Spis treści

1 Wektory losowe

1.1 Definicja

Definicja
Wektorem losowym $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ nazywamy odwzorowanie

$$\boldsymbol{\xi} : (\Omega, \mathcal{F}, P) \to \mathbb{X} \subseteq \mathbb{R}^n$$

takie, że dla dowolnego $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

$$\{\omega \in \Omega : \ \xi_1(\omega) \le x_1, \ldots, \xi_n(\omega) \le x_n\} \in \mathcal{F}$$

1.2 Rozkłady

Rozkład łączny

Określenie
Rozkładem wektora losowego $\boldsymbol{\xi}$ nazywamy rozkład prawdopodobieństwa $P_{\boldsymbol{\xi}}$ na $\mathbb{X}$ określony wzorem

$$P_{\boldsymbol{\xi}}(A) = P(\{\omega \in \Omega : \ \boldsymbol{\xi}(\omega) \in A\}) \quad \text{dla } A \in \mathcal{B}(\mathbb{X})$$

Dystrybuanta
Dystrybuantą wektora losowego $\boldsymbol{\xi}$ nazywamy funkcję $F_{\boldsymbol{\xi}} : \mathbb{R}^n \to [0, 1]$ określoną wzorem

$$F_{\boldsymbol{\xi}}(x_1, \dots, x_n) = P(\xi_1 \le x_1, \dots, \xi_n \le x_n)$$

Rozkład łączny

Rozkład skokowy (dyskretny)
Wektor losowy jest typu skokowego, jeżeli istnieje zbiór skończony lub przeliczalny $\mathbb{X} \subset \mathbb{R}^n$, taki że $P_{\boldsymbol{\xi}}(\mathbb{X}) = 1$

Rozkład łączny

Rozkład ciągły
Wektor losowy jest typu ciągłego, jeżeli istnieje nieujemna funkcja $f_{\boldsymbol{\xi}}(\boldsymbol{x})$ taka, że dla każdego $\boldsymbol{x} \in \mathbb{R}^n$

$$F_{\boldsymbol{\xi}}(\boldsymbol{x}) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{\boldsymbol{\xi}}(u_1, \dots, u_n)\, du_1 \dots du_n$$

Funkcja $f_{\boldsymbol{\xi}}(\boldsymbol{x})$ nazywa się **gęstością**.

Rozkład łączny

Przykład: rzut dwiema sześciennymi kostkami
ξ_1 – wynik rzutu pierwszą kostką; ξ_2 – wynik rzutu drugą kostką (ξ_1, ξ_2) – wynik rzutu parą kostek Funkcja rozkładu prawdopodobieństwa

$$P\{(\xi_1, \xi_2) = (i, j)\} = p_{ij}, \quad i, j = 1, \dots, 6$$

Dystrybuanta

$$F_{(\xi_1,\xi_2)}(x,y) = P\{\xi_1 \leq x, \xi_2 \leq y\} = \sum_{(i,j): i \leq x, j \leq y} p_{ij}$$

Rozkład łączny

Przykład: strzał do okrągłej tarczy
ξ_1 – odcięta punktu trafienia; ξ_2 – rzędna punktu trafienia (ξ_1, ξ_2) – współrzędne punktu trafienia Funkcja gęstości rozkładu prawdopodobieństwa

$$f_{(\xi_1,\xi_2)}(x,y) = \frac{1}{\pi r^2} \mathbf{1}_{\{x^2+y^2 \leq r^2\}}(x,y)$$

Dystrybuanta $F_{(\xi_1,\xi_2)}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(t,u) du dt$

Rozkład brzegowy

Określenie
Niech $f_{\boldsymbol{\xi}}$ będzie gęstością łącznego rozkładu prawdopodobieństwa wektora losowego $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$. Funkcja gęstości rozkładu brzegowego zmiennej losowej ξ_1 określona jest wzorem

$$f_{\xi_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{\xi}}(x_1, \ldots, x_n)\, dx_2 \ldots dx_n$$

Rozkład brzegowy

Określenie
Niech $f_{\boldsymbol{\xi}}$ będzie gęstością łącznego rozkładu prawdopodobieństwa wektora losowego $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$. Funkcja gęstości rozkładu brzegowego wektora losowego (ξ_1, ξ_2) określona jest wzorem

$$f_{(\xi_1,\xi_2)}(x_1, x_2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{\xi}}(x_1, \ldots, x_n)\, dx_3 \ldots dx_n$$

Rozkład brzegowy

Przykład: rzut dwiema sześciennymi kostkami
(ξ_1, ξ_2) – wynik rzutu parą kostek Funkcja rozkładu prawdopodobieństwa zmiennej losowej ξ_1

$$P\{\xi_1 = i\} = \sum_{j=1}^{6} P\{(\xi_1, \xi_2) = (i, j)\} = \sum_{j=1}^{6} p_{ij}, \quad i = 1, \ldots, 6$$

Dystrybuanta rozkładu zmiennej losowej ξ_1

$$F_{\xi_1}(x) = P\{\xi_1 \leq x\} = \sum_{i \leq x} P\{\xi_1 = i\}$$

Rozkład brzegowy

Przykład: strzał do okrągłej tarczy
(ξ_1, ξ_2) – współrzędne punktu trafienia Funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej ξ_1

$$f_{\xi_1}(x) = \int_{-\infty}^{\infty} f_{(\xi_1,\xi_2)}(x, u) du$$

Dystrybuanta

$$F_{\xi_1}(x) = \int_{-\infty}^{x} f_{\xi_1}(t) dt$$

Rozkład brzegowy

Przykład: strzał do okrągłej tarczy

$$\begin{aligned} f_{\xi_1}(x) &= \int_{-\infty}^{\infty} f_{(\xi_1,\xi_2)}(x, u) du \\ &= \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \frac{1}{\pi r^2} du \\ &= \frac{2\sqrt{r^2 - x^2}}{\pi r^2}, \qquad \text{dla } x \in (-r, r) \end{aligned}$$

Rozkład warunkowy

Określenie

Niech $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$. Rozkład warunkowy wektora $\boldsymbol{\xi}_1$ pod warunkiem $\boldsymbol{\xi}_2 = \boldsymbol{x}_2$ opisany jest funkcją gęstości

$$f_{\boldsymbol{\xi}_1|\boldsymbol{\xi}_2=\boldsymbol{x}_2}(\boldsymbol{x}_1) = \frac{f_{\boldsymbol{\xi}}(\boldsymbol{x}_1, \boldsymbol{x}_2)}{f_{\boldsymbol{\xi}_2}(\boldsymbol{x}_2)}$$

Niezależność

Wektory losowe $\boldsymbol{\xi}_1$ oraz $\boldsymbol{\xi}_2$ są niezależne, jeżeli dla wszystkich $\boldsymbol{x}_1$ oraz $\boldsymbol{x}_2$

$$f_{\boldsymbol{\xi}}(\boldsymbol{x}_1, \boldsymbol{x}_2) = f_{\boldsymbol{\xi}_1}(\boldsymbol{x}_1) f_{\boldsymbol{\xi}_2}(\boldsymbol{x}_2)$$

Rozkład warunkowy

Przykład: rzut dwiema sześciennymi kostkami

(ξ_1, ξ_2) – wynik rzutu parą kostek Funkcja rozkładu prawdopodobieństwa zmiennej losowej ξ_2 pod warunkiem $\xi_1 = i$

$$P\{\xi_2 = j | \xi_1 = i\} = \frac{P\{(\xi_1, \xi_2) = (i, j)\}}{P\{\xi_1 = i\}}, \quad j = 1, \ldots, 6$$

Rozkład warunkowy

Przykład: rzut dwiema sześciennymi kostkami

(ξ_1, ξ_2) – wynik rzutu parą kostek Niezależność

$$P\{\xi_2 = j | \xi_1 = i\} = P\{\xi_2 = j\}, \quad j = 1, \ldots, 6,\ i = 1, \ldots, 6$$

$$P\{(\xi_1, \xi_2) = (i, j)\} = P\{\xi_1 = i\} P\{\xi_2 = j\}, \quad i, j = 1, \ldots, 6$$

Rozkład warunkowy

Przykład: strzał do okrągłej tarczy

(ξ_1, ξ_2) – współrzędne punktu trafienia Funkcja gęstości rozkładu prawdopodobieństwa zmiennej losowej ξ_2 pod warunkiem $\xi_1 = x$

$$f_{\xi_2|\xi_1=x}(y) = \frac{f_{(\xi_1,\xi_2)}(x, y)}{f_{\xi_1}(x)}$$

Rozkład warunkowy

Przykład: strzał do okrągłej tarczy

$$\begin{aligned} f_{\xi_2|\xi_1=x}(y) &= \frac{f_{(\xi_1,\xi_2)}(x,y)}{f_{\xi_1}(x)} \\ &= \frac{1/\pi r^2}{2\sqrt{r^2-x^2}/\pi r^2} \\ &= \frac{1}{2\sqrt{r^2-x^2}}, \qquad \text{dla } y \in (-\sqrt{r^2-x^2}, \sqrt{r^2-x^2}) \end{aligned}$$

1.3 Momenty

Momenty

Momenty rzędu k

Niech $k_1, \ldots, k_n$ będą takimi nieujemnymi liczbami całkowitymi, że $k_1 + \cdots + k_n = k$. Momentem rzędu k nazywamy

$$E\xi_1^{k_1} \cdots \xi_n^{k_n} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{k_1} \cdots x_n^{k_n} f_{\boldsymbol{\xi}}(\boldsymbol{x}) d\boldsymbol{x}$$

Moment rzędu pierwszego

Wartość oczekiwana

Jeżeli $k = 1$, to jest n możliwości:

$$\begin{aligned} k_1 = 1, k_2 = 0, \ldots, k_n = 0 &\Rightarrow E\xi_1 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 f_{\boldsymbol{\xi}}(\boldsymbol{x}) d\boldsymbol{x} \\ k_1 = 0, k_2 = 1, \ldots, k_n = 0 &\Rightarrow E\xi_2 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_2 f_{\boldsymbol{\xi}}(\boldsymbol{x}) d\boldsymbol{x} \\ &\vdots \\ k_1 = 0, k_2 = 0, \ldots, k_n = 1 &\Rightarrow E\xi_n = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_n f_{\boldsymbol{\xi}}(\boldsymbol{x}) d\boldsymbol{x} \end{aligned}$$

Moment rzędu pierwszego

Wartość oczekiwana

$$\begin{aligned} E\xi_1 &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 f_{\boldsymbol{\xi}}(\boldsymbol{x}) d\boldsymbol{x} \\ &= \int_{-\infty}^{\infty} x_1 \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{\xi}}(x_1, x_2, \ldots, x_n) dx_2 \ldots dx_n \right] dx_1 \\ &= \int_{-\infty}^{\infty} x_1 f_{\xi_1}(x_1) dx_1 \end{aligned}$$

Moment rzędu pierwszego

Wartość oczekiwana

$$E\boldsymbol{\xi} = \begin{bmatrix} E\xi_1 \\ \vdots \\ E\xi_n \end{bmatrix}$$

Momenty centralne

Momenty centralne rzędu k

Niech $k_1, \ldots, k_n$ będą takimi nieujemnymi liczbami całkowitymi, że $k_1 + \cdots + k_n = k$. Momentem centralnym rzędu k nazywamy

$$E(\xi_1 - E\xi_1)^{k_1} \cdots (\xi_n - E\xi_n)^{k_n}$$

Momenty centralne

Moment centralny rzędu 2

$$k_1 = 2, k_2 = 0, \ldots, k_n = 0 \Rightarrow$$

$$E(\xi_1 - E\xi_1)^2 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 - E\xi_1)^2 f_{\boldsymbol{\xi}}(\boldsymbol{x}) d\boldsymbol{x}$$

$$= \int_{-\infty}^{\infty} (x_1 - E\xi_1)^2 \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{\xi}}(x_1, x_2, \ldots, x_n) dx_2 \ldots dx_n \right] dx_1$$

$$= \int_{-\infty}^{\infty} (x_1 - E\xi_1)^2 f_{\xi_1}(x_1) dx_1 = D^2\xi_1$$

Momenty centralne

Moment centralny rzędu 2

$$k_1 = 2, k_2 = 0, \ldots, k_n = 0 \Rightarrow D^2\xi_1$$
$$k_1 = 0, k_2 = 2, \ldots, k_n = 0 \Rightarrow D^2\xi_2$$
$$\vdots$$
$$k_1 = 0, k_2 = 0, \ldots, k_n = 2 \Rightarrow D^2\xi_n$$

Momenty centralne

Kowariancja

$$\begin{aligned}
&k_1 = 1, k_2 = 1, \ldots, k_n = 0 \Rightarrow E(\xi_1 - E\xi_1)(\xi_2 - E\xi_2) \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 - E\xi_1)(x_2 - E\xi_2) f_{\boldsymbol{\xi}}(\boldsymbol{x}) d\boldsymbol{x} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - E\xi_1)(x_2 - E\xi_2) f_{(\xi_1,\xi_2)}(x_1, x_2) dx_1 dx_2 \\
&= Cov(\xi_1, \xi_2)
\end{aligned}$$

$$Cov(\xi_i, \xi_j) = E\xi_i\xi_j - E\xi_i E\xi_j$$

Momenty centralne

Macierz kowariancji

$$\begin{bmatrix}
D^2\xi_1 & Cov(\xi_1, \xi_2) & \cdots & Cov(\xi_1, \xi_n) \\
Cov(\xi_1, \xi_2) & D^2\xi_2 & \cdots & Cov(\xi_2, \xi_n) \\
\vdots & \vdots & & \vdots \\
Cov(\xi_1, \xi_n) & Cov(\xi_2, \xi_n) & \cdots & D^2\xi_n
\end{bmatrix}$$

Momenty centralne

Przykład: rzut dwiema sześciennymi kostkami

(ξ_1, ξ_2) – wynik rzutu parą kostek Kowariancja

$$\begin{aligned}
Cov(\xi_1, \xi_2) &= \sum_{i,j} (x_i - E\xi_1)(x_j - E\xi_2) P\{(\xi_1, \xi_2) = (i, j)\} \\
&= \sum_{i,j} x_i x_j P\{(\xi_1, \xi_2) = (i, j)\} - E\xi_1 E\xi_2
\end{aligned}$$

Momenty centralne

Przykład: strzał do okrągłej tarczy
(ξ_1, ξ_2) – współrzędne punktu trafienia Kowariancja

$$
\begin{aligned}
Cov(\xi_1, \xi_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - E\xi_1)(y - E\xi_2)f_{(\xi_1,\xi_2)}(x, y)dxdy \\
&= \int\int_{\mathbb{R}^2} xyf_{(\xi_1,\xi_2)}(x, y)dxdy - E\xi_1 E\xi_2
\end{aligned}
$$

Momenty centralne

Ważny fakt
Jeżeli ξ_1 oraz ξ_2 są niezależne, to

$$Cov(\xi_1, \xi_2) = 0$$

$$
\begin{aligned}
Cov(\xi_1, \xi_2) &= \int\int_{\mathbb{R}^2} xyf_{(\xi_1,\xi_2)}(x, y)dxdy - E\xi_1 E\xi_2 \\
&= \int_{\mathbb{R}} xf_{\xi_1}(x)dx \int_{\mathbb{R}} yf_{\xi_2}(y)dy - E\xi_1 E\xi_2 \\
&= E\xi_1 E\xi_2 - E\xi_1 E\xi_2
\end{aligned}
$$

Momenty centralne

Ważny fakt
Jeżeli $Cov(\xi_1, \xi_2) = 0$, to ξ_1 oraz ξ_2 nie muszą być niezależne!

Przykład
Zmienne losowe (ξ_1, ξ_2) mają łączny rozkład

ξ_2 \ ξ_1	-1	0	1
-1	0	0.25	0
0	0.25	0	0.25
1	0	0.25	0

Momenty centralne

Przykład
Rozkłady brzegowe zmiennych losowych ξ_1 oraz ξ_2

-1	0	1
0.25	0.5	0.25

$$E\xi_1 = E\xi_2 = 0$$

Momenty centralne

Przykład

$$\begin{aligned}
&E\xi_1\xi_2 = \\
&(-1)\cdot(-1)\cdot P\{(\xi_1,\xi_2) = (-1,-1)\} + \\
&(-1)\cdot 0 \cdot P\{(\xi_1,\xi_2) = (-1,0)\} + \\
&\ldots \\
&1\cdot 0 \cdot P\{(\xi_1,\xi_2) = (1,0)\} + \\
&1\cdot 1 \cdot P\{(\xi_1,\xi_2) = (1,1)\} \\
&= 0
\end{aligned}$$

Momenty centralne

Przykład
Czyli

$$Cov(\xi_1,\xi_2) = 0$$

Ale

$$P\{(\xi_1,\xi_2) = (-1,-1)\} \neq P\{\xi_1 = -1\} \cdot P\{\xi_2 = -1\}$$

Zmienne losowe ξ_1 oraz ξ_2 nie są niezależne

Momenty centralne

Współczynnik korelacji liniowej

$$\varrho(\xi_i,\xi_j) = \frac{Cov(\xi_i,\xi_j)}{\sqrt{D^2\xi_i D^2\xi_j}}$$

Ważne własności

$$-1 \leq \varrho(\xi_i, \xi_j) \leq 1$$

Jeżeli $|\varrho(\xi_i, \xi_j)| = 1$, to istnieją takie liczby rzeczywiste a i b, że

$$\xi_1 = a\xi_2 + b$$

Własności wynikają z nierówności Schwarza

Momenty centralne

Ważna własność
Jeżeli $\boldsymbol{A}$ jest macierzą liczbową odpowiednich wymiarów, to

$$E\boldsymbol{A\xi} = \boldsymbol{A}E\boldsymbol{\xi}$$

$$D\boldsymbol{A\xi} = \boldsymbol{A}D\boldsymbol{\xi}\boldsymbol{A}^T$$

1.4 Warunkowa wartość oczekiwana

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$

Definicja
Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$ zmiennej losowej ξ_1 pod warunkiem zmiennej losowej ξ_2 to nowa zmienna losowa, która jest postaci $E(\xi_1|\xi_2) = m(\xi_2)$ z prawdopodobieństwem 1 dla borelowskiej funkcji $m(x)$ takiej, że

$$\int_B m(y)dF_{\xi_2}(y) = \int\int_{B\times\mathbb{R}} xdF_{\xi_1,\xi_2}(x,y)$$

dla każdego borelowskiego zbioru B.

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$

Przykład: rzut dwiema sześciennymi kostkami
(ξ_1, ξ_2) – wynik rzutu parą kostek Rozkład zmiennej losowej $E(\xi_1|\xi_2)$

$$\begin{array}{ll} E(\xi_1|\xi_2 = 1) & P\{\xi_2 = 1\} \\ E(\xi_1|\xi_2 = 2) & P\{\xi_2 = 2\} \\ E(\xi_1|\xi_2 = 3) & P\{\xi_2 = 3\} \\ E(\xi_1|\xi_2 = 4) & P\{\xi_2 = 4\} \\ E(\xi_1|\xi_2 = 5) & P\{\xi_2 = 5\} \\ E(\xi_1|\xi_2 = 6) & P\{\xi_2 = 6\} \end{array}$$

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$

Przykład: rzut dwiema sześciennymi kostkami

$$
\begin{aligned}
E(\xi_1|\xi_2 = j) &= \sum_{i=1}^{6} iP\{\xi_1 = i|\xi_2 = j\} \\
&= \sum_{i=1}^{6} i\frac{P\{\xi_1 = i, \xi_2 = j\}}{P\{\xi_2 = j\}} \\
&= \sum_{i=1}^{6} i\frac{p_{ij}}{\sum_{k=1}^{6} p_{kj}}
\end{aligned}
$$

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$

Przykład: strzał do okrągłej tarczy
(ξ_1, ξ_2) – współrzędne punktu trafienia Rozkład zmiennej losowej $E(\xi_1|\xi_2)$

$$
\begin{aligned}
E(\xi_1|\xi_2 = y) &= \int_{-\infty}^{\infty} x f_{\xi_1|\xi_2=y}(x)dx \\
&= \int_{\mathbb{R}} x\frac{f_{(\xi_1,\xi_2)}(x, y)}{f_{\xi_2}(y)}dx \\
&= \int_{\mathbb{R}} x\frac{f_{(\xi_1,\xi_2)}(x, y)}{\int_{\mathbb{R}} f_{(\xi_1,\xi_2)}(t, y)dt}dx
\end{aligned}
$$

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$

Wartość oczekiwana

$$
\begin{aligned}
E(E(\xi_1|\xi_2)) &= \int_{-\infty}^{\infty} E(\xi_1|\xi_2 = x_2) f_{\xi_2}(x_2)dx_2 \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x_1 f_{\xi_1|\xi_2=x_2}(x_1)dx_1\right] f_{\xi_2}(x_2)dx_2 \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x_1\frac{f_{(\xi_1,\xi_2)}(x_1, x_2)}{f_{\xi_2}(x_2)}dx_1\right] f_{\xi_2}(x_2)dx_2 \\
&= E\xi_1
\end{aligned}
$$

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$

Równość wariancyjna

$$D^2\xi_1 = E\left(D^2(\xi_1|\xi_2)\right) + D^2\left(E(\xi_1|\xi_2)\right)$$

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$ **- przykład**

Rozkład łączny zmiennych losowych (ξ_1, ξ_2)

ξ_2 \ ξ_1	-1	0	1
-1	0.1	0.2	0.0
0	0.0	0.2	0.1
1	0.1	0.2	0.1

Rozkłady brzegowe

$$\xi_1 : \frac{-1 \quad 0 \quad 1}{0.2 \quad 0.6 \quad 0.2} \qquad \xi_2 : \frac{-1 \quad 0 \quad 1}{0.3 \quad 0.3 \quad 0.4}$$

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$ **- przykład**

Rozkłady warunkowe

$$\xi_1|\xi_2 = -1 : \frac{-1 \quad 0 \quad 1}{1/3 \quad 2/3 \quad 0}$$

$$\xi_1|\xi_2 = 0 : \frac{-1 \quad 0 \quad 1}{0 \quad 2/3 \quad 1/3}$$

$$\xi_1|\xi_2 = 1 : \frac{-1 \quad 0 \quad 1}{1/4 \quad 1/2 \quad 1/4}$$

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$ **- przykład**

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$

$E(\xi_1|\xi_2=-1)$	$= -1/3$	z prawd.	$P\{\xi_2=-1\}$	$= 0.3$
$E(\xi_1|\xi_2=0)$	$= 1/3$	z prawd.	$P\{\xi_2=0\}$	$= 0.3$
$E(\xi_1|\xi_2=1)$	$= 0$	z prawd.	$P\{\xi_2=1\}$	$= 0.4$

$$E(E(\xi_1|\xi_2)) = 0 = E\xi_1$$

$$D^2(E(\xi_1|\xi_2)) = 2/30$$

Warunkowa wartość oczekiwana $E(\xi_1|\xi_2)$ - przykład

Warunkowa wariancja $D^2(\xi_1|\xi_2)$

$$\begin{array}{llll} D^2(\xi_1|\xi_2=-1) & = 2/9 & \text{z prawd.} \quad P\{\xi_2=-1\} & = 0.3 \\ D^2(\xi_1|\xi_2=0) & = 2/9 & \text{z prawd.} \quad P\{\xi_2=0\} & = 0.3 \\ D^2(\xi_1|\xi_2=1) & = 1/2 & \text{z prawd.} \quad P\{\xi_2=1\} & = 0.4 \end{array}$$

$$E(D^2(\xi_1|\xi_2)) = 1/3$$

$$E(D^2(\xi_1|\xi_2)) + D^2(E(\xi_1|\xi_2)) = 1/3 + 2/30 = 0.4 = D^2\xi_1$$

1.5 Wielowymiarowy rozkład normalny

Wielowymiarowy rozkład normalny $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Określenie

Wektor losowy $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_p]^T$ ma p-wymiarowy rozkład normalny, jeżeli jego funkcja gęstości prawdopodobieństwa wyraża się wzorem

$$f_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{p/2}\sqrt{|\boldsymbol{\Sigma}|}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

dla $\boldsymbol{x} \in \mathbb{R}^p$

Wielowymiarowy rozkład normalny $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Momenty

Wektor wartości oczekiwanych $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_p]^T$ Macierz kowariancji $\boldsymbol{\Sigma} = [\sigma_{ij}]_{i,j=1,\ldots,p}$

$$E\xi_i = \mu_i,\ D^2\xi_i = \sigma_{ii},\ Cov(\xi_i, \xi_j) = \sigma_{ij}$$

$$D^2\xi_i = \sigma_i^2$$

Wielowymiarowy rozkład normalny $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Rozkłady brzegowe

Fakt: jeżeli $\boldsymbol{A}$ jest $r \times p$ macierzą, to

$$\boldsymbol{A\xi} \sim N_r(\boldsymbol{A\mu}, \boldsymbol{A\Sigma A}^T)$$

Każdy rozkład brzegowy jest normalny $\xi_1 = [1, 0, \ldots, 0]\boldsymbol{\xi}$

Wielowymiarowy rozkład normalny $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Rozkłady warunkowe

Niech

$$\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \end{bmatrix} \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^T & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

Rozkład wektora $\boldsymbol{\xi}_1$ pod warunkiem $\boldsymbol{\xi}_2 = \boldsymbol{x}_2$ jest rozkładem normalnym

$$N_m(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{12}^T(\boldsymbol{x}_2 - \mu_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{12}^T)$$

Wielowymiarowy rozkład normalny $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Rozkłady warunkowe

$$E(\boldsymbol{\xi}_1|\boldsymbol{\xi}_2) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{12}^T(\boldsymbol{\xi}_2 - \mu_2)$$

Wielowymiarowy rozkład normalny $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Rozkłady warunkowe

$\boldsymbol{\xi}_1$ oraz $\boldsymbol{\xi}_2$ są niezależne wtedy i tylko wtedy, gdy

$$\boldsymbol{\Sigma}_{12} = \mathbf{0}$$

W szczególności ξ_i oraz ξ_j są niezależne wtedy i tylko wtedy, gdy $\sigma_{ij} = 0$

2 Funkcje wektorów losowych

2.1 Suma zmiennych losowych

Rozkład sumy zmiennych losowych

Niech $\boldsymbol{\xi} = (\xi_1, \xi_2)$ o gęstości $f_{\boldsymbol{\xi}}(x_1, x_2)$. Niech $\eta = \xi_1 + \xi_2$

$$F_\eta(x) = P\{\xi_1 + \xi_2 \leq x\}$$

$$f_\eta(x) = \frac{d}{dx}F_\eta(x)$$

Rozkład sumy zmiennych losowych

$$\begin{aligned}
F_\eta(x) &= P\{\xi_1 + \xi_2 \le x\} \\
&= \int_{-\infty}^{\infty} P\{\xi_1 \le x - x_2 | \xi_2 = x_2\} f_{\xi_2}(x_2) dx_2 \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{x-x_2} f_{\xi_1|\xi_2=x_2}(x_1) dx_1 \right] f_{\xi_2}(x_2) dx_2 \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{x} f_{\xi_1|\xi_2=x_2}(t - x_2) dt \right] f_{\xi_2}(x_2) dx_2 \\
&= \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f_{\xi_1|\xi_2=x_2}(t - x_2) f_{\xi_2}(x_2) dx_2 \right] dt
\end{aligned}$$

Rozkład sumy zmiennych losowych

Funkcja gęstości

$$f_\eta(t) = \int_{-\infty}^{\infty} f_{\xi_1|\xi_2=x_2}(t - x_2) f_{\xi_2}(x_2) dx_2$$

Jeżeli ξ_1 oraz ξ_2 są niezależne, to

$$f_\eta(t) = \int_{-\infty}^{\infty} f_{\xi_1}(t - x_2) f_{\xi_2}(x_2) dx_2$$

Rozkład sumy zmiennych losowych

Wartość oczekiwana

$$E\eta = \int_{\mathbb{R}} x f_\eta(x) dx$$

Rozkład sumy zmiennych losowych

Wartość oczekiwana

$$
\begin{aligned}
& E(\xi_1+\xi_2) \\
&= \int\int_{\mathbb{R}^2} (x+y) f_{(\xi_1,\xi_2)}(x,y)dxdy \\
&= \int_{\mathbb{R}} x \left[\int_{\mathbb{R}} f_{(\xi_1,\xi_2)}(x,y)dy\right] dx + \int_{\mathbb{R}} y \left[\int_{\mathbb{R}} f_{(\xi_1,\xi_2)}(x,y)dx\right] dy \\
&= \int_{\mathbb{R}} x f_{\xi_1}(x)dx + \int_{\mathbb{R}} y f_{\xi_2}(y)dy \\
&= E\xi_1 + E\xi_2
\end{aligned}
$$

Rozkład sumy zmiennych losowych

Wariancja

$$D^2(\xi_1+\xi_2) = D^2\xi_1 + D^2\xi_2 + 2Cov(\xi_1,\xi_2)$$

Jeżeli ξ_1 oraz ξ_2 są niezależne, to

$$D^2(\xi_1+\xi_2) = D^2\xi_1 + D^2\xi_2$$

2.2 Rozkład dwumianowy

Rozkład dwumianowy

Twierdzenie
Jeżeli $\xi_1, \xi_2, \ldots, \xi_n$ są niezależnymi zmiennymi losowymi o rozkładach $B(n_i, p)$, $i = 1, \ldots, n$, to zmienna losowa $\eta = \sum_{i=1}^n \xi_i$ ma rozkład

$$B\left(\sum_{i=1}^n n_i, p\right)$$

Wniosek
Jeżeli $\xi_1, \xi_2, \ldots, \xi_n$ są i.i.d. $D(p)$, $i = 1, \ldots, n$, to $\eta = \sum_{i=1}^n \xi_i \sim B\left(n, p\right)$

Rozkład dwumianowy

Dowód dla $n = 2$
Jeżeli ξ_1, ξ_2 są niezależnymi zmiennymi losowymi o rozkładach $B(n_1, p)$ i $B(n_2, p)$, to zmienna losowa $\eta = \xi_1 + \xi_2$ ma rozkład

$$B\left(n_1 + n_2, p\right)$$

Rozkład dwumianowy

Dowód dla $n = 2$
Zbiorem wartości zmiennej losowej η jest $\{0, 1, \ldots, n_1 + n_2\}$

$$\begin{aligned}P\{\eta = k\} = P\{\xi_1 + \xi_2 = k\} &= P\{\xi_1 = k - \xi_2\} \\ &= \sum_{j=0}^{n_2} P\{\xi_1 = k - j | \xi_2 = j\} P\{\xi_2 = j\} \\ &= \sum_{j=0}^{n_2} P\{\xi_1 = k - j\} P\{\xi_2 = j\}\end{aligned}$$

Rozkład dwumianowy

Dowód dla $n = 2$

$$\begin{aligned}&= \sum_{j=0}^{k} \binom{n_1}{k-j} p^{k-j} (1-p)^{n_1-k+j} p^j (1-p)^{n_2-j} \\ &= p^k (1-p)^{n_1+n_2-k} \sum_{j=0}^{k} \binom{n_1}{k-j} \binom{n_2}{j} \\ &= \binom{n_1+n_2}{k} p^k (1-p)^{n_1+n_2-k}\end{aligned}$$

Dla dowolnego n: zasada indukcji matematycznej

2.3 Rozkład ujemny dwumianowy

Rozkład ujemny dwumianowy

Twierdzenie
Jeżeli $\xi_1, \xi_2, \ldots, \xi_n$ są niezależnymi zmiennymi losowymi o rozkładach $NB(r_i, p)$, $i = 1, \ldots, n$, to zmienna losowa $\eta = \sum_{i=1}^{n} \xi_i$ ma rozkład

$$NB\left(\sum_{i=1}^{n} r_i, p\right)$$

Wniosek
Jeżeli $\xi_1, \xi_2, \ldots, \xi_n$ są i.i.d. $G(p)$, $i = 1, \ldots, n$, to $\eta = \sum_{i=1}^{n} \xi_i \sim NB\,(n, p)$

2.4 Rozkład Poissona

Rozkład Poissona

Twierdzenie
Jeżeli $\xi_1, \xi_2, \ldots, \xi_n$ są niezależnymi zmiennymi losowymi o rozkładach $Po(\lambda_i)$, $i = 1, \ldots, n$, to zmienna losowa $\eta = \sum_{i=1}^{n} \xi_i$ ma rozkład

$$Po\left(\sum_{i=1}^{n} \lambda_i\right)$$

2.5 Rozkład gamma

Rozkład gamma

Twierdzenie
Jeżeli $\xi_1, \xi_2, \ldots, \xi_n$ są niezależnymi zmiennymi losowymi o rozkładach $G(\alpha_i, \lambda)$, $i = 1, \ldots, n$, to zmienna losowa $\eta = \sum_{i=1}^{n} \xi_i$ ma rozkład

$$G\left(\sum_{i=1}^{n} \alpha_i, \lambda\right)$$

Wniosek
Jeżeli $\xi_1, \xi_2, \ldots, \xi_n$ są i.i.d. $E(\lambda)$, $i = 1, \ldots, n$, to $\eta = \sum_{i=1}^{n} \xi_i \sim G(n, \lambda)$

Rozkład gamma

Dowód dla $n = 2$
Jeżeli ξ_1, ξ_2 są niezależnymi zmiennymi losowymi o rozkładach $G(\alpha_1, \lambda)$ i $G(\alpha_2, \lambda)$, to zmienna losowa $\eta = \xi_1 + \xi_2$ ma rozkład

$$G(\alpha_1 + \alpha_2, \lambda)$$

Rozkład gamma

Dowód dla $n = 2$
Zbiorem wartości zmiennej losowej η jest $(0, \infty)$

$$
\begin{aligned}
f_\eta(t) &= \int_{-\infty}^{\infty} f_{\xi_1}(t-u) f_{\xi_2}(u) du \\
&= \int_0^t \frac{1}{\lambda^{\alpha_1}\Gamma(\alpha_1)} (t-u)^{\alpha_1 - 1} e^{-\frac{t-u}{\lambda}} \frac{1}{\lambda^{\alpha_2}\Gamma(\alpha_2)} u^{\alpha_2 - 1} e^{-\frac{u}{\lambda}} du \\
&= \frac{1}{\lambda^{\alpha_1+\alpha_2}\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\frac{t}{\lambda}} \int_0^t (t-u)^{\alpha_1 - 1} u^{\alpha_2 - 1} du
\end{aligned}
$$

Rozkład gamma

Dowód dla $n = 2$

$$
\begin{aligned}
&= \frac{1}{\lambda^{\alpha_1+\alpha_2}\Gamma(\alpha_1)\Gamma(\alpha_2)} t^{\alpha_1+\alpha_2-1} e^{-\frac{t}{\lambda}} \int_0^1 (1-u)^{\alpha_1 - 1} u^{\alpha_2 - 1} du \\
&= \frac{1}{\lambda^{\alpha_1+\alpha_2}\Gamma(\alpha_1+\alpha_2)} t^{\alpha_1+\alpha_2-1} e^{-\frac{t}{\lambda}}
\end{aligned}
$$

Dla dowolnego n: zasada indukcji matematycznej

2.6 Rozkład normalny

Rozkład normalny

Twierdzenie
Jeżeli $\xi_1, \xi_2, \ldots, \xi_n$ są niezależnymi zmiennymi losowymi o rozkładach $N(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$, to zmienna losowa $\eta = \sum_{i=1}^n \xi_i$ ma rozkład

$$
N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)
$$

2.7 Rozkład normalny - inne funkcje

Rozkład chi-kwadrat

Twierdzenie
Jeżeli ξ ma rozkład $N(0,1)$, to ξ^2 ma rozkład chi-kwadrat z jednym stopniem swobody, tzn. ma rozkład $G(1/2,2)$.

Wniosek
Jeżeli $\xi_1,\ldots,\xi_n$ są niezależnymi zmiennymi losowymi o rozkładach $N(0,1)$, to $\eta=\sum_{i=1}^n \xi_i^2$ ma rozkład chi-kwadrat z n stopniami swobody.

Rozkład chi-kwadrat

Formy kwadratowe
Niech $\boldsymbol{\xi}$ ma rozkład $N_n(\mathbf{0},\boldsymbol{I})$ oraz niech $\boldsymbol{A}$ będzie macierzą symetryczną rzędu r. Forma kwadratowa

$$\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}$$

ma rozkład chi-kwadrat z r stopniami swobody wtedy i tylko wtedy, gdy $\boldsymbol{A}$ jest macierzą idempotentną, tzn. $\boldsymbol{A}\boldsymbol{A}=\boldsymbol{A}$.

Rozkład chi-kwadrat

Formy kwadratowe
Niech $\boldsymbol{\xi}$ ma rozkład $N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$ oraz niech $\boldsymbol{A}$ będzie macierzą symetryczną. Forma kwadratowa

$$(\boldsymbol{\xi}-\boldsymbol{\mu})^T\boldsymbol{A}(\boldsymbol{\xi}-\boldsymbol{\mu})$$

ma rozkład chi-kwadrat z tr$(\boldsymbol{A}\boldsymbol{\Sigma})$ stopniami swobody wtedy i tylko wtedy, gdy

$$\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}=\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}$$

Rozkład t

Twierdzenie
Jeżeli ξ ma rozkład $N(0,1)$, η ma rozkład chi-kwadrat z n stopniami swobody, zmienne losowe ξ oraz η są niezależne, to zmienna losowa

$$\zeta=\frac{\xi}{\sqrt{\eta/n}}$$

ma rozkład t z n stopniami swobody.

Rozkład t

Rozkład t

$$f_\zeta = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi}\sqrt{n}} \cdot \frac{1}{(1+x^2/n)^{(n+1)/2}}, \ -\infty < x < \infty.$$

$$E\zeta^{2k-1} = 0 \quad E\zeta^{2k} = \frac{n^k}{\sqrt{\pi}} \frac{\Gamma\left(\frac{n}{2}-k\right)\Gamma\left(k+\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \ (2k < n)$$

$$D^2\zeta = \begin{cases} \frac{n}{n-2}, & n > 2, \\ \infty, & n \le 2. \end{cases}$$

Rozkład F

Twierdzenie
Jeżeli zmienna losowa ξ ma rozkład chi–kwadrat z n stopniami swobody, zmienna losowa η ma rozkład chi–kwadrat z m stopniami swobody i zmienne te są niezależne, to zmienna losowa

$$\zeta = \frac{\xi/n}{\eta/m}$$

ma rozkład F z (n, m) stopniami swobody.

Rozkład F

Rozkład F

$$f_\zeta(x) = \frac{n^{n/2} m^{m/2} \Gamma(\frac{n+m}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} x^{n/2-1}(m+nx)^{-(n+m)/2}, x > 0.$$

$$E\zeta^k = \frac{\Gamma\left(\frac{m}{2}+k\right)\Gamma\left(\frac{m}{2}-k\right) m^k}{\Gamma\left(\frac{n}{2}\right) n^k \Gamma\left(\frac{m}{2}\right)} \ (2k < m)$$

$$D^2\zeta = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)} \ (m > 4)$$

3 Twierdzenia graniczne

3.1 Rodzaje zbieżności

Zbieżność prawie na pewno

Definicja
Ciąg zmiennych losowych $(\xi_n)_{n=1}^{\infty}$ jest zbieżny do zmiennej losowej ξ **prawie na pewno**, jeśli

$$P\left(\left\{\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\}\right) = 1$$

Oznaczenie: $\xi_n \xrightarrow{p.n.} \xi$

Zbieżność według prawdopodobieństwa

Definicja
Ciąg zmiennych losowych $(\xi_n)_{n=1}^{\infty}$ jest zbieżny do zmiennej losowej ξ **według prawdopodobieństwa**, jeśli

$$(\forall \varepsilon > 0)\ \lim_{n\to\infty} P(|\xi_n - \xi| > \varepsilon) = 0$$

Oznaczenie: $\xi_n \xrightarrow{P} \xi$

Zbieżność według rozkładu

Definicja
Ciąg zmiennych losowych $(\xi_n)_{n=1}^{\infty}$ jest zbieżny do zmiennej losowej ξ **według rozkładu**, jeśli ciąg dystrybuant $(F_{\xi_n})_{n=1}^{\infty}$ jest zbieżny do dystrybuanty F_ξ w każdym punkcie jej ciągłości

Oznaczenie:

$$\xi_n \xrightarrow{D} \xi$$

$$\xi_n \xrightarrow{p.n.} \xi \Rightarrow \xi_n \xrightarrow{P} \xi \Rightarrow \xi_n \xrightarrow{D} \xi$$

3.2 Prawa Wielkich Liczb

Oznaczenia

Oznaczenia
$\xi_1, \xi_2, \ldots$ niezależne zmienne losowe o tym samym rozkładzie takim, że $E\xi_1 = \mu$, $D^2\xi_1 = \sigma^2 \in (0, \infty)$

$$S_n = \xi_1 + \xi_2 + \cdots + \xi_n$$

$$\bar{\xi}_n = \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{n} = \frac{S_n}{n}$$

MPWL

Mocne Prawo Wielkich Liczb
Dla ciągu $\xi_1, \xi_2, \ldots$ zachodzi **mocne prawo wielkich liczb**, jeżeli

$$(\forall \varepsilon > 0)\ \ P\left(\lim_{n\to\infty} \bar{\xi}_n = \mu\right) = 1$$

Zapis symboliczny

$$\bar{\xi}_n \xrightarrow{p.n.} \mu$$

SPWL

Słabe Prawo Wielkich Liczb
Dla ciągu $\xi_1, \xi_2, \ldots$ zachodzi **słabe prawo wielkich liczb**, jeżeli

$$(\forall \varepsilon > 0)\ \ \lim_{n\to\infty} P\left(\left|\bar{\xi}_n - \mu\right| < \varepsilon\right) = 1$$

Zapis symboliczny

$$\bar{\xi}_n \xrightarrow{P} \mu$$

Metoda Monte-Carlo

Szacowanie liczby π

Niech $\eta \sim U(0,1)$. Niech $\xi = 4\sqrt{1-\eta^2}$.

$$\begin{aligned} E\xi &= \int_{-\infty}^{\infty} 4\sqrt{1-x^2} f_\eta(x)dx \\ &= 4\int_0^1 \sqrt{1-x^2}dx \\ &= \pi \end{aligned}$$

Metoda Monte-Carlo

Szacowanie liczby π

- Za pomocą komputerowego generatora liczb losowych generujemy n realizacji $\eta_1, \dots, \eta_n$ zmiennej losowej η
- Obliczamy $\xi_1, \dots, \xi_n$.
- Z Prawa Wielkich Liczb: $\bar{\xi}_n \approx \pi$

Nierówności

Nierówność Czebyszewa

$$(\forall \varepsilon > 0) \quad P\{|\bar{\xi}_n| \geq \varepsilon\} \leq \frac{E|\xi_1|}{n\varepsilon}$$

Nierówność Czebyszewa-Bienayme

$$(\forall \varepsilon > 0) \quad P\{|\bar{\xi}_n - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{n\varepsilon^2}$$

Metoda Monte-Carlo

Szacowanie liczby π

- Jak dużo symulacji, by z dużym prawdopodobieństwem γ błąd był nie większy niż zadane ε?
- Z nierówności Czebyszewa-Bienayme: $\frac{\sigma^2}{n\varepsilon^2} \leq 1-\gamma$
- Zatem $n \geq \frac{\sigma^2}{\varepsilon^2(1-\gamma)}$

3.3 Centralne Twierdzenie Graniczne

CTG

Centralne Twierdzenie Graniczne
Niech $\xi_1, \xi_2, \ldots$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie, o wartości średniej μ i wariancji $0 < \sigma^2 < \infty$. Wtedy

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \left| P\left\{ \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x \right\} - \Phi(x) \right| = 0$$

Zapis symboliczny

$$\frac{\bar{\xi}_n - \mu}{\sigma}\sqrt{n} \xrightarrow{D} N(0,1) \quad \text{lub} \quad (S_n)_{n=1}^{\infty} \sim AN(n\mu, n\sigma^2)$$

CTG

Twierdzenie Berry-Esséna'a
Jeżeli $\xi_1, \xi_2, \ldots$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie oraz $E|\xi_1|^3 < \infty$, to

$$\sup_{x\in\mathbb{R}} \left| P\left\{ \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x \right\} - \Phi(x) \right| \leq C\frac{E|\xi_1 - \mu|^3}{\sigma^3\sqrt{n}},$$

gdzie $1/\sqrt{2\pi} \leq C < 0.8$.

Rozkład dwumianowy

Twierdzenie de Moivre-Laplace'a
Niech $\zeta_n \sim B(n,p)$. Wtedy

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \left| P\left\{ \frac{\zeta_n - np}{\sqrt{np(1-p)}} \leq x \right\} - \Phi(x) \right| = 0$$

Zapis symboliczny

$$\zeta_n \sim AN(np, np(1-p))$$

Rozkład dwumianowy

Twierdzenie Berry-Esséna'a
Niech $\zeta_n \sim B(n,p)$. Wtedy

$$\sup_{x\in\mathbb{R}}\left|P\left\{\frac{S_n-np}{\sqrt{np(1-p)}}\le x\right\}-\Phi(x)\right|\le C\frac{p^2+(1-p)^2}{\sqrt{np(1-p)}}$$

Rozkład dwumianowy

Przybliżenie poissonowskie
Niech $\zeta_n \sim B(n,p)$. Wtedy dla każdego zbioru $M\subseteq\mathbb{N}$ mamy

$$\left|P(\zeta_n\in M)-\sum_{k\in M}\frac{(np)^k}{k!}e^{-np}\right|\le\frac{(np)^2}{n}$$

Rozkład dwumianowy

Przykład
Niech $\xi_1,\ldots,\xi_{1000}$ będą niezależnymi zmiennych losowymi o rozkładach $D(3/8)$. Obliczyć w przybliżeniu

$$P\left\{\sum_{i=1}^{1000}\xi_i\le 400\right\}$$

Rozkład dwumianowy

Rozwiązanie
Z twierdzeń asymptotycznych wynika, że

$$\sum_{i=1}^{1000}\xi_i\sim AN(1000\cdot 3/8, 1000\cdot(3/8)\cdot(5/8))$$

lub

$$\sum_{i=1}^{1000}\xi_i\sim Po(1000\cdot 3/8)$$

Rozkład dwumianowy

Błąd przybliżenia rozkładem normalnym
(w twierdzeniu Berry-Esséna przyjmujemy $C = 0.8$)

$$0.8\frac{(3/8)^2 + (5/8)^2}{\sqrt{1000 \cdot (3/8) \cdot (5/8)}} = 0.0276$$

Błąd przybliżenia rozkładem Poissona

$$\frac{(1000 \cdot (3/8))^2}{1000} = 140.7$$

Rozkład dwumianowy

Rozwiązanie

$$\sum_{i=1}^{1000} \xi_i \sim AN(1000 \cdot 3/8, 1000 \cdot (3/8) \cdot (5/8))$$

$$P\left\{\sum_{i=1}^{1000} \xi_i \leq 400\right\}$$

$$= P\left\{\frac{\sum_{i=1}^{1000} \xi_i - 1000 \cdot (3/8)}{\sqrt{1000 \cdot (3/8) \cdot (5/8)}} \leq \frac{400 - 1000 \cdot (3/8)}{\sqrt{1000 \cdot (3/8) \cdot (5/8)}}\right\}$$

$$\approx \Phi\left(1.633\right)$$

Dystrybuanta empiryczna

Definicja
Niech $\xi_1, \xi_2, \ldots$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie o dystrybuancie F. **Dystrybuantą empiryczną** nazywamy funkcję $F_n : \mathbb{R} \to [0, 1]$ określoną wzorem

$$F_n(x)\frac{\#\{1 \leq j \leq n : \xi_j \leq x\}}{n}$$

Dystrybuanta empiryczna

MPWL

$$(\forall x \in \mathbb{R}) \lim_{n\to\infty} F_n(x) = F(x)$$

CTG

$$(\forall x \in \mathbb{R})\, F_n(x) \sim AN\left(F(x), \frac{F(x)(1-F(x))}{n}\right)$$