

UWAGI O ROZKŁADZIE FUNKCJI ZMIENNEJ LOSOWEJ.

X - dana zmienna losowa ciągła o gęstości f .

$Y = g(X)$ g - borelowska, tzn. $g^{-1}(B) \in \mathcal{B}(\mathbb{R})$ dla $B \in \mathcal{B}(\mathbb{R})$,

Wyznaczyć gęstość $g(y)$ zmiennej losowej Y .

1) Jeśli g - ściśle monotoniczna i różniczkowalna w przedziale (a, b) koncentracji X to:

$$g(y) = f(h(y)) |h'(y)|$$

gdzie $h = g^{-1}$.

Należy pamiętać o przekształceniu przedziału koncentracji.

Przykład.

$$Y = aX + b, \text{ wtedy } g(y) = f\left(\frac{y-b}{a}\right) \frac{1}{|a|},$$

Przykład.

$$\text{Jeśli } X \text{ ma rozkład o gęstości } f(x) = \begin{cases} 0 & \text{dla } x \leq 0 \\ e^{-x} & \text{dla } x > 0 \end{cases} \quad Y = \sqrt{X} - 2,$$

$$\text{wtedy } h(y) = (y+2)^2, \quad h'(y) = 2(y+2), \quad g(0) = -2, \quad g(\infty) = \infty,$$

$$g(y) = \begin{cases} 0 & \text{dla } x \leq -2 \\ 2(y+2)e^{-(y+2)^2} & \text{dla } x > -2 \end{cases},$$

2) Jeśli g - przedziałami ściśle monotoniczna i różniczkowalna w przedziale (a, b) koncentracji X to:

$$g(y) = \sum_{i=1}^k f(h_i(y)) |h'_i(y)|$$

gdzie h_i - funkcje odwrotne do g dla poszczególnych przedziałów,

k - liczba wartości funkcji odwrotnej odpowiadających danemu y .

Przykład.

$$Y = |X|, \text{ wtedy } g(y) = f(-y) + f(y) \quad y > 0,$$

Przykład.

$$Y = X^2, \text{ wtedy } g(y) = f(-\sqrt{y}) \frac{1}{2\sqrt{y}} + f(\sqrt{y}) \frac{1}{2\sqrt{y}} \quad y > 0,$$

Jeśli X - skokowa, o funkcji prawdopodobieństwa $P(X = x_i) = p_i$, g - dowolna to funkcja prawdopodobieństwa zmiennej losowej $Y = g(X)$ ma postać:

$g(x_1)$	$g(x_2)$	...	$g(x_k)$
p_1	p_2	...	p_k

Po uporządkowaniu rosnąco wartości $g(x_i)$ i zsumowaniu odpowiednich prawdopodobieństw.

Dokładniej

$$P(Y = y) = P(g(X) = y) = P\left(\bigcup_{\{i: g(x_i)=y\}} (X = x_i)\right) = \sum_{\{i: g(x_i)=y\}} P(X = x_i) = \sum_{\{i: g(x_i)=y\}} p_i$$

Przykład.

X - zmienna losowa skokowa o funkcji prawdopodobieństwa:

-4	-2	-1	0	1	2
0,4	0,1	0,1	0,1	0,1	0,2

wyznamy funkcję prawdopodobieństwa zmiennej losowej $Y = \text{sgn}X$.

$$\text{sgn}(-4) = \text{sgn}(-2) = \text{sgn}(-1) = -1.$$

$$\text{sgn}(0) = 0.$$

$$\text{sgn}(1) = \text{sgn}(2) = 1.$$

Zatem funkcja prawdopodobieństwa zmiennej losowej Y jest następująca

-1	0	1
0,6	0,1	0,3

W niektórych zagadnieniach wyznaczania rozkładu funkcji zmiennej losowej najpierw wyznaczamy dystrybuantę rozkładu zmiennej losowej $Y = g(X)$, wg schematu

$$F_Y(y) = P(Y < y) = P(g(X) < y) = P(X \in g^{-1}((-\infty, y)) < y)$$

następnie jeśli to możliwe, wyznaczamy funkcję prawdopodobieństwa (gdy jest to rozkład skokowy) lub gęstość (gdy jest to rozkład ciągły).

Przykład.

Jeśli X ma rozkład o gęstości

$$f(x) = \begin{cases} 0 & \text{dla } x \notin [0, 3] \\ \frac{1}{3} & \text{dla } x \in [0, 3] \end{cases} \quad (\text{rozkład jednostajny na } [0, 3])$$

$$Y = \max(2, X),$$

$$\text{wtedy } F_Y(y) = P(Y < y) = P(\max(2, X) < y) = \begin{cases} 0 & \text{dla } y \leq 2 \\ \frac{1}{3}y & \text{dla } 2 < y \leq 3 \\ 1 & \text{dla } y > 3 \end{cases}$$

Nie jest to ani rozkład skokowy ani ciągły. Nie można więc wyznaczyć ani funkcji prawdopodobieństwa ani gęstości.

Jest to rozkład mieszany skokowo - ciągły i zgodnie z twierdzeniem o rozkładzie dystrybuanty powyższą dystrybuantę można przedstawić w postaci

$$F_Y = c_1 F_1 + c_2 F_2$$

$$\text{gdzie } c_1 = 2/3, F_1(y) = \begin{cases} 0 & \text{dla } y \leq 2 \\ 1 & \text{dla } y > 2 \end{cases},$$

$$c_2 = 1/3, F_2(y) = \begin{cases} 0 & \text{dla } y \leq 2 \\ y - 2 & \text{dla } 2 < y \leq 3, \\ 1 & \text{dla } y > 3 \end{cases}$$

Funkcje zmiennych losowych 2 wymiarowych.

(X_1, X_2) - dana zmienna losowa ciągła o gęstości f .

$$Y = g(X_1, X_2) \quad g - \text{borelowska,}$$

Dystrybuanta tej zmiennej losowej ma postać

$$G(y) = \iint_{(g(x_1, x_2) \leq y)} f(x_1, x_2) dx_1 dx_2$$

gęstość $g(y)$ wyznaczamy przez różniczkowanie.

Przykład.

$$Y = X_1 \cdot X_2,$$

$$G(y) = \iint_{x_1 \cdot x_2 \leq y} f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^0 \left(\int_{y/x_1}^{\infty} f(x_1, x_2) dx_2 \right) dx_1 + \int_0^{\infty} \left(\int_{-\infty}^{y/x_1} f(x_1, x_2) dx_2 \right) dx_1$$

$$\text{wtedy} \quad g(y) = \int_{-\infty}^0 \frac{1}{x} f(x, \frac{y}{x}) dx + \int_0^{\infty} \frac{1}{x} f(x, \frac{y}{x}) dx$$

Przykład.

(X_1, X_2) - zmienna losowa o rozkładzie jednostajnym w kwadracie $(0, 1) \times (0, 1)$. Wyznaczyć rozkład pola prostokąta o bokach x_1, x_2 tzn. zmiennej losowej

$$Y = X_1 \cdot X_2.$$

$$G(y) = \iint_{x_1 \cdot x_2 \leq y} dx_1 dx_2 = 1 - \iint_{x_1 \cdot x_2 \geq y} dx_1 dx_2 = 1 - \int_y^1 \left(\int_{y/x_1}^1 dx_2 \right) dx_1 = y(1 - \ln y)$$

dla $0 < y \leq 1$

$$\text{stąd} \quad G(y) = \begin{cases} 0 & \text{dla } y \leq 0 \\ y(1 - \ln y) & \text{dla } 0 < y \leq 1 \\ 1 & \text{dla } 1 < y \end{cases}$$

zatem $g(y) = -\ln y$ dla $0 < y \leq 1$

Przykład.

$$Y = X_2/X_1,$$

$$G(y) = \int_{-\infty}^0 \left(\int_{y/x_1}^{\infty} f(x_1, x_2) dx_2 \right) dx_1 + \int_0^{\infty} \left(\int_{-\infty}^{y/x_1} f(x_1, x_2) dx_2 \right) dx_1$$

$$\text{wtedy} \quad g(y) = - \int_{-\infty}^0 x f(x, yx) dx + \int_0^{\infty} x f(x, yx) dx$$

Przykład.

X_1, X_2 - niezależne zmienne losowe.

$$X_1 - N(0, \sigma_1), \quad X_2 - N(0, \sigma_2).$$

Niech $\tilde{X}_1 = \frac{X_1}{\sigma_1}, \quad \tilde{X}_2 = \frac{X_2}{\sigma_2}$ (mają rozkład $N(0, 1)$).

$$\tilde{g}(\tilde{y}) = - \int_{-\infty}^0 x \cdot \frac{e^{-x^2/2}}{\sqrt{2\pi}} \cdot \frac{e^{-\tilde{y}x^2/2}}{\sqrt{2\pi}} dx + \int_0^{\infty} x \cdot \frac{e^{-x^2/2}}{\sqrt{2\pi}} \cdot \frac{e^{-\tilde{y}x^2/2}}{\sqrt{2\pi}} dx = \frac{1}{\pi(1 + \tilde{y}^2)}$$

(rozkład Cauchy'ego)

i korzystając funkcji liniowej od zmiennej losowej $Y = \tilde{Y} \frac{\sigma_1}{\sigma_2}$ mamy

$$g(y) = \frac{1}{\pi \frac{\sigma_2}{\sigma_1} \left(1 + y^2 \left(\frac{\sigma_1}{\sigma_2} \right)^2 \right)}$$

Przykład.

$$Y = X_1 + X_2 ,$$

$$G(y) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{y-x_1} f(x_1, x_2) dx_2 \right) dx_1$$

wtedy

$$g(y) = \int_{-\infty}^{\infty} f(x, y-x) dx = \int_{-\infty}^{\infty} f(y-x, x) dx$$

Uwaga.

Jeśli X_1, X_2 - niezależne zmienne losowe to gęstość sumy wyraża się splotem gęstości brzegowych (p. dalej).

Przykład.

$$Y = X_1 - X_2 ,$$

wtedy

$$g(y) = \int_{-\infty}^{\infty} f(x, x-y) dx = \int_{-\infty}^{\infty} f(x-y, x) dx$$

Suma niezależnych zmiennych losowych.

Własności:

- 1) X, Y niezależne skokowe zmienne losowe o funkcjach prawdopodobieństwa $P(X = x_i), P(Y = y_j)$; wtedy funkcja prawdopodobieństwa zmiennej losowej $X + Y$ wyraża się wzorem:
 $P(Z = z_k) = \sum P(X = x_i)P(Y = z_k - x_i); \quad (z_k = x_i + y_j)$
- 2) X, Y niezależne ciągle zmienne losowe o gęstościach f_1 i f_2 ; wtedy gęstość zmiennej losowej $X + Y$ wyraża się wzorem:

$$f(x) = \int_{-\infty}^{\infty} f_1(t)f_2(x-t)dt = \int_{-\infty}^{\infty} f_1(x-t)f_2(t)dt$$

(splot gęstości składników).

Funkcje zmiennych losowych n - wymiarowych.

Przykład.

System składa się n układów z których każdy ma czas bezawaryjnej pracy określony rozkładem wykładniczym X_i o parametrze a_i niezależnym od pozostałych układów. Wyznaczyć rozkład bezawaryjnego czasu pracy całego systemu (system działa jeśli pracuje chociaż jeden układ).

$$Y = X_1 + X_2 + X_3 + \dots + X_n,$$

Przez indukcję pokazuje się, że Y ma rozkład o gęstości:

$$g(y) = (-1)^{n-1} \prod_{i=1}^n a_i \sum_{j=1}^n \frac{e^{-a_j y}}{\prod_{\substack{k=1 \\ k \neq j}}^n (a_j - a_k)} \quad \text{dla } y > 0$$

wtedy
$$G(y) = (-1)^{n-1} \prod_{i=1}^n a_i \sum_{j=1}^n \frac{1 - e^{-a_j y}}{\prod_{\substack{k=1 \\ k \neq j}}^n (a_j - a_k)}$$

Otrzymany rozkład nazywamy uogólnionym rozkładem Erlanga n - tego rzędu T_n .

$$E(T_n) = E\left(\sum_{k=1}^n T_i\right) = \sum_{k=1}^n \frac{1}{a_k}$$

$$D^2(T_n) = D^2\left(\sum_{k=1}^n T_i\right) = \sum_{k=1}^n \frac{1}{a_k^2}$$

gdzy $a_1 = a_2 = \dots = a_n = \lambda$ to

$$g(y) = \frac{\lambda(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t} = \lambda P(n-1, \lambda t) \quad t > 0 \quad \text{dla } y > 0$$

gdzie $P(n-1, \lambda t)$ jest rozkładem Poissona.

Przykład.

$$Y = \min(X_1, X_2)$$

$$G(y) = F_1(y) + F_2(y) - F(y, y)$$

wtedy
$$g(y) = f_1(y) + f_2(y) - \int_{-\infty}^y f(y, x) dx - \int_{-\infty}^y f(x, y) dx$$

Uwaga.

Jeśli X_1, X_2 - niezależne zmienne losowe to:

$$G(y) = F_1(y) + F_2(y) - F_1(y) \cdot F_2(y)$$

$$g(y) = f_1(y)(1 - F_2(y)) + f_2(y)(1 - F_1(y))$$

Jeśli X_1, X_2 - niezależne zmienne losowe o takim samym rozkładzie to:

$$G(y) = F(y)(2 - F(y))$$

wtedy
$$g(y) = 2f(y)(1 - F(y))$$

Przykład.

$$Y = \max(X_1, X_2)$$

$$G(y) = F(y, y)$$

wtedy
$$g(y) = \int_{-\infty}^y f(y, x) dx + \int_{-\infty}^y f(x, y) dx$$

Uwaga.

Jeśli X_1, X_2 - niezależne zmienne losowe o takim samym rozkładzie to:

$$G(y) = F^2(y)$$

wtedy
$$g(y) = 2f(y)F(y)$$

Przykład.

Jeśli $X_1, X_2, \dots, X_n$ - niezależne zmienne losowe o dystrybucjach $F_1, F_2, \dots, F_n$, wtedy zmienna losowa

a)

$$Y = \max(X_1, X_2, \dots, X_n)$$

ma rozkład o dystrybucji

$$G(y) = P(Y < y) = P(\max(X_1, X_2, \dots, X_n) < y) = P(X_1 < y, X_2 < y, \dots, X_n < y) = F_1(y) \cdot F_2(y) \cdot \dots \cdot F_n(y)$$

Np. gdy $X_1, X_2, \dots, X_n$ mają niezależne **rozkłady jednostajne** na przedziale $(0, 1)$, to

$$G(y) = \begin{cases} 0 & \text{dla } y \leq 0 \\ y^n & \text{dla } 0 < y \leq 1 \\ 1 & \text{dla } y > 1 \end{cases}$$

Gęstość tego rozkładu wyraża się wzorem

$$f(y) = \begin{cases} 0 & \text{dla } y \leq 0 \text{ oraz } y \geq 1 \\ ny^{n-1} & \text{dla } 0 < y < 1 \end{cases}$$

b)

$$Y = \min(X_1, X_2, \dots, X_n)$$

ma rozkład o dystrybuancie

$$G(y) = P(Y < y) = 1 - P(Y \geq y) = 1 - P(X_1 \geq y, X_2 \geq y, \dots, X_n \geq y) = \\ = 1 - [(1 - F_1(y)) \cdot (1 - F_2(y)) \cdot \dots \cdot (1 - F_n(y))]$$

Np. gdy $X_1, X_2, \dots, X_n$ mają niezależne **rozkłady jednostajne** na przedziale $(0, 1)$, to

$$G(y) = \begin{cases} 0 & \text{dla } y \leq 0 \\ 1 - (1 - y)^n & \text{dla } 0 < y \leq 1 \\ 1 & \text{dla } y > 1 \end{cases}$$

Gęstość tego rozkładu wyraża się wzorem

$$f(y) = \begin{cases} 0 & \text{dla } y \leq 0 \text{ oraz } y \geq 1 \\ n(1 - y)^{n-1} & \text{dla } 0 < y < 1 \end{cases}$$

Rozkłady funkcji od rozkładu normalnego.

$(X_1, X_2, X_3, \dots, X_n)$ - rozkład normalny.

$Y = g(X_1, X_2, X_3, \dots, X_n)$, należy wyznaczyć rozkład Y .

Przykład.

$$Y = a_1X_1 + a_2X_2 + a_3X_3 + \dots + a_nX_n + b$$

$n = 2$

$$g(y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}}$$

$$\text{gdzie: } m_y = a_1m_1 + a_2m_2 + b, \quad \sigma_y^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2r\sigma_1\sigma_2,$$

Przez indukcję można pokazać, że dla dowolnego n Y ma rozkład normalny o parametrach:

$$m_y = a_1m_1 + a_2m_2 + \dots + a_nm_n + b, \\ \sigma_y^2 = \sum a_i^2\sigma_i^2 + 2\sum_{i < j} a_i a_j r_{ij} \sigma_i \sigma_j,$$

Przykład.

$X_1, \dots, X_n$ - niezależne, o rozkładzie $N(0, 1)$. $Y_n = X_1^2 + \dots + X_n^2$ ma **rozkład chi kwadrat**

$$n \in N \quad f(y) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$EX = n; \quad D^2X = 2n$$

Przykład.

$X \sim N(m, \sigma)$, $Y = e^X$

Ma rozkład logarytmiczno-normalny.

Nazwa pochodzi stąd, że $X = \ln Y$ ma rozkład normalny.

ZADANIA

Zadanie 1

Wyznaczyć rozkład sumy dwóch niezależnych zmiennych losowych o rozkładzie Poissona z parametrami λ_1, λ_2 .

(odp. Jest to rozkład Poissona z parametrem $\lambda_1 + \lambda_2$)

Zadanie 2

Wyznaczyć rozkład różnicy dwóch niezależnych zmiennych losowych o rozkładzie wykładniczym z parametrem a .

(odp. Jest to rozkład Laplace'a: $g(y) = ae^{\frac{-a|y|}{2}}$)

Zadanie 3

System składa się z 2 układów z których każdy ma czas bezawaryjnej pracy określony rozkładem wykładniczym X_i o parametrze a_i niezależnym od drugiego układu. Wyznaczyć rozkład bezawaryjnego czasu pracy całego systemu (system działa jeśli oba układy pracują). $Y = \min(X_1, X_2)$

Odp. Jest to rozkład wykładniczy o parametrze $a_1 + a_2$.

Zadanie 4

System składa się z 2 układów, czas włączenia tych układów ma rozkład wykładniczy X_i o parametrze a_i niezależnym od drugiego układu. System zaczyna działać dopiero gdy włączą się oba układy. Wyznaczyć rozkład czasu rozpoczęcia pracy całego systemu. $Y = \max(X_1, X_2)$.

(odp. $g(y) = a_1 e^{-a_1 y} (1 - e^{-a_2 y}) + a_2 e^{-a_2 y} (1 - e^{-a_1 y})$ $y > 0$)

Nie jest to rozkład wykładniczy)

Zadanie 5

Wyznaczyć gęstość rozkładu logarytmiczno-normalnego.

(odp. $g(y) = \frac{1}{\sigma y \sqrt{2\pi}} e^{-\frac{(\ln y - m)^2}{2\sigma^2}}$ $y > 0$)

Zadanie 6

Wykonujemy niezależnie pomiar oporności R oraz prądu I (znamy błędy względne $\delta_R = \frac{DR}{R} = a$, $\delta_I = \frac{DI}{I} = b$). Następnie na tej podstawie chcemy ocenić poziom napięcia $U = I \cdot R$, oraz wartość mocy $P = I^2 R$.

Oblicz $D^2 U$, $D^2 P$, $\text{cov}(U, P)$, $\rho(U, P)$.

$$(\text{odp. } Y = (U, P) = T(R, I) \text{ gdzie } T = \begin{bmatrix} \frac{\partial U}{\partial R} & \frac{\partial U}{\partial I} \\ \frac{\partial P}{\partial R} & \frac{\partial P}{\partial I} \end{bmatrix} = \begin{bmatrix} I & R \\ I^2 & 2IR \end{bmatrix},$$

$$K_Y = \begin{bmatrix} (a^2 + b^2)I^2 R^2 & (a^2 + 2b^2)I^3 R^2 \\ (a^2 + 2b^2)I^3 R^2 & (a^2 + 4b^2)I^4 R^2 \end{bmatrix}, \rho = \frac{a^2 + 2b^2}{\sqrt{a^2 + b^2}\sqrt{a^2 + 4b^2}})$$

Zadanie 7

Przy równoległym łączeniu ogniw prąd w obwodzie określony jest wzorem:

$$I = \frac{E}{R + \frac{W}{n}}$$

E - siła elektromotoryczna ogniwa, W - jego opór wewnętrzny, n - ilość ogniw, R - opór zewnętrzny.

Wyznaczyć wartość oczekiwaną i wariancję prądu jeśli E, R, W są niezależne i dane są ich wartości oczekiwane oraz odchylenia standardowe.

Zadanie 8

X, Y - niezależne zmienne losowe o rozkładzie określonym funkcją prawdopodobieństwa

$$P(X = 0) = 1/2, P(X = 1) = 3/8, P(X = 2) = 1/8,$$

Wyznacz funkcję prawdopodobieństwa zmiennej losowej X + Y.

Zadanie 9

a) X, Y - niezależne zmienne losowe o rozkładzie jednostajnym w [0, 1].

Wyznacz gęstość zmiennej losowej X + Y. Wyznacz parametry tej zmiennej losowej.

b) X, Y, Z - niezależne zmienne losowe o rozkładzie jednostajnym w [0, 1].

Wyznacz gęstość zmiennej losowej X + Y + Z. Wyznacz parametry tej zmiennej losowej.

Zadanie wykonaj stosując splot gęstości a wynik sprawdź za pomocą funkcji charakterystycznych.

Naszkicuj i porównaj wykresy gęstości zmiennych losowych

$$X, \quad X + Y, \quad X + Y + Z.$$

Zauważ, że gdy rośnie liczba rozpatrywanych składników, wykres gęstości staje się podobny do krzywej Gaussa.

Zadanie 10

X - zmienna losowa odpowiadająca mierzonej wielkości (zakładamy, że ma rozkład jednostajny w [0, 8]), Y - niezależna od X zmienna losowa opisująca błąd pomiaru (zakładamy, że ma rozkład normalny N(0, 1)).

Wyznacz gęstość zmiennej losowej odpowiadającej wynikowi pomiaru U = X + Y. Wyznacz parametry tej zmiennej losowej.

Zadanie wykonaj stosując splot gęstości a wynik sprawdź za pomocą funkcji charakterystycznych.

Naszkicuj wykres gęstości zmiennej losowej X + Y.