Wykład VIII

Zadanie 1.

Dwuwymiarowa zmienna losowa (X,Y) ma funkcję prawdopodobieństwa określoną tabelą z zadania 1 z poprzedniej pracy domowej (wykład VII)

(a) Oblicz współczynnik korelacji między zmiennymi X, Y.
(b) Oblicz E(X^2Y).

1. $f(X,Y)$: (X,Y) - zmienna losowa dwuwymiarowa

X \ Y	0	1	2
-1	0,1	0,1	0,3
1	0	0,2	0
3	0,1	0	0,2

$\rho = ?$

a) zmienne X i Y są zależne, więc $\rho \neq 0$

z definicji współczynnika korelacji:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{VarX}\sqrt{VarY}}$$, gdzie $Cov(X,Y)$ - kowariancja
$VarX, VarY$ - wariancja
$\sqrt{VarX}, \sqrt{VarY}$ - odchylenie standardowe

z definicji kowariancji:

$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$, co po przekształceniu daje

$Cov(X,Y) = E(XY) - \mu_x \mu_y$

z definicji wartości oczekiwanej brzegowej μ_x dla zmiennej losowej dyskretnej:

$$\mu_x = EX = \sum_x x f_x(x) = -1 \cdot (0,1+0,1+0,3) + 1 \cdot (0+0,2+0) + 3 \cdot (0,1+0+0,2) =$$
$$= -0,5 + 0,2 + 0,9 = 0,6$$

analogicznie dla μ_y:

$$\mu_y = EY = \sum_y y f_y(y) = 0 \cdot (0,1+0+0,1) + 1 \cdot (0,1+0,2+0) + 2(0,3+0+0,2) =$$
$$= 0 + 0,3 + 1 = 1,3$$

$$E(XY) = \sum_x \sum_y x\, y f(x,y) = -1 \cdot \sum_y y \cdot f(-1;y) + 1 \cdot \sum_y y \cdot f(1;y) + 3 \sum_y y \cdot f(3;y) =$$

$$= -1(0\cdot 0,1 + 1\cdot 0,1 + 2\cdot 0,3) + 1(0\cdot 0 + 1\cdot 0,2 + 2\cdot 0) + 3(0\cdot 0,1 + 1\cdot 0 + 2\cdot 0,2) =$$
$$= -1(0+0,1+0,6) + 1(0+0,2+0) + 3(0+0+0,4) = -0,7 + 0,2 + 1,2 = 0,7$$

Po podstawieniu otrzymanych wartości EX, EY i $E(XY)$ otrzymujemy:

$Cov(X,Y) = 0,7$

Z definicji VarX:

$VarX = E(X-\mu_X)^2$, czyli dla zmiennej losowej dyskretnej

$VarX = \sum_x (x-\mu_x)^2 \cdot f(x)$

Należy wyznaczyć rozkłady brzegowe zmiennych X i Y z definicji:

$f_x(x) = \sum_y f(x,y)$ oraz $f_y(y) = \sum_x f(x,y)$

$f_x(x) = \begin{cases} x=-1 & 0,5 \\ x=1 & 0,2 \\ x=3 & 0,3 \end{cases}$ $f_y(y) = \begin{cases} 0,2 & \text{dla } y=0 \\ 0,3 & \text{dla } y=1 \\ 0,5 & \text{dla } y=2 \end{cases}$

stąd $VarX = \sum_x (x-\mu_x)^2 \cdot f_x(x) = \sum_x (x-0,6)^2 \cdot f_x(x) =$

$= (-1-0,6)^2 \cdot 0,5 + (1-0,6)^2 \cdot 0,2 + (3-0,6)^2 \cdot 0,3 = (-1,6)^2 \cdot 0,5 + 0,2 \cdot 0,4^2 + 2,4^2 \cdot 0,3 =$

$= 2,56 \cdot 0,5 + 0,2 \cdot 0,16 + 5,76 \cdot 0,3 = 1,28 + 0,032 + 1,728 = 3,04$

analogicznie dla $VarY = \sum_y (y-\mu_y)^2 \cdot f_y(y) = \sum_y (y-1,3)^2 \cdot f_y(y) =$

$= (0-1,3)^2 \cdot 0,2 + (1-1,3)^2 \cdot 0,3 + (2-1,3)^2 \cdot 0,5 = 1,69 \cdot 0,2 + 0,9 \cdot 0,3 + 0,49 \cdot 0,5 =$

$= 0,845 + 0,27 + 0,245 = 1,36$

po podstawieniu wszystkich otrzymanych wartości:

$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{VarX}\sqrt{VarY}} = \frac{0,7}{\sqrt{3,04}\cdot\sqrt{1,36}} \approx \underline{0,344}$

Odp. Współczynnik korelacji wynosi ok. 0,344.

b) $E(X^2Y) = ?$

z definicji: $E(X,Y) = \sum_x \sum_y xy\, f(x,y)$

$E(X^2Y) = \sum_x \sum_y x^2 y\, f(x,y) = 1 \cdot \sum_y y f(-1,y) + 1 \cdot \sum_y y f(1,y) + 9 \cdot \sum_y y\, f(3,y) =$

$= 1 \cdot (0 \cdot 0,1 + 1 \cdot 0,1 + 2 \cdot 0,3) + 1 \cdot (0 \cdot 0 + 1 \cdot 0,2 + 2 \cdot 0) + 9(0 \cdot 0,1 + 1 \cdot 0 + 2 \cdot 0,2) =$

$= 0,7 + 0,2 + 3,6 = 4,5$

Odp. $E(X^2Y) = 4,5$

Zadanie 2.

Dwuwymiarowa zmienna losowa ciągła (X,Y) ma funkcję gęstości łącznej postaci z zadania 2 z poprzedniej pracy domowej (wykład VII)
Oblicz współczynnik korelacji między zmiennymi X i Y

2.

$$f(x,y) = \begin{cases} \frac{3}{4}x^2 y & -1 \leq x \leq 1 \quad 0 \leq y \leq 2 \\ 0 & \text{przeciwnie} \end{cases}$$

$$\rho(X,Y) = ?$$

z def. współczynnika korelacji:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{VarX}\sqrt{VarY}}$$

z definicji kowariancji dla zmiennej losowej ciągłej

$$Cov(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_x)(y-\mu_y)f(x,y)dxdy$$

na mocy stwierdzenia:

$$Cov(X,Y) = E(XY) - \mu_x\mu_y$$

z def. wartości oczekiwanej dla zmiennej losowej dwuwymiarowej ciągłej:

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y)dxdy = \int_{-1}^{1}dx\int_{0}^{2}\frac{3}{4}x^3y^2dy = \int_{-1}^{1}\left[\frac{3}{4}x^3y^3\right]_0^2 dx = \int_{-1}^{1}2x^3dx =$$

$$= \frac{x^4}{2}\Big|_{-1}^{1} = 0$$

z definicji wartości oczekiwanej brzegowej

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

gdzie $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$

$f_X(x) = \int_0^2 \frac{3}{4}x^2 y\,dy = \frac{3}{8}x^2y^2 \Big|_0^2 = \frac{3x^2}{2}$

stąd $\mu_X = \int_{-1}^{1} x\cdot\frac{3}{2}x^2\,dx = \int_{-1}^{1}\frac{3}{2}x^3\,dx = \frac{3x^4}{8}\Big|_{-1}^{1} = 0$

analogicznie dla μ_Y:

$f_Y(y) = \int_{-1}^{1}\frac{3}{4}x^2y\,dx = \frac{x^3y}{4}\Big|_{-1}^{1} = \frac{y}{2}$

$\mu_Y = \int_0^2 y\cdot\frac{y}{2}\,dy = \int_0^2 \frac{y^2}{2}\,dy = \frac{y^3}{6}\Big|_0^2 = \frac{8}{6} = \frac{4}{3}$

podstawiając do $Cov(X,Y)$:

$Cov(X,Y) = \cancel{0} - 0\cdot\frac{4}{3} = \cancel{0}$

Obliczam VarX i VarY

z definicji: $VarX = \int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)\,dx = \int_{-1}^{1}(x-0)^2\frac{3}{2}x^2\,dx = \int_{-1}^{1} x^2\cdot\frac{3}{2}x^2\,dx = \int_{-1}^{1}\frac{3}{2}x^4\,dx =$

$= \frac{3}{10}x^5\Big|_{-1}^{1} = \frac{3}{5}$

analogicznie dla $VarY = \int_0^2 (y-\mu_Y)^2 f_Y(y)\,dy = \int_0^2\left(y-\frac{4}{3}\right)^2\cdot\frac{y}{2}\,dy =$

$= \int_0^2\left(y^2-\frac{8}{3}y+\frac{16}{9}\right)\frac{y}{2}\,dy = \int_0^2 \frac{y^3}{2}-\frac{4}{3}y^2+\frac{8}{9}y\,dy = \left[\frac{y^4}{8}-\frac{4y^3}{9}+\frac{4y^2}{9}\right]_0^2 = \frac{2}{9}$

stąd $\rho(X,Y) = \frac{0}{\sqrt{\frac{3}{5}}\sqrt{\frac{2}{9}}} = 0$

Odp. Współczynnik korelacji wynosi 0. - zmienne X i Y są niezależne.

Zadanie 3.

Dwuwymiarowa zmienna losowa ciągła (X,Y) ma funkcję gęstości łącznej postaci z zadania 3 z poprzedniej pracy domowej (wykład VII). Oblicz Cov(X, Y).

3.
$$f(x,y)=\begin{cases}2x^2 & 0\le y\le 2x\le 2\\ 0 & \text{przeciwnie}\end{cases}$$

z definicji kowariancji:

$$Cov(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_x)(y-\mu_y)\,f(x,y)\,dx\,dy$$, co po przekształceniu

daje $Cov(X,Y)=E(XY)-\mu_x\mu_y$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xy\,f(x,y)\,dx\,dy=\int_0^1 dx\int_0^{2x}2x^3y\,dy=\int_0^1\left[x^3y^2\right]_0^{2x}dx=$$

$$=\int_0^1 4x^5\,dx=\frac{2x^6}{3}\Big|_0^1=\frac{2}{3}$$

z definicji EX:

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

z poprzedniej pracy domowej:

$$f_X(x)=\begin{cases}4x^3 & \text{dla } 0\le x\le 1\\ 0 & \text{przeciwnie}\end{cases}$$

$$f_Y(y)=\begin{cases}\frac{2}{3}-\frac{y^3}{12} & \text{dla } 0\le y\le 2\\ 0 & \text{przeciwnie}\end{cases}$$

stąd $\mu_X=\int_0^1 x\cdot 4x^3\,dx=\int_0^1 4x^4\,dx=\frac{4x^5}{5}\Big|_0^1=\frac{4}{5}$

analogicznie: $\mu_Y=\int_0^2 y\cdot\left(\frac{2}{3}-\frac{y^3}{12}\right)dy=\int_0^2\frac{2}{3}y-\frac{y^4}{12}\,dy=\frac{1}{12}\left(4y^2-\frac{y^5}{5}\right)\Big|_0^2=\frac{4}{5}$

stąd $Cov(X,Y)=\frac{2}{3}-\frac{4}{5}\cdot\frac{4}{5}=\frac{2}{3}-\frac{16}{25}=\frac{2}{75}\approx 0{,}027$

Odp. $Cov(X,Y)\approx 0{,}027$

Zadanie 4.

X i *Y* są niezależnymi zmiennymi losowymi o rozkładzie *N*(0,1). Wykaż, że zmienna (*Y-2X*, *X+3Y*) ma dwuwymiarowy rozkład normalny. Wyznacz parametry tego rozkładu.

4. X, Y - niezależne zmienne losowe

$X \sim N(0,1)$
$Y \sim N(0,1)$

$(\underset{A}{Y-2X}, \underset{B}{X+3Y})$ - dwuwymiarowy rozkład normalny?
parametry rozkładu: $\mu_A, \mu_B, \sigma_A, \sigma_B, \rho$ =?

$A = Y-2X$
$B = X+3Y$

zgodnie z twierdzeniem (A,B) ma dwuwym. rozkł. normalny tylko gdy zmienna losowa $a \cdot A + b \cdot B$ ma rozkład normalny dla dowolnych stałych a, b

$$a \cdot (Y-2X) + b(X+3Y) = aY - 2Xa + bX + 3bY =$$
$$= Y(a+3b) + X(b-2a)$$

założenie: a, b - dowolne stałe
$X \sim N(0,1)$
$Y \sim N(0,1)$

teza: $Y(a+3b) + X(b-2a) \sim N(\mu, \sigma)$

dowód: na podstawie twierdzenia, dowolna kombinacja liniowa zmiennych losowych o rozkładzie normalnym posiada również rozkład normalny. twierdzenie to kończy dowód. zmienna losowa dwuwymiarowa $(Y-2X, X+3Y)$ ma rozkład normalny,
cbdo.

parametry rozkładu:

$\mu_A = E(Y-2X)$

$\mu_B = E(X+3Y)$

$\sigma_A = \sqrt{VarA}$

$\sigma_B = \sqrt{VarB}$

Na mocy właściwości wartości oczekiwanej i wariancji:

$E(Y-2X) = EY - E(2X) = EY - 2EX = 0-2\cdot 0=0$

$E(X+3Y) = EX+E(3Y) = EX+3EY = 0+3\cdot 0 = 0$

$VarA = Var(Y-2X) = Var(Y) + Var(2X)$ - ponieważ X oraz Y są niezależne pomijam we wzorze $Cov(Y,2X)$

$= VarY + 4Var(X) = 1+4 = 5$

stąd $\sigma_A = \sqrt{VarA} = \sqrt{5} \approx 2{,}24$

$VarB = Var(X+3Y) = VarX + 9VarY = 1+9=10$

$\sigma_B = \sqrt{10} \approx 3{,}16$

z def. wsp. korelacji:

$\rho = \frac{Cov(A,B)}{\sigma_A \cdot \sigma_B}$ z definicji kowariancji:

$Cov(A,B) = Cov(Y-2X, X+3Y) = E[(Y-2X-\mu_A)(X+3Y-\mu_B)] = E[(Y-2X-0)(X+3Y)] =$

$= E[(Y-2X)(X+3Y)] = E(XY+3Y^2-2X^2-6XY) = E(XY)+E(3Y^2)-E(2X^2)-E(6XY) =$

$= E(XY)+3E(Y^2)-2E(X^2)-6E(XY) = 3E(Y^2)-2E(X^2)-5E(XY) = 3-2-0=1$

$\rho = \frac{1}{\sqrt{5}\cdot\sqrt{10}} = \frac{1}{\sqrt{50}} = \frac{1}{5\sqrt{2}} \approx 0{,}14$

Odp. Parametry rozkładu to $\sim N(0{,}0;2{,}24;3{,}16;0{,}14)$.

wykonał
Sławomir Jabłoński,
s14736