

Rozwiązania zadań C09

Zad.1. Rozkład prawdopodobieństwa wektora losowego (X, Y) jest następujący:

	X		
Y	0	1	2
0	$\frac{1}{8} + \varepsilon$	0	$\frac{3}{8} - \varepsilon$
1	$\frac{1}{4} - \varepsilon$	$\frac{1}{4}$	ε

gdzie ε jest dowolną liczbą rzeczywistą z przedziału $[0, 1/4]$.

- Wyznaczyć rozkłady brzegowe zmiennych losowych X i Y .
- Czy zależą one od wartości ε ?
- Wyciągnąć wnioski!.

Rozw.

•

	X			
Y	0	1	2	$f_Y(y)$
0	$\frac{1}{8} + \varepsilon$	0	$\frac{3}{8} - \varepsilon$	$\frac{4}{8}$
1	$\frac{1}{4} - \varepsilon$	$\frac{1}{4}$	ε	$\frac{4}{8}$
$f_X(x)$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	1

•

$$f_X(0) = \frac{1}{8} + \varepsilon + \frac{1}{4} - \varepsilon = \frac{3}{8}, \quad f_X(1) = \frac{1}{4}, \quad f_X(2) = \frac{3}{8} - \varepsilon + \varepsilon = \frac{3}{8}$$

$$f_Y(0) = \frac{1}{8} + \varepsilon + 0 + \frac{3}{8} - \varepsilon = \frac{4}{8} = \frac{1}{2}, \quad f_Y(1) = \frac{1}{4} - \varepsilon + \frac{1}{4} + \varepsilon = \frac{1}{2}$$

x	0	1	2
$f_X(x)$	3/8	1/4	3/8

y	0	1
$f_Y(y)$	1/2	1/2

Rozkłady brzegowe nie zależą od ε .

- Łączny rozkład prawdopodobieństwa wektora losowego (X, Y) zależy od $\varepsilon \in [0, 1/4]$, a rozkłady brzegowe nie zależą. Stąd można oczekiwać, że zmienne losowe X, Y nie są niezależne, czyli są zależne.
- Zauważmy, że nie trzeba analizować zależności rozkładu łącznego od ε , gdyż $0 = f(1, 0) \neq \frac{1}{4} \cdot \frac{4}{8} = f_X(1)f_Y(0)$, a stąd zmienne X, Y są zależne niezależnie od wartości ε .

Zad.2. Rozkład prawdopodobieństwa wektora losowego (X, Y) jest następujący:

	Y		
X	0	1	2
1	0,1	0,1	p
2	0,1	0,2	0,3

- (a) Znaleźć p .
 (b) Obliczyć $F(2, 1)$, $F(1.5, 3)$, gdzie F oznacza dystrybuantę rozkładu wektora losowego (X, Y) .
 (c) Wyznaczyć rozkłady brzegowe zmiennych losowych X i Y oraz sprawdzić czy X i Y są niezależne.
 (d) Czy zmienne losowe X i Y są skorelowane? Jeśli tak, to w jakim stopniu?

Rozw.

- (a) p jest takie, że

$$\sum_{x=0}^2 \sum_{y=1}^2 f(x, y) = 1 \quad \equiv \quad 0,1 + 0,1 + p + 0,1 + 0,2 + 0,3 = 1 \quad \equiv \quad p = 1 - 0,8 = 0,2$$

$$(b) \quad F(2,1) = P(X \leq 2, Y \leq 1) = f(1,0) + f(1,1) + f(2,0) + f(2,1) = 0,1 + 0,1 + 0,1 + 0,2 = 0,5$$

$$F(1,5; 3) = P(X \leq 1,5; Y \leq 3) = f(1,0) + f(1,1) + f(1,2) = 0,1 + 0,1 + 0,2 = 0,4$$

y	0	1	2	
x				$f_X(x)$
1	0,1	0,1	0,2	0,4
2	0,1	0,2	0,3	0,6
$f_Y(y)$	0,2	0,3	0,5	1

(c) X, Y są zależne, bo np. $f(1,1) = 0,1 \neq 0,4 \cdot 0,3 = 0,12 = f_X(1) \cdot f_Y(1)$

(d) $\rho = ?$ (trzeba znaleźć $E(X)$, σ_X , $E(Y)$, σ_Y , $E(XY)$)

$$E(X) = 1 \cdot 0,4 + 2 \cdot 0,6 = 1,6, \quad E(X^2) = 1^2 \cdot 0,4 + 2^2 \cdot 0,6 = 0,4 + 2,4 = 2,8$$

$$Var(X) = E(X^2) - E(X)^2 = 2,8 - (1,6)^2 = 0,24, \quad \sigma_X = \sqrt{Var(X)} = \sqrt{0,24}$$

$$E(Y) = 0 \cdot 0,2 + 1 \cdot 0,3 + 2 \cdot 0,5 = 1,3, \quad E(Y^2) = 0^2 \cdot 0,2 + 1^2 \cdot 0,3 + 2^2 \cdot 0,5 = 2,3$$

$$Var(Y) = E(Y^2) - E(Y)^2 = 2,3 - (1,3)^2 = 2,3 - 1,69 = 0,61, \quad \sigma_Y = \sqrt{Var(Y)} = \sqrt{0,61}$$

$$Cov(X, Y) = E(XY) - E(X)E(Y) = \sum_{x=1}^2 \sum_{y=0}^2 xyf(x, y) - E(X)E(Y) =$$

$$= \sum_{x=1}^2 \sum_{y=1}^2 xyf(x,y) - E(X)E(Y) = f(1,1) + 2f(1,2) + 2f(2,1) + 4f(2,2) - 1,6 \cdot 1,3 =$$

$$= 0,1 + 2 \cdot 0,2 + 2 \cdot 0,2 + 4 \cdot 0,3 - 2,08 = 2,1 - 2,08 = 0,02$$

$$\rho = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{0,02}{\sqrt{0,24}\sqrt{0,61}} = \frac{2}{\sqrt{24 \cdot 61}} = 0,052.$$

Zmienne losowe X, Y są skorelowane, bo współczynnik korelacji jest różny od zera. Wynosi 0,052, czyli można mówić o bardzo niewielkiej współzależności liniowej dodatniej.

Zad.3. Rozkład prawdopodobieństwa wektora losowego (X, Y) jest następujący:

	Y	
X	0	1
0	0,1	0,2
1	0,3	0,1
2	0,2	a

(a) Znaleźć a oraz obliczyć prawdopodobieństwa:

$$P(1 \leq X \leq 2, Y < 4), P(X = 1|Y = 0), P(Y = 1|X = 0).$$

Rozw. (a)

$$a = 1 - (0,1 + 0,2 + 0,3 + 0,1 + 0,2) = 1 - 0,9 = 0,1$$

$$P(1 \leq X \leq 2, Y < 4) = f(1,0) + f(1,1) + f(2,0) + f(2,1) = 0,3 + 0,1 + 0,2 + 0,1 = 0,7$$

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

$$P(Y = 1|X = 0) = \frac{P(X = 0, Y = 1)}{P(X = 0)} = \frac{f(0,1)}{0,1 + 0,2} = \frac{0,2}{0,3} = \frac{2}{3}$$

(b) Wyznaczyć rozkłady brzegowe zmiennych losowych X i Y oraz sprawdzić czy X i Y są niezależne.

	Y	
X	0	1
0	0,1	0,2
1	0,3	0,1
2	0,2	0,1

y	0	1	$f_X(x)$
x			
0	0,1	0,2	0,3
1	0,3	0,1	0,4
2	0,2	0,1	0,3
$f_Y(y)$	0,6	0,4	1

X, Y są niezależne $\equiv f(x, y) = f_X(x)f_Y(y)$, dla wszystkich x, y ?

$$f(0,0) = 0,1 \neq 0,3 \cdot 0,6 = f_X(x)f_Y(y) \Rightarrow X, Y \text{ są zależne}$$

(c) Obliczyć $Cov(X, Y)$. Czy zmienne losowe X i Y są skorelowane?

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

y x	0	1	$f_X(x)$
0	0,1	0,2	0,3
1	0,3	0,1	0,4
2	0,2	0,1	0,3
$f_Y(y)$	0,6	0,4	1

$$E(XY) = 1 \cdot 1 \cdot f(1,1) + 2 \cdot 1 \cdot f(2,1) = 0,1 + 2 \cdot 0,1 = 0,3$$

$$E(X) = 0 \cdot 0,3 + 1 \cdot 0,4 + 2 \cdot 0,3 = 1, \quad E(Y) = 0 \cdot 0,6 + 1 \cdot 0,4 = 0,4$$

$$Cov(X, Y) = E(XY) - E(X)E(Y) = 0,3 - 1 \cdot 0,4 = -0,1.$$

Zmienne losowe X i Y są skorelowane $\equiv Cov(X, Y) \neq 0$

(d) Znaleźć wartość oczekiwaną i wariancję zmiennej losowej $2X + Y$.

$$\color{blue}{+} \quad E(2X + Y) = 2E(X) + E(Y) = 2 + 0,4 = 2,4.$$

$$Var(2X + Y) = Var(2X) + 2Cov(2X, Y) + Var(Y) = 2^2Var(X) + 2 \cdot 2 \cdot Cov(X, Y) + Var(Y). (**)$$

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

$$E(X^2) = 0^2 \cdot 0,3 + 1^2 \cdot 0,4 + 2^2 \cdot 0,3 = 1,6, \quad E(Y^2) = 0^2 \cdot 0,6 + 1^2 \cdot 0,4 = 0,4$$

$$Var(X) = E(X^2) - (E(X))^2 = 1,6 - 1^2 = 0,6,$$

$$Var(Y) = E(Y^2) - (E(Y))^2 = 0,4 - 0,4^2 = 0,4 - 0,16 = 0,24$$

Podstawiając do (**) mamy

$$\color{blue}{+} \quad Var(2X + Y) = 4 \cdot 0,6 + 4 \cdot (-0,1) + 0,24 = 2,4 - 0,4 + 0,24 = 2,24.$$

Uwaga. Wykorzystaliśmy wzory prawdziwe dla dowolnych stałych a, b i dowolnych zmiennych losowych:

$$\color{blue}{+} \quad E(aX + bY) = aE(X) + bE(Y)$$

$$\color{blue}{+} \quad Var(X + Y) = Var(X) + 2Cov(X, Y) + Var(Y)$$

$$\color{blue}{+} \quad Var(aX + b) = a^2Var(X)$$

$$\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$$

Zad.4. Rozkład prawdopodobieństwa wektora losowego (X, Y) jest następujący:

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

- (a) Wyznaczyć rozkłady brzegowe zmiennych losowych X i Y oraz sprawdzić czy X i Y są niezależne.
- (b) Czy zmienne losowe X i Y są nieskorelowane?

	X			
Y	0	2	3	$f_Y(y)$
-1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{4}$
1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{4}$
$f_X(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1

$$f(0, -1) = \frac{1}{4} = f_X(0)f_Y(-1), \quad f(2, -1) = \frac{1}{8} = f_X(2)f_Y(-1), \quad f(3, -1) = \frac{1}{8} = f_X(3)f_Y(-1),$$

$$f(0, 1) = \frac{1}{4} = f_X(0)f_Y(1), \quad f(2, 1) = \frac{1}{8} = f_X(2)f_Y(1), \quad f(3, 1) = \frac{1}{8} = f_X(3)f_Y(1),$$

Zatem, zmienne losowe X i Y są niezależne.

- (b) Zmienne losowe X i Y są nieskorelowane, co wynika z punktu (a) oraz twierdzenia:

Twierdzenie. Jeśli zmienne losowe X i Y są niezależne, to są nieskorelowane, tzn. $\text{Cov}(X, Y) = 0$.

Zad.5. W wyniku przeprowadzonych badań, ustalono, że w pewnej fabryce wytwarzającej sprzęt RTV, liczba wyprodukowanych w ciągu dnia niesprawnych odbiorników telewizyjnych ma rozkład Poissona o średniej 2. Podobnie, liczba niesprawnych zestawów stereo ma również rozkład Poissona o wartości oczekiwanej 3. Wyznaczyć prawdopodobieństwo zdarzenia, że w ciągu 2 dni, liczba niesprawnych produktów nie przekroczy 5.

- $X \sim \text{Poisson}(2)$, liczba wyprod. niesprawnych odbiorników tv w ciągu losowego dnia
- $Y \sim \text{Poisson}(3)$, —" — zestawów stereo —" —
- X_1, X_2 — niezależne zmienne losowe o rozkładach $\text{Poisson}(2)$
- Y_1, Y_2 — niezależne zmienne losowe o rozkładach $\text{Poisson}(3)$

- X_1, X_2, Y_1, Y_2 – niezależne zmienne losowe
- $W = X_1 + X_2 + Y_1 + Y_2 \sim \text{Poisson}(2 + 2 + 3 + 3) = \text{Poisson}(10)$

Uwaga. W ostatnim punkcie skorzystaliśmy z

Twierdzenie. Jeśli zmienne losowe X_1, X_2, \dots, X_n są niezależne i mają rozkłady Poissona, z parametrami, odpowiednio, $\lambda_1, \lambda_2, \dots, \lambda_n$, to

$$S_n = X_1 + X_2 + \dots + X_n \sim \text{Poisson}(\lambda_1 + \lambda_2 + \dots + \lambda_n).$$

$$\begin{aligned} \bullet \quad P(W \leq 5) &= \sum_{k=0}^5 P(W = k) = \sum_{k=0}^5 e^{-10} \frac{10^k}{k!} = \\ &= e^{-10} \left(1 + \frac{10^1}{1!} + \frac{10^2}{2!} + \frac{10^3}{3!} + \frac{10^4}{4!} + \frac{10^5}{5!} \right) = 0,0671. \end{aligned}$$

Zad.6. Cukier pakowany jest w torebki o nominalnej masie 1kg. W rzeczywistości, w wyniku błędów ważenia, masa pojedynczej torebki ma rozkład normalny o wartości oczekiwanej 1kg i wariancji $0,025 \text{ kg}^2$. Torebki są pakowane w paczki po 10 torebek. Jaki jest procent paczek, nominalnie dziesięciokilogramowych, które mają masę mniejszą niż 9,5 kg?

Rozw.

- $X \sim N(1, \sqrt{0,025})$, X – masa 1kg torebki cukru
- $X_1, X_2, \dots, X_9, X_{10}$ – masy 10-ciu 1kg torebek cukru w paczce, niezależne zmienne losowe o jednakowych rozkładach $N(1, \sqrt{0,025})$,
- $S_{10} = X_1 + X_2 + \dots + X_9 + X_{10} \sim N\left(10, \sqrt{10 \cdot 0,025}\right) = N(10; 0,5)$

Uwaga. Ostatni punkt wynika z tw. o dodawaniu dla niezależnych zmiennych losowych o rozkładach normalnych:

Twierdzenie. Jeśli zmienne losowe X_1, X_2, \dots, X_n są niezależne i mają rozkłady normalne, takie, że $X_j \sim N(\mu_j, \sigma_j^2), j = 1, 2, \dots, n$, to

$$S_n = X_1 + X_2 + \dots + X_n \sim N\left(\mu_1 + \mu_2 + \dots + \mu_n, \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}\right)$$

$$\begin{aligned} p = P(S_{10} < 9,5) &= P\left(\frac{S_{10} - 10}{0,5} < \frac{9,5 - 10}{0,5}\right) = \Phi(-1) = 1 - \Phi(1) = 1 - 0,8413 \\ &= 0,1587 \end{aligned}$$

Odp. 15,87% 10-ciu opakowań paczek 1kg cukru ma rzeczywistą wagę mniejszą niż 9,5kg.

Zadanie 7. W pewnym kiosku sprzedawanych jest średnio:

- 1000 egzemplarzy "Gazety Wyborczej" oraz
- 200 egzemplarzy "Przeglądu Sportowego" dziennie,

z odchyleniem standardowym, odpowiednio 100 i 15. Współczynnik korelacji między liczbą sprzedawanych gazet obu tytułów wynosi 0,6. Wiedząc, że zysk ze sprzedaży jednego egzemplarza "Gazety" wynosi 0,1 zł, natomiast ze sprzedaży "Przeglądu" 0,2 zł, obliczyć

Zad. 9 Niech X oznacza liczbę błędów drukarskich w książce liczącej 144 strony. Liczba błędów drukarskich na jednej stronie jest zmienną losową o rozkładzie Poissona. Średnio mamy 1 błąd na 12 stron. Podać wzór na wartość prawdopodobieństwa, że liczba błędów jest nie mniejsza niż 20 i nie większa niż 30. Przybliżyć to prawdopodobieństwo stosując przybliżenie wynikające z Centralnego Twierdzenia Granicznego.

Rozw. Trzeba założyć, że liczby błędów drukarskich na 144 stronach są niezależnymi zmiennymi losowymi.

- X_i – liczba błędów drukarskich na stronie i -tej, $i = 1, 2, \dots, 144$.
- $X_i \sim \text{Poisson}(\lambda)$.

Własność rozkładu Poissona: $X \sim \text{Poisson}(\lambda) \Rightarrow E(X) = \text{Var}(X) = \lambda$

- X_1, X_2, \dots, X_{144} – niezależne zmienne losowe.
- $E(X_1 + X_2 + \dots + X_{144}) = E(X_1) + E(X_2) + \dots + E(X_{144}) = 144E(X_1) = 144\lambda = 12$

$$\text{Stąd } \lambda = \frac{1}{12}.$$

$$S_{144} = X_1 + X_2 + \dots + X_{144} \sim \text{Poisson}(144\lambda) = \text{Poisson}(12)$$

🔧 Dokładny wzór:

$$P(20 \leq S_{144} \leq 30) = \sum_{k=20}^{30} e^{-144\lambda} \frac{(144\lambda)^k}{k!} = \sum_{k=20}^{30} e^{-12} \frac{12^k}{k!} = e^{-12} \sum_{k=20}^{30} \frac{12^k}{k!}$$

$$= 0,0212 \text{ (odczyt z Excela)}$$

- ✓ Z CTG S_{144} ma rozkład bliski rozkładowi normalnemu z wartością średnią $E(S_{144})$ i wariancją $\text{Var}(S_{144})$
- ✓ $E(S_{144}) = 144E(X_1) = 144\lambda = 12$
- ✓ $\text{Var}(S_{144}) = 144\text{Var}(X_1) = 144\lambda = 12$

Przybliżenie z CTG z uwzględnieniem korekty ciągłości:

$$\begin{aligned} \text{🔧 } P(20 \leq S_{144} \leq 30) &= P(19 < S_{144} < 31) = P(19,5 < S_{144} < 30,5) = \\ &= P\left(\frac{19,5 - 12}{\sqrt{12}} < \frac{S_{144} - 12}{\sqrt{12}} < \frac{30,5 - 12}{\sqrt{12}}\right) \approx \Phi\left(\frac{18,5}{\sqrt{12}}\right) - \Phi\left(\frac{7,5}{\sqrt{12}}\right) = \\ &= \Phi(5,34049) - \Phi(2,16506) = 0,0153. \end{aligned}$$