W. KRYSICKI
J. BARTOS
W. DYCZKA
K. KRÓLIKOWSKA
M. WASILEWSKI

RACHUNEK PRAWDOPODOBIEŃSTWA I STATYSTYKA MATEMATYCZNA W ZADANIACH

część I

RACHUNEK PRAWDOPODOBIEŃSTWA

WYDAWNICTWO NAUKOWE PWN

Niniejszy dwuczęściowy podręcznik rachunku prawdopodobieństwa i statystyki matematycznej przeznaczony jest przede wszystkim dla studentów uczelni technicznych; mogą z niego korzystać słuchacze innych kierunków studiów, jak również ci, którzy stosują metody probabilistyczne w swej pracy zawodowej. Poszczególne zagadnienia opracowane zostały teoretycznie z uwzględnieniem podstawowych definicji i twierdzeń, po czym podano szereg zadań rozwiązanych, o wzrastającym stopniu trudności oraz zadań do samodzielnego rozwiązywania wraz z odpowiedziami.

ISBN 83-01-05928-1

RACHUNEK PRAWDOPODOBIEŃSTWA I STATYSTYKA MATEMATYCZNA W ZADANIACH

CZĘŚĆ I. Rachunek prawdopodobieństwa

W. KRYSICKI, J. BARTOS, W. DYCZKA,
K. KRÓLIKOWSKA, M. WASILEWSKI

RACHUNEK PRAWDOPODOBIEŃSTWA I STATYSTYKA MATEMATYCZNA W ZADANIACH

CZĘŚĆ I. Rachunek prawdopodobieństwa

Wydanie VII

WARSZAWA 1999
WYDAWNICTWO NAUKOWE PWN

Książka opracowana na podstawie skryptu uczelnianego Politechniki Łódzkiej, pod tytułem *Elementy probabilistyki w zadaniach*, cz. I. *Rachunek prawdopodobieństwa*.

Okładkę projektował: ANDRZEJ PILICH

Redaktor: ROMANA EHRENFEUCHT

Wydawnictwo Naukowe PWN SA
ul. Miodowa 10, 00-251 Warszawa
Tel.: 69 54 321, e-mail: pwn@pwn.com.pl
www.pwn.com.pl

ISBN 83-01-05928-1

PRZEDMOWA

Książka niniejsza pt. *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach* składa się z dwóch części: *Rachunek prawdopodobieństwa* oraz *Statystyka matematyczna.* Została opracowana przez zespół pięciu autorów na wzór książki: W. Krysicki, L. Włodarski, *Analiza matematyczna w zadaniach.*

Każdy paragraf rozpoczyna się od przytoczenia teorii, podstawowych pojęć i twierdzeń, po czym następuje odpowiednia liczba zadań rozwiązanych całkowicie, o wzrastającej skali trudności, bądź też o różnych zastosowaniach przytoczonych uprzednio pojęć. Na zakończenie każdego rozdziału występuje pewna liczba zadań do samodzielnego rozwiązania, z odpowiedziami i szczegółowymi wskazówkami przy trudniejszych zadaniach.

Oprócz tradycyjnego materiału, który można znaleźć w książkach autorów polskich z probabilistyki, autorzy niniejszego opracowania zamieścili kilka nowych tematów: wyznaczanie obszaru ufności dla obu parametrów rozkładu normalnego, dwa nowe testy, w tym jeden wykorzystujący pełniejszą informację z próby, odnoszące się do rozkładu normalnego, nowsze podejście do tak znanych testów jak chi-kwadrat, Kołmogorowa i porównanie ich oraz łączny obszar ufności dla obu współczynników liniowego równania regresji.

Książka jest przeznaczona przede wszystkim dla słuchaczy studiów technicznych, choć mogą z niej korzystać również słuchacze innych kierunków oraz pracownicy licznych instytutów i wszyscy ci, którzy stosują metody probabilistyczne w zastosowaniach. Poza teorią wspólną dla wszystkich Czytelników, tematyka wielu zadań nie dotyczy wyłącznie techniki.

Autorzy

1 ZDARZENIA LOSOWE I PRAWDOPODOBIEŃSTWO

1.1. ZDARZENIA LOSOWE

1.1.1. Doświadczenia losowe. *Rachunek prawdopodobieństwa* (mówimy również probabilistyka od łacińskiego słowa *probabilitis* oznaczającego prawdopodobny) zajmuje się badaniem praw rządzących zdarzeniami losowymi. W jego aksjomatycznym ujęciu pojęciami pierwotnymi są: zdarzenia elementarne ω i przestrzeń zdarzeń elementarnych Ω związane z doświadczeniem losowym D.

Doświadczenie losowe to realizacja (rzeczywista bądź tylko myślowa) określonego zespołu warunków, wraz z góry określonym zbiorem wyników. Poszczególne wyniki ω doświadczenia losowego (krócej: doświadczenia) traktujemy jako *zdarzenia elementarne*, a zbiór wszystkich zdarzeń elementarnych jako *przestrzeń zdarzeń elementarnych*. Może ona być zbiorem skończonym, przeliczalnym albo nieprzeliczalnym (zadanie 1.1).

1.1.2. Przestrzeń zdarzeń elementarnych i σ-ciało zdarzeń. W zagadnieniach praktycznych najczęściej interesujące są nie pojedyncze zdarzenia elementarne rozpatrywanego doświadczenia D, lecz ich zbiory, czyli podzbiory przestrzeni Ω. Każdy taki podzbiór, gdy przestrzeń Ω jest skończona albo przeliczalna, nazywamy *zdarzeniem losowym* (związanym z rozpatrywanym doświadczeniem D).

Gdy przestrzeń Ω jest nieprzeliczalna, wtedy z różnych względów nie każdy jej podzbiór przyjmuje się jako zdarzenie losowe. Spośród wszystkich jej podzbiorów wyróżnia się pewną klasę [1] $\mathscr{Z}$ podzbiorów, zwaną σ-ciałem zdarzeń, i tylko elementy tej klasy nazywamy zdarzeniami losowymi.

Przeliczalnie addytywnym ciałem zdarzeń (lub *σ-ciałem zdarzeń*) przestrzeni zdarzeń elementarnych nazywamy niepustą klasę $\mathscr{Z}$ jej podzbiorów spełniającą następujące warunki (aksjomaty σ-ciała):

A1. *Cała przestrzeń zdarzeń elementarnych należy do tej klasy*:

$$\Omega \in \mathscr{Z}.$$

---

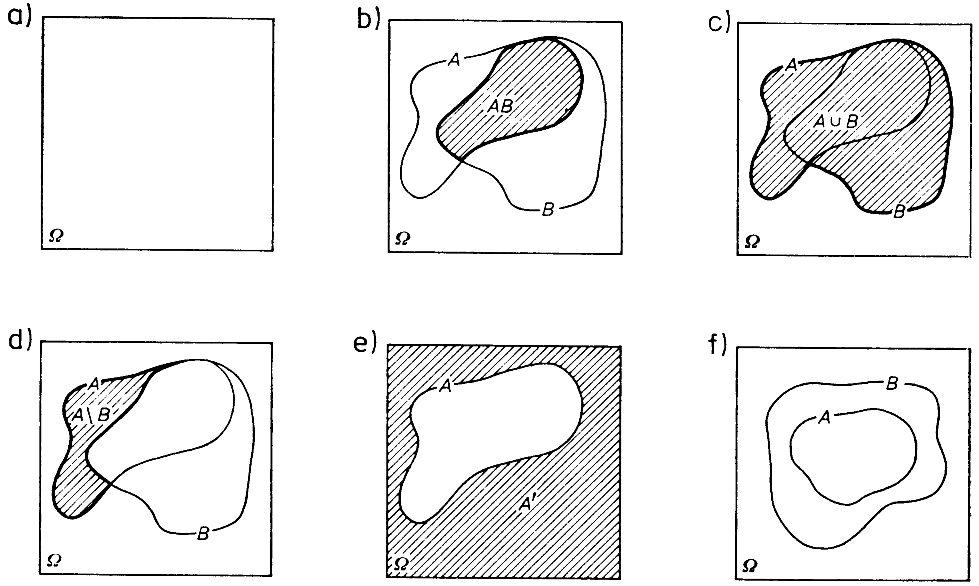
[1] *Klasą* (albo rodziną) nazywamy zbiór zbiorów, a więc zbiór, którego elementami są zbiory.

A2. *Dopełnienie A' (czyli zbiór $\Omega\backslash A$) dowolnego zbioru A należącego do klasy $\mathscr{Z}$ jest elementem tej klasy, czyli*

$$A \in \mathscr{Z} \Rightarrow A' \in \mathscr{Z}.$$

A3. *Suma co najwyżej przeliczalnej (czyli skończonej albo przeliczalnej) liczby zbiorów należących do klasy $\mathscr{Z}$ również należy do tej klasy:*

$$A_1 \in \mathscr{Z}, \ldots, A_n \in \mathscr{Z}, \ldots \Rightarrow (A_1 \cup \ldots \cup A_n \cup \ldots) \in \mathscr{Z}.$$



Rys. 1.1. Ilustracja graficzna działań na zdarzeniach

Każdy zbiór należący do przeliczalnie addytywnego ciała $\mathscr{Z}$ nazywamy *zdarzeniem losowym* (krócej: *zdarzeniem*). Zgodnie z aksjomatami A1 i A2 zdarzeniami losowymi są m. in.: cała przestrzeń zdarzeń elementarnych Ω i dopełnienie A' zdarzenia A. Pierwsze z nich, czyli Ω nazywamy *zdarzeniem pewnym* (rys. 1.1a), drugie, czyli A' — *zdarzeniem przeciwnym* do A (rys. 1.1e). Zdarzeniem jest również zbiór pusty $\emptyset$ — nazywamy go *zdarzeniem niemożliwym*.

1.1.3. Określenie działań na zdarzeniach w języku zbiorów. Na zdarzeniach wykonuje się analogiczne działania jak na zbiorach.

Koniunkcją (iloczynem) dwóch zdarzeń $A, B \in \mathscr{Z}$ nazywamy zdarzenie $A \cap B$ (oznaczone także krócej: AB), składające się z tych wszystkich zdarzeń elementarnych, które należą zarówno do zdarzenia A jak i do zdarzenia B (rys. 1.1b).

Analogicznie określamy koniunkcję co najwyżej przeliczalnej liczby zdarzeń. Niech mianowicie T oznacza dowolny skończony albo przeliczalny podzbiór zbioru liczb naturalnych. Przyporządkujmy każdej liczbie $t \in T$ zdarzenie $A_t \in \mathscr{Z}$. *Koniunkcją (iloczynem) zdarzeń* A_t, $t \in T$ (jest ich skończenie albo przeliczalnie wiele, stosownie do tego, czy zbiór T

jest skończony czy przeliczalny) nazywamy zdarzenie $\bigcap_{t \in T} A_t$, składające się ze wszystkich zdarzeń elementarnych ω, które należą do każdego ze zdarzeń A_t.

Alternatywą (*sumą*) *zdarzeń* $A, B \in \mathscr{Z}$ nazywamy zdarzenie $A \cup B$, składające się ze zdarzeń elementarnych ω, które należą co najmniej do jednego ze zdarzeń A, B (rys. 1.1c).

Alternatywą co najwyżej przeliczalnej liczby zdarzeń $A_t \in \mathscr{Z}$, $t \in T$, nazywamy zdarzenie $\bigcup_{t \in T} A_t$, składające się z tych wszystkich zdarzeń elementarnych ω, które należą co najmniej do jednego ze zdarzeń A_t, $t \in T$.

Różnicą $A \backslash B$ (oznaczaną również przez $A-B$) zdarzeń $A, B \in \mathscr{Z}$ nazywamy zdarzenie składające się z tych zdarzeń elementarnych ω, które należą do zdarzenia A, lecz nie należą do zdarzenia B (rys. 1.1d). W szczególności różnica $\Omega \backslash A$ jest zdarzeniem przeciwnym do zdarzenia A, czyli $\Omega \backslash A = A'$.

Mówimy, że zdarzenie A *pociąga* zdarzenie B (albo, że zdarzenie B jest *następstwem* zdarzenia A), co zapisujemy $A \subset B$, jeśli każde zdarzenie elementarne ω należące do zdarzenia A — również należy do zdarzenia B (rys. 1.1f).

Jeśli $A \subset B$ i $B \subset A$, to mówimy, że zdarzenie A i B są *równe* i zapisujemy $A = B$.

Mówimy, że zdarzenia $A, B \in \mathscr{Z}$ *wykluczają się* (albo: *wyłączają się*), jeśli nie mają wspólnych zdarzeń elementarnych, tj. gdy ich koniunkcja jest zdarzeniem niemożliwym: $AB = \emptyset$.

Mówimy, że zdarzenia $A_1, \ldots, A_n, \ldots$ *wykluczają się parami*, jeśli każde dwa spośród nich wykluczają się, czyli $A_i A_j = \emptyset$, gdy $i \neq j$.

Mówimy, że rodzina $\{A_t, t \in T\}$ zdarzeń $A_t \in \mathscr{Z}$ jest *podziałem* przestrzeni zdarzeń elementarnych Ω, jeżeli zdarzenia A_t spełniają następujące warunki: 1° zdarzenia A_t wykluczają się parami, czyli $A_{t_1} A_{t_2} = \emptyset$, gdy $t_1 \neq t_2$, 2° ich alternatywa jest zdarzeniem pewnym, czyli $\bigcup_{t \in T} A_t = \Omega$. Na przykład podziałem przestrzeni zdarzeń elementarnych jest rodzina $\{A, A'\}$ złożona z dowolnego zdarzenia A i zdarzenia do niego przeciwnego.

1.1.4. Określenie działań na zdarzeniach w języku potocznym. Określone zdarzenie A *zachodzi* (*realizuje się*, *pojawia się*) wtedy i tylko wtedy, gdy zachodzi dokładnie jedno ze zdarzeń elementarnych należących do tego zdarzenia. Stwierdzenie to umożliwia sformułowanie podanych tu określeń przy użyciu zwrotu „zdarzenie zachodzi". Określenia te ułatwiają tłumaczenie zagadnień spotykanych w praktyce na język zdarzeń.

Koniunkcja dowolnej liczby zdarzeń jest to zdarzenie polegające na zajściu wszystkich tych zdarzeń.

Alternatywa dowolnej liczby zdarzeń jest to zdarzenie polegające na zajściu co najmniej jednego z tych zdarzeń.

Różnica $A \backslash B$ zdarzeń A, B jest to zdarzenie polegające na zajściu zdarzenia A i niezajściu zdarzenia B.

Mówimy, że dwa zdarzenia są *równe*, jeśli dowolne z nich zachodzi wtedy i tylko wtedy, gdy zachodzi drugie.

Mówimy, że zdarzenie A *pociąga* zdarzenie B wtedy i tylko wtedy, gdy z zajścia zdarzenia A wynika zajście zdarzenia B.

Mówimy, że dwa zdarzenia *wykluczają się* wtedy i tylko wtedy, gdy zajście dowolnego z tych zdarzeń wyklucza zajście pozostałego zdarzenia.

1.1.5. Interpretacja działań na zdarzeniach. Przestrzeń Ω zdarzeń elementarnych często będziemy interpretowali jako prostokąt na płaszczyźnie (rys. 1.1a), punkty zaś tego prostokąta – wszystkie albo tylko zaznaczone – jako zdarzenia elementarne. Gdy wszystkie punkty prostokąta interpretujemy jako zdarzenia elementarne, wtedy zajście zdarzenia elementarnego możemy traktować jako rezultat losowo rzuconego punktu na ten prostokąt.

Niech A oznacza zdarzenie, że punkt znajdzie się w obszarze A, B – w obszarze B. Wówczas zakreskowane na rysunku 1.1 obszary oznaczają odpowiednio zdarzenie: AB, $A \cup B$, $A \backslash B$, A'; ostatni rysunek ilustruje następstwo zdarzeń: $A \subset B$.

1.1.6. Własności działań na zdarzeniach. Działania na zdarzeniach podlegają następującym prawom (analogicznym do praw rachunku zbiorów):

$$AB = BA \quad \text{— } \textit{przemienność koniunkcji zdarzeń},$$

$$A \cup B = B \cup A \quad \text{— } \textit{przemienność alternatywy zdarzeń},$$

$$A(BC) = (AB)C \quad \text{— } \textit{łączność koniunkcji zdarzeń},$$

$$A \cup (B \cup C) = (A \cup B) \cup C \quad \text{— } \textit{łączność alternatywy zdarzeń},$$

$$\left.\begin{array}{l} A(B \cup C) = AB \cup AC \\ A\left(\bigcup\limits_{t \in T} A_t\right) = \bigcup\limits_{t \in T} (AA_t) \end{array}\right\} \quad \text{— } \textit{rozdzielność koniunkcji zdarzeń względem alternatywy zdarzeń},$$

$$\left.\begin{array}{l} A \cup (BC) = (A \cup B)(A \cup C) \\ A \cup \left(\bigcap\limits_{t \in T} A_t\right) = \bigcap\limits_{t \in T} (A \cup A_t) \end{array}\right\} \quad \text{— } \textit{rozdzielność alternatywy zdarzeń względem koniunkcji zdarzeń},$$

$$\left.\begin{array}{l} (AB)' = A' \cup B' \\ \left(\bigcap\limits_{t \in T} A_t\right)' = \bigcup\limits_{t \in T} A_t' \\ (A \cup B)' = A'B' \\ \left(\bigcup\limits_{t \in T} A_t\right)' = \bigcap\limits_{t \in T} A_t' \end{array}\right\} \quad \text{— } \textit{prawa De Morgana}.$$

Prawa De Morgana wyrażają tzw. *zasadę dwoistości* w odniesieniu do zdarzeń: 1° zdarzenie przeciwne do koniunkcji zdarzeń jest alternatywą zdarzeń do nich przeciwnych i 2° zdarzenie przeciwne do alternatywy zdarzeń jest koniunkcją zdarzeń do nich przeciwnych.

1.1.7. Zadania rozwiązane.

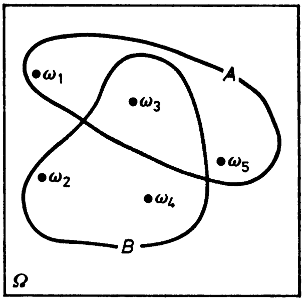
ZADANIE 1.1. Podać przykład doświadczenia losowego dającego opisać się za pomocą przestrzeni zdarzeń elementarnych Ω, która jest: a) skończona, b) przeliczalna, c) nieprzeliczalna.

Rozwiązanie. a) Zbiór $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, gdzie $\omega_i=i$, $i=1, \ldots, 6$, można uważać za przestrzeń zdarzeń elementarnych doświadczenia polegającego na rzucie kostką sześcienną i obserwacji liczby oczek wyrzuconych na górnej ściance. Jest to oczywiście przestrzeń skończona.

b) Niech D_i oznacza wylosowanie, za i-tym razem, z partii towaru sztuki dobrej, W_i — sztuki wadliwej, przy założeniu, że po każdym losowaniu dokonujemy zwrotu wylosowanej sztuki. Losowanie powtarzamy aż do uzyskania sztuki dobrej. Za zdarzenia elementarne ω_n przyjmujemy n-wyrazowe ciągi symboli W_i, D_i postaci $\omega_n=(W_1, W_2, \ldots, W_{n-1}, D_n)$. Ponieważ n może tu przyjmować dowolne wartości naturalne, $n \in N$, których jest przeliczalnie wiele, więc przestrzeń zdarzeń elementarnych $\Omega=\{\omega_1, \ldots, \omega_n, \ldots\}$ rozważanego doświadczenia jest przeliczalna.

c) Nić o długości l poddajemy naprężeniu, aż do jej zerwania, i obserwujemy odciętą w punkcie, w którym nić uległa zerwaniu (punkt zamocowania nici przyjmujemy za początek osi liczbowej o zwrocie zgodnym ze zwrotem rozciągania). Za przestrzeń zdarzeń elementarnych Ω przyjmujemy tu przedział liczbowy $(0, \Delta l+l)$, gdzie Δl jest możliwym maksymalnym wydłużeniem nitki. Przestrzeń ta jest nieprzeliczalna (jej elementów nie daje się ponumerować liczbami naturalnymi).

ZADANIE 1.2. Niech przestrzeń Ω zdarzeń elementarnych doświadczenia składa się z pięciu zdarzeń elementarnych ω_i: $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$. Określamy zdarzenia: $A=\{\omega_1, \omega_3, \omega_5\}$, $B=\{\omega_2, \omega_3, \omega_4\}$ (rys. 1.2). Znaleźć zdarzenia: $A \cup B$, AB, $A \backslash B$, $B \backslash A$.



Rys. 1.2. Do zadania 1.2

Rozwiązanie. Korzystamy kolejno z określenia alternatywy, koniunkcji, różnicy dwu zdarzeń i otrzymujemy:

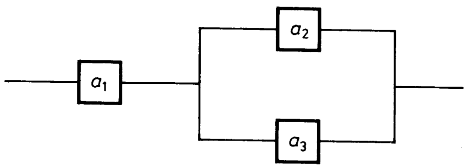
$$A \cup B=\{\omega_1, \omega_3, \omega_5\} \cup \{\omega_2, \omega_3, \omega_4\}=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}=\Omega,$$

$$AB=\{\omega_1, \omega_3, \omega_5\} \cap \{\omega_2, \omega_3, \omega_4\}=\{\omega_3\},$$

$$B \backslash A=\{\omega_2, \omega_3, \omega_4\} \backslash \{\omega_1, \omega_3, \omega_5\}=\{\omega_2, \omega_4\},$$

$$A \backslash B=\{\omega_1, \omega_3, \omega_5\} \backslash \{\omega_2, \omega_3, \omega_4\}=\{\omega_1, \omega_5\}.$$

ZADANIE 1.3. Rysunek 1.3 przedstawia schemat fragmentu sieci elektrycznej. Interesuje nas ciągły przepływ prądu przez ten fragment w odcinku czasu t. Niech A_i, $i=1, 2, 3$, oznacza zdarzenie „element a_i będzie sprawny w czasie t". Za pomocą działań na zdarzeniach A_i opisać zdarzenie A: „w odcinku czasu t przepływ prądu nie ulegnie przerwaniu".



Rys. 1.3. Do zadania 1.3

Rozwiązanie. Przepływ prądu nie ulegnie przerwaniu wtedy i tylko wtedy, gdy nie ulegnie uszkodzeniu element a_1 i nie ulegnie uszkodzeniu co najmniej jeden z elementów a_2, a_3. To samo stwierdzenie wypowiedziane za pomocą zdarzeń A oraz A_i brzmi: zdarzenie A zajdzie wtedy i tylko wtedy, gdy zajdzie zdarzenie A_1 i zajdzie co najmniej jedno ze zdarzeń A_2, A_3. Przy tym zajście co najmniej jednego ze zdarzeń A_2, A_3 oznacza zajście alternatywy $A_2 \cup A_3$. Zatem

$$A = A_1(A_2 \cup A_3) \tag{1}$$

lub, po zastosowaniu prawa rozdzielności koniunkcji zdarzeń względem alternatywy,

$$A = A_1 A_2 \cup A_1 A_3. \tag{2}$$

Proponujemy Czytelnikowi uzyskanie równości (2) bezpośrednio, tj. bez powoływania się na równość (1).

Zadanie 1.4. Praca mechanika polega na zabezpieczeniu technicznej sprawności trzech maszyn M_1, M_2, M_3 w ciągu odcinka czasu. W czasie tym każda z maszyn pracuje niezawodnie albo wymaga interwencji mechanika. Niech A_j, $j=1, 2, 3$, oznacza zdarzenie: maszyna M_j wymaga interwencji mechanika.

a) Określić przestrzeń zdarzeń elementarnych.

b) Za pomocą zdarzeń elementarnych opisać zdarzenia $A_1, A_1', A_2, A_2', A_3, A_3'$.

c) Za pomocą działań $\cap, \cup, '$, wykonanych na zdarzeniach A_j, opisać zdarzenia B_j, jeśli oznaczamy:

B_1 — zajście wszystkich trzech zdarzeń A_1, A_2, A_3,
B_2 — niezajście żadnego ze zdarzeń A_1, A_2, A_3,
B_3 — zajście tylko zdarzenia A_1,
B_4 — zajście tylko jednego spośród zdarzeń A_1, A_2, A_3,
B_5 — zajście co najmniej jednego ze zdarzeń A_1, A_2, A_3,
B_6 — zajście tylko zdarzeń A_1 i A_2,
B_7 — zajście dokładnie dwóch zdarzeń spośród zdarzeń A_1, A_2, A_3,
B_8 — zajście co najmniej dwóch zdarzeń spośród zdarzeń A_1, A_2, A_3,
B_9 — zajście co najwyżej jednego zdarzenia spośród zdarzeń A_1, A_2, A_3.

d) Opisać określone wyżej zdarzenia B_j za pomocą zdarzeń elementarnych.

e) Znaleźć liczbę $n(\mathscr{Z})$ wszystkich zdarzeń w tym doświadczeniu.

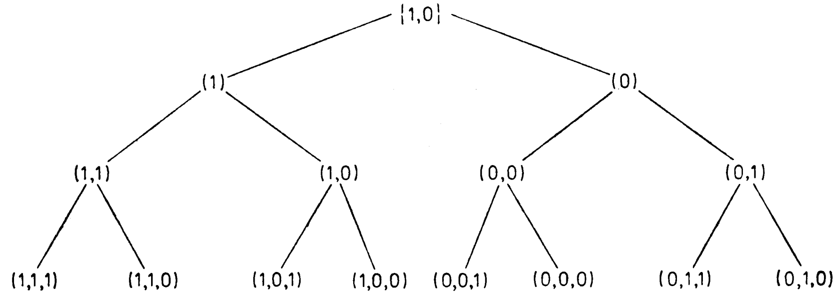
Rozwiązanie. a) W roli zdarzeń elementarnych można przyjąć trzywyrazowe ciągi zer i jedynek: $\omega=(i_1, i_2, i_3)$, gdzie

$$i_j=\begin{cases} 0, & \text{gdy } j\text{-ta maszyna nie wymaga interwencji, } j=1,2,3,\\ 1, & \text{gdy } j\text{-ta maszyna wymaga interwencji.}\end{cases}$$

Na przykład zdarzenie elementarne (0, 1, 0) oznacza: tylko maszyna M_2 wymaga interwencji mechanika. Przestrzenią zdarzeń elementarnych jest tu więc zbiór 8 (bo $2^3=8$) trzywyrazowych ciągów zer i jedynek:

$$\Omega=\{(1,1,1),(1,1,0),(1,0,1),(0,1,1),(1,0,0),(0,1,0),(0,0,1),(0,0,0)\}.$$

Wypisanie w systematyczny sposób wszystkich zdarzeń elementarnych ułatwia rysunek (tzw. *graf*) 1.4.



Rys. 1.4. Do zadania 1.4

b) Zdarzeniami elementarnymi sprzyjającymi zdarzeniu A_j są te zdarzenia elementarne, tj. te ciągi trzywyrazowe, w których j-ty wyraz jest jednością, $i_j=1$. Dlatego:

$$A_1=\{(1,1,1),(1,1,0),(1,0,1),(1,0,0)\},$$
$$A_2=\{(1,1,1),(1,1,0),(0,1,1),(0,1,0)\},$$
$$A_3=\{(1,1,1),(1,0,1),(0,1,1),(0,0,1)\}.$$

Zdarzenie A_i' składa się, zgodnie z definicją, z wszystkich zdarzeń elementarnych, które nie należą do zdarzenia A_i. Zatem

$$A_1'=\{(0,1,1),(0,1,0),(0,0,1),(0,0,0)\},$$
$$A_2'=\{(1,0,1),(1,0,0),(0,0,1),(0,0,0)\},$$
$$A_3'=\{(1,1,0),(1,0,0),(0,1,0),(0,0,0)\}.$$

c) Bezpośrednio z definicji alternatywy i koniunkcji zdarzeń oraz definicji zdarzenia przeciwnego, otrzymujemy:

$$B_1=A_1A_2A_3, \quad B_2=A_1'A_2'A_3', \quad B_3=A_1A_2'A_3',$$
$$B_4=A_1A_2'A_3'\cup A_1'A_2A_3'\cup A_1'A_2'A_3, \quad B_5=A_1\cup A_2\cup A_3,$$
$$B_6=A_1A_2A_3', \quad B_7=A_1A_2A_3'\cup A_1A_2'A_3\cup A_1'A_2A_3,$$
$$B_8=A_1A_2\cup A_1A_3\cup A_2A_3,$$
$$B_9=A_1'A_2'A_3'\cup A_1A_2'A_3'\cup A_1'A_2A_3'\cup A_1'A_2'A_3.$$

d) Korzystając z b) i c), definicji koniunkcji i alternatywy zdarzeń oraz z definicji zdarzenia przeciwnego, otrzymujemy:

$$B_1=\{(1,1,1)\},\quad B_2=\{(0,0,0)\},\quad B_3=\{(1,0,0)\};$$

$$B_4=\{(1,0,0),(0,1,0),(0,0,1)\},$$

$$B_5=\{(1,1,1),(1,1,0),(1,0,1),(0,1,1),(1,0,0),(0,1,0),(0,0,1)\},$$

$$B_6=\{(1,1,0)\},\quad B_7=\{(1,1,0),(1,0,1),(0,1,1)\},$$

$$B_8=\{(1,1,1),(1,1,0),(1,0,1),(0,1,1)\},$$

$$B_9=\{(0,0,0),(1,0,0),(0,1,0),(0,0,1)\}.$$

e) Niech $n(A)$ oznacza liczbę elementów zbioru A. Jak wiadomo liczba wszystkich podzbiorów n-elementowego zbioru A wynosi 2^n. Przestrzeń Ω jest skończona, $n(\Omega)=8$, dlatego każdy jej podzbiór można przyjąć za zdarzenie. Ponieważ $n(\Omega)=8$, więc liczba $n(\mathscr{Z})$ wszystkich zdarzeń tworzących klasę $\mathscr{Z}$ wynosi 2^8, czyli

$$n(\mathscr{Z})=2^8=256.$$

ZADANIE 1.5. Zdarzenia E_1, E_2 są opisane za pomocą danych zdarzeń A, B, C równościami:

a) $E_1=(A\cup B)(A'\cup B)(A\cup B')$, b) $E_2=A\cup AC\cup BC\cup A'B$.

Uprościć prawe strony tych równości.

Rozwiązanie. a) Wykorzystując własności działań na zdarzeniach, otrzymujemy

$$E_1=(AA'\cup AB\cup BA'\cup BB)(A\cup B').$$

Ale

$$AA'=\emptyset,\qquad BB=B,\qquad AB\cup BA'=(A\cup A')B=\Omega B=B.$$

Dlatego

$$E_1=(B\cup B)(A\cup B')=B(A\cup B')=BA\cup BB'=AB\cup\emptyset=AB.$$

Tak więc, otrzymujemy $E_1=AB$. Proponujemy Czytelnikowi rozpocząć upraszczanie prawej strony równości a) od przekształcenia iloczynu $(A'\cup B)(A\cup B')$.

b) Ponieważ

$$\begin{aligned} &A\cup AC=A && (\text{gdyż } AC\subset A),\\ &A=A\cup AB && (\text{dlaczego?}),\\ &AB\cup A'B=B && (\text{p. wyżej}),\\ &BC\cup B=B && (\text{dlaczego?}), \end{aligned}$$

więc

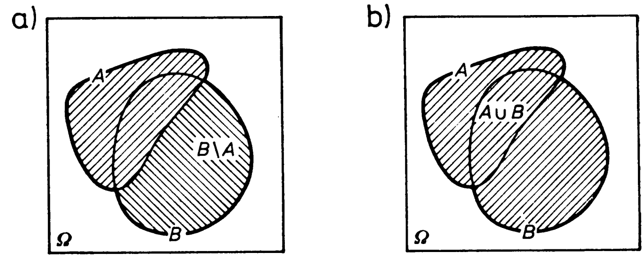
$$E_2=A\cup AB\cup BC\cup A'B=A\cup BC\cup(AB\cup A'B)=A\cup BC\cup B=$$
$$=A\cup(BC\cup B)=A\cup B.$$

czyli

$$E_2 = A \cup B.$$

Widzimy, że opis zdarzeń E_1 i E_2 znacznie się uprościł.

ZADANIE 1.6. Wykazać, że:
a) $(B\backslash A) \cup A = A \cup B$, b) $(B\backslash A) \cup A = B \Leftrightarrow A \subset B$.



Rys. 1.5. Do zadania 1.6

Rozwiązanie. a) Lewą stronę równości ilustruje rysunek 1.5a, prawą rysunek 1.5b. Spełnienie tej równości jest więc widoczne. Przytoczymy jednak jej formalny dowód. Dowód równości zdarzeń przebiega tak jak równości zbiorów (bo zdarzenia, to przecież zbiory) zgodnie z definicją równości zdarzeń. Mianowicie należy wykazać, że (ω, jak zwykle, oznacza zdarzenie elementarne):

1° $\omega \in \big((B\backslash A) \cup A\big) \Rightarrow \omega \in (A \cup B)$,

czyli zachodzenie inkluzji $\big((B\backslash A) \cup A\big) \subset (A \cup B)$ oraz

2° $\omega \in (A \cup B) \Rightarrow \omega \in \big((B\backslash A) \cup A\big)$,

czyli zachodzenie inkluzji $(A \cup B) \subset \big((B\backslash A) \cup A\big)$.

Dowód obydwu implikacji 1° i 2° przeprowadzimy jednocześnie:

$$\omega \in \big((B\backslash A) \cup A\big)$$
$$\Updownarrow$$
$$(\omega \in B \wedge \omega \notin A) \vee \omega \in A$$
$$\Updownarrow$$
$$(\omega \in B \vee \omega \in A) \wedge (\omega \in A' \vee \omega \in A)$$
$$\Updownarrow$$
$$(\omega \in (A \cup B)) \wedge \omega \in \Omega$$
$$\Updownarrow$$
$$\omega \in (A \cup B).$$

Zauważmy, że z udowodnionej równości wynika, iż zbiór dwu zdarzeń $\{B\backslash A, A\}$ stanowi podział zdarzenia $A \cup B$.

b) Występująca tu równość, wobec udowodnionej już równości a), jest równoważna następującej:

$$A \cup B = B.$$

Ale ta równość jest możliwa wtedy i tylko wtedy, gdy $A \subset B$. Równoważność jest więc udowodniona.

W związku z prawami działań na zdarzeniach (str. 10) i przykładami 1.5 i 1.6 nasuwa się następująca uwaga. Aby pomnożyć zdarzenie przez sumę zdarzeń wystarczy to zdarzenie pomnożyć przez każdy składnik sumy i dodać tak otrzymane iloczyny. Podobnie przy mnożeniu sumy zdarzeń przez sumę zdarzeń wystarczy każdy składnik jednej sumy pomnożyć przez każdy składnik drugiej sumy i dodać tak otrzymane iloczyny. Obserwujemy tu pewną analogię z działaniami w arytmetyce. Podkreślamy jednak, że analogia ta nie jest pełna. Świadczy o tym, między innymi, zadanie 1.6a, gdzie wydawałoby się, że zawsze powinno być $(B\backslash A) \cup A=B$. O tylko częściowej analogii świadczą również równości: $AA=A$, $A \cup A=A$. Poza tym, z tych równości wynika, że operując zdarzeniami, nie ma potrzeby używania ani wykładników potęg, ani współczynników.

1.2. PRAWDOPODOBIEŃSTWO. ELEMENTARNE TWIERDZENIA

1.2.1. Aksjomatyczna definicja prawdopodobieństwa. Niech Ω będzie przestrzenią zdarzeń elementarnych doświadczenia losowego D, $\mathscr{Z}$ — jego zbiorem zdarzeń losowych. *Prawdopodobieństwem* nazywamy funkcję P przyporządkowującą każdemu zdarzeniu $A \in \mathscr{Z}$ liczbę $P(A)$ zgodnie z następującymi warunkami:

P1. $P(A)\geqslant 0$ *dla każdego zdarzenia* $A \in \mathscr{Z}$,

P2. $P(\Omega)=1$,

P3. *jeżeli* $A_1, \ldots, A_n, \ldots$ *jest dowolnym ciągiem parami rozłącznych zdarzeń ze zbioru* $\mathscr{Z}$, *to*

$$P(A_1 \cup \ldots \cup A_n \cup \ldots)=P(A_1)+\ldots+P(A_n)+\ldots \tag{1.2.1}$$

Wartość prawdopodobieństwa dla danego zdarzenia $A \in \mathscr{Z}$, czyli liczbę $P(A)$ nazywamy *prawdopodobieństwem zdarzenia A*. Podana tu definicja prawdopodobieństwa jest definicją aksjomatyczną. Sformułował ją w 1931 r. radziecki matematyk A. N. Kołmogorow. Warunki P1, P2, P3 nazywamy *aksjomatami prawdopodobieństwa*. Aksjomat P2 nazywamy *aksjomatem unormowania*, P3 — *aksjomatem przeliczalnej addytywności*.

Aksjomatyczna definicja prawdopodobieństwa nie określa prawdopodobieństw poszczególnych zdarzeń ze zbioru zdarzeń $\mathscr{Z}$ — formułuje tylko warunki, jakie te prawdopodobieństwa muszą spełniać, o jakie musi zadbać badacz, określający prawdopodobieństwa zdarzeń w konkretnym doświadczeniu.

1.2.2. Elementarne własności prawdopodobieństwa. Z definicji prawdopodobieństwa i własności działań na zdarzeniach wynikają następujące, elementarne własności prawdopodobieństwa:

1° *Prawdopodobieństwo zdarzenia niemożliwego równa się zeru:*

$$P(\varnothing)=0. \tag{1.2.2}$$

2° *Jeżeli zdarzenie A pociąga zdarzenie B, $A\subset B$, to*

$$P(A)\leqslant P(B). \tag{1.2.3}$$

3° *Prawdopodobieństwo dowolnego zdarzenia jest nie większe od jedności:*

$$P(A) \leqslant 1, \quad A - \textit{dowolne}. \tag{1.2.4}$$

4° *Jeżeli zdarzenie A pociąga zdarzenie B, $A \subset B$, to*

$$P(B \setminus A) = P(B) - P(A). \tag{1.2.5}$$

5° *Jeżeli zdarzenia $A_1, \ldots, A_n$ są rozłączne parami, to*

$$P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n). \tag{1.2.6}$$

6° *Suma prawdopodobieństw zdarzeń przeciwnych równa się jedności:*

$$P(A) + P(A') = 1. \tag{1.2.7}$$

7° *Prawdopodobieństwo alternatywy dwóch dowolnych zdarzeń (czyli prawdopodobieństwo zajścia co najmniej jednego z tych zdarzeń) jest równe sumie prawdopodobieństw tych zdarzeń zmniejszonej o prawdopodobieństwo ich koniunkcji, czyli*

$$P(A \cup B) = P(A) + P(B) - P(AB). \tag{1.2.8}$$

8° *Jeżeli przestrzeń zdarzeń elementarnych Ω jest co najwyżej przeliczalna i przy tym są określone prawdopodobieństwa p_i poszczególnych zdarzeń jednoelementowych $\{\omega_i\}$, czyli*

$$p_i = P(\{\omega_i\}), \quad p_i \geqslant 0,$$

$$p_1 + \ldots + p_n = 1, \quad \textit{gdy przestrzeń } \Omega \textit{ jest skończona},$$

$$p_1 + \ldots + p_n + \ldots = 1, \quad \textit{gdy przestrzeń } \Omega \textit{ jest przeliczalna},$$

to prawdopodobieństwo zdarzenia A, któremu sprzyjają zdarzenia elementarne $\omega_{i_1}, \ldots, \omega_{i_k}$, jest dane równością:

$$P(A) = p_{i_1} + \ldots + p_{i_k}. \tag{1.2.9}$$

9° (Klasyczna definicja prawdopodobieństwa). *Jeżeli:*

a) *przestrzeń Ω składa się z n zdarzeń elementarnych, czyli $n(\Omega) = n$,*

b) *zdarzenia jednoelementowe $\{\omega_i\}$ są jednakowo prawdopodobne, a więc*

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

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

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{k}{n} = \frac{\text{liczba zdarzeń elementarnych sprzyjających zdarzeniu } A}{\text{liczba wszystkich zdarzeń elementarnych przestrzeni } \Omega}. \tag{1.2.10}$$

W dawniejszym ujęciu wzór (1.2.10) stanowił definicję prawdopodobieństwa zdarzenia i dlatego dla ostatniej własności pozostawiliśmy nazwę „klasyczna definicja prawdopodobieństwa” mimo, że w aksjomatycznym ujęciu własność ta jest twierdzeniem.

1.2.3. Przestrzeń probabilistyczna. Do tej pory z dowolnym doświadczeniem losowym D związaliśmy trzy następujące pojęcia: przestrzeń zdarzeń elementarnych Ω, zbiór zdarzeń $\mathscr{Z}$ i prawdopodobieństwo P (określone na zdarzeniach należących do $\mathscr{Z}$). Tę trójkę $(\Omega, \mathscr{Z}, P)$ *nazywamy przestrzenią probabilistyczną* (doświadczenia losowego D). Przestrzeń probabilistyczna danego doświadczenia losowego stanowi matematyczny opis tego doświadczenia. Większość zadań w rachunku prawdopodobieństwa polega na obliczaniu prawdopodobieństw $P(A)$ tylko dla niektórych zdarzeń A ze zbioru $\mathscr{Z}$.

1.2.4. Zadania rozwiązane.

ZADANIE 1.7. Zbudować przestrzeń probabilistyczną doświadczenia z przykładu 1.4 zakładając, że wszystkie określone tam zdarzenia jednoelementowe są jednakowo prawdopodobne.

Rozwiązanie. Jako zdarzenia elementarne ω przyjmujemy trzywyrazowe ciągi zer i jedynek: $\omega=(i_1, i_2, i_3)$, gdzie $i_j=1$ albo 0 stosownie do tego, czy j-ta, $j=1, 2, 3$, maszyna wymagała interwencji mechanika, czy pracowała bezawaryjnie. W konsekwencji otrzymujemy następującą przestrzeń zdarzeń elementarnych:

$$\Omega=\{(1,1,1),(1,1,0),(1,0,1),(0,1,1),(1,0,0),(0,1,0),(0,0,1),(0,0,0)\}. \tag{1}$$

Jest to pierwszy element budowanej przestrzeni probabilistycznej $(\Omega, \mathscr{Z}, P)$.

Z kolei znajdujemy zbiór zdarzeń $\mathscr{Z}$. Ponieważ jest to skończona przestrzeń zdarzeń elementarnych, więc zdarzeniami, czyli elementami zbioru $\mathscr{Z}$, są tu wszystkie podzbiory przestrzeni Ω.

Jest ich 256 (bo $2^8=256$):

jedno zdarzenie niemożliwe $A_1=\emptyset$,

8 zdarzeń jednoelementowych

$$A_2=\{\omega_1\}, \quad A_3=\{\omega_2\}, \quad \ldots, \quad A_9=\{\omega_8\},$$

28 $\left(\text{tj. } \binom{8}{2}\right)$ zdarzeń dwuelementowych

$$A_{10}=\{\omega_1, \omega_2\}, \quad A_{11}=\{\omega_1, \omega_3\}, \quad \ldots, \quad A_{37}=\{\omega_7, \omega_8\},$$

56 $\left(\text{tj. } \binom{8}{3}\right)$ zdarzeń trójelementowych

$$A_{38}=\{\omega_1, \omega_2, \omega_3\}, \quad \ldots, \quad A_{93}=\{\omega_6, \omega_7, \omega_8\},$$

70 $\left(\text{tj. } \binom{8}{4}\right)$ zdarzeń czteroelementowych

$$A_{94}=\{\omega_1, \omega_2, \omega_3, \omega_4\}, \quad \ldots, \quad A_{163}=\{\omega_5, \omega_6, \omega_7, \omega_8\},$$

56 $\left(\text{tj. } \binom{8}{5}\right)$ zdarzeń pięcioelementowych

$$A_{164}=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}, \quad \ldots, \quad A_{219}=\{\omega_4, \omega_5, \omega_6, \omega_7, \omega_8\},$$

28 $\left(\text{tj. } \binom{8}{6}\right)$ zdarzeń sześcioelementowych

$$A_{220}=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}, \quad \ldots, \quad A_{247}=\{\omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8\},$$

8 $\left(\text{tj. } \binom{8}{7}\right)$ zdarzeń siedmioelementowych

$$A_{248}=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7\}, \quad \ldots, \quad A_{255}=\{\omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8\}$$

oraz jedno zdarzenie ośmioelementowe, czyli

$$A_{256}=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8\}=\Omega.$$

Szukanym zbiorem zdarzeń jest więc zbiór $\mathscr{Z}$ postaci

$$\mathscr{Z}=\{A_1, A_2, \ldots, A_{256}\}. \tag{2}$$

Pozostaje określić trzeci element przestrzeni probabilistycznej, czyli prawdopodobieństwo P. Posłużmy się przy tym własnością (1.2.9) prawdopodobieństwa i wykorzystamy założenie, że zdarzenia jednoelementowe są jednakowo prawdopodobne:

$$P(\{\omega_1\})=P(\{\omega_2\})=\ldots=P(\{\omega_8\})=\tfrac{1}{8}.$$

Mamy zatem:

$$\begin{aligned}
&P(A_1)=P(\emptyset)=0, \quad P(A_{256})=P(\Omega)=1,\\
&P(A_2)=\ldots=P(A_9)=\tfrac{1}{8},\\
&P(A_{10})=\ldots=P(A_{37})=\tfrac{1}{8}+\tfrac{1}{8}=\tfrac{2}{8},\\
&P(A_{38})=\ldots=P(A_{93})=\tfrac{1}{8}+\tfrac{1}{8}+\tfrac{1}{8}=\tfrac{3}{8},\\
&P(A_{94})=\ldots=P(A_{163})=\tfrac{1}{8}+\tfrac{1}{8}+\tfrac{1}{8}+\tfrac{1}{8}=\tfrac{4}{8},\\
&P(A_{164})=\ldots=P(A_{219})=\tfrac{5}{8},\\
&P(A_{220})=\ldots=P(A_{247})=\tfrac{6}{8},\\
&P(A_{248})=\ldots=P(A_{255})=\tfrac{7}{8}.
\end{aligned} \tag{3}$$

Tak więc budowaną przestrzenią probabilistyczną rozpatrywanego doświadczenia jest $(\Omega, \mathscr{Z}, P)$, gdzie elementy Ω, $\mathscr{Z}$, P są dane odpowiednio przez (1), (2), (3).

W konkretnych zagadnieniach praktycznych, jak już wspominaliśmy, nie ma potrzeby wyznaczania prawdopodobieństw wszystkich zdarzeń tworzących zbiór $\mathscr{Z}$ — zwykle interesują nas prawdopodobieństwa $P(A)$ tylko niektórych zdarzeń.

ZADANIE 1.8. Dla danych z przykładu 1.7 znaleźć prawdopodobieństwa tego, że:

a) wszystkie 3 maszyny będą wymagały interwencji mechanika (zdarzenie A),

b) dokładnie 2 maszyny będą wymagały interwencji mechanika (zdarzenie B).

Rozwiązanie. a) Szukane prawdopodobieństwo jest po prostu prawdopodobieństwem zdarzenia jednoelementowego $\{\omega_1\}=\{(1, 1, 1)\}$

$$P(A)=P(\{\omega_1\})=\tfrac{1}{8}.$$

b) Jest to prawdopodobieństwo zdarzenia trójelementowego

$$P(B)=P(\{(1,1,0),(1,0,1),(0,1,1)\})=\tfrac{3}{8}.$$

Jest to jedno z prawdopodobieństw zdarzeń danych równościami (3) z poprzedniego przykładu.

ZADANIE 1.9. Osoba X wykonuje pewną pracę w ciągu 4, 5 albo 6 godzin i może popełnić przy tym 0, 1 albo 2 błędy. Zakładając jednakowe prawdopodobieństwo dla każdego z 9 zdarzeń jednoelementowych, znaleźć prawdopodobieństwa następujących zdarzeń:

a) Praca zostanie wykonana w ciągu 4 godzin (zdarzenie A).

b) Praca zostanie wykonana bezbłędnie w czasie 6 godzin (zdarzenie B).

c) Praca zostanie wykonana w czasie 5 godzin najwyżej z jednym błędem (zdarzenie C).

d) Praca zostanie wykonana z co najwyżej jednym błędem (zdarzenie D).

Rozwiązanie. Jako zdarzenia elementarne rozpatrywanego doświadczenia przyjmujemy dwuwyrazowe ciągi (i_1, i_2), gdzie $i_1=4, 5, 6$ stosownie do czasu wykonania pracy, $i_2=0, 1, 2$ stosownie do liczby popełnionych błędów. Mamy więc następującą przestrzeń zdarzeń elementarnych:

$$\Omega=\{(4,0),(4,1),(4,2),(5,0),(5,1),(5,2),(6,0),(6,1),(6,2)\}.$$

a) Zdarzeniu A sprzyjają 3 zdarzenia elementarne:

$$A=\{(4,0),(4,1),(4,2)\}.$$

Zatem

$$P(A)=\frac{n(A)}{n(\Omega)}=\frac{3}{9}=\frac{1}{3}.$$

b) Zdarzeniu B sprzyja tylko jedno zdarzenie elementarne (6, 0). Zatem

$$P(B)=\frac{n(B)}{n(\Omega)}=\frac{1}{9}.$$

c) Zdarzeniu C sprzyjają 2 zdarzenia elementarne:

$$C=\{(5,0),(5,1)\}.$$

Zatem

$$P(C)=\frac{n(C)}{n(\Omega)}=\frac{2}{9}.$$

d) Zdarzeniu D sprzyja 6 zdarzeń elementarnych:

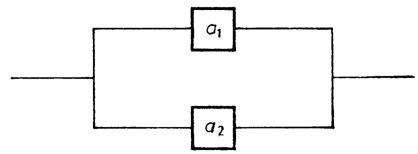
$$D=\{(4,0),(4,1),(5,0),(5,1),(6,0),(6,1)\}$$

i dlatego

$$P(D)=\frac{n(D)}{n(\Omega)}=\frac{6}{9}=\frac{2}{3}.$$

ZADANIE 1.10. Rysunek 1.6 przedstawia schemat fragmentu sieci elektrycznej. Niech A_i, $i=1, 2$, oznacza zdarzenie, że element a_i będzie sprawny co najmniej przez czas t. Obliczyć prawdopodobieństwo ciągłego przepływu prądu przez ten układ co najmniej przez czas t jeżeli $P(A_1)=P(A_2)=p$ oraz $P(A_1A_2)=p^2$.

Rys. 1.6. Do zadania 1.10



Rozwiązanie. Niech A oznacza zdarzenie, którego prawdopodobieństwo należy obliczyć. Układ przedstawiony na rysunku jest sprawny, gdy: 1° sprawny jest tylko element a_1, 2° sprawny jest tylko element a_2, 3° sprawne są obydwa elementy a_1 i a_2. Zatem zdarzenie A jest alternatywą zdarzeń A_1 i A_2: $A=A_1 \cup A_2$. Ponieważ zdarzenia A_1 i A_2 nie wykluczają się, więc na podstawie wzoru (1.2.8), otrzymujemy:

$$P(A)=P(A_1)+P(A_2)-P(A_1A_2)=p+p-p^2=p(2-p).$$

ZADANIE 1.11. Rozpatrujemy ilość (dm^3) wody jaką może mieć do przeprowadzenia w ciągu sekundy betonowy przepust. Dotychczasowe obserwacje pozwalają przyjąć, że: maksymalna możliwa ilość wody wynosi 300 dm^3/s,

$P(A)$ — prawdopodobieństwo, że ilość wody (na sekundę) przyjmie wartość z przedziału $(125, 250\rangle$ wynosi 0,6,

$P(B)$ — prawdopodobieństwo, że ilość wody (na sekundę) przyjmie wartość z przedziału $(200, 300\rangle$ wynosi 0,7 oraz $P(A \cup B)=0{,}8$. Obliczyć prawdopodobieństwa: a) $P(A')$, $P(B')$, b) $P(AB)$, c) $P(A'B')$, d) $P(BA')$, e) $P(B \cup A')$.

Rozwiązanie. a) Suma prawdopodobieństw zdarzeń przeciwnych jest jednością, a więc

$$P(A')=1-P(A)=1-0{,}6=0{,}4,$$

$$P(B')=1-P(B)=1-0{,}7=0{,}3.$$

b) We wzorze (1.2.8) na prawdopodobieństwo alternatywy dwu dowolnych zdarzeń dane są $P(A)$, $P(B)$ oraz $P(A \cup B)$. Wobec tego możemy obliczyć interesujące nas prawdopodobieństwo $P(AB)=P$ (ilość wody przyjmie wartość z przedziału $(200, 250\rangle$):

$$P(AB)=P(A)+P(B)-P(A \cup B)=0{,}6+0{,}7-0{,}8=0{,}5.$$

c) Ponieważ $A'B'=(A \cup B)'$, więc

$$P(A'B')=P\,(\text{ilość wody przyjmie wartość z przedziału } \langle 0, 125\rangle)=$$
$$=1-P(A \cup B)=0{,}2.$$

d) Ponieważ $BA'=B\backslash AB$ (sprawdzić), więc, na podstawie równości (1.2.5),

$$P(BA')=P(B)-P(AB)=0{,}7-0{,}5=0{,}2.$$

e) Postępując podobnie jak w punkcie b), otrzymujemy:

$$P(B \cup A') = P(B) + P(A') - P(BA') = 0{,}7 + 0{,}4 - 0{,}2 = 0{,}9.$$

Zadanie 1.12. Z n-elementowego zbioru A losujemy ze zwrotem k elementów. Obliczyć prawdopodobieństwo, że wszystkie wylosowane elementy są różne.

Rozwiązanie. Zdarzeniami elementarnymi są tu wszystkie k-wyrazowe ciągi o niekoniecznie różnych wyrazach, będących elementami zbioru A. Zdarzeń tych jest n^k (bo tyle jest k-elementowych wariacji z powtórzeniami z n elementów (1.5.6)).

Niech B oznacza zdarzenie, którego prawdopodobieństwo należy obliczyć. Zdarzeniami elementarnymi sprzyjającymi zajściu zdarzenia B są te k-wyrazowe ciągi, których wyrazy są różnymi elementami zbioru A. Jest ich $n^{[k]}$ (bo tyle jest k-wyrazowych wariacji bez powtórzeń z n-elementowego zbioru A (1.5.4')). Zatem szukane prawdopodobieństwo wynosi:

$$P(B) = \frac{n^{[k]}}{n^k}, \quad \text{gdzie} \quad k \in N. \tag{1}$$

Prawdopodobieństwo $P(B)$ obliczyliśmy przy założeniu, że $k \leqslant n$. Gdy $k > n$, wtedy $P(B) = 0$, gdyż wtedy elementy w k-elementowej próbce muszą się powtarzać. Z drugiej strony $n^{[k]} = n(n-1)(n-2) \ldots (n-k+1) = 0$ dla $k > n$. Zatem równość (1) jest prawdziwa dla dowolnego naturalnego k.

1.3. PRAWDOPODOBIEŃSTWO WARUNKOWE I NIEZALEŻNOŚĆ ZDARZEŃ

1.3.1. Określenie prawdopodobieństwa warunkowego. Niech A, B będą dowolnymi zdarzeniami. Prawdopodobieństwo zdarzenia A obliczone przy założeniu, że zaszło zdarzenie B nazywamy *prawdopodobieństwem warunkowym* zdarzenia A pod warunkiem B i oznaczamy symbolem $P(A|B)$ albo $P_B(A)$. Określenie to jest równoważne następującemu, bardziej formalnemu.

Niech $(\Omega, \mathscr{Z}, P)$ będzie przestrzenią probabilistyczną pewnego doświadczenia, B zaś dowolnym ustalonym zdarzeniem o dodatnim prawdopodobieństwie, $P(B) > 0$. *Prawdopodobieństwem warunkowym pod warunkiem B dowolnego zdarzenia A* $\in \mathscr{Z}$ nazywamy liczbę $P(A|B)$ określoną następującą równością:

$$P(A|B) = \frac{P(AB)}{P(B)}, \qquad A, B \in \mathscr{Z}, \qquad P(B) > 0. \tag{1.3.1}$$

Wynika stąd wzór na *prawdopodobieństwo koniunkcji* (łącznego zajścia) *dwu zdarzeń*:

$$\begin{aligned} P(AB) &= P(A)\,P(B|A), \quad \text{gdy} \quad P(A) > 0, \\ P(AB) &= P(B)\,P(A|B), \quad \text{gdy} \quad P(B) > 0, \end{aligned} \tag{1.3.2}$$

czyli jest ono równe iloczynowi prawdopodobieństwa P jednego z tych zdarzeń przez prawdopodobieństwo warunkowe drugiego zdarzenia pod warunkiem, że pierwsze zaszło.

1.3.2. Prawdopodobieństwo koniunkcji n zdarzeń. Ze wzoru (1.3.2) łatwo otrzymujemy wzór na *prawdopodobieństwo koniunkcji trzech zdarzeń A, B, C*:

$$P(ABC)=P(A)P(B|A)P(C|AB), \quad \text{gdy} \quad P(AB)>0. \tag{1.3.3}$$

Ogólnie: metodą indukcji matematycznej wykazuje się, że *prawdopodobieństwo koniunkcji n zdarzeń* $A_1, \ldots, A_n$ wyraża się równością:

$$P(A_1 A_2 \ldots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\ldots P(A_n|A_1A_2\ldots A_{n-1}), \tag{1.3.4}$$

gdy $P(A_1A_2 \ldots A_{n-1})>0$, czyli jest ono równe iloczynowi prawdopodobieństwa dowolnego z tych zdarzeń przez prawdopodobieństwa warunkowe każdego z pozostałych zdarzeń obliczone pod warunkiem, że zaszły wszystkie poprzednie zdarzenia.

1.3.3. Niezależność dwóch zdarzeń. Mówimy, że zdarzenia $A, B \in \mathscr{Z}$ są *niezależne*, gdy

$$P(AB)=P(A)P(B). \tag{1.3.5}$$

Równość ta nie wyklucza sytuacji, gdy $P(A)=0$ lub $P(B)=0$.

Niech $P(A)>0$ i $P(B)>0$. Wówczas każda z równości

$$P(A|B)=P(A), \qquad P(B|A)=P(B) \tag{1.3.6}$$

stanowi warunek konieczny i wystarczający na to, aby zdarzenia były niezależne w sensie definicji (1.3.5). Zatem każda z tych równości, gdy $P(A)>0$ i $P(B)>0$, może być przyjęta jako określenie niezależności zdarzeń A, B.

1.3.4. Niezależność zespołowa n zdarzeń. Bardziej złożone jest pojęcie niezależności n zdarzeń. Mówimy mianowicie, że zdarzenia $A_1, \ldots, A_n$ są *niezależne* (dokładniej: *wzajemnie niezależne* albo *zespołowo niezależne*), gdy prawdopodobieństwo łącznego zajścia dowolnych m, $m \leqslant n$, różnych zdarzeń spośród nich jest równe iloczynowi prawdopodobieństw tych zdarzeń, czyli

$$P(A_{i_1}A_{i_2}\ldots A_{i_m})=P(A_{i_1})P(A_{i_2})\ldots P(A_{i_m}) \tag{1.3.7}$$

dla każdego $m \leqslant n$ i każdego m-wyrazowego rosnącego ciągu $i_1, i_2, \ldots, i_m$ liczb naturalnych $1 \leqslant i_1, i_2, \ldots, i_m \leqslant n$.

1.3.5. Niezależność parami n zdarzeń. W wielu zagadnieniach wystarcza tzw. niezależność parami n zdarzeń. Otóż mówimy, że zdarzenia $A_1, \ldots, A_n$ są *niezależne parami*, gdy każde dwa różne zdarzenia spośród nich są niezależne, czyli

$$P(A_iA_j)=P(A_i)P(A_j), \quad i \neq j, \quad i,j=1,2,\ldots,n. \tag{1.3.8}$$

Zdarzenia wzajemnie niezależne są oczywiście niezależne parami (dlaczego?).

1.3.6. Zadania rozwiązane.

ZADANIE 1.13. Ze zbioru n elementów, wśród których jest n_1 elementów mających cechę C i $n_2=n-n_1$ elementów nie mających tej cechy, losujemy dwukrotnie po jednym elemencie. Obliczyć prawdopodowieństwo tego, że obydwa wylosowane elementy mają cechę C. Przyjąć $n=10$, $n_1=7$.

Rozwiązanie. Niech

A_1 oznacza zdarzenie — pierwszy wylosowany element ma cechę C,

A_2 — drugi wylosowany element ma cechę C,

A — obydwa wylosowane elementy mają cechę C.

Zgodnie z definicją koniunkcji dwu zdarzeń mamy

$$A=A_1A_2.$$

W dalszym ciągu wyróżniamy dwa przypadki: a) przed losowaniem drugiego elementu nie zwracamy pierwszego (losowanie bez zwrotu) oraz b) przed losowaniem drugiego elementu dokonujemy zwrotu pierwszego po jego obejrzeniu (losowanie ze zwrotem).

Przy losowaniu bez zwrotu otrzymujemy

$$P(A)=P(A_1A_2)=P(A_1)P(A_2|A_1)=\frac{n_1}{n}\frac{n_1-1}{n-1}=\frac{7}{10}\cdot\frac{6}{9}=\frac{7}{15}.$$

W drugim przypadku, gdy losujemy ze zwrotem, warunki w drugim losowaniu są identyczne jak w pierwszym i wynik pierwszego losowania nie ma wpływu na prawdopodobieństwo zdarzenia A_2:

$$P(A_2|A_1)=P(A_2)=P(A_1)=\frac{n_1}{n}=\frac{7}{10}.$$

Dlatego

$$P(A_1A_2)=P(A_1)P(A_2|A_1)=P(A_1)P(A_1)=\frac{n_1}{n}\frac{n_1}{n}=\frac{7^2}{10^2}=0{,}49.$$

ZADANIE 1.14. Opiekun 20-osobowej grupy studenckiej, w której (jak się później okazało) było 8 studentów aktywnych społecznie, wytypował na okres dwu tygodni (aby dać czas na poznanie się grupy) trzyosobowy samorząd grupy. Jakie jest prawdopodobieństwo tego, że taki samorząd składać się będzie z samych studentów aktywnych społecznie?

Rozwiązanie. Niech A_i, $i=1, 2, 3$, oznacza zdarzenie: i-ty wytypowany student jest społecznie aktywny. Wówczas zdarzenie A, którego prawdopodobieństwo należy znaleźć, jest koniunkcją zdarzeń A_i:

$$A=A_1A_2A_3,$$

więc, na podstawie równości (1.3.3), otrzymujemy:

$$P(A)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)=\tfrac{8}{20}\cdot\tfrac{7}{19}\cdot\tfrac{6}{18}=\tfrac{14}{285}\approx 0{,}049.$$

ZADANIE 1.15. Spośród cyfr 1, ..., 9 wylosowano bez zwrotu kolejno trzy cyfry C_1, C_2, C_3, układając je w kolejności losowania w liczbę, która w układzie dziesiętnym ma zapis $C_1C_2C_3$. Przyjmując sensowne założenie, że wszystkie możliwe do otrzymania w ten sposób liczby są jednakowo prawdopodobne, obliczyć prawdopodobieństwo tego, że $C_1C_2C_3<444$.

Rozwiązanie. Niech A_i^4 oznacza zdarzenie – otrzymanie w i-tym, $i=1, 2, 3$, losowaniu cyfry 4; $A_i^{<4}$ – otrzymanie w i-tym losowaniu jednej spośród cyfr 1, 2, 3; A oznacza zdarzenie, którego prawdopodobieństwo należy obliczyć.

Zdarzenie A zachodzi wtedy i tylko wtedy, gdy zajdzie: 1° zdarzenie $A_1^{<4}$ (wtedy na II i III miejscu mogą znaleźć się dowolne spośród pozostałych cyfr) lub 2° dwa zdarzenia: A_1^4 i $A_2^{<4}$ (wtedy na III miejscu może się znaleźć dowolna spośród pozostałych cyfr).

Wydawałoby się, że zdarzenie A zachodzi jeszcze w sytuacji gdy zajdą trzy zdarzenia: A_1^4 i A_2^4 i $A_3^{<4}$. Jednak, wobec losowania bez zwrotu jest to niemożliwe. Zatem zdarzenie A można opisać jak następuje:

$$A=A_1^{<4}\cup A_1^4A_2^{<4}.$$

Ponieważ zdarzenia $A_1^{<4}$ i A_1^4 wykluczają się, więc także wykluczają zdarzenia $A_1^{<4}$ i $A_1^4A_2^{<4}$. Wobec tego na podstawie addytywności prawdopodobieństwa i wzoru (1.3.2) na prawdopodobieństwo koniunkcji dwu zdarzeń, otrzymujemy:

$$P(A)=P(A_1^{<4})+P(A_1^4A_2^{<4})=P(A_1^{<4})+P(A_1^4)P(A_2^{<4}|A_1^4)=$$
$$=\tfrac{3}{9}+\tfrac{1}{9}\cdot\tfrac{3}{8}=\tfrac{3}{8}.$$

Tak więc szukane prawdopodobieństwo wynosi $\frac{3}{8}$.

Uwaga. Z metody rozwiązywania zadania widać, że wylosowanie dowolnej liczby trzycyfrowej mniejszej od liczby trzycyfrowej zapisanej w układzie dziesiętnym za pomocą tej samej cyfry $k\in\{1, ..., 9\}$ wynosi $\frac{1}{8}(k-1)$. Istotnie, rozumując analogicznie jak w zadaniu, mamy

$$\frac{k-1}{9}+\frac{1}{9}\cdot\frac{k-1}{8}=\frac{9(k-1)}{72}=\frac{k-1}{8}.$$

ZADANIE 1.16. Na trzech kolejnych zmianach dokonuje się przeglądu technicznego 2 spośród 6 maszyn, które należy poddać temu przeglądowi. Bez ponownego badania nie jest wiadome, czy maszyna została poddana przeglądowi. Obliczyć prawdopodobieństwo, że w ciągu tych trzech zmian wszystkie maszyny zostały poddane przeglądowi, gdyby kolejne zmiany nie przekazywały sobie informacji.

Rozwiązanie. Niech A_i^j, $i=1, ..., 6$, $j=1, 2, 3$ będzie zdarzeniem: i-ta maszyna na j-tej zmianie poddana zostanie przeglądowi, zaś A, jak zwykle, niech oznacza zdarzenie, którego prawdopodobieństwo należy obliczyć. Przy tych oznaczeniach uwzględniając, że zdarzenia A_1^1 i A_2^1 są pewne, zdarzenie A opisuje się następująco:

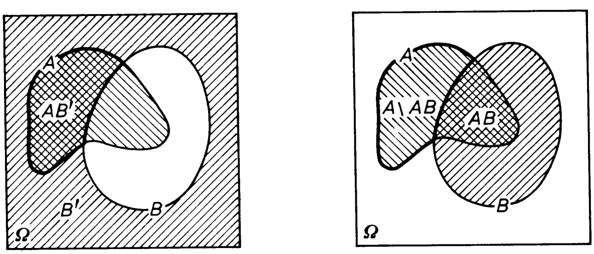
$$A=A_3^2A_4^2A_5^3A_6^3.$$

Stąd, zgodnie ze wzorem (1.3.4) na prawdopodobieństwo koniunkcji n zdarzeń, otrzymujemy:

$$P(A)=P(A_3^2)\,P(A_4^2|A_3^2)\,P(A_5^3|A_3^2A_4^2)\,P(A_6^3|A_3^2A_4^2A_5^3)=$$

$$=\frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{6}\cdot\frac{1}{5}=\frac{2}{75}.$$

Zadanie 1.17. Wykazać, że jeśli zdarzenia A i B są niezależne, to również niezależne są zdarzenia A i B'.



Rys. 1.7. Do zadania 1.17

Rozwiązanie. Należy wykazać, że jest spełniona równość $P(AB')=P(A)P(B')$. W tym celu posłużymy się równością (rys. 1.7):

$$AB'=A\backslash AB.$$

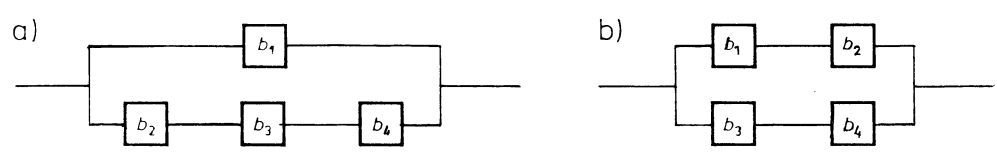
Kolejno otrzymujemy:

$$P(AB')=P(A)-P(AB)=P(A)-P(A)P(B)=P(A)\big(1-P(B)\big)=$$

$$=P(A)P(B').$$

Wykorzystaliśmy tu: wzór (1.2.5), (zauważając, że $AB\subset A$), niezależność zdarzeń A i B oraz związek (1.2.7) na sumę prawdopodobieństw zdarzeń przeciwnych.

Zadanie 1.18. Każde z dwóch urządzeń M_1 i M_2 składa się z 4 jednakowych bloków b_1, b_2, b_3 i b_4. Prawdopodobieństwo bezawaryjnej pracy w odcinku czasu t dla każdego z tych bloków jest jednakowe i wynosi p, $p>0$. Bloki ulegają uszkodzeniu niezależnie od siebie. Urządzenia M_1 i M_2 są montowane z tych bloków odpowiednio według schematów a) i b) na rysunku 1.8. Wyznaczyć prawdopodobieństwo bezawaryjnej pracy w odcinku czasu t dla: a) urządzenia M_1, b) urządzenia M_2 oraz c) porównać te prawdopodobieństwa.



Rys. 1.8. Do zadania 1.18

Rozwiązanie. Określamy zdarzenia A_i, $i=1, 2, 3, 4$, A, B jak następuje:
A_i — i-ty blok będzie pracował niezawodnie przez czas t,
A — urządzenie M_1 będzie pracowało niezawodnie przez czas t,
B — urządzenie M_2 będzie pracowało niezawodnie przez czas t.
Przy tych oznaczeniach mamy:

a) $A=A_1 \cup A_2A_3A_4$,

$$P(A)=P(A_1)+P(A_2A_3A_4)-P(A_1A_2A_3A_4)=$$

$$=P(A_1)+P(A_2)P(A_3)P(A_4)-P(A_1)P(A_2)P(A_3)P(A_4)=p+p^3-p^4.$$

Wykorzystaliśmy tu wzór na prawdopodobieństwo alternatywy dwu dowolnych zdarzeń i założoną niezależność zdarzeń A_1, A_2, A_3, A_4.

b) Podobnie dla zdarzenia B otrzymujemy:

$$B=A_1A_2 \cup A_3A_4,$$

$$P(B)=P(A_1A_2)+P(A_3A_4)-P(A_1A_2A_3A_4)=$$

$$=P(A_1)P(A_2)+P(A_3)P(A_4)-P(A_1)P(A_2)P(A_3)P(A_4)=2p^2-p^4.$$

c) Porównujemy prawdopodobieństwa $P(A)$ i $P(B)$:

$$P(A)\gtrless P(B),$$

$$p+p^3-p^4\gtrless 2p^2-p^4, \qquad p(p^2-2p+1)\gtrless 0.$$

Oczywiście, gdy $0\neq p\neq 1$,

$$p(p^2-2p+1)=p(1-p)^2>0.$$

Ostatecznie: $P(A)>P(B)$. Zatem, o ile nie ma innych przeciwwskazań, bloki należy montować według schematu a), gdyż zapewnia to większe prawdopodobieństwo niezawodnej pracy.

Zadanie (Bernsteina) 1.19. Trzy ściany czworościanu zostały pomalowane na biało, czerwono i zielono, zaś czwarta — w pasy biało-czerwono-zielone. Doświadczenie polega na rzucaniu czworościanu na płaszczyznę i obserwowaniu koloru ściany, na którą upadł czworościan. Zdarzenia B, C, Z określone są następująco:
B — czworościan upadł na ścianę białą,
C — czworościan upadł na ścianę czerwoną,
Z — czworościan upadł na ścianę zieloną.
Zbadać czy zdarzenia B, C, Z są: a) niezależne parami, b) niezależne (wzajemnie).

Rozwiązanie. a) Dwa zdarzenia z trzech zdarzeń B, C, Z można wybrać trzema sposobami (bo $\binom{3}{2}=\binom{3}{1}=3$): $\{B, C\}$, $\{B, Z\}$, $\{C, Z\}$. Mamy:

$$P(BC)=\tfrac{1}{4} \quad \text{i} \quad P(B)P(C)=\tfrac{2}{4}\cdot\tfrac{2}{4}=\tfrac{1}{4},$$

więc

$$P(BC)=P(B)P(C)$$

i tym samym zdarzenia B i C są niezależne. Podobnie sprawdza się, że zdarzenia B i Z oraz C i Z są niezależne. Zatem, zgodnie z definicją, zdarzenia B, C, Z są parami niezależne.

b) Ponieważ $P(BCZ)=\frac{1}{4}$ i $P(B)P(C)P(Z)=\frac{2}{4}\cdot\frac{2}{4}\cdot\frac{2}{4}=\frac{1}{8}$, więc

$$P(BCZ)\neq P(B)P(C)P(Z)$$

i, zgodnie z określeniem, zdarzenia B, C, Z nie są wzajemnie niezależne (mimo, że są niezależne parami).

ZADANIE 1.20. Sprawdzić, że wykluczające się zdarzenia A, B są: a) zależne, gdy $P(A)>0$ i $P(B)>0$, b) niezależne, gdy $P(A)=0$ lub $P(B)=0$. Podać przykłady dwu niezależnych zdarzeń, które: c) wykluczają się, d) nie wykluczają się.

Rozwiązanie. a) Wynika to z relacji:

$$P(AB)=0\neq P(A)P(B).$$

b) Wynika to z równości

$$P(AB)=0=P(A)P(B).$$

c) Przykładem pary zdarzeń niezależnych i wykluczających się może być dowolne zdarzenie A i zdarzenie niemożliwe (dlaczego?).

d) Przykładem pary zdarzeń niezależnych i niewykluczających się mogą być zdarzenia B i C z zadania 1.19.

Z rozwiązanego tu zadania widzimy, że:

– wykluczające się zdarzenia mogą być zarówno zależne jak i niezależne

oraz .

– zdarzenia niezależne mogą wykluczać się, ale również mogą mieć część wspólną.

Zatem: wykluczanie się i niezależność zdarzeń, to zupełnie różne pojęcia. Dodajmy przy tym, że pojęcie wykluczania się zdarzeń wprowadza się bez użycia pojęcia prawdopodobieństwa.

1.4. PRAWDOPODOBIEŃSTWO ZUPEŁNE I TWIERDZENIE BAYESA

1.4.1. Twierdzenie o prawdopodobieństwie zupełnym. Obliczanie prawdopodobieństwa ułatwiają często następujące twierdzenia (o prawdopodobieństwie zupełnym):

Jeżeli B jest dowolnym zdarzeniem, zdarzenia zaś $A_1, \ldots, A_n$ spełniają warunki (rys. 1.9):

1° *wykluczają się parami, czyli*

$$A_iA_j=\emptyset, \quad gdy \quad i\neq j,$$

2° *ich alternatywa jest zdarzeniem pewnym, czyli*

$$A_1\cup\ldots\cup A_n=\Omega,$$

3° *mają dodatnie prawdopodobieństwa,*

$$P(A_i)>0, \quad i=1,\ldots,n,$$

to prawdopodobieństwo zdarzenia B wyraża się równością

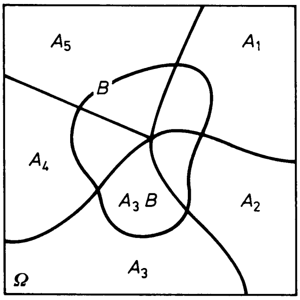
$$P(B)=P(A_1)P(B|A_1)+\ldots+P(A_n)P(B|A_n) \tag{1.4.1}$$

zwaną wzorem na *prawdopodobieństwo zupełne* (albo *prawdopodobieństwo całkowite*).

Warunki 1° i 2° można również wypowiedzieć tak:

Zbiór $\{A_1, \ldots, A_n\}$ *zdarzeń* A_i *stanowi podział* (§ 1.1) *przestrzeni zdarzeń elementarnych* Ω.

Rys. 1.9. Do założeń tw. o prawdopodobieństwie zupełnym i tw. Bayesa



Wyrażamy to również mówiąc, że *zdarzenia* $A_1, A_2, \ldots, A_n$ *są jedynie możliwe i wzajemnie wykluczające się*. W wyniku doświadczenia może zrealizować się jedno i tylko jedno z tych zdarzeń.

Niekiedy zdarzenia $A_1, \ldots, A_n$ nazywamy *przyczynami*, zdarzenie B zaś – *skutkiem*. Wówczas twierdzenie o prawdopodobieństwie zupełnym przyjmuje postać:

Jeżeli skutek B może zajść w wyniku jednej z n przyczyn $A_1, \ldots, A_n$ *jedynie możliwych i wzajemnie wykluczających się, to prawdopodobieństwo zajścia skutku B wyraża się równością* (1.4.1).

1.4.2. Twierdzenie Bayesa. Z ostatnim twierdzeniem wiąże się ściśle następujące *twierdzenie* (*Bayesa*).

Jeżeli B jest dowolnym zdarzeniem o dodatnim prawdopodobieństwie, $P(B)>0$, *zdarzenia* $A_1, \ldots, A_n$ *zaś spełniają warunki* 1°, 2°, 3° *twierdzenia o prawdopodobieństwie zupełnym, to prawdopodobieństwa warunkowe* $P(A_k|B)$ *zdarzeń* A_k *przy warunku B,* $k=1, \ldots, n$, *wyrażają się równościami:*

$$P(A_k|B)=\frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}, \quad k=1,\ldots,n. \tag{1.4.2}$$

Zauważmy, że mianownik jest prawdopodobieństwem zdarzenia B danym wzorem (1.4.1).

Twierdzenie to w terminologii przyczyna-skutek można sformułować jak następuje:

Jeśli skutek B zrealizował się w rezultacie zajścia jednej z n przyczyn $A_1, \ldots, A_n$, jedynie możliwych i wzajemnie wykluczających się, to prawdopodobieństwo $P(A_k|B)$ tego, że A_k, $k=1, \ldots, n$, było przyczyną zajścia skutku B wyraża się równością (1.4.2).

Używając łacińskich słów *a priori* (z góry, przed doświadczeniem) i *a posteriori* (po zbadaniu, po doświadczeniu) widzimy, że twierdzenie Bayesa wyraża prawdopodobieństwa *a posteriori* $P(A_k|B)$ przyczyn A_i za pomocą prawdopodobieństwa *a priori* $P(A_i)$ tych przyczyn.

1.4.3. Zadania rozwiązane.

Zadanie 1.21. Na przenośnik taśmowy trafiają jednakowe produkty wytwarzane przez 2 automaty. Stosunek ilościowy produkcji pierwszego automatu do produkcji drugiego jest równy 3 : 2. Pierwszy automat wytwarza średnio 65% produktów pierwszej jakości, drugi zaś — 85%. a) Spośród produktów na przenośniku pobieramy losowo jeden produkt. Obliczyć prawdopodobieństwo, że będzie to produkt pierwszej jakości. b) Losowo pobrany produkt okazał się pierwszej jakości. Mógł on zostać wyprodukowany przez automat pierwszy albo drugi. Które z tych zdarzeń jest bardziej prawdopodobne?

Rozwiązanie. a) Niech B, A_1, A_2 oznaczają zdarzenia:

B — pobrany losowo produkt z przenośnika jest pierwszej jakości,

A_1 — produkt znajdujący się na przenośniku został wytworzony przez pierwszy automat,

A_2 — produkt znajdujący się na przenośniku został wytworzony przez drugi automat.

Zdarzenie B może zrealizować się tylko łącznie z jednym ze zdarzeń A_1, A_2, które wykluczają się, są jedynie możliwe i mają dodatnie prawdopodobieństwa. Możemy zatem zastosować twierdzenie o prawdopodobieństwie zupełnym:

$$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)=$$
$$=\frac{3}{5}\cdot\frac{65}{100}+\frac{2}{5}\cdot\frac{85}{100}=\frac{3\cdot 13+2\cdot 17}{100}=0{,}73.$$

Tak więc $P(B)=0{,}73$.

b) Należy porównać prawdopodobieństwa $P(A_1|B)$ i $P(A_2|B)$, które znajdujemy na podstawie wzoru Bayesa (1.4.2). Mamy

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{\frac{3}{5}\cdot\frac{65}{100}}{0{,}73}=\frac{39}{73}$$

oraz

$$P(A_2|B)=1-P(A_1|B)=\tfrac{34}{73}$$

lub

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{P(B)}=\frac{\frac{2}{5}\cdot\frac{85}{100}}{0{,}73}=\frac{34}{73}.$$

Zatem $P(A_1|B)>P(A_2|B)$, czyli jest bardziej prawdopodobne, że wzięty z przenośnika produkt, który okazał się pierwszej jakości, jest wyprodukowany przez pierwszy automat.

ZADANIE 1.22. Na linii łączności nadaje się dwa rodzaje sygnałów w postaci kodowych kombinacji 111 albo 000 z prawdopodobieństwami (*a priori*) odpowiednio równymi 0,65 i 0,35. Sygnały podlegają losowym zakłóceniom, w rezultacie czego symbol 1 może być odebrany jako 0 z prawdopodobieństwem 0,2 i z takim samym prawdopodobieństwem symbol 0 może być odebrany jako 1. Zakładamy, że symbole 1 i 0 ulegają zakłóceniom niezależnie jeden od drugiego. a) Obliczyć prawdopodobieństwo odebrania na wyjściu sygnału: a_1) 111, a_2) 000, a_3) 010. b) Na wyjściu odebrano sygnał 010, jakie jest prawdopodobieństwo, że został on nadany jako sygnał 000? c) Na wyjściu odebrano sygnał 111, jakie jest prawdopodobieństwo, że został on nadany również jako 111?

Rozwiązanie. Niech A_1, A_2, B_1, B_2, B_3 oznaczają zdarzenia:

A_1 – na wejściu nadano sygnał 111,
A_2 – na wejściu nadano sygnał 000,
B_1 – na wyjściu odebrano sygnał 111,
B_2 – na wyjściu odebrano sygnał 000,
B_3 – na wyjściu odebrano sygnał 010.

Prawdopodobieństwo, że nadany symbol 1 nie ulegnie zniekształceniu wynosi $0{,}8=1-0{,}2$. Takie samo jest prawdopodobieństwo nie zniekształcenia symbolu 0. Ponieważ zakłócenia pojawiają się niezależnie, więc

$$\begin{aligned} P(B_1|A_1)&=0{,}8\cdot 0{,}8\cdot 0{,}8=0{,}512,\\ P(B_1|A_2)&=0{,}2\cdot 0{,}2\cdot 0{,}2=0{,}008,\\ P(B_2|A_1)&=0{,}2\cdot 0{,}2\cdot 0{,}2=0{,}008,\\ P(B_2|A_2)&=0{,}8\cdot 0{,}8\cdot 0{,}8=0{,}512,\\ P(B_3|A_1)&=0{,}2\cdot 0{,}8\cdot 0{,}2=0{,}032,\\ P(B_3|A_2)&=0{,}8\cdot 0{,}2\cdot 0{,}8=0{,}128. \end{aligned} \tag{1}$$

a_1) Ponieważ $P(A_1)=0{,}65$ i $P(A_2)=0{,}35$ więc, wykorzystując równości (1) i twierdzenie o prawdopodobieństwie zupełnym, otrzymujemy

$$P(B_1)=P(A_1)P(B_1|A_1)+P(A_2)P(B_1|A_2)=$$
$$=0{,}65\cdot 0{,}512+0{,}35\cdot 0{,}008=0{,}3356.$$

Podobnie obliczamy

a_2) $P(B_2)=P(A_1)P(B_2|A_1)+P(A_2)P(B_2|A_2)=$

$=0{,}65\cdot 0{,}008+0{,}35\cdot 0{,}512=0{,}1844,$

a_3) $P(B_3)=P(A_1)P(B_3|A_1)+P(A_2)P(B_3|A_2)=$

$=0{,}65\cdot 0{,}032+0{,}35\cdot 0{,}128=0{,}0656.$

b) Należy obliczyć prawdopodobieństwo *a posteriori* $P(A_2|B_3)$ zdarzenia A_2, które znajdujemy na podstawie wzoru Bayesa (1.4.2):

$$P(A_2|B_3)=\frac{P(A_2)P(B_3|A_2)}{P(B_3)}=\frac{0{,}35\cdot 0{,}128}{0{,}0656}\approx 0{,}683.$$

c) Analogicznie znajdujemy prawdopodobieństwo $P(A_1|B_1)$:

$$P(A_1|B_1)=\frac{P(A_1)P(B_1|A_1)}{P(B_1)}=\frac{0{,}65\cdot 0{,}512}{0{,}3356}\approx 0{,}992.$$

1.5. ELEMENTARNE POJĘCIA Z KOMBINATORYKI

Obliczanie prawdopodobieństw zdarzeń, będących podzbiorami skończonej przestrzeni zdarzeń elementarnych, często ułatwiają pojęcia i twierdzenia z kombinatoryki. W tym paragrafie przypomnimy elementarne pojęcia i twierdzenia z tej tematyki.

1.5.1. Reguła iloczynu. W rozwiązywaniu zagadnień kombinatorycznych wielokrotnie stosuje się następującą *regułę iloczynu*: jeśli pewną czynność wykonuje się w k-etapach, przy czym: etap 1 można wykonać n_1 sposobami, etap 2 — n_2 sposobami, ..., wreszcie k-ty etap — n_k sposobami, to liczba N sposobów, jakimi można wykonać tę czynność, wyraża się wzorem:

$$N=n_1 n_2 \dots n_k. \tag{1.5.1}$$

1.5.2. Permutacje bez powtórzeń. Niech A oznacza dowolny zbiór n różnych elementów, $A=\{a_1, \dots, a_n\}$. *Permutacją bez powtórzeń n-elementowego zbioru A* (lub *permutacją bez powtórzeń n różnych elementów*) nazywamy każdy n-wyrazowy ciąg, w którym każdy element zbioru A występuje dokładnie jeden raz. Jest to więc po prostu uporządkowanie elementów danego zbioru według jakiegoś kryterium. Na przykład ciągi (1, 2, 3), (3, 2, 1) są permutacjami bez powtórzeń zbioru $A=\{1, 2, 3\}$.

Metodą indukcji zupełnej dowodzi się, że liczba P_n wszystkich możliwych permutacji bez powtórzeń n-elementowego zbioru A wynosi:

$$P_n=n!, \tag{1.5.2}$$

gdzie z definicji $n!=n(n-1)(n-2)\dots 2\cdot 1$ oraz $0!=1$.

1.5.3. Permutacje z powtórzeniami. Niech A będzie zbiorem k różnych elementów, $A=\{a_1, \dots, a_k\}$. *Permutacją n-elementową z powtórzeniami*, w której element a_1 powtarza się n_1 razy, ..., element a_k powtarza się n_k razy, $n_1+\dots+n_k=n$, nazywamy każdy n-wyrazowy ciąg, w którym poszczególne elementy zbioru A powtarzają się wskazaną liczbę razy.

Liczba $P_n^{n_1,\dots,n_k}$ wszystkich takich n-wyrazowych permutacji z powtórzeniami jest dana równością:

$$P_n^{n_1,\dots,n_k}=\frac{n!}{n_1!\dots n_k!}. \tag{1.5.3}$$

1.5.4. Wariacje bez powtórzeń. Niech A będzie zbiorem n różnych elementów. Każdy k-wyrazowy ciąg k różnych elementów tego zbioru, $k\leqslant n$, nazywamy *k-wyrazową wariacją bez powtórzeń z n-elementowego zbioru A*. Na przykład 6-cyfrowy numer telefonu o niepowtarzających się cyfrach – 637254 – jest przykładem 6-wyrazowej wariacji bez powtórzeń zbioru cyfr $A=\{0, 1, \dots, 9\}$.

W przypadku, gdy $k=n$, to k-wyrazowa wariacja bez powtórzeń zbioru n-elementowego jest po prostu permutacją bez powtórzeń tego zbioru.

Liczba V_n^k wszystkich k-wyrazowych wariacji bez powtórzeń n-elementowego zbioru wyraża się wzorem:

$$V_n^k=n(n-1)(n-2)\dots(n-k+1), \quad k\leqslant n. \tag{1.5.4}$$

Dla wygody iloczyn ten będziemy oznaczali symbolem $n^{[k]}$:

$$n^{[k]}=n(n-1)(n-2)\dots(n-k+1); \tag{1.5.5}$$

na przykład

$$5^{[3]}=5\cdot 4\cdot 3=60,$$

$$5^{[5]}=5\cdot 4\cdot 3\cdot 2\cdot 1=5!=120.$$

Wzór (1.5.4) w tej symbolice przyjmuje następującą, skondensowaną postać

$$V_n^k=n^{[k]}. \tag{1.5.4'}$$

1.5.5. Wariacje z powtórzeniami. Niech A będzie zbiorem n-elementów. Każdy k-wyrazowy ciąg (mogących się powtarzać) elementów tego zbioru, $k\lesseqgtr n$, nazywamy *k-wyrazową wariacją z powtórzeniami z n-elementowego zbioru A*. Na przykład 55322 jest 5-wyrazową wariacją z powtórzeniami ze zbioru cyfr.

Liczba $\bar{V}_n^k$ wszystkich k-wyrazowych wariacji z powtórzeniami z n-elementowego zbioru A wyraża się wzorem:

$$\bar{V}_n^k=n^k. \tag{1.5.6}$$

1.5.6. Kombinacje bez powtórzeń. Niech A będzie zbiorem n różnych elementów. Każdy k-elementowy ($k\leqslant n$) podzbiór zbioru A nazywamy *k-elementową kombinacją bez powtórzeń n-elementowego zbioru A* (lub *kombinacją z n elementów po k elementów*).

Na przykład zbiór $\{a, c\}$, jako podzbiór $\{a, b, c, d\}$, jest 2-elementową kombinacją 4-elementowego zbioru $\{a, b, c, d\}$. Kolejność elementów jest tu nieistotna. Dlatego dwie k-elementowe kombinacje bez powtórzeń tego samego zbioru są różne wtedy i tylko wtedy, gdy różnią się składem (nie kolejnością) chociaż jednego elementu.

Liczba C_n^k wszystkich k-elementowych kombinacji bez powtórzeń n-elementowego zbioru wyraża się równością:

$$C_n^k = \binom{n}{k} = \frac{n^{[k]}}{k!}. \tag{1.5.7}$$

1.5.7. Kombinacje z powtórzeniami. Rozważamy elementy n różnych rodzajów. Elementy tego samego rodzaju traktujemy jako identyczne. Każdy zbiór (zestaw) składający się z k elementów, $k \lesseqgtr n$, gdy każdy element należy do jednego z tych n rodzajów, nazywamy *k-elementową kombinacją z powtórzeniami z n rodzajów elementów.* Kombinacja z powtórzeniami jest w pełni określona przez podanie liczby elementów poszczególnych rodzajów wchodzących w jej skład. Jest przy tym bez znaczenia, które elementy poszczególnych rodzajów wchodzą do jej składu, ponieważ nie rozróżnia się elementów tego samego rodzaju. Na przykład z elementów trzech rodzajów ($n=3$), gdy elementy poszczególnych rodzajów oznaczymy przez a, b, c, można utworzyć następujące 2-elementowe kombinacje z powtórzeniami

$$\{a,a\}, \quad \{b,b\}, \quad \{c,c\}, \quad \{a,b\}, \quad \{a,c\}, \quad \{b,c\}.$$

Liczba $\bar{C}_n^k$ wszystkich możliwych k-elementowych kombinacji z powtórzeniami z elementów n rodzajów jest równa liczbie k-elementowych kombinacji bez powtórzeń z $n+$ $+k-1$ elementów:

$$\bar{C}_n^k = C_{n+k-1}^k = \binom{n+k-1}{k} = \frac{(n+k-1)^{[k]}}{k!}. \tag{1.5.8}$$

1.6. ZADANIA DO ROZWIĄZANIA

1.23. Uzupełnić równości:

a) $A \cup A = \dots$; b) $AA = \dots$; c) $A \cup \Omega = \dots$; d) $A\Omega = \dots$.

1.24. Uzupełnić implikacje:

a) $\Omega \subset A \Rightarrow A = \dots$; b) $A \subset \emptyset \Rightarrow A = \dots$.

1.25. Zdarzenie A pociąga zdarzenie B, czyli $A \subset B$. Uzupełnić równości:

a) $A \cup B = \dots$; b) $AB = \dots$; c) $A \backslash B = \dots$; d) $ABC = \dots$.

1.26. Uzupełnić równoważności:

a) $AB = A \Leftrightarrow \dots$; b) $A \cup B = B \Leftrightarrow \dots$; c) $A \backslash B = \emptyset \Leftrightarrow \dots$; d) $A' \cup B = \Omega \Leftrightarrow \dots$.

1.27. Uzupełnić równoważności:

a) $AB = A \cup B \Leftrightarrow A = \dots$; b) $(AB = \emptyset \wedge A \cup B = \Omega) \Leftrightarrow B = \dots$;

c) $(A \backslash B) \cup (B \backslash A) = \emptyset \Leftrightarrow A = \dots$; d) $A \backslash B = A \Leftrightarrow AB = \dots$.

1.28. Wyrazić koniunkcję AB za pomocą zdarzenia A i różnicy $A\backslash B$.

1.29. Wyrazić alternatywę $A \cup B$ za pomocą zdarzenia A i różnicy $B\backslash A$.

1.30. Wyrazić różnicę $A\backslash B$ zdarzeń A, B za pomocą zdarzenia A i koniunkcji AB.

1.31. Zdarzenie E jest opisane za pomocą zdarzeń A, B, C. Uprościć ten opis, jeśli:

a) $E=(A\cup B)(A'\cup B)$; b) $E=(A\cup B)C\cup(B\cup C)A\cup(C\cup A)B$;

c) $E=(A\cup B)(B\cup C)(C\cup A)$; d) $E=(A\cup B)(B\cup C)$, gdy $A\subset B$;

e) $E=(A\cup B)(B\cup C)$, gdy $A\subset C$; f) $E=(A\cup B)(B\cup C)$, gdy $A\subset B\subset C$.

1.32. Jakie warunki muszą spełniać zdarzenia A i B, aby

$$(A\cup B)(A'\cup B)(A\cup B')=\emptyset?$$

1.33. Czy możliwe są trzy zdarzenia A, B, C takie, aby

$$AB\neq\emptyset\wedge AC=\emptyset\wedge(AB)\backslash C=\emptyset?$$

1.34. Podać przykłady (różne od podanych w rozwiązaniu zadania 1.1) doświadczeń losowych opisanych: a) skończoną, b) przeliczalną, c) nieprzeliczalną przestrzenią zdarzeń elementarnych.

1.35. Opakowanie zawiera 4 sztuki towaru. W użytkowaniu każda z tych sztuk może spełniać zakładane wymagania albo tych wymagań nie spełniać. Niech doświadczenie polega na obserwacji zachowania się w użytkowaniu wszystkich czterech sztuk.

a) Określić sensownie przestrzeń zdarzeń elementarnych.

Niech $A_i=1, 2, 3, 4$, oznacza zdarzenie: „i-ta sztuka spełnia zakładane wymagania”.

b) za pomocą zdarzeń elementarnych $\omega\in\Omega$ opisać zdarzenie A_2';

c) za pomocą zdarzeń A_i opisać zdarzenie B_i, jeśli:

B_1=zaszły wszystkie cztery zdarzenia A_1, A_2, A_3, A_4;

B_2=nie zaszło żadne ze zdarzeń A_1, A_2, A_3, A_4;

B_3=zaszło dokładnie jedno spośród zdarzeń A_1, A_2, A_3, A_4;

B_4=zaszło co najmniej jedno spośród zdarzeń A_1, A_2, A_3, A_4;

B_5=zaszło co najwyżej jedno spośród zdarzeń A_1, A_2, A_3, A_4.

d) Opisać zdarzenia B_i, określone w c), za pomocą zdarzeń elementarnych ω.

1.36. Przy projektowaniu przepustu odprowadzającego wodę z dwóch oddzielnych obszarów A, B zakłada się, że ilość wody przechodząca z obszaru A może wynieść: 0, 300, 600, 900 dm³/s, z obszaru zaś B – 0, 500, 1000, 1500 dm³/s.

a) Zaproponować sensownie i naszkicować przestrzenie zdarzeń elementarnych Ω_1, Ω_2, Ω_3 dla doświadczeń D_1, D_2, D_3 polegających odpowiednio na:

a_1) obserwacji ilości wody pochodzącej tylko z obszaru A,

a_2) obserwacji ilości wody pochodzącej tylko z obszaru B,

a_3) obserwacji ilości wody pochodzącej łącznie z obydwu obszarów.

b) Naszkicować i opisać za pomocą zdarzeń elementarnych, pochodzących z przestrzeni zdarzeń elementarnych Ω_1 albo Ω_2, albo Ω_3, następujące zdarzenia:

b_1) C=ilość wody pochodzącej z obszaru A przekroczy 300 dm³/s;

b_2) E=ilość wody pochodzącej z obszaru B wyniesie co najmniej 1000 dm³/s;

b_3) $F=C'$; $G=CE$; $H=(C\cup E)'$.

1.37. Inżynier projektuje magazyn do przechowywania kartonów puszek żywności. Kartony mają kształt sześcianów o krawędzi 4 dm i masie 50 kg każdy. Zakłada się, że kartony mogą być układane losowo, ale najwyżej do wysokości 24 dm.

a) Sensownie zaproponować i naszkicować przestrzeń zdarzeń elementarnych Ω_1 dla doświadczenia D_1 polegającego na obserwacji całkowitego obciążenia 16 dm^2 powierzchni podłogi, zakładając, że jest ono wywołane przez jeden stos kartonów.

Naszkicować i opisać za pomocą zdarzeń elementarnych pochodzących z Ω_1 następujące zdarzenia:

A = całkowite obciążenie będzie $\geqslant 150$ kg;
B = całkowite obciążenie będzie $\leqslant 200$ kg;
C = całkowite obciążenie będzie > 250 kg;

$$E = AB; \quad F = B \cup C.$$

b) Sensownie zaproponować i naszkicować przestrzeń zdarzeń elementarnych Ω_2 dla doświadczenia D_2 polegającego na obserwacji całkowitego obciążenia 16 dm^2 powierzchni podłogi przy założeniu, że jest to obciążenie wywołane przez połowę masy każdego z dwu sąsiednich stosów.

Naszkicować i opisać za pomocą zdarzeń elementarnych pochodzących z Ω_2 zdarzenia C, E opisane w punkcie a).

1.38. Niech doświadczenie D polega na obserwacji liczby pojazdów i ich łącznej masy znajdujących się w określonej chwili na moście. Zakładamy, że maksymalna liczba pojazdów, jaką można stwierdzić na tym moście, wynosi 3, maksymalna zaś masa pojedynczego pojazdu wynosi 5 ton, minimalna – 2 tony (masę obserwuje się z dokładnością do 1 tony).

Zaproponować i naszkicować przestrzeń elementarnych zdarzeń Ω opisującej to doświadczenie. Następnie naszkicować i opisać za pomocą zdarzeń elementarnych tej przestrzeni następujące zdarzenia:

A – liczba pojazdów będzie mniejsza niż 2 pojazdy,
B – całkowita masa pojazdów wyniesie co najmniej 10 ton,
C – zaobserwujemy 2 pojazdy o maksymalnej masie,

$$E = BC, \quad F = B \cup C, \quad G = A'B'.$$

1.39. Rozważamy następujące doświadczenie D. Drewniane pale mają losową długość L, przy tym największa długość wynosi 12 m. Pale są przeznaczone do wbijania w ziemię, której skalna warstwa stanowiąca opór, znajduje się na losowej głębokości H, której maksimum wynosi 10 m.

Zaproponować i naszkicować przestrzeń zdarzeń elementarnych dla tego doświadczenia. Następnie zilustrować graficznie następujące zdarzenia:

A – długość losowo wziętego pala będzie większa od głębokości skalnej warstwy,
B – głębokość skalnej warstwy przekroczy 8 m,
C – długość losowo wziętego pala przekroczy 8 m,

$$E = BC, \quad F = B \cup C, \quad G = (B \cup C)A'.$$

1.40. Na szczyt góry prowadzi pięć dróg. Każda z nich nadaje się również do zejścia. Zakładamy ponadto, że wszystkie trasy są równorzędne. Obliczyć prawdopodobieństwo spotkania się dwóch znajomych, z których jeden wchodzi na szczyt, a drugi jest już w drodze powrotnej.

1.41. Pięciu studentów powtarzających dany rok studiów wybiera losowo, każdy niezależnie od pozostałych, jedną z trzech równoległych grup. Zakładając, że wszystkie rozmieszczenia tych studentów są jednakowo prawdopodobne, znaleźć prawdopodobieństwo tego że: a) wszyscy znajdą się w pierwszej grupie, b) wszyscy znajdą się w tej samej grupie, c) w pierwszej grupie znajdzie się dokładnie jeden student, d) w jednej z grup znajdzie się dokładnie jeden student, e) w ustalonej grupie znajdzie się dokładnie trzech studentów.

1.42. Rozwiązać zadanie 1.15, korzystając z kombinatoryki.

1.43. Na dziesięciu klockach wyrzeźbiono litery: a, a, k, s, s, t, t, t, y, y. Bawiąc się nimi dziecko układa je w rząd. Obliczyć prawdopodobieństwo, że przypadkowo złoży ono słowo „statystyka".

1.44. Z dziesięciu pracowników należy utworzyć: a) dwa zespoły liczące po 4 i 6 pracowników, b) trzy zespoły liczące po 5, 3 i 2 pracowników, c) pięć zespołów dwuosobowych. Dla każdego podziału na zespoły znaleźć prawdopodobieństwo tego, że dwóch ustalonych pracowników znajdzie się w tym samym zespole przy założeniu, że podział na zespoły odbywa się losowo.

1.45. Z partii N sztuk towaru, wśród których jest M sztuk zgodnych z normą losujemy n sztuk a) bez zwrotu, b) ze zwrotem. Obliczyć prawdopodobieństwo, że wśród nich znajdzie się k sztuk zgodnych z normą.

1.46. Wśród 10 sztuk towaru 5 odpowiada wymaganiom międzynarodowego znaku jakości, 3 ma usterki w opakowaniu, a 2 ma inne usterki. Znaleźć prawdopodobieństwo, że wśród trzech wylosowanych sztuk: a) znajdą się 2 z tej samej grupy jakości, b) wszystkie będą z różnych grup jakości, jeśli losowanie odbywa się bez zwrotu.

1.47. W fizyce statystycznej rozważa się rozkład (rozmieszczenie) k cząstek w n elementarnych obszarach zwanych *komórkami*. W zależności od postaci tych cząstek przyjmuje się jedno z trzech następujących założeń:

1° cząstki różnią się między sobą (zatem wzajemna zamiana komórek przez dwie cząstki daje nowy rozkład) i liczba cząstek w jednej komórce jest dowolna,

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

3° cząstki nie różnią się między sobą i w każdej komórce może znaleźć się co najwyżej jedna cząstka.

Stosownie do przyjętych założeń mówi się odpowiednio o statystyce: Maxwella-Boltzmanna, Bosego-Einsteina, Fermiego-Diraca. Zakładamy, ponadto, że wszystkie dopuszczalne rozmieszczenia są jednakowo prawdopodobne. Znaleźć prawdopodobieństwo tego, że a) k cząstek rozmieści się po jednej w k ustalonych komórkach dla każdej z rozważanych tam statystyk, b) w przypadku statystyki Bosego-Einsteina znajdzie się dokładnie m cząstek, b_1) w ustalonej komórce, b_2) w jednej z n komórek, c) w przypadku statystyki

Bosego-Einsteina wszystkie komórki będą zajęte, d) w przypadku statystyki Maxwella-Boltzmanna w pierwszej komórce znajdzie się dokładnie k_1 cząstek, w drugiej – k_2 cząstek, ..., w n-tej – k_n cząstek.

1.48. Spośród czcionek a, a, k, s, s, t, t, t, y, y losujemy bez zwrotu po jednej czcionce i układamy je w kolejności losowania w słowa. Obliczyć prawdopodobieństwo otrzymania w ten sposób słowa: a) as, b_1) akt, b_2) kat, b_3) syk, c_1) atak, c_2) kasa, c_3) styk, c_4) takt, c_5) tata, c_6) tyka, d_1) kasta, d_2) taksa, d_3) tasak, e) asysta, f) statyka, g) statysta, h) statystyk, i) statystyka.

1.49. Każda praca pisemna z egzaminu wstępnego jest sprawdzana dwukrotnie. Prawdopodobieństwo niezauważenia błędu przez pierwszą osobę sprawdzającą wynosi 0,08. Dla drugiej osoby prawdopodobieństwo to wynosi 0,05. Obliczyć prawdopodobieństwo, że błąd popełniony w pracy pisemnej nie zostanie zauważony.

1.50. Winda wyposażona jest w dwa układy hamowania włączające się automatycznie (obydwa) w razie zerwania się liny. Przy tym prawdopodobieństwo wyhamowania przez każdy układ z osobna jest jednakowe i wynosi 0,99. Jakie jest prawdopodobieństwo: a) wyhamowania windy w razie zerwania się liny, b) spadnięcie kabiny windy w razie zerwania się liny, jeśli prawdopodobieństwo tego ostatniego zdarzenia wynosi 10^{-5}?

1.51. Rozwiązać zadanie 1.50 przy założeniu, że drugi układ hamowania włączy się tylko w sytuacji, gdy nie zadziała pierwszy.

1.52. Cztery identyczne automaty produkują w tej samej ilości jednakowe wyroby i dostarczają je na przenośnik taśmowy. Przy przejmowaniu wyrobów z przenośnika zauważono, że czwarta ich część ma usterkę spowodowaną rozregulowaniem się automatu. Zakładając, że równoczesne rozregulowanie się dwóch lub więcej automatów jest praktycznie niemożliwe, znaleźć prawdopodobieństwo tego, że przy poszukiwaniu automatu rozregulowanego trzeba będzie sprawdzić wszystkie automaty.

1.53. Kodowa informacja składa się z siedmiu impulsów postaci A, B, C odpowiednio w ilościach: 4, 2, 1. Znaleźć prawdopodobieństwo tego, że: a) pierwszym odebranym impulsem będzie A, b) pierwszym odebranym impulsem będzie A albo C, c) dwoma pierwszymi impulsami będą A, C, d) trzema pierwszymi impulsami będą A, B, C. Zakładamy, że nie ma zakłóceń (zad. 1.22).

1.54. Zdarzenia A_1, A_2, A_3, A_4 są (wzajemnie) niezależne i $P(A_k)=p_k$. Obliczyć prawdopodobieństwo zajścia co najmniej jednego z tych zdarzeń.

1.55. Produkowany wyrób może być zaklasyfikowany z jednakowym prawdopodobieństwem do jednej z trzech klas: I, II, III. Określamy zdarzenia – wylosowany wyrób będzie: I albo II klasy (zdarzenie A), I albo III klasy (zdarzenie B), II albo III klasy (zdarzenie C). Zbadać czy zdarzenia A, B, C są: a) niezależne parami, b) niezależne zespołowo.

1.56. Niech $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4\}$ będzie przestrzenią zdarzeń elementarnych pewnego doświadczenia. Zakładając, że zdarzenia jednoelementowe są jednakowo prawdopodobne, zbadać, czy zdarzenia $A=\{\omega_1, \omega_4\}$, $B=\{\omega_2, \omega_4\}$, $C=\{\omega_3, \omega_4\}$ są: a) niezależne parami, b) niezależne (zespołowo)?

1.57. Czy prawdziwe jest zdanie: jeśli dwa zdarzenia wykluczają się, to są one zależne?

1.58. Wykazać, że gdy: a) zdarzenia A, B są niezależne, wtedy niezależne są również zdarzenia im przeciwne A' i B', b) $P(B)>0$ i $A\subset B$, wtedy $P(A \mid B)=P(A)/P(B)$.

1.59. Oznaczmy $P(AB)-P(A)P(B)=r(A, B)$. Różnica ta może być dodatnia, ujemna albo może równać się zeru. Wartość bezwzględną tej różnicy można przyjąć za miarę zależności między zdarzeniami A i B.

a) Wykazać, że $r(A, B)=0$ wtedy i tylko wtedy, gdy zdarzenia A i B są niezależne.

b) Niech $P(B)>0$. Wykazać, że gdy: b_1) $r(A, B)=0$, wtedy $P(A \mid B)=P(A)$; b_2) $r(A, B)<0$, wtedy $P(A \mid B)<P(A)$ (w tym przypadku mówimy, że zdarzenie B ma negatywny wpływ na zajście zdarzenia A); b_3) $r(A, B)>0$, wtedy $P(A \mid B)>P(B)$ (zdarzenie B ma pozytywny wpływ na zajście zdarzenia A).

c) Obliczyć $r(A_1, A_2)$ dla zdarzeń A_1 i A_2 z zadania 1.13.

d) Obliczyć: d_1) $r(A, A)$ i $\max r(A, A)$, d_2) $r(A, A')$ i $\min r(A, A')$.

1.60. Chcemy rozpalić ognisko mając do dyspozycji tylko dwie zapałki. Wybrać bardziej pewną metodę z dwu następujących: 1° próbujemy rozpalić najpierw jedną, potem drugą zapałką, 2° próbujemy rozpalić dwiema złączonymi zapałkami, jeśli prawdopodobieństwo rozpalenia ogniska pojedynczą zapałką wynosi 0,7; natomiast złączonymi — 0,95.

1.61. Dla pewnego obszaru prawdopodobieństwo trwania przez jedną minutę umiarkowanego trzęsienia ziemi (zdarzenie A) wynosi 10^{-7}, a prawdopodobieństwo wystąpienia w ciągu minuty silnego wiatru (zdarzenie B) wynosi 10^{-5}. Wyjaśnić dlaczego przepisy budowlane obowiązujące na tym obszarze nie wymagają przy projektowaniu budynków uwzględnienia łącznego wystąpienia tych dwu zjawisk?

1.62. Wyjaśnić dlaczego w praktycznych zagadnieniach, gdy prawdopodobieństwa $P(A)$ i $P(B)$ są małe, można przyjąć, że

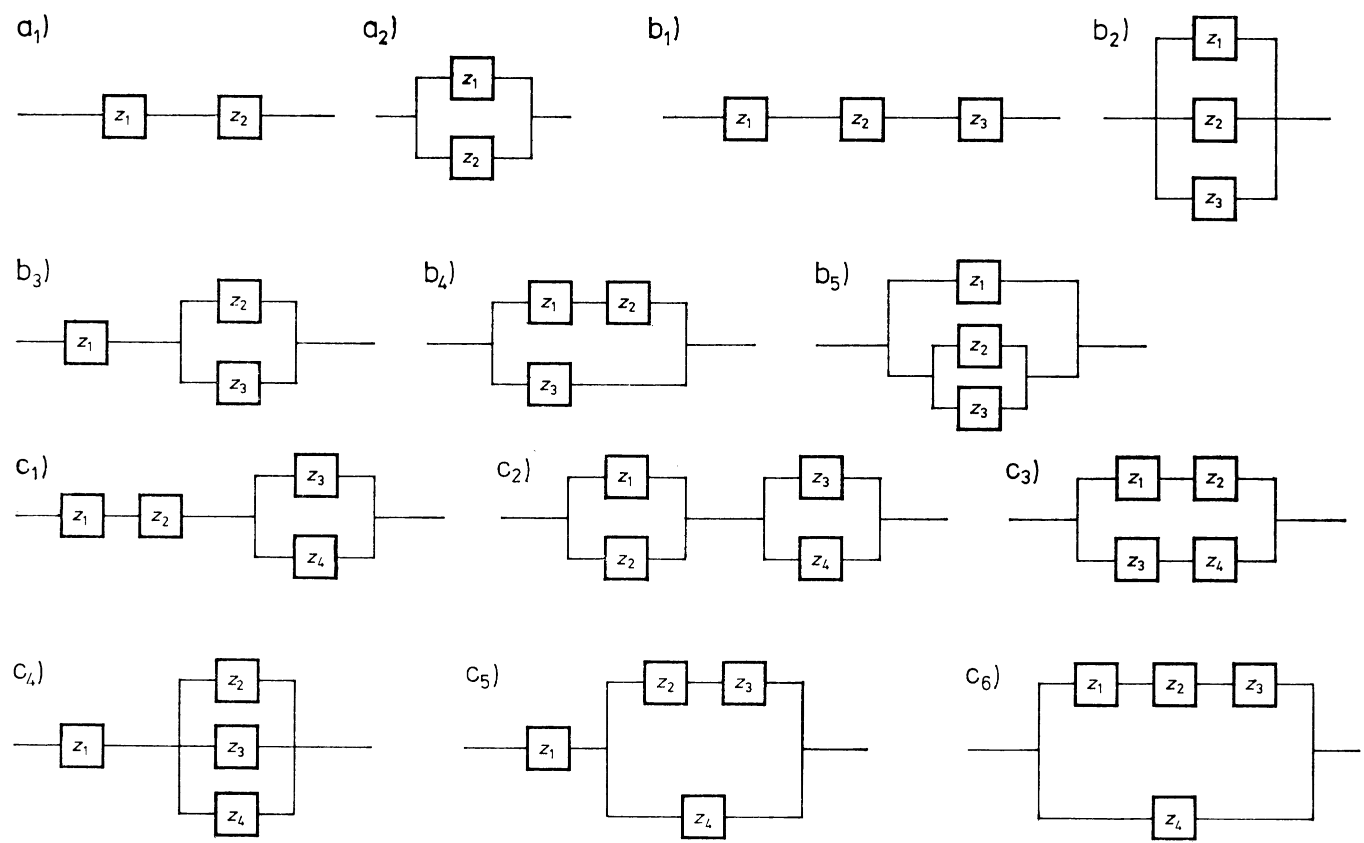
$$P(A \cup B)=P(A)+P(B)?$$

1.63. Na rysunku 1.10, gdzie z_1, z_2, z_3, z_4 oznaczają żarówki, dane są schematy fragmentów sieci elektrycznych. Prawdopodobieństwo nieprzepalania się w czasie t godzin jest dla wszystkich żarówek jednakowe i wynosi p. Zakładając, że żarówki przepalają się niezależnie od siebie, obliczyć prawdopodobieństwo ciągłego przepływu prądu w czasie t dla każdego fragmentu sieci oraz porównać te prawdopodobieństwa w ramach układów złożonych z dwóch albo trzech żarówek i dwóch układów złożonych z czterech żarówek.

1.64. Pewien towar produkują 3 zakłady. Prawdopodobieństwo wyprodukowania przez te zakłady towaru pierwszej jakości wynosi odpowiednio 0,97, 0,90, 0,86. Znaleźć prawdopodobieństwo tego, że losowo wzięta sztuka towaru — spośród trzech sztuk pochodzących z różnych zakładów — jest pierwszej jakości.

1.65. Spośród trzech równorzędnych kandydatów należy wybrać przewodniczącego studenckiej grupy działania. W tym celu na jednej z trzech czystych kartek piszemy słowo „przewodniczący" i wrzucamy je do pudełka. Następnie kandydaci kolejno losują jedną kartkę i ten, który wylosuje „przewodniczącego" zostaje przewodniczącym. Obliczyć prawdopodobieństwo wylosowania „przewodniczącego" przez kandydata losującego jako: a) pierwszy, b) drugi, c) trzeci.

1.66. Zakłady radiowe są zaopatrywane w lampy radiowe tylko przez dwu kooperantów K_1 i K_2. Pierwszy z nich pokrywa zaopatrzenie zakładu w siedemdziesięciu procentach.



Rys. 1.10. Do zadania 1.63

Poza tym wiadomo, że kooperanci produkują średnio 90% i 80% lamp o małym rozrzucie parametrów (tj. dobrych lamp). Znaleźć prawdopodobieństwo tego, że losowo wzięta lampa: a) jest wyprodukowana przez kooperanta K_1, b) jest wyprodukowana przez kooperanta K_1 i charakteryzuje się przy tym małym rozrzutem parametrów, c) charakteryzuje się małym rozrzutem parametrów, d) która okazała się dobrą, pochodzi od pierwszego kooperanta.

1.67. Na przenośnik taśmowy trafiają jednakowe wyroby wytwarzane przez 3 automaty. Stosunek ilościowy produkcji automatów kształtuje się tak jak 2 : 2 : 1. Poza tym wiadomo, że automat pierwszy produkuje 85% wyrobów I gatunku, drugi – 80% I gatunku, trzeci – 90% I gatunku. Obliczyć prawdopodobieństwo tego, że losowo wzięty z przenośnika wyrób: a) wyprodukowany został przez drugi automat, b) jest wyrobem I gatunku wyprodukowanym przez drugi automat, c) jest wyrobem I gatunku, d) który okazał się I gatunku, jest wyprodukowany przez drugi automat.

1.68. Koparka może pracować w warunkach normalnych albo trudnych odpowiednio z prawdopodobieństwami $p_1=0{,}8$ i $p_2=0{,}2$. Prawdopodobieństwo awarii koparki w czasie t wynosi 0,05 przy pracy w warunkach normalnych i 0,25 w warunkach trudnych. a) Ile wynosi prawdopodobieństwo awarii koparki pracującej przez czas t? b) W ciągu czasu t koparka uległa uszkodzeniu, obliczyć prawdopodobieństwo tego, że pracowała wtedy w warunkach normalnych.

1.69. Wykonujemy pomiary trzema przyrządami, z których jeden jest nieco rozregulowany. Przy wykonywaniu pomiaru sprawnym przyrządem prawdopodobieństwo otrzymania błędu pomiaru przewyższającego tolerancję, wynosi 0,03; prawdopodobieństwo to dla przyrządu niesprawnego wynosi 0,3. Znaleźć prawdopodobieństwo tego, że wynik pomiaru losowo wziętym przyrządem: a) przewyższa tolerancję, b) który przewyższał tolerancję, jest wykonany nie w pełni sprawnym przyrządem.

1.70. Prawdopodobieństwo tego, że przy danym reżimie technologicznym wyprodukowany wyrób odpowiada wymaganiom normy wynosi 0,9. Każdy wyrób podlega dwuetapowej kontroli i opuszcza zakład tylko wtedy, gdy przejdzie z wynikiem pozytywnym oba etapy. Obliczyć prawdopodobieństwo tego, że wyrób przeznaczony do sprzedaży odpowiada wymaganiom normy, gdy prawdopodobieństwo pozytywnego przejścia przez poszczególne etapy kontroli jest takie samo dla obu etapów i wynosi 0,95 dla wyrobu odpowiadającego wymaganiom normy i 0,1 dla wyrobu niezgodnego z tymi wymaganiami.

1.71. Kanałem łączności nadaje się tylko 3 rodzaje ciągów liter: AAAA, BBBB, CCCC odpowiednio z prawdopodobieństwami 0,4; 0,3; 0,3. Litery te (sygnały) podlegają niezależnie losowym zakłóceniom (przekłamaniom) w rezultacie czego np. litera A może być odebrana jako B albo C (zamiast A). Prawdopodobieństwa poprawnego przesłania albo przekłamania podaje tablica.

Sygnały nadane	Sygnały odebrane		
	A	B	C
A	0,8	0,1	0,1
B	0,1	0,8	0,1
C	0,1	0,1	0,8

a) Znaleźć prawdopodobieństwo odebrania na wyjściu sygnału: a_1) AAAA, a_2) ACAA, a_3) CCCC, a_4) ABCA. b) Na wyjściu odebrano sygnał ACAA, obliczyć prawdopodobieństwo, że został on nadany jako AAAA. c) Na wyjściu odebrano sygnał AAAA, obliczyć prawdopodobieństwo, że został on wysłany również jako AAAA.

Odpowiedzi

1.23. a) A; b) A; c) Ω; d) A. **1.24.** a) Ω; b) $\emptyset$.

1.25. a) B; b) A; c) $\emptyset$; d) AC.

1.26. a) $A \subset B$; b) $A \subset B$; c) $A \subset B$; d) $A \subset B$.

1.27. a) B; b) A'; c) B; d) $\emptyset$. **1.28.** $AB = A \backslash (A \backslash B)$.

1.29. $A \cup B = A \cup (B \backslash A)$. **1.30.** $A \backslash B = A \backslash AB$.

1.31. a) B; b) $AB \cup AC \cup BC$; c) $AB \cup AC \cup BC$; d) B; e) $A \cup B$; f) B.

1.32. $AB = \emptyset$.

1.33. Nie. Rozwiązanie. $AB \neq \emptyset \Rightarrow$ istnieje zdarzenie elementarne $\omega \in AB$. Ponieważ $AC = \emptyset$, więc $\omega \notin AC$. Zatem $\omega \in (AB \backslash C)$ i tym samym $AB \backslash C \neq \emptyset$.

1.35. a) Np. $\Omega = \{\omega_1, \ldots, \omega_{16}\}$, gdzie $\omega_i = (i_1, i_2, i_3, i_4)$, $i_j = 1$, gdy j-ty detal, $j = 1, 2, 3, 4$, spełnia zakładane wymagania oraz $i_j = 0$ w przypadku przeciwnym. Oczywiście $n(\Omega) = 2^4 = 16$.

b) $A_2' = \{(0,0,0,0), (0,0,0,1), (0,0,1,0), (0,0,1,1), (1,0,1,0), (1,0,1,1), (1,0,0,0), (1,0,0,1)\}$;

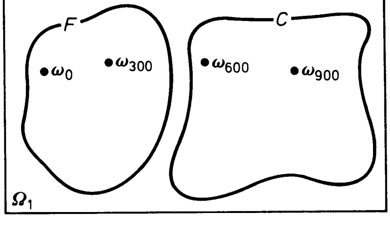
c) $B_1 = A_1 A_2 A_3 A_4$, $B_2 = A_1' A_2' A_3' A_4'$, $B_3 = A_1 A_2' A_3' A_4' \cup A_1' A_2 A_3' A_4' \cup A_1' A_2' A_3 A_4' \cup A_1' A_2' A_3' A_4$, $B_4 = A_1 \cup A_3 \cup A_3 \cup A_4$, $B_5 = A_1' A_2' A_3' A_4' \cup B_3$;

d) $B_1 = \{(1,1,1,1)\}$, $B_2 = \{(0,0,0,0)\}$, $B_3 = \{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)\}$, $B_4 = \Omega - \{(0,0,0,0)\}$, $B_5 = \{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0\ 1), (0,0,0,0)\}$.

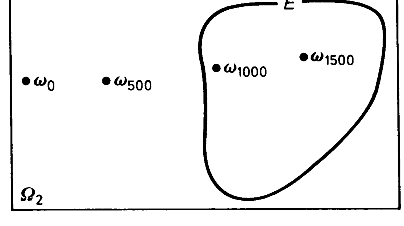
1.36. a) $\Omega_1 = \{\omega_0, \omega_{300}, \omega_{600}, \omega_{900}\}$, gdzie ω_i, $i = 0, 300, 600, 900$, oznacza zdarzenie elementarne: „ilość wody pochodząca z obszaru A wyniesie i dm³/s, (rys. 1.11);

$$\Omega_2 = \{\omega_0, \omega_{500}, \omega_{1000}, \omega_{1500}\} \quad \text{(rys. 1.12)};$$

$$\Omega_3 = \{\omega_{i,j}\}, \quad i = 0, 300, 600, 900, \quad j = 0, 500, 1000, 1500,$$



Rys. 1.11. Do zadania 1.36

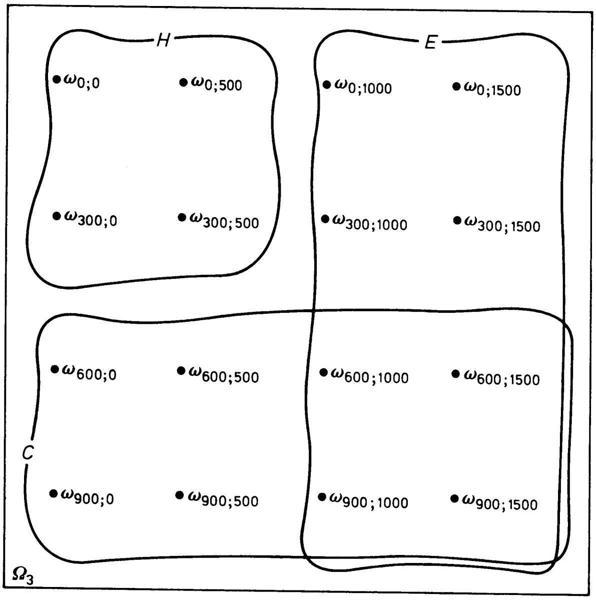


Rys. 1.12. Do zadania 1.36

gdzie ω_{ij} oznacza zdarzenie elementarne: „ilość wody pochodząca z obszaru A wyniesie i dm^3/s" oraz „ilość wody pochodząca z obszaru B w tej samej sekundzie wyniesie j dm^3/s" (rys. 1.13).

b) $C=\{\omega_{600}, \omega_{900}\}$ w przypadku doświadczenia D_1 i przestrzeni zdarzeń elementarnych Ω_1 lub

$C=\{\omega_{600;\,0}, \omega_{600;\,500}, \omega_{600;\,1000}, \omega_{600;\,1500}, \omega_{900;\,0}, \omega_{900;\,500}, \omega_{900;\,1000}, \omega_{900;\,1500}\}$ w przypadku doświadczenia D_3 i przestrzeni zdarzeń elementarnych Ω_3;



Rys. 1.13. Do zadania 1.36

$E=\{\omega_{1000}, \omega_{1500}\}$ w przypadku doświadczenia D_2 i przestrzeni zdarzeń elementarnych Ω_2 lub

$E=\{\omega_{0;\,1000}, \omega_{0;\,1500}, \omega_{300;\,1000}, \omega_{300;\,1500}, \omega_{600;\,1000}, \omega_{600;\,1500}, \omega_{900;\,1000}, \omega_{900;\,1500}\}$ w przypadku doświadczenia D_3 i przestrzeni zdarzeń elementarnych Ω_3;

$F=\{\omega_0, \omega_{300}\}$ lub $F=\{\omega_{0;\,0}, \omega_{0;\,500}, \omega_{0;\,1000}, \omega_{0;\,1500}, \omega_{300;\,0}, \omega_{300;\,500}, \omega_{300;\,1000}, \omega_{300;\,1500}\}$;

$G=\{\omega_{600;\,1000}, \omega_{600;\,1500}, \omega_{900;\,1000}, \omega_{900;\,1500}\}$;

$H=C'E'=\{\omega_{0;\,0}, \omega_{0;\,500}, \omega_{300;\,0}, \omega_{300;\,500}\}$.

1.37. a) $\Omega_1=\{\omega_0, \omega_{50}, \omega_{100}, \omega_{150}, \omega_{200}, \omega_{250}, \omega_{300}\}$, gdzie ω_m oznacza zdarzenie elementarne: „całkowite obciążenie 16 dm^2 powierzchni podłogi wyniesie m kg"; $A=\{\omega_{150}, \omega_{200}, \omega_{250}, \omega_{300}\}$; $B=\{\omega_0, \omega_{50}, \omega_{100}, \omega_{150}, \omega_{200}\}$; $C=\{\omega_{300}\}$; $E=\{\omega_{150}, \omega_{200}\}$; $F=\Omega_1-\{\omega_{250}\}=\{\omega_0, \omega_{50}, \omega_{100}, \omega_{150}, \omega_{200}, \omega_{300}\}$.

b) $\Omega_2=\{\omega_{i,j}: i,j=0, 25, 50, 75, 100, 125, 150\}$, gdzie $\omega_{m,n}$ oznacza zdarzenie elementarne „obciążenie 16 dm^2 powierzchni podłogi pochodzące od pierwszego stosu kartonów wynosi m kg, od drugiego zaś – n kg";

$C=\{\omega_{150;\,150}\}$,

$E=\{\omega_{0;\,150}, \omega_{150;\,0}, \omega_{25;\,125}, \omega_{125;\,25}, \omega_{50;\,100}, \omega_{100;\,50}, \omega_{75;\,75}, \omega_{50;\,150}, \omega_{75;\,125}, \omega_{125;\,75}, \omega_{100;\,100}\}$.

1.38. $\Omega=\{\omega_0, \omega_2, \omega_3, \omega_4, \omega_5, \omega_{22}, \omega_{23}, \omega_{24}, \omega_{25}, \omega_{33}, \omega_{34}, \omega_{44}, \omega_{45}, \omega_{55}, \omega_{222}, \omega_{223}, \omega_{224}, \omega_{225}, \omega_{233}, \omega_{234}, \omega_{235}, \omega_{244}, \omega_{245}, \omega_{255}, \omega_{333}, \omega_{334}, \omega_{335}, \omega_{344}, \omega_{345}, \omega_{355}, \omega_{444}, \omega_{445}, \omega_{455}, \omega_{555}\}$, gdzie np. ω_{245} oznacza zdarzenie elementarne polegające na zaobserwowaniu 3 pojazdów o masie 2, 4, 5 ton (w dowolnej kolejności);

$$A=\{\omega_0, \omega_2, \omega_3, \omega_4, \omega_5\};$$

$$B=\{\omega_{55}, \omega_{235}, \omega_{244}, \omega_{245}, \omega_{255}, \omega_{334}, \omega_{335}, \omega_{344}, \omega_{345}, \omega_{355}, \omega_{444}, \omega_{445}, \omega_{455}, \omega_{555}\};$$

$$C=\{\omega_{55}, \omega_{255}, \omega_{455}\}; \quad E=C \text{ (ponieważ } C\subset B);$$

$$F=B \text{ (ponieważ } C\subset B);$$

$$G=\{\omega_{22}, \omega_{23}, \omega_{24}, \omega_{25}, \omega_{33}, \omega_{34}, \omega_{35}, \omega_{44}, \omega_{45}, \omega_{222}, \omega_{223}, \omega_{224}, \omega_{225}, \omega_{233}, \omega_{234}, \omega_{333}\}.$$

1.39. $\Omega=\{\omega=(L,H):\ 0<L\leqslant 12,\ 0\leqslant H\leqslant 10\}$, rys. 1.14a;

$A=\{\omega=(L,H):\ 0<L\leqslant 12,\ 0\leqslant H<\min(L,10)\}$, rys. 1.14b;

$B=\{\omega=(L,H):\ 0<L\leqslant 12,\ 8<H\leqslant 10\}$, rys. 1.14c;

$C=\{\omega=(L,H):\ 8<L\leqslant 12,\ 0\leqslant H\leqslant 10\}$, rys. 1.14d;

$E=\{\omega=(L,H):\ 8<L\leqslant 12,\ 8<H\leqslant 10\}$, rys. 1.14e;

$F=\{\omega=(L,H):\ (0<L\leqslant 12,\ 8<H\leqslant 10)\cup(8<L\leqslant 12,\ 0\leqslant H\leqslant 10)\}$, rys. 1.14f;

$G=\{\omega=(L,H):\ 0<L\leqslant H,\ 8<H\leqslant 10\}$, rys. 1.14g.

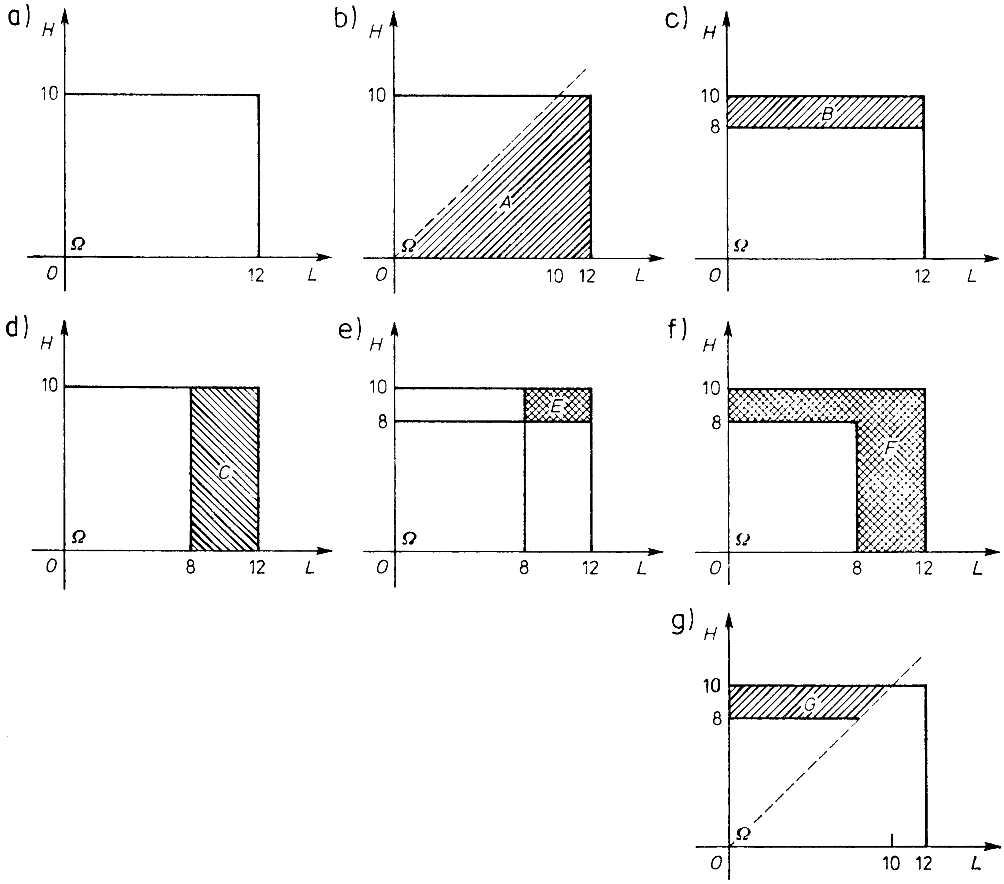
1.40. $\frac{5}{25}=\frac{1}{5}$. **1.41.** a) $\frac{1}{3^5}=\frac{1}{243}$; b) $\frac{3}{3^5}=\frac{1}{81}$; c) $\frac{5\cdot 2^4}{3^5}=\frac{80}{243}$;

d) $\frac{3\cdot 5\cdot 2^4}{3^5}=\frac{80}{81}$; e) $\frac{\binom{5}{3}\cdot 2^2}{3^5}=\frac{40}{241}$.

1.42. a) $$\frac{\binom{3}{1}\cdot V_8^2+1\binom{3}{1}\binom{7}{1}}{V_9^3}=\frac{3}{8};$$

b) $$\frac{\binom{k-1}{1}\cdot V_8^2+1\binom{k-1}{1}\binom{7}{1}}{V_9^3}=\frac{(k-1)\cdot 8\cdot 7+(k-1)\cdot 7}{9\cdot 8\cdot 7}=\frac{k-1}{8}.$$

1.43. $\frac{(2!)^3\cdot 3!}{10!}\approx 0{,}000013$.



Rys. 1.14. Do zadania 1.39

1.44. a) $\dfrac{\binom{2}{2}\binom{8}{2}+\binom{2}{2}\binom{8}{4}}{\binom{10}{4}}=\dfrac{7}{15}$; b) $\dfrac{\binom{2}{2}\binom{8}{3}+\binom{2}{2}\binom{8}{1}+\binom{2}{2}}{\binom{10}{5}\binom{5}{3}}=\dfrac{13}{9\cdot 8\cdot 7}\approx 0{,}026$;

c) $\dfrac{5}{\binom{10}{2}\binom{8}{2}\binom{6}{2}\binom{4}{2}}=\dfrac{5\cdot(2!)^5}{10!}=0{,}000088.$

1.45. a) $\dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$; b) $\dfrac{\binom{n}{k}M^k(N-M)^{n-k}}{N^n}=\binom{n}{k}\left(\dfrac{M}{N}\right)^k\left(1-\dfrac{M}{N}\right)^{n-k}.$

1.46. a) $\dfrac{\binom{5}{2}\binom{5}{1}+\binom{3}{2}\binom{7}{1}+\binom{2}{2}\binom{8}{1}}{\binom{10}{3}}=\dfrac{79}{120}$; b) $\dfrac{5\cdot 3\cdot 2}{\binom{10}{3}}=\dfrac{1}{4}$.

1.47. a) Odpowiednio

$$\frac{k!}{n^k},\quad 1\Big/\binom{n+k-1}{k},\quad 1\Big/\binom{n}{k};\quad \text{b}_1)\ \binom{n+k-m-2}{k-m}\Big/\binom{n+k-1}{k};$$

$$\text{b}_2)\ n\binom{n+k-m-2}{k-m}\Big/\binom{n+k-1}{k};\quad \text{c})\ \binom{k-1}{n-1}\Big/\binom{n+k-1}{k};\quad \text{d})\ \frac{k!}{n^k k_1!\,k_2!\ldots k_n!}.$$

1.48. Wszystkie prawdopodobieństwa obliczamy w ten sam sposób – korzystamy z prawdopodobieństwa koniunkcji n zdarzeń, $n=2,\ldots,10$. Np.: a) $\frac{2}{10}\cdot\frac{2}{9}=\frac{2}{45}$; b$_1$) $\frac{2}{10}\cdot\frac{1}{9}\cdot\frac{3}{8}=\frac{1}{120}$; b$_3$) $\frac{2}{10}\cdot\frac{2}{9}\cdot\frac{1}{8}=\frac{1}{180}$; c$_1$) $\frac{2}{10}\cdot\frac{3}{9}\cdot\frac{1}{8}\cdot\frac{1}{7}=\frac{1}{840}$; g) $\frac{2}{10}\cdot\frac{3}{9}\cdot\frac{2}{8}\cdot\frac{2}{7}\cdot\frac{2}{6}\cdot\frac{1}{5}\cdot\frac{1}{4}\cdot\frac{1}{3}=\frac{1}{37800}$.

1.49. $0{,}08\cdot 0{,}05=0{,}004$.

1.50. a) $0{,}99+0{,}99-0{,}99\cdot 0{,}99=0{,}9999$ albo inaczej: $1-0{,}01\cdot 0{,}01=0{,}9999$; b) $0{,}01\cdot 0{,}01\cdot 10^{-5}=10^{-9}$.

1.51. a) Niech W_i, $i=1,2$ oznacza zdarzenie: winda zostanie wyhamowana przez i-ty układ hamowania, $P(W_1\cup W_1'W_2)=0{,}99+0{,}01\cdot 0{,}99=0{,}9999$; b) $10^{-5}\cdot 0{,}01\cdot 0{,}01=10^{-9}$.

1.52. Niech A_i oznacza zdarzenie: i-ty z kolei sprawdzony automat jest sprawny. Wówczas $P(A_1A_2A_3)=\frac{3}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}=\frac{1}{4}$.

1.53. a) $\frac{4}{7}$; b) $\frac{5}{7}$; c) $\frac{4}{7}\cdot\frac{1}{6}=\frac{2}{21}$; d) $\frac{4}{7}\cdot\frac{2}{6}\cdot\frac{1}{5}=\frac{4}{105}$.

1.54. Ponieważ $(A_1\cup A_2\cup A_3\cup A_4)'=A_1'A_2'A_3'A_4'$, więc $P(A_1\cup A_2\cup A_3\cup A_4)=1-P(A_1'A_2'A_3'A_4')=1-(1-p_1)(1-p_2)(1-p_3)(1-p_4)$.

1.55. $P(AB)=\frac{1}{3}\neq P(A)P(B)=\frac{2}{3}\cdot\frac{2}{3}$, więc zdarzenia A, B są zależne. Zatem zdarzenia A, B, C nie są parami niezależne i tym bardziej nie są zespołowo niezależne.

1.56. a) Tak; b) tak.

1.57. W ogólności nie, bo zdarzenie $A=\varnothing$ i dowolne zdarzenie B wykluczają się i są niezależne, gdyż $P(AB)=0=P(A)P(B)$. Gdy $P(A)>0$ i $P(B)>0$, wtedy zdanie staje się prawdziwe.

1.58. a) $P(A'B')=P\{(A\cup B)'\}=1-P(A\cup B)=\ldots=P(A')P(B')$.

1.59. b) Wskazówka: $P(A\,|\,B)-P(A)=\dfrac{P(AB)}{P(B)}-P(A)=\dfrac{r(A,B)}{P(B)}$; c) $r(A_1,A_2)=\frac{7}{15}-\frac{7}{10}\cdot\frac{7}{10}=-\frac{7}{300}\approx -0{,}023$; d$_1$) $r(A,A)=P(A)P(A')=p(1-p)$, $\max r(A,A)=\frac{1}{4}$ dla $P(A)=p=\frac{1}{2}$; d$_2$) $r(A,A')=-P(A)P(A')$, $\min r(A,A')=-\frac{1}{4}$.

1.60. $0{,}7+0{,}7-0{,}49=0{,}91<0{,}95$, zatem bardziej prawdopodobne jest rozpalenie ogniska dwiema złączonymi zapałkami.

1.61. Można przyjąć, że zdarzenia A i B są niezależne, zatem $P(AB)=P(A)P(B)=10^{-12}$.

1.62. Ponieważ popełniany wtedy błąd $P(AB)$ jest zaniedbywalnie mały.

1.63. $P(A_1)=p^2$; $P(A_2)=p+p-pp=p(2-p)$; $P(A_1)<P(A_2)$ przy każdym p. $P(B_1)=$ $=p^3$; $P(B_2)=3p-3p^2+p^3$.

Wskazówka. Udowodnić i wykorzystać równość $P(A\cup B\cup C)=P(A)+P(B)+$ $+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$; $P(B_3)=p(2p-p^2)$; $P(B_4)=p+p^2-p\cdot p^2$; $P(B_5)=p+(2p-p^2)-p(2p-p^2)=\ldots=P(B_2)$; $P(B_1)<P(B_2)$, $P(B_2)>P(B_3)$, $P(B_3)<P(B_4)$, $P(B_4)<P(B_5)$ przy dowolnym $p\in(0, 1)$, porównać pozostałe pary prawdopodobieństw $P(B_i)$. $P(C_1)=p^2(2p-p^2)$, $P(C_2)=p^2(2-p)^2$, $P(C_3)=p^2+p^2-p^4$, $P(C_4)=p(3p-3p^2+$ $+p^3)$, $P(C_5)=p\cdot p(1+p-p^2)$, $P(C_6)=p+p^3-p^4$; $P(C_3)<P(C_4)$ dla $0<p<\frac{1}{2}$ oraz $P(C_3)>$ $>P(C_4)$ dla $\frac{1}{2}<p<1$.

1.64. $\frac{1}{3}\cdot 0{,}97+\frac{1}{3}\cdot 0{,}90+\frac{1}{3}\cdot 0{,}86=0{,}91$.

1.65. A_i – wylosowanie „przewodniczącego" przez i-tą osobę losującą, $P(A_1)=\frac{1}{3}$, $P(A_2)=P(A_1'A_2)=\frac{2}{3}\cdot\frac{1}{2}=\frac{1}{3}$, $P(A_3)=P(A_1'A_2')=\frac{2}{3}\cdot\frac{1}{2}=\frac{1}{3}$.

1.66. a) 0,7; b) $0{,}7\cdot 0{,}9=0{,}63$; c) $0{,}7\cdot 0{,}9+0{,}3\cdot 0{,}8=0{,}87$; d) stosujemy twierdzenie Bayesa: $0{,}63/0{,}87\approx 0{,}72$.

1.67. a) 0,4; b) $0{,}4\cdot 0{,}8=0{,}32$; c) 0,84; d) $\frac{8}{21}\approx 0{,}38$.

1.68. a) 0,09; b) $\frac{4}{9}\approx 0{,}44$. **1.69.** a) 0,12; b) $\frac{5}{6}\approx 0{,}83$. **1.70.** $\approx 0{,}9988\approx 0{,}999$.

1.71. Niech A_i, B_j, $i=1, 2, 3$, $j=1, 2, 3, 4$ oznaczają następujące zdarzenia: A_i – na wejściu nadano odpowiednio ciąg AAAA, BBBB, CCCC, B_j – na wyjściu odebrano ciąg określony w poleceniu a$_j$). Wówczas: a$_1$) $P(B_1)=\sum_{i=1}^{3} P(A_i)P(B_1\,|\,A_i)=0{,}4\cdot 0{,}8^4+0{,}3\cdot$ $\cdot 0{,}1^4+0{,}3\cdot 0{,}1^4=0{,}1639$; a$_2$) $P(B_2)=0{,}4\cdot 0{,}1\cdot 0{,}8^3+0{,}3\cdot 0{,}1^4+0{,}3\cdot 0{,}8\cdot 0{,}1^3=0{,}02075$; a$_3$) $P(B_3)=0{,}12295$; a$_4$) $P(B_4)=0{,}4\cdot 0{,}8^2\cdot 0{,}1^2+0{,}3\cdot 0{,}8\cdot 0{,}1^3\cdot 2=0{,}00304$; b) $\approx 0{,}987$; c) $\approx 0{,}9996$.

2 ZMIENNE LOSOWE JEDNOWYMIAROWE

2.1. ZMIENNA LOSOWA I JEJ DYSTRYBUANTA

2.1.1. Pojęcie zmiennej losowej. W ujęciu intuicyjnym zmienna losowa (związana z pewnym doświadczeniem), to taka zmienna, która w wyniku doświadczenia przyjmuje wartość liczbową zależną od przypadku, a więc nie dającą się ustalić przed przeprowadzeniem doświadczenia. Użyty zwrot „zależnie od przypadku" oznacza zależnie od zdarzenia elementarnego ω, które zrealizowało się w wyniku doświadczenia. A oto definicja zmiennej losowej.

Niech $(\Omega, \mathscr{Z}, P)$ będzie dowolną przestrzenią probabilistyczną. *Zmienną losową* nazywamy dowolną funkcję X, określoną na przestrzeni zdarzeń elementarnych $\boldsymbol{\Omega}$, o wartościach ze zbioru $\boldsymbol{R}$ liczb rzeczywistych, mającą następujące własności: dla dowolnej, ustalonej liczby rzeczywistej x zbiór zdarzeń elementarnych ω, dla których spełniona jest nierówność $X(\omega)<x$, jest zdarzeniem, czyli

$$\{\omega:\ X(\omega)<x\} \in \mathscr{Z} \text{ dla każdego } x \in \boldsymbol{R}. \tag{2.1.1}$$

Gdy przestrzeń zdarzeń elementarnych jest skończona, a zdarzeniami są wszystkie jej podzbiory, wtedy warunek (2.1.1) nie stanowi żadnego ograniczenia i wobec tego każda funkcja X, odwzorowująca zbiór zdarzeń elementarnych Ω w zbiór $\boldsymbol{R}$ liczb rzeczywistych, jest zmienną losową.

Zmienne losowe oznaczamy dużymi literami np.: S, T, Z, X, Y, ich *wartości* zaś (mówimy także: *realizacje*) – odpowiednimi małymi literami: s, t, z, x, y, często ze wskaźnikami.

2.1.2. Prawdopodobieństwo przyjęcia przez zmienną losową wartości z danego zbioru. Zdarzeniami losowymi są również zbiory zdarzeń elementarnych postaci

$$\{\omega:\ X(\omega) \in A\}, \tag{2.1.2}$$

gdzie A jest dowolnym [1] zbiorem na prostej; np.: $A=\{x_0\}$, $A=\langle a, b\rangle$, $A=(x_0, \infty)$, $A=\boldsymbol{N}$, gdzie $\boldsymbol{N}=\{1, 2, \ldots\}$.

[1] Dokładniej: dowolnym zbiorem borelowskim na prostej (np. [11]).

Prawdopodobieństwo $P(X\in A)$ przyjęcia przez zmienną losową X wartości ze zbioru A, gdzie $A\subset \boldsymbol{R}$, określamy następującą równością

$$P(X\in A)=P[\{\omega:\ X(\omega)\in A\}]. \tag{2.1.3}$$

Jest to więc prawdopodobieństwo zdarzenia (2.1.2). Na przykład gdy $A=\langle a,b)$, wtedy

$$P(a\leqslant X<b)=P[\{\omega:\ a\leqslant X(\omega)<b\}].$$

2.1.3. Dystrybuanta zmiennej losowej i jej własności. Z określenia zmiennej losowej X wynika, że podzbiór (2.1.1) przestrzeni zdarzeń elementarnych Ω jest zdarzeniem. Można więc mówić o jego prawdopodobieństwie $P(X<x)$. Funkcję F_X określoną na całym zbiorze $\boldsymbol{R}=(-\infty,+\infty)$ liczb rzeczywistych wzorem:

$$F_X(x)=P(X<x),\qquad x\in\boldsymbol{R} \tag{2.1.4}$$

nazywamy *dystrybuantą zmiennej losowej X*. Jeżeli nie ma wątpliwości z jaką zmienną losową mamy do czynienia, to jej dystrybuantę będziemy oznaczali literą F zamiast F_X.

Dystrybuanta F dowolnej zmiennej losowej ma następujące własności F1 - F7:

F1. $0\leqslant F(x)\leqslant 1$ *dla każdego* $x\in\boldsymbol{R}$;

F2. $\lim\limits_{x\to-\infty} F(x)\equiv F(-\infty)=0$ *oraz* $\lim\limits_{x\to+\infty} F(x)\equiv F(+\infty)=1$;

F3. *jest funkcją niemalejącą*;

F4. *jest funkcją (co najmniej) lewostronnie ciągłą, czyli*

$$F(x_0-0)=F(x_0)\ \textit{dla każdego}\ x\in\boldsymbol{R},$$

gdzie $F(x_0-0)$ *oznacza granicę lewostronną funkcji* F *w punkcie* x_0: $F(x_0-0)=\lim\limits_{x\to x_0^-}F(x)$;

F5. *prawdopodobieństwo* $P(a\leqslant X<b)$ *przyjęcia przez zmienną losową* X *wartości z przedziału* $\langle a,b)$ *jest równe przyrostowi dystrybuanty* F *między punktami* a, b:

$$P(a\leqslant X<b)=F(b)-F(a); \tag{2.1.5}$$

F6. *prawdopodobieństwo* $P(X=x_0)$ *przyjęcia przez zmienną losową* X *dowolnej, ustalonej wartości* x *wyraża się za pomocą dystrybuanty* F *równością*:

$$P(X=x_0)=F(x_0+0)-F(x_0), \tag{2.1.6}$$

gdzie $F(x_0+0)$ *oznacza granicę prawostronną dystrybuanty* F *w punkcie* x_0, *czyli* $F(x_0+0)=$ $=\lim\limits_{x\to x_0^+}F(x)$;

F7. *jeżeli* G *jest dowolną funkcją o wartościach rzeczywistych mającą własności* F2, F3, F4, *to funkcja* G *jest dystrybuantą zmiennej losowej.*

2.1.4. Zadania rozwiązane.

ZADANIE 2.1. W grupie studenckiej przeprowadzono sprawdzian. Niech X oznacza ocenę (przy czterostopniowej skali ocen) losowo wybranego studenta. Czy X jest zmienną losową?

Rozwiązanie. Tak, gdyż mamy tu do czynienia z przyporządkowaniem poszczególnym studentom ich ocen. Przestrzenią Ω zdarzeń elementarnych jest tu zbiór studentów danej grupy, zbiorem zaś wartości zmiennej losowej X – zbiór $\{5, 4, 3, 2\}$. Gdyby to była grupa dziesięcioosobowa, wtedy zmienna losowa X mogłaby być określona np. następująco:

$$X(\omega_7)=5, \quad X(\omega_2)=X(\omega_6)=X(\omega_{10})=4,$$

$$X(\omega_1)=X(\omega_3)=X(\omega_5)=X(\omega_8)=3, \quad X(\omega_4)=X(\omega_9)=2.$$

Wskaźnik k przy symbolu ω_k można interpretować jako numer studenta na liście obecności.

Zadanie 2.2. Niech A będzie zdarzeniem związanym z pewnym doświadczeniem D_0 (np. A może oznaczać wyrzucenie parzystej liczby oczek w jednorazowym rzucie kostką do gry). Rozważamy bardziej złożone doświadczenie D: doświadczenie D_0 przeprowadzamy aż do zajścia zdarzenia A (mówimy także: aż do zajścia sukcesu) po raz pierwszy. Określić sensownie przestrzeń zdarzeń elementarnych doświadczenia D i podać przykłady zmiennych losowych określonych na tej przestrzeni.

Rozwiązanie. Jako przestrzeń zdarzeń elementarnych Ω przyjmujemy $\Omega=\{\omega_\infty, \omega_1, \omega_2, \ldots, \omega_n, \ldots\}$, gdzie ω_n oznacza zdarzenie elementarne doświadczenia D polegające na tym, że zdarzenie A zrealizowało się po raz pierwszy za n-tym, $n \in N$, przeprowadzeniem doświadczenia D_0, ω_∞ zaś oznacza – przy żadnej liczbie powtórzeń doświadczenia D_0 zdarzenie A nie zrealizowało się. Można tu określić zmienne losowe, powiedzmy T i K, np. następująco:

$$T(\omega_n)=n=\text{liczba przeprowadzonych doświadczeń } D_0, \tag{1}$$

$$K(\omega_n)=cn. \tag{2}$$

Jeżeli założyć, że jedno doświadczenie D_0 jest przeprowadzane w jednej jednostce czasu, to zmienną losową T można interpretować jako *czas oczekiwania na pierwszy sukces*. Podobnie zmienna losowa K oznacza koszt uzyskania sukcesu po raz pierwszy, jeżeli c we wzorze (2) oznacza koszt przeprowadzenia jednego doświadczenia D_0.

2.2. ZMIENNA LOSOWA TYPU SKOKOWEGO

2.2.1. Pojęcie zmiennej losowej typu skokowego. Wyróżniamy dwa zasadnicze typy zmiennych losowych: zmienne losowe typu skokowego i zmienne losowe typu ciągłego. W tym paragrafie zajmujemy się tylko pierwszymi z nich.

Mówimy, że zmienna losowa X jest *typu skokowego* (*dyskretnego*), jeżeli istnieje skończony albo przeliczalny zbiór $W_x=\{x_1, \ldots, x_n, \ldots\}$ jej wartości $x_1, \ldots, x_n, \ldots$ taki, że

$$P(X=x_i)=p_i>0, \quad i \in N, \tag{2.2.1}$$

$$\sum_{i=1} p_i=1, \tag{2.2.2}$$

gdzie górna granica sumowania wynosi n albo ∞ stosownie do tego, czy zbiór W_x jest skończony czy przeliczalny.

Przy tym, równość (2.2.2) nazywamy *warunkiem unormowania*, liczby $x_1, \ldots, x_n, \ldots$ – *punktami skokowymi* (albo *atomami*) zmiennej losowej X, prawdopodobieństwa zaś $p_1, \ldots, p_n, \ldots$ – *skokami*.

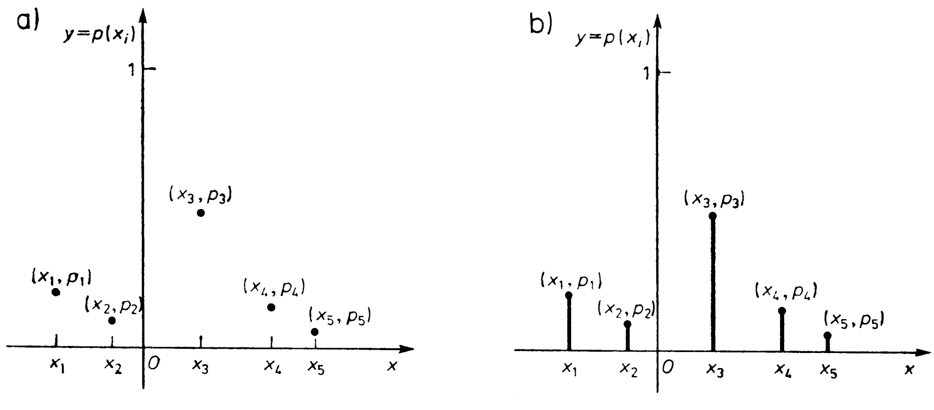
2.2.2. Funkcja rozkładu prawdopodobieństwa. Funkcję p określoną na zbiorze W_x (por. wyżej) równością

$$p(x_i) = P(X = x_i) \equiv p_i, \qquad x_i \in W_x \tag{2.2.3}$$

albo, co jest równoważne, dwuwierszową tablicą

$$\begin{array}{c|c|c|c|c|c} x_i & x_1 & x_2 & \ldots & x_n & \ldots \\ \hline p_i & p_1 & p_2 & \ldots & p_n & \ldots \end{array} \tag{2.2.4}$$

i spełniającą warunek unormowania (2.2.2) nazywamy *funkcją rozkładu prawdopodobieństwa* (krócej: *funkcją prawdopodobieństwa*) zmiennej losowej X.



Rys. 2.1. a) Wykres, b) histogram funkcji rozkładu prawdopodobieństwa

Wykresem w prostokątnym układzie współrzędnych (rys. 2.1) funkcji prawdopodobieństwa jest zbiór punktów (x_i, p_i). Suma długości wszystkich odcinków o końcach $(x_i, 0)$, (x_i, p_i), zgodnie z warunkiem unormowania (2.2.2), jest równa jedności. Wykres funkcji prawdopodobieństwa uzupełniony tymi odcinkami nazywamy *histogramem funkcji prawdopodobieństwa*.

2.2.3. Dystrybuanta zmiennej losowej typu skokowego. Gdy dana jest funkcja prawdopodobieństwa zmiennej losowej X, to – prawdopodobieństwo przyjęcia przez tę zmienną wartości ze zbioru A jest określone równością:

$$P(X \in A) = \sum_{x_i \in A} p_i, \tag{2.2.5}$$

– w szczególności dla dowolnego przedziału (a, b) mamy

$$P(a<X<b)=\sum_{a<x_i<b} p_i, \tag{2.2.6}$$

– podobnie, przyjmując w ostatniej równości $a=-\infty$, otrzymujemy wzór pozwalający wyznaczyć dystrybuantę F skokowej zmiennej losowej X:

$$F(x)=P(X<x)=\sum_{-\infty<x_i<x} p_i. \tag{2.2.7}$$

Liczby x_i w równościach (2.2.5) - (2.2.7) oznaczają oczywiście punkty skokowe zmiennej losowej X.

2.2.4. Wyznaczenie funkcji rozkładu prawdopodobieństwa, gdy dana jest dystrybuanta. Wzór (2.2.7) umożliwia wyznaczenie dystrybuanty zmiennej losowej, gdy znana jest jej funkcja prawdopodobieństwa. Możliwe jest postępowanie w odwróconej kolejności – mając daną dystrybuant F zmiennej losowej X można znaleźć jej funkcję prawdopodobieństwa: punkty skokowe x_i (czyli górny wiersz tabelki (2.2.4)), to punkty nieciągłości dystrybuanty F, odpowiadające im zaś prawdopodobieństwa $p_i=P(X=x_i)$ (mówimy również: masy prawdopodobieństwa umieszczone w tych punktach) obliczamy na podstawie własności F6 dystrybuanty (zad. 2.4).

Jeżeli jest dana funkcja prawdopodobieństwa albo dystrybuanta zmiennej losowej, to mówimy, że dany jest *rozkład prawdopodobieństwa* tej zmiennej.

2.2.5. Zadania rozwiązane.

ZADANIE 2.3. Zakładając, że w zadaniu 2.1 stosunek ocen b. dobrych, dobrych, dostatecznych, niedostatecznych ma się tak, jak 1 : 3 : 4 : 2, wyznaczyć dla określonej tam zmiennej losowej X:

a) funkcję prawdopodobieństwa i jej wykres,

b) dystrybuantę i jej wykres,

c) prawdopodobieństwo $P(X<3,5)$, korzystając:

c_1) z funkcji prawdopodobieństwa,

c_2) z dystrybuanty; zaznaczyć to prawdopodobieństwo na rysunku z wykresem dystrybuanty,

d) prawdopodobieństwo $P(3\leqslant X<4,5)$, korzystając:

d_1) z funkcji prawdopodobieństwa,

d_2) z dystrybuanty; zaznaczyć to prawdopodobieństwo na rysunku z wykresem dystrybuanty,

e) na wykresie dystrybuanty podać interpretację prawdopodobieństwa $P(X=3)$.

Rozwiązanie. a) Niech k będzie współczynnikiem proporcjonalności, N – liczbą studentów danej grupy. Wobec tego liczba studentów z oceną 2, 3, 4, 5 wynosi odpowiednio $2k$, $4k$, $3k$, k; wszystkich studentów jest $N=10k$. Dlatego:

$$p_1=P(X=2)=2k/10k=\tfrac{2}{10}, \quad p_2=P(X=3)=4k/10k=\tfrac{4}{10},$$

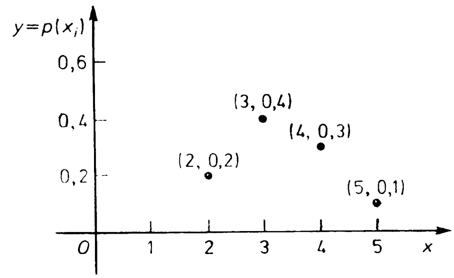
$$p_3=P(X=4)=3k/10k=\tfrac{3}{10}, \quad p_4=P(X=5)=k/10k=\tfrac{1}{10}.$$

Funkcja prawdopodobieństwa zmiennej losowej X jest więc następująca:

x_i	2	3	4	5
p_i	0,2	0,4	0,3	0,1

. Oczywiście $\sum_{i=1}^{4} p_i = 1$.

Wykres funkcji rozkładu prawdopodobieństwa podano na rysunku 2.2.



Rys. 2.2. Wykres funkcji rozkładu prawdopodobieństwa z zad. 2.3

b) Dystrybuantę znajdujemy na podstawie wzoru (2.2.7). Zmienna losowa X ma 4 punkty skokowe, więc dystrybuanta będzie określona pięcioma wzorami. Kolejno otrzymujemy:

dla $x \leqslant 2$ $\quad F(x) = P(X < x) = \sum_{x_i < x} p_i = 0,$

dla $2 < x \leqslant 3$ $\quad F(x) = \sum_{x_i < x} p_i = p_1 = 0{,}2,$

dla $3 < x \leqslant 4$ $\quad F(x) = \sum_{x_i < x} p_i = p_1 + p_2 = 0{,}6,$

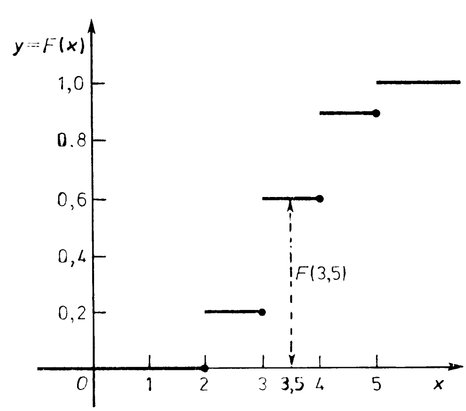
dla $4 < x \leqslant 5$ $\quad F(x) = \sum_{x_i < x} p_i = p_1 + p_2 + p_3 = 0{,}9,$

dla $x > 5$ $\quad F(x) = \sum_{x_i < x} p_i = p_1 + p_2 + p_3 + p_4 = 1.$

Znalezioną dystrybuantę wygodnie jest opisać w postaci dwuwierszowej tabelki:

x	$(-\infty, 2\rangle$	$(2, 3\rangle$	$(3, 4\rangle$	$(4, 5\rangle$	$(5, +\infty)$
$F(x)$	0	$\frac{2}{10}$	$\frac{6}{10}$	$\frac{9}{10}$	1

Jej wykres jest podany na rysunku 2.3.



Rys. 2.3. Wykres dystrybuanty zmiennej losowej z zad. 2.3

c_1) Na podstawie wzoru (2.2.7) otrzymujemy:

$$P(X<3{,}5)=\sum_{-\infty<x_i<3{,}5} P(X=x_i)=P(X=2)+P(X=3)=$$

$$=p_1+p_2=0{,}2+0{,}4=0{,}6.$$

c_2) Prawdopodobieństwo $P(X<3{,}5)$, zgodnie z definicją dystrybuanty, jest po prostu wartością wyznaczonej już dystrybuanty w punkcie $x=3{,}5$ (rys. 2.3):

$$P(X<3{,}5)=F(3{,}5)=0{,}6.$$

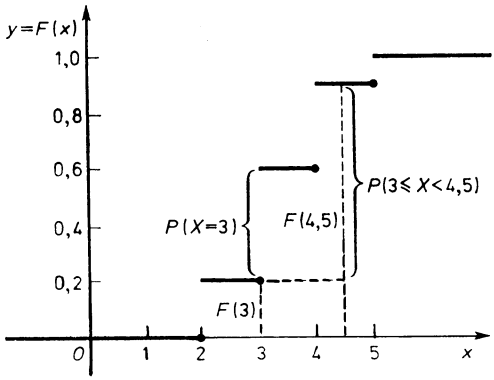
d_1) Na podstawie wzoru (2.2.6) otrzymujemy:

$$P(3\leqslant X<4{,}5)=\sum_{3\leqslant x_i<4{,}5} p_i=p_2+p_3=0{,}4+0{,}3=0{,}7.$$

d_2) Na mocy własności F5 dystrybuanty (2.1.5), mamy:

$$P(3\leqslant X<4{,}5)=F(4{,}5)-F(3)=0{,}9-0{,}2=0{,}7.$$

Interpretację tego prawdopodobieństwa podano na rys. 2.4.



Rys. 2.4. Prawdopodobieństwo $P(3\leqslant X<4{,}5)$ jest przyrostem dystrybuanty

e) W myśl własności (2.1.6) prawdopodobieństwo $P(X=3)$ wyraża się przy użyciu dystrybuanty równością:

$$P(X=3)=F(3+0)-F(3)=\tfrac{6}{10}-\tfrac{2}{10}=0{,}4.$$

Jest zatem ono skokiem dystrybuanty F w punkcie $x=3$ (rys. 2.4).

ZADANIE 2.4. Dystrybuanta F zmiennej losowej X jest określona następującą tabelką:

x	$(-\infty,-2\rangle$	$(-2,3\rangle$	$(3,5\rangle$	$(5,+\infty)$
$F(x)$	0	0,4	0,5	1

Wyznaczyć funkcję prawdopodobieństwa tej zmiennej.

Rozwiązanie. Punkty skokowe x_i, to oczywiście punkty nieciągłości dystrybuanty F, czyli punkty -2, 3, 5. Pozostaje znaleźć odpowiadające im skoki p_1, p_2, p_3. Te zaś znajdujemy na podstawie własności F6 dystrybuanty:

$$p_1 = P(X=-2) = F(-2+0) - F(-2) = 0{,}4 - 0 = 0{,}4,$$

$$p_2 = P(X=3) = F(3+0) - F(3) = 0{,}5 - 0{,}4 = 0{,}1,$$

$$p_3 = P(X=5) = F(5+0) - F(5) = 1 - 0{,}5 = 0{,}5.$$

Tak więc funkcja prawdopodobieństwa zmiennej losowej X jest następująca:

x_i	-2	3	5
p_i	0,4	0,1	0,5

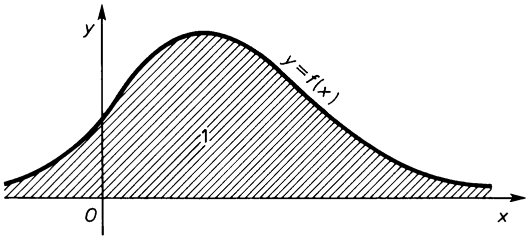
2.3. ZMIENNE LOSOWE TYPU CIĄGŁEGO

Zmienna losowa X przyjmująca wszystkie wartości z pewnego przedziału (albo przedziałów), dla której istnieje nieujemna funkcja f taka, że dystrybuantę F zmiennej losowej X można przedstawić w postaci:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \quad \text{dla} \quad x \in \boldsymbol{R} \tag{2.3.1}$$

nazywamy *zmienną losową absolutnie ciągłą* (inaczej: *zmienną losową ciągłą* lub *typu ciągłego*), a funkcję f jej *gęstością*.

Powiemy, że dany jest *rozkład prawdopodobieństwa zmiennej losowej X typu ciągłego*, jeżeli jest dana jej dystrybuanta F lub gęstość f.



Rys. 2.5. Pole obszaru ograniczonego wykresem gęstości f i osią Ox jest równe 1

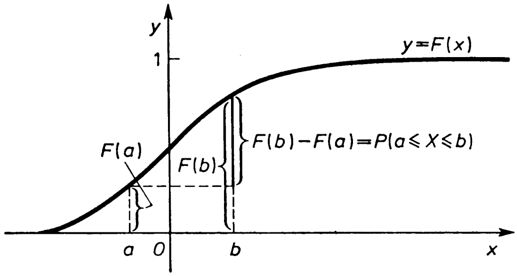
Jeżeli gęstość zmiennej losowej X jest różna od zera tylko w przedziale (a, b), to rozkład nazywamy *skoncentrowanym na przedziale* (a, b). Wykresem dystrybuanty zmiennej losowej typu ciągłego jest linia ciągła.

Własności zmiennej losowej typu ciągłego:

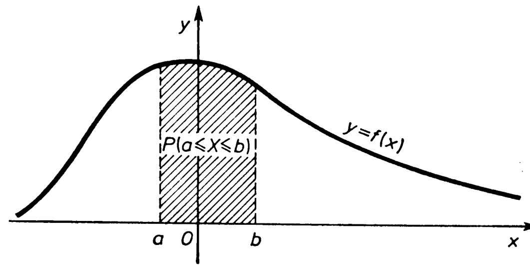
Jeżeli x jest punktem ciągłości gęstości f, to

$$F'(x) = f(x), \tag{2.3.2}$$

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \quad \text{(rys. 2.5)}, \tag{2.3.3}$$



Rys. 2.6. $P(a \leqslant X \leqslant b)$ na rysunku dystrybuanty F zmiennej losowej X typu ciągłego



Rys. 2.7. $P(a \leqslant X \leqslant b)$ jest polem nad przedziałem $\langle a, b\rangle$ pod wykresem gęstości

$$\bigwedge_{c \in R} P(X=c)=0, \tag{2.3.4}$$

$$P(a \leqslant X<b)=P(a<X \leqslant b)=P(a<X<b)=P(a \leqslant X \leqslant b)=F(b)-F(a) \quad \text{(rys. 2.6)}, \tag{2.3.5}$$

$$P(a \leqslant X \leqslant b)=\int_a^b f(x)\,dx \quad \text{(rys. 2.7)}. \tag{2.3.6}$$

2.3.1. Zadania rozwiązane

ZADANIE 2.5. W wielu sytuacjach można przyjąć, że czas X bezawaryjnej pracy badanego urządzenia jest zmienną losową ciągłą o gęstości

$$f(x)=\begin{cases} \dfrac{1}{\lambda}\exp\left(-\dfrac{x}{\lambda}\right) & \text{dla } x>0 \\ 0 & \text{dla pozostałych } x. \end{cases}$$

Niech $\lambda=10$.

a) Obliczyć prawdopodobieństwo $P(5 \leqslant X \leqslant 10)$.

b) Wyznaczyć dystrybuantę zmiennej losowej X.

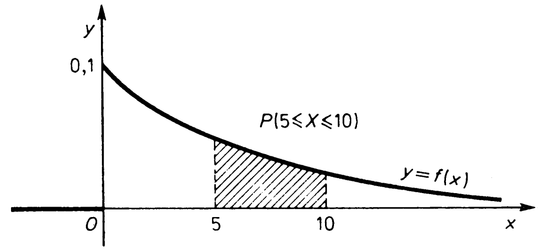
c) Zinterpretować obliczone prawdopodobieństwo za pomocą wykresu gęstości i dystrybuanty.

Rozwiązanie. a) Korzystając ze wzoru (2.3.6), otrzymamy

$$P(5 \leqslant X \leqslant 10)=\tfrac{1}{10}\int_5^{10} \exp\left(-\tfrac{1}{10}x\right) dx=\left[-\exp\left(-\tfrac{1}{10}x\right)\right]_5^{10}=e^{-0,5}-e^{-1} \approx 0,2486.$$

Z interpretacji geometrycznej całki oznaczonej funkcji nieujemnej wynika, że $P(5 \leqslant X \leqslant 10)$ jest równe (liczbowo) polu pod krzywą gęstości nad przedziałem $\langle 5, 10\rangle$ (rys. 2.8).

b) Rozpatrzymy dwa przypadki:



Rys. 2.8. Do zadania 2.5

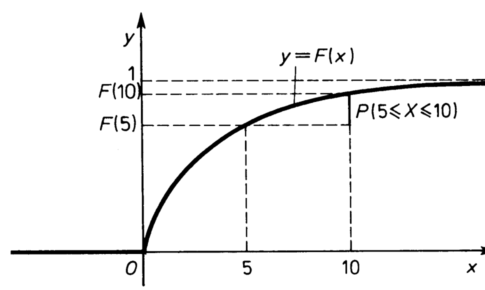
dla $x \leqslant 0$ $\quad F(x) = \int\limits_{-\infty}^{x} 0\,dt = 0,$

dla $x > 0$ $\quad F(x) = \int\limits_{-\infty}^{0} 0\,dt + \int\limits_{0}^{x} \frac{1}{10}\exp(-\frac{1}{10}t)\,dt = [-\exp(-\frac{1}{10}t)]_0^x = 1 - \exp(-\frac{1}{10}x).$

Stąd

$$F(x) = \begin{cases} 0 & \text{dla } x \leqslant 0, \\ 1 - \exp(-\frac{1}{10}x) & \text{dla } x > 0. \end{cases}$$

Korzystając ze związku $P(5 \leqslant X \leqslant 10) = F(10) - F(5)$, obliczone prawdopodobieństwo interpretujemy jako różnicę rzędnych wykresu dystrybuanty w punktach: $x_2 = 10$ i $x_1 = 5$ (rys. 2.9).



Rys. 2.9. Do zadania 2.5

Zadanie 2.6. Dobrać tak stałe A i B by funkcja określona wzorem

$$F(x) = A + B\operatorname{arctg} x \quad \text{dla} \quad -\infty < x < \infty$$

a) była dystrybuantą zmiennej losowej X.

b) Wyznaczyć gęstość zmiennej losowej X.

Rozwiązanie. a) Zauważmy, że dla dowolnych stałych $A, B \in \boldsymbol{R}$, $F(x) = A + B\operatorname{arc\,tg} x$ jest funkcją ciągłą. Na mocy własności 2, p. 2.1 stałe A i B należy tak dobrać, aby F była funkcją niemalejącą spełniającą warunki:

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

W naszym przypadku przyjmują one postać:

$$\lim_{x \to -\infty} (A + B\operatorname{arctg} x) = A - B\tfrac{1}{2}\pi = 0,$$

$$\lim_{x \to \infty} (A + B\operatorname{arctg} x) = A + B\tfrac{1}{2}\pi = 1.$$

Rozwiązaniem otrzymanego układu równań jest para liczb: $A=\frac{1}{2}$, $B=1/\pi$.

Funkcja F określona wzorem

$$F(x)=\frac{1}{\pi}\operatorname{arctg}x+\frac{1}{2}$$

jest rosnąca w przedziale $(-\infty, \infty)$ i z własności 7 z p. 2.1 wynika, że funkcja F jest dystrybuantą zmiennej losowej X.

b) Wobec zależności (2.3.2) gęstość zmiennej losowej X jest postaci:

$$f(x)=\frac{1}{\pi(1+x^2)} \quad \text{dla} \quad x\in \mathbf{R}. \tag{2.3.7}$$

2.4. FUNKCJE ZMIENNEJ LOSOWEJ SKOKOWEJ

2.4.1. Określenie funkcji zmiennej losowej i jej rozkład. Niech X będzie skokową zmienną losową o zbiorze W_x jej punktów skokowych x_i i funkcji prawdopodobieństwa p określonej wzorem (2.2.3) lub tabelką (2.2.4). Niech poza tym g będzie dowolną(¹) funkcją rzeczywistą o wartościach rzeczywistych określoną co najmniej na zbiorze W_x. Wówczas równość

$$U=g(X) \tag{2.4.1}$$

(rozumiana w następujący sposób: $U(\omega)=g[X(\omega)]$, $\omega\in\Omega$) określona na przestrzeni zdarzeń elementarnych Ω jest nową skokową zmienną losową U, zwaną *funkcją zmiennej losowej* X o punktach skokowych u_j, gdzie $u_j=g(x_i)$, tworzących pewien zbiór W_u; gdy g nie jest funkcją różnowartościową, to ten sam punkt skokowy u_j może odpowiadać więcej niż jednemu punktowi skokowemu x_i.

Niech q oznacza funkcję prawdopodobieństwa zmiennej losowej U. Funkcja ta jest wyznaczona przez prawdpodobieństwa p_i następującymi równościami:

$$q_j\equiv q(u_j)=P(U=u_j)=\sum_{\{x_i;\, g(x_i)=u_j\}} p(x_i), \quad x_i\in W_x, \quad u_j\in W_u. \tag{2.4.2}$$

Suma po prawej stronie równości (2.4.2) rozciąga się na te wszystkie punkty skokowe x_i, dla których $q(x_i)=u_j$ (zad. 2.7).

2.4.2. Niezależność zmiennych losowych. Mówimy, że *zmienne losowe* X i Y (określone na tej samej przestrzeni zdarzeń elementarnych) są *niezależne*, gdy dla dowolnych x, y są niezależne zdarzenia $\{X<x\}$, $\{Y<y\}$ tj., gdy

$$P(X<x,\, Y<y)=P(X<x)P(Y<y). \tag{2.4.3}$$

Podobnie określa się niezależność większej skończonej liczby zmiennych losowych określonych na tej samej przestrzeni zdarzeń elementarnych.

(¹) Bardziej dokładnie: niech g będzie dowolną borelowską funkcją rzeczywistą [11].

Mówimy, że *zmienne losowe* $X_1, \ldots, X_n, \ldots$ określone na tej samej przestrzeni zdarzeń elementarnych, tworzące nieskończony ciąg, są *niezależne*, gdy dla $n=2, 3, \ldots$, niezależne są zmienne losowe $X_1, \ldots, X_n$. Dyskretne zmienne losowe X, Y są niezależne wtedy i tylko wtedy, gdy są spełnione równości

$$P(X=x_i,\ Y=y_j)=P(X=x_i)P(Y=y_j) \tag{2.4.4}$$

dla każdej pary punktów skokowych (x_i, y_j), gdzie $x_i \in W_x$, $y_j \in W_y$. Zmienne losowe X, Y odpowiednio o dystrybuantach F_1, F_2 są niezależne wtedy i tylko wtedy, gdy

$$P(X<x,\ Y<y)=F_1(x)F_2(y) \tag{2.4.5}$$

dla każdej pary liczb (x, y).

2.4.3. Zadania rozwiązane.

ZADANIE 2.7. Funkcję prawdopodobieństwa zmiennej losowej X przedstawia tabelka

x_i	-2	-1	0	2	3	4
p_i	0,1	0,3	0,2	0,1	0,2	0,1

Wyznaczyć funkcję prawdopodobieństwa zmiennej losowej U, jeśli:
a) $U=\frac{1}{4}X+1$, b) $U=X^2$.

Rozwiązanie. a) Punkty skokowe u_i zmiennej losowej U są postaci:

$$u_i=\tfrac{1}{4}x_i+1, \quad \text{gdzie} \quad x_i \in W_x=\{-2, -1, 0, 2, 3, 4\}.$$

Stąd otrzymujemy zbiór W_u punktów skokowych u_i zmiennej losowej U:

$$W_u=\{\tfrac{1}{2}, \tfrac{3}{4}, 1, \tfrac{3}{2}, \tfrac{7}{4}, 2\}.$$

Aby znaleźć prawdopodobieństwa $q_i=P(U=u_i)$, zauważmy, że zdarzenia losowe

$$\{\omega\colon U(\omega)=u_i\}, \quad \{\omega\colon X(\omega)=x_i\}$$

z uwagi na różnowartościowość funkcji określonej wzorem $u=\frac{1}{4}x+1$, są identyczne. Zatem

$$q_i=P(U=u_i)=P(X=x_i)=p_i, \quad u_i \in W_u.$$

Tak więc funkcja prawdopodobieństwa q zmiennej losowej U jest następująca:

u_i	$\frac{1}{2}$	$\frac{3}{4}$	1	$\frac{3}{2}$	$\frac{7}{4}$	2
q_i	0,1	0,3	0,2	0,1	0,2	0,1

b) Zauważmy przede wszystkim, że funkcja $u=x^2$ nie jest różnowartościowa w zbiorze W_x, bo dla $x_1=-2$ i $x_2=2$ otrzymujemy ten sam punkt skokowy $u_j=4$. Dlatego konieczny jest teraz inny, ogólniejszy sposób postępowania. Niech mianowicie $\{W_x', W_x''\}$ będzie podziałem zbioru W_x, przy czym $x_i \in W_x''$ wtedy i tylko wtedy, gdy również $-x_i \in W_x''$. W rozważanym zadaniu $W_x'=\{-1, 0, 3, 4\}$, $W_x''=\{-2, 2\}$; oczywiście $W_x=W_x' \cup W_x''$ oraz $W_x' \cap W_x''=\emptyset$.

Podziałowi $\{W_x', W_x''\}$ zbioru W_x odpowiada podział $\{W_u', W_u''\}$ zbioru W_u punktów skokowych zmiennej losowej U. W rozpatrywanym przykładzie

$$W_u' = \{0, 1, 9, 16\}, \quad W_u'' = \{4\}, \quad \text{zatem} \quad W_u = \{0, 1, 4, 9, 16\}.$$

Pozostaje wyznaczyć wartości q_i funkcji prawdopodobieństwa q. Ponieważ na zbiorze W_x' funkcja $u = x^2$ jest różnowartościowa, więc dla $u_i \in W_u'$ otrzymujemy:

$$q_1 = P(U=0) = P(X=0) = 0{,}2,$$

$$q_2 = P(U=1) = P(X=-1) = 0{,}3,$$

$$q_4 = P(U=9) = P(X=3) = 0{,}2,$$

$$q_5 = P(U=16) = P(X=4) = 0{,}1.$$

Gdy natomiast $u_i = 4 \in W_u''$, wtedy

$$q_3 = P(U=4) = P(X^2=4) = P(X=-2 \text{ albo } X=2) =$$
$$= P(X=-2) + P(X=2) = 0{,}1 + 0{,}1 = 0{,}2.$$

Ostatnie prawdopodobieństwo można także znaleźć z warunku unormowania:

$$q_3 = 1 - (q_1 + q_2 + q_4 + q_5) = 1 - 0{,}8 = 0{,}2.$$

Z kolei zauważmy, że przy obliczaniu prawdopodobieństw q_1, q_2, q_4, q_5 również mogliśmy postępować tak, jak przy obliczaniu q_3, uwzględniając, że $P(X=-x_i)=0$, gdy $x_i \in W_x'$. Na przykład

$$q_4 = P(U=9) = P(X^2=9) = P(X=-3 \text{ albo } X=3) =$$
$$= P(X=-3) + P(X=3) = 0 + 0{,}2 = 0{,}2.$$

Jak zwykle, dla większej przejrzystości, szukaną funkcję prawdopodobieństwa q przedstawiamy w tabelce:

u_i	0	1	4	9	16
q_i	0,2	0,3	0,2	0,2	0,1

ZADANIE 2.8. Miesięczny koszt u prowadzenia przyzakładowego laboratorium jest zależny od liczby x zatrudnionych w nim pracowników. Załóżmy, że zależność ta jest postaci

$$u = 15000x + 10000\sqrt{x}. \tag{1}$$

Pierwszy składnik można interpretować jako koszty wynagrodzenia pracowników (jest on wprost proporcjonalny do liczby x zatrudnionych pracowników), drugi — jako inne koszty, wprawdzie również zależne od liczby zatrudnionych pracowników, ale już nie wprost proporcjonalnie. Należy zaplanować przyszłe miesięczne koszty U prowadzenia przyzakładowego laboratorium w budującym się zakładzie. Koszty te traktujemy jako zmienną losową, ponieważ zależą one od liczby zatrudnionych pracowników, ta zaś od wielu czynników, w tym również losowych, jest wobec tego zmienną losową X. Dlatego zamiast (1) mamy

$$U=15000X+10000\sqrt{X}. \tag{2}$$

Wyznaczyć funkcję prawdopodobieństwa q zmiennej losowej U, przyjmując następujący rozkład prawdopodobieństwa zmiennej losowej X:

x_i	2	3	4	5
p_i	0,10	0,25	0,40	0,25

Rozwiązanie. Punkty skokowe u_i zmiennej losowej U są postaci:

$$q_i=15000x_i+10000\sqrt{x_i}, \qquad x_i \in W_x=\{2,3,4,5\}.$$

Stąd (po zaokrągleniu do części całkowitej), mamy

$$W_u=\{44142, 62321, 80000, 97361\}.$$

Ponieważ funkcja (1) jest różnowartościowa, więc

$$q_i=P(U=u_i)=P(X=x_i)=p_i, \qquad u_i \in W_u.$$

Na koniec, uwzględniając zadany rozkład zmiennej losowej X, otrzymujemy szukany rozkład prawdopodobieństwa (w postaci funkcji prawdopodobieństwa) zmiennej losowej U:

u_i	44142	62321	80000	97361
q_i	0,10	0,25	0,40	0,25

2.5. FUNKCJE ZMIENNEJ LOSOWEJ TYPU CIĄGŁEGO

Rozważmy zmienną losową Y określoną równością

$$Y=g(X), \tag{2.5.1}$$

gdzie funkcja $y=g(x)$ jest określona co najmniej na zbiorze wartości zmiennej losowej X. Przypadek dyskretnej zmiennej losowej rozpatrzono w p. 2.4. Obecnie załóżmy, że zmienna losowa X jest typu ciągłego o dystrybuancie F. W prostszych przypadkach rozkład zmiennej losowej (2.5.1) można wyznaczyć bezpośrednio z definicji dystrybuanty G tej zmiennej losowej.

2.5.1. Zadania rozwiązane.

ZADANIE 2.9. Niech X będzie zmienną losową typu ciągłego o dystrybuancie F. Wyznaczyć dystrybuantę zmiennej losowej $Y=-X$.

Rozwiązanie. $G(y)=P(Y<y)=P(-X<y)=P(X>-y)=1-P(X<-y)+$ $+P(X=-y)=1-F(-y)$, ponieważ z własności (2.3.4) dla zmiennej losowej typu ciągłego mamy $P(X=-y)=0$.

W przypadkach bardziej skomplikowanych, jeżeli funkcja g jest ściśle monotoniczna, wygodniej jest skorzystać z następującego twierdzenia:

Jeżeli X jest zmienną losową ciągłą o gęstości f skoncentrowanej na przedziale (a, b) oraz $y=g(x)$ jest funkcją ściśle monotoniczną klasy C^1 o pochodnej $g'(x)\neq 0$ w tym przedziale, przy czym $x=h(y)$ jest funkcją odwrotną do $y=g(x)$, to gęstość k zmiennej losowej ciągłej $Y=g(X)$ jest postaci

$$k(y)=\begin{cases} f[h(y)]\,|h'(y)| & \text{dla } c<y<d, \\ 0 & \text{dla pozostałych } y, \end{cases} \tag{2.5.2}$$

gdzie $c=\min(c_1, d_1)$, $d=\max(c_1, d_1)$, $c_1=\lim\limits_{x\to a+0} g(x)$, $d_1=\lim\limits_{x\to b-0} g(x)$.

Zajmiemy się teraz funkcją g w (2.5.1), która w przedziale (a, b) nie jest ściśle monotoniczna, ale przedział ten daje się rozbić na rozłączne przedziały (a_i, b_i), w których funkcja g jest ściśle monotoniczna.

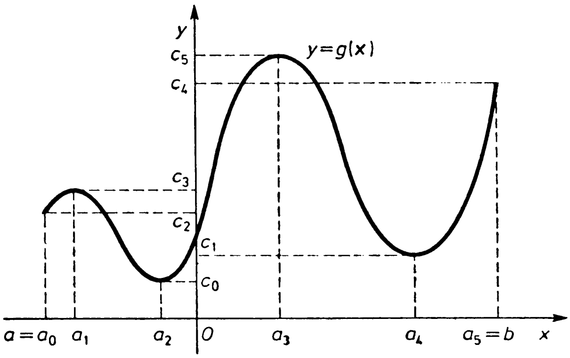
Prawdziwe jest wówczas następujące twierdzenie.

Jeśli X jest zmienną losową ciągłą o gęstości f skoncentrowanej na przedziale (a, b) (w szczególności na przedziale niewłaściwym) oraz funkcja $y=g(x)$ jest przedziałami ściśle monotoniczna, przy czym $g(x)=g_i(x)$ dla $x\in(a_{i-1}, a_i)$, $(a, b)=\bigcup\limits_{i=1}^{n-1}(a_{i-1}, a_i\rangle\cup(a_{n-1}, b)$, gdzie funkcje g_i są ściśle monotoniczne, klasy C^1 o pochodnych $g_i'(x)\neq 0$ dla $x\in(a_{i-1}, a_i)$, $i=1, \ldots, n$, $h_i(y)$ zaś są funkcjami odwrotnymi do $g_i(x)$, to gęstość k zmiennej losowej $Y=g(X)$ jest postaci:

$$k(y)=\begin{cases} \sum\limits_{i=1}^{m(y)} f[h_i(y)]\,|h_i'(y)| & \text{dla } c<y<d, \\ 0 & \text{dla pozostałych } y, \end{cases} \tag{2.5.3}$$

gdzie $c_i=g(a_i)$, $i=0, 1, \ldots, n$, $c=\min\limits_i c_i$, $d=\max\limits_i c_i$.

Liczba składników m w sumie występującej we wzorze (2.5.3) nie jest stała, zależy bowiem od wartości y (rys. 2.10). Dla ustalonego $y\in(c, d)$, $m(y)$ jest liczbą rozwiązań równania $y=g(x)$ w przedziale (a, b).



Rys. 2.10. Interpretacja liczby składników $m(y)$ we wzorze (2.5.3)

ZADANIE 2.10. Niech zmienna losowa X ma rozkład o gęstości $f(x)>0$ dla $x \in \boldsymbol{R}$. Wyznaczyć gęstość zmiennej losowej $Y=X^2$.

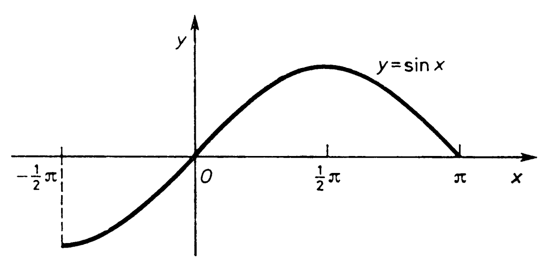
Rozwiązanie. Przedział $(a, b)=(-\infty, \infty)$ daje się rozbić na dwa przedziały rozłączne $(-\infty, \infty)=(-\infty, 0\rangle \cup (0, \infty)$, w których funkcja $y=x^2$ jest ściśle monotoniczna. Wyznaczmy funkcję odwrotną do $y=x^2$ w poszczególnych przedziałach:

1. dla $x \in (-\infty, 0)$ mamy $x=-\sqrt{y}$ oraz $x'=-\dfrac{1}{2\sqrt{y}}$,
2. dla $x \in (0, \infty)$: $x=\sqrt{y}$, $x'=\dfrac{1}{2\sqrt{y}}$.

Korzystając ze wzoru (2.5.3), otrzymamy gęstość zmiennej losowej Y:

$$k(y)=\begin{cases} \dfrac{f(-\sqrt{y})}{2\sqrt{y}}+\dfrac{f(\sqrt{y})}{2\sqrt{y}} & \text{dla } 0<y<\infty, \\ 0 & \text{dla pozostałych } y. \end{cases}$$

ZADANIE 2.11. Niech zmienna losowa X ma rozkład równomierny w przedziale $(-\frac{1}{2}\pi, \pi)$. Wyznaczyć gęstość zmiennej losowej $Y=\sin X$.



Rys 2.11. Do zadania 2.11

Rozwiązanie. Naszkicujemy wykres funkcji $y=\sin x$ (rys. 2.11). Przedział $(-\frac{1}{2}\pi, \pi)$ podzielimy na przedziały $(-\frac{1}{2}\pi, \frac{1}{2}\pi\rangle$, $(\frac{1}{2}\pi, \pi)$, w których funkcja $y=\sin x$ jest ściśle monotoniczna i wyznaczamy w nich funkcje odwrotne do $y=\sin x$:

1) dla $-\frac{1}{2}\pi<x\leqslant\frac{1}{2}\pi$ funkcją odwrotną do $y=\sin x$ jest $x=\arcsin y$ dla $-1<y\leqslant 1$, a jej pochodną $x'=\dfrac{1}{\sqrt{1-y^2}}$ dla $-1<y<1$;

2) dla $\frac{1}{2}\pi<x<\pi$ zauważmy, że $y=\sin x$ przyjmuje takie same wartości jak $y=-\sin x$ dla $-\frac{1}{2}\pi<x<0$, skąd $x=\arcsin(-y)$ dla $0<y<1$ oraz $x'=-\dfrac{1}{\sqrt{1-y^2}}$ dla $0<y<1$.

Zmienna losowa X ma rozkład o gęstości

$$f(x)=\begin{cases} \dfrac{2}{3\pi} & \text{dla } -\frac{1}{2}\pi<x<\pi, \\ 0 & \text{dla pozostałych } x. \end{cases}$$

Stąd i ze wzoru (2.5.3) gęstość zmiennej losowej Y jest określona wzorem:

$$k(y)=\begin{cases} \dfrac{2}{3\pi}\dfrac{1}{\sqrt{1-y^2}} & \text{dla} \quad -1<y\leqslant 0, \\ \dfrac{2}{3\pi}\left(\dfrac{1}{\sqrt{1-y^2}}+\dfrac{1}{\sqrt{1-y^2}}\right)=\dfrac{4}{3\pi\sqrt{1-y^2}} & \text{dla} \quad 0<y<1, \\ 0 \quad \text{dla pozostałych } y. \end{cases}$$

ZADANIE 2.12. Zmienna losowa X ma rozkład o gęstości f. Wyznaczyć gęstość zmiennej losowej: $Y=aX+b$, $a\neq 0$.

Rozwiązanie. Funkcja liniowa $y=ax+b$ jest ściśle monotoniczna na całej osi (dla $a>0$ rosnąca, dla $a<0$ malejąca) i spełnia wszystkie założenia podanego twierdzenia. Wyznaczmy funkcję odwrotną do niej:

$$x=\frac{1}{a}(y-b);$$

jej pochodna $x'=\dfrac{1}{a}$. Jeśli zmienna losowa X przyjmuje wartości z przedziału $(-\infty, \infty)$, to zmienna losowa Y też przyjmuje wartości z przedziału $(-\infty, \infty)$. Zgodnie ze wzorem (2.5.2) gęstość zmiennej losowej Y jest postaci:

$$k(y)=\frac{1}{|a|}f\left(\frac{y-b}{a}\right), \qquad y\in \boldsymbol{R}.$$

ZADANIE 2.13. Przez początek układu współrzędnych przeprowadzono losowo prostą i obrano na niej punkt M odległy od początku układu współrzędnych o jednostkę. Zakładając, że prawdopodobieństwo tego, że miara Φ kąta, jaki tworzy wektor $\overrightarrow{OM}$ z wersorem osi OX, będzie zawierać się w przedziale $(\varphi_0, \varphi_0+\Delta\varphi)$ nie zależy od φ_0, a jedynie od długości $\Delta\varphi$ tego przedziału, wyznaczyć rozkład prawdopodobieństwa odciętej X punktu M i obliczyć $P(0<X\leqslant\frac{1}{2})$.

Rozwiązanie. Na to aby $P(\varphi_0<\Phi<\varphi_0+\Delta\varphi)=F(\varphi_0+\Delta\varphi)-F(\varphi_0)$ nie zależało od φ_0, a jedynie od $\Delta\varphi$, potrzeba i wystarcza, by dystrybuanta zmiennej Φ była funkcją liniową kąta φ. Ponieważ $0<\varphi<\pi$, więc zmienna losowa Φ ma rozkład jednostajny w przedziale $(0, \pi)$, a więc o gęstości określonej wzorem:

$$f(\varphi)=\begin{cases} \dfrac{1}{\pi} & \text{dla } 0<\varphi<\pi, \\ 0 & \text{dla pozostałych } \varphi. \end{cases}$$

Losowa odcięta X punktu M jest funkcją kąta Φ (rys. 2.12): $X=\cos\Phi$ dla $0<\Phi<\pi$, stąd $-1<X<1$.

Pochodna $x' = -\sin\varphi$ funkcji $x = \cos\varphi$ jest różna od zera dla $0 < \varphi < \pi$. Funkcją odwrotną do niej jest $\varphi = \arccos x$, a jej pochodna: $\varphi' = -\dfrac{1}{\sqrt{1-x^2}} \neq 0$ dla $-1 < x < 1$. Stąd i ze wzoru (2.5.2) gęstość zmiennej losowej X jest postaci:

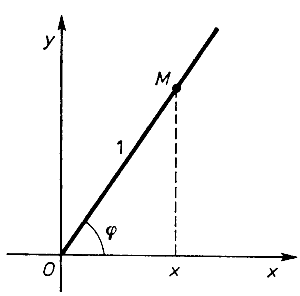
$$k(x) = \frac{1}{\pi\sqrt{1-x^2}} \quad \text{dla} \quad -1 < x < 1.$$

Jest to gęstość tzw. *rozkładu arcusa sinusa*. Korzystając ze wzoru (2.3.6), otrzymamy

$$P(0 < X \leqslant \tfrac{1}{2}) = \frac{1}{\pi}\int_0^{1/2} \frac{dx}{\sqrt{1-x^2}} = \left[\frac{1}{\pi}\arcsin x\right]_0^{1/2} = \tfrac{1}{6}.$$

Dosyć często spotykaną w teorii niezawodności funkcją dodatniej zmiennej losowej X jest tzw. *zmienna logarytmiczno-normalna*. Rozkład zmiennej losowej X, której $\ln X$ ma rozkład normalny (2.8.22) nazywamy *rozkładem logarytmiczno-normalnym*.

Rys. 2.12. Do zadania 2.13



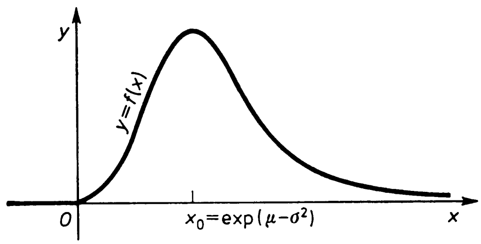
ZADANIE 2.14. Wyznaczyć gęstość rozkładu logarytmiczno-normalnego.

Rozwiązanie. Niech zmienna losowa $Y = \ln X$ ma rozkład $N(\mu, \sigma)$ (2.8.22). Chcemy wyznaczyć gęstość k zmiennej losowej $X = e^Y$. Funkcja $x = e^y$ spełnia założenia podanego ostatnio twierdzenia: jest rosnąca, klasy C^1, o pochodnej $x' = e^y \neq 0$ w przedziale $(-\infty, \infty)$. Funkcją odwrotną do niej jest $y = \ln x$ dla $x \in (0, \infty)$, a jej pochodną $y' = \dfrac{1}{x}$; stąd i ze wzoru (2.5.2) gęstość zmiennej losowej X jest określona wzorem:

$$k(x) = \frac{1}{\sigma\sqrt{2\pi}x}\exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \quad \text{dla} \quad x > 0. \tag{2.5.4}$$

Szkic wykresu tej gęstości podaje rys. 2.13, gdzie $x_0 = \exp(\mu - \sigma^2)$ (zad. 2.166).

Rozkład logarytmiczno-normalny wykorzystuje się przy badaniach: zmęczeniowych materiału, hydrologicznych jako modele przepływu w rzece, plastycznej wytrzymałości prętów, bądź jako modele opadów atmosferycznych.



Rys. 2.13. Gęstość rozkładu logarytmiczno-normalnego

2.6. CHARAKTERYSTYKI LICZBOWE ZMIENNEJ LOSOWEJ

2.6.1. Definicje najważniejszych charakterystyk liczbowych. W celu syntetycznego scharakteryzowania zmiennej losowej, przyporządkowuje się jej pewne liczby charakteryzujące ją pod względem np. wartości najbardziej prawdopodobnej, rozrzutu jej wartości, kształtu histogramu lub krzywej gęstości. Liczby te nazywamy *charakterystykami liczbowymi* zmiennej losowej (lub jej rozkładu prawdopodobieństwa). Charakterystyki te i ich definicje podaje tablica 2.1. Najważniejsze z nich to: wartość przeciętna, wariancja i odchylenie standardowe. Dlatego omówimy je dokładniej.

2.6.2. Wartość przeciętna zmiennej losowej i jej własności. Aby na podstawie definicji (tabl. 2.1) wyznaczyć wartość przeciętną zmiennej losowej $U=g(X)$ należałoby uprzednio wyznaczyć funkcję prawdopodobieństwa albo gęstość prawdopodobieństwa zmiennej losowej U (stosownie do tego czy jest to zmienna typu dyskretnego, czy ciągłego). Obliczanie wartości przeciętnej zmiennej losowej $U=g(X)$ bezpośrednio za pomocą funkcji prawdopodobieństwa albo gęstości prawdopodobieństwa zmiennej losowej X umożliwia następujący wzór:

$$EU=\begin{cases}\sum\limits_{x_i\in W_x} g(x_i)\,p_i,\\ \int\limits_{-\infty}^{\infty} g(x)f(x)\,dx\end{cases} \tag{2.6.15}$$

(przy założeniu, że szereg i całka po prawej stronie są bezwzględnie zbieżne).

Wartość przeciętna, oprócz (2.6.15) ma następujące własności (a, b, c — oznaczają stałe względem zdarzenia elementarnego ω):

$$E(c)=c, \tag{2.6.16}$$

$$E(aX)=aEX, \tag{2.6.17}$$

$$E(X+b)=EX+b, \tag{2.6.18}$$

$$E(X-EX)=0, \tag{2.6.19}$$

$$E(X+Y)=EX+EY, \tag{2.6.20}$$

$$E(XY)=EX\cdot EY, \quad \text{gdy } X,\ Y \text{ są niezależne.} \tag{2.6.21}$$

Tablica 2.1. Charakterystyki liczbowe zmiennej losowej

Nr wzoru	Mierzy wskazuje	Nazwa (nazwy)	Symbol	Definicja dla zmiennej losowej typu dyskretnego	Definicja dla zmiennej losowej typu ciągłego	Uwagi
(2.6.1)	położenie	wartość przeciętna, wartość oczekiwana	EX α_1	$EX=\sum_{x_i \in W} x_i p_i$ W – zbiór punktów skokowych	$EX=\int_{-\infty}^{\infty} xf(x)\,dx$ $f(x)$ – gęstość	założenie: szereg $\Sigma x_i p_i$, całka $\int_{-\infty}^{\infty} xf(x)\,dx$ są bezwzględnie zbieżne
(2.6.2)	położenie	mediana, wartość środkowa	$x_{0,5}$	każda liczba $x_{0,5}$ spełniająca warunki $F(x_{0,5})\leqslant 0,5\leqslant F(x_{0,5}+0)$ $\sum_{x_i<x_{0,5}} p_i\leqslant 0,5\leqslant \sum_{x_i\leqslant x_{0,5}} p_i$	$F(x_{0,5})=0,5$	może istnieć cały przedział median
(2.6.3)	położenie	kwantyl rzędu p	x_p	każda liczba x_p, $0<p<1$ spełniająca warunki $F(x_p)\leqslant p\leqslant F(x_p+0)$ $\sum_{x_i<x_p} p_i\leqslant p\leqslant \sum_{x_i\leqslant x_p} p_i$	$F(x_p)=p$	może istnieć cały przedział kwantyli tego samego rzędu
(2.6.4)	rozrzut	wariancja X dyspersja	D^2X, μ_2 VX σ_x^2, σ^2	$D^2X=E(X-\alpha_1)^2$ $D^2X=\sum_{x_i \in W}(x_i-\alpha_1)^2 p_i$	$D^2X=\int_{-\infty}^{\infty}(x-\alpha_1)^2 f(x)\,dx$	$D^2X\geqslant 0$; $D^2X=0 \Leftrightarrow$ gdy X ma rozkład jednopunktowy
(2.6.5)	rozrzut	odchylenie standardowe, odchylenie średnie	DX, σ_x σ	$DX=\sqrt{D^2X}$		pierwiastek należy rozumieć jako pierwiastek arytmetyczny

Tablica 2.1 (cd.)

Nr wzoru	Mierzy wskazuje	Nazwa (nazwy)	Symbol	Definicja dla zmiennej losowej typu dyskretnego	ciągłego	Uwagi
(2.6.6)	rozproszenie	odchylenie przeciętne od wartości przeciętnej	d_1, δ_1, β_1	$d_1 = E\|X-\alpha_1\|$ $d_1 = \sum_{x_i \in W} \|x_i - \alpha_1\| p_i$	$d_1 = \int_{-\infty}^{\infty} \|x-\alpha_1\| f(x)dx$	
(2.6.7)		odchylenie przeciętne od mediany	d_2, δ_2, β_2	$d_2 = E\|X-x_{0,5}\|$ $d_2 = \sum_{x_i \in W} \|x_i - x_{0,5}\| p_i$	$d_2 = \int_{-\infty}^{\infty} \|x-x_{0,5}\| f(x)dx$	$d_2 < E\|X-c\|$, jeśli $c \neq x_{0,5}$
(2.6.8)		współczynnik zmienności	v	$v = \frac{\sigma}{\alpha_1}$		gdy $\alpha_1 \neq 0$
(2.6.9)		współczynnik nierównomierności	h	$h = \frac{d_1}{\alpha_1}$		gdy $\alpha_1 \neq 0$

(2.6.10)	przy dowolnym r nie ma interpretacji	moment zwykły rzędu r	α_r	$\alpha_r = EX^r$ $\alpha_r = \sum_{x_i \in W} x_i^r p_i$	$\alpha_r = \int_{-\infty}^{\infty} x^r f(x)\,dx$	**Szereg, całka winny być bezwzględnie zbieżne z istnienia α_r wynika istnienie α_k dla $k<r$**
(2.6.11)		moment centralny rzędu r	μ_r	$\mu_r = E(X-\alpha_1)^r$ $\mu_r = \sum_{x_i \in W} (x_i-\alpha_1)^r p_i$	$\mu_r = \int_{-\infty}^{\infty} (x-\alpha_1)^r f(x)\,dx$	**Szereg, całka winny być bezwzględnie zbieżne** $\mu_r = \sum_{k=0}^{r} \binom{r}{k} \alpha_k (-\alpha_1)^{r-k}$
(2.6.12)	asymetrię	współczynnik skośności (asymetrii)	γ_1	$\gamma_1 = \frac{\mu_3}{\sigma^3}$		
(2.6.13)		współczynnik skupienia (kurtoza)	K	$K = \frac{\mu_4}{\sigma^4}$		
(2.6.14)		moda, wartość modalna, dominanta	m_0	punkt skokowy x_k, różny od min (x_i) i max (x_i), dla którego $p(x_k)$ osiąga maksimum absolutne	m_0 – odcięta maksimum absolutnego gęstości	

Własność (2.6.20) należy rozumieć jak następuje: jeżeli istnieją wartości oczekiwane każdej ze zmiennych losowych X i Y, to istnieje również wartość oczekiwana ich sumy $X+Y$ i zachodzi przy tym równość (2.6.20). Podobnie należy rozumieć równość (2.6.21), gdy X, Y są niezależne.

Jeżeli histogram funkcji prawdopodobieństwa lub wykres gęstości prawdopodobieństwa zmiennej losowej jest symetryczny względem pewnej prostej $x=x_0$ i istnieje wartość przeciętna EX, to

$$EX=x_0. \tag{2.6.22}$$

2.6.3. Wariancja zmiennej losowej i jej własności. Zmienną losową $X-\alpha_1$ nazywamy *odchyleniem zmiennej losowej X od jej wartości przeciętnej α_1*. *Wariancja zmiennej losowej* X, zgodnie z definicją (2.6.4), jest wartością przeciętną kwadratu odchylenia zmiennej losowej od jej wartości przeciętnej.

Wariancja zmiennej losowej ma następujące własności (a, b, c — oznaczają stałe względem zdarzenia elementarnego ω):

$$D^2(c)=0, \tag{2.6.23}$$

$$D^2(aX)=a^2D^2X, \tag{2.6.24}$$

$$D^2(X+b)=D^2X, \tag{2.6.25}$$

$$D^2(X\pm Y)=D^2X+D^2Y, \quad \text{gdy } X \text{ i } Y \text{ są niezależne}, \tag{2.6.26}$$

$$D^2X=E(X^2)-(EX)^2=\alpha_2-\alpha_1^2, \tag{2.6.27}$$

$$D^2X=E(X-c)^2-(c-\alpha_1)^2. \tag{2.6.28}$$

Gdy X jest zmienną losową typu skokowego, wtedy stałą c w ostatniej równości dobiera się tak, aby różnice x_i-c były liczbami wygodnymi do dalszych rachunków.

2.6.4. Zadania rozwiązane.

ZADANIE 2.15. Dla zmiennej losowej o funkcji prawdopodobieństwa danej tabelką

x_i	-2	2	4
p_i	0,5	0,3	0,2

wyznaczyć: a) wartość przeciętną EX, b) kwantyl $x_{0,3}$, c) medianę $x_{0,5}$, d) trzema sposobami wariancję D^2X, e) odchylenie standardowe DX, f) odchylenie przeciętne d_1.

Rozwiązanie. a) Bezpośrednio z definicji otrzymujemy:

$$EX=-2\cdot 0{,}5+2\cdot 0{,}3+4\cdot 0{,}2=0{,}4.$$

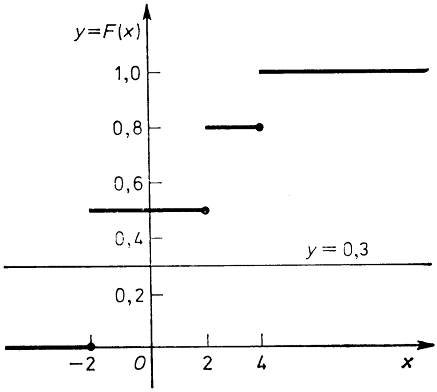
b) Szkicujemy wykres dystrybuanty $y=F(x)$ zmiennej losowej X (rys. 2.14). Prosta $y=0{,}3$ nie przecina wykresu dystrybuanty. Jednak dla punktu $x=-2$ mamy

$$F(-2)=0<0{,}3 \quad \text{oraz} \quad 0{,}3<F(-2+0)=0{,}5$$

i tym samym

$$F(-2) \leqslant 0{,}3 \leqslant F(-2+0).$$

Na podstawie definicji (2.6.3) wnioskujemy, że liczba $x=-2$ jest kwantylem rzędu $p=0{,}3$ rozważanej zmiennej losowej: $x_{0,3}=-2$. Żadna inna liczba nie jest kwantylem tego rzędu.



Rys. 2.14. Dystrybuanta zmiennej losowej z zadania 2.15

c) Na wykresie dystrybuanty (rys. 2.14) prosta $y=0{,}5$ pokrywa się z wykresem dystrybuanty na przedziale $(-2, 2\rangle$. Dlatego dla dowolnego $x \in (-2, 2)$ mamy

$$F(x)=0{,}5 \leqslant 0{,}5 \quad \text{oraz} \quad 0{,}5 \leqslant 0{,}5 = F(x+0).$$

Poza tym

$$F(-2)=0 \leqslant 0{,}5 \quad \text{oraz} \quad 0{,}5 \leqslant 0{,}5 = F(-2+0).$$

Podobnie

$$F(2)=0{,}5 \leqslant 0{,}5 \quad \text{oraz} \quad 0{,}5 \leqslant 0{,}8 = F(2+0).$$

Zatem dla każdego $x \in \langle -2, 2\rangle$ spełniona jest nierówność

$$F(x) \leqslant 0{,}5 \leqslant F(x+0).$$

Oznacza to, że każda liczba x z przedziału domkniętego $\langle -2, 2\rangle$ jest medianą rozważanej tu zmiennej losowej.

d) Najpierw korzystamy z definicji wariancji (2.6.4):

$$D^2X = (-2-0{,}4)^2 \cdot 0{,}5 + (2-0{,}4)^2 \cdot 0{,}3 + (4-0{,}4)^2 \cdot 0{,}2 =$$
$$= 5{,}76 \cdot 0{,}5 + 2{,}56 \cdot 0{,}3 + 12{,}96 \cdot 0{,}2 =$$
$$= 2{,}88 + 0{,}768 + 2{,}592 = 6{,}24.$$

Posłużmy się teraz równością (2.6.28), przyjmując $c=2$. Mamy

$$E(X-c)^2 = \sum_{i=1}^{3} (x_i - c)^2 p_i = (-2-2)^2 \cdot 0{,}5 + (2-2)^2 \cdot 0{,}3 +$$
$$+ (4-2)^2 \cdot 0{,}2 = 8 + 0 + 0{,}8 = 8{,}8$$

oraz

$$(c-EX)^2=(2-0{,}4)^2=2{,}56.$$

Zatem

$$D^2X=8{,}8-2{,}56=6{,}24.$$

Wreszcie skorzystamy ze wzoru (2.6.27). Kolejno otrzymujemy:

$$E(X^2)=\sum_{i=1}^{3} x_i^2 p_i=(-2)^2\cdot 0{,}5+2^2\cdot 0{,}3+4^2\cdot 0{,}2=$$

$$=4(0{,}5+0{,}3+0{,}8)=4\cdot 1{,}6=6{,}4.$$

Uwzględniając znalezioną już wartość przeciętną, ostatecznie otrzymujemy:

$$D^2X=6{,}4-0{,}4^2=6{,}24.$$

e) $DX=D^2X=\sqrt{6{,}24}\approx 2{,}5.$

f) Bezpośrednio z definicji (2.6.6) otrzymujemy:

$$d_1=\sum_{i=1}^{3}|x_i-0{,}4|\,p_i=|-2-0{,}4|\cdot 0{,}5+|2-0{,}4|\cdot 0{,}3+|4-0{,}4|\cdot 0{,}2=$$

$$=\tfrac{1}{10}(12+4{,}8+7{,}2)=2{,}4.$$

W zadaniu tym wariancję obliczyliśmy trzema sposobami. Już w tym prostym zadaniu można zauważyć, że nie wszystkie te sposoby są jednakowo uciążliwe. W związku z tym zanotujemy następującą wskazówkę:

– gdy różnice x_i – EX są całkowite, to stosujemy bezpośrednio definicję wariancji (2.6.4),

– jeżeli liczby x_i są małe co do wartości bezwzględnej, to stosujemy wzór (2.6.27),

– w pozostałych przypadkach, wykorzystujemy własności (2.6.28), obierając przy tym stosownie wartość wyjściową c (najczęściej jako c przyjmuje się liczbę całkowitą najbliższą wartości przeciętnej).

ZADANIE 2.16. Dany jest nieskończony ciąg liczbowy (p_k) postaci

$$p_k=pq^{k-1}, \qquad k\in N, \tag{2.6.29}$$

gdzie $0<p<1$, $q=1-p$. a) Sprawdzić, że ciąg ten jest funkcją prawdopodobieństwa pewnej zmiennej losowej K. Przyjmując $p_k=P(K=k)$ wyznaczyć: b) wartość przeciętną EK, c) wariancję D^2K.

Rozwiązanie. a) Wszystkie $p_k>0$. Pozostaje wobec tego sprawdzić, czy spełniony jest warunek unormowania. Mamy

$$\sum_{k=1}^{\infty} p_k=\sum_{k=1}^{\infty} pq^{k-1}=p(1+q+q^2+\ldots+q^{n-1}+\ldots).$$

Wyrażenie w nawiasie jest zbieżnym szeregiem geometrycznym o ilorazie q (ponieważ $-1<q<1$) i wyrazie pierwszym równym jedności. Sumą takiego szeregu jest $1/(1-q)$.

Otrzymujemy więc

$$\sum_{k=1}^{\infty} pq^{k-1} = p\cdot\frac{1}{1-q} = p\cdot\frac{1}{p} = 1.$$

Przykład zagadnienia, w którym w naturalny sposób pojawia się rozkład (2.6.29) poznamy w zadaniu 2.30.

b) Zgodnie z definicją

$$EK = \sum_{k=1}^{\infty} kp_k = p\sum_{k=1}^{\infty} kq^{k-1} = p\sum_{k=0}^{\infty} kq^{k-1}.$$

Bezwzględna zbieżność jest tu równoważna zwykłej zbieżności, ponieważ rozważana tu zmienna losowa ma tylko dodatnie punkty skokowe $x_k = k$. Szereg jest oczywiście zbieżny w przedziale $(-1, 1)$. Przy obliczaniu jego sumy posłużymy się różniczkowaniem szeregu potęgowego wyraz po wyrazie:

$$EK = p\sum_{k=0}^{\infty} kq^{k-1} = p\sum_{k=0}^{\infty}\frac{d}{dq}q^k = p\frac{d}{dq}\sum_{k=0}^{\infty} q^k =$$

$$= p\frac{d}{dq}\frac{1}{1-q} = p\frac{1}{(1-q)^2} = \frac{1}{p}. \tag{2.6.30}$$

Tak więc wartość przeciętna zmiennej losowej o rozkładzie (2.6.29) wynosi: $EK = 1/p$.

c) Posłużmy się wzorem (2.6.27):

$$D^2K = E(K^2) - (EK)^2 \equiv \alpha_2 - \alpha_1^2. \tag{1}$$

Drugi składnik obliczyliśmy w punkcie b):

$$\alpha_1^2 = \left(\frac{1}{p}\right)^2. \tag{2}$$

Pozostaje obliczyć drugi moment zwykły zmiennej losowej K. Zgodnie z definicją (2.6.10) mamy:

$$\alpha_2 = \sum_{k=1}^{\infty} k^2 p_k = p\sum_{k=1}^{\infty} k^2 q^{k-1}.$$

Aby móc postępować podobnie jak przy znajdowaniu wartości przeciętnej zauważmy najpierw, że

$$k^2 \equiv k(k-1) + k.$$

Zatem

$$\alpha_2 = pq\sum_{k=1}^{\infty} k(k-1)\,q^{k-2} + p\sum_{k=1}^{\infty} kq^{k-1}. \tag{3}$$

Drugi składnik jest znaną wartością przeciętną $\alpha_1 = \dfrac{1}{p}$. W celu obliczenia sumy szeregu w pierwszym składniku, dwukrotnie zróżniczkujemy wyraz po wyrazie odpowiedni szereg potęgowy:

$$\sum_{k=0}^{\infty} k(k-1)q^{k-2} = \sum_{k=0}^{\infty} \frac{d^2}{dq^2} q^k = \frac{d^2}{dq^2} \sum_{k=0}^{\infty} q^k =$$

$$= \frac{d^2}{dq^2} \frac{1}{1-q} = \frac{2}{(1-q)^3} = \frac{2}{p^3}.$$

Wracając kolejno do równości (3), (2), (1), otrzymujemy ostatecznie:

$$D^2K = pq\frac{2}{p^3} + \frac{1}{p} - \frac{1}{p^2} = \frac{q}{p^2}. \tag{2.6.31}$$

ZADANIE 2.17. Pracownik obsługuje m jednakowych maszyn ustawionych w rzędzie w równych odstępach o długości l. Prawdopodobieństwo tego, że maszyna będzie wymagała obsługi jest jednakowe dla każdej maszyny. Maszyny są obsługiwane w kolejności zgłoszeń (nie ma priorytetów). Obliczyć przeciętną długość EL drogi L jaką pokonuje pracownik przy przechodzeniu do następnej maszyny (być może tej samej) wymagającej obsługi, jeżeli nie wiadomo, która maszyna była obsługiwana.

Tablica 2.2.

i	k								
	1	2	3	4	...	k	...	$m-1$	m
1	0	1	2	3	...	$k-1$	...	$m-2$	$m-1$
2	1	0	1	2	...	$k-2$	...	$m-3$	$m-2$
3	2	1	0	1	...	$k-3$	...	$m-4$	$m-3$
4	3	2	1	0	...	$k-4$	...	$m-5$	$m-4$
...	...	...	...,	...	...	...	...	...	...
k	$k-1$	$k-2$	$k-3$	$k-4$	...	0	...	$m-k-1$	$m-k$
...	...	...	...	...	...	...	...	...	...
$m-1$	$m-2$	$m-3$	$m-4$	$m-5$	...	$m-k-1$	...	0	1
m	$m-1$	$m-2$	$m-3$	$m-4$	...	$m-k$	...	1	0

Rozwiązanie. Droga L jaką musi przebyć pracownik jest zależna od numeru maszyny aktualnie obsługiwanej i od numeru maszyny, która będzie obsługiwana w następnej kolejności. Niech więc a_{ik} oznacza drogę (w jednostkach l) jaką pokona pracownik obsługujący i-tą maszynę, gdy następna w kolejności do obsługi jest k-ta maszyna. Ponieważ i, k mogą przyjmować dowolne wartości $1, \ldots, m$, więc wszystkich możliwości jest m^2. Wymieniono je w tablicy 2.2. W tablicy tej występują tylko liczby $0, 1, \ldots, m-1$, przy tym liczba

0 występuje m-krotnie,
1 występuje $2(m-1)$-krotnie,
. . . .
k występuje $2(m-k)$-krotnie,
. . . .
$m-1$ występuje 2-krotnie.

Zgodnie z warunkiem zadania wszystkie możliwe sytuacje (jest ich m^2) są jednakowo prawdopodobne. Zatem otrzymujemy następującą funkcję prawdopodobieństwa zmiennej losowej L:

$$p_k = P(K=kl) = \begin{cases} \dfrac{1}{m}, & \text{gdy } k=0, \\ \dfrac{2(m-k)}{m^2}, & \text{gdy } k=1,\dots,m-1 \end{cases}$$

lub w postaci tabelki

x_k	0	l	$2l$	...	kl	...	$(m-1)l$
p_k	$\frac{1}{m}$	$\frac{2(m-1)}{m^2}$	$\frac{2(m-2)}{m^2}$	...	$\frac{2(m-k)}{m^2}$	...	$\frac{2}{m^2}$

Mając znalezioną funkcję prawdopodobieństwa zmiennej losowej L, możemy znaleźć jej wartość oczekiwaną EL:

$$EL = \sum_{k=0}^{m-1} klp_k = l\sum_{k=0}^{m-1} k\frac{2(m-k)}{m^2} = \frac{2l}{m^2}\Big(m\sum_{k=1}^{m-1} k - \sum_{k=1}^{m-1} k^2\Big),$$

$$m\sum_{k=1}^{m-1} k = m\tfrac{1}{2}(1+m-1)(m-1) = \tfrac{1}{2}m^2(m-1)$$

oraz

$$\sum_{k=1}^{m-1} k^2 = \tfrac{1}{6}(m-1)m(2m-1).$$

Ostatecznie otrzymujemy

$$EL = \frac{2l}{m^2}\left[\tfrac{1}{2}m^2(m-1) + \tfrac{1}{6}(m-1)m(2m-1)\right] = \frac{m^2-1}{3m}l.$$

Na przykład gdy $m=3$, wtedy $EL=\frac{8}{9}l$, jeśli $m=4$, to $EL=\frac{5}{4}l$.

ZADANIE 2.18. Zmienne losowe X i Y mają rozkłady zadane następującymi funkcjami prawdopodobieństwa:

a)

x_i	0	1	2	5
p_i	0,1	0,4	0,3	0,2

; b)

y_i	-1	1	3	5
p_i	0,2	0,3	0,3	0,2

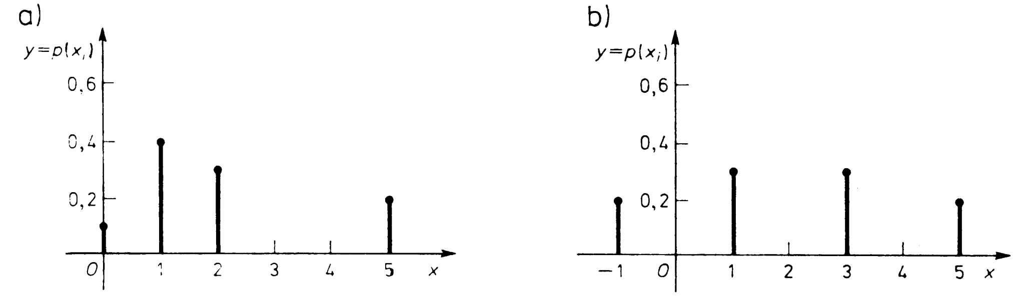
.

Naszkicować histogramy. Na podstawie ich wyglądu bez przeprowadzania rachunków ocenić znak współczynnika asymetrii γ_1 w obu przypadkach. Następnie obliczyć te współczynniki.

Rozwiązanie. Histogramy pokazano na rysunku 2.15. Z ich kształtu przypuszczamy, że współczynnik γ_1 dla przypadku:

a) będzie dodatni, bo znaczna część masy prawdopodobieństwa rozmieszczona jest na prawo w dużej odległości od najbardziej prawdopodobnego punktu skokowego $x_k=1$ (gdy $\gamma_1>0$, wtedy mówimy, że rozkład ma asymetrię prawą);

b) będzie równy zeru, ponieważ histogram jest figurą symetryczną względem punktu $x=2$.



Rys. 2.15. Histogramy funkcji rozkładu prawdopodobieństwa zmiennych z zadania 2.18

Tablica 2.3.

x_i	p_i	$x_i p_i$	$x_i - EX$	$(x_i - EX)^2$	$(x_i - EX)^3$	$(x_i - EX)^2 p_i$	$(x_i - EX)^3 p_i$
0	0,1	0	−2	4	−8	0,4	−0,8
1	0,4	0,4	−1	1	−1	0,4	−0,4
2	0,3	0,6	0	0	0	0	0
5	0,2	1	3	9	27	0,9	5,4
		2				1,7	4,2

W przypadku a) konieczne do obliczenia momentu centralnego μ_3 i wariancji μ_2 wyniki obliczeń pomocniczych podano w tablicy 2.3. Stąd na podstawie definicji (2.6.12) otrzymujemy:

$$\gamma_1 = \frac{\mu_3}{\mu_2\sqrt{\mu_2}} = \frac{4{,}2}{1{,}7\sqrt{1{,}7}} = \frac{4{,}2\sqrt{1{,}7}}{1{,}7^2} \approx 1{,}89,$$

a więc zgodnie z przewidywaniem $\gamma_1 > 0$.

W przypadku b), postępując podobnie, otrzymalibyśmy $\mu_3 = 0$ i tym samym $\gamma_1 = 0$. Ogólnie: moment centralny dowolnego rzędu nieparzystego zmiennej losowej o rozkładzie symetrycznym jest równy zeru (o ile istnieje).

Zadanie 2.19. Zmienna losowa X ma rozkład o gęstości:

$$f(x) = \begin{cases} 6x(1-x) & \text{dla } 0 < x < 1 \\ 0 & \text{dla pozostałych } x. \end{cases}$$

Obliczyć wartość przeciętną i wariancję: a) zmiennej losowej X, b) zmiennej losowej $Y = 2X - 1$.

Rozwiązanie. a) Korzystając z (2.6.1), mamy

$$EX = 6\int_0^1 x^2(1-x)\,dx = 6\left[\tfrac{1}{3}x^3 - \tfrac{1}{4}x^4\right]_0^1 = \tfrac{1}{2}.$$

Aby obliczyć Var X, stosujemy wzór (2.6.27):

$$D^2X = 6\int_0^1 x^3(1-x)\,dx - \tfrac{1}{4} = \tfrac{6}{20} - \tfrac{1}{4} = \tfrac{1}{20}.$$

b) Z własności (2.6.17) - (2.6.18) wartości przeciętnej:

$$EY=E(2X-1)=2EX-1=2\cdot\tfrac{1}{2}-1=0.$$

Podobnie, wykorzystując własności wariancji, otrzymamy

$$D^2Y=D^2(2X-1)=4D^2X=4\cdot\tfrac{1}{20}=\tfrac{1}{5}.$$

ZADANIE 2.20. Wytrzymałość X (w kg) przędzy lnianej w pewnej partii przędzy ma rozkład o gęstości

$$f(x)=\frac{1}{0{,}2\sqrt{2\pi}}\exp\left[-\frac{(x-1)^2}{0{,}08}\right].$$

Obliczyć wartość przeciętną wytrzymałości przędzy.

Rozwiązanie. Z definicji wartości zmiennej losowej ciągłej:

$$EX=\frac{1}{0{,}2\sqrt{2\pi}}\int_{-\infty}^{\infty}x\exp\left[-\frac{(x-1)^2}{0{,}08}\right]dx.$$

Dokonując w całce podstawienia $y=x-1$, otrzymujemy:

$$EX=\frac{1}{0{,}2\sqrt{2\pi}}\int_{-\infty}^{\infty}(y+1)\exp\left(-\frac{y^2}{0{,}08}\right)dy=\frac{1}{0{,}2\sqrt{2\pi}}\int_{-\infty}^{\infty}y\exp\left(-\frac{y^2}{0{,}08}\right)dy+1.$$

Z istnienia ostatniej całki i nieparzystości funkcji podcałkowej wynika zerowanie się pierwszego składnika, a więc

$$EX=1 \text{ kg}.$$

ZADANIE 2.21. Wykazać prawdziwość wzoru (2.6.28).

Rozwiązanie. $D^2X=E(X-EX)^2=E[(X-c)+(c-EX)]^2=E[(X-c)^2+2(X-c)(c-EX)+(c-EX)^2]=E(X-c)^2-2(c-EX)E(c-X)+(c-EX)^2=E(X-c)^2-(c-EX)^2$.

W szczególnym przypadku $c=0$ otrzymujemy wzór (2.6.27).

ZADANIE 2.22. Obliczyć trzeci moment zwykły zmiennej losowej X o gęstości prawdopodobieństwa:

$$f(x)=\begin{cases}4(x-1)^3 & \text{dla } 1<x<2,\\ 0 & \text{dla pozostałych } x.\end{cases}$$

Rozwiązanie. Zastosujemy wzór (2.6.10) dla $r=3$:

$$\alpha_3=4\int_1^2 x^3(x-1)^3dx=4\int_1^2 x^3(x^3-3x^2+3x-1)\,dx=$$

$$=4\int_1^2(x^6-3x^5+3x^4-x^3)\,dx=4\left[\tfrac{1}{7}x^7-\tfrac{1}{2}x^6+\tfrac{3}{5}x^5-\tfrac{1}{4}x^4\right]_1^2=$$

$$=4\left(\tfrac{128}{7}-\tfrac{1}{7}-32+\tfrac{1}{2}+\tfrac{96}{5}-\tfrac{3}{5}-4+\tfrac{1}{4}\right)=5\tfrac{34}{35}.$$

Zadanie 2.23. Niech zmienna losowa X ma rozkład o gęstości

$$f(x)=\begin{cases}\frac{1}{2}x & \text{dla } 0\leqslant x\leqslant 2,\\ 0 & \text{poza tym}.\end{cases}$$

Wyznaczyć modę i medianę zmiennej losowej X.

Rozwiązanie. Jak widać funkcja f nie osiąga maksimum lokalnego w żadnym punkcie przedziału $\langle 0, 2\rangle$, a więc rozkład ten nie ma mody. Dystrybuantą tego rozkładu jest funkcja:

$$F(x)=\begin{cases}0 & \text{dla } x\leqslant 0,\\ \int\limits_0^x \frac{1}{2}t\,dt=\frac{1}{4}x^2 & \text{dla } 0<x\leqslant 2,\\ 1 & \text{dla } x>2.\end{cases}$$

Korzystając ze wzoru (2.6.2), otrzymamy, że mediana jest rozwiązaniem równania: $\frac{1}{4}x^2=\frac{1}{2}$ należącym do przedziału $\langle 0, 2\rangle$, skąd $x=\sqrt{2}=x_{0,5}$.

Zadanie 2.24. Naszkicować wykres gęstości prawdopodobieństwa dla rozkładu: a) jednomodalnego, b) wielomodalnego, c) nie mającego mody, d) o asymetrii prawej ($\gamma_1>0$), e) o asymetrii lewej ($\gamma_1<0$), f) mającego przedział median.

Rozwiązanie. Patrz rys. 2.16.

Zadanie 2.25. Porównać współczynniki skupienia dwóch rozkładów o gęstościach:

$$f_1(x)=\begin{cases}\dfrac{1}{\sqrt{2}} & \text{dla } |x|<\frac{1}{2}\sqrt{2},\\ 0 & \text{dla pozostałych } x,\end{cases}$$

$$f_2(x)=\begin{cases}x+1 & \text{dla } -1<x\leqslant 0,\\ -x+1 & \text{dla } 0<x<1,\\ 0 & \text{dla pozostałych } x.\end{cases}$$

Rozwiązanie. Wykresy gęstości f_1, f_2 przedstawia rysunek 2.17. Wartości przeciętne w obydwu rozkładach są równe zeru:

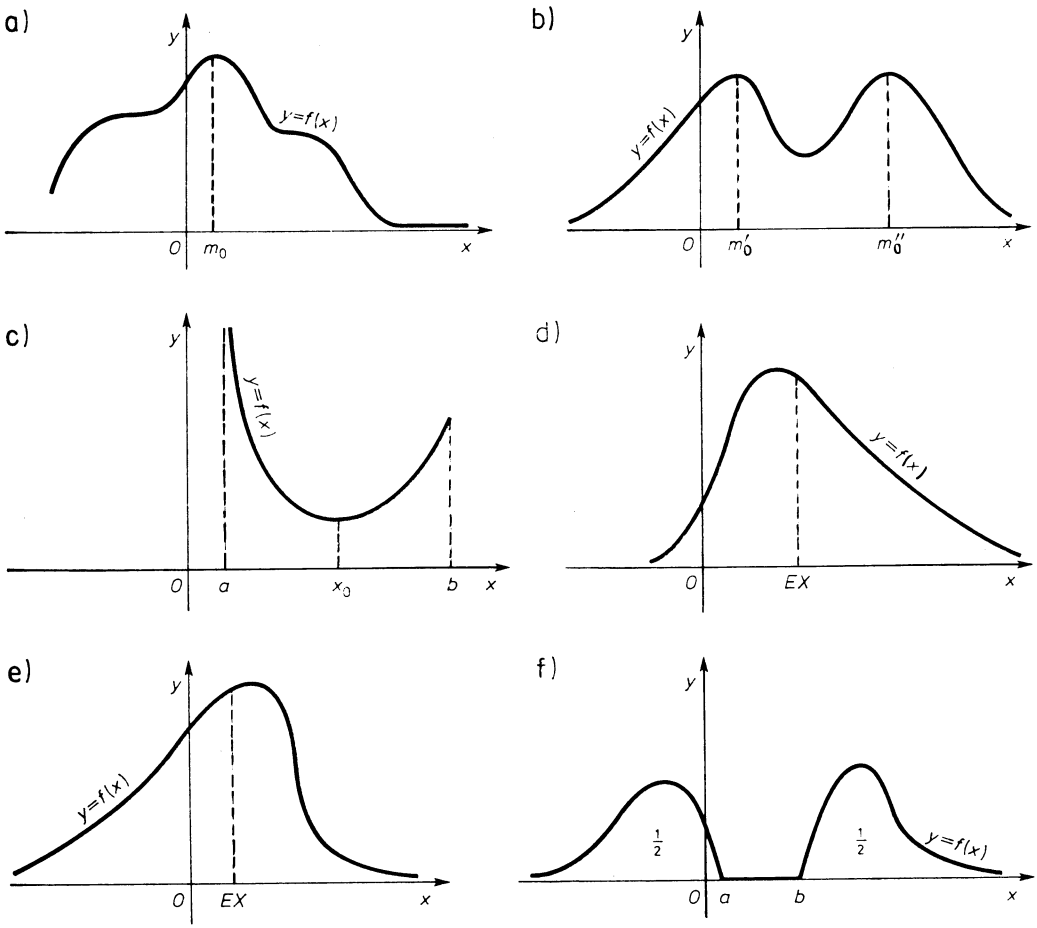
$$EX_1=EX_2=0.$$

Wariancja w rozkładzie o gęstości f_1 symetrycznej względem $x=0$ jest równa drugiemu momentowi zwykłemu:

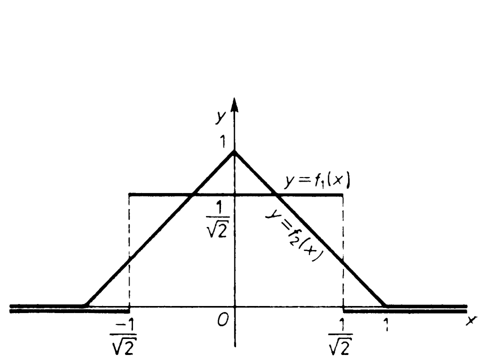
$$D^2X_1=\frac{2}{\sqrt{2}}\int\limits_0^{\sqrt{2}/2}x^2dx=\sqrt{2}\left[\frac{1}{3}x^3\right]_0^{\sqrt{2}/2}=\frac{1}{6};$$

podobnie czwarty moment centralny jest równy czwartemu momentowi zwykłemu:

$$\mu_4=\frac{2}{\sqrt{2}}\int\limits_0^{\sqrt{2}/2}x^4dx=\sqrt{2}\left[\frac{1}{5}x^5\right]_0^{\sqrt{2}/2}=\frac{1}{20}.$$



Rys. 2.16. a) Gęstość rozkładu jednomodalnego, b) gęstość rozkładu dwumodalnego, c) gęstość rozkładu antymodalnego, d) gęstość f rozkładu o asymetrii prawej ($\gamma_1 > 0$), e) gęstość f rozkładu o asymetrii lewej ($\gamma_1 < 0$), f) gęstość f rozkładu prawdopodobieństwa mającego przedział $\langle a, b \rangle$ median



Rys. 2.17. Gęstości rozkładów o tej samej wartości przeciętnej i wariancji, a różnych współczynnikach skupienia (zad. 2.25)

Stąd współczynnik skupienia dla pierwszego rozkładu jest równy

$$K_1=\tfrac{36}{20}=1{,}8\,.$$

Podobne obliczenia wykonujemy dla rozkładu o gęstości f_2 symetrycznej względem $x=0$:

$$D^2X_2=2\int\limits_0^1 x^2(1-x)\,dx=2\int\limits_0^1 (x^2-x^3)\,dx=2(\tfrac{1}{3}-\tfrac{1}{4})=\tfrac{1}{6},$$

$$\mu_4=2\int\limits_0^1 x^4(1-x)\,dx=2\int\limits_0^1 (x^4-x^5)\,dx=2(\tfrac{1}{5}-\tfrac{1}{6})=\tfrac{1}{15},$$

a więc

$$K_2=\tfrac{36}{15}=2{,}4\,.$$

Z porównania współczynników skupienia obydwu rozkładów i z wykresów gęstości widać, że mimo równych wartości przeciętnych i wariancji tych rozkładów, skupienie wartości poszczególnych zmiennych wokół ich wartości przeciętnych jest różne, przy czym dla rozkładu o gęstości f_1 współczynnik skupienia jest mniejszy, a więc skupienie wartości zmiennej X_1 wokół przeciętnej jest mniejsze niż dla drugiego rozkładu.

ZADANIE 2.26. Obliczyć współczynnik asymetrii rozkładu o gęstości

$$f(x)=\begin{cases}12x(1-x)^2 & \text{dla } 0<x<1\,,\\ 0 & \text{dla pozostałych } x\,.\end{cases}$$

Rozwiązanie. Wartość przeciętna zmiennej losowej X

$$EX=12\int\limits_0^1 x^2(1-x)^2dx=12\int\limits_0^1 (x^2-2x^3+x^4)\,dx=\tfrac{2}{5}\,.$$

Wariancję obliczamy ze wzoru (2.6.27):

$$D^2X=12\int\limits_0^1 x^3(1-x)^2dx-\tfrac{4}{25}=12\int\limits_0^1 (x^3-2x^4+x^5)\,dx-0{,}16=$$
$$=0{,}2-0{,}16=0{,}04\,.$$

Stąd $D^2X=0{,}04$, odchylenie zaś standardowe $\sigma=0{,}2$.

Trzeci moment centralny otrzymamy, obliczając całkę

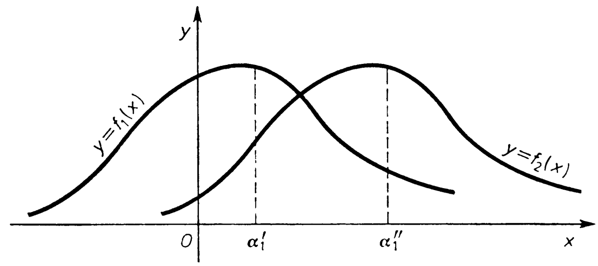
$$\mu_3=12\int\limits_0^1 (x-0{,}4)^3x(1-x)^2dx=0{,}13\,.$$

Stąd współczynnik asymetrii

$$\gamma_1=\frac{0{,}13}{0{,}2^3}=16{,}25>0$$

jest dodatni, a więc mamy rozkład o asymetrii prawej.

Rys. 2.18. Gęstości dwóch rozkładów o tej samej wariancji, a różnych współczynnikach zmienności



ZADANIE 2.27. Wzrost ludzi w pewnej grupie zawodowej jest zmienną losową X typu ciągłego o wartości przeciętnej $EX=1{,}70$ m i odchyleniu standardowym $\sigma_X=0{,}05$ m, masa zaś – zmienną losową Y typu ciągłego, dla której $EY=65$ kg i $\sigma_Y=5$ kg. Porównać współczynniki zmienności dla obydwu zmiennych losowych.

Rozwiązanie.

$$v_1=\frac{\sigma_X}{EX}=\frac{0{,}05}{1{,}70}=\frac{1}{34}\approx 0{,}03, \qquad v_2=\frac{\sigma_Y}{EY}=\frac{5}{65}=\frac{1}{13}\approx 0{,}08.$$

Współczynnik zmienności v_2 jest większy od v_1, to znaczy, że badana grupa ludzi jest bardziej zróżnicowana względnie (względem wartości przeciętnej) na drugą cechę – masę, niż na pierwszą tj. wzrost.

2.7. NIEKTÓRE ROZKŁADY SKOKOWE

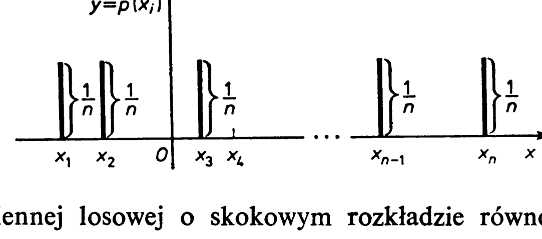
2.7.1. Skokowy rozkład równomierny. Mówimy, że zmienna losowa X ma *skokowy* (*dyskretny*) *rozkład równomierny*, jeżeli jej funkcja prawdopodobieństwa jest postaci:

x_i	x_1	x_2	...	x_n
p_i	$1/n$	$1/n$	...	$1/n$

(2.7.1)

Jest to zatem zmienna losowa typu skokowego mająca skończoną liczbę punktów skokowych x_i i równe skoki p_i. Histogram funkcji prawdopodobieństwa podano na rysunku 2.19.

Rys. 2.19. Histogram funkcji prawdopodobieństwa rozkładu równomiernego



Wartość przeciętna i wariancja zmiennej losowej o skokowym rozkładzie równomiernym, zgodnie z (2.6.1) i (2.6.4), wyrażają się równościami:

$$EX=\frac{1}{n}\sum_{i=1}^{n} x_i, \qquad D^2X=\frac{1}{n}\sum_{i=1}^{n}(x_i-EX)^2. \tag{2.7.2}$$

Z wzorów tych widać, że: wartość oczekiwaną i wariancję zmiennej losowej o skokowym rozkładzie równomiernym można traktować jako średnią arytmetyczną $\bar{x}$ i wariancję s^2 zaobserwowanej próbki $x_1, \ldots, x_n$ z populacji generalnej (część II, wzory (1.3.1) i (1.5.1)).

2.7.2. Rozkład jednopunktowy. Mówimy, że zmienna losowa X typu skokowego ma *rozkład jednopunktowy* (*rozkład zdegenerowany*), jeżeli jej funkcja prawdopodobieństwa jest postaci:

$$\begin{array}{c|c} x_i & x_1 \\ \hline p_i & 1 \end{array} \tag{2.7.3}$$

Ma więc ona tylko jeden punkt skokowy, w którym skupiona jest cała masa prawdopodobieństwa. Jest ona przypadkiem szczególnym (gdy $n=1$) zmiennej losowej o skokowym rozkładzie równomiernym (2.7.1).

Wartość oczekiwana i wariancja zmiennej losowej X o rozkładzie jednopunktowym dane są równościami:

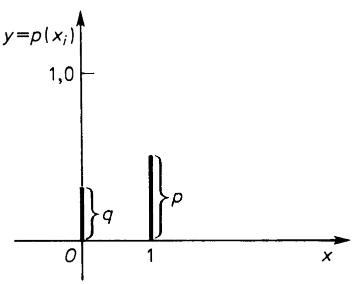
$$EX = x_1, \quad D^2X = 0. \tag{2.7.4}$$

Zmienna losowa o rozkładzie jednopunktowym jest modelem dla wielkości, które przy bliższej analizie okazują się wielkościami zdeterminowanymi. Pozwala ona zaliczyć zmienne zdeterminowane do zmiennych losowych.

2.7.3. Rozkład zero-jedynkowy. Mówimy, że zmienna losowa X typu skokowego ma *rozkład zero-jedynkowy z parametrem p*, $p \in (0, 1)$, jeżeli jej funkcja prawdopodobieństwa jest postaci:

$$\begin{array}{c|c|c} x_i & 0 & 1 \\ \hline p_i & q & p \end{array}, \quad q = 1-p. \tag{2.7.5}$$

Ma więc ona tylko dwa punkty skokowe $x_1 = 0$ i $x_2 = 1$. Histogram funkcji prawdopodobieństwa pokazano na rys. 2.20.



Rys. 2.20. Histogram funkcji prawdopodobieństwa rozkładu zero-jedynkowego

Wartość przeciętna, wariancja oraz moment centralny μ_3 zmiennej losowej X o rozkładzie zero-jedynkowym z parametrem p wyrażają się wzorami:

$$EX = p, \quad D^2X = pq, \quad \mu_3 = pq(1-2p). \tag{2.7.6}$$

Zmienna losowa o rozkładzie (2.7.5) jest związana z doświadczeniem losowym, w którym wyniki są dwojakiego rodzaju: jako mające interesującą nas cechę (wtedy kładziemy $X(\omega)=1$, i nie mające tej cechy (przyjmujemy wtedy $X(\omega)=0$).

2.7.4. Rozkład dwumianowy. Mówimy, że zmienna losowa K typu skokowego ma *rozkład dwumianowy* (rozkład *binomialny*, *rozkład Bernoulliego*) z parametrami (n, p), $n \in N$, $0<p<1$, jeżeli jej funkcja prawdopodobieństwa $p_k=P(k, n, p)=P(K=k)$ jest postaci:

$$P(k,n,p)=\binom{n}{k}p^k q^{n-k}, \quad k=0,1,\ldots,n, \tag{2.7.7}$$

gdzie $q=1-p$.

Przyjmuje więc ona z dodatnimi prawdopodobieństwami $n+1$ wartości: $0, 1, \ldots, n$. Wśród nich jest jedna albo dwie *wartości najbardziej prawdopodobne*:

– gdy $(n+1)p$ jest liczbą całkowitą, wtedy tymi wartościami są liczby

$$k_1=(n+1)p-1, \quad k_2=(n+1)p, \tag{2.7.8a}$$

– jeśli $(n+1)p$ nie jest liczbą całkowitą, to istnieje jedna najbardziej prawdopodobna wartość k_0 dana wzorem

$$k_0=[(n+1)p], \tag{2.7.8b}$$

gdzie $[x]$ oznacza największą liczbę całkowitą nie przekraczającą liczby x, zwaną częścią całkowitą liczby x (całością z x).

Wartość przeciętna, wariancja i moment centralny μ_3 zmiennej losowej K o rozkładzie dwumianowym z parametrami (n, p) wyrażają się wzorami:

$$EK=np, \quad D^2K=npq, \quad \mu_3=npq(1-2p). \tag{2.7.9}$$

Zmienną losową K o rozkładzie dwumianowym z parametrami (n, p) można interpretować jako możliwą liczbę sukcesów (czyli realizacji pewnego zdarzenia A) w dowolnej kolejności w n doświadczeniach $D_1, \ldots, D_n$ przeprowadzonych zgodnie z następującymi warunkami schematu Bernoulliego: 1° doświadczenia $D_1, \ldots, D_n$ są niezależne oraz 2° prawdopodobieństwo p obserwowanego w poszczególnych doświadczeniach D_i zdarzenia A jest jednakowe w każdym doświadczeniu D_i. Zatem

$$P(\text{zdarzenie } A \text{ zrealizuje się } k \text{ razy})=\binom{n}{k}p^k q^{n-k}.$$

Zmienna losowa K o rozkładzie dwumianowym wiąże się ze zmiennymi losowymi o rozkładach: Poissona (2.7.16), normalnym (2.8.22) oraz ze zmienną losową o rozkładzie zero-jedynkowym w następujący sposób:

– gdy $n=1$, wtedy zmienna losowa K staje się zmienną losową o rozkładzie zero-jedynkowym (2.7.5),

– jeżeli $n>1$, to zmienna losowa K o rozkładzie dwumianowym z parametrami (n, p) jest sumą n niezależnych zmiennych losowych X_i o tym samym rozkładzie zero-jedynkowym z parametrem p:

$$K=X_1+\ldots+X_n. \tag{2.7.10}$$

Dla rozkładu dwumianowego zachodzi tzw. *twierdzenie o dodawaniu względem parametru n*: *Jeżeli zmienna losowa K_1 ma rozkład dwumianowy z parametrami (n_1, p) i K_2 jest zmienną losową o tym rozkładzie z parametrami (n_2, p) oraz zmienne losowe K_1 i K_2 są niezależne, to również zmienna losowa $K=K_1+K_2$ ma rozkład dwumianowy z parametrami (n, p), gdzie $n=n_1+n_2$.*

2.7.5. Rozkład hipergeometryczny. Mówimy, że skokowa zmienna losowa K ma *rozkład hipergeometryczny z parametrami* (N, M, n), gdzie N, M, n — liczby naturalne oraz $M, n \leqslant N$, jeśli jej funkcja prawdopodobieństwa $p_k \equiv P(k; N, M, n) = P(K=k)$ jest postaci:

$$P(k; N, M, n) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad \begin{array}{l} k=0,1,\dots,n,\ n\leqslant N, \\ k\leqslant M,\ k\leqslant n,\ n-k\leqslant N-M. \end{array} \tag{2.7.11}$$

Wartość przeciętna, wariancja i moment centralny μ_3 zmiennej losowej K o tym rozkładzie wyrażają się za pomocą jego parametrów następującymi wzorami:

$$\begin{gathered} EK = np, \qquad D^2K = npq\frac{N-n}{N-1}, \\ \mu_3 = npq(1-2p)\frac{(N-n)(N-2n)}{(N-1)(N-2)}, \quad \text{gdzie} \quad p=\frac{M}{N}, \quad q=1-p. \end{gathered} \tag{2.7.12}$$

Zmienna losowa K o rozkładzie hipergeometrycznym z parametrami (N, M, n) ma następującą interpretację: jest ona możliwą liczbą elementów mających wyróżnioną cechę A wśród n elementów wylosowanych bez zwrotu z populacji N elementów, wśród których przed rozpoczęciem losowania znajdowało się M elementów mających tę cechę A.

Gdy $N \to \infty$, $M \to \infty$, ale tak, że $\dfrac{M}{N} \to p$, $0<p<1$, wtedy $P(k; N, M, n) \to P(k; n, p)$,

czyli: rozkład hipergeometryczny z parametrami (N, M, n) jest zbieżny do rozkładu dwumianowego z parametrami (n, p): wynika stąd następujące przybliżenie rozkładu hipergeometrycznego rozkładem dwumianowym:

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \approx \binom{n}{k} p^k (1-p)^{n-k}. \tag{2.7.13}$$

2.7.6. Ujemny rozkład dwumianowy. Mówimy, że skokowa zmienna losowa K ma *ujemny rozkład dwumianowy* (*rozkład Pascala*) *z parametrami* (ν, p), $\nu \in N$, $0<p<1$, jeśli jej funkcja prawdopodobieństwa $p_k \equiv P(k; \nu, p) \equiv P(K=k)$ jest postaci:

$$P(k; \nu, p) = \binom{k-1}{\nu-1} p^\nu q^{k-\nu}, \quad k=\nu, \nu+1, \dots, \quad q=1-p. \tag{2.7.14}$$

Jej wartość przeciętna, wariancja i moment centralny μ_3 dane są wzorami:

$$EK=\frac{v}{p}, \quad D^2K=\frac{vq}{p^2}, \quad \mu_3=\frac{vq(2-p)}{p^3}. \tag{2.7.15}$$

Zmienną losową K o rozkładzie (2.7.14) można interpretować jako liczbę doświadczeń przeprowadzonych według schematu Bernoulliego, z prawdopodobieństwem p sukcesu w pojedynczym doświadczeniu, potrzebną do uzyskania v sukcesów.

Jeśli dodatkowo założyć, że jedno doświadczenie przeprowadza się w jednostce czasu, to rozważana tu zmienna losowa K jest czasem oczekiwania na v sukcesów.

Dla ujemnego rozkładu dwumianowego, podobnie jak dla rozkładu dwumianowego, zachodzi twierdzenie o dodawaniu względem parametru v .

Ujemny rozkład dwumianowy, gdy $v=1$, nosi nazwę *rozkładu geometrycznego z parametrem p.*

2.7.7. Rozkład Poissona. Mówimy, że skokowa zmienna losowa K ma *rozkład Poissona* z parametrem λ, $\lambda>0$, jeśli jej funkcja prawdopodobieństwa $p_k \equiv P(k;\lambda)=P(K=k)$ jest postaci:

$$P(k;\lambda)=e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k\in N_0=N\cup\{0\}. \tag{2.7.16}$$

Przyjmuje więc ona z dodatnimi prawdopodobieństwami przeliczalną liczbę wartości. (Zauważmy, że ułamek $\lambda^k/k!$ jest $(k+1)$-szym wyrazem rozwinięcia funkcji e^λ w szereg Maclaurina.)

Jej wartość oczekiwana, wariancja i moment centralny μ_3 są równe parametrowi λ:

$$EK=\lambda, \quad D^2K=\lambda, \quad \mu_3=\lambda. \tag{2.7.17}$$

Rozkład Poissona zadany funkcją prawdopodobieństwa (2.7.16) wiąże się z rozkładem dwumianowym (2.7.7) następującym twierdzeniem:

Jeżeli $K_1, \ldots, K_n, \ldots$ *jest ciągiem zmiennych losowych o rozkładzie dwumianowym odpowiednio z parametrami* $(1, p_1), \ldots, (n, p_n), \ldots$ *oraz* $np_n \to \lambda$, $\lambda>0$, *gdy* $n\to\infty$, *to*

$$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k\in N_0, \tag{2.7.18}$$

czyli: *ciąg rozkładów dwumianowych jest zbieżny do rozkładu Poissona z parametrem* λ.

Dla dużych n wynika stąd następujące *przybliżenie Poissona rozkładu dwumianowego*:

$$\binom{n}{k}p^kq^{n-k}\approx e^{-\lambda}\frac{\lambda^k}{k!}, \qquad \lambda=np. \tag{2.7.19}$$

Przybliżenie to jest dla celów praktycznych wystarczająco dokładne, gdy: $n\geqslant 50$, $p\leqslant 0{,}1$, $np\leqslant 10$.

Rozkład Poissona jest rozkładem granicznym również dla rozkładu hipergeometrycznego (2.7.11). Mianowicie: gdy $N\to\infty$, $M\to\infty$, $n\to\infty$, ale tak, że $\frac{M}{N}\to 0$ i $n\frac{M}{N}\to\lambda$, $\lambda>0$, wtedy $P(k;\, N,\, M,\, n)\to P(k;\,\lambda)$. Wynika stąd następujące *przybliżenie Poissona rozkładu hipergeometrycznego*:

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\approx e^{-\lambda}\frac{\lambda^k}{k!}. \tag{2.7.20}$$

Dla rozkładu Poissona zachodzi twierdzenie o dodawaniu. Sformułowanie tego twierdzenia pozostawiamy Czytelnikowi. Prawdziwe jest również twierdzenie odwrotne:

Jeżeli skokowe zmienne losowe X_1, X_2, przyjmujące z dodatnimi prawdopodobieństwami wartości 0, 1, 2, ..., *są niezależne oraz ich suma $X=X_1+X_2$ ma rozkład Poissona z parametrem λ, to zmienne te również mają rozkład Poissona odpowiednio z parametrami λ_1, λ_2 takimi, że $\lambda_1+\lambda_2=\lambda$* (udowodnił je Rajkow).

2.7.8. Zadania rozwiązane.

ZADANIE 2.28. Prawdopodobieństwo nieprzekroczenia przez pewien zakład pracy dobowego limitu zużycia energii elektrycznej (bez wyłączenia niektórych maszyn) wynosi $p=0{,}8$. Niech K oznacza możliwą liczbę dni w 5-dniowym tygodniu pracy, w którym nie nastąpiło przekroczenie limitu (przy założeniu, że wyłączanie urządzeń nie wchodziło w rachubę). Wyznaczyć: a) funkcję prawdopodobieństwa zmiennej losowej K i jej histogram, b) dystrybuantę i jej wykres, c) prawdopodobieństwo, że co najmniej w trzech dniach limit nie zostanie przekroczony, d) najbardziej prawdopodobną liczbę dni, w których limit nie zostanie przekroczony i prawdopodobieństwo takiej liczby dni, e) takie p, aby najbardziej prawdopodobna liczba dni, w których limit nie zostanie przekroczony, wynosiła: e_1) 4 albo 5 dni, e_2) 5 dni, f) wartość przeciętną i wariancję zmiennej losowej K, g) współczynnik asymetrii rozkładu zmiennej losowej K.

Rozwiązanie. Ponieważ: 1° obserwujemy tu $n=5$ razy zajście tego samego zdarzenia A = nieprzekroczenie w ciągu doby limitu zużycia energii, 2° prawdopodobieństwo $p=0{,}8$ zdarzenia A w każdej spośród tych pięciu obserwacji jest takie samo oraz 3° obserwacje te są doświadczeniami niezależnymi, więc zmienna losowa K — jako liczba sukcesów w 5 doświadczeniach Bernoulliego — ma rozkład dwumianowy (2.7.7) z parametrami (5; 0, 8):

$$p_k\equiv P(K=k)=P(k\,;5\,;0,8)=\binom{5}{k}0{,}8^k 0{,}2^{5-k},\qquad k=0,1,2,3,4,5.$$

Pozostaje obliczyć prawdopodobieństwo p_k dla poszczególnych wartości k, np. dla $k=3$ otrzymujemy:

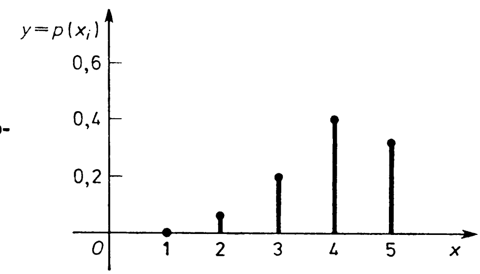
$$p_3\equiv P(K=3)=P(3\,;5\,;0,8)=\binom{5}{3}0{,}8^3 0{,}2^2=10\cdot 0{,}512\cdot 0{,}04=0{,}2048.$$

Rezultaty rachunków dla wszystkich k zawiera tabelka

k	0	1	2	3	4	5
p_k	0,00032	0,0064	0,0512	0,2048	0,4096	0,32768

. (1)

Histogram funkcji prawdopodobieństwa danej tabelką (1) podano na rys. 2.21.



Rys. 2.21. Histogram funkcji rozkładu prawdopodobieństwa z zadania 2.28

b) Korzystając z definicji dystrybuanty ((2.1.4) i (2.2.7)), otrzymujemy:

x	$(-\infty, 0\rangle$	$(0, 1\rangle$	$(1, 2\rangle$	$(2, 3\rangle$	$(3, 4\rangle$	$(4, 5\rangle$	$(5, \infty)$
$F(x)$	0	0,00032	0,00672	0,05792	0,26272	0,67232	1

c) Niech $p_{\geqslant 3}$ oznacza szukane prawdopodobieństwo. Wykorzystując znajomość funkcji prawdopodobieństwa, otrzymujemy

$$p_{\geqslant 3} = p_3 + p_4 + p_5 = 0{,}2048 + 0{,}4096 + 0{,}32768 = 0{,}94208 \approx 0{,}94.$$

Dla znalezienia prawdopodobieństwa $p_{\geqslant 3}$ można też wykorzystać dystrybuantę:

$$p_{\geqslant 3} = P(K \geqslant 3) = 1 - P(K < 3) = 1 - F(3) = 1 - 0{,}05792 = 0{,}94208 \approx 0{,}94.$$

d) Ponieważ $(n+1)p = 6 \cdot 0{,}8 = 4{,}8$ nie jest liczbą całkowitą, więc istnieje tylko jedna najbardziej prawdopodobna wartość i, zgodnie z (2.7.8b), jest nią liczba $k_0 = [4{,}8] = 4$. Z tabelki (1) odczytujemy, że $P(K=4) = p_4 = 0{,}4096$. Zrozumiałe, że wartość najbardziej prawdopodobną mogliśmy odczytać bezpośrednio z tej tabelki.

e_1) W myśl wzorów (2.7.8a) warunki zadania będą spełnione, gdy $k_2 = (5+1)p = 5$, tj. gdy $p = \frac{5}{6}$. Drugą najbardziej prawdopodobną wartością przy $p = \frac{5}{6}$ jest $k_1 = k_2 - 1 = 4$.

e_2) Istnieć będzie tylko jedna najbardziej prawdopodobna i będzie to liczba $k_0 = 5$, gdy

$$(5+1)p = a, \quad \text{gdzie} \quad 5 < a < 6,$$

tj. gdy $\frac{5}{6} < p < 1$.

f) Na podstawie wzorów (2.7.9) otrzymujemy:

$$EK = np = 5 \cdot 0{,}8 = 4 \quad \text{oraz} \quad D^2K = npq = 5 \cdot 0{,}8 \cdot 0{,}2 = 0{,}8.$$

g) Na podstawie kształtu histogramu przypuszczamy, że rozkład zmiennej losowej ma asymetrię lewą, tj. $\gamma_1 < 0$. Sprawdzimy, że rzeczywiście tak jest. Na podstawie definicji (2.6.12) współczynnika asymetrii i wzoru (2.7.9) na trzeci moment centralny otrzymujemy:

$$\gamma_1 = \frac{\mu_3}{\mu_2\sqrt{\mu_2}} = \frac{npq(1-2p)}{npq\sqrt{npq}} = \frac{1-2p}{\sqrt{npq}}. \tag{2.7.21}$$

Stąd widać, że gdy: 1° $p=\frac{1}{2}$, wtedy $\gamma_1=0$ i rozkład dwumianowy jest symetryczny, 2° $p<\frac{1}{2}$, wtedy $\gamma_1>0$ i rozkład ma asymetrię prawą, 3° $p>\frac{1}{2}$, wtedy $\gamma_1<0$ i rozkład ma asymetrię lewą. W rozważanym zadaniu $p=0{,}8>\frac{1}{2}$, więc rozkład ma asymetrię lewą i przy tym

$$\gamma_1 = \frac{1-1{,}6}{\sqrt{5\cdot 0{,}8\cdot 0{,}2}} \approx -0{,}67.$$

ZADANIE 2.29. Przeprowadzamy niezależne doświadczenia $D_1, D_2, \ldots$ i w każdym z nich obserwujemy zajście określonego zdarzenia A. Prawdopodobieństwo zajścia tego zdarzenia w poszczególnych doświadczeniach jest takie samo i wynosi $p=0{,}6$. Niech K oznacza liczbę doświadczeń, gdy przeprowadzamy je aż do pojawienia się sukcesu po raz pierwszy, jednak nie dłużej niż $n=5$ razy. Dla tak określonej zmiennej losowej wyznaczyć: a) funkcję prawdopodobieństwa i jej histogram, b) dystrybuantę i jej wykres, c) wartość przeciętną.

Rozwiązanie. a) Niech A_k oznacza zajście po raz pierwszy zdarzenia A w k-tym doświadczeniu, $k=1, \ldots, n$. Uwzględniając założoną niezależność doświadczeń i stałość prawdopodobieństwa sukcesu w tych doświadczeniach, otrzymujemy:

$$p_1 = P(K=1) = P(A_1) = p,$$

$$p_k = P(K=k) = P(A'_1 A'_2 \ldots A'_{k-1} A_k) = q^{k-1} p, \quad k=2, \ldots, n-1,$$

$$p_n = P(K=n) = P(A'_1 A'_2 \ldots A'_{n-1}) = q^{n-1}$$

albo inaczej:

$$p_n = P(A'_1 A'_2 \ldots A'_{n-1} A'_n \cup A'_1 A'_2 \ldots A'_{n-1} A_n) = q^n + q^{n-1} p = q^{n-1}(q+p) = q^{n-1}.$$

Ostatnie prawdopodobieństwo można wyznaczyć również z warunku unormowania:

$$p_n = 1-(p_1+p_2+\ldots+p_{n-1}) = 1-(p+qp+q^2p+\ldots+q^{n-2}p) =$$
$$= 1 - p\frac{1-q^{n-1}}{1-q} = q^{n-1}.$$

Wykorzystaliśmy tu wzór na sumę skończonej liczby wyrazów ciągu geometrycznego. Szukana funkcja prawdopodobieństwa jest więc postaci:

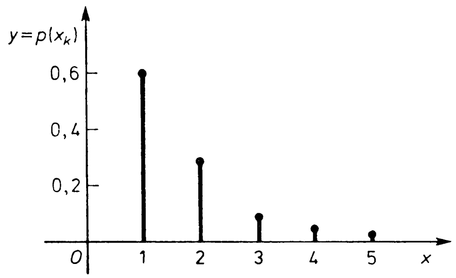
k	1	2	3	...	k	...	$n-1$	n
p_k	p	qp	q^2p	...	$q^{k-1}p$	...	$q^{n-2}p$	q^{n-1}

(1)

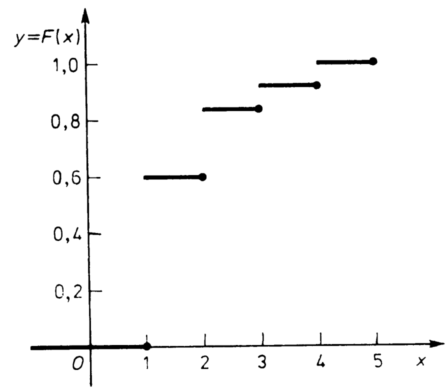
Podstawiając $p=0{,}6$ i $n=5$, otrzymujemy (rys. 2.22):

k	1	2	3	4	5
p_k	0,6	0,24	0,096	0,0384	0,0256

(2)



Rys. 2.22. Histogram funkcji rozkładu prawdopodobieństwa z zadania 2.29



Rys. 2.23. Dystrybuanta zmiennej losowej z zadania 2.29

b) Przy wyznaczaniu dystrybuanty F zmiennej losowej K korzystamy z jej definicji, funkcji prawdopodobieństwa (1) oraz wzoru na sumę skończonej liczby wyrazów ciągu geometrycznego i otrzymujemy:

x	$(-\infty, 1\rangle$	$(1, 2\rangle$	$(2, 3\rangle$	$(3, 4\rangle$	...	$(n-2, n-1\rangle$	$(n-1, n\rangle$	(n, ∞)
$F(x)$	0	$1-q$	$1-q^2$	$1-q^3$	...	$1-q^{n-2}$	$1-q^{n-1}$	1

Dla $p=0{,}6$ i $n=5$ dystrybuanta F przyjmuje następującą postać (rys. 2.23):

x	$(-\infty, 1\rangle$	$(1, 2\rangle$	$(2, 3\rangle$	$(3, 4\rangle$	$(4, 5\rangle$	$(5, \infty)$
$F(x)$	0	0,6	0,84	0,936	0,9744	1

c) Na podstawie definicji wartości przeciętnej i tabeli (1) mamy

$$EK = \sum_{k=1}^{n} kp_k = \sum_{k=1}^{n-1} kpq^{k-1} + nq^{n-1}.$$

Aby obliczyć pierwszy składnik sumy po prawej stronie ostatniej równości, stosujemy wzór na sumę skończonej liczby wyrazów ciągu geometrycznego i regułę różniczkowania sumy:

$$\sum_{k=1}^{n-1} kq^{k-1} = \sum_{k=0}^{n-1} kq^{k-1} = \sum_{k=0}^{n-1} \frac{d}{dq} q^k = \frac{d}{dq} \sum_{k=0}^{n-1} q^k = \frac{d}{dq} \frac{1-q^n}{1-q} =$$

$$= \frac{1}{p^2}(1-q^n-npq^{n-1}).$$

Zatem

$$EK = \frac{1}{p}(1-q^n-npq^{n-1}) + nq^{n-1} = \frac{1}{p}(1-q^n).$$

Stąd dla $p=0{,}6$ i $n=5$ otrzymujemy

$$EK = \frac{1}{0{,}6}(1-0{,}4^5) = 1{,}6496 \approx 1{,}6.$$

ZADANIE 2.30. Prawdopodobieństwo, że w rezultacie przeprowadzenia doświadczenia zajdzie oczekiwany wynik, wynosi $p=0{,}6$. Doświadczenie to możemy powtarzać dowolną liczbę razy. Powtarzamy je aż do otrzymania oczekiwanego wyniku. Oznaczając przez K liczbę powtórzeń tego doświadczenia (i zakładając niezależność powtórzeń) wyznaczyć: a) funkcję prawdopodobieństwa zmiennej losowej K i histogram rozkładu, b) dystrybuantę, c) prawdopodobieństwo, że liczba doświadczeń będzie: c_1) mniejsza niż 3, c_2) nie mniejsza niż 3, c_3) liczbą z przedziału $\langle 2, 5)$, d) wartość przeciętną, wariancję i odchylenie standardowe, e) współczynnik asymetrii.

Rozwiązanie. Zbiorem punktów skokowych zmiennej losowej K jest zbiór $\boldsymbol{N}$ wszystkich liczb naturalnych.

a) Rozumując analogicznie jak w zadaniu 2.29, znajdujemy skoki p_k:

$$p_k \equiv P(K=k)=q^{k-1}p, \quad q=1-p, \quad k\in \boldsymbol{N}. \tag{1}$$

Rozkład prawdopodobieństwa rozważanej tu zmiennej losowej K zadany funkcją prawdopodobieństwa (1), nazywa się *rozkładem geometrycznym z parametrem p*. Spotkaliśmy go już w zadaniu 2.16. Stanowi on przypadek szczególny (gdy $\nu=1$) ujemnego rozkładu dwumianowego (2.7.14). W przypadku $p=0{,}6$

$$p_k=0{,}6\cdot 0{,}4^{k-1}, \quad k\in \boldsymbol{N}. \tag{2}$$

b) Dystrybuantę F zmiennej losowej K znajdujemy w taki sposób jak (dla $k<n$) dystrybuantę w zadaniu 2.29 i otrzymujemy w postaci tabeli:

x	$(-\infty, 1\rangle$	$(1, 2\rangle$	$(2, 3\rangle$	...	$(k-1, k\rangle$	...
$F(x)$	0	$1-q$	$1-q^2$	...	$1-q^{k-1}$	...

lub wzoru

$$F(x)=\begin{cases} 0 & \text{dla } x<1, \\ 1-q^{x-1}=1-0{,}4^{x-1} & \text{dla } x\in \boldsymbol{N}, \\ 1-q^{[x]}=1-0{,}4^{[x]} & \text{dla } x\in \boldsymbol{R}_+\setminus \boldsymbol{N}, \end{cases} \tag{2.7.22}$$

gdzie $[x]$ oznacza część całkowitą liczby x, tj. największą liczbę całkowitą nie przekraczającą x.

c_1) Szukane prawdopodobieństwo jest po prostu wartością dystrybuanty w punkcie $x=3$:

$$P(K<3)=F(3)=1-q^2=1-0{,}4^2=0{,}84.$$

c_2) Ponieważ suma prawdopodobieństw zdarzeń przeciwnych jest równa jedności, więc

$$P(K\geqslant 3)=1-P(K<3)=1-F(3)=1-0{,}84=0{,}16.$$

c_3) Szukane prawdopodobieństwo jest, zgodnie z własnością dystrybuanty daną równością (2.1.5), przyrostem dystrybuanty między punktami 2 i 5:

$$P(2\leqslant K<5)=F(5)-F(2)=1-q^4-(1-q)=q-q^4=$$
$$=0{,}4-0{,}4^4=0{,}3744.$$

d) Wartość przeciętną α_1 i wariancję σ^2 przy dowolnym p obliczono w zadaniu 2.16. Podstawiając $p=0{,}6$, otrzymujemy:

$$\alpha_1=\frac{1}{p}=\frac{1}{0{,}6}\approx 1{,}67, \quad \sigma^2=\frac{q}{p^2}=\frac{0{,}4}{0{,}6^2}\approx 1{,}11, \quad \sigma=\frac{\sqrt{q}}{p}=\frac{\sqrt{0{,}4}}{0{,}6}\approx 1{,}05.$$

e) Konieczne do obliczenia współczynnika asymetrii momenty centralne otrzymujemy z wzorów (2.7.15), przyjmując w nich $\nu=1$ (ponieważ przy $\nu=1$ ujemny rozkład dwumianowy jest rozkładem geometrycznym):

$$\mu_2=D^2K=\frac{q}{p^2}, \quad \mu_3=\frac{q(2-p)}{p^3}. \tag{2.7.23}$$

Zatem

$$\gamma_1=\frac{\mu_3}{\mu_2\sqrt{\mu_2}}=\frac{q(2-p)p^2\sqrt{p^2}}{p^3 q\sqrt{q}}=\frac{2-p}{\sqrt{q}}. \tag{2.7.24}$$

Widać stąd, że przy dowolnej wartości parametru p rozkład geometryczny ma asymetrię prawą: $\gamma_1>0$. Dla $p=0{,}6$ otrzymujemy

$$\gamma_1=\frac{1{,}4}{\sqrt{0{,}4}}=\frac{1}{2}7\sqrt{0{,}4}\approx 2{,}21.$$

Zadanie 2.31. Prawdopodobieństwo awarii aparatury doświadczalnej w jednym doświadczeniu wynosi $p=0{,}02$. Doświadczenia można przeprowadzać dowolną liczbę razy. Obliczyć prawdopodobieństwo, że druga z kolei awaria: a) zdarzy się dokładnie w dziesiątym doświadczeniu, b) zdarzy się najpóźniej w dziesiątym doświadczeniu, c) nie zdarzy się w pierwszych dziesięciu doświadczeniach, przy założeniu, że doświadczenia są niezależne.

Rozwiązanie. Niech K oznacza liczbę doświadczeń, które trzeba będzie przeprowadzić. Zmienna losowa K ma ujemny rozkład dwumianowy (2.7.14) z parametrami $\nu=2$, $p=0{,}02$:

$$p_k\equiv P(K=k)=(k-1)p^2q^{k-2}=(k-1)0{,}02^2\cdot 0{,}98^{k-2}, \tag{1}$$

$$k=2,3,\ldots$$

a) Szukane prawdopodobieństwo jest wartością funkcji prawdopodobieństwa (1) dla $k=10$:

$$P(K=10)=9\cdot 0{,}02^2\cdot 0{,}98^8\approx 0{,}003.$$

b) Postawione zadanie jest pytaniem o prawdopodobieństwo przyjęcia przez zmienną losową K wartości z przedziału $\langle 2, 10\rangle$. Przy sumowaniu prawdopodobieństw p_k postępujemy w znany już sposób (zadanie 2.16) i otrzymujemy:

$$P(2\leqslant K\leqslant 10)=\sum_{k=2}^{10}(k-1)p^2q^{k-2}=p^2\sum_{k=1}^{9}kq^{k-1}=p^2\sum_{k=1}^{9}\frac{d}{dq}q^k=$$

$$=p^2\frac{d}{dq}\sum_{k=1}^{9}q^k=p^2\frac{d}{dq}\left(q\frac{1-q^9}{1-q}\right)=\ldots=$$

$$=1-q^{10}-10pq^9=1-(1+9p)q^9.$$

Dla $p=0{,}02$ mamy

$$P(2\leqslant K\leqslant 10)\approx 0{,}016.$$

c) Postawione pytanie jest pytaniem o prawdopodobieństwo przyjęcia przez zmienną losową wartości z przedziału $\langle 11,\infty)$. Zatem

$$P(11\leqslant K<\infty)=P(K\geqslant 11)=1-P(K\leqslant 10)\approx 1-0{,}016=0{,}984.$$

Zadanie 2.32. Z partii 250 sztuk towaru, zawierającej 18 sztuk wadliwych wylosowano bez zwrotu próbę 10-elementową. W procesie kontroli wyrywkowej partia zostanie odrzucona, gdy w próbce znajdzie się 2 lub więcej sztuk wadliwych. Obliczyć prawdopodobieństwo przyjęcia danej partii towaru.

Rozwiązanie. Liczba K sztuk wadliwych w próbce 10-elementowej jest zmienną losową o rozkładzie hipergeometrycznym z parametrami $N=250$, $M=18$, $n=10$ zadanym funkcją prawdopodobieństwa (2.7.11). Partia towaru zostanie przyjęta, gdy zmienna losowa przyjmuje wartość 0 albo 1. Zatem

$$P(\text{partia zostanie przyjęta})=P(K=0)+P(K=1)=$$

$$=P(0;250,18,10)+P(1;250,18,10)=\frac{\binom{18}{0}\binom{232}{10}}{\binom{250}{10}}+\frac{\binom{18}{1}\binom{232}{9}}{\binom{250}{10}}=$$

$$=\frac{232\cdot 231\cdot 230\ldots 224}{250\cdot 249\cdot 248\ldots 231}(223+10\cdot 18)\approx 0{,}84.$$

Zamiast bezpośrednich obliczeń, które tu przeprowadziliśmy, można posłużyć się rachunkiem logarytmicznym i tablicami logarytmów dziesiętnych silni (tablica 2):

$$\log P(K=0)=\ldots=\bar{1},6693,\quad \text{skąd}\quad P(K=0)=0{,}467,$$

$$\log P(K=1)=\ldots=\bar{1},5763,\quad \text{skąd}\quad P(K=1)=0{,}376,$$

$$P(K=0)+P(K=1)=0{,}843\approx 0{,}84.$$

Gdy $N\leqslant 20$, wtedy przy obliczaniu współczynników dwumianowych można korzystać z tablicy 1 tych współczynników.

Inny sposób obliczenia przybliżonej wartości prawdopodobieństwa przyjęcia danej partii towaru podano w rozwiązaniu zadania 2.36.

Zadanie 2.33. Wyznaczyć współczynnik asymetrii rozkładu hipergeometrycznego z parametrami (N, M, n).

Rozwiązanie. Potrzebne do obliczenia momenty centralne μ_2 i μ_3 tego rozkładu dane są wzorami (2.7.12). Zatem zgodnie z definicją współczynnika γ_1, otrzymujemy:

$$\gamma_1 = \frac{\mu_3}{\mu_2\sqrt{\mu_2}} = \frac{npq(1-2p)(N-n)(N-2n)(N-1)\sqrt{N-1}}{(N-1)(N-2)npq(N-n)\sqrt{npq}\sqrt{N-n}} =$$

$$= \frac{1-2p}{\sqrt{npq}} \frac{N-2n}{N-2} \sqrt{\frac{N-1}{N-n}},$$

gdzie $p = \frac{M}{N}$, $q = 1-p$.

Zadanie 2.34. Zmienna losowa K ma rozkład Poissona (2.7.16) z parametrem $\lambda = 2$. Znając wartość prawdopodobieństwa $P(K=3) = P(3; 2) = 0{,}18$, obliczyć prawdopodobieństwo $P(K=4) = P(4; 2)$.

Rozwiązanie. Rozwiążemy zadanie ogólniejsze. Znajdziemy związek rekurencyjny dla prawdopodobieństw p_k w rozkładzie Poissona. W tym celu rozpatrujemy iloraz

$$\frac{p_{k+1}}{p_k} = \frac{P(k+1\,;\,\lambda)}{P(k\,;\,\lambda)} = \frac{e^{-\lambda}\lambda^{k+1}k!}{(k+1)!\,e^{-\lambda}\lambda^k} = \frac{\lambda}{k+1}.$$

Stąd mamy

$$P(k+1\,;\,\lambda) = \frac{\lambda}{k+1} P(k\,;\,\lambda).$$

Dla danych w zadaniu otrzymujemy:

$$P(4\,;\,2) = \tfrac{2}{4}P(3\,;\,2) = \tfrac{1}{2}\cdot 0{,}18 = 0{,}09.$$

Zadanie 2.35. Prawdopodobieństwo, że produkt poddawany próbie nie wytrzyma tej próby wynosi $p = 0{,}01$. Obliczyć prawdopodobieństwo, że wśród 200 takich produktów co najwyżej 2 nie wytrzyma próby.

Rozwiązanie. Niech K oznacza liczbę tych produktów wśród 200, które nie wytrzymały próby. Zmienna losowa K podlega rozkładowi dwumianowemu (2.7.7) z parametrami $n = 200$, $p = 0{,}01$. Ponieważ jednak n jest duże,

$$n = 200 \geqslant 50 \quad \text{oraz} \quad p = 0{,}01 \leqslant 0{,}1 \quad \text{i} \quad \lambda = np = 2 \leqslant 10,$$

więc stosujemy przybliżenie Poissona (2.7.19) rozkładu dwumianowego. Zdarzenie $\boldsymbol{B}$, którego prawdopodobieństwo należy obliczyć, zachodzi wtedy gdy zmienna losowa K przyjmuje wartości 0 albo 1, albo 2. Zatem

$$P(B) \approx P(0\,;\,2) + P(1\,;\,2) + P(2\,;\,2) = e^{-2}\frac{2^0}{0!} + e^{-2}\frac{2^1}{1!} + e^{-2}\frac{2^2}{2!} =$$
$$= e^{-2}(1+2+2) \approx 0{,}68.$$

Zadanie 2.36. Stosując przybliżenie Poissona (2.7.20) rozkładu hipergeometrycznego, rozwiązać zadanie 2.32.

Rozwiązanie. Ponieważ

$$p = \frac{18}{250}, \quad \lambda = np = 10\cdot\frac{18}{250} = 0{,}72,$$

więc korzystając z tablicy wartości funkcji $P(k;\lambda)$ dla $k=0$ i $k=1$ oraz $\lambda=0{,}7$, otrzymujemy:

$$P(\text{przyjęcia partii towaru}) \approx P(0;\ 0{,}7)+P(1;\ 0{,}7)=$$

$$=e^{-0{,}7}\frac{0{,}7^0}{0!}+e^{-0{,}7}\frac{0{,}7^1}{1!}\approx 0{,}49+0{,}35=0{,}84.$$

2.8. NIEKTÓRE ROZKŁADY TYPU CIĄGŁEGO I ICH PARAMETRY

Najczęściej gęstość prawdopodobieństwa jest funkcją zależną od pewnych stałych zwanych *parametrami*. Wśród nich szczególne znaczenie mają: *parametr skali* i *parametr przesunięcia*.

Jeśli gęstość prawdopodobieństwa zmiennej losowej X jest postaci

$$\frac{1}{\alpha}f\left(\frac{x}{\alpha}\right), \quad \alpha>0, \tag{2.8.1}$$

to przyjmując α jednostek za nową jednostkę skali — co odpowiada wprowadzeniu nowej zmiennej losowej $U=\frac{X}{\alpha}$ — gęstość zmiennej losowej U będzie równa $f(u)$ (p. 2.5), a więc nie będzie zależała od parametru α. Ponieważ uzyskaliśmy to przez zmianę skali, więc dlatego parametr α gęstości (2.8.1) nazywamy *parametrem skali.*

Jeśli gęstość prawdopodobieństwa zmiennej losowej X jest postaci

$$f(x-\lambda), \tag{2.8.2}$$

to przesuwając punkt zerowy skali X do punktu λ — co odpowiada wprowadzeniu nowej zmiennej losowej $X-\lambda=T$ — gęstość prawdopodobieństwa tej nowej zmiennej losowej T będzie równa $f(t)$ (p. 2.5) i nie będzie zależała od parametru λ. Dlatego parametr λ gęstości (2.8.2) nazywamy *parametrem przesunięcia* (*lokacyjnym*).

Zajmiemy się teraz omówieniem tych rozkładów prawdopodobieństwa zmiennej losowej ciągłej, które mają duże znaczenie praktyczne, bądź teoretyczne.

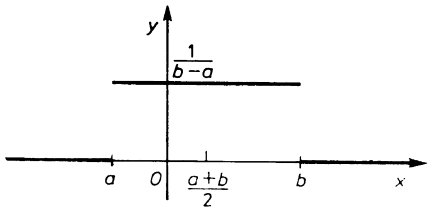
2.8.1. Rozkład równomierny typu ciągłego. Mówimy, że zmienna losowa X ma *rozkład równomierny* (*jednostajny, prostokątny*) skoncentrowany na przedziale $\langle a,b\rangle$ jeżeli jej gęstość prawdopodobieństwa jest określona wzorem

$$f(x)=\begin{cases}\dfrac{1}{b-a} & \text{dla} \quad a\leqslant x\leqslant b,\\ 0 & \text{dla pozostałych } x.\end{cases} \tag{2.8.3}$$

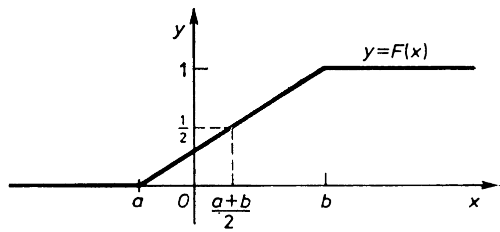
Dystrybuantą tego rozkładu jest funkcja

$$F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant a, \\ \dfrac{x-a}{b-a} & \text{dla} \quad a<x\leqslant b, \\ 1 & \text{dla} \quad x>b. \end{cases} \tag{2.8.4}$$

Wykresy gęstości i dystrybuanty przedstawiają rys. 2.24, 2.25.



Rys. 2.24. Gęstość rozkładu równomiernego skoncentrowanego na $\langle a, b\rangle$



Rys. 2.25. Dystrybuanta rozkładu równomiernego skoncentrowanego na $\langle a, b\rangle$

Wartość przeciętna i mediana zmiennej losowej X o rozkładzie (2.8.3) są równe

$$EX=x_{0,5}=\frac{a+b}{2}.$$

Moment centralny parzystego rzędu $2r$ wyraża się wzorem

$$\bigwedge_{r\in N} \mu_{2r}=\frac{(b-a)^{2r}}{(2r+1)2^{2r}}, \tag{2.8.5}$$

w szczególności gdy $r=1$, otrzymujemy wariancję

$$D^2X=\tfrac{1}{12}(b-a)^2 \tag{2.8.6}$$

równą jednej dwunastej kwadratu długości przedziału $\langle a, b\rangle$. Natomiast wszystkie momenty centralne nieparzystego rzędu wobec symetrii gęstości rozkładu względem wartości przeciętnej są równe zeru.

Liczne zastosowania rozkładu równomiernego wynikają z następującej własności:

Prawdopodobieństwo, że zmienna losowa X o rozkładzie (2.8.3) *przyjmuje wartości z przedziału $\langle x_0, x_0+h\rangle$, gdzie $a\leqslant x_0<x_0+h\leqslant b$, jest wprost proporcjonalne do długości h tego przedziału i nie zależy od x_0.*

2.8.2. Rozkład wykładniczy. Mówimy, że zmienna losowa X ma *rozkład wykładniczy* o parametrze $\lambda>0$, jeżeli jej gęstość f jest postaci

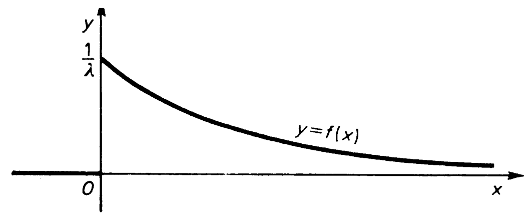
$$f(x)=\begin{cases} \dfrac{1}{\lambda}\exp\left(-\dfrac{x}{\lambda}\right) & \text{dla} \quad x\geqslant 0, \\ 0 & \text{dla pozostałych } x. \end{cases} \tag{2.8.7}$$

Łatwo obliczyć, że $EX=\lambda$, $D^2X=\lambda^2$.

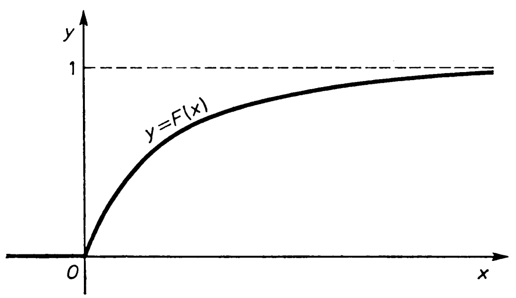
Całkując gęstość (2.8.7), otrzymujemy dystrybuantę tego rozkładu:

$$F(x)=\begin{cases} 1-\exp\left(-\dfrac{x}{\lambda}\right) & \text{dla} \quad x>0, \\ 0 & \text{poza tym}. \end{cases} \tag{2.8.8}$$

Wykresy gęstości i dystrybuanty przedstawiają rysunki 2.26 i 2.27.



Rys. 2.26. Gęstość rozkładu wykładniczego z parametrem λ



Rys. 2.27. Dystrybuanta rozkładu wykładniczego

Zmienna losowa o rozkładzie wykładniczym opisuje wiele często spotykanych zjawisk i tak np.:

a) przyjmuje się często, że czas bezawaryjnej pracy T badanego elementu (tzw. *czas życia elementu*) jest zmienną losową o rozkładzie wykładniczym, wówczas $P(T\geqslant t)=$ $=\exp\left(-\dfrac{t}{\lambda}\right)$ nazywamy *niezawodnością elementu*, zaś $\dfrac{1}{\lambda}$ — *intensywnością awarii*;

b) jeżeli liczba zgłoszeń w centrali telefonicznej w przedziale czasu $(\theta, \theta+\tau)$ o dowolnym początku θ i ustalonym $\tau>0$, jest zmienną losową K o rozkładzie Poissona

$$P(K=k)=\frac{1}{k!}\left(\frac{\tau}{\lambda}\right)^k \exp\left(-\frac{\tau}{\lambda}\right), \qquad k\in N_0,$$

o wartości przeciętnej τ/λ wprost proporcjonalnej do długości τ przedziału (λ — stała dodatnia), oraz liczby zgłoszeń zachodzące w rozłącznych przedziałach czasu są niezależne, to czas T pomiędzy dwoma kolejnymi zgłoszeniami jest zmienną losową o rozkładzie wykładniczym (2.8.7).

Liczne zastosowania rozkładu wykładniczego (2.8.7) wiążą się z jego następującymi własnościami:

1) $\bigwedge\limits_{a,b>0} P(X\geqslant a+b\,|\,X\geqslant a)=P(X\geqslant b).$

Istotnie

$$P(X \geqslant a+b \mid X \geqslant a) = \frac{P(X \geqslant a+b \wedge X \geqslant a)}{P(X \geqslant a)} =$$

$$= \frac{P(X \geqslant a+b)}{P(X \geqslant a)} = \frac{\exp\left(-\frac{a+b}{\lambda}\right)}{\exp\left(-\frac{a}{\lambda}\right)} = \exp\left(-\frac{b}{\lambda}\right) = P(X \geqslant b).$$

Rozkład wykładniczy jest jedynym rozkładem ciągłym, który ma wyżej podaną własność zwaną *brakiem pamięci*. Własność tę można interpretować następująco: jeżeli zmienna losowa X jest czasem bezawaryjnej pracy pewnego elementu o rozkładzie wykładniczym (2.8.7), to niezależnie od dotychczasowego czasu pracy elementu, dalszy czas pracy nie zależy od „przeszłości" i ma taki sam rozkład, co całkowity czas pracy elementu.

2) Suma n niezależnych zmiennych losowych o rozkładzie wykładniczym (2.8.7) ma *rozkład Erlanga* o gęstości

$$f(x) = \begin{cases} \frac{1}{\lambda^n (n-1)!} x^{n-1} \exp\left(-\frac{x}{\lambda}\right) & \text{dla } x \geqslant 0, \\ 0 & \text{poza tym} \end{cases} \qquad \text{(wzór (2.8.15)).}$$

(Rozkład Erlanga jest szczególnym przypadkiem rozkładu gamma (2.8.12) o parametrze $p = n \in N$.)

Jeśli wrócimy do przyjętej w b) interpretacji rozkładu wykładniczego i oznaczymy przez $T_1, \ldots, T_k$ momenty kolejnych pojedynczych zgłoszeń w centrali telefonicznej, to własność 2) oznacza, że dla każdego $n = 1, \ldots, k$ rozkład czasu X, jaki upłynął od początku obserwacji do momentu zgłoszenia się n-tego abonenta, jest rozkładem Erlanga o gęstości (2.8.15).

2.8.3. Rozkład gamma. Rozkład gamma jest związany z funkcją specjalną gamma, którą oznacza się przez Γ.

Funkcją gamma argumentu p zespolonego o części rzeczywistej dodatniej nazywamy całkę postaci

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx, \qquad \operatorname{Re} p > 0, \tag{2.8.9}$$

w szczególności gdy p jest liczbą rzeczywistą dodatnią, całkując (2.8.9) przez części, otrzymujemy

$$\Gamma(p+1) = p\Gamma(p), \qquad p > 0. \tag{2.8.10}$$

Jeśli zaś p jest liczbą naturalną, $p = n$, $n \in N$, to

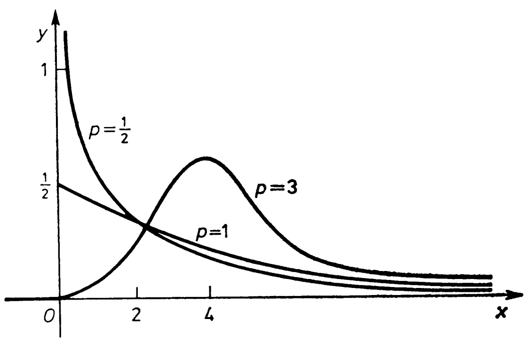
$$\Gamma(n+1) = n!. \tag{2.8.11}$$

Mówimy, że zmienna losowa X ma *rozkład gamma* o parametrach p, $\lambda>0$, jeśli jej gęstość prawdopodobieństwa jest określona wzorem:

$$f(x)=\begin{cases}\dfrac{1}{\lambda^p\Gamma(p)}x^{p-1}\exp\left(-\dfrac{x}{\lambda}\right) & \text{dla } x>0,\\ 0 & \text{dla pozostałych } x,\end{cases} \tag{2.8.12}$$

przy czym λ jest parametrem skali (2.8.1) tego rozkładu, p zaś nazywamy *parametrem kształtu*.

Wykresy gęstości rozkładu (2.8.12) dla $\lambda=2$ i różnych wartości parametru p ($p=1, 3, \frac{1}{2}$) przedstawia rysunek 2.28.



Rys. 2.28. Gęstości **rozkładów gamma o parametrach $\lambda=2$ i $p=\frac{1}{2}, 1, 3$**

Korzystając z określenia funkcji gamma łatwo wykazać, że moment zwykły r-tego rzędu zmiennej losowej X o gęstości (2.8.12) jest postaci

$$\bigwedge_{r\in N}\alpha_r=\frac{\Gamma(p+r)}{\Gamma(p)}\lambda^r. \tag{2.8.13}$$

Stąd

$$EX=\alpha_1=\lambda p, \qquad D^2X=\alpha_2-\alpha_1^2=p\lambda^2.$$

W przypadku $\lambda=1$, mamy

$$EX=D^2X=p.$$

Identyczną własność dla zmiennej losowej typu skokowego ma rozkład Poissona (2.7.16).

Dla $p>1$ rozkład (2.8.12) jest jednomodalny, przy czym

$$m_0=\lambda(p-1). \tag{2.8.14}$$

Rozważamy następujące szczególne przypadki rozkładu gamma:

1) $p=1$, mamy wówczas gęstość

$$f(x)=\begin{cases}\dfrac{1}{\lambda}\exp\left(-\dfrac{x}{\lambda}\right) & \text{dla} \quad x>0,\\ 0 & \text{dla} \quad x\leqslant 0\end{cases}$$

rozkładu wykładniczego (2.8.7) o parametrze $\lambda = EX$.

2) $p=n$, gdzie $n \in N$; otrzymujemy wówczas gęstość

$$f(x)=\begin{cases} \dfrac{1}{(n-1)!\,\lambda^n} x^{n-1} \exp\left(-\dfrac{x}{\lambda}\right) & \text{dla} \quad x>0, \\ 0 & \text{dla} \quad x \leqslant 0 \end{cases} \tag{2.8.15}$$

rozkładu Erlanga z parametrami n i λ (skali).

3) $p=\frac{1}{2}n$, $\lambda=2$. Rozkład zmiennej losowej o gęstości

$$f(x)=\begin{cases} \dfrac{1}{2^{n/2}\Gamma(\frac{1}{2}n)} x^{n/2-1} \exp(-\frac{1}{2}x) & \text{dla } x>0, \\ 0 & \text{dla } x \leqslant 0 \end{cases} \tag{2.8.16}$$

nazywamy *rozkładem* χ^2 *z n stopniami swobody.*

Powyższe rozkłady występują często w takich zastosowaniach rachunku prawdopodobieństwa jak teoria niezawodności i statystyka matematyczna.

2.8.4. Rozkład beta. Rozkład beta jest związany z funkcją specjalną $B(p, q)$. *Funkcję beta* $B(p, q)$ argumentów rzeczywistych dodatnich określamy jako całkę:

$$B(p, q)=\int_0^1 x^{p-1}(1-x)^{q-1}\,dx \quad \text{dla } p, q>0. \tag{2.8.17}$$

Między funkcjami beta i gamma zachodzi następująca zależność:

$$B(p, q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \tag{2.8.18}$$

Mówimy, że zmienna losowa X ma *rozkład beta* I-*go rodzaju* o parametrach p, $q>0$ jeśli jej gęstość jest postaci:

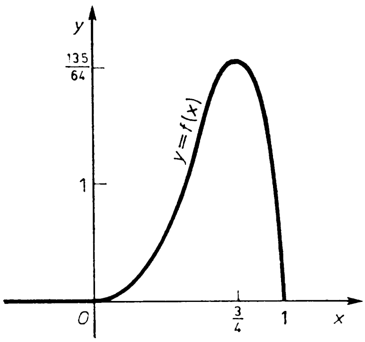
$$f(x)=\begin{cases} \dfrac{1}{B(p, q)} x^{p-1}(1-x)^{q-1} & \text{dla } 0<x<1, \\ 0 & \text{dla pozostałych } x. \end{cases} \tag{2.8.19}$$

Wykres gęstości rozkładu beta dla $p=4$, $q=2$ przedstawia rysunek 2.29. Korzystając ze wzorów (2.8.17) - (2.8.18), można wyprowadzić wzór na momenty zwykłe r-tego rzędu zmiennej losowej X o rozkładzie (2.8.19):

$$\bigwedge_{r \in N} \alpha_r=\frac{B(p+r, q)}{B(p, q)}=\frac{\Gamma(p+r)\Gamma(p+q)}{\Gamma(p+q+r)\Gamma(p)}. \tag{2.8.20}$$

Stąd wartość przeciętna i wariancja tej zmiennej losowej:

$$EX=\frac{p}{p+q}, \qquad D^2X=\frac{pq}{(p+q)^2(p+q+1)}. \tag{2.8.21}$$



Rys. 2.29. Gęstość rozkładu beta o parametrach $p=4$, $q=2$, do zadania 2.177

Jeżeli X jest ograniczoną zmienną losową o momentach α_r określonych wzorem (2.8.20), to ma ona rozkład beta (2.8.19), prawdziwe jest bowiem twierdzenie:

Jeśli zmienna losowa X jest ograniczona(1)*, to jej rozkład jest jednoznacznie wyznaczony przez jej momenty α_r, $r \in N$.*

Jeśli założenie ograniczoności zmiennej losowej nie jest spełnione, to znajomość wszystkich momentów α_r, $r \in N$ nie wyznacza jednoznacznie rozkładu zmiennej losowej.

2.8.5. Rozkład normalny. Najważniejszym i najczęściej występującym w zastosowaniach rozkładem typu ciągłego jest rozkład normalny (gaussowski).

Mówimy, że zmienna losowa X ma *rozkład normalny* o parametrach μ, σ przy czym $\mu \in \mathbf{R}$, $\sigma \in \mathbf{R}_+$, jeśli jej gęstość prawdopodobieństwa jest określona wzorem

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad \text{dla } -\infty<x<\infty, \tag{2.8.22}$$

gdzie μ, σ są odpowiednio parametrami przesunięcia (2.8.2) i skali (2.8.1) tego rozkładu. Rozkład normalny o parametrach μ, σ oznaczamy symbolem $N(\mu, \sigma)$.

ZADANIE 2.37. Zbadać przebieg zmienności i naszkicować wykres gęstości rozkładu $N(\mu, \sigma)$.

Rozwiązanie. Obliczamy pierwszą i drugą pochodną funkcji f określonej wzorem (2.8.22):

$$f'(x)=-\frac{x-\mu}{\sigma^3\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],$$

$$f''(x)=\frac{\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]}{\sigma^3\sqrt{2\pi}}\left[\frac{(x-\mu)^2}{\sigma^2}-1\right].$$

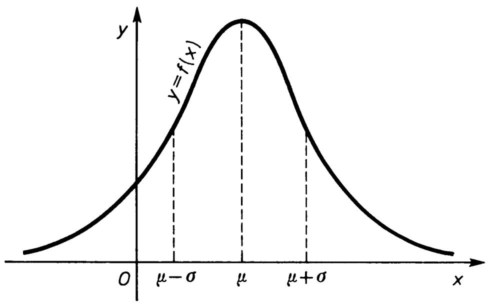
---

(1) Zmienna losowa X typu ciągłego jest ograniczoną, jeśli jej rozkład prawdopodobieństwa jest skoncentrowany na przedziale właściwym.

Badając miejsca zerowe oraz znak obydwu pochodnych, otrzymujemy, że gęstość f ma maksimum w punkcie $x=\mu$ oraz punkty przegięcia $x=\mu\mp\sigma$.

Wykresem gęstości rozkładu normalnego jest tzw. *krzywa Gaussa* (rys. 2.30).

Rys. 2.30. Gęstość rozkładu $N(\mu, \sigma)$, tzw. krzywa Gaussa



Można pokazać, że dla zmiennej losowej o rozkładzie normalnym istnieją skończone momenty dowolnego rzędu, więc wobec symetrii gęstości względem $x=\mu$ jest oczywiste, że wartość przeciętna i mediana są równe μ, ponieważ ponadto dla $x=\mu$ gęstość osiąga jedyne maksimum lokalne, więc również moda jest równa μ, skąd

$$\alpha_1=EX=x_{0,5}=m_0=\mu$$

oraz wszystkie momenty centralne nieparzystego rzędu są równe zeru:

$$\bigwedge_{r\in N}\mu_{2r-1}=0,$$

natomiast momenty centralne rzędu parzystego określa wzór:

$$\bigwedge_{r\in N}\mu_{2r}=1\cdot 3\cdot 5\cdot\ldots\cdot(2r-1)\sigma^{2r}=(2r-1)!!\sigma^{2r}. \tag{2.8.23}$$

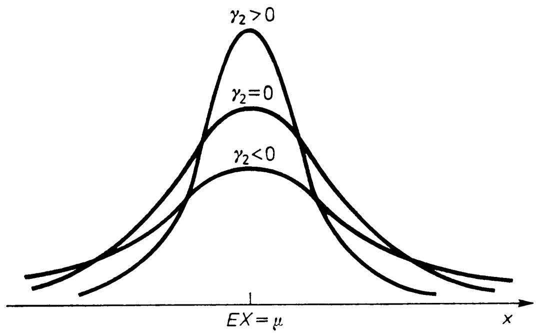
W szczególnym przypadku wariancja

$$D^2X=\mu_2=\sigma^2.$$

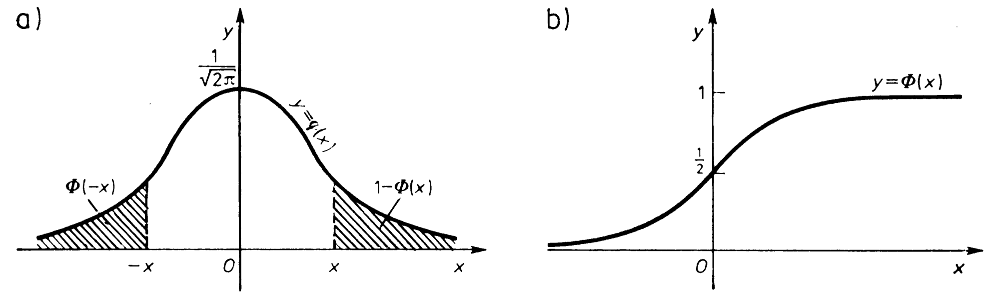
Współczynnik skupienia dla rozkładu normalnego o dowolnych parametrach μ i σ jest równy 3. Istotnie $K=\frac{\mu_4}{\sigma^4}$, a ponieważ dla rozkładu normalnego $\mu_4=3\sigma^4$, więc $K=3$.

Zwróćmy ponadto uwagę, że eksces γ_2 (*współczynnik spłaszczenia*), tj. $\gamma_2=K-3$, dla rozkładu normalnego jest równy zeru. Eksces jest więc współczynnikiem, który porównuje skupienie dowolnego rozkładu wokół jego wartości przeciętnej, ze skupieniem rozkładu $N(\mu, \sigma)$. W większości spotykanych rozkładów obserwujemy, że jeżeli $\gamma_2>0$, to rozkład jest bardziej skupiony (stromy) niż odpowiedni rozkład normalny, a jeśli $\gamma_2<0$ to jest przeciwnie. W ogólności nie jest to jednak prawdziwe (zad. 2.156).

Porównanie gęstości rozkładów o różnych współczynnikach spłaszczenia (ekscesie) znajdziemy na rysunku 2.31.



Rys. 2.31. Gęstości o różnych współczynnikach spłaszczenia (ekscesie). (Zad. 2.156)



Rys. 2.32. a) $\Phi(-x)=1-\Phi(x)$ na rysunku gęstości φ rozkładu $N(0, 1)$, b) Dystrybuanta Φ rozkładu $N(0, 1)$

Jeśli zmienna losowa X ma rozkład $N(\mu, \sigma)$, to standaryzowana zmienna losowa $\dfrac{X-\mu}{\sigma}$ ma rozkład $N(0, 1)$ zwany *standaryzowanym rozkładem normalnym*. Gęstością tego rozkładu jest funkcja (rys. 2.32a)

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\exp(-\tfrac{1}{2}x^2) \quad \text{dla } x\in\boldsymbol{R}, \tag{2.8.24}$$

zaś dystrybuantą Φ (rys. 2.32b)

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-\tfrac{1}{2}t^2)\,dt. \tag{2.8.25}$$

Z symetrii wykresu gęstości względem osi Oy wynika następująca zależność:

$$\Phi(-x)=1-\Phi(x). \tag{2.8.26}$$

Wartości dystrybuanty $\Phi(x)$ stablicowano dla $x>0$ (tabl. 5).

2.8.6. Rozkład Laplace'a. Powiemy, że zmienna losowa X ma *rozkład Laplace'a* (dwustronny wykładniczy) (rys. 2.33), jeżeli jej gęstość jest określona wzorem

$$f(x)=\frac{1}{2\lambda}\exp\left(-\frac{|x-\mu|}{\lambda}\right) \quad \text{dla } x\in\boldsymbol{R},\ \lambda>0,\ \mu\in\boldsymbol{R}. \tag{2.8.27}$$

Parametry λ, μ są odpowiednio parametrami skali (2.8.1) i przesunięcia (2.8.2) tego rozkładu.

Podobnie jak w przypadku rozkładu normalnego mamy

$$EX=x_{0,5}=m_0=\mu$$

oraz wszystkie momenty centralne rzędu nieparzystego są równe zeru:

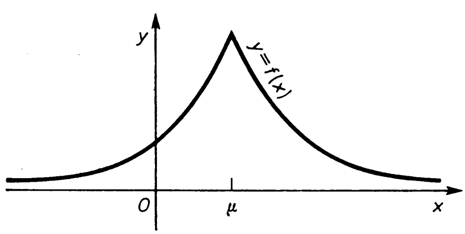
$$\mu_{2r-1}=0 \quad \text{dla } r\in N.$$

Natomiast momenty centralne parzystego rzędu są określone wzorem:

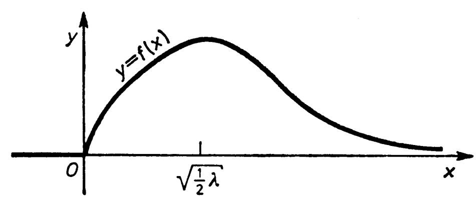
$$\mu_{2r}=\lambda^{2r}\Gamma(2r+1)=\lambda^{2r}(2r)! \quad \text{dla} \quad r\in N, \tag{2.8.28}$$

stąd np. wariancja $\mu_2=D^2X=2\lambda^2$.

Gęstość rozkładu Laplace'a nie jest różniczkowalna dla $x=\mu$.



Rys. 2.33. Gęstość rozkładu Laplace'a



Rys. 2.34. Gęstość rozkładu Rayleigha

2.8.7. Rozkład Rayleigha. Mówimy, że zmienna losowa X ma *rozkład Rayleigha* z parametrem $\lambda>0$, jeśli jej gęstość jest postaci (rys. 2.34)

$$f(x)=\begin{cases} \dfrac{2}{\lambda}x\exp\left(-\dfrac{x^2}{\lambda}\right) & \text{dla } x>0, \\ 0 & \text{dla } x\leqslant 0. \end{cases} \tag{2.8.29}$$

Łatwo zauważyć, że $\sqrt{\lambda}$ jest parametrem skali rozkładu (2.8.29). Momenty zwykłe rzędu r określa wzór:

$$\alpha_r=\lambda^{r/2}\Gamma(\tfrac{1}{2}r+1) \quad \text{dla } r\in N. \tag{2.8.30}$$

Stąd momenty rzędów parzystych:

$$\alpha_{2n}=n!\,\lambda^n \quad \text{dla } n\in N,$$

zaś dla $r=2n-1$, $n\in N$, otrzymujemy:

$$\alpha_{2n-1}=\lambda^{(2n-1)/2}\Gamma\left(\frac{2n-1}{2}+1\right)=\lambda^{(2n-1)/2}\Gamma(n+\tfrac{1}{2})=$$

$$=\frac{1\cdot 3\cdot\ldots\cdot(2n-1)}{2^n}\sqrt{\pi}\,\lambda^{(2n-1)/2}=\sqrt{\pi}\,\frac{(2n-1)!!}{2^n}\lambda^{(2n-1)/2} \quad \text{dla } n\in N,$$

w szczególnym przypadku

$$\alpha_1 = EX = \tfrac{1}{2}\sqrt{\pi\lambda}, \qquad \alpha_2 = EX^2 = \lambda,$$

a więc wariancja zmiennej losowej X

$$\mu_2 = D^2X = \frac{4-\pi}{4}\lambda.$$

2.8.8. Rozkład Maxwella. Mówimy, że zmienna losowa X ma *rozkład Maxwella* z parametrem $\lambda > 0$ (rys. 2.35), jeśli jej gęstość jest określona wzorem

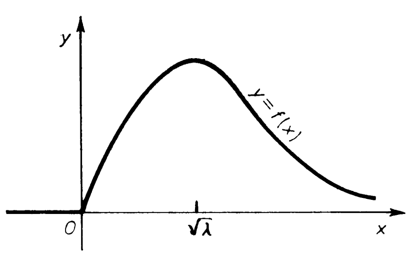
$$f(x) = \begin{cases} \dfrac{4}{\sqrt{\pi}\,\lambda^{3/2}} x^2 \exp\left(-\dfrac{x^2}{\lambda}\right) & \text{dla } x > 0, \\ 0 & \text{dla } x \leqslant 0. \end{cases} \tag{2.8.31}$$

Jest to rozkład jednomodalny: $m_0 = \sqrt{\lambda}$. Tak jak w (2.8.29) $\sqrt{\lambda}$ jest także parametrem skali (2.8.1) rozkładu Maxwella. Momenty zwykłe rzędu r są określone wzorem:

$$\alpha_r = \frac{2\lambda^{r/2}}{\sqrt{\pi}} \Gamma\left(\frac{r+3}{2}\right) \quad \text{dla } r \in N, \tag{2.8.32}$$

przy czym jeśli $r = 2n-1 \wedge n \in N$, to

$$\alpha_{2n-1} = \frac{2 \cdot n!}{\sqrt{\pi}} \lambda^{(2n-1)/2} \quad \text{dla } n \in N,$$



Rys. 2.35. Gęstość rozkładu Maxwella

jeśli zaś $r = 2n \wedge n \in N$, to

$$\alpha_{2n} = \frac{1 \cdot 3 \cdot \ldots \cdot (2n+1)}{2^n} \lambda^n = \frac{(2n+1)!!}{2^n} \lambda^n.$$

Stąd wartość przeciętna $\alpha_1 = EX = 2\sqrt{\dfrac{\lambda}{\pi}}$ oraz drugi moment zwykły $\alpha_2 = EX^2 = \frac{3}{2}\lambda$, wariancja zaś

$$\mu_2 = \operatorname{Var} X = \frac{3\pi - 8}{2\pi}\lambda.$$

Dystrybuantę rozkładu (2.8.31):

$$F(x)=\begin{cases} 0 & \text{dla } x\leqslant 0, \\ \dfrac{4\lambda^{-3/2}}{\sqrt{\pi}}\displaystyle\int_0^x t^2\exp\left(-\frac{t^2}{\lambda}\right)dt & \text{dla } x>0 \end{cases}$$

można inaczej wyrazić za pomocą tzw. *niepełnej funkcji gamma* określonej wzorem $\gamma(a, x)=\int_0^x t^{a-1}e^{-t}dt$:

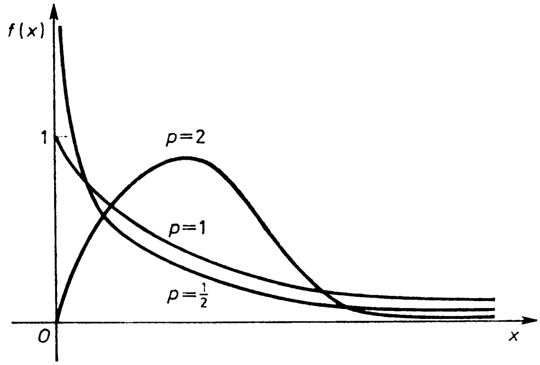
$$F(x)=\begin{cases} \dfrac{2}{\sqrt{\pi}}\gamma\left(\dfrac{3}{2}, \dfrac{x^2}{\lambda}\right) & \text{dla } x>0, \\ 0 & \text{dla } x\leqslant 0. \end{cases}$$

2.8.9. Rozkład Weibulla. Powiemy, że zmienna losowa X ma *rozkład Weibulla* o parametrach p, λ, jeśli jej gęstość jest określona wzorem

$$f(x)=\begin{cases} \dfrac{p}{\lambda}x^{p-1}\exp\left(-\dfrac{x^p}{\lambda}\right) & \text{dla } x>0, \\ 0 & \text{dla } x\leqslant 0, \end{cases} \tag{2.8.33}$$

gdzie p, $\lambda>0$.

Wykresy gęstości dla $\lambda=1$ i $p=\frac{1}{2}$, $p=1$, $p=2$ przedstawiono na rys. 2.36.



Rys. 2.36. Gęstości rozkładów Weibulla o parametrach $\lambda=1$, $p=\frac{1}{2}$, 1, 2

Momenty zwykłe rzędu r zmiennej losowej X o rozkładzie Weibulla są określone wzorem:

$$\bigwedge_{r\in N}\alpha_r=\lambda^{r/p}\Gamma\left(\frac{r}{p}+1\right). \tag{2.8.34}$$

W szczególnym przypadku

dla $r=1$ $\quad \alpha_1=EX=\lambda^{1/p}\Gamma\left(\dfrac{1}{p}+1\right)$,

dla $r=2$ $\quad \alpha_2=E(X^2)=\lambda^{2/p}\Gamma\left(\dfrac{2}{p}+1\right)$, skąd wariancja zmiennej losowej X jest:

$$D^2X = \lambda^{2/p}\left\{\Gamma\left(\frac{2}{p}+1\right)-\left[\Gamma\left(\frac{1}{p}+1\right)\right]^2\right\}.$$

Ostatnie 3 omówione rozkłady: Rayleigha, Maxwella i Weibulla można traktować jako szczególne przypadki tzw. *rozkładu uogólnionego gamma.*

2.8.10. Rozkład uogólniony gamma. Mówimy, że zmienna losowa X ma *rozkład uogólniony* (*trójparametrowy*) *gamma*, jeśli jej gęstość jest określona wzorem:

$$f(x) = \begin{cases} \dfrac{\alpha}{\Gamma\left(\dfrac{p}{\alpha}\right)\lambda^{p/\alpha}}\, x^{p-1}\exp\left(-\dfrac{x^\alpha}{\lambda}\right) & \text{dla } x>0, \\ 0 & \text{dla } x \leqslant 0, \end{cases} \tag{2.8.35}$$

gdzie λ, p, $\alpha > 0$.

Jeżeli

1) $p=1$, $\alpha=1$, to otrzymamy rozkład wykładniczy (2.8.7),
2) $\alpha=1$, to otrzymamy rozkład gamma dwuparametrowy (2.8.12),
3) $p=2$, $\alpha=2$, to otrzymamy rozkład Rayleigha (2.8.29),
4) $p=\alpha$, to otrzymamy rozkład Weibulla (2.8.33),
5) $\alpha=2$ i $p=3$, to otrzymamy rozkład Maxwella (2.8.31).

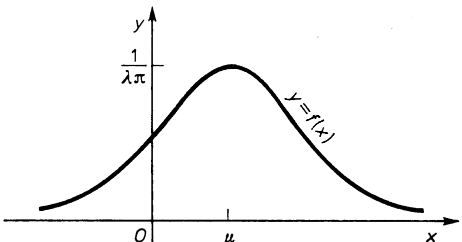
Momenty zwykłe rzędu r zmiennej losowej o rozkładzie (2.8.35) wyrażają się wzorem:

$$\alpha_r = \lambda^{r/\alpha}\frac{\Gamma\left(\dfrac{r+p}{\alpha}\right)}{\Gamma\left(\dfrac{p}{\alpha}\right)}, \quad r\in N. \tag{2.8.36}$$

2.8.11. Rozkład Cauchy'ego. Mówimy, że zmienna losowa X ma *rozkład Cauchy'ego* o parametrach μ, λ (rys. 2.37), jeśli jej gęstość jest określona wzorem

$$f(x) = \frac{1}{\pi}\frac{\lambda}{\lambda^2+(x-\mu)^2}, \quad \text{gdzie } \lambda>0. \tag{2.8.37}$$

λ jest parametrem skali (2.8.1), μ przesunięcia (2.8.2) rozkładu (2.8.37).



Rys. 2.37. Gęstość rozkładu Cauchy'ego

Zmienna losowa o rozkładzie Cauchy'ego nie ma momentów α_r dla żadnego $r \in N$. W szczególności nie istnieje więc wartość przeciętna tej zmiennej losowej (zad. 2.43). Z symetrii f względem $x=\mu$ wynika, że w rozkładzie Cauchy'ego mediana jest równa μ. Ponieważ w tym punkcie gęstość osiąga jedyne maksimum lokalne, więc wartością modalną jest również μ, tak więc EX nie istnieje, natomiast

$$x_{0,5}=m_0=\mu.$$

2.8.12. Zadania rozwiązane.

ZADANIE 2.38. Czas (w minutach) między kolejnymi zgłoszeniami abonentów w pewnej centrali telefonicznej jest zmienną losową o rozkładzie wykładniczym (2.8.7) z parametrem $\lambda=2$. Obliczyć średni czas między kolejnymi zgłoszeniami oraz prawdopodobieństwo, że przed upływem 3 minut nastąpi zgłoszenie.

Rozwiązanie.

$$EX=\tfrac{1}{2}\int_0^{\infty} x\exp(-\tfrac{1}{2}x)\,dx.$$

Oznaczmy przez X czas między kolejnymi zgłoszeniami.

Stosując wzór (2.8.36) dla $p=1$, $\lambda=2$, $r=1$, $\alpha=1$, otrzymamy

$$EX=\alpha_1=2(\min).$$

Obliczmy

$$P(X<3)=F(3)=\tfrac{1}{2}\int_0^3 \exp(-\tfrac{1}{2}x)\,dx=[-\exp(-\tfrac{1}{2}x)]_0^3=1-\exp(-\tfrac{3}{2}).$$

ZADANIE 2.39. Odstęp między kolejnymi podziałkami skali stopera wynosi 0,1 s. Czas na tym stoperze odczytuje się z dokładnością do jednej podziałki, zaokrąglając wynik do bliższej podziałki. Zakładając jednostajny rozkład błędu odczytu czasu, obliczyć prawdopodobieństwo, że zmierzono czas z błędem przekraczającym 0,02 s.

Rozwiązanie. Niech X będzie błędem odczytu popełnionym przy pomiarze czasu. Gęstość tej zmiennej o rozkładzie jednostajnym w przedziale $(-0{,}05, 0{,}05)$ jest:

$$f(x)=\begin{cases}10 & \text{dla } |x|<0{,}05,\\ 0 & \text{dla pozostałych } x.\end{cases}$$

Korzystając z (2.3.6), mamy

$$P(|X|>0{,}02)=1-P(|X|\leqslant 0{,}02)=1-\int_{-0{,}02}^{0{,}02} 10\,dx=0{,}6.$$

ZADANIE 2.40. Automat produkuje odważniki 10-gramowe. Błędy pomiarów masy tych odważników mają rozkład normalny o wartości przeciętnej $\mu=0$ g i odchyleniu standardowemu $\sigma=0{,}01$ g. Znaleźć prawdopodobieństwo tego, że pomiar masy będzie przeprowadzony z błędem nie przekraczającym 0,02 g.

Rozwiązanie. $P(|X|<0{,}02)=P(-0{,}02<X<0{,}02)=$

$$=P\left(-\frac{0{,}02}{0{,}01}<\frac{X}{0{,}01}<\frac{0{,}02}{0{,}01}\right)=P(-2<W<2)=\Phi(2)-\Phi(-2)=$$

$$=\Phi(2)-[1-\Phi(2)]=2\Phi(2)-1=2\cdot 0{,}9773-1=0{,}9546.$$

ZADANIE 2.41. Niech zmienna losowa X ma rozkład $N(\mu, \sigma)$. Obliczyć $P(|X-\mu| < k\sigma)$ dla: a) $k=1{,}96$, b) $k=2{,}58$.

Rozwiązanie.

$$P(|X-\mu|<k\sigma)=P(-k\sigma<X-\mu<k\sigma)=$$

$$=P\left(-k<\frac{X-\mu}{\sigma}<k\right)=\Phi(k)-\Phi(-k)=2\Phi(k)-1,$$

gdzie $\Phi(k)$ jest wartością dystrybuanty rozkładu $N(0, 1)$ w punkcie k.

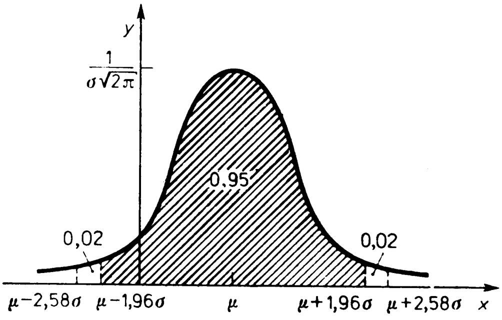
a) Przyjmijmy $k=1{,}96$, wówczas

$$P(|X-\mu|<1{,}96\sigma)=2\Phi(1{,}96)-1=2\cdot 0{,}975-1=0{,}95,$$

b) dla $k=2{,}58$

$$P(|X-\mu|<2{,}58\sigma)=2\Phi(2{,}58)-1=2\cdot 0{,}995-1=0{,}99.$$

Interpretację geometryczną powyższych wyników przedstawia rys. 2.38.



Rys. 2.38. $P(|X-\mu|<1{,}96\sigma)=0{,}95$ i $P(|X-\mu|<2{,}58\sigma)=0{,}99$ na rysunku gęstości f rozkładu $N(\mu, \sigma)$

ZADANIE 2.42. Wyznaczyć i naszkicować dystrybuantę rozkładu (2.8.29) i obliczyć medianę.

Rozwiązanie. Dystrybuanta jest określona wzorem:

$$F(x)=\begin{cases} 0 & \text{dla } x\leqslant 0, \\ \displaystyle\int_0^x \frac{2}{\lambda} t \exp\left(-\frac{t^2}{\lambda}\right) dt = 1-\exp\left(-\frac{x^2}{\lambda}\right) & \text{dla } x>0. \end{cases}$$

Aby wyznaczać punkty przegięcia funkcji F obliczamy jej drugą pochodną:

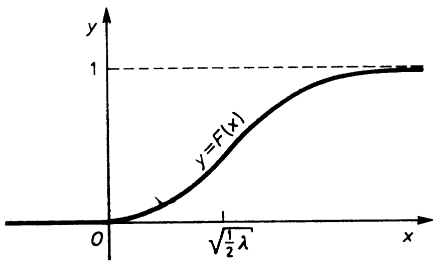
$$F'(x)=f(x)=\frac{2}{\lambda}x\exp\left(-\frac{x^2}{\lambda}\right) \quad \text{dla } x>0,$$

$$F''(x)=\frac{2}{\lambda}\exp\left(-\frac{x^2}{\lambda}\right)\left(1-\frac{2}{\lambda}x^2\right) \quad \text{dla } x>0$$

$$F''(x)=0 \Leftrightarrow x=\sqrt{\tfrac{1}{2}\lambda},$$

$F''(x)>0$ dla $0<x<\sqrt{\frac{1}{2}\lambda}$ oraz $F''(x)<0$ dla $x>\sqrt{\frac{1}{2}\lambda}$.
Stąd F ma punkt przegięcia, gdy $x=\sqrt{\frac{1}{2}\lambda}$.
Wykres dystrybuanty F przedstawia rysunek 2.39.

Rys. 2.39. Dystrybuanta rozkładu Rayleigha



Mediana jest rozwiązaniem równania: $F(x)=\frac{1}{2}$, mamy więc

$$1-\exp\left(-\frac{x^2}{\lambda}\right)=\frac{1}{2}, \quad \exp\left(-\frac{x^2}{\lambda}\right)=\frac{1}{2},$$

skąd mediana: $x_{0,5}=\sqrt{\lambda \ln 2}$.

Zadanie 2.43. Wykazać, że nie istnieje wartość przeciętna w rozkładzie Cauchy'ego o gęstości (2.8.37).

Rozwiązanie. Rozważmy najpierw przypadek szczególny $\lambda=1$, $\mu=0$. Ponieważ $EX=\int\limits_{-\infty}^{\infty} xf(x)\,dx$, więc należy zbadać, czy $\frac{1}{\pi}\int\limits_{-\infty}^{\infty}\frac{x\,dx}{1+x^2}$ jest zbieżna. Zauważmy jednak, że

$$\int\limits_{0}^{\infty}\frac{x\,dx}{1+x^2}=\tfrac{1}{2}[\ln(1+x^2)]_0^\infty$$

jest rozbieżna, a więc EX nie istnieje.

Gęstość (2.8.37) otrzymamy z gęstości:

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

przez zamianę zmiennych $y=\mu+\lambda x$, ponieważ $EY=\mu+\lambda EX$, a EX nie istnieje, więc EY również nie istnieje.

2.9. ZADANIA DO ROZWIĄZANIA

2.44. Dana jest funkcja prawdopodobieństwa zmiennej losowej X:

x_i	-5	-2	0	1	3	8
p_i	0,1	0,2	0,1	0,2	c	0,1

Wyznaczyć: a) stałą c, b) wykres funkcji prawdopodobieństwa i jej histogram, c) dystrybuantę i jej wykres, d) prawdopodobieństwa: d$_1$) $P(X=1)$, d$_2$) $P(X=2)$, d$_3$) $P(X<3)$, d$_4$) $P(X<2)$, d$_5$) $P(X\geqslant 0)$, d$_6$) $P(-2\leqslant X<3)$ dwoma sposobami, korzystając: 1° z danej funkcji prawdopodobieństwa, 2° z wyznaczonej dystrybuanty: znalezione prawdopodobieństwa zilustrować na rysunku dystrybuanty.

2.45. Dana jest dystrybuanta zmiennej losowej X:

x	$(-\infty, -2\rangle$	$(-2, 1\rangle$	$(1, 3\rangle$	$(3, +\infty)$
$F(x)$	0	0,2	0,8	1

Wyznaczyć jej funkcję prawdopodobieństwa.

2.46. Wyrazić za pomocą dystrybuanty następujące prawdopodobieństwa: a) $P(X\leqslant b)$, b) $P(X\geqslant b)$, c) $P(a<X\leqslant b)$, d) $P(a\leqslant X\leqslant b)$, e) $P(a<X<b)$.

2.47. Przeznaczona do odbioru partia towaru zawiera jednakową liczbę sztuk I, II, III gatunku. Niech ω_1, ω_2, ω_3 oznaczają zdarzenia elementarne w doświadczeniu polegającym na wylosowaniu z tej partii towaru (np. przy jej odbiorze) sztuki odpowiednio I, II, III gatunku. Na przestrzeni zdarzeń elementarnych $\Omega=\{\omega_1, \omega_2, \omega_3\}$ określamy dwie zmienne losowe X i Y:

zmienna losowa X: $X(\omega_1)=2$, $X(\omega_2)=1$, $X(\omega_3)=0$,

zmienna losowa Y: $Y(\omega_1)=0$, $Y(\omega_2)=1$, $Y(\omega_3)=2$.

a) Dlaczego zmienne losowe X i Y są różne: $X\neq Y$? b) Wyznaczyć funkcję prawdopodobieństwa zmiennej losowej: b$_1$) X, b$_2$) Y; porównać wyznaczone rozkłady. c) Mając na uwadze a) i b) wyjaśnić, czy zdania: „zmienne losowe X i Y są identyczne (są jednakowe)”, „zmienne losowe X i Y mają jednakowy rozkład”, są równoważne.

2.48. Podać przykład zmiennej losowej typu skokowego określonej na przestrzeni zdarzeń elementarnych, będącej przedziałem (a więc zbiorem, którego elementów nie da się ponumerować liczbami naturalnymi).

2.49. a) Czy zmienne losowe o tym samym rozkładzie prawdopodobieństwa mogą być: a$_1$) niezależne, a$_2$) zależne? b) Sprawdzić, że zmienne losowe X i Y z zadania 2.47 są zależne (zakładając, że zdarzenia jednoelementowe są jednakowo prawdopodobne).

2.50. Obsługa działa artyleryjskiego ma 3 pociski. Prawdopodobieństwo trafienia do celu jednym pociskiem (przy jednym wystrzale) w danych warunkach wynosi 0,7. Strzelanie kończy się z chwilą trafienia celu albo wyczerpania pocisków. Wyznaczyć: a) funkcję prawdopodobieństwa liczby oddanych strzałów, b) przeciętną liczbę oddanych strzałów.

2.51. Na drodze ruchu pociągów znajdują się w znacznej odległości od siebie 4 semafory, z których każdy (wobec znacznej odległości niezależnie od innych) zezwala na przejazd z prawdopodobieństwem $p=0{,}8$. Niech X oznacza liczbę semaforów zezwalających na przejazd i poprzedzających pierwsze zatrzymanie lub stację docelową. Znaleźć: a) funkcję prawdopodobieństwa zmiennej losowej X, b) dystrybuantę zmiennej losowej X, c) prawdopodobieństwo $P(X \geqslant 2)$.

2.52. Robotnik obsługuje 3 maszyny. Prawdopodobieństwo tego, że w ciągu godziny maszyna nie będzie wymagać jego interwencji wynosi 0,6 dla pierwszej oraz 0,7 dla drugiej i trzeciej maszyny. Przy założeniu, że maszyny pracują niezależnie od siebie, wyznaczyć funkcję prawdopodobieństwa liczby X maszyn, które w ciągu godziny ich pracy nie wymagają interwencji robotnika.

2.53. W ramach bieżącej kontroli jakości produkcji okresowo sprawdza się kolejno produkowane sztuki towaru, jednak nie więcej niż 4 i w przypadku stwierdzenia sztuki niezgodnej z wymaganiami jakości proces produkcji poddaje się regulacji. Liczba sztuk sprawdzonych w trakcie jednej takiej kontroli jest więc pewną zmienną losową X. Ponadto przyjmujemy, że: 1° prawdopodobieństwo wyprodukowania sztuki towaru niezgodnej z wymaganiami jakości jest stałe i wynosi $p=0{,}06$, 2° jakości kolejno produkowanych sztuk są zdarzeniami niezależnymi. Dla zmiennej losowej X wyznaczyć: a) funkcję prawdopodobieństwa, b) dystrybuantę, c) wartość przeciętną.

2.54. Zmienna losowa X ma funkcję prawdopodobieństwa postaci:

x_i	-3	-1	3	5
p_i	0,1	0,2	0,5	0,2

Wyznaczyć funkcję prawdopodobieństwa zmiennej losowej U, jeśli:
a) $U=2X+3$, b) $U=X^3$, c) $U=X^2-5$.

2.55. Niech: 1° p oznacza funkcję prawdopodobieństwa zmiennej losowej X, a $W_x=\{x_1, x_2, \ldots\}$ – zbiór jej punktów skokowych, 2° g – funkcję różnowartościową określoną co najmniej na zbiorze W_x, a h – funkcję do niej odwrotną, 3° q – funkcję prawdopodobieństwa zmiennej losowej $U=g(X)$, a $W_u=\{u_1, u_2, \ldots\}$ – zbiór punktów skokowych zmiennej losowej U. Wyjaśnić dlaczego

$$q_i \equiv q(u_i)=P(U=u_i)=P(X=h(u_i))=p(h(u_i)) \equiv p_i, \qquad u_i \in W_u.$$

Znaleźć funkcję g, h i q oraz zbiory W_x i W_u w punktach a) i b) zadania 2.54.

2.56. Zmienna losowa K ma funkcję prawdopodobieństwa postaci: $p_k \equiv P(K=k)=e^{-\lambda}\lambda^k/k!$, $k \in N_0$. Znaleźć funkcję prawdopodobieństwa zmiennej losowej $U=3K$.

2.57. Zmienna losowa K ma funkcję prawdopodobieństwa postaci:

$$p_k \equiv P(K=k)=\binom{4}{k}\left(\tfrac{1}{4}\right)^k\left(\tfrac{3}{4}\right)^{4-k} \qquad k \in \{0, 1, 2, 3, 4\}.$$

Znaleźć funkcję prawdopodobieństwa zmiennej losowej $U=K^2$.

2.58. Dane są funkcje prawdopodobieństwa niezależnych zmiennych losowych X i Y:

(1)

x_i	1	2	4
p_i	0,3	0,5	0,2

y_j	2	3	4
p_j	0,3	0,4	0,3

(2)

x_i	-1	0	1
p_i	0,2	0,6	0,2

y_j	-2	1	2	5
p_j	0,2	0,5	0,2	0,1

W obu przypadkach wyznaczyć funkcję prawdopodobieństwa zmiennej losowej U, jeśli: a) $U=X+Y$, b) $U=XY$.

2.59. Zmienna losowa ma funkcję prawdopodobieństwa postaci:

$$p_k \equiv P(K=k) = \tfrac{1}{3}\left(\tfrac{2}{3}\right)^{k-1}, \qquad k \in N$$

Wyznaczyć funkcję prawdopodobieństwa zmiennej losowej U, jeśli: a) $U=\cos \pi K$, b) $U=\cos \frac{1}{2}\pi K$, c) $U=\cos^2 \frac{1}{2}\pi K$, d) $U=\sin \frac{1}{2}\pi K+\cos \frac{1}{2}\pi K$.

2.60. Zmienna losowa X ma rozkład dwupunktowy: $P(X=x_1)=p$, $P(X=x_2)=1-p$, $p>0$. Dobrać tak funkcję g, aby zmienna losowa $Y=g(X)$ miała rozkład zero-jedynkowy ((2.7.5)).

2.61. Dwie maszyny wykonują ten sam rodzaj produkcji. Niech X oznacza liczbę produktów z usterkami wyprodukowanych w ciągu jednej zmiany przez pierwszą maszynę. Zmienna losowa Y ma analogiczne znaczenie dla drugiej maszyny. Dane są funkcje prawdopodobieństwa tych zmiennych:

x_i	0	1	2
p_i	0,3	0,4	0,3

y_j	0	1	2	3
p_j	0,1	0,4	0,4	0,1

Zakładając brak jakiejkolwiek zależności między jakością produkcji obu maszyn, znaleźć funkcję prawdopodobieństwa zmiennej losowej: a) $U=X+Y$, b) $V=XY$; jaką interpretację ma zmienna losowa U?

2.62. Niech zmienne losowe U_1, U_2 będą funkcjami zmiennych losowych X i Y, określonymi jak następuje: $U_1=\min(X, Y)$, $U_2=\max(X, Y)$. Sprawdzić, że: $U_1=\frac{1}{2}[X+Y-|X-Y|]$, $U_2=\frac{1}{2}[X+Y+|X-Y|]$.

2.63. Dla zmiennej losowej X z zadania 2.54 wyznaczyć: a) wartość przeciętną, b) medianę, c) kwantyl $x_{0,3}$, d) wariancję (dwoma sposobami), e) odchylenie standardowe, f) odchylenie przeciętne, g) współczynnik zmienności, h) współczynnik nierównomierności, i) drugi moment zwykły, j) trzeci moment zwykły, k) trzeci moment centralny, l) współczynnik asymetrii (czy można było przewidzieć jego znak?).

2.64. Charakterystykę liczbową zmiennej losowej nazywamy *miarą położenia*, jeśli dodanie stałej do zmiennej losowej powoduje dodanie tej stałej do charakterystyki. Sprawdzić, że w tym sensie miarami położenia są: wartość przeciętna, mediana i kwantyl dowolnego rzędu.

2.65. Udowodnić własności wariancji dane wzorami: (2.6.23) - (2.6.28).

2.66. Które spośród własności wariancji (2.6.23) - (2.6.28) szczególnie wyraźnie dowodzą, że jest ona miarą rozproszenia?

2.67. Dane są funkcje prawdopodobieństwa zmiennych losowych X i Y:

x_i	0	1
p_i	0,5	0,5

y_i	1/3	2
p_i	0,9	0,1

Są to więc różne rozkłady prawdopodobieństwa (i oczywiście różne zmienne losowe). Sprawdzić, że $EX=EY$, $EX^2=EY^2$ (a więc również $D^2X=D^2Y$) i dopiero $EX^3 \neq EY^3$.

2.68. Wyznaczyć wartość przeciętną i wariancję zmiennych losowych X i Y określonych w zadaniu 2.47 i porównać je. Dokończyć zdanie: nieidentyczne zmienne losowe mające jednakowe rozkłady mają również ... wartości przeciętne i ... wariancje.

2.69. Niezależne zmienne losowe X i Y mają jednakową funkcję prawdopodobieństwa:

x_i	0	1	2
p_i	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Niech: $U_1=X+Y$, $U_2=2X$, $U_3=XY$, $U_4=X^2$. Dla zmiennych losowych U_i wyznaczyć: a) ich funkcje prawdopodobieństwa, b) wartości przeciętne, c) wariancje.

Porównać otrzymane rezultaty dla zmiennych losowych U_1 i U_2 oraz U_3 i U_4.

2.70. Wyznaczyć wartość przeciętną i wariancję zmiennej losowej: a) $U_1=2X+1$, b) $U_2=X^2$, c) $U_3=-X^2+2$, jeżeli $EX=2$, $D^2X=1$, $EX^4=34$.

2.71. Dana jest funkcja prawdopodobieństwa zmiennej losowej X:

x_i	-2	-1	2	5
p_i	0,3	0,1	0,2	0,4

Wyznaczyć dwoma sposobami wartość przeciętną i wariancję zmiennej losowej $U=2X-3$: 1° znajdując najpierw funkcję prawdopodobieństwa zmiennej losowej U oraz 2° korzystając z odpowiednich własności wartości przeciętnej i wariancji.

2.72. Wyznaczyć: a) wartość przeciętną i wariancję zmiennej losowej $U_1=3X-2Y$, b) wartość przeciętną zmiennej losowej $U_2=X^2-Y^2$, jeżeli zmienne losowe X i Y są niezależne oraz $EX=-3$, $EY=4$, $D^2X=0{,}5$; $D^2Y=2$.

2.73. Wyznaczyć funkcję prawdopodobieństwa skokowej zmiennej losowej X mającej tylko dwa punkty skokowe x_1 i x_2, jeśli: $x_1<x_2$, $p_1\equiv P(X=x_1)=0{,}2$, $EX=3$, $D^2X=4$.

2.74. Dana jest funkcja prawdopodobieństwa zmiennej losowej X:

x_i	-2	-1	1	2
p_i	0,25	0,25	0,25	0,25

Niech $Y=X^2$. Sprawdzić, że w tym przypadku mimo, że zmienne losowe X i Y nie są niezależne (bo Y jest funkcją X), to jednak $D^2(X+Y)=D^2X+D^2Y$.

2.75. Niech X i Y będą dowolnymi zmiennymi losowymi mającymi skończone wartości przeciętne. Sprawdzić, że $D^2(X+Y)=D^2X+D^2Y+2E[(X-EX)(Y-EY)]$. Korzystając z tej równości wykazać prawdziwość (2.6.26) dla niezależnych zmiennych losowych X i Y.

2.76. Wykazać, że jeśli: 1° zmienne losowe X, Y są niezależne, 2° ich wartości przeciętne są równe zeru, to $D^2(XY)=D^2X\cdot D^2Y$.

2.77. Uzasadnić, że: wartość przeciętna kwadratu odchylenia zmiennej losowej od jej wartości przeciętnej jest nie większa, niż wartość przeciętna kwadratu odchylenia tej zmiennej losowej od dowolnej liczby c: $E(X-EX)^2\leqslant E(X-c)^2$, przy tym równość zachodzi tylko wtedy, gdy $c=EX$.

2.78. Sprawdzić, że $D^2X=E[X(X-1)]-EX\cdot(EX-1)$.

2.79. Niech $R=X_{\max}-X_{\min}$ oznacza rozstęp zmiennej losowej X. a) Sprawdzić, że jeśli zmienna losowa X ma dwupunktowy rozkład równomierny ((2.7.1) dla $n=2$), to $D^2X=\frac{1}{4}R^2$. b) Wykazać, że dla dowolnej zmiennej losowej X spełniona jest nierówność:

$D^2X \leqslant \frac{1}{4}R^2$. c) Sprawdzić: c_1) bezpośrednim rachunkiem, c_2) wykorzystując nierówność z punktu b) tego zadania, że wariancja w rozkładzie dwumianowym jest nie większa od $\frac{1}{4}n$.

2.80. Wykazać, że jeśli zmienne losowe $X_1, \ldots, X_n$ mają ten sam rozkład prawdopodobieństwa, α_1 i σ^2 oznaczają odpowiednio ich wspólną wartość przeciętną i wariancję, to ich średnia arytmetyczna $\bar{X} = \dfrac{1}{n}(X_1 + \ldots + X_n)$ ma tę samą wartość przeciętną: $E\bar{X} = \alpha_1$; gdy ponadto zmienne losowe $X_1, \ldots, X_n$ są niezależne, wtedy $D^2\bar{X} = \dfrac{1}{n}\sigma^2$ (jest więc n razy mniejsza od wariancji każdej ze zmiennych losowych X_i).

2.81. Wykazać, że jeśli zmienne losowe $X_1, \ldots, X_n$: 1° są niezależne, 2° mają ten sam rozkład prawdopodobieństwa, 3° przyjmują tylko dodatnie wartości oraz 4° $i_1, \ldots$ $\ldots, i_k$ jest k-wyrazowym podciągiem ciągu $1, \ldots, n$, to

a) $E\left(\dfrac{X_i}{X_1 + \ldots + X_n}\right) = \dfrac{1}{n}, \quad i = 1, \ldots, n,$

b) $E\left(\dfrac{X_{i_1} + \ldots + X_{i_k}}{X_1 + \ldots + X_n}\right) = \dfrac{k}{n}, \quad k = 1, \ldots, n.$

2.82. Wyznaczyć wartość przeciętną, wariancję i odchylenie standardowe zmiennej losowej X z zadania 2.51.

2.83. W nawiązaniu do zadania 2.52 znaleźć średnią liczbę maszyn, które w ciągu godziny nie wymagają interwencji obsługującego je robotnika.

2.84. Za pomocą doświadczeń wykonywanych zgodnie z warunkami schematu Bernoulliego (z prawdopodobieństwem p sukcesu w pojedynczym doświadczeniu) określamy następującą grę. Osoba decydująca się na udział w grze przeprowadza jedno doświadczenie Bernoulliego (np. rzuca monetą, rzuca kostką, losuje ze zwrotem). Gdy wynikiem tego doświadczenia jest porażka, wtedy płaci ona m zł. Natomiast w przypadku sukcesu, powtarza to doświadczenie aż do pojawienia się porażki, otrzymując a^{k-1} zł, jeśli porażka ta zaszła w k-tym powtórzeniu, $k = 2, 3, \ldots$ Wygrana osoby, która jeden raz wzięła udział w grze, jest zmienną losową W przyjmującą wartości: $-m, a^1, \ldots, a^k, \ldots$ Gra nazywa się sprawiedliwą, gdy wartość przeciętna wygranej jest równa zeru: $EW = 0$.

a) Przy ustalonych p i a dobrać tak m, aby gra była sprawiedliwa.

b) Przy ustalonych p i m dobrać tak a, aby gra była sprawiedliwa; obliczyć a, jeśli: b_1) $m = 1$, $p = \frac{1}{2}$; b_2) $m = 0{,}2$, $p = \frac{1}{6}$; b_3) $m = 1$, $p = \frac{1}{6}$; b_4) $m = 2$, $p = \frac{1}{6}$; b_5) $m = 5$, $p = \frac{1}{6}$.

c) Przy jakich wartościach a, gdy ustalone są wartości p i m, gra byłaby krzywdząca dla: c_1) osoby podejmującej grę, c_2) dla właściciela gry?

2.85. Wiadomo, że zbiory: a) $\{EX, \operatorname{Var} X, F(0), F(3)\}$ i $\{-5, \frac{1}{2}, \frac{1}{3}, 2\}$ oraz b) $\{x_{0,5}, EX, F(4), F(5)\}$ i $\{\frac{2}{3}, \frac{20}{21}, 1, 2\}$ są identyczne. W obu przypadkach zidentyfikować elementy pierwszego zbioru, jeśli w b) zakładamy ponadto, że zmienna losowa ma tylko dodatnie punkty skokowe, z których 4 i 5 są kolejnymi punktami skokowymi.

2.86. Sprawdzić, że wariancja zmiennej losowej o rozkładzie jednopunktowym jest równa zeru ((2.7.3)). Wykazać prawdziwość twierdzenia odwrotnego.

2.87. Znaleźć wartości funkcji prawdopodobieństwa $P(k;\,5,\,\frac{3}{4})$ rozkładu dwumianowego. Odczytać stąd wartość najbardziej prawdopodobną i porównać zgodność odczytu z odpowiednimi wzorami ((2.7.8)).

2.88. Wyrazić parametry n i p rozkładu dwumianowego za pomocą momentów centralnych μ_2 i μ_3.

2.89. Obliczyć prawdopodobieństwo pojawienia się co najmniej raz zdarzenia A w trzech niezależnych doświadczeniach, jeśli prawdopodobieństwo pojawienia się tego zdarzenia w jednym doświadczeniu wynosi p.

2.90. Bezpośrednim rachunkiem sprawdzić, że suma dwu niezależnych zmiennych losowych o rozkładzie dwumianowym odpowiednio z parametrami (n_1, p), (n_2, p) (a więc z tym samym parametrem p) ma również rozkład dwumianowy z parametrami (n, p), gdzie $n = n_1 + n_2$.

2.91. Przy jakim p, w n niezależnych doświadczeniach z prawdopodobieństwem p sukcesu w każdym doświadczeniu, najbardziej prawdopodobną liczbą sukcesów będą (będzie) liczby (liczba): a) n oraz $n-1$, b) 1 oraz 0, c) n, d) 0?

2.92. Sprawdzić, że dla rozkładu dwumianowego funkcja prawdopodobieństwa $P(k;\,n, p)$ spełnia następujący związek rekurencyjny:

$$P(k+1\,;\,n\,,p) = \frac{n-k}{k+1}\,\frac{p}{1-p}\,P(k\,;\,n\,,p).$$

Obliczyć $P(4;\,7,\,\frac{1}{4})$, jeśli $P(3;\,7,\,\frac{1}{4}) = 0{,}173$.

2.93. Wyobraźmy sobie, że czytamy opis pojedynku przebiegającego według następujących zasad. Każdy z przeciwników otrzymuje rewolwer bębenkowy z jednym nabojem umieszczonym w jednej z pięciu komór rewolweru (pozostałe 4 komory są puste) i wielokrotnie obraca bębenek. Następnie jeden z nich pociąga za język spustowy rewolweru i tylko w przypadku niezranienia próbuje oddać strzał drugi przeciwnik. Obliczyć prawdopodobieństwo, że pojedynek skończy się bezkrwawo, gdy prawdopodobieństwo trafienia do celu dla każdego z przeciwników jest jednakowe i wynosi 0,8.

2.94. Półfabrykaty dostarczane są na linię automatyczną za pomocą podajnika, tj. urządzenia składającego się ze zbiornika, m czerpaków i kanału odprowadzającego. Czerpaki, rozmieszczone równomiernie na obwodzie obracającej się osi, chwytają półfabrykaty przechowywane luzem w zbiorniku i kanałem odprowadzającym przekazują je na linię automatyczną. Zakładając, że: 1° prawdopodobieństwo uchwycenia półfabrykatu i przekazania go na linię automatyczną jest jednakowe dla wszystkich czerpaków i wynosi $p = p_0 c^v$, gdzie $0 < p_0 < 1$, $c > 1$, zaś v jest liczbą obrotów osi czerpaka w ciągu minuty (p jest więc funkcją liczby obrotów), 2° w zbiorniku zawsze znajdują się półfabrykaty w zasięgu wszystkich czerpaków (co pozwala przyjąć, że wynik pracy dowolnego z nich nie ma wpływu na rezultat pracy każdego pozostałego), znaleźć: a) prawdopodobieństwo tego, że liczba K półfabrykatów przekazanych przez podajnik w rezultacie jednego obrotu wyniesie k, b) prawdopodobieństwo tego, że liczba X półfabrykatów przekazanych przez podajnik w czasie jednej minuty wyniesie x, c) średnią liczbę półfabrykatów przekazanych przez podajnik w czasie: c_1) jednego obrotu, c_2) jednej minuty, d) liczbę obrotów v_0, przy której podajnik osiągnie maksymalną średnią wydajność (tj. maksymalną średnią liczbę półfabrykatów przekazanych podajnikiem w czasie minuty). Przeprowadzić obliczenia dla $m=12$, $p_0=0{,}7$, $c=1{,}04$.

2.95. Przyjmujemy, że uszkodzenie urządzenia wytwarzającego produkt (w sztukach) może nastąpić tylko na skutek awarii pewnego jego podzespołu (np. przepalenia się bezpiecznika, złamania noża obrabiarki, itp.). Dla zwiększenia niezawodności tego urządzenia wyposażono go w $v-1=3$ dodatkowe tego samego rodzaju podzespoły, które automatycznie (i pojedynczo) włączają się, gdy ulegnie awarii pracujący podzespół. Tak więc urządzenie przerywa pracę dopiero po v-krotnej awarii tego podzespołu. Zakładając ponadto, że: 1° w ciągu godziny wytwarzana jest jedna sztuka produktu, 2° prawdopodobieństwo powstania uszkodzenia przy produkcji kolejnych sztuk (czyli w kolejnych godzinach) jest stałe i wynosi $p=0{,}005$, 3° zdarzenia polegające na powstaniu uszkodzenia przy produkcji kolejnych sztuk są niezależne, znaleźć: a) funkcję prawdopodobieństwa czasu T (w godzinach) pracy urządzenia, b) dystrybuantę czasu pracy urządzenia, c) prawdopodobieństwo tego, że czas pracy wyniesie: c_1) dokładnie 3 godziny, c_2) co najmniej 3 godziny, c_3) dokładnie 24 godziny, c_4) nie mniej niż 16 godzin i nie więcej niż 24 godziny, d) średni czas pracy tego urządzenia.

2.96. Rozwiązać zadanie 2.95 dla urządzenia pracującego bez podzespołów zapasowych.

2.97. Treść i polecenia jak w zadaniu 2.95 z tym, że: 1° urządzenie pracuje bez podzespołów zapasowych, 2° prawdopodobieństwo α_k powstania uszkodzenia podzespołu w czasie jednej godziny jego pracy rośnie z każdą godziną i w k-tej godzinie wynosi $\alpha_k=1-\dfrac{0{,}999}{k}$.

2.98. Sprawdzić, że dla ujemnego rozkładu dwumianowego funkcja prawdopodobieństwa $P(k;\, v, p)$ spełnia następujący związek rekurencyjny:

$$P(k+1\,;\; v, p)=\frac{k}{k-v+1}\,qP(k\,;\; v, p).$$

Obliczyć $P(7;\,4,\,\frac{1}{2})$, jeśli $P(6;\,4,\,\frac{1}{2})=\frac{5}{32}$.

2.99. Znaleźć wartości funkcji prawdopodobieństwa $P(k;\,10,\,6,\,5)$ rozkładu hipergeometrycznego.

2.100. Obliczyć prawdopodobieństwo przyjęcia partii N sztuk towaru, wśród których jest M sztuk wadliwych, jeśli partię przyjmuje się, gdy w n-elementowej próbce (losowanej bez zwrotu) z tej partii znajdzie się co najwyżej jedna sztuka niedobra. Wykonać rachunki, jeśli: a) $N=50$, $M=8$, $n=5$, b) $N=200$, $M=20$, $n=10$.

2.101. Znaleźć te wartości k, dla których funkcja rozkładu prawdopodobieństwa $P(k;\,N,\,M,\,n)$, dana wzorem (2.7.11) przyjmuje wartości dodatnie.

2.102. Sprawdzić, że dla rozkładu hipergeometrycznego funkcja prawdopodobieństwa $P(k;\,N,\,M,\,n)$ spełnia następujący związek rekurencyjny:

$$P(k+1\,;\; N, M, n)=\frac{n-k}{k+1}\,\frac{M-k}{N-M-n+k+1}\,P(k\,;\; N, M, n).$$

Obliczyć $P(3;\,10,\,4,\,5)$, jeśli $P(2;\,10,\,4,\,5)=0{,}476$.

2.103. Prawdopodobieństwo wyprodukowania sztuki wadliwej wynosi $p=0{,}02$. Obliczyć prawdopodobieństwo, że w partii towaru liczącej 300 sztuk znajdzie się: a) zero sztuk wadliwych, b) jedna sztuka wadliwa, c) dwie sztuki wadliwe, d) co najmniej trzy sztuki wadliwe.

2.104. Urządzenie składa się między innymi z 750 lamp. Prawdopodobieństwo awarii każdej lampy w ciągu jednej doby pracy urządzenia jest jednakowe i wynosi $p=0{,}004$. Obliczyć prawdopodobieństwo, że w ciągu jednej doby pracy urządzenia ulegnie awarii: a) 0 lamp, b) 1 lampa, c) 2 lampy, d) co najmniej 3 lampy.

2.105. Bezpośrednim rachunkiem sprawdzić, że suma dwóch niezależnych zmiennych losowych o rozkładach Poissona odpowiednio z parametrami λ_1 i λ_2 ma również rozkład Poissona z parametrem $\lambda=\lambda_1+\lambda_2$ (także zad. 4.5).

2.106. Weźmy pod uwagę zbiór N elementów, wśród których jest M elementów mających wyróżnioną cechę i $N-M$ elementów nie mających tej cechy. Niech ponadto r będzie dowolną ustaloną liczbą całkowitą. Ze zbioru wszystkich elementów losujemy n elementów zgodnie z następującym schematem (*schemat Polya*). Losujemy jeden element, następnie zwracamy go z powrotem i, w zależności od tego czy $r\geqslant 0$, czy $r<0$, dodatkowo dodajemy albo pobieramy r takich elementów, jak wylosowany; czynność tę powtarzamy n-krotnie. a) Jaki warunek musi spełniać iloczyn rn, gdy $r<0$? b) Niech K oznacza możliwą liczbę elementów mających wyróżnioną cechę wśród n wylosowanych, znaleźć $P(K=k)$, $k=0, 1, \ldots, n$. c) Znaleźć wartość przeciętną zmiennej losowej K. d) Rozważyć przypadki szczególne poleceń b) i c), gdy: c_1) $r=0$, c_2) $r=-1$.

2.107. Zmienna losowa X ma rozkład o gęstości

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

Wyznaczyć jej dystrybuantę oraz obliczyć $P(X\geqslant 2)$.

2.108. Dobrać tak stałą c by funkcja

$$f(x)=\begin{cases} c\sin x & \text{dla} \quad 0\leqslant x\leqslant \pi, \\ 0 & \text{poza tym} \end{cases}$$

była gęstością, a następnie: a) wyznaczyć jej dystrybuantę, b) obliczyć $P(|X|<\frac{1}{3}\pi)$ i zinterpretować za pomocą wykresu gęstości i dystrybuanty.

2.109. Zmienna losowa X ma rozkład o gęstości

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

a) Naszkicować wykres gęstości.

b) Wyznaczyć i naszkicować dystrybuantę tego rozkładu.

c) Obliczyć prawdopodobieństwo zdarzenia A, że w dwóch niezależnych doświadczeniach co najmniej raz zmienna losowa X przyjmie wartość z przedziału $\langle 1, 2)$.

2.110. Z pewnego przystanku autobusy odjeżdżają co 10 minut. Zakładamy, że rozkład czasu przybycia pasażera na przystanek jest jednostajny, obliczyć prawdopodobieństwo, że pasażer będzie czekał co najmniej 4 minuty.

2.111. W pewnej miejscowości temperatura T (w stopniach Celsjusza) mierzona w dniu 1 marca o godz. 8 jest zmienną losową o gęstości:

$$f(t)=\tfrac{1}{2}e^{-|t|}.$$

Wyznaczyć jej dystrybuantę oraz obliczyć $P(|T|<2)$.

2.112. Dobrać tak stałe A i B żeby funkcja

$$F(x)=\begin{cases} A+B\arccos x & \text{dla} \quad |x|<1, \\ 0 & \text{dla} \quad x\leqslant -1, \\ 1 & \text{dla} \quad x\geqslant 1 \end{cases}$$

była dystrybuantą zmiennej losowej X typu ciągłego. Wyznaczyć jej gęstość.

2.113. Wyznaczyć tak stałą a, by funkcja

$$F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 1, \\ 2\left(1-\dfrac{1}{x}\right) & \text{dla} \quad 1<x\leqslant a, \\ 1 & \text{dla} \quad x>a \end{cases}$$

była dystrybuantą zmiennej losowej X typu ciągłego. Obliczyć $P(-1\leqslant X\leqslant 1{,}5)$ i zinterpretować je za pomocą wykresu gęstości.

2.114. Dystrybuanta zmiennej losowej X jest określona wzorem:

$$F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{1}{2}t^2\right)dt.$$

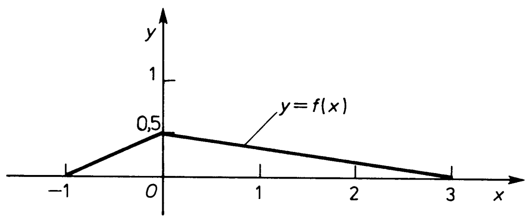
Wyznaczyć gęstość i naszkicować jej wykres.

2.115. Zbadać czy funkcje:

a) $F(x)=c\cos x$ dla $0\leqslant x\leqslant\frac{1}{2}\pi$,

b) $F(x)=c\sin x$ dla $0\leqslant x\leqslant\frac{1}{2}\pi$

mogą być dystrybuantami zmiennej losowej X? Jeśli tak, to przy jakich wartościach c?



Rys. 2.40. Do zadania 2.116

2.116. Wykres gęstości zmiennej losowej X przedstawiono na rysunku 2.40. Wyznaczyć gęstość f zmiennej losowej X oraz jej dystrybuantę.

2.117. Zmienna losowa X ma rozkład o gęstości

$$f(x)=\begin{cases} 3x^2 & \text{dla} \quad 0\leqslant x\leqslant 1, \\ 0 & \text{poza tym}. \end{cases}$$

Obliczyć: a) wartość przeciętną EX, b) medianę $x_{0,5}$, c) odchylenie przeciętne od mediany.

2.118. Zmienna losowa X ma rozkład o gęstości

$$f(x)=\begin{cases}\frac{3}{4}(1-x^2) & \text{dla } |x|\leqslant 1,\\ 0 & \text{poza tym}.\end{cases}$$

Naszkicować wykres gęstości. Wyznaczyć: wartość przeciętną, modę, medianę i trzeci moment centralny zmiennej losowej X.

2.119. Wyznaczyć modę i medianę w rozkładzie o gęstości:

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

2.120. Wyznaczyć modę i medianę dla rozkładu o gęstości:

$$f(x)=\begin{cases}\dfrac{ak(x-x_0)^{a-1}}{[1+k(x-x_0)^a]^2} & \text{dla } x_0\leqslant x<\infty,\\ 0 & \text{poza tym},\end{cases}$$

jeśli a, k są stałe spełniające warunki: $k>0$, $a>1$.

2.121. Wyznaczyć modę, medianę i wartość przeciętną dla rozkładu o gęstości:

$$f(x)=\begin{cases}\frac{1}{4}|\sin x| & \text{dla } 0\leqslant x\leqslant 2\pi,\\ 0 & \text{poza tym}.\end{cases}$$

2.122. Wyznaczyć modę i trzeci moment zwykły dla zmiennej losowej X o gęstości:

$$f(x)=\begin{cases}\dfrac{\exp\left(-\dfrac{1}{x}\right)}{n!\,x^{n+2}} & \text{dla } x>0,\\ 0 & \text{dla } x\leqslant 0,\end{cases}$$

gdy $n=3, 4, \ldots$

2.123. Naszkicować wykres gęstości prawdopodobieństwa zmiennej losowej X:

$$f(x)=\begin{cases}\dfrac{1}{2x^2} & \text{dla } |x|>1,\\ 0 & \text{dla } |x|\leqslant 1.\end{cases}$$

Wyznaczyć medianę, wartość przeciętną i wariancję zmiennej losowej X.

2.124. Automat produkuje kulki metalowe o średnicach X (w cm). Średnica X jest zmienną losową o gęstości

$$f(x)=\begin{cases}5 & \text{dla } 0{,}4\leqslant x\leqslant 0{,}6,\\ 0 & \text{dla pozostałych } x.\end{cases}$$

Wyznaczyć wartość przeciętną objętości tych kulek.

2.125. Pociągi kolejki elektrycznej odjeżdżają ze stacji co 5 minut. Zakładając, że rozkład czasu przybycia pasażera na stację jest jednostajny, obliczyć wartość przeciętną i wariancję czasu oczekiwania na pociąg.

2.126. Obliczyć współczynnik asymetrii i współczynnik skupienia rozkładu o gęstości:

$$f(x)=\begin{cases} 6x(1-x) & \text{dla} \quad 0 \leqslant x \leqslant 1, \\ 0 & \text{poza tym}. \end{cases}$$

2.127. Czas bezawaryjnej pracy pewnego urządzenia (w godz.) jest zmienną losową o rozkładzie wykładniczym (2.8.7) z parametrem $\lambda=2$. Obliczyć współczynnik asymetrii i spłaszczenia dla tego rozkładu.

2.128. Niech X_1 będzie zmienną losową o gęstości

$$f(x)=\begin{cases} \frac{1}{6}(x+2) & \text{dla} \quad -2<x \leqslant 0, \\ -\frac{1}{12}(x-4) & \text{dla} \quad 0<x \leqslant 4, \\ 0 & \text{poza tym} \end{cases}$$

oraz X_2 — zmienną losową o gęstości:

$$g(x)=\begin{cases} \frac{1}{12}(x+2) & \text{dla} \quad -2<x \leqslant 2, \\ -\frac{1}{6}(x-4) & \text{dla} \quad 2<x \leqslant 4, \\ 0 & \text{poza tym}. \end{cases}$$

Naszkicować wykresy obydwu gęstości we wspólnym układzie współrzędnych. Obliczyć współczynniki asymetrii dla każdego z rozkładów i porównać otrzymane wyniki z asymetrią wykresów gęstości.

2.129. Obliczyć wartość przeciętną, wariancję i współczynnik skupienia dla rozkładów o gęstościach:

a) $f_1(x)=\begin{cases} \frac{3}{4}(2x-x^2) & \text{dla} \quad 0 \leqslant x \leqslant 2, \\ 0 & \text{poza tym}, \end{cases}$

b) $f_2(x)=\dfrac{1}{2\sqrt{10}}\exp\left(-\dfrac{|x-1|}{\sqrt{10}}\right).$

Porównać otrzymane charakterystyki dla obydwu rozkładów.

2.130. Zmienna losowa X ma rozkład o dystrybuancie:

$$F(x)=\begin{cases} 0 & \text{dla} \quad x \leqslant 1, \\ \dfrac{x-1}{2} & \text{dla} \quad 1<x \leqslant 3, \\ 1 & \text{dla} \quad x>3. \end{cases}$$

Wyznaczyć odchylenie standardowe i odchylenie przeciętne od wartości średniej tej zmiennej losowej.

2.131. Wyrazić trzeci moment centralny zmiennej losowej X przez jej momenty zwykłe.

2.132. Wyrazić trzeci moment zwykły zmiennej losowej X przez jej momenty centralne.

2.133. Wykazać, że jeśli X jest zmienną losową typu ciągłego przyjmującą tylko wartości dodatnie, to zachodzi nierówność

$$E\left(\frac{1}{X}\right) \geqslant \frac{1}{EX}$$

o ile istnieją $E\left(\dfrac{1}{X}\right)$, EX.

2.134. Korzystając z własności $EX=E(X-a)+a$, obliczyć wartość przeciętną zmiennej losowej X o gęstości:

$$f(x)=\begin{cases} 7(x-1)^6 & \text{dla} \quad 0\leqslant x\leqslant 1, \\ 0 & \text{poza tym}. \end{cases}$$

2.135. Zmienna losowa typu ciągłego ma rozkład jednostajny w przedziale $\langle -5, 15\rangle$. Wyznaczyć i naszkicować jej gęstość. Obliczyć wartość przeciętną i współczynnik skupienia (*kurtozę*) tej zmiennej.

2.136. Zapałkę o długości 5 cm złamano w dowolnym punkcie. Zakładając, że rozkład prawdopodobieństwa długości krótszej części zapałki jest jednostajny, obliczyć prawdopodobieństwo, że długość krótszej części zapałki nie przekracza 0,5 cm.

2.137. Wykazać, że rozkład wykładniczy (2.8.7) jest rozkładem o asymetrii prawej, którego współczynnik asymetrii $\gamma_1=2$. Obliczyć współczynnik skupienia tego rozkładu.

2.138. Podziałka skali woltomierza jest wycechowana co 0,5 V. Wskazania woltomierza zaokrągla się do najbliższego punktu podziału. Znaleźć prawdopodobieństwo tego, że przy odczycie zostanie popełniony błąd przekraczający 0,1 V.

2.139. Naszkicować we wspólnym układzie współrzędnych gęstości rozkładów: $N(1, 1)$, $N(1, 2)$, $N(1, \frac{1}{2})$.

2.140. Niech zmienna losowa X ma rozkład $N(\frac{3}{2}, 2)$. Obliczyć prawdopodobieństwo: $P(X<2{,}5)$, $P(X>-0{,}5)$, $P(0{,}5<X<2)$, $P(|2X-1|<1)$, $P(|X|>0{,}5)$.

2.141. Wykazać, że jeśli X jest zmienną losową o rozkładzie $N(0, \sigma)$, to

$$\bigwedge_{\alpha>0} E(|X|^{\alpha})=\frac{(\sigma\sqrt{2})^{\alpha}}{\sqrt{\pi}}\,\Gamma\left(\frac{\alpha+1}{2}\right).$$

2.142. Pewien automat produkuje części, których długość jest zmienną losową o rozkładzie $N(2; 0,2)$ (w cm). Wyznaczyć prawdopodobieństwo otrzymania braku, jeśli dopuszczalne długości części powinny zawierać się w przedziale $(1{,}7; 2{,}3)$.

2.143. Błąd losowy X pomiaru odległości od drogowskazu jest zmienną losową o gęstości:

$$f(x)=\frac{1}{5\sqrt{2\pi}}\exp\left[-\frac{(x-10)^2}{50}\right].$$

Obliczyć EX, D^2X, μ_4.

2.144. Pewien przyrząd pomiarowy robi błąd systematyczny 1 m w stronę zawyżenia pomiaru i błąd losowy o rozkładzie $N(0; 0{,}5)$.

a) Obliczyć wartość przeciętną błędu pomiaru.

b) Wyznaczyć prawdopodobieństwo tego, że błąd z jakim są mierzone badane przedmioty nie przekracza 2 m.

2.145. Wytrzymałość stalowych lin pochodzących z produkcji masowej jest zmienną losową o rozkładzie $N(1000\ \text{kg/cm}^2, 50\ \text{kg/cm}^2)$. Obliczyć jaki procent lin ma wytrzymałość mniejszą od 900 kg/cm^2.

2.145. Automat produkuje nity. Średnice główek nitów są wartościami zmiennej losowej o rozkładzie $N(2; 0{,}1)$ (w mm). Jakie rozmiary średnicy z przedziału $(2-\varepsilon, 2+\varepsilon)$ można gwarantować z prawdopodobieństwem 0,95?

2.147. Wykazać, że jeśli X jest zmienną losową o rozkładzie $N(\mu, \sigma)$, to: $E|X-\mu| = \sqrt{\frac{2}{\pi}}\sigma$.

2.148. Czas bezawaryjnej pracy X pewnego urządzenia ma rozkład wykładniczy (2.8.7) z parametrem $\lambda = 5$.

Obliczyć:

a) wartość przeciętną bezawaryjnego czasu pracy urządzenia,

b) medianę, c) prawdopodobieństwo, że bezawaryjny czas pracy urządzenia wynosi co najmniej 5 godzin.

2.149. Czas T (w min) pomiędzy przybyciem dwóch taksówek na postój jest zmienną losową o dystrybuancie:

$$F(t) = \begin{cases} 1-\exp(-\frac{1}{3}t) & \text{dla } t>0, \\ 0 & \text{poza tym.} \end{cases}$$

a) Obliczyć prawdopodobieństwo, że $1<T<2$. b) Wyznaczyć gęstość tego rozkładu. c) Obliczyć ET i D^2T.

2.150. Dobrać tak stałą c, by funkcja

$$f(x) = \begin{cases} cxe^{-4x^2} & \text{dla } x>0, \\ 0 & \text{dla pozostałych } x \end{cases}$$

była gęstością prawdopodobieństwa zmiennej losowej X ((2.8.29)). Obliczyć wartość przeciętną i wariancję tej zmiennej losowej.

2.151. Korzystając ze wzoru: $\mu_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3$, obliczyć współczynnik asymetrii rozkładu Erlanga o gęstości (2.8.15).

2.152. Znaleźć modę i wartość przeciętną zmiennej losowej X o rozkładzie Maxwella (2.8.31) z parametrem $\lambda=1$.

2.153. Wykazać, że rozkład gamma (2.8.12) jest dla $p>1$ rozkładem jednomodalnym. Wyznaczyć modę.

2.154. Amplituda X kołysania bocznego statku jest zmienną losową o gęstości (2.8.29):

$$f(x) = \begin{cases} x\exp(-\frac{1}{2}x^2) & \text{dla } x>0, \\ 0 & \text{poza tym.} \end{cases}$$

a) Obliczyć wartość przeciętną, medianę i modę zmiennej losowej X.

b) Obliczyć, z jakim prawdopodobieństwem występują amplitudy większe od mody.

2.155. Gęstość prawdopodobieństwa prędkości V cząsteczek gazu ma postać:

$$f(v) = \begin{cases} \dfrac{4h^3}{\sqrt{\pi}} v^2 \exp(-h^2v^2) & \text{dla } v>0, \\ 0 & \text{poza tym.} \end{cases}$$

(h — stała zależna od temperatury i masy cząsteczki gazu) (2.8.31). Obliczyć:

a) wartość przeciętną długości przebytej drogi przez cząsteczkę w jednostce czasu,

b) modę.

2.156. Dla czterech podanych jednomodalnych symetrycznych gęstości wartości przeciętne są równe 0 i odchylenia standardowe 1. Sprawdzić, że współczynniki skupienia K (koncentracji) i maksymalne wartości gęstości $f(x)$ mają wartości podane w tabelce:

Gęstość rozkładów $f(x)$	$K=\frac{u_4}{\sigma^4}$	$\max f(x)=f(m_0)=f(0)$
$\frac{1}{3\sqrt{\pi}}(\frac{9}{4}+x^4)\exp(-x^2)$	2,75	0,423
$\frac{3}{2\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)-\frac{1}{6\sqrt{\pi}}(\frac{9}{4}+x^4)\exp(-x^2)$	3,125	0,387
$\frac{1}{6\sqrt{\ }}[\exp(-\frac{1}{4}x^2)+4\exp(-x^2)]$	4,5	0,470
$\frac{3\sqrt{3}}{16\sqrt{\pi}}(2+x^2)\exp(-\frac{3}{4}x^2)$	2,667	0,366

Porównać podane wartości z odpowiednimi wartościami dla rozkładu normalnego: $K=3, f(m_0)=f(0)=0{,}399$ oraz stwierdzić, że zarówno w przypadku gdy $f(m_0)$ jest większe od 0,399 jak i gdy jest mniejsze, możliwa jest dodatnia jak i ujemna wartość ekscesu $E=$ $=\gamma_2=K-3$.

2.157. Zmienna losowa X ma rozkład beta o gęstości:

$$f(x)=\begin{cases}12x(1-x)^2 & \text{dla} \quad 0<x<1, \\ 0 & \text{poza tym.}\end{cases}$$

Wyznaczyć gęstość zmiennej losowej Y, jeśli: a) $Y=2X^3$, b) $Y=\arcsin X$, c) $Y=e^X$.

2.158. Zmienna losowa X ma jednoparametrowy rozkład gamma o gęstości

$$f(x)=\begin{cases}\dfrac{1}{\Gamma(p)}x^{p-1}e^{-x} & \text{dla} \quad x>0, \\ 0 & \text{poza tym.}\end{cases}$$

Wyznaczyć rozkład zmiennej losowej Y, jeśli: a) $Y=\dfrac{1}{X}$, b) $Y=\ln X$, c) $Y=X^2$.

2.159. Zmienna losowa X ma rozkład jednostajny w przedziale $(-1, 1)$, wyznaczyć gęstość zmiennej $Y=|X|$.

2.160. Jaka będzie gęstość rozkładu beta (2.8.19) rozszerzonego do dowolnego skończonego przedziału $\langle a, b\rangle$?

Wskazówka. Wyznaczyć funkcję liniową, która dla wartości a przyjmuje wartość 0, a dla b wartość 1 i skorzystać ze wzoru (2.5.2).

2.161. Punkt materialny M porusza się po prostej ruchem jednostajnym od punktu $O(0, 0)$ do punktu $A(2, 0)$. Wyznaczyć rozkład prawdopodobieństwa losowego czasu T trwania ruchu, zakładając, że równe odcinki drogi punkt M przebywa z takim samym prawdopodobieństwem oraz obliczyć ET.

2.162. Czas bezawaryjnej pracy X agregatów spalinowo-elektrycznych w ustalonych warunkach ma rozkład wykładniczy (2.8.7) z parametrem $\lambda=10$. Obliczyć średni czas bezawaryjnej pracy agregatów oraz wyznaczyć gęstość zmiennej losowej $Y=|X-EX|$.

2.163. Moc produkowanych tranzystorów ma rozkład $N(4, 1)$. Wyznaczyć gęstość prawdopodobieństwa pracy jaką wykona prąd elektryczny w tranzystorze w ciągu 2 godzin.

2.164. Dany jest sześcian, którego krawędź X ma losową długość z przedziału $\langle 1, 2\rangle$ o rozkładzie jednostajnym. Obliczyć:

a) Wartość przeciętną objętości i wartość przeciętną pola powierzchni całkowitej tego sześcianu.

b) Wyznaczyć rozkład prawdopodobieństwa objętości tego sześcianu.

2.165. Trwałości lamp mają rozkład Rayleigha o dystrybuancie

$$F(x)=\begin{cases}1-\exp(-\frac{1}{300}x^2) & \text{dla} \quad x>0,\\ 0 & \text{poza tym}.\end{cases}$$

a) Wyznaczyć gęstość zmiennej losowej: $Y=X^2$.

b) Obliczyć EY.

2.166. Wyznaczyć wartość modalną rozkładu logarytmiczno-normalnego o gęstości (2.5.4).

2.167. Zmienna losowa R o wartościach równych odległości punktu trafienia od środka tarczy strzeleckiej ma rozkład Rayleigha o gęstości

$$f(r)=\begin{cases}Are^{-h^2r^2} & \text{dla} \quad r>0,\\ 0 & \text{dla pozostałych } r.\end{cases}$$

Wyznaczyć stałą A, oraz obliczyć prawdopodobieństwo, że odległość punktu trafienia od środka tarczy jest mniejsza od mody zmiennej losowej R.

2.168. Badając zjawisko trzęsienia ziemi, zaobserwowano, że zachodzi związek: $Y=ce^X$, gdzie Y jest intensywnością drgań ziemi w pewnym miejscu, X – siłą trzęsienia ziemi, c – współczynnikiem zależnym od odległości danego miejsca od epicentrum trzęsienia i od doboru jednostek. Zakładając, że X jest zmienną losową o rozkładzie wykładniczym o gęstości

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

wyznaczyć dystrybuantę zmiennej losowej Y.

2.169. Dystrybuanta F pewnej zmiennej losowej X jest funkcją ciągłą dla $x \in \boldsymbol{R}$. Czy stąd wynika, że jej gęstość f jest również funkcją ciągłą? W przypadku odpowiedzi pozytywnej przeprowadzić dowód, w przypadku negatywnej podać kontrprzykład.

2.170. Gęstość prawdopodobieństwa f pewnej zmiennej losowej X jest ciągła dla $x \in \boldsymbol{R}$. Czy stąd wynika, że jej dystrybuanta F jest również ciągła? Polecenie jak w poprzednim zadaniu. Czy gęstość f może być nieograniczona w pewnym otoczeniu (jedno – czy też obustronnym) jednego albo więcej punktów?

2.171. Gęstość prawdopodobieństwa f pewnej zmiennej losowej X ma skończoną liczbę punktów nieciągłości. Czy wynika stąd, że: a) F może być funkcją ciągłą, b) F jest funkcją ciągłą, c) F nie jest funkcją ciągłą?

2.172. Niech F będzie dystrybuantą ciągłej zmiennej losowej X. Wyznaczyć rozkład zmiennej losowej $Y=F(X)$, przy założeniu, że istnieje funkcja odwrotna do $y=F(x)$.

2.173. Niech zmienna losowa Y ma rozkład wykładniczy (2.8.7). Dokonując zamiany zmiennych $Y=X^p$ dla $x>0$, $p>0$, wyznaczyć gęstość zmiennej losowej X.

2.174. Pewne urządzenie składa się z dwóch elementów pracujących niezależnie od siebie połączonych równolegle. Czas bezawaryjnej pracy (w godz.) każdego z nich jest zmienną losową o tym samym rozkładzie wykładniczym o gęstości:

$$f(x)=\begin{cases} 0{,}1e^{-0{,}1x} & \text{dla} \quad x\geqslant 0, \\ 0 & \text{poza tym}. \end{cases}$$

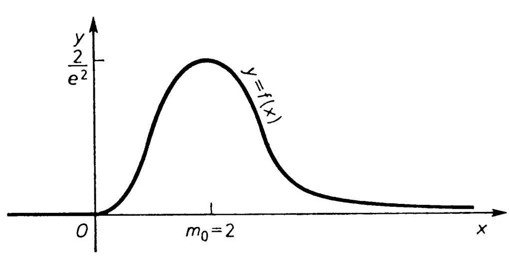
Obliczyć prawdopodobieństwo, że urządzenie będzie działało co najmniej przez 20 godzin.

2.175. Korzystając z zależności $D^2X=D^2(X-a)$, a — stała, obliczyć wariancję zmiennej losowej X o gęstości

$$f(x)=\begin{cases} \frac{1}{24}(x-1)^4\exp[-(x-1)] & \text{dla} \quad x>1, \\ 0 & \text{dla} \quad x\leqslant 1. \end{cases}$$

2.176. Naszkicować gęstość rozkładu gamma (rys. 2.41) w przypadku $\lambda=1$, $p=3$. Wyznaczyć: a) dystrybuantę, b) modę. c) Obliczyć $P(0<X<1)$.

Rys. 2.41. Do zadania 2.176



2.177. Naszkicować wykres gęstości rozkładu beta w przypadku $p=4$, $q=2$ oraz wyznaczyć modę, o ile istnieje.

Odpowiedzi

2.44. a) $c=0{,}3$;

c)

x	$(-\infty,-5\rangle$	$(-5,-2\rangle$	$(-2,0\rangle$	$(0,1\rangle$	$(1,3\rangle$	$(3,8\rangle$	$(8,\infty)$
$F(x)$	0	0,1	0,3	0,4	0,6	0,9	1

;

d₁) 0,2; d₂) 0; d₃) $F(3)=0{,}6$; d₄) $F(2)=0{,}6$; d₅) $1-F(0)=0{,}7$; d₆) 0,5.

2.45.

x_i	-2	1	3
p_i	0,2	0,6	0,2

.

2.46. a) $F(b+0)$; b) $1-F(b)$; c) $F(b+0)-F(a+0)$; d) $F(b+0)-F(a)$; e) $F(b)-F(a+0)$.

2.47. a) Bo np.: $X(\omega_1)\neq Y(\omega_1)$; b$_1$)

x_i	0	1	2
p_i	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

; b$_2$)

y_j	0	1	2
q_j	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

; c) nie, bo różne zmienne losowe mogą mieć jednakowe rozkłady.

2.48. Np. $\Omega=\langle 0,1\rangle$, $X(\omega)=\begin{cases}0, & \text{gdy} \quad 0\leqslant\omega\leqslant\frac{1}{2},\\ 1, & \text{gdy} \quad \frac{1}{2}<\omega\leqslant 1.\end{cases}$

2.49. a$_1$) Tak; bardzo często rozważa się takie zmienne losowe, np. zmienne losowe tworzące próbę prostą (Część II, p. 2.1); a$_2$) tak; b) wystarczy zauważyć, że $P(X=1, Y=1)=P(\{\omega_2\})=\frac{1}{3}\neq P(X=1)\cdot P(Y=1)=\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9}$.

2.50. a)

k	1	2	3
p_k	0,7	0,21	0,09

; b) 1,39; Wskazówka. Por. rozwiązanie zadania 2.29.

2.51. a)

k	0	1	2	3	4
p_k	0,2	0,16	0,128	0,1024	0,4096

;

b)

x	$(-\infty,0\rangle$	$(0,1\rangle$	$(1,2\rangle$	$(2,3\rangle$	$(3,4\rangle$	$(4,\infty)$
$F(x)$	0	0,2	0,36	0,488	0,5904	1

; c) 0,64.

2.52.

k	0	1	2	3
p_k	0,036	0,222	0,448	0,294

.

2.53. a)

k	1	2	3	4
p_k	0,06	$\approx$0,056	$\approx$0,053	$\approx$0,831

;

b)

x	$(-\infty,1\rangle$	$(1,2\rangle$	$(2,3\rangle$	$(3,4\rangle$	$(4,\infty)$
$F(x)$	0	0,06	0,116	0,169	1

c) $\approx$3,65.

2.54. a)

u_i	-3	1	9	13
q_i	0,1	0,2	0,5	0,2

; b)

u_i	-27	-1	27	125
q_i	0,1	0,2	0,5	0,2

;

c)

u_i	-4	4	20
q_i	0,2	0,6	0,2

.

2.55. a) $u=g(x)=2x+3$, $h(u)=\frac{1}{2}(u-3)$, $W_u=\{-3,1,9,13\}$; b) $u=g(x)\equiv x^3$, $h(u)=\sqrt[3]{u}$, $W_u=\{-27,-1,27,125\}$.

2.56. $q_i\equiv P(U=u_i)=P(X=\frac{1}{3}u_i)=e^{-\lambda}\dfrac{\lambda^{u_i/3}}{(\frac{1}{3}u_i)!}$, $u_i\in W_u=\{0,3,6,9,\ldots\}$.

2.57. $q_i\equiv P(U=u_i)=P(X=\sqrt{u_i})=\binom{4}{\sqrt{u_i}}(\frac{1}{4})^{\sqrt{u_i}}(\frac{3}{4})^{4-\sqrt{u_i}}$, $u_i\in W_u=\{0,1,4,9,16\}$.

2.58. a$_1$)

u_i	3	4	5	6	7	8
q_i	0,09	0,27	0,29	0,21	0,08	0,06

, np. $P(U=5)=P((X=1\wedge Y=4)\vee(X=2\wedge Y=3))=P(X=1)P(Y=4)+P(X=2)P(Y=3)=0{,}3\cdot 0{,}3+0{,}5\cdot 0{,}4=0{,}29$;

a$_2$)

u_i	-3	-2	-1	0	1	2	3	4	5	6
q_i	0,04	0,12	0,04	0,1	0,34	0,22	0,04	0,02	0,06	0,02

;

b$_1$)

u_i	2	3	4	6	8	9	12
q_i	0,09	0,12	0,24	0,26	0,15	0,08	0,06

;

b$_2$)

u_i	-5	-2	-1	0	1	2	5
q_i	0,02	0,08	0,1	0,6	0,1	0,08	0,02

.

2.59. a)

u_i	-1	1
q_i	$\frac{3}{5}$	$\frac{2}{5}$

Wskazówka. $\cos \pi K=(-1)^K$, $q_1 \equiv P(U=-1)=P(K=2n-1, n \in N)=\sum_{n=1}^{\infty} \frac{1}{3}(\frac{2}{3})^{2n-1-1}=\frac{3}{5}$;

b)

u_i	-1	0	1
q_i	$\frac{18}{65}$	$\frac{3}{5}$	$\frac{8}{65}$

Wskazówka. $\cos \frac{1}{2}\pi K=\begin{cases} -1 & \text{gdy } K=4n-2, \quad n \in N, \\ 0, & \text{gdy } K=2n-1, \quad n \in N, \\ 1, & \text{gdy } K=4n, \quad n \in N, \end{cases}$

i np. $P(U=-1)=P(K=4n-2, n \in N)=\sum_{n=1}^{\infty} \frac{1}{3}(\frac{2}{3})^{4n-3}=\frac{1}{3} \cdot \frac{27}{8} \sum_{n=1}^{\infty} (\frac{16}{81})^n=\frac{18}{65}$;

c)

u_i	0	1
q_i	$\frac{3}{5}$	$\frac{2}{5}$

Wskazówka. $\cos^2 \frac{1}{2}\pi K \equiv \frac{1}{2}(1+\cos \pi K)=\begin{cases} 0, & \text{gdy } K=2n-1, \\ 1, & \text{gdy } K=2n, \end{cases} \quad n \in N,$
i np. $P(\cos^2 \frac{1}{2}\pi K=0)=P(\cos K=-1)=\ldots=\frac{3}{5}$; można również wykorzystać wynik z punktu b);

d)

u_i	-1	1
q_i	$\frac{6}{13}$	$\frac{7}{13}$

Wskazówka. $\sin \frac{1}{2}\pi K+\cos \frac{1}{2}\pi K=\begin{cases} -1, & \text{gdy } K=4n-1 \text{ albo } K=4n-2, \quad n \in N, \\ 1, & \text{gdy } K=4n \text{ albo } K=4n-3, \quad n \in N, \end{cases}$

i np. $P(U=-1)=\frac{1}{3} \cdot \frac{9}{4} \sum_{n=1}^{\infty} (\frac{16}{81})^n+\frac{1}{3} \cdot \frac{27}{8} \sum_{n=1}^{\infty} (\frac{16}{81})^n=\ldots=\frac{6}{13}$.

2.60. $g(x)=\dfrac{x-x_1}{x_2-x_1}$. **2.61.** a)

v_i	0	1	2	3	4	5
α_i	0,03	0,16	0,31	0,31	0,16	0,03

;

b)

u_i	0	1	2	3	4	6
q_i	0,37	0,16	0,28	0,04	0,12	0,03

.

2.63. a) $\alpha_1=2$, b) $x_{0,5}=3$; c) $x_{0,3}=$dowolna liczba z przedziału $\langle -1, 3 \rangle$; d) $\sigma^2=6{,}6$; e) $\sigma=2{,}57$; f) $d=2{,}2$; g) $v \approx 1{,}28$; h) $h=1{,}2$; i) $\alpha_2=10{,}6$; j) $\alpha_3=35{,}6$; k) $\mu_3=-12$; l) $\gamma_1 \approx -0{,}71$.

2.66. (2.6.23)-(2.6.25).

2.68. $EX=EY=1$, $D^2X=D^2Y=\frac{2}{3}$, ... jednakowe wartości przeciętne i jednakowe wariancje.

2.69. a$_1$)

u_i	0	1	2	3	4
q_i	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

; a$_2$)

u_i	0	2	4
q_i	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

;

a$_3$)

u_i	0	1	2	4
q_i	$\frac{5}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

; a$_4$)

u_i	0	1	4
q_i	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

;

zatem zmienne losowe $X+Y$ oraz $2X$ są różne, różne są również ich funkcje prawdopodobieństwa, podobna uwaga dotyczy zmiennych losowych XY oraz X^2;

b) $EU_1=2$; $EU_2=2$; $EU_3=1$; $EU_4=\frac{5}{3} \approx 1{,}67$; zatem: $E(X+Y)=E(2X)$, ale $EXY \neq EX^2$;

c) $D^2U_1=\frac{4}{3}$; $D^2U_2=\frac{8}{3}$; $D^2U_3=\frac{16}{9}$; $D^2U_4=\frac{20}{9}$, zatem wszystkie wariancje są różne.

2.70. a) $EU_1=5$, $D^2U_1=4$;
b) $EU_2=EX^2=D^2X+(EX)^2=5$, $D^2U_2=D^2X^2=EX^4-(EX^2)^2=34-25=9$;
c) $EU_3=-3$, $D^2U_3=9$.

2.71.

u_i	-7	-5	1	7
q_i	0,3	0,1	0,2	0,4

, $EU=\sum_{i=1}^{4} u_i q_i=\ldots=0{,}4$ albo $EU=2EX-3=2\cdot 1{,}7-3=0{,}4$; podobnie: $D^2U=\sum_{i=1}^{4} u_i^2 q_i-(\sum_{i=1}^{4} u_i q_i)^2=37-0{,}16=36{,}84$ albo $D^2U=4$, $D^2X=4\cdot 9{,}21=36{,}84$.

2.72. a) $EU_1=-17$, $D^2U_1=12{,}5$; b) $EU_2=-8{,}5$. Wskazówka. $EX^2=D^2X+(EX)^2$.

2.73.

x_i	-1	4
p_i	0,2	0,8

.

2.74.

y_i	1	4
q_i	0,5	0,5

, rozkład $U=X+Y$:

u_i	-1	0	2	3	5	6
q_i	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

, $EX=0$, $D^2X=2{,}5$, $EY=2{,}5$, $D^2Y=2{,}25$, $D^2U=4{,}75$.

2.76. $D^2(XY)=E[XY-E(XY)]^2=E(XY-EXEY)^2=E(XY)^2=EX^2EY^2=D^2XD^2Y$.

2.77. Wynika to z własności wariancji wyrażonej równością (2.6.28).

2.79. b) Niech $x'=\min x_i$, $x''=\max x_i$, $R=x''-x'$, $D^2X\leqslant E\left(X-\frac{x'+x''}{2}\right)^2=E(\frac{1}{4}R^2-(x''-X)(X-x'))\leqslant\frac{1}{4}R^2$; c$_2$) Wskazówka. $\max p(1-p)=\frac{1}{4}$.

2.81. b) Wskazówka: Niech $U_i=X_i:(X_1+X_2+\ldots+X_n)$, wtedy $U_1+\ldots+U_n=1$ oraz $EU_1=\ldots=EU_n$.

2.82. $EX\approx 2{,}36$, $D^2X\approx 2{,}58$. **2.83.** $EX\approx 2{,}15$.

2.84. $EW=-mq+q\sum_{k=2}^{\infty} a^{k-1}p^{k-1}=q\left(-m+\frac{ap}{1-ap}\right)$, gdy $0<ap<1$;

a) $EW=0\Leftrightarrow m=\frac{ap}{1-ap}$; b) $a=\frac{m}{(m+1)p}$; b$_1$) $a=1$, b$_2$) $a=1$; b$_3$) $a=3$; b$_4$) $a=4$; b$_5$) $a=5$; c$_1$) $a<\frac{m}{(m+1)p}$; c$_2$) $a>\frac{m}{(m+1)p}$.

2.85. a) $F(0)=\frac{1}{3}$, $F(3)=\frac{1}{2}$, $D^2X=2$, $EX=-5$; b) $F(4)=\frac{2}{3}$, $F(5)=\frac{20}{21}$, $EX=2$ $x_{0,5}=1$. Wskazówka. $p(4)=P(X=4)=F(4+0)-F(4)=F(5)-F(4)=\frac{2}{7}$, skąd $4p(4)=\frac{8}{7}>1$, zatem $EX>1$.

2.87.

k	0	1	2	3	4	5
$p_k=P(k;5,\frac{3}{4})$	0,001	0,015	0,088	0,264	0,396	0,237

, $k_0=4$, $[(n+1)p]=[4{,}5]=4$.

2.88. $n=4\mu_2^3:(\mu_2^2-\mu_3^2)$, $p=\frac{1}{2}\left(1-\frac{\mu_3}{\mu_2}\right)$.

2.89. $1-(1-p)^3$.

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

$$=\sum_{i=0}^{k}\binom{n_1}{i}\binom{n_2}{k-i}p^kq^{n_1+n_2-k}=p^kq^{n_1+n_2-k}\sum_{i=0}^{k}\binom{n_1}{i}\binom{n_2}{k-i}=\binom{n_1+n_2}{k}p^kq^{n_1+n_2-k}.$$

2.91. a) $p=\dfrac{n}{n+1}$; b) $p=\dfrac{1}{n+1}$; c) $\dfrac{n}{n+1}<p<1$; d) $0<p<\dfrac{1}{n+1}$.

2.92. $P(4;7,\frac{1}{4})\approx 0{,}058$. **2.93.** $\approx 0{,}71$.

2.94. a) $\binom{m}{k}p^k(1-p)^{m-k}$; b) $\binom{mv}{x}p^x(1-p)^{mv-x}$; c$_1$) $EK=mp=\dfrac{mp_0}{c^v}$;

c$_2$) $EX=mvp=\dfrac{mvp_0}{c^v}$; d) $\max\alpha_1(v)=\max\dfrac{mvp_0}{c^v}=\alpha_1\left(\dfrac{1}{\ln c}\right)=\dfrac{mp_0}{e\cdot\ln c}\approx 79$ obr./min.

Wskazówka. Posłużyć się rachunkiem różniczkowym, traktując v jako zmienną mogącą przyjmować dowolne dodatnie wartości rzeczywiste.

2.95. a) $p_k=P(T=k)=\binom{k-1}{3}0{,}005^4\cdot 0{,}995^{k-4}$, $k=4,5,\ldots$ Wskazówka. Zastosować (2.7.14); b) $F(x)=\sum_{k<x}p_k$; c$_1$) $P(T=3)=0$; c$_2$) $P(T\geqslant 3)=1$; c$_3$) $P(T=24)\approx 10^{-5}$; c$_4$) $P(16\leqslant T\leqslant 24)=p_{16}+p_{17}+\ldots+p_{24}$; d) $ET=800$ h ((2.7.15)).

2.96. a) $p_k=P(T=k)=0{,}005\cdot 0{,}995^{k-1}, k\in N$; b) $F(x)=\begin{cases}0 & \text{dla } x<1,\\ 1-0{,}995^{x-1} & \text{dla } x\in N,\\ 1-0{,}995^{[x]} & \text{dla } x\in R_+\setminus N;\end{cases}$

c$_1$) $P(T=3)\approx 0{,}005$; c$_2$) $P(T\geqslant 3)=1-F(3)\approx 0{,}990$; c$_3$) $P(T=24)\approx 0{,}004$; c$_4$) $P(16\leqslant T\leqslant 24)=F(24+0)-F(16)\approx 0{,}036$; d) $ET=200$ h.

2.97. $P(T=k)=\dfrac{0{,}999}{1}\cdot\dfrac{0{,}999}{2}\cdots\dfrac{0{,}999}{k-1}\left(1-\dfrac{0{,}999}{k}\right)=\dfrac{0{,}999^{k-1}}{(k-1)!}\left(1-\dfrac{0{,}999}{k}\right)$.

2.98. $P(7;4,\frac{1}{2})=\frac{6}{3}\cdot\frac{1}{2}\cdot\frac{5}{32}=\frac{5}{32}$.

2.99.

k	0	1	2	3	4
p_k	0,024	0,238	0,476	0,238	0,024

. **2.100.** a) 0,824; b) 0,921.

2.101. $\max(0, n+M-N)\leqslant k\leqslant\min(n, M)$. **2.102.** $\approx 0{,}238$.

2.103. a) $\approx P(0;6)=e^{-6}\approx 0{,}0025$; b) $\approx 6e^{-6}\approx 0{,}0149$; c) $\approx 18e^{-6}\approx 0{,}0446$; d) $\approx 1-(0{,}0025+0{,}0149+0{,}0446)=0{,}438$.

2.104. a) $\approx e^{-3}\approx 0{,}0498$; b) $\approx 3e^{-3}\approx 0{,}1494$; c) $\approx 4{,}5e^{-3}\approx 0{,}2240$; d) $\approx 0{,}3528$.

2.106. a) $-rn\leqslant\min(M, N-M)$; b) $P(K=k)=$

$$=\binom{n}{k}\frac{M(M+r)(M+2r)\ldots(M+(k-1)r)(N-M)(N-M+r)(N-M+(n-k-1)r)}{N(N+r)(N+2r)\ldots(N+(n-1)r)};$$

c) $EK=np$, gdzie $p=M/N$; d) dla $r=0$ otrzymuje się rozkład dwumianowy z parametrami (n,p), gdzie $p=M/N$, przy $r=-1$ mamy rozkład hipergeometryczny z parametrami (N, M, n).

2.107. $F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 0, \\ 1-e^{-x}(1+x) & \text{dla} \quad x>0, \end{cases} \qquad P(X\geqslant 2)=\frac{3}{e^2}.$

2.108. $c=\frac{1}{2}$; a) $F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 0, \\ \frac{1}{2}(1-\cos x) & \text{dla} \quad 0<x\leqslant\pi, \\ 1 & \text{dla} \quad x>\pi; \end{cases}$

b) $P(|X|<\frac{1}{3}\pi)=\frac{1}{4}$.

2.109. b) $F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 0, \\ \frac{1}{2}x^2 & \text{dla} \quad 0<x\leqslant 1, \\ 2x-\frac{1}{2}x^2-1 & \text{dla} \quad 1<x\leqslant 2, \\ 1 & \text{dla} \quad x>2. \end{cases}$

c) $P(A)=1-\binom{2}{0}(\frac{1}{2})^0(\frac{1}{2})^2=\frac{3}{4}$, gdyż $P(1\leqslant X<2)=\int_1^2(2-x)\,dx=\frac{1}{2}$.

2.110. $f(t)=\begin{cases} \frac{1}{10} & \text{dla} \quad 0\leqslant t\leqslant 10, \\ 0 & \text{poza tym}, \end{cases} \qquad P(T\geqslant 4)=0{,}6.$

2.111. $F(t)=\begin{cases} \frac{1}{2}e^t & \text{dla} \quad t\leqslant 0, \\ 1-\frac{1}{2}e^{-t} & \text{dla} \quad t>0, \end{cases} \qquad P(|T|<2)=1-e^{-2}.$

2.112. $A=1, \quad B=-\frac{1}{\pi}, \quad f(x)=\begin{cases} \dfrac{1}{\pi\sqrt{1-x^2}} & \text{dla} \quad |x|<1, \\ 0 & \text{poza tym}. \end{cases}$

2.113. $a=2$, $P(-1\leqslant X\leqslant 1{,}5)=\frac{2}{3}$. **2.114.** $f(x)=\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)$.

2.115. a) Nie, ponieważ funkcja $\cos x$ jest w pierwszej ćwiartce malejąca. b) Tak przy $c=1$.

2.116. $f(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant -1, \\ \frac{1}{2}(x+1) & \text{dla} \quad -1<x\leqslant 0, \\ -\frac{1}{6}(x-3) & \text{dla} \quad 0<x\leqslant 3, \\ 0 & \text{dla} \quad x>3, \end{cases}$

$$F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant -1, \\ \frac{1}{2}(\frac{1}{2}x^2+x+\frac{1}{2}) & \text{dla} \quad -1<x\leqslant 0, \\ -\frac{1}{6}(\frac{1}{2}x^2-3x)+\frac{1}{4} & \text{dla} \quad 0<x\leqslant 3, \\ 1 & \text{dla} \quad x>3. \end{cases}$$

2.117. a) $EX=\frac{3}{4}$; b) $x_{0,5}=\frac{1}{\sqrt[3]{2}}$; c) $\delta_2=\frac{3}{4}\left(1-\frac{1}{\sqrt[3]{2}}\right)$.

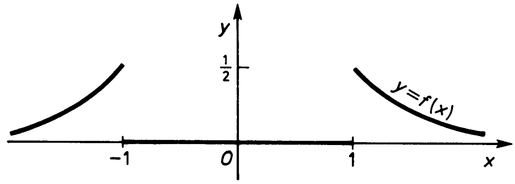
2.118. $EX=m_0=x_{0,5}=0$, $\mu_3=0$. **2.119.** $m_0=0$, $x_{0,5}=0$.

2.120. $m_0=x_0+\left[\frac{1}{k(2a-1)}\right]^{1/a}, \quad x_{0,5}=x_0+\left(\frac{1}{k}\right)^{1/a}.$

2.121. $m_0'=\frac{1}{2}\pi$, $m_0''=\frac{3}{2}\pi$, $x_{0,5}=\pi$, $EX=\pi$.

2.122. $m_0=\frac{1}{n+2}, \quad \alpha_3=\frac{1}{n(n-1)(n-2)}.$

Rys. 2.42. Odpowiedź do zadania 2.123



2.123. EX, D^2X nie istnieją, $x_{0,5} \in \langle -1, 1\rangle$ (rys. 2.42).

2.124. $EV=0{,}022\pi$. **2.125.** $EX=2{,}5$, $D^2X=\frac{25}{12}$.

2.126. $\gamma_1=0$, $\gamma_2=2{,}16$. **2.127.** $\gamma_1=\frac{1}{4}$, $\gamma_2=\frac{5}{8}-3=-2{,}375$.

2.128. a) $\gamma_1=2{,}8$; b) $\gamma_1=-0{,}21$.
asymetria prawa.w a), asymetria lewa w b).

2.129. a) $EX_1=1$, $D^2X_1=\frac{1}{5}$, $\gamma_2=\frac{15}{7}$;
b) $EX_2=1$, $D^2X_2=\frac{1}{5}$, $\gamma_2=6$.

2.130. Odchylenie standardowe $\sigma=\frac{1}{3}\sqrt{3}$, $\delta_1=0{,}5$.

2.131. $\mu_3=\alpha_3-3\alpha_1\alpha_2+2\alpha_1^3$. **2.132.** $\alpha_3=\mu_3+3\mu_1\mu_2+\mu_1^3$.

2.134. Przyjąć $a=1$, to $EX=7\int\limits_0^1(x-1)^7dx+1=\frac{1}{8}$.

2.135. $f(x)=\begin{cases}\frac{1}{20} & \text{dla} \quad -5\leqslant x\leqslant 15,\\ 0 & \text{poza tym,}\end{cases}$ $\quad EX=5, \quad K=1{,}8.$

2.136. $f(x)=\begin{cases}\frac{2}{5} & \text{dla} \quad 0\leqslant x\leqslant 2{,}5,\\ 0 & \text{poza tym,}\end{cases}$ $\quad P(X<0{,}5)=\int\limits_0^{0,5}\frac{2}{5}dx=0{,}2.$

2.137. $K=9$. **2.138.** $P(|X|>0{,}1)=1-\frac{2}{5}=\frac{3}{5}$.

2.140. $P(X<2{,}5)=\Phi(0{,}5)=0{,}6915$, $P(X>-0{,}5)=\Phi(1)=0{,}8413$,

$$P(0{,}5<X<2)=\Phi(0{,}25)+\Phi(0{,}5)-1=0{,}2902,$$
$$P(|2X-1|<1)=\Phi(0{,}75)-\Phi(0{,}25)=0{,}1747,$$
$$P(|X|>0{,}5)=1+\Phi(0{,}5)-\Phi(1)=0{,}8502.$$

2.141. $$E(|X|^\alpha)=\frac{2}{\sqrt{2\pi}\sigma}\int\limits_0^\infty x\exp\left(-\frac{x^2}{2\sigma^2}\right)dx=\frac{\sigma^\alpha(\sqrt{2})^\alpha}{\sqrt{\pi}}\int\limits_0^\infty u^{(\alpha+1)/2-1}e^{-u}du=$$
$$=\frac{(\sigma\sqrt{2})^\alpha}{\sqrt{\pi}}\Gamma\left(\frac{\alpha+1}{2}\right).$$

2.142. $1-P(1{,}7<X<2{,}3)=0{,}1336$.

2.143. $EX=10$, $D^2X=25$, $\mu_4=3(D^2X)^2=1875$.

2.144. a) $EX=1$; b) $P(|X+1|<2)=0{,}9773$.

2.145. $P(0<X<900)=0{,}0227$, tzn. przeciętnie 2,27% lin ma wytrzymałość mniejszą od 900 kg/cm^2.

2.146. $P(|X-2|<k\cdot 0{,}01)=0{,}95 \Rightarrow k=1{,}95$, skąd $1{,}98<X<2{,}02$.

2.147. Oznaczmy $Y=X-\mu$, Y ma rozkład $N(0,\sigma)$, należy wykazać, że $E|Y|=\sqrt{\frac{2}{\pi}}\sigma$.

$$E|Y|=\frac{1}{\sigma\sqrt{2\pi}}\left[-\int_{-\infty}^{0} y\exp\left(-\frac{y^2}{2\sigma^2}\right)dy+\int_{0}^{\infty} y\exp\left(-\frac{y^2}{2\sigma^2}\right)dy\right]=$$

$$=\frac{2}{\sigma\sqrt{2\pi}}\int_{0}^{\infty} y\exp\left(-\frac{y^2}{2\sigma^2}\right)dy=\sigma\sqrt{\frac{2}{\pi}}.$$

2.148. a) $EX=5$; b) $x_{0,5}=5\ln 2$ (rozw. równania $1-\exp(-\frac{1}{5}x)=\frac{1}{2}$);
c) $P(X\geqslant 5)=1-e^{-1}$.

2.149. a) $P(1<T<2)=\exp(-\frac{1}{3})-\exp(-\frac{2}{3})$;
b) $f(t)=\begin{cases}\frac{1}{3}\exp(-\frac{1}{3}t) & \text{dla} \quad t\geqslant 0,\\ 0 & \text{poza tym;}\end{cases}$
c) $ET=3$, $D^2T=9$.

2.150. $c=8$, $EX=\frac{1}{4}\sqrt{\pi}$, $D^2X=\frac{1}{16}(4-\pi)$.

2.151. $\gamma_1=\frac{2}{\sqrt{n}}$ $(\mu_3=2\lambda^3 n)$.

2.152. $m_0=1$, $EX=\frac{2}{\sqrt{\pi}}$.

2.153: $f'(x)=\dfrac{x^{p-2}\exp\left(-\frac{x}{\lambda}\right)}{\lambda^p\Gamma(p)}\left(p-1-\frac{x}{\lambda}\right)$ dla $x>0$,

$$f'(x)=0\Leftrightarrow x=\lambda(p-1)>0 \quad \text{dla} \quad p>1,$$

$$f'(x)>0 \quad \text{dla} \quad x<\lambda(p-1), \quad f'(x)<0 \quad \text{dla} \quad x>\lambda(p-1),$$

więc $f(x)$ dla $p>1$ ma jedno maksimum lokalne w punkcie $x=\lambda(p-1)$, a więc jest to rozkład jednomodalny, przy czym $m_0=\lambda(p-1)$.

2.154. a) $EX=\sqrt{\frac{1}{2}\pi}$, $x_{0,5}=\sqrt{2\ln 2}$, $m_0=1$; b) $P(X>1)=e^{-1/2}$.

2.155. a) $EV=\frac{2}{h\sqrt{\pi}}$; b) $m_0=\frac{1}{h}$.

2.157. a) $g(y)=2\sqrt[3]{2}[1-(\frac{1}{2}y)^{1/3}]^2 y^{-1/3}$ dla $0<y<2$;
b) $g(y)=12\sin y\cos y(1-\sin y)^2$ dla $0<y<\frac{1}{2}\pi$;
c) $g(y)=\frac{12}{y}\ln y(1-\ln y)^2$ dla $0<y<e$.

2.158. a) $g(y)=\frac{1}{\Gamma(p)}y^{-p-1}\exp\left(-\frac{1}{y}\right)$ dla $y>0$;

b) $g(y)=\dfrac{1}{\Gamma(p)}\exp[-e^y+y(p-1)]$ dla $y\in \boldsymbol{R}$;

c) $g(y)=\dfrac{1}{\Gamma(p)}y^{(p-1)/2}e^{-\sqrt{y}}$ dla $y>0$.

2.159. Należy skorzystać z (2.5.3) lub zastosować metodę podaną w zad. **2.9.**

$$g(y)=\begin{cases}1 & \text{dla } 0<y<1,\\ 0 & \text{poza tym}.\end{cases}$$

2.160. Dokonujemy następującej zamiany zmiennych:

$$Y=a+(b-a)X \Leftrightarrow X=\frac{Y-a}{b-a},$$

stosując wzór (2.5.2) otrzymamy gęstość:

$$g(y)=\begin{cases}\dfrac{1}{B(p,q)(b-a)^{p+q-1}}(y-a)^{p-1}(b-y)^{q-1} & \text{dla } a\leqslant y\leqslant b,\\ 0 & \text{poza tym},\end{cases}$$

dla p, $q>0$.

2.161. Niech zmienna losowa X określa położenie punktu M na prostej, jej gęstość **jest** określona wzorem:

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

Czas trwania ruchu $T=1/v\cdot X$, gdzie v jest stałą prędkością ruchu. Gęstość zmiennej losowej T:

$$g(t)=\begin{cases}\dfrac{v}{2} & \text{dla } 0<t<\dfrac{2}{v},\\ 0 & \text{poza tym}.\end{cases}$$

2.162. $EX=10$.

$$g(y)=\begin{cases}\dfrac{1}{10}\exp\left(-\dfrac{y+10}{10}\right)+\dfrac{1}{10}\exp\left(\dfrac{y-10}{10}\right) & \text{dla } 0<y\leqslant 10,\\ \dfrac{1}{10}\exp\left(-\dfrac{y+10}{10}\right) & \text{dla } y>10.\end{cases}$$

2.163. Oznaczmy przez M — moc, L — pracę, t — czas. Wiadomo, że $M=\dfrac{L}{t}$, **u nas** $M=\frac{1}{2}L$: ponieważ gęstość mocy **prądu** jest: $f(m)=\dfrac{1}{\sqrt{2\pi}}\exp\left[-\dfrac{(m-4)^2}{2}\right]$, więc dokonując podstawienia $m=\frac{1}{2}l$, otrzymujemy gęstość pracy prądu postaci: $g(l)=\dfrac{1}{2\sqrt{2\pi}}\exp\left[-\dfrac{(l-8)^2}{8}\right]$, a więc gęstość rozkładu $N(8;2)$.

2.164. a) $EV=E(X^3)=\frac{15}{4}$, $\quad ES=E(6X^2)=14$,

b) $g(v)=\begin{cases} \frac{1}{3}v^{-2/3} & \text{dla} \quad 1<v<8, \\ 0 & \text{poza tym}. \end{cases}$

2.165. a) Gęstość zmiennej losowej X jest postaci

$$f(x)=\begin{cases} \frac{1}{150}x\exp(-\frac{1}{300}x^2) & \text{dla} \quad x>0, \\ 0 & \text{poza tym}. \end{cases}$$

Dokonując podstawienia $x=\sqrt{y}$, otrzymujemy:

$$g(y)=\begin{cases} \frac{1}{300}\exp(-\frac{1}{300}y) & \text{dla} \quad y>0, \\ 0 & \text{poza tym}; \end{cases}$$

b) $EY=300$.

2.166. $f'(x)=\dfrac{1}{\sigma\sqrt{2\pi x^2}}\exp\left[-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\left(1+\dfrac{\ln x-\mu}{\sigma^2}\right)\right]$,

$$m_0=\exp(\mu-\sigma^2).$$

2.167. $A=2h^2$, $\quad m_0=\dfrac{1}{h\sqrt{2}}$, $\quad P(R<m_0)=0{,}393$.

2.168. $F(y)=P(Y<y)=\begin{cases} 1-\left(\dfrac{y}{c}\right)^{-\lambda} & \text{dla} \quad y\geqslant c, \\ & \text{dla pozostałych } y. \end{cases}$

2.169. Odpowiedź negatywna: kontrprzykładem jest rozkład równomierny na przedziale (a, b) (rys. 5.23).

2.170. Odpowiedź pozytywna: ponieważ $F(x)=\int\limits_{-\infty}^{x} f(u)du$, więc z ciągłości funkcji podcałkowej wynika, że F jest nie tylko funkcją ciągłą ale i różniczkowalną i zachodzi $F'(x)=f(x)$ ((2.3.2)).

Przykładem gęstości f nieograniczonej w otoczeniu jednostronnym dwóch punktów jest gęstość arc sinusa (zad. 2.13). Przykładem gęstości nieograniczonej w otoczeniu obustronnym przeliczalnej liczby punktów $a_i<a_{i+1}$, $i=0, \pm 1, \dots, \pm n, \dots$ jest

$$f(x)=\sum_{\substack{i=-\infty \\ i\neq 0}}^{\infty} \frac{1}{2^{|i|+2}\pi\sqrt{(x-a_i)(a_{i+1}-x)}}+\frac{1}{2\pi\sqrt{(x-a_0)(a_1-x)}}.$$

2.171. Zachodzi b) ponieważ dystrybuanta F jako funkcja górnej granicy całkowania w tym przypadku jest funkcją ciągłą.

2.172. Aby wyznaczyć dystrybuantę G zmiennej losowej $Y=F(X)$, rozważmy ją kolejno w trzech wzajemnie wykluczających się przedziałach:

1) dla $y\leqslant 0$ jest oczywiście $G(y)=0$, 2) dla $0<y\leqslant 1$ $G(y)=P(Y<y)=P[F(X)<y]$.

F jako dystrybuanta zmiennej losowej ciągłej jest funkcją ciągłą i wobec przyjętego założenia rosnącą, a więc istnieje do niej funkcja odwrotna $x=F^{-1}(y)$ także ciągła i rosnąca w $\langle 0, 1\rangle$, skąd

$$G(y)=P\,[X<F^{-1}(y)]=F\,[F^{-1}(y)]=y,$$

3) dla $y>1$ $G(y)=1$.

Otrzymaliśmy więc dystrybuantę rozkładu równomiernego w przedziale $\langle 0, 1\rangle$

$$G(y)=\begin{cases} 0 & \text{dla} \quad y\leqslant 0, \\ y & \text{dla} \quad 0<y\leqslant 1, \\ 1 & \text{dla} \quad y>1. \end{cases}$$

2.173. $k(x)=\begin{cases} \dfrac{1}{\lambda}px^{p-1}\exp\left(-\dfrac{x^p}{\lambda}\right) & \text{dla} \quad x>0, \\ 0 & \text{dla pozostałych } x. \end{cases}$

Otrzymaliśmy więc gęstość rozkładu Weibulla (2.8.33).

2.174. $P(Y\geqslant 20)=1-P(X_1<20)P(X_2<20)=1-(1-e^{-2})^2$.

2.175. $EY=\int\limits_{-\infty}^{\infty}(x+b)f(x)\,dx=\int\limits_{-\infty}^{\infty}xf(x)\,dx+b\int\limits_{-\infty}^{\infty}f(x)\,dx=EX+b$. Wobec istnienia EX wartość przeciętna zmiennej $Y=X+b$ istnieje i jest równa $EX+b$.

2.176. Modą zgodnie ze wzorem (2.8.14) jest: $m_0=2$. Wyznaczmy dystrybuantę:

$$F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 0, \\ \frac{1}{2}\int\limits_0^x t^2e^{-t}dt=1-\frac{1}{2}e^{-x}(x^2+2x+2) & \text{dla} \quad x>0. \end{cases}$$

Stąd łatwo obliczyć:

$$P(0<X<1)=F(1)-F(0)=1-\tfrac{5}{2}e^{-1}.$$

2.177. (Rys. 2.29) $m_0=\frac{3}{4}$.

3 ROZKŁADY UCIĘTE I ROZKŁADY MIESZANE. MIESZANINA ROZKŁADÓW

3.1. ROZKŁADY UCIĘTE

W praktyce dość często zdarza się, że jako model badanej cechy przyjmujemy zmienną losową X o wartościach z przedziału (a, b) skończonego albo nie. Z drugiej strony, wiadomo jednak, że badana cecha przyjmuje tylko wartości z przedziału skończonego $\langle\alpha, \beta\rangle \subset (a, b)$. Możliwe są wówczas dwie drogi postępowania:

1) Pominąć zbiór $(a, \alpha) \cup (\beta, b)$, jeśli prawdopodobieństwa $P(a<X<\alpha)$ i $P(\beta<X<b)$ są małe. Tak np. często korzysta się w praktyce z rozkładu $N(\mu, \sigma)$, mimo że rozkład zmiennej losowej X nie jest skoncentrowany na przedziale $(-\infty, +\infty)$, wiadomo jednak, że

$$P(\mu-3\sigma<X<\mu+3\sigma)\approx 0{,}9973,$$

tak więc pomijając prawdopodobieństwa $P(-\infty<X<\mu-3\sigma)$ oraz $P(\mu+3\sigma<X<+\infty)$ i tym samym ograniczając zakres zmienności X z przedziału nieskończonego do przedziału skończonego $\langle\mu-3\sigma, \mu+3\sigma\rangle$, obliczamy prawdopodobieństwo z błędem około 0,0027, co w większości sytuacji praktycznych jest dopuszczalne.

2) Skorzystać z tzw. rozkładów uciętych: *rozkładem uciętym dwustronnie*, jeśli $\alpha>a$ oraz $\beta<b$, nazywamy rozkład zmiennej losowej

$$Y=X\,|\,\alpha\leqslant X\leqslant\beta \tag{3.1.1}$$

jeżeli zaś $\alpha=a$, albo $\beta=b$, mamy wówczas do czynienia z *rozkładem uciętym prawostronnie* bądź *lewostronnie*.

Niech X będzie zmienną losową o dystrybuancie F, Y zaś uciętą zmienną losową (3.1.1) o dystrybuancie G. Korzystając ze wzorów na prawdopodobieństwo warunkowe, otrzymujemy

$$G(y)=P(Y<y)=P(X<y\,|\,\alpha\leqslant X\leqslant\beta)=$$

$$=\frac{P(\alpha\leqslant X<y)}{P(\alpha\leqslant X\leqslant\beta)}=\frac{F(y)-F(\alpha)}{F(\beta+0)-F(\alpha)} \quad \text{dla} \quad \alpha\leqslant y\leqslant\beta. \tag{3.1.2}$$

W szczególnym przypadku, gdy rozpatrzymy rozkłady ucięte jednostronnie, otrzymamy:

1. dla $\alpha=a$, $F(\alpha)=0$, skąd i ze wzoru (3.1.2)

$$G(y)=P(X<y\,|\,X\leqslant\beta)=\frac{F(y)}{F(\beta+0)} \quad \text{dla} \quad y\leqslant\beta, \tag{3.1.3}$$

2. dla $\beta=b$, $F(\beta+0)=1$

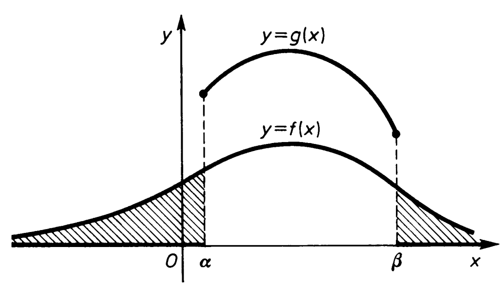
$$G(y)=P(X<y\,|\,X\geqslant\alpha)=\frac{F(y)-F(\alpha)}{1-F(\alpha)} \quad \text{dla} \quad y\geqslant\alpha. \tag{3.1.4}$$

Niech X będzie zmienną losową typu ciągłego o gęstości f, wówczas gęstość g rozkładu uciętego w poszczególnych przypadkach jest postaci:

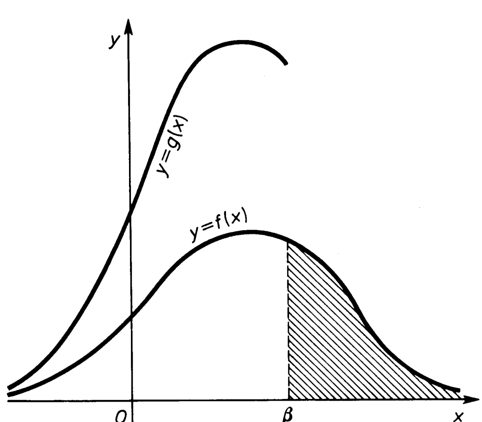
1° dla rozkładu uciętego dwustronnie, a więc o dystrybuancie (3.1.2):

$$g(y)=\frac{f(y)}{F(\beta)-F(\alpha)} \quad \text{dla} \quad \alpha\leqslant y\leqslant\beta. \tag{3.1.5}$$

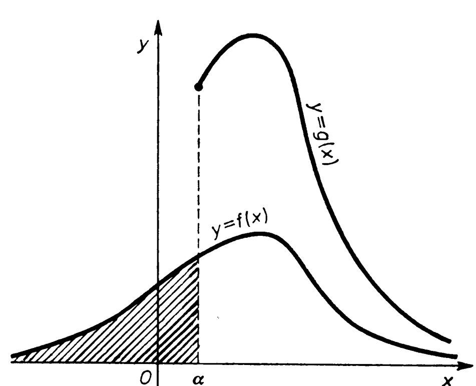
Rysunek 3.1 przedstawia wykresy obydwu gęstości f i g we wspólnym układzie współrzędnych.



Rys. 3.1. Gęstość f i gęstość g rozkładu dwustronnie uciętego we wspólnym układzie współrzędnych



Rys. 3.2. Gęstość f i gęstość g rozkładu uciętego prawostronnie



Rys. 3.3. Gęstość f i gęstość g rozkładu uciętego lewostronnie

2° dla rozkładu uciętego prawostronnie o dystrybuancie (3.1.3):

$$g(y)=\frac{f(y)}{F(\beta)} \quad \text{dla} \quad y\leqslant\beta, \tag{3.1.6}$$

wykresy zaś funkcji f i g przedstawia rysunek 3.2.

3° dla rozkładu uciętego lewostronnie o dystrybuancie (3.1.4):

$$g(y)=\frac{f(y)}{1-F(\alpha)} \quad \text{dla} \quad y\geqslant\alpha. \tag{3.1.7}$$

Podobnie jak w poprzednich przypadkach wykresy gęstości f i g znajdują się na rys. 3.3.

3.1.1. Zadania rozwiązane.

Zadanie 3.1. Czas bezawaryjnej pracy T elementu pewnego urządzenia ma rozkład wykładniczy o gęstości:

$$f(t)=\frac{1}{\lambda}\exp\left(-\frac{t}{\lambda}\right) \quad \text{dla} \quad t>0.$$

Po upływie czasu t_0 wymienia się element na nowy. Obliczyć wartość przeciętną czasu pracy elementu.

Rozwiązanie. Czas pracy elementu jest warunkową zmienną losową

$$Y=T\,|\,T\leqslant t_0,$$

a więc o rozkładzie uciętym prawostronnie, skąd jej gęstość g jest postaci

$$g(y)=\frac{f(y)}{F(t_0)} \quad \text{dla} \quad y\leqslant t_0,$$

gdzie

$$F(t_0)=\frac{1}{\lambda}\int\limits_0^{t_0}\exp\left(-\frac{t}{\lambda}\right)dt=1-\exp\left(-\frac{t_0}{\lambda}\right).$$

Gęstość „uciętego" czasu pracy elementu jest:

$$g(y)=\frac{1}{\lambda}\,\frac{\exp\left(-\frac{y}{\lambda}\right)}{1-\exp\left(-\frac{t_0}{\lambda}\right)} \quad \text{dla} \quad y\leqslant t_0.$$

Średni czas pracy elementu obliczamy, całkując:

$$EY=\frac{1}{\lambda\left[1-\exp\left(-\frac{t_0}{\lambda}\right)\right]}\int\limits_0^{t_0} ye^{-y/\lambda}dy=\lambda-\frac{t_0\exp\left(-\frac{t_0}{\lambda}\right)}{1-\exp\left(-\frac{t_0}{\lambda}\right)}.$$

Korzystając ze wzorów: (3.1.2) - (3.1.4) można także wyznaczyć rozkłady prawdopodobieństwa „uciętych" zmiennych losowych typu skokowego.

Zilustrujemy to na przykładzie:

ZADANIE 3.2. Zmienna losowa X ma rozkład dwumianowy określony wzorem: (2.7.7). Wyznaczyć ucięty rozkład dwumianowy zmiennej losowej Y, przyjmującej wartości: $i=1,\ldots,n$.

Rozwiązanie. $\bigwedge_{i=1,\ldots,n} P(Y=i)=P(X=i\mid X\neq 0)=\dfrac{P(X=i\wedge X\neq 0)}{P(X\neq 0)}=\dfrac{P(X=i\wedge X\neq 0)}{1-P(X=0)}.$

Wykorzystując rozkład dwumianowy, otrzymujemy:

$$\bigwedge_{i=1,\ldots,n} P(Y=i)=\frac{1}{1-(1-p)^n}\binom{n}{i}p^i(1-p)^{n-i}.$$

3.2. MIESZANINA DWÓCH ROZKŁADÓW

Załóżmy, że w magazynie znajdują się elementy pochodzące z produkcji dwóch zakładów, udział elementów pochodzących z produkcji pierwszego zakładu wynosi p_1, z drugiego p_2, gdzie $p_1, p_2>0$ oraz $p_1+p_2=1$.

Cechę X tych elementów traktujemy jako zmienną losową, która ma rozkład o dystrybuancie F_1, jeśli są one wyprodukowane przez zakład pierwszy oraz F_2, jeśli przez zakład drugi.

Cecha X elementów znajdujących się w magazynie ma rozkład o dystrybuancie:

$$F(x)=p_1F_1(x)+p_2F_2(x). \tag{3.2.1}$$

Rozkład o dystrybuancie (3.2.1) nazywamy *mieszaniną rozkładów* o dystrybuantach F_1, F_2, zaś p_1, p_2 — *udziałami* tych *rozkładów* w mieszaninie (3.2.1).

Jeśli f_1 i f_2 są gęstościami dwóch rozkładów typu ciągłego, to gęstość mieszaniny tych rozkładów wyraża się wzorem:

$$f(x)=p_1f_1(x)+p_2f_2(x), \tag{3.2.2}$$

gdzie $p_1, p_2>0$ oraz $p_1+p_2=1$.

Podobnie jeśli: $P_1(x_i)$, $i\in N$, $\sum_i P_1(x_i)=1$ oraz $P_2(x_i)$, $i\in N$, $\sum_i P_2(x_i)=1$ są funkcjami prawdopodobieństwa zmiennych losowych typu skokowego, to mieszaniną tych rozkładów jest rozkład typu skokowego określony wzorem:

$$\bigwedge_{i\in N} P(X=x_i)=p_1P_1(X=x_i)+p_2P_2(X=x_i), \tag{3.2.3}$$

gdzie p_1, $p_2>0$, $p_1+p_2=1$.

Możemy także rozważać mieszaninę rozkładów typu ciągłego o dystrybuancie F_c i typu skokowego o dystrybuancie F_d. Przypadek ten ze względu na specjalne znaczenie omówimy w p. 3.3.

Widać, że o ile istnieją skończone momenty zwykłe rzędu r poszczególnych rozkładów, równe odpowiednio α_r', α_r'', to istnieje skończony moment tego samego rzędu r mieszaniny tych rozkładów i jest on kombinacją liniową momentów α_r', α_r'':

$$\alpha_r = p_1\alpha_r' + p_2\alpha_r'', \qquad r \in N. \tag{3.2.4}$$

Zadanie 3.3. Do systemu obsługi zgłasza się $100p\%$ klientów uprzywilejowanych (pierwszeństwo obsługi) i $100(1-p)\%$ nieuprzywilejowanych. Rozkład czasu obsługi klientów uprzywilejowanych jest wykładniczy (2.8.7) z parametrem λ_1, a nieuprzywilejowanych — z parametrem λ_2. Wyznaczyć rozkład czasu obsługi losowo wybranego klienta i wyznaczyć jego średni czas obsługi.

Rozwiązanie. Czas obsługi losowo wybranego klienta ma rozkład będący mieszaniną dwóch rozkładów wykładniczych, którego gęstość jest postaci:

$$f(x) = \frac{p}{\lambda_1}\exp\left(-\frac{x}{\lambda_1}\right) + \frac{1-p}{\lambda_2}\exp\left(-\frac{x}{\lambda_2}\right) \quad \text{dla} \quad x \geqslant 0.$$

Ponieważ wartości przeciętne rozkładów wykładniczych są odpowiednio równe λ_1 i λ_2, więc (3.2.4) daje

$$EX = p\lambda_1 + (1-p)\lambda_2.$$

3.3. ROZKŁADY MIESZANE

Poza omówionymi w rozdziale 2 zmiennymi losowymi typu skokowego i ciągłego spotykamy w zastosowaniach zmienne losowe typu mieszanego. Na przykład przy badaniu materiałów o dużej wytrzymałości prowadzimy badania wytrzymałości X tylko do pewnej wartości x_0. Zaobserwowane wartości pozwalają opisać rozkład prawdopodobieństwa zmiennej X w badanej populacji tylko dla $x < x_0$, nie precyzując rozkładu dla $x \geqslant x_0$. Wówczas często przyjmuje się, że prawdopodobieństwo tego, że $X \geqslant x_0$ jest skupione w punkcie x_0 i jeżeli rozkład badanej cechy X w populacji jest ciągły, to rozkład prawdopodobieństwa otrzymany w wyżej opisany sposób jest rozkładem mieszanym.

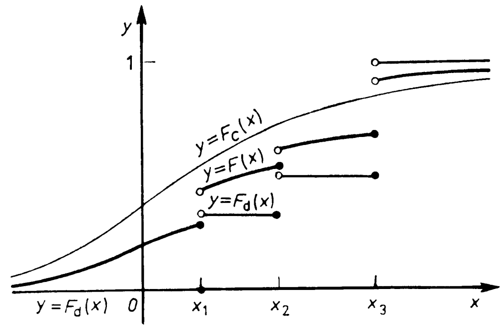
Zmienna losowa X jest *typu mieszanego*, jeśli jej dystrybuanta F jest postaci:

$$F(x) = p_1F_c(x) + p_2F_d(x), \tag{3.3.1}$$

gdzie F_c jest dystrybuantą zmiennej losowej typu ciągłego, F_d — dystrybuantą zmiennej losowej skokowej (dyskretnej), oraz p_1, $p_2 > 0$, $p_1 + p_2 = 1$.

Ze wzoru (3.3.1) i własności dystrybuanty wynika, że dystrybuanta zmiennej losowej typu mieszanego ma skończoną, bądź przeliczalną liczbę punktów nieciągłości, przy czym suma skoków (2.2.1) w tych punktach jest równa p_2, a więc jest mniejsza od 1. Wykresem funkcji (3.3.1) jest linia nieciągła, nie będąca jednak linią schodkową. Wykresy dystrybuant F_c, F_d i F dla $p_1 = p_2 = \frac{1}{2}$ we wspólnym układzie współrzędnych przedstawia rys. 3.4.

Dystrybuanta F — jak to widać ze wzoru (3.3.1) jest dystrybuantą mieszaniny (3.2.1) rozkładów ciągłego i skokowego o udziałach p_1, p_2. Można udowodnić, że wartość prze-



Rys. 3.4. Dystrybuanta F rozkładu typu mieszanego $F(x)=\frac{1}{2}F_c(x)+\frac{1}{2}F_d(x)$ o trzech punktach skokowych

ciętna (o ile istnieje) zmiennej losowej typu mieszanego wyraża się wzorem:

$$EX=\int_0^{\infty}[1-F(x+0)]\,dx-\int_{-\infty}^{0}F(x+0)\,dx, \tag{3.3.2}$$

gdzie $F(x+0)$ oznacza granicę prawostronną dystrybuanty w punkcie x.

3.3.1. Zadania rozwiązane.

ZADANIE 3.4. Zmienna losowa X ma rozkład trójkątny o gęstości:

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

Wyznaczyć rozkład zmiennej losowej $Y=\begin{cases} 0 & \text{dla} \quad X\leqslant 1,\\ X-1 & \text{dla} \quad X>1.\end{cases}$

Rozwiązanie. Zauważmy, że

$$P(Y=0)=P(X\leqslant 1)=\tfrac{1}{2}$$

oraz

$$\bigwedge_{0<y\leqslant 1} F(y)=P(Y<y)=P(X-1<y)=P(X<y+1)=$$

$$=\tfrac{1}{2}+\int_1^{y+1}(2-x)\,dx=\tfrac{1}{2}+\tfrac{1}{2}y(2-y).$$

Mamy więc dystrybuantę zmiennej losowej Y postaci:

$$F(y)=\begin{cases} 0 & \text{dla} \quad y\leqslant 0,\\ \frac{1}{2}+\frac{1}{2}y(2-y) & \text{dla} \quad 0<y\leqslant 1,\\ 1 & \text{poza tym}.\end{cases}$$

Zmienna losowa Y jest zmienną typu mieszanego, ponieważ jej dystrybuanta ma 1 punkt nieciągłości $y_0=0$ o skoku: $F(y_0+0)-F(y_0)=\frac{1}{2}<1$, można ją także przedstawić w postaci (3.3.1), przyjmując:

$$F_d(y)=\begin{cases} 0 & \text{dla} \quad y\leqslant 0,\\ 1 & \text{dla} \quad y>0,\end{cases} \qquad F_c(y)=\begin{cases} 0 & \text{dla} \quad y\leqslant 0,\\ y(y-2) & \text{dla} \quad y>0,\end{cases}$$

oraz $p_1=p_2=\frac{1}{2}$.

ZADANIE 3.5. Czas oczekiwania w pewnej kolejce ma rozkład, który składa się ze skoku $1-p_0$ w punkcie $x=0$ i ciągłej składowej typu wykładniczego o gęstości: $f(x)=\frac{p_0}{\lambda}\exp\left(-\frac{x}{\lambda}\right)$ dla $x>0$. Wyznaczyć dystrybuantę i obliczyć wartość przeciętną czasu oczekiwania T w kolejce.

Rozwiązanie. Wiadomo, że $P(X=0)=1-p_0$, oraz

$$\bigwedge_{x>0} P(0<X<x)=\frac{p_0}{\lambda}\int_0^x \exp\left(-\frac{t}{\lambda}\right)dt.$$

Stąd i z definicji dystrybuanty mamy:

$$F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 0, \\ 1-p_0+\dfrac{p_0}{\lambda}\displaystyle\int_0^x e^{-t/\lambda}dt=1-p_0 e^{-x/\lambda} & \text{dla} \quad x>0. \end{cases}$$

Zauważmy, że przyjmując $p_1=p_0$, $p_2=1-p_0$, oraz

$$F_d(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 0, \\ 1 & \text{dla} \quad x>0, \end{cases} \qquad F_c(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 0, \\ 1-\exp\left(-\dfrac{x}{\lambda}\right) & \text{dla} \quad x>0, \end{cases}$$

możemy F przedstawić w postaci (3.3.1).

Korzystając ze wzoru (3.3.2), otrzymamy

$$EX=p_0\int_0^\infty c^{-x/\lambda}dx=p_0.$$

3.4. ZADANIA DO ROZWIĄZANIA

3.6. Gęstość zaludnienia w pewnym mieście jest postaci:

$$\bigwedge_{r>0} c(r)=ae^{-br}, \qquad a,b>0,$$

gdzie r jest odległością (radialną) od centrum. Traktując r jako zmienną losową, a wyżej podaną funkcję jako jej gęstość prawdopodobieństwa, wyznaczyć największą odległość d od centrum, potrzebną dla planowania sieci komunikacji, tak aby 75% mieszkańców była obsługiwana przez nią. Obliczyć wartość przeciętną odległości od centrum: a) mieszkańca tego miasta, b) mieszkańca korzystającego z sieci komunikacyjnej.

3.7. Zmienna losowa X ma rozkład Poissona określony wzorem

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

Wyznaczyć ucięty rozkład Poissona zmiennej losowej Y przyjmującej wartości $i=2, 3, \ldots$

3.8. Czas T trwania pewnej operacji technicznej jest zmienną losową o gęstości:

$$f(t)=\begin{cases} te^{-t} & \text{dla } t>0, \\ 0 & \text{poza tym.} \end{cases}$$

W praktyce przyjmuje się jednak, że czas T nie może być większy niż $T=10$ h. Wyznaczyć ucięty rozkład czasu T. Obliczyć wartość przeciętną czasu trwania operacji.

3.9. W magazynie znajdują się żarówki pochodzące z produkcji dwóch zakładów: 75% z zakładu A i 25% z B. Zakładamy, że czas świecenia żarówek produkowanych przez zakład A jest zmienną losową o gęstości

$$f_1(x)=\begin{cases} \frac{1}{100}\exp(-\frac{1}{100}x) & \text{dla } x>0, \\ 0 & \text{poza tym,} \end{cases}$$

zaś czas X świecenia żarówek produkowanych przez zakład B jest zmienną losową o gęstości

$$f_2(x)=\begin{cases} \dfrac{1}{2\cdot 100^3}x^2\exp(-\frac{1}{100}x) & \text{dla } x>0. \\ 0 & \text{poza tym}. \end{cases}$$

Wyznaczyć rozkład czasu t świecenia żarówek znajdujących się w magazynie. Obliczyć wartość przeciętną czasu świecenia losowo wybranej żarówki.

3.10. Zmienna losowa X jest typu mieszanego, o wartościach z przedziału $\langle 1, 3\rangle$. Niech $P(X=1)=\frac{1}{3}=P(X=3)$, oraz $\bigwedge_{1<x<3} f(x)=\frac{1}{2}(x-2)^2$. Wyznaczyć i naszkicować dystrybuantę tego rozkładu oraz obliczyć wartość przeciętną zmiennej losowej X.

3.11. Niech X będzie zmienną losową przyjmującą wartości równe liczbie orłów wyrzuconych na dwóch monetach, Y zaś zmienną losową o rozkładzie równomiernym w przedziale $\langle -1, 1\rangle$. Wyznaczyć i naszkicować dystrybuantę mieszaniny tych rozkładów przyjmując ich udziały odpowiednio równe: $p_1=\frac{1}{3}$, $p_2=\frac{2}{3}$. Obliczyć wartość przeciętną w wyznaczonym rozkładzie.

Odpowiedzi

3.6. Na to, aby $c(r)$ było gęstością prawdopodobieństwa, potrzeba i wystarcza, aby $a=b$. Dla mieszkańców korzystających z sieci komunikacji gęstość prawdopodobieństwa $g(r)=\dfrac{f(r)}{\int_0^d f(r)\,dr}$ dla $0<r\leqslant d$, gdzie d jest rozwiązaniem równania

$$P(R<d)=\frac{3}{4}\Leftrightarrow b\int_0^d e^{-br}dr=\frac{3}{4}\Leftrightarrow d=\frac{1}{b}\ln 4.$$

Wartość przeciętna odległości od centrum mieszkańca miasta $ER=1/b$, zaś wartość przeciętna tej odległości dla mieszkańca korzystającego z sieci komunikacyjnej $R_k=\dfrac{1}{b}-\dfrac{de^{-bd}}{1-e^{-db}}$.

3.7. $\bigwedge_{i=2,3,\ldots} P(Y=i)=\dfrac{\lambda^i e^{-\lambda}}{i!\,[1-e^{-\lambda}(1+\lambda)]}.$

3.8. $$g(t)=\begin{cases} \dfrac{te^{-t}}{1-\int\limits_{10}^{\infty} te^{-t}dt}=\dfrac{te^{-t}}{1-11e^{-10}} & \text{dla} \quad 0<t\leqslant 10, \\ 0 & \text{poza tym}, \end{cases}$$

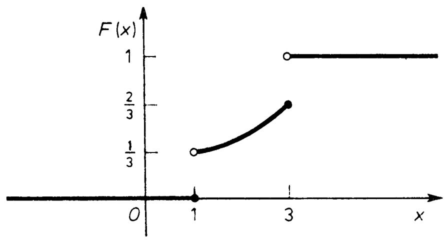
$$ET=\frac{2-82e^{-10}}{1-11e^{-10}}.$$

3.9. $f(x)=\dfrac{0{,}75}{100}\exp\left(-\dfrac{1}{100}x\right)+\dfrac{0{,}25}{2\cdot 100^3}x^2\exp\left(-\dfrac{1}{100}x\right)$ dla $x>0$,

$$EX=\frac{0{,}25}{100}\int\limits_0^{\infty}\left(3x+\frac{x^3}{2\cdot 100^2}\right)\exp\left(-\frac{1}{100}x\right)dx=150.$$

3.10. $$F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 1, \\ \frac{1}{3}+\int\limits_1^x \frac{1}{2}(t-2)^2dt=\frac{1}{3}+\frac{1}{6}(x-2)^3 & \text{dla} \quad 1<x\leqslant 3, \\ \frac{1}{3}+\int\limits_1^3 \frac{1}{2}(x-2)^2dx+\frac{1}{3}=1 & \text{dla} \quad x>3 \end{cases}$$

(rys. 3.5)



Rys. 3.5. Odpowiedź do zadania 3.10

$$EX=1\cdot\tfrac{1}{3}+3\cdot\tfrac{1}{3}+\tfrac{1}{2}\int\limits_1^3 x(x-2)^2dx=2,$$

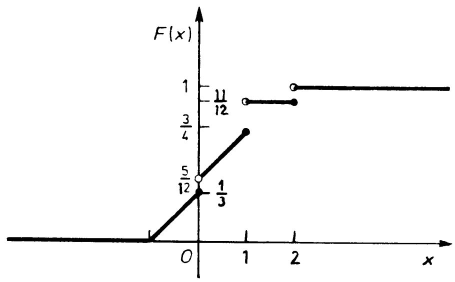
co w prostszy sposób wynika z symetrii rozkładu mieszanego.

3.11. Dystrybuanta rozkładu dyskretnego

$$F_d(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant 0, \\ \frac{1}{4} & \text{dla} \quad 0<x\leqslant 1, \\ \frac{3}{4} & \text{dla} \quad 1<x\leqslant 2, \\ 1 & \text{dla} \quad x>2 \end{cases}$$

oraz dystrybuanta rozkładu ciągłego

$$F_c(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant -1, \\ \frac{1}{2}(x+1) & \text{dla} \quad |x|<1, \\ 1 & \text{dla} \quad x\geqslant 1. \end{cases}$$



Rys. 3.6. Odpowiedź do zadania 3.11

Dystrybuanta mieszaniny tych rozkładów $F(x)=\frac{1}{3}F_d(x)+\frac{2}{3}F_c(x)$, a więc jest określona wzorami:

$$F(x)=\begin{cases} 0 & \text{dla} \quad x\leqslant -1, \\ \frac{1}{3}(x+1) & \text{dla} \quad -1<x\leqslant 0, \\ \frac{1}{12}+\frac{1}{3}(x+1) & \text{dla} \quad 0<x\leqslant 1, \\ \frac{11}{12} & \text{dla} \quad 1<x\leqslant 2, \\ 1 & \text{dla} \quad x>2 \end{cases}$$

(rys. 3.6)

$$EX=-\int_{-1}^{0}\tfrac{1}{3}(x+1)\,dx+\int_{0}^{1}[1-\tfrac{1}{12}-\tfrac{1}{3}(x+1)]\,dx+\int_{1}^{2} 1-\tfrac{11}{12})\,dx=\tfrac{1}{3}.$$

4 FUNKCJE CHARAKTERYSTYCZNE

4.1. WŁASNOŚCI FUNKCJI CHARAKTERYSTYCZNYCH

Funkcją charakterystyczną zmiennej losowej X nazywamy wartość przeciętną funkcji e^{itX}, gdzie t jest zmienną rzeczywistą, a i — tzw. jednostką urojoną; oznaczamy tę funkcję przez $\varphi(t)$:

$$\bigwedge_{t \in \boldsymbol{R}} \varphi(t) = E(\exp itX). \tag{4.1.1}$$

Ponieważ na podstawie wzoru Eulera $|e^{itx}| = |\cos tx + i \sin tx| = \sqrt{\cos^2 tx + \sin^2 tx} = 1$ dla dowolnych $t \in \boldsymbol{R}$, $x \in \boldsymbol{R}$, więc wzór (4.1.1) określa funkcję charakterystyczną dla dowolnej zmiennej losowej. Zgodnie z definicją wartości przeciętnej, funkcja charakterystyczna jest określona wzorem:

$$\varphi(t) = \begin{cases} \sum\limits_k p_k e^{itx_k} & \text{dla zmiennej losowej dyskretnej} \\ & \text{o rozkładzie } P(X = x_k) = p_k, \text{ gdzie } \sum\limits_k p_k = 1, \\ \int\limits_{-\infty}^{\infty} e^{itx} f(x)\,dx & \text{dla zmiennej losowej typu ciągłego o gęstości } f. \\ & \text{Jest to przekształcenie Fouriera funkcji } f(x). \end{cases} \tag{4.1.2}$$

Funkcja charakterystyczna φ zmiennej losowej X jest funkcją zespoloną zmiennej rzeczywistej t i ma następujące własności:

a) $\varphi(0) = 1$,

b) $\bigwedge_{t \in \boldsymbol{R}} \varphi(t) = \overline{\varphi(-t)}$, gdzie $\overline{\varphi(-t)}$ oznacza liczbę zespoloną sprzężoną z $\varphi(-t)$,

c) $\bigwedge_{t \in \boldsymbol{R}} |\varphi(t)| \leqslant 1$,

d) φ jest funkcją ciągłą na całej prostej,

e) φ jest funkcją rzeczywistą wtedy i tylko wtedy, gdy rozkład zmiennej losowej X jest symetryczny względem $x = 0$.

Warunki a) - e) nie są jednak warunkami wystarczającymi do stwierdzenia, że funkcja jest funkcją charakterystyczną pewnej zmiennej losowej; kontrprzykładem jest $\varphi(t) = \dfrac{1}{1+t^4}$.

Zauważmy, że wzór (4.1.2) określa jednoznacznie dla danego rozkładu prawdopodobieństwa funkcję charakterystyczną tego rozkładu; można wykazać, że również odwrotnie: Przy założeniu istnienia skończonej całki $\int\limits_{-\infty}^{\infty} |\varphi(t)|\,dt$, zachodzi wzór:

$$f(x)=F'(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-itx}\varphi(t)\,dt, \tag{4.1.3}$$

przy czym gęstość f zmiennej losowej X jest funkcją ciągłą dla $x \in \boldsymbol{R}$.

Jeśli funkcja charakterystyczna ma okres 2π, to przyjmuje tylko wartości całkowite

$$\bigwedge_{k\in C} p_k=\frac{1}{2\pi}\int\limits_{-\pi}^{\pi} e^{-itk}\varphi(t)\,dt, \tag{4.1.4}$$

gdzie $p_k=P(X=k)$ dla $k=0, \mp 1, \mp 2, \ldots$, oraz $\sum\limits_k p_k=1$, przy czym nie wszystkie p_k muszą być dodatnie.

Powyższe wyniki są szczególnymi przypadkami następującego twierdzenia:

Jeżeli φ jest funkcją charakterystyczną zmiennej losowej X o dystrybuancie F, a x i $x+h$ są punktami ciągłości dystrybuanty F, to

$$F(x+h)-F(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\frac{\exp(-itx)-\exp it(x+h)}{it}\varphi(t)\,dt. \tag{4.1.5}$$

Wzór ten podaje jak można za pomocą funkcji charakterystycznej obliczyć: $P(x\leqslant X<x+h)$, jeśli x i $x+h$ są punktami ciągłości dystrybuanty.

Można więc przy badaniu własności rozkładów prawdopodobieństwa, — tam gdzie jest to wygodniej — posługiwać się funkcjami charakterystycznymi.

Szerokie zastosowanie funkcji charakterystycznych wiąże się z ich następującymi własnościami:

1. *Jeżeli istnieje k-ty moment zmiennej losowej X o funkcji charakterystycznej φ, to φ jest k-krotnie różniczkowalna (w sposób ciągły) i zachodzi związek:*

$$\alpha_k=EX^k=\frac{1}{i^k}\varphi^{(k)}(0) \tag{4.1.6}$$

oraz jeśli można rozwinąć $\varphi(t)$ w szereg Maclaurina, to

$$\varphi(t)=\sum_{k=0}^{\infty} a_k t^k, \quad gdzie \quad a_k=\frac{\varphi^{(k)}(0)}{k!}. \tag{4.1.7}$$

Znając współczynniki a_k rozwinięcia (4.1.7), możemy obliczyć dowolny moment zwykły według wzoru:

$$\alpha_k=\frac{a_k\,k!}{i^k}. \tag{4.1.8}$$

2. *Funkcja charakterystyczna sumy dowolnej skończonej liczby niezależnych zmiennych losowych równa się iloczynowi funkcji charakterystycznych tych zmiennych.*

Warto podkreślić, że również funkcja charakterystyczna sumy zmiennych losowych zależnych może być równa iloczynowi funkcji charakterystycznych.

Z definicji funkcji charakterystycznej i wzoru (2.4.4) wynika, że funkcja charakterystyczna funkcji zmiennej losowej $Y=g(X)$ jest postaci:

$$\varphi(t)=\begin{cases} \sum\limits_k e^{itg(x_k)}p_k & \text{dla skokowej zmiennej losowej o rozkładzie } p_k=P(X=x_k), \\ \int\limits_{-\infty}^{\infty} e^{itg(x)}f(x)\,dx & \text{dla ciągłej zmiennej losowej o gęstości } f. \end{cases} \tag{4.1.9}$$

4.1.1. Zadania rozwiązane.

ZADANIE 4.1. Wyznaczyć funkcję charakterystyczną dla rozkładu dwumianowego: $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$, $k=0,\ldots,n$ dla $0<p<1$.

Rozwiązanie. Korzystamy ze wzoru (4.1.2) dla zmiennej losowej skokowej:

$$\varphi(t)=\sum_{k=0}^{n} e^{itk}\binom{n}{k}p^kq^{n-k}=\sum_{k=0}^{n}\binom{n}{k}(pe^{it})^kq^{n-k}=$$
$$=(pe^{it}+q)^n, \quad \text{gdzie} \quad q=1-p.$$

ZADANIE 4.2. Wyznaczyć gęstość f zmiennej losowej X o funkcji charakterystycznej postaci: $\varphi(t)=\exp(2it-3|t|)$.

Rozwiązanie. Funkcja φ spełnia warunek: $\int\limits_{-\infty}^{\infty}|\varphi(t)|\,dt<\infty$. Istotnie $|\varphi(t)|=$ $=e^{-3|t|}|\cos 2t+i\sin 2t|=e^{-3|t|}$, przy czym

$$\int_{-\infty}^{\infty} e^{-3|t|}dt=2\int_0^{\infty} e^{-3t}dt=\tfrac{2}{3}<\infty.$$

Możemy więc zastosować wzór (4.1.3)

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-itx+2it-3|t|)\,dt.$$

Zbadamy więc istnienie granic:

$$\lim_{N\to\infty}\int_{-N}^{0}\exp(-itx+2it+3t)\,dt+\lim_{M\to\infty}\int_0^M\exp(-itx+2it-3t)\,dt=$$
$$=\lim_{N\to\infty}\left[\frac{\exp[t(3+2i-ix)]}{3+2i-ix}\right]_{-N}^{0}+\lim_{M\to\infty}\left[\frac{\exp[t(-3+2i-ix)]}{-3+2i-ix}\right]_0^M=$$

$$=\frac{1}{3+i(2-x)}+\lim_{N\to\infty} e^{-3N}\frac{\cos[(2-x)N]-i\sin[(2-x)N]}{3+2i-ix}-\frac{1}{-3+i(2-x)}+$$

$$+\lim_{M\to\infty} e^{-3M}\frac{\cos[(2-x)M]+i\sin[(2-x)M]}{-3+2i-ix}.$$

Drugi i czwarty składnik jest zerem, ponieważ są to granice iloczynów funkcji, w których pierwszy czynnik dąży do 0, a drugi jest funkcją ograniczoną. Otrzymujemy więc

$$f(x)=\frac{1}{2\pi}\left[\frac{-3+i(2-x)-3-i(2-x)}{i^2(2-x)^2-9}\right]=\frac{1}{\pi}\frac{3}{(x-2)^2+9}.$$

Jest to gęstość rozkładu Cauchy'ego (2.8.37) o parametrach $\mu=2$, $\lambda=3$.

Zadanie 4.3. Wyznaczyć rozkład prawdopodobieństwa, którego funkcja charakterystyczna jest postaci: $\varphi(t)=e^{2it}$.

Rozwiązanie. Ponieważ $|\varphi(t)|=1$, więc nie istnieje całka skończona: $\int\limits_{-\infty}^{\infty}|\varphi(t)|dt$ i nie możemy skorzystać ze wzoru (4.1.3). Ponieważ e^{2it} jest funkcją okresową, więc zmienna losowa jest typu skokowego; wówczas ze wzoru (4.1.4) mamy

$$p_k=\frac{1}{2\pi}\int\limits_{-\pi}^{\pi} e^{-itk+2it}dt \quad \text{dla} \quad k=0, \mp 1, \ldots$$

$$p_k=\frac{1}{2\pi}\int\limits_{-\pi}^{\pi}[\cos(2-k)t+i\sin(2-k)t]\,dt=$$

$$=\frac{1}{\pi}\int\limits_{0}^{\pi}\cos(2-k)t\,dt=\begin{cases}0 & \text{dla} \quad k\neq 2,\\ 1 & \text{dla} \quad k=2.\end{cases}$$

Stąd funkcja $\varphi(t)=e^{2it}$ jest funkcją charakterystyczną zmiennej losowej o rozkładzie jednopunktowym skupionym w punkcie $X=2$: $P(X=2)=1$. Ten sam wynik można uzyskać nie posługując się wzorem (4.1.4), a opierając się na definicji funkcji charakterystycznej (dla zmiennych dyskretnych), skąd wynika, że winno być

$$e^{2it}\equiv\sum_k e^{itx_k}p_k, \quad \text{gdzie} \quad p_k=P(X=x_k).$$

Ponieważ funkcja charakterystyczna zmiennej losowej wyznacza rozkład jednoznacznie, a tożsamość powyższa może zachodzić wtedy i tylko wtedy, gdy suma (albo szereg) po prawej stronie redukuje się do jednego składnika identycznego z lewą stroną, więc poszukiwanym rozkładem jest rozkład jednopunktowy, określony wzorem: $P(X=2)=1$.

Zadanie 4.4. Zmienna losowa X podlega rozkładowi o funkcji charakterystycznej $\varphi(t)=\exp\left(-\frac{1}{2}t^2\right)$ (jest to funkcja charakterystyczna rozkładu $N(0, 1)$). Wyznaczyć czwarty moment zwykły tej zmiennej.

Rozwiązanie. I sposób. Obliczamy pochodne funkcji

$$\varphi'(t) = -t\exp(-\tfrac{1}{2}t^2),$$

$$\varphi''(t) = (t^2-1)\exp(-\tfrac{1}{2}t^2),$$

$$\varphi'''(t) = (-t^3+3t)\exp(-\tfrac{1}{2}t^2),$$

$$\varphi^{(4)}(t) = (t^4-6t^2+3)\exp(-\tfrac{1}{2}t^2), \qquad \varphi^{(4)}(0)=3.$$

Stąd i ze wzoru (4.1.6) otrzymujemy:

$$\alpha_4 = \frac{1}{i^4}3 = 3.$$

II sposób. Rozwińmy funkcję $\varphi(t) = \exp(-\frac{1}{2}t^2)$ w szereg Maclaurina. Korzystając z rozwinięcia funkcji $\exp t = \sum_{k=0}^{\infty} \frac{t^k}{k!}$ dla $t \in \boldsymbol{R}$, mamy:

$$\exp(-\tfrac{1}{2}t^2) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k} \cdot \frac{t^{2k}}{k!} \qquad \text{dla} \qquad t \in \boldsymbol{R}.$$

W rozwinięciu tym nie występują nieparzyste potęgi t, a więc na podstawie wzoru (4.1.8) wnioskujemy, że wszystkie momenty nieparzystego rzędu są zerami. Współczynniki przy parzystych potęgach t są postaci

$$a_{2k} = \frac{(-1)^k}{2^k k!} \qquad \text{dla} \qquad k \in \boldsymbol{N}_0.$$

Stąd i ze wzoru (4.1.8) otrzymujemy:

$$\alpha_{2k} = \frac{(-1)^k(2k)!}{2^k k!\, i^{2k}} = \frac{(2k)!}{2^k k!} \qquad \text{dla} \qquad k \in \boldsymbol{N}.$$

Dla $k=2$ mamy $\alpha_4 = 3$.

ZADANIE 4.5. Wykazać, że jeżeli X i Y są niezależnymi zmiennymi losowymi, które mają rozkłady gamma o gęstościach równych odpowiednio:

$$f_1(x) = \frac{1}{\Gamma(p_1)} x^{p_1-1}e^{-x} \qquad \text{dla} \qquad x>0,$$

$$f_2(y) = \frac{1}{\Gamma(p_2)} y^{p_2-1}e^{-y} \qquad \text{dla} \qquad y>0,$$

to zmienna losowa $Z=X+Y$ ma rozkład gamma o gęstości

$$f(z) = \frac{1}{\Gamma(p_1+p_2)} z^{p_1+p_2-1}e^{-z} \qquad \text{dla} \qquad z>0,$$

a więc z parametrem $p=p_1+p_2$.

Własność tę ujmujemy krótko mówiąc: *dla rozkładu gamma (jednoparametrowego) zachodzi twierdzenie o dodawaniu względem parametru p, co oznacza, że aby otrzymać gęstość sumy dwóch (albo więcej) niezależnych zmiennych losowych o jednoparametrowych rozkładach gamma wystarczy w gęstości jednego ze składników parametr zastąpić przez sumę parametrów obu (wszystkich) składników.*

Rozwiązanie. Funkcja charakterystyczna zmiennej losowej X jest postaci

$$\varphi_1(t)=\frac{1}{\Gamma(p_1)}\int_0^\infty e^{itx}x^{p_1-1}e^{-x}dx=\frac{1}{\Gamma(p_1)}\int_0^\infty x^{p_1-1}e^{-(1-it)x}dx=$$

$$=\frac{\Gamma(p_1)}{\Gamma(p_1)(1-it)^{p_1}}=\frac{1}{(1-it)^{p_1}}.$$

Zastępując p_1 przez p_2, otrzymujemy funkcję charakterystyczną zmiennej losowej Y:

$$\varphi_2(t)=\frac{1}{(1-it)^{p_2}}.$$

Ponieważ X i Y są niezależnymi zmiennymi losowymi, więc z własności 2) wynika, że funkcja charakterystyczna zmiennej losowej Z jest równa

$$\varphi(t)=\varphi_1(t)\,\varphi_2(t)=\frac{1}{(1-it)^{p_1+p_2}}.$$

Otrzymaliśmy więc funkcję charakterystyczną rozkładu gamma o parametrze $p=p_1+p_2$. Stąd zmienna losowa Z ma rozkład gamma o gęstości f.

ZADANIE 4.6. Zmienna losowa X ma rozkład $N(0, 1)$; wyznaczyć rozkład zmiennej losowej $Y=X^2$.

Rozwiązanie. Wyznaczmy funkcję charakterystyczną zmiennej losowej Y:

$$\varphi(t)=E(\exp itX^2)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(itx^2-\tfrac{1}{2}x^2)\,dx=$$

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

Otrzymaliśmy funkcję charakterystyczną rozkładu gamma (2.8.12) o parametrach $\lambda=2$, $p=\frac{1}{2}$, zmienna losowa Y ma więc rozkład χ^2 o 1 stopniu swobody (2.8.16), którego gęstością jest funkcja: $f(x)=\dfrac{1}{\sqrt{2}\,\Gamma(\frac{1}{2})}x^{-1/2}e^{-x/2}$ dla $x>0$.

Tablica 4.1. Funkcje charakterystyczne podstawowych rozkładów prawdopodobieństwa

Lp.	Nazwa rozkładu	Rozkład prawdopodobieństwa	Funkcja charakterystyczna
1	dwumianowy	$p_k=\binom{n}{k}p^k q^{n-k}$, $k=0,1,\ldots,n$ $p,q>0$, $p+q=1$	$\varphi(t)=(pe^{it}+q)^n$
2	ujemny dwumianowy	$p_k=\binom{k-1}{\nu-1}p^{\nu}q^{k-\nu}, k=\nu,\nu+1,\ldots$ $p,q>0$, $p+q=1$, $\nu\in\boldsymbol{N}$	$\varphi(t)=\left(\frac{pe^{it}}{1-qe^{it}}\right)^{\nu}$
3	Poissona	$p_k=\frac{e^{-\lambda}\lambda^k}{k!}$, $k\in\boldsymbol{N}_0$ $\lambda>0$	$\varphi(t)=\exp[\lambda(e^{it}-1)]$
4	równomierny	$f(x)=\frac{1}{b-a}$ dla $a<x<b$	$\varphi(t)=\frac{1}{b-a}\frac{e^{itb}-e^{ita}}{it}$
5	normalny	$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$, $\sigma>0$, $\mu\in\boldsymbol{R}$	$\varphi(t)=\exp\left(i\mu t-\frac{\sigma^2 t^2}{2}\right)$
6	gamma (dwuparametrowy)	$f(x)=\frac{1}{\lambda^p\Gamma(p)}x^{p-1}\exp\left(-\frac{x}{\lambda}\right)$ $x>0$, $\lambda>0$, $p>0$	$\varphi(t)=\frac{1}{(1-\lambda it)^p}$
7	Laplace'a	$f(x)=\frac{1}{2}e^{-\|x\|}$	$\varphi(t)=\frac{1}{1+t^2}$
8	Cauchy'ego	$f(x)=\frac{\lambda}{\pi[\lambda^2+(x-\mu)^2]}$ $\lambda>0$, $\mu\in\boldsymbol{R}$	$\varphi(t)=\exp(i\mu t-\lambda\|t\|)$
9	wykładniczy	$f(x)=\frac{1}{\lambda}\exp\left(-\frac{x}{\lambda}\right)$, $x>0$	$\varphi(t)=\frac{1}{1-\lambda it}$

4.2. ZADANIA DO ROZWIĄZANIA

4.7. Wyznaczyć funkcję charakterystyczną rozkładu:

a) zero-jedynkowego,

b) geometrycznego,

c) liczby wyrzuconych orłów przy rzucie trzema monetami,

d) wykładniczego o gęstości $f(x)=\begin{cases} \frac{1}{\lambda}\exp\left(-\frac{x-a}{\lambda}\right) & \text{dla} \quad x>a, \\ 0 & \text{dla} \quad x\leqslant a, \end{cases}$

e) Laplace'a o gęstości: $f(x)=\frac{1}{2}e^{-|x|}$.

4.8. Zmienna losowa X ma rozkład prawdopodobieństwa o funkcji charakterystycznej φ. Wyznaczyć funkcję charakterystyczną zmiennej $Y=aX+b$, gdzie a, b są dowolnymi liczbami rzeczywistymi $(a\neq 0)$.

4.9. Wykazać, że niżej podane funkcje nie mogą być funkcjami charakterystycznymi:

a) $\varphi(t)=\exp(-it^2)$, b) $\varphi(t)=\dfrac{1}{1+i|t|}$,

c) $\varphi(t)=\begin{cases} 2-|t| & \text{dla} \quad |t|<2, \\ 0 & \text{dla} \quad |t|\geqslant 2. \end{cases}$

4.10. Wyznaczyć gęstość prawdopodobieństwa zmiennej losowej X, o funkcji charakterystycznej postaci:

a) $\varphi(t)=\exp(-\frac{1}{2}t^2)$,

b) $\varphi(t)=\begin{cases} 1-|t| & \text{dla} \quad |t|\leqslant 1, \\ 0 & \text{dla} \quad |t|>1, \end{cases}$

c) $\varphi(t)=\cos t$,

d) $\varphi(t)=\sum\limits_{k=1} (\frac{1}{2})^k \cos kt$.

4.11. Wyznaczyć za pomocą funkcji charakterystycznych:

a) drugi moment zwykły rozkładu Poissona,

b) trzeci moment zwykły rozkładu dwumianowego,

c) k-ty moment zwykły rozkładu wykładniczego, o gęstości

$$f(x)=\frac{1}{\lambda}\exp\left(-\frac{x}{\lambda}\right) \quad \text{dla} \quad x>0, \quad \lambda>0,$$

d) k-ty moment zwykły dwuparametrowego rozkładu gamma o gęstości postaci:

$$f(x)=\frac{a^p}{\Gamma(p)}x^{p-1}e^{-ax} \quad \text{dla} \quad x>0, \quad a,p>0.$$

4.12. Niech X i Y będą niezależnymi zmiennymi losowymi o rozkładach:

a) Poissona z parametrami λ_1 i λ_2,

b) wykładniczych (2.8.7) z tym samym parametrem λ,

c) gamma o gęstościach odpowiednio równych:

$$f(x)=\frac{a^{p_1}}{\Gamma(p_1)}x^{p_1-1}e^{-ax}, \quad x>0, \qquad g(y)=\frac{a^{p_2}}{\Gamma(p_2)}y^{p_2-1}e^{-ay}, \quad y>0,$$

d) Cauchy'ego o gęstościach:

$$f(x)=\frac{\lambda_1}{\pi[\lambda_1^2+(x-\mu_1)^2]},\qquad g(y)=\frac{\lambda_2}{\pi[\lambda_2^2+(y-\mu_2)^2]}.$$

e) Dla których spośród powyższych rozkładów zachodzi tw. o dodawaniu i jeśli tak, to względem jakich parametrów? Wyznaczyć rozkład zmiennej losowej $Z=X+Y$.

4.13. Zmienna losowa X ma rozkład geometryczny $p_k=P(X=k)=(1-p)p^k$, $k\in N_0$, $0<p<1$. Wyznaczyć rozkład sumy r niezależnych zmiennych losowych o tym samym rozkładzie geometrycznym.

4.14. $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie $N(\mu, \sigma)$. Wyznaczyć rózkład zmiennej losowej $X=\frac{1}{n}\sum_{i=1}^{n} X_i$.

4.15. Wykazać, że jeśli X i Y są niezależnymi zmiennymi losowymi o rozkładach normalnych: $N(\mu_1, \sigma_1)$, $N(\mu_2, \sigma_2)$, to zmienna losowa $Z=aX+bY$, gdzie a i b są dowolnymi liczbami rzeczywistymi spełniającymi warunek $a^2+b^2>0$, ma rozkład: $N(a\mu_1+b\mu_2, \sqrt{a^2\sigma_1^2+b^2\sigma_2^2})$.

4.16. Niech X i Y będą niezależnymi zmiennymi losowymi o tym samym rozkładzie wykładniczym o gęstości:

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

Wyznaczyć rozkład zmiennej losowej $Z=X-Y$.

4.17. Niech X i Y będą niezależnymi zmiennymi losowymi o jednakowych funkcjach charakterystycznych $\varphi(t)$. Wyznaczyć funkcję charakterystyczną zmiennej $Z=X-Y$.

4.18. Wyznaczyć funkcję charakterystyczną rozkładu prawdopodobieństwa o gęstości

$$f(x)=\begin{cases} \frac{1}{2}(1-\frac{1}{2}|x|) & \text{dla} \quad |x|\leqslant 2,\\ 0 & \text{dla pozostałych } x.\end{cases}$$

4.19. Wykazać, że dla dowolnej rzeczywistej funkcji charakterystycznej φ zachodzi nierówność

$$1+\varphi(2t)\geqslant 2[\varphi(t)]^2$$

(dowód przeprowadzić dla zmiennej losowej ciągłej).

Odpowiedzi

4.7. a) $pe^{it}+q$; b) $\frac{pe^{it}}{1-qe^{it}}$; c) $\frac{1}{8}(e^{it}+1)^3$; d) $\frac{e^{ait}}{1-\lambda it}$; e) $\frac{1}{1+t^2}$.

4.8. $\varphi_1(t)=e^{itb}\varphi(at)$.

4.9. a) i b) φ nie spełnia warunku $\varphi(t)=\overline{\varphi(-t)}$, c) $\varphi(t)$ nie spełnia warunku $\varphi(0)=1$.

4.10. a) Zastosować we wzorze (4.1.3) podstawienie $t=u-ix$, $f(x)=\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)$; b) $f(x)=\frac{1-\cos x}{\pi x^2}$, $x\neq 0$; c) $P(X=1)=\frac{1}{2}$, $P(X=-1)=\frac{1}{2}$; d) $P(X=k)=\frac{1}{2}(\frac{1}{2})^{|k|}$, dla $k\in C-\{0\}$.

4.11. a) $\alpha_2=\lambda^2+\lambda$; b) $\alpha_3=np[p^2(n-1)(n-2)+3p(n-1)+1]$;
c) rozwinąć $\varphi(t)$ w szereg Maclaurina, $\alpha_k=\lambda^k k!$, $k\in N$;
d) $\alpha_k=\dfrac{p(p+1)\ldots(p+k-1)}{a^k}$.

4.12. a) $p_k=P(Z=k)=\dfrac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^k}{k!}$, $k\in N_0$;

b) $f(z)=\dfrac{1}{\lambda^2}z\exp\left(-\dfrac{z}{\lambda}\right)$, $z>0$; c) $f(z)=\dfrac{a^{p_1+p_2}}{\Gamma(p_1+p_2)}z^{p_1+p_2-1}e^{-az}$, $z>0$;

d) $f(z)=\dfrac{1}{\pi}\dfrac{\lambda_1+\lambda_2}{(\lambda_1+\lambda_2)^2+(z-\mu_1-\mu_2)^2}$; e) Tw. o dodawaniu zachodzi w przypadku a) rozkładu Poissona, c) dwuparametrowego rozkładu gamma względem p przy wspólnym parametrze a, d) rozkładu Cauchy'ego względem obu parametrów.

4.13. Jest to tzw. rozkład ujemny dwumianowy z parametrami r, p (2.7.14).

4.14. X ma rozkład $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$.

4.16. $f(z)=\frac{1}{2}e^{-|z|}$.

4.17. $\varphi(t)=Ee^{it(X-Y)}=\varphi(t)\varphi(-t)=\varphi(t)\overline{\varphi(t)}=|\varphi(t)|^2$.

4.18. $\varphi(t)=\begin{cases}\dfrac{\sin^2 t}{t^2} & \text{dla } t\neq 0\\ 1 & \text{dla } t=0\end{cases}$ (skorzystać ze wzoru Eulera $e^{ix}=\cos x+i\sin x$).

4.19. $1+\varphi(2t)=1+\int\limits_{-\infty}^{\infty}e^{i2tx}f(x)\,dx=\int\limits_{-\infty}^{\infty}(1+\cos 2tx)f(x)\,dx=2\int\limits_{-\infty}^{\infty}\cos^2 tx f(x)\,dx\geqslant$

$\geqslant 2\left[\int\limits_{-\infty}^{\infty}\cos tx f(x)\,dx\right]^2=2[\varphi(t)]^2$ (skorzystano tu z nierówności Buniakowskiego--Schwarza).

5 DWU- I WIELOWYMIAROWE ZMIENNE LOSOWE

5.1. DWUWYMIAROWA ZMIENNA LOSOWA I JEJ ROZKŁAD PRAWDOPODOBIEŃSTWA

5.1.1. Dystrybuanta dwuwymiarowej zmiennej losowej. Przypuśćmy, że badamy pewną zbiorowość ze względu na dwie cechy np. śrubki ze względu na średnicę przekroju i długość, ludzi ze względu na ciężar i wzrost, włókna bawełny – na długość i wytrzymałość, wiek małżonków itp. Zdarzeniu elementarnemu (losowo wybrana śrubka, człowiek, włókno bawełny, małżeństwo) zostały przyporządkowane pary uporządkowanych liczb rzeczywistych. Podobnie będzie przy rzucie dwiema kostkami sześciennymi, gdzie każdemu rzutowi jest przyporządkowana dokładnie jedna para liczb (i, k), gdzie $i, k=1, \ldots, 6$. Te przykładowo podane pary zmiennych będziemy nazywali *dwuwymiarowymi zmiennymi losowymi.*

Niech X oraz Y będą zmiennymi losowymi określonymi niekoniecznie na tej samej przestrzeni probabilistycznej. Parę (X, Y) zmiennych losowych X, Y nazywamy *dwuwymiarową zmienną losową* lub *dwuwymiarowym wektorem losowym*, a X oraz Y jej *współrzędnymi.*

Dystrybuantą dwuwymiarowej zmiennej losowej (X, Y) nazywamy funkcję F zmiennych x, y, która dla każdej pary liczb rzeczywistych $(x, y) \in \boldsymbol{R}^2$ przyjmuje wartości równe prawdopodobieństwu zdarzenia polegającego na tym, że zmienna losowa X przyjmie wartość mniejszą od x i zmienna losowa Y przyjmie wartość mniejszą od y:

$$F(x, y)=P(X<x,\ Y<y) \quad \text{dla} \quad (x, y) \in \boldsymbol{R}^2. \tag{5.1.1}$$

F nazywamy także *dystrybuantą łącznej zmiennej losowej* (X, Y).

Własności dystrybuanty.

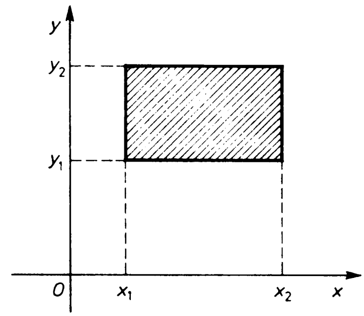
a) $\bigwedge\limits_{x \in \boldsymbol{R}} \lim\limits_{y \to -\infty} F(x, y)=0, \quad \bigwedge\limits_{y \in \boldsymbol{R}} \lim\limits_{x \to -\infty} F(x, y)=0,$

b) $\lim\limits_{\substack{x \to \infty \\ y \to \infty}} F(x, y)=1,$ (5.1.2)

c) Dla dowolnych punktów: (x_1, y_1), (x_2, y_2) takich, że $x_1 \leqslant x_2$ i $y_1 \leqslant y_2$, zachodzi nierówność:

$$F(x_2, y_2)-F(x_2, y_1)-F(x_1, y_2)+F(x_1, y_1) \geqslant 0.$$

Warunek ten okaże się oczywisty, jeśli zauważymy, że lewa strona ostatniej nierówności jest równa $P(x_1 \leqslant X<x_2, y_1 \leqslant Y<y_2)$ (rys. 5.1).



Rys. 5.1. $P(x_1 \leqslant X < x_2, y_1 \leqslant Y < y_2) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geqslant 0$.

d) Dystrybuanta jest funkcją niemalejącą i co najmniej lewostronnie ciągłą względem każdego z argumentów x bądź y.

Każda funkcja dwóch zmiennych spełniająca warunki a) - d) może być traktowana jako dystrybuanta dwuwymiarowej zmiennej losowej (X, Y). W zastosowaniach najczęściej spotykamy dwuwymiarowe zmienne losowe typu skokowego bądź typu ciągłego.

5.1.2. Dwuwymiarowe zmienne losowe typu skokowego. Dwuwymiarową zmienną losową (X, Y), która przyjmuje skończoną, bądź przeliczalną liczbę wartości (x_i, y_k), każdą odpowiednio z prawdopodobieństwem:

$$P(X = x_i, Y = y_k) = p_{ik} \quad \text{dla} \quad i, k \in N, \tag{5.1.3}$$

przy czym $\sum_i \sum_k p_{ik} = 1$, nazywamy *dwuwymiarową zmienną losową skokową* (*dyskretną*) (zad. 5.2). Funkcję (5.1.3), która wartościom (x_i, y_k) przyporządkowuje odpowiednie prawdopodobieństwa p_{ik}, nazywamy *funkcją prawdopodobieństwa dwuwymiarowej zmiennej losowej* (X, Y).

Znając funkcję (5.1.3) można wyznaczyć dystrybuantę dwuwymiarowej zmiennej losowej (X, Y) i odwrotnie. Jeśli dwuwymiarowa zmienna losowa (X, Y) przyjmuje skończoną liczbę wartości, to wygodnie jest umieścić wartości funkcji prawdopodobieństwa (5.1.3) w tabelce dwudzielczej:

	x_i				
y_k	x_1	x_2	...	x_m	$p_{.k}$
y_1	p_{11}	p_{21}	...	p_{m1}	$p_{.1}$
y_2	p_{12}	p_{22}	...	p_{m2}	$p_{.2}$
...	...	...	...	...	...
y_s	p_{1s}	p_{2s}	...	p_{ms}	$p_{.s}$
$p_{i.}$	$p_{1.}$	$p_{2.}$	...	$p_{m.}$	1

Powiemy, że został wyznaczony rozkład prawdopodobieństwa dwuwymiarowej zmiennej losowej skokowej, gdy znana jest jej dystrybuanta albo funkcja prawdopodobieństwa (5.1.3).

ZADANIE 5.1. Z talii 52 kart wylosowano 1 kartę. Niech zmienna losowa X przyjmuje wartości równe liczbie wylosowanych asów, zaś Y – liczbie wylosowanych pików. Wyznaczyć rozkład prawdopodobieństwa dwuwymiarowej zmiennej losowej (X, Y).

Rozwiązanie. Ponieważ losujemy tylko 1 kartę, każda ze zmiennych X, Y może przyjmować z dodatnim prawdopodobieństwem tylko dwie wartości 0 albo 1.

Oznaczmy przez:

A zdarzenie polegające na wylosowaniu asa pik,

B zdarzenie polegające na wylosowaniu karty pikowej, która nie jest asem,

C zdarzenie polegające na wylosowaniu asa, który nie jest pikiem,

D zdarzenie polegające na wylosowaniu karty, która nie jest asem i nie jest pikiem,

$$P(X=0,\ Y=0)=P(D)=\tfrac{36}{52},\quad P(X=1,\ Y=0)=P(C)=\tfrac{3}{52},$$

$$P(X=0,\ Y=1)=P(B)=\tfrac{12}{52},\quad P(X=1,\ Y=1)=P(A)=\tfrac{1}{52}.$$

	x_i	
y_k	0	1
0	$\frac{36}{52}$	$\frac{3}{52}$
1	$\frac{12}{52}$	$\frac{1}{52}$

5.1.3. Rozkłady brzegowe zmiennych losowych typu skokowego. Oznaczmy:

$$p_{i.}=\sum_k p_{ik} \quad \text{dla} \quad i\in N, \tag{5.1.4}$$

$$p_{.k}=\sum_i p_{ik} \quad \text{dla} \quad k\in N. \tag{5.1.5}$$

Zauważmy, że $p_{i.}=P(X=x_i,\ Y=y_1)+P(X=x_i,\ Y=y_2)+\ldots$ jest prawdopodobieństwem tego, że zmienna losowa X przyjmie wartość równą x_i, bez względu na to, którą z wartości: $y_1, y_2, \ldots$ przyjmuje zmienna losowa Y, oraz, że $\sum_i p_{i.}=1$, a więc funkcja:

$$P(X=x_i)=p_{i.},\quad i\in N \tag{5.1.6}$$

wyznacza rozkład prawdopodobieństwa zmiennej losowej X, nazywany rozkładem brzegowym (marginalnym) zmiennej losowej X w rozkładzie dwuwymiarowej zmiennej losowej (X, Y) (jak widać z tabelki – $p_{i.}$ zapisuje się na brzegu tabelki). Podobnie określamy rozkład brzegowy zmiennej losowej Y jako rozkład prawdopodobieństwa określony wzorem:

$$P(Y=y_k)=p_{.k},\quad k\in N. \tag{5.1.7}$$

Oznaczmy przez F_1 i F_2 dystrybuanty rozkładów brzegowych zmiennych losowych X i Y odpowiednio. Jeśli (X, Y) jest dwuwymiarową zmienną losową skokową, to

$$F_1(x) = \sum_{x_i < x} p_{i.} \quad \text{dla} \quad x \in \boldsymbol{R}, \tag{5.1.8}$$

$$F_2(y) = \sum_{y_k < y} p_{.k} \quad \text{dla} \quad y \in \boldsymbol{R}. \tag{5.1.9}$$

ZADANIE 5.2. Pewien mechanizm składa się z dwóch kół zębatych: dużego i małego. Warunki techniczne przy montażu urządzenia zostają naruszone, jeśli w obu kołach występują dodatnie odchylenia grubości zębów od nominalnego wymiaru. Robotnik dysponuje 2 kołami zębatymi dużymi: „plusowym" i „minusowym" i dwoma małymi „plusowym" i „minusowym". Rozważmy zero-jedynkowe zmienne losowe X i Y: zmienna losowa X przyjmuje wartość 1, jeśli robotnik wylosuje duże koło „plusowe" i 0 jeśli duże koło „minusowe". Analogicznie określona jest zmienna losowa Y w przypadku koła małego. a) Wyznaczyć i naszkicować dystrybuantę F dwuwymiarowej zmiennej losowej (X, Y). b) Obliczyć prawdopodobieństwo naruszenia warunków technicznych przy montażu mechanizmu.

Rozwiązanie. a) Dwuwymiarowa zmienna losowa (X, Y) przyjmuje wartości $(0, 0)$, $(1, 0)$, $(0, 1)$, $(1, 1)$ z prawdopodobieństwami:

$$P(X=0,\, Y=0) = P(A \cap B) = P(A)P(B) = \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{4},$$

$$P(X=1,\, Y=0) = P(A' \cap B) = P(A')P(B) = \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{4},$$

$$P(X=0,\, Y=1) = P(A \cap B') = P(A)P(B') = \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{4},$$

$$P(X=1,\, Y=1) = P(A' \cap B') = P(A')P(B') = \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{4},$$

gdzie A jest zdarzeniem polegającym na wylosowaniu dużego koła zębatego „minusowego", B – małego koła „minusowego", zaś A', B' są odpowiednio zdarzeniami przeciwnymi do A i B. Zestawmy otrzymane wyniki w tabelce:

	x_i	
y_k	0	1
0	$\frac{1}{4}$	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$

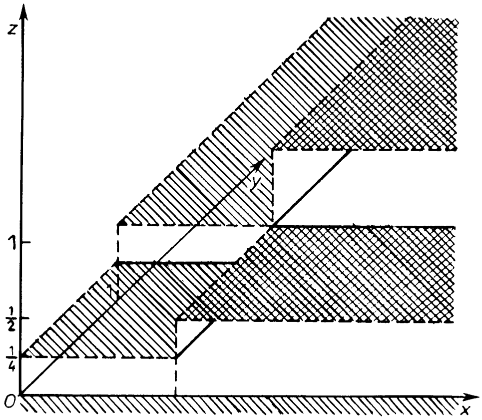
Stosując wzór (5.1.1) wyznaczymy dystrybuantę F dwuwymiarowej zmiennej losowej (X, Y): dla wygody umieśćmy jej wartości w tabelce:

	x		
y	$(-\infty, 0\rangle$	$(0, 1\rangle$	$(1, \infty)$
$(-\infty, 0\rangle$	0	0	0
$(0, 1\rangle$	0	$\frac{1}{4}$	$\frac{1}{2}$
$(1, \infty)$	0	$\frac{1}{2}$	1

Szukana dystrybuanta (rys. 5.2) jest więc określona wzorami:

$$F(x, y) = \begin{cases} 0 & \text{dla} \quad x \leqslant 0 \vee y \leqslant 0, \\ \frac{1}{4} & \text{dla} \quad 0 < x \leqslant 1 \wedge 0 < y \leqslant 1, \\ \frac{1}{2} & \text{dla} \quad x > 1 \wedge 0 < y \leqslant 1, \\ \frac{1}{2} & \text{dla} \quad 0 < x \leqslant 1 \wedge y > 1, \\ 1 & \text{dla} \quad x > 1 \wedge y > 1. \end{cases}$$

b) Naruszenie warunków technicznych montażu nastąpi wtedy, gdy robotnik wybierze losowo duże koło zębate „plusowe" i małe „plusowe"; jest to koniunkcja zdarzeń niezależnych A', B', a więc $P(A' \cap B') = P(A')P(B') = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.



Rys. 5.2. Dystrybuanta dwuwymiarowej zmiennej losowej typu skokowego (zad. 5.2)

ZADANIE 5.3. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład określony w tabelce:

	x_i		
y_k	1	2	3
2	0,1	0,2	0,3
4	0,1	0,1	0,2

Wyznaczyć dystrybuantę rozkładu brzegowego zmiennej losowej Y.

Rozwiązanie. Aby wyznaczyć rozkład brzegowy (5.1.7) zmiennej Y sumujemy prawdopodobieństwa w tabelce dwudzielczej w wierszach i tak:

$$P(Y=2)=0{,}1+0{,}2+0{,}3=0{,}6,$$

$$P(Y=4)=0{,}1+0{,}1+0{,}2=0{,}4.$$

Rozkład prawdopodobieństwa zmiennej Y jest więc rozkładem dwupunktowym:

y_k	2	4
$p_{\cdot k}$	0,6	0,4

Dystrybuanta (5.1.9) tego rozkładu jest określona wzorami:

$$F_2(y)=\begin{cases}0 & \text{dla} \quad y\leqslant 2,\\ 0{,}6 & \text{dla} \quad 2<y\leqslant 4,\\ 1 & \text{dla} \quad y>4.\end{cases}$$

5.1.4. Rozkłady warunkowe zmiennych losowych typu skokowego. Załóżmy, że wszystkie prawdopodobieństwa $p_{.k}$ rozkładu brzegowego zmiennej losowej Y są dodatnie i rozważmy

$$P(X=x_i|Y=y_k)=\frac{p_{ik}}{p_{\cdot k}}. \tag{5.1.10}$$

Zauważmy, że:

1) $\bigwedge\limits_{i,k\in N} 0\leqslant\frac{p_{ik}}{p_{.k}}\leqslant 1$, ponieważ licznik jest jednym ze składników mianownika (5.1.5),

2) $\bigwedge\limits_{k}\sum\limits_{i} P(X=x_i|Y=y_k)=1$, ponieważ

$$\sum_i P(X=x_i|Y=y_k)=\sum_i\frac{p_{ik}}{p_{.k}}=\frac{1}{p_{.k}}\sum_i p_{ik}=\frac{p_{.k}}{p_{.k}}=1.$$

Wzór (5.1.10) określa zatem rozkład prawdopodobieństwa zmiennej losowej X przy ustalonej wartości $Y=y_k$; nazywamy go *rozkładem warunkowym* zmiennej losowej X, pod warunkiem, że Y przyjmuje ustaloną wartość równą y_k. Podobnie można określić rozkład prawdopodobieństwa zmiennej losowej Y przy warunku, że X przyjmuje wartość x_i:

$$P(Y=y_k|X=x_i)=\frac{p_{ik}}{p_{i.}}, \qquad k\in N, \tag{5.1.11}$$

jeśli $\bigwedge\limits_{i} p_{i.}>0$.

Dystrybuanty rozkładów warunkowych oznaczmy odpowiednio przez: $F(x|y_k)$, $F(y|x_i)$. Podobnie jak dla innych rozkładów skokowych można je wyznaczyć według wzorów:

$$F(x|y_k)=P(X<x|Y=y_k)=\sum_{x_i<x}\frac{p_{ik}}{p_{.k}}, \tag{5.1.12}$$

$$F(y|x_i)=P(Y<y|X=x_i)=\sum_{y_k<y}\frac{p_{ik}}{p_{i.}}. \tag{5.1.13}$$

ZADANIE 5.4. Niech dwuwymiarowa zmienna losowa (X, Y) ma rozkład podany w tabelce:

	x_i				
y_k	2	3	3,5	4	5
2	$\frac{3}{35}$	$\frac{2}{35}$	$\frac{1}{35}$	0	0
3	$\frac{1}{35}$	$\frac{10}{35}$	$\frac{2}{35}$	$\frac{1}{35}$	0
3,5	0	$\frac{1}{35}$	$\frac{5}{35}$	$\frac{1}{35}$	$\frac{1}{35}$
4	0	0	0	$\frac{3}{35}$	$\frac{2}{35}$
5	0	0	0	$\frac{1}{35}$	$\frac{1}{35}$

gdzie X jest oceną klasówki z matematyki losowo wybranego ucznia pewnej klasy, Y zaś oceną klasówki z fizyki. Wyznaczyć rozkład warunkowy zmiennej losowej Y pod warunkiem, że $X=4$.

Rozwiązanie. Korzystamy ze wzoru (5.1.4). Sumując w tabelce dwudzielczej prawdopodobieństwa w kolumnie czwartej odpowiadającej $X=4$, otrzymamy: $p_{4.}=P(X=4)=\frac{6}{35}$. Rozkład warunkowy zmiennej losowej Y przy warunku $X=4$ otrzymamy ze wzoru (5.1.11)

$$P(Y=y_k|X=4)=\frac{P(X=4,\, Y=y_k)}{P(X=4)}=\frac{P(X=4,\, Y=y_k)}{\frac{6}{35}},$$

$$P(Y=2|X=4)=0, \quad P(Y=3|X=4)=\tfrac{1}{6}, \quad P(Y=3{,}5|X=4)=\tfrac{1}{6}, \quad P(Y=4|X=4)=\tfrac{3}{6},$$
$$P(Y=5|X=4)=\tfrac{1}{6}.$$

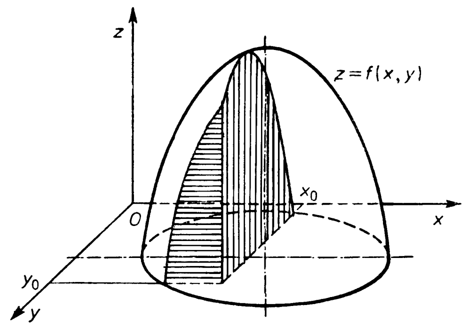
* * *

Innym typem dwuwymiarowych zmiennych losowych są tzw. *zmienne losowe typu ciągłego*, np. wytrzymałość i długość odcinka drutu wybranego losowo z pewnej partii, albo włókna bawełny wylosowanego z określonej beli, wzrost i masa człowieka z pewnej grupy ludzi.

5.1.5. Dwuwymiarowa zmienna losowa typu ciągłego. *Dwuwymiarową zmienną losową* (X, Y) nazywamy *typu ciągłego*, jeśli istnieje nieujemna funkcja f taka, że dystrybuanta tej zmiennej losowej da się przedstawić jako całka

$$F(x,y)=\int_{-\infty}^{x}\Big[\int_{-\infty}^{y} f(u,v)\,dv\Big]du \quad \text{dla} \quad (x,y)\in R^2; \tag{5.1.14}$$

funkcję f nazywamy *gęstością rozkładu prawdopodobieństwa*. Z (5.1.14) i interpretacji całki wynika, że dystrybuantę F w punkcie (x_0, y_0) można traktować jako objętość bryły ograniczonej powierzchnią S o równaniu $z=f(x,y)$, płaszczyznami $x=x_0$, $y=y_0$ i tą częścią płaszczyzny Oxy, dla której $x<x_0$ i $y<y_0$ (rys. 5.3).



Rys. 5.3. Wartość dystrybuanty $F(x_0, y_0)$ na rysunku gęstości f dwuwymiarowej zmiennej losowej (X, Y)

Własności dwuwymiarowej zmiennej losowej typu ciągłego:

1) $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x, y)\,dy\,dx = 1,$ (5.1.15)

2) w punktach (x, y) ciągłości f mamy

$$\frac{\partial^2 F(x, y)}{\partial x\,\partial y} = f(x, y), \tag{5.1.16}$$

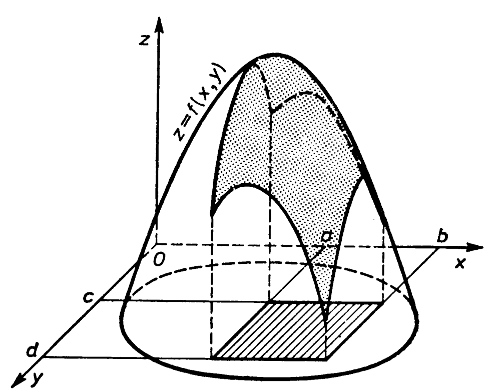
3) dla obszaru regularnego $B \subset \mathbf{R}^2$

$$P[(X, Y) \in B] = \iint\limits_B f(x, y)\,dx\,dy, \tag{5.1.17}$$

w szczególności gdy B jest prostokątem, tzn. $B = \{(x, y): a \leqslant x \leqslant b \wedge c \leqslant y \leqslant d\}$, wtedy

$$P(a \leqslant X \leqslant b \wedge c \leqslant Y \leqslant d) = \int\limits_a^b \Big[\int\limits_c^d f(x, y)\,dy\Big]\,dx \quad \text{(rys. 5.4)}.$$

Powiemy, że dany jest rozkład prawdopodobieństwa dwuwymiarowej zmiennej losowej typu ciągłego, jeśli jest znana jej dystrybuanta lub gęstość.



Rys. 5.4. $P(a \leqslant X \leqslant b \wedge c \leqslant Y \leqslant d)$ na rysunku gęstości dwuwymiarowej zmiennej losowej (X, Y)

ZADANIE 5.5. Dobrać tak stałą c, by funkcja

$$f(x, y) = \begin{cases} cxy & \text{dla} \quad 0 \leqslant x \leqslant 2 \wedge 0 \leqslant y \leqslant 1, \\ 0 & \text{dla pozostałych } (x, y) \end{cases}$$

była gęstością dwuwymiarowej zmiennej losowej (X, Y) oraz wyznaczyć jej dystrybuantę.

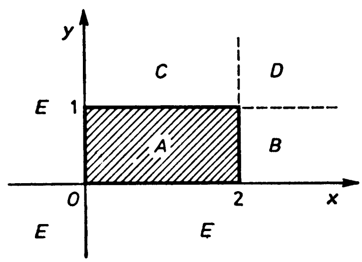
Rozwiązanie. Stałą c należy tak dobrać, aby

$$c\int_0^2\left(\int_0^1 xy\,dy\right)dx=1,$$

skąd $c=1$. Gęstość f dwuwymiarowej zmiennej losowej (X, Y) jest więc postaci:

$$f(x,y)=\begin{cases} xy & \text{dla} \quad 0\leqslant x\leqslant 2\wedge 0\leqslant y\leqslant 1, \\ 0 & \text{dla} \quad \text{pozostałych } (x,y). \end{cases}$$

Podzielmy płaszczyznę Oxy prostymi: $x=0$, $x=2$, $y=0$, $y=1$ na następujące obszary (rys. 5.5), w których korzystając z (5.1.14) wyznaczymy dystrybuantę.



Rys. 5.5. Do zadania 5.5

1) $(x,y)\in A$

$$F(x,y)=\int_0^x\left(\int_0^y uv\,dv\right)du=\tfrac{1}{4}x^2y^2,$$

2) $(x,y)\in B$

$$F(x,y)=\int_0^2\left(\int_0^y uv\,dv\right)du=\int_0^2 u\,du\int_0^y v\,dv=\tfrac{1}{2}\cdot 2^2\cdot\tfrac{1}{2}y^2=y^2,$$

3) $(x,y)\in C$

$$F(x,y)=\int_0^x\left(\int_0^1 uv\,dv\right)du=\tfrac{1}{4}x^2,$$

4) $(x,y)\in D$

$$F(x,y)=\int_0^2\left(\int_0^1 uv\,dv\right)du=1.$$

Dla pozostałych punktów $(x,y)\in E$ jest $F(x,y)=0$, a więc dystrybuanta jest określona wzorami:

$$F(x,y)=\begin{cases} 0 & \text{dla obszaru} \quad E=\{(x,y)\colon x\leqslant 0\vee y\leqslant 0\}, \\ \tfrac{1}{4}x^2y^2 & \text{dla obszaru} \quad A=\{(x,y)\colon 0<x\leqslant 2\wedge 0<y\leqslant 1\}, \\ y^2 & \text{dla obszaru} \quad B=\{(x,y)\colon x>2\wedge 0<y\leqslant 1\}, \\ \tfrac{1}{4}x^2 & \text{dla obszaru} \quad C=\{(x,y)\colon 0<x\leqslant 2\wedge y>1\}, \\ 1 & \text{dla obszaru} \quad D=\{(x,y)\colon x>2\wedge y>1\}. \end{cases}$$

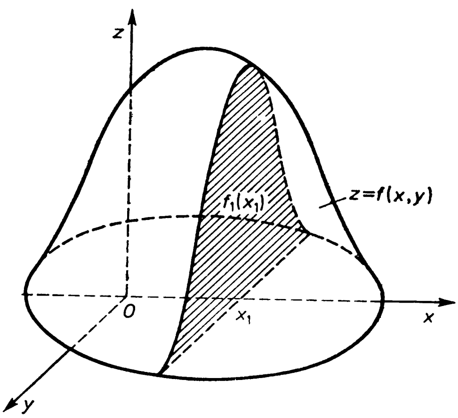
5.1.6. Rozkłady brzegowe zmiennych losowych typu ciągłego. Oznaczmy

$$f_1(x)=\int_{-\infty}^{\infty} f(x,y)\,dy. \tag{5.1.18}$$

Zauważmy, że dla każdego $x\in \boldsymbol{R}$ mamy $f_1(x)\geqslant 0$ oraz

$$\int_{-\infty}^{\infty} f_1(x)\,dx=\int_{-\infty}^{\infty}\Big[\int_{-\infty}^{\infty} f(x,y)\,dy\Big]dx=1.$$

Funkcja f_1 określona wzorem (5.1.18) jest więc gęstością zmiennej losowej X. Rozkład prawdopodobieństwa wyznaczony przez funkcję (5.1.18) nazywamy *rozkładem brzegowym* zmiennej losowej X w rozkładzie dwuwymiarowej zmiennej losowej (X, Y). Jak to wynika ze wzoru (5.1.18) dla każdego ustalonego $x=x_1$ wartość $f_1(x_1)$ jest równa polu obszaru D leżącego w płaszczyźnie $x=x_1$ określonego następująco: $D=\{(y,z): 0\leqslant z\leqslant f(x_1,y)\wedge \wedge y\in \boldsymbol{R}\}$, co ilustruje rysunek 5.6.



Rys. 5.6. Wartość $f_1(x_1)$ gęstości rozkładu brzegowego zmiennej X na rysunku gęstości f dwuwymiarowej zmiennej losowej (X, Y)

Dystrybuantę rozkładu brzegowego zmiennej losowej X (ciągłej) w rozkładzie dwuwymiarowej zmiennej losowej (X, Y) wyznacza wzór:

$$F_1(x)=\int_{-\infty}^{x}\Big[\int_{-\infty}^{\infty} f(u,y)\,dy\Big]du=\int_{-\infty}^{x} f_1(u)\,du \quad \text{dla} \quad x\in\boldsymbol{R}. \tag{5.1.19}$$

W podobny sposób możemy określić rozkład brzegowy zmiennej losowej Y, którego gęstością jest funkcja f_2 wyznaczona następująco:

$$f_2(y)=\int_{-\infty}^{\infty} f(x,y)\,dx, \tag{5.1.20}$$

a dystrybuantą funkcja F_2, przy czym:

$$F_2(y)=\int_{-\infty}^{y}\Big[\int_{-\infty}^{\infty} f(x,v)\,dx\Big]dv=\int_{-\infty}^{y} f_2(v)\,dv \quad \text{dla} \quad y\in\boldsymbol{R}. \tag{5.1.21}$$

Można podać także związek między dystrybuantami rozkładów brzegowych, a dystrybuantą F dwuwymiarowej zmiennej losowej (X, Y), a mianowicie:

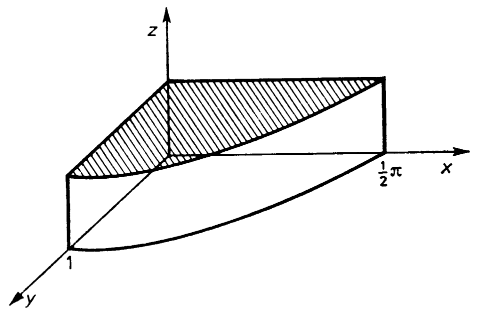
$$F_1(x) = \lim_{y\to\infty} F(x, y),$$
$$F_2(y) = \lim_{x\to\infty} F(x, y). \tag{5.1.22}$$

ZADANIE 5.6. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład równomierny w obszarze $D=\{(x, y): 0 \leqslant x \leqslant \frac{1}{2}\pi \wedge 0 \leqslant y \leqslant \cos x\}$. Wyznaczyć gęstości rozkładów brzegowych.

Rozwiązanie. Gęstość dwuwymiarowego rozkładu równomiernego – podobnie jak w przypadku jednowymiarowym – jest stałą różną od zera w obszarze D. Wobec (5.1.15) stała ta jest równa odwrotności pola obszaru D, w naszym przypadku: $|D| = \int\limits_0^{\pi/2} \cos x\, dx = 1$, stąd gęstość dwuwymiarowej zmiennej losowej (X, Y) jest określona wzorem:

$$f(x, y) = \begin{cases} 1 & \text{dla } (x, y) \in D, \\ 0 & \text{dla pozostałych } (x, y). \end{cases}$$

Jej wykresem (rys. 5.7) jest ta część płaszczyzny $z=1$, której rzutem na płaszczyznę Oxy



Rys. 5.7. Gęstość dwuwymiarowej zmiennej losowej (X, Y) z zadania 5.6

jest obszar D. Korzystając z (5.1.18), otrzymujemy gęstość brzegową zmiennej losowej X postaci:

$$f_1(x) = \begin{cases} \int\limits_0^{\cos x} 1\, dy = \cos x & \text{dla } 0 \leqslant x \leqslant \frac{1}{2}\pi, \\ 0 & \text{dla pozostałych } x. \end{cases}$$

W celu wyznaczenia gęstości brzegowej zmiennej losowej Y zauważmy, że obszar $D= =\{(x, y): 0 \leqslant y \leqslant 1 \wedge 0 \leqslant x \leqslant \arccos y\}$, a więc

$$f_2(y) = \begin{cases} \int\limits_0^{\arccos y} 1\, dx = \arccos y & \text{dla } 0 \leqslant y \leqslant 1, \\ 0 & \text{dla pozostałych } y. \end{cases}$$

5.1.7. Rozkłady warunkowe zmiennych losowych typu ciągłego. Załóżmy, że gęstość brzegowa f_2 jest skoncentrowana na przedziale (c, d) (może to być przedział niewłaściwy), oznaczmy:

$$f(x\,|\,y)=\frac{f(x,y)}{f_2(y)} \quad \text{dla} \quad x\in \boldsymbol{R}\wedge c<y<d. \tag{5.1.23}$$

Funkcja $f(x\,|\,y)$ jest dla $x\in \boldsymbol{R}\wedge c<y<d$ nieujemna oraz spełnia warunek:

$$\int_{-\infty}^{\infty} f(x\,|\,y)\,dx=\frac{1}{f_2(y)}\int_{-\infty}^{\infty} f(x,y)\,dx=\frac{1}{f_2(y)}f_2(y)=1,$$

jest więc gęstością zmiennej losowej X, nazywamy ją *gęstością warunkową* zmiennej losowej X przy warunku, że zmienna Y przyjęła wartość równą y. Dystrybuantę rozkładu warunkowego (5.1.23) wyznaczamy według wzoru:

$$F(x\,|\,y)=\int_{-\infty}^{x} f(u\,|\,y)\,du=\frac{1}{f_2(y)}\int_{-\infty}^{x} f(u,y)\,du \tag{5.1.24}$$

$$\text{dla} \quad x\in \boldsymbol{R}\wedge c<y<d.$$

Podobnie określamy *rozkład warunkowy* zmiennej losowej Y przy warunku $X=x$; *gęstość* tego rozkładu wyrazi się wzorem:

$$f(y\,|\,x)=\frac{f(x,y)}{f_1(x)} \quad \text{dla} \quad y\in \boldsymbol{R}\wedge a<x<b, \tag{5.1.25}$$

jeśli gęstość f_1 rozkładu brzegowego zmiennej losowej X jest skoncentrowana na przedziale (a,b). *Dystrybuanta* tego rozkładu jest określona wzorem:

$$F(y\,|\,x)=\int_{-\infty}^{y} f(v\,|\,x)\,dv=\frac{1}{f_1(x)}\int_{-\infty}^{y} f(x,v)\,dv \quad \text{dla} \quad y\in \boldsymbol{R}\wedge a<x<b. \tag{5.1.26}$$

ZADANIE 5.7. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład o gęstości

$$f(x,y)=\begin{cases}\frac{1}{9}(x+y+2) & \text{dla} \quad 0<x<2\wedge 1<y<2,\\ 0 & \text{dla pozostałych } (x,y).\end{cases}$$

Wyznaczyć gęstość i dystrybuantę rozkładu warunkowego zmiennej losowej Y przy warunku, że $X=x$.

Rozwiązanie. Poszukiwaną gęstość określa wzór (5.1.25), należy więc wyznaczyć najpierw gęstość brzegową zmiennej losowej X:

$$f_1(x)=\begin{cases}\frac{1}{9}\int_1^2 (x+y+2)\,dy=\frac{1}{9}[(x+2)y+\frac{1}{2}y^2]_1^2=\frac{1}{9}(x+\frac{7}{2}) & \text{dla} \quad 0<x<2,\\ 0 & \text{dla pozostałych } x\end{cases}$$

i dla ustalonej wartości $x\in(0,2)$ otrzymamy gęstość:

$$f(y|x)=\begin{cases}\dfrac{x+y+2}{x+3{,}5} & \text{dla}\quad 1<y<2,\\ 0 & \text{dla pozostałych } y\end{cases}$$

oraz dystrybuantę rozkładu warunkowego

$$F(y|x)=\begin{cases}0 & \text{dla}\quad y\leqslant 1,\\ \displaystyle\int_1^y \frac{x+v+2}{x+3{,}5}\,dv & \text{dla}\quad 1<y\leqslant 2,\\ 1 & \text{dla}\quad y>2.\end{cases}$$

Ponieważ $\displaystyle\int_1^y \frac{x+2+v}{x+3{,}5}\,dv=\frac{(x+2)(y-1)+\dfrac{y^2-1}{2}}{x+3{,}5}$, więc otrzymujemy

$$F(y|x)=\begin{cases}0 & \text{dla}\quad y\leqslant 1,\\ \dfrac{(y-1)(2x+y+5)}{2x+7} & \text{dla}\quad 1<y\leqslant 2,\\ 1 & \text{dla}\quad y>2.\end{cases}$$

5.1.8. Niezależność zmiennych losowych. Jednym z ważniejszych pojęć rachunku prawdopodobieństwa jest pojęcie niezależności zmiennych losowych.

Zmienne $X=X(\omega)$ i $Y=Y(\omega)$ nazywamy *niezależnymi zmiennymi losowymi*, jeśli dla dowolnych([1]) zbiorów A i B na prostej zdarzenia losowe $Z_1=\{\omega: X(\omega)\in A\}$ i $Z_2=\{\omega: Y(\omega)\in B\}$ są niezależne (p. 1.3), tzn.

$$P[\{\omega: X(\omega)\in A,\ Y(\omega)\in B\}]=P[\{\omega: X(\omega)\in A\}]\,P[\{\omega: Y(\omega)\in B\}].$$

W szczególnym przypadku przyjmując $A=(-\infty, x)$, $B=(-\infty, y)$, gdzie $x, y\in \boldsymbol{R}$, mamy:

Warunkiem koniecznym i wystarczającym na to, by X i Y były niezależnymi zmiennymi losowymi jest, by dla każdego $(x, y)\in \boldsymbol{R}^2$ dystrybuanta F dwuwymiarowej zmiennej losowej (X, Y) była iloczynem dystrybuant rozkładów brzegowych F_1 i F_2:

$$F(x, y)=F_1(x)F_2(y). \tag{5.1.27}$$

Dla zmiennych losowych typu ciągłego o znanych gęstościach wygodniej jest korzystać z twierdzenia:

Warunkiem koniecznym i wystarczającym niezależności zmiennych losowych X, Y o gęstościach odpowiednio równych f_1, f_2 jest zachodzenie równości:

$$f(x, y)=f_1(x)f_2(y)\quad dla\quad x, y\in \boldsymbol{R}, \tag{5.1.28}$$

gdzie f jest gęstością dwuwymiarowej zmiennej losowej (X, Y).

([1]) A, B są zbiorami borelowskimi na prostej.

Zmienne losowe X i Y typu skokowego są niezależne wtedy i tylko wtedy, gdy

$$P(X=x_i, Y=y_k)=P(X=x_i)P(Y=y_k) \quad \text{dla} \quad i,k\in N, \tag{5.1.29}$$

co zapisujemy również w postaci:

$$p_{ik}=p_{i.}p_{.k} \quad \text{dla} \quad i,k\in N,$$

gdzie wszystkie trzy funkcje p z (5.1.29) są odpowiednio funkcjami prawdopodobieństwa dwuwymiarowej zmiennej losowej i jej rozkładów brzegowych (5.1.4, 5.1.5). Zauważmy, że z (5.1.27), (5.1.29) i określenia rozkładów warunkowych wynika, że:

X i Y są niezależnymi zmiennymi losowymi wtedy i tylko wtedy, gdy rozkłady warunkowe są równe odpowiednim rozkładom brzegowym, co za pomocą np. dystrybuant zapisujemy następująco:

$$F(y\,|\,x)=F_2(y), \qquad F(x\,|\,y)=F_1(x).$$

5.1.9. Zadania rozwiązane.

ZADANIE 5.8. Rzucamy jedną kostką do gry. Niech zmienna losowa X przyjmuje wartość 0, gdy wyrzucimy parzystą liczbę oczek (zdarzenie A), oraz wartość 1, gdy wyrzucimy nieparzystą liczbę oczek (A'), zmienna zaś losowa Y przyjmuje wartość 1, gdy liczba wyrzuconych oczek jest podzielna przez 3 (B), oraz wartość 2, gdy liczba oczek nie jest podzielna przez 3 (B'). Zbadać niezależność zmiennych losowych X i Y.

Rozwiązanie. Dwuwymiarowa zmienna losowa (X, Y) przyjmuje wartości: (0, 1), (0, 2), (1, 1), (1, 2) z prawdopodobieństwami:

$$P(X=0, Y=1)=P(A\cap B)=\tfrac{1}{6}, \qquad P(X=0, Y=2)=P(A\cap B')=\tfrac{2}{6}=\tfrac{1}{3},$$

$$P(X=1, Y=1)=P(A'\cap B)=\tfrac{1}{6}, \qquad P(X=1, Y=2)=P(A'\cap B')=\tfrac{2}{6}=\tfrac{1}{3}.$$

Zestawmy otrzymane wyniki w tabelce dwudzielczej i wyznaczmy rozkłady brzegowe poszczególnych zmiennych.

y_k \ x_i	0	1	$p_{.k}$
1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
$p_{i.}$	$\frac{1}{2}$	$\frac{1}{2}$	1

Zauważmy, że dla $\bigwedge\limits_{\substack{i=1,2\\k=1,2}} p_{ik}=p_{i.}p_{.k}$, a więc X i Y są niezależnymi zmiennymi losowymi.

ZADANIE 5.9. Funkcja zmiennych x, y wyraża się wzorem:

$$f(x,y)=\frac{1}{2\pi}\exp[-\tfrac{1}{2}(x^2+y^2)]+ak(x,y),$$

gdzie

$$k(x,y)=\begin{cases} x^{2r+1}y^{2s+1} & \text{dla } |x|\leqslant 1 \text{ oraz } |y|\leqslant 1, \\ 0 & \text{poza tym,}\end{cases}$$

r, s — dowolnymi liczbami naturalnymi albo zerami, oraz stała a — tak dobrana, że $\bigwedge_{(x,y)\in R^2} f(x,y)\geqslant 0$.

a) Czy tak określona funkcja jest gęstością?

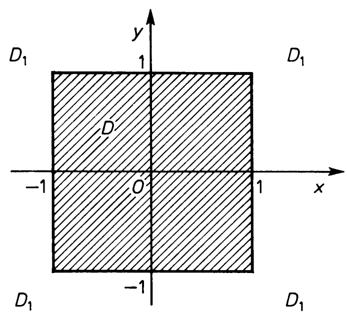
b) Wyznaczyć gęstości rozkładów brzegowych: f_1, f_2.

c) Zbadać czy zmienne losowe X, Y są niezależne.

Rozwiązanie. a) Z danych otrzymujemy

$$f(x,y)=\begin{cases} \dfrac{1}{2\pi}\exp[-\tfrac{1}{2}(x^2+y^2)]+ax^{2r+1}y^{2s+1}\geqslant 0 & \text{dla } |x|\leqslant 1\wedge|y|\leqslant 1, \\ \dfrac{1}{2\pi}\exp[-\tfrac{1}{2}(x^2+y^2)]\geqslant 0 & \text{poza tym.}\end{cases}$$

Oznaczmy: $D=\{(x,y):|x|\leqslant 1\wedge|y|\leqslant 1\}$ oraz $D_1=\{(x,y):|x|>1\vee|y|>1\}$ (rys. 5.8).



Rys. 5.8. Do zadania 5.9

Sprawdźmy, czy jest spełniony warunek (5.1.15),

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy=\frac{1}{2\pi}\iint_D \exp\left[-\frac{1}{2}(x^2+y^2)\right]dx\,dy+a\iint_D x^{2r+1}y^{2s+1}dx\,dy+$$

$$+\frac{1}{2\pi}\iint_{D_1}\exp\left[-\frac{1}{2}(x^2+y^2)\right]dx\,dy=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}(x^2+y^2)\right]dx\,dy=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}x^2\right]dx\,\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}y^2\right]dy=1,$$

ponieważ środkowa całka w środkowej równości, jako iloczyn całek funkcji nieparzystych $\int_{-1}^{1} x^{2r+1}\,dx \cdot \int_{-1}^{1} y^{2s+1}\,dy$ w przedziałach symetrycznych względem zera, jest równa zeru.

b) Korzystając ze wzoru (5.1.18), otrzymujemy:

– dla $|x|\leqslant 1$

$$f_1(x)=\frac{1}{2\pi}\int_{-1}^{1}\exp\left[-\frac{1}{2}(x^2+y^2)\right]dy+a\int_{-1}^{1}x^{2r+1}y^{2s+1}dy+$$

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

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

– dla $|x|>1$

$$f_1(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}(x^2+y^2)\right]dy=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right).$$

Reasumując, $\bigwedge_{x\in R} f_1(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right)$, otrzymaliśmy więc gęstość rozkładu $N(0, 1)$.

W podobny sposób postępując (całkując względem x), otrzymamy gęstość $f_2(y)=$ $=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}y^2\right)$. Mamy więc przykład rozkładu dwuwymiarowej zmiennej losowej różnego od rozkładu normalnego (p. 5.4), którego rozkłady brzegowe są jednowymiarowymi rozkładami normalnymi.

c) Zmienne losowe X, Y są zależne, ponieważ nie dla wszystkich punktów (x, y) zachodzi równość $f(x, y)=f_1(x)f_2(y)$. Równość ta nie zachodzi dla punktów obszaru D.

5.2. FUNKCJE DWUWYMIAROWEJ ZMIENNEJ LOSOWEJ

5.2.1. Gęstość prawdopodobieństwa funkcji zmiennej losowej. Niech

$$u=u(x, y), \quad v=v(x, y) \tag{5.2.1}$$

będą funkcjami ciągłymi i różnowartościowymi o ciągłych pochodnych cząstkowych w pewnym regularnym obszarze płaskim D.

Załóżmy, że (5.2.1) – traktowane jako układ równań – ma dokładnie jedno rozwiązanie

$$x=x(u, v), \quad y=y(u, v) \tag{5.2.2}$$

o ciągłych pochodnych cząstkowych oraz że jakobian

$$\frac{D(x,y)}{D(u,v)}=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \neq 0 \quad \text{dla} \quad (u,v)\in\Delta \tag{5.2.3}$$

w obszarze płaskim Δ płaszczyzny OUV, na który wzory (5.2.1) przekształcają obszar D.

Jeśli gęstość f dwuwymiarowej zmiennej losowej (X, Y) jest ciągła w D poza skończoną liczbą linii (pozostając ograniczoną), to gęstość k dwuwymiarowej zmiennej losowej (U, V) określonej wzorami

$$U=u(X,Y), \qquad V=v(X,Y) \tag{5.2.4}$$

wyraża się wzorem:

$$k(u,v)=\begin{cases} f[x(u,v),y(u,v)]\left|\frac{D(x,y)}{D(u,v)}\right| & \text{dla} \quad (u,v)\in\Delta, \\ 0 & \text{dla} \quad (u,v)\notin\Delta. \end{cases} \tag{5.2.5}$$

5.2.2. Gęstości sumy, iloczynu, ilorazu zmiennych losowych. Gęstość prawdopodobieństwa k_1 zmiennej losowej U (w dwuwymiarowej zmiennej losowej (U, V)), będącej funkcją zmiennych losowych X, Y postaci:

$$U=U(X,Y) \tag{5.2.6}$$

jest *gęstością rozkładu brzegowego dwuwymiarowej zmiennej losowej* (U, V), a więc funkcją określoną wzorem:

$$k_1(u)=\int_{-\infty}^{\infty} k(u,v)\,dv. \tag{5.2.7}$$

Rozważmy następnie szczególne przypadki funkcji (5.2.6):

a) $U=X+Y$; wówczas gęstość k_1 sumy zmiennych losowych X, Y jest postaci:

$$k_1(u)=\int_{-\infty}^{\infty} f(x,u-x)\,dx, \tag{5.2.8}$$

w przypadku gdy X, Y są niezależnymi zmiennymi losowymi o gęstościach odpowiednio f_1, f_2, wtedy

$$k_1(u)=\int_{-\infty}^{\infty} f_1(x)f_2(u-x)\,dx. \tag{5.2.9}$$

Funkcję k_1 określoną wzorem (5.2.9) nazywamy *splotem* (*kompozycją*) funkcji f_1 i f_2 i oznaczamy symbolem $f_1 * f_2$.

b) Gdy $U=XY$, wtedy gęstość k_1 iloczynu zmiennych losowych X, Y jest postaci

$$k_1(u)=\int_{-\infty}^{\infty} f\left(x,\frac{u}{x}\right)\frac{1}{|x|}\,dx. \tag{5.2.10}$$

W przypadku niezależnych zmiennych losowych X, Y:

$$k_1(u)=\int_{-\infty}^{\infty} f_1(x) f_2\left(\frac{u}{x}\right)\frac{1}{|x|}\,dx. \tag{5.2.11}$$

c) $U=\frac{X}{Y}$; wówczas gęstość k_1 ilorazu zmiennych losowych X, Y jest określona wzorem:

$$k_1(u)=\int_{-\infty}^{\infty} f(uy, y)\,|y|\,dy \tag{5.2.12}$$

oraz w przypadku niezależnych zmiennych losowych X, Y:

$$k_1(u)=\int_{-\infty}^{\infty} f_1(uy) f_2(y)\,|y|\,dy. \tag{5.2.13}$$

5.2.3. Bezpośrednie wyznaczenie dystrybuanty funkcji zmiennych losowych. W wielu przypadkach zamiast gęstości wygodniej jest wyznaczyć dystrybuantę K_1 zmiennej losowej $U=g(X, Y)$ ((5.2.6)). Z definicji dystrybuanty mamy:

$$K_1(u)=P(U<u)=P[g(X, Y)<u],$$

gdzie punkty (x, y) należą do płaskiego zbioru określonego nierównością $g(x, y)<u$, skąd

$$K_1(u)=\begin{cases} \iint\limits_D f(x, y)\,dx\,dy & \text{w przypadku gdy } (X, Y) \text{ jest typu ciągłego o gęstości } f, \\ \sum\limits_{(x_i, y_k)\in D_1} P(X=x_i, Y=y_k) & \text{w przypadku dwuwymiarowej zmiennej losowej } (X, Y) \text{ typu skokowego o funkcji prawdopodobieństwa (5.1.3),} \end{cases} \tag{5.2.14}$$

gdzie $D=\{(x, y)\colon g(x, y)<u\}$, zaś $D_1=\{(x_i, y_k)\colon g(x_i, y_k)<u,\ i,\ k\in N\}$.

5.2.4. Rozkład Snedecora. Niech X i Y będą niezależnymi zmiennymi losowymi o rozkładzie χ^2((2.8.16)) odpowiednio z n i m stopniami swobody.

Rozkład zmiennej losowej F określonej wzorem:

$$F=\frac{\frac{1}{n}X}{\frac{1}{m}Y} \tag{5.2.15}$$

nazywamy *rozkładem Snedecora* o parze: (n, m) stopni swobody.

Korzystając z (2.8.1) i (5.2.13), można wykazać, że gęstość k tego rozkładu jest postaci:

$$k(x)=\begin{cases} \dfrac{\Gamma[\frac{1}{2}(n+m)]}{\Gamma(\frac{1}{2}n)\Gamma(\frac{1}{2}m)}\left(\dfrac{m}{n}\right)^{m/2}\dfrac{x^{n/2-1}}{\left(x+\dfrac{m}{n}\right)^{(n+m)/2}} & \text{dla} \quad x>0, \\ 0 & \text{dla pozostałych } x. \end{cases} \tag{5.2.16}$$

Jeśli $X_1, \ldots, X_n$, $Y_1, \ldots, Y_m$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie $N(\mu, \sigma)$, to zmienna losowa

$$F = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}{\frac{1}{m-1}\sum_{i=1}^{m}(Y_i-\bar{Y})^2}, \quad \text{gdzie} \quad \begin{aligned} \bar{X} &= \frac{1}{n}\sum_{i=1}^{n} X_i, \\ \bar{Y} &= \frac{1}{m}\sum_{i=1}^{m} Y_i, \end{aligned} \tag{5.2.17}$$

ma rozkład Snedecora z $(n-1,\ m-1)$ stopniami swobody.

5.2.5. Zadania rozwiązane.

ZADANIE 5.10. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład normalny o gęstości

$$f(x, y) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma^2}+\frac{y^2}{\sigma^2}\right)\right].$$

Wyznaczyć gęstość k dwuwymiarowej zmiennej losowej (U, V), jeśli $U=X+Y$, $V=X-Y$.

Rozwiązanie. Dla każdego $(u, v) \in \mathbf{R}^2$ układ równań:

$$u = x+y, \qquad v = x-y$$

ma dokładnie jedno rozwiązanie: $x=\frac{u+v}{2}$, $y=\frac{u-v}{2}$. Jakobian (5.2.3) jest równy

$$\frac{D(x, y)}{D(u, v)} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\tfrac{1}{2} \neq 0.$$

Gęstość k dwuwymiarowej zmiennej losowej (U, V) zgodnie z (5.2.5) jest więc postaci

$$k(u, v) = \frac{1}{4\pi\sigma^2}\exp\left[-\frac{1}{2}\left(\frac{(u+v)^2}{4\sigma^2}+\frac{(u-v)^2}{4\sigma^2}\right)\right] \quad \text{dla} \quad (u, v) \in \mathbf{R}^2,$$

a po przekształceniu:

$$k(u, v) = \frac{1}{4\pi\sigma^2}\exp\left[-\frac{1}{2}\,\frac{u^2+v^2}{2\sigma^2}\right] \quad \text{dla} \quad (u, v) \in \mathbf{R}^2.$$

ZADANIE 5.11. Niech X, Y będą niezależnymi zmiennymi losowymi o tym samym rozkładzie wykładniczym (2.8.7). Wyznaczyć gęstość ilorazu tych zmiennych losowych.

Rozwiązanie. Gęstości f_1, f_2 zmiennych losowych X, Y są określone wzorami:

$$f_1(x) = \begin{cases} \frac{1}{\lambda}\exp\left(-\frac{x}{\lambda}\right) & \text{dla} \quad x > 0, \\ 0 & \text{dla} \quad \text{pozostałych}\, x, \end{cases}$$

$$f_2(y) = \begin{cases} \frac{1}{\lambda}\exp\left(-\frac{y}{\lambda}\right) & \text{dla} \quad y > 0, \\ 0 & \text{dla pozostałych } y. \end{cases}$$

Gęstość k_1 zmiennej losowej $U=\frac{X}{Y}$ wyznaczamy ze wzoru (5.2.13)

$$k_1(u)=\begin{cases}\int\limits_0^\infty \frac{1}{\lambda}\exp\left(-\frac{uy}{\lambda}\right)\frac{1}{\lambda}\exp\left(-\frac{y}{\lambda}\right)y\,dy & \text{dla} \quad u>0,\\ 0 & \text{dla pozostałych } u.\end{cases}$$

Obliczając całkę $\frac{1}{\lambda^2}\int\limits_0^\infty y\exp\left(-y\frac{u+1}{\lambda}\right)dy=\frac{\lambda^2\Gamma(2)}{(u+1)^2\lambda^2}$, otrzymujemy:

$$k_1(u)=\begin{cases}\frac{1}{(u+1)^2} & \text{dla} \quad u>0,\\ 0 & \text{dla} \quad \text{pozostałych } u.\end{cases}$$

Zadanie 5.12. Dwie osoby z miasta A usiłują na zmianę, co trzy minuty, uzyskać automatyczne połączenie telefoniczne z miastem B, jednak każda z nich nie więcej, niż trzykrotnie. Wyznaczyć funkcję rozkładu prawdopodobieństwa zmiennej losowej: a) X – liczba prób osoby rozpoczynającej, b) Y – liczba prób pozostałej osoby, c) $U=X+Y$. Przyjąć, że prawdopodobieństwo p uzyskania połączenia w czasie trzech minut wynosi 0,1.

Rozwiązanie. Niech A_i, $i=1, 2, 3$ oznacza zdarzenie – osoba rozpoczynająca próby uzyska połączenie za i-tym razem. B_i ma analogiczne zr[illegible]zenie dla drugiej osoby, A_i', B_i' – to zdarzenia przeciwne odpowiednio do A_i, B_i.

a) Mamy $W_x=\{1, 2, 3\}$, przy tym (oznaczając $1-p=q$)

$$p_1'=P(X=1)=P(A_1\cup A_1'B_1)=p+qp=1-q^2=0{,}1900,$$

$$p_2'=P(X=2)=P[A_1'B_1'(A_2\cup A_2'B_2)]=q^2(p+qp)=q^2(1-q^2)=$$
$$=0{,}1539,$$

$$p_3'=P(X=3)=P(A_1'B_1'A_2'B_2')=q^4=0{,}6561.$$

Warunek unormowania mogliśmy wykorzystać dla obliczenia jednego z prawdopodobieństw p_1', p_2', p_3'. Postąpimy inaczej: posłużmy się nim dla potwierdzenia poprawności obliczeń. Mamy

$$p_1'+p_2'+p_3'=1-q^2+q^2(1-q^2)+q^4=(1-q^2)(1+q^2)+q^4=1.$$

Poszukiwaną funkcję rozkładu prawdopodobieństwa zmiennej losowej X przedstawia tabela

x_i	1	2	3
p_i	0,19	0,1539	0,6561

b) Mamy: $W_y=\{0, 1, 2, 3\}$, przy tym:

$$q_0=P(Y=0)=P(A_1)=p=0{,}1,$$

$$q_1=P(Y=1)=P(A_1'B_1\cup A_1'B_1'A_2)=qp+q^2p=q(1-q^2)=0{,}171,$$

$$q_2=P(Y=2)=P(A_1'B_1'A_2'B_2 \cup A_1'B_1'A_2'B_2'A_3)=q^3p+q^4p=$$
$$=pq^3(1+q)=q^3(1-q^2)=0{,}13851,$$
$$q_3=P(Y=3)=P(A_1'B_1'A_2'B_2'A_3')=q^5=0{,}59049.$$

Sprawdzenie warunku unormowania i tym samym potwierdzenia poprawności obliczeń pozostawiamy Czytelnikowi. A oto funkcja rozkładu prawdopodobieństwa zmiennej losowej Y

y_j	0	1	2	3
p_j'	0,1	0,171	0,13851	0,59049

c) Punkty skokowe u_l zmiennej losowej $U=X+Y$, to oczywiście liczby postaci x_r+y_s przy wszystkich możliwych parach (x_r, y_s), gdzie $x_r \in W_x$, $y_s \in W_y$. Zatem $W_u=\{1, 2, 3, 4, 5, 6\}$.

Zmienna losowa U przyjmuje wartości równe liczbie prób jaką wykonałaby jedna osoba, powiedzmy z miasta A, dla uzyskania połączenia, nie większą jednak niż 6. Zatem:

$$q_1=P(U=1)=P(A_1)=p,$$
$$q_2=P(U=2)=P(A_1'A_2)=qp,$$
$$q_3=P(U=3)=P(A_1'A_2'A_3)=q^2p,$$
$$q_4=P(U=4)=P(A_1'A_2'A_3'A_4)=q^3p,$$
$$q_5=P(U=5)=P(A_1'A_2'A_3'A_4'A_5)=q^4p,$$
$$q_6=P(U=6)=P(A_1'A_2'A_3'A_4'A_5')=q^5.$$

Tak więc funkcja rozkładu prawdopodobieństwa sumy $X+Y$ jest następująca

u_l	1	2	3	4	5	6
q_l	0,1	0,09	0,081	0,0729	0,06561	0,59049

Zadanie 5.13. Niech X, Y będą niezależnymi zmiennymi losowymi o tej samej dystrybuancie F_1. Wyznaczyć dystrybuantę zmiennej losowej $Z=\max(X, Y)$.

Rozwiązanie. Oznaczmy przez F dystrybuantę zmiennej losowej Z, mamy wówczas

$$F(z)=P[\max(X, Y)<z]=P(X<z, Y<z),$$

korzystając z niezależności zmiennych losowych X, Y, otrzymujemy

$$F(z)=P(X<z)P(Y<z)=[F_1(z)]^2.$$

Zadanie 5.14. Student jedzie do domu kolejno dwoma autobusami, z których każdy przyjeżdża na przystanek co 10 minut niezależnie jeden od drugiego. Zakładając, że student przychodzi na przystanek w losowo wybranej chwili oraz że czas oczekiwania na poszczególne autobusy ma rozkład jednostajny, obliczyć prawdopodobieństwo tego, że łączny czas oczekiwania na przystankach nie przekracza 17 minut.

Rozwiązanie. Oznaczmy przez X czas oczekiwania na pierwszy autobus, gęstość tej zmiennej losowej jest

$$f_1(x)=\begin{cases}\frac{1}{10} & \text{dla} \quad 0\leqslant x\leqslant 10,\\ 0 & \text{dla pozostałych } x.\end{cases}$$

Czas oczekiwania Y na drugi autobus ma taki sam rozkład jednostajny.

Gęstość f dwuwymiarowej zmiennej losowej (X, Y) wobec niezależności zmiennych X i Y jest postaci:

$$f(x,y)=\begin{cases}\frac{1}{100} & \text{dla} \quad 0\leqslant x\leqslant 10\wedge 0\leqslant y\leqslant 10,\\ 0 & \text{dla pozostałych } (x,y).\end{cases}$$

Łączny czas oczekiwania na przystankach jest zmienną losową $U=X+Y$. Dystrybuantę K_1 zmiennej losowej U wyznaczymy ze wzoru (5.2.14):

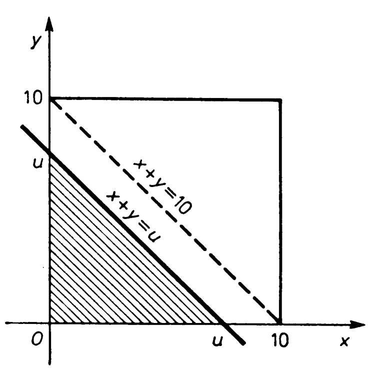
$$K_1(u)=\begin{cases}\iint\limits_D \frac{1}{100}\,dx\,dy & \text{dla} \quad 0\leqslant u\leqslant 20,\\ 0 & \text{dla pozostałych } u,\end{cases}$$

gdzie $D=\{(x,y): x+y<u\wedge 0\leqslant x\leqslant 10\wedge 0\leqslant y\leqslant 10\}$.

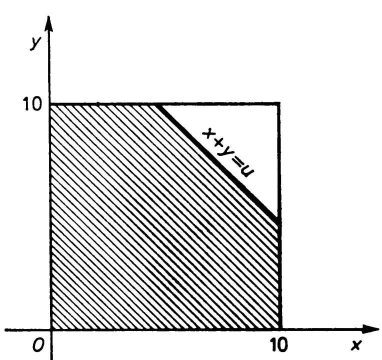
Rozważmy następujące przypadki:

1. $0\leqslant u\leqslant 10$, to:

$$K_1(u)=\iint\limits_{D_1}\frac{1}{100}\,dx\,dy, \quad \text{gdzie} \quad D_1=\{(x,y): 0\leqslant x\leqslant u\wedge 0\leqslant y\leqslant u-x\}$$



Rys. 5.9. Do zadania 5.14



Rys. 5.10. Do zadania 5.14

(rys. 5.9). Stąd

$$K_1(u)=\int_0^u\Big(\int_0^{u-x}\frac{1}{100}\,dy\Big)dx=\frac{1}{200}u^2.$$

2. $10\leqslant u\leqslant 20$, wówczas:

$$K_1(u)=1-\iint\limits_{D_2}\frac{1}{100}\,dx\,dy, \quad \text{gdzie} \quad D_2=\{(x,y): u-10\leqslant x\leqslant 10\wedge u-x\leqslant y\leqslant 10\}$$

(rys. 5.10). Stąd

$$K_1(u)=1-\int_{u-10}^{10}\left(\int_{u-x}^{10}\tfrac{1}{100}\,dy\right)dx=1-\tfrac{1}{100}\int_{u-10}^{10}(10-u+x)\,dx=$$

$$=1-\frac{(20-u)^2}{200}.$$

Zauważmy, że dystrybuantę K_1 w obu przypadkach można również otrzymać jako stosunek pól obszaru zakreskowanego do pola całego kwadratu.

Ostatecznie otrzymujemy więc *dystrybuantę rozkładu Simpsona* (*trójkątnego*) postaci:

$$K_1(u)=\begin{cases}0 & \text{dla} \quad u\leqslant 0,\\ \frac{1}{200}u^2 & \text{dla} \quad 0<u\leqslant 10,\\ 1-\dfrac{(20-u)^2}{200} & \text{dla} \quad 10<u\leqslant 20,\\ 1 & \text{dla} \quad u>20.\end{cases}$$

Szukane prawdopodobieństwo:

$$P(U<17)=K_1(17)=1-\tfrac{9}{200}=\tfrac{191}{200}.$$

5.3. CHARAKTERYSTYKI LICZBOWE DWUWYMIAROWEJ ZMIENNEJ LOSOWEJ

Rozważmy jednowymiarową zmienną losową Z, która jest pewną funkcją zmiennych losowych X, Y (5.2.6). Można wykazać, że wartość przeciętna zmiennej $Z=g(X, Y)$ (o ile istnieje) jest liczbą określoną wzorem:

$$E[g(X,Y)]=\begin{cases}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} g(x,y)f(x,y)\,dx\,dy & \text{dla zmiennej } (X,\ Y) \text{ typu ciągłego.}\\ \sum\limits_i\sum\limits_k g(x_i,y_k)\,p_{ik} & \text{dla } (X,\ Y) \text{ typu skokowego.}\end{cases} \tag{5.3.1}$$

5.3.1. Momenty zwykłe dwuwymiarowej zmiennej losowej. *Momentem zwykłym mieszanym rzędu* $r+s$ *dwuwymiarowej zmiennej losowej* (X, Y) nazywamy wartość przeciętną (o ile istnieje) iloczynu zmiennych losowych X^rY^s. Oznaczamy ją przez

$$\alpha_{rs}=E(X^rY^s) \quad \text{dla} \quad r,s\in N_0. \tag{5.3.2}$$

Dla zmiennej typu ciągłego o gęstości f wzór (5.3.1) przyjmuje postać:

$$\alpha_{rs}=\int\limits_{-\infty}^{\infty}\left[\int\limits_{-\infty}^{\infty} x^ry^sf(x,y)\,dy\right]dx \quad \text{dla} \quad r,s\in N_0,$$

o ile $\int\limits_{-\infty}^{\infty}\left[\int\limits_{-\infty}^{\infty}|x|^r|y|^sf(x,y)\,dy\right]dx<\infty.$ (5.3.3)

W przypadku gdy (X, Y) jest zmienną losową typu skokowego o rozkładzie prawdopodobieństwa (5.1.3), wtedy

$$\alpha_{rs}=\sum_i\sum_k x_i^r y_k^s p_{ik} \quad \text{dla} \quad r, s \in N_0, \tag{5.3.4}$$

o ile $\sum_i\sum_k |x_i|^r |y_k|^s p_{ik} < \infty$.

W szczególnych przypadkach gdy:

a) $s=0$, wzór (5.3.2) określa r-ty moment zwykły (moment zwykły rzędu r) w rozkładzie brzegowym zmiennej losowej X:

$$\alpha_{r0}=EX^r, \tag{5.3.5}$$

dla $r=1$ otrzymujemy z (5.3.5) wartość przeciętną zmiennej losowej X:

$$\alpha_{10}=EX,$$

b) $r=0$, wzór (5.3.2) określa moment zwykły rzędu s w rozkładzie brzegowym zmiennej losowej Y:

$$\alpha_{0s}=EY^s, \tag{5.3.6}$$

dla $s=1$ z (5.3.6) otrzymujemy wartość przeciętną zmiennej losowej Y:

$$\alpha_{01}=EY.$$

Punkt $S(\alpha_{10}, \alpha_{01})$ na płaszczyźnie Oxy nazywamy często *środkiem masy rozkładu prawdopodobieństwa dwuwymiarowej zmiennej losowej* (X, Y).

5.3.2. Momenty centralne dwuwymiarowej zmiennej losowej. *Momentem centralnym mieszanym* μ_{rs} *rzędu* $r+s$, $r, s \in N_0$ dwuwymiarowej zmiennej losowej (X, Y) nazywamy wartość przeciętną iloczynu $(X-EY)^r(Y-EY)^s$ (jeśli istnieje), czyli:

$$\mu_{rs}=E(X-EX)^r(Y-EY)^s \quad \text{dla} \quad r, s \in N_0. \tag{5.3.7}$$

Dla dwuwymiarowej zmiennej losowej typu ciągłego o gęstości f wzór (5.3.7) przyjmuje postać:

$$\mu_{rs}=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}(x-EX)^r(y-EY)^s f(x, y)\,dy\right]dx \quad \text{dla} \quad r, s \in N_0, \tag{5.3.8}$$

o ile $\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}|x-EX|^r|y-EY|^s f(x, y)\,dy\right]dx < \infty$.

Dla zmiennej losowej typu skokowego o rozkładzie prawdopodobieństwa (5.1.3):

$$\mu_{rs}=\sum_i\sum_k (x_i-EX)^r(y_k-EY)^s p_{ik} \quad \text{dla} \quad r, s \in N_0, \tag{5.3.9}$$

jeśli $\sum_i\sum_k |x_i-EX|^r|y_k-EY|^s p_{ik} < \infty$.

Szczególne znaczenie mają momenty centralne rzędu drugiego:

a) gdy $r=2$ i $s=0$, wtedy $\mu_{20}=E(X-EX)^2$ jest wariancją w rozkładzie brzegowym zmiennej losowej X.

b) gdy $r=0$ i $s=2$, wtedy $\mu_{02}=E(Y-EY)^2$ jest wariancją w rozkładzie brzegowym zmiennej losowej Y.

c) gdy $r=1$ i $s=1$, wtedy $\mu_{11}=E[(X-EX)(Y-EY)]$ nazywamy *kowariancją* zmiennych losowych X, Y i oznacza się przez $\operatorname{cov}(X, Y)$, zatem:

$$\operatorname{cov}(X, Y)=E[(X-EX)(Y-EY)]. \tag{5.3.10}$$

Wzór (5.3.7) w przypadku zerowania się jednego z wykładników potęg r albo s wyznacza moment centralny w rozkładach brzegowych zmiennej losowej Y albo X.

Podobnie jak dla jednowymiarowych zmiennych losowych, momenty centralne dowolnego rzędu można wyrazić za pomocą momentów zwykłych, w szczególności dla momentów rzędu drugiego zachodzą następujące wzory:

$$\begin{aligned} &\mu_{20}=\alpha_{20}-\alpha_{10}^2 \quad (D^2X=E(X^2)-(EX)^2),\\ &\mu_{02}=\alpha_{02}-\alpha_{01}^2 \quad (D^2Y=E(Y^2)-(EY)^2), \end{aligned} \tag{5.3.11}$$

$$\mu_{11}=\alpha_{11}-\alpha_{10}\alpha_{01} \quad (\operatorname{cov}(X, Y)=E(XY)-EX\,EY). \tag{5.3.12}$$

Z (2.6.21) i (5.3.12) otrzymujemy, że dla niezależnych zmiennych losowych X, Y:

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

Odpowiednikiem wartości przeciętnej jednowymiarowej zmiennej losowej jest — w przypadku dwuwymiarowej zmiennej losowej wektor wartości przeciętnych $[EX, EY]$ w rozkładach brzegowych, odpowiednikiem wariancji — *macierz momentów centralnych rzędu drugiego*, nazywana też *macierzą kowariancji*; oznacza się ją przez:

$$\mathbf{M}=\begin{bmatrix}\mu_{20} & \mu_{11}\\ \mu_{11} & \mu_{02}\end{bmatrix}=\begin{bmatrix}D^2X & \operatorname{cov}(X, Y)\\ \operatorname{cov}(X, Y) & D^2Y\end{bmatrix}. \tag{5.3.13}$$

Jest to macierz symetryczna i można wykazać, że jej wyznacznik:

$$|\mathbf{M}|=\mu_{20}\mu_{02}-\mu_{11}^2\geqslant 0 \tag{5.3.14}$$

jest nieujemny, co jest równoważne nierówności:

$$[\operatorname{cov}(X, Y)]^2\leqslant D^2X\,D^2Y. \tag{5.3.15}$$

5.3.3. Współczynnik korelacji. Załóżmy, że $D^2X>0$, $D^2Y>0$ i określmy pewien ważny parametr dwuwymiarowej zmiennej losowej zwany *współczynnikiem korelacji*:

$$\rho=\frac{\operatorname{cov}(X, Y)}{\sqrt{D^2X\,D^2Y}}, \quad \text{albo inaczej} \quad \rho=\frac{\mu_{11}}{\sqrt{\mu_{20}\mu_{02}}}. \tag{5.3.16}$$

Często oznacza się $D^2X=\sigma_X^2$ i $D^2Y=\sigma_Y^2$, wówczas

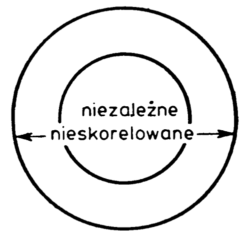
$$\rho=\frac{\operatorname{cov}(X, Y)}{\sigma_X\sigma_Y}.$$

Z nierówności (5.3.15) wynika, że $\rho^2 \leqslant 1$, a więc

$$|\rho| \leqslant 1. \tag{5.3.17}$$

Dowodzi się, że współczynnik korelacji $|\rho| = 1$ wtedy i tylko wtedy, gdy z prawdopodobieństwem 1 zmienne losowe X i Y są związane zależnością liniową, tzn. gdy $P(Y = aX + b) = 1$. Można więc traktować współczynnik korelacji ρ jako miarę współzależności liniowej zmiennych losowych X i Y.

Rys. 5.11. Jeśli zmienne losowe X, Y są niezależne, to są nieskorelowane



W przypadku gdy $\text{cov}(X, Y) = 0$, wtedy zmienne X i Y nazywamy *nieskorelowanymi*. Ze wzoru (5.3.12) wynika, że jeśli X i Y są niezależnymi zmiennymi losowymi, to są one nieskorelowane ($\rho = 0$), odwrotnie niekoniecznie. Rysunek 5.11 ilustruje jakie relacje zachodzą między niezależnymi i nieskorelowanymi oraz zależnymi i skorelowanymi zmiennymi losowymi.

5.3.4. Linie regresji I-go rodzaju. Oznaczmy przez $E(X|Y=y)$ wartość przeciętną zmiennej losowej X pod warunkiem, że zmienna losowa Y przyjmuje wartość równą y, a więc pierwszy moment zwykły w rozkładzie warunkowym zmiennej losowej X((5.1.10), (5.1.23)). Z definicji wartości przeciętnej wynika, że dla każdego y_k, dla którego $p_{.k} \neq 0$,

$$E(X|Y=y_k) = \sum_i x_i P(X=x_i|Y=y_k) = \frac{1}{p_{.k}} \sum_i x_i p_{ik} \tag{5.3.18}$$

w przypadku dwuwymiarowej zmiennej losowej (X, Y) skokowej oraz dla każdego y, dla którego gęstość $f_2(y) \neq 0$,

$$E(X|Y=y) = \int_{-\infty}^{\infty} x f(x|y)\,dx = \frac{1}{f_2(y)} \int_{-\infty}^{\infty} x f(x, y)\,dx \tag{5.3.19}$$

dla zmiennej losowej (X, Y) typu ciągłego.

Podobnie można określić warunkowe wartości przeciętne zmiennej losowej Y, przy warunku, że $X = x$ i tak w przypadku dwuwymiarowej zmiennej losowej skokowej:

$$E(Y|X=x_i) = \sum_k y_k P(Y=y_k|X=x_i) = \frac{1}{p_{i.}} \sum_k y_k p_{ik} \tag{5.3.20}$$

dla tych x_i, dla których $p_{i.} \neq 0$, oraz dla zmiennej losowej typu ciągłego:

$$E(Y|X=x)=\int_{-\infty}^{\infty} yf(y|x)\,dy=\frac{1}{f_1(x)}\int_{-\infty}^{\infty} yf(x,y)\,dy \tag{5.3.21}$$

dla tych x, dla których $f_1(x) \neq 0$.

Postępując jak wyżej, można otrzymać wzory na dowolne momenty warunkowe.

Zauważmy, że $E(X \mid Y=y)$ jest funkcją zmiennej y, $E(Y \mid X=x)$ zaś funkcją x. Oznaczmy

$$E(X|Y=y)=m_1(y), \quad E(Y|X=x)=m_2(x). \tag{5.3.22}$$

Jeśli X i Y są niezależnymi zmiennymi losowymi, to $m_1(y)=EX$ oraz $m_2(x)=EY$.

Zbiór punktów w R^2 o współrzędnych (x, y) spełniających równanie

$$x=m_1(y) \tag{5.3.23}$$

nazywamy *linią regresji I-go rodzaju zmiennej losowej X względem Y*. W przypadku gdy (X, Y) jest dwuwymiarową zmienną losową skokową, wtedy zbiór ten składa się ze skończonej bądź przeliczalnej liczby punktów (x_k, y_k), $k \in N$, których współrzędne spełniają równanie (5.3.23), tzn. $x_k=m_1(y_k)$ dla $k \in N$. Dla zmiennej losowej typu ciągłego, linia regresji I-go rodzaju jest linią o równaniu (5.3.23), mającą co najwyżej przeliczalną liczbę punktów nieciągłości:

Podobnie określamy *linię regresji I-go rodzaju zmiennej losowej Y* względem X jako zbiór punktów w R^2 o współrzędnych (x, y) spełniających równanie

$$y=m_2(x). \tag{5.3.24}$$

Linie regresji I-go rodzaju mają pewną własność, a mianowicie:

Średnie odchylenie kwadratowe zmiennej losowej Y od pewnej funkcji $g(X)$ zmiennej losowej X: $E[Y-g(X)]^2$ jest najmniejsze, gdy funkcja ta z prawdopodobieństwem 1 *jest równa $m_2(X)$, a więc zachodzi*

$$E[Y-m_2(X)]^2=\min_g E[Y-g(X)]^2. \tag{5.3.25}$$

Podobnie dla linii regresji I-go rodzaju zmiennej losowej X względem Y mamy:

$$E[X-m_1(Y)]^2=\min_h E[X-h(Y)]^2. \tag{5.3.26}$$

5.3.5. Zadania rozwiązane.

ZADANIE 5.15. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład o gęstości

$$f(x,y)=\begin{cases} 0{,}2(x+2y) & \text{dla} \quad 0 \leqslant x \leqslant 1 \wedge 0 \leqslant y \leqslant 2, \\ 0 & \text{dla pozostałych } (x,y). \end{cases}$$

Wyznaczyć równanie linii regresji I-go rodzaju zmiennej losowej Y względem X.

Rozwiązanie. Obliczmy wartość przeciętną $E(Y\,|\,X=x)$ według wzoru (5.3.21). W tym celu wyznaczymy gęstość rozkładu brzegowego zmiennej losowej X:

$$f_1(x)=\begin{cases} 0{,}2\int\limits_0^2 (x+2y)\,dy=0{,}4(x+2) & \text{dla} \quad 0\leqslant x\leqslant 1, \\ 0 & \text{poza tym}, \end{cases}$$

skąd

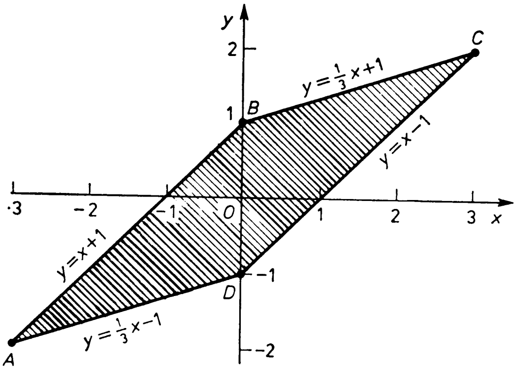
$$\bigwedge_{0\leqslant x\leqslant 1} E(Y\,|\,X=x)=\frac{0{,}2}{0{,}4(x+2)}\int\limits_0^2 y(x+2y)\,dy=\frac{3x+8}{3(x+2)},$$

a więc linią regresji I-go rodzaju zmiennej losowej Y względem X jest

$$y=\frac{3x+8}{3(x+2)} \quad \text{dla} \quad 0\leqslant x\leqslant 1.$$

Jej wykresem jest łuk hiperboli równoosiowej.

Zadanie 5.16. Dany jest dwuwymiarowy rozkład równomierny (X, Y) określony na obwodzie i wewnątrz równoległoboku P ograniczonego prostymi: $y=x-1$, $y=x+1$, $y=\frac{1}{3}x-1$, $y=\frac{1}{3}x+1$. Wykazać, że linia regresji Y względem X jest prostą, natomiast linia regresji X względem Y jest łamaną złożoną z trzech odcinków prostej.



Rys. 5.12. Do zadania 5.16

Rozwiązanie. Z równań boków wynika (rys. 5.12), że $A(-3, -2)$, $B(0, 1)$, $C(3, 2)$ i $D(0, -1)$. Gęstość dwuwymiarowej zmiennej losowej (X, Y) jest postaci

$$f(x,y)=\begin{cases} \dfrac{1}{|P|} & \text{dla} \quad (x,y)\in P, \\ 0 & \text{dla pozostałych } (x,y), \end{cases}$$

gdzie $|P|$ jest polem równoległoboku $ABCD$. Aby je wyznaczyć obliczamy

$$d=AB=\sqrt{9+9}=3\sqrt{2},$$

a następnie wysokość h równoległoboku $ABCD$ jako odległość punktu D od prostej AB ze wzoru: $h=\dfrac{|ax_0+by_0+c|}{\sqrt{a^2+b^2}}$; w naszym przypadku: $h=\dfrac{1+1}{\sqrt{2}}=\sqrt{2}$, skąd $|P|=dh=6$, a więc gęstość:

$$f(x,y)=\begin{cases} \frac{1}{6} & \text{dla} \quad (x,y)\in P, \\ 0 & \text{dla pozostałych } (x,y). \end{cases}$$

Wyznaczmy równanie linii regresji Y względem X: $y=E(Y\,|\,X=x)$.

Aby wyznaczyć rozkład warunkowy, znajdźmy gęstość rozkładu brzegowego zmiennej losowej X:

$$f_1(x)=\begin{cases} \frac{1}{6}\int\limits_{x/3-1}^{x+1} dy=\frac{1}{9}(x+3) & \text{dla} \quad -3\leqslant x<0, \\ \frac{1}{6}\int\limits_{x-1}^{x/3+1} dy=\frac{1}{9}(3-x) & \text{dla} \quad 0\leqslant x\leqslant 3, \\ 0 & \text{dla pozostałych } x. \end{cases}$$

Korzystając ze wzoru $f(y\,|\,x)=\dfrac{f(x,y)}{f_1(x)}$, otrzymujemy

$$f(y\,|\,x)=\begin{cases} \dfrac{3}{2(x+3)} & \text{dla} \quad -3<x<0, \quad \frac{1}{3}x-1<y<x+1, \\ \dfrac{3}{2(3-x)} & \text{dla} \quad 0\leqslant x<3, \quad x-1<y<\frac{1}{3}x+1, \\ 0 & \text{dla pozostałych } x, \end{cases}$$

skąd

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

$$\bigwedge_{-3<x<0} E(Y\,|\,X=x)=\frac{3}{2}\,\frac{1}{x+3}\int\limits_{x/3-1}^{x+1} y\,dy=\frac{2}{3}x$$

oraz

$$\bigwedge_{0\leqslant x<3} E(Y\,|\,X=x)=\frac{3}{3-x}\int\limits_{x-1}^{x/3+1} y\,dy=\frac{2}{3}x,$$

a więc prosta $y=\frac{2}{3}x$ jest linią regresji zmiennej Y względem X: łatwo zauważyć, że zawiera ona przekątną AC. Wyznaczmy teraz równanie linii regresji zmiennej X względem Y, a więc: $x=E(X\,|\,Y=y)$. Rozkład brzegowy zmiennej Y ma gęstość:

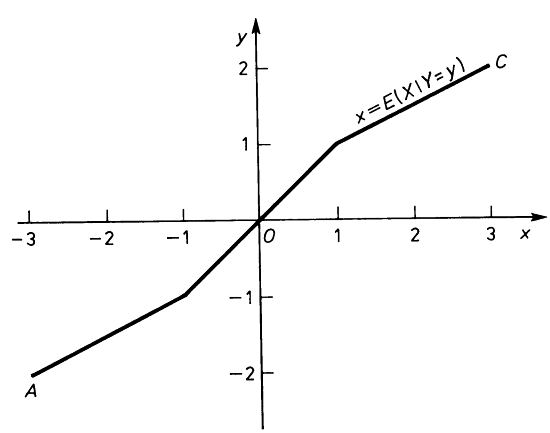
$$f_2(y)=\begin{cases} \frac{1}{6}\int\limits_{y-1}^{3+3y} dx=\frac{1}{3}(y+2) & \text{dla} \quad -2\leqslant y<-1, \\ \frac{1}{6}\int\limits_{y-1}^{y+1} dx=\frac{1}{3} & \text{dla} \quad -1\leqslant y\leqslant 1, \\ \frac{1}{6}\int\limits_{3y-3}^{y+1} dx=\frac{1}{3}(2-y) & \text{dla} \quad 1<y\leqslant 2, \\ 0 & \text{poza tym}. \end{cases}$$

Stąd

$$f(x\,|\,y)=\begin{cases} \dfrac{1}{2(y+2)} & \text{dla} \quad -2<y<-1, \quad y-1<x<3(y+1), \\ \frac{1}{2} & \text{dla} \quad -1\leqslant y<1, \quad y-1<x<y+1, \\ \dfrac{1}{2(2-y)} & \text{dia} \quad 1\leqslant y<2, \quad 3(y-1)<x<y+1, \\ 0 & \text{dla pozostałych } y, \end{cases}$$

$$E(X\,|\,Y=y)=\begin{cases} \dfrac{1}{2(y+2)}\displaystyle\int\limits_{y-1}^{3+3y} x\,dx=2y+1 & \text{dla} \quad -2<y<-1 \\ \frac{1}{2}\int\limits_{y-1}^{y+1} x\,dx=y & \text{dla} \quad -1\leqslant y\leqslant 1, \\ \dfrac{1}{2(2-y)}\displaystyle\int\limits_{3y-3}^{y+1} x\,dx=2y-1 & \text{dla} \quad 1<y<2, \\ 0 & \text{dla pozostałych } y. \end{cases}$$

Stąd linia regresji zmiennej losowej X względem Y jest określona równaniami (rys. 5.13):



Rys. 5.13. Linia regresji I-go rodzaju zmiennej losowej X względem Y (do zad. 5.16)

$$x=\begin{cases} 2y+1 & \text{dla} \quad -2<y\leqslant -1, \\ y & \text{dla} \quad -1<y\leqslant 1, \\ 2y-1 & \text{dla} \quad 1<y<2, \\ 0 & \text{dla pozostałych } y. \end{cases}$$

ZADANIE 5.17. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład podany w tabelce:

	x_i			
y_k	8	9	10	11
1,2	0,1	0,04	0	0
1,25	0,05	0,11	0,2	0
1,3	0	0,1	0,15	0,1
1,35	0	?	0,05	0,1

gdzie X jest wiekiem losowo wybranego dziecka, w pewnej grupie dzieci, Y zaś jego wzrostem (w metrach). Obliczyć a) $P(X=9,\ Y=1{,}35)$, b) współczynnik korelacji między zmiennymi X i Y.

Rozwiązanie. a) Korzystając z warunku, że w podanej tablicy jest określony rozkład prawdopodobieństwa, a więc $\sum\limits_i \sum\limits_k p_{ik}=1$, obliczamy $P(X=9,\ Y=1{,}35)=1-(0{,}1+0{,}04+$ $+0{,}05+0{,}11+0{,}2+0{,}1+0{,}15+0{,}1+0{,}05+0{,}1)=0$.

b) Wyznaczmy rozkłady brzegowe poszczególnych zmiennych, sumując prawdopodobieństwa odpowiednio w kolumnach bądź wierszach. Tak więc rozkładem zmiennej losowej X jest:

x_i	8	9	10	11
$p_{i.}$	0,15	0,25	0,4	0,20

zaś zmiennej losowej Y:

y_k	1,2	1,25	1,3	1,35
$p_{.k}$	0,14	0,36	0,35	0,15

Obliczamy momenty dwuwymiarowej zmiennej losowej potrzebne do wyznaczenia współczynnika korelacji ((5.3.16)).

Wartości przeciętne w rozkładach brzegowych obliczamy według wzorów (5.3.4), (5.3.5):

$$EX=8\cdot 0{,}15+9\cdot 0{,}25+10\cdot 0{,}4+11\cdot 0{,}2=9{,}65,$$

$$EY=1{,}2\cdot 0{,}14+1{,}25\cdot 0{,}36+1{,}3\cdot 0{,}35+1{,}35\cdot 0{,}15=1{,}276.$$

Analogicznie obliczamy drugie momenty zwykłe w tych rozkładach:

$$E(X^2)=8^2\cdot 0{,}15+9^2\cdot 0{,}25+10^2\cdot 0{,}4+11^2\cdot 0{,}2=94{,}05,$$

$$E(Y^2)=1{,}2^2\cdot 0{,}14+1{,}25^2\cdot 0{,}36+1{,}3^2\cdot 0{,}35+1{,}35^2\cdot 0{,}15=1{,}630.$$

Stąd wariancje tych zmiennych ((5.3.11)):

$$D^2X=94{,}05-9{,}65^2=0{,}928,$$

$$D^2Y=1{,}630-1{,}276^2=0{,}002,$$

zaś odchylenia standardowe:

$$\sigma_X=0{,}963, \quad \sigma_Y=0{,}045.$$

Korzystając ze wzoru (5.3.4) dla $r=s=1$, obliczamy:

$$\begin{aligned}\alpha_{11}=E(XY)=&\,8(1{,}2\cdot 0{,}1+1{,}25\cdot 0{,}05)+9(1{,}2\cdot 0{,}04+1{,}25\cdot 0{,}11+1{,}3\cdot 0{,}1)+\\&+10(1{,}25\cdot 0{,}2+1{,}3\cdot 0{,}15+1{,}35\cdot 0{,}05)+\\&+11(1{,}3\cdot 0{,}1+1{,}35\cdot 0{,}1)=12{,}340.\end{aligned}$$

Stąd i ze wzoru (5.3.12):

$$\operatorname{cov}(X, Y)=12{,}340-9{,}65\cdot 1{,}276=12{,}340-12{,}313=0{,}027.$$

Podstawiając otrzymane wyniki do wzoru (5.3.16), otrzymujemy współczynnik korelacji:

$$\rho=\frac{0{,}027}{0{,}963\cdot 0{,}045}\approx 0{,}628.$$

ZADANIE 5.18. Niech X i Y będą niezależnymi zmiennymi losowymi o zerowych wartościach przeciętnych. Wykazać, że dla zależnych zmiennych losowych X, $Z=X^3Y$ zachodzi równość (2.6.21).

Rozwiązanie. Jeśli X i Y są niezależnymi zmiennymi losowymi, to X^n oraz Y dla dowolnego $n\in N$ są niezależnymi zmiennymi losowymi. Stąd i ze wzoru (2.6.21) mamy $E(X^3Y)=E(X^3)EY=0$, a więc również $EZ=0$. Chcemy wykazać, że $E(XZ)=EX\,EZ$; prawa strona tej równości jest zerem, obliczamy $E(XZ)=E(X^4)EY=0$. Wykazaliśmy zatem, że (2.6.21) może zachodzić także dla zależnych zmiennych losowych.

ZADANIE 5.19. Wykazać, że w dwuwymiarowym rozkładzie (X, Y) wartość przeciętna warunkowej wartości przeciętnej zmiennej losowej X przy warunku $Y=y$ jest równa wartości przeciętnej (bezwarunkowej) X (przy założeniu ich istnienia).

Rozwiązanie. Przeprowadźmy przykładowo dowód dla dwuwymiarowej zmiennej losowej ciągłej o gęstości $f(x, y)$:

$$E[E(X|\,Y=y)] = E\int_{-\infty}^{\infty} xf(x|y)\,dx = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} x\,\frac{f(x,y)}{f_2(y)}\,dx\right] f_2(y)\,dy =$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\,dx\,dy=\int_{-\infty}^{\infty} xf_1(x)\,dx=E(X),$$

gdzie f_1, f_2 są brzegowymi gęstościami, a $f(x\,|\,y)$ jest gęstością warunkowego rozkładu zmiennej losowej X pod warunkiem, że $Y=y$.

ZADANIE 5.20. Wykazać słuszność wzoru:

$$D^2X = E[D^2(X \mid Y)] + D^2[E(X \mid Y)]$$

i porównać ten wzór ze wzorem na wariancję połączonych szeregów rozdzielczych ((1.5.17), cz. II).

Rozwiązanie. Rozważmy prawą stronę równości:

$$E[D^2(X \mid Y)] + D^2[E(X \mid Y)] = E[E(X^2 \mid Y) - E^2(X \mid Y)] +$$
$$+ E[E^2(X \mid Y)] - E^2[E(X \mid Y)].$$

Korzystając ze wzoru $E[E(Z \mid Y)] = E(Z)$, otrzymamy

$$E[D^2(X \mid Y)] + D^2[E(X \mid Y)] = E(X^2) - E[E^2(X \mid Y)] + E[E^2(X \mid Y)] - E^2(X) = D^2X.$$

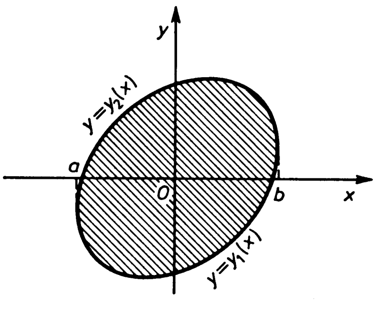
$E[D^2(X \mid Y)]$ jest odpowiednikiem wariancji wewnątrzgrupowej, zaś $D^2[E(X \mid Y)]$ – wariancji międzygrupowej dla połączonych szeregów rozdzielczych (p. 1.5, cz. II).

ZADANIE 5.21. W obszarze ograniczonym elipsą $Ax^2 + 2kxy + By^2 = C$, gdzie $k^2 - AB < 0$, $A, B > 0$, $k \neq 0$, jest określony dwuwymiarowy rozkład równomierny. Wykazać, że linie regresji I-go rodzaju Y względem X oraz X względem Y są liniami prostymi.

Rozwiązanie. Rozkład równomierny określony na w.w. elipsie ma gęstość:

$$f(x, y) = \begin{cases} \dfrac{1}{|D|} & \text{dla} \quad (x, y) \in D, \\ 0 & \text{poza tym}, \end{cases}$$

gdzie przez D (rys. 5.14) oznaczymy elipsę i jej wnętrze, zaś przez $|D|$ – pole obszaru D. Oznaczmy $a = \min\limits_{M(x,y) \in D} x$, zaś $b = \max\limits_{M(x,y) \in D} x$.



Rys. 5.14. Do zadania 5.21

Niech $y = y_1(x)$, $y = y_2(x)$ dla $a \leqslant x \leqslant b$ będą funkcjami określonymi równaniem: $Ax^2 + 2kxy + By^2 = C$, takimi, że obszar $D = \{(x, y) : a \leqslant x \leqslant b \wedge y_1(x) \leqslant y \leqslant y_2(x)\}$. Wyznaczmy równanie linii regresji $y = E(Y \mid X = x)$. W tym celu znajdujemy gęstość warunkową: $f(y \mid x) = \dfrac{f(x, y)}{f_1(x)}$, gdzie f_1 jest gęstością rozkładu brzegowego zmiennej losowej X: $f_1(x) = \int\limits_{-\infty}^{\infty} f(x, y)\, dy$. W naszym przypadku

$$\bigwedge_{a\leqslant x\leqslant b} f_1(x)=\int\limits_{y_1(x)}^{y_2(x)} \frac{1}{|D|}\,dx\,dy=\frac{1}{|D|}[y_2(x)-y_1(x)],$$

a więc $f(y\,|\,x)=\dfrac{1}{y_2(x)-y_1(x)}$, skąd

$$E(Y\,|\,X=x)=\int\limits_{y_1(x)}^{y_2(x)} \frac{y\,dy}{y_2(x)-y_1(x)}=\frac{1}{2}\,\frac{y_2^2(x)-y_1^2(x)}{y_2(x)-y_1(x)}=\frac{1}{2}[y_1(x)+y_2(x)].$$

Ponieważ $\bigwedge_{a\leqslant x\leqslant b} y_1(x),\ y_2(x)$ są rozwiązaniami równania kwadratowego:

$$By^2+2kxy+(Ax^2-C)=0,$$

więc ich suma

$$y_1(x)+y_2(x)=-\frac{2kx}{B},\quad \text{skąd}\quad E(Y\,|\,X=x)=-\frac{xk}{B}.$$

Równanie linii regresji Y względem X jest postaci

$$y=-\frac{xk}{B},$$

a więc jest równaniem prostej przechodzącej przez początek układu współrzędnych.

Ponieważ rozkład prawdopodobieństwa dwuwymiarowej zmiennej losowej (X, Y) jest równomierny oraz równanie elipsy pozostaje niezmienione przy wzajemnej zamianie x i y oraz A i B, więc równaniem linii regresji X względem Y (5.3.23) jest

$$x=-\frac{yk}{A}.$$

5.4. DWUWYMIAROWY ROZKŁAD NORMALNY

Wśród rozkładów dwuwymiarowych zmiennych losowych, szczególnie ważną rolę ze względu na liczne zastosowania (podobnie jak w przypadku jednowymiarowym) spełnia *rozkład normalny*, tj. rozkład typu ciągłego o gęstości:

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

$$\text{dla}\quad (x,y)\in \boldsymbol{R}^2, \qquad (5.4.1)$$

gdzie, jak można wykazać (zad. 5.22), parametry μ_1, μ_2, σ_1, σ_2, są odpowiednio wartościami przeciętnymi i odchyleniami standardowymi w rozkładach brzegowych: $\mu_1=EX$, $\mu_2=EY$, $\sigma_1=\sqrt{D^2X}>0$, $\sigma_2=\sqrt{D^2Y}>0$, zaś ρ współczynnikiem korelacji zmiennych losowych X i Y, przy czym $|\rho|<1$.

5.4.1. Zadania rozwiązane.

Zadanie 5.22. Wykazać, że rozkłady brzegowe dwuwymiarowej zmiennej losowej (X, Y) o rozkładzie (5.4.1) są jednowymiarowymi rozkładami normalnymi o parametrach $\mu_i, \sigma_i, i=1, 2$ odpowiednio.

Rozwiązanie. Korzystając ze wzoru (5.1.18), wyznaczamy gęstość rozkładu brzegowego zmiennej losowej X:

$$f_1(x)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\right.\right.$$
$$\left.\left.-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}dy.$$

Podstawiając $t=\frac{y-\mu_2}{\sigma_2}\Rightarrow y=\mu_2+\sigma_2 t\Rightarrow dy=\sigma_2\,dt$, otrzymujemy

$$f_1(x)=\frac{\exp\left[-\frac{1}{2(1-\rho^2)}\cdot\frac{(x-\mu_1)^2}{\sigma_1^2}\right]}{2\pi\sigma_1\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[-2\rho\frac{x-\mu_1}{\sigma_1}t+t^2\right]\right\}dt,$$

a po dopełnieniu do kwadratu dwumianu wyrażenia pod całką:

$$f_1(x)=\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\rho^2\frac{(x-\mu_1)^2}{\sigma_1^2}\right]\right\}\times$$
$$\times\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2(1-\rho^2)}\left(t-\rho\frac{x-\mu_1}{\sigma_1}\right)^2\right]dt$$

i jeszcze jednej zamianie zmiennych $t-\rho\frac{x-\mu_1}{\sigma_1}=u\sqrt{1-\rho^2}$, mamy

$$f_1(x)=\frac{1}{2\pi\sigma_1}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]\int_{-\infty}^{\infty}e^{-u^2/2}\,du=$$
$$=\frac{1}{\sigma_1\sqrt{2\pi}}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right],$$

tj. gęstość rozkładu $N(\mu_1, \sigma_1)$. Ze wzoru (5.4.1) widać, że postępując analogicznie otrzymamy gęstość f_2 rozkładu brzegowego zmiennej losowej Y postaci:

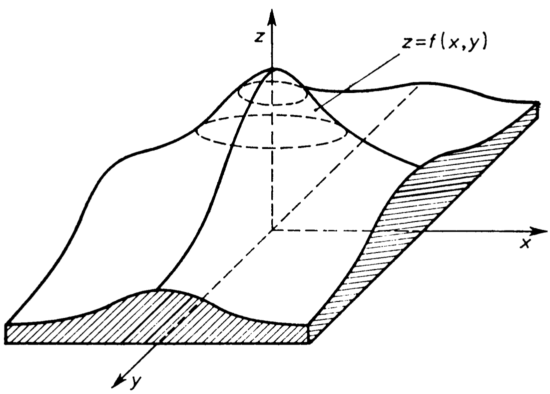
$$f_2(y)=\frac{1}{\sigma_2\sqrt{2\pi}}\exp\left[-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right].$$

Własności dwuwymiarowego rozkładu normalnego:

1. Jeśli dwuwymiarowa zmienna losowa (X, Y) ma rozkład (5.4.1) i zmienne X i Y są nieskorelowane ($\rho=0$), to są niezależne, a gęstość (5.4.1) przyjmuje postać:

$$f(x, y)=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right]. \tag{5.4.2}$$

Szkic wykresu gęstości (5.4.2) dla $\mu_1=\mu_2=0$ przedstawia rys. 5.15.



Rys. 5.15. Gęstość dwuwymiarowego rozkładu normalnego (5.4.2) o parametrach $\mu_1=\mu_2=0=\rho$

2. Rozkłady brzegowe i rozkłady warunkowe w dwuwymiarowym rozkładzie normalnym (5.4.1) są jednowymiarowymi rozkładami normalnymi (zad. 5.22 i 5.23).

3. Linie regresji I-go rodzaju dla rozkładu normalnego (5.4.1) są liniami prostymi (zad. 5.23).

Zadanie 5.23. Wykazać, że równanie linii regresji I-go rodzaju zmiennej losowej X względem Y dla rozkładu (5.4.1) jest postaci

$$\frac{x-\mu_1}{\sigma_1}=\rho\frac{y-\mu_2}{\sigma_2}.$$

Rozwiązanie. Korzystając z (5.1.23) i wyników poprzedniego zadania wyznaczymy gęstość rozkładu warunkowego zmiennej losowej X przy warunku, że $Y=y$,

$$f(x|y)=\frac{1}{\sigma_1\sqrt{2\pi(1-\rho^2)}}\exp\left\{-\frac{1}{2(1-\rho^2)\sigma_1^2}\left[x-\mu_1-\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right]^2\right\},$$

a więc gęstość rozkładu jednowymiarowego normalnego o wartości przeciętnej:

$$E(X|Y=y)=\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2) \tag{5.4.3}$$

i wariancji:

$$D^2(X|Y=y)=(1-\rho^2)\sigma_1^2.$$

Stąd i z (5.3.23) równanie linii regresji I-go rodzaju zmiennej losowej X względem Y jest następujące:

$$x=\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2).$$

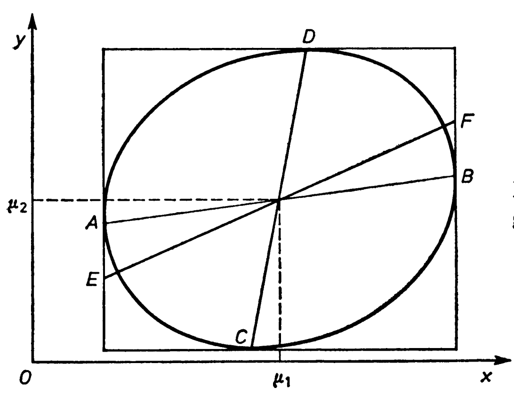
Podobnie można pokazać, że równanie linii regresji I-go rodzaju zmiennej losowej Y względem X dla rozkładu (5.4.1) jest postaci

$$\frac{y-\mu_2}{\sigma_2}=\rho\frac{x-\mu_1}{\sigma_1}.$$

Zauważmy, że — wobec $\frac{\Delta}{4}=\frac{\rho^2}{\sigma_1^2\sigma_2^2}-\frac{1}{\sigma_1^2\sigma_2^2}<0$ oraz $\frac{1}{\sigma_1^2}>0$ — forma kwadratowa $\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}$ jest dodatnio określona dla $0<\rho<1$. Przyjmując więc

$$\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}=c^2, \quad \text{gdzie} \quad c^2\neq 0,$$

otrzymujemy równanie elipsy przy $\rho\neq 0$ o osiach nierównoległych do osi Ox i Oy, w każdym punkcie której gęstość prawdopodobieństwa (5.4.1) dwuwymiarowej zmiennej losowej (X, Y) jest stała; elipsę tę nazywamy *elipsą stałej gęstości*. Wykreślmy następnie prostokąt o bokach równoległych do osi współrzędnych, styczny do elipsy (rys. 5.16). Można wykazać,



Rys. 5.16. Elipsa stałej gęstości dwuwymiarowego rozkładu normalnego (5.4.1) dla $0<\rho<1$

że odcinki łączące przeciwległe punkty styczności przechodzą przez środek (μ_1, μ_2) elipsy, są więc jej średnicami: AB jest średnicą sprzężoną z osią Oy (przechodzi przez środki cięciw elipsy równoległych do osi Oy), a CD jest średnicą sprzężoną z osią Ox; ważniejszymi jednak własnościami są:

1) prosta zawierająca średnicę AB jest prostą regresji Y względem X, a prosta CD jest prostą regresji X względem Y.

2) prosta EF zawierająca dużą oś elipsy jest prostą regresji ortogonalnej (rys. 5.16). (Cz. II, s. 155—157).

ZADANIE 5.24. Dobrać tak stałą c, by

$$f(x, y)=c\exp[-\tfrac{1}{2}(x^2+2xy+5y^2)]$$

była gęstością dwuwymiarowego rozkładu normalnego. Napisać równanie linii regresji I-go rodzaju zmiennej Y względem X.

Rozwiązanie. Funkcja f jest gęstością dwuwymiarowego rozkładu normalnego o parametrach $\mu_1=\mu_2=0$ (5.4.1), jeśli przyjmiemy $c=\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$, gdzie σ_1, σ_2 są odchyleniami standardowymi zmiennych losowych X i Y, ρ zaś – ich współczynnikiem korelacji. Porównując współczynniki przy odpowiednich zmiennych w formie kwadratowej występującej w wykładniku funkcji f oraz funkcji (5.4.1), otrzymamy układ równań:

$$\frac{1}{(1-\rho^2)\sigma_1^2}=1, \qquad \frac{-2\rho}{\sigma_1\sigma_2(1-\rho^2)}=2, \qquad \frac{1}{(1-\rho^2)\sigma_2^2}=5,$$

z którego wyznaczamy ρ, σ_1, σ_2.

W tym celu podzielmy równania pierwsze przez trzecie oraz pierwsze przez drugie stronami, otrzymamy wówczas $\dfrac{\sigma_2^2}{\sigma_1^2}=\dfrac{1}{5}$ oraz $-\dfrac{\sigma_2}{\sigma_1}\dfrac{1}{2\rho}=\dfrac{1}{2}$. Obliczając stąd $\rho=-\dfrac{1}{\sqrt{5}}$ i podstawiając kolejno do równań 1) i 3) mamy $\sigma_1^2=\dfrac{1}{1-\rho^2}=\dfrac{5}{4}$ oraz $\sigma_2^2=\dfrac{1}{5(1-\rho^2)}=\dfrac{1}{4}$. Wobec powyższego stała

$$c=1:\left(2\pi\cdot\frac{\sqrt{5}}{2}\cdot\frac{1}{2}\cdot\frac{2}{\sqrt{5}}\right)=\frac{1}{\pi}.$$

Linia regresji I-go rodzaju zmiennej losowej Y względem X dla dwuwymiarowego rozkładu normalnego jest prostą (zad. 5.23). Podstawiając wyznaczone wartości, otrzymujemy równanie szukanej linii regresji

$$y=-\tfrac{1}{5}x.$$

ZADANIE 5.25. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład o gęstości:

$$f(x, y)=\frac{1}{300\pi}\exp\left\{-\frac{1}{2}\left[\frac{(x-30)^2}{100}+\frac{(y-40)^2}{225}\right]\right\}.$$

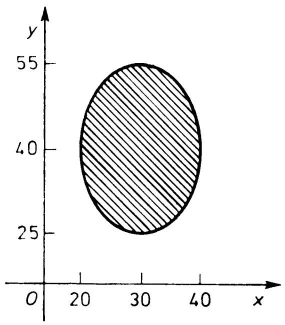
Obliczyć prawdopodobieństwo tego, że suma kwadratów standaryzowanych zmiennych losowych X i Y nie przekracza 1.

Rozwiązanie. Z danej postaci gęstości łatwo odczytać, że $EX=30$, $EY=40$, $D^2X=100$, $D^2Y=225$, $\rho=0$. Stąd szukane prawdopodobieństwo

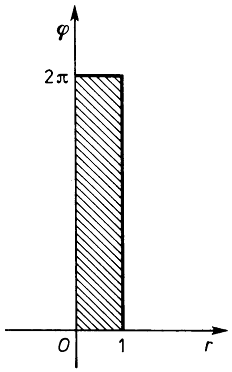
$$P\left[\left(\frac{X-30}{10}\right)^2+\left(\frac{Y-40}{15}\right)^2\leqslant 1\right]=P[(X, Y)\in D]=$$

$$=\frac{1}{300\pi}\iint_D \exp\left\{-\frac{1}{2}\left[\frac{(x-30)^2}{100}+\frac{(y-40)^2}{225}\right]\right\}dx\,dy,$$

gdzie obszar $D=\left\{(x, y):\ \dfrac{(x-30)^2}{100}+\dfrac{(y-40)^2}{225}\leqslant 1\right\}$ (rys. 5.17).



Rys. 5.17. Obszar $D=\{(x,y): \frac{(x-30)^2}{100}+\frac{(y-40)^2}{225}\leqslant 1\}$ (do zad. 5.25)



Rys. 5.18. Obszar $\Delta=\{(r,\varphi): 0\leqslant r\leqslant 1 \wedge 0\leqslant\varphi\leqslant 2\pi\}$ (do zad. 5.25)

Zastosujmy do całki zamianę zmiennych $x=30+10r\cos\varphi$, $y=40+15r\sin\varphi$, o jakobianie $J=150r$. Zauważmy, że przekształcenie odwrotne do podanego odwzorowuje zbiór D na zbiór

$$\Delta=\{(r,\varphi):\ 0\leqslant r\leqslant 1\wedge 0\leqslant\varphi\leqslant 2\pi\} \quad \text{(rys. 5.18)}.$$

Stąd

$$P[(X,Y)\in D]=\frac{1}{300\pi}\iint\limits_D 150re^{-r^2/2}\,dr\,d\varphi=\frac{1}{2\pi}\int_0^1\int_0^{2\pi} re^{-r^2/2}\,d\varphi\,dr=$$

$$=\frac{2\pi}{2\pi}\int_0^1 re^{-r^2/2}\,dr=1-e^{-1/2}\approx 0{,}3935.$$

5.5. PROSTE REGRESJI II-GO RODZAJU

Prostą regresji drugiego rodzaju zmiennej losowej Y względem zmiennej losowej X nazywamy prostą o równaniu $y=ax+b$, którego współczynniki a, b są tak dobrane, aby średnie odchylenie kwadratowe zmiennej losowej Y od zmiennej losowej $aX+b$ było najmniejsze (ze względu na wartości a, b), tzn.

$$E[Y-(aX+b)]^2\equiv K(a,b)=\min. \tag{5.5.1}$$

Okazuje się, że dla dowolnej dwuwymiarowej zmiennej losowej (X, Y), dla której istnieją skończone i dodatnie wariancje σ_X^2, σ_Y^2 w rozkładach brzegowych, istnieje dokładnie jedna prosta $y=ax+b$, która ma własność (5.5.1) o współczynnikach a, b określonych wzorami:

$$a=\rho\frac{\sigma_Y}{\sigma_X}, \qquad b=\alpha_Y-\rho\frac{\sigma_Y}{\sigma_X}\alpha_X, \tag{5.5.2}$$

gdzie $\alpha_X = EX$, $\alpha_Y = EY$. Stąd równanie prostej regresji II-go rodzaju zmiennej losowej Y względem X jest postaci:

$$\frac{y-\alpha_Y}{\sigma_Y} = \rho \frac{x-\alpha_X}{\sigma_X}, \tag{5.5.3}$$

a więc identyczne z równaniem (zad. 5.23) linii regresji I-go rodzaju dla zmiennej losowej o dwuwymiarowym rozkładzie normalnym. W podobny sposób można określić prostą regresji II-go rodzaju zmiennej losowej X względem Y jako prostą o równaniu $x = \alpha y + \beta$, w którym współczynniki α, β są tak dobrane, aby

$$E[X-(\alpha Y+\beta)]^2 = h(\alpha, \beta) = \min. \tag{5.5.4}$$

Otrzymujemy wówczas

$$\alpha = \rho \frac{\sigma_X}{\sigma_Y}, \quad \beta = \alpha_X - \rho \frac{\sigma_X}{\sigma_Y} \alpha_Y.$$

Równanie

$$\frac{x-\alpha_X}{\sigma_X} = \rho \frac{y-\alpha_Y}{\sigma_Y} \tag{5.5.5}$$

jest więc równaniem prostej regresji II-go rodzaju zmiennej X względem Y.

Podstawiając obliczone współczynniki a, b do (5.5.1), zaś α, β do (5.5.4), wyznaczymy minimalne wartości przeciętne odpowiednich odchyleń kwadratowych i tak:

$$E(Y-aX-b)^2 = E\left(Y - \rho \frac{\sigma_Y}{\sigma_X} X - \alpha_Y + \rho \frac{\sigma_Y}{\sigma_X} \alpha_X\right)^2 = \sigma_Y^2 (1-\rho^2),$$

$$E(X-\alpha Y-\beta)^2 = E\left(X - \rho \frac{\sigma_X}{\sigma_Y} Y - \alpha_X + \rho \frac{\sigma_X}{\sigma_Y} \alpha_Y\right)^2 = \sigma_X^2 (1-\rho^2).$$

Prawe strony ostatnich równości noszą nazwę *wariancji resztowych* (*resztkowych*) albo *wariancji regresji Y względem X* i *wariancji regresji X względem Y* odpowiednio. Znaczenie pierwszej wariancji resztowej jest następujące: wariancja Y jest równa σ_Y^2, a po odjęciu od Y „najlepszego" przybliżenia w postaci funkcji liniowej $aX+b$ drugiej zmiennej, wariancja tej reszty (tj. różnicy $Y-aX-b$) — stąd nazwa — przyjmuje wartość najmniejszą $\sigma_Y^2(1-\rho^2)$ (zad. 5.26). Metodę wyznaczania prostych regresji II-go rodzaju (minimalizującą średnie odchylenie kwadratowe) nazywamy *metodą najmniejszych kwadratów*.

5.5.1. Zadania rozwiązane.

ZADANIE 5.26. Wykazać, że wyrażenie $E(Y-aX-b)^2$ w przypadku „najlepszej" prostej jest wariancją zmiennej losowej $Y-aX-b$.

Rozwiązanie. „Najlepszą" (w sensie metody najmniejszych kwadratów) prostą jest prosta regresji II-go rodzaju zmiennej Y względem X o równaniu (5.5.3), której współczynniki: $a = \rho \frac{\sigma_Y}{\sigma_X}$ oraz $b = \alpha_Y - \rho \frac{\sigma_Y}{\sigma_X} \alpha_X$. Wyrażenie $E(Y-aX-b)^2$ jest wariancją zmiennej losowej: $Y-aX-b$, gdy $E(Y-aX-b)=0$. Łatwo sprawdzić, że tak jest:

$$E(Y-aX-b)=E\left(Y-\rho\frac{\sigma_Y}{\sigma_X}X-\alpha_Y+\rho\frac{\sigma_Y}{\sigma_X}\alpha_X\right)=$$

$$=\alpha_Y-\rho\frac{\sigma_Y}{\sigma_X}\alpha_X-\alpha_Y+\rho\frac{\sigma_Y}{\sigma_X}\alpha_X=0.$$

ZADANIE 5.27. Wykazać, że reszta $Y-aX-b$ w przypadku „najlepszej” prostej oraz zmienna losowa X są nieskorelowane.

Rozwiązanie. Zgodnie z definicją nieskorelowanych zmiennych losowych należy wykazać, że $\operatorname{cov}(Y-aX-b, X)=0$. Korzystając z własności wartości przeciętnej i wyników zadania 5.26, mamy:

$$\operatorname{cov}(Y-aX-b, X)=E[(Y-aX-b)(X-EX)]=E[XY-aX^2-bX-(EX)\cdot Y+$$
$$+a(EX)X+bEX]=E(XY)-aE(X^2)-bEX-EX\cdot EY+$$
$$+a(EX)^2+bEX=\operatorname{cov}(X, Y)-a[E(X^2)-(EX)^2]=$$
$$=\operatorname{cov}(X, Y)-\rho\sigma_X^2\frac{\sigma_Y}{\sigma_X}=\operatorname{cov}(X, Y)-\frac{\operatorname{cov}(X, Y)}{\sigma_X\sigma_Y}\sigma_X\sigma_Y=0.$$

Zauważmy, że proste regresji (5.5.3) i (5.5.5) przechodzą przez wspólny punkt (α_X, α_Y) — środek masy rozkładu prawdopodobieństwa, oraz jeśli współczynnik korelacji ρ spełnia równość $\rho^2=1$, to obydwie proste pokrywają się.

ZADANIE 5.28. Niech dwuwymiarowa zmienna losowa (X, Y) ma rozkład podany w tabelce:

	x_i		
y_k	5	6	7
0	0	0	0,1
1	0,1	0,2	0,1
2	0,3	0,1	0,1

Wyznaczyć równania i naszkicować wykresy prostych regresji II-go rodzaju.

Rozwiązanie. Aby wyznaczyć równania prostych regresji (5.5.3), (5.5.5) należy obliczyć wartości przeciętne α_X, α_Y i odchylenia standardowe σ_X, σ_Y w rozkładach brzegowych, oraz współczynnik korelacji ρ.

Sumując prawdopodobieństwa w tabelce w kolumnach, otrzymujemy rozkład brzegowy zmiennej losowej X:

x_i	5	6	7
$p_{i.}$	0,4	0,3	0,3

,

skąd

$$\alpha_X=EX=5\cdot 0{,}4+6\cdot 0{,}3+7\cdot 0{,}3=5{,}9,$$

$$EX^2=5^2\cdot 0{,}4+6^2\cdot 0{,}3+7^2\cdot 0{,}3=35{,}5,$$

$$D^2X=35{,}5-5{,}9^2=0{,}69, \quad \text{a więc} \quad \sigma_X=0{,}83.$$

Podobnie sumując w tabelce prawdopodobieństwa w wierszach, otrzymamy rozkład brzegowy zmiennej losowej Y:

y_k	0	1	2
$p_{.k}$	0,1	0,4	0,5

.

Postępując jak poprzednio, obliczamy:

$$\alpha_Y = EY = 0 \cdot 0{,}1 + 1 \cdot 0{,}4 + 2 \cdot 0{,}5 = 1{,}4,$$

$$EY^2 = 0^2 \cdot 0{,}1 + 1^2 \cdot 0{,}4 + 2^2 \cdot 0{,}5 = 2{,}4,$$

$$D^2 Y = 2{,}4 - 1{,}96 = 0{,}44, \quad \text{skąd} \quad \sigma_Y = 0{,}66.$$

Aby wyznaczyć współczynnik korelacji ρ ((5.3.16)), obliczamy $\operatorname{cov}(X, Y) = E(XY) - EX\,EY$, przy czym $E(XY) = 5(0 \cdot 0 + 1 \cdot 0{,}1 + 2 \cdot 0{,}3) + 6(0 \cdot 0 + 1 \cdot 0{,}2 + 2 \cdot 0{,}1) + 7(0 \cdot 0{,}1 + 1 \cdot 0{,}1 + 2 \cdot$ $\cdot 0{,}1) = 8$,

$$\operatorname{cov}(X, Y) = 8 - 5{,}9 \cdot 1{,}4 = -0{,}26,$$

$$\rho = \frac{-0{,}26}{0{,}83 \cdot 0{,}66} = -0{,}47.$$

Równanie prostej regresji II-go rodzaju (5.5.3) zmiennej losowej Y względem X jest

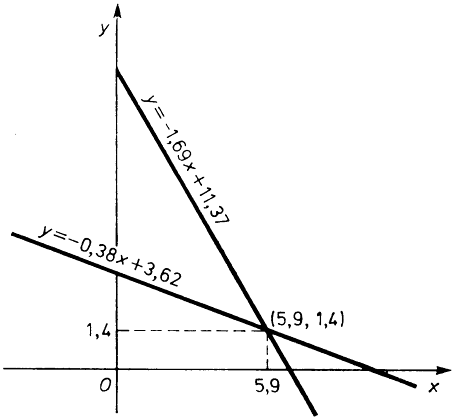
$$\frac{y - 1{,}4}{0{,}66} = -0{,}47 \frac{x - 5{,}9}{0{,}83},$$

po przekształceniu $y = -0{,}38x + 3{,}62$, zaś zmiennej losowej X względem Y (5.5.5) jest postaci:

$$\frac{x - 5{,}9}{0{,}83} = -0{,}47 \frac{y - 1{,}4}{0{,}66},$$

po przekształceniu $y = -1{,}69x + 11{,}37$.

Wykresami są linie proste przechodzące przez punkt (5,9; 1,4) o wyznaczonych współczynnikach kierunkowych (rys. 5.19).



Rys. 5.19. Proste regresji II-go rodzaju (zad. 5.28)

5.6. ROZKŁADY ZŁOŻONE

Niech F będzie dystrybuantą zmiennej losowej X zależną od parametru T. Załóżmy, że T jest zmienną losową o znanym rozkładzie prawdopodobieństwa. Rozważmy dwa przypadki:

1) T — jest zmienną losową typu skokowego o funkcji prawdopodobieństwa:

$$P(T=t_i)=p_i, \qquad i\in N. \tag{5.6.1}$$

Rozkładem złożonym rozkładu F z rozkładem (5.6.1) nazywamy rozkład prawdopodobieństwa o dystrybuancie F_1 określonej wzorem:

$$F_1(x)=\sum_i F(x;\, t_i)P(T=t_i), \tag{5.6.2}$$

2) T — jest zmienną losową typu ciągłego o gęstości g, wówczas *rozkładem złożonym rozkładu F z rozkładem o gęstości g* nazywamy rozkład prawdopodobieństwa o dystrybuancie F_1 określonej wzorem:

$$F_1(x)=\int_{-\infty}^{\infty} F(x,t)g(t)\,dt. \tag{5.6.3}$$

Łatwo zauważyć, że F_1 w obydwu przypadkach jest dystrybuantą rozkładu brzegowego zmiennej losowej X w dwuwymiarowym rozkładzie zmiennej losowej (X, T). Jeśli chcemy wyznaczyć gęstość f_1 albo funkcję prawdopodobieństwa rozkładu złożonego, wystarczy we wzorach (5.6.2), (5.6.3) dystrybuanty F, F_1 zastąpić odpowiednio przez gęstość albo funkcję prawdopodobieństwa, i tak jeśli

1) X jest zmienną losową ciągłą o gęstości f, to

$$f_1(x)=\int_{-\infty}^{\infty} f(x,t)g(t)\,dt, \tag{5.6.4}$$

bądź

$$f_1(x)=\sum_i f(x,\, t_i)P(T=t_i), \tag{5.6.5}$$

w zależności od tego czy T jest zmienną losową typu ciągłego czy skokowego; podobnie w przypadku gdy

2) X jest zmienną losową typu skokowego o funkcji prawdopodobieństwa $P(X=x_i,\, t)$, wówczas

$$P(X=x_k)=\sum_i P(X=x_k,\, t_i)P(T=t_i) \quad \text{dla} \quad k\in N, \tag{5.6.6}$$

albo

$$P(X=x_k)=\int_{-\infty}^{\infty} P(X=x_k,\, t)g(t)\,dt \quad \text{dla} \quad k\in N. \tag{5.6.7}$$

Zauważmy, że rozkłady (5.6.5) i (5.6.6) są mieszaniną rozkładów (p. 3.2).

5.6.1. Zadania rozwiązane.

ZADANIE 5.29. Wyznaczyć złożenie rozkładu Bernoulliego:

$$P(X=i, n)=\binom{n}{i} p^i(1-p)^{n-i} \quad \text{dla} \quad i=0,1,\ldots,n, \quad 0<p<1,$$

z rozkładem Poissona, przyjmując, że

$$P(N=n)=\frac{\lambda^n e^{-\lambda}}{n!} \quad \text{dla} \quad n\in N_0.$$

Rozwiązanie. Stosując wzór (5.6.6) — wobec tego, że $\binom{n}{i}=0$ dla $n<i$ — otrzymujemy

$$P(X=i)=\sum_{n=i}^{\infty}\binom{n}{i} p^i(1-p)^{n-i}\frac{\lambda^n e^{-\lambda}}{n!} \quad \text{dla} \quad i\in N_0,$$

$$P(X=i)=\frac{e^{-\lambda}p^i\lambda^i}{i!}\sum_{n=i}^{\infty}\frac{[\lambda(1-p)]^{n-i}}{(n-i)!}=$$

$$=\frac{e^{-\lambda}(\lambda p)^i}{i!}\sum_{n=0}^{\infty}\frac{[\lambda(1-p)]^n}{n!}=\frac{e^{-\lambda}(\lambda p)^i}{i!}e^{\lambda(1-p)},$$

$$P(X=i)=\frac{(\lambda p)^i e^{-\lambda p}}{i!}, \quad i\in N_0.$$

Jest to rozkład Poissona z parametrem λp.

ZADANIE 5.30. Niech liczba zgłoszeń telefonicznych w pewnej centrali w przedziale czasu o długości t będzie zmienną losową o rozkładzie Poissona

$$P(X=i, t)=e^{-\lambda t}\frac{(\lambda t)^i}{i!} \quad \text{dla} \quad i\in N_0, \tag{1}$$

gdzie λ oznacza przeciętną liczbę zgłoszeń na jednostkę czasu. W pewnej chwili nastąpiła awaria centrali. Zakładając, że czas usuwania awarii jest zmienną losową T o gęstości

$$g(t)=\begin{cases}\mu e^{-\mu t} & \text{dla} \quad t>0,\\ 0 & \text{dla} \quad t\leqslant 0, \quad \mu>0,\end{cases} \tag{2}$$

wyznaczyć funkcję rozkładu prawdopodobieństwa liczby zgłoszeń telefonicznych w czasie usuwania awarii.

Rozwiązanie. Gdyby czas naprawy był zmienną zdeterminowaną t, wtedy szukana funkcja prawdopodobieństwa pokrywałaby się z funkcją (1). Jednakże czas naprawy jest losowy; dlatego czas obserwacji liczby zgłoszeń, jako równy czasowi naprawy, jest zmienną losową T o założonej gęstości (2). Wobec tego, zgodnie z (5.6.7), prawdopodobieństwo $P(X=i)$ tego, że w czasie naprawy pojawi się i zgłoszeń (nieobsłużonych) znajdujemy jak następuje:

$$P(X=i)=\int_{-\infty}^{\infty}P(X=i,T=t)g(t)dt=\int_{0}^{\infty}\frac{e^{-\lambda t}(\lambda t)^i}{i!}\mu e^{-\mu t}dt=$$

$$=\frac{\lambda^i\mu}{i!}\int_{0}^{\infty}t^i e^{-(\lambda+\mu)t}dt=\frac{\lambda^i\mu}{i!}\cdot\frac{\Gamma(i+1)}{(\lambda+\mu)^{i+1}}=$$

$$=\frac{\mu}{\lambda+\mu}\left(\frac{\lambda}{\lambda+\mu}\right)^i=\frac{\mu}{\lambda+\mu}\left(1-\frac{\mu}{\lambda+\mu}\right)^i,\qquad i\in N_0.$$

Złożenie rozkładów Poissona z wykładniczym jest więc rozkładem geometrycznym o parametrze $p=\frac{\mu}{\lambda+\mu}$.

5.7. FUNKCJA CHARAKTERYSTYCZNA DWUWYMIAROWEJ ZMIENNEJ LOSOWEJ

Niech (X, Y) będzie dwuwymiarową zmienną losową o znanym rozkładzie prawdopodobieństwa. Funkcję dwóch zmiennych rzeczywistych t_1, t_2 o wartościach równych wartości przeciętnej zmiennej losowej: $\exp(it_1X+it_2Y)$ nazywamy *funkcją charakterystyczną dwuwymiarowej zmiennej losowej* (X, Y) i oznaczamy przez $\varphi(t_1, t_2)$:

$$\varphi(t_1,t_2)=E\exp[i(t_1X+t_2Y)]\quad \text{dla}\quad (t_1,t_2)\in R^2. \tag{5.7.1}$$

W szczególnym przypadku

$$\varphi(t_1,t_2)=\begin{cases}\sum\limits_j\sum\limits_k p_{jk}\exp[i(x_jt_1+y_kt_2)] & \text{dla zmiennej skokowej o rozkładzie}\\ & P(X=x_j,\ Y=y_k)=p_{jk}>0,\ \text{przy czym}\ \sum\limits_{j,k}p_{jk}=1,\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}f(x,y)\exp[i(xt_1+yt_2)]\,dx\,dy & \text{dla zmiennej typu ciągłego}\\ & \text{o gęstości } f.\end{cases} \tag{5.7.2}$$

Zauważmy, że $\varphi(t_1, 0)$ jest funkcją charakterystyczną rozkładu brzegowego zmiennej losowej X, zaś $\varphi(0, t_2)$ – zmiennej Y. Można wykazać, że:

Warunkiem koniecznym i wystarczającym niezależności zmiennych X i Y jest zachodzenie związku:

$$\bigwedge_{(t_1,t_2)\in R^2}\varphi(t_1,t_2)=\varphi(t_1,0)\,\varphi(0,t_2). \tag{5.7.3}$$

Własności funkcji charakterystycznej:

a) $\varphi(0, 0)=1$,

b) $|\varphi(t_1, t_2)|\leqslant 1$,

c) $\varphi(-t_1, -t_2)=\overline{\varphi(t_1, t_2)}$

($\overline{\varphi(t_1, t_2)}$ oznacza liczbę zespoloną sprzężoną z $\varphi(t_1, t_2)$).

d) Jeśli istnieją skończone momenty mieszane $\alpha_{r_1 r_2}$ rzędu r_1+r_2 zmiennej (X, Y) ((5.3.2)), to istnieje pochodna mieszana funkcji charakterystycznej: $\dfrac{\partial^{r_1+r_2}\varphi(t_1, t_2)}{\partial t_1^{r_1}\partial t_2^{r_2}}$ i zachodzi związek:

$$\alpha_{r_1 r_2}=E(X^{r_1}Y^{r_2})=\frac{1}{i^{r_1+r_2}}\left[\frac{\partial^{r_1+r_2}\varphi(t_1, t_2)}{\partial t_1^{r_1}\partial t_2^{r_2}}\right]_{\substack{t_1=0\\ t_2=0}}. \tag{5.7.4}$$

Funkcja charakterystyczna dwuwymiarowej zmiennej losowej o dwuwymiarowym rozkładzie normalnym (5.4.1) jest postaci:

$$\varphi(t_1, t_2)=\exp\left[i(\mu_1 t_1+\mu_2 t_2)-\tfrac{1}{2}(\sigma_1^2 t_1^2+\sigma_2^2 t_2^2+2\rho\sigma_1\sigma_2 t_1 t_2)\right]. \tag{5.7.5}$$

5.7.1. Zadania rozwiązane.

ZADANIE 5.31. Zmienna (X, Y) ma rozkład o funkcji charakterystycznej (5.7.5). Dla jakich wartości parametrów rozkładu zmienne X i Y są niezależne?

Rozwiązanie. Wyznaczmy funkcje charakterystyczne rozkładów brzegowych:

$$\varphi(t_1, 0)=\exp(i\mu_1 t_1-\tfrac{1}{2}\sigma_1^2 t_1^2),$$

$$\varphi(0, t_2)=\exp(i\mu_2 t_2-\tfrac{1}{2}\sigma_2^2 t_2^2).$$

Wyznaczymy wartości parametrów μ_1, μ_2, σ_1, σ_2, ρ, dla których zachodzi równość (5.7.3):

$$\exp\left[i(\mu_1 t_1+\mu_2 t_2)-\tfrac{1}{2}(\sigma_1^2 t_1^2+\sigma_2^2 t_2^2+2\rho\sigma_1\sigma_2 t_1 t_2)\right]=$$
$$=\exp(i\mu_1 t_1-\tfrac{1}{2}\sigma_1^2 t_1^2)\exp(i\mu_2 t_2-\tfrac{1}{2}\sigma_2^2 t_2^2)\exp(-\rho\sigma_1\sigma_2 t_1 t_2),$$

ponieważ σ_1, $\sigma_2>0$, więc powinno być $\rho=0$.

Otrzymaliśmy znany wynik, że warunkiem koniecznym i wystarczającym niezależności zmiennych losowych o rozkładzie normalnym (5.4.1) jest, by $\rho=0$ (zmienne nieskorelowane).

ZADANIE 5.32. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład podany w tabelce:

	x_i	
y_k	0	1
-1	$\frac{1}{3}$	$\frac{1}{6}$
1	$\frac{1}{3}$	$\frac{1}{6}$

Wyznaczyć funkcję charakterystyczną dwuwymiarowej zmiennej losowej (X, Y) oraz jej moment mieszany α_{12}.

Rozwiązanie. Korzystając ze wzoru (5.7.2), wyznaczmy funkcję charakterystyczną:

$$\varphi(t_1, t_2)=\tfrac{1}{3}e^{-it_2}+\tfrac{1}{3}e^{it_2}+\tfrac{1}{6}e^{it_1-it_2}+\tfrac{1}{6}e^{it_1+it_2}.$$

Stosując wzór Eulera $e^{it}=\cos t+i\sin t$ dla $t\in \boldsymbol{R}$, otrzymamy:

$$\varphi(t_1, t_2)=\tfrac{1}{3}\cos t_2(2+\cos t_1+i\sin t_1).$$

Obliczamy pochodne:

$$\frac{\partial\varphi(t_1, t_2)}{\partial t_1}=\tfrac{1}{3}\cos t_2(-\sin t_1+i\cos t_1),$$

$$\frac{\partial^2\varphi(t_1, t_2)}{\partial t_1\,\partial t_2}=-\tfrac{1}{3}\sin t_2(-\sin t_1+i\cos t_1),$$

$$\frac{\partial^3\varphi(t_1, t_2)}{\partial t_1\,\partial t_2^2}=-\tfrac{1}{3}\cos t_2(-\sin t_1+i\cos t_1).$$

Wartość trzeciej pochodnej dla $t_1=0$, $t_2=0$ wynosi:

$$\left[\frac{\partial^3\varphi(t_1, t_2)}{\partial t_1\,\partial t_2^2}\right]_{\substack{t_1=0\\ t_2=0}}=-\frac{1}{3}i.$$

Stąd i ze wzoru (5.7.4) otrzymujemy

$$\alpha_{12}=-\frac{i}{3i^3}=\frac{1}{3}.$$

5.8. WIELOWYMIAROWE ZMIENNE LOSOWE

Pojęcia związane z dwuwymiarową zmienną losową (5.1) - (5.7) można uogólnić na przypadek n-wymiarowej ($n>2$) zmiennej losowej.

Układ n zmiennych losowych $X_1, \ldots, X_n$ określonych niekoniecznie na tej samej przestrzeni probabilistycznej nazywamy *n-wymiarową zmienną losową*, bądź *wektorem losowym* i zapisujemy $(X_1, \ldots, X_n)$, gdzie X_i, $i=1, \ldots, n$ nazywamy jego *składowymi* (*współrzędnymi*).

Dystrybuantą n-wymiarowej zmiennej losowej $(X_1, \ldots, X_n)$ nazywamy funkcję n zmiennych rzeczywistych określoną wzorem:

$$F(x_1, \ldots, x_n)=P(X_1<x_1, \ldots, X_n<x_n) \quad \text{dla} \quad (x_1, \ldots, x_n)\in\boldsymbol{R}^n. \tag{5.8.1}$$

Łączny rozkład prawdopodobieństwa dowolnego układu k $(1\leqslant k\leqslant n)$ zmiennych $X_{i_1}, \ldots, X_{i_k}$, spośród n zmiennych losowych $X_1, \ldots, X_n$, nazywamy *rozkładem brzegowym k-wymiarowym w n-wymiarowym rozkładzie* zmiennej losowej $(X_1, \ldots, X_n)$. Dystrybuanta takiego rozkładu jest określona wzorem

$$F_{i_1, \ldots, i_k}(x_{i_1}, \ldots, x_{i_k})=\lim_{\substack{x_j\to\infty\\ j\neq i_1, \ldots, i_k}} F(x_1, \ldots, x_n). \tag{5.8.2}$$

Jeśli zmienna losowa $(X_1, \ldots, X_n)$ przyjmuje skończoną bądź przeliczalną liczbę wartości $(x_{1j}, \ldots, x_{nj})$, $j \in N$, każdą z dodatnim prawdopodobieństwem:

$$P(X_1 = x_{1j}, \ldots, X_n = x_{nj}) = p_{1j, \ldots, nj} > 0 \quad \text{dla} \quad j \in N, \tag{5.8.3}$$

przy czym $\sum_j p_{1j, \ldots, nj} = 1$, to $(X_1, \ldots, X_n)$ nazywamy *n-wymiarową zmienną losową skokową* (*dyskretną*), a jej dystrybuanta wyraża się równością:

$$F(x_1, \ldots, x_n) = \sum_{x_{1j} < x_1} \cdots \sum_{x_{nj} < x_n} p_{1j, \ldots, nj}. \tag{5.8.4}$$

Zmienną losową $(X_1, \ldots, X_n)$ nazywamy *typu ciągłego* (*ciągłą*), jeśli istnieje nieujemna funkcja f n zmiennych rzeczywistych $x_1, \ldots, x_n$ taka,że dystrybuanta n-wymiarowej zmiennej losowej $(X_1, \ldots, X_n)$ jest określona wzorem:

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(u_1, \ldots, u_n) du_1 \ldots du_n \quad \text{dla} \quad (x_1, \ldots, x_n) \in \boldsymbol{R}^n. \tag{5.8.5}$$

Funkcję f nazywamy *gęstością prawdopodobieństwa* n-wymiarowej zmiennej losowej $(X_1, \ldots, X_n)$. We wszystkich punktach ciągłości gęstości f zachodzi wzór

$$\frac{\partial^n F(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n} = f(x_1, \ldots, x_n). \tag{5.8.6}$$

Również pojęcie niezależności dwóch zmiennych losowych można uogólnić na przypadek n-wymiarowy i tak powiemy, że $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowymi, jeśli:

$$\bigwedge_{(x_1, \ldots, x_n) \in \boldsymbol{R}^n} F(x_1, \ldots, x_n) = F_1(x_1) \ldots F_n(x_n), \tag{5.8.7}$$

gdzie $F_i(x_i)$, $i = 1, \ldots, n$, są dystrybuantami rozkładów brzegowych zmiennych losowych X_i.

Łatwo zauważyć, że jeśli $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowymi, to dowolna liczba $k<n$ spośród nich stanowi układ k niezależnych zmiennych losowych: $\bigwedge_{k<n} X_{i_1}, X_{i_2}, \ldots, X_{i_k}$, gdzie $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$ są również niezależne.

Jeśli $(X_1, \ldots, X_n)$ jest n-wymiarową zmienną losową skokową, to warunkiem koniecznym i wystarczającym niezależności zmiennych $X_1, \ldots, X_n$ jest zachodzenie równości:

$$\bigwedge_{(x_{1j}, \ldots, x_{nj}) \in \boldsymbol{R}^n} P(X_1 = x_{1j}, \ldots, X_n = x_{nj}) = P_1(X_1 = x_{1j}) \ldots P_n(X_n = x_{nj}), \tag{5.8.8}$$

gdzie P_k, $k = 1, \ldots, n$ są funkcjami prawdopodobieństwa rozkładów brzegowych jednowymiarowych zmiennych losowych $X_1, \ldots, X_n$. Jeżeli natomiast $(X_1, \ldots, X_n)$ jest zmienną losową ciągłą o gęstości f, to korzystając ze wzorów (5.6.5) - (5.6.7) łatwo wykazać, że warunkiem koniecznym i wystarczającym niezależności zmiennych $X_1, \ldots, X_n$ jest spełnienie związku:

$$\bigwedge_{(x_1, \ldots, x_n) \in \boldsymbol{R}^n} f(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n), \tag{5.8.9}$$

gdzie f_i, $i = 1, \ldots, n$ są gęstościami rozkładów brzegowych jednowymiarowych zmiennych losowych: $X_1, \ldots, X_n$.

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

$$P(X_{k+1}=x_{k+1,j},\ldots,X_n=x_{nj}\,|\,X_1=x_{1j},\ldots,X_k=x_{kj})=$$
$$=\frac{P(X_1=x_j,\ldots,X_n=x_{nj})}{P_{1,\ldots,k}(X_1=x_{1j},\ldots,X_k=x_{kj})}, \quad (5.8.10)$$

gdzie $P_{1,\ldots,k}$ jest funkcją prawdopodobieństwa rozkładu brzegowego zmiennej losowej $(X_1,\ldots,X_k)$, nazywamy *rozkładem warunkowym zmiennej losowej* $(X_{k+1},\ldots,X_n)$ pod warunkiem, że $(X_1=x_{1j},\ldots,X_k=x_{kj})$.

Dla zmiennej $(X_1,\ldots,X_n)$ ciągłej o gęstości f, gdy gęstość $f_{1,\ldots,k}$ rozkładu brzegowego zmiennej $(X_1,\ldots,X_k)$ jest dodatnia, wzór:

$$f(x_{k+1},\ldots,x_n\,|\,x_1,\ldots,x_k)=\frac{f(x_1,\ldots,x_n)}{f_{1,\ldots,k}(x_1,\ldots,x_k)} \quad (5.8.11)$$

określa gęstość rozkładu warunkowego zmiennej losowej $(X_{k+1},\ldots,X_n)$ przy założeniu, że $(X_1=x_1,\ldots,X_k=x_k)$.

Wartość przeciętna funkcji g n zmiennych losowych $X_1,\ldots,X_n$ jest liczbą określoną wzorem:

$$Eg(X_1,\ldots,X_n)=$$
$$=\begin{cases}\sum\limits_i g(x_{i1},\ldots,x_{in})p_{i1,\ldots,in} & \text{dla zmiennej skokowej o rozkładzie (5.8.3)}\\ \int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty} g(x_1,\ldots,x_n)f(x_1,\ldots,x_n)\,dx_1\ldots dx_n & \text{dla zmiennej losowej } (X_1,\ldots,X_n) \text{ typu ciągłego o gęstości } f\end{cases} \quad (5.8.12)$$

o ile odpowiednio

$$\sum_i |g(x_{i1},\ldots,x_{in})|\,p_{i1,\ldots,in}<\infty,$$

bądź

$$\int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty}|g(x_1,\ldots,x_n)|f(x_1,\ldots,x_n)\,dx_1\ldots dx_n<\infty.$$

Momentem $\alpha_{r_1\ldots r_n}$ *zwykłym mieszanym rzędu* $r_1+\ldots+r_n$, $r_i\in N_0$, $i=1,\ldots,n$ zmiennej losowej $(X_1,\ldots,X_n)$ nazywamy wartość przeciętną iloczynu $X_1^{r_1}\ldots X_n^{r_n}$, a więc

$$\alpha_{r_1\ldots r_n}=E(X_1^{r_1}\ldots X_n^{r_n}). \quad (5.8.13)$$

W szczególnym przypadku gdy $r_1=\ldots=r_k=0$ dla $k<n$, otrzymujemy moment zwykły mieszany rzędu $r_{k+1}+\ldots+r_n$ rozkładu brzegowego zmiennej losowej $(X_{k+1},\ldots,X_n)$.

Wektor $(EX_1,\ldots,EX_n)$ nazywamy *wektorem wartości przeciętnych*, albo *środkiem masy prawdopodobieństwa w n-wymiarowym rozkładzie.*

Następnie określimy *momenty centralne* $\mu_{r_1 \ldots r_n}$ *mieszane rzędu* $r_1+\ldots+r_n$ zmiennej losowej $(X_1, \ldots, X_n)$ jako wartość przeciętną funkcji: $(X_1-EX_1)^{r_1} \ldots (X_n-EX_n)^{r_n}$, mamy zatem

$$\mu_{r_1 \ldots r_n}=E[(X_1-EX_1)^{r_1} \ldots (X_n-EX_n)^{r_n}]. \tag{5.8.14}$$

W szczególnym przypadku – dla dwóch dowolnych zmiennych losowych $X_i, X_j, i \neq j$, przy $r_i=r_j=1$ – otrzymujemy:

$$E[(X_i-EX_i)(X_j-EX_j)]=\mathrm{cov}(X_i, X_j),$$

kowariancję zmiennych X_i, X_j; często oznacza się ją przez:

$$\sigma_{ij}=\mathrm{cov}(X_i, X_j) \quad \text{dla} \quad i, j=1, \ldots, n. \tag{5.8.15}$$

Jeśli $i=j$, to

$$\sigma_{ii}=D^2 X_i, \quad i=1, \ldots, n.$$

Łatwo można wykazać, że:

a) $E(X_1+\ldots+X_n)=EX_1+\ldots+EX_n$, o ile $EX_1, \ldots, EX_n$ istnieją.

b) $D^2(X_1+\ldots+X_n)=D^2X_1+\ldots+D^2X_n$ dla niezależnych zmiennych losowych $X_1, \ldots, X_n$, przy założeniu istnienia występujących w równości wariancji.

Załóżmy, że σ_{ij}, $i, j=1, \ldots, n$, istnieją, przy czym $\sigma_{ii}=D^2X_i>0$, wówczas pewnym uogólnieniem wariancji na przypadek n-wymiarowy jest tzw. *macierz kowariancji*, albo *macierz momentów rzędu drugiego*, którą oznaczamy przez

$$\mathbf{M}=[\sigma_{ij}]_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}=\begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}=$$

$$=\begin{bmatrix} D^2X_1 & \mathrm{cov}(X_1, X_2) & \ldots & \mathrm{cov}(X_1, X_n) \\ \mathrm{cov}(X_2, X_1) & D^2X_2 & \ldots & \mathrm{cov}(X_2, X_n) \\ \cdot & \cdot & \cdots & \cdot \\ \mathrm{cov}(X_n, X_1) & \mathrm{cov}(X_n, X_2) & \ldots & D^2X_n \end{bmatrix}. \tag{5.8.16}$$

Macierz $\mathbf{M}$ jest macierzą symetryczną o wyznaczniku $|\mathbf{M}| \geqslant 0$. Jeśli jednak $|\mathbf{M}|=0$, a więc rząd macierzy $\mathbf{M}$ jest $k<n$, to $(X_1, \ldots, X_n)$ jest zdegenerowaną n-wymiarową zmienną losową o rozkładzie osobliwym (w rzeczywistości k-wymiarowym), tj. takim którego całkowita masa prawdopodobieństwa jest rozłożona na k-wymiarowej hiperpłaszczyźnie.

Wartością przeciętną zmiennej losowej X_1 o jednowymiarowym rozkładzie warunkowym przy założeniu, że $(X_2=x_{2j}, \ldots, X_n=x_{nj})$ nazywamy liczbę określoną wzorem:

$$E(X_1\,|\,X_2=x_{2j},\ldots,X_n=x_{nj})=$$

$$=\begin{cases}\dfrac{\sum\limits_k x_{1k}P(X_1=x_{1k},X_2=x_{2j},\ldots,X_n=x_{nj})}{\sum\limits_k P(X_1=x_{1k},X_2=x_{2j},\ldots,X_n=x_{nj})} & \text{dla zmiennej skokowej o rozkładzie (5.8.3),}\\[2ex] \dfrac{\int\limits_{-\infty}^{\infty} x_1 f(x_1,\ldots,x_n)\,dx_1}{\int\limits_{-\infty}^{\infty} f(x_1,\ldots,x_n)\,dx_1} & \text{dla zmiennej ciągłej o gęstości } f.\end{cases} \tag{5.8.17}$$

Oznaczmy $E(X_1\,|\,X_2,\ldots,X_n)=m_1(X_2,\ldots,X_n)$.

Zbiór punktów w R^n o współrzędnych $x_1,\ldots,x_n$ spełniających równanie

$$x_1=m_1(x_2,\ldots,x_n) \tag{5.8.18}$$

nazywamy *hiperpowierzchnią regresji I-go rodzaju* (odpowiednik linii regresji I-go rodzaju) zmiennej X_1 względem pozostałych zmiennych $(X_2,\ldots,X_n)$. W podobny sposób można określić hiperpowierzchnię regresji I-go rodzaju zmiennej X_k względem pozostałych zmiennych $(X_1,\ldots,X_{k-1},X_{k+1},\ldots,X_n)$.

Hiperpłaszczyzną regresji II-go rodzaju zmiennej X_1 względem $(X_2,\ldots,X_n)$ (odpowiednik prostej regresji II-go rodzaju) nazywamy zbiór punktów w przestrzeni R^n o współrzędnych $x_1,\ldots,x_n$ spełniających równanie

$$x_1-EX_1=a_{12}(x_2-EX_2)+\ldots+a_{1n}(x_n-EX_n), \tag{5.8.19}$$

w którym współczynniki a_{1k}, $k=2,\ldots,n$, są tak dobrane, aby zachodził związek:

$$E[X_1-EX_1-\sum_{i=2}^{n}a_{1i}(X_i-EX_i)]^2=\min,$$

tzn. hiperpłaszczyzna (5.8.19) ma tę własność, że spośród wszystkich hiperpłaszczyzn w R^n średnie odchylenie kwadratowe zmiennej X_1-EX_1 od tej hiperpłaszczyzny (od zmiennej $\sum\limits_{i=2}^{n}a_{1i}(X_i-EX_i)$) jest najmniejsze (mówi się też, że jest to „najlepsza", w sensie metody najmniejszych kwadratów, hiperpłaszczyzna).

Można wykazać, że

$$a_{1i}=-\frac{M_{1i}^*}{M_{11}^*},\qquad i=2,\ldots,n, \tag{5.8.20}$$

gdzie M_{1i}^* dla $i=1,\ldots,n$ są dopełnieniami algebraicznymi elementów σ_{1i}, $i=1,\ldots,n$, odpowiednio, przy czym wobec $|\mathrm{M}|\neq 0$ i $D^2X_i>0$, mamy $M_{ii}^*\neq 0$ dla $i=1,\ldots,n$.

W podobny sposób można uzyskać równanie hiperpłaszczyzny regresji II-go rodzaju dowolnej zmiennej X_k, $k=1,\ldots,n$, względem pozostałych $n-1$ zmiennych. Mamy wówczas

$$x_k-EX_k=\sum_{i\neq k}^{n}a_{ki}(x_i-EX_i),$$

gdzie $a_{ki}=-\dfrac{M_{ki}^*}{M_{kk}^*}$, $i=1,\ldots,k-1,k+1,\ldots,n$.

Można wykazać, że każda ze zmiennych

$$Y_k = X_k - EX_k + \frac{1}{M_{kk}^*} \sum_{i \neq k}^{n} M_{ki}^*(X_i - EX_i), \quad k=1,\dots,n,$$

nazywana *resztą* jest nieskorelowana z żadną ze zmiennych losowych X_i przy $i \neq k$, a wariancja resztowa

$$D^2 Y_k = \frac{|\mathbf{M}|}{M_{kk}^*} \quad \text{dla} \quad k=1,\dots,n.$$

Podobnie jak wśród jedno- i dwuwymiarowych zmiennych losowych, wśród wielowymiarowych zmiennych losowych szczególne znaczenie ma *wielowymiarowy rozkład normalny*.

n-wymiarowa zmienna losowa $(X_1, \dots, X_n)$ ma *n-wymiarowy rozkład normalny*, jeśli jej gęstość jest określona wzorem:

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{M}|}} \exp\left[-\tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}(x_i - EX_i)(x_j - EX_j)\right], \tag{5.8.21}$$

gdzie c_{ij} są elementami macierzy odwrotnej do macierzy kowariancji $\mathbf{M}$, określonymi wzorami:

$$c_{ij} = \frac{1}{|\mathbf{M}|} M_{ij}, \tag{5.8.22}$$

gdzie M_{ij} są dopełnieniami algebraicznymi elementów σ_{ij}. Można wykazać, że jeśli $(X_1, \dots, X_n)$ ma rozkład (5.8.21) oraz zmienne losowe są parami nieskorelowane, tzn. $\bigwedge_{i \neq j} \sigma_{ij} = 0$, to zmienne $X_1, \dots, X_n$ są niezależne o łącznej gęstości (dla $n = 2$ por. (5.4.2)):

$$f(x_1, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n \sigma_{11}\sigma_{22}\dots\sigma_{nn}}} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} \frac{(x_i - EX_i)^2}{\sigma_{ii}^2}\right].$$

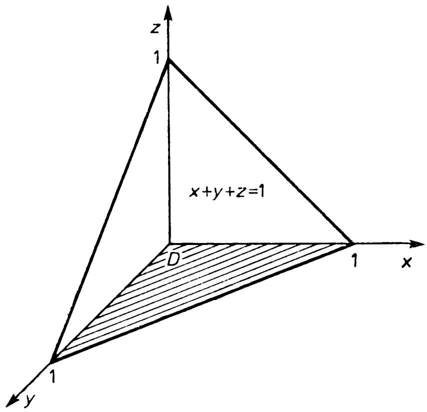
5.8.1. Zadania rozwiązane.

ZADANIE 5.33. Trójwymiarowa zmienna losowa (X, Y, Z) ma rozkład Dirichleta o gęstości:

$$f(x, y, z) = \begin{cases} \dfrac{\Gamma(p_1+p_2+p_3+p_4)}{\Gamma(p_1)\Gamma(p_2)\Gamma(p_3)\Gamma(p_4)} x^{p_1-1} y^{p_2-1} z^{p_3-1} (1-x-y-z)^{p_4-1} & \\ & \text{dla} \quad (x, y, z) \in V, \\ 0 & \text{poza tym}, \end{cases}$$

gdzie $V = \{(x, y, z) : x \geqslant 0,\ y \geqslant 0,\ z \geqslant 0 \wedge x+y+z \leqslant 1\}$ oraz $p_1, p_2, p_3, p_4 > 0$. Jest to uogólnienie rozkładu beta.

Wyznaczyć gęstość rozkładu warunkowego $f(z \mid x, y)$.



Rys. 5.20. Obszar, na którym jest skoncentrowana gęstość rozkładu Dirichleta (zad. 5.33)

Rozwiązanie. Gęstość rozkładu Dirichleta jest dodatnia dla $(x, y, z) \in V$, gdzie V jest ostrosłupem (rys. 5.20).

Korzystając ze wzoru (5.8.11), otrzymamy

$$\bigwedge_{(x,y)\in D} f(z\,|\,x, y)=\frac{f(x, y, z)}{f_{1,2}(x, y)},$$

gdzie $D=\{(x, y): 0\leqslant x\leqslant 1 \wedge 0\leqslant y\leqslant 1-x\}$.

Wyznaczamy gęstość $f_{1,2}(x, y)$ wektora (X, Y):

$$\bigwedge_{(x,y)\in D} f_{1,2}(x, y)=c\int_0^{1-x-y} x^{p_1-1}y^{p_2-1}z^{p_3-1}(1-x-y-z)^{p_4-1}dz,$$

gdzie $c=\dfrac{\Gamma(p_1+\ldots+p_4)}{\Gamma(p_1)\ldots\Gamma(p_4)}$; podstawiając $z=(1-x-y)u$, otrzymamy:

$$\bigwedge_{(x,y)\in D} f_{1,2}(x, y)=cx^{p_1-1}y^{p_2-1}(1-x-y)^{p_3+p_4-1}\int_0^1 u^{p_3-1}(1-u)^{p_4-1}du=$$

$$=cx^{p_1-1}y^{p_2-1}(1-x-y)^{p_3+p_4-1}B(p_3, p_4)=$$

$$=\frac{\Gamma(p_1+p_2+p_3+p_4)\Gamma(p_3)\Gamma(p_4)}{\Gamma(p_1)\Gamma(p_2)\Gamma(p_3)\Gamma(p_4)\Gamma(p_3+p_4)}x^{p_1-1}y^{p_2-1}(1-x-y)^{p_3+p_4-1}.$$

Rozkład brzegowy zmiennej (X, Y) jest więc także rozkładem Dirichleta o gęstości:

$$f_{1,2}(x, y)=\begin{cases}\dfrac{\Gamma(p_1+p_2+p_3+p_4)}{\Gamma(p_1)\Gamma(p_2)\Gamma(p_3+p_4)}x^{p_1-1}y^{p_2-1}(1-x-y)^{p_3+p_4-1} & \\ & \text{dla} \quad x>0,\ y>0\wedge x+y\leqslant 1,\\ 0 & \text{poza tym,}\end{cases}$$

zaś gęstość rozkładu warunkowego zmiennej Z jest określona wzorem:

$$\bigwedge_{(x,y)\in D} f(z\,|\,x, y)=\begin{cases}\dfrac{\Gamma(p_3+p_4)}{\Gamma(p_3)\Gamma(p_4)}z^{p_3-1}\dfrac{(1-x-y-z)^{p_4-1}}{(1-x-y)^{p_3+p_4-1}} & \text{dla} \quad 0\leqslant z\leqslant 1-x-y,\\ 0 & \text{dla pozostałych } z.\end{cases}$$

ZADANIE 5.34. Dane trzy zmienne losowe X, Y, Z są parami niezależne. Czy stąd wynika, że wszystkie trzy zmienne losowe są niezależne?

Rozwiązanie. Załóżmy, że X, Y, Z są zmiennymi losowymi określonymi na tej samej przestrzeni probabilistycznej $\{\Omega, \mathscr{F}, P\}$. Zmienne losowe X, Y, Z są parami niezależne wtedy i tylko wtedy, gdy

$$\bigwedge_{x, y, z \in R} F_{12}(x, y)=F_1(x)F_2(y), \quad F_{23}(y, z)=F_2(y)F_3(z), \quad F_{13}(x, z)=F_1(x)F_3(z),$$

jest to równoważne niezależności parami następujących zdarzeń:

$$A_1=\{\omega: X(\omega)<x\}, \quad A_2=\{\omega: Y(\omega)<y\} \quad \text{i} \quad A_3=\{\omega: Z(\omega)<z\}.$$

Jak wiadomo (zad. 1.19) nie wynika stąd niezależność zespołowa tych zdarzeń. Nie musi więc zachodzić równość:

$$P(A_1 \cap A_2 \cap A_3)=P(A_1)P(A_2)P(A_3).$$

ZADANIE 5.35. Wykazać, że

$$D^2 \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} D^2 X_i + \sum_{i \neq j}^{n}\sum^{n} \operatorname{cov}(X_i, X_j).$$

Rozwiązanie. Korzystając z definicji wariancji i własności wartości przeciętnej, mamy

$$D^2 \sum_{i=1}^{n} X_i = E\Big(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} EX_i\Big)^2 = E\Big[\sum_{i=1}^{n}(X_i-EX_i)\Big]^2=$$

$$=E\Big[\sum_{i=1}^{n}(X_i-EX_i)^2+\sum_{i \neq j}^{n}\sum^{n}(X_i-EX_i)(X_j-EX_j)\Big]=$$

$$=\sum_{i=1}^{n}E(X_i-EX_i)^2+\sum_{i \neq j}^{n}\sum^{n}E(X_i-EX_i)(X_j-EX_j)=$$

$$=\sum_{i=1}^{n}D^2X_i+\sum_{i \neq j}^{n}\sum^{n}\operatorname{cov}(X_i, X_j).$$

ZADANIE 5.36. Trójwymiarowa zmienna losowa (X, Y, Z) ma rozkład o dystrybuancie:

$$F(x, y, z)=\begin{cases} 1-e^{-x}-e^{-2y}-e^{-3z}+e^{-(x+2y)}+e^{-(x+3z)}+e^{-(2y+3z)}-e^{-(x+2y+3z)} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{dla} \quad x>0, y>0, z>0, \\ 0 \qquad\qquad\qquad\qquad\qquad\quad \text{dla pozostałych } (x, y, z). \end{cases}$$

Wyznaczyć jednowymiarowe rozkłady brzegowe zmiennych losowych X, Y, Z oraz zbadać czy zmienne te są niezależne.

Rozwiązanie. Korzystając ze wzoru (5.8.2), otrzymujemy dystrybuanty rozkładów brzegowych:

$$F_1(x)=\lim_{\substack{y\to\infty \\ z\to\infty}} F(x, y, z)=\begin{cases} 1-e^{-x} & \text{dla} \quad x>0, \\ 0 & \text{dla pozostałych } x, \end{cases}$$

$$F_2(y)=\lim_{\substack{x\to\infty\\ z\to\infty}} F(x,y,z)=\begin{cases}1-e^{-2y} & \text{dla} \quad y>0,\\ 0 & \text{dla pozostałych } y,\end{cases}$$

$$F_3(z)=\lim_{\substack{x\to\infty\\ y\to\infty}} F(x,y,z)=\begin{cases}1-e^{-3z} & \text{dla} \quad z>0,\\ 0 & \text{dla pozostałych } z.\end{cases}$$

Łatwo zauważyć, że

$$\bigwedge_{(x,y,z)\in R^3} F(x,y,z)=F_1(x)F_2(y)F_3(z),$$

a więc X, Y, Z są niezależnymi zmiennymi losowymi.

ZADANIE 5.37. Badana cecha X w populacji ma rozkład równomierny na odcinku $(0, b)$. Niech $(X_1, \ldots, X_n)$ będzie próbą prostą n elementową z tej populacji. Wyznaczyć dystrybuantę rozkładu zmiennej losowej $Z=\max(X_1, \ldots, X_n)$.

Rozwiązanie. Jeśli $0<z<b$, to $F(z)=P(Z<z)=P[\max(X_1, \ldots, X_n)<z]=P(X_1<z, \ldots$ $\ldots, X_n<z)$, co wobec niezależności zmiennych jest równe

$$P(X_1<z)\ldots P(X_n<z)=[F_X(z)]^n,$$

ale dystrybuantę zmiennej losowej X na odcinku $(0, b)$ określa wzór $F_X(x)=\dfrac{x}{b}$, skąd $F(z)=\left(\dfrac{z}{b}\right)^n$. Ostatecznie

$$F(z)=\begin{cases}0 & \text{dla} \quad z\leqslant 0,\\ \left(\dfrac{z}{b}\right)^n & \text{dla} \quad 0<z\leqslant b,\\ 1 & \text{dla} \quad z>b.\end{cases}$$

ZADANIE 5.38. Wektor losowy (X, Y, Z) ma rozkład równomierny w obszarze $V=\{(x,y,z): 0\leqslant x\leqslant 1 \wedge 0\leqslant y\leqslant 1-x \wedge 0\leqslant z\leqslant 1\}$. Napisać równanie płaszczyzny regresji II-go rodzaju zmiennej X względem (Y, Z).

Rozwiązanie. Gęstość

$$f(x,y,z)=\begin{cases}\dfrac{1}{|V|} & \text{dla} \quad (x,y,z)\in V,\\ 0 & \text{poza tym},\end{cases}$$

gdzie $|V|$ jest objętością obszaru V (graniastosłupa) naszkicowanego na rysunku 5.21, skąd $|V|=\frac{1}{2}$, zaś

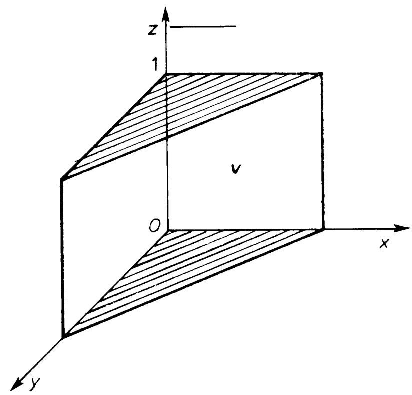
$$f(x,y,z)=\begin{cases}2 & \text{dla} \quad (x,y,z)\in V,\\ 0 & \text{poza tym}.\end{cases}$$

Korzystając ze wzoru (5.8.12), obliczymy wartości przeciętne w jednowymiarowych rozkładach brzegowych:

$$EY=EX=\int_0^1 dx \int_0^{1-x} dy \int_0^1 2x\,dz=2\int_0^1 x(1-x)\,dx=\tfrac{1}{3},$$

$$EZ=\int_0^1 dx \int_0^{1-x} dy \int_0^1 2z\,dz=\int_0^1 (1-x)\,dx=\tfrac{1}{2}.$$

Rys. 5.21. Obszar, nad którym jest skoncentrowana gęstość rozkładu równomiernego (zad. 5.38)



Aby wyznaczyć macierz kowariancji wyznaczymy momenty zwykłe rzędu drugiego

$$EY^2 = EX^2 = \int_0^1 dx \int_0^{1-x} dy \int_0^1 2x^2\,dz = 2\int_0^1 (1-x)x^2\,dx = \tfrac{1}{6},$$

$$EZ^2 = \int_0^1 dx \int_0^{1-x} dy \int_0^1 2z^2\,dz = \tfrac{2}{3}\int_0^1 (1-x)\,dx = \tfrac{1}{3},$$

$$E(XY) = \int_0^1 dx \int_0^{1-x} dy \int_0^1 2xy\,dz = \int_0^1 x(1-x)^2\,dx = \tfrac{1}{12},$$

$$E(YZ) = E(XZ) = \int_0^1 dx \int_0^{1-x} dy \int_0^1 2xz\,dz = \int_0^1 x(1-x)\,dx = \tfrac{1}{6}.$$

Stosując wzór (5.3.10), otrzymamy:

$$D^2X = D^2Y = \tfrac{1}{18}, \qquad D^2Z = \tfrac{1}{12},$$

$$\operatorname{cov}(X, Y) = -\tfrac{1}{36}, \qquad \operatorname{cov}(X, Z) = \operatorname{cov}(Y, Z) = 0.$$

Stąd macierz (5.8.16) jest postaci:

$$\mathbf{M} = \begin{bmatrix} \frac{1}{18} & -\frac{1}{36} & 0 \\ -\frac{1}{36} & \frac{1}{18} & 0 \\ 0 & 0 & \frac{1}{12} \end{bmatrix}.$$

Równanie płaszczyzny regresji II-go rodzaju zmiennej X względem (Y, Z) jest następujące:

$$x - EX = a_{12}(y - EY) + a_{13}(z - EZ),$$

gdzie a_{12}, a_{13} obliczamy według wzoru (5.8.20).

Obliczmy dopełnienia algebraiczne elementów pierwszego wiersza macierzy $\mathbf{M}$:

$$M_{11}^* = (-1)^2 \begin{vmatrix} \frac{1}{18} & 0 \\ 0 & \frac{1}{12} \end{vmatrix} = \frac{1}{18\cdot 12}, \qquad M_{12}^* = (-1)^3 \begin{vmatrix} -\frac{1}{36} & 0 \\ 0 & \frac{1}{12} \end{vmatrix} = \frac{1}{36\cdot 12},$$

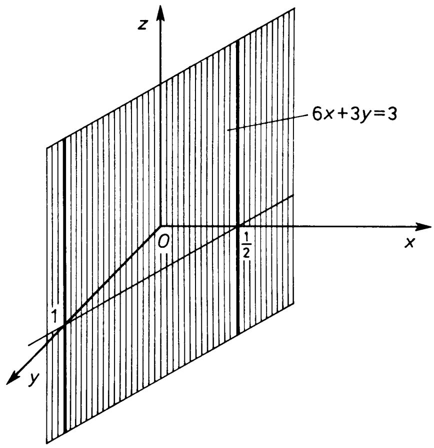
$$M_{13}^* = (-1)^4 \begin{vmatrix} -\frac{1}{36} & \frac{1}{18} \\ 0 & 0 \end{vmatrix} = 0,$$

skąd $a_{12} = -\frac{1}{2}$, $a_{13} = 0$.

Płaszczyzna regresji zmiennej X względem (Y, Z) ma równanie:

$$x-\tfrac{1}{3}=-\tfrac{1}{2}(y-\tfrac{1}{3}),$$

a więc jest płaszczyzną równoległą do osi Oz (rys. 5.22).



Rys. 5.22. Płaszczyzna regresji zmiennej X względem zmiennych (Y, Z) (zad. 5.38)

ZADANIE 5.39. Napisać gęstość (5.8.21) dla $n=2$ i porównać z gęstością dwuwymiarowego rozkładu normalnego (5.4.1).

Rozwiązanie. Rozpiszmy wzór (5.8.21) dla $n=2$

$$f(x_1, x_2)=\frac{1}{2\pi\sqrt{|\mathbf{M}|}}\exp\{-\tfrac{1}{2}[(x_1-EX_1)^2c_{11}+(x_1-EX_1)(x_2-EX_2)c_{12}+$$

$$+(x_2-EX_2)(x_1-EX_1)c_{21}+(x_2-EX_2)^2c_{22}]\},$$

gdzie macierz

$$\mathbf{M}=\begin{bmatrix}\sigma_1^2 & \operatorname{cov}(X_1, X_2)\\ \operatorname{cov}(X_2, X_1) & \sigma_2^2\end{bmatrix},$$

a $\sigma_i^2=D^2X_i$ dla $i=1, 2$.

Jeśli ρ jest współczynnikiem korelacji między zmiennymi X_1, X_2, to $\operatorname{cov}(X_1, X_2)=$ $=\operatorname{cov}(X_2, X_1)=\rho\sigma_1\sigma_2$, wówczas macierz kowariancji

$$\mathbf{M}=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}.$$

Obliczmy jej wyznacznik:

$$|\mathbf{M}|=\sigma_1^2\sigma_2^2-\rho^2\sigma_1^2\sigma_2^2=\sigma_1^2\sigma_2^2(1-\rho^2).$$

Dla $|\rho|<1$ istnieje macierz odwrotna do $\mathbf{M}$ o elementach:

$$c_{11}=\frac{1}{\sigma_1^2(1-\rho^2)},\qquad c_{12}=-\frac{\rho}{\sigma_1\sigma_2(1-\rho^2)},$$

$$c_{21}=-\frac{\rho}{\sigma_1\sigma_2(1-\rho^2)},\qquad c_{22}=\frac{1}{\sigma_2^2(1-\rho^2)}.$$

Stąd dla $|\rho|<1$ gęstość dwuwymiarowego rozkładu normalnego jest postaci:

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-EX_1)^2}{\sigma_1^2}-\right.\right.$$

$$\left.\left.-2\rho\frac{(x_1-EX_1)(x_2-EX_2)}{\sigma_1\sigma_2}+\frac{(x_2-EX_2)^2}{\sigma_2^2}\right]\right\},$$

otrzymaliśmy więc wzór identyczny z (5.4.1).

Zadanie 5.40. Trójwymiarowa zmienna losowa (X, Y, Z) ma rozkład normalny o macierzy kowariancji:

$$\mathbf{M}=\begin{bmatrix}1 & -1 & 1\\ -1 & 2 & -1\\ 1 & -1 & 2\end{bmatrix}$$

i wartościach przeciętnych $EX=EY=EZ=0$. Wyznaczyć jej gęstość prawdopodobieństwa.

Rozwiązanie. Obliczmy wartość wyznacznika

$$|\mathbf{M}|=\begin{vmatrix}1 & -1 & 1\\ -1 & 2 & -1\\ 1 & -1 & 2\end{vmatrix}=1\neq 0.$$

Można więc wyznaczyć macierz odwrotną do $\mathbf{M}$, której elementy obliczamy ze wzoru (5.8.20):

$$\mathbf{M}^{-1}=\begin{bmatrix}3 & 1 & -1\\ 1 & 1 & 0\\ -1 & 0 & 1\end{bmatrix};$$

stąd i ze wzoru (5.8.21) otrzymujemy:

$$f(x,y,z)=\frac{1}{(2\pi)^{3/2}}\exp\left[-\tfrac{1}{2}(3x^2+y^2+z^2+2xy-2xz)\right].$$

Podobnie jak dla dwuwymiarowej zmiennej losowej, w przypadku wielowymiarowym powierzchnie regresji I-go rodzaju dla rozkładu normalnego są odpowiednimi hiperpłaszczyznami regresji II-go rodzaju (zad. 5.23).

5.9. ZADANIA DO ROZWIĄZANIA

5.41. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład prawdopodobieństwa określony następująco: $P(X=1, Y=1)=0{,}2$, $P(X=1, Y=2)=0{,}3$, $P(X=3, Y=1)=0{,}4$, $P(X=3, Y=2)=0{,}1$.

a) Zapisać ten rozkład w tabeli,

b) Zbadać czy zmienne losowe X i Y są niezależne.

c) Wyznaczyć dystrybuantę i wartość przeciętną zmiennej losowej X.

d) Obliczyć wartość dystrybuanty w punkcie (2, 2).

5.42. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład podany w tabelce:

	x_i				
y_k	3	3,5	4	4,5	5
3	0,16	0	0	0	0
3,5	0,12	0,04	0	0	0
4	0,08	0,08	0,08	0	0
4,5	0	0,12	0,06	0,05	0,04
5	0	0	0,01	0,08	0,08

gdzie X jest średnią oceną w sesji egzaminacyjnej w I semestrze dla losowo wybranego studenta, zaś Y – średnią oceną w sesji tego studenta w VI semestrze.

a) Obliczyć współczynnik korelacji tych zmiennych losowych.

b) Napisać równanie prostej regresji II-go rodzaju zmiennej losowej Y względem X.

c) Wyznaczyć linię regresji I-go rodzaju zmiennej losowej Y względem X.

5.43. Rzucamy 5 razy monetą. Niech X oznacza liczbę orłów otrzymanych w tych rzutach, Y – liczbę serii orłów, Z zaś – długość najdłuższej serii orłów. Wypisać wszystkie zdarzenia elementarne w opisanym doświadczeniu. Wyznaczyć:

a) rozkład dwuwymiarowej zmiennej losowej (X, Y),

b) rozkłady brzegowe poszczególnych zmiennych i ich wartości przeciętne,

c) rozkład zmiennej losowej Z i obliczyć przeciętną długość najdłuższej serii orłów,

d) rozkład dwuwymiarowy zmiennej losowej (X, Z) oraz obliczyć $P(X=3, Z\leqslant 2)$.

(w ciągu elementów dwóch rodzajów każdy maksymalny podciąg elementów jednego rodzaju nazywamy serią).

5.44. W 10-cio elementowej partii pewnego towaru są 2 sztuki wadliwe. Wylosowano bez zwrotu 2 sztuki. Niech zmienna losowa X przyjmuje wartości równe liczbie sztuk wadliwych wśród 2 wylosowanych sztuk, Y zaś przyjmuje wartość 1, jeśli pierwsza wylosowana sztuka jest wadliwa, oraz 0, jeśli nie jest wadliwa.

a) Wyznaczyć rozkład dwuwymiarowej zmiennej losowej (X, Y).

b) Zbadać czy zmienne losowe X i Y są niezależne.

c) Obliczyć $E(X|y=1)$ i $D^2(X|y=1)$.

d) Obliczyć $P(X+Y=2)$ oraz $E(X+Y)$.

5.45. Dobrać tak stałą c, by funkcja

$$f(x, y)=\begin{cases} cxy(2-x-y) & \text{dla} \quad 0\leqslant x\leqslant 1, \quad 0\leqslant y\leqslant 1, \\ 0 & \text{dla pozostałych } (x, y) \end{cases}$$

była gęstością dwuwymiarowej zmiennej losowej (X, Y). Wyznaczyć:

a) jej dystrybuantę, b) zbadać czy zmienne losowe X, Y są niezależne, c) wyznaczyć $f(y|x)$.

5.46. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład o gęstości:

$$f(x, y)=\begin{cases} \frac{1}{2}(x+y)e^{-(x+y)} & \text{dla} \quad x\geqslant 0, \quad y\geqslant 0, \\ 0 & \text{dla pozostałych } (x, y). \end{cases}$$

Wyznaczyć: a) dystrybuantę tej zmiennej, b) $P(|X|<1, |Y|>1)$, c) rozkład brzegowy zmiennej losowej X.

5.47. Koszt K pewnej operacji jest wprost proporcjonalny do kwadratu całkowitego czasu jej zakończenia. Operacja jest dwuetapowa, przy czym czasy X zakończenia pierwszego etapu oraz Y drugiego etapu są zmiennymi losowymi skorelowanymi o wartościach przeciętnych: $EX=\alpha_1$, $EY=\alpha_2$, wariancjach $D^2X=\sigma_1^2$, $D^2Y=\sigma_2^2$ oraz współczynniku korelacji ρ. Obliczyć wartość przeciętną kosztu K.

5.48. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład równomierny, tj. o stałej gęstości w obszarze $D=\{(x, y): 0\leqslant x\leqslant 2 \wedge 0\leqslant y\leqslant 2-x\}$.

a) Wyznaczyć dystrybuantę tego rozkładu, b) obliczyć współczynnik korelacji.

5.49. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład równomierny w obszarze ograniczonym osiami współrzędnych oraz wykresem funkcji $y=2\cos\frac{1}{2}x$ dla $0\leqslant x\leqslant \pi$. Wyznaczyć rozkład odległości losowego punktu (X, Y) płaszczyzny od osi odciętych.

5.50. Rozmiary powodzi charakteryzuje się za pomocą dwóch zmiennych: czasu trwania powodzi X i całkowitych opadów atmosferycznych Y. Niech f_1 będzie gęstością zmiennej losowej X określoną wzorem:

$$f_1(x)=\begin{cases} c(x-2)^2 & \text{dla} \quad 2<x\leqslant 7, \\ c(12-x)^2 & \text{dla} \quad 7<x<12, \\ 0 & \text{dla pozostałych } x. \end{cases}$$

Wyznaczyć stałą c, a następnie wiedząc, że

$$f(y|x)=\begin{cases} \frac{1}{3} & \text{dla} \quad \frac{1}{2}x-1\leqslant y\leqslant \frac{1}{2}x+2, \\ 0 & \text{dla pozostałych } y, \end{cases}$$

wyznaczyć gęstość dwuwymiarowej zmiennej losowej (X, Y).

5.51. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład normalny o gęstości:

$$f(x, y)=c\exp\left[-\tfrac{2}{3}\left(x^2-\tfrac{1}{2}xy+\tfrac{1}{4}y^2\right)\right].$$

Wyznaczyć stałą c oraz równanie linii regresji I-go rodzaju zmiennej losowej Y względem X.

5.52. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład o gęstości:

$$f(x, y)=\frac{1}{20\pi}\exp\left[-\tfrac{1}{2}\left(\tfrac{1}{4}x^2+\tfrac{1}{25}y^2\right)\right].$$

a) Zbadać czy zmienne losowe X, Y są niezależne.

b) Obliczyć $P(-1<X<2, 0<Y<3)$.

c) Obliczyć $P[(X, Y)\in D]$, gdzie $D=\{(x, y): \frac{1}{4}x^2+\frac{1}{25}y^2\leqslant 1\}$.

5.53. Zmienna losowa (X, Y) ma dwuwymiarowy rozkład normalny o parametrach: $\mu_1=2$, $\mu_2=1$, $\sigma_1=1{,}5$, $\sigma_2=0{,}8$, $\rho=0{,}6$. Podać gęstość tego rozkładu oraz napisać równanie prostej regresji zmiennej losowej X względem Y.

5.54. Niech X, Y będą niezależnymi zmiennymi losowymi o tym samym rozkładzie $N(m, \sigma)$. Wyznaczyć współczynnik korelacji między zmiennymi $U=aX+bY$, $V=aX-bY$, $a^2+b^2>0$.

5.55. Wykazać, że wariancja iloczynu dwóch niezależnych zmiennych losowych X, Y wyraża się równością:

$$D^2(XY)=\sigma_1^2\sigma_2^2+\alpha_1^2\sigma_2^2+\alpha_2^2\sigma_1^2,$$

gdzie $\alpha_1=EX$, $\alpha_2=EY$, $\sigma_1^2=D^2X$, $\sigma_2^2=D^2Y$.

5.56. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład o gęstości

$$f(x, y)=\frac{\sqrt{5}}{2\pi}\exp\left[-\tfrac{1}{2}(x^2+2xy+5y^2)\right].$$

a) Zbadać czy X i Y są niezależnymi zmiennymi losowymi.

b) Wyznaczyć rozkład prawdopodobieństwa zmiennej losowej $Z=X+Y$.

c) Wyznaczyć gęstość łącznej zmiennej losowej (U, V), gdzie $U=X+Y$, $V=X-Y$.

5.57. Niech (X, Y) ma dwuwymiarowy rozkład normalny (5.4.1). Wykazać, że zmienne losowe $U=\dfrac{X-\mu_1}{\sigma_1\sqrt{1-\sigma^2}}-\dfrac{\rho(Y-\mu_2)}{\sigma_2\sqrt{1-\rho^2}}$, $V=\dfrac{Y-\mu_2}{\sigma_2}$ są niezależne i że ich łączna gęstość ma postać $g(u, v)=\dfrac{1}{2\pi}\exp\left[-\tfrac{1}{2}(u^2+v^2)\right]$.

5.58. Wykazać, że współczynnik korelacji zmiennych losowych X i Y jest równy kowariancji odpowiednich standaryzowanych zmiennych losowych: $\dfrac{X-\alpha_1}{\sigma_1}$, $\dfrac{Y-\alpha_2}{\sigma_2}$.

5.59. Niech X i Y będą dowolnymi zmiennymi losowymi oraz $Z=aX+bY$, $W=cX+$ $+dY$.

a) Wyznaczyć kowariancję zmiennych Z, W za pomocą momentów zmiennych X i Y.

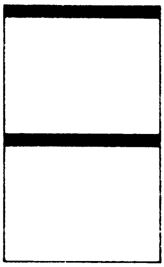
b) Dobrać tak stałe a, b, c i d, by zmienne Z, W były nieskorelowane.

5.60. W konstrukcji budowlanej przedstawionej na rys. 5.23 całkowita deformacja (rys. 5.24) Y punktów najwyżej położonych jest sumą deformacji na poszczególnych piętrach X_1 i X_2 działających niezależnie od siebie. Zmienne losowe X_1, X_2 mają wartości przeciętne i wariancje odpowiednio równe: α_1, σ_1^2 i α_2, σ_2^2.

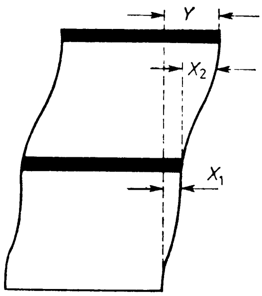
a) Znaleźć wartość przeciętną i wariancję zmiennej losowej Y.

b) Przeprowadzić dyskusję współczynnika korelacji zmiennych losowych Y i X_2 w przypadku, gdy σ_2 jest dużo większe od σ_1.

c) Wyznaczyć współczynnik korelacji zmiennych losowych Y i X_2, jeśli X_1 i X_2 są skorelowane o współczynniku korelacji ρ.



Rys. 5.23. Konstrukcja budowlana niezdeformowana (zad. 5.60)



Rys. 5.24. Konstrukcja budowlana zdeformowana (zad. 5.60)

5.61. Odległość X budynku od epicentrum następnego trzęsienia ziemi w promieniu r_0 km jest zmienną losową, która ma rozkład o gęstości:

$$f(x)=\begin{cases} \dfrac{2}{r_0^2}x & \text{dla} \quad 0\leqslant x\leqslant r_0, \\ 0 & \text{dla pozostałych } x. \end{cases}$$

Intensywność trzęsienia ziemi (w stopniach skali Richtera) ma rozkład o gęstości:

$$f(y)=\begin{cases} \frac{3}{64}(9-y)^2 & \text{dla} \quad 5\leqslant y\leqslant 9, \\ 0 & \text{dla pozostałych } y. \end{cases}$$

Zakładając, że X i Y są niezależnymi zmiennymi losowymi obliczyć prawdopodobieństwo, że następne trzęsienie ziemi będzie o intensywności większej niż 8, a jego epicentrum będzie leżeć nie dalej niż w odległości $\frac{1}{2}r_0$ od budynku.

5.62. Niech zmienne losowe X, Y będą sumami n nieskorelowanych zmiennych losowych postaci: $X=X_1+\ldots+X_k+X_{k+1}+\ldots+X_n$, $Y=X_1+\ldots+X_k+Y_{k+1}+\ldots+Y_n$ dla $1\leqslant k\leqslant n$ o jednakowych wartościach przeciętnych $EX_i=EY_j=\alpha$ i wariancjach $D^2X_i=$ $=D^2Y_j=\sigma^2$ dla $i=1,\ldots,n$, $j=k+1,\ldots,n$. Wykazać, że współczynnik korelacji zmiennych X, Y jest równy k/n.

5.63. Niech zmienne losowe X, Y będą sumami skończonej liczby nieskorelowanych zmiennych losowych postaci: $X=X_1+\ldots+X_k+X_{k+1}+\ldots+X_n$, $Y=X_1+\ldots+X_k+$ $+Y_{k+1}+\ldots+Y_m$ o jednakowej wartości przeciętnej $EX_i=EY_l=\alpha$ i wariancji $D^2X_i=$ $=D^2Y_l=\sigma^2$ dla $i=1,\ldots,n$, $l=k+1,\ldots,m$. Wykazać, że współczynnik korelacji zmiennych losowych X, Y jest równy $k/\sqrt{mn}$.

5.64. Wykazać, że jeśli współczynnik korelacji zmiennych losowych X i Y jest równy ρ, to współczynnik korelacji zmiennych losowych: $U=aX+b$, $V=cY+d$, gdzie a, b, c, d stałe rzeczywiste oraz $ac\neq 0$, jest równy $\rho\,\mathrm{sign}\,(ac)$.

5.65. Czas napełniania X pewnego zbiornika wodą jest zmienną losową o gęstości:

$$f_1(x)=\begin{cases} \lambda e^{-\lambda x} & \text{dla} \quad x\geqslant 0, \quad \lambda>0, \\ 0 & \text{dla pozostałych } x. \end{cases}$$

Czas opróżniania Y tego zbiornika jest zmienną losową o gęstości:

$$f_2(y)=\begin{cases} ve^{-vy} & \text{dla} \quad y\geqslant 0, \quad v>0, \\ 0 & \text{dla pozostałych } y. \end{cases}$$

Zakładając, że X i Y są niezależnymi zmiennymi losowymi wyznaczyć rozkład zmiennej losowej $Z=X-Y$.

5.66. X_1 i X_2 są niezależnymi zmiennymi losowymi o jednoparametrowych rozkładach gamma:

$$f_i(x_i)=\begin{cases} \dfrac{1}{\Gamma(p_i)}x_i^{p_i-1}e^{-x_i} & \text{dla} \quad x_i>0, \\ 0 & \text{dla pozostałych } x_i, \end{cases} \qquad i=1,2,$$

gdzie $p_i>0$. Wyznaczyć gęstość zmiennej losowej $Z=\dfrac{X_1}{X_1+X_2}$.

5.67. X i Y są niezależnymi, standaryzowanymi zmiennymi losowymi o rozkładzie normalnym. Wyznaczyć rozkład zmiennej losowej $Z=X^2+Y^2$.

5.68. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład normalny o gęstości $f(x, y)=$ $=\dfrac{1}{2\pi}\exp\left(-\dfrac{x^2+y^2}{2}\right)$. Wyznaczyć: a) gęstość dwuwymiarowej zmiennej losowej (R, Φ), jeśli $X=R\cos\Phi$, $Y=R\sin\Phi$, b) gęstość rozkładu brzegowego zmiennej losowej R (tj. odległości losowego punktu (X, Y) od początku układu współrzędnych).

5.69. Badamy pewną partię towaru o wadliwości p. Liczbę elementów wylosowaną do momentu trafienia na pierwszą sztukę wadliwą oznaczmy przez X, dodatkową zaś liczbę elementów otrzymaną do momentu wylosowania drugiej sztuki wadliwej przez Y. Zakładając, że X i Y są niezależnymi zmiennymi losowymi o rozkładach geometrycznych:

$$P(X=x)=(1-p)^{x-1}p, \qquad x\in N,$$

$$P(Y=y)=(1-p)^{y-1}p, \qquad y\in N,$$

wyznaczyć rozkład sumy tych zmiennych: $Z=X+Y$.

5.70. Każdy z dwóch odcinków o długości a podzielono losowo wybranym punktem na 2 części. Zakładając, że długości krótszych części są zmiennymi losowymi o rozkładzie jednostajnym, obliczyć prawdopodobieństwo, że suma długości krótszych odcinków spełnia nierówność: $\frac{1}{4}a\leqslant S\leqslant\frac{3}{4}a$.

5.71. Dwuwymiarowa zmienna losowa (X, Y) ma rozkład o gęstości:

$$f(x, y)=\begin{cases} \dfrac{1}{\Gamma(p)\Gamma(r)}x^{p-1}(y-x)^{r-1}e^{-y} & \text{dla} \quad 0<x<y<\infty, \\ 0 & \text{dla pozostałych } x, y, \end{cases}$$

gdzie $p>0$, $r>0$. Wyznaczyć rozkłady brzegowe.

5.72. Niech X i Y będą niezależnymi zmiennymi losowym o jednoparametrowych rozkładach gamma postaci:

$$f_1(x)=\begin{cases}\dfrac{1}{\Gamma(p)}x^{p-1}e^{-x} & \text{dla} \quad x>0,\\ 0 & \text{dla pozostałych } x,\end{cases}$$

$$f_2(y)=\begin{cases}\dfrac{1}{\Gamma(p+0{,}5)}y^{p-0{,}5}e^{-y} & \text{dla } y>0,\\ 0 & \text{dla pozostałych } y.\end{cases}$$

Wyznaczyć gęstość zmiennej losowej: $Z=2\sqrt{XY}$.

5.73. X i Y są niezależnymi zmiennymi losowymi o gęstościach

$$f_1(x)=\begin{cases}\lambda e^{-\lambda x} & \text{dla} \quad x\geqslant 0, \quad \lambda>0,\\ 0 & \text{dla pozostałych } x,\end{cases}$$

$$f_2(y)=\begin{cases}1 & \text{dla} \quad 0\leqslant y\leqslant 1,\\ 0 & \text{dla pozostałych } y.\end{cases}$$

Wyznaczyć rozkład sumy tych zmiennych $Z=X+Y$.

5.74. Gęstość dwuwymiarowej zmiennej losowej (X, Y)

$$f(x,y)=\begin{cases}x+y & \text{dla} \quad 0\leqslant x\leqslant 1, \quad 0\leqslant y\leqslant 1,\\ 0 & \text{poza tym}.\end{cases}$$

Obliczyć współczynnik korelacji tych zmiennych losowych.

5.75. Niech X będzie wielkością zapotrzebowania na pewien towar, Y ceną jednostki tego towaru. Zakładamy, że X i Y są niezależnymi zmiennymi losowymi o gęstościach odpowiednio równych:

$$f_1(x)=\begin{cases}\frac{2}{9}x & \text{dla} \quad 0\leqslant x\leqslant 3,\\ 0 & \text{dla pozostałych } x,\end{cases} \qquad f_2(y)=\begin{cases}\dfrac{1}{\varepsilon} & \text{dla} \quad 0\leqslant y\leqslant \varepsilon,\\ 0 & \text{dla pozostałych } y,\end{cases}$$

gdzie $\varepsilon>0$. a) Wyznaczyć rozkład prawdopodobieństwa zmiennej losowej $Z=XY$, b) obliczyć D^2Z.

5.76. X i Y są niezależnymi zmiennymi losowymi o rozkładach prawdopodobieństwa: $P(X=0)=P(X=1)=\frac{1}{2}$ (rozkład dwupunktowy) oraz

$$f(y)=\begin{cases}1 & \text{dla} \quad 0\leqslant y\leqslant 1,\\ 0 & \text{dla pozostałych } y\end{cases}$$

(rozkład równomierny). Wyznaczyć rozkład zmiennej losowej $Z=X+Y$.

5.77. Wektor losowy (X, Y, Z) ma rozkład o gęstości:

$$f(x,y,z)=\begin{cases}\frac{1}{12}xy^2z^3\exp[-(x+y+z)] & \text{dla} \quad x>0, \quad y>0, \quad z>0,\\ 0 & \text{poza tym}.\end{cases}$$

a) Wyznaczyć gęstości rozkładów brzegowych jednowymiarowych.

b) Zbadać czy zmienne losowe X, Y i Z są niezależne.

c) Obliczyć $E(Z\,|\,X=z,\ Y=y)$.

5.78. Trójwymiarowa zmienna losowa (X, Y, Z) ma rozkład równomierny w obszarze $V=\{(x, y, z): x^2+y^2 \leqslant 4 \wedge 0 \leqslant z \leqslant 1\}$.

a) Wyznaczyć gęstość rozkładu brzegowego zmiennej Z i obliczyć wartość przeciętną $E(Z)$.

b) Wyznaczyć gęstość rozkładu warunkowego $f(x \mid y, z)$, oraz obliczyć $E(X \mid Y=y, Z=z)$.

c) Napisać równanie powierzchni regresji I-go rodzaju zmiennej losowej X względem Y i Z.

5.79. W urnie znajduje się 10 kul ponumerowanych od 1 do 10. Losujemy 3 razy po jednej kuli ze zwrotem. Niech wartością zmiennej losowej X_i, $i=1, 2, 3$, będzie numer kuli wylosowanej za i-tym razem.

a) Wyznaczyć łączny rozkład zmiennych (X_1, X_2, X_3).

b) Obliczyć $E(X_1^2|X_2=x_2, X_3=x_3)$.

5.80. Rzucamy n razy kostką sześcienną. Zmienna losowa X_i, $i=1, \ldots, n$, przyjmuje wartości równe liczbie oczek wyrzuconych za i-tym razem.

a) Wyznaczyć rozkład n-wymiarowej zmiennej losowej $(X_1, \ldots, X_n)$.

b) Obliczyć wariancję $D^2(X_1 X_2|X_3=x_3, \ldots, X_n=x_n)$.

5.81. Trójwymiarowa zmienna losowa (X, Y, Z) ma rozkład normalny o zerowym wektorze wartości przeciętnych i macierzy kowariancji:

$$\mathbf{M}=\begin{bmatrix} 7 & 3 & 2 \\ 3 & 4 & 1 \\ 2 & 1 & 2 \end{bmatrix}.$$

a) Napisać gęstość tego rozkładu.

b) Wyznaczyć równanie płaszczyzny regresji II-go rodzaju zmiennej Y względem X i Z.

c) Obliczyć wartość przeciętną i wariancję zmiennej losowej $W=2X+Z-Y$.

5.82. Trójwymiarowa zmienna losowa (X, Y, Z) ma rozkład o gęstości:

$$f(x, y, z)=\frac{1}{(2\pi)^{3/2}}\exp\left[-(x^2+y^2+\tfrac{1}{2}z^2+xy+yz)\right].$$

a) Wyznaczyć macierz kowariancji.

b) Obliczyć współczynnik korelacji zmiennych X i Y.

5.83. Trójwymiarowa zmienna losowa (X, Y, Z) ma rozkład o gęstości:

$$f(x, y, z)=\frac{1}{4\pi\sqrt{6\pi}}\exp\left\{-\tfrac{1}{12}\left[6x^2+4(y-1)^2+(z+2)^2-2(y-1)(z+2)\right]\right\}.$$

a) Wyznaczyć macierz kowariancji.

b) Obliczyć $E(YZ)$.

c) Obliczyć $E[Y(Y-Z)]$.

5.84. Wyznaczyć wartość przeciętną i wariancję sumy oczek wyrzuconych na n kostkach sześciennych.

5.85. Zmienna losowa (X, Y, Z) ma trójwymiarowy rozkład normalny o gęstości:

$$f(x, y, z) = c\exp\left[-\tfrac{1}{2}(2x^2 + y^2 + 3z^2 - 2xy - 2yz + 4xz)\right].$$

Wyznaczyć stałą c oraz macierz kowariancji.

5.86. Niech zmienna losowa X ma rozkład $N\left(0, \frac{1}{\sqrt{t}}\right)$, $t > 0$. Traktując parametr t jako zmienną losową o rozkładzie gamma o gęstości

$$g(t) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} t^{-1/2} e^{-t/2} & \text{dla} \quad t > 0, \\ 0 & \text{dla pozostałych } t, \end{cases}$$

wyznaczyć rozkład złożony z powyższych rozkładów.

5.87. Czy mieszaninę (3.2.2) dwóch rozkładów typu ciągłego można traktować jako złożenie rozkładów (5.6.5)?

5.88. Pewien aparat elektroniczny dokonuje pomiaru pięciu różnych czasów $X_1, \ldots, X_5$. Zakładając, że są one niezależnymi zmiennymi losowymi o jednakowej gęstości Rayleigha postaci

$$f(x) = \begin{cases} \frac{1}{4} x e^{-x^2/8} & \text{dla} \quad x > 0, \\ 0 & \text{dla pozostałych } x, \end{cases}$$

obliczyć prawdopodobieństwo, że $\max(X_1, \ldots, X_5) > 4$.

5.89. Niech (X, Y) ma rozkład równomierny (o stałej gęstości) w kole: $x^2 + y^2 \leqslant 1$. Wyznaczyć rozkład dwuwymiarowej zmiennej losowej (R, Φ), gdzie $R = \sqrt{X^2 + Y^2}$, $\Phi = \operatorname{arc\,tg} \dfrac{Y}{X}$.

5.90. X, Y są niezależnymi zmiennymi losowymi o jednakowym rozkładzie wykładniczym (2.8.7). Wykazać, że zmienne $X + Y$, $\dfrac{X}{Y}$ są niezależnymi zmiennymi losowymi.

5.91. X, Y są niezależnymi zmiennymi losowymi o rozkładzie χ^2 z n_1 i n_2 stopniami swobody. Wykazać, że zmienne: $X + Y$ oraz $\dfrac{X}{Y}$ są niezależne.

5.92. Składowe V_X, V_Y, V_Z prędkości V cząsteczki gazu są niezależnymi zmiennymi losowymi o jednakowym rozkładzie $N\left(0, \sqrt{k\dfrac{T}{\mu}}\right)$, gdzie k jest stałą Boltzmana, T – temperaturą (w stopniach Kelwina) środowiska, w którym znajduje się cząsteczka gazu, zaś μ – masą cząsteczki. Wyznaczyć gęstość prawdopodobieństwa zmiennej losowej $V = \sqrt{V_X^2 + V_Y^2 + V_Z^2}$.

5.93. Wykazać, że zmienna losowa $Y = \sum_{i=1}^{n} X_i^2$, gdzie X_i są niezależnymi zmiennymi losowymi o tym samym rozkładzie $N(0, 1)$ (2.8.24), ma rozkład χ^2 o n stopniach swobody (2.8.16).

5.94. Wykazać, że jeśli $X_1, \ldots, X_N$ są zmiennymi losowymi o jednakowych wartościach przeciętnych $EX_i = \alpha$ dla $i=1, \ldots, N$, to $E\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \alpha$, niezależnie od tego czy N jest ustaloną liczbą naturalną, czy zmienną losową dyskretną.

5.95. Obliczyć prawdopodobieństwo, że pierwiastki równania

$$x^2 + 2Px + Q = 0$$

są rzeczywiste, przy założeniu, że P i Q są niezależnymi zmiennymi losowymi o rozkładzie równomiernym w przedziale $\langle -a, a\rangle$ bądź $\langle -b, b\rangle$ odpowiednio.

Odpowiedzi

5.41. a)

	x_i	
y_k	1	3
1	0,2	0,4
2	0,3	0,1

b) zależne;

c) $F(x) = \begin{cases} 0 & \text{dla} \quad x \leqslant 1, \\ \frac{1}{2} & \text{dla} \quad 1 < x \leqslant 3, \\ 1 & \text{dla} \quad x > 3, \end{cases}$ $\quad EX = 2$,

d) $F(2, 2) = 0,2$.

5.42. a) $\rho = 0,8$; b) $y - 4,1 = 0,77\,(x - 3,7)$;

c) „krzywa" regresji I-go rodzaju jest tutaj zbiorem 5 punktów o współrzędnych: $x_k, E(Y|x_k)$, $k=1, \ldots, 5$: (3, 3,39), (3,5, 4,17), (4, 4,27), (4,5, 4,81), (5, 4,83).

5.43. a)

	x_i					
y_k	0	1	2	3	4	5
0	$\frac{1}{32}$	0	0	0	0	0
1	0	$\frac{5}{32}$	$\frac{4}{32}$	$\frac{3}{32}$	$\frac{2}{32}$	$\frac{1}{32}$
2	0	0	$\frac{6}{32}$	$\frac{6}{32}$	$\frac{3}{32}$	0
3	0	0	0	$\frac{1}{32}$	0	0

b)

x_i	0	1	2	3	4	5
p_i	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

, $EX = 2,5$.

y_k	0	1	2	3
p_k	$\frac{1}{32}$	$\frac{15}{32}$	$\frac{15}{32}$	$\frac{1}{32}$

, $EY = 1,5$.

c)

z_i	0	1	2	3	4	5
p_i	$\frac{1}{32}$	$\frac{12}{32}$	$\frac{11}{32}$	$\frac{5}{32}$	$\frac{2}{32}$	$\frac{1}{32}$

d)

	x_i					
z_k	0	1	2	3	4	5
0	$\frac{1}{32}$	0	0	0	0	0
1	0	$\frac{5}{32}$	$\frac{6}{32}$	$\frac{1}{32}$	0	0
2	0	0	$\frac{4}{32}$	$\frac{6}{32}$	$\frac{1}{32}$	0
3	0	0	0	$\frac{3}{32}$	$\frac{2}{32}$	0
4	0	0	0	0	$\frac{2}{32}$	0
5	0	0	0	0	0	$\frac{1}{32}$

$$P(X=3, Z\leqslant 2)=\tfrac{1}{32}+\tfrac{6}{32}=\tfrac{7}{32}.$$

5.44. a)

	x_i		
y_k	0	1	2
0	$\frac{56}{90}$	$\frac{16}{90}$	0
1	0	$\frac{16}{90}$	$\frac{2}{90}$

b) X, Y zależne.

c) $E(X|y=1)=\frac{10}{9}$

$D^2(X|y=1)=\frac{8}{81}$,

d) $P(X+Y=2)=\frac{8}{45}$

$E(X+Y)=0{,}6$.

5.45. $c=6$.

a) $F(x,y)=\begin{cases} x^2y^2(3-x-y) & \text{dla} \quad 0<x\leqslant 1 \wedge 0<y\leqslant 1, \\ x^2(2-x) & \text{dla} \quad 0<x\leqslant 1 \wedge y>1, \\ y^2(2-y) & \text{dla} \quad x>1 \wedge 0<y\leqslant 1, \\ 1 & \text{dla} \quad x>1 \wedge y>1, \\ 0 & \text{dla pozostałych } (x,y); \end{cases}$

b) zmienne losowe X, Y są zależne;

c) $\bigwedge\limits_{0\leqslant x\leqslant 1} f(y|x)=\begin{cases} 6y\dfrac{2-x-y}{4-3x} & \text{dla} \quad 0\leqslant y\leqslant 1, \\ 0 & \text{poza tym}. \end{cases}$

5.46.

a) $F(x,y)=\begin{cases} \frac{1}{2}(1-e^{-y})[1-(x+1)e^{-x}]+\frac{1}{2}(1-e^{-x})[1-(y+1)e^{-y}] & \text{dla } x\geqslant 0,\ y\geqslant 0, \\ 0 & \text{dla pozostałych } (x,y); \end{cases}$

b) $P(|X|<1, |Y|>1)=\frac{1}{2}(1-2e^{-2})$;

c) $f_1(x)=\begin{cases} \frac{1}{2}e^{-x}(x+1) & \text{dla} \quad x>0, \\ 0 & \text{dla pozostałych } x. \end{cases}$

5.47. $K=aT^2$, T – całkowity czas zakończenia operacji, a – współczynnik proporcjonalności, $T=X+Y$. $EK=a(\sigma_1^2+\alpha_1^2+\sigma_2^2+\alpha_2^2+2\rho\sigma_1\sigma_2+2\alpha_1\alpha_2)$.

5.48. Ponieważ $f(x,y)=\begin{cases} c & \text{dla} \quad (x,y)\in D, \\ 0 & \text{poza tym}, \end{cases}$ $c\iint\limits_D dx\,dy=1$, przy czym $\iint\limits_D dx\,dy=|D|=2$, więc $c=\frac{1}{2}$, skąd

$$f(x,y)=\begin{cases} \frac{1}{2} & \text{dla} \quad (x,y)\in D, \\ 0 & \text{poza tym}. \end{cases}$$

a) $$F(x,y)=\begin{cases} \frac{1}{2}xy & \text{dla} \quad 0<x\leqslant 2\wedge 0<y\leqslant 2-x, \\ \frac{1}{2}xy-\frac{1}{4}(x+y-2)^2 & \text{dla} \quad 0<x\leqslant 2\wedge 2-x<y\leqslant 2, \\ \frac{1}{4}(4-y)x & \text{dla} \quad 0<x\leqslant 2\wedge y>2, \\ \frac{1}{4}(4-y)y & \text{dla} \quad x>2\wedge y\leqslant 2, \\ 1 & \text{dla} \quad x>2\wedge y>2, \\ 0 & \text{dla pozostałych } (x,y); \end{cases}$$

b) $\alpha_1=\alpha_2=\frac{2}{3}$; $\sigma_1=\sigma_2=\sqrt{\frac{2}{9}}$, $\operatorname{cov}(X,Y)=-\frac{1}{9}$, $\rho=-\frac{1}{2}$.

5.49.

$$f_2(y)=\begin{cases} \frac{1}{2}\arccos\frac{1}{2}y & \text{dla} \quad 0\leqslant y\leqslant 2, \\ 0 & \text{dla pozostałych } y. \end{cases}$$

5.50. $c=0{,}012$. Należy rozpatrzyć dwa obszary, w których $f(x,y)\neq 0$:

$$D_1=\{(x,y):\ 2<x<7\wedge \tfrac{1}{2}x-1\leqslant y\leqslant \tfrac{1}{2}x+2\},$$

$$D_2=\{(x,y):\ 7<x<12\wedge \tfrac{1}{2}x-1\leqslant y\leqslant \tfrac{1}{2}x+2\},$$

mamy wówczas

$$f(x,y)=\begin{cases} 0{,}004(x-2)^2 & \text{dla} \quad (x,y)\in D_1, \\ 0{,}004(12-x)^2 & \text{dla} \quad (x,y)\in D_2, \\ 0 & \text{dla pozostałych } (x,y). \end{cases}$$

5.51. $c=\dfrac{1}{2\pi\sqrt{3}}$, równanie linii regresji $y=x$.

5.52. a) X, Y niezależne o rozkładach odpowiednio $N(0,2)$ i $N(0,5)$;

b) $P(-1<X<2, 0<Y<3)=P(-\frac{1}{2}<\frac{1}{2}X<1)\,P(0<\frac{1}{5}Y<0{,}6)=[\Phi(1)+\Phi(\frac{1}{2})-1][\Phi(0{,}6)-$ $-\Phi(0)]=0{,}1208$ z tablic rozkładu $N(0,1)$;

c) $P[(X,Y)\in D]=\displaystyle\iint\limits_D \frac{1}{2\pi}\exp[-\tfrac{1}{2}(\tfrac{1}{4}x^2+\tfrac{1}{25}y^2)]\,dx\,dy$ po zamianie zmiennych $x=$ $=2r\cos\varphi$, $y=2r\sin\varphi$, $|J|=10r$, $D=\{(r,\varphi):\ 0<r<1\wedge 0\leqslant\varphi<2\pi\}$

$$P[(X,Y)\in D]=\frac{1}{2\pi}\int\limits_0^1\left(\int\limits_0^{2\pi} re^{-r^2/2}\,d\varphi\right)dr=1-e^{-1/2}.$$

5.53. $f(x,y)=\dfrac{1}{,92\pi}\exp\left\{-\dfrac{1}{1{,}28}\left[\dfrac{(x-2)^2}{2{,}25}-(x-2)(y-1)+\dfrac{(y-1)^2}{0{,}64}\right]\right\}$, równanie prostej regresji: $x-2=1{,}125(y-1)$.

5.54. $\rho_{uv}=(a^2-b^2)/(a^2+b^2)$.

5.55. $D^2(XY)=E(XY)^2-E^2(XY)=E(X^2)E(Y^2)-\alpha_1^2\alpha_2^2=(\sigma_1^2+\alpha_1^2)(\sigma_2^2+\alpha_2^2)-\alpha_1^2\alpha_2^2=$ $=\sigma_1^2\sigma_2^2+\alpha_1^2\sigma_2^2+\alpha_2^2\sigma_1^2$.

5.56. a) Nie, ponieważ $\rho=-\dfrac{1}{\sqrt{5}}\neq 0$ i rozkład (X, Y) jest normalny; b) gęstość zmiennej losowej Z jest określona wzorem

$$g(z)=\int_{-\infty}^{\infty} f(x, z-x)\,dx=\frac{1}{\sqrt{2\pi}}\exp(-\tfrac{1}{4}z^2);\quad \text{c) } h(u,v)=\frac{1}{2\pi}\exp[-\tfrac{1}{2}(2u^2-2uv+v^2)].$$

5.57. Zastosować wzór (5.2.5).

5.58. $\operatorname{cov}\left(\dfrac{X-\alpha_1}{\sigma_1}, \dfrac{Y-\alpha_2}{\sigma_2}\right)=E\left[\dfrac{X-\alpha_1}{\sigma_1}\dfrac{Y-\alpha_2}{\sigma_2}\right]=\dfrac{E(X-\alpha_1)(Y-\alpha_2)}{\sigma_1\sigma_2}=\dfrac{\operatorname{cov}(X,Y)}{\sigma_1\sigma_2}=\rho.$

5.59. a) $\operatorname{cov}(Z, W)=acD^2X+bdD^2Y+(ad+bc)\operatorname{cov}(X, Y)$;

b) $a, b, c, d\in R$, warunki: 1) $ad-bc\neq 0$, 2) $ac\,\sigma_X^2+bd\,\sigma_Y^2+(ad+bc)\rho_{XY}\sigma_X\sigma_Y=0$. Dzieląc obustronnie przez $\sigma_Y^2\neq 0$, otrzymujemy równanie kwadratowe względem ilorazu $\sigma_X:\sigma_Y>0$; warunek rzeczywistości $\Delta=(ad+bc)^2\rho^2-4abcd\geqslant 0$: jeśli $abcd\leqslant 0$, to zmienne losowe Z, W zawsze są nieskorelowane przy dodatkowym spełnieniu warunków 1) i 2), jeśli zaś $abcd>0$, to do warunków 1), 2) należy dołączyć warunki $2\sqrt{abcd}/|ad+bc|\leqslant|\rho|$ oraz $ad+bc\neq 0$.

5.60. a) $Y=X_1+X_2$, $EY=\alpha_1+\alpha_2$, $D^2Y=\sigma_1^2+\sigma_2^2$;

b) $\operatorname{cov}(Y, X_2)=E(YX_2)-(EY)(EX_2)=E(X_1+X_2)X_2-(\alpha_1+\alpha_2)\alpha_2=E(X_1X_2)+$ $+E(X_2^2)-\alpha_1\alpha_2-\alpha_2^2=\alpha_1\alpha_2+\sigma_2^2+\alpha_2^2-\alpha_1\alpha_2-\alpha_2^2=\sigma_2^2$, $\rho=\dfrac{\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}$, jeśli σ_2 jest dużo większe od σ_1, to $\rho\approx 1$;

c) $\operatorname{cov}(Y, X_2)=E[(X_1+X_2-\alpha_1-\alpha_2)(X_2, \alpha_2)]=\rho\sigma_1\sigma_2+\sigma_2^2-D^2Y=\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2$,

$$\rho_{YX_2}=\frac{\rho\sigma_1+\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}}.$$

5.61. $P(X<\tfrac{1}{2}r_0, Y>8)=\displaystyle\int_0^{r_0/2}\left[\int_8^9 \frac{2}{r_0^2}x\tfrac{3}{64}(9-y)^2\,dy\right]dx=\tfrac{1}{256}.$

5.62. $EX=EY=n\alpha$, $D^2X=D^2Y=n\sigma^2$ (zmienne nieskorelowane) $\operatorname{cov}(X, Y)=$ $=E[\sum_{i=1}^{n}(X_i-\alpha)\sum_{j=1}^{n}(Y_j-\alpha)]=\sum_{i=1}^{n}\sum_{j=1}^{n}E(X_i-\alpha)(Y_j-\alpha)=k\sigma^2+(n^2-k)\underset{i\neq j}{\operatorname{cov}}(X_i, Y_j)=k\sigma^2$,

$$\rho_{XY}=\frac{k}{n}\frac{\sigma^2}{\sigma^2}=\frac{k}{n}.$$

5.63. $EX=n\alpha$, $EY=m\alpha$, $D^2X=n\sigma^2$, $D^2Y=m\sigma^2$,

$$E(XY)=\sum_{i=1}^{k}EX_i^2+\sum_{i=1}^{k}\sum_{l=k+1}^{m}EX_iEY_l+\sum_{i=1}^{k}\sum_{j=k+1}^{n}EX_iEX_j+\sum_{i\neq j}^{k}EX_iEX_j+$$
$$+\sum_{j=k+1}^{n}\sum_{l=k+1}^{m}EX_jEY_l=\sum_{i=1}^{k}EX_i^2+k(m-k)\alpha^2+k(n-k)\alpha^2+k(k-1)\alpha^2+$$
$$+(n-k)(m-k)\alpha^2,$$

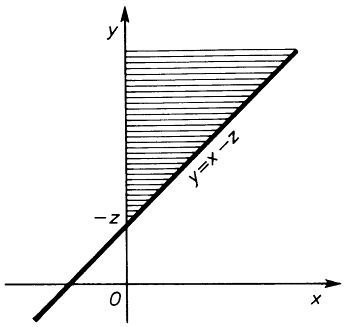
$$\rho_{XY}=\frac{\sum\limits_{i=1}^{k} EX_i^2-k\alpha^2}{\sqrt{mn}\,\sigma^2}=\frac{k}{\sqrt{mn}},$$

w szczególności gdy $n=m$, otrzymujemy wynik zadania poprzedniego.

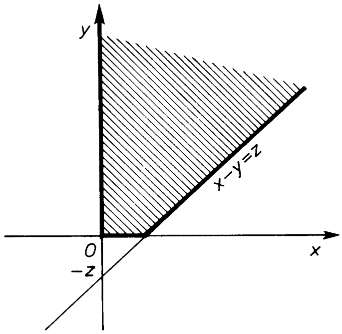
5.64. $\rho_{XY}=\rho$, $\quad \rho_{uv}=\dfrac{\operatorname{cov}(aX+b,\, cY+d)}{\sqrt{D^2(aX+b)D^2(cY+d)}}=$

$$=\frac{E(aX+b)(cY+d)-(aEX+b)(cEY+d)}{|ac|\sqrt{D^2XD^2Y}}=\frac{ac[E(XY)-(EX)(EY)]}{|ac|\sqrt{D^2XD^2Y}}=\rho\operatorname{sign}(ac).$$

5.65. Wyznaczamy dystrybuantę zmiennej losowej $Z=X-Y$, $G(z)=P(Z<z)=$ $=P(X-Y<z)=\iint\limits_{D_z}\lambda\nu e^{-\lambda x}e^{-\nu y}\,dx\,dy.$



Rys. 5.25. Obszar $D_z=\{(x,y): x\geqslant 0 \wedge y\geqslant x-z\}$ (odp. do zad. 5.65)



Rys. 5.26. Obszar $D_z=\{(x,y): y\geqslant 0 \wedge 0\leqslant x\leqslant z+y\}$ (odp. do zad. 5.65)

Należy rozpatrzyć dwa przypadki:

1) $z<0$, wówczas $D_z=\{(x,y)\colon\ x\geqslant 0\wedge y\geqslant x-z\}$ (rys. 5.25)

$$G(z)=\lambda\nu\int\limits_0^{\infty}\left(\int\limits_{x-z}^{\infty}e^{-\lambda x}e^{-\nu y}dy\right)dx=\frac{\lambda}{\lambda+\nu}e^{\nu z},$$

2) $z\geqslant 0$, wówczas $D_z=\{(x,y)\colon\ y\geqslant 0\wedge 0\leqslant x\leqslant y+z\}$ (rys. 5.26)

$$G(z)=\lambda\nu\int\limits_0^{\infty}\left(\int\limits_0^{z+y}e^{-\lambda x}e^{-\nu y}dx\right)dy=1-\frac{\nu}{\lambda+\nu}e^{-\lambda z}.$$

5.66. W gęstości f zmiennej (X_1, X_2) wykonać zamianę zmiennych: $Z=\dfrac{X_1}{X_1+X_2}$, $X=X((5.2.5))$.

$$f(z)=\begin{cases}\dfrac{1}{B(p_1,p_2)}z^{p_1-1}(1-z)^{p_2-1} & \text{dla} \quad 0<z<1,\\ 0 & \text{dla pozostałych } z.\end{cases}$$

5.67. $f(z)=\begin{cases}\frac{1}{2}e^{-z/2} & \text{dla } z>0 \text{ (rozkład } \chi^2 \text{ z } k=2 \text{ stopniami swobody,}\\ 0 & \text{dla pozostałych } z \text{ (jest to jednocześnie rozkład wykładniczy).}\end{cases}$

5.68. a) $k(r,\varphi)=\begin{cases}\frac{1}{2\pi}re^{-r^2/2} & \text{dla} \quad r>0,\ 0\leqslant\varphi\leqslant 2\pi,\\ 0 & \text{dla pozostałych } (r,\varphi);\end{cases}$

b) $k_1(r)$ jest gęstością rozkładu Rayleigha (2.8.29).

5.69. $Z=X+Y$, to $P(Z=z)=\sum\limits_{x} p_1(X=x)p_2(Y=z-x)$ u nas

$$p_1(X=x)=pq^{x-1}, \quad x\in \boldsymbol{N}, \quad p_2(Y=y)=pq^{y-1}, \quad y\in \boldsymbol{N}.$$

$$P(Z=z)=\sum_{x=1}^{z-1} pq^{x-1}pq^{z-x-1}=p^2\sum_{x=1}^{z-1} q^{z-2}=p^2(z-1)q^{z-2}, \quad z=2,3,\ldots,$$

ponieważ $1\leqslant y\Rightarrow y=z-x\Rightarrow 1\leqslant x\leqslant z-1$.

5.70. $P\left(\frac{1}{4}a\leqslant S\leqslant\frac{3}{4}a\right)=\int\limits_{a/4}^{a/2}\frac{4}{a^2}x\,dx+\int\limits_{a/2}^{3a/4}\left(-\frac{4}{a^2}x+\frac{4}{a}\right)dx=1-\frac{1}{4}=\frac{3}{4}$.

Rozkład sumy dwóch niezależnych zmiennych losowych o rozkładzie równomiernym ((2.8.3) przy $a=0$, $b=1$) (rys. 6.2), jest to rozkład Simpsona.

5.71. $f_1(x)=\begin{cases}\dfrac{x^{p-1}}{\Gamma(p)\,\Gamma(r)}\displaystyle\int\limits_x^{\infty} e^{-y}(y-x)^{r-1}dy=\dfrac{x^{p-1}e^{-x}}{\Gamma(p)} & \text{dla} \quad x>0,\\ 0 & \text{dla pozostałych } x,\end{cases}$

$f_2(y)=\begin{cases}\dfrac{e^{-y}}{\Gamma(p)\Gamma(r)}\displaystyle\int\limits_0^{y} x^{p-1}(y-x)^{r-1}dx=\dfrac{1}{\Gamma(p+r)}y^{p+r-1}e^{-y} & \text{dla} \quad y>0,\\ 0 & \text{dla pozostałych } y.\end{cases}$

5.72. $f(z)=\begin{cases}\dfrac{1}{\Gamma(2p)}z^{2p-1}e^{-z} & \text{dla} \quad z>0,\\ 0 & \text{dla pozostałych } z,\end{cases}$

a więc jest to również rozkład gamma o parametrze $2p$.

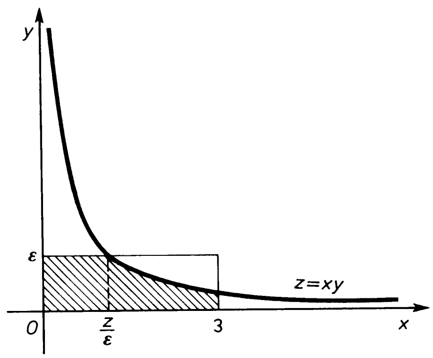
5.73.

$$f(z)=\begin{cases}0 & \text{dla} \quad z\leqslant 0,\\ 1-e^{-\lambda z} & \text{dla} \quad 0<z\leqslant 1,\\ e^{-\lambda z}(e^{\lambda}-1) & \text{dla} \quad z>1.\end{cases}$$

5.74. $\rho=-\frac{1}{144}:\frac{11}{144}=-\frac{1}{11}$.

5.75. a) $F(z)=P(Z<z)=P(XY<z)=\iint\limits_D\frac{2}{9\varepsilon}\,x\,dx\,dy$, gdzie obszar D (rys. 5.27)

dla $0<z<3\varepsilon$,



Rys. 5.27. Odpowiedź do zadania 5.75

$$F(z)=\begin{cases} 0 & \text{dla} \quad z\leqslant 0, \\ \dfrac{2z}{3\varepsilon}-\dfrac{z^2}{9\varepsilon^2} & \text{dla} \quad 0<z\leqslant 3\varepsilon, \\ 1 & \text{dla} \quad z>3\varepsilon; \end{cases}$$

b) $D^2Z=\frac{1}{2}\varepsilon^2$.

5.76. $F(z)=\begin{cases} 0 & \text{dla} \quad z\leqslant 0, \\ \frac{1}{2}z & \text{dla} \quad 0<z\leqslant 2, \\ 1 & \text{dla} \quad z>2. \end{cases}$

5.77. a) $f_1(x)=\begin{cases} xe^{-x} & \text{dla} \quad x>0, \\ 0 & \text{dla pozostałych } x, \end{cases}$

$$f_2(y)=\begin{cases} \frac{1}{2}y^2e^{-y} & \text{dla} \quad y>0, \\ 0 & \text{dla pozostałych } y, \end{cases}$$

$$f_3(z)=\begin{cases} \frac{1}{6}z^3e^{-z} & \text{dla} \quad z>0, \\ 0 & \text{dla pozostałych } z. \end{cases}$$

b) $\bigwedge_{(x,y,z)\in R^3} f(x,y,z)=f_1(x)f_2(y)f_3(z)$, a więc X, Y, Z niezależne;

c) z niezależności zmiennych X, Y, Z wynika, że $E(Z\,|\,X=x,\ Y=y)=E(Z)=$ $=\frac{1}{6}\int_0^\infty z^4e^{-z}dz=4$.

5.78. a) Obszar V jest walcem obrotowym o wysokości $h=1$ i promieniu podstawy $r=2$, skąd jego objętość $|V|=4\pi$ oraz gęstość

$$f(x,y,z)=\begin{cases} \dfrac{1}{4\pi} & \text{dla} \quad (x,y,z)\in V, \\ 0 & \text{dla pozostałych } (x,y,z), \end{cases}$$

$$f_3(z)=\begin{cases} 1 & \text{dla} \quad 0<z<1, \\ 0 & \text{dla pozostałych } z, \end{cases}$$

$E(Z)=\frac{1}{2}$;

b) stosujemy wzór (5.8.11):

$$f(x\,|\,y,z)=\begin{cases}\dfrac{1}{2\sqrt{4-y^2}} & \text{dla} \quad -\sqrt{4-y^2}<x<\sqrt{4-y^2}\\ 0 & \text{dla pozostałych } x,\end{cases}$$

z symetrii tego rozkładu względem $x=0$ wynika bez obliczania, że

$$E(X\,|\,Y=y,Z=z)=0.$$

c) Powierzchnia regresji I-go rodzaju zmiennej X względem Y i Z jest w tym przypadku płaszczyzną OYZ, a więc ma równanie: $x=0$.

5.79. a) $\bigwedge\limits_{i=1,2,3} P(X_i=k)=\frac{1}{10}$ dla $k=1,\ldots,10$,

$$P(X_1=k,X_2=j,X_3=r)=\frac{1}{10^3} \quad \text{dla} \quad k,j,r=1,\ldots,10;$$

b) $E(X_1^2\,|\,X_2=x_2,X_3=x_3)=\frac{1}{10}\sum\limits_{k=1}^{10} k^2=38{,}5$.

5.80. a) $\bigwedge\limits_{i=1,\ldots,n} P(X_i=x_i)=\dfrac{1}{6}$ dla $x_i=1,\ldots,6$, $P(X_1=x_1,\ldots,X_n=x_n)=\dfrac{1}{6^n}$,
X_i – są niezależnymi zmiennymi losowymi,

b) $D^2(X_1X_2\,|\,X_3=x_3,\ldots,X_n=x_n)=D^2(X_1X_2)=E(X_1^2)E(X_2^2)-(EX_1)^2(EX_2)^2=$
$=(\frac{91}{6})^2-(\frac{49}{4})^2=79{,}97$.

5.81. a) Macierz odwrotna do macierzy kowariancji jest:

$$\mathbf{M}^{-1}=\frac{1}{27}\begin{bmatrix}7 & -4 & -5\\ -4 & 10 & -1\\ -5 & -1 & 19\end{bmatrix},$$

$$f(x,y,z)=\frac{1}{(6\pi)^{3/2}}\exp\left[-\frac{1}{54}(7x^2-8xy-10xz-2yz+10y^2+19z^2)\right].$$

b) Równanie płaszczyzny regresji $y=a_{21}x+a_{23}z$, gdzie $a_{21}=-\dfrac{M_{21}^*}{M_{22}^*}=\dfrac{4}{10}$, $a_{23}=$
$=-\dfrac{M_{23}^*}{M_{22}^*}=\dfrac{1}{10}$, skąd $y=0{,}4x+0{,}1z$;

c) $EW=2EX+EZ-EY=0$, $D^2W=E(2X+Z-Y)^2=E(4X^2+Z^2+Y^2+4XZ-$
$-2ZY-4XY)=4D^2X+D^2Y+D^2Z+4\mathrm{cov}(X,Z)-2\mathrm{cov}(Z,Y)-4\mathrm{cov}(X,Y)=28$.

5.82. Macierz kowariancji jest macierzą odwrotną do macierzy:

$$\mathbf{B}=\begin{bmatrix}2 & 1 & 0\\ 1 & 2 & 1\\ 0 & 1 & 1\end{bmatrix}$$

złożonej ze współczynników formy kwadratowej występującej we wzorze na gęstość (5.8.21):

$$f(x, y, z)=\frac{1}{(2\pi)^{3/2}}\exp\left[-\frac{1}{2}(2x^2+2y^2+z^2+2xy+2yz)\right],$$

skąd

$$\mathbf{M}=\begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & -2 \\ 1 & -2 & 3 \end{bmatrix};$$

b) $\rho_{XY}=\dfrac{\operatorname{cov}(X, Y)}{\sqrt{D^2XD^2Y}}=-\dfrac{1}{\sqrt{2}}$.

5.83. a) $\mathbf{M}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 8 \end{bmatrix}$;

b) $\operatorname{cov}(Y, Z)=2$, $E(X)=0$, $E(Y)=1$, $E(Z)=-2$, $E(YZ)-EY\,EZ=2$, skąd $E(YZ)=0$;
c) $EY(Y-Z)=E(Y^2)-E(YZ)=D^2Y+(EY)^2=2+1=3$.

5.84. Niech X_i oznacza liczbę wyrzuconych oczek na i-tej kostce $i=1, \ldots, n$. X_i są niezależnymi zmiennymi losowymi. $E\sum\limits_{i=1}^{n} X_i=\sum\limits_{i=1}^{n} EX_i$ oraz $D^2\sum\limits_{i=1}^{n} X_i=\sum\limits_{i=1}^{n} D^2X_i$. Ponieważ $EX_i=\frac{1}{6}\sum\limits_{k=1}^{6} k=3{,}5$, zaś $D^2X_i=\frac{1}{6}\sum\limits_{k=1}^{6} k^2-3{,}5^2=\frac{35}{12}$, więc $E(\sum\limits_{i=1}^{n} X_i)=3{,}5n$, $D^2(\sum\limits_{i=1}^{n} X_i)=\frac{35}{12}n$.

5.85. $c=\dfrac{1}{(2\pi)^{3/2}}$, $\quad\mathbf{M}=\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}$.

5.86. Korzystając ze wzoru (5.6.4), otrzymuje się gęstość rozkładu Cauchy'ego

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

5.87. Tak; jako złożenie rozkładu o gęstości $f(x|t)$, której parametr t podlega rozkładowi dwupunktowemu, tzn. $P(T=t_1)=p_1>0$, $P(T=t_2)=p_2>0$, przy czym $p_1+p_2=1$, wówczas $f_1(x)=f(x|t_1)$ oraz $f(x|t_2)=f_2(x)$.

5.88. Niech $Y=\max(X_1, \ldots, X_5)$, to $P(Y>4)=1-P(Y\leqslant 4)=1-P(X_1\leqslant 4\wedge\ldots\wedge X_5\leqslant 4)=1-[\frac{1}{4}\int\limits_0^4 xe^{-x^2/8}dx]^5=1-(1-e^{-1/2})^5$.

5.89. Zauważmy, że $X=R\cos\Phi$, $Y=R\sin\Phi$ oraz $|J|=R$. Gęstość g zmiennej losowej (R, Φ) jest postaci:

$$g(r, \varphi)=\begin{cases} \dfrac{r}{\pi} & \text{dla} \quad 0<r<1\wedge 0\leqslant\varphi<2\pi, \\ 0 & \text{dla pozostałych } (r, \varphi). \end{cases}$$

5.90. Gęstość f dwuwymiarowej zmiennej losowej (X, Y) jest określona wzorem:

$$f(x,y)=\begin{cases}\dfrac{1}{\lambda^2}\exp\left(-\dfrac{x+y}{\lambda}\right) & \text{dla} \quad x>0, \quad y>0,\\ 0 & \text{dla pozostałych } (x,y);\end{cases}$$

oznaczając $U=X+Y$, $V=\dfrac{X}{Y}$ i korzystając z (5.2.5), otrzymujemy gęstość k dwuwymiarowej zmiennej (U, V):

$$k(u,v)=\begin{cases}\dfrac{1}{\lambda^2}u\exp\left(-\dfrac{u}{\lambda}\right)\dfrac{1}{(v+1)^2} & \text{dla} \quad u>0, \quad v>0,\\ 0 & \text{dla pozostałych } (u,v).\end{cases}$$

Zauważmy, że jest ona iloczynem gęstości k_1, k_2 rozkładów brzegowych

$$k_1(u)=\begin{cases}\dfrac{1}{\lambda^2}u\exp\left(-\dfrac{u}{\lambda}\right) & \text{dla} \quad u>0,\\ 0 & \text{dla pozostałych } u,\end{cases}$$

$$k_2(v)=\begin{cases}\dfrac{1}{(v+1)^2} & \text{dla} \quad v>0,\\ 0 & \text{dla pozostałych } v,\end{cases}$$

a więc U i V są niezależnymi zmiennymi losowymi.

5.91. Łączna gęstość dwuwymiarowej zmiennej losowej (X, Y) jest określona wzorem:

$$f(x,y)=\begin{cases}\dfrac{1}{2^{(n_1+n_2)/2}\Gamma(\frac{1}{2}n_1)\Gamma(\frac{1}{2}n_2)}x^{n_1/2-1}y^{n_2/2-1}e^{-(x+y)/2} & \text{dla} \quad x>0, \quad y>0,\\ 0 & \text{dla pozostałych } (x,y).\end{cases}$$

Stosując zamianę zmiennych jak w zad. 5.90, otrzymamy gęstość łącznej zmiennej (U, V), gdzie $U=X+Y$, $V=\dfrac{X}{Y}$ postaci:

$$g(u,v)=\begin{cases}\dfrac{1}{\Gamma\left(\dfrac{n_1+n_2-2}{2}\right)}u^{(n_1+n_2-2)/2-1}e^{-u}\cdot\dfrac{\Gamma\left(\dfrac{n_1+n_2-2}{2}\right)}{2^{(n_1+n_2)/2}\Gamma(\frac{1}{2}n_1)\Gamma(\frac{1}{2}n_2)}\dfrac{v^{n_1/2-1}}{(1+v)^{(n_1+n_2)/2-2}} & \\ \qquad\qquad \text{dla} \quad u>0, \quad v>0,\\ 0 \qquad\qquad \text{dla pozostałych } (u,v).\end{cases}$$

Zauważmy, że dla każdego $(u, v)\in \boldsymbol{R}^2$ jest ona iloczynem gęstości rozkładów brzegowych, a więc U i V są niezależnymi zmiennymi losowymi.

5.92. $$k(v)=\begin{cases}\dfrac{2}{\sigma^3\sqrt{\pi}}v^2e^{-v^2/(2\sigma^2)} & \text{dla} \quad v>0,\\ 0 & \text{dla pozostałych } v,\end{cases}$$

gdzie $\sigma=\sqrt{k\dfrac{T}{\mu}}$. Jest to gęstość rozkładu Maxwella (2.8.31).

5.93. Stosując wzór (2.5.3) wykazujemy, że zmienne losowe $Z_i = X_i^2$ dla $i=1,\ldots,n$ mają rozkład gamma (2.8.12) o parametrach $p=\frac{1}{2}$, $\lambda=2$, a następnie korzystając z twierdzenia o dodawaniu (zad. 4.5) względem parametru p, otrzymujemy gęstość k zmiennej losowej Y postaci:

$$k(y)=\begin{cases} \dfrac{1}{2^{n/2}\Gamma(\frac{1}{2}n)} y^{n/2-1}\exp(-\frac{1}{2}y) & \text{dla} \quad y>0, \\ 0 & \text{dla} \quad y\leqslant 0, \end{cases}$$

jest to gęstość rozkładu χ^2 z n stopniami swobody.

5.94. 1) Dla ustalonego N wystarczy skorzystać z własności wartości przeciętnej.

2) Niech N będzie zmienną losową. Oznaczmy przez $\bar{X}=\frac{1}{N}\sum_{i=1}^{N} X_i$. Korzystając ze wzoru: $E\bar{X}=E\,E(\bar{X}|N)$, otrzymamy

$$EX=\sum_{n=1}^{\infty} E(\bar{X}|N=n)P(N=n)=$$

$$=\sum_{n=1}^{\infty} P(N=n)\sum_{i=1}^{n}\frac{1}{n}EX_i=\sum_{n=1}^{\infty} P(N=n)\frac{1}{n}n\alpha=\alpha\sum_{n=1}^{\infty} P(N=n)=\alpha.$$

5.95. Niech A oznacza zdarzenie polegające na tym, że pierwiastki równania x^2+2Px+ $+Q=0$ są rzeczywiste. $P(A)=P(\Delta\geqslant 0)=P(P^2\geqslant Q)$.

Rozpatrzmy dwa przypadki:

1) $b\leqslant a^2$ (rys. 5.28)

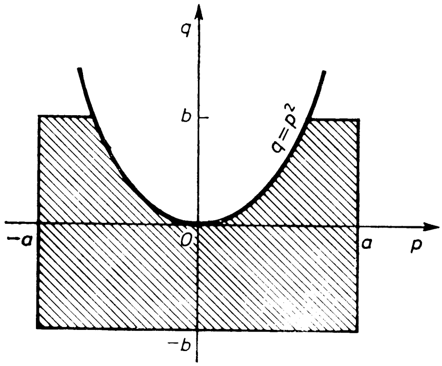
$$P(A)=1-\frac{2}{4ab}\iint_D dp\,dq=1-\frac{\sqrt{b}}{3a},$$

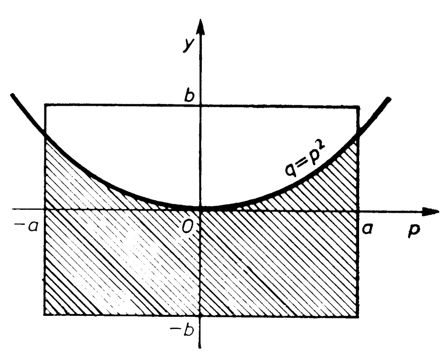
gdzie $D=\{(p,q)\colon\ 0\leqslant q\leqslant b \wedge 0\leqslant p\leqslant\sqrt{q}\}$;

2) $b>a^2$ (rys. 5.29)

$$P(A)=1-\frac{2}{4ab}\iint_{D_1} dp\,dq=\frac{1}{2}+\frac{a^2}{6b},$$

gdzie $D_1=\{(p,q)\colon\ 0\leqslant p\leqslant a \wedge p^2\leqslant q\leqslant b\}$.





Rys. 5.28, 5.29. Odpowiedź do zadania 5.95

6 TWIERDZENIA GRANICZNE

Niech (X_n) będzie ciągiem zmiennych losowych (ciągiem losowym). Rozważmy odpowiadające mu: 1) ciąg funkcji prawdopodobieństwa $P_i(x)=P(X_i=x)$, $i\in N$, w przypadku zmiennych losowych dyskretnych albo ciąg gęstości $(f_n(x))$ w przypadku zmiennych losowych typu ciągłego oraz 2) ciąg dystrybuant $(F_n(x))$. Twierdzenia graniczne, a więc przy $n\to\infty$, dotyczące zbieżności grupy 1) nazywamy *twierdzeniami granicznymi lokalnymi*, a grupy 2) – *twierdzeniami granicznymi integralnymi.*

Przykładem granicznego twierdzenia lokalnego jest twierdzenie (2.7.18) o zbieżności funkcji prawdopodobieństwa rozkładu Bernoulliego do funkcji prawdopodobieństwa rozkładu Poissona.

6.1. CENTRALNE TWIERDZENIA GRANICZNE

Ważniejsze znaczenie – szczególnie w zastosowaniach – mają integralne twierdzenia graniczne, do których należą prawa wielkich liczb rozważone poniżej oraz szereg twierdzeń granicznych, z których do najważniejszych należy:

Centralne twierdzenie graniczne Lindeberga-Levy'ego. *Jeżeli* $\{X_n\}$ *jest losowym ciągiem niezależnych zmiennych o jednakowym rozkładzie, o wartości przeciętnej* α_1 *i skończonej wariancji* $\sigma^2>0$, *to ciąg* (F_n) *dystrybuant standaryzowanych średnich arytmetycznych* $\bar{X}_n$ *(albo – co na jedno wychodzi – standaryzowanych sum* $\sum_{i=1}^{n} X_i$*)*

$$Y_n=\frac{\bar{X}_n-\alpha_1}{\dfrac{\sigma}{\sqrt{n}}}=\frac{\sum\limits_{i=1}^{n} X_i-n\alpha_1}{\sigma\sqrt{n}} \tag{6.1.1}$$

jest zbieżny do dystrybuanty Φ *rozkładu* $N(0, 1)$:

$$\lim_{n\to\infty} F_n(y)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{y} \exp(-\tfrac{1}{2}t^2)\,dt\equiv\Phi(y). \tag{6.1.2}$$

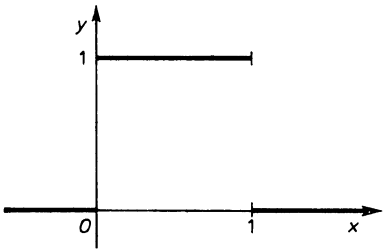
Dowód można znaleźć np. w [12]. Należy tutaj podkreślić fakt, że zmienne losowe mogą mieć zarówno rozkład dyskretny jak i typu ciągłego.

Ze wzoru (6.1.2) wynika, że dla dużych n (w praktyce rzędu kilkunastu, co oczywiście zależy od żądanego przybliżenia) można stosować wzór przybliżony

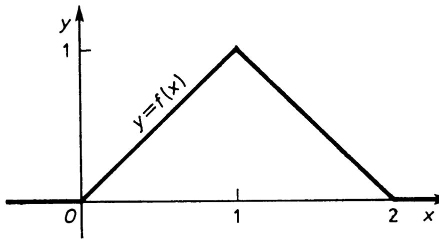
$$P(y_1 < Y_n \leqslant y_2) \approx \Phi(y_2) - \Phi(y_1), \tag{6.1.3}$$

gdzie Φ oznacza dystrybuantę rozkładu $N(0, 1)$.

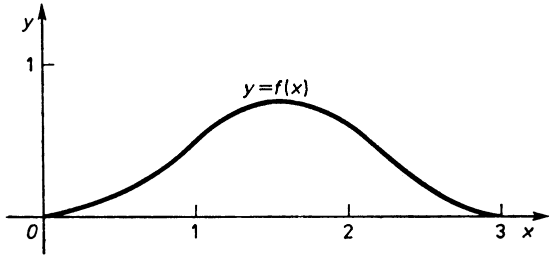
Rysunki 6.1 - 6.4 przedstawiają wykresy gęstości sumy n niezależnych zmiennych losowych o jednakowym rozkładzie równomiernym na przedziale $\langle 0, 1\rangle$ dla $n = 1, 2, 3, 4$.



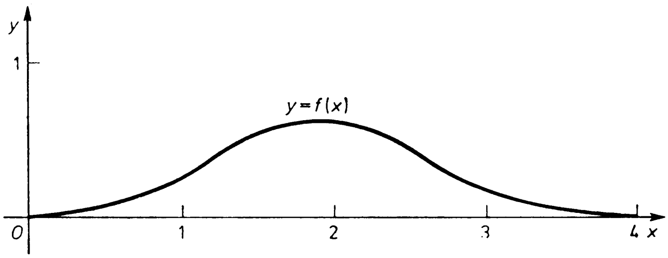
Rys. 6.1. Gęstość zmiennej losowej Y_1 o rozkładzie równomiernym skoncentrowanym na $\langle 0, 1\rangle$



Rys. 6.2. Gęstość f sumy $Y_2 = X_1 + X_2$ niezależnych zmiennych losowych o tym samym rozkładzie równomiernym na $\langle 0, 1\rangle$



Rys. 6.3. Gęstość f sumy $Y_3 = X_1 + X_2 + X_3$ niezależnych zmiennych losowych o tym samym rozkładzie równomiernym na $\langle 0, 1\rangle$



Rys. 6.4. Gęstość f sumy $Y = X_1 + X_2 + X_3 + X_4$ niezależnych zmiennych losowych o tym samym rozkładzie równomiernym na $\langle 0, 1\rangle$

ZADANIE 6.1. Losowy błąd pomiaru pewnej wielkości ma rozkład o wartości przeciętnej $\alpha_1 = 0$ (brak błędu systematycznego) i odchyleniu standardowym 0,08. Obliczyć prawdopodobieństwo, że błąd średniej arytmetycznej 100 pomiarów nie przekroczy (co do wartości bezwzględnej) 0, 1.

Rozwiązanie. Oznaczamy losowy błąd przez X_i, $i=1, \ldots, 100$. Na podstawie przybliżonego wzoru (6.1.3) mamy

$$P(|\bar{X}_{100}|<0{,}1)=P\left(-\frac{0{,}1}{0{,}08}<Y_{100}<\frac{0{,}1}{0{,}08}\right)\approx 2\Phi(1{,}25)-1=0{,}7888.$$

Stosując twierdzenie Lindeberga-Levy'ego do ciągu (X_n) niezależnych zmiennych losowych o tym samym rozkładzie zero-jedynkowym (2.7.5) i oznaczając $S_n = \sum_{k=1}^{n} X_k$, otrzymujemy:

INTEGRALNE TWIERDZENIE MOIVRE'A-LAPLACE'A. *Jeśli* (S_n) *jest ciągiem zmiennych losowych o rozkładzie dwumianowym* (2.7.7) *z parametrami* (n, p), $0<p<1$ (*a więc o wartości przeciętnej* $ES_n = np$ *i wariancji* $\operatorname{Var} S_n = npq$) *oraz* Y_n *jest ciągiem standaryzowanych zmiennych losowych*:

$$Y_n = \frac{S_n - np}{\sqrt{npq}},$$

to dla każdej pary wartości $y_1 < y_2$ *zachodzi wzór*:

$$\lim_{n\to\infty} P\left(y_1 < \frac{S_n - np}{\sqrt{npq}} < y_2\right) = \Phi(y_2) - \Phi(y_1). \tag{6.1.4}$$

6.1.1. Zadania rozwiązane.

ZADANIE 6.2. Prawdopodobieństwo, że w czasie T przestanie świecić jedna żarówka jest równe 0,1. Obliczyć prawdopodobieństwo, że w czasie T spośród 100 przestanie świecić od 7 do 19 żarówek przy założeniu, że żarówki przepalają się niezależnie.

Rozwiązanie. Niech S_{100} będzie liczbą żarówek spośród stu, które w ciągu czasu T przestały świecić: $ES_{100} = np = 10$, $D^2 S_{100} = npq = 10 \cdot 0{,}9 = 9$.

Korzystamy ze wzoru (6.1.4). Ponieważ jednak dla zmiennej losowej ciągłej X zachodzi np. równość $P(a<X<b) = P(a \leqslant X \leqslant b)$, która dla zmiennych skokowych ogólnie nie zachodzi, więc w przypadku rozkładu dwumianowego (a więc skokowego) przy a, b całkowitych nieujemnych postępuje się zazwyczaj następująco:

$$P(a \leqslant S_n \leqslant b) = P(a-0{,}5 < S_n < b+0{,}5) = P\left(\frac{a-0{,}5-np}{\sqrt{npq}} < Y_n < \frac{b+0{,}5-np}{\sqrt{npq}}\right) =$$

$$= \Phi\left(\frac{b+0{,}5-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-0{,}5-np}{\sqrt{npq}}\right).$$

W zadaniu mamy:

$$P(7 \leqslant S_{100} \leqslant 19) = P(6{,}5 < S_{100} < 19{,}5) = P\left(\frac{6{,}5-10}{3} < Y_{100} < \frac{19{,}5-10}{3}\right) \approx$$

$$\approx \Phi\left(\frac{9{,}5}{3}\right) - \Phi\left(-\frac{3{,}5}{3}\right) = \Phi(3{,}17) + \Phi(1{,}17) - 1 = 0{,}8783.$$

ZADANIE 6.3. W centrali telefonicznej znajduje się n linii działających niezależnie. Prawdopodobieństwo, że dowolna ustalona linia jest zajęta, jest równe 0,1. Jakie powinno być n, aby prawdopodobieństwo tego, że co najmniej 7 linii jest zajętych było równe 0,95?

Rozwiązanie. Liczba linii zajętych jest zmienną losową S_n o rozkładzie dwumianowym z parametrami: n, $p = 0{,}1$. Korzystając z tw. Moivre'a-Laplace'a dobieramy n tak, aby zachodziła równość

$$P(S_n \geqslant 0{,}07n) = 0{,}95.$$

Zauważmy, że $ES_n = 0{,}1n$ oraz $D^2X_n = 0{,}1 \cdot 0{,}9n$, skąd odchylenie standardowe: $\sigma = 0{,}3\sqrt{n}$. Standaryzując zmienną losową S_n, otrzymamy

$$P\left(Y_n \geqslant \frac{0{,}07n - 0{,}1n}{0{,}3\sqrt{n}}\right) = 0{,}95,$$

a po przejściu do granicznego rozkładu normalnego:

$$1 - \Phi(-0{,}1\sqrt{n}) \approx 0{,}95, \qquad \Phi(0{,}1\sqrt{n}) \approx 0{,}95.$$

Dla wartości dystrybuanty rozkładu $N(0, 1)$ równej 0,95 odczytujemy z tablic liczbową wartość argumentu

$$0{,}1\sqrt{n} \approx 1{,}64, \qquad n \approx 268{,}96.$$

W centrali telefonicznej powinny być co najmniej 269 linii.

6.2. PRAWA WIELKICH LICZB

Niech (X_n) będzie ciągiem zmiennych losowych, dla których $EX_i = \mu_i < \infty$ dla $i \in N$ oraz

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad E\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} EX_i.$$

Jeżeli dla losowego ciągu (X_n) i dla dowolnego $\varepsilon > 0$

$$\lim_{n\to\infty} P(|\bar{X}_n - E\bar{X}_n| \geqslant \varepsilon) = 0, \tag{6.2.1}$$

to mówimy, że dla tego ciągu zachodzi *słabe prawo wielkich liczb*. Mówimy również, że $\bar{X}_n - E\bar{X}_n \to 0$ *według prawdopodobieństwa* (*stochastycznie, według miary P*), co zapisujemy $\bar{X}_n - E\bar{X}_n \xrightarrow{P} 0$. W przypadku, gdy przy tym samym założeniu

$$P[\lim_{n\to\infty} (\bar{X}_n - E\bar{X}_n) = 0] = 1, \tag{6.2.2}$$

wtedy mówimy, że dla losowego ciągu (X_n) zachodzi *mocne prawo wielkich liczb*, a zbieżność $\bar{X}_n - E\bar{X}_n \to 0$ wyrażoną wzorem (6.2.2) nazywamy *zbieżnością z prawdopodobieństwem* 1 (*prawie na pewno, prawie wszędzie P*), co będziemy zapisywali $\bar{X}_n - E\bar{X}_n \xrightarrow{p=1} 0$.

Pierwsze prawo wielkich liczb (słabe) dla ciągu (X_n) zmiennych losowych o tych samych rozkładach zero-jedynkowych udowodnił Bernoulli (1713 r.). Znacznie silniejsze wyniki osiągnął Kołmogorow: Mocne prawo wielkich liczb Kołmogorowa (I). Dla losowego ciągu (X_n) o wspólnie ograniczonych wariancjach ($D^2X_i \leqslant C$, $i \in N$) zachodzi mocne prawo wielkich liczb.

ZADANIE 6.4. Wykazać, że jeśli w ciągu n niezależnych doświadczeń prawdopodobieństwo zajścia zdarzenia A w i-tym doświadczeniu jest równe p_i, to

$$P\left[\lim_{n\to\infty}\left(\frac{S_n}{n} - \frac{1}{n}\sum_{i=1}^{n} p_i\right) = 0\right] = 1,$$

gdzie S_n oznacza liczbę zajść zdarzenia A w pierwszych n doświadczeniach.

Rozwiązanie. $S_n = \sum_{i=1}^{n} X_i$, gdzie $P(X_i=1)=p_i$, $P(X_i=0)=1-p_i$, skąd $EX_i=p_i$, $D^2X_i = p_i(1-p_i) \leqslant \frac{1}{4}$, a więc są spełnione założenia twierdzenia Kołmogorowa, prawdziwy jest więc wzór (6.2.2). Wykazaliśmy więc, że mocne prawo wielkich liczb zachodzi m. in. dla schematu Poissona.

W dalszych badaniach okazało się, że mocne prawo wielkich liczb może zachodzić bez założenia o ograniczoności wariancji, a nawet bez jej istnienia, przy przyjęciu jednak założenia o jednakowym rozkładzie i niezależności zmiennych.

MOCNE PRAWO WIELKICH LICZB KOŁMOGOROWA (II). *Warunkiem koniecznym i wystarczającym zachodzenia mocnego prawa wielkich liczb dla ciągu (X_n) niezależnych zmiennych losowych o jednakowym rozkładzie jest istnienie skończonej wartości oczekiwanej $E(X_i)=m$ dla $i \in N$.*

Dowód znajduje się w [9].

ZADANIE 6.5. Niezależne zmienne losowe X_i, $i \in N$, mają ten sam rozkład prawdopodobieństwa $P(X_i = 2^k) = 0{,}8 \cdot 0{,}2^k$ dla $k \in N_0$, $i \in N$. Czy dla tego ciągu zachodzi mocne prawo wielkich liczb Kołmogorowa?

Rozwiązanie. Obliczamy $EX_i = \sum_{k=0}^{\infty} 2^k \cdot 0{,}8 \cdot 0{,}2^k = 0{,}8 \sum_{k=0}^{\infty} 0{,}4^k = 0{,}8 \frac{1}{1-0{,}4} = \frac{4}{3}$. A więc zachodzi wzór

$$P[\lim_{n\to\infty}(\bar{X}_n - \tfrac{4}{3}) = 0] = 1.$$

6.3. ZADANIA DO ROZWIĄZANIA

6.6. W pewnym magazynie znajduje się towar o przeciętnej wadliwości 0,1. Korzystając z twierdzenia Moivre'a Laplace'a, obliczyć prawdopodobieństwo, że wśród losowo wybranych 100 sztuk towaru procent sztuk wadliwych różni się od 10 o co najwyżej 0,15.

6.7. W urnie znajduje się 36 kul białych i 64 czarnych. Losujemy kule po jednej ze zwracaniem. Ile losowań należy dokonać, aby prawdopodobieństwo tego, że częstość otrzymywania kuli białej różni się od 0,36 o co najmniej 0,12 było równe 0,1?

6.8. W pewnej grupie ludzi co dziesiąty człowiek jest daltonistą. Obliczyć prawdopodobieństwo, że wśród 100 losowo wybranych ludzi będzie od 5 do 12 daltonistów.

6.9. W zajezdni znajduje się 200 autobusów. Prawdopodobieństwo, że losowo wybrany autobus jest sprawny do jazdy wynosi 0,7. Obliczyć prawdopodobieństwo, że w losowo wybranej chwili co najmniej 160 autobusów jest sprawnych.

6.10. W pewnej szkole uczy się 500 dzieci. Prawdopodobieństwo, że losowo wybrany uczeń ma co najmniej jedną dwójkę jest równe 0,1. Obliczyć prawdopodobieństwo, że w tej szkole liczba dzieci, które mają co najmniej jedną dwójkę różni się od 50 o co najwyżej 10.

6.11. Urządzenie składa się z n elementów. Urządzenie pracuje, jeśli co najmniej 70% elementów jest sprawnych. Prawdopodobieństwo awarii jednego elementu jest równe 0,2. Jak duża powinna być liczba elementów, aby z prawdopodobieństwem 0,95 urządzenie pracowało?

6.12. Strzelec trafia do celu z prawdopodobieństwem 0,5. Jaką liczbę strzałów musi oddać, aby prawdopodobieństwo tego, że częstość trafienia do celu różni się od 0,5 o co najwyżej 0,1 było równe 0,95?

6.13. Niech $X_1, \ldots, X_{100}$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie beta o gęstości $f(x)=12x(1-x)^2$ dla $0<x<1$. Obliczyć prawdopodobieństwo: $P(20 \leqslant \sum_{i=1}^{100} X_i \leqslant 30)$.

6.14. Niech $X_1, \ldots, X_{200}$ będą niezależnymi zmiennymi losowymi o rozkładzie geometrycznym: $P(X_k=i)=(\frac{1}{2})^i$, gdzie $i \in N, k=1, \ldots, 200$. Obliczyć prawdopodobieństwo, że średnia arytmetyczna tych zmiennych przyjmuje wartości z przedziału $\langle 1, 4 \rangle$.

6.15. Niech $X_1, \ldots, X_k, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie: $P(X_k=i)=\frac{2}{3}(\frac{1}{3})^{i-1}$, $i \in N$, $k \in N$. Wykazać że dla zmiennej losowej $Y_n=\frac{1}{n}\sum_{k=1}^{n} X_k$ zachodzi mocne prawo wielkich liczb.

6.16. Niezależne zmienne losowe $X_1, \ldots, X_k, \ldots$ podlegają temu samemu rozkładowi prawdopodobieństwa: $P\left(X_k=\frac{(-1)^i}{i}\right)=\frac{1}{2^i}$, $i, k \in N$. Sprawdzić czy w tym przypadku zachodzi twierdzenie Kołmogorowa.

6.17. Niech $X_1, \ldots, X_k, \ldots$ będą niezależnymi zmiennymi losowymi o rozkładzie Poissona: $P(X_k=i)=\frac{e^{-1}}{i!}$, $i \in N_0$. Zbadać czy dla średniej arytmetycznej $Y_n=\frac{1}{n}\sum_{k=1}^{n} X_k$ zachodzi a) mocne prawo wielkich liczb, b) centralne twierdzenie graniczne.

6.18. Dodajemy 10 000 liczb, każda z nich jest zaokrąglona z dokładnością do 10^{-m}. Zakładając, że błędy zaokrąglenia są niezależnymi zmiennymi losowymi o rozkładzie równomiernym w przedziale $(-\frac{1}{2}\cdot 10^{-m}, \frac{1}{2}\cdot 10^{-m})$, wyznaczyć przedział, w którym z prawdopodobieństwem 0,99 będzie się zawierał błąd sumy.

6.19. Partia towaru zawiera 20% braków. Obliczyć prawdopodobieństwo, że w próbce o liczności: a) $n=100$ sztuk, b) $n=400$ sztuk, c) $n=1600$ sztuk. Stosunek $\frac{k}{n}$ (gdzie k jest liczbą braków) różni się od wadliwości p partii nie więcej niż o 0,02. Przedstawić obliczone prawdopodobieństwo za pomocą wykresów odpowiednich gęstości rozkładu normalnego.

Odpowiedzi

6.6. $P\left(\left|\frac{k}{n}-0{,}1\right|\leqslant 0{,}15\right)\approx 2\Phi(5)-1\approx 1.$

6.7. Należy wykonać co najmniej $n\geqslant 40$ losowań.

6.8. $P(5\leqslant X\leqslant 12)\approx\Phi(0{,}67)+\Phi(1{,}67)-1\approx 0{,}7011.$

6.9. $P(X\geqslant 160)\approx 1-P(X<160)=1-\Phi\left(\frac{20}{\sqrt{42}}\right)\approx 0{,}00097.$

6.10. $P(|X-50|\leqslant 10)\approx 2\Phi\left(\frac{10}{\sqrt{45}}\right)-1\approx 0{,}8788.$

6.11. Urządzenie powinno się składać z co najmniej $n=44$ elementów.

6.12. $n\geqslant 96.$

6.13. $P(20\leqslant\sum\limits_{i=1}^{100}X_i\leqslant 30)\approx 0,\quad \bigwedge\limits_{i=1,\ldots,100} EX_i=\frac{2}{5},\quad D^2X_i=\frac{1}{25}$ ((2.8.21)).

6.14. $P\left(1\leqslant\frac{1}{200}\sum\limits_{i=1}^{200}X_i\leqslant 4\right)\approx 1,\quad \bigwedge\limits_{i=1,\ldots,200} EX_i=2,\quad D^2X_i=2$ ((2.6.30) i (2.6.31)).

6.15. $E(X_k)=\frac{1}{2}<\infty$ dla $k\in N$.

6.16. Tak, ponieważ $E(X_k)=\sum\limits_{i=1}^{\infty}\frac{(-1)^i}{i2^i}=-\ln\frac{3}{2}$ dla $k\in N$.

6.17. a) Tak, ponieważ $\bigwedge\limits_{k\in N}E(X_k)=1;$

b) tak, ponieważ $\bigwedge\limits_{k\in N}E(X_k)=1$ i $D^2(X_k)=1.$

6.18. Niech X_k, $k=1,\ldots,10\,000$ będą błędami zaokrągleń poszczególnych liczb, wtedy dla rozkładu równomiernego w $(-\frac{1}{2}\cdot 10^{-m}, \frac{1}{2}\cdot 10^{-m})$ $EX_k=0$, $D^2X_k=\frac{10^{-2m}}{12}$ ((2.8.6)). Oznaczamy $Y=\sum\limits_{k=1}^{10000}X_k$, wówczas $EY=0$, $D^2Y=10000\cdot\frac{10^{-2m}}{12}$, a odchylenie standardowe $\sigma_Y=\frac{100\cdot 10^{-m}}{2\sqrt{3}}$, skąd i z centralnego tw. granicznego mamy: $P\left(\left|\frac{X-0}{\sigma_Y}\right|<y\right)=0{,}99$, z tablic rozkładu $N(0,1)$ odczytujemy $y=2{,}58$, a więc błąd sumy $Y\in\left(-2{,}58\frac{10^{2-m}}{2\sqrt{3}},\ +2{,}58\frac{10^{2-m}}{2\sqrt{3}}\right)$, a w przybliżeniu $Y\in(-0{,}75\cdot 10^{2-m},\ 0{,}75\cdot 10^{2-m})$.

6.19. Skorzystać z integralnego twierdzenia Moivre'a-Laplace'a.

a) $P\left(\left|\frac{k}{n}-p\right| \leqslant 0{,}02\right)=0{,}383$; b) $P\left(\left|\frac{k}{n}-p\right| \leqslant 0{,}02\right)=0{,}6826$;

c) $P\left(\left|\frac{k}{n}-p\right| \leqslant 0{,}02\right)=0{,}9545$.

TABLICE

Tablica 1. Wartości współczynników $\binom{n}{k}$

$k \backslash n$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	n / k
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1
2		3	6	10	15	21	28	36	45	55	66	78	91	105	120	136	153	171	190	2
3				10	20	35	56	84	120	165	220	286	364	455	560	680	816	969	1 140	3
4						35	70	126	210	330	495	715	1 001	1 365	1 820	2 380	3 060	3 876	4 845	4
5								126	252	462	792	1 287	2 002	3 003	4 368	6 188	8 568	11 628	15 504	5
6										462	924	1 716	3 003	5 005	8 008	12 376	18 564	27 132	38 760	6
7												1 716	3 432	6 435	11 440	19 448	31 824	50 338	77 520	7
8														6 435	12 870	24 310	43 758	75 582	125 970	8
9																24 310	48 620	92 378	167 960	9
10																		92 378	184 756	10

Tablica 2. Silnie i ich logarytmy

n	$n!:10^c$	$\log n! = c+m$	n	$n!:10^c$	$\log n! = c+m$	n	$n!:10^c$	$\log n! = c+m$
1	1,0	0,000 00	41	3,345 25	49,524 43	81	5,797 13	120,763 21
2	2,0	0,301 03	42	1,405 01	51,147 68	82	4,753 64	122,677 03
3	6,0	0,778 15	43	6,041 53	52,781 15	83	3,945 52	124,596 10
4	2,40	1,380 21	44	2,658 27	54,424 60	84	3,314 24	126,520 38
5	1,20	2,079 18	45	1,196 22	56,077 81	85	2,817 10	128,449 80
6	7,20	2,857 33	46	5,502 62	57,740 57	86	2,422 71	130,384 30
7	5,040	3,702 43	47	2,586 23	59,412 67	87	2,107 76	132,323 82
8	4,032 0	4,605 52	48	1,241 39	61,093 91	88	1,854 83	134,268 30
9	3,628 80	5,559 76	49	6,082 82	62,784 10	89	1,650 80	136,217 69
10	3,628 80	6,559 76	50	3,041 41	64,483 07	90	1,485 72	138,171 94
11	3,991 68	7,601 16	51	1,551 12	66,190 65	91	1,352 00	140,130 98
12	4,790 02	8,680 34	52	8,065 82	67,906 65	92	1,243 84	142,094 77
13	6,227 02	9,794 28	53	4,274 88	69,630 92	93	1,156 77	144,063 25
14	8,717 83	10,940 41	54	2,308 44	71,363 32	94	1,087 37	146,036 38
15	1,307 67	12,116 50	55	1,269 64	73,103 68	95	1,033 00	148,014 10
16	2,092 28	13,320 62	56	7,109 99	74,851 87	96	9,916 78	149,996 37
17	3,556 87	14,551 07	57	4,052 69	76,607 74	97	9,619 28	151,983 14
18	6,402 37	15,806 34	58	2,350 56	78,371 17	98	9,426 89	153,974 37
19	1,216 45	17,085 09	59	1,386 83	80,142 02	99	9,332 62	155,970 00
20	2,432 90	18,386 12	60	8,320 99	81,920 17	100	9,332 62	157,970 00
21	5,109 09	19,708 34	61	5,075 80	83,705 50	101	9,425 95	159,974 33
22	1,124 00	21,050 77	62	3,147 00	85,497 90	102	9,614 47	161,982 93
23	2,585 20	22,412 49	63	1,982 61	87,297 24	103	9,902 90	163,995 76
24	6,204 48	23,792 71	64	1,268 87	89,103 42	104	1,029 90	166,012 80
25	1,551 12	25,190 65	65	8,247 65	90,916 33	105	1,081 40	168,033 99
26	4,032 91	26,605 62	66	5,443 45	92,735 87	106	1,146 28	170,059 29
27	1,088 89	28,036 98	67	3,647 11	94,561 95	107	1,226 52	172,088 67
28	3,048 88	29,484 14	68	2,480 04	96,394 46	108	1,324 64	174,122 10
29	8,841 76	30,946 54	69	1,711 22	98,233 31	109	1,443 86	176,159 53
30	2,652 53	32,423 66	70	1,197 86	100,078 41	110	1,588 25	178,200 92
31	8,222 84	33,915 02	71	8,504 79	101,929 66	111	1,762 95	180,246 24
32	2,631 31	35,420 17	72	6,123 45	103 787 00	112	1,974 51	182,295 46
33	8,683 32	36,938 69	73	4,470 12	105,650 32	113	2,231 19	184,348 54
34	2,952 33	38,470 16	74	3,307 89	107,519 55	114	2,543 56	186,405 44
35	1,033 31	40,014 23	75	2,480 91	109,394 61	115	2,925 09	188,466 14
36	3,719 93	41,570 54	76	1,885 49	111,275 43	116	3,393 11	190,530 60
37	1,376 38	43,138 74	77	1,451 83	113,161 92	117	3,969 94	192,598 78
38	5,230 23	44,718 52	78	1,132 43	115,054 01	118	4,684 53	194,670 67
39	2,039 79	46,309 59	79	8,946 18	116,951 64	119	5,574 59	196,746 21
40	8,159 15	47,911 65	80	7,156 95	118,854 73	120	6,689 50	198,825 39

Tablica 2 (cd.)

n	$n!:10^c$	$\log n! = c + m$	n	$n!:10^c$	$\log n! = c + m$	n	$n!:10^c$	$\log n! = c + m$
121	8,094 30	200,908 18	161	7,590 70	286,880 28	201	1,585 20	377,200 08
122	9,875 04	202,994 54	162	1,229 69	289,089 80	202	3,202 11	379,505 44
123	1,214 63	205,084 44	163	2,004 40	291,301 98	203	6,500 28	381,812 93
124	1,506 14	207,177 87	164	3,287 22	293,516 83	204	1,326 06	384,122 56
125	1,882 68	209,274 78	165	5,423 91	295,734 31	205	2,718 42	386,434 32
126	2,372 17	211,375 15	166	9,003 69	297,954 42	206	5,599 94	388,748 18
127	3,012 66	213,478 95	167	1,503 62	300,177 14	207	1,159 19	391,064 15
128	3,856 20	215,586 16	168	2,526 08	302,402 45	208	2,411 11	393,382 22
129	4,974 50	217,696 75	169	4,269 07	304,630 33	209	5,039 22	395,702 36
130	6,466 86	219,810 69	170	7,257 42	306,860 78	210	1,058 24	298,024 58
131	8,471 58	221,927 96	171	1,241 02	309,093 78	211	2,232 88	400,348 87
132	1,118 25	224,048 54	172	2,134 55	311,329 31	212	4,733 70	402,675 20
133	1,487 27	226,172 39	173	3,692 77	313,567 35	213	1,008 28	405,003 58
134	1,992 94	228,299 49	174	6,425 43	315,807 90	214	2,157 72	407,333 99
135	2,690 47	230,429 83	175	1,124 45	318,050 94	215	4,639 09	409,666 43
136	3,659 04	232,563 37	176	1,979 03	320,296 45	216	1,002 04	412,000 89
137	5,012 89	234,700 09	177	3,502 89	322,544 43	217	2,174 43	414,337 35
138	6,917 79	236,839 97	178	6,235 14	324,794 85	218	4,740 27	416,675 80
139	9,615 72	238,982 98	179	1,116 09	327,047 70	219	1,038 12	419,016 25
140	1,346 20	241,129 11	180	2,008 96	329,302 97	220	2,283 86	421,358 67
141	1,898 14	243,278 33	181	3,636 22	331,560 65	221	5,047 33	423,703 06
142	2,695 36	245,430 62	182	6,617 92	333,820 72	222	1,120 51	426,049 41
143	3,854 37	247,585 95	183	1,211 08	336,083 17	223	2,498 73	428,397 72
144	5,550 29	249,744 32	184	2,228 39	338,347 99	224	5,597 16	430,747 97
145	8,047 93	251,905 68	185	4,122 51	340,615 16	225	1,259 36	433,100 15
146	1,175 00	254,070 04	186	7,667 87	342,884 68	226	2,846 16	435,454 26
147	1,727 25	256,237 35	187	1,433 89	345,156 52	227	6,460 77	437,810 28
148	2,556 32	258,407 62	188	2,695 72	347,430 67	228	1,473 06	440,168 22
149	3,808 92	260,580 80	189	5,094 91	349,707 14	229	3,373 30	442,528 05
150	5,713 38	262,756 89	190	9,680 32	351,985 89	230	7,758 59	444,889 78
151	8,627 21	264,935 87	191	1,848 94	354,266 92	231	1,792 23	447,253 39
152	1,311 34	267,117 71	192	3,549 97	356,550 22	232	4,157 98	449,618 88
153	2,006 34	269,302 41	193	6,851 44	358,835 78	233	9,688 10	451,986 24
154	3,089 77	271,489 93	194	1,329 18	361,123 58	234	2,267 02	454,355 45
155	4,789 14	273,680 26	195	2,591 90	363,413 62	235	5,327 49	456,726 52
156	7,471 06	275,873 38	196	5,080 12	365 705 87	236	1,257 29	459,099 43
157	1,172 96	278,069 28	197	1,000 78	368,000 34	237	2,979 77	461,474 18
158	1,853 27	280,267 94	198	1,981 55	370,297 01	238	7,091 85	463,850 76
159	2,946 70	282,469 34	199	3,943 29	372,595 86	239	1,694 95	466,229 16
160	4,714 72	284,673 46	200	7,886 58	374,896 89	240	4,067 89	468,609 37

Tablica 2 (cd.)

n	$n!:10^c$	$\log n! = c+m$	n	$n!:10^c$	$\log n! = c+m$	n	$n!:10^c$	$\log n! = c+m$
241	9,803 60	470,991 39	261	9,996 81	518,999 86	281	4,713 01	567,673 30
242	2,372 47	473,375 20	262	2,619 16	521,418 16	282	1,329 07	570,123 55
243	5,765 11	475,760 81	263	6,888 40	523,838 12	283	3,761 26	572,575 33
244	1,406 69	478,148 20	264	1,818 54	526,259 72	284	1,068 20	575,028 65
245	3,446 38	480,537 36	265	4,819 13	528,682 97	285	3,044 37	577,483 50
246	8,478 10	482,928 30	266	1,281 89	531,107 85	286	8,706 89	579,939 86
247	2,094 09	485, 21 00	267	3,422 64	533,534 36	287	2,498 88	582,397 75
248	5,193 34	487,715 45	268	9,172 68	535,962 50	288	7,196 77	584,857 14
249	1,293 14	490,111 65	269	2,467 45	538,392 25	289	2,079 87	587,318 04
250	3,232 86	492,509 59	270	6,662 11	540,823 61	290	6,031 61	589,780 43
251	8,114 47	494,909 26	271	1,805 43	543,256 58	291	1,755 20	592,244 33
252	2,044 85	497,310 66	272	4,910 78	545,691 15	292	5,125 18	594,709 71
253	5,173 46	499,713 78	273	1,340 64	548,127 31	293	1,501 63	597,176 58
254	1,314 06	502,118 61	274	3,673 36	550,565 06	294	4,414 93	599,644 92
255	3,350 85	504,525 16	275	1,010 17	553,004 40	295	1,302 41	602,114 75
256	8,578 18	506,933 40	276	2,788 08	555,445 31	296	3,855 12	604,586 04
257	2,204 59	509,343 33	277	7,722 98	557,887 79	297	1,144 97	607,058 79
258	5,687 85	511,754 95	278	2,146 99	560,331 83	298	3,412 01	609,533 01
259	1,473 15	514,168 25	279	5,990 10	562,777 43	299	1,020 19	612,008 68
260	3,830 20	516,583 22	280	1,677 23	565,224 59	300	3,060 58	614,485 80

Tablica 3. Wartości funkcji wykładniczej exp $(-x)$

x	0	1	2	3	4	5	6	7	8	9
0,00	1,	0,9990	0,9980	0,9970	0,9960	0,9950	0,9940	0,9930	0,9920	0,9910
0,0	1,	0,9900	0,9802	0,9704	0,9608	0,9512	0,9418	0,9324	0,9231	0,9139
1	0,9048	,8958	,8869	,8781	,8694	,8607	,8521	,8437	,8353	,8270
2	,8187	,8106	,8025	,7945	,7866	,7788	,7711	,7634	,7558	,7483
3	,7408	,7334	,7261	,7189	,7118	,7047	,6977	,6907	,6839	,6771
4	,6703	,6637	,6570	,6505	,6440	,6376	,6313	,6250	,6188	,6126
0,5	0,6065	0,6005	0,5945	0,5886	0,5827	0,5769	0,5712	0,5655	0,5599	0,5543
6	,5488	,5434	,5379	,5326	,5273	,5220	,5169	,5117	,5066	,5016
7	,4966	,4916	,4868	,4819	,4771	,4724	,4677	,4630	,4584	,4538
8	,4493	,4449	,4404	,4360	,4317	,4274	,4232	,4190	,4148	,4107
9	,4066	,4025	,3985	,3946	,3906	,3867	,3829	,3791	,3753	,3716
1,	0,3679	0,3329	0,3012	0,2725	0,2466	0,2231	0,2019	0,1827	0,1653	0,1496
2,	,1353	,1225	,1108	,1003	,0907	,0821	,0743	,0672	,0608	,0550
3,	,0498	0,450	,0408	,0369	,0334	,0302	,0273	,0247	,0224	,0202
4,	,0183	,0166	,0150	,0136	,0123	,0111	,0100	,0091	,0082	,0074
5,	,0067	,0061	,0055	,0050	,0045	,0041	,0037	,0034	,0030	,0027
6,	0,0025	0,0022	0,0020	0,0018	0,0017	0,0015	0,0014	0,0012	0,0011	0,0010
7,	,0009	,0008	,0007	,0007	,0006	,0006	,0005	,0005	,0004	,0004
8,	,0003	,0003	,0003	,0002	,0002	,0002	,0002	,0002	,0002	,0001
9,	,0001	,0001	,0001	,0001	,0001	,0001	,0001	,0001	,0001	,0000

Tablica 4. Wartości $P(k;\ \lambda)$ funkcji rozkładu prawdopodobieństwa Poissona

k	λ						
	0,001	0,002	0,003	0,004	0,005	0,006	0,007
0	0,9990	0,9980	0,9970	0,9960	0,9950	0,9940	0,9930
1	$,0^{3}999$	$,0^{2}200$	$,0^{2}299$	$,0^{2}398$	$,0^{2}498$	$,0^{2}596$	$,0^{2}695$
2	$,0^{5}1$	$,0^{5}200$	$,0^{5}449$	$,0^{5}797$	$,0^{4}124$	$,0^{4}179$	$,0^{4}243$

k	λ						
	0,008	0,009	0,010	0,020	0,030	0,040	0,050
0	0,9920	0,9910	0,9900	0,9802	0,9704	0,9608	0,9512
1	$,0^{2}794$	$,0^{2}892$	$,0^{2}990$	,0196	,0291	,0384	,0476
2	$,0^{4}317$	$,0^{4}401$	$,0^{4}495$	$,0^{3}196$	$,0^{3}437$	$,0^{3}769$	$,0^{2}119$
3	–	–	–	$,0^{5}131$	$,0^{5}437$	$,0^{4}102$	$,0^{4}198$

k	λ						
	0,06	0,07	0,08	0,09	0,10	0,20	0,30
0	0,9418	0,9324	0,9231	0,9139	0,9048	0,8187	0,7408
1	,0565	,0653	,0738	,0823	,0905	,1637	,2222
2	$,0^{2}170$	$,0^{2}228$	$,0^{2}295$	$,0^{2}370$	$,0^{2}452$	,0164	,0333
3	$,0^{4}339$	$,0^{4}533$	$,0^{4}788$	$,0^{3}111$	$,0^{3}151$	$,0^{2}109$	$,0^{2}333$
4	$0,^{5}1$	$,0^{6}93$	$,0^{5}158$	$,0^{5}250$	$,0^{5}377$	$,0^{4}546$	$,0^{3}250$
5	–	–	–	–	–	$,0^{5}218$	$,0^{4}150$

k	λ						
	0,4	0,5	0,6	0,7	0,8	0,9	1,0
0	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	0,3679
1	,2681	,3033	,3293	,3476	,3595	,3659	,3679
2	,0536	,0758	,0988	,1217	,1438	,1647	,1839
3	$,0^{2}715$	,0126	,0198	,0284	,0383	,0494	,0613
4	$,0^{3}715$	$,0^{2}158$	$,0^{2}296$	$,0^{2}497$	$,0^{2}767$	,0111	,0153
5	$,0^{4}572$	$,0^{3}158$	$,0^{3}356$	$,0^{3}696$	$,0^{2}123$	$,0^{2}200$	$,0^{2}307$
6	$,0^{5}381$	$,0^{4}132$	$,0^{4}356$	$,0^{4}811$	$,0^{3}164$	$,0^{3}300$	$,0^{3}511$
7	–	$,0^{5}1$	$,0^{5}305$	$,0^{5}811$	$,0^{4}187$	$,0^{4}386$	$,0^{4}730$
8	–	–	–	$,0^{6}1$	$,0^{5}187$	$,0^{5}434$	$,0^{5}912$
9	–	–	–	–	–	–	$,0^{6}101$

Tablica 4 (cd.)

x	λ									
	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2,0
0	0,332871	0,301194	0,272532	0,246597	0,223130	0,201897	0,182684	0,165299	0,149569	0,135335
1	,366158	,361433	,354291	,345236	,334695	,323034	,310562	,297538	,284180	,270671
2	,201387	,216860	,230289	,241665	,251021	,258428	,263978	,267784	,269971	,270671
3	,073842	,086744	,099792	,112777	,125510	,137828	,149587	,160671	,170982	,180447
4	,020307	,026023	,032432	,039472	,047067	,055131	,063575	,072302	,081216	,090224
5	,004467	,006246	,008432	,011052	,014120	,017642	,021615	,026029	,030862	,036089
6	,000819	,001249	,001827	,002579	,003530	,004705	,006124	,007809	,009773	,012030
7	,000129	,000214	,000339	,000516	,000756	,001075	,001487	,002008	,002653	,003437
8	,000018	,000032	,000055	,000090	,000142	,000215	,000316	,000452	,000630	,000859
9	,000002	,000004	,000008	,000014	,000024	,000038	,000060	,000090	,000133	,000191
10	–	,000001	,000001	,000002	,000004	,000006	,000010	,000016	,000025	,000038
11	–	–	–	–	–	,000001	,000002	,000003	,000004	,000007
12	–	–	–	–	–	–	–	–	,000001	,000001

x	λ									
	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9	3,0
0	0,122456	0,110803	0,100259	0,090718	0,082085	0,074274	0,067206	0,060810	0,055023	0,049787
1	,257159	,243767	,230595	,217723	,205212	,193111	,181455	,170268	,159567	,149361
2	,270016	,268144	,265185	,261268	,256516	,251045	,244964	,238375	,231373	,224042
3	,189012	,196639	,203308	,209014	,213763	,217572	,220468	,222484	,223660	,224042
4	,099231	,108151	,116902	,125409	,133602	,141422	,148816	,155739	,162154	,168031
5	,041677	,047587	,053775	,060196	,066801	,073539	,080360	,087214	,094049	,100819
6	,014587	,017448	,020614	,024078	,027834	,031867	,036162	,040700	,045457	,050409
7	,004376	,005484	,006773	,008255	,009941	,011836	,013948	,016280	,018832	,021604
8	,001149	,001508	,001947	,002477	,003106	,003847	,004708	,005698	,006827	,008102
9	,000268	,000369	,000498	,000660	,000863	,001111	,001412	,001773	,002200	,002701
10	,000056	,000081	,000114	,000158	,000216	,000289	,000381	,000496	,000638	,000810
11	,000011	,000016	,000024	,000035	,000049	,000068	,000094	,000126	,000168	,000221
12	,000002	,000003	,000005	,000007	,000010	,000015	,000021	,000029	,000041	,000055
13	–	,000001	,000001	,000001	,000002	,000003	,000004	,000006	,000009	,000013
14	–	–	–	–	–	,000001	,000001	,000001	,000002	,000003
15	–	–	–	–	–	–	–	–	–	,000001

Tablica 4 (cd.)

x	λ									
	3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	3,9	4,0
0	0,045049	0,040762	0,036883	0,033373	0,030197	0,027324	0,024724	0,022371	0,020242	0,018316
1	,139653	,130439	,121714	,113469	,105691	,098365	,091477	,085009	,078943	,073263
2	,216461	,208702	,200829	,192898	,184959	,177058	,169233	,161517	,153940	,146525
3	,223677	,222616	,220912	,218617	,215785	,212469	,208720	,204588	,200122	,195367
4	,173350	,178093	,182252	,185825	,188812	,191222	,193066	,194359	,195119	,195367
5	,107477	,113979	,120286	,126361	,132169	,137680	,142869	,147713	,152193	,156293
6	,055530	,060789	,066158	,071604	,077098	,082608	,088102	0,93551	0,98925	,104196
7	,024592	,027789	,031189	,034779	,038549	,042484	,046568	,050785	,055115	,059540
8	,009529	,011116	,012865	,014781	,016865	,019118	,021538	,024123	,026869	,029770
9	,003282	,003952	,004717	,005584	,006559	,007647	,008854	,010185	,011643	,013231
10	,001018	,001265	,001557	,001899	,002296	,002753	,003276	,003870	,004541	,005292
11	,000287	,000368	,000467	,000587	,000730	,000901	,001102	,001337	,001610	,001925
12	,000074	,000098	,000128	,000166	,000213	,000270	,000340	,000423	,000523	,000642
13	,000018	,000024	,000033	,000043	,000057	,000075	,000097	,000124	,000157	,000197
14	,000004	,000006	,000008	,000011	,000014	,000019	,000026	,000034	,000044	,000056
15	,000001	,000001	,000002	,000002	,000003	,000005	,000006	,000009	,000011	,000015
16	–	–	–	,000001	,000001	,000001	,000001	,000002	,000003	,000004
17	–	–	–	–	–	–	–	–	,000001	,000001

x	λ									
	4,1	4,2	4,3	4,4	4,5	4,6	4,7	4,8	4,9	5,0
0	0,016573	0,014996	0,013569	0,012277	0,011109	0,010052	0,009095	0,008230	0,007447	0,006738
1	,067948	,062981	,058345	,054020	,049990	,046238	,042748	,039503	,036488	,033690
2	,139293	,132261	,125441	,118845	,112479	,106348	,100457	,094807	,089396	,084224
3	,190368	,185165	,179799	,174305	,168718	,163068	,157383	,151691	,146014	,140374
4	,195127	,194424	,193284	,191736	,189808	,187528	,184925	,182029	,178867	,175467
5	,160004	,163316	,166224	,168728	,170827	,172525	,173830	,174748	,175290	,175467
6	,109336	,114321	,119127	,123734	,128120	,132270	,136167	,139798	,143153	,146223
7	,064040	,068593	,073178	,077775	,082363	,086920	,091426	,095862	,100207	,104445
8	,032820	,036011	,039333	,042776	,046329	,049979	,053713	,057517	,061377	,065278
9	,014951	,016805	,018793	,020913	,023165	,025545	,028050	,030676	,033416	,036266
10	,006130	,007058	,008081	,009202	,010424	,011751	,013184	,014724	,016374	,018133
11	,002285	,002695	,003159	,003681	,004264	,004914	,005633	,006425	,007294	,008242
12	,000781	,000943	,001132	,001350	,001599	,001884	,002206	,002570	,002978	,003434
13	,000246	,000305	,000374	,000457	,000554	,000667	,000798	,000949	,001123	,001321
14	,000072	,000091	,000115	,000144	,000178	,000219	,000268	,000325	,000393	,000472
15	,000020	,000026	,000033	,000042	,000053	,000067	,000084	,000104	,000128	,000157
16	,000005	,000007	,000009	,000012	,000015	,000019	,000025	,000031	,000039	,000049
17	,000001	,000002	,000002	,000003	,000004	,000005	,000007	,000009	,000011	,000014
18	–	–	,000001	,000001	,000001	,000001	,000002	,000002	,000003	,000004
19	–	–	–	–	–	–	–	,000001	,000001	,000001

Tablica 5. Dystrybuanta rozkładu normalnego $N(0, 1)$

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\tfrac{1}{2}u^2\right)du$$

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	,5398	,5438	,5478	,5517	,5557	,5596	,5636	,5675	,5714	,5753
0,2	,5793	,5832	,5861	,5910	,5948	,5987	,6026	,6064	,6103	,6141
0,3	,6179	,6217	,6255	,6293	,6331	,6368	,6406	,6443	,6480	,6517
0,4	,6554	,6591	,6628	,6664	,6700	,6736	,6772	,6808	,6844	,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	,7257	,7291	,7324	,7357	,7389	,7422	,7454	,7486	,7517	,7549
0,7	,7580	,7611	,7642	,7673	,7703	,7734	,7764	,7794	,7823	,7852
0,8	,7881	,7910	,7939	,7967	,7995	,8023	,8051	,8078	,8106	,8133
0,9	,8159	,8186	,8212	,8238	,8264	,8289	,8315	,8340	,8365	,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	,8643	,8665	,8686	,8708	,8729	,8749	,8770	,8790	,8810	,8830
1,2	,8849	,8869	,8888	,8907	,8925	,8944	,8962	,8980	,8997	,90147
1,3	,90320	,90490	,90658	,90824	,90988	,91149	,91309	,91466	,91621	,91774
1,4	,91924	,92073	,92220	,92354	,92507	,92647	,92785	,92922	,93056	,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	,94520	,94630	,94738	,94845	,94950	,95053	,95154	,95254	,95352	,95449
1,7	,95543	,95637	,95728	,95818	,95907	,95994	,96080	,96164	,96246	,96327
1,8	,96407	,96485	,96562	,96638	,96712	,96784	,96856	,96926	,96995	,97062
1,9	,97128	,97193	,97257	,97320	,97381	,97441	,97500	,97558	,97615	,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	,98214	,98257	,98300	,98341	,98382	,98422	,98461	,98500	,98537	,98574
2,2	,98610	,98645	,98679	,98713	,98745	,98778	,98809	,98840	,98870	,98899
2,3	,98928	,98956	,98983	,9^{2}0097	,9^{2}0358	,9^{2}0613	,9^{2}0863	,9^{2}1106	,9^{2}1344	,9^{2}1576
2,4	,9^{2}1802	,9^{2}2024	,9^{2}2240	,9^{2}2451	,9^{2}2656	,9^{2}2857	,9^{2}3053	,9^{2}3244	,9^{2}3431	,9^{2}3613
2,5	0,9^{2}3790	0,9^{2}3963	0,9^{2}4132	0,9^{2}4297	0,9^{2}4457	0,9^{2}4614	0,9^{2}4766	0,9^{2}4915	0,9^{2}5060	0,9^{2}5201
2,6	,9^{2}5339	,9^{2}5473	,9^{2}5604	,9^{2}5731	,9^{2}5844	,9^{2}5975	,9^{2}6093	,9^{2}6207	,9^{2}6319	,9^{2}6427
2,7	,9^{2}6533	,9^{2}6636	,9^{2}6736	,9^{2}6833	,9^{2}6928	,9^{2}7020	,9^{2}7110	,9^{2}7197	,9^{2}7282	,9^{2}7365
2,8	,9^{2}7445	,9^{2}7523	,9^{2}7599	,9^{2}7673	,9^{2}7744	,9^{2}7814	,9^{2}7882	,9^{2}7948	,9^{2}8012	,9^{2}8074
2,9	,9^{2}8134	,9^{2}8193	,9^{2}8250	,9^{2}8305	,9^{2}8359	,9^{2}8411	,9^{2}8462	,9^{2}8511	,9^{2}8559	,9^{2}8605
3,0	0,9^{2}8650	0,9^{2}8694	0,9^{2}8736	0,9^{2}8777	0,9^{2}8817	0,9^{2}8856	0,9^{2}8893	0,9^{2}8930	0,9^{2}8965	0,9^{2}8999
3,1	,9^{3}0324	,9^{3}0646	,9^{3}0957	,9^{3}1260	,9^{3}1553	,9^{3}1836	,9^{3}2112	,9^{3}2378	,9^{3}2636	,9^{3}2886
3,2	,9^{3}3129	,9^{3}3363	,9^{3}3590	,9^{3}3810	,9^{3}4002	,9^{3}4230	,9^{3}4429	,9^{3}4623	,9^{3}4810	,9^{3}4991
3,3	,9^{3}5166	,9^{3}5335	,9^{3}5499	,9^{3}5658	,9^{3}5811	,9^{3}5959	,9^{3}6103	,9^{3}6242	,9^{3}6376	,9^{3}6505
3,4	,9^{3}6631	,9^{3}6752	,9^{3}6869	,9^{3}6982	,9^{3}7091	,9^{3}7197	,9^{3}7299	,9^{3}7398	,9^{3}7493	,9^{3}7585

LITERATURA

[1] Benjamin I. R., Cornell C. A., *Rachunek prawdopodobieństwa, statystyka matematyczna i teoria decyzji*, Warszawa 1977.

[2] Beyer O., Hackal H., Pieper V., Tiedge I., *Wahrscheinlichkeitsrechnung und mathematische Statistik*, Leipzig 1976.

[3] Bobrowski D., *Probabilistyka w zastosowaniach technicznych*, Warszawa 1980.

[4] Cramer H., *Metody matematyczne w statystyce*, PWN, Warszawa 1958.

[5] Diner I. J. i inni, *Rachunek prawdopodobieństwa w problemach i zadaniach*, Warszawa 1979.

[6] Feller W., *Wstęp do rachunku prawdopodobieństwa*, PWN, T1, Warszawa 1980, T2, Warszawa 1981.

[7] Fisz M., *Rachunek prawdopodobieństwa i statystyka matematyczna*, Warszawa 1969.

[8] Gerstenkorn T., Śródka T., *Kombinatoryka i rachunek prawdopodobieństwa*, wyd. 6, Warszawa 1981.

[9] Gniedenko B. W. [Гнеденко Б. В.], *Курс теории вероятностей*, Москва 1969.

[10] Kendall M., Stuart A., *The advanced theory of statistics*, vol 1, Ch. Griffin, London (ros. tłum. pt. *Теория распределений*, Москва 1966).

[11] Kubik L. T., *Rachunek prawdopodobieństwa*, Warszawa 1980.

[12] Plucińska A., Pluciński E., *Elementy probabilistyki*, Warszawa 1981.

[13] Prochorow J. W., Rozanow J. A., *Rachunek prawdopodobieństwa*, Warszawa 1972.

[14] Richter H., *Wahrscheinlichkeitstheorie*, Berlin 1956.

[15] Smirnow N. W., Dunin-Barkowski J. W., *Kurs rachunku prawdopodobieństwa i statystyki matematycznej dla zastosowań technicznych*, Warszawa 1969.

[16] Zieliński R., *Tablice statystyczne*, Warszawa 1972.

[17] Zubrzycki S., *Wykłady z rachunku prawdopodobieństwa i statystyki matematycznej*, Warszawa 1970.

Norma

Rachunek prawdopodobieństwa i statystyka matematyczna PN-74/N-01051.

SKOROWIDZ

SPIS RZECZY

Wydawnictwo Naukowe PWN SA
Wydanie siódme
Arkuszy drukarskich 16,00
Druk ukończono w marcu 1999 r.
Drukarnia Wydawnictw Naukowych SA
Łódź, ul. Żwirki 2