מבוא להסתברות - 80430 - פתרון מועד א'

1 שאלה 1

ראו הוכחה מפורטת בספר.

2 שאלה 2

בשק אטום חמישה כדורים שחורים ושלושה לבנים. בכל פעם מוציאים מתוך השק שני כדורים וממשיכים בכך עד שנשלפים ביחד שני כדורים באותו צבע. נסמן:

X- מספר הכדורים שנשלפו.

Y- משתנה אינדיקטור לכך שבשליפה האחרונה נשלפו שני כדורים לבנים.

א. מהי התפלגותם המשותפת של X ו־Y, והאם הם בלתי תלויים?

<u>פתרון א.</u>

ראשית נשים לב כי $\text{Supp}\left(X\right)=\left\{2,4,6,8\right\},\text{Supp}\left(Y\right)=\left\{0,1\right\}$.

נתאים מרחב הסתברות לסיטואציה: נסמן $\Omega=\left\{\left(x_1,...,x_8\right)\mid\forall i,x_i\in\left\{B,W\right\},\sum_{i=1}^{8}\mathbf{1}_{x_i=B}=5,\sum_{i=1}^{8}\mathbf{1}_{x_i=W}=3\right\}$. כלומר מרחב המדגם הוא סדרות של שליפות של כל שמונת הכדורים על פי הצבעים. בסיטואציה שלנו, נדמיין כי גם לאחר שנשלפו שני כדורים בעלי אותו הצבע, נמשיך לשלוף את יתר הכדורים.

פונקציית ההסתברות המתאימה לסיטואציה היא זו האחידה, שכן כל רצף שליפות כדורים מתקבל בסיכוי שווה.

כמו כן נשים לב כי:

$$\left|\Omega\right|=\binom{8}{5}=56.$$

(יש לבחור את מיקומי הכדורים השחורים על מנת לקבוע באופן יחיד איבר מ־Ω.)

כעת:

$$\mathbb{P}\left(Y=1,X=2\right)=\left\{x\in\Omega\mid x_1=x_2=W\right\}=\frac{\binom{6}{5}}{\left|\Omega\right|}=\frac{6}{56}.$$

שכן x_1,x_2 נקבעו, ויש לבחור כיצד לסדר את ששה הכדורים הנותרים כאשר אחד מהם לבן וחמישה שחורים. בדומה:

$$\mathbb{P}\left(Y=1,X=4\right)=\left\{x\in\Omega\mid\left\{x_1,x_2\right\}=\left\{B,W\right\},x_3=x_4=W\right\}=\frac{2\cdot\binom{4}{4}}{\left|\Omega\right|}=\frac{2}{56}.$$

זאת כיוון שיש לנו בחירה של שתי אפשרויות כיצד יסתדר הזוג הראשון (WB או BW), הכדורים השלישי והרביעי נקבעו, ויש לסדר את ארבעת הכדורים השחורים הנותרים - ויש רק דרך אחת לעשות זאת).

לבסוף נשים לב כי $\mathbb{P}\left(Y=1,X=6\right)=\mathbb{P}\left(Y=1,X=8\right)=0$.

נחשב בדומה את המקרה $Y=0$:

$$\begin{aligned}
\mathbb{P}\left(Y=0,X=2\right)&=\frac{\binom{6}{3}}{|\Omega|}=\frac{20}{56},\\
\mathbb{P}\left(Y=0,X=4\right)&=\frac{2\cdot\binom{4}{2}}{|\Omega|}=\frac{12}{56},\\
\mathbb{P}\left(Y=0,X=6\right)&=\frac{2^2\cdot\binom{2}{1}}{|\Omega|}=\frac{8}{56},\\
\mathbb{P}\left(Y=0,X=8\right)&=\frac{2^3}{|\Omega|}=\frac{8}{56}.
\end{aligned}$$

נסכם הכל בטבלה:

$$\begin{array}{ccc}
\boldsymbol{X\backslash Y} & \mathbf{0} & \mathbf{1}\\
\mathbf{2} & \frac{20}{56} & \frac{6}{56}\\
\mathbf{4} & \frac{12}{56} & \frac{2}{56}\\
\mathbf{6} & \frac{8}{56} & 0\\
\mathbf{8} & \frac{8}{56} & 0
\end{array}
=
\begin{array}{ccc}
\boldsymbol{X\backslash Y} & \mathbf{0} & \mathbf{1}\\
\mathbf{2} & \frac{10}{28} & \frac{3}{28}\\
\mathbf{4} & \frac{12}{56} & \frac{1}{28}\\
\mathbf{6} & \frac{4}{28} & 0\\
\mathbf{8} & \frac{4}{28} & 0
\end{array}$$

קל לראות כי X ו־Y אינם בלתי־תלויים, למשל מכיוון ש:

$$0=P\left(X=6,Y=1\right)\neq P\left(X=6\right)\cdot P\left(Y=1\right)>0$$

הערה. ניתן לפתור את סעיף זה גם באמצעות תרשים עץ או ניסוי רב־שלבי.

לדוגמא:

$$\mathbb{P}\left(X=4,Y=1\right)=\underbrace{2\cdot\frac{5}{8}\cdot\frac{3}{7}}_{\text{BW or WB}}\cdot\underbrace{\frac{2}{6}\cdot\frac{1}{5}}_{\text{WW}}=\frac{2}{56}$$

ובאופן דומה:

$$\mathbb{P}\left(X=8,Y=0\right)=\underbrace{2\cdot\frac{5}{8}\cdot\frac{3}{7}}_{\text{BW or WB}}\cdot\underbrace{2\cdot\frac{4}{6}\cdot\frac{2}{5}}_{\text{BW or WB}}\cdot\underbrace{2\cdot\frac{3}{4}\cdot\frac{1}{3}}_{\text{BW or WB}}\cdot\underbrace{\frac{2}{2}\cdot\frac{1}{1}}_{\text{BB}}=\frac{8}{56}$$

ב. מהי תוחלת מספר הכדורים השחורים שנשלפו?

<u>פתרון סעיף ב.</u>

נסמן ב־Z את מספר הכדורים השחורים שנשלפו. נשים לב כי Z מחזיר ערך שונה בכל אחד מהזוגות הנתמכים בהתפלגות של X,Y. כלומר:

$$\left\{Z=0\right\}=\left\{X=2,Y=1\right\},\left\{Z=1\right\}=\left\{X=4,Y=1\right\}$$

וכן הלאה. על כן נוכל לכתוב את התפלגות Z:

$$\mathbb{P}\left(Z=k\right)=\begin{cases}0 & \frac{6}{56}\\ 1 & \frac{2}{56}\\ 2 & \frac{20}{56}\\ 3 & \frac{12}{56}\\ 4 & \frac{8}{56}\\ 5 & \frac{8}{56}\end{cases}$$

ונחשב את התוחלת:

$$\mathbb{E}\left(Z\right)=\sum_{k=0}^{5}k\cdot\mathbb{P}\left(Z=k\right)=\frac{0+2+40+36+32+40}{56}=\frac{150}{56}$$

ג. מה היא תוחלת מספר הכדורים שנשלפו בהינתן שבשליפה האחרונה נשלפו שני כדורים לבנים?

פתרון סעיף ג.

במילים אחרות, עלינו למצוא את $\mathbb{E}\left(X\mid Y=1\right)$. נחשב:

$$\mathbb{E}\left(X\mid Y=1\right)=\sum_{k=1}^{4}2k\cdot\mathbb{P}\left(X=2k\mid Y=1\right)=\frac{1}{\mathbb{P}\left(Y=1\right)}\sum_{k=1}^{4}2k\cdot\mathbb{P}\left(X=2k,Y=1\right)=$$
$$=\frac{2\cdot\mathbb{P}\left(X=2,Y=1\right)+4\cdot\mathbb{P}\left(X=4,Y=1\right)}{\mathbb{P}\left(Y=1\right)}=\frac{\frac{12}{56}+\frac{8}{56}}{\frac{8}{56}}=\frac{20}{8}=2\frac{1}{2}$$

3 שאלה 3

יהי $X\sim Exp\left(1\right)$. נזכיר כי $F\left(t\right)=\left(1-e^{-t}\right)\cdot\mathbf{1}_{\left(0,\infty\right)}$.

א. יש לחשב את הצפיפות של X^2.

פתרון סעיף א:

נחשב ראשית את פונקציית ההתפלגות המצטברת של X^2:

$$F_{X^2}\left(t\right)=\mathbb{P}\left(X^2\leq t\right)=\mathbb{P}\left(-\sqrt{t}\leq X\leq\sqrt{t}\right)\underset{\underset{\text{a.s.}}{X\geq 0}}{=}\mathbb{P}\left(X\leq\sqrt{t}\right)=F_X\left(\sqrt{t}\right)=\left(1-e^{-\sqrt{t}}\right)\cdot\mathbf{1}_{\left(0,\infty\right)}$$

נגזור ונקבל:

$$f_{X^2}\left(t\right)=F_{X^2}^{'}\left(t\right)=\frac{1}{2\sqrt{t}}e^{-\sqrt{t}}\cdot\mathbf{1}_{\left(0,\infty\right)}$$

ב. עלינו לחשב את:

$$F_{X\mid X\geq 1}\left(t\right)=\mathbb{P}\left(X\leq t\mid X\geq 1\right)$$

פתרון סעיף ב:

ברור כי $\mathbb{P}\left(X\leq t\mid X\geq 1\right)=0$ אם $t\leq 1$. אחרת אם $t>1$ נקבל:

$$F_{X\mid X\geq 1}\left(t\right)=\mathbb{P}\left(X\leq t\mid X\geq 1\right)=\frac{\mathbb{P}\left(1\leq X\leq t\right)}{\mathbb{P}\left(X\geq 1\right)}=\frac{\mathbb{P}\left(X\leq t\right)-\mathbb{P}\left(X\leq 1\right)}{1-\mathbb{P}\left(X\leq 1\right)}=\frac{F_X\left(t\right)-F_X\left(1\right)}{1-F_X\left(1\right)}=$$
$$=\frac{1-e^{-t}-\left(1-e^{-1}\right)}{1-\left(1-e^{-1}\right)}=\frac{e^{-1}-e^{-t}}{e^{-1}}=1-e^{-\left(t-1\right)}$$

כלומר:

$$F_{X\mid X\geq 1}\left(t\right)=\begin{cases}0 & t\leq 1\\ 1-e^{-\left(t-1\right)} & t>1\end{cases}$$

ג. עבור אילו ערכי α קיימת התוחלת $\mathbb{E}\left(e^{aX}\right)$ומה ערכה?

פתרון סעיף ג.

נשים לב כי $\mathbb{E}\left(e^{aX}\right)$ זו הפונקציה יוצרת המומנטים של X בנקודה a. אלא אם במקרה זכרנו את התשובה מהכיתה, עלינו לחשב מתי היא מוגדרת ומה ערכה, כלומר לחשב את האינטגרל:

$$\mathbb{E}\left(e^{aX}\right)=\int_{-\infty}^{\infty}e^{ax}f_X\left(x\right)dx.$$

נזכיר כי $f_X\left(x\right)=e^{-x}\cdot\mathbf{1}_{(0,\infty)}$. על כן:

$$\mathbb{E}\left(e^{aX}\right)=\int_{-\infty}^{\infty}e^{ax}f_X\left(x\right)dx=\int_{0}^{\infty}e^{ax}e^{-x}dx=\int_{0}^{\infty}e^{(a-1)x}dx$$

נשים לב כי אם $a-1\geq 0$, כלומר $a\geq 1$,יתקיים כי $e^{(a-1)x}$ מונוטונית עולה ולכן האינטגרל לא יחזיר ערך סופי.

עבור $a<1$ נקבל:

$$...=\int_{0}^{\infty}e^{(a-1)x}dx=\frac{1}{a-1}e^{(a-1)x}\mid_0^\infty=0-\frac{1}{a-1}\cdot e^0=\frac{1}{1-a}$$

לסיכום: $\mathbb{E}\left(e^{aX}\right)$ מוגדר אם ורק אם $a<1$ ושווה לערך $\frac{1}{1-a}$.

4 שאלה 4

$2n$ נורות מסודרות במעגל. כל נורה דולקת בסיכוי $\frac{1}{3}$ באופן בלתי־תלוי. אם שתי נורות סמוכות כבויות נאמר כי המקטע ביניהן חשוך. נסמן ב־Z את מספר המקטעים החשוכים.

א. יש לחשב את תוחלת ושונות Z.

פתרון סעיף א.

עבור כל מקטע $i\in[2n]$ נסמן X_i אינדיקטור האם המקטע כבוי. נשים לב כי $X_i\sim ber\left(\frac{4}{9}\right)$, וכן $Z=\sum_{i=1}^{2n}X_i$. על כן $E\left(Z\right)=\frac{4}{9}\cdot 2n=\frac{8}{9}n$. על מנת לחשב שונות נחשב ראשית שונויות משותפות.

יהיו $1\leq i\neq j\leq 2n$. אם $|i-j|=_{\mod 2n}1$ אזי נחשב את השונות המשותפת:

$$Cov\left(X_i,X_j\right)=E\left(X_iX_j\right)-E\left(X_i\right)E\left(X_j\right)=\left(\frac{2}{3}\right)^3-\left(\frac{2}{3}\right)^4=\frac{8}{27}\cdot\frac{1}{3}=\frac{8}{81}.$$

עבור כל $i\neq j$ אחרים נקבל כי X_i ו־X_j בלתי־תלויים, ולפיכך $Cov\left(X_i,X_j\right)=0$.

לבסוף נחשב:

$$Var\left(X_i\right)=\frac{4}{9}\cdot\frac{5}{9}=\frac{20}{81}$$

על פי הנוסחה מהכיתה לשונות של משתנה מקרי ברנולי. על כן:

$$Var\left(Z\right)=Var\left(\sum_{i=1}^{2n}X_i\right)=\sum_{i=1}^{2n}Var\left(X_i\right)+2\sum_{1\leq i<j\leq 2n}Cov\left(X_i,X_j\right)=2n\cdot\frac{20}{81}+2\cdot 2n\cdot\frac{8}{81}=\frac{40}{81}n+\frac{32}{81}n=\frac{72}{81}n=\frac{8}{9}n$$

ב. יש להראות כי לכל $\Delta>0$ קיימים $c,C>0$ כך שלכל $n>0$ מתקיים:

$$\mathbb{P}\left(Z-\mathbb{E}\left(Z\right)>\Delta n\right)\leq Ce^{-c\Delta^2n}$$

פתרון סעיף ב.

נפצל את Z לסכום על המקטעים הזוגיים והאי־זוגיים: נסמן

$$Z_1 = X_1 + X_3 + ... + X_{2n-1}, Z_2 = X_2 + X_4 + ... + X_{2n}$$

נשים לב כי $Z = Z_1 + Z_2$, וכן כי Z_i הם סכום של n משתנים מקריים בלתי תלויים. כעת:

$$\mathbb{P}\left(Z - \mathbb{E}\left(Z\right) > \Delta n\right) = \mathbb{P}\left(Z_1 - \mathbb{E}\left(Z_1\right) + Z_2 - \mathbb{E}\left(Z_2\right) > \Delta n\right) \leq \mathbb{P}\left(\left\{Z_1 - \mathbb{E}\left(Z_1\right) > \frac{\Delta}{2}n\right\} \bigcup \left\{Z_2 - \mathbb{E}\left(Z_2\right) > \frac{\Delta}{2}n\right\}\right)$$

כאשר במעבר האחרון השתמשנו במונוטוניות, ובעובדה שאם סכום של שני מספרים גדול מ־Δn, אזי לפחות אחד המספרים צריך להיות גדול מ־$\frac{1}{2}\Delta n$.

כעת על ידי חסם האיחוד, ומסימטריה ($Z_1 - \mathbb{E}\left(Z_1\right)$ ו $Z_2 - \mathbb{E}\left(Z_2\right)$ שווי התפלגות) נקבל כי:

$$... \leq \mathbb{P}\left(Z_1 - \mathbb{E}\left(Z_1\right) > \frac{\Delta}{2}n\right) + \mathbb{P}\left(Z_2 - \mathbb{E}\left(Z_2\right) > \frac{\Delta}{2}n\right) = 2\mathbb{P}\left(Z_1 - \mathbb{E}\left(Z_1\right) > \frac{\Delta}{2}n\right)$$

כעת נשים לב כי

$$Z_1 - \mathbb{E}\left(Z_1\right) = \sum_{i \in [2n] \cap \mathbb{N}_{odd}} \left(X_i - \mathbb{E}\left(X_i\right)\right)$$

כלומר יש לנו כאן סכום של n משתנים מקריים בלתי־תלויים, שכל אחד מהם חסום על ידי 1 (שהרי $0 \leq X_i \leq 1$, והורדנו ממנו מספר בין 0 ל־1).

על כן נוכל להפעיל את אי־שוויון הופדינג ולקבל:

$$\mathbb{P}\left(Z - \mathbb{E}\left(Z\right) > \Delta n\right) \leq 2 \cdot e^{-\frac{\left(\frac{\Delta}{2}n\right)^2}{2n}} = 2 \cdot e^{-\frac{\Delta^2}{8}n}$$

לפיכך הקבועים $C = 2, c = \frac{1}{8}$ מקיימים את המבוקש.

ג. יהי W הוא מספר המקטעים הזוגיים החשוכים. יש למצוא קבועים a, b, c כך שהמשתנה המקרי $\frac{W-a}{bn^c}$ יתכנס בהתפלגות ל־$N\left(0,1\right)$ כאשר n שואף לאינסוף.

<u>פתרון סעיף ג.</u>

נשים לב כי W הוא פשוט Z_2 תחת הסימונים מסעיף ב'. על פי משפט הגבול המרכזי, עלינו לבחור את הקבועים כך שהביטוי $\frac{W-an}{bn^c}$ יהיה סכום של משתנים מקריים שווי־התפלגות מתוחלת 0 ושונות 1, ושחולקו ב־$\sqrt{n}$.

כמקודם, $Z_2 = \sum_{i=1}^{n} X_{2i}$. על כן $\frac{X_{2i} - \mathbb{E}X_{2i}}{\sqrt{\frac{20}{81}}}$ הם משתנים מקריים בלתי־תלויים מתוחלת 0 ושונות 1. אם נסכום אותם נקבל כי $\frac{Z_2 - \frac{4}{9}n}{\sqrt{\frac{20}{81}}}$ הוא סכום של משתנים מקריים מתוחלת 0 ושונות 1. כל שנותר הוא לחלק ב־$\sqrt{n}$ ונקבל כי הסדרה

$$\frac{Z_2 - \frac{4}{9}n}{\sqrt{\frac{20}{81}} \cdot \sqrt{n}}$$

תשאף בהתפלגות ל־$N\left(0,1\right)$. כלומר $a = \frac{4}{9}, b = \sqrt{\frac{20}{81}}, c = \frac{1}{2}$.

5 שאלה 5

H_0 - המדידה מתפלגת $N\left(1, 0.1^2\right)$.

H_1 - המדידה מתפלגת $N\left(0.9, 0.1^2\right)$.

אנו מניחים לצורכי השאלה כי $1 - \Phi\left(-1.65\right) = \Phi\left(1.65\right) = 0.95$, עבור Φ פונקציית ההתפלגות המצטברת של התפלגות נורמלית תקנית.

א. עלינו למצוא מבחן מיטבי בעל מובהקות של 0.05, ולבטא את עוצמת מבחן זה.

<u>פתרון סעיף א.</u>

ממשפט ניימן־פרסון מבחן יחס נראות הוא מבחן מיטבי. פונקציית יחס הנראות עבור מדידה יחידה הינה:

$$r(x) = \frac{f_{N(1,0.1)}(x)}{f_{N(0.9,0.1)}(x)} = \frac{\cancel{\frac{1}{0.1\sqrt{2\pi}}}}{\cancel{\frac{1}{0.1\sqrt{2\pi}}}} \cdot \frac{e^{-\frac{1}{2}\left(\frac{x-1}{0.1}\right)^2}}{e^{-\frac{1}{2}\left(\frac{x-0.9}{0.1}\right)^2}} = e^{-\frac{1}{2}\cdot\frac{1}{0.01}\left((x-1)^2-(x-0.9)^2\right)} =$$
$$= e^{-50\left(x^2-2x+1-x^2+1.8x-0.81\right)} = e^{-50(-0.2x+0.19)} = e^{10x-9.5}$$

יחס הנראות הוא פונקציה מונוטונית עולה ב־x, לכן משפחת מבחני ניימן פירסון היא מהצורה $S = \{x \mid x \le \mu\}$.
על כן נוכל לחשב את רמת המובהקות:

$$\alpha = \mathbb{P}_{H_0}(X \in S) = \mathbb{P}_{H_0}(X \le \mu).$$

עלינו להעביר את המשוואה לתנאי על התפלגות נורמלית תקנית:

$$\mathbb{P}_{H_0}(X \le \mu) = \mathbb{P}_{H_0}\left(\frac{X-1}{0.1} \le \frac{\mu-1}{0.1}\right) = \Phi\left(\frac{\mu-1}{0.1}\right) \stackrel{?}{=} 0.05 = \Phi(-1.65)$$

מכיוון ש־Φ חד־חד־ערכית (למשל כי היא עולה ממש) נוכל להשוות בין הקלטים ולקבל:

$$\frac{\mu-1}{0.1} = -1.65 \iff \mu - 1 = -0.165 \iff \mu = 0.835$$

כעת נוכל לחשב גם את עוצמת המבחן:

$$\beta = \mathbb{P}_{H_1}(X \notin S) = \mathbb{P}_{H_1}(X > \mu) = 1-\mathbb{P}(X \le \mu) = 1-\mathbb{P}_{H_1}\left(\frac{X-0.9}{0.1} \le \frac{\mu-0.9}{0.1}\right) = 1-\Phi\left(\frac{0.835-0.9}{0.1}\right) = 1-\Phi(-0.65)$$

על כן:

$$1-\beta = 1-(1-\Phi(-0.65)) = \Phi(-0.65)$$

ב. כמה מדידות דרושות על מנת להכריע בערך סמך של 0.95 ובעוצמה של 0.95?

<u>פתרון סעיף ב.</u>

על פי ההערה בנוסח השאלה אנו יודעים כי מבחן אופטימלי הוא בעל איזור דחייה מהצורה: $S = \{(x_1, ..., x_n) \in \mathbb{R}^n \mid x_1 + ... + x_n \le \mu\}$.
נסמן $\overline{X} = X_1 + ... + X_n$ משתנה מקרי של סכום הדגימות.
נקבע $\alpha = 0.05$ ונמצא את μ באמצעות α:

$$0.05 = \alpha = \mathbb{P}_{H_0}\left(\overline{X} \le \mu\right).$$

נשים לב כי תחת השערת האפס, $\overline{X}$ מתפלג $N\left(n, 0.1^2 n\right)$, ותחת ההשערה החליפית הוא מתפלג $N\left(0.9n, 0.1^2 n\right)$. כמו כן $0.05 = \Phi(-1.65)$ לפיכך:

$$\Phi(-1.65) = 0.05 = \mathbb{P}_{H_0}\left(\overline{X} \le \mu\right) = \left(\frac{\overline{X}-n}{0.1\sqrt{n}} \le \frac{\mu-n}{0.1\sqrt{n}}\right) = \Phi\left(\frac{\mu-n}{0.1\sqrt{n}}\right)$$

Φ חד־חד־ערכית, לפיכך נסיק:

$$\frac{\mu-n}{0.1\sqrt{n}} = -1.65$$
$$\mu = n - 0.165\sqrt{n}$$

כעת נמצא את n על פי הדרישה $\beta \le 0.05$:

$$\beta = \mathbb{P}_{H_1}\left(\overline{X} > \mu\right) = 1 - \mathbb{P}_{H_1}\left(\overline{X} \le \mu\right) = 1 - \mathbb{P}_{H_1}\left(\frac{\overline{X}-0.9n}{0.1\sqrt{n}} \le \frac{\mu-0.9n}{0.1\sqrt{n}}\right) = 1-\Phi\left(\frac{\mu-0.9n}{0.1\sqrt{n}}\right) \le 0.05 = 1-\Phi(1.65)$$

לפיכך:

$$\Phi\left(\frac{\mu-0.9n}{0.1\sqrt{n}}\right) \ge \Phi(1.65)$$

כלומר:

$$\frac{\mu - 0.9n}{0.1\sqrt{n}} \geq 1.65$$

נציב את μ ונקבל:

$$\frac{n - 0.165\sqrt{n} - 0.9n}{0.1\sqrt{n}} = \frac{0.1n - 0.165\sqrt{n}}{0.1\sqrt{n}} = \sqrt{n} - 1.65 \geq 1.65 \iff \sqrt{n} \geq 3.3 \iff n \geq 3.3^2$$

נשים לב כי $3.3^2 \in (10, 11)$, לפיכך קיבלנו $n \geq 11$.