

1. Pravděpodobnost narození dcery je stejná jako pravděpodobnost narození syna. Náhodná veličina X udává počet dcer v náhodně vybrané rodině se třemi dětmi, veličina Y udává počet starších bratrů nejmladšího dítěte v téže rodině.

Nejprve si udělám tabulku, jak jednotlivé rodiny mohou vypadat. 1 je dcera, 0 je syn.

(0, 0, 0)
(0, 0, 1)
(0, 1, 0)
(0, 1, 1)
(1, 0, 0)
(1, 0, 1)
(1, 1, 0)
(1, 1, 1)

a) Odvoďte rozdělení náhodného vektoru $(X, Y)^T$.

Napišeme si tabulku, uvnitř níž bude $P(X = i, Y = j)$.

$X \setminus Y$	0	1	2	Σ
0	0	0	$\frac{1}{8}$	$\frac{1}{8}$
1	0	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{3}{8}$
2	$\frac{1}{8}$	$\frac{1}{4}$	0	$\frac{3}{8}$
3	$\frac{1}{8}$	0	0	$\frac{1}{8}$
Σ	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

b) Jaká jsou marginální rozdělení veličin X a Y ?

$P(X = i)$ odpovídá pravděpodobnosti ve sloupci označeném Σ .

$P(Y = j)$ odpovídá pravděpodobnosti v řádku označeném Σ .

c) Jsou veličiny X a Y nezávislé?

Pokud by byli nezávislé, muselo by platit $P(X = i, Y = j) = P(X = i) \cdot P(Y = j) \forall i, j$. Když ale vezmeme například $i = 0, j = 0$, zjistíme, že nejsou nezávislé, protože dle tabulky $P(X = 0, Y = 0) = 0 \neq P(X = 0) \cdot P(Y = 0) = \frac{1}{8} \cdot \frac{1}{4}$.

d) Spočítejte kovariaci X a Y . Jaký je vztah mezi nezávislostí dvou veličin a jejich kovariací?

Nejprve si musíme spočítat EXY, EX a EY , abychom mohli určit kovariaci.

$$EXY = \sum_{i=0}^3 \sum_{j=0}^2 P(X = i, Y = j) \cdot i \cdot j = \frac{1}{4} \cdot 1 \cdot 1 + \frac{1}{8} \cdot 1 \cdot 2 + \frac{1}{4} \cdot 2 \cdot 1 = 1$$

$$EX = \sum_{i=0}^3 P(X = i) \cdot i = \frac{3}{8} \cdot 1 + \frac{3}{8} \cdot 2 + \frac{1}{8} \cdot 3 = \frac{3}{2}$$

$$EY = \sum_{j=0}^2 P(Y = j) \cdot j = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 = 1$$

$$\text{Cov}(X, Y) = EXY - (EX)(EY) = 1 - \frac{3}{2} \cdot 1 = -\frac{1}{2}$$

Vztah mezi nezávislostí dvou veličin a jejich kovariací je ten, že pokud jsou nezávislé, tak $\text{Cov}(X, Y) = 0$. Opačně to už platit nemusí.

$$\begin{aligned} A \text{ je to díky } \text{Cov}(X, Y) &= EXY - (EX)(EY) = \sum_i \sum_j P(X = i, Y = j) \cdot i \cdot j - \left(\sum_i P(X = i) \cdot i \right) \cdot \\ &\left(\sum_j P(Y = j) \cdot j \right) \stackrel{\text{nezávislost}}{=} \sum_i \sum_j P(X = i) \cdot P(Y = j) \cdot i \cdot j - \left(\sum_i P(X = i) \cdot i \right) \cdot \left(\sum_j P(Y = j) \cdot j \right) = \\ &\left(\sum_i P(X = i) \cdot i \right) \left(\sum_j P(Y = j) \cdot j \right) - \left(\sum_i P(X = i) \cdot i \right) \cdot \left(\sum_j P(Y = j) \cdot j \right) = 0 \end{aligned}$$

e) Spočítejte korelační koeficient mezi X a Y

Nejprve musím spočítat $\text{Var } X$ a $\text{Var } Y$

$$\text{Var } X = EX^2 - (EX)^2 = \left(\sum_{i=0}^3 P(X = i) \cdot i^2 \right) - \left(\sum_{i=0}^3 P(X = i) \cdot i \right)^2 = 3 - \left(\frac{3}{2} \right)^2 = \frac{3}{4}$$

$$\text{Var } Y = EY^2 - (EY)^2 = \left(\sum_{j=0}^2 P(Y = j) \cdot j^2 \right) - \left(\sum_{j=0}^2 P(Y = j) \cdot j \right)^2 = \frac{3}{2} - 1^2 = \frac{1}{2}$$

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var } X \cdot \text{Var } Y}} = \frac{-\frac{1}{2}}{\sqrt{\frac{3}{4} \cdot \frac{1}{2}}} \approx -0.816$$

2. Náhodná veličina X udává dobu, kterou strávíte čekáním na tramvaj na Malostranském náměstí (v minutách) a náhodná veličina Y udává dobu kterou následně strávíte čekáním na metro A ve stanici Malostranská (také v minutách). Ze zkušenosti víme, že náhodný vektor $(X, Y)^T$ má spojitě rozdělení s hustou.

$$f(x, y) = \frac{1}{2} e^{-x - \frac{y}{2}} \cdot \Pi_{[x > 0, y > 0]}$$

- a) Jaké je rozdělení jednotlivých dob čekání (na tramvaj a na metro zvlášť)?
V této otázce se nás ptají na marginální rozdělení pro veličinu X a Y .

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^{\infty} \frac{1}{2} e^{-x-\frac{y}{2}} dy = \frac{1}{2} e^{-x} \int_0^{\infty} e^{-\frac{y}{2}} dy = e^{-x} \left[e^{-\frac{y}{2}} \right]_0^{\infty} = e^{-x} \cdot \Pi_{[x>0]}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_0^{\infty} \frac{1}{2} e^{-x-\frac{y}{2}} dx = \frac{1}{2} e^{-\frac{y}{2}} \int_0^{\infty} e^{-x} dx = \frac{1}{2} e^{-\frac{y}{2}} \left[e^{-x} \right]_0^{\infty} = \frac{1}{2} e^{-\frac{y}{2}} \cdot \Pi_{[y>0]}$$

- b) Spočítejte kovarianci X a Y .

Nejprve spočítáme EXY , EX a EY .

$$EXY = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cdot x \cdot y dx dy = \int_0^{\infty} \int_0^{\infty} \frac{1}{2} e^{-x-\frac{y}{2}} \cdot x \cdot y dx dy = \frac{1}{2} \int_0^{\infty} e^{-\frac{y}{2}} \cdot y \int_0^{\infty} e^{-x} \cdot x dx dy = 2$$

$$EX = \int_{-\infty}^{\infty} f_X(x) x dx = \int_0^{\infty} e^{-x} x dx = 1$$

$$EY = \int_{-\infty}^{\infty} f_Y(y) y dy = \int_0^{\infty} \frac{1}{2} e^{-\frac{y}{2}} y dy = 2$$

$$\text{Cov}(X, Y) = EXY - (EX)(EY) = 2 - 1 \cdot 2 = 0$$

- c) Jsou doby strávené čekáním na tramvaj a na metro nezávislé?

Jelikož $\text{Cov}(X, Y) = 0$, tak mohou být nezávislé. Musíme tedy zjistit jestli $f(x, y) = f_X(x) \cdot f_Y(y)$.

$$f(x, y) = \frac{1}{2} e^{-x-\frac{y}{2}} \cdot \Pi_{[x>0, y>0]} \stackrel{?}{=} f_X(x) \cdot f_Y(y) = e^{-x} \cdot \Pi_{[x>0]} \cdot \frac{1}{2} e^{-\frac{y}{2}} \cdot \Pi_{[y>0]} = \frac{1}{2} e^{-x-\frac{y}{2}} \cdot \Pi_{[x>0, y>0]}$$

Tedy jsou nezávislé.

- d) Jaká je střední doba Vašeho celkového čekání na dopravní prostředky?

$$E(X + Y) = EX + EY = 2 + 1 = 3$$

3. Náhodná veličina X má rovnoměrné rozdělení na intervalu $(-1, 1)$. Označme $Y = X^2$. Spočítejte kovarianci veličin X a Y a jejich korelační koeficienty ρ_{XY} . Jsou X a Y nezávislé?

Víme, že $f_X(x) = \frac{1}{2} \Pi_{[x \in (-1, 1)]}$

$$EX = \int_{-1}^1 \frac{1}{2} x dx = \frac{1}{2} \int_{-1}^1 x dx = \frac{1}{2} \left[\frac{x^2}{2} \right]_{-1}^1 = \frac{1}{2} \cdot \left(\frac{(1)^2}{2} - \frac{(-1)^2}{2} \right) = 0$$

$$EY = EX^2 = \int_{-1}^1 \frac{1}{2} x^2 dx = \frac{1}{2} \left[\frac{x^3}{3} \right]_{-1}^1 = \frac{1}{2} \cdot \left(\frac{(1)^3}{3} - \frac{(-1)^3}{3} \right) = \frac{1}{3}$$

$$EXY = EXX^2 = EX^3 = \int_{-1}^1 \frac{1}{2} x^3 dx = \frac{1}{2} \int_{-1}^1 x^3 dx = \frac{1}{2} \left[\frac{x^4}{4} \right]_{-1}^1 = \frac{1}{2} \left(\frac{(1)^4}{4} - \frac{(-1)^4}{4} \right) = 0$$

$$\text{Cov}(X, Y) = EXY - (EX)(EY) = 0 - 0 \cdot \frac{1}{3} = 0$$

$$EY^2 = EX^4 = \int_{-1}^1 \frac{1}{2} x^4 dx = \frac{1}{2} \int_{-1}^1 x^4 dx = \frac{1}{2} \left[\frac{x^5}{5} \right]_{-1}^1 = \frac{1}{2} \left(\frac{(1)^5}{5} - \frac{(-1)^5}{5} \right) = \frac{1}{5}$$

$$\text{Var } X = EX^2 - (EX)^2 = \frac{1}{3}$$

$$\text{Var } Y = EY^2 - (EY)^2 = \frac{1}{5} - \frac{1}{3^2} = \frac{4}{45}$$

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var } X \cdot \text{Var } Y}} = \frac{0}{\sqrt{\frac{1}{3} \cdot \frac{4}{45}}} = 0$$

Nejsou nezávislé, protože pro nezávislé veličiny musí platit:

$P(X \in B, Y \in C) = P(X \in B)P(Y \in C) \forall B, C \subset R$ podobně jako pro náhodné jevy.

Potom můžeme zvolit třeba $B = \left\langle -\frac{1}{2}, \frac{1}{2} \right\rangle$, $C = \left\langle 0, \frac{1}{4} \right\rangle$ a dostaneme

$$P(X \in B, Y \in C) = P(X \in B, X^2 \in C) = P(X \in B, X \in B) = \frac{1}{2} \neq \frac{1}{2} \cdot \frac{1}{2} = P(X \in B)P(X \in C)$$

4. Zmatená šatnářka náhodně přiřadí n pánům jejich klobouky. Náhodná veličina X_i je indikátor jevu, zda má i -tý pán správný klobouk, $i = 1, \dots, n$ (tj. $X_i = 1$, pokud i -tý pán má svůj klobouk a $X_i = 0$ jinak).

a) Spočítejte střední hodnotu a rozptyl veličiny X_i pro pevně zvolené i .

$X_i = 0$ s pravděpodobností $P(X_i = 0) = \frac{(n-1) \cdot (n-1)!}{n!} = \frac{n-1}{n}$, protože i -tý klobouk mohou dát na $n-1$ míst a zbytek mohou libovolně rozdělit mezi zbylé pány. Nebo ekvivalentně právě $n-1$ (ne i -tý) mohou dát i -tému pánovi z n klobouků.

$X_i = 1$ s pravděpodobností $P(X_i = 1) = \frac{1 \cdot (n-1)!}{n!} = \frac{1}{n}$, protože i -tý klobouk musí mít i -tý pán a zbytek mohou libovolně rozdělit mezi zbylé pány. Nebo ekvivalentně právě 1 (i -tý) musím dát i -tému pánovi z n klobouků.

$$EX_i = \sum_{j=0}^1 P(X_i = j) \cdot j = \frac{n-1}{n} \cdot 0 + \frac{1}{n} \cdot 1 = \frac{1}{n}$$

$$EX_i^2 = \sum_{j=0}^1 P(X_i = j) \cdot j^2 = \frac{n-1}{n} \cdot 0^2 + \frac{1}{n} \cdot 1^2 = \frac{1}{n}$$

$$\text{Var } X_i = EX_i^2 - (EX_i)^2 = \frac{1}{n} - \frac{1}{n^2} = \frac{n-1}{n^2}$$

b) Jsou veličiny X_i a X_j pro pevně zvolené i, j nezávislé?

Pro výpočty nám stačí znát pravděpodobnost, že $P(X_i = 1, X_j = 1) = \frac{1}{n} \cdot \frac{1}{n-1}$, protože i -tý musím dát i -tému pánovi z n klobouků a j -tý musím dát j -tému z $n-1$ klobouků.

Pokud by byla X_i a X_j nezávislá, pak by platilo $P(X_i = i, X_j = j) = P(X_i = i) \cdot P(X_j = j) \forall i, j$, což neplatí např. pro $i = 1, j = 1$

$$P(X_i = 1, X_j = 1) = \frac{1}{n \cdot (n-1)} \neq P(X_i = 1) \cdot P(X_j = 1) = \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2}$$

c) Spočítejte kovarianci X_i a X_j .

$EX_i X_j = P(X_i = 1, X_j = 1)$, protože pro ostatní hodnoty je i nebo j rovné 0, a tedy $EX_i X_j$ neovlivní.

$$\text{Cov}(X_i, X_j) = EX_i X_j - (EX_i)(EX_j) = \frac{1}{n \cdot (n-1)} - \frac{1}{n^2} = \frac{1}{n^2 - n} - \frac{1}{n^2} = \frac{n^2 - (n^2 - n)}{n^2(n^2 - n)} = \frac{1}{n^3 - n^2}$$

d) Na základě výše uvedených výsledků spočítejte očekávanou hodnotu a rozptyl počtu správně přiřazených klobouků. (Využijte fakt, že počet správně přiřazených klobouků X lze vyjádřit jako $X = \sum_{i=1}^n X_i$)

Střední hodnota součtu je součet středních hodnot.

$$EX = E \sum_{i=1}^n X_i = \sum_{i=1}^n EX_i = n \cdot \frac{1}{n} = 1$$

Dále využijeme poznámky uvedené v „Další užitečné vlastnosti“ pro rozptyl. ($\text{Var}(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i^2 \text{Var } X_i + \sum_{i=1}^n \sum_{i \neq j} \text{Cov}(X_i, X_j)$)

My už z dřívějška víme rozptyly i kovariance. Zbývá je dosadit do výše zmíněné rovnice, kde $a_i = 1 \forall i \in \{1, \dots, n\}$

$$\text{Var } X = \text{Var} \sum_{i=1}^n X_i = \sum_{i=1}^n \text{Var } X_i + \sum_{i=1}^n \sum_{i \neq j} \text{Cov}(X_i, X_j) = n \cdot \frac{1}{n} + n \cdot (n-1) \cdot \frac{1}{n^2(n-1)} = 1 + \frac{1}{n}$$