שונות של משתנה $X\sim Pois(\lambda)$

$$Var(X)=\Sigma P(X=k)\cdot k^2-E^2(X)$$

$$E^2(X)=E(X)E(X)=\lambda\cdot\lambda=\lambda^2$$

$$\Sigma P(X=k)k^2=\Sigma_{k=0}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{k!}\cdot k^2=\Sigma_{k=1}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{k!}\cdot k^2=$$

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

$$\left[\Sigma_{k=0}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{k!}\cdot k=\lambda\right]$$

$$\left[\Sigma_{k=0}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{k!}\cdot k(k-1)=\Sigma_{k=2}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{k!}\cdot k(k-1)=\Sigma_{k=2}^{\infty}e^{-\lambda}\cdot\frac{\lambda^k}{(k-2)!}=\lambda^2\Sigma_{k=2}^{\infty}e^{-\lambda}\cdot\frac{\lambda^{k-2}}{(k-2)!}\right.$$

נציב $t=k-2 \rightarrow$

$$\left.=\lambda^2\Sigma_{t=0}^{\infty}e^{-\lambda}\cdot\frac{\lambda^t}{t!}=\lambda^2\cdot 1=\lambda^2\right]$$

ולכן: נקבל: $\lambda^2+\lambda$

חישוב שונות של משתנה $X\sim G(p)$

בעקרון השונות שווה ל: $V(X)=E(X^2)+E^2(X)$ כאשר :\

$$E^2(X)=\frac{1}{p^2}$$

$$E(X^2)=\Sigma P(X=k)k^2=\Sigma_{k=1}^{\infty}p\cdot q^{k-1}\cdot k^2$$

$$E(X^2)=P(X=1)\cdot 1^2+P(X>1)\cdot E((1+X)^2)=$$

$$p\cdot 1^2+q\cdot E(1+2X+X^2)=$$

$$p+q\cdot 1+q\cdot 2\cdot E(X)+q\cdot E(X^2)=$$

$$1+2q\cdot E(X)+q\cdot E(X^2)=$$

$$1 + 2q \cdot \frac{1}{p} + q \cdot E(X^2) =$$

קיבלנו :

$$E(X^2) = 1 + 2 \cdot \frac{q}{p} + q \cdot E(X^2)$$

$$E(X^2) - q \cdot E(X^2) = 1 + \frac{2q}{p}$$

$$P \cdot E(X^2) = \frac{p + 2q}{p}$$

$$V(X) = E(X^2) + E^2(X) = \frac{p+2q}{p^2} - \frac{1}{p^2} = \frac{p+2(1-p)-1}{p^2} = \frac{p+2-2p-1}{p^2} = \frac{p+1-2p}{p^2} = \frac{1-p}{p^2} = \frac{q}{p^2}$$

זה סביר כי אם p קרוב ל-1 אז q קרוב ל-0 וכפי הנראה נקבל ההצלחה מהר עם פיזור קטן.

אם p קטן אז, אז בכל שלב יש סיכוי קטן להצלחה. נקבל הצלחה בפיזור גבוה והשונות גדולה.

חישוב שונות של משתנה $X \sim NB(n,p)$

משתנה $NB(n,p)$ הוא סכום של n משתנים $G(p)$ ב״ת. לכן שונותו שווה לסכום שונותם שזה $n \cdot \frac{q}{p^2}$.

על השונות של סכום משתנים שווי התפלגות ושל מכפל של משתנה בקבוע

נניח ש $\{X_i\}_{i=1}^{n}$ הם שווי התפלגות בלתי תלויים. $Var(\Sigma_{i=1}^{n} X_i) = n \cdot V(X_2)$ כי השונות המשותפות הן 0.

$$V(nX_1) = n^2 \cdot V(X_1)$$

$V(\Sigma_{i=1}^{n} X_i) = \Sigma V(X_i) + 2\Sigma Cov(X_i, X_j) = \Sigma V(X_i) + 0$ (כי הם בלתי תלויים)

$$V(nX_1) = V(X_1, X_1, \dots, X_1) = \Sigma V(X_1) + 2\Sigma Cov(X_1, X_1)$$

אבל $Cov(X_1, X_1) = V(X_n)$.

שונות של ממוצע היא $Var\left(\frac{\Sigma_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2} \cdot Var(\Sigma_{i=1}^{n} X_i) = \frac{1}{n^2} \cdot \ n \cdot V(X_1) = \frac{V(X_1)}{n}$

$$\lim_{n \to \infty} \frac{V(X_1)}{n} = 0$$

אם היינו עושים ממוצע על פני אותו נשאל אז $Var\left(\frac{\Sigma X_1}{n}\right)=\frac{1}{n^2}\cdot V(nX_1)=\frac{1}{n^2}\cdot\ n^2V(X_1)=V(X_1)$.

- **סוגיה**

נניח שיש משתנה X שאני יודע איך הוא מתפלג. אני צריך לבחור ערך ואני משלם קנס בכל מקרה שלא ניחשתי נכון את הערך ש-X יקבל. מה צריכה להיות מדיניותי?

א. אם בכל מקרה של טעות אשלם אותו קנס של 1 ₪.

ב. נניח שאני משלם קנס ששווה לריבוע הטעות שלי, מה עלי לבחור?

- פתרון:

א. אני צריך לנחש את הערך השכיח שהוא הערך שמתקבל בסיכוי הגבוה ביותר.

ב. טענה: כדי לי לנחש את $E(X)$ ותוחלת הקנס תהיה שווה ל-$Var(X)$.

הוכחה: נניח שאני מנחש את הגודל a.

$$E((X-a)^2)=E\left(\left((X-\mu_X)+(\mu_X-a)\right)^2\right)=$$

$$E\left((X-\mu_X)^2-(\mu_X-a)^2+2(\mu_X-a)(X-\mu_X)\right)=$$

$$E(X-\mu_X)^2-E((\mu_X-a)^2)+2E\left((\mu_X-a)(X-\mu_X)\right)=$$

$$V(X)+(\mu_X-a)^2+2(\mu_X-a)\cdot E(X-\mu_X)=$$

$$V(X)+(\mu_X-a)^2$$

את זה מבינים למינימום אם מאפסים את $(\mu_X-a)^2$. ואז $a=\mu_X$ ותוחלת הקנס היא $Var(X)$.

מקדם המתאם ומשתנים מתוקננים

לכל משתנה מקרי X נגדיר את המשתנה המתוקנן של $\hat{X}$:

$$\hat{X} = \frac{X - \mu_X}{\sigma_X}$$

כאשר: μ_X – זה התוחלת, σ_X – זו סטיית התקן שהיא השורש של השונות.

<u>טענה</u>: התוחלת של כל משתנה מתוקנן היא אפס והשונות שלו היא 1.

$$E\left(\frac{X - \mu_X}{\sigma_X}\right) = \frac{1}{\sigma_X} E(X - \mu_X) = 0$$

$$V\left(\frac{X - \mu_X}{\sigma_X}\right) = \frac{1}{\sigma_X^2} \cdot V(X - \mu_X) = \frac{1}{\sigma_X^2} \cdot V(X) = 1$$

- **דוגמא:**

נניח ש: $X \sim Bin\left(2, \frac{1}{2}\right)$

$$\sigma_X = \sqrt{(2 \cdot 0.5 \cdot 0.5)} = \sqrt{0.5}, \qquad \mu_X = 1, \qquad \hat{X} = \frac{X - 1}{\sqrt{0.5}}$$

$$P\left(\hat{X} = \frac{0 - 1}{\sqrt{0.5}}\right) = P(X \ = 0) = 0.5^2$$

$$P(\hat{X} = \frac{1 - 1}{\sqrt{0.5}} = P(X = 1) = \binom{2}{1} \cdot 0.5 \cdot 0.5$$

$$P\left(\hat{X} = \frac{2 - 1}{\sqrt{0.5}}\right) = P(X = 2) = 0.5^2$$

<u>טענה</u>: אם Y היא פונקציה לינארית של X כך ש: $Y = aX + b$, אז אם $a > 0$ אז יש ל-X ול-Y אותו משתנה מתוקנן,
ואם $a < 0$ אז $\hat{Y} = -\hat{X}$.

נראה את החלק הראשון:

$$Y = aX + b, \qquad \mu_Y = \mu_X + b, \qquad \sigma_Y^2 = a^2 \sigma_X^2$$

$$\hat{Y} = \frac{Y - \mu_Y}{\sigma_Y} = \frac{(ax + b) - (a\mu_X + b)}{a \cdot \sigma_X} = \frac{a(X - \mu_X)}{a\sigma_X} = \frac{X - \mu_X}{\sigma_x} = \hat{X}$$

באופן דומה אם $a < 0$ אז $\hat{Y} = -\hat{X}$.

<u>טענה בכיוון ההפוך</u>: אם ל-X ול-Y יש את אותו משתנה מתוקנן אז Y הוא פונקציה לינארית עולה של X.
אם ל-X ול-Y יש $\hat{Y} = -\hat{X}$ אז Y הוא פונקציה לינארית יורדת של X.

הוכחת החלק הראשון:

$$\hat{Y} = \hat{X}$$

ולכן:

$$\frac{Y-\mu_Y}{\sigma_Y} = \frac{X-\mu_X}{\sigma_X}$$

ולכן:

$$Y = \frac{\sigma_Y}{\sigma_X}(X-\mu_X) + \mu_Y$$

וברור שזו פונקציה לנראית עולה. הוכחת החלק השני היא דומה.

קיבלנו שלשני משתנים יש אותו משתנה מתוקנן אם ורק אם השני הוא פונקציה לינארית עולה של הראשון. קיבלנו שלשני משתנים יש $\hat{Y} = -\hat{X}$ אם ורק אם Y הוא פונקציה לינארית יורדת של X.

נגדיר מקדם מתאם בין 2 משתנים:

<u>הגדרה</u>: $r(X,Y)$ או $\rho(X,Y)$ אלה שני סימונים מקובלים.

$$r(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

<u>טענה</u>: $r(X,Y) = Cov(\hat{X},\hat{Y})$.

$$r(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E\big((X-\mu_X)(Y-\mu_Y)\big)}{\sigma_X \sigma_Y} = E\left(\left(\frac{Y-\mu_Y}{\sigma_Y}\right)\left(\frac{X-\mu_X}{\sigma_X}\right)\right) =$$

$$E(\hat{X}\hat{Y}) = E(\hat{X}\hat{Y}) - E(\hat{X})E(\hat{Y}) = Cov(\hat{X},\hat{Y})$$

*כי הביטוי באדום $= 0$.

תכונות של $r(X,Y)$

א. $r(X,Y) = 0 \Leftrightarrow Cov(X,Y) = 0$

ב. $-1 \le r(X,Y) \le 1$

ג. $r(X,Y) = 1$ אם ורק אם Y פונקציה לינארית עולה של X.

ד. $r(X,Y) = -1$ אם ורק אם Y פונקציה לינארית יורדת של X.

הוכחה:

א. $r(X,Y)=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}$. ברור שכל אחד מהם הוא מכפלה בקבוע של האחר ולכן כל אחד מהם מתאפס אם האחר מתאפס.

ב. נעזר בזה ש-$r(X,Y)=Cov(\hat{X},\hat{Y})$. צריך להראות ש-$-1\le Cov(\hat{X},\hat{Y})\le 1$.

נסתכל על $\hat{X}-\hat{Y}$.

$Var(\hat{X}-\hat{Y})\ge 0$ (כי שונות תמיד גדולה מ-0).

$$V(\hat{X}-\hat{Y})=V(\hat{X})+V(\hat{Y})-2Cov(\hat{X},\hat{Y})=$$

$$1+1-2Cov(\hat{X},\hat{Y})=2-2Cov(\hat{X},\hat{Y})=2\left(1-Cov(\hat{X},\hat{Y})\right)$$

קיבלנו ש-$1-Cov(\hat{X},\hat{Y})\ge 0$. ולכן $Cov(\hat{X},\hat{Y})\ge 1$ ולכן: $r(X,Y)\le 1$.

עכשיו נרצה להראות ש: $r(X,Y)\ge -1$.

נתסכל על $\hat{X}+\hat{Y}$.

$$V(\hat{X}+\hat{Y})=V(\hat{X})+V(\hat{Y})+2Cov(\hat{X},\hat{Y})=$$

$$1+1+2Cov(\hat{X},\hat{Y})=2+2Cov(\hat{X},\hat{Y})=2\left(1+Cov(\hat{X},\hat{Y})\right)$$

קיבלנו ש-$1+Cov(\hat{X},\hat{Y})\ge 0$. ולכן $Cov(\hat{X},\hat{Y})\ge -1$ ולכן: $r(X,Y)\ge -1$.

ג. אם Y הוא פונקציה לינארית עולה של X אז $\hat{Y}=\hat{X}$

ולכן $Var(\hat{X})=1$ $Cov(\hat{Y},\hat{X})=Cov(\hat{X},\hat{X})=$ ולכן $r(X,Y)=1$.

בכיוון ההפוך: אם $r(X,Y)=1$ אז $Cov(\hat{X},\hat{Y})=1$ אז $E(\hat{X}\hat{Y})-E(\hat{X})E(\hat{Y})=1$ אז $E(\hat{X}\hat{Y})=1$

נסתכל על $\hat{X}-\hat{Y}$.

$$Var(\hat{X}-\hat{Y})=Var(\hat{X})+Var(\hat{Y})-2E(\hat{X}\hat{Y})=1+1-2\cdot 1=0$$

לכן $(\hat{X}-\hat{Y})$ הוא קבוע. רק לקבוע יש שונות של 0. הוא לא יכול להיות שווה לקבוע שונה מ-0 כי לשניהם תוחלת 0. לכן $\hat{X}-\hat{Y}$ הוא קבוע ששווה ל-0. ולכן $\hat{Y}=\hat{X}$.

ד. אם Y הוא פונקציה לינארית יורדת של X אז $\hat{Y}=-\hat{X}$ ולכן $V(\hat{X},\hat{Y})=0$.

$$Var(\hat{X}+\hat{Y})=Var(\hat{X})+Var(\hat{Y})+2Cov(\hat{X},\hat{Y})=1+1+2r(X,Y)=2(1+r(X,Y))$$

לכן $r(X,Y)=-1$. נניח ש-$r(X,Y)=-1$ אז $Cov(\hat{X},\hat{Y})=-1$. נסתכל על $V(\hat{X}+\hat{Y})$:

$$Var(\hat{X}+\hat{Y})=Var(\hat{X})+Var(\hat{Y})+2Cov(\hat{X},\hat{Y})=1+1+2r(X,Y)=2(1+(-1))=0$$

לכן $(\hat{X}+\hat{Y})$ הוא קבוע. רק לקבוע יש שונות של 0. הוא לא יכול להיות שווה לקבוע שונה מ-0 כי לשניהם תוחלת 0. לכן $\hat{X}+\hat{Y}=0$ ולכן $\hat{Y}=-\hat{X}$.

<u>הערות</u>:

1. נניח ש $Y=X^2$. האם מקדם המתאם יכול להיות שווה ל-$1\backslash 0\backslash -1$?

מדובר על $r(X,Y)$, אם X מוגדר על קבוצת נקודות: יותר משתי נקודות אז אין קשר לינארי. X^2 אינה פונקציה לינארית. בין 3 נקודות שעל קו ישר לא עוברת פונקציה $Y=X^2$. רק אם X מוגדר על שתי נקודות אז זו יכולה להיות פונקציה לינארית.

❖ **דוגמא**:

פונקציה לינארית:

Y	X
1	1
4	2

$$P(X=1,Y=1)=\frac{1}{3},\qquad P(X=2,Y=4)=\frac{2}{3}$$

במקרה זה יש קו ישר ולכן $r(X,Y)=1$, ה-$\frac{2}{3},\frac{1}{3}$ הם שרירותיים.

2. מתי המתאם יכול להיות -1?

$$P(X=0,Y=0)=\frac{1}{3},\qquad P(X=-1,Y=1)=\frac{2}{3}$$

במקרה זה יש קשר לינארי יורד.