# 9 Parametry rozkładów zmiennych losowych

**Definicja 9.1** (Wartość oczekiwana)**.**
Niech $X$ będzie dyskretną zmienną losową o rozkładzie prawdopodobieństwa $p(x_i)$. Wartością oczekiwaną , $E[X]$, nazywamy

$$E[X] = \mu_X = \sum_i x_i p(x_i)$$

**Przykład 9.1** (Dwa rzuty sześcienną kostką do gry)**.**
Niech $X$ będzie suma wyrzuconych oczek w dwóch rzutach sześcienną kostką. Ile wynosi $E[X]$?

| $x_i$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $p_X(x_i)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |

$$\begin{aligned} E[X] &= 2\cdot\frac{1}{36} + 3\cdot\frac{2}{36} + 4\cdot\frac{3}{36} + 5\cdot\frac{4}{36} + 6\cdot\frac{5}{36} + 7\cdot\frac{6}{36} + \\ & \quad 8\cdot\frac{5}{36} + 9\cdot\frac{4}{36} + 10\cdot\frac{3}{36} + 11\cdot\frac{2}{36} + 12\cdot\frac{1}{36} \\ &= 7 \end{aligned}$$

**Definicja 9.2** (Wartość oczekiwana)**.**
Niech $X$ będzie ciągłą zmienną losową o gęstości pr. $f(x)$. Wartością oczekiwaną ciągłej zmiennej losowej $X$ nazywamy

$$E[X] = \mu_X = \int_{-\infty}^{\infty} x f(x) dx$$

**Przykład 9.2** (Odległość losowo wybranego punktu od środka okręgu)**.**
Wybieramy losowo punkt wewnątrz okręgu o promieniu $R$. Ile wynosi wartość oczekiwana zmiennej losowej $X$ będącej odległością punktu od środka okręgu?

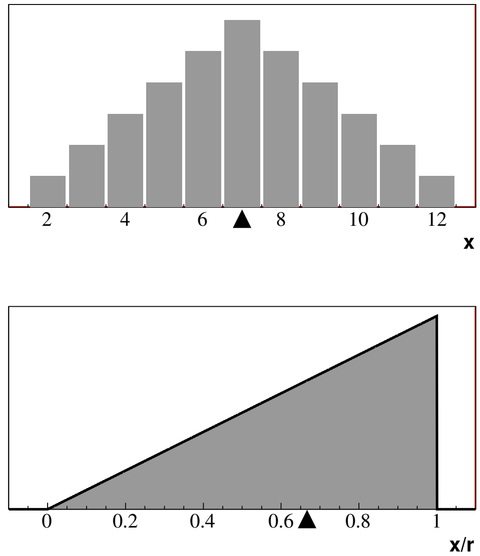
$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^R x \cdot \frac{2x}{R^2} dx = \frac{2}{R^2} \cdot \frac{x^3}{3}\Big|_0^R = \frac{2}{3}R$$

Wartość oczekiwaną możemy intuicyjnie rozumieć jako średnią wartość zmiennej losowej. Gdyby prawdopodobieństwo interpretować jako masę rozłożoną na pewnym zbiorze (dyskretnym lub ciągłym) to wartość oczekiwana będzie oznaczała położenie środka masy prawdopodobieństwa. Interpretację taka obrazuje rysunek 6.
Drugim parametrem obok wartości oczekiwanej charakteryzującym rozkład prawdopodobieństwa dowolnej zmiennej losowej jest wariancja, która jest miarą koncentracji zmiennej losowej wokół swojej wartości oczekiwanej. Rysunek 7 przedstawia wykresy rozkładów prawdopodobieństwa o tych samych wartościach oczekiwanych ale istotnie różnych wariancjach.

**Definicja 9.3** (Wariancja)**.**
Wariancja, $\mathrm{Var}(X)$, zmiennej losowej $X$ to liczba:

$$\mathrm{Var}(X) = \sigma_X^2 = E[(X - E[X])^2].$$

Wielkość $\sigma_X$ mającą wymiar zmiennej losowej nazywamy odchyleniem standardowym.



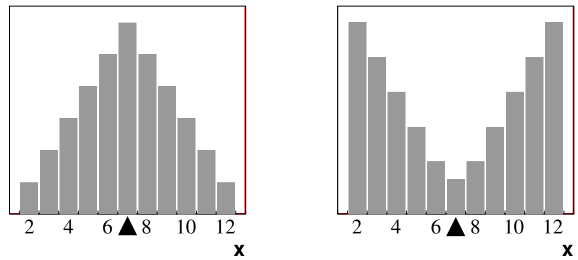

Rysunek 6: Wartość oczekiwana (pełny trójkąt) zmiennej losowej jako położenie środka masy prawdopodobieństwa. Wykres górny odpowiada zmiennej losowej z przykładu 9.1 a dolny z przykładu 9.2.

**Przykład 9.3** (Dwa rzuty sześcienną kostką do gry)**.**
Niech $X$ będzie suma wyrzuconych oczek w dwóch rzutach sześcienną kostką. Ile wynosi wariancja zmiennej losowej $X$ ?

| $x_i$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $p_X(x_i)$ | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |

$$\begin{aligned}
\text{Var}[X] &= E[(X-7)^2] \\
&= \sum_{i=1}^{11}(x_i-7)^2 p(x_i) \\
&= (2-7)^2\cdot\frac{1}{36}+(3-7)^2\cdot\frac{2}{36}+(4-7)^2\cdot\frac{3}{36}+(5-7)^2\cdot\frac{4}{36}+(6-7)^2\cdot\frac{5}{36}+ \\
&\quad (7-7)^2\cdot\frac{6}{36}+(8-7)^2\cdot\frac{5}{36}+(9-7)^2\cdot\frac{4}{36}+(10-7)^2\cdot\frac{3}{36}+(11-7)^2\cdot\frac{2}{36} \\
&\quad +(12-7)^2\cdot\frac{1}{36} \\
&= \frac{105}{18}
\end{aligned}$$

**Przykład 9.4** (Wariancja)**.**




Rysunek 7: Wariancja jako miara koncentracji rozkładu prawdopodobieństwa zmiennej losowej wokół wartość oczekiwanej (pełny trójkąt)

Wybieramy losowo punkt wewnątrz okręgu o promieniu $R$. Ile wynosi wariancja zmiennej losowej $X$ będącej odległością punktu od środka okręgu?

$$\begin{aligned}
\mathrm{Var}[X] &= \int_0^R (x - E[X])^2 f(x)dx \\
&= \int_0^R (x - \frac{2}{3}R)^2 \frac{2x}{R^2} dx \\
&= \frac{2}{R^2}\int_0^R x^3 dx - \frac{8}{3R}\int_0^R x^2 dx + \frac{8}{9}\int_0^R x dx \\
&= \frac{R^2}{2} - \frac{8R^2}{9} + \frac{4R^2}{9} \\
&= \frac{R^2}{18}
\end{aligned}$$

W praktyce do obliczania wariancji korzystamy z poniższego twierdzenia:

**Twierdzenie 9.1** (Wariancja)**.**

$$Var[X] = E[X^2] - (E[X])^2$$

*Dowód.*

$$\begin{aligned}
\mathrm{Var}[X] &= E[(X - \mu_X)^2] \\
&= E[X^2 - 2X\mu_X + \mu_X^2] \\
&= E[X^2] - 2\mu_X E[X] + \mu_X^2 E[1] \\
&= E[X^2] - \mu_X^2
\end{aligned}$$

□

**Twierdzenie 9.2** (Wartość oczekiwana funkcji zmiennej losowej)**.**

*Niech $X$ będzie zmienną losową o rozkładzie $p_X$ lub gęstości pr. $f_X$ a $Y = g(X)$ będzie zmienną losową, gdzie $g : \mathbb{R} \to \mathbb{R}$. Wtedy*

$$E[Y] = E[g(X)] = \begin{cases} \sum_{k=1} g(x_k)p_X(x_k) & \textit{jeśli } X \textit{ jest dyskretna} \\ \int_{-\infty}^{\infty} g(x)f_X(x)dx & \textit{jeśli } X \textit{ jest ciągła} \end{cases}$$

Uwaga:

$$E[g(X)] \neq g(E[X])$$

z jednym wyjątkiem $g(X) = \alpha X + \beta$

**Twierdzenie 9.3** (Liniowa transformacja zm. los.)**.**
*Niech $X$ będzie zmienną losową a $\alpha, \beta \in \mathbb{R}$*

- $E[\alpha X + \beta] = \alpha E[X] + \beta$
- $Var[\alpha X + \beta] = \alpha^2 Var[X]$

*Dowód.*

$$\begin{aligned} E[\alpha X + \beta] &= E[\alpha X] + E[\beta] \\ &= \alpha E[X] + \beta \end{aligned}$$

$$\begin{aligned} \text{Var}\,[\alpha X + \beta] &= E\left[(\alpha X + \beta - \alpha E[X] - \beta)^2\right] \\ &= E\left[(\alpha(X - E[X]))^2\right] \\ &= \alpha^2 E[(X - E[X])^2] \\ &= \alpha^2 \text{Var}\,[X] \end{aligned}$$

□

**Definicja 9.4** (Standaryzacja zmiennej losowej)**.**
Jeśli zmienna losowa $X$ ma wartość oczekiwaną $E[X] = \mu$ i wariancję $\text{Var}[X] = \sigma^2$ to zmienną losową:

$$Z = \frac{X - \mu}{\sigma}$$

Nazywamy standaryzowaną zmienną losową

Uwaga:

$$E[Z] = \frac{E[X] - \mu}{\sigma} = 0$$

$$\text{Var}\,[Z] = \frac{\text{Var}[X]}{\sigma^2} = 1$$

**Definicja 9.5** (Moment centralny rzędu $n$)**.**
Momentem centralnym rzędu $n$ nazywamy

$$\gamma_n = E[(X - E[X])^n]$$

Zatem wariancja jest momentem centralnym rzędu drugiego

$$\gamma_1 = 0, \qquad \gamma_2 = \text{Var}[X]$$

Współczynnik asymetrii rozkładu

$$A = \frac{\gamma_3}{\sigma^3}$$

Kurtoza (miara spłaszczenia rozkładu)

$$K = \frac{\gamma_4}{\sigma^4} - 3$$

**Definicja 9.6** (Mediana)**.**
Medianą rozkładu nazywamy liczbę $m$

$$P(X \geqslant m) \geqslant \frac{1}{2} \ \wedge \ P(X \leqslant m) \geqslant \frac{1}{2}$$

**Twierdzenie 9.4** (Mediana ciągłej zmiennej losowej)**.**
*Niech $X$ będzie ciągłą zmienną losową o dystrybuancie $F$. Wtedy*

$$F(m) = \frac{1}{2}$$

*Dowód.*

$$1 - F(m) \geqslant \frac{1}{2} \ \wedge \ F(m) \geqslant \frac{1}{2}$$

$$F(m) \leqslant \frac{1}{2} \ \wedge \ F(m) \geqslant \frac{1}{2}$$

$$F(m) = \frac{1}{2}$$

□

**Definicja 9.7** (Moda)**.**
Modą rozkładu zmiennej losowej $X$ o rozkładzie pr. $p$ lub gęstości $f$ nazywamy liczbę $x_m$

$$\begin{cases} p(x_m) \geqslant p(x_i) & \text{dla } i = 1, 2, \ldots \\ f(x_m) \geqslant f(x) & \text{dla } x \in \mathbb{R} \end{cases}$$

# 10 Wektor losowy

Często zdarzenie losowe analizujemy ze względu na dwie cechy. Rozważmy dwie zmienne losowe $X$ i $Y$ określone na tej samej przestrzeni zdarzeń elementarnych $\Omega$. Interesuje nas łączny rozkład prawdopodobieństwa pary zmiennych losowych $(X, Y)$.

**Definicja 10.1** (Wektor losowy)**.**
Niech $X$ i $Y$ będą zmiennymi losowymi. Parę $(X, Y)$ nazywamy (dwuwymiarowym) wektorem losowym

**Przykład 10.1** (Rzut dwoma kostkami do gry)**.**
Rzucamy dwoma kostkami do gry.

- $X$ - suma wyrzuconych oczek
- $Y$ - większa z wyrzuconych liczb.

Zdarzenia elementarne oraz odpowiadające im wartości zmiennych losowych $X$ i $Y$ przedstawia tabelka

$$\Omega = \{(\omega_1, \omega_2)\} \quad X = \omega_1 + \omega_2 \quad Y = \text{Max}(\omega_1, \omega_2)$$

| $\omega$ | X | Y | $\omega$ | X | Y | $\omega$ | X | Y | $\omega$ | X | Y | $\omega$ | X | Y | $\omega$ | X | Y |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1,1) | 2 | 1 | (1,2) | 3 | 2 | (1,3) | 4 | 3 | (1,4) | 5 | 4 | (1,5) | 6 | 5 | (1,6) | 7 | 6 |
| (2,1) | 3 | 2 | (2,2) | 4 | 2 | (2,3) | 5 | 3 | (2,4) | 6 | 4 | (2,5) | 7 | 5 | (2,6) | 8 | 6 |
| (3,1) | 4 | 3 | (3,2) | 5 | 3 | (3,3) | 6 | 3 | (3,4) | 7 | 4 | (3,5) | 8 | 5 | (3,6) | 9 | 6 |
| (4,1) | 5 | 4 | (4,2) | 6 | 4 | (4,3) | 7 | 4 | (4,4) | 8 | 4 | (4,5) | 9 | 5 | (4,6) | 10 | 6 |
| (5,1) | 6 | 5 | (5,2) | 7 | 5 | (5,3) | 8 | 5 | (5,4) | 9 | 5 | (5,5) | 10 | 5 | (5,6) | 11 | 6 |
| (6,1) | 7 | 6 | (6,2) | 8 | 6 | (6,3) | 9 | 6 | (6,4) | 10 | 6 | (6,5) | 11 | 6 | (6,6) | 12 | 6 |

Możliwe wartości wektora losowego $(X, Y)$ oraz odpowiadające im prawdopodobieństwa, $P(X = x, Y = y)$, przedstawia tabelka

| $Y \setminus X$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1/36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 2/36 | 1/36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 2/36 | 2/36 | 1/36 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 2/36 | 2/36 | 2/36 | 1/36 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 2/36 | 2/36 | 2/36 | 2/36 | 1/36 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 2/36 | 2/36 | 2/36 | 2/36 | 2/36 | 1/36 |

Powyższa tabela jest przykładem łącznego rozkładu prawdopodobieństwa dyskretnego wektora losowego.

**Definicja 10.2** (Łączny rozkład prawdopodobieństwa)**.**
Niech $(X, Y)$ będzie dyskretnym wektorem losowym o wartościach $\{(x_j, y_k) : j, k = 1, 2, \dots\}$. Łącznym rozkładem prawdopodobieństwa $(X, Y)$ nazywamy funkcję:

$$p(x_j, y_k) = P(X = x_j, Y = y_k) \qquad (j, k = 1, 2, \dots)$$

Własności łącznego rozkładu pr.

- $p(x_j, y_k) \geqslant 0 \quad j, k = 1, 2, \dots$
- $\sum\limits_{j,k=1} p(x_j, y_k) = 1$

**Definicja 10.3** (Łączna dystrybuanta)**.**
Dystrybuantą $F$ wektora losowego $(X, Y)$ nazywamy funkcję $F : \mathbb{R}^2 \to [0, 1]$

$$F(s, t) = P(X \leqslant s, Y \leqslant t) \qquad -\infty < s, t < \infty$$

**Definicja 10.4** (Łączna gęstość prawdopodobieństwa)**.**
Wektor losowy $(X, Y)$ jest łącznie ciągły (lub ma ciągły łączny rozkład), jeśli istnieje funkcja $f(x, y)$ taka, że:

$$F(s, t) = P(X \leqslant s, Y \leqslant t) = \int\limits_{-\infty}^{s} \int\limits_{-\infty}^{t} f(x, y) dx dy$$

Funkcję $f(x, y)$ nazywamy łączną gęstością pr. wektora losowego $(X, Y)$.