# Zmienna losowa dwuwymiarowa

# Zmienna losowa dwuwymiarowa

Zmiennej losowej dwuwymiarowej $(X, Y)$ używamy wówczas, gdy **obserwacji podlegają dwie, a nie jedna cecha jednostek statystycznych**.

# Zmienna losowa dwuwymiarowa

Zmiennej losowej dwuwymiarowej $(X, Y)$ używamy wówczas, gdy **obserwacji podlegają dwie, a nie jedna cecha jednostek statystycznych**.

## ZMIENNA DWUWYMIAROWA DYSKRETNA

- Przyjmuje co najwyżej przeliczalną liczbę wartości $(x_i, y_j)$ dla $i, j \in \{1, 2, \ldots\}$.

# Zmienna losowa dwuwymiarowa

Zmiennej losowej dwuwymiarowej $(X, Y)$ używamy wówczas, gdy **obserwacji podlegają dwie, a nie jedna cecha jednostek statystycznych**.

## ZMIENNA DWUWYMIAROWA DYSKRETNA

- Przyjmuje co najwyżej przeliczalną liczbę wartości $(x_i, y_j)$ dla $i, j \in \{1, 2, \ldots\}$.
- Wartości $(x_i, y_j)$ przyjmowane są z prawdopodobieństwami $p_{ij} = \mathbb{P}((X, Y) = (x_i, y_j))$, gdzie
  - $p_{ij} \geqslant 0$ dla $i, j \in \{1, 2, \ldots\}$,
  - $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1$.

# Zmienna dwuwymiarowa dyskretna

| $(X,Y)$ | $y_1$ | $y_2$ | $y_3$ | $\sum_j p_{ij}$ |
|---|---|---|---|---|
| $x_1$ | $p_{11}$ | $p_{12}$ | $p_{13}$ | $p_{1.}$ |
| $x_2$ | $p_{21}$ | $p_{22}$ | $p_{23}$ | $p_{2.}$ |
| $\sum_i p_{ij}$ | $p_{.1}$ | $p_{.2}$ | $p_{.3}$ | $1$ |

- Rozkład łączny zmiennej losowej dwuwymiarowej $(X,Y)$: $p_{ij} = \mathbb{P}\left((X,Y) = (x_i, y_j)\right)$

# Zmienna dwuwymiarowa dyskretna

| $(X,Y)$ | $y_1$ | $y_2$ | $y_3$ | $\sum_j p_{ij}$ |
|---|---|---|---|---|
| $x_1$ | $p_{11}$ | $p_{12}$ | $p_{13}$ | $p_{1\cdot}$ |
| $x_2$ | $p_{21}$ | $p_{22}$ | $p_{23}$ | $p_{2\cdot}$ |
| $\sum_i p_{ij}$ | $p_{\cdot 1}$ | $p_{\cdot 2}$ | $p_{\cdot 3}$ | $1$ |

- Rozkład łączny zmiennej losowej dwuwymiarowej $(X,Y)$: $p_{ij} = \mathbb{P}\left((X,Y) = (x_i, y_j)\right)$
- Rozkłady brzegowe zmiennych losowych $X$ i $Y$ w rozkładzie dwuwymiarowym:

$$p_{i\cdot} = \mathbb{P}(X = x_i) = \sum_{j=1}^{\infty} p_{ij}, \qquad p_{\cdot j} = \mathbb{P}(Y = y_j) = \sum_{i=1}^{\infty} p_{ij}$$

# Zmienna dwuwymiarowa typu ciągłego

## DWUWYMIAROWA ZMIENNA TYPU CIĄGŁEGO

- Zbiór wartości $(x_i, y_j)$ jest nieprzeliczalny.

# Zmienna dwuwymiarowa typu ciągłego

## DWUWYMIAROWA ZMIENNA TYPU CIĄGŁEGO

- Zbiór wartości $(x_i, y_j)$ jest nieprzeliczalny.
- Rozkład zadaje dwuwymiarowa funkcja gęstości $f_{(X,Y)}$ taka, że
  - $f_{(X,Y)}(x, y) \geqslant 0$ dla $x, y \in \mathbb{R}$,
  - $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{(X,Y)}(x, y)\, dx\, dy = 1$.

# Niezależność zmiennych losowych

# Zmienne losowe niezależne

## Definicja

Zmienne losowe $X$ i $Y$ są **niezależne**, jeśli

$$\mathbb{P}\left(X \in A, Y \in B\right) = \mathbb{P}(X \in A) * \mathbb{P}(Y \in B) \quad \text{dla dowolnych } A, B \subset \mathbb{R}.$$

# Zmienne losowe niezależne

## Definicja

Zmienne losowe $X$ i $Y$ są **niezależne**, jeśli

$$\mathbb{P}\left(X \in A, Y \in B\right) = \mathbb{P}(X \in A) * \mathbb{P}(Y \in B) \quad \text{dla dowolnych } A, B \subset \mathbb{R}.$$

- Dla zmiennych losowych dyskretnych – wystarczy sprawdzić warunek

$$\mathbb{P}\left(X = x_i, Y = y_j\right) = \mathbb{P}\left(X = x_i\right) * \mathbb{P}\left(Y = y_j\right) \quad (p_{ij} = p_i\, p_j) \quad \text{dla wszystkich } i, j.$$

# Zmienne losowe niezależne

## Definicja

Zmienne losowe $X$ i $Y$ są **niezależne**, jeśli

$$\mathbb{P}\left(X \in A, Y \in B\right) = \mathbb{P}(X \in A) * \mathbb{P}(Y \in B) \quad \text{dla dowolnych } A, B \subset \mathbb{R}.$$

- Dla zmiennych losowych typu ciągłego – wystarczy sprawdzić warunek

$$f_{(X,Y)}(x, y) = f_X(x) * f_Y(y) \quad \text{dla} \quad x, y \in \mathbb{R},$$

gdzie $f_{(X,Y)}$ jest gęstością zmiennej dwuwymiarowej $(X, Y)$,
$f_X$ – gęstością zmiennej $X$ oraz $f_Y$ – gęstością zmiennej $Y$.

# Kowariancja i współczynnik korelacji liniowej

# Kowariancja

## Definicja

Kowariancja – określa zależność liniową między zmiennymi losowymi $X$ i $Y$.

# Kowariancja

## Definicja

Kowariancja – określa zależność liniową między zmiennymi losowymi $X$ i $Y$.

- Dla zmiennych dyskretnych:

$$\mathrm{cov}(X,Y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(x_i - \mathbb{E}(X))\,(y_j - \mathbb{E}(Y))\,p_{ij}.$$

- Dla zmiennych typu ciągłego:

$$\mathrm{cov}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - \mathbb{E}X)\,(y - \mathbb{E}Y)\,f_{(X,Y)}(x,y)\,dx\,dy.$$

# Współczynnik korelacji, zmienne losowe skorelowane i nieskorelowane

## Definicja

Współczynnik korelacji:

$$\rho = \frac{\mathrm{cov}(X,Y)}{D(X)\,D(Y)}, \qquad (-1 \leqslant \rho \leqslant 1).$$

# Współczynnik korelacji, zmienne losowe skorelowane i nieskorelowane

## Definicja

Współczynnik korelacji:

$$\rho = \frac{\mathrm{cov}(X,Y)}{D(X)\,D(Y)}, \qquad (-1 \leqslant \rho \leqslant 1).$$

- $\rho = 0$ – zmienne **nieskorelowane**
- $\rho \neq 0$ – zmienne **skorelowane (dodatnio lub ujemnie)**
- $|\rho| \approx 1$ – zmienne **silnie skorelowane**