<u>מבוא להסתברות</u>

© ארזים

8 במאי 2016

0.1 תוחלת של פונקציה של משתנה מקרי

יהי X משתנה מקרי בעל תוחלת, ופונקציה $g:\mathbb{R}\to\mathbb{R}$. אזי

$$\mathbb{E}\left(g\left(x\right)\right)=\sum_{k\in\text{supp}(x)}g\left(k\right)\cdot\mathbb{P}\left(X=k\right)$$

0.2 משתנים מקריים בלתי מתואמים

הגדרה 0.1 יהיו X,Y משתנים מקריים בעלי תוחלת. נניח שהתוחל של XY גם מוגדרת. אזי X,Y בלתי מתואמים אם

$$\mathbb{E}\left(XY\right)=\mathbb{E}\left(X\right)\cdot\mathbb{E}\left(Y\right)$$

טענה 0.2 משתנים מקריים בלתי תלויים הם בלתי מתואמים.

תרגיל יהי מרחב הסתברות $(\Omega,\mathbb{P})$, כאשר $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$, $\mathbb{P}\left(\omega_i\right)=\frac{1}{4}$ לכל i. נגדיר

$$X\sim\begin{cases}-2 & \omega_1\\ -1 & \omega_2\\ 1 & \omega_3\\ 2 & \omega_4\end{cases}$$

ונגדיר $Y=X^2$. האם X,Y בלתי תלויים? בלתי מתואמים?

פתרון המשתנים המקריים תלויים שכן

$$\mathbb{P}\left(X=1,Y=4\right)=0\neq\mathbb{P}\left(X=1\right)\cdot\mathbb{P}\left(Y=4\right)$$

הם כן בלתי מתואמים, שכן $\mathbb{E}X=0$, וגם $\mathbb{E}\left(X\cdot Y\right)=\mathbb{E}\left(X^3\right)=0$, ולכן $\mathbb{E}\left(X\cdot Y\right)=\mathbb{E}\left(X\right)\cdot\mathbb{E}\left(Y\right)$.

1 שונות

יהי X משתנה מקרי. השונות מוגדרת להיות

$$\mathrm{var}\left(X\right)=\mathbb{E}\left(\left(X-\mathbb{E}X\right)^{2}\right)=\mathbb{E}\left(X^{2}\right)-\left(\mathbb{E}X\right)^{2}$$

עבור $a,b\in\mathbb{R}$, מתקיים

$$\mathrm{var}\left(aX+b\right)=a^{2}\mathrm{var}\left(X\right)$$

1.1 שונות של סכום של משתנים מקריים

ראינו בשיעור את מושג השונות המשותפת. כעת, מתקיים

$$\mathrm{var}\left(\sum_{i=1}^{n}X_{i}\right)=\sum_{i=1}^{n}\mathrm{var}\left(X_{i}\right)+\sum_{i\neq j}\mathrm{cov}\left(X_{i},X_{j}\right)=\sum_{i=1}^{n}\mathrm{var}\left(X_{i}\right)+2\sum_{i<j}\mathrm{cov}\left(X_{i},X_{j}\right)$$

1.2 מספר תכונות של שונות משותפת

1. $\mathrm{cov}\left(X,Y\right)=\mathrm{cov}\left(Y,X\right)$
2. $\mathrm{cov}\left(aX+b,Y\right)=a\cdot\mathrm{cov}\left(X,Y\right)$
3. $\mathrm{cov}\left(X+Z,Y\right)=\mathrm{cov}\left(X,Y\right)+\mathrm{cov}\left(Z,Y\right)$

תרגיל במבחן בקורס יש 20 שאלות אמריקאיות, לכל שאלה 4 תשובות אפשריות ורק אחת נכונה. תשובה נכונה מזכה בשש נקודות, תשובה לא נכונה מורידה שתי נקודות. ישנן 8 שאלות קלות עליהן יודע אלון לענות נכון, את התשובות של 12 הנוספות הוא מנחש. נסמן Z - הציון של אלון. חשבו תוחלת ושונות של Z.

פתרון נגדיר X - מספר התשובות הנכונות מתוך השאלות הקשות. $X\sim\mathrm{Bin}\left(12,\frac{1}{4}\right)$. כעת,

$$\begin{aligned}
Z &= 48+6X-2\left(12-X\right)=24+8X\\
\mathbb{E}\left(Z\right) &= \mathbb{E}\left(24+8X\right)=24+8\mathbb{E}\left(X\right)=24+8\cdot12\cdot\frac{1}{4}=48\\
\mathrm{var}\left(Z\right) &= \mathrm{var}\left(24+8X\right)=64\cdot\mathrm{var}\left(X\right)=64\cdot12\cdot\frac{1}{4}\cdot\frac{3}{4}=144\\
\mathrm{std}\left(Z\right) &= \sqrt{\mathrm{var}\left(X\right)}=12
\end{aligned}$$

תרגיל אלון ובן ארגנו מסיבה עם n זוגות וסידרו את $2n$ האנשים בשורה באקראי. יהי X - מספר הזוגות שבהם בני הזוג עומדים אחד לצד השני. מצאו תוחלת ושונות של X.

פתרון נפרק את X לסכום של אינדיקטורים X_i, שמקבלים 1 אם הזוג i צמודים, 0 אחרת. כעת

$$X=\sum_{i=1}^{n}X_i$$

כעת

$$\begin{aligned}
\mathbb{P}\left(X_i=1\right) &= \frac{(2n-1)!\cdot 2!}{(2n)!}=\frac{1}{n}\\
\mathbb{E}\left(X_i\right) &= \frac{1}{n}\\
\mathbb{E}\left(X\right) &= \sum_{i=1}^{n}\mathbb{E}\left(X_i\right)=n\cdot\frac{1}{n}=1\\
\operatorname{var}\left(X\right) &= \sum_{i=1}^{n}\operatorname{var}\left(X_i\right)+\sum_{i\neq j}\operatorname{cov}\left(X_i,X_j\right)\\
\operatorname{var}\left(X_i\right) &= \frac{1}{n}\cdot\frac{n-1}{n}=\frac{n-1}{n^2}\\
\operatorname{var}\left(X\right) &= \frac{n-1}{n}+\sum_{i\neq j}\operatorname{cov}\left(X_i,X_j\right)\\
\operatorname{cov}\left(X_i,X_j\right) &= \mathbb{E}\left(X_i\cdot X_j\right)-\mathbb{E}\left(X_i\right)\cdot\mathbb{E}\left(X_j\right)=\mathbb{E}\left(X_i\cdot X_j\right)-\frac{1}{n^2}\\
X_i\cdot X_j &\sim \begin{cases}1 & X_i=X_j=1\\ 0 & \text{otherwise}\end{cases}\\
\mathbb{P}\left(X_i=1,X_j=1\right) &= \frac{(2n-2)!\cdot(2!)^2}{(2n)!}=\frac{2}{n\left(2n-1\right)}=\mathbb{E}\left(X_i\cdot X_j\right)\\
\operatorname{cov}\left(X_i,X_j\right) &= \frac{2}{n\left(2n-1\right)}-\frac{1}{n^2}=\frac{1}{n}\left(\frac{2}{2n-1}-\frac{1}{n}\right)=\frac{1}{n}\left(\frac{2n-2n+1}{n\left(2n-1\right)}\right)=\frac{1}{n^2\left(n-1\right)}\\
\operatorname{var}\left(X\right) &= \frac{n-1}{n}+\sum_{i\neq j}\frac{1}{n^2\left(2n-1\right)}=\frac{n-1}{n}+\frac{n\left(n-1\right)}{n^2\left(2n-1\right)}=\frac{n-1}{n}\left(1+\frac{1}{2n-1}\right)=\\
&= \frac{n-1}{n}\cdot\frac{2n}{2n-1}=\frac{2n-2}{2n-1}
\end{aligned}$$

תרגיל מטילים קוביה 100 פעמים. X - כמות ההטלות בהן יצא 2 או 3. Y - כמות ההטלות בהן יצא 1 או 2. חשבו את $\operatorname{cov}\left(X,Y\right)$.

פתרון $X,Y\sim\operatorname{Bin}\left(100,\frac{1}{3}\right)$. לכן $\mathbb{E}X=\mathbb{E}Y=\frac{100}{3}$. את X,Y נפרק לאינדיקטורים - X_i אינדיקטור של המאורע בו ההטלה i יצאה 2 או 3, Y_i אינדיקטור של המאורע בו

ההטלה i יצאה 1 או 2. לכן,

$$\begin{aligned} X &= \sum_{i=1}^{100} X_i \\ Y &= \sum_{i=1}^{100} Y_i \end{aligned}$$

נציב ונקבל

$$\text{cov}\left(X,Y\right)=\text{cov}\left(\sum_{i=1}^{100}X_i,\sum_{i=1}^{100}Y_i\right)=\sum_{i,j=1}^{100}\text{cov}\left(X_i,Y_j\right)$$

עבור $i\neq j$, X_i,Y_j מתייחסים להטלות שונות, וכל ההטלות בלתי־תלויות - לכן גם X_i,Y_j בלתי תלויים, ולכן הם בלתי מתואמים, כלומר $\text{cov}\left(X_i,Y_j\right)=0$. לכן

$$\begin{aligned} \text{cov}\left(X,Y\right) &= \sum_{i=1}^{100}\text{cov}\left(X_i,Y_i\right)=100\cdot\text{cov}\left(X_1,Y_1\right)=100\cdot\left(\mathbb{E}\left(X_1\cdot Y_1\right)-\mathbb{E}X_1\cdot\mathbb{E}Y_1\right)= \\ &= 100\left(\mathbb{P}\left(X_1=1,Y_1=1\right)-\frac{1}{3}\cdot\frac{1}{3}\right)=100\left(\mathbb{P}\left(\{2\}\right)-\frac{1}{9}\right)=100\left(\frac{1}{6}-\frac{1}{9}\right)= \\ &= 100\left(\frac{1}{18}\right)=\frac{50}{9} \end{aligned}$$

תרגיל לבן יש סכום כסף שהוא רוצה להשקיע. הוא יכול השקיע באפיק א', שנותן רווח X, או באפיק ב', שנותן רווח Y. בנוסף, הוא יכול לפצל את ההשקעה לכדי $\alpha\in[0,1]$ באפיק א' וכן $1-\alpha$ באפיק ב'. במקרה זה הרווח יהיה $Z\left(\alpha\right)=\alpha X+\left(1-\alpha\right)Y$. נתון כי $0<\mathbb{E}\left(X\right)=\mathbb{E}\left(Y\right)<\infty$, $0<\text{var}\left(X\right)<\text{var}\left(Y\right)<\infty$. כיצד צריך בן להשקיע כדי למקסם תוחלת רווח ולמזער את הסיכון (השונות)?

פתרון נשים לב כי לכל α מתקיים

$$\mathbb{E}\left(\alpha X+\left(1-\alpha\right)Y\right)=\alpha\mathbb{E}\left(X\right)+\left(1-\alpha\right)\mathbb{E}\left(Y\right)=\mathbb{E}\left(X\right)=\mathbb{E}\left(Y\right)$$

לכן התוחלת קבועה בלי קשר לפרמטר α. נמזער את השונות.

$$\begin{aligned} \text{var}\left(\alpha x+\left(1-\alpha\right)Y\right) &= \text{var}\left(\alpha X\right)+\text{var}\left(\left(1-\alpha\right)Y\right)+2\text{cov}\left(\alpha X,\left(1-\alpha\right)Y\right)= \\ &= \alpha^2\text{var}\left(X\right)+\left(1-\alpha\right)^2\text{var}\left(Y\right)+2\alpha\left(1-\alpha\right)\text{cov}\left(X,Y\right)= \\ &= \alpha^2\text{var}\left(X\cdot Y\right)+2\alpha\left(\text{cov}\left(X,Y\right)-\text{var}\left(Y\right)\right) \end{aligned}$$

החישובים נמצאים בעמוד 87 בחוברת התרגולים. זוהי פונקציה ריבועית ביחס אל α. נחפש לה נקודת מינימום. נקבל

$$\alpha_{min}=\frac{\text{var}\left(Y\right)-\text{cov}\left(X,Y\right)}{\text{var}\left(X\cdot Y\right)}$$

צריך לבדוק מתי $0 \leq \alpha_{min} \leq 1$, כי אחרת צריך לבחור את אחד מבין $0, 1$ כי הוא נקודת הקיצון בקטע זה.