Zbiór zadań z metod probabilistycznych

Janina Płaskonka
i Ryszard Rębowski

seria wydawnicza
Państwowej Wyższej Szkoły Zawodowej im. Witelona w Legnicy

Seria wydawnicza
Państwowej Wyższej Szkoły Zawodowej im. Witelona w Legnicy

Janina Płaskonka i Ryszard Rębowski

Zbiór zadań z metod probabilistycznych

Stowarzyszenie „Wspólnota Akademicka”

Legnica 2008

Janina Płaskonka, Wydział Zarządzania i Informatyki PWSZ w Legnicy
e-mail: plaskonkaj@pwsz-legnica.eu
Ryszard Rębowski, Wydział Zarządzania i Informatyki PWSZ w Legnicy
e-mail: rebowskir@pwsz-legnica.eu

Recenzent: prof. dr hab. inż. Marek Kurzyński

Wydawca:
Państwowa Wyższa Szkoła Zawodowa im. Witelona w Legnicy
Redakcja i diapozytywy:
Stowarzyszenie na Rzecz Rozwoju
Państwowej Wyższej Szkoły Zawodowej im. Witelona w Legnicy
,,Wspólnota Akademicka"
ul. Sejmowa 5A, 59-220 Legnica
tel. (076)723 21 20, tel./fax(076)723 29 04.

Skład komputerowy w systemie LaTeX wykonali autorzy.

Wydanie pierwsze.

Nakład: 500 egz.

ISBN 978-83-61389-16-3

Spis treści

Wstęp

Napisany przez nas zbiór zadań stanowi naturalne uzupełnienie książki, która ukazała się w 2006 roku nakładem Serii Wydawniczej PWSZ w Legnicy *Podstawy metod probabilistycznych*. Przedmiot ten wykładany jest na kierunku informatyka PWSZ. Jednocześnie tematyka tej książki jest bliska zagadnieniom omawianym w ramach wykładów ze statystyki inżynierskiej. Dlatego uważamy, że tym bardziej nasza praca była celowa.

Zagadnienia dotyczące teorii prawdopodobieństwa i jej zastosowań, czyli statystyki matematycznej, są różnie przez różnych autorów przedstawiane. Studenci na ogół odbierają ten przedmiot komentarzem: *jest to inna matematyka*. Twierdzimy w sposób zdecydowany, że problem leży w przekazie i przygotowaniu studenta. Dlatego też dużą uwagę zwróciliśmy na usystematyzowanie prezentowanych zagadnień. Nasza koncepcja oparta jest na:

1. założeniu, że podstawą teorii jest pojęcie *przestrzeni probabilistycznej* ustanowione przez A. Kołmogorowa,

2. potrzebie pokazania konsekwencji przyjętej zasady, a więc sposobów realizacji takiej przestrzeni,

3. pokazaniu ewolucji opisu zjawisk losowych poprzez odwołanie się do nowego sposobu ich opisu z wykorzystaniem pojęcia *rozkładu prawdopodobieństwa* oraz pojęcia *zmiennej losowej.*

Dlatego też wyraźnie oddzieliliśmy od siebie te metody. W pierwszych rozdziałach zaprezentowaliśmy sposób postrzegania zjawiska losowego poprzez jego bezpośredni opis w kategoriach losowych, czyli z wykorzystaniem pojęć: *zdarzenia, zdarzenia elementarnego i prawdopodobieństwa zdarzenia.* Przykładami realizacji takich modeli jest: *model dyskretny* z jego odmianami, *model warunkowy*, *model produktowy* czy *model geometryczny.* Celem zadań, które zamieściliśmy w tych rozdziałach jest zrozumienie roli tego opisu i nauczenie się jego stosowania.

Począwszy od rozdziału czwartego, pokazaliśmy w kilku krokach, jak na te same zjawiska można spojrzeć inaczej. Kluczem do zrozumienia tego punktu widzenia jest pojęcie *rozkładu prawdopodobieństwa.* Zaczęliśmy od sytuacji najprostszej, czyli od *rozkładu dyskretnego skończonego*, uogólnoając problem do sytuacji nieskończonej. Jednocześnie zaakcentowaliśmy zaletę tego punktu widzenia na teorię–rolę *zmiennej losowej*, pokazując zmianę w stosunku do sytuacji dotychczasowej, jak i podkreślając związek nowej interpretacji z podstawowym opisem zjawiska losowego, jakim jest pojęcie przestrzeni probabilistycznej. Wreszcie pokazaliśmy, że całość zagadnień można sprowadzić do teorii funkcji rzeczywistych i analizy tych funkcji. Oczywiście mamy tutaj na myśli pojęcie *funkcji dystrybuanty* i *funkcji gęstości rozkładu.*

Przy opracowaniu zbioru przyjęliśmy zasadę, że każdy rozdział będzie zawierał wprowadzenie teoretyczne i przykłady przez nas rozwiązane. Jednocześnie wyszliśmy z założenia, że zbiór ten nie może i nie powinien zastępować ani książki *Podstawy metod probabilistycznych*, ani wykładu. Niewątpliwą atrakcją powinien być rozdział zatytułowany *Przykładowe zadania z egzaminów i kolokwiów.*

Oczywiście nasza praca jest jedną z wielu. Dlatego należy wyraźnie wspomnieć o rozwiązaniach alternatywnych czy też o charakterze uzupełniającym. Odpowiednie informacje zamieściliśmy w zestawieniu literatury.

Wierzymy, że nasza praca umożliwi:

1. studentom - szybsze opanowanie przedmiotu,

2. prowadzącym zajęcia - efektywniejszą realizację zamierzonych celów dydaktycznych.

maj 2008

Janina Płaskonka i Ryszard Rębowski

Rozdział 1

Zbiory i rodziny zbiorów

1.1 Wprowadzenie teoretyczne i przykłady

Zbiory definiujemy poprzez określenie ich elementów. Dwa zbiory, które mają te same elementy, uważamy za identyczne.

$$A = B \Leftrightarrow \forall x \ (x \in A \Leftrightarrow x \in B)$$

Piszemy $A \subset B$ (zbiór A zawiera się w zbiorze B), jeżeli każdy element zbioru A należy również do zbioru B.

$$A \subset B \Leftrightarrow \forall x \ (x \in A \Rightarrow x \in B)$$

Podstawowe działania na zbiorach to: suma, iloczyn i różnica ($A \cup B$, $A \cap B$, $A \setminus B$), które definiujemy nastepująco:

$$x \in A \cup B \Leftrightarrow x \in A \vee x \in B,$$

$$x \in A \cap B \Leftrightarrow x \in A \wedge x \in B,$$

$$x \in A \setminus B \Leftrightarrow x \in A \wedge x \notin B.$$

Zbiór A^C nazywamy dopełnieniem zbioru A.

$$x \in A^C \Leftrightarrow x \notin A$$

Moc zbioru $|A|$ określa ilość elementów danego zbioru.

Iloczyn kartezjański zbiorów A i B ($A \times B$) to taki zbiór par uporządkowych (a, b), że $a \in A \wedge b \in B$.

Zbiór, którego elmentami są zbiory, nazywamy *rodziną zbiorów*.

Rodzina $\mathcal{A}$ jest rodziną indeksowaną, jeżeli istnieje zbiór indeksów T taki, że każdemu indeksowi $t \in T$ jest przyporządkowany w sposób jednoznaczny jeden element A_t rodziny $\mathcal{A}$.

Będziemy wtedy pisali $\mathcal{A} = \{A_t,\ t \in T\}$.

Na rodzinie $\mathcal{A} = \{A_t, t \in T\}$ można wykonywać działania mnogościowe:

$$\bigcup_{t\in T} A_t = \{x \in X : \exists_{t\in T}\ x \in A_t\},$$

$$\bigcap_{t\in T} A_t = \{x \in X : \forall_{t\in T}\ x \in A_t\}.$$

Dla danego zbioru Ω niech Σ oznacza rodzinę złożoną z podzbiorów Ω, która ma następujące własności:

1. zawiera zbiór pusty,
2. zamknięta jest na branie dopełnienia,
3. zamknięta jest na branie sum przeliczalnych, czyli
$$\forall_{n\in\mathbf{N}_o}\ A_n \in \Sigma \Rightarrow \bigcup_{n\in\mathbf{N}_o} A_n \in \Sigma$$
dla dowolnego podzbioru niepustego $\mathbf{N}_o \subset \mathbf{N}$.

Dalej rodzinę Σ będziemy nazywali *σ-algebrą zdarzeń*, zbiór Ω będziemy nazywali *przestrzenią zdarzeń elementarnych*.
Zdarzenie pewne oznaczamy przez Ω, a *zdarzenie niemożliwe* przez $\emptyset$. *Zdarzeniem przeciwnym* do A nazwiemy jego dopełnienie A^C.

Zadanie 1.1.1 *Niech*
A będzie zbiorem liczb naturalnych podzielnych przez 3,
B–zbiorem liczb naturalnych podzielnych przez 5,
C–zbiorem liczb naturalnych podzielnych przez 6.
Znaleźć zbiory:

$A \cup B,\ A \cup C,\ B \cup C,\ A \cup B \cup C,\ A \cap B,\ A \cap C,\ B \cap C,\ A \cap B \cap C,$

$$B \setminus A,\ A \setminus C,\ C \setminus A.$$

Rozwiązanie

Podane w zadaniu zbiory są postaci:

$$A = \{3, 6, 9, 12, ...\},\ B = \{5, 10, 15, 20, ...\},\ C = \{6, 12, 18, 24, ...\}.$$

Sumę dwóch zbiorów stanowią wszystkie elementy, które należą do jednego lub drugiego zbioru, zatem $A \cup B$ stanowią liczby podzielne przez 3 lub podzielne przez 5, $B \cup C$–podzielne przez 5 lub podzielne przez 6, $A \cup C$–podzielne przez 3 lub podzielne przez 6.

$$A \cup B = \{3, 5, 6, 9, 10, 12, 15, ...\},\ B \cup C = \{5, 6, 10, 12, 15, 18, 20, ...\},$$

$$A \cup C = \{3, 6, 9, 12, 15, 18, ...\}.$$

Zbiór C zawiera się w zbiorze A (wszystkie liczby podzielne przez 6 są podzielne również przez 3–wynika to faktu, że 3 jest dzielnikiem 6), więc sumę tych zbiorów stanowi cały zbiór A. $A \cup C = A$, zatem $A \cup B \cup C$ to zbiór liczb podzielnych przez 3 lub przez 5.

$$A \cup B \cup C = A \cup B = \{3, 5, 6, 9, 10, 12, 15, ...\}.$$

Iloczyn dwóch zbiorów stanowią wszystkie elementy, które należą do jednego oraz do drugiego zbioru. Do zbioru $A \cap B$ należą liczby podzielne przez 3 i 5 (będące wielokrotnością NWW(3,5) = 15), do zbioru $B \cap C$–podzielne przez 5 i 6, czyli podzielne przez NWW(5,6) = 30, a do zbioru $A \cap C$–podzielne przez 3 oraz 6 (NWW(3,6)=6).
Ponownie korzystamy z tego, że zbiór C zawiera się w zbiorze A, dlatego iloczyn tych zbiorów stanowi zbiór C.

$$A \cap B = \{15, 30, 45, 60, ...\},\ B \cap C = \{30, 60, 90, 120, ...\},$$

$$A \cap C = \{6, 12, 18, 24, ...\},$$

$$A \cap B \cap C = B \cap C = \{30, 60, 90, 120, ...\}.$$

Zgodnie z określeniem różnicy dla dwóch zbiorów wybieramy te elementy, które należą do zbioru pierwszego i nie należą do drugiego. Zbiór $B \setminus A$ stanowią liczby podzielne przez 5, ale niepodzielne przez 3, $A \setminus C$–podzielne przez 3 oraz niepodzielne przez 6, $C \setminus A$–podzielne przez 6 i niepodzielne przez 3.

$$B \setminus A = \{5, 10, 20, 25, 35, , ...\},\ A \setminus C = \{3, 9, 15, 21, 27, ...\}.$$

Zbiór C zawiera się w zbiorze A, zatem $C \setminus A = \emptyset$.

Zadanie 1.1.2 *Udowodnić następującą równoważność:*
$(B \setminus A) \cup A = B \Leftrightarrow A \subset B.$

Rozwiązanie

Niech x oznacza zdarzenie elementarne.
$x \in (B \setminus A) \cup A \Leftrightarrow (x \in B \ \wedge \ x \notin A) \vee x \in A \Leftrightarrow$
$\Leftrightarrow (x \in B \ \vee \ x \in A) \wedge (x \in A^C \ \vee \ x \in A) \Leftrightarrow (x \in (A \cup B)) \vee x \in \Omega \Leftrightarrow$
$\Leftrightarrow x \in (A \cup B)$
Otrzymujemy zatem równość $A \cup B = B$, która jest prawdziwa wtedy i tylko wtedy, gdy $A \subset B$.

Zadanie 1.1.3 *Zbadano grupę 50 osób i okazało się, że 43 osoby uprawiają sport, 19 osób gra w gry strategiczne, a 20 gra w gry losowe. Wiedząc, że wszyscy grają w jakieś gry, wykazać, że co najwyżej 32 osoby spośród badanych uprawia więcej niż jeden rodzaj gier.*

Rozwiązanie

Niech:
A–zbiór osób uprawiających sport;
B–zbiór osób grających w gry strategiczne;
C–zbiór osób grających w gry losowe.

Z danych w zadaniu wynika, że
$|A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C| = 50.$
Przekształcamy do postaci:
$|A| + |B| + |C| + |A \cap B \cap C| - 50 = |A \cap B| + |A \cap C| + |B \cap C|$ Zbiór $A \cap B \cap C$ nie musi być zbiorem pustym, zatem
$|A| + |B| + |C| - 50 \geq |A \cap B| + |A \cap C| + |B \cap C|$
Wykorzystując dane w zadaniu mamy
$43 + 19 + 20 - 50 \geq |A \cap B| + |A \cap C| + |B \cap C|$
$32 \geq |A \cap B| + |A \cap C| + |B \cap C|$

Zadanie 1.1.4 *Niech* $\Omega = \{0, 2, 4, 6, 8\}$. *Znaleźć najmniejszą* σ*-algebrę* $\mathcal{S}$ *zawierającą rodzinę* $\mathcal{R} = \{\{0\}, \{2, 4, 6\}\}$.

Rozwiązanie

Zauważmy, że ponieważ Ω jest zbiorem skończonym, warunek zamknięcia rodziny na branie sum przeliczalnych oznacza zamknięcie tej rodziny na branie tylko sumy mnogościowej.

Oczywiście rodzina nasza nie jest σ-algebrą. Aby znaleźć najmniejszą σ-algebrę zawierającą $\mathcal{R}$, należy ją powiększyć o brakujące zbiory. Wprost z definicji wynika, że musi to być zbiór pusty oraz dopełnienia elementów tej rodziny. W wyniku powiększenia dostaniemy wtedy rodzinę

$$\{\emptyset,\ \{0\},\ \{2,\ 4,\ 6\},\ \Omega,\ \{2,\ 4,\ 6,\ 8\},\ \{0,\ 8\}\}.$$

Zauważmy, że powstała rodzina dalej nie jest σ-algebrą–nie jest zamknięta na branie sum mnogościowych. Należy ją uzupełnić o zbiór

$$\{0\} \cup \{2,\ 4,\ 6\} = \{0,\ 2,\ 4,\ 6\}$$

i jego dopełnienie $\{8\}$. Z konstrukcji wynika, że tak powstała rodzina $\mathcal{S}$ jest już σ-algebrą. Pozostało uzasadnić, że jest najmniejszą rodziną o tej własności.
W tym celu weźmy σ-algebrę $\mathcal{A}$ zawierającą rodzinę $\mathcal{R}$. Z definicji wynika, że musi ona zawierać również naszą powiększoną rodzinę, która jest σ-algebrą, co kończy dowód.

1.2 Zadania

Zadanie 1.2.1 *Niech*
A będzie zbiorem punktów (x, y), dla których $x^2 + y^2 < 4$,
B–zbiorem punktów (x, y), dla których $x^2 + y^2 < 9$,
C–zbiorem punktów (x, y), dla których $(x-1)^2 + (y+1)^2 < 1$.
Znaleźć zbiory:

$$A \cup B,\ A \cup C,\ B \cup C,\ A \cup B \cup C,\ A \cap B,\ A \cap C,\ B \cap C,\ A \cap B \cap C,$$

$$A \setminus B,\ B \setminus A,\ A \setminus C.$$

Zadanie 1.2.2 *W loterii znajdują się losy puste i wygrywające. Kupujemy trzy losy. Niech*
A oznacza zdarzenie: dokładnie jeden los wygrywa,
B–co najwyżej jeden los wygrywa,
C–co najmniej jeden los wygrywa.
Wyjaśnić, co oznaczają zdarzenia A^C, B^C, C^C, $A \cup B$, $A \cap B$, $B \cup C$, $B \cap C$, $B^C \cap C^C$.

Zadanie 1.2.3 *Niech A będzie zbiorem tych studentów Twojej grupy, których nazwisko zaczyna się od litery K,M lub P, B–zbiorem tych, którzy mają trójkę z matematyki, a C–zbiorem tych, którzy otrzymują stypendium za wyniki w nauce. Zapisać za pomocą działań na zbiorach A, B, C następujące zbiory studentów:*

1. *zbiór studentów, których nazwisko zaczyna się od litery K,M lub P, mają trójkę z matematyki oraz otrzymują stypendium za wyniki w nauce,*

2. *zbiór tych studentów, których nazwisko zaczyna się od litery K,M lub P oraz mają trójkę z matematyki lub otrzymują stypendium za wyniki w nauce,*

3. *zbiór tych studentów, których nazwisko nie zaczyna się od litery K,M lub P, mają trójkę z matematyki i nie otrzymują stypendium za wyniki w nauce,*

4. *zbiór tych studentów, których nazwisko zaczyna się od litery K,M lub P, nie mają trójki z matematyki ani nie otrzymują stypendium za wyniki w nauce.*

Zadanie 1.2.4 *Korzystając z praw de Morgana, zapisz* $A \cap B$ *wyłącznie przy pomocy operacji sumy i dopełnienia.*

Zadanie 1.2.5 *Dane są zbiory* $A = \{a, b, c\}$ *oraz* $B = \{1, 2\}$*. Wyznaczyć* $A \times B$ *oraz* $B \times A$*.*

Zadanie 1.2.6 *Udowodnić, że jeśli zbiór* A *ma* n *elementów, zbiór* B *ma* m *elementów, to* $A \times B$ *ma* $n \cdot m$ *elementów.*

Zadanie 1.2.7 *Zbiory* A *i* B *są zbiorami zawartymi w pewnej 20-to elementowej przestrzeni. Wiemy, że zbiór* A *ma 7 elementów, zbiór* B*–8 elementów, a ich część wspólna ma 3 elementy. Z ilu elementów składają się zbiory:* $A \cup B$, $A^C \cup B^C$, $A \setminus B$, $A^C \cap B$*?*

Zadanie 1.2.8 *Niech* $\Omega = [-1, 1]$*. Znaleźć najmniejszą* σ*-algebrę* $\mathcal{S}$ *zawierającą rodzinę* $\mathcal{R} = \{[-1, -\frac{1}{2}], [-\frac{1}{2}, 0]\}$*.*

Zadanie 1.2.9 *Niech* $X = [0, 1]$ *będzie przestrzenią. Dane są zbiory* $A = [0, \frac{1}{2})$ *oraz* $B = (\frac{1}{4}, 1]$*.*

1. *Wyznaczyć zbiory* $A \cup B$, $A \cap B$, $A - B$, $B - A$, A^C, B^C*.*

2. *Uzupełnić rodzinę* $\mathcal{A} = \{A, B\}$ *tak, aby była algebrą.*

Zadanie 1.2.10 *Ze zbioru cyfr* $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ *losujemy bez zwracania kolejno trzy cyfry i tworzymy liczbę trzycyfrową w ten sposób, że pierwsza wylosowana cyfra oznacza liczbę setek, druga–liczbę dziesiątek i trzecia–liczbę jedności. Za zdarzenie elementarne przyjmujemy liczbę w ten sposób uzyskaną. Wypisać wszystkie zdarzenia elementarne sprzyjające danemu zdarzeniu, jeśli zdarzenie polega na tym, że otrzymana liczba jest:*

1. *wielokrotnością liczby 25;*

2. *mniejsza od 200 i podzielna przez 5;*

3. *kwadratem pewnej liczby naturalnej;*

4. *kwadratem pewnej liczby naturalnej lub jest wielokrotnością liczby 25;*

5. *kwadratem pewnej liczby naturalnej i wielokrotnością liczby 25;*

6. *wielokrotnością liczby 25 i nie jest kwadratem żadnej liczby naturalnej;*

7. *kwadratem pewnej liczby naturalnej i nie jest wielokrotnością liczby 25.*

Zadanie 1.2.11 *Dane są 4 zdarzenia losowe A, B, C i D przestrzeni wszystkich zdarzeń elementarnych Ω. Zapisać za pomocą odpowiednich działań zdarzenie:*

1. *E polegające na tym, że wystąpiły tylko zdarzenia A i B;*

2. *F polegające na tym, że nie wystąpiło żadne zdarzenie;*

3. *G polegające na tym, że wystąpiło zdarzenie A i nie wystąpiły zdarzenia B, C, D;*

4. *H polegające na tym, że wystąpiło tylko zdarzenie spośród zdarzeń A, B, C, D.*

Zadanie 1.2.12 *Dane są zdarzenia losowe A,B i C w przestrzeni wszystkich zdarzeń elementarnych Ω. Wykazać równości:*

1. $A - (B \cup C) = (A - B) - C$;

2. $A - (B \cup C) = (A - B) \cap (A - C)$.

Zadanie 1.2.13 *Wykonujemy trzykrotny rzut symetryczną monetą. Zdarzenie elementarne to trójka (x, y, z), gdzie $x, y, z \in \{o,r\}$, gdzie o oznacza, że wyrzuciliśmy orła, r– wyrzuciliśmy reszkę. Niech A_i będzie zdarzeniem elementarnym polegającym na tym, że reszka wypadła tylko w i-tym rzucie, $i = \{1, 2, 3\}$. Wyznaczyć: Ω, A_1, A_2, A_3.*
Co oznacza zdarzenie:

1. $A_1 \cup A_2 \cup A_3$;

2. $A_1 \cap A_2 \cap A_3$;

3. $A_1^C \cup A_2^C \cup A_3^C$;

4. $A_1^C \cap A_2^C \cap A_3^C$?

Zadanie 1.2.14 *Rzucamy dwa razy symetryczną monetą. W sytuacji gdy otrzymamy dwukrotnie tę samą stronę monety, rzucamy po raz trzeci. Za zdarzenie elementarne uważamy uporządkowane dwójki lub trójki wyników poszczególnych rzutów. Wyznaczyć Ω.*

Zadanie 1.2.15 *Z talii 52 kart wybieramy 4. Niech A oznacza zdarzenie, że wylosowaliśmy co najmniej jednego asa, B–wylosowaliśmy co najwyżej jednego asa czarnego, C–wylosowaliśmy dwa asy.*
Opisać przestrzeń zdarzeń elementarnych dla tego doświadczenia. Co oznaczają zdarzenia: $A \cup B$, $A \cup C$, $A \cap B$, $A \cap C^C$, $A \cup B \cup C$, $A^C \cap B^C$*?*

Zadanie 1.2.16 *Z badań statystycznych wynika, że w pewnej grupie studentów 50% gra w koszykówkę, 60% w siatkówkę, 50% w piłkę nożną, 30% w koszykówkę oraz siatkówkę, 20% w siatkówkę i piłkę nożną, 30% w koszykówkę i piłkę nożną. 10% studentów uprawia wszystkie trzy dyscypliny sportowe. Jaki procent studentów uprawia dokładnie dwie gry zespołowe? Jaki procent nie uprawia żadnej gry?*

Zadanie 1.2.17 *W pewnej wsi mieszka 300 osób, z których każdy śpiewa, tańczy lub gra na gitarze. Połowa grających na gitarze tańczy, połowa tańczących śpiewa, a połowa śpiewających gra na gitarze. Wiemy, że żaden z grających na gitarze nie śpiewa i tańczy. Ile osób śpiewa, tańczy, a ile gra na gitarze?*

Zadanie 1.2.18 *Z badań statystycznych wynika, że wśród 200 ankietowanych 130 posiada psa, 80–kota, a 53–rybki. Nikt nie posiada wszystkich trzech zwierząt, ale mniej niż 40 spośród ankietowanych posiada dwa zwierzątka. Czy wyniki ankiety są rzetelne?*

Zadanie 1.2.19 *Wybieramy losowo studenta drugiego roku. Niech zdarzenie A polega na tym, że jest to kobieta, B–wybrana osoba uczęszcza na lektorat z języka angielskiego, C–wybrany student jest mieszkańcem Legnicy.*
Opisać słownie zdarzenia: $A \cap B^C$, $A \cap B \cap C^C$.
Przy jakich warunkach będzie zachodzić równość, że $A \cap B = A$*? Kiedy zachodzi równość* $A^C = B$*?*

Zadanie 1.2.20 *Z odcinka* $[0, 2]$ *wybieramy losowo i niezależnie dwa punkty. Niech A oznacza zdarzenie polegające na tym, że odległość między nimi jest mniejsza od 1. Opisać przestrzeń zdarzeń elementarnych dla tego doświadczenia. Opisać zdarzenie A.*

Zadanie 1.2.21 *Z odcinka* $[0, k]$ *wybieramy losowo i niezależnie punkty* x *i* y*. Niech A–zdarzenie polegające na tym, że* $x^2 + y^2 > \frac{k^2}{4}$*, B–zdarzenie, że funkcja* $ln(x^2 + y^2 - k)$ *jest dobrze określona.*

1. *Opisać* A, B *i* Ω.

2. *Wyznaczyć* $|A|$, $|B|$ *i* $|\Omega|$.

Rozdział 2

Elementy kombinatoryki oraz techniki zliczania

2.1 Wprowadzenie teoretyczne i przykłady

Niech dane będą zbiory:

$$A_1 = \{a_1, a_2, ..., a_{n_1}\},\ |A_1| = n_1$$

$$A_2 = \{b_1, b_2, ..., b_{n_2}\},\ |A_2| = n_2$$

$$\vdots$$

$$A_k = \{p_1, p_2, ..., p_{n_k}\},\ |A_k| = n_k$$

Ilość ciągów $(a_{j_1}, b_{j_2}, ..., p_{j_k})$ takich, że wyraz pierwszy należy do A_1, drugi do A_2, k-ty do A_k da się wyznaczyć, korzystając z twierdzenia:

Twierdzenie 2.1.1 *(Zasada wielokrotnego wyboru)*
Liczba k-elementowych ciągów, takich że $a_{j_1} \in A_1$, $b_{j_2} \in A_2, ..., x_{j_k} \in A_k$, jest równa $|A_1| \cdot |A_2| \cdot ...|A_k| = n_1 \cdot n_2 \cdot ... \cdot n_k$.

Definicja 2.1.1 *Niech dany będzie zbiór $A = \{a_1, a_2, ..., a_n\}$. Każdą funkcję, która liczbie naturalnej ze zbioru $\{1, 2, ..., k\}$ przyporządkowuje dokładnie jeden element a_m ze zbioru A nazywamy k-elementową wariacją z powtórzeniami zbioru A.*

K-elementowa wariacja z powtórzeniami zbioru A to k-krotny wybór po jednym elemencie ze zwracaniem ze zbioru A.

Twierdzenie 2.1.2 *Liczba k-elementowych wariacji z powtórzeniami zbioru n-elementowego jest równa $|V_n^k| = n^k$.*

Definicja 2.1.2 *Dla danego zbioru* $A = \{a_1, a_2, ..., a_n\}$ *każdą różnowartościową funkcję* $f : \{1, 2, ..., k\} \to \{a_1, a_2, ..., a_n\}$ $(k \leq n)$ *nazywamy k-elementową wariacją bez powtórzeń zbioru* A.

K-elementowa wariacja bez powtórzeń zbioru A to k-krotny wybór bez zwracania jednego elementu ze zbioru A.

Twierdzenie 2.1.3 *Liczba k-elementowych wariacji bez powtórzeń elementów zbioru n-elementowego* $(k \leq n)$ *wynosi* $|W_n^k| = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot (n-k+1) = \frac{n!}{(n-k)!}$.

Definicja 2.1.3 *Permutacją zbioru* $A = \{a_1, a_2, ..., a_n\}$ *nazywamy każdą funkcję różnowartościową* $f : \{1, 2, ..., n\} \to \{a_1, a_2, ..., a_n\}$. *Innymi słowy– permutacja zbioru to każde uporządkowanie elementów tego zbioru.*

Twierdzenie 2.1.4 *Liczba permutacji zbioru n-elementowego jest równa* $|P_n| = n!$.

Definicja 2.1.4 *K-elementową kombinacją n-elementowego zbioru* $(0 \leq k \leq n)$ *nazywamy dowolny k-elementowy podzbiór tego zbioru.*

Twierdzenie 2.1.5 *Liczba k-elementowych kombinacji zbioru n-elementowego jest równa* $|C_n^k| = \binom{n}{k}$.

Definicja 2.1.5 *Niech dany będzie n-elementowy zbiór, w którym mamy* k_1 *elementów typu* a_1, k_2 *elementów typu* a_2,...,k_m *typu* a_m, *przy czym elementy tego samego typu są nierozróżnialne,* $k_1 + k_2 + ... + k_m = n$. *Każde uporządkowanie tego zbioru nazywamy permutacją z powtórzeniami tego zbioru.*

Twierdzenie 2.1.6 *Liczba permutacji z powtórzeniami* k_1 *elementów typu* a_1, k_2 *typu* a_2,...,k_m *typu* a_m $(k_1 + k_2 + ... + k_m = n)$ *jest równa* $\frac{n!}{k_1! \cdot k_2! \cdot ... \cdot k_m!}$.

Definicja 2.1.6 *K-elementową kombinacją z powtórzeniami zbioru n-elementowego nazywamy każdy ciąg* $(k_1, k_2, ..., k_n)$ *taki, że* $k_j \geq 0$ *i całkowite oraz* $k_1 + k_2 + ... + k_n = k$.

Twierdzenie 2.1.7 *Liczba k-elementowych kombinacji z powtórzeniami zbioru n-elementowego jest równa* $\binom{n+k-1}{k}$.

Zadanie 2.1.1 *Studentka ma w szafie jedną bluzkę żółtą i dwie zielone oraz trzy spódnice zielone i dwie żółte. Na ile sposobów może się ubrać, tak aby bluzka oraz spódnica były tego samego koloru?*

Rozwiązanie

Możemy zauważyć, że studentka ma do wyboru dwa zestawy ubioru: żółta bluzka i żółta spódnica lub zielona bluzka i zielona spódnica. Każda z opisanych możliwości da się wyliczyć z wykorzystaniem reguły mnożenia.
Zatem mamy łącznie: $1 \cdot 2 + 2 \cdot 3 = 8$ możliwości.

Zadanie 2.1.2 *Ile możemy otrzymać różnych wyników podczas jednoczesnego rzutu kostką i monetą?*

Rozwiązanie

Wynik doświadczenia opisany jest przez parę (wynik rzutu kostką, wynik rzutu monetą).
A_1–zbiór możliwych wyników podczas rzutu kostką, $A_1 = \{1, 2, 3, 4, 5, 6\}$
A_2–zbiór możliwych wyników podczas rzutu monetą, $A_2 = \{O, R\}$
Mamy zatem $|A_1| \cdot |A_2| = 6 \cdot 2 = 12$ różnych wyników.

Zadanie 2.1.3 *Rzucamy czterema monetami. Ile istnieje wszystkich możliwych wyników rzutu?*

Rozwiązanie

Mamy dwa możliwe wyniki rzutu pojedynczą monetą–$n = 2$, zaś liczba monet $k = 4$. Zatem liczba wszystkich wyników rzutu czterema monetami wynosi $|W_n^k| = n^k = 2^4 = 16$.

Zadanie 2.1.4 *W urnie znajdują się ponumerowane kule $s_1, s_2, ..., s_9$. Zbiór kul oznaczmy $Z_9 = \{s_1, s_2, ..., s_9\}$. Z urny losujemy trzy kule bez zwracania i tworzymy w ten sposób liczby trzycyfrowe. Ile takich liczb możemy uzyskać?*

Rozwiązanie

W wyniku doświadczenia tworzymy trzyelementowe ciągi liczb, przy czym ciągi (s_1, s_2, s_3),(s_2, s_1, s_3) są różne, choć zawierają te same elementy. Utworzone w ten sposób ciągi to trzyelementowe wariacje bez powtórzeń zbioru Z_9.
Zatem liczba tych wariacji wynosi $|W_9^3| = 9 \cdot 8 \cdot 7 = 504$.

Zadanie 2.1.5 *W urnie znajdują się ponumerowane kule $s_1, s_2, ..., s_9$. Zbiór kul oznaczmy $Z_9 = \{s_1, s_2, ..., s_9\}$. Z urny losujemy trzy kule po jednej kuli, wrzucając po każdym losowaniu z powrotem do urny wylosowaną kulę. Wylosowane numery kul zapisujemy w kolejności losowania. Ile różnych liczb uzyskamy w ten sposób?*

Rozwiązanie

W wyniku losowania tworzymy trzyelementowe ciągi liczb, przy czym ciągi (s_1, s_2, s_3),(s_2, s_1, s_3) są różne, choć zawierają te same elementy. Kula po każdym losowaniu wraca do urny, więc możliwe jest uzyskanie liczby, w której cyfry będą się powtarzać (np. (s_6, s_6, s_6)). Utworzone w ten sposób ciągi to trzyelementowe wariacje z powtórzeniami zbioru Z_9. Liczba wszystkich wariacji trzyelementowych zbioru Z_9 jest równa $|V_9^3| = 9^3 = 729$.

Zadanie 2.1.6 *W przedstawieniu przygotowywanym przez kółko teatralne bierze udział trzech chłopców i cztery dziewczynki. Mamy do obsadzenia siedem ról.*

1. *Na ile sposobów zostaną rozdzielone role, gdy nie ma znaczenia płeć młodego aktora?*

2. *Na ile sposobów rodzielimy trzy role męskie i cztery kobiece pomiędzy biorących udział?*

Rozwiązanie

1. Nie mamy podziału na role w zależności od płci, zatem każdy może odegrać każdą rolę, ale jedna osoba nie może odtwarzać dwóch ról. Mamy zatem łacznie $7! = 5040$ możliwości.

2. Role męskie możemy rodzielić na $3!$ sposobów, a role kobiece na $4!$ sposobów. Ponieważ obsadę meską możemy zestawić z dowolną obsadą kobiecą, mamy więc $3! \cdot 4! = 12 \cdot 48 = 576$ możliwości.

Zadanie 2.1.7 *Mamy k kul i rozmieszczamy je w n komórkach. Kule, o których jest mowa w zadaniu, są nierozróżnialne. Na ile sposobów można to zrobić?*

Rozwiązanie

Rozmieszczenie tych kul jest wyznaczone przez podanie liczby kul
w komórkach, czyli ciąg $(k_1, k_2, ..., k_n)$, gdzie k_i–liczba kul w komórce o i-tym numerze. Nie jest istotne, w jakiej kolejności są rozmieszczone kule w danej komórce, istotna jest natomiast ich liczba.

Mamy $k_1 + k_2 + ... + k_n = k$, $k_j \geq 0$, $j = 1, 2, ..., n$. Dwa rozmieszczenia są różne, gdy odpowiednie ciągi $(k_1, k_2, ..., k_n)$ nie są identyczne.

Rozpatrujemy k-elementowe kombinacje z powtórzeniami zbioru $A = \{a_1, a_2, ..., a_n\}$. Kombinacji $(k_1, k_2, ..., k_n)$ przyporządkujemy rozmieszczenie zadane tym samym ciągiem, tzn. k_1 kul w 1. komórce, k_2 kul w 2. komórce itd.

Jeśli w kombinacji z powtórzeniami występuje k_j elementów a_j, tzn., że w komórce o numerze j jest k_j kul. Przyporządkowanie to jest wzajemnie jednoznaczne, więc liczba różnych rozmieszczeń jest równa liczbie kombinacji z powtórzeniami, czyli $\binom{n+k-1}{k}$.

Zadanie 2.1.8 *Z cyfr* $1, 2, 2, 3, 4, 4, 4$ *tworzymy liczby siedmiocyfrowe. Ile różnych liczb możemy tak zrealizować?*

Rozwiązanie

Bezpośrednio ze wzoru na liczbę permutacji n-elementowych z powtórzeniami mamy $\frac{7!}{2!\cdot 3!} = 420$.

Zadanie 2.1.9 *Ile uzyskamy różnych wyników przy rzucie pięcioma nierozróżnialnymi monetami?*

Rozwiązanie

W zadaniu tym wykorzystamy kombinacje z powtórzeniami, gdzie $n = 2$, $k = 5$: $\binom{n+k-1}{k} = \binom{2+5-1}{5} = 6$.

2.2 Zadania

Zadanie 2.2.1 *W kolejce stoi 5 dziewcząt i 5 chłopców. Na ile sposobów mogą oni ustawić się w kolejce, jeśli:*

1. *dziewczęta stoją przed chłopcami;*
2. *w kolejce żadnych dwóch chłopców nie stoi obok siebie;*
3. *nie ma znaczenia kolejność osób?*

Zadanie 2.2.2 *Ile różnych słów mających sens lub nie można utworzyć z liter słowa:*

1. *REGUŁA;*
2. *ZADANIA;*
3. *TATA?*

Zadanie 2.2.3 *Stonoga ma 50 par różnych butów. Rozróżnia tylko buty lewe od prawych. Na ile sposobów stonoga może ubrać buty?*

Zadanie 2.2.4 *Pewna restauracja reklamuje się w sposób następujący: ,,Posiadamy ponad milion zestawów obiadowych". Sprawdzono, że w rzeczywistości było tam 20 zup, 10 drugich dań, 12 przystawek, 20 deserów i 25 gatunków wina. Czy reklama tej restauracji jest błędna?*

Zadanie 2.2.5 *Mały Karolek wkłada buty. Posiada 6 par butów i zawsze kieruje się zasadami:*

1. *nigdy nie wkłada lewego buta na prawą nogę i odwrotnie;*
2. *nigdy nie wkłada dwóch butów z tej samej pary.*

Na ile sposobów może włożyć buty na obie nogi?

Zadanie 2.2.6 *W sali wykładowej jest 150 miejsc. Na wykład przyszło 42 studentów. Na ile sposobów mogą zająć miejsca na tej sali?*

Zadanie 2.2.7 *Ile różnych liczb czterocyfrowych o niepowtarzających się cyfrach można utworzyć, mając do dyspozycji:*

1. *cyfry: 1,2,3,4,5,6,7;*
2. *cyfry: 0,1,2,3,4,5,6,7,8,9?*

Zadanie 2.2.8 *Egzamin z matematyki składa się z 15 pytań testowych. Przy każdym pytaniu podane są trzy odpowiedzi i wiemy, że tylko jedna spośród podanych jest prawidłowa. Na ile sposobów możemy dokonać wyboru odpowiedzi na egzaminie, zakładając, że:*

1. *nie zostaną nam przydzielone punkty ujemne za błędną odpowiedź, więc można zawsze zakreślić jedną odpowiedź;*

2. *za błędną odpowiedź są przydzielane punkty ujemne, więc można nie zakreślać żadnej odpowiedzi?*

Zadanie 2.2.9 *Ile jest liczb siedmiocyfrowych, które są palindromiczne?*

Uwaga 2.2.1 *Liczba jest palindromiczna, jeśli czytana wspak jest tą samą liczbą.*

Zadanie 2.2.10 *Mamy n przedmiotów. Na ile sposobów możemy rozdzielić te przedmioty:*

1. *pomiędzy 3 osoby (przyjmujemy podział skrajnie niesprawiedliwy–tzn. wszystkie przedmioty mogą trafić do jednej osoby);*

2. *na 3 grupy?*

Zadanie 2.2.11 *Wykazać, że* $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$.

Zadanie 2.2.12 *Ile istnieje podzbiorów*

1. *trzyelementowych,*

2. *czteroelementowych,*

3. *siedmioelementowych*

zbioru dziesięcioelementowego?

Zadanie 2.2.13 *Na ile sposobów można rodzielić trzy jednoosobowe zaproszenia na koncert między 12 osób?*

Zadanie 2.2.14 *Przed zajęciami spotkało się 11 studentów. Ile nastąpi powitań?*

Zadanie 2.2.15 *Ile różnych prostych można przeprowadzić przez osiem punktów, z których:*

1. *żadne trzy nie są współliniowe;*

2. *trzy z nich są współliniowe oraz pozostałe 5 tworzy jedną prostą?*

Zadanie 2.2.16 *Na ile sposobów możemy wręczyć karty brydżyście, tak aby otrzymał:*

1. *co najmniej jednego asa;*
2. *5 pików, 4 kiery i 4 kara;*
3. *10,9,8,7 (kolory tych kart mogą być dowolne)?*

Zadanie 2.2.17 *Ile jest różnych rozmieszczeń 5 serwetek w 5 szufladach komody, w których:*

1. *wszystkie szuflady są zajęte;*
2. *co najmniej jedna szuflada jest pusta;*
3. *dokładnie jedna szuflada jest pusta?*

Zadanie 2.2.18 *W pudełku znajduje się 15 kul białych i 5 czarnych. Losujemy bez zwracania 5 kul. Ile istnieje sposobów wylosowania:*

1. *samych kul białych;*
2. *co najmniej jednej kuli czarnej;*
3. *dokładnie 3 kul czarnych?*

Zadanie 2.2.19 *Ile różnych wyników uzyskamy przy rzucie nierozróżnialnymi kostkami do gry, gdy mamy:*

1. *4 kostki;*
2. *n kostek?*

Rozdział 3

Model probabilistyczny Kołmogorowa

3.1 Wprowadzenie teoretyczne i przykłady

Przez *model probabilistyczny Kołmogorowa*, zwany też *przestrzenią probabilistyczną*, będziemy rozumieli następującą trójkę: (Ω, Σ, P), gdzie Ω jest niepustym zbiorem nazywanym *przestrzenią zdarzeń elementarnych*, Σ jest σ-*ciałem zdarzeń*, $P : \Sigma \to R$ jest funkcją zwaną *funkcją prawdopodobieństwa* taką, że:

1. $P(A) \in [0, 1]$, dla każdego $A \in \Sigma$;
2. $P(\emptyset) = 0$;
3. $P(A^C) = 1 - P(A)$, dla każdego $A \in \Sigma$;
4. $P(\bigcup\limits_{n \in \mathbf{N}_o} A_n) = \sum\limits_{n \in \mathbf{N}_o} P(A_n)$ dla dowolnych parami rozłącznych zdarzeń $A_n \in \Sigma$, $\mathbf{N}_o \subset \mathbf{N}$.

Fakt 3.1.1 *Dla każdej przestrzeni probabilistycznej* (Ω, Σ, P) *mamy:*

1. $P(\Omega) = 1$;
2. *jeśli* $A \subset B$, *to* $P(A) \leq P(B)$;
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, *dla dowolnych zdarzeń* A *i* B.

Przypomnimy najważniejsze przykłady przestrzeni probabilistycznych.

Przestrzeń probabilistyczna dyskretna

Przestrzeń probabilistyczną nazywamy *dyskretną*, jeśli

1. $\Omega = \{\omega_n, n \in \mathbf{N}_0\}$, gdzie $\mathbf{N}_0$–zbiór skończony (co najmniej dwuelementowy), albo nieskończony podzbiór liczb naturalnych;
2. dany jest ciąg $p_n \in (0,1)$ taki, że $\sum\limits_{n\in\mathbf{N}_0} p_n = 1$;
3. $\Sigma = \mathcal{P}(\Omega)$–rodzina potęgowa;
4. funkcja prawdopodobieństwa jest zdefiniowana następująco:
   (a) $P(\{\omega_n\}) = p_n,\ n \in \mathbf{N}_0$;
   (b) $P(A) = \sum\limits_{\omega_n\in A} p_n$ dla $A \neq \emptyset$
   (c) $P(\emptyset) = 0$.

Klasyczna definicja prawdopodobieństwa. Model jednorodny

Jeśli przestrzeń Ω składa się z n zdarzeń elementarnych, czyli $|\Omega| = n$, oraz zdarzenia elementarne $\{\omega_i\}$ są jednakowo prawdopodobne, czyli

$$P(\{\omega_1\}) = P(\{\omega_2\}) = ... = P(\{\omega_n\}) = \frac{1}{n},$$

to prawdopodobieństwo dowolnego zdarzenia A składającego się z k zdarzeń elementarnych ($|A| = k$) wyraża się równością

$$P(A) = \frac{|A|}{|\Omega|} = \frac{k}{n} = \frac{\text{liczba zdarzeń elementarnych sprzyjających zdarzeniu A}}{\text{liczba zdarzeń elementarnych przestrzeni }\Omega}$$

Powyższy wzór stanowił dawniej definicję prawdopodobieństwa zdarzenia, dlatego często nazywa się go ,,klasyczną definicją prawdopodobieństwa".

Zadanie 3.1.1 *Spośród czterech kart różnego koloru losujemy jednocześnie dwie. Jakie jest prawdopodobieństwo, że obie będą czarne lub obie czerwone?*

Rozwiązanie

$\Omega = \{\omega = \{x_1, x_2\} : x_i \in \{1,2,3,4\} \wedge i = 1,2\}$, gdzie $\{1,2,3,4\}$ jest zbiorem danych kart.

Wtedy

$$|\Omega| = \binom{4}{2} = \frac{4!}{2! \cdot (4-2)!} = \frac{4!}{2! \cdot 2!} = \frac{2! \cdot 3 \cdot 4}{2!} = 6.$$

Zdefiniujmy zdarzenie A, że obie karty będą czarne lub czerwone. Ponieważ $|A| = \binom{2}{2} + \binom{2}{2} = 1 + 1 = 2$–gdyż albo wyciągniemy pika i trefla albo karo oraz kiera. Dlatego

$$P(A) = \frac{|A|}{|\Omega|} = \frac{2}{6} = \frac{1}{3}.$$

Zadanie 3.1.2 *Sześcian o krawędzi 6 dm pomalowano, a następnie rozcięto w ten sposób, że każdą krawędź podzielono na 6 kawałków o długości 1 dm. Otrzymane w ten sposób sześcianiki wrzucono do pojemnika i wymieszano. Oblicz prawdopodobieństwo, że przy losowaniu jednego sześcianika będzie miał on pomalowane:*

1. *3 ścianki;*
2. *co najmniej 2 ścianki;*
3. *nie będzie pomalowany?*

Rozwiązanie

Przecinając pomalowany sześcian na sześcianiki o krawędzi 1 dm, otrzymamy $6^3 = 216$ sześcianików, z których będziemy mieli 8 sześcianików z pomalowanymi 3 ścianami (te, które były w wierzchołkach dużego sześcianu).

Przy każdej krawędzi będziemy mieli 4 sześcianiki z pomalowanymi 2 ścianami. Sześcian ma 12 krawędzi, więc $12 \cdot 4 = 48$–tyle sześcianików ma pomalowane dwie ściany. Na każdej ścianie dużego sześcianu mamy $4 \cdot 4 = 16$ sześcianików z jedną pomalowaną ścianą. Sześcian ma 6 ścian, więc w sumie mamy $16 \cdot 6 = 96$ sześcianików z jedną pomalowaną ścianą.

Niepomalowane sześciany to te, które leżą w całości we wnętrzu dużego sześcianu. Jest ich zatem $4^3 = 64$. Ponieważ $\Omega = \{x : x \in \{1, ..., 216\}\}$, $|\Omega| = 216$

1. A–zdarzenie polegające na tym, że sześcianik ma pomalowane 3 ściany
   $|A| = 8, \quad P(A) = \frac{|A|}{|\Omega|} = \frac{8}{216} = \frac{1}{27}$,
2. B–zdarzenie polegające na tym, że sześcianik ma pomalowane co najmniej 2 ściany
   $|B| = 48 + 8 = 56, \quad P(B) = \frac{|B|}{|\Omega|} = \frac{56}{216} = \frac{7}{27}$,

3. C–zdarzenie polegające na tym, że sześcianik nie ma pomalowanej żadnej ścianki
$|C| = 64, \quad P(C) = \frac{|C|}{|\Omega|} = \frac{64}{216} = \frac{8}{27}.$

Przestrzeń dwupunktowa

Niech $\Omega = \{0, 1\}$ oraz p będzie liczbą taką, że $p \in (0, 1)$. Określamy funkcję prawdopodobieństwa tak, że $P(\{1\}) = p$ oraz $P(\{0\}) = q$, gdzie $q = 1 - p$. Jest to tzw. *model standardowy dwupunktowy.*

Przestrzeń Bernoulliego

Przestrzeń probabilistyczną nazywamy *przestrzenią Bernoulliego*, jeśli

$$\Omega = \{0, 1, 2, ..., n\}, \ n \in \mathbf{N}$$

oraz p–parametr spełniający warunek $0 \le p \le 1$, z funkcją prawdopodobieństwa określoną ciągiem p_k:

$$p_k = \binom{n}{k} p^k \cdot q^{n-k}, \ k = 0, 1, 2, ..., n$$

gdzie $q = 1 - p$.

Zadanie 3.1.3 *Rzucamy symetryczną monetą. Ile rzutów należy wykonać, aby prawdopodobieństwo uzyskania co najmniej 1 orła było większe niż* $0,99$*?*

Rozwiązanie

Przy rzucie symetryczną monetą prawdopodobieństwo uzyskania orła w jednym rzucie wynosi $0,5$. Niech A oznacza zdarzenie, że wykonaliśmy n rzutów. Skorzystamy tu z własności prawdopodobieństwa, która mówi, że

$$P(A) = 1 - P(A^C).$$

Ponieważ

$$P(A^C) = \binom{n}{0} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^n,$$

gdzie n oznacza ilość rzutów, więc

$$P(A) = 1 - \left(\frac{1}{2}\right)^n.$$

Ale $P(A) > 0,99$, więc

$$1 - \left(\frac{1}{2}\right)^n > 0,99 \Leftrightarrow \left(\frac{1}{2}\right)^n < 0,01 \Leftrightarrow$$

$$2^n > 100 \Leftrightarrow n \geq 7.$$

Przestrzeń Poissona

Niech $\Omega = \mathbf{N} \cup \{0\}$ z funkcją prawdopodobieństwa określoną ciągiem

$$p_n = \frac{\lambda^n}{n!} \cdot e^{-\lambda},\ n \geq 0.$$

Tak określona przestrzeń jest przestrzenią probabalistyczną dyskretną nieskończoną. Nazywamy ją *przestrzenią Poissona.*

Zadanie 3.1.4 *Prawdopodobieństwo, że w pojedynczym rzucie kostką wypadnie 6 oczek wynosi $\frac{1}{6}$. Rzucamy kostką do gry tak długo, aż otrzymamy w rzucie 6 oczek. Niech k oznacza liczbę powtórzeń tego doświadczenia. Oblicz prawdopodobieństwo zdarzenia, że*

1. *doświadczenie zakończymy w piątym rzucie;*

2. *liczba powtórzeń doświadczenia będzie nie mniejsza niż 3.*

Rozwiązanie

W rozwiązaniu zadania posłużymy się funkcją prawdopodobieństwa określoną następująco:

$$p_k = q^{k-1}p,$$

gdzie $k \in \mathbf{N}$ oraz $p + q = 1$.

1. $p = \frac{1}{6}$, zatem $q = 1 - p = \frac{5}{6}$ i dlatego

$$p_5 = \left(\frac{5}{6}\right)^4 \frac{1}{6} = \frac{625}{7776}.$$

2.

$$P(A_{k\geq 3}) = 1 - P(A_{k<3}) = 1 - p_1 - p_2 = 1 - \frac{1}{6} - \frac{5}{6}\cdot\frac{1}{6} = \frac{25}{36},$$

gdzie $A_{k\geq 3}$ oznacza zdarzenie, że liczba powtórzeń wynosi co najmniej 3.

Stochastyczna niezależność zdarzeń

Mówimy, że zdarzenia A i B są *stochastycznie niezależne*, gdy

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

I ogólnie, zdarzenia $A_1, A_2, ..., A_n$ są stochastycznie niezależne, gdy prawdopodobieństwo łącznego zajścia dowolnych m $(m \leq n)$ zdarzeń spośród nich jest równe iloczynowi prawdopodobieństw tych zdarzeń, czyli

$$P(A_{i1}\cap A_{i2}\cap ... \cap A_{im}) = P(A_{i1})\cdot P(A_{i2})\cdot ... \cdot P(A_{im}).$$

Zadanie 3.1.5 *Wykazać, że jeśli zdarzenia A i B są niezależne, to niezależne są również zdarzenia A i B^C.*

Rozwiązanie

Należy zauważyć, że skoro $A\cap B\subset A$, to $A\cap B^C = A - (A\cap B)$. Zatem

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

Z niezależności zdarzeń A i B mamy

$$P(A\cap B^C) = P(A) - P(A)\cdot P(B) \Leftrightarrow P(A\cap B^C) = P(A)(1-P(B)) \Leftrightarrow$$

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

co należało pokazać.

Model warunkowy. Prawdopodobieństwo warunkowe

Niech (Ω, Σ, P) będzie przestrzenią probabilistyczną, B zdarzeniem takim, że $P(B) > 0$. Biorąc $(\Omega_B, \Sigma_B, P_B)$, gdzie

$$\Omega_B = B,\ \ \Sigma_B \text{ jest śladem } \Sigma \ na\ B$$

oraz
$$P_B(\tilde{A}) = \frac{P(A \cap B)}{P(B)}, \ \tilde{A} = A \cap B \in \Sigma_B \ dla\ pewnego\ A \in \Sigma$$
dostaniemy nową przestrzeń probabilistyczną, nazywaną *przestrzenią warunkową*. Dalej będziemy pisali $P(A|B)$ zamiast $P_B(\tilde{A})$ i czytali *prawdopodobieństwo zdarzenia A pod warunkiem zajścia zdarzenia B.*

Zadanie 3.1.6 *Niech $A \subset \Omega$ i $B \subset \Omega$. Udowodnić, że $P(A) + P(A^C \cap B) = P(B) + P(A \cap B^C)$.*

Rozwiązanie

Skorzystamy z pojęcia prawdopodobieństwa warunkowego. Dostaniemy kolejno
$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B) \cdot P(B)$$
oraz
$$P(B|A) = \frac{P(B \cap A)}{P(A)} \Rightarrow P(B \cap A) = P(B|A) \cdot P(A).$$
Stąd $P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$.
Ponieważ
$$P(A|B) = 1 - P(A^C|B)\ i\ P(B|A) = 1 - P(B^C|A)$$
dostaniemy
$$(1 - P(A^C|B)) \cdot P(B) = (1 - P(B^C|A)) \cdot P(A) \Leftrightarrow$$
$$P(B) - P(A^C|B) \cdot P(B) = P(A) - P(B^C|A) \cdot P(A).$$
Wiemy, że $P(A^C|B) \cdot P(B) = P(A^C \cap B)$ oraz $P(B^C|A) \cdot P(A) = P(B^C \cap A)$, czyli $P(B) - P(A^C \cap B) = P(A) - P(B^C \cap A)$ i ostatecznie
$$P(B) + P(B^C \cap A) = P(A) + P(A^C \cap B).$$

Twierdzenie 3.1.1 *(O prawdopodobieństwie zupełnym) Niech A będzie dowolnym zdarzeniem, zaś $\{B_1, B_2, ..., B_n\}$ partycją losową. Wtedy prawdopodobieństwo zdarzenia A wyraża się równością*
$$P(A) = P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2) + ... + P(B_n) \cdot P(A|B_n),$$
czyli
$$P(A) = \sum_{i=1}^{n} P(B_i) \cdot P(A|B_i).$$
Jest to tzw. wzór na prawdopodobieństwo zupełne.

Twierdzenie 3.1.2 *(Wzór Bayesa)*
Niech A będzie zdarzeniem takim, że $P(A) > 0$. Dla każdej partycji losowej $\{B_1, B_2, ..., B_n\}$ prawdopodobieństwa warunkowe $P(B_k|A)$ zdarzeń B_k przy warunku A ($k = 1, ..., n$) wyrażają się wzorem

$$P(B_k|A) = \frac{P(B_k) \cdot P(A|B_k)}{P(A)},$$

gdzie $P(A) = \sum\limits_{i=1}^{n} P(B_i) \cdot P(A|B_i)$.

Zadanie 3.1.7 *Trzy fabryki produkujące pewne detale zaopatrują hurtownię. Z fabryki III pochodzi 35% detali, a z fabryki I cztery razy więcej niż z fabryki II. Wśród wyprodukowanych detali z I fabryki 40% jest pierwszego gatunku, z II fabryki–65%, a z III 70% detali. Z magazynu losowo wybieramy jeden detal. Obliczyć prawdopodobieństwo tego, że*

1. *jest on z I fabryki;*
2. *jest on pierwszego gatunku;*
3. *pochodzi z fabryki III, jeśli jest pierwszego gatunku.*

Rozwiązanie

Zdefiniujmy:
A–zdarzenie polegające na tym, że detal jest pierwszego gatunku,
B_1–zdarzenie polegające na tym, że detal pochodzi z fabryki I,
B_2–zdarzenie polegające na tym, że detal pochodzi z fabryki II,
B_3–zdarzenie polegające na tym, że detal pochodzi z fabryki III.
Z treści zadania wynika, że
$P(A|B_1) = 0,4$, $P(A|B_2) = 0,65$, $P(A|B_3) = 0,7$.

1. Ponieważ

   $$B_1 \cup B_2 \cup B_3 = \Omega,\ B_1 \cap B_2 = \emptyset,\ B_2 \cap B_3 = \emptyset,\ B_1 \cap B_3 = \emptyset,\ P(B_1 \cup B_2 \cup B_3) = 1,$$

   więc dostaniemy kolejno:
   $P(B_1) + P(B_2) + P(B_3) = 1$,
   $P(B_3) = 0,35$,
   $4 \cdot P(B_2) = P(B_1)$,
   $4 \cdot P(B_2) + P(B_2) + 0,35 = 1$,
   $5 \cdot P(B_2) = 0,65 \Rightarrow P(B_2) = 0,13$,
   $P(B_1) = 4 \cdot 0,13 = 0,52$.

2. $P(A) = \sum_{n=1}^{3} P(A|B_n) \cdot P(B_n) = 0,4 \cdot 0,52 + 0,65 \cdot 0,13 + 0,7 \cdot 0,35 = 0,5375.$

3. $P(B_3|A) = \frac{P(A \cap B_3)}{P(A)} = \frac{P(A|B_3) \cdot P(B_3)}{P(A)} = \frac{0,7 \cdot 0,35}{0,5375} = \frac{0,245}{0,5375} \approx 0,456.$

Przestrzeń produktowa

Załóżmy, że dane są dwie przestrzenie probabilistyczne $(\Omega_1, \Sigma_1, P_1)$ i $(\Omega_2, \Sigma_2, P_2)$. Przez *przestrzeń produktową* rozumiemy taką przestrzeń (Ω, Σ, P), dla której $\Omega = \Omega_1 \times \Omega_2$, z σ-ciałem produktowym powstałym z Σ_1 i Σ_2. Funkcja prawdopodobieństwa P określona jest wtedy następująco:

$$P(A_1 \times A_2) = P_1(A_1) \cdot P_2(A_2).$$

I ogólnie, niech

$$(\Omega_j, \Sigma_j, P_j), \ \ j = 1, ..., n$$

będzie ciągiem przestrzeni probabilistycznych. Wtedy (Ω, Σ, P), gdzie

1. $\Omega = \Omega_1 \times ... \times \Omega_n$

2. $\Sigma = \Sigma_1 \otimes ... \otimes \Sigma_n$

3. $P(A_1 \times A_2 \times ... \times A_n) = P_1(A_1) \cdot P_2(A_2) \cdot ... \cdot P_n(A_n)$

dla $A_j \in \Sigma_j$ jest uogólnioną przestrzenią produktową dowolnej skończonej ilości przestrzeni probabilistycznych.

Zadanie 3.1.8 *Pokazać, że zdarzenia $A = A_1 \times \Omega_2$ i $B = \Omega_1 \times B_1$ są stochastycznie niezależne w przestrzeni produktowej.*

Rozwiązanie

Weźmy prawdopodobieństwo produktowe P. Należy pokazać, że $P(A \cap B) = P(A)P(B)$.

Ponieważ

$$A \cap B = (A_1 \times \Omega_2) \cap (\Omega_1 \times B_1) = A_1 \times B_1,$$

więc z definicji prawdopodobieństwa produktowego dostaniemy

$$P(A \cap B) = P_1(A_1)P_2(B_1).$$

Ale

$$P_1(A_1) = P(A) \ oraz \ P_2(B_1) = P(B),$$

co kończy dowód.

Zadanie 3.1.9 *Weźmy n kopii standardowej przestrzeni dwupunktowej oraz utwórzmy z tych kopii ich produkt. Opisać zdarzenia elementarne w tej przestrzeni. Dla $k \in \{0, \dots, n\}$ niech S_k oznacza zdarzenie w σ-ciele produktowym złożone z tych zdarzeń elementarnych, że liczba 1 pojawia się dokładnie k razy. Przyjmując, że prawdopodobieństwo pojedynczego pojawienia się liczby 1 wynosi $p \in (0,1)$, obliczyć $P(S_k)$.*

Rozwiązanie

Niech $(\Omega_o, \Sigma_o, P_o)$ oznacza przestrzeń standardową dwupunktową. Wtedy $\Omega = \Omega_o^n$ z σ-ciałem produktowym Σ i prawdopodobieństwem produktowym P. Dlatego każde zdarzenie elementarne $\omega \in \Omega$ jest ciągiem o wyrazach 0, 1 długości n. Jeśli w takim ciągu cyfra 1 pojawia się k razy, to z definicji prawdopodobieństwa produktowego dostaniemy

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

Ponieważ $\omega \in S_k \Leftrightarrow$ cyfra 1 pojawia się k razy i takich zdarzeń elementarnych w zdarzeniu S_k jest $\binom{n}{k}$, więc $P(S_k) = \binom{n}{k}p^k(1-p)^{n-k}$. Wynik taki już widzieliśmy, omawiając model Bernoulliego.

Zadanie 3.1.10 *Z talii 24 kart (od 9 do asa) losujemy jedną kartę. Niech A_1– zdarzenie polegające na tym, że wylosowaną kartą jest as, A_2–zdarzenie polegające na tym, że wylosowaną kartą jest pik. Oblicz prawdopodobieństwo tego, że wylosowaną kartą jest:*

1. *as pik;*
2. *pik, ale nie as.*

Rozwiązanie

Mamy tu do czynienia z przestrzenią dwuwymiarową określoną następująco:

$$\Omega = \{(x, y) : x \in \Omega_1, y \in \Omega_2\},$$

gdzie:
Ω_1–zbiór figur w kartach, $|\Omega_1| = 6$, Ω_2–zbiór kolorów w kartach, $|\Omega_2| = 4$.

Dostaniemy odpowiednio

1. $P(A_1 \times A_2) = P_1(A_1) \cdot P_2(A_2)$,
   $|A_1| = 1$, $P_1(A_1) = \frac{|A_1|}{|\Omega_1|} = \frac{1}{6}$.
   $|A_2| = 1$, $P_2(A_2) = \frac{|A_2|}{|\Omega_2|} = \frac{1}{4}$,
   $P(A_1 \times A_2) = P_1(A_1) \cdot P_2(A_2) = \frac{1}{24}$.

2. $P(A_1^C \times A_2) = P_1(A_1^C) \cdot P_2(A_2)$,
   $P(A_1^C \times A_2) = (1 - P_1(A_1)) \cdot P_2(A_2) = \frac{5}{24}$.

Prawdopodobieństwo geometryczne

Jeśli $\Omega = [a_1, a_2] \times [a_2, b_2]$, to typowe zdarzenie A w *dwuwymiarowej przestrzeni geometrycznej* ma postać:

$$A = \{(x, y) \in \Omega : x \in [a_1, b_1],\ d(x) \leq y \leq g(x)\},$$

gdzie d i g–funkcje ciągłe określone na odcinku $[a_1, b_1]$.
Prawdopodobieństwo zdarzenia A wynosi wtedy

$$P(A) = \frac{\int\limits_{a_1}^{b_1} (g(x) - d(x))dx}{(b_1 - a_1) \cdot (b_2 - a_2)}.$$

Zadanie 3.1.11 *Z kwadratu* $\Omega = [0, 1] \times [0, 1]$ *wybieramy punkt, którego współrzędne wynoszą* (p, q)*. Jakie jest prawdopodobieństwo, że równanie*

$$x^2 + px + q = 0$$

nie będzie miało pierwiastków rzeczywistych?

Rozwiązanie

Równanie nie ma pierwiastków rzeczywistych, gdy

$$\Delta < 0 \Leftrightarrow p^2 - 4 \cdot q < 0 \Leftrightarrow q > \frac{p^2}{4}$$

Weźmy $\Omega = \{(p, q) : 0 \leq p \leq 1 \wedge 0 \leq q \leq 1\}$. Ponieważ pole $|\Omega| = 1$, więc

$$P(A) = \int\limits_0^1 \frac{p^2}{4} dp = \frac{1}{12}.$$

3.2 Zadania

Zadanie 3.2.1 *W szafce mamy 10 par butów. Wyciągamy losowo 6 butów. Obliczyć prawdopodobieństwo, że wśród nich nie będzie żadnej pary?*

Zadanie 3.2.2 *Rzucamy dwa razy kostką do gry. Jakie jest prawdopodobieństwo, że w sumie otrzymamy*

1. *8 oczek;*
2. *co najmniej 3 oczka?*

Zadanie 3.2.3 *Obliczyć prawdopodobieństwo tego, że 7 losowo wybranych osób urodziło się w różnych dniach tygodnia.*

Zadanie 3.2.4 *Przed kolokwium z matematyki w grupie liczącej 24 studentów prowadzący zaproponował studentom, aby wpisali na kartkach po jednej liczbie naturalnej od 1 do 40. W przypadku gdy wszystkie wpisane liczby będą różne, odwołuje kolokwium i wszyscy otrzymują ocenę bardzo dobrą. Jakie jest prawdopodobieństwo, że kolokwium zostanie odwołane?*

Zadanie 3.2.5 *Uzasadnić, że prawdopodobieństwo warunkowe jest prawdopodobieństwem.*

Zadanie 3.2.6 *Wiemy, że* $P(A \cap B) = \frac{1}{4}, P(A^C) = \frac{1}{3}, P(B) = \frac{1}{2}$.
Obliczyć $P(A \cup B)$ *oraz* $P(A^C \cap B^C)$.

Zadanie 3.2.7 *Przy danych* $P(A|B) = P(B|A), P(A \cup B) = \frac{1}{2}, P(A \cap B) = \frac{1}{4}$ *obliczyć* $P(B)$ *oraz* $P(A^C \cap B^C)$.

Zadanie 3.2.8 *Na egzaminie student otrzymuje jedną z ocen: bardzo dobry, dobry, dostateczny, niedostateczny. W przypadku otrzymania oceny bardzo dobrej, dobrej, dostatecznej egzamin jest zdany. Prawdopodobieństwo tego, że student uzyska ocenę bardzo dobrą lub dobrą jest równe 0,6. Prawdopodobieństwo zdania egzaminu wynosi 0,9. Obliczyć prawdopodobieństwo otrzymania oceny dostatecznej na egzaminie.*

Zadanie 3.2.9 *Liczby* $\{1,2,\ldots,9\}$ *są ustawione w sposób losowy. Obliczyć prawdopodobieństwo tego, że*

1. *liczby 1 i 2 będą stały obok siebie;*
2. *liczba 1 będzie w tym ciągu przed liczbą 5.*

Zadanie 3.2.10 *Do windy, która zatrzymuje się na 10 piętrach, wsiada 8 osób. Obliczyć prawdopodobieństwo, że*

1. *każdy z pasażerów wysiądzie na innym piętrze;*

2. *wszyscy wysiądą na trzech ostatnich piętrach.*

Zadanie 3.2.11 *Z talii ośmiu kart, w której mamy 4 asy i 4 króle, wybieramy losowo 2 karty. Jakie jest prawdopodobieństwo, że*

1. *wybrano dwa asy, jeśli wybrano co najmniej jednego asa;*

2. *wybrano dwa asy, jeśli wiadomo, że wśród kart jest as pik.*

Zadanie 3.2.12 *Wykazać, że dla* $A, B, C \in \Omega$ *takich, że* $P(A \cap B) > 0$*, zachodzi równość* $P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|(A \cap B))$.

Zadanie 3.2.13 *Niech* $\Omega = \{(x, y) : x \in R \wedge y \in R\}$. *Zdarzenia* A *i* B *są określone następująco:*
$A = \{(x, y) : |x - 1| < 1 \wedge |y - 1| < 1\}$
$B = \{(x, y) : x^2 + y^2 \leq 2\}$.
Wyznaczyć $P(A|B)$ *i* $P(B|A)$.

Zadanie 3.2.14 *Wykazać, że jeśli* $A, B \subset \Omega$ *są zdarzeniami niezależnymi, to niezależne są też zdarzenia* A^C *i* B^C.

Zadanie 3.2.15 *Wiadomo, że średnio 4 mężczyzn na 100 i 4 kobiety na 1000 są daltonistami. Z grupy, w której liczba kobiet jest dwa razy większa od liczby mężczyzn, wybrano osobę. Jakie jest prawdopodobieństwo, że jest daltonistą? Jakie jest prawdopodobieństwo, że wybranym daltonistą jest mężczyzna?*

Zadanie 3.2.16 *Ze zbioru* $\{1,2,\ldots,n\}$*, gdzie* $n > 3$*, losujemy dwie liczby. Jakie jest prawdopodobieństwo, że jedna z nich będzie mniejsza, a druga większa od k, gdzie* $1<k<n$ *i* $k \in N$.

Zadanie 3.2.17 *W pewnym mieście mieszka 10 000 osób. Prawdopodobieństwo, że wybrana osoba będzie potrzebowała natychmiastowej pomocy lekarskiej wynosi 0,002. Obliczyć prawdopodobieństwo wezwania pogotowia przez:*

1. *któregokolwiek mieszkańca tego miasta;*

2. *więcej niż 2, ale nie więcej niż 5 mieszkańców tego miasta;*

3. *co najmniej 3 mieszkańców tego miasta.*

Zadanie 3.2.18 *Trzech dostawców dostarcza do hurtowni cytrusy. Dostawy tych dostawców mają się jak 3:5:4. I dostawca dotarcza około 1% skrzynek zepsutych owoców, II–4%, a III–3%. Wybraliśmy skrzynkę z dobrymi owocami. Jakie jest prawdopodobieństwo, że dostarczył ją dostawca I?*

Zadanie 3.2.19 *W dwóch jednakowo wyglądających kopertach znajdują się talie kart, przy czym w pierwszej jest to talia 52 kart, a w drugiej 24 kart (od 9 do asa). Z losowo wybranej kopert wyciągamy kartę i wkładamy do drugiej. Następnie wybieramy jedną kartę z drugiej koperty. Jakie jest prawdopodobieństwo, że wyciągniętą kartą będzie as?*

Zadanie 3.2.20 *Zbiór* $\{1,2,\ldots,4n\}$ *podzielono na dwie równoliczne części. Obliczyć prawdopodobieństwo tego, że w każdej z nich jest taka sama ilość liczb podzielnych przez* n.

Zadanie 3.2.21 *W dwóch jednakowych koszykach jest po 10 białych piłek tenisowych. Wkładamy do tych koszyków 20 czerwonych piłek. Jak rozmieścić te piłki w koszach, aby prawdopodobieństwo wybrania piłki czerwonej z losowo wybranego koszyka było równe* $\frac{7}{15}$*?*

Zadanie 3.2.22 *Z odcinka* $[0,2]$ *wybieramy losowo i niezależnie dwa punkty. Niech* A *oznacza zdarzenie polegające na tym, że odległość między nimi jest mniejsza od 1. Obliczyć prawdopodobieństwo zdarzenia* A.

Zadanie 3.2.23 *Z odcinka* $[0,k]$ *wybieramy losowo i niezależnie punkty* x *i* y. *Niech* A*–zdarzenie polegające na tym, że* $x^2+y^2>\frac{k^2}{4}$, B *- zdarzenie, że funkcja* $ln(x^2+y^2-k)$ *jest dobrze określona.*

1. *Czy zdarzenia* A *i* B *są niezależne?*

2. *Obliczyć prawdopodobieństwo zdarzenia* B *pod warunkiem, że zajdzie zdarzenie* A.

Rozdział 4

Rozkłady dyskretne

4.1 Wprowadzenie teoretyczne i przykłady

Przez *rozkład dyskretny* rozumiemy każdą funkcję

$$d : A \to \mathbf{R}_+,$$

gdzie

1. $$A \text{ jest zbiorem skończonym lub przeliczalnym,}$$

2. $$\sum_{a \in A} d(a) = 1.$$

Jeśli zbiór A jest skończony, to mówimy wtedy o *skończonym rozkładzie dyskretnym*. W przeciwnym razie będziemy mówili, że mamy do czynienia z *rozkładem nieskończonym*.

Przyjmijmy, że

$$A = \{a_n,\ n \in \mathbf{N}_o\},\ \text{gdzie } \mathbf{N_o} \subset \mathbf{N}.$$

Wtedy funkcję rzeczywistą

$$F : \mathbf{R} \to \mathbf{R}$$

określona wzorem

$$F(x) = \sum_{a_n < x} d(a_n) \tag{4.1}$$

nazywamy *dystrybuantą rozkładu* d.

Uwaga 4.1.1 *Sumowanie we wzorze* 4.1 *odbywa się po wszystkich tych wartościach naturalnych* n*, dla których* $a_n < x$.

Zachodzi następujące twierdzenie:

Twierdzenie 4.1.1 *Dystrybuanta* F *każdego dyskretnego rozkładu prawdopodobieństwa ma następujące własności:*

1.
$$\forall_{x \in \mathbf{R}}\ 0 \leq F(x) \leq 1,$$

2.
$$F \text{ jest funkcją niemalejącą,}$$

3.
$$\lim_{x \to -\infty} F(x) = 0,$$

4.
$$\lim_{x \to +\infty} F(x) = 1,$$

5.
$$F \text{ jest lewostronnie ciągła,}$$

6.
$$F \text{ jest przedziałami stała.}$$

Na odwrót, dla każdej funkcji rzeczywistej G mamy:

Twierdzenie 4.1.2 *Jeśli* G *ma własności jak w tezie twierdzenia 4.1.1, to istnieje dokładnie jeden dyskretny rozkład prawdopodobieństwa* $d\colon A \to \mathbf{R}_+$*, gdzie*

$$a \in A \Leftrightarrow G \text{ ma w } a \text{ nieciągłość typu skok,}$$

czyli

$$\lim_{x \to a^+} G(x) - G(a) > 0$$

oraz

$$d(a) = \lim_{x \to a^+} G(x) - G(a).$$

Ponadto dystrybuanta tego rozkładu równa jest funkcji G.

Zadanie 4.1.1 *Sprawdzić, że przyporządkowanie*

$$d(j) = \binom{n}{j} p^j q^{n-j}, \text{ dla } j = 0, 1, \ldots, n,$$

gdzie $0 < p < 1,\ p + q = 1$ *określa dyskretny rozkład prawdopodobieństwa.*

Rozwiązanie

W naszym przypadku $A = \{0, 1, \ldots, n\}$ oraz $d(j) > 0$ dla każdego $j \in A$. Wystarczy więc pokazać, że $\sum_{j \in A} d(j) = 1$. Ze wzoru dwumianowego Newtona mamy

$$(p + q)^n = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j}.$$

Ponieważ $p + q = 1$, dowodzi to równości $\sum_{j \in A} d(j) = 1$.

Zadanie 4.1.2 *Uzasadnić, że przyporządkowanie*

$$d\colon\ \mathbf{N} \cup \{0\} \to \mathbf{R},$$

gdzie

$$d(n) = \exp(-\lambda)\frac{\lambda^n}{n!}$$

definiuje nieskończony dyskretny rozkład prawdopodobieństwa dla każdej wartości $\lambda > 0$.

Rozwiązanie

Skorzystamy z rozwinięcia w szereg potęgowy funkcji $\exp(x)$

$$\forall_{x \in \mathbf{R}}\ \exp(x) = \sum_{n \geq 0} \frac{x^n}{n!}.$$

W naszym przypadku, podstawiając $x = \lambda$, dostaniemy

$$\sum_{k \geq 0} d(k) = \exp(-\lambda)\exp(\lambda) = 1,$$

co pokazuje, że d jest rozkładem.

Zadanie 4.1.3 *Podać przykład nieskończonego rozkładu dyskretnego.*

Rozwiązanie

Weźmy dowolny szereg liczbowy o wyrazach dodatnich i zbieżny, np.

$$s = \sum_{n \geq 1} \frac{1}{n^2}.$$

Wtedy szukany rozkład ma postać

$$d(j) = \frac{1}{sj^2}, \; j \in \mathbf{N}.$$

Zadanie 4.1.4 *Dany jest rozkład*

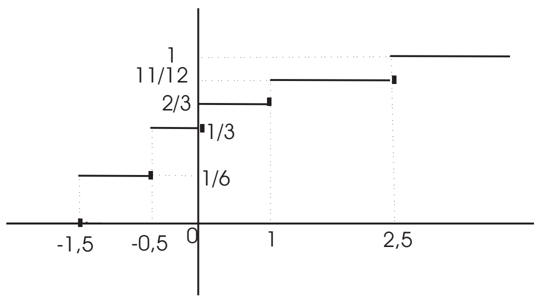
$$d\colon \{-1,5, \; -0,5, \; 0, \; 1, \; 2,5\} \to \mathbf{R},$$

gdzie

$$d(-1,5) = \frac{1}{6}, \; d(-0,5) = \frac{1}{6}, \; d(0) = \frac{1}{3}, \; d(1) = \frac{1}{4}, \; d(2,5) = \frac{1}{12}.$$

Narysować dystrybuantę tego rozkładu.

Rozwiązanie



Rysunek 4.1: Dystrybuanta rozkładu d

Na rysunku tym wyraźnie zaznaczono zjawisko lewostronnej ciągłości dystrybuanty. Ze względu na jej kształt nazywamy ją *funkcją schodkową*.

Możemy więc napisać:

Fakt 4.1.1 *Istnieje wzajemnie jednoznaczna odpowiedniość pomiędzy rozkładami dyskretnymi a dystrybuantami typu schodkowego. Odpowiedniość ta opisana jest w twierdzeniach 4.1.1 i 4.1.2.*

Czas na wyjaśnienie roli terminu *prawdopodobieństwa* pojawiającego się w nazwie rozkładu dyskretnego.

Zachodzi następujące twierdzenie (patrz uwaga powyżej):

Twierdzenie 4.1.3 *Dla każdej dystrybuanty schodkowej F istnieje model probabilistyczny Kołmogorowa $\Big(\Omega, \Sigma, P\Big)$ oraz odwzorowanie $\mathbf{X}$ takie, że:*

1.
$$\mathbf{X}\colon \Omega \to \mathbf{R},$$

2.
$$\mathbf{X}(\Omega) = A,$$
gdzie A jest zbiorem punktów nieciągłości dystrybuanty F,

3.
$$\forall_{t\in\mathbf{R}}\ \{\omega \in \Omega\colon\ \mathbf{X}(\omega) < t\} \in \Sigma,$$

4.
$$P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < t\}\Big) = F(t)$$
dla każdego rzeczywistego t.

Dalej każde odwzorowanie $\mathbf{X}$, o którym mowa jest w powyższym twierdzeniu, w punktach (1) i (3) będziemy nazywali *zmienną losową*. Jak pamiętamy, model probabilistyczny Kołmogorowa opisuje zjawisko losowe, dokładniej, zdarzenie elementarne $\omega \in \Omega$ opisuje stan (wynik) takiego zjawiska, który nie zawsze musi mieć charakter wymierny–czyli nie musi być liczbą. Rola zmiennej losowej polega na tym że pełni ona funkcję *translatora*, tłumacząc stan zjawiska losowego ω na liczbę rzeczywistą $\mathbf{X}(\omega)$.

Co więcej, jeśli d jest dyskretnym rozkładem prawdopodobieństwa odpowiadającym dystrybuancie F, to ponieważ dla każdego t_o

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \leq t_o\}\Big) = \lim_{t\to t_o^+} F(t), \tag{4.2}$$

z własności funkcji prawdopodobieństwa dostaniemy

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = t_o\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \leq t_o\}\Big) - P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < t_o\}\Big),$$

co na mocy twierdzenia 4.1.2 oznacza, że

$$d(t_o) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = t_o\}\Big),\ t_o \in A. \tag{4.3}$$

Na odwrót:

Twierdzenie 4.1.4 *Niech* $\mathbf{Y}$ *będzie zmienną losową zadaną na przestrzeni probabilistycznej* (Ω, Σ, P) *taką, że:*

$$\mathbf{Y}(\Omega) = B \textit{ jest zbiorem przeliczalnym}$$

Wtedy przyporządkowanie d *określone na zbiorze* B *wzorem*

$$d(b) = P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = b\}\Big)$$

jest dyskretnym rozkładem prawdopodobieństwa. Co więcej, biorąc zgodnie z twierdzeniem 4.1.3 zmienną losową $\mathbf{X}$ *odpowiadającą temu rozkładowi, dostaniemy*

$$F_{\mathbf{X}} = F_{\mathbf{Y}},$$

co oznacza równośc dystrybuant dla tych zmiennych.

Twierdzenie 4.1.4 i fakt 4.1.1 pozwalają nam dyskretny rozkład prawdopodobieństwa d traktować równoważnie z dystrybuantą tego rozkładu i zmienną losową, o której mówimy wtedy, że *ma rozkład* d, co zapisujemy

$$d(\mathbf{X}) = F_{\mathbf{X}}. \tag{4.4}$$

Zadanie 4.1.5 *Niech* $\mathbf{X}$ *ma rozkład jak w zadaniu 4.1.4. Bierzemy zmienną losową* $\mathbf{Y}$ *określoną następująco* $\mathbf{Y} = \mathbf{X}^2 + 1$. *Znaleźć rozkład tej zmiennej losowej.*

Rozwiązanie

Z twierdzenia 4.1.3 istnieje przestrzeń probabilistyczna (Ω, Σ, P), dla której

$$\mathbf{X}(\Omega) = \{-1,5,\ -0,5,\ 0,\ 1,\ 2,5\}.$$

Ponieważ złożenie zmiennej losowej z funkcją ciągłą jest zmienną losową (u nas funkcja ta to $x \to x^2 + 1$), więc $\mathbf{Y}$ jest zmienną losową oraz

$$\mathbf{Y}(\Omega) = \{1,\ 1,25,\ 2,\ 3,25,\ 7,25\}.$$

Ponadto

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = 1\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}^2{+}1(\omega) = 1\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 0\}\Big),$$

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = 1,25\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}^2 + 1(\omega) = 1,25\}\Big) =$$
$$P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = -0,5\}\Big),$$

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = 2\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}^2{+}1(\omega) = 2\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 1\}\Big),$$

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = 3,25\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}^2 + 1(\omega) = 3,25\}\Big) =$$
$$P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = -1,5\}\Big),$$

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = 7,25\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}^2 + 1(\omega) = 7,25\}\Big) =$$
$$P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 2,5\}\Big),$$

co wyznacza rozkład zmiennej losowej $\mathbf{Y}$.

Zadanie 4.1.6 *Dane są dwie zmienne losowe* $\mathbf{X}$ *i* $\mathbf{Y}$ *każda o rozkładzie dyskretnym skończonym. Wyznaczyć rozkład sumy tych zmiennych losowych.*

Rozwiązanie

Niech $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$. Wtedy

$$\mathbf{Z}(\Omega) = \mathbf{X}(\Omega) \oplus \mathbf{Y}(\Omega),$$

gdzie symbolem $\oplus$ oznaczyliśmy działanie tzw. *sumy kompleksowej.* W naszym przypadku, zakładając, że

$$\mathbf{X}(\Omega) = \{x_1, \ldots, x_n\},\ \mathbf{Y}(\Omega) = \{y_1, \ldots, y_m\}$$

dostaniemy

$$\mathbf{Z}(\Omega) = \{z_{ij} = x_i + y_j,\ i = 1, 2, \ldots, n,\ j = 1, 2, \ldots,\ m\}.$$

Aby wyznaczyć rozkład zmiennej losowej $\mathbf{Z}$ musimy podać wartości prawdopodobieństw

$$p_{ij} = P\Big(\{\omega \in \Omega\colon\ \mathbf{Z}(\omega) = z_{ij}\}\Big),$$

czyli

$$p_{ij} = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\} \cap \{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = y_j\}\Big).$$

Korzystając z pojęcia *prawdopobieństwa warunkowego* ostatnią równość, możemy zapisać następująco

$$p_{ij} = \frac{P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\} \cap \{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = y_j\}\Big)}{P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = y_j\}\Big)} P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = y_j\}\Big). \tag{4.5}$$

Ostatecznie

$$p_{ij} = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\} | \{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = y_j\}\Big) P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = y_j\}\Big). \tag{4.6}$$

W pewnych sytuacjach prawdopodobieństwo p_{ij} dane 4.6 można wyznaczyć bezpośrednio za pomocą rozkładów zmiennych losowych składowych.

Powiemy, że zmienne losowe $\mathbf{X}$ i $\mathbf{Y}$ o rozkładzie dyskretnym są *stochastycznie niezależne*, jeśli

$$p_{ij} = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\}\Big) P\Big(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = y_j\}\Big)$$

dla dowolnych $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

Zadanie 4.1.7 *Niech $\mathbf{X}_1, \mathbf{X}_2, \ldots\ \mathbf{X}_n$ będzie ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie dwupunktowym. Znaleźć rozkład sumy tych zmiennych losowych.*

Rozwiązanie

Z założenia dla każdego $j = 1, 2, \ldots, n$

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{X}_j(\omega) = 1\}\Big) = p,\ P\Big(\{\omega \in \Omega\colon\ \mathbf{X}_j(\omega) = 0\}\Big) = q,$$

gdzie $0 < p < 1,\ q = 1 - p$.

Niech $\mathbf{X} = \sum_{i=1}^{n} \mathbf{X}_i$. Wtedy $\mathbf{X}(\Omega) = \{0, 1, \ldots, n\}$ oraz z założenia o niezależności zmiennych losowych $\mathbf{X}_i$ (patrz 4.6) dostaniemy dla $0 \le k \le n$

$$P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = k\}\Big) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Istotnie, to że zmienna losowa $\mathbf{X}$ przyjmuje wartość k, oznacza, że dokładnie k spośród n zmiennych losowych składowych przyjmuje wartość 1, pozostałe wartość 0. Zdarzenie takie można opisać jako zbiór wszystkich ciągów $0 - 1$ długości n, w których liczba jeden pojawia się dokładnie k razy. Ponieważ z niezależności zmiennych losowych składowych prawdopodobieństwo każdego takiego ciągu wynosi $p^k(1-p)^{n-k}$ i liczba takich ciągów wynosi $\binom{n}{k}$, więc z własności funkcji prawdopodobieństwa dostaniemy powyższy wzór.

Jak już wiemy, rozkład zmiennej losowej $\mathbf{X}$ jest taki sam jak rozkła Bernoulliego z parametrami n i p.

4.2 Zadania

Zadanie 4.2.1 *Rozważmy rzut kostką symetryczną. Niech* $\mathbf{X}$ *będzie zmienną losową przyjmującą wartość* 1*, gdy liczba oczek na kostce jest parzysta oraz* 0 *gdy jest nieparzysta.*

1. *Opisać* $\mathbf{X}$ *jako funkcję* $\mathbf{X}\colon \Omega \to \mathbf{R}$*. Czy jest to zmienna typu dyskretnego?*
2. *Znaleźć rozkład zmiennej losowej* $\mathbf{X}$*.*
3. *Narysować jej dystrybuantę.*

Zadanie 4.2.2 *Rzucamy w sposób niezależny pięciokrotnie monetą asymetryczną. Niech* $\mathbf{X}$ *zwraca wartość będącą liczbą orłów w pojedynczej serii rzutów.*

1. *Uzasadnić, że* $\mathbf{X}$ *jest zmienną losową typu dyskretnego.*
2. *Znaleźć* $\mathbf{X}(\Omega)$*.*
3. *Wyznaczyć rozkład tej zmiennej losowej.*
4. *Obliczyć prawdopodobieństwo* $P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \leq 3\}\Big)$*.*

Zadanie 4.2.3 *Niech* F *będzie dystrybuantą zmiennej losowej* $\mathbf{X}$ *o rozkładzie dyskretnym. Dla* $a < b$ *obliczyć:*

1. $$P\Big(\{\omega \in \Omega\colon\ a \leq \mathbf{X}(\omega) < b\}\Big),$$
2. $$P\Big(\{\omega \in \Omega\colon\ a < \mathbf{X}(\omega) \leq b\}\Big),$$
3. $$P\Big(\{\omega \in \Omega\colon\ a \leq \mathbf{X}(\omega) \leq b\}\Big).$$

Zadanie 4.2.4 *Niech zmienna losowa* $\mathbf{X}$ *ma rozkład dany funkcją dystrybuanty*

$$F(x) = \begin{cases} 0 & x \leq 1,5 \\ 0,25 & 1,5 < x \leq 3 \\ 0,75 & 3 < x \leq 5 \\ 1 & 5 \leq x. \end{cases}$$

1. *Narysować wykres* F*.*

2. *Obliczyć* $P\Big(\{\omega \in \Omega\colon\ 0,5 \le \mathbf{X}(\omega) < 0,75\}\Big)$.

3. *Obliczyć* $P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > 2,5\}\Big)$.

4. *Obliczyć* $P\Big(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) < 2\}\Big)$.

Zadanie 4.2.5 *Niech* $\mathbf{Y} = \mathbf{X}^2 - 1$, *gdzie* $\mathbf{X}$ *ma rozkład jak w zadaniu powyżej. Znaleźć rozkład zmiennej losowej* $\mathbf{Y}$. *Narysować jej dystrybuantę.*

Zadanie 4.2.6 *Dla* $j \in \Omega = \mathbf{N} \cup \{0\}$ *i* $0 < r < 1$ *definiujemy rozkład* d

$$d(j) = (1-r)^j r.$$

1. *Sprawdzić, że* d *jest rozkładem prawdopodobińswa.*

2. *Obliczyć* $P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \ge 8\}\Big)$.

Zadanie 4.2.7 *Dla* $j \in \Omega = \mathbf{N} \cup \{0\}$ *definiujemy funkcję*

$$dj = \begin{cases} \frac{3}{4}(\frac{1}{4})^j & j \in \mathbf{N} \\ x & j = 0 \end{cases}$$

Dobrać tak wartość x, *aby* d *był rozkładem prawdopodobieństwa.*
Obliczyć $P\Big(\{\omega \in \Omega\colon\ 1 \le \mathbf{X}(\omega) \le 12\}\Big)$.

Zadanie 4.2.8 *Niech* $\mathbf{X}$ *będzie zmienną losową o rozkładzie dyskretnym taką, że*

$$\exists_{a \in \mathbf{R}}\ 0 < P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = a\}\Big) < 1.$$

Uzasadnić, że $\mathbf{X}$ *jest stochastycznie niezależna od zmiennej losowej stałej.*

Zadanie 4.2.9 *Kiedy zmienna losowa o rozkładzie dyskretnym jest stochastycznie niezależna od samej siebie?*

Zadanie 4.2.10 *Dla jakich liczb rzeczywistych* a, b, c

$$F(x) = \begin{cases} 0 & dla\ x \le 1 \\ b(1 - c/x) & dla\ 1 < x \le a \\ 1 & dla\ x > a \end{cases}$$

jest dystrybuantą rozkładu dyskretnego?

Zadanie 4.2.11 *Dany jest rozkład d*

x_i	-2	-1	0	1	2
$p_i =$	1/3	1/6	2/9	1/6	1/9

Narysować dystrybuantę tego rozkładu. Dla rozkładu z powyższego zadania obliczyć $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < 1/2\})$.

Zadanie 4.2.12 *Dany jest rozkład d*

x_i	-3	-1	1	2	3
$p_i =$	1/4	1/4	1/6	1/6	1/6

Obliczyć $P(\{\omega \in \Omega : \ 1,5 \leq X(\omega) < 2,5\})$.

Zadanie 4.2.13 *Dane są dwie zmienne losowe* $\mathbf{X}$, $\mathbf{Y}$, *gdzie*

$$\mathbf{X}(\Omega) = \{1, 2, 3, 4, 5\}, \ \mathbf{Y}(\Omega) = \{0, \ 1,5, \ 2, \ 5,5\}.$$

Definiujemy dwie nowe zmienne losowe:

$$\mathbf{Z}_1 = min\{\mathbf{X}, \mathbf{Y}\}, \ \mathbf{Z}_2 = max\{\mathbf{X}, \mathbf{Y}\}.$$

1. *Znaleźć rozkłady tych zmiennych losowych.*
2. *Obliczyć* $P\Big(\{\omega \in \Omega\colon\ \mathbf{Z}_1(\omega) \leq \mathbf{Z}_2(\omega)\}\Big)$.

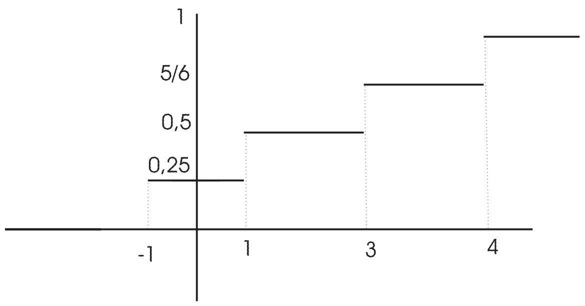
Zadanie 4.2.14 *Narysować dystrybuantę zmiennej losowej* $\mathbf{Z} = \mathbf{XY}$, *gdzie zmienne losowe* $\mathbf{X}$ *i* $\mathbf{Y}$ *określone są jak w powyższym zadaniu.*

Zadanie 4.2.15 *Podać przykład dwóch różnych zmiennych losowych o rozkładach dyskretnych i o jednakowych dystrybuantach.*

Zadanie 4.2.16 *Zmienna losowa* $\mathbf{X}$ *ma dystrybuantę jak na rysunku 4.2*

1. *Opisać rozkład zmiennej losowej* $\mathbf{X}$.
2. *Obliczyć* $P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \leq 4\}\Big)$.

Zadanie 4.2.17 *Zaproponować model probabilistyczny Kołmogorowa* (Ω, Σ, P) *taki, że zmienna losowa* $\Omega \ni \omega \longrightarrow \mathbf{Y}(\omega) \in \mathbf{R}$ *ma rozkład jak w zadaniu powyżej.*



Rysunek 4.2: Dystrybuanta zmiennej losowe $\mathbf{X}$

Zadanie 4.2.18 *Niech* (Ω, Σ, P) *będzie przestrzenią produktową, gdzie*

$$\Omega_1 = \Omega_2 = \{1, 2, 3, 4\},$$

$$P_i(\{j\}) = p_{ij}, \;\; i = 1, 2; \;\; j = 1, 2, 3, 4.$$

Definiujemy zmienne losowe:

$$\mathbf{X}(\omega_1, \omega_2) = min\{\omega_1, \omega_2\},$$

$$\mathbf{Y}(\omega_1, \omega_2) = max\{\omega_1, \omega_2\},$$

$$\mathbf{Z}(\omega_1, \omega_2) = \omega_1 + \omega_2.$$

1. *Znaleźć rozkłady tych zmiennych losowych.*
2. *Czy zmienne losowe* $\mathbf{X}$ *i* $\mathbf{Y}$ *są stochastycznie niezależne?*

Zadanie 4.2.19 *Niech* $(\mathbf{X}_n)$ *będzie ciągiem zmiennych losowych takim, że*

1.
$$\mathbf{X}_n(\Omega) = \{0, 1, 2, \ldots, n\}, \;\; n \geq 1,$$

2.
$$P\Big(\{\omega \in \Omega\colon \; \mathbf{X}_n(\omega) = j\}\Big) = e^{-\lambda}\frac{\lambda^j}{j!} + a_{nj},$$

gdzie

$$a_{nj} \in (0, 1), \;\; \sum_{j=0}^{n} \Big(e^{-\lambda}\frac{\lambda^j}{j!} + a_{nj}\Big) = 1.$$

Obliczyć $\lim_{n\to\infty} P\Big(\{\omega \in \Omega\colon\ \mathbf{X}_n(\omega) = j\}\Big)$ *dla* $j \geq 0$.

Zadanie 4.2.20 *Wiadomo, że w pewnej populacji osób stosunek liczby osób powyżej 70-roku życia do osób w wieku z przedziału* $[50, 70)$ *lat do osób w wieku z przedziału* $[25, 50)$ *lat do osób w wieku z przedziału* $[18, 24)$ *lat ma się jak 1:3:4:2.*

Niech $\mathbf{X}$ *oznacza zmienną losową przyporządkowującą losowo wybranej osobie grupę wiekową: 1, 2, 3, 4.*

1. *Znaleźć rozkład zmiennej losowej* $\mathbf{X}$.
2. *Narysować dystrybuantę tej zmiennej.*
3. *Obliczyć* $P\Big(\{\omega \in \Omega\colon\ 2 < \mathbf{X}(\omega)\} \leq 4\Big)$.

Rozdział 5

Rozkłady ciągłe

5.1 Wprowadzenie teoretyczne i przykłady

Każde odzworowanie $\mathbf{X}$ określone na przestrzeni probabilistycznej $\Big(\Omega, \Sigma, P\Big)$ o wartościach w zbiorze $\mathbf{R}$ będziemy nazywali *zmienną losową*, o ile

$$\forall_{t\in\mathbf{R}}\ \{\omega \in \Omega\colon\ \mathbf{X}(\omega) < t\} \in \Sigma.$$

Pozwala to nam zdefiniować funkcję rzeczywistą

$$F\colon\ \mathbf{R} \to \mathbf{R} \tag{5.1}$$

wzorem

$$F(x) = P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < x\}). \tag{5.2}$$

Wtedy funkcję tę nazywamy *dystrybuantą* zmiennej losowej $\mathbf{X}$. Zachodzi następujące twierdzenie

Twierdzenie 5.1.1 *(o dystrybuancie) Dla każdej zmiennej losowej* $\mathbf{X}$ *funkcja* F *dana wzorem* 5.2 *ma następujące własności:*

1.

$$\forall_{x\in\mathbf{R}}\ 0 \le F(x) \le 1,$$

2.

$$F \textit{ jest funkcją niemalejącą,}$$

3.

$$\lim_{x\to-\infty} F(x) = 0,$$

4.

$$\lim_{x \to +\infty} F(x) = 1,$$

5.

$$F \text{ jest lewostronnie ciągła,}$$

Na odwrót, dla każdej funkcji F o własnościacj (1)–(5) istnieje przestrzeń probabilistyczna $\left(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P}\right)$ oraz zmienna losowa $\tilde{\mathbf{X}}$ określona na tej przestrzeni, że jej dystrybuanta jest równa funkcji F, czyli

$$F_{\tilde{\mathbf{X}}} = F.$$

W przeciwieństwie do sytuacji rozkładu dyskretnego może się zdarzyć, że zbiór $\mathbf{X}(\Omega)$ nie musi być przeliczalny.

Zadanie 5.1.1 *Dla każdego $\omega \in \Omega$ niech $\mathbf{X}(\omega)$ oznacza losowo wybraną liczbę rzeczywistą. Weźmy równanie kwadratowe*

$$x^2 + \mathbf{X}(\omega)x + \mathbf{X}(\omega) = 0.$$

Niech A oznacza zbiór tych $\omega \in \Omega$, że równanie to ma rozwiązanie. Uzasadnić, że A jest zdarzeniem.

Rozwiązanie

Jasne jest, że

$$A = \{\omega \in \Omega\colon\ \mathbf{X}(\omega) \leq 0\} \cup \{\omega \in \Omega\colon\ \mathbf{X}(\omega) \geq 4\},\ \textit{bowiem}\ \mathbf{X}^2(\omega) - 4\mathbf{X}(\omega) \geq 0.$$

Wystarczy pokazać, że oba zbiory występujące w powyższej sumie są zdarzeniami, bowiem suma zdarzeń jest zdarzeniem.

Ponieważ

$$\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \geq 4\} = \{\omega \in \Omega\colon\ \mathbf{X}(\omega) < 4\}^c,$$

z definicji zmiennej losowej zbiór ten jest zdarzeniem. Zajmieny się pierwszym składnikiem.

Można udowodnić, że dla dowolnej liczby rzeczywistej t i zmiennej losowej $\mathbf{X}$ mamy

$$\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \leq t\} = \bigcap_{n \in \mathbf{N}} \{\omega \in \Omega\colon\ \mathbf{X}(\omega) < t + \frac{1}{n}\}. \tag{5.3}$$

Ponieważ przekrój przeliczalnej podrodziny zdarzeń jest zdarzeniem, dowodzi to, że również pierwszy składnik badanej sumy jest zdarzeniem.

Powiemy, że dwie zmienne losowe $\mathbf{X}$ i $\mathbf{Y}$ są *stochastycznie niezależne*, jeśli

$$\forall_{t,s\in\mathbf{R}}\ \{\omega \in \Omega\colon\ \mathbf{X}(\omega) < t\}\ i\ \{\omega \in \Omega\colon\ \mathbf{Y}(\omega) < s\}$$

są zdarzeniami stochastycznie niezależnymi.

Zadanie 5.1.2 *Z odcinka* $[0, 1]$ *wybrano losowo i w sposób niezależny dwie różne liczby, dzieląc ten odcinek na trzy mniejsze odcinki. Opisać zdarzenie* $A \in \Sigma$*, że z powstałych odcinków można zbudować trójkąt.*

Rozwiązanie

Dla $\omega \in \Omega$ niech $\mathbf{X}(\omega)$ i $\mathbf{Y}(\omega)$ będą różnymi liczbami z odcinka $[0, 1]$. Z założenia zmienne te są stochastycznie niezależne.

Mamy wtedy

$$0 \le \mathbf{X}(\omega) < \mathbf{Y}(\omega) \le 1,$$

albo

$$0 \le \mathbf{Y}(\omega) < \mathbf{X}(\omega) \le 1.$$

Weźmy na chwilę parę liczb p, q takich, że

$$0 \le p < q \le 1.$$

Sformułujemy warunek, przy którym z odcinków:

$$[0, p],\ \ [p, q],\ \ [q, 1]$$

można zbudować trójkąt.

Z podstaw geometrii wiadomo, że aby tak było, suma długości dwóch dowolnych odcinków musi być mniejsza od długości pozostałego. Tak z kolei będzie, jeśli każdy z tych odcinków będzie miał długość mniejszą od $\frac{1}{2}$.

Dlatego

$$p < \frac{1}{2},\ q - p < \frac{1}{2}\ i\ 1 - q < \frac{1}{2}.$$

Wracając do naszych zmiennych losowych dostaniemy

$$A = \{\omega \in \Omega\colon\ \mathbf{X}(\omega) < \frac{1}{2}\ \wedge\ \mathbf{Y}(\omega) - \mathbf{X}(\omega) < \frac{1}{2}\ \wedge\ \mathbf{Y}(\omega) > \frac{1}{2}\} \cup$$

$$\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) < \frac{1}{2}\ \wedge\ \mathbf{X}(\omega) - \mathbf{Y}(\omega) < \frac{1}{2}\ \wedge\ \mathbf{X}(\omega) > \frac{1}{2}\}.$$

Jeśli dla każdego rzeczywistego x_o

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_o\}) = P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \leq x_o\}) - P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < x_o\}) = \lim_{x \to x_o^+} F(x) - F(x_o) = 0, \tag{5.4}$$

czyli dystrybuanta zmiennej losowej jest funkcją ciągłą, to będziemy mówili, że zmienna losowa ma *rozkład całkowicie niedyskretny*.

Jeśli dodatkowo funkcja F jest różniczkowalna poza, być może, skończonym podzbiorem, to powiemy, że $\mathbf{X}$ ma *rozkład ciągły*. Ponieważ w standardowym kursie analizy nie wprowadza się pojęcia *całki Stieltjesa*, dalej w większości sytuacji ograniczymy się tylko do rozkładów ciągłych.

Niech F będzie dystrybuantą zmiennej losowej $\mathbf{X}$ o takim rozkładzie. Wtedy ze *wzoru Riemanna-Newtona-Leibnitza* dla całki mamy

$$\int_a^b F'(x)dx = F(b) - F(a). \tag{5.5}$$

Z drugiej strony

$$P(\{\omega \in \Omega\colon\ a \leq \mathbf{X}(\omega) < b\}) = F(b) - F(a).$$

Oznaczjąc przez

$$F'(x) = f(x) \tag{5.6}$$

dostaniemy

$$P(\{\omega \in \Omega\colon\ a \leq \mathbf{X}(\omega) < b\}) = \int_a^b f(x)dx. \tag{5.7}$$

Funkcję f zdefiniowaną wzorem 5.6 nazywamy *funkcją gęstości rozkładu* zmiennej losowej.

Przechodząc w 5.7 z $a \to -\infty$, otrzymamy

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < b\}) = \int_{-\infty}^b f(x)dx. \tag{5.8}$$

Biorąc $b \to +\infty$ w 5.8, dostaniemy

$$1 = P(\Omega) = \int_{-\infty}^{+\infty} f(x)dx. \tag{5.9}$$

I na odwrót, zachodzi następujące twierdzenie:

Twierdzenie 5.1.2 *(o funkcji gęstości) Dla każdej funkcji ciągłej* $f\colon \mathbf{R} \to \mathbf{R}_+ \cup \{0\}$ *takiej, że zachodzi 5.9 istnieje przestrzeń probabilistyczna* $(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P})$ *i zmienna losowa* $\tilde{\mathbf{X}}$ *taka, że* $F'_{\tilde{\mathbf{X}}} = f$.

Niech $f_{\mathbf{X}}$ oznacza gęstość rozkładu zmiennej losowej $\mathbf{X}$. Wtedy dla dowolnej rzeczywistej funkcji ciągłej g odwzorowanie $g(\mathbf{X})$ jest zmienną losową też o rozkładzie ciągłym.

Zadanie 5.1.3 *Weźmy zmienną losową* $\mathbf{Y} = \mathbf{X}^2$ *(czyli* $g(x) = x^2$*). Znaleźć rozkład tej zmiennej.*

Rozwiązanie

Z założenia $\mathbf{X}$ ma gęstość $f_{\mathbf{X}}$. Ponieważ $\mathbf{Y}$ przyjmuje wartości nieujemne, jej dystrybuanta $F_{\mathbf{Y}}(t) = 0$ dla $t \leq 0$.

Dla $t > 0$ dostaniemy

$$F_{\mathbf{Y}}(t) = P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) < t\}) = P(\{\omega \in \Omega\colon\ \mathbf{X}^2(\omega) < t\}).$$

Po obustronnym spierwiastkowaniu otrzymamy

$$F_{\mathbf{Y}}(t) = P(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) < \sqrt{t}\}) = P(\{\omega \in \Omega\colon\ -\sqrt{t} < \mathbf{X}(\omega) < \sqrt{t}\}).$$

Ale $\mathbf{X}$ ma rozkład ciągły, zatem $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = -\sqrt{t}\}) = 0$ i dlatego

$$F_{\mathbf{Y}}(t) = P(\{\omega \in \Omega\colon\ -\sqrt{t} \leq \mathbf{X}(\omega) < \sqrt{t}\}).$$

Mamy więc

$$F_{\mathbf{Y}}(t) = \begin{cases} 0; & t \leq 0, \\ F_{\mathbf{X}}(\sqrt{t}) - F_{\mathbf{X}}(-\sqrt{t}); & t > 0. \end{cases}$$

Po zróżniczkowaniu stronami, z zasady różniczkowania funkcji złożonej otrzymamy poszukiwaną gęstość zmiennej losowej $\mathbf{Y}$

$$f_{\mathbf{Y}}(t) = \begin{cases} 0; & t \leq 0, \\ \frac{1}{2\sqrt{t}}\Big(f_{\mathbf{X}}(\sqrt{t}) + f_{\mathbf{X}}(-\sqrt{t})\Big); & t > 0. \end{cases}$$

Zadanie 5.1.4 *Niech* $\mathbf{X}(\Omega) = [a, b]$ *dla* $a < b$. *Jak wiadomo wtedy,* $\forall_{a \le \alpha < \beta \le b}$ *zbiór* $A = \{\omega \in \Omega \colon \mathbf{X}(\omega) \in [\alpha, \beta]\}$ *jest zdarzeniem.*

Załóżmy, że $P(A) = \frac{\beta - \alpha}{b - a}$. *Znaleźć rozkład zmiennej* $\mathbf{X}$.

Rozwiązanie

Ponieważ $\mathbf{X}(\Omega) = [a, b]$, więc

$$F_{\mathbf{X}}(t) = \begin{cases} 0; & t \le a, \\ 1; & t > b. \end{cases}$$

Dla $t \in (a, b]$ dostaniemy

$$F_{\mathbf{X}}(t) = P(\{\omega \in \Omega \colon \mathbf{X}(\omega) < t\}) = P(\{\omega \in \Omega \colon \mathbf{X}(\omega) \in [a, t)\}) = \frac{t - a}{b - a}.$$

W takim razie $\mathbf{X}$ ma rozkład ciągły i funkcja gęstości ma postać

$$f_{\mathbf{X}}(t) = \begin{cases} 0; & t \notin [a, b] \\ \frac{1}{b-a}; & t \in [a, b]. \end{cases}$$

Dalej rozkład taki będziemy oznaczali przez $\mathcal{J}([a, b])$ i będziemy nazywali *rozkładem jednostajnym na przedziale* $[a, b]$.

Zadanie 5.1.5 *Niech zmienne losowe* $\mathbf{X}$ *i* $\mathbf{Y}$ *będą stochastycznie niezależne o rozkładzie ciągłym. Znaleźć rozkłady zmiennych losowych*

$$\mathbf{Z}_1 = min\{\mathbf{X}, \mathbf{Y}\}, \ \mathbf{Z}_2 = max\{\mathbf{X}, \mathbf{Y}\}.$$

Rozwiązanie

Oznaczmy odpowiednio przez $F_{\mathbf{X}}$, $F_{\mathbf{Y}}$, $f_{\mathbf{X}}$ i $f_{\mathbf{Y}}$ dystrybuanty i funkcje gęstości tych zmiennych. Wtedy z własności funkcji prawdopodobieństwa dostaniemy

$$F_{\mathbf{Z}_1}(t) = P(\{\omega \in \Omega \colon \mathbf{Z}_1(\omega) < t\}) = P(\{\omega \in \Omega \colon min\{\mathbf{X}(\omega), \mathbf{Y}(\omega)\} < t\}) =$$

$$1 - P(\{\omega \in \Omega \colon min\{\mathbf{X}(\omega), \mathbf{Y}(\omega)\} \ge t\}) = 1 - P(\{\omega \in \Omega \colon \mathbf{X}(\omega) \ge t \wedge \mathbf{Y}(\omega) \ge t\}) =$$

$$1 - P(\{\omega \in \Omega \colon \mathbf{X}(\omega) \ge t\}) P(\{\omega \in \Omega \colon \mathbf{Y}(\omega) \ge t\}),$$

gdzie ostatnia równość jest konsekwencją stochastycznej niezależności.

Dlatego

$$F_{\mathbf{Z}_1}(t) = 1 - \Big(1 - P(\{\omega \in \Omega \colon \mathbf{X}(\omega) < t\})\Big)\Big(1 - P(\{\omega \in \Omega \colon \mathbf{Y}(\omega) < t\})\Big)$$

i ostatecznie

$$F_{\mathbf{Z}_1}(t) = 1 - \Big(1 - F_{\mathbf{X}}(t)\Big)\Big(1 - F_{\mathbf{Y}}(t)\Big),$$

co dowodzi, że $\mathbf{Z}_1$ ma rozkład ciągły.

Różniczkując stronami, otrzymamy

$$f_{\mathbf{Z}_1}(t) = -\Big((-f_{\mathbf{X}}(t))(1 - F_{\mathbf{Y}}(t)) + (1 - F_{\mathbf{X}}(t))(-f_{\mathbf{Y}}(t))\Big) =$$

$$f_{\mathbf{X}}(t)(1 - F_{\mathbf{Y}}(t)) + f_{\mathbf{Y}}(t)(1 - F_{\mathbf{X}}(t)).$$

Dla zmiennej $\mathbf{Z}_2$ rachunek będzie wyglądał nieco prościej

$$F_{\mathbf{Z}_2}(t) = P(\{\omega \in \Omega\colon\ \mathbf{Z}_2(\omega) < t\}) = P(\{\omega \in \Omega\colon\ max\{\mathbf{X}(\omega), \mathbf{Y}(\omega)\} < t\}) =$$

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < t \wedge \mathbf{Y}(\omega) < t\}) = P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < t\})P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) < t\}),$$

skąd

$$F_{\mathbf{Z}_2}(t) = F_{\mathbf{X}}(t)F_{\mathbf{Y}}(t).$$

Oznacza to, że $\mathbf{Z}_2$ też ma rozkład ciągły, oraz z zasady różniczkowania iloczynu dostaniemy

$$f_{\mathbf{Z}_2}(t) = f_{\mathbf{X}}(t)F_{\mathbf{Y}}(t) + F_{\mathbf{X}}(t)f_{\mathbf{Y}}(t).$$

Zadanie 5.1.6 *Wiadomo, że f jest gęstością rozkładu pewnej zmiennej losowej. Bierzemy funkcję*

$$g(x) = a + f\Big(\frac{x}{b} + c\Big) \ \ dla\ b \neq 0,\ \ a, c \in \mathbf{R}.$$

1. *Dobrać tak stałe $a, b, c,$ aby g była funkcją gęstości pewnej zmiennej losowej $\mathbf{Y}$.*
2. *Znaleźć dystrybuantę tej zmiennej.*

Rozwiązanie

Z warunku 5.9 wynika, że $a = 0$. Dla takiego a rozwiążemy równanie

$$1 = \int_{-\infty}^{+\infty} g(x)dx = \int_{-\infty}^{0} g(x)dx + \int_{0}^{+\infty} g(x)dx.$$

Ponieważ

$$\int_{-\infty}^{0} g(x)dx = \lim_{T\to-\infty} \int_{T}^{0} f\Big(\frac{x}{b} + c\Big)dx,$$

więc, stosując zamianą zmiennych w ostaniej całce, dostaniemy

$$\int_T^0 f\Big(\frac{x}{b}+c\Big)dx = b\int_{\frac{T}{b}+c}^{c} f(u)du$$

i po powrocie do głównego rachunku

$$\lim_{T\to-\infty}\int_T^0 f\Big(\frac{x}{b}+c\Big)dx = b\lim_{T\to-\infty}\int_{\frac{T}{b}+c}^{c} f(u)du = \begin{cases} b\int_{-\infty}^{c} f(u)du; & b>0,\\ -b\int_{c}^{+\infty} f(u)du; & b<0.\end{cases}$$

Podobnie dla drugiej całki mamy

$$\int_0^{+\infty} g(x)dx = \lim_{S\to+\infty}\int_0^S f\Big(\frac{x}{b}+c\Big)dx =$$

$$b\lim_{S\to+\infty}\int_c^{\frac{S}{b}+c} f(u)du = \begin{cases} b\int_{c}^{+\infty} f(u)du; & b>0,\\ -b\int_{-\infty}^{c} f(u)du; & b<0.\end{cases}$$

Ostatecznie nasze równanie będzie miało postać:

$$1 = \int_{-\infty}^{+\infty} g(x)dx = b\Big(\int_{-\infty}^{c} f(u)du + \int_c^{+\infty} f(u)du\Big),$$

czyli

$$b\int_{-\infty}^{+\infty} f(u)du = 1 \Leftrightarrow b=1,\ \textit{dla } b>0,\ c\in\mathbf{R}.$$

Z kolei dla $b<0$ dostaniemy

$$(-b)\Big(\int_{-\infty}^{c} f(u)du + \int_c^{+\infty} f(u)du\Big) = 1 \Leftrightarrow -b=1,\ c\in\mathbf{R}.$$

Oznacza to, że dla $a=0,\ c\in\mathbf{R},\ b\in\{-1,1\}$ g jest funkcją gęstości pewnej zmiennej losowej $\mathbf{Y}$.
Biorąc na przykład $b=1, c\in\mathbf{R}$, dostaniemy $g(x)=f(x+1)$ i odpowiednio

$$F_{\mathbf{Y}}(t) = \int_{-\infty}^{t} f(x+c)dx = \lim_{T\to-\infty}\int_T^t f(x+c)dx =$$

$$\lim_{T\to-\infty}\int_{T+c}^{t+c} f(u)du = \int_{-\infty}^{t+c} f(u)du = F_{\mathbf{X}}(t+c),$$

gdzie w powyższych całkach po drodze skorzystaliśmy z zasady zamiany zmiennych.

Zadanie 5.1.7 *Sprawdzić, czy*

$$f(x) = \begin{cases} \lambda e^{-\lambda x} \ ; & gdy \ x \geq 0 \\ 0 \ ; & gdy \ x < 0 \end{cases}$$

jest funkcją gęstości prawdopodobieństwa. Jeśli tak, to wyznaczyć dystrybuantę tego rozkładu.

Rozwiązanie

Z własności funkcji gęstości wynika, że $\lambda > 0$. Sprawdzimy, czy dla takich λ zachodzi warunek 5.9.

Ponieważ

$$\int \lambda e^{-\lambda x} dx = -e^{-\lambda x} + C,$$

więc odpowiednie całki będą równe

$$\int_{-\infty}^{0} f(x)dx = 0, \text{ co wynika z definicji } f$$

oraz

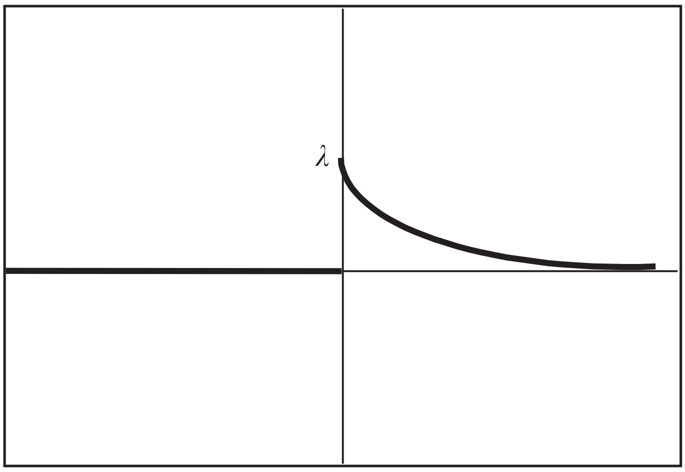
$$\int_{0}^{+\infty} f(x)dx = \lim_{T \to +\infty} \int_{0}^{T} \lambda e^{-\lambda x} dx = \lim_{T \to +\infty} [-e^{-\lambda x}]_0^T =$$
$$\lim_{T \to +\infty} (-e^{-\lambda x} + 1) = 1.$$

Pokazaliśmy, że dla każdej wartości $\lambda > 0$, f definiuje rozkład prawdopodobieństwa. Niech $\mathbf{X}$ oznacza zmienną losową o takim rozkładzie. Wtedy $\mathbf{X}(\Omega) = \mathbf{R}_+ \cup \{0\}$, co oznacza, że $F(t) = 0$ dla $t \leq 0$.

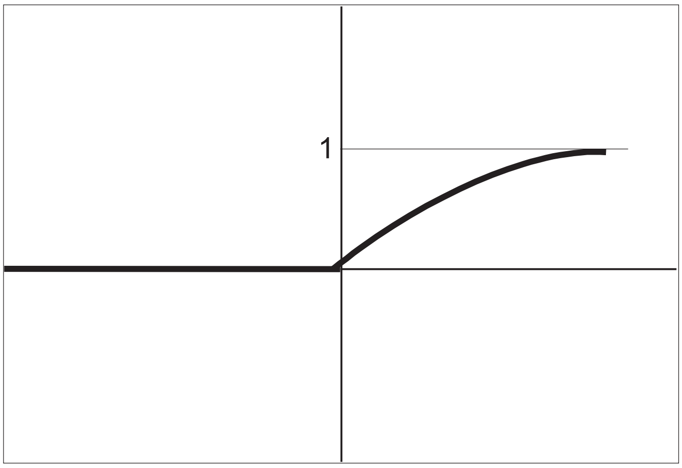
Dla $t > 0$ dostaniemy

$$F(t) = \int_{-\infty}^{t} \lambda e^{-\lambda x} dx = \int_{0}^{t} \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^t = 1 - e^{-\lambda t}.$$

Rozkład rozważany powyżej w literaturze nazywany jest *rozkładem wykładniczym* i oznaczany jest przez $\mathcal{W}(\lambda)$. Poniżej zamieściliśmy wykresy funkcji f i F.



Rysunek 5.1: gęstość rozkładu $\mathcal{W}(\lambda)$



Rysunek 5.2: dystrybuanta rozkładu $\mathcal{W}(\lambda)$

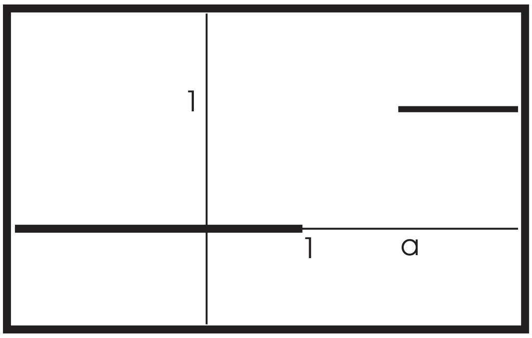
Zadanie 5.1.8 *Dla jakich a, b, c rzeczywistych funkcja*

$$F(x) = \begin{cases} 0; & x \leq 1 \\ -b(1 + \frac{c}{x}); & 1 < x \leq a \\ 1; & x > a \end{cases}$$

jest dystrybuantą rozkładu dyskretnego? Czy F może być dystrybuantą rozkładu całkowicie niedyskretnego i ciągłego?

Rozwiązanie

F jest dystrybuantą rozkładu dyskretnego dokładnie wtedy, gdy jako dystrybuanta jest funkcją schodkową. Z założenia $a > 1$ i dla $x \leq 1$ oraz dla $x \geq a$, F ma postać jak na rysunku 5.3



Rysunek 5.3: dystrybuanta rozkładu

Zatem, aby F mogła być funkcją schodkową na przedziale $(1, a]$, musi przyjmować wartość stałą z przedziału $[0, 1]$. Dlatego b, c muszą przyjmować takie wartości, aby funkcja

$$-b\Big(1 + \frac{c}{x}\Big) = const \in [0, 1] \ dla \ x \in [1, a].$$

Możo to się zdarzyć tylko w dwóch sytuacjach:

1. $$b = 0 \ i \ c \in \mathbf{R}$$

   albo

2. $$c = 0, \ 0 < -b \leq 1, \ czyli \ c = 0 \ i \ -1 \leq b < 0.$$

Jeśli F ma być dystrybuantą rozkładu całkowicie niedyskretnego, to jako funkcja musi być ciągła. Dokładniej, z postaci F chodzi o ciągłość tylko w punktach 1 i a.

Dla liczby 1 dostaniemy

$$\lim_{x \to 1^+} F(x) = -b(1 + c) = 0 \Leftrightarrow b = 0 \ \vee \ c = -1.$$

Ponieważ dla $b = 0$, $F(a) = 0 < \lim_{x \to a^+} F = 1$, dlatego $c = -1$.

Wtedy

$$F(x) = -b\Big(1 - \frac{1}{x}\Big), \; 1 < x \le a \; i \; -1 \le b < 0.$$

Ponadto

$$F(a) = -b\Big(1 - \frac{1}{a}\Big) = 1 \Leftrightarrow b = \frac{a}{1-a} \; i \; a > 1.$$

Ostatecznie dla $c = -1$, $a > 1$ i $b = \frac{a}{1-a}$, F jest dystrybuantą rozkładu całkowicie niedyskretnego. Ponieważ F nie jest różniczkowalna tylko w 1 i a, jest to również rozkład ciągły.

5.2 Zadania

Zadanie 5.2.1 *Zmienna losowa* $\mathbf{X}$ *ma dystrybuantę*

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

Obliczyć : $P(\mathbf{X} = 2),\ \ P(-3 \leq \mathbf{X} < 1),\ \ P(\mathbf{X} > 1/2)$.

Zadanie 5.2.2 *Dobrać tak stałe* $a,\ b \in \mathbf{R}$*, aby*

$$F(x) = \begin{cases} \frac{1}{2}e^{ax} & dla\ x \leq 1 \\ bx + 3/4 & dla\ 1 < x \leq 2 \\ 1 & dla\ x > 2 \end{cases}$$

była dystrybuantą pewnego rozkładu prawdopodobieństwa.

Zadanie 5.2.3 *Dla jakich liczb rzeczywistych* a, b, c

$$F(x) = \begin{cases} 0 & dla\ x \leq 1 \\ b(1 - c/x) & dla\ 1 < x \leq a \\ 1 & dla\ x > a \end{cases}$$

jest dystrybuantą rozkładu dyskretnego? Czy można tak dobrać te stałe, aby F *była dystrybuantą rozkładu całkowicie niedyskretnego?*

Zadanie 5.2.4 *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

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

Znaleźć rozkład zmiennej losowej $\mathbf{Y} = -\mathbf{X} + 2$.

Zadanie 5.2.5 *Dla jakich rzeczywistych* a, b, c

$$f(x) = \begin{cases} a & dla\ x \leq -1 \\ b & dla\ -1 < x \leq 1 \\ c & dla\ x > 1 \end{cases}$$

jest gęstością pewnego rozkładu. Wyznaczyć dystrybuantę tego rozkładu.

Zadanie 5.2.6 *Zmienna losowa* $\mathbf{X}$ *ma gęstość*

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

Obliczyć $P(-1 < \mathbf{X} < 1/2)$.

Zadanie 5.2.7 *Dobrać tak wartość α, aby funkcja*

$$f(x) = \begin{cases} \alpha \sin x & dla\ x \in [0, \pi] \\ 0 & dla\ x \notin [0, \pi] \end{cases}$$

była funkcją gęstości. Wyznaczyć dystrybuantę tego rozkładu.

Zadanie 5.2.8 *Dobrać tak stałe a, b aby funkcja*

$$F(x) = \begin{cases} a + b \arccos x & dla\ |x| < 1 \\ 0 & dla\ x \leq -1 \\ 1 & dla\ x \geq 1 \end{cases}$$

była gęstością zmiennej losowej $\mathbf{X}$ *typu ciągłego. Znaleźć jej funkcję gęstości.*

Zadanie 5.2.9 *Sprawdzić, czy f jest funkcją gęstości pewnego rozkładu. Jeśli tak, to znaleźć jego dystrybuantę, jeśli*

$$f(x) = \begin{cases} 0 & dla\ x \leq 0 \\ x & dla\ x \in (0, 1] \\ 2 - x & dla\ x \in (1, 2] \\ 0 & dla\ 2 < x \end{cases}$$

Zadanie 5.2.10 *Niech* $\mathbf{X}$ *oznacza losowo wybraną liczbę z przedziału* $[1, 5]$ *z funkcją gęstości* $f(x) = \alpha x$.

1. *Znaleźć* α.
2. *Obliczyć* $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \in [2, 4]\})$.
3. *Obliczyć* $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > 4\})$.

Zadanie 5.2.11 *Niech* $\mathbf{X} \in \mathcal{J}([0, 1])$. *Znaleźć rozkłady zmiennych losowych:*

1. $$\mathbf{Y} = \frac{1}{\mathbf{X} + 1},$$

2. $$\mathbf{Y} = ln(\mathbf{X} + 1),$$

3. $$\mathbf{Y} = \left(\mathbf{X} - \frac{1}{2}\right)^2.$$

Zadanie 5.2.12 *Zmienna losowa* $\mathbf{X}$ *ma gęstość*

$$f(x) = \begin{cases} cx(1-x) & dla\ x \in (0,1) \\ 0 & dla\ x \notin (0,1) \end{cases}$$

1. *Znaleźć stałą* c.
2. *Wyznaczyć dystrybuantę* F.
3. *Obliczyć* $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < \frac{1}{4}\})$.

Zadanie 5.2.13 $\mathbf{X}$ *ma dystrybuantę określoną następująco*

$$F(x) = \begin{cases} 0 & dla\ x \leq 0 \\ \sin^2 \frac{\pi x}{2} & dla\ x \in (0,1] \\ 1 & dla\ x > 1 \end{cases}$$

1. *Czy* $\mathbf{X}$ *jest typu ciągłego?*
2. *Jeśli tak, to wyznaczyć* f.
3. *Obliczyć* $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < \frac{1}{4}\})$.

Zadanie 5.2.14 *Niech* $\mathbf{X} \in \mathcal{W}(\lambda)$. *Znaleźć gęstość zmiennej* $\mathbf{Y} = r\mathbf{X}$ *dla* $r > 0$.

Zadanie 5.2.15 *Wiadomo, że zmienna losowa* $\mathbf{X}$ *ma rozkład dany funkcją gęstości*

$$f(x) = \begin{cases} \frac{1}{3} & dla\ x \in (0,1] \\ \frac{2}{3} & dla\ x \in (1,2] \\ 0 & dla\ pozostałych \end{cases}$$

Znaleźć rozkład zmiennej losowej $\mathbf{Y} = -\frac{1}{2}\mathbf{X} + \frac{3}{4}$. *Czy otrzymany rozkład jest ciągły? Jeśli tak, to wyznaczyć* $f_{\mathbf{Y}}$.

Zadanie 5.2.16 *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

$$f(x) = \begin{cases} 12x(1-x)^2 & dla\ x \in (0,1) \\ 0 & dla\ x \notin (0,1) \end{cases}$$

Znaleźć rozkład

1. $\mathbf{Y} = 2\mathbf{X}^2$,
2. $\mathbf{Y} = e^{\mathbf{X}}$.

Zadanie 5.2.17 $\mathbf{X}$ *ma rozkład dany funkcją gęstości*

$$f(x) = \begin{cases} \frac{1}{2x^2} & dla\ |x| > 1 \\ 0 & dla\ |x| \leq 1 \end{cases}$$

Znaleźć dystrybuantę zmiennej $\mathbf{X}$.

Zadanie 5.2.18 *Wiadomo, że w pewnej miejscowości temperatura* $\mathbf{X}$ *mierzona jest według rozkładu o funkcji gęstości* $f(t) = \frac{1}{2}e^{-|t|},\ t \in \mathbf{R}$.

1. *Znaleźć dystrybuantę rozkładu temperatury.*
2. *Obliczyć prawdopodobieństwo* $P(\{\omega \in \Omega \colon\ |\mathbf{X}|(\omega) < 3\})$.

Zadanie 5.2.19 *Czy można tak dobrać stałą* a*, aby funkcja*

$$F(x) = \begin{cases} 0 & \text{dla } x \leq 1 \\ 2\left(1 - \frac{1}{x}\right) & \text{dla } x \in (1, a] \\ 1 & \text{dla } x > a \end{cases}$$

była dystrybuantą zmiennej losowej typu ciągłego? Wyznaczyć gęstość tej zmiennej losowej.

Zadanie 5.2.20 *Zmienna losowa* $\mathbf{X}$ *ma gęstość*

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

Znaleźć rozkład zmiennej losowej $\mathbf{Y} = -\mathbf{X}$. *Czy ten rozkład jest ciągły?*

Rozdział 6

Parametry rozkładów

6.1 Wprowadzenie teoretyczne i przykłady

Na początku przypomnimy podstawy rachunku *całkowania po rozkładzie prawdopodobieństwa.* Niech będą dane:

1. dystrybuanta F rozkładu prawdopodobieństwa,
2. rzeczywista funkcja ciągła g.

Wyjaśnimy znaczenie całki zapisywanej jako

$$\int_{\mathbf{R}} g(x)dF.$$

Jak już sygnalizowaliśmy na wstępie, ten typ całki, zwanej *całką Stieltjesa*, nie mieści się w standardowym kursie analizy. Dlatego zrobimy to z praktycznego punktu widzenia, pokazując zasadę jej liczenia.

W tym celu rozpatrzymy trzy najważniejsze sytuacje:

1. F jest różniczkowalna, poza być może, zbiorem skończonym, czyli dany rozkład jest ciągły. Wtedy
$$\int_{\mathbf{R}} g(x)dF = \int_{-\infty}^{+\infty} g(x)F'(x)dx = \int_{-\infty}^{+\infty} g(x)f(x)dx,$$
o ile funkcja $|g|f$ jest całkowalna, gdzie f jest funkcją gęstości tego rozkładu.

2. F ma przeliczalny zbiór punktów nieciągłości $\{x_n\colon\ n \in N_o\}$, który jest zbiorem ograniczonym. Ponadto załóżmy, że dla każdej pary sąsiednich punktów nieciągłości $x_i <\ x_j$ na przedziale (x_i, x_j), F jest różniczkowalna. Wtedy
$$\int_{\mathbf{R}} g(x)dF = \int_{-\infty}^{a} g(x)f(x)dx + \int_{b}^{+\infty} g(x)f(x)dx + \sum_{x_i<x_j} \int_{x_i}^{x_j} g(x)f(x)dx+$$

$$\sum_{n\in\mathbf{N}_o} g(x_n)p_n,$$

gdzie

$$p_n = \lim_{x\to x_n^+} F(x) - F(x_n) = P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_n\}),$$

o ile funkcja $|g|f$ jest całkowalna i odpowiednie szeregi są zbieżne bezwzględnie. Ponadto liczby a i b oznaczają odpowiednio kres dolny i kres górny zbioru punktów nieciągłości funkcji F i $f = F'$.

3. F ma przeliczalny zbiór punktów nieciągłości $\{x_n\colon\ n \in N_o\}$, który jest zbiorem nieograniczonym. Ponadto załóżmy, że dla każdej pary sąsiednich punktów nieciągłości $x_i < x_j$ na przedziale (x_i, x_j), F jest różniczkowalna. Wtedy

$$\int_{\mathbf{R}} g(x)dF = \sum_{x_i<x_j} \int_{x_i}^{x_j} g(x)f(x)dx + \sum_{n\in\mathbf{N}_o} g(x_n)p_n,$$

gdzie p_n ma znaczenie jak wcześniej, o ile odpowiednie szeregi są zbieżne bezwzględnie.

Jeśli $F = F_{\mathbf{X}}$ i $g(x) = x^r$ dla $r \in \mathbf{N}$, to $\int_{\mathbf{R}} g(x)dF$, o ile jest skończona, nazywamy *r-tym momentem zmiennej losowej* $\mathbf{X}$, co zapisujemy

$$\mathbf{EX}^r. \tag{6.1}$$

Przypadek $r = 1$ nazywamy *wartością średnią* lub *oczekiwaną* zmiennej losowej $\mathbf{X}$.

Zachodzi następujące twierdzenie

Twierdzenie 6.1.1 *Dla każdej zmiennnej losowej* $\mathbf{X}$ *mamy:*

1.

$$\forall_{r\geq 2}\ |\mathbf{EX}^r| < \infty \Rightarrow |\mathbf{EX}^{r-1}| < \infty.$$

2.

$$\forall_{\alpha,\beta\in\mathbf{R}}\ \mathbf{E}(\alpha\mathbf{X} + \beta\mathbf{Y}) = \alpha\mathbf{EX} + \beta\mathbf{EY},$$

co nazywamy własnością liniowości operacji wartości średniej.

Załóżmy teraz, że $\mathbf{X}$ ma r-ty moment ($r \geq 2$). Wtedy, jak wiemy, ma też pierwszy moment oraz można pokazać, że istnieje liczba

$$\mathbf{E}(\mathbf{X} - \mathbf{EX})^r,$$

którą nazywamy *r-tym momentem centralnym* zmiennej losowej $\mathbf{X}$. Wśród takich momentów ważny jest drugi. Nazywamy go *wariancją* zmiennej losowej $\mathbf{X}$ oraz oznaczamy

$$var(\mathbf{X}) = \mathbf{E}(\mathbf{X} - \mathbf{EX})^2.$$

Dla wariancji mamy następujące twierdzenie:

Twierdzenie 6.1.2 *Niech zmienna losowe* $\mathbf{X}$ *i* $\mathbf{Y}$ *mają drugie momenty. Wtedy:*

1.

$$var(\mathbf{X}) \geq 0,$$

2.

$$var(\mathbf{X}) = 0 \Leftrightarrow \ \exists_{c \in \mathbf{R}} \ P(\{\omega \in \Omega \colon \ \mathbf{X}(\omega) = c\}) = 1,$$

3.

$$var(\mathbf{X}) = \mathbf{EX}^2 - (\mathbf{EX})^2,$$

4. jeśli zmienne losowe $\mathbf{X}$ *i* $\mathbf{Y}$ *są stochastycznie niezależne, to*

$$var(\mathbf{X} + \mathbf{Y}) = var(\mathbf{X}) + var(\mathbf{Y}).$$

Zadanie 6.1.1 *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

$$F(x) = \begin{cases} 0; & x \leq -1 \\ -\frac{1}{4}x^2 + \frac{1}{2}x + \frac{3}{4}; & -1 < x \leq 1 \\ 1; & x > 1 \end{cases}$$

Obliczyć wariancję tej zmiennej.

Rozwiązanie

Ponieważ $\mathbf{X}$ ma rozkład ciągły oraz

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

więc

$$\mathbf{EX} = \int_{\mathbf{R}} x dF = \int_{-\infty}^{+\infty} x f(x) dx = \int_{-1}^{1} \Big(-\frac{1}{2}x + \frac{1}{2}\Big) x dx = -\frac{1}{3}.$$

Podobnie

$$\mathbf{EX}^2 = \int_{\mathbf{R}} x^2 dF = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_{-1}^{1} \Big(-\frac{1}{2}x + \frac{1}{2}\Big) x^2 dx = \frac{1}{3}.$$

Dlatego

$$var(\mathbf{X}) = \mathbf{EX}^2 - (\mathbf{EX})^2 = \frac{1}{3} - \frac{1}{9} = \frac{2}{9}.$$

Zadanie 6.1.2 *Niech* $\mathbf{X} \in \mathcal{P}(\lambda)$. *Czy rozkład ten posiada wariancję skończoną?*

Rozwiązanie

Jak wiemy, dystrybuanta F tego rozkładu jest funkcją schodkową i jej punktami nieciągłości są elemety zbioru $\mathbf{N} \cup \{0\}$. Stąd z zasady całkowania dostaniemy

$$\int_{\mathbf{R}} x dF = \int_{-\infty}^{0} x0dx + \sum_{i \geq 0} \int_{i}^{i+1} x0dx + \sum_{i \geq 0} ip_i.$$

Ponieważ

$$p_i = P(\{\omega \in \Omega \colon \ \mathbf{X}(\omega) = i\}) = e^{-\lambda}\frac{\lambda^i}{i!},$$

więc

$$\int_{\mathbf{R}} x dF = \sum_{i \geq 0} ie^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda} \sum_{i \geq 1} \frac{\lambda\lambda^{i-1}}{(i-1)!} = \lambda,$$

co pokazuje, że $\mathbf{EX} = \lambda$.

Sprawdzimy teraz, czy skończona jest wartość drugiego momentu. Dostaniemy odpowiednio

$$\int_{\mathbf{R}} x^2 dF = \sum_{i \geq 0} i^2 e^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda} \sum_{i \geq 1} i\frac{\lambda\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda} \sum_{j \geq 0} (j+1)\frac{\lambda^j}{j!},$$

gdzie w ostatniej sumie zamieniliśmy zmienne podstawiając $j = i - 1$. Kontynuując rachunek, otrzymamy

$$\int_{\mathbf{R}} x^2 dF = \lambda e^{-\lambda}\Big(\sum_{j \geq 1} \lambda \frac{\lambda^{j-1}}{(j-1)!} + \sum_{j \geq 0} \frac{\lambda^j}{j!}\Big) = \lambda e^{-\lambda}(\lambda e^{\lambda} + e^{\lambda}) = \lambda^2 + \lambda.$$

Zatem $\mathbf{X}$ posiada skończony drugi moment, stąd i wariancję równą

$$var(\mathbf{X}) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Zadanie 6.1.3 *Niech* $\mathbf{X}$ *ma rozkład dyskretny skończony, czyli* $\mathbf{X}(\Omega) = \{x_1, \ldots, x_n\}$ *oraz* $P(\{\omega \in \Omega\colon \mathbf{X}(\omega) = x_j\}) = p_j$. *Wyprowadzić wzór na wartość oczekiwaną tej zmiennej losowej.*

Rozwiązanie

Można dalej przyjąć, że elementy x_j zostały uporządkowane rosnąco. Wtedy

$$\int_{\mathbf{R}} xdF = \int_{-\infty}^{x_1} 0xdx + \int_{x_n}^{+\infty} 0xdx + \sum_{i=1}^{n-1} \int_{x_i}^{x_{i+1}} 0xdx + \sum_{i=1}^{n} x_i p_i.$$

Ostatecznie

$$\mathbf{EX} = \sum_{i=1}^{n} x_i p_i.$$

W szczególności, gdy rozkład ten jest jednocześnie jednorodny, czyli $p_i = \frac{1}{n}$, to ostatni wzór przyjmie postać

$$\mathbf{EX} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

i nazywa się *średnią arytmetyczną* z wartości zmiennej losowej.

Zadanie 6.1.4 *Dane są dwie niezależne zmienne losowe* $\mathbf{X}$ *i* $\mathbf{Y}$, *o których wiadomo, że:*

$$\mathbf{EX} = 2,\ \mathbf{EY} = 1,5,\ \mathbf{EX}^2 = 3,\ \mathbf{EY}^2 = 4.$$

Bierzemy zmienną losową $\mathbf{Z} = -2,5\mathbf{X} + 1,5\mathbf{Y}$. *Obliczyć wariancję tej zmiennej losowej.*

Rozwiązanie

Gdybyśmy umieli wyznaczyć rozkład zmiennej losowej $\mathbf{Z}$, postąpilibyśmy jak w przypadku zadań przedstawionych wyżej–całkowalibyśmy po tym rozkładzie. Ponieważ nie możemy tego zrobić, najpierw skorzystamy z liniowości operacji wartości oczekiwanej. Pozwoli to nam obliczyć wartość oczekiwaną $\mathbf{Z}$

$$\mathbf{EZ} = \mathbf{E}(-2,5\mathbf{X} + 1,5\mathbf{Y}) = -2,5\mathbf{EX} + 1,5\mathbf{EX} = -5 + 2,25 = -2,75.$$

Teraz obliczymy jej drugi moment

$$\mathbf{EZ}^2 = \mathbf{E}\Big(-2,5\mathbf{X} + 1,5\mathbf{Y}\Big)^2 = \mathbf{E}\Big(6,25\mathbf{X}^2 + 2,25\mathbf{Y}^2 - 7,5\mathbf{XY}\Big).$$

Korzystając ponownie z liniowości wartości oczekiwanej, otrzymamy

$$\mathbf{EZ}^2 = 6,25\mathbf{EX}^2 + 2,25\mathbf{EY}^2 - 7,5\mathbf{E}\Big(\mathbf{XY}\Big) = 28,75 - 7,5\mathbf{E}\Big(\mathbf{XY}\Big).$$

Można pokazać, że jeśli zmienne losowe $\mathbf{X}$ *i* $\mathbf{Y}$ mają wartości oczekiwane i są niezależne, to

$$\mathbf{E}\Big(\mathbf{XY}\Big) = \mathbf{EX} \cdot \mathbf{EY}.$$

Dlatego

$$\mathbf{EZ}^2 = 28,75 - 7,5 \cdot 2 \cdot 1,5 = 28,75 - 22,5 = 6,25.$$

Wreszcie

$$var(\mathbf{Z}) = 6,25 - (-2,75)^2.$$

6.2 Zadania

Zadanie 6.2.1 *Uzasadnić , że dla zmiennej losowej* $\mathbf{X}$ *mamy*

1.

$$var(\mathbf{X}) = \mathbf{EX}^2 - \Big(\mathbf{EX}\Big)^2.$$

2.

$$var(c\mathbf{X}) = c^2 var(\mathbf{X}), \ \textit{dla } c \in \mathbf{R}.$$

Zadanie 6.2.2 *Wiadomo, że dla pewnej zmiennej losowej* $\mathbf{X}$ *mamy:* $\mathbf{EX} = 100$, $var(\mathbf{X}) = 15$. *Znaleźć:*

1. $\mathbf{EX}^2$,
2. $\mathbf{E}(3\mathbf{X} + 10)$,
3. $\mathbf{E}(-\mathbf{X})$,
4. $var(-\mathbf{X})$.

Zadanie 6.2.3 *Pokazać, że dla niezależnych zmiennych losowych* $\mathbf{X}$ *i* $\mathbf{Y}$ *o rozkładach dyskretnych skończonych zachodzi wzór* $\mathbf{E}(\mathbf{XY}) = (\mathbf{EX})(\mathbf{EY})$.

Zadanie 6.2.4 *Korzystając z faktu, że dla niezależnych zmiennych losowych posiadających drugie momenty* $\mathbf{E}(\mathbf{XY}) = (\mathbf{EX})(\mathbf{EY})$, *uzasadnić, że wariancja sumy tych zmiennych jest sumą wariancji.*

Zadanie 6.2.5 *Korzystając z definicji rozkładu dwumianowego* $\mathbf{X} \in B(n,p)$, *wyprowadzić wzory na wartość oczekiwaną i wariancję.*

Zadanie 6.2.6 *Na przykładzie zmiennej losowej o rozkładzie dyskretnym skończonym zinterpretować pojęcie wartości oczekiwanej i wariancji. Na tej podstawie wywnioskować, że jeśli zmienna losowa* $\mathbf{X}$ *rejestruje stale jednakową wartość, to* $var(\mathbf{X}) = 0$ *oraz* $\mathbf{EX} = \mathbf{X}$.

Zadanie 6.2.7 *Niech zmienna losowa* $\mathbf{X}$ *rejestruje pierwsze pojawienie się porażki w nieskończonej serii powtórzeń pojedynczego eksperymentu, który zwraca* 1 *(sukces) z prawdopodobieństwem* p *i* 0 *(porażkę) z prawdopodobieństwem* q. *Znaleźć rozkład tej zmiennej losowej, a następnie pokazać, że* $\mathbf{EX} = \frac{1}{q}$.

Zadanie 6.2.8 *Niech zmienna losowa* $\mathbf{X}$ *ma rozkład*

$$F(x) = \begin{cases} 0, & gdy \quad x \le -1 \\ 1 - \frac{x^2}{2}, & gdy \quad -1 < x \le 0 \\ 1, & gdy \quad x > 0. \end{cases}$$

Obliczyć jej wariancję.

Zadanie 6.2.9 *Wiadomo,że zmienne losowe* $\mathbf{X}_1,\ \mathbf{X}_2$ *są niezależne i mają rozkład zdefiniowany jak niżej*

$$F(x) = \begin{cases} 0, & gdy \quad x \le 0 \\ \frac{1}{2}x, & gdy \quad 0 < x \le 1 \\ 1, & gdy \quad x > 1. \end{cases}$$

Bierzemy zmienną losową
$\mathbf{X} = 3\mathbf{X}_1 - \mathbf{X}_2 + 1$. *Obliczyć :* $\mathbf{EX},\ \mathbf{EX}^2,\ var(\mathbf{X})$.

Zadanie 6.2.10 *Dany jest rozkład zmiennej* $\mathbf{X}$ *jak poniżej*

x_i	-2	-1	0	1	2
$p_i =$	1/3	1/6	2/9	1/6	1/9

Obliczyć drugi moment i drugi moment centralny zmiennej $\mathbf{Y} = 2\mathbf{X} + 1$.

Zadanie 6.2.11 *Czy dla pewnych a funkcja*

$$f(x) = \begin{cases} \frac{a}{x^2} & dla\ x > 1 \\ 0 & dla\ x \le 1 \end{cases}$$

posiada $\mathbf{EX}$*?*

Zadanie 6.2.12 *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

$$f(x) = \begin{cases} 3x^2 & dla\ x \in (0,1) \\ 0 & dla\ x \notin (0,1) \end{cases}$$

Obliczyć $var(\mathbf{Y})$, *jeśli* $\mathbf{Y} = \mathbf{X}^2$.

Zadanie 6.2.13 *Zmienna losowa* $\mathbf{X}$ *ma rozkład Poissona z parametrem* λ. *Bierzemy zmienną losową* $\mathbf{Y} = \mathbf{X} + 2$. *Obliczyć jej wartość oczekiwaną i wariancję.*

Zadanie 6.2.14 *Dane są dwie niezależne zmienne losowe* $\mathbf{X}_1$ *i* $\mathbf{X}_2$ *o rozkładzie ciągłym oraz o wartościach oczekiwanych równych odpowiednio* m_1 *i* m_2. *Znaleźć wartości oczekiwane zmiennych losowych* $min\{\mathbf{X}_1, \mathbf{X}_2\}$ *i* $max\{\mathbf{X}_1, \mathbf{X}_2\}$.

Zadanie 6.2.15 *Podać przykład zmiennej losowej, która ma wartość oczekiwaną, a nie ma drugiego momentu.*

Zadanie 6.2.16 *Wiadomo, że* $\mathbf{X} \in B(n,p)$. *Obliczyć wartość oczekiwaną i wariancję zmiennej losowej* $\frac{1}{n}\mathbf{X}$.

Zadanie 6.2.17 *Obliczyć wartość oczekiwaną i wariancję zmiennej losowej* $\mathbf{X} \in \mathcal{J}([a,b])$.

Zadanie 6.2.18 *Niech zmienna losowa* $\mathbf{X}$ *ma rozkład*

$$f(x) = \frac{\lambda}{\pi\Big(\lambda^2 + (x-\mu)^2\Big)}, \ dla\ \lambda > 0,\ \ \mu, x \in \mathbf{R}.$$

Sprawdzić, czy $\mathbf{X}$ *ma wartość oczekiwaną.*

Zadanie 6.2.19 *O zmiennej losowej* $\mathbf{X}$ *wiadomo, że*

1. *jej gęstość jest funkcją parzystą,*
2. *ma moment rzędu* 7.

Wywnioskować na tej podstawie, że $\mathbf{EX} = \mathbf{EX}^3 = \mathbf{EX}^5 = \mathbf{EX}^7 = 0$.

Zadanie 6.2.20 *Bierzemy zmienną losową* $\mathbf{X}$ *o rozkładzie* $f(x) = \frac{a}{c}\left(\frac{c}{x}\right)^{a+1}$ *dla* $x > c,\ \ a > 0,\ \ c > 0$. *Uzasadnić, że* f *jest funkcją gęstości. Znaleźć wartość oczekiwaną i wariancję tej zmiennej losowej.*

Rozdział 7

Ważniejsze rozkłady prawdopodobieństwa

7.1 Wprowadzenie teoretyczne i przykłady

Niech $\mathbf{X}$ będzie zmienną losową taką, że $m = \mathbf{EX}$ i $\sigma^2 = var(\mathbf{X}) > 0$. Weźmy zmienną losową

$$\mathbf{Z} = \frac{\mathbf{X} - m}{\sigma}.$$

Wtedy $\mathbf{Z}$ będziemy nazywali efektem *standaryzacji* zmiennej losowej $\mathbf{X}$.

Zauważmy, że wtedy mamy:

Fakt 7.1.1

$$\mathbf{EZ} = 0, \ var(\mathbf{Z}) = 1. \tag{7.1}$$

$$F_{\mathbf{Z}}(t) = F_{\mathbf{X}}(m + t\sigma). \tag{7.2}$$

W szczególności, jeśli $\mathbf{X}$ *jest typu ciągłego, to* $\mathbf{Z}$ *również oraz*

$$f_{\mathbf{Z}}(t) = f_{\mathbf{X}}(m + t\sigma) \cdot \sigma. \tag{7.3}$$

Zadanie 7.1.1 *Udowodnić fakt 7.1.1.*

Rozwiązanie

Wzory 7.1 wynikają z odpowiedniego twierdzenia o wartości oczekiwanej i wariancji. Dla dowodu 7.2 zauważmy, że

$$F_{\mathbf{Z}}(t) = P\Big(\{\omega \in \Omega\colon \frac{\mathbf{X}(\omega) - m}{\sigma} < t\}\Big) = P\Big(\{\omega \in \Omega\colon \mathbf{X}(\omega) < m+t\sigma\}\Big) = F_{\mathbf{X}}(m+t\sigma).$$

Wreszcie, przy założeniu, że $\mathbf{X}$ jest typu ciągłego, z reguły różniczkowania funkcji złożonej ze wzoru 7.2 dostaniemy

$$f_{\mathbf{Z}}(t) = \frac{d}{dt}F_{\mathbf{Z}}(t) = F'_{\mathbf{X}}(m + t\sigma)\frac{d}{dt}(m + t\sigma) = f_{\mathbf{X}}(m + t\sigma) \cdot \sigma,$$

co dowodzi 7.3.

I. **Rozkład dwupunktowy**

Oznaczenie rozkładu	$\mathcal{D}(p),\ p \in (0, 1)$
typ rozkładu	dyskretny, skończony
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 1\}) = p$, $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 0\}) = q = 1 - p$
dystrybuanta	schodkowa
funkcja gęstości	brak
EX	p
$var(\mathbf{X})$	pq

Uwaga 7.1.1 *Rozkład opisany w tabeli I bardzo często nazywany jest standardowym rozkładem dwupunktowym. W sytuacji ogólnej dopuszcza się, że*

$$\mathbf{X}(\Omega) = \{r, s\},\ r, s \in \mathbf{R},$$

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = r\}) = q,\ P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = s\}) = p.$$

II. **Rozkład dwumianowy Newtona-Bernoulliego**

Oznaczenie rozkładu	$B(n,p),\ p \in (0,1),\ n \geq 1$
typ rozkładu	dyskretny, skończony
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = k\}) = \binom{n}{k} p^k (1-p)^{n-k}$, $k \in \{0, 1, \ldots, n\}$
dystrybuanta	schodkowa
funkcja gęstości	brak
EX	np
$var(\mathbf{X})$	npq

Uwaga 7.1.2 *Można pokazać, że jeśli* $\mathbf{X}_j \in \mathcal{D}(p)$ *i są parami stochastycznie niezależne, to*

$$\mathbf{X}_1 + \mathbf{X}_2 + \ldots + \mathbf{X}_n \in B(n,p).$$

III. **Rozkład Poissona**

Oznaczenie rozkładu	$\mathcal{P}(\lambda),\ \lambda > 0$
typ rozkładu	dyskretny, nieskończony
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = k\}) = e^{-\lambda}\frac{\lambda^k}{k!}$, $k \in \mathbf{N} \cup \{0\}$
dystrybuanta	schodkowa
funkcja gęstości	brak
EX	λ
$var(\mathbf{X})$	λ

Twierdzenie 7.1.1 *(o rozkładzie Poissona) Niech $\mathbf{Y}_n \in B(n, p_n)$ będzie ciągiem zmiennych losowych o rozkładzie dwumianowym, gdzie*

$$np_n \longrightarrow \beta > 0.$$

Wtedy

$$\forall_{k\in\mathbf{N}\cup\{0\}}\ \forall_{k\leq n}\ P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = k\}) = \binom{n}{k} p_n^k (1-p_n)^{n-k} \longrightarrow e^{-\beta}\frac{\beta^k}{k!}.$$

Istnieje wtedy zmienna losowa $\mathbf{Z} \in \mathcal{P}(\beta)$ taka, że

$$\forall_{k\in\mathbf{N}\cup\{0\}}\ P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = k\}) \longrightarrow P(\{\omega \in \Omega\colon\ \mathbf{Z}(\omega) = k\}).$$

Zadanie 7.1.2 *Udowodnić twierdzenie 7.1.1.*

Rozwiązanie

Załóżmy, że $np_n \longrightarrow \beta > 0$. Wtedy z twierdzenia o liczbie Eulera dostaniemy

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = 0\}) = (1 - p_n)^n = \left(1 - \frac{np_n}{n}\right)^n =$$

$$\left(\left(1 + \frac{1}{\frac{n}{-np_n}}\right)^{\frac{n}{np_n}}\right)^{np_n} \longrightarrow e^{-\beta}.$$

Weźmy teraz $0 < k \leq n$. Dzieląc przez siebie odpowiednie prawdopodobieństwa, dostaniemy

$$\frac{P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = k\})}{P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = k-1\})} = \frac{\binom{n}{k} p_n^k (1-p_n)^{n-k}}{\binom{n}{k-1} p_n^{k-1} (1-p_n)^{n-k+1}} = \frac{n-k+1}{k} \frac{p_n}{1-p_n}$$

co możemy też zapisać

$$\frac{P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = k\})}{P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = k-1\})} = \frac{np_n}{k} \frac{n-k+1}{n(1-p_n)} = \frac{np_n}{k} \frac{1 - \frac{k}{n} + \frac{1}{n}}{1-p_n} \longrightarrow \frac{\beta}{k},$$

bowiem z założenia $p_n \longrightarrow 0$.

Stąd

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = 1\}) \longrightarrow \beta e^{-\beta}.$$

I ogólnie, z powyższego wzoru na drodze indukcji matematycznej wynika, że dla $k \geq 0$

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}_n(\omega) = k\}) \longrightarrow \frac{\beta^k}{k!} e^{-\beta}.$$

Uwaga 7.1.3 *W zastosowaniach często wykorzystujemy tzw. wersję przybliżoną tego twierdzenia. Mówi ona, że dla* $\mathbf{Y} \in B(n,p)$, *gdzie*

$$n \geq 100,\ p \in (0,\ 0{,}1],\ np = \lambda \in [0{,}1,\ 10]$$

ma miejsce dobre przybliżenie

$$\forall_{0 \leq k \leq n}\ P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) = k\}) = \binom{n}{k} p^k (1-p)^{n-k} \cong e^{-\lambda} \frac{\lambda^k}{k!}. \qquad (7.4)$$

Zadanie 7.1.3 *Błąd w pewnej próbie można wykryć w* $99,8\%$ *przypadków. Oszacować prawdopodobieństwo, że w* 500 *próbach nie wykryto błędu w co najmniej* 5 *przypadkach.*

Rozwiązanie

Niech $\mathbf{X}$ oznacza, że nie wykryto błedu. Z założenia $\mathbf{X} \in B(n,p)$, gdzie $n = 500,\ p = 0,002,\ \lambda = np = 10$. Mamy więc spełnione założenia gwarantujące użycia przybliżenia 7.4, czyli

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = k\}) \cong e^{-10}\frac{10^k}{k!}.$$

Weźmy zdarzenie

$$A = \{\omega \in \Omega\colon\ \mathbf{X}(\omega) \geq 5\}.$$

Wtedy

$$P(A) = 1 - P(A^c) = 1 - \Big(P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 0\}) + P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 1\}) +$$

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 2\}) + P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 3\}) + P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 4\})\Big) \cong$$

$$1 - e^{-10}\Big(1 + 10 + \frac{10^2}{2} + \frac{10^3}{6} + \frac{10^4}{24}\Big).$$

IV. **Rozkład geometryczny**

Oznaczenie rozkładu	$\mathcal{G}(p),\ p \in (0,1)$
typ rozkładu	dyskretny, nieskończony
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = k\}) = pq^k,\ q = 1 - p$ $k \in \mathbf{N} \cup \{0\}$
dystrybuanta	schodkowa
funkcja gęstości	brak
EX	$\frac{p}{q}$
$var(\mathbf{X})$	$\frac{p}{q^2}$

Zadanie 7.1.4 *Zaproponować model probabilistyczny zmiennej losowej o rozkładzie geometrycznym.*

Rozwiązanie

Rozważmy doświadczenie losowe $(\Omega_o, \Sigma_o, P_o)$ polegające na tym, że

$$\Omega_o = \{0,1\},\ P_o(\{1\}) = p,\ P_o(\{0\}) = q = 1 - p,\ \Sigma_o = \mathcal{P}(\Omega_o).$$

Załóżmy, że obserwujemy wynik nieskończonego powtarzania w sposób niezależny naszego pojedynczego doświadczenia. Dowodzi się, że istnieje wtedy tzw. *produktowy model Kołmogorowa* (Ω, Σ, P), gdzie

$$\Omega = \prod \Omega_o = \{\omega = (\omega_n)\colon\ \omega_n \in \Omega_o,\ n \in \mathbf{N}\},\ \Sigma = \bigotimes \Sigma_o$$

z prawdopodobiństwem produktowym P.

Na Ω zdefiniujmy zmienną losową $\mathbf{X}$ o wartościach w zbiorze liczb naturalnych, gdzie

$$\mathbf{X}(\omega) = n \iff \omega \in A_n = \{\omega \in \Omega\colon\ \omega_1 = 0, \ldots, \omega_n = 0,\ \omega_{n+1} = 1\}.$$

Oznacza to, że sukces po raz pierwszy w takiej serii zauważyliśmy przy $n+1$-szym powtórzeniu. Zauważmy, że wtedy dla $n \geq 1$ dostaniemy

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = n\}) = P(A_n) = pq^n.$$

Ponieważ

$$\sum_{n\geq 1} P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = n\}) = p\frac{q}{1-q} = q,$$

więc z definicji musimy przyjąć, że $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = 0\}) = p$.

V. **Rozkład jednostajny**

Oznaczenie rozkładu	$\mathcal{J}([a,b]),\ a < b$
typ rozkładu	ciągły
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	0
dystrybuanta	$F(x) = \begin{cases} 0 & ,\ gdy \quad x \leq a \\ \frac{x-a}{b-a} & ,\ gdy \quad a < x \leq b \\ 1 & ,\ gdy \quad x > b \end{cases}$
funkcja gęstości	$f(x) = \begin{cases} \frac{1}{b-a} & ,\ gdy \quad a \leq x \leq b \\ 0 & ,\ gdy \quad a < x \notin [a,b] \end{cases}$
EX	$\frac{a+b}{2}$
$var(\mathbf{X})$	$\frac{(b-a)^2}{12}$

Zadanie 7.1.5 *Pokazać, że dla zmiennej losowej* $\mathbf{X} \in \mathcal{J}([0,1])$ *i* $a < b$ *rzeczywistych istnieje funkcja rzeczywista* Ψ *taka, że* $\Psi(\mathbf{X}) \in \mathcal{J}([a,b])$.

Rozwiązanie

Definiujemy funkcję $\Psi(x) = a + (b-a)x$ oraz bierzemy zmienną losową

$$\mathbf{Y} = \Psi(\mathbf{X}) = a + (b-a)\mathbf{X}.$$

Zauważmy, że $\mathbf{Y}(\Omega) = [a,b]$, ponieważ $\mathbf{X}(\Omega) = [0,1]$.

Ponadto

$$F_{\mathbf{Y}}(t) = P(\{\omega \in \Omega\colon a+(b-a)\mathbf{X}(\omega) < t\}) = P\Big(\{\omega \in \Omega\colon \mathbf{X}(\omega) < \frac{t-a}{b-a}\}\Big) = F_{\mathbf{X}}\Big(\frac{t-a}{b-a}\Big),$$

więc zmienna losowa $\mathbf{Y}$ ma rozkład ciągły z funkcją gęstości

$$f_{\mathbf{Y}}(t) = F'\Big(\frac{t-a}{b-a}\Big)\frac{d}{dt}\Big(\frac{t-a}{b-a}\Big) = f_{\mathbf{X}}\Big(\frac{t-a}{b-a}\Big)\frac{1}{b-a}.$$

Ale z założenia $f_{\mathbf{X}}(u) = \begin{cases} 1, & gdy \quad u \in [0,1] \\ 0, & gdy \quad u \notin [0,1] \end{cases}$, więc

$$u \in [a,b] \iff \frac{t-a}{b-a} \in [0,1] \iff t \in [a,b] \iff f_{\mathbf{X}}\Big(\frac{t-a}{b-a}\Big) = 1$$

oraz

$$t \notin [a,b] \iff f_{\mathbf{X}}\Big(\frac{t-a}{b-a}\Big) = 0,$$

co dowodzi, że $\mathbf{Y} \in \mathcal{J}([a,b])$.

VI. **Rozkład wykładniczy**

Oznaczenie rozkładu	$\mathcal{W}(\lambda),\ \lambda > 0$
typ rozkładu	ciągły
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	0
dystrybuanta	$F(x) = \begin{cases} 0 & ,\ gdy \quad x \leq 0 \\ 1 - e^{-\lambda x} & ,\ gdy \quad x > 0 \end{cases}$
funkcja gęstości	$f(x) = \begin{cases} \lambda e^{-\lambda x} & ,\ gdy \quad x \geq 0 \\ 0 & ,\ gdy \quad x < 0 \end{cases}$
EX	$\frac{1}{\lambda}$
$var(\mathbf{X})$	$\frac{1}{\lambda^2}$

Zadanie 7.1.6 *Niech* $\mathbf{X} \in \mathcal{J}([0,1])$. *Znaleźć rozkład zmiennej losowej* $\mathbf{Y} = -\frac{1}{\lambda} ln(1 - \mathbf{X})$ *dla* $\lambda > 0$.

Rozwiązanie

Z założenia $\mathbf{Y}(\Omega) \subset \mathbf{R}_+$ i dlatego $F_{\mathbf{Y}}(t) = 0$ dla $t \leq 0$. Dla $t > 0$ dostaniemy

$F_{\mathbf{Y}}(t) = P\Big(\{\omega \in \Omega\colon\ -\dfrac{1}{\lambda} ln(1 - \mathbf{X}(\omega)) < t\}\Big) = P\Big(\{\omega \in \Omega\colon\ ln(1 - \mathbf{X}(\omega)) > -\lambda t\}\Big) =$
$P\Big(\{\omega \in \Omega\colon\ 1 - \mathbf{X}(\omega) > e^{-\lambda t}\}\Big) = P\Big(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < 1 - e^{-\lambda t}\}\Big) = F_{\mathbf{X}}(1 - e^{-\lambda t})$.

Oznacza to, że $\mathbf{Y}$ ma rozkład ciągły oraz dla $t > 0$ funkcja gęstości ma postać (dla $t \leq 0,\ f_{\mathbf{Y}}(t) = 0$)

$$f_{\mathbf{Y}}(t) = F'_{\mathbf{X}}(1 - e^{-\lambda t}) \frac{d}{dt}(1 - e^{-\lambda t}) =$$

$$f_{\mathbf{X}}(1-e^{-\lambda t})\lambda e^{-\lambda t} = \lambda e^{-\lambda t},$$

bowiem

$$t > 0 \iff 1 - e^{-\lambda t} \in [0,1] \iff f_{\mathbf{X}}(1-e^{-\lambda t}) = 1.$$

Dowodzi to, że $\mathbf{Y} \in \mathcal{W}(\lambda)$.

VII. **Rozkład Cauchy'ego**

Oznaczenie rozkładu	$\mathcal{C}(\lambda, \mu),\ \lambda > 0,\ \mu \in \mathbf{R}$
typ rozkładu	ciągły
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	0
dystrybuanta	$F(x) = \frac{1}{\pi}(arctg\frac{t-\mu}{\lambda} + \frac{\pi}{2})$
funkcja gęstości	$f(x) = \frac{\lambda}{\pi(\lambda^2+(x-\mu)^2)}$
EX	nie istnieje
$var(\mathbf{X})$	nie istnieje

Zadanie 7.1.7 *Wyznaczyć dystrybuantę rozkładu Cauchy'ego.*

Rozwiązanie

Z definicji dystrybuanty mamy

$$F(t) = \int_{-\infty}^{t} f(x)dx = \int_{-\infty}^{t} \frac{\lambda}{\pi(\lambda^2 + (x-\mu))^2}dx =$$

$$\frac{\lambda}{\pi}\lim_{T\to-\infty}\int_T^t \frac{1}{\lambda^2}\frac{dx}{1+(\frac{x-\mu}{\lambda})^2} = \frac{1}{\pi\lambda}\lim_{T\to-\infty}\int_T^t \frac{dx}{1+(\frac{x-\mu}{\lambda})^2}.$$

Dla obliczenia ostatniej całki zastosujemy podstawienie $\frac{x-\mu}{\lambda} = u$. Wtedy $dx = \lambda du$ oraz

$$\int_T^t \frac{dx}{1+(\frac{x-\mu}{\lambda})^2} = \int_{\frac{T-\mu}{\lambda}}^{\frac{t-\mu}{\lambda}} \frac{\lambda du}{1+u^2} = \lambda arctg u\Big]_{\frac{T-\mu}{\lambda}}^{\frac{t-\mu}{\lambda}} =$$

$$\lambda\Big(arctg(\frac{t-\mu}{\lambda}) - arctg(\frac{T-\mu}{\lambda})\Big) \to \lambda\Big(arctg(\frac{t-\mu}{\lambda}) + \frac{\pi}{2}\Big),$$

gdy $T \to -\infty$. Ostatecznie

$$F(t) = \frac{1}{\pi}\Big(arctg(\frac{t-\mu}{\lambda}) + \frac{\pi}{2}\Big).$$

VIII. **Rozkład normalny Gaussa**

Oznaczenie rozkładu	$\mathcal{N}(m,\sigma^2),\ \sigma > 0,\ m \in \mathbf{R}$
typ rozkładu	ciągły
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	0
dystrybuanta	$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}}dt,\ x \in \mathbf{R}$
funkcja gęstości	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-m)^2}{2\sigma^2}},\ x \in \mathbf{R}$
EX	m
$var(\mathbf{X})$	σ^2

Twierdzenie 7.1.2 *(o rozkładzie normalnym) Jeśli* $\mathbf{X} \in \mathcal{N}(m, \sigma^2)$*, to* $\frac{\mathbf{X}-m}{\sigma} \in \mathcal{N}(0,1)$*. Dalej rozkład typu* $\mathcal{N}(0,1)$ *będziemy nazywali standardowym rozkładem normalnym, a jego dystrybuantę oznaczymy przez* Φ*.*

Ponadto mamy:

$$\Phi(0) = \frac{1}{2}, \tag{7.5}$$

$$\forall_{t>0}\ P(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) \leq t\}) = P(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) < t\}) = 2\Phi(t) - 1, \tag{7.6}$$

$$\forall_{t>0}\ P(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) \geq t\}) = P(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) > t\}) = 2(1 - \Phi(t)), \tag{7.7}$$

$$P(\{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - m| > 3\sigma\}) = P(\{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - m| \geq 3\sigma\}) = \tag{7.8}$$

$$P(\{\omega \in \Omega\colon\ |\frac{\mathbf{X}(\omega) - m}{\sigma}| \geq 3\}) = 2(1 - \Phi(3)) = 2 \cdot 0{,}00135 < 0{,}01,$$

gdzie ostatnią wartość odczytaliśmy z tablicy rozkładu normalnego.

Uwaga 7.1.4 *Nierówność 7.7 nosi nazwę reguły 3 sigm. W praktyce oznacza ona, że standardowy rozkład normalny skoncentrowany jest na przedziale symetrycznym* $[-3,3]$*. Z kolei wzór 7.5 tłumaczy, że znajomość dystrybuanty* Φ *wystarczy ograniczyć do przedziału* $(0,3]$*.*

Zadanie 7.1.8 *Uzasadnić, że* $\frac{\mathbf{X}-m}{\sigma}$ *jest standardowym rozkładem normalnym dla każdej zmiennej losowej* $\mathbf{X} \in \mathcal{N}(m, \sigma^2)$*.*

Rozwiązanie

Skorzystamy ze wzoru 7.3. W tym celu weźmy $\mathbf{X} \in \mathcal{N}(m, \sigma^2)$. Wtedy

$$f_{\mathbf{Y}}(x) = f_{\mathbf{X}}(m + x\sigma)\sigma = \frac{1}{\sqrt{2\pi\sigma^2}}\sigma e^{-\frac{(x\sigma)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

co dowodzi, że $f_{\mathbf{Y}}$ jest gęstością standardowego rozkładu normalnego.

Zadanie 7.1.9 *Niech* $\mathbf{X} \in \mathcal{N}(-2,4)$*. Obliczyć:*

1.
$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > -3\}),$$

2.
$$P(\{\omega \in \Omega\colon\ -2 < \mathbf{X}(\omega) < 7\}),$$

3.

$$P(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) \leq 4\}).$$

Rozwiązanie

1. Z własności funkcji prawdopodobieństwa mamy
$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > -3\}) = 1 - P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \leq -3\}).$$
Ponieważ zmienna losowa $\mathbf{X}$ ma rozkład ciągły, ostatnie prawdopodobieństwo możemy zapisać
$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > -3\}) = 1 - P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < -3\}) =$$
$$1 - P\Big(\{\omega \in \Omega\colon\ \frac{\mathbf{X}(\omega)+2}{2} < \frac{-1}{2}\}\Big).$$
Korzystając z 7.6 ostatecznie, otrzymamy
$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > -3\}) = 1 - (1 - \Phi(\frac{1}{2})) = \Phi(\frac{1}{2}),$$
z tablic standardowego rozkładu normalnego daje $0,691462$.

2. Z własności rozkładu ciągłego dostaniemy
$$P(\{\omega \in \Omega\colon\ -2 < \mathbf{X}(\omega) < 7\}) = P(\{\omega \in \Omega\colon\ -2 \leq \mathbf{X}(\omega) < 7\}).$$
Po standaryzacji otrzymamy
$$P(\{\omega \in \Omega\colon\ -2 < \mathbf{X}(\omega) < 7\}) = P\Big(\{\omega \in \Omega\colon\ 0 \leq \frac{\mathbf{X}(\omega)+2}{2} < \frac{9}{2}\}\Big)$$
$$= \Phi(\frac{9}{2}) - \Phi(0).$$
Ponowny odczyt z tabeli da wynik $0,999999 - 0,5 = 0,499999$.

3. Korzystając z ciągłości rozkładu, w wyniku jego standaryzacji otrzymamy
$$P(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) \leq 4\}) = P(\{\omega \in \Omega\colon\ -4 \leq \mathbf{X}(\omega) \leq 4\}) =$$
$$P(\{\omega \in \Omega\colon\ -1 \leq \frac{\mathbf{X}(\omega)+2}{2} < 3\}),$$
skąd z własności dystrybuanty możemy zapisać
$$P(\{\omega \in \Omega\colon\ |\mathbf{X}|(\omega) \leq 4\}) = \Phi(3) - \Phi(-1) = \Phi(3) - (1 - \Phi(1)).$$
Po sięgnięciu do tabeli standardowego rozkładu normalnego otrzymamy $0,998650 + 0,841345 - 1 = 0,839995$.

IX. **Rozkład chi-kwadrat Pearsona**

Oznaczenie rozkładu	$\chi_n^2,\ n \geq 2$
typ rozkładu	ciągły
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	0
dystrybuanta	nie podajemy
funkcja gęstości	nie podajemy
EX	n
$var(\mathbf{X})$	$2n$

Uwaga 7.1.5 *Z definicji przyjmuje się, że rozkład zmiennej losowej*

$$\mathbf{X}_1^2(\omega) + \ldots + \mathbf{X}_n^2(\omega),$$

gdzie zmienne losowe $\mathbf{X}_j^2(\omega)$ *są parami stochastycznie niezależne o standardowym rozkładzie normalnym, nazywa się rozkładem chi-kwadrat o n stopniach swobody. Ze względu na ograniczony zakres naszego zbioru zdecydowaliśmy się nie używać jawnie dystrybuanty i gęstości tego rozkładu.*

Wiadomo, że dla $n > 30$ *rozkład ten dobrze jest przybliżany (ich dystrybuanty prawie są sobie równe) standardowym rozkładem normalnym. Natomiast dla* $2 \leq n \leq 30$ *jest stablicowany. W tablicy takiego rozkładu pojawiają się tzw. wartości krytyczne tego rozkładu, czyli takie liczby* χ_α^2, *że*

$$P(\{\omega \in \Omega\colon\ \chi_n^2(\omega) > \chi_\alpha^2\}) = \alpha$$

dla danej wartości α.

Zadanie 7.1.10 *Dla* $n = 16$ *znaleźć wartość krytyczną rozkładu chi-kwadrat, jeśli* $\alpha = 0,05$.

Rozwiązanie

Zgodnie z uwagą mamy rozwiązać równanie $P(\{\omega \in \Omega\colon\ \chi^2_n(\omega) > \chi^2_\alpha\}) = \alpha$ z niewiadomą χ^2_α. W tym celu wykorzystamy tabelę wartości krytycznych rozkładu chi-kwadrat. Na skrzyżowaniu wiersza $n = 16$ i kolumny $\alpha = 0,05$ odczytamy $\chi^2_\alpha = 26,296$.

Zadanie 7.1.11 *Niech* $\mathbf{X}$ *ma standardowy rozkład normalny. Wyznaczyć drugi moment i wariancję kwadratu tej zmiennej losowej.*

Rozwiązanie

Zaczniemy od drugiego momentu. ponieważ $\mathbf{EX}^2 = var(\mathbf{X}) + (\mathbf{EX})^2$, więc $\mathbf{EX}^2 = 1$. Dalej skorzystamy z informacji, że $\mathbf{EX}^4 = 3$ (patrz np. [2]).

Dlatego

$$var(\mathbf{X}^2) = \mathbf{EX}^4 - (\mathbf{EX}^2)^2 = 3 - 1 = 2.$$

Zauważmy, że wyniki tego zadania tłumaczą dlaczego dla rozkładu chi-kwadrat o n stopniach swobody wartość oczekiwana i wariancja odpowiednio są równe n i $2n$.

IX. **Rozkład t-Studenta (Goseta)**

Oznaczenie rozkładu	$t_n,\ n \geq 3$
typ rozkładu	ciągły
$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$	0
dystrybuanta	nie podajemy
funkcja gęstości	nie podajemy
EX	0
$var(\mathbf{X})$	$\frac{n}{n-2}$

Uwaga 7.1.6 *Z definicji przyjmuje się, że rozkład zmiennej losowej*

$$\frac{\mathbf{X}}{\sqrt{\frac{\chi_n^2}{n}}},$$

gdzie zmienne losowe $\mathbf{X}$, χ_n^2 *są stochastycznie niezależne i* $\mathbf{X}$ *ma standardowy rozkład normalny, nazywa się rozkładem t-Studenta o n stopniach swobody. Ze względu na ograniczony zakres naszego zbioru zdecydowaliśmy się nie używać jawnie dystrybuanty i gęstości tego rozkładu.*

Wiadomo, że dla $n > 30$ *rozkład ten dobrze jest przybliżany (ich dystrybuanty prawie są sobie równe) standardowym rozkładem normalnym. Natomiast dla* $3 \leq n \leq 30$ *jest stablicowany. Ponadto jest rozkładem symetrycznym, a więc jego funkcja gęstości jest funkcją parzystą, skąd* $\mathbf{E}t_n = 0$. *W tablicy takiego rozkładu pojawiają się tzw. wartości krytyczne tego rozkładu, czyli takie liczby* t_α, *że*

$$P(\{\omega \in \Omega\colon\ |t_n(\omega)| > t_\alpha\}) = \alpha.$$

Zadanie 7.1.12 *Dla $n = 8$ znaleźć wartość krytyczną rozkładu t-Studenta, jeśli $\alpha = 0,2$.*

Rozwiązanie

Zgodnie z uwagą mamy rozwiązać równanie

$$P(\{\omega \in \Omega\colon\ |t_n(\omega)| > t_\alpha\}) = \alpha$$

z niewiadomą t_α. W tym celu wykorzystamy tabelę wartości krytycznych rozkładu t-Studenta. Na skrzyżowaniu wiersza $n = 8$ i kolumny $\alpha = 0,2$ odczytamy $t_\alpha = 1,397$.

Zadanie 7.1.13 *Dla $n = 8$ znaleźć wartość $\tilde{t_\alpha}$ rozkładu t-Studenta, gdzie*

$$P(\{\omega \in \Omega\colon\ t_n(\omega) > \tilde{t_\alpha}\}),$$

jeśli $\alpha = 0,2$.

Rozwiązanie

Zauważmy, że $\tilde{t_\alpha}$ <u>nie jest wartością krytyczną</u> rozkładu t-Studenta! Aby rozwiązać równanie $P(\{\omega \in \Omega\colon\ t_n(\omega) > \tilde{t_\alpha}\})$, weźmy

$$\{\omega \in \Omega\colon\ |t_n(\omega)| > \tilde{t_\alpha}\} = \{\omega \in \Omega\colon\ t_n(\omega) > \tilde{t_\alpha}\} \cup \{\omega \in \Omega\colon\ t_n(\omega) < -\tilde{t_\alpha}\}$$

oraz z symetrii rozkładu

$$P(\{\omega \in \Omega\colon\ t_n(\omega) < -\tilde{t_\alpha}\}) = 1 - P(\{\omega \in \Omega\colon\ t_n(\omega) \leq \tilde{t_\alpha}\}).$$

Dlatego

$$\begin{aligned} P(\{\omega \in \Omega\colon\ |t_n(\omega)| > \tilde{t_\alpha}\}) &= P(\{\omega \in \Omega\colon\ t_n(\omega) > \tilde{t_\alpha}\})+1-P(\{\omega \in \Omega\colon\ t_n(\omega) < \tilde{t_\alpha}\}) \\ &= P(\{\omega \in \Omega\colon\ t_n(\omega) > \tilde{t_\alpha}\}) + 1 - (1 - P(\{\omega \in \Omega\colon\ t_n(\omega) > \tilde{t_\alpha}\})) = \\ &\quad 2P(\{\omega \in \Omega\colon\ t_n(\omega) > \tilde{t_\alpha}\}). \end{aligned}$$

Wobec tego

$$P(\{\omega \in \Omega\colon\ t_n(\omega) > \tilde{t_\alpha}\}) = \alpha \iff P(\{\omega \in \Omega\colon\ |t_n(\omega)| > \tilde{t_\alpha}\}) = 2\alpha.$$

W naszym przypadku $\tilde{t_\alpha} = 0,889$.

7.2 Zadania

Zadanie 7.2.1 *Znaleźć rozkład zmiennej losowej* $\frac{\mathbf{X}}{n}$, *gdzie* $\mathbf{X} \in B(n,p)$.

Zadanie 7.2.2 *Wiadomo, że prawdopodobieństwo zaobserwowania pojedynczej zmiany układu wynosi* $0,001$. *Układ obserwowano w sposób niezależny* 200 *razy. Oszacować prawdopodobieństwo zaobserwowania mniej niż 5 zmian stanu układu.*

Zadanie 7.2.3 *Bazując na związku pomiędzy* $\mathbf{X} \in \mathcal{J}([a,b])$ *a rozkładem jednostajnym na odcinku jednostkowym, wyprowadzić wzory na wartość średnią i wariancję zmiennej* $\mathbf{X}$.

Zadanie 7.2.4 *Pokazać, że dla każdej zmiennej losowej* $\mathbf{X} \in \mathcal{J}([a,b])$ *i przedziału* $[c,d]$ *istnieje funkcja rzeczywista* g *taka, że* $g(\mathbf{X}) \in \mathcal{J}([c,d])$.

Zadanie 7.2.5 *Czy w wyniku standaryzacji rozkładu jednostajnego dostaniemy rozkład jednostajny?*

Zadanie 7.2.6 *Niech* $\mathbf{Y} \in \mathcal{W}(\lambda)$. *Znaleźć rozkład zmiennej losowej* $\mathbf{X} = 1 - e^{-\lambda\mathbf{Y}}$.

Zadanie 7.2.7 *Czy w wyniku standaryzacji rozkładu wykładniczego dostaniemy rozkład wykładniczy?*

Zadanie 7.2.8 *Uzasadnić, że rozkład Cauchy'ego nie ma wartości oczekiwanej.*

Zadanie 7.2.9 *Niech* $\mathbf{X} \in \mathcal{J}([1,2])$. *Znaleźć rozkład zmiennej losowej* $\mathbf{Y} = \frac{1}{\lambda} ln\mathbf{X}$, *dla* $\lambda > 0$.

Zadanie 7.2.10 *Niech* $\mathbf{X} \in \mathcal{J}([a,b])$, $a < b$. *Znaleźć rozkład zmiennej losowej* $\mathbf{Y} = \mathbf{X}^2$.

Zadanie 7.2.11 *Kwantylem rzędu* α *rozkładu o ciągłej dystrybuancie* F *nazywamy taką liczbę rzeczywistą* x_α, *że* $F(x_\alpha) = \alpha$. *Obliczyć kwantyl rzędu* $\frac{1}{2}$ *dla zmiennej losowej* $\mathbf{X} \in \mathcal{W}(\lambda)$.

Zadanie 7.2.12 *O zmiennych losowych* $\mathbf{X}$, $\mathbf{Y}$ *wiadomo, że są niezależne oraz* $\mathbf{X} \in \mathcal{W}(1/3)$, $\mathbf{Y} \in \mathcal{N}(1,2)$. *Bierzemy zmienną losową* $\mathbf{Z} = 1/2\mathbf{X} - 1/3\mathbf{Y} + 2$. *Znaleźć jej wartość oczekiwaną i wariancję.*

Zadanie 7.2.13 *Niech zmienna losowa* $\mathbf{X}$ *ma rozkład normalny o parametrach* $m = -2,5$, $\sigma = 2$. *Korzystając z tablic standardowego rozkładu normalnego, obliczyć :*

1. $P(\{\omega \in \Omega : |\mathbf{X}(\omega)| \leq 1,5\})$

2. $P(\{\omega \in \Omega : |\mathbf{X}(\omega)| \geq 0,5\})$

3. $P(\{\omega \in \Omega : \mathbf{X}^2(\omega) < 9\})$

Zadanie 7.2.14 *Dla zmiennej losowej* $\mathbf{X} \in \mathcal{N}(m, \ \sigma)$ *niech* $\mathbf{Y} = m\mathbf{X} + \sigma$, $m > 0$. *Znaleźć rozkład zmiennej losowej* $\mathbf{Y}$. *Obliczyć* $\mathbf{EY}$ *i* $var(\mathbf{Y})$.

Zadanie 7.2.15 *Korzystając z tablic rozkładu normalnego, obliczyć kwantyle rzędu* $1 - \alpha$, *dla* $\alpha = 0,1$ *i* $\alpha = 0,05$, *jeśli* $\mathbf{X} \in \mathcal{N}(0,1)$ *oraz* $\mathbf{X} \in \mathcal{N}(-0,5, \ 2)$.

Zadanie 7.2.16 *Dla* $n = 10$ *znaleźć wartość krytyczną rozkładu chi-kwadrat, jeśli* $\alpha = 0,02$.

Zadanie 7.2.17 *Dla* $n = 12$ *znaleźć wartość krytyczną rozkładu t-Studenta, jeśli* $\alpha = 0,02$.

Zadanie 7.2.18 *Dla* $n = 18$ *znaleźć wartość* $\tilde{t_\alpha}$ *rozkładu t-Studenta, gdzie*

$$P(\{\omega \in \Omega\colon \ t_n(\omega) < \tilde{t_\alpha}\}),$$

jeśli $\alpha = 0,2$.

Rozdział 8

Nierówności Markowa i Czebyszewa

8.1 Wprowadzenie teoretyczne i przykłady

Niech F oznacza dystrybuantę zmiennej losowej $\mathbf{X}$. Z zasady całkowania po rozkładzie wynika, że:

Fakt 8.1.1

$$\int_{\mathbf{R}} dF = 1, \tag{8.1}$$

$$\forall_{t_o \in \mathbf{R}} \lim_{t \to t_o^+} F(t) = P(\{\omega \in \Omega \colon \mathbf{X}(\omega) \leq t_o\}) = \int_{-\infty}^{t_o} dF, \tag{8.2}$$

$$\forall_{t_o \in \mathbf{R}} \ 1 - F(t_o) = P(\{\omega \in \Omega \colon \mathbf{X}(\omega) \geq t_o\}) = \int_{t_o}^{+\infty} dF. \tag{8.3}$$

Niech teraz dodatkowo $P(\{\omega \in \Omega \colon \mathbf{X}(\omega) > 0\}) = 1$ oraz załóżmy, że istnieje $\mathbf{EX}$.

Wtedy, ponieważ dla dowolnego dodatniego t

$$\mathbf{EX} = \int_0^{+\infty} x dF = \int_0^t x dF + \int_t^{+\infty} x dF$$

obie całki są nieujemne, co oznacza, że

$$\mathbf{EX} \geq \int_t^{+\infty} x dF.$$

Ponieważ funkcja tożsamościowa x najmniejszą wartość na przedziale $[t, +\infty)$ przyjmuje t, daje nam to nierówność

$$\mathbf{EX} \geq t \int_t^{+\infty} dF.$$

Korzystając z 8.3 możemy napisać

Twierdzenie 8.1.1 *(nierówność Markowa) Niech* $\mathbf{X}$ *będzie zmienną losową posiadającą wartość oczekiwaną i przyjmującą wartości dodatnie z prawdopodobieństwem 1.*

Wtedy

$$\forall_{t>0}\ P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \geq t\}) \leq \frac{1}{t}\mathbf{EX}.$$

Niech teraz zmienna losowa $\mathbf{X}$ ma drugi moment i jak zwykle niech $m = \mathbf{EX}$ oraz $var(\mathbf{X}) = \sigma^2 > 0$. Weźmy nową zmienną losową $\mathbf{Y} = (\mathbf{X} - m)^2$. Zauważmy, że spełnia ona założenia nierówności Markowa. Zatem dla dowolnego $t > 0$ dostaniemy

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \geq t^2\}) \leq \frac{1}{t^2}\mathbf{EY}.$$

Ale

$$\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \geq t^2\} = \{\omega \in \Omega\colon\ (\mathbf{X}(\omega) - m)^2 \geq t^2\} = \{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - m| \geq t\},$$

co daje nam następujące twierdzenie

Twierdzenie 8.1.2 *(nierówność Czebyszewa) Dla każdej zmiennej losowej* $\mathbf{X}$ *posiadającej drugi moment i o wariancji dodatniej zachodzi nierówność*

$$\forall_{t>0}\ P(\{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - m| \geq t\}) \leq \frac{1}{t^2}\sigma^2.$$

Zadanie 8.1.1 *Porównać nierówność Czebyszewa z regułą trzech sigm.*

Rozwiązanie

Jak dobrze wiemy, reguła trzech sigm dotyczy zmiennej losowej o rozkładzie normalnym. Dokładniej mamy wtedy (patrz 7.7)

$$P(\{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - m| \geq 3\sigma\}) < 0,01.$$

Tymczasem zastosowanie w tym przypadku nierówności Czebyszewa oznacza, że $(t = 3\sigma)$

$$P(\{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - m| \geq 3\sigma\}) \leq \frac{1}{9}.$$

Widzimy więc, że w tym przypadku reguła trzech sigm jest o wiele bardziej precyzyjna.

Zadanie 8.1.2 *Podać postać nierówności Markowa i Czebyszewa dla skończonego rozkładu dyskretnego. Zaprezentować wyniki na przykładzie rozkładu Bernoulliego.*

Rozwiązanie

Niech $\mathbf{X}(\Omega) = \{x_1, \ldots, x_n\} \subset \mathbf{R}_+,\ \ p_i = P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\})$.
Wtedy

$$m = \sum_{i=1}^{n} x_i p_i,\ \ \sigma^2 = \sum_{i=1}^{n} (x_i - m)^2 p_i.$$

Nierówność Markowa mówi, że

$$\forall_{t>0}\ \ \sum_{x_i \geq t} p_i \leq \frac{1}{t} \sum_{i=1}^{n} x_i p_i,$$

gdzie sumowanie w sumie po lewej stronie nierówności przebiega po tych i, dla których $x_i \geq t$.

Zauważmy, że biorąc $\mathbf{X} \in B(n, p)$ z tego powodu, że $p_0 = q^n > 0$, zmienna ta nie spełnia założeń twierdzenia Markowa. Z drugiej jednak strony mamy

$$\mathbf{E}\mathbf{X} = \sum_{i=0}^{n} i p_i = \sum_{i=1}^{n} i p_i,$$

a więc w tej szczególnej sytuacji zły przypadek został pominięty i dalej mamy jak w sytuacji ogólnej, czyli

$$\sum_{i=1}^{n} i p_i \geq \sum_{i \geq t} i p_i \geq t \sum_{i \geq t} p_i,$$

co dowodzi, że dla rozkładu dwumianowego nierówność Markowa też zachodzi.

Zajmijmy się teraz przypadkiem nierówności Czebyszewa.

Ponieważ

$$\{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - m| \geq t\} = \{\omega \in \Omega\colon\ \mathbf{X}(\omega) - m \geq t\} \cup \{\omega \in \Omega\colon\ \mathbf{X}(\omega) - m \leq -t\},$$

więc

$$\{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - m| \geq t\} = \{\omega \in \Omega\colon\ \mathbf{X}(\omega) \in [0, m-t] \cup [m+t, n]\},$$

o ile $m + t \leq n$.

Zachodzą dwa przypadki:

1. $m < t$. Wtedy pierwsze ze zdarzeń w powyższym prawdopodobieństwie jest zdarzeniem pustym i dostaniemy postać nierówności Czebyszewa

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) \in [m+t, n]\}) = \sum_{np+t \leq i \leq n} p_i \leq \left(\frac{npq}{t}\right)^2.$$

2. $t \leq m$. Wtedy nierówność Czebyszewa oznacza, że

$$\sum_{0 \leq i \leq np-t} p_i + \sum_{np+t \leq i \leq n} p_i \leq \left(\frac{npq}{t}\right)^2.$$

Zadanie 8.1.3 *Niech dla* $n \geq 2$

$$\mathbf{Y}_n = \frac{1}{n}(\mathbf{X}_1 + \ldots + \mathbf{X}_n),$$

gdzie zmienne losowe $\mathbf{X}_j$ *są parami niezależne i mają jednakowy rozkład standardowy dwupunktowy z parametrami* p, q. *Dla jakiego* n *prawdopodobieństwo zdarzenia, że frekwencja zajścia sukcesu różni się od prawdopodobieństwa teoretycznego o co najmniej* $0,001$ *będzie mniejsze od* $0,05$*?*

Rozwiązanie

Ponieważ $n\mathbf{Y}_n \in B(n, p)$ możemy przyjąć, że zmienna losowa $\mathbf{Y}_n$ rejestruje *frekwencję* pojawienia się sukcesu w serii n niezależnych prób. Wtedy liczbę p możemy nazwać *prawdopodobieństwem teoretycznym* zajścia tego sukcesu.

Weźmy liczbę dodatnią t. Szukamy takiego najmniejszego naturalnego n, że dla $t = 0,001$ będziemy mieli

$$P(\{\omega \in \Omega\colon\ |\mathbf{Y}_n(\omega) - p| \geq t\}) \leq 0,05.$$

Ponieważ z nierówności Czebyszewa

$$P(\{\omega \in \Omega\colon\ |\mathbf{Y}_n(\omega) - p| \geq t\}) \leq \frac{1}{(0,001)^2} \frac{(pq)^2}{n^2},$$

możemy założyć, że szukamy takiego najmniejszego n, które spełnia nierówność

$$\frac{1}{(0,001)^2}\frac{(pq)^2}{n^2} \le 0,05,$$

czyli

$$n^2 \ge 2(pq)^2 10^7.$$

Przyjmując, że rozkład dwupunktowy jest symetryczny, czyli $p = q = 0,5$, dostaniemy

$$n^2 \ge \frac{5}{4} \cdot 10^6.$$

W takim razie należy wykonać co najmniej 1110 powtórzeń.

8.2 Zadania

Zadanie 8.2.1 *Zastosować nierówności Markowa i Czebyszewa dla rozkładu geometrycznego.*

Zadanie 8.2.2 *Niech* $\mathbf{X} \in \mathcal{J}([2,6])$. *Oszacować metodą nierówności Markowa oraz Czebyszewa*

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > 4,5\}).$$

Zadanie 8.2.3 *Niech* $\mathbf{X} \in B(n, 0,75)$. *Jakie musi być* n*, aby z prawdopodobieństwem co najmniej 0,95 prawdopodobieństwo empiryczne różniło się od prawdopodobieństwa teoretycznego o mniej aniżeli 0,001?*

Zadanie 8.2.4 *Niech* $\mathbf{X}_1, \ldots, \mathbf{X}_n$ *będą zmiennymi losowymi o tym samym rozkładzie wykładniczym z parametrm* $\lambda = 20$. *Dla jakiego* n

$$P(\mathbf{X} > 10) < 0,5,$$

gdzie $\mathbf{X} = \sum_{j=1}^{n} \mathbf{X}_j$.

Zadanie 8.2.5 *Korzystając z nierówności Czebyszewa, obliczyć ile razy należy rzucić kostką, aby prawdopodobieństwo spełnienia nierówności*

$$|\frac{k_n}{n} - \frac{1}{6}| < 0,01,$$

gdzie k_n *oznacza liczbę wyrzuconych czwórek w* n *rzutach, było większe od 0,5.*

Zadanie 8.2.6 *Niech* $\mathbf{X}_1, \ldots, \mathbf{X}_{100}$ *są niezależnymi zmiennymi losowymi o rozkładzie* $\mathcal{P}(\lambda)$*, gdzie* $\lambda = 2$. *Oczacować* $P\Big(\sum_{j=1}^{100} \mathbf{X}_j > 190\Big)$.

Zadanie 8.2.7 *Wykonano 1000 serii niezależnych doświadczeń. W każdej serii* $m = 2$, $\sigma^2 = 1,4$. *Oszacować prawdopodobieństwo, że w tych 1000 seriach doświadczeń liczba sukcesów będzie zawsze pomiędzy* 186 *a* 214.

Zadanie 8.2.8 *Rzucamy 15000 razy symetryczną kostką. Oszacować prawdopodobieństwo, że liczba wyrzuconych szóstek będzie różniła się od 2500 o więcej aniżeli 100.*

Zadanie 8.2.9 *Strzelamy 300 razy do tarczy, gdzie prawdopodobieństwo trafienia pojedynczym strzałem wynosi* 0, 25. *Oszacować* $P(|\mathbf{X}-75| < 30)$, *gdzie* $\mathbf{X}$ *oznacza liczbę trafień.*

Zadanie 8.2.10 *Niech* $\mathbf{X} \in \mathcal{N}(0,1)$. *Korzystając z nierówności Czebyszewa oszacować* $P(|\mathbf{X}| \geq 3)$. *Wynik oszacowania porównać z odczytem z tabeli tego rozkładu.*

Zadanie 8.2.11 *Rzucamy 600 razy kostką. Oszacować prawdopodobieństwo, że liczba szóstek będzie się różniła od 100 o co najmniej 20.*

Zadanie 8.2.12 *Automatyczna obrabiarka produkuje w ciągu dnia 100 detali, wśród których 5 jest wadliwych. Oszacuj prawdopodobieństwo, że w ciągu 10 dni obrabiarka wyprodukuje co najmniej 60 sztuk wadliwych.*

Zadanie 8.2.13 *Oszacuj prawdopodobieństwo zdarzenia*

$$\{\omega \in \Omega\colon\ |\mathbf{X}(\omega) - \mathbf{E}\mathbf{X}| < 1\},$$

jeśli $var(\mathbf{X}) = 0{,}02$.

Zadanie 8.2.14 *W dostawie cytrusów do hurtowni mamy 20% koszyków zepsutych owoców. Oszacować prawdopodobieństwo, że częstość występowania koszyków z zepsutymi owocami odchyli się od prawdopodobieństwa ich występowania o mniej niż 0,2.*

Zadanie 8.2.15 *Wariancja każdej niezależnej zmiennej losowej z 1000 wynosi 3. Obliczyć prawdopodobieństwo, że srednia tych zmiennych odchyli się od jej wartości oczekiwanej nie więcej niż 0,2.*

Rozdział 9

Funkcje tworzące

9.1 Wprowadzenie teoretyczne i przykłady

Niech d oznacza skończony rozkład dyskretny, czyli

$$d\colon\ \{x_1, \ldots, x_n\} \to \mathbf{R}, \tag{9.1}$$

gdzie

$$d(x_i) = p_i \in (0,1) \text{ } oraz \text{ } \sum_{i=1}^{n} p_i = 1.$$

Jak dobrze wiadomo, istnieje wtedy przestrzeń probabilistyczna (Ω, Σ, P) i zmienna losowa $\mathbf{X}\colon\ \Omega \to \mathbf{R}$ taka, że

$$X(\Omega) = \{x_1, \ldots, x_n\} \text{ } oraz \text{ } p_i = P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = x_i\}).$$

Przez *funkcję tworzącą* $\mathcal{G}_X$ rozkładu d zmiennej losowej $\mathbf{X}$ rozumiemy wielomian

$$\mathcal{G}_X(x) = \sum_{i=1}^{n} p_i x^i. \tag{9.2}$$

Na odwrót, dla każdego wielomianu

$$W(x) = \sum_{k=1}^{m} a_k x^k,\ a_k > 0,\ W(1) = 1$$

ciąg jego współczynników definiuje na zbiorze $\{1, 2, \ldots m\}$ dyskretny skończony rozkład prawdopodobieństwa pewnej zmiennej losowej $\mathbf{Y}$, że $\mathcal{G}_Y = W$.

Zadanie 9.1.1 *Weźmy skończony rozkład dyskretny określony na zbiorze* $\{1,\ 2,\ 3,\ 4,\ 5\}$*, gdzie* $d(i) = \frac{1}{5}$*. Znaleźć funkcję tworzącą tego rozkładu.*

Rozwiązanie

Niech $\mathbf{X}$ oznacza zmienną losową o tym rozkładzie. Wtedy z definicji funkcji tworzącej 9.2 dostaniemy

$$\mathcal{G}_X(x) = \frac{1}{5}(x + x^2 + \ldots + x^5)$$

lub równoważnie

$$\mathcal{G}_X(x) = \begin{cases} \frac{1}{5}x\frac{1-x^5}{1-x} & , \ gdy \quad x \neq 1 \\ 1 & , \ gdy \quad x = 1. \end{cases}$$

Dalej będziemy zakładali, że rozkład dyskretny 9.1 jest *standardowym rozkładem*, tzn. $x_i = i$ dla $i = 0, 1, \ldots, n-1$.

Zachodzi wówczas następujące twierdzenie

Twierdzenie 9.1.1

$$\mathbf{EX} = \mathcal{G}'_X(1). \tag{9.3}$$

$$\mathbf{EX}^2 = \mathcal{G}''_X(1) + \mathcal{G}'_X(1). \tag{9.4}$$

Zadanie 9.1.2 *Weźmy zmienną losową z zadania 9.1.1. Wykorzystując wzóry 9.3 i 9.4, obliczyć jej wartość oczekiwaną i wariancję.*

Rozwiązanie

Z zadania 9.1.1. wiemy, że

$$\mathcal{G}_X(x) = \frac{1}{5}x\frac{1-x^5}{1-x}, \ x \neq 1.$$

Pokażemy jak w inny sposób aniżeli bezpośrednio poprzez różniczkowanie funkcji tworzącej w punkcie 1 można obliczyć wartości jej kolejnych pochodnych w tym punkcie. W tym celu dokonajmy złożenia naszej funkcji tworzącej z funkcją liniową $x+1$. Jeśli przez V oznaczymy to złożenie, to otrzymamy

$$V(x) = \mathcal{G}_X(x+1) = \frac{1}{5}(x+1)\frac{(x+1)^5 - 1}{x} = \frac{1}{5}\frac{(x+1)^6 - (x+1)}{x}.$$

Stosując wzór dwumianowy Newtona

$$(x+1)^6 = 1 + \binom{6}{1}x + \binom{6}{2}x^2 + \binom{6}{3}x^3 + \binom{6}{4}x^4 + \binom{6}{5}x^5 + x^6$$

dostaniemy

$$V(x) = \frac{1}{5}\frac{5x + \binom{6}{2}x^2 + \binom{6}{3}x^3 + \binom{6}{4}x^4 + \binom{6}{5}x^5 + x^6}{x}$$

i po skróceniu

$$V(x) = \frac{1}{5}\Big(5 + \binom{6}{2}x + \binom{6}{3}x^2 + \binom{6}{4}x^3 + \binom{6}{5}x^4 + x^5\Big), \ x \neq 0.$$

Ponieważ dla każdego naturalnego k dla pochodnych rzędu k mamy

$$\mathcal{G}_X^{(k)}(1) = V^{(k)}(0)$$

oraz z twierdzenia Maclaurina wiadomo, że $\frac{V^{(k)}(0)}{k!}$ jest współczynnikiem wielomianu V stojącym przy x^k, więc w naszym przypadku

$$\mathcal{G}_X'(1) = V'(0) = \frac{1}{5}\binom{6}{2} = 3.$$

Analogicznie

$$\mathcal{G}_X''(1) = V''(0) = 2!\frac{1}{5}\binom{6}{3} = 8.$$

Dlatego ze wzorów 9.3 i 9.4 wynika, że

$$\mathbf{EX} = 2, \ oraz \ \mathbf{EX}^2 = 11,$$

skąd $var(\mathbf{X}) = 7$.

Zadanie 9.1.3 *Zmienna losowa* $\mathbf{X}$ *ma rozkład dyskretny* d *na zbiorze* $\{1, 3, 5, 9, 10\}$ *taki, że* $d(i) = \frac{1}{5}$. *Skonstruować dla niej funkcję tworzącą i na tej podstawie obliczyć jej pierwsze dwa momenty.*

Rozwiązanie

Zauważmy, że rozkład zmiennej losowej $\mathbf{X}$ nie jest rozkładem standardowym dyskretnym. Pokażemy, w jaki sposób można go zmodyfikować, aby moża go było traktować jako rozkład standardowy. W tym celu należy przyjąć, że zmienna losowa $\mathbf{X}$ przyjmuje wartości ze zbioru $\{2, 4, 6, 7, 8\}$ z prawdopodobieństwem 0.

Wtedy jej funkcja tworząca będzie miała postać

$$\mathcal{G}_X(x) = 0 + \frac{1}{5}x + 0x^2 + \frac{1}{5}x^3 + 0x^4 + \frac{1}{5}x^5 + 0x^6 + 0x^7 + 0x^8 + \frac{1}{5}x^9 + \frac{1}{5}x^{10} =$$

$$\frac{1}{5}(x + x^3 + x^5 + x^9 + x^{10}).$$

Wobec tego, ponieważ

$$\mathcal{G}'(x) = \frac{1}{5}(1 + 3x^2 + 5x^4 + 9x^8 + 10x^9)$$

oraz

$$\mathcal{G}''_X(x) = \frac{1}{5}(6x + 20x^2 + 72x^7 + 90x^8)$$

dostaniemy

$$\mathbf{EX} = \mathcal{G}'(1) = \frac{28}{5},$$

$$\mathbf{EX}^2 = \mathcal{G}''_X(1) + \mathcal{G}'(1) = \frac{216}{5}.$$

Przykładem innego zastosowania pojęcia funkcji tworzącej jest następujące twierdzenie

Twierdzenie 9.1.2 *Niech* $\mathbf{X}$ *i* $\mathbf{Y}$ *będą zmiennymi losowymi o standardowym rozkładzie dyskretnym na tym samym zbiorze. Wtedy, jeśli zmienne te są stochastycznie niezależne, to*

$$\mathcal{G}_{X+Y} = \mathcal{G}_X \mathcal{G}_Y. \tag{9.5}$$

Zadanie 9.1.4 *Dana jest zmienna losowa* $\mathbf{X}$ *o standardowym rozkładzie dyskretnym skończonym. Uzasadnić, że* $\mathbf{X}$ *nie zależy stochastycznie od siebie wtedy i tylko wtedy, gdy z prawdopodobieństwem 1 przyjmuje wartość stałą.*

Rozwiązanie

Jeśli $\mathbf{X}$ nie zależy od siebie samej, to ze wzoru 9.5 dostaniemy

$$\mathcal{G}_{2X} = (\mathcal{G}_X)^2.$$

Przyjmując, że $d_X(i) = p_i,\ i = 0, 1, \ldots, n-1$ dostaniemy $d_{2X}(j) = p_j \ dla\ j = 2i,\ i = 0, \ldots, n-1\ i\ d_{2X} = 0$ dla pozostałych wartości j. Dlatego

$$\mathcal{G}_X(x) = \sum_{i=0}^{n-1} p_i x^i,\ \mathcal{G}_{2X}(x) = \sum_{i=0}^{n-1} p_i x^{2i}.$$

Ponieważ

$$(\mathcal{G}_X(x))^2 = \Big(\sum_{i=0}^{n-1} p_i x_i\Big)^2 = \sum_{i+j=0}^{2n-2} p_i p_j x^{i+j}$$

z porównania odpowiednich dwóch wielomianów dostaniemy

$$p_o = p_o^2,\ p_1 = 2p_o p_1,\ itd.$$

Stąd $p_o \in \{0,1\}$ i dlatego $p_o = 1$. Dlatego $p_i = 0$ *dla* $i \geq 1$, co pokazuje, że $\mathbf{X}$ przyjmuje stałą wartość z prawdopodobieństwem 1. Oczywiście, jeśli zmienna losowa przyjmuje stałą wartość z prawdopodobieństwem 1, to jest niezależna od siebie.

9.2 Zadania

Zadanie 9.2.1 *Napisać funkcję tworzącą dla rozkładu zdefiniowanego na zbiorze* $\{3,4,6,7,9\}$*, gdzie* $d(i) = \frac{1}{5}$*. Porównać tę funkcją z fumkcją tworzącą, jeśli wcześniej dokonamy standaryzacji tego rozkładu dyskretnego.*

Zadanie 9.2.2 *Uzasadnić wzór 9.3 z twierdzenia 9.1.1.*

Zadanie 9.2.3 *Uzasadnić wzór 9.4 z twierdzenia 9.1.1.*

Zadanie 9.2.4 *Podać postać funkcji tworzącej dla* $\mathbf{X} \in B(n,p)$*.*

Zadanie 9.2.5 *Korzystając z wyników poprzedniego zadania, wyliczyć dwa pierwsze momenty dla tej zmiennej.*

Zadanie 9.2.6 *Wiadomo, że jeśli* $\mathbf{X} = \mathbf{X}_1 + \ldots + \mathbf{X}_n$*, gdzie zmienne losowe* $\mathbf{X}_i$ *są parami stochastycznie niezależne i o jednakowym standardowym rozkładzie dwupunktowym, to* $\mathbf{X} \in B(n,p)$*. Na tej podstawie skonstruować funkcję tworzącą dla* $\mathbf{X}$*. Wynik porównać z rozwiązaniem zadania poprzedniego.*

Zadanie 9.2.7 *Niech* $\mathbf{X}$ *ma rozkład dyskretny na zbiorze* $\{0,1,\ldots,n-1\}$*, gdzie* $d(i) = \frac{1}{n}$*. Jak wiadomo wtedy* $\mathcal{G}_X(x) = \frac{1}{n}\frac{1-x^n}{1-x}$ *dla* $x \neq 1$*. Wykorzystując technikę przedstawioną w zadaniu 9.1.2, obliczyć dwa pierwsze momenty tej zmiennej, przyjmując* $n = 10$*.*

Zadanie 9.2.8 *Korzystając z pojęcia szeregu liczbowego i zasady różniczkowania szeregu potęgowego*

1. *skonstruować funkcję tworzącą dla* $\mathbf{X} \in \mathcal{P}(\lambda)$*.*

2. *na tej podstawie obliczyć dwa pierwsze momenty tej zmiennej.*

Zadanie 9.2.9 *Powtórzyć polecenie poprzedniego zadania dla rozkładu geometrycznego.*

Zadanie 9.2.10 *Niech* $\mathbf{X}$ *ma standardowy rozkład dyskretny skończony. Wyznaczyć funkcję tworzącą dla zmiennej losowej* $\mathbf{Y} = \mathbf{X} + k$*, gdzie* k *jest daną liczbą naturalną. Na tej podstawie znaleźć dwa pierwsze momenty tej zmiennej losowej.*

Rozdział 10

Twierdzenia graniczne

10.1 Wprowadzenie teoretyczne i przykłady

Niech $\mathbf{X}$ będzie zmienną losową posiadającą wartość oczekiwaną m i wariancję σ^2. Jak już wiemy, wtedy m możemy interpretować jako *średnią wartość* obserwacji zjawiska losowego opisywanego zmienną $\mathbf{X}$. Natomiast wariancja (właściwie σ) jest wtedy miarą *rozproszenia* wartości tej obserwacji względem swojej wartości oczekiwanej. Ponieważ $m = \int_{\mathbf{R}} x dF$, mówimy też, że wartość oczekiwana jest *uśrednieniem po przestrzeni fazowej rozkładu* (w naszym przypadku jest to $\mathbf{R}$).

Podstawowy problem teorii prawdopodobieństwa sprowadza się do pytania:

> w jaki sposób można zrekonstruować nieznaną wartość parametru m rozkładu?

Kluczem do rozwiązania tego problemu okazuje się być zabieg, który zaobserwowaliśmy wcześniej na przykładzie rozkładu dwumianowego. Polega on na tym, aby w sposób niezależny wielokrotnie powtarzać tę zamą obserwację, rejestrując w ten sposób frekwencję zaobserwowanego zjawiska. Dokładniej, jeśli $\mathbf{X}_1, \ldots, \mathbf{X}_n$ jest ciągiem niezależnych zmiennych losowych o tym samym rozkładzie co $\mathbf{X}$, to w celu opisanym wyżej należy wziąć zmienną losową postaci

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i. \tag{10.1}$$

W takim przypadku mówimy, że wynik naszej obserwacji *uśredniliśmy po czasie*. Okazuje się, do czego zmierzamy, że efektem asymptotycznego uśrednienia po czasie jest uśrednienie po przestrzeni fazowej. Prawo to odkrył po raz pierwszy największy z rodziny Bernoullich–Jakub, formułując to w postaci *praw wielkich liczb*.

Twierdzenie 10.1.1 *(Słabe prawo wielkich liczb Bernoulliego) Niech zmienne losowe* $\mathbf{X}$, $\mathbf{X}_1, \ldots, \mathbf{X}_n$ *będą jak wyżej. Wtedy*

$$\forall_{\epsilon>0}\ P\Big(\{\omega \in \Omega\colon\ |\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i(\omega) - m| \geq \epsilon\}\Big) \to 0.$$

Mocniejsza wersja tego twierdzenia powiada, że

Twierdzenie 10.1.2 *(Mocne prawo wielkich liczb Bernoulliego) Niech zmienne losowe* $\mathbf{X}$, $\mathbf{X}_1, \ldots, \mathbf{X}_n$ *będą jak wyżej. Wtedy*

$$P\Big(\{\omega \in \Omega\colon\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i(\omega) \to m\}\Big) = 1.$$

Jako wniosek uzyskujemy wersję prawa wielkich liczb zwaną *Twierdzeniem Moivre'a-Laplace'a*

Wniosek 10.1.1 *Niech* $\mathbf{X}_n \in B(n,p)$*. Wtedy*

$$P\Big(\{\omega \in \Omega\colon\ \frac{1}{n}\mathbf{X}_n(\omega) \to p\}\Big) = 1. \tag{10.2}$$

Możemy iść dalej i zapytać się o sam rozkład zmiennej losowej $\mathbf{X}$, o którym wiadomo tylko, że posiada parametry $\mathbf{m}$ i σ^2. Tę kwestię rozwiązują tzw. *Centralne twierdzenia graniczne.* Jego podstawową wersję zwaną *twierdzeniem Lindenberga-Lévy'ego* zacytujemy poniżej

Twierdzenie 10.1.3 *(Twierdzeniem Lindenberga-Lévy'ego) Niech zmienne loswe* $\mathbf{X}$, $\mathbf{X}_1, \ldots, \mathbf{X}_n$ *będą jak wyżej. Weźmy ciąg zmiennych losowych* $\mathbf{Z}_n$ *będących wynikiem standaryzacji zmiennych* $\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i$*, czyli*

$$\mathbf{Z}_n = \frac{\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i - m}{\frac{\sigma}{\sqrt{n}}}. \tag{10.3}$$

Wtedy

$$\forall_{x\in\mathbf{R}}\ F_{\mathbf{Z}_n}(x) \to \Phi(x), \tag{10.4}$$

co zapisujemy także

$$\mathbf{Z}_n \xrightarrow{d} \mathcal{N}(0,1)$$

i czytamy, że

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i \quad \textit{asymptotycznie ma rozkład bliski } \mathcal{N}\Big(m, \frac{\sigma^2}{n}\Big).$$

Zadanie 10.1.1 *Uzasadnić twierdzenie Moivre'a-Laplace'a.*

Rozwiązanie

Niech $\mathbf{X}_n \in B(n,p)$. Jak dobrze wiemy, istnieje wtedy ciąg zmiennych losowych $(\mathbf{Y}_j)$ parami niezależnych i o jednakowym rozkładzie standardowym dwupunktowym z parametrem p, że

$$\mathbf{X}_n = \mathbf{Y}_1 + \ldots + \mathbf{Y}_n.$$

Ponieważ w naszym przypadku $\mathbf{m} = p$, z mocnego prawa wielkich liczb dostaniemy

$$P\Big(\{\omega \in \Omega\colon\ \frac{1}{n}\mathbf{X}_n(\omega) \to p\}\Big) = 1,$$

co kończy dowód twierdzenia.

Zadanie 10.1.2 *Zastosować centralne twierdzenie graniczne dla rozkładu dwumianowego.*

Rozwiązanie

Skorzystamy ze wstępu do rozwiązania zadania poprzedniego. W wyniku standaryzacji zmiennej losowej $\frac{1}{n}\mathbf{X}_n$ dostaniemy kolejno

$$\frac{\frac{1}{n}\mathbf{X}_n - p}{\sqrt{\frac{pq}{n}}} = \frac{\mathbf{X}_n - np}{\sqrt{npq}},$$

co na mocy twierdzenia 10.1.3 dowodzi, że $\mathbf{X}_n$ asymptotycznie ma rozkład bliski $\mathcal{N}(np, npq)$, czyli

$$\forall_{a<b}\ P\Big(\{\omega \in \Omega\colon\ a \le \frac{\frac{1}{n}\mathbf{X}_n(\omega) - p}{\sqrt{\frac{pq}{n}}} < b\}\Big) \to \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx = \Phi(b) - \Phi(a). \tag{10.5}$$

Zadanie 10.1.3 *Dla* $n = 300$, $p = 0,25$ *oszacować prawdopodobieństwo, że liczba sukcesów będzie mniejsza aniżeli* 250.

Rozwiązanie

Zastosujemy wynik 10.5. Rachunek będzie wyglądał następująco

$$P(\{\omega \in \Omega\colon\ \mathbf{X}_{300} < 250\}) = P\Big(\omega \in \Omega\colon\ \frac{\mathbf{X}_{300}(\omega) - 75}{\sqrt{300 \cdot 0,25 \cdot 0,75}} < \frac{250 - 75}{\sqrt{300 \cdot 0,25 \cdot 0,75}}\Big) \cong$$

$$\Phi(23,3) \cong 1.$$

Zadanie 10.1.4 *Dana jest zmienna losowa* $\mathbf{X} \in \mathcal{W}(\lambda)$, $\lambda = 2$. *Oszacować metodą nierówności Markowa i Czebyszewa oraz metodą twierdzenia granicznego*

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) > 60\}),$$

jeśli

$$\mathbf{Y} = \sum_{j=1}^{100} \mathbf{X}_j \text{ \textit{oraz} } \mathbf{X}_j \text{ \textit{są parami niezależne o rozkładzie} } \mathbf{X}.$$

Rozwiązanie

Ponieważ $\mathbf{Y} > 0$ oraz $\mathbf{EY} = 100\frac{1}{\lambda} = 50$, z nierówności Markowa dostaniemy

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) > 60\}) = P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \geq 60\}) \leq \frac{50}{60} = \frac{5}{6}.$$

Zobaczymy jaki efekt otrzymamy stosując nierówność Czebyszewa. Oczywiście

$$var(\mathbf{Y}) = 100 \cdot var(\mathbf{X}) = \frac{100}{\lambda^2} = 25.$$

Z drugiej strony, ponieważ

$$\{\omega \in \Omega\colon\ |\mathbf{Y}(\omega) - m| \geq t\} = \{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \geq m + t\} \cup \{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \leq m - t\},$$

więc

$$\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) > 60\} \subset \{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \geq 60\} = \{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \geq 50 + 10\} \subset$$

$$\{\omega \in \Omega\colon\ |\mathbf{Y}(\omega) - 50| \geq 10\}.$$

Ostatecznie otrzymamy

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) > 60\}) \leq P(\{\omega \in \Omega\colon\ |\mathbf{Y}(\omega) - 50| \geq 10\}) \leq \frac{1}{100} \cdot 25 = 0,25.$$

Uzyskane oszacowanie jest zatem o wiele lepsze aniżeli w przypadku nierówności Markowa. Na koniec porównamy uzyskane wyniki z metodą centralnego twierdzenia granicznego. Po uśrednieniu i standaryzacji zmiennej losowej $\mathbf{Y}$ dostaniemy

$$\frac{\frac{1}{100}\mathbf{Y} - \frac{1}{2}}{\frac{1}{20}} = \frac{\mathbf{Y} - 50}{5} \overset{d}{\cong} \mathcal{N}(0,1).$$

Dlatego

$$P(\{\omega \in \Omega\colon\ \mathbf{Y} > 60\}) = 1 - P(\{\omega \in \Omega\colon\ \mathbf{Y} \leq 60\}) = 1 - P(\{\omega \in \Omega\colon\ \mathbf{Y} < 60\}) =$$

$$1 - P\Big(\{\omega \in \Omega\colon\ \frac{\mathbf{Y}(\omega) - 50}{5} < \frac{10}{5}\}\Big) \cong 1 - \Phi(2) = 1 - 0,977250 = 0,02275.$$

Oznacza to, że tą metodą otrzymane oszacowanie jest najdokładniejsze, o ile $n = 100$ jest dostatecznie dużą liczbą w rozumieniu centralnego twierdzenia granicznego.

10.2 Zadania

Zadanie 10.2.1 *Niech niezależne zmienne losowe* $\mathbf{X}_1, \ldots, \mathbf{X}_{50}$ *mają jednakowy rozkład typu* $J([1,2])$*. Oszacować metodą twierdzenia granicznego prawdopodobieństwo*

$$P\Big(52 < \sum_{j=1}^{50} \mathbf{X}_j < 98\Big).$$

Zadanie 10.2.2 *Wiadomo, że zmienne losowe* $\mathbf{X}_1, \ldots, \mathbf{X}_{150}$ *są niezależne i o jednakowym rozkładzie wykładniczym z parametrem* $\lambda = 2$*. Bierzemy zmienną losową* $\mathbf{X} = \sum_{j=1}^{150} \mathbf{X}_j$*. Oszacować prawdopodobieństwo*

$$P(\mathbf{X} < 20).$$

Zadanie 10.2.3 *Wykonano* n *niezależnych powtórzeń pewnego doświadczenia, które polega na tym, że zdarzenie* A *zachodzi z prawdopodobieństwem* $\frac{1}{4}$*. Niech* $\mathbf{X}_n$ *oznacza ilość zajść zdarzenia* A *w* n *powtórzeniach. Stosując twierdzenie Moivre'a-Laplace'a, oszacować prawdopodobieństwo*

$$P\Big(|\frac{\mathbf{X}_n}{n} - \frac{1}{4}| \leq 10^{-2}\Big),$$

przyjmując $n = 10^3$*.*

Zadanie 10.2.4 *Niezależne zmienne losowe* $\mathbf{X}_1, \ldots, \mathbf{X}_{80}$ *mają rozkład typu* $\mathcal{P}(\lambda)$*, gdzie*
$\lambda = 3$*. Oszacować prawdopodobieństwo*

$$P\Big(\sum_{j=1}^{j=80} \mathbf{X}_j > 200\Big).$$

Zadanie 10.2.5 *Obserwowane żarówki w ilości* $n = 1000$ *przepalają się w sposób niezależny. Oszacować prawdopodobieństwo, że w czasie* T *spośród* n *żarówek przestanie świecić od 4 do 12, jeśli wiadomo, że prawdopodobieństwo, iż w czasie* T *przestanie świecić pojedyncza żarówka, wynosi* $0,25$.

Zadanie 10.2.6 *Produkt finalny charakteryzuje się swoją wadliwością określoną na poziomie* $0,05$. *Oszacować prawdopodobieństwo, że wśród losowo wybranych 200 sztuk tego produktu procent wadliwych sztuk różni się od 20 o co najmniej* $0,1$.

Zadanie 10.2.7 *Wiadomo, że zmienne losowe* $\mathbf{X}_i,\ i = 1,\ldots,150$ *są niezależne i o jednakowym rozkładzie, takim że* $P(\{\omega \in \Omega\colon\ \mathbf{X}_i(\omega) = k\}) = (\frac{1}{2})^k,\ k \in \mathbf{N}$. *Oszacować prawdopodobieństwo, że średnia arytmetyczna tego ciągu zmiennych losowych przyjmuje wartości z przedziału* $(0,\ 10)$.

Zadanie 10.2.8 *W pewnej populacji ludzkiej co setna osoba jest leworęczna. Oszacować prawdopodobieństwo, że wśród 1000 losowo wybranych ludzi od 10 do 30 osób będzie leworęcznych.*

Zadanie 10.2.9 *Nadzorujemy układ złożony z* n *elementów. Układ pracuje poprawnie, jeśli wiadomo, że co najmniej* 75% *elementów składowych jest sprawna. Z ilu elementów powinien składać się ten układ, aby z prawdopodobieństwem* $0,95$ *mógł pracować poprawnie, jeśli wiadomo, że prawdopodobieństwo awarii pojedynczego elementu wynosi* $0,25$*?*

Zadanie 10.2.10 *Niech* $(\mathbf{X}_n)$ *będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie co zmienna losowa* $\mathbf{X}$, *gdzie* $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) = \frac{1}{k^2}\}) = \frac{1}{2^k}$ *dla* $k \in \mathbf{N}$. *Czy dla tego ciągu można stosować centralne twierdzenie graniczne?*

Rozdział 11

Wprowadzenie do statystyki. Elementy estymacji

11.1 Wprowadzenie teoretyczne i przykłady

Niech R będzie m-elementowym zbiorem. Przypuśćmy, że każdy jego element $r \in R$ można opisać językiem teorii prawdopodobieństwa, a więc za pomocą rozkładu $F_{\mathbf{X}}$ pewnej nieznanej zmiennej losowej $\mathbf{X}$. Wtedy mnogość R będziemy nazywali *populacją generalną*, a zmienną losową $\mathbf{X}$ jej *cechą*.

Niech R_o będzie n-elementowym podzbiorem populacji generalnej R, gdzie $n << m$ (dużo mniejsze). Załóżmy, że każdemu $r_i \in R_o$ odpowiada kopia cechy $\mathbf{X}$, oznaczmy ją przez $\mathbf{X}_i$ taką, że zmienne losowe $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ są parami stochastycznie niezależne.

Wtedy odwzorowanie

$$\Omega \ni \omega \to (\mathbf{X}_1, \ldots, \mathbf{X}_n)(\omega) = (\mathbf{X}_1(\omega), \ldots, \mathbf{X}_n(\omega)) \in \mathbf{R}^n$$

będziemy nazywali *wektorem losowym* odpowiadającym reprezentacji P_o populacji generalnej P. Załóżmy, że obserwacja reprezentacji P_o populacji generalnej da *materiał statystyczny* w postaci ciągu liczbowego $(x_1, \ldots, x_n)$. Powiemy, że taki ciąg jest *próbą prostą* z populacji generalnej P, jeśli

$$\exists_{\omega_o \in \Omega}\ (\mathbf{X}_1, \ldots, \mathbf{X}_n)(\omega_o) = (x_1, \ldots, x_n).$$

Niech teraz $g\colon\ \mathbf{R}^n \to \mathbf{R}$ będzie taką funkcją, że

$$\mathbf{Z}(\omega) = g(\mathbf{X}_1(\omega), \ldots, \mathbf{X}_n(\omega))$$

jest zmienną losową. Wtedy $\mathbf{Z}$ będziemy nazywali *statystyką* populacji generalnej cechy $\mathbf{X}$.

Spośród wielu statystyk ważne są trzy następujące:

1.

$$\overline{\mathbf{X}}_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i$$

zwana *średnią teoretyczną z próby prostej,*

2.

$$\mathbf{S}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbf{X}_i - \overline{\mathbf{X}}_n\right)^2$$

zwana *wariancją teoretyczną z próby prostej,*

3.

$$\widehat{\mathbf{S}}^2 = \frac{n}{n-1}\mathbf{S}^2$$

zwana statystyką *z daszkiem.*

Wtedy wartości tych statystyk w punkcie ω_o nazwiemy *wartościami zaobserwowanymi* albo *empirycznymi* tych statystyk i oznaczymy je odpowiednio małymi literami, czyli

$$\overline{x}_n = \overline{\mathbf{X}}_n(\omega_o),\ s^2 = \mathbf{S}^2(\omega_o),\ \widehat{s^2} = \widehat{\mathbf{S}}^2(\omega_o).$$

Fakt 11.1.1 *Jeśli cecha* $\mathbf{X}$ *populacji generalnej ma wartość oczekiwaną* m *i wariancję* σ^2*, to*

$$\mathbf{E}\overline{\mathbf{X}}_n = m,\ \mathbf{E}\widehat{\mathbf{S}}^2 = \sigma^2.$$

Twierdzenie 11.1.1 *(o trzech statystykach) Niech cecha* $\mathbf{X}$ *populacji generalnej ma rozkład* $\mathcal{N}(m, \sigma^2)$*. Wtedy statystyka*

1.

$$\frac{n\mathbf{S}^2}{\sigma^2}$$

ma rozkład chi-kwadrat Pearsona o $n-1$ *stopniach swobody,*

2.

$$\frac{\overline{\mathbf{X}}_n - m}{\mathbf{S}}\sqrt{n-1}$$

ma rozkład t-Studenta o $n-1$ *stopniach swobody,*

3.

$$\frac{\overline{\mathbf{X}}_n - m}{\sigma}\sqrt{n}$$

ma rozkład standardowy normalny.

Przypuśćmy, że na temat cechy $\mathbf{X}$ wiemy:

1. znamy typ jej rozkładu, czyli jej dystrybuantę $F_{\mathbf{X}}(t, \theta)$, która zależy od parametru θ nieznanej wartości, albo
2. nie znamy typu rozkładu $F_{\mathbf{X}}(t, \theta)$, ale wiemy, że cecha ta ma wartość oczekiwaną i wariancję.

Jeśli potrafimy skonstruować dwie statystyki $\mathbf{Z}_j = f_j(\mathbf{X}_1, \ldots, \mathbf{X}_n),\ j = 1, 2$, takie, że

1.

$$\forall_{\omega \in \Omega}\ \mathbf{Z}_1(\omega) < \mathbf{Z}_2(\omega),$$

2.

$$P(\{\omega \in \Omega\colon\ \theta \in (\mathbf{Z}_1(\omega), \mathbf{Z}_2(\omega))\}) = 1 - \alpha,$$

to powiemy, że dokonaliśmy *estymacji przedziałowej* wartości parametru θ rozkładu cechy $\mathbf{X}$, a zdarzenie $\{\omega \in \Omega\colon\ \theta \in (\mathbf{Z}_1(\omega), \mathbf{Z}_2(\omega))\}$ nazwiemy *przedziałem losowym* odpowiadającym temu parametrowi. Biorąc wartości empiryczne z_1 *i* z_2 odpowiadające statystykom $\mathbf{Z}_1$ i $\mathbf{Z}_2$, będziemy mówili, że

$$\theta \in (z_1, z_2) \text{ z prawdopodobieństwem } 1 - \alpha.$$

Zadanie 11.1.1 *Wiadomo, że cecha* $\mathbf{X}$ *populacji generalnej ma rozkład normalny* $\mathcal{N}(m, \sigma^2)$*, gdzie wartość parametru* m *jest nieznana. Na podstawie próby prostej* $(x_1, \ldots, x_n)$ *skonstruować przedział losowy dla tego parametru.*

Rozwiązanie

Dla skonstruowania odpowiednich statystyk wykorzystamy Twierdzenie 11.1.1 (3). W tym celu niech

$$\mathbf{Z}_1 = \overline{\mathbf{X}}_n - n_\alpha \frac{\sigma}{\sqrt{n}},\ \mathbf{Z}_2 = \overline{\mathbf{X}}_n + n_\alpha \frac{\sigma}{\sqrt{n}},$$

gdzie $\alpha \in (0,1)$ i $2(1-\Phi(n_\alpha)) = \alpha$.

Istotnie, wtedy

$$P(\{\omega \in \Omega\colon\ m \in (\mathbf{Z}_1(\omega), \mathbf{Z}_2(\omega))\}) = P\Big(\{\omega \in \Omega\colon\ -n_\alpha < \frac{\overline{\mathbf{X}}_n(\omega) - m}{\frac{\sigma}{\sqrt{n}}} < n_\alpha\}\Big) =$$

$$P\Big(\{\omega \in \Omega\colon\ |\frac{\overline{\mathbf{X}}_n(\omega) - m}{\frac{\sigma}{\sqrt{n}}}| < n_\alpha\}\Big).$$

Ale z Twierdzenia 1.11.1(3) zmienna losowa $\frac{\overline{\mathbf{X}}_n(\omega)-m}{\frac{\sigma}{\sqrt{n}}} \in \mathcal{N}(0,1)$, zatem dla $\alpha \in (0,1)$ dostaniemy

$$1-\alpha = P(\{\omega \in \Omega\colon\ m \in (\mathbf{Z}_1(\omega), \mathbf{Z}_2(\omega))\}) = 1 - 2(1-\Phi(n_\alpha)),$$

czyli $\Phi(n_\alpha) = 1 - \frac{\alpha}{2}$.

Zadanie 11.1.2 *Przeprowadzić symulację liczbową dla sytuacji opisanej w zadaniu 11.1.1 przyjmując*

$\alpha = 0,05,\ n = 9,\ (1,74,\ 2,01,\ 1,81,\ 1,45,\ 1,78,\ 1,90,\ 1,95,\ 1,99,\ 1,87)$ *i* $\sigma^2 = 6$.

Rozwiązanie

Ponieważ $\overline{x}_9 = 1,83,\ \ \Phi(n_\alpha) = 1 - \frac{\alpha}{2} \Rightarrow n_\alpha = 1,96$ (z tablicy standardowego rozkładu normalnego), z wyników zadania 11.1.1 otrzymamy

$$m \in (1,83 - 1,96\frac{\sqrt{6}}{3},\ 1,83 + 1,96\frac{\sqrt{6}}{3}) = (0,23,\ 3,43)$$

z prawdopodobieństwem $0,95$.

Zadanie 11.1.3 *Na temat cechy* $\mathbf{X}$ *populacji generalnej wiadomo, że ma rozkład normalny, gdzie oba parametry są nieznane. Na podstawie próby prostej skonstruować przedział losowy dla parametru* m.

Rozwiązanie

Rozważmy dwie statystyki:

$$\mathbf{Z}_1 = \overline{\mathbf{X}}_n - t_\alpha \frac{\mathbf{S}}{\sqrt{n-1}},\ \mathbf{Z}_2 = \overline{\mathbf{X}}_n + t_\alpha \frac{\mathbf{S}}{\sqrt{n-1}},$$

dla pewnej liczby t_α zależnej od $\alpha \in (0,1)$ takiej, że

$$P(\{\omega \in \Omega\colon\ m \in (\mathbf{Z}_1(\omega), \mathbf{Z}_2(\omega))\}) = 1 - \alpha.$$

Aby wyznaczyć wartość t_α zauważmy, że

$$1 - \alpha = P\Big(\{\omega \in \Omega\colon\ -t_\alpha < \frac{\overline{\mathbf{X}}_n(\omega) - m}{\mathbf{S}}\sqrt{n-1} < t_\alpha\}\Big) =$$

$$P\Big(\{\omega \in \Omega\colon\ |\frac{\overline{\mathbf{X}}_n(\omega) - m}{\mathbf{S}}\sqrt{n-1}| < t_\alpha\}\Big) = 1 - P(\{\omega \in \Omega\colon\ |t_{n-1}(\omega)| \geq t_\alpha\}),$$

bowiem z twierdzenia 11.1.1(2) ostatnia zmienna losowa ma rozkład t-Studenta o $n-1$ stopniach swobody. Wówczas t_α jest wartością krytyczną dla tego rozkładu, czyli rozwiązaniem równania

$$P(\{\omega \in \Omega\colon\ |t_{n-1}(\omega)| \geq t_\alpha\}) = \alpha.$$

Zadanie 11.1.4 *Wykorzystując dane liczbowe z zadania 11.1.2 przeprowadzić symulację liczbową opisaną w zadaniu 11.1.3.*

Rozwiązanie

Z tabeli wartości krytycznych dla rozkładu t-Studenta o 8 stopniach swobody i $\alpha = 0,05$ mamy

$$P(\{\omega \in \Omega\colon\ |t_8(\omega)| \geq t_\alpha\}) = 0,006 \Rightarrow t_\alpha = 2,306.$$

Ponadto

$$s^2 = \frac{1}{9}\sum_{i=1}^{9}(x_i - \overline{x}_n)^2 = \frac{1}{9}\Big(1,74 - 1,83)^2 + (2,01 - 1,83)^2 + (1,81 - 1,83)^2 +$$

$$(1,45 - 1,83)^2 + (1,78 - 1,83)^2 + (1,90 - 1,83)^2 + (1,95 - 1,83)^2$$

$$(1,99 - 1,83)^2 + (1,87 - 1,83)^2\Big) = 0,5343 \Rightarrow s = 0,731.$$

Ostatecznie

$$m \in \Big(\overline{x}_9 - t_\alpha\frac{s}{\sqrt{8}},\ \overline{x}_9 + t_\alpha\frac{s}{\sqrt{8}}\Big) = \Big(1,83 - 2,306\frac{0,731}{2,828},\ 1,83 + 2,306\frac{0,731}{2,828}\Big).$$

Dlatego $m \in (1,24,\ 2,43)$ z prawdopodobieństwem $0,95$.

Zadanie 11.1.5 *Na temat cechy* $\mathbf{X}$ *populacji generalnej wiemy tylko tyle, że ma wartość oczekiwaną i wariancję. Skonstruować przedział losowy dla nieznanej wartości oczekiwanej na podstawie próby prostej.*

Rozwiązanie

Tym razem wykorzystamy Twierdzenie 11.1.1(3). W tym celu weźmy dwie statystyki

$$\mathbf{Z}_1 = \overline{\mathbf{X}}_n - n_\alpha \frac{\widehat{\mathbf{S}}}{\sqrt{n}},\ \mathbf{Z}_2 = \overline{\mathbf{X}}_n + n_\alpha \frac{\widehat{\mathbf{S}}}{\sqrt{n}},$$

gdzie liczba n_α jest taka, że

$$1 - \alpha = P(\{\omega \in \Omega\colon\ m \in (\mathbf{Z}_1(\omega), \mathbf{Z}_2(\omega))\}).$$

Zaprezentujemy metodę pozwalającą wyliczyć wartość n_α. Przede wszystkim zauważmy, że

$$1 - \alpha = P\Big(\{\omega \in \Omega\colon\ -n_\alpha < \frac{\overline{\mathbf{X}}_n(\omega) - m}{\widehat{\mathbf{S}}(\omega)}\sqrt{n} < n_\alpha\}\Big).$$

Z Mocnego prawa wielkich liczb wiadomo, że $P(\{\omega \in \Omega\colon\ \widehat{\mathbf{S}}^2(\omega) \to \sigma^2\}) = 1$, ponieważ $\mathbf{E}\widehat{\mathbf{S}}^2 = \sigma^2$. Dlatego z centralnego twierdzenia granicznego

$$P\Big(\{\omega \in \Omega\colon\ -n_\alpha < \frac{\overline{\mathbf{X}}_n(\omega) - m}{\widehat{\mathbf{S}}(\omega)}\sqrt{n} < n_\alpha\}\Big) \cong \Phi(n_\alpha) - \Phi(-n_\alpha)$$

dla dostatecznie dużych n.

Stąd dla takich n dostaniemy

$$1 - \alpha \cong 1 - 2(1 - \Phi(n_\alpha)) \Leftrightarrow \Phi(n_\alpha) \cong 1 - \frac{\alpha}{2}.$$

Zadanie 11.1.6 *Przeprowadzić symulację liczbową dla sytuacji opisanej w zadaniu 11.1.5.*

Rozwiązanie

W tym przypadku próba prosta musi być duża. Załóżmy, że $n = 150$. Dla ułatwienia obliczeń przyjmijmy, że dla tej próby: $\overline{x}_{150} = 2,015$, $\widehat{s} = 0,181$ i $\alpha = 0,05$. Ponieważ wtedy $n_\alpha \cong 1,96$, to dostaniemy

$$m \in \Big(2,015 - 1,96\frac{0,181}{\sqrt{150}},\ 2,015 + 1,96\frac{0,181}{\sqrt{150}}\Big) = (1,989,\ 2,044)$$

z prawdopodobieństwem $0,95$.

Zadanie 11.1.7 *Wiadomo, że cecha* $\mathbf{X}$ *populacji generalnej ma rozkład normalny o nieznanej wartości wariancji. Na podstawie próby prostej skonstruować przedział losowy dla wariancji.*

Rozwiązanie

Skonstruujemy takie statystyki $\mathbf{Z}_1$ i $\mathbf{Z}_2$, że dla $\alpha \in (0,1)$

$$1-\alpha = P(\{\omega \in \Omega\colon\ \sigma^2 \in (\mathbf{Z}_1(\omega), \mathbf{Z}_2(\omega))\}).$$

W tym celu weźmy:

$$\mathbf{Z}_1 = \frac{n\mathbf{S}^2}{a},\ \mathbf{Z}_2 = \frac{n\mathbf{S}^2}{b},$$

gdzie n oznacza długość próby prostej, natomiast liczby a *i* b są tak dobrane, że

$$1-\alpha = P\Big(\{\omega \in \Omega\colon\ \sigma^2 \in \Big(\frac{n\mathbf{S}^2(\omega)}{a}, \frac{n\mathbf{S}^2(\omega)}{b}\Big)\}\Big) =$$

$$P\Big(\{\omega \in \Omega\colon\ b < \frac{n\mathbf{S}^2(\omega)}{\sigma^2} < a\}\Big).$$

Z twierdzenia 11.1.1(1) zmienna losowa $\frac{n\mathbf{S}^2(\omega)}{\sigma^2}$ ma rozkład typu χ^2_{n-1}, więc ostatnia równość oznacza, że

$$1-\alpha = P(\{\omega \in \Omega\colon\ \chi^2_{n-1}(\omega) < a\}) - P(\{\omega \in \Omega\colon\ \chi^2_{n-1}(\omega) < b\}).$$

Ze względu na to, że w tabeli rozkładu ch-kwadrat podane są jego wartości krytyczne, ostanią równość zapiszemy następująco

$$1-\alpha = P(\{\omega \in \Omega\colon\ \chi^2_{n-1}(\omega) \geq b\}) - P(\{\omega \in \Omega\colon\ \chi^2_{n-1}(\omega) \geq a\}).$$

Aby znaleźć a i b, zauważmy, że wystarczy przyjąć, że

$$P(\{\omega \in \Omega\colon\ \chi^2_{n-1}(\omega) \geq b\}) = 1 - \frac{\alpha}{2},\ P(\{\omega \in \Omega\colon\ \chi^2_{n-1}(\omega) \geq a\}) = \frac{\alpha}{2}.$$

Zadanie 11.1.8 *Przeprowadzić symulację liczbową dla sytuacji omówionej w zadaniu 11.1.7, biorąc:* $\alpha = 0,05$, $n = 6$ *i próbę prostą*

$$(0,01,\ 0,05,\ 0,21,\ 0,02,\ 0,07,\ 0,02).$$

Z tabeli rozkładu ch-kwadrat o 5 stopniach swobody i wyników zadania 11.1.7 dostaniemy

$$P(\{\omega \in \Omega\colon\ \chi^2_5(\omega) \geq a\}) = 0,025 \Rightarrow a = 12,832.$$

Podobnie

$$P(\{\omega \in \Omega\colon\ \chi^2_5(\omega) \geq b\}) = 1 - 0,025 \Rightarrow b = 0,831.$$

Ponieważ dla wybranej próby prostej $\overline{x}_6 = 0,0633$ oraz $s^2 = 0,0046$, dostaniemy

$$\sigma^2 \in (2,15 \cdot 10^{-3},\ 3,32 \cdot 10^{-2}) \text{ z prawdopodobieństwem } 0,95.$$

11.2 Zadania

Zadanie 11.2.1 *Podać przykłady populacji generalnych i ich cech.*

Zadanie 11.2.2 *Niech* $\mathbf{X}$ *oznacza cechę populacji generalnej. Pobrano próbę prostą*

$$(-0,95,\ -0,35,\ 0,\ 0,20,\ 0,25).$$

Obliczyć wartości empiryczne następujących statystyk:

$$\overline{\mathbf{X}},\ \mathbf{S},\ \widehat{\mathbf{S}}.$$

Zadanie 11.2.3 *Niech* $\mathbf{X}$ *będzie cechą populacji generalnej o nieznanej dystrybuancie* F, $(x_1, \ldots, x_n) = (\mathbf{X}_1(\omega_o), \ldots, \mathbf{X}_n(\omega_o))$ *próbą prostą. Dystrybuantą empiryczną cechy* $\mathbf{X}$ *nazywamy odwzorowanie*

$$\Omega \times \mathbf{R} \ni (\omega, x) \to F(\omega, x) = \frac{1}{n}|\{i\colon\ \mathbf{X}_i(\omega) < x\}|.$$

Wiadomo, że wtedy $P(\{\omega \in \Omega\colon\ F(\omega, x) \to F(x)\}) = 1$. *Możemy więc przyjąć, że* $F(\omega_o, x) \cong F(x)$ *dla każdego* x. *Wyznaczyć* $F(\omega_o, x)$ *dla próby prostej* $(1,23,\ 1,34,\ 1,54,\ 1,32,\ 1,67,\ 1,45)$.

Zadanie 11.2.4 *Kontroli podlega partia produkcji dziennej. Kontrola przeprowadzana jest wyrywkowo według zasady: detal wadliwy jest odrzucany.*

1. *Skonstruować model statystyczny opisanego zjawiska.*
2. *Przyjmując, że pobrano próbę prostą długości 1500, w której zarejestrowano 340 wadliwych produktów skonstruować przedział losowy dla wartości oczekiwanej cechy tej populacji przyjmując* $\alpha = 0,02$.

Zadanie 11.2.5 *Cecha* X *populacji generalnej ma rozkład* $\mathcal{N}(m, \sigma)$. *Pobrano próbę prostą*

$$(3,1,\ 3,5,\ 2,9,\ 2,7,\ 2,8).$$

Wiedząc, że $\sigma = 2$ *na poziomie istotności* $1 - \alpha = 0,95$, *skonstruować przedział dla wartości oczekiwanej.*

Zadanie 11.2.6 *Niech cecha* $\mathbf{X}$ *i próba prosta będą jak w zadaniu 11.2.5. Przyjmując, że* σ *jest nieznane na poziomie* $1 - \alpha = 0,95$, *skonstruować przedział dla wartości oczekiwanej.*

Zadanie 11.2.7 *Niech cecha* $\mathbf{X}$ *i próba prosta będą jak w zadaniu 11.2.5. Przyjmując, że* σ *jest nieznane na poziomie* $1 - \alpha = 0,95$, *skonstruować przedział dla wariancji.*

Rozdział 12

Testowanie hipotez statystycznych

12.1 Wprowadzenie teoretyczne i przykłady

Niech $\mathbf{X}$ będzie cechą populacji generalnej P o rozkładzie znanego typu $F(x, \theta)$ i nieznanej wartości parametru θ. Metoda estymacji przedziałowej pozwala na skonstruowanie przedziału losowego, czyli zdarzenia $\{\omega \in \Omega\colon\ \theta \in (\mathbf{Z}_1(\omega), \mathbf{Z}_2(\omega))\}$, którego prawdopodobieństwo wynosi $1 - \alpha$ dla $\alpha \in (0, 1)$. Mając teraz próbę prostą $(x_1, \ldots, x_n) = (\mathbf{X}_1(\omega_o), \ldots, \mathbf{X}_n(\omega_o))$, możemy otrzymać następujący wynik

$$\theta \in (\mathbf{Z}_1(\omega_o), \mathbf{Z}_2(\omega_o)) \text{ z prawdopodobieństwem } 1 - \alpha.$$

Bardzo często w takiej sytuacji przyjmuje się, że $\theta = \theta_o = \frac{1}{2}(\mathbf{Z}_1(\omega_o) + \mathbf{Z}_2(\omega_o))$.

Metoda testowania hipotez statystycznych wspiera proces wyboru wartości parametru θ_o. Polega ona na tym, że formułuje się tzw. *hipotezę zerową*

$$H_o\colon\ \theta = \theta_o$$

przeciwko tzw. *hipotezie alternatywnej*

$$H_1\colon\ \theta \triangle \theta_o,$$

gdzie symbol $\triangle$ na ogół reprezentuje jedną z trzech relacji: $=$, $<$, $>$. Podstawą procesu weryfikacji poprawności hipotezy zerowej, czyli ewentualnego jej wyboru, jest podzbiór $Q \subset \mathbf{R}$, zwany *obszarem krytycznym*.

Przypuśćmy, że $(x_1, \ldots, x_n) = (\mathbf{X}_1(\omega_o), \ldots, \mathbf{X}_n(\omega_o))$ jest próbą prostą tej populacji generalnej. Konstruujemy statystykę $\mathbf{Z} = f(\mathbf{X}_1, \ldots, \mathbf{X}_n)$, która dla $\alpha \in (0, 1)$ wyznacza obszar krytyczny Q według zasady

$$P(\{\omega \in \Omega\colon\ \mathbf{Z}(\omega) \in Q\}) = \alpha \text{ przy założeniu, że } \theta = \theta_o$$

czyli, że zaszła hipoteza H_o.

Jeśli teraz $\mathbf{Z}(\omega_o)$–wartość empiryczna tej statystyki, spełnia warunek

$$\mathbf{Z}(\omega_o) \in Q,$$

to hipotezę H_o odrzucamy na korzyść hipotezy alternatywnej H_1. W przeciwnym razie mówimy, że *nie ma powodów do odrzucenia* hipotezy zerowej i możemy ją przyjąć. W taki przypadku proces weryfikacji (czy $\theta = \theta_o$) zakończy się pomyślnie.

Zadanie 12.1.1 *Wiadomo, że cecha* $\mathbf{X}$ *populacji generalnej ma rozkład normalny z nieznaną wartością oczekiwaną* m*. Przeprowadzić procedurę weryfikacji hipotezy* $H_o : m = m_o$ *przeciwko hipotezie* $H_1 : m \triangle m_o$*.*

Rozwiązanie

Niech $(x_1, \ldots, x_n)$ będzie próbą prostą tej populacji generalnej. Weźmy statystyką

$$\mathbf{Z} = \frac{\overline{\mathbf{X}}_n - m}{\sigma}\sqrt{n} \textit{ oraz } \alpha \in (0, 1). \tag{12.1}$$

Postać obszaru krytycznego zależała będzie od typu relacji $\triangle$.
Rozpatrzymy trzy przypadki:

1.

$$\triangle \;\Leftrightarrow \neq, \textit{ czyli } H_1 : m \neq m_o.$$

Wtedy przy założeniu prawdziwości hipotezy H_o statystyka $\mathbf{Y}$ powstała ze statystyki danej wzorem 12.1 w wyniku podstawienia $m = m_o$ ma standardowy rozkład normalny oraz

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \in Q\}) = \alpha \Leftrightarrow$$

$$P(\{\omega \in \Omega\colon\ |\mathbf{Y}(\omega)| \geq n_\alpha\}) = \alpha \Leftrightarrow 2(1 - \Phi(n_\alpha)) = \alpha,$$

co oznacza, że

$$Q = (-\infty, -n_\alpha) \cup (n_\alpha, +\infty). \tag{12.2}$$

2.

$$\triangle \;\Leftrightarrow >, \textit{ czyli } H_1 : m > m_o.$$

Wtedy przy założeniu prawdziwości hipotezy zerowej dostaniemy

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \in Q\}) = \alpha \Leftrightarrow$$

$$P(\{\omega \in \Omega\colon\ \mathbf{Y}(\omega) \geq n_\alpha^+\}) = \alpha \Leftrightarrow 1 - \Phi(n_\alpha^+) = \alpha,$$

co oznacza, że

$$Q = (n_\alpha^+, +\infty). \tag{12.3}$$

3.

$$\triangle \;\Leftrightarrow <, \; czyli \; H_1 : \; m < m_o.$$

Wtedy przy założeniu prawdziwości hipotezy zerowej dostaniemy

$$P(\{\omega \in \Omega\colon \; \mathbf{Y}(\omega) \in Q\}) = \alpha \Leftrightarrow P(\{\omega \in \Omega\colon \; \mathbf{Y}(\omega) < -n_\alpha^+\}) = \alpha,$$

gdzie n_α^+ wyznaczamy jak wyżej. Oznacza to, że

$$Q = (-\infty, -n_\alpha^+). \tag{12.4}$$

Zadanie 12.1.2 *Z zadania 11.1.2 wiadomo, że dla* $\alpha = 0,05$ *i* $\sigma^2 = 6, \;\; m \in (0,23, \;\; 3,43)$ *z prawdopodobieństwem* $0,95$. *Przeprowadzić test statystyczny weryfikujący, czy należy przyjąć hipotezę*

$$H_o \; : \; m = m_o = \frac{1}{2}(0,23 + 3,43) \; przeciwko \; H_1 \; : \; m \neq m_o$$

dla $\alpha = 0,05$, *jeśli wiadomo, że pobrano próbę prostą taką, że* $\overline{x}_9 = 1,83$.

Rozwiązanie

Wyznaczymy obszar krytyczny Q dla tej sytuacji. Z zadania 12.1.1(1) dostaniemy

$$Q = (-\infty, -n_\alpha) \cup (n_\alpha, +\infty),$$

gdzie

$$2(1 - \Phi(n_\alpha)) = 0,05 \Leftrightarrow \Phi(n_\alpha) = 0,975 \Leftrightarrow n_\alpha = 1,96.$$

Przy założeniu hipotezy $H_o \; : \; m = m_o = 1,83$ dostaniemy

$$\mathbf{Z}(\omega_o) = \frac{\overline{x}_9 - m_o}{\sigma}\sqrt{n} = \frac{1,83 - 1,83}{\sqrt{6}}\sqrt{9} = 0.$$

Ponieważ $\mathbf{Z}(\omega_o) \;\; \notin \;\; Q$, więc nie ma powodu do odrzucenia hipotezy zakładającej, że $m = 1,83$ z prawdopodobieństwem $0,95$.

Zadanie 12.1.3 *Cecha* $\mathbf{X}$ *populacji generalnej ma rozkład normalny z nieznanymi parametrami* m *i* σ^2. *Przeprowadzić weryfikację hipotezy* $H_o \; : \; m = m_o$ *przeciwko hipotezie* $H_1 \; : \; m \triangle m_o$.

Rozwiązanie

Niech $(x_1, \ldots, x_n)$ będzie próbą prostą. W tym przypadku weźmiemy statystykę

$$\mathbf{Z} = \frac{\overline{\mathbf{X}}_n - m}{\mathbf{S}} \sqrt{n-1}. \tag{12.5}$$

Przy założeniu hipotezy $H_o : m = m_o$ statystyka ta ma rozkład t-Studenta o $n-1$ stopniach swobody, dlatego

$$t_{n-1} = \frac{\overline{\mathbf{X}}_n - m_o}{\mathbf{S}} \sqrt{n-1}. \tag{12.6}$$

Skonstruujemy teraz obszar krytyczny Q. W tym celu rozpatrzymy trzy przypadki:

1.

$$\triangle = \neq, \ czyli \ H_1 : m \neq m_o.$$

Wtedy przy założeniu hipotezy H_o i dla $\alpha \in (0,1)$ dostaniemy

$$P(\{\omega \in \Omega\colon \mathbf{Z}(\omega) \in Q\}) = \alpha \Leftrightarrow P(\{\omega \in \Omega\colon \mathbf{t}_{n-1}(\omega) \in Q\}) = \alpha,$$

co oznacza, że

$$Q = (-\infty, -t_\alpha) \cup (t_\alpha, +\infty), \tag{12.7}$$

gdzie t_α jest wartością krytyczną rozkładu t-Studenta o $n-1$ stopniach swobody, czyli $P(\{\omega \in \Omega\colon |\mathbf{t}_{n-1}|(\omega) \geq t_\alpha\}) = \alpha$.

2.

$$\triangle = >, \ czyli \ H_1 : m > m_o.$$

Wtedy przy założeniu hipotezy H_o i dla $\alpha \in (0,1)$ dostaniemy

$$P(\{\omega \in \Omega\colon \mathbf{Z}(\omega) \in Q\}) = \alpha \Leftrightarrow P(\{\omega \in \Omega\colon \mathbf{t}_{n-1}(\omega) \in Q\}) = \alpha,$$

co oznacza, że

$$Q = (t_\alpha^+, +\infty), \tag{12.8}$$

gdzie t_α^+ spełnia warunek $P(\{\omega \in \Omega\colon \mathbf{t}_{n-1}(\omega) \geq t_\alpha^+\}) = \alpha$.
Z zadania 7.1.13 wiemy, że wtedy $P(\{\omega \in \Omega\colon |\mathbf{t}_{n-1}|(\omega) \geq t_\alpha^+\}) = 2\alpha$.

3.

$$\triangle = <, \ czyli \ H_1 : m < m_o.$$

Wtedy przy założeniu hipotezy H_o i dla $\alpha \in (0,1)$ dostaniemy

$$P(\{\omega \in \Omega\colon \mathbf{Z}(\omega) \in Q\}) = \alpha \Leftrightarrow P(\{\omega \in \Omega\colon \mathbf{t}_{n-1}(\omega) \in Q\}) = \alpha,$$

co oznacza, że

$$Q = (-\infty, -t_\alpha^+), \tag{12.9}$$

gdzie t_α^+ spełnia warunek opisany w punkcie 2.

Zadanie 12.1.4 *Z zadania 11.1.4 wiemy, że* $m \in (1,24,\ 2,43)$ *z prawdopodobieństwem* $0,95$. *Przeprowadzić test weryfikujący przyjęcia hipotezy* $H_o : m = m_o = \frac{1}{2}(1,24 + 2,43) = 1,835$ *przeciwko hipotezie alternatywnej* $H_1 : m > m_o$ *dla* $\alpha = 0,05$.

Rozwiązanie

Przy założeniu hipotezy H_o zmienna losowa $\mathbf{Z}$ dana wzorem 12.5 ma rozkład t-Studenta o 8 stopniach swobody. Skonstruujemy obszar krytyczny dla hipotezy alternatywnej $H_1 : m > m_o$. Z sytuacji (2) poprzedniego zadania

$$Q = (t_\alpha^+, +\infty),$$

gdzie

$$P(\{\omega \in \Omega\colon\ |\mathbf{t}_8|(\omega) \geq t_\alpha^+\}) = 2 \cdot 0,05 = 0,1,$$

skąd $t_\alpha^+ = 1,86$ i ostatecznie $Q = (1,86,\ +\infty)$.

Ponieważ

$$\mathbf{Z}(\omega_o) = \frac{1,83 - 1,835}{0,731}\sqrt{8} = -6,839 \cdot 10^{-3},$$

więc $\mathbf{Z}(\omega_o) \notin Q$, co oznacza, że hipotezę H_o przyjmujemy.

Zadanie 12.1.5 *Na temat cechy* $\mathbf{X}$ *populacji generalnej wiemy tylko tyle, że ma rozkład normalny. Zweryfikować hipotezę* $H_o : \sigma^2 = \sigma_o^2$ *przeciwko hipotezie* $H_1 : \sigma^2 > \sigma_o^2$.

Rozwiązanie

Niech $(x_1, \ldots, x_n)$ *będzie próbą prostą. Wtedy obszar krytyczny* Q *musi spełniać warunek*

$$P(\{\omega \in \Omega\colon\ \mathbf{Z}(\omega) \in Q\}) = \alpha \textit{ przy założeniu, że } \sigma^2 = \sigma_o^2.$$

Bierzemy

$$\mathbf{Z} = \frac{n\mathbf{S}^2}{\sigma^2}. \tag{12.10}$$

Wtedy przy założeniu hipotezy H_o

$$\frac{n\mathbf{S}^2}{\sigma_o^2} = \chi_{n-1}^2 \tag{12.11}$$

oraz

$$Q = (\chi_\alpha^2, +\infty), \textit{ gdzie } P(\{\omega \in \Omega\colon\ \chi_{n-1}^2(\omega) \geq \chi_\alpha^2\}) = \alpha. \tag{12.12}$$

Zadanie 12.1.6 *Dla danych z zadania 11.1.8 zweryfikować hipotezę* $H_o : \sigma^2 = \sigma_o^2 = \frac{1}{2}(2,15 \cdot 10^{-3} + 3,32 \cdot 10^{-2}) = 0,17675$ *przeciwko hipotezie* $H_1 : \sigma^2 > \sigma_o^2$ *dla* $\alpha = 0,05$.

Rozwiązanie

Z tabeli wartości krytycznych dla rozkładu chi-kwadrat dla 5 stopni swobody mamy

$$P(\{\omega \in \Omega\colon\ \chi_5^2(\omega) \geq \chi_\alpha^2\}) = 0,05 \Leftrightarrow \chi_\alpha^2 = 11,070.$$

Wartość empiryczna statystyki 12.10 przy założeniu hipotezy zerowej wyniesie

$$\mathbf{Z}(\omega_o) = \frac{ns^2}{\sigma_o^2} = \frac{6 \cdot 0,0046}{0,17675} = 0,156,$$

co oznacza, że $\mathbf{Z}(\omega_o) \notin Q = (11,070,\ +\infty)$. Dlatego nie ma powodów do odrzucenia hipotezy, że $\sigma^2 = 0,17675$ z prawdopodobieństwem $0,95$.

12.2 Zadania

Zadanie 12.2.1 *Niech cecha* $\mathbf{X}$ *i próba prosta będą jak w zadaniu 11.2.5. Dla* $\alpha = 0,05$ *zweryfikować hipotezę* $H_o: \ m = 3$ *przeciwko hipotezie alternatywnej* H_1*:*

1. $m \neq 3$,
2. $m < 3$,
3. $m > 3$.

Zadanie 12.2.2 *Cecha* $\mathbf{X}$ *populacji generalnej ma rozkład normalny. Pobrano próbę prostą*

$$-1,5, \ 0, \ 0,51, \ 0,48, \ 0,62, \ 0,33, \ -0,33, \ 0,22, \ 0,49, \ 0,12.$$

Dla $\alpha = 0,05$ *zweryfikować hipotezę, że* $m = 0,35$ *przeciwko hipotezie alternatywnej* H_1

1. $m \neq 0,35$,
2. $m > 0,35$,
3. $m < 0,35$.

Zadanie 12.2.3 *Cecha* $\mathbf{X}$ *populacji generalnej ma rozkład normalny. Pobrano próbę prostą*

$$2,01, \ 1,99, \ 2,20, \ 3,51, \ 1,75, \ 1,66, \ 2,25, \ 2,31, \ 2,0.$$

Zweryfikować hipotezę, że $\sigma^2 = 0,2$ *przeciwko hipotezie* $\sigma^2 > 0,2$ *przyjmując* $\alpha = 0,05$.

Zadanie 12.2.4 *Z pobranej próby prostej cechy* $\mathbf{X}$ *populacji generalnej o rozkładzie normalnym wynika, że:*

$$m \in (1,888, \ 2,257), \ \overline{x}_4 = 2,072, \ s = 0,11.$$

Sprawdzić, czy dla tej populacji można przyjąć $m = 1,9$ *(*$\alpha = 0,05$*).*

Zadanie 12.2.5 *Na uczelni wylosowano 1500 studentów, których poddano ankietowaniu. Pytanie brzmiało, czy po skończonych studiach podejmą pracę zgodną z profilem swojego wykształcenia?. Na pytanie odpowiedziało 800 ankietowanych studentów. Na tej podstawie zweryfikować hipotezę, że* 75% *studiujących na tej uczelni podejmie pracę zgodnie z kierunkiem swojego wykształcenia. Przyjąć* $\alpha = 0,05$.

Zadanie 12.2.6 *Dla cechy* **X** *populacji generalnej o rozkładzie normalnym pobrano 20 elementową próbę prostą. Stwierdzono, że* $s^2 = 0,07$. *Przyjmując* $\alpha = 0,05$, *rozstrzygnąć, czy można stwierdzić, że* $\sigma^2 = 0,075$*?.*

Zadanie 12.2.7 *Dla cechy* **X** *populacji generalnej o rozkładzie normalnym pobrano próbę prostą, z której wynika, że:* $\overline{x}_{25} = 0,06$, $s^2 = 0,08$. *Czy dla* $\alpha = 0,05$ *można przyjąć, że wartość oczekiwana tej cechy wynosi* $0,035$*?*

Zadanie 12.2.8 *Przyrząd zarejestrował 6 niezależnie wykonanych pomiarów pewnej wielkości fizycznej:* $10,1$, $10,2$, $9,99$, $9,74$, $10,0$, $9,98$. *Zakładając, że rozkład pomiarów sporządzonych tym przyrządem jest normalny, zweryfikować hipotezę, że średnia wartość pomiaru tej wielkości fizycznej wyniesie* $9,99$ *dla* $\alpha = 0,05$.

Zadanie 12.2.9 *Wykorzystując dane liczbowe z poprzedniego zadania, zweryfikować hipotezę, że niepewność pomiaru wykonanego tym przyrządem (jego klasa) wynosi* $0,05$. *Przyjąć, że* $\alpha = 0,05$.

Zadanie 12.2.10 *Dla cechy* **X** *populacji generalnej o rozkładzie normalnym pobrano próbę prostą:* $-0,75$, $-0,61$, $-0,33$, $0,1$. *Dla* $\alpha = 0,05$ *zweryfikować hipotezę, że* $m = -0,5$.

Rozdział 13

Przykładowe zadania z kolokwiów i egzaminów

Zadania z kolokwiów

Zestaw 1

1. *Ile jest wszystkich możliwych sposobów podziału zbioru 5-elementowego na co najwyżej 3 podzbiory?*

2. *Winda rusza z 7 pasażerami i zatrzymuje się na 10 piętrach. Jakie jest prawdopodobieństwo $P(A)$ zdarzenia A, że żadnych dwóch pasażerów nie opuści windy na tym samym piętrze?*

3. *Ze zbioru $\{5, 6, 7, 8, 9\}$ losujemy bez zwracania trzy cyfry i zgodnie z kolejnością losowania tworzymy z nich liczbę trzycyfrową. Jakie jest prawdopodobieństwo tego, że ta liczba będzie większa od 750?*

4. *Ile liczb dwucyfrowych o niepowtarzających się cyfrach można utworzyć z cyfr $\{1, 2, 3, 4\}$?*

5. *W sposób losowy ustawiamy litery znajdujące się w słowie MAMA. Jakie jest prawdopodobieństwo tego, że dwie litery A będą stały obok siebie?*

6. *Na nieskończoną szachownicę o boku a rzucono monetę o średnicy $2r < a$. Obliczyć prawdopodobieństwo tego, że moneta przetnie co najwyżej jeden bok kwadratu (pola tej szachownicy)?*

Zestaw 2

1. *Ze zbioru $\{1, 2, 3, 4, 5\}$ losujemy bez zwracania trzy cyfry i zgodnie z kolejnością losowania tworzymy z nich liczbę trzycyfrową. Jakie jest prawdopodobieństwo tego, że ta liczba będzie mniejsza od 230?*

2. *Niech A będzie zbiorem punktów (x, y), dla których $x^2 + y^2 < 1$, B– zbiorem punktów (x, y), dla których $x^2 + y^2 < 4$, C– zbiorem punktów (x, y), dla których $(x - 1)^2 + y^2 < 1$. Znaleźć zbiory $A \cup B$, $B \cup C$, $A \cap B \cap C$ oraz $A - C$.*

3. *Na ile sposobów można ustawić na półce 30 książek, spośród których 20 jest mniejszego formatu, a 10 większego, tak by mniejsze książki nie były przemieszane z większymi?*

4. *Co jest bardziej prawdopodobne: w rzucie dwiema kostkami wyrzucenie sumy oczek podzielnej przez 4 czy w rzucie czterema monetami wyrzucenie więcej orłów niż reszek?*

5. *Ze zbioru $\{1, 2, 3, 4\}$ losujemy dwie cyfry bez zwracania. Jakie jest prawdopodobieństwo tego, że liczba utworzona z tych cyfr (zgodnie z kolejnością losowania) będzie parzysta?*

6. *Z odcinka $[0, 1]$ wybrano losowo i niezależnie dwa punkty x i y, które dzielą ten odcinek na trzy odcinki. Znaleźć prawdopodobieństwo tego, że z tych odcinków można zbudować trójkąt.*

Zestaw 3

1. *Windą 10-piętrowego bloku jedzie n osób. Jaka musi być liczba n, by prawdpodobieństwo tego, że wszystie osoby wysiądą na ostatnich trzech piętrach było mniejsze niż $0,03$?*

2. *Z talii 52 kart losujemy jedną kartę. Jakie jest prawdopodobieństwo tego, że wylosowana karta będzie królem, damą lub pikiem?*

3. *Autobus wiozący 20 pasażerów zatrzymuje się na 5 przystankach. Na ile różnych sposobów pasażerowie mogą wysiadać z autobusu, zakładając, że na każdym przystanku wysiądzie przynajmniej jeden pasażer?*

4. *Rozwiąż równanie $2 \cdot |C_n^2| = |C_{n+1}^3|$, gdzie $|C_n^k|$ oznacza liczbę wszystkich k-elementowych kombinacji zbioru n-elementowego.*

5. *Rzucamy dwiema kostkami. Jakie jest prawdopodobieństwo tego, że suma oczeka na tych kostkach będzie równa co najmniej 7?*

6. *W poczekalni przychodni lekarskiej znajdują się dwie kobiety i dwaj mężczyźni. Poszczególne osoby są proszone do gabinetu w sposób losowy. Jakie jest prawdopodobieństwo tego, że jako pierwsze będą zbadane kobiety?*

Zestaw 4

1. *Rzucamy jednocześnie czterema monetami. Wypisz zdarzenia elementarne sprzyjające zdarzeniu* $A \cup B$, *gdy:* A*–zdarzenie polegające na wyrzuceniu co najmniej dwóch reszek,* B*–zdarzenie polegające na wyrzuceniu nieparzystej liczby orłów.*

2. *Liczby* $\{1, 2, 3, 4\}$ *porządkujemy w spoób losowy. Jakie jest prawdopodobieństwo tego, że liczba 1 będzie stała jako pierwsza?*

3. *Na ile sposobów można rozmieścić 5 ponumerowanych kul w 3 urnach?*

4. *W urnie jest 6 kul: dwie białe, dwie czarne, dwie czerwone. Jakie jest prawdopodobieństwo tego, że losując kolejno 3 kule bez zwracania, wylosujemy jedną kulę czarną, jedną czerowną i jedną białą?*

5. *Wiadomo, że 30% śrub ma dodatnie (+) odchylenia wymiarów średnicy od nominalnego wymiaru, a 70%–ujemne (-). Spośród* $n = 100$ *wybrano trzy sztuki. Obliczyć prawdopodobieństwo tego, że:*
   a) jedna śruba jest ,,plusowa";
   b) nie ma żadnej śruby ,,plusowej"wśród trzech wybranych.

6. *Na odcinku* $[0, 1]$ *umieszczamy losowo i niezależnie punty* x *i* y*. Niech* A *oznacza zdarzenie polegające na tym, że* $x < y$*, natomiast* B *będzie zdarzeniem polegającym na tym, że* $y < 0,5$*. Czy zdarzenia* A *i* B *są niezależne?*

Zestaw 5

1. *Dziecko bawi się literami* $A, A, A, E, K, M, M, T, T, Y$*. Znaleźć prawdopodobieństwo tego, że przypadkowo złoży słowo MATEMATYKA.*

2. *Cztery osoby, w tym jedno małżeństwo, siadają losowo przy stoliku brydżowym. Jakie jest prawdopodobieństwo tego, że małżonkowie będą siedzieli naprzeciw siebie?*

3. *Na kartkach wrzuconych do pudełka napisane są odpowiednio numery* $1, 2, 3, 4$*. Losujemy w sposób przypadkowy jedną kartkę. Niech* A *oznacza zdarzenie polegające na wylosowaniu kartki z numerem 1;* B *oznacza zdarzenie polegające na wylosowaniu kartki z numerem parzystym. Obliczyć* $P(A), P(A^C), P(B), P(B^C)$*.*

4. *20-osobowa grupa studencka, w której jest 6 kobiet, otrzymała 5 biletów do teatru. Bilety rodziela się drogą losowania. Jakie jest prawdopodobieństwo tego, że wśród posiadaczy biletów znajdą się dokładnie 3 kobiety?*

5. *Ile różnych liczb czterocyfrowych o niepowtarzająch się cyfrach można utworzyć z cyfr* $5, 6, 7, 8$*?*

6. *Rozwiąż równanie* $20 \cdot |P_{n-2}| = |P_n|$, *gdzie* $|P_n|$ *oznacza liczbę wszystkich permutacji zbioru n-elementowego.*

Zestaw 6

1. *Na kartkach wrzuconych do pudełka napisane są odpowiednio liczby* $\{1, 2, 3, 4\}$. *Losujemy w sposób przypadkowy dwie kartki. Niech* A *oznacza zdarzenie polegające na wylowaniu pary liczb, których suma jest mniejsza od 5,* B*–zdarzenie polegające na wylosowaniu pary liczb, których suma jest większa od 4, a* C*–zdarzenie polegające na wylosowaniu pary liczb, z których przynajmniej jedna jest większa od 1. Obliczyć* $P(A), P(B), P(C)$.

2. *Na ile różnych sposobów można rozdzielić medale w zawodach sportowych pośród czterech finalistów?*

3. *Liczby* $\{1, 2, ..., n\}$ *zostały ustawione przypadkowo. Znaleźć prawdopodobieństwo, że a) cyfry 1 i 2, b) cyfry 1,2 i 3 pojawiły się w sąsiedztwie i w wymienionej kolejności?*

4. *Ze zbioru ośmiu kart składającego się z 4 dam i 4 waletów losujemy dwie karty. Jakie jest prawdopodobieństwo tego, że wśród wylosowanych kart jest jedna dama i jeden walet?*

5. *Na ile sposobów można wyznaczyć delegację złożoną z dwóch dziewczynek oraz dwóch chłopców z klasy, w której jest 20 dziewczynek i 15 chłopców?*

6. *Z partii towarów zawierającej sztuki dobre i wadliwe losujemy 3 sztuki. Niech* A *oznacza zdarzenie: dokładnie jedna sztuka dobra w trzech sztukach wylosowanych,* B*–co najwyżej jedna sztuka dobra w trzech wylosowanych,* C*–co najmniej jedna sztuka w trzech wylosowanych. Wyjaśnić, co oznaczają zdarzenia:* $A^C, A \cup B, B \cap C, B^C \cap C^C$.

Zestaw 7

1. *Zmienna losowa* X *ma rozkład dyskretny taki, że* $p(-3) = 0$, $p(-1) = a$, $p(3) = 0,5$, $p(5) = 0,2$.

   (a) *Wyznacz a;*

   (b) *Wyznacz dystrybuantę i narysuj jej wykres;*

   (c) *Oblicz* $P(0 \leq X \leq 5)$;

   (d) *Wyznacz* EX, D^2X.

2. *Prawdopodobieństwo wyprodukowania sztuki wadliwej wynosi 0,002. Oblicz prawdopodobieństwo, że w partii liczącej 500 sztuk znajdą się:*

   (a) *2 sztuki wadliwe;*

(b) *0 sztuk wadliwych;*

(c) *co najmniej 3 sztuki wadliwe.*

3. *Zmienna losowa X ma rozkład o gęstości*

$$F(x) = \begin{cases} x & \text{dla } 0 \leq x \leq 1 \\ a - x & \text{dla } 1 \leq x \leq 2 \\ 0 & \text{poza tym} \end{cases}$$

Wyznaczyć:

(a) *Stałą a;*

(b) *Dystrybuantę i jej wykres;*

(c) *Var(X).*

Zestaw 8

1. *Zmienna losowa X ma rozkład dyskretny taki, że $p(-1) = 0,2$, $p(2) = 0,3$, $p(3) = 0,4$, $p(4) = a$.*

   (a) *Wyznacz a;*

   (b) *Wyznacz dystrybuantę i narysuj jej wykres;*

   (c) *Oblicz $P(0 \leq X \leq 5)$;*

   (d) *Wyznacz EX, D^2X.*

2. *Urządzenie składa się między innymi z 750 lamp. Prawdopodobieństwo awarii lampy w ciągu doby wynosi $p = 0,004$. Obliczyć prawdopodobieństwo, że w ciągu doby pracy urządzenia ulegnie awarii:*

   (a) *0 lamp;*

   (b) *2 lampy;*

   (c) *co najmniej 3 lampy.*

3. *Dobrać tak stałą c, aby funkcja*

$$f(x) = \begin{cases} c \cdot sinx & \text{dla } 0 \leq x \leq \pi \\ 0 & \text{poza tym} \end{cases}$$

była gęstością pewnej zmiennej losowej X. Wyznaczyć:

   (a) *$P(|X| < \frac{\pi}{3})$;*

   (b) *EX;*

   (c) *Var(X).*

Zestaw 9

1. *Wiadomo, że zmienna* $\mathbf{X} \in \mathcal{J}([1,5])$. *Znaleźć rozkład zmiennej losowej* $\mathbf{Y} = \frac{1}{4}(\mathbf{X} - 1)$.

2. *Niech zmienna losowa* $\mathbf{X}$ *ma rozkład*

x_i	-2	-1	0	2	3
p_i	$0,1$	$0,2$	$0,3$	$0,2$	$0,2$

Wyznaczyć rozkład zmiennej losowej $\mathbf{Y} = -\mathbf{X} + 3$. *Narysować dystrybuantę* $F_{\mathbf{Y}}$.

3. *Z partii 100 przedmiotów, wśród których jest 10 wykonanych wadliwie wybrano losowo bez zwracania 8 sztuk. Niech* $\mathbf{X}$ *oznacza liczbę sztuk wadliwych. Wyznaczyć rozkład* $\mathbf{X}$ *i jej wartość oczekiwaną.*

4. *Wiadomo, że zmienna losowa* $\mathbf{X}$ *ma gęstość*

$$f(x) = \begin{cases} 2 - 2x, & gdy \quad x \in (0,1) \\ 0, & gdy \quad x \notin (0,1) \end{cases}$$

Znaleźć dystrybuantę tego rozkładu.

Zestaw 10

1. *Na odcinku* $[0,1]$ *umieszczono losowo i niezależnie punkty* x *i* y. *Niech*

$$A = \{(x,y) \in [0,1] \times [0,1]\colon\ x^2 + y^2 \leq 1\},$$

$$B = \{(x,y) \in [0,1] \times [0,1]\colon\ x < y\}.$$

Czy zdarzenia te są stochastycznie niezależne?

2. *Dla jakich wartości* $a,\ b \in \mathbf{R}$ *funkcja*

$$F(x) = \begin{cases} 0, & gdy \qquad x \leq 1 \\ x^3 + a, & gdy \quad -1 < x \leq b \\ 1, & gdy \qquad x > b. \end{cases}$$

jest dystrybuantą pewnej zmiennej losowej?
Obliczyć $P(\{\omega \in \Omega\colon\ -1 < \mathbf{X}(\omega) < 1\})$.

3. *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

x_i	1	2	4	$4,5$
p_i	$0,2$	$0,4$	$0,3$	α

Znaleźć α *i narysować dystrybuantę tej zmiennej.*

4. *Wiadomo, że zmienna losowa* $\mathbf{X}$ *ma gęstość*

$$f(x) = \begin{cases} 2-2x, & gdy \quad x \in (0,1) \\ 0, & gdy \quad -x \notin (0,1) \end{cases}$$

Znaleźć jej dystrybuantę. Obliczyć $P(\{\omega \in \Omega\colon\ -5 < \mathbf{X}(\omega) < 0,75\})$.

Zestaw 11

1. *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

x_i	-2	-1	0	2	3
p_i	$0,1$	$0,2$	$0,3$	$0,2$	$0,2$

Znaleźć rozkład zmiennej losowej $\mathbf{Y} = \mathbf{X}^2 - 1$. *Narysować wykres jej dystrybuanty.*

2. *Wiadomo, że zmienna losowa ma rozkład jednostajny na odcinku* $[a,b]$. *Wyznaczyć rozkład zmiennej losowej* $\mathbf{Y} = \frac{\mathbf{X}-a}{b-a}$ *podając jej dystrybuantę i gęstość.*

3. *Obliczyć w przybliżeniu prawdopodobieństwo, że partia 200 elemntów zawiera co najmniej 1 elemnt wadliwy, jeśli wiadomo, że prawdopodobieństwo wytworzenia wadliwego elemntu wynosi* $p = 0,01$.

4. *Zmienna losowa* $\mathbf{X}$ *ma gęstość*

$$f(x) = \begin{cases} \alpha e^x, & gdy \quad x \in [0, ln3] \\ 0, & gdy \quad x \notin [0, ln3]. \end{cases}$$

Znaleźć wartość α *i dystrybuantę tej zmiennej losowej.*

Zestaw 12

1. *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

x_i	-3	-2	0	1	2
p_i	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{3}{12}$

Obliczyć wartość oczekiwaną zmiennej losowej $\mathbf{Y} = 2\mathbf{X}$.

2. *Wiadomo, że zmienna losowa* $\mathbf{X}$ *ma gęstość* f. *Znaleźć gęstość zmiennej losowej* $\mathbf{Y} = \mathbf{X} + 1$.

3. *Z partii 100 przedmiotów, wśród których jest 10 wykonanych wadliwie, wybrano losowo bez zwracania 5 sztuk. Niech* $\mathbf{X}$ *oznacza liczbę sztuk wadliwych w próbie. Znaleźć rozkład zmiennej losowej* $\mathbf{X}$.

4. *Dane są trzy zdarzenia:* $A, B, C \in \Sigma$. *Wyrazić* $P(A \cup B \cup C)$ *za pomocą prawdopodobieństw zdarzeń:* $A, B, C, A \cap B, B \cap C, A \cap C$ *i* $A \cap B \cap C$.

Zestaw 13

1. *Dobrać tak stałe* $a, b, c, d \in \mathbf{R}$, *aby funkcja*

$$f(x) = \begin{cases} a, & gdy \quad x \le 0 \\ bx^2, & gdy \quad 0 < x \le 1 \\ \frac{x^2}{2} - x + c, & gdy \quad x > 2. \end{cases}$$

była dystrybuantą pewnej zmiennej losowej.

2. *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

$$f(x) = \begin{cases} \frac{1}{4}, & gdy \quad x \in [1,5] \\ 0, & gdy \quad -x \notin [1,5]. \end{cases}$$

Dokonać standaryzacji tej zmiennej. Na tej podstawie obliczyć jej wartość oczekiwaną i wariancję.

3. *Niech* $\mathbf{X} \in \mathcal{N}(-2{,}5,\ 2)$. *Obliczyć* $P(\{\omega \in \Omega\colon\ \mathbf{X}^2(\omega) < 9\})$.

4. *Wiadomo, że* $\mathbf{X}$ *ma rozkład jednostajny na odcinku* $[2,6]$. *Metodą nierówności Markowa i Czebyszewa oszacować* $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > 4{,}5\})$.

Zestaw 14

1. *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

$$F(x) = \begin{cases} 0, & gdy \quad x \le -1 \\ \frac{x+1}{4}, & gdy \quad -1 < x \le 3 \\ 1, & gdy \quad x > 3. \end{cases}$$

Znaleźć rozkład zmiennej losowej $|\mathbf{X}|$.

2. *Błąd w pewnej próbie można wykryć w* $99{,}8\%$ *przypadków. Oszacować prawdopodobieństwo, że w 500 próbach nie wykryto błędu w co najmniej 5 przypadkach.*

3. *Dane są zmienne losowe* $\mathbf{X}_1, \ldots, \mathbf{X}_n$ *każda o rozkładzie wykładniczym z parametrem* $\lambda = 2$. *Bierzemy zmienną losową* $\mathbf{X} = \sum_{i=1}^{n} \mathbf{X}_i$. *Dla jakich* n *zachodzi nierówność*

$$P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > 10\}) < 0{,}5.$$

4. *Wykonano* n *niezależnych powtórzeń doświadczenia, które polega na tym, że zdarzenie* A *zachodzi z prawdopodobieństwem* $0{,}25$. *Niech* $\mathbf{X}_n$ *oznacza ilość zajść zdarzenia* A *w* n *powtórzeniach. Oszacować prawdopodobieństwo*

$$P\Big(\{\omega \in \Omega\colon\ |\frac{\mathbf{X}_n(\omega)}{n} - 0{,}25| \le 10^{-3}\}\Big).$$

Przyjąć, że $n = 10^3$,

Zestaw 15

1. *8 osób posadzono obok siebie na ławce. Opisać to zjawisko w terminach kombinatoryki. Na ile sposobów można rozmieścić te osoby, aby wybrane dwie z nich siedziały obok siebie.*

2. *Z talii 52 kart losujemy 6. Niech A będzie zdarzeniem, że wylosowano co najmniej jednego asa czarnego, B, że wylosowano 2 asy. Czy zdarzenia te są stochastycznie niezależne?*

3. *Dany jest rozkład zmiennej losowej* $\mathbf{X}$

x_i	-3	-1	0	1	2
p_i	?	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{3}{12}$

Narysować dystrybuantę tej zmiennej losowej.
Obliczyć $P(\{\omega \in \Omega\colon\ -\frac{5}{2} < \mathbf{X}(\omega) < \frac{3}{2}\})$.

4. *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

$$F(x) = \begin{cases} 0, & \text{gdy} \quad x \leq 0 \\ \frac{1}{4}x, & \text{gdy} \quad 0 < x \leq 4 \\ 1, & \text{gdy} \quad x > 4. \end{cases}$$

Znaleźć rozkład zmiennej $\mathbf{Y} = -3\mathbf{X} + 1$.

Zadania z egzaminów

Zestaw 16

1. *Dane są zdarzenia* $A, B, C \in \Sigma$. *Wyrazić* $P(A \cap B \cap C)$ *za pomocą prawdopodobieństw zdarzeń:* $A, B, C, A \cap B, B \cap C, A \cap C, A \cap B \cap C$.

2. *Zmienna losowa* $\mathbf{X}$ *ma rozkład*

x_i	-2	-1	0	1	2
p_i	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{2}{9}$	$\frac{1}{6}$	$\frac{1}{9}$

   Znaleźć rozkład zmiennej losowej $\mathbf{Y} = e^{\mathbf{X}}$. *Narysować dystrybuantę tego rozkładu.*

3. *Dla zmiennej losowej* $\mathbf{Y}$ *z zadania 2 obliczyć*

   (a)
$$P(\{\omega \in \Omega\colon\ 0 \leq \mathbf{Y}(\omega) \leq \frac{5}{2}\}),$$

   (b) $var(\mathbf{Y})$.

4. *Błąd w pewnej próbie można wykryć w* $99,86\%$ *przypadków. Oszacować prawdopodobieństwo, że w 600 przypadkach nie wykryto błędu w co najmniej 2 przypadkach.*

5. *Zmienna losowa* $\mathbf{X}$ *ma rozkład*
$$f(x) = \begin{cases} \frac{1}{4} & , \ gdy \quad x \in [1,5] \\ 0 & , \ gdy \quad x \notin [1,5]. \end{cases}$$

   Dokonać jej standaryzacji. Na tej podstwie obliczyć jej wariancję.

6. *Wiadomo, że* $\mathbf{X} \in \mathcal{N}(-2,5,\ 2)$. *Obliczyć* $P(\{\omega \in \Omega\colon\ 5 \leq \mathbf{X}(\omega) \leq 10\})$.

Zestaw 17

1. *Niech* A, B, C *będą zdarzeniami losowymi. Podać postać zdarzenia* D *o tej własności, że*

   (a) *zachodzi co najmniej jedno z tych zdarzeń,*

   (b) *zachodzą co najmniej dwa z nich,*

   (c) *zachodzi tylko zdarzenie* B,

   (d) *nie zachodzi żadne z nich.*

2. *Z talii 52 kart losujemy 4. Niech A oznacza wylosowanie 3 asów i króla, B wylosowanie 2 asów. Sprawdzić, czy zdarzenia te są stochastycznie niezależne.*

3. $\mathbf{X}$ *ma rozkład*

x_i	-3	-1	1	2	3
p_i	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Obliczyć $P(\{\omega \in \Omega\colon\ 0 \leq \mathbf{X}(\omega) \leq \frac{5}{2}\})$.

4. *Dla zmiennej losowej z zadania 3 obliczyć jej wariancję.*

5. *Zmienna losowa ma rozkład*

$$f(x) = \begin{cases} 4x^3 & ,\ gdy \quad x \in (0,1) \\ 0 & ,\ gdy \quad x \notin (0,1). \end{cases}$$

Obliczyć jej wariancję.

6. *Wiadomo, że* $\mathbf{X} \in \mathcal{N}(-3,\ 4)$. *Obliczyć* $P(\{\omega \in \Omega\colon\ |\mathbf{X}(\omega)| \leq \frac{5}{2}\})$.

Zestaw 18

1. *Z odcinka* $[0,1]$ *wybrano losowo i w sposób niezależny dwie liczby a i b. Niech* $A = \{(a,b) \in [0,1]^2\colon\ a^2 + b^2 > 1\}$, $B = \{(a,b) \in [0,1]^2\colon\ a < b\}$. *Czy zdarzenia te są stochastycznie niezależne?*

2. *Niech zmienna losowa* $\mathbf{X}$ *ma rozkład*

x_i	-3	-1	0	1	3
p_i	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{9}$

Znaleźć rozkład zmiennej losowej $ln|\mathbf{X}(\omega)+1|$. *Narysować jej dystrybuantę.*

3. *Zmienna losowa* $\mathbf{X}$ *ma dystrybuantę*

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

Czy $\mathbf{X}$ *jest typu ciągłego? Jeśli tak, to*

(a) *znaleźć gęstość tego rozkładu,*

(b) *obliczyć wariancję.*

4. *Wiadomo, że* $\mathbf{X} \in \mathcal{J}([3,7])$. *Oszacować prawdopodobieństwo* $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) > 5,5\})$.

5. *Niech* $\mathbf{X}$ *ma rozkład standardowy normalny. Porównać odczyt z tablic* $P(\{\omega \in \Omega\colon\ |\mathbf{X}(\omega)| \geq 3\})$ *z oszacowaniem tego prawdopodobieństwa metodą nierówności Czebyszewa.*

6. *Cecha* $\mathbf{X}$ *populacji generalnej ma rozkład* $\mathcal{N}(m, \sigma^2)$, *gdzie* $\sigma^2 = 4$. *Pobrano próbę prostą:* $2,9,\ \ 2,7,\ \ 2,8,\ \ 3,1$. *Dla* $\alpha = 0,05$ *skonstruować przedział losowy dla wartości oczekiwanej. Zweryfikować hipotezę* $H_o : m = 2,9$ *przeciwko hipotezie* $H_1 : m > 2,9$ *dla* $\alpha = 0,05$.

Zestaw 19

1. *Dane są zdarzenia* $A, B, C \in \Sigma$. *Wyrazić* $P(A \cap B \cap C)$ *za pomocą prawdopodobieństw zdarzeń:* $A, B, C, A \cap B, B \cap C, A \cap C, A \cup B \cup C$.

2. *Niech* $\mathbf{X}$ *ma rozkład*

x_i	-1	0	1	2	3
p_i	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{9}$

   Znaleźć rozkład zmiennej $-\mathbf{X}$. *Narysować jej dystrybuantę.*

3. *Obliczyć wariancję zmiennej losowej z zadania 2.*

4. *Dla zmiennej losowej* $\mathbf{X}$ *o rozkładzie jednostajnym na przedziale* $[2, 6]$ *oszacować prawdopodobieństwo* $P(\{\omega \in \Omega\colon\ \mathbf{X}(\omega) < 4\})$.

5. *Wiadomo, że* $\mathbf{X} \in \mathcal{N}(-2,\ 4)$. *Obliczyć* $P(\{\omega \in \Omega\colon\ |\mathbf{X}(\omega)| > 0,5\})$.

6. *Cecha* $\mathbf{X}$ *populacji generalnej ma rozkład normalny. Pobrano próbę prostą:* $2,6,\ \ 3,0,\ \ 3,1,\ \ 2,8,\ \ 2,7$. *Dla* $\alpha = 0,05$ *skonstruować przedział losowy dla wartości oczekiwanej. Zweryfikować hipotezę* $H_o : m = 3$ *przeciwko* $H_1 : m \neq 3$ *dla* $\alpha = 0,05$.

Rozdział 14

Odpowiedzi do zadań

Zad. 1.2.1
Wskazówka: wykonaj ilustrację graficzną w układzie kartezjańskim.

Zad. 1.2.2
Wskazówka: $A^C \cap C^C = (A \cup B)^C$.

Zad. 1.2.3

1. $A \cap B \cap C$;

2. $A \cap B \cup C$;

3. $A^C \cap B \cap C^C$;

4. $A \cap B^C \cap C^C$.

Zad. 1.2.4
$A \cap B = (A^C \cup B^C)^C$

Zad. 1.2.5
$A \times B = \{(a,1),(a,2),(b,1),(b,2),(c,1),(c,2)\}$
$B \times A = \{(1,a),(2,a),(1,b),(2,b),(1,c),(2,c)\}$

Zad. 1.2.7
$|A \cup B| = 12$, $|A^C \cup B^C| = 3$, $|A - B| = 4$

Zad. 1.2.8
$\mathcal{A} = \{\ \emptyset,\ \ \Omega,\ \ [-1,0],\ \ (0,1],\ \ [-1,-\frac{1}{2}] \cup (0,1],\ \ (-\frac{1}{2},0],\ \ (-\frac{1}{2},1],\ \ [-1,-\frac{1}{2}],$
$[-\frac{1}{2},0],\ \ \{-\frac{1}{2}\},\ \ [-1,-\frac{1}{2}) \cup (-\frac{1}{2},1],\ \ [-1,-\frac{1}{2}),\ \ [-\frac{1}{2},1]\}$

Zad. 1.2.9

1. $A \cup B = [0,1]$, $A \cap B = (\frac{1}{4}, \frac{1}{2})$, $A - B = [0, \frac{1}{4}]$, $B - A = [\frac{1}{2}, 1]$, $A^C = [\frac{1}{2}, 1]$, $B^C = [0, \frac{1}{4}]$
2. $\mathcal{A} = \{X, \ \emptyset, \ [0, \frac{1}{2}], \ [\frac{1}{2}, 1], \ (\frac{1}{4}, 1], \ [0, \frac{1}{4}], \ (\frac{1}{4}, \frac{1}{2}), \ [0, \frac{1}{4}] \cup [\frac{1}{2}, 1]\}$

Zad. 1.2.10
Wskazówka: liczba dzieli się przez 25, jeśli jej dwie ostatnie cyfry to 25, 50 lub 75 (przypadek 00 odrzucamy ze względu na to, że losujemy bez zwracania).

Zad. 1.2.11
$E = A \cap B$
$F = A^C \cap B^C \cap C^C \cap D^C = (A \cup B \cup C \cup D)^C$
$G = A \cap B^C \cap C^C \cap D^C$
$H = (A \cap B^C \cap C^C \cap D^C) \cup (A^C \cap B \cap C^C \cap D^C) \cup (A^C \cap B^C \cap C \cap D^C) \cup (A^C \cap B^C \cap C^C \cap D)$

Zad. 1.2.12
Wskazówka: skorzystać z praw rachunku zbiorów.

Zad. 1.2.13
$\Omega = \{(o,o,o), (o,o,r), (o,r,o), (r,o,o), (o,r,r), (r,o,r), (r,r,o), (r,r,r)\}$
$A_1 = \{(r,o,o), (r,o,r), (r,r,o), (r,r,r)\}$
$A_2 = \{(o,r,o), (o,r,r), (r,r,o), (r,r,r)\}$
$A_3 = \{(o,o,r), (o,r,r), (r,o,r), (r,r,r)\}$

Zad. 1.2.14
$\Omega = \{(o,r), (r,o), (o,o,r), (o,o,o), (r,r,o), (r,r,r)\}$

Zad. 1.2.15
$\Omega = \{(\omega_1, \omega_2, \omega_3, \omega_4) : \omega_1, \omega_2, \omega_3, \omega_4 \in \{1, 2, ..., 52\}\}$
$A \cup B = A, \quad A \cup C = A, \quad A \cap B = B, \quad A \cup B \cup C = A$
$A \cap C^C$–zdarzenie polegające na tym, że wylosowaliśmy karty tak, że nie mamy dwóch asów
$(A^C \cap B^C) = (A \cup B)^C = A^C$

Zad. 1.2.16
$50\%, 10\%$

Zad. 1.2.17
200

Zad. 1.2.18
Nie.

Zad. 1.2.19
$A \cap B^C$–studentka nieuczęszczająca na lektorat z języka angielskiego
$A \cap B \cap C^C$–studentka uczęszczająca na lektorat z języka angielskiego i niemieszkająca w Legnicy
$A \cap B = A$–zachodzi, gdy wszystkie studentki uczestniczą w lektoracie z języka angielskiego
$A^C = B$–żadna studentka nie uczestniczy w lektoracie z języka angielskiego

Zad. 1.2.20
$\Omega = \{(\omega_1, \omega_2) : 0 \leq \omega_1 \leq 2 \wedge 0 \leq \omega_2 \leq 2\}$
$A = \{(\omega_1, \omega_2) \in \Omega : |\omega_1 - \omega_2| < 1\}$

Zad. 1.2.21
$\Omega = \{(x, y) : 0 \leq x \leq k \wedge 0 \leq y \leq k\}$
$A = \{(x, y) \in \Omega : x^2 + y^2 \geq \frac{k^2}{4}\}$
$B = \{(x, y) \in \Omega : x^2 + y^2 - k > 0\}$
$|\Omega| = k^2, \quad |A| = k^2 - \frac{\pi k^2}{16}, \quad |B| = k^2 - \frac{\pi k}{4}$

Zad. 2.2.1
14400; 28800; 3628800.

Zad. 2.2.2
720; 840; 12.

Zad. 2.2.3
$2 \cdot 50!$

Zad. 2.2.4
Nie.

Zad. 2.2.5
20

Zad. 2.2.6
$\frac{150!}{(150-42)!}$

Zad. 2.2.7
840; 4536.

Zad. 2.2.8
3^{15}; 4^{15}.

Zad. 2.2.9
9000

Zad. 2.2.10
3^n; $3^n - 2^n - 3$.

Zad. 2.2.12
120; 210; 120.

Zad. 2.2.13
220

Zad. 2.2.14
55

Zad. 2.2.15
28; 17.

Zad. 2.2.16
$\binom{52}{13} - \binom{48}{13}$; $\binom{13}{4}\binom{13}{4}\binom{13}{4}$; $4^4\binom{48}{9}$.

Zad. 2.2.17
$5!$; $5^5 - 5!$; 1200.

Zad. 2.2.18
3003; 12501; 1050.

Zad. 2.2.19
21; $\binom{5+n}{n}$.

Zad. 3.2.1
$0,35$

Zad. 3.2.2
$\frac{5}{36}$; $\frac{35}{36}$.

Zad. 3.2.3
$0,0009$

Zad. 3.2.4
$\frac{40!}{10! \cdot 24^{40}}$

Zad. 3.2.5
$\frac{11}{12}$, $\frac{1}{12}$

Zad. 3.2.6
$\frac{1}{8}$, $\frac{1}{2}$

Zad. 3.2.7
$0,7$

Zad. 3.2.8
$\frac{2}{9}$, $\frac{1}{2}$

Zad. 3.2.9

1. $0,018$;
2. $6,6 \cdot 10^{-5}$.

Zad. 3.2.10

1. $\frac{3}{11}$;
2. $\frac{3}{7}$.

Zad. 3.2.11
Wskazówka: wykorzystać łączność iloczynu zbiorów.

Zad. 3.2.12
$\frac{\pi}{4}$, $\frac{1}{4}$

Zad. 3.2.14
$0,016$, $\frac{5}{6}$

Zad. 3.2.15
$\frac{2(k-1)(n-k)}{n(n-1)}$

Zad. 3.2.16
$1 - \frac{1}{e^2}$; $\frac{34}{15e^2}$; $1 - \frac{5}{e^2}$.

Zad. 3.2.17
$0,25$

Zad. 3.2.18
$0,12$

Zad. 3.2.19
$\frac{3n(2n-1)}{(4n-1)(4n-3)}$

Zad. 3.2.20
5 i 15

Zad. 3.2.21
$\frac{3}{4}$

Zad. 3.2.22

1. nie są niezależne;
2. $\frac{\pi k(k-4)}{4k-\pi}$.

Zad. 4.2.1

1. $X : \{1, 2, 3, 4, 5, 6\} \to R$;
2. $F(x) = \begin{cases} 1 & \omega = 2n \\ 0 & \omega = 2n-1 \end{cases}$;
3. $p(0) = \frac{1}{2}, \quad p(1) = \frac{1}{2}$.

Zad. 4.2.2

1. Wskazówka: zbiór A jest skończony i dystrybuanta jest przedziałami stała;
2. $X = \{0, 1, 2, 3, 4, 5\}$;
3. $p(0) = \frac{1}{32}, \quad p(1) = \frac{5}{32}, \quad p(2) = \frac{10}{32}, \quad p(3) = \frac{10}{32}, \quad p(4) = \frac{5}{32}, \quad p(5) = \frac{1}{32}$;
4. $\frac{26}{32}$.

Zad. 4.2.3

1. $F(b) - F(a) = \sum\limits_{a \le x < b} P(x)$;
2. $F(b) - F(a) - P(a) + P(b) = \sum\limits_{a < x \le b} P(x)$;
3. $F(b) - F(a) + P(b) = \sum\limits_{a \le x \le b} P(x)$.

Zad. 4.2.5
$p(1,25) = 0,25, \quad p(8) = 0,5, \quad p(24) = 0,25$

Zad. 4.2.6

1. $d(j) = (1-r)^j r \quad r \in (0,1)$;
2. $(1-r)^7$.

Zad. 4.2.7
$x = \frac{3}{4}, \quad 0,25$

Zad. 4.2.10
$a > 1, \quad c = 0, \quad 0 \leq b \leq 1$

Zad. 4.2.11
$\frac{13}{18}$

Zad. 4.2.12
$\frac{1}{6}$

Zad. 4.2.13
Dla zmiennej z_1:
$p(0) = \frac{5}{20}, \quad p(1) = \frac{3}{20}, \quad p(1,5) = \frac{4}{20}, \quad p(2) = \frac{5}{20}, \quad p(3) = \frac{1}{20}, \quad p(4) = \frac{1}{20}, \quad p(5) = \frac{1}{20}, \quad p(5,5) = 0$
Dla zmiennej z_2:
$p(0) = 0, \quad p(1) = \frac{1}{20}, \quad p(1,5) = \frac{1}{20}, \quad p(2) = \frac{4}{20}, \quad p(3) = \frac{3}{20}, \quad p(4) = \frac{3}{20}, \quad p(5) = \frac{4}{20}, \quad p(5,5) = \frac{4}{20}$

Zad. 4.2.16

1. $p(-1) = \frac{1}{4}, \quad p(1) = \frac{1}{4}, \quad p(3) = \frac{1}{3}, p(4) = \frac{1}{6}$;
2. 1.

Zad. 4.2.19
0

Zad. 4.2.20
$p(1) = 0,1, \quad p(2) = 0,3, \quad p(3) = 0,4, \quad p(4) = 0,2; \quad 0,6$

Zad. 5.2.1
$0; \quad 1; \quad 0$

Zad. 5.2.2
$0 < a \leq ln(2b+1,5), \quad 0 \leq b \leq \frac{1}{8}$

Zad. 5.2.3
$c = 0, \quad 0 \leq b \leq 1 \quad c = 1, \quad a > 1, \quad \frac{1}{a} + \frac{1}{b} = 1$

Zad. 5.2.4

$$F(y) = \begin{cases} 0 & y \leq 1 \\ \frac{1}{2}y & 1 < y \leq 2 \\ 1 & y > 2 \end{cases}$$

Zad. 5.2.5

$$a = 0, \quad b = \tfrac{1}{2}, \quad c = 0 \quad F(x) = \begin{cases} 0 & x \leq -1 \\ \frac{1}{2}x + \frac{1}{2} & -1 < x \leq 1 \\ 1 & x > 1 \end{cases}$$

Zad. 5.2.6

$\frac{3}{4}$

Zad. 5.2.7

$\alpha = \frac{1}{2}$

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{2}(1 - cos x) & 0 < x \leq \pi \\ 1 & x > \pi \end{cases}$$

Zad. 5.2.8

$a = 1, \quad b = -\frac{1}{\pi}$

$$F(x) = \begin{cases} \frac{1}{\pi\sqrt{1-x^2}} & |x| < 1 \\ 0 & \text{poza tym} \end{cases}$$

Zad. 5.2.9

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

Zad. 5.2.10

$\alpha = \frac{1}{12}$

$\frac{1}{2}, \quad \frac{3}{8}$

Zad. 5.2.11

1. $G(y) = \begin{cases} 0 & y \leq \frac{1}{2} \\ 2 - \frac{1}{y} & \frac{1}{2} < y \leq 1 \\ 1 & y > 1 \end{cases}$

2. $G(y) = \begin{cases} 0 & y \leq 0 \\ e^y - 1 & 0 < y \leq ln2 \\ 1 & y > ln2 \end{cases}$

3. $G(y) = \begin{cases} 0 & y \leq 0 \\ 2\sqrt{y} & 0 < y \leq \frac{1}{4} \\ 1 & y > \frac{1}{4} \end{cases}$

Zad. 5.2.12

1. $c = 6$
2. $F(x) = \begin{cases} 0 & x \leq 0 \\ 3x^2 - 2x^3 & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}$
3. $\frac{5}{12}$

Zad. 5.2.13

1. Tak
2. $F(x) = \begin{cases} \frac{\pi}{2} \cdot sin\pi x & x \in (0,1) \\ 0 & \text{poza tym} \end{cases}$
3. $\frac{2-\sqrt{2}}{2}$

Zad. 5.2.14

$f(y) = \begin{cases} 0 & y < 0 \\ \frac{\lambda}{r} e^{-\frac{\lambda}{r} y} & y \geq 0 \end{cases}$

Zad. 5.2.15

$G(y) = \begin{cases} 0 & y \leq -\frac{1}{4} \\ \frac{4}{3}y + \frac{1}{3} & -\frac{1}{4} < y \leq \frac{1}{4} \\ \frac{2}{3}y + \frac{1}{2} & \frac{1}{4} < y \leq \frac{3}{4} \\ 1 & y > \frac{3}{4} \end{cases}$

$f(y) = \begin{cases} \frac{4}{3} & y \in (-\frac{1}{4}, \frac{1}{4}] \\ \frac{2}{3} & y \in (\frac{1}{4}, \frac{3}{4}] \\ 0 & \text{dla pozostałych} \end{cases}$

Zad. 5.2.16

1. $f(y) = \begin{cases} 3(1 - \sqrt{0,5y})^2 & y \in (0,2) \\ 0 & \text{dla pozostałych} \end{cases}$
2. $f(y) = \begin{cases} \frac{12}{y} lny(1 - lny)^2 & y \in (0,e) \\ 0 & \text{dla pozostałych} \end{cases}$

Zad. 5.2.17

$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{3}x & 0 < x \leq 1 \\ \frac{2}{3}x & 1 < x \leq 2 \\ 1 & x > 2 \end{cases}$

Zad. 5.2.18

1. $F(t) = \begin{cases} \frac{1}{2}e^t & t \leq 0 \\ 1 - \frac{1}{2}e^{-t} & t > 0 \end{cases}$

2. $1 - e^{-3}$

Zad. 5.2.19
$a = 2$
$$f(x) = \begin{cases} \frac{2}{x^2} & x \in (1, 2] \\ 0 & \text{dla pozostałych} \end{cases}$$

Zad. 5.2.20
$$G(y) = \begin{cases} e^y(1-y) & y \leq 0 \\ 1 & y > 0 \end{cases}$$

Zad. 6.2.5
Wskazówka: skorzystać ze wzoru dwumianowanego Newtona.

Zad. 6.2.7
$P(x = k) = p^{k-1}q$

Zad. 6.2.8
$-\frac{5}{6}$

Zad. 6.2.9
$\frac{9}{4}$; $\frac{751}{96}$; $\frac{265}{96}$

Zad. 6.2.10
$\frac{23}{3}$; $\frac{620}{81}$

Zad. 6.2.11
Nie.

Zad. 6.2.12
$\frac{12}{175}$

Zad. 6.2.16
p; $\frac{pq}{n}$

Zad. 6.2.17
$\frac{a+b}{2}$; $\frac{(b-a)^2}{12}$

Zad. 6.2.18
Nie.

Zad. 6.2.20
$\frac{ac}{1-a}$; $ac^2 \cdot \frac{1-4a+2a^2}{(2-a)(1-a)^2}$

Zad. 7.2.1
$\binom{n}{k}\frac{p^k q^{n-k}}{n}$

Zad. 7.2.2
$\frac{6}{e^2}$

Zad. 7.2.5
Tak.

Zad. 7.2.6
$f(x) = \frac{1}{\lambda(1-x)}$

Zad. 7.2.9
$$f(y) = \begin{cases} \lambda e^{\lambda y} & y \in [0, \frac{1}{\lambda} ln2] \\ 0 & \text{poza tym} \end{cases}$$

Zad. 7.2.10
$$f(y) = \begin{cases} \frac{1}{2(b-a)\sqrt{y}} & y \in [a^2, b^2] \\ 0 & \text{poza tym} \end{cases}$$

Zad. 7.2.11
$x_{\frac{1}{2}} = \frac{1}{\lambda} ln2$

Zad. 7.2.12
$\frac{35}{6}; \quad \frac{89}{36}$

Zad. 7.2.13

1. $0,2807$
2. $0,9081$
3. $0,5957$

Zad. 7.2.14
$\frac{y-\sigma}{m}; \quad EY = m^2 + \sigma, \quad Var(Y) = m^2\sigma^2$

Zad. 7.2.15
$x_{0,9} = 1,29, \quad x_{0,95} = 1,65$
$x_{0,9} = 2,08, \quad x_{0,95} = 2,8$

Zad. 7.2.16
$21,1608$

Zad. 7.2.17
$2,6810$

Zad. 7.2.18
$0,6652$

Zad. 8.2.1
$\leq \frac{p}{aq}, \quad \leq \frac{p}{a^2q^4}$

Zad. 8.2.2
$\leq \frac{8}{9}, \quad \leq \frac{4}{3}$

Zad. 8.2.3
$n \geq 3750000$

Zad. 8.2.4
$n < 50$

Zad. 8.2.5
$n \geq 2778$

Zad. 8.2.6
$< \frac{5}{19}$

Zad. 8.2.7
$0,0038$

Zad. 8.2.8
$\leq \frac{1}{48}$

Zad. 8.2.9
$\leq \frac{15}{16}$

Zad. 8.2.10
$\leq \frac{1}{9}$

Zad. 8.2.11
$\leq \frac{5}{24}$

Zad. 8.2.12
$\leq 0,475$

Zad. 8.2.13
$\geq 0,98$

Zad. 8.2.14
$\geq 0,996$

Zad. 8.2.15
$0,925$

Zad. 9.2.1
$g_x(x) = \frac{1}{5}(x^3 + x^4 + x^6 + x^7 + x^9)$

Zad. 9.2.4
$(q + px)^n$

Zad. 9.2.5
$np, \quad np(np + q)$

Zad. 9.2.6
$g_x(x) = (q + px)^n$

Zad. 9.2.7
$4,5; \quad 24$

Zad. 9.2.8

1. $g_x(x) = e^{-\lambda + \lambda x}$
2. $\lambda, \quad \lambda^2$

Zad. 9.2.9

1. $g_x(x) = \left(\frac{p}{1-qx}\right)^n$
2. $\frac{nq}{p}, \quad (n^2 + n)\frac{q^2}{p^2}$

Zad. 9.2.10
$\sum\limits_{i=1}^{n} ip_{ki}(-1)^{i-1}$
$\sum\limits_{i=2}^{n}(i^2 - i)p_{ki}(-1)^{i-2}$

Zad. 10.2.1
$0,9672$

Zad. 10.2.2
$0,342$

Zad. 10.2.3
$0,5346$

Zad. 10.2.4
$0,355$

Zad. 10.2.5
0

Zad. 10.2.6
$0,9993$

Zad. 10.2.7
$0,92$

Zad. 10.2.8
$0,5$

Zad. 10.2.9
$n \geq 5$

Zad. 10.2.10
Tak.

Zad. 11.2.2
$-0,17; 0,44; 0,5$

Zad. 11.2.3
$N(1,425; \quad 0,14)$
$F(x) = \frac{x-1,425}{0,14}$

Zad. 11.2.4
$(-0,77; \quad 2,31)$

Zad. 11.2.5
$(1,274; \quad 4,753)$

Zad. 11.2.6
$(2,752; \quad 3,248)$

Zad. 11.2.7
$(0,042; \quad 0,536)$

Zad. 12.2.1

1. nie ma powodu do odrzucenia hipotezy H_0;
2. nie ma powodu do odrzucenia hipotezy H_0;
3. nie ma powodu do odrzucenia hipotezy H_0.

Zad. 12.2.2

1. nie ma powodu do odrzucenia hipotezy H_0;
2. nie ma powodu do odrzucenia hipotezy H_0;
3. nie ma powodu do odrzucenia hipotezy H_0.

Zad. 12.2.3
Nie ma powodu do odrzucenia hipotezy H_0.

Zad. 12.2.4
Należy odrzucić hipotezę H_0 o tym, że $m = 1,9$.

Zad. 12.2.5
Należy odrzucić hipotezę H_0.

Zad. 12.2.6
Nie ma postaw do odrzucenia hipotezy H_0.

Zad. 12.2.7
Tak.

Zad. 12.2.8
Należy odrzucić hipotezę H_0.

Zad. 12.2.9
Nie ma postaw do odrzucenia hipotezy H_0.

Zad. 12.2.10
Nie ma postaw do odrzucenia hipotezy H_0.

Tablice statystyczne

Tablica 1. Dystrybuanta standardowego rozkładu normalnego $\mathcal{N}(0,\ 1)$

u	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,500000	0,503989	0,507978	0,511967	0,515953	0,519939	0,523922	0,527903	0,531881	0,535856
0,1	0,539828	0,543795	0,547758	0,551717	0,555670	0,559618	0,563559	0,567495	0,571424	0,575345
0,2	0,579260	0,583166	0,587064	0,590954	0,594835	0,598706	0,602568	0,606420	0,610261	0,614092
0,3	0,617911	0,621719	0,625516	0,629300	0,633072	0,636831	0,640576	0,644309	0,648027	0,651732
0,4	0,655422	0,659097	0,662757	0,666402	0,670031	0,673645	0,677242	0,680822	0,684386	0,687933
0,5	0,691462	0,694974	0,698468	0,701944	0,705402	0,708840	0,712260	0,715661	0,719043	0,722405
0,6	0,725747	0,729069	0,732371	0,735653	0,738914	0,742154	0,745373	0,748571	0,751748	0,754903
0,7	0,758036	0,761148	0,764238	0,767305	0,770350	0,773373	0,776373	0,779350	0,782305	0,785236
0,8	0,788145	0,791030	0,793892	0,796731	0,799546	0,802338	0,805106	0,807850	0,810570	0,813267
0,9	0,815940	0,818589	0,821214	0,823814	0,826391	0,828944	0,831472	0,833977	0,836457	0,838913
1	0,841345	0,843752	0,846136	0,848495	0,850830	0,853141	0,855428	0,857690	0,859929	0,862143
1,1	0,864334	0,866500	0,868643	0,870762	0,872857	0,874928	0,876976	0,878999	0,881000	0,882977
1,2	0,884930	0,886860	0,888767	0,890651	0,892512	0,894350	0,896165	0,897958	0,899727	0,901475
1,3	0,903199	0,904902	0,906582	0,908241	0,909877	0,911492	0,913085	0,914656	0,916207	0,917736
1,4	0,919243	0,920730	0,922196	0,923641	0,925066	0,926471	0,927855	0,929219	0,930563	0,931888
1,5	0,933193	0,934478	0,935744	0,936992	0,938220	0,939429	0,940620	0,941792	0,942947	0,944083
1,6	0,945201	0,946301	0,947384	0,948449	0,949497	0,950529	0,951543	0,952540	0,953521	0,954486
1,7	0,955435	0,956367	0,957284	0,958185	0,959071	0,959941	0,960796	0,961636	0,962462	0,963273
1,8	0,964070	0,964852	0,965621	0,966375	0,967116	0,967843	0,968557	0,969258	0,969946	0,970621
1,9	0,971284	0,971933	0,972571	0,973197	0,973810	0,974412	0,975002	0,975581	0,976148	0,976705
2	0,977250	0,977784	0,978308	0,978822	0,979325	0,979818	0,980301	0,980774	0,981237	0,981691
2,1	0,982136	0,982571	0,982997	0,983414	0,983823	0,984222	0,984614	0,984997	0,985371	0,985738
2,2	0,986097	0,986447	0,986791	0,987126	0,987455	0,987776	0,988089	0,988396	0,988696	0,988989
2,3	0,989276	0,989556	0,989830	0,990097	0,990358	0,990613	0,990863	0,991106	0,991344	0,991576
2,4	0,991802	0,992024	0,992240	0,992451	0,992656	0,992857	0,993053	0,993244	0,993431	0,993613
2,5	0,993790	0,993963	0,994132	0,994297	0,994457	0,994614	0,994766	0,994915	0,995060	0,995201
2,6	0,995339	0,995473	0,995603	0,995731	0,995855	0,995975	0,996093	0,996207	0,996319	0,996427
2,7	0,996533	0,996636	0,996736	0,996833	0,996928	0,997020	0,997110	0,997197	0,997282	0,997365
2,8	0,997445	0,997523	0,997599	0,997673	0,997744	0,997814	0,997882	0,997948	0,998012	0,998074
2,9	0,998134	0,998193	0,998250	0,998305	0,998359	0,998411	0,998462	0,998511	0,998559	0,998605
3	0,998650	0,998694	0,998736	0,998777	0,998817	0,998856	0,998893	0,998930	0,998965	0,998999
3,1	0,999032	0,999064	0,999096	0,999126	0,999155	0,999184	0,999211	0,999238	0,999264	0,999289
3,2	0,999313	0,999336	0,999359	0,999381	0,999402	0,999423	0,999443	0,999462	0,999481	0,999499
3,3	0,999517	0,999533	0,999550	0,999566	0,999581	0,999596	0,999610	0,999624	0,999638	0,999650
3,4	0,999663	0,999675	0,999687	0,999698	0,999709	0,999720	0,999730	0,999740	0,999749	0,999758
3,5	0,999767	0,999776	0,999784	0,999792	0,999800	0,999807	0,999815	0,999821	0,999828	0,999835
3,6	0,999841	0,999847	0,999853	0,999858	0,999864	0,999869	0,999874	0,999879	0,999883	0,999888

Tablica 2. Wartości krytyczne rozkładu t–Studenta $P(\mid t_n \mid > t_\alpha) = \alpha$

	0,9	**0,8**	**0,7**	**0,6**	**0,5**	**0,4**	**0,3**	**0,2**	**0,1**	**0,05**	**0,04**	**0,02**	**0,01**
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	15,894	31,821	63,656
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	4,849	6,965	9,925
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	3,482	4,541	5,841
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	2,999	3,747	4,604
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	2,757	3,365	4,032
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	2,612	3,143	3,707
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,517	2,998	3,499
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,449	2,896	3,355
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,398	2,821	3,250
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,359	2,764	3,169
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,328	2,718	3,106
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,303	2,681	3,055
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,282	2,650	3,012
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,264	2,624	2,977
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,249	2,602	2,947
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,235	2,583	2,921
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,224	2,567	2,898
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,214	2,552	2,878
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,205	2,539	2,861
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,197	2,528	2,845
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,189	2,518	2,831
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,183	2,508	2,819
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,177	2,500	2,807
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,172	2,492	2,797
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,167	2,485	2,787
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,162	2,479	2,779
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,158	2,473	2,771
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,154	2,467	2,763
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,150	2,462	2,756
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,147	2,457	2,750

Tablica 3. Wartości krytyczne rozkładu chi–kwadrat $P(\chi_n^2 > \chi_\alpha^2) = \alpha$

α / l.ss	0,99	0,975	0,95	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,025	0,01	0,001
1	0,000	0,001	0,004	0,016	0,064	0,148	0,275	0,455	0,708	1,074	1,642	2,706	3,841	5,024	6,635	5,024
2	0,020	0,051	0,103	0,211	0,446	0,713	1,022	1,386	1,833	2,408	3,219	4,605	5,991	7,378	9,210	7,378
3	0,115	0,216	0,352	0,584	1,005	1,424	1,869	2,366	2,946	3,665	4,642	6,251	7,815	9,348	11,345	9,348
4	0,297	0,484	0,711	1,064	1,649	2,195	2,753	3,357	4,045	4,878	5,989	7,779	9,488	11,143	13,277	11,143
5	0,554	0,831	1,145	1,610	2,343	3,000	3,656	4,351	5,132	6,064	7,289	9,236	11,070	12,832	15,086	12,832
6	0,872	1,237	1,635	2,204	3,070	3,828	4,570	5,348	6,211	7,231	8,558	10,645	12,592	14,449	16,812	14,449
7	1,239	1,690	2,167	2,833	3,822	4,671	5,493	6,346	7,283	8,383	9,803	12,017	14,067	16,013	18,475	16,013
8	1,647	2,180	2,733	3,490	4,594	5,527	6,423	7,344	8,351	9,524	11,030	13,362	15,507	17,535	20,090	17,535
9	2,088	2,700	3,325	4,168	5,380	6,393	7,357	8,343	9,414	10,656	12,242	14,684	16,919	19,023	21,666	19,023
10	2,558	3,247	3,940	4,865	6,179	7,267	8,295	9,342	10,473	11,781	13,442	15,987	18,307	20,483	23,209	20,483
11	3,053	3,816	4,575	5,578	6,989	8,148	9,237	10,341	11,530	12,899	14,631	17,275	19,675	21,920	24,725	21,920
12	3,571	4,404	5,226	6,304	7,807	9,034	10,182	11,340	12,584	14,011	15,812	18,549	21,026	23,337	26,217	23,337
13	4,107	5,009	5,892	7,041	8,634	9,926	11,129	12,340	13,636	15,119	16,985	19,812	22,362	24,736	27,688	24,736
14	4,660	5,629	6,571	7,790	9,467	10,821	12,078	13,339	14,685	16,222	18,151	21,064	23,685	26,119	29,141	26,119
15	5,229	6,262	7,261	8,547	10,307	11,721	13,030	14,339	15,733	17,322	19,311	22,307	24,996	27,488	30,578	27,488
16	5,812	6,908	7,962	9,312	11,152	12,624	13,983	15,338	16,780	18,418	20,465	23,542	26,296	28,845	32,000	28,845
17	6,408	7,564	8,672	10,085	12,002	13,531	14,937	16,338	17,824	19,511	21,615	24,769	27,587	30,191	33,409	30,191
18	7,015	8,231	9,390	10,865	12,857	14,440	15,893	17,338	18,868	20,601	22,760	25,989	28,869	31,526	34,805	31,526
19	7,633	8,907	10,117	11,651	13,716	15,352	16,850	18,338	19,910	21,689	23,900	27,204	30,144	32,852	36,191	32,852
20	8,260	9,591	10,851	12,443	14,578	16,266	17,809	19,337	20,951	22,775	25,038	28,412	31,410	34,170	37,566	34,170
21	8,897	10,283	11,591	13,240	15,445	17,182	18,768	20,337	21,992	23,858	26,171	29,615	32,671	35,479	38,932	35,479
22	9,542	10,982	12,338	14,041	16,314	18,101	19,729	21,337	23,031	24,939	27,301	30,813	33,924	36,781	40,289	36,781
23	10,196	11,689	13,091	14,848	17,187	19,021	20,690	22,337	24,069	26,018	28,429	32,007	35,172	38,076	41,638	38,076
24	10,856	12,401	13,848	15,659	18,062	19,943	21,652	23,337	25,106	27,096	29,553	33,196	36,415	39,364	42,980	39,364
25	11,524	13,120	14,611	16,473	18,940	20,867	22,616	24,337	26,143	28,172	30,675	34,382	37,652	40,646	44,314	40,646
26	12,198	13,844	15,379	17,292	19,820	21,792	23,579	25,336	27,179	29,246	31,795	35,563	38,885	41,923	45,642	41,923
27	12,878	14,573	16,151	18,114	20,703	22,719	24,544	26,336	28,214	30,319	32,912	36,741	40,113	43,195	46,963	43,195
28	13,565	15,308	16,928	18,939	21,588	23,647	25,509	27,336	29,249	31,391	34,027	37,916	41,337	44,461	48,278	44,461
29	14,256	16,047	17,708	19,768	22,475	24,577	26,475	28,336	30,283	32,461	35,139	39,087	42,557	45,722	49,588	45,722
30	14,953	16,791	18,493	20,599	23,364	25,508	27,442	29,336	31,316	33,530	36,250	40,256	43,773	46,979	50,892	46,979

Literatura

[1] W. Feller, Wstęp do rachunku prawdopodobieństwa. PWN, Warszawa 1969.

[2] M. Fisz, Rachunek Prawdopodobieństwa i Statystyka Matematyczna. PWN, Warszawa 1969.

[3] J. Greń, Statystyka Matematyczna. Modele i zadania. PWN, Warszawa 1976.

[4] H. Jasiulewicz, W. Kordecki, Rachunek prawdopodobieństwa i statystyka matematyczna. Przykłady i zadania. Oficyna Wydawnicza GiS, Wrocław 2002.

[5] W. Krysicki, J. Bartos, W. Dyczka, K. Królikowska, M. Wasilewski, Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Część I. Rachunek Prawdopodobieństwa. PWN, Warszawa 1986.

[6] W. Krysicki, J. Bartos, W. Dyczka, K. Królikowska, M. Wasilewski, Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Część II. Statystyka matematyczna. PWN, Warszawa 1986.

[7] Marek Wiktor, J. Onyszkiewicz, Elementy logiki i teorii mnogości w zadaniach. PWN, Warszawa 1972.

[8] A. Plucińska, E. Pluciński, Zadania z rachunku prawdopodobieństwa i statystyki matematycznej dla studentów politechnik. PWN, Warszawa 1982.

[9] R. Rębowski, Podstawy metod probabilistycznych. Seria Wydawnicza Państwowej Wyższej Szkoły Zawodowej im. Witelona w Legnicy, 2006.

[10] R. Rębowski, Matematyka dyskretna dla informatyków, w przygotowaniu do wydania.

[11] R. Zieliński, Rachunek prawdopodobieństwa z elementami statystyki matematycznej. Wydawnictwa szkolne i pedagogiczne. Warszawa 1976.

Zbiór zadań z metod probabilistycznych

Janina Płaskonka
i Ryszard Rębowski

seria wydawnicza
Państwowej Wyższej Szkoły Zawodowej im. Witelona w Legnicy