הסתברות מ - 094412 - HW7

גור תלם - 206631848

29.12.2020

שאלה 1

בכד n כדורים ממוספרים מ1 עד n.
מוציאים באקראי 2 ללא החזרה.
X הערך הימינימלי מבין שתי ההוצאות.
Y הערך המקסימלי.

סעיף א - חשב את $E\left(X \mid Y=y\right)$ ואת $E\left(Y \mid X=x\right)$.

הוכחה: נתחיל עם $E\left(X \mid Y=y\right)$. כלומר, צריך לחשב את התוחלת של X בהינתן ערך מקסימלי כלשהו.
לכן

$$E\left(X \mid Y=y\right)=\sum_{x=1}^{y-1} x \cdot P\left(X=x \mid Y=y\right)=\sum_{x=1}^{y-1} x \cdot \frac{P\left(X=x, Y=y\right)}{P\left(Y=y\right)}=\sum_{x=1}^{y-1} x \cdot \frac{P\left(\{first=x, second=y\} \cup \{first=y, second=x\}\right.}{P\left(Y \leq y\right)-P\left(Y \leq y-1\right)}$$

כיוון ש$X \neq Y$ תמיד (כי מדובר בשליפה ללא החזרה) אז אנחנו נקבל ש:

$$=\sum_{x=1}^{y-1} x \cdot \frac{P\left(\{first=x, second=y\}\right)+P\left(\{first=y, second=x\}\right)}{\frac{y}{n} \cdot \frac{y-1}{n-1}-\frac{y-1}{n} \cdot \frac{y-2}{n-1}}=\sum_{x=1}^{y-1} x \cdot \frac{2 \cdot \cancel{\frac{1}{n}} \cdot \cancel{\frac{1}{n-1}}}{\frac{y}{\cancel{n}} \cdot \frac{y-1}{\cancel{n-1}}-\frac{y-1}{\cancel{n}} \cdot \frac{y-2}{\cancel{n-1}}}=\sum_{x=1}^{y-1} x \cdot \frac{2}{y \cdot (y-1)-(y-2) \cdot (y-}$$

$$=\sum_{x=1}^{y-1} x \cdot \frac{\cancel{2}}{\left(\cancel{y-y}+\cancel{2}\right) \cdot (y-1)}=\frac{1}{y-1} \cdot \sum_{x=1}^{y-1} x=\frac{(y\cancel{-1+1})\cancel{(y-1)}}{2\cancel{(y-1)}}=\frac{y}{2}$$

נעבוד ל$E\left(Y \mid X=x\right)$. כלומר לחשב את התוחלת של Y בהינתן שידוע מה הערך המינימלי.
לכן

$$E\left(Y \mid X=x\right)=\sum_{y=x+1}^{n} y \cdot P\left(Y=y \mid X=x\right)=\sum_{y=x+1}^{n} y \cdot \frac{P\left(Y=y, X=x\right)}{P\left(X=x\right)}=\sum_{y=x+1}^{n} y \cdot \frac{2 \cdot \frac{1}{n} \cdot \frac{1}{n-1}}{\frac{n-x+1}{n} \cdot \frac{n-x}{n-1}-\frac{n-x}{n} \cdot \frac{n-x-1}{n-1}}=\frac{1}{n-x} \cdot \frac{(n+x+}{}$$

$$=\frac{1}{\cancel{n-x}} \cdot \frac{\cancel{(n-x)} \cdot (n+x+1)}{2}=\frac{n+x+1}{2}$$

■

סעיף ב - מוציאים עם החזרות. צריך שוב לחשב את התוחלות.

הוכחה: נחשב הסתברות תחילה עבור המקרה ש$x \neq y$:

$$P\left(X=x \mid Y=y\right)=\frac{P\left(X=x, Y=y\right)}{P\left(Y=y\right)}=\frac{2 \frac{1}{\cancel{n^2}}}{\frac{y}{\cancel{n}} \cdot \frac{y}{\cancel{n}}-\frac{y-1}{\cancel{n}} \cdot \frac{y-1}{\cancel{n}}}=\frac{2}{y^2-(y-1)^2}=$$

$$=\frac{2}{2y-1}$$

התווסף לנו מקרה נוסף, המקרה בו $x=y$ שלא היה אפשרי בסעיף א. נחשב את ההסתברות במקרה זה:

$$P\left(X=y \mid Y=y\right)=\frac{P\left(X=y, Y=y\right)}{P\left(Y=y\right)}=\frac{2 \frac{1}{\cancel{n^2}}}{\frac{y}{\cancel{n}} \cdot \frac{y}{\cancel{n}}-\frac{y-1}{\cancel{n}} \cdot \frac{y-1}{\cancel{n}}}=\frac{1}{2y-1}$$

ולכן התוחלת

$$E\left(X \mid Y=y\right)=y \cdot \frac{1}{2y-1}+\sum_{y=1}^{y-1} x \cdot \frac{2}{2y-1}=\frac{1}{2y-1} \cdot \left(y+2 \cdot \frac{(1+y-1) \cdot (y-1)}{2}\right)=\frac{y^2}{2y-1}$$

ובצורה אנלוגית נחשב $P\left(Y=y\mid X=x\right)$ (עבור $x\neq y$):

$$P\left(Y=y\mid X=x\right)=\frac{P\left(Y=y,X=x\right)}{P\left(X=x\right)}=\frac{2\cancel{\frac{1}{n^2}}}{\frac{(n-x+1)^2-(n-x)^2}{\cancel{n^2}}}=\frac{2}{2n-2x+1}$$

ועבור $x=y$

$$P\left(Y=x\mid X=x\right)=\frac{1}{2n-2x+1}$$

ועכשיו לתוחלת

$$E\left(Y\mid X=x\right)=x\cdot\frac{1}{2n-2x+1}+\frac{2}{2n-2x+1}\cdot\sum_{y=x+1}^{n}y=x\cdot\frac{1}{2n-2x+1}+\frac{\cancel{2}}{2n-2x+1}\cdot\frac{(n+x+1)\cdot(n-x)}{\cancel{2}}=\frac{n^2-x^2+n}{2n-2x+1}$$

■

שאלה 2

מבצעים ניסוי עם 2 שלבים. מטילים קוביה הוגנת עד שמתקבלת הספרה 6.
N מספר ההטלות בשלב הראשון.
בשלב השני מטילים N קוביות.

סעיף א - X סכום הספרות שהתקבלו בכל ההטלות בשלב השני של הניסוי. חשב $E\left(X \mid N=n\right)$ ואת $E\left(X\right)$.

הוכחה: כלומר, אם ידוע שבשלב הראשון הוטלו N קוביות, אז בשלב השני הוטלו עוד N קוביות. בשלב הראשון אנחנו מטילים בהתפלגות גיאומטרית כאשר הצלחה היא קבלה של המספר 6. כלומר, התפלגות גיאומטרית עם פרמטר $\frac{1}{6}$. ולכן התוחלת של כמות ההטלות היא $E\left(N\right)=6$.
נתחיל מחישוב $E\left(X \mid N=n\right)$.
נחשב את התוחלת של הספרה של קוביה כלשהי. נסמן D_iהספרה על הקוביה הi

$$E\left(D_i\right)=3.5$$

(כבר חישבנו והוכחנו זאת בהרצאה באחת הדוגמות).
ידוע כי

$$X=\sum_{i=1}^{N} D_i=3.5N$$

ולכן, לפי לינאריות של תוחלת

$$E\left(X \mid N=n\right)=\sum_{i=1}^{n} E\left(D_i\right)=3.5n$$

ומנוסחת ההחלקה, אנחנו נקבל

$$E\left(X\right)=E\left(E\left(X \mid N\right)\right)=E\left(3.5N\right)=3.5\cdot 6=21$$

■

סעיף ב - Y_i מ"מ מספר הפעמים שיצאה i בשלב השני. צ"ל $E\left(Y_i \mid N=n\right)$ ואת $E\left(Y_i\right)$.

הוכחה: כיוון שהקוביה הוגנתת התוחלת וכיוון שכבר אמרנו שתוחלת כמות ההטלות בשלב השני הוא 6 שכן הכמות שווה לכמות ההטלות בשלב הראשון. נסמן I_k אינדיקטור שבהטלה הk בשלב השני יצא i.
לכן התוחלת של

$$Y_i=\sum_{k=1}^{N} I_k$$

לכן מלינאריות התוחלת נקבל

$$E\left(Y_i \mid N=n\right)=\sum_{k=1}^{n} E\left(I_k\right)=\frac{1}{6}n$$

מהוגנות הקוביה. ומנוסחת ההחלקה, אנחנו נקבל

$$E\left(Y_i\right)=E\left(E\left(Y_i \mid N\right)\right)=E\left(\frac{1}{6}N\right)=\frac{1}{6}\cdot E\left(N\right)=1$$

■

סעיף ג - בשלב השני של הניסוי, מקבלים נק' עבור הטלה עם תוצאה זוגית ו2 נקודות עבור הטלה עם תוצאה א"ז. מה התוחלת של סכום הנקודות?

הוכחה: הקוביה הוגנת ויש את המספרים 1 עד 6 כולל, כלומר, בדיוק חצי מהם זוגיים וחצי א"ז. כלומר, השאלה שקולה להטלת מטבע H=זוגי וT=א"ז. ושואלים מה התוחלת של $2T+H$ כאשר H הינו כמות הפלי וT הינו כמות העץ.
מתכונת הלינאריות של תוחלת

$$E\left(2T+H\right)=2\cdot E\left(T\right)+E\left(H\right)$$

אבל ההתפלגויות הן בינומיות עם פרמטר $\frac{1}{2}$ ועם N הטלות. לפי אי תלות ונוסחת ההחלקה, נקבל

$$E\left(2T+H\right)=E\left(2\cdot E\left(T \mid N\right)+E\left(H \mid N\right)\right)=E\left(2\cdot\frac{1}{2}\cdot N+\frac{1}{2}\cdot N\right)=2\cdot\frac{1}{2}\cdot 6+\frac{1}{2}\cdot 6=6+3=9$$

■

שאלה 3

בוחרים ספרות דצימליות בצורה בלתי תלויה אחת אחרי השניה. מפסיקים אחרי שכל ספרה מופיע לפחות פעם אחת.

סעיף א - מה תוחלת אורך הסדרה?

הוכחה: נסמן מ"מ X אורך הסדרה.
נסמן Y_i מ"מ המיצג את מספר ההגרלות אחרי שכבר יצאו i ספרות שונות, עד לקבלת ספרה חדשה. לכן

$$P\left(Y_i = y\right) = \left(\frac{i}{10}\right)^y \cdot \frac{10-i}{10}$$

כאשר $9 \geq i \geq 1$.
לכן, התוחלת של X הינה סכום התוחלות של Y_i ועוד 1 (עבור הספרה הראשונה)

$$E\left(X\right) = 1 + \sum_{i=1}^{9} E\left(Y_i\right) =$$

* השתמשנו פה בלינאריות של תוחלת.
לפי התוחלת של התפלגות גאומטרית

$$= 1 + \sum_{i=1}^{9} \frac{10}{10-i} = \frac{7381}{252} \approx 29.2897$$

■

עבור כל ספרה d שהתקבלה בסדרה, מטילים סדרת הטלות מטבע עם סיכוי $\frac{1}{d+1}$ לקבל H. כאשר מתקבל H, עוצרים וממשיכים לסדרה הבאה.
כל ההטלות ב"ת.

סעיף ב - מה תוחלת מספר הטלות המטבע בסה"כ בסדרה?

הוכחה: d_i היא הספרה הi שהתקבלה בסדרת הספרות. נסמן מ"מים Z_i כמות ההטלות עבור הספרה הd_i שהתקבלה. מההגדרה קל לראות בנקל כי מדובר בהתפלגות גיאומטרית עבור Z_i עם פרמטר $\frac{1}{d_i+1}$. לכן עבור הספרה d_i התוחלת היא $d_i + 1$.
לכן, אם ידוע לנו אורך הסדרה, התוחלת של כמות הטלות המטבע (נסמן מ"מ $T = Total\,tosses$)

$$E\left(T \mid X = x\right) = \sum_{i=1}^{x} E\left(Z_i\right)$$

עכשיו נשתמש בנוסחת ההחלקה

$$E\left(Z_i\right) = E\left(E\left(Z_i \mid d_i\right)\right) = E\left(d_i + 1\right) = E\left(d_i\right) + 1 = 5.5$$

ולכן נוכל לקבל

$$E\left(T \mid X = x\right) = \sum_{i=1}^{x} E\left(Z_i\right) = \sum_{i=1}^{x} 5.5 = 5.5x$$

ושוב נוסחת ההחלקה

$$E\left(T\right) = E\left(E\left(T \mid X = x\right)\right) = E\left(5.5x\right) = 5.5 \cdot \frac{7381}{252} \approx 161.093$$

■

שאלה 4

סעיף א - נתונים X,Y מ"מ ב"ת ושווי התפלגות. יהי $S=X+Y$ סכומם. הוכיחו ש $E\left(X\mid S\right)=\frac{S}{2}$.

הוכחה: נראה כי $X\mid S=s,Y\mid S=s$ בעלי התפלגות זהה.

$$P\left(X=x\mid S=s\right)=\frac{P\left(X=x,S=s\right)}{P\left(S=s\right)}=\frac{P\left(X=x,Y=s-x\right)}{P\left(S=s\right)}=\frac{P\left(Y=s-x\right)\cdot P\left(X=x\right)}{P\left(S=s\right)}=$$

המעבר האחרון נובע מאי תלות. עכשיו נשתמש בכך שהמשתנים שווי התפלגות.

$$=\frac{P\left(Y=x\right)\cdot P\left(X=s-x\right)}{P\left(S=s\right)}=\frac{P\left(Y=x,S=s\right)}{P\left(S=s\right)}=P\left(Y=x\mid S=s\right)$$

וזה כמובן נכון לכל x,s.

אבל גם

$$E\left(X\mid S=s\right)=\sum_{x\in R}x\cdot P\left(X=x\mid S=s\right)=\sum_{x\in R}x\cdot P\left(Y=x\mid S=s\right)=E\left(Y\mid S=s\right)$$

כלומר שהתוחלות שוות.

כמו כן,

$$E\left(X+Y\mid S=s\right)=s$$

שכן $X+Y=S$.

ולכן מלניאריות התוחלת

$$E\left(X+Y\mid S=s\right)=E\left(X\mid S=s\right)+E\left(Y\mid S=s\right)=2E\left(X\mid S=s\right)=s$$

ולכן קיבלנו את אשר רצינו להוכיח

$$E\left(X\mid S=s\right)=\frac{s}{2}$$

■

סעיף ב - U,V מ"מ. $E\left(U\mid V\right)=V+2$ וכמו כן נתון $V\sim Pois\left(1\right)$. צ"ל $E\left(\frac{U}{V+1}\right)$.

הוכחה: נשתמש בנוסחת ההחלקה

$$E\left(U\right)=E\left(E\left(U\mid V\right)\right)=E\left(V+2\right)=E\left(V\right)+2=3$$

בגלל התפלגות פואסונית.

ולכן

$$E\left(\frac{U}{V+1}\right)=E\left(E\left(\frac{U}{V+1}\mid V\right)\right)=$$

אם נגדיר $h(V)=\frac{1}{V+1}$ אז=

$$=E\left(E\left(U\cdot h\left(V\right)\mid V\right)\right)=E\left(\frac{1}{V+1}\cdot E\left(U\mid V\right)\right)=E\left(\frac{V+2}{V+1}\right)=E\left(1+\frac{1}{V+1}\right)$$

לפי הנתון שמדובר בהתפלגות פואסונית וביחד עם תוחלת של טרנספורמציה

$$E\left(\frac{1}{V+1}\right)=\sum_{i=0}^{\infty}\frac{1}{k+1}\cdot P\left(V\right)=\sum_{i=0}^{\infty}\frac{1}{i+1}\cdot e^{-1}\cdot\frac{1}{i!}=\frac{1}{e}\cdot\sum_{i=0}^{\infty}\frac{1}{(i+1)!}=\frac{1}{e}\cdot\left(e-1\right)=1-\frac{1}{e}$$

ולכן נקבל לבסוף מלינאריות

$$E\left(1+\frac{1}{V+1}\right)=E\left(\frac{1}{V+1}\right)+1=2-\frac{1}{e}$$

■

סעיף ג - נתון מ"מ $Z \sim Bin\left(n,p\right)$ ופונקציה $f\left(z\right)=\frac{z}{p^{z}\left(1-p\right)^{n-z}}$. צ"ל ש$E\left(f\left(Z\right)\right)=n\cdot2^{n-1}$.

הוכחה: לפי נוסחה של תוחלת של טרנספורמציה

$$E\left(f\left(Z\right)\right)=\sum_{z=0}^{n}f\left(z\right)\cdot P\left(Z=z\right)=\sum_{z=0}^{n}f\left(z\right)\cdot\binom{n}{z}\cdot p^{z}\left(1-p\right)^{n-z}=\sum_{z=0}^{n}z\cdot\binom{n}{z}=2^{n}\cdot\sum_{z=0}^{n}z\cdot\binom{n}{z}\cdot0.5^{n}=2^{n}\cdot n\cdot\frac{1}{2}$$

המעבר האחרון נובע מתוחלת של התפלגות בינומית והגדרת התוחלת הישירה. ולכן קיבלנו

$$E\left(f\left(Z\right)\right)=n\cdot2^{n-1}$$

■

שאלה 5

הוכחה: n זוגות משובצים ל־n חדשים. חולקו $2n$ כרטיסים לחדרים באופן אקראי (טעויות קורות), ללא קשר לחדרים בשיבוץ.

X מ"מ שהוא מספר האורחים שקיבלו כרטיס הפותח את דלת חדרם.

נחשב $E\left(X\right)$ ואת $E\left(X^2\right)$.

מ"מ אינדיקטור I_i משמעו שהאורח ה־i קיבל מפתח שפותח את חדרו (כאשר $1 \leq i \leq 2n$). לכן

$$X = \sum_{i=1}^{2n} I_i$$

והתוחלת ניתנת לחישוב

$$E\left(X\right) = \sum_{i=1}^{2n} 1 \cdot P\left(I_i = 1\right) = \sum_{i=1}^{2n} \frac{2}{2n} = \sum_{i=1}^{2n} \frac{1}{n} = 2$$

ועתה נחשב $E\left(X^2\right)$.

$$E\left(X^2\right) = E\left(\sum_{i=1}^{2n} I_i \cdot \sum_{i=1}^{2n} I_i\right) = E\left(\sum_{i=1}^{2n}\sum_{j=1}^{2n} I_j \cdot I_i\right) =$$

המעבר האחרון כיוון שמדובר באינדיקטורים שמקבלים 0 או 1. ועכשיו מלינאריות של תוחלת

$$= \sum_{i=1}^{2n}\sum_{j=1}^{2n} E\left(I_j \cdot I_i\right) =$$

עכשיו ניתן להפריד למקרים

$$= \sum_{i=1}^{2n} E\left(I_i^2\right) + \sum_{1 \leq i \neq j \leq 2n} E\left(I_i \cdot I_j\right) = 2 + \sum_{1 \leq i \neq j \leq 2n} P\left(I_i = 1, I_j = 1\right) =$$

לפי תוחלת של אינדיקטורים. ועכשיו שוב נפצל למקרים (כאשר i, j באותו חדר או בחדרים שונים)

$$= 2 + \sum_{\substack{1 \leq i \neq j \leq 2n \\ j, i\, same\, room}} P\left(I_i = 1, I_j = 1\right) + \sum_{\substack{1 \leq i \neq j \leq 2n \\ j, i\, different\, rooms}} P\left(I_i = 1, I_j = 1\right) =$$

$$= 2 + \sum_{\substack{1 \leq i \neq j \leq 2n \\ j, i\, same\, room}} \frac{2}{2n} \cdot \frac{1}{2n-1} + \sum_{\substack{1 \leq i \neq j \leq 2n \\ j, i\, different\, rooms}} \frac{2}{2n} \cdot \frac{2}{2n-1} =$$

עכשיו נותר לספור כמה צמדים i, j יש שבחדרים שונים ושבאותו חדר. צמדים שבאותו חדר יש כמובן n לפי הגדרת השאלה. צמדים שלא באותו חדר יש $\left(2n\right)^2 - 2n - 2n$ (האפשרויות לבחור אנשים עם חזרות פחות הזוגות העצמיים פחות הזוגות שבאותו חדר). ולכן ניתן לקבל

$$= 2 + 2n \cdot \frac{2}{2n} \cdot \frac{1}{2n-1} + \left(\left(2n\right)^2 - 4n\right) \cdot \frac{2}{2n} \cdot \frac{2}{2n-1} = 2 + \frac{2}{2n-1} + \left(4n-4\right) \cdot \frac{2}{2n-1} =$$

$$= 2 + \left(4n-3\right) \cdot \frac{2}{2n-1} = \frac{2\left(2n-1\right) + 2 \cdot \left(4n-3\right)}{2n-1} = \frac{4n-2+8n-6}{2n-1} = \frac{12n-6-2}{2n-1} = \frac{2}{2n-1} + 6$$

ואפילו $wolfram\, alpha$ מאשר את התשובה. ■

שאלה 6

40% קונים של חוברות חדו"א (סוג 1)
25% קונים של חוברות אלגברה (סוג 2)
20% קונים של חוברות הסתברות (סוג 3)
15% קונים של חוברות אחרות (סוג 4)
לקוח מסוג i קונה $Y_i \sim Geo\left(\frac{1}{2^i}\right)$ חוברות ללא תלות בלקוחות האחרים. הלקוחות שומרים על התו הסגור ונכנסים לחנות אחד אחד.

סעיף א - מ"מ X_k ומ"מ T_k הוא מספר החוברות וסוג של הלקוח הk, בהתאמה. צ"ל $E\left(X_k \mid T_k = i\right)$ ואת $E\left(X_k\right)$.

הוכחה: כלומר, מה כמות החוברות שקנה לקוח k בהינתן הסוג שלו וכמו כן, אם לא נתון הסוג שלו.
לפי תוחלת של התפלגות גיאומטרית

$$E\left(X_k \mid T_k = i\right) = 2^i$$

ועכשיו לפי נוסחת ההחלקה והגדרת תוחלת ישירה

$$E\left(X_k\right) = E\left(E\left(X_k \mid T_k = i\right)\right) = \sum_{i=1}^{4} 2^i \cdot P\left(T_k = i\right) = 2 \cdot 0.4 + 4 \cdot 0.25 + 8 \cdot 0.2 + 16 \cdot 0.15 = 5.8$$

■

סעיף ב - מה תוחלת מספר החוברות שנמכרו לשלושת הלקוחות הראשונים?

הוכחה: כלומר, מה $X_1 + X_2 + X_3$. אבל גילינו שהתוחלת של כל לקוח, ללא קשר לאיזה מספר לקוח הוא, היא מספר קבוע (שזה הגיוני כי אין תלות בין הלקוחות הריי)

$$E\left(X_2 + X_2 + X_3\right) = E\left(X_1\right) + E\left(X_2\right) + E\left(X_3\right) = 3 \cdot 5.8 = 17.4$$

■

סעיף ג - מה התוחלת של מספר חוברות האלגברה שנמכרו עד שנכנס בפעם הראשונה לקוח מסוג 3 לחנות?

הוכחה: נשים לב כי מעניין אותנו רק לקוחות שקונים חוברות אלגברה (סוג 2) ולקוחות מסוג 3. כלומר, ההסתברות היחסית לכניסה של לקוח מסוג 2 אם ידוע שנכנסים רק לקוחות מסוג 2 או 3 היא

$$P\left(T_k = 2 \mid T_k = 2, 3\right) = \frac{5}{9}$$

כלומר מדובר בהתפלגות גאומטרית כאשר כניסה של לקוח מסוג 3 נחשבת להצלחה. כלומר פרמטר $\frac{4}{9}$.
נסמן C_2 כמות הלקוחות מסוג 2 עד שנכנס לקוח מסוג 3. התוחלת של C_2 היא $E\left(C_2\right) = \frac{9}{4} - 1 = \frac{5}{4}$ (התפלגות גיאומטרית של כמות הלקוחות מסוג 2 עד לקוח מסוג 3 כולל פחות הלקוח האחרון שאינו מסוג 2).
כמו כן, תוחלת כמות החוברות שלקוח מסוג 2 קונה הינה $E\left(Y_2\right) = 2^2 = 4$. כיוון ש$C_2$ וY_2 ב"ת (שכן כמות החוברות לא תלויה בכמות הלקוחות מהנתון), אז נקבל ש

$$E\left(C_2 \cdot Y_2\right) = E\left(C_2\right) \cdot E\left(Y_2\right) = 5$$

■