

## Poprzedni wykład

- Niezależność zmiennych losowych  $f_{X,Y}(x,y) = f_X(x)f_Y(y)$
- Kowariancja  $\text{cov}(X,Y) = E[(X-E(X))(Y-E(Y))]$   

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-E(X))(y-E(Y))f_{X,Y}(x,y) dx dy$$
- Współczynnik korelacji  $\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}$   
 $-1 \leq \rho_{X,Y} \leq 1$   
 $\rho_{X,Y} = 1 \quad \text{dla} \quad Y = aX + b \quad a > 0$   
 $\rho_{X,Y} = -1 \quad \text{dla} \quad Y = aX + b \quad a < 0$

1

## Współczynnik korelacji zmiennych losowych

- 5. Jeżeli  $\rho_{X,Y} = 0$  to zmienne  $X$  i  $Y$  nie muszą być niezależne; nazywamy je wtedy **nieskorelowanymi**.  
 Ważny wyjątek: jeżeli  $X$  i  $Y$  mają rozkłady normalne i  $\rho_{X,Y} = 0$  to zmienne  $X$  i  $Y$  są niezależne.

- 6. Jeżeli  $\rho_{X,Y} \neq 0$  to zmienne  $X$  i  $Y$  nie są niezależne.

## Macierz kowariancji

$$K_{X_1, \dots, X_n} = \begin{pmatrix} \text{var}(X_1) & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{var}(X_2) & & \vdots \\ \vdots & & \ddots & \\ \text{cov}(X_n, X_1) & \dots & \dots & \text{var}(X_n) \end{pmatrix} \xrightarrow{2\text{-dim}} K_{X,Y} = \begin{pmatrix} \text{var}(X) & \text{cov}(X,Y) \\ \text{cov}(Y,X) & \text{var}(Y) \end{pmatrix}$$

Jest symetryczna, bo  $\text{cov}(X_i, X_j) = \text{cov}(X_j, X_i)$ .

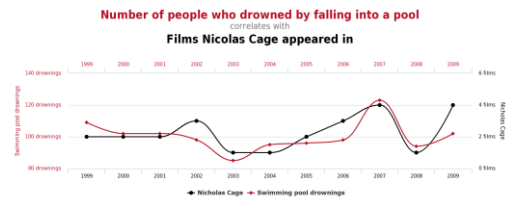
$$V_N = \frac{2}{N} \frac{\pi^{N/2}}{\Gamma(N/2)}$$

- Korelacja jako narzędzie testowania generatorów liczb pseudolosowych ( $\rho_{X,Y} = 0$ ), zadania kontrolne, np. N-wymiarowa kula o  $R=1$ .

2

<http://tylervigen.com/spurious-correlations>


3

<http://tylervigen.com/spurious-correlations>


Korelacje przypadkowe (falszywe) czy prawdziwe?

Analiza ekspercka

Zmienne współzależne:  $X \rightarrow Y, X \rightarrow Z$  prowadzi do korelacji  $(Y,Z)$ Zmienne pośredniczące:  $X \rightarrow Y \rightarrow Z$  prowadzi do korelacji  $(X,Z)$ 

Analiza zebrania próby: zmienne kontrolne, badania z randomizacją, dobór próby

Analiza wyników: grupowanie, analiza regresji, analiza wariancji, sieci zależności

4

## Łączna dystrybuanta

## Łączna dystrybuanta

$$F_{X,Y}(x,y) = P[X \leq x, Y \leq y] = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u,v) du dv$$

## Własności

- $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$
- $F_{X,Y}(+\infty, +\infty) = 1$
- dla  $x_1 \leq x_2$  i  $y_1 \leq y_2$   $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)$
- $\lim_{\delta \rightarrow 0^+} F_{X,Y}(x + \delta, y) = \lim_{\delta \rightarrow 0^+} F_{X,Y}(x, y + \delta) = F_{X,Y}(x, y)$
- $f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$

5

6

## Łączna dystrybuanta

### Własności

$$6. P[a < X \leq b, c < Y \leq d] = F_{X,Y}(b,d) - F_{X,Y}(b,c) - F_{X,Y}(a,d) + F_{X,Y}(a,c)$$

Dowód:

$$\begin{aligned} P[a < X \leq b, c < Y \leq d] &= \int_a^b \int_c^d f_{X,Y}(x,y) dy dx = \\ &= \int_a^b \int_c^d f_{X,Y}(x,y) dy dx - \int_a^b \int_c^c f_{X,Y}(x,y) dy dx = \\ &= \int_a^b \int_{-\infty}^d f_{X,Y}(x,y) dy dx - \int_a^b \int_{-\infty}^c f_{X,Y}(x,y) dy dx + \\ &\quad - \int_a^b \int_{-\infty}^c f_{X,Y}(x,y) dy dx + \int_a^b \int_{-\infty}^c f_{X,Y}(x,y) dy dx = \\ &= F_{X,Y}(b,d) - F_{X,Y}(b,c) - F_{X,Y}(a,d) + F_{X,Y}(a,c) \end{aligned}$$

7

## Łączna dystrybuanta

$$7. F_X(x) = P[X \leq x] = P[X \leq x, Y < +\infty] = F_{X,Y}(x, +\infty)$$

$$F_Y(y) = F_{X,Y}(+\infty, y)$$

■ Prowadzi to do **dystrybucyj brzegowej**

■ **Dystrybuanta brzegowa** zmiennej  $X$  to

$$F_X(x) = \lim_{y \rightarrow +\infty} F_{X,Y}(x, y)$$

a stąd **brzegowe funkcje gęstości prawdopodobieństwa**

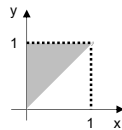
$$\begin{aligned} f_X(x) &= \frac{d}{dx} F_X(x) = \frac{d}{dx} F_{X,Y}(x, +\infty) = \\ &= \frac{d}{dx} \left[ \int_{-\infty}^{+\infty} f_{X,Y}(x, v) dv \right] = \int_{-\infty}^{+\infty} f_{X,Y}(x, v) dv \end{aligned}$$

8

## Łączna funkcja gęstości prawdopodobieństwa i dystrybuanta - przykład

Syt.  $X, Y$  – ciągle zmienne losowe

$$f_{X,Y}(x,y) = \begin{cases} c & \text{dla } 0 \leq x \leq y \leq 1 \\ 0 & \text{poza tym obszarem} \end{cases}$$



### Normalizacja

$$\begin{aligned} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dy dx &= 1 \rightarrow \int_0^1 \int_x^1 c dy dx = 1 \\ \int_0^1 \int_x^1 c dy dx &= c \int_0^1 \int_x^1 1 dy dx = c \int_0^1 y \Big|_x^1 dx = c \int_0^1 (1-x) dx = c \left( x - \frac{1}{2} x^2 \Big|_0^1 \right) = \\ &= c \left( 1 - 0 - \frac{1}{2} + 0 \right) = \frac{1}{2} c \\ \frac{1}{2} c &= 1 \rightarrow c = 2 \end{aligned}$$

9

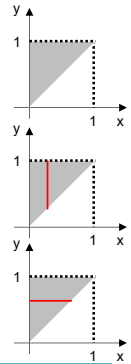
## Łączna funkcja gęstości prawdopodobieństwa i dystrybuanta - przykład

$$f_{X,Y}(x,y) = \begin{cases} 2 & \text{dla } 0 \leq x \leq y \leq 1 \\ 0 & \text{poza tym obszarem} \end{cases}$$

■ Brzegowe funkcje gęstości prawdopodobieństwa

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dy = \int_x^1 2 dy = 2y \Big|_x^1 = 2(1-x)$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dx = \int_0^y 2 dx = 2x \Big|_0^y = 2(y-0) = 2y$$



10

## Łączna funkcja gęstości prawdopodobieństwa i dystrybuanta - przykład

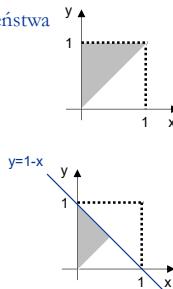
$$f_{X,Y}(x,y) = \begin{cases} 2 & \text{dla } 0 \leq x \leq y \leq 1 \\ 0 & \text{poza tym obszarem} \end{cases}$$

### Prawdopodobieństwo $P(X+Y < 1)$

$$P(X+Y < 1) = \iint_{\text{obszar}} f_{X,Y}(x,y) dy dx$$

$$X+Y < 1 \rightarrow Y < 1-X$$

$$\begin{aligned} P(X+Y < 1) &= \int_0^{0.5} \int_x^{1-x} 2 dy dx = \int_0^{0.5} 2(1-x) dx = \\ &= 2 \int_0^{0.5} (1-x) dx = 2 \left( x - \frac{1}{2} x^2 \right) \Big|_0^{0.5} = 2 \left( \frac{1}{2} - \frac{1}{2} \left( \frac{1}{2} \right)^2 - 0 + 0 \right) = 2 \left( \frac{1}{2} - \frac{1}{8} \right) = \frac{1}{2} \end{aligned}$$



11

## Łączna funkcja gęstości prawdopodobieństwa i dystrybuanta - przykład

$$f_{X,Y}(x,y) = \begin{cases} 2 & \text{dla } 0 \leq x \leq y \leq 1 \\ 0 & \text{poza tym obszarem} \end{cases}$$

### Dystrybuanta $F_{X,Y}(x,y)$

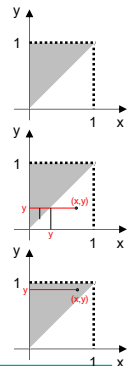
$$1. x < 0 \text{ lub } y < 0: F_{X,Y}(x,y) = 0$$

$$2. 0 < x < 1 \text{ i } 0 < y < x$$

$$\begin{aligned} F_{X,Y}(x,y) &= \int_0^y \int_x^y 2 dy dx = 2 \int_0^y y dy = 2 \int_0^y (y-x) dy = \\ &= 2 \left( yx - \frac{1}{2} y^2 \right) \Big|_0^y = 2 \left( yx - \frac{1}{2} y^2 - 0 + 0 \right) = y^2 \end{aligned}$$

$$3. 0 < x < 1 \text{ i } x < y < 1$$

$$\begin{aligned} F_{X,Y}(x,y) &= \int_0^y \int_x^y 2 dy dx = 2 \int_0^y y dy = 2 \int_0^y (y-x) dy = \\ &= 2 \left( yx - \frac{1}{2} y^2 \right) \Big|_0^y = 2 \left( yx - \frac{1}{2} y^2 - 0 + 0 \right) = 2xy - x^2 \end{aligned}$$



12

### Łączna funkcja gęstości prawdopodobieństwa i dystrybuanta - przykład

4.  $0 < x < 1$  i  $y \geq 1$

$$F_{X,Y}(x,y) = \int_0^x \int_y^1 2 \, dy \, dx = 2 \int_0^x \left[ y \right]_y^1 dx = 2 \int_0^x (1 - y) \, dx =$$

$$= 2 \left( x - \frac{1}{2} x^2 \right) \Big|_0^x = 2 \left( x - \frac{1}{2} x^2 - 0 + 0 \right) = 2x - x^2$$

5.  $x \geq 1$  i  $0 \leq y < 1$

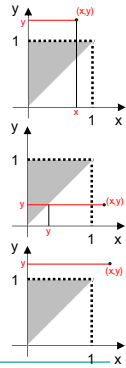
$$F_{X,Y}(x,y) = \int_0^y \int_x^1 2 \, dx \, dy = 2 \int_0^y \left[ x \right]_x^1 dy = 2 \int_0^y (1 - x) \, dy =$$

$$= 2 \left( yx - \frac{1}{2} x^2 \right) \Big|_0^y = 2 \left( y^2 - \frac{1}{2} y^2 - 0 + 0 \right) = y^2$$

6.  $x \geq 1$  i  $y \geq 1$

$$F_{X,Y}(x,y) = \int_0^1 \int_x^1 2 \, dx \, dy = 2 \int_0^1 \left[ x \right]_x^1 dy = 2 \int_0^1 (1 - x) \, dy =$$

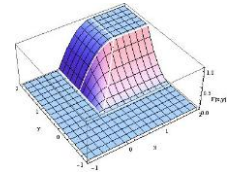
$$= 2 \left( x - \frac{1}{2} x^2 \right) \Big|_0^1 = 2 \left( 1 - \frac{1}{2} 1^2 - 0 + 0 \right) = 1$$



13

### Łączna funkcja gęstości prawdopodobieństwa i dystrybuanta - przykład

Wynik końcowy  $F_{X,Y}(x,y)$ :



14