פתרון מועד א

17 במרץ 2025

פתרון שאלה 1

1. הגדרה 4.27 בספר הקורס.
2. הגדרה 4.24.
3. טענה 4.28.
4. טענה 4.30.

הערות לפתרון:

1. לסעיפים א' ו- ב'
   (א) כשאתם מציעים הגדרה, עמדו בפיתוי להסביר את עצמכם. בין הטעויות הנפוצות במבחן היו שגיאות שבהם בתשובה נכתבה הגדרה (נכונה) ואחריה "כלומר..." שהיה שגוי ולא שקול להגדרה המקורית.
   (ב) אם בהוכחה השתמשתם בהגדרה מסעיף קודם, ודאו שיש התאמה בין השימושים! לעיתים הוכחה שאתם זוכרים יותר טוב מאשר ההגדרה, תעזור לכם לנסח את ההגדרה בצורה נכונה.
   (ג) הקפידו על סימונים עקביים וקוהרנטיים. הם יצילו אתכם משגיאות לוגיות מיותרות.
2. לסעיפים ג' ו-ד'
   (א) כל מעבר וכל נימוק (בשאלת הוכחה במיוחד) חייבים להיות מפורשים. אם כתבתם נימוק והבודק צריך להבין או להשלים *איך* הנימוק רלוונטי, לא עניתם תשובה מלאה.
   (ב) אם אתם אומרים "למדנו בכיתה" עליכם לצטט משפט מפורש. אם אינכם יכולים - אולי לא למדנו את זה בכיתה. או שתוכיחו בעצמכם ליתר ביטחון, או שתמצאו דרך לעקוף את מה שאינכם בטוחים בו.

פתרון שאלה 2

1. נסמן Y_i את תוצאת הטלת הקוביה ה-i, כפי שראינו רבות בקורס $X \sim \text{Geo}\left(\frac{1}{2}\right)$ ו- $Y_i \sim \text{Unif}\left([6]\right)$ לכל i קוביה שהוטלה. בנוסף ניזכר שתוחלת זריקת קוביה הוגנת היא 3.5, כלומר $\mathbb{E}\left(Y_i\right) = 3.5$ לכל i קוביה שהוטלה. כעת נשתמש בנוסחת התוחלת השלמה עם החלוקה $\{X = n\}_{n\in\mathbb{N}}$ בכדי לחשב את התוחלת של Y

$$\begin{aligned}\mathbb{E}\left(Y\right) &= \sum_{n=1}^{\infty}\mathbb{P}\left(X=n\right)\mathbb{E}\left(Y|X=n\right) = \sum_{n=1}^{\infty}\mathbb{P}\left(X=n\right)\mathbb{E}\left(\sum_{i=1}^{i=n}Y_i|X=n\right)\\ &= \sum_{n=1}^{\infty}\mathbb{P}\left(X=n\right)\sum_{i=1}^{i=n}\left(\mathbb{E}\left(Y_i\right)|X=n\right) = \sum_{n=1}^{\infty}\mathbb{P}\left(X=n\right)3.5\cdot n =\\ &3.5\cdot\mathbb{E}\left(X\right) \underbrace{=}_{X\sim\text{Geo}\left(\frac{1}{2}\right)} 7\end{aligned}$$

כאשר השוויון הראשון הוא נוסחת התוחלת השלמה, השוויון השלישי הוא לינאריות התוחלת, השוויון הרביעי הוא תוחלת של הטלת קוביה בהינתן שהוטלה, השוויון החמישי הוא הגדרה של התוחלת של X אחרי שהוצאנו 3.5 מחוץ לסכום.

2. ראשית נחשב את התפלגות מספר ההטלות בהינתן המאורע $X=Y$. יהי $n\in\mathbb{N}$ אז מתקיים

$$\begin{aligned}\mathbb{P}\left(X=n|X=Y\right)&=\frac{\mathbb{P}\left(X=n,X=Y\right)}{\mathbb{P}\left(X=Y\right)}=\frac{\mathbb{P}\left(X=n,X=Y\right)}{\sum_{k=1}^{\infty}\mathbb{P}\left(X=k,X=Y\right)}\\&=\frac{\left(\frac{1}{6}\right)^n\cdot\left(\frac{1}{2}\right)^n}{\sum_{k=1}^{\infty}\left(\frac{1}{6}\right)^k\cdot\left(\frac{1}{2}\right)^k}=\left(\frac{1}{12}\right)^{n-1}\cdot\frac{11}{12}\end{aligned}$$

השוויון הראשון נובע מהגדרה של הסתברות מותנית, השוויון השני נובע מנוסחת הסתברות השלמה עם החלוקה $\{X=k\}_{k=1}^{\infty}$, השוויון השלישי נובע מכך שההסתברות שהמטבע הוטל בדיוק n פעמים היא $\left(\frac{1}{2}\right)^n$, כפול ההסתברות של המאורע הבלתי תלוי שיצא רק 1 בקוביה מתוך n זריקות, $\left(\frac{1}{6}\right)^n$. השוויון האחרון הוא נוסחת טור הנדסי. לכן $X|X=Y\sim\text{Geo}\left(\frac{11}{12}\right)$ ומכך נקבל

$$\mathbb{E}\left(X|X=Y\right)=\frac{12}{11}$$

פתרון שאלה 3

1. עבור $i,j\in[n]$ שונים, נסמן $Z_{i,j}$ המציין של המאורע "X_i ו- X_j עוקבים מעגלית".

$$\mathbb{P}\left(Z_{i.j}=1\right)=\sum_{k=1}^{m}\mathbb{P}\left(X_i\in\{k-1,k+1\}\mod m\mid X_j=k\right)\mathbb{P}\left(X_j=k\right)=\sum_{k=1}^{m}\frac{2}{m}\cdot\frac{1}{m}=\frac{2}{m}$$

לכן $Z_{i,j}\sim\text{Ber}\left(\frac{2}{m}\right)$. נבחין כי $Y_n=\sum_{1\le i<j\le n}Z_{i,j}$. מלינאריות התוחלת נקבל

$$\mathbb{E}\left(Y_n\right)=\sum_{1\le i<j\le n}\mathbb{E}\left(Z_{i,j}\right)=\binom{n}{2}\cdot\frac{2}{m}=\frac{n\left(n-1\right)}{m}$$

נראה כי אוסף המשתנים $\{Z_{i,j}\}_{1\le i<j\le n}$ בלתי תלוי **בזוגות**. אם אין אינדקס משותף אז $Z_{i,j}$ ו- $Z_{k,l}$ הם בלתי תלויים כי $Z_{i,j}$ הוא פונקציה של משתנים מקריים בלתי תלויים במשתנים ש- $Z_{k,l}$ הוא פונקציה שלהם. במקרה שקיים בדיוק אינדקס אחד משותף j. ראינו משפט שכדי לבדוק שמשתנים ברנולי בלתי תלויים מספיק לבדוק אי תלות עבור המקרה ששני המשתנים שווים לאחד.

$$\mathbb{P}\left(Z_{i,j}=1,Z_{j,k}=1\right)=\frac{m\cdot2\cdot2}{m^3}=\left(\frac{2}{m}\right)^2=\mathbb{P}\left(Z_{i,j}=1\right)\mathbb{P}\left(Z_{j,k}=1\right)$$

כאשר השוויון הראשון נובע מכך שיש m^3 אופציות שונות לתוצאות של X_i,X_j,X_k, כדי שיתקיים המאורע הנתון, לכל תוצאה ש- X_jיכול לקבל, ל- X_k,X_iיש שתי אופציות אפשריות כל אחד. מכך נקבל

$$\text{Var}\left(Y_n\right)=\sum_{1\le i<j\le n}\text{Var}\left(Z_{i,j}\right)=\binom{n}{2}\cdot\frac{2}{m}\cdot\frac{m-2}{m}=\frac{n\left(n-1\right)\left(m-2\right)}{m^2}$$

2. אם קיים קבוע כזה אז מתקיים

$$c=\mathbb{E}\left(c\right)=\lim_{n\to\infty}\mathbb{E}\left(\frac{Y_n}{n^2}\right)=\lim_{n\to\infty}\frac{n\left(n-1\right)}{n^2m}=\frac{1}{m}$$

מטענה 6.19 בספר הקורס כל שנותר להראות הוא שהשונות שאופת לאפס, ואכן

$$\lim_{n\to\infty}\text{Var}\left(\frac{Y_n}{n^2}\right)=\lim_{n\to\infty}\frac{\text{Var}\left(Y_n\right)}{n^4}=\lim_{n\to\infty}\frac{n\left(n-1\right)\left(m-2\right)}{n^4m^2}=0$$

פתרון שאלה 4

1. התוחלת מוגדרת כי האינטגרל $\int_{-\infty}^{\infty}|x|\,e^{-|x|}dx$ מתכנס, נראה זאת

$$\int_0^\infty |x|\,e^{-|x|} = \lim_{M\to\infty}\int_0^M |x|\,e^{-|x|} = \lim_{M\to\infty}\int_0^M xe^{-x}$$
$$= \lim_{M\to\infty}\left[-xe^{-x}\right]_0^M - \int_0^M -e^{-x} = \lim_{M\to\infty} -Me^{-M} - \left(e^{-M}-1\right) = 1$$

יכולים גם להשתמש באינטגרל הנתון במקום לחשב במפורש. באופן דומה

$$\int_{-\infty}^0 |x|\,e^{-|x|} = 1$$

התוחלת היא אפס כי פונקצית הצפיפות סימטרית, אפשר גם להראות זאת באופן מפורש כך

$$\int_{-\infty}^\infty xe^{-|x|}dx = \int_0^\infty xe^{-|x|}dx + \int_{-\infty}^0 xe^{-|x|}dx$$
$$= \int_0^\infty xe^{-x}dx + \int_{-\infty}^0 xe^{x}dx \underbrace{=}_{x=-y} 1 + \int_\infty^0 -ye^{-y}\left(-dy\right)$$
$$= 1 - \int_0^\infty ye^{-y}dy = 1-1=0$$

כאשר השוויון השני בשורה השניה הוא שינוי משתנה ומחישוב אינטגרל הנתון בשאלה, והשוויון אחריו הוא הפיכת הגבולות של האינטגרל. השונות היא

$$\mathrm{Var}\left(X\right) = \int_0^\infty \frac{x^2}{2}e^{-x}dx + \int_{-\infty}^0 \frac{x^2}{2}e^{x}dx = 1 + \int_{-\infty}^0 \frac{y^2}{2}e^{-y}\left(-dy\right)$$
$$1 + \int_0^\infty \frac{y^2}{2}e^{-y}dy = 1+1=2$$

כאשר כמו מקודם נעזרנו בשינוי משתנה $x=-y$, בהפיכת גבולות האינטגרציה ובחישוב אינטגרל הנתון בשאלה.

2. ראשית שימו לב שאי שוויון של הופדינג לא עובד כאן כי התומך של המשתנה המקרי הנתון לא חסום. נחשב את הפונקציה יוצרת מומנטים של X

$$M_X\left(t\right) = \int_0^\infty \frac{e^{tx}}{2}e^{-|x|}dx = \int_0^\infty \frac{e^{tx}}{2}e^{-x}dx + \int_{-\infty}^0 \frac{e^{tx}}{2}e^{x}dx \underbrace{=}_{-1<t<1} \left[\frac{e^{x(t-1)}}{2\left(t-1\right)}\right]_0^\infty + \left[\frac{e^{x(t+1)}}{2\left(t+1\right)}\right]_{-\infty}^0$$
$$= \frac{1}{2\left(1-t\right)} + \frac{1}{2\left(1+t\right)} = \frac{1}{1-t^2}$$

תחום הגדרה של הפונקציה הוא $-1<t<1$. מאחר והמשתנים המקרים בלתי תלויים, פונקציה יוצרת מומנטים של הסכום היא מכפלה של הפונקציות יוצרת מומנטים כלומר

$$M_{\sum_{i=1}^n X_i}\left(t\right) = \left(1-t^2\right)^{-n}$$

מצרנוף נקבל כי

$$\mathbb{P}\left(\sum_{i=1}^n X_i - \mathbb{E}\left(X_i\right) < n\right) = \mathbb{P}\left(\sum_{i=1}^n X_i < n\right) \le \left(1-t^2\right)^{-n}e^{-nt} = e^{-n\left(t+ln\left(1-t^2\right)\right)}$$

בחירה למשל של $t=\frac{1}{2}$ יגדיר לנו c מתאים, לא בהכרח אופטימלי.

פתרון שאלה 5

1. ראשית נבחין כי לכל תוצאה $(x_1, \ldots, x_8)$ מתקיים

$$L_{H_0}(x_1, \ldots x_8) = \left(\frac{1}{2}\right)^8$$

נסמן ב- $S \subseteq \mathbb{R}^8$ את הקבוצה המתארת את המבחן. כדי שהתקיים $\alpha = \frac{37}{256}$ צריך להתקיים

$$\frac{37}{256} = \alpha = \mathbb{P}_{H_0}\left((x_1, \ldots, x_8) \in S\right) = \frac{|S|}{2^8} \Rightarrow |S| = 37$$

נזכר בלמה של פירסון שאומרת שמבחן יחס נראות הוא מבחן מיטבי עבור המובהקות והעוצמה שלו. מבחן יחס נראות הוא מבחן מהצורה הבאה

$$S = \left\{(x_1, \ldots, x_8) \mid \Lambda_{H_1:H_0}(x_1, \ldots x_8) > \eta\right\}$$

עבור $\eta > 0$ כלשהי. מהבחנה הראשונה בפתרון נוכל לרשום את S בצורה הבאה

$$S = \left\{(x_1, \ldots, x_8) \mid L_{H_1}(x_1, \ldots x_8) > \frac{\eta}{2^8}\right\}$$

לכן אם ניקח S להיות קבוצת התוצאות שבהן שיצא לפחות 6 עצים, נקבל שהמבחן הוא מבחן עם רף $\eta = 2^8 \cdot \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^6$ ולכן מיטבי ומתקיים $|S| = 37$, וראינו למעלה כי מכך נובע $\alpha = \frac{37}{256}$ כפי שנדרש מהמבחן.

2. נסמן ב- X את מספר העצים שיצאו. נבחין כי ניתן לרשום את פונקצית הנראות באופן הבא

$$\Lambda_{H_1:H_0}(x_1, \ldots x_8) = 2^8 \left(\frac{1}{4}\right)^{8-X} \left(\frac{3}{4}\right)^X = \frac{3^X}{2^8}$$

ולכן יחס הנראות הוא פונקציה של מספר העצים X, ולכן מספר העצים הוא סטטיסטי מספיק.

3. נחשב את עוצמת המבחן שמצאנו,

$$\beta = \mathbb{P}_{H_1}\left((x_1, \ldots, x_8) \in S^c\right) = \mathbb{P}_{X \sim \text{Bin}\left(8, \frac{3}{4}\right)}(X \leq 5)$$

ולכן העוצמה היא

$$\pi = 1 - \beta = \mathbb{P}_{X \sim \text{Bin}\left(8, \frac{3}{4}\right)}(X \geq 6) = \left(\frac{3}{4}\right)^8 + 8\left(\frac{3}{4}\right)^7 \cdot \frac{1}{4} + \binom{8}{2}\left(\frac{3}{4}\right)^6 \cdot \left(\frac{1}{4}\right)^2$$