הסתברות וסטטיסטיקה לדו־חוגי ־ פתרון תרגול 11
אי שוויונות מרקוב וצ'בישב ושאלת חזרה

יותם חרובי

סמסטר א' תשע"ח

1. ממבחן של ד"ר שלומי רובינשטיין, מרץ 2007
מבצעים סדרה של 9 הטלות בלתי תלויות של מטבע הוגן ומתעניינים ברצפים של שתי תוצאות "עץ". מצאו חסם על ההסתברות שמספר הרצפים לקבל לפחות ארבעה רצפים בסה"כ.
פתרון:
נמצא חסם בעזרת אי שוויון מרקוב.
נסמן ב X את מספר הרצפים שהתקבלו. X מקבל ערכים בין 0ל 8. יהי Y_iאינדיקטור לכך שבהטלה ה iובהטלה ה $i+1$ מתקבלות תוצאות "עץ". מכאן:

$$X = \sum_{i=1}^{8} Y_i$$

ולכן

$$\mathbb{E}X = \sum_{i=1}^{8} \mathbb{E}Y_i = 8 \cdot \mathbb{E}Y_1 = 8 \cdot \frac{1}{2} \cdot \frac{1}{2} = 2$$

כעת לפי אי שוויון מרקוב:

$$P(X \geq 4) \leq \frac{\mathbb{E}X}{4} = \frac{2}{4} = \frac{1}{2}$$

2. ממבחן של ד"ר שלומי רובינשטיין, ינואר 2016
מבצעים סידרה אינסופית של הטלות מטבע. על צד אחד של המטבע מופיע 0, ועל הצד השני מופיע 1.
המטבע נופל על 0 בהסתברות $\frac{2}{3}$ ועל 1 בהסתברות $\frac{1}{3}$ באופן בלתי תלוי בהטלות האחרות.
עבור $1 \leq i \leq \infty$ יהי X_i אינדיקטור לקבלת התוצאה 1 בהטלה ה ־ i.
עבור $1 \leq i \leq \infty$ יהי Z_i אינדיקטור לקבלת תוצאות שונות זו מזו בהטלות ה ־ i וה ־ $i+1$.
עבור $1 \leq n \leq \infty$ יהי $S_n = \Sigma_{i=1}^{n} X_i$
עבור $1 \leq n \leq \infty$ יהי $T_n = \Sigma_{i=1}^{n} Z_i$
סמנו את התשובה הנכונה.

(א) מאי שוויון צ'בישב ניתן להסיק ש $P\left(|S_{90} - 30| > 10\right) < 0.05$
(ב) מאי שוויון צ'בישב ניתן להסיק ש $P\left(|S_{90} - 30| < 10\right) < 0.5$

(ג) מאי שוויון צ'בישב ניתן להסיק ש $\lim_{n\to\infty} P\left(|S_n - \frac{n}{3}| < 10\right) = 1$
(ד) מאי שוויון מרקוב ניתן להסיק ש $P\left(T_{90} \geq 60\right) \leq \frac{2}{3}$
(ה) אף תשובה אינה נכונה.

פתרון:

ראשית שימו לב שכדי לפסול תשובה, מספיק להראות שהטענה שמופיעה בה לא נובעת מאי השוויון המוזכר בה. אין צורך להראות שהטענה עצמה איננה נכונה!

כדי לבחון את טענות א' ב' ו - ג', נחשב את התוחלת והשונות של המשתנה S_n:

$$\mathbb{E}S_n = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\left(\mathbb{E}X_i\right) = \sum_{i=1}^{n}\frac{1}{3} = \frac{n}{3}$$

$$Var(S_n) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\left(Var(X_i)\right) = n\cdot\frac{1}{3}\cdot\frac{2}{3} = \frac{2n}{9}$$

כאשר הסתמכנו על כך ש $\forall i:\ X_i \sim Ber\left(\frac{1}{9}\right)$ ב"ת.
עבור S_{90} אי שוויון צ'בישב נותן:

$$P\left(|S_{90} - \mathbb{E}S_{90}| \geq 10\right) = P\left(|S_{90} - 30| \geq 10\right) \leq \frac{Var(S_{90})}{10^2} = \frac{20}{100} = 0.2$$

$$P\left(|S_{90} - 30| > 10\right) < P\left(|S_{90} - 30| \geq 10\right) \leq 0.2$$

כלומר אי שוויון צ'בישב מבטיח חסם הרבה פחות הדוק מכפי שנטען בתשובה א', ולכן היא נפסלת.

תשובה ב' מדברת על חסם להסתברות שנקבל ערכים **קרובים** לתוחלת, אבל אי שוויון צ'בישב נותן חסם על ההסתברות לקבל ערכים **הרחוקים** מהתוחלת ביותר מאשר גודל מסויים. מכאן שלא ייתכן שטענה זו נובעת מאי שוויון צ'בישב, ודי בכך כדי לפסול את תשובה ב'.
הערה: ניתן אפילו להראות ש $P\left(|S_{90} - 30| < 10\right) > 0.5$: ראינו כבר ש $P\left(|S_{90} - 30| \geq 10\right) \leq 0.2$ ולכן הסתברות המאורע המשלים היא: $P\left(|S_{90} - 30| < 10\right) \geq 0.8$ ולא כפי שנטען בסעיף ב'.

תשובה ג' טוענת שאי שוויון צ'בישב מבטיח שההסתברות ש S_n יתרחק מתוחלתו ביותר מ 10 , שואפת לאפס כאשר n הולך וגדל. זה לא ייתכן משום שהשונות של S_n עולה (באופן לינארי) ב n. במילים אחרות, אם ננסה להפעיל את אי שוויון צ'בישב, נצטרך ראשית לעבור למאורע המשלים (שימו לב לכיוון אי השוויון), ונקבל:

$$P\left(|S_n - \frac{n}{3}| < 10\right) = 1 - P\left(|S_n - \frac{n}{3}| \geq 10\right) \geq 1 - \frac{\frac{2n}{9}}{10^2} = 1 - \frac{2n}{900}$$

והביטוי הזה בוודאי שאינו מתכנס ל 1.

נראה כעת שתשובה ד' נכונה. ראשית נחשב את התוחלת של T_n:

$$\mathbb{E}T_n = \mathbb{E}\left(\sum_{i=1}^{n} Z_i\right) = \sum_{i=1}^{n}\left(\mathbb{E}Z_i\right) = \sum_{i=1}^{n} 2\cdot\frac{1}{3}\cdot\frac{2}{3} = n\cdot\frac{4}{9}$$

$$\mathbb{E}T_{90} = 90 \cdot \frac{4}{9} = 40$$

T_{90} הוא מ"מ אי שלילי ולכן ניתן להפעיל עליו את אי שוויון מרקוב:

$$P\left(T_{90} \geq 60\right) \leq \frac{\mathbb{E}T_{90}}{60} = \frac{40}{60} = \frac{2}{3}$$

3. שאלת חזרה בנושא התפלגות דו מימדית רציפה.
ממבחן של ד"ר אסף כהן, ינואר 2013
יהיו X ו Y מ"מ שצפיפותם המשותפת נתונה על ידי

$$f_{XY}(x,y) = \begin{cases} (c+6)e^{-cx} & x > y > 0 \\ (0 & otherwise \end{cases}$$

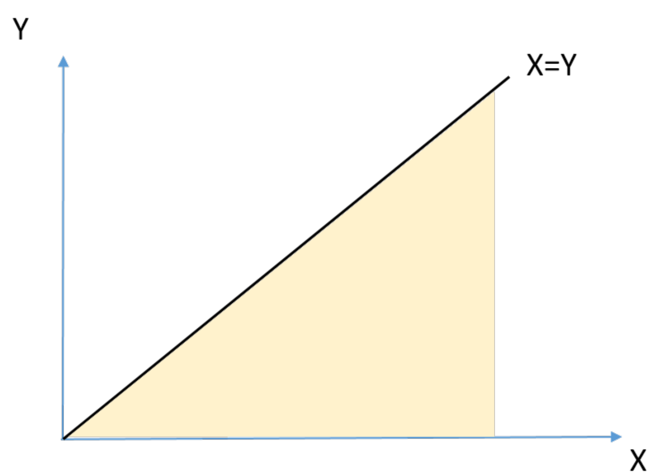
(א) חשבו את הפרמטר c

(ב) מצאו את התפלגות המ"מ $Y \mid X = x$

(ג) חשבו את $\mathbb{E}(\frac{Y}{X})$

פתרון:

(א) הצפיפות המשותפת מקבלת ערכים שונים מאפס בתחום הבא:



כדי למצוא את c נדרוש שהאינטגרל הכפול של הצפיפות על התחום הזה יהיה שווה ל 1ונקבל משוואה ב c:

$$1 = \int_0^\infty \int_0^x (c{+}6)e^{-cx}\mathrm{d}y\mathrm{d}x = \int_0^\infty (c{+}6)\cdot xe^{-cx}\mathrm{d}x = \frac{c+6}{c}\int_0^\infty c\cdot xe^{-cx}\mathrm{d}x = \frac{c+6}{c}\cdot\frac{1}{c} = \frac{c+6}{c^2}$$

כאשר האינטגרל האחרון הוא התוחלת של מ"מ מעריכי עם פרמטר c. קיבלנו אם כך:

$$c = 3$$

(ב) אם קובעים $X = x$ פונקצית הצפיפות המשותפת מקבלת ערך קבוע חיובי ממש עבור $y \in (0, x)$ ואפס עבור כל ערך אחר של y. לכן:

$$Y \mid X = x \sim U(0, x)$$

ניתן גם לחשב את הצפיפות לפי ההגדרה:

$$f_{Y|X=x}(y) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{(c+6) \cdot e^{-cx}}{\int_0^x (c+6)e^{-cx}\mathrm{d}y} = \frac{(c+6) \cdot e^{-cx}}{(c+6) \cdot xe^{-cx}} = \frac{1}{x}$$

שימו לב שניתן היה לפתור את הסעיף הזה גם ללא הסעיף הקודם (כלומר לא היינו חייבים להציב את c)

(ג) בסעיף הקודם מצאנו את ההתפלגות של $Y \mid X = x$. נוכל להעזר בכך כדי לחשב תוחלת שלמה:

$$\mathbb{E}(\frac{Y}{X}) = \mathbb{E}\mathbb{E}(\frac{Y}{X} \mid X) = \mathbb{E}\mathbb{E}(\frac{Y \mid X}{X}) = \mathbb{E}(\frac{X/2}{X}) = \mathbb{E}(\frac{1}{2}) = 1/2$$