Alicja Jokiel-Rokita
Ryszard Magiera

MODELE I METODY STATYSTYKI MATEMATYCZNEJ W ZADANIACH

Wydanie IV rozszerzone

Oficyna Wydawnicza GiS
Wrocław 2018

alicja.jokiel-rokita@pwr.edu.pl, ryszard.magiera@pwr.edu.pl

Autor projektu okładki *Ryszard Magiera.*

Printed in Poland

Skład komputerowy książki w systemie LaTeX wykonali autorzy. Do sporządzenia wykresów wykorzystano procedury pakietu komputerowego *Mathematica* 11 (licencja L4691-7788).

ISBN 978–83–62780–58–7

Wydanie IV, Wrocław 2018
Oficyna Wydawnicza GiS, s.c.
Druk i oprawa: Drukarnia I-BiS sp. z o.o., A.Bieroński, P.Bieroński s.j.

Przedmowa

Książka zawiera zadania i problemy dotyczące podstawowych wnioskowań statystyki matematycznej. Zadania spełniają swą praktyczną i uzupełniającą rolę w zdobywaniu doświadczenia w stosowaniu metod wnioskowania statystycznego. Zadania przedstawione w książce wykorzystywane były na kursach *Wstęp do statystyki matematycznej* i *Statystyka matematyczna* prowadzonych przez autorów tej książki w Instytucie Matematyki, obecnie Wydziale Matematyki Politechniki Wrocławskiej.

Zadania umieszczone są w pięciu rozdziałach, z których każdy składa się z dwóch części. Zgodnie z tematem każdego rozdziału, w pierwszej części przedstawiono podstawowe pojęcia, fakty i twierdzenia, a w drugiej części (*Zadania*) – odpowiedni zestaw zadań. Rozdziały 1 – 5 zawierają zadania dotyczące odpowiednio: (1) rozkładów prawdopodobieństwa najczęściej występujących w statystyce matematycznej (Rozdział 1 *Rozkłady*), (2) podstawowych własności statystyk służących do konstrukcji narzędzi wnioskowań statystycznych (Rozdział 2 *Statystyki i rodziny rozkładów prawdopodobieństwa*), (3) metod estymacji i optymalnych własności estymatorów (Rozdział 3 *Estymacja*), (4) konstrukcji testów parametrycznych i nieparametrycznych (Rozdział 4 *Testowanie hipotez*), (5) wyznaczania optymalnych funkcji decyzyjnych przy różnych funkcjach straty (Rozdział 5 *Statystyczne funkcje decyzyjne*).

Rozdział 6 *Rozwiązania, wskazówki, odpowiedzi* zawiera pełne rozwiązania, wskazówki lub odpowiedzi do 238 zadań. W wielu przypadkach pełne rozwiązania przedstawiają szczegółowo metody rozwiązania dla innych zadań. Wszystkich zadań jest 470.

W Rozdziale 7 podano tablice zawierające oznaczenia i podstawowe własności rozkładów rozpatrywanych w zadaniach.

Rozdział 8 zawiera tablice statystyczne podające wartości dystrybuanty i kwantyle rozkładu normalnego, wartości dystrybuanty rozkładu Kołmogorowa, kwantyle rozkładu statystyki Kołmogorowa i kwantyle rozkładu t Studenta. Są to najczęściej wykorzystywane wartości w zadaniach dotyczących weryfikacji hipotez i przedziałów ufności.

Książka może być wykorzystana na wszystkich matematycznych kierunkach studiów jako pomoc do ćwiczeń i uzupełnienie wykładu ze statystyki matematycznej. Mogą z niego również korzystać słuchacze studiów doktoranckich na studiach ekonomicznych i technicznych pragnący poszerzyć swą wiedzę dotyczącą podstawowych wnioskowań statystyki matematycznej. Pewna część zadań może być również wykorzystana na kursach statystyki dla studentów kierunków technicznych. Zadania te zostały oznaczone symbolem $^{\bullet}$. Według autorów, materiał potrzebny do ich rozwiązania nie wykracza poza ramy standardowego kursu statystyki. Własności rozkładów i związki między rozkładami będące treścią zadań z Rozdziału 1 mogą być również wykorzystane w algorytmach symulacji stochastycznych.

Niniejsze wydanie jest rozszerzeniem poprzedniego wydania III. Wydanie III książki zawierało 408 wszystkich zadań, a pełne rozwiązania, wskazówki lub odpowiedzi podane były do 225 zadań.

W wydaniu niniejszym zmieniono układ graficzny tekstu, poprawiono zauważone błędy redakcyjne i typograficzne.

Książkę można traktować jako uzupełnienie (o zadania i przykłady) dwóch tomów książki *Modele i metody statystyki matematycznej* (pozycje [22] i [23] spisu literatury) oraz książki *Statystyczne funkcje decyzyjne* [24]. Niniejsza książka wraz z wymienionymi powyżej zachowują ten sam matematyczny sposób prezentacji tematów związanych z wnioskowaniem statystycznym oraz takie same oznaczenia.

Wrocław, wrzesień 2018 *Alicja Jokiel-Rokita* *Ryszard Magiera*

Spis treści

1

Rozkłady

1.1 Rozkład i dystrybuanta zmiennej losowej

DEFINICJA **1.1** Niech $\mathcal{P}$ będzie miarą probabilistyczną na przestrzeni $(\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k})$. *Dystrybuantą miary* $\mathcal{P}$ nazywamy funkcję $F : \mathbb{R}^k \mapsto [0, 1]$ określoną wzorem

$$F(x_1, \ldots, x_k) = \mathcal{P}\big((-\infty, x_1] \times \cdots \times (-\infty, x_k]\big), \quad x_i \in \mathbb{R}. \tag{1.1}$$

□

DEFINICJA **1.2** Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną i niech $X : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k})$ będzie zmienną losową. *Rozkładem zmiennej losowej* X nazywamy miarę indukowaną przez X, tzn. miarę P^X na $(\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k})$ określoną wzorem

$$P^X(B) = P(\{\omega \in \Omega : X(\omega) \in B\}) = P(X^{-1}(B)), \quad B \in \mathcal{B}_{\mathbb{R}^k}.$$

Dystrybuanta miary P^X, określona wzorem (1.1), nazywa się *dystrybuantą rozkładu zmiennej losowej* X i oznaczana jest przez F_X. □

1.2 Rozkład funkcji wektora losowego

Niech $\mathbf{X} = (X_1, \ldots, X_n)$ będzie wektorem losowym typu ciągłego o gęstości $f_{\mathbf{X}}(\mathbf{x})$, a A niech będzie otwartym podzbiorem $\mathbb{R}^n$ takim, że $P(\mathbf{X} \in A) = 1$. Załóżmy, że funkcje $g_i : \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, \ldots, n$, spełniają następujące warunki:

(i) przekształcenie $\mathbf{g} = (g_1, \ldots, g_n)$ jest wzajemnie jednoznaczne na zbiorze A;

(ii) przekształcenie $\mathbf{g}$ ma na zbiorze A ciągłe pierwsze pochodne cząstkowe względem wszystkich zmiennych;

(iii) jakobian $J_{\mathbf{g}}(\mathbf{x}) = \dfrac{\partial(g_1, \ldots, g_n)}{\partial(x_1, \ldots, x_n)}$ jest różny od zera na zbiorze A.

Oznaczmy przez $\mathbf{h} = (h_1, \ldots, h_n)$ przekształcenie odwrotne do $\mathbf{g}$, tzn. $\mathbf{x} = \mathbf{h}(\mathbf{y}) = (h_1(\mathbf{y}), \ldots, h_n(\mathbf{y}))$ dla każdego $\mathbf{y} \in \mathbf{g}(A)$. Dla każdego zbioru mierzalnego B zachodzi wzór

$$P(\mathbf{X} \in B) = \int_B f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = \int_{\mathbf{g}(B)} f_{\mathbf{X}}(\mathbf{h}(\mathbf{y})|J_{\mathbf{h}}(\mathbf{y})|d\mathbf{y} = P(\mathbf{Y} \in \mathbf{g}(B)),$$

gdzie $J_{\mathbf{h}}(\mathbf{y}) = \dfrac{\partial(h_1, \ldots, h_n)}{\partial(y_1, \ldots, y_n)}$ jest jakobianem przekształcenia odwrotnego.

TWIERDZENIE **1.1** Jeżeli wektor $\mathbf{X}$ ma gęstość $f_{\mathbf{X}}$, a przekształcenie $\mathbf{g} : \mathbb{R}^n \mapsto \mathbb{R}^n$ spełnia warunki (i) – (iii), to wektor $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ ma gęstość

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(h_1(\mathbf{y}), \ldots, h_n(\mathbf{y}))|J_{\mathbf{h}}(\mathbf{y})|.$$

⊡

1.3 Funkcja charakterystyczna zmiennej losowej

DEFINICJA **1.3** *Funkcją charakterystyczną* zmiennej losowej X o rozkładzie określonym przez dystrybuantę $F(x) = P(X \leqslant x)$ nazywamy funkcję $\phi : \mathbb{R} \mapsto \mathbb{C}$ określoną wzorem

$$\begin{aligned}\phi(t) &= E\left(\mathrm{e}^{\iota tX}\right) = \int_{-\infty}^{\infty} \mathrm{e}^{\iota tx} dF(x) \\ &= \begin{cases} \displaystyle\int_{-\infty}^{\infty} \mathrm{e}^{\iota tx} f(x)dx, & \text{gdy } X \text{ jest typu ciągłego o gęstości } f(x), \\ \displaystyle\sum_k \mathrm{e}^{\iota tx_k} p_k, & \text{gdy } X \text{ jest typu dyskretnego, } p_k = P(X = x_k) \end{cases}\end{aligned}$$

(ι oznacza jednostkę urojoną, tzn. $\iota = \sqrt{-1}$). □

Jeżeli $E(|X|^n) < \infty$ dla pewnego $n \in \mathbb{N}$, to dla każdego $k = 0, 1, \ldots, n$,

$$E(X^k) = (-\iota)^k \left.\frac{d^k\phi(t)}{dt^k}\right|_{t=0}.$$

Niech $\mathbf{X} = (X_1, \ldots, X_k)^T$ będzie wektorem losowym o dystrybuancie $F(\mathbf{x}) = P(X_1 \leqslant x_1, \ldots, X_k \leqslant x_k)$, $\mathbf{x} = (x_1, \ldots, x_k)^T \in \mathbb{R}^k$. Funkcja charakterystyczna wektora losowego $\mathbf{X}$ określona jest wzorem

$$\phi(\mathbf{t}) = E\left(\mathrm{e}^{\iota \mathbf{t}^T \mathbf{X}}\right) = \int_{\mathbf{R}^k} \mathrm{e}^{\iota \sum_{j=1}^{k} t_j X_j} dF(\mathbf{x}), \quad \mathbf{t} = (t_1, \ldots, t_k)^T \in \mathbb{R}^k.$$

Jeżeli istnieją tzw. *mieszane momenty* $E\left(X_1^{j_1} X_2^{j_2} \cdots X_k^{j_k}\right)$ wektora losowego $\mathbf{X}$, to zachodzi wzór

$$E\left(X_1^{j_1} X_2^{j_2} \cdots X_k^{j_k}\right) = (-\iota)^j \left. \frac{\partial^j \phi(\mathbf{t})}{\partial t_1^{j_1} \partial t_2^{j_2} \ldots \partial t_k^{j_k}} \right|_{\mathbf{t}=\mathbf{0}},$$

gdzie $j_1, j_2, \ldots, j_k \in \overline{\mathbb{N}}; j = j_1 + j_2 + \cdots + j_k$.

Niech $\mathbf{X}$ i $\mathbf{Y}$ będą k-wymiarowymi niezależnymi wektorami losowymi. Wtedy

$$\phi_{\mathbf{X}+\mathbf{Y}}(\mathbf{t}) = \phi_{\mathbf{X}}(\mathbf{t}) \phi_{\mathbf{Y}}(\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^k.$$

Niech $\mathbf{Y} = \mathbf{C}^T \mathbf{X} + \mathbf{c}$, gdzie $\mathbf{C}$ jest macierzą wymiaru $k \times m$, a $\mathbf{c} \in \mathbf{R}^m$. Wtedy

$$\phi_{\mathbf{Y}}(\mathbf{u}) = \mathrm{e}^{\iota \mathbf{u}^T \mathbf{c}} \phi_{\mathbf{X}}(\mathbf{C}\mathbf{u}), \quad \mathbf{u} \in \mathbb{R}^m.$$

1.4 Funkcja tworząca momenty zmiennej losowej

DEFINICJA **1.4** *Funkcją tworzącą momenty* zmiennej losowej X o rozkładzie określonym przez dystrybuantę $F(x) = P(X \leqslant x)$ nazywamy funkcję

$$\begin{aligned} \psi(t) &= E\left(\mathrm{e}^{tX}\right) = \int_{-\infty}^{\infty} \mathrm{e}^{tx} dF(x) \\ &= \begin{cases} \displaystyle\int_{-\infty}^{\infty} \mathrm{e}^{tx} f(x) dx, & \text{gdy } X \text{ jest typu ciągłego o gęstości } f(x), \\ \displaystyle\sum_k \mathrm{e}^{t x_k} p_k, & \text{gdy } X \text{ jest typu dyskretnego, } p_k = P(X = x_k). \end{cases} \end{aligned}$$

□

Funkcja tworząca momenty jest nieujemna, ale może być równa ∞ wszędzie z wyjątkiem $t = 0$. Jeżeli $\psi(t) < \infty$ dla $-\varepsilon \leqslant t \leqslant \varepsilon$, to

$$E(X^k) = \left. \frac{d^k \psi(t)}{dt^k} \right|_{t=0}.$$

1.5 Wielowymiarowy rozkład normalny

DEFINICJA **1.5** Wektor losowy $\mathbf{X} = (X_1, \ldots, X_p)^T$ typu ciągłego ma *p-wymiarowy rozkład normalny*, jeżeli można go przedstawić w postaci $\mathbf{X} = \boldsymbol{\mu} + \mathbf{AZ}$, gdzie $\boldsymbol{\mu}$ jest p-wymiarowym wektorem, $\mathbf{A}$ jest macierzą wymiaru $p \times m$ rzędu m oraz $\mathbf{Z} = (Z_1, \ldots, Z_m)^T$ jest wektorem m niezależnych (jednowymiarowych) zmiennych losowych takich, że $Z_i \sim \mathcal{N}(0, 1)$, $i = 1, \ldots, m$. □

Reprezentacja $\mathbf{X} = \boldsymbol{\mu} + \mathbf{AZ}$ nie jest wyznaczona jednoznacznie. Jeżeli $\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}_1\mathbf{Z}_1$ i $\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}_2\mathbf{Z}_2$, to $\mathbf{A}_1\mathbf{A}_1^T = \mathbf{A}_2\mathbf{A}_2^T$.

Często p-wymiarowy rozkład normalny definiuje się w oparciu o następujące kryterium.

TWIERDZENIE **1.2** (*Craméra-Wolda*) Wektor losowy $\mathbf{X} = (X_1, \ldots, X_p)^T$ ma p-wymiarowy rozkład normalny wtedy i tylko wtedy, gdy dla każdego $\mathbf{u} \in \mathbb{R}^p$ zmienna losowa $Y = \mathbf{u}^T\mathbf{X}$ ma jednowymiarowy rozkład normalny. ▣

Jeżeli wektor losowy $\mathbf{X}$ ma p-wymiarowy rozkład normalny w sensie Twierdzenia 1.2, to prawdziwe są następujące twierdzenia.

TWIERDZENIE **1.3** Wartość oczekiwana $E(\mathbf{X})$ oraz macierz kowariancji $\mathrm{Var}(\mathbf{X})$ istnieją. Ponadto $E(\mathbf{X}) = \boldsymbol{\mu}$, $\mathrm{Var}(\mathbf{X}) = \mathbf{AA}^T$. ▣

Macierz kowariancji $\mathrm{Var}(\mathbf{X}) = \mathbf{AA}^T$ oznaczamy przez $\boldsymbol{\Sigma}$.

TWIERDZENIE **1.4** Funkcja charakterystyczna wektora losowego $\mathbf{X}$ o p-wymiarowym rozkładzie normalnym jest postaci

$$\phi(\mathbf{t}) = \exp\left[\imath\mathbf{t}^T\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}\right], \quad \mathbf{t} = (t_1, \ldots, t_p)^T. \tag{1.2}$$

Na odwrót, jeżeli $\boldsymbol{\Sigma}$ jest nieujemnie określoną macierzą symetryczną wymiaru $p \times p$, to funkcja postaci (1.2), gdzie $\boldsymbol{\mu} \in \mathbb{R}^p$, jest funkcją charakterystyczną wektora losowego o p-wymiarowym rozkładzie normalnym. ▣

Rozkład normalny p-wymiarowy jest całkowicie określony przez wektor średnich $\boldsymbol{\mu}$ oraz macierz kowariancji $\boldsymbol{\Sigma}$, gdyż funkcja charakterystyczna zależy wyłącznie od tych dwóch wielkości. Rozkład normalny p-wymiarowy z wektorem wartości oczekiwanych $\boldsymbol{\mu}$ i macierzą kowariancji $\boldsymbol{\Sigma}$ oznaczamy symbolem $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Wielowymiarowego rozkładu normalnego nie definiuje się przez podanie postaci gęstości rozkładu, tak jak zwykle w przypadku innych rozkładów, ponieważ nie zawsze istnieje gęstość względem miary Lebesgue'a p-wymiarowego rozkładu normalnego w przestrzeni $\mathbb{R}^p$. Zachodzą następujące dwa przypadki:

(1) $\mathrm{rz}(\boldsymbol{\Sigma}) = p$ ($|\boldsymbol{\Sigma}| := \det(\boldsymbol{\Sigma}) \neq 0$). Wówczas rozkład wektora losowego $\mathbf{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ma gęstość względem miary Lebesgue'a w przestrzeni $\mathbb{R}^p$.

Funkcja gęstości tego rozkładu wyraża się wzorem

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\mathbf{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right];$$

w tym przypadku mówimy, że $\mathbf{X}$ ma nieosobliwy rozkład normalny;

(2) $\mathrm{rz}(\mathbf{\Sigma}) < p$. Ponieważ macierz $\mathbf{\Sigma}$ jest osobliwa, więc składowe $X_1, \ldots, X_p$ muszą być liniowo zależne. Dlatego wektor $(X_1, \ldots, X_p)^T$ nie może mieć łącznej p-wymiarowej gęstości. Istnieje jednak gęstość na pewnej podprzestrzeni. Mianowicie, jeżeli niektóre ze składowych wyeliminujemy tak, że pozostałe będą liniowo niezależne, to te pozostałe składowe będą miały nieosobliwy wielowymiarowy rozkład normalny.

TWIERDZENIE **1.5** Jeżeli wektor losowy $\mathbf{X} = (X_1, \ldots, X_p)^T$ ma p-wymiarowy rozkład normalny $\mathcal{N}_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, to $X_1, \ldots, X_p$ są niezależnymi zmiennymi losowymi wtedy i tylko wtedy, gdy $\mathbf{\Sigma}$ jest macierzą diagonalną. ⊡

TWIERDZENIE **1.6** Jeżeli wektor losowy $\mathbf{X}$ ma p-wymiarowy rozkład normalny $\mathcal{N}_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, to dla każdego wektora $\mathbf{a} \in \mathbb{R}^k$ i każdej stałej macierzy $\mathbf{A}$ wymiaru $k \times p$ wektor losowy $\mathbf{Y} = \mathbf{AX} + \mathbf{a}$ ma k-wymiarowy rozkład normalny $\mathcal{N}_k(\mathbf{A}\boldsymbol{\mu} + \mathbf{a}, \mathbf{A\Sigma A}^T)$. ⊡

WNIOSEK **1.1** Rozkłady brzegowe wektora losowego $\mathbf{X}$ o p-wymiarowym rozkładzie normalnym są rozkładami normalnymi w odpowiednich podprzestrzeniach przestrzeni $\mathbb{R}^p$. □

Następujące twierdzenie pokazuje, że ortogonalne przekształcenie liniowe zachowuje niezależność współrzędnych wektora losowego o rozkładzie normalnym, przy założeniu, że współrzędne mają jednakowe wariancje.

TWIERDZENIE **1.7** Jeżeli $\mathbf{X} = (X_1, \ldots, X_n)^T$ jest wektorem losowym o p-wymiarowym rozkładzie normalnym $\mathcal{N}_p(\boldsymbol{\mu}, \sigma^2\mathbf{I}_p)$, a $\mathbf{C}$ jest macierzą ortogonalną rozmiaru $p \times p$, to wektor losowy $\mathbf{Y} = \mathbf{CX}$ ma rozkład normalny $\mathcal{N}_p(\mathbf{C}\boldsymbol{\mu}, \sigma^2\mathbf{I}_p)$. ⊡

Wielowymiarowy rozkład normalny ma następującą własność addytywności.

TWIERDZENIE **1.8** Jeżeli $\mathbf{X}_i$ są niezależnymi wektorami losowymi $\mathbf{X}_i \sim \mathcal{N}_p(\boldsymbol{\mu}_i, \mathbf{\Sigma}_i)$, $i = 1, \ldots, n$, oraz $a_1, \ldots, a_n \in \mathbb{R}$, to wektor losowy

$$\mathbf{Y} = a_1\mathbf{X}_1 + \ldots + a_n\mathbf{X}_n$$

ma p-wymiarowy rozkład normalny $\mathcal{N}_p(\sum_{i=1}^n a_i\boldsymbol{\mu}_i, \sum_{i=1}^n a_i^2\mathbf{\Sigma}_i)$. ⊡

Jeżeli wektor losowy $\mathbf{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ podzielimy na dwa podwektory: pierwszy $\mathbf{X}^{(1)}$ zawierający p_1 pierwszych współrzędnych i drugi – $\mathbf{X}^{(2)}$ – zawierający kolejne $p_2 = p - p_1$ współrzędnych oraz stosując analogiczny podział wektora $\boldsymbol{\mu}$ oraz macierzy kowariancji $\boldsymbol{\Sigma}$ postaci

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}^{(11)} & \boldsymbol{\Sigma}^{(12)} \\ \boldsymbol{\Sigma}^{(21)} & \boldsymbol{\Sigma}^{(22)} \end{pmatrix},$$

to otrzymujemy następujące twierdzenia dotyczące rozkładów brzegowych wektorów $\mathbf{X}^{(1)}$ i $\mathbf{X}^{(2)}$ oraz rozkładów warunkowych $\mathbf{X}^{(1)}|\mathbf{X}^{(2)} = \mathbf{x}^{(2)}$ i $\mathbf{X}^{(2)}|\mathbf{X}^{(1)} = \mathbf{x}^{(1)}$.

TWIERDZENIE **1.9** Wektor losowy $\mathbf{X}^{(i)}$ ma p_i-wymiarowy rozkład normalny $\mathcal{N}_{p_i}(\boldsymbol{\mu}^{(i)}, \boldsymbol{\Sigma}^{(ii)})$, $i = 1, 2$. ◻

TWIERDZENIE **1.10** Wektory losowe $\mathbf{X}^{(1)}$ i $\mathbf{X}^{(2)}$ są niezależne wtedy i tylko wtedy, gdy $\boldsymbol{\Sigma}^{(12)} = \mathbf{0}$. ◻

TWIERDZENIE **1.11** Jeżeli $\boldsymbol{\Sigma}^{(22)}$ jest macierzą nieosobliwą, to wektor losowy $\mathbf{X}^{(1)}|\mathbf{X}^{(2)} = \mathbf{x}^{(2)}$ ma p_1-wymiarowy rozkład normalny $\mathcal{N}_{p_1}(\boldsymbol{\mu}_{1\cdot 2}, \boldsymbol{\Sigma}_{11\cdot 2})$, gdzie

$$\begin{aligned} \boldsymbol{\mu}_{1\cdot 2} &= \boldsymbol{\mu}^{(1)} + \boldsymbol{\Sigma}^{(12)}(\boldsymbol{\Sigma}^{(22)})^{-1}(\mathbf{x}^{(2)} - \boldsymbol{\mu}^{(2)}), \\ \boldsymbol{\Sigma}_{11\cdot 2} &= \boldsymbol{\Sigma}^{(11)} - \boldsymbol{\Sigma}^{(12)}(\boldsymbol{\Sigma}^{(22)})^{-1}\boldsymbol{\Sigma}^{(21)}. \end{aligned}$$

◻

Analogiczne twierdzenie zachodzi dla rozkładu $\mathbf{X}^{(2)}|\mathbf{X}^{(1)} = \mathbf{x}^{(1)}$.

Zadania

Oznaczenia dotyczące rozkładów podane są w Tablicach 7.1 i 7.2
na stronach 266 – 273

1.1 • Niech $X_1, \dots, X_k$ będą niezależnymi zmiennymi losowymi i niech $Y = \sum_{i=1}^{k} X_i$. Udowodnić następujące stwierdzenia.

(a) Jeżeli X_i ma rozkład dwumianowy $\mathcal{B}(n_i, p)$, $i = 1, \dots, k$, to Y ma rozkład dwumianowy $\mathcal{B}(\sum_{i=1}^{k} n_i, p)$.

(b) Jeżeli X_i ma rozkład Poissona $\mathcal{P}(\lambda_i)$, $i = 1, \dots, k$, to Y ma rozkład Poissona $\mathcal{P}(\sum_{i=1}^{k} \lambda_i)$.

(c) Jeżeli X_i ma rozkład dwumianowy ujemny $\mathcal{NB}(r_i, p)$, $i = 1, \dots, k$, to Y ma rozkład dwumianowy ujemny $\mathcal{NB}(\sum_{i=1}^{k} r_i, p)$.

(d) Jeżeli X_i ma rozkład wykładniczy $\mathcal{E}(\lambda)$, $i = 1, \dots, k$, to Y ma rozkład gamma $\mathcal{G}(k, \lambda)$.

(e) Jeżeli X_i ma rozkład Cauchy'ego $\mathcal{C}(\alpha_i, \lambda_i)$, $i = 1, \dots, k$, to Y ma rozkład $\mathcal{C}(\sum_{i=1}^{k} \alpha_i, \sum_{i=1}^{k} \lambda_i)$.

1.2 • Udowodnić, że jeżeli zmienne losowe $X_1, \dots, X_n$ są niezależne o jednakowym rozkładzie wykładniczym $\mathcal{E}(\lambda)$, to zmienna losowa

$$T(X_1, \dots, X_n) = \frac{2}{\lambda} \sum_{i=1}^{n} X_i$$

ma rozkład $\chi^2(2n)$.

1.3 Pokazać, że jeżeli $\mathbf{X}_1, \dots, \mathbf{X}_k$ są niezależnymi wektorami losowymi o rozkładach wielomianowych, odpowiednio $\mathcal{M}_r(n_1, \mathbf{p}), \dots, \mathcal{M}_r(n_k, \mathbf{p})$, gdzie $\mathbf{p}= (p_1, \dots, p_r)$, to wektor $\mathbf{Y} = \sum_{i=1}^{k} \mathbf{X}_i$ ma rozkład wielomianowy $\mathcal{M}_r(n, \mathbf{p})$, gdzie $n = \sum_{i=1}^{k} n_i$.

1.4 Niech $\mathbf{X} = (X_1, \ldots, X_r)$ będzie wektorem losowym o rozkładzie wielomianowym $\mathcal{M}_r(n, \mathbf{p})$, $\mathbf{p} = (p_1, \ldots, p_r)$. Pokazać, że dla każdego $i = 1, \ldots, r$, rozkładem brzegowym zmiennej X_i jest rozkład dwumianowy $\mathcal{B}(n, p_i)$.

1.5 • Niech F_X będzie dystrybuantą zmiennej losowej X, a f_X jej gęstością. Wyznaczyć dystrybuanty i gęstości następujących zmiennych losowych:

(a) $aX + b$, $a \neq 0$;

(b) $|X|$;

(c) X^2;

(d) $\sqrt{X}$, przy założeniu, że $P(X \geqslant 0) = 1$;

(e) $\sin(X)$.

1.6 • Niech X będzie zmienną losową o rozkładzie z ciągłą i ściśle rosnącą dystrybuantą F. Pokazać, że zmienna losowa $Y = F(X)$ ma rozkład jednostajny $\mathcal{U}(0,1)$.

1.7 • Niech U będzie zmienną losową o rozkładzie jednostajnym $\mathcal{U}(0,1)$ i niech F będzie dystrybuantą pewnego rozkładu. Oznaczmy

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geqslant t\}, \quad 0 < t < 1.$$

Pokazać, że zmienna losowa $Y = F^{-1}(U)$ ma rozkład o dystrybuancie F.

1.8 • Niech U oznacza zmienną losową o rozkładzie jednostajnym $\mathcal{U}(0,1)$. Pokazać, że zmienne losowe

$$Y = -\vartheta \ln(1-U), \quad Z = -\vartheta \ln(U), \ \vartheta > 0,$$

mają rozkład wykładniczy $\mathcal{E}(\vartheta)$.

1.9 • Wykazać, że jeżeli zmienna losowa X ma rozkład t Studenta $\mathcal{T}(n)$, to zmienna losowa $Y = X^2$ ma rozkład F Snedecora $\mathcal{F}(1, n)$.

1.10 Wykazać, że jeżeli zmienna losowa X ma rozkład F Snedecora $\mathcal{F}(m, n)$, to zmienna losowa

$$Y = \frac{mX}{n + mX}$$

ma rozkład beta $\mathcal{B}e(\alpha, \beta)$ o parametrach $\alpha = m/2$, $\beta = n/2$.

1.11 Wykazać, że jeżeli zmienna losowa X ma rozkład $\mathcal{U}(-1/2\pi,\ 1/2\pi)$, to zmienna losowa $Y = \mathrm{tg}(X)$ ma rozkład Cauchy'ego $\mathcal{C}(0,1)$.

1.12 Wykazać, że jeżeli zmienne losowe X i Y są niezależne o rozkładach odpowiednio $\mathcal{G}(1, \lambda_1)$ i $\mathcal{G}(1, \lambda_2)$, to zmienna losowa $X/(X+Y)$ ma rozkład beta $\mathcal{B}e(\lambda_1, \lambda_2)$.

1.13 Wykazać, że jeżeli zmienna losowa X ma rozkład t Studenta $\mathcal{T}(n)$, to zmienna losowa $Y = (1 + X^2/n)^{-1}$ ma rozkład beta $\mathcal{B}e(1/2n, 1/2)$.

1.14 Wykazać, że jeżeli zmienna losowa X ma rozkład Pareto $\mathcal{P}a(x_0, \alpha)$, to zmienna losowa $1/X$ ma rozkład potęgowy $\mathcal{P}o(1/x_0, \alpha)$, a zmienna losowa $\ln(X/x_0)$ ma rozkład wykładniczy $\mathcal{E}(1/\alpha)$.

1.15 Wykazać, że jeżeli zmienna losowa U ma rozkład jednostajny $\mathcal{U}(0, 1)$, to zmienna losowa $X = x_0 U^{-1/\alpha}$ ma rozkład Pareto $\mathcal{P}a(x_0, \alpha)$.

1.16 Wykazać, że jeżeli X i Y są niezależnymi zmiennymi losowymi o rozkładzie normalnym $\mathcal{N}(0, 1)$, to zmienna losowa X/Y ma rozkład Cauchy'ego $\mathcal{C}(0, 1)$.

1.17 • Wykazać, że jeżeli zmienna losowa X ma rozkład normalny $\mathcal{N}(0, 1)$, zmienna losowa Y ma rozkład $\chi^2(n)$ i zmienne te są niezależne, to zmienna losowa

$$\frac{X}{\sqrt{Y/n}}$$

ma rozkład t Studenta $\mathcal{T}(n)$.

1.18 • Wykazać, że jeżeli X i Y są niezależnymi zmiennymi losowymi o rozkładzie jednostajnym $\mathcal{U}(a/2, b/2)$, to zmienna losowa $Z = X + Y$ ma rozkład trójkątny na przedziale (a, b) o gęstości

$$f(x) = \frac{2}{b-a} - \frac{2}{(b-a)^2}|a + b - 2x|, \quad x \in (a, b).$$

1.19 • Wykazać, że jeżeli zmienna losowa X ma rozkład wykładniczy $\mathcal{E}(\lambda)$, to zmienna losowa $Y = X^{1/\alpha}, \alpha > 0,$ ma rozkład Weibulla $\mathcal{W}e(\alpha, \lambda^{1/\alpha})$.

1.20 • Wykazać, że jeżeli zmienna losowa X ma rozkład normalny $\mathcal{N}(\mu, \sigma^2)$, to zmienna losowa $Y = \exp(X)$ ma rozkład logarytmiczno-normalny $\mathcal{LN}(\mu, \sigma^2)$.

1.21 Pokazać, że jeżeli X ma rozkład t Studenta trójparametrowy $\mathcal{T}_1(\nu, \mu, \sigma^2)$, to $(X - \mu)^2/\sigma^2$ ma rozkład F Fishera $\mathcal{F}(1, \nu)$.

1.22 Udowodnić, że jeżeli X i Y są niezależnymi zmiennymi losowymi o rozkładzie jednostajnym $\mathcal{U}(0, 1)$, to

$$U = \sqrt{-2\ln(X)}\cos(2\pi Y), \quad V = \sqrt{-2\ln(X)}\sin(2\pi Y)$$

są niezależnymi zmiennymi losowymi o rozkładzie normalnym $\mathcal{N}(0,1)$.

1.23 • Niech $X_1, \ldots, X_n$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie o dystrybuancie F. Wyznaczyć rozkład zmiennych losowych $X_{1:n} = \min\{X_1, \ldots, X_n\}$, $X_{n:n} = \max\{X_1, \ldots, X_n\}$. Podać postać funkcji gęstości rozkładu zmiennych $X_{1:n}$ i $X_{n:n}$ w przypadku, gdy F jest dystrybuantą rozkładu wykładniczego.

1.24 Niech X_1, X_2 będą niezależnymi zmiennymi losowymi o rozkładzie normalnym $\mathcal{N}(0,1)$. Wyznaczyć gęstość rozkładu wektora losowego (Y_1, Y_2), gdzie $Y_1 = \sqrt{X_1^2 + X_2^2}$ i $Y_2 = X_1/X_2$. Czy Y_1 i Y_2 są niezależne?

1.25 • Niech $\mathbf{X} = (X_1, X_2)$ będzie wektorem losowym typu ciągłego z gęstością $f_{\mathbf{X}}$. Wyznaczyć gęstość wektora losowego $\mathbf{Y} = (Y_1, Y_2)$, gdzie $Y_1 = X_1 X_2$ i $Y_2 = X_1/X_2$ oraz gęstości brzegowe zmiennych losowych Y_1 i Y_2.

1.26 • Niech X_1 i X_2 będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie wykładniczym $\mathcal{E}(\lambda)$. Niech $Y = X_1 - X_2$ i $Z = X_2$.

(a) Wyznaczyć gęstość łącznego rozkładu wektora losowego (Y, Z).

(b) Wykazać, że zmienna losowa Y ma rozkład Laplace'a $\mathcal{L}a(0,\lambda)$.

1.27 Niech Y będzie zmienną losową o rozkładzie gamma $\mathcal{G}(r, 1/a)$. Niech X oznacza zmienną losową, której rozkład warunkowy przy ustalonej wartości $Y = \lambda$ jest rozkładem Poissona $\mathcal{P}(\lambda)$. Wykazać, że dla $a = (1-p)/p$ rozkład bezwarunkowy zmiennej losowej X jest ujemnym rozkładem dwumianowym $\mathcal{NB}(r,p)$.

1.28 Niech X i Y będą niezależnymi zmiennymi losowymi o rozkładach dwumianowych odpowiednio $\mathcal{B}(n,p)$ i $\mathcal{B}(r,p)$. Wyznaczyć warunkowy rozkład zmiennej losowej X pod warunkiem $X + Y = t$.

1.29 Niech $\mathbf{X} = (X_1, \ldots, X_n)$ będzie próbą z rozkładu Poissona $\mathcal{P}(\lambda)$ i niech $T = X_1 + \ldots + X_n$. Wyznaczyć warunkowy rozkład wektora losowego $\mathbf{X}$ pod warunkiem $T = t$.

1.30 Niech X_1 i X_2 będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie wykładniczym $\mathcal{E}(\lambda)$. Wyznaczyć warunkową wartość oczekiwaną

$$E(X_1|X_1 + X_2).$$

1.31 Niech X będzie zmienną losową taką, że $E(X^2) < \infty$ i niech $Y = |X|$. Załóżmy, że funkcja gęstości rozkładu zmiennej losowej X jest symetryczna względem 0. Pokazać, że X i Y są nieskorelowane, ale nie są niezależne.

1.32 Zmienna losowa X typu ciągłego ma rozkład beta 1 rodzaju na przedziale (a, b) z parametrami α, β ($\alpha > 0, \beta > 0$), który oznaczamy przez $\mathcal{B}e^{(1)}(\alpha, \beta; a, b)$, jeżeli jej gęstość wyraża się wzorem

$$f(x) = \frac{(b-a)^{1-\alpha-\beta}}{B(\alpha, \beta)}(x-a)^{\alpha-1}(b-x)^{\beta-1}\mathbf{1}_{(a,b)}(x),$$

gdzie $B(\alpha, \beta)$ oznacza funkcję beta. Pokazać, że rozpatrywany zwykle rozkład beta 1 rodzaju $\mathcal{B}e^{(1)}(\alpha, \beta; 0, 1)$ na odcinku $(0, 1)$ można otrzymać z rozkładu $\mathcal{B}e^{(1)}(\alpha, \beta; a, b)$ w wyniku liniowego przekształcenia $X \to (X-a)/(b-a)$.

1.33 • Zaproponować algorytm generowania liczb losowych z rozkładu

(a) dwumianowego $\mathcal{B}(n, p)$;

(b) Poissona $\mathcal{P}(\lambda)$.

1.34 • Zaproponować algorytm generowania liczb losowych z rozkładu

(a) wykładniczego $\mathcal{E}(\lambda)$;

(b) Weibulla $\mathcal{W}e(\alpha, \beta)$;

(c) beta $\mathcal{B}e(\alpha, \beta)$;

(d) Pareto $\mathcal{P}a(x_0, \alpha)$;

(e) normalnego $\mathcal{N}(\mu, \sigma^2)$.

1.35 • Korzystając z prawa wielkich liczb i przeprowadzając odpowiednie symulacje, oszacować wartość oczekiwaną zmiennej losowej X o rozkładzie

(a) Weibulla $\mathcal{W}e(2, 5)$;

(b) beta $\mathcal{B}e(3, 4)$;

(c) Pareto $\mathcal{P}a(1, 2)$.

Porównać otrzymane oszacowania z dokładnymi wartościami $E(X)$.

1.36 • Niech $U_1, U_2, \ldots$ będą niezależnymi zmiennymi losowymi o rozkładzie jednostajnym $\mathcal{U}(0, 1)$ i niech zmienna losowa N będzie określona następująco

$$N = \min\left\{n : \sum_{i=1}^{n} U_i > 1\right\}.$$

Oszacować wartość $E(N)$ przez wygenerowanie 100, 1000 i 10000 wartości N.

1.37 • Niech U będzie zmienną losową o rozkładzie jednostajnym $\mathcal{U}(0, 1)$. Przeprowadzając odpowiednie symulacje oszacować

(a) $\mathrm{Cov}(U, \exp(U))$;

(b) $\mathrm{Cov}(U^2, \sqrt{1-U^2})$.

W przypadku (a) porównać otrzymane oszacowanie z dokładną wartością.