Rachunek prawdopodobieństwa

Stanisław Jaworski

Rachunek prawdopodobieństwa: dział matematyki zajmujący się badaniem modeli zjawisk losowych (przypadkowych) i praw nimi rządzących (*Encyklopedia Popularna PWN*, 1998)

Rachunek prawdopodobieństwa zajmuje się zdarzeniami, pojawiającymi się przy wykonywaniu doświadczeń, których wyniku nie da się z góry przewidzieć, a jednocześnie dających się powtarzać w tych samych warunkach.

Pojęciem pierwotnym w rachunku prawdopodobieństwa jest **przestrzeń zdarzeń elementarnych**. Będziemy ją oznaczać przez Ω.

Przykład. Rzut monetą.

$$\Omega = \{\mathrm{O}, \mathrm{R}\}$$

■

Przykład. Rzut kostką.

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

■

Przykład. Rzut monetą do chwili pojawienia się orła.

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\},$$

gdzie ω_n oznacza, że w pierwszych $n - 1$ rzutach wypadły reszki, a za n – tym razem wypadł orzeł. Możliwych wyników jest nieskończenie wiele. Dadzą się ustawić w ciąg, tzn. że jest ich **przeliczalnie wiele.** ■

Przykład. Ze zbioru n ponumerowanych elementów losujemy dwa elementy.

$$\Omega = \{(\omega_i, \omega_j)|\ i, j = 1, 2, \ldots, n,\ i < j\}$$

ω_i oznacza wylosowanie elementu o numerze i. ■

Przykład. Czas oczekiwania na autobus.

$$\Omega = [0, \infty)$$

■

Przykład. Niech $T_k \in \langle[0, 45],\ k = 1, 2, \ldots, 10$, oznacza spóźnienie k – tego studenta na losowo wybrany wykład (w minutach).

$$(T_1, T_2, \ldots, T_{10}) \in \Omega$$

$$\Omega = [0, 45] \times [0, 45] \times \ldots \times [0, 45] = [0, 45]^{10}$$

■

Podstawowe pojęcia rachunku prawdopodobieństwa

Definicja. Rodzinę $\mathcal{F}$ spełniającą warunki

1. $\mathcal{F} \neq \emptyset$
2. Jeśli $A \in \mathcal{F}$, to $\Omega \setminus A \in \mathcal{F}$
3. Jeśli $A_i \in \mathcal{F}$ dla $i = 1, 2, \dots$, to $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

nazywamy σ – ciałem podzbiorów zbioru Ω.

Zdarzenie losowe jest elementem rodziny $\mathcal{F}$

Definicja. Prawdopodobieństwem nazywamy dowolną funkcję P, określoną na $\sigma-$ciele zdarzeń $\mathcal{F} \subseteq 2^{\Omega}$, spełniającą warunki

A1. $P : \mathcal{F} \to \mathbf{R}_+$;

A2. $P(\Omega) = 1$

A3. Jeśli $A_i \in \mathcal{F},\ i = 1, 2, \dots$ oraz $A_i \cap A_j = \emptyset$ dla $i \neq j$, to

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Mówimy, że matematyczny model doświadczenia losowego to trójka $(\Omega, \mathcal{F}, P)$, którą nazywamy **przestrzenią probabilistyczną**

Przykład. Rozkład prawdopodobieństwa w skończonej przestrzeni zdarzeń

Niech

$$\Omega = \{\omega_1,\ \omega_2, \dots\ ,\ \omega_n\}.$$

Niech

$$p_i \geqslant 0,\ i = 1, 2, \dots , n,$$

będą tak dobrane, że

$$\sum_{i=1}^{n} p_i = 1$$

Wówczas funkcję P określamy w następujący sposób:

$$P(\{\omega_i\}) = p_i \text{ oraz}$$

dla $A \subseteq \Omega$ postaci $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}\}$

$$P(A) = p_{i_1} + p_{i_2} + \dots + p_{i_k}$$

Tak określona funkcja spełnia układ aksjomatów Kołmogorowa dla $\mathcal{F} = 2^{\Omega}$ ■

Przykład. Rzut kostką.

ω_i	1	2	3	4	5	6
p_i	1/6	1/6	1/6	1/6	1/6	1/6

$$P(\{1,2,5\}) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2$$

ω_i	1	2	3	4	5	6
p_i	1/12	1/12	1/12	3/12	3/12	3/12

$$P(\{1,2,5\}) = 1/12 + 1/12 + 3/12 = 5/12 < 1/2$$

■

Przykład. Przeliczalna przestrzeni zdarzeń

$$\Omega = \{\omega_1,\ \omega_2, \dots\}$$

$$p_i \geqslant 0, \quad \sum_{i=1}^{\infty} p_i = 1$$

$$P(\{\omega_i\}) = p_i,\ P(A) = \sum_{j:\omega_j \in A} p_j$$

(?) Tak określona funkcja spełnia układ aksjomatów Kołmogorowa dla $\mathcal{F} = 2^{\Omega}$ ■

Przykład. Liczba zarejestrowanych cząstek w odcinku czasu $[0, t]$.

$$\Omega = \{0, 1, \dots\}$$

$$P(\{k\}) = e^{-\alpha t} \frac{(\alpha t)^k}{k!}, \quad k = 0, 1, \dots$$

(?) Zachodzi $\sum_{k=0}^{\infty} e^{-\alpha t} \frac{(\alpha t)^k}{k!} = 1$ ■

Ciągła przestrzeń zdarzeń

Przykład. Czas oczekiwania na pierwszą cząstkę.

$$\Omega = [0, \infty)$$

Zdarzenie (t, ∞): pierwsza cząstka pojawi się później niż w chwili t

$$P((t, \infty)) = e^{-\alpha t} \frac{(\alpha t)^0}{0!} = e^{-\alpha t}$$

Stąd dla dowolnych $s < t$

$$P((s, t]) = e^{-\alpha s} - e^{-\alpha t}$$

■

Przykład. Rzut strzałką do tarczy o promieniu 1.

Model 1.

$$\Omega = \{(x,y) : x^2 + y^2 \leq 1\}$$

(x,y)– współrzędne kartezjańskie punktu trafienia strzałki w tarczę

Szansa trafienia w zbiór $A \subseteq \Omega$

$$P(A) = \frac{\text{pole } A}{\text{pole } \Omega} = \frac{|A|}{|\Omega|} = \frac{|A|}{\pi}$$

Zdarzenie $A_r = \{(x,y) : \ x^2 + y^2 \leq r^2\}$: trafienie w dziesiątkę

$$P(A_r) = \frac{\pi r^2}{\pi} = r^2$$

Model 2.

$$\Omega = \{(\varrho, \phi) : 0 \leq \varrho \leq 1, 0 \leq \phi \leq 2\pi\} = [0,1] \times [0, 2\pi]$$

(ϱ, ϕ)– współrzędne biegunowe punktu trafienia strzałki w tarczę

Szansa trafienia w zbiór $A \subseteq \Omega$:

$$P(A) = \frac{\text{pole } A}{\text{pole } \Omega} = \frac{|A|}{|\Omega|} = \frac{|A|}{2\pi}$$

Zdarzenie $A_r = \{(\varrho, \phi) : \varrho \leq r\}$: trafienie w dziesiątkę

$$P(A_r) = \frac{2\pi r}{2\pi} = r$$

■

Model 3.

$$\Omega = \{\varrho : 0 \leq \varrho \leq 1\} = [0,1]$$

ϱ– odległość punktu trafienia od środka tarczy

Zdarzenie $A_r = \{\varrho : \varrho \leq r\}$: trafienie w dziesiątkę

$$P(A_r) = \frac{\pi r^2}{\pi} = r^2$$

Zdarzenie $A_{rk} = \{\varrho : r < \varrho \leq k\}$: trafienie w dziewiątkę

$$P(A_{rk}) = k^2 - r^2 = 2(k-r)\frac{k+r}{2}$$

Co łączy podane przykłady dla przestrzeni ciągłych?

$$P(A) = \int_A f, \quad \text{gdzie } f \geq 0$$

Czas oczekiwania na pierwszą cząstkę

$$f(x) = \alpha x e^{-\alpha x}, \quad P((s,t]) \stackrel{?}{=} \int\limits_s^t f(x)\, dx$$

Rzut strzałką do tarczy (Model 1.)

$$f(x,y) = \frac{1}{\pi}, \quad P(A_r) \stackrel{?}{=} \int\limits_{A_r} f(x,y)\, dx\, dy$$

Rzut strzałką do tarczy (Model 2.)

$$f(\varrho,\phi) = \frac{1}{2\pi}, \quad P(A_r) \stackrel{?}{=} \int\limits_{A_r} f(\varrho,\phi)\, d\varrho\, d\phi$$

Rzut strzałką do tarczy (Model 3.)

$$f(\varrho) = 2\varrho, \quad P(A_r) \stackrel{?}{=} \int\limits_{A_r} f(\varrho)\, d\varrho$$

Problem: Jak określić $\mathcal{F}$?

Czas oczekiwania na pierwszą cząstkę

$$\mathcal{F} = \mathcal{B}(R_+)$$

Rzut strzałką do tarczy (Model 1.)

$$\mathcal{F} = \mathcal{B}(K(0,1))$$

Rzut strzałką do tarczy (Model 2.)

$$\mathcal{F} = \mathcal{B}([0,1] \times [0,2\pi])$$

Rzut strzałką do tarczy (Model 3.)

$$\mathcal{F} = \mathcal{B}([0,1])$$

Własności prawdopodobieństwa

Twierdzenie 1. Jeśli $(\Omega, \mathcal{F}, P)$ jest przestrzenią probabilistyczną i $A,\ B,\ A_1,\ A_2, \ldots,\ A_n \in \mathcal{F}$, to:

W1. $P(\emptyset) = 0$

W2. Jeśli $A_1,\ A_2, \ldots,\ A_n$ wykluczają się wzajemnie, tj. $A_i \cap A_j = \emptyset$ dla $i \neq j$, to

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

W3. $P(A') = 1 - P(A)$, gdzie $A' = \Omega \setminus A$

W4. Jeśli $A \subseteq B$, to $P(B \setminus A) = P(B) - P(A)$

W5. Jeśli $A \subseteq B$, to $P(A) \leqslant P(B)$

W6. $P(A) \leqslant 1$

W7. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Dowód.

ad W1.

$$\text{Niech } A_1 = \Omega,\ \ A_i = \emptyset \text{ dla } i = 2, 3, \ldots$$
$$\Downarrow \text{ aksjomat A3.}$$
$$P(\Omega) = P(\Omega) + \sum_{i=2}^{\infty} P(\emptyset)$$
$$\Downarrow \text{ aksjomat A1.}$$
$$P(\emptyset) = 0$$

ad W2.

$$\text{Niech } A_k = \emptyset,\ \text{dla } k \geqslant n$$
$$\Downarrow \text{ aksjomat A3. oraz własność W1.}$$
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

ad W3.

$$1 = P(\Omega) = P(A \cup A') \overset{\text{W2.}}{=} P(A) + P(A')$$

ad W4.

Jeśli $A \subset B$, to $B = A \cup (B \setminus A)$. Zatem

$$P(B) \overset{\text{W2.}}{=} P(A) + P(B \setminus A)$$

ad W5.

$$P(B) - P(A) \overset{\text{W4.}}{=} P(B \setminus A) \overset{\text{A1.}}{\geqslant} 0$$

ad W6.

Wystarczy zastosować W5. dla $B = \Omega$

ad W7.

$$\begin{aligned} A \cup B &= [A \setminus (A \cap B)] \cup (A \cap B) \cup [B \setminus (A \cap B)] \\ &\Downarrow \ \text{W2, W4.} \\ P(A \cup B) &= P(A) - P(A \cup B) + \\ &+ P(A \cap B) + P(B) - P(A \cap B) = \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

■

Zauważmy, że

$$A \cup B = \underbrace{[A \cap B'] \cup [A \cap B] \cup [A' \cap B]}_{\text{trzy składowe sumy}}$$

Zatem każda składowa sumy $A_1 \cup A_2 \cup \ldots \cup A_n$ da się przedstawić, po odpowiednim przenumerowaniu zbiorów, w postaci

$$A_1 \cap A_2 \cap \ldots \cap A_k \cap A'_{k+1} \cap A'_{k+2} \cap \ldots \cap A'_n,$$
$$\text{gdzie } k \geqslant 1$$

Twierdzenie 2. (Wzór włączeń i wyłączeń)

$$P(A_1 \cup A_2 \cup \ldots \cup A_n)$$
$$= \sum_{1 \leqslant i \leqslant n} P(A_i) - \sum_{1 \leqslant i_1 \leqslant i_2 \leqslant n} P(A_{i_1} \cap A_{i_2}) +$$
$$\ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n)$$

Dowód. Zbiór $A_1 \cup A_2 \cup \ldots \cup A_n$ daje się zapisać w postaci sumy rozłącznych składowych. Zatem Lewa strona równania włącza każdą składową dokładnie raz. Musimy pokazać, że prawa strona równania wprowadza każdą składową też dokładnie raz.
W pierwszym składniku wzoru, czyli

$$\sum_{1 \leqslant i \leqslant n} P(A_i)$$

każda składowa postaci

$$A_1 \cap A_2 \cap \ldots \cap A_k \cap A'_{k+1} \cap A'_{k+2} \cap \ldots \cap A'_n$$

zostanie włączona k razy, w drugim, czyli

$$\sum_{1 \leqslant i_1 \leqslant i_2 \leqslant n} P(A_{i_1} \cap A_{i_2}),$$

wyłączona $\binom{k}{2}$ razy, itd. Ostatecznie liczba włączeń wyniesie

$$\binom{k}{1} - \binom{k}{2} + \binom{k}{3} + \ldots (-1)^{k+1}\binom{k}{k} = 1.$$

Uwaga. Korzystam ze wzoru dwumianowego Newtona:

$$(a+b)^k = \sum_{i=0}^{k} \binom{k}{i} a^{k-i} b^i$$

■

Przykład. n listów losowo wkładamy do kopert. Jakie jest prawdopodobieństwo, że choć jeden list dotrze do adresata?

Niech A_i oznacza zdarzenie, że $i-$ty list dotrze do adresata. Zatem

$$P\left(\bigcup_{i=1}^{n} A_i\right) = n\,\overbrace{\frac{(n-1)!}{n!}}^{P(A_1)} - \binom{n}{2}\,\overbrace{\frac{(n-2)!}{n!}}^{P(A_1 \cap A_2)} + \ldots$$

$$+ (-1)^n \binom{n}{n-1}\frac{1}{n!} + (-1)^{n+1}\frac{1}{n!} =$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} + \ldots + (-1)^n \frac{1}{(n-1)!} + (-1)^{n+1} \frac{1}{n!}$$

$$= \sum_{i=1}^{n} \frac{(-1)^{i+1}}{i!} = 1 + \sum_{i=0}^{n} \frac{(-1)^{i+1}}{i!} = 1 - \sum_{i=0}^{n} \frac{(-1)^i}{i!}$$

$$\approx 1 - e^{-1}$$

Błąd oszacowania

$$\left| P\left(\bigcup_{i=1}^{n} A_i\right) - (1 - e^{-1}) \right| \leqslant \frac{1}{(n+1)!}$$

Skorzystałem z oszacowania:

$$\left| e^x - \sum_{i=0}^{n} \frac{x^i}{i!} \right| \leqslant \frac{|x|^{n+1}}{(n+1)!}$$

$\blacksquare$

Twierdzenie 3. (O ciągłości). Niech$(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną.

(i) Jeśli $(A_n)_{n=1}^{\infty}$ jest wstępującą rodziną zdarzeń oraz $\bigcup\limits_{n=1}^{\infty} A_n = A$, to

$$P(A) = \lim_{n\to\infty} P(A_n).$$

(ii) Jeśli $(A_n)_{n=1}^{\infty}$ jest zstępującą rodziną zdarzeń oraz $\bigcap\limits_{n=1}^{\infty} A_n = A$, to

$$P(A) = \lim_{n\to\infty} P(A_n).$$

Rodzinę zdarzeń A_i nazywamy wstępującą, jeśli

$$A_1 \subseteq A_2 \subseteq \ldots \subset A_n \subseteq A_{n+1} \cdots$$

i zstępującą, jeśli

$$A_1 \supseteq A_2 \supseteq \ldots \supset A_n \supseteq A_{n+1} \cdots$$

Dowód. (i) Niech

$$B_1 = A_1,\ B_2 = A_2 \setminus A_1 \text{ i ogólnie: } B_n = A_n \setminus A_{n-1}$$

Wtedy zdarzenia B_i wykluczają się,

$$\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i = A_n,$$

a także $\bigcup_{i=1}^{\infty} B_i = A$. Z przeliczalnej addytywności wynika, że

$$P(A) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) =$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} P(A_n)$$

(ii) Rozpatrzmy rodzinę wstępującą $(C_n)_{n=1}^{\infty}$, gdzie $C_n = A'_n$. Wtedy

$$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} A'_n = \left[\bigcap_{n=1}^{\infty} A_n\right]' = A'$$

i wystarczy skorzystać z (i) ■

Prawdopodobieństwo warunkowe

Definicja. Prawdopodobieństwem warunkowym zajścia zdarzenia A pod warunkiem zajścia zdarzenia B, gdzie $P(B) > 0$, nazywamy liczbę

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

(?) **Uwaga.** Przy ustalonym B prawdopodobieństwo warunkowe $P(A|B)$ jest zwykłym prawdopodobieństwem na $(\Omega, \mathcal{F})$, a także na $(B, \mathcal{F}_B)$, gdzie

$$\mathcal{F}_B = \{A \cap B : \ A \in \mathcal{F}\}$$

(?) **Wzór łańcuchowy.** Jeśli $P(A_1 \cap \ldots \cap A_{n-1}) > 0$, to

$$\begin{aligned} P(A_1 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1) \times \\ \times P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P(A_n|A_1 \cap \ldots \cap A_{n-1}) \end{aligned}$$

Definicja. Rozbiciem przestrzeni Ω nazywamy rodzinę zdarzeń $\{H_i\}_{i\in I}$, które wzajemnie wykluczają się, zaś ich suma jest równa Ω.

Twierdzenie 4.
Jeżeli $\{H_1,\ H_2, \dots ,\ H_n\}$ jest rozbiciem Ω na zdarzenia o dodatnim prawdopodobieństwie, to dla dowolnego zdarzenia A

$$P(A) = \sum_{i=1}^{n} P(A|H_i)P(H_i)$$

Dowód.

$$P(A) = P(\bigcup_{i=1}^{n}(A \cap H_i)) = \sum_{i=1}^{n} P(A|H_i)P(H_i)$$

$\blacksquare$

(?) **Uwaga.**Twierdzenie jest prawdziwe i dla rozbicia Ω na przeliczalną liczbę zdarzeń $H_i,\ i = 1, 2, \dots$.

Przykład. W loterii fantowej szansa wygranej jest równa p, przegranej – q, a z prawdopodobieństwem r wyciągamy los „graj dalej”. Los „graj dalej” wrzucamy z powrotem do urny i dokonujemy ponownego losowania. Jakie jest prawdopodobieństwo wygranej?

$A-$wyciągneliśmy los wygrywający
$B-$wyciągneliśmy los przegrywający
$C-$wyciągneliśmy los „graj dalej”
$W-$wygraliśmy na loterii

$$\begin{aligned}P(W) = {} & P(W|A)P(A) + P(W|B)P(B) + \\ & + P(W|C)P(C) = 1 \cdot p + 0 \cdot q + P(W) \cdot r\end{aligned}$$

Stąd

$$P(W) = \frac{p}{1-r} = \frac{p}{p+q}$$

■

Twierdzenie 5. Niech $\{H_i\}_{i \in I}$ będzie rozbiciem Ω na zdarzenia o dodatnim prawdopodobieństwie. Gdy $P(B) > 0$, to

$$P(A|B) = \sum_{i \in I} P(A|B \cap H_i) P(H_i|B),$$

gdzie zbiór indeksów I jest skończony lub przeliczalny.

Przykład. Grześ i Jaś rzucają na przemian monetą. Jaś wygrywa, gdy pojawią się kolejno OOR, Grześ – gdy ROR. Jakie są prawdopodobieństwa wygranej dla obu chłopców?

Niech
W_1 – wygra Jaś, W_2 – wygra Grześ,
O_k – w k-tym rzucie wypadł orzeł,
R_k – w k-tym rzucie wypadła reszka.

$$\begin{aligned} x &= P(W_1|O_1 \cap O_2) \quad & y &= P(W_1|O_1 \cap R_2) \\ z &= P(W_1|R_1 \cap O_2) \quad & w &= P(W_1|R_1 \cap R_2) \end{aligned}$$

Zatem

$$\begin{aligned} y =& P(W_1|O_1 \cap R_2 \cap O_3)P(O_3|O_1 \cap R_2) + \\ &+ P(W_1|O_1 \cap R_2 \cap R_3)P(R_3|O_1 \cap R_2) \\ =& z\frac{1}{2} + w\frac{1}{2} \end{aligned}$$

Analogicznie

$$x = \frac{1}{2}x + \frac{1}{2} \cdot 1, \; z = \frac{1}{2}x + 0, \; w = \frac{1}{2}w + \frac{1}{2}z$$

Stąd $P(W_1) = (x + y + z + w)/4 = 5/8$. ■

Twierdzenie 6. *Wzór Bayesa.* Niech $\{H_i\}_{i\in I}$ będzie rozbiciem Ω na zdarzenia o dodatnim prawdopodobieństwie i $P(A)>0$, to dla dowolnego $j\in I$ mamy

$$P(H_j|A)=\frac{P(A|H_j)P(H_j)}{\sum_{i\in I}P(A|H_i)P(H_i)}$$

Przykład. Amperomierze pochodzą z trzech taśm produkcyjnych w stosunku 1:1:1. Dostawy z pierwszej taśmy zawierają 0.5% braków, z drugiej 0.7%, a z trzeciej 1%. Wybrany losowo amperomierz okazał się brakiem. Obliczyć prawdopodobieństwo, że został on wyprodukowany na taśmie drugiej.

$A-$amperomierz jest brakiem
H_i-amperomierz pochodzi z $i-$tej taśmy

$$P(H_1)=P(H_2)=P(H_3)=1/3$$

$$P(A|H_1)=0.005;\ P(A|H_2)=0.007;\ P(A|H_3)=0.01$$

Stąd

$$P(A)=\frac{1}{3}(0.005+0.007+0.01)=\frac{0.022}{3}$$

$$P(H_2|A)=\frac{P(H_2)P(A|H_2)}{P(A)}=\frac{\frac{1}{3}\,0.007}{\frac{1}{3}\,0.022}=\frac{7}{22}$$

■

Niezależność zdarzeń.

Zdarzenie B nie zależy od zdarzenia A, gdy wiedza o tym, że zaszło A nie wpływa na prawdopodobieństwo zajścia B.

$$P(B|A) = P(B), \quad P(A) > 0$$
$$\Downarrow$$
$$P(A \cap B) = P(A)P(B)$$

Definicja. Zdarzenia A oraz B nazywamy niezależnymi, gdy

$$P(A \cap B) = P(A)P(B)$$

Definicja. Zdarzenia $A_1,\ A_2, \dots,\ A_n$ nazywamy niezależnymi, gdy

$$P(A_{i1} \cap A_{i2} \cap \ldots \cap A_{ik}) = P(A_{i1}) \ldots P(A_{ik})$$

dla $1 \leq i_i < i_2, \ldots < i_k \leq n,\ k = 2, 3, \dots, n$

Przykład. Spośród rodzin mających n dzieci wybieramy jedną rodzinę. Niech zdarzenie A polega na tym, że w losowo wybranej rodzinie jest co najwyżej jedna dziewczynka, B – w rodzinie są dziewczynki i chłopcy. Czy zdarzenia A i B są niezależne?

Przestrzeń probabilistyczną tworzą ciągi $n-$ elementowe – uporządkowane według starszeństwa dzieci.

$$P(A \cap B) = P(A)P(B) \Leftrightarrow \frac{n}{2^n} = \left(\frac{n+1}{2^n}\right)\left(\frac{2^n - 2}{2^n}\right)$$
$$\Leftrightarrow n = 3$$

■

Przykład. W urnie są cztery kule – niebieska, zielona, czerwona i pstrokata (niebiesko-zielono-czerwona).
Zdarzenia
A_n – wyciągneliśmy kulę z kolorem niebieskim
A_z – wyciągneliśmy kulę z kolorem zielonym
A_n – wyciągneliśmy kulę z kolorem czerwonym
Mamy

$$P(A_n) = P(A_z) = P(A_c) = 1/2$$

$$P(A_n \cap A_z) = P(A_n \cap A_c) = P(A_z \cap A_c) = 1/4$$

Zatem rozważane zdarzenia są parami niezależne.

Zauważmy jednak, że

$$P(A_n \cap A_z \cap A_c) = \frac{1}{4} \neq \frac{1}{8} = P(A_n)P(A_z)P(A_c)$$

■

Przykład. $\Omega = [0,1]^2$, $\mathcal{F} = \mathcal{B}([0,1]^2)$, $P-$ rozkład równomierny na $[0,1]^2$.
Zdarzenia

$$A = B = \{(x,y) \in [0,1]^2 : \ x > y\}$$
$$C = \{(x,y) \in [0,1]^2 : \ x < 0.5\}$$

Zauważmy, że

$$P(A \cap B \cap C) = \frac{1}{8} = P(A)P(B)P(C)$$

natomiast żadne dwa nie są niezależne ■

Przyjmijmy konwencję: $A^0 = A, \ A^1 = A'$

Twierdzenie 7.
Następujące warunki są równoważne:

(i) Zdarzenia $A_1, \ A_2, \ldots, \ A_n$ są niezależne;

(ii) Dla każdego ciągu $\varepsilon_1, \ \varepsilon_2, \ldots, \ \varepsilon_n$, gdzie $\varepsilon_i \in \{0,1\}, \ i = 1,2,\ldots,n$, zdarzenia $A_1^{\varepsilon_1}, \ldots, A_n^{\varepsilon_n}$ są niezależne;

(iii) Dla każdego ciągu $\varepsilon_1, \ \varepsilon_2, \ldots, \ \varepsilon_n$, gdzie $\varepsilon_i \in \{0,1\}, \ i = 1,2,\ldots,n$, zachodzi równość

$$P(A_1^{\varepsilon_1} \cap \ldots \cap A_n^{\varepsilon_n}) = P(A_1^{\varepsilon_1}) \ldots P(A_n^{\varepsilon_n})$$

Dowód. $(i) \Rightarrow (ii)$ (indukcja względem n)

(1^o) Pokażemy dla $n = 2$

(2^o) Założymy, że tw. jest prawdziwe dla $n-1$

(3^o) Pokażemy, że

$$\begin{gathered} A_1, \ldots, A_{n-1}, A_n \text{ niezależne} \\ \Downarrow \\ A_1, \ldots, A_{n-1}, A_n' \text{ niezależne} \end{gathered}$$

(4^o) Zauważymy, że z 3^o wynika

$$A_1^{\varepsilon_1}, \ldots, A_{n-1}^{\varepsilon_{n-1}}, A_n^{\varepsilon_n} \text{ niezależne}$$

Dla $n = 2$

$$\begin{aligned} P(A_1 \cap A_2') &= P(A_1 \setminus A_1 \cap A_2) = \\ &= P(A_1) - P(A_1 \cap A_2) = \\ &= P(A_1)[1 - P(A_2)] = P(A_1)P(A_2') \end{aligned}$$

Zatem $A_1,\ A_2'$ są niezależne. Na mocy symetrii także $A_1',\ A_2$ są niezależne. Stosując jeszcze raz powyższe rozumowanie do $A_1',\ A_2$, otrzymujemy niezależność $A_1',\ A_2'$

Zakładamy, że tw. jest prawdziwe dla $n - 1$ i dowodzimy dla n.

W tym celu wystarczy pokazać:

$$\begin{aligned} &P(A_1 \cap \ldots \cap A_{n-1} \cap A_n') = \\ &\quad = P(A_1 \cap \ldots \cap A_{n-1} \setminus A_1 \cap \ldots \cap A_{n-1} \cap A_n) = \\ &\quad = P(A_1 \cap \ldots \cap A_{n-1}) - P(A_1 \cap \ldots \cap A_n) = \\ &\quad = P(A_1 \cap \ldots \cap A_{n-1})[1 - P(A_n)] \\ &\quad = P(A_1) \ldots P(A_{n-1})P(A_n') \end{aligned}$$

■

Definicja. Zdarzenia $A_1, A_2, \ldots$ nazywamy niezależnymi, gdy dla każdego n zdarzenia $A_1,\ A_2, \ldots ,\ A_n$ są niezależne.

Zmienne losowe.

Cel: Ujednolicić sposób rozważań dla różnych przestrzeni zdarzeń elementarnych.

Definicja. Zmienna losowa jest to funkcja rzeczywista

$$X : \Omega \to \mathcal{X}$$

o własności:

$$\bigwedge_{x \in \mathbf{R}} \{\omega \in \Omega : \ X(\omega) \leq x\} \in \mathcal{F}$$

$\mathcal{X}$ – zbiór wartości zmiennej losowej

Często
$\mathcal{X} = \{0, 1, \ldots\}$, $\mathcal{X} = [0, \infty)$, $\mathcal{X} = [a, b]$, $\mathcal{X} = R$

Definicja. Rozkładem prawdopodobieństwa zmiennej losowej X nazywamy rozkład prawdopodobieństwa P_X określony wzorem

$$\begin{aligned} P_X(A) &= P(\{\omega \in \Omega : \ X(\omega) \in A\}) \quad \text{dla } A \subset \mathcal{X} \\ &= P(X^{-1}(A)) \end{aligned}$$

! dokładniej dla $A \in \mathcal{B}(\mathcal{X})$

Definicja. Trójkę $(\mathcal{X}, \mathcal{B}(\mathcal{X}), P_X)$ nazywamy modelem probabilistycznym.

Przykład. Ze zbioru pięciu ponumerowanych elementów losujemy jeden element

$$\Omega = \{\omega_1, \ldots, \omega_5\} \quad P(\{\omega_i\} = 1/5)$$

ω_i–wylosowano i–ty element

Wtedy dla $X(\omega_i) = i$ mamy $\mathcal{X} = \{1, 2, 3, 4, 5\}$ oraz

$$\begin{gathered} P_X(i) = 1/5, \quad i = 1, 2, 3, 4, 5 \\ P_X(A) = \sum_{i \in A} P_X(i), \quad \text{dla } A \subset \mathcal{X} \end{gathered}$$

■

Definicja. Dystrybuanta zmiennej losowej X, jest to funkcj $F : \mathbf{R} \to [0,1]$ określona wzorem

$$F_X(x) = P(X \leq x)$$

Własności dystrybuanty

W1. F jest niemalejąca

$$x_1 < x_2,\ A = (-\infty, x_1],\ B = (-\infty, x_2],\ A \subset B$$

$$F(x_1) = P(A) \leqslant P(B) = F(x_2)$$

W2. $\lim\limits_{x\to-\infty} F(x) = 0, \quad \lim\limits_{x\to\infty} F(x) = 1$

$$\{x_n\} \nearrow \infty$$

$$\begin{aligned}\lim_{x\to\infty} F(x) = \lim_{n\to\infty} F(x_n) &= P\left(\bigcup_n (-\infty, x_n]\right)\\ &= P((-\infty,\infty)) = 1\end{aligned}$$

$$\{x_n\} \searrow \infty$$

$$\lim_{x\to-\infty} F(x) = \lim_{n\to\infty} F(x_n) = P\left(\bigcap_n (-\infty, x_n]\right)$$
$$= P(\emptyset) = 0.$$

W3. F jest prawostronnie ciągła

$$\{x_n\} \searrow x_0$$

$$\lim_{x\to x_0^+} F(x) = \lim_{n\to\infty} F(x_n) = P\left(\bigcap_n (-\infty, x_n]\right)$$
$$= P((-\infty, x_0])$$
$$= F(x_0),$$

Twierdzenie 8. Każda funkcja $F : \mathbf{R} \to [0, 1]$ o własnościach 1–3 jest dystrybuantą pewnej zmiennej losowej.

Dowód.

$$F^{-1}(u) := \inf\{x : \ F(x) \geq u\} \quad \text{dla } 0 < u < 1$$

$$F^{-1}(u) \leq x \quad \Leftrightarrow \quad u \leq F(x)$$

Niech U oznacza zmienną losową o rozkładzie równomiernym na zbiorze $(0, 1)$:

$$F_U(u) = P(U \leq u) = u$$

Niech $X = F^{-1}(U)$.

$$\begin{aligned} F_X(x) = P(X \leq x) = P(F^{-1}(U) \leq x) \\ = P(U \leq F(x)) = F(x) \end{aligned}$$

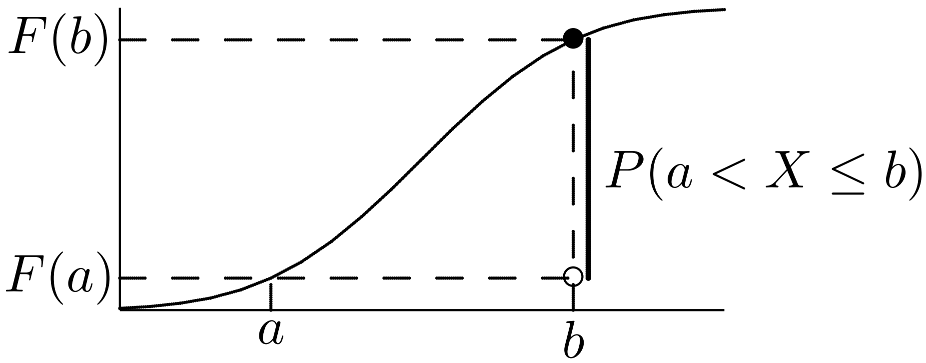
■

Własności dystrybuanty, ciąg dalszy

oznaczmy $F(a+) := \lim\limits_{x \to a^+} F(x)$

(?)

$$\begin{aligned} &\text{(i)} \quad P(a < X \le b) = F(b) - F(a) \\ &\text{(ii)} \quad P(X = a) = F(a) - F(a-) \\ &\text{(iii)} \quad P(a \le X \le b) = F(b) - F(a-) \\ &\text{(iv)} \quad P(a < X < b) = F(b-) - F(a) \end{aligned}$$



Zmienne losowe typu skokowego

Definicja. Mówimy, że zmienna losowa jest typu skokowego (dyskretna), jeżeli istnieje zbiór skończony lub przeliczalny $\mathcal{X} \subset \mathbf{R}$ taki, że

$$P_X(\mathcal{X}) = 1$$

Przykłady zmiennych losowych
typu skokowego:

- *rozkład dwumianowy*
- *rozkład Poissona*
- *rozkład ujemny dwumianowy*
- *rozkład wielomianowy*

Rozkład dwumianowy

Powtarzające się i niezależne próby nazywamy **próbami Bernoulliego**, jeżeli każda próba ma tylko dwa możliwe wyniki: *„sukces”* z prawdopodobieństwem p oraz *„porażka”* z prawdopodobieństwem q

Niech X oznacza ilość sukcesów osiągniętych w ciągu n prób Bernoulliego.

Zmienna losowa X ma następujący rozkład prawdopodobieństwa:

$$P(X = k) = \binom{n}{k} p^k q^{n-k},$$

gdzie $p \in (0, 1)$ oraz $k = 0, 1, \ldots, n$.

O zmiennej losowej X mówimy, że ma *rozkład dwumianowy* ($X \sim B(n, p)$).

Przykład. Dziesięciu robotników używa z przerwami energię elektryczną. Jakiego należy oczekiwać obciążenia, jeżeli

1. W każdej danej chwili każdy robotnik ma to samo prawdopodobieństwo p zapotrzebowania na jednostkę energii.
2. Robotnicy pracują niezależnie od siebie.
3. Przeciętnie jeden robotnik używa dostarczanej energii w ciągu 12 minut na godzinę.

Niech X oznacza liczbę robotników, którzy potrzebują energii w tym samym czasie.

$$X \sim B(10, 1/5).$$

Wówczas, jeżeli dopływ energii jest ustalony na poziomie sześciu jednostek, to przeciążenie ma szanse:

$$P(X \geqslant 7) = 0.0008643584$$

(?) W ciągu 20 godzin powinno trwać łącznie przez około minutę.

■

Rozkład Poissona

Zmienna losowa X *ma rozkład Poissona* z parametrem $\lambda > 0$ ($X \sim P_0(\lambda)$), jeżeli:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, \ldots$$

Rozkład Poissona a rozkład dwumianowy.

Załóżmy, że liczba doświadczeń n w poszczególnych seriach schematu Bernoulliego wzrasta dążąc do nieskończoności a prawdopodobieństwo p dąży do zera tak, że iloczyn np jest wielkością stałą równą $\lambda > 0$. Wtedy zachodzi

$$\lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} = \frac{e^{-\lambda}\lambda^k}{k!}.$$

Wynika to z rozpisania: $\binom{n}{k} p^k (1-p)^{n-k} =$

$$= \frac{1}{k!}(n-k+1)(n-k+2)\ldots n \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$
$$= \frac{\lambda^k}{k!}\left(1-\frac{k-1}{n}\right)\left(1-\frac{k-2}{n}\right)\ldots$$
$$\ldots\left(1-\frac{1}{n}\right)1\left(1-\frac{\lambda}{n}\right)^{n-k}$$

Przykład. Jakie jest prawdopodobieństwo p_k, że wśród 500 ludzi dokładnie k będzie miało urodziny w dniu Nowego Roku?

Jeżeli 500 ludzi zostało wybranych losowo, to możemy zastosować schemat 500 prób Bernoulliego z prawdopodobieństwem sukcesu 1/365. Wówczas

$$p_0 = (364/365)^{500} = 0.2537\ldots$$

Dla przybliżenia Poissona bierzemy

$$\lambda = 500/365 = 1.3699\ldots$$

Wtedy

$$p_0 \approx \frac{e^{-1.3699} 1.3699^0}{0!} \approx 0.2541$$

$\blacksquare$

Ujemny rozkład dwumianowy.

Prowadzimy doświadczenia według schematu Bernoulliego do momentu pojawienia się $r-$tego sukcesu. Niech X oznacza liczbę porażek poprzedzających $r-$ty sukces.

$$P(X=k)=\binom{r+k-1}{k}p^{r-1}q^k\cdot p=\binom{r+k-1}{k}p^r q^k$$

gdzie $q=1-p$, $k=0,1,\ldots$

O zmiennej losowej X mówimy, że ma *ujemny rozkład dwumianowy* ($X\sim f(r,p)$). Zakładamy, że $r>0$ oraz $0<p<1$.

Uwaga. Możemy przyjąć, że $r>0$ nie musi być liczbą całkowitą. Wtedy przyjmujemy następującą definicję **symbolu Newtona** (dla $a\in\mathbf{R}$ oraz $k\geqslant 0$)

$$\binom{a}{k}:=\begin{cases}\dfrac{a(a-1)(a-2)\ldots(a-(k-1))}{k!} & ,k\in\mathbf{N}\\ 1 & ,k=0\\ 0 & ,k\notin\mathbf{Z}\end{cases}$$

Przykład. *Zadanie Banacha o pudełkach zapałek.* Mamy dwa pudełka zapałek – jedno w prawej kieszeni i jedno w lewej. Kiedy potrzebujemy zapałkę, wybieramy jedną z kieszeni losowo. Przypuśćmy, że początkowo każde z pudełek zawiera N zapałek. Ile wynosi prawdopodobieństwo, że gdy wyciągniemy puste pudełko, w drugim będzie dokładnie m zapałek.

X – liczba wyciągnięć pudełka z prawej kieszeni do momentu aż w drugim pudełku będzie m zapałek

Y – ... z lewej kieszeni ...

$$X \sim f(N-m, 0.5), \quad Y \sim f(N-m, 0.5)$$

Poszukiwane prawdopodobieństwo wynosi

$$\begin{aligned} P(\{X = N+1\} \cup \{Y = N+1\}) = \\ = P(X = N+1) + P(Y = N+1) \end{aligned}$$

■

Rozkład wielomianowy

uogólnienie rozkładu dwumianowego

Wykonujemy serię n niezależnych prób. Każda próba może mieć jeden z kilku wyników, np. $E1, E2, \ldots, E_r$.

Prawdopodobieństwo realizacji E_i w każdej próbie wynosi p_i, $i = 1, 2, \ldots, r$.

Prawdopodobieństwo, że w n próbach E_1 występuje k_1 razy, E_2 występuje k_2 razy itd. wynosi

$$\frac{n!}{k_1!k_2!\ldots k_r!} p_1^{k_1} p_2^{k_2} \ldots p_r^{k_r}$$

Zmienne losowe typu ciągłego

Definicja. Mówimy, że zmienna losowa o dystrybuancie F jest typu ciągłego, jeżeli istnieje taka funkcja $f \geq 0$, że dla każdego x zachodzi równość

$$F(x) = \int\limits_{-\infty}^{x} f(u)\, du$$

Funkcję f nazywamy gęstością prawdopodobieństwa zmiennej losowej X lub w skrócie **gęstością**

Uwagi

(1) W punktach, w których f jest ciągła zachodzi

$$\frac{d}{dx} F(x) = f(x)$$

(2) $\int_{-\infty}^{\infty} f(x)\, dx = 1$

(3) Każda funkcja f nieujemna i spełniająca (2) wyznacza dystrybuantę F za pomocą wzoru

$$F(x) = \int\limits_{-\infty}^{x} f(u)\, du$$

Przykład. Sprawdzić, czy funkcja f określona wzorem

$$f(x) = \begin{cases} 0 & \text{dla } x < 0 \\ e^{-x} & \text{dla } x \geq 0 \end{cases}$$

jest gęstością.

$$\int\limits_{-\infty}^{\infty} f(x)\,dx = \int\limits_{0}^{\infty} e^{-x}\,dx = \left[-e^{-x}\right]_0^{\infty} = 1$$

■

Przykłady zmiennych losowych ciągłych:

- *rozkład normalny* $N(\mu, \sigma^2)$
- *rozkład jednostajny* $U(a, b)$
- *rozkład gamma* $G(b, p)$
- *rozkład beta* $B(p, q)$
- *rozkład Cauchyego* $C(\mu, \lambda)$

$N(\mu, \sigma^2),\ \ \sigma > 0$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$$

$U(a, b),\ \ a < b$

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b] \end{cases}$$

$G(b, p),\ \ b > 0,\ \ p > 0$

$$f(x) = \begin{cases} \dfrac{b^p}{\Gamma(p)} x^{p-1} e^{-bx}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

gdzie

$$\Gamma(p) = \int\limits_0^{\infty} x^{p-1} e^{-x}\, dx$$

$B(p,q),\ b>0,\ p>0$

$$f(x) = \begin{cases} \dfrac{1}{B(p,q)} x^{p-1}(1-x)^{q-1}, & x \in (0,1) \\ 0, & x \notin (0,1) \end{cases}$$

gdzie

$$B(p,q) = \int\limits_0^1 x^{p-1}(1-x)^{q-1}\,dx$$

a także

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

$C(\mu,\lambda),\ \lambda>0$

$$f(x) = \frac{1}{\pi}\frac{\lambda}{\lambda^2+(x-\mu)^2}$$

Przykład. Sprawdzimy, że rozkład $N(\mu, \sigma^2)$ jest rzeczywiście rozkładem prawdopodobieństwa:

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] dx =$$

podstawienie: $y = (x-\mu)/\sigma$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy =$$

Należy zatem sprawdzić, że ostatnia całka równa jest $\sqrt{2\pi}$. Ponadto zauważmy, że przy okazji otrzymaliśmy następujący fakt

$$X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$\left(\int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right)\, dy\right)^2 =$$

$$= \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right)\, dx \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right)\, dy =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2+y^2}{2}\right)\, dx\, dy =$$

przejście na współrzędne biegunowe:

$\varphi(r,t) = (r\cos t, r\sin t)$

$J\varphi(r,t) = \begin{vmatrix} \cos(t) & -r\sin(t) \\ \sin(t) & r\cos(t) \end{vmatrix} = r$

$$= \int_0^{\infty} \int_0^{2\pi} \exp\left(-\frac{r^2}{2}\right) r\, dr\, dt = 2\pi \int_0^{\infty} \exp\left(-\frac{r^2}{2}\right) r\, dr =$$

$$= 2\pi \left[-\exp\frac{-r^2}{2}\right]_0^{\infty} = 2\pi$$

$\blacksquare$

Funkcje zmiennej losowej

Przykład. Niech $Y = aX + b$, gdzie $a \neq 0$ oraz X jest zmienną losową o rozkładzie

$$P(X = 0) = 1/4,\ P(X = 1) = 3/4.$$

Chcemy znaleźć rozkład zmiennej losowej Y.

$$\begin{aligned} P(X = 0) &= P(Y = b) = 1/4 \\ P(X = 1) &= P(Y = a + b) = 3/4 \end{aligned}$$

■

Przykład. Niech X będzie zmienną losową typu ciągłego o gęstości f_X, dystrybuancie F_X oraz niech $Y = aX + b,\ a < 0$. Chcemy znaleźć rozkład Y

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X \geq \frac{y-b}{a}) = \\ &= 1 - P\left(X < \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right) \end{aligned}$$

Zatem $f_Y(y) = \dfrac{d}{dy} F_Y(y) = -\dfrac{1}{a} f_X\left(\dfrac{y-b}{a}\right)$ ■

Przykład. Niech X oznacza zmienną losową ciągłą o dystrybuancie F_X oraz gęstości f_X. Niech f_X jest funkcją ciągłą, a g funkcją ściśle monotoniczną oraz niech $h = g^{-1}$. Wtedy dystrybuantą zmiennej losowej $Y = g(X)$ jest:

(dla g - rosnącej)

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P(X \leq h(y)) = F_X(h(y)) \end{aligned}$$

Jeżeli h jest funkcją różniczkowalną, to

$$\frac{d}{dy} F_Y(y) = f_X(h(y))h'(y)$$

jest gęstością zmiennej losowej $Y = g(X)$

(dla g - malejącej)

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P(X \geq h(y)) = 1 - F_X(h(y)) \end{aligned}$$

Jeżeli h jest funkcją różniczkowalną, to

$$\frac{d}{dy} F_Y(y) = f_X(h(y))(-h'(y))$$

jest gęstością zmiennej losowej $Y = g(X)$

Zatem w obu przypadkach

$$f_Y(y) = f_X(h(y))|h'(y)|$$

$\blacksquare$

Przykład. Niech X – nieujemna zmienna losowa typu ciągłego oraz $Y = \sqrt{X}$. Zatem $h(y) = y^2$ oraz

$$f_Y(y) = 2y \cdot f_X(y^2) \cdot I_{(0,\infty)}(y)$$

Uwaga. $I_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$ $\blacksquare$

Przykład. Niech X – zmienna losowa typu ciągłego oraz $Y = X^2$.

$$\begin{aligned} F_Y(y) &= P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \\ f_Y(y) &= \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left(F_X(\sqrt{y}) - F(-\sqrt{y})\right) = \\ &= \frac{1}{2\sqrt{y}}\left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right) \end{aligned}$$

$\blacksquare$

Twierdzenie 9.

Niech X będzie zmienną losową typu ciągłego. Niech g będzie funkcją określoną na zbiorze

$$\bigcup_{k=1}^{n} [a_k, b_k],$$

która na każdym przedziale otwartym (a_k, b_k) jest funkcją ściśle monotoniczną oraz ma ciągłą pochodną $g(x)' \neq 0$. Niech $h_k(y)$ będzie funkcją odwrotną do funkcji $g(x)$ na przedziale

$$I_k = g((a_k, b_k)) = \{y : \ x \in (a_k, b_k), g(x) = y\}.$$

Wówczas funkcja gęstości zmiennej losowej $Y = g(X)$ ma następującą postać

$$f_Y(y) = \sum_{k=1}^{n} f_X(h_k(y)) \cdot |h'(y)| \cdot I_{I_k}(y)$$

Przykład. $X-$ ciągła, $Y = X^2$. Wtedy $g(x) = x^2$, $h_1(y) = -\sqrt{y}$, $h_2(y) = \sqrt{y}$, $I_1 = I_2 = (0, \infty)$. ■

Dowód. Niech $A \in \mathcal{B}(\mathbf{R})$

$$\begin{aligned}
P(Y \in A) &= P(g(X) \in A) = P(X \in g^{-1}(A)) \\
&= \sum_{k=1}^{n} P(X \in (a_k, b_k) \cap g^{-1}(A)) \\
&= \sum_{k=1}^{n} P(X \in g^{-1}(I_k) \cap g^{-1}(A)) \\
&= \sum_{k=1}^{n} P(X \in g^{-1}(I_k \cap A)) \\
&= \sum_{k=1}^{n} P(X \in h_k(I_k \cap A)) \\
&= \sum_{k=1}^{n} \int\limits_{h_k(I_k \cap A)} f_X(x)\, dx \\
&= \sum_{k=1}^{n} \int\limits_{I_k \cap A} f_X(h_k(y)) \cdot |h_k'(y)|\, dy \\
&= \int\limits_{A} \sum_{k=1}^{n} f_X(h_k(y)) \cdot |h_k'(y)| \cdot I_{I_k}\, dy
\end{aligned}$$

Pytanie: Czy coś by się zmieniło, gdyby $n = \infty$? $\blacksquare$

Wektory losowe

Definicja. Wektor losowy $\boldsymbol{X} = (X_1, \ldots, X_n)$ to odwzorowanie

$$\boldsymbol{X} : \Omega \to \mathcal{X} \subseteq \mathbf{R}^n$$

o własności:

$$\{\omega \in \Omega : \; X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n\} \in \mathcal{F}$$

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

$\mathcal{X}$ – zbiór wartości wektora losowego

Często

$\mathcal{X} = \{0, 1, \ldots\}^n$, $\mathcal{X} = [0, \infty)^n$, $\mathcal{X} = [a, b]^n$, $\mathcal{X} = R^n$

Definicja. Rozkładem prawdopodobieństwa wektora losowego $\boldsymbol{X}$ nazywamy rozkład prawdopodobieństwa $P_{\boldsymbol{X}}$ określony wzorem

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

Definicja. Trójkę $(\mathcal{X}, \mathcal{B}(\mathcal{X}), P_{\boldsymbol{X}})$ nazywamy modelem probabilistycznym.

Definicja. Funkcja $F_{\boldsymbol{X}} : \mathbf{R}^n \to [0, 1]$ postaci

$$F_{\boldsymbol{X}}(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n)$$

nazywamy dystrybuantą wektora losowego $\boldsymbol{X}$

Definicja. Wektor losowy jest typu skokowego, jeżeli istnieje zbiór przeliczalny $\mathcal{X} \subset \mathbf{R}^n$, taki że $P_{\boldsymbol{X}}(\mathcal{X}) = 1$

Definicja. Wektor losowy jest typu ciągłego, jeżeli istnieje nieujemna funkcja $f_{\boldsymbol{X}}(x_1,\ x_2, \ldots,\ x_n)$, zwana gęstością, taka że dla każdego $\boldsymbol{x} = (x_1,\ x_2, \ldots,\ x_n) \in \mathbf{R}^n$

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \int\limits_{-\infty}^{x_1} \cdots \int\limits_{-\infty}^{x_2} f_{\boldsymbol{X}}(u_1, \ldots, u_n)\, du_1 \ldots du_n$$

Uwagi

Prawie wszędzie ma miejsce równość

$$\frac{\partial F_{\boldsymbol{X}}(x_1, \dots, x_n)}{\partial x_1, \dots, \partial x_n} = f_{\boldsymbol{X}}(x_1, \dots, x_n)$$

Dla dowolnego $A \in \mathcal{B}(\mathbf{R}^n)$ zachodzi

$$\int\limits_A f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}$$

Zauważmy, że

$$\begin{aligned}
P(X_1 \in A) &= P(X_1 \in A, X_2 \in \mathbf{R}, \dots, X_n \in \mathbf{R}) \\
&= \int\limits_A \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1, \dots, x_n)\, dx_1 \dots dx_n \\
&= \int\limits_A \left(\int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1, \dots, x_n)\, dx_2 \dots dx_n \right) dx_1
\end{aligned}$$

Zatem

$$f_{X_1}(x_1) = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1, \dots, x_n)\, dx_2 \dots dx_n$$

Jest to tzw. *brzegowa gęstość prawdopodobieństwa.*

Dla rozkładów brzegowych wielowymiarowych mamy:

$$f_{(X_1,X_2)}(x_1,x_2) = \\ = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1,\ldots,x_n)\,dx_3\ldots dx_n$$

$$f_{(X_1,X_2,X_3)}(x_1,x_2,x_3) = \\ = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1,\ldots,x_n)\,dx_4\ldots dx_n$$

itd.

Podobnie postępuje się przy rozkładach skokowych:

Przykład. Niech wektor losowy (X, Y) ma rozkład określony liczbami

$$p_{ik} = P(X = x_i, Y = y_k), \text{ gdzie } i \in I, \ k \in K.$$

Wówczas rozkład zmiennej losowej X określają liczby

$$p_i = P(X = x_i) = \sum_{k \in K} p_{ik}, \text{ gdzie } i \in I$$

■

Przykład.

Niech (X, Y) ma rozkład równomierny na $\Omega = [0,2] \times [0,3]$:

$$f_{X,Y}(x,y) = \frac{1}{6} I_\Omega(x,y).$$

Wówczas

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \frac{1}{6}\int_{-\infty}^{\infty} I_\Omega(x,y)\,dy =$$

$$= \frac{1}{6}\int_{-\infty}^{\infty} I_{[0,2]}(x) \cdot I_{[0,3]}(y)\,dy =$$

$$= \frac{1}{6} I_{[0,2]}(x) \int_{-\infty}^{\infty} I_{[0,3]}(y) = \frac{1}{2} I_{[0,2]}(x)$$

■

Przykład. Niech (X_1, X_2) ma dwuwymiarowy rozkład normalny, tzn:

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2(1-\varrho^2)^{\frac{1}{2}}} \cdot$$
$$\cdot \exp\left\{-\frac{1}{2(1-\varrho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \right.\right.$$
$$\left.\left. + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\varrho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right]\right\}$$

gdzie $\sigma_1, \sigma_2 > 0$ oraz $\varrho \in (-1, 1)$

Rozpisujemy wyrażenie w nawiasie kwadratowym:

$$\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\varrho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} =$$
$$= \left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2$$
$$+ \varrho^2\frac{(x_2-\mu_2)^2}{\sigma_2^2} - \varrho^2\frac{(x_2-\mu_2)^2}{\sigma_2^2}$$
$$- 2\varrho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} =$$

$$= \left(\frac{x_1-\mu_1}{\sigma_1} - \varrho\frac{x_2-\mu_2}{\sigma_2}\right)^2 + (1-\varrho^2)\frac{(x_2-\mu_2)^2}{\sigma_2^2} =$$

$$= \frac{1}{\sigma_1^2}\left(x_1-\mu_1-\varrho\frac{\sigma_1}{\sigma_2}(x_2-\mu_2)\right)^2 + (1-\varrho^2)\frac{(x_2-\mu_2)^2}{\sigma_2^2}$$

Zatem wyrażenie w nawiasie klamrowym ma postać:

$$\overbrace{-\frac{1}{2(1-\varrho^2)\sigma_1^2}\left(x_1-\mu_1-\varrho\frac{\sigma_1}{\sigma_2}(x_2-\mu_2)\right)^2}^{h(x_1,x_2)} - \\ - \frac{1}{2\sigma_2^2}(x_2-\mu_2)^2$$

Zatem

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2(1-\varrho^2)^{\frac{1}{2}}} \cdot \\ \cdot \exp\left\{h(x_1,x_2) - \frac{1}{2\sigma_2^2}(x_2-\mu_2)^2\right\}$$

Zauważmy, że

$$g(x_1) := \frac{1}{\sqrt{2\pi(1-\varrho^2)}\sigma_1} \exp(h(x_1, x_2))$$

jest gęstością rozkładu

$$N\left(\mu_1 + \varrho\frac{\sigma_1}{\sigma_2}(x_2-\mu_2), (1-\varrho^2)\sigma_1^2\right)$$

Zatem

$$f_{X_2}(x_2) = \int\limits_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)\, dx_1 =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{\frac{1}{2\sigma_2^2}(x_2-\mu_2)^2\right\} \underbrace{\int\limits_{-\infty}^{\infty} g(x_1)\, dx_1}_{=1}$$

Wniosek: Rozkład brzegowy dwuwymiarowego rozkładu normalnego jest jednowymiarowym rozkładem normalnym ■

Niezależność zmiennych losowych

Definicja. Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną, a $X_1,\ X_2, \dots ,\ X_n$ będą zmiennymi losowymi określonymi na tej przestrzeni. Mówimy, że te zmienne losowe są niezależne, jeżeli dla dowolnych zbiorów borelowskich $A_1,\ A_2, \dots ,\ A_n$ zachodzi:

$$P(X_1 \in A_1, \dots X_n, \in A_n) = \\ = P(X_1 \in A_1) \dots P(X_n \in A_n)$$

Definicja. Mówimy, że zmienne losowe $X_1,\ X_2, \dots$ są niezależne, jeżeli każdy skończony podciąg ciągu $X_1,\ X_2, \dots$ składa się z niezależnych zmiennych losowych

Twierdzenie 10.
Dla zmiennych losowych $X_1,\ X_2, \dots ,\ X_n$ następujące warunki są równoważne

(i) zmienne losowe są niezależne

(ii) dla $\boldsymbol{x} = (x_1,\ x_2, \dots ,\ x_n) \in R^n$

$$F_{\mathbf{X}}(\boldsymbol{x}) = F_{X_1}(x_1) \dots F_{X_n}(x_n)$$

Twierdzenie 11. Jeżeli $\boldsymbol{X} = (X_1,\ X_2, \dots,\ X_n)$ jest wektorem losowym typu skokowego to warunkiem koniecznym i wystarczającym niezależności zmiennych losowych $X_1,\ X_2, \dots,\ X_n$ jest:

$$P(X_1 = x_1, \dots, X_n = x_n) = \\ = P_1(X_1 = x_1) \dots P_n(X_n = x_n),$$

dla każdego $(x_1, \dots, x_n) \in \mathbf{R}^n$, gdzie P_k oznacza brzegowy rozkład prawdopodobieństwa zmiennej losowej X_k $(k = 1, 2, \dots, n)$.

Twierdzenie 12. Jeżeli $\boldsymbol{X} = (X_1,\ X_2, \dots,\ X_n)$ jest wektorem losowym typu ciągłego o gęstości $f_{\boldsymbol{X}}$, to warunkiem koniecznym i wystarczającym niezależności zmiennych losowych $X_1,\ X_2, \dots,\ X_n$ jest:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = f_{X_1}(x_1) \dots f_{X_n}(x_n),$$

dla każdego $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbf{R}^n$, gdzie f_{X_k} jest gęstością rozkładu brzegowego zmiennej losowej X_k $(k = 1, \dots, n\ )$

Przykład. Niech X_1, X_2 ma łączny rozkład normalny. Chcemy znaleźć warunek konieczny i wystarczający na niezależność zmiennych X_1 oraz X_2. Z twierdzenia mamy, że powinno zachodzić

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$$

Ponieważ

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2(1-\varrho^2)^{\frac{1}{2}}} \cdot$$
$$\cdot \exp\left\{-\frac{1}{2(1-\varrho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \right.\right.$$
$$\left.\left. + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\varrho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right]\right\}$$

oraz

$$f_{X_1}(x_1) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[\frac{-(x_1-\mu_1)^2}{2\sigma_1^2}\right]$$

$$f_{X_2}(x_2) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left[\frac{-(x_2-\mu_2)^2}{2\sigma_2^2}\right]$$

zauważamy, że warunkiem tym jest $\varrho = 0$ ■

Przykład. Niech $X_i \sim N(\mu_i, \sigma_i^2),\ i = 1, 2, \cdots, n.$

Wówczas

$$\prod_{i=1}^{n} f_{X_i}(x_i) =$$

$$= \frac{1}{\prod\limits_{i=1}^{n}(\sigma_i\sqrt{2\pi})} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu_i)^2}{\sigma_i^2}\right] =$$

$$= \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})'\right],$$

gdzie $\boldsymbol{x} = (x_1,\ x_2, \ldots,\ x_n)$, $\boldsymbol{\mu} = (\mu_1,\ \mu_2, \ldots,\ \mu_n)$ oraz

$$\Sigma = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

Wniosek: Jeżeli $\boldsymbol{X} = (X_1,\ X_2, \ldots,\ X_n) \sim N_n(\boldsymbol{\mu}, \Sigma)$, to warunkiem koniecznym i dostatecznym niezależności zmiennych losowych $X_i,\ i = 1, 2, \ldots, n$ jest to, aby macierz Σ była diagonalna. ■

Twierdzenie 13.

(a) Jeżeli zmienne losowe $X_1,\ X_2,\dots,\ X_n$ są niezależne oraz $g_1,\ g_2,\dots,\ g_n$ są funkcjami borelowskimi, to zmienne losowe

$$Y_1 = g_1(X_1), \dots, Y_n = g_n(X_n)$$

są również niezależne.

(b) Jeżeli $X_1, \dots, X_m, Y_1, \dots, Y_n$ są niezależnymi zmiennymi losowymi oraz

$$f : \mathbf{R}^m \to \mathbf{R} \quad \text{i} \quad g : \mathbf{R}^n \to \mathbf{R}$$

są funkcjami borelowskimi, to

$$U = f(X_1, \dots, X_m) \quad \text{i} \quad V = g(Y_1, \dots, Y_n)$$

są niezależnymi zmiennymi losowymi, a także

$$U, Y_1, \dots, Y_n$$

są niezależnymi zmiennymi losowymi.

Przykład. Niech $X_i \sim N(0,1),\ i = 1,2$ będą zmiennymi niezależnymi.

Chcemy znaleźć rozkład zmiennej losowej $X_1^2 + X_2^2$.

Ponieważ zmienne X_1, X_2 są niezależne, to zmienne $Y_1 = X_1^2, Y_2 = X_2^2$ też są niezależne. Zatem

$$f_{Y_1,Y_2}(y_1, y_2) = f_{Y_1}(y_1) f_{Y_2}(y_2)$$

Ponieważ

$$f_{Y_i}(y_i) = \frac{1}{2\sqrt{y_i}} \left(f_{X_i}(\sqrt{y_i}) + f_{X_i}(-\sqrt{y_i}) \right) I_{(0,\infty)}(y_i)$$

oraz

$$f_{X_i}(x_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-x_i^2}{2}\right]$$

mamy

$$f_{Y_i}(y_i) = \frac{1}{\sqrt{2\pi y_i}} \exp\left[-\frac{y_i}{2}\right] I_{(0,\infty)}(y_i)$$

Niech $Z = X_1^2 + X_2^2 = Y_1 + Y_2$.

$$F_Z(z) = P(Y_1 + Y_2 \leq z) = \int\limits_{Y_1+Y_2 \leq z} f_{Y_1,Y_2}(y_1, y_2)\, dy_1 dy_2 =$$

$$= \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{z-y_2} f_{Y_1,Y_2}(y_1, y_2)\, dy_1 \right) dy_2 =$$

$$= \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{z-y_2} f_{Y_1}(y_1)\, dy_1 \right) f_{Y_2}(y_2) dy_2 =$$

$$= \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{z} f_{Y_1}(y_1 - y_2)\, dy_1 \right) f_{Y_2}(y_2) dy_2 =$$

$$= \int\limits_{-\infty}^{z} \left(\int\limits_{-\infty}^{\infty} f_{Y_1}(y_1 - y_2) f_{Y_2}(y_2)\, dy_2 \right) dy_1$$

Zmiana oznaczeń dla funkcji w nawiasach:

$$z := y_1,\ x := y_2$$

Zatem $f_Z(z) = \int\limits_{-\infty}^{\infty} f_{Y_1}(z - x) f_{Y_2}(x)\, dx$

Robimy odpowiednie podstawienie i otrzymujemy dla $z > 0$:

$$f_Z(z) = \frac{1}{2\pi} \int\limits_0^z \frac{1}{\sqrt{(z-x)x}} \exp\left[-\frac{z-x+x}{2}\right] dx =$$

$$= \frac{1}{2\pi} \exp\left[-\frac{z}{2}\right] \int\limits_0^z \frac{1}{\sqrt{(z-x)x}}\, dx =$$

podstawienie $t := x/z$

$$= \frac{1}{2\pi} \exp\left[-\frac{z}{2}\right] \int\limits_0^1 t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\, dt =$$

$$= \frac{1}{2\pi} \exp\left[-\frac{z}{2}\right] B(1/2, 1/2) =$$

$$= \frac{1}{2\pi} \exp\left[-\frac{z}{2}\right] \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2}+\frac{1}{2})} =$$

$$= \frac{1}{2} \exp\left[-\frac{z}{2}\right]$$

Zatem

$$f_Z(z) = \frac{1}{2} \exp\left[-\frac{z}{2}\right] I_{(0,\infty)}(z)$$

Można pokazać przez indukcję ze względu na n, że zmienna losowa $Z = X_1^2 + \ldots + X_n^2$ ma rozkład o gęstości

$$f_Z(z) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1} e^{-z/2} I_{(0,\infty)}(z)$$

Jest to tzw. rozkład chi-kwadrat o n stopniach swobody. Symbolicznie piszemy

$$X_1^2 + \ldots + X_n^2 \sim \chi^2(n)$$

Fakt.
$\sum\limits_{i=1}^{m} \lambda_i X_i^2 \sim \chi^2(n),\ X_i^2 \sim \chi^2(1) \quad \Rightarrow \quad m = n, \lambda_i = 1$ ■

Parametry rozkładów

Wartość oczekiwaną (wartość przeciętna, nadzieję matematyczną) zmiennej losowej X oznaczamy symbolem $E(X)$ i określamy w następujący sposób:

Dla zmiennej losowej skokowej

Jeżeli X jest zmienną losową typu skokowego, $\mathcal{X} = \{x_1, x_2, \ldots\}$, przy czym szereg

$$\sum_k |x_k| P(X = x_k)$$

jest zbieżny, to

$$E(X) = \sum_k x_k P(X = x_k)$$

Dla zmiennej losowej ciągłej

Jeżeli X jest zmienną losową typu ciągłego o gęstości f i zbieżna jest całka

$$\int_{\mathbf{R}} |x| f(x)\, dx,$$

to

$$E(X) = \int_{\mathbf{R}} x\, f(x)\, dx$$

Ogólnie: $E(X) = \int_\Omega X(\omega) dP(\omega)$

Przykład. Niech $\mathcal{X} = \{0, 1\}$, $P(X = 0) = q$, $P(X = 1) = p = 1 - q$. Wówczas

$$E(X) = 0 \cdot q + 1 \cdot p = p$$

$\blacksquare$

Przykład. Niech $X \sim B(n, p)$. Wówczas

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = \\ &= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^{k-1} q^{n-k} = \\ &= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} q^{n-k} = \\ &= np\,(p+q)^{n-1} = np \end{aligned}$$

$\blacksquare$

Przykład. Niech $X \sim P_o(\lambda)$. Wówczas

$$E(X) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} =$$

$$= \lambda e^{-\lambda} \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$\blacksquare$

Przykład. Niech $X \sim N(\mu, \sigma^2)$. Wówczas

$$E(X) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} x \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] dx$$

Stosujemy podstawienie $z = \dfrac{x-\mu}{\sigma}$ i otrzymujemy

$$E(X) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} (\mu + \sigma z) e^{-\frac{z^2}{2}}\, dz =$$

$$= \frac{\mu}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\, dz + \frac{\sigma}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} z e^{-\frac{z^2}{2}}\, dz =$$

$$= \frac{\mu}{\sqrt{2\pi}} \cdot \sqrt{2\pi} + \frac{\sigma}{\sqrt{2\pi}} \cdot 0 = \mu$$

■

Przykład. Niech $X \sim C(0,1)$. Wówczas

$$\int\limits_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)}\, dx = 2 \lim_{A\to\infty} \int\limits_{0}^{A} \frac{x}{\pi(1+x^2)}\, dx =$$

$$= 2 \lim_{A\to\infty} \frac{1}{2\pi} \ln(1 + A^2) = \infty$$

Wniosek: Dla rozkładu Cauchy'ego wartość oczekiwana nie istnieje. ■

Własności wartości oczekiwanej

Jeżeli $E(X) < \infty,\ E(Y) < \infty$, to

(i) $E(X+Y) = E(X) + E(Y)$

(ii) $E(aX+b) = aE(X) + b, \quad$ dla $a, b \in \mathbf{R}$

(iii) Jeżeli $X \geq 0$, to $E(X) = \int\limits_0^\infty P(X > t)\, dt$

(iv) Jeżeli X oraz Y są niezależne, to

$$E(XY) = E(X)E(Y)$$

Przykład. Niech $X_i,\ i = 1, 2, \ldots n$, ma rozkład dwupunktowy:

$$P(X_i = 0) = q, \quad P(X_i = 1) = p$$

Jeżeli zdarzenia $A_i = \{X_i = 1\}$ są niezależne, to

$$X = \sum_{k=1}^{n} X_i \sim B(n, p)$$

Zatem

$$E(X) = \sum_{k=1}^{n} E(X_k) = \sum_{k=1}^{n} p = np$$

■

Twierdzenie 14. Jeżeli funkcja φ jest borelowska, to

(i) Dla X z rozkładu skokowego

$$E(\varphi(X)) = \sum_k \varphi(x_k) P(X = x_k)$$

(ii) Dla X z rozkładu ciągłego o gęstości $f(x)$

$$E(\varphi(X)) = \int_{\mathbf{R}} \varphi(x) f(x)\, dx$$

Przykład. Znaleźć wartość oczekiwaną pola prostokąta, którego obwód jest równy 10, a jeden bok jest zmienną losową X o rozkładzie $U[1, 10]$.

Pole $= X(10 - X)$, $f_X(x) = \frac{1}{9} I_{[1,10]}(x)$

$$E(X(10 - X)) = \int_{-\infty}^{\infty} x(10 - x) f_X(x)\, dx =$$

$$= \frac{1}{9} \int_1^{10} x(1 - x)\, dx = 18$$

■

Problem. Jak możliwie najdokładniej zmierzyć długości dwóch prętów za pomocą zwykłej miarki, jeśli wolno mierzyć tylko dwa razy?

Propozycje

1. Mierzymy osobno każdy pręt.
2. Mierzymy sumę długości prętów, składając je razem, a potem – różnicę.

Miernik precyzji pomiaru.

$$\text{Wynik pomiaru} = \text{rzeczywista długość} + \text{błąd}$$
$$X = x + \varepsilon$$

$$\boxed{E(X - x)^2 = E(\varepsilon)^2}$$

ad 1.

Niech X_i oznacza pomiar i – tego pręta, $i = 1, 2$. Zatem $X_i = x_i + \varepsilon_i$. Wielkość błędu pomiaru pierwszego pręta wynosi $E(\varepsilon_1)^2$, a drugiego $E(\varepsilon_2)^2$. Rozsądnie jest przyjąć

$$E(\varepsilon_1)^2 = E(\varepsilon_2)^2 = \sigma^2$$

ad 2.

Niech S oznacza pomiar sumy długości prętów oraz R różnicę.

$$S = x_1 + x_2 + \varepsilon_1$$

$$R = x_1 - x_2 + \varepsilon_2$$

Jako oszacowanie x_1 przyjmujemy

$$\frac{S+R}{2} = x_1 + \frac{\varepsilon_1 + \varepsilon_2}{2}$$

Jako oszacowanie x_2 przyjmujemy

$$\frac{S-R}{2} = x_2 + \frac{\varepsilon_1 - \varepsilon_2}{2}$$

Rozsądnie jest przyjąć, że

$$E(\varepsilon_1) = E(\varepsilon_2) = 0, \quad \text{oraz } \varepsilon_1, \varepsilon_2 \text{ niezależne}$$

Na mocy twierdzenia 19, 20 oraz założeń:

$$\begin{aligned} E\left(\frac{\varepsilon_1 \pm \varepsilon_2}{2}\right)^2 &= \frac{1}{4}E(\varepsilon_1)^2 + \frac{1}{4}E(\varepsilon_2)^2 \pm \frac{1}{2}E(\varepsilon_1)E(\varepsilon_2) \\ &= \frac{1}{4}E(\varepsilon_1)^2 + \frac{1}{4}E(\varepsilon_2)^2 \pm 0 \cdot 0 = \frac{\sigma^2}{2} \end{aligned}$$

Średni kwadrat błędu jest dwa razy mniejszy niż poprzednio.

Zauważmy, że jeżeli $E(\varepsilon) = 0$ to $E(X) = x$. Zatem

$$\boxed{E(X - x)^2 = E(X - E(X))^2}$$

Definicja. Jeżeli $E(X - EX)^2 < \infty$, to tę liczbę nazywamy wariancją zmiennej losowej X i oznaczamy:

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

Uwaga.

$$\begin{aligned} D^2X = E(X - EX)^2 &= E(X^2 - 2X \cdot EX + (EX)^2) \\ &= EX^2 - (EX)^2 \end{aligned}$$

Definicja. Pierwiastek z wariancji nazywamy odchyleniem standardowym i oznaczamy przez DX.

Własności wariancji

Jeżeli X jest zmienną losową, dla której $EX^2 < \infty$, to istnieje D^2X oraz:

(i) $D^2X \geq 0$

(ii) $D^2(cX) = c^2D^2X$

(iii) $D^2(X + a) = D^2X$

(iv) $D^2X = 0$ wtedy i tylko wtedy, gdy zmienna losowa X jest z prawdopodobieństwem 1 stała

Uwaga.

$$
\begin{aligned}
E(X-t)^2 =& E(X - EX + EX - t)^2 \\
=& E(X - EX)^2 + E(X-t)^2 - \\
& - 2E((X - EX)(EX - t)) \\
=& E(X - EX)^2 + E(X-t)^2 - \\
& - 2E(X - EX) \cdot E(EX - t) \\
\geq& E(X - EX)^2
\end{aligned}
$$

Zatem funkcja $f(t) = E(X-t)^2$ przyjmuje minimum – równe wariancji – dla $t = EX$.

Przykład. Zagadnienie regresji liniowej. Chcemy zmienną Y w rozsądny sposób przybliżyć przy pomocy funkcji liniowej zmiennej X. Za kryterium jakości przybliżenia przyjmiemy średni kwadrat błędu:

wyznaczyć takie liczby a i b, ażeby $E(Y - aX - b))^2$ była minimalna.

Na podstawie uwagi

$$b = E(Y - aX) = EY - aEX$$

Zatem szukamy takiego a, które minimalizuje

$$\begin{aligned} &E(Y - aX - (EY - aEX)))^2 = \\ &\quad = E(Y - EY - a(X - EX))^2 \\ &\quad = D^2Y + a^2D^2X - 2aE((Y - EY)(X - EX)) \end{aligned}$$

Mamy tu funkcję kwadratową względem a. Zatem

$$a = \frac{E((Y - EY)(X - EX))}{D^2X}$$

Oznaczając

$$\varrho(X,Y) = \frac{E((Y-EY)(X-EX))}{\sqrt{D^2X \cdot D^2Y}}$$

mamy

$$aX + b = \varrho(X,Y)\frac{DY}{DX}(X-EX) + EY$$

oraz

$$\min_{a,b} E(Y - aX - b)^2 = (1 - \varrho(X,Y)^2)D^2Y$$

■

Definicja. Kowariancją całkowalnych zmiennych losowych X, Y, spełniających warunek $E|XY| < \infty$, nazywamy wielkość

$$\mathrm{Cov}(X,Y) = E((Y-EY)(X-EX)).$$

Definicja. Współczynnikiem korelacji zmiennych X, Y nazywamy wielkość

$$\varrho(X,Y) = \frac{E((Y-EY)(X-EX))}{\sqrt{D^2X \cdot D^2Y}}.$$

Uwaga. Z ostatniej równości w przykładzie wynika:

(i) $-1 \leq \varrho(X,Y) \leq 1$

(ii) $|\varrho(X,Y)| = 1$ wtedy i tylko wtedy, gdy istnieją liczby $a \neq 0$ oraz b takie, że $P(Y = aX + b) = 1$

Wariancja sumy zmiennych losowych

Jeżeli każda ze zmiennych losowych $X_1,\ X_2, \ldots\ ,\ X_n$ ma wariancję, to istnieje wariancja sumy i

$$D^2(X_1+\ldots+X_n) = \sum_{i=1}^{n} D^2 X_i + 2 \sum_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j)$$

Definicja. Zmienne losowe X, Y, dla których

$$\mathrm{Cov}(X,Y) = 0, \quad \text{czyli } \varrho(X,Y) = 0,$$

nazywamy nieskorelowanymi.

Wniosek. Jeśli zmienne losowe $X_1,\ X_2, \ldots\ ,\ X_n$ mają wariancję i są parami nieskorelowane, to

$$D^2(X_1 + \ldots + X_n) = \sum_{i=1}^{n} D^2 X_i$$

(?) **Uwaga.** Jeżeli X, Y są niezależne, to są nieskorelowane.

Odwrotny fakt nie zachodzi (chyba, że mamy do czynienia z rozkładem normalnym)

Przykład. Niech (X_1, X_2) ma dwuwymiarowy rozkład normalny. Policzmy $\mathrm{Cov}(X_1, X_2)$. Zgodnie z przekształceniami z przykładu na rozkład brzegowy dwywymiarowego rozkładu normalnego mamy:

$$f_{X_1,X_2}(x_1, x_2) = f_{X_2}(x_2)\,\underbrace{g(x_1, x_2)}_{g(x_1)-\text{prz.}}$$

gdzie $f_{X_2}(x_2)$ jest gęstością rozkładu $N(\mu_2, \sigma_2^2)$ oraz $g(x_1, x_2)$ traktowana jako funkcja zmiennej x_1 z parametrem x_2, jest funkcją gęstości zmiennej

$$N\left(\mu_1 + \varrho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), (1 - \varrho^2)\sigma_1^2\right)$$

Zatem

$$\mathrm{Cov}(X_1, X_2) = E((X_1 - \mu_1)(X_2 - \mu_2)) =$$

$$= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (x_1 - \mu_1)(x_2 - \mu_2) f_{X_2}(x_2) g(x_1, x_2)\, dx_1 dx_2 =$$

$$= \left(\int_{-\infty}^{\infty} (x_2 - \mu_2) f_{X_2}(x_2) \cdot \right.$$

$$\left. \cdot \left(\int_{-\infty}^{\infty} (x_1 - \mu_1) g(x_1, x_2)\, dx_1 \right) \right) dx_2$$

Zatem

$$= \left(\int_{-\infty}^{\infty} (x_2 - \mu_2) f_{X_2}(x_2) \cdot \right.$$

$$\left. \cdot \left(\mu_1 + \varrho \frac{\sigma_1}{\sigma_2}(x_2 - \mu_2) - \mu_1 \right) \right) dx_2$$

A zatem

$$= \varrho \frac{\sigma_1}{\sigma_2} \int_{-\infty}^{\infty} (x_2 - \mu_2)^2 f_{X_2}(x_2) dx_2 = \varrho \frac{\sigma_1}{\sigma_2} \sigma_2^2 = \varrho \sigma_1 \sigma_2$$

Stąd $\varrho(X, Y) = \varrho$. Zatem X, Y niezależne $\Leftrightarrow$ X, Y nieskorelowane. ■

Rozkłady warunkowe

Przykład. Niech (X, Y) – dwuwymiarowy wektor losowy typu skokowego

$$X \in \{x_1, x_2, \dots\}, \quad Y \in \{y_1, y_2, \dots\}$$

Rozkład

$$p_{ij} := P(X = x_i, Y = y_j)$$

Prawdopodobieństwa brzegowe

$$P(X = x_i) = \sum_k p_{ik}, \quad P(Y = y_k) = \sum_i p_{ik}$$

Zachodzi

$$P(X = x_i | Y = y_k) \geq 0, \quad \sum_i P(X = x_i | Y = y_k) = 1$$

$$P(Y = y_k | X = x_i) \geq 0, \quad \sum_k P(Y = y_k | X = x_i) = 1$$

■

Zatem dla ustalonego y_k

$$P(\,\cdot\,| Y = y_k)$$

jest rozkładem prawdopodobieństwa.

Podobnie

$$P(\,\cdot\,| X = x_i)$$

Przykład. Rzut dwiema kostkami.

X – wynik rzutu pierwszą kostką
Y – wynik rzutu drugą kostką

$U := \min\{X, Y\}, \quad V := \max\{X, Y\}$

$u \setminus v$	1	2	3	4	5	6	$P(U = u)$
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{11}{36}$
2	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
3	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{7}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{5}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$
6	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
$P(V = v)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	

v	1	2	3	4	5	6	suma
$P(V = v \vert U = 3)$	0	0	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{2}{7}$	1

$$E(V|U = 3) = \frac{33}{7} \quad F(4|U = 3) = \frac{3}{7}$$

■

Przykład. Jaka jest średnia liczba sukcesów w pierwszej próbie, jeżeli wiemy, ile zaszło sukcesów w całej serii n doświadczeń według schematu Bernoulliego?

Oznaczenia

S_n – łączna liczba sukcesów
Y – liczba sukcesów w pierwszej próbie

$A_k := \{S_n = k\}, \quad B_k := A_k \cap \{Y = 1\}$

$$
\begin{aligned}
E(Y|A_k) &= \sum_{\omega \in A_k} Y(\omega) P(\omega|A_k) = \\
&= \frac{1}{P(A_k)} \sum_{\omega \in A_k} Y(\omega) P(\omega) = \\
&= \frac{1}{P(A_k)} \sum_{\omega \in B_k} P(\omega) = \frac{P(B_k)}{P(A_k)} = \\
&= \frac{p\binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)}}{\binom{n}{k} p^k q^{n-k}} = \frac{k}{n}
\end{aligned}
$$

Zatem $E(Y|S_n) = \frac{S_n}{n}$ oraz

$$
E(E(Y|S_n)) = E(S_n/n) = \frac{E(S_n)}{n} = \frac{np}{n} = p = E(Y)
$$

$\blacksquare$

Przykład. Niech $f(x,y)$ – gęstość wektora (X,Y).

Rozkłady brzegowe

$$\text{dla zmiennej } X: \quad f_1(x) := \int\limits_{-\infty}^{\infty} f(x,y)\,dy$$

$$\text{dla zmiennej } Y: \quad f_2(y) := \int\limits_{-\infty}^{\infty} f(x,y)\,dx$$

Niech $P(x < X \leq x+h) > 0$.

Wtedy

$$P(Y \leq y | x < X \leq x+h) = \frac{\int\limits_{x}^{x+h} \left(\int\limits_{-\infty}^{y} f(x,y)\,dy \right) dx}{\int\limits_{x}^{x+h} f_1(x)\,dx}$$

(?) Przy założeniu, że $f(x,y)\ldots$ oraz $f_1(x)\ldots$

$$\begin{aligned} P(Y \le y | X = x) :&= \lim_{h\to 0^+} P(Y \le y | x < X \le x+h) \\ &= \lim_{h\to 0^+} \frac{\frac{1}{h}\int\limits_x^{x+h}\left(\int\limits_{-\infty}^{y} f(x,y)\,dy\right)dx}{\frac{1}{h}\int\limits_x^{x+h} f_1(x)\,dx} \\ &= \frac{\int\limits_{-\infty}^{y} f(x,y)\,dy}{f_1(x)} = \int\limits_{-\infty}^{y} \frac{f(x,y)}{f_1(x)}\,dy \end{aligned}$$

Oznaczając

$$F(y|x) = P(Y \le y | X = x), \quad f(y|x) = f(x,y)/f_1(x)$$

mamy

$$\boxed{F(y|x) = \int\limits_{-\infty}^{y} f(y|x)\,dy}$$

Zauważamy, że

$$f_1(x)F(y|x) = \int\limits_{-\infty}^{y} f(x,y)\,dy$$

Po scałkowaniu obu stron

$$\boxed{\int\limits_{-\infty}^{\infty} f_1(x)F(y|x)\,dx = F_Y(y)}$$

Przyjmując

$$E(Y|x) := \int\limits_{-\infty}^{\infty} yf(y|x)\,dy$$

mamy

$$\begin{aligned}\int\limits_{-\infty}^{\infty} E(Y|x)f_1(x)\,dx &= \int\left(\int yf(y|x)\,dy\right)f_1(x)\,dx\\ &= \int\left(\int y\frac{f(x,y)}{f_1(x)}\,dy\right)f_1(x)\,dx\end{aligned}$$

$$\int \left(\int y f(x,y)\, dy \right) dx = \int y \left(\int f(x,y)\, dx \right) dy =$$

$$= \int\limits_{-\infty}^{\infty} y f_2(y)\, dy = E(Y)$$

Otrzymaliśmy $\boxed{E(E(Y|X)) = E(Y)}$ ■

Definicja. Jeżeli (X, Y) jest wektorem losowym o gęstości $f(x, y)$ to funkcję

$$f(y|x) = \begin{cases} \frac{f(x,y)}{f_1(x)} & \text{gdy } f_1(x) > 0 \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

nazywamy gęstością warunkową zmiennej Y dla danego $X = x$.

$$(?) \int\limits_{-\infty}^{\infty} f(y|x)\, dy = 1$$

Nadal zachodzi $\quad E(E(Y|X)) = E(Y)$

Ponadto rozumiemy, że

$$P(Y \in B|x) = \int\limits_B f(y|x)\,dy\text{– rozkład warunkowy}$$

Przyjmując $Z(\omega) = I_B(Y(\omega))$ mamy

$$\begin{aligned} E(Z) &= \int\limits_{-\infty}^{\infty} I_B(y) f_2(y)\,dy = \int\limits_B f_2(y)\,dy \\ &= P(Y \in B) \\ E(Z|x) &= \int\limits_{-\infty}^{\infty} I_B(y) f(y|x)\,dy = \int\limits_B f(y|x)\,dy \\ &= P(Y \in B|x) \end{aligned}$$

Zatem

$$\boxed{E(P(Y \in B|X)) = P(Y \in B)}$$

Dla $B = (-\infty, y]$ mamy

dystrybuantę zmiennej Y

$$F_Y(y) = P(Y \in B)$$

dystrybuantę zmiennej Y pod warunkiem $X = x$

$$F(y|x) = P(Y \in B|x)$$

oraz wzór

$$\boxed{E(F(y|X)) = F_Y(y)}$$

Przykład. Z odcinka [0,1] wybrano losowo (zgodnie z rozkładem równomiernym) punk X, a następnie z odcinka $[0, X]$, również losowo, punkt Y. Jaka jest średnia długość odcinka $[0, Y]$?

$$E(Y|X) = \frac{1}{2}X$$

$$E(Y) = E(E(Y|X)) = E(\frac{1}{2}X) = \frac{1}{4}$$

■

Przykład. Owad składa X jajeczek zgodnie z rozkładem Poissona z parametrem λ, a owad z jajeczka wylęga się z prawdopodobieństwem p, niezależnie od innych. Znaleźć średnią liczbę potomków.

Niech Y oznacza liczbę potomków owada. Zatem

$$E(Y|X) = Xp$$

Stąd

$$EY = E(E(Y|X)) = E(Xp) = \lambda p$$

Ten przykład pokazuje, jak można obliczać wartość oczekiwaną, korzystając z warunkowej wartości oczekiwanej. Właściwy wybór zmiennej losowej X często bardzo upraszcza rachunki. ■

Uwaga. Skorzystaliśmy ze wzoru $EY = E(E(Y|X))$, gdy X typu skokowego.

Ja w takim przypadku rozumieć „gęstość łączną”?

Umowa: $\int\limits_a^b f(x,y)\,dy = P(X = x, a \le Y \le b)$

Przy takiej umowie możemy zachować bez zmian określenia „gęstości warunkowych”

$$f(x|y) = \frac{f(x,y)}{f_2(y)}, \quad f(y|x) = \frac{f(x,y)}{f_1(x)}$$

gdzie

$$f_1(x) = \int f(x,y)\,dy, \quad f_2(y) = \sum_x f(x,y)$$

Rodzaje zbieżności

Przykład. Niech P -rozkład jednostajny na $[0,1]$ oraz

$$X_{kn}(\omega) = \begin{cases} 1 & \text{dla } \omega \in \left[\frac{k}{n}; \frac{k+1}{n}\right); \\ 0 & \text{dla } \omega \in \Omega \setminus \left[\frac{k}{n}; \frac{k+1}{n}\right) \end{cases}$$

dla $0 \le k \le n-1, \quad n = 1, 2, \ldots$

$$P(X_{nk} = 0) = 1 - \frac{1}{n}, \quad P(X_{nk} = 1) = \frac{1}{n}$$

$$P(|X_{nk}| > \varepsilon) = \frac{1}{n} \xrightarrow[n \to \infty]{} 0, \quad \text{dla } 0 < \varepsilon < 1$$

O ciągu X_{01}, $\boldsymbol{X}_{02}, \boldsymbol{X}_{12}$, $X_{03}, X_{13}, X_{23}, \ldots$ powiemy, że jest zbieżny do zera według prawdopodobieństwa.

Ciąg ten jest rozbieżny w każdym punkcie przedziału.

Na przykład dla $\omega = 1/2$ mamy ciąg: $0, 0, 1, 0, \ldots$, który na dowolnie dalekich miejscach ma zera i jedynki. ■

Definicja. Ciąg zmiennych losowych $(X_n)_{n=1}^{\infty}$ jest zbieżny do zmiennej losowej X:

według prawdopodobieństwa, jeśli

$$\text{dla każdego } \varepsilon > 0 \quad \lim_n P(|X_n - X| > \varepsilon) = 0,$$

co oznaczamy $X_n \xrightarrow{P} X$,

prawie na pewno, jeśli

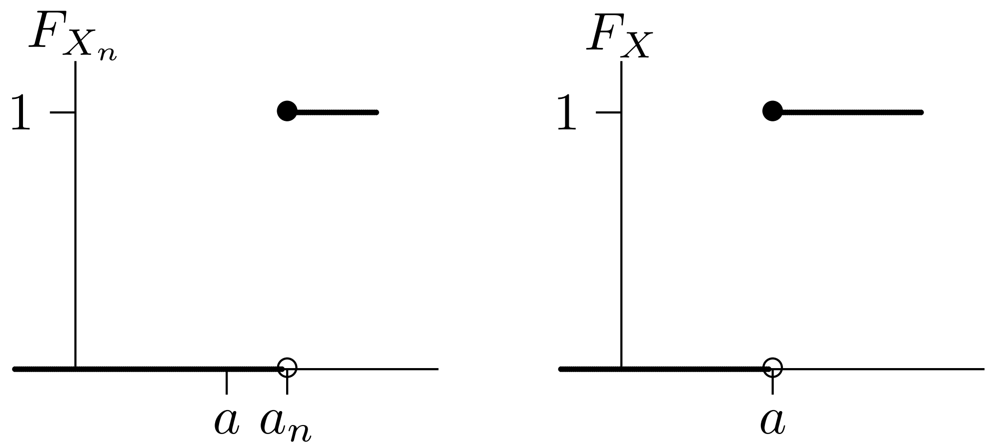
$$P\left(\left\{\omega : \lim_n X_n(\omega) = X(\omega)\right\}\right) = 1$$

co oznaczamy $X_n \xrightarrow{p.n.} X$

$$\begin{aligned}
X_n \xrightarrow{p.n.} X \Leftrightarrow P\left(\bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{|X_n - X| \le \varepsilon\}\right) = 1 \Leftrightarrow \\
\Leftrightarrow \lim_{N\to\infty} P\left(\bigcap_{n=N}^{\infty} \{|X_n - X| \le \varepsilon\}\right) = 1 \Leftrightarrow \\
\Leftrightarrow \lim_{N\to\infty} P\left(\bigcup_{n=N}^{\infty} \{|X_n - X| > \varepsilon\}\right) = 0 \Rightarrow \\
\Rightarrow \lim_{N\to\infty} P\left(\{|X_N - X| > \varepsilon\}\right) = 0 \Leftrightarrow \\
\Leftrightarrow X_n \xrightarrow{P} X
\end{aligned}$$

Wniosek. Zbieżność prawie na pewno pociąga zbieżność według prawdopodobieństwa.

Przykład. Niech X_n ma rozkład $P(X_n = a_n) = 1$. Zatem $F_{X_n}(t) = I_{[a_n,\infty)}(t)$



Gdy $a_n \downarrow a$ okazuje się, że $F_{X_n}(a) \equiv 0 \neq 1 = F_X(a)$ ■

Przykład. Niech F będzie dowolną dystrybuantą. Zdefiniujmy dystrybuantę $F_n(t)=F(t-\frac{1}{n})$, $n=1,2,\ldots$ Wtedy

$$F_n(t) \to F(t-).$$

Zauważmy, że $F(t-) = F(t)$ tylko wtedy, gdy t jest punktem ciągłości t. ■

Definicja. Ciąg zmiennych losowych $(X_n)_{n=1}^{\infty}$ jest zbieżny do zmiennej losowej X według dystrybuant, jeśli ciąg dystrybuant $(F_{X_n})_{n=1}^{\infty}$ jest zbieżny do dystrybuanty F_X w każdym punkcie jej ciągłości, co oznaczamy:

$$X_n \xrightarrow{D} X$$

Można pokazać

$$(\, X_n \overset{p.n.}{\to} X) \Rightarrow (X_n \overset{P}{\to} X) \Rightarrow (X_n \overset{D}{\to} X)$$

Prawa wielkich liczb

Oznaczmy

$$S_n = X_1 + X_2 + \cdots + X_n, \quad \bar{X}_n = \frac{X_1 + X_2 + \ldots +, X_n}{n}$$

Niech $X_1, X_2, \ldots$ będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie, o wartości średniej μ i wariancji $0 < \sigma^2 < \infty$. Wtedy dla każdego $\varepsilon > 0$ mamy

Słabe prawo wielkich liczb

$$\lim_{n\to\infty} P\left(\left| \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right| < \varepsilon \right) = 1$$

$$\boxed{\bar{X}_n \xrightarrow{P} \mu}$$

Mocne prawo wielkich liczb

$$P\left(\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu \right) = 1$$

$$\boxed{\bar{X}_n \xrightarrow{p.n.} \mu}$$

Wniosek. Prawdopodobieństwo jest odpowiednikiem teoretycznym częstości.

Faktycznie, jeżeli w wyniku powtórzenia niezależnie n razy doświadczenia otrzymaliśmy $\omega_1,\ \omega_2, \dots,\ \omega_n$, to

$$\frac{I_A(\omega_1) + I_A(\omega_2) + \cdots + I_A(\omega_n)}{n} \xrightarrow{p.n.} EI_A = P(A)$$

Metoda Monte Carlo obliczania całek.

Niech X_i będą niezależnymi zmiennymi losowymi o wartościach w (0,1) i o gęstości g. Wtedy z MPWL

$$S_n = \frac{1}{n} \sum_{i=1}^{n} \frac{f(X_i)}{g(X_i)} \xrightarrow[n \to \infty]{} E\left[\frac{f(X_1)}{g(X_1)}\right] =$$

$$= \int_0^1 \frac{f(x)}{g(x)} \cdot g(x)\, dx = \int_0^1 f(x)\, dx$$

W szczególności, gdy $X_i \sim U(0,1)$, to

$$\frac{1}{n} \sum_{i=1}^{n} f(X_i) \xrightarrow[n \to \infty]{} \int_0^1 f(x)\, dx$$

Przykład. Obliczanie liczby π przy pomocy komputera: generujemy 50 wartości z rozkładu $U(0,1)$ (kolumna x). Następnie wyliczamy $y = \sqrt{1-x^2}$. Z kolumny y wyliczamy średnią i mnożymy ją przez cztery. Otrzymujemy wartość 3.155. Jeśli przybliżenia to nie nie jest zadowalające, można wygenerować na przykład 1000 wartości.

x	y	x	y	x	y
0.382	0.924	0.017	1.000	0.952	0.307
0.101	0.995	0.285	0.959	0.053	0.999
0.596	0.803	0.343	0.939	0.705	0.709
0.899	0.438	0.554	0.833	0.817	0.577
0.885	0.466	0.357	0.934	0.973	0.233
0.958	0.285	0.372	0.928	0.466	0.885
0.014	1.000	0.356	0.935	0.300	0.954
0.407	0.913	0.910	0.414	0.750	0.661
0.863	0.505	0.466	0.885	0.351	0.936
0.139	0.990	0.426	0.905	0.776	0.631
0.245	0.970	0.304	0.953	0.074	0.997
0.045	0.999	0.976	0.219	0.198	0.980
0.032	0.999	0.807	0.591	0.064	0.998
0.164	0.986	0.991	0.132	0.358	0.934
0.220	0.976	0.256	0.967	0.487	0.873
0.511	0.859	0.373	0.928	0.986	0.167
0.041	0.999	0.231	0.973		

■

Dystrybuanta empiryczna $F_n(x)$

Powtarzamy pewne doświadczenie niezależnie n razy. W wyniku tego otrzymujemy ciąg

$$X_1,\ X_2, \dots ,\ X_n$$

niezależnych zmiennych losowych o nieznanej dystrybuancie F.

Chcemy odtworzyć F. W tym celu dla każdego $x \in \mathbf{R}$ definiujemy

$$F_n(x)(\omega) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le x\}}(\omega)$$

Ponieważ

$$E[I_{\{X_1 \le x\}}] = P(X_1 \le x) = F(x),$$

to z MPWL

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le x\}} \xrightarrow[n\to\infty]{} F(x)$$

Centralne twierdzenie graniczne

Centralne Twierdzenie Graniczne
Niech $X_1,\ X_2,\ldots$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie, o wartości średniej μ i wariancji $0 < \sigma^2 < \infty$. Wtedy

$$\sup_{x\in\mathbf{R}}\left|P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) - \Phi(x)\right| \xrightarrow[n\longrightarrow\infty]{} 0$$

$$\boxed{\frac{\bar{X}_n - \mu}{\sigma}\sqrt{n} \xrightarrow{D} N(0,1)}$$

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

$$\sup_{x\in\mathbf{R}}\left|P\left(\frac{Y_n - np}{\sqrt{npq}} \le x\right) - \Phi(x)\right| \xrightarrow[n\longrightarrow\infty]{} 0$$

Przykład. Wykonano $n = 100$ niezależnych rzutów monetą. Oznaczmy przez Y_n liczbę orłów w n rzutach.

Obliczymy $P(Y_n \geq 61)$

$$
\begin{aligned}
P(Y_n \geq 61) &= 1 - P(Y_n \leq 60) = \\
&= 1 - P\left(\frac{Y_n - 100 \cdot 0.5}{10 \cdot 0.5} \leq \frac{60 - 100 \cdot 0.5}{10 \cdot 0.5}\right) = \\
&= 1 - P\left(\frac{Y_n - 100 \cdot 0.5}{10 \cdot 0.5} \leq 2\right) \approx \\
&\approx 1 - \Phi(2) \approx 0.0228
\end{aligned}
$$

■

Uwaga. Dość dobre przybliżenie uzyskujemy ze wzoru:

$$
P\left(a \leq \frac{Y_n - np}{\sqrt{npq}} \leq b\right) \sim \Phi(b + \frac{1}{2}h) - \Phi(a - \frac{1}{2}h),
$$

gdzie $h = \dfrac{1}{\sqrt{npq}}$

Szybkość zbieżności w centralnym twierdzeniu granicznym

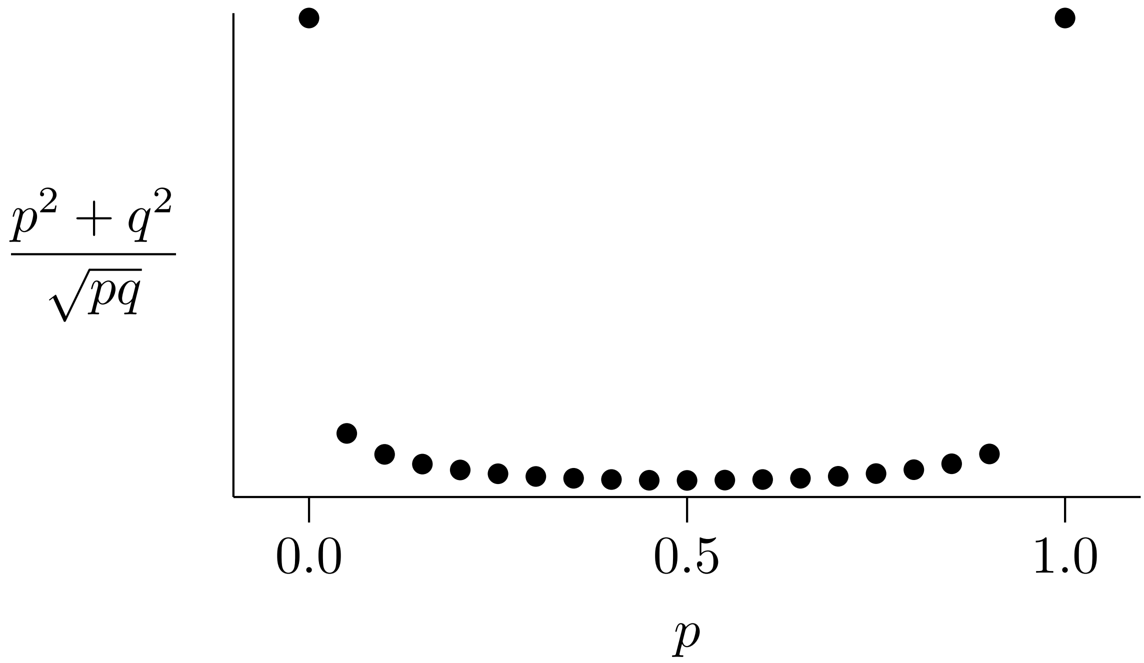
Twierdzenie Berry–Esséen'a
Jeżeli $X_1, X_2, \ldots$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie oraz $E|X_1|^3 < \infty$, to

$$\sup_{x \in \mathbf{R}} \left| P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x \right) - \Phi(x) \right| \le C \frac{E|X_1 - EX_1|^3}{\sigma^3\sqrt{n}},$$

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

Dla rozkładu dwumianowego:

$$\sup_{x \in \mathbf{R}} \left| P\left(\frac{Y_n - np}{\sqrt{npq}} \le x \right) - \Phi(x) \right| \le C \frac{p^2 + q^2}{\sqrt{npq}}$$



Dla $p \approx 1$ lub $p \approx 0$ przybliżenie rozkładem normalnym nie musi być zadowalające. Alternatywą jest przybliżenie rozkładem Poissona:

Twierdzenie 15. Niech $Y_n \sim B(n,p)$ oraz $\lambda = np$. Wtedy dla każdego zbioru $M \subseteq \mathbf{N}$ mamy

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

Przykład. Prawdopodobieństwo trafienia „szóstki” w Toto-Lotku jest równe

$$1/\binom{49}{6} = 1/13983816 \approx 7.151 \cdot 10^{-8}.$$

Ilu „szóstek” można się spodziewać w każdym tygodniu, jeżeli grający wypełniają kupony całkowicie losowo i niezależnie od siebie, a kuponów jest $n = 10^7$.

Liczba „szóstek” ma rozkład dwumianowy, w przybliżeniu rozkład Poissona z parametrem $\lambda = np \approx 0.7151$.

k	$\frac{\lambda^k}{k!}e^{-\lambda}$	k	$\frac{\lambda^k}{k!}e^{-\lambda}$
0	0.4891	3	0.0298
1	0.3498	4	0.0053
2	0.1251	5	0.0008

Błąd przybliżenia rozkładem Poissona:

$$\lambda^2/n \leq 0.5 \cdot 10^{-7}.$$

■

Twierdzenie Słuckiego

Niech $X_n \xrightarrow{D} X$ oraz $Y_n \xrightarrow{D} c$, gdzie c jest pewną skończoną stałą. Wówczas:

(i) $X_n + Y_n \xrightarrow{D} X + c$

(ii) $X_n Y_n \xrightarrow{D} cX$

(iii) $X_n / Y_n \xrightarrow{D} X/c$

Z twierdzenia Słuckiego wynika, że ciąg zmiennych losowych $(X_n)_n$ zbiega według rozkładu do $N(\mu, \sigma^2)$, jeżeli równoważnie ciąg $\dfrac{X_n - \mu}{\sigma}$ zbiega do rozkładu $N(0,1)$.

Asymptotyczna normalność

Mówimy, że ciąg zmiennych $(X_n)_n$ jest asymptotycznie normalny o średniej μ_n i wariancji σ_n^2, jeżeli $\sigma_n^2 > 0$ dla dostatecznie dużych n oraz

$$\frac{X_n - \mu_n}{\sigma_n} \to N(0,1).$$

Zapisujemy to jako: X_n jest $AN(\mu_n, \sigma_n^2)$.

Asymptotyczna normalność przy przekształaceniach

Niech X_n będzie $AN(\mu, \sigma_n^2)$, $\sigma_n \to 0$. Niech g będzie funkcją różniczkowalną w punkcie $x = \mu$ oraz niech $g'(\mu) \neq 0$. Wówczas

$$g(X_n) \text{ jest } AN(g(\mu), (g'(\mu))^2\sigma_n^2)$$

Przykład. Niech X_n ma rozkład Poissona o wartości oczekiwanej θn, gdzie $\theta > 0$. Wówczas X_n jest

$$AN(\theta n, \theta n)$$

(wariancja rozkładu Poissona jest równa wartości średniej) lub równoważnie

$$\frac{X_n}{n} \text{ jest } AN(\theta, \frac{\theta}{n}).$$

Niech g będzie rozwiązaniem równania

$$\frac{dg(\theta)}{d\theta} = \frac{1/2}{\theta^{1/2}}.$$

To znaczy $g(x) = x^{1/2}$. Zatem $(X_n/n)^{1/2}$ jest

$$AN(\theta^{1/2}, 1/(4n))$$

lub równoważnie $X_n^{1/2}$ jest

$$AN((\theta n)^{1/2}, 1/4).$$

■

Własności rozkładów

Rozkład sumy niezależnych zmiennych losowych.

Niech X, Y mają rozkłady dyskretne:

$$\sum_{x \in \mathcal{X}} P(X = x) = 1, \quad \sum_{y \in \mathcal{Y}} P(Y = y) = 1$$

Szukamy rozkładu zmiennej losowej $Z = X + Y$:

$$\begin{aligned} P(Z = z) &= P(X + Y = z) = \\ &= \sum_{x \in \mathcal{X}} P(X = x, Y = z - x) = \\ &= \sum_{x \in \mathcal{X},\ z - x \in \mathcal{Y}} P(X = x) P(Y = z - x) = \end{aligned}$$

W przypadku, gdy $\mathcal{X} = \{0, 1, \dots\}$ oraz $\mathcal{Y} = \{0, 1, \dots\}$ mamy

$$P(X + Y = r) = \sum_{i=0}^{r} P(X = i) P(Y = r - i)$$

Przykład. $X \sim B(n_1, p), \quad Y \sim B(n_2, p).$

$$\begin{aligned}
&P(X+Y=r)= \\
&=\sum_{i=0}^{r}\binom{n_1}{i}p^i(1-p)^{n_1-i}\binom{n_2}{r-i}p^{r-i}(1-p)^{n_1-r+i} \\
&=p^r(1-p)^{n_1+n_2-r}\sum_{i=0}^{r}\binom{n_1}{i}\binom{n_2}{r-i} \\
&=\binom{n_1+n_2}{r}p^r(1-p)^{n_1+n_2-r}
\end{aligned}$$

Zatem $X+Y \sim B(n_1+n_2, p)$ ■

Przykład. $X \sim Po(\lambda_1), \quad Y \sim Po(\lambda_2)$

$$\begin{aligned}
&P(X+Y=r)= \\
&=\sum_{i=0}^{r}\frac{\lambda_1^i}{i!}e^{-\lambda_1}\cdot\frac{\lambda^{r-i}}{(r-i)!}e^{-\lambda_2} \\
&=e^{-(\lambda_1+\lambda_2)}\frac{1}{r!}\sum_{i=0}^{r}\binom{r}{i}\lambda_1^i\lambda_2^{r-i} \\
&=\frac{(\lambda_1+\lambda_2)^r}{r!}e^{-(\lambda_1+\lambda_2)}
\end{aligned}$$

Zatem $X+Y \sim Po(\lambda_1+\lambda_2)$ ■

Niech X, Y mają rozkłady ciągłe:

$$X \sim f_X(x),\ Y \sim f_Y(y)$$

Wówczas (porównać – strona 86)

$$F_Z(z) = P(X+Y \leq z) = \int\limits_{X+Y \leq z} f_{X,Y}(x,y)\,dxdy =$$

$$= \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{z-y} f_{X,Y}(x,y)\,dx \right) dy =$$

$$= \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{z-y} f_X(x)\,dx \right) f_Y(y)dy =$$

$$= \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{z} f_X(x-y)\,dx \right) f_Y(y)dy =$$

$$= \int\limits_{-\infty}^{z} \left(\int\limits_{-\infty}^{\infty} f_X(x-y) f_Y(y)\,dy \right) dx$$

Zatem $f_Z(z) = \int\limits_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,dy$

Przykład. $X \sim U[0,1]$, $Y \sim U[0,1]$

Ponieważ $I_{[0,1]}(z-y) = I_{[-1,0]}(y-z) = I_{[z-1,z]}(y)$, mamy

$$f_{X+Y}(z) = \int\limits_{-\infty}^{\infty} I_{[0,1]}(z-y) I_{[0,1]}(y)\, dy =$$

$$= \int\limits_{0}^{1} I_{[z-1,z]}(y)\, dy = \begin{cases} z & \text{dla } 0 \le z \le 1, \\ 2-z & \text{dla } 1 \le z \le 2 \\ 0 & \text{dla } z \notin [0,2] \end{cases}$$

Jest to rozkład trójkątny ■

Przykład. Niech $X_0, X_1, \ldots, X_n$ mają rozkład wykładniczy:

$$\text{tzn. o gęstości } f(x) = \lambda e^{-\lambda x} \text{ dla } x > 0$$

Wtedy $X_0 + X_1 + \cdots + X_n$ ma rozkład o gęstości

$$g_n(x) = \lambda \frac{(\lambda x)^n}{n!} e^{-\lambda x} \text{ dla } x > 0$$

Jest to rozkład gamma $G(1, n+1)$

Dowód. $n = 0$

$$g_0(x) = f(x)$$

$n = k, \quad n = k+1$

$$
\begin{aligned}
g_{k+1}(x) &= \int_0^\infty f(x-y) g_k(y)\, dy = \\
&= \int_0^x \lambda e^{-\lambda(x-y)} \cdot \lambda \frac{(\lambda y)k}{k!} e^{-\lambda y}\, dy \\
&= \frac{\lambda^{k+2}}{k!} e^{-\lambda x} \int_0^x y^k\, dy = \lambda \frac{(\lambda x)^{k+1}}{(k+1)!} e^{-\lambda x}
\end{aligned}
$$

■

Ponadto można pokazać, że

$$G_n(x) = 1 - e^{-\lambda x}(1 + \frac{\lambda x}{1!} + \cdots + \frac{(\lambda x)^n}{n!}), \quad x > 0$$

jest dystrybuantą rozkładu gamma $G(\lambda, n+1)$. ■

Przykład. Proces Poissona

Oznaczmy przez $X_1, X_2, \ldots$ niezależne zmienne losowe o wspólnym rozkładzie wykładniczym oraz przyjmijmy

$$S_0 = 0, \quad S_n = X_1 + X_2 + \ldots + X_n$$

Niech $N(t)$ oznacza liczbę wskaźników $k \geq 1$ takich, że $S_k \leq t$. Zdarzenie $\{N(t) = n\}$ następuje wtedy i tylko wtedy, gdy

$$S_n \leq t, \; S_{n+1} > t.$$

Ponieważ S_n ma rozkład G_{n-1}, to

$$P(N(t) = n) = G_{n-1}(t) - G_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Interpretacja

$$X_n \qquad\qquad X_{n-1} \qquad\qquad \cdots \qquad\qquad X_2 \qquad X_1$$

X_i – czas oczekiwania na klienta „$i + 1$” od chwili przybycia klienta „i”

$N(t)$ – liczba przybyłych klientów do chwili t

Problem. Czy w praktyce X_i może mieć rozkład wykładniczy?

Niech T oznacza czas oczekiwania na klienta.

Zakładamy, że prawdopodobieństwo tego, że klient, na którego czekamy już t jednostek czasu, przybędzie w ciągu czasu Δt jest równe $\lambda \Delta t + o(\Delta t)$, gdzie

$$o(\Delta t): \quad \lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$$

(niezależnie od tego, jak długo czekamy).

Zatem

$$\begin{aligned} P(T > t + \Delta t) &= P(T > t + \Delta t,\ T > t) = \\ &= P(T > t + \Delta t | T > t) P(T > t) = \\ &= (1 - \lambda \Delta t - o(\Delta t)) P(T > t) \end{aligned}$$

Zatem

$$\begin{aligned}\frac{P(T > t + \Delta t) - P(T > t)}{\Delta t} &= \\ &= -\lambda P(T > t) - \frac{o(\Delta t)}{\Delta t} P(T > t)\end{aligned}$$

Oznaczając $P(t) := P(T > t)$, dla $\Delta t \to 0$ mamy:

$$P'(t) = -\lambda P(t)$$

Stąd $P(t) = ce^{-\lambda t}$.

Zatem

$$F_T(t) = \begin{cases} 1 - P(t) = 1 - ce^{-\lambda t} & \text{dla } t \geq 0, \\ 0 & \text{dla } t < 0 \end{cases}$$

Ponieważ musi zachodzić $F_T(0) = 0$, więc $c = 1$. Zatem T ma rozkład wykładniczy. ■

Przykład. Rozkład wykładniczy a własność braku pamięci. Niech T ma rozkład wykładniczy z parametrem λ. Zauważmy, że

$$\begin{aligned} P(T > t + s | T > t) &= \frac{P(T > t + s, T > t)}{P(T > t)} = \\ &= \frac{P(T > t + s)}{P(T > t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = \\ &= e^{-\lambda s} = P(T > s) \end{aligned}$$

Zatem $P(T > t+s|T > t) = P(T > s)$, z czego wynika następująca równość

$$P(T > s + t) = P(T > s)P(T > t)$$

Załóżmy teraz, że nie wiemy jaki rozkład ma zmienna T, ale niech to będzie zmienna losowa ciągła, która spełnia powyższa równość.
Jeżeli $u(t) = P(T > t)$ nie jest tożsamościowo równe zeru, to istnieje punkt x taki, że $u(x) > 0$.
Niech $\alpha = -\ln u(x)$ i niech $v(t) = e^{\alpha t} u(xt)$. Wówczas

$$v(t + s) = v(t)v(s), \quad v(1) = 1$$

Pokażemy, że $v(t) = 1$ dla wszystkich $t > 0$.

Zauważmy

$$
\begin{aligned}
v^2(\frac{1}{2}) &= v(1) = 1; \\
v^n(1/n) &= v(1) = 1 \text{ dla } n \in \mathbf{N}; \\
v(m/n) &= v^m(1/n) = 1 \text{ dla } m, n \in \mathbf{N}
\end{aligned}
$$

Zatem $v(w) = 1$ dla wszystkich w wymiernych dodatnich. Z ciągłości v wynika, że jest to prawda dla każdej rzeczywistej dodatniej. Zatem

$$v(t) = e^{\alpha t} u(xt) = 1$$

Przyjmując $y = xt \in (0, \infty)$ oraz $\lambda = \alpha/x$ mamy

$$P(T > y) = u(y) = e^{-\lambda y}$$

Zatem zmienna T ma rozkład wykładniczy. Powyższe przekształcenia pokazały, że rozkład wykładniczy jest jedynym rozkładem ciągłym (nieujemnym) o własności braku pamięci. ■

Przykład. Własność braku pamięci dla rozkładu dyskretnego.

Skorzystamy z poprzedniego wyniku:

$$P(T > y) = (e^{\lambda})^{y}$$

Niech $k \in \mathbf{N}$ oraz $1 - p = e^{\lambda}$. Wówczas

$$P(T = k) = P(T > k - 1) - P(T > k) = (p - 1)^{k-1} p$$

Otrzymaliśmy **rozkład geometryczny**, który interpretujemy jako liczbę doświadczeń, które należy wykonać, by doczekać się sukcesu. Przy czym doświadczenia wykonujemy według schematu Bernoulliego z prawdopodobieństwem sukcesu p. ■

Parametry wektorów losowych

Wielowymiarowy rozkład normalny

Rozkłady form kwadratowych

Oznaczenia

$$\boldsymbol{X} = (X_1,\ X_2, \dots\ ,\ X_n)'$$
$$\boldsymbol{x} = (x_1,\ x_2, \dots\ ,\ x_n)'$$
$$\boldsymbol{\mu} = (\mu_1,\ \mu_2, \dots\ ,\ \mu_n)'$$

Wartość oczekiwana wektora losowego

$$E(\boldsymbol{X}) = (EX_1,\ EX_2, \dots\ ,\ EX_n)'.$$

Macierz kowariancji wektora losowego

$$\begin{aligned} D^2(\boldsymbol{X}) &= E[(\boldsymbol{X} - E\boldsymbol{X})(\boldsymbol{X} - E\boldsymbol{X})'] \\ &= [\mathrm{Cov}(X_i, X_j)]_{i,j=1,\dots,n} \end{aligned},$$

o ile $D^2 X_i < \infty$ dla każdego $i = 1, \cdots, n$

Zauważmy

$$0 \leq D^2 \left(\sum_{i=1}^{n} t_i X_i \right) = E \left(\sum_{i=1}^{n} t_i (X_i - EX_i) \right)^2$$
$$= \sum_{i,j} t_i t_j \mathrm{Cov}(X_i, X_j)$$

Zatem macierz kowariancji jest symetryczna i nieujemnie określona, co na przykład daje

$$\begin{vmatrix} D^2(X_i) & \mathrm{Cov}(X_i X_j) \\ \mathrm{Cov}(X_i X_j) & D^2 X_j \end{vmatrix} \geq 0$$

a po przekształceniu

$$|\mathrm{Cov}(X_i, X_j)| \leq \sqrt{D^2 X_i \cdot D^2 X_j}$$

i w konsekwencji $|\varrho(X_i, X_j)| \leq 1$

Podstawowe własności

Jeżeli Ajest macierzą $p \times n$, $B-$macierzą $n \times n$ to

$E(A\boldsymbol{X}) = AE(\boldsymbol{X}), \quad E(A\boldsymbol{X}B) = AE(\boldsymbol{X})B$
$D^2(A\boldsymbol{X}) = AD^2(\boldsymbol{X})A'$

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

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]$$

Niech $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$, A – macierz $(n \times n)$ nieosobliwa oraz $Y = A\boldsymbol{X}$.

$$P(\boldsymbol{Y} \in B) = P(\boldsymbol{X} \in A^{-1}B) = \int\limits_{A^{-1}B} f(\boldsymbol{x})\, d\boldsymbol{x}$$

$$= \int\limits_{B} f(A^{-1}\boldsymbol{y}) \Big| \, |A^{-1}| \, \Big| \, d\boldsymbol{y}$$

Łatwo sprawdzić, że $g(\boldsymbol{y}) := f(A^{-1}\boldsymbol{y}) \Big| \, |A^{-1}| \, \Big|$ jest gęstością rozkładu $N(A\boldsymbol{\mu}, A\Sigma A')$

Niech teraz A_1 – macierz $(k \times n)$, $r(A_1) = k$. Bierzemy macierz A_2 taką, że

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$$

jest macierzą $(n \times n)$ nieosobliwą.

Mamy $\boldsymbol{Y} = A\boldsymbol{X} = \begin{pmatrix} A_1\boldsymbol{X} \\ A_2\boldsymbol{X} \end{pmatrix} \sim N(A\boldsymbol{\mu}, A\Sigma A')$, gdzie

$$A\mu = \begin{pmatrix} A_1\boldsymbol{\mu} \\ A_2\boldsymbol{\mu} \end{pmatrix} \quad A\mu = \begin{pmatrix} A_1\Sigma A_1' & A_1\Sigma A_2' \\ A_2\Sigma A_1' & A_2\Sigma A_2' \end{pmatrix}$$

Zatem $A_1\boldsymbol{X} \sim N(A_1\boldsymbol{\mu}, A_1\Sigma A_1')$ o ile zachodzi taki

Fakt. Jeżeli $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \Sigma)$, gdzie

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

to $\boldsymbol{Y}_1 \sim N(\boldsymbol{\mu}_1, \Sigma_{11})$ oraz $\boldsymbol{Y}_2 \sim N(\boldsymbol{\mu}_2, \Sigma_{22})$

Niech $f(\boldsymbol{y}_1, \boldsymbol{y}_2)$ oznacza gęstość rozkładu $N(\boldsymbol{\mu}, \Sigma)$. Przedstawimy tę funkcję w postaci

$$f(\boldsymbol{y}_1, \boldsymbol{y}_2) = g(\boldsymbol{y}_1|\boldsymbol{y}_2)f_2(\boldsymbol{y}_2)$$

Przyjmijmy $\boldsymbol{Y}_1\ (m \times 1), \boldsymbol{Y}_2\ (k \times 1)$ oraz oznaczmy

$$\Sigma^{-1} = R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$$

Wtedy $(\boldsymbol{y}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) = (\boldsymbol{y}-\boldsymbol{\mu})'R(\boldsymbol{y}-\boldsymbol{\mu}) =$

$$= (\boldsymbol{y}-\boldsymbol{\mu})' \begin{bmatrix} R_{11}(\boldsymbol{y}_1-\boldsymbol{\mu}_1) + R_{12}(\boldsymbol{y}_2-\boldsymbol{\mu}_2) \\ R_{21}(\boldsymbol{y}_1-\boldsymbol{\mu}_1) + R_{22}(\boldsymbol{y}_2-\boldsymbol{\mu}_2) \end{bmatrix}$$

$$= (\boldsymbol{y}_1-\boldsymbol{\mu}_1)'[R_{11}(\boldsymbol{y}_1-\boldsymbol{\mu}_1) + R_{12}(\boldsymbol{y}_2-\boldsymbol{\mu}_2)] + \\ +(\boldsymbol{y}_2-\boldsymbol{\mu}_2)'[R_{21}(\boldsymbol{y}_1-\boldsymbol{\mu}_1) + R_{22}(\boldsymbol{y}_2-\boldsymbol{\mu}_2)]$$

$$= (\boldsymbol{y}_1-\boldsymbol{\mu}_1)'R_{11}(\boldsymbol{y}_1-\boldsymbol{\mu}_1) + (\boldsymbol{y}_1-\boldsymbol{\mu}_1)'R_{12}(\boldsymbol{y}_2-\boldsymbol{\mu}_2) + \\ +(\boldsymbol{y}_2-\boldsymbol{\mu}_2)'R_{22}(\boldsymbol{y}_2-\boldsymbol{\mu}_2) + (\boldsymbol{y}_2-\boldsymbol{\mu}_2)'R_{21}(\boldsymbol{y}_1-\boldsymbol{\mu}_1)$$

(liczba)$'$ =liczba oraz $(ABC)' = C'B'A', \quad R'_{12} = R_{12}$

$$= (\boldsymbol{y}_1-\boldsymbol{\mu}_1)'R_{11}(\boldsymbol{y}_1-\boldsymbol{\mu}_1) + 2(\boldsymbol{y}_1-\boldsymbol{\mu}_1)'R_{12}(\boldsymbol{y}_2-\boldsymbol{\mu}_2) + \\ +(\boldsymbol{y}_2-\boldsymbol{\mu}_2)'R_{21}(\boldsymbol{y}_1-\boldsymbol{\mu}_1)$$

$$
\begin{aligned}
&= \boldsymbol{y}_1' R_{11} \boldsymbol{y}_1 - 2\boldsymbol{y}_1' R_{11} \boldsymbol{\mu}_1 + \boldsymbol{\mu}_1' R_{11} \boldsymbol{\mu}_1 + 2\boldsymbol{y}_1' R_{12}(\boldsymbol{y}_2 - \mu_2) - \\
&\qquad -2\boldsymbol{\mu}_1' R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2) + (\boldsymbol{y}_2 - \boldsymbol{\mu}_2)' R_{21}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)
\end{aligned}
$$

$$
\begin{aligned}
&= \boldsymbol{y}_1' R_{11} \boldsymbol{y}_1 - 2\boldsymbol{y}_1' (R_{11} \boldsymbol{\mu}_1 - R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2)) + \boldsymbol{\mu}_1' R_{11} \boldsymbol{\mu}_1 - \\
&\qquad -2\boldsymbol{\mu}_1' R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2) + (\boldsymbol{y}_2 - \boldsymbol{\mu}_2)' R_{21}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)
\end{aligned}
$$

$$
\begin{aligned}
&= \boldsymbol{y}_1' R_{11} \boldsymbol{y}_1 - 2\boldsymbol{y}_1' R_{11} (\boldsymbol{\mu}_1 - R_{11}^{-1} R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2)) + \boldsymbol{\mu}_1' R_{11} \boldsymbol{\mu}_1 - \\
&\qquad -2\boldsymbol{\mu}_1' R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2) + (\boldsymbol{y}_2 - \boldsymbol{\mu}_2)' R_{21}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)
\end{aligned}
$$

$$
\boldsymbol{\mu}_1^\star := \boldsymbol{\mu}_1 - R_{11}^{-1} R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2)
$$

$$
\begin{aligned}
&= \boldsymbol{y}_1' R_{11} \boldsymbol{y}_1 - 2\boldsymbol{y}_1' R_{11} \boldsymbol{\mu}_1^\star + \boldsymbol{\mu}_1' R_{11} \boldsymbol{\mu}_1 - \\
&\qquad -2\boldsymbol{\mu}_1' R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2) + (\boldsymbol{y}_2 - \boldsymbol{\mu}_2)' R_{21}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)
\end{aligned}
$$

$$
\begin{aligned}
(\boldsymbol{\mu}^\star)' R_{11} \boldsymbol{\mu}^\star &= \boldsymbol{\mu}_1' R_{11} \boldsymbol{\mu}_1 - 2\boldsymbol{\mu}_1' R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2) + \\
&\quad + (\boldsymbol{y}_2 - \boldsymbol{\mu}_2)' R_{21} R_{11}^{-1} R_{12}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)
\end{aligned}
$$

$$= \boldsymbol{y}_1' R_{11} \boldsymbol{y}_1 - 2\boldsymbol{y}_1' R_{11} \boldsymbol{\mu}_1^\star + (\boldsymbol{\mu}_1^\star)' R_{11} \boldsymbol{\mu}_1^\star +$$
$$+ (\boldsymbol{y}_2 - \boldsymbol{\mu}_2)'(R_{22} - R_{21} R_{11}^{-1} R_{12})(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)$$

$$= (\boldsymbol{y}_1 - \boldsymbol{\mu}^\star)' R_{11} (\boldsymbol{y}_1 - \boldsymbol{\mu}^\star) +$$
$$+ (\boldsymbol{y}_2 - \boldsymbol{\mu}_2)'(R_{22} - R_{21} R_{11}^{-1} R_{12})(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)$$

Zatem możemy przyjąć

$$f(\boldsymbol{y}_1 | \boldsymbol{y}_2) =$$
$$= \frac{1}{\sqrt{(2\pi)^k |R_{11}^{-1}|}} \exp\left[-\frac{1}{2} (\boldsymbol{y}_1 - \boldsymbol{\mu}^\star)' R_{11} (\boldsymbol{y}_1 - \boldsymbol{\mu}^\star) \right]$$

$$f(\boldsymbol{y}_2) = \frac{1}{\sqrt{(2\pi)^m |R_{11}^{-1}|^{-1} |\Sigma|}} \times$$
$$\times \exp\left[-\frac{1}{2} (\boldsymbol{y}_2 - \boldsymbol{\mu}_2)'(R_{22} - R_{21} R_{11}^{-1} R_{12})(\boldsymbol{y}_1 - \boldsymbol{\mu}_1) \right]$$

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = \begin{pmatrix} I & O \\ O & I \end{pmatrix}$$

$$\begin{aligned} \Sigma_{11}R_{11} + \Sigma_{12}R_{21} &= I \\ \Sigma_{21}R_{11} + \Sigma_{22}R_{21} &= O \\ &\Downarrow \\ \Sigma_{11} + \Sigma_{12}\underline{R_{21}}R_{11}^{-1} &= R_{11}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{21}R_{11} &= \underline{R_{21}} \\ &\Downarrow \\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} &= R_{11}^{-1} \end{aligned}$$

Ponadto

$$\begin{aligned} \boldsymbol{\mu}_1^{\star} &= \boldsymbol{\mu}_1 - R_{11}^{-1}R_{12}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2) \\ &= \boldsymbol{\mu}_1 - R_{11}^{-1}(-\Sigma_{22}^{-1}\Sigma_{21}R_{11})'(\boldsymbol{y}_2 - \boldsymbol{\mu}_2) \\ &= \boldsymbol{\mu}_1 + \Sigma_{22}^{-1}\Sigma_{21}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

$$\Sigma_{21}R_{12} + \Sigma_{22}R_{22} = I$$

$$\Downarrow$$

$$\underline{\Sigma_{21}R_{12}} = I - \Sigma_{22}R_{22}$$

Ponieważ

$$-\Sigma_{22}^{-1}\Sigma_{21}R_{11} = R_{21}$$

Zatem

$$-\Sigma_{22}^{-1}\underline{\Sigma_{21}R_{12}} = R_{21}R_{11}^{-1}R_{12}$$

$$-\Sigma_{22}^{-1}(I - \Sigma_{22}R_{22}) = R_{21}R_{11}^{-1}R_{12}$$

$$R_{22} - R_{21}R_{11}^{-1}R_{12} = \Sigma_{22}^{-1}$$

Otrzymaliśmy

$$Y_2 \sim N(\boldsymbol{\mu}_2, \Sigma_{22})$$

oraz

$$Y_1|Y_2 \sim N(\boldsymbol{\mu}_1 + \Sigma_{22}^{-1}\Sigma_{21}(\boldsymbol{y}_2 - \boldsymbol{\mu}_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$$

Przykład. Niech $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$, gdzie $\Sigma = (\sigma_{ij})$. Przyjmijmy $A = [1, 0, \dots, 0]$. Mamy

$$X_1 = AX \sim N(\mu_1, \sigma_{11})$$

Analogicznie

$$\begin{aligned} X_2 &\sim N(\mu_2, \sigma_{22}) \\ X_3 &\sim N(\mu_3, \sigma_{33}) \\ &\dots\dots \\ X_n &\sim N(\mu_n, \sigma_{nn}) \end{aligned}$$

Zatem $E\boldsymbol{X} = \boldsymbol{\mu}$ oraz $D^2 X_i = \sigma_{ii}$

Ponadto dla

$$e_{ij} = [\dots, 1, \dots, 1, \dots]$$

jedynka na i – tym oraz j – ym miejscu,

na pozostałych zera,

mamy $e_{ij}\boldsymbol{X} \sim N(\mu_i + \mu_j, e_{ij}\Sigma e'_{ij} = \sigma_{ii} + \sigma_{jj} + 2\sigma_{ij})$.

Ponieważ $D^2(X_i + X_j) = D^2 X_i + D^2_j + 2\mathrm{Cov}(X_i, X_j)$, to $\mathrm{Cov}(X_i, X_j) = \sigma_{ij}$.

Stąd $D^2\boldsymbol{X} = \Sigma$ ■

Przykład. Niech $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \sigma^2 I)$ oraz niech A będzie macierzą ortogonalną. Wtedy

$$\boldsymbol{Y} = A\boldsymbol{X} \sim N(A\boldsymbol{\mu}, A\sigma^2 I A' = \sigma^2 I).$$

Zatem jeżeli $X_1,\ X_2, \dots,\ X_n$ niezależne o rozkładzie $N(\ \cdot\ , \sigma^2)$, to $Y_1,\ Y_2, \dots,\ Y_n$ też są niezależne. ■

Twierdzenie Fishera – Cochrana.

Niech $\boldsymbol{X} \sim N(0, I)$ oraz $Q_1, \dots, Q_k$ będą formami kwadratowymi rzędu $n_1, \cdots, n_k$ takimi, że

$$\boldsymbol{X}'\boldsymbol{X} = Q_1 + \dots, Q_k$$

Wówczas warunkiem koniecznym i dostatecznym na to, by zmienne losowe Q_i miały rozkłady $\chi^2(n_i)$ i były niezależne, jest $\sum n_i = n$

Dowód. Istnieje $B_i = \begin{bmatrix} w_1^i \\ \cdots \\ w_{n_i}^i \end{bmatrix}$ $(n \times n_i)$ taka, że

$$Q_i = \pm(w_1^i \boldsymbol{X})^2 \pm \cdots \pm (w_{n_i}^i \boldsymbol{X})^2$$

Ponieważ $n = \sum n_i$, to przyjmując $B = \begin{bmatrix} B_1 \\ \cdots \\ B_{n_k} \end{bmatrix}$ mamy

$$\boldsymbol{X}'\boldsymbol{X} = \sum Q_i = \boldsymbol{X}'B'\Delta B\boldsymbol{X}$$

gdzie

$$\Delta = \begin{bmatrix} \pm 1 & & & \\ & \ddots & & \\ & & \pm 1 & \\ & & & \ddots \end{bmatrix}$$

Ponieważ $\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}'B'\Delta B\boldsymbol{X}$ jest spełniona dla dowolnych $\boldsymbol{X}$, mamy

$$I = B'\Delta B$$

Ponieważ $n = r(I) = r(B'\Delta B) \leq r(B) \leq n$, to B jest macierzą nieosobliwą. Zatem $\Delta = (B^{-1})'B^{-1}$ jest macierzą dodatnio określoną. W konsekwencji $\Delta = I$ oraz macierz B jest ortogonalna. Zatem dla $\boldsymbol{Y} = B\boldsymbol{X}$

$$\begin{aligned} &Q_1 = y_1^2 + \ldots + y_{n_1}^2 \\ &Q_2 = y_{n_1+1}^2 + \ldots + y_{n_1+n_2}^2 \\ &\ldots\ldots \end{aligned}$$

są niezależne i mają rozkłady $\chi^2(n_1), \chi^2(n_2), \ldots$.
W ten sposób została udowodniona dostateczność warunku. Konieczność jest oczywista. ■

Twierdzenie 16. Warunkiem koniecznym i dostatecznym na to, aby zmienna losowa $\boldsymbol{X}'A\boldsymbol{X}$ miała rozkład $\chi^2(\nu)$ jest, by macierz A była idempotenta. Wtedy $\nu = r(A) = tr(A)$.

Dowód.
dostateczność

$$\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}'A\boldsymbol{X} + \boldsymbol{X}'(I - A)\boldsymbol{X}$$
$$A^2 = A \Leftrightarrow r(A) + r(I - A) = n$$

Teza wynika z tw. F–C.

konieczność
Istnieje macierz ortogonalna C, że przy przekształceniu $\boldsymbol{X} = C\boldsymbol{Y}$

$$\boldsymbol{X}'A\boldsymbol{X} = \boldsymbol{Y}'C'AC\boldsymbol{Y} = \lambda_1 y_1^2 + \ldots + \lambda_m y_m^2$$

gdzie λ_i są niezerowymi wartościami własnymi macierzy A.

Ponieważ $Y_i \sim \chi^2(1)$, więc $m = k$, $\lambda_i = 1$. Zatem $C'AC$ jest macierzą diagonalną o elementach 0 lub 1. Zatem A jest idempotentna, bo

$$C'AC = C'ACC'AC = C'A^2C \quad \Rightarrow \quad A = A^2$$

■

Twierdzenie 17. Jeżeli $\boldsymbol{X}'\boldsymbol{X} = Q_1 + Q_2$ oraz $Q_1 \sim \chi^2(k)$, to $Q_2 \sim \chi^2(n-k)$

Dowód. Niech $Q_1 = \boldsymbol{X}'A\boldsymbol{X}$. Wtedy

$$Q_2 = \boldsymbol{X}'(I - A)\boldsymbol{X}$$

oraz $(I-A)^2 = I^2 + A^2 - IA - AI = I - A$. ■

Twierdzenie 18. Jeżeli $Q = Q_1 + Q_2$, $Q \sim \chi^2(m)$, $Q_1 \sim \chi^2(k)$ oraz $Q_2 \geq 0$, to $Q_2 \sim \chi^2(m-k)$.

Dowód. Niech $Q = \boldsymbol{X}'A\boldsymbol{X}, \quad Q_1 = \boldsymbol{X}'B\boldsymbol{X}$

$$Q = \boldsymbol{X}'A\boldsymbol{X} = \boldsymbol{X}'B\boldsymbol{X} + \boldsymbol{X}'(A-B)\boldsymbol{X}$$

Istnieje macierz ortogonalna C, że przy przekształceniu $\boldsymbol{X} = C\boldsymbol{Y}$

$$\begin{aligned}
\boldsymbol{X}'A\boldsymbol{X} &= \boldsymbol{Y}'C'AC\boldsymbol{Y} = Y_1^2 + \ldots + Y_m^2 \\
\boldsymbol{X}'B\boldsymbol{X} &= \boldsymbol{Y}'C'BC\boldsymbol{Y} \\
\boldsymbol{X}'(I-B)\boldsymbol{X} &= \boldsymbol{Y}'C'(A-B)C\boldsymbol{Y}
\end{aligned}$$

Oznaczmy $B_1 := C'AC, \quad B_2 := C'(A-B)C$. Zatem

$$Y_1^2 + Y_2^2 + \ldots + Y_m^2 = \boldsymbol{Y}'B_1\boldsymbol{Y} + \boldsymbol{Y}'B_2\boldsymbol{Y}$$

Ponieważ $\boldsymbol{Y}'B_1\boldsymbol{Y}$, $\boldsymbol{Y}'B_2\boldsymbol{Y}$ są nieujemne, każda z form zawiera wyłącznie elementy $Y_1, \ldots, Y_m$. Z poprzedniego twierdzenia $Q_2 \sim \chi^2(m-k)$ ■

Twierdzenie 19. Niech

$$\boldsymbol{X}'A_1\boldsymbol{X} \sim \chi^2(k), \ \boldsymbol{X}'A_2\boldsymbol{X} \sim \chi^2(m).$$

Warunkiem koniecznym i dostatecznym na to, aby obie zmienne losowe były niezależne, jest, by

$$A_1A_2 = 0$$

Dowód. Ponieważ

$$A_1(I - A_1 - A_2) = A_2(I - A_1 - A_2) = 0,$$

to

$$r(A_1) + r(A_2) + r(I - A_1 - A_2) = n$$

Ale

$$\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}'A_1\boldsymbol{X} + \boldsymbol{X}'A_2\boldsymbol{X} + \boldsymbol{X}'(I - A_1 - A_2)\boldsymbol{X}$$

Zatem z tw. F – C są niezależne.

Jeżeli są niezależne, to $\boldsymbol{X}'(A_1 + A_2)\boldsymbol{X} \sim \chi^2(k+m)$

Zatem $A_1 A_2 = 0$, ponieważ

$$A_1 + A_2 = (A_1 + A_2)^2 = A_1 + A_2 + A_1 A_2 + A_2 A_1$$
$$0 = A_1 A_2 + A_2 A_1$$

$$\begin{cases} A_1 0 = A_1 A_2 + A_1 A_2 A_1 \\ 0 A_1 = A_1 A_2 A_1 + A_2 A_1 \end{cases}$$

dodaję stronami

$$0 = 2A_1 A_2 A_1$$
podstawiam $A_2 A_1 = -A_1 A_2$.
$$0 = -2A_1 A_2$$

■

Twierdzenie 20. Niech

$$\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}'A_1\boldsymbol{X} + \ldots + \boldsymbol{X}'A_k\boldsymbol{X}.$$

Każdy z następujących warunków jest warunkiem koniecznym i dostatecznym na to, aby zmienne losowe

$$\boldsymbol{X}'A_1\boldsymbol{X}, \ldots, \boldsymbol{X}'A_k\boldsymbol{X}$$

były niezależne i aby $\boldsymbol{X}'A_i\boldsymbol{X} \sim \chi^2(n_i)$, gdzie n_i jest rzędem macierzy A_i:

(a) Macierze $A_1, \ldots, A_k$ są idempotentne,

(b) $A_iA_j = 0$ dla wszystkich $i \neq j$.

Dowód. Ponieważ $A_i^2 = A_i$, więc $tr(A_i) = r(A_i)$. Ale

$$I = A_1 + \ldots A_k,$$

więc

$$tr(I) = n = \sum tr(A_i) = \sum n_i$$

i z tw. F – C wynika konieczność i dostateczność warunku (a). Ponieważ

$$I = I = A_1 + \ldots A_k \text{ oraz } A_iA_j = 0,$$

więc $A_i(I - A_i) = 0$, zatem macierze A_i są indempotentne. Wynika stąd, że $(a) \Leftrightarrow (b)$ ∎

Twierdzenie 21. Niech $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$. Warunkiem koniecznym i dostatecznym na to, aby zmienna losowa

$$(\boldsymbol{X} - \boldsymbol{\mu})' A (\boldsymbol{X} - \boldsymbol{\mu})$$

miała rozkład χ^2, jest, by

$$A \Sigma A = A.$$

Liczba stopni swobody jest wtedy równa $tr(A\Sigma)$.

Dowód.

$$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma) \Rightarrow (\boldsymbol{X} - \boldsymbol{\mu}) \sim N(0, \Sigma) \Rightarrow$$
$$\boldsymbol{Y} = B^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \sim N(0, I)$$
$$\text{gdzie } \Sigma = BB',\ B \text{ nieosobliwa}$$

Zatem

$$\boldsymbol{X} = \boldsymbol{\mu} + B\boldsymbol{Y},\ \boldsymbol{Y} \sim N(0, I)$$
$$(\boldsymbol{X} - \boldsymbol{\mu})' A (\boldsymbol{X} - \boldsymbol{\mu}) = \boldsymbol{Y}' B' A B \boldsymbol{Y}$$

Zatem

$$(\boldsymbol{X} - \boldsymbol{\mu})' A (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2 \Leftrightarrow B'AB \text{ idempotentna}$$

$$B'AB \cdot B'AB = B'AB \Leftrightarrow A\Sigma A = A$$

Ponadto $tr(B'AB) = tr(ABB') = tr(A\Sigma)$ ■

Twierdzenie 22. Niech $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$. Warunkiem koniecznym i dostatecznym na to, aby zmienne losowe

$$P'\boldsymbol{X} \text{ oraz } (\boldsymbol{X}-\boldsymbol{\mu})'A(\boldsymbol{X}-\boldsymbol{\mu})$$

były niezależne, jest, by

$$A\Sigma P = 0$$

Twierdzenie 23. Niech $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$. Warunkiem koniecznym i dostatecznym na to, aby zmienne losowe

$$(\boldsymbol{X}-\boldsymbol{\mu})'A(\boldsymbol{X}-\boldsymbol{\mu}) \text{ oraz } (\boldsymbol{X}-\boldsymbol{\mu})'B(\boldsymbol{X}-\boldsymbol{\mu})$$

były niezależne, jest, by

$$A\Sigma B = 0$$