פתרון מועד ג'

28 באפריל 2022

שאלה 1 פתרון אפשרי לשני הסעיפים מופיע בספר הלימוד.

שאלה 2 בכד אטום כדור שחור, כדור לבן וכדור אפור. שולפים כדורים מתוך הכד שוב ושוב עד לקבלת כדור שחור. בכל פעם ששולפים כדור לבן - מחזירים אותו לכד, אך אם שולפים אפור - משאירים אותו מחוץ לכד.

א. מה ההסתברות שהכדור האפור נותר בכד בסוף התהליך?

פתרון: דרך א': נשים לב שהכדור האפור ישאר בכד בסוף התהליך אם ורק אם הכדור השחור יישלף לפניו. היות והכדורים סימטריים, ההסתברות שאחד מהם יישלף לפני השני היא חצי ולכן חצי היא ההסתברות המבוקשת.

דרך ב': נבחין כי הכדור האפור ישאר בכד בסוף התהליך רק אם לאורך כל המשחק, לא כולל השליפה האחרונה, הוצא רק הכדור הלבן מן הכד. נסמן את המאורע הנ"ל ב-C. נגדיר בנוסף לכל $n \in \mathbb{N}$ את המאורע A_n בו הכדור השחור נשלף לאחר בדיוק n שליפות. נחשב:

$$\begin{aligned}\mathbb{P}(A_n \cap C) &= \mathbb{P}(\text{The first n-1 balls are white and the nth is black})\\ &= (\prod_{k=1}^{n-1} \frac{1}{3}) \cdot \frac{1}{3}\\ &= \frac{1}{3^n}\end{aligned}$$

כאשר המעבר השני נובע מכך שההסתברות לשלוף את הכדור הלבן בשליפה הראשנה היא $\frac{1}{3}$ וכך גם ההסתברות לשלוף כדור לבן לאחר שבכל השליפות לפני כן נשלף כדור לבן (נותרו שלושה כדורים בכד לאורך כל התהליך). לבסוף, היות והקבוצה $\{A_n\}_{n\in\mathbb{N}}$ מהווה חלוקה של מרחב ההסתברות, מנוסחת ההסתברות השלמה (או

סיגמא-אדטיביות):

$$\begin{aligned}\mathbb{P}(C) &= \sum_{n=1}^{\infty} \mathbb{P}(A_n \cap C)\\ &= \sum_{n=1}^{\infty} \frac{1}{3^n}\\ &= \frac{\frac{1}{3}}{1-\frac{1}{3}}\\ &= \frac{1}{2}\end{aligned}$$

ולכן ההסתברות המבוקשת היא חצי.

ב. מה ההסתברות שתדרשנה k שליפות עד לשליפה של הכדור השחור (כולל השליפה האחרונה)?

פתרון: צריך לחשב את הסתברותו של המאורע A_k מהסעיף הקודם (דרך ב'). נגדיר קבוצת מאורעות $\{B_l\}_{l=1}^{k-1}$ כך ש-B_l המאורע בו הכדור האפור נשלף בשלב ה-l. מאדטיוביות סופית (המאורעות זרים):

$$\begin{aligned}\mathbb{P}(A_k) &= \mathbb{P}(C \cap A_k) + \sum_{l=1}^{k-1} \mathbb{P}(B_l \cap A_k)\\ &= \frac{1}{3^k} + \sum_{l=1}^{k-1} (\frac{1}{3})^l \cdot (\frac{1}{2})^{k-l}\\ &= \sum_{l=1}^{k} \frac{1}{3^l} \cdot \frac{1}{2^{k-l}}\\ &= \frac{1}{2^k} \sum_{l=1}^{k} (\frac{2}{3})^l\\ &= \frac{1}{2^k} \cdot \frac{\frac{2}{3}((\frac{2}{3})^k - 1)}{\frac{2}{3} - 1}\\ &= \frac{1}{2^k}(2(1-(\frac{2}{3})^k))\\ &= \frac{1}{2^{k-1}} - \frac{2}{3^k}.\end{aligned}$$

ג. בהנתן שהכדור האפור נותר בכד בסוף התהליך, מה ההסתברות שתדרשנה k שליפות עד לשליפה של הכדור השחור (כולל השליפה האחרונה)?

פתרון: בסימוני הסעיפים הקודמים, צריך לחשב את $\mathbb{P}(A_k|C)$. למעשה עשינו כבר את העבודה בסעיפים הקודמים ולכן כל שנותר לעשות הוא להציב בנוסחת ההתניה:

$$\mathbb{P}(A_k|C) = \frac{\mathbb{P}(A_k \cap C)}{\mathbb{P}(C)} = \frac{\frac{1}{3^k}}{\frac{1}{2}} = \frac{2}{3^k}.$$

שאלה 3

יהיו X ו-Y משתני מקריים בלתי תלויים המתפלגים מעריכית עם פרמטר **1**.

א. מה היא הסתברות המאורע $X+Y<2$?

פתרון: המשתנים מתפלגים מעריכית עם פרמטר 1 ולכן פונקציות הצפיפות שלהם מקיימות:

$$f_Y(x)=f_X(x)=\begin{cases}e^{-x} & x\geq 0\\ 0 & x<0\end{cases}$$

נשתמש בנוסחת ההסתברות (הרציפה) השלמה:

$$\begin{aligned}\mathbb{P}(X+Y<2)=&\mathbb{P}(X+Y\leq 2)\\ =&\int_{-\infty}^{\infty}\mathbb{P}(X+Y\leq 2|X=x)f_X(x)dx\\ =&\int_{-\infty}^{\infty}\mathbb{P}(Y\leq 2-x|X=x)f_X(x)dx\\ =&\int_{0}^{\infty}\mathbb{P}(Y\leq 2-x)e^{-x}dx\\ =&\int_{0}^{\infty}F_Y(2-x)e^{-x}dx\\ =&\int_{0}^{2}(1-e^{-(2-x)})e^{-x}dx\\ =&\int_{0}^{2}(e^{-x}-e^{-2})dx\\ =&[-e^{-x}-e^{-2}x]_0^2\\ =&-e^{-2}-2e^{-2}+1\\ =&1-3e^{-2}\end{aligned}$$

כאשר השתמשנו ברציפות הסכום במעבר הראשון ובאי תלות כאשר הורדנו את ההתניה.

ב. מהי תוחלת המכפלה XY?

פתרון: המשתנים X,Y בלתי תלויים ולכן תוחלת המכפלה שלהם היא מכפלת התוחלות. ניזכר בנוסף כי תוחלתו של משתנה מעריכי עם פרמטר λ היא $\frac{1}{\lambda}$ ולכן תוחלתם של X ושל Y היא 1. נסיק כי תוחלת המכפלה אף היא 1.

ג. יש לחשב את צפיפותו המותנית של Y בהינתן $X+Y=a$.

פתרון: מנוסחת הצפיפות המותנית:

$$f_{Y|X+Y=a}(y)=\frac{f_{X+Y,Y}(a,y)}{f_{X+Y}(a)}$$

לכל a בו המכנה לא מתאפס. בפרט ניתן להניח ש- $a>0$.

אמנם לא צריך לעשות את זה אך נראה תחילה איך מחשבים את צפיפותם המשותפת של $X+Y$ ו-Y. עבור $0 \leq y \leq z$ נקבל את הביטוי הבא (לכל זוג y, z אחרים ההסתברות היא כמובן אפס):

$$\begin{aligned}
F_{X+Y,Y}(z,y) =& \mathbb{P}(X+Y \leq z, Y \leq y) \\
=& \int_0^\infty \mathbb{P}(X+Y \leq z, Y \leq y | Y=t) \cdot f_Y(t) dt \\
=& \int_0^y \mathbb{P}(X+Y \leq z | Y=t) \cdot e^{-t} dt \\
=& \int_0^y \mathbb{P}(X \leq z-t | Y=t) \cdot e^{-t} dt \\
=& \int_0^y \mathbb{P}(X \leq z-t) \cdot e^{-t} dt \\
=& \int_0^y (1-e^{t-z}) \cdot e^{-t} dt \\
=& \int_0^y (e^{-t} - e^{-z}) dt \\
=& [-e^{-t} - e^{-z} t]_0^y \\
=& -e^{-y} - y \cdot e^{-z} + 1
\end{aligned}$$

נגזור לפי שני המשתנים ונקבל: $f_{X+Y,Y}(z,y) = e^{-z} \cdot \chi_{0 \leq y \leq z}$. כעת נוכל גם להסיק את צפיפותו של המשתנה $X+Y$. עבור $z > 0$:

$$\begin{aligned}
f_{X+Y}(z) =& \int_0^\infty f_{X+Y,Y}(z,y) dy \\
=& \int_0^z e^{-z} dy \\
=& z e^{-z}
\end{aligned}$$

ובדומה לנאמר קודם, עבור $z \leq 0$ הצפיפות היא אפס. נציב בנוסחא בתחילת הסעיף ונקבל:

$$f_{Y|X+Y=a}(y) = \frac{e^{-a} \cdot \chi_{0 \leq y \leq a}}{a \cdot e^{-a}} = \begin{cases} \frac{1}{a} & 0 < y \leq a \\ 0 & otherwise \end{cases}.$$

שאלה 4

יהיו $\{X_i\}_{i \in \mathbb{N}}$ משתנים מקריים בלתי-תלויים שכל אחד מהם מתפלג אחיד על $\{0,1,2\}$. נסמן $Y_i = \chi_{X_i > X_{i+1}}$ ונסמן $Z_n = \sum_{i=1}^n Y_i$.

א. יש לחשב את תוחלת Z_n (כתלות ב-n).

פתרון: כל Y_i הינו משתנה ברנולי, נמצא את הפרמטר שלו. היות וכל המשתנים $\{X_i\}$ בלתי תלויים ושווי התפלגות, כך גם המשתנים $\{Y_i\}$ ולכן מספיק לחשב את הפרמטר של

Y_1. מתקיים:

$$\mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 > X_2) = \mathbb{P}((X_1, X_2) \in \{(2,1),(2,0),(1,0)\}) = \frac{3}{9} = \frac{1}{3}$$

כאשר חישוב ההסתברות נובע מכך שכל זוג תוצאות עבור המשתנים X_1, X_2 מתקבל בהסתברות $\frac{1}{9}$ (מאי-תלות $\frac{1}{3} \cdot \frac{1}{3}$). בפרט קיבלנו כי $Y_i \sim Ber(\frac{1}{3})$ לכל i. כעת, מליניאריות התוחלת:

$$\mathbb{E}(Z_n) = \mathbb{E}(\sum_{i=1}^{n} Y_i) = \sum_{i=1}^{n} \mathbb{E}(Y_i) = \sum_{i=1}^{n} \frac{1}{3} = \frac{1}{3}n.$$

ב. יש לחשב את שונות Z_n (כתלות ב-n).
פתרון: נרצה להשתמש בנוסחה לשונות של סכום שראינו בכיתה:

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

נבחין כי לכל i, j שאינם עוקבים, המשתנים Y_i, Y_j בלתי-תלויים ולכן גם $Cov(Y_i, Y_j) = 0$. נחשב עבור i, j עוקבים:

$$\begin{aligned}
Cov(Y_1, Y_2) =& \mathbb{E}(Y_1 \cdot Y_2) - \mathbb{E}(Y_1)\mathbb{E}(Y_2) \\
=& \mathbb{P}(Y_1 \cdot Y_2 = 1) - (\frac{1}{3})^2 \\
=& \mathbb{P}(Y_1 = 1, Y_2 = 1) - \frac{1}{9} \\
=& \mathbb{P}(X_1 > X_2 > X_3) - \frac{1}{9} \\
=& \mathbb{P}(X_1 = 2, X_2 = 1, X_3 = 0) - \frac{1}{9} \\
=& (\frac{1}{3})^3 - \frac{1}{9} \\
=& -\frac{2}{27}.
\end{aligned}$$

נציב בנוסחא הנ"ל:

$$Var(Z_n) = \sum_{i=1}^{n}(\frac{1}{3} - \frac{1}{9}) + 2\sum_{i=1}^{n-1} Cov(Y_i, Y_{i+1}) = \frac{2}{9}n - 2(n-1)\cdot\frac{2}{27} = \frac{2n+4}{27}.$$

ג. יש להראות כי קיימים $a, b > 0$ כך שמתקיים לכל n כי:

$$\mathbb{P}(Z_n - \mathbb{E}(Z_n) > \frac{n}{100}) \le be^{-an}.$$

פתרון: המשתנים $\{Y_{2i}-\mathbb{E}(Y_{2i})\}_{i\in\mathbb{N}}$ כולם בלתי-תלויים (הזזה בקבוע של משתנים בלתי תלויים מותירה אותם כאלו), בעלי תוחלת אפס וחסומים בקטע $[-1,1]$. כך גם המשתנים $\{Y_{2i-1}-\mathbb{E}(Y_{2i-1})\}_{i\in\mathbb{N}}$. נחתור להשתמש באי-שוויון הופדינג:

$$\begin{aligned}
\mathbb{P}(Z_n-\mathbb{E}(Z_n)>\frac{n}{100}) \le&\mathbb{P}(\sum_{i=1}^{n}(Y_i-\mathbb{E}(Y_i))\ge\frac{n}{100})\\
\le&\mathbb{P}(\{\sum_{i=1}^{\frac{n}{2}}(Y_{2i}-\mathbb{E}(Y_{2i}))\ge\frac{n}{200}\}\cup\{\sum_{i=1}^{\frac{n}{2}}(Y_{2i-1}-\mathbb{E}(Y_{2i-1}))\ge\frac{n}{200}\})\\
\le&\mathbb{P}(\sum_{i=1}^{\frac{n}{2}}(Y_{2i}-\mathbb{E}(Y_{2i}))\ge\frac{n}{200})+\mathbb{P}(\sum_{i=1}^{\frac{n}{2}}(Y_{2i-1}-\mathbb{E}(Y_{2i-1}))\ge\frac{n}{200})\\
=&2\mathbb{P}(\sum_{i=1}^{\frac{n}{2}}(Y_{2i}-\mathbb{E}(Y_{2i}))\ge\frac{n}{200})\\
\le&2\cdot exp(-\frac{(\frac{n}{200})^2}{2n})\\
=&2\cdot exp(-\frac{1}{80,000}n).
\end{aligned}$$

בפרט אם נבחר $a=\frac{1}{80,000}$ ו- $b=2$ נקבל את הנדרש.

שאלה 5

אלכס מצא במכרה בסיביר אבן שעשויה להיות אחד משני גבישים רדיו-אקטיביים. הזמן (בדקות) עד לפליטה של פרוטון מגביש מסוג א' מתפלג מעריכית עם פרמטר 1 ואילו מסוג ב' - מעריכית עם פרמטר $\frac{1}{100}$. הוא ממתין עד לפליטת פרוטון ראשונה שמתרחשת בזמן X. כלומר:

השערת האפס היא $X\sim Exp(1)$, ההשערה החליפית היא $X\sim Exp(0.01)$.

א. יש לנסח משפחה של מבחנים אופטימליים (ניימן-פירסון), להכרעה איזה סוג גביש מצא אלכס, על סמך זמן ההמתנה עד לפליטה הראשונה.

פתרון: לפי הלמה של ניימן-פירסון, מבחן יחס נראות הוא מיטבי. נחשב את פונקציית יחס הנראות:

$$r(x)=\frac{f_0(x)}{f_1(x)}=\frac{e^{-x}}{0.01\cdot e^{-0.01x}}=100\cdot e^{-0.99x}.$$

לכל $\eta>0$ מבחן יחס נראות עם רף η הוא:

$$f_\eta(x)=\begin{cases}1 & 100\cdot e^{-0.99x}\le\eta\\ 0 & otherwise\end{cases}=\begin{cases}1 & x\ge-\frac{1}{0.99}ln\frac{\eta}{100}\\ 0 & otherwise\end{cases}.$$

כל η כנ"ל מגדיר מבחן מיטבי וקיבלנו משפחה של מבחנים אופטימליים.

ב. יש לזהות את המבחן האופטימלי עבור מובהקות $\alpha=e^{-1}$ ולקבוע את עוצמתו β.

פתרון: עלינו למצוא η כך ש- $\mathbb{P}_{\mathcal{H}_0}(f_\eta(x)=1)=e^{-1}$. נחשב:

$$\begin{aligned}\mathbb{P}_{\mathcal{H}_0}(f_\eta(x)=1) &= \mathbb{P}_{\mathcal{H}_0}(x \geq -\frac{1}{0.99}ln\frac{\eta}{100}) \\ &= 1-\mathbb{P}_{\mathcal{H}_0}(x < -\frac{1}{0.99}ln\frac{\eta}{100}) \\ &= 1-(1-e^{\frac{1}{0.99}ln\frac{\eta}{100}}) \\ &= e^{\frac{1}{0.99}ln\frac{\eta}{100}} \\ &= e^{-1}\end{aligned}$$

כאשר המעבר השלישי הוא הצבה בפונקציה הצוברת של משתנה מעריכי עם פרמטר 1. סה"כ קיבלנו:

$$-1=\frac{1}{0.99}ln\frac{\eta}{100}$$

$$e^{-0.99}=\frac{\eta}{100}$$

$$\eta=100\cdot e^{-0.99}.$$

עוצמת המבחן עבור הרף שחישבנו:

$$\begin{aligned}\mathbb{P}_{\mathcal{H}_1}(f_\eta(x)=0) &= \mathbb{P}_{\mathcal{H}_1}(x<-\frac{1}{0.99}ln\frac{\eta}{100}) \\ &= 1-e^{-0.01\cdot(-\frac{1}{0.99}ln\frac{\eta}{100})} \\ &= 1-e^{\frac{1}{99}\cdot ln(e^{-0.99})} \\ &= 1-e^{-0.01}.\end{aligned}$$

ג. אלכס בירר וגילה שבמכרה שבו כרה אחוז אחד מהאבנים הן מסוג א' והיתר מסוג ב'. כעת הבעיה הפכה מבעיה סטטיסטית לבעיה הסתברותית. בהינתן שב-50 הדקות הראשונות לא ארעה פליטה, מה ההסתברות שמדובר בגביש מסוג א'.

פתרון: נחשב תחילה את ההסתברות שלא ארעה פליטה ב-50 הדקות הראשונות (נסמן את המאורע ב-A):

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A|\mathcal{H}_0)\cdot\mathbb{P}(\mathcal{H}_0)+\mathbb{P}(A|\mathcal{H}_1)\cdot\mathbb{P}(\mathcal{H}_1) \\ &= (1-(1-e^{-50}))\cdot 0.01+(1-(1-e^{-0.01\cdot 50}))\cdot 0.99 \\ &= 0.01\cdot e^{-50}+0.99\cdot e^{-0.5}.\end{aligned}$$

כעת, מנוסחת בייס:

$$\begin{aligned}
\mathbb{P}(\mathcal{H}_0|A) &= \mathbb{P}(A|\mathcal{H}_0)\cdot\frac{\mathbb{P}(\mathcal{H}_0)}{\mathbb{P}(A)}\\
&= e^{-50}\cdot\frac{0.01}{0.01\cdot e^{-50}+0.99\cdot e^{-0.5}}\\
&= \frac{0.01}{0.01+0.99\cdot e^{49.5}}\\
&= \frac{1}{1+99\cdot e^{49.5}}.
\end{aligned}$$