14.12.14

מבוא להסתברות - שיעור שמיני

סעיף 2 מהשעורים בא"ל.

איך מתפלג מספר הפונים לכספומט ביממה?
סכום פואסונים זה עוד פואסוני, ולכן מספר
הפונים ביממה מתפלג פואסוני עם $P(24\cdot 2)$
שזה $P(48)$.

ג. כל אחד מהם אמין בסיכוי $\frac{2}{3}$. לכן מספר
המושכים אמין מתפלג פואסונית עם $P(\frac{2}{3}\cdot 48)$
שזה $P(32)$.

תוחלת:

ממוצע על פי ההתפלגות.

הגדרה:

את התוחלת של משתנה X נסמן ב $E(X)$

ומוגדר: $E(X)=\sum_{X=x} P(X=x)\cdot \underbrace{x}_{\text{ערך } x}$.

$$\boxed{E(X)=\sum_{X=x} P(X=x)\cdot x}$$

דוגמא:

$$P(X=1)=P(X=2)=P(X=5)=\frac{1}{3}$$

אז $E(X)=\frac{1}{3}\cdot 1+\frac{1}{3}\cdot 2+\frac{1}{3}\cdot 5$

דוגמא נוספת:

$$P(X=35)=\frac{1}{8},\ P(X=20)=\frac{1}{4},\ P(X=3)=\frac{1}{4},\ P(X=5)=\frac{3}{8}$$

$$E(x)=\frac{1}{4}\cdot 3+\frac{3}{8}\cdot 5+\frac{1}{4}\cdot 20+\frac{1}{8}\cdot 35$$

הערה:

התוחלת זה מה שהממוצע מתקרב אליו אם עושים הרבה פעמים (זה לא תמיד קורה, אבל הרבה פעמים כן).

דוגמא:

X - הטלת קוביה

$$E(X)=\frac{1}{6}\cdot 1+\frac{1}{6}\cdot 2+\frac{1}{6}\cdot 3+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 6=$$

$$=\frac{1}{6}(1+2+3+4+5+6)=\frac{1}{6}\cdot\frac{(1+6)6}{2}=3.5$$

תוחלת של משתנה אחיד $U[a,b]$:

$$E(X)=\sum_{k=a}^{b}P(X=k)\cdot k=\sum_{k=a}^{b}\frac{1}{b-a+1}\cdot k=$$

$$=\frac{1}{b-a+1}\sum_{k=a}^{b}k=\frac{1}{b-a+1}\cdot\frac{(a+b)(b-a+1)}{2}=$$

$$=\frac{a+b}{2}$$

שאלה:

X ~~מתפלג~~ משתנה אינדיקטורי עם סיכוי p להצלחה.

אז $X\sim B(p)$. מה התוחלת של X?

$$E(X)=P(X=0)\cdot 0+P(X=1)\cdot 1=(1-p)\cdot 0+p\cdot 1=p$$

תוחלת של משתנה $X\sim P(\lambda)$?

$$E(X)=\sum_{k=0}^{\infty}P(X=k)\cdot k=\sum_{k=0}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{k!}\cdot k$$

$$=\sum_{k=1}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{k!}\cdot k=\sum_{k=1}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{(k-1)!}=$$

$$=\lambda\cdot\sum_{k=1}^{\infty}e^{-\lambda}\cdot\frac{\lambda^{k-1}}{(k-1)!}=\lambda\cdot\sum_{m=0}^{\infty}e^{-\lambda}\cdot\frac{\lambda^m}{m!}=\lambda\cdot 1=\lambda$$

מה שבתוך ה-Σ זה סכום ההסתברויות של משתנה $P(\lambda)$ ולכן זה מסתכם ב-1.

סכום ההסתברויות של משתנה מסתכם ב-1.

שאלה: (תוחלת של סכום משתנים).

X ו Y שני משתנים מקריים

$$P(X=1)=P(X=3)=P(X=5)=\frac{1}{3}$$

$\cancel{P(X=}$

$$P(Y=2)=P(Y=4)=\frac{1}{2}$$

מהו $E(X+Y)$?

פתרון:

מתקיים תמיד $E(X+Y)=E(X)+E(Y)$

ולכן במקרה זה

$$E(X+Y)=E(X)+E(Y)=\frac{1}{3}(1+3+5)+\frac{1}{2}(2+4)=$$

$$=3+3=6$$

צריך לשים לב שלא נדרשנו לדעת התפלגות משותפת.
לא נדרשנו גם להסתמכות על אי-תלות. הטענה נכונה
תמיד בין אם המשתנים תלויים או בלתי תלויים.

נוכיח שעבור שני משתנים Y, X מתקיים תמיד:

$$E(X+Y)=E(X)+E(Y)$$

$$E(X+Y)=\sum_{w\in\Omega}P(w)\cdot(X+Y)(w)=$$

$$=\left(\sum_{w\in\Omega}P(w)\cdot X(w)\right)+\left(\sum_{w\in\Omega}P(w)\cdot Y(w)\right)$$

$$=E(X)+E(Y)$$

לא הנחנו כלל שום אי-תלות!

הוכחה:

$X\sim Bin(n,p)$ מהי תוחלתו?

מהו $E(X)$?

X הוא סכום של n משתנים אינדיקטורים

עם פרמטר p. לכל אחד מהם יש תוחלת p

לכן תוחלת הסכום זו $E(X)$ היא $n\cdot p$.

$X\sim Bin(n,p)$ אזי $X=\sum_{i=1}^{n} X_i$

כאשר X_i משתנים אינדיקטורים עם פרמטר p.

$$E(X)=\sum_{i=1}^{n} E(X_i)=n\cdot p$$

כאשר $E(X_i)=p$ לכל $1\le i\le n$

אינדיקטור עם הסתברות p זה משתנה שמקבל את הערך 1 (הצלחה) בסיכוי p ואת הערך 0 (כישלון) בסיכוי $1-p$.

חישוב תוחלת של משתנה $X\sim HG(n,a,b)$

X הוא סכום של n נסיונות שלכל אחד מהם יש סיכוי $\frac{a}{a+b}$ $\quad X=\sum_{i=1}^{n} X_i$

$E(X_i)=\frac{a}{a+b}$ לכן $E(X)=\frac{n\cdot a}{a+b}$

אותו דבר כי $EX=\sum E(X_i)$ למרות ש-X_i תלויים.

X_i זה תוצאת ההוצאה ה-i.

הסבר השני זה הוצאה של כדור אחרון.

$$P(X_2=1)=P(X_1=0)\cdot\frac{a}{a+b-1}+P(X_1=1)\cdot\frac{a-1}{a+b-1}=$$

$$=\frac{b}{a+b}\cdot\frac{a}{a+b-1}+\frac{a}{a+b}\cdot\frac{a-1}{a+b-1}=\frac{1}{(a+b-1)(a+b)}(ab+a(a-1))$$

$$=\frac{1}{(a+b-1)(a+b)}(ab+a^2-a)=\frac{a(a+b-1)}{(a+b-1)(a+b)}=\frac{a}{a+b}$$

והסיכוי הצלחה לא השתנה.

בעיית המזכירה הרשלנית:

יש n אנשים ו-n מכתבים שכל אחד מהם מיועד בדיוק לאחד מהאנשים. מזכירה שולחת את המכתבים באקראי. יהי X מספר האנשים שמקבלים מכתב נכון. מהו $E(X)$?

$X=\sum_{i=1}^{n} X_i$ הינו אינדיקטור לכך שהאדם ה-i קיבל מכתב נכון.

$$E(X_1)=\frac{1}{n}$$

$$E(X_i)=\frac{1}{n}$$

$$E(X)=\sum E(X_i)=n\cdot\frac{1}{n}=1$$

זה אומר שבממוצע רק אדם אחד יקבל את המכתב הנכון שלו.

שאלה:

נציג לו 2 מכתבים ולשאר האחרים נציג מכתב אחד בדיוק. יש 9 אנשים ו-10 מכתבים. מה תוחלת מספר המכתבים שיגיעו לכתובת הנכונה?

X - מספרם

$$X=x_1+x_2+\dots+x_{10}$$

כאשר x_1 ו-x_2 הם אינדיקטורים למכתבים שמיועדים לו.

(*) הסיכוי של אינדיקטור הוא בעצם התוחלת שלו

$$E(X_3)=E(X_4)=\dots=E(X_{10})=\frac{1}{10}$$

$$E(X_1)=E(X_2)=\frac{2}{10}$$

$$E(X)=2\cdot\frac{2}{10}+8\cdot\frac{1}{10}$$

שאלה:

מטילים מטבע הוגן n הטלות ב"ת.

מהי תוחלת מספר השונות שיתקבלו?

אפשר לקבל תוצאה ~~א~~ כל הפעמים מתקבל עץ

או כל הפעמים פלי, אחרת מתקבלים שתי תוצאות.

פתרון על פי ההתפלגות

X - מספר התוצאות שנקבל.

$$P(X=1)=\left(\tfrac{1}{2}\right)^n+\left(\tfrac{1}{2}\right)^n=2\left(\tfrac{1}{2}\right)^n=\left(\tfrac{1}{2}\right)^{n-1}$$

$$P(X=2)=1-\left(\tfrac{1}{2}\right)^{n-1}$$

$$E(X)=P(X=1)\cdot 1+P(X=2)\cdot 2=\left(\tfrac{1}{2}\right)^{n-1}\cdot 1+\left(1-\left(\tfrac{1}{2}\right)^{n-1}\right)\cdot 2$$

$$=2-0.5^{n-1}$$

פתרון ע"י אינדיקטורים

X_1 - אינדיקטור לקבלת התוצאה עץ לפחות פעם אחת.

X_2 - אינדיקטור לקבלת התוצאה פלי לפחות פעם אחת.

$$X=X_1+X_2$$

$$E(X)=E(X_1)+E(X_2)$$

$$P(X_1=0)=0.5^n \qquad P(X_2=0)=0.5^n$$

הסיכוי שלא לקבל עץ או פלי.

$$P(X_1=1)=1-0.5^n \qquad P(X_2=1)=1-0.5^n$$

$$E(X_1)=1-0.5^n \qquad E(X_2)=1-0.5^n$$

$$E(X)=E(X_1)+E(X_2)=(1+0.5^n)+(1+0.5^n)=2-0.5^{n-1}$$

כעת נשנה את השאלה ונטיל קוביה הוגנת n

הטלות ב"ת. מהי תוחלת מספר התוצאות שנקבל?

פתרון:

X_i - אינדיקטור למספר i לפחות פעם אחת.

$$P(X_i=0)=\left(\frac{5}{6}\right)^n$$

$$P(X_i=1)=1-\left(\frac{5}{6}\right)^n$$

$$E(X_i)=P(X_i=1)=1-\left(\frac{5}{6}\right)^n$$

$$E(X)=\sum E(X_i)=6\left[1-\left(\frac{5}{6}\right)^n\right]$$

לינאריות התוחלת:

X - משתנה מקרי

$$E(aX+b)=a\cdot E(X)+b$$

כאשר a ו b הם מספרים ממשיים.

הא חסר:

$$P(X=1)=P(X=3)=0.5$$

a שווה 5, b שווה 6.

$$E(X)=0.5\cdot 1+0.5\cdot 3=2$$

נקרא למשתנה $y=ax+b$.

$$Y=ax+b=5x+6$$

~~P(Y)=~~

$$P(Y=11)=P(X=1)=0.5$$

$$P(Y=21)=P(X=3)=0.5$$

$$E(Y)=\frac{1}{2}\cdot 11+\frac{1}{2}\cdot 21=16$$ לכן

לפי הנוסחה של לינאריות התוחלת נקבל את אותה תוצאה.

$$E(aX+b)=a\cdot E(X)+b \qquad E(X)=2$$

$$E(Y)=5\cdot E(X)+6=5\cdot 2+6=16$$

סעיף:

במבחן אמריקאי יש 20 שאלות. הסיכוי של תלמיד לענות נכון כל שאלה הוא $\frac{1}{4}$ והסיכוי $\frac{3}{4}$ הוא טעות. על תשובה נכונה הוא מקבל 5 נקודות ועל שגויה מורידים לו 3 נקודות. מה תוחלת הציון של תלמיד זה שעונה את כל התשובות?

פתרון ראשון:

נשתמש על לינאריות התוחלת.

X - מספר התשובות הנכונות של התלמיד.

Y - הציון של התלמיד.

$$Y = 5X - 3(20-X) = 8X - 60$$

$$X \sim Bin(20, \tfrac{1}{4}) \Rightarrow E(X) = 20 \cdot \tfrac{1}{4} = 5$$

$$E(Y) = 8 \cdot E(X) - 60 = 8 \cdot 5 - 60 = -20$$

תוחלת שלילית - לא משנה.

פתרון שני:

נשתמש על תוחלת סכום.

יהי W_i - מספר הנקודות שהתלמיד מקבל בשאלה i.

$$E(W_i) = \tfrac{1}{4} \cdot 5 - \tfrac{3}{4} \cdot 3 = -1$$

W - הציון הכולל של התלמיד.

$$E(W) = E(\sum W_i) = \sum E(W_i) = 20 \cdot (-1) = -20$$

שאלה:

ענה שחקנים יחד תחרות לנחש 11 שאלות. בין

עשר הוא 60. מה כדאי לשחק לפחות?

שווה לו. למה תשובה אחת ויחס. לדעתי כסיכוי

של $\frac{1}{2}$. לא נראה שכדאי לו למה הרבה כי אין חוק

תלוי על תוחלת שאלות לניחושים.

תוחלת של משתנה סימטרי:

דוגמא:

$$P(X=7)=0.1 \qquad P(X=13)=0.1$$

$$P(X=8)=0.4 \qquad P(X=12)=0.4$$

יש סימטריה סביב 10. הסתברות שמשמאל שווה
להסתברות להיות שמאלה. לכן ~~...~~ סביר שהתוחלת
(הערך הממוצע של ההתפלגות) הוא 10.

טענה:

אם X סימטרי סביב $\cancel{a}$ אז תוחלתו $E(X)=a$.

נוכיח ראשון נניח שאם X סימטרי סביב 0 אז

$$E(X)=0$$

X סימטרי סביב 0, כלומר שעבור כל ערך x
מתקיים $P(X=x)=P(X=-x)$.

$$E(X)=\sum_x P(X=x)\cdot x = \left(\sum_{x<0} P(X=x)\cdot x\right)+\left(\sum_{x>0} P(X=x)\cdot x\right)$$

$$=\left(\sum_{x<0} P(X=-x)\cdot x\right)+\left(\sum_{x>0} P(X=x)\cdot x\right)=$$

$$=\left(\sum_{x>0} P(X=x)\cdot(-x)\right)+\left(\sum_{x>0} P(X=x)\cdot x\right)=$$

$$=-\left(\sum_{x>0} P(X=x)\cdot x\right)+\left(\sum_{x>0} P(X=x)\cdot x\right)=0$$

הראינו שאם X סימטרי סביב 0 אז $E(X)=0$.
כעת נניח ש Y סימטרי סביב b ונרצה להראות
ש $E(Y)=b$.

נגדיר משתנה $X=Y-b$. כך הערך b הופך
ל-0, כלומר מזיזים בגודל b את X והופכת להיות
בגודל b ל-0.

נסתכל על משתנה Y כאשר הוא סימטרי סביב b.
משמעות זה שמתקבלים הערכים $b+d$ ו $b-d$
באותה הסתברות.

אם נגדיר משתנה $X=Y-b$ אז הוא יתקבל
הערכים d ו $-d$ בסיכויים שווים.
לכן X סימטרי סביב 0.

X סימטרי סביב 0 ולכן מתקיים $E(X)=0$
$E(Y-b)=0$ לכן לפי לינאריות התוחלת מתקיים
$E(Y)-b=0$ ולכן $E(Y)=b$ כנדרש. משמעות שהוא
סימטרי סביב b ונעזרנו בלינאריות התוחלת וקבלנו את
התוצאה.

נוסחה: הצגה לתוחלת חלופית:

של משתנה שמקבל ערכים שלמים אי שליליים בלבד.

X - משתנה שמקבל רק ערכים שלמים אי שליליים.

נראה שמתקיים $E(X)=\sum_{k=1}^{\infty} P(X\geq k)$.

זו נוסחה מתמטית רק למציאת ערך (נוסחה

לחישוב תוחלת).

דוגמא:

$$P(X=0)=\frac{1}{6}$$

$$P(X=1)=\frac{1}{2}$$

$$P(X=3)=\frac{1}{3}$$

מתקיים לפי נוסחה זו:

$$E(X)=0\cdot\frac{1}{6}+\frac{1}{2}\cdot 1+\frac{1}{3}\cdot 3=1.5$$

נראה כיצד לחשב בנוסחה זו.

$$E(X)=\sum_{k=1}^{\infty} P(X\geq k)=P(X\geq 1)+P(X\geq 2)+P(X\geq 3)+P(X\geq 4)+\dots$$

$$=\left(\frac{1}{2}+\frac{1}{3}\right)+\left(\frac{1}{3}\right)+\frac{1}{3}+0+0+\dots=1.5$$

הוכחת הנוסחה:

$$E(X) = \sum_{k=1}^{\infty} P(X=k)\cdot k = \sum_{k=1}^{\infty}\sum_{i=1}^{k}(X=k)$$

$$= \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} P(X=k) = \sum_{i=1}^{\infty} P(X \geq i)$$

וזו הנוסחה למשתנה שמקבל רק ערכים שלמים
אי שליליים.