} E[X|Y=l] \Pr(Y=l)$. התוחלת השלמה היא אנלוגיה לנוסחת ההסתברות השלמה. ההוכחה שלה דומה מאוד להוכחת נוסחת ההסתברות השלמה, והטענה המרכזית העולה ממנה היא - ניתן לחשב את התוחלת של X בעזרת חלוקת המרחב למאורעות לפי Y , ושקלול כלל הערכים המתקבלים בהתאם להתפלגות של Y . לרוב, נהוג לסמן את התוחלת השלמה על ידי $E[X] = E_Y[E_X(X|Y)]$ (אם יהיה זמן בהמשך, נעמיק יותר בניסוח המדובר).

תרגיל 4.8 ניקח צמד מספרים $0 < p, q < 1$ ונניח כי מספר שנות המיתון בפונטי-פנדי הוא $X \sim \text{Bin}(n, p)$. מספר החברות שנסגרות בשנת המיתון ה- k הוא משתנה מקרי בינומי $Y_k \sim \text{Bin}(k, q)$ עם פרמטרים Y_k . נסמן ב- Y את מספר החברות הנסגרות במהלך המיתון. חשבו את התוחלת של Y .

פתרון: בכדי לחשב את התוחלת של מספר החברות שנסגרות במהלך מיתון, צריך לדעת תחילה כמה שנות מיתון מתרחשות בפועל. קרי, צריך להתנות על ערכי X . בהינתן $X = k$, נתון $Y = \sum_{l=0}^k Y_l$ לכן

$$E[Y|X=k] = E\left[\sum_{l=0}^k Y_l\right] = \sum_{l=0}^k E[Y_l] = \sum_{l=0}^k lq = q \sum_{l=0}^k l = q \frac{k(k+1)}{2}.$$

מכאן נובע

$$\begin{aligned} E[Y] &= \sum_{k=0}^{\infty} E[Y|X=k] \Pr(X=k) \\ &= \sum_{k=0}^{\infty} q \frac{k(k+1)}{2} \Pr(X=k) \\ &= \frac{q}{2} \sum_{k=0}^{\infty} k(k+1) \Pr(X=k) \\ &= \frac{q}{2} \sum_{k=0}^{\infty} (k^2 + k) \Pr(X=k). \end{aligned}$$

נפתח סוגרים ונפצל לשני סכומים,

$$\begin{aligned} E[Y] &= \frac{q}{2} \left[\sum_{k=0}^{\infty} k^2 \Pr(X=k) + \sum_{k=0}^{\infty} k \Pr(X=k) \right] \\ &= \frac{q}{2} [E[X^2] + E[X]] \\ &= \frac{q}{2} [V(X) + (E[X])^2 + E[X]] \\ &= \frac{q}{2} [np(1-p) + n^2p^2 + np] \\ &= \frac{qnp}{2} [np - p + 2]. \end{aligned}$$

4.4 אי תלות בין משתנים מקריים

בחלק הראשון של הקורס, בעודנו לומדים הסתברות בסיסית, עסקנו בנושא של אי-תלות בין מאורעות. אי תלות בין צמד מאורעות קובעת כי קיומו או אי-קיומו של מאורע אחד איננה משפיעה, מבחינה הסתברותית, על קיומו או אי-קיומו של מאורע אחר. הגדרה זאת ניתנת להרחבה גם בהקשר של משתנים מקריים. באופן דומה לאי-תלות של צמד מאורעות, משתנים מקריים הינם בלתי תלויים אם הערך שמשתנה אחד מקבל אינו משפיע, הסתברותית, על הערך שהשני מקבל. לדוגמה: הטלת צמד קוביות. למרות שהקוביות בעלות אותה התפלגות ומוגדרות על אותו מרחב מדגם, הערך שמתקבל מהטלה אחת אינו משפיע על הערך של ההטלה האחרת. אותו הדבר מתקבל מהטלת מספר מטבעות, וכן הלאה. נעבור להגדרה המדויקת.

הגדרה 4.9 (אי-תלות של צמד משתנים מקריים) מ"מ X ו- Y הם בלתי תלויים אם ורק אם לכל צמד מספרים $k, l \in \mathbb{R}$ מתקיים $\Pr(X=k, Y=l) = \Pr(X=k) \cdot \Pr(Y=l)$.

למעשה, משתנים מקריים הם בלתי-תלויים אם ההתפלגות המשותפת היא מכפלה של ההתפלגויות השוליות, $P_{X,Y}(k,l) = P_X(k) P_Y(l)$, בכל נקודה (k,l) . דרישות האי-תלות היא מאוד חזקה, השוויון לעיל צריך להתקיים עבור **כל צמד מספרים** שנבחר, לכן אם קיים זוג אחד של מספרים k, l עבורם השוויון $P_{X,Y}(k,l) = P_X(k) P_Y(l)$ אינו מתקיים, אזי המשתנים תלויים! בנוסף, ניתן להגדיר אי תלות בין משתנים מקריים בדרך אחרת - X ו- Y הם ב"ת אם לכל $k, l \in \mathbb{R}$ מתקיים $\Pr(X=k|Y=l) = \Pr(X=k)$. ההגדרות הנ"ל תקפות עבור צמד משתנים מקריים, והכללה עבור מספר כלשהו של משתנים מקריים היא ישירה.

תרגיל 4.10 מספר ביצי החופש שמטילה תרנגולת מתפלג אחיד בין 1 ל-4. מכל ביצה עתיד לבקוע אפרוח מאושר בהסתברות $\frac{1}{3}$ באופן ב"ת בביצים האחרות ובמספר הביצים שהוטלו. יהי X מספר הביצים שהוטלו ו- Y מספר האפרוחים שבקעו.

1. מצאו את ההתפלגות המשותפת של (X, Y) .

2. האם X, Y תלויים?

3. מצאו כיצד מתפלג $X|Y=3$.

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

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

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

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

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

שכן $\Pr(X = 1) = \frac{1}{4} > 0$ ו- $\Pr(Y = 4) = \Pr(X = 4, Y = 4) = \frac{1}{4} \cdot \left(\frac{1}{3}\right)^4 > 0$. על כן, המ"מ הללו תלויים.

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

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

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

תרגיל 4.11 יהיו X, Y זוג מ"מ בדידים שההתפלגות המשותפת שלהם נתונה על ידי $\Pr(X = x, Y = y) = \begin{cases} \frac{xy}{27} & x = 1, 2; y = 2, 3, 4 \\ 0 & otherwise \end{cases}$

האם הם בלתי תלויים?

פתרון: חישוב ישיר מלמד כי

$Y \backslash X$	1	2	P_Y
2	$\frac{2}{27}$	$\frac{4}{27}$	$\frac{2}{9}$
3	$\frac{3}{27}$	$\frac{6}{27}$	$\frac{1}{3}$
4	$\frac{4}{27}$	$\frac{8}{27}$	$\frac{4}{9}$
P_X	$\frac{1}{3}$	$\frac{2}{3}$	1

טבלה 3: התפלגות משותפת ל- (X, Y)

ומתקיים $\Pr(X = x, Y = y) = \Pr(X = x) \Pr(Y = y)$, לכן המשתנים ב"ת.

דוגמה 4.12 אלון מטיל שתי קוביות הוגנות, פעם אחת כל אחת. X - התוצאה של הקוביה הראשונה, Y - התוצאה המקסימאלית. נמצא את ההתפלגויות של $Y | X = 2$ ו- $X | Y = 3$. האם X, Y תלויים?

פתרון: נחשב ישירות, בעזרת התוצאות שכבר מצאנו בתרגיל קודם.

$$\begin{aligned}\Pr(Y = 1 | X = 2) &= \frac{\Pr(X = 2, Y = 1)}{\Pr(X = 2)} = \frac{0}{1/6} = 0, \\ \Pr(Y = 2 | X = 2) &= \frac{\Pr(X = 2, Y = 2)}{\Pr(X = 2)} = \frac{2/36}{1/6} = \frac{1}{3}, \\ \Pr(Y = k | X = 2) &= \frac{\Pr(X = 2, Y = k)}{\Pr(X = 2)} = \frac{1/36}{1/6} = \frac{1}{6}, \quad k \in \{3, 4, 5, 6\}.\end{aligned}$$

נעזר בטבלה לחישוב ונקבל

$$\begin{aligned}\Pr(X = 1 | Y = 3) &= \frac{\Pr(X = 1, Y = 3)}{\Pr(Y = 3)} = \frac{1/36}{5/36} = \frac{1}{5}, \\ \Pr(X = 2 | Y = 3) &= \frac{\Pr(X = 2, Y = 3)}{\Pr(Y = 3)} = \frac{1/36}{5/36} = \frac{1}{5}, \\ \Pr(X = 3 | Y = 3) &= \frac{\Pr(X = 3, Y = 3)}{\Pr(Y = 3)} = \frac{3/36}{5/36} = \frac{3}{5}, \\ \Pr(X = k | Y = 3) &= \frac{\Pr(X = k, Y = 3)}{\Pr(Y = 3)} = \frac{0}{5/36} = 0, \quad k \in \{4, 5, 6\}.\end{aligned}$$

נראה שההגדרה לאי-תלות אינה מתקיימת בעזרת שני ערכים ממשיים, $\Pr(X = 1 | Y = 3) = \frac{1}{5} \neq \frac{1}{6} = \Pr(X = 1)$. במקרה זה השתמשנו בדרך האלטרנטיבית להוכיח תלות בין משתנים.

תרגיל 4.13 נגדיר מרחב מדגם Ω באופן הבא: המרחב כולל את כל התמורות האפשריות של $\{1, 2, 3\}$ ובנוסף את השלושת $(1, 1, 1)$, $(2, 2, 2)$, ו- $(3, 3, 3)$ ונניח כי זהו מרחב הסתברות סימטרי. נסמן X_i - הערך של הקוארדינטה ה- i בוקטור שנבחר. עבור כל $i \neq j$, האם X_i, X_j בלתי תלויים? האם X_1, X_2, X_3 בלתי תלויים?

פתרון: מרחב המדגם מונה בסה"כ 9 אפשרויות. כל קוארדינטה מקבלת כל ערך 3 פעמים ולכן עבור כל קוארדינטה k וכל ערך אפשרי l נקבל $\Pr(X_k = l) = \frac{1}{3}$. כל צירוף של שני ערכים מתאים לשלישיה יחידה מתוך 9 השלישיות האפשריות ולכן $\Pr(X_i = k, X_j = l) = \frac{1}{9}$ וזו בדיוק מכפלת ההסתברויות. לכן לכל $i \neq j$ המשתנים הם ב"ת. לעומת זאת, עבור שלושה משתנים מקריים המצב משתנה.

$$\Pr(X_1 = 1, X_2 = 1, X_3 = 2) = 0 \neq \Pr(X_1 = 1) \Pr(X_2 = 1) \Pr(X_3 = 2),$$

וקיימת תלות.

4.5 שונות משותפת ומקדם מתאם

בחלק הקודם ראינו כי הדרישה לאי תלות היא נוקשה שכן תנאי השיוויון אמור להתקיים עבור כל צמד ערכים אפשריים. בפועל ישנם לא מעט מקריים בהם קיימת תלות כלשהי בין המשתנים, ועדיין אנו מעוניינים לנתח את הקשר בין המשתנים תחת הגדרה יותר עדינה. לשם כך מגדירים מתאם (קורולציה) בין משתנים מקריים. נתחיל בדוגמא ע"ב תרגילי ביצי החופש שראינו בפרק הקודם.

איש מכירות משווק מכונות כביסה דרך הטלפון. בכל שעה הוא פונה ל- X אנשים, כאשר $X \sim U, [0, 3]$, ומכל שיחה נוצרת מכירה בסיכוי 0.5, באופן בלתי תלוי במספר השיחות או במכירות האחרות. נסמן ב- Y את מספר המכונות הנמכרות בשעה. נוכל די בקלות למצוא את ההתפלגות המשותפת של צמד המשתנים,

$$\begin{aligned} P_{X,Y}(k,l) &= \Pr(X=k, Y=l) \\ &= \Pr(Y=l|X=k) \Pr(X=k) \\ &= \binom{k}{l} \left(\frac{1}{2}\right)^l \left(\frac{1}{2}\right)^{k-l} \cdot \frac{1}{4}, \quad \forall k=0,1,2,3, 0 \leq l \leq k. \end{aligned}$$

הוכחת התלות בין המשתנים היא טריוויאלית. כל שנדרש זה למצוא זוג ערכים עבורם השיוויון שמגדיר אי תלות אינו מתקיים. לדוגמא, $\Pr(X=0, Y=1) = 0$ בעוד $\Pr(X=0) \Pr(Y=1) > 0$. הסבר: לא ייתכן מצב בו מספר המכירות גדול ממש ממספר השיחות ($l=1 > k=0$), בעוד בהחלט ייתכן כי מספר השיחות של איש המכירות יהיה 0, או לחילופין, מספר המכירות יהיה 1 (לא ב־זמנית!). על כן המשתנים הנ"ל תלויים.

קעת נוכל לשאול שאלה אחרת בנוגע לממוצעים. בפרט, כיצד בממוצע משתנה Y כאשר X גדל? ובכן, $E[Y|X=0] = 0$ באופן מיידי, אבל $E[Y|X=3] = 1.5$ כי $Y|X=3 \sim \text{Bin}(3, 0.5)$. נשים לב ש- $E[Y|X=3] > E[Y] < E[Y|X=0]$. זאת אומרת, יש בממוצע קשר חיובי בין המשתנים. הקשר הזה יחסית אינטואיטיבי. ככל שיש יותר שיחות, גם מספר המכירות הממוצע גדל. חשוב להבין שמספר השיחות לא בהכרח ייגדל, אלא שהוא גדל בממוצע.

4.5.1 שונות משותפת

הדוגמא האחרונה ממחישה את הרעיון שבבסיס השונות המשותפת ומקדם המתאם בין המשתנים. השונות המשותפת ומקדם המתאם מהווים מדד למידת התיאום הממוצע בין המשתנים. בפועל הם בודקים **חמישה סוגי קשרים בין המשתנים**: קשר חיובי, קשר שלילי, קשר לינארי חיובי, קשר לינארי שלילי, או שאין קשר בממוצע בין המשתנים. נעבור להגדרה הפורמאלית.

הגדרה 4.14 (שונות משותפת) Covariance, השונות המשותפת, של צמד מ"מ X ו- Y היא $Cov(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$.

במילים, השונות המשותפת מוגדרת ע"י תוחלת המכפלה $(X - \mu_x)(Y - \mu_y)$, ולפני שנעסוק במשמעות שלה נתחיל בשאלה טכנית - מדוע בוחנים את המשתנים בחיסור תוחלת שלהם? ובכן, המשתנה $X - \mu_x$ הוא בעל תוחלת אפס (בשל לינאריות התוחלת), אבל התפלגותו דומה להתפלגות של X , למעט הזזה בקבוע. אותו הדבר נכון עבור $Y - \mu_y$. השמטת הקבוע מגיעה כחלק מתהליך נרמול המשתנים. בכדי לייצר סקאלה אחידה, אנחנו מחסרים את התוחלת ואז כל המשתנים הם בעלי תוחלת אפס. בנוסף, בעזרת מספר פעולות אלגבריות ושימוש בלינאריות התוחלת ניתן להראות ש-

$$\begin{aligned} Cov(X, Y) &= E((X - E(X))(Y - E(Y))) = \\ &= E(XY - XE(Y) - YE(X) + E(X)E(Y)) = \\ &= E(XY) - E(XE(Y)) - E(YE(X)) + E(E(X)E(Y)) = \\ &= E(XY) - 2E(X)E(Y) + E(X)E(Y) = \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

הנוסחה האחרונה מאפשרת לנו בקלות לחשב את השונות המשותפת. אם נתייחס לדוגמא איש המכירות איתה התחלנו, אזי $E[X] = 1.5$, $E[Y] = \frac{3}{4}$ ו-

$$E[XY] = 1 \cdot 1 \cdot \frac{1}{8} + 1 \cdot 2 \cdot \frac{1}{8} + 2 \cdot 2 \cdot \frac{1}{16} + 1 \cdot 3 \cdot \frac{3}{32} + 2 \cdot 3 \cdot \frac{3}{32} + 3 \cdot 3 \cdot \frac{1}{32} = \frac{7}{4},$$

לכן $Cov(X, Y) = \frac{7}{4} - \frac{3}{2} \cdot \frac{3}{4} = \frac{5}{8} > 0$ וקיים בין המשתנים מתאם חיובי. נעבור על מספר הגדרות הנוגעות לשונות המשותפת.

- אם $Cov(X, Y) > 0$, אז המשתנים **מתואמים חיובית**, או בעלי מתאם חיובי.
- אם $Cov(X, Y) < 0$, אז המשתנים **מתואמים שלילית**, או בעלי מתאם שלילי.
- אם $Cov(X, Y) = 0$, אז המשתנים **מבלתי מתואמים**, או בעלי מתאם אפס. שימו לב שבמקרה כזה, $E[XY] = E[X]E[Y]$, באנלוגיה מסויימת להגדרת אי-תלות, כאשר ההתפלגויות הוחלפו בתוחלות.

הנקודה האחרונה בנוגע לחוסר תיאום קשורה בצורה די הדוקה לאי תלות בין משתנים. למעשה, ניתן להוכיח באופן ישיר כי צמד משתנים בלתי תלויים הם בהכרח בלתי מתואמים, כפי שמוצג בטענה הבאה.

טענה 4.15 משתנים מקריים בלתי-תלויים הם גם בלתי מתואמים.

לעומת זאת, הטענה ההפוכה איננה נכונה. ז"א, ישנן אינספור דוגמאות של משתנים בלתי מתואמים ותלויים! לדוגמא, נניח ש- X הוא משתנה בעל התפלגות אחידה על הקבוצה $\{-2, -1, 1, 2\}$. נסתכל על $Y = X^2$. המשתנים X ו- Y תלויים שכן $\Pr(X = 2, Y = 1) = 0$ בעוד $\Pr(X = 2) \Pr(Y = 1) > 0$. יחד עם זאת, המשתנים בלתי מתואמים - $E[XY] = E[X^3] = E[X] = 0$ (ההתפלגות של X ו- X^3 הן סימטריות סביב 0 ולכן התוחלת שלהן היא 0). מצד שני, אם ניקח Y , בלתי תלוי ב- X , המתפלג אחיד על הקבוצה $\{-1, 1\}$, אזי $E[XY] = 0$ משיקולי סימטריה סביב 0, והמשתנים הבלתי תלויים יהיו גם בלתי מתואמים. בדומה לפונקציית השונות, גם לשונות המשותפת יש מספר תכונות חשובות וקלות להוכחה. ניקח צמד משתנים מקריים X ו- Y , אזי:

- **שונות משותפת היא הכללה של שונות** - $Cov(X, X) = E[XX] - E[X]E[X] = V(X)$.
- **השונות המשותפת היא סימטרית**, $Cov(X, Y) = Cov(Y, X)$.
- **שונות משותפת של טרנספורמציה ליניארית** - עבור קבועים $a, b \in \mathbb{R}$ כלשהם מתקיים $Cov(aX + b, Y) = aCov(X, Y)$. כמובן שתכונת הסימטריה מרמזת כי אותה הנוסחה נכונה גם עבור טרנספורמציה ליניארית של Y , קרי $Cov(X, aY + b) = aCov(X, Y)$. ההוכחה היא ישירה,

$$\begin{aligned} Cov(aX + b, Y) &= E[(aX + b - \mu_{aX+b})(E[Y] - \mu_Y)] \\ &= E[(aX + b - a\mu_X - b)(E[Y] - \mu_Y)] \\ &= E[a(X - \mu_X)(E[Y] - \mu_Y)] \\ &= aE[(X - \mu_X)(E[Y] - \mu_Y)] = aCov(X, Y). \end{aligned}$$

- **שונות משותפת של סכום משתנים** - עבור משתנה מקרי נוסף Z מתקיים $Cov(X + Z, Y) = Cov(X, Y) + Cov(Z, Y)$. ההוכחה היא ישירה גם כן,

$$\begin{aligned} Cov(X + Z, Y) &= E[(X + Z - \mu_{X+Z})(Y - \mu_Y)] \\ &= E[(X + Z - \mu_X - \mu_Z)(Y - \mu_Y)] \\ &= E[((X - \mu_X) + (Z - \mu_Z))(Y - \mu_Y)] \\ &= E[(X - \mu_X)(Y - \mu_Y) + (Z - \mu_Z)(Y - \mu_Y)] \\ &= E[(X - \mu_X)(Y - \mu_Y)] + E[(Z - \mu_Z)(Y - \mu_Y)] \\ &= Cov(X, Y) + Cov(Z, Y). \end{aligned}$$

נקודה זאת ניתנת להכללה עבור $n \in \mathbb{N}$ משתנים מקריים כלשהם $(X_i)_{i=1}^n$ ומ"מ Y על ידי $Cov(\sum_{i=1}^n X_i, Y) = \sum_{i=1}^n Cov(X_i, Y)$.

4.5.2 מקדם מתאם

לאחר שהגדרנו שונות משותפת, נוכל להגדיר מקדם מתאם של צמד משתנים מקריים X ו- Y . מקדם המתאם מוגדר על ידי השונות המשותפת של צמד משתנים מקריים לאחר שתוקנו כך שהשונות שלהם תהיה שווה ל-1. נזכור כי ההגדרה לשונות משותפת מבטיחה כי המשתנים הם בעלי תוחלת אפסית, $E[X - \mu_X] = 0$. בכדי להבטיח כי השונות תהיה שווה ל-1, נחלק בסטיית התקן $\sigma_X = \sqrt{V(X)}$ ונקבל

$$V\left(\frac{X - \mu_X}{\sigma_X}\right) = V\left(\frac{X}{\sigma_X} - \frac{\mu_X}{\sigma_X}\right) = V\left(\frac{X}{\sigma_X}\right) = \left(\frac{1}{\sigma_X}\right)^2 V(X) = 1.$$

נגדיר באופן רשמי את מקדם המתאם.

הגדרה 4.16 (מקדם מתאם) מקדם המתאם של צמד משתנים מקריים X ו- Y הוא

$$\rho(X, Y) = \text{Cov}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

אם כן, מקדם המתאם הוא שונות המשותפת של משתנים מקריים לאחר חיסור התוחלת וחלוקה בסטיית התקן (עבור כל משתנה בנפרד). החלוקה בסטיית התקן מבטיחה שהתוצאה שנקבל לא תהיה תלויה בפיזור של כל משתנה בפני עצמו, אלא רק ביחס בין המשתנים. כתוצאה מכך, מקדם המתאם הוא פונקציה מנורמלת שמקבלת ערכים בין -1 ל-1. הנקודות הבאות מסכמות את הערכים האפשריים של מקדם המתאם והמשמעות של כל ערך:

- $\rho(X, Y) = 1$ אם ורק אם $Y = aX + b$ כאשר $a > 0$. זאת אומרת, כאשר קיים קשר ליניארי חיובי בין צמד המשתנים המקריים, ומשתנה אחד מתקבל ע"י שינוי קנה המידה של המשתנה האחר.
- $\rho(X, Y) = -1$ אם ורק אם $Y = aX + b$ כאשר $a < 0$. זאת אומרת, כאשר קיים קשר ליניארי שלילי בין צמד המשתנים המקריים, ומשתנה אחד מתקבל ע"י שינוי קנה המידה של המשתנה האחר.
- $0 < \rho(X, Y) < 1$ אם ורק אם המשתנים מתואמים חיובית.
- $-1 < \rho(X, Y) < 0$ אם ורק אם המשתנים מתואמים שלילית.
- $\rho(X, Y) = 0$ אם ורק אם המשתנים בלתי מתואמים.

נוכיח את הנקודות בטענה הבאה.

טענה 4.17 יהיו צמד משתנים מקריים X ו- Y ונתון מקדם המתאם ביניהם $\rho(X, Y)$. אזי,

$$1. |\rho(X, Y)| \leq 1.$$

$$2. \rho(X, Y) = 1 \text{ אם ורק אם } \Pr(Y = aX + b) = 1 \text{ כאשר } a, b \text{ קבועים ו-} a > 0.$$

$$3. \rho(X, Y) = -1 \text{ אם ורק אם } \Pr(Y = aX + b) = 1 \text{ כאשר } a, b \text{ קבועים ו-} a < 0.$$

הוכחה: ראינו כי מקדם המתאם מוגדר בעזרת שונות משותפת תחת נרמול של משתנים לתוחלת 0 וסטיית תקן 1. לכן מלכתחילה נניח ש- $E[X] = E[Y] = 0$ ו- $V(X) = V(Y) = 1$ (זו הנחה שאיננה מגבילה את הכלליות שכן עבור תוחלת ושונות שונות, הגדרת מקדם

מהתאם תאפס את התוחלת ותנרמל את השונות). בעזרת פתיחת סוגריים והעברת אגפים נקבל

$$\begin{aligned} E[(X \pm Y)^2] &\geq 0 \\ E[X^2 \pm 2XY + Y^2] &\geq 0 \\ E(X^2) + E(Y^2) &\geq \pm 2E[XY] \\ V(X) + V(Y) &\geq \pm 2(E[XY] - E[X]E[Y]) \\ 2 &\geq \pm 2Cov(X, Y) \\ 2 &\geq \pm 2\rho(X, Y) \\ -1 &\leq \rho(X, Y) \leq 1 \end{aligned}$$

כאשר במעבר משורה שלישית לרביעית השתמשנו בכך ש- $E[X] = E[Y] = 0$ ובמעבר מהשורה החמישית לשישית השתמשנו ב- $\sigma_X = \sigma_Y = 1$. נמשיך להוכיח יתר הטענה. שיוויון מתקיים אם ורק אם $E[(X \pm Y)^2] = 0$ וזאת אומרת ש- $V(X \pm Y) = 0$, או במילים אחרות, $X \pm Y = \text{const}$, מה שמוכיח את הטענה. ■

תרגיל 4.18 נתונים צמד משתנים מקריים $X \sim Y \sim \text{Ber}(0.5)$. רשמו את ההתפלגות המשותפת שלהם בהנחה שהם בלתי מתואמים; מתואמים חיובית; מתואמים שלילית.

פתרון: נרשום את ההתפלגות המשותפת במקרה והם בלתי מתואמים. בכדי לקבל חוסר תיאום, נדרוש $E[XY] = E[X]E[Y] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. בנוסף, המשתנה XY הוא גם משתנה ברנולי (מכפלה של משתני ברנולי תמיד תניב משתנה ברנולי חדש), לכן $E[XY] = \Pr(XY = 1)$ כמו בכל משתנה ברנולי כלשהו. השילוב של תובנה זאת יחד עם הדרישה לחוסר תיאום גוררת כי $\Pr(XY = 1) = \frac{1}{4}$. המאורע $XY = 1$ מתקיים אם ורק אם $X = Y = 1$ ונוכל לרשום זאת בטבלת ההתפלגות המשותפת. בהתאם נוכל במהירות להשלים את יתר התאים בטבלה בעזרת ההתפלגויות השוליות. ראו טבלה 4. שימו לב שקיבלנו משתנים בלתי תלויים. נמשיך לדרישה לייצר מתאם חיובי. הדרישה למתאם

$Y \backslash X$	0	1	P_Y
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
P_X	$\frac{1}{2}$	$\frac{1}{2}$	1

טבלה 4: התפלגות משותפת כאשר המשתנים בלתי מתואמים.

חיובי קובעת כי $E[XY] > E[X]E[Y] = \frac{1}{4}$. נוכל לבחור אינספור ערכים גדולים מרבע לטובת הפתרון. נבחר $\Pr(XY = 1) = \frac{1}{3}$. בדומה לחישוב הקודם, נרשום זאת בטבלה המשותפת ונשלים את יתר התאים בעזרת התפלגויות שוליות. ראו טבלה 5. שימו לב כי יכולנו

$Y \backslash X$	0	1	P_Y
0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$
1	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$
P_X	$\frac{1}{2}$	$\frac{1}{2}$	1

טבלה 5: התפלגות משותפת כאשר המשתנים מתואמים חיובית.

באותה המידה לבחור כל ערך אחר בין רבע לחצי, ולא ניתן לחרוג מעל לחצי שכן היינו מקבלים סתירה להתפלגויות השוליות. נעבור לדרישה האחרונה לייצר מתאם שלילי. באופן דומה לפתרונות קודמים, נדרוש $\Pr(XY = 1) = E[XY] < E[X]E[Y] = \frac{1}{4}$ ואנו חופשיים לבחור

$Y \backslash X$	0	1	P_Y
0	0	$\frac{1}{2}$	$\frac{1}{2}$
1	$\frac{1}{2}$	0	$\frac{1}{2}$
P_X	$\frac{1}{2}$	$\frac{1}{2}$	1

טבלה 6: התפלגות משותפת כאשר המשתנים מתואמים שלילית.

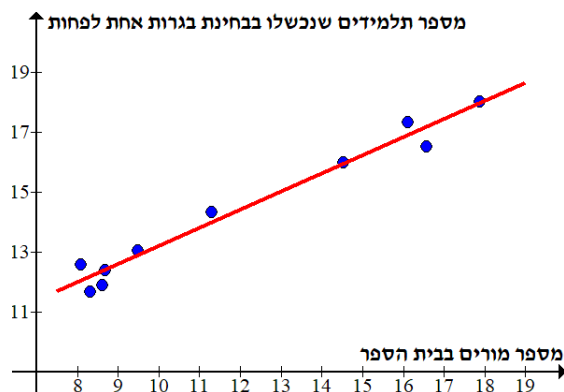
כל מספר אי שלילי קטן מרבע. נניח כי $\Pr(XY = 1) = 0$ ונמלא את הטבלה בהתאם כפי שעשינו בסעיפים הקודמים. ראו טבלה 6. שימו לב שבדוגמא האחרונה קיבלנו מתאם

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[XY] - E[X]E[Y]}{\sqrt{\frac{1}{2} \cdot \frac{1}{2}} \cdot \sqrt{\frac{1}{2} \cdot \frac{1}{2}}} = \frac{0 - \frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{4}} = -1,$$

קרי, יחס לינארי שלילי בין המשתנים. תוצאה זו איננה מפתיעה שכן המשוואה $Y = 1 - X$ מתארת את היחס בין המשתנים עם ההתפלגות המשותפת בטבלה 6.

4.5.3 דוגמאות גרפיות וקו רגרסיה

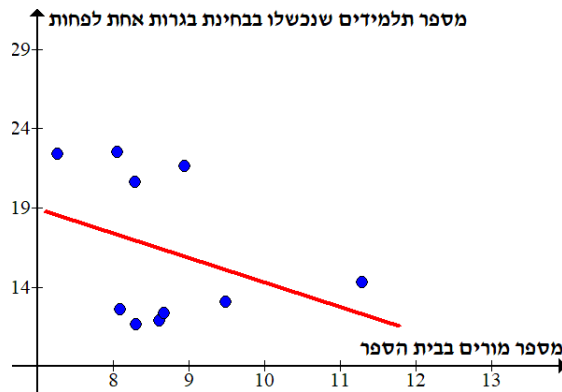
ההבדל בין מתאם ותלות איננה רק עניין טכני בתורה ההסתברותית. דווקא במדעי החיים והחברה, ההבנה של הקשרים הללו וההבדלים ביניהם היא חיונית לכל מחקר או מדיניות בהן ננקוט. נתחיל עם דוגמא פשוטה להמחשת הבעיה. בפונטיפאנדי בדקו את היחס בין מספר המורים בבית ספר למספר התלמידים שנכשלו בבחינת בגרות אחת לפחות. הגדירו משתנה X בתור מספר המורים בבית ספר (סה"כ ישנם 10 בתי ספר) ו- Y כמספר התלמידים שנכשלו בבחינת בגרות אחת לפחות. התוצאות מוצגות באיור 9. המסקנה המיידית של אלביס, החוקר



איור 9: מתאם בין מספר המורים בבתי הספר לבין מספר התלמידים שנכשלו בבחינת בגרות אחת לפחות.

הראשי בפונטיפאנדי, היא שמורים משפיעים לרעה על התלמידים וגורמים לכשלונות רבים. המדיניות החדשה עליה החליט היא קיצוץ של 50% במספר המורים בארבעת בתי הספר עם מספר הנכשלים הגבוה ביותר. כעבור שנה, הגיעו התוצאות והתבררה גודל הטעות. הקיצוץ במספר המורים באותם בתי הספר רק החמיר את הבעיה. הנתונים המעודכנים מוצגים באיור 10. בעוד המצב ביתר בתי הספר לא השתנה כלל, המצב בבתי הספר בהם מספר המורים קוצץ רק החמיר ומספר הנכשלים גדל ב-20% בממוצע. ז"א, קיצוץ מספר המורים לא שיפר את מצב התלמידים.

הדוגמא הזאת באה להמחיש את ההבדל בין קשר סטטיסטי לקשר סיבתי. קשר סיבתי הוא קשר של סיבה ותוצאה ("הפלתי את האגרטל ולכן הוא נשבר"), בעוד קשר סטטיסטי אינו בהכרח מעיד על יחס של סיבה ותוצאה בשל תלות בעוד אינספור גורמים. בדוגמא



איור 10: מתאם בין מספר המורים בבתי הספר לבין מספר התלמידים שנכשלו בבחינת בגרות אחת לפחות.

לעיל, ייתכן שבתי הספר החלשים יותר (במונחי נכשלים) מקבלים תלמידים פחות טובים, לכן מצבם הבסיסי קשה יותר ומספר המורים הגבוה בה לתקן זאת. הסתכלו באיור 11 ותראו כיצד נתונים אשר אינם קשורים אחד לשני בשום צורה (מעבר לכל ספק סביר!) עדיין מציגים מתאם כמעט מושלם.¹⁴ כלל האצבע שצריך לקחת מן הדוגמאות הללו הוא: קורולציה איננה סיבתיות (Correlation is not causation). מתאם חיובי או שלילי בין פרמטרים שונים אינו בהכרח מעיד על יחס של סיבה ותוצאה, ולעיתים יכול להעיד על יחס הפוך (לעומת המתאם) בין הגורמים השונים (כמו בדוגמת הנכשלים בבתי הספר).

לסיכום, אם תרצו להשתעשע בזמנכם הפנוי, תיכנסו לאתר <http://guessthecorrelation.com/>. האתר מציג משחק פשוט במהלכו צריך לנחש את הקורולציה בין סדרות הנתונים המוצגות על גרף. המטרה היא לנחש כמה שיותר קרוב לערך האמיתי. אומנם לא משחק מורכב ואין לו אנימציה מרשימה, אבל זו כנראה הדרך הכי קלה ומשעשעת ללמוד חומר הסתברותי וסטטיסטי (וגם להעביר את הזמן בעודכם מחפשים צידוק עצמי לכך שאינכם פותרים תרגילי בית כרגע...).

4.5.4 שונות סכום משתנים מקריים

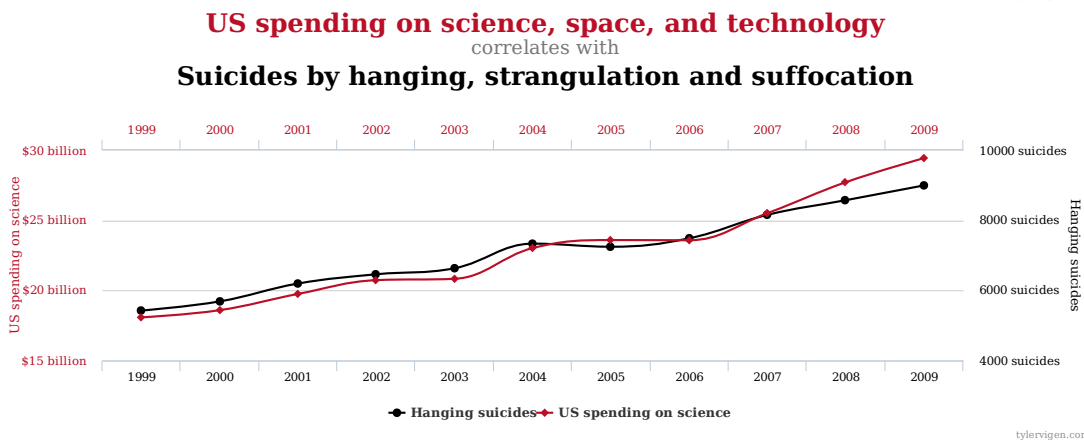
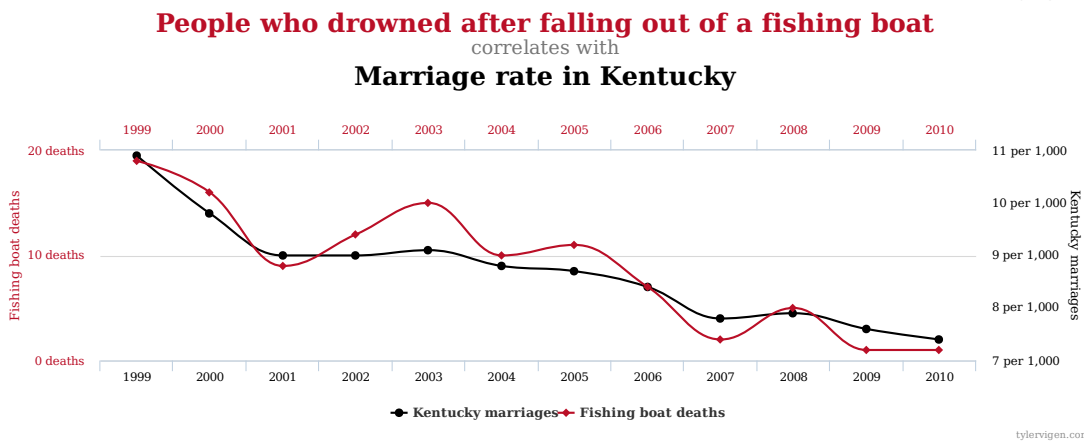
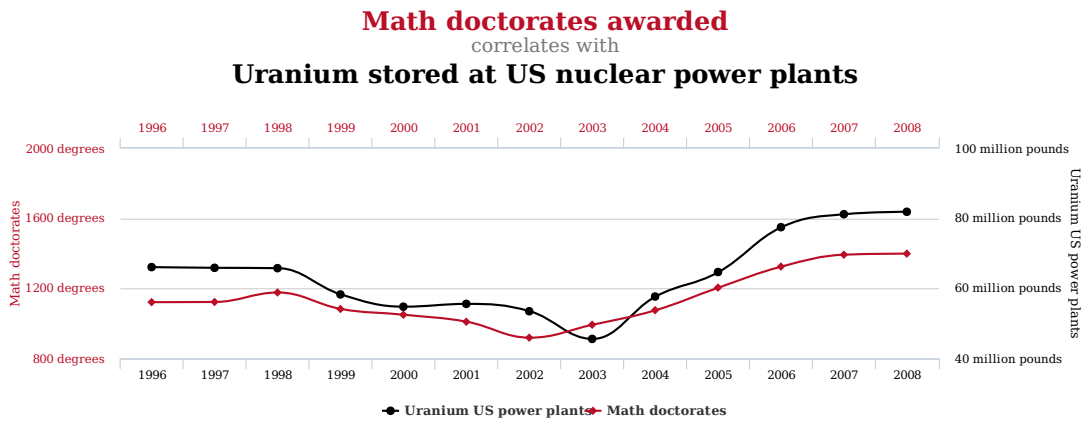
בדיון על תוחלת של סכום משתנים מקריים ראינו כי ישנה תכונת לינאריות ביחס למשתנים המקריים בלי קשר לתלות ביניהם. טענה דומה לא תהיה נכונה, באופן כללי, עבור השונות. יחד עם זאת, בתנאים מסויימים כן נוכל לטעון שהשונות היא לינאריות ביחס לסכום משתנים מקריים - כאשר המשתנים בלתי מתואמים. ז"א, השונות של סכום משתנים מקריים תהיה לינארית אם ורק אם המשתנים המקריים בלתי-מתואמים כפי שהמשפט הבא מדגים.

משפט 4.19 (שונות של סכום מ"מ) עבור משתנים מקריים $(X_i)_{i=1}^n$ מתקיים

$$\begin{aligned} V\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n V(X_i) + \sum_{i \neq j} Cov(X_i, X_j) = \\ &= \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} Cov(X_i, X_j). \end{aligned}$$

לכן, אם המשתנים בלתי מתואמים, אז השונות המשותפת שווה ל-0, ונקבל ששונות סכום משתנים מקריים שווה לסכום השונותיות,

¹⁴מקור הגרפים - tylervigen.com. ראו את האתר למקור המידע המתאים לסדרות הנתונים השונות.



איור 11: מתאם גבוה בין נתונים חסרי כל קשר. מקדם המתאם בכל הדוגמאות הנ"ל גבוה מ-0.95, ובחלקם כמעט מגיע ל-0.99.

$V(\sum_{i=1}^n X_i) = \sum_{i=1}^n V(X_i)$. ההוכחה עבור צמד משתנים מקריים היא ישירה וכך גם ההכללה לסכום כללי.

$$\begin{aligned} V(X \pm Y) &= E[(X \pm Y)^2] - (E[X \pm Y])^2 \\ &= E[X^2 \pm 2XY + Y^2] - [(E[X])^2 + (E[Y])^2 \pm 2E[X]E[Y]] \\ &= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 \pm 2(E[XY] - E[X]E[Y]) \\ &= V(X) + V(Y) \pm 2Cov(X, Y). \end{aligned}$$

שימו לב, משתנים מקריים בלתי תלויים הם גם בלתי מתואמים, בפרט עבור משתנים בלתי תלויים מתקיים $V(\sum_{i=1}^n X_i) = \sum_{i=1}^n V(X_i)$.

תרגיל 4.20 בכד כדורים אדומים שחורים ולבנים (4 מכל צבע). מוציאים מהכד 4 כדורים עם החזרה. יהיו X, Y מספר הכדורים האדומים, השחורים והלבנים שהוצאו, בהתאמה.

1. כיצד מ"מ אלו מתפלגים?
2. כיצד מתפלג $X + Y$? מהי שונותו?
3. מהו מקדם המתאם בין X ל- Y ?
4. מהו מקדם המתאם בין $X + Y$ לבין Z ? (לחשוב לפני שמחשבים!)

פתרון:

1. שלושת המשתנים המקריים הם שווי התפלגות. ההסתברות לכל הוצאה היא $\frac{1}{3}$ ולכן כל אחד מתפלג $Bin(4, \frac{1}{3})$. לכן $V(X) = V(Y) = V(Z) = 4 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{8}{9}$.
2. $X + Y$ הוא מספר האדומים והשחורים שמוציאים. מאחר וההסתברות להוציא אדום או שחור היא $\frac{2}{3}$ הרי ש- $X + Y \sim Bin(4, \frac{2}{3})$. לכן $Var(X + Y) = 4 \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{8}{9}$.
3. נשתמש בנוסחת שונות הסכום: $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$ ונקבל $Cov(X, Y) = -\frac{4}{9}$. לפיכך: $\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{-4/9}{8/9} = -0.5$ וזהו ערך שלילי, כצפוי.
4. מתקיים $X + Y + Z = 4$ לכן בין $X + Y$ לבין Z יש קשר לינארי מושלם שכן $Z = 4 - (X + Y)$. על כן $\rho(X + Y, Z) = -1$.

תרגיל 4.21 בבית חרושת פועלות שלוש מכונות בלתי תלויות במקביל. התפוקה הממוצעת ליום של כל מכונה היא 50 יחידות עם סטיית תקן של 5 יחידות. אלו, בן, ובר התווכחו ביניהם לגבי התפוקה הכוללת הממוצעת בבית החרושת וסטיית התקן שלה. חלק טענו שהערך הממוצע הוא 150 וסטיית התקן היא 15, בעוד האחרים טענו שהערך הממוצע הוא 150 וסטיית התקן היא $5\sqrt{3}$. קבעו מי צודק.

פתרון: לכל מכונה $i = 1, 2, 3$ נגדיר משתנה מקרי X_i שייסמן את התפוקה שלה. מבחינת התפוקה הכוללת הממוצעת במפעל, החישוב הוא קל - $E[X_1 + X_2 + X_3] = 3 \cdot 50 = 150$. אבל חשוב לשים לב לתכונה הבאה - $E[3X_1] = E[X_1 + X_2 + X_3] = 150$ וזאת אנחנו יודעים מלינאריות, אבל $3X_1 \neq X_1 + X_2 + X_3$ באופן כללי. המכונות הן בלתי תלויות, לכן התפוקה של מכונה אחת יכולה להיות שונה מתפוקה של מכונה אחרת. להלן שני חישובים לבחון מי צדק.

$$\begin{aligned} V(X_1 + X_2 + X_3) &= V(3X_1) = 9V(X_1) = 9 \cdot 5^2 = 225, \\ V(X_1 + X_2 + X_3) &= V(X_1) + V(X_2) + V(X_3) = 3V(X_1) = 3 \cdot 5^2 = 75, \end{aligned}$$

השורה הראשונה מציג את החישוב השגוי תחת ההנחה הלא נכונה כי $X_1 + X_2 + X_3 = 3X_1$, והשורה השנייה מציגה את החישוב הנכון. לכן התפוקה הכוללת הממוצעת היא 150 וסטיית התקן היא $5\sqrt{3}$.

תרגיל 4.22 לבן יש סכום כסף אותו הוא רוצה להשקיע. הוא צריך לבחור בין אפיק השקעות א' שמפיק רווח הנתון על ידי מ"מ X לבין אפיק ב' שמפיק רווח הנתון על ידי מ"מ Y (בשני המקרים מדובר ברווח עבור השקעת כל הכסף). בנוסף, הוא יכול לפצל את ההשקעה כאשר $\alpha \in [0, 1]$ מהכסף מושקע באפיק הראשון ו- $1-\alpha$ לשני, במקרה כזה הרווח יהיה $Z(\alpha) = \alpha X + (1-\alpha)Y$. הניחו כי $0 < E(X) = E(Y)$ ו- $0 < V(X) < V(Y)$. מצאו כיצד בן צריך להשקיע את כספו אם הוא מעוניין למקסם את תוחלת הרווח שלו ובזמנית למזער את הסיכון? (רמת הסיכון נמדדת על ידי השונות).

פתרון: ליניאריות התוחלת קובעת שלכל $\alpha \in [0, 1]$ נקבל

$$\begin{aligned} E(Z(\alpha)) &= E(\alpha X + (1-\alpha)Y) = \\ &= \alpha E(X) + (1-\alpha)E(Y) = \\ &= \alpha E(X) + (1-\alpha)E(X) = \\ &= E(X) = E(Y) \end{aligned}$$

לכן תוחלת הרווח קבועה, בלי קשר לאופן ההשקעה. נבחן את השונות.

$$\begin{aligned} V(Z(\alpha)) &= V(\alpha X + (1-\alpha)Y) = \\ &= V(\alpha X) + V((1-\alpha)Y) + 2Cov(\alpha X, (1-\alpha)Y) = \\ &= \alpha^2 V(X) + (1-\alpha)^2 V(Y) + 2\alpha(1-\alpha)Cov(X, Y) = \\ &= \alpha^2 (V(X) + V(Y) - 2Cov(X, Y)) + 2\alpha(Cov(X, Y) - V(Y)) + V(Y) = \\ &= \alpha^2 V(X - Y) + 2\alpha(Cov(X, Y) - V(Y)) + V(Y) \end{aligned}$$

וקיבלנו פולינום ב- α . נשים לב ש- $X \neq Y$ ולכן $V(X - Y) > 0$. זאת אומרת, יש לנו פרבולה עם מקדם חיובי, ולכן יש לה נקודת מינימום. נגזור ונשווה לאפס ונקבל

$$\alpha_{min} = \frac{V(Y) - Cov(X, Y)}{V(X - Y)} \quad (3)$$

נשים לב ש- $\alpha_{min} \in [0, 1]$ אם ורק אם

$$1. V(Y) \geq Cov(X, Y)$$

$$2. V(Y) - Cov(X, Y) \leq V(X - Y)$$

$$\begin{aligned} V(Y) - Cov(X, Y) &\leq V(X) + V(Y) - 2Cov(X, Y) \\ Cov(X, Y) &\leq V(X) \end{aligned}$$

שקלול צמד התנאים נותן לנו את התנאי $Cov(X, Y) \leq V(X)$ (התנאי הראשון מתקיים כתוצאה מהשני כי נתון $V(X) < V(Y)$). לסיכום, ממשוואה (3) נקבל שני תנאים:

- אם $V(X) < Cov(X, Y)$, אז $\alpha_{min} \notin [0, 1]$ ובפרט $\alpha_{min} > 1$. מאחר ומדובר בפרבולה, המינימום האפשרי יתקבל כאשר $\alpha = 1$, על כן צריך להשקיע את כל הכסף באפיק א'.

- אם $V(X) \geq Cov(X, Y)$, אז יש לפזר את ההשקעה על פי משוואה (3).

תרגיל 4.23 בקערה יש 10 אגסים, 10 תפוחים ו-10 בננות. מוצאים באקראי וללא החזרה 10 פירות. יהיו Y, X מספר התפוחים והבננות שהוצאו מהכד בהתאמה. חשבו את $\rho(X, Y)$.

פתרון: נגדיר את Z להיות מספר האגסים שהוצאו ונשים לב ש- $X + Y = 10 - Z$, לכן $V(X + Y) = V(10 - Z) = V(Z)$. בנוסף, משיקולי סימטריה נובע ש- $V(X) = V(Y) = V(Z)$. נשתמש בנוסחה לשונות של סכום מ"מ,

$$\begin{aligned} V(Z) &= V(X + Y) = \\ &= V(X) + V(Y) + 2Cov(X, Y) = \\ &= 2V(X) + 2Cov(X, Y), \\ Cov(X, Y) &= \frac{V(Z) - 2V(X)}{2} = \frac{V(X) - 2V(X)}{2} = -\frac{V(X)}{2}. \end{aligned}$$

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{-\frac{V(X)}{2}}{\sqrt{V(X)^2}} = -\frac{1}{2}.$$

נחשב את מקדם המתאם לפי הגדרתו:

תרגיל 4.24 ישנם 100 כדורים בכד ועליהם המספרים 1 עד 100 כאשר על כל כדור מספר שונה. מוציאים בזה אחר זה ללא החזרה 10 כדורים. יהי X_i הערך המופיע על הכדור ה- i שהוצא. חשבו את $Cov(X_i, X_j)$.

פתרון: נניח שממשיכים להוציא את כלל הכדורים ונגדיר משתנים מקריים $X_i, i = 1, \dots, 100$ כאשר X_i הוא הערך המופיע על הכדור ה- i שהוצא. אנחנו יודעים כי הסכום של כלל המ"מ הללו הוא קבוע ולכן השונות של הסכום שווה ל-0,

$$V\left(\sum_{i=1}^{100} X_i\right) = V(\text{Const}) = 0.$$

בנוסף, נזכור שניתן לבצע הקבלה בין הוצאות מכד ללא החזרה עם סידור עצמים בשורה, כפי שראינו בעבר בבעיות קומבינטוריות, לכן ישנה סימטריה מלאה בין המ"מ הללו. זאת אומרת, כלל המשתנים בעלי אותה ההתפלגות ובנוסף

$$Cov(X_i, X_j) = Cov(X_k, X_l),$$

לכל $1 \leq i, j, k, l \leq 100$ השונים זה מזה. אם $i = j$ אזי $Cov(X_i, X_j) = V(X_i)$. מאחר ו- $X_1 \sim X_i \sim U[1, 100]$ נוכל להיעזר בנוסחה לשונות של מ"מ בעל התפלגות אחידה (אם $X \sim U[a, b]$ אז $V(X) = \frac{(b-a+1)^2-1}{12}$) ו-

$$\begin{aligned} V(X_i) &= \frac{(b-a+1)^2-1}{12} = \\ &= \frac{100^2-1}{12} = \frac{9999}{12}. \end{aligned}$$

בנוסף,

$$\begin{aligned} 0 &= V\left(\sum_{i=1}^{100} X_i\right) = \\ &= \sum_{i=1}^{100} V(X_i) + 2 \sum_{i=1}^{100} \sum_{1 \leq j \neq i}^{100} Cov(X_i, X_j) = \\ &= 100V(X_1) + 100 \cdot 99Cov(X_1, X_2), \\ Cov(X_i, X_j) &= -\frac{100V(X_1)}{100 \cdot 99} = \\ &= -\frac{\frac{9999}{12}}{99} = -\frac{101}{12}, \end{aligned}$$

לכל $1 \leq i \neq j \leq 100$ שלמים.

5 משתנים מקריים רציפים

בפרקים הקודמים דנו במשתנים מקריים בדידים, כלומר משתנים המקבלים מספר סופי, או לכל היותר בן מניה, של ערכים. בפרק זה נדון במשתנים המקבלים רצף של ערכים. נפתח בדוגמא.

דוגמה 5.1 נתון מקל באורך מטר 1. שוברים את המקל לשני חלקים בנקודה שנבחרת באקראי. יהא X משתנה מקרי המתאר את אורך חלק המקל השמאלי במטרים. אילו ערכים יכול X לקבל? מה ההסתברות שאורך המקל השמאלי הוא לכל היותר חצי? מה ההסתברות ש- X הוא בין 0.5 ל-0.75? מה ההסתברות שאורך המקל הוא בדיוק 0.5? התשובה לשאלות הללו איננה פשוטה. באופן כללי, כאשר נאמר שהבחירה היא "אקראית" הכוונה לבחירה אחידה במובן זה שההסתברות לבחור נקודה בקטע מסויים פרופורציונית לאורך הקטע, לכן ככל שאורך הקטע קטן יותר כך ההסתברות לבחור נקודה בתוכו קטנה. בדוגמא הזו, X מקבל ערכים בין 0 ל-1 ו- $\Pr(X \leq 0.5) = 0.5$ שכן הקטע הנבחר הוא באורך חצי. באופן דומה, $\Pr(\frac{1}{2} \leq X \leq \frac{3}{4}) = 0.25$ כי אורך הקטע הנבחר הוא רבע. מצד שני, $\Pr(X = \frac{1}{2}) = 0$ כי אורך הקטע הנבחר הוא 0! אותה תוצאה אפסית תתקבל עבור כל ערך ספציפי, שכן ערך ספציפי מייצג קטע באורך 0.

מן הדוגמא אנו למדים שלא נוכל לתאר התפלגות של משתנה רציף בצורה דומה להתפלגות של משתנה בדיד. עבור משתנה בדיד יכולנו לקבוע במדויק את ההסתברות לקבל כל ערך ואז לסכום בהתאם את הערכים רצויים. במשתנים רציפים, ההסתברות לערך ספציפי היא אפס, לכן אנו נדרשים למצוא דרך אחרת לתאר את המשתנים הרציפים ולשם כך נעזר בפונקציה מוכרת - פונקציית ההתפלגות המצטברת.

5.1 פונקציית ההתפלגות מצטברת ופונקציית הצפיפות

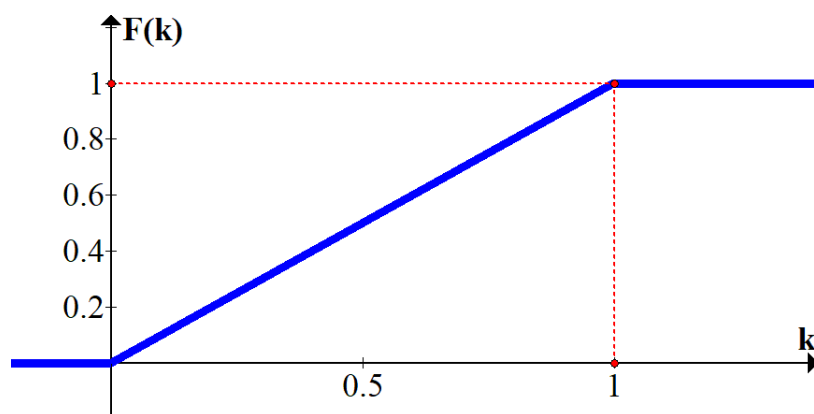
5.1.1 פונקציית ההתפלגות המצטברת

פונקציית ההתפלגות המצטברת של משתנה רציף מוגדרת באופן דומה לזאת של משתנה בדיד. **פונקציית ההתפלגות המצטברת של F_X של משתנה רציף X** היא פונקציה ממשיית מ- $\mathbb{R}$ לקטע $[0, 1]$ כך ש- $F_X(k) = \Pr(X \leq k)$. זו הגדרה זהה לזאת של משתנה הבדיד. למעשה, פונקציית ההתפלגות מצטברת יכולה לתאר באופן מדויק כל משתנה - בדיד, רציף, גם בדיד וגם רציף וכן הלאה. במקרה הבדיד היה לנו יותר קל לתאר את המשתנה המקרי בעזרת ההתפלגות P_X , אבל יכולנו באותה המידה להשתמש רק ב- F_X .

התכונות של פונקציית ההתפלגות המצטברת של משתנה רציף דומות לאלו של משתנה בדיד עם שינוי אחד קטן, אך מהותי. נתחיל בתכונות הדומות. הפה"מ היא **אי-שלילית** מהגדרתה, היא **מקיימת את הגבולות** $\lim_{k \rightarrow -\infty} F_X(k) = 0$, $\lim_{k \rightarrow \infty} F_X(k) = 1$ ובנוסף היא **מונוטונית לא-יורדת**. בנוסף, במקרה הרציף ובניגוד לבדיד, הפה"מ היא **רציפה ואף גזירה כמעט בכל מקום** (ייתכנו מספר קטן של נקודות בהן היא לא גזירה). במקרה הבדיד הפה"מ לא הייתה רציפה, אלא פונקציית מדרגות אשר קופצת בערכים אשר נמצאים בתומך של X , אבל זה לא המצב במקרה הרציף. במקרה הרציף, הפה"מ היא רציפה.

נרשום אלגברית ונשרטט את פונקציית ההתפלגות המצטברת של דוגמת המקל. ראו איור 12.

$$F_X(k) = \Pr(X \leq k) = \begin{cases} 0, & k < 0, \\ \frac{k}{1}, & 0 \leq k < 1, \\ 1, & 1 \leq k. \end{cases}$$



איור 12: שרטוט פונקציית התפלגות מצטברת עבור דוגמא 5.1

ניתן לראות כי כלל התכונות מתקיימות: הפונקציה נעה מ-0 ל-1 באופן רציף, מונוטונית עולה חלש, וגזירה כמעט בכל מקום למעט בנקודות $k = 0$ ו- $k = 1$.

5.1.2 פונקציית הצפיפות

בשלב זה לרוב עולה השאלה - האם יש פונקציית התפלגות למשתנה רציף, בדומה לפונקציית ההתפלגות של משתנה בדיד? התשובה, לשמחתנו, היא כן. ישנה פונקציה נוספת המתארת משתנה מקרי רציף, ויש לה דמיון מסויים לפונקציית ההתפלגות של המשתנה הבדיד. הפונקציה הנוספת נקראת - פונקציית הצפיפות.

פונקציית הצפיפות מגדירה משתנה רציף כשם שפונקציית התפלגות מגדירה משתנה בדיד, ולכן נתחיל מההגדרות. משתנה מקרי X הוא רציף אם קיימת פונקציה f_X ממשיית ואי-שלילית מ- $\mathbb{R}$ ל- $[0, \infty)$ כך ש- $\Pr(a \leq X \leq b) = \int_a^b f_X(k) dk$ לכל $a < b$. הפונקציה f_X נקראת **פונקציית הצפיפות של משתנה מקרי רציף** X והיא אנלוגית לפונקציית ההתפלגות של משתנה מקרי בדיד. זאת אומרת, למשתנה מקרי בדיד יש פונקציית התפלגות ולרציף - פונקציית צפיפות. שימו לב לתכונה חשובה של פונקציית הצפיפות - ההסתברות לקבל ערך בקטע $[a, b]$ שווה לאינטגרל על הפונקציה בקטע. לכן, ברגע שנמצא את פונקציית הצפיפות של משתנה כלשהו, מיד נוכל לחשב את ההסתברות לקבל ערך בכל קטע נתון.

מהתיאור לעיל ניתן לזהות תכונה חשובה אשר יוצרת הקבלה בין פונקציית הצפיפות ופונקציית ההתפלגות. עבור משתנה בדיד,

$$F_X(b) - F_X(a) = \Pr(a < X \leq b) = \sum_{k>a}^b P_X(k),$$

בעוד עבור משתנה רציף מתקבל השיויון הבא

$$F_X(b) - F_X(a) = \Pr(a < X \leq b) = \int_a^b f_X(k) dk.$$

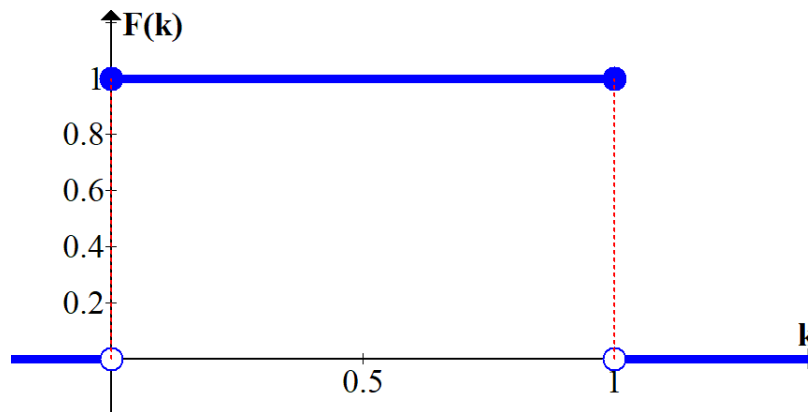
זאת אומרת, בשני המקרים ניתן לסכום על פונקציית ההתפלגות/הצפיפות בכדי לקבל את פונקציית ההתפלגות המצטברת. במקרה הבדיד מדובר בסכום ובמקרה הרציף מדובר באינטגרל.

הערה 5.2 אם X הוא משתנה מקרי רציף, אזי ההסתברות לקבל ערך ספציפי היא אפסית, ולכן

$$\Pr(a < X \leq b) = \Pr(a \leq X \leq b) = \Pr(a \leq X < b).$$

נמשיך את דוגמא 5.1 ונתאר את פונקציית הצפיפות של המקל הנשבר באקראי, בצורה אלגברית ובשרטוט (ראו איור 13). בניגוד לפונקציית ההתפלגות המצטברת, אשר הנוסחה שלה הייתה מיידית (פרופורציונאלית לאורך הקטע), במקרה של פונקציית הצפיפות אנחנו נדרשים למצוא פונקציה אשר האינטגרל שלה על קטע נתון מניב את ההסתברות המתאימה. זו לא נראית משימה פשוטה באופן כללי. לכן נרצה למצוא דרך לקשור בין הפה"מ לצפיפות, ומסתבר שהדרך הזו קלה ביותר.

$$f_X(k) = \begin{cases} 1, & 0 \leq k \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$



איור 13: שרטוט פונקציית הצפיפות עבור דוגמא 5.1

5.1.3 הקשר בין הפה"מ לפונקציית הצפיפות

אנו מעוניינים לאתר דרך קלה לחישוב פונקציית הצפיפות מתוך פונקציית ההתפלגות המצטברת. הדרך ההפוכה היא קלה, משום ש-
 $F_X(b) - F_X(a) = \int_a^b f_X(k) dk$ מצד שני, אם נחזור רגע למשפט היסודי של החדו"א או נוסחת ניוטון-לייבניץ, אז נזכור שיש קשר הדוק בין האינטגרל $\int_a^b f_X(k) dk$ לבין הפונקציה f_X עליה מבצעים את האינטגרציה, וקשר זה קובע כי

$$f_X(k) = F'_X(k),$$

$$F_X(k) = \int_{-\infty}^k f_X(t) dt.$$

במילים, פונקציה הצפיפות היא הנגזרת של פונקציה ההתפלגות המצטברת ופונקציית ההתפלגות המצטברת היא האינטגרל על פונקציית הצפיפות (בגבולות המתאימים). זהו הקשר שמגדיר את היחס בין הצפיפות לפה"מ של משתנה רציף.

הקשר בין הפה"מ לצפיפות גם נותן מעט אינטואיציה על פונקציית הצפיפות. פונקציית הצפיפות היא הנגזרת של הפה"מ, ז"א היא השינוי בפונקציית ההתפלגות המצטברת. למעשה, הצפיפות היא "מעין" ריכוז של הסתברות בנקודה, או השינוי בהסתברות באותה נקודה. כמובן, זו רק האינטואיציה וההסתברות לכל נקודה במשתנה הרציף היא עדיין אפס.

הערה 5.3 אם פונקציית ההתפלגות המצטברת נתונה, אזי מציאת הצפיפות איננה בעייתית - נגזרת של הפה"מ. מצד שני, אם נתונה פונקציית הצפיפות ורוצים לחשב את פונקציית ההתפלגות המצטברת המשימה מעט יותר מסובכת, שכן צריך לבצע אינטגרציה. בחלק זה ישנם סטודנטים הנוטים להתבלבל. הערך $F_X(k)$ מתקבל על ידי האינטגרל $\int_{-\infty}^k f_X(t) dt$ כאשר הגבולות נעים ממינוס אינסוף ועד k . מרבית הסטודנטים מתבלבלים בגבול התחתון - כאשר פונקציית הצפיפות מוגדרת במקטעים שונים בצורה שונה, אז יש נטייה להציב בתור הגבול

התחתון את תחילת המקטע הרלוונטי. זו טעות! הגבול התחתון הוא תמיד $-\infty$ ובפונקציה המוגדרת במקטעים שונים, צריך לחלק את האינטגרל לעיל למספר אינטגרלים שונים בהתאם לתיאור פונקצית הצפיפות כך שכל הקטע $(-\infty, k]$ נלקח בחשבון. ראו תרגיל 5.6 בהמשך.

תרגיל 5.4 אוטובוס מגיע לתחנה כל 30 דקות. בר מגיעה לתחנה באקראי, מבלי לתכנן מראש ומבלי לבדוק את לוחות הזמנים של האוטובוס. נסמן ב- T את זמן ההמתנה המדויק לאוטובוס, בדקות. רשמו את פונקצית ההתפלגות המצטברת ואת פונקצית הצפיפות של T . חשבו את ההסתברויות $\Pr(20 \leq T \leq 40)$ ו- $\Pr(T > 23)$. ענו שוב על השאלה עבור אוטובוס המגיע כל 40 דקות.

פתרון: במקרה הנוכחי, בדומה לדוגמא עם המקל, זמן ההמתנה הוא אחיד משום שמועד ההגעה של בר הוא אחיד בזמן - אין העדפה למועד אחד ע"פ מועד אחר וההסתברות פרפורציונאלית לאורך מקטע הזמן. לכן,

$$F_T(t) = \begin{cases} 0, & t < 0, \\ \frac{t}{30}, & 0 \leq t \leq 30, \\ 1, & 30 \leq t, \end{cases}$$

$$f_T(t) = F'_T(t) = \begin{cases} \frac{1}{30}, & 0 \leq t \leq 30, \\ 0, & \text{otherwise,} \end{cases}$$

$$\Pr(20 \leq T \leq 40) = F_T(40) - F_T(20) = 1 - \frac{20}{30} = \frac{1}{3},$$

$$\Pr(T > 23) = 1 - \Pr(T \leq 23) = 1 - F_T(23) = 1 - \frac{23}{30} = \frac{7}{30}.$$

עבור אוטובוס המגיע כל 40 דקות נקבל

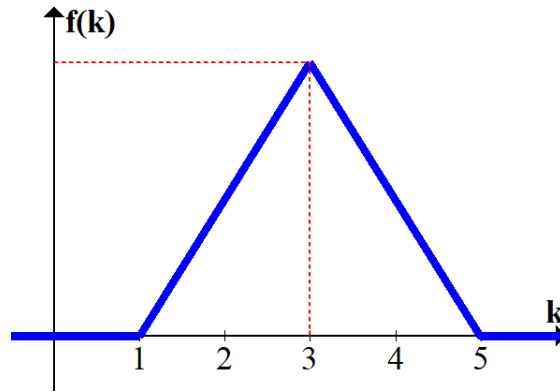
$$F_T(t) = \begin{cases} 0, & t < 0, \\ \frac{t}{40}, & 0 \leq t \leq 40, \\ 1, & 40 \leq t, \end{cases}$$

$$f_T(t) = F'_T(t) = \begin{cases} \frac{1}{40}, & 0 \leq t \leq 40, \\ 0, & \text{otherwise,} \end{cases}$$

$$\Pr(20 \leq T \leq 40) = F_T(40) - F_T(20) = 1 - \frac{20}{40} = \frac{1}{2},$$

$$\Pr(T > 23) = 1 - \Pr(T \leq 23) = 1 - F_T(23) = 1 - \frac{23}{40} = \frac{17}{40}.$$

תרגיל 5.5 מכונה במפעל ברגים אמורה לייצר ברגים בקוטר 3 מ"מ. בפועל, קוטר הברגים אינו מדויק וההתפלגות שלו נתונה על ידי איור 14. רשמו במדויק את פונקצית הצפיפות. מה ההסתברות שקוטר בורג יעלה על 3 מ"מ? על 4 מ"מ?



איור 14: צפיפות קוטר הברגים, תרגיל 5.5

פתרון: בכדי לענות על השאלה נצטרך למצוא את פונקציית הצפיפות בצורה מדויקת. נשתמש תחילה בתנאי הנרמול. אנחנו יודעים כי השטח שבין הגרף לציר x שווה ל-1 כי כלל השטח מייצג את כל ההסתברות, $\int_1^3 f_X(k) dk = \Pr(1 \leq X \leq 5) = 1$. נשתמש בנוסחה לשטח משולש (שטח משולש שווה למכפלת הבסיס בחצי הגובה) ונקבל שגובה המשולש h מקיים $\frac{h \cdot 4}{2} = 1$, ולכן $h = 0.5$. עם כן, יש לנו את הנקודות $(1, 0)$, $(5, 0)$, ו- $(3, 0.5)$ ונוכל למצוא את משוואות הקווים בהתאם (בצעו רענון קצר למשוואות קו ישר העובר דרך 2 נקודות מהתיכון). לכן הצפיפות נתונה על ידי

$$f_X(k) = \begin{cases} 0, & k < 1, \\ \frac{1}{4}(k-1), & 1 \leq k \leq 3, \\ -\frac{1}{4}(k-5), & 3 \leq k \leq 5, \\ 0, & 5 < k. \end{cases}$$

בעזרת הצפיפות נוכל לחשב ישירות את ההסתברויות המבוקשות.

$$\Pr(X > 3) = \int_3^5 -\frac{1}{4}(k-5) dk = -\frac{1}{4} \left[\frac{k^2}{2} - 5k \right]_3^5 = \frac{1}{2},$$

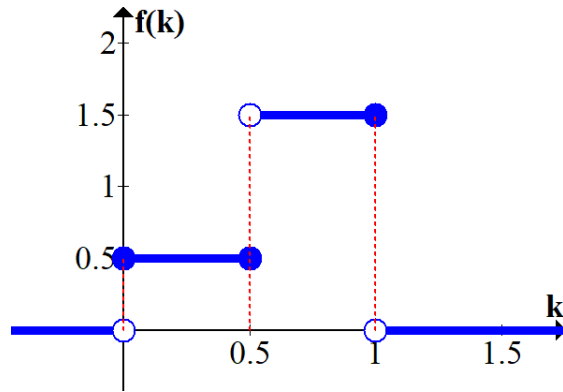
$$\Pr(X > 4) = \int_4^5 -\frac{1}{4}(k-5) dk = -\frac{1}{4} \left[\frac{k^2}{2} - 5k \right]_4^5 = \frac{1}{8}.$$

תרגיל 5.6 משתנה מקרי רציף X בעל פונקציית צפיפות

$$f_X(k) = \begin{cases} \frac{1}{2}, & 0 \leq k \leq \frac{1}{2}, \\ \frac{3}{2}, & \frac{1}{2} < k \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

חשבו את ההסתברויות הבאות: $\Pr(0 \leq X \leq \frac{1}{4})$, $\Pr(\frac{1}{2} \leq X \leq \frac{3}{4})$, $\Pr(\frac{1}{2} \leq X \leq 1)$, $\Pr(0 \leq X \leq \frac{1}{2})$, $\Pr(0 \leq x \leq 1)$. מצאו את פונקציית ההתפלגות המצטברת F_X .

פתרון: חישוב ההסתברויות הוא מיידי מנוסחת האינטגרציה וחישוב שטחים. רק צריך לאתר את נקודות הקצה בכל חישוב, ולמצוא את השטח שכלוא בין הגרף לציר x . שרטוט הפונקציה יסייע בנייתו, ראו איור 15.



איור 15: צפיפות משתנה רציף X , תרגיל 5.6

נתחיל לפתור את התרגילים לפי הסדר. הראשון הוא מיידי שכן ההסתברויות נסכמות ל-1. נבדוק זאת,

$$\Pr(0 \leq X \leq 1) = \left(\frac{1}{2} - 0\right) \cdot \frac{1}{2} + \left(1 - \frac{1}{2}\right) \cdot \frac{3}{2} = 1,$$

כנדרש. נמשיך באותו האופן לכל ההסתברויות המבוקשות.

$$\Pr\left(0 \leq X \leq \frac{1}{2}\right) = \left(\frac{1}{2} - 0\right) \cdot \frac{1}{2} = \frac{1}{4},$$

$$\Pr\left(0 \leq X \leq \frac{1}{4}\right) = \left(\frac{1}{4} - 0\right) \cdot \frac{1}{2} = \frac{1}{8},$$

$$\Pr\left(\frac{1}{2} \leq X \leq 1\right) = \left(1 - \frac{1}{2}\right) \cdot \frac{3}{2} = \frac{3}{4},$$

$$\Pr\left(\frac{1}{2} \leq X \leq \frac{3}{4}\right) = \left(\frac{3}{4} - \frac{1}{2}\right) \cdot \frac{3}{2} = \frac{3}{8}.$$

כעת נעבור לחישוב פונקציית ההתפלגות המצטברת. ניזכר בהערה קודמת לגבי חישוב פונקציית התפלגות מצטברת מתוך הצפיפות בעזרת $F_X(k) = \int_{-\infty}^k f_X(t) dt$, בדגש על הגבולות. נתחיל דווקא מן ההגדרה של פונקציית ההתפלגות המצטברת ומקרי הקיצון. $F_X(k) = 0$ לכן עבור $k < 0$, הערך של הפונקציה הוא 0 כי X מקבל ערכים בקטע $[0, 1]$. באופן דומה, עבור $k \geq 1$ נקבל כי $F_X(k) = 1$ מאותם השיקולים בדיוק. כעת נותר למצוא את הערכים המתאימים עבור k בקטע $[0, 0.5]$ ובקטע $[0.5, 1]$.

עבור $k \in [0, 0.5]$, נפרק את האינטגרל לשני חלקים: אינטגרל על הקטע $(-\infty, 0)$ ועל הקטע $[0, k]$. הסיבה לפירוק הזה היא הגדרת הצפיפות. הערך של פונקציית הצפיפות שונה בין צמד הקטעים הללו, לכן לא ניתן לבצע אינטגרל על ביטוי אלגברי יחיד, אלא צריך לבצע את החישוב בשלבים עבור כל קטע בנפרד ולסכום. נציב באינטגרל את הביטוי המתאים בכל אחד מהקטעים.

$$\begin{aligned} F_X(k) &= \int_{-\infty}^k f_X(t) dt = \int_{-\infty}^0 f_X(t) dt + \int_0^k f_X(t) dt \\ &= \int_{-\infty}^0 0 dt + \int_0^k \frac{1}{2} dt = \frac{1}{2}k. \end{aligned}$$

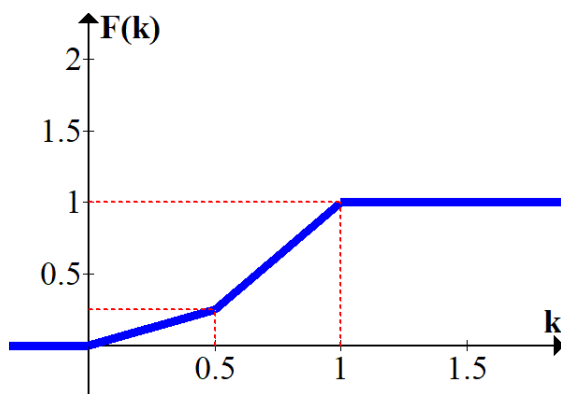
נעבור ל- $k \in [0.5, 1]$. במקרה זה נפרק את האינטגרל לשלושה חלקים: אינטגרל על הקטע $(-\infty, 0)$, אינטגרל על הקטע $[0, \frac{1}{2}]$, ואינטגרל

על הקטע $[\frac{1}{2}, k]$. נציב באינטגרל את הביטוי האלגברי המתאים לפונקציית הצפיפות בכל אחד מהקטעים ונחשב.

$$\begin{aligned} F_X(k) &= \int_{-\infty}^k f_X(t) dt = \int_{-\infty}^0 f_X(t) dt + \int_0^{0.5} f_X(t) dt + \int_{0.5}^k f_X(t) dt \\ &= \int_{-\infty}^0 0 dt + \int_0^{0.5} \frac{1}{2} dt + \int_{0.5}^k \frac{3}{2} dt = \frac{1}{2} \cdot \frac{1}{2} + \frac{3}{2} \left(k - \frac{1}{2} \right) = \frac{3}{2}k - \frac{1}{2}. \end{aligned}$$

נסכם את התוצאות שקיבלנו, ונשרטט על גרף, ראו איור 16.

$$F_X(k) = \begin{cases} 0, & k < 0, \\ \frac{k}{2}, & 0 \leq k \leq \frac{1}{2}, \\ \frac{3k}{2} - \frac{1}{2}, & \frac{1}{2} \leq k \leq 1, \\ 1, & 1 < k. \end{cases}$$



איור 16: פונקציית התפלגות מצטברת משתנה רציף X , תרגיל 5.6

קל לבדוק שלא ביצענו טעויות בחישוב. קל לראות שהפונקציית הצפיפה ומונוטוניות עולה חלש. בנוסף, גזירה של הפונקציית צריכה לייצר את הצפיפות ממנה התחלנו. בדקו שזה אכן המצב.

תרגיל 5.7 משתנה מקרי רציף X בעל פונקציית צפיפות

$$f_X(k) = \begin{cases} ck, & 0 \leq k \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

חשבו את c , ומצאו את פונקציית ההתפלגות המצטברת F_X .

פתרון: בכדי למצוא את הקבוע c נשתמש בתנאי הנרמול. ידוע כי $\int_{-\infty}^{\infty} f_X(k) dk = 1$

$$1 = \int_{-\infty}^{\infty} f_X(k) dk = \int_0^2 c k dk = c \left[\frac{k^2}{2} \right]_0^2 = 2c$$

לכן $c = \frac{1}{2}$ ו- $2c = 1$. נעבור לחישוב הפה"מ. נתחיל שוב ממקרי הקיצון. עבור $k < 0$, ההסתברות $\Pr(X \leq k) = 0$ כי המשתנה נע בין 0 ל-2. עבור $k \geq 2$, ההסתברות $\Pr(X \leq k) = 1$ מאותה הסיבה בדיוק. כעת נותר למצוא את הערך עבור $k \in (0, 2)$.

$$\begin{aligned} F_X(k) &= \int_{-\infty}^k f_X(t) dt = \int_{-\infty}^0 f_X(t) dt + \int_0^k f_X(t) dt \\ &= \int_{-\infty}^0 0 dt + \int_0^k \frac{1}{2} t dt = \left[\frac{t^2}{4} \right]_0^k = \frac{k^2}{4}. \end{aligned}$$

לסיכום, פונקציית ההתפלגות המצטברת היא

$$F_X(k) = \begin{cases} 0, & k < 0, \\ \frac{k^2}{4}, & 0 \leq k \leq 2, \\ 1, & 2 < k. \end{cases}$$

בדקו כי קיבלנו פונקציה רציפה, הנעה מ-0 ל-1 ומונוטונית עולה חלש, בנוסף לכך שגזירה מניבה את הצפיפות המקורית.

תרגיל 5.8 נתונה פונקציית הצפיפות הבאה:

$$f_X(x) = \begin{cases} -x & -1 \leq x \leq 0 \\ cx^2 & 0 < x \leq 1 \\ 0 & |x| > 1 \end{cases}$$

1. מצאו את ערכו של c .

2. חשבו את פונקציית ההתפלגות המצטברת של X .

3. חשבו את ההסתברויות:

(א) $\Pr(X \leq -0.5)$

(ב) $\Pr(X < -0.5)$

(ג) $\Pr(X \leq 0.5)$

(ד) $\Pr(-0.2 \leq X \leq 0.3)$

פתרון: נפתור כל סעיף בנפרד.

1. מתכונת הנרמול נקבל $1 = \int_{-\infty}^{\infty} f_X(x) dx = \int_{-1}^0 -x dx + \int_0^1 cx^2 dx = \frac{1}{2} + c \left(\frac{1}{3} - 0 \right)$ ולכן $c = 1.5$.

2.
$$F_X(t) = \int_{-\infty}^t f_X(x) dx = \begin{cases} 0 & t < -1 \\ \frac{1}{2} - \frac{t^2}{2} & -1 \leq t < 0 \\ \frac{1}{2} + \frac{t^3}{2} & 0 \leq t \leq 1 \\ 1 & t > 1 \end{cases}$$

3. נחשב את ההסתברויות באמצעות הצבה בפונקציית ההתפלגות המצטברת:

$$\Pr(X \leq -0.5) = F_X(-0.5) = 0.375 \quad (\text{א})$$

(ב) $\Pr(X < -0.5) = \Pr(X \leq -0.5) = 0.375$ שכן כזכור, נקודה אחת אינה משפיעה על תוצאת אינטגרל וההסתברות להיות שווה בדיוק ל-0.5 היא אפס.

$$\Pr(X \leq 0.5) = F_X(0.5) = 0.5625 \quad (\text{ג})$$

$$\Pr(-0.2 \leq X \leq 0.3) = F_X(0.3) - F_X(-0.2) = 0.0335 \quad (\text{ד})$$

תרגיל 5.9 אספקת הדלק למאגר של תחנת דלק מתבצעת אחת לשבוע. כמות הדלק (באלפי ליטרים) שתחנה זו מוכרת בשבוע היא משתנה מקרי המתפלג בהתאם לפונקציית הצפיפות:

$$f_X(x) = \begin{cases} c(1-x)^4 & x \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

1. חשבו את c .

2. מה צריכה להיות קיבולת המאגר כדי שההסתברות שהוא יתרוקן בשבוע תהיה קטנה מ-5%?

פתרון: נפתור כל סעיף בנפרד.

$$1. \text{ מתכונת הנרמול נקבל } \frac{c}{5} = -\frac{c}{5}(1-x)^5 \Big|_{x=0}^{x=1} = \int_0^1 c(1-x)^4 dx = 1 \text{ ולכן } c = 5.$$

2. נסמן את קיבולת המאגר ב- M . אנו דורשים שההסתברות שצריכת הדלק תהיה גדולה מהקיבולת תהיה קטנה מ-5%, כלומר $\Pr(X > M) \leq 0.05$. נחשב את ההסתברות המתוארת. $\Pr(X > M) = \int_M^1 f_X(x) dx = (1-M)^5 \leq 0.05$ ולכן $M \geq 0.4507$. המספר שקיבלנו הוא בדיוק האחוזון ה-95% של ההתפלגות, שכן ב-95% מהשבועות צריכת הדלק נמוכה ממספר זה.

תרגיל 5.10 תרביית חיידקים מפוזרת באופן אחיד על פני צלחת שרדיוסה 10 ס"מ. יהא R המרחק של חיידק אקראי ממרכז הצלחת.

1. מה ההסתברות שהמרחק הוא בדיוק 3 ס"מ?

2. מצאו את פונקציית ההתפלגות המצטברת של R .

3. מהי ההסתברות שהמרחק הוא 3 ס"מ לפחות אם ידוע שהוא 6 ס"מ לכל היותר?

4. מצאו את פונקציית הצפיפות של R .

פתרון:

1. מאחר וזהו משתנה מקרי רציף, ההסתברות שיקבל בדיוק ערך מסוים היא אפס.

2. מאחר והחיידקים מפולגים באופן אחיד על הצלחת, הסיכוי שחיידק מסוים נמצא במרחק קטן מ- r היא היחס בין השטח של האזורים

$$F_R(r) = \begin{cases} 0, & r < 0, \\ \frac{r^2}{10^2}, & 0 \leq r \leq 10, \text{ לכן } r \text{ ביחס לשטח הכולל.} \\ 1, & r > 10. \end{cases}$$

$$3. \Pr(R > 3 | R < 6) = \frac{\Pr(3 < R < 6)}{\Pr(R < 6)} = \frac{F_R(6) - F_R(3)}{F_R(6)} = 0.75.$$

$$4. f_R(r) = \frac{dF_R}{dr} = \begin{cases} \frac{r}{50}, & 0 \leq r \leq 10, \\ 0, & \text{otherwise.} \end{cases}$$

5.2 תוחלת ושונות

התוחלות והשונות של משתנה מקרי רציף מוגדרות בצורה אנלוגית למשתנה הבדיד. מבחינה מהותית, אין הבדל בין הגדלים. תוחלת היא עדיין מדד מרכזי (מדד אמצע) שמטרתו לאמוד את גודל המשתנה בצורה גסה, והשונות עדיין מוגדרת בתור ממוצע ריבוע הפיזור סביב התוחלת. למעשה, השונות מוגדרת באופן זהה לחלוטין במקרה הבדיד ובמקרה הרציף, שכן היא מבוססת על הגדרת התוחלת, וכל מה שנצטרך זה לדון בהגדרת התוחלת.

כפי שראינו בחלק הקודם, משתנה מקרי רציף מאופיין על ידי הצפיפות שלו או על ידי פונקציית ההתפלגות המצטברת שלו, כאשר הצפיפות של המשתנה הרציף אנלוגית להתפלגות של המשתנה הבדיד. לכן נשתמש בצפיפות להגדרת התוחלת, בדומה לשימוש בהתפלגות במקרה הבדיד. באופן מפורש, **התוחלת של משתנה מקרי רציף** X עם צפיפות f_X היא $E[X] = \int_{-\infty}^{\infty} k f_X(k) dk$. זאת אומרת, לוקחים את כל הנקודות שמקבל המשתנה, כופלים כל נקודה במשקל היחסי שלה (ע"פ הצפיפות), וסוכמים את כלל הגורמים (בעזרת אינטגרציה). השונות מוגדרת באותו האופן כמו משתנה בדיד. אם X הוא משתנה מקרי רציף, אזי השונות שלו היא $V(X) = E[(X - \mu_X)^2]$. כאשר $\mu_X = E[X]$. בדומה למקרה הבדיד, נוכל להשתמש בתכונות הלינאריות ולקבל $V(X) = E[X^2] - (\mu_X)^2$. כעת נשאלת השאלה - כיצד לחשב את $E[X^2]$? מאחר וזאת תוחלת, אז ניתן למצוא את הצפיפות של X^2 ולהשתמש בהגדרה המקורית של התוחלת למשתנה רציף. מצד שני ניתן להשתמש בנוסחה לתוחלת של פונקציה של משתנה מקרי, אשר גם היא אנלוגית למקרה הבדיד. בפרט, אם g היא פונקציה ממשית כלשהי ו- X משתנה מקרי רציף, אזי $E[g(X)] = \int_{-\infty}^{\infty} g(k) f_X(k) dk$. כדאי לציין כי כלל תכונות התוחלת והשונות שראינו עבור משתנה בדיד נשמרות גם במקרה הרציף (לינאריות התוחלת, היחס לקבועים, וכן הלאה). נמשיך עם מספר תרגילי חישוב.

תרגיל 5.11 מצאו את התוחלת והשונות של T המוגדר בתרגיל 5.4.

פתרון: נשתמש בהגדרה של תוחלת ושונות בהתאם לפונקציית הצפיפות שמצאנו.

$$\begin{aligned} E[T] &= \int_{-\infty}^{\infty} t f_T(t) dt = \int_0^{30} t \frac{1}{30} dt = \frac{1}{30} \left[\frac{t^2}{2} \right]_0^{30} = 15, \\ E[T^2] &= \int_{-\infty}^{\infty} t^2 f_T(t) dt = \int_0^{30} t^2 \frac{1}{30} dt = \frac{1}{30} \left[\frac{t^3}{3} \right]_0^{30} = 300, \\ V(T) &= E[T^2] - (\mu_T)^2 = 300 - 225 = 75. \end{aligned}$$

תרגיל 5.12 מצאו את התוחלת והשונות של X המוגדר בתרגיל 5.6.

פתרון: נשתמש בהגדרה של תוחלת ושונות בהתאם לפונקציית הצפיפות שמצאנו.

$$f_X(k) = \begin{cases} \frac{1}{2}, & 0 \leq k \leq \frac{1}{2}, \\ \frac{3}{2}, & \frac{1}{2} < k \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

$$\begin{aligned}
E[X] &= \int_{-\infty}^{\infty} k f_X(k) dk = \int_0^{\frac{1}{2}} k \frac{1}{2} dk + \int_{\frac{1}{2}}^1 k \frac{3}{2} dk = \frac{1}{2} \left[\frac{k^2}{2} \right]_0^{\frac{1}{2}} + \frac{3}{2} \left[\frac{k^2}{2} \right]_{\frac{1}{2}}^1 = \frac{5}{8}, \\
E[X^2] &= \int_{-\infty}^{\infty} k^2 f_X(k) dk = \int_0^{\frac{1}{2}} k^2 \frac{1}{2} dk + \int_{\frac{1}{2}}^1 k^2 \frac{3}{2} dk = \frac{1}{2} \left[\frac{k^3}{3} \right]_0^{\frac{1}{2}} + \frac{3}{2} \left[\frac{k^3}{3} \right]_{\frac{1}{2}}^1 = \frac{11}{24}, \\
V(X) &= E[X^2] - (\mu_X)^2 = \frac{11}{24} - \frac{25}{64} = \frac{13}{192}.
\end{aligned}$$

5.3 משפחות התפלגויות רציפות

אנו נעסוק בשלוש התפלגויות נפוצות במקרה הרציף: אחידה, מעריכית, ונורמאלית, כאשר הדיון על התפלגות אחידה החל כבר בפרק הקודם.

5.3.1 התפלגות אחידה רציפה

דוגמת המקל היא הדוגמא בסיסית ביותר למשתנה מקרי אחיד רציף. הרעיון בבסיס המשתנה האחיד היא שההסתברות לקבל ערך בקטע מסויים פורפוציונאלית לאורכו היחסי של אותו הקטע. במקרה של המקל, ישנה התפלגות אחידה על הקטע $(0, 1)$, ולכן הסיכוי לקבל ערך בכל תת קטע שווה לאורכו של תת הקטע הרלוונטי. לדוגמא, נסמן את המשתנה האחיד על הקטע ב- $X \sim U(0, 1)$ ונשים לב ש- $\Pr(0.2 \leq X \leq 0.56) = 0.56 - 0.2 = 0.36$.

באופן כללי, ניקח קטע סופי (a, b) ונגיד ש- X מתפלג אחיד בקטע זה, $X \sim U(a, b)$ אם ורק אם

$$f_X(k) = \begin{cases} \frac{1}{b-a}, & a \leq k \leq b, \\ 0, & \text{otherwise.} \end{cases}$$

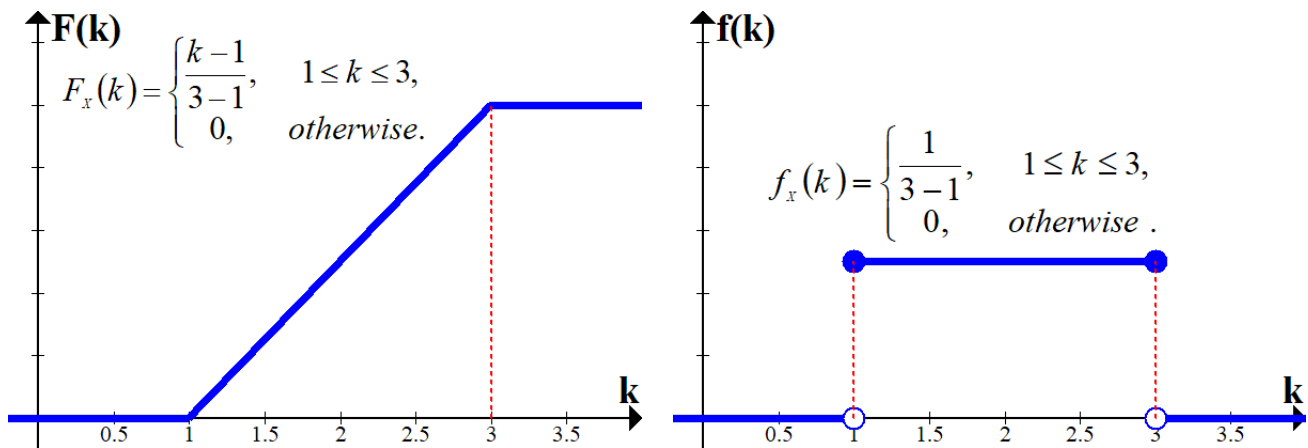
במילים, **משתנה מקרי אחיד רציף הוא משתנה בעלת צפיפות קבועה** על קטע כלשהו (a, b) , וביתר הנקודות הצפיפות שווה לאפס. בעזרת הצפיפות נוכל למצוא את פונקציית ההתפלגות המצטברת של משתנה אחיד (כפי שעשינו בתרגילים בפרק הקודם) ונמצא שפונקציית ההתפלגות המצטברת של משתנה אחיד רציף היא לינארית (קו ישר),

$$F_X(k) = \begin{cases} 0, & k < a, \\ \frac{k-a}{b-a}, & a \leq k \leq b, \\ 1, & b < k. \end{cases}$$

תוכלו לבדוק שכל תנאי הפה"מ מתקיימים: מונוטוניות עולה חלש, נעה מ-0 ל-1, הנגזרת שלה שווה לצפיפות (פונקציה קבועה בקטע הנתון), ורציפה. ראו איור 17.

חישובי התוחלת והשונות זהים לחישובים שעשינו קודם לכן. התוחלת והשונות של מ"מ אחיד רציף הן

$$\begin{aligned}
E[X] &= \int_a^b \frac{k}{b-a} dk = \frac{1}{b-a} \left[\frac{k^2}{2} \right]_a^b = \frac{b+a}{2}, \\
E[X^2] &= \int_a^b \frac{k^2}{b-a} dk = \frac{1}{b-a} \left[\frac{k^3}{3} \right]_a^b = \frac{(b-a)(b^2+ba+a^2)}{3(b-a)} = \frac{b^2+ba+a^2}{3}, \\
V(X) &= E[X^2] - (\mu_X)^2 = \frac{(b-a)^2}{12}.
\end{aligned}$$



איור 17: פונקציות התפלגות מצטברת וצפיפות של משתנה אחיד על הקטע (1, 3)

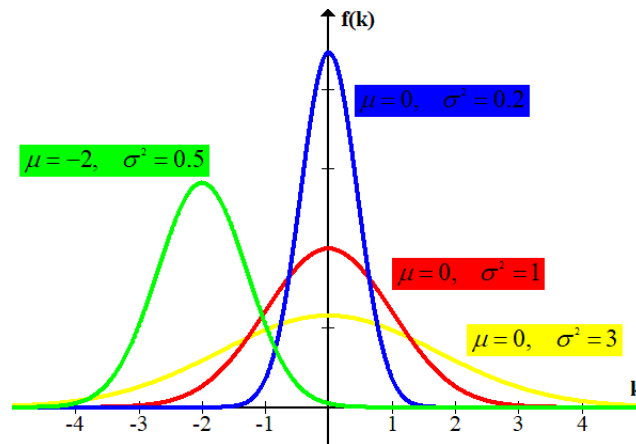
תרגיל 5.13 בר היא כלכלנית בכירה. זמן היציאה שלה מהעבודה בסוף יום מתפלג אחיד בין 6 ל-8 בערב. מה תוחלת זמן היציאה? מה השונות? מה הסיכוי שבר תצא מן העבודה עד 18:15? עד 19:15?

פתרון: נענה על כל סעיף בנפרד. נסמן את זמן היציאה ב- $T \sim U(6, 8)$. לכן $E[T] = \frac{8+6}{2} = 7$ ו- $V(T) = \frac{(8-6)^2}{12} = \frac{1}{3}$. $\Pr(T \leq 7.25) = \frac{7.25-6}{8-6} = \frac{5}{8}$ ו- $\Pr(T \leq 6.25) = \frac{6.25-6}{8-6} = \frac{1}{8}$.

5.3.2 התפלגות נורמאלית

התפלגות נורמאלית היא, ככל הנראה, אחת ההתפלגויות הנפוצות ביותר בעולם הסטטיסטי וההסתברותי. הסיבות המלאות לכך נעוצות בנושאים אשר נלמד לעומק בקורס ההמשך בסטטיסטיקה, אבל נוכל כעת לתאר מספר תופעות המייצרות, בקירוב טוב, התפלגות נורמאלית. לדוגמא, המשקל או הגובה של אוכלוסיה מסויימת, לחץ הדם של קבוצת אנשים גדולה, הקוטר של ברגים המיוצרים במפעל, אורך החיים של מכוניות במדינה כלשהי, ועוד. בקורס סטטיסטיקה הבא אנחנו נראה שהתוצאה הממוצעת של סדרת ניסויים בלתי תלויים מתפלגת נורמאלית, וזאת בלי שום תלות בהתפלגות או בתוצאות האפשריות בניסויים! זאת אומרת, לא משנה איזה ניסוי נבצע, כל עוד נבצע אותו באופן בלתי תלוי מספיק פעמים, התוצאה הממוצעת תניב התפלגות נורמאלית. זאת תוצאה (תיאורטית!) די מדהימה שמאפשרת לנו ללמוד דברים מהותיים על האוכלוסיה מבלי לדעת כיצד ערכים מסויימים מתפלגים. יחד עם זאת, השימוש בפועל בתכונות הללו שונה מהותית מן העולם התיאורטי ומשמעותית יותר בעיני, ולכן ייתכנו לא מעט טעויות. אבל, דיה לצרה בשעתה, כעת נתמקד בהתפלגות עצמה. נתחיל בתיאור ההתפלגות. הדרך הטובה ביותר לתאר התפלגות נורמאלית היא באמצעות עקומת פעמון. הסתכלו באיור 18. באיור מוצגים גרפים אשר נראים כמו פעמונים. הכל אחד מהם מייצג את **הצפיפות של ההתפלגות הנורמאלית** והנוסחה האלגברית שלה היא

$$f_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right).$$



איור 18: צפיפות התפלגות נורמאלית.

הפרמטרים המיידיים של ההתפלגות הנורמאלית הם μ אשר מייצגת את התוחלת, מרכז המסה של העקומה - מרכז הפעמון, ו- σ אשר מייצגת את סטיית התקן של ההתפלגות ובאה לידי ביטוי ברוחב הפעמון (מידת הפיזור). משתנה מקרי נורמאלי כזה נסמן ב- $X \sim N(\mu, \sigma^2)$. באיור 18 מוצגות מספר התפלגויות בעלות פרמטרים שונים. ניתן לראות כיצד שינוי של התוחלת וסטיית התקן משנות את מבנה ההתפלגות. צפיפות ההתפלגות הנורמאלית בעלת מספר תכונות חשובות. בראש ובראשונה הצפיפות היא **תמיד חיובית** ושואפת לאפס בגבולות כאשר $t \rightarrow \pm\infty$. בנוסף, היא **סימטרית סביב** μ (התוחלת), והיא קמורה בקטע $[\mu - \sigma, \mu + \sigma]$, ואחרת קעורה. מן הסתם, ובדומה לכל פונקציות צפיפות, השטח התחום בין גרף הפונקציה לציר x שווה ל-1 כנדרש מתנאי הנרמול.

האם מכל התכונות הללו ניתן לבצע חישובים בסיסיים עם התפלגות נורמאלית? לצערנו, התשובה איננה חד משמעית, כן ולא. בפרט, ניתן לבצע חישובים אבל לא בצורה מדויקת והם אינם חישובים אנליטיים, אלא חישובים נומריים. נסביר. הבעיה המרכזית עם חישובי הסתברות והתפלגות נורמאלית היא ביצוע אינטגרציה על פונקציה הצפיפות. כבר ראינו כי $\Pr(a < X \leq b) = \int_a^b f_X(t) dt$ ולכן בכדי לחשב הסתברויות (ובכדי למצוא את הפה"מ), צריך לבצע אינטגרציה על הצפיפות. הבעיה עם הנקודה הזאת היא שאין פונקציה קדומה לפונקציה הצפיפות, זאת אומרת לא ניתן לחשב את האינטגרל הנ"ל להתפלגות הנורמלית. אין לו פתרון אנליטי. הדרך היחידה לבצע את החישוב היא נומרית, באמצעות מחשב, וזאת עבור כל שתי נקודות a ו- b ספציפיות. חוסר היכולת למצוא פתרון אלגברי לאינטגרל הנ"ל איננו עניין זמני. יש הוכחה לכך שאין לאינטגרל פתרון! במילים אחרות, אין אף פונקציה $F(t)$ שהנגזרת שלה שווה לפונקציה הצפיפות של התפלגות נורמאלית. לשמחתנו, נמצא לכך פתרון והוא נעוץ בהתפלגות נורמאלית ספציפית, אשר מכונה "**התפלגות נורמאלית סטנדרטית**".

משתנה מקרי נורמאלי סטנדרטי הוא משתנה מקרי נורמאלי עם צמד פרמטרים ספציפיים $\mu = 0$ ו- $\sigma = 1$. זאת אומרת, המשתנה הנורמאלי הוא בעל תוחלת אפסית וסטיית תקן השווה ל-1. נהוג לסמן אותו ב- $Z \sim N(0, 1)$. אפריורית, אין שום דבר מיוחד בהתפלגות הזאת, למעט העבודה שהפרמטרים שלה נוחים מבחינה חישובית והצפיפות שלה נתונה על ידי $f_Z(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$. אם כך, מה החידוש הגדול? ובכן, אין באמת חידוש, אלא עניין טכני פעוט של חישובים קלים. ננסה לפשט.

לא ניתן לפתור את בעיית האינטגרציה על פונקציה הפעמון (זאת אומרת, אנחנו יודעים להוכיח שאין פתרון אנליטי לאינטגרל), לכן חייבים לבצע חישובים נומריים. אבל אין סיבה שכל סטודנט אשר מבקש לבצע חישוב קטן יהפוך למתכנת. הפתרון הפשוט הוא שימוש בטבלאות. מייצרים טבלה גדולה עם הערכים המתאימים להסתברויות חשובות ומשתמשים בטבלה הזאת בכדי לבצע חישובים. מה הטבלה מכילה? הטבלה מכילה את הערכים של פונקציה ההתפלגות המצטברת של Z . נדון בפונקציה זאת בקצרה.

פונקציה ההתפלגות המצטברת של Z , משתנה מקרי נורמאלי סטנדרטי, מסומנת ב- $\Phi(t) = \Pr(Z \leq t)$, מתקבלת מאינטגרציה על הצפיפות בקטע $(-\infty, t)$. הטבלה המוצגת באיור 19 מציגה את ערכי הפונקציה בנקודות שונות. שימו לב לסקאלה העליונה והסקאלה השמאלית בטבלה. השמאלית היא סקאלה גסה בעלת קפיצות של 0.1 והעליונה יותר עדינה עם קפיצות של 0.01. הסקאלות משלימות אחת את השנייה בכדי להניב את הערכים בתוך הטבלה. לדוגמא, בכדי למצוא את $\Phi(1.23)$ נסתכל בשורה המסומנת בערך 1.2 משמאל ובעמודה המתחילה בערך 0.03 ממעלה. בתא המתאים בטבלה, במקרה זה התא עם הספרות 0.8907, נתון הערך של פונקציה ההתפלגות המצטברת

בנקודה 1.23, קרי $\Phi(1.23) = 0.8907$. באופן דומה, $\Phi(0.35) = 0.6368$ וכן הלאה. בדקו כי אתם מבינים את אופן השימוש בטבלה. אם כך, השימוש בטבלה הוא חלק קריטי בחישובים המבוצעים להתפלגות נורמאלית, אבל הטבלה לבדה איננה מספיקה. לא ניתן לתאר את כל ערכי הפה"מ, ולכן גם בטבלה מוצגים רק ערכי הפונקציה מ-0 ועד 3.5. קל לראות כי עבור ערכים גבוהים מ-3.5 הפה"מ היא כמעט 1 ולכן אין טעם לכתוב ערכים אלו וניתן להניח כי הם שווים ל-1, בקירוב טוב. אבל מה לגבי הערכים השליליים? למה שווה $\Phi(-t)$, לדוגמא? בכדי לחשב את ערכי הפונקציה בחלק השלילי של הציר, נשתמש בסימטריה של הצפיפות סביב התוחלת (השווה לאפס), ונבין כי $\Pr(Z \leq -t) = \Pr(Z \geq t)$. במילים, ההסתברות של הזנב הימני שווה להסתברות של הזנב השמאלי המתחילים בנקודה t ו- $-t$, בהתאמה. השיוויון הזה נובע מסימטריה סביב התוחלת האפסית. לכן,

$$\Phi(-t) = \Pr(Z \leq -t) = \Pr(Z \geq t) = 1 - \Pr(Z < t) = 1 - \Phi(t),$$

וקיבלנו את הנוסחה $\Phi(-t) = 1 - \Phi(t)$. בעזרת נוסחה זאת נוכל לחשב כל ערך של הפה"מ, גם בתחום השלילי, בעזרת הטבלה המקורית. הנוסחה מראה כי מספיק לדעת את ערכי הפה"מ רק בתחום החיובי בכדי להסיק לגבי ערכיו בתחום השלילי (זה נובע ישירות מהסימטריה של פונקצית הצפיפות). לדוגמא, $\Phi(-0.2) = 1 - \Phi(0.2) = 1 - 0.5793$.

תרגיל 5.14 חשבו את ההסתברויות הבאות: $\Pr(Z > 0.45)$, $\Pr(Z < -1.32)$, $\Pr(Z > -1.32)$, $\Pr(-1 \leq Z \leq 1)$, $\Pr(-2 \leq Z \leq 2)$.

פתרון: נבצע חישוב ישיר בעזרת התכונות שראינו לעיל והטבלה המצורפת.

$$\Pr(Z > 0.45) = 1 - \Pr(Z \leq 0.45) = 1 - \Phi(0.45) = 1 - 0.6736 = 0.3264,$$

$$\Pr(Z < -1.32) = \Phi(-1.32) = 1 - \Phi(1.32) = 1 - 0.9066 = 0.0934,$$

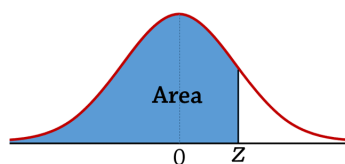
$$\Pr(Z > -1.32) = 1 - \Pr(Z < -1.32) = 1 - 0.0934 = 0.9066,$$

$$\Pr(-1 \leq Z \leq 1) = \Phi(1) - \Phi(-1) = \Phi(1) - [1 - \Phi(1)] = 2\Phi(1) - 1 = 2 \cdot 0.8413 - 1 = 0.6826,$$

שימו לב, החישוב האחרון מראה שההסתברות להיות במרחק של לכל היותר סטיית תקן אחת מהממוצע הוא 0.68 לערך.

$$\Pr(-2 \leq Z \leq 2) = \Phi(2) - \Phi(-2) = \Phi(2) - [1 - \Phi(2)] = 2\Phi(2) - 1 = 2 \cdot 0.9772 - 1 = 0.9544,$$

קרי, ההסתברות להיות במרחק של לכל היותר שתי סטיות תקן מן הממוצע הוא 0.95. אם נסתכל על המשלים, הסיכוי של דגימה אקראית מהתפלגות נורמאלית להיות רחוקה ביותר משתי סטיות תקן מן הממוצע קטן מ-5%!



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

איור 19: טבלת התפלגות נורמאלית סטנדרטית.

לאחר הדיון הממצה בנוגע להתפלגות נורמאלית סטנדרטית, הגיע הזמן לחזור חזרה לשאלה המקורית - כיצד נבצע חישוב עבור התפלגות נורמאלית כלשהי? הפתרון קצת פשוט וכבר ראינו אותו בעבר. נניח כי $X \sim N(\mu, \sigma^2)$ הוא משתנה נורמאלי סטנדרטי כלשהו, ואנו מעוניינים לחשב את ההסתברות $\Pr(a \leq X \leq b)$. החישוב יתחיל בביצוע תקנון למשתנה ולכל יתר האגפים במקביל בכדי לשמר את האי־שיוויונות. התקנון הופך את המשתנה הנורמאלי לתצורה של משתנה נורמאלי סטנדרטי. ברגע שנקבל משתנה סטנדרטי, נשתמש בטבלה כפי שכבר ראינו. למה הכוונה? נתחיל מההסתברות שאנו מעוניינים לחשב $\Pr(a \leq X \leq b)$ ונחסר מכל האגפים את התוחלת בכדי לקבל את הביטוי $\Pr(a - \mu \leq X - \mu \leq b - \mu)$. כעת נחלק את כל האגפים בסטיית התקן σ , $\Pr\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right)$. קיבלנו לבסוף בדיק

את מה שרצינו. בתהליך התקנון, המשתנה X הפך ל- $\frac{X-\mu}{\sigma}$, אבל כבר ראינו כי תקנון שכזה מייצר משתנה מקרי עם תוחלת אפסית ושונות מנורמלת (שווה ל-1). יחד עם זאת, לא שינינו את אופי ההתפלגות, אלא רק את יחידות המידה (הזזה וכפל בקבוע), לכן זאת עדיין התפלגות נורמאלית, אלא עם תוחלת אפסית ושונות השווה לאחת. אם כן, $\frac{X-\mu}{\sigma} = Z$ ו-

$$\begin{aligned}\Pr(a \leq X \leq b) &= \Pr\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) \\ &= \Pr\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).\end{aligned}$$

כך "מתרגמים" בעית חישוב של התפלגות נורמאלית כלשהי, להתפלגות נורמאלית סטנדרטית, ולבסוף משתמשים בטבלה.

תרגיל 5.15 ציוני הפסיכומטרי נקבעים כך שהתפלגותם נורמאלית עם תוחלת 500 וסטטיית תקן של 100. נבחר תלמיד אקראי ונסמן את ציונו ב- $X \sim N(500, 100^2)$. חשבו את הסיכוי שציון התלמיד מעל 700; מתחת ל-400; ייסטה בסטיית תקן אחת לכל היותר מהתוחלת; בין 650 ל-500. מה ציון הפסיכומטרי כך שציוני 60% מהסטודנטים נמוכים ממנו?

פתרון: נתחיל בחישובים הבסיסיים. נתון $X \sim N(500, 100^2)$ ואנו מעוניינים לחשב את $\Pr(X > 700)$. נבצע תקנון לטובת החישוב כפי שהדגמנו קודם לכן.

$$\Pr(X > 700) = \Pr\left(\frac{X-500}{100} > \frac{700-500}{100}\right) = \Pr(Z > 2) = 1 - \Phi(2) = 1 - 0.9772 = 0.0228,$$

ז"א רק 2.28% מההתלמידים מקבלים ציון הגבוה מ-700. נבצע חישוב דומה ליתר הסעיפים,

$$\Pr(X < 400) = \Pr\left(\frac{X-500}{100} < \frac{400-500}{100}\right) = \Pr(Z < -1) = 1 - \Phi(1) = 0.1578.$$

את הסעיף הבא נפתור באופן כללי עבור כל משתנה נורמאלי. נחשב את הסיכוי שהמשתנה נמצא במרחק סטיית תקן אחת, לכל היותר, מהתוחלת שלו - $\Pr(|X - \mu| < \sigma)$.

$$\begin{aligned}\Pr(|X - \mu| < \sigma) &= \Pr(-\sigma < X - \mu < \sigma) = \Pr\left(-\frac{\sigma}{\sigma} < \frac{X-\mu}{\sigma} < \frac{\sigma}{\sigma}\right) = \Pr(-1 < Z < 1) \\ &= \Phi(1) - \Phi(-1) = \Phi(1) - [1 - \Phi(1)] = 2\Phi(1) - 1 = 0.6826.\end{aligned}$$

כמובן שקיבלנו תשובה זהה לחישוב שבצענו קודם לכן עבור משתנה נורמאלי סטנדרטי.

$$\begin{aligned}\Pr(500 < X < 650) &= \Pr\left(\frac{500-500}{100} < \frac{X-500}{100} < \frac{650-500}{100}\right) = \Pr(0 < Z < 1.5) \\ &= \Phi(1.5) - \Phi(0) = 0.9332 - 0.5 = 0.4332.\end{aligned}$$

הסעיף הבא נראה קצת יותר מסובך, אבל החישוב הוא זהה. כעת אנחנו נדרשים למצוא מספר a , כך ש- $\Pr(X < a) = 0.6$. במילים, צריכים לחשב את הציון a כך ש-60% מהתלמידים יקבלו פחות ממנו. כמובן שציון זה יהיה גדול מ-500 שהוא מרכז ההתפלגות. נבצע את החישוב ישירות.

$$0.6 = \Pr(X < a) = \Pr\left(\frac{X-500}{100} < \frac{a-500}{100}\right) = \Pr\left(Z < \frac{a-500}{100}\right) = \Phi\left(\frac{a-500}{100}\right),$$

נבדוק בטבלה איזה ערך מייצר לנו הסתברות של 0.6, ונקבל שמדובר בערך 0.25. כלומר, $\Phi(0.25) = 0.6$. לכן נדרוש ש- $\frac{a-500}{100} = 0.25$ או $a = 500 + 100 \cdot 0.25 = 525$. זה הציון ש-60% מהתלמידים יקבלו פחות ממנו בבחינה.

הסעיף האחרון הוא הזדמנות לדון באחוזונים של משתנה מקרי נורמאלי. האחוזון ה-60 של התפלגות נורמאלית כלשהי, או משתנה נורמאלי $X \sim N(\mu, \sigma)$ זה הערך $x_{.60}$ כך ש- $\Pr(X < x_{.60}) = 0.6$. באותו האופן מגדירים כל אחוזון אחר $\Pr(X < x_{.35}) = 0.35$ ו- $x_{.35}$ זה הערך אשר הסיכוי להיות קטן ממנו הוא 35%. באופן כללי, האחוזון ה- p , מסומן ב- x_p זה מספר כך ש- $\Pr(X < x_p) = p$ הסיכוי להיות קטן ממנו הוא p .

כיצד מוצאים את האחוזונים במקרה של התפלגות נורמאלית? שוב, משתמשים בטבלה. בכדי למצוא אחוזון כלשהו נבצע תקנון ונשתמש באחוזונים המתאימים של התפלגות נורמאלית סטנדרטית. ז"א,

$$p = \Pr(X < x_p) = \Pr\left(\frac{X - \mu}{\sigma} < \frac{x_p - \mu}{\sigma}\right) = \Pr\left(Z < \frac{x_p - \mu}{\sigma}\right) = \Phi\left(\frac{x_p - \mu}{\sigma}\right)$$

$$p = \Phi\left(\frac{x_p - \mu}{\sigma}\right),$$

לכן $\frac{x_p - \mu}{\sigma}$ זה הערך אשר מניב הסתברות p בטבלה. למעשה, $\frac{x_p - \mu}{\sigma}$ הוא האחוזון ה- p של התפלגות נורמאלית סטנדרטית. נוהג לסמן את האחוזון הזה ב- z_p . על כן, $z_p = \frac{x_p - \mu}{\sigma}$, ולחילופין $x_p = \mu + \sigma \cdot z_p$. במילים, הנוסחה האחרונה קובעת שהאחוזון ה- p של התפלגות נורמאלית שווה לסכום התוחלת μ עם המכפלה $\sigma \cdot z_p$ של סטיית התקן עם האחוזון ה- p של ההתפלגות הנורמאלית הסטנדרטית. כמובן ש- μ ו- σ נתונים ולכן נותר לחשב את z_p . לשם כך יש את הטבלה. נחפש בטבלה את הערך אשר מניב הסתברות p , וזהו z_p . לדוגמא, $z_{0.65} = 0.385$, $z_{0.99} = 2.325$, $z_{0.95} = 1.945$, $z_{0.9} = 1.285$, $z_{0.78} = 0.775$.

מה נעשה במקרה של אחוזון $p < 0.5$ אשר אינו מופיע בטבלה, לדוגמא $z_{0.1}$? נבין שמדובר בערך שלילי שכן בכדי שההסתברות תהיה קטנה מחצי, צריכים להיות בחצייה השמאלית של ההתפלגות (משמאל לציר y). בנוסף, משיקולי סימטריה

$$0.1 = \Pr(Z < z_{0.1}) = \Phi(z_{0.1}) = 1 - \Phi(-z_{0.1})$$

$$0.1 = 1 - \Phi(-z_{0.1})$$

$$\Phi(-z_{0.1}) = 1 - 0.1 = 0.9$$

$$-z_{0.1} = z_{0.9},$$

ובאופן כללי קיבלנו את הנוסחה, $z_p = -z_{1-p}$. על כן, עבור כל הסתברות הקטנה מחצי נוכל להיעזר בנוסחה המדוברת. כעת נוכל לחשב כל אחוזון של התפלגות נורמאלית בעזרת הטבלה, והנוסחאות $x_p = \mu + \sigma \cdot z_p$, יחד עם $z_p = -z_{1-p}$.

תרגיל 5.16 בבנק אוצר הסטודנט בוחנים את החשבונות של הלקוחות ואת מדיניות הריבית. הם מעריכים כי יתרת החשבון (בש"ח) של לקוח מתפלגת נורמלית עם $\mu = 2200$ ו- $\sigma = 600$.

1. לאיזה אחוז מהלקוחות יש יתרה של למעלה מ-3000 ש"ח? לאיזה אחוז יתרה של פחות מ-1400?

2. הבנק מעוניין להציע ריבית גבוהה יותר ל-5% מהלקוחות - בעלי היתרה הגבוהה ביותר. מהי היתרה המינימלית הנדרשת כדי לקבל את הריבית המיוחדת?

3. בבוקר יום א' נמצאים בבנק 30 לקוחות. מהי תוחלת מספר הלקוחות בעלי הריבית המיוחדת מהסעיף הקודם?

פתרון: נסמן $X \sim N(2200, 600^2)$.

1. $\Pr(X \leq 1400) =$ באותו האופן, $\Pr(X \geq 3000) = \Pr\left(\frac{X-2200}{600} \geq \frac{800}{600}\right) = \Pr\left(Z \geq \frac{4}{3}\right) = 1 - \Phi\left(\frac{4}{3}\right) = 1 - 0.9066 = 0.0934$.
 $\Pr\left(Z \leq \frac{1400-2200}{600}\right) = \Phi\left(-\frac{4}{3}\right) = 1 - \Phi\left(\frac{4}{3}\right) = 0.0934$.

2. אנחנו מחפשים את האחוזון ה-95 של ההתפלגות, כלומר את X_{95} כך ש- $\Pr(X \leq X_{95}) = 0.95$. נתקן לקבל מ"מ נורמלי סטנדרטי ונקבל $\Pr\left(Z \leq \frac{X_{95}-2200}{600}\right) = 0.95$. מהטבלה אנו רואים שהאחוזון ה-95 של Z הינו 1.65 ולכן $\frac{X_{95}-2200}{600} = 1.65$ ולכן $X_{95} = 1.65 \cdot 600 + 2200 = 3190$.

3. ההסתברות של לקוח מקרי יש את הריבית הנ"ל היא 0.05. לכן מספר הלקוחות עם הריבית המיוחדת מתפלג $Bin(30, 0.05)$ ולכן התוחלת היא $30 \cdot 0.05 = 1.5$.

תרגיל 5.17 מפעל מייצר אריזות שוקולד במשקל 100 גרם. בפועל משקל כל אריזה נקבע על פי התפלגות נורמאלית $X \sim N(100, 5^2)$.

1. מה אחוז האריזות ששוקלות מעל 95 גרם?

2. במפעל מעוניינים לשפר את הנתון הקודם כך שההסתברות של חבילה להיות גבוהה מ-95 גרם תהיה 0.9. בפני מנהל המפעל עומדות שתי דרכי פעולה: העלאת התוחלת תוך שימור סטיית התקן והקטנת סטיית התקן תוך שימור התוחלת. מצאו כיצד עליו לשנות כל פרמטר בנפרד בכדי להגיע לתוצאה הרצויה.

פתרון: בסעיף הראשון אנחנו מעוניינים לחשב את ההסתברות

$$\Pr(X > 95) = \Pr\left(\frac{X - 100}{5} > \frac{95 - 100}{5}\right) = \Pr(Z > -1) = \Pr(Z < 1) = \Phi(1) = 0.8413.$$

נשתמש בתוצאה זו עבור הסעיף השני. נתחיל בעדכון התוחלת. אנחנו מעוניינים לייצר מצב בו האחוזון ה-10 הוא 95 גרם, קרי $\Pr(X > 95) = 0.9$ אשר שקול ל- $\Pr(X < 95) = 1 - 0.9 = 0.1$. לכן,

$$x_{0.1} = 95 = \mu + \sigma \cdot z_{0.1} = \mu + 5 \cdot (-1.28),$$

$$\mu = 95 + 5 \cdot 1.28 = 101.4.$$

התוחלת החדשה צריכה להיות 101.4, גבוהה ב-1.4 גרם מהמקורית. חישוב דומה עבור סטיית התקן מניב

$$x_{0.1} = 95 = 100 + \sigma \cdot z_{0.1} = 100 + \sigma \cdot (-1.28),$$

$$1.28 \cdot \sigma = 100 - 95 \Rightarrow \sigma = \frac{5}{1.28} \approx 3.9.$$

סטיית התקן צריכה לקטון לגודל של 3.9 בכדי לקיים את התנאי המתבקש בשאלה.

תכונה חשובה בהתפלגות נורמאלית מתייחסת לסכום של משתנים נורמאליים בלתי תלויים. סכום של משתנים נורמאליים ב"ת הוא בעצמו נורמאלי ואת התוחלת והשונות נוכל למצוא בעזרת תכונת הלינאריות. בפרט, נניח כי $X_i \sim N(\mu_i, \sigma_i^2)$ עבור כל $i = 1, \dots, n$. בנוסף, ניקח מספרים $a_1, a_2, \dots, a_n$, אזי $\sum_{i=1}^n a_i X_i$ גם מתפלג נורמאלית ו-

$$E\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i E[X_i] = \sum_{i=1}^n a_i \mu_i,$$

$$V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n V(a_i X_i) = \sum_{i=1}^n a_i^2 V(X_i) = \sum_{i=1}^n a_i^2 \sigma_i^2,$$

אלו התוחלת וסטיית התקן של הסכום. קרי, $\sum_{i=1}^n a_i X_i \sim N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2\right)$.

תרגיל 5.18 אלון, בן, ובר הם סוחרי מט"ח והתשואה השנתית שלהם מתפלגת נורמאלית ע"פ $X_1 \sim N(5, 10^2)$, $X_2 \sim N(6, 12^2)$, $X_3 \sim N(7, 14^2)$ בהתאמה ובאופן ב"ת בסוחרים האחרים. סמי וספדי החליטו להשקיע יחד אצל אלון, בן, ובר באופן שווה (שליש אצל כל אחד). מה הסיכוי שהתשואה של אלון תהיה גבוהה מהתשואה של בר? מה הסיכוי שהתשואה השנתית של סמי וספדי תהיה גבוהה מ-6.5?

פתרון: נתחיל ברישום הבעיה בצורה מתמטית. השאלה הראשונה דורשת חישוב $\Pr(X_1 > X_3)$ והשאלה השנייה מצריכה חישוב של $\Pr(X_1 - X_3 > 0)$. נפתח בחישוב הראשון. נפתח $\Pr\left(\frac{1}{3}(X_1 + X_2 + X_3) > 6.5\right)$. כפי שראינו

$X_1 - X_3$ מתפלג נורמאלי עם פרמטרים $\mu_{X_1 - X_3} = \mu_1 - \mu_2 = 5 - 7 = -2$ ו- $V(X_1 - X_3) = V(X_1) + V(-X_3) = 244$. כאמור, $X_1 - X_3 \sim N(-2, 244)$. נסמן את המשתנה ב- X ונחשב $\Pr(X > 0)$.

$$\begin{aligned}\Pr(X > 0) &= \Pr\left(\frac{X - (-2)}{\sqrt{244}} > \frac{0 - (-2)}{\sqrt{244}}\right) = \Pr\left(Z > \frac{1}{\sqrt{61}}\right) = 1 - \Pr\left(Z < \frac{1}{\sqrt{61}}\right) \\ &= 1 - \Phi(0.128) \approx 1 - 0.55 = 0.45.\end{aligned}$$

נעבור לחישוב השני. נסמן את המשתנה ב- Y , אשר מתפלג נורמאלי גם כן.

$$\begin{aligned}E\left[\frac{1}{3}(X_1 + X_2 + X_3)\right] &= \frac{1}{3}\left(\sum_{i=1}^3 E[X_i]\right) = 6, \\ V\left[\frac{1}{3}(X_1 + X_2 + X_3)\right] &= \frac{1}{9}\left(\sum_{i=1}^3 V[X_i]\right) = \frac{440}{9}, \\ \sigma_{\frac{1}{3}(X_1 + X_2 + X_3)} &= \sqrt{\frac{440}{9}} \approx 7.\end{aligned}$$

לכן, עבור $Y \sim N(6, 7^2)$ נקבל את התוצאה הבאה

$$\begin{aligned}\Pr(Y > 6.5) &= \Pr\left(\frac{Y - 6}{7} > \frac{6.5 - 6}{7}\right) = \Pr\left(Z > \frac{1}{14}\right) = 1 - \Pr\left(Z \leq \frac{1}{14}\right) \\ &= 1 - \Phi(0.07142) \approx 1 - 0.5279 = 0.4721.\end{aligned}$$

5.3.3 התפלגות מעריכית

הדיון על התפלגות מעריכית מתחיל דווקא בהתפלגות אחרת מן העבר - התפלגות פואסון. הדוגמא הראשונה שראינו עבור התפלגות פואסון היא מוקד שירות הלקוחות. במוקד שירות לקוחות מספקים שירות לכל לקוחות חברת הסלולר הגדולה של פונטי-פאנדי. בכל שעה בממוצע מגיעות למוקד $\lambda = 5$ פניות שנזקקות למענה. משתנה מקרי פואסוני מתאר את מספר השיחות שנכנסות למוקד בשעה אחת, $X \sim P(5)$. בשעתיים נכנסות בממוצע 10 שיחות והמשתנה המתאר אותם יהיה $Y \sim P(10)$. יש לנו נוסחה סגורה עבור כל אחד מהמשתנים הבדידים הללו, והתומך שלהם הוא כל המספרים השלמים האי שליליים.

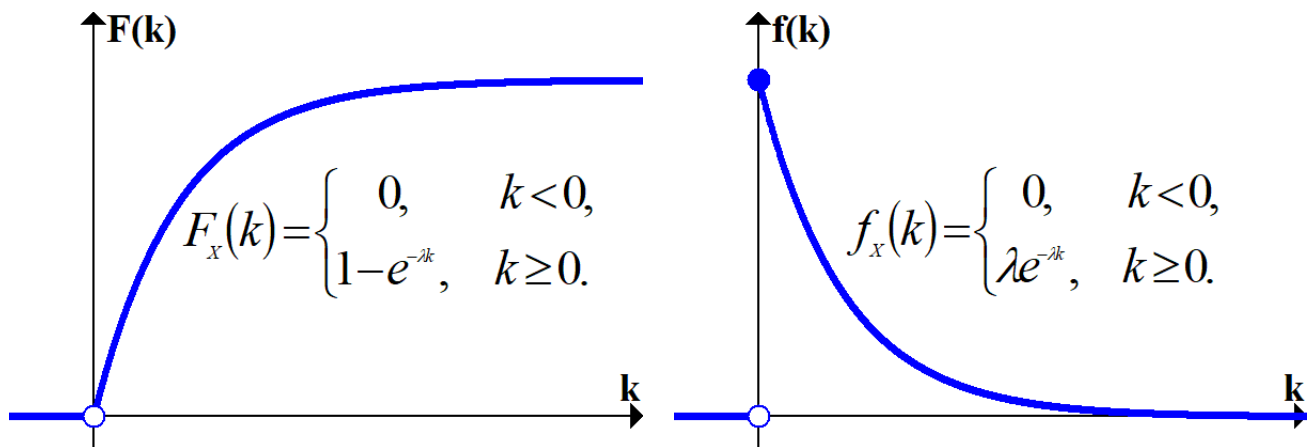
כעת ניקח את אותה הדוגמא ונציג אותה בצורה אחרת. במקום לחשוב על מספר השיחות הנכנסות בשעה, נחשוב על הזמן בין כל שיחה לשיחה. לדוגמא, אם בממוצע נכנסות 5 שיחות בשעה, אזי הזמן הממוצע בין השיחות הוא 12 דקות. אנחנו בפועל מתארים את אותו הניסוי, אבל עושים זאת בצורה שונה בעזרת מדידת הזמן בין האירועים. הזמן הוא רציף ולכן הזמן בין השיחות השונות הוא משתנה מקרי רציף, הנקרא משתנה מקרי מעריכי. באופן רשמי, משתנה מקרי מעריכי מודד את הזמן (או המרחק) בין אירועים שונים המתרחשים לפי תהליך פואסון (התפלגות פואסונית). נעבור לדוגמא נוספת - מספר טיפות הגשם הנופלות ביחידת שטח נתונה מתפלגות על פי התפלגות פואסון. ניתן להציג את אותן הטיפות בעזרת המרחק ביניהן. לדוגמא, במקרה החד מימדי ניקח מטר אחד ונניח כי בממוצע נפלו לאורכו 10 טיפות. באופן דומה ניתן לקבוע כי המרחק הממוצע בין כל טיפה הוא 10 ס"מ, ובכך להציג את אותה הסיטואציה רק בעזרת פרמטר אחר, רציף. נגדיר את ההתפלגות המעריכית בצורה מפורשת. התפלגות מעריכית מתקבלת על ידי מדידת אורכי הזמן (או השטח) בין אירועים המתרחשים לפי התפלגות פואסון. משתנה מקרי מעריכי מסומן ב- $X \sim \exp(\lambda)$, כאשר $0 < \lambda$ הוא הפרמטר של התפלגות פואסון המתאימה, ז"א קצב האירועים הממוצע ליחידת זמן (או שטח). הצפיפות ופונקציית ההתפלגות המצטברת של X הן

$$f_X(k) = \begin{cases} 0, & k < 0, \\ \lambda e^{-\lambda k}, & k \geq 0, \end{cases}, \quad F_X(k) = \begin{cases} 0, & k < 0, \\ 1 - e^{-\lambda k}, & k \geq 0. \end{cases}$$

ננסה להסביר את פונקצית ההתפלגות המצטברת של המשתנה המעריכי. אנחנו מעוניינים לחשב את $F_X(k) = \Pr(X \leq k)$, במילים - ההסתברות שהאירוע הראשון יתרחש בזמן הקטן מ- k . מאחר וזמן הוא פרמטר חיובי נתחיל את המדידות בזמן 0. נסמן ב- Y את המשתנה הפואסוני $Y \sim P(\lambda k)$. המשתנה Y מודד מספר האירועים ביחידת הזמן $[0, k]$ בעלת אורך k . מה הסיכוי שיתרחש לפחות אירוע אחד ביחידת הזמן הזאת? ההסתברות הזאת היא למעשה מה שאנחנו מעוניינים לחשב, כי אם יתרחש לפחות אירוע אחת ביחידת הזמן הזאת, אז הזמן עד לאירוע הראשון יהיה קטן מ- k . ובכן,

$$\Pr(Y \geq 1) = 1 - \Pr(Y = 0) = 1 - e^{-\lambda k} \frac{(\lambda k)^0}{0!} = 1 - e^{-\lambda k},$$

כנדרש. נשרטט את צמד הפונקציות לעיל באיור 20. בנוסף לדוגמאות שראינו לעיל עבור התפלגות מעריכית, גם אורך החיים של רכבים



איור 20: פונקצית התפלגות מצטברת וצפיפות של משתנה מעריכי

חשמליים מתפלג מעריכית. לדוגמא: מכונות כביסה, טלפונים סלולריים, טלוויזיות, נורות חשמל, וכן הלאה.

תרגיל 5.19 בתחנת כיבוי האש של פונטי-פאנדי מתקבלת שיחה אחת בממוצע בכל 3 דקות. מה הסיכוי שהשיחה הראשונה תגיע בין הדקה השנייה לדקה הרביעית? מה הסיכוי שהשיחה הראשונה תגיע אחרי 5 דקות? בהנחה שבחמש הדקות הראשונות לא התקבלה שיחה, מה הסיכוי שנמתין לפחות עוד דקה עד לשיחה הראשונה?

פתרון: הזמן עד השיחה הראשונה מתפלג פואסונית $Y \sim \exp\left(\frac{1}{3}\right)$ כאשר הזמן נמדד בדקות. נחשב את ההסתברויות $\Pr(2 \leq Y \leq 4)$, $\Pr(Y > 5)$, ו- $\Pr(Y > 5 + 1 | Y > 5)$.

$$\Pr(2 \leq Y \leq 4) = F_Y(4) - F_Y(2) = 1 - e^{-\frac{1}{3} \cdot 4} - \left(1 - e^{-\frac{1}{3} \cdot 2}\right) = e^{-2/3} - e^{-4/3} \approx 0.25,$$

$$\Pr(Y > 5) = 1 - \Pr(Y \leq 5) = 1 - F_Y(5) = 1 - \left(1 - e^{-\frac{1}{3} \cdot 5}\right) = e^{-5/3} \approx 0.189,$$

$$\Pr(Y > 5 + 1 | Y > 5) = \frac{\Pr(Y > 6, Y > 5)}{\Pr(Y > 5)} = \frac{\Pr(Y > 6)}{\Pr(Y > 5)} = \frac{1 - F_Y(6)}{1 - F_Y(5)} = \frac{e^{-6/3}}{e^{-5/3}} = e^{-1/3} = \Pr(Y > 1).$$

נשים לב שבחישוב האחרון קיבלנו תכונה מוכרת - **חוסר זיכרון**. הסיכוי שנמתין 6 דקות לפחות עד לשיחה הראשונה, בהנחה שכבר המתנו 5 דקות ללא שיחה, היא בדיוק הסיכוי שנמתין דקה אחת נוספת - זו תכונת חוסר הזיכרון שראינו עבור משתנה מקרי גיאומטרי. נסכם בטענה הבאה.

טענה 5.20 עבור משתנה מקרי מעריכי, $X \sim \exp(\lambda)$, וכל צמד מספרים $s, t > 0$ מתקיים

$$\Pr(X > s + t | X > s) = \Pr(X > t).$$

ההוכחה היא חישוב ישיר.

$$\begin{aligned} \Pr(X > s + t | X > s) &= \frac{\Pr(X > s + t, X > s)}{\Pr(X > s)} = \frac{\Pr(X > s + t)}{\Pr(X > s)} = \frac{1 - F_X(s + t)}{1 - F_X(s)} = \frac{1 - (1 - e^{-\lambda(s+t)})}{1 - (1 - e^{-\lambda s})} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = 1 - (1 - e^{-\lambda t}) = 1 - F_X(t) = \Pr(X > t). \end{aligned}$$

חשוב לציין שתכונת חוסר זיכרון לא קיימת בהתפלגות אחידה או נורמאלית. לכן היא ייחודית להתפלגות מעריכית מבין משפחות ההתפלגות שראינו.

חישובי התוחלת והשונות של משתנה מקרי מעריכי הם ישירים מן ההגדרה.

$$\begin{aligned} E[X] &= \int_0^\infty k f_X(k) dk = \int_0^\infty k \lambda e^{-\lambda k} dk = \left[\frac{\lambda k e^{-\lambda k}}{-\lambda} \right]_0^\infty - \int_0^\infty e^{-\lambda k} dk \\ &= 0 - \left[\frac{e^{-\lambda k}}{-\lambda} \right]_0^\infty = \frac{1}{\lambda}, \\ E[X^2] &= \int_0^\infty k^2 f_X(k) dk = \int_0^\infty k^2 \lambda e^{-\lambda k} dk = \left[\frac{\lambda k^2 e^{-\lambda k}}{-\lambda} \right]_0^\infty - \frac{2\lambda}{-\lambda} \int_0^\infty k e^{-\lambda k} dk \\ &= 0 + 2 \left[\frac{k e^{-\lambda k}}{-\lambda} \right]_0^\infty - \frac{2}{-\lambda} \int_0^\infty e^{-\lambda k} dk = \frac{2}{\lambda} \cdot \left[\frac{e^{-\lambda k}}{-\lambda} \right]_0^\infty = \frac{2}{\lambda^2}, \\ V(X) &= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda} \right)^2 = \frac{1}{\lambda^2}. \end{aligned}$$

המשמעות מאחורי התוחלת $\frac{1}{\lambda}$ היא שהזמן הממוצע עד לאירוע הבא בעל יחס הפוך למספר האירועים ליחידות זמן. במילים אחרות, ככל שיש יותר אירועים ליחידת זמן, כך הזמן הממוצע בין האירועים קטן.

תרגיל 5.21 בשוק שני יצרני נורות. מפעל 1 שולט על 60% מהשוק ומפעל 2 על היתר. אורך חיים ממוצע של נורה ממפעל 1 הוא חודשיים, בעוד מפעל 2 מייצר נורות עם אורך חיים ממוצע של ארבעה חודשים. עבור נורה שנקנית בשוק (בהתאם לנתח השוק המתאים לכל חברה), מה הסיכוי שהנורה לא תתקלקל בחודש הראשון? בהנחה והנורה שקנינו דולקת כבר חודש ללא תקלה, מה הסיכוי שמדובר בנורה ממפעל 2?

פתרון: נסמן את זמן החיים של הנורות, בכל מפעל בנפרד, לפי $X_1 \sim \exp\left(\frac{1}{2}\right)$ ו- $X_2 \sim \exp\left(\frac{1}{4}\right)$. נסמן את אורך החיים של הנורה שנקנתה ב- Y .

$$\begin{aligned} \Pr(Y > 1) &= \Pr(Y > 1 | Y = X_1) \Pr(Y = X_1) + \Pr(Y > 1 | Y = X_2) \Pr(Y = X_2) \\ &= \Pr(X_1 > 1) \cdot 0.6 + \Pr(X_2 > 1) \cdot 0.4 \\ &= [1 - F_{X_1}(1)] \cdot 0.6 + [1 - F_{X_2}(1)] \cdot 0.4 \\ &= e^{-\frac{1}{2} \cdot 1} \cdot 0.6 + e^{-\frac{1}{4} \cdot 1} \cdot 0.4 \approx 0.675. \end{aligned}$$

נענה על הסעיף השני בעזרת חוק בייס.

$$\Pr(Y = X_2 | Y > 1) = \frac{\Pr(Y > 1 | Y = X_2) \Pr(Y = X_2)}{\Pr(Y > 1)} = \frac{e^{-\frac{1}{4} \cdot 1} \cdot 0.4}{e^{-\frac{1}{2} \cdot 1} \cdot 0.6 + e^{-\frac{1}{4} \cdot 1} \cdot 0.4} \approx 0.462 > 0.4,$$

זאת אומרת שזמן החיים הארוך (בפועל) מגדיל את הסיכוי שמדובר בנורה ממפעל 2, שכן אורך החיים הממוצע של הנורות ממפעל 2 הוא ארוך יותר מאלו של מפעל 1.

חלק II

חלק ב' - סטטיסטיקה

6 סטטיסטיקה תיאורית

6.1 הסטטיסטיקה בעולם האמיתי

סטטיסטיקה נמצאת בכל שלב בחיי היומיום שלנו, מהדברים הטריטוריאליים ביותר ועד המורכבים והחשובים. נחשוב על דוגמא פשוטה ביותר. סטודנט קם בבוקר, מתארגן בביתו, וצועד להנאתו לקורס סטטיסטיקה. בדרך לאוניברסיטה הוא חוצה צומת מרומזרת ומיד שואל את עצמו - איך קובעים את זמני האור הירוק להולכי הרגל בכדי שההולכים יחצו בבטחה? מצד אחד, צריך לאפשר להולכי הרגל מספיק זמן לחצות, ולא כולם הולכים באותו הקצב. מצד שני, אם הולכי הרגל יקבלו זמן בלתי מוגבל, לעולם המכוניות לא יוכלו לחצות את הצומת. התשובה לכך נעוצה בסטטיסטיקה. אוספים נתונים על קצב המעבר של אנשים שונים, במקביל לאיסוף הנתונים לגבי עומסי התנועה בכביש בשעות היום השונות. בהתאם לנתונים מחליטים על מדיניות שתניב תועלת מקסימאלית לחברה - גם תעניק מספיק זמן להולכי הרגל ומנגד לא תפגע בתנועת הרכבים. זו דוגמא אשר מתארת באופן שלם את **התורה הסטטיסטית: איסוף הנתונים, עיבודם, הצגתם, והסקת מסקנות מהם**.

נמשיך הלאה לעולם ה-**Big Data**. זהו מונח שהועצם בשנים האחרונות עם איסוף המידע האדיר אשר מתבצע בחסות ענקיות האינטרנט (Google, Facebook, Amazon, וכד'). כיום, כמעט כל שבב מידע מתועד באינספור דרכים: חיפוש באינטרנט, תמונות מצולמות, סטאטוס ברשת חברתית, לייק, שיחות וואטספ ועוד. כל המידע הזה נאגר ומסייע לבנות פרופילים עבור כל פרט ופרט באוכלוסייה. כעת נשאלת השאלה - מה עושים עם המידע הזה? המידע עובר עיבוד על ידי חברות מסחריות באמצעות מודלים סטטיסטיים אשר מעניקים לחברות יכולות לשווק לנו בדיוק את מה שנדרש לנו. להלן ניסוי פשוט. תבחרו עיר אקראית בה לעולם לא ביקרתם (ואולי אתם מאוד רוצים לראות...). תיכנסו לגוגל ותתחילו לחפש טיסות ליעד הזה או לחילופין מלונות. תיכנסו לאתר כגון Booking.com ותסתכלו על חדרים במלונות ביעד הנכסף. לאחר כמה דקות תעשו הפסקה ותיכנסו לפרופיל שלכם ב-Facebook. ייתכן שזה ייקח מעט זמן, אבל ככל הנראה מהר מאוד אתם תראו פרסומות הנוגעות לאתרים שחיפשתם וראיתם. הפרסומות הללו מוצגות בגלל שהאלגוריתם הסטטיסטי מרמז על כך שאתם מחפשים לטייל ביעד חדש, והחברות מעוניינות לשווק לכם פרסומות מתאימות.

המודל העסקי של רוב הענקיות האינטרנט (Google, Facebook, Twitter, אתרי חדשות, אתרי ספורט, וכיו"ב) מבוסס על פרסום. זו כלכלה ענפה אשר מייצרות מאות מיליארדי דולרים ואף יותר. עשרות המיליארדים שחברות אלו מרוויחות בשנה מגיעים מפרסום אשר הן משווקות למשתמשי הקצה, לכולנו. כל המידע הזה חייב לעבור עיבוד סטטיסטי. המידע הוא כה רב (שכן אנחנו כצרכנים מבצעים המון חיפושים ומייצרים המון סיגנלים), ובסופו של דבר צריכים להחליט על מספר מצומצם של פרסומות להציג לנו. לכן המידע כולו צריך להשתקלל להחלטה בודדת ותהליך הזיקוק הזה הוא סטטיסטי. המכונה הסטטיסטית מקבלת מידע עצום כקלט, ומסוגלת לפלוט את הפרמטרים החשובים לאחר עבודה מאומצת.

הסטטיסטיקה לא נמצאת רק העולם הוירטואלי, אלא בעלת משמעויות אדירות בעולם הפיזי, ובכדי להמחיש זאת נסתכל על אחד התחומים הכי פיזיים כיום - ספורט. כיום, כמעט **בכל קבוצה, בכל ליגה, בכל ענף ספורט, יש שימוש אינטנסיבי בכלים סטטיסטיים** בכדי להעצים את השחקן והקבוצה. לפני שני עשורים ומעלה, הסטטיסטיקה על שחקני ליגת הכדורסל האמריקאית, ה-NBA, הייתה בסיסית ביותר: אחוזי קליעה ל-2 נקודות/3 נקודות/קו העונשין, חסימות, איבודים, חטיפות, וכן הלאה. בשנים האחרונות התחום הזה עבר מהפכה אדירה. כיום ישנה סטטיסטיקה על כל דבר ששחקן מבצע, הן ברמה האישית, והן ברמה הקבוצתית. כל שחקן מנוח בכל פרמטר אפשרי, ולכן יודעים לקבוע מה הנקודות הטובות ביותר מהן הוא צריך לזרוק עד לרמת הסנטימטר; איפה מומלץ לו לעמוד בהגנה ובאילו זמנים; כיצד שהותו על המגרש משפיעה על המשחק הקבוצתי במונחי הגנה/התקפה; והרשימה נמשכת עוד ועוד.

נבחן את הסרטון החביב מהסרט Moneyball (<http://www.math.harvard.edu/~knill/mathmovies/swf/statisticsmoneyball.html>). גיוס שחקנים משלב בתוכו אינספור אלמנטים (איכות השחקן, פוטנציאל, מחיר, תרומה לקבוצה וכד'), והמטרה העיקרית היא מציאת אותו שחקן אנונימי, אך משמעותי, אשר יהפוך לדבר הגדול הבא. באנלוגיה לשווקים הפיננסיים, השקעה בשחקן דומה להשקעה בחברה בע"מ אשר יכולה להניב דיבידנדים גדולים בעתיד, ומנגד יכולה להתרסק. צוותי הגיוס בכל ענף עובדים עם סטטיסטיקאים צמודים אשר מנתחים כל נתון אפשרי לאורך הקריירה של השחקנים הצעירים והבוגרים ומסייעים בקבלת החלטות הגיוס.

נמשיך לעולם הכלכלי. **ההשפעה של סטטיסטיקה על העולם הכלכלי היא משמעותית, בטח ובטח לשיקולי מדיניות**. ניקח לדוגמא החלטות שנדונות באופן רציף במדינתנו ובעולם: פיקוח על שכר דירה והשפעתו בטווח הקצר והטווח הארוך על הדירות והשוק; שינויים בשכר המינימום והשפעתו (קצרה וארוכת הטווח); הורדה או העלאה של מס חברות והשפעתו על הצמיחה והאבטלה; השפעה של לימודים אקדמיים על הכנסה וכד'. מצד אחד, ישנה תיאוריה בסיסית אשר מניבה לנו לא מעט ניבויים להמשך. מצד שני, הניבויים הללו עומדים לבחינה באמצעות ניתוח

השפעות בפועל בכלים סטטיסטיים. דוגמא לא פשוטה היא שינויים בשכר המינימום. כיום המחקר האמפירי לא מסוגל לקבוע בצורה חד משמעית כיצד העלאת שכר המינימום תשפיע על התעסוקה. זהו ויכוח ארוך שנים והתוצאות הסטטיסטיות משתנות בכל עת. כמובן שהעלאה מאוד גדלה או מאוד קטנה הן זניחות, אבל טווח הביניים הוא רחב וההשפעות מגוונות.

6.2 עיבוד נתונים

קיבוץ הנתונים זה החלק הראשון בעבודה הסטטיסטית, ומיד לאחר מכן יש צורך בעיבוד הנתונים ובהצגתם בצורה קלה להבנה וברורה, אשר תאפשר ניתוח והסקת מסקנות. להמחשה נחשוב על הדוגמא הבאה.

דוגמה 6.1 הטמפרטורות שנמדדו בחודש מאי בשעה 14:00 בבאר שבע מוצגות בטבלה 7.

יום בחודש	טמפרטורה (צלזיוס)	יום בחודש	טמפרטורה (צלזיוס)
1	30	17	34
2	30	18	31
3	29	19	30
4	27	20	30
5	28	21	30
6	29	22	31
7	30	23	30
8	32	24	29
9	30	25	30
10	26	26	30
11	25	27	29
12	26	28	28
13	26	29	29
14	27	30	29
15	27	31	27
16	33		

טבלה 7: הטמפרטורות שנמדדו בחודש מאי בבאר שבע בשעה 14:00

לאחר קיבוץ הנתונים בטבלה אחת, נרצה לעבור לשלב העיבוד. הבעיה המרכזית שתעמוד בפנינו היא התמודדות עם מספר גדול של נתונים. לכן נידרש למספר סימונים ופישוטים טכניים.

6.2.1 סכימה

תחילה נגדיר משתנה x_i בתור הטמפרטורה ביום ה- i כאשר i הוא אינדקס במציין את היום בחודש, קרי אינדקס שנע מ-1 עד 31. לדוגמא, $x_1 = 30$ ו- $x_6 = 29$. בצורה כזאת יהיה לנו נוח להתייחס לכל נתון בנפרד בצורה קצרה ומדויקת ובכך להימנע מרישום מסורבל ולא מדויק. הדרך הנוחה לחשוב על המשתנים הללו היא באמצעות סדרות. אם הזדמן לכם ללמוד סדרות בעבר, אז אתם יודעים שלכל איבר בסדרה יש ערך כלשהו. במידה ואינכם מכירים סדרות כלל, אז הדבר הטוב ביותר הוא לחשוב על תוכנת אקסל. בתוכנת אקסל יש תאים ובכל תא ניתן להזין מספרים. זאת בעצם סדרה. סדרה היא רשימה של מספרים ולכל מספר יש אינדקס שמאפשר לנו להתייחס אליו ישירות.

לדוגמא, נסתכל על הסדרה

$$\begin{aligned}x_1 &= 1, \\x_2 &= 1, \\x_3 &= 2, \\x_4 &= 3, \\x_5 &= 5,\end{aligned}$$

ובאופן כללי, $x_i = x_{i-1} + x_{i-2}$. במילים, האיבר ה- i הוא סכום של צמד האיברים הקודמים לו x_{i-1} ו- x_{i-2} . הסדרה הזאת נקראת סדרת פיבונצ'י. הסדרה מייצגת רשימה של ערכים כלשהם ובעת הניתוח הסטטיסטי נפגוש הרבה רשימות מספרים שכאלה. בחלק מן המקרים הרשימות יהיו יותר מורכבות (טבלאות עם הרבה עמודות, תאים, ונתונים), אך כרגע אנחנו מתייחסים לסדרות פשוטות של מספרים. הסדרה לא חייבת להתחיל מהאינדקס $i = 1$, ובאותה המידה היינו יכולים לרשום את סדרת פיבונצ'י לעיל מהאיבר השלישי באופן הבא:

$$\begin{aligned}y_3 &= 1, \\y_4 &= 1, \\y_5 &= 2, \\y_6 &= 3, \\y_7 &= 5,\end{aligned}$$

וכן הלאה. זאת אומרת, האינדקסים הם פשוט דרך לסמן את האיברים. לערכם של האינדקסים אין באמת משמעות נשגבת. זה פשוט השם של אותו איבר. בסדרה הראשונה, לאיבר הראשון קראו x_1 והוא היה שווה ל-1, ובסדרה השנייה לאיבר הראשון קראו y_3 והוא היה שווה ל-1. הסימון הזה נוח מאוד כאשר עובדים עם נתונים רבים, שכן הרישום של "האיבר הראשון בסדרה ועוד האיבר השלישי בסדרה" הוא מסורבל, בעוד $x_1 + x_3$ זה סימון קל ופשוט.

קעת אנחנו נרצה לבצע פעולות אלגבריות על הנתונים - חיבור, חיסור, כפל, חילוק וכן הלאה. בעבודה עם מספר גדול של גורמים יש צורך להיות מדויקים וקצרים (במיוחד כאשר משתמשים בתוכנות מחשוב). לשם כך נגדיר סימונים ופעולות מתמטיות לטיפול במספר גדול של איברים.

נניח ואנו מעוניינים לחשב את הטמפרטורה הממוצעת בחודש מאי בשעה 14:00 בב"ש. לשם כך עלינו לסכום את כל הנתונים ולחלק במספר הימים, זהו **ממוצע פשוט**. חישוב הממוצע מתבצע על פי הנוסחה $\frac{x_1 + x_2 + \dots + x_{30} + x_{31}}{31} = 29.096$. בכדי לפשט את הסימון אנחנו נוהגים לסמן סכומים באות היוונית $\sum$, סיגמא. לדוגמא, את הסכום הנ"ל נסמן ב-

$$\frac{1}{31} \sum_{i=1}^{31} x_i = 29.096.$$

הסימון מציג את הסכימה של 31 האיברים, כאשר האינדקס i מציין את האיבר עליו סוכמים. במקרה הזה, i נע מ-1 ועד 31, וזה אומר שאנחנו סוכמים על כל ערכי ה- x_i מהערך הראשון ועד ה-31. באופן כללי, $\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$. לדוגמא, ממוצע הטמפרטורות בשבוע הראשון של מאי בבאר שבע, כפי שנמדדו בטבלה 7, הוא

$$\frac{1}{7} \sum_{i=1}^7 x_i = \frac{1}{7} [30 + 30 + 29 + 27 + 28 + 29 + 30] = 29.$$

כללי הסכימה הם פשוטים ביותר ונובעים ישירות מכללי חיבור וחסור שאתם מכירים:

$$\begin{aligned}\sum_{i=1}^n a &= a + a + a + \cdots + a = na, \\ \sum_{i=1}^n ax_i &= ax_1 + ax_2 + \cdots + ax_n = a[x_1 + x_2 + \cdots + x_n] = a \sum_{i=1}^n x_i, \\ \sum_{i=1}^n (x_i + y_i) &= (x_1 + y_1) + (x_2 + y_2) + \cdots + (x_n + y_n) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i.\end{aligned}$$

השורה הראשונה מייצגת את הסכום של n איברים זהים השווים כולם ל- a , ולכן סכומם הוא $a \cdot n$, בעוד השורה השנייה מייצגת סכום כללי כאשר כל המחוברים מוכפלים ב- a , ולכן ניתן להוציא את a כגורם משותף מן הסכום. בשורה השלישית לקחנו צמד סדרות וסכמנו את האיברים שלהן במקביל. הסכום המתקבל שווה לסכום של כל סדרה בנפרד.

תרגיל 6.2 נתונה סדרה $x_i = 2i$ לכל מספר שלם $i = 1, 2, \dots$. חשבו את הביטויים הבאים: $\sum_{i=1}^5 x_i$, הממוצע של ארבעת האיברים הראשונים בסדרה, $\sum_{i=1}^5 \frac{1}{2}x_i$.

פתרון: נענה על כל סעיף בנפרד.

$$\begin{aligned}\sum_{i=1}^5 x_i &= \sum_{i=1}^5 2i = 2 \sum_{i=1}^5 i = 2[1 + 2 + 3 + 4 + 5] = 2 \cdot 15 = 30, \\ \sum_{i=1}^5 \frac{1}{2}x_i &= \frac{1}{2} \sum_{i=1}^5 x_i = \frac{1}{2} \cdot 30 = 15, \\ \frac{1}{4} \sum_{i=1}^4 x_i &= \frac{1}{4} \left[\sum_{i=1}^5 x_i - x_5 \right] = \frac{1}{4} [30 - 10] = \frac{20}{4} = 5.\end{aligned}$$

תרגיל 6.3 נתונה סדרה $x_i = \frac{1}{4}i + 4$ לכל מספר שלם $i = 3, 4, 5, \dots$. חשבו את הביטויים הבאים: $\sum_{i=6}^{10} x_i$, $\sum_{i=7}^{11} x_i$.

פתרון: נענה על כל סעיף בנפרד.

$$\begin{aligned}\sum_{i=6}^{10} x_i &= \sum_{i=6}^{10} \left(\frac{1}{4}i + 4 \right) = \sum_{i=6}^{10} \frac{1}{4}i + \sum_{i=6}^{10} 4 = \frac{1}{4} \sum_{i=6}^{10} i + 4 \cdot 5 = \frac{1}{4} [6 + 7 + 8 + 9 + 10] + 4 \cdot 5 = 30, \\ \sum_{i=7}^{11} x_i &= \sum_{i=7}^{11} \left(\frac{1}{4}i + 4 \right) = \sum_{i=7}^{11} \frac{1}{4}i + \sum_{i=7}^{11} 4 = \frac{1}{4} \sum_{i=7}^{11} i + 4 \cdot 5 = \frac{1}{4} [7 + 8 + 9 + 10 + 11] + 4 \cdot 5 = \frac{125}{4}.\end{aligned}$$

6.2.2 טרנספורמציה לינארית

כעת נניח ופונים אלינו מארה"ב ומבקשים שנשלח אליהם את הנתונים. כידוע, האמריקאים עובדים ביחידות פרנהייט ולכן נצטרך להמיר את כל הערכים בטבלה. כיצד נבצע זאת? הנוסחה להמרה בין צלזיוס C לפרנהייט F היא $F = \frac{9}{5}C + 32$. הנוסחה ההפוכה היא כמובן $C = \frac{5(F-32)}{9}$. הנוסחאות הללו מציגות **טרנספורמציה לינארית**, קרי לוקחים נתונים כלשהם ומעבירים אותם באמצעות כפל בקבוע וחיבור של קבוע לסקאלה אחרת.

פורמאלית, ניקח נתון x ונמיר אותו בעזרת טרנספורמציה לינארית $y = ax + b$ לערך y , כאשר a, b קבועים. כיצד שינוי זה משפיע על הסכום או על הממוצע? קל לבצע את החישוב ולראות. נסמן $y_i = ax_i + b$ לכל איבר i בסדרה.

$$\begin{aligned}\sum_{i=1}^n y_i &= \sum_{i=1}^n (ax_i + b) = \sum_{i=1}^n ax_i + \sum_{i=1}^n b = a \left[\sum_{i=1}^n x_i \right] + b \cdot n, \\ \frac{1}{n} \sum_{i=1}^n y_i &= \frac{1}{n} \sum_{i=1}^n (ax_i + b) = \frac{1}{n} \sum_{i=1}^n ax_i + \frac{1}{n} \sum_{i=1}^n b \\ &= \frac{a}{n} \left[\sum_{i=1}^n x_i \right] + \frac{b \cdot n}{n} = \frac{a}{n} \left[\sum_{i=1}^n x_i \right] + b = a \left[\frac{1}{n} \sum_{i=1}^n x_i \right] + b,\end{aligned}$$

זאת אומרת, הממוצע זה בדיוק באופן בו השתנה כל גורם בסכום - הממוצע הוכפל ב- a ונוסף לו האיבר b .

תרגיל 6.4 חזרו על החישובים בתרגיל 6.2 עבור הטרנספורמציה $y_i = \frac{1}{4}x_i + 1$.

פתרון: הפתרון הוא מיידי.

$$\begin{aligned}\sum_{i=1}^5 y_i &= \sum_{i=1}^5 \left(\frac{1}{4}x_i + 1 \right) = \sum_{i=1}^5 \frac{1}{4}x_i + \sum_{i=1}^5 1 = \frac{1}{4} \sum_{i=1}^5 x_i + 5 = \frac{30}{4} + 5 = 12.5, \\ \sum_{i=1}^5 \frac{1}{2} y_i &= \frac{1}{2} \sum_{i=1}^5 y_i = 6.25, \\ \frac{1}{4} \sum_{i=1}^4 y_i &= \frac{1}{4} \sum_{i=1}^4 \left(\frac{1}{4}x_i + 1 \right) = \frac{1}{4} \cdot \frac{1}{4} \sum_{i=1}^4 x_i + 1 = \frac{5}{4} + 1 = \frac{9}{4}.\end{aligned}$$

שימו לב כי כל התוצאות לעיל עיקביות עם החישובים הכללים שביצענו קודם לכן.

6.2.3 טבלת שכיחויות

נרצה לסכם את הנתונים בצורה נוחה, ולשם כך נשתמש בטבלת שכיחויות. **השכיחות של ערך מסוים** היא מספר הפעמים שאותו הערך מופיע בנתונים. לדוגמא, השכיחות של הערך 25 היא 1 כי $x_{11} = 25$, והשכיחות של הערך 31 היא 2 כי $x_{18} = x_{22} = 31$. נסמן את השכיחות של ערך x בתור $f(x)$ (frequency). ראו עמודה ראשונה ושיייה בטבלה 8.

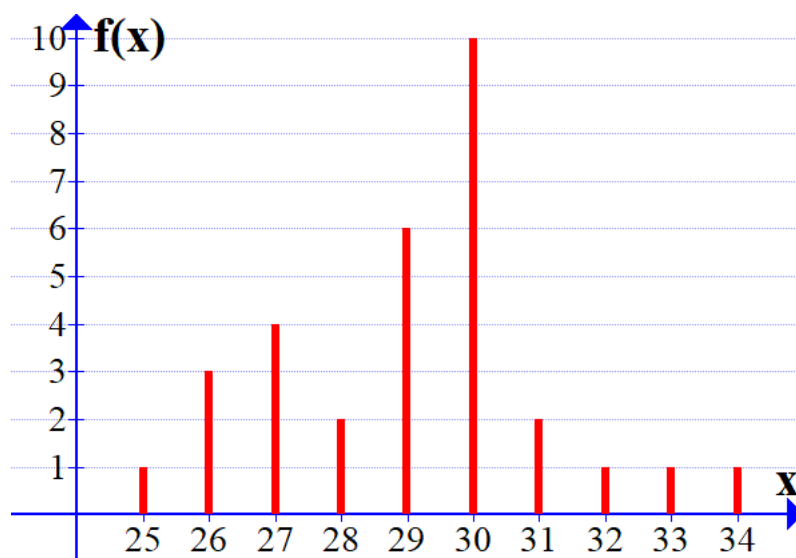
לאחר מציאת השכיחות של כל ערך, נגדיר את **השכיחות המצטברת** בתור השכיחות עד לנקודה מסוימת. לדוגמא, $f(25) = 1$, $f(26) = 3$, $f(27) = 4$, לכן השכיחות המצטברת של הערך 27 היא 8, ונסמן זאת ב- $F(27) = 8$. בנוסף, השכיחות המצטברת של הערך 28 היא $F(28) = 10$ בתור הסכום של $f(25) + f(26) + f(27) + f(28) = 10$. נרשום את כלל הנתונים בעמודה השלישית בטבלה 8. לאחר חישוב הגדלים המקוריים, נחשב את **הגדלים היחסיים**. הגדלים היחסיים דומים לגדלים המקוריים שרשמנו בעמודות הראשונות, עם תוספת של חלוקה במספר הימים הכולל (כמות הנתונים הכוללת). סה"כ דגמנו 31 ימים שונים ולכן ניקח את הגדלים הקודמים, הן את השכיחות והן את השכיחות המצטברת, ונחלק בגודל הדגימה הכללי. נרשום את התוצאות בעמודות האחרונות.

6.2.4 הצגת הנתונים

ישנן מספר דרכים להציג את הנתונים שקיבלנו. הדרך הראשונה היא **דיאגרמת מקלות** כאשר ציר x הוא המשתנה (הטמפרטורה) וציר y הוא ציר שכיחויות $f(x)$. גובה המקלות מציין את השכיחות של כל ערך בנפרד. נציג דיאגרמת מקלות באיור 21.

טמפרטורה x	שכיחות $f(x)$	שכיחות מצטברת $F(x)$	שכיחות יחסית $\frac{f(x)}{n}$	שכיחות מצטברת יחסית $\frac{F(x)}{n}$
25	1	1	(3.2%) $\frac{1}{31}$	(3.2%) $\frac{1}{31}$
26	3	4	(9.6%) $\frac{3}{31}$	(12.9%) $\frac{4}{31}$
27	4	8	(12.9%) $\frac{4}{31}$	(25.8%) $\frac{8}{31}$
28	2	10	(6.5%) $\frac{2}{31}$	(32.3%) $\frac{10}{31}$
29	6	16	(19.6%) $\frac{6}{31}$	(51.6%) $\frac{16}{31}$
30	10	26	(32.2%) $\frac{10}{31}$	(83.8%) $\frac{26}{31}$
31	2	28	(6.5%) $\frac{2}{31}$	(90%) $\frac{28}{31}$
32	1	29	(3.2%) $\frac{1}{31}$	(93.5%) $\frac{29}{31}$
33	1	30	(3.2%) $\frac{1}{31}$	(96.8%) $\frac{30}{31}$
34	1	31	(3.2%) $\frac{1}{31}$	(100%) $\frac{31}{31}$

טבלה 8: טבלת שכיחויות עבור נתוני טמפרטורה בבאר שבע

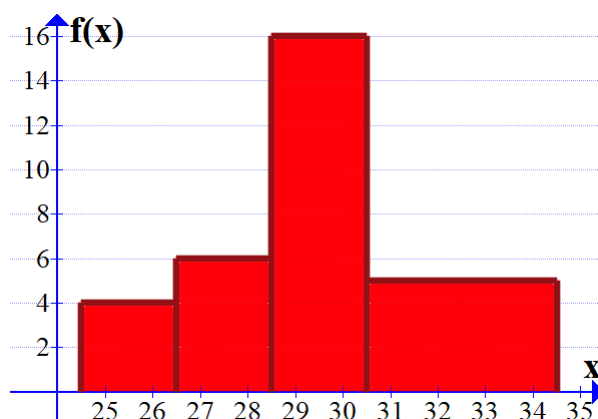


איור 21: דיאגרמת מקלות לנתוני הטמפרטורה

דרך אחרת להציג את אותם הנתונים היא באמצעות היסטוגרמה. **ההיסטוגרמה** היא גרף המייצג את השכיחות של טווח ערכים. טווח הערכים מסומן על ציר ה- x וגובה המלבן מייצג את השכיחות של אותם הימים. שימו לב שטווחי הערכים לא חייבים להיות אחידים. זאת אומרת, טווח ערכים אחד יכול להיות יחסית גדול בעוד השני יחסית קטן. ראו את טבלה 9 המציגה את טווחי הערכים והשכיחויות הרלוונטיות ואת איור 22 עבור ההיסטוגרמה.

שכיחות $f(x)$	טמפרטורה x
4	24.5-26.5
6	26.5-28.5
16	28.5-30.5
5	30.5-34.5

טבלה 9: טבלת שכיחויות עבור טווחי טמפרטורה בבאר שבע



איור 22: היסטוגרמה לטווחי הטמפרטורה

6.3 מדדי מרכז ופיזור

6.3.1 מדדי מרכז

המטרה הראשונית בניתוח סדרת הנתונים היא יצירת מדד כללי שמשקף את הערך של כלל התצפיות. המדד הוא מעין אינדיקטור אשר מתמצת את כלל הנתונים לכדי מספר בודד ומטרתו לשקף את ערכם. כמובן שמספר בודד לא יכול להכיל את כל האינפורמציה הנתונה, לכן חלק מן המידע "יילך לאיבוד" בתהליך העיבוד. מדדי מרכז באים לשקף את הערך המרכזי או הקריטי של סדרת הנתונים. כמובן שאין ערך יחיד שכזה, ובהתאם ישנם לא מעט מדדי מרכז. נסקור את החשובים:

- **שכיח (Mode)** - הערך בעל השכיחות הגבוהה ביותר. בדוגמא עימה אנו עובדים, הערך השכיח הוא 30 שכן השכיחות שלו היא 16, הגבוהה מכולן.

- **אמצע הטווח (Midrange)** - ממוצע שני הערכים הקיצוניים, הגדול ביותר והקטן ביותר. בדוגמא שלנו, אמצע הטווח הוא $\frac{34+25}{2} = 29.5$.

- **חציון (Median)** - הערך שמחלק את התצפיות לשתי קבוצות כך שלפחות מחצית מהתצפיות מקבלות ערך זה ומטה, ולפחות מחצית מקבלות ערך זה ומעלה. בדוגמא שלנו, החציון הוא 29. מספר התצפיות המקבלות 29 ומטה הוא 16, ומספר התצפיות המקבלות ערך זה ומעלה הן 21. שימו לב ש-30 לא יוכל להיות החציון שכן מספר התצפיות שמקבלות ערך זה ומעלה הוא 15 וזה פחות ממחצית התצפיות.

מה קורה אם ישנם שני ערכים אשר עונים להגדרה של חציון? ייתכן מצב בו יש צמד ערכים שעונים להגדרה של חציון ולא ערך בודד.

לדוגמא, ניקח את המספרים מ-1 עד 16 בתור הדגימות שלנו, ז"א, סדרת הנתונים שלנו היא כל המספרים השלמים 1, 2, 3, ..., 15, 16. החציון יכול להיות המספר 8, שכן מספר התצפיות מ-8 (כולל) ומטה הוא, מן הסתם, 8 (התצפיות הרלוונטיות הן 1, 2, 3, 4, 5, 6, 7, 8), ומספר התצפיות מ-8 ומעלה הוא 9 (התצפיות הרלוונטיות הן 8, 9, 10, 11, 12, 13, 14, 15, 16). לכן הערך 8 עונה להגדרה של חציון. אבל אותו הטיעון נכון עבור הערך 9: התצפיות מן הערך 9 ומטה הוא 9 (תצפיות רלוונטיות - 1, 2, 3, 4, 5, 6, 7, 8, 9), ומספר התצפיות מ-9 ומעלה הוא 8 (תצפיות רלוונטיות - 9, 10, 11, 12, 13, 14, 15, 16). במקרה כזה לוקחים את הממוצע בין צמד הערכים האפשריים. ז"א, נקבע את החציון להיות 8.5 בתור הממוצע של 8 ו-9.

• **ממוצע (Average)** - סכימה של כל התצפיות וחלוקה במספר התצפיות. במקרה שלנו הממוצע הוא $\bar{x} = \frac{1}{31} \sum_{i=1}^{31} x_i = 29.096$.

כדאי לשים לב שרוב המדדים הניבוי תוצאה דומה, אבל זה לא המצב באופן כללי. ייתכנו מצבים בהם הערך הממוצע והחציון, לדוגמא, יהיו שונים מאוד זה מזה. הדוגמא הכלכלית הפשוטה ביותר היא נתוני הכנסה. ישנו פער יחסית גדול בין השכר הממוצע לשכר החציוני במדינת ישראל, לטובת השכר הממוצע. זה מעיד על מצב בו יש הרבה אנשים אשר מרוויחים שכר יחסית נמוך, בעוד אחרים מקבלים שכר די גבוה.

6.3.2 מדדי פיזור

מדדי פיזור הם מדדים משלימים למדדי מרכז. מדדי מרכז משקפים את הערך או המיקום של התצפיות, בעוד מדדי פיזור מנסים להציג את המידה בה התצפיות מפוזרות על פני הערכים השונים. ישנם מספר מדדי פיזור חשובים אותם נסקור:

• **טווח (Range)** - ערך התצפית המקסימאלי פחות ערך התצפית המינימאלי. בדוגמא לעיל, הטווח הוא $34 - 25 = 9$.

• **טווח בין-רבעוני (Interquartile range)** - הטווח בין הרבעון הראשון לרבעון השלישי כך שלפחות מחצית מן התצפיות נמצאות בטווח זה. חישוב הטווח הבין-רבעוני נפתח בחישוב החציון. לאחר מציאת החציון מחלקים את התצפיות לשני חלקים: התצפיות הקטנות מן החציון והתצפיות הגדולות מן החציון. שימו לב, התצפית החציונית (במידה וקיימת, כאשר מספר התצפיות אי זוגי) איננה בקבוצת ה"ל". עבור כל קבוצה בנפרד מחשבים את החציון, הן עבור התצפיות הנמוכות והן עבור התצפיות הגבוהות. הטווח הבין-רבעוני מוגדר כהפרש הערכים שהתקבלו.

נדגים את החישוב בעזרת ערכי הטמפרטורה בב"ש מן הטבלה לעיל. החציון הוא 29. ניקח את קבוצת התצפיות שקטנה מן החציון, המכילה את הערכים 25, 26, 27, 28, 29, ומצד שני ניקח את קבוצת התצפיות מעל החציון אשר מכילה את הערכים 30, 31, 32, 33, 34. נדגיש כי קבוצת הערכים הגבוהים לא מכילה תצפית בעלת ערך 29, שכן החציון מתקבל בדיוק בתצפית ה-16 ($x_{16} = 29$), והתצפית ה-17 היא בעלת ערך 30 ($x_{17} = 30$). החציון בקבוצת הערכים הנמוכים הוא 27, בעוד החציון בקבוצת הערכים הגבוהים הוא 30. לכן הטווח הבין-רבעוני הוא $30 - 27 = 3$.¹⁵

• **שונות (Variance)** - הטווחים לעיל לא לוקחים בחשבון את כל אחת מן התצפיות בנפרד, לכן התוצאה המתקבלת היא יחסית גסה. מדד השונות, לעומת זאת, מנסה למדוד את הפיזור של התצפיות סביב הממוצע. חישוב השונות מתחיל בחישוב הממוצע. כאמור, בדוגמא לעיל הערך הזה הוא 29.096 ונסמן אותו ב- $\bar{x} = 29.096$. כעת נבחן את המרחק הריבועי של כל תצפית מן הממוצע ונסכום את כלל המרחקים הללו, $\sum_{i=1}^{31} (x_i - \bar{x})^2$. נחלק את הגודל ה"ל" במספר התצפיות ונקבל את השונות,

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

זאת ההגדרה לשונות, הנקראת גם Variance, מסומנת ב- S_x^2 או ב- Var . השונות היא המרחק הריבועי הממוצע של התצפיות מהערך הממוצע. זאת אומרת, אנחנו בודקים מה המרחק הריבועי של כל תצפית מן הממוצע, מה שנותן לנו אמת מידה לפיזור של הערכים סביב הממוצע, ואז אנחנו ממצעים על הערכים שקיבלנו. בניגוד למדדים הקודמים, השונות מושפעת מערכי קיצון וההשפעה גדלה ככל שערכי הקיצון יותר רחוקים מן הממוצע.

השונות היא בעלת יחידות ריבועיות, שכן אנחנו לוקחים מרחק ריבועי ממוצע. לכן נהוג להוציא שורש לביטוי לעיל, ובכך מגדירים את

$$\text{סטיית התקן, } S_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

¹⁵חשוב לציין שישנן מספר הגדרות עבור הטווח הבין-רבעוני, ולכן בחלק מספרי הלימוד ייתכן ותראו הגדרות וחישובים שונים.

הערה 6.5 (העשרה - חישוב רבעונים) למרות שאין לנו צורך בחישוב רבעונים לטובת מדדי מרכז או פיזור, ייתכן וחלקכם מעוניינים לדעת כיצד מתבצע החישוב. ובכן, **הרבעון הראשון** הוא הערך שלפחות רבע מן התצפיות שוות לו או קטנות ממנו, ולפחות שלוש רבעי מהתצפיות שוות לו או גדולות ממנו. **הרבעון השלישי** הוא הערך שלפחות רבע מן התצפיות גדולות או שוות לו, בעוד שלוש רבעי מהתצפיות שוות לו או קטנות ממנו. בדוגמא לעיל הרבעון הראשון הוא 27 (מספר התצפיות מ-27 ומטה הן 8, ומספר התצפיות מ-27 ומעלה הן 27). הרבעון השלישי הוא 30 (מספר התצפיות מ-30 ומעלה הוא 15, ומספר התצפיות מ-30 ומטה הוא 26). לכן הטווח הבין-רבעוני הוא $30 - 27 = 3$. קל לראות שטווח זה, כולל ערכי הקצוות, מכיל 22 תצפיות.

שימו לב! גם בחישוב הרבעונים יכולים להתקבל שני ערכים העונים להגדרה של רבעון ראשון או רבעון שלישי. לדוגמא, אם התצפיות שלנו היו 1, 2, 3, 4, 5, 6, 7, 8, אז הערך 2 והערך 3 עונים להגדרה של רבעון ראשון, שכן מספר התצפיות מ-2 ומטה הוא 2 ומספר התצפיות מ-2 ומעלה הוא 7 (עונה להגדרה של לפחות רבע ולפחות שלוש רבעי מכלל התצפיות). בעוד מספר התצפיות מן הערך 3 ומטה הוא 3 וממנו ומעלה הוא 6, ולכן גם עונה על דרישות הרבעון הראשון. שוב, בכדי לפתור את הבעיה ניקח את הממוצע בין צמד הערכים האפשריים בתור הרבעון המתאים.

תרגיל 6.6 נתונה סדרות הנתונים הבאות, חשבו את החציון, הרבעון הראשון, הרבעון השלישי, ואת הטווח הבין-רבעוני של כל סדרה בנפרד.

Series #1 = 2, 4, 4, 5, 6, 7, 8, 9.

Series #2 = 2, 3, 4, 4, 5, 6, 7, 8, 9.

פתרון: נתחיל עם הסדרה הראשונה. ישנם סה"כ 8 איברים בסדרה. הן הערך 5 והן הערך 6 עונים להגדרה של החציון ולכן ניקח את הממוצע ביניהם. משמע, החציון הוא 5.5. הרבעון הראשון מתקבל בערך 4 והרבעון השלישי מתקבל בערך 7 או בערך 8 (בשני המצבים ההגדרה מתקיימת). לכן ניקח את הממוצע ביניהם ונקבל 7.5. הטווח הבין-רבעוני מתקבל על ידי החציון של הערכים הנמוכים והחציון של הערכים הגבוהים. החציון של 2, 4, 4, 5 הוא 4, בעוד החציון של 6, 7, 8, 9 הוא 7.5. לכן הטווח הבין-רבעוני יהיה $7.5 - 4 = 3.5$. נעבור לסדרה השנייה המכילה 9 דגימות. החציון מתקבל בערך המרכזי, 5. הרבעון הראשון מתקבל בערך 4, הרבעון השלישי מתקבל בערך 7. הטווח הבין-רבעוני של מספר אי-זוגי של איברים מחושב בעזרת השמטה של החציון, 5. במילים אחרות, נסתכל על הסדרה החדשה 2, 3, 4, 4, 6, 7, 8, 9 ונחשב עבורה טווח בין-רבעוני. החציון של מחצית הערכים הנמוכים הוא 3.5, בעוד החציון של מחצית הערכים הגבוהים הוא 7.5, לכן הטווח הבין-רבעוני הוא $7.5 - 3.5 = 4$.

תרגיל 6.7 כיצד משפיעה הטראנספורמציה הליניארית $y_i = ax_i + b$ על הממוצע, על השונות, ועל סטיית התקן.

פתרון: נחשב את הגדלים באופן ישיר.

$$\begin{aligned}\frac{1}{n} \sum_{i=1}^n y_i &= \frac{1}{n} \sum_{i=1}^n (ax_i + b) = \frac{1}{n} \left[\sum_{i=1}^n ax_i + \sum_{i=1}^n b \right] = \frac{1}{n} \left[a \sum_{i=1}^n x_i + n \cdot b \right] = a \frac{1}{n} \sum_{i=1}^n x_i + \frac{n \cdot b}{n} \\ &= a\bar{x} + b = \bar{y}, \\ \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 &= \frac{1}{n} \sum_{i=1}^n (ax_i + b - (a\bar{x} + b))^2 = \frac{1}{n} \sum_{i=1}^n (ax_i + b - a\bar{x} - b)^2 = \frac{1}{n} \sum_{i=1}^n (ax_i - a\bar{x})^2 \\ &= \frac{1}{n} \sum_{i=1}^n a^2 (x_i - \bar{x})^2 = a^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = a^2 S_x^2, \\ S_y &= \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} = \sqrt{a^2 S_x^2} = a S_x,\end{aligned}$$

במילים, הממוצע משתנה בדיוק ע"פ הטרנספורמציה $\bar{y} = a\bar{x} + b$, השונות מוכפלת בקבוע הכפלי a^2 כך ש- $S_y^2 = a^2 S_x^2$, סטיית התקן מוכפלת ב- a כך ש- $S_y = a S_x$.

6.3.3 מדד מיקום יחסי

המדד האחרון בו נעסוק בחלק זה יהיה מדד המיקום היחסי. מדד המיקום היחסי הוא שילוב של מדד המרכז ומדד הפיזור והוא מתייחס לכל תצפית בנפרד. ניקח תצפית x_i כאשר הממוצע של כלל התצפיות הוא $\bar{x}$ וסטיית התקן היא S . נגדיר את מדד מיקום היחסי של x_i על ידי

$$z_i = \frac{x_i - \bar{x}}{S_x}.$$

זאת אומרת, מדד המיקום היחסי של x_i הוא ההפרש בין x_i לממוצע, בחלוקה בסטיית התקן. המדד הזה מניב לנו את המרחק של התצפית מן הממוצע כאשר היחידות בהן אנו מודדים הן סטיות תקן. קרי, סטיית התקן זו הסקאלה שאנחנו מודדים לפיה. מדוע זה טוב? בלא מעט מקרים נרצה להשוות גדלים שאינם מגיעים מאותו התחום או תצפיות, כגון - טמפרטורה בצלזיוס לעומת טמפרטורה בפרנהייט; ציונים בספרות לעומת ציונים בסטטיסטיקה; גבהים של אנשים לעומת משקלים של אנשים וכן הלאה. מדד המיקום היחסי מייצר סקאלה אחידה, ולכן אם אדם נמצא מרחק סטיית תקן אחת מעל הממוצע או סטיית תקן אחת מתחת לממוצע, אנחנו מבינים את המשמעות של הנתון, בלא צורך לרדת לפרטים של אופן המדידה או הסקאלה שלפיה עובדים. הטרנספורמציה ההפוכה היא

$$x_i = \bar{x} + z_i \cdot S_x.$$

תרגיל 6.8 נתונה סדרת נתונים כך ש- $\bar{x} = 1200$ ו- $S_x = 600$. מצאו את מדד המיקום היחסי של $x_i = 1000, 1400, 1800$ ומצאו את x_i כך ש- $z_i = 1, -1, 0.5$.

פתרון: נבצע חישוב ישיר.

$$\begin{aligned} x_i = 1000 &\Rightarrow z_i = \frac{1000 - 1200}{600} = \frac{-200}{600} = -\frac{1}{3}, \\ x_i = 1400 &\Rightarrow z_i = \frac{1400 - 1200}{600} = \frac{200}{600} = \frac{1}{3}, \\ x_i = 1800 &\Rightarrow z_i = \frac{1800 - 1200}{600} = \frac{600}{600} = 1, \\ z_i = 1 &\Rightarrow x_i = 1200 + 1 \cdot 600 = 1800, \\ z_i = -1 &\Rightarrow x_i = 1200 + (-1) \cdot 600 = 600, \\ z_i = 0.5 &\Rightarrow x_i = 1200 + 0.5 \cdot 600 = 1500. \end{aligned}$$

בתוצאות שקיבלנו רואים כיצד מדד מיקום יחסי של סטיית תקן אחת הוא למעשה סטיית תקן אחת מעל הממוצע, ובאופן עקבי דומה מדד מיקום יחסי של -1 הינו ערך המרוחק בסטיית תקן אחת מתחת לממוצע.