מבוא להסתברות וסטטיסטיקה 80430
תרגול 1

לאורך הקורס, תמיד נתכוון ל"זרים בזוגות" גם כאשר נאמר "זרים" בלבד.

1 מרחבי הסתברות

הגדרה 1. מרחב המדגם מסומן לרוב באות Ω, ואיבר המדגם באות ω. פונקציה $p:\Omega\to\mathbb{R}_+$ המקיימת $\sum_{\omega\in\Omega}p(w)=1$ נקראת פונקציית הסתברות נקודתית הקבוצה $\operatorname{Supp}(p)=\{\omega\in\Omega\mid p(\omega)>0\}$ נקראת התומך של p. תת-קבוצה של מרחב המדגם $A\subseteq\Omega$ תיקרא מאורע. אוסף כל המאורעות יסומן ב-$\mathcal{F}$. פונקצית הסתברות מוגדרת להיות פונקציה $\mathbb{P}:\mathcal{F}\to\mathbb{R}_+$ המקיימת:

- $\mathbb{P}(\Omega)=1$.
- $\mathbb{P}(\emptyset)=0$.
- לכל סדרת מאורעות זרים בזוגות $\{A_n\}_{n\in\mathbb{N}}$ מתקיים

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

$\mathbb{P}$ תיקרא פונקציית הסתברות על $(\Omega,\mathcal{F})$ נאמר גם ש- $\mathbb{P}$ נתמכת ע"י A אם $\mathbb{P}(A)=1$. בהינתן פונקציית הסתברות נקודתית p על Ω ניתן להגדיר פונקציית הסתברות על $\left(\Omega,2^{\Omega}\right)$, באופן הבא $\mathbb{P}_p(A):=\sum_{\omega\in A}p(\omega)$.
אם עבור מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$ קיימת פונקציית הסתברות נקודתית $p:\Omega\to\mathbb{R}_+$ כך ש $\mathbb{P}=\mathbb{P}_p$ אזי $\mathbb{P}$ תיקרא פונקציית הסתברות בדידה והמרחב ייקרא מרחב הסתברות בדידה.

דוגמה 2. מרחב המדגם $\Omega=\{1,2,...,6\}$ יחד עם פונקציית ההסתברות הנקודתית המקיימת $p(\omega)=\frac{1}{6}$ לכל $\omega\in\Omega$ יכול לייצג הטלה של קוביה תקנית. אם נגדיר את $\mathbb{P}_p$ על $\Omega=[6]$ כנ"ל, נקבל למשל כי ההסתברות לקבל מספר גדול מ-4 בהטלת קוביה הינו

$$\mathbb{P}_p(\{5,6\})=p(5)+p(6)=\frac{2}{6}=\frac{1}{3}.$$

מרחבי הסתברות אחידים - דוגמאות

הגדרה 3. תהא Ω קבוצה סופית ולא ריקה. תמיד ניתן להגדיר על Ω את פונקציית ההסתברות $\mathbb{P}$ המוגדרת ע"י $\mathbb{P}(A)=\frac{|A|}{|\Omega|}$ כאשר A מאורע, ו-$|A|$ מסמן את מספר האיברים של A. למרחב ההסתברות שנקבל באופן הזה $(\Omega,2^{\Omega},\mathbb{P})$ קוראים מרחב הסתברות אחיד, ול-$\mathbb{P}$ קוראים פונקציית הסתברות אחידה.

הערה. בשאלות הנוגעות למרחבי הסתברות אחידים האתגר הראשון הוא לייצג את הבעיה נכון ולזהות מה אחיד, אחר כך נשאר רק קושי קומבינטורי.

שימו לב למוסכמת הסימונים שלנו – אותיות גדולות $(\Omega, A, ...)$ תסמנה לרוב קבוצות, אותיות קטנות $(\omega, n, ...)$ תסמנה לרוב איברים בקבוצות, ואותיות מסולסלות $(\mathcal{F})$ תסמנה לרוב קבוצות של קבוצות.

דוגמאות

1. מניחים 4 כדורים שחורים ב-6 מגירות, כאשר בכל הנחה שמים כל כדור בהסתברות שווה בכל אחת מהמגירות. מה ההסתברות שבמגירה האחרונה אין כדורים?

פתרון (שגוי!): הכדורים זהים בצבעם (כולם שחורים) לפיכך אין חשיבות לסדר. על כן מרחב המדגם שלנו יהיה:

$$\Omega = \left\{ (x_1, ..., x_6) \mid \sum_{i=1}^{6} x_i = 4, \forall i \, x_i \geq 0 \right\},$$

ופונקציית ההסתברות הנקודתית תהיה ההסתברות האחידה $\forall \omega \in \Omega, p(\omega) = \frac{1}{|\Omega|}$. כעת אנו מעוניינים במאורע $A = \{(x_1, ..., x_6) \in \Omega \mid x_6 = 0\}$. ברור כי קבוצה זו בעלת אותו מספר איברים כמו

$$B = \left\{ (x_1, ..., x_5) \mid \sum_{i=1}^{5} x_i = 4, \forall i \, x_i \geq 0 \right\}$$

שכן נוכל להתאים בין איברי A ל-B על ידי $(x_1, ..., x_5) \mapsto (x_1, ..., x_5, 0)$. כעת יש רק לחשב את:

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{|B|}{|\Omega|} = \frac{\binom{4+4}{4}}{\binom{4+5}{5}} = \frac{\frac{8!}{4!4!}}{\frac{9!}{4!5!}} = \frac{5}{9} \approx 0.555.$$

מה הבעיה? **<u>הפתרון שגוי!</u>** הבעיה היא שדווקא יש חשיבות לסדר, כיוון שהכדורים נכנסים למגירות בזה אחר זה. אמנם הם כולם שחורים, אך ברגע שהכנסנו אותם בזה אחר זה נוכל להבדיל ביניהם – אנו יודעים מי הכדור הראשון שנכנס, השני שנכנס, וכו'...

על מנת להבין את ההבדל נחשוב על דוגמא פשוטה יותר – מכניסים 2 כדורים ל-2 מגירות בזה אחר זה. מה ההסתברות שבמגירה האחרונה אין כדורים?

<u>פתרון:</u> זה ייתכן רק אם שני הכדורים נכנסו למגירה הראשונה, מה שקורה בהסתברות $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$. לעומת זאת, אם ננסה לפתור בשיטה הקודמת נקבל כי

$$\Omega = \{(x_1, x_2) \mid x_1 + x_2 = 2\} = \{(2,0), (0,2), (1,1)\},$$

ולפיכך ההסתברות שהמגירה האחרונה ריקה הוא $\frac{1}{3}$.

בחזרה לבעיה המקורית – מרחב ההסתברות הנכון צריך להיות:

$$\Omega = \{(x_1, x_2, x_3, x_4) \mid \forall i, x_i \in [6]\} = [6]^4,$$
$$A = \{(x_1, x_2, x_3, x_4) \in \Omega \mid \forall i, x_i \neq 6\} = [5]^4.$$

לפיכך:

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{5^4}{6^4} = \left(\frac{5}{6}\right)^4 \approx 0.482.$$

הבעיה שלנו כאן היא שאיננו יודעים עדיין להצדיק מתי מותר להשתמש בכל אחד ממרחבי ההסתברות הללו. הסבר מנומק ופורמלי יותר ינתן בפרק מרחבי מכפלה.

2. נציג שתי בעיות:

(א) בוחרים באקראי סידור של 2 כדורים שחורים ו–6 כדורים לבנים בשורה. מה הסיכוי שהכדורים השחורים יהיו זה לצד זה?

(ב) בכד 2 כדורים שחורים ו–6 כדורים לבנים. שולפים בזה אחר זה את הכדורים ומסדרים אותם בשורה לפי סדר השליפה. מה הסיכוי ששני הכדורים השחורים יהיו זה לצד זה?

הנה דרך אחת לפתור את הבעיה הראשונה – מספר הסידורים הכללי של 2 כדורים שחורים ו-6 כדורים לבנים הוא $\binom{8}{2}$. מספר הסידורים בהם השחורים זה לצד זה הוא $7 = \binom{7}{1}$, שכן ניתן להשאיר את 6 הכדורים במקום, ורק לבחור היכן יהיה זוג השחורים (כולל הקצוות). לפיכך הסיכוי שיהיו זה לצד זה הינו $\frac{7}{\binom{8}{2}} = \frac{1}{4}$.

את הבעיה השנייה נוכל לפתור כדלקמן: ראשית, נוכל לתייג את הכדורים עם שמות (נדמיין שאנחנו מדביקים לכל כדור פתק עם שם אחר). לאחר שנתנו להם שמות, מובן כי כל שליפה וסידור של הכדורים נותן סידור אחר של השמות בשורה, ושכל סידור כזה יתקבל בסיכוי שווה – הרי מבחינת השליפה, אין הבדל בין שם של כדור שחור או של כדור לבן. לפיכך, הבעיה הפכה להיות הבעיה הבאה: מסדרים 8 אנשים (כדורים עם שמות) בשורה, מה הסיכוי שיואל ויוסי (שמות הכדורים השחורים) יימצאו זה לצד זה? את זה אנו יודעים לפתור – מספר הסידורים הכולל הוא $8!$, ומספר הסידורים בהם יוסי ליד יואל הוא $2 \cdot 7!$. לפיכך ההסתברות היא $\frac{7!\cdot 2}{8!} = \frac{1}{4}$. **יצאה לנו אותה התשובה!**

זה מרמז שאולי יכולנו להוכיח מראש שהבעיות שקולות, גם מבלי לחשב את ההסתברות המדויקת. כיצד? נשים לב שכל ההבדל בין הבעיות הוא שפעם אחת אנו בוחרים סידור עם שמות באופן אחיד ופעם סידור בלי שמות. קבוצות אלה בעלות גודל שונה – יש הרבה יותר סידורים עם שמות מאשר ללא שמות, לפיכך מרחבי המדגם באמת שונים, ולא נוכל להתאים ביניהם ישירות. אולם, לכל סידור בלי שמות מתאימים בדיוק אותו מספר של סידורים עם שמות. כמה? בהינתן סידור של הכדורים בלי שמות, יש $6! \cdot 2!$ סידורים עם שמות המתאימים לו – יש לכפול את הסידור הפנימי של הכדורים הלבנים בסידור הפנימי של הכדורים השחורים. לפיכך החלק היחסי בין הסידורים של אלה שבהם השחורים נמצאים זה לצד זה בסידור בלי שמות שווה בדיוק לחלקם היחסי בין סידורים עם שמות, שכן כל הגורמים (במונה ובמכנה) נכפלים באותו הקבוע.

3. נתונה חפיסת קלפים תקנית בעלת 52 קלפים. מחלקים 10 קלפים מתוך החבילה. חשבו מה ההסתברות שבדיוק 6 מהקלפים הם מסוג תלתן.

פתרון: נסמן ב-C את אוסף הקלפים בחבילה תקנית. כאמור $|C| = 52$, מתוכם $\frac{52}{4} = 13$ הם תלתנים. נגדיר את מרחב ההסתברות

$$\Omega = \{(a_1, ..., a_{10}) \in C \mid \forall i \neq j \in [10] \quad a_i \neq a_j\}.$$

זהו מרחב הסדרות של 10 קלפים שונים מתוך חבילת תקנית. גודל המרחב הוא $|\Omega| = \frac{52!}{(52-10)!} = \frac{52!}{42!}$. נגדיר עליו את פונקציית ההסתברות האחידה, כלומר לכל מאורע $A \subseteq \Omega$

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{42!}{52!}\,|A|.$$

הסדרות הרלוונטיות לנו הן הסדרות בהן 6 קלפים יצאו תלתן. ראשית יש לבחור את המיקומים בסדרה בהן יופיע תלתן, לכך יש $\binom{10}{6}$ אפשרויות. שנית, בתוך אותם מקומות יש לבחור 6 קלפים מתוך 13

התלתנים הקיימים בחפיסה עם חשיבות לסדר ($\frac{13!}{7!}$ אפשרויות). ביתר המקומות יש לבחור 4 קלפים מתוך 39 הקלפים שאינם תלתנים בחפיסה עם חשיבות לסדר ($\frac{39!}{35!}$ אפשרויות) סך הכל

$$|A| = \binom{10}{6}\frac{13!\cdot 39!}{7!\cdot 35!},$$

אזי

$$\mathbb{P}(A) = \frac{42!}{52!}\binom{10}{6}\frac{13!\cdot 39!}{7!\cdot 35!} \approx \frac{1}{100}.$$

דוגמאות – מרחבי הסתברות אינסופיים (בני-מניה)

מאחר והערנו שמרחבי הסתברות סופיים ובני-מניה הם דומים, וטרחנו להגדיר את פונקציית ההסתברות בצורה שהגיונית גם למרחבי הסתברות בני-מניה, רצוי שנראה מה זה מרחב הסתברות אינסופי.

1. יהי $\Omega = \mathbb{N}$. אם נגדיר

$$\mathbb{P}(A) = \frac{|A\cap\{1,2,3,4,5,6\}|}{6},$$

נקבל מודל נוסף לבעיית הקובייה ההוגנת.

נשים לב שפונקצית ההסתברות הזו **נתמכת** על הקבוצה $\{1,...,6\}$, למרות שהיא **מוגדרת** על כל $\mathbb{N}$. זו דוגמא לכך שמרחבי הסתברות יכולים להכיל מאורעות בעלי הסתברות אפס.

2. יהי $\Omega = \mathbb{N}$. נגדיר את $\mathbb{P} : 2^{\Omega} \to \mathbb{R}$ באופן הבא

$$\mathbb{P}(A) = \sum_{n\in A} 2^{-n}.$$

וכמקרה פרטי,

$$\mathbb{P}(\{n\}) = 2^{-n}.$$

- מתקבל מיד ש-$0 \geq \mathbb{P}(A)$ לכל $A \subseteq \mathbb{N} = \Omega$.
- נבדוק נרמול – הסתברות כל המרחב היא אכן 1:

$$\mathbb{P}(\Omega) = \sum_{n=1}^{\infty} 2^{-n} = \frac{1/2}{1-1/2} = \frac{1/2}{1/2} = 1.$$

כשהשתמשנו בנוסחה של סכום טור הנדסי $\sum_{n=1}^{\infty} q^n = \frac{q}{1-q}$.

- הערה פורמלית – נשים לב כי ההסתברות של כל מאורע $A \subseteq \Omega$ היא טור חלקי לטור הנדסי הנ"ל, לכן מתכנס ממבחן ההשוואה לטורים חיוביים. לפיכך, הפונקציה $\mathbb{P}$ מוגדרת היטב.
- סכימות בת-מניה עובדת מההגדרה, כי בעצם הגדרנו את ההסתברות של קבוצה כסכום על ההסתברויות של היחידונים שבה. לכן אם $\{A_i\}_{i=1}^{\infty}$ מאורעות זרים בזוגות, נסמן $A := \bigcup_{i=1}^{\infty} A_i$ ונקבל:

$$\mathbb{P}(A) = \sum_{n\in A} 2^{-n} = \sum_{i=1}^{\infty}\sum_{n\in A_i} 2^{-n} = \sum_{i=1}^{\infty}\mathbb{P}(A_i),$$

כאשר המעבר המרכזי השתמש בכך שכל $n \in A$ נמצא רק ב-A_i יחיד, וניתן לשנות סדר סכימה כיוון שהטור מתכנס בהחלט.

- דוגמא – נסמן $\mathbb{N}_{even}=\{2n\mid n\in\mathbb{N}\}$. נחשב את $\mathbb{P}\left(\mathbb{N}_{even}\right)$:

$$\mathbb{P}\left(\mathbb{N}_{even}\right)=\sum_{n=1}^{\infty}2^{-2n}=\sum_{n=1}^{\infty}4^{-n}=\frac{1/4}{1-1/4}=\frac{1/4}{3/4}=\frac{1}{3}.$$

הערה. הגדרה זו של מרחב אינסופי עולה באופן מאד טבעי. המרחב הזה ממדל למשל את הבעיה הבאה: מטילים מטבע הוגן עד שיוצא פלי. מה ההסתברות שהטלנו את המטבע n פעמים? לעת עתה לא נראה כי זהו המודל המתאים לניסוי זה. נטפל בנושא זה בפרק 4 של הקורס. מה שאפשר ללמוד כבר עכשיו הוא שהכוונה במרחב אינסופי אינה ש"אינסוף" היא תוצאה במרחב המדגם, אלא שיש אינסוף תוצאות אפשריות. אם נטיל מטבע עד שיוצא פלי, נעצור אחרי מספר הטלות סופי. אנחנו לעולם לא נטיל עד אינסוף (טעון הוכחה, אגב, אבל נכון). אבל אין מספר הטלות, גדול ככל שיהיה, שהוא בלתי אפשרי.

3. יהי $\Omega=\mathbb{N}$ ו $p:\Omega\to\mathbb{R}_+$ מוגדרת על ידי $p\left(n\right)=\frac{1}{n(n+1)}$. ראשית נוכיח כי p היא פונקציית הסתברות נקודתית על Ω.

<u>הוכחה:</u> ניתן להוכיח כי $\sum_{n\in\mathbb{N}}\frac{1}{n(n+1)}=1$. כלומר, $\sum_{n\in\mathbb{N}}p\left(n\right)=\sum_{n\in\mathbb{N}}\frac{1}{n(n+1)}=1$.
כעת, משהוכחנו כי זוהי פונקציית הסתברות נקודתית, נוכל להגדיר את מרחב ההסתברות $\left(\Omega,2^{\Omega},\mathbb{P}\right)$ על ידי

$$\mathbb{P}\left(A\right):=\sum_{\omega\in A}p\left(\omega\right).$$

נוכל לשאול שאלות הסתברותיות, למשל – מה ההסתברות לקבל מספר גדול מ-2?
<u>תשובה:</u> המאורע אותו אנו מחפשים הוא $A=\{n\in\mathbb{N}\mid n>2\}=\{3,4,5,...\}$. נשים לב כי במקרה זה יהיה קל יותר לחשב את ההסתברות לקבל את $A^c=\{1,2\}$. יתקיים:

$$\mathbb{P}\left(A^c\right)=p\left(1\right)+p\left(2\right)=\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}=\frac{1}{2}+\frac{1}{6}=\frac{2}{3}.$$

לפיכך

$$\mathbb{P}\left(A\right)=1-\mathbb{P}\left(A^c\right)=1-\frac{2}{3}=\frac{1}{3}.$$

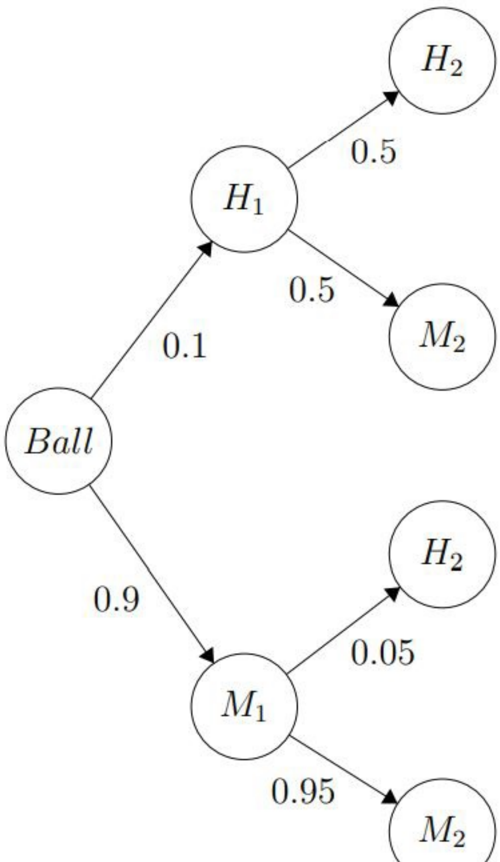
2 ניסויים רב-שלביים

לפעמים ניסוי הסתברותי מחולק באופן טבעי לשלבים התלויים זה בזה. חשבו למשל על הניסוי הבא: בשלב הראשון מוטל מטבע הוגן. בשלב השני, אם תוצאת ההטלה היא עץ - מוטלת קוביה בת שש פאות, ואחרת מוטלת קוביה בת שמונה פאות. נשים לב שההגרלה בשלב השני תלויה בתוצאת ההגרלה בשלב הראשון.

הגדרה 4. יהיו Ω_1, Ω_2 מרחבי מדגם. ניסוי דו-שלבי מתואר על ידי פונקציית הסתברות נקודתית p על Ω_1, וכן לכל $\omega_1 \in \Omega_1$ פונקציית הסתברות נקודתית p_{ω_1} על Ω_2 המגדירה את ההסתברות של כל תוצאה בשלב השני לאחר שהתקבלה בשלב הראשון ω_1. לניסוי הדו-שלבי מתאים מרחב הסתברות על $\Omega := \Omega_1 \times \Omega_2$ כאשר פונקציית ההסתברות הנקודתית המתאימה לתוצאת הניסוי הדו-שלבי $q : \Omega \to \mathbb{R}_+$ מוגדרת באמצעות

$$q\left(\left(\omega_1, \omega_2\right)\right) = p\left(\omega_1\right) p_{\omega_1}\left(\omega_2\right).$$

ניסוי n-שלבי יוגדר בדומה עבור $\Omega = \Omega_1 \times \ldots \times \Omega_n$.



דוגמא נניח כי מתרגל בקורס "מבוא להסתברות" הולך למרכז קוסל לבצע שתי הטלות לסל. ההסתברות של המתרגל לקלוע לסל בזריקת הכדור הראשונה הוא $\frac{1}{10}$. אם המתרגל קלע לסל בזריקה הראשונה, הוא משולהב ובזריקה השנייה יקלע בהסתברות $\frac{1}{2}$. אחרת, אם פספס בראשונה, הוא יקלע בשנייה רק בהסתברות $\frac{1}{20}$. מה ההסתברות שהמתרגל קלע בדיוק בזריקה אחת מבין השתיים?

תשובה: נבנה את מרחב ההסתברות. נסמן $\Omega_1 :=$ $\{H_1, M_1\}$ (Hit,Miss), ו $\Omega_2 = \{H_2, M_2\}$. נגדיר את פונקציות ההסתברות הנקודתיות:

$$p\left(H_1\right) = \frac{1}{10}, \quad p\left(M_1\right) = \frac{9}{10}.$$
$$p_{H_1}\left(H_2\right) = \frac{1}{2}, \quad p_{H_1}\left(M_2\right) = \frac{1}{2}.$$
$$p_{M_1}\left(H_2\right) = \frac{1}{20}, \quad p_{M_1}\left(M_2\right) = \frac{19}{20}.$$

ניתן גם לרשום את ההסתברויות הנ"ל בעץ (מופיע משמאל). כעת, כמו בהגדרה מרחב המדגם של הניסוי הדו-שלבי הוא $\Omega_1 \times \Omega_2$ עם פונק' הסתברות $q\left(\left(\omega_1, \omega_2\right)\right) = p\left(\omega_1\right) p_{\omega_1}\left(\omega_2\right)$.

אנו מחפשים את הסתברות המאורע $A = \{(H_1, M_2), (M_1, H_2)\}$. נחשב:

$$q\left(\left(H_1, M_2\right)\right) = p\left(H_1\right) p_{H_1}\left(M_2\right) = \frac{1}{10} \cdot \frac{1}{2} = \frac{1}{20},$$
$$q\left(\left(M_1, H_2\right)\right) = p\left(M_1\right) p_{M_1}\left(H_2\right) = \frac{9}{10} \cdot \frac{1}{20} = \frac{9}{200}.$$

לפיכך

$$\mathbb{P}_q\left(A\right)=q\left(\left(H_1,M_2\right)\right)+q\left(\left(M_1,H_2\right)\right)=\frac{1}{20}+\frac{9}{200}=\frac{19}{200}.$$

מבוא להסתברות וסטטיסטיקה 80430

תרגול 2

1 ניסויים רב-שלביים-המשך

דוגמא הביטו בניסוי הבא: בשלב הראשון מוגרל מספר טבעי $n\in\mathbb{N}$ על פי הניסוי שתואר בתרגול קודם במרחבים אינסופיים, בו $p\left(n\right)=2^{-n}$. בשלב השני מוגרל מספר באקראי באופן אחיד מתוך $[n]$. מה ההסתברות שהוגרל בסוף השלב השני המספר 1?

נבנה את הניסוי הדו שלבי. מרחב המדגם יהיה $\Omega=\mathbb{N}\times\mathbb{N}$. הפונקציה הקובעת את ההגרלה הראשונה היא כאמור $p\left(n\right)=2^{-n}$. לאחר מכן יתקיים

$$p_{n}\left(m\right)=\begin{cases}\frac{1}{n} & m\in[n]\\ 0 & \text{otherwise.}\end{cases}$$

תשובה: המאורע אותו אנו מחפשים הוא $A=\{(n,1)\mid n\in\mathbb{N}\}\subset\Omega$. מסיגמא-אדיטיביות נקבל

$$q\left(A\right)=\sum_{n=1}^{\infty}q\left((n,1)\right)=\sum_{n=1}^{\infty}p\left(n\right)\cdot p_{n}\left(1\right)=\sum_{n=1}^{\infty}2^{-n}\cdot\frac{1}{n}=\sum_{n=1}^{\infty}\frac{1}{2^{n}\cdot n}.$$

הטור הזה אכן מתכנס! למעשה, הביטוי הנ"ל שווה ל-$\log 2$.

דוגמא מתוך חפיסת קלפים מוציאים 4 קלפים בזה אחר זה (ללא החזרה), מהי ההסתברות שהקלף השלישי שנוציא יהיה קלף מלכה?

תשובה: נשים לב שמרחב המדגם שלנו Ω הוא אוסף כל הסדרות מאורך 4 של קלפים <u>שונים</u> מתוך החפיסה. מכיוון שמדובר בהסתברות אחידה, נגדיר $A=\{(card_{1},card_{2},card_{3},card_{4})\in\Omega:card_{3}\quad\text{is numbered Q}\}$ אז ההסתברות המבוקשת תהיה $p\left(A\right)=\frac{|A|}{|\Omega|}$.

כעת נגדיר את הקבוצה $B=\{(card_{1},card_{2},card_{3},card_{4})\in\Omega:card_{1}\quad\text{is numbered Q}\}$, ונבנה התאמה חח"ע ועל $f:A\to B$ המוגדרת על ידי:

$$f\left(card_{1},card_{2},card_{3},card_{4}\right)=\left(card_{3},card_{4},card_{1},card_{2}\right)$$

בדקו שזו התאמה חח"ע ועל, ומכאן נובע ש-$|A|=|B|$ ולכן כמובן $p\left(A\right)=p\left(B\right)$ (משום שאנחנו במרחב הסתברות אחידה), אבל לא מסובך לחשב את ההסתברות של המאורע B שכן $p\left(B\right)$ היא ההסתברות שהקלף <u>הראשון</u> שנבחר הוא מסוג מלכה שהיא כמובן $\frac{1}{13}$, ומכאן $p\left(A\right)=\frac{1}{13}$.

שימו לב שהבעיה הזו מתארת גם ניסוי רב שלבי, אבל אם ננסה לפתור אותו בצורה הזו ניתקל בבעיה שכן הסיכוי שהקלף השלישי יהיה מסוג מלכה מושפע משני הקלפים הראשונים שהוצאנו ולכן נצטרך לשקול 4 מקרים (שני הראשונים לא יצאו מלכה, אחד מהם יצא מלכה, אף אחד מהם לא יצא מלכה).

2 מרחבי מכפלה

מקרה פרטי מעניין של ניסויים רב שלביים, הוא כאשר אופי הניסוי בכל שלב נקבע מראש ואינו מושפע או תלוי בתוצאות השלבים הקודמים לו. למשל הטלת מטבע מס' פעמים, או ניסוי המורכב מהטלת מטבע והטלת קוביה. כדי לבנות מרחב הסתברות מתאים לניסויים מסוג זה, נשתמש במרחבי ההסתברות המתאימים לכל אחד מהשלבים בנפרד.

הגדרה 2.1 יהיו $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1 = \mathbb{P}_{p_1}), (\Omega_2, \mathcal{F}_2, \mathbb{P}_2 = \mathbb{P}_{p_2})$ מרחבי הסתברות, עם פונקציות הסתברות נקודתיות p_1, p_2. מרחב המכפלה שלהם הוא $(\Omega_1 \times \Omega_2, \mathcal{F}, \mathbb{P} = \mathbb{P}_{p_1 \times p_2})$, כאשר $p_1 \times p_2((\omega_1, \omega_2)) = p_1(\omega_1) \cdot p_2(\omega_2)$.

ראיתם בכיתה כי לכל זוג מאורעות $A \in \mathcal{F}_1, B \in \mathcal{F}_2$ יתקיים

$$\mathbb{P}(A_1 \times B) = \mathbb{P}_1(A_1) \cdot \mathbb{P}_2(B).$$

הגדרה 2.2 יהיו $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1 = \mathbb{P}_{p_1}), (\Omega_2, \mathcal{F}_2, \mathbb{P}_2 = \mathbb{P}_{p_2})$ מרחבי הסתברות, ויהי $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mathbb{P})$ מרחב המכפלה שלהם. מאורעות מהסוג $\Omega_1 \times B$ או $A \times \Omega_2$ עבור $A \in \mathcal{F}_1, B \in \mathcal{F}_2$, מכונים מאורעות שוליים. מאורע מן הסוג $A \times B$ מכונה מאורע מכפלה.

הערה 2.3 חשוב! מאורע מכפלה מתאר את הדרישה שהתוצאה בשלב הראשון שייכת לקבוצה מסוימת, והתוצאה בשלב השני שייכת לקבוצה מסוימת (אולי אחרת מהראשונה) וכן הלאה.

טענה 2.4 יהיו $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1 = \mathbb{P}_{p_1}), (\Omega_2, \mathcal{F}_2, \mathbb{P}_2 = \mathbb{P}_{p_2})$ מרחבי הסתברות, ויהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב המכפלה שלהם. יהיו A ו-B מאורעות ב-Ω_1 ו-Ω_2 בהתאמה. אזי $\mathbb{P}(A \times B) = \mathbb{P}_1(A) \cdot \mathbb{P}_2(B)$.

דוגמא נבנה את מרחב המכפלה עבור ניסוי שמכיל הטלת מטבע הוגן $\left(\Omega_1 = \{H, T\}, \mathcal{F}_1 = 2^{\{H,T\}}, \mathbb{P}_{p_1}\right)$ והטלת קוביה הוגנת $\left(\Omega_2 = \{1,2,3,4,5,6\}, \mathcal{F}_2 = 2^{\{1,2,3,4,5,6\}}, \mathbb{P}_{p_2}\right)$.

	H	T
1	$(H,1) \to p_1 \times p_2((H,1)) = p_1(H) \cdot p_2(1) = \frac{1}{12}$	$(T,1) \to p_1 \times p_2((T,1)) = p_1(T) \cdot p_2(1) = \frac{1}{12}$
2	$(H,2) \to p_1 \times p_2((H,2)) = p_1(H) \cdot p_2(2) = \frac{1}{12}$	$(T,2) \to p_1 \times p_2((T,2)) = p_1(T) \cdot p_2(2) = \frac{1}{12}$
3	$(H,3) \to p_1 \times p_2((H,3)) = p_1(H) \cdot p_2(3) = \frac{1}{12}$	$(T,3) \to p_1 \times p_2((T,3)) = p_1(T) \cdot p_2(3) = \frac{1}{12}$
4	$(H,4) \to p_1 \times p_2((H,4)) = p_1(H) \cdot p_2(4) = \frac{1}{12}$	$(T,4) \to p_1 \times p_2((T,4)) = p_1(T) \cdot p_2(4) = \frac{1}{12}$
5	$(H,5) \to p_1 \times p_2((H,5)) = p_1(H) \cdot p_2(5) = \frac{1}{12}$	$(T,5) \to p_1 \times p_2((T,5)) = p_1(T) \cdot p_2(5) = \frac{1}{12}$
6	$(H,6) \to p_1 \times p_2((H,6)) = p_1(H) \cdot p_2(6) = \frac{1}{12}$	$(T,6) \to p_1 \times p_2((T,6)) = p_1(T) \cdot p_2(6) = \frac{1}{12}$

הבה נחשב את ההסתברות של המאורע "יצא בקוביה מספר גדול מ-4 (שהוא המאורע $A = \{5,6\}$ במרחב המקורי של הטלות הקוביה) וגם המטבע הוטל עם עץ כלפי מעלה" (שהוא האירוע $B = \{H\}$ במרחב המקורי של הטלות המטבע). במרחב המכפלה, המאורעות השוליים המתאימים לשניים הנ"ל יתוארו באופן הבא:

$$A \times \{H,T\} = \{(a,b) : a \in \{5,6\}, b \in \{H,T\}\}, \{1,2,3,4,5,6\} \times B = \{(a,b) : a \in \{1,...,6\}, b \in \{H\}\}$$

לפי הטענה האחרונה, המאורע "יצא בקוביה מספר גדול מ-4 וגם המטבע הוטל עם עץ כלפי מעלה" הוא מאורע המכפלה $A \times B = (A \times \{H,T\}) \cap (\{1,...,6\} \times B)$ והסתברותו תהיה

$$\mathbb{P}(A \times B) = \mathbb{P}_1(A) \cdot \mathbb{P}_2(B) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

.

נשים לב שהאירוע המשלים $(A \times B)^c$ הוא אינו מאורע מכפלה: נניח בשלילה כי $(A \times B)^c = U \times V$ עבור מאורעות מתאימים U, V. מאחר ו-$A \times B = \{(5,H), (6,H)\}$ נובע כי

$$(3,H), (6,T) \in (A \times B)^c = U \times V$$

כלומר $6 \in U,\ H \in V$ ולכן $(6,H) \in (A \times B)^c$, אבל $(6,H) \in A \times B$ וקיבלנו סתירה.

3 נוסחאת ההסתברות השלמה

הגדרה 3.1 חלוקה בת מניה של מרחב מדגם Ω היא אוסף בן מניה של קבוצות זרות בזוגות שאיחודן הוא Ω. באופן שקול חלוקה בת מניה של Ω היא סדרה $(\Omega_i)_{i=1}^{\infty}$ של קבוצות זרות בזוגות שאיחודן הוא Ω.

טענה 3.2 נוסחת ההסתברות השלמה: תהי $\{A_1, A_2, ...\}$ חלוקה של Ω ומאורע $B \subseteq \Omega$. אזי

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

הערה נוסחת ההסתברות השלמה היא הכללה של השיטה הקומבינטורית של ספירה לפי מקרים: אם B קבוצה הניתנת לכתיבה כאיחוד זר של תתי-קבוצות $B = \bigcup_{i=1}^{\infty} (B \cap A_i)$, אז הגודל של הקבוצה הוא פשוט סכום הגדלים של תתי-הקבוצות $|B| = \sum_{i=1}^{\infty} |B \cap A_i|$. מכאן נובע באופן מיידי שאם B מאורע במרחב הסתברות אחיד, אז ההסתברות של B שווה לסכום ההסתברויות של $B \cap A_i$. נוסחת ההסתברות השלמה שלמדתם בהרצאה היא הכללה של אותו עקרון למרחבי הסתברות כלליים.

שימו לב כי ניתן להשתמש בנוסחאת ההסתברות השלמה לחישוב מאורעות שאינם מאורעות מכפלה, נראה זאת בדוגמא הבאה:

תרגיל מגרילים מספר באקראי מ-$[20]$, אז מטילים שלושה מטבעות, ומוסיפים את כמות הפלי שיצאו למספר. מה הסיכוי שהסכום המתקבל שייך לקבוצה $\{2, 7, 8\}$?

פתרון: את הניסוי הראשון נמדל על ידי $\Omega_1 = [20]$, וכן $p_1(\omega) = \frac{1}{20} \quad \forall \omega \in \Omega_1$. את הניסוי השני על ידי $\Omega_2 = \{0,1\}^3$, וכן $p_2(\omega) = \frac{1}{8} \quad \forall \omega \in \Omega_2$. נשים לב שעל אף שהתוצאה אותה אנו מחפשים קשורה למספר ולהטלות המטבעות, הניסויים עצמם אינם קשורים זה בזה! לכן נוכל למדל את השאלה באמצעות מרחב מכפלה. אנו מעוניינים במאורע

$$A = \{(n, (x_1, x_2, x_3)) \in \Omega_1 \times \Omega_2 \mid n + x_1 + x_2 + x_3 \in B\},$$

כאשר $B = \{2, 7, 8\}$. שימו לב – המאורע A איננו מאורע מכפלה, כלומר לא ניתן לכתוב אותו כמכפלה קרטזית של שתי קבוצות $A_1 \times A_2$. במרחב מכפלה, קל לחשב את ההסתברות של מאורעות מכפלה (זוהי פשוט מכפלת ההסתברויות של הנכפלים, כפי שראיתם בכיתה), אך לרוב יש לעבוד קשה יותר עבור מאורעות שאינם מאורעות מכפלה, נשתמש כאן בנוסחאת ההסתברות השלמה ונכתוב את A כאיחוד זר של מאורעות מכפלה. לכל $k \in \{0, 1, 2, 3\}$ נסמן

$$C_k = \left\{(n, x_1, x_2, x_3) \in \Omega_1 \times \Omega_2 \mid \sum_{i=1}^{3} x_i = k\right\},$$

כלומר המאורע שהתקבלו k פלי בדיוק, ו-

$$D_k = \left\{(n, (x_1, x_2, x_3)) \in A \mid \sum_{i=1}^{3} x_i = k\right\} = A \cap C_k.$$

כעת מכיוון ש-C_k הם חלוקה של המרחב $\Omega_1 \times \Omega_2$ כולו ניתן להשתמש בנוסחאת ההסתברות השלמה:

$$\mathbb{P}(A) = \sum_{k=0}^{3} \mathbb{P}(A \cap C_k) = \sum_{k=0}^{3} \mathbb{P}(D_k).$$

ראשית $D_0 = \{(n, (x_1, x_2, x_3)) \in \Omega_1 \times \Omega_2 \mid n \in B, \sum x_i = 0\} = \{n \in B\} \times C_0$. מנוסחת המכפלה נקבל:

$$\mathbb{P}(D_0) = \mathbb{P}(\{n \in B\}) \cdot \mathbb{P}(C_0) = \frac{3}{20} \cdot \frac{1}{8}.$$

עבור $k = 1$ נקבל בדומה

$$D_1 = \left\{(n_1, (x_1, x_2, x_3)) \in \Omega_1 \times \Omega_2 \mid n + 1 \in B, \sum x_i = 1\right\} = \{1, 6, 7\} \times C_1.$$

לפיכך

$$\mathbb{P}(D_1) = \frac{3}{20} \cdot \frac{3}{8}.$$

נשים לב שכאשר $k = 2$, המספר 2 לעולם לא יתקבל, לכן אנו נותרים עם $D_2 = \{5, 6\} \times C_2$, כלומר $\mathbb{P}(D_2) = \frac{2}{20} \cdot \frac{3}{8}$. לבסוף $D_3 = \{4, 5\} \times C_3$, לפיכך $\mathbb{P}(D_3) = \frac{2}{20} \cdot \frac{1}{8}$. סה"כ:

$$\mathbb{P}(A) = \frac{3}{160} + \frac{9}{160} + \frac{6}{160} + \frac{2}{160} = \frac{20}{160} = \frac{1}{8}.$$

דוגמא בוחרים סידור אקראי בשורה של 3 כדורים אדומים, 5 לבנים, ו-7 שחורים. מה הסיכוי ששני הקצוות יהיו באותו הצבע?

פתרון: נסמן Ω – אוסף כל הסידורים, B – אוסף כל הסידורים כך ששני הקצוות באותו הצבע, וכן A_c – אוסף כל הסידורים כך שהכדור הראשון הוא בצבע c, עבור $c \in \{B, W, R\}$ (שימו לב כי זהו לא מרחב מכפלה!). ברור כי $\Omega = A_B \cup A_W \cup A_R$, וכן שהקבוצות A_B, A_W, A_R זרות, שכן לא ייתכן שלכדור הראשון שני צבעים שונים. על כן עלינו לחשב את הסכום

$$\mathbb{P}(B) = \mathbb{P}(B \cap A_W) + \mathbb{P}(B \cap A_B) + \mathbb{P}(B \cap A_R).$$

נחשב כל אחד בנפרד – $B \cap A_W$ הוא המאורע בו שני הכדורים בקצוות הם לבנים. על כן עלינו לסדר רק את השורה הפנימית בה יש 3 אדומים, 3 לבנים, ו-7 שחורים:

$$|B \cap A_W| = \binom{13}{3} \cdot \binom{10}{3}.$$

החישוב הוא על ידי כך שבחרנו ראשית את מיקום הכדורים האדומים מתוך 13 המקומות, ואז את מיקום הלבנים מתוך 10 המקומות שנותרו.

בדומה נוכל לחשב את שאר ההסתברויות:

$$|B \cap A_B| = \binom{13}{3} \cdot \binom{10}{5},$$
$$|B \cap A_R| = \binom{13}{1} \cdot \binom{12}{5},$$
$$|\Omega| = \binom{15}{3} \cdot \binom{12}{5}.$$

לפיכך:

$$\mathbb{P}(B) = \frac{|B \cap A_W| + |B \cap A_B| + |B \cap A_R|}{|\Omega|} = \frac{\binom{13}{3} \cdot \binom{10}{3} + \binom{13}{3} \cdot \binom{10}{5} + \binom{13}{1} \cdot \binom{12}{5}}{\binom{15}{3} \cdot \binom{12}{5}} = ... = \frac{34}{105}.$$

4 חסם האיחוד (אי שוויון בול)

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. הוכחתם בתרגיל הראשון כי לשני מאורעות $A, B \in \mathcal{F}$ מתקיים:

$$\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B),$$

ובאופן כללי לסדרה של מאורעות יתקיים:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i).$$

למעשה זה נכון גם למספר אינסופי של מאורעות:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

מתי האי-שוויון שימושי? כאשר מסובך מדי לחשב את ההסברות המדויקת של המאורע, או כאשר מספיק לנו למצוא חסם מלעיל של ההסתברות.

תרגיל יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, ותהא סדרה $(A_n)_{n=1}^{\infty}$ של מאורעות שמתקיימים כמעט תמיד (כלומר $\mathbb{P}(A_n) = 1$). נוכיח כי $\bigcap_{n=1}^{\infty} A_n$ גם מתקיים כמעט תמיד.

פתרון: נתבונן במאורע המשלים של החיתוך:

$$\mathbb{P}\left(\left(\bigcap_{n=1}^{\infty}A_n\right)^c\right)=\mathbb{P}\left(\bigcup_{n=1}^{\infty}A_n^c\right)$$

ולפי אי-שוויון בול נקבל:

$$\mathbb{P}\left(\left(\bigcap_{n=1}^{\infty}A_n\right)^c\right)\le\sum_{i=1}^{\infty}\mathbb{P}\left(A_i^c\right)=\sum_{i=1}^{\infty}\left(1-\mathbb{P}\left(A_i\right)\right)=0$$

ולכן $\mathbb{P}\left(\left(\bigcap_{n=1}^{\infty}A_n\right)^c\right)=0$ ונובע כי $\mathbb{P}\left(\bigcap_{n=1}^{\infty}A_n\right)=1$.

תרגיל נביט בניסוי הבא: ניקח לוח משבצות ריבועי 4×4, ונגריל על כל משבצת מספר באקראי מתוך $[100]$. חסמו מלמעלה את ההסתברות שיהיו זוג משבצות סמוכות (לא כולל אלכסונים) המציגות את אותו המספר.

פתרון: נסמן את הלוח $V=[4]\times[4]$, נסדר את ריבועי הלוח בסדר כלשהוא $u_1,...,u_{16}$, ומרחב המדגם מתאים לכל משבצת בלוח מספר ב-$[100]$, כלומר $\Omega=[100]^V=[100]^{[4]\times[4]}$. נסמן לכל $u,v\in V$ את $A_{u,v}$ – המאורע ש-u ו-v הגרילו את אותו המספר, ונסמן $u\sim v$ אם שתי המשבצות סמוכות. אנו רוצים לחסום מלמעלה את המאורע $A=\bigcup_{u_i\sim u_j,i<j}A_{u_i,u_j}$, ונוכל להשתמש בחסם האיחוד:

$$\mathbb{P}\left(A\right)=\mathbb{P}\left(\bigcup_{u_i\sim u_j,i<j}A_{u_i,u_j}\right)\le\sum_{u_i\sim u_j,i<j}\mathbb{P}\left(A_{u_i,u_j}\right).$$

חישוב ההסתברויות למאורעות הפרטיים איננו קשה: $P\left(A_{u,v}\right)=\frac{1}{100}$, כיוון שניתן לבחור שרירותית את הערך של v, ואז יש לקבוע ביחידות את הערך של u מתוך $[100]$. כעת יש להבין מה מספר זוגות המשבצות הסמוכות. נוכל למשל לספור סמיכויות אנכיות וסמיכויות אופקיות – יש $3\cdot 4=12$ זוגות של משבצות סמוכות באופן אנכי, וכמובן אותו מספר אופקי. בסך הכל 24 זוגות של משבצות סמוכות. לפיכך

$$\mathbb{P}\left(A\right)\le\sum_{u_i\sim u_j,i<j}\mathbb{P}\left(A_{u_i,u_j}\right)=24\cdot\frac{1}{100}=0.24.$$

תרגיל יהי Ω מרחב כל הסדרות של $2n$ ביטים כך במופיעים בהן בדיוק n 1-ים, עם הסתברות אחידה. חסמו מלמעלה את ההסתברות שיש k 1-ים ברצף. הוכיחו בעזרת החסם כי כאשר $k=n-1$ ו-n שואף לאינסוף, ההסתברות שואפת ל-0.

פתרון: נסמן את מרחב המדגם $\Omega=\left\{(x_1,...,x_{2n})\mid x_i\in\{0,1\},\sum x_i=n\right\}$. לכל $1\le\ell\le 2n-k+1$ נסמן ב-E_ℓ את המאורע $E_\ell=\left\{(x_1,...,x_{2n})\in\Omega\mid x_\ell=x_{\ell+1}=...=x_{\ell+k-1}=1\right\}$.

נשים לב כי המאורע אותו אנו מחפשים הוא $\bigcup_{\ell=1}^{2n-k+1}E_\ell$. נחשב ראשית את הסיכוי של כל E_ℓ בנפרד – נשים לב שלאחר שקבענו k 1-ים, יש לקבוע את מיקומם של $n-k$ ה-1-ים הנותרים מתוך $2n-k$ המקומות הנותרים. לפיכך $|E_\ell|=\binom{2n-k}{n-k}$, לכן

$$\mathbb{P}\left(E_\ell\right)=\frac{\binom{2n-k}{n-k}}{\binom{2n}{n}}=\frac{(2n-k)!n!n!}{(n-k)!n!\,(2n)!}=\frac{(2n-k)!n!}{(n-k)!\,(2n)!}.$$

נשתמש בחסם האיחוד

$$\mathbb{P}\left(E\right)=\mathbb{P}\left(\bigcup_{\ell=1}^{2n-k+1}E_\ell\right)\le\sum_{\ell=1}^{2n-k+1}\mathbb{P}\left(E_\ell\right)=(2n-k+1)\cdot\frac{(2n-k)!n!}{(n-k)!\,(2n)!}.$$

עבור החלק השני של התרגיל נציב $k=n-1$. נקבל

$$\mathbb{P}\left(E\right)\le(n+2)\cdot\frac{(n+1)!n!}{(2n)!}=(n+2)\left(n+1\right)\cdot\frac{n!n!}{(2n)!}=\frac{(n+2)\left(n+1\right)}{\binom{2n}{n}}\le\frac{(n+2)\left(n+1\right)}{2^n}\to 0.$$

השתמשנו כאן בחסם $2^n\le\binom{2n}{n}$.

5 נוסחת ההכלה וההפרדה

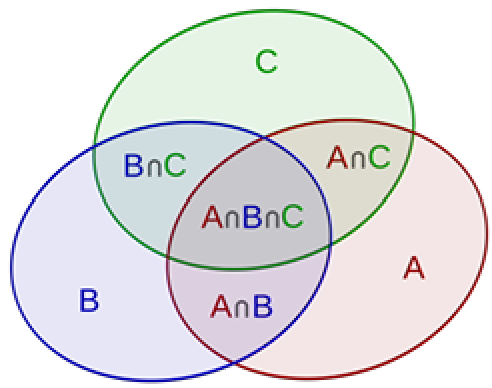
ראיתם בכיתה את הנוסחאות הבאות - יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות, ויהיו $A,B,C\in\mathcal{F}$ מאורעות. אזי

$$\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\cap B),$$
$$\mathbb{P}(A\cup B\cup C)=\mathbb{P}(A)+\mathbb{P}(B)+\mathbb{P}(C)-\mathbb{P}(A\cap B)-\mathbb{P}(A\cap C)-\mathbb{P}(B\cap C)+\mathbb{P}(A\cap B\cap C).$$

ובאופן כללי עבור מאורעות $A_1,...,A_n$ אם נגדיר $A_I:=\bigcap_{i\in I}A_i$ (עבור $I\subseteq[n]$) אז מתקיים:

$$\mathbb{P}\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\mathbb{P}(A_i)-\sum_{1\le i<j\le n}\mathbb{P}(A_i\cap A_j)+\sum_{1\le i<j<k\le n}\mathbb{P}(A_i\cap A_j\cap A_k)-...+(-1)^{n+1}\mathbb{P}\left(\bigcap_{i=1}^{n}A_i\right)$$
$$=\sum_{l=1}^{n}\sum_{I\subseteq[n],|I|=l}(-1)^{l+1}|A_I|$$

הנוסחא נעשית ברורה יותר כאשר מסתכלים על דיאגרמת וון הבאה:



נוסחת ההכלה וההפרדה שימושית כאשר עלינו לחשב הסתברות איחוד של מאורעות שאינם זרים, כאשר קל יותר לחשב את הסתברות החיתוכים שלהם.

תרגיל בוחרים חלוקה מקרית של 21 תפוזים **זהים** ל-5 ילדים. מה ההסתברות שאף ילד לא יקבל יותר מ-6 תפוזים?

פתרון: נסמן $\Omega=\{(x_1,...,x_5)\mid\sum x_i=21\}$.
מרחב ההסתברות שלנו אחיד על Ω. המאורע המבוקש הוא $A=\{(x_1,...,x_5)\in\Omega\mid\forall i,x_i\le 6\}$.
נסמן לכל $k\in[5]$ את המאורע $A_k=\{(x_1,...,x_5)\in\Omega\mid x_k\le 6\}$. אנו מבקשים למצוא את $\mathbb{P}(A_1\cap...\cap A_5)$.
כדי להעביר את זה לצורה של איחוד נחשב את המאורע המשלים

$$\mathbb{P}(A_1\cap...\cap A_5)=1-\mathbb{P}((A_1\cap...\cap A_5)^c)=1-\mathbb{P}(A_1^c\cup...\cup A_5^c).$$

כמובן $A_k^c=\{(x_1,...,x_5)\in\Omega\mid x_k\ge 7\}$. מאורעות מהצורה הזו אנו כבר יודעים לחשב! לכל k יתקיים

$$|A_k^c|=\left\{\sum_{i=1}^{5}y_i=14,\quad y_i\ge 0\right\}=\binom{14+4}{4}.$$

לכל $1\le k_1<k_2\le 5$ יתקיים

$$\left|A_{k_1}^c\cap A_{k_2}^c\right|=\left|\left\{\sum_{i=1}^{5}y_i=7,\quad y_i\ge 0\right\}\right|=\binom{7+4}{4}.$$

בדומה עבור שלושה אינדקסים

$$\left|A_{k_1}^c\cap...\cap A_{k_3}^c\right|=\left|\left\{\sum_{i=1}^{5}y_i=0,\quad y_i\ge 0\right\}\right|=|(0,0,...,0)|=1.$$

לא ייתכנו חיתוכים של יותר מ-3 אינדקסים. לפיכך לפי נוסחת ההכלה וההפרדה

$$\mathbb{P}(A)=5\cdot\frac{\binom{14+4}{4}}{\binom{21+4}{4}}-\binom{5}{2}\cdot\frac{\binom{7+4}{4}}{\binom{21+4}{4}}+\binom{5}{3}\cdot\frac{1}{\binom{21+4}{4}}=\frac{5\cdot\binom{18}{4}-10\cdot\binom{11}{4}+10\cdot 1}{\binom{25}{4}}.$$

6 חומר נוסף[1]

הוכחת $\sum_{n=1}^{\infty}\frac{1}{2^n\cdot n}=\log 2$

בטורים מהסוג הזה (טור הנדסי בתוספת פולינום כלשהו ב-n בביטוי שבתוך הסכום), לרוב ניתן לטפל באמצעות פונקציות יוצרות.

נתבונן בטור $\sum_{n=1}^{\infty}x^n$. זהו טור חזקות המתכנס בהחלט בתחום $|x|<1$, ותוצאתו

$$\sum_{n=0}^{\infty}x^n=\frac{1}{1-x}.$$

בתוך תחום ההתכנסות נוכל לבצע אינטגרל איבר-איבר (והזזה ב-1) ולקבל

$$\sum_{n=1}^{\infty}\frac{x^n}{n}=\sum_{n=0}^{\infty}\frac{x^{n+1}}{n+1}=-\log\left(1-x\right).$$

כעת נציב $x=\frac{1}{2}$ ונקבל

$$\sum_{n=1}^{\infty}\frac{\left(\frac{1}{2}\right)^n}{n}=\sum_{n=1}^{\infty}\frac{1}{2^n\cdot n}=-\log\left(\frac{1}{2}\right)=-\left(\log 1-\log 2\right)=\log 2,$$

כרצוי.

לחלופין – מאינפי אנו יודעים כי טור טיילור של $\log\left(1-x\right)$ הוא

$$\log\left(1-x\right)=-\sum_{n=1}^{\infty}\frac{x^n}{n},$$

והוא מתכנס לכל $|x|<1$. על כן נוכל להמשיך כמקודם.

דוגמא (לקריאה בלבד) הביטו בניסוי הבא: בשלב הראשון מוגרל מספר באופן אחיד מהקבוצה $[3]=\{1,2,3\}$, נסמנו a. בשלב השני מוגרל מספר באופן אחיד מהקבוצה $[a]=\{1,...,a\}$, אותו נסמן ב-b. בשלב השלישי, מוגדל מספר באופן אחיד מהקבוצה $[a+b]=\{1,...,a+b\}$. מה ההסתברות שהוגרל בסוף השלב השלישי המספר 1?

תשובה: זהו ניסוי תלת-שלבי. נבחר את מרחב המדגם להיות $\Omega=\mathbb{N}^3$. הפונקציה הקובעת את ההגרלה הראשונה היא $p(a)=\begin{cases}\frac{1}{3} & x\in[3]\\ 0 & x\notin[3]\end{cases}$. לאחר מכן, הפונקציה הקובעת את ההגרלה מוגדרת על ידי $p_a(b)=\begin{cases}\frac{1}{a} & b\in[a]\\ 0 & b\notin[a]\end{cases}$. לבסוף, הפונקציה הקובעת את ההגרלה השלישית היא $p_{a,b}(c)=\begin{cases}\frac{1}{a+b} & c\in[a+b]\\ 0 & c\notin[a+b]\end{cases}$.
על כן פונקצית ההסתברות הנקודתית המתאימה למרחב ההסתברות של הניסוי התלת-שלבי כולו היא

$$q\left((a,b,c)\right)=p(a)p_a(b)p_{a,b}(c)=\begin{cases}\frac{1}{3}\cdot\frac{1}{a}\cdot\frac{1}{a+b} & a\in[3],\quad b\in[a],\quad c\in[a+b]\\ 0 & \text{otherwise.}\end{cases}$$

המאורע אותו אנו מחפשים הוא $A=\{(a,b,1)\mid a\in[3],\quad b\in[a]\}$. נרשום במפורש את כל האיברים במאורע

$$A=\left\{(1,1,1),(2,1,1),(2,2,1),(3,1,1),(3,2,1),(3,3,1)\right\}.$$

אזי

$$q(A)=\frac{1}{3}\left(\frac{1}{1}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{4}+\frac{1}{3}\cdot\frac{1}{4}+\frac{1}{3}\cdot\frac{1}{5}+\frac{1}{3}\cdot\frac{1}{6}\right)=\frac{359}{1080}.$$

תרגיל (לקריאה בלבד) מגרילים פונקציה מקרית באופן אחיד מ-$[n]$ ל-$[m]$. מה ההסתברות שהיא על?

[1] חומר לקריאה עצמאית שלא הופיע בתרגול.

פתרון: מרחב המדגם שלנו הוא $\Omega = [m]^{[n]}$. ראשית נשים לב שאם $m > n$ לא קיימות כלל פונקציות על מ-$[n]$ ל-$[m]$ על כן ההסתברות היא 0. נניח אם כן $m \leq n$. אנו נחשב באמצעות המאורע המשלים! כלומר נסמן $A = \{f \in \Omega \mid f \text{surjective not is}\}$. מה זה אומר שפונקציה איננה על? זה אומר שיש איבר ב-$[m]$ שהיא לא מחזירה כפלט. זה מוביל אותנו לסמן את המאורעות הבאים: לכל $y \in [m]$ נסמן

$$A_y = \{f \in \Omega \mid y \notin f([n])\}.$$

מתקיים כמובן $A = \bigcup_{y \in [m]} A_y$. אם נציב זאת בנוסחת ההכלה וההפרדה נקבל

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{y \in [m]} A_y\right) = \sum_{l=1}^{m} (-1)^{l+1} \sum_{1 \leq y_1 < y_2 < \cdots < y_l \leq m} \mathbb{P}\left(\bigcap_{j=1}^{l} A_{y_j}\right).$$

<u>נחשב את ההסתברויות:</u>

לכל $1 \leq y \leq m$ מתקיים

$$\mathbb{P}(A_y) = \mathbb{P}(\{f(1), f(2), ..., f(n) \neq y\}) = \mathbb{P}(\{f(1), ..., f(n) \in [m] \setminus \{y\}\}) = \left(1 - \frac{1}{m}\right)^n.$$

לכל $1 \leq y_1 < y_2 \leq m$ מתקיים

$$\mathbb{P}(A_{y_1} \cap A_{y_2}) = \mathbb{P}(\{f(1), ..., f(n) \in [m] \setminus \{y_1, y_2\}\}) = \left(1 - \frac{2}{m}\right)^n.$$

וכן הלאה לכל $1 \leq y_1 < y_2 < \cdots < y_l \leq m$ נקבל

$$\mathbb{P}\left(\bigcap_{j=1}^{l} A_{y_j}\right) = \left(1 - \frac{l}{m}\right)^n.$$

אם כן:

$$\mathbb{P}(A) = \sum_{l=1}^{m} (-1)^{l+1} \cdot \binom{m}{l} \cdot \left(1 - \frac{l}{m}\right)^n$$

$$\Downarrow$$

$$\mathbb{P}(\{f \in \Omega \mid f \text{surjective is}\}) = 1 - \sum_{l=1}^{m} (-1)^{l+1} \cdot \binom{m}{l} \cdot \left(1 - \frac{l}{m}\right)^n.$$

מבוא להסתברות וסטטיסטיקה 80430

תרגול 3

1 הסתברות מותנית

יהא $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. יהיו A, B שני מאורעות כך ש $\mathbb{P}(A) > 0$ אז

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

היא ההסתברות של המאורע B, בהינתן שידוע לנו ש-A קרה. ראיתם בכיתה כי הפונקציה $\mathbb{P}(\cdot \mid A) \to [0,1]$ היא פונקציית הסתברות על $(\Omega, \mathcal{F})$. ניתן לחשוב עליה כעל "עדכון" של ההסתברויות של פונקציית ההסתברות המקורית, בהינתן המידע החדש כי A התקיים.

תרגיל 1. מטילים שתי קוביות הוגנות שונות. נתון שסכום התוצאות שהתקבלו הוא לפחות 7.

1. מה הסיכוי שהסכום הוא לפחות 10?
2. כעת נתון גם כי בקוביה הראשונה יצא לפחות 5. מה הסיכוי כעת שהסכום הוא לפחות 10?
3. מרחב ההסתברות הוא $\Omega = [6]^2$. נסמן $B = \{(x,y) \in \Omega \mid x+y \geq 10\}, A = \{(x,y) \in \Omega \mid x+y \geq 7\}$.
עלינו לחשב את:
$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}.$$
נשים לב כי $B \subseteq A$ לפיכך $B \cap A = B$. נחשב:

$$B = \{(4,6), (5,5), (5,6), (6,4), (6,5), (6,6)\} \Longrightarrow |B| = 1+2+3 = 6.$$
$$A = \{(1,6), (2,5), (2,6), ..., (6,1), (6,2), (6,3), ..., (6,6)\} \Longrightarrow |A| = 1+2+3+...+6 = \frac{6 \cdot 7}{2} = 21.$$

על כן:
$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{\frac{|B|}{|\Omega|}}{\frac{|A|}{|\Omega|}} = \frac{|B|}{|A|} = \frac{6}{21} = \frac{2}{7}.$$

4. כעת ידוע גם כי מתקיים $C = \{(x,y) \in \Omega \mid x \geq 5\} = \{5,6\} \times [6]$. אם כן המאורע בו אנו מתנים הוא כי מתקיים A וגם C, במילים אחרות $C \cap A$. על כן:
$$\mathbb{P}(B \mid C \cap A) = \frac{\mathbb{P}(B \cap C \cap A)}{\mathbb{P}(C \cap A)}.$$

כמקודם, מכיוון ש-A $B \subset$ מתקיים

$$B \cap C \cap A = B \cap C = \{(x,y) \in \Omega \mid x \geq 5, x+y \geq 10\} = \{(5,5),(5,6),(6,4),(6,5),(6,6)\}$$
$$\implies |B \cap C \cap A| = 5,\ A \cap C = \{(x,y) \in \Omega \mid x \geq 5, x+y \geq 7\} = C \setminus \{(5,1)\} \implies$$
$$\implies |A \cap C| = |C| - 1 = 12 - 1 = 11.$$

לפיכך:

$$\mathbb{P}(B \mid C \cap A) = \frac{5}{11}.$$

העובדה כי ההסתברות עלתה לא צריכה להפתיע אותנו – אם ידוע כי ערך הקוביה הראשונה לפחות 5, הסיכוי כי סכום הערכים יהיה גבוה עלה. שימו לב שכאן בעצם יש לנו ניסוי דו שלבי, ובאופן כללי ניתן לחשוב על ניסוי דו שלבי כעל הסתברות מותנית כפי שנראה בהמשך.

2 נוסחת ההסתברות השלמה

מהנוסחה להסתברות מותנית אנו מקבלים מיידית את הנוסחה הבאה, אשר נקראת כלל השרשרת:

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

באופן מילולי – "ההסתברות ש-A וגם B קורים שווה להסתברות ש-A קורה כפול ההסתברות ש-B קורה בהינתן A".

ניזכר גם בנוסחת ההסתברות השלמה מהתרגול הקודם – בהינתן חלוקה $\Omega = \bigcup_{k=1}^{n} A_k$ ומאורע $B \in \mathcal{F}$ מתקיים:

$$\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}(A_k \cap B).$$

אם מתקיים $\mathbb{P}(A_k) > 0$ לכל k, נציב פנימה את הנוסחה שלמעלה ונקבל

$$\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}(A_k) \cdot \mathbb{P}(B \mid A_k).$$

נוסחה זו גם כן נקראת נוסחת ההסתברות השלמה.

תרגיל 2. 6 ילדים משחקים כדורגל בשכונה, וצריכים להחליט מי מהם יהיה שוער. יהושע (ילד שאיננו מעורב במשחק) מציע את הדרך הבאה להחליט מי יהיה שוער: יהושע יקח שישה גפרורים זהים, וישבור לאחד מהם את החלק התחתון. לאחר מכן, בזה אחר זה, הילדים ישלפו גפרורים ללא החזרה מתוך ידו של יהושע (כך שהם לא יכולים לדעת האם החלק התחתון של הגפרור שבור או לא). הילד שישלוף את הגפרור השבור יהיה שוער. האם זוהי דרך הוגנת לבחור שוער?

פתרון. על השאלה הזו נענה בלי לבנות מרחב מדגם מפורש. ראשית – מהי "דרך הוגנת לבחור שוער"? זוהי דרך בה ההסתברות לכל ילד להיות שוער שווה, כלומר $\frac{1}{6}$.

נסמן את הילדים לפי המיקום שלהם בשליפה (ילד 1, ילד 2 וכן הלאה). נחשב לכל ילד i את ההסתברות של המאורע A_i– המאורע שילד i ייבחר להיות שוער. ראשית $\mathbb{P}(A_1) = \frac{1}{6}$, שכן לילד הראשון יש בחירה של גפרור שבור 1 מתוך 6 גפרורים. כעת, $\Omega = A_1 \cup A_1^c$ ולכן לפי נוסחת ההסתברות השלמה:

$$\mathbb{P}(A_2) = \mathbb{P}(A_2|A_1)\mathbb{P}(A_1) + \mathbb{P}(A_2|A_1^c)\mathbb{P}(A_1^c).$$

מכיוון ש $\mathbb{P}(A_2|A_1) = 0$ (לא ייתכן שהילד השני ייבחר אם הילד הראשון כבר נבחר), וכן $\mathbb{P}(A_1^c) = \frac{5}{6}$ ו-$\mathbb{P}(A_2|A_1^c) = \frac{1}{5}$ (כי הילד השני בוחר גפרור אחד מתוך החמישה שנותרו), נקבל

$$\mathbb{P}(A_2) = \mathbb{P}(A_2|A_1^C)\mathbb{P}(A_1^C) = \frac{1}{5} \cdot \frac{5}{6} = \frac{1}{6}.$$

באופן דומה, עבור הילד השלישי נבחין ש $\Omega = (A_1 \cup A_2) \cup (A_1 \cup A_2)^c$. לכן

$$\mathbb{P}(A_3) = \mathbb{P}(A_3 \mid A_1 \cup A_2) \cdot \mathbb{P}(A_1 \cup A_2) + \mathbb{P}(A_3 \mid (A_1 \cup A_2)^c) \cdot \mathbb{P}((A_1 \cup A_2)^c).$$

המאורע $A_1 \cup A_2$ משמעותו שהילד הראשון נבחר או שהילד השני נבחר. לכן $\mathbb{P}(A_3|A_1 \cup A_2) = 0$. כמו כן, נשים לב ש-$A_1 \cup A_2$ הוא איחוד זר, לפיכך מסכימות

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

לכן $\mathbb{P}((A_1 \cup A_2)^c) = 1 - \frac{1}{3} = \frac{2}{3}$. לבסוף,

$$\mathbb{P}(A_3|(A_1 \cup A_2)^c) = \frac{1}{4},$$

כי לילד השלישי נותרו 4 גפרורים לבחור מתוכם. לכן,

$$\mathbb{P}(A_3) = \mathbb{P}(A_3 \mid (A_1 \cup A_2)^c) \cdot \mathbb{P}((A_1 \cup A_2)^c) = \frac{1}{4} \cdot \frac{2}{3} = \frac{1}{6}.$$

באופן דומה, ניתן לחשב כי לכל i, $\mathbb{P}(A_i) = \frac{1}{6}$ ולכן זו דרך הוגנת לבחור שוער, הידד ליהושע.

תרגיל 3. בדיקת פוליגרף משטרתית מגלה שקרן בהסתברות 0.8 ומגלה שאדם דובר אמת בהסתברות 0.9. ידוע שבהסתברות 0.7 הנבדק משקר.
מה ההסתברות שאדם יוכרז כ"דובר אמת" במכונה?

פתרון. נסמן ב-H_0 את המאורע שהאדם דובר אמת, ב-H_1 את המאורע שהוא דובר שקר, ב-H_0' את המאורע שהאדם הוכרז כדובר אמת וב-H_1' את המאורע שהאדם הוכרז כדובר שקר. לפי נוסחת ההסתברות השלמה, היות והאדם הוא דובר אמת או דובר שקר (כלומר $\Omega = H_0 \cup H_1$),

$$\mathbb{P}(H_0') = \mathbb{P}(H_0' \mid H_0) \cdot \mathbb{P}(H_0) + \mathbb{P}(H_0' \mid H_1) \cdot \mathbb{P}(H_1).$$

ידוע כי

$$\mathbb{P}(H_0' \mid H_0) = 0.9, \qquad \mathbb{P}(H_1' \mid H_1) = 0.8, \qquad \mathbb{P}(H_1) = 0.7.$$

מהסתברות מאורעות משלימים נסיק $\mathbb{P}(H_0) = 1 - \mathbb{P}(H_1) = 0.3$, וכן $\mathbb{P}(H_0' \mid H_1) = 1 - \mathbb{P}(H_1' \mid H_1) = 1 - 0.8 = 0.2$. לפיכך

$$\mathbb{P}(H_0') = 0.9 \cdot 0.3 + 0.2 \cdot 0.7 = 0.41.$$

במציאות, מה שהיינו רוצים לדעת הוא למעשה את $\mathbb{P}(H_0 \mid H_0')$, כלומר הסיכוי שאדם דובר אמת אם הוא הוכרז ככזה. אך זוהי ההתניה ההפוכה למידע שיש לנו! מה נעשה? נשתמש בחוק בייס.

3 חוק בייס

טענה (חוק בייס) יהא $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. עבור שני מאורעות A, B בעלי הסתברות חיובית (כלומר $\mathbb{P}(A) > 0, \quad \mathbb{P}(B) > 0$), מתקיים

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

חוק בייס נובע פחות או יותר מההגדרה, אכך ההשלכות שלו "בעולם האמיתי" יכולות להפתיע.

תרגיל 4. המשך השאלה הקודמת. מה הסיכוי שאדם דובר אמת אם מוכרז כי הוא דובר אמת?

פתרון. לפי חוק בייס נקבל:

$$\mathbb{P}\left(H_0 \mid H_0'\right) = \mathbb{P}\left(H_0' \mid H_0\right) \cdot \frac{\mathbb{P}\left(H_0\right)}{\mathbb{P}\left(H_0'\right)} = 0.9 \cdot \frac{0.3}{0.41} \approx 0.66.$$

תרגיל 5. סטודנט פותר שאלה במבחן אמריקאי בה יש ארבע תשובות אפשריות. אם הסטודנט אינו יודע את הפתרון, הוא בוחר תשובה באקראי. נתון שהסטודנט יודע בוודאות את התשובה ל-60% מהשאלות במבחן (כלומר ההסתברות שהסטודנט יודע את התשובה לשאלה במבחן היא 0.6).

1. תהי שאלה מהמבחן. מה ההסתברות שהסטודנט יענה עליה נכון?
2. נניח שידוע שהסטודנט ענה נכון על השאלה. מה ההסתברות שהוא ידע את התשובה מראש ולא ניחש?
3. נסמן ב-B את המאורע שהסטודנט יודע לענות על השאלה, וב-A את המאורע שהסטודנט ענה נכון על השאלה. לפי הנתון $\mathbb{P}(B) = 0.6$ וכן $\mathbb{P}(A|B) = 1$ ו-$\mathbb{P}(A|B^c) = \frac{1}{4}$. מכאן ההסתברות של המאורע A לפי נוסחת ההסתברות השלמה היא
$$\mathbb{P}(A) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c) = 1 \cdot 0.6 + \frac{1}{4} \cdot 0.4 = 0.7.$$
4. נציב בחוק בייס ונקבל
$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B) \cdot \mathbb{P}(B)}{\mathbb{P}(A)} = \frac{6}{7}.$$

תרגיל 6. נניח שקיימת מחלה נדירה D אשר שכיחותה באוכלוסייה היא $\gamma \ll 1$, תהי f בדיקה אשר מחזירה שלושה ערכים אפשריים:

- $f = 1$ – אין חשד למחלה
- $f = 2$ – חשד נמוך למחלה
- $f = 3$ – חשד גבוה למחלה

בנוסף, בדומה לדוגמא הקודמת נסמן ב-H_0 את המאורע שאדם הוא בריא וב-H_1 את המאורע שהוא חולה. נגדיר $\mathbb{P}(f=i|H_0)=a_i$ ו- $\mathbb{P}(f=i|H_1)=b_i$ עבור $i\in\{1,2,3\}$, ונתבונן בשני המבחנים הבאים לאיתור חולים.

1. המבחן המחמיר: מטופל מאובחן כחולה אם $f=3$, נקרא למבחן זה test_1
2. המבחן המקל: מטופל מאובחן כחולה אם $f\in\{2,3\}$ נקרא למבחן זה test_2

עבור כל מבחן חשבו, בהינתן שהמבחן מצא חשד כבד למחלה אצל נבדק, מהי ההסתברות שהוא בריא? ובהינתן שהמבחן מצא שאין חשד למחלה אצל נבדק, מהי ההסתברות שהוא חולה? הציבו $a_1=\frac{4}{5}, b_1=10^{-5}, a_3=\frac{1}{25}, b_3=0.9, \gamma=0.01$.

נסמן $\alpha_i=\mathbb{P}\left(H_0\middle|\text{test}_i=T\right), \beta_i=\mathbb{P}\left(H_1\middle|\text{test}_i=F\right)$, α_i נקראת שגיאה מסוג ראשון (שלילי כוזב) ו-β_i נקראת שגיאה מסוג שני (חיובי כוזב). מנוסחאת ההסתברות השלמה ברור כי $\mathbb{P}\left(f=i\right)=\gamma b_i+(1-\gamma)a_i$.

1. עבור המבחן הראשון, הטעות מסוג ראשון היא:

$$\begin{aligned}\mathbb{P}\left(H_0\middle|\text{test}_1=T\right)&=\mathbb{P}\left(H_0\middle|f=3\right)\\&=\frac{1-\gamma}{\mathbb{P}\left(f=3\right)}\cdot\mathbb{P}\left(f=3\middle|H_0\right)\\&=\frac{1-\gamma}{\gamma b_3+(1-\gamma)a_3}a_3\end{aligned}$$

והטעות מסוג שני היא:

$$\begin{aligned}\mathbb{P}\left(H_1\middle|\text{test}_1=F\right)&=\mathbb{P}\left(H_1\middle|f\in\{1,2\}\right)\\&=\frac{\gamma}{\mathbb{P}\left(f\in\{1,2\}\right)}\cdot\mathbb{P}\left(f\in\{1,2\}\middle|H_1\right)\\&=\frac{\gamma}{1-(\gamma b_3+(1-\gamma)a_3)}\cdot(1-b_3)\end{aligned}$$

נציב את הנתונים ונקבל ש-

$$\beta_1\approx 1.05\cdot 10^{-3}, \alpha_1\approx 0.81$$

2. עבור המבחן השני, הטעות מסוג ראשון היא:

$$\begin{aligned}\mathbb{P}\left(H_0\middle|\text{test}_2=F\right)&=\mathbb{P}\left(H_0\middle|f\in\{2,3\}\right)\\&=\frac{1-\gamma}{\mathbb{P}\left(f\in\{2,3\}\right)}\cdot\mathbb{P}\left(f\in\{2,3\}\middle|H_0\right)\\&=\frac{1-\gamma}{1-(\gamma b_1+(1-\gamma)a_1)}(1-a_1)\end{aligned}$$

והטעות מסוג שני היא:

$$\begin{aligned}\mathbb{P}\left(H_1\middle|\text{test}_2=F\right)&=\mathbb{P}(H_1|f=1)\\&=\frac{\gamma}{\mathbb{P}\left(f=1\right)}\cdot\mathbb{P}\left(f=1\middle|H_1\right)\\&=\frac{\gamma}{\gamma b_1+(1-\gamma)a_1}\cdot b_1\end{aligned}$$

אם נציב את הנתונים נקבל:

$$\beta_2 \approx 1.26 \cdot 10^{-7}, \alpha_2 \approx 0.95$$

נשים לב כי כשאנו משווים בין המבחן ניתן לראות שהשגיאה מסוג ראשון במבחן הראשון נמוכה יותר, אבל השגיאה מהסוג השני גבוהה יותר, במילים פשוטות הסיבה שזה קורה היא שבניסוי השני אנחנו מאבחנים קבוצה גדולה יותר של אנשים כחולים, ולכן יש סיכוי גבוה יותר שנאבחן אנשים בריאים כחולים אבל סיכוי נמוך יותר שנאבחן אנשים בריאים כחולים. סה"כ יש יתרונות וחסרונות לכל מבחן שכן כל אחד מקטין שגיאה אחת ומגדיל את השגיאה השנייה - יש "יחס המרה" (tradeoff) בין סוגי השגיאות.
אם היינו מעוניינים לערוך מבחן לאוכלוסייה הכללית (מבחן סקר) אז עדיף לבחור במבחן הראשון משום שבמבחן סקר אנו נרצה שהשגיאה מהסוג הראשון תהיה קטנה יותר (שכן רוב האוכלוסייה בריאה ולכן נרצה להקטין את ההסתברות שלנו לחיזוי חיובי כוזב ככל האפשר). לעומת זאת, אם היינו מעוניינים לערוך את המבחן לאוכלוסייה ספציפית שבה אחוז האנשים הבריאים נמוך יותר (מבחן אימות) אז עדיף לבחור במבחן השני מכיוון שבמבחן כזה נרצה להקטין את השגיאה מהסוג השני ככל האפשר (נרצה פחות תוצאות לחיזוי שלילי כוזב).

4 פרדוקס מונטי הול

שאלה: אתם בשעשועון טלויזיה כשבסופו אתם נדרשים לבחור פרס מאחורי אחד מבין שלושה וילונות. מאחורי אחד הוילונות ישנה מכונית חדשה ומאחורי הוילונות הנותרים שתי עיזים. בחרתם וילון אחד באקראי ובשלב זה מנחה התוכנית מצביע על אחד משני הוילונות הנותרים, חושף מאחוריו עז, ומאפשר לכם לבחור האם לשנות את בחירתכם. בהנחה שאתם מעדיפים לקבל מכונית על פני עז, כיצד עליכם לפעול?
(מה אתם חושבים? האם שווה תמיד להישאר בבחירה המקורית? האם שווה תמיד להחליף? האם זה לא משנה?)

פתרון: כפי שהצגנו את השאלה, ניתן לחשוב על כמה דרכים שונות לפרש אותה, שנבדלות בהתנהגותו של המנחה. התוצאה משתנה בהתאם לפרשנות.

1. **פרשנות 1:** איננו יודעים איך מתנהג המנחה. למשל, המנחה יכול להחליט לפתוח את הוילון האחר רק אם מאחורי החלון שבחרתם יש מכונית. למקרה זה (בו לא ידוע לנו כיצד המנחה מתנהג) לא נוכל להציע בקורס זה פתרון, שכן הוא נוגע יותר לפסיכולוגיה ופחות להסתברות.

2. **פרשנות 2:** המנחה **תמיד** פותח וילון שיש מאחוריו עז. אם מאחורי הדלת שבחרנו יש מכונית, הוא פותח באקראי את אחת משתי הדלתות הנותרות. אם מאחורי הדלת שבחרנו יש עז, הוא פותח את הדלת שמאחוריה נמצאת העז השניה. זו הפרשנות המקובלת לבעיה.
חישוב ההסתברויות: מאחר שכל הוילונות זהים נניח שבחרנו את וילון 1. ננתח את האפשרויות השונות, כל אחת מהן קורית בהסתברות $\frac{1}{3}$:

   (א) אילו המכונית מאחורי וילון 1, המנחה פותח את וילון 2 או 3. **לא כדאי להחליף**.

   (ב) אילו המכונית מאחורי וילון 2, המנחה פותח את וילון 3. **כדאי להחליף**.

   (ג) אילו המכונית מאחורי וילון 3, המנחה פותח את וילון 2. **כדאי להחליף**.

מניתוח זה ניתן לראות כי בשניים מהמקרים (שמתאימים להסתברות $\frac{2}{3}$, מפני שההסתברות אחידה) שינוי הבחירה יוביל לזכייה בפרס בעוד רק במקרה אחד (כלומר בהסתברות $\frac{1}{3}$) יוביל לויתור על הפרס. מכאן שעדיף לכם לשנות את בחירתכם ולהחליף וילון.
<u>הסבר נוסף:</u>

נניח בלי הגבלת הכלליות כי השחקן בחר בוילון מספר 1. נסמן $\Omega=\{(1,2),(1,3),(2,3),(3,2)\}$, כאשר האיבר הראשון בזוג מסמן את הוילון שמאחוריו המכונית והשני את הוילון שהמנחה פתח. נסמן ב-A_i את המאורע שהמכונית נמצאת מאחורי הוילון ה-i. מההנחה $\mathbb{P}(A_i)=\frac{1}{3}$ לכל i. כמו כן

$$\mathbb{P}(\{(1,2)\}\mid A_1)=\mathbb{P}(\{(1,3)\}\mid A_1)=\frac{1}{2},$$

כי אם השחקן בחר בוילון עם המכונית, המנחה פותח כל וילון אחר בסיכוי שווה. מכאן

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

כעת נדמיין את השחקן – נניח שהוא ראה את המנחה פותח את וילון 2, האם שווה להחליף? במילים אחרות, מה הסיכוי שהמכונית בוילון 3 אם וילון 2 נפתח? נחשב

$$\mathbb{P}(A_3\mid\{(\cdot,2)\})=\frac{\mathbb{P}(A_3\cap\{(\cdot,2)\})}{\mathbb{P}(\{(\cdot,2)\})}=\frac{\mathbb{P}(\{(3,2)\})}{\mathbb{P}(\{(3,2)\})+\mathbb{P}(\{(1,2)\})}=\frac{\frac{1}{3}}{\frac{1}{3}+\frac{1}{6}}=\frac{\frac{1}{3}}{\frac{1}{2}}=\frac{2}{3}.$$

החישוב יהיה זהה אם נניח שהמנחה פתח את וילון 3, ונשאל מה הסיכוי שהמכונית בוילון 2.

הערה: אם התוצאה עדיין לא אינטואיטיבית בעיניכם, חשבו על המקרה בו היו 100 וילונות. לאחר בחירת וילון אחד, המנחה היה חושף 98 וילונות אחרים שמאחוריהם עז, ושואל אם ברצונכם להחליף את הוילון שבחרתם בוילון האחרון שנותר סגור. מובן שצדקתם בבחירתם הראשונית בהסתברות $\frac{1}{100}$ בלבד, או במילים אחרות, המכונית **לא** נמצאת מאחורי הוילון שבחרתם בהסתברות $\frac{99}{100}$. מכאן נובע שבהסתברות $\frac{99}{100}$ המכונית נמצאת מאחורי הוילון שהמנחה בחר לא לפתוח, אז כמובן שכדאי לכם להחליף את הבחירה.

מבוא להסתברות וסטטיסטיקה 80430
תרגול 4

1 אי תלות

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. שני מאורעות A, B הם בלתי תלויים (נכתוב בדרך כלל ב״ת) אם מתקיים:

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

אי תלות של מאורעות לא גוררת שהם זרים, ואפילו להיפך! אם שני מאורעות A, B מהסתברות חיובית הם ב״ת, אז הם בהכרח לא זרים מכיוון ש-$\mathbb{P}(A \cap B) > 0$.

תרגיל בסיסי: שולפים קלף באופן מקרי מתוך חפיסת קלפים בת 52 קלפים. יהא E המאורע ״קיבלנו אס״ ויהא F המאורע ״קיבלנו לב״. האם אלו מאורעות ב״ת?

פתרון: נבחר את Ω להיות חפיסת הקלפים, ולכן $|\Omega| = 52$. מכיוון שיש בחבילה 4 קלפים מסוג אס, אז $|E| = 4$ ולכן $\mathbb{P}(E) = \frac{4}{52}$. באופן דומה, בחפיסה יש 13 קלפים מסוג לב ולכן $\mathbb{P}(F) = \frac{13}{52}$. מצד שני, המאורע ״קיבלנו אס מסוג לב״ שווה בדיוק למאורע $E \cap F$ וזהו מאורע בגודל 1 ולכן $\mathbb{P}(E \cap F) = \frac{1}{52}$. קיבלנו

$$\mathbb{P}(F) \cdot \mathbb{P}(E) = \frac{4}{52} \cdot \frac{13}{52} = \frac{1}{52} = \mathbb{P}(F \cap E)$$

ולכן המאורעות F ו-E הינם בלתי תלויים.

תרגיל: בכד נמצאים 3 מטבעות. שניים מהם הוגנים, ואילו על השלישי מסומן H (כלומר עץ) על שני הצדדים. נניח שבוחרים מטבע אחד באופן מקרי ומטילים אותו פעמיים. עבור $i = 1, 2$ נסמן את המאורע A_i להיות המאורע שבו יצא בהטלה ה-i עץ. האם המאורעות הנ״ל הם בלתי תלויים?

פתרון: כדי לענות על השאלה הזאת, נחשב את שני הביטויים בהגדרה בנפרד ונבדוק האם הם שווים. נסמן ב-D את המאורע ששלפנו מטבע הוגן מהכד. לפי הנתון $\mathbb{P}(D) = \frac{2}{3}$. כמו כן,

$$\mathbb{P}(A_1 | D) = \frac{1}{2}, \quad \mathbb{P}(A_1 | D^c) = 1.$$

לפי נוסחת ההסתברות השלמה,

$$\mathbb{P}(A_1) = \mathbb{P}(A_1|D)\mathbb{P}(D) + \mathbb{P}(A_1|D^c)\mathbb{P}(D^c) = \frac{1}{2} \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}.$$

נקבל בדיוק מאותו החישוב ש-$\mathbb{P}(A_2) = \frac{2}{3}$. מצד שני,

$$\begin{aligned}\mathbb{P}(A_1 \cap A_2) &= \mathbb{P}(A_1 \cap A_2 | D)P(D) + P(A_1 \cap A_2 | D^c)P(D^c) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} + 1 \cdot 1 \cdot \frac{1}{3} = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}.\end{aligned}$$

אבל $\frac{1}{2} < \frac{4}{9} = (\frac{2}{3})^2 = \mathbb{P}(A_1)\cdot\mathbb{P}(A_2)$. לכן, המאורעות הנ"ל תלויים זה בזה. שימו לב כי $\mathbb{P}\left(A_2 \mid A_1\right) = \frac{\mathbb{P}(A_1\cap A_2)}{\mathbb{P}(A_1)} = \frac{\frac{1}{2}}{\frac{2}{3}} = \frac{3}{4} > \frac{2}{3} = \mathbb{P}\left(A_2\right)$. אינטואיטיבית, אם בהטלה הראשונה התקבל עץ, זה מעלה את הסיכוי שהמטבע שהטלנו הוא המטבע המזויף, ולפיכך מעלה את הסיכוי שההטלה השנייה היא גם כן עץ.

תרגיל: שלושה שופטים מכריעים את גורלו של נאשם על פי דעת רוב. שניים מהשופטים מנוסים ומזהים נכונה את אשמתו של הנאשם בסיכוי 90%. האחרון אינו מנוסה ומזהה נכונה את אשמתו של הנאשם בסיכוי 51%. השוו בין ההסתברויות לפסק דין נכון במקרים הבאים:

1. החלטת כל שופט בלתי-תלויה.
2. השופטים המנוסים מחליטים באופן בלתי תלוי והשופט שאינו מנוסה בוחר אחד מהם באקראי ומצטרף לדעתו.

פתרון: נסמן ב-A את המאורע בו התקבל פסק דין נכון וב-A_i את המאורע בו השופט ה-i פסק נכון, כאשר השופט הבלתי מנוסה הוא השלישי. A יתקיים אם לפחות שניים מבין $\{A_1,A_2,A_3\}$ יתקיימו. עוד דרך לומר זאת היא ששני השופטים המנוסים יפסקו נכון, או שבדיוק אחד מבין המנוסים יפסוק נכון, וגם הבלתי מנוסה. כלומר:

$$A = (A_1\cap A_2)\cup(A_1^c\cap A_2\cap A_3)\cup(A_1\cap A_2^c\cap A_3)$$

וזהו איחוד זר.

נניח ראשית את המקרה הראשון בו <u>החלטת כל שופט היא בלתי תלויה</u>. נחשב:

$$\begin{aligned}\mathbb{P}\left(A_1\cap A_2\right) &= \mathbb{P}\left(A_1\right)\cdot\mathbb{P}\left(A_2\right) = 0.9^2,\\ \mathbb{P}\left(A_1\cap A_2^c\cap A_3\right) &= \mathbb{P}\left(A_1^c\cap A_2\cap A_3\right) = 0.9\cdot 0.1\cdot 0.51.\end{aligned}$$

סה"כ:

$$\mathbb{P}(A) = (0.9)^2 + 2\cdot 0.9\cdot 0.1\cdot 0.51 = 0.9018.$$

כעת נבדוק מה קורה במקרה השני בו השופט שאינו מנוסה מצטרף באקראי לאחת הדעות:

$$\begin{aligned}\mathbb{P}\left(A_1\cap A_2\right) &= \mathbb{P}\left(A_1\right)\cdot\mathbb{P}\left(A_2\right) = 0.9^2,\\ \mathbb{P}\left(A_1\cap A_2^c\cap A_3\right) &= \mathbb{P}\left(A_1^c\cap A_2\cap A_3\right) = \mathbb{P}\left(A_1\cap A_2^c\right)\cdot\mathbb{P}\left(A_3\mid A_1\cap A_2^c\right) = 0.9\cdot 0.1\cdot 0.5.\end{aligned}$$

סה"כ

$$\mathbb{P}\left(A\right) = 0.9^2 + 2\cdot 0.9\cdot 0.1\cdot 0.5 = 0.9,$$

כלומר עדיף שהשופטים יכריעו באופן בלתי תלוי!

הערה: הבדל יצא כל כך קטן מכיוון שהחלטת השופט הבלתי מנוסה כמעט שקולה להטלת מטבע גם אם הוא מחליט לבד. ככל שהניסיון שלו עולה (ההסתברות לתת פס"ד נכון גבוהה יותר) העדיפות לאי-תלות נראית הרבה יותר מובהקת.

2 אי-תלות בין מספר מאורעות

אוסף מאורעות $A_1,...,A_n$ נקרא <u>אוסף מאורעות בלתי תלויים</u> אם **לכל** תת אוסף $A_{i_1},...,A_{i_k}$ מתקיים

$$\mathbb{P}\left(A_{i_1}\cap A_{i_2}\cap\ldots\cap A_{i_k}\right) = \mathbb{P}\left(A_{i_1}\right)\mathbb{P}\left(A_{i_2}\right)\cdot\ldots\cdot\mathbb{P}\left(A_{i_k}\right).$$

ובאופן דומה אוסף מאורעות אינסופי $\mathcal{A}=\{A_i\}_{i\in I}$ ייקרא אוסף מאורעות בלתי תלויים אם לכל תת אוסף סופי $A_{i_1},...,A_{i_k}$ תתקיים המשוואה הנ"ל.

אוסף מאורעות ייקרא בלתי תלוי בזוגות אם לכל שני מאורעות $A_i,A_j\in\mathcal{A}$ יתקיים

$$\mathbb{P}(A_i\cap A_j)=\mathbb{P}(A_i)\cdot\mathbb{P}(A_j).$$

ראיתם בכיתה כי אי-תלות בזוגות זו תכונה **חלשה יותר** מאשר אי תלות.

תרגיל: נביט בניסוי של הטלת N קוביות הוגנות בזו אחר זו. לכל $1\le i\le N$ נסמן ב-$x_i\in[6]$ את תוצאת ההטלה ה-i. הוכיחו כי אוסף המאורעות

$$1\le n\le N\quad A_n:=\left\{\sum_{i=1}^{n}x_i=0 \bmod 6\right\}$$

הוא אוסף מאורעות בלתי תלויים.

פתרון: צריך להוכיח כי לכל תת-אוסף סופי $A_{n_1},...,A_{n_k}$ מתקיים

$$\mathbb{P}(A_{n_1}\cap...\cap A_{n_k})=\mathbb{P}(A_{n_1})\cdot...\cdot\mathbb{P}(A_{n_k}).$$

ראשית נחשב את ההסתברות של המאורעות הבודדים. לכל $1\le n\le N$ מתקיים

$$A_n=\left\{\sum_{i=1}^{n}x_i=0 \bmod 6\right\}=\left\{x_n=-\sum_{i=1}^{n-1}x_i \bmod 6\right\}.$$

פונקצית ההסתברות על x_n היא אחידה, לכן ההסתברות לקבל בדיוק את המספר $-\sum_{i=1}^{n-1}x_i \bmod 6$ היא $\frac{1}{6}$. אזי לכל $1\le n\le N$ מצאנו $\mathbb{P}(A_n)=\frac{1}{6}$. כעת נחשב את ההסתברות לחיתוך מאורעות. נניח בלי הגבלת הכלליות ש-$n_1<n_2<...<n_k$. אזי

$$\begin{aligned}A_{n_1}\cap...\cap A_{n_k}&=\left\{\sum_{i=1}^{n_1}x_i=0 \bmod 6 \text{ and } \sum_{i=1}^{n_2}x_i=0 \bmod 6 \text{ and ... and } \sum_{i=1}^{n_k}x_i=0 \bmod 6\right\}\\&=\left\{\sum_{i=1}^{n_1}x_i=0 \bmod 6 \text{ and } \sum_{i=n_1+1}^{n_2}x_i=0 \bmod 6 \text{ and ... and } \sum_{i=n_{k-1}+1}^{n_k}x_i=0 \bmod 6\right\}\\&=\left\{x_{n_1}=-\sum_{i=1}^{n_1-1}x_i \bmod 6\right\}\cap\left\{x_{n_2}=-\sum_{i=n_1+1}^{n_2-1}x_i \bmod 6\right\}\cap...\cap\left\{x_{n_k}=-\sum_{i=n_{k-1}+1}^{n_k-1}x_i \bmod 6\right\}.\end{aligned}$$

כאן החיתוך הוא של מאורעות בלתי תלויים משום שתוצאת כל הטלת קובייה היא בלתי-תלויה בהטלות האחרות. יחד עם העובדה שכל תוצאה של הטלה מתקבלת בהסתברות $\frac{1}{6}$, נסיק

$$\begin{aligned}\mathbb{P}(A_{n_1}\cap...\cap A_{n_k})&=\mathbb{P}\left(\left\{x_{n_1}=-\sum_{i=1}^{n_1-1}x_i \bmod 6\right\}\right)\cdot...\cdot\mathbb{P}\left(\left\{x_{n_k}=-\sum_{i=n_{k-1}+1}^{n_k-1}x_i \bmod 6\right\}\right)\\&=\frac{1}{6}\cdot...\cdot\frac{1}{6}=\mathbb{P}(A_{n_1})\cdot...\cdot\mathbb{P}(A_{n_k})\end{aligned}$$

כנדרש ■.

תרגיל: מטילים עשרה מטבעות הוגנים. נסמן בA_i את המאורע שבמטבע ה-i התקבלה תוצאה של עץ ונסמן ב-B את המאורע שמספר הטלות שהתקבל בהן עץ בסך הכל היה זוגי. נראה כי לכל $I \subsetneq [10]$ המאורעות $\{B\} \cap \{A_i : i \in I\}$ בלתי תלויים, אך המאורעות $\{B\} \cap \{A_i : i \in [10]\}$ תלוים.

פתרון: מרחב המדגם המתאים הוא $\{0,1\}^{10}$ (כאשר 0 מייצג עץ), ופונקציית ההסתברות היא האחידה. עלינו להראות כי לכל $I \subsetneq [10]$

$$\mathbb{P}\left(B \cap \bigcap_{i \in I} A_i\right) = \mathbb{P}(B) \cdot \prod_{i \in I} \mathbb{P}(A_i).$$

ראשית נשים לב ש-$|B| = |B^c|$, מאחר שישנה העתקה חד-חד-ערכית ועל בין B ל-B^c (ישנן מספר דרכים להגדיר העתקה כזו, למשל ההעתקה π המוגדרת למטה). לכן $\mathbb{P}(B) = \mathbb{P}(B^c) = \frac{1}{2}$. כעת תהא $I \subsetneq [10]$. מאחר ו-$I \neq [10]$, קיים $j \in [10] \setminus I$, והפונקציה המתאימה

$$\begin{aligned}
&\pi : B \cap \bigcap_{i \in I} A_i \to B^c \cap \bigcap_{i \in I} A_i, \\
&\pi(\omega_1, ..., \omega_{10}) = (\omega_1, ..., \omega_{j-1}, 1 - \omega_j, ..., \omega_{10})
\end{aligned}$$

הינה פונקציה חד-חד-ערכית ועל! כלומר $\left|B \cap \bigcap_{i \in I} A_i\right| = \left|B^c \cap \bigcap_{i \in I} A_i\right|$ ולכן.

$$\mathbb{P}\left(B \cap \bigcap_{i \in I} A_i\right) = \mathbb{P}\left(B^c \cap \bigcap_{i \in I} A_i\right).$$

מנוסחאת ההסתברות השלמה

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \mathbb{P}\left(B \cap \bigcap_{i \in I} A_i\right) + \mathbb{P}\left(B^c \cap \bigcap_{i \in I} A_i\right),$$

כלומר קיבלנו

$$\mathbb{P}\left(B \cap \bigcap_{i \in I} A_i\right) = \frac{1}{2}\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \mathbb{P}(B) \cdot \mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \mathbb{P}(B) \cdot \prod_{i \in I} \mathbb{P}(A_i),$$

כאשר בשלב האחרון השתמשנו בכך שהקבוצות A_i בלתי תלויות. כעת, נשים לב ש-$\bigcap_{i \in [10]} A_i \subset B$ ולכן

$$\mathbb{P}\left(B \mid \bigcap_{i \in [10]} A_i\right) = 1 \neq \mathbb{P}(B),$$

כלומר המאורעות $\{A_1, ..., A_{10}, B\}$ תלויים.

3 משתנים מקריים

יהא $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות בדיד. משתנה מקרי (מ״מ) הוא פונקציה $X : \Omega \to \mathbb{R}$ ממרחב המדגם לממשיים. משתנה מקרי הינה ״סתם״ פונקציה $X : \Omega \to \mathbb{R}$ (בהמשך ניתן תנאי במרחבים כלליים, אך גם הוא יראה ״מיותר״), אז ההגדרה על פניו נראת מיותרת. למעשה ההגדרה מצהירה על הכוונות שלנו : להשתמש במשתנה מקרי על מנת להגדיר פונקציית הסתברות על $\mathbb{R}$ שלה קשר הדוק לפונקציית ההסתברות המקורית על Ω (״דוחפים״ את $\mathbb{P}$ על $\mathbb{R}$ באמצעות X). עד כה, כשרצינו לפתור בעיה הסתברותית, בנינו לה מרחב מדגם וחישבנו את ההסתברויות של מאורעות באותו המרחב, ומשתנים מקריים יאפשרו לנו לחשב את ההסתברויות ב-$\mathbb{R}$ במקום. כך נוכל להשתמש בתכונות של המספרים הממשיים שאנו כבר מכירים מאינפי 1-2 כדי להגיע לתוצאות על מרחבי הסתברות כלליים.

3.1 דוגמאות למ"מ:

1. פוגשים באקראיות אדם ברחוב ורוצים לדעת מה ההסתברות שגובהו מעל 1.70 מטר. במודל ההסתברותי, Ω הוא אוסף כל האפשרויות לפגישת אדם ברחוב (נאמר – כל בני האדם האפשריים בעולם), $\mathbb{P}$ תהיה איזו פונקציית הסתברות שתתאר נכון את מי יותר סביר שנפגוש ואת מי לא. בשביל הגובה, נגדיר $X:\Omega\to\mathbb{R}$ על ידי:

$$X(\omega)=\text{ Height of }\omega$$

והשאלה הומרה לחישוב ההסתברות של המאורע $\{X>1.70\}$, כלומר:

$$\mathbb{P}(\{X>1.70\})=\mathbb{P}(\{\omega\in\Omega:X(\omega)>1.70\}).$$

2. מטילים פעמיים קוביה, מה ההסתברות שתוצאת ההטלה הראשונה גדול מערך ההטלה השניה? אם $(\Omega,\mathcal{F},\mathbb{P})$ הוא מרחב ההסתברות שמתאים לשאלה, עבור $i=1,2$ נגדיר את X_i להיות הערך שיצא בהטלה ה–i:

$$X_i(\omega)=\text{Output of the }i^{th}\text{ roll in }\omega$$

והשאלה הומרה לחישוב הערך $\mathbb{P}(X_1>X_2)$. לדוגמא, אם בחרנו את מרחב ההסתברות להיות כמו שאנו פותרים בדרך כלל שאלה מסוג זה – $\Omega=\{1,...,6\}^2$, אז נוכל להציג במפורש את הפונקציות X_i:

$$X_1(a,b)=a$$
$$X_2(a,b)=b.$$

3.2 התפלגות של משתנה מקרי

נבחין כי לא עשינו שימוש בפונקציית ההסתברות הנתונה $\mathbb{P}$ בהגדרת המשתנה המקרי. זה עלול להשאיר את הרושם השגוי שאם נחליף את פונקציית ההסתברות $\mathbb{P}$ בפונקציית הסתברות אחרת על Ω נקבל את אותו המשתנה המקרי. מייד נראה כי אמנם דבר זה אינו משפיע על המשתנה המקרי כפונקציה מ–Ω ל–$\mathbb{R}$, אבל הוא משפיע באופן דרמטי על **ההתפלגות** של המשתנה המקרי שבה נדון מייד. יהי X משתנה מקרי על מרחב הסתברות $(\Omega,\mathbb{P})$. עבור קבוצה $E\subset\mathbb{R}$ נשתמש בקיצור

$$.\{X\in E\}=X^{-1}(E)=\{\omega\in\Omega:X(\omega)\in E\}.$$

ההתפלגות של X היא פונקציית הסתברות חדשה על המרחב $\mathbb{R}$, המוגדרת על ידי

$$.\forall E\subset\mathbb{R}\ ,\ \mathbb{P}_X(E)=\mathbb{P}(X\in E).$$

אם $\mathbb{P}_X$ היא פונקציית הסתברות בדידה, נאמר כי X הוא **משתנה מקרי בדיד**. אם X משתנה מקרי בדיד נגדיר את **התומך** של X להיות הקבוצה בת-המנייה

$$\text{Supp}(X)=\{x\in\mathbb{R}\ :\ p_X(x)>0\}.$$

עד עתה הדרך הפורמלית לחשוב על בעיות הסתברותיות הייתה לבנות מרחב הסתברות מתאים. משתנים מקריים מאפשרים פרספקטיבה שונה בה אנו לא מתעניינים במרחב ההסתברות עצמו, אלא בהתפלגות של פונקציה ממנו אל המספרים הממשיים. יהיו $(\Omega_1,\mathcal{F}_1,\mathbb{P}_1)$, $(\Omega_2,\mathcal{F}_2,\mathbb{P}_2)$ שני מרחבי הסתברות (זהים או שונים). עבור שני משתנים מקריים $X:\Omega_1\to\mathbb{R},\ Y:\Omega_2\to\mathbb{R}$, נאמר שהם **שווי התפלגות** אם $\mathbb{P}_X\equiv\mathbb{P}_Y$ (שוויון פונקציות ההתפלגות). במקרה כזה נסמן $X\overset{\mathrm{d}}{=}Y$ (d בשביל distribution).

תרגיל: בוחרים פעמיים מספר ב–$[2]$, בזה אחר זה ובאופן בלתי תלוי. נתבונן במשתנה המקרי שתוצאתו היא סכום המספרים שבחרנו. זהו משתנה מקרי שמקבל את הערכים $2,3,4$ (באופן פורמלי, הכוונה היא כי ישנה הסתברות חיובית שתוצאת המשתנה המקרי תהיה $2,3,4$ ובמובלע ההסתברות לקבל כל ערך אחר $x \in \mathbb{R} \setminus \{2,3,4\}$ היא אפס).

כעת נראה כיצד ניתן להציג משתנה מקרי זה בצורות שונות. נזכור כי כפי שכבר ראינו הניסוי המתואר ניתן להצגה באמצעות שני מודלים הסתברותיים שונים.

1. מודל ראשון: $\Omega_1 = [2]^2$ עם $\mathbb{P}_1$ פונקציית הסתברות אחידה. במודל זה $X : \Omega_1 \to \mathbb{R}$ מוגדר על ידי
$$X(i,j) = i + j.$$
כעת נובע כי
$$\mathbb{P}_X(2) = \mathbb{P}_1(X=2) = \mathbb{P}_1(\{(1,1)\}) = 1/4,$$
ובאופן דומה גם $\mathbb{P}_X(4) = 1/4$. לעומת זאת,
$$\mathbb{P}_X(3) = \mathbb{P}_1(X=3) = \mathbb{P}_1(\{(1,2),(2,1)\}) = 2/4 = 1/2.$$

2. מודל שני: נסמן ב-x_i את מספר הפעמים שבחרנו i. $\Omega_2 = \left\{(x_1,x_2) \in \{0,1,2\}^2 : x_1 + x_2 = 2\right\}$ עם פונקציית ההסתברות הבאה
$$\mathbb{P}_2(0,2) = \mathbb{P}_2(2,0) = 1/4\,, \quad \mathbb{P}_2(1,1) = 1/2.$$
במודל זה $Y : \Omega_2 \to \mathbb{R}$ מוגדר על ידי
$$Y(2,0) = 2\,, \quad Y(0,2) = 4\,, \quad Y(1,1) = 3.$$
כעת נובע כי
$$\mathbb{P}_Y(2) = \mathbb{P}_2(Y=2) = \mathbb{P}_2(2,0) = 1/4,$$
ובאופן דומה גם $\mathbb{P}_Y(4) = 1/4$. לעומת זאת,
$$\mathbb{P}_Y(3) = \mathbb{P}_2(Y=3) = \mathbb{P}_2(1,1) = 1/2.$$

הפונקציה Y במודל השני הוגדרה באופן שונה מהפונקציה X במודל הראשון, וברור כי הן לא שוות כפונקציות ממשיות. עם זאת, מתקיים $X \overset{\mathrm{d}}{=} Y$.

תרגיל: מטילים 2 קוביות הוגנות. יהי N הפרש הקוביות (הראשונה פחות השנייה). מה ההתפלגות של N?

נשים לב שבשלב השאלה, המספר N מופיע כמספר לא ידוע שנקבע באקראיות. במודל המתמטי, המספר הזה מתואר כפונקציה $N : \Omega \to \mathbb{R}$ ממרחב המדגם לממשיים, לכן יש היגיון בלדבר על התפלגותו בהקשר של משתנים מקריים.

פתרון: יהי $\Omega = \{1,...,6\}^2$, $\mathbb{P}$ התפלגות אחידה. כל איבר ב–Ω יסומן ב-$\omega = (\omega_1,\omega_2)$. נסמן ב-$X_i$ את התוצאה בהטלת הקוביה ה–i, ואת N נגדיר להיות הפרשם:
$$\begin{aligned} X_1(\omega) &= \omega_1 \\ X_2(\omega) &= \omega_2 \\ N(\omega) &= X_1(\omega) - X_2(\omega) \end{aligned}$$

פתרון לשאלה שקול למציאת פונקציית ההתפלגות של N:

$$p_N(0) = \mathbb{P}(N=0) = \mathbb{P}\left(\{(n,n) : n \in [6]\}\right) = \frac{1}{6}$$
$$p_N(1) = \mathbb{P}(N=1) = \mathbb{P}(\{(n+1,n) : n \in [5]\}) = \frac{5}{36}$$

ובאופן כללי

$$\forall k \geq 0 \; p_N(k) = \mathbb{P}(\{(n+k,n) : n \in [6-k]\}) = \frac{6-k}{36}$$

ועבור $k < 0$ נקבל $p_N(k) = \frac{6+k}{36}$.

3.3 שוויון כמעט תמיד של משתנים מקריים

יהא $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, ו-$X, Y : \Omega \to \mathbb{R}$ משתנים מקריים. נאמר כי X שווה ל-Y כמעט תמיד אם

$$\mathbb{P}\left(\{X=Y\}\right) = \mathbb{P}\left(\{\omega \in \Omega : X\left(\omega\right) = Y\left(\omega\right)\}\right) = 1$$

במקרה זה נסמן $X \overset{\text{a.s.}}{=} Y$ (מלשון $X = Y$ almosth surely). אם X, Y משתנים מקריים ששווים כמעט תמיד, אז X, Y שווי התפלגות. תהא $E \subset \mathbb{R}$. לפי ההגדרה של ההתפלגויות מתקיים

$$\mathbb{P}_X\left(E\right) - \mathbb{P}_Y\left(E\right) = \mathbb{P}\left(\{X \in E\}\right) - \mathbb{P}\left(\{Y \in E\}\right)$$

נסמן $A = \{X \in E\}, B = \{Y \in E\}$ ואז

$$\mathbb{P}\left(A\right) - \mathbb{P}\left(B\right) = \mathbb{P}\left(\left(A \backslash B\right) \cup A \cap B\right) - \mathbb{P}\left(\left(B \backslash A\right) \cup A \cap B\right) =$$
$$\mathbb{P}\left(\left(A \backslash B\right)\right) + \mathbb{P}\left(A \cap B\right) - \mathbb{P}\left(\left(B \backslash A\right)\right) - \mathbb{P}\left(A \cap B\right) = \mathbb{P}\left(\left(A \backslash B\right)\right) - \mathbb{P}\left(\left(B \backslash A\right)\right)$$

כעת נשים לב:

$$A \backslash B = \{X \in E \wedge Y \notin E\} \subset \{X \neq Y\}$$
$$B \backslash A = \{Y \in E \wedge X \notin E\} \subset \{X \neq Y\}$$

והרי $\mathbb{P}\left(\{X \neq Y\}\right) = 0$ ולכן ממונוטוניות נובע כי $\mathbb{P}\left(\left(A \backslash B\right)\right) = \mathbb{P}\left(\left(B \backslash A\right)\right) = 0$ וסיימנו. (שימו לב כי שוויון כמעט תמיד אינו גורר שוויון).

דוגמא: תוצאות ההטלה של שתי קוביות הוגנות שונות הן שוות התפלגות (כי הן מקבלות כל תוצאה בהסתברות זהה), אך הן אינן שוות כמעט-תמיד (מפני שבהסתברות חיובית מתקבלת תוצאה שונה בכל קוביה).

דוגמא לקריאה בלבד: יהא $\Omega = [10]^2$ מרחב המדגם, ונגדיר את פונקציית ההסתברות הנקודתית

$$\forall \left(m,n\right) \in \Omega \; p\left(m,n\right) = \begin{cases} \frac{1}{10m} & n \leq m \\ 0 & \text{else} \end{cases}$$

כלומר מרחב ההסתברות מתאר את הניסוי הדו שלבי הבא: מגרילים באופן אחיד מספר m בין 1 ל-10, ואז מגרילים באופן אחיד מספר n בין 1 ל-m.

נגדיר את שני המשתנים המקריים הבאים:

$$X\left(m,n\right) = m$$
$$Y\left(m,n\right) = \max\left\{m,n\right\}$$

נשים לב כי

$$\mathbb{P}\left(X=Y\right)=\mathbb{P}\left(\left\{(m,n)\in\Omega:n\leq m\right\}\right)=1$$

ולכן המשתנים המקריים הללו שווים כמעט תמיד, אך אינם שווים.

מבוא להסתברות וסטטיסטיקה 80430
תרגול 5

1 תזכורת להתפלגויות

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות.

- נאמר שמ״מ X המוגדר על Ω **מתפלג ברנולי** עם פרמטר p (ונסמן $X \sim \mathrm{Ber}(p)$), אם מתקיים:

$$p_X(1) = p$$
$$p_X(0) = 1 - p$$

- נאמר שמשתנה מקרי X **מתפלג אחיד** על קבוצה סופית $S \subset \mathbb{R}$ אם $\mathbb{P}_X(\{t\}) = \frac{1}{|S|}$ לכל $t \in S$. במקרה זה נרשום $X \sim \mathrm{Unif}(S)$.

2 וקטורים מקריים

אוסף סופי של משתנים מקריים $X = (X_1, ..., X_n)$ המוגדרים על אותו מרחב הסתברות נקרא **וקטור מקרי**. עבור וקטור מקרי $X = (X_1, ..., X_n)$ על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$, הפונקציה $\mathbb{P}_X = \mathbb{P}_{X_1,...,X_n} : \mathcal{F}_{\mathbb{R}^n} \to [0,1]$ המוגדרת על ידי

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

נקראת **ההתפלגות המשותפת** של $X = (X_1, ..., X_n)$. ההתפלגויות של המשתנים המקריים $X_1, ..., X_n$ נקראות **התפלגויות שוליות**.

דוגמא: מטילים 2 קוביות הוגנות. נסמן ב-X את סכום הקוביות וב-Y את הפרש הקוביות (ראשונה פחות שנייה). ההתפלגויות השוליות של (X, Y) ניתנות על ידי

$$\forall k \in \{2, ..., 12\} \ \mathbb{P}_X(\{k\}) = \frac{6 - |7-k|}{36}$$
$$\forall m \in \{-5, ..., 5\} \ \mathbb{P}_Y(\{k\}) = \frac{6 - |m|}{36}$$

ההתפלגות המשותפת ניתנת על ידי: לכל $(x,y) \in \mathbb{R}^2$ כך ש-$(\frac{x+y}{2}, \frac{x-y}{2}) \in [6]^2$.

$$\mathbb{P}_{X,Y}(\{x,y\}) = \mathbb{P}\left(\left\{\frac{x+y}{2}, \frac{x-y}{2}\right\}\right) = \frac{1}{36}$$

ולכל נקודה אחרת ההסתברות היא אפס.

תרגיל: מטילים 3 קוביות הוגנות, נסמן ב-X_i את המשתנה המקרי שמחזיר את התוצאה בקוביה ה-i. נגדיר את הוקטור המקרי $X = (X_1, \frac{X_1+X_2}{2}, \frac{X_1+X_2+X_3}{3})$. נסמן ב- $S \subseteq \mathbb{R}^3$ את הקבוצה $S = \{(a,a,a) : a \in \mathbb{R}\}$. מה ההסתברות למאורע $\{X \in S\}$?

פתרון: מרחב המדגם הינו $\Omega = [6]^3$, ופונקציית ההסתברות $\mathbb{P}$ אחידה.

$$\mathbb{P}_X(S) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in S\}) = \mathbb{P}(\{(x_1,x_2,x_3) \in \Omega : x_1 = \frac{x_1+x_2}{2} = \frac{x_1+x_2+x_3}{3}\})$$
$$= \mathbb{P}(\{(x_1,x_2,x_3) \in \Omega : x_1 = x_2 = x_3\}) = \mathbb{P}(\{(a,a,a) \in \Omega : a \in [6]\}) = \frac{6}{6^3} = \frac{1}{36}.$$

תרגיל: נמשיך עם הניסוי מהתרגיל הקודם, אך הפעם נסתכל על הוקטור המקרי $Y = (X_1, X_1+X_2, X_1+X_2-X_3)$. נסמן ב- $T \subseteq \mathbb{R}^3$ את הקבוצה $T = \{(a, 2a+1, 0) : a \in \mathbb{R}\}$. מה ההסתברות למאורע $\{Y \in T\}$?

פתרון: שוב מרחב המדגם הינו $\Omega = [6]^3$, ופונקציית הסתברות $\mathbb{P}$ אחידה.

$$\begin{aligned}\mathbb{P}_Y(T) &= \mathbb{P}(Y^{-1}(T)) \\ &= \mathbb{P}(\{(y_1,y_2,y_3) \in \Omega : y_1+y_2 = 2y_1+1, y_1+y_2 = y_3\}) \\ &= \mathbb{P}(\{(y_1,y_2,y_3) \in \Omega : y_2 = y_1+1, 2y_1+1 = y_3\}) = \mathbb{P}(\{(1,2,3), (2,3,5)\}) = \frac{2}{6^3}.\end{aligned}$$

3 התפלגות בהינתן מאורע

יהי X מ"מ על מרחב ההסתברות $(\Omega, \mathcal{F}, \mathbb{P})$, ויהי $A \in \mathcal{F}$ מאורע עם הסתברות חיובית. נסמן ב– $\mathbb{P}_A = \mathbb{P}((\cdot) \mid A)$ את פונקציית ההסתברות המותנית על Ω:

$$\forall B \in \mathcal{F} \; \mathbb{P}_A(B) = \mathbb{P}(B \mid A) := \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$$

בנוסף נסמן ב–$(X|A)$ את הפונקציה $X : \Omega \to \mathbb{R}$ כמשתנה מקרי על מרחב ההסתברות $(\Omega, \mathcal{F}, \mathbb{P}((\cdot) \mid A))$. שימו לב שכפונקציה $X|A$ הוא אותה הפונקציה. עם זאת, כמו בהגדרה של משתנה מקרי, אנחנו מצהירים שבכוונתנו להתבונן בהתפלגות של X ביחס לפונקציית ההסתברות $\mathbb{P}((\cdot) \mid A)$, כפי שנראה בהגדרה הבאה. עם זאת, לא ניתן לשאול למשל האם $X, X|A$ שווים כמעט תמיד, מאחר והם מוגדרים במרחבי הסתברות שונים. יהי X מ"מ על מרחב ההסתברות $(\Omega, \mathcal{F}, \mathbb{P})$, ויהי $A \in \mathcal{F}$ מאורע עם הסתברות חיובית. ההתפלגות של $X|A$ (נקראת גם "ההתפלגות של X בהינתן A") היא:

$$\forall E \subset \mathbb{R} \; \mathbb{P}_{X|A}(E) = \mathbb{P}(\{X \in E\} \mid A)$$

תרגיל: מטילים שלושה מטבעות מוטים עם פרמטר p בזה אחר זה, ונסמן ב-X_1, X_2, X_3 את תוצאות ההטלה של המטבעות בהתאמה
(בשפה הסתברותית ניתן לחשוב על X_1, X_2, X_3 כעל שלושה משתני ברנולי עם פרמטר p שהינם בלתי תלויים). כיצד מתפלג וקטור ההטלות $X = (X_1, X_2, X_3)$ אם ידוע כי $X_1 + X_2 + X_3 = 2$?

פתרון: נגדיר את המאורע A על ידי:

$$A = \{X_1 + X_2 + X_3 = 2\} = \{(x_1,x_2,x_3) : x_i \in \{0,1\}, x_1+x_2+x_3 = 2\}.$$

נשים לב ש-$A=\{(1,1,0),(1,0,1),(0,1,1)\}$ ושההסתברות למאורע A היא $3p^2(1-p)$. כעת, לכל $(x_1,x_2,x_3)\in A$, מתקיים:

$$\mathbb{P}_{X|A}(X=(x_1,x_2,x_3))=\mathbb{P}(X=(x_1,x_2,x_3)|A)=\mathbb{P}_A(X=(x_1,x_2,x_3))=\frac{\mathbb{P}((x_1,x_2,x_3)\cap A)}{\mathbb{P}(A)}$$

$$=\frac{\mathbb{P}(x_1,x_2,x_3)}{\mathbb{P}(A)}=\frac{p^2(1-p)}{3p^2(1-p)}=\frac{1}{3}.$$

כלומר $X|A$ מתפלג אחיד ואינו תלוי בערך של p.

תרגיל: מטילים מטבע הוגן. אם יצא 0 מטילים שוב מטבע הוגן ואם יצא 1 מטילים מטבע מוטה בעל הסתברות p לקבל 1. מהי התפלגות ההטלה השנייה?

פתרון: יהי $X\sim\text{Ber}(1/2)$ תוצאת ההטלה הראשונה ויהי Y תוצאת ההטלה השנייה. מהנתון ידוע כי

$$Y\mid\{X=0\}\sim\text{Ber}(1/2)\ ,\ \ Y\mid\{X=1\}\sim\text{Ber}(p).$$

מנוסחת ההסתברות השלמה נקבל כי

$$\begin{aligned}\mathbb{P}(Y=1)&=\mathbb{P}(Y=1\mid X=0)\,\mathbb{P}(X=0)+\mathbb{P}(Y=1\mid X=1)\,\mathbb{P}(X=1)\\&=1/2\cdot 1/2+p\cdot 1/2=1/4+p/2\end{aligned}$$

כלומר $Y\sim\text{Ber}(1/4+p/2)$.

דוגמא: נטיל שלושה מטבעות הוגנים. נגדיר את המשתנה המקרי X להחזיר את המיקום הראשון בו מופיע עץ (אם באף הטלה לא יצא עץ המשתנה המקרי יחזיר 0). מה ההתפלגות של X בהינתן המאורע "יצא רק עץ אחד"?

פתרון: מרחב המדגם הרלוונטי הוא $\Omega=[\{0,1\}]^3$ (עץ מיוצג על ידי 1) עם פונקציית ההסתברות האחידה על Ω. המאורע "יצא עץ" הוא $A=\{(1,0,0),(0,1,0),(0,0,1)\}$. נבצע חישוב ישיר:

$$\mathbb{P}_{X|A}(\{1\})=\mathbb{P}(X=1|A)=\frac{\mathbb{P}(\{1,0,0\})}{\mathbb{P}(A)}=\frac{1}{3}$$

$$\mathbb{P}_{X|A}(\{2\})=\mathbb{P}(X=1|A)=\frac{\mathbb{P}(\{0,1,0\})}{\mathbb{P}(A)}=\frac{1}{3}$$

$$\mathbb{P}_{X|A}(\{3\})=\mathbb{P}(X=1|A)=\frac{\mathbb{P}(\{0,0,1\})}{\mathbb{P}(A)}=\frac{1}{3}$$

כלומר $X|A$ מתפלג אחיד על הקבוצה $\{1,2,3\}$.

4 אי-תלות של משתנים מקריים

תזכורת: משתנים מקריים X,Y על אותו מרחב הסתברות ייקראו **בלתי תלויים** אם לכל שתי קבוצות $S,T\in\mathcal{F}_{\mathbb{R}}$, המאורעות $\{X\in S\},\{Y\in T\}$ בלתי תלויים, כלומר $\mathbb{P}(\{X\in S\}\cap\{Y\in T\})=\mathbb{P}(\{X\in S\})\mathbb{P}(\{Y\in T\})$. באופן שקול X,Y בלתי תלויים אם לכל $S\in\mathcal{F}_{\mathbb{R}}$ כך ש-$\mathbb{P}(\{X\in S\})>0$ מתקיים $\mathbb{P}_{(Y|\{X\in S\})}=\mathbb{P}_Y$.

אם X,Y משתנים מקריים בלתי תלויים, אז **ההתפלגות המשותפת** של $Z=(X,Y)$ נקבעת על ידי ההתפלגויות השוליות של X,Y.

תרגיל: יהיו X,Y שני משתנים מקריים בלתי תלויים המתפלגים אחיד על $\{0,1,2\}$ ונסמן $Z=(X+Y)\bmod 3$. האם X ו-Z בלתי תלויים?

פתרון: כיוון שX ו-Z בדידים, די שנבדוק כי לכל $a,b\in\{0,1,2\}$ מתקיים $\mathbb{P}(X=a,Z=b)=\mathbb{P}(X=a)\mathbb{P}(Z=b)$. נחשב לכל $a,b\in\{0,1,2\}$:

$$\mathbb{P}(X=a,Z=b)=\mathbb{P}(X=a,a+Y\overset{3\text{ mod}}{=}b)=\mathbb{P}(X=a)\mathbb{P}(Y\overset{3\text{ mod}}{=}b-a)=\frac{1}{9}$$

מנוסחת ההסתברות השלמה,

$$\mathbb{P}(Z=b)=\sum_{x\in\mathbb{R}}\mathbb{P}(X=x.Z=b)=\sum_{x\in\{0,1,2\}}\mathbb{P}(X=x.Z=b)=\frac{1}{3}$$

אז $\mathbb{P}(X=a,Z=b)=\mathbb{P}(X=a)\mathbb{P}(Z=b)$ ולכן המשתנים הם בלתי תלויים.

דוגמה לאי תלות מותנית- אם מספיקים: בוחרים מטבע אחת מתוך סל שיש בו מטבע אחד הוגן ומטבע אחד שאינה הוגן (למשל ניקח מטבע שההסתברות עבור קבלת עץ בה היא $\frac{1}{3}$) ומטילים אותו פעמיים. בהינתן שיצא מטבע הוגן, וכן בהינתן שיצא מטבע שאינו הוגן, תוצאת ההטלות היא בלתי תלויה. אחרת, היא תלויה.

תרגיל: יהיו $X\sim\text{Ber}(p)$, $Y\sim\text{Ber}(q)$ בלתי תלויים. הראו כי $XY\sim\text{Ber}(pq)$.

פתרון: נסמן ב–A את המאורע $\{X=1\}$ וב–B את המאורע $\{Y=1\}$. השתכנעו ש–XY מקבל אך ורק את הערכים $0,1$.

לכל ω, מתקיים $XY(\omega)=1$ אם ורק אם $X(\omega)=Y(\omega)=1$

ולכן:

$$\mathbb{P}(XY=1)=\mathbb{P}(X=1,Y=1)=\mathbb{P}(X=1)\mathbb{P}(Y=1)=pq$$

כאשר השוויון השני נובע מאי-תלות. נסיק כי $XY\sim\text{Ber}(pq)$ כנדרש.

(תרגיל- פיתוח נוסחת הקונבולוציה): יהיו X,Y משתנים מקריים בלתי תלויים הנתמכים על $\mathbb{Z}$ ונסמן $Z=X+Y$. אזי:

$$\mathbb{P}(Z=n)=\sum_{i\in\mathbb{Z}}\mathbb{P}(X=i)\mathbb{P}(Y=n-i).$$

נחשב בעזרת נוסחת ההסתברות השלמה:

$$\begin{aligned}\mathbb{P}(Z=n)&=\sum_{r\in\mathbb{R}}\mathbb{P}(Z=n|X=r)\mathbb{P}(X=r)\\&=\sum_{i\in\mathbb{Z}}\mathbb{P}(X+Y=n|X=i)\mathbb{P}(X=i)\\&=\sum_{i\in\mathbb{Z}}\mathbb{P}(Y=n-i|X=i)\mathbb{P}(X=i)\\&=\sum_{i\in\mathbb{Z}}\mathbb{P}(Y=n-i)\mathbb{P}(X=i),\end{aligned}$$

כאשר השוויון האחרון נובע מאי-תלות.

תרגיל: יהיו $X,Y,B:\Omega\to\mathbb{R}$ משתנים מקריים כך שקיימות $S,T\subset\mathbb{R}$ זרות המקיימות:

$$\begin{aligned}&X\sim\text{Unif}\,(S)\\&Y\sim\text{Unif}\,(T)\\&B\sim\text{Ber}\left(\frac{|S|}{|T\cup S|}\right)\end{aligned}$$

הראו כי אם זוגות המשתנים $\{B,Y\},\{B,X\}$ בלתי תלויים אז $BX+(1-B)\,Y\sim\text{Unif}\,(T\cup S)$.

הערה: המשתנה המקרי $BX+(1-B)\,Y$ הוא התוצאה של הניסוי הבא: ישנן שתי קוביות הוגנות, אחת עם מספר פאות $|S|$ והשנייה עם מספר פאות $|T|$. השלב הראשון בניסוי הוא להגריל את אחת הקוביות כך שההסתברות לקבל את הראשונה

הוא $\frac{|S|}{|S|+|T|}$, ולאחר מכן מטילים את הקוביה הנבחרת וניקח את ההטלה הזו כתוצאת הניסוי.
פתרון: יהא $x \in T \cup S$. נסמן $Z = BX + (1-B)Y$ ואז מתקיים

$$\mathbb{P}(Z=x) = \mathbb{P}(Z=x|B=1)\,\mathbb{P}(B=1) + \mathbb{P}(Z=x|B=0)\,\mathbb{P}(B=0)$$

מאי-תלות של זוגות המשתנים נקבל

$$\begin{aligned}\mathbb{P}(Z=x) &= \mathbb{P}(Z=x|B=1)\,\mathbb{P}(B=1) + \mathbb{P}(Z=x|B=0)\,\mathbb{P}(B=0) = \\ &= \mathbb{P}(X=x)\,\mathbb{P}(B=1) + \mathbb{P}(Y=x)\,\mathbb{P}(B=0)\end{aligned}$$

כעת נחלק למקרים: אם $x \in T$ אז מתקיים $\mathbb{P}(Y=x) = \frac{1}{|T|}$ אבל מהנתון ש-S,T זרות נקבל $\mathbb{P}(X=x) = 0$ ולכן

$$\mathbb{P}(Z=x) = \mathbb{P}(Y=x)\,\mathbb{P}(B=0) = \frac{1}{|T|}\frac{|T\cup S| - |S|}{|T\cup S|} = \frac{1}{|T\cup S|}$$

ואם $x \in S$ נקבל באופן דומה

$$\mathbb{P}(Z=x) = \mathbb{P}(X=x)\,\mathbb{P}(B=1) = \frac{1}{|S|}\frac{|S|}{|T\cup S|} = \frac{1}{|T\cup S|}$$

כנדרש.
תרגיל: יהא X משתנה מקרי בדיד. הוכיחו כי X בלתי תלוי בעצמו אם ורק אם X קבוע כמעט תמיד.
פתרון: X בלתי תלוי בעצמו אם ורק אם לכל A מתקיים $P(X\in A) = P(X\in A)P(X\in A) = P(X\in A)^2$ ולכן אם ורק אם $P(X\in A)\in\{0,1\}$. אם X קבוע זה מובן מאליו. בכיוון השני -
נניח ש-X כמעט תמיד ב-A, ותהי B כלשהי אז נסיק כי או ש-$X \in (A\cap B)$ כמעט תמיד, או ש-$X\in(A\cap B^c)$ כמעט תמיד.
נעשה את התהליך באופן מקונן על קטעים דיאדיים ונסיק ש-X נמצא בחיתוך המקונן של סדרת קטעים דיאדיים ולכן יש ערך שהוא מקבל כמעט תמיד.

מבוא להסתברות וסטטיסטיקה 80430
תרגול 6

1 התפלגות בינומית

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. נאמר שמשתנה מקרי X המוגדר על Ω **מתפלג בינומי** עם פרמטרים (n,p), אם X מקבל ערכים בקבוצה $\{0,...,n\}$ ופונקציית ההסתברות הנקודתית שלו היא, עבור $0 \le k \le n$

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

במקרה זה נסמן $X \sim \mathrm{Bin}(n,p)$.

ראינו בכיתה כי אם $X_1, \ldots, X_n$ משתנים מקריים בלתי תלויים ובעלי התפלגות $\mathrm{Ber}(p)$, אזי $X_1 + \ldots + X_n \sim \mathrm{Bin}(n,p)$.

תרגיל 1: שיכור עומד על ציר המספרים השלמים מעל 0. בכל דקה הוא הולך באקראי צעד אחד ימינה או צעד אחד שמאלה, כאשר הצעדים בלתי תלויים. נחשב את ההסתברות שהשיכור חוזר ל-0 לאחר n צעדים.

פתרון: יהי A_n המאורע שלאחר n צעדים השיכור חזר ל-0. ברור כי $\mathbb{P}(A_n) = 0$ אם n אי זוגי. עבור $k \in \mathbb{N}$ יהי X_k המשתנה המקרי שמתאר האם השיכור בחר לצעוד ימינה או שמאלה. קרי, $\mathbb{P}(X_k = -1) = \mathbb{P}(X_k = 1) = 1/2$. נבחין כי

$$A_n = \left\{ \sum_{k=1}^{n} X_k = 0 \right\}.$$

כעת, אמנם X_k אינו בעל התפלגות ברנולי, אבל המשתנה המקרי $Y_k = (X_k + 1)/2$ הוא בעל התפלגות ברנולי (למה? תרגיל קטן). לכן נציג

$$A_n = \left\{ \sum_{k=1}^{n} X_k = 0 \right\} = \left\{ \sum_{k=1}^{n} (2Y_k - 1) = 0 \right\} = \left\{ 2\sum_{k=1}^{n} Y_k - n = 0 \right\} = \left\{ \sum_{k=1}^{n} Y_k = n/2 \right\}$$

כעת מהעובדה ש-$\sum_{k=1}^{n} Y_k \sim \mathrm{Bin}(n, 1/2)$ ובשימוש בקירוב סטירלינג נסיק כי

$$\begin{aligned}\mathbb{P}(A_n) &= \binom{n}{n/2} \cdot 1/2^{n/2} \cdot (1-1/2)^{n-n/2} = \frac{n!}{(n/2)!^2} \cdot 1/2^n \\ &\approx \frac{\sqrt{2\pi n}\,(n/\mathrm{e})^n}{\pi n\,(n/2\mathrm{e})^n} \cdot 1/2^n = \sqrt{2/\pi} \cdot 1/\sqrt{n} = O\left(1/\sqrt{n}\right).\end{aligned}$$

נזכיר כי קירוב סטירלינג הוא

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

תרגיל 2: רובין בובין פותח n עוגיות מזל שבכל אחת יש נבואה משמחת בהסתברות p. נסמן ב–k את מספר העוגיות בהן הייתה נבואה משמחת. לאחר שסיים, רובין פותח עוד k עוגיות מזל מחברה אחרת שבכל אחת יש נבואה משמחת בהסתברות q. מהי התפלגות מספר הנבואות המשמחות בערימת העוגיות השניה?

פתרון: נסמן ב–X את המשתנה המקרי שסופר את כמות הנבואות המשמחות בערימת העוגיות הראשונה. כפי שכבר ראינו – $\mathrm{X} \sim \mathrm{Bin}\left(n,p\right)$. נגדיר משתנה מקרי נוסף Y שסופר את מספר הנבואות המשמחות בערימת העוגיות השניה. נשים לב שלכל k, $\mathrm{Y}|\{\mathrm{X}=k\} \sim \mathrm{Bin}\left(k,q\right)$, ו–$Supp(\mathrm{Y}) \subseteq [n]$. נחשב את התפלגות Y:

$$\mathbb{P}_{\mathrm{Y}}\left(i\right) = \sum_{k=1}^{n} \mathbb{P}\left(\mathrm{Y}=i \mid \mathrm{X}=k\right)\mathbb{P}\left(\mathrm{X}=k\right)$$

$$= \sum_{k=i}^{n} \binom{k}{i} q^{i}\left(1-q\right)^{k-i}\binom{n}{k} p^{k}\left(1-p\right)^{n-k}$$

$$(*) = \binom{n}{i}\sum_{k=i}^{n}\binom{n-i}{k-i} q^{i}\left(1-q\right)^{k-i} p^{k}\left(1-p\right)^{n-k}$$

$$= \binom{n}{i}\left(pq\right)^{i}\sum_{k=0}^{n-i}\binom{n-i}{k}\left(p\left(1-q\right)\right)^{k}\left(1-p\right)^{n-i-k}$$

$$= \binom{n}{i}\left(pq\right)^{i}\left(p\left(1-q\right)+\left(1-p\right)\right)^{n-i}$$

$$= \binom{n}{i}\left(pq\right)^{i}\left(1-pq\right)^{n-i}$$

כאשר השוויון ב–$(*)$ נובע מכך ש–

$$\binom{k}{i}\binom{n}{k} = \frac{k!n!}{i!\left(k-i\right)!k!\left(n-k\right)!} = \frac{n!}{i!\left(k-i\right)!\left(n-k\right)!}$$

$$= \frac{n!}{i!\left(n-i\right)!}\binom{n-i}{k-i} = \binom{n}{i}\binom{n-i}{k-i}.$$

סה"כ קיבלנו ש–$\mathrm{Y} \sim \mathrm{Bin}\left(n,pq\right)$.

2 קשר בין ברנולי לבינומי

תרגיל 1: יהיו $\mathrm{X}_1,\mathrm{X}_2,\mathrm{X}_3$ משתנים מקריים בלתי תלויים מתפלגים ברנולי $0<p<1$. נגדיר לכל $1 \le i \le 2$ משתנה מקרי Y_i להיות האינדיקטור של הקבוצה $\mathrm{X}_i=\mathrm{X}_{i+1}$, עבור איזה p מתקיים Y_1,Y_2 בלתי תלויים?

פתרון: משאלה 4 בתרגיל 5 נובע שמספיק להראות שמתקיים

$$\mathbb{P}\left(\mathrm{Y}_1=1\right)\mathbb{P}\left(\mathrm{Y}_2=1\right) = \mathbb{P}\left(\mathrm{Y}_1=1,\mathrm{Y}_2=1\right) \tag{1}$$

נשים לב שמתקיים

$$\mathbb{P}\left(\mathrm{Y}_i=1\right)=\mathbb{P}\left(\mathrm{X}_i=\mathrm{X}_{i+1}=1\right)+\mathbb{P}\left(\mathrm{X}_i=\mathrm{X}_{i+1}=0\right)=p^2+(1-p)^2$$

לכן $\mathrm{Y}_1\sim\mathrm{Y}_2\sim\mathrm{Ber}\left(p^2+(1-p)^2\right)$. כעת נחשב את הביטוי הימני במשוואה אחת

$$\mathbb{P}\left(\mathrm{Y}_1=1,\mathrm{Y}_2=1\right)=\mathbb{P}\left(\mathrm{X}_1=\mathrm{X}_2=\mathrm{X}_3=1\right)+\mathbb{P}\left(\mathrm{X}_1=\mathrm{X}_2=\mathrm{X}_3=0\right)=p^3+(1-p)^3$$

כלומר משוואה אחת מתקיימת אם ורק אם

$$\left(p^2+(1-p)^2\right)^2=p^3+(1-p)^3$$

פתרונות למשוואה הן $p\in\left\{0,\frac{1}{2},1\right\}$, לכן $p=\frac{1}{2}$ הוא היחיד המקיים ש- $\mathrm{Y}_1,\mathrm{Y}_2$ בלתי תלויים.

תרגיל 2: נראה כי אם $\mathrm{X}\sim\mathrm{Bin}\left(n,p\right)$ אז $(n-\mathrm{X})\sim\mathrm{Bin}\left(n,1-p\right)$.

פתרון: ניקח $\mathrm{Y}=\{\mathrm{X}_i'\}_{i\in[n]}$ וקטור של משתני ברנולי עם הסתברות הצלחה p, בלתי-תלויים. מהטענה לעיל נקבל ש-

$$\mathrm{Y}=\sum_{i\in[n]}\mathrm{X}_i'\sim\mathrm{X}$$

לכן

$$(n-\mathrm{X})\sim n-\sum_{i\in[n]}\mathrm{X}_i'=\sum_{i\in[n]}\left(1-\mathrm{X}_i'\right)\sim\mathrm{Bin}\left(n,1-p\right)$$

שוויון ההתפלגות האחרון נובע מהטענה המוזכרת, מכך ש- $\left(1-\mathrm{X}_i'\right)\sim\mathrm{Ber}\left(1-p\right)$ ו- $\{1-\mathrm{X}_i'\}_{i\in[n]}$ בלתי תלויים (אי תלות נשמרת תחת הפעלת פונקציות).

3 התפלגות גאומטרית

נמשיך להכיר התפלגויות מיוחדות נוספות. כבר הכרנו משתנה מקרי ברנולי, אחיד ובינומי. נגדיר שתי התפלגויות נוספות וניתן דוגמאות.

הגדרה: יהא $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות, $\mathrm{X}:\Omega\to\mathbb{R}$ משתנה מקרי עליו. נאמר ש-X **מתפלג גיאומטרית** עם סיכוי הצלחה p עבור $p\in(0,1)$ כלשהו, אם $Supp(\mathrm{X})=\mathbb{N}$ והתפלגותו:

$$\mathbb{P}(\mathrm{X}=n)=(1-p)^{n-1}p.$$

במקרה זה נרשום $\mathrm{X}\sim\mathrm{Geo}\left(p\right)$.

הערה: משתנה גיאומטרי מתאר ניסוי חוזר ונשנה עם הסתברות p להצליח בכל ניסיון (הניסיונות בלתי תלויים), המשתנה מונה את מספר הניסויים עד להצלחה הראשונה (כולל).

הערה: אם $X\sim\mathrm{Geo}\left(p\right)$ אז לכל $k\in\mathbb{N}$ מתקיים

$$\mathbb{P}(\mathrm{X}>k)=(1-p)^k.$$

טענה: אם $\mathrm{X}\sim\mathrm{Geo}\left(p\right)$ אז $\mathbb{P}(\mathrm{X}\in\mathbb{N})=1$.

הוכחה: סדרת ההסתברויות $\{\mathbb{P}(\mathrm{X}=n)\}_{n\in\mathbb{N}}$ היא סדרה הנדסית עם איבר ראשון p ומנה $1-p$, לכן:

$$\mathbb{P}(\mathrm{X}\in\mathbb{N})=\sum_{n=1}^{\infty}\mathbb{P}(\mathrm{X}=n)$$

$$= \sum_{n=1}^{\infty} (1-p)^{n-1} p = \frac{p}{1-(1-p)} = \frac{p}{p} = 1.$$

דוגמא: מבדיקה שנערכה התברר ש–40% מבין תלמידי תואר שני למתמטיקה באוניברסיטה העברית הם מתרגלים פעילים. רועי רוצה לפגוש מתרגל פעיל ומתחיל לפנות לסטודנטים בתואר שני באקראי. נסמן ב–X את מספר הסטודנטים אליהם פנה רועי עד אשר הגיע למתרגל פעיל. ההסתברות שיפנה לכל היותר ל–3 סטודנטים היא:

$$\mathbb{P}(\mathrm{X} \le 3) = 1 - \mathbb{P}(\mathrm{X} > 3) = 1 - 0.6^3$$

כאשר השוויון השני נובע מההערה.
טענה: יהי $\mathrm{X} \sim \mathrm{Geo}\,(p)$ משתנה מקרי גיאומטרי, אזי:

$$\mathbb{P}(\mathrm{X} = n + k | \mathrm{X} > k) = \mathbb{P}(\mathrm{X} = n)$$

תכונה זו נקראת **תכונת חוסר הזיכרון**. למעשה משתנה מקרי הוא גיאומטרי אם ורק אם הוא חסר זיכרון.
תרגיל 1: איתמר ואלעד משחקים זוג או פרט שוב ושוב. הם מפסיקים לשחק כאשר אחד מהם ניצח פעמיים יותר מהשני. נסמן ב–X את מספר הפעמים בסדרת המשחקים שבהן הם בתיקו, כיצד מתפלג X?
פתרון: התפלגות X היא גיאומטרית. בכל פעם שהם מגיעים לתיקו, המצב זהה, ואז אנחנו שואלים מה יקרה קודם – שאחד מהם ינצח פעמיים יותר מהשני, או שהם יחזרו שוב לתיקו. כדי לחזור למצב של תיקו, לא משנה מי מנצח במשחק הראשון, המנצח במשחק השני צריך להיות השחקן האחר. היות שההסתברות של כל שחקן לנצח משחק היא $\frac{1}{2}$, ההסתברות לחזור לתיקו היא $\frac{1}{2}$ (וההסתברות שהמשחק יסתיים גם היא חצי, בהתאם למה שנאמר קודם), לכן $\mathrm{X} \sim \mathrm{Geo}\left(\frac{1}{2}\right)$.
תרגיל 2: עליסה נמצאת בחדר מבואה עם דלת כניסה ושלוש דלתות יציאה. דלת אחת מוליכה לארץ הפלאות, דלת שניה מוליכה לשפרינצק והדלת השלישית מוליכה בחזרה לחדר המבואה. בכל ביקור בחדר המבואה עליסה בוחרת בדלת המוליכה לשפרינצק בסיכוי p_1 ובדלת שמובילה לארץ הפלאות בהסתברות p_2 עבור $p_1 + p_2 < 1$ נתונים.

1. מה הסיכוי שעליסה תגיע בסופו של דבר לארץ הפלאות?
2. כיצד מתפלג מספר הביקורים של עליסה בחדר המבואה בהינתן שבסופו של דבר היא הגיעה לארץ הפלאות?

פתרון:

1. כדי שעליסה תגיע לארץ הפלאות, אסור שהיא תיכנס אי פעם בדלת של שפרינצק, כלומר כל סדרת דלתות שתביא אותה לארץ הפלאות היא סדרה שבה יש כניסות רצופות בדלת השלישית, ולאחריהן כניסה בדלת המוליכה לארץ הפלאות. נסמן ב–ω_n את המאורע שבו עליסה מגיעה לארץ הפלאות בדיוק בפעם ה–n, אזי $p(\omega_n) = (1-(p_1+p_2))^{n-1} p_2$. נסמן ב A את המאורע שבו עליסה מגיעה בסופו של דבר לארץ הפלאות. אזי $A = \{\omega_n | \, n \in \mathbb{N}\}$, וההסתברות של המאורע הזה היא:

$$\mathbb{P}(A) = \sum_{n=1}^{\infty} (1-(p_1+p_2))^{n-1} p_2 = \frac{p_2}{p_1+p_2}.$$

2. נסמן ב X את מספר הביקורים של עליסה בחדר המבואה, ונחשב את ההתפלגות של $(\mathrm{X}|A)$:

$$\mathbb{P}(\mathrm{X} = n|A) = \frac{\mathbb{P}(\{\mathrm{X} = n\} \cap A)}{\mathbb{P}(A)} = \frac{p(\omega_n)}{\frac{p_2}{p_1+p_2}} = (1-(p_1+p_2))^{n-1}(p_1+p_2)$$

ולכן $(\mathrm{X}|A) \sim \mathrm{Geo}\,(p_1+p_2)$.

תרגיל 3: מטילים מטבע מוטה p, n פעמים. יהי X משתנה מקרי המתאר את כמות העצים שיצאה ב-n ההטלות, ו-Y הוא המשתנה המקרי המתאר את מספר ההטלה הראשונה בה יצא עץ.

1. חשבו את $\mathbb{P}\left(\mathrm{X}=0\right)$.
2. חשבו את $\mathbb{P}\left(\mathrm{Y}>n\right)$.

פתרון:

1. נשים לב כי $\mathrm{X}\sim\mathrm{Bin}\left(n,p\right)$ (מדובר בסכום הצלחות של n ניסויי ברנולי ב"ת), לכן:

$$\mathbb{P}\left(\mathrm{X}=0\right)=\binom{n}{0}\cdot p^{0}\cdot\left(1-p\right)^{n-0}=\left(1-p\right)^{n}.$$

2. שכנעו את עצמכם כי $\mathrm{Y}\sim\mathrm{Geo}\left(p\right)$. לכן:

$$\mathbb{P}\left(\mathrm{Y}>n\right)=\left(1-p\right)^{n}.$$

קיבלנו הסתברויות שוות! אל לנו להיות מופתעים, מאחר שמדובר באותו המאורע: $\left\{\mathrm{X}=0\right\}$ הוא המאורע "כמות העצים ב-n ההטלות היא 0", בעוד ש-$\left\{\mathrm{Y}>n\right\}$ הוא המאורע "לפחות n פעמים לא יצא עץ בניסוי", ובהקשר של השאלה שלנו אלו אכן תיאורים שקולים של אותו המאורע.

תרגיל 4: יהיו $X_1,X_2\sim Geo(p)$ בלתי תלויים. הוכיחו כי $\mathrm{X}_1|\mathrm{X}_1+\mathrm{X}_2=k+1\sim\mathrm{Unif}\left(\left[k\right]\right)$.
פתרון: נשים לב שלכל $l\in\left[k\right]$ מתקיים:

$$\mathbb{P}\left(\mathrm{X}_1=l|\mathrm{X}_1+\mathrm{X}_2=k+1\right)=\frac{\mathbb{P}\left(\mathrm{X}_1=l\wedge\mathrm{X}_1+\mathrm{X}_2=k+1\right)}{\mathbb{P}\left(\mathrm{X}_1+\mathrm{X}_2=k+1\right)}=\frac{\mathbb{P}\left(\mathrm{X}_1=l\wedge\mathrm{X}_2=k+1-l\right)}{\mathbb{P}\left(\mathrm{X}_1+\mathrm{X}_2=k+1\right)}.$$

בתרגיל הבית תראו כי עבור $X_1,X_2\sim Geo(p)$ בלתי תלויים, מתקיים:

$$\mathbb{P}(X_1+X_2=2+k)=(k+1)(1-p)^kp^2.$$

נשתמש בעובדה זו, ובכך ש$\mathrm{X}_1,\mathrm{X}_2$ בלתי-תלויים:

$$=\frac{\mathbb{P}\left(\mathrm{X}_1=l\right)\mathbb{P}\left(\mathrm{X}_2=k+1-l\right)}{k\left(1-p\right)^{k-1}p^2}=\frac{\left(1-p\right)^{l-1}p\cdot\left(1-p\right)^{k-l}p}{k\left(1-p\right)^{k-1}p^2}=\frac{\left(1-p\right)^{k-1}}{\left(1-p\right)^{k-1}k}=\frac{1}{k}.$$

ולכן $\mathrm{X}_1|\mathrm{X}_1+\mathrm{X}_2=k+1\sim\mathrm{Unif}\left(\left[k\right]\right)$.

תרגול 7

1 התפלגות פואסון

הגדרה: נאמר ש־X **מתפלג פואסון** (או **פואסונית**) עם שכיחות $\lambda \geq 0$ אם לכל $n \in \mathbb{N}_0$ מתקיים:

$$p_{\mathrm{X}}(n) = \frac{e^{-\lambda}\lambda^n}{n!}.$$

במקרה זה, נרשום $X \sim \mathrm{Pois}(\lambda)$.

התפלגות פואסונית מתאימה למשתנה מקרי שסופר את מספר ההופעות של אירוע מסויים בפרק זמן נתון כאשר ממוצע ההופעות באותו פרק זמן הוא λ.
הערה: ניתן לחשוב על התפלגות פואסונית כעל התפלגות בינומית עם מס׳ הניסויים שואף לאינסוף. (הוכחתם בכיתה)
דוגמאות לדברים המתפלגים פואסונית: מס׳ האטומים המתפרקים בחומר רדיואקטיבי בפרק זמן נתון, מס׳ הפוטונים המגיעים לגלאי בפרק זמן נתון.
טענה: אם $X \sim \mathrm{Pois}(\lambda)$ ו־$Y \sim \mathrm{Pois}(\eta)$ בלתי תלויים אז $X + Y \sim \mathrm{Pois}(\lambda + \eta)$.
הוכחה:
נעזר בנוסחת הקונבולוציה מתרגול קודם

$$\begin{aligned}\mathbb{P}(X+Y=n) &= \sum_{i\in\mathbb{Z}} \mathbb{P}(X=i)\,\mathbb{P}(Y=n-i) = \sum_{i=0}^{n} \frac{e^{-\lambda}\lambda^i}{i!}\frac{e^{-\eta}\eta^{n-i}}{(n-i)!} \\ &= \frac{e^{-\lambda-\eta}}{n!}\sum_{i=0}^{n}\binom{n}{i}\lambda^i\eta^{n-i} \underbrace{=}_{Binomial} \frac{(\lambda+\eta)^n e^{-(\lambda+\eta)}}{n!}\end{aligned}$$

דוגמא: לקפיטריה של שפרינצק יש בממוצע 10 לקוחות כל שעה, מה ההסתברות שבשעה מסויימת יגיעו לקפיטריה 20 לקוחות אם ידוע שמספר הלקוחות מתפלג פואסונית?
יהא X משתנה מקרי המתאר את מספר הלקוחות שהגיעו לקפיטריה בשעת המדידה. כאמור, X מתפלג פואסונית עם פרמטר 10 ולכן:

$$p_{\mathrm{X}}(20) = \frac{e^{-10}10^{20}}{20!} \approx 0.002.$$

תרגיל 1:
ידוע שבכל דקה מדקות היום, מספר הפניות למוקד 144 מתפלג $\mathrm{Pois}(5) \sim$ באופן בלתי תלוי ביתר הדקות. מה ההסתברות שבין $10:00$ ל־$10:01$ לא התקבלה אף פניה? מה ההסתברות שבין $10:00$ ל־$10:05$ התקבלו 4 פניות בדיוק?

פתרון:
נסמן ב–X_1 את המשתנה המקרי שסופר את כמות השיחות בדקה הראשונה, נרצה לחשב את ההסתברות של המאורע $X_1 = 0$. נחשב:

$$\mathbb{P}(\mathrm{X}_1 = 0) = \frac{e^{-5}5^0}{0!} = e^{-5}.$$

עבור החישוב השני, נגדיר בדומה משתנים מקריים נוספים X_2, X_3, X_4, X_5, לפי הנתון כל החמישה בלתי תלויים. מהתזכורת, סכומם של משתני פואסון בלתי תלויים אף הוא פואסוני ושכיחותו היא סכום השכיחויות, ולכן מתפלג $\sim \mathrm{Pois}(25)$. בפרט:

$$\mathbb{P}(\sum_{i=1}^{5} X_i = 4) = \frac{e^{-25}25^4}{4!}.$$

תרגיל 2: מספר הביצים שלטאה מטילה ביום מתפלג פואסונית עם פרמטר $\lambda > 0$. כל ביצה בוקעת בהסתברות $0 < p < 1$ באופן בלתי-תלוי באחרות. נגדיר את המשתנה המקרי X להיות מספר הביצים שהוטלו בשעה מסוימת ויבקעו (מתישהו, הכוונה שאנחנו לא סופרים ביצים מתות). כיצד מתפלג X?

פתרון: נחשב את $\mathbb{P}\left(X = k\right)$ לכל $k \in \mathbb{N} \cup \{0\}$. נחלק את המאורע $\{X = k\}$ בהתאם למספר ההטלות n ע"י משתנה מקרי Y שמחזיר את מספר ההטלות (כולל אלו שלא בקעו):

$$\begin{aligned}\mathbb{P}\left(X = k\right) &= \sum_{n=k}^{\infty} \mathbb{P}\left(\{X = k\} \cap \{Y = n\}\right) = \\ &= \sum_{n=k}^{\infty} \mathbb{P}\left(X = k \middle| \{Y = n\}\right) \cdot \mathbb{P}\left(\{Y = n\}\right) = \\ &= \sum_{n=k}^{\infty} \mathbb{P}\left(X = k \middle| \{Y = n\}\right) \cdot e^{-\lambda}\frac{\lambda^n}{n!} = \ldots\end{aligned}$$

נבחין שבהינתן כך שהיו n הטלות, ההתפלגות של מספר הביצים שבקעו הינו בינומי$\sim \mathrm{Bin}\left(n, p\right)$. לכן נוכל להמשיך בקלות את החישוב:

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

כלומר המשתנה המקרי מתפלג כמו משתנה פואסוני בעל פרמטר λp: $X \sim \mathrm{Pois}(\lambda p)$.

הסבר: עדיין מתקיימות ההנחות של משתנה מקרי פואסוני, וממוצע הביצים החיות שהוטלו הוא p (אחוז הביצים החיות) כפול ממוצע הביצים שהוטלו λ.

2 תוחלת של משתנה מקרי בדיד

הגדרה: יהי X משתנה מקרי בדיד על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. ה**תוחלת** של X היא

$$\mathbb{E}(X) = \sum_{x \in Supp(X)} x \mathbb{P}_X(x).$$

כאשר הטור מתכנס בהחלט. אחרת נאמר כי ל־X אין תוחלת (סופית).
נזכיר בקצרה כמה מהתכונות היסודיות של התוחלת וכמה טענות שראיתם בכיתה:

- תוחלת היא תכונה של פונקציית ההתפלגות של המשתנה המקרי. כלומר, למשתנים מקריים שווי-התפלגות יש את אותה התוחלת, גם אם הם מוגדרים מעל מרחבי הסתברות שונים.
- **ליניאריות:** $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.
- **חיוביות:** אם $X \overset{a.s.}{\geq} 0$ אז $\mathbb{E}(X) \geq 0$. אם בנוסף $\mathbb{P}(X > 0) > 0$ אז $\mathbb{E}(X) > 0$.
- **מונוטוניות:** אם $X \overset{a.s}{\geq} Y$ אז $\mathbb{E}(X) \geq \mathbb{E}(Y)$.
- אם X, Y מ״מ ב״ת ובעלי תוחלת, אזי $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.
- יהי X מ״מ ותהי $f : \mathbb{R} \to \mathbb{R}$ פונקציה כך שלמשתנה המקרי $f(X)$ יש תוחלת. אזי

$$\mathbb{E}(f(X)) = \sum_{x \in Supp(X)} f(x) \mathbb{P}_X(x).$$

דוגמא: סטודנטים בספריה הולכים לנמנם בחדר הפופים, כאשר אורך תנומה הוא אקראי בטווח $\{30, ..., 60\}$ דקות. מהי תוחלת אורכה של תנומה מקרית?
נתון כי האורך של תנומה הוא $X \sim Unif(\{30, ..., 60\})$ ולכן:

$$\mathbb{E}(X) = \sum_{n=30}^{60} n\mathbb{P}(X = n) = \sum_{n=30}^{60} n \cdot \frac{1}{31} = \frac{1395}{31} = 45.$$

טענה: (נוסחת הזנב) יהא X מ״מ הנתמך על הטבעיים. אזי:

$$\mathbb{E}(X) = \sum_{n \in \mathbb{N}} \mathbb{P}(X \geq n).$$

הוכחה: ננצל את העובדה שהמחוברים בטור אי-שליליים כדי להחליף סדר סכימה באמצעות משפט פוביני.

$$\begin{aligned}\mathbb{E}(X) &= \sum_{n=0}^{\infty} n\mathbb{P}(X = n) = \sum_{n=1}^{\infty}\sum_{k=1}^{n} \mathbb{P}(X = n) = \sum_{k,n \in \mathbb{N}, k \leq n} \mathbb{P}(X = n) \\ &= \sum_{k=1}^{\infty}\sum_{n=k}^{\infty} \mathbb{P}(X = n) = \sum_{n \in \mathbb{N}} \mathbb{P}(X \geq n).\end{aligned}$$

דוגמא: n חברים משחקים בהטלת קוביה הוגנת. מהי תוחלת ערך ההטלה הנמוכה ביותר?

יהי X משתנה מקרי שתוצאתו היא ערך ההטלה הנמוכה ביותר. נבחין כי:

$$\mathbb{P}\left(X \geq k\right) = \left(\frac{7-k}{6}\right)^{n}$$

ולכן, מנוסחת הזנב:

$$\mathbb{E}\left(X\right) = \sum_{k=1}^{6} \mathbb{P}\left(X \geq k\right) = \sum_{k=1}^{6} \left(\frac{7-k}{6}\right)^{n}.$$

3 תוחלת משתנים מקריים מוכרים

3.1 משתנה מקרי ברנולי

תוחלת משתנה מקרי **ברנולי** $X \sim \textbf{Ber}\left(p\right)$ היא

$$\mathbb{E}\left(X\right) = p \cdot 1 + \left(1-p\right) \cdot 0 = \boldsymbol{p}.$$

הערה: מכך נסיק שהתוחלת של פונקציית המציין של המאורע A היא $\mathbb{P}\left(A\right)$.

3.2 משתנה מקרי אחיד

תוחלת משתנה מקרי **אחיד** $X \sim \textbf{Unif}\left(\left[n\right]\right)$ היא

$$\mathbb{E}\left(X\right) = \sum_{k\in[n]} k\mathbb{P}\left(X=k\right) = \sum_{k\in[n]} \frac{k}{n} = \frac{n\left(n+1\right)}{2n} = \frac{\boldsymbol{n+1}}{\boldsymbol{2}}.$$

3.3 משתנה מקרי בינומי

(תראו בהרצאה) תוחלת משתנה מקרי **בינומי** $X \sim \text{Bin}\left(n,p\right)$ היא

$$\begin{aligned}
\mathbb{E}\left(X\right) &= \sum_{k=0}^{n} k\binom{n}{k} p^{k}\left(1-p\right)^{n-k} = \sum_{k=0}^{n} k\frac{n!}{(n-k)!k!} p^{k}\left(1-p\right)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(n-k)!\left(k-1\right)!} p^{k}\left(1-p\right)^{n-k} \underbrace{=}_{m=k-1} \sum_{m=0}^{n-1} \frac{n!}{(n-m-1)!\left(m\right)!} p^{m+1}\left(1-p\right)^{n-m-1} \\
&= np\sum_{m=0}^{n-1} \binom{n-1}{m} p^{m}\left(1-p\right)^{n-m-1} \underbrace{=}_{Binomial} np\left(p+\left(1-p\right)\right)^{n-1} = \boldsymbol{np}.
\end{aligned}$$

3.4 משתנה מקרי פואסוני

תוחלת משתנה מקרי **פואסוני** $X \sim \text{Pois}\left(\lambda\right)$ היא

$$\mathbb{E}\left(X\right)=\sum_{k=0}^{\infty}ke^{-\lambda}\frac{\lambda^{k}}{k!}=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\underbrace{=}_{m=k-1}\lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^{m}}{(m)!}=\lambda e^{-\lambda}e^{\lambda}=\boldsymbol{\lambda}.$$

3.5 משתנה מקרי גיאומטרי

תוחלת משתנה מקרי **גיאומטרי** $X\sim\textbf{Geo}\left(p\right)$ היא

$$\mathbb{E}\left(X\right)=\sum_{n\in\mathbb{N}}\mathbb{P}(X\geq n)=\sum_{k=0}^{\infty}\left(1-p\right)^{k}=\frac{1}{1-(1-p)}=\frac{\mathbf{1}}{\boldsymbol{p}}\mathbf{.}$$

כאשר בשוויון הראשון נעזרנו בנוסחת הזנב והשוויון השני נעזרנו בהערה הבאה
הערה: אם $X\sim\text{Geo}(p)$ אז לכל $k\in\mathbb{N}$ מתקיים

$$\mathbb{P}(X>k)=(1-p)^{k}.$$

דוגמא: אם נבצע ניסוי שבכל ניסיון יש לנו הסתברות שליש להצלחה, אז קיבלנו מחישוב התוחלת שידרשו בממוצע שלושה ניסיונות עד שאמנם נצליח. וזה מה שהיינו מצפים.

4 חישוב תוחלת פונקציה של משתנה מקרי

דוגמה: נחשב את תוחלת השטח של ריבוע בעל צלע המתפלגת אחיד על השלמים מ-0 עד n.
יהי $X\sim\text{Unif}(\{0,1,...,n\})$ המשתנה המקרי שמתאים לאורך צלע. אנו רוצים לחשב את $\mathbb{E}(X^{2})$. נשתמש בנוסחה ונקבל:

$$\mathbb{E}(X^{2})=\sum_{k\in[n]}k^{2}\mathbb{P}\left(X=k\right)=\sum_{k\in[n]}\frac{k^{2}}{n}=\frac{n(n+1)(2n+1)}{6}\frac{1}{n}=\frac{(n+1)(2n+1)}{6}$$

שימו לב, התוצאה לעיל שונה מ$\mathbb{E}(X)^{2}$.

5 גרף ארדש–רני $Erdos-Renyi$

הגדרה: נגדיר את $G_{n,p}$ להיות גרף ארדש-רני על $[n]$ קודקודים. כלומר גרף מקרי שבו לכל שני קודקודים שונים $u,v\in[n]$ הקשת (u,v) נמצאת בגרף בהסתברות p, ואין תלות בין הימצאותן של קשתות שונות בגרף. נסמן את קבוצת הקשתות שהוגרלו להמצא בגרף ב-E.

הגדרה: גרף עם n קודקודים שבו קיימת קשת בין כל שני קודקודים שונים נקרא <u>גרף שלם</u>. תת גרף (כלומר קבוצת קודקודים וקבוצת קשתות המוכלות בקבוצות בגרף המקורי) המהווה גרף שלם נקרא <u>קליקה</u>. נסמן ב-K_n קליקה עם n קודקודים. מס' הקשתות בK_n הינו $|\mathbb{E}\left(K_{n}\right)|=\frac{n(n-1)}{2}$.

תרגיל: יהי $G_{n,p}$ גרף ארדש-רני. חשבו את התוחלת של מס' הקליקות עם 4 קודקודים.

פתרון: נתחיל בלמצוא את מספר K_4 המוכלים בגרף. נסמן את מס' זה ב-:

$$k=\left|\left\{ \left\{ i,j,k,l\right\} \subset[n]:\left(i,j\right),\left(i,k\right),\left(i,l\right),\left(j,k\right),\left(j,l\right),\left(k,l\right)\in E\left(G_{n,p}\right)\right\} \right|.$$

נתאר את הגרף באמצעות $\binom{n}{2}$ משתנים מציינים. לכל שני משתנים שונים $u,v\in[n]$ נסמן $X_{uv}=\mathbb{1}_{(u,v)\in E(G_{n,p})}$ המציין האם הקשת הוגרלה או לא. נשים לב שזהו משתנה ברנולי-p. ונסמן ב $Y_{i,j,k,l}=\mathbb{1}_{\{i,j,k,l \text{ a form } K_4\}}$ את המשתנה המציין את הקליקה הנוצרת ע"י $1\leq i<j<k<l\leq n$. בסימונים אלו נקבל כי:

$$k=\sum_{1\leq i<j<k<l\leq n}Y_{i,j,k,l}=\sum_{1\leq i<j<k<l\leq n}X_{i,j}X_{i,k}X_{i,l}X_{j,k}X_{j,l}X_{k,l}.$$

נחשב את התוחלת:

$$\begin{aligned}\mathbb{E}\left(k\right)&=\sum_{1\leq i<j<k<l\leq n}\mathbb{E}\left(X_{i,j}X_{i,k}X_{i,l}X_{j,k}X_{j,l}X_{k,l}\right)\\&=\sum_{1\leq i<j<k<l\leq n}\mathbb{E}\left(X_{i,j}\right)\mathbb{E}\left(X_{i,k}\right)\mathbb{E}\left(X_{i,l}\right)\mathbb{E}\left(X_{j,k}\right)\mathbb{E}\left(X_{j,l}\right)\mathbb{E}\left(X_{k,l}\right)\\&=\binom{n}{4}p^6.\end{aligned}$$

כאשר השתמשנו בליניאריות התוחלת, כפליות התוחלת עבור משתנים בלתי-תלויים ובמעבר האחרון ספרנו את כמות האפשרויות לבחור ארבעה קודקודים מתוך n ולהגריל את 6 הקשתות הנחוצות כדי ליצור קליקה.

מבוא להסתברות וסטטיסטיקה 80430
תרגול חזרה

1 תזכורת – קבוצות ופונקציות

(נספחים א.1 - א.2)

הגדרה תהיינה S קבוצה כלשהי (של אינדקסים), ו-$M = \{A_\alpha : \alpha \in S\}$ משפחה של קבוצות. נגדיר:

$$\bigcap_{\alpha\in S} A_\alpha := \{x : \forall\alpha \in S, x \in A_\alpha\}, \bigcup_{\alpha\in S} A_\alpha = \{x : \exists\alpha \in S, x \in A_\alpha\}.$$

נאמר כי הקבוצות בעלות חיתוך ריק אם $\bigcap_{\alpha\in S} A_\alpha = \emptyset$. נאמר כי הקבוצות זרות (בזוגות) אם לכל $\alpha, \beta \in S$ שונים מתקיים $A_\alpha \cap A_\beta = \emptyset$.

דוגמא הקבוצות $\{\{1,2\},\{3,4\},\{4,5\}\}$ בעלות חיתוך ריק אך אינן זרות.

הגדרה תהיינה A, B קבוצות. ההפרש בין שתי הקבוצות A ל-B מסומן ב-$A \setminus B$ והוא אוסף כל האיברים של A שאינם ב-B:

$$A \setminus B = \{a : a \in A \wedge a \notin B\}.$$

לעיתים קרובות אנו מגבילים את עצמנו בדיון לתת-קבוצות של קבוצה אוניברסלית Ω. במקרה זה (כאשר $A \subseteq \Omega$) נסמן $A^c = \Omega \setminus A$ ונאמר ש-A^c הוא המשלים של A. למשל כעת נוכל לקבל את השוויון: $A \setminus B = A \cap B^c$.

טענה תהיינה $A, B \subset \Omega$ קבוצות. אזי $(A \cap B)^c = (A^c \cup B^c)$. (כלל דה-מורגן)

הוכחה נוכיח על ידי הכלה דו-כיוונית. יהי $x \in (A \cap B)^c$. כלומר $x \notin A \cap B$. זה אומר שלא מתקיים $x \in A \wedge x \in B$. כלומר $x \notin A$ או $x \notin B$, לפיכך $x \in A^c \cup B^c$.
צד שני – יהי $x \in A^c \cup B^c$. אזי $x \in A^c$ או $x \in B^c$, ונוכל להניח בלי הגבלת הכלליות כי $x \in A^c$. לפיכך $x \notin A$ ובפרט $x \notin A \cap B$, כלומר $x \in A \cap B$.

1.1 פונקציה מציינת וקבוצת החזקה

הגדרה תהיינה Ω קבוצה לא-ריקה, ו $A \subseteq \Omega$ תת-קבוצה שלה. הפונקציה המציינת של A (ביחס ל-Ω), המסומנת $\mathbb{1}_A : \Omega \to \{0,1\}$ נתונה ע"י:

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

הגדרה קבוצת החזקה של Ω מסומנת על ידי 2^Ω ומוגדרת להיות:

$$2^\Omega = \{A : A \subseteq \Omega\}.$$

כלומר 2^Ω היא אוסף כל התת-קבוצות של Ω.

דוגמאות

- $2^\emptyset = \{\emptyset\}$, (יש לשים לב שקבוצת החזקה של הקבוצה הריקה היא קבוצה לא ריקה).
- $2^{\{x,y\}} = \{\emptyset, \{x\}, \{y\}, \{x,y\}\}$.

טענה תהי Ω קבוצה כלשהי. ההתאמה $A \mapsto \mathbb{1}_A$ לכל $A \subset \Omega$ היא התאמה חח"ע ועל מ-2^Ω ל-$\{f : \Omega \to \{0,1\}\}$ (קבוצת הפונקציות מ-Ω לקבוצה $\{0,1\}$)

הוכחה ההתאמה **חח"ע**: תהיינה $A, B \in 2^\Omega$ כך ש $A \neq B$. לפיכך קיים בלי הגבלת הכלליות $x \in A$ כך ש $x \notin B$. אז יתקיים $\mathbb{1}_A(x) = 1, \mathbb{1}_B(x) = 0$, כלומר $\mathbb{1}_A \neq \mathbb{1}_B$.
ההתאמה **על**: תהי $f : \Omega \to \{0,1\}$. נסמן $A = f^{-1}(\{1\}) = \{x \in \Omega \mid f(x) = 1\}$.
כעת מתקיים $f = \mathbb{1}_A$, כיוון שלכל $x \in A$ מתקיים $f(x) = \mathbb{1}_A(x) = 1$, ולכל $x \in A^c$ מתקיים $f(x) = \mathbb{1}_A(x) = 0$.

מסקנה תהי Ω סופית. אזי $\left|2^{\Omega}\right| = \left|\{f:\Omega\to\{0,1\}\}\right| = 2^{|\Omega|}$.

דוגמא מה מספר התת-קבוצות של $[7]$? תשובה: $2^7 = 128$.

1.2 קבוצות בנות מנייה

הגדרה קבוצה נקראת בת מניה אם יש פונקציה חח"ע מ-A ל-$\mathbb{N}$ או פונקציה על מ- $\mathbb{N}$ ל-A.
קבוצות בנות מניה תשחקנה תפקיד מרכזי בקורס שלנו.

דוגמא הקבוצות $\mathbb{N},\mathbb{Z},\mathbb{Q},\mathbb{N}^7$ הן בנות מניה אך הקבוצה $\mathbb{R}$ איננה כזו.

סימון אם A קבוצה סופית, נסמן ב-$|A|$ את מספר האיברים בקבוצה.

דוגמא הוכיחו שהקבוצה $A=\{a\in\mathbb{R}:a^2\in\mathbb{N}\}$ היא בת מנייה.

פתרון נגדיר פונקציה $f:A\to\mathbb{N}$ על ידי $f(a)=2a^2+1$ אם $a\geq 0$ ו-$f(a)=2a^2$ אם $a<0$, נוכיח שהפונקציה הזו היא חח"ע, נניח ש-$f(a)=f(b)$, אם a,b בעלי סימן שונה אז אגף אחד יהיה מספר זוגי והאגף השני יהיה מספר אי זוגי לכן $a,b\geq 0$ או $a,b<0$. נניח ללא הגבלת הכלליות כי $a,b\geq 0$ מההנחה נקבל כי $2a^2+1=2b^2+1$ מה שאומר ש-$a^2=b^2$ ומכיוון שהם שווי סימן נקבל $a=b$, לכן f באמת חח"ע והקבוצה A היא בת מנייה.

1.3 מכפלה קרטזית

הגדרה קבוצת המכפלה הקרטזית בין הקבוצה A ל-B מסומנת ע"י $A\times B$ והיא קבוצת כל הזוגות הסדורים כך שהאיבר הראשון בכל זוג סדור כזה הוא איבר מתוך A, והאיבר השני בכל זוג סדור כזה הוא איבר מתוך B:

$$.A\times B=\{(a,b):a\in A,b\in B\}$$

מכפלה קרטזית נקראת לעתים גם מכפלה ישרה.

הגדרה קבוצת המכפלה הקרטזית של n קבוצות מוגדרת ע"י:

$$.\bigtimes_{k=1}^{n}A_k=A_1\times A_2\times\cdots\times A_n=\{(a_1,a_2,\ldots,a_n),\forall k\in[n],a_k\in A_k\}$$

הסימן $[n]$ הוא אוסף כל המספרים הטבעיים עד n. כלומר $[n] := \{1, 2, .., n\}$.
מכפלה של n עותקים של אותה קבוצה A, מכונה החזקה הקרטזית מסדר n של A, ומסומנת A^n. נשים לב שזו היא קבוצת כל הסדרות הסופיות באורך n של איברים מ-A:

$$A^n = \{(a_1, a_{2,} \ldots, a_n) \mid \forall k \in [n], a_k \in A\}.$$

נוכל לזהות כל סדרה כנ"ל כפונקציה $f : \{1, 2, .., n\} \longrightarrow A$ כאשר לכל k, $k \in [n]$, $f(k) = a_k$, לכן נוכל להתייחס ל-A^n כקבוצת כל הפונקציות f מ-$\{1, 2, ..., n\}$ ל-A.
קבוצת המכפלה הקרטזית של מספר בן מניה של קבוצות מוגדרת באופן דומה על ידי:

$$\bigtimes_{i=1}^{\infty} A_i = \{(a_1, a_2, ...) \mid \forall k \in \mathbb{N}, a_k \in A_k\}.$$

מכפלה של מספר בן מניה של עותקים של אותה קבוצה A, מכונה החזקה הקרטזית מסדר $\mathbb{N}$ של A, ומסומנת $A^{\mathbb{N}}$. נשים לב שזו קבוצת כל הסדרות עם איברים ב-A, לכן נוכל לזהות אותה עם הקבוצה $\{f : \mathbb{N} \to A\}$.
באופן כללי בהינתן קבוצות A, B נסמן ב-A^B את <u>קבוצת כל הפונקציות מ-B ל-A</u>.

תזכורת נתון $A_1, A_2, ..., A_n$ קבוצות סופיות ולא ריקות. נסמן: $A = A_1 \times A_2 \times ... \times A_n$.
אזי $|A| = \prod_{i=1}^{n} |A_i|$.
בפרט, אם כל הקבוצות במכפלה זהות, נקבל $|B^n| = |B|^n$. באופן כללי יותר, אם S, T קבוצות סופיות, יתקיים $\left|S^T\right| = |S|^{|T|}$.

תרגיל

א) כמה פונקציות $g : \{x, y\} \to \{a_1, ..., a_n\}$ שונות קיימות?

ב) כמה מתוך הפונקציות שבסעיף א' הן פונקציות חד-חד-ערכיות?

פתרון נסמן $S = \{a_1, a_2, ..., a_n\}$ ו- $T = \{x, y\}$.

א) מדובר בקבוצה $S^T = \{a_1, ..., a_n\}^{\{x,y\}}$. על פי התזכורת יתקיים:

$$\left|S^T\right| = |\{a_1, ..., a_n\}|^{|\{x,y\}|} = n^2$$

ב) נסמן ב-G את משפחת הפונקציות $\{x, y\} \to \{a_1, ..., a_n\}$ החח"ע, וב-B את משפחת הפונקציות $\{x, y\} \to \{a_1, ..., a_n\}$ שאינן חח"ע. נשים לב כי מתקיים

$G = S^T \setminus B$ ו $B \subseteq S^T$, ולפיכך $|G| = \left|S^T\right| - |B|$. נותר למצוא את $|B|$. נשים לב כי $B = \{g : \{x, y\} \to \{a_1, ..., a_n\} \mid g(x) = g(y)\}$. לפיכך $g \mapsto g(x)$ היא התאמה חח״ע ועל מ-B ל-$\{a_1, ..., a_n\}$. לפיכך $|B| = n$.
מכאן נסיק כי:

$$|G| = \left|S^T\right| - |B| = n^2 - n = n(n-1).$$

2 על הסתברות וספירה

כאשר התהליך המקרי מערב בחירה של איבר מתוך קבוצה, וכל האיברים בעלי אותו סיכוי להבחר, השאלה ההסתברותית "מה הסיכוי שהבחירה תוביל לתוצאה רצויה?" הופכת לשאלה הקומבינטורית "מה הוא היחס בין מספר אפשרויות הבחירה הרצויות לבין מספר האפשרויות הכוללי?".

המושג המדויק לתיאור הסיטואציה הנ״ל הוא <u>מרחב הסתברות אחיד</u>, אותו נלמד בתרגול הבא.

דוגמא מסדרים 6 אנשים בשורה באקראי. מה הסיכוי שהם הסתדרו לפי גובה עולה (בהנחה שכולם בעלי גבהים שונים)? מה הסיכוי שיוסי עומד ליד יואל?

תשובה יש $6! = 720$ דרכים לסדר 6 אנשים בשורה. ישנה רק דרך אחת לסדר אותם לפי גובה עולה, לפיכך הסיכוי לכך הוא $\frac{1}{720}$.
על מנת לענות על השאלה השנייה, נמצא כמה דרכים יש לסדר את האנשים כך שיוסי עומד ליד יואל. נחשוב על הזוג יוסי-יואל כאדם אחד אותו אנו רוצים לסדר. לסידור חמשת ה"אנשים" ישנן $5! = 120$ אפשרויות, וכעת יש לבחור את הסידור הפנימי של הזוג יוסי-יואל, עבורו יש שתי אפשרויות. לפיכך הסיכוי להתרחשות יהיה $\frac{120 \cdot 2}{720} = \frac{1}{3}$.

3 קומבינטוריקה

(נספח א.3.1)

עבור מי שלמד קומבינטוריקה או מתמטיקה דיסקרטית, חלק זה יהווה חזרה. למי שלא, זו תהיה הצגה די חפוזה של כמה נוסחאות וטענות חשובות. מי שרוצה לדעת יותר או לתרגל יותר, הספר של נתי ליניאל ומיכל פרנס "מתמטיקה בדידה", ספציפית פרק 4, הוא מקור מצוין לדיון מסודר ולתרגילים, והוא אפילו בעברית.

3.1 בעיות מנייה

אנו מעוניינים למצוא את מספר הדרכים לבחור k איברים מתוך קבוצה בת n איברים ($k \leq n$). יש לנו שני פרמטרים שמשנים לנו:

1. האם הבחירה עם או בלי החזרה – דהיינו, האם אחרי שבחרנו איבר אנחנו יכולים לבחור בו שוב – או לא.

2. האם הסדר (שבו בחרנו את האיברים) משנה או לא.

במילים אחרות אנו רוצים למלא את הטבלה הבאה:

	עם החזרה	ללא החזרה
סדר משנה		
סדר לא משנה		

הערה הטבלה המלאה מופיעה בפרק הלמידה העצמית.

3.1.1 עם החזרות, סדר משנה.

טענה מספר הדרכים לבחור k איברים מתוך קבוצה בת n איברים, **עם חזרות כשהסדר משנה** שווה ל-n^k.

משום שהבחירה מתבצעת עם חזרות, בכל k השלבים של בחירה כנ״ל אנו בוחרים איבר מתוך **אותה** קבוצה.

הוכחה נסמן ב-A, את הקבוצה בגודל n שמתוכה נבחר k איברים. כל דרך בחירה כנ״ל מתאימה איבר $a_i \in A$ לשלב ה-i של הבחירה. כלומר כל דרך בחירה כנ״ל היא בעצם פונקציה $f : \{1, 2, .., k\} \to A$ – או סדרה $(a_1, a_2, ..., a_k) \in A$. אז הבעיה שלנו שקולה לבעיית מציאת מספר הפונקציות $A^{\{1,2,...,k\}}$. מהתזכורת נקבל $\left|A^{\{1,2,...,k\}}\right| = |A|^{|\{1,2,..,k\}|} = n^k$.

דוגמא כמה מספרים בינריים עם n ספרות יש?
במקרה שלנו $A_i = \{0, 1\}$ לכל $1 \leq i \leq n$. ולכן יש 2^n מספרים שונים. שימו לב שהמכפלה הקרטזית אכן מתארת את פתרון הבעיה שלנו, כי איברי $\{0, 1\}^n$ הם n-יות של 0 ו-1-ים שהסדר שלהם משנה.

3.1.2 ללא החזרות, סדר משנה.

טענה מספר הדרכים לבחור k איברים מתוך קבוצה בת n איברים, **ללא חזרות כשהסדר משנה** שווה ל-

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

משום שהבחירה מתבצעת ללא חזרות, בכל שלב מתוך k השלבים של בחירה כנ״ל אנו בוחרים איבר מתוך קבוצת אחרת. ראשית מתוך קבוצה בגודל n, לאחר מכן מתוך קבוצה בגודל $n-1$ של איברים שנותרו, וכו'.

דוגמא בכיתה 52 תלמידים המתחרים בתחרות ריצה עבור חמשת המקומות הראשונים. כמה אפשרויות יש לתוצאת התחרות?
עלינו לבחור חמישה תלמידים מתוך קבוצה של 52, ללא חזרות כשהסדר משנה, לכן התשובה היא $\frac{52!}{47!}$.

3.1.3 ללא החזרות, סדר לא משנה.

תזכורת מספר הדרכים לבחור k איברים מתוך קבוצה בעלת n איברים, ללא חזרות וללא חשיבות לסדר היא $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, הביטוי הזה נקרא מקדם בינומי.

דוגמא מה מספר הדרכים לבחור 3 סטודנטים מתוך כיתה של 17? התשובה: $\binom{17}{3} = 680$.

3.1.4 עם החזרות, סדר לא משנה.

מופיע בפרק הלמידה העצמית.

3.2 סידורים בשורה ובמעגל

תרגיל בכמה דרכים ניתן לסדר חבילת קלפים תקנית, כך שכל הקלפים מאותו המספר מופיעים ברצף זה אחר זה? מה הסיכוי שחבילה אקראית תסודר באופן הזה?

פתרון בחבילת קלפים תקנית יש 52 קלפים: 13 מספרים – $\{2, 3, 4, 5, ..., 10, J, Q, K, A\}$, וכל מספר מופיע ב-4 צורות שונות – $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$.
נוכל לספור את הסידורים באופן הבא: ראשית עלינו לבחור את הסידור של המספרים

עצמם. יש $13!$ דרכים לעשות זאת.

כעת, יש לבחור את הסידור הפנימי של כל רביעייה - לכל סידור כזה יש $4!$ אפשרויות, ועלינו לעשות זאת 13 פעמים. לפיכך מספר הסידורים הוא $13!\cdot(4!)^{13} \approx 5.48\cdot 10^{27}$.
מספר הסידורים באופן כללי של חבילה בעלת 52 קלפים הוא $52!$, לכן הסיכוי שחבילה אקראית תסודר כך שכל הקלפים מאותו המספר מופיעים ברצף זה אחר זה הוא

$$\frac{13!\cdot(4!)^{13}}{52!} \approx 6.76\cdot 10^{-41}.$$

4 טורים חיוביים

(נספח א.4.5)

הגדרה תהי $(a_n)_{n=1}^{\infty}$ סדרה של מספרים אי-שליליים. נגדיר סדרה חדשה $(S_k)_{k=1}^{\infty}$ הנקראת **סדרת הסכומיים החלקיים** על ידי

$$S_k = \sum_{n=1}^{k} a_n = a_1 + \cdots + a_k \qquad \forall k \in \mathbb{N}.$$

כאשר סדרת הסכומיים החלקיים מתכנסת, נאמר ש**הטור** $\sum\limits_{n=1}^{\infty} a_n$ מתכנס, וסכומו שווה ל-$\lim\limits_{k\to\infty} S_k$.

טענה אם הטור $\sum\limits_{n=1}^{\infty} a_n$ מתכנס אז הסדרה $(a_n)_{n=1}^{\infty}$ שואפת לאפס.

הוכחה נסמן ב-L את סכום הטור, כלומר $L := \lim\limits_{k\to\infty} S_k$. נשים לב שמתקיים

$$a_k = \sum_{n=1}^{k} a_n - \sum_{n=1}^{k-1} a_n = S_k - S_{k-1}$$

ומאריתמטיקה של גבולות

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} S_k - \lim_{k\to\infty} S_{k-1} = L - L = 0$$

הערה השאיפה של a_n לאפס היא תנאי הכרחי להתכנסות הטור $\sum\limits_{n=1}^{\infty} a_n$, אך לא תנאי מספיק!

דוגמאות

1. הטור ההרמוני $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ מתבדר.
2. הטור $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ מתכנס.
3. הטור הגיאומטרי $\sum\limits_{n=1}^{\infty} q^n$ מתכנס אם ורק אם $|q| < 1$.

תזכורת הטענות הבאות לגבי טורים חיוביים, אותן הוכחתם באינפי, ישמשו אותנו במהלך הקורס.

משפט (מבחן ההשוואה) יהיו $(a_n), (b_n)$ זוג סדרות כך ש- $0 \leq a_n \leq b_n$ כמעט תמיד. אם $\sum\limits_{n=1}^{\infty} b_n$ מתכנס אז גם $\sum\limits_{n=1}^{\infty} a_n$ מתכנס.

משפט (שינוי סדר) בטורים חיוביים, שינוי סדר הסכימה לא משפיע על היותו של טור מתכנס או מתבדר, ובמקרה של התכנסות, לא משפיע על ערכו.
באופן פורמלי – תהי (a_n) סדרת מספרים אי-שליליים, ותהי $\sigma : \mathbb{N} \to \mathbb{N}$ פונקציה הפיכה (נקראת גם "תמורה"). הטור $\sum\limits_{n=1}^{\infty} a_n$ מתכנס אם ורק אם הטור $\sum\limits_{n=1}^{\infty} a_{\sigma(n)}$ מתכנס, ובמקרה של התכנסות $\sum\limits_{n=1}^{\infty} a_n = \sum\limits_{n=1}^{\infty} a_{\sigma(n)}$.

הערה ההגדרה של טורים מאפשרת לנו לטפל באופן מדויק בסכום של אוסף בן-מניה של מספרים. במהלך הקורס, לעיתים נרצה גם לסכום אוסף לא בן-מניה של מספרים.

הגדרה תהי I קבוצה (לא בהכרח בת-מניה). נחשוב עליה כקבוצת אינדקסים, ולכל $j \in I$ יהי a_j מספר אי-שלילי. נגדיר את הסכום של הקבוצה $\{a_j\}_{j\in I}$ על ידי

$$\sum_{j\in I} a_j := \sup\{a_{j_1} + ... + a_{j_n} \mid j_1, ..., j_n \in I, \quad j_1 \neq ... \neq j_n, \quad n \in \mathbb{R}\}.$$

כלומר, הסופרמום של כל הסכומים הסופיים האפשריים של איברים ב-$\{a_j\}_{j\in I}$.

הסימון $\sum\limits_{j\in I} a_j$ יהיה לנו נוח לשימוש בהמשך, אבל מסתבר שהוא לא מגיע עם עומק מתמטי חדש מעבר לטורים:

אם $\sum\limits_{j\in I} a_j < \infty$, אז בהכרח קבוצת האינדקסים j עבורם $a_j \neq 0$ היא בת-מניה. במקרה זה, ניתן לסדר את אותם איברים ששונים מאפס בסדר שרירותי כלשהו, ולסכום אותם לפי ההגדרה הרגילה של טור חיובי. משום שסדר הסכימה לא משפיע, ניתן להוכיח שסכום הטור הנ״ל יהיה שווה ל-$\sum\limits_{j\in I} a_j$.

הגדרה טור $\sum\limits_{n=1}^{\infty} a_n$ נקרא מתכנס בהחלט אם הטור $\sum\limits_{n=1}^{\infty} |a_n|$ מתכנס.

מבוא להסתברות וסטטיסטיקה 80430 - תרגול 8

1 נוסחת התוחלת השלמה

הגדרה. יהא X משתנה מקרי בדיד על מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$ ויהי $A\in\mathcal{F}$ מאורע עם $\mathbb{P}(A)>0$. ה**תוחלת** של X **בהינתן** A היא

$$\mathbb{E}(X\mid A)=\sum_{\omega\in\Omega}X(\omega)\,\mathbb{P}(\{\omega\}\mid A)=\sum_{x\in\text{Supp}(X)}x\mathbb{P}_X(x\mid A).$$

טענה. **(נוסחת התוחלת השלמה)** יהא X משתנה מקרי בעל תוחלת ו– $\{A_1,A_2,\dots\}$ חלוקה של מרחב ההסתברות, אזי

$$\mathbb{E}(X)=\sum_{n\in\mathbb{N}}\mathbb{E}(X\mid A_n)\,\mathbb{P}(A_n).$$

כאשר מפרשים איברים בסכום כאפסים אם $\mathbb{P}(A_n)=0$ (מותר להתנות רק במאורעות עם הסתברות חיובית).

תרגיל 1. מטילים קוביה הוגנת עד שיוצא המספר 6 , נסמן ב- X את המשתנה המקרי שמחזיר את מספר ההטלות, מה התוחלת של X אם ידוע שיצאו רק מספרים זוגים בקוביה? (נסמן ב- A את המאורע השיצאו רק מספרים זוגיים בקוביה). כעת מטילים קוביה הוגנת עם שלוש פאות עם המספרים $2,4,6$ עד שיוצא המספר 6 נסמן ב- Y את המשתנה המקרי שמחזיר את מספר ההטלות, האם מתקיים $Y\overset{d}{=}(X|A)$?

פתרון: ראשית נסמן ב- A את המאורע שיצאו רק מספרים זוגים, כלומר

$$A=\left\{[2,4]^{n-1}\times\{6\}\;:\;n\in\mathbb{N}\right\}$$

נרצה לחשב את התוחלת של $(X|A)$, קודם נחשב את התפלגותו, הוא כמובן נתמך על הטבעים. יהי $n\in\mathbb{N}$, אזי

$$\begin{aligned}\mathbb{P}\left(X=n|A\right)&=\frac{\mathbb{P}\left(X=n,A\right)}{\mathbb{P}\left(A\right)}=\frac{\mathbb{P}\left([2,4]^{n-1}\times\{6\}\right)}{\sum_{k=1}^{\infty}\mathbb{P}\left([2,4]^{k-1}\times\{6\}\right)}\\&=\frac{\left(\frac{1}{3}\right)^{n-1}\frac{1}{6}}{\sum_{k=1}^{\infty}\left(\frac{1}{3}\right)^{k-1}\frac{1}{6}}=\left(\frac{1}{3}\right)^{n-1}\left(\sum_{k=0}^{\infty}\left(\frac{1}{3}\right)^{k}\right)^{-1}\\&=\left(\frac{1}{3}\right)^{n-1}\left(\frac{1}{1-\frac{1}{3}}\right)^{-1}=\left(\frac{1}{3}\right)^{n-1}\frac{2}{3}\end{aligned}$$

קיבלנו ש- $(X|A)\sim Geo\left(\frac{2}{3}\right)$ ולכן התוחלת של $(X|A)$ הינה

$$\mathbb{E}\left((X|A)\right)=\frac{1}{\frac{2}{3}}=\frac{3}{2}$$

בעניין השאלה השנייה אנחנו כבר יודעים להגיד ש- $Y\sim Geo\left(\frac{1}{3}\right)$ לכן $Y\overset{d}{\neq}(X|A)$.

תרגיל 2. מטילים קוביה פעם אחר פעם ועוצרים כאשר התקבל בפעם הראשונה המספר 1. מהי תוחלת סכום ערכי ההטלות שהתקבלו?

פתרון: יהי $X\sim Geo\left(\frac{1}{6}\right)$ משתנה מקרי שתוצאתו היא מספר ההטלות עד שהתקבל המספר 1. לכל $i\in\mathbb{N}$ נגדיר $Y_i\sim Unif([6])$ משתנה מקרי שתוצאתו היא ערך ההטלה ה–i, ו– Y משתנה מקרי שתוצאתו היא סכום ערכי ההטלות עד (כולל) לקבלת הערך 1. אזי מנוסחת התוחלת השלמה נקבל:

$$\begin{aligned}\mathbb{E}\left(Y\right)&=\sum_{n=1}^{\infty}\mathbb{E}\left(Y\mid X=n\right)\mathbb{P}\left(X=n\right)\\&=\sum_{n=1}^{\infty}\mathbb{E}\left(\sum_{i=1}^{n}Y_i\mid X=n\right)\left(\frac{5}{6}\right)^{n-1}\frac{1}{6}\\&=\sum_{n=1}^{\infty}\sum_{i=1}^{n}\mathbb{E}\left(Y_i\mid X=n\right)\left(\frac{5}{6}\right)^{n-1}\frac{1}{6}.\end{aligned}$$

לא קשה לראות כי עבור $1\le i\le n-1$ מתקיים $(Y_i\mid X=n)\sim Unif(\{2,3,4,5,6\})$, ולכן $\mathbb{E}\left(Y_i\mid X=n\right)=$

$4 = \frac{2+6}{2}$. כמו כן $(Y_n \mid X = n) \equiv 1$ (פונקציה קבועה) ולכן התוחלת שלו היא 1. אזי:

$$\begin{aligned}
\mathbb{E}\left(Y\right) &= \sum_{n=1}^{\infty}\left(\left(n-1\right)4+1\right)\left(\frac{5}{6}\right)^{n-1}\frac{1}{6} \\
&= \sum_{n=1}^{\infty}\left(4n-3\right)\left(\frac{5}{6}\right)^{n-1}\frac{1}{6} \\
&= 4\sum_{n=1}^{\infty}n\left(\frac{5}{6}\right)^{n-1}\frac{1}{6}-\frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{5}{6}\right)^{n-1}.
\end{aligned}$$

קיימות מספר דרכים לחשב את סכום הטור הראשון, אבל במקום לחשב טורים כאלה מחדש, נבחין כי זו למעשה התוחלת של משתנה מקרי $\mathrm{Geo}\left(\frac{1}{6}\right)$, שאותה כבר חישבנו והיא שווה ל–$6 = \frac{1}{\frac{1}{6}}$. הטור השני הוא פשוט טור גאומטרי שסכומו הוא $6 = \frac{1}{1-\frac{5}{6}}$. כלומר בסך הכל:

$$\mathbb{E}\left(Y\right) = 4\cdot 6-\frac{1}{2}\cdot 6 = 24-3 = 21.$$

תרגיל 3. איתמר ואלעד, נרגשים אחרי ששמעו על הצלחתה של עליסה להגיע לארץ הפלאות, מנסים להגיע לשם גם הם. השניים מגיעים לחדר מבואה ובו 3 דלתות. אם ייכנסו בדלת הראשונה, יגיעו לארץ הפלאות תוך 30 דקות. אם ייכנסו בדלת השניה, יחזרו חזרה לחדר המבואה תוך 50 דקות. אם ייכנסו בדלת האחרונה, הם יחזרו חזרה לחדר המבואה תוך 70 דקות. אחרי ששתו מהשיקוי הסגול בחדר, איתמר ואלעד מאוד מבולבלים ועייפים ולכן בכל פעם שהם חוזרים לחדר המבואה הם בוחרים באקראי דלת לעבור דרכה. מהי תוחלת הזמן שייקח להם להגיע לארץ הפלאות?

פתרון: נסמן ב–X את הזמן שלוקח להם להגיע לארץ הפלאות וב–Y את הדלת בה הם בוחרים בפעם הראשונה שהם

מגיעים לחדר המבואה. נרצה להשתמש בנוסחת ההסתברות השלמה ולשם כך נבחין ראשית כי:

$$\begin{aligned}
\mathbb{E}(X|Y=2) &= \sum_{x\in supp(X)} x\mathbb{P}(X=x|Y=2) \\
&= \sum_{x\in supp(X)} x\frac{\mathbb{P}(X=x,Y=2)}{\mathbb{P}(Y=2)} \\
&= \sum_{x\in supp(X)} x\frac{\mathbb{P}(X=x-50)\cdot\mathbb{P}(Y=2)}{\mathbb{P}(Y=2)} \\
&= \sum_{x\in supp(X)} x\mathbb{P}(X=x-50) \\
&= \sum_{z\in supp(X)} (z+50)\mathbb{P}(X=z) \\
&= \mathbb{E}(X+50) \\
&= \mathbb{E}(X)+50.
\end{aligned}$$

מחישוב דומה נסיק כי:

$$\mathbb{E}(X|Y=3) = \mathbb{E}(X)+70$$

לכן, מנוסחת ההסתברות השלמה:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{n=1}^{3}\mathbb{E}(X|Y=n)\mathbb{P}(Y=n) \\
&= \sum_{n=1}^{3}\mathbb{E}(X|Y=n)\frac{1}{3} \\
&= \frac{1}{3}(30+(50+\mathbb{E}(X))+(70+\mathbb{E}(X)) \\
&= 50+\frac{2}{3}\mathbb{E}(X)
\end{aligned}$$

נכפול ב–3 ונעביר אגפים. מתקבל השוויון: $\mathbb{E}(X)=150$.

הערה. שימו לב שצריך להראות שהתוחלת של X סופית כדי שנוכל לבצע את החישוב האחרון. סופיותה נובעת למשל מכך שאם נגדיר את Z להיות מספר הביקורים של השניים בחדר המבואה, אז $Z\sim Geo(\frac{1}{3})$ ו–$X\le 70Z$ (כל ביקור

תורם לכל היותר 70 דקות לתוחלת), לכן ממונוטוניות:

$$\mathbb{E}(X) \le \mathbb{E}(70Z) = 70\mathbb{E}(Z) = 70 \cdot 3 = 210.$$

2 אי שוויון מרקוב

ביסוד של אי שוויון מרקוב עומדת האבחנה שלא ייתכן שבהסתברות מלאה כל הערכים של משתנה מקרי ירוכזו רק בצד אחד של התוחלת. באופן אינטואיטיבי יותר, עבור ממוצע נתון, חייבים להיות ערכים הן מתחת לממוצע והן מעל לממוצע.

טענה. (אי-שוויון מרקוב) עבור X מ״מ בדיד המקבל **ערכים אי שליליים** ובעל תוחלת סופית, לכל $c > 0$,

$$\mathbb{P}\left(X \ge c\right) \le \frac{\mathbb{E}\left(X\right)}{c}.$$

תרגיל 4. ילד מקבל סכום כסף מהוריו בכל בוקר לקנות חטיף. בכל חטיף פוג אחד מבין n סוגים שונים באקראי. נחשב את תוחלת הזמן שייקח לילד להשיג את כל הפוגים.

פתרון: עבור $i \in [n]$ יהי X_i משתנה מקרי שתוצאתו היא מספר הימים שלוקח לילד להשיג פוג חדש לאחר שהשיג $i-1$ פוגים מסוגים שונים. מתקיים כי

$$.X_i \sim Geo\left(\frac{n-(i-1)}{n}\right)$$

הזמן בו ייקח לילד להשיג את כל הפוגים הוא $X = \sum_{i=1}^{n} X_i$. מלינאריות התוחלת ומנוסחת התוחלת למשתנה מקרי גאומטרי מתקיים:

$$\begin{aligned}\mathbb{E}\left(X\right) &= \sum_{i=1}^{n} \mathbb{E}\left(X_i\right) = \sum_{i=1}^{n} \frac{n}{n-(i-1)} \\ &\overset{(j=n-(i-1))}{=} n\sum_{j=1}^{n} \frac{1}{j} \le n\left(log\left(n\right)+1\right)\end{aligned}$$

נוכל להשתמש בתוצאה זו כדי לחסום מאורעות שונים באמצעות אי-שוויון מרקוב. נתבונן למשל בשאלה הבאה: לאחר כמה ימים ניתן לקבוע כי ההסתברות שהילד עוד לא ישיג את כל הפוגים תהיה לכל היותר $\frac{1}{2}$? כלומר, עלינו למצוא

$t \in \mathbb{N}$ שעבורו $\mathbb{P}\left(X \geq t\right) \leq \frac{1}{2}$. לפי אי-שוויון מרקוב מתקיים:

$$\mathbb{P}\left(X \geq t\right) \leq \frac{\mathbb{E}\left(X\right)}{t} \leq \frac{n\left(log(n)+1\right)}{t} \leq \frac{1}{2}$$

ולכן נוכל לקבוע זאת עבור כל $t \geq 2n\left(log\left(n\right)+1\right)$, זה כמובן לא אומר שמדובר בחסם אופטימלי, כלומר, ייתכן שמסיבות אחרות אפשר להבטיח כי $\mathbb{P}\left(X \geq t\right) \leq \frac{1}{2}$ גם עבור זמנים קצרים יותר.

תרגיל 5. (גרף ארדש–רני) נקבע $n \in \mathbb{N}$ ונקבע $0 < p < 1$. נגדיר "גרף מקרי" $G\left(n,p\right)$ באופן הבא: נקבע n קודקודים, ובין כל זוג קודקודים שונים נמתח צלע (בלתי מכוונת) בהסתברות p, באופן בלתי תלוי בין הקודקודים והצלעות.

באופן יותר פורמלי, יהי Ω מרחב כל הגרפים על n קודקודים. נבחין כי כל גרף נקבע על ידי הצלעות שלו ושכל צלע נקבעת על ידי זוג קודקודים, ולכן $|\Omega| = 2^{\binom{n}{2}}$. נקבע $0 < p < 1$ ונגדיר פונקציית הסתברות (נקודתית) על Ω על ידי

$$\mathbb{P}\left(\{G\}\right) = p^{|E|}\left(1-p\right)^{\binom{n}{2}-|E|},$$

כאשר E היא קבוצת הצלעות בגרף G. שתי אבחנות יסודיות:

- יהי X משתנה מקרי שתוצאתו היא מספר הצלעות בגרף. אזי $Supp\left(X\right) = \left\{0,1,\ldots,\binom{n}{2}\right\}$ וכן $X \sim Bin\left(\binom{n}{2},p\right)$, לכן $\mathbb{E}\left(X\right) = \binom{n}{2}p$.
- לכל קודקוד $i = 1,\ldots,n$, יהי D_i משתנה מקרי שתוצאתו היא דרגת הקודקוד בגרף, כלומר מספר הצלעות שיוצאות ממנו. אזי $\mathrm{Supp}\left(D_i\right) = \{0,1,\ldots,n-1\}$ וכן $D_i \sim \mathrm{Bin}\left(n-1,p\right)$.

נשתמש באי-שוויון מרקוב כדי להעריך את ההסתברות שקיים משולש בגרף $G\left(n,p\right)$. נזכיר כי "משולש" הוא שלישיית קודקודים שכל שניים מהם מחוברים בצלע. נבחין כי חישוב מדויק של ההסתברות הזאת הוא קשה, היות והמשולשים השונים תלויים זה בזה ולכן נסתפק בחסם. כל $i,j,k \in \{1,\ldots,n\}$ שונים זה מזה מהווים משולש בהסתברות p^3 (צריכות לעבור 3 צלעות בין הקודקודים). נגדיר אם כך משתנה מקרי $X_{i,j,k} \sim Ber\left(p^3\right)$ שמקבל את הערך 1 אם ורק אם i,j,k מהווים משולש. מספר המשולשים בגרף $G\left(n,p\right)$ מתואר על ידי המשתנה המקרי:

$$T = \sum_{i,j,k} X_{i,j,k}$$

כאשר סכום זה הוא על פני כל השלשות של קודקודים שונים. לפי אי-שוויון מרקוב נקבל

$$\mathbb{P}\left(T \geq 1\right) \leq \frac{\mathbb{E}\left(T\right)}{1} = \sum_{i,j,k} \mathbb{E}\left(X_{i,j,k}\right) = \sum_{i,j,k} p^3 = \binom{n}{3}p^3.$$

3 שונות

הגדרה: עבור מ"מ בדיד X המקיים $\mathbb{E}\left(X^2\right)<\infty$, **השונות** היא

$$\mathrm{Var}\left(X\right)=\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^2\right)=\mathbb{E}\left(X^2\right)-\mathbb{E}\left(X\right)^2.$$

תכונות: נזכיר כמה תכונות יסודיות של השונות: עבור מ"מ X בעל שונות סופית, לכל קבוע $c\in\mathbb{R}$ מתקיים:

1. השונות היא תמיד אי-שלילית: $0\le\mathrm{Var}\left(X\right)$. כמו כן $\mathrm{Var}\left(X\right)=0$ אם ורק אם X מ"מ קבוע, כלומר $\mathbb{P}\left(X=\mathbb{E}\left(X\right)\right)=1$.
2. מתקיים $\mathrm{Var}\left(X+c\right)=\mathrm{Var}\left(X\right)$.
3. מתקיים $\mathrm{Var}\left(cX\right)=c^2\mathrm{Var}\left(X\right)$.
4. אם X,Y מ"מ בלתי תלויים ובעלי שונות סופית, אזי

$$\mathrm{Var}\left(X+Y\right)=\mathrm{Var}\left(X\right)+\mathrm{Var}\left(Y\right)$$

תכונה זו נכונה כמובן גם לכל סכום סופי של מ"מ ב"ת ומספיק גם ב"ת בזוגות כפי שנבחין בהמשך התרגול.

3.1 שונות משתנה מקרי ברנולי

יהי $\mathrm{X}\sim\mathrm{Ber}\left(p\right)$. אזי:

$$\mathrm{Var}\left(\mathrm{X}\right)=\mathbb{E}\left(\mathrm{X}^2\right)-\mathbb{E}\left(\mathrm{X}\right)^2=\mathbb{E}\left(\mathrm{X}\right)-\mathbb{E}\left(\mathrm{X}\right)^2=p-p^2=p\left(1-p\right).$$

כאשר השתמשנו בכך שאם X מתפלג ברנולי, אז $\mathrm{X}^2\overset{a.s}{=}\mathrm{X}$.

3.2 שונות משתנה מקרי בעל התפלגות אחידה

יהי $\mathrm{X}\sim\mathrm{Unif}\left(\left[n\right]\right)$. אזי:

$$\mathrm{Var}\left(\mathrm{X}\right)=\sum_{k\in\left[n\right]}k^2\mathbb{P}\left(\mathrm{X}=k\right)-\left(\frac{n+1}{2}\right)^2=\frac{n\left(n+1\right)\left(2n+1\right)}{6n}-\frac{\left(n+1\right)^2}{4}=\frac{n^2-1}{12}.$$

כאשר השיוויון השני הוא סכום של סדרת n הריבועים הראשונים.

3.3 שונות משתנה מקרי בינומי

יהי $X \sim \mathrm{Bin}(n,p)$. על מנת לחשב את השונות, נזכר כי $X \overset{d}{=} \sum_{k\in[n]} Y_k$ כאשר $\forall k \in [n] : Y_k \sim \mathrm{Ber}(p)$ משתנים מקריים בלתי תלויים. לפי תכונה 4 נקבל:

$$\mathrm{Var}(X) = \mathrm{Var}\left(\sum_{k=1}^{k} Y_k\right) = \sum_{k=1}^{n} \mathrm{Var}(Y_k) = np(1-p).$$

3.4 שונות משתנה מקרי פואסוני

יהי $X \sim \mathrm{Poi}(\lambda)$. נחשב תחילה את $\mathbb{E}\left(X^2\right)$

$$\begin{aligned}\mathbb{E}\left(X^2\right) &= \sum_{n=0}^{\infty} \frac{n^2\lambda^n}{n!e^{\lambda}} = \sum_{n=1}^{\infty} \frac{n^2\lambda^n}{n!e^{\lambda}} = \lambda \sum_{n=1}^{\infty} \frac{n\lambda^{n-1}}{(n-1)!e^{\lambda}} \overset{m=n-1}{=} \lambda \sum_{m=0}^{\infty} \frac{(m+1)\lambda^m}{m!e^{\lambda}} \\ &= \lambda\mathbb{E}(X+1) = \lambda^2 + \lambda\end{aligned}$$

וכעת נוכל לחשב את השונות:

$$\mathrm{Var}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

3.5 שונות משתנה מקרי גאומטרי

יהי $X \sim \mathrm{Geo}(p)$. נזכר בנוסחה לסכום של טור הנדסי עבור $|q| < 1$:

$$\sum_{n=1}^{\infty} q^n = \frac{1}{1-q}.$$

נגזור את שני צדדים פעמיים (מותר לעשות זאת מאחר שטור הנגזרות הוא טור חזקות המתכנס במידה שווה בסביבת q):

$$\sum_{n=2}^{\infty} n(n-1)q^{n-2} = \frac{2}{(1-q)^3}.$$

כעת, נחשב את $\mathbb{E}\left(\mathrm{X}^2\right)$:

$$\mathbb{E}\left(\mathrm{X}^2\right)=\sum_{n=1}^{\infty} n^2 p\,(1-p)^{n-1}=\sum_{n=1}^{\infty} n\,(n-1)\,p\,(1-p)^{n-1}+\sum_{n=1}^{\infty} np\,(1-p)^{n-1}\,.$$

נשים לב שעל פי הגדרה מתקיים $\sum_{n=1}^{\infty} np\,(1-p)^{n-1}=\mathbb{E}\,(\mathrm{X})=\frac{1}{p}$. נחשב את הסכום השני:

$$\sum_{n=1}^{\infty} n\,(n-1)\,p\,(1-p)^{n-1}=p\,(1-p)\sum_{n=2}^{\infty} n\,(n-1)\,(1-p)^{n-2}\overset{q=1-p}{=}p\,(1-p)\,\frac{2}{p^3}=\frac{2\,(1-p)}{p^2}.$$

סה"כ:

$$\mathrm{Var}\,(\mathrm{X})=\frac{2\,(1-p)}{p^2}+\frac{1}{p}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

מבוא להסתברות וסטטיסטיקה 80430
תרגול 9

1 שונות

הגדרה: עבור מ״מ בדיד X המקיים $\mathbb{E}\left(X^2\right) < \infty$, **השונות** היא

$$\mathrm{Var}\left(X\right) = \mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^2\right) = \mathbb{E}\left(X^2\right) - \mathbb{E}\left(X\right)^2.$$

תכונות: נזכיר כמה תכונות יסודיות של השונות: עבור מ״מ X בעל שונות סופית, לכל קבוע $c \in \mathbb{R}$ מתקיים:

1. השונות היא תמיד אי-שלילית: $0 \le \mathrm{Var}\left(X\right)$. כמו כן $\mathrm{Var}\left(X\right) = 0$ אם ורק אם X מ"מ קבוע, כלומר $\mathbb{P}\left(X = \mathbb{E}\left(X\right)\right) = 1$.
2. מתקיים $\mathrm{Var}\left(X+c\right) = \mathrm{Var}\left(X\right)$.
3. מתקיים $\mathrm{Var}\left(cX\right) = c^2\mathrm{Var}\left(X\right)$.
4. אם X,Y מ״מ בלתי תלויים ובעלי שונות סופית, אזי

$$\mathrm{Var}\left(X+Y\right) = \mathrm{Var}\left(X\right) + \mathrm{Var}\left(Y\right)$$

תכונה זו נכונה כמובן גם לכל סכום סופי של מ״מ ב״ת ומספיק גם ב״ת בזוגות כפי שנבחין בהמשך התרגול.

תרגיל 1. בכיתה n תלמידים. בהנחה שימי ההולדת של כל התלמיד מתפלגים אחיד על פני שנה נתונה והם בלתי-תלויים, מהי התוחלת והשונות של מספר הזוגות של תלמידים שיש להם יום-הולדת באותו היום?

פתרון: יהי X משתנה מקרי שסופר את הזוגות שיש להם יום-הולדת באותו היום. נגדיר $Y_i \sim \mathrm{Unif}\left\{1,\ldots,365\right\}$ משתנה מקרי שמחזיר את יום ההולדת של הילד ה i. לכל $i,j \in [n]$ יהי $X_{i,j} \sim \mathrm{Ber}\left(1/365\right)$ משתנה מקרי שבודק

האם לזוג התלמידים i,j יש אותו יום-הולדת. נבחין כי $X=\sum_{1\le i<j\le n}X_{i,j}$ (כי $X_{i,j}=X_{j,i}$). מכאן נובע כי

$$.\mathbb{E}\left(X\right)=\sum_{1\le i<j\le n}\mathbb{E}\left(X_{i,j}\right)=\binom{n}{2}/365=\frac{n\left(n-1\right)}{730}$$

נרצה לחשב את השונות על ידי סכימה של השונות של המשתנים המקריים $\left\{X_{i,j}\right\}$ בדומה לחישוב התוחלת. בשונות בנגוד לתוחלת לא תמיד אפשר לעשות זאת ולכן נרצה להראות שהמשתנים ב״ת בזוגות. כלומר לכל 2 משתנים עם אינדקס אחד לפחות שונה (בלי הגבלת הכלליות $n\notin\{i,j\}$) מתקיים[1]

$$\mathbb{P}\left(X_{ij}=1,X_{mn}=1\right)=\mathbb{P}\left(X_{ij}=1\right)\mathbb{P}\left(X_{mn}=1\right) \tag{1}$$

מאחר ו- $X_{i,j}\sim\text{Ber}\left(1/365\right)$ אנחנו יודעים מהו הביטוי הימני במשוואה, נחשב את הביטוי השמאלי

$$\mathbb{P}\left(X_{ij}=1,X_{mn}=1\right)=\mathbb{P}\left(Y_j=Y_i,Y_n=Y_m\right)=\begin{cases}\mathbb{P}\left(Y_j=Y_i\right)\cdot\mathbb{P}\left(Y_n=Y_m\right) & \{j,i\}\cap\{n,m\}=\emptyset\\ \mathbb{P}\left(Y_j=Y_i=Y_n\right) & m\in\{j,i\}\end{cases}$$

$$=\begin{cases}\left(\frac{1}{365}\right)^2 & \{j,i\}\cap\{n,m\}=\emptyset\\ \sum_{k=1}^{365}\mathbb{P}\left(Y_j=Y_i=Y_n=k\right) & m\in\{j,i\}\end{cases}$$

$$=\begin{cases}\left(\frac{1}{365}\right)^2 & \{j,i\}\cap\{n,m\}=\emptyset\\ \sum_{k=1}^{365}\mathbb{P}\left(Y_j=k\right)\mathbb{P}\left(Y_i=k\right)\mathbb{P}\left(Y_n=k\right) & m\in\{j,i\}\end{cases}$$

$$=\begin{cases}\left(\frac{1}{365}\right)^2 & \{j,i\}\cap\{n,m\}=\emptyset\\ \left(\frac{1}{365}\right)^2 & m\in\{j,i\}\end{cases}$$

כאשר השתמשנו בנתון שימי ההולדת בלתי תלויים, לכן משוואה 1 מתקיימת ומכך המשתנים המקריים $\left\{X_{i,j}\right\}$ בלתי-תלויים. אם כך נקבל

$$\text{Var}\left(X\right)=\sum_{1\le i<j\le n}\text{Var}\left(X_{i,j}\right)$$

נוכל להשתמש בשונות שחישבנו קודם עבור משתנה ברנולי ולחשב: $\text{Var}\left(X_{i,j}\right)=p\left(1-p\right)=\frac{1}{365}\left(1-\frac{1}{365}\right)$.

[1] אנחנו נעזרים פה בטענה שמשתנים מקריים שהם מציינים הם בלתי-תלויים אם ורק אם המאורעות שמגדירים אותם בלתי-תלויים

2 אי-שוויון צ׳בישב

טענה. **(אי-שוויון צ׳בישב)** יהי X מ״מ בדיד בעל תוחלת ושונות סופיות. אזי לכל $0 < c$,

$$\mathbb{P}\left(\left|X - \mathbb{E}\left(X\right)\right| \geq c\right) \leq \mathrm{Var}\left(X\right)/c^{2}.$$

אחת הדרכים להשתמש באי-שוויון צ׳בישב היא להבחין כי באופן כללי אם X הוא משתנה מקרי בעל תוחלת, ואם $c \in \mathbb{R}$ מקיים $c - \mathbb{E}\left(X\right) > 0$, אז

$$\begin{aligned}
\mathbb{P}\left(X \geq c\right) &= \mathbb{P}\left(X - \mathbb{E}\left(X\right) \geq c - \mathbb{E}\left(X\right)\right) \\
&\leq \mathbb{P}\left(\left|X - \mathbb{E}\left(X\right)\right| \geq c - \mathbb{E}\left(X\right)\right) \\
&\leq \mathrm{Var}\left(X\right)/\left(c - \mathbb{E}\left(X\right)\right)^{2}
\end{aligned}$$

תרגיל 2. עליסה חזרה לחדר מבואה אחר, בו יש שתי דלתות. בסיכוי p בוחרת בדלת המובילה אותה לארץ הפלאות, ובסיכוי $1-p$ לדלת המחזירה אותה חזרה לדלת המבואה. חסמו את הסיכוי של עליסה להגיע לארץ הפלאות לפחות לאחר עשרה נסיונות.

פתרון: יהי X משתנה מקרי המתאר את כמות הניסיונות של עליסה עד שסוף סוף היא פותחת את הדלת לארץ הפלאות (כולל). ע״פ הנתון, $\mathrm{X} \sim \mathrm{Geo}\left(p\right)$. נשתמש באי-שיוויון צ׳בישב:

$$\begin{aligned}
\mathbb{P}\left(\mathrm{X} \geq 10\right) \leq \mathrm{Var}\left(\mathrm{X}\right)/\left(10 - \mathbb{E}\left(\mathrm{X}\right)\right)^{2} &= \frac{1-p}{p^{2}}/\left(10 - \frac{1}{p}\right)^{2} \\
&= \frac{\frac{1-p}{p^{2}}}{\frac{(10p-1)^{2}}{p^{2}}} = \frac{1-p}{\left(10p-1\right)^{2}}.
\end{aligned}$$

3 שונות משותפת

הגדרה: עבור מ״מ בדידים X, Y בעלי שונות סופית, ה**שונות המשותפת** היא

$$\mathrm{Cov}\left(X,Y\right) = \mathbb{E}\left(\left(X - \mathbb{E}\left(X\right)\right)\left(Y - \mathbb{E}\left(Y\right)\right)\right) = \mathbb{E}\left(XY\right) - \mathbb{E}\left(X\right)\mathbb{E}\left(Y\right).$$

על שונות משותפת ניתן לחשוב כך: אם המאורעות $X > \mathbb{E}(X)$, $Y > \mathbb{E}(Y)$ מתרחשים ביחד, כלומר על אותה קבוצה במרחב המדגם, וכך גם המאורעות $X < \mathbb{E}(X)$, $Y < \mathbb{E}(Y)$ מתרחשים ביחד, אזי $\mathrm{Cov}(X,Y)$ מקבל ערכים גדולים. אם לעומת זאת על אותה הקבוצה במרחב המדגם שבה $X > \mathbb{E}(X)$ מתקיים $Y < \mathbb{E}(Y)$, וכן להיפך, אזי $\mathrm{Cov}(X,Y)$ מקבלת ערכים שליליים גדולים (כלומר, גדולים בערך מוחלט).

תכונות:

1. לכל משתנה מקרי X מתקיים $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$.
2. סימטריות: $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.
3. בילינאריות: $\mathrm{Cov}(aX+bY,Z) = a\mathrm{Cov}(X,Z) + b\mathrm{Cov}(Y,Z)$, (וכך גם במשתנה השני).
4. אדישות להזזה בקבוע: $\mathrm{Cov}(X+a,Y) = \mathrm{Cov}(X,Y)$
5. אם X,Y מ״מ בעלי תוחלת סופית, אז

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y).$$

עבור $X_1,\ldots,X_n$ מ״מ בעלי תוחלת סופית, הנוסחה הכללית נראית כך:

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i,j=1}^{n} \mathrm{Cov}(X_i,X_j)\\ &= \sum_{i=1}^{n}\mathrm{Var}(X_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\mathrm{Cov}(X_i,X_j)\\ &= \sum_{i=1}^{n}\mathrm{Var}(X_i) + 2\sum_{1\le i<j\le n}\mathrm{Cov}(X_i,X_j).\end{aligned}$$

נבחין כי אם X,Y ב״ת אז $\mathrm{Cov}(X,Y)=0$ ולכן שונות סכום של מ״מ בלתי-תלויים בזוגות היא סכום השונויות. בפרט השיוויון

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\mathrm{Var}(X_i)$$

מתקיים כאשר המשתנים המקרים ב״ת.

דוגמה: מטילים שתי קוביות הוגנות. נחשב את השונות המשותפת בין סכומן לבין התוצאה של אחת מהן. נסמן $X,Y \sim \mathrm{Unif}\{1,...,6\}$, ונחשב את $\mathrm{Cov}(X+Y,X)$.

$$\mathrm{Cov}(X+Y,X) = \mathrm{Cov}(X,X) + \mathrm{Cov}(Y,X) = \mathrm{Var}(X) + 0 = \frac{35}{12}$$

תרגיל 3. (מספר המשולשים בגרף ארדש רני): יהי $G_{n,p}$ גרף ארדש רני על n קודקודים (גרף המקרי מתרגול קודם). כלומר גרף מקרי כך שהשתתפותה בגרף של כל קשת אפשרית בין שני קודקודים שונים ב- $[n]$ נקבעת בהגרלה נפרדת ומתרחשת בהסתברות p. נסמן את מספר המשולשים בגרף ב- $\triangle_n$. תרגול קודם חשבנו כי $\mathbb{E}\left(\triangle_n\right) = \binom{n}{3}p^3 = \frac{pn^3}{6}\left(1+o\left(1\right)\right)$. כעת נראה כי $\triangle_n$ מרוכז סביב תחולתו ביחס לתחולתו כאשר מספר הקודקודים בגרף שואף לאינסוף, קרי לכל $\varepsilon > 0$ מתקיים:

$$\lim_{n\to\infty}\mathbb{P}\left(\triangle_n - \mathbb{E}\left(\triangle_n\right) \geq \varepsilon\mathbb{E}\left(\triangle_n\right)\right) = 0$$

פתרון: תרגול קודם הגדרנו משתנה מקרי $X_{i,j,k} \sim Ber\left(p^3\right)$ שמקבל את הערך 1 אם ורק אם i, j, k מהווים משולש. מספר המשולשים בגרף $G_{n.p}$ מתואר על ידי המשתנה המקרי:

$$\triangle_n = \sum_{i,j,k} X_{i,j,k}$$

כאשר סכום זה הוא על פני כל השלשות של קודקודים שונים.

כעת, נחשב את השונות של $\triangle_n$ ונשים לב כי לפי נוסחת השונות לסכום נקבל:

$$\mathrm{Var}\left(\triangle_n\right) = \sum_{i,j,k\in[n],\ i',j',k'\in[n]} \mathrm{Cov}\left(X_{i,j,k}, X_{i',j',k'}\right)$$

נשים לב שמספר הקשתות המשותפות של שני משולשים נמצא בטווח של $\{0,1,3\}$ ולכן נוכל להפריד את הסכימה לפי מספר הקשתות המשותפות ולהבחין כי:

$$\begin{aligned}\mathrm{Var}\left(\triangle_n\right) &= \sum_{i,j,k\in[n]} \mathrm{Var}\left(X_{i,j,k}\right) + \sum_{i,j,k,l\in[n]} 6\cdot\mathrm{Cov}\left(X_{i,j,k}, X_{i,j,l}\right) + \sum_{i,j,k,i',j',k'\in[n]} \mathrm{Cov}\left(X_{i,j,k}, X_{i',j',k'}\right) \\ &= \binom{n}{3}p^3\left(1-p^3\right) + \sum_{i,j,k,l\in[n]} 6\cdot\mathrm{Cov}\left(X_{ij}X_{jk}X_{ki}, X_{ij}X_{jl}X_{li}\right) + 0 = \star\end{aligned}$$

ערך המחובר הראשון ידוע לנו כשונות של משתנה ברנולי ערך המחובר השלישי אפס כי המשתנים בלתי תלויים (כי לא חולקים קשתות משותפות יכול להיות שחולקים קודקוד משותף), הכפלנו את האבירם בסכום השני ב -6 כי עבור 4 קודקודים בגרף ניתן ליצור 6 אפוציות לשני משולשים עם חפיפה של קשת. נותר לחשב את המחובר השני.

נשים לב כעת, שאת האיברים בסכום השני ניתן לחשב לפי ההגדרה:

$$\text{Cov}\left(X_{ij}X_{jk}X_{ki},X_{ij}X_{jl}X_{li}\right)=\mathbb{E}\left(X_{ij}X_{jk}X_{ki}X_{jl}X_{li}\right)-\mathbb{E}\left(X_{ij}X_{jk}X_{ki}\right)\mathbb{E}\left(X_{ij}X_{jl}X_{li}\right)=p^{5}-p^{6}$$

כלומר נקבל כי:

$$\star=\binom{n}{3}p^{3}\left(1-p\right)^{3}+2\binom{n}{4}\cdot6\left(p^{5}-p^{6}\right)=O\left(n^{4}\right)$$

וכעת, כדי לקבל את התוצאה הנכספת, נשתמש באי שיוויון צ׳בישב עם $a=\varepsilon\mathbb{E}\left(\triangle_{n}\right)$ ונסיק כי:

$$\mathbb{P}\left(\triangle_{n}-\mathbb{E}\left(\triangle_{n}\right)\geq\varepsilon\mathbb{E}\left(\triangle_{n}\right)\right)\leq\frac{\text{Var}\left(\triangle_{n}\right)}{\left(\varepsilon\mathbb{E}\left(\triangle_{n}\right)\right)^{2}}=\frac{O\left(n^{4}\right)}{\varepsilon^{2}O\left(n^{6}\right)}$$

וכאשר נשאיף את n לאינסוף נקבל את הנדרש.

תרגיל 4. מסדרים את המספרים $[n]$ בשורה בסדר אקראי שנבחר באופן אחיד. יהי X משתנה מקרי שתוצאתו היא מספר האיברים $i\in[n]$ שנותרו במקומם.

- נחשב את השונות של X.

עבור $i\in[n]$, יהי $X_i\sim\text{Ber}(p)$ משתנה מקרי שתוצאתו היא 1 אם ורק אם המספר i נשאר במקום. נבחין כי מספר הסידורים הכולל הוא $n!$ ומספר הסידורים בהם i נשאר במקומות הוא $(n-1)!$, ולכן $p=(n-1)!/n!=1/n$ לכל i. מתקיים כי $X=\sum_{i=1}^{n}X_i$ ומלינאריות התוחלת

$$\mathbb{E}\left(X\right)=\sum_{i=1}^{n}\mathbb{E}\left(X_{i}\right)=\sum_{i=1}^{n}\frac{1}{n}=1.$$

כדי לחשב את השונות, עלינו להבחין שהמשתנים המקריים X_i תלויים ולכן השונות של X אינה סכום השונויות של X_i. למשל, מתקיים כי

$$\mathbb{P}\left(X_{1}=1,X_{2}=1\right)=\frac{(n-2)!}{n!}=\frac{1}{n\left(n-1\right)}\neq\frac{1}{n^{2}}=\mathbb{P}\left(X_{1}=1\right)\mathbb{P}\left(X_{2}=1\right).$$

לכן נשתמש בנוסחה הכללית לשונות:

$$.\mathrm{Var}\left(X\right)=\sum_{i=1}^{n}\mathrm{Var}\left(X_i\right)+2\sum_{i<j}\mathrm{Cov}\left(X_i,X_j\right)$$

נבחין כי $X_i=X_i^2$ לכל i, ולכן

$$\begin{aligned}\mathrm{Var}\left(X_i\right)&=\mathbb{E}\left(X_i^2\right)-\mathbb{E}\left(X_i\right)^2=\mathbb{E}\left(X_i\right)-\mathbb{E}\left(X_i\right)^2\\&=1/n-1/n^2=\left(n-1\right)/n^2\end{aligned}$$

נבחין עוד כי מהחישוב שערכנו לעיל, $X_iX_j\sim\mathrm{Ber}\left(\frac{1}{n(n-1)}\right)$ לכל $i<j$, ולכן

$$\begin{aligned}.\mathrm{Cov}\left(X_i,X_j\right)&=\mathbb{E}\left(X_iX_j\right)-\mathbb{E}\left(X_i\right)\mathbb{E}\left(X_j\right)\\&=\frac{1}{n\left(n-1\right)}-\frac{1}{n^2}=\frac{1}{n^2\left(n-1\right)}\end{aligned}$$

בסך הכל

$$\begin{aligned}\mathrm{Var}\left(X\right)&=\sum_{i=1}^{n}\frac{n-1}{n^2}+2\sum_{i<j}\frac{1}{n}\left(\frac{1}{n-1}-\frac{1}{n}\right)\\&=\frac{n-1}{n}+2\binom{n}{2}\frac{1}{n^2\left(n-1\right)}\\&=\frac{n-1}{n}+2\frac{n\left(n-1\right)}{2}\frac{1}{n^2\left(n-1\right)}=1\end{aligned}$$

– עבור $1\leq k\leq n$ כלשהו, נחסום את הסתברות שלפחות k מספרים נותרו במקומם. נשתמש תחילה באי-שוויון מרקוב:

$$.\mathbb{P}\left(X\geq k\right)\leq\mathbb{E}\left(X\right)/k=1/k$$

עבור $1<k$, ניעזר באי-שוויון צ׳בישב כדי לקבל חסם טוב יותר:

$$\begin{aligned}\mathbb{P}\left(X\geq k\right)&=\mathbb{P}\left(X-\mathbb{E}\left(X\right)\geq k-1\right)\\&\leq\mathbb{P}\left(\left|X-\mathbb{E}\left(X\right)\right|\geq k-1\right)\\&\leq\mathrm{Var}\left(X\right)/\left(k-1\right)^2=1/\left(k-1\right)^2\end{aligned}$$

וניכר שעבור $k \geq 3$ החסם באי-שויון צ'בישב טוב יותר.

הערה: מ״מ X, Y המקיימים $\mathrm{Cov}\left(X, Y\right) = 0$ מכונים "בלתי-מתואמים". שימו לב: אם X, Y בלתי-תלויים אז הם בלתי-מתואמים, אבל ההיפך אינו נכון.

4 החוק החלש של המספרים הגדולים

טענה: יהי X מ״מ בעל שונות סופית. יהיו $X_1, X_2, \ldots$ משתנים מקריים שווי התפלגות ל–X ובלתי-תלויים. אזי לכל $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}\left(X\right)\right| < \epsilon\right) = 1 .$$

במהלך ההוכחה של החוק החלש השתמשתם באי-שוויון צ׳בישב, ובהערכה הפשוטה

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}\left(X_i\right) = n\mathrm{Var}\left(X\right)$$

שמתבססת על כך שהמשתנים המקריים כולם בלתי-תלויים ושווי-התפלגות, שממנה נובע כי

$$\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\mathrm{Var}\left(X\right) \to 0,\ n\to\infty .$$

אבל למעשה ברור כי הדרישה שכל המשתנים המקריים יהיו שווי-התפלגות ובלתי-תלויים חזקה מידי. כל שעלינו לדרוש הוא

$$\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) \to 0,\ n\to\infty ,$$

ואם נוכל להעריך את הביטוי $\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right)$ באמצעות חישוב השונויות והשונויות המשותפות, לעתים נוכל לקבל את החוק החלש של המספרים הגדולים גם אם המשתנים המקריים תלויים ובעלי התפלגויות שונות.

מבוא להסתברות וסטטיסטיקה 80430
תרגול 10

1 פונקציה יוצרת מומנטים

הגדרה. הפונקציה יוצרת המומנטים של X היא $M_X(t) = \mathbb{E}\left(e^{tX}\right)$. משתנה מקרי בעל פונקציה יוצרת מומנטים המוגדרת בסביבה כלשהי של הראשית, מכונה **מ״מ בעל מומנט מעריכי.**

הערה. הפונקציה יוצרת המומנטים מוגדרת באמצעות תוחלת של פונקציה של משתנה מקרי. את החישוב הזה מתבצע בעזרת המשפט הבא:
יהיו X מ״מ ו-$g : \mathbb{R} \to \mathbb{R}$ פונקציה כך ש-$g(X)$ בעל תוחלת. אזי

$$\mathbb{E}\left(g\left(x\right)\right) = \sum_{x \in \mathrm{Supp}X} g\left(x\right) p_X\left(x\right).$$

כשאתם מתבקשים לחשב פונקציה יוצרת מומנטים של מ״מ כלשהו, תחום ההגדרה הוא חלק אינטגרלי מהתשובה! דהיינו, עליכם לומר עבור אלו ערכי t הפונקציה $M_X(t)$ מוגדרת. בדוגמאות לעיל תחום ההגדרה בכל מקום שלא הזכרנו אותו, היה כל $\mathbb{R}$. אין בעיה שבמבחן או תרגיל תאמרו פעם אחת בהתחלה שכשאתם לא אומרים כלום אז תחום ההגדרה הוא כל $\mathbb{R}$ - אבל זו לא קונבנציה של עולם המתמטיקה ועליכם להצהיר זאת במפורש, כדי שהבודק ידע שחשבתם על זה.

1.1 אי-שיוויון צ׳רנוף

כמו שאי״ש צ׳בישב נבע מאי״ש מרקוב ושיפר אותו בעזרת מידע נוסף (השונות) כך גם אי״ש צ׳רנוף נובע ממנו ומשפר אותו בעזרת מידע נוסף (פונקציה יוצרת מומנטים).

משפט. (אי-שוויון צ׳רנוף) יהי X מ״מ בעל מומנט מעריכי, אזי לכל $t > 0$ עבורו $M_X(t)$ סופי, ולכל $a \in \mathbb{R}$, מתקיים

$$\mathbb{P}\left(X \geq a\right) \leq M_X\left(t\right) e^{-ta}.$$

הערה. שימו לב שלכל a קיבלנו חסם מלעיל להסתברות הנ"ל עבור כל $t > 0$ (עבורו $M_X(t)$ סופי). אנו כמובן מעוניינים בחסם הטוב ביותר (הקטן ביותר) האפשרי שנותן צ'רנוף. דרך שימושית להשיג זאת, היא כאשר אגף ימין הוא פונקציה (של t) גזירה שיש לה מינימום.

תרגיל 1. יהי $X \sim \mathrm{Poi}(\lambda)$. ניזכר כי $\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$ ולכן מאי-שוויון צ'בישב מתקיים

$$\mathbb{P}(|X-\lambda| \geq \lambda) \leq \lambda/\lambda^2 = 1/\lambda.$$

נבחין כי היות ש-X מקבל רק ערכים אי-שליליים נובע כי $\{|X-\lambda| \geq \lambda\} = \{X=0\} \cup \{X \geq 2\lambda\}$. אם כך נקבל מאי-שוויון צ'רנוף כי לכל $t \in \mathbb{R}$,

$$\mathbb{P}(|X-\lambda| \geq \lambda) - \mathbb{P}(\{X=0\}) = \mathbb{P}(X \geq 2\lambda) \leq M_X(t)\, e^{-2\lambda t}.$$

נחשב:

$$\begin{aligned} M_X(t) = \mathbb{E}\left(e^{tX}\right) &= \sum_{n=0}^{\infty} e^{tn} \mathbb{P}(X=n) \\ &= \sum_{n=0}^{\infty} e^{tn} e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\left(e^t \lambda\right)^n}{n!} \\ &= e^{-\lambda} e^{e^t \lambda} = e^{\lambda\left(e^t - 1\right)} \end{aligned}$$

ונובע כי

$$\mathbb{P}(|X-\lambda| \geq \lambda) \leq e^{\lambda\left(e^t-1\right)-2\lambda t} + \mathbb{P}(\{X=0\}) = e^{\lambda\left(e^t-2t-1\right)} + e^{-\lambda}.$$

מה שקיבלנו הוא למעשה משפחה גדולה מאוד של חסמים לביטוי $\mathbb{P}(|X-\lambda| \geq \lambda)$. האם יש אחד מהם שהוא אופטימלי?

ברור שאם נמזער את הביטוי $e^{\lambda\left(e^t-2t-1\right)}$ נקבל את החסם הכי טוב. היות ש-$\lambda > 0$ ושהפונקציה האקספוננציאלית מונוטונית עולה, די למזער את $f(t) := e^t - 2t - 1$. נגזור ונשווה לאפס

$$f'(t) = e^t - 2 = 0 \Rightarrow t = \log 2$$

וקיבלנו כי החסם הכי טוב שמתקבל מאי-שוויון צ'רנוף הוא

$$\mathbb{P}(|X-\lambda| \geq \lambda) \leq e^{\lambda(2-2\log 2-1)} + e^{-\lambda} \approx e^{-0.39\cdot\lambda} + e^{-\lambda}.$$

1.2 אי-שוויון הופדינג

א״ש הופדינג הוא שוב א״ש המבוסס על א״ש מרקוב (דרך א״ש צ׳רנוף) והוספת מידע. המידע הנוסף עכשיו הוא אי-תלות:

משפט. יהיו $\{X_i\}_{i=1}^{n}$ מ״מ ב״ת בעלי תומך $\text{Supp}\left(X_i\right)\subseteq[-1,1]$ ו- $\mathbb{E}\left(X_i\right)=0$ לכל $i\in[N]$. אזי לכל $a>0$:

$$\mathbb{P}\left(\sum_{i=1}^{n}X_i\geq a\right)\leq\exp\left(-\frac{a^2}{2n}\right)$$

הערה. כדי להשתמש באי-שוויון הופדינג מספיק ש- $\{X_i\}_{i=1}^{n}$ מ״מ ב״ת וחסומים!

משפט. אי-שיויון הופדינג גרסא כללית: $\{X_i\}_{i=1}^{n}$ מ״מ ב״ת אשר מקיימים $M \overset{a.s.}{\leq} |X_k-\mathbb{E}(X_k)|$ לכל $k\in[n]$. נסמן $X=\sum_{k\in[n]}X_k$. אזי לכל $a>0$ מתקיים:

$$\mathbb{P}\left(X-\mathbb{E}\left(X\right)\geq a\right)\leq\exp\left(-\frac{a^2}{2nM^2}\right)$$
$$\mathbb{P}\left(\left|X-\mathbb{E}\left(X\right)\right|\geq a\right)\leq 2\exp\left(-\frac{a^2}{2nM^2}\right)$$

תרגיל 2. מטילים מטבע הוגן n פעמים, כאשר n אי-זוגי. יהי X מספר הפעמים שיצאו שני עצים ברצף (HH). השתמשו באי-שיוויון הופדינג כדי לקבל חסם טוב על

$$\mathbb{P}\left(\left|X-\frac{n-1}{4}\right|\geq a\right)$$

כפונקציה של a. הדרכה: אם X_i הוא האינדיקטור של המאורע ״ההטלה ה-i וגם ההטלה ה-$i+1$ יצאו עץ״, אז $X=\sum_{i=1}^{n-1}X_i$. לא ניתן לישם את אי-שיוויון הופדינג על X, אבל אפשר על $\sum_{i=1}^{(n-1)/2}X_{2i}$ ועל $\sum_{i=1}^{(n-1)/2}X_{2i-1}$ בנפרד.

נעזר בהדרכה. ראשית נשים לב שבעזרת אי-שוויון המשולש מתקיים

$$\left|X-\frac{n-1}{4}\right|\leq\left|\sum_{i=1}^{(n-1)/2}X_{2i}-\frac{n-1}{8}\right|+\left|\sum_{i=1}^{(n-1)/2}X_{2i-1}-\frac{n-1}{8}\right|$$

כדי שצד שמאל יהיה גדול מ- a, צריך שלפחות אחת הנסכמים בצד ימין יהיה גדול מ- $\frac{a}{2}$, לכן מתקיימת ההכלה הבאה

$$\left\{\left|X-\frac{n-1}{4}\right|\geq a\right\}\subseteq\left\{\left|\sum_{i=1}^{(n-1)/2}X_{2i}-\frac{n-1}{8}\right|\geq\frac{a}{2}\right\}\cup\left\{\left|\sum_{i=1}^{(n-1)/2}X_{2i-1}-\frac{n-1}{8}\right|\geq\frac{a}{2}\right\}$$

מחסם איחוד נקבל

$$\mathbb{P}\left(\left|X-\frac{n-1}{4}\right|\geq a\right)\leq\mathbb{P}\left(\left|\sum_{i=1}^{(n-1)/2}X_{2i}-\frac{n-1}{8}\right|\geq\frac{a}{2}\right)+\mathbb{P}\left(\left|\sum_{i=1}^{(n-1)/2}X_{2i-1}-\frac{n-1}{8}\right|\geq\frac{a}{2}\right)$$

נבחין כי $X_k\sim\text{Ber}\left(\frac{1}{4}\right)$ לכל k, לכן $\mathbb{E}\left(X_k\right)=\frac{1}{4}$. מכך נקבל

$$\mathbb{E}\left(\sum_{i=1}^{(n-1)/2}X_{2i-1}\right)=\sum_{i=1}^{(n-1)/2}\underbrace{\mathbb{E}\left(X_{2i-1}\right)}_{=\frac{1}{4}}=\frac{n-1}{2}\cdot\frac{1}{4}=\frac{n-1}{8}$$

ובאותו אופן

$$\mathbb{E}\left(\sum_{i=1}^{(n-1)/2}X_{2i}\right)=\frac{n-1}{8}$$

נשים לב ש- $\left\{X_{2i}\right\}_{i=1}^{\frac{n-1}{2}}$ הוא אוסף של משתנים מקריים בלתי תלויים (כי אין הטלות חופפות במשתנים המקריים). באותו אופן $\left\{X_{2i-1}\right\}_{i=1}^{\frac{n-1}{2}}$ בלתי תלויים. בנוסף נבחין כי $\left|X_k-\mathbb{E}\left(X_k\right)\right|\overset{a.s.}{\leq}1$ לכל k. לכן אפשר להפעיל את אי-שוויון הופדינג הגרסא הכללית ולקבל

$$\mathbb{P}\left(\left|\sum_{i=1}^{(n-1)/2}X_{2i}-\frac{n-1}{8}\right|\geq\frac{a}{2}\right)\leq 2\exp\left(\frac{-a^2}{4\left(n-1\right)}\right)$$

$$\mathbb{P}\left(\left|\sum_{i=1}^{(n-1)/2}X_{2i-1}-\frac{n-1}{8}\right|\geq\frac{a}{2}\right)\leq 2\exp\left(\frac{-a^2}{4\left(n-1\right)}\right)$$

ולכן

$$\mathbb{P}\left(\left|X-\frac{n-1}{4}\right|\geq a\right)\leq\mathbb{P}\left(\left|\sum_{i=1}^{(n-1)/2}X_{2i}-\frac{n-1}{8}\right|\geq\frac{a}{2}\right)+\mathbb{P}\left(\left|\sum_{i=1}^{(n-1)/2}X_{2i-1}-\frac{n-1}{8}\right|\geq\frac{a}{2}\right)$$
$$\leq 4\exp\left(\frac{-a^{2}}{4\left(n-1\right)}\right)$$

נבחין שעבור $a=\beta\sqrt{n-1}$ הדעיכה של ההסתברות היא כמו $e^{-\beta^{2}/4}$.

2 משתנים מקריים רציפים

הגדרה. יהא X מ״מ. נאמר ש-X הוא מ״מ **רציף בהחלט** אם קיימת פונקציה $f_X:\mathbb{R}\to[0,\infty)$ כך שלכל $-\infty\leq a\leq b\leq\infty$ מתקיים

$$\mathbb{P}\left(a\leq X\leq b\right)=\int_{a}^{b}f_{X}\left(t\right)dt$$

הפונקציה f_X נקראת **פונקציית הצפיפות** של המשתנה המקרי X.

טענה. יהי X משתנה מקרי רציף בהחלט. אז מתקיים:

1. $\mathbb{P}(X=a)=\int_{a}^{a}f_{X}\left(t\right)dt=0$ לכל $a\in\mathbb{R}$.
2. לכל $-\infty\leq a<b\leq\infty$ מתקיים

$$\mathbb{P}(a<X\leq b)=\mathbb{P}(a\leq X<b)=\mathbb{P}(a<X<b)=\int_{a}^{b}f_{X}\left(t\right)dt$$

3. סך ההסתברות היא 1: מתקיים $\int_{-\infty}^{\infty}f_{X}\left(t\right)dt=1$.
4. תהי $F\left(t\right)=\mathbb{P}\left(X\leq t\right)$ הפונקציה המצטברת של X. אזי מתקיים

$$F'\left(t\right)=f_{X}\left(t\right).$$

תרגיל 3. נתון X משתנה מקרי בעל פונקציית צפיפות מהצורה הבאה:

$$f_X(t) = \begin{cases} 2C\left(2t - t^2\right) & t \in [0,2] \\ 0 & else \end{cases}$$

כאשר C קבוע ממשי כלשהו.

1. חשבו את הערך של C.

מתכונה 3 של פונקציית הצפיפות נקבל את האילוץ הבא

$$1 = \int_{-\infty}^{\infty} f_X(t)dt = \int_0^2 2C(2t - t^2)dt$$

ולכן נסיק כי

$$1 = C\frac{8}{3} \Rightarrow C = \frac{3}{8}.$$

2. חשבו את הפונקציה המצטברת $F(t) = \mathbb{P}(X \leq t)$ עבור $t \in \mathbb{R}$.

לאחר שחישבנו את פונקציית הצפיפות נקבל בעזרת תכונות 3 ו-4

$$\mathbb{P}(X \leq t) = \begin{cases} 0 & t < 0 \\ \int_0^t \frac{3}{4}(2x - x^2)dx & 0 \leq t \leq 2 \\ 1 & t > 2 \end{cases} = \begin{cases} 0 & t < 0 \\ \frac{3}{4}(t^2 - \frac{t^3}{3}) & 0 \leq t \leq 2 \\ 1 & t > 2 \end{cases}$$

הגדרה. יהא X משתנה מקרי רציף. התוחלת של X מוגדרת ע"י הביטוי:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t f_X(t)\,dt$$

הגדרה. נאמר שמשתנה מקרי רציף בהחלט X מתפלג **אחיד** בקטע $[a,b]$, ונסמן $X \sim U([a,b])$, אם פונקציית הצפיפות שלו מקיימת

$$f_X(t) = \begin{cases} \frac{1}{b-a} & t \in [a,b] \\ 0 & t \notin [a,b] \end{cases}$$

במקרה זה פונקציית ההתפלגות המצטברת של X תקיים

$$F_X\left(t\right)=\begin{cases}\frac{t-a}{b-a} & t\in\left[a,b\right]\\ 0 & t<a\\ 1 & t>b\end{cases}$$

עבור התוחלת, נשים לב שלפי ההגדרה מתקיים:

$$\mathbb{E}(X)=\int_{-\infty}^{\infty}tf_X(t)dt=\int_a^b\frac{t}{b-a}dt=\frac{1}{b-a}\int_a^b t=\frac{b^2-a^2}{2}\cdot\frac{1}{b-a}=\frac{a+b}{2}$$

תרגיל 4. קו 68 מגיע לתחנה כל 15 דקות החל משעה עגולה (כולל). גברת עם סלים מגיעה לתחנה בין 7:00 ל-7:30 כאשר זמן ההגעה שלה מתפלג באופן אחיד. מהי ההסתברות שהגברת תחכה פחות מ-5 דקות?

יהא X המשתנה המקרי שמתאר את זמן ההגעה של הגברת בדקות אחרי 7:00. אז $X\sim U([0,30])$. יהא Y המשתנה המקרי שמתאר את זמן ההמתנה שלה. אז

$$\begin{aligned}\mathbb{P}\left(Y\le 5\right)&=\mathbb{P}\left(\{X=0\}\bigcup\{10\le X\le 15\}\bigcup\{25\le X\le 30\}\right)\\&=\mathbb{P}\left(X=0\right)+\mathbb{P}\left(10\le X\le 15\right)+\mathbb{P}\left(25\le X\le 30\right)\\&=0+\int_{10}^{15}\frac{1}{30}dx+\int_{25}^{30}\frac{1}{30}dx=\frac{10}{30}=\frac{1}{3}\end{aligned}$$

2.1 פונקציה של משתנה מקרי רציף

בחלק זה של התרגול נדון בבעיה הבאה: נתון משתנה מקרי רציף בהחלט X בעל התפלגות ידועה ונתונה פונקציה $g:\mathbb{R}\to\mathbb{R}$. במקרים רבים המשתנה המקרי $g\left(X\right)$ הוא רציף בהחלט ונוכל למצוא את הצפיפות שלו על ידי כך שנחשב את פונקציית ההתפלגות המצטברת $F_{g(X)}$ ונגזור אותה.

תרגיל 5. יהא $X\sim U([0,1])$ ויהא $n\in\mathbb{N}$. מהי ההתפלגות של $Y=X^n$?

ראשית, מתקיים ש $\mathbb{P}\left(X\in[0,1]\right)=1$ ולכן גם $\mathbb{P}\left(Y\in[0,1]\right)=1$. כעת, לכל $t\in[0,1]$

$$F_Y(t)=\mathbb{P}(Y\le t)=\mathbb{P}(X^n\le t)=\mathbb{P}(X\le t^{\frac{1}{n}})=F_X(t^{\frac{1}{n}})$$

את F_X אנחנו כבר יודעים ולכן נקבל

$$F_Y(t) = \begin{cases} 0 & t < 0 \\ t^{\frac{1}{n}} & 0 \leq t \leq 1 \\ 1 & t > 1 \end{cases}$$

לבסוף, נגזור את $F_Y(t)$ ונקבל את פונקציית הצפיפות של Y:

$$f_Y(t) = F'_Y(t) = \begin{cases} \frac{t^{\frac{1}{n}-1}}{n} & 0 < t < 1 \\ 0 & else \end{cases}$$

מבוא להסתברות וסטטיסטיקה 80430
תרגול 11

1 התפלגות משותפת

הגדרה יהיו X, Y זוג משתנים מקריים על אותו מרחב הסתברות. **פונקציית ההתפלגות המצטברת המשותפת** של (X, Y) היא $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ שמוגדרת ע״י

$$F_{X,Y}(s,t) = \mathbb{P}(X \leq s, Y \leq t)$$

הגדרה יהיו X, Y שני משתנים מקריים המוגדרים על אותו מרחב הסתברות. נאמר של-(X, Y) יש **צפיפות משותפת** אם קיימת פונקציה אינטגרבילית אי שלילית $f_{X,Y}(x,y) : \mathbb{R}^2 \to \mathbb{R}$, כך שלכל $t, s \in \mathbb{R}$ מתקיים

$$.F_{X,Y}(t,s) = \int_{-\infty}^{t} \int_{-\infty}^{s} f_{X,Y}(x,y)dydx$$

נסכם את התכונות של הצפיפות המשותפת בטענה הבאה:

טענה נניח כי ל-(X, Y) יש צפיפות משותפת. אזי:

1. לכל זוג מספרים $a < b$ ו-$c < d$, מתקיים

$$\mathbb{P}(a \leq X \leq b, c \leq Y \leq d) = \int_{a}^{b} \int_{c}^{d} f_{X,Y}(x,y)dydx$$

2. **סך ההסתברות היא 1:** מתקיים

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = 1$$

3. **רציפות וגזירות:** פונקציית ההתפלגות המצטברת המשותפת $F_{X,Y}$ היא רציפה. בנוסף, בכל נקודה $(s,t) \in \mathbb{R}^2$ שבה $f_{X,Y}$ רציפה מתקיים

$$.\frac{\partial^2}{\partial s \partial t} F_{X,Y}(s,t) = f_{X,Y}(s,t)$$

4. **צפיפות שולית:** X, Y הם משתנים מקריים רציפים בהחלט בעלי פונקציות צפיפות שולית

$$f_X\left(x\right) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$
$$f_Y\left(y\right) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$$

5. עבור X, Y מ״מ רציפים בהחלט בעלי תוחלת, לכל פונקציה $g : \mathbb{R}^2 \to \mathbb{R}$ מתקיים

$$\mathbb{E}\left(g\left(X,Y\right)\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g\left(x,y\right) f_{X,Y}\left(x,y\right) dydx$$

6. אם X, Y מ״מ רציפים בהחלט ב״ת אז פונקציית הצפיפות המשותפת של (X,Y) היא מכפלת פונקציות צפיפות השוליות של X, Y, כלומר

$$f_{X,Y}\left(x,y\right) = f_X\left(x\right) \cdot f_Y\left(y\right)$$

הגדרה: תהא $D \subset \mathbb{R}^2$ תת קבוצה בעלת שטח S. נאמר שלזוג מ״מ (X,Y) צפיפות משותפת **אחידה** ב- D אם מתקיים

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{S} & (x,y) \in D \\ 0 & else \end{cases}$$

1.1 המקרה הדו-מימדי

הגדרה: תהי $f:\mathbb{R}^2\to\mathbb{R}$ פונ' ויהא $D\subset\mathbb{R}^2$ תחום. **האינטגרל** של f על D, $\int\int_D f(x,y)\,dxdy$, הוא הנפח שכלוא בין הציר $z=0$ לגרף של f בתחום $(x,y)\in D$.

משפט: (פוביני) יהי D תחום מהצורה

$$D=\{(x,y): a\le x\le b, g(x)\le y\le h(x)\}$$

כאשר $h(x), g(x)$ פונקציות כלשהן. אזי מתקיים

$$\int\int_D f(x,y)\,dxdy=\int_a^b\left[\int_{g(x)}^{h(x)} f(x,y)\,dy\right]dx.$$

כלומר אפשר לחשב את האינטגרל הדו-מימדי על ידי חישוב של שני אינטגרלים חד-מימדיים, אחד אחרי השני.
המשפט נכון גם אם התחום הוא מהצורה

$$D=\{(x,y): a\le y\le b, g(y)\le x\le h(y)\},$$

אז מקבלים

$$\int\int_D f(x,y)\,dxdy=\int_a^b\left[\int_{g(y)}^{h(y)} f(x,y)\,dx\right]dy.$$

תרגיל 1. אישה מחכה לפקיד שיתפנה בבנק. בבנק שני פקידים, והזמן עד שכל אחד מהם יתפנה הוא משתנה מקרי שמתפלג אחיד על $[0,10]$, באופן ב״ת. מה תוחלת זמן ההמתנה שלה?

פתרון: אם נסמן ב X,Y את הזמנים עד שכל אחד מהפקידים מתפנה, הזמן שעליה לחכות הוא

$$Z=\min\{X,Y\}$$

נתון לנו ש - $X,Y\sim U([0,10])$ ב״ת, ולכן פונקציית הצפיפות המשותפת שלהם היא

$$f_{X,Y}(x,y)=f_X(x)f_Y(y)=\begin{cases}\frac{1}{100} & 0\le x,y\le 10\\ 0 & else\end{cases}$$

ע״פ הטענה מתקיים

$$\begin{aligned}\mathbb{E}(Z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\min\{x,y\}\, f_{X,Y}(x,y)\,dxdy \\ &= \frac{1}{100}\int_0^{10}\left(\int_0^x ydy + \int_x^{10} xdy\right)dx \\ &= \frac{1}{100}\int_0^{10}\left(\frac{x^2}{2} + x(10-x)\right)dx \\ &= \frac{1}{100}\int_0^{10}\left(10x - \frac{x^2}{2}\right)dx \\ &= \frac{1}{100}\left(5\cdot 10^2 - \frac{10^3}{6}\right) = 5 - \frac{5}{3} = 3\frac{1}{3}\end{aligned}$$

תרגיל 2. (לקריאה בלבד, לא לתרגול הפרונטלי) סינדרלה והנסיך קובעים להיפגש באקראי בין 22:00 ל 23:00. כדי לוודא שסינדרלה תוכל להימלט לפני חצות ולא תבזבז יותר מידי זמן, היא מסכימה עם הנסיך שאף אחד מהם לא ימתין לשני יותר מ 15 דק'. מה הסיכוי שהשניים יפגשו?

פתרון: יהיו X,Y משתנים מקריים שמתארים את זמן ההגעה של סינדרלה והנסיך ביחידות של שברי שעות. בפרט מהנתון: $X,Y\sim U([0,1])$ בלתי תלויים. לכן פונקציות הצפיפות של כל אחד מהם היא:

$$f_Y(t) = f_X(t) = \begin{cases}1 & 0\le t\le 1\\ 0 & else\end{cases}$$

ומשום שהמשתנים המקריים בלתי תלוים פונקציית הצפיפות המשותפת שלהם היא:

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y) = \begin{cases}1 & 0\le x,y\le 1\\ 0 & else\end{cases}$$

משום ש 15 דק הן $\frac{1}{4}$ שעה, ההסתברות שאנו רוצים לחשב היא:

$$\mathbb{P}\left(|X-Y|\le\frac{1}{4}\right) = \int\int_{|X-Y|\le\frac{1}{4}} dxdy$$

אילו נחשוב על זמני ההגעה של השניים כזוגות סדורים $(x,y)\in[0,1]\times[0,1]$, נשים לב שהאינטגרל שקיבלנו הוא שטח הריבוע בגודל 1×1, פחות שטח שני משולשים ישרי זווית עם שוקיים בגודל $\frac{3}{4}\times\frac{3}{4}$ שמייצגים זמני הגעה בהפרש גדול מ 15 דק', כלומר מהצורה (x,y) כאשר $|x-y|\geq\frac{1}{4}$.
לכן ההסתברות המבוקשת היא $\frac{7}{16}=\left(\frac{3}{4}\right)\left(\frac{3}{4}\right)-1$.

תרגיל 3. נניח של-X ו-Y התפלגות אחידה בעיגול ברדיוס 1 סביב $(0,0)$ (כלומר בעיגול היחידה), כלומר, פונקציית הצפיפות שלהם מקיימת

$$f_{X,Y}(x,y)=\begin{cases}\frac{1}{\pi} & x^2+y^2<1\\ 0 & else\end{cases}$$

חשבו את הצפיפות של X, ואת ההסתברות $P\left(X^2+Y^2\leq 1/2\right)$.
פתרון: הצפיפות של X נתונה על ידי

$$f_X(x)=\int_{-\infty}^{\infty}f_{X,Y}(x,y)dy=\int_{\{y:x^2+y^2\leq 1\}}f_{X,Y}(x,y)dy$$

אם $|x|\geq 1$ אז לכל $y\in\mathbb{R}$ מתקיים $x^2+y^2\geq 1$ ולכן

$$f_X(x)=\int_{-\infty}^{\infty}0dy=0$$

אם $|x|<1$ אז $x^2+y^2<1$ כאשר $-\sqrt{1-x^2}<y<\sqrt{1-x^2}$ ולכן

$$f_X(x)=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}dy=\frac{2\sqrt{1-x^2}}{\pi}$$

$$\mathbb{P}(X^2+Y^2\leq\frac{1}{2})=\frac{1}{\pi}\int\int_{\{(x,y)|x^2+y^2\leq 1/2\}}dxdy$$

האינטגרל שקיבלנו הוא שטח העיגול ברדיוס $\frac{1}{\sqrt{2}}$ סביב $(0,0)$. אנחנו יודעים ששטח עיגול זה הוא $\frac{\pi}{2}$, ולכן

$$\mathbb{P}(X^2+Y^2\leq\frac{1}{2})\leq\frac{1}{\pi}\cdot\frac{\pi}{2}=\frac{1}{2}.$$

תרגיל 4. יהא $X_1\sim U([0,1])$, $X_2=\mathrm{Ber}\left(1/3\right)$ שני משתנים מקריים בלתי תלויים. הראו ש-$Y=X_1+X_2$ משתנה מקרי רציף בהחלט וחשבו את צפיפותו. מהי הצפיפות של $Z=Y^3$?
פתרון: נשים לב שניתן פשוט לכתוב במפורש את $F_Y\left(t\right)=\mathbb{P}\left(Y\leq t\right)$: מכיוון ש-$X_1,X_2$ מקבלים ערכים בתחום

$[0,1]$, המשתנה המקרי Y נתמך בתחום $[0,2]$.

עבור $0 \le t < 1$: כדי שיתקיים $Y \le t$ חייב להתקיים $X_2 = 0, X_1 \le t$ ולכן בתחום $0 \le t < 1$ מתקיים מאי תלות

$$F_Y(t) = \mathbb{P}(Y \le t) = \mathbb{P}(X_2 = 0)\,\mathbb{P}(X_1 \le t) = \frac{2}{3}t.$$

עבור $1 \le t \le 2$: כדי שיתקיים $Y \le t$ חייב להתקיים $X_2 = 0$ או $X_2 = 1$ וגם $X_1 \le t - 1$. בסה"כ נקבל בתחום $1 \le t \le 2$ מאי תלות:

$$F_Y(t) = \mathbb{P}(Y \le t) = \mathbb{P}(X_2 = 0) + \mathbb{P}(X_2 = 1)\,\mathbb{P}(X_1 \le t - 1) = \frac{2}{3} + \frac{1}{3}(t-1) = \frac{t+1}{3}$$

לכן

$$F_Y(t) = \begin{cases} 0 & t \le 0 \\ \frac{2}{3}t & 0 \le t \le 1 \\ \frac{t+1}{3} & 1 \le t \le 2 \\ 1 & t > 2 \end{cases}$$

וזו פונקציה רציפה וגזירה בכל $\mathbb{R}$ מלבד ב-$\{0,1,2\}$ ולכן מגדירה משתנה מקרי רציף בהחלט. הצפיפות היא הנגזרת שלה בכל נקודות הגזירות:

$$f_Y(t) = \frac{d}{dt}F_Y(t) = \begin{cases} 0 & t \le 0 \\ \frac{2}{3} & 0 < t \le 1 \\ \frac{1}{3} & 1 < t \le 2 \\ 0 & t > 2 \end{cases}.$$

כדי לחשב את הצפיפות של Y^3 שמים לב כי

$$\begin{aligned} F_{Y^3}(t) &= \mathbb{P}\left(Y^3 \le t\right) = \mathbb{P}\left(Y \le t^{1/3}\right) \\ &= F_Y\left(t^{1/3}\right) \end{aligned}$$

לכן

$$F_{Y^3}(t)=\begin{cases}0 & t\le 0\\ \frac{2}{3}t^{1/3} & 0\le t\le 1\\ \frac{t^{1/3}+1}{3} & 1\le t\le 8\\ 1 & t>8\end{cases}$$

ומגזירה מקבלים

$$.f_{Y^3}(t)=\frac{d}{dt}F_{Y^3}(t)=\begin{cases}0 & t\le 0\\ \frac{2}{9}t^{-2/3} & 0<t\le 1\\ \frac{t^{-2/3}}{9} & 1<t\le 2\\ 0 & t>2\end{cases}$$

שימו לב כי f_{Y^3} אינטגרבילית רק במובן הלא אמיתי, מכיוון שהיא אינה חסומה בסביבת 0.

תרגיל 5. בתחרות חץ וקשת בין 3 אנשים, המתחרים יורים חץ אל עבר לוח מטרה אינסופי. ההצלחה של קשת נמדדת לפי מרחק הפגיעה מראשית הצירים. נתון שקואורדינטות X,Y של הנקודה בה יפגע רובין הוד הן בלתי-תלויות, וכן שיש לכל אחת מהן התפלגות נורמלית תקנית $X,Y\sim N(1,0)$, כלומר $f_X(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ וכדומה עבור Y.

1. מתחרים $1,2$ פגעו במטרה במרחקים $r_1<r_2$ מהראשית. מה ההסתברות שרובין הוד יסיים במקום השני?
2. מה ההסתברות שנקודת הפגיעה תהיה בגזרה האינסופית שבין הקרניים המוגדרות על ידי הזוויות θ_1,θ_2 (נניח לשם הנוחות $0<\theta_1<\theta_2<2\pi$)?
3. מה תוחלת המרחק מראשית הצירים של החץ של רובין הוד?
4. נגדיר $Z=X^2+Y^2$. איך Z מתפלג? מצאו את f_Z וזהו את ההתפלגות – היא מוכרת לנו היטב.

__פתרון__:

1. הוא יסיים במקום השני אם יפגע ברדיוס $r_1<R<r_2$. ההסתברות לכך היא:

$$\begin{aligned}\mathbb{P}(r_1<R<r_2)&=\mathbb{P}\left(r_1^2<X^2+Y^2<r_2^2\right)\\&=\iint_D f_{XY}(x,y)dxdy\end{aligned}$$

כאשר $D=\left\{\binom{x}{y}|r_1^2<x^2+y^2<r_2^2\right\}$. אך מהי פונקציית ההתפלגות המשותפת? אמנם לרוב לא ניתן למצוא את $f_{XY}(x,y)$ בעזרת $f_X(x)$ ו- $f_Y(y)$, אולם נתון לנו כי X,Y ב"ת. לכן:

$$f_{XY}(x,y)=f_X(x)f_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$$

נשים לב לכך שההתפלגות תלויה רק באחת מבין שתי הקואורדינטות הפולאריות (ב- r ולא ב- θ). זה מפשט לנו מאוד את האינטגרל:

$$\begin{aligned}\mathbb{P}\left(r_1<R<r_2\right)&=\frac{1}{2\pi}\iint_D e^{-\frac{x^2+y^2}{2}}dxdy\\&=\frac{1}{2\pi}\int_0^{2\pi}\int_{r_1}^{r_2}e^{-\frac{\rho^2}{2}}\,d\rho d\theta\\&=\int_{r_1}^{r_2}e^{-\frac{\rho^2}{2}}\rho\,d\rho\\&=e^{-\frac{r_1^2}{2}}-e^{-\frac{r_2^2}{2}}\end{aligned}$$

2. הרעיון דומה, רק צריך להחליף את תחום האינטגרציה. הפעם, בזכות האי-תלות ב- θ האינטגרציה עוד יותר פשוטה:

$$\begin{aligned}\mathbb{P}(\theta_1<\theta<\theta_2)&=\iint_{\theta\in(\theta_1,\theta_2)}f_{XY}(x,y)dxdy\\&=\frac{1}{2\pi}\int_{\theta_1}^{\theta_2}\int_0^{\infty}e^{-\frac{r^2}{2}}rdrd\theta\\&=\frac{1}{2\pi}\int_{\theta_1}^{\theta_2}d\theta\\&=\frac{\theta_2-\theta_1}{2\pi}\end{aligned}$$

במובן מסוים התחום $\theta_1<\theta<\theta_2$ מכיל $\frac{\theta_2-\theta_1}{2\pi}$ מתוך המישור, ובשל הסימטריה לסיבובים יש ל- θ התפלגות אחידה של על פני $[0,2\pi]$.

3. צריך למצוא את $\mathbb{E}\left(R\right)$.

$$\begin{aligned}\mathbb{E}\left(R\right) &= \iint_D R(x,y)\cdot f_{XY}(x,y)dxdy\\ &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty}\rho^2 e^{-\frac{\rho^2}{2}}\,d\rho d\theta\\ &= \int_0^{\infty}\rho^2 e^{-\frac{\rho^2}{2}}\,d\rho\\ &= \sqrt{\frac{\pi}{2}}\end{aligned}$$

כאשר את השלב האחרון עושים בעזרת אינטגרציה בחלקים.

הערה: ההתפלגות של R בבעיה זו נקראת התפלגות $Rayleigh$ או התפלגות $Maxwell-Boltzmann$ הדו-מימדית. התפלגות $Maxwell-Boltzmann$ התלת-מימדית (כאשר $R^2=X^2+Y^2+Z^2$) מאוד חשובה בפיזיקה, שכן היא מתארת את התפלגות המהירויות של חלקיקים בגז[1] - המהירות בכל ציר מתפלגת נורמלית ולכן המהירות הכוללת מתפלגת על פי התפלגות זו.

4. נשים לב לכך ש- $Z=R^2$ כאשר $R=\sqrt{X^2+Y^2}$. עבור $r>0$ מתקיים

$$F_R\left(r\right)=\frac{1}{2\pi}\int_0^{2\pi}\int_0^{r}e^{-\frac{\rho^2}{2}}\rho\,d\rho d\theta=\int_0^{r}e^{-\frac{\rho^2}{2}}\rho\,d\rho=\left[-e^{-\frac{\rho^2}{2}}\right]_0^r=1-e^{-\frac{r^2}{2}}$$

על כן

$$F_Z\left(z\right)=\mathbb{P}\left(Z\leq z\right)=\mathbb{P}\left(R\leq\sqrt{z}\right)=1-e^{-\frac{z}{2}}$$

זוהי פונקציית ההתפלגות המצטברת של מ"מ מעריכי עם פרמטר $\lambda=\frac{1}{2}$. בהתאם:

$$f_Z\left(z\right)=\frac{d}{dz}F_Z\left(z\right)=\frac{1}{2}e^{-\frac{z}{2}}$$

[1] בהזנחת אפקטים יחסותיים

מבוא להסתברות וסטטיסטיקה 80430
תרגול 12 – התכנסות בהתפלגות, משפט הגבול המרכזי וסטטיסטיקה

1 התכנסות בהתפלגות

הגדרה: יהיו X , $(X_n)_{n=1}^{\infty}$ משתנים מקריים, לאו דווקא על אותו מרחב הסתברות. נאמר כי הסדרה $(X_n)_{n=1}^{\infty}$ מתכנסת ל-X **בהתפלגות**, ונסמן

$$X_n \xrightarrow{d} X$$

אם לכל $a \in \mathbb{R}$ שהיא **נקודת רציפות** של F_X מתקיים

$$\lim_{n\to\infty} F_{X_n}(a) = F_X(a)$$

1.1 דוגמא:

תהי $(X_n)_{n=1}^{\infty}$ סדרה של משתנים מקריים כך ש-

$$\forall n \in \mathbb{N}, X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$$

עבור $\lambda > 0$ כלשהו. נרצה להראות ש- X_n מתכנס בהתפלגות ל $X \sim Pois(\lambda)$.

פתרון: מספיק להראות שלכל $k \in \mathbb{N}$ מתקיים:

$$\lim_{n\to\infty} \mathbb{P}_{X_n}(k) = \mathbb{P}_X(k)$$

בפרט,

$$\lim_{n\to\infty} \mathbb{P}_{X_n}(k) = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$\lambda^k \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{1}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$\frac{\lambda^k}{k!} \lim_{n\to\infty} \left(\left[\frac{n(n-1)(n-1)...(n-k+1)}{n^k}\right] \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}\right) =$$

$$\frac{\lambda^k e^{-\lambda}}{k!} = \mathbb{P}_X(k)$$

$\square$

2 משפט הגבול המרכזי

משפט: תהי $(X_n : n \geq 1)$ סדרת משתנים מקריים בלתי-תלויים ושווי-התפלגות, המקיימים $\mathbb{E}(X_n) = 0$, $\mathrm{Var}(X_n) = 1$. אזי

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{\mathrm{d}} Z ,$$

כאשר $Z \sim \mathcal{N}(0,1)$ הוא מ"מ נורמלי סטנדרטי בעל הצפיפות $f_Z(z) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}$. באופן מפורש, היות שפונקציית ההתפלגות המצטברת של Z רציפה בכל מקום ב–$\mathbb{R}$, המשפט קובע כי לכל $z \in \mathbb{R}$,

$$\mathbb{P}\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}} \leq z\right) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-x^2/2} dx$$

נהוג לסמן את פונקציית ההתפלגות המצטברת של Z על ידי

$$\Phi(z) = F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-x^2/2} dx.$$

מסקנה: תהי $(X_n : n \geq 1)$ סדרת משתנים מקריים בלתי-תלויים ושווי-התפלגות ובעלי שונות סופית. נסמן $\mu = \mathbb{E}(X_n)$, $\sigma^2 = \mathrm{Var}(X_n)$. נגדיר $Y_n = \frac{X_n - \mu}{\sigma}$ ונבחין כי $\mathbb{E}(Y_n) = 0$ וכי $\mathrm{Var}(Y_n) = 1$. לכן ממשפט הגבול המרכזי לסדרה $(Y_n : n \geq 1)$ נקבל

$$\mathbb{P}\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq z\right) = \mathbb{P}\left(\frac{Y_1 + \cdots + Y_n}{\sqrt{n}} \leq z\right) \xrightarrow[n\to\infty]{} \Phi(z).$$

דוגמה: מטילים n קוביות הוגנות ובלתי-תלויות. העריכו את ההסתברות שסכום ההטלות הוא בטווח $[3.5n - \sqrt{n}, 3.5n + \sqrt{n}]$ באמצעות משפט הגבול המרכזי. עבור $1 \leq i \leq n$ יהי $X_i \sim \mathrm{Unif}\{1, \ldots, 6\}$, כולם בלתי-תלויים. אזי $\mu = \mathbb{E}(X_i) = 3.5$, $\sigma^2 = \mathrm{Var}(X_i) = \frac{6^2-1}{12} = \frac{35}{12}$. לכן מהמסקנה הקודמת,

$$\begin{aligned}
\mathbb{P}\left(3.5n - \sqrt{n} \leq \sum_{i=1}^{n} X_i \leq 3.5n + \sqrt{n}\right) &= \mathbb{P}\left(-1 \leq \frac{X_1 + \cdots + X_n - 3.5n}{\sqrt{n}} \leq 1\right) \\
&= \mathbb{P}\left(-\sqrt{12/35} \leq \frac{X_1 + \cdots + X_n - 3.5n}{\sqrt{n}\sqrt{35/12}} \leq \sqrt{12/35}\right) \\
&\approx \Phi\left(\sqrt{12/35}\right) - \Phi\left(-\sqrt{12/35}\right) \\
&= 2\Phi\left(\sqrt{12/35}\right) - 1 \\
&\approx 2\Phi(0.58) - 1 \approx 2 \cdot 0.72 - 1 = 0.44
\end{aligned}$$

למשל אם נציב $n = 100$ נקבל כי הסיכוי שסכום ההטלות הוא בטווח $[340, 360]$ הוא בערך 0.44.

הערה. מתקיימת הזהות $\Phi(a) = 1 - \Phi(-a)$.

3 סטטיסטיקה

3.1 הפרדת השערות פשוטות

הגדרה: תהי $\mathcal{D}$ התפלגות כלשהי שאיננה ידועה. **דגימות בלתי-תלויות** מהתפלגות $\mathcal{D}$ היא סדרת ערכים $(x_1, \ldots, x_n)$ המתקבלת מסדרת משתנים-מקריים בלתי-תלויים $(X_1, \ldots, X_n)$ המקיימים $X_i \sim \mathcal{D}$ לכל $1 \le i \le n$.

למשל, איננו יודעים כיצד מתפלג הגובה של תושבי ירושלים, אבל אנחנו מסוגלים ללכת ברחוב ולמדוד את גובהם של n אנשים אקראיים שהגבהים שלהם בלתי-תלויים. עלינו להיזהר כשמבצעים דגימה כזאת מפני שני סוגים של טעויות:

1. יתכן שההתפלגות לא מתאימה: למשל, לא נרצה לבחור n אנשים מתוך משתתפי תחרות "האיש הגבוה ביותר", כי התפלגות הגבהים אינה זו שבאוכלוסייה.
2. ייתכן שהמ"מ שבדגימה תלויים: למשל, לא נרצה לבחור n אנשים בני אותה משפחה (אפילו אם המשפחה נבחרת באקראי), כי יש תלות בין הגבהים של בני אותה המשפחה.

הגדרה: נניח כי ההתפלגות $\mathcal{D}$ היא אחת משתי התפלגויות ידועות $\mathcal{D}_0, \mathcal{D}_1$, אבל איננו יודעים איזו מהן. בפנינו שתי השערות: נסמן ב-H_0 את ההשערה כי $\mathcal{D} = \mathcal{D}_0$ ונסמן ב-H_1 את ההשערה כי $\mathcal{D} = \mathcal{D}_1$.

כדי להכריע בין שתי ההשערות עלינו לקבוע **מבחן**: כלומר, עלינו לקבוע קבוצה $S \subset \mathbb{R}^n$ ולפעול לפי הכלל הבא: נדגום $(X_1, \ldots, X_n)$ בלתי תלויים מההתפלגות הלא ידועה, אם $(X_1, \ldots, X_n) \in S$ נחליט לדחות את H_0 (כלומר $\mathcal{D} = \mathcal{D}_1$). אחרת נחליט לקבל את ההשערה H_0 (כלומר $\mathcal{D} = \mathcal{D}_0$). דרך נוספת היא להגדיר את המבחן בעזרת אינדיקטור $\mathbf{1}_S$.

כמובן שמבחן יכול להיקבע באופן שרירותי, ועלינו להגדיר פרמטרים לפיהם נוכל להשוות בין שני מבחנים כדי להעדיף אחד על פני אחר:

הגדרה: ישנם שני סוגים של מסקנות שגויות עבור מבחן S להתפלגות לא ידועה $\mathcal{D}$:

- **טעות מסוג ראשון:** מתקיים $(X_1, \ldots, X_n) \in S$ למרות ש-$\mathcal{D} = \mathcal{D}_0$. נסמן את ההסתברות לטעות כזאת על ידי

$$\alpha := \mathbb{P}_{H_0}\left((X_1, \ldots, X_n) \in S\right).$$

הסתברות זו מכונה **רמת המובהקות** של המבחן. בהתאם, **רמת הסמך** (p-value) של מבחן היא $1 - \alpha$ (ההסתברות לאמץ את H_0 אם היא נכונה).

- **טעות מסוג שני:** מתקיים $(X_1, \ldots, X_n) \in S^c$ למרות ש-$\mathcal{D} = \mathcal{D}_1$. נסמן את ההסתברות לטעות כזאת על ידי

$$\beta := \mathbb{P}_{H_1}\left((X_1, \ldots, X_n) \in S^c\right).$$

- בהתאם, **העוצמה** (power) של מבחן היא $1 - \beta$ (ההסתברות לאמץ את H_1 אם היא נכונה).

היינו שמחים למזער הן את α הן את β במקביל, אולם לעתים כדי להקטין את אחד מהם ניאלץ להתפשר על השני.

הערות חשובות:

1. חשוב להבחין כי הסימון $\mathbb{P}_{H_0}$ ו-$\mathbb{P}_{H_1}$ מתארים פונקציות הסתברות שונות. הרי אם H_0 נכונה ההתפלגות האמתית היא $\mathcal{D}_0$ ואם H_1 נכונה ההתפלגות האמתית היא $\mathcal{D}_1$, ולכן כל אחד מחישובי ההסתברות הללו נעשה מעל מרחב הסתברות שונה. כך למשל, לחלוטין <u>לא נכון</u> ש-$\alpha + \beta = 1$.

2. חשוב להבחין שאמנם טעות מסוג ראשון וטעות מסוג שני הן סימטריות מבחינה מתמטית, היות שאין הבדל בין ההשערה H_0 להשערה H_1, אולם כאן אנחנו עוסקים בסטטיסטיקה ובהקשר זה נהוג לייחס להן משמעות שונה: ההשערה H_0 מכונה "השערת האפס" והיא ההשערה השמרנית, בעוד שההשערה H_1 מכונה "השערה אלטרנטיבית" או "אלטרנטיבה" והיא ההשערה הנועזת מבין השתיים. להבחנה זו אין חשיבות מתמטית אולם יש לה חשיבות סטטיסטית כפי שאפשר לראות בדוגמה הבאה:

דוגמה: פותח חיסון חדש למחלת השפעת בעונה חורפית נתונה. נתון כי לפני החיסון כל אדם מאושפז עקב המחלה בהסתברות של 0.01 באופן בלתי-תלוי באחרים. לטענת החברה שפיתחה את החיסון כל אדם שמתחסן מאושפז עקב המחלה בהסתברות של 0.004. רופא לשכת הבריאות המחוזית אינו יודע איזו מהטענות נכונה, אז הוא מסמן ב-$\mathcal{D}$ את ההתפלגות האמתית שאיננה ידועה ומגדיר : H_0 $\mathcal{D} = \text{Bin}\left(n, 0.01\right)$ (השערת האפס השמרנית היא שהחיסון לא משפיע כל עוד אין לנו הוכחה לכך), $H_1 : \mathcal{D} = \text{Bin}\left(n, 0.004\right)$ (האלטרנטיבה היא שהחברה המפתחת צודקת בטענתה). בוחרים באקראי $n = 1000$ אנשים, מחסנים אותם ובסוף החורף בודקים כמה מהם אושפזו. נשתמש בקירוב הפואסוני, לפיו

$$\text{Bin}\left(1000, 0.01\right) \approx \text{Poiss}\left(1000 \cdot 0.01\right) = \text{Poiss}\left(10\right)$$

$$.\text{Bin}\left(1000, 0.004\right) \approx \text{Poiss}\left(1000 \cdot 0.004\right) = \text{Poiss}\left(4\right)$$

נסמן את הדגימה המתאימה לפי מ"מ $\left(X_1, \ldots, X_{1000}\right)$ שכל אחד מהם מתפלג ברנולי עם פרמטר לא ידוע, ויהי $X = \sum_{i=1}^{1000} X_i$ מ"מ שסופר את המאושפזים בקרב 1000 האנשים ושהתפלגותו נקבעת על ידי H_0, H_1 והיא בקירוב פואסונית. כדי להכריע בין ההשערות נציע את המבחן הבא: $S = \left\{\left(x_1, \ldots, x_{1000}\right) \mid \sum_{i=1}^{1000} x_i \leq 5\right\}$. הסיבה לבחירת מבחן זה היא שהתוחלת של מ"מ פואסוני היא הפרמטר "λ" המתאים לו. אנחנו מרשים לחברת התרופות לשגות ב-1 כלפי מעלה, כלומר שהשגיאה מהתוחלת שמיוחסת לפי החברה (4) תהיה לכל היותר 1, אבל אם השגיאה גדולה מ-1, כלומר היו לפחות 6 מאושפזים, אנחנו לא מקבלים את הטענה של חברת התרופות. כמובן שיכולנו לבחור הרבה מבחנים אחרים.

נחשב טעות מסוג ראשון ושני:

$$\begin{aligned}
\alpha_T &= \mathbb{P}_{H_0}\left(\left(X_1, \ldots, X_n\right) \in S\right) \\
&= \mathbb{P}_{H_0}\left(\sum_{i=1}^{1000} X_i \leq 5\right) \\
&= \mathrm{e}^{-10}\left(1 + 10 + \frac{10^2}{2} + \frac{10^3}{3!} + \frac{10^4}{4!} + \frac{10^5}{5!}\right) \approx 0.067
\end{aligned}$$

$$
\begin{aligned}
\beta_T &= \mathbb{P}_{H_1}\left((X_1,\ldots,X_n)\notin S\right)\\
&= \mathbb{P}_{H_1}\left(\sum_{i=1}^{1000} X_i \geq 6\right)\\
&= 1-\mathbb{P}_{H_1}\left(\sum_{i=1}^{1000} X_i \leq 5\right)\\
&= 1-\mathrm{e}^{-4}\left(1+4+\frac{4^2}{2}+\frac{4^3}{3!}+\frac{4^4}{4!}+\frac{4^5}{5!}\right)\approx 0.21
\end{aligned}
$$

במובן מסוים, ההסתברויות הללו לטעויות משקפות את האופן שבו אנחנו מעוניינים לקבל תרופות חדשות: אנחנו מציבים רף גבוה של הוכחה כדי לשווק לציבור חיסון חדש, תמורת סיכון כלכלי גדול על חברת התרופות. מצד אחד, יש הסתברות של 6.7% בלבד לטעות שעלולה לסכן את הציבור (להחליט שהחיסון עובד למרות שהוא לא עובד), בעוד שיש הסתברות גבוהה יותר של 21% בלבד לטעות שתפגע כלכלית בעיקר בחברת התרופות (להחליט שהחיסון לא עובד למרות שהוא עובד). כמובן שגם לטעות השנייה עלול להיות מחיר ציבורי, ובעניין זה אמורים לקבוע קובעי המדיניות ולא הסטטיסטיקאים.

3.2 השוואת מבחנים והלמה של ניימן-פירסון

הגדרה: יהיו H_0, H_1 השערות פשוטות ויהיו C, C' זוג מבחנים. נאמר כי מבחן C הוא **טוב** מהמבחן C', אם גם $\alpha_C \leq \alpha_{C'}$ וגם $\beta_C \leq \beta_{C'}$. נאמר כי המבחן C **טוב ממש** מהמבחן C', אם לפחות אחד משני אי השוויונים הללו הוא אי-שוויון חזק. מבחן נקרא **מיטבי** אם אין אף מבחן אחר שטוב ממש ממנו.

הגדרה: תהי $\mathcal{D}$ התפלגות בדידה או רציפה בהחלט כלשהי (ידועה) ותהי $(X_1,\ldots,X_n)$ דגימה ממנה. נגדיר את **פונקציית הנראות** במקרה <u>הבדיד</u> על ידי

$$\mathcal{L}_{\mathcal{D}}\left(x_1,\ldots,x_n\right)=\mathbb{P}_{\mathcal{D}}\left(X_1=x_1,\ldots,X_n=x_n\right)=\prod_{i=1}^{n}\mathbb{P}_{\mathcal{D}}\left(X_i=x_i\right)$$

במקרה הרציף בהחלט פונקציית הנראות מוגדרת על ידי

$$\mathcal{L}_{\mathcal{D}}\left(x_1,\ldots,x_n\right)=\prod_{i=1}^{n} f_{\mathcal{D}}(x_i)$$

נעיר שלעיתים פונקצית הנראות מסומנת על ידי $\mathcal{L}\left(x_1,\ldots,x_n;\mathcal{D}\right)$.

יהיו $H_0:\mathcal{D}=\mathcal{D}_0, H_1:\mathcal{D}=\mathcal{D}_1$ השערות עבור $\mathcal{D}$. נגדיר את **פונקציית יחס הנראות** על ידי

$$\Lambda_{H_1:H_0}\left(x_1,\ldots,x_n\right)=\frac{\mathcal{L}_{\mathcal{D}_1}\left(x_1,\ldots,x_n\right)}{\mathcal{L}_{\mathcal{D}_0}\left(x_1,\ldots,x_n\right)}.$$

נאמר ש $C=\{X\in S\}$ הוא **מבחן יחס נראות** עם רף η אם לכל $(x_1,\ldots,x_n)$ עבורו $\Lambda_{H_1:H_0}\left(x_1,\ldots,x_n\right)>\eta$ מתקיים $(x_1,\ldots,x_n)\in S$, ולכל $(x_1,\ldots,x_n)$ עבורו $\Lambda_{H_1:H_0}\left(x_1,\ldots,x_n\right)\leq\eta$ מתקיים $(x_1,\ldots,x_n)\notin S$.

הערות

1. המשמעות של יחס הנראות עבור $(x_1,x_2,...,x_n)$, היא "פי כמה סביר שתתקבל (נקודה זו) תחת H_1 מאשר תחת H_0". לכן ככל שיחס זה גדול יותר נשתכנע יותר לדחות את H_0.

2. כזכור,אנו מעוניינים במבחן C עם שגיאה $\alpha_S = P_{H_0}(S)$ קטנה ככל האפשר ועוצמה $1-\beta_S = \pi_S =$ $P_{H_1}(S)$ גדולה ככל האפשר. איך נעשה זאת? נגדיל את אזור הדחייה בדרך שבה נגדיל מעט את α_S, ובאותו זמן נגדיל בהרבה את π_S(ככל האפשר). ולכן בהינתן ערכי יחס הנראות, נעדיף להכניס לאזור הדחייה S את הנקודות $(x_1, x_2, ..., x_n)$ שעבורם $\Lambda_{H_1:H_0}(x_1, x_2, ..., x_n)$ גדול יותר(למה ?!). וזה בדיוק מה שעושה מבחן יחס הנראות

משפט: (**הלמה של ניימן-פירסון**) יהי C מבחן יחס נראות (עם רף η כלשהו) אז הוא מיטבי; **<u>מבחן כזה נקרא גם מבחן ניימן-פירסון</u>**. כלומר, אם יש למבחן ניימן-פירסון C יש רמת מובהקות α_C, אזי לכל מבחן אחר C' שרמת המובהקות שלו $\alpha_{C'} \leq \alpha_C$ מתקיים $\beta_{C'} \geq \beta_C$.

ניתן לחשוב על משפט זה כך: נניח שאנחנו "מוכנים לסבול" הסתברות לטעות מסוג ראשון שהיא לכל היותר α כלשהו, כלומר אנחנו מחפשים רק מבחנים C שעבורם $\alpha_C \leq \alpha$. נניח שהצלחנו למצוא מבחן יחס נראות C עם רף η שעבורו אכן $\alpha_C = \alpha$. אזי מובטח לנו שכל מבחן אחר C' שיקיים $\alpha_{C'} \leq \alpha$ בהכרח יקיים $\beta_{C'} \geq \beta_C$.

הערה. קיימים גם מבחני ניימן-פירסון סטוכסטיים דטרמיניסטיים שעליהם תדברו בהרצאה.

דוגמה: נתון כי ההסתברות לזכות בלוטו היא 0.01. בדוכן מסוים המוכר מבטיח לנו כי ההסתברות לזכות אצלו היא 0.1. כדי לבדוק האם טענת המוכר נכונה, אנחנו ממלאים טפסים בדוכן עד שזוכים בלוטו בפעם הראשונה. נמצא מבחן מיטבי שעבורו ההסתברות לטעות מסוג ראשון היא לכל היותר $\alpha = 0.05$, ונחשב את ההסתברות לטעות מסוג שני במבחן המיטבי שמצאנו.

נכתוב את ההשערות $H_0 : \mathcal{D} = \text{Geo}(0.01)$, $H_1 : \mathcal{D} = \text{Geo}(0.1)$. מהלמה של ניימן-פירסון עלינו לבחור מבחן יחס נראות. נחשב את פונקציית יחס הנראות: עבור כל $1 \leq n$,

$$\Lambda_{H_1:H_0}(n) = \frac{\mathcal{L}_{\mathcal{D}_1}(n)}{\mathcal{L}_{\mathcal{D}_0}(n)} = \frac{(1-0.1)^{n-1} \cdot 0.1}{(1-0.01)^{n-1} \cdot 0.01} = \frac{10}{1.1^{n-1}}$$

מבחן יחס נראות עם רף η הוא

$$f(n) = \begin{cases} 1 & \frac{10}{1.1^{n-1}} > \eta \\ 0 & \text{otherwise} \end{cases} = \begin{cases} 1 & n < 1 + \log_{1.1}\left(\frac{10}{\eta}\right) \\ 0 & \text{otherwise} \end{cases}.$$

נסמן $N = 1 + \log_{1.1}\left(\frac{10}{\eta}\right)$ ונקבל סה"כ שמשפחת המבחנים שלנו היא מהצורה:

$$f(n) = \begin{cases} 1 & n < N \\ 0 & \text{otherwise} \end{cases}$$

עבור $N \in [1, \infty)$ (הערכים האחרים לא רלוונטים). נחשב את ההסתברות לטעות מסוג ראשון:

$$\begin{aligned} \alpha &= \mathbb{P}_{H_0}(f(n) = 1) = \mathbb{P}_{H_0}(n < N) \\ &= 1 - 0.99^{N-1} \end{aligned}$$

-דורשים כי $\alpha \leq 0.05$ ולכן

$$1 - 0.99^{N-1} \leq 0.05 \iff 0.99^{N-1} \geq 0.95 \iff N \leq 1 + \log_{0.99} 0.95 \approx 6.1$$

ולכן N שהוא לכל היותר 6 יתאים ו-N גדול יותר יגדיל את השגיעה אל מעבר ממה שדרשנו. נחשב את ההסתברות לטעות מסוג שני עבור מבחן יחס נראות עם $N=5$:

$$\begin{aligned}\beta &= \mathbb{P}_{H_1}\left(f\left(n\right)=0\right)=\mathbb{P}_{H_1}\left(n>5\right)\\ &=\left(1-0.1\right)^5=0.9^5\approx 0.59\end{aligned}$$

הערה. שימו לב כי לא תמיד ניתן יהיה לחשב את אלפא בצורה מפורשת, לפעמים נצטרך לחסום אותו בעזרת חסמים כמו מרקוב/צ'בישב וכו'. כמו כן, שימו לב שניתן לחשוב על הניסוי הגאומטרי כמספר ניסויי ברנולי בלתי תלויים, ואנחנו מצאנו את כמות ה"דגימות" הנחוצה בכדי להקטין את α.

דוגמה: יהי X משתנה מקרי רציף בהחלט עם פונקציית צפיפות שנתונה על ידי

$$f_X\left(x;\theta\right)=\begin{cases}\theta x^{\theta-1}, & 0\le x\le 1\\ 0, & otherwise\end{cases}.$$

נדרש למצוא את המבחן המיטבי עבור $\alpha=0.05$ בכדי לבחון את $H_0:\theta=3$ אל מול $H_1:\theta=2$. נניח שיש לנו תצפית יחידה x. אז מתקיים

$$\frac{\mathcal{L}_{\mathcal{D}_1}\left(x\right)}{\mathcal{L}_{\mathcal{D}_0}\left(x\right)}=\frac{2x^{2-1}}{3x^{3-1}}=\frac{2}{3x}$$

מכאן, משפחת המבחנים היא:

$$g\left(x\right)=\begin{cases}1 & \frac{2}{3x}>\eta\\ 0 & \text{otherwise}\end{cases}$$

שזה שקול (על ידי פעולות אלגבריות) למשפחה

$$g\left(x\right)=\begin{cases}1 & x<\eta\\ 0 & \text{otherwise}\end{cases}$$

עבור $\eta\in\left(0,\infty\right)$
הלמה תבטיח לנו שמבחן יחס נראות הוא מיטבי ולכן נחשב מהו η שנדרש

$$\alpha=P\left(X\le\eta\text{ when }\theta=3\right)=\int_0^\eta 3x^2dx=0.05$$

נפתור בכדי למצוא את η ונקבל

$$\eta=\left(0.05\right)^{1/3}\approx 0.368$$

ניתן לחשב גם את β:

$$\beta=P\left(X>\eta\text{ when }\theta=2\right)=\int_\eta^1 2xdx=1-\eta^2\approx 0.864$$

3.3 סטטיסטי מספיק

הגדרה. סטטיסטי הוא פונקציה כלשהי של המדגם, למשל אם המדגם נתון על ידי סדרת משתנים-מקריים $X = (X_1, ..., X_n)$.

$$\sin(X) = (\sin(X_1), ..., \sin(X_n)$$
$$|X| = (|X_1|, ..., |X_n|)$$

הם סטטיסטים.

סטטיסטי מספיק: סטטיסטי $f(X)$ נקרא סטטיסטי מספיק ביחס למבחן T אם קיימת פונקציה h כך ש-$T(X) = h(f(X))$. בפרט, יחס הנראות הוא תמיד סטטיסטי מספיק.

דוגמא: ניח שאמיר ירון רוצה לבדוק האם מבטע כלשהו הוא מאוזן. לשם כך הוא מחליט להטיל את המטבע 100 פעמים ועל סמך התוצאות יקבע האם המטבע מאוזן או לא. נגדיר משתנה מקרי X שיספור את מספר העצים ב-100 הטלות. נשים לב ש $X \sim Bin(100, p)$ ואת ההשערות הנבדקות נרשום על ידי:

$$H_0 : p = \frac{1}{2}$$
$$H_1 : p \neq \frac{1}{2}$$

אמיר קובע את המבחן הבא: אם מספר העצים שהתקבל ב-100 הטלות יהיה קטן מ-40 הוא יחליט שהמטבע לא מאוזן. אחרת יחליט שהמטבע מאוזן. נשים לב שאזור הדחיה של המבחן הוא

$$S = \{X < 40\}$$

משום ש-X הוא משתנה מקרי בינומי, הוא יכול לקבל את כל הערכים בין 0 ל-100. נעריך את ההסתברות לטעות מסוג ראשון.

$\alpha = \mathbb{P}_{H_0}(S) = \mathbb{P}_{p=\frac{1}{2}}(X < 40) = \star$ משפט הגבול המרכזי (על-ידי נרמול) נקבל:

$\star = \mathbb{P}_{p=\frac{1}{2}}(X < 40) \simeq \Phi\left(\frac{39.5-50}{\sqrt{25}}\right) = \Phi(-2.1) = 0.0179$ כעת, ההסתברות לטעות מסוג שני היא ההסתברות שהשערת האפס לא תידחה כאשר היא לא נכונה. ההסתברות הזו תלויה במודל האלטרנטיבי והיא תהיה שונה עבור ערכי p שונים. למשל עבור $p = 0.35$ נצטרך לשם הנרמול את

$$Var(X) = 22.75$$
$$\mathbb{E}(X) = 35$$

ולכן

$$\beta = \mathbb{P}_{H_1}(S^c) = \mathbb{P}_{H_1}(40 \leq X) = 1 - \mathbb{P}_{H_1}(X < 40) \simeq$$
$$1 - \Phi\left(\frac{39.5 - 35}{\sqrt{22.75}}\right) = 1 - \Phi(0.94) = 0.5868$$

כלומר העוצמה של המבחן עבור $p = 0.35$ היא $1 - \beta = 0.4132$. באותו האופן ניתן לחשב את הטעות מסדר שני עבור p מגוונים. אילו נרצה שהמבחן יהיה ברמת מובהקות של 0.01 נרצה שההסתברות לטעות מסוג ראשון לא תעלה על 0.01. כלומר נרצה למצוא C עם $S_C = \{x < C\}$ כך שיתקיים:

$$0.01 \geq \alpha = \mathbb{P}_{p=\frac{1}{2}}(S_C) = \mathbb{P}_{p=\frac{1}{2}}(X < C) \simeq \Phi\left(\frac{C - 0.5 - 50}{\sqrt{25}}\right)$$

כלומר

$$\Phi\left(\frac{C - 0.5 - 50}{\sqrt{25}}\right) \leq 0.01 \implies C = 37$$

יתר על כן, נטען שהמבחן S_C הוא מבחן ניימן-פירסון, ולכן הוא מיטבי. ואמנם, נחשב את פונקציות הנראות

$$L_0(x_1,...,x_{100}) = \prod_{k=1}^{100} \mathbb{P}_{Ber(0.5)}(X_k = x_k) = 0.5^{100}$$
$$L_1(x_1,...,x_{100}) = \prod_{k=1}^{100} \mathbb{P}_{Ber(0.35)}(X_k = x_k) = (0.35)^{\sum_{k=1}^{100} x_k} (0.65)^{100-\sum_{k=1}^{100} x_k}$$

אז יחס הנראות הוא

$$\Lambda_{H_1:H_0}(x_1,...,x_{100}) = \frac{L_1(x_1,...,x_{100})}{L_0(x_1,...,x_{100})} = \frac{(0.35)^{\sum_{k=1}^{100} x_k} (0.65)^{100-\sum_{k=1}^{100} x_k}}{0.5^{100}} = \left(\frac{0.5}{0.65}\right)^{-100} \left(\frac{0.35}{0.65}\right)^{-\sum_{k=1}^{100} x_k}$$

קיבלנו שיחס הנראות הוא **פונקציה מונוטונית יורדת בסכום** $\sum_{k=1}^{100} x_k$**.** מכאן נובע שחסם מלעיל על הסכום שקול לחסם מלערע על יחס הנראות.

$$\Lambda_{H_1:H_0}(x_1,...,x_{100}) > \eta \quad \Leftrightarrow \quad \sum_{k=1}^{100} x_k < C$$

הטיעון הזה מספיק כדי לומר ש-S_C הוא מבחן ניימן-פירסון ולכן מיטבי. אם בכל זאת רוצים לראות זאת באופן מפורש, נניח שיחס הנראות חסום על ידי η, אז על ידי העברת אגפים

$$\left(\frac{0.5}{0.65}\right)^{100} \left(\frac{0.35}{0.65}\right)^{\sum_{k=1}^{100} x_k} < \frac{1}{\eta}$$
$$\Updownarrow$$
$$\sum_{k=1}^{100} x_k < \log_{\left(\frac{0.65}{0.35}\right)} \left(\frac{\frac{1}{\eta}}{\left(\frac{0.5}{0.65}\right)^{100}}\right)$$

הכיוון השני (לעבור מחסם על הסכם לחסם על יחס הנראות) דומה מאוד. שימו לב שקיבלנו בעצם שהסטטיסטי $\sum_{k=1}^{100} x_k$ (זה סטטיסטי כי זו פונקציה של הדגימות) הוא סטטיסטי מספיק ביחס למבחני ניימן - פירסון.