<u>מבוא להסתברות</u>

4 במאי 2016

נוכיח את המשפט שראינו בסוף השיעור שעבר:

משפט 0.1 יהיו X,Y משתנים מקריים בעלי תוחלות מוגדרות, יהי c מספר ממשי.

1. התוחלת של cX מוגדרת ומתקיים $\mathbb{E}\left(cX\right)=c\mathbb{E}X$.
2. אם לא מתקיים $\mathbb{E}X=\infty,\mathbb{E}Y=-\infty$ או $\mathbb{E}X=-\infty,\mathbb{E}Y=\infty$, אזי התוחלת של $X+Y$ מוגדרת ומתקיים $\mathbb{E}\left(X+Y\right)=\mathbb{E}X+\mathbb{E}Y$.

הוכחה: נוכיח בסעיפים.

1. עבור $c=0$, ברור. נוכיח עבור $c>0$ (המקרה $c<0$ דומה). כעת,

$$\left(cX\right)^{+}=c\left(X^{+}\right)$$

ולכן

$$\mathbb{E}\left(\left(cX\right)^{+}\right)=\sum_{\omega\in\Omega}\mathbb{P}\left(\omega\right)\cdot\left(cX\right)^{+}\left(\omega\right)=\sum_{\omega\in\Omega}\mathbb{P}\left(\omega\right)\cdot c\cdot X^{+}\left(\omega\right)=c\sum_{\omega\in\Omega}\mathbb{P}\left(\omega\right)X^{+}\left(\omega\right)=c\mathbb{E}X^{+}$$

באותו אופן נקבל

$$\mathbb{E}\left(cX\right)^{-}=c\mathbb{E}X^{-}$$

ובסך הכל, התוחלת של cX מוגדרת ומתקיים

$$\mathbb{E}\left(cX\right)=\mathbb{E}\left(cX\right)^{+}-\mathbb{E}\left(cX\right)^{-}=c\left(\mathbb{E}X^{+}-\mathbb{E}X^{-}\right)=c\mathbb{E}X$$

2. נוכיח במקרה שבו $\mathbb{E}X,\mathbb{E}Y$ סופיות. מקרים אחרים - תרגיל. נקבל

$$\begin{aligned}\mathbb{E}\left(X+Y\right) &= \sum_{\omega\in\Omega}\mathbb{P}\left(\omega\right)\cdot\left(X+Y\right)\left(\omega\right)=\sum_{\omega\in\Omega}\mathbb{P}\left(\omega\right)\cdot X\left(\omega\right)+\sum_{\omega\in\Omega}\mathbb{P}\left(\omega\right)\cdot Y\left(\omega\right)=\\ &= \mathbb{E}X+\mathbb{E}Y\end{aligned}$$

כאשר הטור עבור $X+Y$ מתכנס בהחלט בגלל שהוא סכום של שני טורים מתכנסים בהחלט.

■

דוגמא בעיית ימי ההולדת
בכיתה יש n תלמידים. נניח כי ימי ההולדת שלהם בלתי תלויים, וכל אחד מתפלג אחיד מבין 365 האפשרויות.
לדיון: האם סביר שיש שני תלמידים עם אותו יום הולדת? כמה גדול צריך להיות n כדי שההסתברות תהיה לפחות $\frac{1}{2}$? מה ההסתברות לכך?

$$\mathbb{P}\left(2\text{ students share a birthday}\right)=1-\frac{365}{365}\cdot\frac{364}{365}\cdot\dots\cdot\frac{365-n+1}{365}=1-\binom{365}{n}\cdot\frac{n!}{365^{n}}$$

ההסתברות עוברת את $\frac{1}{2}$ לראשונה עבור $n=23$. כדי לקבל אינטואיציה, נסמן את X להיות כמות זוגות התלמידים עם אותו יום הולדת. כעת, השאלה האם יש זוג תלמידים עם אותו יום הולדת היא השאלה האם $X>0$. לכן, נחשב את התוחלת של X. את זה נעשה בעזרת הלינאריות: נסמן עבור $1\le i<j\le n$ את המשתנה המקרי $X_{i,j}$ להיות האינדיקטור של המאורע בו תלמיד i ותלמיד j נולדו באותו יום. אזי

$$X=\sum_{1\le i<j\le n}X_{i,j}$$

ולכן

$$\mathbb{E}X=\sum_{1\le i<j\le n}\mathbb{E}X=\binom{n}{2}\cdot\frac{1}{365}=\frac{n\left(n-1\right)}{2\cdot365}$$

וכעת מתקיים $\mathbb{E}X\ge1$ לראשונה כאשר $n=28$.

טענה 0.2 (מונוטוניות התוחלת) אם X,Y משתנים מקריים ממשיים המקיימים $\mathbb{P}\left(X\le Y\right)=1$, והתוחלות של X,Y מוגדרות, אזי

$$\mathbb{E}X\le\mathbb{E}Y$$

ושוויון מתקבל אם ורק אם $\mathbb{P}\left(X=Y\right)=1$.

הוכחה: נראה כי $\mathbb{E}X^{+}\le\mathbb{E}Y^{+},\mathbb{E}X^{-}\ge\mathbb{E}Y^{-}$, ומכאן נסיק את הנדרש. נראה עבור X^{+} - עבור X^{-} ההוכחה דומה.

$$\mathbb{E}X^{+}=\sum_{\omega\in\Omega}\mathbb{P}\left(\omega\right)\cdot X^{+}\left(\omega\right)\le\sum_{\omega\in\Omega}\mathbb{P}\left(\omega\right)\cdot Y^{+}=\mathbb{E}Y^{+}$$

כאשר המעבר עם אי השוויון נובע משום שמתקיים $X\left(\omega\right)\le Y\left(\omega\right)$. השוויון מתקבל אם ורק אם תמיד מתקיים שוויון, כלומר $X\left(\omega\right)=Y\left(\omega\right)$ תמיד, כלומר $\mathbb{P}\left(X=Y\right)=1$.
■

0.1 תוחלת של פונקציה של משתנה מקרי

נתון משתנה מקרי שמקבל ערכים בקבוצה S ופונקציה $f : S \to \mathbb{R}$. נדון בבעיית חישוב $\mathbb{E}f(X)$ (למשל $\mathbb{E}X^2, \mathbb{E}\left(\frac{1}{X}\right)$ וכולי). בדרך כלל,

$$\mathbb{E}f(X) \neq f(\mathbb{E}X)$$

טענה 0.3 אם f אי־שלילית, כלומר $f : S \to [0, \infty)$ אזי

$$\mathbb{E}f(X) = \sum_x \mathbb{P}(X = x) \cdot f(x)$$

כמו כן, עבור פונקציה כללית, $\mathbb{E}f(X)$ סופית אם ורק אם הטור למעלה מתכנס בהחלט, ואז הטור שווה לתוחלת.

ההוכחה כתרגיל.

דוגמא $X \sim \text{Geom}\left(\frac{1}{2}\right)$, $Y \sim 2^X$.

$$\begin{aligned}
\mathbb{E}X &= 2 \\
\mathbb{E}Y &= ? \\
X &\sim \begin{cases} \frac{1}{2^k} & k \end{cases} \\
Y &\sim \begin{cases} \frac{1}{2^k} & 2^k \end{cases} \\
\mathbb{E}Y &= 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + \cdots = \infty \\
\mathbb{E}Y &= \mathbb{P}(X = 1) \cdot 2 + \mathbb{P}(X = 2) \cdot 4 + \cdots = \infty
\end{aligned}$$

ראינו כאן שתי דרכים שונות לישוב התוחלת ־ וראינו שהן מזדהות.

נזכר שעבור פונקציה לינארית f מתקיים $\mathbb{E}f(X) = f(\mathbb{E}X)$.

0.2 תוחלת של מכפלה של משתנים מקריים בלתי־תלויים

יהיו X, Y משתנים מקריים ממשיים. גם כאשר X, Y בעלי תוחלת סופית, לא נובע שהמשתנה $X \cdot Y$ הוא בעל תוחלת, וממילא בדרך כלל

$$\mathbb{E}X \cdot Y \neq \mathbb{E}X \cdot \mathbb{E}Y$$

יחד עם זאת, יש לנו התוצאה הבאה:

טענה 0.4 יהיו X, Y משתנים מקריים ממשיים <u>בלתי־תלויים</u>. אם X, Y אי־שליליים או שלשניהם תוחלת סופית, אז התוחלת של $X \cdot Y$ מוגדרת ומתקיים $\mathbb{E}(X \cdot Y) = \mathbb{E}X \cdot \mathbb{E}Y$.

הוכחה: נתחיל במקרה האי־שלילי. נשתמש בטענה הקודמת:

$$\begin{aligned}\mathbb{E}(X\cdot Y) &= \sum_{x,y}\mathbb{P}(X=x,Y=y)\cdot xy=\sum_{x,y}\mathbb{P}(X=x)\cdot x\cdot\mathbb{P}(Y=y)\cdot y=\\ &= \left(\sum_{x}\mathbb{P}(X=x)\cdot x\right)\cdot\left(\sum_{y}\mathbb{P}(Y=y)\cdot y\right)=\mathbb{E}X\cdot\mathbb{E}Y\end{aligned}$$

כאשר יש לשניהם תוחלות סופיות, נרשום:

$$\mathbb{E}(X\cdot Y)=\mathbb{E}\left(X^+-X^-\right)\left(Y^+-Y^-\right)=\mathbb{E}\left(X^+Y^+\right)-\mathbb{E}\left(X^+Y^-\right)-\mathbb{E}\left(X^-Y^+\right)+\mathbb{E}\left(X^-Y^-\right)$$

המעבר האחרון נבע מלינאריות. כעת, ממה שהראינו עבור אי־שליליים, השוויון הקודם שווה גם

$$\begin{aligned}&= \mathbb{E}\left(X^+\right)\cdot\mathbb{E}\left(Y^+\right)-\mathbb{E}\left(X^+\right)\cdot\mathbb{E}\left(Y^-\right)-\mathbb{E}\left(X^-\right)\cdot\mathbb{E}\left(Y^+\right)+\mathbb{E}\left(X^-\right)\cdot\mathbb{E}\left(Y^-\right)=\\ &= \left(\mathbb{E}X^+-\mathbb{E}X^-\right)\cdot\left(\mathbb{E}Y^+-\mathbb{E}Y^-\right)=\mathbb{E}X\cdot\mathbb{E}Y\end{aligned}$$

■

1 שונות

הגדרה 1.1 יהי X משתנה ממשי בעלת תוחלת סופית. השונות של X היא

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

סטיית התקן של X היא

$$\mathrm{Std}(X)=\sqrt{\mathrm{Var}(X)}$$

הערה 1.2 השונות מוגדרת לכל משתנה מקרי עם תוחלת סופית, אבל עשויה להיות אינסוף. יש אלטרנטיבות להגדרה של סטיית התקן, אבל נבחרה זו שלנו ־ נדון בכך.