Notatki z Rachunku Prawdopodobieństwa

2024-10-08 – 2025-01-21

Spis treści

Przedmowa

To są notatki z przedmiotu rachunek prawdopodobieństwa prowadzonego na kierunku IAD w 2024/2025 roku przez dr hab. Annę Kuczmaszewską, prof. PL. Treści obejmują 14 wykładów i dodatkowo zawierają twierdzenie z ćwiczeń w prostszej formie niż było na wykładzie.

Niektóre dowody, które istnieją w wykładach Pani Profesor nie zostały zanotowane ponieważ były pomijane na wykładach. W wykładach również zdarzały się niedoprecyzowania lub błędy i jeśli były zauważone, były poprawione, jednak to nie oznacza że treści notatek są bezbłędne (w tym gramatycznie).

Notatki znajdują się w **domenie publicznej** na warunkach licencji CC0 1.0 Universal[1]

[1] https://creativecommons.org/publicdomain/zero/1.0/deed.pl

2024-10-08

1 Trochę kombinatoryki

1.1 Wariacja z powtórzeniami (funkcja)

Równoważne definicje:

k-elementową wariacją z powtórzeniami zbioru Y nazywamy każdą funkcję $f : \{1, 2, \ldots, k\} \to Y$

k-elementową wariacją z powtórzeniami zbioru Y (złożonego z n różnych elementów) nazywamy każde uporządkowanie zbioru złożonego z k elementów wybranych ze zbioru Y, przy czym wybrane elementy mogą się powtarzać.

k-elementową wariacją z powtórzeniami zbioru Y (złożonego z n różnych elementów) nazywamy każdy wynik k-krotnego losowania ze zwracaniem ze zbioru Y uwzględniający kolejność wylosowanych elementów

1.2 Wariacja bez powtórzeń (funkcja różnowartościowa)

Równoważne definicje:

k-elementową wariacją bez powtórzeń zbioru Y nazywamy każdą różnowartościową funkcję $f : \{1, 2, 3, \ldots, k\} \to Y$

k-elementową wariacją bez powtórzeń zbioru Y (złożonego z n-różnych elementów) nazywamy każde uporządkowanie k-elementowego podzbioru zbioru Y

k-elementową wariacją bez powtórzeń zbioru Y (złożonego z n różnych elementów) nazywamy każdy wynik k-krotnego losowania bez zwracania ze zbioru Y uwzględniający kolejność wylosowanych elementów

1.3 Permutacja bez powtórzeń (bijekcja, permutacja)

Równoważne definicje:

Permutacją n-elementowego zbioru nazywamy każdą funkcję f odwzorowującą zbiór $Y = \{y_1, y_2, \ldots, y_n\}$ na zbiór Y

$$f : \{y_1, y_2, \ldots, y_n\} \overset{\text{na}}{\to} Y$$

Permutacją n-elementowego zbioru Y nazywamy każde uporządkowanie zbioru $Y = \{y_1, y_2, \ldots, y_n\}$

1.4 Kombinacje bez powtórzeń (podzbiór)

k**-elementową kombinacją bez powtórzeń zbioru** Y (złożonego z n-różnych elementów) nazywamy każdy k-elementowy podzbiór zbioru Y

Achtung! 1

W przypadku kombinacji, w przeciwieństwie do wariacji i permutacji, nie jest istotna kolejność elementów, ważne są tylko elementy w zbiorze

1.5 Kombinacja z powtórzeniami (krata)

k**-elementową kombinacją z powtórzeniami z** n nazywamy każde rozmieszczenie n nierozróżnialnych kul w k komórkach

1.6 Permutacje z powtórzeniami (podział zbioru)

n**-elementową permutacją z powtórzeniami** nazywamy każde uporządkowanie zbioru n-elementowego, gdzie k różnych elementów powtarza się odpowiednio $r_1, r_2, \ldots, r_k$ razy

1.7 Oznaczenia

$\overline{V}_n^k$ liczba k-elementowych wariacji z powtórzeniami ze zbioru n różnych elementów
V_n^k liczba k-elementowych wariacji bez powtórzeń ze zbioru n różnych elementów
P_n liczba n-elementowych permutacji
C_n^k liczba k-elementowych kombinacji ze zbioru n różnych elementów
$\overline{C}_n^k$ liczba k-elementowych kombinacji z powtórzeniami z n
$\overline{P}_n(r_1, r_2, \ldots, r_k)$ liczba podziałów zbioru n-elementowego na zbiory o liczebnościach $r_1, r_2, \ldots, r_k$

Twierdzenie 1.1

1. $\overline{V}_n^k = n^k$
2. $V_n^k = \dfrac{n!}{(n-k)!}$
3. $P_n = n!$
4. $C_n^k = \dbinom{n}{k}$
5. $\overline{C}_n^k = \binom{n+k-1}{k}$
6. $\overline{P}_n(r_1, r_2, \ldots, r_k) = \frac{n!}{r_1! r_2! \cdots r_k!}$

Twierdzenie 1.2

Każdy ciąg k-elementowy można otrzymać wybierając $k-$elementowy podzbiór i ustawiając jego elementy w ciąg (permutując je mamy zatem) $V_n^k = C_n^k \cdot P_n$

2 Wstęp do rachunku prawdopodobieństwa

2.1 Zdarzenie elementarne

Zdarzenie elementarne jest pojęciem, którego nie definiujemy. Uważamy je za pojęcie pierwotne.

Przestrzeń zdarzeń elementarnych Zbiór wszystkich zdarzeń elementanych nazywamy przestrzenią zdarzeń elementarnych i oznaczamy symbolem Ω.

2.2 Przykłady zdarzeń elementarnych

Przykład 2.1

Rzucamy monetą. Zdarzenia elementarne: wypadnięcie orła, wypadnięcie reszki.
Zbiór zdarzeń elementarnych Ω to $\Omega = \{O, R\}$

Przykład 2.2

Rzucamy kostką do gry: $\Omega = \{1, 2, 3, 4, 5, 6\}$

Przykład 2.3

Mamy odcinek OA o długości l na osi OX. Wybieramy punkt P o współrzędnej x.
Zdarzeniem elementarnym jest «wylosowanie dowolnego punktu na odcinku OA».
Zbiór zdarzeń elementarnych Ω składa się ze wszystkich punktów na odcinku OA

$$\Omega = \{x \in \mathbb{R} : 0 \leq x \leq l\}$$

Zdarzenie elementarne Zdarzenia elementarne możemy traktować jako najprosztsze nierozkładalne, elementarne wyniki doświadczenia losowego, charakteryzujące się tym, że każde powtórzenie tego doświadczenia kończy się jednym i tylko jednym z nich.
Jednak nie jest to definicja.

Zdarzenie pewne Zdarzeniem pewnym nazywamy całą przestrzeń zdarzeń elementarnych Ω

Zdarzenie niemożliwe Zdarzeniem niemożliwym nazywamy podzbiór pusty $\emptyset$ przestrzeni Ω

Zdarzenie przeciwne Zdarzeniem przeciwnym do zdarzenia A nazywamy zdarzenie C składające się z tych zdarzeń elementarnych, które nie należą do zdarzenia A. Zdarzenie przeciwne do zdarzenia A oznaczamy symbolem

$$A' = \{\omega : (\omega \notin A)\}$$

2.3 Sigma-ciało zdarzeń losowych

Rodzinę $\mathcal{F}$ podzbiorów zbioru Ω spełniającą następujące warunki:

1. $\Omega \in \mathcal{F}$

   do rodziny $\mathcal{F}$ należy zdarzenie pewne

2. $\bigwedge\limits_{A \in \mathcal{F}} A' \in \mathcal{F}$

   jeżeli zdarzenie A należy do rodziny $\mathcal{F}$, to zdarzenie przeciwne A' też należy do $\mathcal{F}$

3. $\bigwedge\limits_{A_1, A_2, \ldots, \in \mathcal{F}} \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$

   jeżeli zdarzenia $A_1, A_2, \ldots$ należą do rodziny $\mathcal{F}$, to ich suma również należy do rodziny $\mathcal{F}$

nazywamy σ**-ciałem zdarzeń losowych** (używa się również pojęcia borelowskie σ-ciał zdarzeń lub niekiedy σ-algebra), a elementy tej rodziny nazywamy zdarzeniami losowymi

Achtung! 2

- Jeżeli przestrzeń zdarzeń elementarnych Ω jest co najwyżej przeliczalna, to jako klasę $\mathcal{F}$ (zbiór wszystkich zdarzeń losowych) przyjmuje się klasę wszystkich podzbiorów zbioru Ω
- Jeżeli Ω jest przestrzenią zdarzeń elementarnych położoną w przestrzeni euklidesowej $\mathbb{R}^n$, to jako klasę $\mathcal{F}$ zdarzeń losowych przyjmujemy klasę podzbiorów borelowskich przestrzeni Ω

2.4 Zbiory borelowskie

Klasę zbiorów borelowskich na przestrzeni topologicznej definiujemy jako klasę wszystkich zbiorów które można otrzymać ze zbiorów otwartych (tę samą klasę otrzymamy z przedziałów domkniętych) za pomocą skończonej lub przeliczalnej liczby działań teriomnogościowych (dodawanie, mnożenie, uzupełnianie zbiorów).

2.4.1 Zbiory borelowskie na prostej

W szczególności zbiorami borelowskimi na prostej są wszystkie predziały postaci:

1. $(a, b), \langle a, b \rangle, (a, b\rangle, \langle a, b)$
2. $(-\infty, b), (-\infty, b\rangle, (a, +\infty), \langle a, +\infty)$
3. wszystkie zbiory jednopunktowe, przeliczalne, otwarte, domknięte
4. cała prosta $(-\infty, \infty)$, zbiór pusty

2.5 Definicja prawdopodobieństwa (Kołmogorow, 1933)

Niech

- Ω będzie zbiorem zdarzeń elementarnych
- $\mathcal{F}$ – σ-ciałem zdarzeń losowych

Prawdopodobieństwem P nazywamy funkcję określoną na σ-ciele $\mathcal{F}$ przyjmującą wartości rzeczywiste

$$P : \mathcal{F} \to \mathbb{R}$$

i spełniającą następujące aksjomaty:

1. Dla każdego zdarzenia losowego $A \in \mathcal{F}$ prawdopodobieństwo $P(A)$ jest liczbą z przedziału $\langle 0, 1\rangle$ (to, że jest mniejsze od 1 można wyprowadzić z aksjomatów)

$$0 \le P(A)$$

2) Prawdopodobieństwo zdarzenia pewnego Ω jest równe jeden,

   $P(\Omega) = 1$

3) Prawdopodobieństwo przeliczalnej sumy parami wykluczających się zdarzeń ($A_i \cap A_j = \varnothing, i \neq j$ dla $A_i \in \mathcal{F} i \in \mathbb{N}$) jest równe sumie prawdopodobieństw tych zdarzeń:

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

2.6 Przestrzeń probabilistyczna

Dokonujemy pewnego doświadczenia losowego. Niech:

- Ω będzie zbiorem zdarzeń elementarnych w tym doświadczeniu

- $\mathcal{F} - \sigma$ - ciałem zdarzeń losowych w tym doświadczeniu
- P - funkcja prawdopodobieństwa określoną na $\mathcal{F}$

Przestrzenią probabilistyczną (przestrzenią prawdopodobieństwa) nazywamy trójkę $(\Omega, \mathcal{F}, P)$

2024-10-11:

3 Podstawowe pojęcia probabilistyczne

Twierdzenie 3.1

Jeżeli $(\Omega, \mathcal{F}, P)$ jest przestrzenią probabilistyczną i $A, B, A_1, A_2, \ldots, A_n \in \mathcal{F}$, to

1. $P(\varnothing) = 0$
2. jeżeli $A_1, A_2, \ldots, A_n$ wykluczają się wzajemnie ($A_i \cap A_j = \varnothing$, $i \neq j$), to

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

3. $P(A') = 1 - P(A)$
4. Jeżeli $A \subset B$, to $P(B \setminus A) = P(B) - P(A)$
5. Jeżeli $A \subset B$, to $P(A) \leq P(B)$
6. $P(A) \leq 1$
7. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Dowód 3.1

Teza: $P(\varnothing) = 0$
Na mocy pierwszego aksjomatu $P(\varnothing) \geq 0$
Niech $P(\varnothing) = a \geq 0$. Wtedy mamy

$$a = P(\varnothing) = P(\bigcup_{n=1}^{\infty} \varnothing) = \sum_{n=1}^{\infty} P(\varnothing) = \sum_{n=1}^{\infty} a = \lim_{n\to\infty} \sum_{i=1}^{n} a = \lim_{n\to\infty} n \cdot a$$

Zbiór pusty wyrażamy w postaci przeliczalnej sumy zbiorów pustych (które są zbiorami parami rozłącznymi)
Otrzymujemy równanie $a = \lim_{n\to\infty} n \cdot a$
Jedynym rozwiązaniem równania jest $a = 0$. CND

Dowód 3.2

Teza: jeżeli $A_1, A_2, \dots, A_n$ wykluczają się wzajemnie ($A_i \cap A_j = \varnothing, i \neq j$), to

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

Zdefiniujmy następująco ciąg zdarzeń losowych $\{B_i, i \geq 1\}$

$$B_i = \begin{cases} A_i, & i = 1, 2, \dots, n \\ \varnothing, & i = n+1, n+2, \dots \end{cases}$$

Zdarzenia te są parami wykluczające, a zatem na mocy III aksjomatu i własności 1 mamy:

$$\begin{aligned} P(\bigcup_{i=1}^{n} A_i) &= P(\bigcup_{i=1}^{n} A_i \cup \bigcup_{i=n+1}^{\infty} A_i) \\ &= P(\bigcup_{i=1}^{\infty} B_i) \\ &\overset{\text{III aksjomat}}{=} \sum_{i=1}^{\infty} P(B_i) \\ &= \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{n} P(\varnothing) \\ &= \sum_{i=1}^{n} P(A_i) \end{aligned}$$

CND

Dowód 3.3

Teza: $P(A') = 1 - P(A)$
Zauważmy, że dla dowolnego zdarzenia A mamy

$$A \cup A' = \Omega \quad A \cap A' = \varnothing$$

Na mocy aksjomatu II i własności 2 mamy:

$$1 = P(\Omega) = P(A \cup A') = P(A) + P(A')$$

Stąd $P(A') = 1 - P(A)$. CND

Dowód 3.4

Teza: Jeżeli $A \subset B$, to $P(B \setminus A) = P(B) - P(A)$
Dla $A \subset B$ mamy:

$$A \cup (B \setminus A) = B \quad \text{oraz} \quad A \cap (B \setminus A) = \varnothing$$

Wtedy na mocy wlasności 2:

$$P(A) + P(B \setminus A) = P(B)$$

Stąd

$$P(B \setminus A) = P(B) - P(A)$$

CND

Dowód 3.5

Teza: Jeżeli $A \subset B$, to $P(A) \leq P(B)$
Na mocy aksjomatu I oraz własności 4 dla $A \subset B$ mamy
$0 \leq P(B \setminus A) = P(B) - P(A)$
Stąd $P(A) \leq P(B)$. CND

Dowód 3.6

Teza: $P(A) \leq 1$
Zauważmy, że $A \subset \Omega$
Wtedy na mocy własności 5 i aksjomatu II mamy:
$P(A) \leq P(\Omega) = 1$
Stąd $P(A) \leq 1$. CND

Dowód 3.7

Teza: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Korzystamy z następujących zależności:

$$A \cup B = A \cup (B \setminus A), \quad \text{i} \quad B = (B \setminus A) \cup (A \cap B)$$

Korzystając z własności 2 dostajemy układ dwóch zależności:

$$\begin{cases} P(A \cup B) = P(A) + P(B \setminus A) \\ P(B) = P(B \setminus A) + P(A \cap B) \end{cases}$$

Odejmując stronami dostajemy:

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

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

Twierdzenie 3.2: Prawdopodobieństwo sumy trzech zdarzeń

Niech $A, B, C \in \mathcal{F}$.

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ - P(B - C) + P(A \cap B \cap C) \end{aligned}$$

Dowód jak dla mocy sumy trzech zbiorów na mocy własności 7

Twierdzenie 3.3: Wzór włączeń i wyłączeń

Niech $A_1, A_2, \ldots, A_n \in \mathcal{F}$

$$\begin{aligned} P(A_1 \cup A_2 \cup \ldots \cup A_n) &= \sum_{i=1}^{n} P(A_i) \\ &- \sum_{1 \le i_1 < i_2 \le n} P(A_{i_1} \cap A_{i_2}) \\ &+ \ldots \\ &+ (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n) \end{aligned}$$

Twierdzenie 3.4: o ciągłości miary

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

1. Jeżeli $\{A_n, n \ge 1\}$ jest wstępującą rodziną zdarzeń losowych

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

   Oraz

   $$\bigcup_{n=1}^{\infty} A_n = A$$

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

2) i $\{A_n, n \geq 1\}$ jest zstępującą rodziną zdarzeń losowych

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

Oraz

$$\bigcap_{n=1}^{\infty} A_n = A$$

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

Dowód 3.8: Dowód dla rodziny wstępującej

Niech

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_2 \setminus A_1 \\ &\vdots \\ B_n &= A_n \setminus A_{n-1} \\ &\vdots \end{aligned}$$

Wtedy

1. zdarzenia $B_i, i \geq 1$ wykluczają się
2. $\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i = A_n$

   $P(A_n) \overset{2.}{=} \sum_{i=1}^{n} P(B_i)$

   $\bigcup_{i=1}^{\infty} B_i = A$

$$
\begin{aligned}
P(A) &= P(\bigcup_{i=1}^{\infty} B_i) \\
&\overset{\text{III aksjomat}}{=} \sum_{i=1}^{\infty} P(B_i) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) \\
&= \lim_{n\to\infty} P(A_n)
\end{aligned}
$$

Dowód 3.9: Dowód dla rodziny zstępującej

Rozważamy rodzinę zstępującą $C_n, n \geq 1 \quad C_n = A_n'$. Wtedy

$$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} A_n' = \left(\bigcap_{n=1}^{\infty} A_n\right)' = A'$$

$$P(A') \overset{(i)}{=} \lim_{n\to\infty} P(C_n) = \lim_{n\to\infty} (1 - P(A_n))$$

Stąd mamy

$$1 - P(A) = 1 - \lim_{n\to\infty} P(A_n)$$

i ostatecznie

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

Twierdzenie 3.5: o przedłużaniu miary

Jeżeli P jest skończenie addytywną i nieujemną funkcją na pewnym ciele $\mathcal{A}$ podzbiorów Ω, przy czym $P(\Omega) = 1$, i spełniony jest

- warunek (i) z twierdzenia o ciągłości miary LUB
- warunek (ii) dla $A = \varnothing$

to P przedłuża się jednoznacznie do prawdopodobieństwa na $\sigma(\mathcal{A})$, czyli σ-ciele generowanym przez $\mathcal{A}$

4 Różne miary probabilistyczne

4.1 Przeliczalny zbiór zdarzeń elementarnych

- Ω - przeliczalny zbiór zdarzeń
- $\mathcal{F}$ - zbiór wszystkich podzbiorów zbioru Ω

Możemy podać kompletny, zgodny z intuicją i zdrowym rozsądkiem, opis wszystkich możliwych prawdopodobieństw.

Twierdzenie 4.1

Jeśli

- $\Omega = \{\omega_1, \omega_2, \ldots\}$
- dla każdego $i \in \mathbb{N}$ zbiór $\{\omega_i\} \in \mathcal{F}$
- P jest prawdopodobieństwem

to dla każdego $A \subset \Omega$ mamy

$$P(A) = \sum_{i \in \{i : \omega_i \in A\}} p_i$$

gdzie $p_i = P(\{\omega_i\})$

Przykład 4.1

Niech

- $\Omega = \{\omega_1, \omega_2, \ldots\}$ oraz
- $\mathcal{F}$ – rodzina wszystkich podzbiorów Ω.
- $P : \mathcal{F} - > \mathbb{R}$
- $$p(\{\omega_1\}) = p_1$$
  $$p(\{\omega_2\}) = p_2$$
  $$\vdots$$
  $$p(\{\omega_n\}) = p_n$$
- $$\sum_{i=1}^{n} p_i = 1$$

- $P(\Omega) = P(\bigcup_{i=1}^{\infty}\{\omega_i\}) = \sum_{i=1}^{n} p_i = 1$

Weźmy zdarzenia losowe $A_1, A_2, \ldots$, takie, że $A_i \cap A_j = \varnothing$
$A_i = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_n}\}$
$A_j = \{\omega_{j_1}, \omega_{j_2}, \ldots, \omega_{j_n}\}$
$P(\bigcup_{i=1}^{\infty} A_i) = P(\bigcup_{\substack{k:\omega_k \in A_j \\ j=1,2,\ldots}} \{\omega_k\}) = \sum_{\substack{k:\omega_k \in A_j \\ j=1,2,\ldots}} P(\{\omega_k\})$

Dowód 4.1

Dowolne zdarzenie $A \in \mathcal{F}$ wyraża się wzorem

$$A = \bigcup_{i \in \{i:\omega_i \in A\}} \{\omega_i\}$$

- spełniony jest warunek $\{\omega_i\} \cap \{\omega_j\} = \varnothing$ gdy $i \neq j$
- $A = \bigcup_{i \in \{i:\omega_i \in A\}} \{\omega_i\}$ jest skończony lub przeliczalny

Z III aksjomatu lub drugiej własności

$$\begin{aligned} P(A) &= P(\bigcup_{i \in \{i:\omega_i \in A\}} \{\omega_i\}) \\ &= \sum_{i \in \{i:\omega_i \in A\}} P(\{\omega_i\}) \\ &= \sum_{i \in \{i:\omega_i \in A\}} p_i \end{aligned}$$

4.1.1 Uwagi

1. Prawdopodobieństwo jest jednoznacznie wyznaczone przez prawdopodobieństwo zdarzeń elementarnych
2. $\sum_{i=1}^{\infty} p_i = 1 \implies$ na zbiorze przeliczalnym i nieskończonym prawdopodobieństwo nie może być rozłożone «równo»
3. Tylko w przypadku zbioru Ω skończonego często mamy do czynienia z rozkładem równomiernym.

4.2 Prawdopodobieństwo geometryczne

Niech

- $\Omega \subset \mathbb{R}^n$
- μ - naturalna miara na $\mathbb{R}^n$ (np. miara Lebesque'a [lebega])

Wtedy możemy mówić o prawdopodobieństwie geometrycznym, a zadanie sprowadza się do znalezienia miary (długości, pola, objętości itd.) podzbiorów $\mathbb{R}^n$.

$$\Omega \subset \mathbb{R}^n, \quad \mu \text{ – miara}$$

Każde zdarzenie elementarne (w tym przypadku każdy punkt zbioru $\Omega \subset \mathbb{R}^n$) jest «jednakowo możliwe».

$$A \subset \Omega, \quad A \in \mathcal{F}, \quad \text{A jest zdarzeniem losowym}$$

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}$$

Przykład 4.2

> Pani X i pani Y umówiły się na spotkanie między godz. 16 a 17 w centrum miasta. Komunikacja w godzinach szczytu działa jak działa. Osoba, która przyjdzie pierwsza czeka na drugą 20 minut. Jaka jest szansa, że dojdzie do spotkania?

Rozwiązanie
Niech

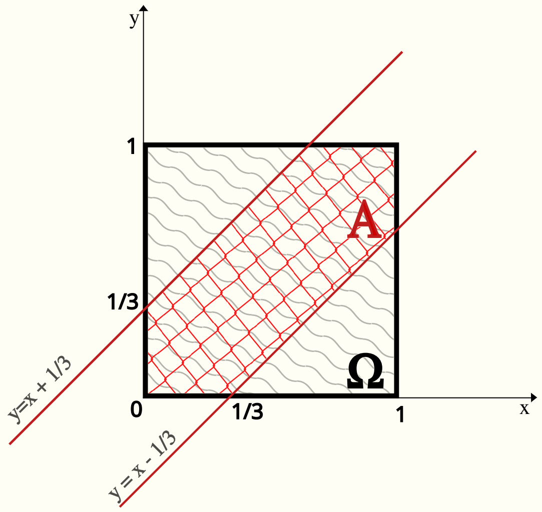
- x – godzina przybycia na umówione miejsce p. X
- y – godzina przybycia na umówione miejsce p. Y

Dla uproszczenia przyjmujemy, że panie spotkają się między godzinami 0.00 i 1.00. Zauważmy, że wtedy

$$x \in \langle 0; 1 \rangle \quad \text{i} \quad y \in \langle 0; 1 \rangle$$

Zbiór zdarzeń elementarnych Ω możemy przedstawić następująco

$$\Omega = \{(x, y) : 0 \leq x, y \leq 1\} \subset \mathbb{R}^2$$



Rysunek 1: prawdopodobieństwo geometryczne

A – zdarzenie polegające na tym, że panie spotkają się (panie spotkają się gdy różnica czasów ich pojawienia się w umówionym miejscu będzie mniejsza niż $\frac{1}{3}$ godziny)

$$A = \{(x, y) : 0 \leq x, y \leq 1, |x - y| \leq \frac{1}{3}\}$$

μ – pole

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{1 - (\frac{2}{3})^2}{1} = \frac{5}{9}$$

2024-10-22

5 Prawdopodobieństwo warunkowe

Bardzo wiele zadań i problemów rachunku prawdopodobieństwa jest formułowanych w terminach prawdopodobieństwa warunkowego.

- firmy ubezpieczeniowe i pytania zadawane przy zawieraniu ubezpieczeń -> skutek: wysokość składki ubezpieczeniowej ustalana na podstawie prawdopodobieństwa przeżycia

Przykład 5.1

W jednej urnie są same białe kule, w drugiej same czarne. Wybieramy losowo urnę i wyciągamy z niej kolejno 2 kule. Niech A(odp. B) oznacza zdarzenie, że pierwsza (odp. druga) kula jest biała. Wyznaczyć $P(A), P(B), P(B|A)$
$P(A) = P(B) = \frac{1}{2}$ – bo wybór urny determinuje wybór koloru kuli
$P(B|A) = 1$ - bo jeżeli zaszło zdarzenie A, to druga będzie biała

5.1 Definicja

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną

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

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

Uwaga: przy ustalnomy $B \in \mathcal{F}$ prawdopodobieństwo warunkowe $P(A|B)$ jest zwykłym prawdopodobieństwem na $(\Omega, \mathcal{F})$ oraz na $(B, \mathcal{F}_B)$, gdzie $\mathcal{F}_B = \{A \cap B : A \in \mathcal{F}\}$

Przykład 5.2

Wybieramy jedną rodzinę spośród rodzin z dwójką dzieci. Obliczyć prawdopodobieństwo, że wybierzemy rodzinę z dwoma chłopcami, jeśli wiemy, że w tej rodzinie

- starsze dziecko jest chłopcem
- jest co najmniej jeden chłopiec

Rozwiązanie.

$$\Omega = \{(c,c), (c,d), (d,c), (d,d)\}$$

pierwszy element pary oznacza płeć starszego dziecka, drugi drugiego.

- $$\begin{aligned} P(\{(c,c)\}|\{(c,c),(c,d)\}) &= \frac{P(\{(c,c)\} \cap \{(c,c),(c,d)\})}{P(\{(c,c),(c,d)\})} \\ &= \frac{P(\{(c,c)\})}{P(\{(c,c),(c,d)\})} \\ &= \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} \end{aligned}$$

- $$\begin{aligned} P(\{(c,c)\}|\{(c,c),(c,d),(d,c)\}) &= \frac{P(\{(c,c)\} \cap \{(c,c),(c,d),(d,c)\})}{P(\{(c,c),(c,d),(d,c)\})} \\ &= \frac{P(\{c,c\})}{P(\{(c,c),(c,d),(d,c)\})} \\ &= \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3} \end{aligned}$$

Przykład 5.3

Udowodnić, że dla dowolnych $A, B \in \mathcal{F}$, takich, że $P(A) > 0$, $P(B) > 0$

$$P(A|B) > P(A) \iff P(B|A) > P(B)$$

Rozwiązanie

$$\begin{aligned} P(A|B) > P(A) &\iff \frac{P(A \cap B)}{P(B)} > P(A) \\ &\iff P(A \cap B) > P(A) \cdot P(B) \\ &\iff P(B|A) > P(B) \end{aligned}$$

Interpretacja: zajście zdarzenia B zwiększa szanse zajścia zdarzenia A wtw, gdy zajście zdarzenia A zwiększa szanse zajścia zdarzenia B

Zastosowanie ostatniego przykładu Prawdopodobieństwo warunkowe, że brydżysta ma asa pik (zdarzenie A), jeśli wiemy, że ma co najmniej 1 asa (zdarzenie B) jest większe od prawdopodobieństwa, że brydżysta ma asa pik

$P(A|B) = 1 \geq P(B)$ ma co najmniej jednego asa gdy ma asa pik, zatem $P(A|B) >= P(A)$

Twierdzenie 5.1

Jeżeli zdarzenia $A_1, A_2, \ldots, A_n$ spełniają warunek

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) > 0$$

to

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot \ldots \cdot P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

Dowód 5.1

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) > 0 \implies$$

wszystkie prawdopodobieństwa warunkowe są dobrze określone

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

5.2 Rozbicie przestrzeni Ω

Rozbiciem przestrzeni Ω nazywamy rodzinę zdarzeń $\{A_i\}_{i \in I}$, które wzajemnie się wykluczają

- $\bigwedge\limits_{\substack{i,j \in I \\ i \neq j}} A_i \cap A_j = \varnothing$

a ich suma daje zdarzenie pewne

- $\bigcup\limits_{i \in I} A_i = \Omega$

5.3 Prawdopodobieństwo całkowite

Jeżeli zdarzenia $\{A_1, A_2, \ldots, A_n\}$ tworzą rozbicie przestrzeni Ω i $P(A_i) > 0$ dla $i \in \{1, 2, \ldots, n\}$, to dla dowolnego zdarzenia B

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

Dowód 5.2

Dla $i = \{1, 2, \ldots, n\}$:

- $P(A_i) > 0$
- $A_i \cap A_j = \varnothing, i \neq j$
- $\bigcup\limits_{i=1}^{n} A_i = \Omega$

$$
\begin{aligned}
B &= B \cap \Omega \\
&= B \cap (A_1 \cup A_2 \cup \ldots \cup A_n) \\
&= \underbrace{(B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n)}_{\text{suma zbiorów rozłącznych}}
\end{aligned}
$$

$$
\begin{aligned}
P(B) &= P((B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n)) \\
&= \sum_{i=1}^{n} P(B \cap A_i) \\
&= \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)
\end{aligned}
$$

5.3.1 Prawdopodobieństwo całkowite dla przeliczalnego rozbicia

Niech $\{A_1, A_2, \ldots, A_n, \ldots\}$ – rozbicie przeliczalne przy czym $\bigwedge_{i \in \mathbb{N}} P(A_i) > 0$, to

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

(dowód analogiczny jak powyżej)

Przykład 5.4

> W loterii fantowej zorganizowane na balu szansa wygranej w jednym losowaniu jest równe p, przegranej q, a z prawdopodobieństwem r wyciągany jest los «graj dalej». Wtedy wrzucamy go z powrotem do urny i losujemy jeszcze raz. Jakie jest prawdopodobieństwo wygranej?

Rozwiązanie. Wprowadzamy następujące oznaczenia:

- A_1 – wyciągnięto los «wygrana», $P(A_1) = p$
- A_2 – wyciągnięto los «przegrana», $P(A_2) = q$
- A_3 – wyciągnięto los «graj dalej», $P(A_3) = r$

gdzie $p + q + r = 1$
Niech B oznacza wygraną. Mamy wtedy:

- $P(B|A_1) = 1$
- $P(B|A_2) = 0$
- $P(B|A_3) = P(B)$

Zakładamy, że loteria liczy tak dużo losów, że pierwsze losowania nie zmieniają prawdopodobieństw.

$$\begin{aligned} P(B) &= P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3) \\ &= 1 \cdot p + 0 \cdot q + P(B) \cdot r \end{aligned}$$

Stąd

$$P(B) \cdot (1 - r) = p \quad P(B) = \frac{p}{1-r}$$

5.4 Wzór Bayesa

Jeżeli zdarzenia $\{A_1, A_2, \dots, A_n\}$ tworzą rozbicie przestrzeni Ω i $P(A_i) > 0$ dla $i \in \{1, \dots, n\}$, to dla dowolnego zdarzenia B takiego, że $P(B) > 0$ mamy

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum\limits_{i=1}^{n} P(B|A_i) \cdot P(A_i)}, \quad i \in \{1, \dots, n\}$$

Dowód 5.3

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i) \cdot P(A_i)}{\sum\limits_{i=1}^{n} P(B|A_i) \cdot P(A_i)}$$

$$P(B|A_i) = \frac{P(B \cap A_i)}{P(A_i)} \implies P(B \cap A_i) = P(B|A_i) \cdot P(A_i)$$

6 Niezależność zdarzeń

6.1 Zdarzenia niezależne

Niezależność - jedno z podstawowych pojęć odróżniających rachunek prawdopodobieństwa od innych działów matematyki badających własności przestrzeni mierzalnych

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią prawdopodobieństwa oraz niech $A,B \in \mathcal{F}$

Naturalnym jest powiedzieć, że zdarzenie B **nie zależy od zdarzenia** A, gdy wiedza o tym, że zaszło zdarzenie A nie ma wpływu na prawdopodobieństwo zajścia zdarzenia B

Załóżmy, że $P(A) > 0$

$$P(B|A) = P(B) \qquad P(A \cap B) = P(A) \cdot P(B)$$

6.2 Zdarzenie niezależne - definicja

Zdarzenia niezależne Zdarzenia A i B są niezależne (stochastycznie niezależne) gdy zachodzi warunek

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

- Zdarzenie niezałeżne i rozłączne ($A \cap B = \varnothing$) to dwa różne pojęcia
- Zdarzenia rozłączne są niezależne wtw, gdy $P(A) = 0 \vee P(B) = 0$
- Relacja niezależności jest symetryczna

Przykład 6.1

Wybieramy jedną rodzinę spośród rodzin mających n dzieci. Niech:

- A – zdarzenie, że w losowo wybranej rodzinie jest co najwyżej jedna dziewczynka
- B – zdarzenie, że w rodzinie są dziewczynki i chłopcy

Czy zdarzenia A i B są niezależne
Rozwiązanie. Tworzymy przestrzeń probabilistyczną:

- Ω – zbiór n elementowych ciągów $(a_1, a_2, \dots, a_n)$ $a_i \in \{\text{chłopiec}, \text{dziewczynka}\}$ uporządkowanych w/g starszeństwa
- $\mathcal{F}$ – zbiór wszystkich podzbiorów zbioru Ω
- P – miara prawdopodobieństwa:

$$P(A) = \frac{\text{liczba elementów w zbiorze } A}{\text{liczba elementów zbioru } \Omega} \overset{\text{ozn.}}{=} \frac{n(A)}{n(\Omega)}$$

Mamy

- $n(\Omega) = 2^n$ (n elementowa wariacja z powtórzeniami z 2)
- $n(A) = n + 1$ (1 dziewczynka na kolejnych miejscach lub sami chłopcy)
- $n(B) = 2^n - 2$ (odejmujemy same dziewczynki i samych chłopców)

Zdarzenia są niezależne $\iff P(A \cap B) = P(A) \cdot P(B)$
Mamy

$$P(A) = \frac{n+1}{2^n} \quad P(B) = \frac{2^n - 2}{2^n} \quad P(A) \cdot P(B) = \frac{(n+1)(2^n-2)}{2^{2n}}$$

$$P(A \cap B) = \frac{n}{2^n} \quad \text{i} \quad P(A \cap B) = \frac{n}{2^n}$$

Warunek:

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

po przekształceniach:

$$n2^n - 2n + 2^n - 2 = n2^n, \quad 2^n = 2 + 2n, \quad 2^{n-1} = 1 + n$$

Mamy

$$2^{n-1} = 1 + n$$

Warunek jest spełniony tylko dla $n = 3$

6.3 Niezależność trzech zdarzeń

Niech $A, B, C \in \mathcal{F}$.

Zdarzenia A, B, C są niezależne wtw, gdy

- $P(A \cap B) = P(A) \cdot P(B)$
- $P(A \cap C) = P(A) \cdot P(C)$
- $P(B \cap C) = P(B) \cdot P(C)$
- $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

6.4 Niezależność n zdarzeń

Niech $A_1, A_2, \ldots, A_n \in \mathcal{F}$.

Zdarzenia $A_1, A_2, \ldots, A_n$ są niezależne wtw, gdy

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_k})$$

dla

$$1 \leq i_1 < i_2 < \ldots < i_k < n \quad \text{i} \quad k = 2, 3, \ldots, n$$

Achtung! 3

Żeby wykazać, że zdarzenia $A_1, A_2, \ldots, A_n \in \mathcal{F}$ są niezależne należy sprawdzić $2^n - n - 1$ równości

Przykład 6.2

Dla $n \geq 3$ z niezależności zdarzeń parami nie wynika ich niezależność.
Niech $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ i $P(\{\omega_i\}) = \frac{1}{4}$
Dla $A_1 = \{\omega_1, \omega_2\}$, $A_2 = \{\omega_1, \omega_3\}$, $A_3 = \{\omega_1, \omega_4\}$ mamy:

$$P(A_i) = \frac{1}{2}, P(A_i \cap A_j) = \frac{1}{4} \quad i \neq j$$

Mamy zatem niezależność parami:

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j) \quad i \neq j$$

Z drugiej strony:

$$P(A_1 \cap A_2 \cap A_3) = P(\{\omega_1\}) = \frac{1}{4} \neq \frac{1}{8} = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

Zatem zdarzenia nie są niezależne.

Twierdzenie 6.1

Jeżeli A i B są zdarzeniami niezależnymi

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

to niezależne są również następujące zdarzenia:

i) A i B'
ii) A' i B
iii) A' i B'

Dowód 6.1

Rozwiązanie. Zakładamy, że A i B są zdarzeniami niezależnymi.
Wtedy:

i) $A = A \cap \Omega = A \cap (B \cup B') = \underbrace{(A \cap B) \cup (A \cap B')}_{\text{suma zdarzeń rozłącznych}}$

$$P(A) = P(A \cap B) + P(A \cap B') = P(A) \cdot P(B) + P(A \cap B')$$

$$P(A \cap B') = P(A) - P(A) \cdot P(B) = P(A)(1 - P(B)) = P(A) \cdot P(B')$$

Zdarzenia A i B' są niezależne

ii) $B = B \cap \Omega = B \cap (A \cup A')$

iii)

$$B' = B' \cap \Omega = B' \cap (A \cup A') = (A \cap B') \cup (A' \cap B')$$

$$P(B') = P(A \cap B') + P(A' \cap B') = P(A) \cdot P(B') + P(A' \cap B')$$

$$P(A' \cap B') = P(B') - P(A) \cdot P(B') = P(B')(1 - P(A)) = P(A')P(B')$$

Twierdzenie 6.2: uogólnienie na n zdarzeń

Przyjmujemy oznaczenia

- $A^0 = A$ i $A^1 = A'$

Twierdzenie
Następujące warunki są równoważne:

- Zdarzenia $A_1, A_2, \ldots, A_n$ są niezależne
- Dla każdego ciągu $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, gdy $\varepsilon_i \in \{0, 1\}$, $i \in \{1, 2, \ldots, n\}$ zdarzenia $B_1 = A_1^{\varepsilon_1}, B_2 = A_2^{\varepsilon_2}, \ldots, B_n = A_n^{\varepsilon_n}$ są niezależne
- Dla każdego ciągu $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, gdy $\varepsilon_i \in \{0, 1\}$, $i \in \{1, 2, \ldots, n\}$ zachodzi równość $P(A_1^{\varepsilon_1} \cap A_2^{\varepsilon_2} \cap \ldots A_n^{\varepsilon_n}) = P(A_1^{\varepsilon_1}) \cdot P(A_2^{\varepsilon_2}) \cdot \ldots \cdot P(A_n^{\varepsilon_n})$

6.5 Niezależność σ-ciał

Niech $(\Omega, \mathcal{F}, P)$ – przestrzeń probabilistyczna

Niezależność σ-ciał σ-ciała $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ ($\mathcal{F}_j \subseteq \mathcal{F}$ dla każdego $j \in \{1, 2, \ldots, n\}$) są niezależne wtw, gdy

$$\bigwedge_{j \in \{1,2,\ldots,n\}} \bigwedge_{A_j \in \mathcal{F}_j} P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2) \cdot \ldots \cdot P(A_n)$$

Twierdzenie 6.3

Zdarzenia $A_1, A_2, \ldots, A_n$ są niezależne wtw, gdy $\sigma(A_1), \sigma(A_2), \ldots, \sigma(A_n)$ są niezależne, gdzie

$$\sigma(A_i) = \{A_i, A_i', \varnothing, \Omega\}$$

jest σ-ciałem generowanym przez $A_i, i \in \{1, 2, \dots, n\}$

6.6 Prawdopodobieństwo sumy niezależnych zdarzeń

Niech $A_1, A_2, \dots, A_n$ – zdarzenia niezależne

$$P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - P\left(\bigcap_{i=1}^{n} A_i'\right) = 1 - \prod_{i=1}^{n} P(A_i') = 1 - \prod_{i=1}^{n}(1 - P(A_i))$$

6.7 Nieskończony ciąg niezależnych zdarzeń

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

2024-10-29

7 Schematy

7.1 Schemat ogólny

Rozważamy sytuację gdy:

- dokonujemy n doświadczeń w sposób niezależny
- z j-tym doświadczeniem związana jest przestrzeń probabilistyczna $(\Omega_j, \mathcal{F}_j, P_j)$(z każdym doświadczeniem związana różna przestrzeń probabilistyczna)
- wynikiem n doświadczeń jest ciąg $(\omega_1, \omega_2, \dots, \omega_n)$, gdzie ω_j - wynik j-tego doświadczenia

Konstruujemy nową przestrzeń $(\Omega, \mathcal{F}, P)$

- $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$
- $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes \ldots \otimes \mathcal{F}_n$ – σ-ciało produktowe, jest to σ-ciało generowane przez cylindry, tzn. zbiory postaci $A_1 \times A_2 \times \ldots \times A_n, A_i \in \mathcal{F}_i$
- $P = P_1 \otimes P_2 \otimes \ldots \otimes P_n$ – miara produktowa, która jest rozszerzeniem na σ-ciało $\mathcal{F}$ miary $P(A_1 \times A_2 \times \ldots \times A_n) = \prod_{i=1}^{n} P_i(A_i)$

7.2 Schemat Bernoulliego

Schemat Bernoulliego jest szczególnym przypadkiem wyżej opisanej sytuacji gdy

- dokonujemy niezależnie n razy pewnego doświadczenia
- każde doświadczenie kończy się jednym spośród dwóch możliwych wyników "sukces", "porażka"
- znamy prawdopodobieństwo p "sukcesu" (a zatem i q "porażki")
- poszczególne doświadczenia nazywamy próbami Bernoulliego

Twierdzenie 7.1: Prawdopodobieństwo zajścia k sukcesów

Prawdopodobieństwo zajścia dokładnie k sukcesów w schemacie Bernoulliego przy n próbach z prawdopodobieństwem sukcesu w pojedynczym doświadczeniu równym p wynosi

$$\binom{n}{k} p^k (1-p)^{n-k}$$

7.3 Granica górna ciągu zdarzeń

Granicą górną ciągu zdarzeń $\{A_n, n \geq 1\}$ nazywamy zdarzenie

$$\limsup A_n := \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$$

interpretacja: $\omega \in \limsup A_n \iff \omega$ należy do nieskończenie wielu zdarzeń z ciągu $\{A_n, n \geq 1\}$

Dowód 7.1

$$\begin{aligned} \omega \in \limsup A_n &\iff \omega \in \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n \\ &\iff \bigwedge_{m \in \mathbb{N}_+} \bigvee_{n \geq m} \omega \in A_n \end{aligned}$$

7.4 Granica dolna ciągu zdarzeń

Granicą dolną ciągu zdarzeń $\{a_n, n \geq 1\}$ nazywamy zdarzenie

$$\liminf A_n := \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n$$

interpretacja: $\omega \in \liminf A_n \iff \omega$ należy do wszystkich zdarzeń z ciągu $\{A_n, n \geq 1\}$, poczynając od pewnego miejsca

Dowód 7.2

$$\begin{aligned}\omega \in \liminf A_n &\iff \omega \in \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n \\ &\iff \bigvee_{m \in \mathbb{N}_+} \bigwedge_{n \geq m} \omega \in A_n\end{aligned}$$

7.4.1 Własności

$$(\liminf A_n)' = \limsup(A_n') \qquad \begin{aligned}(\liminf A_n)' &= \left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n\right)' \\ &= \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n' \\ &= \limsup A_n'\end{aligned}$$

$$(\limsup A_n)' = \liminf(A_n') \qquad \begin{aligned}(\limsup A_n)' &= \left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n\right)' \\ &= \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n' \\ &= \liminf A_n'\end{aligned}$$

7.5 Lemat Borela-Cantellego

- ważne narzędzie do badania własności zachodzących z prawdopodobieństwem 1
- dostarcza informacji o prawdopodobieństwie zdarzeń typu "zaszło nieskończenie wiele zdarzeń $A_n, n \geq 1$"

Twierdzenie 7.2: Lemat Borela-Cantellego

Niech $\{A_n, n \geq 1\}$ będzie ciągiem zdarzeń losowych

i) Jeśli $\sum_{n=1}^{\infty} P(A_n) < \infty$, to $P(\limsup_{n\to\infty} A_n) = 0$

ii) Jeżeli zdarzenia $A_1, A_2, \ldots$ są niezależne i $\sum_{n=1}^{\infty} P(A_n) = \infty$, to $P(\limsup_{n\to\infty} A_n) = 1$

8 Zmienne losowe

Niech $(\Omega, \mathcal{F}, P)$ – przestrzeń probabilistyczna

Zmienna losowa Odwzorowanie $X : \Omega \to \mathbb{R}^n$ nazywamy **zmienną losową** o wartościach w $\mathbb{R}^n$, jeśli

$$\bigwedge_{A \in \mathcal{B}(\mathbb{R}^n)} X^{-1}(A) \in \mathcal{F}$$

gdzie

- $\mathcal{B}(\mathbb{R}^n)$ – wszystkie zbiory borelowskie na $\mathbb{R}^n$
- $\underbrace{X^{-1}(A)}_{\text{przeciwobraz}} = \{\omega \in \Omega : X(\omega) \in A\}$

8.1 Zmienna losowa - uwagi

Inaczej: X jest zmienną losową, jeśli jest odwzorowaniem mierzalnym $(\Omega, \mathcal{F})$ w $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

Powyższa definicja gwarantuje, że prawdopodobieństwo wszystkich zdarzeń związanych ze zmienną losową jest dobrze określone, czyli zdarzeń postaci $X^{-1}(A)$, gdzie $A \in \mathcal{B}(\mathbb{R}^n)$

Wektor losowy Jeśli zmienna losowa odwzorowuje na przestrzeń euklidesową, to taką zmienną nazywamy wektorem losowym.

8.2 Zbiory borelowskie

Zbiór borelowski podzbiór przestrzeni topologicznej, który można uzyskać ze zbiorów otwartych tej przestrzeni (lub równoważnie, ze zbiorów domkniętych) za pomocą przeliczalnych sum, przekrojów bądź dopełnień.

σ-ciało borelowskie Klasa zbiorów uzyskanych za pomocą tych operacji tworzy σ-ciało nazywamy σ-ciałem zbiorów borelowskich lub σ-ciałem borelowskim danej przestrzeni topologicznej

8.2.1 Przykłady zbiorów borelowskich

- zbiór liczb wymiernych na prostej rzeczywistej uzyskany jako przeliczalna suma przeliczalnego iloczynu zbiorów otwartych
- pojedynczy punkt będący dopełnieniem sumy dwóch zbiorów otwartych np. $(-\infty, 1) \cup (1, \infty)$
- rodzina zbiorów borelowskich na prostej jest generowana przez wszystkie przedziały otwarte (równoważnie: domknięte) o końcach wymiernych

- rodzina zbiorów borelowskich na płaszczyźnie jest generowana przez wszystkie prostokąty otwarte o wierzchołkach wymiernych (wystarczą prostokąty o bokach równoległych do osi współtrzędnych)
- nie ma naturalnego przykładu podzbioru prostej rzeczywistej, który nie byłby borelowski (intuicyjnie wszystkie zbiory, które można opisać wzorem są borelowskie)
- istnieją konstrukcje zbiorów korzystające z pewnika wyboru, które dają zbiory nie należące do tej klasy, np. zbiór Vitalego lub zbiór Bernsteina.

Twierdzenie 8.1: Warunek równoważny

Jeżeli odwzorowanie $X : \Omega \to \mathbb{R}^n$ spełnia następujący warunek

$$\bigwedge_{(t_1,t_2,\dots,t_n)\in\mathbb{R}^n} X^{-1}((-\infty,t_1\rangle \times (-\infty,t_2\rangle \times \ldots \times (-\infty,t_n\rangle) \in \mathcal{F}$$

to X jest zmienną losową (wektorem losowym) o wartościach w $\mathbb{R}^n$

Achtung! 4

Wystarczy zauważyć, że rodzina zbiorów

$$\{A \subset \mathbb{R}^n : X^{-1}(A) \in \mathcal{F}\}$$

jest σ-ciałem zawierającym wszystkie zbiory postaci

$$(-\infty,t_1\rangle \times (-\infty,t_2\rangle \times \ldots \times (-\infty,t_n\rangle,$$

a więc wszystkie zbiory borelowskie

Funkcja borelowska Funkcją borelowską nazywamy taką funkcję

$$\varphi : \mathbb{R}^n \to \mathbb{R}^m$$

że przeciwobrazy zbiorów borelowskich w $\mathbb{R}^m$ są zbiorami borelowskimi w $\mathbb{R}^n$

Twierdzenie 8.2

Niech

- $X : \Omega \to \mathbb{R}^n$ będzie zmienną losową
- $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ – funkcją borelowską

Wtedy $(\varphi \circ X)$ jest zmienną losową o wartościach w $\mathbb{R}^m$

Wniosek: Złożenie wektora losowego z funkcją borelowską jest wektorem losowym.

Dowód 8.1

Niech

- $B \in \mathcal{B}(\mathbb{R}^m)$.
- $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ jest funkcją borelowską.

Wtedy $\varphi^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$
Oznaczmy $C = \varphi^{-1}(B)$. Widzimy, że $C \in \mathcal{B}(\mathbb{R}^n)$
Ponieważ X jest zmienną losową, to $X^{-1}(C) \in \mathcal{F}$
Mamy zatem

$$(\varphi \circ X)^{-1}(B) = X^{-1}(\varphi^{-1}(B)) = X^{-1}(C) \in \mathcal{F}$$

8.2.2 Uwagi

Na zmiennych losowych można wykonywać operacje, które są wykonywane na funkcjach mierzalnych i dalej pozostawać w klasie zmiennych losowych.

Każda funkcja

$$Y : \Omega \to \mathbb{R}^k$$

która jest zdefiniowana za pomocą funkcji elementarnych i działań przeliczalnych (w tym $\lim, \limsup, \liminf$) na wektorach losowych $X_1, X_2, \ldots$ jest k-wymiarową zmienną losową.

8.3 σ-ciało generowane przez zmienną losową

σ-ciałem generowanym przez zmienną losową o wartościach w $\mathbb{R}^n$ nazywamy najmniejsze σ-ciało podzbiorów Ω, względem którego X jest mierzalna.

Achtung! 5

Najmniejsze podzbiorów Ω, względem którego X jest mierzalna jest częścią wspólną wszystkich σ-ciał, względem których X jest mierzalna.

Twierdzenie 8.3

$$\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^n)\}$$

9 Rozkład prawdopodobieństwa

Rozkładem prawdopodobieństwa na $\mathbb{R}^n$ nazywamy każdą miarę probabilistyczną $\mu : \mathcal{B}(\mathbb{R}^n) \to \mathbb{R}$ spełniającą warunki

I. $\bigwedge\limits_{B \in \mathcal{B}(\mathbb{R}^n)} \mu(B) \geq 0$

II. $\mu(\mathbb{R}^n) = 1$

III. $\bigwedge\limits_{\substack{B_1, B_2, \ldots \in \mathcal{B}(\mathbb{R}^n) \\ B_i \cap B_j = \varnothing}} \mu(\bigcup\limits_{i=1}^{\infty} B_i) = \sum\limits_{i=1}^{\infty} \mu(B_i)$

Rozkładem prawdopodobieństwa zmiennej losowej X o wartościach w $\mathbb{R}^n$ nazywamy rozkład μ_X określony na $\mathcal{B}(\mathbb{R}^n)$ w następujący sposób

$$\mu_X(B) = P[X^{-1}(B)] = P[\omega : X(\omega) \in B]$$

9.1 Rozkład ciągły

Gęstość rozkładu Jeżeli μ jest rozkładem prawdopodobieństwa na $\mathbb{R}^n$ i dla pewnej funkcji $f : \mathbb{R}^n \to \mathbb{R}$ całkowalnej w sensie Lebesgue'a mamy

$$\mu(A) = \int_A f(x)dx, \quad A \in \mathcal{B}(\mathbb{R}^n)$$

to f nazywamy gęstością rozkładu μ

Rozkład ciągły Rozkład, który ma gęstość nazywamy rozkładem ciągłym.

Twierdzenie 9.1

Niech f będzie gęstością prawdopodobieństwa rozkładu μ na $\mathbb{R}^n$.
Wtedy

1) $\int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x} = 1$
2) $f \geq 0$ prawie wszędzie (poza zbiorem miary Lebesgue'a zero)
3) Gęstość jest wyznaczona jednoznacznie z dokładnością do zbiorów miary Lebesqgue'a 0.

Każda funkcja, spełniająca warunki 1 i 2 jest gęstością prawdopodobieństwa jakiegoś rozkładu.

9.2 Rozkład dyskretny

Rozkład μ na $\mathbb{R}^n$ nazywamy dyskretnym, jeżeli istnieje co najwyżej przeliczalny zbiór $S \subset \mathbb{R}^n$, dla którego $\mu(S) = 1$.

Niech

- $S = \{s_i, i \in I\}$
- $p_i = \mu(\{s_i\})$ dla $i \in I$
- I jest przeliczalnym zbiorem wskaźników, tzn. $I_A(s_i) = \begin{cases} 0, & s_i \notin A \\ 1, & s_i \in A \end{cases}$

Wtedy

$$\mu(A) = \mu(A \cap S) = \sum_{i \in I} p_i I_A(s_i), \quad \forall A \in \mathcal{B}(\mathbb{R}^n)$$

Zbiór $((s_i, p_i), i \in I)$ jednoznacznie wyznacza rozkład μ

9.3 Dystrybuanta

Dystrybuantą rozkładu prawdopodobieństwa μ na $\mathbb{R}^n$ nazywamy funkcję

$$F_\mu : \mathbb{R}^n \to \mathbb{R}$$

określoną zależnością

$$F_\mu(t_1, t_2, \ldots, t_n) = \mu((-\infty, t_1\rangle \times (-\infty, t_2\rangle \times \ldots \times (-\infty, t_n\rangle)$$

Jeżeli μ_X jest rozkładem zmiennej losowej $X = (X_1, X_2, \ldots, X_n)$ o wartościach w $\mathbb{R}^n$, dystrybuantę $F_{\mu X}$ nazywamy dystrybuantą zmiennej losowej X i oznaczamy przez F_X. Wtedy

$$\begin{aligned} F_X(t_1, t_2, \ldots, t_n) &= P(X_1 \le t_1, X_2 \le t_2, \ldots, X_n \le t_n) \\ &= P[\omega \in \Omega : X_1(\omega) \le t_1, X_2(\omega) \le t_2, \ldots, X_n(\omega) \le t_n] \end{aligned}$$

9.3.1 Dystrybuanta na $\mathbb{R}$

Dystrybuanta zmiennej losowej X o wartościach w $\mathbb{R}^1$.

$$F_X : \mathbb{R} \to \mathbb{R} \qquad F_X(t) = P(X \leq t)$$

Własności dystrybuanty na $\mathbb{R}^1$:

1) F_X jest funkcją niemalejącą:

$$\bigwedge_{t_1 < t_2} F_X(t_1) \leq F_X(t_2)$$

2)

$$\lim_{t \to \infty} F_X(t) = 1 \qquad \lim_{t \to -\infty} F_X(t) = 0$$

3) F_X jest funkcją prawostronnie ciągłą:

$$\bigwedge_{t_0 \in \mathbb{R}} \lim_{t \to t_0^+} F_X(t) = F_X(t_0)$$

Twierdzenie 9.2

Jeżeli funkcja $F : \mathbb{R} \to \mathbb{R}$ spełnia warunki (1), (2) i (3), to jest to dystrybuanta pewnego rozkładu

9.3.2 Własności dystrybuanty na $\mathbb{R}^n$

Dystrybuanta F zmiennej losowej X (wektora losowego) na $\mathbb{R}^n$ ma następujące własności:

1) F_X jest funkcją niemalejącą względem każdego argumentu

2)

$$F_X(t_1, t_2, \dots, t_n) \to 0, \text{ jeśli } \inf_{i \in \{1,\dots,n\}} t_i \to -\infty$$

(przynajmniej jeden z argumentów dąży do $-\infty$)

$$F_X(t_1, t_2, \dots, t_n) \to 1, \text{ jeśli } \inf_{i \in \{1,\dots,n\}} t_i \to \infty$$

(każdy z argumentów dąży do ∞)

3) F_X jest funkcją prawostronnie ciągłą ze względu na każdy z argumentów

9.3.3 Rozkład zmiennej losowej $Y = \varphi(X)$

- X - zmienna losowa o ciągłym rozkładzie
- φ – funkcja klasy C^1

Wtedy zmienna losowa $Y = \varphi(X)$ jest zmienną losową o ciągłym rozkładzie.

Twierdzenie 9.3

Jeżeli zmienna losowa X ma rozkład ciągły o gęstości f i $X(\Omega) \subset (a, b)$, funkcja $\varphi : (a, b) \to \mathbb{R}$ jest klasy C^1 oraz $\bigwedge_{x\in(a,b)} \varphi'(x) \neq 0$, to zmienna losowa $Y = \varphi(X)$ ma rozkład ciągły o gęstości

$$f_Y(y) = f(\varphi^{-1}(y)) \cdot \left|(\varphi^{-1}(y))'\right| \cdot I_{\varphi((a,b))}(y)$$

Przykład 9.1

X ma rozkład jednostajny $U([0, 2])$, tzn. ma gęstość $f_X(t) = \begin{cases} 0, & t \notin [0, 2] \\ \frac{1}{2}, & t \in [0, 2] \end{cases}$

$$\varphi : \mathbb{R} \to \mathbb{R} \qquad \varphi(x) = 2x + 5$$

$$y = 2x + 5$$
$$y - 5 = 2x$$
$$x = \frac{y-5}{2}$$
$$x = \varphi^{-1}(y) = \frac{y-5}{2}$$

$$\left(\varphi^{-1}(y)\right)' = \left(\frac{y-5}{2}\right) = \frac{1}{2}$$

$$\begin{aligned} f_Y(y) &= f_X(\varphi^{-2}(y)) \cdot \left|\varphi^{-1}(y)'\right| \cdot I_{[5;9]}(y) \\ &= \begin{cases} \frac{1}{2} \cdot \frac{1}{2}, & y \in [5;9] \\ 0, & y \notin [5;9] \end{cases} \\ &= \begin{cases} \frac{1}{4}, & y \in [5;9] \\ 0, & y \notin [5;9] \end{cases} \end{aligned}$$

2024-11-05

9.4 Wartość oczekiwana

9.4.1 Przypadek dyskretny

Niech

- $(\Omega, \mathcal{F}, P)$ – przestrzeń probabilistyczna
- Zmienna losowa X o wartościach w $\mathbb{R}$ ma rozkład dyskretny $\{(x_i, p_i) : i \in I\}$, gdzie I jest co najwyżej przeliczalnym
- $\sum\limits_{i \in I} |x_i| p_i < \infty$, gdzie $p_i = P[X = x_i]$

Wartością oczekiwaną zmiennej losowej X nazywamy liczbę

$$EX = \sum_{i \in I} x_i p_i = \sum_{i \in I} x_i P[X = x_i]$$

Achtung! 6

Jeżeli I jest zbiorem skończonym, to powyższa definicja pokrywa się z definicją całki z funkcji prostej przyjmującej wartości x_i na zbiorach

$$A_i = \{\omega \in \Omega : X(\omega) = x_i\}$$

względem miary P
Mamy

$$X(\omega) = \sum_{i \in I} x_i I_{A_i}(\omega)$$

oraz

$$\int_\Omega X(\omega) dP(\omega) = \sum_{i \in I} x_i P(A_i) = \sum_{i \in I} x_i P(X = x_i)$$

9.5 Wartość oczekiwana uogólniona

Zmienna losowa X o wartościach w $\mathbb{R}$ ma wartość oczekiwaną EX, jeżeli jest całkowalna, czyli

$$\int_\Omega |X(\omega)| dP(\omega) < \infty$$

Wtedy wartością oczekiwaną zmiennej losowej X nazywamy liczbę

$$EX = \int_\Omega X(\omega)dP(\omega)$$

Achtung! 7

Jeżeli zmienna losowa nie jest całkowalna, to nie ma wartości oczekiwanej (wartość oczekiwana nie istnieje)

9.6 Przypadek ciągły

Jeżeli zmienna losowa X ma rozkład typu ciągłego o funkcji gęstości f, to

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

pod warunkiem, że $\int_{-\infty}^{\infty} |x|f(x)dx < \infty$

9.6.1 Wartość oczekiwana wektora losowego

Wartością oczekiwaną zmiennej losowej (wektora losowego) $X = (X_1, X_2, \dots, X_n)$ o wartościach w $\mathbb{R}^n$ nazywamy wektor

$$EX = (EX_1, EX_2, \dots, EX_n)$$

o ile wszystkie współrzędne wektora X mają wartości oczekiwane.

9.7 Własności wartości oczekiwanej

Niech X, Y i $\{X_n : n \geq 1\}$ będą zmiennymi losowymi o wartościach oczekiwanych odpowiednio równych EX, EY i $\{EX_n, n \geq 1\}$.

Wtedy:

1. Jeżeli $X \geq 0$, to $EX \geq 0$
2. $|EX| \leq E|X|$
3. $E(aX + b) = aEX + b$
4. Jeżeli $X_n \geq 0$, to $E(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} EX_n$

5. Jeżeli $\{X_n, n \geq 1\}$ jest niemalejącym ciągiem nieujemnych zmiennych losowych, to $E\left(\lim\limits_{n\to\infty} X_n\right) = \lim\limits_{n\to\infty} EX_n$ (tw. Lebesgue'a o zbieżności monotonicznej)
6. Jeżeli $\{X_n, n \geq 1\}$ jest ciągiem zmiennych losowych takich, że

$$|X_n| \leq Z$$

dla pewnej całkowalnej zmiennej losowej Z, to

$$E \lim_{n\to\infty} X_n = \lim_{n\to\infty} EX_n$$

(tw. Lebesgue'a o zbieżności zmajoryzowanej)

Przykład 9.2: rozkład dyskretny

$X : \{(x_i, p_i) : i \in I\}$
$P[X = -1] = P[X = 1] = \frac{1}{4}$ oraz $P[X = 0] = \frac{1}{2}$
$\{(-1, \frac{1}{4}), (0, \frac{1}{2}), (1, \frac{1}{4})\}$
$EX = -1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0$
$Y = 2X + 5$

p_i	Y
$\frac{1}{4}$	3
$\frac{1}{2}$	5
$\frac{1}{4}$	7

$$EY = 3 \cdot \frac{1}{4} + 5 \cdot \frac{1}{2} + 7 \cdot \frac{1}{4} = 5$$

$$EY = E(2X + 5) = 2 \cdot EX + 5 = 2 \cdot 0 + 5 = 5$$

$P[Y = 3] = P[2X + 5 = 3] = P[X = -1] = \frac{1}{4}$

Przykład 9.3: rozkład ciągły

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

$$EX = \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{0} x \underbrace{f(x)}_{0} dx + \int_{0}^{2} x \underbrace{f(x)}_{\frac{1}{2}} dx + \int_{2}^{\infty} x \underbrace{f(x)}_{0} dx = \ldots = 1$$

Twierdzenie 9.4

Niech

- $\varphi : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją borelowską
- X – zmienna losowa typu ciągłego o wartościach w $\mathbb{R}^n$

Wtedy

$$E_{\varphi}(X) = \int_{\mathbb{R}^n} \varphi(x) f(x) dx$$

gdzie $f(x)$ jest funkcją gęstości prawdopodobieństwa i $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

Twierdzenie 9.5

Jeżeli zmienna losowa X ma rozkład dyskretny $\{(x_i, p_i), i \in I\}$, to wartość oczekiwana zmiennej losowej $Y = \varphi(X)$ istnieje wtw, gdy zbieżny jest szereg

$$\sum_{i \in I} |\varphi(x_i)| p_i < \infty$$

i wyraża się wzorem

$$E_{\varphi}(X) = \sum_{i \in I} \varphi(x_i) p_i$$

9.8 Wariancja

Wariancja jest miarą rozrzutu wartości zmiennej wokół wartości średniej (wartości oczekiwanej)

Niech X będzie zmienną losową o wartościach rzeczywistych.

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

$D^2(X)$ lub($\mathrm{Var}(X)$)

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

Achtung! 8

Warunkiem koniecznym istnienia wariancji jest istnienie wartości oczekiwanej EX zmiennej losowej X

9.8.1 Własności wariancji

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

1) $D^2(X) = E(X^2) - (EX)^2$
2) $D^2(X) \geq 0$
3) $\bigwedge_{c\in\mathbb{R}} D^2(cX) = c^2 D^2(X)$
4) $\bigwedge_{x\in\mathbb{R}} D^2(X + a) = D^2(X)$
5) $D^2(X) = 0 \iff \bigvee_{c\in\mathbb{R}} P(X = c) = 1$

Dowód 9.1

1) Na mocy założenia $EX^2 < \infty$ i nierówności

$$|t| < t^2 + 1$$

mamy

$$|EX| \leq E|X| < EX^2 + 1 < \infty$$

Zatem wartość oczekiwana istnieje i spełniony jest warunek konieczny istnienia wariancji

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

co dowodzi istnienia wariancji i prawdziwości (1)

2)

$$(X - EX)^2 \geq 0 \implies E(X - EX)^2 \geq 0 \iff D^2(X) \geq 0$$

3)

$$\begin{aligned} D^2(cX) &= E(cX - E(cX))^2 \\ &= E(cx - cEX)^2 \\ &= E(c^2(X - EX)^2) \\ &= c^2 E(X - EX)^2 \end{aligned}$$

4)

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

5)

$$\begin{aligned} D^2(X) = 0 &\iff E(X - EX)^2 = 0 \\ &\iff P(X - EX = 0) = 1 \\ &\iff P(X = EX) = 1 \end{aligned}$$

Wtedy $c = EX$ i mamy

$$P(X = c) = 1$$

9.8.2 Odchylenie standardowe

Odchyleniem standardowym zmiennej X o wariancji $D^2(X)$ nazywamy liczbę $D(X) = \sigma_X$ zdefiniowaną następująco:

$$D(X) = \sigma_X = \sqrt{D^2(X)}$$

9.8.3 Wyznaczanie wariancji i odchylenia standardowego

- dla rozkładu dyskretnego $\{(x_i, p_i), i \in I\}$:

$$D^2(X) = \sum_{i \in I} (x_i - EX)^2 p_i = \sum_{i \in I} x_i^2 p_i - \left(\sum_{i \in I} x_i p_i \right)^2,$$

- dla rozkładu ciągłego (jednowymiarowego) z funkcją gęstości f

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

9.9 Momenty zwykłe i absolutne

Momentem zwykłym rzędu $r > 0$ **zmiennej** X nazywamy liczbę

$$EX^r$$

Momentem absolutnym rzędu $r > 0$ **zmiennej** X nazywamy liczbę

$$E|X|^r$$

pod warunkiem, że wartości oczekiwane istnieją.

9.10 Momenty centralne

- $E(X - EX)^r$ – moment zwykły centralny rzędu $r > 0$
- $E|X - EX|^r$ – moment absolutny centralny rzędu $r > 0$

9.10.1 Interpretacja fizyczna

- Rozkład prawdopodobieństwa - rozkład masy jednostkowej (rozkład dyskretny), zwykła gęstość (gęstość rozkładu ciągłego)
- Wartość oczekiwana EX – położenie środka ciężkości
- Wariancja $D^2(X)$ – moment bezwładności względem środka masy
- Momenty wyższych rzędów – wykorzystywane w statystyce do mierzenia:
  - asymetrii rozkładu (skośności)
  - stopnia koncentracji rozkładu wokół średniej (kurtoza)

9.11 Skośność i koncentracja rozkładu

- współczynnik asymetrii (skośność):

$$\alpha_3 = \frac{E(X - EX)^3}{(D^2(X))^{\frac{3}{2}}} = \frac{E(X - EX)^3}{\sigma_X^3}$$

- współczynnik spłaszczenia (kurtoza):

$$\alpha_4 = \frac{E(X - EX)^4}{(D^2(X))^2} = \frac{E(X - EX)^4}{\sigma_X^4}$$

9.12 Mediana zmiennej losowej

Medianą zmiennej losowej X nazywamy liczbę oznaczaną symbolem Me zdefiniowaną następującym wzorem

$$P[X \leq \mathrm{Me}] \geq \frac{1}{2} \quad \text{i} \quad P[X \geq \mathrm{Me}] \geq \frac{1}{2}$$

Dla rozkładu zmiennej losowej typu ciągłego o gęstości $f(x)$ medianę wyznaczamy ze wzoru

$$\int_{-\infty}^{\mathrm{Me}} f(x)dx = \frac{1}{2} \quad \text{lub} \quad \int_{\mathrm{Me}}^{\infty} f(x)dx = \frac{1}{2}$$

Widać, że jak zsumujemy lewą i prawą całkę, to dostaniemy $\int_{-\infty}^{\infty} f(x)dx = 1$

Achtung! 9

- Mediana istnieje dla każdego zmiennej losowej, ale może nie być wyznaczona w sposób jednoznaczny
- Mediana jest jednym z parametrów służącym do charakteryzowania położenia środka rozkładu zmiennej losowej

9.13 Moda zmiennej losowej

Modą (dominantą) Mo **zmiennej losowej typu skokowego** nazywamy tę wartość zmiennej losowej X, która przyjmuje ona z największym prawdopodobieństwem, pod warunkiem, że nie jest to jej wartość najmniejsza ani największa

Modą (dominantą) Mo **zmiennej losowej typu ciągłego** nazywamy odciętą ekstremum maximum gęstości f zmiennej losowej

Achtung! 10

- Istnieją rozkłady prawdopodobieństwa, które nie mają mody
- Istnieją rozkłady prawdopodobieństwa, które mają więcej niż jedną modę – nazywamy je rozkładami wielomodalnymi

---

2024-11-12

10 Przegląd wybranych dyskretnych rozkładów prawdopodobieństwa

10.1 Rozkład jednopunktowy

Zmienna losowa X ma rozkład **jednopunktowy**, jeżeli istnieje stała $c \in \mathbb{R}$, dla której $P[X = c] = 1$

10.1.1 Wartość oczekiwana

Wtedy $EX = c \cdot 1 = c$, oraz moment rzędu drugiego $E(X^2) = c^2$

Wariancja $D^2(X) = EX^2 - (EX)^2 = 0$, oraz mediana $\mathrm{Me} = c$

Dystrybuanta tej zmiennej losowej dana jest wzorem

$$F(x) = \begin{cases} 0, & x < c \\ 1, & c \leq x \end{cases}$$

10.2 Rozkład dwupunktowy

Zmienna losowa X ma rozkład dwupunktowy, gdy istnieją stałe $c_1, c_2 \in \mathbb{R}$ dla których

$$P[X = c_1] = p_1 \qquad P[X = c_2] = p_2$$

$$0 < p_1 < 1 \qquad 0 < p_2 < 1 \qquad p_1 + p_2 = 1$$

10.2.1 Rozkład zero-jedynkowy

Szczególny przypadek rozkładu dwupunktowego

W przypadku, gdy $c_1 = 0, c_2 = 1$

oraz $p_1 = q, p_2 = p$

rozkład ten przyjmuje postać

$$P[X = 1] = p, \qquad P[X = 0] = q = 1 - p$$

Wartość oczekiwana

$$\begin{aligned} EX &= 1 \cdot p + 0 \cdot q = p \\ EX^2 &= 1^2 \cdot p + 0^2 \cdot q = p \\ D^2(X) &= p - p^2 = pq \end{aligned}$$

Dystrybuanta

$$F(x) = \begin{cases} 0, & x < 0 \\ q, & 0 \leq x < 1 \\ 1, & 1 \leq x \end{cases}$$

10.3 Rozkład dwumianowy

Zmienna losowa X ma rozkład dwumianowy($\mathcal{B}(n,p)$) (zwany także rozkładem Bernoulliego), jeżeli jej funkcja prawdopodobieństwa ma następującą postać:

$$P[X = k] = \binom{n}{k} p^k q^{n-k}, \quad k \in \{0, 1, \ldots, n\}$$

gdzie n jest pewną liczbą naturalną, $0 < p < 1, q = 1 - p$

$$(x_k, p_k)_{k \in \{0,1,\ldots,n\}}, \text{ gdzie } x_k = k, p_k = \binom{n}{k} p^k q^{n-k}$$

10.3.1 Jak rozpoznać rozkład dwumianowy

Dokonujemy n niezależnych doświadczeń losowych:

- zbiór zdarzeń elementarnych w każdym z tych doświadczeń jest dwuelementowy $\Omega = \{\omega_1, \omega_2\}$
- wyróżniamy jedno z tych zdarzeń elementarnych i nazywamy je "sukcesem", a drugie "porażką"
- prawdopodobieństwo "sukcesu" w każdym doświadczeniu jest jednakowe i wynosi p
- zmienna losowa X wyraża liczbę sukcesów w ciągu n niezależnych doświadczeń losowych

10.3.2 wartość oczekiwana

$$
\begin{aligned}
EX &= \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=1}^{n} k \cdot \frac{n!}{k! \cdot (n-k)!} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} q^{n-k} | \text{ podstawiamy } i = k-1 \\
&= np \cdot \sum_{i=0}^{n-1} \frac{(n-1)!}{i! \cdot (n-1-i)!} p^i q^{n-1-i} \\
&= np(p+q)^{n-1} \\
&= np
\end{aligned}
$$

10.3.3 Wariancja

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

$$
\begin{aligned}
E(X^2) &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=1}^{n} \binom{n}{k} k^2 p^k q^{n-k} \\
&= \sum_{k=1}^{n} k \frac{n!}{(k-1)! \cdot (n-k)!} p^k q^{n-k} \\
&\overset{\text{niech } i:=k-1}{=} \sum_{i=0}^{n-1} (i+1) \frac{n!}{i!(n-1-i)!} p^{i+1} q^{n-1-i} \\
&= \sum_{i=0}^{n} i \cdot \frac{n!}{i!(n-1-i)!} p^{i+1} q^{n-1-i} + \sum_{i=0}^{n} \cdot \frac{n!}{i!(n-1-i)!} p^{i+1} q^{n-1-i} \\
&= np \sum_{i=0}^{n-1} i \frac{(n-1)!}{i!(n-1-i)!} p^i q^{n-1-i} + np \sum_{i=0}^{n-1} \frac{(n-1)!}{i!(n-1-i)!} p^i q^{n-1-i} \\
&= np \sum_{i=0}^{n-1} i \binom{n-1}{i} p^i q^{n-1-i} + np \sum_{i=0}^{n-1} \binom{n-1}{i} p^i q^{n-1-i} \\
&= np(n-1)p + np(p+q)^{n-1} \\
&= n^2p^2 - np^2 + np
\end{aligned}
$$

$$E(X^2) = n^2p^2 - np^2 + np$$

$$\begin{aligned} D^2(X) &= E(X^2) - (EX)^2 \\ &= n^2p^2 - np^2 + np - (np)^2 \\ &= np - np^2 \\ &= np(1-p) = npq \end{aligned}$$

10.4 Rozkład Poissona[2]

Zmienna losowa X ma rozkład Poissona $\mathcal{P}(\lambda)$, jeżeli jej funkcja prawdopodobieństwa, jeżeli jej funkcja prawdopodobieństwa jest następująca

$$P[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \{0, 1, \dots\}$$

gdzie $\lambda > 0$

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}, \text{ zatem } \sum_{k=0}^{\infty} P[X = k] = 1$$

$$\begin{aligned} \sum_{k=0}^{\infty} P[X = k] &= \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} e^{\lambda} \\ &= 1 \end{aligned}$$

10.4.1 Wartość oczekiwana

$$\begin{aligned} EX &= \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \text{ niech } i := k - 1 \\ &= e^{-\lambda} \lambda \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \\ &= e^{-\lambda} \lambda e^{\lambda} \\ &= \lambda \end{aligned}$$

[2] /płasona/

10.4.2 Wariancja

$$
\begin{aligned}
E(X^2) &= \sum_{k=0}^{\infty} k^2 \cdot e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{\infty} k(k-1+1) \frac{\lambda^{k-1}}{k!} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{\infty} k(k-1) \frac{\lambda^{k-1}}{k!} + \lambda e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{k!} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{\infty} (k-1) \frac{\lambda^{k-1}}{(k-1)!} + \underbrace{\lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{\lambda} \\
&= e^{-\lambda} \lambda \sum_{k=2}^{\infty} (k-1) \cdot \frac{\lambda^{k-1}}{(k-1)!} + \lambda \\
&= e^{-\lambda} \lambda \sum_{k=2}^{\infty} \frac{\lambda^{k-1}}{(k-2)!} + \lambda \\
&= e^{-\lambda} \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda \\
&= e^{-\lambda} \lambda^2 \sum_{j=0}^{\infty} \frac{\lambda^{j}}{(j)!} + \lambda \\
&= e^{-\lambda} \lambda^2 e^{\lambda} + \lambda \\
&= \lambda^2 + \lambda
\end{aligned}
$$

Zatem

$$
D^2(X) = \underbrace{\lambda^2 + \lambda}_{E(X^2)} - \underbrace{\lambda^2}_{(EX)^2} = \lambda
$$

10.5 Związek rozkładu Bernoulliego i rozkładu Poissona

Jeżeli zmienna losowa X_n ma rozkład dwumianowy $\mathcal{B}(n, p_n)$:

$$
P[X_n = k] = \binom{n}{k} p_n^k q_n^{n-k}, \qquad k \in \{0, 1, \dots, n\}
$$

przy czym $\lim\limits_{n\to\infty} np_n = \lambda < 20$

$$
\lim_{n\to\infty} P[X_n = k] = e^{-\lambda} \frac{\lambda^k}{k!}
$$

Twierdzenie to, dla dostatecznie dużej liczby doświadczeń, pozwala traktować rozkład Bernoulliego jako rozkład w przybliżeniu Poissona z parametrem $\lambda = np$

$$P[X_n = k] \approx e^{-\lambda} \frac{\lambda^k}{k!}$$

10.6 Rozkład geometryczny

Zmienna losowa X ma rozkład geometryczny, jeżeli jej funkcja prawdopodobieństwa jest następująca

$$P[X = k] = (1 - p)^{k-1} p \qquad k \in \mathbb{N}_+$$

gdzie $0 < p < 1$ i $q = 1 - p$

Achtung! 11

- Jest to rozkład czasu oczekiwania na pierwszy sukces w ciągu doświadczeń Bernoulliego
- ma własność "braku pamięci"
- jego odpowiednikiem ciągłym jest rozkład wykładniczy

10.6.1 Wartość oczekiwana

$$\begin{aligned} EX &= \sum_{k=1}^{\infty} k \cdot P[X = k] \\ &= \sum_{k=1}^{\infty} k \cdot (1 - p)^{k-1} p \\ &= \sum_{k=1}^{\infty} k \cdot q^{k-1} p \\ &= p \sum_{k=1}^{\infty} k \cdot q^{k-1} \end{aligned}$$

$$p(1 + q + q^2 + \ldots) = \frac{p}{1 - q} \qquad \text{różniczkujmey po } q$$

$$p + 2pq + 3pq^2 + \ldots = \frac{p}{(1 - q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

Zatem

$$EX = \sum_{k=1}^{\infty} kq^{k-1}p = p + 2pq + 3pq^2 + \ldots = \frac{1}{p}$$

10.6.2 Wariancja

$$\begin{aligned} EX^2 &= \sum_{k=1}^{\infty} k^2 P[X = k] \\ &= \sum_{k=1}^{\infty} k^2 (1-p)^{k-1} p \\ &= \sum_{k=1}^{\infty} k^2 q^{k-1} p \end{aligned}$$

$$p + pq + 3pq^2 + \ldots = \frac{p}{(1-q)^2} \qquad \text{mnożymy obustronnie przez } q$$

$$pq + 2pq^2 + 3pq^3 + \ldots = \frac{pq}{(1-q)^2} \qquad \text{różniczkujemy po } q$$

$$p + 2^2pq + 3^2pq^2 + \ldots = \frac{p(1-q)^2 + 2(1-q)pq}{(1-q)^4} = \frac{p(1-q^2)}{(1-q)^4}$$

Zatem

$$EX^2 = \frac{p(1-q^2)}{p^4} = \frac{1-q^2}{p^3}$$

Na końcu mamy wariancję

$$\begin{aligned} D^2(X) &= E(X^2) - (EX)^2 \\ &= \frac{1-q^2}{p^3} - \frac{1}{p^2} \\ &= \frac{1-p}{p^2} \\ &= \frac{q}{p^2} \end{aligned}$$

Achtung! 12

Czasami rozkład geometryczny jest definiowany następująco:

$$P[X = k] = (1-p)^k p, \qquad k \in \mathbb{N}$$

gdzie $p \in (0, 1)$ i $q = 1 - p$

$$EX = \frac{1}{p} - 1 = \frac{q}{p}$$

10.7 Rozkład ujemnie dwumianowy

Zmienna losowa X ma rozkład ujemnie dwumianowy, jeżeli jej funkcja prawdopodobieństwa jset następująca

$$P[X = k] = \binom{r+k-1}{k}(1-p)^k p^r, \qquad k \in \mathbb{N}$$

gdzie $p \in (0, 1)$ i $r > 0$

Nazwa rozkładu nawiązuje do zależności

$$\begin{aligned}\binom{r+k-1}{k} &= (-1)^k \binom{-r}{k} \\ &= (-1)^k \cdot \frac{r(r-1)\dots(r-k+1)}{k!}\end{aligned}$$

Achtung! 13

- Jeśli r jest liczbą całkowitą, to wyraża ona czas oczekiwania na mr -ty sukces w ciągu prób Bernoulliego - **rozkład Pascala**
- X – liczba porażek poprzedzających r-ty sukces - **rozkład Pascala**
- Dla $r = 1$ otrzymujemy **rozkład geometryczny**

10.7.1 parametry

- Wartość oczekiwana:

$$EX = \frac{r(1-p)}{p} = \frac{rq}{p}$$

- Wariancja

$$D^2(X) = \frac{r(1-p)}{p^2} = \frac{rq}{p^2}$$

10.8 Rozkład hipergeometryczny

Zmienna losowa X ma rozkład hipergeometryczny, jeżeli jej funkcja prawdopodobieństwa jest następująca

$$P[X = k] = \frac{\binom{n}{k}\binom{m-n}{r-k}}{\binom{m}{r}}$$

gdzie $\max(0, n + r - m) \leq k \leq \min(n, r)$

$$\begin{aligned}
&\binom{n}{k} && 0 \leq k \leq n \\
&\binom{m-n}{r-k} \quad 0 \leq r-k \leq m-n && \implies n - m + r \leq k \leq r \\
&\binom{m}{r} && 0 \leq r \leq m
\end{aligned}$$

Zatem $\max(0, n - m + r) \leq k \leq \min(n, r)$

10.8.1 Interpretacja

Z populacji m elementów, gdzie n są wyróżnione wybieramy r elementów, przy czym chcemy żeby k były wyróżnione. Zmienna X wyraża liczbę wyróżnionych elementów w próbie

$$|\Omega| = \binom{m}{r} \qquad |A_k| = \binom{n}{k}\binom{m-n}{r-k}$$

Zatem $P[X = k] = P(A_k) = \frac{|A_k|}{|\Omega|} = \frac{\binom{n}{k}\binom{m-n}{r-k}}{\binom{m}{r}}$

10.8.2 Parametry

$$0 \leq n \leq m \qquad 0 \leq r \leq m$$

- wartość oczekiwana:

$$EX = \frac{nr}{m}$$

- wariancja:

$$D^2(X) = \frac{nr(m-r)(m-n)}{m^2(m-1)}$$

10.9 Tabela rozkładów dyskretnych i ich parametrów

Nazwa	Wzór	EX	$D^2(X)$
Rozkład jednopunktowy	$P[X=c]=1$	c	0
Rozkład dwupunktowy	$P[X=c_1]=p_1$ $P[X=c_2]=p_2$ $p_1+p_2=1$	$p_1c_1+p_2c_2$	$p_1p_2(c_1-c_2)^2$
Rozkład zero-jedynkowy	$P[X=0]=1-p$ $P[X=1]=p$	p	$p-p^2$
Rozkład dwumianowy $\mathcal{B}(n,p)$	$P[X=k]=\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
Rozkład Poissona $\mathrm{Pois}(\lambda)$	$P[X=k]=e^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ
Rozkład geometryczny	$P[X=k]=(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Rozkład ujemnie dwumianowy	$P[X=k]=\binom{r+k-1}{k}(1-p)^kp^r$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$
Rozkład hipergeometryczny	$P[X=k]=\frac{\binom{n}{k}\binom{m-n}{r-k}}{\binom{m}{r}}$	$\frac{nr}{m}$	$\frac{nr(m-r)(m-n)}{m^2(m-1)}$

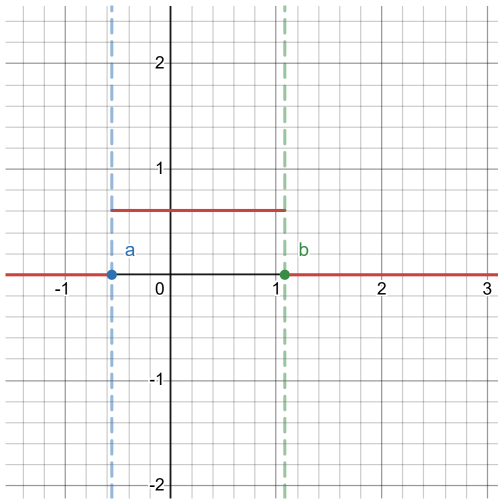
2024-11-19

11 Rozkłady ciągłe

11.1 Rozkład jednostajny

Zmienna X ma rozkład jednostajny(prostokątny/równomierny) w przedziale skończonym $\langle a, b\rangle$ (lub $(a, b\rangle$, $\langle a, b)$, (a, b)), jeżeli ma gęstość $f(x)$ określoną wzorem

$$f(x) = \begin{cases} 0, & x < a \\ \frac{1}{b-a}, & a \le x \le b \\ 0, & b < x \end{cases}$$



Rysunek 2: Wykres funkcji gęstości [3]

dla (a, b) gęstość rozkładu jednostajnego ma postać $f(x) = \begin{cases} 0, & x \le a \\ \frac{1}{b-a}, & a < x < b \\ 0, & x \le b \end{cases}$

11.1.1 Dystrybuanta

Dla $a \le x < b$ mamy

[3] https://www.desmos.com/calculator/1hj8epygtb

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(x)dx \\ &= \int_{-\infty}^{a} f(x)dx + \int_{a}^{x} f(x)dx \\ &= 0 + x \cdot \frac{1}{b-a} \\ &= \frac{x}{b-a}\Big|_a^x \\ &= \frac{x-a}{b-a} \end{aligned}$$

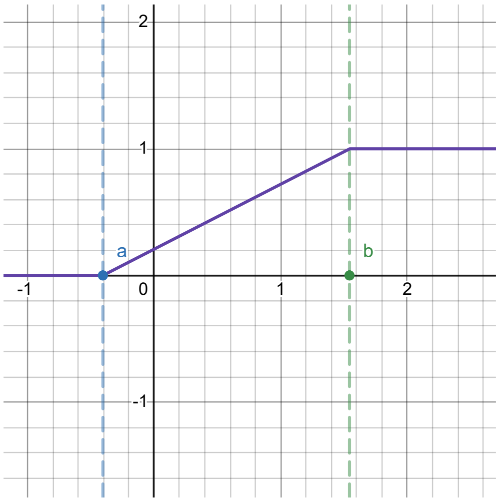
Dla $x < a$:

$$F(x) = \int_{-\infty}^{x} f(x)dx = 0$$

Dla $b < x$:

$$F(x) = \int_{-\infty}^{x} f(x)dx = \ldots = 1$$

Zatem ostatecznie $F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & b \leq x \end{cases}$



Rysunek 3: Wykres dystrybuanty [4]

[4] https://www.desmos.com/calculator/1hj8epygtb

11.1.2 Wartość oczekiwana

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_a^b \frac{x}{b-a}dx \\ &= \frac{x^2}{2(b-a)}\Bigg|_a^b \\ &= \frac{a+b}{2} \end{aligned}$$

11.1.3 Drugi moment zwykły

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx \\ &= \int_a^b \frac{x^2}{b-a}dx \\ &= \frac{x^3}{3(b-a)}\Big|_a^b \\ &= \frac{a^2+ab+b^2}{3} \end{aligned}$$

11.1.4 Wariancja

$$D^2(X) = E(X^2) - (EX)^2 = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

11.1.5 Mediana

$$\mathrm{Me} = \frac{a+b}{2}$$

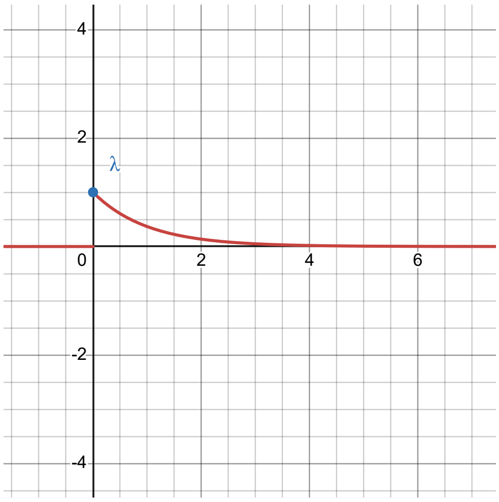
11.1.6 Dominanta

Rozkład jednostajnie ciągły nie ma dominanty.

11.2 Rozkład wykładniczy

Zmienna X ma rozkład wykładniczy z parametrem $\lambda > 0$, jeżeli ma gęstość prawdopodobieństwa $f(x)$ określoną wzorem

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \lambda e^{-\lambda x}, & 0 < x \end{cases}$$

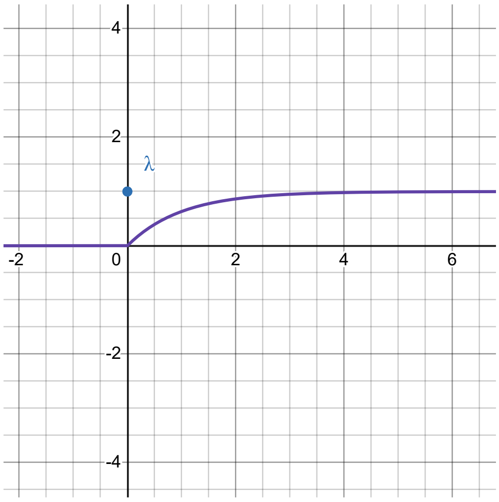


Rysunek 4: Wykres funkcji gęstości [5]

11.2.1 Dystrybuanta

$$F(x) = \int_{-\infty}^{x} f(x)dx = \begin{cases} \int_{-\infty}^{x} 0dx, & x < 0 \\ \int_{0}^{x} \lambda x^{-\lambda x} dx, & 0 \leq x \end{cases} = \begin{cases} 0, & x < 0 \\ -e^{-\lambda x}\Big|_{0}^{x}, & 0 \leq x \end{cases} = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & 0 \leq x \end{cases}$$

[5]https://www.desmos.com/calculator/nyv0n0b128



Rysunek 5: Wykres dystrybuanty [6]

11.2.2 Wartość oczekiwana

$$
\begin{aligned}
EX &= \int_{-\infty}^{\infty} x f(x) dx \\
&= \int_{0}^{\infty} \lambda x e^{-\lambda x} dx \\
&= -xe^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x} \\
&= -\frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty \\
&= \frac{1}{\lambda}
\end{aligned}
$$

11.2.3 Drugi moment zwykły

$$
\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx \\
&= \int_{0}^{\infty} \lambda x^2 e^{-\lambda x} dx \\
&= -x^2 e^{-\lambda x}\Big|_0^\infty + \frac{2}{\lambda}\int_0^\infty \lambda x^{\lambda x} \\
&= \frac{2}{\lambda^2}
\end{aligned}
$$

[6] https://www.desmos.com/calculator/nyv0n0b128

11.2.4 Wariancja

$$D^2(X) = E(X^2) - (EX)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

11.2.5 Zastosowania

Rozkład wykładniczy ma zastosowanie do pewnych zagadnień praktycznych

- czas świecenia żarówki,
- czas niezawodnego działania urządzeń elektrycznych można traktować jako zmienną losową o rozkładzie wykładniczym.

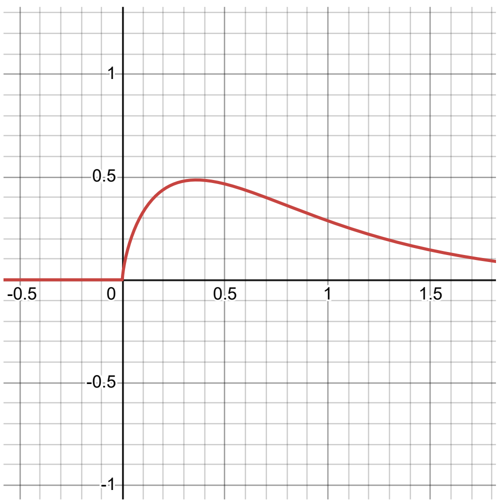
11.3 Rozkład gamma

Zmienna X ma rozkład gamma z parametrami $b, p > 0 \, \Gamma(b, p)$, jeżeli jej gęstość prawdopodobieństwa dana jest wzorem:

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

gdzie

- $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$
- p – parametr kształtu
- b – parametr skali



Rysunek 6: Wykres funkcji gęstości [7]

Achtung! 14

- Funkcja $\Gamma(p)$ jest uogólnieniem znanego pojęcia funkcji silnia: $n \mapsto n!$
- $\Gamma(p+1) = p\Gamma(p)$, bo np. odpowiednio $(n+1)! = n \cdot n!$
- $\int_0^\infty x^{p-1}e^{-bx}dx = \frac{\Gamma(p)}{b^p}$

11.3.1 Dystrybuanta

Nie wyliczymy jej dla wszystkich parametrów

11.3.2 Wartość oczekiwana

$$
\begin{aligned}
EX &= \int_{-\infty}^{\infty} xf(x)dx \\
&= \int_0^\infty \frac{b^p}{\Gamma(p)} x^p e^{-bx} dx \\
&= \frac{\Gamma(p+1)}{b \cdot \Gamma(p)} \int_0^\infty \underbrace{\frac{b^{p+1}}{\Gamma(p+1)} x^p e^{-bx}}_{\text{gęstość rozkładu } \Gamma(b,p+1)} dx \\
&= \frac{p \cdot \Gamma(p)}{b \cdot \Gamma(p)} \cdot 1 \\
&= \frac{p}{b}
\end{aligned}
$$

11.3.3 Drugi moment zwykły

$$
\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx \\
&= \int_0^\infty \frac{b^p}{\Gamma(p)} x^{p+1} e^{-bx} dx \\
&= \frac{\Gamma(p+2)}{b^2 \cdot \Gamma(p)} \int_0^\infty \underbrace{\frac{b^{p+2}}{\Gamma(p+2)} x^{p+1} e^{-bx}}_{\text{gęstość rozkładu } \Gamma(b,p+2)} dx \\
&= \frac{(p+1)p \cdot \Gamma(p)}{b^2 \cdot \Gamma(p)} \cdot 1 \\
&= \frac{p^2+p}{b^2}
\end{aligned}
$$

[7] https://www.desmos.com/calculator/pncczxnwdt

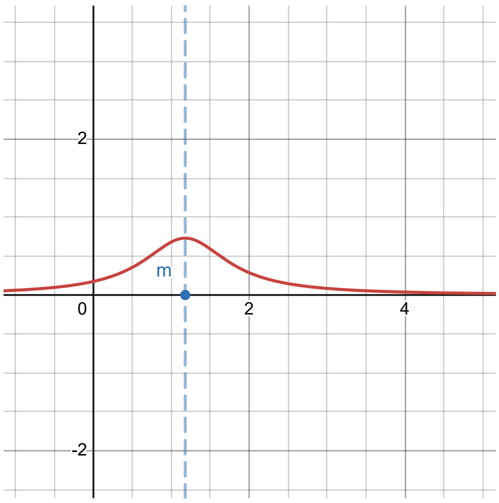
11.3.4 Wariancja

$$D^2(X) = E(X^2) - (EX)^2 = \frac{p^2 + p}{b^2} - \frac{p^2}{b^2} = \frac{p}{b^2}$$

11.4 Rozkład Cauchy'ego

Zmienna X ma rozkład Cauchy'ego z parametrami $a > 0, m \in \mathbb{R}$, jeśli funkcja gęstości dana jest wzorem

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



Rysunek 7: Wykres funkcji gęstości [8]

- Postać standardowa - gdy $a = 1, m = 0$

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

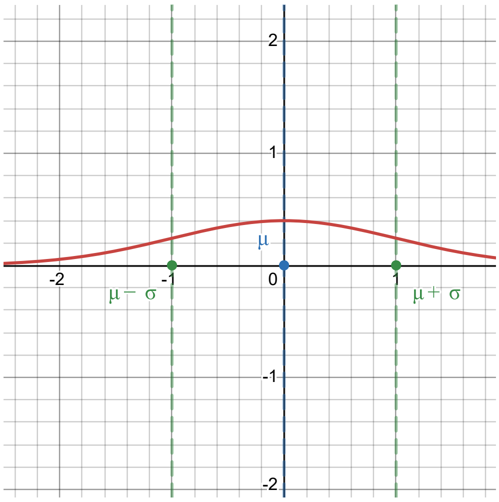
- Momenty: wartość oczekiwana, wariancja i momenty wyższych rzędów nie istnieją

[8] https://www.desmos.com/calculator/k7em2gm1st

11.5 Rozkład normalny

Zmienna losowa X ma rozkład normalny $N(\mu, \sigma)$ z parametrami $\mu \in \mathbb{R}$ i $\sigma \in \mathbb{R}_{+}$, jeżeli ma gęstość prawdopodobieństwa $f(x)$ określoną wzorem

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$



Rysunek 8: Wykres funkcji gęstości [9]

11.5.1 Uwagi

- Badając tę funkcję metodami rachunku różniczkowego mamy:
  1) funkcja $f(x)$ osiąga maksimum w punkcie $x = \mu$
  2) oś x jest asymptotą funkcji $f(x)$ gdy x dąży do $-\infty$ lub $+\infty$
  3) $f(x)$ posiada punkty przegięcia w punktach: $x = \mu - \sigma$ i $x = \mu + \sigma$
  4) krzywa $f(x)$ jest symetryczna względem prostej $x = \mu$
- Badając krzywe $f(x)$ przy różnych wartościach parametru σ stwierdzamy, że im σ jest większe tym krzywa jest bardziej spłaszczona
- Parametr σ jest miarą rozproszenia (odchyleniem standardowym) wartości zmiennej losowej X wokół wartości $x = \mu$ (wartości średniej)

[9] https://www.desmos.com/calculator/ykoq0sgvgl

- Obliczając parametry tego rozkładu skorzystamy z następującego wzoru

$$\int_{-\infty}^{\infty} e^{\frac{1}{2}x^2}dx = \sqrt{2\pi}$$

11.5.2 Wartość oczekiwana

Niech $z = \frac{x-\mu}{\sigma}, dx = \sigma dz, \begin{cases} x = -\infty & \Rightarrow z = -\infty \\ x = \infty & \Rightarrow z = \infty \end{cases}$

$$\begin{aligned} EX &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z + \mu)e^{-\frac{z^2}{2}}dz \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ze^{-\frac{z^2}{2}}dz + \mu \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz \\ &= 0 + \mu \cdot 1 \\ &= \mu \end{aligned}$$

11.5.3 Wariancja

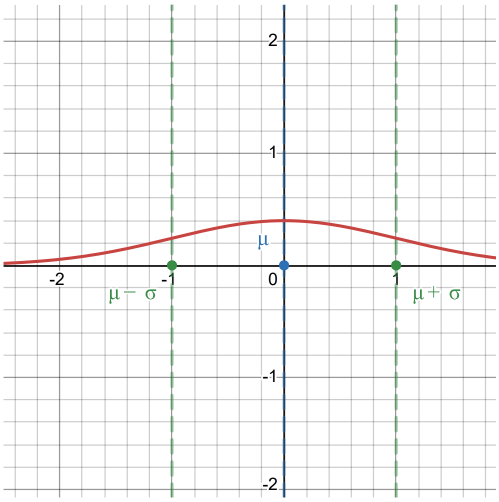
Podstawiamy $z = \frac{x-\mu}{\sigma}$ jak poprzednio

$$\begin{aligned} D^2(X) &= E(X-\mu)^2 \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}}dz \\ &\text{Całkujemy przez części} \left| \begin{matrix} u = z & u' = 1 \\ v' = ze^{-\frac{z^2}{2}} & v = -e^{-\frac{z^2}{2}} \end{matrix} \right| \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \left(\left[-ze^{-\frac{z^2}{2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2}dz \right) \\ &= \frac{\sigma^2}{\sqrt{2\pi}}(0 + \sqrt{2\pi}) = \sigma^2 \end{aligned}$$

11.6 Rozkład normalny $N(0, 1)$

Dla $\mu = 0$ i $\sigma = 1$ to funkcja gęstości jest oznaczana symbolem $\varphi(x)$:

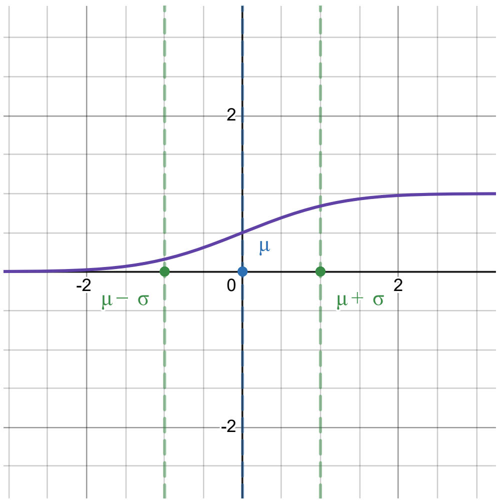
$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$



Rysunek 9: Wykres funkcji gęstości [10]

Dystrybuanta rozkładu $N(0,1)$ ma postać

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} dx$$



Rysunek 10: Wykres dystrybuanty [11]

[10] https://www.desmos.com/calculator/ykoq0sgvgl

Funkcja $\Phi(x)$ jest stabelaryzowana

Ponadto funkcja $\Phi(x)$ spełnia następujący warunek

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

11.6.1 Standaryzacja rozkładu normalnego

X ma rozkład $N(\mu, \sigma) \Rightarrow Y = \frac{X-\mu}{\sigma}$ ma rozkład $N(0, 1)$

Zatem dla dowolnych liczb rzeczywistych a i b takich, że $a < b$ mamy

$$\begin{aligned} P[a < x < b] &= P\left[\frac{a-\mu}{\sigma} < Y < \frac{b-\mu}{\sigma}\right] \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

Możemy zatem od dowolnego rozkładu normalnego $N(\mu, \sigma)$ przejść do rozkładu normalnego $N(0, 1)$

11.6.2 Zastosowania

- rozkład normalny odgrywa ważną rolę w zastosowaniach,
- wiele spotykanych na codzień wielkości(zmiennych losowych) ma rozkład normalny lub bardzo zbliżony do normalnego.

[11] https://www.desmos.com/calculator/ykoq0sgvgl

11.7 Tabela parametrów rozkładów ciągłych

Nazwa	$f(x)$	$F(x)$	EX	$E(X^2)$	$D^2(X)$
Jednostajny $U(a,b)$	$\begin{cases} 0, & x < a \\ \frac{1}{b-a}, & a \le x \le b \\ 0, & b < x \end{cases}$	$\begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x < b \\ 1, & b \le x \end{cases}$	$\frac{a+b}{2}$	$\frac{a^2+ab+b^2}{3}$	$\frac{(b-a)^2}{12}$
Wykładniczy $\mathrm{Exp}(\lambda)$	$\begin{cases} 0, & x \le 0 \\ \lambda e^{-\lambda x}, & 0 < x \end{cases}$	$\begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & 0 \le x \end{cases}$	$\frac{1}{\lambda}$	$\frac{2}{\lambda^2}$	$\frac{1}{\lambda^2}$
Gamma $\Gamma(b,p)$	$\begin{cases} 0, & x \le 0 \\ \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx}, & 0 < x \end{cases}$	–	$\frac{p}{b}$	$\frac{p^2+p}{b^2}$	$\frac{p}{b^2}$
Cauchy'ego	$\dfrac{1}{\pi(a^2 + (x-m)^2)}$	–	–	–	–
Normalny $N(\mu,\sigma)$	$\dfrac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$	–	μ	$\mu^2 + \sigma^2$	σ^2
Normalny $N(0,1)$	$\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}}\, dx$	0	1	1

2024-11-26

12 Rozkłady dwuwymiarowe i wielowymiarowe

12.1 Dwuwymiarowa zmienna losowa (przypomnienie)

Niech

- $(\Omega, \mathcal{F}, P)$ – przestrzeń probabilistyczna
- (X,Y) – dwuwymiarowy wektor losowy (dwuwymiarowa zmienna losowa)

  $(X,Y) : \Omega \to \mathbb{R} \times \mathbb{R}$

  i $\bigwedge\limits_{t_1, t_2 \in \mathbb{R}} \{\omega \in \Omega : X(\omega) < t, Y(\omega) < t_2\} \in \mathcal{F}$

12.2 Moment zwykły mieszany

Momentem zwykłym mieszanym $m_{r,s}$ rzędu $r+s$ zmiennej losowej (X,Y) nazywamy wartość oczekiwaną iloczynu $X^r \cdot Y^s$ (o ile istnieje)

$$m_{r,s} = E(X^r \cdot Y^s) \quad \text{dla } r,s \in \mathbb{N}$$

12.2.1 Rozkład skokowy

Niech (X,Y) – zmienna typu skokowego o rozkładzie prawdopodobieństwa
$p_{ij} = P[X = x_i, Y = y_j] \qquad i,j \in \mathbb{N}_+$

$$\boldsymbol{m_{r,s} = \sum_i \sum_j x_i^r \cdot y_j^s \cdot p_{ij}} \quad \text{o ile} \quad \sum_i \sum_j |x_i|^r \cdot |y_j|^s \cdot p_{ij} < \infty$$

12.2.2 Rozkład ciągły

Niech (X,Y) – zmienna typu ciągłego o rozkładzie prawdopodobieństwa danym za pomocą funkcji gęstości: $f(x,y)$

$$\boldsymbol{m_{r,s} = \iint_{\mathbb{R}^2} x^r \cdot y^s \cdot f(x,y)\, dx\, dy} \quad \text{o ile} \quad m_{r,s} = \iint_{\mathbb{R}^2} |x|^r \cdot |y|^s \cdot f(x,y)\, dx\, dy < \infty$$

W szczególnym przypadku

- dla $r = 0$: $m_{0s} = EY^s$
- dla $s = 0$: $m_{r0} = EX^r$

12.3 Moment centralny mieszany rzędu $r+s$

Momentem centralnym mieszanym $\mu_{r,s}$ rzędu $r+s$ zmiennej losowej (X,Y) nazywamy wartość oczekiwaną iloczynu $(X - EX)^r \cdot (Y - EY)^s$ (o ile istnieje):

$$\mu_{r,s} = E\left[(X - EX)^r \cdot (Y - EY)^s\right]$$

dla $r,s \in \mathbb{N}$

12.3.1 Moment centralny mieszany dla rozkładu skokowego

Niech (X, Y) – zmienna typu skokowego o rozkładzie
$p_{ij} = P[X = x_i, Y = y_j] \quad i,j \in \mathbb{N}_+$

$$\boldsymbol{\mu_{r,s} = \sum_i \sum_j (x_i - EX)^r \cdot (y_j - EY)^s \cdot p_{ij}} \quad \text{o ile} \quad \sum_i \sum_j |x_i - EX|^r \cdot |y_j - EY|^s \cdot p_{ij} < \infty$$

12.4 Moment centralny mieszany dla rozkładu ciągłego

Niech (X, Y) – zmienna losowa typu ciągłego o gęstości prawdopodobieństwa $f(x, y)$

$$\boldsymbol{\mu_{r,s} = \iint_{\mathbb{R}^2} (x - EX)^r \cdot (y - EY)^s \cdot f(x,y)\; dx\; dy} \quad \text{o ile} \quad \iint_{\mathbb{R}^2} |x - EX|^r \cdot |y - EY|^s \cdot f(x,y)\; dx\; dy < \infty$$

W szczególnym przypadku

- dla $r = 2, s = 0$: $\mu_{20} = E(X - EX)^2 = D^2(X)$
- dla $r = 0, s = 2$: $\mu_{02} = E(Y - EY)^2 = D^2(Y)$

12.5 Kowariancja

Kowariancją całkowalnych zmiennych losowych X, Y spełniających warunek

$$E|X \cdot Y| < \infty$$

nazywamy moment centralny mieszany rzędu $1 + 1$

Achtung! 15

Kowariancja zawsze jest dla dwóch zmiennych losowych.

$$\operatorname{cov}(X, Y) = E\left[(X - EX)(Y - EY)\right]$$

12.5.1 Własności

- $\operatorname{cov}(X, Y) = E(XY) - EX \cdot EY$
- $|\operatorname{cov}(X, Y)| \leq \sqrt{D^2(X) \cdot D^2(Y)}$

Dowód 12.1

1) Zauważmy, że

$$\begin{aligned}\operatorname{cov}(X,Y) &= E[(X-EX)(Y-EY)] \\ &= E(XY - XEY - YEX + EX\cdot EY) \\ &= E(XY) - EY\cdot EX - EX\cdot EY + EX\cdot EY \\ &= E(XY) - EX\cdot EY\end{aligned}$$

Czyli $\mu_{11} = m_{11} - m_{10}\cdot m_{01}$

2) Na mocy nierówności Höldera dla

$$\widetilde{X} = X - EX, \quad \widetilde{Y} = Y - EY$$

mamy

$$E(\widetilde{X}\widetilde{Y}) \le \sqrt{E(\widetilde{X}^2)E(\widetilde{Y}^2)}$$

Zatem

$$|E\left((X-EX)\cdot(Y-EY)\right)| \le \sqrt{D^2(X)\cdot D^2(Y)}$$

$$|\operatorname{cov}(X,Y)| \le \sqrt{D^2(X)\cdot D^2(Y)}$$

12.5.2 Nierówność Höldera

Niech

- $p > 1$
- $q > 1$
- $\frac{1}{p} + \frac{1}{q} = 1$

Wtedy

$$E|X\cdot Y| \le \left[E(|X|^p)\right]^{\frac{1}{p}}\left[E(|Y|^q)\right]^{\frac{1}{q}}$$

Czyli

- dla przypadku dyskretnego:

$$\sum_i a_i b_i \leq (\sum_i a_i^p)^{\frac{1}{p}} \cdot (\sum_i b_i^q)^{\frac{1}{q}}$$

- dla przypadku ciągłego:

$$\int f(x)g(x)dx \leq \left(\int (f(x))^p \; dx\right)^{\frac{1}{p}} \left(\int (g(x))^q \; dx\right)^{\frac{1}{q}}$$

12.5.3 Uwagi i interpretacja

- Z zależności 2. wynika, że kowariancja istnieje gdy istnieją odpowiednio wariancje $D^2(X)$ i $D^2(Y)$.
- Kowariancja $\operatorname{cov}(X, Y)$ jest wartością średnią(oczekiwaną) iloczynu odchyleń obu zmiennych od ich wartości oczekiwanych
- może być traktowana jako pewna miara "zgodności" dwu zmiennych losowych:
  - im większa kowariancja tym większa "zgodność" zmiennych (równocześnie przyjmują wartości "duże" i równocześnie przyjmują wartości "małe")

12.6 Korelacja

Niech (X, Y) – dwuwymiarowa zmienna losowa taka, że

- $0 < D^2(X) < \infty$ oraz $0 < D^2(Y) < \infty$

Dwie zmienne losowe X, Y nazywamy **nieskorelowanymi** wtw, gdy

$$\operatorname{cov}(X, Y) = 0$$

Współczynnikiem korelacji ϱ_{XY} zmiennych losowych X, Y nazywamy wyrażenie

$$\begin{aligned} \varrho_{XY} &= \frac{E\left[(X - EX)(Y - EY)\right]}{\sqrt{D^2(X) \cdot D^2(Y)}} \\ &= \frac{\operatorname{cov}(X, Y)}{\sqrt{D^2(X) \cdot D^2(Y)}} \end{aligned}$$

- Jeżeli $\varrho_{X,Y} = 0$, to zmienne losowe są nieskorelowane
- Jeżeli $\varrho_{X,Y} > 0$, to zmienne losowe są dodatnio skorelowane
- Jeżeli $\varrho_{X,Y} < 0$, to zmienne losowe są ujemnie skorelowane

12.6.1 Własności

1. Dla dowolnego wektora losowego (X, Y) o skończonych dodatnich wariancjach $0 < D^2(X) < \infty$, $0 < D^2(Y) < \infty$ zachodzi warunek

$$-1 \leq \varrho_{XY} \leq 1$$

Dowód 12.2

Z własności 2 mamy

$$\begin{aligned} |\operatorname{cov}(X,Y)| \leq \sqrt{D^2(X) \cdot D^2(Y)} &\iff \left|\frac{\operatorname{cov}(X,Y)}{\sqrt{D^2(X) \cdot D^2(Y)}}\right| \leq 1 \\ &\iff |\varrho_{XY}| \leq 1 \end{aligned}$$

2. Niech X, Y będą dowolnymi zmiennymi losowymi, dla których istnieje współczynnik korelacji ϱ_{XY}.

   Warunkiem koniecznym i dostatecznym na to, by zmienne losowe X, Y były z prawdopodobieństwem 1 związane zależnością liniową, tzn. $P[\omega \in \Omega : Y(\omega) = aX(\omega) + b] = 1$ dla $a \neq 0$, jest

$$\varrho^2_{XY} = 1$$

12.7 Regresja

Niech (X, Y) – dwuwymiarowa zmienna losowa, dla której istnieje kowariancja $\operatorname{cov}(X, Y)$

Szukamy funkcji wyrażającej zależność jednej zmiennej od drugiej.

Najprostszą zależnością, najwygodniejszą do dalszych badań jest **zależność liniowa**.

Dlatego często, ze świadomością popełnienia pewnych błędów, przyjmujemy, że zależność pomiędzy badanymi zmiennymi jest liniowa i szukamy prostej, która najlepiej wyraża jedną zmienną jako **funkcję liniową** drugiej zmiennej – szukamy prostej regresji lub inaczej regresji II rodzaju.

12.7.1 Regresja II rodzaju Y względem X

Prostą regresji (regresją II rodzaju) zmiennej Y względem zmiennej X nazywamy prostą

$$y = \alpha x + \beta$$

gdzie α i β są liczbami, dla których spełniona jest zależność

$$E\left((Y - \alpha X - \beta)^2\right) = \min$$

Warunek ten jest spełniony gdy

$$\alpha = \varrho \frac{\sigma_Y}{\sigma_X}, \quad \beta = EY - \varrho \frac{\sigma_Y}{\sigma_X} \cdot EX$$

Prosta regresji zmiennej Y względem zmiennej X ma postać

$$y - EY = \varrho \frac{\sigma_Y}{\sigma_X}(x - EX)$$

12.7.2 Regresja II rodzaju X względem Y

Prostą regresji (regresją II rodzaju) zmiennej X względem zmiennej Y nazywamy prostą

$$x = \gamma y + \delta$$

gdzie γ i δ są liczbami, dla których spełniona jest zależność

$$E\left((X - \gamma Y - \delta)^2\right) = \min$$

Warunek ten jest spełniony gdy

$$\gamma = \varrho \frac{\sigma_X}{\sigma_Y}, \quad \delta = EX - \varrho \frac{\sigma_X}{\sigma_Y} \cdot EY$$

Prosta regresji zmiennej Y względem zmiennej X ma postać

$$x - EX = \varrho \frac{\sigma_X}{\sigma_Y}(y - EY)$$

12.8 Wariancja sumy zmiennych losowych

Jeżeli zmienne losowe $X_1, X_2, \ldots, X_n$ mają wariancje ($\sigma_i^2 = D^2(X_i) < \infty$, $i \in \{1, 2, \ldots, n\}$), to istnieje wariancja sumy $X_1 + X_2 + \ldots + X_n$ i

$$D^2(X_1 + X_2 + \ldots + X_n) = \sum_{i=1}^{n} D^2(X_i) + 2 \sum_{1 \le i < j \le n} \operatorname{cov}(X_i, X_j)$$

Dowód 12.3

$$\begin{aligned} D^2(X_1 + X_2 + \ldots + X_n) &= E[(X_1 + X_2 + \ldots + X_n)^2] - [E(X_1 + X_2 + \ldots + X_n)]^2 \\ &= \sum_{i=1}^{n} E(X_i^2) + 2 \sum_{j=2}^{n} \sum_{i=1}^{j-1} E(X_i X_j) - \sum_{i=1}^{n} (EX_i)^2 - 2 \sum_{j=2}^{n} \sum_{i=1}^{j-1} EX_i \cdot EX_j \\ &= \sum_{i=1}^{n} [E(X_i)^2 - (EX_i)^2] + 2 \sum_{j=1}^{n} \sum_{i=1}^{j-1} [E(X_i X_j) - EX_i \cdot EX_j] \\ &= \sum_{i=1}^{n} D^2(X_i) + 2 \sum_{j=2}^{n} \sum_{i=1}^{j-1} \operatorname{cov}(X_i, X_j) \end{aligned}$$

12.8.1 Wniosek

jeżeli zmienne losowe $X_1, X_2, \ldots, X_n$ mają wariancje ($\sigma_i^2 = D^2(X_i) < \infty$, $i \in \{1, 2, \ldots, n\}$) i są parami nieskorelowane, to

$$D^2(X_1 + X_2 + \ldots + X_n) = \sum_{i=1}^{n} D^2(X_i)$$

12.9 Macierz kowariancji

Dla wektora losowego $X = (X_1, X_2, \ldots, X_n)$ odpowiednikiem wariancji jest macierz kowariancji.

Jeżeli $D^2(X_i) < \infty$ dla każdego $i \in \{1, 2, \ldots, n\}$, to macierz

$$
\begin{aligned}
Q_X &= \Big[\operatorname{cov}(X_i, X_j)\Big]_{i,j\in\{1,2,\ldots,n\}} \\
&= \begin{bmatrix}
D^2(X_1) & \operatorname{cov}(X_1, X_2) & \cdots & \operatorname{cov}(X_1, X_n) \\
\operatorname{cov}(X_2, X_1) & D^2(X_2) & \cdots & \operatorname{cov}(X_2, X_n) \\
\vdots & \vdots & \ddots & \vdots \\
\operatorname{cov}(X_n, X_1) & \operatorname{cov}(X_n, X_2) & \cdots & D^2(X_n)
\end{bmatrix}
\end{aligned}
$$

nazywamy macierzą kowariancji wektora losowego $X = (X_1, X_2, \ldots, X_n)$

12.10 Rozkład wielomianowy

Rozkład wielomianowy jest uogólnieniem rozkładu dwumianowego i opisuje rozkład wyników przy n-krotnym powtórzeniu doświadczenia o k możliwych wynikach

Oznaczamy

- X_i – liczba wyników i-tego typu w serii n doświadczeń.

$$
\begin{aligned}
P(X_1 = n_1, X_2 = n_2, \ldots, X_k = n_k) &= \frac{n!}{n_1! n_2! \ldots n_k!} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k} \\
&= \binom{n}{n_1, n_2, \ldots, n_k} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}
\end{aligned}
$$

Dla takich p i n, że

$$
\bigwedge_{i\in\{1..k\}} p_i \in (0,1) \qquad \sum_{i=1}^{k} p_i = 1 \qquad \sum_{i=1}^{k} n_i = n
$$

Achtung! 16

- Zmienna losowa X_i ma rozkład $\mathcal{B}(n, p_i)$
- $\operatorname{cov}(X_i, X_j) = -p_i p_j, i \neq j$

12.11 Rozkład normalny

Niech

- $X : \Omega \to \mathbb{R}^n$

- $X = (X_1, X_2, \ldots, X_n)$ – zmienna losowa ciągła określona na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$

Mówimy, że zmienna losowa $X = (X_1, X_2, \ldots, X_n)$ ma rozkład normalny z parametrami m, C, gdzie $m \in \mathbb{R}^n$ i C jest macierzą symetryczną, dodatnio określoną stopnia n, jeżeli jej funkcja gęstości prawdopodobieństwa określona jest wzorem

$$f_X(\underbrace{x_1, x_2, \ldots, x_n}_{\mathbf{x}}) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det C}} \cdot \exp\left(-\frac{1}{2}\left[(\mathbf{x} - m) \times C^{-1} \times (\mathbf{x} - m)^T\right]\right)$$

2024-12-03

13 Niezależne zmienne losowe

- $(\Omega, \mathcal{F}, P)$ – przestrzeń probabilistyczna
- $X_i, i \in \{1..n\}$ – zmienne losowe określone na $(\Omega, \mathcal{F}, P)$

Zmienne losowe $X_1, X_2, \ldots, X_n$ są niezależne, jeśli dla dowolnego wyboru zbiorów borelowskich

$$B_1, B_2, \ldots, B_n \in \mathcal{B}(\mathbb{R})$$

zachodzi równość:

$$P[X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n] = P[X_1 \in B_1] \cdot P[X_2 \in B_2] \cdots P[X_n \in B_n]$$

13.1 Uwagi

- rozkład łączny – gęstość dla całego wektora $(X_1, X_2, \ldots, X_n)$, rozkład brzegowy – rozkład poszczególnej zmiennej X_i wektora $(X_1, X_2, \ldots, X_n)$
- Łączny rozkład niezależnych zmiennych losowych jest wyznaczony przez rozkłady brzegowe:

  $$P[X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n] = P[X_1 \in B_1] \cdot P[X_2 \in B_2] \cdots P[X_n \in B_n]$$

- Definicja równoważna:

  Zmienne losowe są niezależne wtw, gdy niezależne są σ-ciała generowane przez te zmienne losowe.

Twierdzenie 13.1

Dla zmiennych losowych $X_1, X_2, \ldots, X_n$ o wartościach w $\mathbb{R}$ następujące warunki są równoważne:

1) zmienne losowe są niezależne
2) $F_{X_1,X_2,\ldots,X_n}(t_1, t_2, \ldots, t_n) = F_{X_1}(t_1) \cdot F_{X_2}(t_2) \cdots F_{X_n}(t_n)$

13.2 Własności

Załóżmy, że zmienne losowe $X_1, X_2, \ldots, X_n$ są niezależne, a funkcje $\varphi_j : \mathbb{R} \to \mathbb{R}$ dla $j \in \{1..n\}$ są mierzalne względem σ-ciała zbiorów borelowskich (są funkcjami borelowskimi).

Wtedy zmienne losowe

$$Y_j = \varphi_j(X_j)$$

są też niezależne.

Dowód 13.1

- φ_j są funkcjami borelowskimi, stąd dla dowolnych zbiorów borelowskich $B_j \in \mathcal{B}(\mathbb{R})$ dla $j \in \{1..n\}$, zbiory

  $$\varphi_j^{-1}(B_j), \quad j \in \{1..n\}$$

  też są zbiorami borelowskimi

Mamy zatem:

$$\begin{aligned}
&P[Y_1 \in B_1, Y_2 \in B_2, \ldots, Y_n \in B_n] \\
&= P[X_1 \in \varphi_1^{-1}(B_1), X_2 \in \varphi_2^{-1}(B_2), \ldots, X_n \in \varphi_n^{-1}(B_n)] \\
&= P[X_1 \in \varphi_1^{-1}(B_1)] \cdot P[X_2 \in \varphi_2^{-1}(B_2)] \cdots P[X_n \in \varphi_n^{-1}(B_n)] \\
&= P[Y_1 \in B_1] \cdot P[Y_2 \in B_2] \cdots P[Y_n \in B_n]
\end{aligned}$$

co kończy dowód

13.3 Zmienne losowe o rozkładzie dyskretnym

Twierdzenie 13.2

Zmienne losowe $X_1, X_2, \ldots, X_n$ o rozkładach dyskretnych są niezależne wtw, gdy dla każdego ciągu

$$x_1, x_2, \ldots, x_n$$

gdzie

- $x_i \in S_{X_i}$ dla $i \in \{1..n\}$
- S_{X_i} – zbiór wartości zmiennej X_i

spełniony jest warunek

$$P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n]$$
$$= P[X_1 = x_1] \cdot P[X_2 = x_2] \cdots P[X_n = x_n]$$

Dowód 13.2

1. $\implies$ Oczywiste z definicji:

$$B_1 = \{x_1\}, B_2 = \{x_2\}, \ldots, B_n = \{x_n\},$$

2. $\impliedby$ Skorzystamy z warunku niezależności dla dystrybuant.

$$
\begin{aligned}
&F_{X_1,X_2,\ldots,X_n}(t_1,t_2,\ldots,t_n) \\
&= \sum_{\substack{y_j^{(i)} \le t_i \\ y_j^{(i)} \in S_{X_i} \\ i \in \{1..n\}}} P[X_1 = y_j^{(1)}, X_2 = y_j^{(2)}, \ldots, X_n = y_j^{(n)}] \\
&= \sum_{\substack{y_j^{(i)} \le t_i \\ y_j^{(i)} \in S_{X_i} \\ i \in \{1..n\}}} P[X_1 = y_j^{(1)}] \cdot P[X_2 = y_j^{(2)}] \cdots P[X_2 = y_j^{(n)}] \\
&= \sum_{i=1}^{n} \sum_{\substack{y_j^{(i)} \le t_i \\ y_j^{(i)} \in S_{X_i}}} P[X_1 = y_j^{(1)}] \cdot P[X_2 = y_j^{(2)}] \cdots P[X_n = y_j^{(n)}] \\
&= \left[\sum_{\substack{y_j^{(1)} \le t_1 \\ y_j^{(1)} \in S_{X_1}}} P[X_1 = y_j^{(1)}] \right] \cdot \left[\sum_{\substack{y_j^{(2)} \le t_2 \\ y_j^{(2)} \in S_{X_2}}} P[X_2 = y_j^{(2)}] \right] \cdots \left[\sum_{\substack{y_j^{(n)} \le t_n \\ y_j^{(n)} \in S_{X_n}}} P[X_n = y_j^{(n)}] \right] \\
&= F_{X_1}(t_1) \cdot F_{X_2}(t_2) \cdots F_{X_n}(t_n)
\end{aligned}
$$

13.4 Zmienne losowe o rozkładzie ciągłym

Twierdzenie 13.3

Zmienne losowe $X_1, X_2, \ldots, X_n$ o rozkładach ciągłych danych za pomocą funkcji gęstości

$$f_1, f_2, \ldots, f_n$$

i rozkładzie łącznym o funkcji gęstości $f_{X_1,X_2,\ldots,X_n}$ są niezależne wtw, gdy

$$\boldsymbol{f_{X_1,X_2,\ldots,X_n}(t_1,t_2,\ldots,t_n) = f_1(t_1) \cdot f_2(t_2) \cdots f_n(t_n)}$$

Dowód 13.3

1. $\implies$ Zakładamy, że zmienne są niezależne. Wtedy:

$$\underbrace{F_{X_1,X_2,\ldots,X_n}(t_1,t_2,\ldots,t_n)}_{L} = \underbrace{F_{X_1}(t_1) \cdot F_{X_2}(t_2) \cdots F_{X_n}(t_n)}_{P}$$

$$\begin{aligned} P &= \int_{-\infty}^{t_1} f_1(x_1)\ dx_1 \cdot \int_{-\infty}^{t_2} f_2(x_2)\ dx_2 \cdots \int_{-\infty}^{t_n} f_n(x_n)\ dx_n \\ &= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \cdots \int_{-\infty}^{t_n} f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)\ dx_1\ dx_2 \cdots\ dx_n \end{aligned}$$

$$L = \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \cdots \int_{-\infty}^{t_n} f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)\ dx_1\ dx_2 \cdots\ dx_n$$

Ostatecznie

$$f_1(x_1) \cdot f_2(x_2) \cdots f_n(x_n) = f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)$$

2. $\Longleftarrow$ Zakładamy, że

$$f_1(x_1) \cdot f_2(x_2) \cdots f_n(x_n) = f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)$$

Wtedy

$$\begin{aligned} &F_{X_1,X_2,\ldots,X_n}(t_1, t_2, \ldots, t_n) \\ &= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \cdots \int_{-\infty}^{t_n} f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)\ dx_1\ dx_2 \cdots\ dx_n \\ &= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \cdots \int_{-\infty}^{t_n} f_1(x_1) \cdot f_2(x_2) \cdots f_n(x_n)\ dx_1\ dx_2\ \ldots\ dx_n \\ &= \int_{-\infty}^{t_1} f_1(x_1)\ dx_1 \cdot \int_{-\infty}^{t_2} f_2(x_2)\ dx_2 \cdots \int_{-\infty}^{t_n} f_n(x_n)\ dx_n \\ &= F_{X_1}(t_1) \cdot F_{X_2}(t_2) \cdots F_{X_n}(t_n) \end{aligned}$$

Co dowodzi niezależności zmiennych

13.4.1 Własności

Twierdzenie 13.4

Załóżmy, że

- zmienne losowe $\overbrace{X_1, X_2, \ldots, X_n}^{n}, \overbrace{X_{n+1}, \ldots, X_{n+m}}^{m}$ są niezależne
- funkcje $\varphi : \mathbb{R}^n \to \mathbb{R}$ oraz $\psi : \mathbb{R}^m \to \mathbb{R}$

Wtedy zmienne losowe

$$Y_1 = \varphi(X_1, X_2, \dots, X_n) \quad \text{oraz} \quad Y_2 = \psi(X_{n+1}, X_{n+2}, \dots, X_{n+m})$$

są niezależne

Twierdzenie 13.5

Jeżeli $X_1, X_2, \dots, X_n$ są niezależnymi zmiennymi losowymi, które mają wartości oczekiwane $EX_1, EX_2, \dots, EX_n$, to istnieje wartość oczekiwana ich iloczynu $X_1 \cdot X_2 \cdots X_n$ i zachodzi warunek

$$E(X_1 \cdot X_2 \cdots X_n) = EX_1 \cdot EX_2 \cdots EX_n$$

Twierdzenie 13.6

Jeżeli $X_1, X_2, \dots, X_n$ są niezależnymi zmiennymi losowymi posiadającymi wariancje $D^2(X_i)$ dla $i \in \{1..n\}$, to istnieje wariancja ich sumy i mamy

$$D^2(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} D^2(X_i)$$

Twierdzenie 13.7

Jeśli X, Y są niezależne, to X, Y są nieskorelowane.
Ale w inną stronę warunek może nie zachodzić

13.5 Rozkład sumy niezależnych zmiennych losowych

Jeżeli zmienne losowe X, Y są niezależne, to rozkład ich sumy $X + Y$ nazywamy splotem rozkładów zmiennej losowej X i zmiennej Y

$$\mathcal{L}(X + Y) = \mathcal{L}(X) * \mathcal{L}(Y)$$

13.5.1 Rozkład sumy niezależnych zmiennych losowych o rozkładach ciągłych

Niech

- X, Y – niezależne zmienne losowe typu ciągłego
- f – funkcja gęstości zmiennej X

- g – funkcja gęstości zmiennej Y

Wyznaczamy dystrybuantę zmiennej losowej $Z = X + Y$:

$$\begin{aligned} F_Z(t) &= P[Z \leq t] \\ &= P[X + Y \leq t] \\ &= \iint_{x+y \leq t} f(x, y) \; dxdy \\ &= \iint_{x+y \leq t} f(x) g(y) \; dxdy \end{aligned}$$

stosujemy podstawienie $x + y = z$ oraz $y = u$.

Stąd

- $x = z - u$ oraz $y = u$
- $\dfrac{\partial x}{\partial z} = 1$ oraz $\frac{\partial x}{\partial u} = -1$
- $\dfrac{\partial y}{\partial z} = 0$ oraz $\frac{\partial y}{\partial u} = 1$
- $\begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$
- $z \leq t, u \in \mathbb{R}$

$$\begin{aligned} F_Z(t) &= \iint_{x+y \leq t} f(x) g(y) \; dxdy \\ &= \int_{-\infty}^{t} \left[\int_{\mathbb{R}} f(z - u) g(u) \; du \right] \; dz \end{aligned}$$

Z drugiej strony:

$$F_Z(t) = \int_{-\infty}^{t} h(z) \; dz, \quad h \text{ – funkcja gęstości zmiennej } Z$$

Stąd

$$h(z) = (f * g)(z) = \int_{\mathbb{R}} f(z - u) g(u) du$$

14 Funkcja charakterystyczna

Niech

- X – zmienna losowa określona na przestrzeni $(\Omega, \mathcal{F}, P)$

Funkcja

$$\varphi : \mathbb{R} \to \mathbb{C}$$

jest funkcją charakterystyczną zmiennej losowej X, jeśli dla $t \in \mathbb{R}$

$$\begin{aligned}\varphi(t) &= Ee^{itX(\omega)} \\ &\stackrel{def}{=} E\cos[tX(\omega)] + iE\sin[tX(\omega)] \\ &= \int_\Omega e^{itX(\omega)} P(d\omega)\end{aligned}$$

14.1 Rozkład dyskretny

Niech

- X – zmienna losowa o rozkładzie dyskretnym
- $\{x_k, k \in \{1..\}\}$ – zbiór wartości zmiennej X
- $p_k = P[X = x_k]$

Funkcja charakterystyczna ma postać

$$\varphi(t) = \sum_{k=1}^{\infty} e^{itx_k} p_k$$

Przykład 14.1

Funkcja charakterystyczna rozkładu «zero-jedynkowego».

$$P[X = 0] = 1 - p = q \qquad P[X = 1] = p$$

$$\varphi(t) = Ee^{itX} = pe^{it} + qe^{it\cdot 0} = pe^{it} + q$$

14.2 Rozkład ciągły

Niech

- X – zmienna losowa o rozkładzie ciągłym
- f – funkcja gęstości zmiennej X

Funkcja charakterystyczna jest równa

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx$$

Przykład 14.2

Funkcja charakterystyczna rozkładu ciągłego o funkcji gęstości $f(x) = \frac{1}{2}e^{-|x|}$

$$\begin{aligned}
\varphi(t) &= Ee^{itX} \\
&= \int_{-\infty}^{+\infty} e^{itx}\frac{1}{2}e^{-|x|}\, dx \\
&= \frac{1}{2}\int_{-\infty}^{+\infty} \cos(tx)e^{-|x|}\, dx + \frac{1}{2}i\int_{-\infty}^{+\infty} \sin(tx)e^{-|x|}\, dx \\
&= \frac{1}{2}\int_{-\infty}^{+\infty} \cos(tx)e^{-|x|}\, dx \\
&= 2 \cdot \frac{1}{2}\int_{0}^{+\infty} \cos(tx)e^{-x}\, dx
\end{aligned}$$

$$\begin{aligned}
\frac{1}{2}\int_0^{+\infty} \cos(tx)e^{-x}\, dx &= \begin{vmatrix} u = \cos(tx) & u' = -t\sin(tx) \\ v' = e^{-x} & v = -e^{-x} \end{vmatrix} \\
&= -e^{-x}\cos tx\Big|_0^{\infty} - t\int_0^{+\infty} e^{-x}\sin(tx)\, dx \\
&= 1 - t\int_0^{\infty} e^{-x}\sin(tx)\, dx \\
&= \begin{vmatrix} u = \sin(tx) & u' = t\cos(tx) \\ v' = e^{-x} & v = -e^{-x} \end{vmatrix} \\
&= 1 - t\left(-e^{-x}\sin(tx)\Big|_0^{\infty} + t\int_0^{+\infty} e^{-x}\cos(tx)\, dx\right) \\
&= 1 - t(0 + t\int_0^{+\infty} e^{-x}\cos(tx)\, dx) \\
&= 1 - t^2\int_0^{+\infty} e^{-x}\cos(tx)\, dx
\end{aligned}$$

$$(1+t^2)\int_0^{+\infty} \cos(tx)e^{-x}\, dx = 1$$

$$\int_0^{+\infty} \cos(tx)e^{-x}\, dx = \frac{1}{1+t^2}$$

Stąd

$$\varphi(t) = \frac{1}{1+t^2}$$

14.3 Własności funkcji charakterystycznej

Funkcja charakterystyczna zmiennej losowej X ma następujące własności

1) $|\varphi_X(t)| \leq \varphi_X(0) = 1$
2) $\varphi_X(-t) = \overline{\varphi_X(t)} = \varphi_{-X}(t)$
3) $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$
4) Jeżeli X, Y są niezależne, to

$$\varphi_{aX+bY} = \varphi_X(at) \cdot \varphi_Y(bt) \qquad \forall a, b \in \mathbb{R}$$

Dowód 14.1

$\varphi_X(0) = Ee^{i0 \cdot X} = 1$
Stąd
$|\varphi_X(t)| = |Ee^{itX}| \leq E|e^{itX}| = E\sqrt{\cos^2(tX) + \sin^2(tX)} = 1$

Dowód 14.2

$\varphi_X(-t) = E\cos(-tX) + iE\sin(-tX) = E\cos(tX) - iE\sin(tX) = \overline{E\cos(tX) + iE\sin(tX)} = \overline{\varphi_X(t)}$
Ponadto
$\varphi_X(-t) = Ee^{-itX} = Ee^{it(-X)} = \varphi_{-X}(t)$

Dowód 14.3

$\varphi_{aX+b}(t) = Ee^{it(aX+b)} = E(e^{it(aX)} \cdot e^{itb}) = e^{itb}Ee^{i(ta)X} = e^{itb}\varphi_X(at)$

Dowód 14.4

X, Y są niezależne $\implies$ $e^{it(aX)}, e^{it(bY)}$ są niezależne

$$\begin{aligned}
\varphi_{aX+bY} &= Ee^{it(aX+bY)} \\
&= E(e^{it(aX)} \cdot e^{it(bY)}) \\
&= Ee^{i(ta)X} \cdot Ee^{i(tb)Y} \\
&= \varphi_X(at) \cdot \varphi_Y(bt)
\end{aligned}$$

14.4 Własności funkcji charakterystycznej

- Każda funkcja charakterystyczna jest funkcją jednostajnie ciągłą na $\mathbb{R}$
- Funkcja charakterystyczna zmiennej X jest rzeczywista wtw, gdy rozkład zmiennej losowej jest symetryczny

  Uwaga: Rozkład jest symetryczny, gdy:

  - $P[X = a] = P[X = -a], a \in \mathbb{R}$, X ma rozkład dyskretny
  - $f(-x) = f(x)$, X ma rozkład ciągły
- Jeżeli $E|X^n| < \infty$, $n \in \mathbb{N}$, to n-ta pochodna funkcji charakterystycznej $\varphi_X^{(n)}$ istnieje i jest jednostajnie ciągła, a ponadto

$$\varphi_X^{(n)}(0) = i^n E(X^n)$$

- Zmienne losowe o różnych rozkładach mają różne funkcje charakterystyczne
- Funkcja $\varphi : \mathbb{R} \to \mathbb{C}$ jest funkcją charakterystyczną pewnego rozkładu prawdopodobieństwa na $\mathbb{R}$ wtw, gdy jest:
  - ciągła
  - dodatnio określona(tzn. $\bigwedge\limits_{x \in \mathbb{R}} \varphi(x) \neq 0$)
  - $\varphi(0) = 1$

2024-12-09 (ćwiczenia)

15 Rozkład ciągły

15.1 Związek gęstości zmiennej X i zmiennej $g(X)$

Jeżeli

- X jest zmienną losową ciągłą o gęstości f skoncentrowanej na przedziale (a, b)
- $y = g(x)$ jest funkcją ściśle monotoniczną klasy $C^{(1)}$ o pochodnej $g'(x) \neq 0$ w tym przedziale
- $x = h(y)$ jest funkcją odwrotną do $y = g(x)$,

to gęstość k zmiennej $Y = g(X)$ jest postaci

$$k(y) = \begin{cases} f[h(y)] \cdot h'(y), & c < y < d \\ 0, & \text{poza} \end{cases}$$

gdzie

- $c = \min(c_1, d_1)$ oraz $d = \max(c_1, d_1)$
- $c_1 = \lim\limits_{x \to a^+} g(x)$ oraz $d_1 = \lim\limits_{x \to b^-} g(x)$

Przykład 15.1

> Zmienna losowa X ma gęstość prawdopodobieństwa $f(x) = \alpha(1-x)I_{(0,1)}$. Wyznaczyć parametr α. Wyznaczyć funkcję gęstości i dystrybuantę zmiennej losowej $Y = e^X$

Gęstość w postaci jawnej:

$$f(x) = \begin{cases} \alpha(1-x), & 0 < x < 1 \\ 0, & \text{poza} \end{cases}$$

Dla gęstości musi być spełniony warunek $\int_{-\infty}^{+\infty} f(x)\ dx = 1$

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)\ dx &= \int_0^1 \alpha(1-x)\ dx \\ &= \left[\alpha x - \frac{\alpha x^2}{2}\right]_0^1 \\ &= \alpha - \frac{\alpha}{2} \\ &= \frac{\alpha}{2} = 1 \implies \alpha = 2 \end{aligned}$$

Zatem gęstość ma postać

$$f(x) = \begin{cases} 2 - 2x, & 0 < x < 1 \\ 0, & \text{poza} \end{cases}$$

Oraz dystrybuanta:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 2x - x^2, & 0 < x < 1 \\ 1, & 1 \leq x \end{cases}$$

Wyznaczmy gęstość zmiennej $Y = e^X$ na dwa sposoby:

(1) Wyznaczymy dystrybuantę i zróżniczkujemy

$$
\begin{aligned}
F_X(t) &= P[Y \leq t] = P[e^X \leq t] \\
&= \begin{cases} 0, & t \leq 0 \\ P[X \leq \ln t], & 0 < t \end{cases} \\
&= \begin{cases} 0, & t \leq 0 \\ F_X(\ln t), & 0 < t \end{cases} \\
&= \begin{cases} 0, & t \leq 0 \\ 0, & \ln t \leq 0 \\ 2\ln t - \ln^2 t, & 0 < \ln t < 1 \\ 1, & 1 \leq \ln t \end{cases} \\
&= \begin{cases} 0, & t \leq 1 \\ 2\ln t - \ln^2 t, & 1 < t < e \\ 1, e \leq t \end{cases}
\end{aligned}
$$

Teraz różniczkując dostajemy gęstość

$$
f_X(t) = \begin{cases} 0, & t \leq 1 \\ \frac{2}{t} - 2 \cdot \frac{\ln t}{t}, & 1 < t < e \\ 0, & e \leq t \end{cases}
$$

(1) Wykorzystując twierdzenie

- $(0;1) = (0;1)$
- $a = 0 \quad b = 1$
- $y = e^x, x \in (0;1)$
- $\ln y = x$
- $g(x) = e^x \quad h(y) = \ln y$

Wyznaczymy c, d

$$c_1 = \lim_{x\to a^+} g(x) = \lim_{x\to 0^+} g(x) = 1$$
$$d_1 = \lim_{x\to b^-} g(x) = \lim_{x\to 1^-} g(x) = e$$
$$c = \min(c_1, d_1) = \min(1, e) = 1$$
$$d = \max(c_1, d_1) = \max(1, e) = e$$

Ostatecznie mamy gęstość

$$k(y) = \begin{cases} (2 - 2\ln y) \cdot \left|\frac{1}{y}\right|, & 1 < y < e \\ 0, & \text{poza} \end{cases}$$
$$= \begin{cases} \frac{2}{y} - \frac{2\ln y}{y}, & 1 < y < e \\ 0, & \text{poza} \end{cases}$$

Co się zgadza z ręcznie obliczoną gęstością.

2024-12-10

16 Nierówności związane z momentami.

16.1 Nierówność Schwarza

Jeżeli zmienne losowe X, Y mają skończone wariancje tzn. $E(X^2) < \infty$, $E(Y^2) < \infty$, to

$$(E|XY|)^2 \leq E(X^2) \cdot E(Y^2)$$

Dowód 16.1

Załóżmy, że $EX^2 > 0$ i $EY^2 > 0$.
Zauważmy ponadto, że

$$(|a| - |b|)^2 \geq 0 \iff a^2 - 2|a||b| + b^2 \geq 0$$

Stąd

$$2|a \cdot b| = 2|a| \cdot |b| \leq a^2 + b^2$$

Podstawmy $a = \overline{X} = \frac{X}{\sqrt{E(X^2)}}$ i $b = \overline{Y} = \frac{Y}{\sqrt{E(Y^2)}}$
Wtedy

$$2|\overline{X}||\overline{Y}| \leq \overline{X}^2 + \overline{Y}^2$$

Zatem

$$2E|\overline{XY}| \leq E(\overline{X}^2) + E(\overline{Y}^2)$$

$$2E\frac{|X \cdot Y|}{\sqrt{E(X^2)}\sqrt{E(Y^2)}} \leq 1 + 1 = 2$$

$$E|X \cdot Y| \leq \sqrt{E(X^2) \cdot E(Y^2)}$$

Ostatecznie:

$$(E|XY|)^2 \leq E(X^2) \cdot E(Y^2)$$

16.2 Nierówność Jensena

Niech X będzie zmienną losową taką, że $E|X| < \infty$ i niech g będzie funkcją wypukłą, taką, że $E|g(X)| < \infty$. Wtedy

$$g(EX) \leq Eg(X)$$

Achtung! 17

Dla funkcji wypukłej jej styczna w dowolnym punkcie zawsze jest nie większa niż sama funkcja.

Przykład 16.1

$$g(x) = x^2 \qquad \underbrace{(EX)^2}_{g(EX)} \leq \underbrace{E(X^2)}_{E(g(X))}$$

$$g(x) = e^x \qquad E(e^X) \geq e^{EX}$$

Dowód 16.2

- Ponieważ funkcja g jest wypukła, tzn., że w każdym punkcie jej wykres ma prostą podpierającą(wykres funkcji leży nad styczną do wykresu w dowolnym punkcie).
- Wybieramy punkt $(x_0, g(x_0))$ na krzywej $y = g(x)$ i prowadzimy prostą podpierającą (styczną):

$$y - g(x_0) = \lambda(x_0)(x - x_0)$$

gdzie $\lambda(x_0)$ – współczynnik kierunkowy prostej

lub inaczej $y = g(x_0) + (x - x_0)\lambda(x_0)$

Z wypukłości funkcji g wynika warunek $g(x) \geq y$
Mamy zatem

$$g(x) \geq g(x_0) + (x - x_0)\lambda(x_0)$$

Dokonujemy podstawienia: $x_0 = EX, x = \overbrace{X}^{\text{zmienna losowa}}$
Mamy:

$$g(X) \geq g(EX) + (X - EX)\lambda(EX)$$

$$Eg(X) \geq E\left[g(EX) + (X - EX)\lambda(EX)\right]$$

$$Eg(X) \geq \underbrace{Eg(EX)}_{\text{wartość oczekiwana liczby}} + E(X - EX) \cdot \lambda(EX)$$

$$Eg(X) \geq g(EX) + \lambda(EX)\underbrace{E(X - EX)}_{=EX-EX=0}$$

Stąd otrzymujemy

$$Eg(X) \geq g(EX)$$

16.3 Nierówność Höldera

Niech

- $p > 1, q > 1$
- $\frac{1}{p} + \frac{1}{q} = 1$
- X, Y będą zmiennymi losowymi takimi, że $E(|X|^p) < \infty$, $E(|Y|^q) < \infty$.

Wtedy $E|X \cdot Y| < \infty$ oraz

$$E|X \cdot Y| \leq (E|X|^p)^{\frac{1}{p}} \cdot (|Y|^q)^{\frac{1}{q}}$$

Dowód 16.3

Jeden z warunków na funkcję wypukłą mówi, że

$$f \text{ jest wypukła} \iff f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

gdzie $\lambda \in (0,1)$
Dla funkcji wklęsłej mamy warunek ze znakiem przeciwnym.
Zatem prawdziwa jest nierówność:

$$\log(\lambda x + (1-\lambda)y) \geq \lambda \log x + (1-\lambda)\log y$$

$$\log(\lambda x + (1-\lambda)y) \geq \log x^{\lambda} + \log y^{1-\lambda}$$

$$\log(\lambda x + (1-\lambda)y) \geq \log(x^{\lambda} \cdot y^{1-\lambda})$$

Stąd otrzymujemy:

$$x^{\lambda} \cdot y^{1-\lambda} \leq \lambda x + (1-\lambda)y$$

Podstawiamy: $\lambda := \frac{1}{p}, 1-\lambda := \frac{1}{q}$
oraz $x := w^p$ i $y := z^q$

$$(w^p)^{\frac{1}{p}} \cdot (z^q)^{\frac{1}{q}} \leq \frac{1}{p} \cdot w^p + \frac{1}{q} z^q$$

$$w \cdot z \leq \frac{w^p}{p} + \frac{z^q}{q} \text{ – nierówność Younga}$$

Niech teraz $\overline{X} := \frac{|X|}{(E|X|^p)^{\frac{1}{p}}}$ oraz $\overline{Y} := \frac{|Y|}{(E|Y|^q)^{\frac{1}{q}}}$

$$\overline{X} \cdot \overline{Y} \leq \frac{\overline{X}^p}{p} + \frac{\overline{Y}^q}{q}$$

$$\frac{|X|}{(E|X|^p)^{\frac{1}{p}}} \cdot \frac{|Y|}{(E|Y|^q)^{\frac{1}{q}}} \leq \frac{|X|^p}{p(E|X|^p)} \cdot \frac{|Y|}{q(E|Y|^q)}$$

$$\frac{E|X \cdot Y|}{(E|X|^p)^{\frac{1}{p}} \cdot (E|Y|^q)^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1$$

$$E|X \cdot Y| \leq (E|X|^p)^{\frac{1}{p}} \cdot (E|Y|^q)^{\frac{1}{q}}$$

16.4 Nierówność Czebyszewa

Niech $X \geq 0$ będzie zmienną losową o wartości oczekiwanej EX.

Wtedy dla każdego $\varepsilon > 0$ prawdziwa jest nierówność:

$$P[X \geq \varepsilon] \leq \frac{EX}{\varepsilon}$$

Dowód 16.4

$$EX = E(XI[X < \varepsilon] + XI[X \geq \varepsilon])$$

gdzie I jest identykatorem:

$$I[X \geq \varepsilon] = \begin{cases} 1, & X \geq \varepsilon \\ 0, & X < \varepsilon \end{cases} \qquad \begin{aligned} &E(I[X \geq \varepsilon]) \\ &= 1 \cdot P[X \geq \varepsilon] + 0 \cdot P[X < \varepsilon] \\ &= P[X \geq \varepsilon] \end{aligned}$$

$$\begin{aligned} EX &= E(XI[X < \varepsilon] + XI[X \geq \varepsilon]) \\ &= E(XI[X < \varepsilon]) + E(XI[X \geq \varepsilon]) \\ &\geq E(XI[X \geq \varepsilon]) \\ &\geq E(\varepsilon I[X \geq \varepsilon]) \\ &= \varepsilon P[X \geq \varepsilon] \end{aligned}$$

Ostatecznie

$$P[X \geq \varepsilon] \leq \frac{EX}{\varepsilon}$$

16.5 Nierówność Markowa

Niech $p > 0$. Wtedy dla zmiennej losowej X takiej, że $E|X|^p < \infty$ mamy:

$$P[|X| \geq \varepsilon] \leq \frac{E|X|^p}{\varepsilon^p} \qquad \text{dla dowolnego } \varepsilon > 0$$

Dowód 16.5

Wprowadzamy oznaczenie $Y := |X|^p$.

$$\begin{aligned} P[|X| \geq \varepsilon] &= P[|X|^p \geq \varepsilon^p] \\ &= P[Y \geq \varepsilon^p] \\ &\leq \frac{EY}{\varepsilon^p} \\ &= \frac{E|X|^p}{\varepsilon^p} \end{aligned}$$

16.6 Nierówność Czebyszewa-Bienaymé

Niech X będzie zmienną losową o skończonej wariancji $D^2(X) < \infty$. Wtedy dla dowolnego $\varepsilon > 0$ mamy

$$P[|X - EX| \geq \varepsilon] \leq \frac{D^2(X)}{\varepsilon^2}$$

Dowód 16.6

Wprowadzamy oznaczenie $Y := |X - EX|^2$

$$\begin{aligned} &P[|X - EX| \geq \varepsilon] \\ &\quad = P[|X - EX|^2 \geq \varepsilon^2] \\ &\quad = P[Y \geq \varepsilon^2] \\ &\quad \leq \frac{EY}{\varepsilon^2} \\ &\quad = \frac{E|X - EX|^2}{\varepsilon^2} \\ &\quad = \frac{D^2(X)}{\varepsilon^2} \end{aligned}$$

16.7 Nierówność Czebyszewa – wykładnicza

Jeśli $Ee^{pX} < \infty$ dla pewnej zmiennej losowej X i $p > 0$, to dla $\lambda \in [0, p]$ mamy:

$$P[X \geq \varepsilon] \leq \frac{Ee^{\lambda X}}{e^{\lambda \varepsilon}}$$

dla dowolnego $\varepsilon > 0$

Dowód 16.7

Wprowadzamy oznaczenie: $Y := e^{\lambda X}$

$$\begin{aligned} P[X \geq \varepsilon] &= P[e^{\lambda X} \geq e^{\lambda \varepsilon}] \\ &= P[Y \geq e^{\lambda \varepsilon}] \\ &\leq \frac{EY}{e^{\lambda \varepsilon}} \\ &= \frac{Ee^{\lambda X}}{e^{\lambda \varepsilon}} \end{aligned}$$

17 Rodzaje zbieżności zmiennych losowych

Ciąg zmiennych losowych $\{X_n, n \geq 1\}$ jest zbieżny do zmiennej losowej X

1) **prawie na pewno**, jeśli

$$P[\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)] = 1 \qquad (X_n \overset{\text{p.n.}}{\to} X)$$

2) **według prawdopodobieństwa** (według miary probabilistycznej), jeśli dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} P[\omega : |X_n(\omega) - X(\omega)| > \varepsilon] = 0 \qquad (X_n \overset{\text{p}}{\to} X)$$

3) **według p-tego momentu** (w L^p), $0 \in (0, \infty)$, jeśli $E|X| < \infty$, $E|X_n|^p < \infty$ dla $n \geq 1$ oraz

$$\lim_{n\to\infty} E|X_n - X|^p = 0 \qquad (X_n \overset{L^p}{\to} X)$$

Achtung! 18

Każda z tych zbieżności ma swój odpowiednik w analizie:

1) zbieżność prawie wszędzie
2) zbieżność według miary
3) zbieżność w L^p

17.1 Zbieżność «prawie na pewno» – własności

Jeśli $X_n \overset{\text{p.n.}}{\to} X$ oraz $Y_n \overset{\text{p.n.}}{\to} Y$, to

1) $\bigwedge_{a,b\in\mathbb{R}} aX_n + bY_n \overset{\text{p.n.}}{\to} aX + bY$
2) $X_n \cdot Y_n \overset{\text{p.n.}}{\to} X \cdot Y$
3) jeśli $P[X \neq 0] = 1$, to

$$I[X_n \neq 0] \cdot \frac{1}{X^n} \overset{\text{p.n.}}{\to} \frac{1}{X}$$

17.2 Charakteryzacja zbieżności «prawie na pewno»

Niech $\{X_n, n \geq 1\}$ będzie ciągiem zmiennych losowych, a X zmienną losową.

Następujące warunki są równoważne:

1) $X_n \overset{\text{p.n.}}{\to} X$
2) dla każdego $\varepsilon > 0$

$$\lim_{N\to\infty} P[\bigcap_{n=N}^{\infty} \{|X_n - X| \leq \varepsilon\}] = 1$$

3) dla każdego $\varepsilon > 0$

$$\lim_{N\to\infty} P[\bigcup_{n=N}^{\infty} \{|X_n - X| > \varepsilon\}] = 0$$

Dowód 17.1

Równoważność warunków (2) i (3) jest oczywista[a]. Wystarczy zatem pokazać, że równoważne są warunki (1) i (2).

$$\begin{aligned} X_n \overset{\text{p.n.}}{\to} X &\iff P[\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega) = 1] \\ &\iff \bigwedge_{\varepsilon>0} P[\omega : \{\bigvee_N \bigwedge_{n\geq N} |X_n(\omega) - X(\omega)| \leq \varepsilon\}] = 1 \\ &\iff \bigwedge_{\varepsilon>0} P[\omega : \bigcup_{N=1}^{\infty} \bigcap_{n=N} \{|X_n(\omega) - X(\omega)| \leq \varepsilon\}] = 1 \end{aligned}$$

Wprowadzamy oznaczenie:

$$B_N = \bigcap_{n=N}^{\infty} \{|X_n(\omega) - X(\omega)| \leq \varepsilon\}$$

Ciąg $\{B_N, N \geq 1\}$ tworzy ciąg wstępujący, więc na mocy twierdzenia o ciągłości miary mamy

$$\bigwedge_{\varepsilon>0} P[\bigcup_{N=1}^{\infty} B_N] = 1 \iff \bigwedge_{\varepsilon>0} \lim_{N\to\infty} P[B_N] = 1$$
$$\iff \bigwedge_{\varepsilon>0} P[\omega : \bigcap_{n=N}^{\infty} \{|X_n(\omega) - X(\omega)| \le \varepsilon\}] = 1 \quad \text{CND}$$

[a] $\left[\bigcup_{n=N}^{\infty} \{|X_n - X| > \varepsilon\}\right]' = \bigcap_{n=N}^{\infty} \{|X_n - X| \le \varepsilon\}$

17.2.1 Wniosek

Jeżeli dla każdego $\varepsilon > 0$ mamy

$$\sum_{n=1}^{\infty} P[|X_n - X| > \varepsilon] < \infty \quad \text{to} \quad X_n \overset{\text{p.n.}}{\to} X$$

Dowód 17.2

Zauważmy, że

$$P[\bigcup_{n=N}^{\infty} \{|X_n - X| > \varepsilon\}] \le \sum_{n=N}^{\infty} P[|X_n - X| > \varepsilon]$$

Wtedy

$$\lim_{N\to\infty} P\left[\bigcup_{n=N}^{\infty} \{|X_n - X| > \varepsilon\}\right]$$
$$\le \lim_{N\to\infty} \sum_{n=N}^{\infty} P[|X_n - X| > \varepsilon] \to 0 \quad \text{jako reszta szeregu zbieżnego}$$

17.3 Zbieżność «prawie na pewno» i zbieżność w/g prawdopodobieństwa

Zbieżność prawie na pewno implikuje zbieżność według prawdopodobieństwa:

$$X_n \overset{\text{p.n.}}{\to} X \implies X_n \overset{\text{p}}{\to} X$$

Dowód 17.3

Niech $\varepsilon > 0$. Zauważmy, że

$$P\left[\bigcap_{k=n}^{\infty}\{|X_k - X| \leq \varepsilon\}\right] \leq P[|X_n - X| \leq \varepsilon]$$

oraz

$$P[|X_n - X| > \varepsilon] = 1 - P[|X_n - X| \leq \varepsilon] \leq 1 - P\left[\bigcap_{k=n}^{\infty}\{|X_k - X| \leq \varepsilon\}\right]$$

Stąd mamy:

$$0 \leq \lim_{n\to\infty} P[|X_n - X| > \varepsilon] \leq 1 - \underbrace{\lim_{n\to\infty} P\left[\bigcap_{k=n}^{\infty}\{|X_k - X| \leq \varepsilon\}\right]}_{=1 \quad \text{bo } X_n \overset{\text{p.n.}}{\to} X} = 1 - 1 = 0$$

Na mocy twierdzenia o trzech ciągach mamy:

$$\lim_{n\to\infty} P[|X_n - X| > \varepsilon] = 0$$

17.4 Zbieżność w/g p-tego momentu i zbieżność w/g prawdopodobieństwa

Zbieżność w/g p-tego momentu implikuje zbieżność według prawdopodobieństwa:

$$X_n \overset{L^p}{\to} X \implies X_n \overset{\text{p}}{\to} X$$

Dowód 17.4

Niech $\varepsilon > 0$ Zauważmy, że na mocy nierówności Markowa mamy:

$$\lim_{n\to\infty}[|X_n - X| > \varepsilon] \leq \lim_{n\to\infty} \frac{E|X_n - X|^p}{\varepsilon^p} \underbrace{=}_{X_n \overset{L^p}{\to} X} 0$$

17.5 Twierdzenie Riesza[12]

Jeżeli $X_n \overset{\text{p}}{\to} X$, to istnieje podciąg $\{X_{n_k}, k \geq 1\}$ taki, że

$$X_{n_k} \overset{\text{p.n.}}{\to} X$$

2024-12-17

18 Twierdzenia graniczne

Twierdzenia graniczne formułują warunki, przy których dla ciągu zmiennych losowych $\{X_n, n \geq 1\}$ istnieje asymptotyczny rozkład i określają postać tego rozkładu.

- **Prawa wielkich liczb** – zbieżność do rozkładu jednopunktowego
- **Lokalne twierdzenia graniczne** – zbieżność funkcji gęstości lub rozkładu prawdopodobieństwa
- **Integralne twierdzenia graniczne** – zbieżność dystrybuant

18.1 Prawo wielkich liczb Bernoulliego

Spostrzeżenie: *Przy wielokrotnym rzucie monetą symetryczną częstość wystąpienia orła stabilizuje się w końcu w pobliżu $\frac{1}{2}$.*

Jakob Bernoulli w książce Ars Conjectandi (1713 r).

> Z prawdopodobieństwem dowolnie bliskim 1 można się spodziewać, iż przy dostatecznie wielkiej liczbie prób częstość danego zdarzenia losowego będzie się dowolnie mało różniła od jego prawdopodobieństwa

W dzisiejszych czasach:

Jeśli S_n oznacza liczbę sukcesów w schemacie n niezależnych prób Bernoulliego, gdzie prawdopodobieństwo sukcesu w pojedynczej próbie jest równe p, to

$$\bigwedge_{\varepsilon > 0} \lim_{n \to \infty} P[|\frac{S_n}{n} - p| \leq \varepsilon] = 1$$

[12] /Risa/

Dowód 18.1

$$P[|\frac{S_n}{n} - p| > \varepsilon] \leq \frac{E(\frac{S_n}{n} - p)^2}{\varepsilon^2} = \frac{E(S_n - np)^2}{n^2\varepsilon^2} = \frac{npq}{n^2\varepsilon^2} = \frac{pq}{n\varepsilon^2} \overset{n\to\infty}{\to} 0$$

Zatem

$$\lim_{n\to\infty} P[|\frac{S_n}{n} - p| \leq \varepsilon] = 1$$

Jako granica prawdopodobieństwa zdarzenia przeciwnego

18.1.1 Interpretacja

- Sens tego twierdzenia polega na tym, że wprowadzone przez nas określenie prawdopodobieństwa odpowiada intuicyjnemu rozumieniu prawdopodobieństwa jako granicy częstości pojawiania się zdarzenia.
- $\frac{S_n}{n}$ – można rozpatrywać jako częstość pojawiania się zdarzenia A – sukcesu, dla którego $P(A) = p$.
- Twierdzenie mówi, że w sensie zbieżności słabej (w/g prawdopodobieństwa) $\frac{S_n}{n}$ zbliża się nieograniczenie do p.

18.2 Mocne prawo wielkich liczb Bernoulliego

Przy założeniach z twierdzenia poprzedniego

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{\text{p.n.}} p$$

Dowód 18.2

W dowodzie korzystać będziemy z lematu Borela-Cantelliego i następującej nierówności Bernsteina:

$$P[|\frac{S_n}{n} - p| > \varepsilon] \leq 2e^{-\frac{n\varepsilon^2}{4}}$$

$$\sum_{n=1}^{\infty} P[|\frac{S_n}{n} - p| > \varepsilon] \leq 2\sum_{n=1}^{\infty} e^{-\frac{n\varepsilon^2}{4}} < \infty$$

Skoro

$$\sum_{n=1}^{\infty} P[|\frac{S_n}{n} - p| > \varepsilon] < \infty$$

To na mocy lematu Borela-Cantelliego mamy

$$P[\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{|\frac{S_n}{n} - p| > \varepsilon\}] = 0$$

Na podstawie warunku równoważnego zbieżności prawie na pewno i tw. o ciągłości miary mamy

$$\lim_{m\to\infty} P\left[\bigcup_{n=m}^{\infty} \{|\frac{S_n}{n} - p| > \varepsilon\}\right]$$
$$= P\left[\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{|\frac{S_n}{n} - p| > \varepsilon\}\right] = 0$$

18.3 Prawo wielkich liczb

Niech

- $\{X_n, n \geq 1\}$ będzie ciągiem zmiennych losowych o wartościach oczekiwanych $EX_n, n \geq 1$
- $S_n = X_1 + X_2 + \ldots + X_n$

Mówimy, że ciąg $\{X_n, n \geq 1\}$ spełnia **mocne prawo wielkich liczb**, jeżeli

$$\frac{S_n - ES_n}{n} \xrightarrow[n\to\infty]{\text{p.n.}} 0$$

Mówimy, że ciąg $\{X_n, n \geq 1\}$ spełnia **słabe prawo wielkich liczb**, jeżeli

$$\frac{S_n - ES_n}{n} \xrightarrow[n\to\infty]{\text{p}} 0$$

Problem prawa wielkich liczb można uogólnić i rozważać problem zbieżności

$$\frac{S_n - a_n}{b_n}$$

gdzie $\{a_n, n \geq 1\}$ jest dowolnym ciągiem liczb rzeczywistych, a $\{b_n, n \geq 1\}$ jest ciągiem liczb rzeczywistych dodatnich rozbieżnym do $+\infty$.

Historycznie rzecz biorąc pierwszym słabym prawem wielkich liczb było prawo wielkich liczb Bernoulliego.

18.4 Słabe prawo wielkich liczb Markowa

Jeżeli $\{X_n, n \geq 1\}$ jest ciągiem zmiennych losowych takich, że

$$\lim_{n\to\infty} \frac{D^2(S_n)}{n^2} = 0$$

to dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} P\left[\omega : \left|\frac{1}{n}\sum_{k=1}^{n}(X_k(\omega) - EX_k)\right| < \varepsilon\right] = 1$$

Dowód 18.3

Niech $Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$. Z założenia mamy, że

$$\lim_{n\to\infty} D^2(Y_n) = \lim_{n\to\infty} \frac{D^2(S_n)}{n^2} = 0$$

Na mocy nierówności Czebyszewa mamy

$$\begin{aligned} 1 \geq P\left[\omega : \left|\frac{1}{n}\sum_{k=1}^{n}(X_k(\omega) - EX_k)\right| < \varepsilon\right] &= P[\omega : |Y_n(\omega) - EY_n| < \varepsilon] \\ &= 1 - P[\omega : |Y_n(\omega) - EY_n| \geq \varepsilon] \\ &\geq 1 - \frac{1}{\varepsilon^2} D^2(Y_n) \to_{n\to\infty} 1 \end{aligned}$$

18.5 Słabe prawo wielkich liczb Czebyszewa

Niech $\{X_n, n \geq 1\}$ będzie ciągiem zmiennych losowych parami nieskorelowanych, o wspólnie ograniczonych wariancjach. Wtedy dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} P\left[\omega : \left|\frac{1}{n}\sum_{k=1}^{n}(X_k(\omega) - EX_k)\right| < \varepsilon\right] = 1$$

Dowód 18.4

$$\begin{aligned}1 &\geq \lim_{n\to\infty} P\left[\omega : \left|\frac{1}{n}\sum_{k=1}^{n}(X_k(\omega) - EX_k)\right| < \varepsilon\right] \\ &= 1 - \lim_{n\to\infty} P\left[\omega : \left|\frac{1}{n}\sum_{k=1}^{n}(X_k(\omega) - EX_k)\right| \geq \varepsilon\right] \\ &\geq 1 - \frac{D^2(\sum_{k=1}^{n} X_k)}{n^2\varepsilon^2} \\ &= 1 - \frac{\sum_{k=1}^{n} D^2(X_k)}{n^2\varepsilon^2} \\ &\geq 1 - \frac{nC}{n^2\varepsilon^2} \to_{n\to\infty} 1\end{aligned}$$

Achtung! 19

Twierdzenia Markowa, Bernoulliego oraz Czebyszewa wymagają by zmienne losowe $\{X_n, n \geq 1\}$ miały skończone wariancje.

18.6 Twierdzenie Kołmogorowa, mocne prawo wielkich liczb Kołmogorowa

Niech $\{X_n, n \geq 1\}$ będzie ciągiem niezależnych zmiennych losowych takich, że $D^2(X_n) < \infty, n \in \{1, 2, \dots\}$ i niech $\{b_n, n \geq 1\}$ będzie ciągiem rosnącym liczb dodatnich, takim, że

$$\lim_{n\to\infty} b_n = \infty \quad \text{oraz} \quad \sum_{n=1}^{\infty} \frac{D^2(X_n)}{b_n^2} < \infty$$

Wtedy

$$\lim_{n\to\infty} \frac{S_n - ES_n}{b_n} = 0 \quad \text{p.n.}$$

W szczególności, jeżeli

$$\sum_{n=1}^{\infty} \frac{D^2(X_n)}{n^2} < \infty$$

to $\{X_n, n \geq 1\}$ spełnia mocne prawo wielkich liczb.

18.7 Mocne prawo wielkich liczb Kołmogorowa

Jeżeli $\{X_n, n \geq 1\}$ jest ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie i $E|X_1| < \infty$, to $\{X_n, n \geq 1\}$ spełnia **mocne prawo wielkich liczb**, czyli

$$\lim_{n\to\infty} \frac{S_n}{n} = EX_1 \quad \text{p.n.}$$

19 Lokalne twierdzenia graniczne

Lokalne twierdzenia graniczne dotyczą granicznego zachowania się ciągów rozkładów zmiennych losowych wyrażonych za pomocą:

- funkcji prawdopodobieństwa $p_i = P[X = x_i]$ (dla rozkładów dyskretnych)
- funkcji gęstości $f(x)$ (dla rozkładów ciągłych)

19.1 Twierdzenie Poissona

Jeśli $X_n \sim \mathcal{B}(n, p_n)$ dla $n \in \mathbb{N}$ i spełnione są warunki:

$$p_n \to \infty \qquad np_n \to \lambda > 0, \quad \text{gdy } n \to \infty$$

to

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} \to \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \quad \text{gdy } n \to \infty$$

Dowód 19.1

Niech $\lambda_n = n \cdot p_n$, $\lambda_n \to_{n\to\infty} \lambda$, $p_n = \frac{\lambda_n}{n}$

$$\begin{aligned}
\binom{n}{k} p_n^k (1-p_n)^{n-k} &= \frac{n!}{k!(n-k)!} \cdot p_n^k (1-p_n)^{n-k} \\
&= \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\
&= \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \dots \frac{n-k+1}{n} \cdot \frac{\lambda_n^k}{k!} \left(1 - \frac{\lambda_n}{n}\right)^n \cdot \left(1 - \frac{\lambda_n}{n}\right)^{-k} \\
&= \frac{n-1}{n} \cdot \frac{n-2}{n} \dots \frac{n-k+1}{n} \cdot \frac{\lambda_n^k}{k!} \left(1 - \frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}(-\lambda_n)} \cdot \left(1 - \frac{\lambda_n}{n}\right)^{-k} \\
&\to_{n\to\infty} \frac{\lambda^k}{k!} e^{-\lambda}
\end{aligned}$$

19.2 Twierdzenie Moivre'a-Laplace'a

Niech S_n ma rozkład dwumianowy $\mathcal{B}(n,p)$, a $Y_n = \underbrace{\dfrac{S_n - np}{\sqrt{np(1-p)}}}_{\text{Odejmujemy } ES_n \text{ i dzielimy przez } D^2(S_n)}$

Wtedy

$$\lim_{n\to\infty} \sqrt{np(1-p)} P\left[Y_n = \frac{k-np}{\sqrt{np(1-p)}}\right] = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

gdzie

$$y = \lim_{n\to\infty} \frac{k-np}{\sqrt{np(1-p)}} \quad \text{i gdy } 0 \le k \le n \wedge n \to \infty$$

19.2.1 Wzory praktyczne

W szczególności dla dostatecznie dużych n mamy wzór przybliżony:

$$\begin{aligned} P[S_n = k] &\approx \frac{1}{\sqrt{np(1-p)}} \cdot \varphi\left(\frac{k-np}{\sqrt{np(1-p)}}\right) \\ &= \frac{1}{\sqrt{2\pi \cdot np(1-p)}} e^{-\frac{(k-np)^2}{2np(1-p)}} \end{aligned}$$

19.3 Centralne Twierdzenie Graniczne Lindeberga-Lévy'ego – integralne

Niech $\{X_n, n \ge 1\}$ będzie ciągiem:

- niezależnych zmiennych losowych
- o jednakowych rozkładach
- z wartością oczekiwaną $EX_1 = m$
- ze skończoną wariancją $D^2(X_1) = \sigma^2 < \infty$

Wtedy dla dowolnej liczby $t \in \mathbb{R}$

$$\lim_{n\to\infty} P\left[\frac{\sum_{i=1}^n X_i - n \cdot m}{\sigma\sqrt{n}} \le t\right] = \Phi(t)$$

gdzie $\Phi(t)$ jest dystrybuantą rozkładu $N(0,1)$

Dowód 19.2

Bez utraty ogólności możemy założyć, że $EX_i = m = 0$.
W przeciwnym przypadku rozważamy zmienne losowe

$$Y_n = X_n - m \quad \text{(wtedy } \sum_{i=1}^{n} X_i - n \cdot m = \sum_{i=1}^{n} Y_i\text{)}$$

Niech $\varphi(t) = Ee^{itX_1}$
Wiemy, że jeśli $E|X|^n < \infty$, to $\varphi_X^{(n)}(0) = i^n EX^n$
Zatem

$$\varphi'(0) = iEX_1 = 0, \quad \varphi'' = i^2 EX_1^2 = -EX_1^2 = -D^2(X_1) = -\sigma^2$$

oraz

$$\varphi(0) = 1$$

Z rozwinięcia Taylora drugiego rzędu mamy

$$\varphi(t) = \varphi(0) + \frac{\varphi'(0)}{1!}t + \frac{\varphi''(0)}{2!} + o(t^2)$$

Zatem

$$\varphi(t) = 1 + 0 - \frac{\sigma^2}{2}t^2 + o(t^2)$$

Wyznaczamy teraz funkcję charakterystyczną zmiennej

$$\frac{\sum_{i=1}^{n} X_i - n \cdot m}{\sigma\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i}{\sigma\sqrt{n}} \quad \text{bo umówiliśmy się, że } m = 0$$

$$
\begin{aligned}
Ee^{it\cdot\frac{\sum_{i=1}^n X_i}{\sigma\sqrt{n}}} &= Ee^{i\frac{t}{\sigma\sqrt{n}}\sum_{i=1}^n X_i} \\
&= E\left(e^{i\frac{t}{\sigma\sqrt{n}}X_1}\cdot e^{i\frac{t}{\sigma\sqrt{n}}X_2}\cdots e^{i\frac{t}{\sigma\sqrt{n}}X_n}\right) \\
&\overset{\text{z niezależności}}{=} E\left(e^{i\frac{t}{\sigma\sqrt{n}}X_1}\right)\cdot E\left(e^{i\frac{t}{\sigma\sqrt{n}}X_2}\right)\cdots E\left(e^{i\frac{t}{\sigma\sqrt{n}}X_n}\right) \\
&= \prod_{i=1}^n \varphi_{X_i}(\frac{t}{\sigma\sqrt{n}}) \\
&\overset{\text{jednakowe rozkłady}}{=} \prod_{i=1}^n \varphi(\frac{t}{\sigma\sqrt{n}}) \\
&= \left(\varphi(\frac{t}{\sigma\sqrt{n}})\right)^n \\
&\overset{\text{z rozwinięcia Taylora}}{=} (1-\frac{\sigma^2}{2}(\frac{t}{\sigma\sqrt{n}})^2+o((\frac{t}{\sigma\sqrt{n}})^2))^n \\
&= \left(1-\frac{t^2}{2n}+o(\frac{t^2}{\sigma n})\right)^2 \\
&= \left[\left(1-\frac{t^2}{2n}+o(\frac{t^2}{\sigma^2 n})\right)^{-\frac{2n}{t^2}}\right]^{-\frac{t^2}{2}} \to_{n\to\infty} e^{-\frac{t^2}{2}}
\end{aligned}
$$

Otrzymaliśmy postać funkcji charakterystycznej zmiennej losowej

$$Z_n = \frac{\sum_{i=1}^n X_i}{\sigma\sqrt{n}}$$

$$\varphi_{Z_n}(t) \to_{n\to\infty} e^{-\frac{t^2}{2}}$$

Rozważmy teraz zmienną losową X o rozkładzie $N(0,1)$. Wyznaczamy jej funkcję charakterystyczną

$$\begin{aligned}
\varphi_X(t) = Ee^{itX} &= \int_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\ dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx} e^{-\frac{x^2}{2}}\ dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}+itx}\ dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2-2itx}{2}}\ dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2-2itx-t^2}{2}} \cdot e^{\frac{t^2}{2}}\ dx \\
&= e^{-\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{+\infty} e^{-\frac{(x-it)^2}{2}}\ dx}_{\sqrt{2\pi}} \\
&= e^{-\frac{t^2}{2}}
\end{aligned}$$

Mamy zatem postać funkcji rozkładu normalnego $N(0,1)$, tj.

$$\varphi_X(t) = e^{-\frac{t^2}{2}}$$

co dowodzi, że rozkładem granicznym zmiennej losowej

$$\frac{\sum_{i=1}^{n} X_i - n \cdot m}{\sigma\sqrt{n}}$$

jest rozkład normalny standaryzowany $N(0,1)$

19.4 Twierdzenie Moivre'a-Laplace'a – integralne

Twierdzenie Moivre'a-Laplace'a jest wnioskiem z Centralnego Twierdzenia Granicznego Lindeberga-Levy'ego

Niech S_n będzie liczbą sukcesów w n próbach Bernoulligo z prawdopodobieństwem sukcesu p.

Wtedy dla dowolnych $x \in \mathbb{R}$

$$P\left[\frac{S_n - np}{\sqrt{np(1-p)}} \leq x\right] \to_{n\to\infty} P(Y \leq x) = \Phi(x)$$

gdzie $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\ dx$ to dystrybuanta standaryzowanego rozkładu normalnego $N(0,1)$

Dowód 19.3

Niech $S_n = X_1 + X_2 + \ldots + X_n$
gdzie zmienne losowe $\{X_n, n \geq 1\}$ są:

- niezależne
- mają jednakowy rozkład: $P[X_i = 1] = p$ i $P[X_i = 0] = 1 - p$

Wtedy

$$EX_1 = p = m \quad \text{i} \quad D^2(X_1) = p(1-p) = \sigma^2$$

Zatem na mocy Centralnego Twierdzenia Granicznego Lindeberga-Levy'ego mamy

$$P\left[\frac{S_n - np}{\sqrt{np(1-p)}} \leq x\right] = P\left[\frac{S_n - nm}{\sigma\sqrt{n}} \leq x\right] \overset{n\to\infty}{\to} \Phi(x)$$

2025-01-07

20 Warunkowa wartość oczekiwana

20.1 Warunkowa wartość oczekiwana względem zdarzenia A

- $(\Omega, \mathcal{F}, P)$ – dana przestrzeń probabilistyczna
- $A \in \mathcal{F}$, $P(A) > 0)$
- P_A – prawdopodobieństwo warunkowe względem A
- X – całkowalna zmienna losowa określona na $(\Omega, \mathcal{F}, P)$

Warunkową wartością oczekiwaną zmiennej losowej X pod warunkiem zdarzenia losowego A nazywamy liczbę $E(X|A)$:

$$E(X|A) = \int_\Omega X(\omega) dP_A$$

Jednak znacznie poręczniejszy w użyciu jest następujący, równoważny wzór:

$$E(X|A) = \frac{1}{P(A)} \int_A X(\omega) dP$$

20.2 Warunkowa wartość oczekiwana względem σ-ciała

- $(\Omega, \mathcal{F}, P)$ – dana przestrzeń probabilistyczna

Niech X będzie zmienną losową na $(\Omega, \mathcal{F}, P)$ i niech $\mathcal{A}$ będzie σ-ciałem zawartym w σ-ciele $\mathcal{F}$

Warunkową wartością oczekiwaną zmiennej losowej X pod warunkiem σ-ciała $\mathcal{A}$ nazywamy **zmienną losową** $E(X|\mathcal{A})$ spełniającą warunki:

1. $E(X|\mathcal{A})$ jest funkcją $\mathcal{A}$-mierzalną
2. dla dowolnego $A \in \mathcal{A}$

$$\int_A XdP = \int_A E(X|\mathcal{A})dP$$

Twierdzenie 20.1

Dla dowolnej zmiennej losowej całkowalnej na przestrzeni $(\Omega, \mathcal{F}, P)$ i dowolnego σ-ciała $\mathcal{A}$, $\mathcal{A} \subset \mathcal{F}$ istnieje warunkowa wartość oczekiwana $E(X|\mathcal{A})$.
Ponadto $E(X|\mathcal{A})$ jest wyznaczona jednoznacznie z dokładnością do zbioru miary zero

Dowód 20.1

Dowód jednoznaczności.
Załóżmy, że Y_1, Y_2 spełniają warunki 1) i 2) oraz

$$Y_1 \neq Y_2$$

Zauważmy, że zdarzenie $\{Y_1 > Y_2\} \in \mathcal{A}$, bo Y_1 i Y_2 są $\mathcal{A}$-mierzalne
Zatem

$$\int_{Y_1 > Y_2} Y_1 dP = \int_{Y_1 > Y_2} XdP = \int_{Y_1 > Y_2} Y_2 dP$$

Stąd

$$\int_{Y_1 > Y_2} (Y_1 - Y_2)dP = 0$$

$Y_1 - Y_2 > 0$, zatem miara zbioru $\{Y_1 > Y_2\}$ musi być równa zero:

$$P[Y_1 > Y_2] = 0$$

To samo rozumowanie zastosowane do $\{Y_2 > Y_1\}$ pokazuje ostatecznie, że

$$P[Y_1 \neq Y_2] = 0$$

20.2.1 Własność 1

Jeśli $A = \{\varnothing, \Omega\}$, to

$$E(X|\mathcal{A}) = EX \quad \text{p.n.}$$

Dowód 20.2

$A = \{\varnothing, \Omega\}$ – jest σ-ciałem atomowym.
Wtedy każda zmienna losowa X, $\mathcal{A}$-mierzalna jest postaci:

$$X = a_1 I_\varnothing + a_2 I_\Omega = a_2$$

Zatem

$$\int_\Omega E(X|\mathcal{A}) = \int_\Omega E(a_2|\mathcal{A}) = \int_\Omega a_2 dP$$

Stąd

$$\int_\Omega \left(E(X|\mathcal{A}) - a_2\right) dP = 0$$

co oznacza, że

$$E(X|\mathcal{A}) = a_2 \quad \text{p.n.}$$

Ponadto mamy:

$$\int_\Omega E(X|\mathcal{A}) dP = \underbrace{\int_\Omega X dP}_{EX} = \int_\Omega a_2 dP = a_2$$

Stąd

$$EX = a_2$$

co dowodzi ostatecznie, że

$$E(X|\mathcal{A}) = EX \quad \text{p.n.}$$

20.2.2 Własność 2

Jeśli X jest $\mathcal{A}$-mierzalna, to

$$E(X|\mathcal{A}) = X \quad \text{p.n.}$$

Dowód 20.3

Z definicji, dla dowolnego $A \in \mathcal{A}$ mamy:

$$\int_A E(X|\mathcal{A})dP = \int_A XdP \iff \int_A E(X|\mathcal{A})dP - \int_A XdP$$
$$\int_A (E(X|\mathcal{A}) - X)\, dP = 0 \iff E(X|\mathcal{A}) - X = 0 \quad \text{p.n.}$$

$$E(X|\mathcal{A}) = X \quad \text{p.n.}$$

20.2.3 Własność 3

Jeśli $X \geq 0$, to

$$E(X|\mathcal{A}) \geq 0 \quad \text{p.n.}$$

Dowód 20.4

Niech $\varepsilon > 0$. Wprowadzamy oznaczenie

$$A_\varepsilon = \{\omega : E(X|\mathcal{A}) < -\varepsilon\}$$

$A_\varepsilon \in \mathcal{A}$ zatem

$$\bigwedge_{\varepsilon>0} 0 \leq \int_{A_\varepsilon} XdP = \int_{A_\varepsilon} E(X|\mathcal{A})dP \leq -\varepsilon P[A_\varepsilon] \leq 0$$

co dowodzi, że $\int_{A_\varepsilon} XdP = 0$ i stąd, (bo $X \geq 0$),

$$A[A_\varepsilon] = 0 \iff P[\omega : E(X|\mathcal{A}) < -\varepsilon] = 0$$
$$\iff E(X|\mathcal{A}) \geq 0 \quad \text{p.n.}$$

20.2.4 Własność 4

$$|E(X|\mathcal{A})| \leq E(|X| \mid \mathcal{A}) \quad \text{p.n.}$$

Dowód 20.5

Zauważmy, że

$$X \geq Y \implies E(X|\mathcal{A}) \geq E(Y|\mathcal{A}) \quad \text{p.n.}$$

Rzeczywiście, jeżeli $X \geq Y$, to $Z = X - Y \geq 0$.
Wtedy na mocy Własności 3, $E(Z|\mathcal{A}) \geq 0$ p.n.

$$\begin{aligned}
\bigwedge_{A\in\mathcal{A}} 0 \leq \int_A E(Z|\mathcal{A})dP &= \int_A ZdP \\
&= \int_A (X-Y)dP \\
&= \int_A XdP - \int_A YdP \\
&= \int_A E(X|\mathcal{A})dP - \int_A E(Y|\mathcal{A})dP \\
&= \int_A \left(E(X|\mathcal{A}) - E(Y|\mathcal{A})\right) dP
\end{aligned}$$

$$\implies E(X|\mathcal{A}) - E(Y|\mathcal{A}) \geq 0 \text{ p.n.} \iff E(X|\mathcal{A}) \geq E(Y|\mathcal{A}) \text{ p.n.}$$

Korzystając z powyższego faktu i nierówności

$$-|X| <= X \leq |X|$$

mamy

$$E(-|X| \mid \mathcal{A}) \leq E(X|\mathcal{A}) \leq E(|X| \mid \mathcal{A}) \quad \text{p.n.}$$

Ponadto dla każdego $A \in \mathcal{A}$ mamy

$$\int_A E(-|X| \mid \mathcal{A})dP = \int_A (-|X|)dP = \int_A E(|X| \mid \mathcal{A})dP$$

Stąd

$$\begin{aligned}
&\int_A \left[E(-|X| \mid \mathcal{A}) + E(|X| \mid \mathcal{A})\right] dP = 0 \quad \text{p.n.} \\
&\implies E(-|X| \mid \mathcal{A}) + E(|X| \mid \mathcal{A}) = 0 \quad \text{p.n.} \\
&\iff E(-|X| \mid \mathcal{A}) = -E(|X| \mid \mathcal{A}) \quad \text{p.n.}
\end{aligned}$$

Łącząc zatem dwie zależności:

$$E(-|X| \mid \mathcal{A}) \leq E(X|\mathcal{A}) \leq E(|X| \mid \mathcal{A}) \quad \text{p.n.}$$

i

$$E(-|X| \mid \mathcal{A}) = -E(|X| \mid \mathcal{A}) \quad \text{p.n.}$$

dostajemy

$$-E(|X| \mid \mathcal{A}) \leq E(X|\mathcal{A}) \leq E(|X| \mid \mathcal{A}) \quad \text{p.n.}$$

co jest równoważne

$$|E(X|\mathcal{A}) \leq E(|X| \mid \mathcal{A}) \quad \text{p.n.}|$$

20.2.5 Własność 5

Dla dowolnych $a, b \in \mathbb{R}$ mamy

$$E\left((aX + bY) \mid \mathcal{A}\right) = aE(X|\mathcal{A}) + bE(Y|\mathcal{A}) \quad \text{p.n.}$$

Dowód 20.6

$$\begin{aligned}\bigwedge_{A \in \mathcal{A}} \int_A E\left((aX + bY) \mid \mathcal{A}\right) dP &= \int_A (aX + bY) dP \\ &= a \int_A X dP + b \int_A Y dP \\ &= a \int_A E(X|\mathcal{A}) dP + b \int_A E(Y|\mathcal{A}) dP \\ &= \int_A \left(aE(X|\mathcal{A}) + bE(Y|\mathcal{A})\right) dP\end{aligned}$$

Mamy zatem

$$\begin{aligned}&\int_A E\left((aX + bY)|\mathcal{A}\right) dP = \int_A \left(aE(X|\mathcal{A}) + bE(Y|\mathcal{A})\right) dP \\ &\implies E\left((aX + bY)|\mathcal{A}\right) = aE(X|\mathcal{A}) + bE(Y|\mathcal{A}) \quad \text{p.n.}\end{aligned}$$

2025-01-14

20.2.6 Własność 6

Jeżeli $X_1 \le X_2 \le \dots$ jest niemalejącym ciągiem zmiennych losowych zbieżnych do zmiennej losowej X ($X_n \nearrow X$), to

$$E(X_n|\mathcal{A}) \nearrow_{n\to\infty} E(X|\mathcal{A}) \quad \text{p.n.}$$

Dowód 20.7

Jeżeli $X_1 \le X_2 \le \dots$, to $E(X_1|\mathcal{A}) \le E(X_2|\mathcal{A}) \le \dots$
Zatem na mocy tw. Lebesquea o monotonicznej zbieżności granica $\lim\limits_{n\to\infty} E(X_n|\mathcal{A})$ istnieje i jest funkcją $\mathcal{A}$-mierzalną oraz

$$E(X|\mathcal{A}) = E(\lim_{n\to\infty} X_n|\mathcal{A}) \quad \text{jest również } \mathcal{A}\text{-mierzalną}$$

$$\begin{aligned}
\bigwedge_{A\in\mathcal{A}} \int_A \lim_{n\to\infty} E(X_n|\mathcal{A})\, dP &= \lim_{n\to\infty} \int_A E(X_n|\mathcal{A})\, dP \\
&= \lim_{n\to\infty} \int_A X_n\, dP \\
&= \int_A \lim_{n\to\infty} X_n\, dP \\
&= \int_A E(\lim_{n\to\infty} X_n|\mathcal{A})\, dP \\
&= \int_A E(X|\mathcal{A})\, dP
\end{aligned}$$

$$\iff \lim_{n\to\infty} E(X_n|\mathcal{A}) = E(X|A) \quad \text{p.n.}$$

20.2.7 Własność 7

Jeśli $\mathcal{A}_1 \subset \mathcal{A}_2$, to

$$E(X|\mathcal{A}_1) = E\left(E(X|\mathcal{A}_2)|\mathcal{A}_1\right) = E\left(E(X|\mathcal{A}_1)|\mathcal{A}_2\right)$$

Dowód 20.8

$E(X|\mathcal{A}_1)$ jest zmienną losową $\mathcal{A}_1$-mierzalną, zatem $E(X|\mathcal{A}_1)$ jest również zmienną losową $\mathcal{A}_2$-mierzalną, bo $\mathcal{A}_1 \subset \mathcal{A}_2$.

Stąd

$$E(E(X|\mathcal{A}_1)|\mathcal{A}_2) = E(X|\mathcal{A}_1) \quad \text{p.n.}$$

Ponadto $E(E(X|\mathcal{A}_2)|\mathcal{A}_1)$ jest zmienną losową $\mathcal{A}_1$-mierzalną.
Zatem

$$\bigwedge_{A\in\mathcal{A}_1} \int_A E(E(X|\mathcal{A}_2)|\mathcal{A}_1)\ dP = \int_A E(X|\mathcal{A}_2)\ dP$$
$$\underbrace{=}_{A\in\mathcal{A}_2} \int_A X\ dP$$
$$= \int_A E(X|\mathcal{A}_1)\ dP$$

Zatem

$$E(E(X|\mathcal{A}_2)|\mathcal{A}_1) = E(X|\mathcal{A}_1) \quad \text{p.n.}$$

20.2.8 Własność 8

$$EX = E(E(X|\mathcal{A}))$$

Dowód 20.9

$$E(E(X|\mathcal{A})) = \int_\Omega E(X|\mathcal{A})\ dP = \int_\Omega X\ dP = EX$$

20.2.9 Własność 9

Jeżeli zmienna losowa X jest niezależna od σ-ciała, to

$$E(X|\mathcal{A}) = EX \quad \text{p.n.}$$

Dowód 20.10

- EX jest stałą, jest więc funkcją $\mathcal{A}$-mierzalną
- Niech $A \in \mathcal{A}$
  X jest niezależna od σ-ciała $\mathcal{A}$,
  zatem zmienne losowe X i I_A są niezależne

Stąd dla dowolnego $A \in \mathcal{A}$ mamy

$$\begin{aligned}\int_A E(X|\mathcal{A})\, dP &= \int_A X\, dP \\ &= \int_\Omega XI_A\, dP \\ &= E(X \cdot I_A) \\ &= EX \cdot EI_A \\ &= EX \cdot \int_\Omega I_A\, dP \\ &= \int_\Omega EX \cdot I_A\, dP \\ &= \int_A EX\, dP\end{aligned}$$

czyli

$$E(X|\mathcal{A}) = EX \quad \text{p.n.}$$

20.2.10 Własność 10

Jeżeli Y jest ograniczoną zmienną losową $\mathcal{A}$-mierzalną, to

$$E(X \cdot Y \mid \mathcal{A}) = YE(X|\mathcal{A}) \quad \text{p.n.}$$

$YE(X|\mathcal{A})$ jest $\mathcal{A}$-mierzalna

Twierdzenie 20.2

Niech

- $\{A_i : i \in \mathbf{I}\}$ jest rozbiciem Ω, tzn.
  - $\Omega = \bigcup_{i \in \mathbf{I}} A_i$,
  - $A_i \cap A_j = \varnothing, i \neq j$
- $\mathcal{G} = \sigma(\{A_i, i \in I\})$

Wtedy

$$E(X|\mathcal{G}) = \sum_{i \in \mathbf{I}} E(X|A_i) \cdot I_{A_i}$$

21 Moment zatrzymania

21.1 Filtracja

- Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną – modelem doświadczenia losowego
- Martyngały – narzędzie do analizy zjawisk rozwijających się w czasie.
- W tym celu musimy wzbogacić model doświadczenia losowego o niemalejącą rodzinę σ-ciał $\mathcal{F}_t, t \in T$
- T – zbiór indeksów

Filtracją nazywamy niemalejącą rodzinę σ-ciał $\{\mathcal{F}_t, t \in T\}$, $\mathcal{F}_t \subset \mathcal{F}$, dla $t \in T$

21.2 Interpretacja i przykład

Każde takie σ-ciało $\mathcal{F}_t$ interpretujemy jako rodzinę zdarzeń, o których wiemy, czy zaszły, czy też nie zaszły do chwili t

$$T = \mathbb{N}, \qquad \mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$$

21.3 Rodzina zmiennych losowych adaptowana do filtracji

Rodzina zmiennych losowych $\{X_t, t \in T\}$ jest **adaptowana** do filtracji $\{\mathcal{F}_t, t \in T\}$, jeżeli zmienne losowe X_t są $\mathcal{F}_t$-mierzalne dla $t \in T$

W szczególności rodzina zmiennych losowych $\{X_t, t \in T\}$ jest adaptowana do naturalnej filtracji $\{\mathcal{F}_t, t \in T\}$, gdzie

$$\mathcal{F}_t = \sigma(\{X_s, s \leq t\}), \quad t \in T$$

21.4 Moment stopu

Momentem stopu względem filtracji $\{\mathcal{F}_t, t \in T\}$ nazywamy zmienną losową τ:

$$\tau : \Omega \to T \cup \{+\infty\}$$

spełniającą warunek

$$\bigwedge_{t \in T} \{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$$

Moment stopu nazywa się czasem momentem Markowa lub momentem zatrzymania (ang.: stopping time)

Przykład 21.1

- Rozważamy grę, w której 2 graczy A, B rzuca monetą.
- Jeśli wypadnie «orzeł», gracz A wygrywa od gracza B złotówkę, w przeciwnym razie gracz B wygrywa złotówkę
- Gra toczy się tak długo, jak długo obaj gracze mają pieniądze.
- Zakładamy, że gracz A postanawia wycować się, gdy po raz pierwszy wygra 7 złotych
- Notuje więc swoje kolejne wygrane – są one zmiennymi losowymi

$$\xi_1, \xi_2, \dots,$$

  o wartościach ze zbioru $\{-1, 1\}$
- Łączna wygrana w chwili n to:

$$X_n = \sum_{i=1}^{n} \xi_i, \quad n = 1, 2, \dots$$

- przyjmujemy, że $X_0 \equiv 0$
- Chwila wycofania się z gry gracza A jest zmienną losową, która definiujemy następująco:

$$\tau_7 = \inf\{n : X_n = 7\}$$

- Zmienna losowa $\boldsymbol{\tau_7}$ jest **czasem zatrzymania**
- Zauważmy, że gracz A może nie doczekać się chwili gdy wygra 7 zł. (gdy np. moneta jest niesymetrczna lub gracz A ma ograniczony kapitał)

  Oznacza to, że $P[\tau_7 = +\infty] > 0$. Zatem

$$\tau_7 : \Omega \to \mathbb{R} \cup \{+\infty\}$$

  – τ_7 jest zmienną losową niewłaściwą.
- Ponadto w każdej chwili wiadomo, czy moment τ_7 już nastąpił, czy też nie, zatem wiemy, czy zdarzenie $\{\tau_7 \leq t\}$ już nastąpiło, co oznacza, że

$$\{\tau_7 < t\} \in \mathcal{F}_t$$

- τ_7 jest zatem czasem zatrzymania

22 Martyngały, nadmartyngały i podmartyngały

Niech

- $(\Omega, \mathcal{F}, P)$ – przestrzeń probabilistyczna
- $\{\mathcal{F}_t, t \in T\}$ – filtracja
- $\{X_t, t \in T\}$ – adaptowana rodzina zmiennych losowych

Martyngał Rodzinę $(X_t, \mathcal{F}_t)_{t\in T}$, gdzie zmienne losowe X_t są całkowalne dla $t \in T$, jest

- **martyngałem** , jeżeli dla $s, t \in T$ i $s \leq t$

$$E(X_t|\mathcal{F}_s) = X_s \quad \text{p.n.}$$

- **nadmartyngałem** , jeżeli dla $s, t \in T$ i $s \leq t$

$$E(X_t|\mathcal{F}_s) \leq X_s \quad \text{p.n.}$$

- **podmartyngałem** , jeżeli dla $s, t \in T$ i $s \leq t$

$$E(X_t|\mathcal{F}_s) \geq X_s \quad \text{p.n.}$$

22.1 Interpretacja

- Martyngał jest modelem gry sprawiedliwej w takim sensie, że:

  średnia wygranej w chwili t, gdy znamy przebieg gry do chwili s, $E(X_t|\mathcal{F}_s)$(zdarzenia, o których wiemy, czy zaszły, czy nie, jeżeli obserwujemy grę do chwili s tworzą σ-ciało $\mathcal{F}_s$) jest równa X_s, czyli wygranej w chwili s.

- Nadmartyngał jest modelem gry niekorzystnej dla gracza.

- Podmartyngał jest modelem gry korzystnej dla gracza

22.2 Własności

Dla martyngału:

$s, t \in T$, $s \le t$ i $A \in \mathcal{F}_s$

$$\int_A X_t \, dP = \int_A X_s \, dP$$

Dowód 22.1

Dla martyngału mamy warunek $E(X_t | \mathcal{F}_s) = X_s$

$$\bigwedge_{A \in \mathcal{F}_s} \int_A X_s \, dP = \int_A E(X_t | \mathcal{F}_s) \, dP = \int_A X_t \, dP$$

Jeśli $T = \mathbb{N}$, to warunek równoważny przyjmuje postać

$$E(X_{n+1} | \mathcal{F}_n) = X_n \quad n = 1, 2, \ldots \quad \text{p.n.}$$

lub

$$E((X_{n+1} - X_n) \mid \mathcal{F}_n) = 0 \quad n = 1, 2, \ldots \quad \text{p.n.}$$

22.3 Ciąg prognozowalny

Niech $(X_n, \mathcal{F}_n)_{n=1}^{\infty}$ będzie martyngałem i $\mathcal{F}_{-1} = \{\varnothing, \Omega\}$

Ciąg zmiennych losowych $\{V_n, n \ge 0\}$ nazywamy ciągiem prognozowalnym, jeżeli zmienna losowa V_n jest $\mathcal{F}_{n-1}$-mierzalna dla $n = 0, 1, 2, \ldots$

22.4 Transformata martyngałowa

- Niech $(X_n, \mathcal{F}_n)_{n=1}^{\infty}$ będzie martyngałem i $\mathcal{F}_{-1} = \{\varnothing, \Omega\}$
- $\{V_n, n \ge 0\}$ – ciąg prognozowalny

Jeżeli założymy dodatkowo, że $V_n, n \ge 0$ są ograniczone i zdefiniujemy

$$Z_n = V_0 X_0 + V_1 (X_1 - X_0) + \ldots + V_n (X_n - X_{n-1})$$

to $(Z_n, \mathcal{F}_n, n \ge 0)$ jest martyngałem.

$$
\begin{aligned}
E((Z_n - Z_{n-1})|\mathcal{F}_{n-1}) &= E(V_n(X_n - X_{n-1})|\mathcal{F}_{n-1}) \\
&\overset{\text{własność 10}}{=} V_n E((X_n - X_{n-1})|\mathcal{F}_{n-1}) \\
&= 0
\end{aligned}
$$

Transformatą martyngałową nazywamy ciąg $\{Z_n, n \geq 1\}$

22.4.1 Interpretacja

- Niech $X_n, n \geq 1$ oznacza łączną wygraną gracza A do chwili n.
- Grę obserwuje gracz C, który stawia na gracza A w n-tej partii V_n złotych (jego decyzja o wysokości stawki zależy od historii gry do chwili $n-1$, stąd warunek, że V_n jest **ciągiem prognozowalnym**) i wygrywa

  $$V_n(X_n - X_{n-1}) \text{ złotych}$$

- $Z_n = V_0X_0 + V_1(X_1 - X_0) + \ldots + V_n(X_n - X_{n-1})$ – wygrana gracza C

Gracz C uważa, że wygra, choć jego wygrana i wygrana gracza A tworzą martyngał.

2025-01-21

22.5 Jak wygrać milion dolarów

22.5.1 Strategia

Dla uproszczenia zakładamy, że chcemy wygrać 1 zł.

- Niech $\{Y_n, n \geq 1\}$ będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie

  $$P[Y_n = 1] = p = 1 - P[Y_n = -1]$$

- Jeżeli $Y_n = 1$, to uznajemy, że w n-tej kolejce wygraliśmy.
- Rozważamy grę z następującym sposobem obstawiania:

  najpierw stawiamy 1 zł, a potem

  - podwajamy stawkę, gdy przegrywamy
  - wycofujemy się z gry, gdy po raz pierwszy wygramy

- Stawka w grze

$$W_1 = 1, \quad W_n = \begin{cases} 2^{n-1}, & \text{gdy } Y_i = -1 \quad \forall i \in \{1, \dots, n-1\} \\ 0, & \text{w pozostałych przypadkach} \end{cases}, \ n \geq 2$$

- $W_n, n \geq 1$ jest ciągiem prognozowalnym
- $S_n = \sum_{i=1}^{n} Y_i$ jest martyngałem gdy $p = \frac{1}{2}$
- zgromadzony kapitał w grze (wygrana) – transformata:

$$X_n = \sum_{k=1}^{n} W_k Y_k = \sum_{k=1}^{n} W_k (S_k - S_{k-1}) = X_{n-1} + W_n X_n$$

- $X_n = \sum_{k=1}^{n} W_k Y_k = \sum_{k=1}^{n} W_k (S_k - S_{k-1}) = X_{n-1} + W_n X_n$
- X_n – martyngał dla $p = \frac{1}{2}$
- Gdy $\bigwedge_{i \in \{1,2,\dots,n\}} Y_i = -1$, to gracz przegrał

$$X_n = -\sum_{i=1}^{n} W^{i-1} = -(1 + 2 + 2^2 + \dots + 2^n) = -(2^n - 1)$$

- Gdy $Y_{n+1} = 1$, to

$$X_{n+1} = X_n + W_{n+1} \cdot 1 = -(2^n - 1) + 2^n = 1$$

- W tym momencie kończymy grę (wygraliśmy)

22.5.2 Analiza gry – Czy na pewno wygramy?

Niech

$$\tau = \inf\{n \geq 1, X_n = 1\} \quad \text{– moment stopu}$$

Dla $p = \frac{1}{2}$ mamy

$$P[\tau = n] = \frac{1}{2^n}$$

$$\begin{aligned} P[\tau < \infty] &= P\left[\bigcup_{i=1}^{\infty} \{\tau = n\}\right] \\ &= \sum_{i=1}^{\infty} P[\tau = n] \\ &= \sum_{i=1}^{\infty} \frac{1}{2^n} \\ &= 1 \end{aligned}$$

Z prawdopodobieństwem równym 1 wygramy.

22.5.3 Kiedy wygramy?

$$\begin{aligned} E_\tau &= \sum_{n=1}^{\infty} nP[\tau = n] \\ &= \sum_{n=1}^{\infty} n\frac{1}{2^n} \\ &= \frac{1}{2}\sum_{n=1}^{\infty} n\frac{1}{2^{n-1}} \\ &= \frac{1}{2}\frac{1}{(1-\frac{1}{2})^2} \\ &= 2 \end{aligned}$$

Przeciętnie, powinniśmy wygrać już w drugiej rundzie obstawiania.

Korzystamy ze wzoru

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

22.5.4 Ile wygramy?

Wiemy, że

- τ – moment wygranej
- $X_\tau \equiv 1$ – wielkość wygranej w chwili τ
- Przeciętna wielkości wygranej:

$$EX_\tau = 1 \cdot P[X_\tau = 1] = 1$$

22.5.5 Wnioski

- W grze sprawiedliwej istnieje strategia gry, w której
  - w skończonym czasie
  - z prawdopodobieństwem 1

  gracz wygrywa 1 zł
- Choć średni czas oczekiwania τ na wygraną jest równy 2, nie jest jednak z góry ograniczony.
- Dlatego pewnym wygranej można być tylko w przypadku, gdy mamy nieograniczony kapitał