פתרון מטלה 08 — תורת ההסתברות 1, 80420

27 בדצמבר 2025

שאלה 1

יהי $1 \leq n$ ויהי $X \sim Unif(\{-n, ..., n\})$.

נוכיח שאם $f : \mathbb{R} \to \mathbb{R}$ היא פונקציה זוגית (כלומר $f(x) = f(-x)$ לכל $x \in \mathbb{R}$), אז המשתנים המקריים $X, f(X)$ הם בלתי־מתואמים ונבחן האם הם בלתי־תלויים.

הוכחה: ראשית נגיד כי שני משתנים מקריים X, Y הם בלתי־מתואמים אם מתקיים

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0$$

ראשית נבחין

$$\mathbb{E}(X) = \sum_{i \in \{-n, \cdots, n\}} i\mathbb{P}(X = i) \underset{X \sim Unif(\{-n,...,n\})}{=} \sum_{i \in \{-n, \cdots, n\}} \frac{i}{2n} = \frac{1}{2n} \sum_{i \in \{-n, \cdots, n\}} i = 0$$

ולכן בהכרח יתקיים

$$\mathrm{Cov}(X, Y) = \mathbb{E}(Xf(X)) - \mathbb{E}(X)\mathbb{E}(f(X)) = \mathbb{E}(Xf(X))$$

נשאר לחשב את $\mathbb{E}(Xf(X))$, ניזכר שראינו שאם $Y = f(X)$ אז

$$\mathbb{E}(Y) = \sum_{k \in \mathrm{supp}} f(k)\mathbb{P}(X = k)$$

ולכן אם נגדיר $g : \mathbb{R} \to \mathbb{R}$ על־ידי $g(X) = X$, אז g היא פונקציה אי־זוגית כי $g(-x) = -x$ לכל $x \in \mathbb{R}$.

אז $h(X) = g(X) \cdot f(X)$ היא פונקציה אי־זוגית שכן

$$h(-X) = g(X) \cdot f(-X) \underset{\substack{f \text{ זוגית} \\ g \text{ אי־זוגית}}}{=} -g(X) \cdot f(X) = -X \cdot f(X) = -g(X)$$

ולכן אם נסמן $Y = h(X)$

$$\mathbb{E}(Xf(X)) = \mathbb{E}(Y) = \sum_{i \in \{-n, \cdots, n\}} h(i)\mathbb{P}(X = i) = \sum_{i=-n}^{n} if(i)\mathbb{P}(X = i) \underset{X \sim Unif(\{-n,...,n\})}{=} \frac{1}{2n+1} \sum_{i=-n}^{n} if(i)$$

$$= \frac{1}{2n+1} \sum_{i=1}^{n} (-if(-i) + if(i)) \underset{f(-i)=f(i)}{=} \frac{1}{2n+1} \sum_{i=1}^{n} (-if(i) + if(i)) = 0$$

אז $X, f(X)$ בלתי־מתואמים.

נטען שלא בהכרח שהם בלתי־תלויים: ניקח $f(x) = x^2 = Y$ ונניח שהם בלתי־תלויים ולכן צריך להתקיים

$$\mathbb{P}(X = n, Y = n^2) = \mathbb{P}(X = n)\mathbb{P}(Y = n^2)$$

נשים לב שההסתברות $\mathbb{P}(X = n, Y = n^2)$ זו ההסתברות שגם $X = n$ וגם $Y = n^2$, אבל אם $X = n$ אז בהכרח $Y = n^2$ (כי $Y = f(x)$) אז

$$\mathbb{P}(X = n, Y = n^2) = \mathbb{P}(X = n) = \frac{1}{2n+1}$$

מצד שני

$$\mathbb{P}(X = n)\mathbb{P}(Y = n^2) = \mathbb{P}(X = n)\mathbb{P}(X = n \text{ or } X = -n) \underset{\text{אי־תלות}}{=} \mathbb{P}(X = n)(\mathbb{P}(X = n) + \mathbb{P}(X = -n))$$

$$= \frac{1}{2n+1} \cdot \frac{2}{2n+1} = \frac{2}{(2n+1)^2}$$

אבל לכל $n \in \mathbb{N}$ מתקיים

$$\frac{1}{2n+1} \neq \frac{2}{(2n+1)^2}$$

ולכן לא בהכרח שהמאורעות בלתי־תלויים. $\square$

שאלה 2

מטילים באופן בלתי־תלוי מטבע על כל משבצת בלוח משבצות בגודל $n \times n$.
יהי X_n המשתנה המקרי הסופר את מספר הריבועים מהצורה $\begin{bmatrix} H & H \\ H & H \end{bmatrix}$.
נחשב את התוחלת והשונות של X_n ונראה כי $\frac{X_n}{\mathbb{E}(X_n)} \xrightarrow[n\to\infty]{} 1$.
פתרון: נגדיר את המשתנה המקרי $X_{i,j}$ להיות המשתנה המקרי המציין שבמשבצת ה־(i,j) הטלת המטבע יצאה H, אז $X_{i,j} \sim Ber\left(\frac{1}{2}\right)$ (המטבע הוגן).
נגדיר $Y_{i,j} = X_{i,j}X_{i+1,j}X_{i,j+1}X_{i+1,j+1}$ המשתנה המקרי שאומר שיש לנו ריבוע שכולו H שמתחיל בקורדינאטה (i,j) (כלומר, קורדינאטה זו היא הפינה השמאלית שלו).
מהיות ההטלות בלתי־תלויות, $Y_{i,j} \sim Ber\left(\frac{1}{2^4}\right) = Ber\left(\frac{1}{16}\right)$ וכן $X_n = \sum_{1\le i,j\le n-1} Y_{i,j}$.
כאשר האינדקסים נובעים מגבולות אפשריים של ריבועים על לוח מסדר $n \times n$.
מלינאריות התוחלת נקבל

$$\mathbb{E}(X_n) = \mathbb{E}\left(\sum_{1\le i,j\le n-1} \mathbb{E}\big(Y_{i,j}\big)\right) = (n-1)^2\mathbb{E}\big(Y_{1,1}\big) \underset{Y_{i,j}\sim Ber\left(\frac{1}{16}\right)}{=} \frac{(n-1)^2}{16}$$

בשביל חישוב השונות עלינו למצוא את כל החפיפות האפשריות (כמובן שאם שני ריבועים של H־ים אינם חופפים אז הם מייצגים הטלות בלתי־תלויות ולכן בלתי־מתואמות).
בדומה למטלה הקודמת, יש לנו חפיפה בין שני מרובעים בשתי צורות אפשריות:

1. יש לנו צורת L ולכן חולקים משבצת אחת

$$\begin{matrix} HHT \\ HHH \\ THH \end{matrix}$$

2. חולקים 2 משבצות

$$\begin{matrix} H & H & H \\ H & H & H \end{matrix}$$

נתייחס קודם למקרה השני כי הוא פשוט יותר, הוא קורה כאשר כל 6 הקורדינאטות הן הטלה של H כלומר $\mathbb{E}\big(Y_{i,j}Y_{k,l}\big) = \frac{1}{2^6}$ (שכן כמובן משתנה ברנולי נתמך על $\{0,1\}$ ובריבוע הם אותו הדבר) ולכן

$$\mathrm{Cov}\big(Y_{i,j}, Y_{k,l}\big) = \frac{1}{64} - \frac{1}{16^2} = \frac{1}{64} - \frac{1}{256} = \frac{3}{256}$$

מבחינת כמות חפיפות, יש לנו $(n-1)(n-2)$ חפיפות אופקיות וכן הנחיות ובסך־הכל $2(n-1)(n-2)$ חפיפות כאלו מהסוג הזה.
עבור המקרה הראשון, הוא קורה כאשר יש לנו 7 קורדינאטות שכולן קיבלו H, כלומר $\mathbb{E}\big(Y_{i,j}Y_{k,l}\big) = \frac{1}{2^7}$ ולכן

$$\mathrm{Cov}\big(Y_{i,j}Y_{k,l}\big) = \frac{1}{128} - \frac{1}{256} = \frac{1}{256}$$

נשאר לבחון כמה פעמים מקרה כזה קורה:
אם יש לנו מרובע שמתחיל ב־(i,j) אז ריבוע שחופף לו בידיוק בקורדינאטה אחת מתחיל ב־$(i+1,j+1)$, כלומר הגבולות שלנו מצטמצמים באחד מהקורדינאטות שאנחנו יכולים לקבל,כלומר יש $(n-2)^2$ אפשרויות כאלו (כמובן צריך לכפול ב־2 כי זה רק לחפיפה תחתונה לצורך העניין), ולכן יש לנו $2(n-2)^2$ אפשרויות לחיתוכים של L.
אם־כך, מנוסחה לחישוב סכום שנויות מהיות

$$\mathrm{Var}(X_n) = \mathrm{Var}\left(\sum_{1\le i,j\le n-1} \mathrm{Var}\big(Y_{i,j}\big)\right) = \sum_{1\le i,j\le n-1} \mathrm{Var}\big(Y_{i,j}\big) + 2\sum_{(i,j)<(k,l)} \mathrm{Cov}\big(Y_{i,j}, Y_{k,l}\big)$$

ראשית

$$Y_{i,j} \sim Ber\left(\frac{1}{16}\right) \Longrightarrow \mathrm{Var}\big(Y_{i,j}\big) = \frac{1}{16}\left(1 - \frac{1}{16}\right) = \frac{15}{256}$$

ולכן

$$\mathrm{Var}(X_n) = (n-1)^2 \cdot \frac{15}{256} + 2\left(2(n-1)(n-2) \cdot \frac{3}{256} + 2(n-2)^2 \cdot \frac{1}{256}\right)$$

$$= \frac{15n^2 - 30n + 15}{256} + 2\left(\frac{3n^2 - 9n + 6}{128} + \frac{n^2 - 4n + 4}{128}\right) = \frac{15n^2 - 30n + 15}{256} + \frac{4n^2 - 13n + 10}{64}$$

$$= \frac{15n^2 - 30n + 15}{256} + \frac{4\left(4n^2 - 13n + 10\right)}{256} = \frac{15n^2 - 30n + 15 + 16n^2 - 52n + 40}{256}$$

$$= \frac{31n^2 - 82n + 55}{256}$$

נשאר להראות $\frac{X_n}{\mathbb{E}(X_n)} \xrightarrow[n\to\infty]{} 1$, יהי $\varepsilon > 0$ ונראה כי $\mathbb{P}\left(\left|\frac{X_n}{\mathbb{E}(X_n)} - 1\right| \geq \varepsilon\right) \xrightarrow[n\to\infty]{} 0$.
ממה שמצאנו אודות $\mathrm{Var}(X_n)$ נובע כי קיים $C \in \mathbb{R}$ כך ש־$\mathrm{Var}(X_n) \leq Cn^2$ ונקבל אם כך

$$\mathbb{P}\left(\left|\frac{X_n}{\mathbb{E}(X_n)} - 1\right| \geq \varepsilon\right) = \mathbb{P}(|X_n - \mathbb{E}(X_n)| \geq \varepsilon\mathbb{E}(X_n)) \underset{\text{אי־שיוויון צ'בישב}}{\leq} \frac{\mathrm{Var}(X_n)}{\varepsilon^2\mathbb{E}(X_n)^2} \leq \frac{Cn^2}{\varepsilon^2\left(\frac{(n-1)^2}{16}\right)^2} = \frac{16Cn^2}{\varepsilon^2(n-1)^4} = 0$$

שכן יש לנו ביטוי מהצורה

$$\frac{n^2}{n^4} = \frac{1}{n^2} \xrightarrow[n\to\infty]{} 0$$

$\square$

שאלה 3

יהיו $X_1, X_2, \ldots$ משתנים מקריים בלתי־תלויים כך ש־$X_i \sim Ber\left(\frac{1}{2}\right)$.
נגדיר $Y_i = \max(X_i, X_{i+1})$ ו־$Z_n = \sum_{i=1}^{n} Y_i$.

סעיף א'

נחשב את התוחלת והשונות של Z_n.

פתרון: ראשית נבחין כי $Y_i \in \{0,1\}$ שכן $X_i \sim Ber\left(\frac{1}{2}\right)$ ונפרט את המקרים האפשריים:

$$Y_i = 1 \iff \{\{X_i = X_{i+1} = 1\}, \{X_i = 1, = X_{i+1} = 0\}, \{X_i = 0, = X_{i+1} = 1\}\}$$
$$Y_i = 0 \iff \{X_i = X_{i+1} = 0\}$$

כלומר $Y_i = 0$ אם ורק אם $X_i = X_{i+1} = 0$ ו־$Y_i = 1$ אם לפחות אחד מ־X_i, X_{i+1} הם 1.
מהאי־תלות אנחנו יכולים להסיק שמתקיים $Y_i = 0$ בהסתברות של $\frac{1}{4}$ ו־$Y_i = 1$ בהסתברות של $\frac{3}{4}$, כלומר $Y_i \sim Ber\left(\frac{3}{4}\right)$.
מלינאריות התוחלת נקבל

$$\mathbb{E}(Z_n) = \mathbb{E}\left(\sum_{i=1}^{n} Y_i\right) = n \cdot \mathbb{E}(Y_1) = \frac{3n}{4}$$

בשביל השונות נשתמש כמובן בנוסחה לחישוב סכום שנויות

$$\mathrm{Var}(Z_n) = \mathrm{Var}\left(\sum_{i=1}^{n} \mathrm{Var}(Y_i)\right) = \sum_{i=1}^{n} \mathrm{Var}(Y_i) + 2\sum_{i<j} \mathrm{Cov}(Y_i, Y_j)$$

ראשית נבחין מהיות $Y_i \sim Ber\left(\frac{3}{4}\right)$

$$\mathrm{Var}(Y_i) = \frac{3}{4}\left(1 - \frac{3}{4}\right) = \frac{3}{16}$$

כעת, נבחין כי המאורעות Y_i, Y_j בלתי־תלויים אם $j = i + 2$ שכן אז הם ערכים של פונקציות שמייצגות מאורעות בלתי־תלויים.
אז המקרה המעניין היחידי עבורו לחשב שונות זה המקרה של Y_i, Y_{i+1}:

$$\mathrm{Cov}(Y_i, Y_{i+1}) = \mathbb{E}(Y_i Y_{i+1}) - \mathbb{E}(Y_i)\mathbb{E}(Y_{i+1})$$

נשים לב שלחשב את $\{Y_i Y_{i+1} = 1\}$ יותר קל לחשב עם המשלים, כלומר להוריד את המקרים הרעים:
$\{Y_i Y_{i+1} = 0\}$ קורה אם $Y_i = 0$ או $Y_{i+1} = 0$ כאשר שניהם קורים אם ורק אם $X_i = X_{i+1} = 0$ או $X_{i+1} = X_{i+2} = 0$ כלומר אם ורק אם יש לנו את אחת מהשלשות הבאות

$$(X_i, X_{i+1}, X_{i+2}) \in \{000, 001, 100\}$$

נשים לב שכל מאורע כזה הוא מהסתברות של $\frac{1}{8}$, ולכן $\mathbb{P}(Y_i Y_{i+1} = 1) = 1 - \frac{3}{8} = \frac{5}{8}$ ונקבל

$$\mathrm{Cov}(Y_i, Y_{i+1}) = \frac{5}{8} - \left(\frac{3}{4}\right)^2 = \frac{5}{8} - \frac{9}{16} = \frac{1}{16}$$

כלומר הנוסחה שאנחנו מחשבים לפיה היא

$$\mathrm{Var}(Z_n) = \mathrm{Var}\left(\sum_{i=1}^{n} \mathrm{Var}(Y_i)\right) = \sum_{i=1}^{n} \mathrm{Var}(Y_i) + 2\sum_{i=1}^{n-1} \mathrm{Cov}(Y_i, Y_{i+1}) = \frac{3n}{16} + \frac{2(n-1)}{16} = \frac{5n-2}{16}$$

$\square$

סעיף ב'

נוכיח שלכל $\varepsilon > 0$ מתקיים $\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{2n}\sum_{i=1}^{2n} Y_i - \frac{3}{4}\right| \geq \varepsilon\right) = 0$.

הוכחה: מהסעיף הקודם אנחנו יכולים להסיק

$$2n \cdot \frac{3}{4} = \mathbb{E}(Z_{2n})$$

אז מהגדרת

$$Z_n = \sum_{i=1}^{n} Y_i \Longrightarrow \sum_{i=1}^{2n} Y_i = Z_{2n}$$

ולכן נקבל

$$\mathrm{Var}(Z_{2n}) = \mathrm{Var}\left(\sum_{i=1}^{2n} \mathrm{Var}(Y_i)\right) = \sum_{i=1}^{2n} \mathrm{Var}(Y_i) + 2\sum_{i=1}^{2n-1} \mathrm{Cov}(Y_i, Y_{i+1}) = \frac{6n}{16} + \frac{2(2n-1)}{16} = \frac{10n+2}{16}$$

ובסך־הכל

$$\mathbb{P}\left(\left|\frac{1}{2n}\sum_{i=1}^{2n} Y_i - \frac{3}{4}\right| \geq \varepsilon\right) = \mathbb{P}(|Z_{2n} - \mathbb{E}(Z_{2n})|) \geq 2n\varepsilon) \underset{\text{אי־שיוויון צ'בישב}}{\leq} \frac{\mathrm{Var}(Z_{2n})}{4n^2\varepsilon^2} = \frac{\frac{10n+2}{16}}{4n^2\varepsilon^2} = \frac{10n+2}{64n^2\varepsilon^2} \xrightarrow[n\to\infty]{} 0$$

מאריתמטיקה של גבולות. $\square$

שאלה 4

יהיו X, Y משתנים מקריים ב־L^2 בעלתי תוחלת μ_X, μ_Y וסטיות תקן σ_X, σ_Y בהתאמה.
נראה כי

$$\min\{\mathbb{E}((X-Z)^2) \mid Z \in \mathrm{Span}(1,Y))\} = \mathbb{E}\Big(\big(X - \mathrm{Proj}_{1,Y}(X)\big)^2\Big) = \sigma_X^2 - \frac{\mathrm{Cov}(X,Y)^2}{\sigma_Y^2}$$

כאשר ההטלה של X על $\mathrm{Span}(1,Y)$ נתונה במפורש על־ידי $\mathrm{Proj}_{1,Y}(X) = \mu_X + \frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}(Y-\mu_Y)$.
הוכחה: נשים לב ש־$\mathrm{Span}(1,Y)$ מגדיר את כל הפונקציות הלינאריות של Y, כלומר $\{a+bY \mid a,b \in \mathbb{R}\}$ ולפי טענה שראינו בהרצאה (6.51 בספר), הערך Z שממעזר את הסטייה הריבועית של $\mathbb{E}((X-Z)^2)$ היא ההטלה האורתוגונלית של X על ה־Span הזה.
נגדיר $\beta = \frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}$ ונגדיר

$$\hat{X} = \mathrm{Proj}_{\mathrm{Span}(1,Y)}(X) = \mu_X + \frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}(Y-\mu_Y)$$

ואנחנו רוצים לחשב את $\mathbb{E}\Big(\big(X-\hat{X}\big)^2\Big)$ כאשר

$$X - \hat{X} = X - \mu_X - \beta(Y-\mu_Y)$$

נגדיר משתנים מקריים ממורכזים

$$\tilde{X} = X - \mu_X,\ \tilde{Y} = Y - \mu_Y$$

ומתקיים מהגדרות

$$\mathbb{E}(\tilde{X}) = 0,\ \mathbb{E}(\tilde{Y}) = 0,\ \mathbb{E}(\tilde{X}^2) = \mu_X^2,\ \mathbb{E}(\tilde{Y}^2) = \mu_Y^2, \mathbb{E}(\tilde{X}\tilde{Y}) = \mathrm{Cov}(X,Y)$$

ונקבל

$$\begin{aligned}\mathbb{E}\Big(\big(X-\hat{X}\big)^2\Big) &= \mathbb{E}\Big(\big(\tilde{X}-\beta\tilde{Y}\big)^2\Big) = \mathbb{E}(\tilde{X}^2 - 2\beta\tilde{X}\tilde{Y} + \beta^2\tilde{Y}^2) = \mathbb{E}(\tilde{X}^2) - 2\beta\mathbb{E}(\tilde{X}\tilde{Y}) + \beta^2\mathbb{E}(\tilde{Y}^2)\\ &= \sigma_X^2 - 2\beta\,\mathrm{Cov}(X,Y) + \beta^2\sigma_Y^2\end{aligned}$$

נציב בחזרה את β ונקבל

$$\begin{aligned}\mathbb{E}\Big(\big(X-\hat{X}\big)^2\Big) &= \sigma_X^2 - 2\left(\frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}\right)\mathrm{Cov}(X,Y) + \left(\frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}\right)^2\sigma_Y^2 = \sigma_X^2 - 2\frac{\mathrm{Cov}(X,Y)^2}{\sigma_Y^2} + \frac{\mathrm{Cov}(X,Y)^2}{\sigma_Y^2}\\ &= \sigma_X^2 - \frac{\mathrm{Cov}(X,Y)^2}{\sigma_Y^2}\end{aligned}$$

כלומר

$$\min_{Z\in\mathrm{Span}(1,Y)} \mathbb{E}((X-Z)^2) = \sigma_X^2 - \frac{\mathrm{Cov}(X,Y)^2}{\sigma_Y^2}$$

□

שאלה 5

סעיף א'

יהי X משתנה מקרי בעל תוחלת 0 המקיים $|X| \leq M$ כמעט־תמיד עבור $M > 0$. נוכיח שלכל $t \in \mathbb{R}$ מתקיים

$$\mathbb{E}(e^{tX}) \leq \exp\left(\frac{M^2t^2}{2}\right)$$

הוכחה: נשחזר את ההוכחה מההרצאה בהתאם לשינויים הנדרשים.
ראשית מהיות $|X| \leq M$ מתקיים $-M \leq X \leq M$ ואם נגדיר את הפונקציה $f(x) = e^{tx}$ כפי שראינו בהרצאה היא פונקציה קמורה.
אנחנו רוצים למתוח קו ישר בין $(-M, e^{-tM})$ ולבין (M, e^{tM}) בקטע $[-M, M]$, והמשוואה לקו־ישר נתונה על־ידי

$$y = \frac{e^{tM} - e^{-tM}}{2M}(x - (-M)) + e^{-tM} = \frac{x}{2M}(e^{tM} - e^{-tM}) + \frac{e^{tM} - e^{-tM}}{2} + e^{-tM} = \left(\frac{M+x}{2M}\right)e^{tM} + \left(\frac{M-x}{2M}\right)e^{-tM}$$

ומקמירות אנחנו יודעים שמתקיים

$$e^{tx} \leq \left(\frac{M+x}{2M}\right)e^{tM} + \left(\frac{M-x}{2M}\right)e^{-tM}$$

נפעיל על הביטוי הזה את התוחלת

$$\begin{aligned}
\mathbb{E}(e^{tX}) &\leq \mathbb{E}\left(\left(\frac{M+X}{2M}\right)e^{tM} + \left(\frac{M-X}{2M}\right)e^{-tM}\right) \\
&\underset{(\star)}{=} \left(\frac{M+\mathbb{E}(X)}{2M}\right)e^{tM} + \left(\frac{M-\mathbb{E}(x)}{2M}\right)e^{-tM} \\
&\underset{(\star\star)}{=} \left(\frac{M}{2M}\right)e^{tM} + \left(\frac{M}{2M}\right)e^{-tM} \\
&= \frac{1}{2}e^{tM} + \frac{1}{2}e^{-tM} \\
&= \cosh(tM)
\end{aligned}$$

כאשר $(\star)$ נובע מלינאריות התוחלת, $(\star\star)$ נובע מהיות $\mathbb{E}(X) = 0$.
אבל בהרצאה ראינו שלכל $t \in \mathbb{R}$ מתקיים $\cosh(x) = \frac{e^x + e^{-x}}{2} \leq e^{\frac{x^2}{2}}$ (נובע מטור טיילור), אז בפרט עבור $x = tM$ נקבל

$$\cosh(tM) \leq e^{\frac{(tM)^2}{2}} = e^{\frac{t^2M^2}{2}} = \exp\left(\frac{t^2M^2}{2}\right)$$

כפי שהתבקשנו להראות. $\square$

סעיף ב'

יהיו $X_1, X_2, ..., X_n$ משתנים מקריים בלתי־תלויים ויהיו $M_1, M_2, ..., M_n$ מספרים חיוביים כך שלכל $i \in [n]$ מתקיים $|X_i| \leq M_i$ כמעט־תמיד ו־$\mathbb{E}(X_i) = 0$.
נוכיח כי לכל $a > 0$ מתקיים

$$\mathbb{P}(X_1 + X_2 + \cdots + X_n \geq a) \leq \exp\left(-\frac{a^2}{2M}\right)$$

כאשר $M = M_1 + M_2 + \cdots + M_n$.
הוכחה: נסמן $X = \sum_{i=1}^n X_i$, אז לכל $t > 0$ מתקיים

$$\mathbb{P}(X \leq a) = \mathbb{P}(e^{tX} \geq e^{ta}) \underset{\text{אי־שיוויון מרקוב}}{\leq} e^{-ta}\mathbb{E}(e^{tX})$$

מהיות המשתנים בלתי־תלויים

$$\mathbb{E}(e^{tX}) = \prod_{i=1}^{n} \mathbb{E}(e^{tX_i}) \underset{\text{סעיף קודם}}{\leq} \prod_{i=1}^{n} \exp\left(\frac{M_i^2 t^2}{2}\right) = \exp\left(\frac{t^2}{2}\sum_{i=1}^{n} M_i^2\right)$$

נרצה לחסום בצורה הדוקה יותר ולשים כך נשתמש בעובדה ש־$M = \sum_{i=1}^{n} M_i$ ומהיות כל $M_i > 0$ נקבל

$$\sum_{i=1}^{n} M_i^2 \leq \sum_{i=1}^{n} M_i M = M^2$$

אזי

$$\mathbb{E}(e^{tX}) \leq \exp\left(\frac{M^2 t^2}{2}\right)$$

ולכן

$$\mathbb{P}(X \leq a) \leq e^{-ta}\mathbb{E}(e^{tX}) \leq \exp\left(-ta + \frac{M^2 t^2}{2}\right)$$

בדומה להרצאה, כדי למקסם על ערך ה־t נגזור לפי t ונקבל

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(-ta + \frac{M^2 t^2}{2}\right) = -a + M^2 t \Longrightarrow t = \frac{a}{M^2}$$

כלומר

$$-ta + \frac{M^2 t^2}{2} \Longrightarrow -\frac{a^2}{M^2} + \frac{\cancel{M^2}}{2} \cdot \frac{a^2}{\cancel{M^4}} = -\frac{a^2}{M^2} + \frac{a^2}{2M^2} = -\frac{a^2}{2M^2}$$

ולכן

$$\mathbb{P}(X \leq a) \leq \exp\left(-\frac{a^2}{2M^2}\right)$$

□

סעיף ג'

נסיק את אי־שיוויון הופדינג המוכלל: יהיו $X_1, X_2, ..., X_n$ משתנים מקריים בלתי־תלויים ויהיו $M_1, M_2, ..., M_n$ מספרים חיוביים כך שלכל $i \in [n]$ מתקיים $|X_i - \mathbb{E}(X_i)| \leq M_i$ כמעט־תמיד.
נסמן $X = X_1 + \cdots + X_n, M = M_1 + \cdots + M_n$ אז לכל $a > 0$ מתקיים

$$\mathbb{P}(X - \mathbb{E}(X) \geq a) \leq \exp\left(-\frac{a^2}{2M^2}\right)$$

הוכחה: נגדיר $Y_i = X_i - \mathbb{E}(X_i)$ לכל $i \in [n]$ אז $Y_1, Y_2, \cdots, Y_n$ הם משתנים מקריים בלתי־תלויים כך ש־$\mathbb{E}(Y_i) = 0$.
נתון כי $|Y_i| = |X_i - \mathbb{E}(X_i)| \overset{a.s}{\leq} M_i$ ולכן אם נגדיר $Y = \sum_{i=1}^{n} Y_i$ נוכל להשתמש בסעיף הקודם ולקבל

$$\mathbb{P}(Y \geq a) \leq \exp\left(-\frac{a^2}{2M^2}\right) \Longrightarrow \mathbb{P}(X - \mathbb{E}(X) \geq a) \leq \exp\left(-\frac{a^2}{2M^2}\right)$$

כאשר $\Longrightarrow$ נובע מלינאריות התוחלת, שכן

$$Y = \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n}(X_i - \mathbb{E}(X_i)) = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\mathbb{E}(X_i) = \sum_{i=1}^{n} X_i - \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = X - \mathbb{E}(X)$$

□

סעיף ד'

אורנה בוחרת באקראי ובאופן אחיד מספר בין 0 ל־10, דפנה בוחרת באקראי מספר בין 0 ל־20 ודורית בוחרת באקראי מספר בין 0 ל־30, באופן בלתי־תלוי.

נשתמש בתוצאות הסעיף הקודם כדי לתת חסם מלרע להסתברות שסכום המספרים ששלושתן בחרו נמצא בין 0 ל־49.

פתרון: נסמן X_1, X_2, X_3 המספר שאורנה, דפנה ודורית בחרו בהתאמה ונסמן $X = X_1 + X_2 + X_3$.

ניזכר כי אם $Y \sim Unif([n])$ אז $\mathbb{E}(Y) = \frac{n+1}{2}$, הספירה שלנו כוללת את הקצוות ולכן

$$\mathbb{E}(X_1) \underset{X_1 \sim Unif([11])}{=} \frac{12}{2} = 6,\ \mathbb{E}(X_2) \underset{X_2 \sim Unif([21])}{=} \frac{22}{2} = 11,\ \mathbb{E}(X_3) \underset{X_3 \sim Unif([31])}{=} \frac{32}{2} = 16$$

לכן כמעט־תמיד מתקיים

$$|X_1 - \mathbb{E}(X_1)| \leq 6,\ |X_2 - \mathbb{E}(X_2)| \leq 11,\ |X_3 - \mathbb{E}(X_3)| \leq 16$$

ומהסעיף הקודם נקבל

$$\mathbb{P}(X \geq 50) = \mathbb{P}(X - \mathbb{E}(X) \geq 17) \leq \exp\left(-\frac{17^2}{2(6^2 + 11^2 + 16^2)}\right) = \exp\left(-\frac{289}{826}\right)$$

$\square$

שאלה 6

שיכור עומד על ציר המספרים השלמים. הוא מטיל מטבע שנופל על עץ בהסתברות q. אם יוצא לו עץ הוא הולך שני צעדים ימינה ואם יוצא לו פלי הוא הולך צד אחד שמאלה. הוא מבצע n הטלות מטבע כאלו באופן בלתי־תלוי. יהי X_n מיקומו אחרי n הצעדים האלו.
נמצא עבור אילו ערכים של q ניתן לחסום בצורה לא טריוויאלית את $\mathbb{P}(X_n \geq 0)$ בעזרת אי־שיוויון הופדינג ונמצא מהו החסם שמתקבל.
פתרון: נגדיר Y_i המשתנה המקרי של כמות הצעדים שהשיכור זז בצעד ה־i, קרי זה משתנה מקרי מתפלג ברנולי והתומך שלו הוא $\{2,-1\}$, כלומר

$$\mathbb{P}(Y_i = 2) = q,\ \ \mathbb{P}(Y_i = -1) = 1 - q$$

ו־$X_n = \sum_{i=1}^{n} Y_i$ סך כל הצעדים, כאשר כל Y_i בלתי־תלויים מהנתון.
בהתאם התוחלת של משתנה ברנולי מוזז

$$\mathbb{E}(Y_i) = (2 \cdot q) + (-1 \cdot (1-q)) = 2q - 1 + q = 3q - 1$$

בשביל להשתמש באי־שיוויון הופדינג עלינו לנרמל את המשתנה המקרי לעיל כאשר המשתנה המקרי המנורמל, נסמנו Z_i, יהיה בעל תוחלת 0 ויתקיים $|Z_i| \overset{a.s.}{\leq} 1$.
בשביל התנאי של התוחלת, נגדיר $Z_i = Y_i - \mathbb{E}(Y_i) = Y_i - 3q + 1$ אך עדיין לא מתקיים $|Z_i| \overset{a.s.}{\leq} 1$:
אנחנו כבר יודעים ש־$\text{supp}(Y_i) = \{2,-1\}$, אז נחלק למקרים הללו ונקבל מהחישוב הלא מדוייק לעיל

$$Y_i = 2 \Longrightarrow Z_i = Y_i - (3q-1) = 2 - 3q + 1 = 3(1-q)$$
$$Y_1 = -1 \Longrightarrow Z_i = Y_i - (3q-1) = -1 - 3q + 1 = -3q$$

ונקבל אם כך

$$|Y_i - (3q-1)| \leq \max\{3(1-q), 3q\} \leq 3$$

ולכן עלינו לנרמל בחילוק ב־3 ולכן בעצם $Z_i = \frac{Y_i - (3q-1)}{3}$ וזה יביא לנו $\mathbb{E}(Z_i) = 0, |Z_i| \overset{a.s.}{\leq} 1$.
עוד נשים לב שמתקיים $\text{supp}(Z_i) = \{-q, 1-q\} \subseteq [-1,1]$ שכן $q \leq 1$ (מהגדרת ההסתברות) ואנחנו עומדים בכל תנאי אי־שיוויון הופדינג.
נפעיל אותו ונקבל אם כך

$$\begin{aligned}
\mathbb{P}(X_n \geq 0) &= \mathbb{P}\left(\sum_{i=1}^{n} Y_i \geq 0\right) \\
&= \mathbb{P}\left(\sum_{i=1}^{n} (3Z_i + (3q-1)) \geq 0\right) \\
&= \mathbb{P}\left(\sum_{i=1}^{n} Z_i \geq \frac{n(1-3q)}{3}\right) \\
&\underset{(\star)}{\leq} \exp\left(-\frac{n^2(1-3q)^2}{9 \cdot 2n}\right) = \exp\left(-\frac{n(1-3q)^2}{18}\right)
\end{aligned}$$

נשים לב שכדי שהמהלך $(\star)$ לעיל יהיה מוגדר היטב (כי $d > 0$ בתנאי אי־שיוויון הופדינג), צריך להתקיים $1 - 3q > 0 \Longrightarrow q < \frac{1}{3}$. $\square$