

Uniwersytet Marii Curie Skłodowskiej w Lublinie

Wydział Matematyki, Fizyki i Informatyki

Rachunek prawdopodobieństwa

I stopień studiów na kierunku Matematyka

Piotr Pawlas

Spis treści

1	Wstęp	6
2	Krótko o historii probabilistyki	7
3	Elementy kombinatoryki	8
3.1	Wariacje z powtórzeniami	9
3.2	Wariacje bez powtórzeń	9
3.3	Kombinacje bez powtórzeń	10
3.4	Kombinacje z powtórzeniami	11
3.5	Zadania	12
4	Przestrzenie probabilistyczne	13
4.1	Wprowadzenie	13
4.2	Zdarzenie	14
4.3	σ -ciało	15
4.4	Miara probabilistyczna	16
4.5	Różne definicje prawdopodobieństwa	21
4.6	Zadania	23
5	Prawdopodobieństwo warunkowe	24
5.1	Definicja i własności	24
5.2	Prawdopodobieństwo całkowite i wzór Bayesa	27
5.3	Zadania	30
6	Niezależność	32
6.1	Niezależność dwóch zdarzeń	32
6.2	Niezależność zbioru zdarzeń	37
6.3	Schemat Bernoulli'ego	38
6.4	Lematy Borela-Cantelliego	39

6.5	Zadania	40
7	Zmienne losowe jednowymiarowe	41
7.1	Pojęcie zmiennej losowej	41
7.2	Zmienne losowe dyskretne	43
7.3	Zmienne losowe absolutnie ciągłe	45
7.4	Zadania	46
8	Dystrybuanta zmiennej losowej	48
8.1	Dystrybuanta i jej własności	48
8.2	Przypadek dyskretny	50
8.3	Przypadek ciągły	50
8.4	Zadania	52
9	Przykłady rozkładów dyskretnych	54
9.1	Rozkład Bernoulliego	54
9.2	Rozkład jednostajny na zbiorze skończonym	54
9.3	Rozkład dwumianowy	54
9.4	Rozkład hipergeometryczny	55
9.5	Rozkład geometryczny	56
9.6	Rozkład ujemny dwumianowy. Rozkład Pascala.	57
9.7	Rozkład Poissona	57
9.8	Zadania	58
10	Przykłady rozkładów ciągłych	61
10.1	Rozkład jednostajny	61
10.2	Rozkład wykładniczy	62
10.3	Rozkład normalny	63
10.4	Rozkład gamma	63
10.5	Rozkład beta	64

10.6 Zadania	64
11 Zmienne losowe wielowymiarowe	65
11.1 Pojęcie wektora losowego	65
11.2 Niezależność zmiennych losowych	68
11.3 Zadania	69
12 Funkcje zmiennych losowych	70
12.1 Funkcje jednowymiarowych zmiennych losowych	70
12.2 Przypadek wielowymiarowy	71
12.3 Zadania	72
13 Charakterystyki zmiennych losowych	73
13.1 Wartość oczekiwana. Momenty zwykłe	73
13.2 Ryzyko w ujęciu inwestycyjnym	79
13.2.1 Wartość oczekiwana-oczekiwana stopa zwrotu	79
13.2.2 Miara ryzyka	81
13.2.3 Wariancja stopy zwrotu. Ujemna semiwariancja	82
13.2.4 Odchylenie standardowe. Ujemne semiodchylenie stan- dardowe	83
13.2.5 Odchylenie przeciętne. Ujemne semiodchylenie przeciętne	85
13.3 Wariancja, kowariancja i współczynnik korelacji	85
13.3.1 Związki między inwestycjami	85
13.4 Zadania	89
14 Funkcja tworząca prawdopodobieństwa i funkcja charaktery- styczna	93
14.1 Funkcja tworząca prawdopodobieństwa	93
14.2 Funkcja charakterystyczna	95
14.3 Zadania	98

15 Rodzaje zbieżności ciągów zmiennych losowych	100
15.1 Podstawowe rodzaje zbieżności	100
15.2 Zadania	102
16 Prawa wielkich liczb	103
16.1 Słabe prawa wielkich liczb	103
16.2 Mocne prawa wielkich liczb	107
16.3 Zadania	108
17 Centralne twierdzenie graniczne	110
17.1 Słaba zbieżność ciągów zmiennych losowych	110
17.2 Centralne twierdzenie graniczne	111
17.3 Zadania	113
18 Proces Poissona	115

1 Wstęp

W języku potocznym pojęcie prawdopodobieństwa może mieć dwie różne interpretacje.

- Interpretacja częstościowa.

Założmy, że w tych samych warunkach wykonujemy pewną liczbę doświadczeń losowych (na przykład rzucamy dwukrotnie kostką sześcienną do gry). Przez prawdopodobieństwo realizacji pewnego zdarzenia (dla przykładu uzyskania dwóch szóstek) możemy rozumieć iloraz liczby sukcesów i liczby wszystkich możliwych wyników tego doświadczenia.

- Interpretacja subiektywna.

Przypuśćmy, że pewien pacjent przed operacją pyta swojego lekarza o prawdopodobieństwo tego, że operacja przebiegnie pomyślnie. Odpowiedź zależy będzie wtedy od wielu czynników (stanu pacjenta, rodzaju choroby, wieku, itd.) Nie ma więc sensu pytać o częstość sukcesów w liczbie podobnych doświadczeń.

2 Krótko o historii probabilistyki

Rachunek prawdopodobieństwa, zwany inaczej probabilistyką jest działem matematyki nieco odmiennym od geometrii, algebry czy analizy. Kojarzy się często z hazardem, zabawą i rozrywką. Nie jest to całkowity przypadek, bowiem spore zasługi w rozwoju tej dziedziny matematyki miały warstwy społeczeństwa szukające rozrywki w hazardzie. Z początkami rozwoju niezwykle ważnego dziś działu matematyki jakim jest probabilistyka, związane są modne w XVII wieku gry hazardowe; w szczególności gra w kości. Mimo, że gra w kości znana była już w starożytności, pierwsze teoretyczne analizy związane z tą grą pochodzą od matematyków francuskich z połowy XVII wieku: Pascala i Fermata. Znamienne w rozwoju dialogu pomiędzy tymi uczonymi wydaje się rola Kawalera de Méré (prawdziwe nazwisko Antoine Gombaud), który sam nie będąc matematykiem niejako zainspirował Pascala a w konsekwencji Fermata do szczegółowych badań związanych z różnymi problemami pojawiającymi się w analizie gier w kości. W 1657 roku Christiaan Huygens w swojej książce “De ratiociniis in ludo aleae” opisuje fundamentalne pojęcia rachunku prawdopodobieństwa, podaje także wzór na wartość oczekiwaną zmiennej losowej przyjmującej skończoną liczbę wartości. Następnym ważnym dziełem jest “Ars conjectandi” (1713), w którym Jaques Bernoulli rozszerza wyniki Huygansa. Udowadnia między innymi, za pomocą rachunku kombinatorycznego Prawo Wielkich Liczb. W 1733, w pracy “The doctrine of chances”, Abraham de Moivre (1667-1754) szacuje szybkość zbieżności w Prawie Wielkich Liczb, jest to pierwsza znana wersja centralnego twierdzenia granicznego. Następnie te rezultaty uogólnia Pierre-Simon Laplace rozwijając w dziele “Théorie analytique des probabilités” (1812) teorię funkcji tworzących i funkcji charakterystycznych czym uczynił olbrzymi krok dla dalszego rozwoju probabilistyki. Dalsze wyniki dotyczące Praw Wielkich Liczb i Centralnego Twierdzenia Granicznego były dziełem takich matematyków jak: Denis Poisson (1781-1840), Pafnouti

Tchebychev (1821-1894), Andriej Markov (1856-1922) i Alexandre Liapounov (1857-1918). W XVIII i XIX wieku rachunek prawdopodobieństwa zaczęto także stosować w różnych dziedzinach związanych z matematyką tj.:

- statystyce,
- zagadnieniach ubezpieczeniowych,
- rachunku błędów,
- demografii.

W XX wieku, teoria miary i całki pozwoliła sformalizować rachunek prawdopodobieństwa. Nadano nowy sens pojęciom takim jak: miara probabilistyczna, zmienna losowa, wartość oczekiwana, rozkład, rozkład warunkowy, itd.

W tym miejscu należy przytoczyć nazwisko Andrieja Nikołajewicza Kołmogorowa-rosyjskiego matematyka, uznawanego za twórcę współczesnej teorii prawdopodobieństwa. W 1933 roku podał on w pracy "Podstawowe pojęcia teorii prawdopodobieństwa" aksjomatyczną definicję miary prawdopodobieństwa.

Od pierwszej połowy XX wieku rozwinęła się nowa gałąź rachunku prawdopodobieństwa: procesy stochastyczne. Dała ona nowy bodziec do rozwoju nie tylko teorii ale także zastosowań probabilistyki w takich dziedzinach jak ekonomia, biologia, fizyka oraz informatyka.

3 Elementy kombinatoryki

Kombinatoryka to dział matematyki zajmujący się teorią obliczania liczby elementów zbiorów skończonych. Poniżej przedstawimy podstawowe pojęcia kombinatoryczne wykorzystywane w rachunku prawdopodobieństwa.

Punktem wyjścia naszych rozważań będzie zbiór skończony Ω , który nazywać będziemy populacją oraz k -elementowa próba pobrana z elementów tego zbioru. Od tej chwili zakładamy, że $|\Omega| = n$, gdzie symbol $|X|$ oznacza moc zbioru X .

3.1 Wariacje z powtórzeniami

Definicja 3.1. *K-elementową próbą uporządkowaną z powtórzeniami wybraną z populacji Ω nazywamy dowolny k-elementowy ciąg (x_1, \dots, x_k) złożony z elementów zbioru Ω .*

Twierdzenie 3.1. *Liczba k-elementowych prób uporządkowanych z powtórzeniami wybranych z n-elementowego zbioru Ω jest równa*

$$n^k.$$

Dowód wynika z własności iloczynu kartezjańskiego dla zbiorów skończonych.

Liczbę k-elementowych prób uporządkowanych z powtórzeniami wybranych ze zbioru n-elementowego nazywamy także liczbą k-elementowych wariacji z powtórzeniami ze zbioru n-elementowego i oznaczamy symbolem W_n^k .

Przykład 3.1. *Ile trzyliterowych słów możemy utworzyć z 26 liter alfabetu?*

Odp. $W_{26}^3 = 26^3$.

3.2 Wariacje bez powtórzeń

Definicja 3.2. *K-elementową próbą uporządkowaną bez powtórzeń z populacji Ω nazywamy dowolny k-elementowy ciąg (x_1, \dots, x_k) różnych elementów wybrany ze zbioru Ω .*

Twierdzenie 3.2. *Liczba k-elementowych prób uporządkowanych bez powtórzeń wybranych z n-elementowego zbioru Ω jest równa:*

1. 0, jeżeli $k > n$;
2. $\frac{n!}{(n-k)!}$, jeżeli $k \leq n$.

Dowód. Istnieje n możliwości wyboru pierwszego elementu, $n - 1$ możliwości wyboru drugiego elementu, itd., aż do $(n - k + 1)$ możliwości wyboru k -tego elementu. Stosując regułę iloczynu otrzymujemy żądany wzór.

Liczbę k -elementowych prób uporządkowanych bez powtórzeń wybranych ze zbioru n -elementowego nazywamy także liczbą k -elementowych wariacji bez powtórzeń ze zbioru n -elementowego i oznaczamy symbolem V_n^k .

Uwaga 3.1. Dla $k = n$ otrzymujemy permutację n -elementową:

$$P_n = n!.$$

Przykład 3.2. Ile jest pięciocyfrowych numerów telefonicznych złożonych z cyfr $\{1, 2, \dots, 9\}$, w których każda cyfra może wystąpić co najwyżej raz?

Odp. Liczba takich numerów wynosi $V_9^5 = \frac{9!}{4!}$.

3.3 Kombinacje bez powtórzeń

Definicja 3.3. K -elementową próbą nieuporządkowaną bez powtórzeń z populacji Ω nazywamy dowolny k -elementowy podzbiór (x_1, \dots, x_k) różnych elementów zbioru Ω .

Twierdzenie 3.3. Liczba k -elementowych prób nieuporządkowanych bez powtórzeń wybranych z n -elementowego zbioru Ω jest równa:

1. 0, jeżeli $k > n$;
2. $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, jeżeli $k \leq n$.

Dowód. Czynniki $k!$ odpowiada permutacji w obrębie wybranego podzbioru k -elementowego.

Przykład 3.3. Ile jest możliwości skreśleń w Dużym Lotku?

Odp. Liczba wszystkich możliwości to $\binom{49}{6} = 13983816$.

Liczbę k -elementowych prób nieuporządkowanych bez powtórzeń wybranych ze zbioru n -elementowego nazywamy także liczbą k -elementowych kombinacji bez powtórzeń ze zbioru n -elementowego i oznaczamy symbolem C_n^k .

3.4 Kombinacje z powtórzeniami

Definicja 3.4. K -elementową próbą nieuporządkowaną z powtórzeniami z populacji Ω nazywamy dowolny k -elementowy podzbiór (x_1, \dots, x_k) elementów zbioru Ω .

Twierdzenie 3.4. Liczba k -elementowych prób nieuporządkowanych bez powtórzeń wybranych z n -elementowego zbioru Ω jest równa:

$$\binom{n+k-1}{k}.$$

Dowód. Rozważmy $n+k-1$ miejsc. Zajmijmy $n-1$ różnych miejsc na $\binom{n+k-1}{n-1}$ sposobów. Istnieje wzajemnie jednoznaczne odwzorowanie pomiędzy wyborami miejsc a k -elementowymi próbami nieuporządkowanymi z powtórzeniami wybranymi ze zbioru n -elementowego zbioru Ω .

Przykład 3.4. Na ile sposobów możemy rozdzielić cztery jednakowe paczki pomiędzy trzy osoby?

Odp. $\binom{3+4-1}{4} = \binom{6}{4} = 15$. Liczbę k -elementowych prób nieuporządkowanych z powtórzeniami wybranych ze zbioru n -elementowego nazywamy także liczbą k -elementowych kombinacji z powtórzeniami ze zbioru n -elementowego i oznaczamy symbolem \bar{C}_n^k .

Przykład 3.5. W fizyce statystycznej rozważa się rozkład (rozmiesczenie) k cząstek w n elementarnych obszarach zwanych „komórkami”. W zależności od postaci tych cząstek przyjmuje się jedno z trzech następujących założeń:

1. cząstki różnią się między sobą (zatem wzajemna zamiana komórek przez

dwie cząstki daje nowy rozkład) i liczba cząstek w danej komórce jest dowolna (statystyka Maxwella – Boltzmannna),

2. cząstki są nierozróżnialne między sobą (zatem wzajemna zamiana komórek przez dwie cząstki daje ten sam rozkład – istotne jest tylko ile cząstek trafiło do poszczególnych komórek, a nie to jakie to są cząstki) i liczba cząstek w jednej komórce jest dowolna (statystyka Bosego – Einsteina),

3. cząstki nie różnią się między sobą i w każdej komórce może się znaleźć co najwyżej jedna cząstka (statystyka Fermiego – Diraca).

Zakładamy ponadto, że wszystkie dopuszczalne rozmieszczenia są jednakowo prawdopodobne. Na ile sposobów możemy rozmieścić k cząstek w n komórkach w każdym z tych przypadków?

3.5 Zadania

Zadanie 1. W urnie znajdują się kartki o numerach: 1, 2, 3, 4, 5. Losujemy 3 kartki

(a) ze zwracaniem

(b) bez zwracania.

Obliczyć prawdopodobieństwo, że wylosujemy 2 kartki o numerach nieparzystych.

Rozwiązanie. B – zdarzenie określone w zadaniu

(a)

$$N(\Omega) = 5^3 = 125$$

$$N(B) = \binom{3}{1} \binom{3}{1} \binom{2}{1} \cdot \frac{3!}{2!1!} = 3 \cdot 18 = 54$$

$$P(B) = \frac{54}{125} = 0,432$$

(b)

$$N(\Omega) = 5 \cdot 4 \cdot 3 = 60$$

$$N(B) = 2 \cdot 3 \cdot 2 \cdot 3 = 36$$

$$P(B) = \frac{36}{60} = 0,6$$

Zadanie 2. Chór składa się z 8 sopranów, 6 altów, 7 tenorów i 4 basów. Jakie jest prawdopodobieństwo, że:

(a) dwie losowo wybrane osoby mają ten sam rodzaj głosu;

(b) trzy losowo wybrane osoby są tej samej płci?

Rozwiązanie.

(a) A – zdarzenie polegające na tym, że dwie losowo wybrane osoby mają ten sam rodzaj głosu

$$N(\Omega) = \binom{25}{2} = 300$$

$$N(A) = \binom{8}{2} + \binom{6}{2} + \binom{7}{2} + \binom{4}{2} = 28 + 15 + 21 + 6 = 70$$

$$P(A) = \frac{N(A)}{N(\Omega)} = \frac{7}{30}.$$

(b) B – zdarzenie polegające na tym, że trzy losowo wybrane osoby są tej samej płci

$$N(\Omega) = \binom{25}{3} = 2300$$

$$N(B) = \binom{14}{3} + \binom{11}{3} = 529$$

$$P(B) = \frac{N(B)}{N(\Omega)} = \frac{529}{2300} = \frac{23}{100} = 0,23.$$

4 Przestrzenie probabilistyczne

4.1 Wprowadzenie

Teoria prawdopodobieństwa dostarcza modele matematyczne pozwalające opisywać eksperymenty, których wyników nie możemy w pełni przewidzieć. Oto

niektóre przykłady.

Eksperyment	Obserwowany rezultat
Rzut kostką	Liczba całkowita $k \in \{1, 2, 3, 4, 5, 6\}$
Obserwacja próbki n -produktów	Liczba produktów uszkodzonych
Kwestionariusz 100 pytań binarnych	Ciąg 100 odpowiedzi $\omega \in \{tak, nie\}^{100}$
Strzał z łuku	Punkt na tarczy

Pomimo, że precyzyjnych wyników żadnego z tych eksperymentów nie można przewidzieć, ich obserwacja i intuicja utwierdzają nas w przekonaniu, że te zjawiska podlegają pewnym prawom. Na przykład, jeżeli wykonamy 60000 rzutów kostką oczekujemy, że liczba wyników, w których pojawi się szóstka wyniesie około 10000.

Teoria prawdopodobieństwa pozwala nadać precyzyjny sens tym nieco "niejasnym" rozważaniom.

4.2 Zdarzenie

Aby modelować doświadczenia losowe nowoczesna teoria prawdopodobieństwa używa języka zbiorów.

Oznaczmy przez Ω zbiór, którego elementy reprezentują wszystkie możliwe wyniki danego doświadczenia losowego. Zbiór ten nazywamy przestrzenią zdarzeń elementarnych a wyniki zdarzeniami elementarnymi. Natomiast zdarzenia będą reprezentowane przez podzbiory zbioru Ω . Na przykład rozważając rzut kostką zdarzenie A —otrzymamy cyfrę parzystą, nie jest elementarne. Jest ono złożone z trzech zdarzeń elementarnych 2, 4, 6 : $A = \{2, 4, 6\}$. W tym przypadku $\Omega = \{1, 2, 3, 4, 5, 6\}$. Poniżej przedstawimy podstawowe pojęcia z teorii zbiorów wykorzystywane w probabilistyce wraz z z odpowiednimi oznaczeniami i nazewnictwem charakterystycznym dla tej teorii.

Oznaczenie	Nazwa
Ω	Zdarzenie pewne
\emptyset	Zdarzenie niemożliwe
ω	Zdarzenie elementarne
A	Zdarzenie
$\omega \in A$	ω należy do zbioru realizacji A
$A \subset B$	A implikuje B
$A \cap B$	A i B
$A \cup B$	A lub B
A^c	zdarzenie przeciwne do A

Oczywiście operacje logiczne mogą dotyczyć więcej niż dwóch zbiorów. Dla przykładu

$$\bigcup_{i=1}^n A_i = A_1 \cup A_2 \dots \cup A_n$$

oznacza realizację przynajmniej jednego ze zdarzeń A_i dla $i = 1, 2, \dots, n$, natomiast

$$\bigcap_{i=1}^n A_i = A_1 \cap A_2 \dots \cap A_n$$

oznacza realizację wszystkich zdarzeń A_i dla $i = 1, 2, \dots, n$. Analogicznie wprowadzamy sumę i iloczyn niekończonej ilości zbiorów.

4.3 σ -ciało

W przypadku, kiedy przestrzeń Ω jest skończona, zdefiniowaliśmy zdarzenie jako podzbiór Ω i to jest intuicyjnie jasne. Ale gdy chcemy na przykład modelować eksperyment polegający na nieskończenie wielu rzutach monetą naturalnym jest przyjąć przestrzeń zdarzeń elementarnych jako $\Omega = \{O, R\}^{\mathbb{N}}$ a więc przestrzeń, która nie jest przeliczalna ponieważ nie istnieje surjekcja ze zbioru liczb naturalnych na Ω . Gdy przestrzeń Ω jest nieprzeliczalna aby skonstruować przestrzeń probabilistyczną konieczne jest zdefiniowanie pewnej

podklasy $\mathcal{P}(\Omega)$ zwanej σ -ciałem. Niezbędne jest aby rodzina ta była zamknięta ze względu na zdefiniowane wcześniej operacje teoriomnogościowe.

Definicja 4.1. σ -ciałem \mathcal{A} zbioru Ω nazywamy rodzinę podzbiorów Ω spełniającą następujące warunki:

- 1) $\Omega \in \mathcal{A}$.
- 2) Jeżeli $A \in \mathcal{A}$ to $A^c \in \mathcal{A}$.
- 3) Jeżeli $A_1, A_2, \dots, \in \mathcal{A}$ to $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Zauważmy, że z poniższej definicji wynika wiele ciekawych własności dotyczących σ -ciała zbiorów:

- $\emptyset \in \mathcal{A}$.
- Jeżeli $A \in \mathcal{A}$ oraz $B \in \mathcal{A}$, to $A \cap B \in \mathcal{A}$.
- Jeżeli $A \in \mathcal{A}$ oraz $B \in \mathcal{A}$, to $A \setminus B \in \mathcal{A}$.
- Jeżeli $A \in \mathcal{A}$ oraz $B \in \mathcal{A}$, to $A \triangle B \in \mathcal{A}$.

Dowód. (ĆWICZENIE).

Warto także zanotować, że powyższa definicja absolutnie nie wymaga aby w σ -ciele znajdowały się wszystkie podzbiory zbioru Ω . Zauważmy na przykład, że:

zbiór $\{\emptyset, \Omega\}$ jest σ -ciałem (jest to najmniejsze σ -ciało);

zbiór $\{\emptyset, A, A^c, \Omega\}$ dla dowolnego zdarzenia A jest σ -ciałem (jest to najmniejsze σ -ciało generowane przez A).

4.4 Miara probabilistyczna

Mając zdefiniowaną przestrzeń (Ω, \mathcal{A}) złożoną ze zbioru zdarzeń elementarnych oraz σ -ciała możemy zdefiniować miarę probabilistyczną.

Definicja 4.2. *Miarą probabilistyczną na przestrzeni (Ω, \mathcal{A}) nazywamy funkcję $P : \mathcal{A} \rightarrow [0, 1]$ taką, że:*

1. $P(\Omega) = 1$;
2. $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, jeżeli $A_i \cap A_j = \emptyset$ dla dowolnych wskaźników i, j takich, że $i \neq j$. (własność σ -addytywności).

Przestrzeń (Ω, \mathcal{A}, P) nazywamy przestrzenią probabilistyczną.

Krótko mówiąc definicja prawdopodobieństwa na przestrzeni (Ω, \mathcal{A}) jest niczym innym jak przyporządkowaniem każdemu zdarzeniu pewnej „masy” przy założeniu, że masa całego zbioru Ω wynosi 1.

Przykład 4.1. *Bolek i Lolek rzucają kolejno kostką sześcienną do gry. Wygrywa ten, który pierwszy uzyska szóstkę. Bolek zaczyna grę. Interesują nas trzy zdarzenia:*

$$A = \{\text{zwycięstwo Bolka}\}.$$

$$B = \{\text{zwycięstwo Lolka}\}.$$

$$C = \{\text{nie ma zwycięzcy}\}.$$

Zauważmy, że ze względu na kolejność rzutów Bolek może zostać zwycięzcą po nieparzystej ilości rzutów, natomiast Lolek po parzystej ilości rzutów. Dlatego też możemy zapisać, że

$$A = \bigcup_{k=0}^{\infty} F_{2k+1},$$

$$B = \bigcup_{k=1}^{\infty} F_{2k},$$

gdzie zdarzenie F_n oznacza, że gra zostanie zakończona po n -tym rzucie. Ponadto

$$P(F_n) = (5/6)^{n-1}(1/6).$$

Zanotujmy także, że jeżeli $i \neq j$, to $F_i \cap F_j = \emptyset$, ponieważ gra nie może zostać zakończona jednocześnie w innym czasie. Wykorzystując własność

σ -addytywności otrzymujemy:

$$P(A) = \sum_{k=0}^{\infty} P(F_{2k+1}) = \sum_{k=0}^{\infty} \frac{1}{6} \left(\frac{5}{6}\right)^{2k} = \frac{1}{6} \frac{1}{1 - \frac{25}{36}} = \frac{6}{11},$$

$$P(B) = \sum_{k=1}^{\infty} P(F_{2k}) = \sum_{k=1}^{\infty} \frac{1}{6} \left(\frac{5}{6}\right)^{2k-1} = \frac{5}{36} \frac{1}{1 - \frac{25}{36}} = \frac{5}{11}.$$

Na podstawie powyższych wyników stwierdzamy, że Bolek ma minimalnie większe szanse na wygraną, co jest zgodne z intuicją, ponieważ rozpoczyna grę. Ponadto z własności σ -addytywności (A, B, C są rozłączne) wnioskujemy, że $P(C) = 0$, zatem prawdopodobieństwo, że gra nie przyniesie zwycięzcy jest równe zero, co także nie jest zaskoczeniem. Stwierdzamy ponadto, że w naszym modelu zbiór $\{1, 2, 3, 4, 5\}^{\mathbb{N}}$, odpowiadający zdarzeniu niewyłonienia zwycięzcy jest nieprzeliczalny.

Poniżej podamy w formie twierdzenia podstawowe własności miary probabilistycznej.

Twierdzenie 4.1. *{Podstawowe własności prawdopodobieństwa.}*

Miara prawdopodobieństwa określona w definicji 4.2 na przestrzeni (Ω, \mathcal{A}) spełnia następujące własności:

1. $P(\emptyset) = 0$.
2. $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$, jeżeli $A_i \cap A_j = \emptyset$ dla dowolnych wskaźników i, j takich, że $1 \leq i, j \leq n$ oraz $i \neq j$. (własność addytywności).
3. $\forall A \in \mathcal{A}, P(A^c) = 1 - P(A)$.
4. $\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, A \subseteq B, P(A) \leq P(B)$.
5. $\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
6. (Lemat o ciągłości)
 - a) Jeżeli $(B_n)_{n \geq 1}$ jest rosnącym ciągiem zdarzeń zbieżnym do zdarzenia

B , to $P(B) = \lim_{n \rightarrow \infty} P(B_n)$.

b) Jeżeli $(C_n)_{n \geq 1}$ jest malejącym ciągiem zdarzeń zbieżnym do zdarzenia C , to $P(C) = \lim_{n \rightarrow \infty} P(C_n)$.

7. a) $\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, P(A \cup B) \leq P(A) + P(B)$.

b) $\forall A_1, \dots, A_n \in \mathcal{A}, P(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$.

c) $\forall A_1, \dots, A_n, \dots \in \mathcal{A}, P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$.

8. Zasada włączeń i wyłączeń

Dla dowolnych zdarzeń A_1, \dots, A_n oraz $n \geq 2$ prawdziwy jest wzór

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}).$$

Dowód. Niech P będzie miarą prawdopodobieństwa spełniającą warunki definicji 4.2. Pokażemy, że P spełnia warunki 1 – 8.

Ad 1. $1 = P(\Omega) = P(\emptyset \cup \Omega) = P(\Omega) + P(\emptyset) = 1 + P(\emptyset) \Rightarrow P(\emptyset) = 0$.

Ad 2. Niech A_1, \dots, A_n będą parami rozłączne. Połóżmy $A_j = \emptyset$ dla $j > n$.

Wtedy

$$P\left(\bigcup_{j=1}^n A_j\right) = P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^n P(A_j) + \sum_{j=n+1}^{\infty} P(A_j).$$

Druga suma po prawej stronie jest równa zero co dowodzi 2.

Ad 3. $1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$ co dowodzi (3).

Ad 4. Jeżeli $A \subset B$ to $B = A \cup (B \cap A^c)$. Stosując 2 otrzymujemy tezę.

Ad 5. $A \cup B = (A \cap B^c) \cup (A \cap B) \cup (B \cap A^c)$. Z addytywności miary mamy $P(A \cup B) = P((A \cap B^c) \cup (A \cap B) \cup (B \cap A^c)) = P(A \cap B^c) + P(A \cap B) + P(B \cap A^c) = (P(A \cap B^c) + P(A \cap B)) + (P(B \cap A^c) + P(A \cap B)) - P(A \cap B) = P(A) + P(B) - P(A \cap B)$.

Ad 6. Wystarczy udowodnić pierwszą część, ponieważ druga wynika z 6(a) po podstawieniu $B_n = C_n^c$. Zauważmy, że dla dowolnego n zdarzenie B_n może

zostać zapisane jako suma zdarzeń rozłącznych następująco:

$$B_n = B_0 \cup \left(\bigcup_{i=1}^n (B_i \setminus B_{i-1}) \right).$$

Korzystając z σ -addytywności miary otrzymujemy kolejno

$$P(B_n) = P(B_0) + \sum_{i=1}^n P(B_i \setminus B_{i-1}).$$

$$P(B) = P(B_0) + \sum_{i=1}^{\infty} P(B_i \setminus B_{i-1}).$$

Stąd

$$P(B) = \lim_{n \rightarrow \infty} \left(P(B_0) + \sum_{i=1}^n P(B_i \setminus B_{i-1}) \right) = \lim_{n \rightarrow \infty} P(B_n).$$

Ad 7. (ĆWICZENIE)

Ad 8. Udowodnimy ten wzór indukcyjnie. Dla $n = 2$ redukuje się on do udowodnionego już wzoru

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

Załóżmy, że wzór jest prawdziwy dla dowolnego n i spróbujmy pokazać, że zachodzi dla $n + 1$ zbiorów.

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) - P\left(\bigcup_{i=1}^n (A_i \cap A_{n+1})\right).$$

Dla pierwszej i trzeciej części prawej strony powyższego wzoru skorzystamy teraz z założenia o prawdziwości wzoru dla n zbiorów. Otrzymujemy:

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}) + \\ &P(A_{n+1}) - \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k} \cap A_{n+1}) = \\ &\sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}) + \\ &P(A_{n+1}) + \sum_{k=1}^n (-1)^{k+2} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k} \cap A_{n+1}) = \\ &\sum_{k=1}^{n+1} (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n+1} P(A_{i_1} \cap \dots \cap A_{i_k}). \end{aligned}$$

Wzór został tym samym udowodniony.

4.5 Różne definicje prawdopodobieństwa

Definicja miary prawdopodobieństwa podana przez Kołmogorowa jest układem aksjomatów, który nie wyznacza konkretnych wartości liczbowych funkcji zbioru P . Podstawowym etapem w matematycznym opisie danego zjawiska jest stworzenie przestrzeni probabilistycznej (Ω, \mathcal{A}, P) . Poniżej podamy trzy różne definicje prawdopodobieństwa bazujące na różnych przestrzeniach Ω .

Definicja 4.3. *{Klasyczna definicja prawdopodobieństwa}*

Załóżmy, że przestrzeń Ω składa się z n jednakowo prawdopodobnych zdarzeń elementarnych. Wtedy prawdopodobieństwo zajścia zdarzenia A jest równe ilorazowi liczby zdarzeń sprzyjających temu zdarzeniu do liczby wszystkich możliwych zdarzeń, tzn.:

$$P(A) = \frac{N(A)}{N(\Omega)},$$

gdzie symbol $N(\cdot)$ oznacza liczbę elementów danego zbioru.

Jak łatwo zauważyć, stosowanie powyższej definicji ogranicza się jedynie do skończonych przestrzeni Ω . Dlatego często istnieje potrzeba rozszerzenia tej definicji na przypadki, kiedy zbiór zdarzeń elementarnych jest nieskończony. Przykładem takiego rozszerzenia jest definicja geometryczna.

Definicja 4.4. *(Geometryczna definicja prawdopodobieństwa)*

Niech \mathcal{A} będzie σ -ciałem określonym na przestrzeni Ω . Każdemu zdarzeniu A z \mathcal{A} przypisujemy w sposób jednoznaczny miarę $\mu(A)$, spełniającą warunki:

1. $\mu(A) \geq 0$,
2. $\mu(\emptyset) = 0$;
3. $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$,

jeżeli

$$A_i \cap A_j = \emptyset, \quad \text{dla } i \neq j.$$

Prawdopodobieństwo geometryczne zdarzenia A ($P(A)$) definiujemy jako iloraz miary A do miary zbioru Ω , tzn.:

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

Przykład 4.2. Robotnik obsługuje dwie maszyny. Długotrwałe obserwacje wykazały, że każdej z maszyn poświęca 8 minut w ciągu godziny. Znaleźć prawdopodobieństwo, że w ciągu godziny maszyna wymaga interwencji robotnika wtedy i tylko wtedy, gdy jest zajęty drugą maszyną.

$$\Omega = \left\{ \begin{array}{l} (x, y) : \\ 0 \leq x \leq 60 \\ 0 \leq y \leq 60 \end{array} \right. , \quad A = \{(x, y) : |x - y| < 8\}.$$

Stąd

$$P(A) = \frac{60^2 - 52 \cdot 52}{60^2} = \frac{225 - 169}{225} = \frac{56}{225}.$$

Przykład 4.3. Rzucamy monetę o średnicy 2 cm na podłogę pokrytą kwadratowymi płytkami o boku 10 cm. Jakie jest prawdopodobieństwo, że moneta w całości leży na jednej płytce?

Przykład 4.4. Paradoks Bertranda.

Aby obliczyć prawdopodobieństwo geometryczne, musimy znać miary zbiorów, co nie zawsze jest możliwe. Dlatego statystycy podali propozycję innej definicji.

Definicja 4.5. {Statystyczna definicja prawdopodobieństwa}

Załóżmy, że w długiej serii doświadczeń obserwujemy pojawienie się zdarzenia A . Jeśli częstość $\frac{m(n)}{n}$ pojawienia się zdarzenia A , gdzie n jest długością serii, a $m(n)$ liczbą doświadczeń, w których pojawiło się zdarzenie A , przy wzrastaniu długości serii zbliża się do pewnej liczby p oscylując wokół tej liczby i jeśli wahania częstości zdarzenia A przejawiają tendencję malejącą przy wzrastającym n , to liczba p nazywa się prawdopodobieństwem zdarzenia A .

4.6 Zadania

Zadanie 3. Niech A i B należą do pewnego σ -ciała \mathcal{A} zdarzeń. Wykazać, że \mathcal{A} zawiera zdarzenia: $A \cap B$, $A \setminus B$ i $A \Delta B$.

Rozwiązanie.

1. $A \cap B = \overline{(\overline{A \cup B})} \in \mathcal{A}$.
2. $A \setminus B = A \cap \overline{B} = \overline{(\overline{A \cup B})} \in \mathcal{A}$.
3. $A \Delta B = (A \setminus B) \cup (B \setminus A) = \overline{(\overline{A \cup B})} \cup \overline{(A \cup \overline{B})} \in \mathcal{A}$.

Zadanie 4. Załóżmy, że $P(A) = 0,4$, $P(\overline{B}) = 0,6$, $P(\overline{A \cap B}) = 0,1$. Znaleźć $P(A \cup \overline{B})$.

Rozwiązanie. $P(A \cup \overline{B}) = P(A) + P(\overline{B}) - P(A \cap \overline{B})$. Dalej

$$B = B \cap A \cup B \cap \overline{A}$$

$$P(B) = P(A \cap B) + P(\overline{A} \cap B)$$

$$0,4 = P(A \cap B) + 0,1 \implies P(A \cap B) = 0,3$$

$$A = A \cap B \cup A \cap \overline{B}$$

$$P(A) = P(A \cap B) + P(A \cap \overline{B})$$

$$0,4 = 0,3 + P(A \cap \overline{B}) \implies P(A \cap \overline{B}) = 0,1.$$

Stąd

$$P(A \cup \overline{B}) = 0,4 + 0,6 - 0,1 = 0,9.$$

Zadanie 5. Niech $\Omega = \mathbf{R}_+$ a \mathcal{A} będzie σ -ciałem podzbiorów Ω zawierającym wszystkie przedziały. Załóżmy, że funkcja (zbioru) P określona na \mathcal{A} spełnia aksjomat przeliczalnej addytywności. Ponadto zakładamy, że dla przedziału $I \in \mathcal{A}$ mamy $P(I) > 0$ oraz dla

$$I_k = (k, k+1), \quad P(I_k) = \frac{1}{k(k+1)}, \quad k \geq 1.$$

Wykazać, że P nie jest prawdopodobieństwem.

Na mocy założenia $P(I) > 0$ dla $I = (0, 1)$. Stąd

$$P(\Omega) > P\left(\bigcup_{k=1}^{\infty} I_k\right) = \sum_{k=1}^{\infty} P(I_k) = \sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1.$$

5 Prawdopodobieństwo warunkowe

5.1 Definicja i własności

Pytanie, na jakie będziemy się obecnie starali odpowiedzieć, jest następujące: w jaki sposób dodatkowa informacja wpłynie na wartość prawdopodobieństwa? Okazuje się, że odpowiedź możemy uzyskać rozważając tzw. prawdopodobieństwo warunkowe. Na przykład rozważmy sytuację, kiedy w pewnej grupie liczącej N osób jest N_c chłopców i N_d dziewcząt. Ponadto wiemy, że grupa ta zawiera N_l osób leworęcznych. Przyjmijmy następujące oznaczenia:

$$L = \{\text{przypadkowo wybrana osoba jest leworęczna}\}.$$

$$C = \{\text{przypadkowo wybrana osoba jest chłopcem}\}.$$

Przez N_{cl} oznaczmy liczbę chłopców leworęcznych.

Wtedy oczywiście $P(C) = \frac{N_c}{N}$ a $P(C \cap L) = \frac{N_{cl}}{N}$. Możemy sobie jednak zadać pytanie jakie jest prawdopodobieństwo tego, że wybrany losowo chłopiec jest leworęczny? Zauważmy, że dysponujemy tutaj dodatkową informacją: wiemy, że wybraną osobą jest chłopiec. Zatem szukając leworęcznych chłopców wśród wszystkich chłopców otrzymujemy wartość prawdopodobieństwa równą $\frac{N_{cl}}{N_c}$. Ale zauważmy, że wartość ta może być także otrzymana za pomocą $P(C)$ i $P(C \cap L)$. rzeczywiście

$$\frac{N_{cl}}{N_c} = \frac{NP(C \cap L)}{NP(C)} = \frac{P(C \cap L)}{P(C)}.$$

W ten sposób dochodzimy do zdefiniowania pojęcia prawdopodobieństwa warunkowego.

Definicja 5.1. Niech (Ω, \mathcal{A}, P) będzie przestrzenią probabilistyczną a $B \subset \mathcal{A}$ zdarzeniem takim, że $P(B) \neq 0$. Dla dowolnego zdarzenia A prawdopodobieństwo warunkowe A pod warunkiem zdarzenia B definiujemy następująco:

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

Okazuje się, co wykażemy poniżej, że zdefiniowaliśmy tym samym nową miarę probabilistyczną.

Stwierdzenie 1. Niech (Ω, \mathcal{A}, P) będzie przestrzenią probabilistyczną, a $B \subset \mathcal{A}$, ustalonym zdarzeniem takim, że $P(B) \neq 0$. Wtedy funkcja zbioru $P(\cdot|B)$ zdefiniowana wzorem:

$$P(\cdot|B) : \mathcal{A} \rightarrow [0, 1], \quad A \rightarrow P(A|B)$$

jest miarą prawdopodobieństwa zdefiniowaną na (Ω, \mathcal{A}) .

Dowód tego faktu pozostawimy jako ćwiczenie. Wystarczy sprawdzić, że są spełnione aksjomaty definicji prawdopodobieństwa.

Podana własność uzasadnia wprowadzenie nowego oznaczenia, często wykorzystywanego w literaturze a mianowicie

$$P(A|B) := P_B(A).$$

Oczywiście, podobnie jak w przypadku miary zdefiniowanej w definicji 4.2, dla miary warunkowej $P_B(\cdot)$ możemy podać własności analogiczne do zaprezentowanych wcześniej.

Wniosek 5.1. Funkcja zbioru $P(\cdot|B)$ spełnia następujące własności:

1. $P(\emptyset|B) = 0$.
2. $P(\bigcup_{i=1}^n A_i|B) = \sum_{i=1}^n P(A_i|B)$, jeżeli $A_i \cap A_j = \emptyset$ dla dowolnych wskaźników i, j takich, że $1 \leq i, j \leq n$ oraz $i \neq j$. (własność addytywności).

3. $\forall A \in \mathcal{A}, P(A^c|B) = 1 - P(A|B)$.
4. $\forall A \in \mathcal{A}, \forall C \in \mathcal{A}, A \subseteq B, P(A|B) \leq P(C|B)$.
5. $\forall A \in \mathcal{A}, \forall C \in \mathcal{A}, P(A \cup C|B) = P(A|B) + P(C|B) - P(A \cap C|B)$.
6. (*Lemat o ciągłości*)
 - a) Jeżeli $(A_n)_{n \geq 1}$ jest rosnącym ciągiem zdarzeń zbieżnym do zdarzenia A , to $P(A|B) = \lim_{n \rightarrow \infty} P(A_n|B)$.
 - b) Jeżeli $(C_n)_{n \geq 1}$ jest malejącym ciągiem zdarzeń zbieżnym do zdarzenia C , to $P(C|B) = \lim_{n \rightarrow \infty} P(C_n|B)$.
7. a) $\forall A_1, \dots, A_n \in \mathcal{A}, P(\bigcup_{i=1}^n A_i|B) \leq \sum_{i=1}^n P(A_i|B)$.
 b) $\forall A_1, \dots, A_n \dots \in \mathcal{A}, P(\bigcup_{i=1}^{\infty} A_i|B) \leq \sum_{i=1}^{\infty} P(A_i|B)$.

Zauważmy, że powyższe wyniki nie zawierają żadnego wzoru pozwalającego wyrazić wartość prawdopodobieństwa iloczynu zdarzeń za pomocą prawdopodobieństw tych zdarzeń. W ogólnym przypadku taki wzór nie istnieje. Natomiast przy pomocy prawdopodobieństw warunkowych wzór taki istnieje i nazywany jest często wzorem łańcuchowym.

Twierdzenie 5.1. *Jeżeli A_1, \dots, A_n są zdarzeniami losowymi takimi, że $P(A_1 \cap \dots \cap A_{n-1}) \neq 0$, to*

$$P(A_1 \cap \dots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \times \dots \times P(A_n|A_1 \cap \dots \cap A_{n-1}).$$

Dowód. Dla $1 \leq i \leq n - 1$ mamy $\bigcap_{j=1}^{n-1} A_j \subset \bigcap_{j=1}^i A_j$, skąd

$$0 < P\left(\bigcap_{j=1}^{n-1} A_j\right) \leq P\left(\bigcap_{j=1}^i A_j\right).$$

Zatem $P(\bigcap_{j=1}^i A_j) \neq 0$ dla każdego $i \leq n - 1$ i możemy „warunkować” po $\bigcap_{j=1}^i A_j$. To uzasadnia następujące równości:

$$P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \times \dots \times P(A_n|A_1 \cap \dots \cap A_{n-1})$$

$$\begin{aligned}
&= P(A_1) \times \frac{P(A_1 \cap A_2)}{P(A_1)} \times \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \times \dots \times \frac{P(A_1 \cap \dots \cap A_n)}{P(A_1 \cap \dots \cap A_{n-1})} \\
&= P(A_1 \cap \dots \cap A_n).
\end{aligned}$$

5.2 Prawdopodobieństwo całkowite i wzór Bayesa

Następny wzór, zwany wzorem na prawdopodobieństwo całkowite, pozwala wyrazić prawdopodobieństwo dowolnego zdarzenia przy pomocy prawdopodobieństw warunkowych. Zanim go podamy, przytoczymy definicję rodziny zupełnej zdarzeń.

Założmy, że $I = \{1, 2, \dots, n\}$.

Definicja 5.2. *Mówimy, że rodzina zdarzeń $(H_i)_{i \in I}$ tworzy rodzinę zupełną na przestrzeni Ω (tworzy podział zbioru Ω), jeżeli są spełnione następujące warunki:*

1. $\forall i \in I, H_i \neq \emptyset$.
2. $\Omega = \bigcup_{i \in I} H_i$.
3. $H_i \cap H_j = \emptyset$ dla $i \neq j, 1 \leq i, j \leq n$.

Twierdzenie 5.2. Wzór na prawdopodobieństwo całkowite

Niech (Ω, \mathcal{A}, P) będzie przestrzenią probabilistyczną. Jeżeli $P(B) \neq 0$ i $P(B^c) \neq 0$, to

1. $\forall A \in \mathcal{A}, P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$.
2. Jeżeli $(H_i)_{i \in I}$ tworzy zupełną rodzinę zdarzeń, to

$$\forall A \in \mathcal{A}, P(A) = \sum_{i=1}^n P(A|H_i)P(H_i).$$

Dowód. Udowodnimy wzór (2).

Ponieważ $(H_i)_{i \in I}$ tworzy rodzinę zupełną zdarzeń, otrzymujemy

$$A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^n H_i \right) = \bigcup_{i=1}^n (A \cap H_i).$$

Stąd

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

Wzór został zatem udowodniony.

Przykład 5.1. *Na wejściu pewnego układu pojawiają się sygnały: "0" z prawdopodobieństwem 0.1 i "1" z prawdopodobieństwem 0.9. Układ przekazuje daną wejściową bez zakłóceń z prawdopodobieństwem 0.7 jeśli nadano sygnał "1" i z prawdopodobieństwem 0.9 jeśli nadano sygnał "0". Pojawienie się zakłóceń zamienia "0" na "1" i odwrotnie. Jak często na wyjściu pojawia się sygnał "1", a jak często "0"?*

Rozwiązanie.

Niech

- zdarzenie B – na wyjściu sygnał 0,
- zdarzenie C – na wyjściu sygnał 1,
- zdarzenie A_0 – nadano sygnał 0,
- zdarzenie A_1 – nadano sygnał 1.

Wtedy mamy

$$P(B) = P(B|A_0)P(A_0) + P(B|A_1)P(A_1) = 0.9 \cdot 0.1 + 0.3 \cdot 0.9 = 0.36.$$

Przykład 5.2. (Problem ruiny gracza.)

Rozważmy gracza, który startuje z kapitałem początkowym i zł. W każdej grze może z prawdopodobieństwem p wygrać 1 lub z prawdopodobieństwem $q = 1 - p$ stracić 1zł. Znaleźć prawdopodobieństwo tego, że gracz osiągnie kapitał N .

Wzór na prawdopodobieństwo całkowite pozwala obliczyć prawdopodobieństwo dowolnego zdarzenia przy pomocy prawdopodobieństw warunkowych pod warunkiem zdarzeń $(H_i)_{i \in I}$. Problem można postawić trochę inaczej. Jak wyrazić kolejne prawdopodobieństwa podzbiorów H_i przy warunku zdarzenia A . Odpowiedź podaje tzw. wzór Bayesa.

Twierdzenie 5.3. Wzór Bayesa

Niech (Ω, \mathcal{A}, P) będzie przestrzenią probabilistyczną a A dowolnym zdarzeniem takim, że $P(A) \neq 0$. Jeżeli $(H_i)_{i \in I}$ tworzy zupełną rodzinę zdarzeń, to dla $i = 1, 2, \dots, n$

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

Dowód. Z definicji prawdopodobieństwa warunkowego oraz wzoru na prawdopodobieństwo całkowite otrzymujemy:

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

Ciąg dalszy poprzedniego przykładu:

-na wyjściu z naszego urzędu odebraliśmy sygnał "1". Jakie jest prawdopodobieństwo, że nadano faktycznie sygnał "1"?

-na wyjściu z urzędu odebraliśmy sygnał "0". Jakie jest prawdopodobieństwo, że faktycznie nadano sygnał "0"?

Rozwiązanie.

$$P(A_0|B) = \frac{P(B|A_0)P(A_0)}{P(B|A_0)P(A_0) + P(B|A_1)P(A_1)} = \frac{0.9 \cdot 0.1}{0.9 \cdot 0.1 + 0.3 \cdot 0.9} = \frac{1}{4}.$$

Przykład 5.3. Test na rzadką chorobę, która dotyka średnio 1 osobę na 1000, daje tzw. fałszywą pozytywną odpowiedź u 4% zdrowych, przy czym u chorych wynik pozytywny występuje zawsze. Jaka jest szansa, że osoba, u której test dał pozytywną odpowiedź, jest rzeczywiście chora? Zakładamy, że u chorej osoby nie występują jakiegokolwiek objawy choroby.

Rozwiązanie.

Niech

-zdarzenie B – pozytywna odpowiedź testu,

-zdarzenie A_1 – osoba chora,

-zdarzenie A_2 – osoba zdrowa.

Ze wzoru Bayesa mamy:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} =$$
$$\frac{1 \cdot \frac{1}{1000}}{1 \cdot \frac{1}{1000} + \frac{1}{25} \cdot \frac{999}{1000}} = 0,244.$$

Wynik może wydawać się dziwny i sprzeczny z intuicją – szansa, że osoba, u której test wykazał odpowiedź pozytywną jest rzeczywiście chora wynosi 2,44%.

Przykład 5.4. (Doświadczenie Poly'a).

Zacniemy od pewnego doświadczenia, które opiszemy metodami teorii prawdopodobieństwa. Pozwoli to nam doświadczenie to nazywać zjawiskiem losowym. Wyobraźmy sobie, że mamy urnę zawierającą n kul białych i m kul czarnych, a więc łącznie $s = n+m$ kul. Niech $k \geq 1$ i $r \in \mathbb{Z}$. Losujemy z urny kulę. Po odczytaniu jej koloru wrzucamy do urny kulę dodając r kul tego samego koloru, co wylosowana kula. Powtarzamy tę czynność k -razy. Wyznaczyć prawdopodobieństwo, że w k losowaniach zgodnie ze schematem Poly'a otrzymano j kul białych.

Przykład 5.5. Urna zawiera k kul białych i $n-k$ kul czarnych. Losujemy kolejno po 1 kuli bez zwracania z urny. Jakie jest prawdopodobieństwo wylosowania za i -tym razem kuli białej? (Wniosek).

5.3 Zadania

Zadanie 6. Niech $P(A) = p$, $P(B) = 1 - \varepsilon$, gdzie ε jest dowolnie małą liczbą. Oszacować $P(A|B)$ z góry i z dołu.

Rozwiązanie.

$$\begin{aligned} P(A|B) &= \frac{P(AB)}{P(B)} = \frac{1 - P(\overline{AB})}{P(B)} = \frac{1 - P(\overline{A}) - P(\overline{B}) + P(\overline{A}\overline{B})}{P(B)} = \\ &= 1 - \frac{P(\overline{A})}{P(B)} + \frac{P(\overline{A}\overline{B})}{P(B)} \geq 1 - \frac{1-p}{1-\varepsilon} = \frac{p-\varepsilon}{1-\varepsilon}. \end{aligned}$$

Z drugiej strony

$$P(A|B) = \frac{P(A) - P(\overline{B}) + P(\overline{A}\overline{B})}{P(B)} \leq \frac{P(A)}{P(B)} = \frac{p}{1-\varepsilon}.$$

Zadanie 7. Dane są $P(A)$, $P(B)$, $P(C)$, $P(AB)$, $P(AC)$, $P(BC)$, $P(ABC)$.

Znaleźć

$$P(C|\overline{A}\overline{B}).$$

Rozwiązanie.

$$\begin{aligned} P(C|\overline{A}\overline{B}) &= 1 - P(\overline{C}|\overline{A}\overline{B}) = 1 - \frac{P(\overline{A}\overline{B}\overline{C})}{P(\overline{A}\overline{B})} = 1 - \frac{P(\overline{A}\overline{B}\overline{C})}{P(\overline{A}\overline{B})} \\ &= 1 - \frac{P(\overline{A}\overline{B}\overline{C})}{1 - P(A \cup B)} = 1 - \frac{P(\overline{A}\overline{B}\overline{C})}{1 - P(A) - P(B) + P(AB)} \\ &= 1 - \frac{P(\overline{A}\overline{B}\overline{C})}{1 - P(A) - P(B) + P(AB)} \\ &= 1 - \frac{1 - (P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC))}{1 - P(A) - P(B) + P(AB)}. \end{aligned}$$

Zadanie 8. Urna zawiera n kul, spośród których część jest białych, a pozostałe czarne. Do urny tej dokładamy jedną kulę białą, a następnie losujemy jedną kulę. Jakie jest prawdopodobieństwo wylosowania kuli białej, gdy założymy, że każda liczba kul białych w początkowym zestawie jest jednakowo możliwa?

Rozwiązanie. Niech A_k oznacza zdarzenie, że w urnie znajdowało się k kul białych przed dodaniem $n + 1$ kuli białej. Wtedy prawdopodobieństwo

wylosowania kuli białej (zdarzenie B)

$$P(B) = P(A_0)P(B|A_0) + P(A_1)P(B|A_1) + \dots + P(A_n)P(B|A_n)$$

$$\text{Ale } P(A_0) = P(A_1) = \dots = P(A_n) = \frac{1}{n+1},$$

$$P(B|A_k) = \frac{k+1}{n+1}, \quad k = 0, 1, \dots, n.$$

Zatem

$$P(B) = \frac{n+2}{2n+2}.$$

6 Niezależność

6.1 Niezależność dwóch zdarzeń

Niech A i B będą dwoma zdarzeniami o niezerowym prawdopodobieństwie. Zdarzyć się może, że wiedza o zdarzeniu A nie zmienia prawdopodobieństwa zdarzenia B , inaczej mówiąc $P(B|A) = P(B)$. Dzieje się tak na przykład wtedy, gdy losujemy kartę z talii ze zwracaniem i zdarzenie A zależy tylko od wyniku pierwszego losowania natomiast zdarzenie B tylko od wyniku drugiego losowania. Oczywiście w tym przypadku zachodzi symetria tzn.

$P(A|B) = P(A)$. Powyższe rozważania możemy uogólnić i zapisać w formie następującego twierdzenia.

Twierdzenie 6.1. *Niech (Ω, \mathcal{A}, P) będzie przestrzenią probabilistyczną a A i B dowolnymi zdarzeniami takimi, że $P(A) \neq 0$ i $P(B) \neq 0$. Wtedy następujące równości są równoważne:*

1. $P(B|A) = P(B)$.
2. $P(A|B) = P(A)$.
3. $P(A \cap B) = P(A)P(B)$.

Dowód. Rozważmy następujący ciąg równości:

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

Tym samym dowód równoważności jest zakończony.

Zauważmy, że trzecia równość jest troszeczkę ogólniejsza od dwóch pozostałych, ponieważ pozostaje w mocy nawet w przypadku, gdy $P(A) = P(B) = 0$. Poza tym ma ona nad pozostałymi przewagę symetrii. Prowadzi to do definicji niezależności zdarzeń.

Definicja 6.1. *{Niezależność dwóch zdarzeń.}*

Niech (Ω, \mathcal{A}, P) będzie przestrzenią probabilistyczną. Zdarzenia A i B nazywamy niezależnymi, jeżeli

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

Z niezależności zdarzeń A i B wynika szereg ciekawych własności.

Twierdzenie 6.2. *Jeżeli zdarzenia A i B są niezależne, to niezależne są także zdarzenia:*

1. A i B^c .
2. A^c i B .
3. A^c i B^c .

Dowód.

$$(1) P(A \cap B^c) = P(A) - P(A \cap B) = P(A)(1 - P(B)) = P(A)P(B^c).$$

Analogicznie dowodzimy (2) i (3).

Przykład 6.1. Wykonujemy jeden rzut symetryczną kostką do gry. Niech zdarzenie A polega na wyrzuceniu parzystej liczby oczek a zdarzenie B na wyrzuceniu liczby oczek większej niż 2. Mamy zatem:

$$A = \{2, 4, 6\}, \quad B = \{3, 4, 5, 6\}, \quad \Omega = \{1, 2, 3, 4, 5, 6\}, \quad A \cap B = \{4, 6\}.$$

Stąd

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{2}{3}, \quad P(A \cap B) = \frac{1}{3},$$

i w konsekwencji

$$\frac{1}{3} = P(A \cap B) = P(A) \cdot P(B) = \frac{1}{3}.$$

Równość jest prawdziwa, więc zdarzenia A i B są zdarzeniami niezależnymi.

Korzystając z definicji niezależności zauważmy, że warunek

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

zachodzi, jeżeli $P(B) = 0$, ponieważ $A \cap B \subseteq B$. Jeżeli $P(B) \neq 0$ to dzieląc obie strony wyrażenia przez $P(B)$ otrzymujemy:

$$P(A|B) = P(A)$$

Powyższa równość oznacza, że $P(A)$ i $P(B)$ są niezależne, jeżeli zdarzenie B nie wpływa na A . Jest to bardziej intuicyjne określenie niezależności zdarzeń, które możemy jednak stosować tylko w przypadku zdarzeń o prawdopodobieństwie niezerowym.

Założmy teraz, że A i B są zdarzeniami rozłącznymi. Zastanówmy się czy te zdarzenia są niezależne? Z jednej strony wiemy że

$$P(A \cap B) = 0,$$

ponieważ $(A \cap B) = \emptyset$. Z drugiej strony

$$P(A) \cdot P(B) > 0$$

z wyjątkiem zdegenerowanego przypadku kiedy $P(A) = 0$ lub $P(B) = 0$. Zatem w ogólnym przypadku z tego, że zdarzenia są rozłączne nie wynikają że są niezależne.

Przykład 6.2 (Eksperyment z dwoma monetami). *Założmy że rzucamy dwoma sprawiedliwymi monetami. Oznaczmy zdarzenia:*

A- wypadną dwa orły lub dwie reszki,

B- pierwszy wypadnie orzeł.

Czy zdarzenia A i B są niezależne?

Intuicyjnie odpowiedź brzmi „nie”. Wydaje się bowiem, że wynik pierwszego rzutu ma wpływ na zdarzenie A. Aby się o tym przekonać, sprawdźmy warunek niezależności.

Mamy:

$$P(A) = P(OO) + P(RR) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(B) = P(OO) + P(OR) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(A \cap B) = P(OO) = \frac{1}{4}$$

Zatem zdarzenia A i B są zdarzeniami niezależnymi.

Zastanówmy się teraz, czy zdarzenia A i B w eksperymencie przeprowadzonym w poprzednim przykładzie będą dalej niezależne, jeżeli prawdopodobieństwo wyrzucenia reszki będzie różnić się od prawdopodobieństwa wyrzucenia orła.

Wykonujemy dwa niezależne rzuty moneta niesymetryczną z prawdopodobieństwem wyrzucenia orła równym p i reszki $1 - p$. Niech A i B będą zdarzeniami zdefiniowanymi w poprzednim przykładzie.

Mamy wtedy:

$$P(A) = P(OO) + P(RR) = p^2 + (1 - p)^2 = 2p^2 - 2p + 1$$

$$P(B) = P(OO) + P(OR) = p^2 + p(1 - p) = p$$

$$P(A \cap B) = P(OO) = p^2.$$

Warunek niezależności $P(A \cap B) = P(A) \cdot P(B)$ przyjmuje postać:

$$(2p^2 - 2p + 1) \cdot p = p^2.$$

Stąd $p = 0$ lub $2p^2 - 3p - 1 = 0$, co daje rozwiązania $p = 0$, $p = 1$ i $p = \frac{1}{2}$.

Zatem zdarzenia A i B są niezależne wtedy i tylko wtedy, gdy monety są albo sprawiedliwe albo całkowicie obciążone w kierunku dwóch reszek lub dwóch orłów.

6.2 Niezależność zbioru zdarzeń

Zastanówmy się jak uogólnić definicję niezależności zaproponowaną dla dwóch zdarzeń i dlaczego nie przenosi się ona bezpośrednio na dowolną ilość zdarzeń.

W tym celu rozważmy następujący przykład.

Przykład 6.3. *Urna zawiera cztery żetony: niebieski, biały, czerwony i niebiesko-biało-czerwony. Losujemy jeden żeton. Rozważmy następujące zdarzenia:*

$A = \{\text{wylosowany żeton zawiera kolor niebieski}\},$

$B = \{\text{wylosowany żeton zawiera kolor biały}\},$

$C = \{\text{wylosowany żeton zawiera kolor czerwony}\}.$

Jasne jest, że $P(A) = P(B) = P(C) = 1/2$.

Z drugiej strony:

$$P(A \cap B) = 1/4 = P(A)P(B),$$

$$P(A \cap C) = 1/4 = P(A)P(C),$$

$$P(B \cap C) = 1/4 = P(B)P(C).$$

W ten sposób zdarzenia A, B, C są parami niezależne. Z drugiej jednak strony $P(A \cap B \cap C) = 1/4 \neq 1/8 = P(A)P(B)P(C)$. Ponadto $P(A) = 1/2 \neq 1 = P(A|B \cap C)$, czyli informacja o zdarzeniu B i C zmienia prawdopodobieństwo zdarzenia A . Tak więc A nie jest niezależne od B i C .

Widzimy więc, że niezależność parami jest niewystarczająca aby przenieść ideę niezależności na większą ilość zdarzeń. To uzasadnia następującą definicję.

Definicja 6.2. *Mówimy, że zdarzenia A_1, \dots, A_n są wzajemnie niezależne, jeżeli dla każdego podzbioru A_{i_1}, \dots, A_{i_k} , gdzie $1 \leq i_1 < \dots < i_k \leq n$,*

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \times \dots \times P(A_{i_k}).$$

Rozszerzmy teraz definicję na nieskończoną ilość zdarzeń.

Definicja 6.3. *Mówimy, że nieskończony ciąg zdarzeń A_1, A_2, \dots jest ciągiem niezależnym, jeżeli wszystkie podciągi skończone tego ciągu są niezależne.*

6.3 Schemat Bernoulli'ego

Rozważmy ciąg n powtórzeń jednego doświadczenia mającego jedynie dwa możliwe wyniki. Jeden z tych wyników nazywamy zwyczajowo "sukcesem".

Twierdzenie 6.3 (Bernoulli'ego). *Prawdopodobieństwo uzyskania dokładnie k sukcesów w serii n powtórzeń doświadczenia o prawdopodobieństwie "sukcesu" w pojedynczej próbie równym p wynosi:*

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}; \quad k = 0, 1, \dots, n;$$

gdzie $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Przykład 6.4. *Co jest bardziej prawdopodobne przy grze z przeciwnikiem równej klasy z wykluczeniem remisów:*

- wygranie 3 z 4 partii, czy wygranie 5 z 8 partii?
- wygranie co najmniej 3 z 4 partii, czy wygranie co najmniej 5 z 8 partii?

Ile razy należy rzucić symetryczną monetą, aby prawdopodobieństwo tego, że wypadł przynajmniej jeden orzeł było większe od 0,99?

6.4 Lematy Borela-Cantelliego

Lematy Borela-Cantelliego są szczególnie użyteczne w rozważaniach dotyczących granicznych zachowań ciągów zmiennych losowych.

Dotyczą one nieskończonych ciągów zdarzeń losowych.

Przykład 6.5. Niech (A_n) będzie ciągiem zdarzeń losowych takich, że

$$A_n = \begin{cases} \left(\frac{1}{n}, \frac{2}{3} - \frac{1}{n}\right) & \text{dla } n = 1, 3, 5, \dots, \\ \left(\frac{1}{3} - \frac{1}{n}, 1 + \frac{1}{n}\right) & \text{dla } n = 2, 4, 6, \dots \end{cases}$$

Obliczyć

$$P(\liminf_{n \rightarrow \infty} A_n), \quad \text{oraz} \quad P(\limsup_{n \rightarrow \infty} A_n).$$

Twierdzenie 6.4. *{Lematy Borela-Cantelliego}*

1. Niech (A_n) będzie ciągiem zdarzeń. Jeżeli $\sum_{n=1}^{\infty} P(A_n) < \infty$, to

$$P(\limsup_{n \rightarrow \infty} A_n) = 0;$$

2. Niech (A_n) będzie niezależnym ciągiem zdarzeń. Jeżeli

$$\sum_{n=1}^{\infty} P(A_n) = \infty, \quad \text{to}$$

$$P(\limsup_{n \rightarrow \infty} A_n) = 1.$$

Dowód.(1) Korzystając z twierdzenia o ciągłości oraz z subaddytywności miary probabilistycznej otrzymujemy:

$$\begin{aligned} P(\limsup_{n \rightarrow \infty} A_n) &= P\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n\right) = \lim_{m \rightarrow \infty} P\left(\bigcup_{n=m}^{\infty} A_n\right) \\ &\leq \lim_{m \rightarrow \infty} \sum_{n=m}^{\infty} P(A_n) = 0, \end{aligned}$$

na mocy warunku koniecznego zbieżności szeregu.

Dowód.(2) Korzystając z twierdzenia o ciągłości, niezależności oraz nierówności $1 - x \leq e^{-x}$ otrzymujemy:

$$P\left(\bigcap_{n=m}^{\infty} A'_n\right) = \lim_{N \rightarrow \infty} P\left(\bigcap_{n=m}^N A'_n\right)$$

$$\begin{aligned}
&= \lim_{N \rightarrow \infty} \prod_{n=m}^N P(A'_n) = \lim_{N \rightarrow \infty} \prod_{n=m}^N (1 - P(A_n)) \leq \lim_{N \rightarrow \infty} \prod_{n=m}^N e^{-P(A_n)} \\
&= \lim_{N \rightarrow \infty} e^{-\sum_{n=m}^N P(A_n)} = 0,
\end{aligned}$$

na mocy warunku o rozbieżności szeregu.

Przykład 6.6. *Rzucamy nieskończenie wiele razy niesymetryczną monetą. Niech A_n oznacza zdarzenie, że w pierwszych n rzutach było tyle samo orłów co reszek. Wtedy z prawdopodobieństwem 1 zachodzi skończona liczba zdarzeń A_n .*

Wskazówka: skorzystać ze wzoru Stirlinga:

$$n! \sim n^n e^{-n} \sqrt{2\pi n}.$$

6.5 Zadania

Zadanie 9. Wykazać, że jeżeli zdarzenia A_1, \dots, A_{n+1} są niezależne, to zdarzenia

$$A_1 \cup A_2 \cup \dots \cup A_n, A_{n+1}$$

są niezależne.

Dowód indukcyjny. Niech $n = 2$. Mamy

$$\begin{aligned}
&P((A_1 \cup A_2) \cap A_3) = P((A_1 \cap A_3) \cup (A_2 \cap A_3)) \\
&P((A_1 \cap A_3) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3)) \\
&= P(A_1)P(A_3) + P(A_2)P(A_3) - P(A_1 \cap A_2)P(A_3) \\
&= P(A_3)(P(A_1) + P(A_2) - P(A_1 \cap A_2)) \\
&= P(A_3)(P(A_1 \cup A_2)).
\end{aligned}$$

Zakładając teraz, że zdarzenia $A_1 \cup A_2 \cup \dots \cup A_{n-1}, A_n$ są niezależne podobnie wykazujemy, że zdarzenia

$$A_1 \cup A_2 \cup \dots \cup A_n, A_{n+1}$$

są niezależne.

Zadanie 10. Niech (A_n) będzie ciągiem niezależnych zdarzeń losowych. Wykazać, że

1. (\bar{A}_n) jest ciągiem niezależnych zdarzeń losowych
- 2.

$$P\left(\bigcup_{k=1}^n A_k\right) = 1 - \prod_{k=1}^n (1 - P(A_k)).$$

Rozwiązanie.

1. Dla dwóch niezależnych zdarzeń A_1 i A_2

$$\begin{aligned} P(\bar{A}_1 \cap \bar{A}_2) &= P(\overline{A_1 \cup A_2}) = 1 - P(A_1 \cup A_2) = 1 - P(A_1) - P(A_2) + P(A_1 A_2) \\ &= 1 - P(A_1) - P(A_2) + P(A_1)P(A_2) = 1 - P(A_1) - P(A_2)(1 - P(A_1)) \\ &= (1 - P(A_1))(1 - P(A_2)) = P(\bar{A}_1)P(\bar{A}_2). \end{aligned}$$

Równość

$$P(\bar{A}_{i_1} \cdot \bar{A}_{i_2} \cdot \dots \cdot \bar{A}_{i_n}) = P(\bar{A}_{i_1})P(\bar{A}_{i_2}) \dots P(\bar{A}_{i_n})$$

wynika z zasady indukcji.

- 2.

$$\begin{aligned} P\left(\bigcup_{k=1}^n A_k\right) &= 1 - P\left(\overline{\bigcup_{k=1}^n A_k}\right) = 1 - P\left(\bigcap_{k=1}^n \bar{A}_k\right) \\ &= 1 - \prod_{k=1}^n P(\bar{A}_k) = 1 - \prod_{k=1}^n (1 - P(A_k)). \end{aligned}$$

7 Zmienne losowe jednowymiarowe

7.1 Pojęcie zmiennej losowej

Podstawowym pojęciem rachunku prawdopodobieństwa jest zmienna losowa.

Jak zwykle rozważamy przestrzeń probabilistyczną, złożoną z przestrzeni zdarzeń elementarnych Ω , określonego na niej σ -ciała \mathcal{A} , którego elementy są nazywane zdarzeniami losowymi, oraz miary probabilistycznej P , przyporządkowującej zdarzeniom liczby zwane prawdopodobieństwami. Tak określone prawdopodobieństwa są jednak niewygodne do badania, gdyż Ω może być dowolnym zbiorem, nawet bez zadanych jakichkolwiek relacji między jego elementami. Wprowadza się zatem funkcję mierzalną zwaną zmienną losową, która przyporządkowuje elementom przestrzeni zdarzeń elementarnych Ω elementy pewnej przestrzeni mierzalnej. Wykorzystywać będziemy przestrzeń euklidesową $Y = \mathbb{R}$.

Definicja 7.1. *Zmienną losową (rzeczywistą) określoną na przestrzeni probabilistycznej (Ω, \mathcal{A}, P) nazywamy dowolną rzeczywistą funkcję mierzalną $X : \Omega \rightarrow \mathbb{R}$, tzn. funkcję X spełniającą warunek $X^{-1}(B) \in \mathcal{A}$ dla każdego zbioru borelowskiego $B \subset \mathbb{R}$.*

Zmienna losowa wyznacza nową miarę prawdopodobieństwa. Rzeczywiście, obrazem każdego zdarzenia losowego (elementu rodziny \mathcal{A}) poprzez zmienną losową X jest mierzalny podzbiór Y . Mierzalne podzbiory Y tworzą σ -ciało $\mathcal{B}(Y)$. Ponieważ zmienna losowa nie musi być funkcją różnowartościową, więc ten sam zbiór mierzalny $A \in \mathcal{B}(Y)$ można w ogólnym przypadku otrzymać z wielu różnych zdarzeń o różnych prawdopodobieństwach. Aksjomaty σ -ciała zapewniają, że wśród tych zdarzeń jest także ich suma i do niej jest przypisane największe prawdopodobieństwo. Suma ta jest równa przeciwobrazowi zbioru A , czyli $X^{-1}(A)$. Zatem rozkład zmiennej losowej X to funkcja P_X , określona na $\mathcal{B}(Y)$ wzorem $P_X(A) = P(X^{-1}(A))$, $A \in \mathcal{B}(Y)$.

Rozkład P_X jest właśnie nową miarą probabilistyczną. Jest on w przestrzeni stanów Y odpowiednikiem miary probabilistycznej P .

Wyróżnia się niżej omówione rozkłady ciągłe i dyskretne, jednak należy pamiętać, iż oprócz nich istnieją także rozkłady nie mieszczące się w żadnej z

tych kategorii – na przykład rozkład o dystrybucanie Cantora.

7.2 Zmienne losowe dyskretne

Rozważmy rzut dwiema sześciennymi kostkami: niebieską i czerwoną. Za każdym razem notujemy sumę oczek uzyskaną w obu rzutach. W tym przypadku przestrzeń zdarzeń elementarnych $\Omega = \{1, 2, 3, 4, 5, 6\}^2$ jest przestrzenią skończoną. Zdarzenie elementarne w tym doświadczeniu jest parą (n, c) , gdzie n oznacza liczbę oczek na kostce niebieskiej a c na czerwonej. Wszystkie wyniki tego doświadczenia możemy przedstawić za pomocą poniższej tabelki:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

W ten sposób zdefiniowaliśmy odwzorowanie X ze zbioru Ω w zbiór wszystkich możliwych sum: $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. Oczywiście X jest zmienną losową na (Ω, \mathcal{A}) . Zauważmy, że nie interesuje nas zdarzenie elementarne (n, c) , czyli ω , lecz $S(\omega)$. Ponadto chcielibyśmy znać prawdopodobieństwa, że suma przybiera określoną wartość, tzn. $P(S = k)$ dla całkowitych wartości między 2 a 12. Przedstawia je poniższa tabelka.

k	2	3	4	5	6	7	8	9	10	11	12
P(S=k)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Pozwala to zdefiniować nową przestrzeń zdarzeń elementarnych postaci

$$\Omega' = X(\Omega) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\},$$

i wyposażyć ją w wartości $P(S = k)$, zwane rozkładem zmiennej losowej X . Zauważmy, że otrzymaliśmy nową przestrzeń $(\Omega', \mathcal{A}, P)$ oraz zmienną losową, określoną na tej przestrzeni, spełniającą warunki

$$X : \Omega \rightarrow \mathbf{R} \quad (\omega \rightarrow X(\omega)),$$

oraz

1. Zbiór wartości $X(\Omega' = \{S(\omega), \omega \in \Omega'\})$ jest co najwyżej przeliczalny.
2. Dla każdego $x_k \in X(\Omega)$, zbiór $A_k = \{\omega \in \Omega, X(\omega) = x_k\}$ należy do σ -ciała \mathcal{A} .

Powyższe warunki definiują klasę tzw. **dyskretnych zmiennych losowych** a więc takich, które przyjmują co najwyżej przeliczalnie wiele wartości i są wyznaczone przez te wartości oraz prawdopodobieństwa ich otrzymania.

Zdefiniujemy rozkład prawdopodobieństwa dla dyskretnych zmiennych losowych.

Definicja 7.2. Niech X będzie dyskretną zmienną losową określoną na przestrzeni probabilistycznej (Ω, \mathcal{A}, P) . Funkcję zbioru P_X , zdefiniowaną dla dowolnej wartości x_k przez

$$p_k = P_X(\{x_k\}) = P(A_k) = P(X = x_k),$$

oraz dla dowolnego zbioru wartości $B \subset \mathbf{R}$ przez

$$P_X(B) = \sum_{x_k \in B} P(X = x_k) = \sum_{x_k \in B} p_k,$$

nazywamy rozkładem prawdopodobieństwa zmiennej losowej X .

Uwaga 7.1. Dwie zmienne losowe, które nie są równe z prawdopodobieństwem 1 mogą mieć ten sam rozkład prawdopodobieństwa.

Dla przykładu rozważmy jak poprzednio rzut dwiema sześciennymi kostkami do gry, niebieską i czerwoną. Niech X oznacza liczbę oczek uzyskaną na

pierwszej a Y na drugiej kostce. Zmienne losowe X i Y są określone na tej samej przestrzeni probabilistycznej i

$$\forall k \in \{1, 2, 3, 4, 5, 6\} \quad P(X = k) = P(Y = k) = 1/6.$$

Mają one zatem ten sam rozkład prawdopodobieństwa: $P_X = P_Y$. Nie ma jednak pomiędzy nimi równości w sensie $P(X = Y) = 1$. Rzeczywiście

$$P(X = Y) = P\left(\bigcup_{k=1}^6 \{(X, Y) = (k, k)\}\right) = 6/36 = 1/6,$$

zatem $P(X \neq Y) = 5/6$.

7.3 Zmienne losowe absolutnie ciągłe

W poprzednim paragrafie przedmiotem naszego zainteresowania były zmienne losowe, które przyjmowały co najwyżej przeliczalnie wiele wartości. Ale w wielu przypadkach takie zmienne okazują się niewystarczające do opisu pewnych zjawisk losowych. Rozważmy na przykład eksperyment polegający na mierzeniu czasu działania żarówki lub mierzeniu wartości temperatury. W tych przypadkach otrzymane wartości należą do nieprzeliczalnego zbioru, dlatego nasze rozważania musimy rozszerzyć na przestrzenie nieprzeliczalne. Będziemy mieli do czynienia z rodzajem zmiennych losowych, które posiadają funkcję gęstości.

Rozważmy pewną funkcję całkowalną f określoną na zbiorze liczb rzeczywistych, spełniającą następujące warunki:

1. $\forall x \in \mathbf{R} : f(x) \geq 0$;
2. $\int_{-\infty}^{\infty} f(x) dx = 1$.

Funkcję f nazywamy funkcją gęstości na \mathbf{R} .

Możemy teraz zdefiniować klasę zmiennych losowych ciągłych. Podobnie jak w przypadku wcześniejszym są to mierzalne odwzorowania postaci

$X : \Omega \rightarrow \mathbf{R}$, $\omega \rightarrow X(\omega)$, ale w tym przypadku dla dowolnego przedziału $[a, b]$ mamy

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

Takie zmienne losowe nazywamy absolutnie ciągłymi.

7.4 Zadania

Zadanie 11. Załóżmy, że X_1, X_2, \dots są zmiennymi losowymi. Wykazać, że następujące wielkości są zmiennymi losowymi:

(a) $\max\{X_1, X_2\}, \min\{X_1, X_2\}$

(b) $\sup_n X_n, \inf_n X_n$

(c) $\limsup_{n \rightarrow \infty} X_n$ i $\liminf_{n \rightarrow \infty} X_n$

(d) Jeżeli $(X_n(\omega))$ jest zbieżny dla każdego $\omega \in \Omega$, gdy $n \rightarrow \infty$, to $\lim_{n \rightarrow \infty} X_n$ jest zmienną losową.

Rozwiązanie.

(a) Dla każdego x

$$\begin{aligned}\{\omega : \max\{X_1, X_2\}(\omega) \leq x\} &= \{\omega : \max\{X_1(\omega), X_2(\omega)\} \leq x\} \\ &= \{\omega : X_1(\omega) \leq x\} \cap \{\omega : X_2(\omega) \leq x\}, \\ \{\omega : \min\{X_1, X_2\}(\omega) \leq x\} &= \{\omega : \min\{X_1(\omega), X_2(\omega)\} \leq x\} \\ &= \{\omega : X_1(\omega) \leq x\} \cup \{\omega : X_2(\omega) \leq x\},\end{aligned}$$

co dowodzi (a), bo iloczyn i suma zbiorów mierzalnych są mierzalnymi zbiorami.

(b) Podobnie jak wyżej

$$\{\omega : \sup_n X_n(\omega) \leq x\} = \bigcap_n \{\omega : X_n(\omega) \leq x\} \in \mathcal{A}$$

bo przeliczalny iloczyn zbiorów mierzalnych jest zbiorem mierzalnym i

$$\{\omega : \inf_n X_n(\omega) < x\} = \bigcup_n \{\omega : X_n(\omega) < x\} \in \mathcal{A}$$

bo przeliczalna suma zbiorów mierzalnych jest mierzalna.

(c) W tym przypadku

$$\begin{aligned}\{\omega : \limsup_{n \rightarrow \infty} X_n(\omega) \leq x\} &= \{\omega : \inf_n \sup_{m \geq n} X_m(\omega) \leq x\} \\ &= \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{\omega : X_m(\omega) \leq x\} \in \mathcal{A}\end{aligned}$$

i

$$\begin{aligned}\{\omega : \liminf_{n \rightarrow \infty} X_n(\omega) < x\} &= \{\omega : \sup_n \inf_{m \geq n} X_m(\omega) < x\} \\ &= \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{\omega : X_m(\omega) < x\} \in \mathcal{A}.\end{aligned}$$

Inny dowód.

Ponieważ $\sup_{n \geq m} X_n(\omega)$ jest zmienną losową na mocy (b), to stąd wynika również na mocy (b), że $\inf_m (\sup_{n \geq m} X_n(\omega))$ jest zmienną losową. Podobnie dla $\liminf_{n \rightarrow \infty} X_n(\omega)$.

(d) Teza wynika z faktu, że $\limsup_{n \rightarrow \infty} X_n$ i $\liminf_{n \rightarrow \infty} X_n$ są zmiennymi losowymi.

Zadanie 12. Wykazać, że funkcja

$$f_X(x) = \frac{\alpha}{2} e^{-\alpha|x|}, \quad \alpha > 0, \quad -\infty < x < \infty,$$

jest gęstością prawdopodobieństwa.

Rozwiązanie.

$$\begin{aligned} f_X(x) &= \frac{\alpha}{2} e^{-\alpha|x|} \geq 0. \\ \int_{-\infty}^{\infty} f_X(x) dx &= \frac{\alpha}{2} \int_{-\infty}^{\infty} e^{-\alpha|x|} dx = \frac{\alpha}{2} \left[\int_{-\infty}^0 e^{\alpha x} dx + \int_0^{\infty} e^{-\alpha x} dx \right] \\ &= \frac{\alpha}{2} \left[\frac{1}{\alpha} e^{\alpha x} \Big|_{-\infty}^0 - \frac{1}{\alpha} e^{-\alpha x} \Big|_0^{\infty} \right] \\ &= \frac{\alpha}{2} \left[\frac{1}{\alpha} + \frac{1}{\alpha} \right] = 1 \end{aligned}$$

8 Dystrybuanta zmiennej losowej

8.1 Dystrybuanta i jej własności

Istnieje wiele narzędzi do badania własności zmiennych losowych. Jednym z nich jest dystrybuanta.

Definicja 8.1. *Dystrybuantą zmiennej losowej X nazywamy funkcję F_X określoną na zbiorze liczb rzeczywistych taką, że*

$$\forall x \in \mathbf{R}, \quad F_X(x) = P_X\left((-\infty, x]\right) = P(X \leq x).$$

Twierdzenie 8.1. *Niech F_X będzie dystrybuantą zmiennej losowej. Wtedy*

1. F_X jest niemalejąca.
2. F_X jest prawostronnie ciągła.
3. F_X ma granice lewostronne.

4. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ oraz $\lim_{x \rightarrow \infty} F_X(x) = 1$.
5. $F_X(x)$ jednoznacznie wyznacza rozkład prawdopodobieństwa, tzn $F_x = F_y$ wtedy i tylko wtedy, gdy X i Y mają jednakowy rozkład prawdopodobieństwa.

Dowód. Pokażemy, że dystrybuanta jest niemalejąca.

Niech s i t będą liczbami rzeczywistymi takimi, że $s \leq t$. Wtedy

$\{X \leq s\} \subseteq \{X \leq t\}$, skąd $P(X \leq s) \leq P(X \leq t)$ i w konsekwencji $F_X(s) \leq F_X(t)$.

Aby pokazać, że F_X jest prawostronnie ciągła skorzystamy z udowodnionego wcześniej lematu o ciągłości. Wystarczy pokazać, że

$$F_X(x) = \lim_{n \rightarrow \infty} F_X(x + 1/n).$$

Aby to zrobić stosujemy lemat o ciągłości dla ciągu malejącego postaci:

$$B_n = \{X \leq x + 1/n\}.$$

Stąd

$$\bigcap_{n \in \mathbf{N}} \{X \leq x + 1/n\} = \{X \leq x\},$$

czego rezultatem jest prawostronna ciągłość dystrybuanty. Podobnie pokazujemy, że ma lewostronne granice.

Dalej, wykorzystując lemat o ciągłości i fakt, że

$$\bigcap_{n \in \mathbf{N}} \{X \leq -n\} = \emptyset,$$

oraz

$$\bigcup_{n \in \mathbf{N}} \{X \leq n\} = \Omega,$$

otrzymujemy (4). Załóżmy teraz, że zmienne losowe X i Y mają jednakowy rozkład prawdopodobieństwa. To znaczy, że dla dowolnego $B \subset \mathbf{R}$, $P_X(B) = P_Y(B)$. Wybierając $B = (-\infty, t]$ dostajemy $F_X(t) = F_Y(t)$. Drugą część (5) pozostawiamy jako ćwiczenie.

8.2 Przypadek dyskretny

Pojęcie dystrybuanty wprowadzone zostało dla dowolnej zmiennej losowej. W zależności jednak od charakteru tej zmiennej warto rozróżnić klasę dystrybuant dla zmiennych losowych dyskretnych i ciągłych.

W przypadku, gdy X jest zmienną losową dyskretną, jej dystrybuantę możemy zapisać następująco:

$$F_X(x) = P_X((-\infty, x]) = P(X \leq x) = \sum_{x_k \in X, x_k \leq x} P(X = x_k).$$

W tym wypadku $F_X(x)$ jest funkcją "schodkową", spełniającą oczywiście wszystkie własności dystrybuanty.

8.3 Przypadek ciągły

Ze względu na to, że dla zmiennej losowej ciągłej prawdopodobieństwo wyraża się za pomocą całki z funkcji gęstości, dodatkowe własności dystrybuanty opiszemy w następującym twierdzeniu.

Twierdzenie 8.2. *Jeżeli zmienna losowa X ma funkcję gęstości f , to jej dystrybuanta F spełnia następujące warunki:*

1. $\forall x \in \mathbf{R}, \quad F(x) = \int_{-\infty}^x f(t)dt;$
2. F jest ciągła na \mathbf{R} ;
3. Jeżeli f jest ciągła w x_0 , to F jest różniczkowalna w tym punkcie i $F'(x_0) = f(x_0)$.

Dowód.

Ad 1. Ponieważ X ma funkcję gęstości f , dla dowolnych liczb rzeczywistych $a < b$ mamy

$$P(X \in (a, b]) = F(b) - F(a) = \int_a^b f(t)dt.$$

Aby udowodnić (1) wystarczy zastosować powyższy wzór przyjmując $b = x$ oraz $a = -n$ takie, że $-n < x$. Zauważmy, że ciąg zdarzeń

$$A_n = \{X \in (-n, x]\}, \quad n > -x,$$

jest ciągiem rosnącym ze względu na inkluzję i nieskończona suma dla tego ciągu wynosi $A = \{X \in (-\infty, x]\}$. Z lematu o ciągłości $P(A_n) \uparrow P(A)$, skąd

$$F(x) = P(A) = \lim_{n \rightarrow \infty} P(X \in (-n, x]) = \lim_{n \rightarrow \infty} \int_{-n}^x f(t) dt = \int_{-\infty}^x f(t) dt.$$

Dowód (2). Ustalmy $x_0 \in \mathbf{R}$. Jak wiadomo F jest prawostronnie ciągła w dowolnym punkcie. Wystarczy więc udowodnić lewostronną ciągłość w x_0 . Z własności całek i funkcji gęstości wiemy jednak, że

$$\lim_{x \uparrow x_0} \int_a^x f(t) dt = \int_a^{x_0} f(t) dt,$$

co w języku dystrybuant możemy zapisać następująco

$$\lim_{x \uparrow x_0} (F(x) - F(a)) = F(x_0) - F(a).$$

Stąd dystrybuanta jest lewostronnie ciągła a więc ciągła.

Dowód (3). Z tego, że f jest ciągła w punkcie x_0 ,

$$\forall \epsilon > 0, \exists \delta > 0 : \quad \forall t \in (x_0 - \delta; x_0 + \delta), \quad |f(t) - f(x_0)| < \epsilon.$$

Dla każdego h takiego, że $0 < |h| < \delta$ z tego, że

$$F(x_0 + h) - F(x_0) = \int_{x_0}^{x_0+h} f(t) dt$$

mamy:

$$|F(x_0 + h) - F(x_0) - hf(x_0)| = \left| \int_{x_0}^{x_0+h} (f(t) - f(x_0)) dt \right| \leq h\epsilon.$$

Dzieląc obie strony przez h widzimy, że F ma pochodną w punkcie x_0 równą $f(x_0)$.

8.4 Zadania

Zadanie 13. Zmienna losowa X podlega rozkładowi według gęstości danej wzorem

$$f(x) = \begin{cases} 0 & \text{dla } x < 1; \\ \ln x & \text{dla } 1 \leq x < a; \\ 0 & \text{dla } x \geq a. \end{cases}$$

- (a) Obliczyć a .
- (b) Podać dystrybuantę.
- (c) Obliczyć $P(2 \leq X \leq e)$.

Rozwiązanie.

(a)

$$\int_1^a \ln x dx = [x \ln x - x]_1^a = a \ln a - a - \ln 1 + 1;$$

$$a(\ln a - 1) + 1 = 1;$$

$$a(\ln a - 1) = 0;$$

$$a = 0 \vee \ln a = 1 \Leftrightarrow \ln a = \ln e \Leftrightarrow a = e.$$

Zatem

$$f(x) = \begin{cases} 0, & x < 1; \\ \ln x, & 1 \leq x < e; \\ 1, & x \geq e. \end{cases}$$

(b) Dla $x \in (1, e]$ mamy

$$F(x) = \int_1^x \ln x dx = [x \ln x - x]_1^x = x(\ln x - 1) + 1.$$

Stąd

$$F(x) = \begin{cases} 0, & \text{dla } x < 1 \\ x(\ln x - 1) + 1, & \text{dla } 1 \leq x < e \\ 1, & \text{dla } x \geq e \end{cases}$$

(c)

$$\begin{aligned} P(2 \leq X \leq e) &= \int_2^e \ln x dx = [x \ln x - x]_2^e = e \ln e - e - 2 \ln 2 + 2 \\ &= e - e - 2 \ln 2 + 2 = 2(1 - \ln 2). \end{aligned}$$

Zadanie 14. Zmienna losowa X ma rozkład zadany gęstością

$$f(x) = \begin{cases} \frac{a}{\sqrt{1-x^2}}, & |x| \leq 1; \\ 0, & |x| > 1. \end{cases}$$

1. Wyznaczyć stałą a tak, aby funkcja f była gęstością zmiennej losowej o rozkładzie absolutnie ciągłym.
2. Wyznaczyć dystrybuantę zmiennej X .
3. Wyznaczyć $P\left(0 < X \leq \frac{\sqrt{3}}{2}\right)$, $P\left(|X| < \frac{\sqrt{2}}{2}\right)$.

a)

Aby f była gęstością musi być spełniony warunek $\int_{-1}^1 f(x) dx = 1$, zatem

$$a \int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} = a \cdot \arcsin x \Big|_{-1}^1 = a \left(\frac{\pi}{2} + \frac{\pi}{2} \right) = a\pi$$

a stąd dostajemy $a = \frac{1}{\pi}$.

b)

$$F(x) = \begin{cases} 0 & \text{dla } x < -1; \\ \frac{1}{2} + \frac{\arcsin x}{\pi} & \text{dla } x \in \langle -1, 1 \rangle; \\ 1 & \text{dla } x \geq 1. \end{cases}$$

c)

$$\begin{aligned} P\left(0 < X \leq \frac{\sqrt{3}}{2}\right) &= \frac{\arcsin \frac{\sqrt{3}}{2} - \arcsin 0}{\pi} = \frac{\frac{\pi}{3} - 0}{\pi} = \frac{1}{3}, \\ P\left(|X| < \frac{\sqrt{2}}{2}\right) &= \frac{\arcsin \frac{\sqrt{2}}{2} - \arcsin \left(-\frac{\sqrt{2}}{2}\right)}{\pi} = \frac{\frac{\pi}{4} + \frac{\pi}{4}}{\pi} = \frac{1}{2}. \end{aligned}$$

9 Przykłady rozkładów dyskretnych

9.1 Rozkład Bernoulliego

Definicja 9.1. Zmienna losowa X ma rozkład Bernoulliego z parametrem p ($p \in (0, 1)$), jeżeli przyjmuje jedynie dwie wartości 0 oraz 1 z prawdopodobieństwami:

$$P(X = 1) = p, \quad P(X = 0) = 1 - p = q.$$

Oznaczenie rozkładu: $X \sim B(p)$.

Zauważmy, że jeżeli A jest zdarzeniem takim, że $P(A) = p$, to jego indykator określony wzorem:

$$I_A(\omega) = \begin{cases} 1 & \text{dla } \omega \in A, \\ 0 & \text{dla } \omega \in A^c, \end{cases}$$

jest zmienną losową o rozkładzie Bernoulliego z parametrem p .

9.2 Rozkład jednostajny na zbiorze skończonym

Definicja 9.2. Zmienna losowa X ma rozkład jednostajny na skończonym zbiorze liczb rzeczywistych $\{x_1, \dots, x_n\}$, jeżeli X jest równorozłożona na tym zbiorze, tzn.

$$\forall k = 1, \dots, n \quad P(X = x_k) = \frac{1}{n}.$$

Na przykład liczba oczek wylosowanych w jednym rzucie kostką sześcienną ma rozkład jednostajny na zbiorze $\{1, 2, 3, 4, 5, 6\}$.

9.3 Rozkład dwumianowy

Definicja 9.3. Zmienna losowa X ma rozkład dwumianowy z parametrami n i p ($n \in \mathbb{N}$, $p \in (0, 1)$), jeżeli jej rozkład prawdopodobieństwa dany jest wzorem:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad k = 0, 1, \dots, n.$$

Oznaczenie rozkładu: $X \sim B(n, p)$.

Rozkład dwumianowy jest rozkładem liczby sukcesów otrzymanych w n niezależnych próbach wykonanych zgodnie ze schematem Bernoulliego, tzn. w każdej próbie prawdopodobieństwo sukcesu jest równe p a porażki $q = 1 - p$. Inaczej mówiąc, jeżeli A_1, \dots, A_n są niezależnymi zdarzeniami takimi, że dla każdego $k \in \{1, \dots, n\}$, $P(A_k) = p$ i jeżeli zmienna losowa X_k jest indykatorem zdarzenia A_k ($X_k \sim B(p)$), to zmienna losowa $S_n = \sum_{k=1}^n X_k$ ma rozkład dwumianowy.

9.4 Rozkład hipergeometryczny

Podczas gdy rozkład dwumianowy możemy w języku „urnowym” traktować jako rozkład ze zwracaniem, to rozkład hipergeometryczny jest rozkładem bez zwracania.

Przykład 9.1. Załóżmy, że mamy N obiektów, wśród których M jest wadliwych. Losujemy bez zwracania n obiektów. Niech X oznacza liczbę wylosowanych obiektów, które są wadliwe. Widzimy, że

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}.$$

Oczywiście musimy jeszcze nałożyć pewne założenia na parametry rozkładu, tzn. $0 \leq k \leq M$ oraz $0 \leq n - k \leq N - M$.

Definicja 9.4. Zmienna losowa X ma rozkład hipergeometryczny z parametrami N , M i n jeżeli jej rozkład prawdopodobieństwa dany jest wzorem:

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}},$$

dla $0 \leq k \leq M$ oraz $0 \leq n - k \leq N - M$.

Oznaczenie rozkładu: $X \sim H(N, M, n)$.

9.5 Rozkład geometryczny

Przykład 9.2. Rozważmy niekończony ciąg niezależnych prób Bernoulliego z jednakowym prawdopodobieństwem sukcesu równym $p \in (0, 1)$. Niech X oznacza numer pierwszego uzyskanego sukcesu w ciągu tych doświadczeń. Jeżeli sukces nie zostanie osiągnięty oznaczmy $X = \infty$. Obliczmy $P(X = k)$. W tym celu wprowadźmy następujące zdarzenia:

$$R_i = \{\text{sukces w } i\text{-tej próbie}\}.$$

Mamy zatem

$$\begin{aligned} \{X = k\} &= \{\text{porażki w } (k-1) \text{ pierwszych próbach i sukces w } k\text{-tej próbie}\} \\ &= \left(\bigcap_{i=1}^{k-1} R_i^c \right) \cap R_k. \end{aligned}$$

Stosując założenie o niezależności otrzymujemy:

$$P(X = k) = \left(\prod_{i=1}^{k-1} P(R_i^c) \right) \times P(R_k) = (1-p)^{k-1}p.$$

Zauważmy, że

$$\sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} (1-p)^{k-1}p = p \sum_{k=0}^{\infty} (1-p)^k = p \frac{1}{1-(1-p)} = 1.$$

Stąd, z prawdopodobieństwem 1 pierwszy sukces pojawi się w skończonej próbie ($P(X = \infty) = 0$). Można zatem traktować X jako zmienną losową o wartościach w zbiorze $\{1, 2, \dots\}$.

Definicja 9.5. Zmienna losowa X ma rozkład geometryczny z parametrem $p \in (0, 1)$, jeżeli jej rozkład prawdopodobieństwa dany jest wzorem:

$$P(X = k) = (1-p)^{k-1}p,$$

dla $k \in \{1, 2, \dots\}$.

Oznaczenie rozkładu: $X \sim G(p)$.

Przykład 9.3. Dla zmiennej losowej X o rozkładzie geometrycznym z parametrem p obliczmy prawdopodobieństwo $P(X > n)$.

Skorzystajmy bezpośrednio z funkcji rozkładu zmiennej X . Mamy ($q = 1 - p$):

$$\begin{aligned} P(X > n) &= \sum_{k=n+1}^{\infty} q^{k-1} p = \sum_{k=n}^{\infty} q^k p \\ &= pq^n \sum_{k=n}^{\infty} q^{k-n} = pq^n \sum_{k=0}^{\infty} q^k = \frac{pq^n}{1-q} = q^n. \end{aligned}$$

9.6 Rozkład ujemny dwumianowy. Rozkład Pascala.

Definicja 9.6. Zmienna losowa X ma rozkład ujemny dwumianowy z parametrami r i p ($r > 0, p \in (0, 1)$), jeżeli jej rozkład prawdopodobieństwa dany jest wzorem:

$$P(X = k) = \binom{r+k-1}{k} (1-p)^k p^r, \quad k = 0, 1, \dots$$

Oznaczenie rozkładu: $X \sim NB(n, p)$. Jeżeli parametr r jest nieujemny i całkowity, zmienną X możemy interpretować jako liczbę porażek poprzedzających r -ty sukces a jej rozkład nazywamy wtedy rozkładem Pascala.

9.7 Rozkład Poissona

Definicja 9.7. Zmienna losowa X ma rozkład Poissona z parametrem $\lambda > 0$, jeżeli jej rozkład prawdopodobieństwa dany jest wzorem:

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

dla $k \in \{0, 1, 2, \dots\}$.

Oznaczenie rozkładu: $X \sim P(\lambda)$.

Wiadomo z analizy, że funkcja wykładnicza ma rozwinięcie w szereg nieskończony na całej osi rzeczywistej postaci:

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}, \quad \lambda > 0.$$

Możemy zatem z łatwością sprawdzić, że X jest rozkładem prawdopodobieństwa.

Rzeczywiście:

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

Twierdzenie 9.1. *{Twierdzenie Poissona}*

Niech S_n będzie ciągiem zmiennych losowych o rozkładach dwumianowych $B(n, p_n)$.

Wówczas jeżeli $\lim_{n \rightarrow \infty} np_n = \lambda > 0$, to

$$\lim_{n \rightarrow \infty} P(S_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Dowód. Z definicji rozkładu dwumianowego dostajemy, że dla $k = 0, \dots, n$,

$$P(S_n = k) = \binom{n}{k} p_n^k (1 - p_n)^{n-k}.$$

Niech $\lambda_n = np_n$, wtedy mamy z założenia, że $\lambda_n \rightarrow \lambda$. Mamy zatem

$$\begin{aligned} P(S_n = k) &= \binom{n}{k} p_n^k (1 - p_n)^{n-k} \\ &= n \cdot (n-1) \cdot (n-2) \dots (n-k+1) \frac{(np_n)^k}{n^k k!} \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \\ &= \frac{n-1}{n} \dots \frac{n-k+1}{n} \cdot \frac{(np_n)^k}{k!} \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

9.8 Zadania

Zadanie 15. Wykazać, że przy spełnieniu założeń dotyczących rozkładów: dwumianowego, ujemnie dwumianowego, hipergeometrycznego i Poissona zachodzą następujące związki:

1.

$$p(x) = \frac{\binom{Np}{x} \binom{Nq}{n-x}}{\binom{N}{n}} \rightarrow \frac{n!}{x!(n-x)!} p^x q^{n-x}, \quad \text{gdy} \quad N \rightarrow \infty.$$

2.

$$\binom{n+k-1}{k} p^n q^k \longrightarrow \frac{\lambda^k}{k!} e^{-\lambda},$$

gdy $p \rightarrow 1$ i $n \rightarrow \infty$ w taki sposób, że $nq = \lambda$.

(1)

$$p(x) = \frac{\binom{Np}{x} \binom{Nq}{n-x}}{\binom{N}{n}} = \frac{\frac{(Np)!}{x!(Np-x)!} \frac{(Nq)!}{(n-x)!(Nq-n+x)!}}{\frac{N!}{n!(N-n)!}}$$

$$= \frac{n!}{x!(n-x)!} \cdot \frac{Np(Np-1) \dots (Np-x+1) Nq(Nq-1) \dots (Nq-n+x+1)}{N(N-1) \dots (N-n+1)}.$$

Ponieważ w liczniku drugiego ułamka występuje $x + (n-x) = n$ czynników i w mianowniku również n czynników, to po podzieleniu przez N otrzymujemy

$$p(x) = \frac{n!}{x!(n-x)!} \cdot \frac{p \left(p - \frac{1}{N}\right) \dots \left(p - \frac{x-1}{N}\right) q \left(q - \frac{1}{N}\right) \dots \left(q - \frac{n-x-1}{N}\right)}{1 \cdot \left(1 - \frac{1}{N}\right) \dots \left(1 - \frac{n-1}{N}\right)}$$

$$\longrightarrow \frac{n!}{x!(n-x)!} p^x q^{n-x}, \quad \text{gdym } N \rightarrow \infty.$$

(2)

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

$$= \frac{1}{k!} \cdot \frac{1 \cdot 2 \cdot \dots \cdot (n-1)n(n+1) \cdot \dots \cdot (n+k-1)}{(n-1)!} p^n q^k$$

$$= \frac{1}{k!} \left(1 + \frac{1}{n}\right) \left(1 + \frac{2}{n}\right) \dots \left(1 + \frac{k-1}{n}\right) p^n (nq)^k$$

$$= \left[nq = \lambda, q = \frac{\lambda}{n}, p = 1 - \frac{\lambda}{n} \right]$$

$$= \frac{1}{k!} \left(1 + \frac{1}{n}\right) \left(1 + \frac{2}{n}\right) \dots \left(1 + \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^n \lambda^k \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}, \quad n \rightarrow \infty.$$

Zadanie 16. Pewne urządzenie kontrolne składa się z n niezależnie pracujących czujników przy czym prawdopodobieństwo wykrycia awarii przez każdy z czujników wynosi $p > 0$. Opisanym urządzeniem kontrolnym N -krotnie sprawdzono stan maszyn w fabryce. Niech X_k ($k = 0, 1, \dots, n$) oznacza liczbę testów przy których dokładnie k czujników sygnalizowało awarię. Podać rozkład zmiennej losowej X_k .

Rozwiązanie. Oznaczmy:

A_k – zdarzenie polegające na tym, że podczas kontroli k czujników zasygnalizowało awarię.

Mamy wtedy

$$P(A_k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Ponieważ kolejne sprawdzanie maszyn jest dokonywane niezależnie od poprzednich, to mamy do czynienia z N -krotnym powtarzaniem tego samego doświadczenia, a zatem

$$P(X_k = i) = \binom{N}{i} P(A_k)^i (1 - P(A_k))^{N-i}, \quad i = 0, 1, \dots, N.$$

Wobec tego zmienna losowa X_k ma rozkład dwumianowy $B\left(N, \binom{n}{k} p^k (1-p)^{n-k}\right)$.

10 Przykłady rozkładów ciągłych

Przedstawimy teraz krótką listę najbardziej znanych i wykorzystywanych rozkładów absolutnie ciągłych. Rozpocznemy od rozkładu jednostajnego.

10.1 Rozkład jednostajny

Definicja 10.1. *Zmienna losowa X ma rozkład jednostajny na przedziale $[a, b]$, jeżeli jej funkcja gęstości jest postaci*

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{dla } x \in [a, b], \\ 0 & \text{poza.} \end{cases}$$

Oznaczenie rozkładu: $X \sim U(a, b)$.

Następujący wzór pozwala w ogólnym przypadku uniknąć dosyć uciążliwych rachunków w celu wyznaczenia dla rozkładu jednostajnego wartości $P(X \in B)$, gdzie B jest skończoną lub przeliczalną sumą przedziałów liczbowych. (Dowód pozostawiamy jako ćwiczenie).

Stwierdzenie. Jeżeli X ma rozkład jednostajny na przedziale $[a, b]$, to dla dowolnego przedziału $I \in \mathbf{R}$,

$$P(X \in I) = \frac{l([a, b] \cap I)}{l([a, b])},$$

gdzie $l(J)$ oznacza długość przedziału J .

Dużą wagę rozkładu jednostajnego w zastosowaniach uzasadnia następujące twierdzenie (prawdziwe także w ogólnym przypadku dla dowolnej dystrybuanty):

Twierdzenie 10.1. *Jeżeli X jest zmienną losową o ciągłej ściśle rosnącej dystrybuancie F i jeżeli U jest zmienną o rozkładzie jednostajnym na przedziale $[0, 1]$, to zmienna losowa $Y := F^{-1}(U)$ ma rozkład taki jak zmienna X .*

Dowód. Z tego, że F jest ciągła i ściśle rosnąca, jest bijekcją z \mathbf{R} na przedział $(0, 1)$. Dlatego funkcja odwrotna F^{-1} jest dobrze zdefiniowana i mamy:

$$\forall u \in (0, 1), \quad \forall x \in \mathbf{R}, \quad F^{-1}(u) \leq x$$

wtedy i tylko wtedy, gdy $u \leq F(x)$.

Dla dowolnego $x \in \mathbf{R}$:

$$P(Y \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = \frac{l([0, F(x)])}{l([0, 1])} = F(x).$$

10.2 Rozkład wykładniczy

Definicja 10.2. Zmienna losowa X ma rozkład wykładniczy z parametrem $\lambda > 0$, jeżeli jej funkcja gęstości jest postaci

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{dla } x > 0, \\ 0 & \text{poza.} \end{cases}$$

Oznaczenie rozkładu: $X \sim E(\lambda)$. W praktyce, w przypadku rozkładu wykładniczego, zamiast dystrybuanty F używa się tak zwanej dystrybuanty ogonowej danej wzorem $G(x) = 1 - F(x)$. Wtedy

$$G(x) = \begin{cases} 1 & \text{dla } x \leq 0, \\ e^{-\lambda x} & \text{dla } x > 0. \end{cases}$$

W praktyce rozkład wykładniczy jest często używany do modelowania takich zdarzeń jak na przykład czas oczekiwania (czas oczekiwania na trzęsienie ziemi, czas oczekiwania na pierwszą szkodę itp.). Wiele zastosowań dla tego rozkładu wynika z własności zwanej własnością braku pamięci, która charakteryzuje rozkład wykładniczy w klasie rozkładów ciągłych.

Założmy, że zmienna losowa X jest absolutnie ciągła. Zachodzi wtedy następujące twierdzenie, które przytoczymy bez dowodu:

Twierdzenie 10.2. *{Własność braku pamięci} Zmienna losowa X ma rozkład wykładniczy wtedy i tylko wtedy, gdy spełnia własność braku pamięci tzn.*

$$\forall s \in \mathbf{R}^+, \forall t \in \mathbf{R}^+, P(X > t + s) | X > t = P(X > s).$$

10.3 Rozkład normalny

Zmienna losowa X ma rozkład normalny z parametrami μ i σ , jeżeli jej funkcja gęstości ma postać

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbf{R}; \quad \mu \in \mathbf{R}, \quad \sigma > 0.$$

Oznaczenie rozkładu: $X \sim N(\mu, \sigma)$.

Twierdzenie 10.3. *Jeżeli zmienna losowa X ma rozkład normalny $N(\mu, \sigma)$, to zmienna losowa $Y = \frac{X-\mu}{\sigma}$ ma rozkład normalny $N(0, 1)$.*

Dowód. Obliczmy $P(a < Y \leq b)$ dla dowolnych a i b ($a < b$). Mamy

$$\begin{aligned} P\left(a < \frac{X-\mu}{\sigma} \leq b\right) &= P(\sigma a + \mu < X \leq \sigma b + \mu) \\ &= \int_{\sigma a + \mu}^{\sigma b + \mu} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx. \end{aligned}$$

Podstawiając $y = (x - \mu)/\sigma$ otrzymujemy tezę.

10.4 Rozkład gamma

Niech $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$. Funkcję $\Gamma(\alpha)$ nazywamy funkcją gamma. Można zauważyć, że $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$, a w konsekwencji $\Gamma(n + 1) = n!$.

Definicja 10.3. *Zmienna losowa X ma rozkład gamma z parametrami α i β , jeżeli jej funkcja gęstości jest postaci*

$$f(x) = \begin{cases} \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)} & \text{dla } x > 0; \alpha > 0, \beta > 0, \\ 0 & \text{poza.} \end{cases}$$

Oznaczenie rozkładu: $X \sim \Gamma(\alpha, \beta)$.

10.5 Rozkład beta

Funkcję $\beta(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$, nazywamy funkcją beta.

Można zauważyć, że $\beta(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

Definicja 10.4. Zmienna losowa X ma rozkład beta z parametrami α i β , jeżeli jej funkcja gęstości jest postaci

$$f(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\beta(\alpha, \beta)} & \text{dla } x \in (0, 1); \alpha > 0, \beta > 0, \\ 0 & \text{poza.} \end{cases}$$

Oznaczenie rozkładu: $X \sim \beta(\alpha, \beta)$.

10.6 Zadania

Zadanie 17. Zmienna losowa X podlega rozkładowi normalnemu $N(m, \sigma)$.

Obliczyć $P(|X - m| < k\sigma)$, $k > 0$. Podać wartości dla $k = 1, 2, 3$.

Rozwiązanie. Mamy

$$\begin{aligned} P(|X - m| < k\sigma) &= P\left(\left|\frac{X - m}{\sigma}\right| < k\right) = P(|T| < k) = P(-k < T < k) \\ &= \Phi(k) - \Phi(-k) = 2\Phi(k) - 1. \end{aligned}$$

Wykonujemy obliczenia dla $k = 1, 2, 3$:

$$k = 1 : P(|X - m| < \sigma) \approx 0,6827,$$

$$k = 2 : P(|X - m| < 2\sigma) \approx 0,9545,$$

$$k = 3 : P(|X - m| < 3\sigma) \approx 0,9973.$$

Ostatni przypadek ($k = 3$) jest znany pod nazwą reguły trzysigmowej, w myśl której prawdopodobieństwo, że wartość zmiennej losowej o rozkładzie $N(m, \sigma)$ różni się od wartości przeciętnej o mniej niż 3σ jest równe w przybliżeniu 0,9973.

Zadanie 18. Niech $X \sim N(-1, 5; 2)$. Korzystając z tablic rozkładu normalnego $N(0, 1)$ obliczyć prawdopodobieństwa

1. $P[|X| > 0,5]$,
2. $P[X^2 < 4]$,
3. $P[e^X > 1]$.

Rozwiązanie. Mamy:

1. 0,8502,
2. 0,5586,
3. 0.2266.

11 Zmienne losowe wielowymiarowe

11.1 Pojęcie wektora losowego

W przypadku, gdy obserwujemy pewną liczbę zmiennych losowych, wyznaczenie wartości prawdopodobieństwa jednej zmiennej często zależy od wartości pozostałych zmiennych. W ten sposób wprowadzimy pojęcie wektora losowego.

Przykład 11.1. *Rzucamy jednocześnie dwiema symetrycznymi monetami. Niech X oznacza zmienną losową przyjmującą wartość jeden gdy na pierwszej kostce wypadnie orzeł i wartość 0, gdy wypadnie reszka. Analogicznie Y oznacza zmienną losową przyjmującą wartość jeden, gdy na drugiej kostce wypadnie orzeł i wartość 0, gdy wypadnie reszka. Wyznaczmy $P(X = i, Y = j)$, dla $i, j = 0, 1$.*

Mamy

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

Wyznaczone prawdopodobieństwa tworzą tak zwany rozkład łączny wektora losowego (X, Y) . Na podstawie tego rozkładu możemy wyznaczyć rozkłady brzegowe zmiennych X i Y . Mamy bowiem

$$P(X = 0) = P(X = 0, Y = 0) + P(X = 0, Y = 1) = 1/2,$$

oraz

$$P(X = 1) = P(X = 1, Y = 0) + P(X = 1, Y = 1) = 1/2.$$

Analogiczny rezultat otrzymamy dla drugiej zmiennej. Jak widzimy oba rozkłady są rozkładami Bernoulliego z parametrem $p = 1/2$.

Definicja 11.1. Niech X oraz Y będą zmiennymi losowymi, zdefiniowanymi na tej samej przestrzeni probabilistycznej (Ω, \mathcal{A}, P) . Odwzorowanie mierzalne:

$$\Omega \rightarrow \mathbf{R}^2, \quad \omega \rightarrow (X(\omega), Y(\omega))$$

nazywamy dwuwymiarowym wektorem losowym i oznaczamy przez (X, Y) .

Definicja 11.2. Niech X_1, \dots, X_n będą zmiennymi losowymi, zdefiniowanymi na tej samej przestrzeni probabilistycznej (Ω, \mathcal{A}, P) . Odwzorowanie mierzalne:

$$\Omega \rightarrow \mathbf{R}^n, \quad \omega \rightarrow (X_1(\omega), \dots, X_n(\omega))$$

nazywamy n -wymiarowym wektorem losowym i oznaczamy przez (X_1, \dots, X_n) .

Wprowadźmy teraz definicję rozkładu dwuwymiarowego wektora losowego. Rozkład wektora n -wymiarowego definiujemy analogicznie.

Definicja 11.3. Rozkład $P_{X,Y}$ wektora losowego (X, Y) definiujemy następująco:

$$\forall B \subset \mathbf{R}^2, \quad P_{X,Y}(B) = P(\{\omega \in \Omega, (X(\omega), Y(\omega)) \in B\}).$$

Rozkłady P_X oraz P_Y nazywamy rozkładami brzegowymi.

Przypadek dyskretny

Jeżeli (X, Y) jest wektorem losowym, przyjmującym wartości z co najwyżej przeliczalnego zbioru to P_X oraz P_Y wyznaczamy następująco:

$$\forall x_i \in X(\Omega), \quad P(X = x_i) = \sum_{y_j \in Y(\Omega)} P(X = x_i, Y = y_j),$$

$$\forall y_i \in Y(\Omega), \quad P(Y = y_i) = \sum_{x_j \in X(\Omega)} P(X = x_i, Y = y_j).$$

Przypadek ciągły

W tym przypadku zakładamy, że wektory losowe mają funkcję gęstości względem miary Lebesgue'a w R^n . Spełnia ona identyczne warunki jak w przypadku jednowymiarowym.

Definicja 11.4. *Funkcją gęstości prawdopodobieństwa określoną na R^n nazywamy każdą funkcję spełniającą następujące warunki:*

1. *f jest nieujemna na R^n .*
2. *f jest całkowna.*
3. $\int_{R^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1$.

Definicja 11.5. *Niech f będzie gęstością prawdopodobieństwa na R^n . Wektor losowy (X_1, \dots, X_n) ma funkcję gęstości f , jeżeli dla każdego zbioru borelowskiego C :*

$$P(X \in C) = \int_C f(x_1 \dots x_n) dx_1 \dots dx_n.$$

Oczywiście mając gęstość wektora losowego możemy wyznaczyć jego dystrybuantę wzorem:

$$F_{(X_1, \dots, X_n)}(x_1 \dots x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n), \quad (x_1, \dots, x_n) \in R^n.$$

Możemy także wyznaczyć rozkład dowolnej zmiennej losowej wchodzącej w skład wektora losowego. Mówi o tym następujące twierdzenie.

Twierdzenie 11.1. *Niech (X_1, \dots, X_n) będzie absolutnie ciągłym wektorem losowym. Wtedy zmienna losowa X_i jest absolutnie ciągła i jej rozkład prawdopodobieństwa dany jest przez funkcję gęstości postaci:*

$$f_{X_i}(x_i) = \int_{R_{n-1}} f_{(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n.$$

Podobnie obliczamy rozkłady dowolnej ilości zmiennych losowych wchodzących w skład wektora losowego.

11.2 Niezależność zmiennych losowych

Spróbujmy przenieść definicję niezależności ze zdarzeń na zmienne losowe.

Definicja 11.6. *Mówimy, że zmienne losowe X_1, \dots, X_n są niezależne, jeżeli dla dowolnych zbiorów borelowskich $A_1, \dots, A_n \subset \mathbf{R}$, zdarzenia $\{X_1 \in A_1\}, \dots, \{X_n \in A_n\}$ są niezależne.*

Dla $n=2$ otrzymujemy definicję niezależności dla dwuwymiarowego wektora losowego (X, Y) . Ze względu na to, że zbiór $(X_i \leq x_i) = \{\omega : X_i(\omega) \leq x_i\}$ dla $i = 1, \dots, n$ jest zdarzeniem losowym, to definicję niezależności ze zdarzeń można przenieść bezpośrednio na dystrybuanty. Zatem jeżeli zmienne losowe X_1, \dots, X_n są niezależne, to

$$F_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = \prod_{k=1}^n F_{X_k}(x_k), \quad x_k \in R, \quad k = 1, 2, \dots, n.$$

Oczywiście warunek ten możemy zapisać dla zmiennych losowych dyskretnych w postaci:

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

natomiast dla zmiennych losowych ciągłych w postaci:

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n).$$

11.3 Zadania

Zadanie 19. Wektor losowy (X, Y) ma funkcję gęstości

$$f_{XY}(x, y) = \begin{cases} xe^{-x(1+y)}, & x > 0, y > 0, \\ 0, & \text{poza.} \end{cases}$$

Czy zmienne losowe X i Y są niezależne?

Rozwiązanie. Zmienna losowa X ma gęstość

$$f_X(x) = \int_0^\infty xe^{-x(1+y)} dy = -e^{-x(1+y)} \Big|_0^\infty = e^{-x}, \quad x > 0.$$

Zmienna losowa Y ma gęstość

$$f_Y(y) = \int_0^\infty xe^{-x(1+y)} dx = \frac{\Gamma(2)}{(1+y)^2} = \frac{1}{(1+y)^2}, \quad y > 0.$$

Zatem

$$f_X(x)f_Y(y) = \frac{e^{-x}}{(1+y)^2} \neq f_{XY}(x, y),$$

a więc X i Y nie są niezależne.

Zadanie 20. Gęstość prawdopodobieństwa dana jest wzorem

$$f(x, y) = \begin{cases} Cx(x+y) & \text{jeśli } x \geq 0, y \geq 0, x+y \leq 1, \\ 0 & \text{poza.} \end{cases}$$

Znaleźć $P\left[X < \frac{1}{2}\right]$.

Rozwiązanie.

$$\begin{aligned} 1 &= \iint Cx(x+y) dx dy = C \int_0^1 \left[\int_0^{1-y} (x^2 + xy) dx \right] dy = \\ &= C \int_0^1 \left(\frac{x^3}{3} + \frac{x^2}{2} y \Big|_0^{1-y} \right) dy = C \int_0^1 \left[\frac{(1-y)^3}{3} + \frac{(1-y)y}{2} \right] dy = \frac{C}{8} \\ &\implies C = 8 \end{aligned}$$

$$\begin{aligned} P\left[X \leq \frac{1}{2}\right] &= \iint_D 8x(x+y) dx dy = 8 \int_0^{1/2} \left[\int_0^{1-x} (x^2 + xy) dy \right] dx = \\ &= 8 \int_0^{1/2} \left(x^2 y + \frac{xy^2}{2} \Big|_0^{1-x} \right) dx = \\ &= 8 \int_0^{1/2} \left[x^2(1-x) + \frac{1}{2}x(1-x)^2 \right] dx = \frac{7}{16} \end{aligned}$$

12 Funkcje zmiennych losowych

Czasami interesujemy się nie tyle rozkładem zmiennej losowej lecz rozkładem jej funkcji. Na przykład dla dwóch zmiennych losowych X oraz Y chcemy wyznaczyć rozkład sumy tych zmiennych. Na początku zajmiemy się funkcjami jednowymiarowymi zmiennych losowych, a następnie rozważymy przypadek wielowymiarowy.

12.1 Funkcje jednowymiarowych zmiennych losowych

Jeżeli zmienna X ma rozkład ciągły, to $Y = h(X)$ może, lecz nie musi być typu ciągłego. Zajmiemy się sytuacją, kiedy h jest odwzorowaniem gładkim i wyznaczymy rozkład $h(X)$.

Twierdzenie 12.1. *Jeżeli zmienna X ma funkcję gęstości f oraz $X \subset (a, b)$, funkcja $h : (a, b) \rightarrow \mathbb{R}$ jest klasy C^1 oraz $h'(x) \neq 0$ dla $x \in (a, b)$, to zmienna $Y = h(X)$ ma rozkład ciągły o gęstości*

$$g(y) = f(h^{-1}(y)) |(h^{-1}(y))'| I_{g(I)}(y),$$

dla $I = (a, b)$.

Przykład 12.1. *Niech $X \sim N(m, \sigma^2)$. Znaleźć rozkład zmiennej $Y = aX + b$ dla $a \neq 0$. Mamy $h(x) = ax + b$ i dla $u \in \mathbb{R}$*

$$f_Y(u) = f_x\left(\frac{u-b}{a}\right) \cdot \frac{1}{|a|} = \frac{1}{\sqrt{2\pi\sigma|a|}} e^{-\frac{(u-b-am)^2}{2\sigma^2 a^2}}.$$

a zatem $Y \sim N(b + am, a^2\sigma^2)$.

12.2 Przypadek wielowymiarowy

Rozważmy najpierw bardzo ważny szczególny przypadek, gdy X i Y są niezależnymi zmiennymi losowymi oraz

$$g : \mathbb{R}^2 \rightarrow \mathbb{R},$$

gdzie $g(X, Y) = X + Y$. Wyznamy dystrybuantę zmiennej losowej $Z = X + Y$.

Mamy:

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(X + Y \leq z) = P(\omega : X(\omega) + Y(\omega) \leq z) \\ &= P_{(X,Y)}((x, y) : x + y \leq z) = \int \int_{(x,y):x+y \leq z} dP_{(X,Y)}(x, y) = \int \int_{(x,y):x+y \leq z} dF_{(X,Y)}(x, y) \\ &= \int \int_{(x,y):x+y \leq z} dF_X(x) dF_Y(y) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-x} dF_Y(y) \right) dF_X(x) \\ &= \int_{-\infty}^{\infty} F_Y(z - x) dF_X(x). \end{aligned}$$

Całkę

$$\int_{-\infty}^{\infty} F_Y(z - x) dF_X(x)$$

oznaczamy symbolem

$$(F_Y * F_X)(z)$$

i nazywamy splotem dystrybuant Y i X . Oczywiście

$$F_Y * F_X = F_X * F_Y,$$

oraz

$$(F_X * F_Y) * F_Z = F_X * (F_Y * F_Z).$$

W przypadku, gdy dystrybuantom F_X i F_Y odpowiadają funkcje gęstości f_X oraz f_Y możemy mówić o splotcie funkcji gęstości, mianowicie:

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(X + Y \leq z) = \int_{-\infty}^{\infty} F_Y(z - x) dF_X(x) \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-x} f_Y(v) dv \right) dF_X(x) \end{aligned}$$

$$\begin{aligned}
&= \{v = t - x\} = \int_{-\infty}^{\infty} \left(\int_{-\infty}^z f_Y(t - x) dt \right) dF_X(x) \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^z f_Y(t - x) dt \right) f_X(x) dx \\
&= \int_{-\infty}^z \left(\int_{-\infty}^{\infty} f_Y(t - x) f_X(x) dx \right) dt.
\end{aligned}$$

Zatem

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_Y(t - x) f_X(x) dx.$$

Całkę

$$\int_{-\infty}^{\infty} f_Y(t - x) f_X(x) dx$$

nazywamy splotem funkcji gęstości f_X i f_Y i oznaczamy

$$(f_Y * f_X)(t) = (f_X * f_Y)(t).$$

Niech teraz (X_1, \dots, X_n) będzie n -wymiarowym wektorem losowym o wartościach w pewnym podzbiore zbioru R^n .

Twierdzenie 12.2. *Jeżeli X ma rozkład ciągły z gęstością f , S jest otwartym podzbiorem R^n , takim, że $P(X \in S) = 1$, $h : S \rightarrow R^n$ jest wzajemnie jednoznaczny odwzorowaniem klasy C^1 , z jacobianem $J_h \neq 0$, to gęstość zmiennej $Y = h(X)$ ma postać:*

$$g_y(y) = f(h^{-1}(y))(|J_{h^{-1}}(y)|), \quad y \in h(S).$$

12.3 Zadania

Zadanie 21. Łączny rozkład prawdopodobieństwa (X, Y) ma postać

$$f_{X,Y}(x, y) = \begin{cases} 24x(1 - y) & 0 \leq x \leq y, 0 \leq y \leq 1, \\ 0 & \text{poza.} \end{cases}$$

Znaleźć rozkład $U = X + Y$.

Rozwiązanie. Podstawmy $V = Y$. W płaszczyźnie (u, v)

$$v = u - x, \quad 0 \leq x \leq v.$$

Stąd obszar, w którym $f_{U,V}$ jest nieujemna, leży między prostymi $v = u$, $v = u/2$.

$$f_U(u) = \begin{cases} \int_{u/2}^u 24(u-v)(1-v)dv & 0 < u \leq 1, \\ \int_{u/2}^1 24(u-v)(1-v)dv & 1 \leq u \leq 2, \\ 0 & \text{poza.} \end{cases}$$

Po wykonaniu całkowania otrzymujemy

$$f_U(u) = \begin{cases} -2u^3 + 3u^2 & 0 \leq u \leq 1, \\ 2u^3 - 9u^2 + 12u - 4 & 1 \leq u \leq 2, \\ 0 & \text{poza.} \end{cases}$$

13 Charakterystyki zmiennych losowych

13.1 Wartość oczekiwana. Momenty zwykłe

W wielu sytuacjach życiowych jesteśmy zmuszeni do podejmowania ryzyka. W niektórych przypadkach, takich jak na przykład wybór kariery zawodowej czy chociażby kupno samochodu nie jesteśmy w stanie ustalić za pomocą metod matematycznych, czy taka a nie inna decyzja jest słuszna czy też nie. Są jednak przykłady sytuacji, kiedy z pomocą przychodzi nam matematyka, a w szczególności rachunek prawdopodobieństwa. Od razu trzeba zaznaczyć, że podane tutaj reguły oceny opłacalności na podstawie wartości oczekiwanej są jednymi z wielu stosowanych i w niektórych przypadkach można stosować inne, lepsze metody.

Przykład 13.1. *Załóżmy, że bierzemy udział w następującej grze. Rzucamy kostką do gry. Jeżeli wypadnie 1 przegrywamy 2zł, za dwójkę przegrywamy 1 zł, w przypadku trójki wygrywamy 5 zł, czwórki przegrywamy 2 zł a za piątkę i szóstkę nie dostajemy nic. Pytanie jakie możemy sobie postawić, to czy ta gra*

jest na dłuższą metę opłacalna. Korzystając z tego, że kostka jest symetryczna możemy założyć, że po n grach średnio każda liczba oczek wypadła $n/6$ razy. Wtedy łączna wygrana po n grach wyniesie:

$$W(n) = -2 \cdot \frac{n}{6} - \frac{n}{6} + 5 \frac{n}{6} - 2 \cdot \frac{n}{6} + 0 \cdot \frac{n}{6} + 0 \cdot \frac{n}{6} = 0$$

Jak widać trudno jest ustalić opłacalność gry, można powiedzieć, że ta gra jest sprawiedliwa, ponieważ średnia wygrana jest równa zero.

Zauważmy, że w powyższym wzorze pojawiły się wartości wygranej oraz prawdopodobieństwa ich wystąpienia. Na tej podstawie średnią zmiennej losowej możemy zdefiniować jako średnią wartości tej zmiennej, ważoną przez jej prawdopodobieństwa. Takie rozumienie średniej pozwala na wyrażenie wartości oczekiwanej (średniej) dla zmiennej losowej dyskretnej wzorem:

$$EX = \sum_{x \in X(\Omega)} xP(X = x).$$

Ten wzór ma sens tylko wtedy, kiedy prawa strona jest skończona, co tłumaczymy za pomocą następującego warunku na istnienie wartości oczekiwanej:

$$\sum_{x \in X(\Omega)} |x|P(X = x) < \infty.$$

Spróbujmy teraz wyrazić pojęcie wartości oczekiwanej w przypadku zmiennej losowej absolutnie ciągłej, dla której istnieje funkcja gęstości prawdopodobieństwa. Zastępując sumę całką a prawdopodobieństwa funkcją gęstości otrzymujemy:

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

z warunkiem na istnienie wartości oczekiwanej w formie

$$\int_{-\infty}^{\infty} |x|f(x)dx < \infty.$$

W tym momencie należy się krótkie wyjaśnienie. O ile definicja dla zmiennych losowych dyskretnych jest poprawna w ogólnym przypadku i na jej podstawie będziemy udowadniać szereg własności dla wartości oczekiwanej, o tyle

definicja dla zmiennej losowej ciągłej wymaga uściślenia. Już bowiem badanie sumy zmiennych losowych ciągłych może spowodować trudności wynikające chociażby z faktu, że suma dwóch ciągłych zmiennych losowych nie musi być zmienną losową ciągłą. Dlatego uzasadnione jest podanie ogólnej definicji wartości oczekiwanej.

Zauważmy, że dowolna dyskretna zmienna losowa przyjmująca wartości ze zbioru x_1, \dots, x_n może zostać zapisana w postaci

$$X(\omega) = \sum_{k=1}^n I_{A_k}(\omega),$$

gdzie zbiory $A_k = \{\omega : X(\omega) = x_k\}$, dla $k = 1, \dots, n$ tworzą podział przestrzeni Ω . W języku analizy matematycznej powyższą zmienną nazywamy zmienną losową prostą. Wtedy nasza definicja przyjmuje postać:

$$E(X) = \sum_{k=1}^n x_k P(A_k).$$

Wiadomo jednak, że dowolna nieujemna zmienna losowa może zostać przedstawiona jako granica niemalejącego ciągu nieujemnych zmiennych losowych prostych. To prowadzi nas do definicji wartości oczekiwanej dla dowolnej nieujemnej zmiennej losowej.

Definicja 13.1. *Wartością oczekiwaną nieujemnej zmiennej losowej X po zbiorze Ω względem miary probabilistycznej P nazywamy liczbę*

$$E(X) = \lim_{n \rightarrow \infty} E(X_n),$$

gdzie ciąg (X_n) jest ciągiem niemalejącym nieujemnych zmiennych losowych prostych zbieżnych do X .

Teraz już krok od ogólnej definicji wartości oczekiwanej.

Rozważmy dowolną zmienną losową X określoną na przestrzeni (Ω, \mathcal{A}, P) . Zmienną X możemy zapisać następująco:

$$X(\omega) = X^+(\omega) - X^-(\omega),$$

gdzie

$$X^+(\omega) = \begin{cases} X(\omega) & \text{jeżeli } X(\omega) > 0, \\ 0 & \text{poza,} \end{cases}$$

natomiast

$$X^-(\omega) = \begin{cases} -X(\omega) & \text{jeżeli } X(\omega) < 0, \\ 0 & \text{poza.} \end{cases}$$

Definicja 13.2. *Jeżeli X jest dowolną zmienną losową określoną na przestrzeni (Ω, \mathcal{A}, P) , to jej wartość oczekiwaną $E(X)$ względem miary P definiujemy wzorem:*

$$E(X) = E(X^+) - E(X^-),$$

o ile przynajmniej jedna z wielkości występujących po prawej stronie jest skończona.

Od tego momentu wszystkie dowody dotyczące wartości oczekiwanej i innych charakterystyk będą przeprowadzane dla zmiennych losowych prostych. Analogiczne dowody dla zmiennych ciągłych wymagają użycia reprezentacji dowolnej zmiennej za pomocą granicy ciągów zmiennych prostych i zastosowania twierdzeń granicznych.

Twierdzenie 13.1. *{Własności wartości oczekiwanej}*

Wartość oczekiwana $E(X)$ zmiennej losowej X ma następujące własności:

1. Jeżeli $X = I_A$, to $E(X) = P(A)$.
2. Jeżeli $X \geq 0$, to $E(X) \geq 0$.
3. $E(aX + bY) = aE(X) + bE(Y)$.
4. Jeżeli $X \geq Y$, to $E(X) \geq E(Y)$.
5. $|E(X)| \leq E(|X|)$.
6. Jeżeli X i Y są niezależnymi zmiennymi losowymi, to $E(XY) = E(X)E(Y)$.

Udowodnimy dla przykładu własność (6). W dowodach pozostałych własności używamy podobnych argumentów.

Dowód. (6)

Zauważmy, że

$$\begin{aligned} E(XY) &= E\left[\left(\sum_k x_k I_{A_k}\right)\left(\sum_l y_l I_{B_l}\right)\right] \\ &= E\left(\sum_{k,l} x_k \cdot y_l I_{A_k \cap B_l}\right) = \sum_{k,l} x_k \cdot y_l P(A_k \cap B_l) = \\ &= \sum_{k,l} x_k \cdot y_l P(A_k)P(B_l) = \left(\sum_k x_k P(A_k)\right)\left(\sum_l y_l P(B_l)\right) = E(X)E(Y). \end{aligned}$$

Podamy teraz kilka przykładów obliczania wartości oczekiwanej dla zmiennych losowych.

Przykład 13.2. *{Rozkład równomierny}* Niech X ma rozkład równomierny (jednostajny) na zbiorze x_1, \dots, x_n . Wtedy

$$E(X) = \sum_{k=1}^n x_k P(X = x_k) = \frac{1}{n} \sum_{k=1}^n x_k.$$

Jak widzimy wynik jest średnią arytmetyczną wartości zmiennej losowej.

Przykład 13.3. *{Rozkład geometryczny}* Niech X ma rozkład geometryczny z parametrem p . Wtedy

$$\sum_{k=1}^{\infty} kq^{k-1}p = p \sum_{k=1}^{\infty} (k-1+1)q^{k-1} = p \sum_{k=1}^{\infty} (k-1)q^{k-1} + p \sum_{k=1}^{\infty} q^{k-1}.$$

Stąd

$$E(X) = qE(X) + 1 \Rightarrow E(X) = \frac{1}{p}.$$

Uwaga. W tym przypadku nie badaliśmy warunku na istnienie, ponieważ wartości zmiennej losowej są nieujemne.

Przykład 13.4. Niech X ma rozkład prawdopodobieństwa dany wzorem:

$$P\left(X = \frac{(-2)^k}{k}\right) = \frac{1}{2^k}, \quad k = 1, 2, \dots$$

Zauważmy, że

$$\sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} \frac{(-2)^k}{k} \frac{1}{2^k} = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -\ln 2.$$

Pomimo tego wartość oczekiwana nie istnieje, ponieważ

$$\sum_{k=1}^{\infty} |x_k| p_k = \sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

Przykład 13.5. {Rozkład jednostajny} Niech X ma rozkład jednostajny na przedziale (a, b) . Wtedy

$$E(X) = \frac{1}{b-a} \int_a^b x dx = \frac{a+b}{2}.$$

Jak widzimy wynik jest (podobnie jak w przypadku rozkładu równomiernego) średnią arytmetyczną wartości granicznych zmiennej losowej.

Przykład 13.6. {Rozkład normalny} Niech X ma rozkład normalny $N(\mu, \sigma)$. Wtedy

$$E(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Podstawmy $x = \sigma t + \mu$. Wtedy

$$\begin{aligned} E(X) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma t) e^{-\frac{t^2}{2}} dx. \\ &= \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dx + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-\frac{t^2}{2}} dx = \mu. \end{aligned}$$

Jeżeli zmienna losowa Y jest pewną funkcją mierzalną zmiennej losowej X , tzn $Y = g(X)$ i znamy rozkład zmiennej X , to możemy obliczyć wartość oczekiwaną zmiennej Y (o ile ona istnieje) ze wzoru:

$$E(Y) = E(g(X)).$$

Zajmiemy się teraz szczególnymi przypadkami funkcji $g(x)$.

Definicja 13.3. Jeżeli X jest dowolną zmienną losową określoną na przestrzeni (Ω, \mathcal{A}, P) , to momentem zwykłym rzędu k zmiennej losowej X nazywamy liczbę

$$m_k = E(X^k).$$

Liczbę $\beta_k = E(|X|^k)$ nazywamy momentem absolutnym rzędu k .

Twierdzenie 13.2. *Jeżeli moment zwykły rzędu s zmiennej losowej X jest skończony, to wszystkie momenty zwykłe rzędu $r < s$ są skończone.*

Dowód. Dla dowolnego $a \in R$ i dla $r < s$ mamy $|a|^r < 1 + |a|^s$. Zastępując a przez zmienną losową X i stosując własność (4) dla wartości oczekiwanych otrzymujemy tezę.

13.2 Ryzyko w ujęciu inwestycyjnym

13.2.1 Wartość oczekiwana-oczekiwana stopa zwrotu

Celem porównywania inwestycji jest wybranie tej, która da nam lepszy dochód. Nie jesteśmy jednak w stanie przewidzieć jaka będzie cena akcji za tydzień, a co dopiero za rok czy dwa. Pojawia się więc pytanie w jaki sposób porównać dwie inwestycje nie znając przecież ich przyszłych wartości. Zazwyczaj w tym celu możemy skorzystać z przewidywań analityków lub z danych historycznych danej spółki. W celu obliczenia oczekiwanej stopy zwrotu posłużymy się tutaj wzorem dla rozkładu dyskretnego stopy zwrotu:

$$E(R) = \sum_{k=1}^n p_k r_k, \quad (1)$$

gdzie: $E(R)$ - oczekiwana stopa zwrotu, p_k - prawdopodobieństwo, że stopa zwrotu wyniesie r_k .

Dla rozkładu ciągłego mamy natomiast wzór:

$$E(R) = \int_{-\infty}^{+\infty} r f_R(r) dr, \quad (2)$$

gdzie $f_R(r)$ - funkcja gęstości stopy zwrotu.

W przypadku ciągłym, prawdopodobieństwo że stopa zwrotu r będzie z przedziału $[a; b]$ (tzn. $a \leq r \leq b$) obliczamy ze wzoru:

$$P(R \in [a; b]) = \int_a^b f_R(r) dr. \quad (3)$$

Definiujemy również dystrybuantę rozkładu stopy zwrotu wzorem:

$$F_R(x) = P(R \leq x).$$

W przypadku rozkładu ciągłego zależność między funkcją gęstości a dystrybuantą zadana jest wzorem:

$$F_R(r) = \int_{-\infty}^r f_R(x)dx.$$

Przykład 13.7. *Pan Kowalski zakupił jedną akcję firmy KRZAK za cenę 1000zł. Po roku została wypłacona dywidenda w wysokości 30zł na akcję. Po dwóch latach wypłacono dywidendę w wysokości 50zł na akcję. Akcja została sprzedana po trzech latach, po otrzymaniu dywidendy w wysokości 70zł za cenę 1200zł.*

Zakładając brak reinwestowania środków z dywidend, na koniec trzech lat Pan Kowalski zgromadził 1350zł. Zgodnie ze wzorem na stopę zwrotu:

$$r = \left(\frac{1350}{1000}\right)^{\frac{1}{3}} - 1,$$

a więc

$$r = 10,52\%.$$

Przykład 13.8. *Analitycy przewidują, że stopa zwrotu z akcji firmy KRZAK za rok 2011 ma rozkład jak w tabeli poniżej:*

i	p_i	r_i
1	0,1	-0,1
2	0,1	0
3	0,1	0,5
4	0,2	0,1
5	0,3	0,15
6	0,15	0,2
7	0,05	0,3

Znajdź oczekiwaną stopę zwrotu z akcji tej firmy.

Wstawiając dane z tabelki do wzoru (1.5) otrzymujemy:

$$E(R) = 0,1 \cdot (-0,1) + 0,1 \cdot 0 + 0,1 \cdot 0,5 + 0,2 \cdot 0,1 + 0,3 \cdot 0,15 + 0,15 \cdot 0,2 + 0,05 \cdot 0,3,$$

a to daje nam

$$E(R) = 0,15.$$

13.2.2 Miara ryzyka

Intuicyjnie ryzyko rozumiane jest jako możliwość straty. Zatem miara ryzyka musi być proporcjonalna do wielkości straty i prawdopodobieństwa jej wystąpienia.

Definicja 13.4. *Ryzyko inwestycyjne - ryzyko, że zrealizowana stopa zwrotu z inwestycji będzie niższa od oczekiwanej przez inwestora. Miara ryzyka ponadto jest proporcjonalna do wielkości straty i prawdopodobieństwa jej wystąpienia.*

Niestety poniższy przykład dobitnie obrazuje, jak mało informacji taka miara ryzyka nam daje.

Przykład 13.9. *Analitycy przewidują stopy zwrotu z akcji firmy KRZAK i KRZYK na rok 2010, a wyniki analiz przedstawili w tabeli poniżej:*

	KRZAK		KRZYK	
i	p_i	r_i	p_i	r_i
1	0,3	0%	0,3	0%
2	0,4	+10%	0,4	+10%
3	0,25	+12%	0,1	+20%
4	0,05	+80%	0,2	+25%

Jak widać są to całkiem różne wyniki analiz i inwestycja w akcje firmy KRZAK jest inna niż inwestycja w akcje spółki KRZYK. Wartości oczekiwane inwestycji w obie akcje są równe i wynoszą 11%. Zauważmy, że prawdopodobieństwo, że stopa zwrotu będzie niższa niż oczekiwana jest w obu przypadkach równe 70%. Tak samo dla dowolnej stopy zwrotu $r_0 \in (-\infty; 11\%]$ mamy:

$$P[R_{KRZAK} < r_0] = P[R_{KRZYK} < r_0] = \begin{cases} 0\% & \text{dla } r_0 \leq 0\% \\ 30\% & \text{dla } r_0 \leq 10\% \\ 70\% & \text{dla } r_0 \leq 11\% \end{cases} .$$

Zatem, przy takiej definicji zarówno wartość oczekiwana jak i ryzyko były by równe, a więc obie inwestycje można by stwierdzić są jednakowe, co oczywiście jest nieprawdą. Rozważmy troszkę zmodyfikowaną definicję, a mianowicie:

Definicja 13.5. *Ryzyko inwestycyjne - ryzyko, że zrealizowana stopa zwrotu z inwestycji będzie różna (niższa lub wyższa) od oczekiwanej przez inwestora. Miara ryzyka jest proporcjonalna do wielkości odchylenia wartości stopy zwrotu od oczekiwanej, oraz do prawdopodobieństwa ich zajścia.*

W ten sposób rozumiane ryzyko mierzy zarówno negatywne jak i pozytywne scenariusze danego zdarzenia. W dalszej części pracy wprowadzone zostaną miary ryzyka. Do konstrukcji portfeli będziemy używać wariancji, odchylenia standardowego i odchylenia przeciętnego zgodnych z tą definicją ryzyka. Wprowadzone jednak zostaną definicje analogicznych wskaźników ryzyka skonstruowanych zgodnie z pierwszą intuicyjną definicją (tj. ujemna semiwariancja, ujemne semiodchylenie standardowe, oraz ujemne semiodchylenie przeciętne).

13.2.3 Wariancja stopy zwrotu. Ujemna semiwariancja

Wiemy już co to jest ryzyko. Wariancja stopy zwrotu (krótko: wariancja) jest jednym ze sposobów mierzenia ryzyka. Do obliczenia wariancji stosujemy wzór:

$$Var(R) = \sum_{i=1}^n p_i [r_i - E(R)]^2. \quad (4)$$

Łatwo sprawdzić, że wzór ten jest równoważny poniższym wzorom:

$$Var(R) = E((R - E(R))^2),$$

$$Var(R) = E(R^2) - [E(R)]^2.$$

Tak określona wariancja jest nieujemna. Wartość zero przyjmowana jest jedynie dla zdarzenia które przyjmuje jedną wartość z prawdopodobieństwem równym 1. Dodatkowo im większe odchylenie możliwych wartości stopy zwrotu

tym większa wartość wariancji. Z definicji wynika poniższa równość:

$$\text{Var}(\alpha R + \beta) = \alpha^2 \text{Var}(R).$$

Dla niezależnych inwestycji X i Y mamy ponadto:

$$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y).$$

Osoby preferujące pierwszą przytoczoną definicję ryzyka 13.4 korzystają natomiast z analogicznej miary jaką jest ujemna semiwariancja:

$$\text{SVar}^-(R) = \sum_{i=1}^n p_i [(r_i - E(R))^-]^2, \quad (5)$$

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

Wtedy oczywiście zachodzi równość:

$$\text{SVar}^-(R) = E(((R - E(R))^-)^2).$$

13.2.4 Odchylenie standardowe. Ujemne semiodchylenie standardowe

Niestety wariancja jest mało praktyczną miarą ryzyka. Wynika to z tego, że ma inny rząd wielkości niż zmienna losowa. Chęć stworzenia lepszej miary ryzyka doprowadziła do powstania obecnie najpopularniejszego wskaźnika. Odchylenie standardowe stopy zwrotu liczymy ze wzoru:

$$\sigma(R) = \sqrt{\sum_{i=1}^n p_i [r_i - E(R)]^2}. \quad (6)$$

Łatwo zauważyć prostą zależność:

$$\sigma(R) = \sqrt{\text{Var}(R)}.$$

Odchylenie standardowe jest więc unormowaną wariancją wyrażoną w tych samych jednostkach. Dlatego jest z punktu widzenia użytkownika lepszą miarą niż wariancja.

Przykład 13.10. *Analitycy przewidują stopy zwrotu z akcji firmy KRZAK i KRZYK za rok 2011, a wyniki analiz przedstawili w tabeli poniżej:*

	KRZAK		KRZYK	
i	p_i	r_i	p_i	r_i
1	0,1	-15%	0,1	0%
2	0,2	-5%	0,2	2,5%
3	0,4	+5%	0,4	5%
4	0,2	+15%	0,2	+7,5%
5	0,1	+25%	0,1	+10%

Łatwo się przekonać, że wartości oczekiwane stopy zwrotu akcji obu tych spółek są równe 5%. Jednak ich odchylenia standardowe różnią się.

$$\sigma(R_{KRZAK}) = \sqrt{0,1(-20\%)^2 + 0,2(-10\%)^2 + 0,2(10\%)^2 + 0,1(20\%)^2}$$

$$\sigma(R_{KRZAK}) = 10,95\%$$

$$\sigma(R_{KRZYK}) = \sqrt{0,1(-5\%)^2 + 0,2(-2,5\%)^2 + 0,2(2,5\%)^2 + 0,1(5\%)^2}$$

$$\sigma(R_{KRZYK}) = 2,74\%$$

Stąd wniosek, że akcje firmy KRZYK mają mniejsze ryzyko.

Analogicznie, jak w przypadku ujemnej semiwariancji, definiujemy ujemne semiodchylenie standardowe za pomocą wzoru:

$$\sigma^-(R) = \sqrt{\sum_{i=1}^n p_i [(r_i - E(R))^-]^2}. \quad (7)$$

Łatwo zauważyć, że zachodzi analogiczna zależność jak w przypadku odchylenia standardowego:

$$\sigma^-(R) = \sqrt{SVar^-(R)}.$$

13.2.5 Odchylenie przeciętne. Ujemne semiodchylenie przeciętne

Odchylenie przeciętne $d(R)$ jest alternatywną metodą badania ryzyka związanego ze stopą zwrotu. Mierzy ona średnią ważoną odległości stóp zwrotu od wartości oczekiwanej. Stosujemy następujący wzór:

$$d(R) = \sqrt{\sum_{i=1}^n p_i |r_i - E(R)|}. \quad (8)$$

Co jest równoważne:

$$d(R) = E(|R - E(R)|).$$

Ujemne semiodchylenie przeciętne mierzy natomiast wielkość możliwej straty w stosunku do przewidywanej stopy zwrotu

$$sd^-(R) = \sqrt{\sum_{i=1}^n p_i ((r_i - E(R))^-)}, \quad (9)$$

czyli po prostu:

$$sd^-(R) = E((R - E(R))^-).$$

13.3 Wariancja, kowariancja i współczynnik korelacji

13.3.1 Związki między inwestycjami

Analizując ceny akcji dwóch spółek z tego samego sektora rynkowego można zauważyć, że zmiany ich cen są często powiązane różnymi zależnościami. Ich genezy nie są jednoznaczne. Najczęściej wynikają one z kilku powodów:

- wzrost/spadek dochodowości danego sektora spowodowany preferencjami konsumentckimi lub zmianą prawa
- zmiana cen towarów substytucyjnych
- ogólny kryzys w branży.

Dla zobrazowania zależności rozważmy następujący przykład:

Przykład 13.11. *Jeżeli założymy, że każdy obywatel niezależnie od swojej sytuacji kupuje dziennie jeden bochenek chleba, a do wyboru ma chleb z pieczywa jasnego lub ciemnego, to wzrost sprzedaży pieczywa jasnego towarzyszy jednakowemu spadkowi sprzedaży pieczywa ciemnego.*

Jednak nie często powiązania są aż tak silne, często wręcz trudno o doszukanie się jakiegokolwiek zależności między cenami. Na przykład ciężko jest powiązać zmianę cen samochodów luksusowych ze zmianą cen papierosów, czy usług fryzjerskich.

Do mierzenia zależności zmian stóp zwrotu stworzono kowariancję.

Jak mówiliśmy wcześniej, średnia jest zaliczana do parametrów położenia zmiennej losowej. W wielu sytuacjach jest ona miarą niewystarczającą do oceny danego eksperymentu i często w takich wypadkach posługujemy się tzw. miarami rozproszenia. Podstawową miarą rozproszenia jest wariancja, której definicję związaną ze stopą zwrotu przytoczyliśmy wcześniej. Formalnie wariancję definiujemy następująco.

Definicja 13.6. *Liczbę $Var(X) = E(X - E(X))^2$ nazywamy wariancją zmiennej losowej X , jeżeli wartość po prawej stronie istnieje. Liczbę $\sigma(X) = \sqrt{Var(X)}$ nazywamy odchyleniem standardowym zmiennej X .*

Zauważmy, że $Var(X) = E(X - E(X))^2 = E(X^2 - 2XE(X) + (E(X))^2) = E(X)^2 - (E(X))^2$.

Wzór ten nosi nazwę formuły Koeniga.

Twierdzenie 13.3. *Dla dowolnych liczb a oraz b zachodzi wzór:*

$$Var(aX + b) = a^2Var(X).$$

Dowód pozostawiamy jako ćwiczenie.

Przykład 13.12. *{Rozkład równomierny}* Niech X ma rozkład równomierny (jednostajny) na zbiorz $1, \dots, n$. Jak pokazano wcześniej $E(X) = \frac{n+1}{2}$. Ponadto

$$E(X^2) = \frac{1}{n} \sum_{k=1}^n k^2 = \frac{(n+1)(2n+1)}{6}.$$

Stosując formułę Koeniga mamy $Var(X) = \frac{(n+1)(n-1)}{12}$.

Wariancja jest szczególnym przypadkiem momentu centralnego.

Definicja 13.7. *Dla każdego k naturalnego, liczbę*

$$\mu_k = E(X - E(X))^k$$

nazywamy momentem centralnym rzędu k zmiennej losowej X .

W poprzednim paragrafie udowodniliśmy dwie równości. Po pierwsze wartość oczekiwana sumy jest równa sumie wartości oczekiwanych a po drugie dla niezależnych zmiennych losowych wartość oczekiwana iloczynu jest równa iloczynowi wartości oczekiwanych. Zastanówmy się czy wariancja sumy zmiennych losowych jest równa sumie ich wariancji. Mamy

$$\begin{aligned} Var(X + Y) &= E[(X - E(X)) + (Y - E(Y))]^2 \\ &= Var(X) + Var(Y) + 2E[(X - E(X))(Y - E(Y))]. \end{aligned}$$

Widzimy zatem, że w ogólnym przypadku

$$Var(X + Y) \neq Var(X) + Var(Y).$$

Czynnik $E[(X - E(X))(Y - E(Y))]$, który decyduje o tym, że równość nie zachodzi nosi nazwę kowariancji zmiennych losowych.

Definicja 13.8. *Niech zmienne X oraz Y mają momenty drugiego rzędu. Kowariancję tych zmiennych definiujemy wzorem:*

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Podobnie jak w przypadku wariancji, kowariancję możemy zapisać za pomocą formuły Koeniga. Zauważmy, że $Cov(X, Y) = E(XY) - E(X)E(Y)$. Poniżej podamy własności kowariancji. Wcześniej jednak przypomnijmy znaną z analizy a stosowaną także dla zmiennych losowych nierówność Cauchy'ego-Schwarza.

Twierdzenie 13.4. *Niech zmienne X oraz Y mają momenty drugiego rzędu. wtedy*

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

Dowód. Zauważmy, że dla każdego $t \in R$

$$t^2 E(Y)^2 + 2tE(XY) + E(X)^2 = E(X + tY)^2 \geq 0.$$

Stąd

$$\Delta = 4(E(XY))^2 - 4E(X)^2 E(Y)^2 \leq 0$$

co daje tezę.

Twierdzenie 13.5. *{Własności kowariancji} Dla dowolnej pary zmiennych losowych mających momenty rzędu drugiego:*

1. $Cov(X, Y) = Cov(Y, X)$.
2. Dla dowolnych, rzeczywistych a, b, c, d : $Cov(aX+b, cY+d) = acCov(X, Y)$.
3. $|Cov(X, Y)| \leq \sigma(X)\sigma(Y)$.
4. Jeżeli X i Y są niezależne, to $Cov(X, Y) = 0$.

Dowód. Własność (1) oraz (2) wynika bezpośrednio z definicji i własności wartości oczekiwanej. Własność (3) to konsekwencja nierówności Cauchy'ego-Schwarza, zaś własność (4) wynika z formuły Koeniga i została już udowodniona wcześniej.

Definicja 13.9. *{Współczynnik korelacji}* Niech zmienne X oraz Y mają momenty drugiego rzędu. Współczynnik korelacji tych zmiennych definiujemy wzorem:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

Zauważmy, że z nierówności Cauchy'ego-Szwarza wynika, że

$$|\rho(X, Y)| \leq 1.$$

13.4 Zadania

Zadanie 22. Zmienna losowa X ma rozkład prawdopodobieństwa dany wzorem

$$P(X = x) = p_X(x) = \begin{cases} \frac{1}{x(x+1)} & \text{jeśli } x = 1, 2, \dots, \\ 0 & \text{poza.} \end{cases}$$

Dla jakich wartości α istnieje EX^α ?

Rozwiązanie.

$$EX^\alpha = \sum_{x=1}^{\infty} x^\alpha \frac{1}{x(x+1)} < \infty, \quad \text{gdy } \alpha < 1.$$

Zadanie 23. Eksperyment polega na rzucie 3 razy monetą. Zakładamy, że prawdopodobieństwo wyrzucenia orła równa się $2/3$ a reszki $1/3$. Niech zmienna losowa X oznacza liczbę otrzymanych orłów. Wyznaczyć rozkład zmiennej losowej X , jej dystrybuantę oraz EX i $\sigma^2 X$.

Rozwiązanie.

ω	$X = X(\omega)$	$P(X) = P[X(\omega) = x]$	$xp(x)$	$x^2p(x)$
OOO	3	8/27	24/27	72/27
OOR	2	4/27	8/27	16/27
ORO	2	4/27	8/27	16/27
ROO	2	4/27	8/27	16/27
ORR	1	2/27	2/27	2/27
ROR	1	2/27	2/27	2/27
RRO	1	2/27	2/27	2/27
RRR	0	1/27	0	0

Stąd otrzymujemy:

$$F(x) = P[X \leq x] = \begin{cases} 0 & x \leq 0; \\ 1/27 & 0 < x \leq 1; \\ 7/27 & 1 < x \leq 2; \\ 19/27 & 2 < x \leq 3; \\ 1 & x > 3. \end{cases}$$

$$EX = 2.$$

$$EX^2 = \frac{14}{3}.$$

$$\sigma^2 X = EX^2 - E^2 X = \frac{14}{3} - 4 = \frac{2}{3}.$$

Zadanie 24. Załóżmy, że X ma gęstość $f(x) = ax + bx^3$ dla $0 \leq x \leq 2$ i $f(x) = 0$ poza. Znaleźć a i b , jeśli:

(a) $EX = 1$

(b) $EX^3 = 4$.

Rozwiązanie.

$$\int_0^2 (ax + bx^3) dx = a \frac{x^2}{2} \Big|_0^2 + b \frac{x^4}{4} \Big|_0^2 = 2a + 4b = 1$$

(a)

$$EX = \int_0^2 (ax^2 + bx^4) dx = a \frac{x^3}{3} \Big|_0^2 + b \frac{x^5}{5} \Big|_0^2 = \frac{8}{3}a + \frac{32}{5}b = 1$$

$$\begin{cases} 2a + 4b = 1 & | \cdot (-20) \\ 40a + 96b = 15 \end{cases}$$

$$\begin{cases} -40a - 80b = -20 \\ 40a + 96b = 15 \end{cases}$$

$$16b = -5 \Rightarrow b = -\frac{5}{16},$$

$$2a = 1 - 4 \cdot \left(-\frac{5}{16}\right) = 1 + \frac{5}{4} = \frac{9}{4} \Rightarrow a = \frac{9}{8}$$

$$\begin{cases} a = \frac{9}{8} \\ b = -\frac{5}{16} \end{cases}$$

(b)

$$\begin{aligned} EX^3 &= \int_0^2 x^3(ax + bx^3) dx = \int_0^2 (ax^4 + bx^6) dx = a \frac{x^5}{5} \Big|_0^2 + b \frac{x^7}{7} \Big|_0^2 = \\ &= \frac{32}{5}a + \frac{128}{7}b = 4 \end{aligned}$$

$$\begin{cases} 2a + 4b = 1 \\ \frac{32}{5}a + \frac{128}{7}b = 4 \end{cases}$$

$$\begin{cases} 2a + 4b = 1 & \implies 2a = 1 - 4b, \quad a = \frac{1-4b}{2} \\ 224a + 640b = 140 \end{cases}$$

$$224 \cdot \frac{1-4b}{2} + 640b = 140$$

$$112 - 448b + 640b = 140$$

$$192b = 28$$

$$b = \frac{28}{192} = \frac{7}{48}$$

$$a = \frac{1}{2} \left(1 - 4 \cdot \frac{7}{48} \right) = \frac{1}{2} \left(1 - \frac{7}{12} \right) = \frac{5}{24}$$

$$\begin{cases} a = \frac{5}{24} \\ b = \frac{7}{48} \end{cases}$$

Zadanie 25. Czas T rozmowy telefonicznej jest zmienną losową spełniającą warunek

$$P(T > t) = pe^{-\lambda t} + (1-p)e^{-\mu t}, \quad t > 0,$$

gdzie $p \in (0, 1)$ oraz $\lambda, \mu > 0$. Obliczyć ET i $\text{Var } T$.

Rozwiązanie. Mamy

$$ET^k = \int_0^{\infty} t^{k-1} P(T > t) dt.$$

Stąd

$$\begin{aligned} ET &= \int_0^{\infty} P(T > t) dt \\ &= \int_0^{\infty} (pe^{-\lambda t} + (1-p)e^{-\mu t}) dt \\ &= \frac{p}{\lambda} + \frac{1-p}{\mu}. \end{aligned}$$

Dalej podobnie

$$\begin{aligned} ET^2 &= \int_0^\infty tP(T > t)dt \\ &= \frac{2p}{\lambda^2} + \frac{2(1-p)}{\mu^2}. \end{aligned}$$

Stąd

$$\begin{aligned} \text{Var}T &= ET^2 - (ET)^2 \\ &= 2p(1-p) \left(\frac{1}{\lambda^2} + \frac{1}{\mu^2} + \frac{1}{\lambda\mu} \right). \end{aligned}$$

14 Funkcja tworząca prawdopodobieństwa i funkcja charakterystyczna

14.1 Funkcja tworząca prawdopodobieństwa

Jak wspominaliśmy wcześniej istnieje wiele narzędzi do badania własności zmiennych losowych. Jednym z nich jest funkcja tworząca prawdopodobieństwa. Zanim omówimy jej własności należy wspomnieć, że funkcję tę stosujemy jedynie w przypadku tzw. zmiennych losowych arytmetycznych czyli takich, które przyjmują nieujemne wartości całkowite. Funkcje służące do badania dowolnych zmiennych losowych to funkcja tworząca momenty oraz omówiona w następnym paragrafie funkcja charakterystyczna.

Definicja 14.1. Niech X będzie zmienną losową przyjmującą wartości w zbiorze $\{0, 1, 2, \dots\}$. Funkcję tworzącą prawdopodobieństwa definiujemy wzorem:

$$G_X(s) = E(s^X) = \sum_{n=0}^{\infty} s^n P(X = n) = \sum_{n=0}^{\infty} s^n p_n.$$

Zauważmy, że $G_X(0) = 0$. Ponadto funkcja tworząca prawdopodobieństwa jest określona co najmniej na zbiorze $|s| \leq 1$, ponieważ

$$\sum_{n=0}^{\infty} |s|^n p_n \leq \sum_{n=0}^{\infty} p_n = 1 = G_X(1).$$

Tak więc na mocy własności szeregu potęgowego dla $|s| < 1$ funkcję $G_X(s)$ można różniczkować pod znakiem sumy i otrzymujemy:

$$G'_X(s) = \sum_{n=1}^{\infty} n s^{n-1} p_n, \quad G''_X(s) = \sum_{n=2}^{\infty} n(n-1) s^{n-2} p_n, \quad (10)$$

oraz

$$G_X^{(k)}(s) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} s^{n-k} p_n.$$

Zatem dla $s = 0$ otrzymujemy:

$$p_n = \frac{G_X^{(n)}(0)}{n!}.$$

Prowadzi to do następującego twierdzenia.

Twierdzenie 14.1. *Rozkład prawdopodobieństwa zmiennej losowej o nieujemnych wartościach całkowitych jest jednoznacznie wyznaczony przez funkcję tworzącą prawdopodobieństwa.*

Aby wyprowadzić dalsze własności funkcji $G_X(s)$ przypomnijmy znane z analizy twierdzenie Abela.

Twierdzenie 14.2. *{Twierdzenie Abela o szeregach potęgowych}*

Niech $(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Jeżeli szereg $\sum_{n=0}^{\infty} a_n x^n$ jest zbieżny oraz $f : (-1, 1) \rightarrow \mathbb{R}$ jest dana wzorem $f(x) = \sum_{n=0}^{\infty} a_n x^n$ to wówczas $\sum_{n=0}^{\infty} a_n = \lim_{x \rightarrow 1^-} f(x)$.

Powróćmy teraz do wzoru(10). Jeżeli EX jest skończone, to spełnione jest założenie twierdzenia Abela, tzn

$$\lim_{s \rightarrow 1^-} G'_X(s) = \sum_{n=1}^{\infty} n p_n = EX.$$

Otrzymujemy zatem

$$EX = G'_X(s), \quad \text{oraz} \quad EX(X-1) = G''_X(1).$$

Przykład 14.1. Niech $X \sim B(n, p)$. Wtedy

$$G_X(s) = E(s^X) = \sum_{k=0}^n s^k \binom{n}{k} p^k (1-p)^{n-k} = (ps + q)^n.$$

Przykład 14.2. Niech $X \sim G(p)$. Wtedy

$$G_X(s) = E(s^X) = \sum_{k=0}^{\infty} s^k p(1-p)^{k-1} = \frac{ps}{1-qs}.$$

Przykład 14.3. Niech $X \sim P(\lambda)$. Wtedy

$$G_X(s) = E(s^X) = \sum_{k=1}^{\infty} s^k e^{-\lambda} \frac{\lambda^k}{k!} = e^{\lambda(s-1)}.$$

Konstrukcja funkcji tworzącej prawdopodobieństwa pozwala badać własności sum niezależnych zmiennych losowych. Mówi o tym następujące twierdzenie.

Twierdzenie 14.3. Jeżeli X_1, \dots, X_n są niezależnymi zmiennymi losowymi o funkcjach tworzących G_1, \dots, G_n odpowiednio, to suma $X_1 + \dots + X_n$ ma funkcję tworzącą postaci

$$G_{X_1 + \dots + X_n}(s) = \prod_{k=1}^n G_k(s).$$

Dowód. Z niezależności zmiennych X_1, \dots, X_n otrzymujemy, że zmienne s^{X_1}, \dots, s^{X_n} są niezależne i w konsekwencji

$$G_{X_1 + \dots + X_n}(s) = E s^{X_1 + \dots + X_n} = \prod_{k=1}^n E s^{X_k}.$$

Przykład 14.4. Z postaci funkcji tworzącej prawdopodobieństwa dla rozkładu Poissona widzimy, że jeżeli zmienne X oraz Y są niezależnymi zmiennymi losowymi o rozkładach Poissona takimi, że $X \sim P(\lambda)$ a $Y \sim P(\mu)$, to

$$G_{X+Y}(s) = e^{\lambda(s-1)} e^{\mu(s-1)} = e^{(\lambda+\mu)(s-1)},$$

tzn. suma $X + Y \sim P(\lambda + \mu)$.

14.2 Funkcja charakterystyczna

W poprzednim rozdziale badaliśmy własności funkcji tworzącej prawdopodobieństwa. Jej stosowanie ograniczone było jedynie do zmiennych losowych przyjmujących nieujemne wartości całkowite. Powstaje zatem pytanie, czy istnieje aparat matematyczny służący do analizy dowolnych zmiennych losowych? Okazuje się, że takim narzędziem jest funkcja charakterystyczna.

Definicja 14.2. Niech (Ω, \mathcal{A}, P) będzie przestrzenią probabilistyczną. Funkcję charakterystyczną miary P definiujemy wzorem

$$\phi(t) = \int_{\mathbb{R}} e^{itx} dP(x), \quad t \in \mathbb{R}.$$

Jeżeli X jest zmienną losową to wiemy, że generuje ona miarę P_X . Wtedy funkcję charakterystyczną zmiennej losowej X definiujemy przez

$$\phi_X(t) = \int_{\mathbb{R}} e^{itx} dP_X(x), \quad t \in \mathbb{R}.$$

Zauważmy, że dla dowolnego $t \in \mathbb{R}$ funkcja charakterystyczna jest dobrze określona, ponieważ szereg

$$\phi_X(t) = \int_{\mathbb{R}} e^{itx} dP_X(x),$$

jest bezwzględnie zbieżny. Rzeczywiście

$$\int_{\mathbb{R}} |e^{itx}| dP_X(x) \leq \int_{\mathbb{R}} dP_X(x) = 1.$$

Zauważmy, że jeżeli X jest zmienną losową dyskretną, to

$$\phi_X(t) = \sum_{x_k \in X} p_k e^{itx_k}, \quad t \in \mathbb{R},$$

natomiast jeżeli X jest zmienną losową absolutnie ciągłą, to

$$\phi_X(t) = \int_{\mathbb{R}} e^{itx} f(x) dx, \quad t \in \mathbb{R}.$$

Dalsze własności funkcji charakterystycznej.

Twierdzenie 14.4. Dla dowolnych liczb rzeczywistych a i b i zmiennej losowej X o wartościach rzeczywistych

$$\phi_{aX+b}(t) = e^{itb} \phi_X(at), \quad t \in \mathbb{R}.$$

Dowód.

$$\phi_{aX+b}(t) = E e^{it(aX+b)} = e^{itb} E e^{itaX} = e^{itb} \phi_X(at), \quad t \in \mathbb{R}.$$

Twierdzenie 14.5. *Jeżeli X_1, \dots, X_n będą niezależnymi zmiennymi losowymi. Wtedy suma $X_1 + \dots + X_n$ ma funkcję charakterystyczną postaci*

$$\phi_{X_1+\dots+X_n}(t) = \prod_{k=1}^n \phi_{X_k}(t), \quad t \in R.$$

Dowód analogiczny jak w przypadku funkcji tworzących prawdopodobieństwa.

Twierdzenie 14.6. *Jeżeli zmienna losowa X posiada moment rzędu $n \in N$, to ϕ_X jest funkcją klasy C^n oraz*

$$E(X^n) = \frac{\phi_X^{(n)}(0)}{i^n}.$$

Dowód. Dla $n = 1$, mamy dla każdego $t \in R$, $\frac{d}{dx}e^{itx} = ixe^{itx}$. Ponadto $|ixe^{itx}| \leq |x|$ a $E|X| < \infty$ na mocy założenia. Z twierdzenia Lebesgue' a o zbieżności zmajoryzowanej $\phi_X(t)$ jest różniczkowalna i

$$\phi_X'(t) = \int_R ixe^{itx} dP_X(x) = E(iXe^{itX}).$$

Dla $t = 0$ otrzymujemy tezę. Analogicznie dowodzimy równości dla dowolnego $n \in N$.

Przykład 14.5. *Znaleźć funkcję charakterystyczną dla zmiennej losowej o rozkładzie Poissona z parametrem λ .*

Mamy

$$\phi_X(t) = \sum_{k=0}^{\infty} e^{itk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{\lambda(e^{it}-1)}.$$

Przykład 14.6. *Znaleźć funkcję charakterystyczną dla zmiennej losowej o rozkładzie normalnym $N(0, 1)$. Zauważmy, że*

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{\infty} \cos(tx) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Zatem

$$\begin{aligned}\phi'_X(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(-\sin(tx))e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sin(tx)e^{-\frac{x^2}{2}} d\left(-\frac{x^2}{2}\right) \\ &= \frac{1}{\sqrt{2\pi}} \left(\sin(tx)e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} t \cos(tx) \cdot e^{-\frac{x^2}{2}} dx \right) = -t\phi_X(t).\end{aligned}$$

Stąd

$$\frac{\phi'_X(t)}{\phi_X(t)} = -t,$$

i w konsekwencji

$$\phi_X(t) = c \cdot e^{-\frac{t^2}{2}}.$$

Dla $t = 0$ mamy $c = 1$ co daje

$$\phi_X(t) = e^{-\frac{t^2}{2}}.$$

14.3 Zadania

Zadanie 26. Zmienna losowa X ma rozkład prawdopodobieństwa postaci

$$P[X = k] = \frac{a^k}{(1+a)^{k+1}}, \quad k = 1, 2, \dots, a > 0.$$

Obliczyć EX i $\sigma^2 X$ korzystając z własności funkcji tworzącej prawdopodobieństwa.

Rozwiązanie.

$$G_X(s) = \sum_{k=1}^{\infty} s^k \frac{a^k}{(1+a)^{k+1}} = \frac{1}{1+a} \sum_{k=1}^{\infty} \left(\frac{as}{1+a}\right)^k = \frac{1}{1+a} \cdot \frac{as}{1+a-as},$$

$$EX = a, \quad \sigma^2 X = a(1+a).$$

Zadanie 27. Wyznaczyć funkcję charakterystyczną zmiennej X o rozkładzie zadanym gęstością

$$f(x) = \begin{cases} \frac{1}{6}, & x \in (-2, -1) \cup (1, 2); \\ \frac{1}{3}, & x \in \langle -1, 1 \rangle; \\ 0, & \text{poza.} \end{cases}$$

Następnie korzystając z postaci funkcji charakterystycznej wyznacz EX .

Rozwiązanie. Wyznaczamy najpierw funkcję charakterystyczną

$$\begin{aligned}\varphi(t) &= Ee^{itX} = \int_{-2}^2 e^{itx} \frac{1}{6} dx + \int_{-1}^1 e^{itx} \frac{1}{6} dx \\ &= \frac{1}{3t} (\sin 2t + \sin t).\end{aligned}$$

Teraz $EX = \frac{\varphi'(0)}{i}$.

$$\varphi'(t) = \frac{1}{3t^2} ((2 \cos 2t + \cos t)t - (\sin 2t + \sin t)).$$

Teraz obliczmy granicę tego wyrażenia przy $t \rightarrow 0$:

$$\lim_{t \rightarrow 0} \varphi'(t) \stackrel{H}{=} \lim_{t \rightarrow 0} \frac{1}{6t} (4 \sin 2t + \sin t) t = 0.$$

Stąd ostatecznie $EX = 0$.

Zadanie 28. Zmienna losowa ma rozkład prawdopodobieństwa zadany równością

$$P \left[X_n = \frac{k\pi}{n} \right] = \frac{1}{n}, \quad k = 1, 2, \dots, n; \quad n = 1, 2, \dots$$

Znaleźć funkcję charakterystyczną tej zmiennej losowej oraz obliczyć $\lim_{n \rightarrow \infty} \varphi_{X_n}(t)$.

Rozwiązanie.

$$\begin{aligned}\varphi_{X_n}(t) &= Ee^{itX} = \sum_{k=1}^n e^{it \frac{k\pi}{n}} \frac{1}{n} = \frac{1}{n} \sum_{k=1}^n \left(e^{\frac{it\pi}{n}} \right)^k = \frac{1}{n} e^{\frac{it\pi}{n}} \frac{1 - \left(e^{\frac{it\pi}{n}} \right)^n}{e^{\frac{it\pi}{n}} - 1} \\ &= \frac{1}{n} \frac{1 - e^{it\pi}}{e^{-\frac{it\pi}{n}} - 1}.\end{aligned}$$

Aby znaleźć $\lim_{n \rightarrow \infty} \varphi_{X_n}(t)$ wystarczy obliczyć

$$\lim_{n \rightarrow \infty} n \left(e^{-\frac{it\pi}{n}} - 1 \right)$$

stosując regułę de L'Hospitala:

$$\lim_{n \rightarrow \infty} n \left(e^{-\frac{it\pi}{n}} - 1 \right) = \lim_{n \rightarrow \infty} \frac{e^{-\frac{it\pi}{n}} - 1}{\frac{1}{n}} = \lim_{n \rightarrow \infty} \frac{e^{-\frac{it\pi}{n}} (-it\pi) \left(-\frac{1}{n^2} \right)}{\left(-\frac{1}{n^2} \right)} = -it\pi.$$

Stąd

$$\lim_{n \rightarrow \infty} \varphi_{X_n}(t) = \frac{e^{it\pi} - 1}{it\pi}.$$

15 Rodzaje zbieżności ciągów zmiennych losowych

15.1 Podstawowe rodzaje zbieżności

Przedmiotem tego paragrafu jest badanie granicznego zachowania się ciągów zmiennych losowych. Podamy najbardziej znane w teorii prawdopodobieństwa rodzaje zbieżności ciągów zmiennych losowych. Następnie przedstawimy związki między tymi zbieżnościami ilustrując je przykładami. Pierwszym rodzajem zbieżności jest zbieżność prawie pewna.

Niech X_1, X_2, \dots , oraz X będą zmiennymi losowymi określonymi na tej samej przestrzeni probabilistycznej (Ω, \mathcal{A}, P) .

Definicja 15.1. Ciąg (X_n) jest zbieżny do zmiennej losowej X z prawdopodobieństwem jeden (prawie pewnie), jeżeli

$$P(\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1.$$

Jeżeli zachodzi powyższa zbieżność, to oznaczamy ją symbolem $X_n \xrightarrow{p.w.} X$. Do badania związku pomiędzy zbieżnością z prawdopodobieństwem jeden a następnym rodzajem zbieżności jaki zdefiniujemy pomocne jest następujące twierdzenie.

Twierdzenie 15.1. Ciąg (X_n) jest zbieżny do zmiennej losowej X z prawdopodobieństwem jeden (prawie pewnie) wtedy i tylko wtedy, gdy dla każdego $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P(\{\omega : \sup_{k \geq n} |X_k(\omega) - X(\omega)| > \epsilon\}) = 0. \quad (11)$$

Przejdziemy teraz do zdefiniowania znacznie słabszej zbieżności: zbieżności według prawdopodobieństwa.

Definicja 15.2. Ciąg (X_n) jest zbieżny do zmiennej losowej X według prawdopodobieństwa $(X_n \xrightarrow{p} X)$, jeżeli

$$\forall \epsilon > 0 \quad \lim_{n \rightarrow \infty} P(\{\omega : |X_k(\omega) - X(\omega)| \geq \epsilon\}) = 0.$$

Zauważmy, że z powyższej definicji oraz z twierdzenia 15.1 wynika, że jeżeli ciąg jest zbieżny prawie pewnie, to jest zbieżny według prawdopodobieństwa.

Przed zdefiniowaniem następnego typu zbieżności założymy, że dla $0 < p < \infty$, $E(|X_n|^p) < \infty$, $n = 1, 2, \dots$ oraz $E(|X|^p) < \infty$.

Definicja 15.3. Ciąg (X_n) jest zbieżny do zmiennej losowej X według p -tej średniej $(X_n \xrightarrow{L^p} X)$, jeżeli

$$\lim_{n \rightarrow \infty} E(|X_n - X|^p) = 0, \quad 0 < p < \infty.$$

Twierdzenie 15.2. Zachodzą następujące implikacje:

$$X_n \xrightarrow{p.w.} \implies X_n \xrightarrow{p} X.$$

$$X_n \xrightarrow{L^p} \implies X_n \xrightarrow{p} X.$$

Dowód. Pierwsza część została uzasadniona, natomiast część druga wynika z nierówności Makowa, która zostanie udowodniona w następnym paragrafie.

15.2 Zadania

Zadanie 29. Niech (X_n) będzie ciągiem niezależnych zmiennych losowych o gęstości

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

Określmy $Z_n = \min(X_1, \dots, X_n)$. Wykazać, że

$$Z_n \xrightarrow{P} a, \quad n \rightarrow \infty.$$

Rozwiązanie.

$$\begin{aligned} F_{Z_n}(z) &= P[Z_n < z] = P[\min(X_1, \dots, X_n) < z] = \\ &= 1 - P[\min(X_1, \dots, X_n) > z] = 1 - e^{-n(z-a)} \quad \text{dla } z \geq a \\ F_{Z_n}(z) &= 0 \quad \text{dla } z < a. \end{aligned}$$

Stąd

$$\begin{aligned} P[|Z_n - a| \geq \varepsilon] &= P[Z_n < a - \varepsilon] + P[Z_n > a + \varepsilon] = \\ &= e^{-n(a+\varepsilon-a)} = e^{-n\varepsilon} \rightarrow 0, \quad n \rightarrow \infty. \end{aligned}$$

Zadanie 30. Niech (X_n) będzie ciągiem niezależnych zmiennych losowych takich, że

$$X_n = \begin{cases} n^{1/2r} & \text{z prawdopodobieństwem } 1/n \\ 0 & \text{z prawdopodobieństwem } 1 - 1/n \end{cases}, \quad r > 0.$$

Zbadać zbieżność (według średniej, według prawdopodobieństwa, prawie pewną) tego ciągu do 0.

Rozwiązanie.

$$E|X_n - 0|^r = (n^{1/2r})^r \frac{1}{n} = \frac{1}{\sqrt{n}} \rightarrow 0,$$

czyli $X_n \xrightarrow{L^r} 0$, a więc również $X_n \xrightarrow{P} 0$.

$$\begin{aligned} P \left[\bigcup_{k=n}^{\infty} [|X_k - 0| \geq \varepsilon] \right] &= 1 - P \left[\bigcap_{k=n}^{\infty} |X_k| < \varepsilon \right] \\ 1 - P \left[\bigcap_{k=n}^{\infty} [X_k = 0] \right] &= 1 - \lim_{N \rightarrow \infty} P \left[\bigcap_{k=n}^N [X_k = 0] \right] \\ &= 1 - \lim_{N \rightarrow \infty} \prod_{k=n}^N \left(1 - \frac{1}{k} \right) = 1, \end{aligned}$$

bo

$$\begin{aligned} \lim_{N \rightarrow \infty} \prod_{k=n}^N \left(1 - \frac{1}{k} \right) &= 0 \\ \left(\prod_{n=2}^{\infty} \left(1 - \frac{1}{n} \right) = 0 \quad \text{z teorii szeregów} \right). \end{aligned}$$

Zatem

$$X_n \xrightarrow{p.p.} 0, \quad n \rightarrow \infty.$$

16 Prawa wielkich liczb

16.1 Słabe prawa wielkich liczb

W tym paragrafie podamy pewną interpretację wartości oczekiwanej zmiennej losowej. Załóżmy, że X_1, \dots, X_n są zmiennymi losowymi o tej samej wartości oczekiwanej ($EX_i = \mu$). Wiadomo z poprzedniego paragrafu, że wartość oczekiwana średniej $\frac{X_1 + \dots + X_n}{n}$ jest równa μ . Chcielibyśmy wiedzieć więcej: czy wartości tej średniej są z dużym prawdopodobieństwem bliskie μ a jeżeli tak to jakie warunki muszą być spełnione. O tym właśnie mówi prawo wielkich liczb. Zajmiemy się na początku słabymi prawami wielkich liczb.

Niech (X_n) będzie takim ciągiem zmiennych losowych, że dla każdej zmiennej losowej X_n istnieje wartość oczekiwana $m_n = E(X_n)$.

Definicja 16.1. *Jeżeli*

$$\frac{1}{n} \sum_{k=1}^n (X_k - m_k) \xrightarrow{p} 0,$$

to mówimy, że ciąg (X_n) spełnia słabe prawo wielkich liczb (SPWL). Jeżeli zachodzi zbieżność z prawdopodobieństwem jeden to mówimy, że ciąg (X_n) spełnia mocne prawo wielkich liczb (MPWL).

W dowodach związanych z SPWL wykorzystamy nierówność zwaną nierównością Bienaymé-Chebyszeva. Aby ją jednak udowodnić posłużymy się następującym lematem.

Lemat 16.1. *{Nierówność Markowa} Niech X będzie zmienną losową o skończonej wartości oczekiwanej. Wtedy dla każdego $\epsilon > 0$,*

$$P(|X| \geq \epsilon) \leq \frac{E(|X|)}{\epsilon}.$$

Dowód.

$$\begin{aligned} P(|X| \geq \epsilon) &= \sum_{x \in X(\Omega), |x| \geq \epsilon} P(X = x) \leq \sum_{x \in X(\Omega), |x| \geq \epsilon} \frac{|x|}{\epsilon} P(X = x) \\ &\leq \frac{1}{\epsilon} \sum_{x \in X(\Omega)} |x| P(X = x) = \frac{1}{\epsilon} E(|X|). \end{aligned}$$

Wniosek 16.1. *{Nierówność Bienaymé-Chebyszeva} Niech X będzie zmienną losową o skończonej wartości oczekiwanej. Wtedy dla każdego $\epsilon > 0$,*

$$P(|X - EX| \geq \epsilon) \leq \frac{\text{Var}(X)}{\epsilon^2}.$$

Dowód. *Wystarczy podstawić $Y = |X - EX|^2$. Wtedy*

$$P(|X - EX| \geq \epsilon) = P(Y \geq \epsilon^2) \leq \frac{EY}{\epsilon^2} = \frac{\text{Var}(X)}{\epsilon^2}.$$

Poniżej podamy jedną z klasycznych wersji SPWL. Zwróćmy uwagę na założenia dotyczące zmiennych losowych.

Twierdzenie 16.1. *{Słabe prawo wielkich liczb}* Niech X_1, \dots, X_n, \dots będzie ciągiem nieskorelowanych zmiennych losowych o jednakowym rozkładzie ze skończoną wartością oczekiwaną μ i wariancją σ^2 . Niech ponadto $S_n = \sum_{i=1}^n X_i$. Wtedy $\forall \epsilon > 0$,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0.$$

Dowód. Z nierówności Bienaymé-Czebyszeva mamy:

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\text{Var}\left(\frac{S_n}{n}\right)}{\epsilon^2} = \frac{\text{Var}(S_n)}{\epsilon^2 n^2} = \frac{n\sigma^2}{\epsilon^2 n^2} = \frac{\sigma^2}{n\epsilon^2},$$

co dąży do 0, gdy $n \rightarrow \infty$.

Bezpośrednio z tego twierdzenia wynika tzw. Słabe Prawo Wielkich Liczb Bernoulliego.

Twierdzenie 16.2. *{Słabe Prawo Wielkich Liczb Bernoulliego}*

Jeżeli S_n jest liczbą sukcesów w n próbach wykonanych według schematu Bernoulliego, to dla S_n zachodzi SPWL, tzn.

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - p\right| > \epsilon\right) = 0 \quad \left(\frac{S_n}{n} \xrightarrow{p} p\right).$$

Dowód. Wystarczy zauważyć, że $\text{Var}(S_n) = npq$ i zastosować wcześniejsze twierdzenie.

Zauważmy, że w powyższym twierdzeniu założyliśmy, że zmienne są nieskorelowane i mają jednakowe rozkłady. Powstaje pytanie czy można osłabić te założenia? Odpowiedź uzyskamy w następującym twierdzeniu.

Twierdzenie 16.3. *{Prawo wielkich liczb Markowa}* Jeżeli (X_n) jest ciągiem zmiennych losowych takich, że

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \text{Var}\left(\sum_{k=1}^n X_k\right) = 0, \quad (12)$$

to (X_n) spełnia słabe prawo wielkich liczb.

Dowód. Przyjmijmy

$$Y_n = \frac{1}{n} \sum_{k=1}^n X_k.$$

Wtedy $E(Y_n) = \frac{1}{n} \sum_{k=1}^n m_k$. Z nierówności Czebyszewa wynika, że dla każdego $\epsilon > 0$ mamy

$$\begin{aligned} P(|Y_n - EY_n| > \epsilon) &= P\left(\left|\frac{1}{n} \sum_{k=1}^n (X_k - m_k)\right| \geq \epsilon\right) \\ &\leq \frac{1}{\epsilon^2} \cdot \frac{1}{n^2} \text{Var}\left(\sum_{k=1}^n X_k\right) \rightarrow 0, \end{aligned}$$

gdy $n \rightarrow \infty$, na mocy założenia.

Wniosek 16.2. *Jeżeli zmienne losowe X_1, X_2, \dots są niezależne, to warunek (12) przyjmuje postać*

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{k=1}^n \sigma_k^2 = 0.$$

Szczególnym przypadkiem prawa wielkich liczb Markowa jest prawo wielkich liczb Czebyszewa.

Twierdzenie 16.4. *{Prawo wielkich liczb Czebyszewa} Jeżeli zmienne losowe X_1, X_2, \dots są niezależne i ich wariancje są wspólnie ograniczone, tzn. istnieje takie σ^2 , że $\sigma_k^2 < \sigma^2$ dla $k = 1, 2, \dots$, to ciąg (X_n) spełnia SPWL.*

Dowód. Zauważmy, że

$$\frac{1}{n^2} \sum_{k=1}^n \sigma_k^2 \leq \sigma^2 \frac{1}{n} \rightarrow 0, \quad n \rightarrow \infty.$$

Prawa wielkich liczb Markowa i Czebyszewa dotyczą takich ciągów, dla których istnieją wariancje. Podamy teraz twierdzenie, w którym nie zakładamy istnienia wariancji zmiennych losowych, ale nie możemy zrezygnować z ich niezależności.

Twierdzenie 16.5. *{Prawo wielkich liczb Chinczyna} Jeżeli zmienne losowe X_1, X_2, \dots są niezależne i mają jednakowe rozkłady ze skończoną wartością oczekiwaną m , to ciąg (X_n) spełnia SPWL.*

16.2 Mocne prawa wielkich liczb

Różnica pomiędzy słabymi a mocnymi prawami wielkich liczb jest taka, że te drugie dotyczą zbieżności z prawdopodobieństwem jeden.

Zastanówmy się nad warunkiem mocnego prawa wielkich liczb (MPWL) Bernoulliego. Formalnie wyglądałby on następująco:

$$P\left(\lim_{n \rightarrow \infty} \frac{S_n}{n} = p\right) = 1,$$

lub korzystając z warunku równoważnego (11)

$$\lim_{n \rightarrow \infty} P\left(\{\omega : \sup_{k \geq n} \left| \frac{S_k(\omega)}{k} - p \right| > \epsilon\}\right) = 0,$$

dla każdego $\epsilon > 0$. Zauważmy, że chcąc oszacować to prawdopodobieństwo nie możemy już skorzystać z nierówności Markowa. Dlatego posłużymy się nierównością Bernsteina.

Można pokazać, że zachodzi następująca nierówność:

Twierdzenie 16.6. *{Nierówność Bernsteina}*

Jeżeli S_n oznacza liczbę sukcesów w n doświadczeniach wykonanych zgodnie ze schematem Bernoulliego z prawdopodobieństwem sukcesu p , to

$$P(\{\omega : \left| \frac{S_n(\omega)}{n} - p \right| \geq \epsilon\}) \leq 2e^{-2n\epsilon^2},$$

dla dowolnego $\epsilon > 0$.

Przejdźmy teraz do sformułowania twierdzenia.

Twierdzenie 16.7. *{MPWL Bernoulliego}* *Jeżeli S_n oznacza liczbę sukcesów w n doświadczeniach wykonanych zgodnie ze schematem Bernoulliego z prawdopodobieństwem sukcesu p , to dla każdego $\epsilon > 0$*

$$\lim_{n \rightarrow \infty} P(\{\omega : \sup_{k \geq n} \left| \frac{S_k(\omega)}{k} - p \right| > \epsilon\}) = 0,$$

Dowód. Mamy:

$$\begin{aligned} P\left(\{\omega : \sup_{k \geq n} \left| \frac{S_k(\omega)}{k} - p \right| > \epsilon\}\right) &= P\left(\bigcup_{k \geq n} \{\omega \in \Omega : \left| \frac{S_k(\omega)}{k} - p \right| > \epsilon\}\right) \\ &\leq \sum_{k \geq n} P\left(\{\omega : \left| \frac{S_k(\omega)}{k} - p \right| > \epsilon\}\right) \leq 2 \sum_{k \geq n} e^{-2k\epsilon^2} \\ &= 2e^{-2n\epsilon^2} \frac{1}{1 - e^{-2\epsilon^2}} \rightarrow 0, \quad n \rightarrow \infty. \end{aligned}$$

Poniżej podamy dwa twierdzenia Kołmogorowa dotyczące MPWL.

Twierdzenie 16.8. *{Prawo wielkich liczb Kołmogorowa}* Jeżeli (X_n) jest ciągiem niezależnych zmiennych losowych takich, że dla $n = 1, 2, \dots$,

$\text{Var}(X_n) < \infty$ oraz

$$\sum_{n=1}^{\infty} \frac{\text{Var}(X_n)}{n^2} < \infty, \quad (13)$$

to ciąg (X_n) spełnia MPWL

Twierdzenie 16.9. *{Prawo wielkich liczb Kołmogorowa}* Jeżeli zmienne losowe X_1, X_2, \dots są niezależne i mają jednakowe rozkłady ze skończoną wartością oczekiwaną m , to ciąg (X_n) spełnia MPWL.

16.3 Zadania

Zadanie 31. Niech (X_n) będzie ciągiem zmiennych losowych takich, że $\sigma^2 X_n \leq C < \infty$, $n = 1, 2, \dots$ oraz istnieje taka liczba naturalna k , że współczynnik korelacji $\rho(X_i, X_j) = 0$ dla $|i - j| > k$. Wykazać, że zachodzi SPWL.

Rozwiązanie. Chcemy pokazać, że

$$P\left(\left| \frac{S_n - ES_n}{n} \right| > \epsilon\right) \rightarrow 0 \quad n \rightarrow \infty.$$

Stosując nierówność Czebyszewa otrzymujemy

$$P(|S_n - ES_n| > n\epsilon) \leq \frac{E(S_n - ES_n)^2}{n^2\epsilon^2},$$

Licząc teraz występujący w liczniku prawej strony moment centralny otrzymujemy:

$$\begin{aligned}
E(S_n - ES_n)^2 &= \sum_{i=1}^n \text{Var } X_n + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(X_i, X_j) \\
&= \sum_{i=1}^n \text{Var } X_n + 2 \sum_{i=1}^n \sum_{j=i+1}^{i+k} \text{Cov}(X_i, X_j) \\
&\leq nC + 2 \sum_{i=1}^n \sum_{j=i+1}^{i+k} \sqrt{\text{Var } X_i \text{Var } X_j} \\
&\leq nC + 2 \sum_{i=1}^n \sum_{j=i+1}^{i+k} C \\
&= nC + 2nkC = nC(1 + 2k).
\end{aligned}$$

Stąd ostatecznie

$$P\left(\left|\frac{S_n - ES_n}{n}\right| > \varepsilon\right) \leq \frac{E(S_n - ES_n)^2}{n^2\varepsilon^2} \leq \frac{nC(1 + 2k)}{n^2\varepsilon} \rightarrow 0.$$

Zadanie 32. Dany jest ciąg nieskorelowanych zmiennych losowych (X_n) , przy czym X_n może przyjmować wartości: $-na$, 0 , na z prawdopodobieństwami odpowiednio równymi $\frac{1}{2n^2}$, $1 - \frac{1}{n^2}$, $\frac{1}{2n^2}$. Zbadać, czy do tego ciągu zmiennych losowych stosuje się prawo wielkich liczb.

Rozwiązanie. Sprawdzamy po kolei założenia twierdzenia 16.1:

1. Zmienne X_n są nieskorelowane.
2. $E(X_n) = -na \cdot \frac{1}{2n^2} + na \cdot \frac{1}{2n^2} = 0$, a zatem istnieje skończona wartość przeciętna jednakowa dla wszystkich zmiennych.
3. $E(X_n^2) = a^2$, więc $\sigma_n^2 = a^2$, czyli zmienne X_n mają tę samą skończoną wariancję.

Zatem ciąg $\left(\frac{1}{n} \sum_{i=1}^n X_i\right)$ jest zbieżny do wspólnej wartości przeciętnej $m = 0$, tj.

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i\right| \geq \varepsilon\right) = 0 \quad \forall \varepsilon > 0.$$

17 Centralne twierdzenie graniczne

17.1 Słaba zbieżność ciągów zmiennych losowych

W poprzednich paragrafach podaliśmy podstawowe typy zbieżności zmiennych losowych tj: zbieżność z prawdopodobieństwem jeden, według prawdopodobieństwa i według średniej. W tym rozdziale uzupełnimy tą listę o zbieżność według rozkładu.

Niech F_1, F_2, \dots będzie ciągiem dystrybuant odpowiadających zmiennym losowym X_1, X_2, \dots , a F dystrybuantą zmiennej losowej X , tzn.

$$F_n(x) = P(X_n \leq x), \text{ oraz } F(x) = P(X \leq x).$$

Definicja 17.1. Ciąg (X_n) jest zbieżny według rozkładu do zmiennej X , jeżeli

$$\lim_{n \rightarrow \infty} F_n(x) = F(x),$$

dla dowolnego punktu ciągłości dystrybuanty $F(x)$.

Ponieważ $F(x) = P(X \leq x)$ to znaczy, że prawdopodobieństwo, że X należy do pewnego przedziału jest bardzo bliskie prawdopodobieństwu, że dla dużych wartości n X_n należy do tego przedziału. Zbieżność według rozkładu oznaczamy symbolem

$$X_n \xrightarrow{D} X.$$

Zbieżność według rozkładu jest najslabszym rodzajem zbieżności z dotychczas poznanych, tak więc dowolny rodzaj jednej z wprowadzonych wcześniej zbieżności implikuje tą zbieżność. Ten typ rozkładu jest używany przy okazji badania centralnych twierdzeń granicznych.

Podamy teraz dwa ważne twierdzenia dotyczące zbieżności według rozkładu.

Twierdzenie 17.1. Ciąg (X_n) jest zbieżny według rozkładu wtedy i tylko wtedy, gdy dla dowolnej, ograniczonej funkcji ciągłej f

$$\lim_{n \rightarrow \infty} E[f(X_n)] = E[f(X)].$$

Twierdzenie 17.2. {Twierdzenie Levy'ego o ciągłości} Niech $\phi_n(t)$ będzie ciągiem funkcji charakterystycznych ciągu zmiennych losowych X_n a $\phi(t)$ funkcją charakterystyczną zmiennej X . Wtedy

$$\{\forall t \in \mathbb{R} : \phi_n(t) \rightarrow \phi(t)\} \Leftrightarrow X_n \xrightarrow{\mathcal{D}} X.$$

17.2 Centralne twierdzenie graniczne

Twierdzenie 17.3. {Centralne twierdzenie graniczne Lindeberga-Lévy'ego.} Rozważmy ciąg X_1, X_2, \dots , niezależnych zmiennych losowych o jednakowym rozkładzie, ze skończoną wartością oczekiwaną μ i skończoną wariancją $\sigma^2 > 0$.

Niech $S_n = X_1 + X_2 + \dots + X_n$ i $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$. Wtedy zachodzi wzór:

$$\forall x \in \mathbb{R} \quad \lim_{n \rightarrow \infty} P(Z_n \leq x) = \Phi(x),$$

gdzie $\Phi(x)$ jest dystrybuantą rozkładu $N(0, 1)$.

Dowód. Jeżeli zmienna losowa Y ma wartość oczekiwaną równą 0 a wariancję 1, to jej funkcja charakterystyczna da się zapisać w postaci

$$\varphi_Y(t) = 1 - \frac{t^2}{2} + o(t^2), \quad t \rightarrow 0.$$

Niech $Y_i = \frac{X_i - \mu}{\sigma}$. Wtedy można zapisać:

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \sum_{i=1}^n \frac{Y_i}{\sqrt{n}},$$

gdzie $Y_i \sim Y$. Zatem funkcja charakterystyczna zmiennej Z_n ma postać

$$\left[\varphi_Y \left(\frac{t}{\sqrt{n}} \right) \right]^n = \left[1 - \frac{t^2}{2n} + o \left(\frac{t^2}{n} \right) \right]^n \rightarrow e^{-t^2/2}, n \rightarrow +\infty.$$

Jest to funkcja charakterystyczna rozkładu $N(0, 1)$.

Szczególnym przypadkiem powyższego twierdzenia jest Twierdzenie de Moivre'a Laplace'a.

Twierdzenie 17.4. *Niech X_1, \dots, X_n będzie ciągiem niezależnych prób Bernoulliego z jednakowym prawdopodobieństwem sukcesu w każdej próbie równym p ($p \in (0, 1)$) i porażki $q = 1 - p$. Wtedy*

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - np}{\sqrt{npq}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \quad x \in R,$$

gdzie $S_n = X_1 + \dots + X_n$.

Nie zawsze są jednak spełnione założenia powyższego twierdzenia. Wtedy musimy korzystać z innych warunków na zachodzenie CTG.

Twierdzenie 17.5. *{Warunek Lapunowa} Niech X_n będzie ciągiem niezależnych zmiennych losowych o skończonych średnich i skończonych wariancjach. zdefiniujmy*

$$s_n^2 = \sum_{i=1}^n \sigma_i^2.$$

Załóżmy ponadto, że momenty centralne trzeciego rzędu są skończone i połączmy

$$r_n^3 = \sum_{i=1}^n E(|X_i - \mu_i|^3).$$

Jeżeli

$$\lim_{n \rightarrow \infty} \frac{r_n}{s_n} = 0,$$

to ciąg

$$Z_n = \frac{S_n - m_n}{s_n},$$

gdzie

$$m_n = \sum_{i=1}^n \mu_i$$

jest zbieżny według prawdopodobieństwa do zmiennej $N(0, 1)$.

Ponadto przy powyższych oznaczeniach zachodzi

Twierdzenie 17.6. *{Warunek Lindeberga (1920)}*

Jeżeli dla każdego $\epsilon > 0$

$$\lim_{n \rightarrow +\infty} \frac{1}{s_n^2} \sum_{i=1}^n \mathbb{E} \left[(X_i - \mu_i)^2 \cdot \mathbf{I}_{\{|X_i - \mu_i| > \epsilon s_n\}} \right] = 0,$$

to ciąg Z_n jest zbieżny według rozkładu do zmiennej $N(0, 1)$.

17.3 Zadania

Zadanie 33. Niech (X_n) będzie ciągiem niezależnych zmiennych losowych takich, że

$$P[X_k = k^{1/2}] = \frac{1}{2k} = P[X_k = -k^{1/2}] \quad \text{i} \quad P[X_k = 0] = 1 - \frac{1}{k}.$$

Sprawdzić, czy ciąg (X_n) spełnia

(a) warunek Lapunowa?

Rozwiązanie.

$$\begin{aligned} EX_k &= 0, \quad \sigma^2 X_k = 1, \quad s_n^2 = \sigma^2(S_n) = n \\ r_n^3 &= \sum_{k=1}^n E|X_k - EX_k|^3 = \sum_{k=1}^n k^{3/2} \sim n^{5/2}. \end{aligned}$$

Stąd

$$L_n = \frac{r_n^3}{s_n^3} \sim \frac{n^{5/2}}{n^{3/2}} \rightarrow 0, \quad n \rightarrow \infty.$$

Warunek Lapunowa nie jest spełniony.

Zadanie 34. Niech X będzie zmienną losową taką, że $EX = 0$,

$E|X|^p < \infty$, dla każdego $p > 0$. Rozważmy ciąg niezależnych zmiennych losowych $\{X_n, n \geq 1\}$ takich, że $X_k \stackrel{D}{=} \sqrt{k}X$. Wykazać, że ciąg zmiennych losowych $\{S_n, n \geq 1\}$, gdzie

$$S_n = \frac{1}{n} \sum_{k=1}^n X_k$$

jest zbieżny do rozkładu normalnego (według rozkładu).

Rozwiązanie. Jeżeli ciąg ma być zbieżny do rozkładu normalnego tzn. ma zachodzić CTG. Sprawdźmy wobec tego warunek Lapunowa.

$$E|X_k|^3 = Ek^{\frac{3}{2}}|X|^3 = k^{\frac{3}{2}}E|X|^3.$$

Podobnie

$$s_n = \sqrt{\sum_{k=1}^n \text{Var} X_k} = \sqrt{\sum_{k=1}^n EX_k^2} = \sqrt{\sum_{k=1}^n kEX^2} = \sqrt{\frac{n(n+1)}{2}EX^2}.$$

Zatem

$$\frac{1}{s_n^3} \sum_{k=1}^n E|X_k|^3 = \frac{\sum_{k=1}^n k^{\frac{3}{2}}E|X|^3}{\left(\frac{n(n+1)}{2}EX^2\right)^{\frac{3}{2}}} = \frac{2^{\frac{3}{2}}E|X|^3}{(EX^2)^{\frac{3}{2}}} \frac{\sum_{k=1}^n k^{\frac{3}{2}}}{n^{\frac{3}{2}}(n+1)^{\frac{3}{2}}} \sim C \frac{n^{\frac{5}{2}}}{n^3} \rightarrow 0, \quad n \rightarrow \infty.$$

Stąd wiemy, że

$$\frac{S_n}{s_n} \xrightarrow{D} N(0, 1)$$

ale $s_n = n\sqrt{\frac{n+1}{2n}EX^2}$. Zatem

$$\frac{S_n}{n} \xrightarrow{D} N\left(0, \sqrt{\frac{EX^2}{2}}\right).$$

Zadanie 35. Znaleźć liczbę doświadczeń, które należy wykonać, aby błąd oszacowania nieznannej wielkości $\mu = EX$ przez wartość średniej \bar{X}_n z próby był nie większy niż 10% odchylenia standardowego σ z prawdopodobieństwem wynoszącym co najmniej 0,99. (Podać dwie różne metody.)

Rozwiązanie. Należy znaleźć n takie, aby

$$P\left[|\bar{X}_n - \mu| < 0,1\sigma\right] \geq 0,99.$$

Na mocy nierówności Czebyszewa

$$P\left[|\bar{X}_n - \mu| \leq 0,1\sigma\right] = P\left[|S_n - n\mu| \leq 0,1\sigma n\right] \geq 1 - \frac{\sigma^2 S_n}{n^2 \sigma^2 (0,1)^2} = 1 - \frac{1}{n \cdot (0,1)^2}.$$

Stąd

$$1 - \frac{1}{n \cdot \frac{1}{100}} \geq 0,99 \implies \frac{100}{n} \leq \frac{1}{100} \implies n \geq 10000.$$

Na podstawie centralnego twierdzenia granicznego

$$P \left[\left| \bar{X}_n - \mu \right| < 0,1\sigma \right] \approx 1 - 2\Phi \left(\frac{\sqrt{n}}{10} \right) \leq 0,01.$$

Stąd

$$\begin{aligned} \Phi \left(\frac{\sqrt{n}}{10} \right) &\geq 0,495, & \Phi^{-1}(0,495) &= 2.58 \\ n &\geq (25,8)^2, & n &= 666. \end{aligned}$$

18 Proces Poissona

W poprzednich paragrafach badaliśmy własności ciągów zmiennych losowych, to znaczy zbiorów zmiennych losowych indeksowanych liczbami naturalnymi. Okazuje się jednak, że takie ciągi są niewystarczające na przykład do opisu zjawisk zachodzących w czasie. Gdy notujemy indeksy giełdowe, obserwujemy rozkład temperatury w danym miejscu w określonym przedziale czasowym mamy do czynienia z wartościami zmiennych losowych indeksowanych na pewnym zbiorze nieprzeliczalnym, na przykład przedziale. Taką rodzinę zmiennych losowych określonych na pewnej przestrzeni probabilistycznej o wartościach w pewnej przestrzeni mierzalnej nazywamy procesem stochastycznym. Celem tego wykładu nie jest jednak systematyczny kurs dotyczący teorii procesów stochastycznych. Przedstawiony zostanie jedynie pewien szczególny przypadek procesu stochastycznego, mający szerokie zastosowania w opisie różnych zjawisk losowych zachodzących w czasie. Nosi on nazwę procesu Poissona.

Rozważmy zmienne losowe T_1, T_2, T_3, \dots spełniające warunek $P[T_j > 0] = 1$. Połóżmy $S_0 = 0$ i dla $n \geq 1$, $S_n = \sum_{j=1}^n T_j$. Ponadto niech $N(t) = \max\{n \geq 0 : S_n \leq t\}$. Inaczej mówiąc

$$N(t) = \left\{ \begin{array}{l} 0 \quad \text{gdy } S_0 \leq t < S_1 \\ 1 \quad \text{gdy } S_1 \leq t < S_2 \\ 2 \quad \text{gdy } S_2 \leq t < S_3 \\ \cdot \\ \cdot \\ \cdot \\ k \quad \text{gdy } S_k \leq t < S_{k+1} \\ \cdot \end{array} \right. .$$

Jeżeli S_1, S_2, S_3, \dots reprezentują czasy, w których zachodzą pewne zdarzenia, to $N(t)$ reprezentuje całkowitą liczbę zdarzeń, które zaszły w przedziale czasu $[0, t]$. To uzasadnia następującą definicję.

Określenie procesu liczącego

Rozważmy zmienne losowe T_1, T_2, T_3, \dots spełniające warunek $P[T_j > 0] = 1$ dla każdego $j \geq 1$. Połóżmy $S_0 = 0$ oraz dla $n \geq 1$, $S_n = \sum_{j=1}^n T_j$. Wreszcie dla $t \geq 0$ zdefiniujemy proces

$$N(t) = \max\{n \geq 0 : S_n \leq t\}.$$

Proces losowy $(N(t); t \geq 0)$ nazywamy procesem liczącym.

W zastosowaniach zmienne losowe T_1, T_2, T_3, \dots reprezentować mogą na przykład długości życia podzespołów elektronicznych, które używane są jeden po drugim. Dla zilustrowania tego założymy, że dysponujemy dużą liczbą żarówek elektrycznych, powiedzmy żarówką numer 1, numer 2 numer 3 itd. Używamy żarówki jedna po drugiej w pewnym systemie elektronicznym. Tak więc w chwili 0 instalujemy żarówkę numer 1. Czas świecenia tej żarówki to

właśnie zmienna T_1 , po czym zostaje ona zastąpiona przez żarówkę numer 2, dla której czas działania wynosi T_2 itd. Zmienna losowa $N(t)$ reprezentuje zatem liczbę zmian żarówek, jaka odbyła się w przedziale czasowym $[0, t]$. Inaczej zmiennie T_1, T_2, T_3, \dots interpretuje się czasami jako czasy oczekiwania pomiędzy przybywającymi klientami w kolejce.

Wtedy zmienna $N(t)$ oznacza liczbę klientów, którzy przybyli w przedziale $[0, t]$. W analizie procesów liczących zmiennie losowe T_1, T_2, T_3, \dots są często nazywane czasami trwania życia procesu $N(t); t \geq 0$, lub czasami oczekiwania tego procesu.

Szczególny przypadek, kiedy czasy trwania życia są niezależnymi zmiennymi losowymi o jednakowym rozkładzie jest łatwy do analizowania. Weźmy przykład żarówek rozważany poprzednio z czasami życia T_1, T_2, T_3, \dots , które są niezależne i mają jednakowe rozkłady.

W momencie, kiedy instalujemy w systemie nową żarówkę mówimy, że ma miejsce odnowa systemu czyli tak jakby system zaczął swoją pracę od nowa. Zmienna losowa $N(t)$ reprezentuje wtedy liczbę "odnów" systemu w czasie $[0, t]$.

Określenie procesu odnowy

Procesem odnowy nazywamy proces liczący, dla którego czasy trwania życia są niezależnymi zmiennymi losowymi o jednakowym rozkładzie.

Bardzo ważnym przykładem procesu odnowy jest rozważany w pracy proces Poissona.

Pierwsze określenie procesu Poissona

Proces Poissona z intensywnością λ jest to proces odnowy, w którym rozkład długości życia jest rozkładem wykładniczym z parametrem λ .

Takie określenie procesu Poissona jest uzasadnione przez następujące twierdzenie.

Twierdzenie 18.1. *Jeżeli $(N(t); t \geq 0)$ jest procesem odnowy z rozkładem trwania życia wykładniczym z parametrem λ , to dla $t \geq 0$ $N(t)$ ma rozkład Poissona z parametrem λt .*

Dowód:

Dowód twierdzenia opiera się na dwóch następujących spostrzeżeniach:

(1) Dla dowolnej liczby rzeczywistej $t \geq 0$ i liczby całkowitej $n \geq 0$, mamy $N(t) \geq n$ wtedy i tylko wtedy, gdy $S_n \leq t$. otrzymujemy więc:

$$P[N(t) \geq n] = P[S_n \leq t].$$

(2) Jeżeli czasy trwania życia T_1, T_2, T_3, \dots są niezależnymi zmiennymi losowymi o jednakowym rozkładzie wykładniczym z parametrem λ , to znaczy $T_i \sim \Gamma(1, \lambda)$, to zgodnie ze znanym twierdzeniem o addytywności mamy:

$$S_n \sim \Gamma(n, \lambda).$$

Przejdźmy teraz do szczegółowego dowodu twierdzenia.

Dla dowolnego $n \geq 0$ mamy:

$$P[N(t) = n] = P[N(t) \geq n] - P[N(t) \geq n + 1] =$$

$$= P[S_n \leq t] - P[S_{n+1} \leq t] =$$

$$\int_0^t \frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s} ds - \int_0^t \frac{\lambda^{n+1}}{n!} s^n e^{-\lambda s} ds =$$

$$\int_0^t \left(\frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s} - \frac{\lambda^{n+1}}{n!} s^n e^{-\lambda s} \right) ds =$$

$$\int_0^t \frac{d}{ds} \left(\frac{\lambda^n}{n!} s^n e^{-\lambda s} \right) ds = \left(\frac{\lambda^n}{n!} s^n e^{-\lambda s} \right) \Big|_{s=0}^{s=t} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

To pokazuje, że $N(t)$ ma rozkład Poissona z parametrem λt .

Używając własności braku pamięci dla rozkładu wykładniczego, można uogólnić poprzednie twierdzenie i pokazać, że jeżeli $(N(t); t \geq 0)$ jest procesem Poissona z intensywnością λ , to

(a) dla dowolnego wyboru liczb rzeczywistych $0 \leq s \leq t < \infty$, mamy

$$N(t) - N(s) \sim P(\lambda(t-s))$$

(b) dla dowolnej dodatniej liczby całkowitej n i dla dowolnego wyboru liczb rzeczywistych $0 = t_1 \leq t_2 \dots \leq t_n < \infty$, zmienne losowe

$$N(t_1) - N(t_0), N(t_2) - N(t_1), \dots, N(t_n) - N(t_{n-1})$$

są wzajemnie niezależne. Zobaczmy teraz, że zachodzi także własność odwrotna. Niech $(N(t); t \geq 0)$, będzie procesem liczącym spełniającym warunki (a) i (b) podane powyżej. Wyznamy rozkład czasu pierwszej odnowy $T_1 = \min\{t > 0 : N(t) = 1\}$. Dla $t > 0$ mamy:

$$P[T_1 > t] = P[N(t) = 0] = e^{-\lambda t} \frac{(\lambda t)^0}{0!} = e^{-\lambda t}.$$

Mamy więc:

$$F_{T_1}(t) = \begin{cases} 1 - e^{-\lambda t} & \text{gdy } t \geq 0 \\ 0 & \text{gdy } t < 0 \end{cases}.$$

To pokazuje, że T_1 ma rozkład wykładniczy z parametrem λ . W ten sam sposób można pokazać, że każda ze zmiennych losowych $T_k = \min\{t > 0 : N(t) = k\} - \min\{t > 0 : N(t) = k - 1\}$ ma rozkład wykładniczy z parametrem λ i zmienne te są wzajemnie niezależne. Tak więc, gdy $(N(t); t \geq 0)$ jest procesem liczącym spełniającym warunki (a) i (b), to proces jest procesem Poissona z intensywnością λ .

Przejdźmy teraz do następnej własności procesów stochastycznych zwanej własnością stacjonarności. Niech $(X_t; t \geq 0)$ będzie procesem losowym. Wtedy zmienną losową $X_\nu - X_u$ nazywamy przyrostem procesu $(X_t; t \geq 0)$ na prze-

dziale $(u, \nu]$. Mówimy, że $(X_t; t \geq 0)$ jest procesem o przyrostach stacjonarnych, jeżeli rozkład przyrostu na pewnym przedziale zależy jedynie od długości tego przedziału, to znaczy

$$X_{t+s} - X_t \sim X_s - X_0$$

dla dowolnego $t \geq 0$.

Mówimy natomiast, że $(X_t; t \geq 0)$ jest procesem o przyrostach niezależnych, jeżeli przyrosty odpowiadające rozłącznym przedziałom są zmiennymi losowymi niezależnymi, to znaczy dla $0 = t_1 \leq t_2 \leq \dots \leq t_n < \infty$ zmienne losowe

$$X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$$

są wzajemnie niezależne.

Tak więc warunek (b) mówi o tym, że $(N(t); t \geq 0)$ jest procesem o przyrostach niezależnych.

Natomiast warunek (a) mówi, że $(N(t); t \geq 0)$ jest procesem o przyrostach stacjonarnych. Pierwsze określenie dla procesu Poissona jest więc równoważne z następującą definicją:

Drugie określenie procesu Poissona

Proces Poissona $(N(t); t \geq 0)$ z intensywnością λ jest to proces liczący o przyrostach stacjonarnych i niezależnych taki, że

$$N(t) - N(s) \sim P(\lambda(t - s))$$

dla dowolnych $0 \leq s \leq t < \infty$.

Zaprezentujemy teraz trzecią definicję dla procesu Poissona i pokażemy, że jest ona równoważna dwóm poprzednim definicjom. Ale na początku musimy wprowadzić notację nieskończenie małej rzędu t a mianowicie notację $o(t)$.

Notacja $o(t)$.

Założmy, że g jest pewną funkcją zdefiniowaną na przedziale $(0, \infty)$ o wartościach rzeczywistych. Mówimy, że funkcja g jest nieskończenie małą rzędu u , gdy $u \rightarrow 0$ i piszemy

$$g(u) = o(u) \quad u \rightarrow 0,$$

jeżeli

$$\lim_{u \rightarrow 0} \frac{g(u)}{u} = 0.$$

Notacji $o(u)$ używamy reprezentując dowolną funkcję $g(u)$ spełniającą warunek

$\lim_{u \rightarrow 0} \frac{g(u)}{u} = 0$. Warto odnotować, że o ile notacja $o(u)$ może być stosowana wiele razy w różnych równaniach lub nawet w tym samym, to wartość $o(u)$ nie musi reprezentować tej samej wielkości.

Rozważmy teraz proces Poissona $(N(t); t \geq 0)$ z intensywnością λ . Ustalmy wartość rzeczywistą dodatnią t oraz nieujemną całkowitą k . Wykonując proste obliczenia otrzymujemy następujące rezultaty:

(a) $P[N(t+h) - N(t) = 0 | N(t) = k] = 1 - \lambda h + o(h)$, gdy $h \rightarrow 0$.

(b) $P[N(t+h) - N(t) = 1 | N(t) = k] = \lambda h + o(h)$, gdy $h \rightarrow 0$.

(c) $P[N(t+h) - N(t) \geq 2 | N(t) = k] = o(h)$, gdy $h \rightarrow 0$.

Dla przykładu punkt (a) otrzymujemy w następujący sposób:

$$P[N(t+h) - N(t) = 0 | N(t) = k] = P[N(t+h) - N(t) = 0] = e^{-\lambda h}.$$

$$= 1 - \lambda h + \frac{(\lambda h)^2}{2!} - \frac{(\lambda h)^3}{3!} + \frac{(\lambda h)^4}{4!} - \frac{(\lambda h)^4}{4!} + \dots =$$

$$= 1 - \lambda h + (\lambda h)^2 \left(\frac{1}{2!} - \frac{\lambda h}{3!} + \frac{(\lambda h)^2}{4!} - \dots \right) =$$

$$1 - \lambda h + o(h).$$

Pierwsza równość wynika z faktu, że przyrosty procesu Poissona są niezależne. Druga równość wynika z faktu, że $N(t+h) - N(t)$ ma rozkład Poissona z parametrem λh . Następnie użyliśmy rozwinięcia funkcji wykładniczej w szereg Taylora a na końcu skorzystaliśmy z tego, że:

$$\lim_{h \rightarrow 0} \frac{(\lambda h)^2 \left(\frac{1}{2!} - \frac{\lambda h}{3!} + \frac{(\lambda h)^2}{4!} - \dots \right)}{h} = 0.$$

Punkt (b) oraz (c) otrzymujemy analogicznie.

Udowodnimy teraz następujący rezultat: jeżeli $(N(t); t \geq 0)$ jest procesem liczącym spełniającym warunki (a)-(c), to $(N(t); t \geq 0)$ jest procesem Poissona z intensywnością λ .

Zanotujmy na początku fakt jasny intuicyjnie, że jeżeli $(N(t); t \geq 0)$ spełnia warunki (a)-(c), to $(N(t); t \geq 0)$ jest procesem o przyrostach stacjonarnych

i niezależnych.

Aby jeszcze pokazać, że $(N(t); t \geq 0)$ jest procesem Poissona, zostaje nam jedynie udowodnić, że $N(t) \sim P(\lambda t)$ dla każdego $t \geq 0$. Dla $t \geq 0$ i dla $n = 0, 1, 2, \dots$, połóżmy

$$f_n(t) = P[N(t) = n].$$

Trzeba pokazać, że

$$f_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Rozważmy na wstępie funkcję $f_0(t)$. Z (a) mamy:

$$\begin{aligned} f_0(t+h) &= P[N(t+h) = 0] \\ &= P[N(t) = 0]P[N(t+h) - N(t) = 0 | N(t) = 0] = \\ &= f_0(t)(1 - \lambda h + o(h)). \end{aligned}$$

Mamy więc

$$\frac{f_0(t+h) - f_0(t)}{h} = f_0(t) \left(-\lambda + \frac{o(h)}{h} \right)$$

Gdy $h \rightarrow 0$ otrzymujemy:

$$f_0'(t) = -\lambda f_0(t).$$

Z warunkiem początkowym $f_0(0) = P[N(0) = 0] = 1$, rozwiązaniem tego równania różniczkowego jest $f_0(t) = e^{-\lambda t}$. Mamy więc:

$$P[N(t) = 0] = e^{-\lambda t}.$$

Rozważmy teraz funkcję $f_1(t)$. Używając (a) oraz (b), otrzymujemy:

$$\begin{aligned}
 f_1(t+h) &= P[N(t+h) = 1] = \\
 &= P[N(t) = 0]P[N(t+h) - N(t) = 1 | N(t) = 0] + \\
 &= P[N(t) = 1]P[N(t+h) - N(t) = 0 | N(t) = 1] = \\
 &= f_0(t)(\lambda h + o(h) + f_1(t))(1 - \lambda h + o(h)).
 \end{aligned}$$

Mamy więc:

$$\frac{f_1(t+h) - f_1(t)}{h} = e^{-\lambda t} \left(\lambda + \frac{o(h)}{h} \right) + f_1(t) \left(-\lambda + \frac{o(h)}{h} \right).$$

Gdy $h \rightarrow 0$ otrzymujemy:

$$f_1'(t) = -\lambda f_1(t) + \lambda e^{-\lambda t}.$$

Z warunkiem początkowym $f_1(0) = P[N(0) = 1] = 0$, rozwiązaniem tego równania różniczkowego jest $f_1(t) = \lambda t e^{-\lambda t}$. Mamy więc

$$P[N(t) = 1] = \lambda t e^{-\lambda t}.$$

W ten sam sposób kontynuując nasze rozumowanie z funkcjami $f_2(t), f_3(t), \dots$ otrzymujemy

$$P[N(t) = k] = \frac{e^{-\lambda t} (\lambda t)^k}{k!}.$$

To pokazuje, że

$$N(t) \sim P(\lambda t)$$

dla dowolnego $t \geq 0$. Pokazaliśmy tym samym, że następująca definicja procesu Poissona jest równoważna z dwoma poprzednimi.

Trzecie określenie procesu Poissona

Proces Poissona $(N(t); t \geq 0)$ z intensywnością λ jest to proces liczący taki, że:

$$(a) P[N(t+h) - N(t) = 0 | N(t) = k] = 1 - \lambda h + o(h), \text{ gdy } h \rightarrow 0.$$

$$(b) P[N(t+h) - N(t) = 1 | N(t) = k] = \lambda h + o(h), \text{ gdy } h \rightarrow 0.$$

$$(c) P[N(t+h) - N(t) \geq 2 | N(t) = k] = o(h), \text{ gdy } h \rightarrow 0.$$

Literatura

- [1] Billingsley P. *Prawdopodobieństwo i miara*. Warszawa, PWN 1987.
- [2] Fisz M. *Rachunek prawdopodobieństwa i statystyka matematyczna*. Warszawa, PWN 1967.
- [3] Jakubowski J. Sztencel R. *Wstęp do teorii prawdopodobieństwa*. Warszawa, SCRIPT 2000.
- [4] Jakubowski J. Sztencel R. *Rachunek prawdopodobieństwa dla (prawie) każdego*. Warszawa, SCRIPT 2002.
- [5] Krzyśko M. *Wykłady z teorii prawdopodobieństwa*. Warszawa, WNT 2000.
- [6] Foryś U. *Matematyka w biologii*. Warszawa, WNT 2005.

- [7] Murray J. D. *Wprowadzenie do biomatematyki*. Warszawa, PWN 2006.
- [8] Wrzosek D. *Matematyka dla biologów*. Warszawa, Wydawnictwo Uniwersytetu Warszawskiego, 2008.