HW9 - 094412 - הסתברות מ

גור תלם - 206631848

13.01.2021

שאלה 1

יהי $X\sim N\left(\mu,\sigma^{2}\right)$. מה ההסתברות ש$X$ יסטה מתוחלתו בלפחות:

סעיף א - סטיית תקן אחת?

הוכחה: ננרמל $Z=\frac{X-\mu}{\sigma}\sim N\left(0,1\right)$. נניח כי $\sigma\geq0$ (כמידה והיה קטן מ0 היה אפשר לרשום $-\sigma$ בכל הפתרון והכל היה זהה).
ובמונחים של X: $X=\sigma Z+\mu$
המשמעות שX יסטה בלפחות a סטיות תקן אחת הינה

$$P\left(\left|X-\mu\right|>a\sqrt{\sigma^{2}}\right)=1-P\left(-a\sigma\leq X-\mu\leq a\sigma\right)$$

נשים לב שאין משמעות ככ לאי שיוויון חזק או חלש שכן ההסתברות לנקודה ספציפית הינה כמעט 0. נמשיך

$$=1-P\left(-a\sigma\leq\sigma Z\cancel{+\mu-\mu}\leq a\sigma\right)=1-P\left(-a\leq Z\leq a\right)=1-P\left(-a\leq Z\leq a\right)=1-\left(2\Phi\left(a\right)-1\right)=$$

$$=2-2\Phi\left(a\right)$$

נציב $a=1,2,3$:

$$2-2\Phi\left(1\right)\approx2-2\cdot0.841244746=0.317510508$$

■

סעיף ב - שתי סטיות תקן?

הוכחה:

$$2-2\Phi\left(2\right)\approx2-2\cdot0.97724986805182=0.04550026389636$$

■

סעיף ג - שלוש סטיות תקן?

הוכחה:

$$2-2\Phi\left(3\right)\approx2-2\cdot0.9986501019683699=0.0026997960632602$$

■

שאלה 2

נניח שגובה גבר אקראי מתפלג $M \sim N\left(178.4, 7.6^2\right)$ ושל אישה $W \sim N\left(164.7, 7.1^2\right)$.

הערה 0.1 הערת הפותר: אני רק רוצה לומר שהתקופה הזו גם ככה מדכאת, להזכיר לי שכמעט 95% מהגברים בעולם גבוהים ממני זה לא עוזר (אני מקבל בונוס שחישבתי איזה אחוזון אני?).

סעיף א - מה ההסתברות שגבר בגובה לפחות 190? מה ההסתברות שאישה?

הוכחה: אז אתחיל בלומר שההסתברות שאישה היא 0.496 נכון לאחוזי הנשים בעולם (אני צוחק על עצמי, אתם ניסחתם את השאלה סבבה).
אבל ידוע כי $Z_M = \frac{M-178.4}{7.6} \Longrightarrow M = 7.6Z_M + 178.4$
ובדומה $Z_W = \frac{W-164.7}{7.1} \Longrightarrow W = 7.1Z_W + 164.7$
כמובן שהZים הם בעצם אותו משתנה Z שכן מדובר באותה התפלגות עם אותם פרמטרים.

$$P\left(M > 190\right) = P\left(7.6Z_M + 178.4 > 190\right) = P\left(7.6Z_M > 11.6\right) = P\left(Z_M > \frac{116}{76}\right) =$$

אבל Z_M כבר מתפלג נורמלית סטנדרטית ולכן

$$= 1 - P\left(Z_M \le \frac{116}{76}\right) = 1 - \Phi\left(\frac{116}{76}\right) \approx 1 - 0.9365 = 0.0635$$

ובאופן דומה

$$P\left(W > 190\right) = 1 - \Phi\left(\frac{253}{71}\right) \approx 1 - 0.99982 = 0.00018 = 1.8 \cdot 10^{-4}$$

■

סעיף ב - מה הסטייה מהתוחלת בה נמצאים 90% מהגברים?

הוכחה: כלומר, נמצא את m המקיים

$$P\left(M \in [178.4 - m, 178.4 + m]\right) = 0.9$$

אבל מהחישובים בסעיף הקודם

$$= P\left(178.4 - m \le M \le 178.4 + m\right) = P\left(\cancel{178.4} - m \le 7.6Z_M + \cancel{178.4} \le \cancel{178.4} + m\right) =$$

$$= P\left(-m \le 7.6Z_M \le m\right) = P\left(-\frac{m}{7.6} \le Z_M \le \frac{m}{7.6}\right) = 2\Phi\left(\frac{m}{7.6}\right) - 1 = 0.9$$

ולכן

$$\Phi\left(\frac{m}{7.6}\right) = 0.95$$

ועכשיו לפי הטבלה המטופשת או חישוב דרך הנוסחה ההפוכה

$$\frac{m}{7.6} \approx 1.64485$$

ומכאן

$$m \approx 12.50086$$

כלומר, 90% מהגברים בגובה של לכל היותר $m \approx 12.5$ ס"מ מהממוצע (שהוא כאמור 178.4). ■

סעיף ג - אותו דבר עבור נשים.

הוכחה: כלומר, נמצא את m המקיים

$$P\left(M\in\left[164.7-m,164.7+m\right]\right)=0.9$$

אבל מהחישובים בסעיף הקודם

$$=P\left(164.7-m\le M\le164.7+m\right)=P\left(\cancel{164.7}-m\le7.1Z+\cancel{164.7}\le\cancel{164.7}+m\right)=$$

$$=P\left(-m\le7.1Z_{M}\le m\right)=P\left(-\frac{m}{7.1}\le Z_{M}\le\frac{m}{7.1}\right)=2\Phi\left(\frac{m}{7.1}\right)-1=0.9$$

ולכן

$$\Phi\left(\frac{m}{7.1}\right)=0.95$$

ועכשיו לפי הטבלה המטופשת או חישוב דרך הנוסחה ההפוכה

$$\frac{m}{7.1}\approx1.64485$$

ומכאן

$$m\approx11.678435$$

כלומר, 90% מהמהנשים בגובה של לכל היותר $m\approx11.678435$ ס"מ מהממוצע (שהוא כאמור 164.7).

אגב, נשים לב שלא משנה פה מה התוחלת והתוצאה תלויה רק בסטיית התקן. שזה הגיוני, כי הμ רק מזיזה לנו את הגרם של ההתפלגות ימינה ושמאלה אבל לא "מותחת" אותו. ■

סעיף ד - נתונה התכונה $X\sim N\left(\mu_{x},\sigma_{x}^{2}\right),Y\sim N\left(\mu_{y},\sigma_{y}^{2}\right)$ **ב"ת אז** $X-Y\sim X\sim N\left(\mu_{x}-\mu_{y},\sigma_{x}^{2}+\sigma_{y}^{2}\right)$. **מה ההסתברות שגובה אישה אקראית גדול מזה של גבר אקראי?**

הוכחה: כלומר, עבור $X=W,Y=M$, אנחנו רוצים למצוא $G=X-Y>0$. עכשיו, מהתכונה, אנחנו יודעים איך מתפלג G

$$G\sim N\left(164.7-178.4,7.1^{2}+7.6^{2}\right)$$

ובקירוב טוב

$$G\sim N\left(-13.7,10.4^{2}\right)$$

ועכשיו, אנחנו רוצים לראות מתי $G>0$

$$P\left(G>0\right)=1-P\left(G\le0\right)=1-\Phi\left(\frac{0-\left(-13.7\right)}{10.4}\right)\approx1-0.90612=0.09388$$

אפילו לאישה אקראית יש סיכוי טוב יותר להיות גבוהה יותר מגבר אקראי מאשר שאני אהיה גבוהה מגבר אקראי. ■

שאלה 3

10% מהאוכלוסיית התיישים הם לבנים.
אורך זקן תייש שחור $B \sim exp\left(0.125\right)$
אורך זקן תייש לבן $W \sim exp\left(0.25\right)$

סעיף א - אחוז התיישים שזקנם

1. ארוך מ4 ס"מ.

הוכחה: נתחיל בלבדוק את אחוז התיישים הלבנים שזקנם ארוך מ4 ס"מ.

$$P\left(W>4\right)=e^{-0.25\cdot 4}=\frac{1}{e}\approx 0.3678794412$$

ועכשיו לאחוז התיישים השחורים

$$P\left(B>4\right)=e^{-0.125\cdot 4}=\frac{1}{\sqrt{e}}\approx 0.60653065971$$

אבל אוכלוסיית התיישים הכוללת מורכבת מ10% לבנים והשאר שחורים. ולכן נעשה ממוצע משוקלל

$$0.1\cdot P\left(W>4\right)+0.9\cdot P\left(B>4\right)\approx 0.582665537859$$

פה אנחנו הנחנו הנחה כבדה שאוכלוסיית התיישים מורכבת אך ורק מתיישים שהם שחורים בלבד או לבנים בלבד. ואז מדובר בקבוצות משלימות (ובעצם יש פה הסתברות מותנית במסווה). ■

2. ארוך מ10 ס"מ

הוכחה: נחשב את אחוז התיישים עם זקן קצר מ10 ס"מ.
הפעם לא נסווה את ההסתברות המותנית (סתם בשביל הגיוון).
נסמן L מ"מ אורך זקן של תייש כלשהו (לבן או שחור). כמו כן, נסמן I_W מ"מ אינדיקטור האם התייש לבן.
לכן

$$P\left(L\leq 10\right)=P\left(L<10\mid I'_W\right)\cdot P\left(I'_W\right)+P\left(L<10\mid I_W\right)\cdot P\left(I_W\right)=0.9\cdot P\left(B<10\right)+0.1\cdot P\left(W<10\right)=$$

$$=0.9\cdot\int_0^{10}0.125e^{-0.125b}db+0.1\cdot\int_0^{10}0.25e^{-0.25w}dw=-\left(0.9\cdot\left.e^{-0.125b}\right|_{b=0}^{10}+0.1\cdot\left.e^{-0.25w}\right|_{w=0}^{10}\right)=$$

$$=-\left(0.9e^{-1.25}-0.9+0.1e^{-2.5}-0.1\right)=1-0.9e^{-1.25}-0.1e^{-2.5}\approx 0.733937182963439$$

■

סעיף ב - מה פונקציית ההתפלגות המצטברת של אורך זקן תיש שנבחר מקרית? מה פונקציית הצפיפות?

הוכחה: בדומה לחישובים מקודם

$$F_L\left(l\right)=P\left(L\leq l\right)=0.9\cdot P\left(B\leq l\right)+0.1\cdot P\left(W\leq l\right)=0.9\cdot\left(1-e^{-0.125l}\right)+0.1\cdot\left(1-e^{-0.25l}\right)=$$

$$=1-\left(0.9e^{-0.125l}+0.1e^{-0.25l}\right)$$

וכדי למצוא את הצפיפות, אנחנו יכולים פשוט לגזור

$$f_L\left(l\right)=0.1125\cdot e^{-0.125l}+0.025e^{-0.25l}$$

וכמובן שכל זה בתחום $l\geq 0$. ■

סעיף ג - נתון שלתייש זקן ארוך מ5 ס"מ. מה ההסתברות שהוא ארוך מ10 ס"מ?

הוכחה: נחשב ישירות

$$P\left(L>10\mid L>5\right)=\frac{P\left(L>10,L>5\right)}{P\left(L>5\right)}=\frac{P\left(L>10\right)}{P\left(L>5\right)}=\frac{1-F_{L}\left(10\right)}{1-F_{L}\left(5\right)}=\frac{0.9e^{-0.125\cdot10}+0.1e^{-0.25\cdot10}}{0.9e^{-0.125\cdot5}+0.1e^{-0.25\cdot5}}=$$

$$\approx0.521297487308419535$$

■

סעיף ד - האם התפלגות אורך הזקן של תיש מקיימת את תכונת חוסר הזיכרון?

הוכחה: ניתן לראות כי הדבר אינו מתקיים. כי אחרת היינו מצפים שהביטוי

$$P\left(L>5\right)=0.9e^{-0.125\cdot5}+0.1e^{-0.25\cdot5}\approx0.51038576535311$$

יהיה שווה לתוצאה הקודמת. וזה כמובן לא מתקיים (ראה את התוצאה הקודמת). ■

שאלה 4

סעיף א - כיצד מתפלג Y?

הוכחה: Y מתפלג גיאומטרית. שכן

$$P(Y=y)=P(y-1<X\leq y)=P(X>y-1,X\leq y)=P(X\leq y\mid X>y-1)\cdot P(X>y-1)=$$

השתמשנו בנוסחת בייס (היה אפשר גם לעשות חיסור הסתברויות ולהגיע כנראה לאותה תוצאה – $P(y-1<X\leq y)=P(X\leq y)$ $.(P(X<y-1)$

$$=(1-P(X>y\mid X>y-1))\cdot e^{-\lambda(y-1)}=(1-P(X>1))\cdot e^{-\lambda(y-1)}=$$

השתמשנו פה בתכונת חוסר זיכרון של מ"מ רציף מתפלג אקספוננציאלית.

$$=\left(1-e^{-\lambda}\right)\cdot e^{-\lambda(y-1)}=\left(1-\frac{1}{e^{\lambda}}\right)\cdot\left(\frac{1}{e^{\lambda}}\right)^{y-1}$$

כלומר, שY מתפלג גאומטרית עם פרמטר $p=1-e^{-\lambda}$. ■

סעיף ב - דני חישב כי התוחלת של Y היא 2. צ"ל את λ.

הוכחה: אבל לפי החישוב שלנו, התוחלת של Y מ"מ מתפלג גאומטרית היא

$$E(Y)=\frac{1}{1-e^{-\lambda}}=2\Longrightarrow e^{-\lambda}=\frac{1}{2}$$

ולכן

$$\lambda=-\ln 0.5\approx 0.69314718056$$

והתוחלת של X הינה

$$E(X)=\frac{1}{\lambda}\approx 1.4426950409$$

■

סעיף ג - עכשיו יש 2 קווי אוטובוס שמאחרים בהתפלגויות $exp(\lambda_1),exp(\lambda_2)$ בהתאמה (כמובן בלתי תלוי).

הוכחה: נסמן Z את כמות הזמן שדני מחכה בתחנה. נסמן X_i את האיחור של קו האוטובוס i (שמתפלג λ_i).
נגדיר קבוצת אינדקסים (סופית) A הקבוצה של האינדקסים של כל הקווים (במקרה שלנו רק 2 קווים).
משמעות הדבר שדני עדיין מחכה בתחנה אחרי z היא שכל האוטובוסים מאחרים לפחות z. לכן

$$P(Z>z)=P\left(\bigcap_{i\in A}X_i>z\right)=$$

עתה נעבור למקרה של 2 קווים כדי לשמור על פשטות הפתרון (ניתן להמשיך גם עם הקו של כמות סופית כלשהי של קווים אבל למה בעצם?).

$$=P(X_1>z\cap X_2>z)=P(X_1>z)\cdot P(X_2>z)=e^{-\lambda_1 z}\cdot e^{-\lambda_2 z}=e^{-(\lambda_1+\lambda_2)z}$$

כלומר, מצאנו את המבוקש

$$P(Z>z)=e^{-(\lambda_1+\lambda_2)z}$$

כלומר, שהזמן אותו דני מחכה, מתפלג אקספו עם פרמטר $\lambda_1+\lambda_2$. ■

שאלה 5

הוכחה: נתון $X\sim Uni\left(a,b\right)$ וגם $Y\sim exp\left(\lambda\right)$ ב"ת. מה הצפיפות של $Z=Y-X$?

X מתפלג יוניפורמית ולכן $-X$ מתפלג יוניפורמית $-X\sim Uni\left(-b,-a\right)$ (שיקוף סביב ה0). נגדיר $T=-X$. לכן $Z=Y+T$. מנוסחת הקונבולוציה:

$$f_Z\left(z\right)=f_{Y+T}\left(z\right)=\int_{-\infty}^{\infty}f_Y\left(y\right)f_T\left(z-y\right)dy=$$

אבל $z-y\in\left[-b,-a\right]$. כלומר $z+a\leq y$ וגם $y\leq z+b$ (כדי שהצפיפות היוני תהה שונה מ0).

כמו כן, הצפיפות האקספו תהיה שונה מ0 עבור $y\geq 0$. לכן $m=\max\left\{z+a,0\right\}\leq y\leq z+b$

$$=\int_m^{z+b}\lambda e^{-\lambda y}\cdot\frac{1}{-a-(-b)}dy=\frac{1}{a-b}\int_m^{z+b}-\lambda e^{-\lambda y}\cdot dy=\frac{1}{a-b}\left(e^{-\lambda(z+b)}-e^{-\lambda m}\right)$$

אז נסכם

$$f_Z\left(z\right)=\begin{cases}\frac{1}{a-b}\left(e^{-\lambda(z+b)}-e^{-\lambda m}\right) & z\geq -b\\ 0 & ow\end{cases}$$

כאשר $m=\max\left\{z+a,0\right\}$. ■

שאלה 6

הוכחה: יהיו $X_i \sim exp\left(\lambda_i\right)$ ב"ת כאשר $i \in \{1,2,3\}$. צ"ל $P\left(X_1 < X_2 < X_3\right)$.
המשתנים ב"ת ולכן הצפיפות המשותפת היא

$$f_{X_1,X_2,X_3}\left(x,y,z\right) = f_{X_1}\left(x\right) \cdot f_{X_2}\left(y\right) \cdot f_{X_3}\left(z\right) = \begin{cases} \lambda_1\lambda_2\lambda_3 e^{-\lambda_1 x} e^{-\lambda_2 y} e^{-\lambda_3 z} & x,y,z \geq 0 \\ 0 & ow \end{cases}$$

ולכן

$$P\left(X_1 < X_2 < X_3\right) = \int_0^\infty \int_x^\infty \int_y^\infty \lambda_1\lambda_2\lambda_3 e^{-\left(\lambda_1 x + \lambda_2 y + \lambda_3 z\right)} dzdydx =$$

נוציא קבועים ביחס לכל אינטגרל

$$= \int_0^\infty \lambda_1 e^{-\lambda_1 x} \int_x^\infty \lambda_2 e^{-\lambda_2 y} \int_y^\infty \lambda_3 e^{-\lambda_3 z} dzdydx = \int_0^\infty \lambda_1 e^{-\lambda_1 x} \int_x^\infty \lambda_2 e^{-\lambda_2 y} \left(e^{-\lambda_3 y}\right) dydx =$$

ועכשיו בעזרת התבנית הידועה של אינטגרל של אקספוננט המון פעמים עד שנקבל תשובה

$$= \int_0^\infty \lambda_1 e^{-\lambda_1 x} \frac{\lambda_2}{\lambda_2 + \lambda_3} \int_x^\infty \left(\lambda_2 + \lambda_3\right) e^{-\left(\lambda_2+\lambda_3\right) y} dydx =$$

$$= \int_0^\infty \lambda_1 e^{-\lambda_1 x} \frac{\lambda_2}{\lambda_2 + \lambda_3} e^{-\left(\lambda_2+\lambda_3\right) x} dx =$$

$$= \frac{\lambda_1\lambda_2}{\lambda_2 + \lambda_3} \int_0^\infty e^{-\left(\lambda_1+\lambda_2+\lambda_3\right) x} dx =$$

$$= \frac{\lambda_1\lambda_2}{\left(\lambda_2 + \lambda_3\right) \cdot \left(\lambda_1 + \lambda_2 + \lambda_3\right)} \int_0^\infty \left(\lambda_1 + \lambda_2 + \lambda_3\right) e^{-\left(\lambda_1+\lambda_2+\lambda_3\right) x} dx =$$

$$= \frac{\lambda_1\lambda_2}{\left(\lambda_2 + \lambda_3\right) \cdot \left(\lambda_1 + \lambda_2 + \lambda_3\right)} \cdot e^{-\left(\lambda_1+\lambda_2+\lambda_3\right) \cdot 0} =$$

$$= \frac{\lambda_1\lambda_2}{\left(\lambda_2 + \lambda_3\right) \cdot \left(\lambda_1 + \lambda_2 + \lambda_3\right)}$$

■