משתנה אקראי (מקרי)

מטרה: לאפיין בצורה מתמטית תוצאה של ניסוי אקראי

הסתברות: שכיחות יחסית של תופעה כלשהי אחרי אינסוף ניסויים אקראיים

משתנה אקראי: משתנה אקראי (נקרא גם משתנה מקרי) זה מודל שמתאר קשר בין תוצאה של ניסוי אקראי למספר ממשי[א]. הסימון הוא ע"י אות גדולה, לדוגמה X, Y וכו'.

כ 9.3 ← מעבר על קבע (או מספר קבע)

בדיד ← ניתן למספר את התוצאות

דוגמה 1.1: (משתנה אקראי בדיד עם מספר תוצאות ניסוי סופיות) משתנה אקראי מסויים המקשר בין תוצאת זריקת קוביה, $\{1,2,3,4,5,6\}$, לבין הסתברות לתוצאה זו.

דוגמה 1.2: (משתנה אקראי בדיד עם מספר תוצאות ניסוי אינסופיות) משתנה אקראי מסויים המקשר בין מספר זריקות קוביה לבין הסתברות לקבל התוצאה "6".

דוגמה 1.3: (משתנה אקראי רציף) משתנה אקראי מסויים המקשר בין זמן העבודה של מכונה מסויימת להסתברות התקלה שלה.

משתנה אקראי בדיד

$$0 \leqslant p_X[x_i] \leqslant 1 \ \forall i$$

סה"כ ההסתברויות אפשריות של X

$$\sum_i p_X[x_i] = 1$$

Probability density function (PDF) (הגדרה 1.5):

קשר בין תוצאה של הניסוי x_k להסתברות של תוצאה זו:

$$p_X[x_k] = \Pr[X = x_k] \tag{1.1}$$

א מספר סופי
או אינסופי.

השימוש בסוגריים $[\cdot]$ הוא כדי להדגיש, שמדובר בערכים בדידים, לעומת $(\cdot)$, המשמש בהמשך לערכים רציפים.

שם של מ"א
את גדולה

תוצאת הניסוי
מס' K

דוגמה: X מתאר תוצאה
של זריקת קובייה

$$F_X(3.5) = P(X \leq 3.5)$$

$$= p_X[1] + p_X[2] + p_X[3] = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

Cumulative distribution function (CDF) (הגדרה 1.6):

הסתברות, שערך משתנה אקראי X קטן או שווה מערך הנתון x,

$$F_X(x) = \Pr(X \leqslant x), \quad x \in \mathbb{R}. \tag{1.4}$$

לא דווקא בדיד

ערך x הוא ממשי ורציף.

פונ' CDF היא לא יורדת,

$$\Pr(a < X \leqslant b) = F_X(b) - F_X(a)$$

$$F_X(-\infty) = 0, \quad F_X(\infty) = 1 \qquad 0 \leqslant F_X(x) \leqslant 1$$

תוחלת

* חישוב ממוסס מ"א - חישוב הסתברות Expectation
* ממוצע - חישוב מספרי

הגדרה:

סימון

$$E[X] = \sum_i x_i p_X[x_i]$$

מ"א

ערך

הסתברות
לערך

תכונות:

$$E[aX] = aE[X]$$

$$E[b] = b$$

$$E[g(X)] = \sum_i g(x_i) p_X[x_i]$$

$$g(x) = x^2 \Rightarrow E[x^2]$$

שונות (Variance)

$$\mathrm{Var}[X] = E\Big[(X - E[X])^2\Big]$$

(תוחלת של ריבוע מרחק
מ התוחלת)

* פיתוח נוסחה נוספת

$$Var[x] = E[x^2 - 2x\,E[x] + E^2[x]]$$

(1) $E[x^2]$

(2) $-2E[x E[x]] = -2E[x] \cdot E[x] = -2E^2[x]$

$$\mathrm{Var}[X] = E[X^2] - E^2[X]$$

$① \; E[x]$

$② \; -2E[x E[x]] = -2E[x]\cdot E[x] = -2E^2[x]$

מספר קבוע

$③ \; E[E^2[x]] = E^2[x]$

$$\mathrm{Var}[X] = E\left[X^2\right] - E^2[X]$$

$$\mathrm{Var}[aX+b] = a^2\mathrm{Var}[X]$$

$$\mathrm{Var}[b] = 0$$

$$E[x^2] = \underbrace{E[x\cdot x] \neq E[x]\cdot E[x]}= E^2[x]$$

התפלגויות נפוצות

1.3.1 התפלגות Bernoulli

מטרה: לתאר ניסוי בעל 2 תוצאות אפשריות:

- תוצאה "0" עם הסתברות $1-p$. כשלון
- תוצאה "1" עם הסתברות p. הצלחה

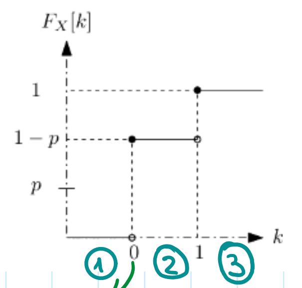
דוגמא: בריקת מטבע $p=\frac{1}{2}$

PDF

$$p_X[k] = \begin{cases} 1-p & k=0 \\ p & k=1 \end{cases}$$

CDF

$$F_X(x) = \begin{cases} 0 & x<0 & ① \\ 1-p & 0\leqslant x<1 & ② \\ 1 & 1\leqslant x & ③ \end{cases}$$



רציפות מימין

$F_x(0) = 1-p$

$F_x(0^-) = 0$

תוחלת

$$E[X] = \sum_i x_i p_X[x_i] = (k=0)p_X[0] + (k=1)p_X[1] = p$$

$0\cdot(1-p) + 1\cdot p$

שונות

$$\mathrm{Var}[X] = E[X^2] - E^2[X] = \underbrace{(k=0)^2 p_X[0] + (k=1)^2 p_X[1]}_{\text{משוואה (1.11)}} - p^2 = p - p^2 = p(1-p)$$

$0^2 \quad 1^2$

התפלגות בינומית

- נתונים n ניסויים בלתי תלויים.
- כל אחד בעל התפלגות Bernoulli עם סיכוי הצלחה p, $Ber(p)$.
- נדרשת ההתסברות ל-k תוצאות "1" בדיוק, מתוך סה"כ n תוצאות.

PDF

עבור $X\sim Bin(n,p)$ ניתן לפרק את החישוב ל-3 חלקים:

1. ישנם $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ אפשרויות ל-k-הצלחות, מתוך n ניסויים.
2. הסתברות של k הצלחות (תוצאות "1") היא p^k.
3. הסתברות של $n-k$ כשלונות (תוצאות "0") היא $(1-p)^{n-k}$.

מתוך איחוד של שלושת ההסתברויות,

$$p_X[k] = \underbrace{\binom{n}{k}}_{①\,②} \underbrace{p^k}_{②} \underbrace{(1-p)^{n-k}}_{③}$$

דוגמא:

זורקים קובייה 6 פעמים. חשב הסתברות לבדיוק שתי תוצאות "6"

$n=6 \quad p=\frac{1}{6}$

הסתברות תוצאת 6

$$X\sim Bin\left(6,\frac{1}{6}\right) \quad k=2$$

$$p_X[2] = \binom{6}{2}\left(\frac{1}{6}\right)^2\left(1-\frac{1}{6}\right)^{6-2} \approx 0.20$$

התפלגות גאומטרית

k ניסיונות $Ber(p)$

הצלחה היא רק בניסוי k (אחרי $k-1$ כשלונות)

$X\sim Geo(p)$

$$p_x[k] = (1-p)^{k-1}\cdot p \quad \text{PDF}$$

$$F_X[k] = 1-(1-p)^k \quad \text{CDF}$$

$X \sim Geo(p)$

$p_X[k] = (1-p) \cdot p$ PDF

$F_X[k] = 1-(1-p)^k$ CDF

דוגמא:

מהי הסתברות לא לקבל תוצאה 6 ב-6 זריקות הראשונות של קוביה?

דרך א' $X \sim Geo(1/6)$

- סיכוי ל-6 דווקא בזריקה ראשונה הוא $p_X[1] = \frac{1}{6}$.
- סיכוי ל-6 דווקא בזריקה שניה הוא $p_X[2] = \frac{5}{6} \cdot \frac{1}{6}$.
- סיכוי ל-6 דווקא בזריקה שלישית הוא $p_X[3] = \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6}$.
- סיכוי ל-6 דווקא בזריקה שישית הוא $p_X[6] = \left(\frac{5}{6}\right)^5 \cdot \frac{1}{6}$.

סיכוי לתוצאה שונה מ-6

סיכוי ל-6

לסיכום,

$$\Pr(X>6) = 1-\Pr(X \leqslant 6)$$
$$= 1-\left(p_X[1]+p_X[2]+\cdots+p_X[6]\right) \cong 0.33$$

דרך ב'

$$\Pr(X>6) = 1-F_X(6) = \left(1-\frac{1}{6}\right)^6 \approx 0.3349$$

$\Pr(X>x)+\Pr(X\leqslant x)=1$,

$\Pr(X>x) = 1-F_X(x)$

משתנה אקראי רציף

$F_X(x) = \Pr(X \leqslant x)$ זהה- CDF למ"א בדיד

$$f_X(x) = \frac{dF_X(x)}{dx} \quad \text{PDF}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(p)\,dp \quad \text{CDF}$$

תכונות

$\cancel{1 \geq}\ f_X(x) \geqslant 0$ — PDF כ.צ.פ היא לא "הסתברות"

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

הסתברות ל קטע

$$\Pr(a < X \leqslant b) = \int_a^b f_X(x)dx = F_X(b)-F_X(a)$$

הסתברות של ערך יחיד

$p(X=a)=0 \quad \forall a \in \mathbb{R}$ ← יש אינסוף ערכים אפשריים

שונות: זהה למ"א בדיד

תוחלת

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$

התפלגות אקספוננציאלית $X \sim Exp(\lambda)$

לאפיין תופעות עבורם הסתברות משתנה באופן אקספוננציאלי

PDF

$$f_X(x) = \begin{cases} \lambda \exp(-\lambda x) & x \geqslant 0 \\ 0 & \text{אחרת} \end{cases}$$

$$E[X] = \lambda \int_0^{\infty} x \exp(-\lambda x)dx = \frac{1}{\lambda}$$

$$E[X^2] = \int_0^{\infty} \lambda x^2 \exp(-\lambda x)dx = \frac{2}{\lambda^2}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(x)dx = \begin{cases} 1-\exp(-\lambda x) & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

דוגמה: יש למצוא $F_X(x)$, $E[X]$, $Var[X]$

דוגמה 2.3: מודל של הופעת תקלות במערכת מסויימת מתוארת ע"י התפלגות $X \sim Exp(\lambda)$, כאשר $\lambda = \frac{1}{1000}\left[\frac{1}{\text{שעה}}\right]$. שאלות:

(א) מהו זמן ממוצע לתקלה? $E[X] = \frac{1}{\lambda} = 1000[\text{שעות}]$

$X \sim Exp(\lambda)$, כאשר $\lambda = \frac{1}{1000}\left[\frac{1}{\text{שעה}}\right]$. שאלות:

(א) מהו זמן ממוצע לתקלה?

$$E[X] = \frac{1}{\lambda} = 1000[\text{שעות}]$$

(ג) מהו סיכוי לתקלה ב-100 שעות הראשונות של הפעילות?

$$p(X \leqslant 100) = F_X(100) \approx 0.0953$$

(ב) מהו סיכוי לתקלה במרווח זמן $[1000, \infty]$? (החל מ-1000)

$$p(X > 1000) = 1 - F_X(1000)$$

(ה) מהו משך הזמן עבורו הסיכוי לתקלה הוא $\frac{1}{2}$?

t_0

$\Pr(X > t_0) \pm \Pr(X \leqslant t_0) = 1/2$,

חציוני ערך $\Pr(X \leqslant x) = \Pr(X > x) = \frac{1}{2}$

$$F_X(x) = 1 - F_X(x) = \frac{1}{2}$$

שינוי נושא נוסחא

$$\Rightarrow x = \frac{\ln\left(\frac{1}{2}\right)}{-\frac{1}{1000}} \approx 693[\text{שעה}]$$

משכורת חציונית

⇐ חצי מקבלים פחות מזה

חצי " יותר מזה