משתנים מקריים

משתנים מקריים

משתנה מקרי הוא פונק' $X:\Omega\to\mathbb{R}$. אם $X=a$ בהרתי, מאורע שניתן a.

דוגמה: נניח שהניסוי הוא הטלת 3 מטבעות תקינים. אם X מסמן את מס' הפעמים שהתקבל H, אז X הוא משתנה מקרי שמקבל את אחד הערכים $0, 1, 2, 3, 4$ בהסתברויות:

$$P\{X=0\}=P\{(T,T,T)\}=\frac{1}{8}$$

$$P\{X=1\}=P\{(T,T,H),(T,H,T),(H,T,T)\}=\frac{3}{8}$$

$$P\{X=2\}=P\{(T,H,H),(H,T,H),(H,H,T)\}=\frac{3}{8}$$

$$P\{X=3\}=P\{(H,H,H)\}=\frac{1}{8}$$

מכיוון ש X חייב לקבל את אחד מהערכים הללו, בהכרח מתקיים

$$P\left(\bigcup_{i=1}^{3}\{X=i\}\right)=1$$

סימון: הפונק' $P(X=a)$ נקראת פונק' ההסתברות/ההתפלגות

הפונק' $P(X^{-1}(a))$ היא פונק' שמחזירה מאורע

הערה: כל מ"מ משרה אירוע בר של Ω

פונקציית ההתפלגות המצטברת

פונק' ההצטברות: $F_X(t)=\sum_{a\le t} P(X=a) = P(X\le t)$

כי X משרה איחוד זר.

תכונות:

1. פונק' עולה
2. $\lim_{t\to\infty} F_X(t) = 1$
3. $\lim_{t\to-\infty} F_X(t) = 0$

הגדרה: השכיח של מ"מ הוא $\underset{a}{\operatorname{argmax}}(P(X=a))$
כלומר הערך ב $\mathbb{R}$ שמתקבל בהסתברות הגבוהה ביותר.

הגדרה: חציון של מ"מ. $Med(X)=m \iff F_X(m)=\frac{1}{2}$

תוחלת

הגדרה: התוחלת של מ"מ X היא $E(X)=\sum_{a} a\cdot P(X=a)$

דוגמה: הניסוי הוא הטלת קובייה, X הוא ערך הקובייה.

$$E(X)=1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{1}{6}+6\cdot\frac{1}{6}=\frac{7}{2}=3.5$$

תוחלת של פונקציה של משתנה מקרי

<u>טענה:</u> יהי X משתנה מקרי המקבל את הערכים $x_1, x_2, \dots$ בהסתברויות $P(x_1), P(x_2), \dots$ אזי. לכל פונקציה ממשית f מתקיים

$$E(f(X)) = \sum_i g(x_i)\, P(x_i)$$

<u>דוגמה:</u> מטילים שני מטבעות הוגנים, X הוא סכום ההצלות

e	$p(e)$	X	X^2
0,0	$\frac{1}{4}$	0	0
0,1	$\frac{1}{4}$	1	1
1,0	$\frac{1}{4}$	1	1
1,1	$\frac{1}{4}$	2	4

$$E(X) = 0\,\tfrac{1}{4} + 1\,\tfrac{1}{4} + 1\cdot\tfrac{1}{4} + 2\cdot\tfrac{1}{4} = 1$$

$$E(X) = 0\cdot\tfrac{1}{4} + 1\,\tfrac{1}{4} + 1\cdot\tfrac{1}{4} + 4\cdot\tfrac{1}{4} = \tfrac{3}{2}$$

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

<u>הערה:</u> זה בדיוק נכון כאשר הפונק' לינארית

<u>טענה (לינאריות התוחלת):</u>

$$E(aX+b) = aE(X) + b$$
$$E(X+Y) = E(X) + E(Y)$$
$$E(aX+bY) = aE(X) + bE(Y)$$

__הערה:__ במקרה הכללי, $E(f(X)) = \sum_{a} f(a) P(X=a)$

__הערה:__ התוחלת של מספר הינו המספר עצמו

$$E(a) = a \cdot P(X=a) = a \cdot 1 = a$$

__מסקנה:__ $E(E(X)) = E(X)$

שונות

<u>מוטיבציה:</u> חשוב לראות מה ההסתברות של מ"מ להיות קרוב לתוחלת בהגדרה אחת.

<u>הגדרה:</u> יהי משתנה מקרי X נגדיר משתנה מקרי חדש $Z=(X-E(X))^2$. השונות של X היא $Var(X)=E(Z)$

<u>מסקנה:</u> מפעולות אלגבריות פשוטות ומלינאריות התוחלת קיבלנו

$$Var(X)=E(X^2)-E^2(X)$$

<u>תכונות:</u>

1. $(X-E(X))^2\geq 0$ לכן $Var(X)\geq 0$

2. $V(X)=0 \iff P(X=E(X))=1$

<u>דוגמה:</u> התפלגות גבהים

X - גובה אופן רנדומלי

$X = \quad 160, \; 180, \; 200$

$P(X=a) = \quad \frac{1}{4} \quad \frac{1}{2} \quad \frac{1}{4}$

תוחלת: $E(X)=\frac{1}{4}\cdot 160+\frac{1}{2}\cdot 180+\frac{1}{4}\cdot 200=180$

שונות: $E(X^2)=\sum_a a^2P(X=a)=\frac{1}{4}\cdot 160^2+\frac{1}{2}\cdot 180^2+\frac{1}{4}\, 200^2=$

$=32,600$

$E^2(X) = 180^2$

$Var(X) = 32600 - 180^2 = 200$

הגדרה: סטיית התקן היא $\sigma = \sqrt{Var(X)}$ סיגמא.

טענה: $Var(aX+b) = a^2\,Var(X)$

התפלגות משותפת של משתנים מקריים

$$P(X=a, Y=b) = P(X^{-1}(a) \cap Y^{-1}(b))$$

פונק' של
a ו b

התפלגות משותפת מתלקת את Ω למטריצה של מאורעות זרים

<u>כללים:</u>

1. $$P(X=a) = \sum_b P(X=a, Y=b)$$

התפלגות שולית

2. $$\sum_a \sum_b P(X=a, Y=b) = 1$$

3. $$E(X) = \sum_a \sum_b a\, P(X=a, Y=b)$$

4. $$E(f(X,Y)) = \sum_a \sum_b f(a,b) P(X=a, Y=b)$$

5. הסתברות מותנית:
$$P(X=a \mid Y=b) = \frac{P(X=a, Y=b)}{P(Y=b)}$$

<u>דוגמה:</u> בכובע יש שלושה כרטיסים: 1,2,3. שולפים באקראי 2 בלי החזרה.

X - ערך הכרטיס הראשון

Y - ערך הכרטיס השני.

$P(X=a, Y=b) =$

Y \ X	1	2	3
1	0	1/6	1/6
2	1/6	0	1/6
3	1/6	1/6	0

$$P(X=2)=\sum_b P(X=2,Y=b) = \frac{1}{6}+0+\frac{1}{6}=\frac{1}{3}$$

$$P(X=2|Y=1)=\frac{P(X=2,Y=1)}{P(Y=1)} = \frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$$

משמעות: המאורעות $X=2$ ו $Y=1$ תלויים.

לפיכך X ו Y הם משתנים מקריים תלויים

הגדרה: משתנים מקריים X, Y בלתי תלויים, אם מתקיים שלכל a,b: המאורעות $X=a$, $Y=b$ בלתי תלויים

נסתכל על הגדרה $P(X=a|Y=1)$. זו בעצם פונק' התפלגות של $X|Y=1$ זה משתנה מקרי! יש לו גם תוחלת.

$$E(X|Y=1)=\sum_a aP(X=a|Y=1)=1\cdot 0+2\cdot\frac{1}{2}+2\cdot\frac{1}{2}+3\frac{1}{2}=\frac{5}{2}$$

הערה: מ"מ תלויים $\Longleftarrow$ תוחלות שונות בהינתן אחד מהם.
(לא הפוך)

נוסחה (נוסחת התוחלת השלמה):
$$E(X)=\sum_b E(X|Y=b)\cdot P(Y)$$

עבור X,Y ב"ת:
$$E(XY)=E(X)E(Y)$$

הגדרה: שונות משותפת של מ"מ X, Y הינו

$$cov(X,Y)=E((X-E(X))(Y-E(Y)))$$

או בקיצור

$$cov(X,Y)=E(XY)-E(X)E(Y)$$

הערה: $cov(X,X)=E((X-E(X))^2)=Var(X)$

הערה: X, Y ב"ת $\Longrightarrow$ $cov(X,Y)=0$

נחזור לדוגמה שלנו.

$$E(X)=E(Y)=2$$

$$E(XY)=\sum_a\sum_b ab\,P(X=a, Y=b)=\frac{19}{6}$$

$$cov(X,Y)=\frac{19}{6}-4=-\frac{5}{6}$$

מתאם שלילי

הגדרה: מקדם המתאם הוא

האות רו

$$\rho(X,Y)=\frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{cov(X,Y)}{\sigma(X)\sigma(Y)}$$

הערה: מקדם המתאם (בערך מוחלט) גדל ככל שהתלות יותר "לינארית"

הערה: $cov(X,Y)\le Var(X), Var(Y)$

שונות של סכום:

$$Var(X+Y) = E((X+Y)^2) - E^2(X+Y) =$$

$$= E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2 =$$

$$= E(X^2) + 2E(XY) + E(Y^2) - E^2(X) - 2E(X)E(Y) - E^2(Y) =$$

$$= \boxed{Var(X) + Var(Y) + 2cov(X,Y)}$$

עבור X,Y בלתי מתואמים: $Var(X+Y) = Var(X) + Var(Y)$

באופן כללי של $\{X_i\}$ בלתי מתואמים $Var(\sum X_i) = \sum Var(X_i)$

התפלגויות בדידות

התפלגות אחידה: כל אחד מהערכים מתקבל בהסתברות שווה.

לדוג': $X=1,\dots,n$.

$$P(X=a)=\begin{cases}\frac{1}{n}, & a\in[1,n]\\ 0, & \text{אחרת}\end{cases}$$

סימון: $X\sim U(1,n)$

מ"מ X — נוצר מהתפלגות — אחידה (uniform) — עם פרמטרים $1,\dots,n$

תוחלת:

$$E(X)=\sum_a aP(X=a)=\frac{1}{n}\sum_a a=\frac{1}{n}\cdot\frac{(1+n)\cdot n}{2}=\frac{n+1}{2}$$

שונות:

$$E(X^2)=\sum_a a^2P(X=a)=\frac{1}{n}\cdot\sum a^2=\frac{1}{n}\cdot\frac{n(n+1)(2n+1)}{6}=$$

$$=\frac{(n+1)(2n+1)}{6}$$

$$Var(X)=E(X^2)-E(X)=\frac{(n+1)(2n+1)}{6}-\left(\frac{n+1}{2}\right)^2=\frac{n^2-1}{12}$$

התפלגות ברנולי: מקבל רק 0 או 1. מסמן כישלון או הצלחה.

סימון: $X\sim b(p)$

p — הסתברות ההצלחה.

דוגמא: הטלת מטבע. נסמן פלי 0, עץ 1. $\Leftarrow X\sim b\left(\frac{1}{2}\right)$

אפשר גם $X\sim U(0,1)$

תוחלת:

$$E(X)=\sum_a aP(X=a)=0\cdot(1-p)+1\cdot p=p$$

$$E(X^2)=\sum_a a^2P(X=a)=0^2(1-p)+1^2\cdot p=p$$

$$Var(X)=E(X^2)-E^2(X)=p-p^2=p(1-p)$$

התפלגות בינומית

<u>הגדרה:</u> יהיו $Y_1,\dots,Y_n$ מ"מ ברנולי ב"ת, $Y_i\sim b(p)$. המ"מ $X=\sum Y_i$

מתפלג בינומית, וסימונו $X\sim Bin(n,p)$

X סופר את מס' ההצלחות

<u>דוגמה:</u> כמה פעמים נקבל 6 בהטלת 10 קוביות

נגדיר $X\sim Bin(\frac{1}{6},10)$. $X=0,\dots,10$

תוחלת: $$E(X)=E\left(\sum Y_i\right)=\sum E(Y_i)=\sum_{i=1}^n p=np$$

שונות: $$Var(X)=Var\left(\sum Y_i\right)\underset{\text{ב"ת}}{=}\sum Var(Y_i)=\sum p(1-p)=np(1-p)$$

<u>טענה:</u> יהי $X\sim Bin(n,p)$.

$$P(X=a)=\binom{n}{a}p^a(1-p)^{n-a}$$

$$E(X)=np \qquad Var(X)=np(1-p)$$

התפלגות גיאומטרית

הגדרה: ההתפלגות הגאומטרית של X - מס' ניסוי "ברנולי" הבלתי תלויים הנדרשים עד להשגת הצלחה אחת. ערכים אפשריים הם
$X = 1, 2, \ldots$

סימון: $X \sim G(p)$

פונק' ההתפלגות:
$$P(X=a) = (1-p)^{a-1}\, p$$

תוחלת:
$$E(X) = \sum_{a=1}^{\infty} a\, P(X=a) = \sum_{a=1}^{\infty} a(1-p)^{a-1} \cdot p = p\sum_{a=1}^{\infty} a(1-p)^{a-1} =$$

$$= p\sum_{a=1}^{\infty} \frac{d}{dq}(1-q)^{a} = p\sum_{a=0}^{\infty} \frac{d}{dq}(1-q)^{a} = p \cdot \frac{d}{dq}\left(\sum_{a=0}^{\infty}(1-q)^{a}\right) =$$

מתכנס

$$= p\,\frac{d}{dq}\left(\frac{1}{p}\right) = p \cdot \frac{1}{p^2} = \frac{1}{p}$$

שונות:
$$E(X^2 - X) = \sum_{a=1}^{\infty}(a^2-a)(1-p)^{a-1}p = p(1-p)\underbrace{\sum_{a=0}^{\infty} a(a-1)(1-p)^{a-2}}_{\left(\frac{1}{p}\right)''} =$$

$$= p(1-p)\left(\frac{1}{p}\right)'' \ldots p(1-p)\cdot\left(\frac{2}{p^3}\right) = \frac{2(1-p)}{p^2}$$

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

$$Var(X) = E(X^2) - E^2(X) = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2}$$

סיכום: $P(X=a)=(1-p)^{a-1}p \qquad E(X)=\frac{1}{p} \qquad Var(X)=\frac{1-p}{p^2}$

דוגמה: דני קונה כל יום כרטיס לוטו עם הסתברות זכייה של $p=10^{-6}$. בתוחלת תוך כמה זמן דני יזכה?

נגדיר $X\sim G(10^{-6})$

$$E(X)=\frac{1}{p}=10^6$$

מליון ימים לזכות.

מה הסיכוי שדני יזכה תוך יומיים או פחות?

$$P(X\le 2)=F_X(2)=P(X=1)+P(X=2)=10^{-6}+(1-10^{-6})10^{-6}$$
$$\approx 2\cdot 10^{-6}$$

אם מליון קונים מדי יום, מהי ההסתברות שמישהו יזכה תוך יומיים?

פתרון: יש 10^6 ניסויי ברנולי $b(2\cdot 10^{-6})$

נגדיר $Y\sim Bin(10^6, 2\cdot 10^{-6})$

$$P(Y=a)=\binom{n}{a}(1-p)^{n-a}p^a$$

$$P(Y\ge 1)=1-P(Y=0)=1-\binom{10^6}{0}(1-2\cdot 10^{-6})^0\cdot(2\cdot 10^{-6})^{10^6}=$$

הערה חשובה: מ"מ גאומטרי הוא חסר זיכרון. זה לא משנה כמה כשלונות כבר היו, ההתחלה של כמה ניסיונות עד להצלחה הבאה לא משתנה כלל

התפלגות פואסון

מוטיבציה: בעולם רגיל אי אפשר לדבר על ניסוי בכל דקה, אלא יש צורך לחלק את הזמן לחתיכות אינפיניטסימליות

התפלגות:

$$P(X=a)=\frac{\lambda^a}{a!}e^{-\lambda}$$

סימון: $X \sim Pois(\lambda)$

תוחלת: $E(X)=\lambda$

שונות: $Var(X)=\lambda$

דוגמה: בקילו בצק סרבבו 1000 פצפוצי שוקולד. הכינו מהבצק 100 עוגיות. דני בחר עוגיה באקראי. מה ההסתברות שיקבל עוגיה ללא פצפוצים?