Procesy stochastyczne

WYKŁAD 0

Literatura

- A. Plucińska, E. Pluciński, Probabilistyka (WNT), 2009
- D. Bobrowski, Probabilistyka w zastosowaniach technicznych
- O. Tikhonenko, Metody probabilistyczne analizy systemów informacyjnych, 2006
- M. Dędys, S. Dorosiewicz, Procesy stochastyczne, 2005
- L. Kowalski, Statystyka, skrypt WAT, 2021
- Strona: Statystyka.rezolwenta.eu.org

ZMIENNA LOSOWA I JEJ PARAMETRY -powtórzenie

(Ω, S, P) – **przestrzeń probabilistyczna**
(matematyczny model zjawiska losowego),

Ω – zbiór wszystkich zdarzeń elementarnych,
S – zbiór zdarzeń, (podzbiory zbioru Ω,
(dokładnie σ – ciało podzbiorów)),
P – prawdopodobieństwo (funkcja przyporządkowująca zdarzeniom szansę ich zajścia).

$$P : S \to R$$

Prawdopodobieństwo warunkowe

Jeśli $P(B) > 0,\ B \in S$ to określamy **prawdopodobieństwo warunkowe** dowolnego zdarzenia *A* pod warunkiem, że zaszło zdarzenie *B*:

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad A \in S}$$

Twierdzenie (o prawdopodobieństwie całkowitym)

Niech zdarzenia $A_1, A_2, ..., A_n$, parami wykluczają się, oraz $A_1 \cup A_2 \cup ... \cup A_n = \Omega$ i $P(A_i) > 0$, i = 1,2, ..., n. wtedy dla dowolnego zdarzenia B

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

Zmienną losową *X* nazywamy funkcję (borelowską czyli praktycznie każdą) przyporządkowującą zdarzeniom elementarnym liczby rzeczywiste.

$$X : \Omega \longrightarrow R$$

Dystrybuantą zmiennej losowej *X* nazywamy funkcję $F\colon R \longrightarrow R$ określoną wzorem:

$$F(x) = P(X < x) = P_X((-\infty, x))$$

Własności dystrybuanty:

a) F jest funkcją niemalejącą,

b) F jest funkcją lewostronnie ciągłą,

c) $F(-\infty)=0; \quad F(\infty)=1$,

d) dystrybuanta zmiennej losowej wyznacza jednoznacznie jej rozkład,

e) $P(a \leq X < b) = F(b) - F(a); \quad a < b$

f) $P(X = a) = F(a^{+}) - F(a);$ gdzie $F(a^{+})$ oznacza granicę prawostronną, (jeśli a jest punktem ciągłości dystrybuanty to $P(X = a) = 0$).

Zmienna losowa jest **skokowa** (**dyskretna**) jeśli zbiór wszystkich jej wartości jest skończony lub przeliczalny.
Rozkład zmiennej losowej skokowej często określamy za pomocą **funkcji prawdopodobieństwa:**

$$P(X = x_k) = p_k$$

(własność: $\sum_k p_k = 1;\ p_k > 0$)

Liczby p_k nazywamy **skokami**, a wartości x_k **punktami skokowymi**.

Zmienna losowa X o dystrybuancie F jest **ciągła** jeśli jej dystrybuanta da się przedstawić w postaci

$$F(x) = \int_{-\infty}^{x} f(t)dt \qquad x \in R$$

gdzie f jest funkcją spełniającą warunki:

$$f(x) \geq 0; \quad x \in R; \qquad \int_{-\infty}^{\infty} f(t)dt = 1$$

i nazywamy ją **gęstością prawdopodobieństwa** zmiennej losowej X.

Własności rozkładu zmiennej losowej często charakteryzujemy jej **parametrami**.

Jednym z podstawowych parametrów jest wartość oczekiwana.

Wartość oczekiwana. Oznaczenie *EX* lub *m*.

Dla zmiennej losowej skokowej

$$EX = \sum_i x_i p_i$$

Dla zmiennej losowej ciągłej

$$EX = \int_{-\infty}^{\infty} x f(x) dx$$

Przykład

Dla zmiennej losowej o funkcji prawdopodobieństwa

x_k	-1	2	3
p_k	0,2	0,6	0,2

$$EX = -1 \cdot 0{,}2 + 2 \cdot 0{,}6 + 3 \cdot 0{,}2 = 1{,}6 \ .$$

Przykład

Dla zmiennej losowej o gęstości

$$f(x)=\begin{cases}2x & x\in<0,1>\\ 0 & x\notin<0,1>\end{cases}$$

$$EX=\int_0^1 x\cdot 2x dx=2\int_0^1 x^2 dx=2\frac{x^3}{3}\Big|_0^1=\frac{2}{3}$$

Własności wartości oczekiwanej

a) $Ec = c; \quad c$ – stała,

b) $E(aX) = aE(X)$,

c) $E(X + Y) = EX + EY$,

d) Jeśli $a \leq X \leq b$, to $a \leq EX \leq b$,

e) Jeśli $X \leq Y$, to $EX \leq EY$,

f) $EX \leq E|X|$, $|EX| \leq E|X|$

g) X, Y – niezależne, to $E(XY) = EX \cdot EY$.

h) X, Y – **dowolne**, to $E(XY) = EX \cdot EY + cov(X,Y)$.

Miarą rozrzutu wartości zmiennej losowej jest wariancja.

Wariancja. Oznaczenie D^2X lub σ^2 .

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

Dla zmiennej losowej skokowej

$$D^2 X = \sum (x_i - EX)^2 p_i$$

Dla zmiennej losowej ciągłej

$$D^2 X = \int_{-\infty}^{\infty} (x - EX)^2 f(x) dx$$

Własności wariancji

a) $D^2 c = 0$; c – stała,

b) $D^2(aX) = a^2 D^2(X)$,

c) $D^2(X + b) = D^2 X$, b – stała,

d) X, Y – niezależne, to $D^2(X \pm Y) = D^2 X + D^2 Y$

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

f) X, Y – **dowolne** zmienne losowe
$D^2(X + Y) = D^2 X + D^2 Y + 2Cov(X, Y)$,

Odchylenie standardowe (K. Pearson).

Oznaczenie *DX* lub σ.

$$DX = \sqrt{D^2 X}$$

Podstawowe rozkłady.

Rozkład dwupunktowy (zerojedynkowy)

Niech $p \in (0, 1)$ będzie ustaloną liczbą.
Określamy:
$P(X = 0) = q$, $P(X = 1) = p$; gdzie $q = 1 - p$.

Umowa: 0 – porażka, 1 - sukces

$$EX = p, \qquad D^2X = pq$$

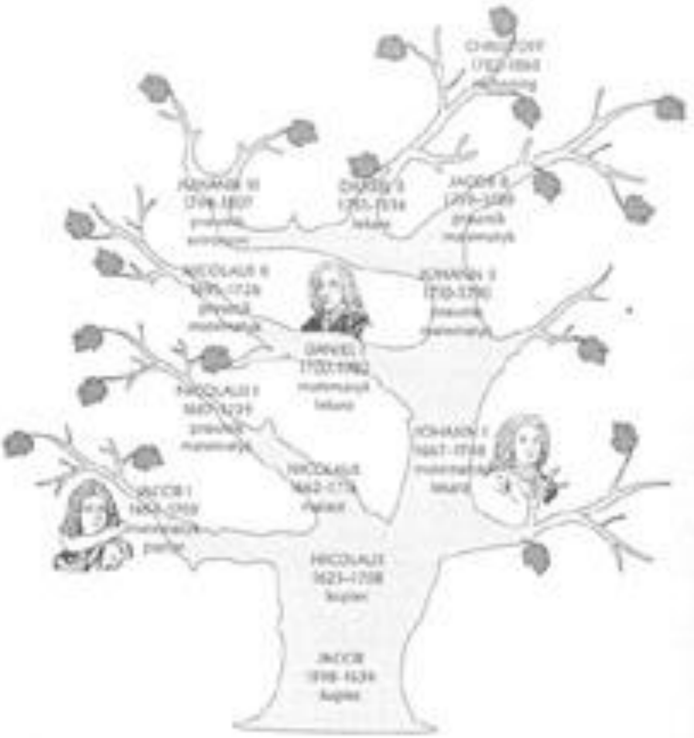
Rozkład dwumianowy

Dla danych 0 < p < 1 , $n \in N$ określamy funkcję prawdopodobieństwa

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \qquad \text{gdzie } q = 1 - p$$

$$k = 0, 1, 2, \ldots, n.$$

Zauważmy, że gdy $n = 1$ to rozkład dwumianowy jest rozkładem zerojedynkowym.



Płyta nagrobna Jakuba Bernoulliego [1654, 1705] w katedrze w Bazylei (Szwajcaria)

Rozkład geometryczny

X - liczba prób Bernoulliego **poprzedzających** pierwszy sukces

$$\boxed{P(X = k) = pq^k}$$

$q = 1 - p$ $\qquad$ $k = 0, 1, 2, \ldots$

$$EX = q/p; \qquad D^2X = q/p^2$$

Dystrybuanta

F(x) = 0 dla x ≤ 0

dla k < x ≤ k + 1

$$F(x) = P(X < x) = \sum_{i=0}^{k} P(X = i) = \sum_{i=0}^{k} pq^i =$$

$$= \frac{p(1-q^{k+1})}{1-q} = 1-q^{k+1}$$

(Funkcja schodkowa)

Dla dowolnych całkowitych m, n $\geq$ 0 mamy

$$P(X \geq m+n \mid X \geq m) = P(X \geq n)$$

(własność braku pamięci).

Rozkład Poissona

Dla $\lambda > 0$ określamy funkcję prawdopodobieństwa

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad k = 0, 1, 2, \ldots$$

(wartości tych prawdopodobieństw zawiera tablica rozkładu Poissona)

$$EX = \lambda \qquad D^2X = \lambda$$

Siméon Denis Poisson (1781–1840), francuski mechanik teoretyk, fizyk i matematyk.

Rozkład Poissona jest związany z rozkładem dwumianowym.

X - rozkład dwumianowy o parametrach n, p.
Niech λ = np, wtedy

$$P(X=k)=\binom{n}{k}p^k q^{n-k}=\frac{n!}{(n-k)!k!}p^k(1-p)^{n-k}=$$

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

Dla ustalonego λ i dużych n mamy

$$\left(1-\frac{\lambda}{n}\right)^{n} \approx e^{-\lambda} \qquad \left(1-\frac{\lambda}{n}\right)^{k} \approx 1 \qquad \frac{n(n-1)\cdots(n-k+1)}{n^{k}} \approx 1$$

zatem

$$P(X=k)=\binom{n}{k}p^{k}q^{n-k} \approx \frac{\lambda^{k}}{k!}e^{-\lambda} \qquad \text{(przybliżenie Poissona)}$$

Suma niezależnych rozkładów Poissona.

X, Y – niezależne rozkłady Poissona o parametrach odpowiednio λ_1, $\lambda_2 > 0$.

Wtedy X + Y ma też rozkład Poissona o parametrze $\lambda_1 + \lambda_2$.

Dowód

$$P(X+Y=n)=\sum_{k=0}^{n}P(X=k,Y=n-k)=$$

$$=\sum_{k=0}^{n}P(X=k)(Y=n-k)=\sum_{k=0}^{n}\frac{\lambda_1^k}{k!}e^{-\lambda_1}\frac{\lambda_2^{n-k}}{(n-k)!}e^{-\lambda_2}=$$

$$=e^{-(\lambda_1+\lambda_2)}\sum_{k=0}^{n}\frac{\lambda_1^k}{k!}\frac{\lambda_2^{n-k}}{(n-k)!}=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k}=$$

$$=\frac{(\lambda_1+\lambda_2)^n}{n!}e^{-(\lambda_1+\lambda_2)}$$

Rozkłady ciągłe

Rozkład jednostajny

Rozkład którego gęstość jest stała w pewnym przedziale nazywamy **jednostajnym**.

Gęstość rozkładu jednostajnego w (a, b)

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

Ponieważ gęstość ta ma oś symetrii w punkcie $x = (a + b)/2$ to $EX = (a+b)/2$

$$D^2X = (b - a)^2/12$$

Rozkład wykładniczy

Rozkład ten występuje często w zagadnieniach rozkładu czasu między zgłoszeniami (awariami) lub czasu oczekiwania na obsługę w systemach kolejkowych.

Gęstość rozkładu wykładniczego o parametrze a > 0 ma postać

$$f(x) = \begin{cases} ae^{-ax} & x > 0 \\ 0 & x \le 0 \end{cases}$$

dystrybuantą tego rozkładu jest funkcja

$$F(x) = \begin{cases} 1 - e^{-ax} & x > 0 \\ 0 & x \le 0 \end{cases}$$

(uzasadnienie: F'(x) = f(x))

$EX = 1/a$ $\qquad$ $D^2X = 1/a^2$

Własność.

1) Jeśli liczba zgłoszeń w systemie kolejkowym w przedziale czasu (t, t + T) ma rozkład Poissona o parametrze λT, oraz liczby zgłoszeń przychodzące w rozłącznych przedziałach czasu są niezależne to czas X między kolejnymi zgłoszeniami ma rozkład wykładniczy o parametrze a = 1/λ.
2) Dla dowolnych t, T > 0 mamy $P(X \geq t+T \mid X \geq t) = P(X \geq T)$ (własność braku pamięci).

Rozkład normalny

Dla $m \in R, \ \sigma \in (0, +\infty)$

Określamy gęstość rozkładu

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \qquad x \in R$$

Wartości dystrybuanty dla argumentów ujemnych wyznaczamy na podstawie zależności

$$\Phi(-x) = 1 - \Phi(x)$$

Carl Friedrich Gauss (1777 - 1855) – niemiecki matematyk i fizyk. Jego badania związane z teorią błędów doprowadziły do odkrycia rozkładu normalnego zmiennej losowej (nazywany także rozkładem Gaussa), który jest najważniejszym rozkładem w teorii prawdopodobieństwa.


GD9674175N9
DEUTSCHE BUNDESBANK
Banknote
10
ZEHN DEUTSCHE MARK
1777–1855 Carl Friedr. Gauß
Deutsche Bundesbank
Frankfurt am Main
1.Oktober 1993
GD9674175N9

Uwaga
Jeśli X ma rozkład $N(m, \sigma)$ to zmienna losowa $Y = (X - m)/\sigma$ ma rozkład $N(0, 1)$ (takie przekształcenie nazywamy **standaryzacją**).

ROZKŁAD ZMIENNEJ LOSOWEJ n-WYMIAROWEJ. CIĄGI LOSOWE

(Ω, S, P)- ustalona przestrzeń probabilistyczna.

X = (X_1, X_2, ..., X_n) - zmienna losowa n - wymiarowa (wektor losowy, ciąg losowy).

$X : \Omega \to R^n$ (funkcja borelowska)

$P_X : \mathrm{B}(R^n) \to [0,\ 1]$ - rozkład zmiennej losowej X.

Dystrybuanta

$$F(x_1,\ \ ...,x_n) = P\left(X_1 < x_1,\ \ ...,X_n < x_n\right)$$

X nazywamy **zmienną losową skokową** jeśli jej zbiór wartości jest skończony lub przeliczalny.

X nazywamy **zmienną losową ciągłą** jeśli jej dystrybuanta da się przedstawić w postaci

$$F(x_1, \ ..., x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(u_1, \ ..., u_n) du_1 ... du_n$$

dla pewnej nieujemnej funkcji f zwanej gęstością.

Uwaga.

1.W punktach ciągłości funkcji f zachodzi:

$$\frac{\partial^{(n)} F(x_1, \ ..., x_n)}{\partial x_1 \ ...\partial x_n} = f(x_1, \ ..., x_n)$$

2.Dla $A \in \mathrm{B}(R^n)$ mamy $P_X(A) = \int\int\limits_A ... \int f(x_1, \ ..., x_n) dx_1 ... dx_n$.

Funkcja charakterystyczna zmiennej losowej n - wymiarowej.

$$\varphi(t) = \varphi(t_1,\ \ldots, t_n) = E\left(e^{itX}\right) = E\left(\exp(i(t_1 X_1 + \ldots + t_n X_n))\right).$$

Rozkłady warunkowe.

Jeśli $P_{1,\dots,k}(X_1 = x_{1j},\ \dots, X_k = x_{kj}) > 0$ to rozkład zmiennej losowej skokowej (n - k) wymiarowej określonej wzorem:

$$P\,(X_{k+1} = x_{k+1,j},\ \dots, X_n = x_{nj} \mid X_1 = x_{1j},\ \dots, X_k = x_{kj}) = \frac{P\,(X_1 = x_{1j},\ \dots, X_n = x_{nj})}{P_{1,\dots,k}(X_1 = x_{1j},\ \dots, X_k = x_{kj})}$$

nazywamy rozkładem warunkowym zmiennej losowej $\left(X_{k+1},\ \dots, X_n\right)$ pod warunkiem, $\left(X_1 = x_{1j},\ \dots, X_k = x_{kj}\right)$.

Jeśli gęstość $f_{1,\dots,k} > 0$ to rozkład zmiennej losowej ciągłej (n - k) wymiarowej określonej wzorem:

$$f\,(x_{k+1},\ \dots, x_n \mid x_1,\ \dots, x_k) = \frac{f\,(x_1,\ \dots, x_n)}{f\,(x_1,\ \dots, x_k)}$$

nazywamy rozkładem warunkowym zmiennej losowej $\left(X_{k+1},\ \dots, X_n\right)$ pod warunkiem, że $\left(X_1 = x_1,\ \dots, X_k = x_k\right)$.

Niezależność zmiennych losowych.
Zmienne losowe $X_1, X_2, ..., X_n$ są niezależne jeśli

$$F(x_1, \ ..., x_n) = F_1(x_1) \cdot F_2(x_2) ... \cdot F_n(x_n)$$

dla dowolnych $x_1, x_2, ..., x_n \in R^n$.
gdzie F_i - dystrybuanty rozkładów brzegowych jednowymiarowych.

Dla zmiennych losowych skokowych odpowiedni warunek ma postać:

$$P(X_1 = x_{1j}, \ ..., X_n = x_{nj}) = P_1(X_1 = x_{1j}) \cdot ... \cdot P_n(X_n = x_{nj})$$

dla dowolnych $x_{1j}, \ ..., x_{nj} \in R^n$

Dla zmiennych losowych ciągłych odpowiedni warunek ma postać:

$$f(x_1, \ ..., x_n) = f_1(x_1) \cdot f_2(x_2) ... \cdot f_n(x_n)$$

dla dowolnych $x_1, x_2, ..., x_n \in R^n$.

Parametry (mogą nie istnieć)

Wartość oczekiwana $E(X) = \left[EX_1,\ EX_2, ..., EX_n \right]$.

Wariancja $\mathrm{D}^2(X) = \left[\mathrm{D}^2 X_1,\ \mathrm{D}^2 X_2, ..., \mathrm{D}^2 X_n \right]$.

Moment (zwyczajny) rzędu $l_1 + l_2 + ... + l_n$

$$m_{l_1 l_2 \dots l_n} = E\left(X_1^{l_1} X_2^{l_2} \dots X_n^{l_n}\right),$$

Moment centralny rzędu $l_1 + l_2 + ... + l_n$

$$\mu_{l_1 l_2 \dots l_n} = E\left(\left(X_1 - EX_1\right)^{l_1} \dots \left(X_n - EX_n\right)^{l_n}\right),$$

Macierz kowariancji $K = [k_{ij}]$, gdzie

$$k_{ij} = \operatorname{cov}(X_i, X_j) = E\left[\left(X_i - EX_i\right)\left(X_j - EX_j\right)\right] =$$
$$= E\left(X_i X_j\right) - E\left(X_i\right)E\left(X_j\right)$$

Uwaga $k_{ii} = D^2X_i$, jest wariancją i - tej składowej.

Macierz K jest kwadratowa, symetryczna i słabo dodatnio określona (w szczególności ma wyznacznik nieujemny).

Macierz korelacji $R = [\rho_{ij}]$, gdzie

$$\rho_{ij} = \frac{\operatorname{cov}(X_i, X_j)}{DX_i \cdot DX_j}$$

Uwaga

$\rho_{ii} = 1$,

$$\forall i, j \quad -1 \leq \rho_{ij} \leq 1$$

Rozkład normalny n - wymiarowy.

K - macierz kowariancyjna, niech detK $\neq$ 0.

Zmienna losowa n - wymiarowa ma rozkład normalny n - wymiarowy gdy gęstość tej zmiennej losowej wyraża się wzorem:

$$f(x) = f(x_1, x_2, ..., x_n) = \frac{\sqrt{|L|}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \sum_{j,k=1}^{n} l_{jk}(x_j - m_j)(x_k - m_k) \right) =$$

$$= \frac{\sqrt{|L|}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}(x - m)^T L(x - m) \right)$$

gdzie

$m_i = E(X_i)$ dla i = 1, 2, ..., n

L = $[l_{jk}]$ j, k = 1, 2, ..., n jest macierzą odwrotną do K.

Dla n = 2 warunek |K| $\neq$ 0 jest równoważny warunkowi $\rho^2 \neq 1$.

Ponieważ macierz K ma wtedy postać

$$K = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \text{ to}$$

$$L = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}$$

Zatem gęstość rozkładu normalnego 2-wymiarowego N(m_1, m_2, σ_1, σ_2, ρ) można zapisać następująco:

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2} - 2\rho\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2} + \frac{(y-m_2)^2}{\sigma_2^2}\right]\right\}$$

Powyższa funkcja gęstości ma stałą wartość f(x, y) = h na elipsie:

$$\frac{(x-m_1)^2}{\sigma_1^2} - 2\rho\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2} + \frac{(y-m_2)^2}{\sigma_2^2} = const = \lambda^2$$

o środku w punkcie (m_1, m_2).

gdzie $\lambda^2 = -2(1-\rho^2)\ln\left(h2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}\right)$.

Dla $\rho \neq 0$ osie główne mają równania:

$$y - m_2 = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2 \pm \sqrt{\left(\sigma_1^2 - \sigma_2^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2}}\left(x - m_1\right)$$

Dla ρ = 0 osie rozpatrywanej elipsy są równoległe do osi układu współrzędnych.

Zauważmy, że gdy $\rho^2 \to 1$ to jedna oś się wydłuża, a druga skraca, zależność między zmiennymi staje się ściśle liniowa.

Osie powyższej elipsy tworzą z osią OX kąty α i α + π/2 gdzie

$$\operatorname{tg} 2\alpha = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}$$

Funkcja charakterystyczna:

$$\varphi(t) = \exp\left(im^T t - \frac{1}{2} t^T K t \right)$$

gdy n = 2 to

$$\varphi(t_1, t_2) = \exp\left(i(t_1 m_1 + t_2 m_2) - \frac{1}{2}\left(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2\right)\right)$$

Twierdzenie.

Dowolny rozkład brzegowy normalnego rozkładu n-wymiarowego jest rozkładem normalnym.

Twierdzenie.

Jeśli składowe normalnego rozkładu n-wymiarowego są parami nieskorelowane to są niezależne.

Zbieżność ciągów losowych

Zbieżność ciągu zmiennych losowych z prawdopodobieństwem 1 (prawie napewno)

Ciąg zmiennych losowych (X_n) jest zbieżny do zmiennej losowej X z prawdopodobieństwem 1 jeśli

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

Średniokwadratowa zbieżność ciągu zmiennych losowych

Ciąg zmiennych losowych (X_n) jest średniokwadratowo zbieżny do zmiennej losowej X jeśli

$$\lim_{n\to\infty} E\left(\left|X_n - X\right|^2\right) = 0$$

Rozpatrując ten rodzaj zbieżności zakładamy, że dla występujących tu zmiennych losowych (X_n), X istnieje skończony moment rzędu 2.

Niekiedy stosuje się zapis $\text{l.i.m.}\ X_n = X$ (skrót od „limit in mean”).

Stochastyczna zbieżność ciągu zmiennych losowych

Ciąg zmiennych losowych (X_n) jest stochastycznie (wg prawdopodobieństwa) zbieżny do zmiennej losowej X jeśli

$$\bigwedge_{\varepsilon>0} \lim_{n\to\infty} P\left(\left|X_n - X\right| < \varepsilon\right) = 1$$

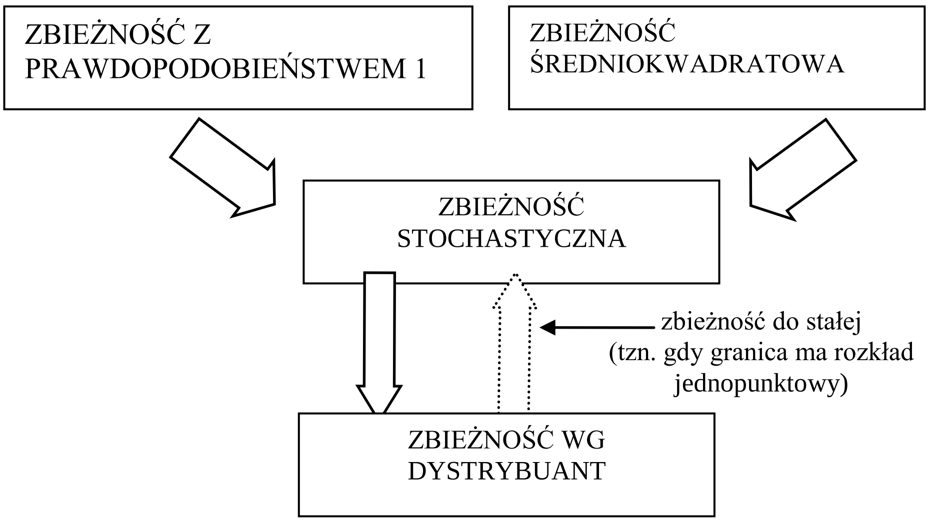
lub równoważnie

$$\bigwedge_{\varepsilon>0} \lim_{n\to\infty} P\left(\left|X_n - X\right| \geq \varepsilon\right) = 0$$

Zbieżność ciągu zmiennych losowych wg dystrybuant (wg rozkładu)

Ciąg zmiennych losowych (X_n) jest zbieżny do zmiennej losowej X wg dystrybuant jeśli ciąg ich dystrybuant F_n jest zbieżny do dystrybuanty F w każdym punkcie jej ciągłości (F jest dystrybuantą zmiennej losowej X).

Zależności miedzy zbieżnościami.



Przykład.
Rozpatrzmy ciąg zmiennych losowych skokowych określonych na przedziale [0, 1) w następujący sposób

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

$$P(X_{kn} = 1) = \frac{1}{n}; \qquad P(X_{kn} = 0) = 1 - \frac{1}{n}$$

Ciąg X_{01}, X_{02}, X_{12}, X_{03}, X_{13}, X_{23}, ... jest zbieżny stochastycznie do zera bo

$$\bigwedge_{0<\varepsilon<1} \lim_{n\to\infty} P\left(\left|X_n\right| \geq \varepsilon\right) = \lim_{n\to\infty} \frac{1}{n} = 0$$

Natomiast ciąg ten nie jest zbieżny w żadnym punkcie przedziale [0, 1) bowiem dla każdego ustalonego punktu otrzymujemy rozbieżny ciąg zer i jedynek (zera i jedynki występują na dowolnie dalekich miejscach).

Przykład.

Ciąg zmiennych losowych X_n ciągłych o rozkładach jednostajnych na przedziałach (0, 1/n) jest zbieżny do rozkładu jednopunktowego X ($P(X = 0) = 1$) wg dystrybuant.

Uwaga.

Punktowa granica ciągu dystrybuant nie musi być dystrybuantą.

Jeśli ciąg funkcji charakterystycznych odpowiadających rozpatrywanemu ciągowi dystrybuant jest punktowo zbieżny do funkcji ciągłej to granica tych dystrybuant jest dystrybuantą.

Przykład

Granica dystrybuant

$$F_n(x) = \begin{cases} 0 & x \le -\mathrm{n} \\ \dfrac{n+x}{2n} & -n < x \le \mathrm{n} \\ 1 & x > \mathrm{n} \end{cases}$$

Nie jest dystrybuantą.

$\lim_{n\to\infty} F_n(x) = 1/2 \qquad x \in \mathrm{R}$

Ciągi zmiennych losowych

(Ω, S, P) - ustalona **przestrzeń probabilistyczna**.

N = {0, 1, 2, ….}.

Def.

Dla każdego n ∊ N określamy zmienną losową

$$X_n : \Omega \to R$$

Otrzymany ciąg (X_n) = X_0, X_1, X_2, …nazywamy **ciągiem losowym.**

Zbiór wartości tych zmiennych losowych nazywamy **zbiorem stanów.**

Dla ustalonego $\omega \in \Omega$ otrzymujemy ciąg liczbowy $(X_n(\omega))$ zwany **realizacją** lub **trajektorią** ciągu losowego.

Realizacje nie mają charakteru losowego

Przykład

X – zmienna losowa o rozkładzie zerojedynkowym

Określamy

$$X_n = \begin{cases} X & \text{gdy n parzyste} \\ -X & \text{gdy n nieparzyste} \end{cases}$$

Jakie realizacje ma ten ciąg losowy?

Momenty ciągu losowego
Zmienne losowe tworzące ciąg losowy (X_n) można charakteryzować za pomocą ciągów charakterystyk (o ile istnieją).

Moment (zwyczajny) rzędu $l_1 + l_2 + ...+ l_s$

$$m_{l_1 l_2 ... l_s} = E\left(X_{n_1}^{l_1} X_{n_2}^{l_2} ... X_{n_s}^{l_s}\right),$$

Przypadek szczególny – **ciąg wartości oczekiwanych.**

Moment centralny rzędu $l_1 + l_2 + ...+ l_s$

$$\mu_{l_1 l_2 ... l_s} = E\left(\left(X_{n_1} - EX_{n_1}\right)^{l_1} ... \left(X_{n_s} - EX_{n_s}\right)^{l_s}\right),$$

Przypadek szczególny – **ciąg wariancji, ciąg kowariancji (autokowariancji).**

Przykład
Wyznaczyć $E(X_n)$, $cov(X_n, X_m)$, $D^2 (X_n)$, dla ciągu losowego $X_n = An+B$,

gdzie A, B to zmienne losowe o parametrach: $EA = 0$; $EB = 0$, $D^2A = 2$, $D^2B = 2$,

$\rho = 0{,}5$.

L.Kowalski 02.03.2022