# Тринаести час

★ Нека су $X$ и $Y$ случајне величине. **Коваријација (коваријанса) случајних величина** $X$ и $Y$ је број:

$$cov(X,Y) = E[(X-EX)(Y-EY)] = E(XY) - EXEY.$$

- Ако је коваријација једнака нули, онда су случајне величине **некорелисане**.
- Ако су случајне величине независне, онда је коваријација једнака нули. Обрнуто тврђење не важи.

★ **Коефицијент корелације** случајних величина $X$ и $Y$ је:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{DXDY}}.$$

Особине коефицијента корелације:

- $-1 \le \rho_{X,Y} \le 1$
- $|\rho_{X,Y}| = 1 \Leftrightarrow Y = aX + b$
- $\rho_{X,Y} = 0 \Rightarrow X, Y$ некорелисане
- $X, Y$ независне $\Rightarrow \rho_{X,Y} = 0$

**1.** *Случајна величина $X$ има униформну $U[-1,1]$ расподелу. Ако је $Y = sgnX$, израчунати коефицијент корелације случајних величина $X$ и $Y$.*

Тражи се да се одреди коефицијент корелације случајних величина $X$ и $Y$. Он се одређује према формули:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{DXDY}}.$$

Дакле, треба одредити: $EX, EY, DX, DY, EXY$.

Случајна величина $X : U[-1,1]$, због тога је густина расподеле: $f_X(x) = \begin{cases} \frac{1}{2}, & x \in [-1,1] \\ 0, & \text{иначе.} \end{cases}$

Према формулама за очекивање и дисперзију добија се: $EX = 0, DX = \frac{1}{3}$.

С друге стране, пошто је $Y = sgnX$, то значи да је: $Y = sgnX = \begin{cases} -1, & X < 0 \\ 0, & X = 0 \\ 1, & X > 0 \end{cases}$

Одређујемо расподелу случајне величине $Y$:

- $P\{Y=-1\} = P\{X<0\} = \int\limits_{-1}^{0} \frac{1}{2}dx = \frac{1}{2}$
- $P\{Y=0\} = P\{X=0\} = 0$ [1]
- $P\{Y=1\} = P\{X>0\} = \int\limits_{0}^{1} \frac{1}{2}dx = \frac{1}{2}$

Дакле, $Y : \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$. Из расподеле за $Y$ добијамо да је: $EY = -\frac{1}{2} + \frac{1}{2} = 0, DY = EX^2 - (EX)^2 = EX^2 = \frac{1}{2} + \frac{1}{2} = 1$.

Још треба одредити $EXY$: $EXY = E(XsgnX) = \int\limits_{-1}^{1} xsgnx\frac{1}{2}dx = 2\int\limits_{0}^{1} xsgnx\frac{1}{2}dx = \frac{1}{2}$. [2]

Дакле,

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{DXDY}} = \frac{\frac{1}{2} - 0}{\sqrt{\frac{1}{3} \cdot 1}} = \frac{\sqrt{3}}{2}.$$

[1] вероватноћа да непрекидна случајна величина узме вредност у конкретној тачки је 0

[2] при решавању интеграла користи се парност подинтегралне функције; други начин за решавање је раздвајање на два интеграла

★ **Расподеле за минимум и максимум:**

Нека су $X_1, X_2, ..., X_n$ независне случајне величине са истом апсолутно непрекидном расподелом, чија је функција расподеле $F(x)$ и густином $f(x) = F'(x)$. Нека су $X_{(1)} = \min\{X_1, X_2, ..., X_n\}$ и $X_{(n)} = \max\{X_1, X_2, ..., X_n\}$. Одредити расподеле за $X_{(1)}, X_{(2)}$ и дводимензиону расподелу за $(X_{(1)}, X_{(2)})$.

- $$\begin{aligned} F_{X_{(1)}}(x) &= P\{X_{(1)} \le x\} \\ &= P\{\min\{X_1, X_2, ..., X_n\} \le x\} \\ &= 1 - P\{\min\{X_1, X_2, ..., X_n\} > x\} \\ &\overset{3}{=} 1 - P\{X_1 > x, X_2 > x, ..., X_n > x\} \\ &\overset{4}{=} 1 - P\{X_1 > x\}P\{X_2 > x\} \cdots P\{X_n > x\} \\ &\overset{5}{=} 1 - (P\{X_1 > x\})^n \\ &= 1 - (1 - F(x))^n \end{aligned}$$
  $\Rightarrow f_{X_{(1)}}(x) = F'_{X_{(1)}}(x) = n(1 - F(x))^{n-1} f(x)$
- $$\begin{aligned} F_{X_{(n)}}(x) &= P\{X_{(n)} \le x\} \\ &= P\{\max\{X_1, X_2, ..., X_n\} \le x\} \\ &\overset{6}{=} P\{X_1 \le x, X_2 \le x, ..., X_n \le x\} \\ &\overset{7}{=} P\{X_1 \le x\}P\{X_2 \le x\} \cdots P\{X_n \le x\} \\ &\overset{8}{=} (P\{X_1 \le x\})^n \\ &= (F(x))^n \end{aligned}$$
  $\Rightarrow f_{X_{(n)}}(x) = F'_{X_{(n)}}(x) = n(F(x))^{n-1} f(x)$
- $$\begin{aligned} F_{X_{(1)},X_{(n)}}(x,y) &= P\{X_{(1)} \le x, X_{(n)} \le y\} \\ &\overset{9}{=} P\{X_{(n)} \le y\} - P\{X_{(1)} > x, X_{(n)} \le y\} \\ &= (F(n))^n - P\{X_1 > x, ..., X_n > x, X_1 \le y, ..., X_n \le y\} \\ &= (F(n))^n - P\{x < X_1 \le y, ..., x < X_n \le y\} \\ &\overset{10}{=} (F(n))^n - P\{x < X_1 \le y\} \cdots P\{x < X_n \le y\} \\ &\overset{11}{=} (F(n))^n - (F(y) - F(x))^n, \text{ за } x \le y \end{aligned}$$
  $\Rightarrow f_{X_1,X_{(n)}}(x,y) = \frac{\partial^2 F_{X_{(1)},X_{(n)}}(x,y)}{\partial x \partial y} = \frac{\partial}{\partial y}(n(F(y) - F(x))^{n-1} f(x)) = n(n-1)(F(y) - F(x))^{n-2} f(x) f(y)$, за $x \le y$.

[3] ако је минимум већи од неке вредности, онда су све вредности од којих се рачуна минимум веће од те вредности

[4] случајне величине су независне

[5] све случајне величине имају исту расподелу

[6] ако је максимим мањи од неке вредности, онда су све вредности од којих се рачуна максимум мање од те вредности

[7] случајне величине су независне

[8] све случајне величине имају исту расподелу

[9] формула потпуне вероватноће: $P\{X_{(n)} \le y\} = P\{X_{(1)} \le x, X_{(n)} \le y\} + P\{X_{(1)} > x, X_{(n)} \le y\}$

[10] случајне величине су независне

[11] све случајне величине имају исту расподелу

**2.** *Ако су $X_1, X_2, ..., X_n$ независне случајне величине са истом униформном $U[0,1]$ расподелом и ако је $X_{(1)} = \min\{X_1, X_2, ..., X_n\}$, а $X_{(n)} = \max\{X_1, X_2, ..., X_n\}$, израчунати коефицијент корелације случајних величина $X_{(1)}$ и $X_{(n)}$.*

Тражи се да се одреди коефицијент корелације случајних величина $X$ и $Y$. Он се одређује према формули:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{DXDY}}.$$

Дакле, треба одредити: $EX, EY, DX, DY, EXY$.
Случајне величине $X_1, X_2, ..., X_n : U[0,1]$, због тога је функција расподеле сваке од њих:

$F_X(x) = \begin{cases} 0, & x < 0 \\ x, & x \in [0,1] \\ 1, & x > 1 \end{cases}$, а густина: $f_X(x) \begin{cases} 1, & x \in [0,1] \\ 0, & \text{иначе.} \end{cases}$

Према претходно изведеним формулама (уз замену датих функција густине и расподеле) следи:

$f_{X_{(1)}} = n(1-x)^{n-1}, 0 \le x \le 1$

$f_{X_{(n)}} = nx^{n-1}, 0 \le x \le 1$

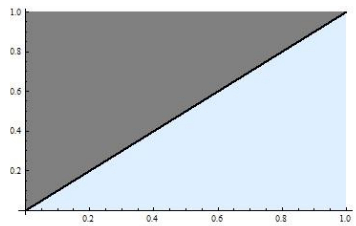
Израчунаћемо потребне вредности:

- $EX_{(1)} = \int\limits_0^1 xn(1-x)^{n-1}dx = \left| \begin{array}{c} 1-x=t \\ -dx=dt \end{array} \right| = n\int\limits_0^1 (1-t)t^{n-1}dt = 1 - \frac{n}{n+1} = \frac{1}{n+1}$
- $DX_{(1)} = EX_{(1)}^2 - (EX_{(1)})^2$

$$\begin{aligned}
&= \int\limits_0^1 x^2 n(1-x)^{n-1}dx - (\tfrac{1}{n+1})^2 = \left| \begin{array}{c} 1-x=t \\ -dx=dt \end{array} \right| \\
&= n\int\limits_0^1 (1-t)^2 t^{n-1}dt - (\tfrac{1}{n+1})^2 \\
&= n\int\limits_0^1 (1-2t+t^2)t^{n-1}dt - (\tfrac{1}{n+1})^2 \\
&= 1 - \tfrac{2n}{n+1} + \tfrac{n}{n+2} - (\tfrac{1}{n+1})^2 \\
&= \tfrac{2}{(n+1)(n+2)} - (\tfrac{1}{n+1})^2 \\
&= \tfrac{n}{(n+1)^2(n+2)}
\end{aligned}$$

- $EX_{(n)} = \int\limits_0^1 xnx^{n-1}dx = n\int\limits_0^1 x^n dt = \frac{n}{n+1}$
- $DX_{(n)} = EX_{(n)}^2 - (EX_{(n)})^2 = \int\limits_0^1 x^2 nx^{n-1}dx - (\frac{n}{n+1})^2 = \frac{n}{n+2} - (\frac{n}{n+1})^2 = \frac{n}{(n+1)^2(n+2)}$
- $E(X_{(1)}, X_{(n)}) = \int\limits_D\int xyf(x,y)dxdy$

$$\begin{aligned}
&\overset{12}{=} \int\limits_0^1\int\limits_x^1 xyn(n-1)(y-x)^{n-2}dydx = \left| \begin{array}{c} y-x=t \\ dy=dt \end{array} \right| \\
&\overset{13}{=} \int\limits_0^1\int\limits_0^{1-x} x(x+t)n(n-1)t^{n-2}dtdx \\
&= n(n-1)\int\limits_0^1 (x^2\tfrac{(1-x)^{n-1}}{n-1} + x\tfrac{(1-x)^n}{n})dx = \left| \begin{array}{c} 1-x=t \\ -dx=dt \end{array} \right| \\
&= n\int\limits_0^1 (1-t)^2 t^{n-1}dt + (n-1)\int\limits_0^1 (1-t)t^n dt \\
&= \tfrac{2}{(n+1)(n+2)} + (n-1)\tfrac{1}{n+1} + (n-1)\tfrac{1}{n+2} \\
&= \tfrac{1}{n+2}
\end{aligned}$$




Дакле,

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{DXDY}} = \frac{\frac{1}{n+2} - \frac{1}{n+1}\frac{n}{n+1}}{\sqrt{\frac{n}{(n+1)^2(n+2)}\frac{n}{(n+1)^2(n+2)}}} = \frac{1}{n}.$$

---

[12]користи се раније израчуната формула; област $D$ је тамна област на графикону
[13]у унутрашњем интегралу $x$ се посматра као константа

**3.** *Случајна величина $X$ има униформну $U[0,1]$ расподелу. Ако је $Y_n = \sum\limits_{k=1}^{n} \frac{k-1}{n} I\{\frac{k-1}{n} < X \leq \frac{k}{n}\}, n \in N, n > 1$. Израчунати коефицијент корелације случајних величина $X$ и $Y_n$.*

Тражи се да се одреди коефицијент корелације случајних величина $X$ и $Y$. Он се одређује према формули:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{DXDY}}.$$

Дакле, треба одредити: $EX, EY, DX, DY, EXY$.

Случајна величина $X : U[0,1]$, због тога је густина расподеле: $f_X(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & \text{иначе.} \end{cases}$, а функција расподеле $F_X(x) = \begin{cases} 0, & x < 0 \\ x, & x \in [0,1] \\ 1, & x > 1 \end{cases}$. Према формулама за очекивање и дисперзију добија се: $EX = \frac{1}{2}, DX = \frac{1}{12}$.

С друге стране:

- $$\begin{aligned} EY_n &= E(\sum_{k=1}^{n} \tfrac{k-1}{n} I\{\tfrac{k-1}{n} < X \leq \tfrac{k}{n}\}) \\ &\overset{14}{=} \sum_{k=1}^{n} \tfrac{k-1}{n} E(I\{\tfrac{k-1}{n} < X \leq \tfrac{k}{n}\}) \\ &\overset{15}{=} \sum_{k=1}^{n} \tfrac{k-1}{n} P\{\tfrac{k-1}{n} < X \leq \tfrac{k}{n}\} \\ &= \sum_{k=1}^{n} \tfrac{k-1}{n} (F(\tfrac{k}{n}) - F(\tfrac{k-1}{n})) \\ &\overset{16}{=} \sum_{k=1}^{n} \tfrac{k-1}{n} (\tfrac{k}{n} - \tfrac{k-1}{n}) \\ &= \sum_{k=1}^{n} \tfrac{k-1}{n^2} \\ &= \tfrac{1}{n^2} \sum_{k=1}^{n-1} k \\ &= \tfrac{1}{n^2} \tfrac{(n-1)n}{2} \\ &= \tfrac{n-1}{2n} \end{aligned}$$

- $$\begin{aligned} EY_n^2 &= E(\sum_{k=1}^{n} \tfrac{k-1}{n} I\{\tfrac{k-1}{n} < X \leq \tfrac{k}{n}\})^2 \\ &\overset{17}{=} E(\sum_{k=1}^{n} (\tfrac{k-1}{n})^2 I\{\tfrac{k-1}{n} < X \leq \tfrac{k}{n}\}) \\ &= \sum_{k=1}^{n} (\tfrac{k-1}{n})^2 EI\{\tfrac{k-1}{n} < X \leq \tfrac{k}{n}\} \\ &= \sum_{k=1}^{n} \tfrac{(k-1)^2}{n^3} \\ &= \tfrac{1}{n^3} \sum_{k=1}^{n} (k-1)^2 \\ &= \tfrac{1}{n^3} \sum_{k=0}^{n-1} k^2 \\ &= \tfrac{1}{n^3} \tfrac{(n-1)n(2n-1)}{6} \\ &= \tfrac{2n^2-3n+1}{6n^2} \end{aligned}$$

[14] особине очекивања: очекивање суме је сума очекивања и очекивање од константе је та константа

[15] $I_A : \begin{pmatrix} 0 & 1 \\ 1-P(A) & P(A) \end{pmatrix}$, $EI_A = P(A)$

[16] $\frac{k}{n} \in (0,1)$ јер је $k \leq n$ због сумирања по к

[17] због дисјунктности интервала у индикатору када се помноже различити индикатори добија се 0, јер није могуће да $x$ буде у два различита интервала; квадрат индикатора има исту расподелу као индикатор

- $DY_n = EY_n^2 - (EY_n)^2 = \frac{2n^2-3n+1}{6n^2} - (\frac{n-1}{2n})^2 = \frac{n^2-1}{12n^2}$
- $EXY_n = E(X\sum\limits_{k=1}^{n}\frac{k-1}{n}I\{\frac{k-1}{n} < X \leq \frac{k}{n}\})$

  $= \sum\limits_{k=1}^{n}\frac{k-1}{n}E(XI\{\frac{k-1}{n} < X \leq \frac{k}{n}\})$

  $= \sum\limits_{k=1}^{n}\frac{k-1}{n}\int\limits_{0}^{1}xI\{\frac{k-1}{n} < X \leq \frac{k}{n}\}dx$

  $= \sum\limits_{k=1}^{n}\frac{k-1}{n}\int\limits_{\frac{k-1}{n}}^{\frac{k}{n}}xdx$

  $= \sum\limits_{k=1}^{n}\frac{k-1}{n}((\frac{k}{n})^2 - (\frac{k-1}{n})^2)$

  $= \frac{1}{2n^3}\sum\limits_{k=1}^{n}(2k^2 - 3k + 1)$

  $= \frac{1}{2n^3}(2\frac{(n-1)n(2n-1)}{6} - 3\frac{(n-1)n}{2} + n)$

  $= \frac{4n^2-3n-1}{12n^2}$

  Дакле,

$$\rho_{X,Y_n} = \frac{cov(X,Y_n)}{\sqrt{DXDY}} = \frac{\frac{4n^2-3n-1}{12n^2} - \frac{n-1}{4n}}{\sqrt{\frac{1}{12}\frac{n^2-1}{12n^2}}} = \frac{\sqrt{n^2-1}}{\sqrt{n^2}} = \sqrt{1-\frac{1}{n^2}} \to 1, \text{при } n \to \infty$$