

# RACHUNEK PRAWDOPODOBIENSTWA

## **WYKŁAD 4.**

### **PARAMETRY ZMIENNEJ LOSOWEJ JEDNOWYMIAROWEJ**

## **Wartość oczekiwana.**

Oznaczenie  $EX$  lub  $m$ .

Dla zmiennej losowej skokowej

$$EX = \sum_i x_i p_i$$

(jeśli ewentualny szereg jest zbieżny  
**bezwzględnie**, takie szeregi są "odporne"  
np. na zmianę kolejności wyrazów).

Dla zmiennej losowej ciągłej

$$EX = \int_{-\infty}^{\infty} xf(x)dx$$

(jeśli ewentualna całka niewłaściwa jest **zbieżna bezwzględnie**).

## Przykład

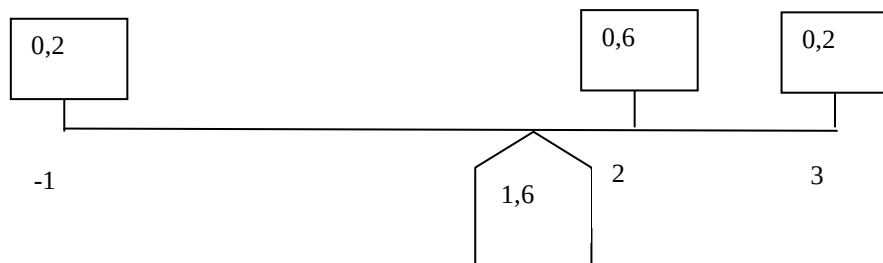
Dla zmiennej losowej o funkcji prawdopodobieństwa

|       |     |     |     |
|-------|-----|-----|-----|
| $x_k$ | -1  | 2   | 3   |
| $p_k$ | 0,2 | 0,6 | 0,2 |

$$EX = -1 \cdot 0,2 + 2 \cdot 0,6 + 3 \cdot 0,2 = 1,6 .$$

## Interpretacja.

Wartość oczekiwana wyznacza środek ciężkości masy jednostkowej rozłożonej w punktach skokowych.



## Przykład

Dla zmiennej losowej o gęstości

$$f(x) = \begin{cases} 2x & x \in < 0,1 > \\ 0 & x \notin < 0,1 > \end{cases}$$

$$EX = \int_0^1 x \cdot 2x dx = 2 \int_0^1 x^2 dx = 2 \frac{x^3}{3} \Big|_0^1 = \frac{2}{3}$$

## Przykład

Dla zmiennej losowej o gęstości

$$f(x) = \begin{cases} \frac{1}{x^2} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

wartość oczekiwana nie istnieje bo

$$\int_1^{\infty} x \cdot \frac{1}{x^2} dx = \int_1^{\infty} \frac{1}{x} dx = \ln x \Big|_1^{\infty} = \infty.$$

## Własności wartości oczekiwanej

- a)  $Ec = c$ ;  $c$  – stała,
- b)  $E(aX) = aE(X)$ ,
- c)  $E(X + Y) = EX + EY$ ,
- d) Jeśli  $a \leq X \leq b$ , to  $a \leq EX \leq b$ ,  
jeśli  $X \leq Y$ , to  $EX \leq EY$ ,
- e)  $EX \leq E|X|$ ,  $|EX| \leq E|X|$
- f)  $X, Y$  – niezależne, to  $E(XY) = EX \cdot EY$
- g) Jeśli  $Y = g(X)$ , to

$$EY = \begin{cases} \sum_i g(x_i) p_i & \text{gdy } X \text{ skokowa} \\ \int_{-\infty}^{\infty} g(x) f(x) dx & \text{gdy } X \text{ ciągła} \end{cases}$$



**Uwaga.**

$$EX = \int_0^{\infty} (1 - F(x)) dx - \int_{-\infty}^0 F(x) dx$$

Bardziej ogólną definicję wartości oczekiwanej otrzymamy stosując **całkę Stieltiesa**

$$EX = \int_{-\infty}^{\infty} x dF(x)$$

przy założeniu, że powyższa całka jest bezwzględnie zbieżna.

Ponieważ sumy całkowe definiujące całkę

$$\int_a^b x dF(x) \quad \text{mają postać}$$

$$S_n = \sum_{i=1}^n \hat{x}_i (F(x_i) - F(x_{i-1}))$$

to dla dystrybuanty kawałkami stałej wartość oczekiwana określona za pomocą całki Stieltiesa jest równa sumie iloczynów skoków dystrybuanty przez punkty skokowe i pokrywa się z określeniem podanym wcześniej.

Natomiast dla dystrybuanty zmiennej losowej ciągłej mamy

$$F(x_i) - F(x_{i-1}) = F'(\hat{x}_i)(x_i - x_{i-1}) = f(\hat{x}_i)(x_i - x_{i-1})$$

$$\text{czyli } dF(x) = f'dx$$

co również daje definicję podaną wcześniej.

Z addytywności całki Stieltjesa względem funkcji  $F$  mamy następującą własność:

Jeśli zmienna losowa  $X$  jest skokowo - ciągła tzn. jej dystrybuanta da się przedstawić

w postaci  $F(x) = cF_1(x) + (1-c)F_2(x)$ ,  $c > 0$ ,  
 $F_1$ - dystrybuanta zmiennej losowej skokowej  $X_1$ ,  
 $F_2$ - dystrybuanta zmiennej losowej ciągłej  $X_2$ .  
Wtedy

$$EX = cEX_1 + (1-c) EX_2$$

Inny sposób liczenia wartości oczekiwanych zmiennych losowych skokowo - ciągłych jest następujący:

Jeśli dystrybuenta  $F$  jest przedziałami ciągła i różniczkowalna wewnątrz przedziałów ciągłości to

$$EX = \int_{-\infty}^{\infty} xF'(x)dx + \sum_i x_i (F(x_i^+) - F(x_i)),$$

gdzie  $x_i$  - punkty skokowe dystrybuenty. Wzór ten pozwala też na wyznaczanie wartości oczekiwanej zmiennej losowej  $Y = g(X)$  bez znajomości rozkładu  $Y$ , mianowicie

$$EY = E(g(X)) = \int_{-\infty}^{\infty} g(x)F'(x)dx + \sum_i g(x_i)(F(x_i^+) - F(x_i))$$

gdzie  $x_i$  - punkty skokowe dystrybuenty  $F$  zmiennej losowej  $X$ .

## Przykład.

Jeśli zmienna losowa ma dystrybuantę

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 0,25 + 0,5x & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}$$

to  $F(x) = 0,5F_1(x) + 0,5F_2(x)$ , gdzie

$$F_1(x) = \begin{cases} 0 & x \leq 0 \\ 0,5 & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}$$

$$F_2(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}$$

Zatem  $EX = 0,5 \cdot 0,5 + 0,5 \cdot 0,5 = 0,5$ .

II sposób.

$$EX = \int_0^1 x \cdot 0,5 dx + (0 \cdot 0,25 + 1 \cdot 0,25) = 0,25 + 0,25 = 0,5$$

Miarą rozrzutu wartości zmiennej losowej jest wariancja.

### **Wariancja.**

Oznaczenie  $D^2X$  lub  $\sigma^2$  lub  $VX$ .

$$D^2X = E(X - EX)^2$$

Dla zmiennej losowej skokowej

$$D^2X = \sum (x_i - EX)^2 p_i$$

Dla zmiennej losowej ciągłej

$$D^2X = \int_{-\infty}^{\infty} (x - EX)^2 f(x) dx$$

## Własności wariancji

a)  $D^2 c = 0$ ;  $c$  – stała,

b)  $D^2(aX) = a^2 D^2(X)$ ,

c)  $D^2(X + b) = D^2 X$ ,  $b$  – stała,

d)  $X, Y$  – niezależne, to  $D^2(X \pm Y) = D^2 X + D^2 Y$

e)  $D^2 X = E(X^2) - (EX)^2$ .



## Wykorzystanie własności e)

$$D^2 X = \begin{cases} \sum_i x_i^2 p_i - (EX)^2 & \text{gdy } X \text{ skokowa} \\ \int_{-\infty}^{\infty} x^2 f(x) dx - (EX)^2 & \text{gdy } X \text{ ciągła} \end{cases}$$

## **Odchylenie standardowe.**

Oznaczenie  $DX$  lub  $\sigma$ .

$$DX = \sqrt{D^2 X}$$

## Własność

Jeśli  $X$  ma wartość oczekiwaną  $m$  i odchylenie standardowe  $\sigma > 0$

to zmienna losowa  $Y = \frac{X - m}{\sigma}$  ma

$$EY = 0 \text{ i } DX = 1.$$

Zmienną losową  $Y$  nazywamy zmienną losową **standaryzowaną**.

## Przykład

Dla zmiennej losowej o gęstości

$$f(x) = \begin{cases} 2x & x \in \langle 0,1 \rangle \\ 0 & x \notin \langle 0,1 \rangle \end{cases} \quad \text{mamy} \quad EX = \frac{2}{3} \quad \text{zatem}$$

$$D^2 X = \int_0^1 \left( x - \frac{2}{3} \right)^2 \cdot 2x dx = 2 \int_0^1 \left( x^3 - \frac{4}{3} x^2 + \frac{4}{9} x \right) dx = \frac{1}{18}$$

II sposób (na podstawie własności e))

$$D^2 X = \int_0^1 x^2 \cdot 2x dx - \left(\frac{2}{3}\right)^2 = 2 \int_0^1 x^3 dx - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

## Przykład

Jeśli niezależne zmienne losowe  $X_i$  ( $i = 1, 2, \dots, n$ ) mają taką samą wartość oczekiwaną  $m$  i takie samo odchylenie standardowe  $\sigma > 0$  to zmienna losowa  $\bar{X}$

będąca ich średnią  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  ma

$$E\bar{X} = m; \quad D\bar{X} = \frac{\sigma}{\sqrt{n}}$$

## **Nierówność Czebyszewa.**

Jeśli zmienna losowa  $X$  ma wartość oczekiwaną  $m$  i odchylenie standardowe  $\sigma > 0$  to dla dowolnego  $\varepsilon > 0$  mamy

$$P(|X - m| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

**Moment rzędu k** ( k - liczba naturalna)

$$m_k = E(X^k)$$

własność e) wariacji można zapisać

$$D^2X = m_2 - m^2.$$



**Własność.** Jeśli istnieje  $m_k$  to istnieje  $m_s$   
dla każdego  $s < k$ .

**Moment centralny rzędu  $k$**  (  $k$  - liczba naturalna)

$$\mu_k = E\left(\left(X - EX\right)^k\right)$$

Zauważmy, że w szczególności

$$\mu_1 = 0,$$

$$\mu_2 = D^2X.$$

**współczynnik asymetrii (skośności)**

$$a = \frac{\mu_3}{\sigma^3}$$

**i współczynnik skupienia (kurtozę)**

$$k = \frac{\mu_4}{\sigma^4}$$

**Zależności między momentami zwykłymi i centralnymi.**

$$m_k = \sum_{i=0}^k \binom{k}{i} \mu_{k-i} m_1^i$$

w szczególności

$$m_2 = \mu_2 + m_1^2,$$

$$m_3 = \mu_3 + 3\mu_2 m_1 + m_1^3,$$

$$m_4 = \mu_4 + 4\mu_3 m_1 + 6\mu_2 m_1^2 + m_1^4,$$

$$\mu_k = \sum_{i=0}^k (-1)^i \binom{k}{i} m_{k-i} m_1^i$$

w szczególności

$$\mu_2 = m_2 - m_1^2,$$

$$\mu_3 = m_3 - 3m_2 m_1 + 2m_1^3,$$

$$\mu_4 = m_4 - 4m_3 m_1 + 6m_2 m_1^2 - 3m_1^4,$$

**Kwantylem rzędu  $p$**  ( $0 < p < 1$ ) zmiennej losowej  $X$  o dystrybuancie  $F$  nazywamy liczbę  $x_p$ , taką, że

$$F(x_p) \leq p \leq F(x_p^+)$$



Zauważmy, że dla zmiennej losowej ciągłej  $x_p$  wyznaczymy z równości

$$F(x_p) = p$$

Kwantyl rzędu 0,5 nazywamy **medianą**.

Kwantyle rzędu 0,25 ; 0,5; 0,75 nazywamy **kwartylami** (drugi kwantyl jest medianą).

Kwantyle istnieją dla każdej zmiennej losowej, lecz nie zawsze są wyznaczone jednoznacznie.

### **Przykład.**

Dla zmiennej losowej o gęstości

$$f(x) = \begin{cases} 0,5 & x \in <0, 1> \cup <2, 3> \\ 0 & x \notin <0, 1> \cup <2, 3> \end{cases}$$
 medianą jest dowolna liczba z przedziału  $[1, 2]$ .