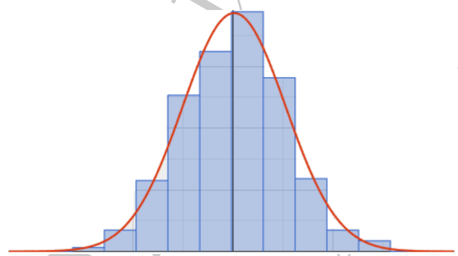
ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ НАУК ТА
КІБЕРНЕТИКИ КИЇВСЬКОГО НАЦІОНАЛЬНОГО
УНІВЕРСИТЕТУ
ІМЕНІ ТАРАСА ШЕВЧЕНКА

ЗБІРНИК ЗАДАЧ З ТЕОРІЇ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ

Навчальний посібник



Шарапов М.М., Розора І.В., Лівінська Г.В., Пономарьов В.Д.

2023

УДК 519.2 (075.8)
ББК 22.1я73

Рецензенти:
д-р ф.-м.н., доцентка, професорка кафедри математичного аналізу
та теорії ймовірностей НТУ "КПІ ім. І.Сікорського" О.І. Василик;
к.ф.-м.н., доцентка кафедри теорії ймовірностей, статистики
та актуарної математики КНУ ім. Т.Шевченка Т.О. Яневич.

Рекомендовано до друку вченою радою факультету комп'ютерних наук та кібернетики
(протокол № 10 від 25 квітня 2023)

Шарапов М.М.

Збірник задач з теорії ймовірностей та математичної статистики: навч. посіб. для студ. ф-ту КНК КНУ ім. Т. Шевченка. / М.М. Шарапов, І.В. Розора, Г.В. Лівінська , В.Д. Пономарьов – К. : 2023. – 326 с.

У посібнику зібрано задачі з теорії ймовірностей та математичної статистики, які пропонувались студентам факультету комп'ютерних наук та кібернетики Київського національного університету ім. Тараса Шевченка. До кожного розділу подано необхідний теоретичний матеріал.

Для студентів факультету комп'ютерних наук та кібернетики КНУ ім. Т. Шевченка.

1 СТОХАСТИЧНИЙ ЕКСПЕРИМЕНТ І ПРОСТІР ЕЛЕМЕНТАРНИХ ПОДІЙ

Стохастичним експериментом називають експеримент, результат якого не можна передбачити напевне. Результати ω експерименту називають **елементарними подіями**. Сукупність усіх можливих результатів експерименту, тобто сукупність усіх елементарних подій, будемо називати **простором елементарних подій** і позначати літерою Ω. Простір елементарних подій називається дискретним, якщо множина Ω скінченна або зліченна. Підмножини Ω називаються **подіями**. Сама множина Ω називається **достовірною подією**, а порожня множина $\varnothing$ **неможливою подією**.

Будемо говорити, що при здійсненні експерименту відбулася подія A, якщо як результат ми отримали ω_0 і $\omega_0 \in A$. При цьому про елементарну подію ω_0 говорять як про таку, що **сприяє події A** або **тягне за собою подію A**.

Сумою подій A і B називається подія C, яка відбувається лише тоді, коли відбувається подія A або подія B. Позначення: $C = A \cup B$.

Добутком подій A і B називається подія C, яка відбувається лише тоді, коли відбувається і подія A, і подія B. Позначення: $C = A \cap B$ або $C = AB$.

Різницею подій A і B називається подія C, яка відбувається лише тоді, коли відбувається подія A, і не відбувається подія B. В цьому випадку пишуть $C = A \backslash B$.

Подія $\Omega \backslash A$ називається **протилежною до події A** (доповненням до події A, запереченням події A) і позначається як $\bar{A}$ (чи $\neg A$).

Події A та B називаються **несумісними**, якщо вони не можуть трапитися одночасно, тобто $A \cap B = \varnothing$. В противному випадку події A та B називаються **сумісними**.

Якщо кожна елементарна подія, яка сприяє події A, сприяє і події B, то говорять, що подія B випливає з події A, або подія A тягне за собою подію B. Це відношення між подіями записують у вигляді $A \subset B$ (або $B \supset A$).

Оскільки події є множинами, то для операцій над подіями справедливі ті ж самі правила, що і для операцій над множинами:

а) $A \cup B = B \cup A$, $A \cap B = B \cap A$ комутативність суми та добутку;

б) $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$ - асоціативність суми та добутку;

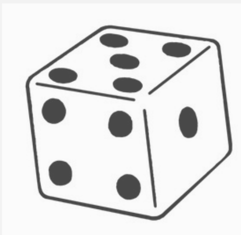
в) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ - розподільні закони добутку відносно додавання та додавання відносно добутку;

г) $\overline{A \cup B} = \overline{A} \cap \overline{B}$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$ - правила де Моргана.

Послідовність подій A_n називається **зростаючою**, якщо $A_1 \subset A_2 \subset A_3 \subset \dots$

Послідовність подій B_n називається **спадною**, якщо $B_1 \supset B_2 \supset B_3 \supset \dots$

Приклад 1.1



Стохастичний експеримент полягає в тому, що підкидається симетричний гральний кубик. Експеримент можна повторювати як завгодно довго. При цьому передбачити напевно, як впаде кубик, неможливо. Результатом цього експерименту є число від 1 до 6, яке випадає в результаті підкидання.

Отже, простором елементарних подій буде множина

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Для цього стохастичного експерименту можна визначити, наприклад, такі події:

$A = \{\text{Випала парна кількість очок}\} = \{2, 4, 6\}$;

$B = \{\text{Кількість очок, що випала, менша чотирьох}\} = \{1, 2, 3\}$;

$C = \{\text{Кількість очок, що випала, не менша 5}\} = \{5, 6\}$.

Зауважимо, що події B та C не можуть трапитися одночасно, вони взаємно виключають одна одну, отже є несумісними подіями. Натомість події A та C можуть трапитись одночасно в результаті експерименту, якщо випаде 6. Отже, ці події є сумісними.

Приклад 1.2

Експеримент полягає в проведенні серії пострілів по мішені (качці) до першого влучення. Результатом експерименту буде серія пострілів, останній з яких – вдалий. В даному випадку маємо справу з нескінченним (але зліченим) простором елементарних подій, оскільки теоретично стріляти можна нескінченну кількість разів залежно від вправності стрільця, близькості мішені, тощо. Отже, простір елементарних подій можна записати таким чином:

$$\Omega = \{\text{У, НУ, ННУ, ННУ, ...}\},$$

де У – це "успіх" (влучення в мішень), а Н – це "невдача", тобто невлучення в мішень при одному пострілі.

Сформулюємо деякі події:

$A = \{$Качку вбито при першому пострілі$\} = \{$У$\}$;

$B = \{$Качку вбито не більше ніж за три постріли$\} = \{$У, НУ, ННУ$\}$;

$C = \{$Не вдалося вполювати качку$\} = \Omega \setminus \{$У, НУ, ННУ, ННУ, ...$\}$.

В останньому випадку подія C є протилежною до події, яку можна сформулювати як $\{$Мисливець на якомусь кроці таки влучить в качку$\}$.

Приклад 1.3

Експериментом є фіксація кількості ДТП в Києві впродовж доби (тижня, місяця, року). Елементарними подіями є будь-які цілі невід'ємні числа:

$$\Omega = \{0, 1, 2, ...\}.$$

Елементарними подіями є події

$$\omega_i = \{\text{Протягом доби трапилось рівно } i \text{ ДТП }\}, \quad i = 0, 1, 2, ...$$

Кількість елементарних подій формально є нескінченною. Аналітично обчислити ці ймовірності не вдасться, в даному випадку необхідна статистична інформація за попередні періоди (причому з урахуванням того, чи є день вихідним, святковим, робочим, тощо).

Можна розглядати, наприклад, такі події:

$A = \{$Не трапилось жодної ДТП протягом доби$\} = \{i = 0\}$;

$B = \{$Трапилось не більше 5 ДТП протягом доби$\} = \{i \leq 5\}$;

$C = \{$Трапилось більше 10 ДТП протягом доби$\} = \{i > 10\}$.

Приклад 1.4

Експеримент полягає в рівномірному киданні точки на інтервал $[0, 1]$.

Елементарною подією тут буде число, яке є координатою точки, а простором елементарних подій – сам відрізок $[0, 1]$. Всі елементарні події рівноможливі. Проте простір елементарних подій в даному випадку складається з незліченої кількості елементарних подій.

Подією є, наприклад, те, що точка потрапила на інтервал $[0; 1/3]$.

Так, зокрема, можна генерувати послідовності «успіхів» та «невдач», якщо використовувати генератор псевдовипадкових чисел, а «успіхом» вважати значення отриманого числа, меншого за певну величину (залежно від того, що в експерименті є «успіхом»).

Приклад 1.5

Два клієнти страхової компанії уклали угоди страхування на випадок крадіжки власних авто терміном на один рік. Страхові події з кожним із них можуть статися незалежним чином у будь-який день року. Описати
а) простір елементарних подій Ω;
б) подію A, яка полягає в тому, що відшкодування клієнтам доведеться заплатити в один день;
в) подію B, яка полягає в тому, що термін між сплатою відшкодувань буде не більшим ніж 30 днів.

Розв'язок:

а) Припустимо, що мова йде про рік, який не є високосним. Тоді за простір елементарних подій можна обрати множину пар (i, j), де кожна компонента i та j приймає значення від 1 до 365, тобто

$$\Omega = \left\{(i, j) : i = \overline{1, 365},\ j = \overline{1, 365}\right\}.$$

б) В цих позначеннях подія A буде виглядати як множина пар з однаковими компонентами, а саме

$$A = \left\{(i, j) : i = j,\ j = \overline{1, 365}\right\}.$$

в) Подібним же чином подія B запишеться так:

$$B = \left\{(i,j) : |i-j| \le 30,\ i=\overline{1,365},\ j=\overline{1,365}\right\}.$$

Задача 1.1. Перевірити чи спростувати справедливість наступних співвідношень:

1. $\overline{A \cup B} = \overline{A}\,\overline{B}$;
2. $\overline{A} \cup \overline{B} = \overline{AB}$;
3. $A \cup B \backslash B = A \backslash AB = A\overline{B}$;
4. $AA = A \cup A = A$;
5. $A \cup B \backslash AB = A\overline{B} \cup \overline{A}B$;
6. $(A \cup B)\, C = AC \cup BC$;
7. $\overline{A \cup B \cup C} = \overline{A}\,\overline{B}\,\overline{C}$;
8. $\left(\overline{A \cup B}\right) C = \overline{A}\,\overline{B}\,\overline{C} = C \backslash C\,(A \cup B)$;
9. $A\overline{B}C \subset A \cup B$;
10. $(A \backslash AB) \cup B = A \cup B$;
11. $AB \cup BC \cup CA \supset ABC$;
12. $AB \cup BC \cup CA \subset C\,(A \cup B \cup C)$;
13. $\overline{A} \cup \overline{B} = \overline{AB}$;
14. $A \cup AB = A$;
15. $A \cup B = (A \backslash B) \cup (B \backslash A) \cup AB$;
16. $A \cup B \backslash B = A \backslash B$;
17. $(A \backslash B) \cup B = A \cup B$;
18. $(A \cup B)\, C = AC \cup BC$;
19. $(A \cup C)\,(B \cup C) = AB \cup C$;
20. $AC \backslash B = AC \backslash BC$;
21. $(A \backslash B) \cup (A \backslash C) = A \backslash BC$;
22. $\left(\overline{A} \cup BC\right)\left(\overline{B} \cup AC\right)\left(\overline{C} \cup AB\right) = ABC \cup \overline{A}\,\overline{B}\,\overline{C}$;
23. $(A \cup B)\left(\overline{A} \cup B\right)\left(A \cup \overline{B}\right) = AB$;
24. $(A \cup B)\left(A \cup \overline{B}\right) = A$.

Задача 1.2. Нехай A, B, C – три довільні події. Знайти вирази для подій, що полягають у тому, що з A, B, C:

1. відбулася тільки подія A;
2. відбулися A та B, але C не відбулася;
3. відбулися усі три події;
4. відбулась принаймні одна із цих подій;
5. відбулись принаймні дві події;
6. відбулася одна і тільки одна подія;
7. відбулися дві і тільки дві події;
8. жодна з подій не відбулася;
9. відбулося не більше двох подій.

Задача 1.3. Нехай A_i, $i = 1, 2, 3$, – події, які полягають в тому, що i-тий магазин зачинено. Що означають події

1. $\overline{A_1 \cap A_2 \cap A_3}$;
2. $\overline{A_1} \cap \overline{A_2} \cup \overline{A_1} \cap \overline{A_3} \cup \overline{A_2} \cap \overline{A_3}$;
3. $A_1 \cap \overline{A_2} \cap A_3$;
4. $\overline{\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}}$;
5. $A_1 \cup A_2 \cup A_3$;
6. $A_1 \cap A_2 \cap A_3$.

Задача 1.4. Об'єднання $A \cup B$ двох подій може бути подане як об'єднання двох несумісних подій: $A \cup B = A \cup (B \backslash AB)$. Подати аналогічним чином об'єднання трьох подій A, B, C.

Задача 1.5. Витягнута навмання куля може виявитися або червоною (подія A, або білою (подія B), або чорною (подія C). Описати наступні події:
1) $A \cup B$; 2) $\overline{A \cup C}$; 3) AC; 4) $AB \cup C$.

Задача 1.6. Спростити вирази:

1. $(A \cup B)\left(A \cup \overline{B}\right)$;
2. $(A \cup B)(B \cup C)(C \cup A)$;
3. $A \cup (B \backslash AB) \cup (C \backslash AC)$;
4. $(A \cup B) B \cup A (AB)$.

Задача 1.7. Які з наступних співвідношень вірні?

1. $\overline{ABC} \subset AB \cup BC \cup CA$;
2. $\overline{A} \cup \overline{B} = \overline{A \cup B}$;
3. $\overline{A \cup B \cup C} = \overline{A}\ \overline{B}\ \overline{C}$;
4. $\overline{A \cup B \cup C} = \overline{A}\ \overline{BC}$;
5. $(A \cup B) \backslash C = A \cup (B \backslash C)$;
6. $ABC = AB (C \cup B)$;
7. $(A \cup B) \backslash A = B$;
8. $\left(\overline{A \cup B}\right) C = \overline{A}C \cup \overline{B}C$;
9. $\overline{A}BC \cup A\overline{B}C \cup AB\overline{C} = (AB \cup AC \cup BC) \backslash ABC$.

Задача 1.8. Знайти випадкову подію X із наступної рівності:

$$\left(A \cup \overline{X}\right) \cap \left(\overline{A} \cup \overline{X}\right) \cup \overline{X \cup A} \cup \overline{X \cup \overline{A}} = B.$$

Задача 1.9. Знайти випадкову подію X із наступної рівності:

$$\overline{X \cup A} \cup \overline{X \cup \overline{A}} = B.$$

Задача 1.10. Відновити наступні формули:

а) $A \cup A = ...$
б) $A \cap A = ...$
в) $A \cup \Omega = ...$
г) $A \cap \Omega = ...$
д) $\Omega \subset A \Rightarrow A = ...$
е) $A \subset \varnothing \Rightarrow A = ...$
ж) $A \cap B = A \Rightarrow ...$
и) $A \cup B = B \Rightarrow ...$
к) $A \backslash B = \varnothing \Rightarrow ...$
л) $\overline{A} \cup B = \Omega \Rightarrow ...$
м) $(B \backslash A) \cup A = B \Rightarrow ...$
н) $(A \cup B) \cap \left(A \cup \overline{B}\right) = ...$
п) $(A \backslash A \cap B) \cup B = ...$
р) $(A \cup B) \cap \left(A \cup \overline{B}\right) \cup \left(\overline{A} \cup B\right) \cap \left(\overline{A} \cup \overline{B}\right) = ...$
с) $(A \cup B) \cap \left(A \cup \overline{B}\right) \cap \left(\overline{A} \cup B\right) \cap \left(\overline{A} \cup \overline{B}\right) = ...$

Задача 1.11. Довести, що:

а) $A \backslash B = A \backslash (A \cap B)$;

б) $A = (A \cap B) \cup (A \backslash B)$;

в) $(A \cup B) \backslash C = (A \backslash C) \cup (B \backslash C)$;

г) $A \backslash (B \backslash C) = (A \backslash B) \cup (A \cap C)$;

д) $(A \backslash B) \cup B = A \cup B$;

е) $(A \cup B) \backslash B = A \backslash B$.

Задача 1.12. Відомо, що $A \subset B$. Потрібно дописати праву частину рівностей:
a) $A \cup B = ...$; b) $A \cap B = ...$; c) $A \backslash B = ...$; d) $A \cap B \cap C = ...$.

Задача 1.13. Подія E подається за допомогою подій A, B, C. Спростити це подання, якщо:

а) $E = (A \cup B)\left(\overline{A} \cup B\right)$;

б) $E = (A \cup B)\, C \cup (B \cup C)\, A \cup (C \cup A)\, B$;

в) $E = (A \cup B)(B \cup C)$ і $A \subset C$;

г) $E = (A \cup B)\left(\overline{A} \cup B\right)\left(A \cup \overline{B}\right)$.

Задача 1.14. Множини A, B, C є підмножинами ряду натуральних чисел. A є множиною чисел, які діляться на 6, B - множиною парних чисел, C - множиною чисел від 1 до 100. Описати наступні множини:
a) $A \backslash B$, b) $A \cup C$, c) $(A \cap B) \cup C$, d) $C \backslash (A \cap B)$, e) $A \backslash A \cap B \cap C$.

Задача 1.15. Підкидають два гральних кубики 3 рази. Описати простір елементарних подій.

Задача 1.16. Гральний кубик кидається до тих пір, доки підряд не випаде одиниця і двійка. Описати простір елементарних подій.

Задача 1.17. Нитку довжини l навмання розірвано на дві частини. Описати простір елементарних подій.

Задача 1.18. За допомогою операцій ”$\cup$”, ”$\cap$” та ”$^{-}$” описати наступні події:

a) відбулися всі три події A, B, C одночасно;

b) не відбулось жодної з подій A, B, C;

c) відбулась лише подія A;

d) відбулась лише одна з подій A, B, C;

e) відбулась принаймні одна з подій A, B, C;

f) відбулися лише події A і B;

g) відбулися лише дві події з A, B, C;

h) відбулися принаймні дві події з A, B, C;

i) відбулася не більше ніж одна з подій A, B, C.

Задача 1.19. Експеримент полягає у тому, що на піввісь $[0, \infty)$ кидається точка. Результатом експерименту будемо вважати $\omega = [x]$, де x - відстань від точки до початку координат, $[x]$ - ціла частина числа x. Подати приклади зростаючої і спадної послідовності подій, що пов'язані з цим експериментом.

(а) (б)

Рис. 1.1: до задачі 1.20

Задача 1.20. Операції над подіями можна демонструвати на діаграмах Ейлера. Так, наприклад, для подій A, B, C подію $D = A \cap B \cap C$ можна зобразити так, як показано на рис. 1.1а. Спробуйте розв'язати "обернену" задачу – за діаграмою 1.1б запишіть відповідну подію. Якщо це можна зробити різними способами, то зробіть це найменшою кількістю операцій над подіями.

(а) (б)

Рис. 1.2: до задачі 1.21

Задача 1.21. Підкидають один гральний кубик. Для такого стохастичного експерименту наведіть приклад трьох подій A, B, C, діаграма Ейлера для яких буде:

а) рис. 1.2а,

б) рис. 1.2б.

2 СКІНЧЕННИЙ ПРОСТІР ЕЛЕМЕНТАРНИХ ПОДІЙ

Коли простір елементарних подій є скінченним, тобто $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, то довільну підмножину з Ω називають **подією**, а кількість цих подій є, очевидно, 2^n. Кожній елементарній події ω_i з простору Ω ставиться у відповідність число $\mathrm{P}(\omega_i) = p_i$ (**ймовірність елементарного наслідку**) таке, що $0 \le p_i \le 1$ та $\sum\limits_{i=1}^{n} p_i = 1$. Тоді ймовірність довільної події A підраховують за формулою:

$$\mathrm{P}(A) = \sum_{\omega \in A} \mathrm{P}(\omega). \tag{2.1}$$

Для визначеної таким чином імовірності мають місце наступні властивості:

1) $\forall A \subset \Omega \quad 0 \le \mathrm{P}(A) \le 1$,

2) $\mathrm{P}(\Omega) = 1$, $\mathrm{P}(\varnothing) = 0$,

3) якщо $A \cap B = \varnothing$, то $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$.

Якщо у визначенні імовірності $\forall i\ p_i = 1/n$ (тобто всі елементарні наслідки рівноможливі), то таке визначення називають **класичним**. У цьому випадку

$$\mathrm{P}(A) = \frac{|A|}{|\Omega|}, \tag{2.2}$$

де $|A|$ – кількість елементів множини A.

При застосуванні останньої формули використовується, як правило, такий розділ математики як комбінаторика. Нижче наведені деякі формули та позначення з цього розділу, які найчастіше використовуються при підрахунках:

1) $n!$ – число різних можливих перестановок з n елементів;

2) $\dfrac{n!}{n_1!n_2!...n_m!}$ – число різних можливих перестановок з повтореннями, коли серед n елементів є n_1 елементів першого типу, n_2 елементів другого типу, ... , n_m елементів m-го типу;

3) $C_n^k = \dfrac{n!}{k!(n-k)!}$ – число комбінацій з n елементів по k, в яких порядок елементів не враховується;

4) $A_n^k = \dfrac{n!}{(n-k)!}$ – число комбінацій з n елементів по k, в яких враховується порядок елементів;

5) $\overline{C}_n^k = C_{k+n-1}^{n-1} = \dfrac{(k+n-1)!}{k!\,(n-1)!}$ – число комбінацій (без урахування порядку) з n типів елементів по k з повтореннями.

Правило добутку комбінаторики: якщо подія A_1 може відбутися n_1 різними способами, подія A_2 незалежно від цього може відбутися n_2 різними способами, подія A_3 незалежно від цього може відбутися n_3 різними способами... подія A_m незалежно від цього може відбутися n_m різними способами, то послідовність подій $A_1, A_2, \ldots, A_m$ може відбутися $n_1 \cdot n_2 \cdot \ldots \cdot n_m$ різними способами.

Приклад 2.1

В номері автомобіля спочатку записані дві букви, потім 4 цифри, потім ще дві букви.

а) Скільки різних номерних знаків може бути сформовано в такий спосіб, якщо використовувати 17 літер українського алфавіту та будь-які з 10 цифр?

б) Скільки може бути номерних знаків, в яких всі букви та цифри різні?

в) всі цифри однакові, а всі букви різні?

г) всі цифри однакові і букви однакові?

д) всі цифри різні?

е) цифрова частина складається з цифр 1, 2, 3, 4, перші дві букви однакові, останні дві букви різні і не такі, як перші?

Розв'язок: а) В загальному випадку, якщо немає обмежень щодо наборів цифр та букв, для формування номеру треба незалежним чином чотири рази обрати одну з 17 цифр та чотири рази – одну з 10 цифр. Згідно правилу добутку загальна кількість способів, якими це можна зробити, дорівнює

$$n_1 = 17^4 \cdot 10^4.$$

б) В цьому випадку для формування номеру треба обрати впорядковану підмножину з чотирьох букв (з множини, що містить 17 елементів-літер) та впорядковану підмножину з чотирьох цифр (з 10 можливих). Для цього двічі скористаємось формулою 4) для кількості розміщень (кількості впорядкованих підмножин). Оскільки вибір букв і вибір цифр для номерного знаку – дві незалежні події, які проводяться одночасно, то згідно правилу добутку треба перемножити можливі кількості способів:

$$n_2 = A_{17}^4 \cdot A_{10}^4.$$

в) В цьому випадку ми маємо вибрати лише одну з десяти цифр, яка і буде повторюватись в номері автомобіля (кількість способів – $C_{10}^1 = 10$), букви обираємо, як і в п. б), загальна кількість способів є добутком способів вибору цифри і букв:

$$n_3 = 10 \cdot A_{17}^4.$$

г) Тут треба обрати одну цифру і одну букву. Маємо

$$n_4 = 10 \cdot 17 = 170.$$

д) Для цифрової частини обираємо впорядковану підмножину (A_{10}^4 способами), кожну з чотирьох букв обираємо одним з 17 способів:

$$n_5 = A_{10}^4 \cdot 17^4.$$

е) Кількість можливих варіантів цифрової частини дорівнює кількості можливих перестановок чотирьох цифр (4!), для першої групи букв можна обрати літеру 17 способами, для другої групи треба з 16 можливих букв (окрім тої, яка була використана в першій частині номеру) вибрати впорядковану підмножину, що містить дві літери, це можна зробити A_{16}^2 способами. Згідно правилу добутку всі кількості способів сформувати різні групи перемножаємо. Остаточно маємо

$$n_6 = 4! \cdot 17 \cdot A_{16}^2.$$

Приклад 2.2

Сергійко лише почав вивчати теорію ймовірностей і йому здається, що ймовірність зустріти динозавра має дорівнювати $\frac{1}{2}$, бо є всього два можливі варіанти – ми або зустрінемо динозавра (ω_1), або не зустрінемо (ω_2), тож $\Omega = \{\omega_1, \omega_2\}$, а оскільки нас цікавить подія $A = \{\omega_1\}$, то за формулою (2.2) її ймовірність складе

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

Чи не помилився Сергійко?

Вказівка: якщо уважно прочитати класичне визначення ймовірності, то легко зрозуміти, в чому саме помилився Сергійко :)

Приклад 2.3

Нехай експеримент полягає в підкиданні двох симетричних монет, тоді простір елементарних подій буде складатись з чотирьох рівноможливих результатів:

$$\Omega = \{\text{ГГ, ГР, РГ, РР}\}.$$

В цьому випадку ми потрапляємо в умови класичної ймовірнісної схеми і можемо отримати ймовірності подій за формулою (2.2). Наприклад, для події

$$B = \{\text{Випав хоча б один герб}\} = \{\text{ГГ, ГР, РГ}\}$$

маємо $\mathrm{P}(B) = 3/4$.

В історії відомий парадокс, який впирається в некоректне визначення простору елементарних подій цього стохастичного експерименту:

$$\Omega = \{\text{ГГ, ГР або РГ, РР}\},$$

тобто подія $A = \{\text{ГР, РГ}\}$ трактувалася як елементарна і, слідуючи "аксіомі симетрії", стверджувалось, що $\mathrm{P}\{\text{ГГ}\} = \mathrm{P}\{\text{РР}\} = \mathrm{P}(A) = 1/3$. (Помилка Д'Аламбера)

Оскільки результати дослідів протирічили такій ймовірнісній моделі, (спостереження давали $\mathrm{P}\{\text{ГГ}\} = \mathrm{P}\{\text{РР}\} = 1/4,\ \mathrm{P}(A) = 1/2$), то описаний феномен був оголошений парадоксом, над вирішенням якого билися багато відомих науковців. Прояснилося все пізніше, після чіткого формулювання *незалежності подій*.

Насправді, якщо є бажання не враховувати порядок і обмежитись трьома елементарними подіями, оголосивши елементарні події $\{\text{ГР}\}$ та $\{\text{РГ}\}$ одним результатом, то в цьому випадку класична ймовірнісна схема не має місця, отже ймовірності подій треба шукати за більш загальною формулою (2.1).

Приклад 2.4

Знайти ймовірність того, що при підкиданні двох симетричних гральних кубиків
а) сума очок, що випала на двох кубиках, буде не менша 10 (подія A);
б) добуток буде не більший за 3 (подія B);
в) випаде парна кількість очок (подія C).

Розв'язок: Простір елементарних подій Ω даного стохастичного експерименту складається з 36 рівноможливих елементарних подій, кожна з яких є парою чисел від 1 до 6:

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

Подія A трапиться, якщо випадуть такі пари: $\{(4,6),(6,4),(5,5),(5,6),(6,5),(6,6)\}$. Згідно (2.2) маємо $\mathrm{P}(A) = 6/36 = 1/6$.

Аналогічно подія $B = \{(1,1),(1,2),(1,3),(2,1),(3,1)\}$, тобто її появі сприяють 5 рівноможливих елементарних подій, отже $\mathrm{P}(B) = 5/36$.

Для того, щоб трапилась подія C, треба, щоб обидві цифри, що випали, були або парні, або непарні. Отже, якщо на перше місце може бути обрана будь-яка з 6 цифр, то на друге – лише одна з трьох цифр тої ж парності, що й перша цифра. Відповідно кількість пар, які забезпечують появу події C дорівнює $m = 6 \cdot 3 = 18$, а ймовірність $\mathrm{P}(C) = 18/36 = 1/2$.

Приклад 2.5

Симетричний гральний кубик підкидають 6 разів. Яка ймовірність того, що випадуть всі 6 граней?

Розв'язок: Елементарною подією цього експерименту є набір з шести чисел, значення кожного з яких від 1 до 6. Загальна кількість таких наборів (кількість елементів простору елементарних подій) 6^6. Тепер треба порахувати кількість таких наборів, в яких присутні всі 6 чисел. Це – кількість перестановок з 6 елементів, яка дорівнює $6!$. Відповідно шукана ймовірність $\mathrm{P}(A) = 6!/6^6$.

Приклад 2.6

Серед 10 угод страхування авто 5 укладено на випадок крадіжки, 3 – на випадок ушкоджень з вини автовласника та 2 – на випадок ушкоджень не з вини автовласника. Всі страхові випадки рівноможливі. Знайти ймовірність того, що з трьох страхових випадків:

а) будуть рівно два, що мають одну і ту саму причину;
б) всі випадки будуть мати різні причини.

Розв'язок: нехай A та B означають події: A = {з трьох страхових випадків будуть рівно два, що мають одну і ту саму причину}, B = {всі випадки мають різні причини}. Тоді

$$\mathrm{P}(A) = \frac{C_5^2 C_5^1 + C_3^2 C_7^1 + C_2^2 C_8^1}{C_{10}^3} = \frac{79}{120},$$

$$\mathrm{P}(B) = \frac{5 \cdot 3 \cdot 2}{C_{10}^3} = \frac{1}{4}.$$

Приклад 2.7

Із 7 секретних файлів агент ФБР Фокс Малдер закодував 3. Наступного дня агент ФБР Дана Скалі взяла 5 секретних файлів. Знайти ймовірність того, що 2 взятих файли виявляться закодованими.

Розв'язок: Ймовірність шуканої події A = {2 взятих файли виявляться закодованим} складе

$$\mathrm{P}(A) = \frac{C_3^2 C_4^3}{C_7^5} = \frac{4}{7},$$

бо всього існує C_7^5 способів обрати 5 файлів із 7, серед яких є саме $C_3^2 C_4^3$ варіантів, коли 2 файли будуть закодовані (бо всього закодованих 3) і 3 = 5 - 2 файли будуть незакодовані (бо незакодованих файлів, очевидно, 7 - 3 = 4).

Відповідь: 4/7.

Приклад 2.8

При грі в бридж повна колода (52 карти) роздається чотирьом гравцям по 13 карт кожному. Яка ймовірність того, що
а) один з гравців отримає всі чотири туза;
б) один з гравців отримає всі чотири туза, а інший – всі чотири короля?

Розв'язок: Елементарними подіями стохастичного експерименту, який полягає у відповідній роздачі гральних карт, є всі можливі набори з 13 карт у чотирьох гравців.

Причому порядок гравців є суттєвим, порядок карт на руках у кожного окремого гравця – ні. Отже, загальна кількість всіх таких наборів (загальна кількість елементарних подій цього експерименту) можна обчислити або користуючись формулою перестановок з повтореннями, або формулами комбінацій:

$$n = |\Omega| = \frac{52!}{13!13!13!13!} = C_{52}^{13} C_{39}^{13} C_{26}^{13} C_{13}^{13} = \frac{52!}{(13!)^4}.$$

Для того, щоб знайти потрібні ймовірності, знайдемо, скільки елементарних подій забезпечують появу події у випадках а), б) (будемо позначати їх A і B відповідно).

а) Гравець, у якого мають бути всі тузи, може бути обраний 4 способами. Окрім тузів у нього є ще 9 карт. Отже, кількість способів розподілити колоду без тузів можна підрахувати знову за допомогою формули перестановок з повтореннями, а кількість роздач, що відповідають події A

$$m_A = |A| = 4 \cdot \frac{48!}{9!(13!)^3},$$

а ймовірність цієї події

$$P(A) = \frac{m_A}{n} = \frac{4 \cdot \frac{48!}{9!(13!)^3}}{\frac{52!}{(13!)^4}} \approx 0.01.$$

Якщо йдеться про випадіння всіх тузів конкретному гравцю, ця ймовірність буде в 4 рази менша.

б) Гравці, які мають королів та тузів, можуть бути обрані A_4^2 способами, оскільки в даному випадку порядок має значення. Розподілити карти вказаним чином можна

$$m_B = |B| = A_4^2 \cdot \frac{44!}{(9!)^2(13!)^2},$$

а ймовірність цієї події

$$P(B) = \frac{m_B}{n} = \frac{A_4^2 \cdot \frac{44!}{(9!)^2(13!)^2}}{\frac{52!}{(13!)^4}} \approx 0.00012.$$

Приклад 2.9

Після прання з 10 різних пар шкарпеток зникло 6 шкарпеток.
а) Яка ймовірність оптимістичного сценарію: лишилося 7 повних пар шкарпеток?
б) Яка ймовірність найгіршого випадку: лишилося лише 4 пари шкарпеток?

Розв'язок: В результаті стохастичного експерименту маємо C_{20}^{6} можливих способів вибрати 6 шкарпеток з 20.
а) З того, що лишилося в наявності, можна зібрати 7 повних пар, якщо втрачено обидві шкарпетки з трьох пар. Ці три пари можна обрати C_{10}^{3} способами. Отже, ймовірність оптимістичного сценарію

$$\mathrm{P}(A) = \frac{C_{10}^{3}}{C_{20}^{6}} \approx 0.0031.$$

б) З того, що лишилося, можна обрати лише чотири повних пари, якщо загублено по одній шкарпетці з різних шістьох пар. Ці 6 пар можна обрати C_{10}^{6} способами, причому є по дві можливі шкарпетки в кожній парі, які є можливість згубити. Отже, є $C_{10}^{6} \cdot 2^6$ способів вибрати з 20 шкарпеток таким чином, щоб лишилося лише 4 повних пари. Ймовірність цього

$$\mathrm{P}(B) = \frac{C_{10}^{6} \cdot 2^6}{C_{20}^{6}} \approx 0.3467.$$

Зауважимо, що найгірший сценарій більш, ніж в 100 разів ймовірніший за найкращий.

Приклад 2.10

В ряд з $2N$ місць довільним чином розсаджуються N хлопців та N дівчат. Яка ймовірність того, що ніякі дівчата не сидітимуть поруч?

Розв'язок: Загальна кількість способів розсадити $2N$ осіб дорівнює кількості їх перестановок: $n = (2N)!$

Знайдемо кількість перестановок, в яких ніякі дві дівчини не сидять поруч. Це означає, що між усіма дівчатами має сидіти хоча б один хлопець. Мінімальна необхідна для цього кількість хлопців – $(N-1)$. N-тий хлопець може сидіти або поруч з іншим хлопцем (таких можливостей $N-1$), або з лівого чи з правого краю. Отже, загальна кількість послідовностей з дівчат та хлопців, де дівчата не сидять поруч – $N+1$. В кожній з цих послідовностей дівчата можуть мінятися місцями, а отже розміщатися $N!$ способами, хлопці також можуть в кожній з послідовностей розміщатися $N!$ способами. Відповідно, загальна кількість елементарних подій, які задовольняють нашу умову, $(N+1)(N!)^2$. Шукана ймовірність

$$\mathrm{P}(A) = \frac{(N+1)(N!)^2}{(2N)!}.$$

Задача 2.1. В групі 25 студентів. Скількома способами можна обрати серед них старосту та двох заступників?

Задача 2.2. В понеділок 8 уроків. Скількома способами можна скласти розклад з різних уроків на цей день, якщо викладається 12 предметів? Скількома способами можна скласти розклад, щоб кожен предмет повторювався двічі?

Задача 2.3. Наявні тканини 6 кольорів. Скількома способами можна зшити прапор

а) з двох горизонтальних смуг різного кольору?

б) з трьох горизонтальних смуг різного кольору?

в) з трьох смуг, колір яких може повторюватися?

Задача 2.4. В конкурсі беруть участь 20 студентських робіт. Скількома способами можуть бути розподілені три премії, якщо розмір премії однаковий? Якщо премії відрізняються?

Задача 2.5. 12-томне видання творів Джека Лондона розставляють на полиці у випадковому порядку. Скількома способами можна це зробити? Скількома способами можна це зробити таким чином, щоб томи 1-4 стояли поруч в порядку зростання? поруч в порядку зростання чи спадання?

Задача 2.6. В програмі курсу 25 тем. Студент знає 20 з них. Скількома способами може бути сформований екзаменаційний білет з п'яти питань, в якому студент знатиме принаймні 4?

Задача 2.7. При грі в кості для виграшу треба, щоб при підкиданні трьох кубиків принаймні двічі випало 5 або 6 очків. Скількома способами це може трапитись? Скільки всього результатів можливі при підкиданні трьох симетричних гральних кубиків?

Задача 2.8. По дорозі до університету професор може на першій ділянці шляху сісти в одну з трьох маршруток, а потім пересісти на одну з шести маршруток. Скількома різними способами він може доїхати до університету?

Задача 2.9. Скільки тризначних чисел можна написати за допомогою цифр 0, 1, 2, 3, 4, 5? Скільки чисел серед них буде таких, що цифри не повторюються? Якщо число має бути кратним 5?

Задача 2.10. В номері автомобіля спочатку записані дві букви, потім 4 цифри, потім ще дві букви. Скільки різних номерних знаків може бути сформовано в такий спосіб, якщо використовувати 17 літер українського алфавіту?

Задача 2.11. В шаховому турнірі беруть участь n учасників. Скільки партій буде зіграно, якщо кожні 2 учасника мають зустрітися по одному разу?

Задача 2.12. Скількома способами можна розмістити на шахівниці розміром $n \times n$ m тур таким чином, щоб вони не били одна одну ($m \leq n$)?

Задача 2.13. В пиріжковій продається 7 видів солодких пиріжків та 6 видів солоних. Скількома способами можна вибрати 5 пиріжків? Скількома спобобами можна вибрати пиріжки так, щоб серед них було два солоних і три солодких?

Задача 2.14. Скільки кісток доміно можна утворити, використовуючи числа $1, 2, ..., n$?

Задача 2.15. В групі навчається 25 студентів. Скількома способами можна обрати серед них 4 куратора молодшого курсу? 2 офіційних куратора і 4 тіньових куратора молодшого курсу? 4 куратора для чотирьох різних груп молодшого курсу?

Задача 2.16. В ліфт готелю на нульовому поверсі заходить 10 пасажирів. Скількома способами вони можуть вийти з ліфту, який зупинятиметься на 15 поверхах? Скільки серед них способів, коли всі пасажири виходять на різних поверхах?

Задача 2.17. На курсі 10 груп по 25 студентів в кожній. Скількома способами можна обрати по 2 представника з кожної групи в студентський парламент?

Задача 2.18. Якщо розвернути лист паперу на 180°, то цифри 0, 1, 8 не змінюються, а 6 та 9 переходять одна в одну. Скільки існує семизначних чисел, величина яких не зміниться при розвороті листа паперу на 180°?

Задача 2.19. В чемпіонаті країни по футболу в вищій лізі беруть участь 18 команд. Команди, які зайняли перші три місця, отримують золоту, срібну та бронзову медалі. Команди, які опинилися на останніх двох місцях, залишають вищу лігу. Скільки може бути різних результатів чемпіонату?

Задача 2.20. Скільки діагоналей можна провести в опуклому n-кутнику?

Задача 2.21. Розглянемо прямокутну сітку квадратів ("шахове місто"), яка складається з $m \times n$ квадратів, розділених $n-1$ горизонтальними та $m-1$ вертикальними "вулицями". Скільки різних найкоротших шляхів існує з лівої нижньої точки (з координатами $(0,0)$) в праву верхню точку (з координатами (m,n))?

Задача 2.22. Скільки підмножин має множина, що містить n елементів (порожня множина також є підмножиною будь-якої множини)?

Задача 2.23. В кімнаті є 10 джерел світла. Скількома способами можна освітити кімнату?

Задача 2.24. Скількома способами можна впорядкувати множину $\{1, 2, ..., 2n\}$ таким чином, щоб кожне парне число мало парний номер?

Задача 2.25. Скількома способами можна впорядкувати множину $\{1, 2, ..., 2n\}$ таким чином, щоб кожне парне число та кожне число, кратне трьом, мало номер, кратний 2 та 3 відповідно?

Задача 2.26. Скільки є перестановок з n елементів, в яких задані два елементи не стоять поруч?

Задача 2.27. Студенту треба протягом 12 днів здати 4 іспити. Скількома способами можна це зробити, якщо в один день не може бути більше одного іспиту?

Задача 2.28. Студенту протягом 14 днів треба здати 4 іспити та 6 заліків. Скількома способами це можна зробити, якщо в один день може бути або один іспит, або будь-яка кількість заліків?

Задача 2.29. Скільки існує перестановок з n елементів, в яких між двома певними елементами стоять рівно r елементів?

Задача 2.30. На семінар приготували доповіді 4 студенти: A, B, C і D. Скількома способами можна впорядкувати їхні доповіді, якщо доповідь студента D не може бути раніше за доповідь студента A?

Задача 2.31. Скільки різних слів можна утворити перестановкою букв в слові "математика"?

Задача 2.32. Скількома способами можна розділити 100 студентів на 4 групи по 25 студентів в кожній?

Задача 2.33. Скількома способами можна розділити 3 варіанти контрольної роботи між 21 студентом так, щоб 6 студентів писало варіант 1, 7 – варіант 2, 8 – варіант 3? Так, щоб кожен варіант писало по 7 студентів?

Задача 2.34. n однакових куль розкладають навмання в N урнах. Довести, що
а) кількість різних способів розміщення дорівнює $C_{N+n-1}^{n} = C_{N+n-1}^{N-1}$;
б) кількість розміщень, коли в кожній урні є принаймні одна куля, дорівнює C_{n-1}^{N-1}.

Задача 2.35. Скількома способами можна розподілити n однакових подарунків серед N дітей? Скільки серед них способів, коли кожна дитина отримає принаймні один подарунок?

Задача 2.36. 12-томне видання творів Джека Лондона поставили на полку у випадковому порядку. Яка ймовірність того, що томи будуть стояти або в порядку зростання, або в порядку спадання?

Задача 2.37. На 10 картках написано літери А, А, А, Е, К, М, М, Т, Т, И. Знайти ймовірність того, що при випадковому розміщенні цих карток отримаємо слово МАТЕМАТИКА.

Задача 2.38. Дано 5 відрізків довжиною 1, 3, 5, 7 та 9 см. Яка ймовірність того, що випадково обрані три відрізки утворять трикутник?

Задача 2.39. Шість чоловіків пробують відгадати невідомий їм результат підкидання грального кубика. Яка ймовірність того, що це вдасться зробити хоча б одному з них?

Задача 2.40. Чотиритомну збірку творів розміщують на полиці навмання. Обчислити ймовірність того, що томи будуть стояти в порядку зростання.

Задача 2.41. Якщо у цьому збірнику задач обрати навмання одну зі сторінок, на яких є номер сторінки, то яка ймовірність того, що

а) номер сторінки буде меншим за 55?

б) номер сторінки буде містити цифру 0?

в) усі цифри номеру сторінки будуть парні?

г) усі цифри номеру сторінки будуть різні?

Задача 2.42. Якщо у цьому запитані навманя обрати одне слово, то яка ймовірність того, що слово написано з помилкою?

Задача 2.43. Обчислити ймовірність того, що чотиризначний номер навмання обраного у великому місті автомобіля:

а) складається з різних цифр;

б) має тільки дві однакові цифри;

в) має дві пари однакових цифр;

г) має тільки три однакові цифри;

д) складається з однакових цифр.

Задача 2.44. У записаному телефонному номері 259-0... останні три цифри стерлись. Знайти ймовірності наступних подій: A = {стерлись різні цифри, відмінні від 0 та 1}, B = {стерлись однакові цифри}.

Задача 2.45. З множини $\{1, 2, ..., N\}$ випадковим чином вибирають k чисел. Знайти ймовірності таких подій: а) всі вибрані значення кратні q; б) рівно два числа кратні q; в) хоча б одне з вибраних ділиться націло на q.

Задача 2.46. Знайти ймовірність p_N того, що навмання вибране число з множини {1, 2, ..., N} кратне q ($q < N$). Обчислити $\lim\limits_{n\to\infty} p_N$.

Задача 2.47. Підкидають n гральних кубиків. Яка ймовірність того, що випадуть n_1 одиниць, n_2 двійок, ... ,n_6 шісток ($n_1 + n_2 + \ldots + n_6 = n$)?

Задача 2.48. У ящику лежать червоні та чорні шкарпетки. Якщо із ящика навмання вийняти дві шкарпетки, то ймовірність того, що обидві з них червоні, дорівнює $1/2$. Яке найменше можливе число шкарпеток у ящику?

Задача 2.49. Марійка розкладає число $a = 10^{2023}$ на два множника $a = b \cdot c$, обираючи можливі множники навмання. Яка ймовірність того, що жодне з чисел b, c не закінчується на 0?

Задача 2.50. У групі є n студентів. Яка ймовірність того, що принаймні у двох з них збігаються дні народження? Розв'язати задачу в загальному вигляді, та, зокрема, для $n = 23$.

Задача 2.51. Знайти ймовірність того, що:

а) дні народження 12 чоловік припадають на 12 різних місяців року;

б) дні народження 6 чоловік припадуть точно на 2 місяці року.

Задача 2.52. Обчислити ймовірність того, що для даних тридцяти осіб з 12 місяців року на 6 місяців припадає по два дні народження, а на 6 - по 3 дні народження.

Задача 2.53. У ліфті знаходиться сім пасажирів. Ліфт зупиняється на десяти поверхах. Яка ймовірність того, що жодні два пасажири не вийдуть на одному поверсі?

Задача 2.54. Яка ймовірність того, що при підкиданні 5 симетричних монет випадуть і герби, і решки?

Задача 2.55. У три вагони заходять дев'ять пасажирів. Яка ймовірність того, що:

а) у перший вагон зайде три пасажири?

б) у кожен вагон зайде по три пасажири?

в) в один з вагонів зайде чотири, в другий – три і в третій – два пасажири?

Задача 2.56. Знайти ймовірність того, що при підкиданні шести гральних кубиків одиниця з'явиться хоча б один раз.

Задача 2.57. Гравець A кидає шість гральних кубиків і виграє, якщо випадає хоча б одна одиниця. Гравець B кидає 12 гральних кубиків і виграє, якщо випаде хоча б дві одиниці. У кого більша ймовірність виграти?

Задача 2.58. Що більш ймовірно: отримати хоча б одну одиницю при підкиданні чотирьох гральних кубиків, чи хоча б одну пару одиниць при 24 кидках двох гральних кубиків?

Задача 2.59. Кидають 12 гральних кубиків. Яка ймовірність того, що кожне число з'явиться 2 рази?

Задача 2.60. Серед n осіб виділено особу A і особу B. Усі n осіб шикуються у шеренгу в будь-якому порядку. Яка ймовірність того, що між A і B буде рівно r осіб?

Задача 2.61. Десять чоловіків і десять жінок випадковим чином поділено на пари. Знайти ймовірність того, що кожна пара складається з осіб протилежної статі.

Задача 2.62. У n урнах випадковим чином розміщено n куль. Знайти ймовірність того, що кожна урна буде зайнята.

Задача 2.63. Кожен з 50 штатів представлено двома сенаторами. Розглянемо події, які полягають у тому, що у комітеті з 50 випадково обраних сенаторів:

а) представлено фіксований штат; б) представлені усі штати.

Підрахувати ймовірності подій а) і б).

Задача 2.64.* а) Знайти ймовірність того, що при випадковому розміщенні n куль по n урнах рівно одна урна залишиться порожня.
б) Дати відповідь в загальному випадку та при $n = 5$ та $n = 3$.
в) Знайти помилку в наступному розв'язку цієї задачі при $n = 3$: з усіх 10 варіантів (0, 1, 2), (0, 2, 1), (1, 0, 2), (2, 0, 1), (1, 2, 0), (2, 1, 0), (0, 0, 3), (0, 3, 0), (3, 0, 0), (1, 1, 1) нам підходять лише шість перших, тому шукана ймовірність складе $6/10 = 0.6$.

Задача 2.65. Випадковим чином з чисел 1, 2, ..., n обрано два різні числа (порядок їх обирання значення не має, тобто обрану пару (2, 5) та (5, 2) вважаємо одним і тим же елементарним наслідком). Яка ймовірність того, що одне з них буде строго менше, а друге строго більше від заданого числа $k \in \{1, 2, \dots, n\}$? Дати відповідь, зокрема, при $n = 3$, $k = 2$.

Задача 2.66.* В черзі до каси стоїть $2n$ осіб, з яких n осіб має лише банкноти по 10 гривень, а решта - по 5 гривень. На початку грошей у касі немає. Кожна особа повинна купити один білет ціною у 5 гривень. Яка ймовірність того, що жодна особа не буде чекати на решту?

Задача 2.67. Випадковим чином розміщено r куль по n урнах. Знайти ймовірність того, що фіксована урна буде містити рівно k куль.

Задача 2.68.* В урні є m білих та n чорних куль ($m > n$). Кожного разу з урни випадково виймають по одній кулі. Яка ймовірність того, що в деякий момент число вийнятих білих і чорних куль буде однаковим?

Задача 2.69. З колоди карт у 32 карти два рази виймають по одній карті без повернення. Яка ймовірність того, що:
а) лише одна з цих карт буде валет? б) принаймні одна з цих карт буде валет?

Задача 2.70. З колоди у N карт ($N = 32$, 36, 52) гравець отримує п'ять карт. Знайти ймовірність того, що цей гравець буде мати п'ять карт різних значень.

Задача 2.71. При грі у бридж між чотирма гравцями роздається колода у 52 карти. Знайти ймовірність того, що фіксований гравець буде мати тринадцять карт різних значень.

Задача 2.72. При грі у бридж між чотирма гравцями роздається колода у 52 карти. Знайти ймовірність того, що фіксований гравець буде мати тринадцять карт однієї масті.

Задача 2.73. При грі у бридж між чотирма гравцями роздається колода у 52 карти. Знайти ймовірність того, що у кожного з гравців будуть всі найменування карт від двійки до туза.

Задача 2.74. Знайти ймовірність того, що при роздаванні 52 карт при грі у бридж (коли гравці отримують по 13 карт) кожний гравець отримає туза.

Задача 2.75. Знайти ймовірність того, що у деякого гравця у бридж буде 5 пік, 4 черви, 3 бубни і 1 хрест.

Задача 2.76. Знайти ймовірність того, що при роздаванні карт у бридж (52 карти на 4 гравців) фіксований гравець отримає рівно k тузів.

Задача 2.77. Знайти ймовірність того, що при роздаванні карт у бридж гравець A отримає m, а гравець B - n пік.

Задача 2.78. Чому дорівнює ймовірність того, що при роздаванні карт у бридж у гравців A і B разом буде рівно k тузів?

Задача 2.79. Нехай a, b, c, d - чотири невід'ємних цілих числа такі, що $a + b + c + d = 13$. Знайти ймовірність того, що фіксований гравець у бридж при роздаванні карт буде мати a пік, b черв, c бубен і d хрест.

Задача 2.80.* Нехай a, b, c, d - чотири невід'ємних цілих числа такі, що $a + b + c + d = 13$. Знайти ймовірність того, що при роздаванні карт у бридж гравці A, B, C, D мають a, b, c, d пік відповідно.

Задача 2.81. Знайти ймовірність того, що серед $r = 16$ карт, випадково обраних з колоди у $N = 52$ листа, буде рівно $k = 1$ туз.

Задача 2.82. Знайти ймовірність того, що кожен з двох наборів карт має рівно k тузів, якщо кожен набір складається з r карт і вибір провадиться:

а) з однієї колоди для гри у бридж;

б) з двох таких колод.

Задача 2.83. Колода карт для гри у бридж (52 карти) ретельно перемішується і далі послідовно зверху беруть по одній карті. Знайти ймовірності $p_1(r), p_2(r), p_3(r), p_4(r)$ того, що, відповідно, перший, другий, третій і четвертий туз з'явиться на r-ому кроці.

Задача 2.84. Знайти імовірність того, що при випадковому розташуванні 52-ох карт для гри у бридж ніякі два тузи не будуть знаходитись поруч.

Задача 2.85. Кожна з n палиць ламається на дві частини – довгу і коротку. Далі ці $2n$ уламків з'єднуються у n пар, кожна з яких утворює нову палицю. Знайти імовірність того, що:

а) частини будуть об'єднані у початковому порядку;

б) усі довгі частини будуть об'єднані з короткими.

Задача 2.86. На полиці стоять 20 книжок. З них випадковим чином вибирається 7. Яка ймовірність, що хоча б якісь з вибраних книжок стояли поруч?

Задача 2.87. Статистика Максвелла-Больцмана. Кожна з n різних частинок потрапляє в один з N лічильників. Знайти ймовірність того, що

а) перший, другий,..., N-й лічильник зареєструє відповідно $n_1, n_2, ..., n_N$ частинок, якщо $n_1 + n_2 + ... + n_N = n$;

б) даний лічильник зареєструє k частинок.

Задача 2.88. Статистика Бозе-Ейнштейна. Кожна з n однакових частинок потрапляє в один з N лічильників. Знайти ймовірність того, що перший, другий,$\ldots$, N-й лічильник зареєструє відповідно n_1, $n_2, ..., n_N$ частинок ($n_1 + n_2 + ... + n_N = n$), якщо рівноможливими вважаються розміщення, які відрізняються кількістю частинок, зареєстрованих лічильниками. Довести, що ймовірність того, що даний лічильник зареєструє k частинок, дорівнює

$$p_k = \frac{C_{N+n-k-2}^{n-k}}{C_{N+n-1}^{n}}.$$

Довести, що при $N > 2$ послідовність $\{p_k,\ k \geq 0\}$ спадна. Знайти ймовірність того, що рівно m лічильників не зареєструє жодної частинки.

Задача 2.89. Статистика Фермі-Дірака. Кожна з n однакових частинок реєструється одним з N лічильників ($n < N$). Яка ймовірність того, що перший, другий,$\ldots$, N-й лічильник зареєструє відповідно n_1, $n_2, ..., n_N$ частинок ($n_1 + n_2 + ... + n_N = n$), якщо рівноможливими вважаються розміщення, які задовольняють «принцип Паулі» (кожен лічильник реєструє не більше однієї частинки).

Задача 2.90. Група з $2N$ хлопців і $2N$ дівчат розділена на дві рівні частини. Знайти ймовірність того, що кожна частина буде мати однакове число хлопців і дівчат.

Задача 2.91. У продаж надійшло n лотерейних білетів, серед яких m виграшних. Яка ймовірність хоча б одного виграшу при купівлі k білетів?

Задача 2.92. Учасник лотереї «Спортлото» 6 із 49 заповнив одну картку. Яка ймовірність того, що він вгадає чотири номери?

Задача 2.93. Учасник лотереї «Спортлото» 6 з 49 заповнив дві картки таким чином, що усі закреслені ним номери на обох картках різні. Знайти ймовірність того, що він не вгадав жодного номера.

Задача 2.94. Дві картки «Спортлото» 6 із 49 заповнені наступним чином: $(4.12, 38.20, 41.46)$, $(4.12, 38.20, 41.49)$. Яка ймовірність того, що на кожній карточці буде вгадано рівно три номери?

Задача 2.95.* У комірчині лежить n пар чобіт. З них випадковим чином обирають $2r$ чобіт ($2r < n$). Знайти ймовірність того, що серед обраних чобіт:

а) немає жодної пари;

б) є рівно одна пара.

Задача 2.96. Вибирається навмання один член визначника n-го порядку. Яка ймовірність того, що він не містить елементів головної діагоналі?

Задача 2.97.* Написано n листів, але адреси на конвертах написані навмання. Яка ймовірність того, що:

а) принаймні один з адресатів одержить призначений для нього лист?

б) У пункті а) підрахувати граничне значення ймовірності при $n \to \infty$.

в) m адресатів одержать призначені для них листи?

Задача 2.98. Маємо 6 олівців різного кольору і 6 футлярів до них тих самих кольорів. Олівці навмання розкладають по футлярах. Знайти імовірність того, що жоден із них не буде у своєму футлярі.

Задача 2.99. Група з 24 студентів, серед яких 5 відмінників, довільно розбивається порівну на дві підгрупи. Знайти ймовірність того, що три відмінники будуть у першій підгрупі (подія A).

Задача 2.100. Фокусник пропонує трьом глядачам задумати число від 1 до 10. Вважаючи вибір числа кожним з учасників будь-якого числа рівноймовірним та незалежним від інших учасників, знайти ймовірність того, що у когось з них задумані числа співпадуть.

Задача 2.101. Чи однаковими є ймовірності того, що випадково обрана особа народилася в неділю та того, що дві випадково обрані особи народилися в один день тижня?

Задача 2.102. Набираючи номер телефона, абонент забув дві останні цифри. Проте він пам'ятає, що одна з них – непарна, а одна – нуль. Знайти ймовірність того, що він набере правильний номер.

Задача 2.103. Учасник лотереї «Спортлото» закреслює 6 чисел з 49. Повний виграш отримує той, хто вгадає всі 6 цифр. Виграш отримають також ті, хто вгадав принаймні три числа. Знайти ймовірність

а) повного виграшу в «Спортлото»;

б) вгадати 5, 4 та 3 числа;

в) отримати якийсь виграш;

г) не вгадати жодного числа.

Задача 2.104. Навмання обирається чотиризначне число. Яка ймовірність того, що

а) число є симетричним (наприклад, 2552);

б) число кратне 5;

в) число складене лише з парних цифр;

г) число складене лише з непарних цифр.

Задача 2.105. Десять студентів домовились їхати разом одним поїздом, в якому 10 вагонів, проте не домовились про вагон. Вважаючи вибір вагона студентами випадковим, яка ймовірність того, що всі вони будуть в різних вагонах.

Задача 2.106. Два гравця, рівних по силі гри, грають в певну гру (нічиєї бути не може). Той, хто першим виграє 4 партії, отримає 1 млн грн. Гра була перервана, коли рахунок був 2:1. Продовжити гру неможливо. Як розділити гроші між гравцями?

Задача 2.107. В лотереї розігруються 500 автомобілів серед 2 млн учасників за номером реєстрації. Виграшні номери визначаються комп'ютерною програмою, причому ця програма не виключає номер, який вже виграв, зі списку при подальших розіграшах. Яка ймовірність того, що хтось виграє два або більше автомобіля?

Задача 2.108. Рядок, в якому записано лише дужки, що відкриваються чи закриваються, назвемо *правильною дужковою послідовністю* (ПДП), якщо виконуються такі три умови:

- порожній рядок є ПДП;
- якщо ПДП взяти в дужки, то вона залишиться ПДП;
- якщо зліва чи справа до ПДП дописати якусь ПДП, то також отримаємо ПДП.

Приклад ПДП: (( ) ( )) ( )( )

Петрик записує рядок із n символів, кожен з яких обирає навмання – або ліва дужка (відкривається), або права дужка (закривається). Яка ймовірність того, що в результаті Петрик отримає ПДП, якщо

а) $n = 4$; б) $n = 5$; в) $n = 2024$?

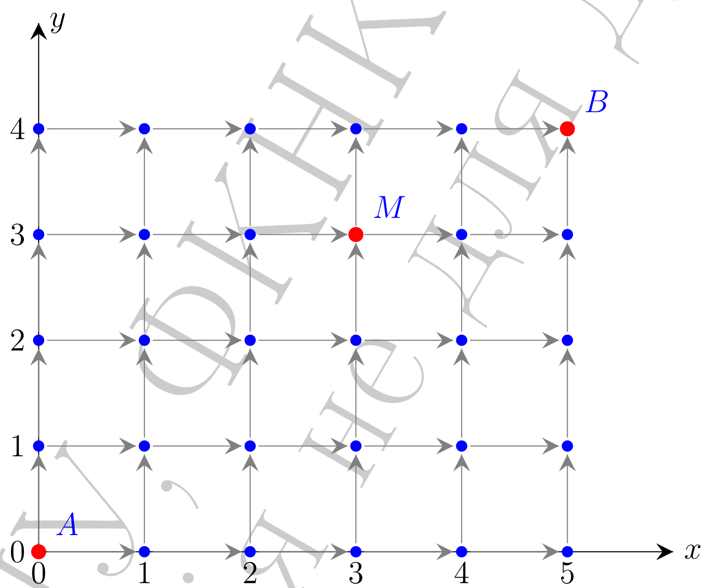


Рис. 2.1: до задачі 2.109

Задача 2.109. Рух на множині точок $\{(i,j)|\ i = 0\ldots 5,\ j = 0\ldots 4\}$ (див. 2.1) починається у точці $A(0,0)$ і у кожній точці рух здійснюється навмання у одному з двох дозволених напрямків (вгору або праворуч, – наприклад з точки $(0,0)$ перейти можна або в точку $(0,1)$, або в точку $(1,0)$) або лише в одному з них (якщо вибору немає, – наприклад з точки $(0,4)$ перейти можна лише в точку $(1,4)$). Врешті маршрут закінчиться в точці $B(5,4)$. Далі під маршрутом будемо мати на увазі лише маршрут з $A(0,0)$ до $B(5,4)$.

а) Яка довжина кожного маршруту?

б) Скільки усього є різних маршрутів?

в) Скільки усього є різних маршрутів, які проходять через точку $M(3,3)$?

г) Яка ймовірність, що маршрут пройде через точку $M(3,3)$?

3 ЗЛІЧЕННА ЙМОВІРНІСНА СХЕМА

Побудуємо ймовірнісну модель для стохастичного експерименту із зліченною множиною елементарних подій $\Omega = \{\omega_1, \omega_2, ..., \omega_n, ...\}$. Як і у випадку скінченного Ω, **подією** будемо називати довільну підмножину з Ω, і нехай $\Re$ - множина усіх подій. Для будь-якої події $A = \{\omega_{i_1}, \omega_{i_2}, ..., \omega_{i_m}, ...\} \subseteq \Omega$ ймовірність $\mathrm{P}(A)$ визначимо наступним чином. Візьмемо послідовність чисел $p_1, p_2, ..., p_n, ...$ таку, що $p_i \geq 0$, $\sum\limits_{i=1}^{\infty} p_i = 1$. Тоді за визначенням

$$\mathrm{P}(A) = \sum_{m=1}^{\infty} p_{i_m}.$$

Трійку $(\Omega, \Re, \mathrm{P})$ будемо називати **зліченною ймовірнісною схемою**. Як і в скінченній схемі p_n, $n = 1, 2, ...$ є ймовірностями елементарних подій. Ймовірність $\mathrm{P}(\cdot)$ має наступні властивості:

1) $\mathrm{P}(A) \geq 0$ для будь-якої події $A \in \Re$;

2) $\mathrm{P}(\Omega) = 1$;

3) якщо $A_1, ..., A_n, ... \in \Re$ такі, що $A_i \bigcap A_j = \varnothing$, $i \neq j$, то

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathrm{P}(A_n).$$

Наведемо два приклади визначення ймовірності у зліченній схемі:

1) **геометрична ймовірність** з параметром $0 < p < 1$

$$p_n = (1-p)p^{n-1}, \qquad n = 1,\ 2,\ ...$$

2) **ймовірність Пуассона** з параметром $\lambda > 0$

$$p_n = \frac{\lambda^n}{n!} e^{-\lambda}, \qquad n = 0,\ 1,\ ...$$

Приклад 3.1

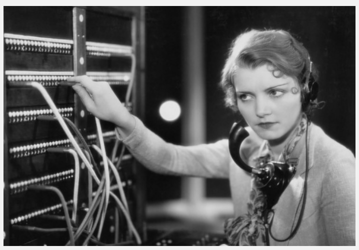
Велике значення для страхової компанії має час з моменту укладення угоди до виникнення страхового випадку і пов'язаного з цим відшкодуванням клієнтові. Припустимо, що $\Omega = \{\omega_n = n : n = 1, 2, ...\}$, $\mathrm{P}(\omega_n) = qp^{n-1}$, $0 < p < 1$, $q = 1 - p$. Знайти ймовірність того, що час до страхового випадку буде кратним деякому числу $m > 1$.

Розв'язок: Нехай $A = \{\omega_n : \ n = km, \ k = 1, \, 2, \, 3, \ ...\}$. Тоді

$$\mathrm{P}(A) = \sum_{k=1}^{+\infty} \mathrm{P}(\omega_{km}) = \sum_{k=1}^{+\infty} qp^{km-1} = \frac{q}{p} \frac{p^m}{1-p^m}.$$

Відповідь: $\dfrac{q}{p}\dfrac{p^m}{1-p^m}$.

Приклад 3.2



Припустимо, що кількість викликів (дзвінків), що приходять на автоматизовану телефонну станцію (АТС) упродовж години, розподілені за Пуассонівським законом з деяким параметром $\lambda > 0$, тобто $\Omega = \{\omega_n = n : n = 0, \, 1, \, 2, ...\}$,

$$\mathrm{P}(\omega_n) = p_n = \frac{\lambda^n}{n!} e^{-\lambda}.$$

Яку кількість викликів більш імовірно зареєструвати на АТС за деяку годину – парну чи непарну?

Розв'язок: Нехай

$$A = \{\omega_n : \ n = 2k, \ k = 0, \, 1, \ 2, \ 3, \ ...\}$$

та

$$B = \{\omega_n : \ n = 2k+1, \ k = 0, \, 1, \ 2, \ 3, \ ...\}.$$

Тоді

$$\mathrm{P}(A) = \sum_{k=0}^{+\infty} \mathrm{P}(\omega_{2k}) = \sum_{k=0}^{+\infty} \frac{\lambda^{2k}}{(2k)!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{1}{2} \left(\frac{\lambda^k}{k!} + (-1)^k \frac{\lambda^k}{k!} \right) =$$

$$= \frac{e^{-\lambda}}{2} \left(\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} + \sum_{k=0}^{+\infty} \frac{(-\lambda)^k}{k!} \right) = \frac{e^{-\lambda}}{2} \left(\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} + \sum_{k=0}^{+\infty} \frac{(-\lambda)^k}{k!} \right) = \frac{e^{-\lambda}}{2} \left(e^{\lambda} + e^{-\lambda} \right) = \frac{1 + e^{-2\lambda}}{2}$$

і аналогічно

$$\mathrm{P}(B) = \sum_{k=0}^{+\infty} \mathrm{P}(\omega_{2k+1}) = \frac{1 - e^{-2\lambda}}{2},$$

тому

$$\mathrm{P}(A) > \mathrm{P}(B).$$

Задача 3.1. Два гравці по черзі підкидають симетричну монету. Виграє той, у кого вперше випаде герб. Знайти ймовірності виграшу для кожного гравця.

Задача 3.2. Гральний кубик підкидається до тих пір, доки не випаде шістка. Знайти ймовірності таких подій:
а) шістка випаде за перші два підкидання; б) число підкидань непарне.

Задача 3.3. Симетрична монета підкидається до тих пір, доки вона не випаде два рази підряд однією стороною. Знайти ймовірність того, що число підкидань буде парним.

Задача 3.4.* Симетрична монета підкидається доки двічі підряд не випаде гербом. Знайти імовірність того, що число підкидань буде парним.

Задача 3.5. Гравці А і В грають у шахи. За виграш партії зараховується одне очко. Імовірність того, що партію виграє гравець А - α, гравець В - β ($\alpha > \beta$, $\alpha + \beta = 1$). Гру виграє той, хто випередить суперника на два очки. Яка ймовірність того, що гру виграє А, В? Що вигідніше для А: грати одну партію чи цілий матч?

Задача 3.6. (продовження попередньої задачі) Розв'язати попередню задачу, якщо умова "випередить суперника на два очки" замінюється на "виграє підряд дві партії". За якими з двох правил вигідніше для А грати матч?

Задача 3.7. Три спортсмена А, В та С однакової майстерності приймають участь у "круговому" змаганні: спочатку змагаються А та В, потім спортсмен С змагається з переможцем, і так далі, переможець кожного разу змагається з тим, хто відпочивав. Змагання триває доки один з спортсменів не переможе підряд двічі (нічийних результатів немає). Знайти ймовірності p_A, p_B, p_C того, що переможцем буде, відповідно А, В чи С. Як зміняться ці ж ймовірності, якщо відомо, що перше змагання виграв А?

Задача 3.8. Припустимо, що кількість викликів (дзвінків), що приходять на АТС упродовж години, розподілені за Пуассонівським законом. Ймовірність того, що за годину прийде парна кількість викликів, у $k > 1$ разів більша за ймовірність того, що за годину прийде непарна кількість викликів. Знайти параметр Пуассонівського розподілу λ. `Примітка: див. приклад 3.2 на стор. 30.`

Задача 3.9. Кількість виклик ів (дзвінків), що приходять на АТС упродовж години, розподілені за Пуассонівським законом. Частка тих годин, упродовж яких викликів немає, становить $\frac{1}{n}$, де n – деяке натуральне число. Яку частку складають ті години, коли приходить тільки один виклик? `Примітка: див. приклад 3.2 на стор. 30.`

Задача 3.10. Два симетричних гральних кубика підкидається до тих пір, доки в сумі не випаде 7. Знайти ймовірність того, що доведеться зробити більше трьох підкидань.

4 АКСІОМАТИКА ТЕОРІЇ ЙМОВІРНОСТЕЙ

Нехай маємо довільний простір елементарних подій $\Omega = \{\omega\}$. Сукупність підмножин $\Re$ множини Ω називається σ-алгеброю подій, якщо виконуються наступні умови:

1) $\Omega \in \Re$;

2) Якщо $A \in \Re$, то $\overline{A} \in \Re$;

3) Якщо $A_i \in \Re$ $(i = 1, 2 \ldots)$, то $\bigcup\limits_{i=1}^{\infty} A_i \in \Re$.

Властивість 3) може бути замінена наступною:

3') Якщо $A_i \in \Re$ $(i = 1, 2 \ldots)$, то $\bigcap\limits_{i=1}^{\infty} A_i \in \Re$.

Елементи множини $\Re$ називаються **подіями**. Числова функція $\mathrm{P}(\cdot)$, яка визначена на множині $\Re$, називається **ймовірністю**, якщо виконуються наступні умови (**аксіоми**):

1) **Аксіома невід'ємності**: $\mathrm{P}(A) \geq 0$ для кожної події A;

2) **Аксіома нормованості**: $\mathrm{P}(\Omega) = 1$;

3) **Аксіома сигма-адитивності**: якщо $A_i \in \Re$ $(i = 1, 2...)$, та $A_i A_j = \varnothing$ для $i \neq j$, то

$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(A_i).$$

Трійка $(\Omega, \Re, \mathrm{P})$ називається **ймовірнісним простором**.

Наслідками аксіом є наступні властивості:

1. $\mathrm{P}(\varnothing) = 0$;

2. $0 \leq \mathrm{P}(A) \leq 1$ для кожної події $A \in \Re$;

3. $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$ для кожної події $A \in \Re$;

4. Якщо $A \subset B$, то $\mathrm{P}(B \backslash A) = \mathrm{P}(B) - \mathrm{P}(A)$;

5. Якщо $A \subset B$, то $\mathrm{P}(A) \leq \mathrm{P}(B)$;

6. $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(AB)$ для кожних подій A, B (це так звана **теорема суми для двох подій**).

Приклад 4.1

Доведемо властивість 6. Дійсно події $A \cup B$ та B можна представити так: $A \cup B = A \cup \overline{A} \cap B$, $B = A \cap B \cup \overline{A} \cap B$. Очевидно доданки в правих частинах несумісні. Тому за аксіомою 3:

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(\overline{A} \cap B), \qquad \mathrm{P}(B) = \mathrm{P}(A \cap B) + \mathrm{P}(\overline{A} \cap B).$$

Віднімемо від першої рівності другу: $\mathrm{P}(A \cup B) - \mathrm{P}(B) = \mathrm{P}(A) - \mathrm{P}(A \cap B)$. Звідси випливає властивість 6.

Приклад 4.2

Нехай $\mathrm{P}(A) = \mathrm{P}(B) = \frac{1}{2}$. Довести, що

$$\mathrm{P}(AB) = \mathrm{P}(\bar{A}\bar{B}).$$

Доведення: у рівності

$$\mathrm{P}(\bar{A}\bar{B}) = 1 - \mathrm{P}(A \cup B).$$

застосуємо до від'ємника теорему суми для двох подій:

$$\mathrm{P}(\bar{A}\bar{B}) = 1 - (\mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(AB)). \tag{$*$}$$

Оскільки за умовою $\mathrm{P}(A) = \mathrm{P}(B) = \frac{1}{2}$, то $\mathrm{P}(A) + \mathrm{P}(B) = 1$ і формула $(*)$ перетворюється на

$$\mathrm{P}(\bar{A}\bar{B}) = \mathrm{P}(AB),$$

що і треба було довести.

Приклад 4.3

Довести, що

$$\mathrm{P}^2(AB) + \mathrm{P}^2(\overline{A}B) + \mathrm{P}^2(A\overline{B}) + \mathrm{P}^2(\overline{A}\,\overline{B}) \geq \frac{1}{4}.$$

Доведення: Оскільки події $AB,\ \overline{A}B,\ A\overline{B},\ \overline{A}\,\overline{B}$ є несумісними, а їхня сума утворює достовірну подію, то

$$\mathrm{P}(AB)+\mathrm{P}\left(\overline{A}B\right)+\mathrm{P}\left(A\overline{B}\right)+\mathrm{P}\left(\overline{A}\,\overline{B}\right)=1.$$

Покладемо

$$\mathrm{P}(AB)=x+\frac{1}{4},\qquad \mathrm{P}\left(\overline{A}B\right)=y+\frac{1}{4},$$
$$\mathrm{P}\left(A\overline{B}\right)=z+\frac{1}{4},\qquad \mathrm{P}\left(\overline{A}\,\overline{B}\right)=w+\frac{1}{4}.$$

Тоді $x+y+z+w=0$, а

$$\mathrm{P}^2(AB)+\mathrm{P}^2\left(\overline{A}B\right)+\mathrm{P}^2\left(A\overline{B}\right)+\mathrm{P}^2\left(\overline{A}\,\overline{B}\right)=\left(x+\frac{1}{4}\right)^2+\left(y+\frac{1}{4}\right)^2+\left(z+\frac{1}{4}\right)^2+$$

$$+\left(w+\frac{1}{4}\right)^2=x^2+y^2+z^2+w^2+\frac{x+y+z+w}{2}+\frac{1}{4}=x^2+y^2+z^2+w^2+\frac{1}{4}\geq\frac{1}{4},$$

що і треба було довести.

Приклад 4.4

П'ятеро дівчат прийшли на вечірку зі схожими сумочками, і після вечірки кожна взяла з собою випадково обрану сумочку. Яка ймовірність того, що хоча б одна з них взяла власну сумочку? Яка ймовірність того, що жодна з них не взяла власної сумочки?

Розв'язок: Позначимо події A_i – i-та дівчина взяла власну сумочку, $i=1,...,5$, A – хоча б одна дівчина взяла власну сумочку. Тоді $A=A_1\cup...\cup A_5$. Для даного типу задач найзручнішим способом розв'язання є безпосереднє застосування формули ймовірності об'єднання подій:

$$P(A) = P(A_1 \cup ... \cup A_5) = \underbrace{\sum_{i=1}^{5} P(A_i)}_{(1)} - \underbrace{\sum_{1 \le i < j \le 5} P(A_i \cap A_j)}_{(2)} +$$

$$+ \underbrace{\sum_{1 \le i < j < k \le 5} P(A_i \cap A_j \cap A_k)}_{(3)} - \underbrace{\sum_{1 \le i < j < k < l \le 5} P(A_i \cap A_j \cap A_k \cap A_l)}_{(4)} +$$

$$+ P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5).$$

Знайдемо ймовірності, які фігурують в правій частині формули і підрахуємо кількість доданків в кожній сумі. Кількість всіх елементарних подій – кількість способів розподілити сумочки між дівчатами – дорівнює 5!. Ймовірність того, що i-та дівчина взяла свою власну сумку (власниці підходить лише одна сумка, а інші сумки можна розподілити 4! способами) дорівнює

$$P(A_i) = \frac{4!}{5!} = \frac{1}{5}, \quad i = 1, ..., 5.$$

Якщо дві дівчини взяли правильні сумки (таких пар дівчат є $C_5^2 = 10$, отже в сумі (2) – 10 доданків), то решту можна розподілити 3! способами, відповідно ймовірність для двох дівчат (i-тої та j-тої) взяли власні сумочки

$$P(A_i \cap A_j) = \frac{3!}{5!} = \frac{1}{20}, \quad i \ne j, \;\; i, j = 1, ..., 5.$$

Аналогічно

$$P(A_i \cap A_j \cap A_k) = \frac{2!}{5!} = \frac{1}{60}, \quad i \ne j \ne k, \;\; i, j, k = 1, ..., 5;$$

$$P(A_i \cap A_j \cap A_k \cap A_l) = \frac{1}{5!} = \frac{1}{120}, \quad i \ne j \ne k \ne l, \;\; i, j, k = 1, ..., 5;$$

$$P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = \frac{1}{120}.$$

Причому різних потрійних перетинів може бути $C_5^3 = 10$, а різних перетинів з чотирма елементами – $C_5^4 = 5$, отже в сумі (3) маємо 10 доданків, а в сумі (4) – 5 доданків.

Підставляючи всі отримані значення в формулу для $P(A)$ маємо:

$$P(A) = 5 \cdot \frac{1}{5} - 10 \cdot \frac{1}{20} + 10 \cdot \frac{1}{60} - 5 \cdot \frac{1}{120} + \frac{1}{120} = \frac{19}{30}.$$

Подія "жодна з дівчат не взяла власної сумочки" є протилежною подією до події A. Отже, її ймовірність

$$P(\overline{A}) = 1 - P(A) = \frac{11}{30}.$$

Задача 4.1. Відомі ймовірності подій A, B та $A \cap B$. Знайти ймовірність події

$$\overline{A} \cap (A \cup B).$$

Задача 4.2. Нехай $A \cap B = \varnothing$. Довести, що

$$\mathrm{P}(A) \le \mathrm{P}\left(\overline{B}\right).$$

Задача 4.3. В офісі компанії "Odds.ltd" є два однаковісінькі автомати, які продають каву. Зранку обидва автомати заправлені найкращою кавою. Ймовірність того, що до вечора в одному автоматі закінчиться кава, складає 0.1, а ймовірність того, що в обох автоматах до вечора закінчиться кава, складає 0.05. Яка ймовірність того, що до вечора кава все ще буде в обох автоматах?

Задача 4.4. Нехай $\mathrm{P}(A) \ge 0.8$ та $\mathrm{P}(B) \ge 0.8$. Довести, що

$$\mathrm{P}(A \cap B) \ge 0.6.$$

Задача 4.5. Довести, що

$$\mathrm{P}(A \cup B \cup C) = \mathrm{P}(A) + \mathrm{P}(B) + \mathrm{P}(C) -$$

$$- \mathrm{P}(A \cap B) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C) + \mathrm{P}(A \cap B \cap C).$$

Задача 4.6. Довести, що для будь-яких подій A та B має місце:

$$\max\{\mathrm{P}(A), \mathrm{P}(B)\} \le \mathrm{P}(A \cup B) \le 2\max\{\mathrm{P}(A), \mathrm{P}(B)\}.$$

Задача 4.7. Нехай p_1, p_2, p_{12} – дійсні числа. Довести, що для того, щоб існували випадкові події A і B такі, що $\mathrm{P}(A) = p_1$, $\mathrm{P}(B) = p_2$, $\mathrm{P}(A \cap B) = p_{12}$, необхідно і достатньо, щоб виконувались нерівності:

$$1 - p_1 - p_2 + p_{12} \ge 0, \qquad p_i - p_{12} \ge 0 \quad (i = 1.2), \qquad p_{12} \ge 0.$$

Задача 4.8. Нехай $A_1, \ldots, A_n$ - випадкові події. Довести, що

а) $\mathrm{P}\left(\bigcup\limits_{i=1}^{n} A_i\right) \le \sum\limits_{i=1}^{n} \mathrm{P}(A_i)$. б) $\mathrm{P}\left(\bigcap\limits_{i=1}^{n} A_i\right) \ge 1 - \sum\limits_{i=1}^{n} \mathrm{P}(\overline{A}_i)$.

Задача 4.9. Довести, що

$$\mathrm{P}\left(\bigcap_{i=1}^{n} A_i\right) \ge \sum_{i=1}^{n} \mathrm{P}(A_i) - (n-1).$$

Задача 4.10. Довести, що для будь-яких трьох подій A, B та C

$$|\mathrm{P}\left(A \cap B\right) - \mathrm{P}\left(A \cap C\right)| \leq \mathrm{P}\left(B \Delta C\right),$$

де Δ – операція симетричної різниці двох подій, тобто $B\Delta C = (B \backslash C) \cup (C \backslash B)$.

Задача 4.11. Довести нерівність

$$\mathrm{P}\left(A \Delta B\right) \leq \mathrm{P}\left(A \Delta C\right) + \mathrm{P}\left(C \Delta B\right).$$

Задача 4.12. Довести, що для будь-яких двох подій A та B

$$|\mathrm{P}\left(A \cap B\right) - \mathrm{P}(A)\mathrm{P}(B)| \leq \frac{1}{4}.$$

Задача 4.13. Довести, що для довільних випадкових подій A і B виконуються співвідношення

а) $\mathrm{P}(A \cup B) \cdot \mathrm{P}(A \cap B) \leq \mathrm{P}(A) \cdot \mathrm{P}(B)$;
б) $\mathrm{P}^2(A \cup B) + \mathrm{P}^2(A \cap B) = \mathrm{P}^2(A) + \mathrm{P}^2(B) + 2 \cdot \mathrm{P}(A \cap \bar{B}) \cdot \mathrm{P}(\bar{A} \cap B)$.

Задача 4.14. Нехай

$$S_0^{(n)} = 1,$$

$$S_k^{(n)} = \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \mathrm{P}\left(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}\right), \qquad k = 1, 2, \ldots, n.$$

Довести, що
а) $\dfrac{S_{k+1}^{(n)}}{C_n^{k+1}} \leq \dfrac{S_k^{(n)}}{C_n^k}$, $k = 0.1, \ldots, n-1$; б) $\mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} (-1)^{k-1} S_k^{(n)}$.

5 ГЕОМЕТРИЧНЕ ВИЗНАЧЕННЯ ЙМОВІРНОСТІ

Нехай стохастичний експеримент полягає у тому, що ми кидаємо навмання точку у деяку область M евклідового простору R^n, причому $\lambda(M) < \infty$, де $\lambda(\circ)$ - міра Лебега в просторі R^n. В цьому випадку $\Omega = M$, а за σ-алгебру подій $\Re$ візьмемо сукупність усіх підмножин з M, вимірних за Борелем. Якщо $A \in \Re$, то за визначенням ймовірність події A визначається так

$$\mathrm{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)}. \tag{5.1}$$

Формула (5.1) називається **геометричним визначенням ймовірності**.

В частинному випадку, коли $n = 1$, а $M = [a, b]$, $-\infty < a < b < \infty$, маємо

$$\mathrm{P}(A) = \frac{\lambda(A)}{b - a},$$

де $\lambda(\circ)$ - міра Лебега на прямій.

Приклад 5.1

На відрізку M_1M_2, довжина якого дорівнює L, навмання обирається точка K. Знайти імовірність того, що відстань від точки K до точки M_1 перевищує її відстань до точки M_2.

Розв'язок: Позначимо через G множину точок відрізку M_1M_2, а через x – координату точки K відрізку M_1M_2, вважаючи точку M_1 початком відліку. Тоді множина $g \subset G$, що відповідає наведеній у прикладі випадкової події A, складе

$$g = \{K \in G | x > L - x\} = \left\{K \in G | x > \frac{L}{2}\right\}.$$

На основі геометричного визначення ймовірності (5.1) маємо $\mathrm{P}(A) = \frac{L}{2L} = \frac{1}{2}$.

Відповідь: $1/2$.

Приклад 5.2

На відрізок довжини l навмання кидають дві точки. Знайти ймовірність того, що з утворених частин можна скласти трикутник.

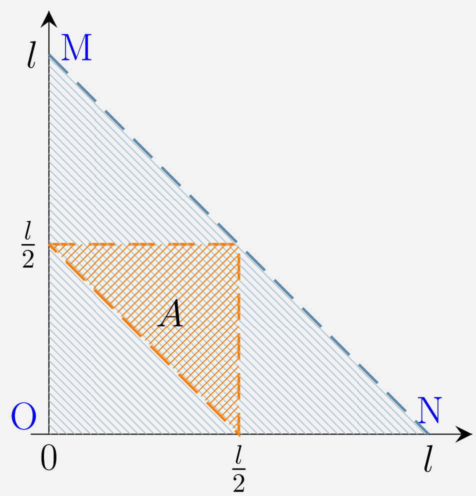
Розв'язок: Позначимо через x, y, $l-x-y$ довжини частин відрізка що утворилися. Тоді за простір елементарних подій беремо множину $\Omega = \{(x,y) : x > 0,\ y > 0,\ x+y < l\}$. У декартовій системі координат на площині простором елементарних подій Ω буде трикутник OMN. Нехай A – подія, яка полягає в тому, що з утворених частин можна створити трикутник. Подія A описується наступним чином

$$A = \{(x,y) : x < y + l - x - y,\ y < x + l - x - y,\ l - x - y < x + y\}$$

або

$$A = \{(x,y) : x < l/2,\ y < l/2,\ x + y > l/2\}.$$

На координатній площині множині A відповідає область, що заштрихована на наступному рисунку.



Згідно з геометричним визначенням ймовірності маємо

$$\mathrm{P}(A) = \frac{l^2/8}{l^2/2} = \frac{1}{4}.$$

Відповідь: $1/4$.

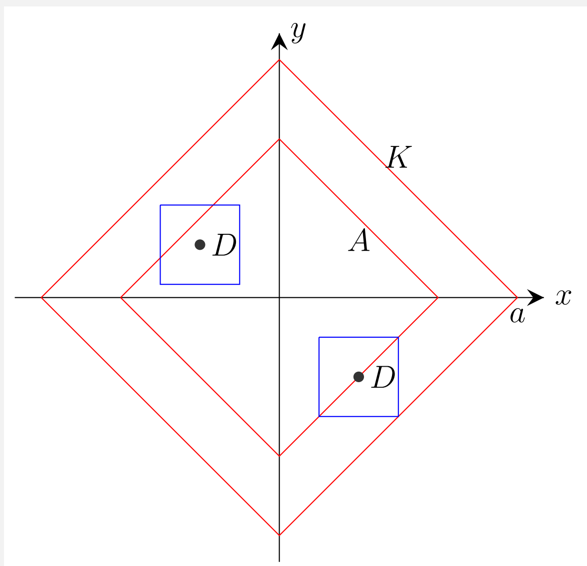
Приклад 5.3

Точка D випадковим чином розміщена в квадраті

$$K = \{(x, y) : |x| + |y| \le a\}.$$

Знайти ймовірність того, що квадрат з центром в точці D і сторонами довжини d $(d < a)$, паралельними осям координат, повністю знаходиться в квадраті K.

Розв'язок: Простором елементарних подій даного експерименту є квадрат K, сторона цього квадрату дорівнює $a\sqrt{2}$, а його площа відповідно $\lambda(\Omega) = S(K) = 2a^2$.



Всередині квадрата K побудуємо квадрат A, сторони якого знаходяться на відстані $d/\sqrt{2}$ від його сторін. Якщо точка D потрапить в цей внутрішній квадрат, то умова задачі буде мати місце. Якщо ж точка D потрапить в смугу між квадратами, то квадрат з центром в D перетне принаймні одну зі сторін квадрата K.

Таким чином, квадрат A є множиною елементарних подій, які забезпечують появу події A. Довжина сторони цього квадрата дорівнює

$$\sqrt{2}a - 2\frac{d}{\sqrt{2}} = \sqrt{2}(a - d).$$

Відповідно шукана ймовірність дорівнює

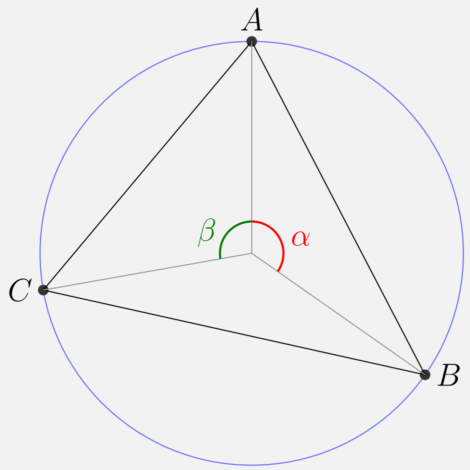
$$\mathrm{P}(A) = \frac{S(A)}{S(\Omega} = \frac{2(a-d)^2}{2a^2} = \frac{(a-d)^2}{a^2}.$$

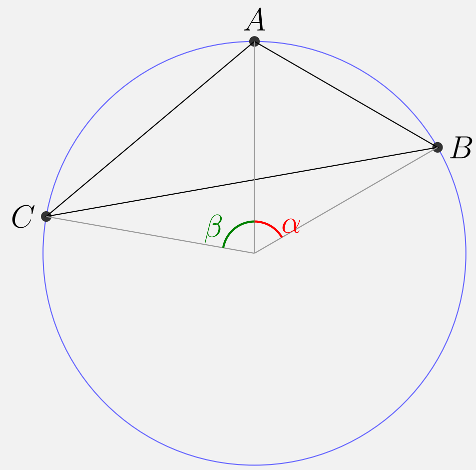
Відповідь: $\dfrac{(a-d)^2}{a^2}$.

Приклад 5.4

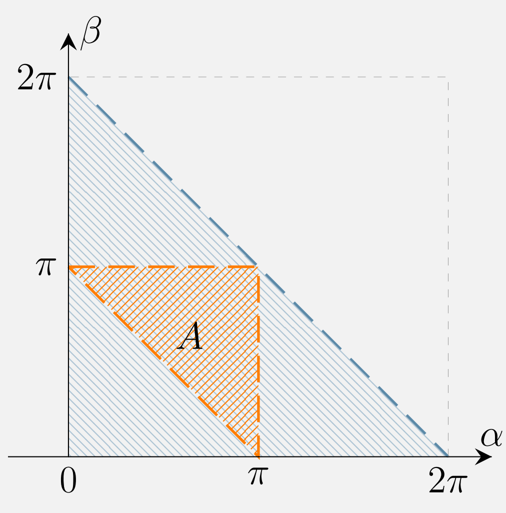
Три точки A, B, C навмання обираються на колі одиничного радіуса. Яка ймовірність того, що трикутник ABC буде гострокутним?

Розв'язок: Зафіксуємо точку A в будь-якій точці кола. Тоді положення двох інших точок визначатиметься величиною секторів величиною α та β, що розходяться по різні боки від точки A, відповідно $0 < \alpha+\beta < 2\pi$. Ця нерівність визначає область, яка є простором елементарних подій. Для того, щоб трикутник був гострокутним, необхідно, щоб $0 < \alpha < \pi$ та $0 < \beta < \pi$. Окрім того, їхня сума не може бути менша за π, оскільки в цьому разі кут A трикутника буде тупим, отже, маємо ще умову $\alpha + \beta > \pi$.





Зобразимо область Ω та область, що відповідає шуканій події: Ω – це трикутник, обмежений нерівностями $\alpha > 0$, $\beta > 0$ та $0 < \alpha + \beta < 2\pi$. Область, яка відповідає нашій події (нехай подія A) – це трикутник, обмежений нерівностями $\alpha < \pi$, $\beta < \pi$ та $\alpha+\beta > \pi$.



Відношення площ зображених фігур дорівнює $1/4$.

Відповідь: 1/4.

Зауважимо, що якщо випадковий трикутник не є гострокутним, його можна вважати тупокутним, оскільки ймовірність того, що дві з трьох точок A, B, C утворюють діаметр, дорівнює 0. (Щоб BC був діаметром, потрібно мати $\alpha + \beta = \pi$, яка є на рис. прямою лінією з нульовим значенням площі.) Таким чином, ми можемо сказати, що ймовірність того, що трикутник ABC буде тупокутним, дорівнює 3/4.

Для тупокутного трикутника коло можна розділити на дві половини, причому трикутник повністю лежить в одній з половин. Це, зокрема, означає, що 3/4 – це розв'язок задачі:

Три точки A, B, C навмання розміщені на колі радіуса 1. Яка ймовірність того, що всі три точки лежать у півколі?

Задача 5.1. У круг вписано квадрат. Точку навмання кинуто у круг. Знайти ймовірність того, що вона потрапить у квадрат.

Задача 5.2. На відрізку PQ довжини l вибрані навмання дві точки A і B. Знайти ймовірність того, що:
а) точка A буде ближче до точки P, ніж до B;
б) точка A буде ближче до точки B, ніж до P.

Задача 5.3. На паркетну підлогу навмання кидають монету діаметра d. Паркет має форму квадратів зі стороною a ($a > d$). Яка ймовірність того, що
а) монета не перетне жодну з сторін квадратів паркету?
б) монета перетне рівно одну сторону квадрату паркету?
в) монета перетне рівно дві сторони квадрату паркету?
г) монета накриє кут квадрату?

Задача 5.4. На паркетну підлогу навмання кидають монету діаметра 1. Паркет має форму прямокутних трикутників з катетами 3 та 4. Яка ймовірність того, що монета не перетне жодного катету?

Задача 5.5. На площині проведено паралельні прямі, відстані між якими дорівнюють почергово 2 та 8 см. На площину кидають навмання коло радіуса 4. Яка ймовірність того, що коло не перетне жодну з ліній?

Задача 5.6. У сфері радіусу R навмання обрано N точок. Знайти ймовірність того, що відстань від центру до найближчої точки не перевищуватиме r. За яких умов границя цієї ймовірності буде додатною, якщо N і R прямують до нескінченності?

Задача 5.7. Дві особи вирішили зустрітися між восьмою і дев'ятою годиною вечора. Чекати один одного вони домовилися не більше десяти хвилин. Яка ймовірність їхньої зустрічі?

Задача 5.8.* Три особи вирішили зустрітися за умовами попередньої задачі.
а) Яка ймовірність зустрічі трьох осіб? б) Яка ймовірність того, що зустрінуться принаймні двоє? в) Узагальнити відповідь пункту а) на випадок часу очікування в $a \in [0; 60]$ хвилин.

Задача 5.9. Хлопець домовився про зустріч з дівчиною між 6 та 7 годинами вечора. Причому хлопець пообіцяв чекати дівчину до 7 вечора, а дівчина, якщо прийде раніше, чекатиме 10 хвилин. Яка ймовірність того, що вони зустрінуться, якщо кожен з них може прийти в будь-який момент упродовж години.

Задача 5.10. На відрізку довжини m навмання узяли дві точки. Яка ймовірність того, що відстань між ними менша за αm, де $0 < \alpha < 1$?

Задача 5.11. Стрижень довжини m навмання розламали на дві частини. Яка ймовірність того, що довжина меншої частини менша за $m/3$?

Задача 5.12. Шматок дроту довжиною 7 см зігнуто під прямим кутом у випадково обраному місці. Яка ймовірність того, що відстань між кінцями дроту більша за 5 см?

Задача 5.13. Точка A рівномірно розподілена у квадраті із стороною 1. Знайти імовірності наступних подій:
а) відстань від точки A до фіксованої сторони квадрату не перевищує x;
б) відстань від точки A до найближчої сторони квадрату не перевищує x;
в) відстань від точки A до центра квадрату не перевищує x; зобразити залежність цієї ймовірності від x.

Задача 5.14. У квадрат навмання кидається точка. Яка ймовірність того, що відстань до найближчої сторони квадрату буде менша, ніж до найближчої діагоналі?

Задача 5.15. Два судна рухаються незалежним чином і повинні підійти до одного і того ж причалу. Моменти появи суден рівноможливі упродовж доби. Знайти ймовірність того, що одному з суден прийдеться чекати на звільнення причалу, якщо час зупинки першого судна одна година, а другого – дві години.

Задача 5.16. У випадковий момент часу $\alpha \in [0, T]$ з'являється радіосигнал тривалості t_1. Для його прийому у випадковий момент часу $\beta \in [0, T]$ включається приймач на час t_2. Знайти ймовірність прийому сигналу.

Задача 5.17. З відрізка [–1, 2] навмання узяли два числа. Яка ймовірність того, що їхня сума більша за одиницю, а добуток менший?

Задача 5.18. Значення α і β рівноможливі у квадраті $|\alpha| \le 1$, $|\beta| \le 1$. Знайти ймовірності наступних подій:
$A = \{$корені рівняння $x^2 + 2\alpha x + \beta = 0$ є дійсними числами$\}$,
$B = \{$корені рівняння $x^2 + 2\alpha x + \beta = 0$ є додатними числами$\}$.

Задача 5.19. Знайти ймовірність того, що корені рівняння $x^2 + 2ax + b = 0$ є дійсними, якщо для деякого $k > 0$ точка (a, b) рівноможлива у прямокутнику:
a) $-k \le a \le k$, $-k^2 \le b \le k^2$;
b) $-k \le a \le k$, $-m \le b \le m$, де $0 < m < k^2$;
c) $-k \le a \le k$, $-m \le b \le m$, де $m > k^2$.

Задача 5.20. Дано два концентричних кола радіусів $r_2 > r_1$. На більшому колі навмання обрано дві точки A і B . Яка ймовірність того, що відрізок AB не перетне мале коло?

Задача 5.21. Якої товщини повинна бути монета, щоб ймовірність падіння її на ребро дорівнювала $1/3$?

Задача 5.22. Парадокс Бертрана. У колі радіуса R випадковим чином проводиться хорда. Нехай ξ – її довжина. Знайти ймовірність того, що $\xi > R$, якщо:
а) середина хорди рівномірно розподілена у колі;
б) напрям хорди задано, а її середина рівномірно розподілена на діаметрі, який перпендикулярний до цього напряму;
в) один кінець хорди закріплений, а другий рівномірно розподілений на колі.

Задача 5.23. З відрізка $[0.1]$ навмання, незалежним чином обрали три числа. Яка ймовірність того, що з трьох відрізків, довжинами яких є обрані числа, можна побудувати трикутник?

Задача 5.24. Задача Бюффона. Площина розкреслена паралельними прямими, відстань між якими дорівнює $2m$. На цю площину навмання кидається голка довжини m. Яка ймовірність того, що вона перетне одну з паралельних прямих?

Задача 5.25.* Площина розкреслена паралельними прямими, відстань між якими дорівнює m. На цю площину кидається правильний n-кутник діаметром менше m (діаметром геометричної фігури називається найбільша відстань між її точками) та стороною a. Знайти ймовірність того, що n- кутник перетне одну з паралельних прямих.

Задача 5.26.* У квадрат навмання кинуто дві точки A і B. Знайти ймовірність того, що
а) коло, діаметр якого є відрізок AB, знаходиться у квадраті.
б) квадрат, діагональ якого є відрізок AB, знаходиться у квадраті.

Задача 5.27.* На колі навмання береться точка A (рівномірний розподіл на колі), а всередині кола навмання береться точка B (рівномірний розподіл в крузі). На відрізку AB як на діагоналі будується прямокутник, сторони якого паралельні осям координат. Яка ймовірність того, що прямокутник не перетне коло?

Задача 5.28.* Дві точки (A та B) навмання кинуто всередину кола, яке розглядаємо в деякій прямокутній декартовій системі координат. Потім будуємо прямокутник AA_1BB_1 зі сторонами, паралельними до осей координат. Знайти ймовірність того, що прямокутник AA_1BB_1 буде повністю знаходитися всередині вихідного кола.

Задача 5.29. Число x навмання обирають з відрізка $[0.5;\ 0.8]$. Яка ймовірність того, що четвертий доданок $T_4 = 560 \cdot 8 \cdot x^3 \cdot 513$ біноміального розкладу $(5+2x)^{16}$ більший за два сусідніх доданки?

Задача 5.30. Число x навмання обирають з відрізка $[0.6;\ 1]$. Яка ймовірність того, що єдиним найбільшим доданком біноміального розкладу $(5+3x)^{10}$ є четвертий?

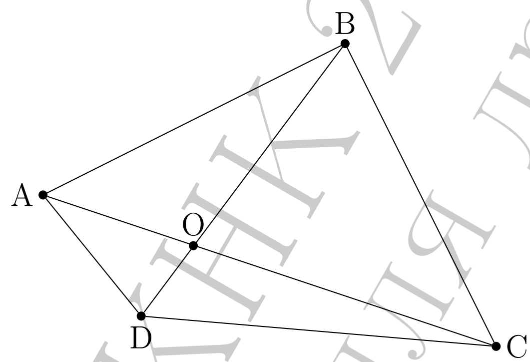
Задача 5.31. Знайти номер найбільшого члена $T_k = C_n^{k-1} q^{k-1} p^{n-k+1}$ біноміального розкладу $(p+q)^n$ за степенями p, якщо $p > 0$, $q > 0$, $p + q = 1$. При яких умовах

а) найбільший член буде останнім,

б) найбільший член буде першим,

в) розклад буде містити два однакових найбільших члени?

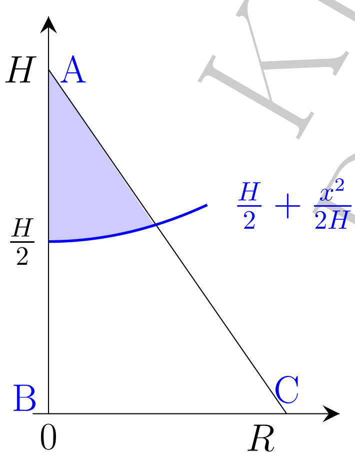
Задача 5.32.



У опуклому чотирикутнику $ABCD$ діагоналі AC та BD перетинаються в точці O. Відомі площі наступних трикутників: $S_{AOB} = 3$ см^2, $S_{AOD} = 1$ см^2, $S_{DOC} = 2$ см^2. Із чотирикутника $ABCD$ навмання обирають одну точку. Яка ймовірність того, що ця точка буде належати трикутнику BOC?

Задача 5.33. У конусі з висотою H і радіусом основи R навмання обирають точку M. Знайти ймовірність того, що точка M буде ближчою до вершини конуса, ніж до його основи.

`Примітка: подумайте, чому наступний план розв'язання` ***не є правильним***.

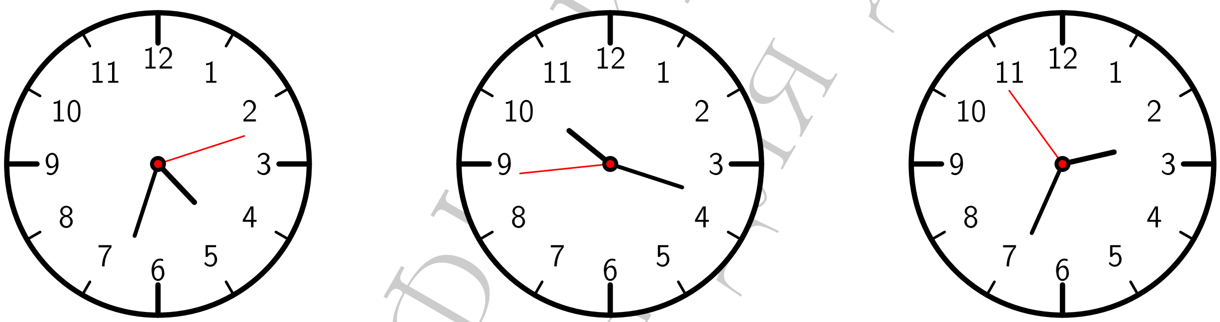


Оскільки точка M завжди потрапляє до центрального перерізу конуса, то розглянемо той центральний переріз, який містить точку M. Цей переріз буде рівнобедреним трикутником і, не порушуючи загальності, можемо вважати, що точка M потрапляє в "праву" частину цього трикутника, тому завжди можемо розглянути лише цю "праву" частину, яка буде прямокутним трикутником з катетами R і H. Для зручності вважаємо, що цей прямокутний трикутник ABC розміщено у прямокутній системі координат так, що $A(0, H)$, $B(0, 0)$, $C(R, 0)$ (див. рис.) Геометричне місце точок, рівновіддалених від вершини A та від основи BC за визначенням параболи є саме параболою, для якої нескладно записати і явний вид: $y(x) = \dfrac{H}{2} + \dfrac{x^2}{2H}$. Парабола перетинає гіпотенузу AC в точці з координатою $x_0 = \dfrac{-H^2 + H\sqrt{H^2 + R^2}}{R}$, а шукана ймовірність є відношенням площі заштрихованої фігури (її шукаємо через інтегрування) до площі трикутника ABC.

Задача 5.34. До станції метро можна доїхати автобусом за 10 хв. або дістатися пішки за 13 хв. Інтервал руху автобусів становить 7 хв. Будемо вважати, що варто їхати автобусом, якщо ймовірність випадкової події A $=$ {швидше доїхати до метро автобусом} перевищує $1/2$. Чи варто чекати автобуса?

Задача 5.35. На відрізку OA числової осі Ox, довжина якого дорівнює L, навмання поставлено дві точки: $B(x)$ та $C(y)$, x – координата точки B, тобто точки, яка опинилась ближче до O, а y – координата точки C, тобто точки, яка опинилася далі від O, отже $y \geq x$. Знайти ймовірність того, що довжина відрізку BC менша за довжину відрізка OB.

Задача 5.36. Яка ймовірність того, що із трьох узятих навмання відрізків довжиною не більше за l (кожний) можна побудувати трикутник?



(а) (б) (в)

Рис. 5.1: до задачі 5.37

Задача 5.37. Три стрілки годинника (H – годинна, M – хвилинна і S – секундна) рухаються плавно і рівномірно. Знайти ймовірності наступних подій:

1) $A =$ {кут між H та M буде гострим} (Рис. 5.1а);

2) $B =$ {кут між S та M буде тупим} (Рис. 5.1б);

3) $C =$ {усі три кути (між H та M, між H та S, між S та M) будуть однакові} (Рис. 5.1в).

6 УМОВНІ ЙМОВІРНОСТІ. НЕЗАЛЕЖНІСТЬ ПОДІЙ

Нехай дано ймовірнісний простір $(\Omega, \Re, \mathrm{P})$. Події A та B називаються **незалежними**, якщо

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

Події $A_1, A_2, \ldots, A_n$ називаються **незалежними у сукупності**, якщо

$$\mathrm{P}\left(\bigcap_{j=1}^{k} A_{n_j}\right) = \prod_{j=1}^{k} \mathrm{P}(A_{n_j}) \tag{6.1}$$

для кожної послідовності індексів $n_1 < n_2 < \ldots < n_k$, $1 < k \leq n$.

Якщо подія B така, що $\mathrm{P}(B) > 0$, то **умовна ймовірність** події A при умові, що відбулась подія B, визначається наступним чином

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(AB)}{\mathrm{P}(B)}. \tag{6.2}$$

З (6.2) випливає, що

$$\mathrm{P}(AB) = \mathrm{P}(A|B)\mathrm{P}(B) \tag{6.3}$$

та

$$\mathrm{P}(BA) = \mathrm{P}(B|A)\mathrm{P}(A). \tag{6.4}$$

Формули (6.3), (6.4) називаються **формулами добутку** або **теоремами добутку**. Формула (6.3) може бути узагальнена наступним чином:

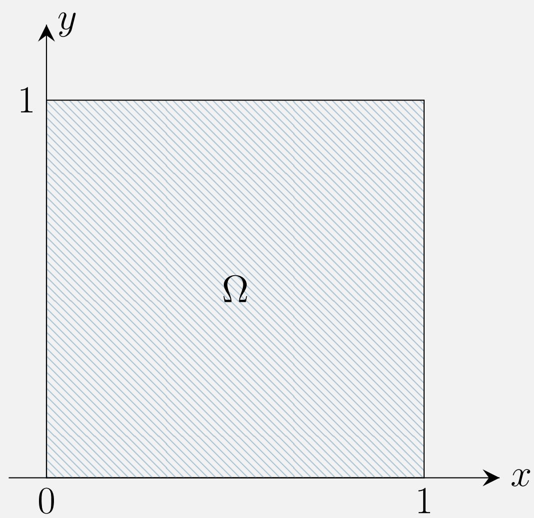
$$\mathrm{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathrm{P}(A_n|A_{n-1}A_{n-2}\ldots A_1)\cdot \mathrm{P}(A_{n-1}|A_{n-2}\ldots A_1)\cdot\ldots\cdot \mathrm{P}(A_2|A_1)\cdot \mathrm{P}(A_1).$$

Умовна ймовірність має наступні властивості.

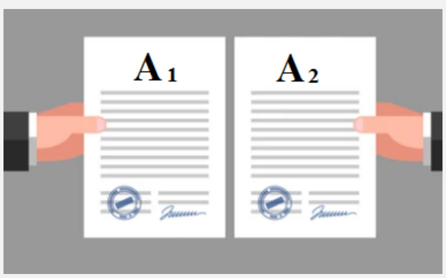
1. $0 \leq \mathrm{P}(A|B) \leq 1$;
2. $\mathrm{P}(\Omega|B) = 1$;
3. $\mathrm{P}(B|B) = 1$;
4. Якщо події A та B є незалежними, то $\mathrm{P}(A|B) = \mathrm{P}(A)$;
5. Якщо $A_i \in \Re$ $i = 1, 2, \ldots$ та $A_i \cap A_j = \varnothing$ $(i \neq j)$, то $\mathrm{P}\left(\bigcup\limits_{k=1}^{\infty} A_k \middle| B\right) = \sum\limits_{k=1}^{\infty} \mathrm{P}\left(A_k|B\right)$.

Приклад 6.1

Розглянемо приклад незалежних подій. Нехай подія A полягає у тому, що навмання кинута у квадрат зі стороною 1 точка впала в область, що лежить вправо від абсциси a, подія B – у тому, що точка впала в область, що лежить вище ординати b. Події A та B незалежні, бо $\mathrm{P}(AB) = \mathrm{P}(A)\mathrm{P}(B)$, що підтверджується рисунком



Приклад 6.2



Припустимо, що страхова компанія має два типи угод A_1 та A_2, кількість яких дорівнює відповідно n та m. Всі страхові випадки однаково ймовірні. Нехай C_i – подія, яка полягає в тому, що i-ий позов до компанії належить типу A_1. Знайти ймовірність того, що перші три позови будуть за угодами типу A_1.

Розв'язок:

$$\mathrm{P}(C_1 \cap C_2 \cap C_3) = \mathrm{P}(C_1)\mathrm{P}(C_2|C_1)\mathrm{P}(C_3|C_1 \cap C_2) =$$

$$= \frac{n}{n+m} \cdot \frac{n-1}{n+m-1} \cdot \frac{n-2}{n+m-2} = \frac{n(n-1)(n-2)}{(n+m)(n-1+m)(n-2+m)}.$$

Відповідь: $\dfrac{n(n-1)(n-2)}{(n+m)(n-1+m)(n-2+m)}$.

Приклад 6.3

З колоди карт у 36 листів навмання витягають одну карту. Розглянемо такі події: $A = \{$карта чорної масті$\}$, $B = \{$туз$\}$, $C = \{$пікова дама$\}$. Покажемо, що події A та B незалежні, а події A та C залежні.

Дійсно, $\mathrm{P}(A) = \frac{18}{36} = \frac{1}{2}$, $\mathrm{P}(B) = \frac{4}{36} = \frac{1}{9}$, $\mathrm{P}(C) = \frac{1}{36}$,

$$\mathrm{P}(AB) = \frac{2}{36} = \frac{1}{18} = \mathrm{P}(A)\mathrm{P}(B) = \frac{1}{2} \cdot \frac{1}{9},$$

$$\mathrm{P}(AC) = \mathrm{P}(C) = \frac{1}{36} \neq \mathrm{P}(A)\mathrm{P}(C) = \frac{1}{2} \cdot \frac{1}{36},$$

що і треба було довести.

Приклад 6.4

Підкинули два гральних кубики. Яка ймовірність того, що на обох кубиках випали парні числа, якщо відомо, що сума чисел, що випали, дорівнює 6?

Розв'язок: $\Omega = \{(i,j) \mid i,j = 1,2,\dots,6\}$, $|\Omega| = 6 \cdot 6 = 36$,

$$A = \{(i,j) \in \Omega |\ i \text{ та } j \text{ парні}\},$$

$$B = \{(i,j) \in \Omega |\ \ i+j=6\} = \{(1;5),\ (2;4),\ (3;3),\ (4;2),\ (5;1)\},$$

$$|B| = 5, \quad \mathrm{P}(B) = \frac{5}{36},$$

$$AB = \{(i,j) \in \Omega |\ \ i \text{ та } j \text{ парні}, \ \ i+j=6\} = \{(2;4),\ (4;2)\},$$

$$|AB| = 2, \quad \mathrm{P}(AB) = \frac{2}{36},$$

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(AB)}{\mathrm{P}(B)} = \frac{2/36}{5/36} = \frac{2}{5}.$$

Відповідь: $\frac{2}{5}$.

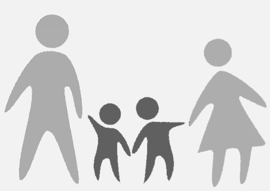
Приклад 6.5

В урні міститься 6 кульок, серед яких дві білі та чотири чорних. На кожному кроці випадковим чином витягується кулька, фіксується її колір, і вона повертається назад в урну разом з двома кульками того самого кольору. Яка ймовірність того, що в перших трьох експериментах кожен раз буде витягнуто білу кульку?

Розв'язок: Позначимо через A_i, $i = 1, 2, 3$, подію, яка полягає в тому, що на i-тому кроці витягується біла куля. Тоді шукана ймовірність – це $\mathrm{P}(A_1 \cap A_2 \cap A_3)$. Згідно правилу добутку маємо

$$\mathrm{P}(A_1 \cap A_2 \cap A_3) = \mathrm{P}(A_1) \cdot \mathrm{P}(A_2|A_1) \cdot \mathrm{P}(A_3|A_1 \cap A_2) = \frac{2}{6} \cdot \frac{4}{8} \cdot \frac{6}{10} = \frac{1}{10}.$$

Приклад 6.6



З усіх сімей з двома дітьми обрано одну. Описати простір елементарних подій та випадкові події: A – «в сім'ї є хлопчик та дівчинка», B – «в сім'ї не більше однієї дівчинки». Всі елементарні події рівноймовірні. Знайти $\mathrm{P}(A)$, $\mathrm{P}(B)$, $\mathrm{P}(A \cap B)$, та довести, що події A та B залежні.

Розв'язок: Простір елементарних подій $\Omega = \{\text{ХХ, ДД, ХД, ДХ}\}$. Випишемо події та визначимо їхні ймовірності:

$$A = \{\text{ХД, ДХ}\}, \quad B = \{\text{ХХ, ХД, ДХ}\}, \quad A \cap B = \{\text{ХД, ДХ}\};$$

$$\mathrm{P}(A) = \frac{1}{2}; \quad \mathrm{P}(B) = \frac{3}{4}; \quad \mathrm{P}(A \cap B) = \frac{1}{2} \neq \mathrm{P}(A)\mathrm{P}(B).$$

Отже, події A та B залежні.

Приклад 6.7

З усіх сімей з трьома дітьми обрано одну. Описати простір елементарних подій та випадкові події: A – «в сім'ї є хлопчик та дівчинка», B – «в сім'ї не більше однієї дівчинки». Всі елементарні події рівноймовірні. Знайти $\mathrm{P}(A)$, $\mathrm{P}(B)$, $\mathrm{P}(A \cap B)$, та довести, що події A та B незалежні.

Розв'язок: Простір елементарних подій

$$\Omega = \{\text{ХХХ, ДДД, ХХД, ДХХ, ХДХ, ДДХ, ХДД, ДХД}\}.$$

Випишемо події та визначимо їхні ймовірності:

$$A = \{\text{ХХД, ДХХ, ХДХ, ДДХ, ХДД, ДХД}\}, \quad B = \{\text{ХХХ, ХХД, ДХХ, ХДХ}\},$$

$$A \cap B = \{\text{ХХД, ДХХ, ХДХ}\};$$

$$\mathrm{P}(A) = \frac{6}{8} = \frac{3}{4}; \quad \mathrm{P}(B) = \frac{4}{8} = \frac{1}{2}; \quad \mathrm{P}(A \cap B) = \frac{3}{8} = \mathrm{P}(A)\mathrm{P}(B).$$

Отже, події A та B в даному випадку незалежні.

Приклад 6.8

Довести нерівність $\mathrm{P}(A|B) \geq 1 - \dfrac{\mathrm{P}(\overline{A})}{\mathrm{P}(B)}$.

Розв'язок: Перепишемо нерівність у вигляді

$$\frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \geq \frac{\mathrm{P}(B) - (1 - \mathrm{P}(A))}{\mathrm{P}(B)}.$$

Воно справедливе лише коли виконується

$$\mathrm{P}(A \cap B) \geq \mathrm{P}(B) + \mathrm{P}(A) - 1,$$

або

$$\mathrm{P}(B) + \mathrm{P}(A) - \mathrm{P}(A \cap B) \leq 1.$$

В лівій частині – ймовірність об'єднання подій:

$$\mathrm{P}(A \cup B) \leq 1.$$

Ця нерівність є справедливою згідно аксіомам ймовірності.

Приклад 6.9

Два мисливця стріляють по мішені. Вони влучають в мішень незалежно один від одного з ймовірностями 0.9 та 0.8 відповідно. Яка ймовірність того, що влучить лише один з них?

Розв'язок: Позначимо шукану подію через A. Вона матиме місце, якщо одночасно перший мисливець влучить, а другий не влучить, або другий влучить, а перший не влучить. Якщо позначити події A_1 – перший мисливець влучив при одному пострілі, A_2 – другий мисливець влучив при одному пострілі, то

$$A = (A_1 \cap \overline{A_2}) \cup (\overline{A_1} \cap A_2).$$

Маємо об'єднання несумісних подій. Зважаючи на незалежність подій A_1 та A_2 і, як наслідок, незалежність їхніх заперечень, можемо знайти ймовірність події A:

$$\begin{aligned} \mathrm{P}(A) = \mathrm{P}((A_1 \cap \overline{A_2}) \cup (\overline{A_1} \cap A_2)) = \mathrm{P}(A_1)\mathrm{P}(\overline{A_2}) + \mathrm{P}(\overline{A_1})\mathrm{P}(A_2) = \\ = 0.9(1-0.8) + (1-0.9)0.8 = 0.26. \end{aligned}$$

Відповідь: 0.26.

Приклад 6.10

Два стрільці двічі стріляють по мішені. Вони влучають в мішень при одному пострілі незалежно один від одного з ймовірностями p_1 та p_2 відповідно. Той, хто влучить більше разів, отримає грошову нагороду. Яка ймовірність того, що нагороду отримає перший стрілець?

Розв'язок: Позначимо подію A – виграє перший стрілець. Вона матиме місце, якщо одночасно перший влучить двічі, а другий не влучить жодного разу або влучить 1 раз, або якщо перший влучить один раз, а другий не влучить жодного. Якщо позначити події A_i – перший стрілець влучив рівно i разів, $i = 0, 1, 2$, B_i – другий стрілець влучив рівно i разів, $i = 0, 1, 2$, то

$$A = (A_2 \cap B_0) \cup (A_2 \cap B_1) \cup (A_1 \cap B_0).$$

Зважаючи на несумісність подій, які об'єднуються, та на незалежність подій A_i та B_i, $i = 0, 1, 2$, можемо знайти ймовірність події A:

$$P(A) = P((A_2 \cap B_0) \cup (A_2 \cap B_1) \cup (A_1 \cap B_0)) = P(A_2)P(B_0) + P(A_2)P(B_1) + P(A_1)P(B_0) =$$
$$= p_1^2(1-p_2)^2 + 2p_1^2p_2(1-p_2) + 2p_1(1-p_1)(1-p_2)^2.$$

Відповідь: $p_1^2(1-p_2)^2 + 2p_1^2p_2(1-p_2) + 2p_1(1-p_1)(1-p_2)^2$.

Приклад 6.11

Нехай A та B – незалежні події. Довести, що якщо $A \cup B$ та $A \cap B$ незалежні, то або $P(A) = 1$, або $P(B) = 1$, або $P(A) = 0$, або $P(B) = 0$.

Розв'язок:

Оскільки вказані пари подій A та B і $A \cup B$ та $A \cap B$ незалежні, то

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

З іншого боку, $(A \cap B) \cap (A \cup B) = A \cap B$. Отже,

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

Прирівнюючи ці вирази, маємо:

$$P(A)P(B)P(A \cup B) = P(A)P(B); \quad P(A)P(B)(1 - P(A \cup B)) = 0;$$

$$P(A)P(B)(1 - P(A) - P(B) + P(A)P(B)) = 0; \quad P(A)P(B)(1 - P(A))(1 - P(B)) = 0,$$

звідки випливає твердження задачі.

Задача 6.1. З колоди карт у 52 листи навмання виймають одну карту. Нехай подія A означає, що витягнута карта є пікової масті, а подія B означає, що витягнута карта є туз. Чи є ці події незалежними? Відповідь обґрунтувати підрахунками.

Задача 6.2. Два рази підкидається гральний кубик. Подія A означає, що за першим разом випало 5, подія B означає, що сума очок більше 8. Чи є ці події незалежними? Відповідь обґрунтувати підрахунками.

Задача 6.3. Тричі підкидається гральний кубик. Знайти ймовірність того, що випало три одиниці, якщо відомо, що
а) на одному з кубиків випала одиниця;

б) на першому кубику випала одиниця;

в) на двох кубиках випали одиниці;

г) принаймні на двох кубиках випала однакова кількість очок;

д) на всіх кубиках випала однакова кількість очок.

Задача 6.4. Відомо, що при підкиданні п'яти монет принаймні на одній з них випав орел. Яка ймовірність того, що випало п'ять орлів?

Задача 6.5. Сім пасажирів сідають в чотири вагони метро, кожен обирає вагон навмання. Яка ймовірність того, що всі вони зайшли в один вагон, якщо відомо, що один з вагонів лишився порожнім?

Задача 6.6. Довести, що якщо події A і B незалежні, і $A \subset B$, то або $\mathrm{P}(A) = 0$, або $\mathrm{P}(B) = 1$. Отже якщо подія A не залежить сама від себе , то $\mathrm{P}(A) = 0$ або $\mathrm{P}(A) = 1$.

Задача 6.7. Спадковість незалежності подій. Довести, що якщо події A і B незалежні, то A і $\overline{B}$ теж незалежні.

Задача 6.8. Події A, B_1 та A, B_2 є незалежними, причому $B_1 \cap B_2 = \varnothing$. Довести, що події A і $B_1 \cup B_2$ теж незалежні.

Задача 6.9. Довести, що коли події A, B, C незалежні у сукупності, то події A і $B \cup C$, а також A і $B \setminus C$ – незалежні.

Задача 6.10. Дано $\mathrm{P}(A|B) = 0.7$, $\mathrm{P}\left(A|\overline{B}\right) = 0.3$, $\mathrm{P}(B|A) = 0.6$. Знайти $\mathrm{P}(A)$.

Задача 6.11. Довести, що якщо $\mathrm{P}(A) = 0.9$, $\mathrm{P}(B) = 0.8$, то $\mathrm{P}(A|B) \geq 0.875$.

Задача 6.12. Нехай $\mathrm{P}(A) = p$, $\mathrm{P}(B) = 1 - \varepsilon$, де $0 < \varepsilon < 1$. Побудувати оцінки зверху і знизу для $\mathrm{P}(A|B)$.

Задача 6.13.* Є три попарно незалежні події A, B, C, які усі три разом відбутися не можуть. Припустимо, що усі вони мають одну і ту саму ймовірність p . Визначити найбільше можливе значення p.

Задача 6.14.* Відомо, що події A і B незалежні, $A \cap B \subset C$, $\overline{C} \supset \overline{A} \cap \overline{B}$. Довести, що

$$\mathrm{P}(A \cap C) \geq \mathrm{P}(A) \cdot \mathrm{P}(C).$$

Задача 6.15. Стрілець A влучає у мішень з ймовірністю 0.6, стрілець B – з ймовірністю 0.5, стрілець C – з ймовірністю 0.4. Влучення чи промах одного стрільця не впливають на результати інших. Стрільці зробили залп по мішені. Відомо, що є два влучення. Що більш ймовірно: влучив C у мішень чи ні?

Задача 6.16. Три мисливці одночасно вистрілили по ведмедю. Відомо, що ведмедя вбито однією кулею. Яка ймовірність того, що ведмедя вбив перший, другий або третій мисливець, якщо ймовірності влучення для них дорівнюють відповідно $p_1 = 0.2$, $p_2 = 0.4$, $p_3 = 0.6$?

Задача 6.17. Ймовірності подій A, B і C дорівнюють відповідно p_1, p_2, p_3. Після проведення експерименту виявилось, що дві події відбулися, а одна ні. Довести, що при цій умові ймовірність події C буде більшою за $1/2$ тоді, коли

$$\frac{1}{p_1}+\frac{1}{p_2}-\frac{1}{p_3}>1.$$

Задача 6.18. Два однаково влучних стрільці по черзі стріляють по мішені. Кожен має право зробити не більше двох пострілів. Перший, хто влучає у мішень, отримує приз.

а) Якщо ймовірність влучення $p=1/5$, то що більш ймовірно: отримають стрільці приз чи ні?

б) Нехай p_1 і p_2 будуть ймовірності отримати приз для першого і другого стрільця відповідно. Знайти p_1/p_2, якщо $p=1/3$.

в) Яким буде це відношення, якщо число пострілів не обмежується?

Задача 6.19. Радіолокаційна станція веде спостереження за n об'єктами. За час спостереження k-ий об'єкт може бути загублений з ймовірністю p_k. Знайти ймовірність того, що: жодного з об'єктів не буде загублено; буде загублено один об'єкт; буде загублено не більш ніж один об'єкт.

Задача 6.20. Підкидають два гральних кубики. Яка ймовірність того, що принаймні один раз випаде "шістка", якщо відомо, що сума очок дорівнює восьми?

Задача 6.21. Підкидають три гральних кубики. Яка ймовірність того, що принаймні один раз випаде "шістка", якщо на усіх трьох кубиках випали різні грані?

Задача 6.22. Гральний кубик А має чотири червоних і дві білі грані, а кубик В - дві червоні і чотири білих. Один раз кидається монета. Якщо випав герб, то увесь час кидається тільки кубик А, а якщо решка – кубик В.

а) Відомо, що перші два кидки кубика дали червону грань. Яка ймовірність того, що і третій кидок також дасть червону грань?

б) Якщо перші n кидків дали червону грань, то яка ймовірність того, що кидали кубик А?

Задача 6.23. Для підвищення надійності приладу він дублюється другим таким самим приладом. Надійність кожного приладу p (надійність – це ймовірність безвідмовної роботи на деякому фіксованому інтервалі часу). При виході з ладу першого приладу відбувається миттєве підключення другого приладу.

а) Знайти надійність роботи цієї системи приладів.

б) Знайти надійність системи приладів, якщо в свою чергу пристрій підключення працює з надійністю p_1.

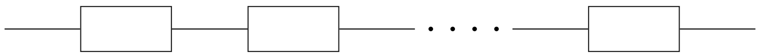
Задача 6.24. Події A_i, $i=1,2,\ldots,n$, незалежні у сукупності, причому $\mathrm{P}(A_1)=\ldots=\mathrm{P}(A_n)=p$. Знайти ймовірність того, що жодна з цих подій не відбудеться.

Задача 6.25. Для підвищення надійності приладу він дублюється n іншими приладами такого ж типу. Надійність кожного приладу дорівнює p.

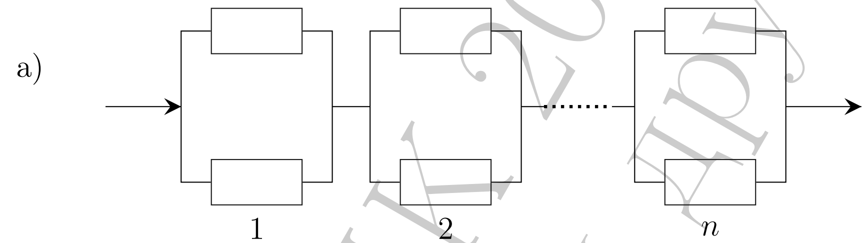
а) Знайти надійність роботи цієї системи приладів.

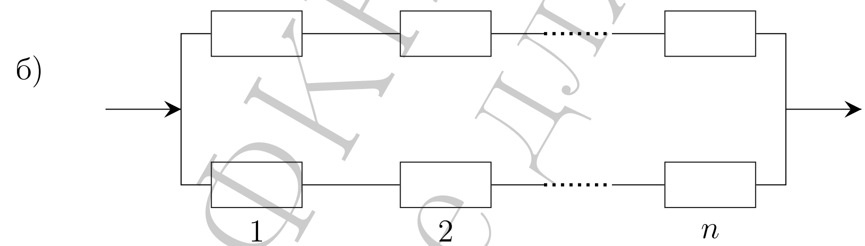
б) Знайти надійність системи приладів, якщо пристрій, який підключає дублюючий прилад, має надійність p_1.

Задача 6.26. Система складається з n елементів, з'єднаних послідовно, надійність кожного з яких дорівнює p:

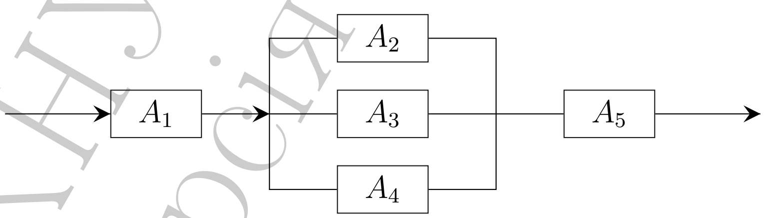


Вихід з ладу принаймні одного з блоків призводить до виходу з ладу всієї системи. Для підвищення надійності системи вона дублюється n блоками такого ж типу. Який з наведених нижче варіантів дублювання дає більш високу надійність?

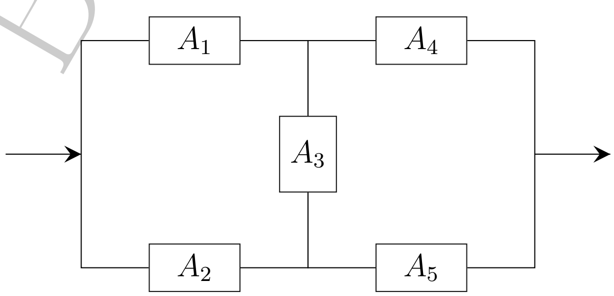




Задача 6.27. Нехай p_i – надійність елементу A_i, $i = 1, 2, \ldots, 5$. Знайти надійність наступної схеми



Задача 6.28. Нехай p_i – надійність елементу A_i, $i = 1, 2, \ldots, 5$. Знайти надійність наступної схеми



Задача 6.29. Серед семи виробів є три бракованих. Знайти ймовірність події A, яка полягає в тому, що один за одним без повернення будуть вийняті три вироби у такій послідовності: бракований – не бракований – бракований.

Задача 6.30. Знайти ймовірність влучити в мішень принаймні один раз при трьох пострілах (подія A), якщо ймовірність влучити в мішень при першому пострілі (подія A_1) становить 0.7, при другому (подія A_2) – 0.8, при третьому (подія A_3) – 0.85.

Задача 6.31.* Несиметрична монета підкидається до тих пір, доки N раз підряд не випаде герб. Знайти ймовірність того, що підкидання припиняться на парному кроці.

Задача 6.32. Двічі підкидають симетричний гральний кубик. Нехай подія A – "сума очок дорівнює 4", B – "випала принаймні одна трійка". Знайти $\mathrm{P}(A|B)$. Чи є події A та B незалежними?

Задача 6.33. З колоди, що містить 52 карти, витягують послідовно без повернення дві карти. Нехай подія A_1 – "перша карта пікової масті", A_2 – "друга карта пікової масті". Знайти $\mathrm{P}(A_2|A_1)$, $\mathrm{P}(A_2|\overline{A_1})$, $\mathrm{P}(A_2 \cap A_1)$, $\mathrm{P}(A_2 \cup A_1)$, $\mathrm{P}(A_2)$.

Задача 6.34. З урни, що містить одну білу та одну червону кулю, навмання виймають кулю. Якщо вона виявилася білою, її кладуть назад в урну. Якщо вона виявилася червоною, її кладуть в урну разом з ще двома червоними кулями. Кулі перемішують і знову виймають одну. Яка ймовірність того, що обидва рази витягнули червону кулю?

Задача 6.35. Мускатні сорти винограду, які вирощуються в Херсонській та Одеській областях, час від часу вражаються певною грибковою хворобою. Припустимо, що ймовірність ураження винограду в Херсонській області $3/4$, в Одеський – $2/5$, а ймовірність того, що враження є принаймні в одній з областей – $4/5$. Виявилось, що в певний сезон є враження винограду в Херсонській області. Яка ймовірність того, що в Одеській області виноград також уражений?

Задача 6.36. (Задача де-Мере) Скільки разів треба підкинути два гральні кубика, щоб ймовірність принаймні один раз отримати шістки була більша за $1/2$?

Задача 6.37. Події $A_1, ..., A_n$ незалежні в сукупності, $\mathrm{P}(A_k) = p_k$. Яка ймовірність того, що
а) не трапиться жодна з подій $A_1, ..., A_n$;
б) трапиться принаймні одна з подій $A_1, ..., A_n$;
в) трапиться рівно одна з подій $A_1, ..., A_n$?

Задача 6.38. При одному циклі огляду радіолокаційної станції, що веде спостереження за космічним об'єктом, об'єкт може бути виявлено з ймовірністю p. Виявлення об'єкта в кожному циклі можливе незалежно від інших. Проведено n циклів огляду. Яка ймовірність того, що об'єкт буде виявлено?

Задача 6.39. Працюють m радіолокаційних станцій, кожна з яких за один цикл огляду виявляє об'єкт з ймовірністю p незалежно від інших циклів та інших станцій. За час T кожна станція встигає провести n циклів спостережень. Знайти ймовірність таких подій: A – за час T об'єкт буде виявлено принаймні однією станцією; B – за час T об'єкт буде виявлено кожною станцією.

Задача 6.40. По мішені стріляють ракетами. Ймовірність влучення кожної ракети в мішень p, влучання окремих ракет – незалежні події. Стріляють до влучення в мішень або до використання всього боєзапасу ракет (в наявності n ракет). Яка ймовірність того, що
а) не весь боєзапас буде використано;
б) після влучення в мішень лишиться ще принаймні дві ракети;
в) буде використано не більше двох ракет?

Задача 6.41. Тричі підкидають монету. Описати простір елементарних подій. Описати події: A – двічі випав герб; B – принаймні один раз випав герб. Обчислити $\mathrm{P}(A \cup B)$, $\mathrm{P}(B)$, $\mathrm{P}(A|B)$.

Задача 6.42. Тричі підкидають монету. Описати простір елементарних подій. Описати події: A – двічі випав герб; B – принаймні один раз випав герб. Обчислити $\mathrm{P}(A \cup B)$, $\mathrm{P}(B)$, $\mathrm{P}(A|B)$.

Задача 6.43. З урни, яка містить m білих та n чорних куль, послідовно беруть 2 кулі. Відомо, що перша куля біла. Яка ймовірність того, що друга куля також біла?

Задача 6.44. Відомо, що при підкиданні 10 гральних кубиків випала принаймні одна одиниця. Яка ймовірність того, що випало 2 або більше одиниць?

Задача 6.45. Довести, що якщо A та B – несумісні події з додатними ймовірностями, то A та B – залежні.

Задача 6.46. Кидають два гральних кубика. Розглянемо випадкові події:
A_1 – на першому кубику випала парна кількість очок;
A_2 – на другому кубику випала непарна кількість очок;
A_3 – сума очок на двох кубиках непарна.
Довести, що події A_1, A_2, A_3 попарно незалежні , але не є незалежними в сукупності.

Задача 6.47. Чотири особи тестують наосліп три види пива (A, B та C) і присуджують їм ранг: 1 – для найкращого пива, потім 2 та 3. Потім ранги для кожного виду пива сумуються. Припустимо, що особи не розбираються в пиві і присуджують ранги випадковим чином.
а) Яка ймовірність того, що пиво A отримає загальну оцінку 4?
б) Яка ймовірність того, що якийсь з видів пива отримає оцінку 4?
в) Яка ймовірність того, що якийсь з видів пива отримає оцінку 5 або менше?

Задача 6.48. Двоє гравців A і B грають матч. Ймовірність того, що A переможе в одній грі, дорівнює $2/3$, ймовірність перемоги B – $1/3$. За перемогу в одній грі нараховується 1 очко, за програш – 0 очок. Для перемоги в усьому матчі гравцю A треба набрати 4 очка, а гравцю B – 2 очка. Чи рівні у гравців шанси на успіх?

Задача 6.49. Підводний човен атакує корабель, випускаючи по ньому послідовно одну за одною n торпед. Кожна торпеда незалежно від інших влучає в корабель з ймовірністю p і, при влучанні, – з однаковою ймовірністю в один з k відсіків, на які розділена підводна частина корабля. Торпеда, що потрапила у відсік, призводить до заповнення його водою. Корабель йде на дно, якщо водою заповнено принаймні два відсіка. З якою ймовірністю корабель буде затоплено?

7 ФОРМУЛА ПОВНОЇ ЙМОВІРНОСТІ ТА ФОРМУЛА БАЄСА

Нехай для події A існує послідовність подій B_i, $i = 1, 2, \dots$ таких, що $B_i \cap B_j = \varnothing$ для $i \neq j$, $\mathrm{P}(B_i) > 0$ для $i = 1, 2, \dots$, і, крім того, $A \subset \bigcup\limits_{i=1}^{\infty} B_i$. Тоді

$$\mathrm{P}(A) = \sum_{i=1}^{\infty} \mathrm{P}(B_i)\mathrm{P}(A|B_i). \tag{7.1}$$

Формула (7.1) називається **формулою повної ймовірності**.

Якщо до того ж $\mathrm{P}(A) > 0$, то

$$\mathrm{P}(B_i|A) = \frac{\mathrm{P}(B_i)\mathrm{P}(A|B_i)}{\mathrm{P}(A)} = \frac{\mathrm{P}(B_i)\mathrm{P}(A|B_i)}{\sum\limits_{j=1}^{\infty} \mathrm{P}(B_j)\mathrm{P}(A|B_j)}. \tag{7.2}$$

Формула (7.2) називається **формулою Баєса**.

Приклад 7.1

Два стрільці підкидають монету та обирають, хто з них стріляє по мішені (однією кулею). Перший стрілець влучає по мішені з ймовірністю 1, другий стрілець – з ймовірністю 0.00001. Можна зробити два висновки про експеримент: $H_1 = \{$стріляє 1-й стрілець$\}$ та $H_2 = \{$стріляє 2-й стрілець$\}$. Апріорні (a'priori – «до досліду») ймовірності цих гіпотез однакові: $\mathrm{P}(H_1) = \mathrm{P}(H_2) = 1/2$. Розглянемо подію $A = \{$куля влучила у мішень$\}$. Відомо, що $\mathrm{P}(A|H_1) = 1, \quad \mathrm{P}(A|H_2) = 0.00001$. Тому ймовірність події, що куля влучить у мішень, за формулою повної ймовірності складе

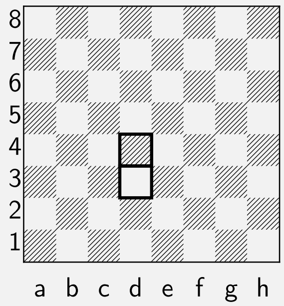
$$\mathrm{P}(A) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0.00001.$$

Припустимо що подія A сталася. Яка тепер апостеріорна (a'posteriori – «після досліду») ймовірність кожної з гіпотез H_i? Очевидно, що перша з цих гіпотез більш ймовірна за другу (а саме, у 100000 раз). Дійсно, за формулою Баєса

$$\mathrm{P}(H_1|A) = \frac{1/2 \cdot 1}{1/2 \cdot 1 + 1/2 \cdot 0.00001} = \frac{1}{1 + 0.00001} \approx 0.99999;$$

$$\mathrm{P}(H_2|A) = \frac{1/2 \cdot 0.00001}{1/2 \cdot 1 + 1/2 \cdot 0.00001} = \frac{0.00001}{1 + 0.00001} \approx 0.0000099999.$$

Приклад 7.2



На шаховій дошці навмання обирають два поля (дві клітинки). Яка ймовірність, що це будуть два сусідні поля (дві клітинки, що мають спільну сторону)?

Розв'язок: розглянемо такі три гіпотези:

$H_1 =$\{першим обрано одне з 4 кутових полів a1, a8, h1, h8\},

$H_2 =$\{першим обрано одне з 24 граничних не-кутових полів a2, a3,… a7; b8, c8,… , g8; h7, h6, … , h2; g1, f1, … , b1\},

$H_3 =$\{першим обрано одне з решти 36 не-граничних полів\}.

Для події $A =$\{обрані поля будуть сусідніми\} нескладно порахувати за класичним визначенням ймовірності наступні значення: $\mathrm{P}(H_1) = \frac{4}{64} = \frac{1}{16}$, $\mathrm{P}(H_2) = \frac{24}{64} = \frac{3}{8}$, $\mathrm{P}(H_3) = \frac{36}{64} = \frac{9}{16}$, $\mathrm{P}(A|H_1) = \frac{2}{64-1} = \frac{2}{63}$, $\mathrm{P}(A|H_2) = \frac{3}{64-1} = \frac{1}{21}$, $\mathrm{P}(A|H_3) = \frac{4}{64-1} = \frac{4}{63}$, звідки за формулою повної ймовірності маємо

$$\mathrm{P}(A) = \sum_{k=1}^{3} (A|H_k)\,\mathrm{P}(H_k) = \frac{2}{63} \cdot \frac{1}{16} + \frac{1}{21} \cdot \frac{3}{8} + \frac{4}{63} \cdot \frac{9}{16} = \frac{1}{18} \approx 5.6\ \%.$$

Відповідь: $\dfrac{1}{18} \approx 5.6\ \%$.

Приклад 7.3

Авто експлуатується двома особами: чоловіком та жінкою по черзі. Ймовірність дорожньо-транспортної пригоди при керуванні авто чоловіком є 0.1, а ймовірність дорожньо-транспортної пригоди при керуванні авто жінкою є 0.03. Чоловік користується транспортним засобом удвічі більше ніж жінка.

а) Підрахувати ймовірність дорожньо-транспортної пригоди для даного авто;

б) Відомо, що сталася дорожня пригода. Підрахувати ймовірність того, що за кермом був чоловік.

Розв'язок:

Нехай A, A_1, A_2 є наступними подіями: A – сталася дорожня пригода, A_1 – за кермом був чоловік, A_2 – за кермом була жінка. Використовуючи умову задачі, можемо записати, що

$$\mathrm{P}(A_1) = \frac{2}{3}, \qquad \mathrm{P}(A_2) = \frac{1}{3}.$$

А також: $\mathrm{P}(A|A_1) \ = \frac{10}{100} = 0.1$, $\mathrm{P}(A|A_2) = \frac{3}{100} = 0.03$.

а) Оскільки $A \subset A_1 \cup A_2$, то на підставі формули повної ймовірності маємо:

$$\mathrm{P}(A) = \mathrm{P}(A_1)\mathrm{P}(A|A_1) + \mathrm{P}(A_2)\mathrm{P}(A|A_2) == \frac{2}{3} \cdot \frac{1}{10} + \frac{1}{3} \cdot \frac{3}{100} = \frac{23}{300}.$$

б) За формулою Баєса маємо:

$$\mathrm{P}(A_1|A) = \frac{\mathrm{P}(A_1)\mathrm{P}(A|A_1)}{\mathrm{P}(A)} = \frac{\frac{2}{3} \cdot \frac{1}{10}}{23/300} = \frac{20}{23}.$$

Відповідь: а) $\dfrac{23}{300}$, б) $\dfrac{20}{23}$.

Приклад 7.4

В 2001 році Європейська комісія запровадила масові перевірки крупної рогатої худоби на наявність форми коров'ячого сказу. Звичайно, тести не дають 100 % точності, практично всі вони з певною ймовірністю дають хибнопозитивний (здорова корова визнана хворою) або хибнонегативний результат (хвора корова визнана здоровою). Для визначення точності тесту Європейська комісія проводила дослідження, аналізуючи зразки гарантовано хворих та гарантовано здорових корів. В результаті було заявлено, що один з тестів визнавав хворими 70 % інфікованих корів та 10 % здорових корів.

Нехай подія A – корова хвора, B – тест визначив, що корова хвора. Тоді ймовірності хибнопозитивного та хибнонегативного результатів

$$\mathrm{P}(B|A) = 0.7, \quad \mathrm{P}(B|\overline{A}) = 0.1.$$

Припустимо, в певному регіоні доля корів, які є носіями хвороби, є 0.01 (цифра взята для зручності підрахунків. Справжні цифри менші і є різними для різних країн, наприклад для Голландії орієнтовна доля хворих корів в 2003 році складала 0.000013). Визначимо ймовірність того, що корова буде визнана хворою в результаті перевірки. Ця ймовірність залежить від того, чи є корова дійсно хворою, чи ні. Згідно формулі повної ймовірності маємо

$$\mathrm{P}(B) = \mathrm{P}(B|A)\mathrm{P}(A) + \mathrm{P}(B|\overline{A})\mathrm{P}(\overline{A}) = 0.7 \cdot 0.01 + 0.1 \cdot (1 - 0.01) = 0.106.$$

Припустимо тепер, що корова визнана хворою в результаті тестування. Яка ймовірність того, що вона дійсно є хворою? Скористаємось формулою Баєса:

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(B|A)\mathrm{P}(A)}{\mathrm{P}(B|A)\mathrm{P}(A) + \mathrm{P}(B|\overline{A})\mathrm{P}(\overline{A})} = \frac{0.7 \cdot 0.01}{0.7 \cdot 0.01 + 0.1 \cdot (1 - 0.01)} \approx 0.066.$$

Аналогічно можна знайти ймовірність того, що корова хвора, якщо вона визнана здоровою

$$\mathrm{P}(A|\overline{B}) = \frac{\mathrm{P}(\overline{B}|A)\mathrm{P}(A)}{\mathrm{P}(\overline{B}|A)\mathrm{P}(A) + \mathrm{P}(\overline{B}|\overline{A})\mathrm{P}(\overline{A})} = \frac{(1 - 0.7) \cdot 0.01}{(1 - 0.7) \cdot 0.01 + (1 - 0.1) \cdot (1 - 0.01)} \approx 0.0034.$$

Ці обчислення демонструють, що цей тест не є якісним. Було б добре, щоб ймовірність $\mathrm{P}(A|B)$ була близькою до 1.

Отримані значення ймовірностей ми можемо прокоментувати так: якщо про якусь випадкову корову нічого не відомо, то ймовірність того, що вона є носієм хвороби, 0.01. Якщо ж проведено тест, який виявився позитивним, ймовірність того, що корова дійсно є носієм хвороби, 0.066. Проте, якщо тест дав негативний результат, то ймовірність цієї ж події лише 0.0034. З цього ми зокрема бачимо, що ймовірність події A залежить від того, чи трапилась подія B.

Приклад 7.5 Л.Керол, Заплутана історія: A Tangled Tale (1880-1885)

В урні знаходиться одна куля, про яку відомо, що вона або біла, або чорна. В урну поклали білу кулю, а потім, після ретельного перемішування, вийняли навмання одну кулю, яка виявилася білою. Яка ймовірність того, що після цього з урни виймуть білу кулю?

Розв'язок: Стохастичний експеримент в даному випадку полягає у витягуванні двох куль з урни, що містить після описаних процедур дві кулі. Причому відомо, що одна з них точно біла, друга – або біла, або чорна (в таких випадках, коли немає підстав вважати інакше, вважають, що ймовірність того, що куля була білою, дорівнює 1/2).

Опишемо події:

$$A = \{\text{перша витягнута куля – біла}\}; \quad B = \{\text{друга витягнута куля – біла}\};$$

$$B \cap A = \{\text{обидві витягнуті кулі – білі}\}.$$

Згідно формулі умовної ймовірності маємо (знаменник обчислюємо за формулою повної ймовірності, вводячи гіпотези про колір кулі, що лежала в урні до початку експерименту)

$$\mathrm{P}(B|A) = \frac{\mathrm{P}(B \cap A)}{\mathrm{P}(A)} = \frac{1/2}{1 \cdot 1/2 + 1/2 \cdot 1/2} = \frac{2}{3}.$$

Приклад 7.6

Студент здає іспит у вигляді тесту, кожне питання якого має п'ять альтернативних відповідей. На частину питань студент знає відповідь, на частину – ні. Якщо студент знає відповідь на питання, він відповідає правильно, якщо не знає – обирає відповідь випадковим чином. Викладач, який спостерігає за студентом, бачить, що на певне питання студент дав правильну відповідь, і йому цікаво, чи студент дійсно знав відповідь, чи просто вгадав. Нехай p – доля питань, на які студент знав відповіді, і припустимо, що на ті питання, які студент не знає, він дає вірну відповідь з ймовірністю $1/5$ (насправді ситуація дещо краща, оскільки часто студент може, принаймні, виявити гарантовано неправильні відповіді і ймовірність вгадати правильну буде вища). Нехай подія A – студент відповів правильно на певне питання, що цікавить викладача, подія H_1 (гіпотеза викладача) – це те, що студент знав відповідь на питання, H_2 – вгадав. Нас цікавить $\mathrm{P}(H_1|A)$.

Розв'язок: Згідно формулі Баєса,

$$\mathrm{P}(H_1|A) = \frac{\mathrm{P}(A|H_1)\mathrm{P}(H_1)}{\mathrm{P}(A|H_1)\mathrm{P}(H_1) + \mathrm{P}(A|H_1)\mathrm{P}(H_1)} = \frac{1 \cdot p}{1 \cdot p + \frac{1}{5}(1-p)}.$$

Зауважимо, що

$$\frac{p}{p + \frac{1}{5}(1-p)} \geq p.$$

Для значення $p = 0.75$ маємо

$$\mathrm{P}(H_1|A) = \frac{0.75}{0.75 + 0.25/5} = 0.9375,$$

а для значення $p = 0.25$

$$\mathrm{P}(H_1|A) = \frac{0.25}{0.25 + 0.75/5} = 0.625.$$

Тобто навіть якщо студент знає лише чверть матеріалу, якщо він відповів правильно, ймовірність того, що він знає питання, доволі висока.

Задача 7.1. Серед N екзаменаційних білетів є n "щасливих". За білетами студенти підходять один за одним. У кого більша ймовірність узяти щасливий білет: у того, хто підійшов першим, чи у того, хто підійшов другим?

Задача 7.2. Страхова компанія займається страхуванням життя. 30 % застрахованих в цій компанії палять. Якщо застрахований не палить, ймовірність його смерті упродовж року дорівнює 0.01. Якщо ж він палить, то ця ймовірність дорівнює 0.05. Яка частина курців серед тих застрахованих, що померли упродовж року?

Задача 7.3. Під час рентгенівського обстеження ймовірність виявити захворювання у хворого на туберкульоз дорівнює $1 - b$. Ймовірність визнати здорову людину хворою дорівнює a. Нехай частина хворих на туберкульоз відносно всього населення становить r. Знайти ймовірність того, що людина здорова, якщо вона визначена хворою.

Задача 7.4. В урні n куль, білі або чорні. Усі припущення про число білих куль в урні рівноможливі. З урни навмання беруть одну кулю. Яка ймовірність того, що вона біла?

Задача 7.5. В першій урні знаходиться 4 білих кулі і 6 чорних куль, у другій урні – 7 білих куль і 8 чорних куль. З кожної урни навмання виймають по одній кулі і перекладають їх у третю урну. Потім з третьої урни виймають одну кулю.
а) Яка ймовірність того, що вона біла?
б) Нехай відомо, що на другому етапі експерименту з третьої урни витягли білу кулю. Яка ймовірність того, що вона належить першій урні?

Задача 7.6. На шаховій дошці навмання обирають два поля (два квадратики). Яка ймовірність, що це будуть два поля, які мають спільний кут (два квадратики, що мають спільну точку)? Такими є, наприклад, поля b4 та c5. `Примітка: див. приклад 7.2 на стор. 61`

Задача 7.7. Відбулось 3 випробування, в кожному з яких подія A відбувається з ймовірністю 0.2. Ймовірність появи другої події B залежить від числа появ події A: при одноразовій появі події A ця ймовірність дорівнює 0.1, при двоpазовому – 0.3, при тpиpазовому – 0.7; якщо подія A не відбулась, то подія B неможлива. Знайти найбільш ймовірне число появ події A, якщо подія B відбулась.

Задача 7.8. Дві фабрики виробляють деяку продукцію. Частина браку на першій фабриці складає 3 % , а на другій - 5 % . Навмання обрано фабрику і придбано 100 одиниць продукції. Яка ймовірність того, що серед 100 виробів буде 2 бракованих?

Задача 7.9. У двох урнах знаходиться n_1 і n_2 куль відповідно. У першій урні m_1 білих куль і $n_1 - m_1$ чорних куль, у другій – m_2 і $n_2 - m_2$. З першої урни переклали у другу одну кулю, колір якої невідомий. Після цього з другої урни беруть одну кулю. Яка ймовірність, що ця куля біла?

Задача 7.10. Із чисел $\{1, 2, \ldots, n\}$ навмання одне за одним обирають два числа. Яка ймовірність того, що різниця між першим обраним числом і другим буде не менша m $(0 < m < n)$?

Задача 7.11. З колоди карт у 32 листа навмання одна за одною без повернення витягли дві карти. Яка ймовірність того, що другою картою можна накрити першу?

Задача 7.12. У ящику знаходиться 15 тенісних м'ячів, з яких 9 є новими. Для першої гри навмання беруть 3 м'ячі, які після цього повертають у ящик. Для другої гри також навмання беруть 3 м'ячі. Знайти ймовірність того, що усі м'ячі, які узяли для другої гри, є новими.

Задача 7.13. Партія транзисторів, серед яких 10 % дефектних, надійшла на перевірку. Схема перевірки така, що з ймовірністю 0.95 виявляється дефект (якщо він є) і з імовірністю 0.03 справний транзистор буде признаний дефектним. Яка ймовірність того, що навмання обраний з партії транзистор буде признаний дефектним?

Задача 7.14. В партії з N виробів є M дефектних. Для контролю з цієї партії навмання обрано $n < N$ виробів. При перевірці можливі похибки двох типів:
а) дефектний виріб визнається справним (ймовірність цього p_1);
б) справний виріб з ймовірністю p_2 визнається дефектним.
Знайти ймовірність того, що серед n обраних виробів m виробів будуть визнані дефектними.

Задача 7.15. Для пошуку родовища нафти на заданій території організовано n геологорозвідувальних партій, кожна з яких незалежно від інших знаходить поклад з ймовірністю p. Після аналізу додаткової інформації уся територія поділена на два райони, причому у першому районі родовище нафти може бути з ймовірністю p_1, а у другому - з ймовірністю $1 - p_1$. Як слід розподілити n партій між двома районами так, щоб ймовірність знаходження нафти була максимальною?

Задача 7.16. Маємо дві монети – справжню і фальшиву. Фальшива монета випадає гербом у два рази частіше, ніж решіткою. Підкидаємо навмання обрану монету. Вона випала гербом. Яка ймовірність того, що ця монета фальшива?

Задача 7.17. Деталі виробляють на двох заводах. Об'єм продукції на другому заводі у n раз більший, ніж на першому. Частина браку на першому заводі – p_1, на другому заводі – p_2. Навмання обрана деталь виявилась бракованою. Яка ймовірність того, що вона зроблена на другому заводі?

Задача 7.18. Відомо, що 5 % усіх чоловіків і 0.25 % усіх жінок є дальтоніками. Навмання обрана особа виявилась дальтоніком. Яка ймовірність того, що це чоловік? Вважати, що чоловіків та жінок однакова кількість.

Задача 7.19. Урна містить n куль. Усі припущення про число білих куль в урні рівноймовірні. Навмання обрана з урни куля виявилася білою. Обчислити ймовірності усіх припущень про склад куль в урні. Яке припущення найбільш ймовірне?

Задача 7.20. В урні 30 куль – білі і чорні. Всі припущення про склад куль рівноможливі. Із урни навмання обрали 3 кулі, вони виявилися: одна – білою та дві – чорними.
а) Яка ймовірність саме таких трьох витягнутих куль?
б) Яке припущення про кількість білих куль в урні найімовірніше?

Задача 7.21. З урни, яка містить m білих куль і n чорних куль, загубили одну кулю. Для того, щоб визначити склад куль в урні, з урни узяли дві кулі. Вони виявилися білими. Обчислити ймовірність того, що загублена куля була білою.

Задача 7.22. В урні знаходиться одна куля, про яку відомо, що вона або біла або чорна. В цю урну кладуть білу кулю і після перемішування навмання виймають одну кулю. Вона виявляється білою. Яка ймовірність того, що в урні залишилася біла куля?

Задача 7.23. З 18 стрільців п'ять влучають у мішень з ймовірністю 0.8, сім влучають з ймовірністю 0.7 і два стрільці з ймовірністю 0.5. Навмання обраний стрілець зробив постріл, але у мішень не влучив. До якої групи найбільш ймовірно він належить?

Задача 7.24. У рибалки є три місця для ловлі риби, які він відвідує рівноймовірно. Відомо, що якщо він закидає вудку на першому місці, то риба клює з ймовірністю p_1, на другому – з ймовірністю p_2, на третьому – з ймовірністю p_3. На ловлі риби рибак три рази закинув вудку, а риба клюнула тільки один раз. Треба знайти ймовірність того, що це трапилося на першому місці.

Задача 7.25. Імовірність поразки команди ДК у матчі з командою ЮМ при дощовій погоді становить 0.5, а при відсутності дощу – 0.6. Імовірність дощу у день матчу становить 0.2.
а) Знайти ймовірність уникнення поразки командою ДК.
б) Відомо, що команда ДК зазнала поразки. Яка ймовірність того, що матч відбувався при дощовій погоді?

Задача 7.26. Нехай деяка комаха з ймовірністю $\dfrac{e^{-\lambda}\lambda^k}{k!}$ кладе k яєць ($\lambda > 0$), а ймовірність розвитку комахи з яйця дорівнює p. Припускаючи взаємну незалежність розвитку яєць, знайти ймовірність того, що у комахи буде рівно n нащадків.

Задача 7.27. Перша урна містить три червоних, дві білих та одну синю кулю, друга – одну червону, дві білих і три синіх кулі. Випадковим чином з кожної урни витягнули по одній кулі.
а) Знайти ймовірність того, що обидві кулі одного кольору;

б) що більш ймовірно: мати в результаті експерименту дві червоні чи дві білі кулі?
в) кулі, які витягли в результаті експерименту, поклали в третю урну, перемішали і потім навмання витягли по одній кулі з усіх трьох урн. Знайти ймовірність того, що будуть витягнуті кулі всіх трьох кольорів.

Задача 7.28. В певному регіоні 0.1 % людей хворіють на туберкульоз. Тест на туберкульоз виявляє туберкульоз з ймовірністю 0.999. Якщо людина здорова, то тест дасть хибнопозитивний результат з ймовірністю 0.02. Для випадково обраної людини тест дав позитивний результат. Яка ймовірність того, що людина дійсно хвора на туберкульоз?

Задача 7.29.* Чотири злодія, A, B, C та D, на допиті кажуть правду з ймовірністю $1/3$. Злодій A зробив заяву, що він не винен в злочині, що розслідується, після чого D сказав, що C сказав, що B сказав, що A сказав правду. Яка ймовірність того, що A дійсно не винен?

Задача 7.30. В штучному лабіринті піддослідній миші дається вибір піти ліворуч і отримати їжу, чи піти праворуч, де вона отримає слабкий електричний удар. Нехай до початку всіх експериментів миша з однаковою ймовірністю може піти ліворуч чи праворуч. Після отримання їжі в одному експерименті, ймовірність піти ліворуч та праворуч в наступному експерименті стає 0.6 та 0.4, Після отримання розряду ймовірності піти ліворуч та праворуч в наступному експерименті стають 0.8 та 0.2 відповідно. Яка ймовірність того, що миша піде ліворуч в другому експерименті? В третьому експерименті?

Задача 7.31. Перша урна містить дві білих та три чорних кулі, друга – шість білих та чотири чорних. З ймовірністю p обирається перша урна або з ймовірністю $1 - p$ – друга, з обраної урни навмання виймається куля. За якого значення p ймовірність вийняти в такий спосіб чорну кулю буде такою ж, як вийняти чорну кулю з однієї урни, де змішані всі наявні кулі (6 білих та 7 чорних)?

Задача 7.32. 5 % людей має підвищений кров'яний тиск. Серед тих, хто має підвищений тиск, 75 % вживають алкоголь, а серед тих, хто не має високого тиску, алкоголь вживають 50 % . Який відсоток тих, хто вживає алкоголь, має високий тиск?

Задача 7.33. Особа, яка займається продажем насіння сортових гарбузів, визначила, що 4 % насіння не проросте. Вона продає насіння в пакетах по 50 насінин та гарантує 90 % проростання. Яка ймовірність того, що для випадково обраного пакетика буде порушена гарантія?

Задача 7.34. Пан Багатько, заможний продавець діамантів, хоче нагородити свого сина, запропонувавши йому на вибір дві коробочки, в кожній з яких міститься три каменя. В першій коробочці – два справжніх діаманти, і одна нічого не варта підробка, в другій – один діамант і два підробних каменя. Син обирає коробочку навмання, причому він має можливість вийняти випадковим чином з обраної коробочки один камінь і перевірити, чи він є справжнім діамантом. Син вирішив, що якщо перевірений камінь виявиться справжнім, то він обере ту коробочку, з якої він її витягнув. Якщо камінь виявиться підробкою, він обере собі іншу коробочку. Яка ймовірність того, що він обере коробочку з двома діамантами?

Задача 7.35. Умовна ймовірність того, що у подружжя народиться хлопчик, дорівнює $\frac{1}{2} + m\varepsilon_1 - f\varepsilon_2$, де ε_1 та ε_2 – деякі малі константи, m – кількість хлопців, а f – кількість дівчат, що вже народилися у подружжя. Знайти ймовірність того, що

а) третя дитина буде хлопчиком, якщо відомо, що перші дві – дівчата;

б) перші три дитини будуть хлопцями;

в) серед перших трьох дітей буде принаймні один хлопчик.

Задача 7.36. Виробник приладів, які перевіряють якість певних компонент, заявляє, що прилад має високу надійність і ймовірність хибнонегативного і хибнопозитивного результату – лише $p = 0.05$. Покупець планує перевірити за допомогою цього приладу велику партію, в якій 5 % браку.

а) Яка ймовірність того, що компонент бракований, якщо прилад визнав його бракованим?

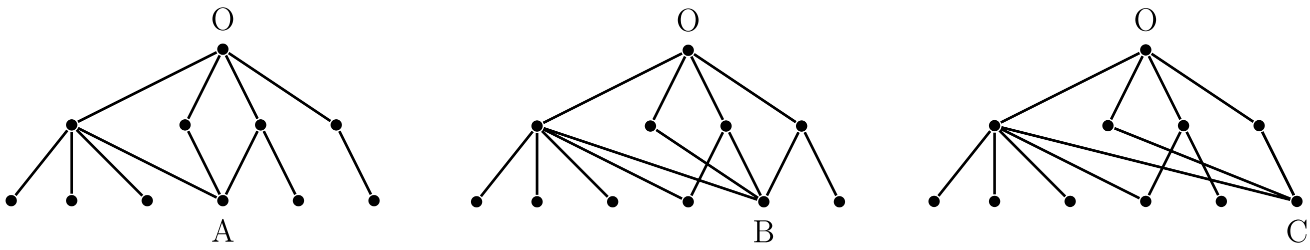
б) Якщо покупець хоче, щоб ймовірність в пункті а) була 0.9, якою повинна бути ймовірність хибнопозитивного та хибнонегативного результату p приладу?

Задача 7.37. На шахову дошку в різні клітини ставлять навмання чорного та білого короля. Яка ймовірність того, що при цьому вийде допустима ситуація (тобто королі не стоять в сусідніх клітинах, зокрема по діагоналі)?

Задача 7.38. Дві станції надсилають сигнал в спільний канал зв'язку, причому перша з них посилає вдвічі більше сигналів, ніж друга. Ймовірність отримати спотворений сигнал з першої станції дорівнює 0.05, з другої – 0.02. Яка ймовірність отримати спотворений сигнал в загальному каналі зв'язку?

Задача 7.39. Ймовірність правильного виявлення наявності шкідливих речовин в продукті для кожного з контролерів дорівнює 4/5, 3/4 та 2/3. При одночасному контролі трьох проб трьома контролерами правильно виявлено шкідливі речовини у двох пробах. Яка ймовірність того, що помилився третій контролер?

Задача 7.40. Тато поклав до однієї коробки 5 чорних кульок та 8 білих, а до другої коробки – 19 чорних кульок та 12 білих. За таких умов він пропонує сину обрати навмання коробку, а з неї навмання обрати кульку. Якщо кулька виявиться білою, то посуд митиме син, а якщо чорна, то посуд митиме тато. У скільки разів син міг би підвищити свої шанси на "перемогу", якби йому дозволили до початку гри розкласти усі кульки у ці дві урни на свій розсуд?



(а) (б) (в)

Рис. 7.1: до задачі 7.41

Задача 7.41. Мандрівник вирушає в подорож з пункту O і на кожному роздоріжжі обирає навмання один з можливих шляхів (окрім зворотного). Яка ймовірність того, що він потрапить в пункт

а) A (див. рис. 7.1 (а))? б) B (див. рис. 7.1 (б))? в) C (див. рис. 7.1 (в))?

Задача 7.42. Урна містить n куль, всі припущення про кількість білих куль в урні рівноможливі. Навмання обрана з урни куля виявилася білою. Обчислити ймовірність всіх припущень про склад куль в урні.

Задача 7.43. З урни, яка містить n куль невідомого кольору, вийняли одну кулю, яка виявилася білою. Після цього виймають ще одну кулю. Яка ймовірність того, що вона також біла (Всі припущення про початкову кількість білих куль в урні рівноможливі)?

Задача 7.44. На вхід радіолокаційного приладу надходить сигнал з шумом з ймовірністю p та лише шум з ймовірністю $1-p$. Якщо надходить сигнал з шумом, то прилад реєструє наявність сигналу з ймовірністю p_1, якщо надходить лише шум, то прилад помилково реєструє наявність сигналу з ймовірністю p_2. Відомо, що прилад зареєстрував сигнал. Яка ймовірність того, що сигнал надійшов на вхід приладу?

Задача 7.45. Відомо, що 34 % людей має першу групу крові, 37 % – другу, 21 % – третю і 8 % – четверту. Хворому з першою групою можна переливати лише кров першої групи, з другою – кров першої та другої груп, з третьою – кров першої та третьої груп, з четвертою можна переливати кров будь-якої групи. Яка ймовірність того, що довільно взятому хворому можна переливати кров довільно обраного донора?

8 СХЕМА НЕЗАЛЕЖНИХ ВИПРОБУВАНЬ БЕРНУЛЛІ

Уявімо, що проводяться n незалежних випробувань (дослідів), результатом кожного з яких може бути або успіх (з ймовірністю $0 < p < 1$), або невдача (з ймовірністю $q = 1 - p$). Часто настання успіху позначають одиничкою, а невдачу – нулем, тоді простір елементарних подій Ω такого експерименту є n-мірний вектор з нулів та одиниць. Очевидно, що $|\Omega| = 2^n$, а оскільки множина Ω є скінченною, то за σ-алгебру $\Re$ беруть множину всіх її підмножин (булеан), а ймовірність визначають так: $\forall A \in \Re$

$$\mathrm{P}(A) = \sum_{\omega \in A} \mathrm{P}(\omega) = \sum_{\omega \in A} p^{\sum\limits_{k=1}^{n} \omega_k} q^{n - \sum\limits_{k=1}^{n} \omega_k},$$

бо кожен елементарний наслідок ω складається саме з $\sum\limits_{k=1}^{n} \omega_k$ успіхів (одиничок) і саме з $n - \sum\limits_{k=1}^{n} \omega_k$ невдач (нулів). Побудовану таким чином скінченну ймовірнісну схему $(\Omega, \Re, \mathrm{P})$ називають **схемою незалежних випробувань Бернуллі з параметрами n та p.**

Позначимо через $\mathrm{P}_{n,p}(m)$ ймовірність того, що в цій схемі відбулося рівно m успіхів (це так звана **біноміальна ймовірність**). Відомо, що:

$$\mathrm{P}_{n,p}(m) = C_n^m p^m (1 - p)^{n-m}, \qquad 0 \le m \le n. \tag{8.1}$$

У випадках, коли кількість випробувань в схемі Бернуллі n дуже велика (порядку десятків, сотень, тисяч), рахувати ймовірності отримати певну кількість успіхів за формулою (8.1) дуже незручно. Тому виникає необхідність в отриманні зручних наближених формул для обчислення цих ймовірностей для великих n.

Розглянемо два наближення для формули (8.1) в різних припущеннях щодо значень n і p, задані теоремою Пуассона і локальною та інтегральною теоремами Муавра-Лапласа.

Теорема 8.1 (Пуассона) Якщо $n \to \infty$, $p = p(n) \to 0$, $n \cdot p(n) \to \lambda > 0$, то

$$\mathrm{P}_{n,p}(m) \xrightarrow[n \to \infty]{} e^{-\lambda} \frac{\lambda^m}{m!}$$

для кожного $m \ge 0$.

Ця теорема (її ще називають законом рідкісних подій) дає змогу наближено підраховувати біноміальну ймовірність для великих n та малих p.

Зауваження. Щоб похибка в результаті заміни $\mathrm{P}\{\xi_n = k\} = C_n^k p^k q^{n-k}$ на $\frac{\lambda^k}{k!}e^{-\lambda}$ була незначною, n має бути не менше за кілька десятків, а краще сотень, а добуток $\lambda = np$ не повинен перевищувати 10. При інших значеннях λ варто застосовувати локальну або інтегральну теорему Муавра-Лапласа.

Теореми Муавра-Лапласа стверджують, що дискретний біноміальний розподіл може бути наближений неперервним нормальним розподілом при великих значеннях n.

Нехай ν_n позначає кількість успіхів у схемі Бернуллі з параметрами n та p.

Теорема 8.2 (Інтегральна теорема Муавра-Лапласа) Для довільних $a < b$ справедливе співвідношення

$$\lim_{n\to\infty} \mathrm{P}\left(a < \frac{\nu_n - np}{\sqrt{npq}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{y^2}{2}}dy.$$

З цієї теореми випливає наступна формула для обчислення ймовірності кількості успіхів у схемі Бернуллі

$$\mathrm{P}\left(A \le \nu_n \le B\right) \approx \Phi\left(\frac{B - np}{\sqrt{npq}}\right) - \Phi\left(\frac{A - np}{\sqrt{npq}}\right), \tag{8.2}$$

де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2}}dy$, $x \in (-\infty, \infty)$, – функція розподілу стандартного нормального закону, значення якої можна знайти в табл. 3 на стор. 302. Таблиці значень функції Φ даються зазвичай лише для додатних значень x, які не перевищують значення 3. Значення $\Phi(x)$ для $x > 3$ можна вважати рівними 1. Для знаходження значень Φ для від'ємного аргумента, варто скористатися симетричністю та нормованістю цієї функції, внаслідок яких $\Phi(-x) = 1 - \Phi(x)$, $x > 0$.

Порядок похибки в формулі (8.2) є $O(1/\sqrt{npq})$.

Теорема 8.3 (Локальна теорема Муавра-Лапласа) Нехай $n \to \infty$ та $k \to \infty$ так, що $\dfrac{k - np}{\sqrt{npq}} \to y$. Тоді справедливе наступне співвідношення

$$\lim_{\substack{n \to \infty \\ k \to \infty}} \sqrt{npq}\,\mathrm{P}\left(\nu_n = k\right) = \frac{1}{\sqrt{2\pi}} e^{-\dfrac{y^2}{2}}.$$

З цієї теореми як наслідок маємо наступну наближену формулу для обчислення ймовірності кількості успіхів у схемі Бернуллі для великих n, k:

$$\mathrm{P}\left(\nu_n = k\right) \approx \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-(k-np)^2 / 2npq},$$

причому похибка є теж порядку $O(1/\sqrt{npq})$.

Зауваження. Наближені формули, які фігурують в теоремах Муавра-Лапласа доцільно використовувати (теореми дають задовільне наближення) при $n > 100$, $np > 20$, $npq > 9$ (локальна) ($np > 5$ та $n(1-p) > 5$ – інтегральна). Причому чим ближче p та q до $1/2$, тим точніші дані формули. При великих або малих значеннях ймовірності p (близьких до 0 або 1) формули дають велику похибку (порівняно з базовою формулою Бернуллі). В цих випадках краще використовувати наближення Пуассона.

Приклад 8.1

Скільки випробувань Бернуллі з ймовірністю успіху $p = 0.01$ треба провести, щоб ймовірність хоча б одного успіху була не менша $1/2$?

Розв'язок:

Оскільки ймовірність того, що відбудеться хоча б один успіх є $1 - \mathrm{P}_{n,p}(0)$, то за умовою задачі ми шукаємо таке n, для якого $1 - \mathrm{P}_{n,p}(0) \geq \frac{1}{2}$, тобто: $1 - (0.99)^n \geq 1/2$, $(0.99)^n \leq 1/2$, $n \geq \log_{0.99} 0.5 \approx 68.96756$. Враховуючи, що шукане n є натуральним числом, маємо таку **відповідь:** $n \geq 69$.

Приклад 8.2

У певного студента ймовірність здати кожен з п'яти іспитів сесії дорівнює 0.7 незалежно від результатів здачі інших іспитів. Яка ймовірність того, що
а) студент здасть рівно три іспити;
б) студент здасть хоча б два іспити.

Розв'язок: Маємо схему незалежних випробувань Бернуллі, одне випробування полягає в успішній здачі іспиту. Відповідно $n = 5$, $p = 0.7$, $q = 1 - p = 0.3$, випадкова величина ξ_5 – кількість успішно зданих іспитів упродовж сесії. Скористаємось формулою (8.1):

а) $\mathrm{P}\{\xi_5 = 3\} = C_5^3 0.7^3 0.3^2 \approx 0.3$;

б) $\mathrm{P}\{2 \leq \xi_5 \leq 5\} = 1 - \mathrm{P}\{\xi_5 \leq 1\} = 1 - C_5^0 0.7^0 0.3^5 - C_5^1 0.7^1 0.3^4 \approx 0.97$.

Приклад 8.3

Іван Гуляйполе є постійним клієнтом корчми. Він пропонує іншим клієнтам таку гру: клієнт вибирає число від 1 до 6, і симетричний гральний кубик підкидається тричі. Якщо число, яке назвав клієнт, жодного разу не випаде, клієнт платить Івану 100 грн. Якщо вказане число випадає один, два чи три рази, Іван платить клієнту 100, 200 чи 300 грн відповідно. Наскільки вигідною є ця гра для Івана?

Розв'язок: На перший погляд ця гра є більш вигідною для клієнта. Порахуємо.

Маємо схему незалежних випробувань Бернуллі з трьома випробуваннями ($n = 3$) та ймовірністю успіху в одному випробуванні $p = 1/6$. Отже, можемо знайти ймовірність того, що вказане число випаде k разів ($k = 0, 1, 2, 3$):

$$\mathrm{P}\{\xi_3 = k\} = C_3^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{3-k}.$$

Звідси можна підрахувати математичне сподівання[a] Іванового виграшу:

$$100 \cdot \mathrm{P}\{\xi_3 = 0\} - 100 \cdot \mathrm{P}\{\xi_3 = 1\} - 200 \cdot \mathrm{P}\{\xi_3 = 2\} - 300 \cdot \mathrm{P}\{\xi_3 = 3\} =$$
$$= 100 \cdot \frac{125}{216} - 100 \cdot 3 \cdot \frac{25}{216} - 200 \cdot 3 \cdot \frac{5}{216} - 300 \cdot \frac{1}{216} = 100 \cdot \frac{17}{216} = 7.87 \text{ грн.}$$

[a] Визначення математичного сподівання див. на ст. 84.

Приклад 8.4

В схемі незалежних випробувань Бернуллі з параметрами n та p знайти найбільш імовірну кількість успіхів.

Розв'язок: Потрібно знайти k, для якого ймовірність $\mathrm{P}_{n,p}(k)$ досягає свого найбільшого значення. Із очевидних рівностей

$$\frac{\mathrm{P}_{n,p}(k)}{\mathrm{P}_{n,p}(k-1)} = \frac{p}{(1-p)} \frac{n-k+1}{k} = 1 + \frac{(n+1)\,p - k}{k\,(1-p)}$$

маємо: якщо $k < (n+1)\,p$ то

$$\mathrm{P}_{n,p}(k) > \mathrm{P}_{n,p}(k-1)$$

(з ростом k ймовірності $\mathrm{P}_{n,p}(k)$ збільшуються), а якщо $k > (n+1)\,p$, то

$$\mathrm{P}_{n,p}(k) < \mathrm{P}_{n,p}(k-1)$$

(з ростом k ймовірності $\mathrm{P}_{n,p}(k)$ зменшуються). Нехай $x = (n+1)\,p$, а через $[x]$ позначимо найбільше ціле число, яке не перевищує x. Якщо x не ціле, то шуканим k буде $k = [x]$. Якщо ж x ціле, то шуканих k буде два: $[x]$ та $[x]-1$. Цей же результат інколи формулюють так: шукані k є цілими розв'язками нерівності

$$np - q \leq k \leq np + p.$$

Приклад 8.5

За даними метеослужби ймовірність того, що 8 березня в Києві піде дощ, дорівнює 0.31. Знайти найбільш ймовірну кількість дощових днів 8 березня в Києві упродовж 50 років.

Розв'язок: В даному випадку $k = np + p = 51 \cdot 0.31 = 15.81$ – не ціле. Отже, найбільш ймовірною кількістю дощових днів є 15.

Приклад 8.6

Нехай ймовірність наявності специфічної хвороби у людини 0.003. Яка ймовірність того, що в спільноті, що складається з 1000 осіб, є не менше 10 осіб з цією хворобою?

Розв'язок: Маємо схему з $n = 1000$ випробувань Бернуллі з ймовірністю "успіху" $p = 0.003$. Шукана ймовірність дорівнює

$$\mathrm{P}\{\xi_{1000} \geq 10\} = \sum_{k=10}^{1000} C_{1000}^{k}(0.003)^{k}(0.997)^{1000-k} = 1 - \sum_{k=0}^{9} C_{1000}^{k}(0.003)^{k}(0.997)^{1000-k}.$$

Обчислення навіть одного доданку в цій сумі є вельми проблематичним. В даному випадку, оскільки n велике, а p – мале, варто скористатися наближенням, яке дає теорема Пуассона. Маємо $\lambda = 1000 \cdot 0.003 = 3$, і шукана ймовірність

$$\mathrm{P}\{\xi_{1000} \geq 10\} = 1 - \sum_{k=0}^{9} C_{1000}^{k}(0.003)^{k}(0.997)^{1000-k} \approx 1 - \sum_{k=0}^{9} \frac{3^k}{k!}e^{-3} \approx 0.001.$$

Приклад 8.7 (Дні народження. Феллер, т.1, ст. 171.)

Яка ймовірність p_k того, що в групі з 500 осіб рівно k осіб народилися 1 січня?

Розв'язок: Якщо ці особи вибрані випадково, то можна вважати, що маємо справу зі схемою незалежних випробувань Бернуллі з $n = 500$ (достатньо велике) та $p = 1/365$ (дуже мале). Можемо використати наближення розподілом Пуассона для підрахунку ймовірностей зі значенням $\lambda = 500/365 \approx 1.3699$:

$$\mathrm{P}\{\xi_{500} = k\} = \frac{(1.3699)^k}{k!} e^{-1.3699}.$$

Кілька ймовірностей p_k, підрахованих за точними формулами для біноміального розподілу та їхні пуассонівські наближення наведені в таблиці.

k	0	1	2	3	4	5	6
Біноміальні ймовірності	.2537	.3484	.2388	.1089	.0372	.0101	.0023
Пуассонівські ймовірності	.2541	.3481	.2385	.1089	.0373	.0102	.0023

Приклад 8.8

Страхова компанія має таку політику: вона покриває певні страхові випадки стандартною виплатою 100000 грн. Щомісячний внесок застрахованої особи – 25 грн. Кількість вимог на покриття страхового випадку – в середньому 100. Компанія має більше ніж 1 млн. застрахованих клієнтів. Яка ймовірність того, що компанії за місяць доведеться виплатити більше, ніж 15 млн. грн.?

Розв'язок: Фактично, кожна застрахована особа є одним випробуванням Бернуллі, «успіхом» можна вважати настання страхового випадку та відповідно отримання страхової премії упродовж місяця. Маємо дуже велику кількість випробувань з дуже малою ймовірністю «успіху» в одному з них. Отже, доцільним є наближення випадкової величини, яка є загальною кількістю страхових вимог упродовж місяця, розподілом Пуассона з параметром $\lambda = 100$ (параметр розподілу Пуассона λ є математичним сподіванням[a]).

Ймовірність того, що компанії доведеться виплатити більше ніж 15 млн. грн. упродовж місяця – це ймовірність того, що $\mathrm{P}\{\xi_{1000000} > 150\}$. Якщо підрахувати це число згідно наближенню Пуассона, матимемо $P\{\xi_{1000000} > 150\} = 1.23 \cdot 10^{-6}$. Отже, керівництву компанії не варто хвилюватися.

[a]Визначення математичного сподівання див. на ст. 84.

Приклад 8.9

Знайти ймовірність того, що в результаті 1000 підкидань монети число випадання герба буде знаходитися в інтервалі (475, 525).

Розв'язок: Маємо схему незалежних випробувань Бернуллі з $p = 0.5$, $n = 1000$. Отже, $np = 500$, $npq = 250$.

Підставляючи всі дані в формулу (8.2), маємо:

$$P\{475 \leq \xi_{100} \leq 525\} \approx \Phi\left(\frac{525-500}{\sqrt{250}}\right) - \Phi\left(\frac{475-500}{\sqrt{250}}\right) =$$
$$= \Phi(1.58) - \Phi(-1.58) = 2 \cdot \Phi(1.58) - 1 \approx 2 \cdot 0.9429 - 1 = 0.8858.$$

Значення функції $\Phi(x)$ беремо з таблиці 3 на стор. 302.

Приклад 8.10

Невелике місто щодня відвідують 100 туристів, які вдень йдуть обідати. Кожен з них обирає для обіду один з двох міських ресторанів з рівними ймовірностями та незалежно один від одного. Власник одного з ресторанів хоче, щоб з ймовірністю приблизно 0.99 всі туристи, які прийшли в його ресторан, могли там одночасно пообідати. Скільки місць має бути в ресторані, щоб це було можливо?

Розв'язок: Будемо вважати, що подія A трапилась, якщо турист вибрав ресторан зацікавленого власника. Відповідно за умовою маємо $n = 100$, $p = P(A) = q = 1/2$. Нас цікавить найменша кількість відвідувачів m, для якої ймовірність одночасного приходу не менш ніж m туристів зі 100 з ймовірністю "успіху" для кожного $1/2$ приблизно дорівнює ймовірності переповнення ресторану, тобто 1-0.99=0.01.

Таким чином, нас цікавить таке найменше число m, що

$$P\{m \leq \xi_{100} \leq 100\} \approx 0.01.$$

Застосуємо інтегральну теорему Муавра-Лапласа (формулу (8.2)) В нашому випадку m – невідоме, $\sqrt{npq} = 5$, $\dfrac{B - np}{\sqrt{npq}} = \dfrac{100 - 50}{5} = 10$, $\dfrac{A - np}{\sqrt{npq}} = \dfrac{m - 50}{5} = \dfrac{m}{5} - 10$. Тоді

$$0.01 \approx \mathrm{P}\left\{m \le \xi_{100} \le 100\right\} \approx \left(\Phi(10) - \Phi\left(\frac{m}{5} - 10\right)\right) \approx 1 - \Phi\left(\frac{m}{5} - 10\right),$$

$$\Phi\left(\frac{m}{5} - 10\right) \approx 0.99.$$

В таблицях значень функції $\Phi(x)$ (див. табл. 3 на стор. 302) знаходимо $\dfrac{m}{5} - 10 \approx 2.33$, отже $m \approx (2.33 + 10) \cdot 5 = 61.65$.

Отже, в ресторані має бути 62 місця.

Задача 8.1. Скільки раз треба підкинути три монети, щоб ймовірність появи принаймні один раз трьох орлів була більшою за 0.8?

Задача 8.2. Скільки раз потрібно підкинути два гральних кубики, щоб ймовірність випадання хоча б один раз суми чисел 5 була більшою за 0.95?

Задача 8.3. Ймовірність взяття пенальті воротарем дорівнює p. Яка ймовірність того, що воротар візьме хоча б один м'яч з 5? При $p = 0.5$; $p = \frac{1}{6}$ знайти найбільш ймовірне число взятих пенальті.

Задача 8.4. Компанія встановлює ціну за страхування збитків від ураганів, використовуючи наступні припущення:
1) упродовж одного календарного року не може бути більше одного урагану;
2) ймовірність того, що упродовж одного календарного року буде ураган, дорівнює 0.05;
3) число ураганів упродовж будь-якого року не залежить від числа ураганів упродовж будь-якого іншого календарного року.
Використовуючи припущення компанії, знайти розподіл числа ураганів за 20 років та підрахувати ймовірність того, що за 20 років буде менше трьох ураганів.

Задача 8.5. Що більш ймовірно: виграти у гравця рівного собі за силою 4 партії з 8, чи 3 партії з 5?

Задача 8.6. Гральний кубик підкидають п'ять раз. Знайти ймовірність того, що два рази з'явилося число очок, яке кратне трьом.

Задача 8.7. Знайти умовну ймовірність того, що у n випробуваннях Бернуллі герб випав при k-ому випробуванні, якщо відомо, що герб випав рівно один раз.

Задача 8.8. Зроблено 20 пострілів у ціль. Ймовірність влучення при одному пострілі дорівнює 0.7. Обчислити:

а) ймовірність того, що буде принаймні одне влучення;

б) ймовірність того, що буде не більше двох влучень;

в) найбільш ймовірне число влучень.

Задача 8.9. Батарея зробила 14 пострілів по об'єкту. Ймовірність влучення дорівнює 0.2. Обчислити:

а) найбільш ймовірне число влучень;

б) ймовірність знищення об'єкту, якщо для знищення потрібно не менше чотирьох влучень.

Задача 8.10. Проблема Джона Сміта. У 1693 році Джоном Смітом було поставлене таке питання: чи однакові шанси на успіх у трьох чоловіків, якщо першому треба отримати хоча б одну "шістку" при підкиданні грального кубика шість раз, другому – не менше двох "шісток" при дванадцяти підкиданнях грального кубика, третьому – не менше трьох "шісток" при вісімнадцяти кидках грального кубика?

Задача 8.11. Ймовірність успіху в схемі Бернуллі втричі більша за ймовірність невдачі. Для кожного цілого невід'ємного k знайти ймовірність того, що серед чотирьох випробувань буде рівно k успіхів.

Задача 8.12. Нехай ξ – випадкова величина, що має біноміальний розподіл з параметрами $n = 25$ та $p = 0.2$. Оцінити $\mathrm{P}\xi < \mathbf{M}\xi - 2\sqrt{\mathbf{D}\xi}$.

Задача 8.13. Нехай ξ – випадкова величина, що має біноміальний розподіл з параметрами $n = 100$ та $p = 0.1$. Оцінити $\mathrm{P}\xi < \mathbf{M}\xi - 3\sqrt{\mathbf{D}\xi}$.

Задача 8.14. Нехай ξ – випадкова величина, що має біноміальний розподіл з параметрами n та p. За яких значень параметра p дисперсія $\mathbf{D}\xi$ досягатиме свого максимального значення при фіксованому n?

Задача 8.15. П'яниця, маючи бажання дійти до кінця вулиці, робить крок у вірному напрямку або в протилежному напрямку з однаковою ймовірністю. Нехай ξ – кількість кроків в правильному напрямку, на які він просунеться за n спроб. Знайти розподіл ξ та її математичне сподівання[1].

Задача 8.16. Нехай U_n – ймовірність того, що число успіхів у n випробуваннях Бернуллі парне. Знайти рекурентне співвідношення і генератрису послідовності U_n $(n = 1, 2, \ldots)$. Генератриса послідовності φ_n, $(n = 1, 2, \ldots)$ визначається так:

$$\psi(s) = \sum_{n \geq 1} \varphi_n s^n.$$

Знайти явний вид ймовірності U_n.

[1] Визначення математичного сподівання див. на ст. 84.

Задача 8.17.* Нехай V_n – це ймовірність того, що число успіхів у n випробуваннях Бернуллі ділиться на три. Знайти рекурентне співвідношення і генератрису для послідовності V_n ($n = 1, 2, \ldots$).

Задача 8.18. Знайти ймовірність того, що у випробуваннях Бернуллі з ймовірністю успіху p (і невдачі $q = 1 - p$) r-ий успіх відбудеться у випробуванні з номером $r + k$.

Задача 8.19. Задача Банаха про сірники. Математик носить у правій і лівій кишені по коробці сірників, які містять по N сірників. Коли йому потрібен сірник, він навмання обирає одну з кишень. У деякий момент часу математик вперше виймає порожню коробку. Інша коробка у цей момент часу може мати $r \in \{0.1, \ldots, N\}$ сірників. Треба знайти ймовірності відповідних подій u_r ($r = 0.1, 2 \ldots, N$) і ймовірність того, що відсутність сірників було виявлено вперше не в тій коробці, яка спорожніла першою.

Задача 8.20. Гра двох гравців А і В складається з n партій, причому ймовірність виграшу в одній партії гравця А дорівнює p, а гравця В – $q = 1 - p$. Гру зупинили у момент, коли А не вистачає до перемоги a очок, а В – b очок. Яким чином поділити приз?

Задача 8.21.* Підрахувати ймовірність того, що у випробуваннях Бернуллі (з ймовірністю успіху p в окремому випробуванні) серія з a послідовних успіхів з'явиться раніше, ніж серія з b послідовних невдач.

Задача 8.22. Знайти ймовірність того, що з n незалежних випробувань в схемі Бернуллі буде парна кількість успіхів, якщо ймовірність успіху є p, а ймовірність невдачі є $q = 1 - p$.

Задача 8.23. Ви уклали парі з другом, що серед наступних 15 машин, які проїдуть вулицею, буде принаймні дві, у яких співпадають перша та остання цифри номерного знаку. Яка ймовірність того, що Ви виграєте парі?

Задача 8.24. Яка ймовірність того, що одне й те саме число випаде принаймні двічі в 10 обертаннях рулетки?

Задача 8.25. Двоє підкидають монету n разів кожен. Яка ймовірність того, що у них випаде однакова кількість гербів?

Задача 8.26. Знайти ймовірність влучити в мішень не менше двох разів при десяти незалежних пострілах, якщо ймовірність влучення при одному пострілі $1/5$? Знайти ймовірність принаймні двох влучень, якщо відомо, що є одне влучення.

Задача 8.27. Знайти ймовірність влучити в мішень не менше двох разів при 500 незалежних пострілах, якщо ймовірність влучення при одному пострілі 0.002? Знайти ймовірність принаймні двох влучень.

Задача 8.28. В магазин надійшла партія з 1000 ламп. Доля бракованих ламп цього типу складає 0.1 %. Яка ймовірність того, що

а) в партії не буде жодної бракованої лампи;

б) в партії буде рівно одна бракована лампа;
в) буде більше двох бракованих ламп?

Задача 8.29. Яка ймовірність того, що при 1000 підкиданнях симетричної монети хоча б 5 разів випаде 10 гербів?

Задача 8.30. Студент їде маршруткою до університету та назад 20 днів упродовж місяця. Час очікування маршрутки є випадковою величиною, рівномірно розподіленою від 0 до 15 хвилин. Яка ймовірність того, що студент буде чекати маршрутку менше 30 секунд хоча б 5 разів упродовж осені?

Задача 8.31. Ймовірність того, що мешканці квартири замовлять піццу упродовж дня, дорівнює 0.001. Яка ймовірність того, що мешканці кварталу на 500 квартир замовлять більше, ніж одну піццу?

Задача 8.32. Страхова компанія виявила, що приблизно 0.005 відсотка населення помирає щороку внаслідок специфічної аварії. Яка ймовірність того, що в наступному році компанія буде змушена виплатити більше трьох страхових виплат, пов'язаних з цією аварією, серед 10000 застрахованих осіб?

Задача 8.33. Процес виробництва електричних чайників має забезпечувати долю бракованих виробів в межах 1 %. Кожну годину з виробництва випадковим чином беруть 10 чайників на перевірку. Якщо серед них буде хоча б один бракований чайник, всю партію, що вироблена упродовж години, відправляють на додаткову перевірку. Якщо доля браку відповідає заявленій, яка ймовірність того, що партія буде відправлена на перевірку? Яку кількість чайників у вибірці (замість 10) треба перевіряти, якщо виробник бажає, щоб ймовірність того, що партія відправлятиметься на перевірку, була приблизно 0.95 при долі браку 10 %?

Задача 8.34. Одна лампа при вмиканні перегорає з ймовірністю 0.001. Знайти ймовірність того, що п'яти ламп вистачить на 2000 вмикань.

Задача 8.35. Ймовірність вгадати всі шість номерів в лотереї "6 з 49" дорівнює $7.15 \cdot 10^{-8}$. В розіграші було продано 5 млн білетів. Яка ймовірність того, що хоча б хтось вгадав всі шість номерів? Яка мінімальна кількість білетів має бути продана, щоб з ймовірністю не менше, ніж 0.9, був би хоча б один повний виграш?

Задача 8.36. Насінина соняшника проростає в певному регіоні з ймовірністю 0.8. Знайти ймовірність того, що серед 500 посаджених насінин проросте хоча б 450.

Задача 8.37. Знайти ймовірність того, що серед 100 незалежних пострілів влучними будуть від 80 до 90, якщо ймовірність влучення при кожному пострілі незалежно від інших дорівнює 0.75. Якщо для знищення об'єкту потрібно принаймні 90 влучень, яку мінімальну кількість пострілів треба зробити, щоб ймовірність знищення об'єкта була принаймні 0.95?

Задача 8.38. В містечку, де проживає 5000 дорослих, проведено соціальне опитування щодо певних змін. У виборці з 100 осіб 60 осіб висловилось за зміни, 40 – проти. Якби б зміни підтримувало лише половина населення, яка ймовірність того, що при вибірковому опитуванні отримано 60 голосів "за" ?

Задача 8.39. Експеримент полягає в підкиданні двох симетричних гральних кубиків. Яка ймовірність того, що якщо підкинути їх 600 разів, то кількість разів, коли сума очок на двох кубиках дорівнює 7, буде від 90 до 110?

Задача 8.40. В певному регіоні лише 75 відсотків водіїв регулярно використовує пасок безпеки. Випадковим чином обрано 500 водіїв. Яка ймовірність того, що
а) з обраних водіїв регулярно користуються пасками безпеки від 360 до 400 водіїв?
б) менше 400 з цих водіїв регулярно користуються пасками безпеки?

9 ДИСКРЕТНІ ВИПАДКОВІ ВЕЛИЧИНИ

Нехай $(\Omega, \Re, \mathrm{P})$ ймовірнісний простір, а $A_1, A_2, ..., A_k, .. \in \Re$ скінченна або зліченна множина подій таких, що $\bigcup\limits_k A_k = \Omega$, $A_i \cap A_j = \varnothing$, $i \neq j$. Через $\chi_{A_k}(\omega)$ будемо позначати **індикатор події** A_k:

$$\chi_{A_k}(\omega) = \begin{cases} 1, & \text{якщо } \omega \in A_k, \\ 0, & \text{якщо } \omega \notin A_k. \end{cases}$$

Дискретною випадковою величиною $\xi(\omega)$ будемо називати лінійну комбінацію індикаторів

$$\xi(\omega) = \sum_k x_k \chi_{A_k}(\omega), \qquad x_k \in R.$$

Очевидно, що для $\omega \in A_k$, $\xi(\omega) = x_k$ і множина усіх значень $X = \{x_1, ..., x_k, ...\}$ не більш ніж зліченна.

Законом розподілу випадкової величини ξ будемо називати ймовірність $\mathrm{P}(\xi \in B)$, що розглядається як функція числової множини B. Закон розподілу визначається значеннями $x_1, x_2, ..., x_k, ..$, які приймає ξ, і ймовірностями цих значень $p_k = P\{\omega : \xi(\omega) = x_k\}$, $k = 1, 2, ...$ Так

$$\mathrm{P}(\xi \in B) = \sum_{k: x_k \in B} p_k.$$

Нижче наведено приклади важливих дискретних розподілів.

1. **Біноміальний** закон розподілу з параметрами n і p ($n \in N$, $0 < p < 1$):

$$p_k = C_n^k p^k (1-p)^{n-k}, \qquad k \in X = \{0, 1, ..., n\}.$$

Для скорочення цей розподіл часто позначають як $Bi(n,p)$, також можна зустріти позначення $Bin(n,p)$ та $B(n,p)$. Фразу "*випадкова величина* ξ *має біноміальний розподіл з параметрами* n *і* p" можна скорочено записати як $\xi \sim Bi(n,p)$.

2. **Геометричний** закон розподілу з параметром p ($0 < p < 1$):

$$p_k = (1-p)^k p, \qquad k \in X = \{0, 1, 2, ...\}.$$

Цей розподіл інколи позначають як $Geom(p)$ або $G(p)$. Для випадкової величини ξ з таким розподілом можна писати $\xi \sim G(p)$.

3. **Пуассонівський** закон розподілу з параметром λ ($\lambda > 0$):

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k \in X = \{0, 1, ...\}.$$

Цей розподіл інколи позначають як $\Pi(\lambda)$ або $Pois(\lambda)$. Для випадкової величини ξ з таким розподілом можна писати $\xi \sim \Pi(\lambda)$.

4. **Гіпергеометричний** закон розподілу з параметрами n, N, M (n, N, M – натуральні числа, $n, M \le N$):

$$p_k = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \qquad k \in X = \{\max(0, n-(N-M)), ..., \min(M,n)\}.$$

5. **Розподіл Паскаля** з параметром $a > 0$:

$$\mathrm{P}(\xi = n) = \frac{a^n}{(1+a)^{n+1}}, \qquad n = 0, 1, ...$$

Дискретна випадкова величина ξ називається **сумовною**, якщо ряд $\sum\limits_k x_k p_k$ збігається абсолютно. Тоді сума цього ряду називається **математичним сподіванням** і позначається через $\mathbf{M}\xi$. Таким чином,

$$\mathbf{M}\xi = \sum_k x_k p_k.$$

Якщо ξ і η сумовні випадкові величини, то $\xi + \eta$ теж сумовна і

$$\mathbf{M}\,(\xi + \eta) = \mathbf{M}\xi + \mathbf{M}\eta.$$

Коли існує $\mathbf{M}\xi^n$, то його називають моментом n-го порядку випадкової величини ξ. Момент n-го порядку $(\xi - \mathbf{M}\xi)$ називається центральним моментом n-го порядку для ξ. Центральний момент другого порядку називається **дисперсією** і позначається через $\mathbf{D}\xi$:

$$\mathbf{D}\xi = \mathbf{M}(\xi - \mathbf{M}\xi)^2.$$

Наведемо основні властивості дисперсії:

1) $\mathbf{D}\xi = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2$;

2) для будь-якої константи c $\mathbf{D}(c\xi) = c^2\mathbf{D}\xi$, $\mathbf{D}(\xi + c) = \mathbf{D}\xi$;

3) $\mathbf{D}\xi = 0$ тоді і тільки тоді, коли $P\{\omega : \xi(\omega) = \mathbf{M}\xi(\omega)\} = 1$.

Нехай на $(\Omega, \Re, \mathrm{P})$ задано n дискретних випадкових величин $\xi_1, ..., \xi_n$. Їх можна розглядати як n-вимірний **випадковий вектор** $\xi = (\xi_1, ..., \xi_n)$. **Сумісним законом розподілу** $\xi_1, ..., \xi_n$ (або вектора ξ) будемо називати ймовірність

$$P\{\xi = (\xi_1, ..., \xi_n) \in B\},$$

яка розглядається як функція множини $B \subseteq R^n$.
Цей закон визначається значеннями $(x_{1j_1}, x_{2j_2}, ..., x_{nj_n})$, які приймає $\xi = (\xi_1, ..., \xi_n)$, і ймовірностями цих значень

$$p_{j_1,j_2,...,j_n} = \mathrm{P}(\xi_1 = x_{1j_1}, \xi_2 = x_{2j_2}, ..., \xi_n = x_{nj_n}).$$

Очевидно, $p_{j_1,j_2,...,j_n} \ge 0$ і $\sum\limits_{j_1,j_2,...,j_n} p_{j_1,j_2,...,j_n} = 1$.

Випадкові величини $\xi_1, ..., \xi_n$ будемо називати **незалежними**, якщо для довільних $x_{1j_1}, x_{2j_2}, ..., x_{nj_n}$

$$\mathrm{P}(\xi_1 = x_{1j_1}, \xi_2 = x_{2j_2}, ..., \xi_n = x_{nj_n}) = \prod_{i=1}^{n} \mathrm{P}(\xi_i = x_{ij_i}).$$

Якщо на ймовірнісному просторі $(\Omega, \Re, \mathrm{P})$ задана послідовність випадкових величин $\{\xi_n\}_1^\infty$, то їх незалежність визначається як незалежність будь-якої скінченної підмножини випадкових величин з цієї послідовності.

Наведемо додаткові властивості математичного сподівання і дисперсії для незалежних випадкових величин.

1) Якщо $\xi_1, ..., \xi_n$ незалежні і сумовні, то $\prod\limits_{i=1}^{n} \xi_i$ теж сумовна і

$$\mathbf{M}\left(\prod_{i=1}^{n} \xi_i\right) = \prod_{i=1}^{n} \mathbf{M}\xi_i.$$

2) Якщо $\xi_1, ..., \xi_n$ незалежні і $\mathbf{M}\xi_i^2 < +\infty$, $i = 1, 2, ..., n$, то

$$\mathbf{D}(\xi_1 + ... + \xi_n) = \mathbf{D}\xi_1 + ... + \mathbf{D}\xi_n.$$

Розглянемо дві випадкові величини ξ і η, для яких $\mathbf{M}\xi^2, \mathbf{M}\eta^2 < +\infty$.

Коваріацією випадкових величин ξ та η називається

$$\mathrm{cov}(\xi, \eta) = \mathbf{M}(\xi - \mathbf{M}\xi)(\eta - \mathbf{M}\eta) = \mathbf{M}\xi\eta - \mathbf{M}\xi\mathbf{M}\eta.$$

Коефіцієнтом кореляції випадкових величин ξ та η називається

$$\rho(\xi, \eta) = \frac{\mathrm{cov}(\xi, \eta)}{\sqrt{\mathbf{D}\xi}\sqrt{\mathbf{D}\eta}}.$$

Теорема 9.1 Коефіцієнт кореляції має наступні властивості

1. $|\rho(\xi, \eta)| \leq 1$.
2. Якщо $\rho(\xi, \eta) = \pm 1$, то існують дійсні числа a, b, c такі що $a\xi + b\eta = c$.
3. Якщо випадкові величини ξ та η є незалежними, то $\rho(\xi, \eta) = 0$.
4. Для довільних дійсних a_1 та a_2 має місце $\rho(a_1\xi + b_1, a_2\eta + b_2) = sign\,(a_1 a_2)\,\rho(\xi, \eta)$.

Дискретну випадкову величину, яка приймає цілі невід'ємні значення, будемо називати **цілочисловою**. Цілочислову випадкову величину зручно вивчати за допомогою **генератриси** $\psi_\xi(s)$, що визначається як

$$\psi_\xi(s) = \mathbf{M}s^\xi = \sum_{k\geq 0} p_k s^k \tag{9.1}$$

і існує для $|s| \leq 1$. Оскільки

$$p_k = \frac{1}{k!}\psi_\xi^{(k)}(0), \quad k = 0, 1, ... \tag{9.2}$$

то між законами розподілу і генератрисами рівності (9.1), (9.2) встановлюють взаємно однозначну відповідність.

Замість моментів $\mathbf{M}\xi^n$ у випадку цілочислових випадкових величин більш зручно мати справу з **факторіальними моментами** $\mathbf{M}\xi^{[n]}$, де $\xi^{[n]} = \xi(\xi-1)...(\xi-n+1)$, $\xi^{[0]} = 1$. Очевидно, через факторіальні моменти можна виписати $\mathbf{M}\xi^n$ і навпаки. Має місце рівність

$$\mathbf{M}\xi^{[n]} = \psi_\xi^{(n)}(1) \tag{9.3}$$

де права частина (9.3) розуміється як $\lim\limits_{s\uparrow 1}\psi_\xi^{(n)}(s)$ і може дорівнювати $+\infty$.

Для підрахунку $\mathbf{M}\xi$ і $\mathbf{D}\xi$ часто користуються формулами

$$\mathbf{M}\xi = \psi'_\xi(1), \qquad \mathbf{D}\xi = \psi''_\xi(1) + \psi'_\xi(1) - [\psi'_\xi(1)]^2.$$

Якщо $\xi_1, ..., \xi_n$ незалежні, то

$$\psi_{\xi_1+...+\xi_n}(s) = \prod_{i=1}^{n}\psi_{\xi_i}(s) \tag{9.4}$$

Формулу (9.4) можна узагальнити. Якщо $\nu, \xi_1, ..., \xi_n, ...$ незалежні випадкові величини, $S_\nu = \xi_1 + ... + \xi_\nu$ і $\xi_1, ..., \xi_n, ...$ однаково розподілені з генератрисою $\psi_\xi(s)$, то

$$\psi_{S_\nu}(s) = \psi_\nu(\psi_\xi(s)) \tag{9.5}$$

Наведені вище поняття та формули можна узагальнити на випадок, коли випадкова величина приймає цілі значення (а не лише невід'ємні). В цьому випадку випадкова величина теж називається цілочисловою а її генератриса визначається тепер так

$$\psi_\xi(s) = \mathbf{M}s^\xi = \sum_{k=-\infty}^{\infty} p_k s^k, \quad |s| \leq 1.$$

Аналогом формули (9.2) тепер є

$$p_k = \frac{1}{2\pi i}\oint\limits_{|s|=1}\frac{\phi_\xi(s)}{s^{k+1}}ds.$$

Формули (9.3)-(9.5) зберігаються і для таких цілочислових випадкових величин. (Звичайно вважаємо, що у випадку формули (9.5) випадкова величина ν приймає натуральні значення).

Нехай тепер (ξ_1, ξ_2) – пара випадкових величин з цілочисловими компонентами ξ_i, $i = 1, 2$, та сумісним розподілом

$$\mathrm{P}\{\xi_1 = i, \xi_2 = j\} = p_{ij}, \quad i, j = 0, 1, 2, ...$$

Двовимірною генератрисою вектора (ξ_1, ξ_2) називається функція від двох змінних

$$P_{\xi_1,\xi_2}(s_1, s_2) = \sum_{i,j} p_{ij} s_1^i s_2^j = \mathbf{M}(s_1^{\xi_1} s_2^{\xi_2})$$

Двовимірний випадок може бути очевидним чином розширений на $\xi' = (\xi_1, ..., \xi_m)$ – випадковий вектор з цілочисловими компонентами ξ_i, $i = 1, ..., m$.

Нехай є однотипні частки, які можуть породжувати подібні собі нові частки. Одна початкова частка утворює нульове покоління. Вона живе одиницю часу (покоління), після чого гине, породжуючи випадкову кількість ξ_0 подібних собі нащадків: рівно k нових часток з ймовірністю p_k, $k = 0, 1, 2, ...$ Кожна з новонароджених часток також живе одиницю часу і в кінці її породжує незалежно від інших часток випадкову кількість ξ_1 нащадків k з ймовірністю p_k, $k = 0, 1, 2, ...,$, і т.д. Безпосередні нащадки часток n-го покоління утворюють $(n + 1)$-ше покоління. Яка ймовірність того, що нащадки однієї частки вимруть?

Для всіх номерів покоління часток $n = 0, 1, ...$ визначимо випадкові величини ξ_n – кількість нащадків однієї частки n- го покоління. Ми припускаємо, що ξ_n, $n = 0, 1, ...,$ – незалежні в сукупності однаково розподілені невід'ємні цілочисельні випадкові величини, $p_k = \mathrm{P}\{\xi = k\}$ – ймовірність того, що одна частка матиме k нащадків.

Простим гіллястим процесом (процесом Гальтона-Ватсона) називається послідовність випадкових величин ν_n, $n = 0, 1, ...$, яка визначає загальну кількість часток в n-му поколінні та задається рекурентно

$$\nu_0 = 1, \quad \nu_n = \sum_{k=1}^{\nu_{n-1}} \xi_k = \sum_{k=1}^{\xi_1} \nu_{n-1}, \quad n = 0, 1, 2, ...$$

Позначимо через

$$P(s) = \mathbf{M}s^{\xi} = \sum_{k=0}^{\infty} \mathrm{P}\{\xi = k\} s^k = \sum_{k=0}^{\infty} p_k s^k, \quad |s| \le 1,$$

генератрису, що визначає розподіл кількості нащадків однієї частки ξ_n. Генератриса кількості часток в n-му поколінні може бути записана як суперпозиція

$$P_n(s) = P_{\nu(n)}(s) = P_{\nu(n-1)}(P(s)) = P(P(\nu(n-1))) = \underbrace{P(P(...P(s)))}_{n \text{ раз}}.$$

Ймовірністю виродження гіллястого процесу в n-му поколінні називають ймовірність того, що в цьому поколінні кількість часток дорівнює нулю:

$$\pi_n = \mathrm{P}\{\nu_n = 0\}.$$

Оскільки генератриса за визначенням є сумою ряду

$$P_n(s) = \mathbf{M}s^{\nu_n} = \sum_{k=0}^{\infty} \mathrm{P}\{\nu_n = k\}s^k,$$

то ймовірність виродження процесу в n-му поколінні дорівнює коефіцієнту p_0 цього ряду, або значенню генератриси в нулі:

$$\pi_n = \mathrm{P}\{\nu_n = 0\} = P_n(0).$$

Ймовірністю виродження гіллястого процесу називається границя ймовірності виродження процесу в n-му поколінні при $n \to \infty$:

$$\pi = \lim_{n\to\infty} \mathrm{P}\{\nu_n = 0\} = \lim_{n\to\infty} \pi_n.$$

Якщо $\mathbf{M}\xi = \mu$ середня кількість безпосередніх нащадків однієї частки, то, згідно властивості генератриси, маємо:

$$\mathbf{M}\xi = \mu = P'(1).$$

Теорема 9.2 Теорема про ймовірність виродження гіллястого процесу Якщо для гіллястого процесу $\mu \le 1$, то ймовірність виродження процесу $\pi = 1$. Якщо $\mu > 1$, то ймовірність виродження процесу $\pi < 1$ є єдиним на інтервалі $[0, 1)$ розв'язком рівняння

$$\pi = P(\pi).$$

Приклад 9.1

В двох урнах лежать білі та чорні кулі: в першій – 3 білих, 4 чорних, в другій – 4 білих, 3 чорних. З першої урни навмання витягли дві кулі і переклали в другу урну. Записати розподіл випадкової величини ξ – кількості білих куль в другій урні після експерименту.

Розв'язок: Після експерименту кількість білих куль в другій урні може лишитися рівною 4, якщо переклали дві чорні кулі (ймовірність цього $\frac{4}{7} \cdot \frac{3}{6} = \frac{2}{7}$), може стати рівною 5, якщо переклали 1 білу, 1 чорну кулі (з ймовірністю $\frac{4}{7} \cdot \frac{3}{6} + \frac{3}{7} \cdot \frac{4}{6} = \frac{4}{7}$), або стати рівною 6, якщо переклали дві білі кулі (з ймовірністю $\frac{3}{7} \cdot \frac{2}{6} = \frac{1}{7}$). Отже, ряд розподілу цієї випадкової величини зручно записати у вигляді таблиці:

x_i	4	5	6
p_i	2/7	4/7	1/7

Зауважимо, що сума всіх ймовірностей в нижньому ряду таблиці дорівнює 1, що є доброю ознакою.

Приклад 9.2

Розподіл випадкової величини ξ визначено таким чином:

$$\mathrm{P}\{\xi = k\} = \frac{C}{k(k+1)}, \quad k = 1, 2, \dots .$$

Знайти значення константи C та ймовірності $\mathrm{P}\{\xi \le 3\}$, $\mathrm{P}\{k_1 \le \xi \le k_2\}$.

Розв'язок: Значення C знаходимо з умови $\sum_{k=1}^{\infty} \mathrm{P}\{\xi = k\} = 1$:

$$\sum_{k=1}^{\infty} \mathrm{P}\{\xi = k\} = \sum_{k=1}^{\infty} \frac{C}{k(k+1)} = C \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+1} \right) =$$
$$= C \left(\sum_{k=1}^{\infty} \frac{1}{k} - \sum_{k=1}^{\infty} \frac{1}{k+1} \right) = C \left(\sum_{k=1}^{\infty} \frac{1}{k} - \sum_{k=2}^{\infty} \frac{1}{k} \right) = C = 1.$$

Знайдемо вказані ймовірності:

$$\mathrm{P}\{\xi \le 3\} = \mathrm{P}\{\xi = 1\} + \mathrm{P}\{\xi = 2\} + \mathrm{P}\{\xi = 3\} = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} = \frac{3}{4};$$

$$\mathrm{P}\{k_1 \le \xi \le k_2\} = \sum_{k=k_1}^{k_2} \frac{1}{k(k+1)} = \sum_{k=k_1}^{k_2} \frac{1}{k} - \sum_{k=k_1}^{k_2} \frac{1}{k+1} =$$
$$= \sum_{k=k_1}^{k_2} \frac{1}{k} - \sum_{k=k_1+1}^{k_2+1} \frac{1}{k} = \frac{1}{k_1} - \frac{1}{k_2+1} = \frac{k_2 + 1 - k_1}{k_1(k_2+1)}.$$

Приклад 9.3

Екзаменатор задає студенту додаткові питання. На кожне з питань студент може відповісти незалежно від інших з ймовірністю 0.9. Екзаменатор припиняє співбесіду, щойно студент не знатиме відповіді на питання. Записати розподіл кількості питань, які екзаменатор задасть студенту. Скільки в середньому питань буде задано? Яка дисперсія цієї випадкової величини?

Розв'язок: Нехай випадкова величина ξ – кількість питань, заданих студенту. Очевидно, що буде задано принаймні одне питання, отже, множина значень цієї випадкової величини $\{1, 2, ...\}$. Знайдемо відповідні ймовірності.

$$\mathrm{P}\{\xi = 1\} = \mathrm{P}\{\text{Студент на перше ж питання не дав відповіді}\} = 1 - 0.9 = 0.1;$$

$$\mathrm{P}\{\xi = 2\} = \mathrm{P}\{\text{Студент на перше питання дав відповідь, на друге – ні}\} = \\ = 0.9 \cdot (1 - 0.9) = 0.09;$$

$$\dots$$

$$\mathrm{P}\{\xi = k\} = \mathrm{P}\{\text{Студент дав відповідь на перші } k-1 \text{ питання, на } k\text{-те – ні}\} = \\ = (0.9)^{k-1} \cdot (1 - 0.9).$$

Отже, кількість заданих питань є геометрично розподіленою випадковою величиною з параметром (ймовірністю "успіху" в одному випробуванні) $p = 0.1$. Математичне сподівання такої величини дорівнює $\mathbf{M}\xi = 1/p = 10$, дисперсія – $\mathbf{D}\xi = q/p^2 = 0.9/0.01 = 90$.

Приклад 9.4

Припустимо, що число пошкоджень транспортного засобу описується випадковою величиною ν, яка має розподіл Паскаля з параметром $a > 0$. Звернення до страхової компанії за відшкодуванням при цьому є незалежними і мають ймовірність p. Показати, що число звернень буде мати розподіл Паскаля з параметром pa.

Розв'язок: Число звернень до страхової компанії будемо описуватися випадковою величиною η, яка може бути представлена наступним чином: $\eta = \xi_1 + \xi_2 + ... + \xi_\nu$, де ξ_i є незалежними випадковими величинами, що мають розподіл Бернуллі:

$$\xi_i = \begin{cases} 1 \text{ з ймовірністю } & p, \\ 0 \text{ з ймовірністю } & 1-p. \end{cases}$$

Знайдемо генератрису випадкових величин ν та ξ_i:

$$\psi_\nu(s) = \sum_{n=0}^{+\infty} s^n \frac{a^n}{(1+a)^{n+1}} = \frac{1}{1+a(1-s)},$$

$$\psi_{\xi_i}(s) = sp + 1 - p = 1 - \mathrm{P}(1-s).$$

Оскільки генератриса випадкової величини η дорівнює суперпозиції генератрис випадкових величин ν та ξ_i, то

$$\psi_\eta(s) = \psi_\nu(\psi_{\xi_i}(s)) = \frac{1}{1+a\mathrm{P}(1-s)}.$$

Останній вираз є генератрисою розподілу Паскаля з параметром pa.

Приклад 9.5

Володимир і Олексій навмання виймають по одній кулі з урни, в якій 5 білих, 1 чорна та 2 сірі кулі. Володимир першим виймає одну кулю без повернення. Потім кулю бере Олексій. Нехай ξ – число білих куль у Володимира, а η – число чорних куль у Олексія. Знайти коефіцієнт кореляції $\rho(\xi, \eta)$.

Розв'язок: Оскільки таблиця розподілу вектора $(\xi; \eta)$ має вид

ξ \ η	0	1	Σ
0	$19/56$	$2/56$	$3/8$
1	$30/56$	$5/56$	$5/8$
Σ	$7/8$	$1/8$	1

то $\mathbf{M}\xi = \mathbf{M}\xi^2 = \frac{5}{8}$, $\mathbf{M}\eta = \mathbf{M}\eta^2 = \frac{1}{8}$, $\mathbf{D}\xi = \frac{15}{64}$, $\mathbf{D}\eta = \frac{7}{64}$, $\mathbf{M}\xi\eta = \frac{5}{56}$ і тому

$$\rho(\xi, \eta) = \frac{\mathbf{M}\xi\eta - \mathbf{M}\xi\mathbf{M}\eta}{\sqrt{\mathbf{D}\xi\mathbf{D}\eta}} = \frac{\frac{5}{56} - \frac{5}{8}\cdot\frac{1}{8}}{\sqrt{\frac{15}{64}\cdot\frac{7}{64}}} = \frac{\sqrt{105}}{147} \approx 0.0697.$$

Відповідь: $\rho(\xi, \eta) = \dfrac{\sqrt{105}}{147} \approx 0.0697.$

Приклад 9.6

Задано розподіл ймовірностей дискретної двовимірної випадкової величини (випадкового вектора) (X, Y):

$Y \backslash X$	-1	0	1
0	0	0.1	0.4
1	0.2	0.2	0.1

Треба знайти розподіл компонент X та Y. Підрахувати ймовірність $\mathrm{P}\{X < Y\}$.

Розв'язок: Спочатку знайдемо розподіл випадкових величин X та Y. Випадкова величина X приймає значення -1, 0 та 1 з ймовірностями $0.2 = 0 + 0.2$, $0.3 = 0.1 + 0.2$ і $0.5 = 0.4 + 0.1$ відповідно, тому випадкова величина X має розподіл

X	-1	0	1
P	0.2	0.3	0.5

Аналогічно отримаємо розподіл випадкової величини Y:

Y	0	1
P	0.5	0.5

Ймовірність

$$\mathrm{P}\{X < Y\} = \mathrm{P}\{(X = -1) \cap (Y = 0)\} + \mathrm{P}\{(X = -1) \cap (Y = 1)\} +$$
$$+\mathrm{P}\{(X = 0) \cap (Y = 1)\} = 0 + 0.2 + 0.2 = 0.4.$$

Відповідь: $\mathrm{P}\{X < Y\} = 0.4$.

Приклад 9.7

Розглянемо бінарне розщеплення часток, тобто випадок, коли в кінці життя частка або гине з ймовірністю q, або розщеплюється на дві нові частки (генерує двох нащадків) з ймовірністю $p = 1 - q$. Така модель часто використовується в біології.

Знайдемо ймовірність того, що нащадки однієї частки вимруть за два, три покоління, та ймовірність того, що вони вимруть коли-небудь.

Розв'язок: Генератриса випадкової величини ξ – кількості нащадків частки в одному поколінні – має вигляд $P(s) = q + ps^2$.

Ймовірність для популяції вимерти за n кроків є значенням $\pi_n = P_n(0)$ ($\pi_1 = P(0) = P_1(0) = q$). Отже, спочатку знайдемо генератриси кількості часток в другому та третьому поколінні.

$$P_2(s) = P(P(s)) = q + pP(s)^2, \quad P_3(s) = P(P_2(s)) = q + pP_2(s)^2.$$

Відповідно ймовірність вимирання в другому та третьому поколінні дорівнює відповідно

$$\pi_2 = P_2(0) = q + pq^2, \quad \pi_3 = P_3(0) = P(P_2(0)) = q + p(q + pq^2)^2,$$

Для того, щоб знайти ймовірність виродження, треба розв'язати рівняння $q + ps^2 = s$. В результаті матимемо $\pi = q/p$, якщо $p > 1/2$ (середня кількість нащадків однієї частки > 1), та $\pi = 1$, якщо $p \leq 1/2$.

Задача 9.1. Випадкова величина ξ приймає значення -1, 0, 1 з ймовірностями $1/3$. Знайти розподіл випадкової величини $\eta = |\xi|$, $\mathbf{M}\eta$ і $\mathbf{D}\eta$.

Задача 9.2. Випадкова величина ξ приймає значення -1, 0, 1, 2 з ймовірностями 0.2, 0.1, 0.3 і 0.4 відповідно. Знайти розподіл випадкової величини $\eta = 2^\xi$, $\mathbf{M}\eta$ і $\mathbf{D}\eta$.

Задача 9.3. Випадкова величина ξ приймає значення -4, -3, -2, -1, 0, 1, 2, 3 з ймовірностями $1/16$, $1/16$, $1/8$, $1/4$, $1/4$, $1/8$, $1/16$, $1/16$ відповідно. Треба знайти розподіл випадкових величин $\eta = \xi - 3$ та $\zeta = \xi^2 - 1$, їхні математичні сподівання та дисперсії.

Задача 9.4. Які з поданих нижче послідовностей є розподілами випадкових величин:
а) $qp^{k/2}$, $k = 1, 2, 3, \ldots$, $0 < p < 1$, $q = 1 - p$;
б) $qp^{k/2}$, $k = 2, 4, 6, \ldots$, $0 < p < 1$, $q = 1 - p$;
в) $k/\left(k^2 + 1\right)$, $k = 1, 2, 3, \ldots$;
г) $(k-1)p^2q^{k-2}$, $k = 2.4, 6, \ldots$, $0 < p < 1$, $q = 1 - p$;

д) $1/k(k+1)$, $k = 1, 2, ...$;

е) $\frac{2^k}{k!}e^{-2}$, $k = 1, 2, ...$

Для кожної такої послідовності (якщо вона задає розподіл випадкової величини) побудувати ймовірнісний простір і задати на ньому випадкову величину.

Задача 9.5. Підкидаються два гральних кубики. Випадкова величина ξ визначається як сума очок, що випали при підкиданні. Знайти розподіл ξ.

Задача 9.6. Підкидають два гральних кубики 3 рази. Нехай ξ – число появ шістки на першому кубику, а η – число появ шістки на другому кубику. Довести, що випадкові величини ξ і η є незалежними.

Задача 9.7. Підкидають два гральних кубики. Нехай ξ– число очок, що випало на першому кубику, а η – мінімальне з двох очок. Чи будуть випадкові величини ξ і η незалежними?

Задача 9.8. Підкидають два гральних кубики. Треба знайти:

а) математичне сподівання суми очок, що випали;

б) математичне сподівання суми очок, що випали, якщо відомо, що випали різні грані.

Задача 9.9. Нехай ξ – випадкова величина, яка набуває значень $0,\ \pm 1,\ \pm 2, \dots,\ \pm n$ з ймовірностями $\dfrac{1}{2n+1}$. Обчислити $\mathbf{M}\xi$ і $\mathbf{D}\xi$.

Задача 9.10. Гральний кубик підкидають n раз. Нехай ξ – число появ одиниці. Знайти $\mathbf{M}\xi$.

Задача 9.11. Гральний кубик підкидають до n-ої появи одиниці. Треба знайти математичне сподівання числа підкидань.

Задача 9.12. Для біноміального, геометричного та пуассонівського розподілу знайти математичне сподівання і дисперсію.

Задача 9.13. Довести, що якщо ξ – випадкова величина з розподілом $\mathrm{P}\{\xi = k\}$, $k = 1, 2, ...$, то $\mathbf{M}\xi = \sum_{k=1}^{\infty} \mathrm{P}\{\xi \geq k\}$.

Задача 9.14. П'яниця має 5 ключів, з яких лише один відкриває двері в квартиру. Він навмання обирає ключ і намагається відкрити двері. Нехай ξ – кількість спроб. Записати розподіл випадкової величини ξ, якщо

а) кожен раз п'яниця обирає навмання один з п'яти ключів;

б) перевірений ключ відкладається і в подальших випробуваннях участі не бере.

Знайти математичне сподівання та дисперсію кількості спроб в кожному з випадків.

Задача 9.15. Певна частка (перше покоління) може з однаковою ймовірністю мати 1, 2 чи 3 нащадки, які, в свою чергу, також з однаковими ймовірностями можуть мати по 1, 2 чи 3 нащадки. Знайти розподіл кількості часток третього покоління.

Задача 9.16. Два гравця підкидають одночасно симетричний гральний кубик. При кожному підкиданні виграє той, у кого кількість очок буде більшою. Нехай ξ – кількість виграшних підкидань для першого гравця при n послідовних експериментах. Записати розподіл випадкової величини ξ, знайти її математичне сподівання.

Задача 9.17. При грі з автоматом на барабані випадають випадковим чином номери від 000 до 999. Якщо випадають дві однакові цифри, гравець отримує 10 грн, якщо три однакові цифри – 100 грн. В противному випадку гравець нічого не виграє. Записати ряд розподілу виграша гравця при одній грі. Знайти математичне сподівання цієї величини.

Задача 9.18. Довести, що випадкова величина ξ, яка набуває значень $0, 1, 2, \ldots$ має геометричний розподіл тоді і тільки тоді, коли для довільного $r \geq 0$ виконується співвідношення

$$P\{\xi = k + r | \xi \geq k\} = P\{\xi = r\}.$$

Задача 9.19. Випадкові величини ξ і η – незалежні і однаково розподілені за геометричним законом. Довести, що

$$\mathrm{P}(\xi = k | \xi + \eta = n) = \frac{1}{n+1}, \quad k = 0, 1, ..., n.$$

Задача 9.20. Випадкові величини ξ і η - незалежні і однаково розподілені за геометричним законом. Нехай $\zeta = \max\{\xi, \eta\}$. Треба знайти розподіл ζ а також сумісний розподіл ζ і ξ.

Задача 9.21. Нехай ξ – число випробувань в схемі незалежних випробувань Бернуллі з ймовірністю успіху p, які треба провести до r-го успіху. Знайти $\mathrm{P}\{\xi = n\}$ для $n \geq r$.

Задача 9.22. Подружжя вирішило, що вони народжуватимуть дітей до тих пір, доки у них не буде дві дівчинки. Вважаючи народження дівчинки та хлопчика рівноможливими подіями (без двоєнь, троєнь, тощо),

а) знайти розподіл кількості хлопчиків, яких матиме в результаті це подружжя;

б) знайти ймовірність того, що подружжя матиме чотирьох дітей;

в) знайти ймовірність того, що подружжя матиме принаймні чотирьох дітей;

г) яка в середньому кількість дітей народиться в цій сім'ї?

Задача 9.23. Подружжя вирішило, що вони народжуватимуть дітей до тих пір, доки у них не буде три дитини одної статі. Вважаючи народження дівчинки та хлопчика рівноможливими подіями (без двоєнь, троєнь, тощо), записати розподіл кількості дітей в родині та знайти її математичне сподівання.

Задача 9.24. Нехай ξ – випадкова величина, яка має біноміальний розподіл з параметрами n і p. Відомо, що $\mathbf{M}\xi = 12$, $\mathbf{D}\xi = 4$. Знайти n і p.

Задача 9.25. Нехай ξ– випадкова величина, яка приймає значення $k = r,\ r+1,\ r+2,\ \ldots$ з ймовірностями

$$\mathrm{P}(\xi = k) = C_{k-1}^{r-1} p^r q^{k-r}, \quad p, q > 0, \quad p + q = 1.$$

(**Від'ємний біноміальний розподіл** з параметрами r і p). Потрібно знайти $\mathbf{M}\xi$ і $\mathbf{D}\xi$.

Задача 9.26. Нехай $\mathbf{M}\xi = 10$ та $\mathbf{D}\xi = 4$. Чи може вона мати від'ємний біноміальний розподіл?

Задача 9.27. Нехай ξ має від'ємний біноміальний розподіл з $\mathbf{M}\xi = 2$. Знайти $\mathrm{P}\{\xi < 1 | \xi < 2\}$.

Задача 9.28. Нехай ξ має від'ємний біноміальний розподіл з параметром λ. Знайти $\mathbf{D}\xi$, якщо $\mathrm{P}\{\xi \le 1\} = \mathrm{P}\{\xi > 1\}$.

Задача 9.29. Нехай ξ_1 і ξ_2 незалежні випадкові величини, які мають пуассонівський розподіл з параметрами λ_1 і λ_2. Довести, що:
а) випадкова величина $\eta = \xi_1 + \xi_2$ має розподіл Пуассона з параметром $\lambda_1 + \lambda_2$;
б) розподіл ξ_1 при умові $\xi_1 + \xi_2 = n$ є біноміальним розподілом з параметрами n і $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$, тобто

$$\mathrm{P}(\xi_1 = k | \xi_1 + \xi_2 = n) = C_n^k \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{n-k}, \quad k = 0, 1, ..., n.$$

Задача 9.30. Випадкова величина ξ має розподіл Пуассона з параметром λ. Знайти $\mathbf{M}\frac{1}{1+\xi}$.

Задача 9.31. Випадкова величина ξ має розподіл Пуассона з параметром $\lambda = 1$. Показати, що $\mathbf{M}|\xi - 1| = \frac{2\sqrt{\mathbf{D}\xi}}{e}$.

Задача 9.32. Нехай ξ – випадкова величина, що має розподіл Пуассона, для якої $\mathrm{P}\{\xi = 0\} = \mathrm{P}\{\xi = 1\}$. Чому дорівнює $\mathbf{M}\xi$?

Задача 9.33.* У наявності є η куль, де η – випадкова величина, яка приймає цілі невід'ємні значення. Кожну з η куль кладуть або в урну A, або в урну B. (ймовірності відповідних подій дорівнюють p і $(1-p)$. Нехай η_A – число куль в урні A, а η_B – число куль в урні B. Треба знайти розподіли η_A, η_B і довести, що випадкові величини η_A і η_B незалежні тоді і тільки тоді, коли η має розподіл Пуассона.

Задача 9.34. Знайти математичне сподівання і дисперсію випадкової величини з розподілом Паскаля.

Задача 9.35. $\chi_1, \ldots, \chi_n, \chi_{n+1}$ – незалежні випадкові величини, які приймають два значення 0 і 1 з ймовірностями $\mathrm{P}\{\chi_i = 1\} = p$, $\mathrm{P}\{\chi_i = 0\} = q = 1 - p$, $i = 1, 2, \ldots, n, n+1$. Для $i = 1, 2, \ldots, n$ покладемо

$$\xi_i = \begin{cases} 0, & \text{якщо } \chi_i + \chi_{i+1} \text{ непарне}, \\ 1, & \text{якщо } \chi_i + \chi_{i+1} \text{ парне}. \end{cases}$$

Треба знайти $\mathbf{M}\eta$ і $\mathbf{D}\eta$, де $\eta = \sum\limits_{i=1}^{n} \xi_i$.

Задача 9.36. Урна містить N куль з номерами від 1 до N. Послідовно виймають n куль, повертаючи кожен раз взяту кулю назад. Нехай ξ – найбільший номер, який було при цьому одержано. Знайти розподіл ξ і математичне сподівання.

Задача 9.37.* В урні знаходиться N куль, серед яких M білих. З урни навмання взято n куль, $n \ \leq M$. Нехай ξ – число білих куль серед n. Знайти $\mathbf{M}\xi$ і $\mathbf{D}\xi$.

Задача 9.38. Двоє підкидають монету n раз кожний. Знайти ймовірність того, що у них випаде однакова кількість гербів.

Задача 9.39. Припустимо, що випадкові величини $\xi_1, \xi_2, \ldots, \xi_n$ незалежні і кожна з них приймає значення $+1$ і -1 з ймовірностями p і $q = 1-p$ відповідно. Знайти розподіл випадкової величини $\eta_n = \xi_1 \cdot \xi_2 \cdot \ldots \cdot \xi_n$.

Задача 9.40. Підкидають два гральних кубики. Нехай ξ_1, ξ_2 – числа очок, які випали на першому і другому кубику відповідно. Треба:
а) знайти сумісний розподіл ξ_1 і $\eta = \max(\xi_1, \xi_2)$;
б) обчислити $\mathbf{M}\xi_1$, $\mathbf{D}\xi_1$, $\mathbf{M}\eta$, $\mathbf{D}\eta$ та коваріацію ξ_1 і η.

Задача 9.41. Нехай ξ і η відповідно сума і різниця очок, які з'явилися при підкиданні двох гральних кубиків. Довести, що величини ξ і η залежні, але некорельовані.

Задача 9.42. Нехай α і β незалежні випадкові величини з однаковими розподілами і

$$\xi = \alpha + \beta, \qquad \eta = \alpha - \beta.$$

Довести, що ξ і η некорельовані.

Задача 9.43. Нехай ξ набуває значень ± 1 та ± 2 кожне з ймовірністю $1/4$, а $\eta = \xi^2$. Треба:
а) знайти сумісний розподіл ξ і η; б) довести, що ξ і η залежні, але некорельовані.

Задача 9.44. Нехай випадкові величини ξ і η набувають тільки по два значення. Довести, що в цьому випадку з некорельованості випливає незалежність.

Задача 9.45.* Випадкова величина ξ приймає n різних значень, а випадкова величина η – m різних значень. Довести, що якщо

$$\mathbf{M}\xi^k\eta^s = \mathbf{M}\xi^k\mathbf{M}\eta^s$$

для всіх $k = 1, 2, \ldots n-1$, $s = 1, 2, \ldots, m-1$, то випадкові величини ξ і η є незалежними.

Задача 9.46. Гральний кубик підкидають n раз. Обчислити:
а) ймовірність того, що n_1 раз випаде одиниця, n_2 раз випаде двійка, ... , n_6 раз випаде шістка;
б) ймовірність того, що шістка не випаде жодного разу.

Задача 9.47. У коло вписано правильний трикутник. Навмання у круг кинуто 13 точок. Яка імовірність того, що у кожному з сегментів буде по 3 точки, а всередині трикутника – 4 точки?

Задача 9.48. Вектор $(\nu_1, \ldots, \nu_r)$ має **поліноміальний розподіл**, що означає:

$$\mathrm{P}(\nu_1 = m_1, ..., \nu_r = m_r) = \frac{n!}{m_1! \cdot ... \cdot m_r!} p_1^{m_1} \cdot ... \cdot p_r^{m_r},$$

де $m_i \geq 0$ $(i = 1, ..., r)$ та $m_1 + ... + m_r = n$. Довести, що:
а) $\mathbf{M}\nu_i = np_i$, $\mathbf{D}\nu_i = np_i(1 - p_i)$ $(\ i = 1, ..., r)$;
б) $r(\nu_i, \nu_j) = -\sqrt{\dfrac{p_i p_j}{(1 - p_i)(1 - p_j)}}$, $i \neq j$, де $r(\nu_i, \nu_j)$ – коефіцієнт кореляції.

Задача 9.49. Знайти коефіцієнт кореляції між числом появ одиниці та числом появ шістки при n підкиданнях симетричного кубика.

Задача 9.50. За багаторічними статистичними даними знайдено, що ймовірність народження хлопчиків дорівнює 0.515. Скласти закон розподілу випадкової величини ξ – кількості хлопчиків, які народилися в родині, де четверо дітей. Знайти числові характеристики цієї випадкової величини.

Задача 9.51.* Написані n листів, але адреси на конвертах написано у випадковому порядку. Нехай ξ_n – число листів, які будуть одержані тими адресатами, кому вони призначені. Обчислити $\mathbf{M}\xi_n$ та $\mathbf{D}\xi_n$.

Задача 9.52.* Рекламний агент має за мету переконати m клієнтів придбати товар своєї фірми, для чого він телефонує кожному з них. Ймовірність переконати клієнта за одну телефонну розмову (незалежно від інших спроб та від інших клієнтів) складає p. Якщо клієнта не вдалося переконати за одну розмову, то агент телефонує йому ще раз, доки не переконає його.
а) Знайти розподіл ймовірностей випадкової величини ξ, яка є кількістю всіх телефонних дзвінків, які зробить рекламний агент.
б) Знайти математичне сподівання та дисперсію випадкової величини ξ.
в) Перевірити характеристичну властивість знайденого розподілу, тобто, що знайдені числа невід'ємні та в сумі дають одиницю.

Задача 9.53. Переконання упертого гравця. Браун завжди ставить один долар на номер 13 у американській рулетці всупереч порадам свого товариша. Щоб відвернути Брауна від цієї гри, його товариш пропонує йому парі на 20 доларів, стверджуючи, що Браун не виграє в жодній з 36 ігор. Чи є сенс Брауну прийняти це парі? З якою вигодою чи невигодою? (Американська рулетка має 38 однаково ймовірних номерів. Якщо випадає номер гравця, то він отримує свою ставку назад плюс таку ж суму у 35-кратному розмірі, якщо ні – губить свою ставку).

Задача 9.54. Довести наступне твердження: якщо випадкова величина X має гіпергеометричний розподіл з параметрами n, N, M і при цьому $N \to \infty$, $M \to \infty$ та $\frac{M}{N} \to p$, то для всіх фіксованих m та n
$$\mathrm{P}\left(X = m\right) \longrightarrow C_n^m p^m (1-p)^{n-m}.$$

Задача 9.55. З множини $\{1, 2, ..., n\}$ навмання обирають число η, після чого з множини

$\{1, 2, ..., \eta\}$ навмання обирають число ξ. Знайти $\mathbf{M}\xi$ та $\mathbf{D}\xi$, якщо $n \geq 2$.

Задача 9.56. Авіакомпанія виконує два рейси на добу. Імовірність затримки першого рейсу через технічні причини дорівнює 0.1, а другого – 0.05.

1. Скласти закон розподілу системи (ξ, η), де ξ – кількість затримок першого рейсу, η – сумарна кількість затримок двох рейсів.
2. Записати одновимірні закони розподілів кожної компоненти.
3. Знайти основні числові характеристики системи випадкових величин (ξ, η).
4. Записати умовні закони розподілу $\{\xi|\eta = 1\}$; $\{\eta|\xi = 1\}$.
5. Обчислити відповідні умовні сподівання для компонент системи (ξ, η).

Задача 9.57. Геолог зібрав 10 зразків базальту та 10 зразків граніту та віддав їх в лабораторію на аналіз, попросивши лаборанта проаналізувати 6 випадково обраних зразків.

а) Записати закон розподілу кількості зразків граніту, які потрапили на аналіз.

б) Яка ймовірність того, що на аналіз потрапило однакова кількість зразків граніту та базальту?

в) Яка ймовірність того, що на аналіз потрапили лише зразки одного виду?

Задача 9.58. Урна містить невідому кількість N однакових кульок. Для того, щоб оцінити кількість N, вибирається навмання кулька, маркується та повертається в урну. Знову навмання вибирається кулька з урни. Якщо витягли марковану кульку, експеримент припиняється. В противному випадку кулька повертається в урну і після перемішування процес повторюється до тих пір, доки не буде витягнута маркована куля. Нехай ξ – це кількість спроб. Який розподіл має ξ? Якщо маркована куля була витягнута на k-му кроці, як можна було б оцінити N?

б) Для оцінки можна використати також інший підхід: після повернення маркованої кулі в урну, навмання обирається куля. Якщо вона маркована, експеримент припиняється, якщо ні – з урни навмання обирається друга куля (перша куля в урну не повертається!). Зупиняється процес тоді, коли буде витягнута маркована куля. Нехай η – це кількість спроб. Який розподіл має η?

в) Замість маркування одної кулі, маркується m обраних куль ($m \leq N$). Після повернення маркованих куль в урну та перемішування навмання обирають r куль. Нехай ζ – кількість маркованих куль в цій вибірці. Який розподіл має випадкова величина ζ?

Задача 9.59. Для того, щоб зруйнувати міст, треба двічі влучити в нього з певної зброї. Ймовірність влучення при одному пострілі p. Нехай ξ – кількість пострілів, необхідних для того, щоб зруйнувати міст.

а) Знайти $\mathrm{P}\{\xi = 2\}$, $\mathrm{P}\{\xi = 3\}$, $\mathrm{P}\{\xi = 4\}$.

б) Показати, що $\mathrm{P}\{\xi = k\} = (k-1)p^2(1-p)^{k-2}$, для $k \geq 2$.

Задача 9.60. N осіб сідає в ліфт на нульовому поверсі офісного центру, де офіси розташовані на 21 поверхах. Кожна особа вибирає випадковим чином (з точки зору спостерігача) поверх незалежно від інших з ймовірністю $1/21$. Нехай ξ_N – кількість зупинок, які робить

ліфт, щоб висадити цих осіб. Введемо випадкові величини η_i, $i = 1, ..., 21$, які дорівнюють 1, якщо ліфт зупиняється на i-тому поверсі, і 0, якщо не зупиняється.
а) Кожна з випадкових величин η_i має розподіл Бернуллі з параметром (ймовірністю "успіху") p. Показати, що $p = 1 - \left(\frac{20}{21}\right)^N$.
б) З визначення ξ_N випливає, що $\xi_N = \eta_1 + \eta_2 + ... + \eta_{21}$. Чи можна вважати, що ξ_N має біноміальний розподіл з параметрами $(21, p)$, де p визначений в пункті а)? Відповідь обґрунтувати.
в) Очевидно, якщо $N = 1$, то $\mathrm{P}\{\xi_1 = 1\} = 1$. Показати, що для $N = 2$

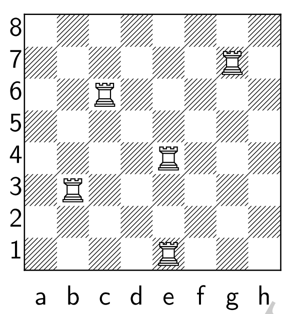
$$\mathrm{P}\{\xi_2 = 1\} = \frac{1}{21} = 1 - \mathrm{P}\{\xi_2 = 2\},$$

і що ξ_3 має розподіл

k	1	2	3
$\mathrm{P}\{\xi_3 = k\}$	$1/441$	$60/441$	$380/441$

Задача 9.61. Тричі випадковим чином обирається число серед чисел 1, 2, 3. Нехай ξ_i, $i = 1, 2, 3$, – результати першого, другого та третього вибору. Знайти розподіл випадкової величини $\overline{\xi} = \dfrac{\xi_1 + \xi_2 + \xi_3}{3}$ – середнього арифметичного значення обраних чисел. Знайти ймовірність того, що рівно два обраних числа дорівнюють 1.

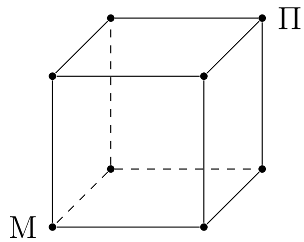
Задача 9.62. Симетричний гральний кубик підкидають до тих пір, доки сума очок, що випали, не досягне 6. Записати розподіл випадкової величини ξ – кількості необхідних підкидань, та знайти її математичне сподівання.



Задача 9.63. Домовимось казати, що одна тура на шаховій дошці б'є іншу туру, якщо вони стоять або на одній горизонталі, або на одній вертикалі, колір тур не важливий. Василь на порожню шахову дошку ставить тури, обираючи кожного разу навмання одне з вільних полів, аж допоки чергова виставлена тура не буде бити одну з раніше поставлених тур. Випадкова величина ξ є кількістю виставлених таким чином тур на шаховій дошці. Записати ряд розподілу випадкової величини ξ, знайти $\mathbf{M}\xi$, $\mathbf{D}\xi$ та $\mathrm{P}\{\xi < 6\}$. На рисунку наведено приклад, коли тури були виставлені у такому порядку: c6, b3, e4, g7, e1. Остання тура (на e1) б'є туру на e4.

Задача 9.64. Гравець, маючи 50 євро, вирішив піти в казино та збільшити свій капітал принаймні до 200 євро. Він вирішив грати лише в рулетку і ставити лише на червоне при кожному обертанні. Яка ймовірність того, що гравець досягне своєї мети, якщо він вирішив на кожному кроці подвоювати ставку незалежно від результатів і ставити послідовно 5, 10, 25, 50 та 100 євро? Знайти розподіл та математичне сподівання кількості грошей, які у нього будуть в наявності після відвідування казино, якщо він уходить за умови, що має принаймні 200 євро або не має достатньо грошей, щоб зробити наступну ставку. Вважати, що червоне випаде з ймовірністю 1/2.

Задача 9.65. Гра має такі правила: гравець кидає симетричний гральний кубик. Якщо на кубику випало i очок, він підкидає одночасно i кубиків. Гравець виграє 100 грн, якщо сума на всіх кубиків буде більша за 12, програє 100 грн, якщо сума буде меншою за 12, і не виграє (програє) нічого, якщо сума дорівнює 12. Знайдіть розподіл та математичне сподівання виграшу гравця.



Задача 9.66.* В одній з вершин куба сидить муха (точка М на рисунку), у протилежній вершині куба чатує павук (точка П на рисунку). Кожного разу муха обирає навмання одне з трьох ребер, інцидентних вершині її перебування і по обраному ребру за 1 секунду переповзає до наступної вершини. Так муха повзає, поки не натрапить на нерухомого павука. Нехай ξ – час існування мухи від початку її мандрування ребрами куба. Скласти ряд розподілу випадкової величини ξ, знайти $\mathbf{M}\xi$, $\mathbf{D}\xi$ та $\mathrm{P}\{\xi \le 7\}$. Примітка: див. також задачу 13.25 на стор. 187.

Задача 9.67.* Грають колодою з 32 карт, масть карти не має значення, важливе лише найменування карти (ранг). Гравець і круп'є отримують по одній карті з колоди. Якщо карта гравця вища рангом за карту круп'є, то гравець виграє суму, яка вдвічі більша за його ставку. Якщо карта гравця нижча рангом за карту круп'є, то гравець програє гроші, які він поставив. Якщо карти однакового рангу, гра продовжується і гравець подвоює свою ставку. Круп'є знову виймає дві карти з колоди: одну для гравця, одну для себе. Якщо карта гравця вища рангом за карту круп'є, то гравець виграє суму, яка вдвічі більша за його стартову ставку, в противному випадку він програє всі гроші (стартову ставку та її підвищення). Знайти математичне сподівання виграшу в цій грі.

Задача 9.68. Знайти генератрису випадкової величини, розподіленої за

а) геометричним законом з параметром p

$$\mathrm{P}\{\xi = n\} = (1-p)^n p, \quad n = 0, 1, ...; \quad p \in (0,1);$$

б) біноміальним законом з параметрами m, p

$$\mathrm{P}\{\xi = n\} = C_m^n p^n (1-p)^{m-n}, \quad n = 0, 1, ..., m, \quad p \in (0,1);$$

в) законом Пуассона з параметром λ

$$\mathrm{P}\{\xi = n\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad n = 0, 1, ...; \quad \lambda > 0;$$

г) дискретним рівномірним законом

$$\mathrm{P}\{\xi = n\} = \frac{1}{N+1}, \quad n = 0, 1, ..., N.$$

Задача 9.69. Знайти послідовності, яким відповідають генератриси

$$\text{а) } \left(1 - s + \frac{1}{4}s^2\right)^{-1}; \quad \text{б) } a(1-ps^l)^{-1}; \quad \text{в) } se^{s^{l\lambda}-\lambda}.$$

Задача 9.70. Нехай ξ – випадкова величина, яка розподілена за геометричним законом з параметром $p \in (0,1)$. Знайти генератриси випадкових величин $\xi_n^+ = \max(N,\xi)$, $\xi_n^- = \min(N,\xi)$.

Задача 9.71. Нехай ξ – цілочисельна випадкова величина, з генератрисою $\psi(s)$. Знайти генератрису випадкової величини $\eta = a\xi + b$, де a, b – невід'ємні цілі числа.

Задача 9.72. Нехай $\phi(s)$ – генератриса випадкової величини ξ. Довести, що генератриса послідовності $\mathrm{P}\{\xi \geq n\}$

$$\psi(s) = \frac{1 - s\phi(s)}{1-s}.$$

Задача 9.73. За допомогою генератрис перевірити властивість стійкості до сумування розподілів випадкових величин, вказаних в задачі 9.68.

Задача 9.74. Послідовність P_n, $n = 0,1,2,...$, будується рекурентно

$$P_{n+2} = aP_{n+1} + bP_n, \;\; n = 0,1,2,...$$

Знайти генератрису цієї послідовності.

Задача 9.75. Нехай u_n – ймовірність того, що за n підкидань симетричного грального кубика шістка випала парну кількість очок ($n = 1,2,\ldots$). Побудувати рекурентне співвідношення для послідовності u_n ($u_0 = 1$) та знайти її генератрису.

Задача 9.76. Сумісний розподіл випадкових величин ξ та η, які набувають цілих невід'ємних значень, задається співвідношенням:

$$\mathrm{P}\{\xi = n, \eta = k\} = \begin{cases} \dfrac{\lambda^n}{n!} e^{-\lambda} C_n^k p^k (1-p)^{n-k}, & n \geq k \geq 0, \\ 0, & k > n, \end{cases}$$

$\lambda > 0$, $p \in (0,1)$. Знайти сумісну генератрису випадкових величин ξ та η.

Задача 9.77. Випадкові величини ξ та η набувають цілих невід'ємних значень, причому ξ має геометричний розподіл з параметром $p \in (0,1)$, а умовний розподіл величини η відносно ξ

$$\mathrm{P}\{\eta = k | \xi = n\} = \begin{cases} u(1-u)^k, \; k = 0,1,2,..., & \text{якщо } n \leq N, \\ v(1-v)^k, \; k = 0,1,2,...,, & \text{якщо } n > N, \end{cases}$$

де $u, v \in (0,1)$. Знайти сумісну генератрису випадкових величин ξ та η.

Задача 9.78. Нехай незалежні випадкові величини ξ та η набувають цілих невід'ємних значень, і їхні генератриси дорівнюють відповідно $\psi_1(s)$ та $\psi_2(s)$. Знайти сумісну генератрису та сумісний розподіл випадкових величин ξ та $\zeta = \xi + \eta$.

Задача 9.79. Знайти ймовірність виродження для гіллястого процесу такого, що кількість нащадків $\xi_1 \in \{0,1,2\}$ (з однаковими ймовірностями).

Задача 9.80. Знайти ймовірність виродження для гіллястого процесу такого, що кількість нащадків $\xi_1 \in \{0, 1, 2, 3\}$ (з однаковими ймовірностями).

Задача 9.81. Нехай популяція бактерій починається з одного представника. У наступному поколінні бактерія може загинути, не лишивши нащадків, з ймовірністю $1/4$, або розщепитися на дві подібні собі бактерії з ймовірністю $3/4$. Яка ймовірність того, що популяція бактерій вимре в третьому поколінні? Що вона вимре коли-небудь? Чому дорівнює ця ймовірність, якщо початкова популяція складалася з двох бактерій?

Задача 9.82. 4 % чоловіків певного регіону не мають дітей. З решти чоловіків чверть має одну дитину, і три чверті – двоє дітей. Кожна дитина з однаковою ймовірністю може бути хлопчиком чи дівчинкою. Якщо прізвище передається лише по чоловічій лінії, яка ймовірність того, що прізвище одного з засновників окремої династії врешті-решт зникне?

10 ВИПАДКОВІ ВЕЛИЧИНИ ТА ФУНКЦІЇ РОЗПОДІЛУ У ЗАГАЛЬНОМУ ВИПАДКУ

Нехай дано ймовірнісний простір $(\Omega, \Re, \mathrm{P})$ і σ-алгебру $\mathscr{B}$ борелівських підмножин з $\mathbb{R} = (-\infty; +\infty)$. Відображення

$$\xi = \xi(\omega) : \ \Omega \to \mathbb{R}$$

називається **випадковою величиною**, якщо $\forall B \in \mathscr{B} \ \ \xi^{-1}(B) \in \Re$, де

$$\xi^{-1}(B) = \{\omega \in \Omega : \ \xi(\omega) \in B\}.$$

Такі відображення в теорії функцій називають вимірними. Тому можемо просто сказати, що випадкова величина – це вимірне відображення вимірного простору $(\Omega, \Re)$ у вимірний простір $(\Re, \mathscr{B})$. Відомо, що для вимірності ξ досить, щоб

$$\forall x \in \mathbb{R} \quad \{\omega \in \Omega : \ \xi(\omega) \le x\} \in \Re,$$

причому знак $\le$ може бути замінено на $<$, $>$ або $\ge$.

Функція $F(x)$, яка задається рівністю

$$F(x) = \mathrm{P}\{\omega : \xi(\omega) \le x\},$$

або скорочено

$$F(x) = \mathrm{P}\{\xi \le x\},$$

називається **функцією розподілу** випадкової величини ξ.

Наведемо для функції розподілу основні результати.

Теорема 10.1 Функція розподілу має наступні характеристичні властивості:

1. Якщо $x < y$, то $F(x) \le F(y)$ (неспадність);
2. $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to \infty} F(x) = 1$ (нормованість);
3. $\lim\limits_{x \to x_0 + 0} F(x) = F(x_0)$ (неперервність справа);

Якщо для деякої функції $F(x)$, заданої на $\mathbb{R}$, справедливі пункти 1-3 попередньої теореми, то існує ймовірнісний простір і випадкова величина, задана на цьому просторі, такі, що $F(x)$ буде функцією розподілу цієї випадкової величини.

Якщо в точці $x = x_0$ функція розподілу $F(x)$ має стрибок, тобто

$$F(x_0) - F(x_0 - 0) = p_0 > 0,$$

то говорять що $F(x)$ має **атом** величини p_0 в точці x_0. Функція розподілу $F(x)$ випадкової величини ξ має атом величини p_0 в точці x_0 тоді і лише тоді, коли $\mathrm{P}(\xi = x_0) = p_0$. Якщо функція розподілу $F(x)$ не має атомів, тобто є неперервною, то говорять, що $F(x)$ є **неперервного типу**. Відомо, що функція розподілу $F(x)$ випадкової величини ξ є неперервного типу тоді і лише тоді, коли $\mathrm{P}(\xi = x) = 0$ для кожного $x \in \mathbb{R}$.

Оскільки кожна неспадна функція може мати не більше ніж зліченну кількість точок розриву, то кожна функція розподілу може мати не більше ніж зліченну кількість атомів.

Нехай $A = \{x_i,\ i = 1.2, \ldots\}$ будуть усі атоми функції розподілу $F(x)$, p_i, $i = 1, 2, \ldots$ – відповідні величини цих атомів. Якщо

$$\sum_{i=1}^{\infty} p_i = 1,$$

то говоримо, що функція розподілу $F(x)$ є **дискретного типу**. Для функції розподілу дискретного типу можемо записати

$$F(x) = \sum_{i:\, x_i \le x} p_i.$$

Ця функція розподілу є сходинкоподібною (стрибкоподібною) функцією: $F(x)$ має стрибки в точках x_i, які дорівнюють $p_i = F(x_i) - F(x_i - 0)$.

Математичним сподіванням випадкової величини ξ називається число

$$\mathbf{M}\xi = \int\limits_{\Omega} \xi(\omega) \mathrm{P}(d\omega), \tag{10.1}$$

якщо інтеграл Лебега, який стоїть в правій частині рівності, існує. Тоді, як відомо, математичне сподівання може бути подане і інтегралом Рімана-Стілтьєса:

$$\mathbf{M}\xi = \int\limits_{-\infty}^{+\infty} x dF(x). \tag{10.2}$$

Дисперсія випадкової величини визначається рівністю

$$\mathbf{D}\xi = \mathbf{M}(\xi - \mathbf{M}\xi)^2 = \int\limits_{-\infty}^{+\infty} (x - \mathbf{M}\xi)^2 dF(x). \tag{10.3}$$

Для математичного сподівання та дисперсії випадкових величин у загальному випадку справедливі ті ж самі властивості, що і для дискретних випадкових величин. Для зручності користування наведемо їх ще раз, припускаючи, що всі приведені нижче математичні сподівання скінченні.

1. $\mathbf{M}c = c$;

2. $\mathbf{M}(a\xi + b\eta) = a\mathbf{M}\xi + b\mathbf{M}\eta$ для довільних чисел a, b;

3. Якщо $\mathrm{P}(\xi \le \eta) = 1$, то $\mathbf{M}\xi \le \mathbf{M}\eta$;

4. $|\mathbf{M}\xi| \le \mathbf{M}|\xi|$;

5. $\mathbf{D}c = 0$;

6. Якщо $\mathbf{D}\xi = 0$, то $\mathrm{P}(\xi = \mathbf{M}\xi) = 1$.

Кажуть, що функція розподілу $F(x)$ випадкової величини ξ є **абсолютно неперервного типу**, якщо існує вимірна функція $f(x)$ така, що $\forall x \in \mathbb{R}$

$$F(x) = \int\limits_{-\infty}^{x} f(y)dy. \tag{10.4}$$

Функція $f(x)$ називається **щільністю** функції розподілу $F(x)$.

Вимірна функція $f(x)$, $-\infty < x < \infty$ є щільністю деякої функції розподілу тоді і лише тоді, коли виконуються наступні умови

$$\text{а) } f(x) \ge 0 \text{ майже для всіх } x \in \mathbb{R}; \qquad \text{б) } \int\limits_{-\infty}^{\infty} f(x)dx = 1.$$

Для будь-якої борелівської множини A справедлива формула

$$\mathrm{P}(\xi \in A) = \int\limits_{A} f(x)dx.$$

Якщо щільність $f(x)$ є неперервною в точці x_0, то

$$f(x_0) = \frac{dF(x)}{dx}\,\Big|_{x=x_0} .$$

В загальному випадку ця формула справедлива для майже всіх (відносно міри Лебега в просторі $\mathbb{R}$) точок x.

Точка x називається **точкою росту** функції розподілу, якщо для будь-якого $\varepsilon > 0$

$$F(x+\varepsilon) - F(x-\varepsilon) > 0.$$

Якщо функція розподілу неперервна і множина всіх точок росту цієї функції має міру Лебега нуль, то така функція розподілу називається **сингулярною**. Для сингулярної функції розподілу маємо

$$\frac{dF(x)}{dx} = 0$$

для майже всіх (за мірою Лебега) точок $x \in \mathbb{R}$.

Теорема 10.2 (Лебега) Кожна функція розподілу $F(x)$ може бути єдиним чином подана у вигляді

$$F(x) = qF_d(x) + pF_{ac}(x) + rF_s(x),$$

де $0 \le q, p, r \le 1$, $q + p + r = 1$, а $F_d(x)$, $F_{ac}(x)$, $F_s(x)$ є відповідно функції розподілу дискретного, абсолютно неперервного та сингулярного типу.

Наведемо деякі приклади абсолютно неперервних розподілів (наводимо формули лише для щільності).

1. **Рівномірний розподіл** на відрізку $[a, b]$:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b; \\ 0, & \text{інакше.} \end{cases}$$

Для скорочення цей розподіл часто позначають як $U([a, b])$ або $U(a, b)$. Фразу "*випадкова величина* ξ *має рівномірний розподіл на відрізку* $[a, b]$" можна скорочено записати як $\xi \sim U([a, b])$.

2. **Показниковий (або експоненціальний) розподіл** з параметром $\lambda > 0$:

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

Цей розподіл часто позначають як $Exp(\lambda)$ або $\epsilon(\lambda)$. Для випадкової величини ξ з таким розподілом можна писати $\xi \sim Exp(\lambda)$.

3. **Нормальний (або гауссівський) розподіл** з параметрами m і σ^2:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(x-m)^2}{2\sigma^2}}, \quad \sigma > 0.$$

Позначення: $N(m, \sigma^2)$, хоча інколи можна зустріти позначення $N(m, \sigma)$. Для випадкової величини ξ з розподілом $N(m, \sigma^2)$ можна писати $\xi \sim N(m, \sigma^2)$.
Гауссівський розподіл з параметрами $m = 0$ і $\sigma^2 = 1$ називається **стандартним гауссівським розподілом**.

4. **Розподіл Коші** з параметрами a та $b > 0$:

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

Позначення: $C(a, b)$. Для випадкової величини ξ з таким розподілом можна писати $\xi \sim C(a, b)$.

Якщо випадкова величина ξ має експоненціальний розподіл, то для довільних $x \geq 0$ та $y \geq 0$має місце така властивість

$$P(\xi > x + y | \, \xi > x) = P(\xi > y),$$

яка називається **відсутністю пам'яті** експоненціального розподілу. Серед неперервних розподілів лише експоненціальний має таку властивість.

Модою випадкової величини називається точка D, в якій щільність має максимум. Якщо мода єдина, то функція розподілу називається **унімодальною**.

Квантиллю k_p порядку p функції розподілу $F(x)$ називається число

$$k_p = \inf\{x : F(x) \geq p\}.$$

Якщо функція розподілу є неперервною, тоді квантиль k_p є найменшим коренем рівняння $F(x) = p$.

Квантиль порядку $1/2$ називається **медіаною** випадкової величини ξ або функції розподілу $F(x)$ і позначається m_0.

Функція

$$\varphi(t) = \mathbf{M}e^{it\xi} = \int_{-\infty}^{\infty} e^{itx} dF(x), \qquad t \in (-\infty, \infty).$$

називається **характеристичною функцією** випадкової величини ξ.

Теорема 10.3 Характеристична функція має наступні властивості

1) $|\varphi(t)| \leq 1$, $\varphi(0) = 1$;

2) $\varphi_{a\xi+b}(t) = e^{itb}\varphi(at)$, для довільних чисел a, b;

3) функція $\varphi(t)$ рівномірно неперервна відносно t.

Кожній характеристичній функції відповідає лише одна функція розподілу і якщо $\varphi(t)$ характеристична функція для функції розподілу $F(x)$, то має місце така формула

$$F(y) - F(x) = \frac{1}{2\pi} \lim_{A \to \infty} \int_{-A}^{A} \frac{e^{-ity} - e^{-itx}}{it} \varphi(t) dt,$$

де $x < y$ є точками неперервності функції $F(\cdot)$.

Приклад 10.1

Нехай майнові збитки описуються випадковою величиною ξ, яка має розподіл Парето з параметрами $\alpha > 0$ та $\lambda > 0$.
а) Перевірити умову нормування щільності розподілу Парето.
б) Знайти математичне сподівання можливих збитків.
в) Знайти ймовірність того, що збитки перевищать рівень $c > \alpha$.
г) Знайти ймовірність того, що збитки будуть у відрізку $[c, d]$, $\alpha < c < d$.

Розв'язок: **Розподіл Парето** є абсолютно неперервним розподілом, щільність якого має вигляд:

$$f(x) = \begin{cases} \dfrac{\lambda}{\alpha}\left(\dfrac{\alpha}{x}\right)^{\lambda+1}, & x \geq \alpha, \\ 0\,, & x < \alpha. \end{cases}$$

а) Перевіримо, чи інтеграл від щільності по всій області значень дорівнює одиниці:

$$\int_{\alpha}^{+\infty} \frac{\lambda}{\alpha}\left(\frac{\alpha}{x}\right)^{\lambda+1} dx = \alpha^{\lambda}\lambda \int_{\alpha}^{+\infty} x^{-(\lambda+1)} dx = \alpha^{\lambda}(-x^{-\lambda})\Big|_{\alpha}^{\infty} = 1.$$

б) За визначенням математичного сподівання маємо

$$\mathbf{M}\xi = \int\limits_{\alpha}^{+\infty} x\, f(x)\, dx = \alpha^{\lambda}\lambda \int\limits_{\alpha}^{+\infty} x^{\,-\lambda} dx = \lambda\alpha^{\lambda}\frac{1}{1-\lambda}\; x^{1-\lambda} \quad \Big|_{\alpha}^{+\infty} = \frac{\lambda}{\lambda - 1}\alpha,$$

якщо значення параметра $\lambda > 1$. Якщо $\lambda \leq 1$, то значення інтегралу в $+\infty$ дорівнює нескінченності, і скінченне математичне сподівання в цьому випадку не існує.

в) Ймовірність того, що збитки перевищать заданий рівень c:

$$\mathrm{P}\{\xi > c\} = \int_{c}^{+\infty} \frac{\lambda}{\alpha}\left(\frac{\alpha}{x}\right)^{\lambda+1} dx = \alpha^{\lambda}\lambda \int_{c}^{+\infty} x^{-(\lambda+1)} dx = \alpha^{\lambda}(-x^{-\lambda})\Big|_{c}^{\infty} = \left(\frac{\alpha}{c}\right)^{\lambda}.$$

г) Аналогічно п. в), маємо

$$\mathrm{P}\{c < \xi \leq d\} = \int_{c}^{d} \frac{\lambda}{\alpha}\left(\frac{\alpha}{x}\right)^{\lambda+1} dx = \alpha^{\lambda}(-x^{-\lambda})\Big|_{c}^{d} = \left(\frac{\alpha}{c}\right)^{\lambda} - \left(\frac{\alpha}{d}\right)^{\lambda}.$$

Відповідь: б) $\mathbf{M}\xi = \dfrac{\lambda}{\lambda - 1}\alpha$; в) $\left(\dfrac{\alpha}{c}\right)^{\lambda}$; г) $\left(\dfrac{\alpha}{c}\right)^{\lambda} - \left(\dfrac{\alpha}{d}\right)^{\lambda}$.

Приклад 10.2

Довести, що
$$\mathbf{D}\xi = \mathbf{M}(\xi - \mathbf{M}\xi)^2 = \min_a \mathbf{M}(\xi - a)^2.$$

Доведення:

$$\mathbf{M}(\xi - a)^2 = \mathbf{M}((\xi - \mathbf{M}\xi) + (\mathbf{M}\xi - a))^2 =$$

$$= \mathbf{D}\xi + (\mathbf{M}\xi - a)^2 + 2\mathbf{M}(\xi - \mathbf{M}\xi)(\mathbf{M}\xi - a) =$$

$$= \mathbf{D}\xi + (\mathbf{M}\xi - a)^2 \geq \mathbf{D}\xi,$$

причому рівність досягається тільки при $a = \mathbf{M}\xi$, що і треба було довести.

Приклад 10.3

Довести, що якщо випадкові величини $\xi_1, \xi_2, ..., \xi_n$ додатні, незалежні та однаково розподілені, то
$$\mathbf{M}\left\{\frac{\xi_1}{\xi_1 + \xi_2 + ... + \xi_n}\right\} = \frac{1}{n}.$$

Доведення: Розглянемо випадкові величини

$$\eta_1 = \frac{\xi_1}{\xi_1 + \xi_2 + ... + \xi_n},\ \eta_2 = \frac{\xi_2}{\xi_1 + \xi_2 + ... + \xi_n},\ ...,\ \eta_n = \frac{\xi_n}{\xi_1 + \xi_2 + ... + \xi_n}.$$

Оскільки за умовою $\xi_1, \xi_2, ..., \xi_n$ є однаково розподіленими, то й $\eta_1, \eta_2, ..., \eta_n$ є однаково розподіленими. Зокрема, їхні математичні сподівання однакові. Розглянемо математичне сподівання суми введених величин. Згідно властивостям математичного сподівання, маємо
$$\sum_{i=1}^{n} \mathbf{M}\eta_i = \mathbf{M}\left\{\sum_{i=1}^{n} \eta_i\right\} = \mathbf{M}\left\{\sum_{i=1}^{n} \frac{\xi_i}{\sum_{i=1}^{n} \xi_i}\right\} = \mathbf{M}\left\{\frac{\sum_{i=1}^{n} \xi_i}{\sum_{i=1}^{n} \xi_i}\right\} = 1.$$
Отже, маємо
$$\sum_{i=1}^{n} \mathbf{M}\eta_i = n\mathbf{M}\eta_1 = 1, \text{ звідки } \mathbf{M}\eta_1 = \frac{1}{n},$$
що і треба було довести.

Приклад 10.4

На відрізок $[0;1]$ навмання кидають точку, яка ділить його на дві частини. Знайти функцію розподілу меншої частини відрізку.

Розв'язок: Нехай ξ – координата цієї точки. Тоді довжина меншої частини відрізку η дорівнює $\min\{\xi, 1-\xi\}$.

Зауважимо, що $F_\eta(x) = 0$ при $x < 0$, і $F_\eta(x) = 1$ при $x \geq 1/2$.

Для $0 \leq x < 1/2$ маємо

$$F_\eta(x) = \mathrm{P}\{\eta \leq x\} = \mathrm{P}\{\min\{\xi, 1-\xi\} \leq x\} = 1 - \mathrm{P}\{\min\{\xi, 1-\xi\} > x\} =$$
$$= 1 - \mathrm{P}\{\xi > x, 1-\xi > x\} = 1 - \mathrm{P}\{x < \xi < 1-x\}.$$

Ймовірність $\mathrm{P}\{x < \xi < 1-x\}$ можна знайти як різницю функцій розподілу рівномірного закону або як геометричну ймовірність (всі можливі положення точки рівноможливі): $\mathrm{P}\{x < \xi < 1-x\} = 1 - 2x$.

Отже, остаточно маємо:

$$F_\eta(x) = \begin{cases} 0, & x < 0; \\ 2x, & x \in [0, 1/2); \\ 1, & x \geq 1/2. \end{cases}$$

Приклад 10.5

Щільність розподілу випадкової величини ξ має вид

$$f_\xi(x) = \begin{cases} x - \frac{1}{2}, & \text{якщо } x \in (1,2]; \\ 0, & \text{якщо } x \notin (1,2]. \end{cases}$$

Знайти функцію розподілу та медіану розподілу.

Розв'язок: За визначенням, $F_\xi(x) = \mathrm{P}\{\xi \leq x\} = \int_{-\infty}^{x} f_\xi(u)du$. Очевидно, ліворуч і праворуч від проміжку $(1,2]$ функція розподілу дорівнює 0 та 1 відповідно. Для значень x, які належать проміжку $(1,2]$, значення $F_\xi(x)$ можна знайти або геометрично (як відношення площ трапецій з висотами $x-1$ та $2-1=1$), або як інтеграл:

$$F_\xi(x) = \int_1^x \left(u - \frac{1}{2}\right) du = \frac{x^2 - x}{2}.$$

Отже, остаточно

$$F_\xi(x) = \begin{cases} 0, & \text{якщо } x \le 1; \\ \dfrac{x^2 - x}{2}, & \text{якщо } 1 < x \le 2; \\ 1, & \text{якщо } x > 2. \end{cases}$$

Знайдемо медіану розподілу. За визначенням це – значення $x_{1/2}$, для якого $\mathrm{P}\{\xi \le x_{1/2}\} = \mathrm{P}\{\xi > x_{1/2}\}$, тобто $F_\xi(x_{1/2}) = 1 - F_\xi(x_{1/2}) = 1/2$.

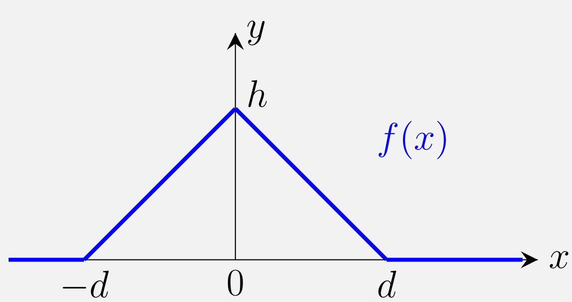
Для нашого випадку медіана буде розв'язком рівняння

$$\frac{x_{1/2}^2 - x_{1/2}}{2} = \frac{1}{2},$$

звідки $x_{1/2} = (1+\sqrt{5})/2$. Другий корінь квадратного рівняння від'ємний, отже, він не може бути медіаною.

Приклад 10.6

Випадкова величина ξ розподілена за **законом Сімпсона**, графік щільності якого має вид



Знайти аналітичний вираз для щільності, функцію розподілу, $\mathbf{M}\xi$ та $\mathbf{D}\xi$.

Розв'язок: будь-яка функція може бути щільністю розподілу, якщо вона задовольняє дві властивості: $f_\xi(x) \ge 0$ та $\int_{-\infty}^{+\infty} f_\xi(x)dx = 1$. Для функції, зображеної на графіку, перша властивість очевидна. Геометричний сенс другої властивості полягає в тому, що площа фігури, обмеженої щільністю, дорівнює одиниці. В нашому випадку фігурою є трикутник, площа якого дорівнює $d \cdot h$. Відповідно, з умови нормування $d \cdot h = 1$ маємо, що лише якщо $h = 1/d$, то зображена функція буде функцією щільності.

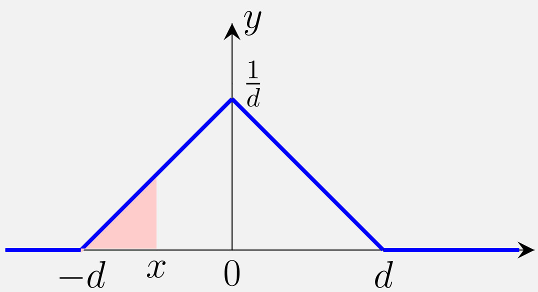
Тепер знайдемо аналітичний вираз щільності. $f_\xi(x) = 0$ при $x < -d$ та $x > d$.

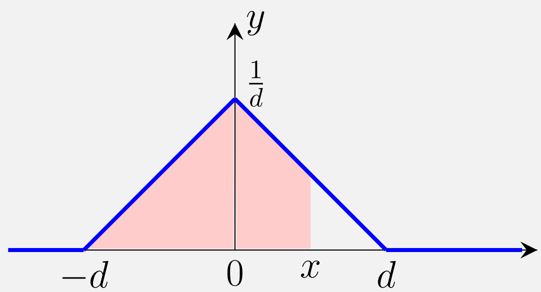
При $x \in [-d, 0]$ щільністю є пряма, неважко знайти її рівняння: $f_\xi(x) = \dfrac{1}{d}\left(1 + \dfrac{x}{d}\right)$. Аналогічно для $x \in [0, d]$: $f_\xi(x) = \dfrac{1}{d}\left(1 - \dfrac{x}{d}\right)$.

Отже, остаточно маємо

$$f_\xi(x) = \begin{cases} \dfrac{1}{d}\left(1 - \dfrac{|x|}{d}\right), & \text{якщо } x \in [-d, d]; \\ 0, & \text{якщо } x \notin [-d, d]. \end{cases}$$

Функцію розподілу в даному випадку простіше знайти не обчислюючи інтеграл, а геометрично. Очевидно, її значення на $(-\infty, -d)$ дорівнює 0, на $(d, +\infty)$ – 1. Якщо $x \in [-d, 0]$, то функція розподілу, $F_\xi(x) = \mathrm{P}\{\xi \le x\}$, є площею трикутника, який лежить лівіше за x.





З подібності трикутників та з того, що половина графіка під щільністю (над проміжком $[-d, 0]$) дорівнює $1/2$, можемо легко обчислити його площу: $F_\xi(x) = \frac{1}{2}\left(\frac{d+x}{d}\right)^2$, для $x \in [0, d]$ потрібна площа – це $F_\xi(x) = 1 - \frac{1}{2}\left(\frac{d-x}{d}\right)^2$. Остаточно

$$F_\xi(x) = \begin{cases} 0, & \text{якщо } x \in (-\infty, -d); \\ \dfrac{1}{2}\left(\dfrac{d+x}{d}\right)^2, & \text{якщо } x \in [-d, 0]; \\ 1 - \dfrac{1}{2}\left(\dfrac{d-x}{d}\right)^2, & \text{якщо } x \in (0, d]; \\ 1, & \text{якщо } x \in (d, +\infty). \end{cases}$$

Математичне сподівання заданої випадкової величини, очевидно, дорівнює нулю, оскільки графік щільності симетричний відносно вісі ординат. Для того, щоб знайти дисперсію, нам треба знайти $\mathbf{M}\xi^2$. Його ми знайдемо за визначенням, з урахуванням симетричності підінтегральної функції:

$$\mathbf{D}\xi = \mathbf{M}\xi^2 = \int_{-d}^{d} x^2 \frac{1}{d}\left(1 - \frac{|x|}{d}\right) dx = 2\int_{0}^{d} x^2 \frac{1}{d}\left(1 - \frac{x}{d}\right) dx = \left(\frac{2x^3}{3d} - \frac{2x^4}{4d^2}\right)\Big|_0^d = \frac{d^2}{6}.$$

Приклад 10.7

Випадкова величина ξ має щільність розподілу

$$f_\xi(x) = \begin{cases} e^{-x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Знайти щільність розподілу випадкової величини $\eta = \sqrt{\xi}$.

Розв'язок: очевидно, що $\eta \geq 0$, тому при $x < 0$ $f_\eta(x) = 0$. Якщо ж $x \geq 0$, то $F_\eta(x) = \mathrm{P}\left\{\eta \leq x\right\} = \mathrm{P}\left\{\sqrt{\xi} \leq x\right\} = \mathrm{P}\left\{\xi \leq x^2\right\} = F_\xi\left(x^2\right)$, звідки

$$f_\eta(x) = \frac{d}{dx}F_\eta(x) = \frac{d}{dx}F_\xi\left(x^2\right) = f_\xi\left(x^2\right)\frac{d}{dx}\left(x^2\right) = 2xf_\xi\left(x^2\right) = 2xe^{-x^2}.$$

Остаточно маємо:

$$f_\eta(x) = \begin{cases} 2xe^{-x^2}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Приклад 10.8

Випадкова величина ξ має рівномірний розподіл на відрізку $\left[-\frac{\pi}{2};\ \frac{\pi}{2}\right]$. Знайти щільність розподілу випадкової величини $\eta = \cos\left(\xi\right)$.

Розв'язок: очевидно, що при $x \in \left[-\frac{\pi}{2};\ \frac{\pi}{2}\right]$ $\eta = \cos\left(\xi\right) \in [0;\ 1]$, тому шукана щільність матиме нульові значення поза відрізком $[0;\ 1]$. Якщо ж $x \in [0;\ 1]$, то

$$F_\eta(x) = \mathrm{P}\left\{\eta \leq x\right\} = \mathrm{P}\left\{\cos\left(\xi\right) \leq x\right\} =$$

$$= \mathrm{P}\left\{-\frac{\pi}{2} \leq \xi \leq -\arccos(x)\right\} + \mathrm{P}\left\{\arccos(x) \leq \xi \leq \frac{\pi}{2}\right\} =$$

$$= \int\limits_{-\frac{\pi}{2}}^{-\arccos(x)} \frac{1}{\pi} dx + \int\limits_{\arccos(x)}^{\frac{\pi}{2}} \frac{1}{\pi} dx = 1 - \frac{2}{\pi}\arccos(x).$$

Остаточно маємо:

$$f_\eta(x) = \frac{d}{dx}F_\eta(x) = \begin{cases} \dfrac{2}{\pi\sqrt{1-x^2}}, & \text{якщо } x \in [0;\ 1), \\ 0, & \text{якщо } x \notin [0;\ 1). \end{cases}$$

Приклад 10.9

Нехай випадкова величина ξ має показниковий розподіл з параметром $\lambda > 0$. Знайти функцію розподілу і щільність розподілу випадкових величин а) $\eta = 3 - \xi$, б) $\zeta = \ln \xi$.

Розв'язок: Функція розподілу показникового закону задається рівнянням

$$F_\xi(x) = \mathrm{P}\{\xi \le x\} = \begin{cases} 1 - e^{-\lambda x}, & \text{якщо } x \ge 0; \\ 0, & \text{якщо } x < 0. \end{cases}$$

Для того, щоб знайти функції розподілу вказаних величин, запишемо за визначенням їхні функцію розподілу, виразимо через відому нам функцію розподілу випадкової величини ξ. Для знаходження щільностей треба взяти похідну від отриманих функції розподілу.

а) Отже, за визначенням,

$$F_\eta(x) = \mathrm{P}\{\eta \le x\} = \mathrm{P}\{3 - \xi \le x\} = \mathrm{P}\{\xi \ge 3 - x\} =$$

$$= 1 - \mathrm{P}\{\xi < 3 - x\} = 1 - F_\xi(3 - x) = e^{-\lambda(3-x)},$$

для $3 - x \ge 0$. Запишемо остаточно функцію розподілу η :

$$F_\eta(x) = \begin{cases} e^{-\lambda(3-x)}, & \text{якщо } x \le 3; \\ 1, & \text{якщо } x > 3. \end{cases}$$

Відповідно щільність розподілу η :

$$f_\eta(x) = \begin{cases} \lambda e^{-\lambda(3-x)}, & \text{якщо } x \le 3; \\ 0, & \text{якщо } x > 3. \end{cases}$$

б) Для випадкової величини ζ функція розподілу за визначенням дорівнює

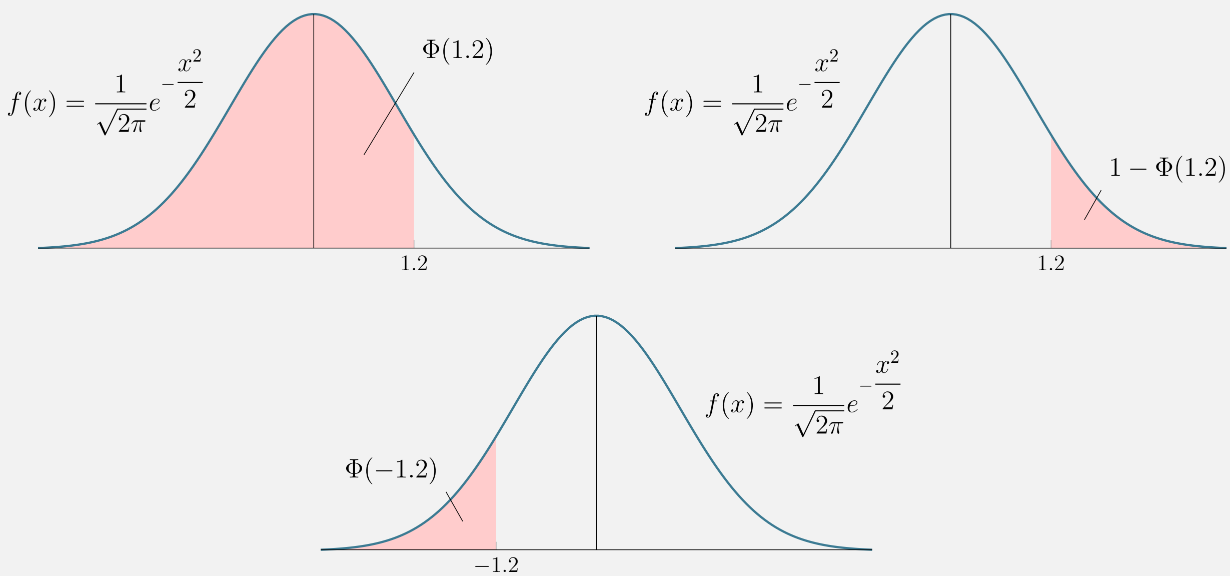
$$F_\zeta(x) = \mathrm{P}\{\zeta \le x\} = \mathrm{P}\{\ln\xi \le x\} = \mathrm{P}\{\xi \le e^x\} = F_\xi(e^x) = 1 - e^{-\lambda e^x},$$

а щільність розподілу

$$f_\zeta(x) = F'_\zeta(x) = (1 - e^{-\lambda e^x})' = \lambda e^{x-\lambda e^x}.$$

Приклад 10.10

Для стандартної нормальної випадкової величини η визначимо ймовірності $\mathrm{P}\{\eta \le 1.2\}$, $\mathrm{P}\{\eta > 1.2\}$ та $\mathrm{P}\{\eta \le -1.2\}$.



В таблиці значень функції розподілу стандартного нормального закону (див. табл. 3 на стор. 302) знаходимо значення $\Phi(1.2) = 0.8849$. Відповідно, маємо:

$$\mathrm{P}\{\eta \le 1.2\} = 0.8849, \quad \mathrm{P}\{\eta > 1.2\} = \mathrm{P}\{\eta \le -1.2\} = 1 - 0.8849 = 0.1151.$$

Приклад 10.11

Для того, щоб знайти подібні ймовірності для довільної нормальної випадкової величини з параметрами m, σ^2, скористаємось співвідношенням між довільною та стандартною нормальними величинами та таблицею значень функції розподілу стандартного нормального закону.

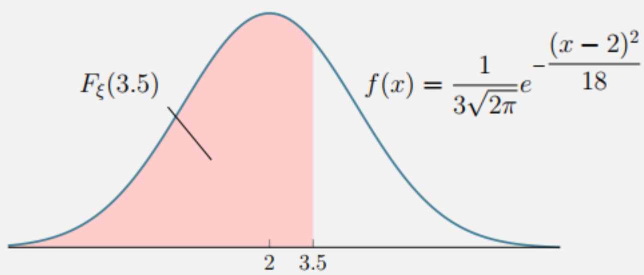
Наприклад, для $\xi \sim N(2, 9)$

$$F_\xi(x) = \Phi\left(\frac{x-m}{\sigma}\right) = \Phi\left(\frac{x-2}{\sqrt{9}}\right) = \Phi\left(\frac{x-2}{3}\right),$$

відповідно

$$\mathrm{P}\{\xi \le 3.5\} = F_\xi(3.5) = \Phi\left(\frac{3.5-2}{3}\right) = \Phi(0.5) \approx 0.6915,$$

$$\mathrm{P}\{\xi \le -3.5\} = \Phi\left(\frac{-3.5-2}{3}\right) \approx \Phi(-1.8333) \approx 0.0333.$$



Приклад 10.12

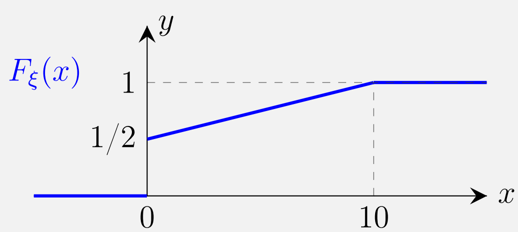
Автомобіль заїжджає на АЗС з метою заправитись пальним. З ймовірністю $1/2$ буде хоча б одна вільна колонка і водій одразу почне заправляти авто. Відповідно час очікування в черзі дорівнює 0 з ймовірністю $1/2$. Якщо все ж доводиться чекати в черзі, час очікування є рівномірно розподіленою випадковою величиною на відрізку $[0, 10]$. Записати функцію розподілу часу очікування і знайти ймовірність того, що доведеться чекати більше 5 хвилин. Обчислити математичне сподівання часу очікування.

Розв'язок: Випадкова величина ξ, час очікування, є випадковою величиною змішаного типу. Позначимо через η випадкову величину, яка має рівномірний розподіл на відрізку $[0,10]$. Тоді можемо виписати вигляд функції розподілу ξ, використовуючи формулу повної ймовірності:

$$F_\xi(x) = \mathrm{P}\{\xi \le x\} = \begin{cases} 0, & \text{з ймовірністю } 1/2; \\ \mathrm{P}\{\eta \le x\} = F_\eta(x), & \text{з ймовірністю } 1/2; \end{cases} =$$

$$= \begin{cases} 0, & \text{якщо } x < 0 \\ \frac{1}{2} + \frac{1}{2}\int_0^x \frac{1}{10}du, & \text{якщо } x \in [0,10]; \\ 1, & \text{якщо } x > 10; \end{cases} = \begin{cases} 0, & \text{якщо } x < 0 \\ \frac{1}{2} + \frac{x}{20}, & \text{якщо } x \in [0,10]; \\ 1, & \text{якщо } x > 10. \end{cases}$$

Ймовірність того, що не доведеться чекати більше 5 хвилин, дорівнює $\mathrm{P}\{\xi > 5\} = 1 - F_\xi(5) = 1 - 3/4 = 1/4$.



Математичне сподівання часу очікування не може бути обчислене за допомогою базової формули, оскільки випадкова величина ξ не є абсолютно неперервною.

В таких випадках його можна знайти за більш загальною формулою:

$$\mathbf{M}\xi = \int_0^{+\infty} [1 - F_\xi(x)]dx - \int_{-\infty}^{0} F_\xi(x)dx.$$

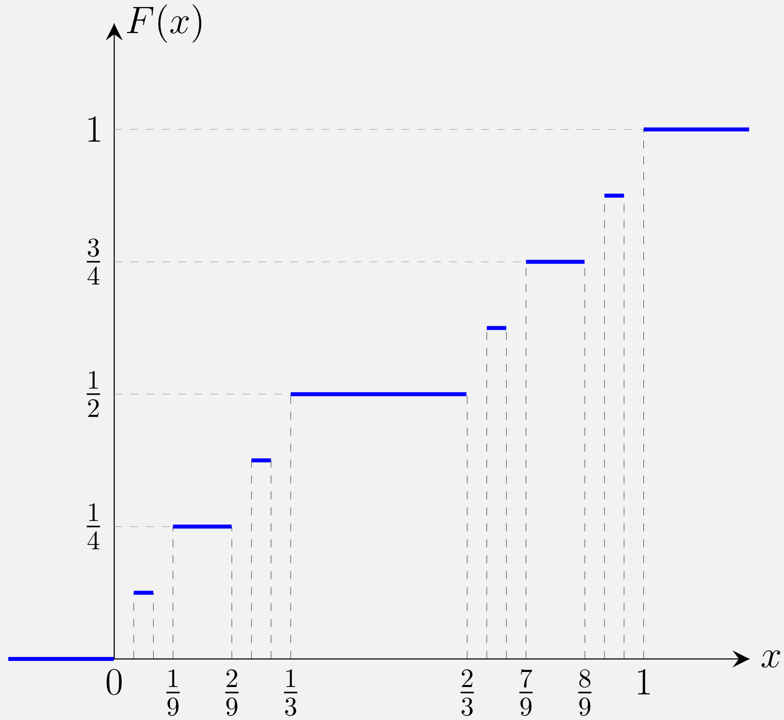
Для нашого випадку

$$\mathbf{M}\xi = \int_0^{10} \left[1 - \left(\frac{1}{2} + \frac{x}{20}\right)\right] dx = 2.5.$$

Отже, в середньому доведеться чекати 2.5 хвилини.

Приклад 10.13

Прикладом сингулярної функції розподілу є **кривa Кантора** $F(x)$. Для $x < 0$ $F(x) = 0$ і $F(x) = 1$ при $x > 1$. На відрізку $[0.1]$ $F(x)$ будується наступним чином.



Відрізок $[0,1]$ розбивається на три рівні частини $[0,1/3]$, $[1/3,2/3]$, $[2/3,1]$. На внутрішньому сегменті покладемо $F(x) = 1/2$. Два сегменти, що залишилися, знову розбиваємо на три рівні частини кожний і на внутрішніх сегментах $F(x)$ задаємо рівною $1/4$ і $3/4$ відповідно. Кожен з тих сегментів, що залишилися, знову ділимо на три рівні частини і на внутрішніх сегментах $F(x)$ визначаємо як сталу, що дорівнює середньому арифметичному між сусідніми значеннями, що вже визначені для $F(x)$, і т.д. В точках, що не належать внутрішнім сегментам, $F(x)$ визначимо за властивістю неперервності.

Неважко бачити, що сумарна довжина внутрішніх сегментів, на яких $F(x)$ стала, дорівнює

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + ... + \frac{2^n}{3^{n+1}} + ... = \frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^n = \frac{1}{3}\frac{1}{1-\frac{2}{3}} = 1.$$

Таким чином, функція розподілу $F(x)$ зростає на множині міри Лебега 0, але без стрибків: $\dfrac{dF(x)}{dx} = 0$ майже всюди.

Задача 10.1. Побудуємо ймовірнісний простір $(\Omega,\ \Re,\ \mathrm{P})$ наступним чином: Ω є квадрат на площині XOY з вершинами (0.0), (0.1), (1.0), (1.1); U – множина усіх підмножин квадрату Ω, для яких поняття площі має зміст; $\mathrm{P}(\cdot)$ визначена на U і дорівнює площі підмножини. Елементарні події $\omega \in \Omega$ будемо позначати $\omega = (x, y)$. Треба знайти функції розподілу і щільності випадкових величин, заданих на ймовірнісному просторі $(\Omega,\ \Re,\ \mathrm{P})$ та побудувати їхні графіки:

a) $\xi(\omega) = x + y$;

b) $\xi(\omega) = x - y$;

c) $\xi(\omega) = xy$;

d) $\xi(\omega) = x^2 + y^2$;

e) $\xi(\omega) = x/y$;

f) $\xi(\omega) = \min(x, y)$;

g) $\xi(\omega) = \max(x, y)$;

h) $\xi(\omega) = |x - y|$;

i) $\xi(\omega)$ є відстань точки (x, y) до діагоналі, що сполучає точки (0,0) і (1,1).

Задача 10.2. Які з поданих нижче функцій є функціями розподілу?

a) $F(x) = \dfrac{3}{4} + \dfrac{1}{2\pi} arctg x$;

в) $F(x) = e^{-e^{-x}}$;

c) $F(x) = \begin{cases} \dfrac{x}{1+x}, & \text{якщо } x \geq 0, \\ 0, & \text{якщо } x < 0. \end{cases}$

d) $F(x) = \begin{cases} 1 - \dfrac{1 - e^{-x}}{x}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \leq 0. \end{cases}$

e) $F(x) = \begin{cases} 0, & \text{якщо } x < 0, \\ \dfrac{[x]}{2}, & \text{якщо } 0 \leq x < 2, \\ 1, & \text{якщо } x \geq 2. \end{cases}$

Задача 10.3. Щільність розподілу випадкової величини ξ дорівнює

$$p_\xi(x) = a \cdot e^{-\lambda |x|},$$

де $a, \lambda > 0$. Обчислити: а) коефіцієнт a; б) функцію розподілу ξ, $\mathbf{M}\xi$ і $\mathbf{D}\xi$; с) побудувати графіки щільності і функції розподілу для $\lambda = 2$.

Задача 10.4. Задана функція

$$F(x) = \begin{cases} 0, & x \leq 1, \\ ax^2 + bx, & 1 < x \leq 4, \\ 1, & x > 4. \end{cases}$$

Підібрати коефіцієнти a та b таким чином, щоб $F(x)$ була функцією розподілу деякої випадкової величини ξ. Знайти $\mathrm{P}\{2 < \xi \leq 3\}$.

Задача 10.5. Нижче наведено функції, які залежать від певних параметрів. Визначити значення параметрів, для яких ці функції будуть щільностями розподілу.

а) $f(x) = \begin{cases} c, & \text{якщо } a \leq x \leq b, \\ 0, & \text{інакше}; \end{cases}$

б) $f(x) = \begin{cases} k|x - a|, & \text{якщо } c \leq x \leq d, \\ 0, & \text{інакше}; \end{cases}$

в) $f(x) = \begin{cases} ax^2 + bx + c, & ax^2 + bx + c \geq 0, \\ 0, & ax^2 + bx + c < 0; \end{cases}$

г) $f(x) = \begin{cases} cx^\alpha e^{-\lambda x}, & \text{якщо } x \geq 0, \\ 0, & \text{якщо } x < 0; \end{cases}$

д) $f(x) = \begin{cases} \dfrac{c}{ax + b}, & \text{якщо } d \leq x < \infty, \\ 0, & \text{якщо } x \leq d. \end{cases}$

Задача 10.6. Чи можна підібрати сталу C так, щоб функція C/x^3 визначала щільність розподілу ймовірностей а) на промені $[1; +\infty)$; б) на промені $[0, \infty)$; в) на відрізку $[-2, -1]$?

Задача 10.7. Нехай $F(x)$ – функція розподілу випадкової величини ξ. Знайти функцію розподілу випадкової величини $\eta = -\xi$.

Задача 10.8. Нехай $F(x)$ – функція розподілу випадкової величини ξ. Знайти функцію розподілу та щільність розподілу випадкової величини $\eta = -e^{\xi}$.

Задача 10.9. Професор ніколи не завершує лекцію в межах пари, натомість завершує її впродовж трьох хвилин після дзвінка на перерву. Нехай ξ – час, який пройшов від дзвінка до завершення лекції, і нехай щільність ξ задається таким чином:

$$f_\xi(x) = \begin{cases} kx^2, & 0 \le x \le 3; \\ 0, & \text{інакше.} \end{cases}$$

а) Знайти константу k;

б) знайти функцію розподілу випадкової величини ξ;

в) побудувати графіки щільності розподілу та функції розподілу випадкової величини ξ;

г) знайти середнє значення часу затримки лекції;

д) яка ймовірність того, що лектор затримає студентів не більш, ніж на хвилину; на час від 30 секунд до двох хвилин?

Задача 10.10. Двоє друзів домовились зустрітися в певному місці з п'ятої до шостої вечора. Знайти функцію розподілу та щільність розподілу часу очікування того, хто прийде першим, якщо моменти приходу обох друзів рівноможливі впродовж години.

Задача 10.11. Випадкова величина ξ має показниковий розподіл з параметром $\lambda = 1/3$. Обчислити ймовірності

$$\mathrm{P}\{\xi > 3\}; \qquad \mathrm{P}\{\xi > 6 | \xi > 3\}; \qquad \mathrm{P}\{\xi > t + 3 | \xi > t\}.$$

Задача 10.12. Знайти математичне сподівання, дисперсію та четвертий момент випадкової величини ξ, яка має рівномірний розподіл на відрізку $[a, b]$.

Задача 10.13. Нехай ξ має рівномірний розподіл на відрізку $[2, 4]$. Знайти таке x_0, що $\mathrm{P}\{\xi > x_0 + \mathbf{M}\xi\} = 1/4$.

Задача 10.14. Нехай ξ має рівномірний розподіл з математичним сподіванням 6 та дисперсією 4. Знайти ймовірність того, що $\xi < 5$.

Задача 10.15. Нехай $F(x)$ – функція розподілу випадкової величини ξ. Знайти функцію розподілу та щільність розподілу таких величин:

а) $a\xi + b,\ a \ne 0$; б) ξ^2;

в) $g(\xi)$, де $g(\cdot)$ – монотонна диференційована функція.

Задача 10.16. Випадкова величина ξ рівномірно розподілена на відрізку $[0, 2]$. Знайти функцію розподілу випадкової величини а) $\eta = |\xi - 1|$, б) $\zeta = |\xi|$.

Задача 10.17. Випадкова величина ξ рівномірно розподілена на відрізку $[-1, 1]$. Знайти функцію розподілу випадкової величини $\eta = |\xi|$.

Задача 10.18. Випадкова величина ξ рівномірно розподілена на відрізку $[0, 1]$. Знайти щільність розподілу випадкових величин а) $\eta = \xi^2$, б) $\eta = 1/\xi$, в) $\eta = e^{\xi}$ та побудувати їхні графіки.

Задача 10.19. Випадкова величина ξ має показниковий розподіл з параметром $\lambda > 0$. Знайти функцію розподілу та щільність розподілу випадкової величини $\eta = \xi^2$.

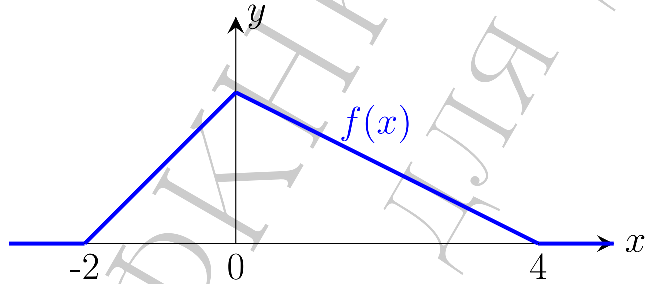
Задача 10.20. Випадкова величина ξ має щільність розподілу

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

Знайти щільність розподілу випадкової величини $\eta = \xi^2$.

Задача 10.21. Випадкова величина ξ задана на проміжку $[a, 4]$ щільністю $f(x) = Ax^2$. Треба знайти a, A та $F(2)$, якщо відомо $\mathbf{M}\xi = 0$.

Задача 10.22. Графік щільності розподілу випадкової величини ξ має вид



Знайти аналітичний вираз для щільності, функцію розподілу, $\mathbf{M}\xi$ та $\mathbf{D}\xi$.

Задача 10.23. Нехай $F(x) = a \cdot \sin x + b \cdot \cos x + 1/2$. Довести, що якщо x приймає значення з відрізку $[0; \pi/2]$, то можна знайти параметри a і b такі, що $F(x)$ буде функцією розподілу. Знайти $F(\pi/4)$.

Задача 10.24. Нехай функція $F(x)$ задається формулою

$$F(x) = \begin{cases} 0, & \text{якщо } x \le 1, \\ a + b \cdot arctg(x), & \text{якщо } x > 1. \end{cases}$$

Треба підібрати параметри a і b таким чином, щоб $F(x)$ була функцією розподілу.

Задача 10.25. Випадкова величина ξ приймає значення у відрізку $[a, 2]$ і має щільність розподілу

$$f(x) = \frac{3}{16} x^{n+1},$$

де n – деяке натуральне число. Знайти a і n, якщо відомо, що $\mathbf{M}\xi = 0$.

Задача 10.26. Випадкова величина ξ приймає значення у відрізку $[a, b]$, на якому задається щільністю
$$f(x) = \frac{3}{16}(x - c)^2.$$
Знайти a і b, якщо відомо, що $f(a) = f(b)$ і $\mathbf{M}\xi = 2$.

Задача 10.27. Випадкова величина ξ приймає значення у відрізку $[1, b]$, на якому задається щільністю
$$f(x) = A\ln(x).$$
Знайти A і b, якщо $\mathbf{M}\xi = \dfrac{b^2+1}{4}$.

Задача 10.28. Випадкова величина ξ задається щільністю
$$f_\xi(x) = \begin{cases} 0, & \text{якщо } x \le 0 \text{ або } x > 2, \\ 1/2, & \text{якщо } 0 < x \le 1, \\ x/3, & \text{якщо } 1 < x \le 2. \end{cases}$$
Знайти функцію розподілу і щільність випадкової величини $\eta = \xi^2$. Знайти $\mathbf{M}\eta$ та $\mathbf{D}\eta$.

Задача 10.29. Випадкова величина ξ має розподіл Коші зі щільністю
$$f(x) = \frac{1}{\pi\,(1 + x^2)}.$$
Знайти а) $\mathrm{P}\left\{|\xi| > 1\right\}$, б) $\mathbf{M}\,|\xi|\,/\,(1+\xi^2)$.

Задача 10.30. Колесо вагону має тріщину на зовнішньому краю. Нехай ξ – висота тріщини над землею при випадковій зупинці вагону. Знайти функцію розподілу ξ, щільність ξ та $\mathbf{M}\xi$.

Задача 10.31. Точка А рівномірно розподілена всередині кола радіуса R. Нехай ξ – відстань від точки А до центра кола. Знайти функцію розподілу та щільність випадкової величини ξ. Побудувати їхні графіки. Обчислити $\mathbf{M}\xi$ та $\mathbf{D}\xi$.

Задача 10.32. Точка А рівномірно розподілена на колі одиничного радіуса з центром в точці (0,0). Нехай ξ – проєкція точки А на вісь Ox. Знайти функцію розподілу та щільність випадкової величини ξ. Обчислити $\mathbf{M}\xi$ та $\mathrm{P}\{|\xi| \ge \frac{1}{2}\}$.

Задача 10.33. Хвилинна стрілка механічного годинника пересувається стрибком в кінці кожної хвилини. Знайти ймовірність того, що в даний момент годинник покаже час, який відрізняється від справжнього більше, ніж на 20 с.

Задача 10.34. Студент хотів під'їхати одну зупинку на автобусі, але автобус від'їхав безпосередньо перед ним. Нехай студент вирішив, що він піде пішки, якщо наступний автобус не прийде упродовж 5 хвилин. Інтервал між автобусами є випадковою величиною, що має рівномірний розподіл на відрізку $[4, 6]$. Яка ймовірність того, що студент чекатиме менше, ніж 4.5 хвилини? Яка ймовірність того, що студент чекатиме 5 хвилин? Записати розподіл випадкової величини ξ – часу очікування студента.

Задача 10.35. Точка A рівномірно розподілена на колі одиничного радіуса з центром в точці (0,0). Нехай ξ – довжина відрізку дотичної, проведеної в точці A до точки перетину з віссю Ox. Знайти функцію розподілу та щільність випадкової величини ξ. Чи існує $\mathbf{M}\xi$?

Задача 10.36. При яких параметрах функція

$$f_\xi(x) = \begin{cases} cx^\alpha \exp\left\{-\beta x\right\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \le 0. \end{cases}$$

є щільністю випадкової величини ξ? Знайти $\mathbf{M}\xi^n$, $\mathbf{M}\exp\left(-\lambda\xi\right)$.

Задача 10.37. Нехай ξ та η – незалежні випадкові величини рівномірно розподілені на [0,2]. Знайти функцію розподілу для величин
а) $\min\left(\xi,\eta\right)$; б) $\max\left(\xi,\eta\right)$.

Задача 10.38. Нехай ξ та η – незалежні випадкові величини, які мають показниковий розподіл з параметром $\lambda > 0$. Знайти функцію розподілу для величин
а) $\min\left(\xi,\eta\right)$; б) $\max\left(\xi,\eta\right)$; в) $\max\left(2\xi,\eta\right)$; г) $\min\left(\xi^3,\eta\right)$.

Задача 10.39. У невеликому приморському місті річні збитки від штормів, пожеж та розкрадань майна є незалежними випадковими величинами, що мають показниковий розподіл з середніми значеннями 1, 1.5 та 2.4 відповідно. Знайти ймовірність того, що максимальний з цих збитків буде більше 3.

Задача 10.40. Випадкова величина ξ задається щільністю

$$f_\xi(x) = \begin{cases} \dfrac{1}{\sigma}\sqrt{\dfrac{2}{\pi}}\exp\left\{-\dfrac{x^2}{2\sigma^2}\right\}, & \text{якщо } x \ge 0, \\ 0, & \text{якщо } x < 0. \end{cases}$$

Знайти щільність розподілу випадкової величини $\eta = 1/\xi$. Чи скінченне $\mathbf{M}\eta$?

Задача 10.41. Випадкова величина ξ задається щільністю

$$f_\xi(x) = \begin{cases} \frac{1}{5}, & \text{якщо } x \in [0,5], \\ 0, & \text{якщо } x \notin [0,5]. \end{cases}$$

Знайти щільність розподілу випадкової величини $\eta = \exp\left\{\xi^2 - 4\xi\right\}$.

Задача 10.42. Знайти функцію розподілу, математичне сподівання $\mathbf{M}l$ та дисперсію $\mathbf{D}l$ довжини хорди l, що з'єднує фіксовану точку кола радіусу R з іншою точкою кола, всі положення якої на колі рівноможливі.

Задача 10.43. Нехай ξ – випадкова величина, яка має показниковий розподіл з параметром λ. Знайти розподіл випадкової величини $\eta = [\xi]$ та $\mathbf{M}\eta$.

Задача 10.44. Нехай ξ – випадкова величина, яка має рівномірний розподіл у відрізку $[0,1]$. Треба знайти щільність розподілу випадкової величини $\eta = -\frac{1}{\lambda}\ln\xi$.

Задача 10.45. Нехай ξ – випадкова величина, яка має неперервну строго монотонну функцію розподілу $F(x)$. Довести, що випадкова величина $\eta = F(\xi)$ має рівномірний розподіл у проміжку $[0, 1]$.

Задача 10.46. Випадкова величина ξ розподілена за показниковим законом з параметром λ. При якому значенні λ ймовірність події $\{\xi \in [\alpha, \beta]\}$ буде максимальною?

Задача 10.47. Випадкова величина ξ розподілена за показниковим законом з параметром λ. Яка з подій більш ймовірна: випадкова величина більша свого математичного сподівання чи менша свого математичного сподівання? Чи залежить ця ймовірність від λ?

Задача 10.48. Випадкова величина ξ розподілена за показниковим законом з параметром λ. Обчислити $\mathbf{M}\xi^k$, $\mathbf{D}\xi$, $\mathrm{P}\{\xi > 1\}$.

Задача 10.49. Випадкова величина ξ має щільність розподілу $f_\xi(x) = \lambda e^{-\lambda x}$, $x \geq 0$, з параметром $\lambda > 0$. Для яких значень параметра μ існує математичне сподівання випадкової величини $\eta = e^{\mu\xi}$?

Задача 10.50. Випадкова величина ξ розподілена за нормальним законом з параметрами m і σ^2. Знайти $\mathbf{M}|\xi - m|$.

Задача 10.51. Випадкова величина ξ розподілена за нормальним законом з параметрами $m = 0$ і σ^2. Задано інтервал (α, β), який не включає початок координат. Визначити, при якому σ ймовірність події $\{\xi \in (\alpha, \beta)\}$ буде максимальною.

Задача 10.52. Випадкова величина ξ задається щільністю

$$f_\xi(x) = \begin{cases} 1 - |x - 1|, & \text{якщо } x \in [0, 2], \\ 0, & \text{якщо } x \notin [0, 2]. \end{cases}$$

Треба знайти $\mathbf{M}\xi$ і $\mathbf{D}\xi$.

Задача 10.53. Довести, що

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-a)^2}{2\sigma^2}}, \qquad x \in \mathbb{R},$$

є щільністю розподілу.

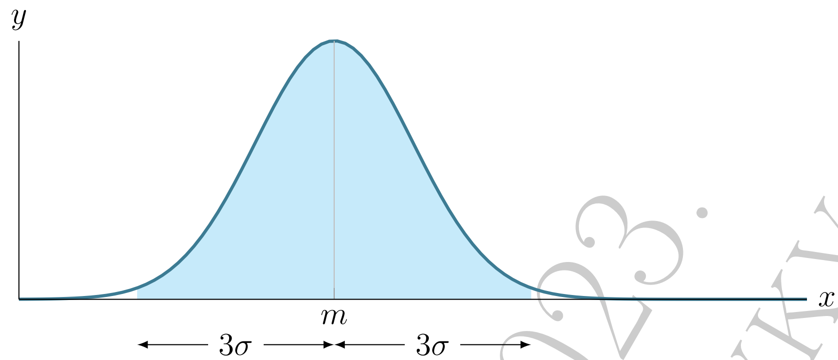
Задача 10.54. Нехай ξ має розподіл $N(a, \sigma^2)$, а $\Phi(x)$ – функція розподілу стандартної гауссівської випадкової величини. Довести, що для довільного $x \in \mathbb{R}$

$$F_\xi(x) = \Phi\left(\frac{x - a}{\sigma}\right).$$

Задача 10.55. Знайти математичне сподівання та дисперсію випадкової величини, що має розподіл $N(a, \sigma^2)$.

Задача 10.56. Правило «трьох сигм». Показати, що для випадкової величини ξ, що має розподіл $N(a, \sigma^2)$, має місце

$$P\{|\xi - a| > 3\sigma\} < 0.003.$$



Задача 10.57. Місячний прибуток підприємства є нормально розподіленою випадковою величиною з середнім значенням 5 млн. дол. та стандартним відхиленням 0.5 млн. дол. Знайти ймовірність того, що в даному місяці прибуток підприємства перевищить 4 млн. дол. Записати формулу щільності розподілу цієї випадкової величини, зобразити її графік та вказати на ньому обчислену ймовірність.

Задача 10.58. Місячний прибуток підприємства є нормально розподіленою випадковою величиною зі стандартним відхиленням 0.5 млн. дол. Відомо, що в 70 % випадків прибуток підприємства перевищує 4 млн. дол. Знайти середній прибуток підприємства.

Задача 10.59. Розглядається питання закупівлі машин для нарізання корків для пляшок з вином. Перша нарізає корки з діаметром, який є нормально розподіленою випадковою величиною з математичним сподіванням 3 см та стандартним відхиленням 0.1 см. Друга продукує корки з діаметром, який є нормально розподіленою випадковою величиною з математичним сподіванням 3.04 см та стандартним відхиленням 0.02 см. Прийнятними є корки з діаметром від 2.9 до 3.1 см. Яка машина виробляє більше бракованих корків?

Задача 10.60. Прилад для автоматичного розкривання військового вантажного парашута спроєктований таким чином, щоб відкрити парашут на висоті 200 м над поверхнею землі. Насправді висота відкриття є випадковою величиною, яка має нормальний розподіл з математичним сподіванням 200 м і стандартним відхиленням 30 м. Вантаж буде ушкоджено, якщо парашут розкриється на висоті меншій, ніж 100 м. Яка ймовірність того, що при незалежному зкиданні п'яти парашутів з вантажем хоча б один з них буде ушкоджено?

Задача 10.61. Завод виготовляє кульки для підшипників з номінальним діаметром 10 мм. Фактичний діаметр є випадковою величиною, яка має нормальний розподіл з математичним сподіванням 10 мм та стандартним відхиленням 0.4 мм. При контролі якості відбраковуються всі кульки, діаметр яких більший за 10.7 мм, або менший за 9.3 мм. Знайти відсоток бракованих кульок.

Задача 10.62. Нехай випадкова величина ξ набуває значення 1 та -1 з ймовірностями 1/2.

Обчислити характеристичну функцію цієї випадкової величини.

Задача 10.63. Довести, що функція $\varphi(t) = \cos^2 t$ є характеристичною функцією та знайти відповідний їй розподіл ймовірностей.

Задача 10.64. Нехай випадкова величина ξ набуває значення -1, 0 та 1 з ймовірностями $1/3$ кожне. Обчислити характеристичну функцію цієї випадкової величини.

Задача 10.65. а) Нехай $\xi_1, \xi_2, ..., \xi_n$ – незалежні випадкові величини, кожна з яких набуває значення 1 та -1 з ймовірностями $1/2$. Обчислити характеристичну функцію випадкової величини $\nu_n = \xi_1 + ... + \xi_n$.

б) Довести, що для будь-якого натурального n функція $\varphi(t) = \cos^n t$ є характеристичною функцією

Задача 10.66. Обчислити характеристичну функцію біноміальної випадкової величини з параметрами n та p.

Задача 10.67. Обчислити характеристичну функцію випадкової величини, яка має

а) розподіл Пуассона з параметром λ;

б) геометричний розподіл з параметром p.

Задача 10.68. Обчислити характеристичну функцію випадкової величини, яка має

а) рівномірний розподіл на відрізку $[-a; a]$;

б) рівномірний розподіл на відрізку $[a; b]$;

в) показниковий розподіл з параметром λ;

г) гамма-розподіл зі щільністю

$$f_\xi(x) = \begin{cases} \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

$\Gamma(\alpha) = \int_0^{+\infty} y^{\alpha-1} e^{-x} dx = (\alpha - 1)\Gamma(\alpha - 1)$ – гамма-функція Ейлера.

Задача 10.69.* Довести, що характеристична функція випадкової величини η, що розподілена за законом $N(a, \sigma^2)$, має вид

$$\varphi_\eta(t) = e^{ita - \sigma^2 t^2/2}.$$

Задача 10.70. Нехай ξ має стандартний гауссівський розподіл. Знайти моменти $\mathbf{M}\xi^n$, $n = 1, 2, \ldots$

Задача 10.71.* На ймовірнісному просторі $(\Omega, \Re, \mathrm{P})$ (де $\Omega = [0.1]$, $\Re$ – σ-алгебра вимірних за Лебегом підмножин [0, 1], P – міра Лебега) визначена випадкова величина $\xi = \sqrt{\omega}$. Знайти характеристичну функцію ξ.

Задача 10.72. Використовуючи характеристичні функції, показати, що сума $\xi_1 + \xi_2$ незалежних випадкових величин ξ_1 та ξ_2, які мають розподіл Пуассона з параметрами λ_1 та λ_2 відповідно, також має розподіл Пуассона з параметром $\lambda_1 + \lambda_2$.

Задача 10.73. Використовуючи характеристичні функції, показати, що якщо випадкові величини $\xi_1 \sim \Gamma(\lambda, \alpha)$ та $\xi_2 \sim \Gamma(\lambda, \beta)$ незалежні, то їхня сума $\xi_1 + \xi_2 \sim \Gamma(\lambda, \alpha + \beta)$ (див. задачу 10.68).

Задача 10.74. Використовуючи характеристичні функції, показати, що сума $\xi_1 + \xi_2$ незалежних випадкових величин ξ_1 та ξ_2, які мають нормальний розподіл $N(\mu_1, \sigma_1^2)$ та $N(\mu_2, \sigma_2^2)$ відповідно, також має нормальний розподіл $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Задача 10.75. Нехай $\varphi(t)$ є характеристичною функцією випадкової величини ξ. Знайти характеристичну функцію випадкової величини $\eta = a + b\xi$.

Задача 10.76. Використовуючи характеристичну функцію, знайти всі моменти показникового (експоненційного) розподілу з параметром $\alpha > 0$, зокрема, порахувати математичне сподівання та дисперсію.

Задача 10.77. Час безвідмовної роботи у годинах деякого приладу є випадковою величиною ξ абсолютно неперервного типу, яка задана щільністю розподілу $f_\xi(x) = 0.08e^{-0.08t}$. Знайти наступні характеристики:

а) середній час безвідмовної роботи приладу;

б) ймовірність безвідмовної роботи приладу упродовж 4 год.;

в) ймовірність відмови приладу в інтервалі часі від 4 до 20 год.;

г) надійність (ймовірність безвідмовної роботи) системи упродовж 4 год., якщо вона складається з п'яти послідовно підключених однакових приладів. Вважати відмови пристроїв незалежними.

11 ВИПАДКОВІ ВЕКТОРИ ТА ФУНКЦІЇ ВІД ВИПАДКОВИХ ВЕЛИЧИН

Нехай випадкові величини ξ_i, $i = 1.2, ..., n$, задані на одному і тому ж ймовірнісному просторі. **Випадковим вектором** будемо називати $\xi' = (\xi_1, \xi_2, ..., \xi_n)$. Очевидно, випадковий вектор задає вимірне відображення вимірного простору $(\Omega, \Re)$ у вимірний простір $(\mathbb{R}^n, \mathscr{B}^n)$, де $\mathbb{R}^n$ – n-вимірний Евклідів простір, $\mathscr{B}^n$ – σ-алгебра борелівських множин з R^n.

Функція

$$F_{\xi'}(x_1, x_2, ..., x_n) \stackrel{def}{=} \mathrm{P}(\xi_1 \leq x_1, \xi_2 \leq x_2, ..., \xi_n \leq x_n),$$
$$x_i \in \mathbb{R}^1, \quad i = 1, 2, ..., n$$

називається **функцією розподілу випадкового вектора** ξ або **сумісною функцією розподілу випадкових величин** $\xi_1, \xi_2, ..., \xi_n$.

Сумісна функція розподілу має наступні властивості:

1. $\lim\limits_{x_i \to -\infty} F_\xi(x_1, x_2, ..., x_n) = 0, \quad i = \overline{1, n},$

   $\lim\limits_{x_i \to \infty} F_\xi(x_1, x_2, ..., x_n) = F_{\xi_1,..,\xi_{i-1},\xi_{i+1},...,\xi_n}(x_1, ...x_{i-1}, x_{i+1}..., x_n).$

2. Якщо оператор T означає перестановку координат вектора $x' = (x_1, ..., x_n)$, то

$$F_{T\xi}(Tx) = F_\xi(x).$$

3. Функція $F_\xi(x_1, x_2, ..., x_n)$ є неперервною справа і неспадною по кожному аргументу.

Будемо говорити, що функція розподілу $F_{\xi'}(x_1, x_2, ..., x_n)$ випадкового вектора $\xi' = (\xi_1, \xi_2, ..., \xi_n)$ є **абсолютно неперервного типу**, якщо вона може бути подана у вигляді

$$F_\xi(x_1, x_2, ..., x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} ... \int_{-\infty}^{x_n} f(y_1, y_2, ...y_n) dy_1 dy_2 ... dy_n \tag{11.1}$$

Функція $f(x_1, x_2, ..., x_n)$ називається **щільністю функції розподілу** $F_\xi(x_1, x_2, ..., x_n)$.

Вимірна функція $f(x_1, x_2, ..., x_n)$, $x_i \in \mathbb{R}^1$, $i = 1, 2, ..., n$, є щільністю деякої функції розподілу тоді і лише тоді, коли виконуються наступні умови:

1. $f(x_1, x_2, ..., x_n) \geq 0$, для всіх $x_i \in \mathbb{R}^1$, $i = 1, 2, ..., n$.
2. $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} ... \int\limits_{-\infty}^{\infty} f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n = 1.$

Для будь-якої борелівської множини $A \in \mathscr{B}^n$ справедлива формула

$$\mathrm{P}(\xi \in A) = \int\int_A ... \int f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n.$$

Якщо щільність $f(x_1, x_2, ..., x_n)$, є неперервною в точці $x_0 = (x_1^0, x_2^0, ..., x_n^0)$, то

$$f(x_0) = \left. \frac{\partial^n F_\xi(x)}{\partial x_1 \partial x_2 ... \partial x_n} \right|_{x=x_0}. \tag{11.2}$$

В загальному випадку формула (11.2) справедлива для майже всіх відносно міри Лебега в просторі $\mathbb{R}^n$ точок $(x_1, x_2, ..., x_n)$.

Випадкові величини $\xi_1, \xi_2, ..., \xi_n$ називаються **незалежними**, якщо

$$F_\xi(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} F_{\xi_i}(x_i)$$

для довільних $x_i \in \mathbb{R}^1$, $i = 1, 2, ..., n$.

Наведемо властивості чисельних характеристик випадкових величин, що пов'язані з поняттям незалежності. Якщо випадкові величини ξ та η є незалежними, то:

1. $\mathbf{M}\xi\eta = \mathbf{M}\xi\mathbf{M}\eta$ при умові, що $\mathbf{M}\xi$ та $\mathbf{M}\eta$ скінченні;
2. $\mathbf{D}(a\xi + b\eta) = a^2\mathbf{D}\xi + b^2\mathbf{D}\eta$ при умові, що $\mathbf{D}\xi$ та $\mathbf{D}\eta$ скінченні;
3. $\varphi_{\xi+\eta}(t) = \varphi_\xi(t)\varphi_\eta(t)$.

Розглянемо дві випадкові величини з сумісною функцією розподілу

$$F(x, y) = \mathrm{P}(\xi \leq x, \eta \leq y).$$

Функції

$$F_\xi(x) = \mathrm{P}(\xi \leq x) = F(x, \infty), \qquad F_\eta(x) = \mathrm{P}(\eta \leq y) = F(\infty, y)$$

називаються **маргінальними функціями розподілу** для функції $F(x, y)$.

Теорема 11.1 Нехай функція розподілу $F(x, y)$ має щільність $f(x, y)$. Тоді
а) функції розподілу випадкових величини ξ та η теж мають щільності:

$$f_\xi(x) = \int_{-\infty}^{\infty} f(x, y)dy, \qquad f_\eta(y) = \int_{-\infty}^{\infty} f(x, y)dx,$$

б) випадкові величини ξ та η незалежні тоді і лише тоді, коли

$$f(x, y) = f_\xi(x)f_\eta(y).$$

Функція

$$F(x|y) = \mathrm{P}(\xi \leq x|\eta = y) \stackrel{def}{=} \begin{cases} \int\limits_{-\infty}^{x} \dfrac{f(z, y)}{f_\eta(y)}dz, & f_\eta(y) > 0; \\ 0, & f_\eta(y) = 0, \end{cases}$$

називається **умовною функцією розподілу** для ξ при умові $\eta = y$, а функція

$$f(x|y) \stackrel{def}{=} \begin{cases} \dfrac{f(x,y)}{f_\eta(y)}, & f_\eta(y) > 0; \\ 0, & f_\eta(y) = 0, \end{cases}$$

називається **умовною щільністю** для ξ при умові $\eta = y$.

Теорема 11.2 Нехай випадкові величини ξ_i, $i = 1, 2, ..., n$, визначені на одному і тому ж ймовірнісному просторі. Тоді $f(\xi_1, \xi_2, ..., \xi_n)$ є випадковою величиною для кожної борелівської функції $f(x_1, x_2, ..., x_n)$. Якщо $A_x = \{(x_1, x_2, ..., x_n) : f(x_1, x_2, ..., x_n) \le x\}$, то $\mathrm{P}\{f(\xi_1, \xi_2, ..., \xi_n) \le x\} = \mathrm{P}(\xi \in A_x)$.

Теорема 11.3 Нехай випадкові величини ξ та η будуть незалежними і $F_\xi(x)$ та $F_\eta(x)$ позначають їхні функції розподілу. Тоді

$$F_{\xi+\eta}(x) = \mathrm{P}(\xi + \eta \le x) = \int\limits_{-\infty}^{\infty} F_\xi(x-y)dF_\eta(y) = \int\limits_{-\infty}^{\infty} F_\eta(x-y)dF_\xi(y). \tag{11.3}$$

$$F_{\xi\eta}(x) = \mathrm{P}(\xi\eta \le x) = \int\limits_{0}^{\infty} F_\xi(xy^{-1})dF_\eta(y) + \int\limits_{-\infty}^{0} (1 - F_\xi(xy^{-1} - 0))dF_\eta(y), \tag{11.4}$$

або

$$F_{\xi\eta}(x) = \int\limits_{0}^{\infty} F_\eta(xy^{-1})dF_\xi(y) + \int\limits_{-\infty}^{0} (1 - F_\eta(xy^{-1} - 0))dF_\xi(y).$$

Якщо додатково $\mathrm{P}(\eta = 0) = 0$, то

$$F_{\xi|\eta}(x) = \mathrm{P}(\xi|\eta \le x) = \int\limits_{0}^{\infty} F_\xi(xy)dF_\eta(y) + \int\limits_{-\infty}^{0} (1 - F_\xi(xy - 0))dF_\eta(y) \tag{11.5}$$

Функція розподілу яка визначається інтегралом (першим або другим) в формулі (11.3) називається **згорткою функцій розподілу** $F_\xi(x)$ та $F_\eta(x)$.

Якщо випадкові величини ξ та η незалежні, і одна з них (наприклад ξ) має щільність $f_\xi(x)$, то випадкові величини $\xi + \eta$, $\xi\eta$ та $\xi|\eta$ теж мають щільності, які визначаються формулами

$$f_{\xi+\eta}(x) = \int\limits_{-\infty}^{\infty} f_\xi(x-y)dF_\eta(y), \tag{11.6}$$

$$f_{\xi\eta}(x) = \int\limits_{-\infty}^{\infty} |y|^{-1} f_\xi(xy^{-1})dF_\eta(y) = \int\limits_{-\infty}^{\infty} |y^{-1}| f_\eta(xy^{-1})dF_\xi(y). \tag{11.7}$$

Якщо $\mathrm{P}(\eta = 0) = 0$, то

$$f_{\xi|\eta}(x) = \int\limits_{-\infty}^{\infty} |y| f_\xi(xy)dF_\eta(y). \tag{11.8}$$

Формули для коваріації та коефіцієнта кореляції залишаються тими ж, що і в дискретному випадку:

$$\mathrm{cov}\,(\xi,\eta) = \mathbf{M}\,(\xi - \mathbf{M}\xi)\,(\eta - \mathbf{M}\eta) = \mathbf{M}\xi\eta - \mathbf{M}\xi\mathbf{M}\eta,$$

$$\rho = \frac{\mathrm{cov}\,(\xi,\eta)}{\sqrt{\mathbf{D}\xi\mathbf{D}\eta}},$$

тільки тепер $\mathbf{M}\xi\eta$ розраховується за формулою

$$\mathbf{M}\xi\eta = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xy dF_{\xi,\eta}(x,y),$$

або, якщо існує щільність $f_{\xi,\eta}(x,y)$, за формулою

$$\mathbf{M}\xi\eta = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xy f_{\xi,\eta}(x,y)dxdy.$$

Властивості кореляції та коефіцієнта кореляції, наведені для випадку дискретних випадкових величин, зберігаються і в загальному випадку.

Приклад 11.1

Двовимірна випадкова величина задана таблицею

$\xi \setminus \eta$	0	1
0	$\frac{15}{28}$	$\frac{6}{28}$
1	$\frac{6}{28}$	$\frac{1}{28}$

Побудуємо для неї сумісну функцію розподілу.

Якщо $x < 0$ або $y < 0$, то $F_{\xi,\eta}(x, y) = 0$. Якщо $0 \le x < 1$ або $0 \le y < 1$, то

$$F_{\xi,\eta}(x, y) = \mathrm{P}\{\xi \le x, \eta \le y\} = \mathrm{P}\{\xi = 0, \eta = 0\} = \frac{15}{28}.$$

Якщо $0 \le x < 1$, $y \ge 1$, то

$$F_{\xi,\eta}(x, y) = \mathrm{P}\{\xi \le x, \eta \le y\} = \mathrm{P}\{\xi = 0, \eta = 0 \text{ або } \eta = 1\} = \frac{15}{28} + \frac{6}{28} = \frac{21}{28}.$$

Якщо $x \ge 1$, $0 \le y < 1$, то

$$F_{\xi,\eta}(x, y) = \mathrm{P}\{\xi \le x, \eta \le y\} = \mathrm{P}\{\xi = 0 \text{ або } \xi = 1, \eta = 0\} = \frac{15}{28} + \frac{6}{28} = \frac{21}{28}.$$

Якщо $x \ge 1$, $y \ge 1$, то

$$F_{\xi,\eta}(x, y) = \mathrm{P}\{\xi \le x, \eta \le y\} = \mathrm{P}\{\xi = 0 \text{ або } \xi = 1, \eta = 0 \text{ або } \eta = 1\} = \\ = \frac{15}{28} + \frac{6}{28} + \frac{6}{28} + \frac{1}{28} = 1.$$

Приклад 11.2

Нехай ξ_1 та ξ_2 – незалежні випадкові величини, що мають рівномірний на відрізку $[0, 1]$. Знайти щільність розподілу суми $\xi_1 + \xi_2$.

Розв'язок: $f_{\xi_1}(x) = f_{\xi_2}(x) = \begin{cases} 1, & x \in [0.1] \\ 0, & x \notin [0.1] \end{cases}$. Очевидно, що $(\xi_1 + \xi_2) \in [0.2]$, і оскільки ξ_1 та ξ_2 незалежні, то $\forall x \in [0.2]$

$$f_{\xi_1+\xi_2}(x) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(x - y) f_{\xi_2}(y) dy = \int\limits_0^1 f_{\xi_1}(x - y) 1 dy =$$

$$= \int\limits_0^1 f_{\xi_1}(x - y) dy = \int\limits_{\max(0,x-1)}^{\min(x,1)} 1 dy = \begin{cases} x, & x < 1 \\ 2 - x, & x \ge 1 \end{cases}$$

Остаточно маємо наступну **відповідь:**

$$f_{\xi_1+\xi_2}(x) = \begin{cases} 0, & \text{якщо } x < 0, \\ x, & \text{якщо } 0 \le x < 1, \\ 2 - x, & \text{якщо } 1 \le x < 2, \\ 0, & \text{якщо } x \ge 2. \end{cases}$$

Приклад 11.3

Нехай незалежні випадкові величини ξ_1 та ξ_2 мають стандартний нормальний розподіл. Доведемо, що їхня сума має нормальний розподіл з параметрами $\mu = 0$, $\sigma^2 = 2$.

Щільність нормального розподілу задається формулою $f_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$, $x \in \mathbb{R}^1$.

Згідно формули згортки щільність суми дорівнює

$$f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{+\infty} f_{\xi_1}(u) f_{\xi_2}(x-u)du = \int_{-\infty}^{+\infty} \frac{1}{2\pi} e^{-u^2/2} e^{-(x-u)^2/2} du =$$
$$= \int_{-\infty}^{+\infty} \frac{1}{2\pi} e^{-u^2+\frac{x^2}{2}-xu} du = e^{-x^2/4} \int_{-\infty}^{+\infty} \frac{1}{2\pi} e^{-\left(u-\frac{x}{2}\right)^2} du =$$
$$= \frac{1}{2\sqrt{\pi}} e^{-x^2/4} \int_{-\infty}^{+\infty} \frac{1}{\pi} e^{-v^2} dv = \frac{e^{-x^2/4}}{2\sqrt{\pi}}.$$

Останній інтеграл дорівнює 1, оскільки під інтегралом стоїть щільність нормального закону з параметрами $\mu = 0$, $\sigma^2 = 1/2$.

Отже, ми отримали, що шукана щільність суми є щільністю нормального розподілу з параметрами 0 та 2.

Приклад 11.4

Сумісний розподіл випадкових величин ξ_1 та ξ_2 задається наступною щільністю:

$$f(x,y) = \begin{cases} 0, & \text{якщо } (x,y) \notin [0.1] \times [0.1], \\ \frac{12}{13}\left(x^2 + 3xy\right), & \text{якщо } (x,y) \in [0.1] \times [0.1]. \end{cases}$$

Знайти щільність розподілу суми $\xi_1 + \xi_2$.

Розв'язок: Очевидно, що $(\xi_1 + \xi_2) \in [0.2]$, тому $\forall t \in [0.2]$

$$F_{\xi_1+\xi_2}(t) = \mathrm{P}\left\{\xi_1 + \xi_2 \le t\right\} = \begin{cases} \int\limits_0^t \int\limits_0^{t-x} f(x,y)dydx, \quad t \in [0,1) \\ \int\limits_0^{t-1} \int\limits_0^1 f(x,y)dydx + \int\limits_{t-1}^1 \int\limits_0^{t-x} f(x,y)dydx, \ t \in [1,2] \end{cases} =$$

$$= \begin{cases} \frac{5}{26}t^4, \quad t \in [0,1) \\ \frac{1}{13}\left(4t^3 + 2t^2 - 16t + 5 - 5t^4\right), \ t \in [1,2] \end{cases}$$

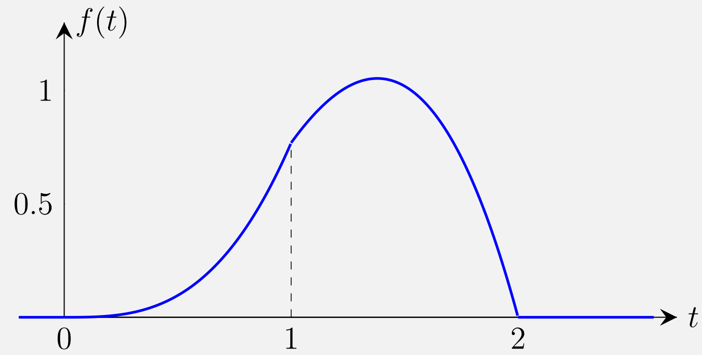
Таким чином,

$$F_{\xi_1+\xi_2}(t) = \mathrm{P}\left\{\xi_1 + \xi_2 \le t\right\} = \begin{cases} 0, \quad t < 0 \\ \frac{5}{26}t^4, \quad t \in [0,1) \\ \frac{1}{13}\left(4t^3 + 2t^2 - 16t + 5 - 5t^4\right), \ t \in [1,2] \\ 1, \quad t > 2 \end{cases}$$

і

$$f_{\xi_1+\xi_2}(t) = \frac{d}{dt}F_{\xi_1+\xi_2}(t) = \begin{cases} 0, \quad t < 0 \\ \frac{10}{13}t^3, \quad t \in [0,1) \\ \frac{2}{13}\left(6t^2 + 12t - 8 - 5t^3\right), \ t \in [1,2] \\ 0, \quad t > 2 \end{cases}$$

Графік цієї щільності має наступний вид:



Приклад 11.5

Знайдемо щільність а) добутку та б) частки незалежних випадкових величин ξ_1, ξ_2, кожна з яких має рівномірний розподіл на $(0,1)$.

а) Добуток ξ_1, ξ_2, очевидно, може набувати значень лише на проміжку $(0,1)$. Відповідно, за межами цього проміжку щільність добутку дорівнює 0. Знайдемо її значення для $x \in (0,1)$.

Аргументи обох щільностей в (11.7) мають належати $(0,1)$, щоб кожна з підінтегральних щільностей не дорівнювала нулю. Отже, для $x \in (0,1)$

$$\begin{cases} 0 < u < 1; \\ 0 < x/u < 1; \end{cases} \quad \begin{cases} 0 < u < 1; \\ x < u < +\infty; \end{cases} \quad \text{отже } \; x < u < 1.$$

Маємо:

$$f_{\xi_1 \cdot \xi_2}(x) = \int_{-\infty}^{+\infty} \frac{1}{|u|} f_{\xi_1,\xi_2}\left(u, \frac{x}{u}\right) du = \int_x^1 \frac{1}{u} du = \ln u\big|_x^1 = -\ln x.$$

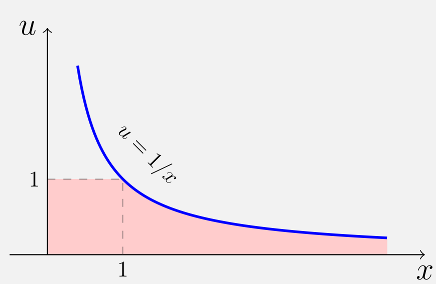
Остаточно

$$f_{\xi_1 \cdot \xi_2}(x) = \begin{cases} -\ln x, & x \in (0,1); \\ 0, & x \notin (0,1). \end{cases}$$

б) Частка ξ_1/ξ_2 може набувати значень на $(0,+\infty)$. Тому для $x \le 0$ її щільність дорівнює 0. Для $x > 0$, міркуючи аналогічно попередньому випадку, маємо, що

$$\begin{cases} 0 < u < 1; \\ 0 < xu < 1; \end{cases} \quad \begin{cases} 0 < u < 1; \\ 0 < u < 1/x \end{cases}$$

Відповідно, якщо $x < 1$, то межі інтегрування по u в (11.8) від 0 до 1, якщо $x \ge 1$, то $u \in (0, 1/x)$.



Маємо:

$$f_{\xi_1/\xi_2}(x) = \int_{-\infty}^{+\infty} |u| f_{\xi_1,\xi_2}(ux, u)\, du = \begin{cases} 0, & x \le 0; \\ \int_0^1 u\, du = \frac{1}{2}, & x \in (0, 1]; \\ \int_0^{1/x} u\, du = \frac{1}{2x^2}, & x \in (1, +\infty). \end{cases}$$

Зауважимо, що математичне сподівання частки

$$\mathbf{M}(\xi_1/\xi_2) = \int_0^1 x\frac{1}{2}\, dx + \int_1^{+\infty} x\frac{1}{2x^2}\, dx = \left.\frac{x^2}{4}\right|_0^1 + \left.\frac{1}{2}\ln x\right|_1^{+\infty} = +\infty$$

суттєво відрізняється від частки математичних сподівань

$$\mathbf{M}\xi_1/\mathbf{M}\xi_2 = 1.$$

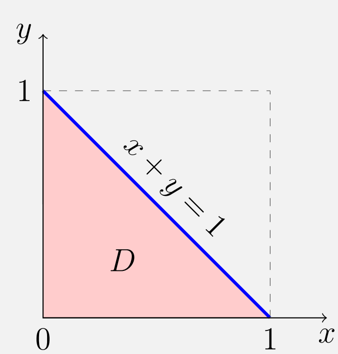
Приклад 11.6

Кондитерська фабрика пакує новорічні подарункові набори, в кожному з яких міститься кілограм суміші цукерок трьох видів: карамельні, желейні та шоколадні, причому кількість кожного виду є випадковою. Оскільки сума трьох типів цукерок є фіксованою, то для побудови сумісної ймовірнісної моделі достатньо розглянути лише дві випадкові величини, наприклад, вагу в наборі карамельок та желейок. Маючи інформацію щодо цих двох типів цукерок, ми зможемо визначити розподіл ваги шоколадок.

Позначимо через ξ та η вагу відповідно карамельок та желейок в одному подарунковому наборі. Їхня сумісна щільність буде додатньою в області

$$D = \{(x, y) : 0 \le x \le 1, 0 \le y \le 1, x + y \le 1\}.$$

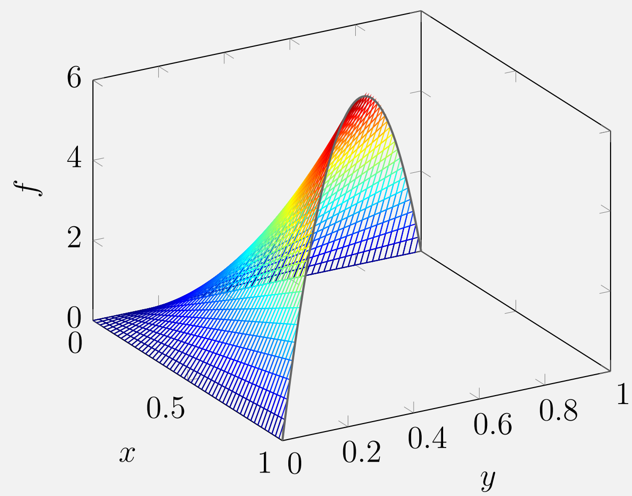
Ця область зображена на рисунку.



Нехай тепер сумісна щільність випадкових величин ξ та η має вигляд

$$f(x,y) = \begin{cases} 24xy, & x \in [0,1], y \in [0,1], x+y \leq 1; \\ 0, & \text{інакше.} \end{cases}$$

Для будь-яких фіксованих значень x функція зростає по y, при фіксованих y функція зростає по x.



Така форма щільності демонструє, що компанія економить: ймовірність того, що в наборі буде багато шоколадних цукерок, невелика, тобто пакунок більш ймовірно буде заповнений переважно карамельками та желейками: ймовірність того, що кількість цих цукерок набуває малих значень (а отже, шоколадні цукерки складають більшу частину набору), близька до нуля. А ймовірність коло межі (коли набір складається майже навпіл з желейок та карамельок) – висока.

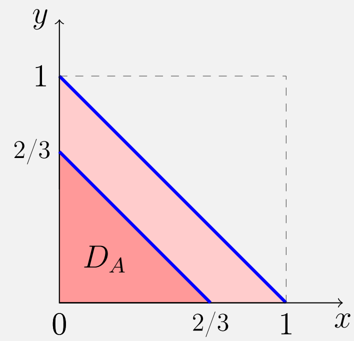
Очевидно, $f_{\xi,\eta}(x,y) \geq 0 \ \ \forall (x,y) \in \mathbb{R}^2$. Перевіримо умову нормування:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{\xi,\eta}(x,y)\, dx\, dy = \iint_D f_{\xi,\eta}(x,y)\, dx\, dy = \int_0^1 \left[\int_0^{1-x} 24xy\, dy \right] dx =$$

$$= \int_0^1 \left[24x \left. \frac{y^2}{2} \right|_0^{1-x} \right] dx = \int_0^1 12x(1-x)^2 dx = 1.$$

Знайдемо ймовірність того, що карамельки та желейки складають менше, ніж дві третини ваги пакунка (це означатиме, що в пакунку міститься принаймні третина шоколадних цукерок). Цій події відповідає область

$$D_A = \{(x,y) : 0 \leq x \leq 1, 0 \leq y \leq 1, x+y \leq 2/3\}.$$



Відповідно

$$P\{(\xi,\eta) \in D_A\} = \iint_{D_A} f_{\xi,\eta}(x,y)\,dx\,dy = \int_0^{2/3}\int_0^{2/3-x} 24xy\,dy\,dx = \frac{16}{81}.$$

Отже, ймовірність того, що в наборі буде хоча б третина шоколадних цукерок, невелика.

Інтегруванням сумісної щільності по "зайвій" змінній можна знайти маргінальні щільності розподілу ваги карамельок:

$$f_\xi(x) = \begin{cases} \int_0^{1-x} 24xy\,dy = 12x(1-x)^2, & x \in [0,1]; \\ 0, & x \notin [0,1]. \end{cases}$$

Маргінальна щільність випадкової величини η буде мати такий самий вигляд (із заміною x на y) в зв'язку з симетричністю сумісної щільності та області D.

Можемо тепер знайти, наприклад, ймовірність того, що карамельок в наборі буде не менше половини:

$$P\{\xi \geq 1/2\} = \int_{1/2}^{1} 12x(1-x)^2 dx = \frac{3}{16}.$$

Приклад 11.7

Нехай сумісна щільність випадкових величин ξ, η задана формулою:

$$f_{(\xi,\eta)}(x,y) = \begin{cases} C(xy^2 + yx^2), & x \in [0,1], y \in [1,2]; \\ 0, & \text{інакше}. \end{cases}$$

Знайдемо

а) константу C,

б) ймовірність того, що точка (ξ, η) потрапить у прямокутник $x \in [0, 1/2]$, $y \in [1, 3/2]$,
в) сумісну функцію розподілу випадкових величин ξ, η.

Розв'язок. а) Для визначення константи C скористаємось умовою нормування щільності:

$$\int_0^1 \int_1^2 C(xy^2 + yx^2)\, dy\, dx = C \int_0^1 \left[\frac{xy^3}{3}\bigg|_{y=1}^2 + \frac{x^2y^2}{2}\bigg|_{y=1}^2 \right] dx =$$
$$= C \left[\frac{7}{3} \cdot \frac{x^2}{2}\bigg|_{x=0}^1 + \frac{3}{2}\frac{x^3}{3} \cdot \bigg|_{x=0}^1 \right] = C\frac{5}{3} = 1$$

Звідси $C = 3/5$.

б)

$$\mathrm{P}\{\xi \in [0, 1/2], \eta \in [1, 3/2]\} = \int_0^{1/2} \left[\int_1^{3/2} \frac{3}{5}(xy^2 + yx^2) dy \right] dx =$$
$$\frac{3}{5} \int_0^{1/2} \left[\frac{xy^3}{3}\bigg|_{y=1}^{3/2} + \frac{x^2y^2}{2}\bigg|_{y=1}^{3/2} \right] dx = \frac{19}{40} \cdot \frac{x^2}{2}\bigg|_{x=0}^{1/2} + \frac{15}{40} \cdot \frac{x^3}{3}\bigg|_{x=0}^{1/2} = \frac{3}{40}.$$

в) Визначимо значення функції розподілу для значень (x, y) всередині прямокутника $x \in [0, 1]$, $y \in [1, 2]$:

$$F_{(\xi,\eta)}(x, y) = \int_0^x \left[\int_1^y \frac{3}{5}(uv^2 + vu^2) dv \right] du = \frac{3}{5} \int_0^x \left[\frac{uy^3}{3} - \frac{u}{3} + \frac{u^2y^2}{2} - \frac{u^2}{2} \right] du =$$
$$... = \frac{1}{10}(x^2 + x^3)(y^2 - 1).$$

Для x або y, які набувають значень за межами прямокутника $[0, 1] \times [1, 2]$, вирази для функції розподілу будуть різними. Наприклад, нехай $x \in [0, 1]$, а $y > 2$. Тоді, оскільки на цьому проміжку значення $f_{\xi,\eta)}(x, y) = 0$, то для будь-якого $y > 2$ значення сумісної функції розподілу

$$F_{(\xi,\eta)}(x, y) = \mathrm{P}\{\xi \leq x, \eta \leq y\} = \mathrm{P}\{\xi \leq x, \eta \leq 2\} = F_{\xi,\eta)}(x, 2) = \frac{3}{10}(x^2 + x^3).$$

Це є маргінальною функцією розподілу випадкової величини ξ.

Приклад 11.8

Нехай сумісна щільність ξ та η задана рівнянням

$$f_{\xi,\eta}(x,y) = \begin{cases} Cxy^2, & \text{якщо } x \geq 0, y \geq 0, x+y \leq 1; \\ 0, & \text{інакше.} \end{cases}$$

Перевіримо, чи незалежні випадкові величини ξ та η. Для цього знайдемо маргінальні щільності:

$$f_{\xi}(x) = \begin{cases} \int_0^{1-x} Cxy^2 dy = \frac{Cx(1-x)^3}{3}, & \text{якщо } 0 \leq x \leq 1; \\ 0, & x < 0,\ x > 1. \end{cases}$$

$$f_{\eta}(y) = \begin{cases} \int_0^{1-y} Cxy^2 dy = \frac{Cy^2(1-y)^2}{2}, & \text{якщо } 0 \leq y \leq 1; \\ 0, & y < 0,\ y > 1. \end{cases}$$

Оскільки сумісна щільність не дорівнює добутку магрінальних щільностей, випадкові величини є залежними.

Приклад 11.9

Ганнуся уклала парі, що два дні підряд вона зможе з'їсти по Київському торту протягом певного часу кожен. Частина торту, яку вона в стані з'їсти в перший день, є випадковою величиною з щільністю $f_1(x) = 12x^2(1-x)$, $x \in [0,1]$. Здатність з'їсти торт в другий день не залежить від того, скільки вона з'їла в перший день, але дещо нижча: її щільність $f_2(y) = 6y(1-y)$, $y \in [0,1]$. Якщо вона з'їсть менше, ніж 3/4 торта обидва рази, вона програє парі. Яка ймовірність того, що вона виграє парі?

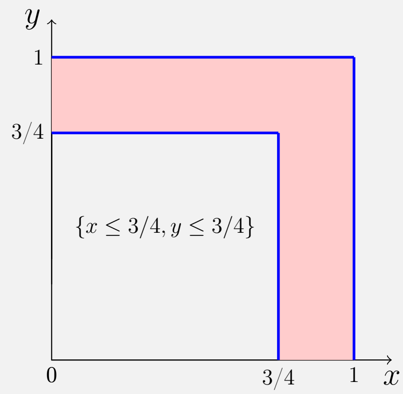
Розв'язок. Оскільки випадкові величини, які розглядаються, є незалежними, то їхня сумісна щільність дорівнює добутку маргінальних щільностей:

$$f_{1,2}(x,y) = \begin{cases} 72x^2(1-x)y(1-y), & x \in [0,1],\ y \in [0,1]; \\ 0, & \text{інакше.} \end{cases}$$

Шукана ймовірність

$$\mathrm{P}\{(\xi_1 \geq 3/4) \cup (\xi_2 \geq 3/4)\} = 1 - \mathrm{P}\{\xi_1 \leq 3/4, \xi_2 \leq 3/4\},$$

її можна знайти як інтеграл від сумісної щільності над зафарбованою областю, або, простіше, – через протилежну подію.



Отже, остаточно маємо

$$1 - \int_0^{3/4}\int_0^{3/4} 72x^2(1-x)y(1-y)dx\,dy \approx 1 - 0.623 = 0.377.$$

Приклад 11.10

Гравець в дартс гарантовано влучає дротиком в мішень у формі кола з радіусом R, але не точно: дротик може влучити в будь-яку область кола з ймовірністю, пропорційній цій області. Розмістимо центр кола в початку координат, випадкова величина ξ – це абсциса точки влучення, η – її ордината. Знайдемо
а) сумісну щільність випадкових величин ξ та η;
б) маргінальну щільність ξ.

Розв'язок. а) Рівняння кола, яке обмежує мішень, може бути записане як $x^2 + y^2 = R^2$. За умовою, точка влучення має рівномірний розподіл в колі, отже, сумісна щільність її координат є сталою в області кола і може бути записана як

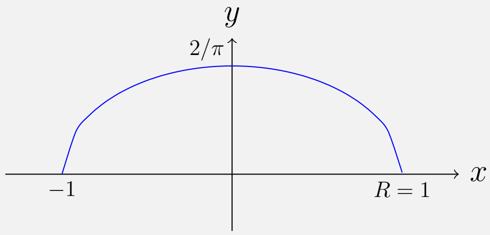
$$f_{\xi,\eta}(x,y) = \begin{cases} \dfrac{1}{\pi R^2}, & \text{якщо } x^2+y^2 \le R^2; \\ 0, & \text{інакше.} \end{cases}$$

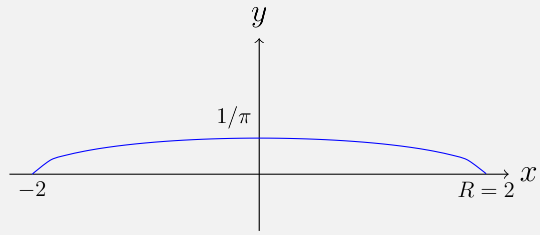
б) Очевидно, щільність $f_{\xi,\eta}(x,y)$ дорівнює 0, якщо $|x| > R$. Якщо ж $|x| < R$, то $f_{\xi,\eta}(x,y)$ дорівнює 0, якщо $x^2+y^2 > R^2$, тобто якщо $|y| > \sqrt{R^2-x^2}$, або $f_{\xi,\eta}(x,y)$ дорівнює $\dfrac{1}{\pi R^2}$, якщо $|y| \le \sqrt{R^2-x^2}$.

Враховуючи ці міркування, можемо обчислити маргінальну щільність $f_\xi(x)$ для значень $|x| \le R$:

$$f_\xi(x) = \int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} \frac{1}{\pi R^2} dy = \frac{2\sqrt{R^2-x^2}}{\pi R^2}.$$

Графік цієї функції щільності для значень $R = 1$ та $R = 2$ зображено на рисунку.





Приклад 11.11

Нехай ξ_1 та ξ_2 – випадкові величини, що мають показниковий розподіл з параметром λ. Знайти функцію розподілу випадкової величини $\eta = \dfrac{\xi_1}{\xi_1+\xi_2}$.

Розв'язок. Оскільки за умовою ξ_1 та ξ_2 – випадкові величини, що мають показниковий розподіл, то вони є додатними випадковими величинами, а отже з ймовірністю 1 маємо

$$0 < \frac{\xi_1}{\xi_1+\xi_2} < 1.$$

Тому для значень $x \notin [0,1]$

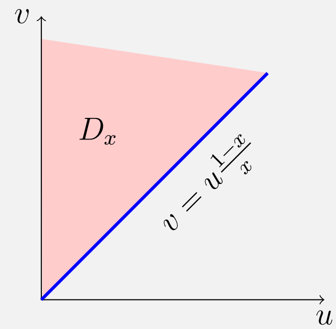
$$F_\eta(x) = \mathrm{P}\left\{\frac{\xi_1}{\xi_1+\xi_2} \le x\right\} = \begin{cases} 1, & x > 1; \\ 0, & x < 0. \end{cases}$$

Залишилось визначити значення цієї ймовірності (функції розподілу) для $x \in [0,1]$. Для таких значень аргументу

$$F_\eta(x) = \mathrm{P}\left\{\frac{\xi_1}{\xi_1+\xi_2} \le x\right\} = \mathrm{P}\{(\xi_1,\xi_2) \in D_x\},$$

де область D_x визначена такими нерівностями:

$$D_x = \left\{(u,v) : \frac{u}{u+v} \le x;\ u > 0,\ v > 0\right\} = \left\{(u,v) : v \ge \frac{u(1-x)}{x};\ u > 0,\ v > 0\right\}.$$



Функцію розподілу η для $x \in [0,1]$ тепер можемо знайти шляхом інтегрування щільності розподілу вектора (ξ_1, ξ_2) по області D_x :

$$F_\eta(x) = \mathrm{P}\left\{\frac{\xi_1}{\xi_1+\xi_2} \le x\right\} = \iint_{D_x} f_{\xi_1,\xi_2}(u,v)dudv =$$

$$\lambda^2 \iint_{D_x} e^{-\lambda u - \lambda v} dudv = \lambda^2 \int_0^{+\infty} du \int_{u\frac{1-x}{x}}^{+\infty} e^{-\lambda u-\lambda v} dv =$$

$$= \lambda^2 \int_0^{+\infty} e^{-\lambda u} \left.\frac{e^{-\lambda v}}{-\lambda}\right|_{v=u\frac{1-x}{x}}^{+\infty} du = \lambda \int_0^{+\infty} e^{-\lambda u - \lambda u \frac{1-x}{x}} du = \lambda \int_0^{+\infty} e^{-\frac{\lambda u}{x}} du = x.$$

Таким чином, випадкова величина η має рівномірний розподіл на інтервалі [0, 1].

Відповідь:$F_\eta(x) = \begin{cases} 0, & x < 0; \\ x, & x \in [0,1]; \\ 1, & x > 1. \end{cases}$

Приклад 11.12

Розглянемо стохастичний експеримент, який полягає в підкиданні двох симетричних тетраедрів з гранями, що пронумеровані цифрами від 1 до 4. Нехай випадкова величина ξ – це кількість очок, що випали на першому тетраедрі, η – максимальне з того, що випало на обох тетраедрах.

а) Записати сумісний розподіл випадкових величин ξ та η.

б) Знайти математичне сподівання добутку ξ та η.

в) Обчислити коефіцієнт кореляції ξ та η.

г) Записати умовний розподіл η за умови, що $\xi = 2$ та за умови, що $\xi = 3$.

д) Знайти умовне математичне сподівання η за умови, що $\xi = 2$.

Розв'язок. а) Випишемо сумісний розподіл цих випадкових величин:

$x\backslash y$	1	2	3	4	p_i
1	1/16	1/16	1/16	1/16	1/4
2	0	2/16	1/16	1/16	1/4
3	0	0	3/16	1/16	1/4
4	0	0	0	4/16	1/4
q_j	1/16	3/16	5/16	7/16	1

б) Тепер можемо знайти математичне сподівання добутку цих випадкових величин:

$$\mathbf{M}(\xi\eta) = \sum_{i,j=1}^{4} ij\mathrm{P}\{\xi = i, \eta = j\} = 1 \cdot 1\frac{1}{16} + 1 \cdot 2\frac{1}{16} + 1 \cdot 3\frac{1}{16} + 1 \cdot 4\frac{1}{16} +$$
$$+ 2 \cdot 2\frac{2}{16} + 2 \cdot 3\frac{1}{16} + 2 \cdot 4\frac{1}{16} + 3 \cdot 3\frac{3}{16} + 3 \cdot 4\frac{1}{16} + 4 \cdot 4\frac{4}{16} = \frac{135}{16}.$$

в) Знайдемо коефіцієнт кореляції для випадкових величин ξ та η. Очевидно, він має бути додатним, оскільки зі збільшенням ξ в.в. η також буде збільшуватись.

Ми обчислили $\mathbf{M}\xi\eta = 135/16$, з маргінальних розподілів легко можна отримати $\mathbf{M}\xi = 5/2$, $\mathbf{M}\eta = 50/16$, $\mathbf{D}\xi = 5/4$ та $\mathbf{D}\eta = 55/64$. (Перевірте!) Таким чином,

$$\mathrm{cov}(\xi, \eta) = \frac{135}{16} - \frac{5}{2} \cdot \frac{50}{16} = \frac{5}{8},$$
$$\rho(\xi, \eta) = \frac{5/8}{\sqrt{5/4}\sqrt{55/64}} = \frac{2}{\sqrt{11}}.$$

г) Тепер випишемо умовний розподіл η за умови, що $\xi = 2$.

$$\mathrm{P}\{\eta = 2|\xi = 2\} = \frac{\mathrm{P}\{\xi = 2, \eta = 2\}}{\mathrm{P}\{\xi = 2\}} = \frac{2/16}{4/16} = \frac{1}{2};$$
$$\mathrm{P}\{\eta = 3|\xi = 2\} = \frac{\mathrm{P}\{\xi = 2, \eta = 3\}}{\mathrm{P}\{\xi = 2\}} = \frac{1/16}{4/16} = \frac{1}{4};$$
$$\mathrm{P}\{\eta = 4|\xi = 2\} = \frac{\mathrm{P}\{\xi = 2, \eta = 4\}}{\mathrm{P}\{\xi = 2\}} = \frac{1/16}{4/16} = \frac{1}{4}.$$

Аналогічно, умовний розподіл η за умови, що $\xi = 3:$

$$\mathrm{P}\{\eta = y|\xi = 3\} = \begin{cases} 3/4, & \text{якщо } y = 3; \\ 1/4, & \text{якщо } y = 4. \end{cases}$$

д) Маючи умовний розподіл, можемо обчислити умовне математичне сподівання:

$$\mathbf{M}(\eta|\xi = 2) = \sum_{i=2}^{4} i \cdot \mathrm{P}_{\eta|\xi}(\eta = i|\xi = 2) = 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} = \frac{11}{4}.$$

Приклад 11.13

Зі звичайної колоди з 52 карт виймають без повернення 12 карт. Нехай ξ_1 – кількість отриманих при цьому тузів, ξ_2 – кількість двійок, ξ_3 – кількість трійок, ξ_4 – кількість четвірок. Записати приклад умовного розподілу для цього експерименту.

Розв'язок. Сумісний розподіл цих випадкових величин задається формулою

$$\mathrm{P}\{\xi_1 = x_1, \xi_2 = x_2, \xi_3 = x_3, \xi_4 = x_4\} = \frac{C_4^{x_1} C_4^{x_2} C_4^{x_3} C_4^{x_4} C_{36}^{12-x_1-x_2-x_3-x_4}}{C_{52}^{12}},$$

$x_i = 0, 1, 2, 3, 4,\ \ i = 1, 2, 3, 4$, за обмеження $\sum_i x_i \le 12$.

Можна виписати досить багато варіантів умовних розподілів для цієї моделі. Наприклад:

$$\mathrm{P}\{\xi_2 = x_2, \xi_4 = x_4 | \xi_1 = x_1, \xi_3 = x_3\} = \frac{C_4^{x_1} C_4^{x_2} C_4^{x_3} C_4^{x_4} C_{36}^{12-x_1-x_2-x_3-x_4} / C_{52}^{12}}{C_4^{x_1} C_4^{x_3} C_{44}^{12-x_1-x_3} / C_{52}^{12}},$$

зокрема

$$\mathrm{P}\{\xi_2 = x_2, \xi_4 = x_4 | \xi_1 = 2, \xi_3 = 1\} = \frac{C_4^{x_2} C_4^{x_4} C_{36}^{9-x_2-x_4}}{C_{44}^{9}}.$$

Приклад 11.14

Точка ξ обрана навмання на відрізку $[a, b]$, а потім точка η обрана навмання з проміжку $[\xi, b]$. Знайти розподіл випадкової величини η.

Розв'язок. Оскільки

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

та

$$f_{\eta|\xi}(y|\xi = x) = \begin{cases} \frac{1}{b-x}, & y \in [x, b]; \\ 0, & y \notin [x, b], \end{cases}$$

то маємо

$$f_{\xi,\eta}(x, y) = f_\xi(x) f_{\eta|\xi}(y|\xi = x) = \begin{cases} \frac{1}{b-a} \cdot \frac{1}{b-x}, & a \le x \le y \le b; \\ 0, & \text{інакше}. \end{cases}$$

Отже, для $y \in [a, b]$ маємо:

$$f_\eta(y) = \int_a^y \frac{1}{b-a} \cdot \frac{1}{b-x} dx = \frac{1}{b-a} \ln \frac{b-a}{b-y}.$$

Бачимо, що значення, ближчі до b більш ймовірні, ніж ті, що ближчі до a.

Приклад 11.15

Нехай сумісна щільність випадкових величин ξ та η задана таким чином:

$$f_{\xi,\eta}(x, y) = \begin{cases} x + y, & x, y \in [0, 1]; \\ 0 & \text{інакше.} \end{cases}$$

а) Знайти математичне сподівання добутку ξ та η.
б) Обчислити коефіцієнт кореляції ξ та η.
в) Умовну щільність та умовну функцію розподілу випадкової величини $\eta|\xi$ г) Знайти умовне математичне сподівання η за умови, що $\xi = x$.

Розв'язок. а) За визначенням маємо

$$\mathbf{M}(\xi\eta) = \int_0^1 \int_0^1 xy(x + y) dx dy = \frac{1}{3},$$

б) Неважко показати, що $\mathbf{M}\xi = \mathbf{M}\eta = 7/12$, $\mathbf{M}\xi^2 = \mathbf{M}\eta^2 = 5/12$, $\mathbf{D}\xi = \mathbf{D}\eta = 11/144$. (Перевірте!)

За визначенням маємо

$$\rho(\xi, \eta) = \frac{1/3 - 49/144}{11/144} = -\frac{1}{11}.$$

В даному випадку маємо від'ємний коефіцієнт кореляції.

в) Знайдемо спочатку маргінальну функцію щільності випадкової величини ξ, інтегруючи сумісну щільність по змінній y :

$$f_\xi(x) = \int_0^1 (x + y) dy = x + 1/2, \;\; x \in [0, 1].$$

Тепер можемо знайти умовну щільність та умовну функцію розподілу випадкової величини η :

$$f_{\eta|\xi}(y|x) = \begin{cases} \dfrac{x + y}{x + 1/2}, & y \in [0, 1], \;\; x \in [0, 1]; \\ 0, & \text{інакше,} \end{cases}$$

$$F_{\eta|\xi}(y|x) = \int_{-\infty}^{y} f_{\eta|\xi}(u|x)du = \int_0^y \frac{x+u}{x+1/2}du =$$

$$= \frac{1}{x+1/2}\int_0^y (x+u)du = \frac{xy + y^2/2}{x+1/2}, \ \ y \in [0,1].$$

Зауважимо, що випадкові величини ξ та η не є незалежними вже тому, що умовна щільність залежить від x, а отже не може бути функцією $f_\eta(y)$.

г) Умовне математичне сподівання

$$\mathbf{M}(\eta|\xi = x) = \int_0^1 y\ f_{\eta|\xi}(y|x)\ dy = \int_0^1 y\frac{x+y}{x+1/2}dy = \frac{1}{x+1/2}\left(\frac{x}{2}+\frac{1}{3}\right),\ \text{для } x \in [0,1].$$

Приклад 11.16 (залежних випадкових величин з нульовим коефіцієнтом кореляції)

Нехай ξ та η – координати точки, кинутої навмання в колі радіуса R. Покажемо, що ці випадкові величини залежні, але некорельовані.

Дійсно, маємо рівномірний двомірний розподіл точки всередині кола, отже сумісна щільність випадкових величин ξ та η не рівна нулеві лише в колі $x^2 + y^2 \le R^2$ та набуває сталих значень, рівних $1/\pi R^2$ всередині цього кола:

$$f_{\xi\eta}(x,y) = \begin{cases} 1/\pi R^2, & x^2+y^2 \le R^2; \\ 0, & x^2+y^2 > R^2. \end{cases}$$

Знайдемо маргінальні щільності координат, проінтегрувавши сумісну щільність по "зайвій" змінній:

$$f_\xi(x) = \begin{cases} \int\limits_{x^2+y^2\le R^2} \frac{1}{\pi R^2}dy, & |x| \le R; \\ 0, & |x| > R; \end{cases} = \begin{cases} \frac{2\sqrt{R^2-x^2}}{\pi R^2}, & |x| \le R; \\ 0, & |x| > R; \end{cases}$$

$$f_\eta(y) = \begin{cases} \frac{2\sqrt{R^2-y^2}}{\pi R^2}, & |y| \le R; \\ 0, & |y| > R; \end{cases}$$

Для точок, що належать кругу, маємо:

$$f_\xi(x)\cdot f_\eta(y) = \frac{4\sqrt{(R^2-x^2)(R^2-y^2)}}{\pi^2 R^4},$$

що, очевидно, не збігається зі значенням сумісної щільності $f_{\xi\eta}(x,y) = 1/\pi R^2$ в цій області. Таким чином, випадкові величини ξ та η залежні.

Покажемо, що вони некорельовані. Величини ξ та η – симетричні відносно нуля, отже $\mathbf{M}\xi = \mathbf{M}\eta = 0$. Змішаний момент знаходимо за визначенням:

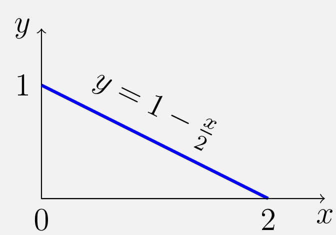
$$\mathbf{M}(\xi\eta) = \frac{1}{\pi R^2}\int\limits_{-R}^{R} xdx \int\limits_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} ydy = 0$$

Отже, координати точки, рівномірно розподіленої всередині кола, є залежними, проте некорельованими.

Приклад 11.17

Точка навмання кидається в трикутник з вершинами в точках (2,0), (0,0), та (0,1), її координати – ξ та η. Знайти коваріацію між цими величинами.

Розв'язок.



$$f_{\xi,\eta}(x,y) = \begin{cases} \frac{1}{S_\Delta}, & (x,y) \in S_\Delta; \\ 0, & (x,y) \notin S_\Delta; \end{cases} = \begin{cases} 1, & (x,y) \in S_\Delta; \\ 0, & (x,y) \notin S_\Delta. \end{cases}$$

Зауваживши, що гіпотенуза трикутника йде по прямій $y = 1 - x/2$, або $x = 2 - 2y$, легко знайти маргінальні щільності координат, проінтегрувавши сумісну щільність по "зайвій" змінній.

$$f_\xi(x) = \begin{cases} \int_0^{1-x/2} dy, & x \in [0,2]; \\ 0, & x \notin [0,2]; \end{cases} = \begin{cases} 1 - x/2, & x \in [0,2]; \\ 0, & x \notin [0,2]. \end{cases}$$

$$f_\eta(y) = \begin{cases} \int_0^{2-2y} dx, & y \in [0,1]; \\ 0, & y \notin [0,1]; \end{cases} = \begin{cases} 2 - 2y, & y \in [0,1]; \\ 0, & y \notin [0,1]. \end{cases}$$

Нам потрібні ще математичні сподівання окремих координат та їхній добутку:

$$\mathbf{M}\xi = \int_0^2 x\left(1 - \frac{x}{2}\right)dx = \frac{2}{3}$$

$$\mathbf{M}\eta = \int_0^1 y\,(2 - 2y)\,dy = \frac{1}{3}$$

$$\mathbf{M}(\xi\eta) = \iint\limits_{\Delta} xyf_{\xi\eta}(x,y)dx\,dy = \int_0^2 dx \int_0^{1-x/2} xydy = \frac{1}{6}.$$

Відповідно, коваріація дорівнює

$$\operatorname{cov}(\xi,\eta) = \mathbf{M}(\xi\eta) - \mathbf{M}\xi\mathbf{M}\eta = \frac{1}{6} - \frac{2}{3}\cdot\frac{1}{3} = -\frac{1}{18}.$$

Отже, координати є від'ємно корельованими величинами. "На пальцях" це можна пояснити таким чином: чим більша ξ, тим менше у η можливості бути великою (в середньому ордината буде меншою зі збільшенням абсциси).

Приклад 11.18

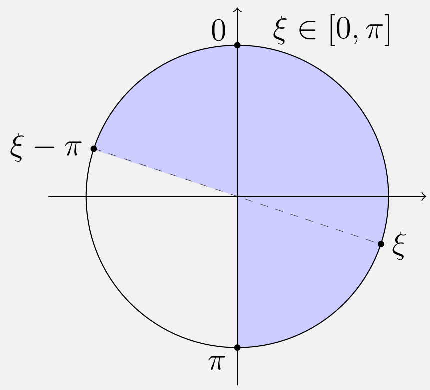
На колі навмання обрано три точки. Яка ймовірність того, що всі три точки потраплять в півколо?

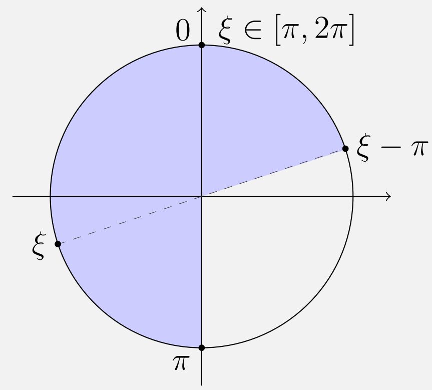
Розв'язок. Вибір точок "навмання" має на увазі, що кожна точок з однаковою ймовірністю може потрапити в будь-яку точку кола, тобто є рівномірно розподіленою на колі. Зафіксуємо положення першої точки, наприклад, таким чином, щоб вона лежала на позитивній півосі Oy, вважаючи, що центр кола збігається з початком координат. Нехай ξ – випадкова величина, яка характеризує положення другої точки, а подія A означає, що всі три точки потрапили в одне півколо. ξ має рівномірний розподіл на інтервалі $(0, 2\pi)$. Скористаємось узагальненим варіантом формули повної ймовірності:

$$\mathrm{P}\{A\} = \int_{-\infty}^{\infty} \mathrm{P}\{A|\xi = x\}\, f_\xi(x)dx.$$

Зауважимо, що для $0 < x < \pi$ $\mathrm{P}\{A|\xi = x\} = (2\pi - x)/2\pi$, оскільки, якщо положення другої точки фіксоване і рівне x, подія A матиме місце, якщо третя точка потрапить в область між $x - \pi$ та π.

Аналогічно, $\mathrm{P}\{A|\xi = x\} = (x - \pi + \pi)/2\pi = x/2\pi$ для $x \in [\pi, 2\pi)$.





Звідси маємо

$$\mathrm{P}\left\{A\right\} = \int_0^{2\pi} \mathrm{P}\left\{A|\xi = x\right\} \frac{1}{2\pi}dx = \frac{1}{2\pi}\left(\int_0^{\pi} \frac{2\pi - x}{2\pi}dx + \int_{\pi}^{2\pi} \frac{x}{2\pi}dx\right) = \frac{3}{4}.$$

Відповідь: $3/4$

Зауважимо, що цю задачу можна розв'язати в термінах геометричного визначення ймовірності, зокрема через ймовірність протилежної події (див. приклад гл. 5).

Приклад 11.19

Нехай випадкова величина η – це діаметр болта, а випадкова величина ξ – внутрішній діаметр гайки, куди повинен входити болт. Болт повинен мати діаметр 99.5 одиниць, а внутрішній діаметр гайки – 100 одиниць. Нехай процес виробництва цих деталей не бездоганний, і внаслідок цього фактично в.в. η рівномірно розподілена на інтервалі (98.5, 100.5), а ξ рівномірно розподілена на інтервалі (99, 101). Яка ймовірність того, що окремо взятий болт підійде до випадково обраної гайки, якщо "підійде" означає, що $\xi - h < \eta < \xi$ для деякої сталої $h > 0$?

Розв'язок. Припустимо, що ξ та η є незалежними випадковими величинами. Тоді

$$\mathrm{P}\left\{\xi - h < \eta < \xi\right\} = \int_{-\infty}^{+\infty} \mathrm{P}\left\{\xi - h < \eta < \xi|\xi = x\right\} f_\xi(x)dx =$$

$$= \int_{99}^{101} \mathrm{P}\left\{\xi - h < \eta < \xi|\xi = x\right\} \frac{1}{2}dx.$$

Тут

$$\mathrm{P}\left\{\xi - h < \eta < \xi | \xi = x\right\} = F_\eta(\xi|\xi = x) - F_\eta(\xi - h|\xi = x) = F_\eta(x) - F_\eta(x - h),$$

$F_\eta(x)$ – функція розподілу рівномірного закону $U[98.5; 100.5]$.

Нехай для визначеності $h = 1$. Тоді

$$\mathrm{P}\left\{\xi - 1 < \eta < \xi | \xi = x\right\} = \begin{cases} \frac{x-98.5}{2}, & \text{якщо } x \in (99, 99.5]; \\ 1/2, & \text{якщо } x \in (99.5, 100.5]; \\ \frac{100.5-(x-1)}{2}, & \text{якщо } x \in (100.5, 101]. \end{cases}$$

Звідси маємо

$$\mathrm{P}\left\{\xi - 1 < \eta < \xi\right\} = \int_{99}^{101} \mathrm{P}\left\{\xi - 1 < \eta < \xi | \xi = x\right\} \frac{1}{2} dx =$$
$$= \int_{99}^{99.5} \frac{x - 98.5}{2} \cdot \frac{1}{2} dx + \int_{99.5}^{100.5} \frac{1}{2} \cdot \frac{1}{2} dx + \int_{100.5}^{101} \frac{100.5 - (x - 1)}{2} \cdot \frac{1}{2} dx = \frac{7}{16}.$$

Відповідь: $\frac{7}{16}$

Задача 11.1. Знайти щільність розподілу суми $\xi+\eta$, якщо ξ і η незалежні, ξ має рівномірний розподіл у проміжку $[a, b]$, а η – рівномірний розподіл у проміжку $[c, d]$. Зобразити графік знайденої щільності у випадках:

а) $a = 0,\ b = 1,\ c = 0,\ d = 2$;
б) $a = 0,\ b = 1,\ c = 0,\ d = 1$;
в) $a = 0,\ b = 1,\ c = 1,\ d = 2$;
г) $a = 0,\ b = 2,\ c = 0,\ d = 3$.

Задача 11.2. Знайти щільність розподілу суми $\xi+\eta$, якщо ξ і η незалежні, ξ має рівномірний розподіл у проміжку [–1, 1], а η – показниковий розподіл з параметром l. Навести графік знайденої щільності при $\lambda = 2$.

Задача 11.3. Випадкові величини $\xi_1,\ \xi_2$ – незалежні і мають геометричний розподіл з однаковими параметрами: $\mathrm{P}\left\{\xi_i = k\right\} = q^k p$, $k = 0.1, \ldots$ Нехай $\eta = \max\left\{\xi_1,\ \xi_2\right\}$. Знайти розподіл η і сумісний розподіл η та ξ_1.

Задача 11.4. Щільність розподілу випадкового вектора (ξ, η) дорівнює

$$p(x, y) = \begin{cases} x + y, \quad 0 \le x \le 1, \quad 0 \le y \le 1, \\ 0 \quad \text{в іншому випадку.} \end{cases}$$

а) Знайти щільність кожної компоненти і довести, що випадкові величини ξ та η залежні.
б) Знайти щільність та функцію розподілу суми $\xi + \eta$, побудувати їхні графіки.

в) Знайти $\mathbf{M}(\xi+\eta)$ трьома способами: 1) використати щільність $p(x,y)$; 2) використати лінійність математичного сподівання та знайдені щільності випадкових величин ξ, η; 3) використати знайдену щільність суми $\xi+\eta$.

Задача 11.5. Випадкові величини ξ та η незалежні та мають щільності

$$f_\xi(x)=\left(x+\frac{1}{2}\right)\cdot I_{x\in[0,1]},\qquad f_\eta(y)=\left(y+\frac{1}{2}\right)\cdot I_{y\in[0,1]},$$

де

$$I_{x\in A}=\begin{cases}1, & x\in A,\\ 0, & x\notin A.\end{cases}$$

Знайти щільність суми $\xi+\eta$ та побудувати її графік. Знайти також $\mathbf{M}(\xi+\eta)$.

Задача 11.6. Нехай ξ_1 та ξ_2 – незалежні нормально розподілені $N(0.1)$ випадкові величини. Обчислити $\mathrm{P}\left\{{\xi_1}^2+{\xi_2}^2\le R^2\right\}$.

Задача 11.7. Щільність розподілу випадкового вектора (ξ,η) дорівнює

$$p(x,y)=\begin{cases}\lambda^2e^{-\lambda x}, & 0<y<x,\\ 0 & \text{в інших випадках.}\end{cases}$$

Довести, що випадкові величини ξ та η залежні. Знайти щільність та функцію розподілу суми $\xi+\eta$ та побудувати їхні графіки при $\lambda=2$.

Задача 11.8. Випадкові величини ξ та η незалежні та мають щільності

$$f_\xi(x)=x\lambda^2e^{-\lambda x}I_{x>0},\qquad f_\eta(y)=\lambda e^{-\lambda y}I_{y>0}.$$

Знайти щільність суми $\xi+\eta$ та побудувати її графік.

Задача 11.9. Бігуни A та B стартували одночасно і фінішують у моменти часу ξ та η, які є незалежними, однаково розподіленими за показниковим законом з параметром λ випадковими величинами. Знайти ймовірність того, що час переможця перевищує той запас часу, на який він випередив супротивника.

Задача 11.10. Щільність розподілу випадкового вектора (ξ,η) дорівнює

$$p(x,y)=\begin{cases}8xy(1-x^2), & 0<x<1,\\ 0 & \text{в інших випадках.}\end{cases}$$

Довести, що випадкові величини ξ та η незалежні. Знайти щільність та функцію розподілу добутку $\xi\eta$, побудувати їхні графіки.

Задача 11.11. Щільність розподілу випадкового вектора (ξ,η) дорівнює

$$p(x,y)=\begin{cases}xe^{-x(1+y)}, & x>0,\\ 0 & \text{в інших випадках.}\end{cases}$$

Довести, що випадкові величини ξ та η залежні. Знайти розподіл добутку $\xi\eta$.

Задача 11.12. Щільність розподілу випадкового вектора (ξ, η) має вид

$$f(x,y) = \begin{cases} [(1+ax)(1+ay)-a]e^{-x-y-axy}, & x>0, y>0, \\ 0 & \text{в інших випадках.} \end{cases}$$

Знайти функцію розподілу (ξ, η) і щільності ξ, η.

Задача 11.13. Щільність розподілу випадкового вектора (ξ, η) дорівнює

$$f(x,y) = \begin{cases} 24x^2y(1-x), & 0<x<1, 0<y<1, \\ 0 & \text{в інших випадках.} \end{cases}$$

Знайти щільності розподілу ξ і η. Довести, що ξ і η – незалежні.

Задача 11.14. Страхова компанія вивчає страхові випадки, спричинені торнадо, за угодами страхування ферм. Нехай X– частина збитків, що пов'язана з пошкодженням будинку, а Y – частина збитків, що пов'язана з пошкодженням іншого майна. Сумісний розподіл випадкових величин X та Y має щільність

$$f(x,y) = \begin{cases} 6(1-(x+y)), & x>0, y>0, x+y<1, \\ 0 & \text{в інших випадках.} \end{cases}$$

Визначити ймовірність того, що збитки, спричинені будинку, складатимуть менше 20 % від загальних збитків.

Задача 11.15. На інтервалі (0, 1) зафіксовано точку a. Випадкова величина ξ рівномірно розподілена у відрізку [0.1]. Нехай η – відстань від точки ξ до a. Знайти $\operatorname{cov}(\xi, \eta)$.

Задача 11.16. Випадкова величина ξ має показниковий розподіл з параметром λ, $\eta = a\xi + b$. Знайти коефіцієнт кореляції між ξ і η.

Задача 11.17. Випадкові величини ξ_1 і ξ_2 незалежні і мають нормальний розподіл з параметрами (a_1, σ_1^2) і (a_2, σ_2^2) відповідно. Знайти коваріацію випадкових величин $\eta_1 = a\xi_1 - b\xi_2$ і $\eta_2 = a\xi_1 + b\xi_2$.

Задача 11.18. Знайти коваріацію між компонентами випадкового вектора (ξ, η), щільність розподілу якого є

$$f(x,y) = \begin{cases} \frac{1}{2}, & |x|+|y| \le 1, \\ 0 & \text{в інших випадках.} \end{cases}$$

та знайти дисперсії компонент випадкового вектора (ξ, η).

Задача 11.19. Випадковий вектор (ξ, η) має рівномірний розподіл в області, яка обмежена трикутником з вершинами у точках (0.0), (1.0), (0.1). Знайти $\operatorname{cov}(\xi, \eta)$.

Задача 11.20. Частковий випадок стійкості гауссівського розподілу Нехай незалежні випадкові величини η та ξ мають стандартний нормальний розподіл. Довести за теоремою згортки, що їхня сума має нормальний розподіл з параметрами 0 та 2.

Задача 11.21. Виразити для двовимірного випадкового вектора (ξ_1, ξ_2) ймовірність потрапити у прямокутник $[a_1, b_1] \times [a_2, b_2]$ через його сумісну функцію розподілу.

Задача 11.22. Нехай ξ та η – незалежні випадкові величини, причому дисперсія ξ відмінна від нуля. Довести, що ξ та $\xi + \eta$ залежні.

Задача 11.23. (Продовження попередньої задачі) Нехай ξ та η будуть не тільки незалежними, але й однаково розподіленими випадковими величинами із ненульовими дисперсіями. Знайти коефіцієнт кореляції між ξ та $\xi + \eta$.

Задача 11.24. Нехай випадкова величина φ має рівномірний розподіл в $[0, 2\pi]$, $\xi = \cos\varphi$, $\eta = \sin\varphi$. Довести, що ξ та η залежні, але некорельовані.

Задача 11.25.* Кажуть, що випадкова величина ξ має **гамма-розподіл** $\Gamma_{a,\lambda}$ з параметрами $a > 0$, $\lambda > 0$, якщо вона має щільність розподілу

$$f_\xi(x) = \begin{cases} 0, & x \le 0, \\ c \cdot x^{\lambda-1} e^{-ax} & x < 0. \end{cases}$$

де $c = \dfrac{a^\lambda}{\Gamma(\lambda)}$, $\Gamma(\lambda) = \int\limits_0^\infty x^{\lambda-1} e^{-1} dx$. Нехай незалежні випадкові величини $\xi_1, \ldots, \xi_n$ мають розподіл $\Gamma_{a,1}$. Довести, що тоді $\xi_1 + \ldots + \xi_n$ має розподіл $\Gamma_{a,n}$.

Задача 11.26.* Довести стійкість таких розподілів: біноміального, пуассонівського, Коші, гауссівського та гамма-розподілу (**стійкість розподілу** – це коли сума незалежних випадкових величин із розподілами цього типу є випадковою величиною цього ж типу).

Задача 11.27.* Нехай $\xi_0, \xi_1, \ldots, \xi_n$ – незалежні випадкові величини, які мають рівномірний розподіл у відрізку [0, 1]. Треба знайти щільність розподілу випадкової величини

$$\eta_n = \prod_{k=0}^{n} \xi_k$$

та її математичне сподівання.

Задача 11.28. (Продовження попередньої задачі) Нехай $\xi_0, \xi_1, \ldots, \xi_n$ – незалежні випадкові величини, які мають рівномірний розподіл у відрізку [0, 1]. Треба знайти розподіл випадкової величини

$$\eta = \min\left\{ n : \prod_{k=0}^{n} \xi_k < e^{-\lambda} \right\}, \quad \lambda > 0.$$

Задача 11.29. Випадкові величини $\xi_0, \xi_1, \ldots, \xi_n$ незалежні і мають однаковий показниковий розподіл з параметром $\lambda > 0$. Довести, що випадкові величини

$$\max\{\xi_1, \ldots, \xi_n\} \quad \text{і} \quad \sum_{k=1}^{n} \frac{\xi_k}{k}$$

однаково розподілені.

Задача 11.30. Знайти ймовірність того, що рівняння

$$x^2 - 2ax + b = 0$$

має комплексні корені, якщо коефіцієнти a і b є незалежними однаково розподіленими випадковими величинами, які:

а) мають рівномірний розподіл у відрізку [0.1];

б) мають показниковий розподіл з параметром $\lambda > 0$.

Задача 11.31. Нехай ξ і η незалежні випадкові величини, причому ξ має рівномірний розподіл у проміжку [0.1], а η має пуассонівський розподіл з параметром $\lambda > 0$. Треба знайти щільність розподілу $\xi + \eta$. Побудувати графік знайденої щільності при $\lambda = 3$.

Задача 11.32. Незалежні величини $\xi_1, \ldots, \xi_n$ додатні і мають однаковий невироджений розподіл. Нехай $\eta_k = \dfrac{\xi_k}{\xi_1 + ... + \xi_n}$ $(k = 1, 2, \ldots, n)$. Треба знайти:

а) математичне сподівання η_k; б) коефіцієнт кореляції між η_k і η_m;

в) коефіцієнт кореляції між $\eta_1 + ... + \eta_k$ і $\eta_1 + ... + \eta_m$.

Задача 11.33. Випадкові величини $\xi_1, \ldots, \xi_n$ незалежні і рівномірно розподілені у відрізку [0, 1]. Треба знайти $\mathrm{P}\{\xi_1 + \xi_2 + \ldots + \xi_n \le x\}$ для $0 \le x \le 1$.

Задача 11.34. Нехай $\xi_1, \xi_2, \ldots$ – послідовність незалежних, рівномірно розподілених у проміжку [0.1] випадкових величин. Нехай η – випадкова величина, яка дорівнює тому значенню k, при якому вперше сума $\xi_1 + \ldots + \xi_k$ буде перевищувати рівень 1. Треба знайти $\mathbf{M}\eta$.

Задача 11.35. Випадковий вектор $\xi = (\xi_1, \xi_2)$ має рівномірний розподіл у трикутнику з вершинами в точках (1, 0), (0, -1), (-1, 0). Знайти:

а) розподіл $(\xi_1 + \xi_2)/2$, б) коефіцієнт кореляції $\rho(\xi_1, \xi_2)$.

Задача 11.36. "Задача про зустріч" – узагальнення. N осіб домовились про зустріч. Кожен з них приходить на місце зустрічі у випадковий момент з проміжку $[t, t+T]$ і чекає не більше часу a $(a < T)$. Яка ймовірність того, що всі N осіб зустрінуться, якщо моменти їхнього приходу на місце зустрічі незалежні і рівномірно розподілені в $[t, t+T]$?

Задача 11.37. Нехай ξ та η – випадкові величини з сумісним розподілом, заданим таблицею:

$\xi \backslash \eta$	2	3	4
0	1/32	0	b
1	0	5/32	8/32
2	a	0	2/32
3	3/32	6/32	0

Знайти

а) ймовірності a та b, якщо відомо, що $\mathrm{P}\{\xi = \eta\} = 1/4$;

б) $\mathrm{P}\{\xi\eta = 0\}$;

в) $\mathbf{M}(\xi\eta)$;

г) якщо $F_{\xi,\eta}(x, y)$ – функція розподілу ξ та η, знайти $F_{\xi,\eta}(-1, 3)$, $F_{\xi,\eta}(1.5, 2.5)$, $F_{\xi,\eta}(6, 3)$.

Задача 11.38. Двічі підкидається симетричний гральний кубик. Нехай випадкова величина ξ – сума очок на двох кубиках, η – максимум з того, що випало на двох кубиках. Знайти

а) сумісний розподіл $\xi,\ \eta$;

б) ймовірності $\mathrm{P}\{\xi \le 7, \eta \le 4\}$; $\mathrm{P}\{\xi = 8, \eta \le 2\}$; $\mathrm{P}\{3 \le \xi \le 8, 1 \le \eta \le 5\}$;

в) $\mathrm{P}\{\eta = 3|\xi = 6\}$; $\mathrm{P}\{\eta \le 4|\xi = 6\}$; $\mathrm{P}\{4 \le \eta \le 6|\xi \ge 10\}$.

Задача 11.39. Обчислити $\mathbf{M}(\xi+\eta)$ та $\mathbf{M}(\xi\eta)$ для випадкових величин з сумісним розподілом, заданим таблицею

$\xi\backslash\eta$	0	1	2
0	0	1/4	0
1	1/4	0	1/4
2	0	1/4	0

Задача 11.40. Для випадкових величин з сумісним розподілом, заданим таблицею

$\eta\backslash\xi$	0	1	2
-1	1/6	1/6	1/6
1	0	1/2	0

а) знайти $\mathbf{M}(\xi\eta)$;

б) чи є випадкові величини ξ та η незалежними? некорельованими?

в) обчислити $\mathbf{D}(\xi + \eta)$ та $\mathbf{D}(\xi - \eta)$.

Задача 11.41. Невелике виробництво робить вази ручної роботи у формі циліндра висотою h та радіусом r см, причому h та r є незалежними випадковими величинами, що мають рівномірний розподіл на інтервалах $[9, 11]$ та $[28, 32]$ відповідно. Об'єм вази також є випадковою величиною. Знайти математичне сподівання об'єму вази.

Задача 11.42. З колоди, що містить 36 карт, обрано навмання 2 карти. Нехай в.в. ξ – це кількість тузів, η – кількість карт чирвової масті. Знайти сумісний розподіл випадкових величин (ξ, η) та маргінальні розподіли в.в. ξ та η. Чи є ξ та η незалежними випадковими величинами?

Задача 11.43. В урні 5 кульок, з яких 2 червоних та 3 зелені. Без повернення навмання витягують 3 кулі. Нехай ξ – кількість червоних, η – кількість зелених витягнутих куль. Знайти

а) сумісний розподіл випадкових величин (ξ, η);

б) маргінальні розподіли в.в. ξ та η;

в) чи залежні в.в. ξ та η?

г) сумісний розподіл випадкових величин $(\xi, \eta - \xi)$.

Задача 11.44. Нехай ξ та η – час роботи до відмови двох частин механізму. Їхня сумісна щільність

$$f_{\xi,\eta}(x,y) = \begin{cases} xe^{-x(1+y)}, & x \geq 0, y \geq 0; \\ 0, & \text{інакше.} \end{cases}$$

Знайти:

а) $\mathrm{P}\{\xi > 5\}$; б) $\mathrm{P}\{\max(\xi,\eta) > 2\}$;

в) чи є ξ та η незалежними випадковими величинами?

Задача 11.45. Відділення банку має працівника (першого), що працює з пересічними клієнтами, та працівника (другого) для обслуговування VIP-клієнтів. Нехай ξ та η – долі часу (робочого дня), коли обидва працівника зайняті. Сумісна щільність цих випадкових величин дорівнює

$$f(x,y) = \begin{cases} C(2x^2 + y^2), & x \in [0,1],\ y \in [0,1]; \\ 0, & \text{інакше.} \end{cases}$$

Знайти:

а) сталу C; б) маргінальні щільності ξ та η;

в) ймовірність того, що другий працівник буде зайнятий більше, ніж 80 % часу, а перший буде зайнятий менше 20 % часу.

Задача 11.46. Фразу "палку навмання розламали на три частини" можна розуміти порізному. Нехай маємо палку довжиною 1 (відрізок $[0,1]$) і нехай випадкові величини $0 < \xi < \eta < 1$ – це точки зламів. Можливі, наприклад, такі варіанти процедури поділу:

1) Точка з координатами (ζ_1, ζ_2) навмання обирається в одиничному квадраті, і в якості точок поділу беруться $\xi = \min(\zeta_1, \zeta_2)$, $\eta = \max(\zeta_1, \zeta_2)$.

2) Точка ζ_1 обирається навмання з відрізку $[0,1]$; якщо $\zeta_1 < 1/2$, то ζ_2 вибирається навмання на відрізку $[\zeta_1, 1]$, якщо $\zeta_1 \geq 1/2$, то з відрізку $[0, \zeta_1]$. Далі ξ та η визначаються як в п.1.

3) ξ вибирається навмання з $[0,1]$, а потім η обирається навмання з $[\xi, 1]$.

4) ζ_1 вибирається навмання на $[0,1]$; з ймовірностями ζ_1 та $1 - \zeta_1$ обирається відрізок $[0, \zeta_1]$ або $[\zeta_1, 1]$ відповідно. Потім ζ_2 обирається навмання з обраного інтервалу. Далі ξ та η визначаються як в п.1.

Для кожного з варіантів знайти сумісну щільність ξ та η та їхні маргінальні щільності. Чи є серед способів 1-4 еквівалентні?

Задача 11.47. Нехай (ξ, η) – координати точки, кинутої навмання в трикутник з вершинами (0,0), (0,1), (1,1).

а) Знайти їхню сумісну функцію розподілу.

б) Знайти сумісну щільність $f_{\xi,\eta}(x,y)$ та маргінальні щільності $f_{\xi,\eta}(x,y)$, $f_{\xi,\eta}(x,y)$.

в) Точка з координатами (ζ, δ) навмання кидається в квадрат з вершинами (0,0), (0,1), (1,0), (1,1). Показати, що $\min\{\zeta, \delta\}$ має такий самий розподіл, що й випадкова величина ξ, а $\max\{\zeta, \delta\}$ – розподіл, що збігається з розподілом η.

г) Знайти $\mathbf{M}\xi$, $\mathbf{D}\xi$, $\mathbf{M}\eta$, $\mathbf{D}\eta$, $\mathbf{D}(\xi+\eta)$.

д) Обчислити коваріацію та коефіцієнт кореляції випадкових величин ξ та η.

Задача 11.48. Обчислити коефіцієнт кореляції $\rho(\xi, \xi^2)$, якщо випадкова величина ξ має рівномірний розподіл на відрізку $[0, a]$.

Задача 11.49. Нехай ξ та η – випадкові величини. Виразіть $\operatorname{cov}(\xi, \xi+\eta)$ через $\mathbf{D}\xi$ та $\operatorname{cov}(\xi, \eta)$. Чи є ξ та $\xi+\eta$ додатньокорельованими, від'ємнокорельованими, некорельованими, чи може бути будь-що? Те саме питання, якщо ξ та η – некорельовані випадкові величини.

Задача 11.50. Коефіцієнт кореляції між випадковими величинами ξ та η дорівнює 0.8. Знайти коефіцієнт кореляції між величинами $\alpha = 2\xi - 1$ та $\beta = -\eta + 3$.

Задача 11.51. Тричі підкидається симетричний гральний кубик. Нехай ξ – кількість одиниць, які при цьому випали, η – кількість шісток. Знайти

а) сумісний розподіл ξ та η;

б) коефіцієнт кореляції між випадковими величинами ξ та η.

Задача 11.52. Сумісна щільність випадкових величин ξ та η задана формулою

$$f(x,y) = \begin{cases} y^2(xy^3+1), & 0 \le x \le C,\ 0 \le y \le 1; \\ 0, & \text{інакше.} \end{cases}$$

Знайти:

а) сталу C; б) $\mathrm{P}\{\xi \le \eta\}$; в) маргінальні щільності ξ та η;

г) чи є ξ та η незалежними випадковими величинами?

Задача 11.53. Сумісна щільність випадкових величин ξ та η задана формулою

$$f(x,y) = \begin{cases} x+y, & 0 \le x \le 1,\ 0 \le y \le 1; \\ 0, & \text{інакше.} \end{cases}$$

Знайти $\mathrm{P}\{\eta \le \sqrt{\xi} - 1/2\}$.

Задача 11.54. Сумісна щільність випадкових величин ξ та η задана формулою

$$f(x,y) = \begin{cases} Cxy^2, & 0 \le x \le 1,\ 0 \le y \le 1; \\ 0, & \text{інакше.} \end{cases}$$

Знайти:

а) сталу C; б) $\mathrm{P}\{\xi^2 \le \eta \le \xi\}$;

в) сумісну функцію розподілу ξ та η.

Задача 11.55. Сумісна щільність випадкових величин ξ та η задана формулою

$$f(x,y) = \begin{cases} Cx^2y, & 0 \le x \le 1,\ x^2 \le y \le 1; \\ 0, & \text{інакше.} \end{cases}$$

Знайти:

а) Знайти сталу C; б) $\mathrm{P}\{\xi \geq \eta\}$.

Задача 11.56. Двічі підкидають симетричний гральний кубик. Нехай випадкова величина ξ – кількість очок, що випали на першому кубику, η – модуль різниці того, що випало на двох кубиках. Знайти $\mathrm{P}\left\{\xi - \text{непарна} | \eta = 5\right\}$, $\mathrm{P}\left\{\eta > 2 | \xi = 5\right\}$.

Задача 11.57. Кількість ДТП, що трапились упродовж дня в двох районах Києва є незалежними випадковими величинами ξ та η, що мають розподіл Пуассона з параметрами λ_1 та λ_2 відповідно. Знайти ймовірність того, що

а) в двох районах трапилось не менше двох ДТП;

б) загальна кількість ДТП у двох районах дорівнює m;

в) в першому районі трапилось k ДТП, якщо загальна кількість ДТП в обох районах n.

Задача 11.58. Нехай сумісна щільність випадкових величин ξ та η

$$f_{\xi,\eta}(x,y) = \begin{cases} Cxy^2, & 0 \leq x \leq 1,\ 0 \leq x \leq 1,\ x + y \geq 1; \\ 0, & \text{інакше.} \end{cases}$$

Знайти маргінальні щільності випадкових величин ξ та η та умовну щільність $f_{\xi|\eta}(x|y)$.

Задача 11.59. Нехай ξ та η мають сумісну щільність

$$f_{\xi,\eta}(x,y) = \begin{cases} \lambda^2 e^{-\lambda y}, & 0 \leq x \leq y; \\ 0, & \text{інакше.} \end{cases}$$

Знайти

а) сумісну функцію розподілу (ξ, η);

б) маргінальні щільності випадкових величин ξ та η;

в) умовну щільність $\eta|\xi$.

Задача 11.60. Випадкові величини ξ та η мають сумісну щільність

$$f_{\xi,\eta}(x,y) = \begin{cases} Cx(3x + 2y), & x, y \in [0,1]; \\ 0, & \text{інакше.} \end{cases}$$

Знайти:

а) сталу C; б) $\mathbf{M}(\eta|\xi = x)$; в) $\mathbf{M}(\eta)$.

Задача 11.61. Кількість ДТП, що трапляються упродовж тижня в деякому місті, є випадковою величиною з середнім μ та дисперсією σ^2. Кількість людей, що постраждали під час ДТП – незалежні випадкові величини з середнім m та дисперсією κ^2. Знайти середнє та дисперсію кількості людей, що постраждали в ДТП упродовж тижня.

Задача 11.62. Нехай ξ та η – координати точки, кинутої навмання в трикутник з вершинами в точках (0,0), (2,0) та (3,1). Знайти

а) $\mathbf{M}(\xi|\eta)$ $\mathbf{M}(\eta|\xi)$; , б) $\mathbf{D}(\xi|\eta)$, $\mathbf{D}(\eta|\xi)$;

в) $\mathbf{M}(\xi)$ та $\mathbf{M}(\eta)$, використовуючи властивості умовних матем. сподівання та дисперсії.

12 ГРАНИЧНІ ТЕОРЕМИ

Розглянемо послідовність випадкових величин $\xi_1, \xi_2, \ldots$ та випадкову величину ξ, які задані на одному й тому ж ймовірнісному просторі $(\Omega,\ \Re,\ \mathrm{P})$. Нехай $A \in \Re$ і $\mathrm{P}(A) = 0$. Кажуть, що послідовність ξ_n, $n \geq 1$, **збігається до випадкової величини ξ з ймовірністю 1 (або майже напевно)**, якщо для кожного $\omega \in \Omega \backslash A$

$$\lim_{n\to\infty} \xi_n(\omega) = \xi(\omega).$$

Збіжність з ймовірністю 1 позначають ще так: $\lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega) \pmod{\mathrm{P}}$, або $\xi_n \underset{n\to\infty}{\overset{\text{м.н.}}{\Longrightarrow}} \xi$.

Послідовність ξ_n, $n \geq 1$, **збігається до випадкової величини ξ за ймовірністю**, якщо для будь-якого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\left(|\xi_n(\omega) - \xi(\omega)| > \varepsilon\right) = 0.$$

Для збіжності за імовірністю використовують наступні позначення:

$$\xi_n \underset{n\to\infty}{\overset{P}{\Longrightarrow}} \xi \quad \text{або} \quad \mathrm{P}\text{-}\lim_{n\to\infty} \xi_n(\omega) = \xi(\omega).$$

Нехай $L_p,\ p \geq 1$, множина випадкових величин, для яких $\mathbf{M}|\xi|^p < \infty$. Кажуть, що послідовність $\xi_n \in L_1,\ n \geq 1$, **збігається до випадкової величини $\xi \in L_1$ в середньому**, якщо

$$\lim_{n\to\infty} \mathbf{M}|\xi_n - \xi| = 0.$$

Запис $\xi_n \underset{n\to\infty}{\overset{c}{\Longrightarrow}} \xi$ використовують для позначення збіжності в середньому.

Послідовність випадкових величин $\xi_n \in L_2,\ n \geq 1$, **збігається до випадкової величини $\xi \in L_2$ в середньоквадратичному**, якщо

$$\lim_{n\to\infty} \mathbf{M}(\xi_n - \xi)^2 = 0.$$

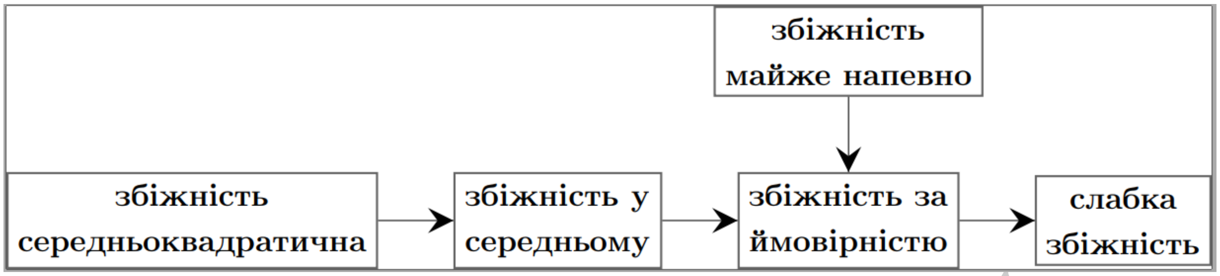
Цей тип збіжності позначають символом $\xi_n \underset{n\to\infty}{\overset{\text{с.к.}}{\Longrightarrow}} \xi$.

Кажуть, що послідовність ξ_n, $n \geq 1$, випадкових величин **збігається до випадкової величини ξ слабко**, або **за розподілом**, якщо

$$\lim_{n\to\infty} \mathrm{P}(\xi_n \leq x) = \mathrm{P}(\xi \leq x).$$

для кожного x, яке є точкою неперервності функції розподілу $F(x) = \mathrm{P}(\xi \leq x)$. Слаба збіжність позначається так: $\xi_n \underset{n\to\infty}{\overset{\text{сл.}}{\Longrightarrow}} \xi$ або $\xi_n \underset{n\to\infty}{\overset{d}{\Longrightarrow}} \xi$. Відзначимо, що цей тип збіжності визначається через відповідну збіжність функції розподілу. Тому не обов'язково, щоб ξ_n, $n \geq 1$, і ξ були задані на одному імовірнісному просторі.

Співвідношення між різними типами збіжності ілюструє наступна діаграма



Кажуть, що для послідовності випадкових величин ξ_n, $n \geq 1$, виконується **слабкий закон великих чисел (ЗВЧ)**, якщо

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i \xrightarrow[n\to\infty]{P} 0. \tag{12.1}$$

Якщо в (12.1) збіжність за ймовірністю замінити на збіжність з імовірністю 1, то отримаємо **посилений ЗВЧ.**

Наступні теореми дають достатні умови для виконання слабкого ЗВЧ.

Теорема 12.1 (ЗВЧ у формі Чебишова) Нехай ξ_n, $n \geq 1$, – послідовність незалежних випадкових величин, дисперсії яких обмежені зверху: $\mathbf{D}\xi_i \leq C$, $i = 1, 2, \ldots$ Тоді для ξ_n, $n \geq 1$, має місце слабкий ЗВЧ.

Доведення цієї теореми спирається на **нерівність Чебишова**: для довільної випадкової величини ξ із скінченним математичним сподіванням та дисперсією і довільного $\forall \varepsilon > 0$

$$\mathrm{P}\left\{|\xi - \mathbf{M}\xi| > \varepsilon\right\} \leq \frac{\mathbf{D}\xi}{\varepsilon^2}.$$

Теорема 12.2 (ЗВЧ у формі Хінчина) Нехай ξ_n, $n \geq 1$, – послідовність незалежних однаково розподілених випадкових величин зі скінченним математичним сподіванням $\mathbf{M}\xi_1 = a$. Тоді $\forall \varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - a\right| < \varepsilon\right\} = 1.$$

Теорема 12.3 (ЗВЧ у формі Маркова) Нехай $\xi_n,\ n \geq 1$, – будь-яка послідовність випадкових величин. Для виконання слабкого ЗВЧ достатньо, щоб при $n \to \infty$

$$\frac{1}{n^2}\mathbf{D}\sum_{k=1}^{n}\xi_k \longrightarrow 0.$$

У випадку незалежних, однаково розподілених випадкових величин $\xi_n,\ n \geq 1$, наступна теорема дає необхідні та достатні умови, щоб ця послідовність підкорялась слабому закону великих чисел.

Теорема 12.4 (О.М. Колмогоров) Для послідовності незалежних, однаково розподілених випадкових величин $\xi_n,\ n \geq 1$, посилений закон великих чисел виконується тоді та лише тоді, коли існує скінченне математичне сподівання $\mathbf{M}\xi_i = a$.

Наступна теорема теж належить О.М. Колмогорову:

Теорема 12.5 (О.М. Колмогоров) Якщо $\xi_n,\ n \geq 1$, незалежні та задовольняють умові

$$\sum_{n=1}^{\infty}\frac{\mathbf{D}\xi_n}{n^2} < \infty,$$

тоді для $\xi_n,\ n \geq 1$, виконується посилений ЗВЧ.

Для послідовності випадкових величин $\xi_n,\ n \geq 1$, виконується **центральне граничне твердження (ЦГТ)**, якщо можна вказати послідовності чисел A_n, B_n такі, що

$$\mathrm{P}\left(\frac{S_n - A_n}{B_n} \leq x\right) \xrightarrow[n\to\infty]{} \quad \Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy, \qquad x \in (-\infty, \infty),$$

де $S_n = \xi_1 + \xi_2 + \ldots + \xi_n$.

Теорема 12.6 (ЦГТ за умови Ліндеберга) Нехай випадкові величини ξ_n, $n \geq 1$, будуть незалежними і такими, що існують $\mathbf{M}\xi_k = a_k$, $\mathbf{D}\xi_k = \sigma_k^2 < \infty$. Позначимо $A_n = \sum\limits_{k=1}^{n} a_k$, $B_n^2 = \sum\limits_{k=1}^{n} \sigma_k^2$. Якщо для кожного $\varepsilon > 0$

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\varepsilon B_n} (x-a_k)^2 dF_k(x) = 0, \tag{12.2}$$

де $F_k(x)$ функція розподілу для ξ_k, то

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{S_n - A_n}{B_n} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2}} dy$$

рівномірно по x.

Умова 12.2 називається **умовою Ліндеберга**. З попередньої теореми випливає такий результат:

Теорема 12.7 Нехай випадкові величини ξ_n, $n \geq 1$, будуть незалежними однаково розподіленими і такими, що $\mathbf{M}\xi_k = a$, $\mathbf{D}\xi_k = \sigma^2 < \infty$. Тоді

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\sum\limits_{k=1}^{n} \xi_k - na}{\sqrt{n}\sigma} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2}} dy$$

рівномірно по х.

Наступна теорема дає достатні моментні умови для того, щоб для послідовності незалежних випадкових величин виконувалось центральне граничне твердження

Теорема 12.8 (ЦГТ за умови Ляпунова) Нехай випадкові величини ξ_n, $n \geq 1$, будуть незалежними і такими, що існують $\mathbf{M}\xi_k = a_k$, $\mathbf{D}\xi_k = \sigma_k^2 < \infty$. Позначимо $A_n = \sum\limits_{k=1}^{n} a_k$, $B_n^2 = \sum\limits_{k=1}^{n} \sigma_k^2$. Якщо $\mathbf{M}|\xi_k - a_k|^{2+\delta} < \infty$ для кожного $k \geq 1$ та деякого $\delta > 0$ і виконується умова

$$\lim_{n\to\infty} \frac{\sum\limits_{k=1}^{n} \mathbf{M}|\xi_k - a_k|^{2+\delta}}{B_n^{2+\delta}} = 0,$$

то

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{S_n - A_n}{B_n} \le x\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2}}\, dy.$$

Приклад 12.1

Припустимо, що розмір одного кроку пішохода рівномірно розподілений в інтервалі від 70 см до 80 см і розміри різних кроків незалежні. Знайти ймовірність p_0 того, що за 10 000 кроків пішохід пройде відстань не менше 7.49 км і не більше 7.51 км.

Розв'язок: Позначимо через ξ_i довжину i-го кроку та розглянемо

$$S = \sum_{i=1}^{10^4} \xi_i.$$

Очевидно, що

$$\mathbf{M}S = \sum_{i=1}^{10^4} \mathbf{M}\xi_i = 10^4 \cdot 0.75 = 7500$$

та

$$\mathbf{D}S = \sum_{i=1}^{10^4} \mathbf{D}\xi_i = 10^4 \cdot \frac{1}{1200} = \frac{25}{3}.$$

За центральною граничною теоремою:

$$p_0 = \mathrm{P}\left\{7490 \le S \le 7510\right\} = \mathrm{P}\left\{\frac{7490 - 7500}{\sqrt{25/3}} \le \frac{S - 7500}{\sqrt{25/3}} \le \frac{7510 - 7500}{\sqrt{25/3}}\right\} =$$

$$= \mathrm{P}\left\{-2\sqrt{3} \le \frac{S - 7500}{\sqrt{25/3}} \le 2\sqrt{3}\right\} \approx \Phi\left(2\sqrt{3}\right) - \Phi\left(-2\sqrt{3}\right) =$$

$$= \Phi\left(2\sqrt{3}\right) - \left(1 - \Phi\left(2\sqrt{3}\right)\right) = 2\Phi\left(2\sqrt{3}\right) - 1 \approx 2 \cdot 0.99973 - 1 \approx 0.99946$$

Відповідь: $p_0 \approx 0,99946.$

Приклад 12.2

За допомогою центральної граничної теореми довести, що

$$\lim_{n\to\infty} e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\frac{1}{2}.$$

Доведення: Нехай η_n - випадкова величина, яка має розподіл Пуассона з параметром n. Тоді

$$\mathrm{P}\left\{\eta_n \le n\right\}=\sum_{k=0}^{n}\mathrm{P}\left\{\eta_n = k\right\}=e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}.$$

З іншого боку, η_n можна подати у вигляді суми $\eta_n=\xi_1+\ldots+\xi_n$, де всі ξ_k є незалежними і однаково розподіленими за законом Пуассона з $\mathbf{M}\xi_k=\mathbf{D}\xi_k=1$. Тому за центральною граничною теоремою η_n є асимптотично нормальною з параметрами $(n,\ \sqrt{n})$, тому $\mathrm{P}\left\{\eta_n\le n\right\}\to 1/2$, що і треба було довести.

Приклад 12.3

Нехай $\{\xi_n\}_{n=1}^{\infty}$ – послідовність випадкових величин і

$$\mathrm{P}\left\{\xi_n=\pm\ln n\right\}=\frac{1}{2n},\qquad \mathrm{P}\left\{\xi_n=0\right\}=1-\frac{1}{n}.$$

Чи має місце для даної послідовності закон великих чисел?

Розв'язок: $\mathbf{M}\xi_n=\frac{\ln n}{2n}-\frac{\ln n}{2n}+0\cdot\left(1-\frac{1}{n}\right)=0,$

$$\mathrm{P}\left\{\xi_n^2=\ln^2 n\right\}=\frac{1}{n},\qquad \mathrm{P}\left\{\xi_n^2=0\right\}=1-\frac{1}{n},$$

$$\mathbf{D}\xi_n=\mathbf{M}\xi_n^2-(\mathbf{M}\xi_n)^2=\mathbf{M}\xi_n^2-0=\mathbf{M}\xi_n^2=\frac{\ln^2 n}{n}+0=\frac{\ln^2 n}{n}.$$

Оскільки для будь-якого натурального n

$$\frac{\ln^2 n}{n}\le 1,$$

то всі дисперсії $\mathbf{D}\xi_n$ рівномірно обмежені, а тому за теоремою Чебишова 12.1 (стор. 162) маємо, що для вихідної послідовності випадкових величин має місце закон великих чисел.

Приклад 12.4

Розглянемо послідовність $\xi_1, \xi_2, \ldots, \xi_n, \ldots$, в якій ξ_n має розподіл:

$$\mathrm{P}\left\{\xi_n = 0\right\} = 1 - \frac{1}{n}, \quad \mathrm{P}\left\{\xi_n = n^7\right\} = \frac{1}{n}.$$

Довести, що послідовність $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ збігається за ймовірністю до нуля.

Доведення: Зафіксуємо довільне $\varepsilon > 0$. Для всіх n, починаючи з деякого n_0 такого, що $n_0^7 > \varepsilon$, має місце рівність $\mathrm{P}\left\{\xi_n \geq \varepsilon\right\} = \mathrm{P}\left\{\xi_n = n^7\right\} = \dfrac{1}{n}$. Тому

$$\mathrm{P}\left\{|\xi_n - 0| \geq \varepsilon\right\} = \mathrm{P}\left\{\xi_n \geq \varepsilon\right\} = \mathrm{P}\left\{\xi_n = n^7\right\} = \frac{1}{n} \to 0 \quad \text{при } n \to \infty,$$

що і треба було довести.

Приклад 12.5

Нехай випадкові величини $\xi_1, \xi_2, \ldots$ незалежні у сукупності і мають рівномірний розподіл на відрізку $[0, 1]$. Довести, що послідовність випадкових величин

$$\varphi_1 = \xi_1, \quad \varphi_2 = \max\left\{\xi_1, \xi_2\right\}, \quad \varphi_3 = \max\left\{\xi_1, \xi_2, \xi_3\right\}, \ldots$$

збігається за ймовірністю до одиниці.

Доведення: збіжність за ймовірністю до константи рівносильна слабкій збіжності, тому досить довести, що послідовність функцій розподілу $F_{\varphi_n}(x)$ слабко збігається до $F_1(x) = \begin{cases} 0, \ x < 1, \\ 1, \ x \geq 1 \end{cases}$.

Для довільного $x \leq 0$ $F_{\varphi_n}(x) = 0 \to F_1(x) = 0$ при $n \to \infty$. Для $0 < x < 1$ маємо $F_{\varphi_n}(x) = x^n \to F_1(x) = 0$ при $n \to \infty$. Нарешті для $x \geq 1$ $F_{\varphi_n}(x) = 1 \to F_1(x) = 1$ при $n \to \infty$.

Таким чином, φ_n слабко збігається до 1, що і треба було довести.

Приклад 12.6

Монета підкидається 10^4 раз. На основі нерівності Чебишова оцінити ймовірність того, що частота появ герба відрізняється від $1/2$ не менше, ніж на 0.01.

Розв'язок: Нехай $\xi_1, \xi_2, \ldots$ – незалежні випадкові величини, кожна з яких приймає два значення 0 і 1 з ймовірністю $1/2$. Величина ξ_n дорівнює 1, якщо при n-ому підкиданні випав герб, і дорівнює 0, якщо випала решка. Потрібно оцінити

$$\mathrm{P}\left\{\left|\frac{\mu_n}{n} - \frac{1}{2}\right| \geq 0.01\right\},$$

де $n = 10^4$, $\mu_n = \sum\limits_{i=1}^{n} \xi_i$ – число появ герба при n підкиданнях монети.

Оскільки $\mathbf{D}\xi_1 = 0.5 \cdot 0.5 = 0.25$, то шукана оцінка за нерівністю Чебишова (див. теорему 12.1 на стор. 162) набуває виду

$$\mathrm{P}\left\{\left|\frac{\mu_n}{n} - \frac{1}{2}\right| \geq 0.01\right\} \leq \frac{\mathbf{D}\xi_1}{n \cdot 0.01^2} = \frac{1}{4}.$$

Відзначимо, що ця оцінка дуже груба. Так, якщо застосувати інтегральну теорему Муавра-Лапласа (теорема 8.2 на стор. 72), то можна показати, що

$$\mathrm{P}\left\{\left|\frac{\mu_n}{n} - \frac{1}{2}\right| \geq 0.01\right\} \approx 2\left(1 - \Phi(2)\right) \approx 0.046.$$

Приклад 12.7

Правильна гральний кубик підкидається 500 разів. За допомогою нерівності Чебишова оцініть ймовірність того, що середнє арифметичне кількості очок, що випали, відхилиться від математичного сподівання за абсолютною величиною не більше, ніж на 0.2.

Розв'язок. Підставимо середнє арифметичне кількості очок, що випали, в нерівність Чебишова (див. теорему 12.1 на стор. 162):

$$\mathbf{P}\left\{|\bar{\xi} - \mathbf{M}\bar{\xi}| \leq \varepsilon\right\} \geq 1 - \frac{\mathbf{D}\bar{\xi}}{\varepsilon^2}.$$

Тут ξ_i – кількість очок, що випали при i-му підкиданні кубика, $\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i$ – середнє арифметичне, причому легко показати (покажіть!), що $\mathbf{M}\bar{\xi} = \mathbf{M}\xi_1$, $\mathbf{D}\bar{\xi} = \mathbf{D}\xi_1/n$.

Для кількості очок при одному підкиданні кубика маємо

$$\mathbf{M}\xi_1 = 3.5, \;\; \mathbf{M}\xi_1^2 = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6},$$

$$\mathbf{D}\xi_1 = \mathbf{M}\xi_1^2 - (\mathbf{M}\xi_1)^2 = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}, \;\; \mathbf{D}\bar{\xi} = \frac{35}{12\cdot 500} = \frac{7}{1200},$$

звідки маємо

$$\mathbf{P}\left\{|\bar{\xi} - 3.5| \le 0.2\right\} \ge 1 - \frac{7}{1200\cdot(0.2)^2} = 1 - \frac{25\cdot 7}{1200} = 1 - \frac{7}{48} \approx 0.8542.$$

Приклад 12.8

Нехай в магазині плата за всі покупки округляються до найближчого цілого числа. Яка сума накопичиться після 100 покупок, якщо вважати похибки заокруглення незалежними випадковими величинами з рівномірним розподілом на $[-0.5, 0.5)$?

Розв'язок. Спробуємо знайти розподіл і якісь ймовірності. Маємо 100 покупок, 100 випадкових величин $\xi_1, ..., \xi_{100}$ – похибки заокруглення у відповідних покупках, $\xi_i \sim U_{[-0.5,0.5)}$. $S_{100} = \xi_1 + ... + \xi_{100}$ – накопичена сума, $S_n = 100\cdot\overline{\xi_{100}}$.

Згідно ЦГТ

$$\overline{\xi_{100}} \sim N\left(\mu, \frac{\sigma^2}{100}\right),$$

де $\mu = \mathbf{M}\xi_1 = 0$, $\sigma^2 = \mathbf{D}\xi_1 = 1/12$, отже

$$S_n \sim N\left(100\cdot\mu, 100^2\frac{\sigma^2}{100}\right) = N\left(0, \frac{100}{12}\right).$$

Отримали апроксимацію розподілу для величини накопиченої суми S_n.

Розподіл потрібен для того, щоб визначати практично значущі величини, наприклад, ймовірність того, що на 100 клієнтах сума заокруглень буде більше, ніж 10 грошових одиниць.

$$\mathbf{P}\{S_n > 10\} = 1 - F_{S_n}(10) \approx 0.0003.$$

Це дуже мала величина, тому така стратегія заокруглення не загрожує магазину великими збитками (про те, як знайти функцію розподілу нормального закону $F_{S_n}(10)$, див. приклади розділу 10).

Задача 12.1. Нехай $\{\xi_n\}_{n=1}^{\infty}$ - послідовність випадкових величин, така, що ξ_n може залежати тільки від ξ_{n-1} і ξ_{n+1}, і не залежить від інших ξ_k. Показати, що для цієї послідовності виконується закон великих чисел, якщо:

а) $\mathbf{D}\xi_n \le C$ (використати закон великих чисел у формі Чебишова);

б) $\mathbf{D}\xi_n = o(n)$ (використати закон великих чисел у формі Маркова).

в) Навести приклад таких ξ_1, ξ_2 та ξ_3, щоб ξ_1 і ξ_2 були залежними, ξ_2 і ξ_3 були залежними, а ξ_1 і ξ_3 були незалежними.

Задача 12.2. Використовуючи нерівність Чебишова, оцінити ймовірність того, що при 1000 підкиданнях монети число появ герба буде у межах між 450 і 550.

Задача 12.3. Скільки треба підкидати правильну монету, щоб з ймовірністю 0.99 відносна частота випадіння герба відрізнялась від 1/2 не більше, ніж на 0.01?

Задача 12.4. У таблиці випадкових чисел кожна цифра з'являється незалежно від інших з ймовірністю 1/10. Скільки треба набрати таких випадкових чисел, щоб з ймовірністю 0.999 серед них з'явилося не менше 100 нулів?

Задача 12.5. За допомогою варіанту нерівності Чебишова оцінити ймовірність того, що в результаті підкидання гральної кістки 400 разів відносна частота появи шістки на верхній грані відхилиться від ймовірності цієї події за абсолютною величиною не більше, ніж на 0.03.

Задача 12.6. Знайти $\mathbf{P}\left\{|\xi - \mathbf{M}\xi| > 3\sqrt{\mathbf{D}\xi}\right\}$, якщо випадкова величина ξ має

1) рівномірний розподіл на відрізку $[a, b]$;

2) показниковий розподіл з параметром λ;

3) розподіл Бернуллі з параметром 1/2.

Задача 12.7. У певній місцевості середня річна кількість сонячних днів дорівнює 100. Оцінити ймовірність того, що упродовж року в цій місцевості буде не більше ніж 125 сонячних днів.

Задача 12.8. Скільки потрібно провести незалежних випробувань, щоб ймовірність нерівності $\left|\frac{S_n}{n} - p\right| \le 0.06$ не перевищувала б 0.78, якщо ймовірність появи події в окремому випробуванні $p = 0.7$?

Задача 12.9. В деякій місцевості середня швидкість вітру на даній висоті дорівнює 20 км/год, а середнє квадратичне відхилення – 4 км/год. Оцінити відхилення швидкості вітру на цій висоті від середнього значення з ймовірністю не меншою, ніж 0.9.

Задача 12.10. Нехай $\mathbf{M}\xi = 12$, $\mathbf{P}\{\xi \ge 14\} = 0.12$, $\mathbf{P}\{\xi \le 10\} = 0.18$. Доведіть, що дисперсія ξ не менше за 1.2.

Задача 12.11. Словник має 1500 сторінок. Ймовірність друкарської помилки на одній сторінці дорівнює 0,001. Знайти ймовірність того, що в словнику: а) буде точно три помилки; б) не буде жодної помилки; в) буде хоча б одна помилка.

Задача 12.12. Імовірність влучити в мішень при одному пострілі дорівнює 0,8. Знайти ймовірність того, що із 100 пострілів число влучень буде: а) точно 85; б) від 75 до 85.

Задача 12.13. Знайти наступні границі:

а) $\lim\limits_{n\to\infty} \int\limits_0^{0.99999n} \frac{y^{n-1}e^{-y}}{(n-1)!}dy$, б) $\lim\limits_{n\to\infty} \int\limits_0^{n} \frac{y^{n-1}e^{-y}}{(n-1)!}dy$, в) $\lim\limits_{n\to\infty} \int\limits_0^{1.000001n} \frac{y^{n-1}e^{-y}}{(n-1)!}dy$.

Задача 12.14. Довести, що для послідовності незалежних випадкових величин $\{\xi_n\}_{n=1}^{\infty}$ такої, що

$$\mathrm{P}(\xi_n = n^{\alpha}) = \mathrm{P}(\xi_n = -n^{\alpha}) = 1/2,$$

виконується закон великих чисел тоді і тільки тоді, коли $\alpha < 1/2$.

Задача 12.15. Розглянемо послідовність незалежних випадкових величин $\{\xi_n\}_{n=1}^{\infty}$, яка визначається рівністю

$$\xi_n = \begin{cases} \pm 1 & \text{з ймовірністю } (1-2^{-n})/2, \\ \pm 2^{n/2} & \text{з ймовірністю } 2^{-n-1} \end{cases}$$

Довести, що для цієї послідовності виконується закон великих чисел.

Задача 12.16. Послідовність незалежних, однаково розподілених випадкових величин $\{\xi_n\}$ визначена рівностями:

а) $\mathrm{P}(\xi_n = 2^{k-\ln k - 2\ln\ln k}) = 1/2^{k-1}$, $k = 2.3, ...$;

б) $\mathrm{P}(\xi_n = k) = \frac{c}{k^2 \ln^2 k}$, $k = 2.3, ...$, $c^{-1} = \sum\limits_{k=2}^{\infty} \frac{1}{k^2 \ln^2 k}$.

Перевірити, чи виконується для цих послідовностей закон великих чисел.

Задача 12.17. Розглянемо $\{\xi_n\}_{n=1}^{\infty}$ - послідовність невід'ємних випадкових величин, які приймають цілі значення. Через $m_k(n)$ будемо позначати k-ий **факторіальний момент** випадкової величини ξ_n:

$$m_k(n) = \mathbf{M}\xi_n\,(\xi_n - 1)\cdot \ldots \cdot (\xi_n - k + 1)\,.$$

Нехай для кожного $k = 1, 2, \ldots$

$$\lim_{n\to\infty} m_k(n) = \lambda^k, \quad \text{де} \quad 0 < \lambda < \infty.$$

Довести, що тоді $\forall m \in \{0.1, 2, \ldots\}$

$$\lim_{n\to\infty} \mathrm{P}\left(\xi_n = m\right) = \frac{\lambda^m}{m!}e^{-\lambda}.$$

Задача 12.18. Випадкові величини ξ_n, η_n незалежні і мають пуассонівський розподіл з $\mathbf{M}\xi_n = \mathbf{M}\eta_n = \lambda$. Знайти

$$\lim_{n\to\infty} \mathrm{P}\left\{\frac{\xi_n - \eta_n}{\sqrt{n}} \le x\right\}.$$

Задача 12.19. Якщо випадкові величини $\xi_1, \xi_2, \ldots, \xi_n$ незалежні і мають стандартний нормальний розподіл, то розподіл випадкової величини

$$\chi_n^2 = \xi_1^2 + ... + \xi_n^2$$

називають **розподілом χ^2 (хі-квадрат) з n степенями свободи**.

а) Довести, що $\forall \varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\left\{\left|\frac{\chi_n^2}{n} - 1\right| > \varepsilon\right\} = 0.$$

б) Знайти

$$\lim_{n\to\infty} \mathrm{P}\left\{\frac{\chi_n^2 - \mathbf{M}\chi_n^2}{\sqrt{\mathbf{D}\chi_n^2}} \le x\right\}, \quad -\infty < x < \infty.$$

Задача 12.20. Якщо випадкові величини $\xi_1, \xi_2, \ldots, \xi_n$ незалежні і мають стандартний нормальний розподіл, то розподіл випадкової величини

$$\tau_n = \frac{\xi_0}{\sqrt{(\xi_1^2 + \ldots + \xi_n^2)/n}}$$

називають **розподілом Стьюдента з n степенями свободи**. Знайти

$$\lim_{n\to\infty} \mathrm{P}(\tau_n \le x), \quad -\infty < x < \infty.$$

Задача 12.21. Випадкові величини $\xi_1, \xi_2, \ldots, \xi_n$ незалежні і однаково розподілені, причому існують $\mathbf{M}\xi_n = a$, $\mathbf{D}\xi_n = \sigma^2 < \infty$. Покладемо $\eta_n = \xi_n + \xi_{n+1} + \xi_{n+2}$, $n = 1, 2, \ldots$ Знайти

$$\lim_{n\to\infty} \mathrm{P}\left\{\frac{\eta_1 + \ldots + \eta_n - 3na}{\sigma\sqrt{3n}} \le x\right\}.$$

Задача 12.22. Нехай випадкові величини $\xi_1, \xi_2, \ldots, \xi_n \ldots$ незалежні і однаково розподілені, $\mathbf{M}\xi_n = 0$, $\mathbf{M}\xi_n^2 = 1$ і нехай λ_n, $n = 1, 2, \ldots$ числова послідовність, яка задовольняє умові

$$\frac{\max\limits_{1\le k\le n} \lambda_k^2}{\sum\limits_{k=1}^{n} \lambda_k^2} \xrightarrow[n\to\infty]{} 0.$$

Довести, що для послідовності $\lambda_1\xi_1, \lambda_2\xi_2, \ldots, \lambda_n\xi_n, \ldots$ справедливе центральне граничне твердження.

Задача 12.23. Випадкові величини $\xi_1, \xi_2, \ldots, \xi_n \ldots$ незалежні і мають наступний розподіл

$$\xi_n = \begin{cases} \pm n^\alpha & \text{з ймовірністю } \dfrac{1}{2n^\beta}, \\ 0 & \text{з ймовірністю } 1 - \dfrac{1}{n^\beta} \end{cases}$$

При яких α і β виконуються умови теореми Ляпунова?

Задача 12.24. Нехай p_i є ймовірність появи події A у i- ому випробуванні, а μ_n – число появ події A у n незалежних випробуваннях. Довести, що для

$$\left(\mu_n - \sum_{k=1}^{n} p_k\right) \Big/ \sqrt{\sum_{i=1}^{n} p_i q_i}$$

справедливе центральне граничне твердження тоді і тільки тоді, коли $\sum\limits_{i=1}^{\infty} p_i q_i = \infty$.

Задача 12.25. Для послідовності випадкових величин $\xi_1, \xi_2, \ldots$ через σ_i^2 позначимо дисперсію $\mathbf{D}\xi_i$, а через r_{ij} – коефіцієнт кореляції між ξ_i і ξ_j. Довести, що, якщо $\sigma_i^2 \le C$ і $r_{ij} \to 0$ при $|i-j| \to \infty$, то для цієї послідовності виконується слабкий закон великих чисел.

Задача 12.26.* Нехай $\{\xi_n\}_{n=1}^{\infty}$ - послідовність незалежних випадкових величин, яка визначається рівністю

$$\xi_n = \begin{cases} \pm 1 & \text{з ймовірністю } \dfrac{n-1}{2n}, \\ \pm\sqrt{n} & \text{з ймовірністю } \dfrac{1}{2n} \end{cases}$$

Позначимо через $S_n = \xi_1 + \ldots + \xi_n$. Довести, що ні при яких нормуючих сталих α_n, розподіл S_n/α_n не збігається слабко до нормального розподілу.

Задача 12.27.* Випадкові величини $\xi_1, \xi_2, \ldots, \xi_{100}$ незалежні і рівномірно розподілені у відрізку [0, 1]. Знайти ймовірність того, що

$$\prod_{k=1}^{100} \xi_k \le 10/2^{100}.$$

Задача 12.28.* Нехай маємо суму 24 незалежних випадкових величин, кожна з яких рівномірно розподілена на інтервалі $[0, 1]$. Написати наближений вираз для щільності розподілу суми цих випадкових величин і обчислити ймовірність того, що їхня сума лежить в межах від 6 до 8.

Задача 12.29. Розглянемо випадок короткострокового страхування життя, коли страхова компанія сплачує 25 000 гривень у випадку смерті застрахованого упродовж року, і нічого не платить у протилежному випадку. Припустимо, що на цих умовах застраховано 3 000 клієнтів в одному віці. З таблиць смертності відомо, що ймовірність смерті упродовж року дорівнює 0.002984. Використовуючи гауссівське або пуассонівське наближення для сумарного позову, треба знайти страхову премію, яка забезпечить ймовірність нерозорення страхової компанії 0.95. Чому дорівнювала б ймовірність розорення, якщо б страховий поліс коштував нетто-премію?

Задача 12.30. В страховій компанії застраховано $N_1 = 3000$ клієнтів у віці 38 років і $N_2 = 1000$ клієнтів у віці 18 років. З таблиць смертності відомо, що для клієнтів першої групи ймовірність смертності упродовж року дорівнює 0.003, а для клієнтів другої групи – 0.001. Компанія сплачує спадкоємцям 10 000 гривень у випадку смерті застрахованого упродовж року, і не сплачує нічого, якщо клієнт доживає до кінця року.

Треба підрахувати страхову премію для клієнтів кожної групи, яка б гарантувала ймовірність нерозорення страхової компанії 0.98. Який середній прибуток страхової компанії у цьому випадку? При підрахунках використати пуассонівське наближення сумарного позову.

Задача 12.31. Припустимо, що страхова компанія уклала 10000 угод страхування життя строком на один рік на наступних умовах. У випадку смерті застрахованого упродовж року від нещасного випадку компанія сплачує 10000 гривень. У випадку смерті від природних причин компанія сплачує 2500 гривень. Компанія нічого не платить, якщо клієнт доживає до кінця року. Ймовірність смерті від нещасного випадку дорівнює $5 \cdot 10^{-4}$, а від природних причин – $3 \cdot 10^{-3}$.

Знайти величину страхової премії, яка забезпечить ймовірність нерозорення страхової компанії 0.96. При підрахунках використати гауссівське наближення сумарного позову.

Задача 12.32. (Продовження попередньої задачі) Нехай виконуються всі умови попередньої задачі, крім однієї. Будемо вважати, що ймовірність смерті від природних причин залежить від віку. У зв'язку з цим 10 000 застрахованих розбивається на дві вікові групи, які містять $N_1 = 4000$ і $N_2 = 6000$. Ймовірність смерті від природних причин в кожній групі дорівнює $q_1 = 0.004$ і $q_2 = 0.002$ відповідно. Треба підрахувати величину страхової премії для кожної вікової групи, яка забезпечить ймовірність виконання компанією своїх зобов'язань на рівні 95 %. При підрахунках використати гауссівське наближення сумарного позову, а страхову надбавку за ризик розділити між віковими групами пропорційно дисперсіям.

13 ЛАНЦЮГИ МАРКОВА З ДИСКРЕТНИМ ЧАСОМ

Нехай задана послідовність випадкових величин $\xi_0, \xi_1, \xi_2, \ldots$, які визначені на одному й тому ж ймовірнісному просторі $(\Omega,\ \Re,\ \mathrm{P})$ і які набувають значень з множини $E = \{x_1, x_2, \ldots\}$.

Говорять, що послідовність $\xi_0, \xi_1, \xi_2, \ldots$ утворює **ланцюг Маркова**, якщо для будь-якого натурального n та довільних $i_0,\ i_1,\ \ldots,\ i_n$ таких, що

$$\mathrm{P}\{\xi_{n-1} = x_{i_{n-1}}, \xi_{n-2} = x_{i_{n-2}},\ \ldots, \xi_1 = x_{i_1}, \xi_0 = x_{i_0}\} > 0$$

виконується рівність

$$\mathrm{P}\left\{\xi_n = x_{i_n} | \xi_{n-1} = x_{i_{n-1}}, \ldots, \xi_0 = x_{i_0}\right\} = \mathrm{P}\left\{\xi_n = x_{i_n} | \xi_{n-1} = x_{i_{n-1}}\right\} \tag{13.1}$$

Властивість (13.1) називається **властивістю Маркова**.

Якщо ймовірність

$$p_{ij}(n) = \mathrm{P}\left(\xi_n = x_j | \xi_{n-1} = x_i\right)$$

не залежить від n, то ланцюг Маркова називається **однорідним**. Матриця $P = \|p_{ij}\|$, де $p_{ij} = \mathrm{P}\left(\xi_1 = x_j | \xi_0 = x_i\right)$, називається **матрицею перехідних ймовірностей за один крок** або просто **перехідною матрицею** ланцюга Маркова ξ_n.

Ймовірності $p_i = \mathrm{P}\left(\xi_0 = x_i\right)$ називаються **початковим розподілом** ланцюга Маркова ξ_n. Оскільки

$$\mathrm{P}\left(\xi_0 = x_{i_0},\ \xi_1 = x_{i_1}, \ldots,\ \xi_{n-1} = x_{i_{n-1}},\ \xi_n = x_{i_n}\right) = p_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdot \ldots \cdot p_{i_{n-1} i_n},$$

то початковий розподіл $\{p_i\}_{i=0}^{\infty}$ і матриця перехідних імовірностей $P = \|p_{ij}\|$ визначають скінченновимірні розподіли ланцюга Маркова.

Ймовірність

$$p_{ij}^{(n)} = \mathrm{P}\left(\xi_n = x_j | \xi_0 = x_i\right)$$

називають **ймовірністю переходу за n кроків** і якщо $P^{(n)} = \left\|p_{ij}^{(n)}\right\|$, то $P^{(n)} = P^n$. Для довільних натуральних n, m рівність

$$P^{(n+m)} = P^n P^m$$

називається **рівнянням Чепмена-Колмогорова**, яке в розгорнутому вигляді можна подати так:

$$p_{ij}^{(n+m)} = \sum_k p_{ik}^{(n)} p_{kj}^{(m)}.$$

Стан x_i називається **неістотним**, якщо знайдеться такий стан x_{j_o}, що $p_{ij_0}^{(n_0)} > 0$ для деякого n_0, але $p_{j_0 i}^{(n)} = 0$ для всіх $n > 1$. У протилежному випадку стан будемо називати **істотним**.

Стан x_j **досягається** зі стану x_i якщо $p_{ij}^{(n)} > 0$ для деякого $n > 1$. В цьому випадку будемо писати $x_i \to x_j$.

Якщо $x_i \to x_j$ і $x_j \to x_i$, то говорять, що стани x_i,x_j **сполучаються**. В цьому випадку будемо писати $x_i \leftrightarrow x_j$. За визначенням вважаємо, що $x_i \leftrightarrow x_i$.

Теорема 13.1 Множина станів E ланцюга Маркова ξ_n може бути подана у вигляді

$$E = E_0 \cup E_1 \cup E_2 \cup \ldots$$

де E_0 клас усіх неістотних станів, а E_i, $i \geq 1$ класи істотних станів, що сполучаються, причому $E_i \cap E_j = \varnothing$ для $i \neq j$.

Ланцюг Маркова, всі стани якого сполучаються, називається **незвідним**.

Нехай

$$F_i = \sum_{n=1}^{\infty} \mathrm{P}\left(\xi_1 \neq x_i,\ \xi_2 \neq x_i, ..., \xi_{n-1} \neq x_i,\ \xi_n = x_i \middle|\ \xi_0 = x_i\right).$$

Стан x_i називають **рекурентним**, якщо $F_i = 1$, і **нерекурентним**, якщо $F_i < 1$.

Якщо $\xi_0 = x_i$, то випадкову величину $\tau_i = inf\{n > 0 : \xi_n = x_i\}$ називають **часом першого повернення в стан** x_i.

Стан x_i називається **рекурентним додатнім**, якщо він є рекурентним і $\mathbf{M}\tau_i < \infty$. Якщо $\mathbf{M}\tau_i = \infty$, то такий стан називається **рекурентним нульовим**.

Теорема 13.2 Стан x_i є рекурентним тоді і лише тоді, коли

$$P_i = \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty.$$

Якщо стан x_i не є рекурентним, то

$$F_i = \frac{P_i}{1 + P_i}.$$

Стан x_i називають **періодичним** з періодом d> 1, якщо НСД$\{n : p_{ii}^{(n)} > 0\} = d$.

Теорема 13.3 (про солідарність) В класі станів, що сполучаються:
а) якщо один стан є рекурентним, тоді всі стани будуть рекурентними;
б) якщо один стан є рекурентно нульовим, тоді всі стани будуть рекурентно нульовими;
в) якщо один стан є рекурентно додатнім, тоді всі стани будуть рекурентно додатними;
г) якщо один стан є періодичним з періодом d, тоді всі стани будуть періодичні з тим самим періодом.

Таким чином, якщо ланцюг є незвідним і хоча б один із його станів має період, то всі його стани мають цей же період, який називають періодом ланцюга.

Теорема 13.4 Простір станів E незвідного періодичного з періодом d ланцюга Маркова ξ_n може бути поданий у вигляді

$$E = E_0 \cup E_1 \cup E_2 \cup E_{d-1} \tag{13.2}$$

де $E_i \cap E_j = \varnothing$ для $i \neq j$, до того ж за один крок ланцюг переходить з класу E_i до класу E_{i+1}, а з класу E_{d-1} до класу E_0.

Класи E_i в (13.2) називаються **циклічними підкласами ланцюга ξ_n.**

Розподіл $\{\pi_i\}_{i=1}^{\infty}$ називають **стаціонарним розподілом** ланцюга ξ_n,$n \geq 0$, якщо

$$\pi_i = \sum_k \pi_k p_{ki}$$

або в матричній формі

$$\pi = \pi P,$$

де π позначає вектор-рядок з координатами π_i. Якщо для марковсього ланцюга існують границі

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j > 0,$$

то $\{\pi_j\}_{j=1}^{\infty}$ **називають ергодичним розподілом.**

Наступні три теореми дають умови для існування ергодичного розподілу для скінченних ланцюгів Маркова.

Теорема 13.5 Нехай ланцюг маркова ξ_n, $n \geq 0$, скінченний. Умова:

$$\min_{i,j} p_{ij}^{(n)} > 0 \text{ для деякого } n$$

є необхідною і достатньою для існування ергодичного розподілу $\pi = (\pi_1, \ldots, \pi_N)$, причому цей ергодичний розподіл буде і єдиним стаціонарним розподілом.

Теорема 13.6 Якщо ланцюг Маркова ξ_n, $n \geq 0$, незвідний, скінченний та неперіодичний, то для нього існує ергодичний розподіл $\pi = (\pi_1, \ldots, \pi_N)$, який буде і єдиним стаціонарним розподілом, і

$$\forall j = \overline{1, N} \quad \pi_j = \frac{1}{\mathbf{M}\tau_j}.$$

Аналог попередньої теореми для періодичних ланцюгів виглядає так:

Теорема 13.7 Якщо ланцюг Маркова ξ_n, $n \geq 0$, незвідний, скінченний та періодичний з періодом d, стани x, y належать циклічним підкласам E_i та E_j відповідно ($x \in E_i$, $y \in E_j$), то

$$\lim_{n\to\infty} p_{xy}^{(nd+m)} = \frac{d}{\mathbf{M}\tau_y}$$

при умові $m = j - i \pmod d$ та

$$\lim_{n\to\infty} p_{xy}^{(nd+m)} = 0$$

при умові $m \neq j - i \pmod d$.

Наступні дві теореми дають умови для існування ергодичних розподілів для нескінченних ланцюгів Маркова.

Теорема 13.8 Для того, щоб ланцюг Маркова ξ_n, $n \geq 0$, був ергодичним, необхідно і достатньо, щоб цей ланцюг був незвідним, неперіодичним, та щоб існував стан i_0 такий, що $\mathbf{M}\tau_{i_0} < \infty$. При цьому єдиний стаціонарний розподіл буде співпадати з ергодичним.

Теорема 13.9 Якщо ланцюг Маркова ξ_n, $n \geq 0$, незвідний та неперіодичний і такий, що для деякого j_0 та $\varepsilon > 0$ виконується умова

$$\inf_i p_{ij_0} \geq \varepsilon,$$

тоді існує ергодичний розподіл $\pi_1, \ldots, \pi_N, \ldots$, який буде єдиним стаціонарним розподілом, і

$$\pi_j = \frac{1}{\mathbf{M}\tau_j}.$$

Приклад 13.1

Розглянемо ланцюг Маркова з двома станами 0 та 1 і матрицею переходів

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}.$$

Знайти границі ймовірності ${p_{i0}}^{(n)}$ та ${p_{i1}}^{(n)}$.

Розв'язок 1: Після підрахунку матриці

$$P^2 = \begin{pmatrix} {p_{00}}^2 + p_{01}p_{10} & p_{01}(p_{00} + p_{11}) \\ {p_{00}}^2 + p_{01}p_{10} & p_{01}(p_{00} + p_{11}) \end{pmatrix}$$

за індукцією можна показати, що матриця P^n має вигляд

$$P^n = \frac{1}{2 - p_{00} - p_{11}} \begin{pmatrix} 1 - p_{11} & 1 - p_{00} \\ 1 - p_{11} & 1 - p_{00} \end{pmatrix} + \frac{(p_{00} + p_{11} - 1)^n}{2 - p_{00} - p_{11}} \begin{pmatrix} 1 - p_{00} & -(1 - p_{00}) \\ -(1 - p_{11}) & 1 - p_{11} \end{pmatrix}.$$

Якщо елементи матриці P є такими, що

$$|p_{00} + p_{11} - 1| < 1, \tag{13.3}$$

то

$$P^n \longrightarrow \frac{1}{2 - p_{00} - p_{11}} \begin{pmatrix} 1 - p_{11} & 1 - p_{00} \\ 1 - p_{11} & 1 - p_{00} \end{pmatrix}.$$

Звідси отримаємо:

$$\lim_{n\to\infty} {p_{i0}}^{(n)} = \frac{1 - p_{11}}{2 - p_{00} - p_{11}}, \qquad \lim_{n\to\infty} {p_{i1}}^{(n)} = \frac{1 - p_{00}}{2 - p_{00} - p_{11}}.$$

Розв'язок 2: Якщо посилити умову (13.3) і припустити, що всі $p_{ij} > 0$, то ланцюг Маркова, що розглядається, буде незвідним та неперіодичним. Тоді за теоремою 6 існують границі $\lim\limits_{n\to\infty} {p_{ij}}^{(n)} = \pi_j > 0$, $i, j = 0, 1$, які є розв'язком системи рівнянь:

$$\begin{cases} \pi_0 = \pi_0 p_{00} + \pi_1 p_{10} \\ \pi_1 = \pi_0 p_{01} + \pi_1 p_{11} \end{cases}$$

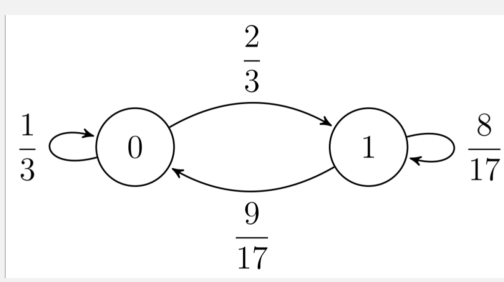
з умовою нормування $\pi_0 + \pi_1 = 1$. Розв'язуючи цю систему, знаходимо:

$$\pi_0 = \frac{1 - p_{11}}{2 - p_{00} - p_{11}}\ , \quad \pi_1 = \frac{1 - p_{00}}{2 - p_{00} - p_{11}}.$$

Приклад 13.2

Припустимо, що після зростання (зелена добова "свічка") ціна на Bitcoin наступного дня з ймовірністю $\frac{1}{3}$ зростатиме (або з ймовірністю спадатиме $\frac{2}{3}$), а після спадання ціна наступного дня з ймовірністю $\frac{9}{17}$ зростатиме (або з ймовірністю спадатиме $\frac{8}{17}$). Яка ймовірність того, що через 3 дні після зростання ціна знову буде зростати?

Розв'язок: Виходячи з умови задачі йдеться про 2 можливі стани – зростання (позначимо цей стан як 0) та спадання (позначимо цей стан як 1). Переходи між станами відповідно до умови можна зобразити у вигляді такого графа



та записати відповідну матрицю переходів за 1 день

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{9}{17} & \frac{8}{17} \end{pmatrix}.$$

Обчислимо матрицю переходів за 3 дні, підносячи матрицю P до степеня 3:

$$P^{(3)} = P^3 = \begin{pmatrix} \frac{3421}{7803} & \frac{4382}{7803} \\ \frac{2191}{4913} & \frac{2722}{4913} \end{pmatrix}.$$

Шуканою ймовірністю тоді буде число $P_{00}^{(3)} = \dfrac{3421}{7803} \approx 0.438$.

Відповідь: $\dfrac{3421}{7803} \approx 0.438$.

Приклад 13.3

Припустимо, що поведінка двох гравців в шахи А та В в турнірній грі можна описати таким чином. Шахист А кожну партію незалежно від результатів попередніх партій виграє з ймовірністю p, програє з ймовірністю q і грає внічию з ймовірністю $r = 1 - p - q$.

Шахист В менш врівноважений, він виграє партію з ймовірностями $p + \varepsilon$, p та $p - \varepsilon$ відповідно, якщо попередня партія була їм виграна, зіграна внічию та програна. Ймовірність програшу в цих випадках дорівнює $q - \varepsilon$, q, $q + \varepsilon$ відповідно. Хто з гравців А та В в тривалому турнірі набере більше очок?

Розв'язок: Щоб відповісти на це питання, треба обчислити для кожного шахиста стаціонарні ймовірності q_1, q_2, q_3 попадання в стани E_1, E_2, E_3, які відповідають виграшу, нічиїй та програшу.

Для гравця А ЛМ з станами E_1, E_2, E_3, які описують хід його гри, буде мати матрицю перехідних ймовірностей

$$P_A = \begin{pmatrix} p & r & q \\ p & r & q \\ p & r & q \end{pmatrix}.$$

Для гравця B

$$P_B = \begin{pmatrix} p+\varepsilon & r & q-\varepsilon \\ p & r & q \\ p-\varepsilon & r & q+\varepsilon \end{pmatrix}.$$

Рівняння для стаціонарних ймовірностей в цьому випадку матимуть вид

$$\begin{aligned} q_1(p+\varepsilon) + q_2 p + q_3(p-\varepsilon) = q_1, \\ q_1 r + q_2 r + q_3 r = q_2, \\ q_1 + q_2 + q_3 = 1. \end{aligned}$$

Розв'язуючи цю систему, знаходимо

$$q_2 - r = 0, \quad q_1 - p = \varepsilon \frac{p-q}{1-2\varepsilon}.$$

Таким чином, долі нічиїх при великій кількості партій у гравців A та B співпадатимуть, доля виграшів у гравця B буде більша, ніж у A, якщо $\varepsilon > 0$, $p > q$, або $\varepsilon < 0$, $p < q$. Якщо $p = q$, то стаціонарні розподіли для результатів ігор у A та B однакові.

Задача 13.1. Гральний кубик рівноймовірно перекладається на будь-яку з чотирьох сусідніх граней. Нехай X_n – число очок на верхній грані після n-го перекладання. Треба знайти матрицю P ймовірностей переходів за один крок для ланцюга Маркова X_n.

Задача 13.2. Розглянемо ланцюг Маркова X_n, пов'язаний зі схемою випробувань Бернуллі наступним чином: X_n знаходиться у стані 1, якщо $(n-1)$-ше та n-те випробування мали результатом “успіх”; аналогічно стани 2, 3, 4 відповідають парам результатів “успіх - невдача“, “невдача - успіх“, “невдача - невдача“. Знайти матрицю переходу за один крок для ланцюга X_n і стаціонарні ймовірності, якщо ймовірність “успіху“ дорівнює p.

Задача 13.3. Урна містить 5 куль: білі і чорні. На кожному кроці з урни навмання витягають одну кулю і замість неї повертають кулю, але іншого кольору: замість білої чорну і навпаки. Нехай X_n число білих куль в урні після n-го кроку. Знайти матрицю P ймовірностей переходів за один крок для ланцюга X_n та стаціонарні ймовірності.

Задача 13.4. Муха повзає по ребрам тетраедра, обираючи в кожній вершині напрямок свого подальшого руху навмання та проповзаючи ребро за одиницю часу. Описати цей процес за допомогою ланцюга Маркова, зобразити граф переходів та навести класифікацію його станів. Знайти ймовірності переходів за 2 кроки, а також розподіл ЛМ після 1, 2, 3 кроків, якщо початкове положення мухи а) достовірно відоме; б) рівномірно розподілене на множині вершин. Як зміниться ситуація, якщо в одній з вершин сидить павук?

Задача 13.5. Через фіксовані проміжки часу проводиться контроль технічного стану приладу, який може знаходитися в одному з трьох станів: E_0 – працює, E_1 – не працює і чекає ремонту, E_2 – в ремонті. Нехай X_n – номер стану приладу при n -й перевірці, є однорідним ЛМ з перехідною матрицею

$$P = \begin{pmatrix} 0,8 & 0,1 & p_{02} \\ 0,2 & p_{11} & 0,6 \\ p_{20} & 0,1 & 0,3 \end{pmatrix}.$$

Знайти невідомі елементи матриці P та обчислити $p(2)$ – розподіл по станах в момент другої перевірки, якщо в початковий момент часу прилад працював. В якому стані прилад перебуває найчастіше?

Задача 13.6. Нехай $\xi_1, \xi_2, \ldots$ – цифрова послідовність, у якій цифри з'являються рівноймовірно і незалежно одна від одної. Лічильник X_n у момент n показує, скільки різних цифр зустрілося серед перших n членів послідовності $\xi_1, \xi_2, \ldots$ Довести, що показання лічильника X_n утворюють ланцюг Маркова. Знайти матрицю ймовірностей переходу за один крок цього ланцюга. Вказати істотні і неістотні стани.

Задача 13.7. Числа обираються навмання з таблиці, яка містить усі цілі числа від 1 до m. Лічильник у момент часу n знаходиться у стані $j \in \{1, 2, \ldots, n\}$, якщо найбільше з n обраних чисел дорівнює j. Знайти матрицю переходу за один крок ланцюга Маркова, який зв'язаний з показаннями лічильника. В явному вигляді виписати матрицю ймовірностей переходів за один крок для випадку $m = 5$.

Задача 13.8. За виглядом матриці перехідних ймовірностей за один крок ланцюга Маркова вказати класи істотних і неістотних станів:

а) $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$ б) $\begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ в) $\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

г) $\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}$ д) $\begin{pmatrix} 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ е) $\begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$

Задача 13.9. З'ясувати, чи буде періодичним ланцюг Маркова з матрицею ймовірностей переходів за один крок:

$$\begin{pmatrix} 1/3 & 0 & 0 & 1/3 & 1/3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Задача 13.10. Знайти період ланцюга Маркова з наступною матрицею ймовірностей переходів за один крок:

а) $\begin{pmatrix} 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, б) $\begin{pmatrix} 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \\ 1/3 & 1/3 & 0 & 1/3 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \end{pmatrix}$, в) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \end{pmatrix}$.

Задача 13.11. Матриця ймовірностей переходу за один крок ланцюга Маркова має вигляд

$$P = \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{pmatrix}.$$

Розподіл по станах у момент часу $n = 0$ визначається вектором $p_0 = (0.7,\ 0.2,\ 0.1)$. Знайти:
а) розподіл по станах у момент часу $n = 2$;
б) ймовірність того, що у моменти $n = 0,\ 1,\ 2,\ 3$ станами ланцюга будуть відповідно $1, 3, 3, 2$;
в) стаціонарний розподіл.

Задача 13.12. Подати через відповідні перехідні ймовірності та початковий розподіл наступні умовні ймовірності:
а) $\mathrm{P}(\xi_0 = i | \xi_n = j)$; б) $\mathrm{P}(\xi_r = k | \xi_0 = i, \xi_n = j)$, $0 < r < n$.

Задача 13.13. Нехай для ланцюга Маркова ξ_n, $n \geq 0$, матриця ймовірностей переходу має вигляд

$$\begin{pmatrix} 3/7 & 3/7 & 1/7 \\ 1/11 & 4/11 & 6/11 \\ 1/11 & 4/11 & 6/11 \end{pmatrix}.$$

Покладемо

$$\eta_n = \begin{cases} 1, & \text{якщо } \xi_n = 1, \\ 2, & \text{якщо } \xi_n \neq 1. \end{cases}$$

Показати, що послідовність η_n, $n \geq 0$, також є ланцюгом Маркова. Знайти його матрицю ймовірностей переходу за один крок.

Задача 13.14. Послідовність $\{\xi_0,\ \xi_1,\ \ldots\}$ є ланцюгом Маркова з множиною станів $E = \{1,\ 2,\ 3\}$ та матрицею ймовірностей переходів за один крок

$$P = \begin{pmatrix} 0 & \alpha & 1-\alpha \\ \alpha & 0 & 1-\alpha \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, \qquad 0 < \alpha < 1.$$

Будується нова послідовність випадкових величин

$$\eta_n = \begin{cases} 1, & \text{якщо } \xi_n = 1 \text{ чи } \xi_n = 2, \\ 2, & \text{якщо } \xi_n = 3. \end{cases}$$

Довести, що послідовність $\{\eta_0,\ \eta_1, \ldots\}$ є ланцюгом Маркова та знайти матрицю ймовірностей переходів за один крок.

Задача 13.15. Нехай η_n, $n \geq 0$ – ланцюг Маркова з множиною станів 1, 2,3 і матрицею перехідних ймовірностей

$$P = \begin{pmatrix} 0 & 1-\alpha & \alpha \\ 1-\beta & 0 & \beta \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

Визначимо послідовність ξ_n, $n \geq 0$, наступним чином:

$$\xi_n = \begin{cases} 1, & \text{якщо } \eta_n = 1 \text{ чи } \xi_n = 2, \\ 2, & \text{якщо } \eta_n = 3. \end{cases}$$

При якій умові послідовність ξ_n, $n \geq 0$, також є ланцюгом Маркова?

Задача 13.16.* Розглянемо ланцюг Маркова з двома станами і матрицею перехідних ймовірностей

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \qquad \text{де} \quad 0 < \alpha, \beta < 1.$$

Знайти генератриси послідовностей

$$\left\{p_{11}^{(n)} |\ n = 0.1, \ldots\right\}, \qquad \left\{p_{21}^{(n)} |\ n = 0.1, \ldots\right\},$$

$$\left\{p_{12}^{(n)} |\ n = 0.1, \ldots\right\}, \qquad \left\{p_{11}^{(n)} |\ n = 0.1, \ldots\right\}$$

та точні формули для $p_{ij}^{(n)}$. Знайти також $\lim\limits_{n\to\infty} p_{ij}^{(n)} = q_{ij}$.

Задача 13.17. Ланцюг Маркова з множиною станів $\{1, 2, \ldots, N\}$ має таку матрицю перехідних ймовірностей:

$$P = \begin{pmatrix} q & p & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & q & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & q & p \end{pmatrix}.$$

Знайти стаціонарний розподіл $(q_1, \dots, q_N)$ цього ланцюга.

Задача 13.18. Нехай матриця перехідних ймовірностей $P = \|p_{ij}\|_1^N$, що відповідає неперіодичному ланцюгу Маркова з одним істотним класом станів, задовольняє умовам

$$\sum_{i=1}^{N} p_{ij} = 1, \qquad j = 1, 2, ..., N.$$

Довести, що стаціонарні ймовірності $\pi_k = 1/N$, $k = 1, 2, ..., N$.

Задача 13.19.* Кожна з двох урн містить N білих і N чорних куль. Число чорних куль у першій урні визначає стан відповідного ланцюга Маркова. На кожному кроці навмання виймають по одній кулі з кожної урни і міняють їх місцями. Знайти перехідні та стаціонарні ймовірності.

Задача 13.20.* Матриця переходу за один крок ланцюга Маркова має тридіагональний вид

$$\begin{pmatrix} r_0 & p_0 & 0 & 0 & ... & 0 & 0 & 0 \\ q_1 & r_1 & p_1 & 0 & ... & 0 & 0 & 0 \\ 0 & q_2 & r_2 & p_2 & ... & 0 & 0 & 0 \\ & & & & ... & & & \\ 0 & & & & ... & q_{N-1} & r_{N-1} & p_{N-1} \\ 0 & & & & ... & 0 & q_N & r_N \end{pmatrix}.$$

Треба знайти стаціонарний розподіл $\pi_0, \pi_1, ..., \pi_N$.

Задача 13.21. Розглянемо ланцюг Маркова з матрицею перехідних ймовірностей за один крок $P = \begin{pmatrix} 0 & 1-\alpha & \alpha \\ 1 & 0 & 0 \\ 1-\beta & \beta & 0 \end{pmatrix}$, $0 \le \alpha, \beta \le 1$. Нехай $\pi_1,\ \pi_2,\ \pi_3$ – стаціонарні ймовірності введеного ланцюга. При яких значеннях α і β $\pi_1 = \pi_2 = \pi_3$?

Задача 13.22. Нехай $\xi_n,\ n \ge 0$ – ланцюг Маркова з матрицею імовірностей переходу за один крок

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ & & \dots & & \ddots \end{pmatrix}.$$

Для зазначеного ланцюга знайти умови ергодичності, а також знайти ергодичний розподіл.

Задача 13.23. Розглянемо ланцюг Маркова з матрицею ймовірностей переходу за один крок

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 & \cdots \\ p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & p_0 & p_1 & p_2 & \cdots \\ 0 & 0 & p_0 & p_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \text{де } \forall k \geq 0 \; p_k > 0.$$

Знайти умови ергодичності та генератрису ергодичного розподілу для введеного ланцюга.

Задача 13.24. Нехай кожна людина, яка почула новину, може передати її іншому. При цьому ймовірність спотворення змісту на протилежний однакова та дорівнює $p = 0.0001$. Яка ймовірність почути новину у неспотвореному виді після того, як вона пройшла через велику кількість людей?

Задача 13.25.* Для задачі 9.66 (див. стор. 101) побудувати матрицю ймовірностей переходів за 1 крок (попередньо розглянувши задачу мінімізації кількості станів відповідного ланцюга) і розв'язати задачу методами ланцюгів Маркова.

14 ЛАНЦЮГИ МАРКОВА З НЕПЕРЕРВНИМ ЧАСОМ

Нехай задано деякий ймовірнісний простір $(\Omega, \Re, P)$. Дійснозначна функція двох аргументів $\xi(t,\omega)$, $t \in [0,\infty)$, $\omega \in \Omega$ називається **випадковим процесом**, якщо для кожного фіксованого t $\xi(t,\omega)$ є випадковою величиною. Для кожного фіксованого $\omega \in \Omega$ функція $\xi(t,\omega)$ називається **траєкторією** процесу $\xi(t,\omega)$. Аналогічно до того, як випадкову величину $\xi(\omega)$ часто позначають просто через ξ, так і випадковий процес $\xi(t,\omega)$ часто позначають через $\xi(t)$. Нехай надалі множина $X = \{x_1, x_2, \ldots\}$ можливих значень процесу $\xi(t)$ не більш ніж зліченна. Набір ймовірностей

$$\mathrm{P}\left\{\xi\left(t_1\right) = x_{i_1}, \ldots, \xi\left(t_n\right) = x_{i_n}\right\}$$

представляє собою **сумісний розподіл випадкового процесу $\xi(t)$**. Випадковий процес $\xi(t)$ називається **марковським** (або **ланцюгом Маркова з неперервним часом**), якщо для довільної зростаючої послідовності моментів часу $0 \le t_1 < t_2 < \ldots < t_n < \ldots$ послідовність $\xi_1(t)$, $\xi_2(t), \ldots, \xi_n(t), \ldots$ є ланцюгом Маркова з дискретним часом.

Функція

$$p_{ij}\left(s,t\right) = \mathrm{P}\left\{\xi(t) = x_j \middle|\, \xi(s) = x_i\right\}$$

називається **перехідною імовірністю** процесу Маркова. Якщо ця функція залежить від різниці $t-s$, то такий процес Маркова називають **однорідним**. В цьому випадку для перехідних імовірностей $p_{ij}\left(s, s+t\right) = p_{ij}\left(0,t\right) = \mathrm{P}\left\{\xi(t) = x_j \middle|\, \xi(0) = x_i\right\},\ s,t > 0$, прийняті такі позначення:

$$p_{ij}\left(0,t\right) = p_{ij}\left(t\right).$$

Перехідні ймовірності мають наступні властивості:

а) $p_{ij}(t) \ge 0$,

б) $\sum\limits_j p_{ij}(t) = 1$,

в) $p_{ij}(t+s) = \sum\limits_k p_{ik}(t)p_{kj}(s)$, $i,j = 1,2,\ldots$, $s,t \ge 0$.

Якщо розглянути матрицю перехідних імовірностей $P(t) = \|p_{ij}(t)\|$, то остання властивість в матричному вигляді може бути подана так:

$$P(t+s) = P(t)P(s).$$

Будемо називати ланцюг Маркова $\xi(t)$ **неперервним справа**, якщо з імовірністю 1 його траєкторії неперервні справа. Такий ланцюг Маркова є **стохастично неперервним**, тобто для нього виконується:

г) $\lim\limits_{t\downarrow 0} p_{ij}(t) = \delta_{ij} = \begin{cases} 0, & i \ne j \\ 1, & i = j \end{cases}.$

Крім того, для неперервного справа ланцюга Маркова існують границі

$$\lim_{h\downarrow 0} \frac{p_{ij}(h) - \delta_{ij}}{h} = a_{ij},$$

які називаються **інфінітезимальними характеристиками** (або **інтенсивностями переходів**) процесу $\xi(t)$, а матриця $A = \|a_{ij}\|$ називається **інфінітезимальною матрицею** (або **матрицею інтенсивностей**). Із визначення випливає, що

$$\forall i, j \quad a_{ij} = \frac{d}{dt} p_{ij}(t)\Big|_{t=0}.$$

Властивості інтенсивностей:

1. $\forall i \neq j\ a_{ij} \geq 0$ та $\forall i\ a_{ii} \leq 0$;
2. $\forall i \sum\limits_j a_{ij} = 0$;
3. $\forall i\ \dfrac{1}{-a_{ii}}$ є середнім часом перебування процесу в стані x_i;
4. $\forall i, j\ \hat{p}_{ij} = \dfrac{a_{ij}}{-a_{ii}}$ є ймовірністю того, що після виходу зі стану x_i процес перейде відразу в стан x_j $(x_i \neq x_j)$. Якщо $a_{ii} = 0$, то стан x_i називається поглинаючим (із нього процес вже ніколи не вийде).

Позначимо через τ_k час перебування процесу в k-му за порядком відвідування стані. Ланцюг Маркова ξ_n, $n \geq 1$, з дискретним часом, називається **вкладеним ланцюгом** для процесу Маркова $\xi(t)$, якщо

$$\xi_n = \xi\left(\tau_1 + \tau_2 + \ldots + \tau_n - 0\right).$$

Тоді $\hat{P} = \|\hat{p}_{ij}\|$ буде матрицею перехідних імовірностей для ξ_n, $n \geq 1$ (див. властивість 4).

Теорема 14.1 Для однорідного локально регулярного процесу Маркова має місце **перша система рівнянь Колмогорова**:

$$\forall i, j \quad \frac{d}{dt} p_{ij}(t) = \sum_k a_{ik} p_{kj}(t),$$

яка в матричному виді може бути подана так:

$$P'(t) = AP(t).$$

Теорема 14.2 Якщо $\sup\limits_i |a_{ii}| < \infty$, то перша система рівнянь Колмогорова має єдиний розв'язок за початкових умов $p_{ij}(0) = \delta_{ij}$.

Теорема 14.3 Для однорідного неперервного справа процесу Маркова при умові

$$\sup_i |a_{ii}| < \infty$$

має місце друга система рівнянь Колмогорова:

$$\forall i, j \quad \frac{d}{dt} p_{ij}(t) = \sum_k p_{ik}(t) a_{kj}.$$

яка в матричному виді може бути подана так:

$$P'(t) = P(t)A.$$

Із останньої теореми можна отримати систему диференціальних рівнянь для безумовних імовірностей $p_j(t) = \mathrm{P}\{\xi(t) = x_j\}$:

$$\forall j \quad \frac{d}{dt} p_j(t) = \sum_k p_k(t) a_{kj},$$

яка розв'язується за деяких початкових умов $p_j(0) = p_j^0$, де $\left\{p_j^0\right\}$ є початковим розподілом процесу, тобто $p_j^0 = \mathrm{P}\{\xi(0) = x_j\}$.

Початковий розподіл $q = \{q_j\}_{j=1}^{\infty}$ називається **стаціонарним**, якщо безумовні імовірності $p_j(t)$ не змінюються з перебігом часу t: $p_j(t) = \sum\limits_{i=1}^{\infty} q_i p_{ij}(t) \equiv q_j, \quad j = 1, 2, \ldots$ Його шукають як розв'язок системи лінійних рівнянь

$$\begin{cases} \sum\limits_{i=1}^{\infty} q_i a_{ij} = 0, \quad j = 1, 2, \ldots \\ \sum\limits_{i=1}^{\infty} q_i = 1 \quad \text{(умова нормування).} \end{cases} \tag{14.1}$$

Однорідний процес Маркова називається **ергодичним**, якщо існує імовірнісний розподіл $\pi = (\pi_1, \ldots, \pi_j, \ldots)$, $\pi_j > 0$ такий, що для будь-яких $i, j = 1, 2, \ldots$ існує незалежна від i границя

$$\lim_{t \to \infty} p_{ij}(t) = \pi_j.$$

В цьому випадку розподіл π називається **ергодичним**.

Теорема 14.4 (Ергодична теорема для процесу Маркова) Якщо для неперервного справа процесу Маркова вкладений ланцюг незвідний, то для будь-яких i, j існують

$$\lim_{t\to\infty} p_{ij}(t) = \pi_j > 0.$$

Якщо до того ж $p_{ij}(t)$ задовольняють другій системі рівнянь Колмогорова, а числа π_j такі, що

$$\sum_j \pi_j \left(-a_{jj}\right) < \infty,$$

то для усіх i

$$\sum_j \pi_j a_{ji} = 0,$$

тобто ергодичний розподіл π збігається із стаціонарним.

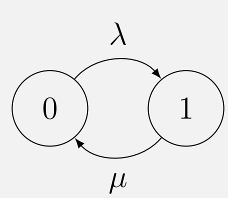
Приклад 14.1

Нехай задано ланцюг Маркова з неперервним часом, двома станами 0, 1 та інфінітезимальною матрицею A вигляду:

$$A = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}.$$

а) Знайти перехідні ймовірності $p_{ij}(t)$, $i, j = 0, 1$.
б) Знайти ймовірності станів в довільний момент часу t, якщо початковий розподіл $\bar{p}(0) = (p_0^0, p_1^0)$.
в) Чи існують граничні ймовірності $p_0(\infty)$, $p_1(\infty)$? Якщо існують, то чи залежать вони від початкового розподілу?
г) Показати, що вкладений ланцюг Маркова ергодичний і знайти (не використовуючи результат п. в)) ергодичний розподіл ймовірностей та середню долю часу, що процес проводить в кожному зі станів.

Розв'язок: а) Граф переходів процесу виглядає таким чином:



Оскільки $p_{01}(t) = 1 - p_{00}(t)$ і $p_{10}(t) = 1 - p_{11}(t)$, то перша система Колмогорова буде мати вигляд:

$$\begin{cases} p'_{00}(t) + (\lambda+\mu)p_{00}(t) = \mu \\ p'_{11}(t) + (\lambda+\mu)p_{11}(t) = \lambda \end{cases}.$$

Кожне рівняння цієї системи розв'язуємо методом варіації довільної сталої і одержуємо:

$$p_{00}(t) = \left(1 - \frac{\mu}{\lambda+\mu}\right)e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda+\mu}, \qquad p_{11}(t) = \left(1 - \frac{\lambda}{\lambda+\mu}\right)e^{-(\lambda+\mu)t} + \frac{\lambda}{\lambda+\mu}$$

б) СДР Колмогорова для безумовних ймовірностей має вигляд

$$\begin{cases} p'_0(t) = -\lambda p_0(t) + \mu p_1(t); \\ p'_1(t) = \lambda p_0(t) - \mu p_1(t). \end{cases}$$

Знову, оскільки $p_0(t) + p_1(t) = 1$, маємо

$$p'_0(t) = -(\lambda+\mu)p_0(t) + \mu.$$

Розв'язуючи це рівняння з урахуванням початкового розподілу, маємо

$$p_0(t) = \frac{\mu}{\lambda+\mu} + \frac{(\lambda+\mu)p_0^0 - \mu}{\lambda+\mu}e^{-(\lambda+\mu)t}$$

$$p_1(t) = 1 - p_0(t) = \frac{\lambda}{\lambda+\mu} + \frac{(\lambda+\mu)p_1^0 - \lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}$$

в) Легко бачити, що існують границі цих ймовірностей при $t \to \infty$:

$$\lim_{t\to\infty} p_0(t) = \frac{\mu}{\lambda+\mu};$$

$$\lim_{t\to\infty} p_1(t) = \frac{\lambda}{\lambda+\mu},$$

які не залежать від початкового розподілу.

г) Оскільки кількість станів скінченна і вони сполучаються, то ланцюг є ергодичним, отже, існує єдиний граничний розподіл ймовірностей $\bar{\pi} = (\pi_0, \pi_1)$, який співпадає зі стаціонарним. Знаходимо його з системи (14.1)

$$\begin{cases} -\lambda\pi_0 + \mu\pi_1 = 0; \\ \lambda\pi_0 - \mu\pi_1 = 0; \\ \pi_0 + \pi_1 = 1. \end{cases}$$

Звідси маємо

$$\pi_0 = \frac{\mu}{\lambda+\mu}, \quad \pi_1 = \frac{\lambda}{\lambda+\mu}.$$

Середня доля часу, який процес проводить в стані 0, дорівнює $\frac{\mu}{\lambda+\mu}$, а в стані 1 – $\frac{\lambda}{\lambda+\mu}$.

Процесом загибелі та народження називається однорідний процес Маркова із множиною станів $X = \{0, 1, 2, \ldots\}$, для вкладеного марковського ланцюга якого із кожного стану n можливими є переходи лише в стан $n-1$ та $n+1$, а зі стану 0 – лише в стан 1. Якщо $\xi(t)$ інтерпретувати як число індивідів деякої популяції, то перехід $n \to n+1$ означає народження (а перехід $n \to n-1$ означає загибель) деякого індивіда, причому в загальному випадку не виключено самозародження $(0 \to 1)$.

Для неперервного справа процесу загибелі та народження існують інтенсивності переходів. Позначимо через $\lambda_i = a_{i\,i+1}$ $(i = 0, 1, \ldots)$, $\mu_i = a_{i\,i-1}$ $(i = 1, 2, \ldots)$. Величини λ_i та μ_i є відповідно **інтенсивностями народження та загибелі**. Інфінітезимальна матриця (матриця інтенсивностей) набуває вигляду:

$$A = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \ldots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \ldots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \ldots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

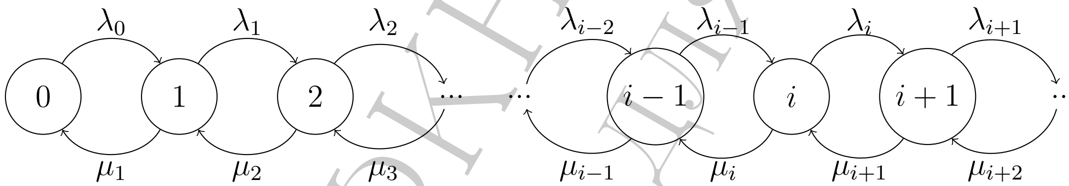


Рис. 14.1: Граф переходів процесу народження та загибелі

Система диференційних рівнянь Колмогорова для безумовних ймовірностей станів для процесів народження та загибелі набуває вигляду:

$$\frac{dp_0(t)}{dt} = -\lambda_0 p_0(t) + \mu_1 p_1(t),$$

$$\frac{dp_j(t)}{dt} = \lambda_{j-1} p_{j-1}(t) - (\lambda_j + \mu_j) p_j(t) + \mu_{j+1} p_{j+1}(t), \quad j \geq 1. \tag{14.2}$$

Приклад 14.2 Процес Пуассона

Розглянемо найпростіший процес чистого народження, який визначається як процес, для якого $\mu_i = 0 \; \forall i > 0$, а $\lambda_i = \lambda$ для всіх $i = 0, 1, 2, ...$ Підставляючи ці значення в рівняння (14.2), матимемо

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t),$$

$$\frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - \lambda_k p_k(t), \quad k \geq 1. \tag{14.3}$$

Для зручності припустимо, що процес починається в нульовий момент часу з нульового стану:

$$p_k(0) = \begin{cases} 1, & j = 0; \\ 0, & j \neq 0. \end{cases}$$

Звідси для $p_0(t)$ маємо розв'язок

$$p_0(t) = e^{-\lambda t},$$

і підставляючи цей розв'язок в рівняння (14.3) при $j = 1$ приходимо до рівняння

$$\frac{dp_1(t)}{dt} = \lambda e^{-\lambda t} - \lambda p_1(t).$$

Розв'язок цього диференційного рівняння має вид

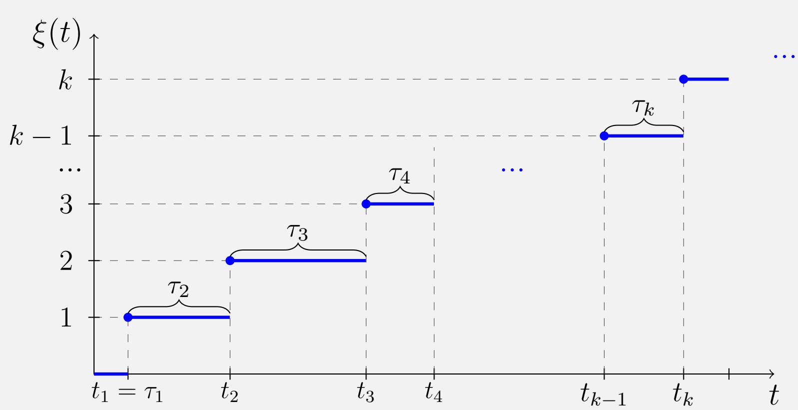
$$p_1(t) = \lambda t e^{-\lambda t},$$

Далі по індукції знаходимо розв'язок рівняння (14.3)

$$p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \;\; k \geq 0, \;\; t \geq 0.$$

Ми отримали розподіл Пуассона. Таким чином, процес чистого народження з постійною інтенсивністю λ приводить до послідовності "народжень", що утворює пуассонівський процес (потік).

Потік подій $\xi(t)$, $t \geq o$, який є процесом Пуассона (кількість подій, які трапились на інтервалі часу $[0, t)$), має властивість незалежності приростів, яка забезпучує марковську властивість: кількість подій, яка трапиться в наступному інтервалі часу, не залежить від того, яка кількість подій трапилась в минулому. Станами процесу є множина $\{0, 1, 2, ...\}$. В кожному стані процес проводить час τ_i, які є н.о.р. випадковими величинами, що мають показниковий розподіл з параметром λ. Траєкторія процесу може бути зображена сходинковою лінією, де всі стрибки дорівнюють 1, а довжина кожної сходинки дорівнює часу перебування процесу у відповідному стані.



Задача 14.1. Розглянемо $\xi(t)$, $t \geq 0$ – процес чистого народження з інтенсивностями народження λ_k, $k = 0, 1, ...$, які залежать від стану процесу. Виписати першу та другу систему диференціальних рівнянь Колмогорова для перехідних ймовірностей.

Задача 14.2. Нехай $\xi(t)$, $t \geq 0$ – процес чистого народження і $\lambda_{2n} = a$, $\lambda_{2n+1} = b$, $n = 0.1, ...$ Треба знайти $P_1(t) = \mathrm{P}\{\xi(t)$ є непарним $\}$, $P_2(t) = \mathrm{P}\{\xi(t)$ є парним $\}$, якщо $\xi(0) = 0$.

Задача 14.3. Розглянемо процес чистої загибелі $\xi(t)$, для якого $\mu_n = n \cdot \mu$, $n = 1, 2, ...$ Треба знайти $p_{ij}(t) = \mathrm{P}\{\xi(t) = j | \xi(0) = i\}$ для $j = 0.1, ..., i$.

Задача 14.4. Для пуассонівського процесу $\xi(t)$ з параметром λ знайти коваріацію між $\xi(t)$ і $\xi(t + \tau)$, де $t, \tau > 0$.

Задача 14.5. Нехай $\xi^{(1)}(t), \xi^{(2)}(t), ..., \xi^{(k)}(t)$ незалежні пуассонівські процеси з інтенсивностями $\lambda_1, \lambda_2, ..., \lambda_k$ відповідно. Довести, що $\xi(t) = \xi^{(1)}(t) + \xi^{(2)}(t) + ... + \xi^{(k)}(t)$ є пуассонівським процесом з інтенсивністю $\lambda = \lambda_1 + \lambda_2 + ... + \lambda_k$.

Задача 14.6. Розглянемо пуассонівський процес з інтенсивністю λ як потік вимог. Кожну вимогу з ймовірністю p_i, $i = 1, 2, ..., k$, $\sum_{i=1}^{k} p_i = 1$, віднесемо до i-го підпотоку незалежно від інших вимог. Довести, що i-ий підпотік є пуассонівським процесом з інтенсивністю λp_i.

Задача 14.7. Для пуассонівського процесу $\xi(t)$ та для $T > t$ і $n \geq k$ довести, що

$$\mathrm{P}(\xi(t) = k | \xi(T) = n) = C_n^k (t/T)^k (1 - t/T)^{n-k}.$$

Задача 14.8. В пологовому будинку народжується в середньому 28 дітей за тиждень. Нехай в певний день народилось 6 дітей. Припускаючи, що кількість дітей, які народилися, утворюють пуассонівський потік, знайти ймовірність того, що

а) в наступні два дні народиться по 3 дитини;

б) протігом наступних двох днів народиться 6 дітей;

в) математичне сподівання кількості днів в липні, коли народиться рівно одна дитина.

Задача 14.9. Надходження електронних листів на сервер може бути змодельовано процесом Пуассона. Нехай в середньому надходить 330 листів за хвилину. Яка інтенсивність надходження листів за секунду? Яке математичне сподівання часу між надходженням двох послідовних листів?

Задача 14.10. Відвідувачі офісу утворюють пуассонівський потік, причому в середньому приходять 4 особи впродовж години. Знайти ймовірність того, що

а) впродовж години ніхто не прийде;

б) впродовж 30 хвилин прийде 3 особи;

в) за 45 хв. прийде більше двох осіб.

Задача 14.11. Для пуассонівського потоку подій з параметром λ розглянемо подію {Трапилась рівно одна подія на інтервалі часу $[0, 2s]$}. Ймовірність цієї події

$$\mathrm{P}\{\xi_{2s} = 1\} = \lambda \cdot 2s \cdot e^{-\lambda \cdot 2s}. \tag{14.4}$$

Вказана подія еквівалентна такій події: {На проміжку $[0, s)$ не трапилось жодної події, а на проміжку $[s, 2s]$ трапилась рівно одна подія, АБО на проміжку $[s, 2s]$ не трапилось жодної події, а на проміжку $[0, s)$ трапилась рівно одна подія}. Перевірити, використовуючи властивості пуассонівського процесу, що ймовірність цієї події буде дорівнювати (14.4).

Задача 14.12. ***k-вимірним процесом Пуассона*** *з інтенсивністю λ називають набір $X_1, X_2, \ldots$ випадкових точок, які володіють такими властивостями: якщо $\nu(A)$ – це кількість таких точок в множині A, то*

1. *(однорідність) Випадкова величина $\nu(A)$ має розподіл Пуассона з параметром $\lambda m(A)$;*
2. *(незалежність) Якщо множини $A_1, A_2, \ldots, A_n$ не перетинаються, то випадкові величини $\nu(A_1)$, $\nu(A_2)$, ..., $\nu(A_n)$ є незалежними в сукупності.*

Нехай локалізація дефектів певної поверхні (з певного матеріалу) узгоджується з двовимірною моделлю пуассонівського процесу. Для цього матеріалу відомо, що він містить в середньому 5 дефектів на квадратний метр. Яка ймовірність того, що смужка довжиною 2м та шириною 5см не буде містити дефектів?

Задача 14.13. Нехай ξ – випадкова величина, яка має розподіл Пуассона з параметром λ. Довести, що якщо $\lambda < 1$, то ймовірності $\mathrm{P}\{\xi = k\}$ будуть строго спадати з ростом k, якщо $\lambda > 1$, то $\mathrm{P}\{\xi = k\}$ спочатку зростає, потім спадає. Як будуть себе поводити ці ймовірності при $\lambda = 1$?

Задача 14.14. Нехай ξ – випадкова величина, що має розподіл Пуассона, для якої $\mathrm{P}\{\xi = 0\} = \mathrm{P}\{\xi = 1\}$. Чому дорівнює $\mathbf{M}\xi$?

Задача 14.15. Знайти функцію розподілу моменту τ_1 появи першої події пуассонівського потоку з інтенсивністю λ, якщо відомо, що до моменту часу T трапилась рівно одна подія.

Задача 14.16. Модель Пуассона використовують для дослідження потоку машин, якщо вони їдуть повільно. Нехай 20-хвилинний інтервал розділено на підінтервали по 10 с. В певній точці автотраси реєструється кількість машин, які проїхали протягом кожного підінтервалу. Нехай ν_i – кількість підінтервалів, в яких проїхало i машин, $i = 0, 1, \ldots, 9$. Припустимо, що зафіксовано такі результати спостереження:

i	0	1	2	3	4	5	6	7	8	9
ν_i	19	38	28	20	7	3	4	0	0	1

Загальна кількість машин, які проїхали повз реєстратор за ці 20 хвилин, дорівнює 230.
а) Як можна було б обчислити параметр інтенсивності потоку λ?
б) Припустимо, що ймовірність того, що не проїхало жодної машини за 10 секунд, можна

оцінити величиною ν_0, діленою на загальну кількість підінтервалів спостережень. Чи узгоджується це (логічно) зі значенням, яке Ви запропонували для п. а)?
в) Що можна було б запропонували в якості оцінки ймовірності того, що за 10 секунд проїде 10 машин?

Задача 14.17. Розглянемо один з засобів визначення багатовимірного пуассонівського процесу. Нехай $(\xi(t), \eta(t))$ – двовимірний вектор, де $\xi(t) = \alpha(t) + \delta(t)$, $\eta(t) = \beta(t) + \delta(t)$, а $\alpha(t)$, $\beta(t)$, $\delta(t)$ – три незалежних пуассонівські процеси з параметрами λ_1, λ_2 і λ_3 відповідно. Треба знайти генератрису $(\xi(t), \eta(t))$.

Задача 14.18. Для ланцюга Маркова $\xi(t) \in \{1, 2\}$, який визначений у прикладі на стор. 191, позначимо через $\tau_k(t)$, $k = 1, 2$ сумарний час перебування у стані "k" на проміжку $(0, t)$. Треба знайти $m_{ij}(t) = \mathbf{M}(\tau_j(t)|\, \xi_0 = i)$, $i, j = 1, 2$.

Задача 14.19. Нехай інфінітезимальна матриця ланцюга Маркова з неперервним часом має вигляд

$$\begin{pmatrix} -\lambda & \lambda & 0 \\ \mu & -(\lambda+\mu) & \lambda \\ \mu & 0 & -\mu \end{pmatrix}.$$

Треба знайти матрицю переходу за один крок вкладеного ланцюга і його стаціонарні ймовірності.

Задача 14.20. На телефонну лінію зв'язку можуть надходити виклики двох типів: прості і термінові. Будь-який виклик, що надійшов у вільну лінію, займає її. Простий виклик, що надійшов у зайняту лінію, отримує відмову і губиться. Терміновий виклик, що надійшов у зайняту лінію, перериває розмову, яку ведуть у цей момент часу, і сам займає лінію. Моменти надходження простих і термінових викликів утворюють пуассонівські процеси з інтенсивностями λ_1 і λ_2 відповідно. Випадковий час розмови має показниковий розподіл з параметром μ. Треба побудувати ланцюг Маркова з трьома станами, який відповідає описаному вище процесу обслуговування, і знайти стаціонарні ймовірності того, що лінія вільна, зайнята простим викликом, зайнята терміновим викликом.

Задача 14.21. Є два трансатлантичних кабелі, кожен з яких може передавати тільки одне телеграфне повідомлення. Час роботи до відмови кожного кабелю розподілений за показниковим законом з параметром λ. Час ремонту кожного кабелю має один і той самий показниковий розподіл з параметром μ. Треба знайти ймовірність того, що у момент часу t обидва кабелі справні, якщо така ж ситуація була у початковий момент.

Задача 14.22. У молекулярній біології виникає наступна задача. Поверхня бактерії містить декілька точок, де можуть прикріплюватись молекули, що надійшли ззовні, якщо вони мають правильну будову. Молекули, що мають правильну будову, будемо називати допустимими. Розглянемо фіксовану точку і будемо вважати, що молекули надходять у цю точку у відповідності до пуассонівського процесу з параметром μ. Долю p $(0 < p < 1)$ цих молекул складають допустимі. Недопустимі молекули перебувають у цій точці випадковий час, розподілений за показниковим законом з параметром λ. Поки вони прикріплені до бактерії, вони перешкоджають іншим молекулам. Допустима молекула займає це місце назавжди і також перешкоджає іншим молекулам. Знайти ймовірність того, що точка вільна від молекул у момент часу t.

Задача 14.23. Телефонний вузол має m каналів зв'язку. Моменти надходження викликів утворюють пуассонівський процес з параметром λ. Виклики обслуговуються, якщо є вільний канал зв'язку. У протилежному випадку вони губляться. Тривалість кожної розмови є випадковою величиною, розподіленою за показниковим законом з параметром μ. Тривалості окремих розмов – незалежні випадкові величини. Треба знайти стаціонарний розподіл числа зайнятих каналів.

Задача 14.24. Система складається з N ідентичних компонент, кожна з яких працює незалежно від інших випадковий час до відмови. Час роботи до відмови має показниковий розподіл з параметром λ. Коли яка-небудь компонента відмовляє, вона ремонтується. Час ремонту є випадковою величиною, розподіленою за показниковим законом з параметром μ. Треба знайти середнє число працюючих компонент системи у стаціонарному режимі.

Задача 14.25. Нехай $\xi(t) \in \{0.1, ..., N\}$ - ланцюг Маркова з неперервним часом, який задається інфінітезимальною матрицею

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 & 0 & 0 \\ \mu & -(\lambda+\mu) & \lambda & \dots & 0 & 0 & 0 \\ & & & \dots & & & \\ 0 & 0 & 0 & \dots & \mu & -(\lambda+\mu) & \lambda \\ & & & \dots & & \mu & -\mu \end{pmatrix}.$$

Треба знайти стаціонарні ймовірності $\pi_i^{(N)}$, $i = 0.1, ..., N$ як функції $\rho = \lambda/\mu$. У випадку $\rho < 1$ знайти границі

$$\pi_j = \lim_{N\to\infty} \pi_j^{(N)},\ j = 0.1, ...$$

Для $\rho > 1$ знайти границі

$$\pi_j^* = \lim_{N\to\infty} \pi_{N-j}^{(N)},\ j = 0.1, ...$$

Задача 14.26. Розглянемо процес народження і загибелі, для якого інтенсивності народження і загибелі дорівнюють відповідно

$$\lambda_k = (k+2)\lambda\ \ k = 0, 1, 2, ... \qquad \text{і} \qquad \mu_k = k\cdot\mu,\ k = 1, 2, ...$$

Треба знайти стаціонарні ймовірності π_k, $k = 0.1, 2, ...$

Задача 14.27. Розглянемо процес народження і загибелі, для якого інтенсивності народження і загибелі дорівнюють відповідно

$$\lambda_k = \begin{cases} \lambda, 0 \le k \le N, \\ 2\lambda, k > N, \end{cases} \qquad k = 0, 1, 2, ... \qquad \mu_k = \mu, \quad k = 1, 2, ...$$

З'ясувати, для яких λ і μ існують стаціонарні ймовірності π_k, $k = 0.1, 2, ...$ і знайти їх.

Задача 14.28. Нехай $\xi(t) \in \{0.1, ...\}$ є процесом народження і загибелі з інтенсивностями загибелі $\mu_n > 0$, $n = 1, 2, ...$ і інтенсивностями народження $\lambda_n = \lambda > 0$, $n = 0.1, ...$ Будемо припускати, що для $\xi(t)$ виконуються умови існування стаціонарного розподілу і початковий розподіл збігається зі стаціонарним. Довести, що число випадків загибелі на проміжку часу $[0, t]$ має розподіл Пуассона з параметром $\lambda \cdot t$.

Задача 14.29. Розглянемо систему масового обслуговування, на вхід якої ззовні надходить пуассонівський потік вимог інтенсивності λ. Для обслуговування вимог є необмежена кількість однотипних приладів. Час обслуговування розподілений за показниковим законом з параметром μ. Нехай $\xi(t)$ – число зайнятих приладів у момент часу t. Треба знайти:

а) стаціонарний розподіл для процесу $\xi(t)$;

б) при умові, що у початковий момент система порожня, ймовірності $p_k(t) = \mathrm{P}(\xi(t) = k)$, $k = 0.1, ...$, як функції від t.

Задача 14.30. Пуассонівський потік вимог інтенсивності λ обслуговується одним приладом. Час обслуговування розподілений за показниковим законом з параметром μ. Якщо прилад зайнятий, то утворюється черга. Через $\xi(t)$ будемо позначати число вимог у системі. При якій умові на параметри λ і μ для $\xi(t)$ існує стаціонарний розподіл? Якщо стаціонарний розподіл існує, знайти явний вигляд стаціонарних ймовірностей.

Задача 14.31. Розглянемо марковську систему масового обслуговування з одним приладом і повторними викликами. Це означає, що на обслуговуючий прилад надходить пуассонівський потік вимог інтенсивності λ. Час обслуговування розподілений за показниковим законом з параметром μ. За умови, що у момент надходження вимоги, прилад вільний, вимога займає його і одразу починається її обслуговування. У протилежному випадку вимога стає джерелом повторних викликів. Незалежно від інших вимог вона через випадковий час, розподілений за показниковим законом з параметром ν, опитує прилад. Якщо він вільний, то займає його. Якщо прилад зайнятий, то продовжує опитування. Процес обслуговування у системі, що розглядається, буде двовимірним

$$\xi(t) = (\xi_1(t), \xi_2(t)).$$

Перша компонента $\xi_1(t)$ приймає два значення – 0 або 1 в залежності від того, вільний обслуговуючий прилад, чи ні. Друга компонента $\xi_2(t) \in \{0.1, ...\}$ дорівнює числу джерел повторних викликів. Треба з'ясувати, за якої умови для $\xi(t)$ існує стаціонарний розподіл і виписати стаціонарні ймовірності через параметри λ, μ і ν.

Задача 14.32. Розглянемо ланцюг Маркова $\xi(t)$, $t \geq 0$, який визначено у задачі 14.30. Методом генератрис знайти

$$P_{0k}(t) = \mathrm{P}(\xi(t) = k|\, \xi(0) = 0), \quad k = 0.1, ...$$

як функції від t.

Задача 14.33.* При умові $\xi(0) = 0$ для процесу $\xi(t)$, який визначено у задачі 14.30, розглянемо випадкову величину

$$\tau_n = \inf\{t : \xi(t) = n\},$$

n – натуральне число. Виписати $m_n = \mathbf{M}\tau_n$ через параметри λ, μ, і у кожному з трьох випадків
а) $\lambda/\mu < 1$, б) $\lambda/\mu = 1$, в) $\lambda/\mu > 1$
і знайти слабку границю τ_n/m_n при $n \to \infty$.

15 ВИКОРИСТАННЯ ЙМОВІРНІСНИХ ХАРАКТЕРИСТИК В АКТУАРНИХ РОЗРАХУНКАХ

Вивчення математичних моделей і методів у страхуванні прийнято називати актуарною математикою (actuarial mathematics). Актуарна математика широко використовує закони та поняття теорії ймовірностей. Зокрема, значною мірою це пов'язано з таким напрямком актуарної науки, яким є теорія страхування життя. Невизначеність моменту смерті є основним джерелом випадковості при страхуванні життя. Таким чином, ми можемо моделювати **тривалість людського життя** невід'ємною випадковою величиною X, стохастична природа якої описується функцією розподілу $F(x)$. В актуарній математиці зазвичай працюють не з функцією розподілу, а з функцією

$$s(x) = 1 - F(x),$$

яка називається **функцією виживання**. Отже, значення функції $s(x)$ – це є ймовірність того, що людина доживе до віку x років. З визначення випливає, що функція виживання має наступні властивості:

- $s(x)$ спадає (нестрого);
- $s(0) = 1$, $s(+\infty) = 0$;
- $s(x)$ неперервна справа.

Крім того, природно передбачити, що $s(x)$ є неперервною, позаяк інакше існував би деякий фіксований невипадковий момент у житті людини, в який вона помирала би з додатньою ймовірністю.

Друге природне обмеження, пов'язане зі спаданням функції виживання. Зрозуміло, що $s(x)$ має бути строго спадною функцією, інакше існував би фіксований невипадковий період у житті людини, коли її смерть є неможливою. В актуарній математиці передбачається, що випадкова величина X може приймати як необмежені значення, та і те, що існує граничний вік ω (як правило, $\omega \in [100;\ 120]$ років), для якого $s(x) = 0$ при $x > \omega$. Наведемо ще деякі імовірнісні поняття, що використовуються в актуарних розрахунках. Функція

$$f(x) = -s'(x) = F'(x)$$

в точках своєї неперервності називається в актуарній математиці **кривою смертей**. Величина

$$\mu_x = \frac{f(x)}{1 - F(x)}$$

називається **інтенсивністю смертності**, оскільки при малих значеннях t величина $\mu_x t$ наближено описує ймовірність смерті в інтервалі $(x,\ x+t)$ людини, що дожила до x років.

Функція виживання пов'язана з інтенсивністю смерті наступним чином:

$$s(x) = \exp\left\{-\int\limits_0^x \mu_u du\right\}.$$

Вкрай важливим з практичної точки зору є такі характеристики тривалості життя як його **математичне сподівання** $\overset{0}{e}_0$, дисперсія, та **математичне сподівання залишкового часу життя людини** $\overset{0}{e}_x$, що дожила до віку x років. Ці величини для актуарних задач можна обчислити за наступними формулами:

$$\overset{0}{e}_0 = \mathbf{M}X = \int\limits_0^{+\infty} s(x)dx,$$

$$\mathbf{D}X = \mathbf{M}X^2 - (\mathbf{M}X)^2, \qquad \mathbf{M}X^2 = 2\int\limits_0^{+\infty} xs(x)dx,$$

$$\overset{0}{e}_x = \frac{1}{s(x)}\int\limits_x^{+\infty} s(u)du.$$

Задача 15.1. Показати, що середню тривалість життя $\overset{0}{\underset{o}{e}} = \mathbf{M}X$ можна розрахувати за формулою:

$$\overset{0}{\underset{o}{e}} = \int\limits_0^{+\infty} s(x)dx.$$

Задача 15.2. Показати, що величину $\mathbf{M}X^2$ можна обчислити за формулою

$$\mathbf{M}X^2 = 2\int\limits_0^{+\infty} xs(x)dx.$$

Задача 15.3. Крива смертей задається формулою

$$f(x) = \frac{2}{\lambda}x\exp\left(-\frac{x^2}{\lambda}\right), \quad \lambda > 0.$$

Знайти μ_x.

Задача 15.4. Підрахувати математичне сподівання часу життя, що залишився людині, яка дожила до віку 40 років, якщо крива смертей задається формулою:

$$f(x) = \frac{1}{80}\,e^{-\frac{x}{40}} + \frac{1}{160}\,e^{-\frac{x}{80}}.$$

Задача 15.5. Нехай інтенсивність смертності задається формулою: $\mu_x = \dfrac{x}{a(x+a)}$. Знайти $\overset{o}{\underset{x}{e}}$.

Задача 15.6. При якому значенні константи c функція $f(x) = \dfrac{x}{1089}e^{-\frac{x}{c}}$ може бути кривою смертей?

Задача 15.7. Знайти інтенсивність смертності μ_x, якщо крива смертей має вигляд

$$f(x) = \frac{1}{60}e^{-\frac{x}{30}} + \frac{1}{180}e^{-\frac{x}{90}}.$$

Задача 15.8. Нехай крива смертей задається формулою

$$f(x) = 2x^5 \exp\left(-\frac{x^6}{3}\right), \quad x > 0.$$

Знайти функцію виживання $s(x)$, функцію розподілу тривалості життя $F(x)$ та інтенсивність смертності μ_x.

Задача 15.9. Нехай крива смертей задається формулою

$$f(x) = \exp\left(2x - \frac{1}{2}(e^{2x} - 1)\right), \quad x > 0.$$

Знайти функцію виживання $s(x)$, функцію розподілу тривалості життя $F(x)$ та інтенсивність смертності μ_x.

Задача 15.10. Нехай інтенсивність смертності задається формулою $\mu_x = kx^n$, $x > 0$, $k > 0$, $n > 0$. Знайти функцію розподілу тривалості життя $F(x)$, криву смертей $f(x)$ та функцію виживання $s(x)$.

Задача 15.11. Нехай крива смертей задається формулою

$$f(x) = kx^n \exp\left(-\frac{kx^{n+1}}{n+1}\right), \quad x > 0.$$

Знайти функцію виживання $s(x)$, функцію розподілу тривалості життя $F(x)$.

Задача 15.12. Нехай функція розподілу тривалості життя задається формулою

$$F(x) = 0.01x, \quad 0 \le x \le 100.$$

Знайти функцію виживання $s(x)$, криву смертей $f(x)$ та інтенсивність смертності μ_x.

16 ОЦІНЮВАННЯ НЕВІДОМИХ ПАРАМЕТРІВ РОЗПОДІЛІВ

Вибіркою будемо називати сукупність випадкових величин $\xi' = (\xi_1, ..., \xi_n)$, що спостерігаються в експерименті. Величини $\xi_i,\ i = 1, 2, ..., n$ – це елементи вибірки, n – об'єм вибірки. Реалізації вибірки ξ будемо позначати через $x' = (x_1, ..., x_n)$.

Вибірковий простір – множина усіх можливих значень вибірки ξ:

$$X = \{x\}.$$

Вибірковий простір може бути або усім n-вимірним евклідовим простором Re^n , або його частиною. У кожному конкретному випадку задається σ-алгебра U на вибірковому просторі X. Під **статистичною моделлю експерименту** будемо розуміти набір

$$(\Omega, U, \mathscr{P}),$$

де $\mathscr{P}$ – клас усіх допустимих розподілів випадкового вектора ξ, які можуть бути заданими на X.

Розподіл ймовірностей довільного випадкового вектора однозначно визначається його функцією розподілу. Тому статистична модель експерименту визначається вибірковим простором X і сімейством функцій розподілу F, якому належить невідома функція розподілу

$$F_\xi(x_1, ..., x_n) = \mathrm{P}\{\xi_1 \le x_1, ..., \xi_n \le x_n\}$$

вибірки $\xi' = (\xi_1, ..., \xi_n)$. Таким чином, статистичну модель можна задавати у вигляді (X, U, F).

Часто виникають ситуації, коли компоненти $\xi_1, ..., \xi_n$ – незалежні і розподілені так, як деяка випадкова величина ξ_0. Це відповідає експерименту, у якому провадяться повторні незалежні спостереження над випадковою величиною ξ_0. Таку модель можна задавати в термінах функції розподілу $F_{\xi_0}(\cdot)$, оскільки $F_\xi(x) = F_{\xi_1}(x_1) \cdot ... \cdot F_{\xi_n}(x_n)$, і $F_{\xi_i}(\cdot) = F_{\xi_0}(\cdot)$ для $i = 1, 2, ..., n$. У цьому випадку говорять, що $\xi' = (\xi_1, ..., \xi_n)$ – вибірка з розподілу випадкової величини ξ_0.

Іноді множину можливих значень ξ_0 з розподілом $F_{\xi_0}(\cdot)$ називають **генеральною сукупністю**, яка має функцію розподілу $F_{\xi_0}(\cdot)$, а вектор ξ – **вибіркою з генеральної сукупності**. Розподіл ξ_0 позначають $\mathscr{L}(\xi_0)$.

Параметрична модель – це така модель, у якій функції розподілу з класу $\mathscr{F}$ задані з точністю до значень деякого параметру θ, який приймає значення з множини Θ:

$$\mathscr{F} = \{F(x; \theta),\ \theta \in \Theta\}.$$

Говорять, що в цьому випадку відомий тип розподілу випадкової величини, що спостерігається. Невідомим є тільки параметр, від якого залежить розподіл.

16.1 Оцінювання невідомих параметрів

Розглянемо довільну параметричну ймовірнісно-статистичну модель

$$\mathscr{F} = \{F(x;\theta), \quad \theta \in \Theta\}.$$

яка відповідає схемі повторних незалежних спостережень над деякою випадковою величиною ξ_0. Нехай $\xi' = (\xi_1, ..., \xi_n)$ – вибірка з розподілу $\mathscr{L}(\xi_0) \in \mathscr{F}$.

У загальному вигляді задача оцінки параметра θ формулюється так: використовуючи статистичну інформацію, яка міститься у виборці ξ, зробити статистичні висновки про дійсне значення θ^0 невідомого параметра θ, тобто оцінити точку θ^0. Таким чином, задача оцінювання параметра θ полягає у побудові наближених формул

$$\theta^0 \approx T(\xi) \tag{16.1}$$

де $T(x) = T(x_1, ..., x_n)$ – вимірна функція на вибірковому просторі X, яка приймає значення з множини Θ. При цьому функція $T(x)$ повинна бути однією і тією ж самою для усіх розподілів $F(x,\theta)$ з даного сімейства $\mathscr{F}$.

Статистикою називається довільна випадкова величина

$$T = T(\xi),$$

яка є функцією лише від вибірки $\xi' = (\xi_1, ..., \xi_n)$. При точковому оцінюванні шукають статистику $T = T(\xi)$, значення якої при заданій реалізації $x' = (x_1, ..., x_n)$ вибірки ξ приймають за наближене значення параметру θ^0. У цьому випадку говорять, що статистика $T(\xi)$ оцінює θ або що статистика $T(\xi)$ є оцінкою θ. Як правило (але не завжди) область значень оцінки $T(\cdot)$ збігається з Θ.

Ясно, що для оцінювання θ можна використовувати різні оцінки і, щоб вибрати найкращу з них, треба мати критерій якості оцінок. Цей критерій визначається вибором міри близькості оцінки до дійсного значення параметра. Якщо зафіксований клас оцінок і вибрана міра близькості, то оцінка, що мінімізує цю міру близькості, і буде оптимальною.

Припустимо, що параметр θ – скалярний. Статистика $T(\xi)$ називається **незсуненою** оцінкою параметра θ, якщо виконується умова

$$\mathbf{M}_\theta T(\xi) = \theta \text{ для будь-якого } \theta \in \Theta. \tag{16.2}$$

Для оцінок, які не задовольняють умові (16.2), величину

$$b(\theta) = \mathbf{M}_\theta T(\xi) - \theta \tag{16.3}$$

називають **зсувом** оцінки $T(\xi)$. **Середньоквадратичною похибкою** оцінки $T(\xi)$ за визначенням є

$$\mathbf{M}_\theta (T(\xi) - \theta)^2 = \mathbf{D}_\theta T + b^2(\theta). \tag{16.4}$$

Для незсунених оцінок середньоквадратична похибка збігається з дисперсією $\mathbf{D}_\theta T(\xi)$.

Для розширення застосувань передбачимо можливість оцінювання не тільки параметра θ, а і параметричної функції від нього $\tau(\theta)$. У цьому випадку статистика $T = T(\xi)$ є **незсуненою оцінкою** для $\tau(\theta)$, якщо виконується співвідношення

$$\mathbf{M}_\theta T(\xi) = \tau(\theta) \quad \text{для будь-якого } \theta \in \Theta. \tag{16.5}$$

Дамо формальне визначення поняття оптимальної оцінки. Нехай треба оцінити задану параметричну функцію $\tau = \tau(\theta)$ у моделі $\mathscr{F} = \{F(x;\theta), \theta \in \Theta\}$ за статистичною інформацією, яка міститься у виборці $\xi' = (\xi_1, ..., \xi_n)$.

Припустимо, що для даної задачі існують незсунені оцінки, тобто статистики $T(\xi)$, що задовольняють (16.5). Позначимо через $\Im_\tau$ – клас усіх незсунених оцінок. Додатково припустимо, що дисперсії усіх оцінок з класу $\Im_\tau$ скінченні. Якщо для $T^*(\cdot) \in \Im_\tau$ виконується співвідношення

$$\mathbf{D}_\theta T^* \le \mathbf{D}_\theta T \text{ для усіх } T(\cdot) \in \Im_\tau \text{ і } \theta \in \Theta, \tag{16.6}$$

то оцінку $T^*(\cdot)$ називають **незсуненою оцінкою з рівномірно мінімальною дисперсією** або **оптимальною** оцінкою. Щоб підкреслити, що вона відноситься до функції $\tau(\theta)$, ми будемо також використовувати позначення τ^*.

Будемо казати, що дві статистики T_1 і T_2 дорівнюють одна одній $T_1 = T_2$, якщо

$$P_\theta(\xi \in \{x : T_1(x) \ne T_2(x)\}) = 0 \text{ для усіх } \theta \in \Theta.$$

Справедлива наступна теорема про єдиність оптимальної оцінки.

Теорема 16.1 Нехай $T_i = T_i(\xi)$, $i = 1, 2$ – дві оптимальні оцінки для $\tau = \tau(\theta)$. Тоді

$$T_1 = T_2.$$

Разом з властивостями незсуненості і оптимальності, що були введені при фіксованому об'ємі вибірки, розглянемо асимптотичні властивості оцінок.

Оцінку $T(\xi) = T(\xi_1, ..., \xi_n) = T_n$ можна розглядати як послідовність оцінок при $n \to \infty$. Послідовність оцінок T_n будемо називати **асимптотично незсуненою** при оцінюванні параметричної функції $\tau(\theta)$, якщо

$$\lim_{n\to\infty} \mathbf{M}_\theta T_n = \tau(\theta) \quad \text{для усіх } \theta \in \Theta.$$

Послідовність оцінок $T_n(\xi)$ називається **асимптотично нормальною** при оцінюванні параметра θ з коефіцієнтом $\sigma^2(\theta)$, якщо

$$\sqrt{n}(T_n - \theta) \underset{n\to\infty}{\Longrightarrow} \eta,$$

де випадкова величина η має нормальний розподіл $N(0, \sigma^2(\theta))$.

Асимптотична нормальність оцінок є важливою властивістю послідовності оцінок. В подальшому ми побачимо, що ця властивість використовується, зокрема, при побудові надійних інтервалів для невідомих параметрів та в задачах перевірки гіпотез про значення цих параметрів. Вона демонструє, окрім іншого, що швидкість збіжності до випадкової величини має порядок $1/\sqrt{n}$. Проте відсутність такої властивості не є чимось поганим. Натомість для оцінок, які не є асимптотично ненормальними, порядок збіжності навіть може виявитися вищим.

Приклад 16.1

Нехай $\xi_1, ..., \xi_n$ – вибірка розміру n з рівномірного розподілу на інтервалі $[0, \theta]$ з невідомим параметром $\theta > 0$. Чи є оцінки $\hat\theta_1 = 2\bar\xi$ та $\hat\theta_2 = \xi_{(n)}$ асимптотично нормальними?

Розв'язок. Згідно ЦГТ (див. стор. 163),

$$\sqrt{n}(\hat\theta_1 - \theta) = \sqrt{n}(2\bar\xi - \theta) = \sqrt{n}\left(2\frac{\sum_{i=1}^n \xi_i}{n} - \theta\right) =$$
$$= \frac{\sum_{i=1}^n (2\xi_i) - n\theta}{\sqrt{n}} = \frac{\sum_{i=1}^n (2\xi_i) - n\mathbf{M}(2\xi_1)}{\sqrt{n}} \Rightarrow \mathrm{N}(0, \mathbf{D}(2\xi_1)) = \mathrm{N}(0, 4\mathbf{D}\xi_1).$$

Отже, оцінка $\hat\theta_1 = 2\bar\xi$ є асимптотично нормальною з коефіцієнтом

$$\sigma^2(\theta) = 4\mathbf{D}\xi_1 = 4 \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3}.$$

Для того, щоб перевірити асимптотичну нормальність оцінки $\hat\theta_2 = \xi_{(n)}$, скористаємося визначенням слабкої збіжності: послідовність випадкових величин $\xi_n \Rightarrow \xi_0$ збігається за розподілом (слабко) до випадкової величини ξ_0 з функцією розподілу $F(x)$, якщо для будь-якої точки x, яка є точкою неперервності функції розподілу $F(x)$, має місце збіжність

$$\mathbf{P}\{\xi_n \le x\} \to F(x) \ \text{ при } \ n \to \infty.$$

Відповідно, слабка збіжність величини $\sqrt{n}(\xi_{(n)} - \theta)$ до випадкової величини, що має нормальний розподіл $\mathrm{N}(0, \sigma^2(\theta))$ матиме місце, якщо *для будь-якого* $x \in \mathbb{R}$

$$\mathbf{P}\{\sqrt{n}(\xi_{(n)} - \theta) \le x\} \to \Phi_{(0,\sigma^2(\theta))}(x) \ \text{ при } \ n \to \infty.$$

Проте в точці $x = 0$ функція розподілу величини $\sqrt{n}(\xi_{(n)} - \theta)$ дорівнює 1. А для нормального розподілу $\mathrm{N}(0, \sigma^2(\theta))$ функція розподілу в нулі дорівнює 1/2. Оскільки 1 не збігається до 1/2 при $n \to \infty$, то слабка збіжність $\sqrt{n}(\xi_{(n)} - \theta)$ до випадкової величини, що має нормальний розподіл $\mathrm{N}(0, \sigma^2(\theta))$ не має місця, а отже оцінка $\hat{\theta}_2 = \xi_{(n)}$ не є асимптотично нормальною.

При дослідженні оцінок більш вагомою є наступна властивість. Послідовність оцінок T_n параметричної функції $\tau(\theta)$ називається **консистентною** для $\tau(\theta)$, якщо

$$T_n \underset{n\to\infty}{\overset{P_\theta}{\Longrightarrow}} \tau(\theta) \qquad \text{для усіх } \theta \in \Theta.$$

Якщо збіжність за ймовірністю замінити на збіжність з імовірністю 1, то будемо мати **сильну консистентність**.

Для того, щоб перевірити консистентність оцінки, зазвичай використовують один з трьох наступних способів.

1) Консистентність доводиться на основі безпосереднього обчислення функції розподілу оцінки та перевірки збіжності за визначенням (див. приклад 16.2 на стор. 209):

$$\mathbf{P}\{|T_n - \tau(\theta)| > \varepsilon\} \underset{n\to\infty}{\longrightarrow} 0.$$

2) Для оцінок, побудованих по усереднених сумах, перевірка консистентності спирається на використання закону великих чисел (див. стор 162) (сильна консистентність – на використання посиленого ЗВЧ) та Твердження 1.

Твердження 1. *Нехай випадкові величини $\eta_1(n), .., \eta_r(n)$ збігаються за ймовірністю при $n \to \infty$ до деяких сталих $c_1, ..., c_r$. Тоді для довільної функції $\phi(x_1, ..., x_r)$, неперервної в точці $(c_1, ..., c_r)$,*

$$\phi(\eta_1(n), .., \eta_r(n)) \underset{n\to\infty}{\overset{\mathbf{P}}{\longrightarrow}} \phi(c_1, ..., c_r).$$

3) При доведенні консистентності використовується наступний допоміжний результат.

Твердження 2. *Якщо зсув $b_n(\theta) = \mathbf{M}_\theta T_n - \tau(\theta)$ і дисперсія $\mathbf{D}_\theta T_n$ прямують до нуля при $n \to \infty$, то оцінка T_n є консистентною оцінкою $\tau(\theta)$.*

Приклад 16.2

Для випадкових величин ξ_i, $i = 1, ..., n$, рівномірно розподілених на відрізку $[0, \theta]$, доведемо консистентність оцінки $\hat{\theta}_1 = \xi_{(n)} = \max\{\xi_1, ..., \xi_n\}$ параметра θ
а) безпосередньо обчислюючи функцію розподілу;
б) застосовуючи Твердження 2.

Розв'язок. а) Функція розподілу n-тої порядкової статистики дорівнює $F_{\xi_{(n)}}(x) = (x/\theta)^n$ при $0 \le x \le \theta$ (доведіть це аналогічно п.1 прикладу 16.4). Оскільки $\mathbf{P}\{\xi_{(n)} \le \theta\} = 1$, то для будь-якого $\varepsilon \in (0, \theta)$ маємо

$$\mathbf{P}\{|\hat{\theta}_1 - \theta| > \varepsilon\} = \mathbf{P}\{\xi_{(n)} \le \theta - \varepsilon\} = (1 - \varepsilon/\theta)^n \to 0 \ \text{ при } \ n \to \infty,$$

що означає, що оцінка консистентна.

б) Використовуючи функцію розподілу n-тої порядкової статистики $\xi_{(n)}$, можна знайти $\mathbf{M}\hat{\theta}_1 = \mathbf{M}\xi_{(n)} = \frac{n\theta}{n+1} \to \theta$ та $\mathbf{D}\hat{\theta}_1 = \mathbf{D}\xi_{(n)} = \frac{n\theta^2}{(n+1)^2(n+2)} \to 0$ при $n \to \infty$. Отже, оцінка консистентна.

Приклад 16.3

Для статистичної моделі прикладу 16.2 розглянемо оцінку $\hat{\theta}_2 = (n+1)\xi_{(1)}$, де $\xi_{(1)} = \min\{\xi_1, ..., \xi_n\}$. Використовуючи функцію розподілу першої порядкової статистики $\xi_{(1)}$ рівномірного на $[0, \theta]$ розподілу, можемо знайти її математичне сподівання (знайдіть його аналогічно п.1 прикладу 16.4), і в результаті матимемо

$$\mathbf{M}\hat{\theta}_2 = (n+1)\mathbf{M}\xi_{(1)} = (n+1)\left(\theta - \frac{n\theta}{n+1}\right) = \theta.$$

З незалежності випадкових величин $\xi_1, .., \xi_n$ випливає, що при $n \to \infty$

$$\mathrm{P}\left\{\hat{\theta}_2 > \theta + \varepsilon\right\} = \mathrm{P}\left\{(n+1)\xi_{(1)} > \theta + \varepsilon\right\} = \prod_{i=1}^{n} \mathrm{P}\left\{\xi_i > \frac{\theta+\varepsilon}{n+1}\right\} =$$
$$\left(1 - F_{\xi_1}\left(\frac{\theta+\varepsilon}{n+1}\right)\right)^n = \left(1 - \frac{\theta+\varepsilon}{\theta(n+1)}\right)^n \to e^{-(\theta+\varepsilon)/\theta} \not\to 0.$$

Отже, оцінка $\hat{\theta}_2 = (n+1)\xi_{(1)}$ є незсуненою, але не є консистентною оцінкою параметра θ.

Приклад 16.4

Нехай $\xi_1, ..., \xi_n$ – вибірка з розподілу зі щільністю

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

Розглянемо оцінки $\hat{\theta}_1 = \max\{\xi_1, ..., \xi_n\}$, $\hat{\theta}_2 = \min\{\xi_1, ..., \xi_n\}$, $\hat{\theta}_3 = \frac{1}{n}\sum_{i=1}^{n}\xi_i$, $\hat{\theta}_4 = \frac{1}{2}(\xi_n + \xi_{n-1})$. Які з оцінок $\hat{\theta}_1, ..., \hat{\theta}_4$ та яких параметрів (можливо, відмінних від a і b) є незсуненими оцінками? Консистентними оцінками?

Розв'язок. 1) $\hat{\theta}_1 = \max\{\xi_1, ..., \xi_n\}$. Знайдемо розподіл оцінки $\hat{\theta}_1$.

$$F_{\hat{\theta}_1}(x) = \mathrm{P}\{\hat{\theta}_1 \le x\} = \mathrm{P}\{\max\{\xi_1, ..., \xi_n\} \le x\} = \\ = \prod_{i=1}^{n} \mathrm{P}\{\xi_i \le x\} = \prod_{i=1}^{n} \int_{-\infty}^{x} f(t)dt = \left(\int_{-\infty}^{x} f(t)dt\right)^n,$$

звідки випливає, що $\hat{\theta}_1$ – абсолютно неперервна випадкова величина зі щільністю

$$f_{\hat{\theta}_1}(x) = \frac{d}{dx}F_{\hat{\theta}_1}(x) = nf(x)\left(\int_{-\infty}^{x} f(t)dt\right)^{n-1} = \\ = \begin{cases} \dfrac{n(x-a)^{n-1}}{(b-a)^n}, & \text{якщо } x \in [a,b]; \\ 0, & \text{якщо } x \notin [a,b]. \end{cases}$$

Тоді математичне сподівання оцінки $\hat{\theta}_1$

$$\mathbf{M}\hat{\theta}_1 = \int_{-\infty}^{+\infty} x f_{\hat{\theta}_1}(x)dx = \frac{a+nb}{n+1}.$$

Отже, $\hat{\theta}_1$ не є незсуненою оцінкою ні параметра a, ні параметра b, але $\mathbf{M}\hat{\theta}_1 \to b$ при $n \to \infty$. Таким чином $\hat{\theta}_1$ є асимптотично незсуненою оцінкою параметра b. З'ясуємо, чи є $\hat{\theta}_1$ консистентною оцінкою параметра b. Для довільного $\varepsilon > 0$

$$\mathrm{P}\{|\hat{\theta}_1 - b| > \varepsilon\} = \int_{-\infty}^{b-\varepsilon} f_{\hat{\theta}_1}(x)dx + \int_{b+\varepsilon}^{+\infty} f_{\hat{\theta}_1}(x)dx = \\ \int_{a}^{b-\varepsilon} f_{\hat{\theta}_1}(x)dx = \left(1 - \frac{\varepsilon}{b-a}\right)^n \underset{n\to\infty}{\to} 0,$$

тобто $\hat{\theta}_1$ є консистентною оцінкою параметра b.

2) $\hat{\theta}_2 = \min\{\xi_1, ..., \xi_n\}$. Ця оцінка досліджується аналогічно попередній. Її щільність дорівнює

$$f_{\hat{\theta}_2}(x) = \begin{cases} \dfrac{n(b-x)^{n-1}}{(b-a)^n}, & \text{якщо } x \in [a,b]; \\ 0, & \text{якщо } x \notin [a,b]. \end{cases}$$

Звідси $\mathbf{M}\hat{\theta}_2 = \dfrac{an+b}{n+1}$, тобто $\hat{\theta}_2$ – асимтотично незсунена оцінка параметра a.

Аналогічно випадку з $\hat{\theta}_1$ перевіряється те, що $\hat{\theta}_2$ є консистентною оцінкою параметра a.

3) $\hat{\theta}_3 = \frac{1}{n}\sum_{i=1}^{n}\xi_i$. Обчислимо математичне сподівання $\hat{\theta}_3$:

$$\mathbf{M}\hat{\theta}_3 = \mathbf{M}\frac{1}{n}\sum_{i=1}^{n}\xi_i = \frac{1}{n}\sum_{i=1}^{n}\mathbf{M}\xi_i = \frac{a+b}{2}.$$

отже $\hat{\theta}_3$ – незсунена оцінка параметра $m = \frac{a+b}{2}$, середини інтервалу. Далі, згідно з законом великих чисел $\hat{\theta}_3$ збігається до m, тобто $\hat{\theta}_3$ є консистентною оцінкою параметра m.

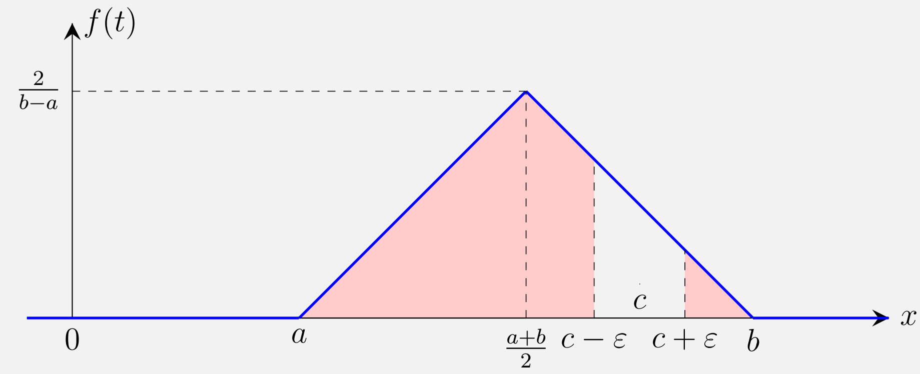
4) $\hat{\theta}_4 = \frac{1}{2}(\xi_n + \xi_{n-1})$. Для цієї оцінки маємо

$$\mathbf{M}\hat{\theta}_4 = \frac{1}{2}(\mathbf{M}\xi_n + \mathbf{M}\xi_{n-1}) = \frac{a+b}{2}.$$

отже $\hat{\theta}_4$ також є незсуненою оцінкою параметра $m = \frac{a+b}{2}$. Щільність $f_{\hat{\theta}_4}(x)$ оцінки $\hat{\theta}_4$ шукаємо як згортку щільностей рівномірних розподілів (див. розділ 11), в результаті одержуємо:

$$f_{\hat{\theta}_4}(x) = \begin{cases} \frac{x-a}{(b-a)^2}, & \text{якщо } x \in \left[a, \frac{a+b}{2}\right]; \\ \frac{4(b-x)}{(b-a)^2}, & \text{якщо } x \in \left[\frac{a+b}{2}, b\right]; \\ 0, & \text{якщо } x \notin [a,b]. \end{cases}$$

Графік щільності $f_{\hat{\theta}_4}(x)$:



Якщо припустити, що $f_{\hat{\theta}_4}(x)$ збігається за ймовірністю до деякої константи c, то для досить малого $\varepsilon > 0$

$$\mathrm{P}\{|\hat{\theta}_4 - c| > \varepsilon\} = \int_{\{x:|x-c|>\varepsilon\}} f_{\hat{\theta}_4}(x)dx$$

є константою, що не залежить від n (заштрихована площа на рисунку), не збігається до 0 при $n \to \infty$і, отже, не є консистентною оцінкою.

Оцінки математичного сподівання та дисперсії. Для вибірки $\xi_1, ..., \xi_n$ з генеральної сукупності ξ_0 позначимо $m = \mathbf{M}\xi$, $\sigma^2 = \mathbf{D}\xi$. Введемо позначення для вибіркових оцінок математичного сподівання та дисперсії:

$$\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n}\xi_i, \quad S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2,$$

$$\hat{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2, \quad \bar{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - m)^2.$$

Теорема 16.2 Нехай $\mathbf{M}\xi_0^2 < \infty$. Тоді:
а) статистика $\bar{\xi}$ є незсуненою та сильно консистентною оцінкою параметра m;
б) статистики $\hat{S}_n^2$, $\bar{S}_n^2$ є незсуненими та сильно консистентними оцінками параметра σ^2;
в) статистика S_n^2 є асимптотично незсуненою та сильно консистентною оцінкою параметра σ^2.

Зауваження (Щодо незсуненості функції від оцінки). *В загальному випадку незсуненість не зберігається при функціональному перетворенні параметра, тобто якщо* $\hat{\theta}_n$ *є незсуненою оцінкою параметра* θ, *то* $f(\hat{\theta}_n)$ *не обов'язково буде незсуненою оцінкою для параметричної функції* $f(\theta)$. *Проте якщо функція лінійна по* θ, $f(\theta) = a\theta + b$, *то легко показати, що* $a\hat{\theta}_n + b$ *є незсуненою оцінкою* $a\theta + b$, *якщо* $\mathbf{M}\hat{\theta}_n = \theta$.

Приклад 16.5

Оцінка $\bar{S}_n^2 = \frac{1}{n}\sum_{i=1}^n (\xi_i - m)^2$ є незсуненою оцінкою дисперсії, параметра σ^2 нормального розподілу $N(m, \sigma^2)$, якщо параметр m відомий. Доведемо, що $\bar{S}_n$ не є незсуненою оцінкою для параметра σ – стандартного відхилення.

Оскільки елементи вибірки $\xi_i \sim N(m, \sigma^2)$, то $(\xi_i - m)/\sigma \sim N(0, 1)$, а випадкова величина $\eta = \sum_{i=1}^n (\xi_i - m)^2/\sigma^2$ в свою чергу має розподіл $\chi^2(n)$ з n ступенями свободи.

Виразивши оцінку $\bar{S}_n = \frac{\sigma}{\sqrt{n}}\sqrt{\eta}$, маючи вираз для щільності $\chi^2(n)$-розподілу, можемо знайти математичне сподівання

$$\mathbf{M}\bar{S}_n = \frac{\sigma}{\sqrt{n}2^{n/2}\Gamma(n/2)} \int_0^{+\infty} \sqrt{x} e^{-x/2} x^{n/2-1} dx =$$

$$= \frac{\sigma\sqrt{2}}{\sqrt{n}\Gamma(n/2)} \int_0^{+\infty} e^{-x} x^{(n-1)/2} dx = \frac{\sigma\sqrt{2}\Gamma((n+1)/2)}{\sqrt{n}\Gamma(n/2)} \neq \sigma.$$

Отже, оцінка $\bar{S}_n$ є зсуненою.

Приклад 16.6

Для вибіркового контролю з партії готової продукції відібрано n приладів. Нехай $\xi_1, ..., \xi_n$ – їхній час роботи до відмови, причому ξ_i – незалежні однаково розподілені випадкові величини, які мають показниковий розподіл з невідомим параметром $\theta > 0$: $F_\theta(x) = 1 - e^{-\theta x}$, $x \geq 0$. Треба оцінити середній час до відмови приладу

$$\tau(\theta) = \mathbf{M}\xi_1 = \frac{1}{\theta}.$$

Згідно теоремі 16.2, вибіркове середнє $\bar{\xi}$ буде незсуненою оцінкою для функції $\tau(\theta)$: $\mathbf{M}\bar{\xi} = \mathbf{M}\xi_1 = \tau(\theta)$. Але якщо спробувати оцінити сам параметр θ за допомогою $\hat{\theta} = 1/\bar{\xi}$, то отримаємо зсунену оцінку.

16.2 Нерівність Рао-Крамера і ефективні оцінки

Розглянемо схему повторних незалежних спостережень над випадковою величиною ξ_0. Нехай $f(x,\theta)$, $\theta \in \Theta$, – щільність розподілу ξ_0, $\xi' = (\xi_1, ..., \xi_n)$ – вибірка з $\mathscr{L}(\xi_0) \in F = \{f(x,\theta),\ \theta \in \Theta\}$ і $x' = (x_1, ..., x_n)$ – реалізація ξ.

Функція

$$L(\xi,\theta) = f(\xi_1,\theta) \times ... \times f(\xi_n,\theta)$$

називається **функцією вірогідності**. Реалізація функції вірогідності

$$L(x,\theta) = f(x_1,\theta) \times ... \times f(x_n,\theta)$$

представляє собою щільність розподілу випадкового вектора ξ.

Далі будемо припускати, що $L(x,\theta)$ диференційована за скалярним параметром θ. Випадкова величина

$$U(\xi;\theta) = \frac{\partial \ln L(\xi;\theta)}{\partial\theta} = \sum_{i=1}^{n} \frac{\partial \ln L(\xi_i;\theta)}{\partial\theta} \tag{16.7}$$

називається **внеском вибірки** ξ, а i-ий доданок $\dfrac{\partial \ln L(\xi_i;\theta)}{\partial\theta}$ називається **внеском** i**-ого спостереження**. Будемо вважати, що

$$0 < \mathbf{M}_\theta U^2(\xi,\theta) < \infty \quad \text{для усіх } \theta \in \Theta.$$

У подальшому нам прийдеться диференціювати по θ інтеграли від функцій на вибірковому просторі X і міняти місцями порядок інтегрування і диференціювання. Моделі, для яких ця операція коректна, називають коротко **регулярними**. Таким чином, для регулярної моделі

$$\mathbf{M}_\theta U(\xi,\theta) = 0 \quad \text{для усіх } \theta \in \Theta.$$

Приклад 16.7 Нерегулярні моделі

Точні аналітичні умови, які забезпечують регулярність моделі відомі з математичного аналізу і вид їх визначається у кожному конкретному випадку. При оцінюванні параметрів обов'язкова умова регулярності моделі полягає у тому, що вибірковий простір $\mathbb{X}$ не повинен залежати від невідомого параметра θ. Зокрема, рівномірний розподіл зазвичай не володіє умовою регулярності.

1) Нехай ξ_0 має рівномірний розподіл на $(0,\theta)$. Щільність її розподілу

$$f_{\xi_0}(x,\theta) = \begin{cases} 1/\theta, & x \in (0,\theta); \\ 0, & x \notin (0,\theta). \end{cases}$$

При диференціюванні тотожності $\int_0^\theta \frac{1}{\theta}dx \equiv 1$ не випливає, що $\int_0^\theta \frac{\partial}{\partial\theta}\left(\frac{1}{\theta}\right)dx = 0$, оскільки при диференціюванні інтеграла по верхні границі з'явиться ще один доданок.

Отже, в даному випадку причиною нерегулярності є те, що вибірковий простір залежить від невідомого параметра θ.

2) Прикладом ще одного нерегулярного сімейства є зсунений показниковий розподіл з параметром зсуву θ, який задається щільністю

$$f_{\xi_0}(x,\theta) = \begin{cases} e^{\theta - x}, & x > \theta; \\ 0, & x \le \theta. \end{cases}$$

Тут також вибірковий простір залежить від невідомого параметра, отже модель не є регулярною.

Функцію

$$I_n(\theta) = \mathbf{D}_\theta U(\xi;\theta) = \mathbf{M}_\theta U^2(\xi;\theta)$$

називають **функцією інформації Фішера** про параметр θ, яка міститься у виборці ξ. Величина

$$I_1(\theta) = I(\theta) = \mathbf{M}\left(\frac{\partial \ln f(\xi_1;\theta)}{\partial\theta}\right)^2$$

– це кількість фішерівської інформації, яка міститься в одному спостереженні. Для моделі повторних незалежних спостережень

$$I_n(\theta) = nI(\theta).$$

Якщо функція $f(x,\theta)$ двічі диференційована по θ, то

$$I(\theta) = -\mathbf{M}_\theta\left(\frac{\partial^2 \ln f(\xi_1,\theta)}{\partial\theta^2}\right). \qquad (16.8)$$

Розглянемо тепер задачу оцінювання заданої параметричної функції $\tau(\theta)$ у моделі $F = \{F(x,\theta),\ \theta \in \Theta\}$. Припустимо, що модель F регулярна, функція $\tau(\theta)$ диференційована і нехай $\Im_\tau$ – клас усіх незсунених оцінок $\tau(\theta)$, дисперсія яких скінченна. Тоді має місце таке твердження.

Теорема 16.3 (Критерій Рао-Крамера) Для довільної оцінки $T = T(\xi) \in \Im_\tau$ справедлива нерівність

$$\mathbf{D}_\theta T \geq \frac{\left[\tau'(\theta)\right]^2}{nI(\theta)}. \tag{16.9}$$

Рівність у формулі (16.9) має місце тоді і тільки тоді, коли T – лінійна функція внеску вибірки:

$$T(\xi) - \tau(\theta) = a(\theta)\, U(\xi;\theta), \tag{16.10}$$

де $a(\theta)$ – деяка функція від θ. Нерівність (16.9) називається **нерівністю Рао-Крамера**. Вона визначає нижню границю дисперсій усіх незсунених оцінок заданої параметричної функції $\tau(\theta)$ для регулярної моделі.

Якщо існує оцінка $T^* \in \Im_\tau$, для якої нижня границя Рао-Крамера досягається, то її називають **ефективною**. Ефективна оцінка є оптимальною і згідно раніше доведеної теореми вона єдина. Критерієм ефективності оцінки є подання (16.10). Будемо називати цей критерій оптимальності оцінки **критерієм Рао-Крамера**.

16.3 Оцінки максимальної вірогідності

Нехай задана вибірка $\xi' = (\xi_1, \xi_2, ..., \xi_n)$ з розподілу $\mathscr{L}(\xi_0) \in F \in \{F(x,\theta),\ \theta \in \Theta\}$ і $L(\xi,\theta) = f(\xi_1,\theta) \times ... \times f(\xi_n,\theta)$ – функція вірогідності, де $f(x,\theta)$ – щільність розподілу ξ_0 для абсолютно неперервної моделі F і $f(x,\theta) = P_\theta(\xi_0 = x)$ для дискретної моделі.

Оцінкою максимальної вірогідності (о.м.в.) $\hat\theta = \hat\theta(\xi)$ параметра θ називається точка параметричної множини Θ, в якій функція вірогідності досягає максимуму.

Якщо $\hat\theta = \hat\theta(x)$ - реалізація о.м.в., то

$$L\left(x,\hat\theta\right) \geq L(x,\theta) \text{ для усіх } \theta \in \Theta \text{ або } L\left(x,\hat\theta\right) = \sup_{\theta\in\Theta} L(x,\theta).$$

Розглянемо спочатку випадок, коли $\theta \in \Theta \subseteq R^1$ – одновимірний параметр. При умові, що для кожного $x \in X$ максимум $L(x,\theta)$ досягається у внутрішній точці Θ і $L(x,\theta)$ диференційована по θ, то о.м.в. $\hat\theta$ задовольняє рівнянню

$$\frac{\partial L(x,\theta)}{\partial\theta} = 0 \quad \text{або} \quad \frac{\partial \ln L(x,\theta)}{\partial\theta} = 0. \tag{16.11}$$

Якщо θ – векторний параметр, $\theta' = (\theta_1, ..., \theta_r)$, то рівняння (16.11) замінюється на систему рівнянь

$$\frac{\partial \ln L(x,\theta)}{\partial\theta_i} = 0, \quad i = 1, 2, ..., r. \tag{16.12}$$

Рівняння (16.11) або (16.12) називаються **рівняннями вірогідності**.

16.4 Оцінки методу моментів

Нехай $\xi' = (\xi_1, \xi_2, ..., \xi_n)$ – вибірка з розподілу $\mathscr{L}(\xi_0) \in F = \{F(x, \theta), \ \theta \in \Theta\}$, де $\theta' = (\theta_1, \theta_2, ..., \theta_r)$ і множина $\Theta \subseteq R^r$. Для випадкової величини ξ_0, що спостерігається, існують перші r моментів $a_k = \mathbf{M}_\theta \xi_0^k, \quad k = 1, 2, ..., r$. Ці моменти є функціями від невідомих параметрів $a_k = a_k(\theta) = a_k(\theta_1, \theta_2, ..., \theta_r)$.

Розглянемо відображення

$$a_k(\theta) = y_k, \quad k = 1, 2, ..., r, \tag{16.13}$$

$a : \Theta \to Y = \{(y_1, ..., y_r) \ : \ a_k(\theta) = y_k, \quad k = 1, 2, ..., r, \text{ для деякого } \theta\}$. Будемо вважати, що це відображення взаємно однозначне і взаємно неперервне. **Метод моментів** полягає у прирівнюванні значень теоретичних і вибіркових моментів (точніше реалізацій вибіркових моментів)

$$a_k(\theta) = A_{nk}(x), \quad k = 1, 2, ..., r, \quad \text{ де } A_{nk}(x) = \frac{1}{n}\sum_{i=1}^{n} x_i^k. \tag{16.14}$$

Таким чином, ми підставляємо в (16.14) $A_{nk}(x)$ замість y_k.

В силу зроблених припущень існує неперервна функція

$$a^{-1}(y) = \left(a_1^{-1}(y), a_2^{-1}(y), ..., a_r^{-1}(y)\right).$$

Методом моментів ми отримаємо наступні реалізації для оцінок параметрів $\theta_1, \theta_2, ..., \theta_r$

$$\hat{\theta}_i = \hat{\theta}_i(x_1, ..., x_n) = a_i^{-1}(A_{n1}(x), ..., A_{nr}(x)), \quad i = 1, 2, ..., r \tag{16.15}$$

Теорема 16.4 Оцінки параметрів $\theta_1, \theta_2, ..., \theta_r$ методу моментів у випадку, коли система (16.13) встановлює взаємно однозначне і взаємно неперервне відображення, є консистентними.

Відзначимо, що метод моментів не можна застосувати, коли теоретичні моменти необхідного порядку не існують. Окрім цього, оцінки методу моментів, взагалі кажучи, неефективні.

Приклад 16.8

Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю, яка приймає ненульові значення на відрізку $[2, \theta]$ і має на цьому відрізку вид "рівнобедреного трикутника". Оцінити параметр θ методом моментів.

Розв'язок: зважаючи на опис щільності з умови та те, що повний інтеграл від щільності має бути рівним 1, нескладно зрозуміти, що щільність формально можна записати так:

$$f(x;\theta) = \begin{cases} \frac{4(x-2)}{(\theta-2)^2}, & x \in \left[2, \frac{2+\theta}{2}\right], \\ \frac{4(\theta-x)}{(\theta-2)^2}, & x \in \left(\frac{2+\theta}{2}, \theta\right], \\ 0, & x \notin [2, \theta]. \end{cases}$$

Для застосування методу моментів обчислимо математичне сподівання

$$\mathbf{M}\xi = \int_{-\infty}^{\infty} xf(x;\theta)dx \ = \int_2^{\theta} xf(x;\theta)dx \ = \int_2^{\frac{2+\theta}{2}} x\frac{4(x-2)}{(\theta-2)^2}dx + \int_{\frac{2+\theta}{2}}^{\theta} x\frac{4(\theta-x)}{(\theta-2)^2}dx \ = \frac{2+\theta}{2}$$

і прирівняємо його до вибіркового середнього:

$$\bar{\xi} = \frac{2+\theta}{2},$$

звідки отримуємо

$$\hat{\theta} = 2\bar{\xi} - 2.$$

Відповідь: $\hat{\theta} = 2\bar{\xi} - 2$..

Примітка: математичне сподівання можна було знайти простіше, навіть не формалізуючи щільність, якщо зважити на те, що за умовою графік щільності симетричний відносно прямої $x = \frac{2+\theta}{2}$. Подумайте, як шукати в цій же задачі оцінку методом максимальної вірогідності.

Приклад 16.9

Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з гауссівського розподілу із щільністю

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\dfrac{(x-\theta)^2}{2\sigma^2}}.$$

Припустимо, що значення σ є відомим, а параметр θ є невідомим. Оцінити параметр θ

методом максимальної вірогідності.

Розв'язок: для зазначеної щільності функція вірогідності матиме вид

$$L(\theta)=\prod_{i=1}^{n} f\left(x_i;\theta\right)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left\{-\sum_{i=1}^{n}\frac{\left(x_i-\theta\right)^2}{2\sigma^2}\right\}$$

і тоді логарифм від неї

$$\ln L(\theta)=\ln\left(\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left\{-\sum_{i=1}^{n}\frac{\left(x_i-\theta\right)^2}{2\sigma^2}\right\}\right)=\ln\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n-\sum_{i=1}^{n}\frac{\left(x_i-\theta\right)^2}{2\sigma^2},$$

а похідна від цього логарифма

$$\frac{\partial}{\partial\theta}\ln L(\theta)=\frac{\partial}{\partial\theta}\left(\ln\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n-\sum_{i=1}^{n}\frac{\left(x_i-\theta\right)^2}{2\sigma^2}\right)=\left(0-\sum_{i=1}^{n}\frac{-(x_i-\theta)}{\sigma^2}\right)=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\theta)$$

і ця похідна буде дорівнювати нулю тільки коли

$$\sum_{i=1}^{n}(x_i-\theta)=0\Leftrightarrow\sum_{i=1}^{n}x_i-n\theta=0\Leftrightarrow\sum_{i=1}^{n}x_i=n\theta\Leftrightarrow\theta=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}.$$

Відповідь: $\hat{\theta}=\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i=\bar{\xi}$.

Примітка: очевидно, що цю ж оцінку ми б отримали і методом моментів, причому набагато простіше.

Задача 16.1. Нехай $\xi_1,\xi_2,...,\xi_n$ – вибірка з розподілу із щільністю

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

Розглядаються оцінки

$$\hat{\theta}_1=\min\{\xi_1,\xi_2,...,\xi_n\},\qquad \hat{\theta}_2=\max\{\xi_1,\xi_2,...,\xi_n\},$$

$$\hat{\theta}_3=\frac{1}{n}\sum_{i=1}^{n}\xi_i,\qquad \hat{\theta}_4=\frac{\xi_{n-1}+\xi_n}{2}.$$

Які з оцінок $\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3,\hat{\theta}_4$ та для яких параметрів (можливо, відмінних від a і b) є незсуненими? Конзистентними оцінками?

Задача 16.2. Нехай $\xi_1,\xi_2,...,\xi_n$ – вибірка з нормального розподілу з параметрами $(\theta,1)$. Чи є $\hat{\theta}=\bar{\xi}=\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ ефективною оцінкою параметра θ?

Задача 16.3. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з пуассонівського розподілу з параметром λ. Чи є $\hat{\lambda} = \bar{\xi}$ ефективною оцінкою параметра λ?

Задача 16.4. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з нормального розподілу з параметрами $(0, \sigma^2)$. Чи є $\hat{\sigma}^2 = \frac{1}{n} \sum\limits_{i=1}^{n} \xi_i^2$ ефективною оцінкою параметра σ^2?

Задача 16.5. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу зі щільністю

$$f(x; \theta) = \begin{cases} 0, & \text{якщо } x < \theta, \\ \exp\{\theta - x\}, & \text{якщо } x \geq \theta. \end{cases}$$

Чи є $\hat{\theta} = \frac{1}{n} + \min\{\xi_i\}$ незсуненою оцінкою параметра θ? Конзистентною оцінкою θ?

Задача 16.6. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу

$$\mathrm{P}(k; \theta) = \frac{1}{1+\theta} \left(\frac{\theta}{1+\theta} \right)^k, \quad k = 0, 1, ...$$

Чи є $\hat{\theta} = \frac{1}{n} \sum\limits_{i=1}^{n} \xi_i$ ефективною оцінкою параметра θ?

Задача 16.7. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу Ерланга з параметрами m, θ (m – відоме), тобто має щільність

$$f(x; m, \theta) = \begin{cases} 0, & \text{якщо } x < 0, \\ \frac{1}{\theta^m (m-1)!} x^{m-1} \exp\left\{ -\frac{x}{\theta} \right\}, & \text{якщо } x \geq 0. \end{cases}$$

Чи є $\hat{\theta} = \frac{1}{nm} \sum\limits_{i=1}^{n} \xi_i$ ефективною оцінкою параметра θ?

Задача 16.8. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x; m, h) = \begin{cases} \dfrac{1}{2h}, & \text{якщо } x \in [\theta - h, \theta + h], \\ 0, & \text{якщо } x \notin [\theta - h, \theta + h]. \end{cases}$$

Розглядаються оцінки

$$\hat{\theta}_1 = \frac{1}{2}(\max\{\xi_i\} - \min\{\xi_i\}), \qquad \hat{\theta}_2 = \frac{1}{2}(\min\{\xi_i\} + \max\{\xi_i\})$$

$$\hat{\theta}_3 = \min\{\xi_i\} - \frac{1}{n-1}(\max\{\xi_i\} - \min\{\xi_i\}),$$

$$\hat{\theta}_4 = \max\{\xi_i\} + \frac{1}{n-1}(\max\{\xi_i\} - \min\{\xi_i\}).$$

Визначити серед $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4$ незсунені, конзистентні оцінки параметрів θ та h. Можливо, серед них є незсунені, конзистентні оцінки інших параметрів. Яких саме?

Задача 16.9. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x; m, \theta) = \begin{cases} 0, & \text{якщо } x < m, \\ \frac{1}{\theta} \exp\left\{-\frac{1}{\theta}(x - m)\right\}, & \text{якщо } x \geq m. \end{cases}$$

$\theta > 0$, m – відоме. Чи є $\hat{\theta} = \bar{\xi} - m$ ефективною оцінкою параметра θ?

Задача 16.10. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x; \alpha) = \begin{cases} 0, & \text{якщо } x < 0, \\ \frac{1}{\alpha} \exp\left\{-\frac{1}{\alpha}x\right\}, & \text{якщо } x \geq 0. \end{cases}$$

Довести, що $\hat{\theta}_1 = \bar{\xi}$ є незсуненою та консистентною оцінкою параметра α. Чи є $\hat{\theta}_2 = \frac{1}{2}(\xi_{n-1} + \xi_n)$ незсуненою оцінкою параметра α? Консистентною оцінкою цього ж параметра?

Задача 16.11. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із біноміального розподілу

$$\mathrm{P}(k; \theta) = C_N^k \theta^k (1 - \theta)^{N-k}, \quad k = 0, 1, ..., N,$$

де N – відоме ціле число. Чи є $\hat{\theta} = \dfrac{1}{Nn} \sum\limits_{i=1}^{n} \xi_i$ ефективною, консистентною оцінкою параметра θ?

Задача 16.12. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x; b, \theta) = \begin{cases} 0, & \text{якщо } x < b, \\ \frac{1}{\theta} \exp\left\{-\frac{1}{\theta}(x - b)\right\}, & \text{якщо } x \geq b. \end{cases}$$

Розглядаються оцінки

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^{n} \xi_i, \qquad \hat{\theta}_2 = \min\{\xi_i\}, \qquad \hat{\theta}_3 = \hat{\theta}_1 - \hat{\theta}_2.$$

Чи є серед $\hat{\theta}_1$, $\hat{\theta}_2, \hat{\theta}_3$ незсунені, консистентні оцінки параметрів θ та b? Можливо, серед них є незсунені та консистентні оцінки інших параметрів? Яких саме?

Задача 16.13. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу **Релея**, тобто із щільністю

$$f(x; \theta) = \begin{cases} \frac{x}{\theta} \exp\left\{-\frac{x^2}{2\theta}\right\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \leq 0. \end{cases}$$

Чи є $\hat{\theta} = \frac{1}{2n} \sum\limits_{i=1}^{n} \xi_i^2$ ефективною оцінкою параметра θ?

Задача 16.14. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із **логарифмічно нормального розподілу** із щільністю

$$f(x; \mu) = \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma_0^2}x} \exp\left\{-\dfrac{(\ln x - \mu)^2}{2\sigma_0^2}\right\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \leq 0. \end{cases}$$

$\sigma_0 > 0$ – відоме. Чи є $\hat{\theta} = \frac{1}{n}\sum\limits_{i=1}^{n} \ln \xi_i$ ефективною оцінкою параметра μ?

Задача 16.15. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу Парето зі щільністю

$$f(x;\lambda) = \begin{cases} \frac{\theta\lambda^{\theta}}{x^{\theta+1}}, & \text{якщо } x > \lambda, \\ 0, & \text{якщо } x \le \lambda. \end{cases}$$

$\lambda > 0$, $\theta > 2$, θ – відоме. Розглядаються оцінки

$$\hat{\lambda}_1 = \frac{\theta - 1}{\theta n}\sum_{i=1}^{n} \xi_i, \qquad \hat{\lambda}_2 = \frac{\theta - 2}{\theta n}\sum_{i=1}^{n} \xi_i^2.$$

Незсуненими і консистентними оцінками яких параметрів є оцінки $\hat{\lambda}_1$, $\hat{\lambda}_2$?

Задача 16.16. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;b,\theta) = \begin{cases} \frac{1}{\theta^{\nu}\Gamma(\nu)} x^{\nu-1}\exp\{-\frac{x}{\theta}\}, & \text{якщо } x > 0, \\ \frac{1}{\theta}\exp\left\{-\frac{1}{\theta}(x-b)\right\}, & \text{якщо } x \le 0. \end{cases}$$

ν – відомий параметр. Чи є $\hat{\theta} = \dfrac{1}{n\nu}\sum\limits_{i=1}^{n} \xi_i$ ефективною оцінкою параметра θ?

Задача 16.17. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з рівномірного на відрізку $[a,\ a+1]$ розподілу. Розглядаються дві оцінки параметра a :

$$\hat{a}_1 = \frac{1}{n}\sum_{i=1}^{n} \xi_i - \frac{1}{2}, \qquad \hat{a}_2 = \max\{\xi_i\} - \frac{n}{n+1}.$$

Довести, що $\hat{a}_1$ та $\hat{a}_2$ є незсуненими оцінками параметра a. Знайти $\mathbf{D}\hat{a}_1$, $\mathbf{D}\hat{a}_2$. Довести, що $\mathbf{D}\hat{a}_2 = o(\mathbf{D}\hat{a}_1)$, коли $n \to \infty$.

Задача 16.18. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;\theta) = \frac{1}{2\theta}\exp\left\{-\frac{1}{\theta}|x|\right\}, \quad \theta > 0.$$

Чи є $\hat{\theta} = \frac{1}{n}\sum\limits_{i=1}^{n} |\xi_i|$ ефективною оцінкою параметра θ?

Задача 16.19. Знайти оцінки параметрів a і σ^2 нормального розподілу методом моментів.

Задача 16.20. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;\theta) = \begin{cases} \theta\exp\{-\theta x\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \le 0. \end{cases}$$

знайти оцінку максимальної вірогідності параметра θ.

Задача 16.21. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із логарифмічно нормального розподілу із щільністю

$$f(x;\mu,\sigma^2) = \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma^2}x} \exp\left\{-\frac{(\ln x-\mu)^2}{2\sigma^2}\right\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \le 0. \end{cases}$$

Знайти оцінку параметрів μ та σ^2методом моментів.

Задача 16.22. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;\theta) = \begin{cases} \dfrac{1}{2h_0}, & \text{якщо } x \in [\theta - h_0, \theta + h_0], \\ 0, & \text{якщо } x \notin [\theta - h_0, \theta + h_0]. \end{cases}$$

Знайти оцінку $\hat{\theta}$ параметра θ методом максимальної вірогідності. З'ясувати, чи буде $\hat{\theta}$ незсуненою та консистентною оцінкою параметра θ.

Задача 16.23. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;\theta,b) = \frac{1}{2\sqrt{\theta}} \exp\left\{-\frac{1}{\sqrt{\theta}}|x-b|\right\}, \quad \theta > 0.$$

Знайти оцінки $\hat{\theta}$ та $\hat{b}$ відповідно параметрів θ та b методом моментів. З'ясувати, чи є $\hat{\theta}$ та $\hat{b}$ незсуненими і консистентними оцінками параметрів θ та b.

Задача 16.24. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$\mathrm{P}(k;\theta) = C_{r-1+k}^{r-1}\left(\frac{1}{1+\theta}\right)^r \left(\frac{\theta}{1+\theta}\right)^k, \; k = 0, 1, ...,$$

де $\theta > 0$, r – відоме. Знайти оцінку параметра θ методом максимальної вірогідності. З'ясувати, чи є оцінка $\hat{\theta}$ незсуненою, консистентною, ефективною оцінкою параметра θ.

Задача 16.25. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із геометричного розподілу із параметром p:

$$\mathrm{P}(k;p) = (1-p)^k p, \;\; k = 0, 1, ...$$

Оцінити параметр p методом моментів.

Задача 16.26. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;a,b) = \frac{1}{2a} \exp\left\{-\frac{1}{a}|x-b|\right\}, \qquad a > 0.$$

Знайти оцінки параметрів a та b методом максимальної вірогідності.

Задача 16.27. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;a,b) = \begin{cases} (x - (a - \sqrt{b}))/b, & \text{якщо } x \in [a - \sqrt{b}, a], \\ -(x - (a + \sqrt{b}))/b, & \text{якщо } x \in [a, a + \sqrt{b}], \\ 0, & \text{якщо } x \notin [a - \sqrt{b}, a + \sqrt{b}]. \end{cases}$$

Знайти оцінки $\hat{a}$ та $\hat{b}$ параметрів a та b методом моментів. З'ясувати, чи є $\hat{a}$ і $\hat{b}$ незсуненими і консистентними оцінками параметрів a та b відповідно.

Задача 16.28. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;\theta) = \frac{1}{2\theta} \exp\left\{-\frac{1}{\theta}|x|\right\}, \quad \theta > 0.$$

Знайти оцінку дисперсії методом моментів і з'ясувати, чи є вона незсуненою і консистентною оцінкою дисперсії.

Задача 16.29. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x;h) = \begin{cases} \dfrac{1}{2h}, & \text{якщо } x \in [\theta_0 - h, \theta_0 + h], \\ 0, & \text{якщо } x \notin [\theta_0 - h, \theta_0 + h]. \end{cases}$$

Знайти оцінку параметра h методом максимальної вірогідності. З'ясувати, чи буде ця оцінка незсуненою та консистентною оцінкою параметра h.

Задача 16.30. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу **Вейбулла** із щільністю

$$f(x;\theta) = \begin{cases} \dfrac{p}{\theta^p} x^{p-1} \exp\left\{-\left(\dfrac{x}{\theta}\right)^p\right\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \le 0. \end{cases}$$

p – відомий параметр. Знайти оцінку параметра θ методом максимальної вірогідності.

Задача 16.31. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з рівномірного на відрізку $\left[\theta - \sigma\sqrt{3}; \theta + \sigma\sqrt{3}\right]$ розподілу, $\sigma > 0$. Знайти оцінки $\hat{\theta}$ та $\hat{\sigma}$ параметрів θ та σ відповідно методом моментів. З'ясувати, чи є $\hat{\theta}$ та $\hat{\sigma}$ незсуненими і консистентними оцінками параметрів θ та σ.

Задача 16.32. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу Релея, тобто із щільністю

$$f(x;\theta) = \begin{cases} \dfrac{x}{\theta} \exp\left\{-\dfrac{x^2}{2\theta}\right\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \le 0. \end{cases}$$

Знайти оцінку параметра θ методом максимальної вірогідності. З'ясувати, чи є оцінка $\hat{\theta}$ незсуненою, консистентною, ефективною оцінкою параметра θ.

Задача 16.33. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу **Ерланга** з параметрами m і λ (λ – відоме). Оцінити параметр m методом моментів. Щільність розподілу Ерланга з параметрами m і λ має вигляд

$$f(x;\lambda,m) = \begin{cases} 0, & \text{якщо } x < 0, \\ \dfrac{\lambda^m}{(m-1)!} x^{m-1} \exp\{-\lambda x\}, & \text{якщо } x \ge 0. \end{cases}$$

Задача 16.34. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x; a, b) = \begin{cases} 0, & \text{якщо } x < b, \\ \dfrac{1}{a} \exp\left\{-\dfrac{1}{a}(x-b)\right\}, & \text{якщо } x \geq b. \end{cases}$$

Знайти оцінки параметрів a і b методом максимальної вірогідності. З'ясувати, чи будуть ці оцінки незсуненими та конзистентними оцінками параметрів a і b.

Задача 16.35. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із **гамма-розподілу** із щільністю

$$f(x; \nu, \alpha) = \begin{cases} \dfrac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} \exp\{-\alpha x\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \leq 0. \end{cases}$$

Знайти оцінки параметрів ν та α методом моментів.

Задача 16.36. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу

$$\mathrm{P}(k; \theta) = \frac{1}{1+\theta}\left(\frac{\theta}{1+\theta}\right)^k, \quad k = 0, 1, ..., \theta > 0.$$

Знайти оцінку параметра θ методом максимальної вірогідності. З'ясувати, чи є оцінка $\hat{\theta}$ незсуненою, конзистентною, ефективною оцінкою параметра θ.

Задача 16.37. Позначимо через ξ випадкову величину – число невдач до появи r-го успіху в необмеженій послідовності незалежних випробувань Бернуллі з ймовірністю успіху p в одному випробуванні. Випадкова величина ξ має розподіл

$$\mathrm{P}(\xi = k) = C_{r-1+k}^{r-1} p^r (1-p)^k, \quad k = 0, 1, ...$$

Цей розподіл називають від'ємним біноміальним з параметрами r, p або розподілом Паскаля. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із від'ємного біноміального розподілу з параметрами r, p; r – відомо. Оцінити параметр p методом моментів. (Вказівка. Загальну кількість невдач до r-го успіху можна подати у вигляді суми $\xi = \eta_1 + \eta_2 + ... + \eta_r$, де $\eta_1, \eta_2, ...\eta_r$ – незалежні випадкові величини, кожна з яких має геометричний розподіл з параметром p).

Задача 16.38. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x; a, b) = \begin{cases} 0, & \text{якщо } x < b, \\ \dfrac{1}{a} \exp\left\{-\dfrac{1}{a}(x-b)\right\}, & \text{якщо } x \geq b. \end{cases}$$

Знайти оцінки параметрів a та b методом моментів.

Задача 16.39. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу із щільністю

$$f(x; a, b) = \begin{cases} \dfrac{1}{b-a}, & \text{якщо } x \in [a, b], \\ 0, & \text{якщо } x \notin [a, b]. \end{cases}$$

Знайти оцінки параметрів a та b методом максимальної вірогідності. З'ясувати, чи будуть ці оцінки незсуненими та конзистентними оцінками відповідних параметрів.

Задача 16.40. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із біноміального розподілу

$$\mathrm{P}(k;p) = C_m^k p^k (1-p)^{m-k}, \quad k = 0, 1, ..., m,$$

де m – відоме. Знайти оцінку параметра p методом моментів. З'ясувати, чи є оцінка $\hat{p}$ незсуненою, конзистентною, ефективною оцінкою параметра p.

Задача 16.41. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із **логарифмічно нормального** розподілу із щільністю

$$f(x;\mu,\sigma^2) = \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma^2}x} \exp\left\{-\dfrac{(\ln x - \mu)^2}{2\sigma^2}\right\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \le 0. \end{cases}$$

Знайти оцінки параметрів μ та σ^2 методом максимальної вірогідності.

Задача 16.42. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу Пуассона

$$\mathrm{P}(k;\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, ...$$

Знайти оцінку параметра λ методом максимальної вірогідності. З'ясувати, чи є оцінка $\hat{\lambda}$ незсуненою, конзистентною, ефективною оцінкою параметра λ.

Задача 16.43. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка з розподілу

$$\mathrm{P}(k;\theta) = C_{r-1+k}^{r-1} \left(\frac{1}{1+\theta}\right)^r \left(\frac{\theta}{1+\theta}\right)^k, \quad k = 0, 1, ...,$$

де $\theta > 0$, r – відоме. Знайти оцінку параметра θ методом моментів. З'ясувати, чи є оцінка $\hat{\theta}$ незсуненою, конзистентною, ефективною оцінкою параметра θ.

Задача 16.44. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із біноміального розподілу

$$\mathrm{P}(k;p) = C_m^k p^k (1-p)^{m-k}, \quad k = 0, 1, ..., m,$$

де m – відоме. Знайти оцінку параметра p методом максимальної вірогідності. З'ясувати, чи є оцінка $\hat{p}$ незсуненою, конзистентною, ефективною оцінкою параметра p.

Задача 16.45. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу Релея, тобто із щільністю

$$f(x;\theta) = \begin{cases} \dfrac{x}{\theta} \exp\left\{-\frac{x^2}{2\theta}\right\}, & \text{якщо } x > 0, \\ 0, & \text{якщо } x \le 0. \end{cases}$$

Знайти оцінку параметра θ методом моментів. З'ясувати, чи є оцінка $\hat{\theta}$ незсуненою, конзистентною оцінкою параметра θ.

Задача 16.46. Нехай $\xi_1, \xi_2, ..., \xi_n$ – вибірка із розподілу Парето із щільністю

$$f(x; \theta, \lambda) = \begin{cases} \dfrac{\theta \lambda^\theta}{x^{\theta+1}}, & \text{якщо } x > \lambda, \\ 0, & \text{якщо } x \le \lambda. \end{cases}$$

$\lambda > 0,\ \theta > 0$. Знайти оцінки параметрів θ та λ методом максимальної вірогідності.

17 ІНТЕРВАЛЬНЕ ОЦІНЮВАННЯ

17.1 Визначення надійного інтервалу

Раніше ми розглядали точкові оцінки параметра θ у моделі

$$\mathscr{F} = \{F(x, \theta), \theta \in \Theta\}.$$

Довільна точкова оцінка представляє собою функцію $T = T(\xi)$ вибірки $\xi' = (\xi_1, ..., \xi_n)$ і при кожній реалізації x вибірки ξ ця функція визначає єдине значення параметра θ.

Можна задачу оцінювання поставити інакше: необхідно вказати такий інтервал, в середині якого з високою ймовірністю γ знаходиться значення параметра, що оцінюється. Ймовірність γ відбиває ступінь готовності миритися з можливістю похибки. При заданому γ довжина надійного інтервалу характеризує точність локалізації параметра, тому бажано обирати найменший інтервал.

При інтервальному оцінюванні шукають такі дві статистики

$$T_1 = T_1(\xi) \text{ і } T_2 = T_2(\xi), \quad T_1 < T_2,$$

для яких при заданому γ виконується умова

$$\mathrm{P}_\theta \{T_1(\xi) < \theta < T_2(\xi)\} \geq \gamma \text{ для усіх } \theta \in \Theta. \tag{17.1}$$

Інтервал $(T_1(\xi), T_2(\xi))$ називають **γ-надійним інтервалом для θ**, γ – **рівень надійності**, T_1, T_2 - **нижня і верхня границя надійності.**

Таким чином γ- надійний інтервал – це випадковий інтервал в параметричній множині Θ $((T_1, T_2) \subset \Theta)$, який залежить тільки від вибірки ξ (не від θ) і який накриває значення невідомого параметра θ з імовірністю, яка не менше γ.

Приклад 17.1

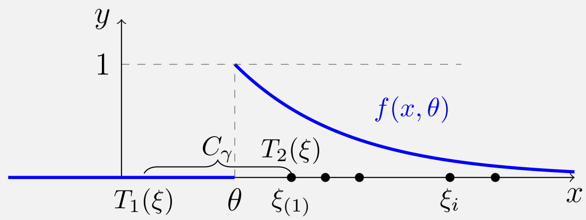
Розглянемо модель зсуненого показникового закону розподілу з щільністю

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

Носієм розподілу є проміжок $[\theta, +\infty)$, тому, очевидно, елементи вибірки не можуть бути меншими за справжнє значення параметру θ. Отже, доцільно в якості верхньої межі $T_2(\xi)$ надійного інтервалу для цього параметру взяти мінімальний елемент вибірки $\xi_{(1)}$ (зауважимо, що $\hat{\theta}_n = \xi_{(1)} = \min\{\xi_1, ..., \xi_n\}$ є оцінкою максимальної вірогідності параметру θ, див. [12], приклад 2.18, ст. 62). Матимемо $\theta \le \xi_{(1)}$.

Тоді ліва межа надійного інтервалу з рівнем довіри γ знаходиться в точці $T_1(\xi) = T_2(\xi) - C_\gamma = \xi_{(1)} - C_\gamma$, де C_γ – деяка стала. Її значення можна знайти з визначення надійного інтервалу: для будь-яких значень $\theta \in \Theta$

$$\mathrm{P}\{\xi_{(1)} - C_\gamma \le \theta \le \xi_{(1)}\} = \mathrm{P}\{\xi_{(1)} - C_\gamma \le \theta\} = \gamma.$$



Зважаючи на те, що елементи вибірки є незалежними і мають зсунений показниковий розподіл, то величини $\xi_i - \theta$ є незалежними і показниково розподіленими випадковими величинами. Відповідно, останню рівність можемо записати таким чином:

$$1-\gamma = \mathrm{P}\{\xi_{(1)}-\theta \ge C_\gamma\} = \mathrm{P}\{\min\{\xi_1, ..., \xi_n\}-\theta \ge C_\gamma\} = \mathrm{P}\{\xi_1-\theta \ge C_\gamma, ..., \xi_n-\theta \ge C_\gamma\} =$$
$$= \prod_{i=1}^{n} \mathrm{P}\{\xi_i - \theta \ge C_\gamma\} = \exp\{-nC_\gamma\}.$$

Звідси маємо значення довжини надійного інтервалу $C_\gamma = (-\ln(1-\gamma))/n$. Отже, остаточно, для справжнього значення параметра θ маємо надійний інтервал рівня γ :

$$\xi_{(1)} + \ln(1-\gamma))/n \le \theta \le \xi_{(1)}.$$

Зауважимо, що довжина надійного інтервалу $C_\gamma \to \infty$ при $1-\gamma \to 0$, і $C_\gamma \to 0$ при $n \to \infty$.

Якщо ймовірність у лівій частині нерівності (17.1) прямує до γ при $n \to \infty$, то інтервал $(T_1(\xi), T_2(\xi))$ називається **асимптотичним надійним інтервалом.** Зазвичай довжина надійного інтервалу зростає при збільшенні коефіцієнта надійності γ та прямує до нуля зі збільшенням розміру вибірки n.

17.2 Побудова надійного інтервалу за допомогою центральної статистики

Нехай $\mathscr{F} = \{F(z,\theta), \theta \in \Theta\}$ - абсолютно неперервна модель і існує випадкова величина $G(\xi, \theta)$, яка залежить від θ така, що

1) розподіл $G(\xi, \theta)$ не залежить від θ;

2) при кожному $x \in X$ функція $G(\xi, \theta)$ неперервна і строго монотонна по θ.

Таку випадкову величину $G(\xi, \theta)$ називають **центральною статистикою** для θ.

Приклад 17.2

Для загальної нормальної моделі, якщо $\xi_i \sim N(\theta_1, \theta_2^2)$, підходять в якості центральних статистик $\sqrt{n-1}(\bar{\xi} - \theta_1)/S \sim t(n-1)$ для параметру зсуву θ_1, та $nS^2/\theta_2^2 \sim \chi^2(n-1)$ для параметру масштабу θ_2.

Нехай для моделі $\mathscr{F}$ побудована центральна статистика $G(\xi, \theta)$ і $f_G(g)$ - її щільність розподілу. Функція $f_G(g)$ від параметра θ) не залежить (умова 1), тому для довільного $\gamma \in (0,1)$ можна обрати $g_1 < g_2$ (багатьма способами) так, що

$$\mathrm{P}_\theta \left\{ g_1 < G(\xi, \theta) < g_2 \right\} = \int\limits_{g_1}^{g_2} f_G(g) dg = \gamma \text{ для усіх } \theta \in \Theta. \tag{17.2}$$

Визначимо при кожному $x \in X$ числа $T_1(x), T_2(x)$ $(T_1(x) < T_2(x))$ як розв'язок за θ рівнянь

$$G(x, \theta) = g_i, i = 1, 2. \tag{17.3}$$

Однозначність визначення цих чисел забезпечує умова 2, яка накладається на функцію $G(\xi, \theta)$. Тоді нерівності

$$g_1 < G(x, \theta) < g_2 \tag{17.4}$$

еквівалентні нерівностям

$$T_1(x) < \theta < T_2(x)$$

і (17.2) можна переписати у вигляді

$$\mathrm{P}_\theta \left\{ T_1(\xi) < \theta < T_2(\xi) \right\} = \gamma \ \text{ для усіх } \theta \in \Theta.$$

Отже інтервал $(T_1(\xi), T_2(\xi))$ є γ-надійним інтервалом для θ.

Зауважимо, що в (17.4) g_1 та g_2 є квантилями розподілу центральної статистики $G(\xi, \theta)$ рівнів α_1 та $1 - \alpha_2$, де $\alpha_1 + \alpha_2 = 1 - \gamma$. І, власне, числа α_1, α_2 можна вибрати безліччю способів. Але оскільки метою є побудова надійного інтервалу мінімальної довжини, то в якості цих чисел треба вибрати такі α_1^*, α_2^*, для яких довжина надійного інтервалу $G^{-1}(\xi, g(1 - \alpha_2*)) - G^{-1}(\xi, g(\alpha_1^*))$ досягає мінімуму по α_1, α_2.

У випадках, коли розподіл $F_G(g)$ є симетричним (наприклад, нормальний розподіл, розподіл Стьюдента), тобто $F_G(g) = 1 - F_G(-g)$, точкою, в якій довжина інтервалу досягає свого мінімуму є $\alpha_1 = \alpha_2 = (1-\gamma)/2$.

Приклад 17.3

Для параметра θ зсуненого показникового закону з прикладу 17.1 можна взяти в якості центральної статистики $G(\xi, \theta) = \xi_{(1)} - \theta$.

Дійсно, знаючи, що функція розподілу мінімуму

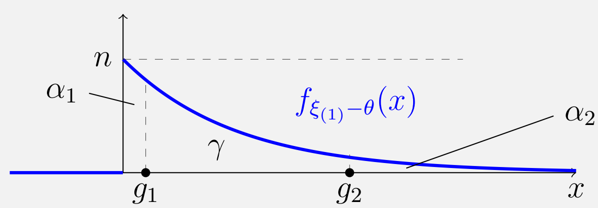
$$F_{\xi_{(1)}}(x) = 1 - (1 - F_\xi(x))^n,$$

легко знайти розподіл статистики $G(\xi, \theta)$:

$$F_{\xi_{(1)}-\theta}(x) = \mathrm{P}\{\xi_{(1)} - \theta \le x\} = \mathrm{P}\{\xi_{(1)} \le x + \theta\} = F_{\xi_{(1)}}(x+\theta) = 1 - e^{-nx}, \quad x \ge 0.$$

Відповідно щільність має вид $f_{\xi_{(1)}-\theta}(x) = ne^{-nx}$, $x \ge 0$, – це показниковий розподіл з параметром n, який не залежить від параметру θ.

Тепер знайдемо межі надійного інтервалу з урахуванням (17.2). Для цього спочатку треба визначити α_1 та α_2 таким чином, щоб $\alpha_1 + \alpha_2 = 1 - \gamma$, і таких способів безліч.



Проте, з рисунка видно, що для того, щоб отримати надійний інтервал найменшої довжини, треба максимально зсунути його ліворуч, тобто покласти $g_1 = 0$. В цьому випадку $\alpha_1 = 0$, $\alpha_2 = 1 - \gamma$, а $g_2 = g_\gamma$ – квантиль рівня γ показникового розподілу з параметром n. З рівняння

$$F_G(g_\gamma) = F_{\xi_{(1)}-\theta}(g_2) = 1 - e^{-ng_2} = \gamma$$

маємо

$$g_2 = -\frac{\ln(1-\gamma)}{n}.$$

Тепер (17.2) можна переписати таким чином:

$$\mathrm{P}\left\{0 < G(\xi, \theta) < -\frac{\ln(1-\gamma)}{n}\right\} = \mathrm{P}\left\{0 < \xi_{(1)} - \theta < -\frac{\ln(1-\gamma)}{n}\right\} = \gamma,$$

Звідки маємо найкоротший γ-надійний інтервал для параметру θ :

$$\mathrm{P}\left\{\xi_{(1)} + \frac{\ln(1-\gamma)}{n} < \theta < \xi_{(1)}\right\} = \gamma$$

Приклад 17.4 Побудова асимптотичного надійного інтервалу

Пошук центральної статистики – це окрема, часом непроста задача, і навіть якщо вдалося її визначити, в свою чергу точний надійний інтервал не завжди можна побудувати. Універсальний вид центральної статистики для побудови *асимптотичних надійних інтервалів* дає центральна гранична теорема.

Нехай $\xi_1, ..., \xi_n$ – вибірка з генеральної сукупності з показникового розподілу з невідомим параметром $\theta > 0$. Побудуємо асимптотичний (асимптотично точний) γ-надійний інтервал для параметра θ.

Згідно ЦГТ

$$\frac{\sum_{i=1}^{n} \xi_i - n\mathbf{M}\xi_1}{\sqrt{n\mathbf{D}\xi_1}} = \sqrt{n}\frac{\bar{\xi} - 1/\theta}{1/\theta} = \sqrt{n}(\theta\bar{\xi} - 1) \Rightarrow \eta \sim N(0,1).$$

Отже, статистика $G(\xi,\theta) = \sqrt{n}(\theta\bar{\xi} - 1)$ має розподіл, який *збігається* до розподілу, який не залежить від параметра θ.

Візьмемо в якості константи $g_2 = z_{\frac{1+\gamma}{2}}$ квантиль стандартного нормального розподілу рівня $\frac{1+\gamma}{2}$ (в цьому випадку $g_1 = -g_2$ в силу симетричності розподілу). Згідно визначенню слабкої збіжності, при $n \to \infty$

$$\mathrm{P}\{-g_2 < \sqrt{n}(\theta\bar{\xi} - 1) < g_2\} \to \mathrm{P}\{-g_2 < \eta < g_2\} = \gamma.$$

Розв'язуючи нерівність всередині ймовірності відносно θ, отримаємо для нього асимптотичний довірчий інтервал:

$$\mathrm{P}\left\{\frac{1}{\bar{\xi}} - \frac{z_{\frac{1+\gamma}{2}}}{\sqrt{n}\bar{\xi}} < \theta < \frac{1}{\bar{\xi}} + \frac{z_{\frac{1+\gamma}{2}}}{\sqrt{n}\bar{\xi}}\right\} \to \gamma, \text{ при } n \to \infty.$$

17.3 Інтервальне оцінювання в нормальній моделі

17.3.1 Надійний інтервал для середнього, коли відома дисперсія

Нехай $\xi' = (\xi_1, ..., \xi_n)$ – вибірка з нормального розподілу $N(\theta, \sigma^2)$. Параметр θ невідомий, а σ^2 – відоме. Цю модель часто застосовують до даних, отриманих при незалежних вимірюваннях деякої величини θ за допомогою приладу (або методу), який має відому середню похибку (стандартну) σ. Для цієї моделі випадкова величина

$$G(\xi,\theta) = \sqrt{n}\frac{\bar{\xi} - \theta}{\sigma}, \quad \text{де } \bar{\xi} = \frac{\xi_1 + \cdots + \xi_n}{n},$$

буде центральною статистикою.

Дійсно, $G(\xi,\theta) = \dfrac{\sum\limits_{i=1}^{n}(\xi_i-\theta)}{\sqrt{n}\sigma}$ як сума незалежних нормально розподілених випадкових величин буде розподілена нормально. Оскільки

$$\mathbf{M}_\theta G(\xi,\theta) = 0, \;\; \mathbf{D}_\theta G(\xi,\theta) = 1,$$

то $G(\xi,\theta)$ має нормальний розподіл з параметрами (0,1), який, очевидно, не залежить від параметра θ. При фіксованому x, функція $G(\xi,\theta)$ неперервна і монотонно спадна за θ. Таким чином умови 1),2) виконуються.

Розв'язком рівнянь (17.3) будуть функції

$$T_1(x) = \bar{x} - \frac{\sigma}{\sqrt{n}}g_2, \quad T_2(x) = \bar{x} - \frac{\sigma}{\sqrt{n}}g_1,$$

а γ-надійний інтервал для θ можна записати формулою

$$\Delta_\gamma(\xi) = \left(\bar{\xi} - \frac{\sigma}{\sqrt{n}}g_2, \bar{\xi} - \frac{\sigma}{\sqrt{n}}g_1\right),$$

де $g_1 < g_2$ – довільні числа, що задовольняють умові

$$\Phi(g_2) - \Phi(g_1) = \gamma,$$

де $\Phi(\cdot)$ – функція розподілу стандартного нормального закону.

Хоч інтервал $\Delta_\gamma(\xi)$ випадковий, його довжина стала і дорівнює

$$l_\gamma(g_1,g_2) = \frac{\sigma(g_2-g_1)}{\sqrt{n}}.$$

Тому, щоб серед множини інтервалів $\Delta_\gamma(\xi)$ знайти інтервал мінімальної довжини, треба розв'язати задачу на умовний екстремум

$$\begin{cases} \dfrac{\sigma(g_2-g_1)}{\sqrt{n}} \to \min, \\ \Phi(g_2) - \Phi(g_1) = \gamma. \end{cases}$$

Використовуючи метод невизначених множників Лагранжа, можна довести, що мінімум досягається при $g_2 = -g_1$. Тепер g_2 визначається однозначно

$$\Phi(g_2) - \Phi(-g_2) = \Phi(g_2) - (1-\Phi(g_2)) = \gamma,$$

$$\Phi(g_2) = \frac{1+\gamma}{2},$$

$$g_2 = \Phi^{-1}\left(\frac{1+\gamma}{2}\right),$$

де $z_{\frac{1+\gamma}{2}} = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$ – квантиль порядку $\frac{1+\gamma}{2}$ для стандартного нормального розподілу (див. табл. 3 на стор. 302).

У результаті отримуємо оптимальний γ-надійний інтервал

$$\Delta_\gamma(\xi) = \left(\bar{\xi} - \frac{\sigma}{\sqrt{n}}z_{\frac{1+\gamma}{2}}, \bar{\xi} + \frac{\sigma}{\sqrt{n}}z_{\frac{1+\gamma}{2}}\right).$$

17.3.2 Надійний інтервал для дисперсії, коли відоме середнє

Нехай $\xi' = (\xi_1, ..., \xi_n)$ – вибірка з нормального розподілу $N(\mu, \theta^2)$. Параметр μ тепер вважатимемо відомим, а θ^2 – невідомим. Таку модель можна використовувати для визначення середньої точності приладу (або методу) шляхом багатократних вимірювань еталона.

За центральну статистику візьмемо $G(\xi, \theta) = \frac{1}{\theta^2}\sum\limits_{i=1}^{n}(\xi_i - \mu)^2$. Розподіл випадкової величини $G(\xi, \theta)$ збігається, очевидно, з розподілом суми квадратів n незалежних випадкових величин, кожна з яких розподілена згідно стандартного нормального закону розподілу.

За параметром θ $G(\xi, \theta)$ неперервна й монотонно спадна.

Застосовуючи метод центральної статистики, знаходимо нижню і верхню границю надійності у вигляді

$$T_1(\xi) = \frac{1}{g_2}\sum_{i=1}^{n}(\xi_i - \mu)^2, \quad T_2(\xi) = \frac{1}{g_1}\sum_{i=1}^{n}(\xi_i - \mu)^2,$$

де числа $0 < g_1 < g_2 < \infty$ задовольняють умові

$$\gamma = \int_{g_1}^{g_2} k_n(x)dx = \mathrm{P}_\theta\{g_1 < G(\xi, \theta) < g_2\},$$

де $k_n(x) = \dfrac{x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}$ – щільність χ^2 розподілу з n степенями свободи, $\Gamma(\lambda) = \int\limits_0^\infty x^{\lambda-1}e^{-x}dx,\ \lambda > 0$ – гамма-функція.

Зазвичай g_1 і g_2 обирають так, щоб

$$\int_0^{g_1} k_n(x)dx = \frac{1-\gamma}{2}, \int_{g_2}^{\infty} k_n(x)dx = \frac{1-\gamma}{2}. \tag{17.5}$$

В цьому випадку площа, що залишилася за межами криволінійної трапеції з основою $[g_1, g_2]$ ділиться навпіл. Тепер g_1 і g_2 визначаються однозначно через γ:

$$g_1 = \chi^2_{\frac{1-\gamma}{2}}(n), \quad g_2 = \chi^2_{\frac{1+\gamma}{2}}(n),$$

де $\chi^2_p(n)$ – p-квантиль розподілу $\chi^2(n)$ (див. табл. 7 на стор. 308).

Надійний інтервал, побудований за умови (17.4), називають **центральним**. Для моделі $N(\mu, \theta^2)$ центральний надійний інтервал має вигляд

$$\Delta_\gamma(\xi) = \left(\frac{1}{\chi^2_{\frac{1+\gamma}{2}}(n)}\sum_{i=1}^{n}(\xi_i - \mu)^2, \frac{1}{\chi^2_{\frac{1-\gamma}{2}}(n)}\sum_{i=1}^{n}(\xi_i - \mu)^2\right).$$

17.3.3 Загальна нормальна модель. Надійний інтервал для дисперсії

Нехай $\xi' = (\xi_1, \xi_2, \ldots, \xi_n)$ – вибірка з розподілу $N(\theta_1, \theta_2^2)$. Щоб побудувати центральну статистику для дисперсії, потрібен такий результат.

Теорема 17.1 Нехай $\xi' = (\xi_1, \xi_2, \ldots, \xi_n)$ – вибірка з розподілу $N(\theta_1, \theta_2^2)$. Тоді $\dfrac{nS^2}{\theta^2}$, де $S^2(\xi) = \frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - \bar{\xi}\right)^2$, має χ^2-розподіл з $(n-1)$ ступенями свободи.

Звідси випливає, що розподіл випадкової величини

$$G\left(\xi, \theta_2^2\right) = \frac{1}{\theta_2^2}\sum_{i=1}^{n}\left(\xi_i - \bar{\xi}\right)^2$$

збігається з розподілом $\chi^2(n-1)$ і не залежить від параметру $\theta = (\theta_1, \theta_2^2)$. Неперервність і монотонність за θ_2 очевидна. Отже $G(\xi, \theta_2^2)$ – центральна статистика для θ_2^2. Враховуючи попередній аналіз $N(\mu, \theta^2)$-моделі, робимо висновок, що центральним γ - надійним інтервалом для θ_2^2 є інтервал

$$\Delta_\gamma(\xi) = \left(\frac{nS^2(\xi)}{\chi^2_{\frac{1+\gamma}{2},n-1}}, \frac{nS^2(\xi)}{\chi^2_{\frac{1-\gamma}{2},n-1}}\right), \tag{17.6}$$

де $\chi^2_{\frac{1+\gamma}{2},n-1}$ – $\frac{1+\gamma}{2}$-квантиль розподілу $\chi^2(n-1)$ (див. табл. 7 на стор. 308).

17.3.4 Загальна нормальна модель. Надійний інтервал для середнього

Побудуємо центральну статистику для середнього θ_1 у загальній моделі $N(\theta_1, \theta_2^2)$.

Розглянемо випадкову величину

$$G(\xi, \theta_1) = \sqrt{n-1}\frac{\bar{\xi} - \theta_1}{S(\xi)} = \frac{\sqrt{n}(\bar{\xi} - \theta_1)/\theta_2}{\sqrt{\frac{1}{\theta_2^2}\sum\limits_{i=1}^{n}(\xi_i - \bar{\xi})^2/(n-1)}}.$$

Оскільки вибіркове середнє $\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n}\xi_i$ і дисперсія $S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - \bar{\xi}\right)^2$ незалежні, то $G(\xi, \theta_1)$ має розподіл Стьюдента з $(n-1)$ степенями свободи. Таким чином $G(\xi, \theta_1)$ – центральна статистика для θ_1. Враховуючи схожість розподілу Стьюдента і стандартного нормального розподілу, маємо для θ_1 надійний інтервал мінімальної довжини

$$\left(\bar{\xi} - \frac{S(\xi)}{\sqrt{n-1}}t_{\frac{1+\gamma}{2},n-1},\ \bar{\xi} + \frac{S(\xi)}{\sqrt{n-1}}t_{\frac{1+\gamma}{2},n-1}\right), \tag{17.7}$$

де $t_{\frac{1+\gamma}{2},n-1}$ – $\frac{1+\gamma}{2}$-квантиль розподілу Стьюдента з $(n-1)$ степенями свободи (див. табл. 6 на стор. 306).

Приклад 17.5

Серед даних, з якими працював Вільям Госсет (Стьюдент) є масив спостережень над групою з 10 пацієнтів, яким давали два снодійні препарати для порівняння їхньої ефективності і фіксували збільшення годин сну порівняно з контрольним рівнем. Цей датасет "sleep" є, зокрема, в базовому пакеті "datasets" R (інформацію про R див. в розділі 20 посібника).

Розглянемо дію першого препарату. Маємо 10 спостережень:

0.7, -1.6, -0.2, -1.2, -0.1, 3.4, 3.7, 0.8, 0.0, 2.0.

З певних міркувань можна вважати ці дані нормально розподіленими. Обчислимо надійні інтервали рівня $\gamma = 0.95$ для середнього ефекту першого препарату та його дисперсії.

В даному випадку параметри розподілу невідомі, тобто ми маємо справу з загальною нормальною моделлю. Спочатку знайдемо значення оцінок середнього та дисперсії $\bar{\xi}_n$ та S_n^2 за формулами $\bar{\xi} = \frac{1}{n}\sum_{i=1}^n \xi_i$ і $S^2 = \frac{1}{n}\sum_{i=1}^n \left(\xi_i - \bar{\xi}\right)^2$:

$$\bar{\xi}_{10} = \frac{1}{10}(0.7 - 1.6 - 0.2 - 1.2 - 0.1 + 3.4 + 3.7 + 0.8 + 0.0 + 2.0) = 0.75;$$

$$\sum_{i=1}^n (\xi_i - \bar{\xi}_{10})^2 = (0.7 - 0.75)^2 + (-1.6 - 0.75)^2 + ... + (2.0 - 0.75)^2 = 28.805;$$

$$S_{10}^2 = \frac{1}{10}\sum_{i=1}^n (\xi_i - \bar{\xi}_n)^2 = 2.8805; \quad S_{10} \approx 1.7.$$

Зауважимо, що в більшості статистичних пакетів (зокрема в R) в якості оцінки дисперсії за замовчуванням обчислюється незсунена оцінка дисперсії $\hat{S}_n^2$, яка відрізняється від S_n^2 : $S_n^2 = \frac{n-1}{n}\hat{S}_n^2$. Для дуже великих значень n різниця між цими оцінками дисперсії буде незначна, проте при роботі з невеликими датасетами, як в цьому прикладі, різниця доволі суттєва. Отже, при використанні подібних програм треба пильнувати, що саме вона обчислює.

В формулах (17.6) та (17.7) для обчислення надійних інтервалів для загальної нормальної моделі фігурують також квантилі розподілу Стьюдента та χ^2-розподілу. Їх візьмемо в таблицях 6 та 7 додатку для $n - 1 = 9$, $(1 + \gamma)/2 = 0.975$, $(1 - \gamma)/2 = 0.025$:

$$t_{0.975;9} = 2.262; \quad \chi^2_{0.975;9} = 19.023; \quad \chi^2_{0.0.025;9} = 2.7.$$

Підставляючи отримані значення в (17.7) та (17.6), матимемо надійний інтервал для середнього ефекту снодійного №1 (параметра θ_1 нормального розподілу)

$$\left(0.75 - \frac{1.7}{\sqrt{9}} \cdot 2.262; 0.75 + \frac{1.7}{\sqrt{9}} \cdot 2.262\right) = (-0.532; 2.032); \tag{17.8}$$

та для його дисперсії (θ_2^2)

$$\left(\frac{1}{19.023}\cdot 28.805; \frac{1}{2.7}\cdot 28.805\right) = (1.514; 10.669).$$

Бачимо, що обидва інтервали дуже широкі. Це пов'язано з тим, що даних для аналізу, насправді, дуже мало, і точний висновок на їхній основі зробити неможливо.

Забігаючи наперед, зауважимо, що надійний інтервал (17.8) для ефекту першого снодійного містить в собі значення 0. Це означає, що на основі наявних даних не можна вважати ефект цього препарату значущим (відмінним від 0). Детальніше про подібні висновки буде сказано в розділі "Параметричні гіпотези".

17.4 Побудова надійних інтервалів на основі точкових оцінок

Якщо є деяка точкова оцінка $T(\xi)$ параметра θ і відома її функція розподілу $F_T(t,\theta)$, то надійний інтервал можна побудувати при умові, що $F_T(t,\theta)$ неперервна і монотонна по θ.

Обираючи різні оцінки $T(\xi)$, отримаємо різні надійні інтервали. Кінцева мета – при фіксованому рівні надійності γ отримати якомога коротший інтервал. Припустимо, що використовуються незсунені і приблизно нормальні оцінки. Тоді інтервали тим коротші, чим менше дисперсія оцінки. Таким чином ефективні і асимптотично ефективні оцінки призводять до найменших або асимптотично найменших інтервалів. Цим вимогам задовольняють оцінки максимальної вірогідності $\hat{\theta}_n$.

Зафіксуємо $0 < \gamma < 1$ і $\hat{c}_\gamma$ визначимо з рівняння $2\Phi(\hat{c}_\gamma) - 1 = \gamma$. Отже, $\hat{c}_\gamma = z_{\frac{1+\gamma}{2}} = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$ – квантиль порядку $\frac{1+\gamma}{2}$ для стандартного нормального розподілу (див. табл. 3 на стор. 302). При виконанні умов теореми про асимптотичну нормальність оцінок максимальної вірогідності

$$\mathrm{P}_\theta\left\{|\hat{\theta}_n - \theta|\sqrt{I_n\left(\hat{\theta}_n\right)} \le \hat{c}_\gamma\right\} \xrightarrow[n\to\infty]{} \Phi(\hat{c}_\gamma) - \Phi(-\hat{c}_\gamma) = 2\Phi(\hat{c}_\gamma) - 1 = \gamma.$$

Отже,

$$\left(\hat{\theta}_n - \frac{z_{\frac{1+\gamma}{2}}}{\sqrt{I_n\left(\hat{\theta}_n\right)}},\ \hat{\theta}_n + \frac{z_{\frac{1+\gamma}{2}}}{\sqrt{I_n\left(\hat{\theta}_n\right)}}\right)$$

є асимптотично найменшим γ-надійним інтервалом для параметра θ.

Приклад 17.6

Побудувати асимтотичний γ-надійний інтервал для невідомого параметра θ за вибіркою з генеральної сукупності з розподілом Пуассона.

Розв'язок: Нехай $\xi' = (\xi_1, \xi_2, \ldots, \xi_n)$ – вибірка з генеральної сукупності з розподілом Пуассона з параметром θ. Відповідно

$$\mathrm{P}\left\{\xi_0 = x\right\} = \frac{\theta^x}{x!} e^{-\theta},\ x = 0, 1, \ldots$$

Якщо $x' = (x_1, x_2, \ldots, x_n)$ – реалізація вектора ξ', то функція вірогідності

$$L(x, \theta) = e^{-n\theta} \frac{\theta^{\sum_{k=1}^{n} x_k}}{\prod_{k=1}^{n} x_k!}.$$

Далі,

$$\frac{\partial \ln L(x,\theta)}{\partial \theta} = \frac{n}{\theta}\left(\bar{x} - \theta\right), \quad \frac{\partial^2 \ln L(x,\theta)}{\partial \theta^2} = -\frac{n}{\theta^2}\bar{x}.$$

З рівняння $\dfrac{\partial \ln L(x,\theta)}{\partial \theta} = 0$ знаходимо оцінку максимальної вірогідності

$$\hat{\theta}_n = \bar{\xi} = \frac{1}{n}\sum_{k=1}^{n} \xi_k.$$

$$I\left(\theta\right) = -\mathbf{M}\frac{\partial^2 \ln L(\xi,\theta)}{\partial \theta^2} = \frac{n\theta}{\theta^2} = \frac{n}{\theta}.$$

Тоді асимптотично найкоротший γ-надійний інтервал для параметра θ буде таким:

$$\left(\bar{\xi} - z_{\frac{1+\gamma}{2}}\sqrt{\frac{\bar{\xi}}{n}},\ \bar{\xi} + z_{\frac{1+\gamma}{2}}\sqrt{\frac{\bar{\xi}}{n}}\right)$$

Задача 17.1. Нехай $\xi_1, \ldots, \xi_n$ – вибірка розміру n з генеральної сукупності з рівномірного розподілу на інтервалі $[\theta, \theta+1]$, $\theta \in \mathbb{R}$. Яка довжина інтервалу c_γ підійде для побудови надійного інтервалу на базі статистики $\xi_{(1)}$?

Задача 17.2. Надійний інтервал для параметра біноміального розподілу Нехай

$$\mathrm{P}\left\{\xi_0 = 0\right\} = C_n^x \theta^x \left(1-\theta\right)^{n-x}, \quad x = 0, 1, \ldots, n, \quad \theta \in (0; 1).$$

Побудувати надійний інтервал для параметра θ.

Задача 17.3. Для вибірки $(x_1, x_2, x_3) = (1, 2, 3)$ з генеральної сукупності з розподілом $N(m, \sigma^2)$ побудувати довірчі інтервали для m та для σ^2 при $\gamma = 0.1$.

Задача 17.4. Знайти довірчий інтервал для ймовірності попадання снаряду в ціль з довірчою ймовірністю $\gamma = 0.95$, якщо після 220 пострілів у ціль попало 75 снарядів.

Задача 17.5. Виконано 100 незалежних випробувань, у результаті яких подія A спостерігалась 40 разів. Побудувати надійний інтервал для ймовірності події A за рівнів надійності 0.95 та 0.99, якщо кількість появ події A має біноміальний розподіл.

Задача 17.6. На телефонній станції проводились спостереження за кількістю невірних з'єднань за хвилину. Спостереження упродовж години дали такі результати:

x_i	0	1	2	3	4	5	7
n_i^*	8	17	16	10	6	2	1

Припускаючи, що кількість невірних з'єднань за хвилину має пуассонівський розподіл, знайти надійний інтервал для невідомого параметра з надійністю 0.99.

Задача 17.7. Побудуйте надійний інтервал для параметра θ розподілу $N(\theta, 4\theta^2)$.

Задача 17.8. Побудувати надійний інтервал для параметра θ за вибіркою $\xi_1, \xi_2, \ldots, \xi_n$ із генеральної сукупності з розподілом:

а) $\mathrm{P}_\theta\{\xi_0 = x\} = \dfrac{\theta^x}{(1+\theta)^{x+1}}$, $x = 0, 1, \ldots, \theta > 0$;

б) $\mathrm{P}_\theta\{\xi_0 = x\} = \dfrac{(\theta-1)^x}{\theta^{x+1}}$, $x = 0, 1, \ldots, \theta > 1$.

Задача 17.9. Знайти надійний інтервал для параметрів a та σ^2 нормального розподілу за вибіркою:

а) 0.6, 2.4, 2.1, 1.4, 1.2, 4.8, 0.9, 1.1, 3.5, 3.0;

б) 0.2, 0.5, 1.0, 1.5, 0.8, 1.0, 2.0.

Для параметра a покласти $\gamma = 0.95$, для σ^2 – $\gamma = 0.9$.

Задача 17.10. Знайти надійний інтервал для математичного сподівання нормального розподілу $N(a, \sigma^2)$, якщо $n = 25$, $\bar{x} = 16.8$, $\sigma^2 = 25$, $\gamma = 0.99$.

Задача 17.11. Знайти надійний інтервал для дисперсії нормального розподілу $N(a, \theta^2)$, якщо $n = 20$, $\hat{S}^2 = 10$, $\gamma = 0.99$.

Задача 17.12. За групованою вибіркою з нормального розподілу

x_i	[-2; 0)	[0;1)	[1;2)	[2;3)	[3;4)	[4;5)	[5;6)
n_i^*	3	3	4	6	2	1	1

з надійністю $\gamma = 0.95$ знайти інтервальні оцінки для параметра a, коли:

а) $\sigma^2 = 4$;

б) σ^2 - невідоме.

Знайти також надійний інтервал для σ^2, коли:

а) $a = 2$;

б) a - невідоме. Надійність γ покласти 0.9.

Задача 17.13. Припускається, що ємність конденсатора має розподіл $N(m, 16)$. За результатами вимірювання партії з 16 конденсаторів, було визначено середню ємність $\bar{x} = 20$ мкФ. Знайти довірчий інтервал для невідомого математичного сподівання m з рівнем надійності 0.9.

Задача 17.14. З великої партії електроламп випадковим чином було відібрано 400 штук для вимірювання середньої тривалості горіння. Вибіркове середнє тривалості горіння ламп виявилася рівним $\bar{x} = 500$ год. Знайти з довірчою ймовірністю $\gamma = 0.99$ довірчий інтервал для середньої тривалості горіння електроламп по всій партії, якщо середнє квадратичне відхилення тривалості горіння $\sigma = 10$ год.

Задача 17.15. Оцінка величини опору для партії резисторів, підрахована за результатами 100 тестів, склала $\bar{x} = 10$ кОм. Вважаючи, що опір резистора – нормально розподілена випадкова величина з дисперсією $\sigma^2 = 1$ кОм^2, знайти:

а) ймовірність того, що для всієї партії опір знаходиться в межах 10 ± 0.1 кОм;

б) скільки вимірів (тестів) треба зробити, щоб з імовірністю, не меншою за 0.95 стверджувати, що опір знаходиться в межах 10 ± 0.1 кОм.

Задача 17.16. Результати 10 вимірів ємності конденсаторів приладом, який не має систематичної помилки, склали такі відхилення від номіналу (мкФ): 5.4, –13.9, –11, 7.2, –15.6, 29.2, 1.4, –0.3, 6.6, –9.9. Знайти 90%-ий довірчий інтервал для дисперсії та середньоквадратичного відхилення.

Задача 17.17. Для визначення вертикального кута орієнтира використовують середнє арифметичне декількох вимірів кута за допомогою секстанта. Для виміряних кутів покладають $\sigma = 1.5'$ (1.5 хвилини). Знайти кількість вимірів, які треба зробити, щоб

а) похибка результату з імовірністю 0.99 не перевищувала $1'$;

б) похибка результату з імовірністю 0.95 не перевищувала $1.5'$.

Задача 17.18. Середня квадратична похибка висотоміра $\sigma = 15$ м. Скільки потрібно мати таких приладів на літаку, щоб з імовірністю 0.99 похибка виміру середньої висоти $\bar{\xi}$ була менша ніж 30 м? При цьому випадкові помилки розподілені за нормальним законом, а систематичні помилки відсутні.

Задача 17.19. Із партії валів відібрали $n_1 = 9$ штук. Значення вибіркового середнього діаметру вала $\bar{x}_1 = 30$ мм, вибіркової дисперсії $S_1^2 = 9$ мм^2. Потім зробили повторний експеримент, відібравши $n_2 = 16$ шт. і отримали значення вибіркових оцінок $\bar{x}_2 = 29$ мм, $S_1^2 = 4.5$ мм^2. Використовуючи об'єднані вибіркові оцінки, знайдіть 99%-ий довірчий інтервал для середнього.

Задача 17.20. Відомо, що вимірювальний пристрій не має систематичних похибок, а випадкові похибки виміру підпорядковуються нормальному закону розподілу. Скільки потрібно провести вимірів для визначення оцінки середнього квадратичного відхилення приладу, щоб з довірчою ймовірністю 0.7 абсолютна величина похибки визначення цієї величини була не більша 20 % від $\bar{\sigma}$?

Задача 17.21. З автоматичної лінії, що виробляє підшипники, було відібрано 100 шт., причому 10 виявилися бракованими. Знайдіть:

а) 90%-ий довірчий інтервал для ймовірності того, що навмання вибраний підшипник виявиться бракованим;

б) кількість підшипників, які потрібно перевірити, щоб з ймовірністю 0.9973 можна було стверджувати, що доля браку відрізняється від частоти не більше ніж на 0.05.

Задача 17.22. При перевірці 100 деталей із великої партії виявлено 10 бракованих. Знайдіть 95%-ий довірчий інтервал для кількості бракованих деталей у всій партії.

Задача 17.23. При перевірці впливу снодійного №2 вимірювалась величина збільшення кількості годин сну порівняно з контрольним рівнем для 10 пацієнтів (дані масиву "sleep", див. приклад на стор. 236) отримано такі результати: 1.9, 0.8, 1.1, 0.1, -0.1, 4.4, 5.5, 1.6, 4.6, 3.4. Побудувати 0.99-надійний інтервал для дисперсії та 0.9-надійний інтервал для середнього ефекту цього препарату, вважаючи дані нормально розподіленими.

18 ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

18.1 Поняття статистичної гіпотези та статистичного критерію

Статистична гіпотеза – це довільне твердження про тип або властивості розподілів випадкових величин, що спостерігаються в експерименті.

Приклад 18.1

Нехай експеримент полягає в багаторазовому вимірюванні деякої фізичної величини, точне значення якої a невідоме й у процесі вимірювань не змінюється. На результати вимірювань впливають багато факторів: точність налагодження приладу, похибка заокруглення тощо, тому результат i-го вимірювання ξ_i можна записати у вигляді

$$\xi_i = a + \varepsilon_i,$$

де ε_i – випадкова похибка вимірювання.

Будемо вважати, що загальна похибка ε_i складається з великої кількості похибок, кожна з яких невелика. На основі центральної граничної теореми припускатимемо, що випадкові величини ξ_i мають нормальний розподіл. Таке припущення є статистичною гіпотезою про тип розподілу випадкових величин, що спостерігаються. Наведемо кілька типів статистичних гіпотез.

1. Гіпотеза про тип розподілу. Нехай виконано n незалежних спостережень над деякою випадковою величиною ξ_0 з невідомою функцією розподілу $F_{\xi_0}(z)$. Гіпотеза, яка підлягає перевірці:

$$H_0:\ F_{\xi_0}(z) = F(z),$$

де функція $F(z)$ повністю задана, або $H_0:\ F_{\xi_0}(z) \in \mathscr{F} = \{F(z,\theta), \theta \in \Theta\}$ – задане сімейство функцій розподілу.

2. Гіпотеза однорідності. Нехай виконано k серій незалежних спостережень

$$(\xi_{i1}, \xi_{i2}, \ldots, \xi_{in}) \quad i = 1, 2, \ldots, k$$

з генеральних сукупностей з функціями розподілу $F_i(z),\ i = 1, 2, \ldots, k$ (узагалі кажучи, невідомими). Чи є підстави розглядати ці дані як результати спостережень над тією самою випадковою величиною? Якщо це так, то кажуть, що статистичні дані однорідні. Відповідно перевіряється гіпотеза однорідності

$$H_0 : F_1(z) = \ldots = F_k(z).$$

3. Гіпотеза незалежності.

В експерименті спостерігається двовимірна випадкова величина (ξ, η) з невідомою сумісною функцією розподілу $F_{\xi,\eta}(z_1, z_2)$, і є підстави вважати, що компоненти ξ та η незалежні. У цьому випадку треба перевірити гіпотезу незалежності, тобто

$$H_0 : F_{\xi,\eta}(z_1, z_2) = F_\xi(z_1)F_\eta(z_2).$$

Якщо гіпотеза H_0 однозначно фіксує розподіл спостережень, то її називають **простою**, у протилежному випадку – **складною**. У наведених вище прикладах лише гіпотеза про тип розподілу $H_0 :\ F_{\xi_0}(z) = F(z)$ є простою.

Статистичний критерій – це правило, згідно з яким гіпотеза, що перевіряється, приймається або відхиляється.

Розглянемо методи перевірки гіпотез описаних вище типів. Нехай про розподіл вибірки $\xi' = (\xi_1, ..., \xi_n)$, що описує результати експерименту, сформульована гіпотеза H_0. Необхідно перевірити, узгоджуються чи ні статистичні дані з цією гіпотезою. Відповідні критерії називаються **критеріями згоди**. Наведемо методику побудови критеріїв згоди.

Обирається статистика $T = T(\xi)$, яка характеризує відхилення емпіричних даних від гіпотетичних значень, що відповідають гіпотезі H_0. Вона є мірою розбіжності статистичного та гіпотетичного законів розподілу і називається **статистикою критерію**. Розподіл статистики $T = T(\xi)$ треба знати точно або наближено в припущенні, що розподіл спостережень збігається з гіпотетичним.

Нехай таку статистику знайдено. Визначимо для фіксованого достатньо малого числа $\alpha > 0$ число t_α так, щоб у випадку справедливості гіпотези H_0 імовірність настання події $\mathrm{P}\{T(\xi) > t_\alpha | H_0\} = \alpha$. Число α називається **рівнем значущості критерію**.

Нехай $x' = (x_1, \ldots, x_n)$ – реалізація вибірки, а $\hat{T} = T(x)$ – відповідне значення статистики T , обчислене за статистичними даними. Якщо $\hat{T} \geq t_\alpha$, то відхилення від гіпотетичного закону розподілу вважається значущим і гіпотеза відхиляється. У протилежному випадку немає підстав відмовлятися від висунутої гіпотези і слід вважати, що спостереження не суперечать гіпотезі (на рівні α). Область $\{\hat{T} \geq t_\alpha\}$ називається **критичною областю** для гіпотези H_0.

18.2 Гіпотези про тип розподілу

18.2.1 Критерій згоди Колмогорова

Нехай $\xi' = (\xi_1, ..., \xi_n)$ – вибірка з генеральної сукупності з невідомою функцією розподілу $F_{\xi_0}(z)$, про яку висунута проста гіпотеза $H_0 :\ F_{\xi_0}(z) = F(z)$, де $F(z)$ – неперервна функція. Статистикою критерію Колмогорова є величина

$$D_n = D_n(\xi) = \sup_{-\infty<z<+\infty} |F_n(z) - F(z)|,$$

яка є максимальним відхиленням емпіричної функції розподілу $F_n(z)$ від гіпотетичної $F(z)$.

У тих випадках, коли гіпотеза H_0 справедлива, зі збільшенням розміру вибірки n відбувається зближення $F_n(z)$ із $F(z)$. Тому принаймні для великих n значення D_n не повинно істотно відрізнятися від 0.

Особливості статистики D_n:

1) Її розподіл за справедливості гіпотези H_0 не залежить від вигляду функції $F(z)$:

$$D_n = \sup_{-\infty<z<+\infty} |F_n(z) - F(z)| \overset{d}{=} \sup_{0<u<1} |\Phi_n(u) - u|, \tag{18.1}$$

де $\Phi_n(u)$ – емпірична функція розподілу для вибірки з рівномірного на інтервалі $(0, 1)$ розподілу, $\overset{d}{=}$ – рівність за розподілом.

Дійсно, поклавши у (18.1) $z = F^{-1}(u)$, $0 \le u \le 1$, де $F^{-1}(u)$ – функція, обернена до $F(z)$, отримаємо

$$D_n = \sup_{0<u<1} |F_n(F^{-1}(u)) - u|.$$

Перейдемо до нових випадкових величин, використовуючи формулу

$$U_i = F(\xi_i),\ i = 1, \ldots, n.$$

Нехай $U_{(1)} \le U_{(2)} \le \ldots \le U_{(n)}$ – їхній варіаційний ряд. Функція $F(z)$ монотонна, тому $U_{(k)} = F(\xi_k)$, $k = 1, \ldots, n$ і нерівності $F^{(-1)}(u) \ge \xi_{(k)}$, $k = 1, \ldots, n$ еквівалентні нерівностям $u \ge U_{(n)}$. Оскільки $F_n(x) = \frac{1}{n}\sum\limits_{k=1}^{n} \chi(x - \xi_{(k)})$, то маємо

$$F_n(F^{(-1)}(u)) = \frac{1}{n}\sum_{k=1}^{n} \chi(F^{(-1)}(u) - \xi_{(k)}) = \frac{1}{n}\sum_{k=1}^{n} \chi(u - U_{(k)}) = \Phi_n(u).$$

Тут функція $\chi(u)$ є **функцією Гевісайда**:

$$\chi(u) = \begin{cases} 1, & \text{якщо } u \ge 0, \\ 0, & \text{якщо } u < 0. \end{cases}$$

Розподіл U_i збігається з рівномірним розподілом на $(0, 1)$ і $\Phi_n(u)$ – емпірична функція розподілу для вибірки з рівномірного розподілу на $(0, 1)$.

Цей факт дуже корисний, оскільки достатньо обчислити розподіл D_n тільки один раз, а саме для вибірки з рівномірного на $(0, 1)$ розподілу, і використовувати його для перевірки гіпотези щодо довільної неперервної функції розподілу $F(z)$.

2) Друга особливість полягає в тому, що

$$\sqrt{n} D_n \underset{n\to\infty}{\overset{\text{сл}}{\Rightarrow}} \eta,$$

де випадкова величина η має функцію розподілу Колмогорова

$$K(t) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2t^2}.$$

Цей граничний розподіл і використовується як розподіл D_n уже при $n \geq 20$.

Використаємо останній факт. За заданим рівнем значущості α підбираємо число λ_α так, що

$$P\left\{\sqrt{n}D_n \geq \lambda_\alpha | H_0\right\} \approx 1 - K(\lambda_\alpha) = \alpha.$$

У табл.4 додатка наведені значення $K(x)$ для різних x. А в табл.5 можна знайти критичні значення λ_α для розподілу Колмогорова. Базуючись на цьому, будуємо **правило перевірки гіпотези H_0**.

Нехай $\lambda_n = \sqrt{n}D_n$ – значення статистики критерію, обчислене за реалізацією вибірки $x' = (x_1, x_2 \ldots, x_n)$. Якщо $\lambda_n \geq \lambda_\alpha$, то гіпотеза відхиляється, а при $\lambda_n < \lambda_\alpha$ – приймається, тобто робиться висновок, що статистичні дані не суперечать гіпотезі.

Приклад 18.2

Нехай маємо реалізацію вибірки $x: 0.7; 2.3; 4.8; 9.7; 5.3; 6.8; 5.9; 8.7; 1.4; 3.2$. Треба перевірити гіпотезу про те, що величина, яка спостерігається, має рівномірний розподіл на $[0; 10]$ з рівнем значущості $\alpha = 0.05$:

$$H_0: \qquad F_{\xi_0}(z) = \begin{cases} 0, & \text{якщо } z < 0, \\ \frac{z}{10}, & \text{якщо } z \in [0, 10], \\ 1, & \text{якщо } z > 10. \end{cases}$$

Значення емпіричної та гіпотетичної функцій розподілу наведемо у таблиці ($n = 10$):

Значення	0.7	1.4	2.3	3.2	4.8	5.3	5.9	6.8	8.7	9.7
$F_n(z_i)$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$F_{\xi_0}(z_i)$	0.07	0.14	0.23	0.32	0.48	0.53	0.59	0.68	0.87	0.97
$\lvert F_n(z_i) - F_{\xi_0}(z_i)\rvert$	0.03	0.06	0.07	0.08	0.02	0.07	0.11	**0.12**	0.03	0.03

Значення статистики критерію

$$\lambda_n = \sqrt{n} \sup_z |F_n(z) - F_{\xi_0}(z)| = \sqrt{10} \cdot 0.12 \approx 0.38,$$

критична область визначається значенням $\lambda_\alpha = 1.358$, яке взято з табл. 5. Оскільки $\lambda_n < \lambda_\alpha$, то гіпотеза H_0 приймається.

Розв'язування даного прикладу з використання мови програмування R див. на стор. 291.

18.2.2 Критерій χ^2 К. Пірсона

Нехай, як і раніше, $\xi' = (\xi_1, ..., \xi_n)$ – вибірка з генеральної сукупності з невідомою функцією розподілу $F_{\xi_0}(z)$, про яку висунута проста гіпотеза

$$H_0 : F_{\xi_0}(z) = F(z).$$

Про властивості гіпотетичної функції $F(z)$ у даному випадку нічого не відомо, тобто цей критерій можна використовувати як для неперервних, так і для дискретних розподілів.

Задамо $E_1, E_2, \ldots, E_N$ – інтервали групування даних, що не перетинаються. Якщо спостерігається дискретна випадкова величина, то $E_1, E_2, \ldots, E_N$ – її різні значення. Нехай $\nu' = (\nu_1, \nu_2, \ldots, \nu_N)$ – вектор частот потрапляння елементів вибірки до відповідних інтервалів групування. Позначимо

$$p_i = P\{\xi_0 \in E_i | H_0\}, \quad i = \overline{1, N}.$$

Очевидно, що $M(\nu_i|H_0) = np_i$ як середнє біноміального розподілу з параметрами n і p_i. За міру відхилення емпіричних даних від їхній гіпотетичних значень візьмемо статистику

$$\hat{\chi}_n^2 = \sum_{i=1}^{N} \frac{(\nu_i - np_i)^2}{np_i}.$$

> **Теорема 18.1** Якщо $0 < p_i < 1$, $i = 1, \ldots, N$, то при $n \to \infty$ розподіл величини $\hat{\chi}_n^2$ слабко збігається до χ^2-розподілу з $(N-1)$ ступенями свободи.

Використовуючи апроксимацію розподілу статистики $\hat{\chi}_n^2$ розподілом χ^2, маємо

$$\alpha = \int\limits_{\chi^2_{1-\alpha,N-1}}^{\infty} k_{N-1}(x)dx \approx P\left\{\hat{\chi}_n^2 \geq \chi^2_{1-\alpha,N-1} | H_0\right\}.$$

Таким чином, критична область – це множина $\left\{t : t \geq \chi^2_{1-\alpha,N-1}\right\}$.

На практиці граничний розподіл $\chi^2(N-1)$ можна використовувати з непоганим наближенням уже при $n \geq 50$ і $\nu_i \geq 5$.

Критерій перевірки гіпотези H_0 будується таким чином. Обчисливши значення статистики критерію $\hat{\chi}_n^2 = \sum_{i=1}^{N} \frac{(\nu_i - np_i)^2}{np_i}$ і вибравши рівень значущості α, за таблицею значень квантилів χ^2-розподілу (табл. 7) визначимо величину $\chi^2_{1-\alpha,N-1}$ таку, що

$$P\left\{\chi^2(N-1) \geq \chi^2_{1-\alpha,N-1}\right\} = \alpha$$

(або $P\left\{\chi^2(N-1) < \chi^2_{1-\alpha,N-1}\right\} = 1 - \alpha$). Якщо $\hat{\chi}_n^2 \geq \chi^2_{1-\alpha,N-1}$, то гіпотеза H_0 відхиляється, якщо ж $\hat{\chi}_n^2 < \chi^2_{1-\alpha,N-1}$, – то приймається.

Приклад 18.3

При 50 підкиданнях монети герб з'явився 20 раз. Чи можна вважати, що монета симетрична? Прийняти $\alpha = 0.1$.

Розв'язок: Експеримент із підкиданням монети можна описати в термінах незалежних спостережень випадкової величини ξ_0, яка набуває двох значень: $x = 1$, якщо випав герб, та 0, якщо випала решка. Гіпотеза про симетричність монети в термінах розподілу ξ_0 формулюється так: розподілом ξ_0 є

$$P\{\xi = x\} = \frac{1}{2^x} \cdot \frac{1}{2^{1-x}} = \frac{1}{2}, \; x = 0, 1.$$

Імовірності потрапляння вибіркових значень у підмножини вибіркових значень, якими є дві підмножини $X_0 = \{0\}$ та $X_1 = \{1\}$, підраховані за гіпотетичним розподілом, становлять $p_0 = p_1 = 1/2$.

Підрахуємо значення статистики критерію χ^2:

$$\chi^2 = \frac{\left(30 - \frac{1}{2}50\right)^2}{\frac{1}{2}50} + \frac{\left(20 - \frac{1}{2}50\right)^2}{\frac{1}{2}50} = 2.$$

У таблиці значень χ^2-розподілу (див. табл. 7 додатка) шукаємо значення

$$\chi^2_{1-\alpha,n-1} = \chi^2_{0.9,1} = 2.71.$$

Оскільки значення статистики критерію не перевищує табличне, то гіпотеза приймається. Таким чином, гіпотеза узгоджується з даними експерименту, або, іншими словами, гіпотеза про симетричність монети не суперечить експериментальним даним.

Як розв'язати дану задачу за допомогою мови програмування R див. на стор. 292

Критерій Пірсона для перевірки гіпотези про вид розподілу у випадку наявності невідомих параметрів. Гіпотетичний розподіл, що не залежить від параметрів, на практиці зустрічається рідко. Зазвичай гіпотетичний розподіл залежить від невідомих параметрів, стосовно значень яких є лише інформація, що міститься у виборці. У такому випадку маємо задачу перевірки складних гіпотез, для яких також можна використовувати критерій χ^2, проте в дещо зкоригованій формі.

Нехай за вибіркою $\xi' = (\xi_1, \ldots, \xi_n)$ треба перевірити гіпотезу

$$H_0' : F_{\xi_0}(x) \in F,$$

де $F = \{F(x,\theta), \theta \in \Theta \subseteq \mathbb{R}^m\}$ – задане сімейство функцій розподілу. Значення параметрів, а отже, і ймовірностей $p_i(\theta)$, невідомі. Природно оцінити невідомий параметр θ за вибіркою й у статистику підставити ймовірності, підраховані через $F(x,\hat{\theta}_n)$, де $\hat{\theta}_n$ – оцінка θ. Проте в цьому випадку величини $p_i(\hat{\theta}_n)$ уже не є сталі, вони є функціями від вибірки, а отже, випадковими величинами. Окрім того, граничний розподіл статистики критерію залежить від методу побудови оцінки $\hat{\theta}_n$. Тому теорема 18.1 не може бути застосована, оскільки розподіл статистики критерію вже буде інакшим. Р. Фішер показав, що існують методи оцінювання параметра θ, за яких граничним розподілом для статистики критерію буде χ^2-розподіл, але з $(N-m-1)$ ступенями свободи, де m – розмірність вектору невідомих параметрів θ. Одним з таких методів є метод мінімуму χ^2.

Отже, нехай $\hat{\theta}^* = \hat{\theta}^*(\xi)$ – значення параметра θ, за якого статистика

$$\hat{\chi}_n^2(\theta) = \sum_{i=1}^{N} \frac{(\nu_i - np_i(\theta))^2}{np_i(\theta)}. \tag{18.2}$$

досягає мінімуму при заданій вибірці ξ.

Теорема 18.2 (Р. Фішер) Якщо справедлива гіпотеза H_0', m – розмірність невідомого параметра θ, то при фіксованому N та при $n \to \infty$ розподіл величини $\hat{\chi}_n^2(\hat{\theta}^*)$ слабко збігається до χ^2-розподілу з $(N-m-1)$ ступенями свободи

$$\hat{\chi}_n^2(\hat{\theta}^*) = \sum_{i=1}^{N} \frac{(\nu_i - np_i(\hat{\theta}^*))^2}{np_i(\hat{\theta}^*)} \Rightarrow \chi^2(N-m-1).$$

Далі критерій перевірки гіпотези H_0' будується аналогічно наведеному вище алгоритму. Якщо значення статистики критерію, обчислене за реалізацією вибірки, $\hat{\chi}_n^2(\hat{\theta}^*) \geq \chi^2_{1-\alpha,N-m-1}$, то гіпотеза H_0' відхиляється, якщо ж $\hat{\chi}_n^2(\hat{\theta}^*) < \chi^2_{1-\alpha,N-m-1}$, то немає підстав відхилити гіпотезу, отже, H_0' приймається.

Зауваження 1. Обчислення точки мінімуму функції (18.2) в загальному випадку можливе лише чисельно. Проте доведено, що оцінки параметрів λ розподілу Пуассона та μ нормального розподілу близькі до вибіркового середнього, а оцінкою σ^2 нормального розподілу є вибіркова дисперсія $\bar{S}^2 = \frac{1}{n}\sum_{i=1}^{N} \nu_i(\xi_i - \bar{\xi})^2$, обчислені по групованій вибірці.

Зауваження 2. На практиці можливий випадок, коли вибірка згрупована, проте доступними також є значення $x_1, ..., x_n$. В такій ситуації можна оцінити невідомі параметри більш ефективно, скажімо, за допомогою методу максимальної вірогідності, на базі значень $x_1, ..., x_n$ вибірки. При використанні таких оцінок граничний розподіл статистики критерію знаходиться між χ^2-розподілом з $N-1$ ступенем свободи та χ^2-розподілом з $N-m-1$

ступенем свободи. Кілька "втрачених" ступенів свободи повертаються завдяки ефективному оцінюванню невідомих параметрів.

Приклад 18.4 [18], ст. 160.

За час другої світової війни в південній частині Лондона впало 535 балістичних ракет. Уся ця територія була розділена на 576 ділянок площею по 0.25 км^2. Нижче наведені числа ділянок n_k на які впало k ракет

k	0	1	2	3	4	5
n_k	229	211	93	35	7	1

Чи узгоджуються ці дані з гіпотезою про те, що число ракет, які впали на кожну ділянку, мають розподіл Пуассона? Прийняти $\alpha = 0.05$.

Розв'язок: Маємо $n = \sum\limits_{k=0}^{5} n_k = 576$ незалежних спостережень випадкової величини ξ_0 – кількості балістичних ракет, що впали на ділянку.

Відносно невідомого розподілу випадкової величини ξ_0 висувається гіпотеза

$$H_0 : P\{\xi_0 = k\} = \frac{\lambda^k}{k!} e^{-\lambda},\ k = 0, 1, \dots,$$

яку треба перевірити.

Параметр λ гіпотетичного розподілу невідомий. За його оцінку ми можемо взяти вибіркове середнє:

$$\hat{\lambda} = \frac{1}{576}(0 \cdot 229 + 1 \cdot 211 + 2 \cdot 93 + 3 \cdot 35 + 4 \cdot 7 + 5 \cdot 1) \approx 0.9,$$

отже, гіпотетичний розподіл має вигляд

$$P\{\xi_0 = k\} = \frac{0.9^k}{k!} e^{-0.9},\ k = 0, 1, \dots.$$

Користуючись наведеною вище методикою, шукаємо значення статистики критерію $\hat{\chi}^2_n$:

$$\hat{\chi}^2_n = \frac{(229 - 576e^{-0.9})^2}{576e^{-0.9}} + \frac{(211 - 576 \cdot 0.9e^{-0.9})^2}{576 \cdot 0.9e^{-0.9}} +$$
$$+ \frac{(93 - 576\frac{(0.9)^2}{2!}e^{-0.9})^2}{576\frac{(0.9)^2}{2!}e^{-0.9}} + \frac{(35 - 576\frac{(0.9)^3}{3!}e^{-0.9})^2}{576\frac{(0.9)^3}{3!}e^{-0.9}} +$$
$$+ \frac{\left(7 - 576\frac{(0.9)^4}{4!}e^{-0.9}\right)^2}{576\frac{(0.9)^4}{4!}e^{-0.9}} + \frac{\left(1 - 576\left(1 - \sum\limits_{i=0}^{4}\frac{(0.9)^i}{i!}e^{-0.9}\right)\right)^2}{576\left(1 - \sum\limits_{i=0}^{4}\frac{(0.9)^i}{i!}e^{-0.9}\right)} \approx 1.171.$$

Отримане значення порівнюємо з табличним значенням $\chi^2_{1-\alpha,N-1-m}$, де m – кількість параметрів, оцінених за вибіркою, N – кількість підмножин, на які розбито вибірковий простір. Маємо $\chi^2_{1-\alpha,N-2} = \chi^2_{0,95;4} = 9.49$. Таким чином, експериментальні дані узгоджуються з гіпотезою.

Приклад розв'язування цієї задачі за допомогою R можна знайти на стор. 293.

Приклад 18.5 Перевірка узгодженості з нормальним розподілом (параметри невідомі)

За спостереженнями, наведеними у таблиці, за допомогою критерію χ^2 з рівнем значущості $\alpha = 0.05$ перевірити гіпотезу про те, що випадкова величина має нормальний розподіл.

Інтервал	[-4;0)	[0;2)	[2;4)	[4;6)
ν_i	20	40	30	10

Розв'язок. Маємо груповану вибірку з $N = 4$ інтервали групування. Ймовірності того, що випадкова величина, яка має нормальний розподіл з параметрами (μ, σ^2), набуде значення з i-го інтервалу групування $[z_{i-1}; z_i)$, $i = 1, 2, 3, 4$, може бути обчислена за допомогою функції розподілу стандартного нормального розподілу $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$, значення якої наведені в таблиці 3 додатку:

$$p_i = \mathrm{P}\{\xi_0 \in [z_{i-1}; z_i)\} = \int_{z_{i-1}}^{z_i} \phi_{\mu,\sigma^2}(t)dt = \Phi\left(\frac{z_i - \mu}{\sigma}\right) - \Phi\left(\frac{z_{i-1} - \mu}{\sigma}\right)$$

Для обчислення цих ймовірностей нам будуть потрібні оцінки параметрів $\hat{\mu}$ та $\hat{\sigma}^2$, обчислені згідно зауваженню 2. Тоді величини

$$\hat{p}_i = \Phi\left(\frac{z_i - \hat{\mu}}{\hat{\sigma}}\right) - \Phi\left(\frac{z_{i-1} - \hat{\mu}}{\hat{\sigma}}\right)$$

можуть бути обчислені на основі інформації, що міститься у виборці, і статистика критерію χ^2 обчислена за формулою

$$\hat{\chi}^2_n(\hat{\mu}, \hat{\sigma}^2) = \sum_{i=1}^{4} \frac{(\nu_i - n\hat{p}_i)^2}{n\hat{p}_i} \Rightarrow \chi^2(4 - 2 - 1) = \chi^2(1).$$

Гіпотеза про згоду з нормальним розподілом має бути відхилена, якщо обчислене значення цієї статистики критерію буде більшим за відповідне критичне значення для χ^2-розподілу з одним ступенем свободи.

Отже, обчислимо значення оцінок параметрів нормального розподілу:

$$\hat{\mu} = \bar{x} = \frac{1}{100}(-40 + 40 + 90 + 50) = 1.4;$$

$$\hat{\sigma}^2 = \bar{S}^2 = \frac{1}{n}\sum_{i=1}^{N} \nu_i (x_i^* - \bar{x})^2 =$$

$$= \frac{1}{100}\left((-2-1.4)^2 \cdot 20 + (1-1.4)^2 \cdot 40 + (3-1.4)^2 \cdot 30 + (5-1.4)^2 \cdot 10\right) =$$

$$= 4.48, \quad \hat{\sigma} = \bar{S} \approx 2.1.$$

Тут для обчислень в якості типового значення кожного інтервалу береться x_i^* – його середина.

Тепер підрахуємо оцінки ймовірностей $\hat{p}_i$, $i = \overline{1,4}$, користуючись таблицями для функції розподілу стандартного нормального закону:

$$\hat{p}_1 = \mathrm{P}\{-4 \le \xi_0 < 0\} = \mathrm{P}\left\{\frac{-4-1.4}{2.1} \le \frac{\xi_0 - \hat{\mu}}{\hat{\sigma}} < \frac{0-1.4}{2.1}\right\} =$$

$$= \Phi(-0.66) - \Phi(-2.57) = 0.2546 - 0.0051 = 0.2495;$$

$$\hat{p}_2 = \mathrm{P}\{0 \le \xi_0 < 2\} = \mathrm{P}\left\{\frac{0-1.4}{2.1} \le \frac{\xi_0 - \hat{\mu}}{\hat{\sigma}} < \frac{2-1.4}{2.1}\right\} =$$

$$= \Phi(0.29) - \Phi(-0.66) = 0.6141 - 0.2546 = 0.3595;$$

$$\hat{p}_3 = \mathrm{P}\{2 \le \xi_0 < 4\} = \mathrm{P}\left\{\frac{2-1.4}{2.1} \le \frac{\xi_0 - \hat{\mu}}{\hat{\sigma}} < \frac{4-1.4}{2.1}\right\} =$$

$$= \Phi(1.24) - \Phi(0.29) = 0.8925 - 0.6141 = 0.2784;$$

$$\hat{p}_4 = \mathrm{P}\{4 \le \xi_0 < 6\} = \mathrm{P}\left\{\frac{4-1.4}{2.1} \le \frac{\xi_0 - \hat{\mu}}{\hat{\sigma}} < \frac{6-1.4}{2.1}\right\} =$$

$$= \Phi(2.19) - \Phi(1.24) = 0.9857 - 0.8925 = 0.0932.$$

Обчислені результати зручно занести в таблицю:

Інтервал	[-4;0)	[0;2)	[2;4)	[4;6)
ν_i	20	40	30	10
$\hat{p}_i$	0.2495	0.3595	0.2784	0.0932
$n\hat{p}_i$	24.95	35.95	27.84	9.32
$(\nu_i - n\hat{p}_i)^2$	24.5	16.4	4.67	0.64
$\frac{(\nu_i - n\hat{p}_i)^2}{n\hat{p}_i}$	0.98	0.46	0.17	0.07

Значення статистики критерію є сумою нижнього рядка таблиці: $\hat{\chi}_n^2 = 1.68$. Кількість інтервалів $N = 4$, кількість невідомих параметрів $m = 2$, отже кількість ступенів свободи граничного χ^2-розподілу $4 - 2 - 1 = 1$. В таблиці 7 знаходимо критичне значення $\chi^2_{0.95;1} = 3.841$. Оскільки $1.68 < 3.841$, то ми не маємо підстав відхилити нульову гіпотезу про те, що спостережувана величина має нормальний розподіл. Отже, гіпотеза про узгодження з нормальним розподілом приймається.

18.3 Гіпотези однорідності

Нехай є дві незалежні вибірки $\xi' = (\xi_1, \ldots, \xi_n)$ і $\eta' = (\eta_1, \ldots, \eta_m)$, які описують одне й те саме явище, однак отримані вони в різний час і в різних умовах. Треба перевірити чи будуть ці вибірки з одного розподілу, чи закон розподілу від вибірки до вибірки змінювався.

Подібна задача може виникнути при контролі якості деякої продукції. У загальному вигляді її можна сформулювати так.

Нехай $\xi' = (\xi_1, \ldots, \xi_n)$ – вибірка з деякою невідомою функцією розподілу $F_1(x)$, а $\eta' = (\eta_1, \ldots, \eta_m)$ – вибірка з деякою невідомою функцією розподілу $F_2(x)$. Треба перевірити гіпотезу однорідності

$$H_0 : F_1(x) = F_2(x).$$

18.3.1 Критерій Смірнова-Колмогорова

Одним з критеріїв перевірки гіпотези однорідності є критерій Смирнова–Колмогорова, який застосовують у випадку неперервних розподілів. Цей критерій базується на статистиці $D_{nm} = \sup_{-\infty<x<\infty} |F_{1n}(x) - F_{2m}(x)|$, де $F_{1n}(x)$ і $F_{2m}(x)$ – емпіричні функції розподілу, побудовані за першою та другою вибірками.

Критичну границю знаходять на основі відомого при гіпотезі H_0 граничного розподілу статистики D_{nm}.

Теорема 18.3 Нехай $F_{1n}(x)$ і $F_{2m}(x)$ – дві емпіричні функції розподілу, які побудовані на основі двох незалежних вибірок об'ємів n і m одного і того ж розподілу, який має неперервну функцію розподілу $F(x)$. Тоді для довільного фіксованого $t > 0$

$$\lim_{n,m\to\infty} P\left\{\sqrt{nm/(n+m)}D_{nm} \le t\right\} = K(t) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2t^2}.$$

Відповідно за міру розбіжності приймається величина

$$\lambda_{nm} = \sqrt{\frac{nm}{n+m}} D_{nm} = \sqrt{\frac{nm}{n+m}} \sup_{-\infty<x<\infty} |F_{1n}(x) - F_{2m}(x)|,$$

розподіл якої збігається, згідно з теоремою, до розподілу Колмогорова. Далі критерій будується аналогічно критерію Колмогорова: за заданим рівнем значущості α знайдемо за таблицею критичне значення λ_α (табл. 4 додатка). Якщо $\lambda_{nm} \geq \lambda_\alpha$, то гіпотеза H_0 відхиляється, якщо ж $\lambda_{nm} < \lambda_\alpha$ - приймається.

18.3.2 Критерій однорідності χ^2

Цей критерій можна використовувати для перевірки дискретних даних. Окрім того, за його допомогою можна перевіряти однорідність будь-якої скінченної кількості вибірок (критерій Смірнова-Колмогорова може аналізувати лише дві вибірки).

Припустимо, що виконано k послідовних серій незалежних спостережень, які включають $n_1, \ldots, n_k$ спостережень відповідно. При цьому в кожному експерименті може з'являтися один із s результатів.

Нехай ν_{ij} – число реалізацій i-го результату в j–ій серії, так що $\sum\limits_{i=1}^{s} \nu_{ij} = n_j$, $j = 1, \ldots, k$, $n_1 + n_2 + \cdots + n_k = n$ - загальний обсяг спостережень. Треба перевірити гіпотезу H_0, що всі спостереження були над одною і тою самою випадковою величиною.

Так як $M\{\nu_{ij}|H_0\} = n_j p_i$, то, спираючись на принцип χ^2, за міру відхилення емпіричних даних від їхній гіпотетичних значень у даному випадку слід брати статистику

$$\hat{\chi}_n^2 = \sum_{i=1}^{s} \sum_{j=1}^{k} \frac{\left(\nu_{ij} - n_j \frac{\nu_{i\bullet}}{n}\right)^2}{n_j \frac{\nu_{i\bullet}}{n}}, \quad \nu_{i\bullet} = \sum_{j=1}^{k} \nu_{ij}.$$

Для знаходження критичної границі застосовують таку граничну теорему:

$$\chi_n^2 \overset{\text{сл.}}{\underset{n\to\infty}{\Rightarrow}} \chi^2((s-1)(k-1)).$$

У таблиці χ^2-розподілу (табл. 7 додатка) за заданим рівнем значущості α та кількістю ступенів свободи $(s-1)(k-1)$ знаходимо число $\chi^2_{1-\alpha,(s-1)(k-1)}$ таке, що

$$P\left\{\chi^2((s-1)(k-1)) \geq \chi^2_{1-\alpha,(s-1)(k-1)}\right\} = \alpha$$

або $P\left\{\chi^2((s-1)(k-1)) < \chi^2_{1-\alpha,(s-1)(k-1)}\right\} = 1 - \alpha$. Якщо значення статистики критерію

$$\hat{\chi}_n^2 \geq \chi^2_{1-\alpha,(s-1)(k-1)},$$

то гіпотеза H_0 відхиляється. У протилежному випадку – приймається.

Приклад 18.6

За допомогою критерію χ^2 перевірити гіпотезу однорідності двох вибірок при $\alpha = 0.05$.

x_i	1	2	3	4
ν_{i1}	40	26	24	10
ν_{i2}	30	20	30	20
$\nu_{i\bullet}$	70	46	54	30

Розв'язок: $n_1 = 100$, $n_2 = 100$, $n = 200$.

$$\chi_n^2 = \frac{(40-35)^2}{35} + \frac{(26-23)^2}{23} + \frac{(24-27)^2}{27} + \frac{(10-15)^2}{15} + \frac{(30-35)^2}{35} +$$

$$+\frac{(20-23)^2}{23} + \frac{(30-27)^2}{27} + \frac{(20-15)^2}{15} = 6.2;$$

$$\chi^2_{0.95,3} = 7.815.$$

Оскільки $6.2 < 7.815$, то гіпотеза приймається.

18.3.3 Гіпотези незалежності. Критерій незалежності χ^2

Є n незалежних спостережень $(\xi_1, \eta_1), \ldots, (\xi_n, \eta_n)$ випадкової величини (ξ_0, η_0) з невідомою функцією розподілу $F_{\xi_0,\eta_0}(x, y)$, для якої треба перевірити гіпотезу

$$H_0 : F_{\xi_0,\eta_0}(x, y) = F_{\xi_0}(x) F_{\eta_0}(y),$$

де $F_{\xi_0}(\cdot)$, $F_{\eta_0}(\cdot)$ - одновимірні функції розподілу.

Будемо припускати, що випадкова величина ξ_0 приймає скінченне число значень s, які будемо позначати літерами $a_1, \ldots, a_s$, а друга компонента η_0 - k значень $b_1, \ldots, b_k$.

Якщо модель має іншу структуру, то групують усі можливі значення випадкових величин окремо по першій і другій компонентам. В цьому випадку множина значень ξ_0 розбивається на s інтервалів $E_1^{(1)}, \ldots, E_s^{(1)}$, а множина значень η_0 на k інтервалів $E_1^{(2)}, \ldots, E_k^{(2)}$, множина значень вектора (ξ_0, η_0) на $N = s \cdot k$ прямокутників $E_i^{(1)} \times E_j^{(2)}$.

Позначимо через ν_{ij} - число спостережень пари (a_i, b_j) (або число елементів вибірки, які належать прямокутнику $E_i^{(1)} \times E_j^{(2)}$, якщо дані групуються), так що $\sum\limits_{i=1}^{s} \sum\limits_{j=1}^{k} \nu_{ij} = n$.

Результати спостережень зручно подати у вигляді таблиці спряженості двох ознак:

	b_1	$\dots$	b_k	Сума
a_1	ν_{11}	$\dots$	ν_{1k}	$\nu_{1\bullet}$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
a_s	ν_{s1}	$\dots$	ν_{sk}	$\nu_{s\bullet}$
Сума	$\nu_{\bullet 1}$	$\dots$	$\nu_{\bullet k}$	n

Відстань між емпіричними даними і їхніми гіпотетичними значеннями має вигляд

$$\hat{\chi}_n^2 = \sum_{i=1}^{s} \sum_{j=1}^{k} \frac{\left(\nu_{ij} - n\frac{\nu_{i\bullet}}{n}\frac{\nu_{\bullet j}}{n}\right)^2}{n_j \frac{\nu_{i\bullet}}{n}\frac{\nu_{\bullet j}}{n}}, \ \nu_{\bullet j} = \sum_{i=1}^{s} \nu_{ij}, \ \nu_{i\bullet} = \sum_{j=1}^{k} \nu_{ij}.$$

При $n \to \infty$ розподіл відхилення $\hat{\chi}_n^2$ збігається до χ^2-розподілу із $(s-1)(k-1)$ ступенями свободи.

Вибір табличного значення $\chi^2_{1-\alpha,(s-1)(k-1)}$ і прийняття рішення про допустимість гіпотези робиться аналогічно описаній вище процедурі для критерію однорідності χ^2. При цьому з імовірністю α гіпотеза H_0 буде відхилятися, коли вона вірна.

Приклад 18.7

Нижче наведені результати опитування 100 студентів перших трьох курсів, яким ставилося одне питання: “Чи вважаєте ви, що куріння заважає навчанню?”

З'ясувати чи підтведжують ці дані припущення про те, що відношення до куріння студентів на різних курсах різне? Прийняти $\alpha = 0.05$.

Відповідь/Курс	1	2	3	Сума
Так	-	30	25	55
Не знаю	8	5	7	20
Ні	15	10	-	25
Сума	23	45	32	100

Розв'язок:

$$\hat{\chi}_n^2 = \frac{\left(0 - 100\frac{23}{100}\frac{55}{100}\right)^2}{100\frac{23}{100}\frac{55}{100}} + \frac{\left(30 - 100\frac{55}{100}\frac{45}{100}\right)^2}{100\frac{55}{100}\frac{45}{100}} + \frac{\left(25 - 100\frac{55}{100}\frac{32}{100}\right)^2}{100\frac{55}{100}\frac{32}{100}} + \dots = 44.2$$

$$\chi^2_{1-\alpha,(s-1)(k-1)} = \chi^2_{0.95,4} \approx 9.49.$$

Гіпотеза про незалежність відхиляється і сформульована теза приймається.

Приклад розв'язування цієї задачі за допомогою R можна знайти на стор. 294.

Задача 18.1. За спостереженнями, наведеними у таблиці, за допомогою критерію χ^2 перевірити гіпотезу, що випадкова величина має пуассонівський розподіл.

а) $\alpha = 0.05$,

x_i	0	1	2	3	4
m_i	109	65	22	3	1

б) $\alpha = 0.05$,

x_i	0	1	2	3	4	5	6	7
m_i	112	168	120	61	32	5	1	1

в) $\alpha = 0.05$,

x_i	0	1	2	3	4	5
m_i	229	211	93	35	7	1

г)$\alpha = 0.01$,

x_i	0	1	2	3	4	5	6	7
m_i	8	17	16	10	6	2	0	1

д) $\alpha = 0.1$,

x_i	0	1	2	3	4	5
m_i	376	100	81	35	7	1

Задача 18.2. За спостереженнями, наведеними у таблиці, за допомогою критерію χ^2 перевірити згоду з рівномірним розподілом. У першому рядку таблиці вказана ліва границя інтервалу (i – номер інтервалу $[i, i+1)$).

а) $\alpha = 0.05$,

x_i	0	1	2	3	4	5	6	7	8	9	10	11
m_i	48	42	36	54	39	43	41	33	37	41	47	39

б) $\alpha = 0.1$,

x_i	0	1	2	3	4	5	6	7	8	9
m_i	69	89	83	79	80	73	77	75	76	91

в) $\alpha = 0.01$,

x_i	0	1	2	3	4	5	6	7	8	9
m_i	16	15	19	13	14	19	14	12	17	13

Задача 18.3. За спостереженнями, наведеними у таблиці, за допомогою критерію χ^2 перевірити згоду з нормальним розподілом.

а) $\alpha = 0.05$,

Інтервал	[0;5)	[5;10)	[10;15)	[15;20)	[20;25)
m_i	15	75	100	50	10

б) $\alpha = 0.01$,

Інтервал	[3.0;3.6)	[3.6;4.2)	[4.2;4.8)	[4.8;5.4)	[5.4;6.0)	[6.0;6.6)
m_i	2	8	35	43	22	10

в) $\alpha = 0.05$,

Інтервал	[-3;-1)	[-1;0)	[0;1)	[1;2)	[2;3)	[3;5)
m_i	13	15	24	25	13	10

,

г) $\alpha = 0.1$,

Інтервал	[-8;-2)	[-2;4)	[4;10)	[10;16)
m_i	10	50	30	10

ґ) $\alpha = 0.05$,

Інтервал	[-4;0)	[0;2)	[2;4)	[4; 6)
m_i	25	40	25	10

Задача 18.4. Для наступних 50 реалізацій випадкової величини за критерієм Колмогорова-Смірнова або χ^2 перевірити гіпотезу про рівномірний розподіл на проміжку $[1; 3]$ на рівні значущості 0.1.

2.03	2.25	2.94	2.30	1.00	2.18	1.93	1.60	1.52	2.42
2.32	1.43	1.79	2.07	1.89	1.49	1.31	2.58	2.17	1.53
2.55	2.46	2.65	1.68	1.81	1.21	2.34	2.00	1.35	2.53
2.49	1.30	2.79	2.76	2.60	1.25	1.71	2.57	1.70	1.65
1.58	1.93	2.84	1.03	2.85	1.25	2.85	2.45	1.37	1.90

Задача 18.5. Для наступних 50 реалізацій випадкової величини за критерієм χ^2 перевірити а) гіпотезу про пуассонівський розподіл, б) гіпотезу про пуассонівський розподіл з параметром $\lambda = 7$ з надійністю 0,9.

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

Задача 18.6. Для наступних 50 реалізацій випадкової величини за критерієм χ^2 перевірити а) гіпотезу про нормальний розподіл, б) гіпотезу про нормальний розподіл $N(2; 1)$ з надійністю 0.95.

2.30	2.04	3.62	2.21	1.82	2.82	2.77	1.44	1.72	2.69
1.72	4.08	2.11	0.91	2.82	1.82	2.91	0.14	0.76	1.45
1.59	1.80	2.33	2.25	1.76	3.31	1.92	3.32	0.60	1.81
2.96	1.33	3.01	-1.23	0.76	2.19	2.82	2.95	3.22	2.48
1.56	5.06	0.61	1.83	2.04	1.57	1.49	1.07	3.06	1.77

Задача 18.7. Для наступних 50 реалізацій випадкової величини за критерієм χ^2 перевірити: а) гіпотезу про логнормальний розподіл, б) гіпотезу про логнормальний розподіл $logN(2;1)$ з надійністю 0.9.

9.68	4.31	27.89	5.67	12.62	0.68	16.79	7.05	9.10	6.59
10.87	2.07	4.84	6.65	3.42	25.28	3.40	61.42	3.51	5.53
26.68	9.85	13.96	1.34	5.72	3.12	17.00	17.03	2.79	8.36
33.51	7.21	129.97	47.03	3.22	15.63	7.78	22.26	5.26	8.26
1.53	33.77	0.87	2.91	19.83	36.54	4.93	5.64	3.44	2.86

Задача 18.8. Для наступних 50 реалізацій випадкової величини за критерієм χ^2 перевірити: а) гіпотезу про геометричний розподіл, б) гіпотезу про геометричний розподіл з параметром $p = \frac{1}{3}$ з надійністю 0.95.

0	7	0	8	0	4	0	0	0	2	1	0	0	2	1	0	2	2	0	0
1	0	3	10	1	0	4	3	1	0	1	3	0	1	2	0	0	0	0	3
0	1	6	0	4	6	4	3	2	0										

Задача 18.9. Для наступних 50 реалізацій випадкової величини за критерієм χ^2 перевірити: а) гіпотезу про експоненціальний розподіл; б) гіпотезу про експоненціальний розподіл з параметром $\lambda = 2$ з надійністю 0.95.

0.57	0.95	0.15	0.40	0.56	0.29	1.64	0.18	0.62	0.00
0.92	0.43	0.16	1.22	0.56	0.21	0.02	0.22	0.15	0.49
0.27	0.38	0.05	0.31	0.20	0.09	0.28	0.05	0.30	0.32
0.41	0.49	0.81	0.33	0.71	1.59	0.58	0.59	0.18	0.14
0.04	0.07	0.03	0.45	0.16	0.78	0.25	0.08	0.02	0.31

Задача 18.10. Рентгенівське випромінення викликає в органічних клітинах певну перебудову хромосом. В таблиці наведено результати експеримента (підраховувалась кількість перебудов хромосом під впливом рентгенівських променів).

i	0	1	2	3	4 і більше	Всього
n_i	434	195	44	9	0	682

Тут i – кількість змін в клітині, n_i – кількість клітин зі змінами. Чи узгоджується з наведеними даними гіпотеза про пуассонівський розподіл кількості перебудов у клітині?

Задача 18.11. З продукції двох верстатів зробили дві вибірки по 40 виробів:

Розмір деталі	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56
$m_i^{(1)}$	2	1	2	2	1	1	4	1	1	2	5	0	0	6	4	3	5
$m_i^{(2)}$	2	0	0	2	3	0	1	6	0	5	3	1	5	3	3	2	4

Перевірити, використовуючи критерій Смірнова-Колмогорова, гіпотезу про те, що ці вибірки належать одній і тій самій генеральній сукупності при рівні значущості $\alpha = 0.1$.

Задача 18.12. У першому потоці з 300 абітурієнтів оцінку "2" отримало 33 особи, "3" – 43 особи, "4" – 80 осіб, "5" – 144. У другому потоці інші 300 абітурієнтів мали такий результат: "2" – 39 осіб, "3" – 35, "4" – 72, "5" – 154. Чи можна вважати обидва потоки однорідними при рівні значущості 0.05?

Задача 18.13. У таблиці наведено результати обстеження 697 школярів. Хлопці були впорядковані за рівнем IQ та відповідно до умов їхнього проживання вдома. При цьому використано позначення: A – дуже здібний, B – досить здібний, C – має середні здібності, D – недостатньо розвинутий, E – розумово відсталий. Чи можна вважати, що умови життя (забезпеченість) дітей впливають на їхні здібності?

Забезпеченість	Здібність хлопців					**Всього**
	A	B	C	D	E	
Хороша	33	137	125	47	8	**350**
Погана	21	127	129	61	9	**347**
Всього	**54**	**264**	**254**	**108**	**17**	**697**

Задача 18.14. За допомогою критерія χ^2 для $\alpha = 0.05$ перевірити гіпотезу однорідності двох вибірок, наведених у таблиці.

x_i	1	2	3	4	5	6	7	8
$m_i^{(1)}$	4	4	15	51	22	3	1	0
$m_i^{(2)}$	1	1	8	43	34	7	3	3

Задача 18.15. Двовимірна випадкова величина (ξ_0, η_0) може приймати 4 значення: (0;0), (0;1), (1;0), (1;1). 180 незалежних спостережень дали такі результати: значення (0;0) з'явилося 39 раз, (0;1) – 50, (1;0) – 53, (1;1) – 38. Чи можна вважати, що ξ_0 і η_0 – незалежні? Рівень значущості прийняти $\alpha = 0.05$.

Задача 18.16. Проведено 200 спостережень над випадковими величинами ξ_0 та η_0, які приймають значення 1, 2 та 1, 2, 3 відповідно. Результати спостережень наведені у таблиці:

ξ_0/η_0	1	2	3	$\nu_{i\bullet}$
1	25	50	25	100
2	51	42	7	100
$\nu_{\bullet j}$	76	92	32	200

Перевірити за допомогою критерію χ^2, чи будуть незалежними випадкові величини ξ_0 та η_0 при $\alpha = 0.05$.

Задача 18.17. Серед 300 осіб, які поступали в університет, 97 мали оцінку "5" в школі, 48 отримали "5" на вступних іспитах по тому ж самому предмету, причому лише 18 осіб мали "5" і в школі, і на вступних іспитах. На рівні значущості 0.1 перевірити гіпотезу незалежності оцінок "5" в школі і на вступних іспитах.

Задача 18.18. В таблиці (Грінвуд, Юл, 1915р.) наведено дані про 818 випадків, класифікованих за двома ознаками: наявність щеплення проти холери та відсутність захворювання. Чи можна на основі цих даних прийти до висновку про залежність між відсутністю захворювання та наявністю щеплення?

Наявність щеплення	Наявність захворювання		
	Не захворіли	Захворіли	**Всього**
Щеплені	276	3	**279**
Не щеплені	473	66	**539**
Всього	**749**	**69**	**818**

19 ПАРАМЕТРИЧНІ ГІПОТЕЗИ

19.1 Поняття параметричної гіпотези

Нехай $\xi' = (\xi_1, ..., \xi_n)$ – незалежні спостереження випадкової величини ξ_0. **Параметричні гіпотези** – це гіпотези про справжнє значення невідомого параметра, який визначає сімейство розподілів $F_{\xi_0}(x) \in \mathscr{F} = \{F(x, \theta), \theta \in \Theta\}$, $\theta' = (\theta_1, \theta_2, \ldots, \theta_r) \in \Theta \in R^r$.

В загальному випадку параметричну гіпотезу можна подати наступним чином: $H_0 : \theta \in \Theta_0$. Альтернативна гіпотеза має вигляд $H_1 : \theta \in \Theta_1 = \Theta \setminus \Theta_0$. Точки $\theta \in \Theta_1$ називають **альтернативами**. Якщо множина Θ_0 складається з однієї точки, то гіпотеза H_0 називається **простою**, у протилежному випадку – **складною**. За вибіркою $\xi' = (\xi_1, ..., \xi_n)$ треба перевірити, чи вірна гіпотеза H_0 відносно альтернативи H_1, чи ні.

Приклади параметричних гіпотез:

1) $H_0 : \theta = \theta_0$, де $\theta_0 \in \Theta$ – деяке фіксоване значення параметра;

2) $H_0 : \theta_1 = \theta_2 = \ldots = \theta_r$;

3) $H_0 : g(\theta) = g_0$, де g_0 – фіксоване значення, а $g(\theta)$ – функція параметра θ.

Тут 1) – проста гіпотеза; 2) – складна; 3) – може бути і простою, і складною.

Приклад 19.1

Розглянемо нормальний розподіл $\mathscr{F} = N(\theta_1, \theta_2^2)$. Тоді

H_0: $\theta_1 = \theta_{10}$, $\theta_2 = \theta_{20}$ – проста гіпотеза;

H_0: $\theta_1 = \theta_{10}$ – складна гіпотеза.

19.2 Критерії перевірки гіпотези

Для перевірки сформульованої гіпотези $H_0 : \theta \in \Theta_0$ потрібен критерій (правило), який давав би можливість для кожної реалізації x вибірки ξ прийняти одне з двох рішень: прийняти гіпотезу H_0 (відхилити H_1) або відхилити її (прийняти H_1). В зв'язку з цим вибірковий простір X розбивається на дві підмножини X_0 і X_1:

$$X_0 \cup X_1 = X, \quad X_0 \cap X_1 = \varnothing.$$

При $x \in X_0$ приймається H_0, при $x \in X_1$ приймається H_1 (тобто H_0 відхиляється). X_0 - область прийняття (ухвалення) гіпотези. X_1 - область її відхилення (критична область).

Отже, критерій повністю визначається заданням критичної області X_1. Його часто називають **X_1-критерієм**.

Загальний принцип вибору критичної області критерію. При виборі критичної області треба мати на увазі, що, приймаючи або відхиляючи гіпотезу, можна допустити похибки двох видів.

Похибка першого роду: відхилення H_0, коли вона справедлива. Будемо позначати ймовірність цієї похибки через $\alpha = P\left(H_1|H_0\right)$ і називати **рівнем значущості критерію**.

Похибка другого роду: прийняти H_0, коли вона несправедлива. $\beta = P\left(H_0|H_1\right)$– імовірність похибки другого роду. Імовірність $1-\beta$ називається **потужністю критерію**.

Бажано було б побудувати такий критерій перевірки гіпотези, щоб імовірності α та β були мінімальними. Очевидно, що за заданої кількості випробувань (спостережень) n одночасно зменшити похибки першого та другого роду неможливо.

Введемо $W(\theta) = W(X_1;\theta) = P_\theta\left(\xi \in X_1\right)$, $\theta \in \Theta$ - функцію потужності критерія X_1. $P(H_1|H_0) = W(\theta)$, $\theta \in \Theta_0$ - імовірність похибки першого роду; $P(H_0|H_1) = 1 - W(\theta)$, $\theta \in \Theta_1$ - імовірність похибки другого роду.

Раціональний принцип: при заданому числі випробувань n фіксується ймовірність похибки першого роду. При цьому обирається та критична область X_1, для якої імовірність похибки другого роду мінімальна. Таким чином при фіксованому α обирається така критична область X_1, що

$$\begin{cases} W(\theta) \le \alpha, & \text{для усіх } \theta \in \Theta_0, \\ 1 - W(\theta) \to \min, & \text{для усіх } \theta \in \Theta_1; \end{cases}$$

або

$$\begin{cases} W(\theta) \le \alpha, & \text{для усіх } \theta \in \Theta_0, \\ W(\theta) \to \max, & \text{для усіх } \theta \in \Theta_1. \end{cases}$$

Зазвичай $\alpha = 0.005; 0.01; 0.05$.

Похибки другого роду на практиці призводять до більших втрат, ніж похибки першого роду. Це слідураховувати, обираючи гіпотезу H_0 або H_1.

Нехай $X_{1\alpha}$ і $X^*_{1\alpha}$ - два критерії однакового рівня значущості α для гіпотези H_0. Якщо

$$W(X^*_{1\alpha};\theta) \le W(X_{1\alpha};\theta) \text{ для усіх } \theta \in \Theta_0$$

і

$$W(X^*_{1\alpha};\theta) \ge W(X_{1\alpha};\theta) \text{ для усіх } \theta \in \Theta_1$$

причому строга нерівність у останньому виразі має місце хоча б при одному значенні θ, то говорять, що критерій $X^*_{1\alpha}$ рівномірно більш потужний, ніж критерій $X_{1\alpha}$.

Якщо наведені співвідношення виконуються для довільного критерія $X_{1\alpha}$, то $X^*_{1\alpha}$ називають рівномірно найбільш потужним критерієм для перевірки гіпотези H_0.

Введемо клас незсунених критеріїв:

$$W(\theta) \le \alpha, \text{ для усіх } \theta \in \Theta_0; W(\theta) \ge \alpha, \text{ для усіх } \theta \in \Theta_1.$$

У деяких задачах, для яких рівномірно найбільш потужні критерії не існують, можуть мати місце рівномірно найбільш потужні незсунені критерії.

Часто критична область X_1 задається у вигляді:

$$X_1 = \{x : T(x) \ge c\}.$$

В цьому випадку $T(\xi)$ - **статистика критерію**.

19.3 Вибір із двох простих гіпотез. Критерій Неймана-Пірсона

Розглянемо випадок $\Theta = \{\theta_0, \theta_1\}$,

$$H_0 : \theta = \theta_0, \ H_1 : \theta = \theta_1,$$

$\mathscr{F} = \{F(x, \theta_0), F(x, \theta_1)\}$, α - задане,

$$\begin{cases} W(\mathrm{X}_{1\alpha}; \theta_0) = \alpha, \\ W(\mathrm{X}_{1\alpha}; \theta_1) \to \max. \end{cases}$$

З цих співвідношень шукають $X^*_{1\alpha}$.

Припустимо, що розподіли $F(x, \theta_0)$ і $F(x, \theta_1)$ абсолютно неперервні і відповідні щільності $f_0(x)$ і $f_1(x)$ задовольняють умові $f_j(x) > 0$, $j = 0, 1$.

Розглянемо статистику **відношення вірогідності**:

$$l(\xi) = \frac{L(\xi, \theta_1)}{L(\xi, \theta_0)} = \frac{\prod\limits_{i=1}^{n} f_1(\xi_i)}{\prod\limits_{i=1}^{n} f_0(\xi_i)}$$

і визначимо функцію $\psi(c) = P_{\theta_0}\{l(\xi) \ge c\}$: 1) $\psi(0) = 1$, 2) $\psi(0) \to 0$ при $c \to \infty$.

Дійсно

$$P_{\theta_1}\{l(\xi) \ge c\} = \int\limits_{x:l(x) \ge c} L(x; \theta_1) dx \ge$$

$$\ge c \int\limits_{x:l(x) \ge c} L(x; \theta_0) dx = c P_{\theta_0}\{l(\xi) \ge c\} = c\psi(c),$$

тому $\psi(c) \le 1/c$, звідки й випливає 2).

Теорема 19.1 За зроблених припущень існує найбільш потужний критерій перевірки гіпотези H_0. Цей критерій задається критичною областю $X^*_{1\alpha} = \{x : l(x) \geq c\}$, де критична границя c визначається з умови $\psi(c) = \alpha$.

Побудований критерій називають **критерієм Неймана–Пірсона**.

19.4 Перевірка гіпотез про параметри нормальної моделі

Розглянемо ситуацію, коли про нормально розподілену випадкову величину ξ_0 з невідомим середнім θ і відомою дисперсією σ^2 є дві гіпотези:

$$H_0 : \theta = \theta_0,\ \ H_1 : \theta = \theta_1.$$

Для визначеності будемо вважати, що $\theta_1 > \theta_0$.

Нехай $\xi' = (\xi_1, \xi_2, \dots, \xi_n)$ вибірка з розподілу випадкової величини ξ_0 і x - реалізація ξ, що спостерігалася. Тоді

$$l(x) = \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[(x_i-\theta_1)^2-(x_i-\theta_0)^2]\right\} =$$
$$= \exp\left\{\frac{n}{\sigma^2}(\theta_1-\theta_0)\bar{x} - \frac{n}{2\sigma^2}({\theta_1}^2-{\theta_0}^2)\right\}.$$

Нерівність $l(x) \geq c$ еквівалентна нерівності

$$\bar{x} \geq \frac{\sigma^2 \ln c}{n(\theta_1-\theta_0)} + \frac{\theta_1+\theta_0}{2},$$

яку можна подати у вигляді

$$\frac{\sqrt{n}}{\sigma}(\bar{x}-\theta_0) \geq \frac{\sigma}{\sqrt{n}(\theta_1-\theta_0)}\ln c + \frac{\sigma}{2\sqrt{n}}(\theta_1-\theta_0) = t(c).$$

При $\theta = \theta_0$ випадкова величина $\frac{\sqrt{n}}{\sigma}(\bar{\xi}-\theta_0)$ має стандартний нормальний розподіл, тому

$$\psi(c) = P_{\theta_0}\{l(\xi) \geq c\} = P_{\theta_0}\left\{\frac{\sqrt{n}}{\sigma}(\bar{\xi}-\theta_0) \geq t(c)\right\} =$$
$$= 1 - \Phi(t(c)) = \Phi(-t(c)).$$

Для довільного $\alpha \in (0,1)$ послідовно визначаємо величини t_α і c_α такі, що $\Phi(-t_\alpha) = \alpha$ і $t(c_\alpha) = t_\alpha$.

Умови теореми Неймана-Пірсона виконуються і отже найбільш потужний критерій для перевірки гіпотези H_0 проти альтернативи H_1 задається критичною областю

$$X_{1\alpha}^* = \left\{x : \frac{\sqrt{n}}{\sigma}(\bar{x} - \theta_0) \geq t_\alpha\right\}, \ \Phi(-t_\alpha) = \alpha.$$

Обчислимо потужність критерію

$$W(X_{1\alpha}^*; \theta_1) = P_{\theta_1}\left\{\frac{\sqrt{n}}{\sigma}(\bar{\xi} - \theta_0) \geq t_\alpha\right\} = P_{\theta_1}\left\{\bar{\xi} \geq \theta_0 + \frac{\sigma}{\sqrt{n}} t_\alpha\right\} =$$

$$= P_{\theta_1}\left\{\frac{\sqrt{n}}{\sigma}(\bar{\xi} - \theta_1) \geq -\frac{\sqrt{n}}{\sigma}(\theta_1 - \theta_0) + t_\alpha\right\} =$$

$$= 1 - \Phi\left(-\frac{\sqrt{n}}{\sigma}(\theta_1 - \theta_0) + t_\alpha\right) = \Phi\left(\frac{\sqrt{n}}{\sigma}(\theta_1 - \theta_0) - t_\alpha\right).$$

Звідси випливає, що імовірність похибки другого роду дорівнює

$$\beta = \beta(\alpha, n) = \Phi\left(t_\alpha - \sqrt{n}(\theta_1 - \theta_0)/\sigma\right).$$

Приклад 19.2

Нехай заздалегідь задані ймовірності похибок першого та другого роду. Визначити, якими повинно бути мінімальне число $n^* = n^*(\alpha, \beta)$ випробувань, щоб хибні висновки могли бути зроблені з ймовірностями, які не перевищують α і β).

Розв'язок: Для визначення n маємо два рівняння:

$$\Phi(-t_\alpha) = \alpha, \ \Phi\left(t_\alpha - \sqrt{n}(\theta_1 - \theta_0)/\sigma\right) = \beta.$$

Нехай k_p - це квантіль p-го порядку, або розв'язок рівняння $\Phi(k_p) = p$. Тоді

$$-t_\alpha = k_\alpha, \ t_\alpha - \sqrt{n}(\theta_1 - \theta_0)/\sigma = k_\beta,$$

і

$$n = \frac{\sigma^2(k_\alpha + k_\beta)^2}{(\theta_1 - \theta_0)^2} \text{ або } n^* = \left[\frac{\sigma^2(k_\alpha + k_\beta)^2}{(\theta_1 - \theta_0)^2}\right] + 1.$$

Звідси випливає, що при фіксованих похибках число спостережень пропорційне дисперсії і обернено пропорційне квадрату різниці між середніми значеннями.

Підсумовуючи вищевказане, наведемо деякі алгоритми перевірки гіпотез про параметри нормальної моделі.

Гіпотеза $H_0 : m = m_0$ (σ^2 невідома).

Зазначимо, що така гіпотеза може мати різні альтернативи ($m \neq m_0$, $m > m_0$, $m < m_0$).

Оскільки незалежно від того, справедлива гіпотеза $H_0 : m = m_0$ чи ні, оцінка $\bar{\xi} = \frac{1}{n}\sum\limits_{k=1}^{n} \xi_k$ є консистентною та незсуненою для m, то логічно прийняти гіпотезу $H_0 : m = m_0$ за невеликих значень $\bar{\xi} - m_0$ і відхилити її за великих значень $\bar{\xi} - m_0$. Межі, що відділяють великі значення $\bar{\xi} - m_0$ від невеликих, будуються на основі того факту, що для вибірки $\xi' = (\xi_1, ..., \xi_n)$ величина $\frac{\bar{\xi} - m_0}{\hat{S}/\sqrt{n}}$ має розподіл Стьюдента із $n - 1$ ступенем свободи, де $\hat{S}^2 = \frac{1}{n-1}\sum\limits_{k=1}^{n} \left(\xi_k - \bar{\xi}\right)^2$.

Критерій Стьюдента перевірки гіпотези $H_0 : m = m_0$.

а) За альтернативи $H_1 : m \neq m_0$ гіпотеза $H_0 : m = m_0$ відхиляється при

$$\frac{|\bar{\xi} - m_0|}{\hat{S}/\sqrt{n}} > t_{1-\frac{\alpha}{2}, n-1},$$

де $t_{1-\frac{\alpha}{2}, n-1}$– квантиль рівня $1 - \frac{\alpha}{2}$ розподілу Стьюдента з $n - 1$ ступенями свободи. У протилежному випадку гіпотезу $H_0 : m = m_0$ приймаємо. При цьому з імовірністю α (рівень значущості) гіпотеза H_0 буде відхилятися, коли вона справедлива.

б) За альтернативи $H_1 : m > m_0$ гіпотеза $H_0 : m = m_0$ відхиляється при

$$\frac{\bar{\xi} - m_0}{\hat{S}/\sqrt{n}} > t_{1-\alpha, n-1}.$$

У протилежному випадку гіпотезу $H_0 : m = m_0$ приймаємо (рівень значущості α).

в) За альтернативи $H_1 : m < m_0$ гіпотеза $H_0 : m = m_0$ відхиляється при

$$\frac{\bar{\xi} - m_0}{\hat{S}/\sqrt{n}} < t_{\alpha, n-1}.$$

У протилежному випадку гіпотезу $H_0 : m = m_0$ приймаємо (рівень значущості α).

Приклад 19.3

Для перевірки твердження виробника про те, що генератор за зміну споживає в середньому 20 л пального, здійснили 10 випробувань. За 10 змін споживання генератора встановили:

$$19.1, 18.6, 19.1, 18.1, 16.6, 20.1, 19.8, 21.1, 24.4, 21.6.$$

Перевірити твердження виробника при рівні значущості 0.05.

Розв'язок: $H_0 : m = m_0 = 20,\ H_1 : m \neq m_0$.

$$\bar{\xi} = \frac{1}{10}\,(19.1 + 18.6 + \cdots + 21.6) = 19.85.$$

$$\hat{S}^2 = \frac{1}{9}\left((19.1 - 19.85)^2 + (18.6 - 19.85)^2 + \cdots + (21.6 - 19.85)^2\right) =$$

$$= \frac{41.705}{9} = \frac{8341}{1800}.$$

$$\frac{|\bar{\xi} - m_0|}{\hat{S}/\sqrt{n}} \approx 0.22 < t_{1-\frac{\alpha}{2},n-1} = t_{0.975,9} = 2.262.$$

Це означає, що дані не суперечать твердженню виробника про те, що генератор за зміну споживає в середньому 20 л пального, і гіпотеза $H_0 : m = m_0 = 20$ приймається.

Приклад розв'язання цієї задачі за допомогою R див. на стор. 295.

Гіпотеза $H_0 : \sigma^2 = \sigma_0^2$ (m невідоме). Альтернативна гіпотеза при цьому може бути як однобічною ($H_1 : \sigma^2 < \sigma_0^2$ чи $H_1 : \sigma^2 > \sigma_0^2$),так і двобічною ($H_1 : \sigma^2 \neq \sigma_0^2$).

Незалежно від того, справедлива чи ні гіпотеза $H_0 : \sigma^2 = \sigma_0^2$, оцінка $\hat{S}^2 = \frac{1}{n-1}\sum\limits_{k=1}^{n} \xi_k - \bar{\xi}$ є консистентною та незсуненою оцінкою параметра σ^2, тому гіпотезу $H_0 : \sigma^2 = \sigma_0^2$ слід приймати, якщо $\frac{\hat{S}^2}{\sigma_0^2}$ не дуже відхиляться від 1. Межі, що відділяють значення $\frac{\hat{S}^2}{\sigma_0^2}$, які дуже відхиляються від 1, від значень $\frac{\hat{S}^2}{\sigma_0^2}$, які мало відхиляються від 1, будують на базі того, що $\frac{(n-1)\hat{S}^2}{\sigma_0^2}$ має χ^2-розподіл з $n-1$ ступенями свободи.

Критерій перевірки гіпотези $H_0 : \sigma^2 = \sigma_0^2$.

а) За альтернативи $H_1 : \sigma^2 \neq \sigma_0^2$ гіпотеза $H_0 : \sigma^2 = \sigma_0^2$ приймається при

$$\hat{S}^2 \in \left(\frac{\sigma_0^2}{n-1}\chi^2_{\frac{\alpha}{2},n-1}; \frac{\sigma_0^2}{n-1}\chi^2_{1-\frac{\alpha}{2},n-1}\right).$$

У протилежному випадку гіпотеза H_0 відхиляється (рівень значущості α).

б) За альтернативи $H_1 : \sigma^2 > \sigma_0^2$ гіпотеза $H_0 : \sigma^2 = \sigma_0^2$ приймається при

$$\hat{S}^2 < \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha,n-1}.$$

У протилежному випадку гіпотеза H_0 відхиляється (рівень значущості α).

в) За альтернативи $H_1 : \sigma^2 < \sigma_0^2$ гіпотеза $H_0 : \sigma^2 = \sigma_0^2$ приймається при

$$\hat{S}^2 > \frac{\sigma_0^2}{n-1}\chi^2_{\alpha,n-1}.$$

У протилежному випадку гіпотеза H_0 відхиляється (рівень значущості α).

19.5 Перевірка гіпотез про рівність математичних сподівань і дисперсій двох нормальних вибірок

Нехай $\xi^{(1)'} = \left(\xi_1^{(1)}, \xi_2^{(1)}, \ldots, \xi_{n_1}^{(1)}\right)$ вибірка з нормального розподілу $N(m_1, \sigma_1^2)$, а $\xi^{(2)'} = \left(\xi_1^{(2)}, \xi_2^{(2)}, \ldots, \xi_{n_2}^{(2)}\right)$ - вибірка з нормального розподілу $N(m_2, \sigma_2^2)$.

Розглянемо ряд гіпотез про параметри m_1, m_2, σ_1^2, σ_2^2.

Гіпотеза про рівність математичних сподівань за відомих дисперсій $H_0 : m_1 = m_2$ (коли σ_1^2, σ_2^2 - відомі).

Випадкові величини $\bar{\xi}^{(1)} = \frac{1}{n_1}\sum\limits_{k=1}^{n_1}\xi_k^{(1)}$, $\bar{\xi}^{(2)} = \frac{1}{n_2}\sum\limits_{k=1}^{n_2}\xi_k^{(2)}$ мають нормальний розподіл $N(m_1, \frac{\sigma_1^2}{n_1})$ і $N(m_2, \frac{\sigma_2^2}{n_2})$, відповідно. Тоді $\bar{\xi}^{(1)} - \bar{\xi}^{(2)}$ має нормальний розподіл $N(m_1 - m_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$. Якщо справедлива гіпотеза $H_0 : m_1 = m_2$, то величина $\frac{\bar{\xi}^{(1)} - \bar{\xi}^{(2)}}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})}$ має стандартний нормальний розподіл $N(0,1)$.

Критерій перевірки гіпотези $H_0 : m_1 = m_2$.

а) За альтернативи $H_1 : m_1 \neq m_2$ гіпотеза $H_0 : m_1 = m_2$ відхиляється при

$$\frac{|\bar{\xi}^{(1)} - \bar{\xi}^{(2)}|}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})} \geq c_{1-\frac{\alpha}{2}}.$$

У протилежному випадку гіпотезу $H_0 : m_1 = m_2 = c$ приймаємо (рівень значущості α).

б) За альтернативи $H_1 : m_1 > m_2$ гіпотеза $H_0 : m_1 = m_2$ відхиляється при

$$\frac{\bar{\xi}^{(1)} - \bar{\xi}^{(2)}}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} > c_{\frac{1-2\alpha}{2}}.$$

У протилежному випадку гіпотезу $H_0 : m_1 = m_2$ приймаємо (рівень значущості α).

в) За альтернативи $H_1 : m_1 < m_2$ гіпотеза $H_0 : m_1 = m_2$ відхиляється при

$$\frac{\bar{\xi}^{(1)} - \bar{\xi}^{(2)}}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < -c_{\frac{1-2\alpha}{2}}.$$

У протилежному випадку гіпотезу $H_0 : m_1 = m_2$ приймаємо (рівень значущості α).

Гіпотеза про рівність математичних сподівань за невідомих дисперсій $H_0 : m_1 - m_2 = c$ (коли $\sigma_1^2 = \sigma_2^2 = \sigma^2$ і σ^2 не відома).

Позначимо $\Delta = \bar{\xi}^{(1)} - \bar{\xi}^{(2)}$. Незалежно від того, справедлива чи ні гіпотеза $H_0 : m_1 - m_2 = c$, Δ є конзистентною та незсуненою оцінкою для c, тому гіпотезу H_0 слід приймати за незначних відхилень $\Delta - c$. Межі, які відділяють великі відхилення від малих, будуються на основі того, що $\frac{\Delta - c}{\hat{S}\sqrt{\frac{n_1+n_2}{n_1 n_2}}}$ має розподіл Стьюдента з $n_1 + n_2 - 2$ ступенями свободи, де $\hat{S}^2 = \frac{(n_1-1)\hat{S}_1^2 + (n_2-1)\hat{S}_2^2}{n_1+n_2-2}$, $\hat{S}_i^2 = \frac{1}{n_i - 1}\sum\limits_{k=1}^{n_i}\left(\xi_k^{(i)} - \bar{\xi}^{(i)}\right)^2$, $i = 1, 2$.

Критерій Стьюдента перевірки гіпотези $H_0 : m_1 - m_2 = c$.

а) За альтернативи $H_1 : m_1 - m_2 \neq c$ гіпотеза $H_0 : m_1 - m_2 = c$ відхиляється при

$$\frac{|\Delta - c|}{\hat{S}\sqrt{\frac{n_1+n_2}{n_1 n_2}}} > t_{1-\frac{\alpha}{2}, n_1+n_2-2}.$$

У протилежному випадку гіпотезу $H_0 : m_1 - m_2 = c$ приймаємо (рівень значущості α).

б) За альтернативи $H_1 : m_1 - m_2 > c$ гіпотеза $H_0 : m_1 - m_2 = c$ відхиляється при

$$\frac{\Delta - c}{\hat{S}\sqrt{\frac{n_1+n_2}{n_1 n_2}}} > t_{1-\alpha, n_1+n_2-2}.$$

У протилежному випадку гіпотезу $H_0 : m_1 - m_2 = c$ приймаємо (рівень значущості α).

в) За альтернативи $H_1 : m_1 - m_2 < c$ гіпотеза $H_0 : m_1 - m_2 = c$ відхиляється при

$$\frac{\Delta - c}{\hat{S}\sqrt{\frac{n_1+n_2}{n_1 n_2}}} < t_{\alpha, n_1+n_2-2}.$$

У протилежному випадку гіпотезу $H_0 : m_1 - m_2 = c$ приймаємо (рівень значущості α).

Приклад 19.4

Нижче наведено дані вимірювання рівня чутливості однорідного фотоматеріалу за допомогою двох фотосенсометрів. Чи можна вважати, що між показниками приладів немає систематичного розходження (рівень значущості 0,05)?

Перший прилад: 5.0, 5.1, 5.5, 6.0, 5.5, 3.6, 1.8, 7.8, 6.9, 2.7;

Другий прилад: 6.1, 5.4, 5.8, 2.7, 2.8, 2.8, 4.2, 2.4.

Розв'язок: $H_0 : m_1 - m_2 = 0$, $H_1 : m_1 - m_2 \neq 0$.

$$\bar{\xi}^{(1)} = \frac{1}{10}(5.0 + 5.1 + \cdots + 2.7) = 4.99, \bar{\xi}^{(2)} = \frac{1}{8}(6.1 + 5.4 + \cdots + 2.4) = 4.025,$$

$$\Delta = \bar{\xi}^{(1)} - \bar{\xi}^{(2)} = 4.99 - 4.025 = 0.965,$$

$$\hat{S}_1^2 = \frac{1}{9}\left((5.0 - 4.99)^2 + (5.1 - 4.99)^2 + \cdots + (2.7 - 4.99)^2\right) \approx 3.38,$$

$$\hat{S}_2^2 = \frac{1}{7}\left((6.1 - 4.025)^2 + (5.4 - 4.025)^2 + \cdots + (2.4 - 4.025)^2\right) \approx 2.4,$$

$$\hat{S}^2 = \frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{n_1 + n_2 - 2} = \frac{9 \cdot 3.38 + 7 \cdot 2.4}{10 + 8 - 2} \approx 2.95,$$

$$\frac{|\Delta - c|}{\hat{S}\sqrt{\frac{n_1+n_2}{n_1 n_2}}} = \frac{|0.965 - 0|}{\sqrt{2.95\frac{10+8}{10\cdot 8}}} \approx 1.1844 < t_{1-\frac{0.05}{2},10+8-2} = t_{0.975,16} = 2.12,$$

а це означає, що дані не суперечать гіпотезі про те, що між показниками приладів немає систематичного розходження, і ми приймаємо цю гіпотезу.

Приклад розв'язання цієї задачі за допомогою R див. на стор. 297.

Гіпотеза про рівність дисперсій за відомих математичних сподівань $H_0 : \sigma_1^2 = \sigma_2^2$ (коли m_1, m_2 - відомі).

Незалежно від того, справедлива чи ні гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$, $\bar{S}_i^2 = \frac{1}{n_i}\sum_{k=1}^{n_i}\left(\xi_k^{(i)} - m_i\right)^2$, $i = 1, 2$ є консистентними та незсуненими оцінками для σ_1^2 та σ_1^2, відповідно, тому, як відомо, $\frac{\bar{S}_1^2}{\bar{S}_2^2}$ збігається за ймовірністю до $\frac{\sigma_1^2}{\sigma_2^2}$. Отже, за справедливості гіпотези $H_0 : \sigma_1^2 = \sigma_2^2$, $\frac{\bar{S}_1^2}{\bar{S}_2^2}$ має не дуже відхилятися від 1. Межі, що відділяють значні відхилення від незначних, отримують, виходячи з того, що $\frac{\bar{S}_1^2}{\bar{S}_2^2}$ має $F-$розподіл (Фішера) із n_1 та n_2 ступенями свободи.

Критерій перевірки гіпотези $H_0 : \sigma_1^2 = \sigma_2^2$.

а) За альтернативи $H_1 : \sigma_1^2 \neq \sigma_2^2$ гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$ приймається при

$$\frac{\bar{S}_1^2}{\bar{S}_2^2} \in \left[\frac{1}{F_{1-\frac{\alpha}{2},n_2,n_1}}; F_{1-\frac{\alpha}{2},n_1,n_2}\right],$$

де $F_{\alpha,m,n}$ – квантиль рівня α розподілу Фішера з m, n ступенями свободи.

У протилежному випадку гіпотезу $H_0 : \sigma_1^2 = \sigma_2^2$ відхиляємо (рівень значущості α).

б) За альтернативи $H_1 : \sigma_1^2 > \sigma_2^2$ гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$ приймається при

$$\frac{\bar{S}_1^2}{\bar{S}_2^2} \leq F_{1-\alpha,n_1,n_2}.$$

У протилежному випадку гіпотезу $H_0 : \sigma_1^2 = \sigma_2^2$ відхиляємо (рівень значущості α).

в) За альтернативи $H_1 : \sigma_1^2 < \sigma_2^2$ гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$ приймається при

$$\frac{\bar{S}_1^2}{\bar{S}_2^2} \geq \frac{1}{F_{1-\alpha,n_2,n_1}}.$$

У протилежному випадку гіпотезу $H_0 : \sigma_1^2 = \sigma_2^2$ відхиляємо (рівень значущості α).

Гіпотеза про рівність дисперсій за невідомих математичних сподівань $H_0 : \sigma_1^2 = \sigma_2^2$ (коли m_1, m_2 - невідомі).

Ця гіпотеза перевіряється аналогічно попередній, але в даному випадку в якості статистики критерію беруть величину $\frac{\hat{S}_1^2}{\hat{S}_2^2}$, де $\hat{S}_i^2 = \frac{1}{n_i - 1} \sum\limits_{k=1}^{n_i} \left(\xi_k^{(i)} - \bar{\xi}^{(i)} \right)^2$, $i = 1, 2$. Якщо справедлива гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$, то випадкова величина $\frac{\hat{S}_1^2}{\hat{S}_2^2}$ має розподіл Фішера з $n_1 - 1$ та $n_2 - 1$ ступенями свободи.

Критерій перевірки гіпотези $H_0 : \sigma_1^2 = \sigma_2^2$.

а) За альтернативи $H_1 : \sigma_1^2 \neq \sigma_2^2$ гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$ приймається при

$$\frac{\hat{S}_1^2}{\hat{S}_2^2} \in \left[\frac{1}{F_{1-\frac{\alpha}{2}, n_2-1, n_1-1}} ; F_{1-\frac{\alpha}{2}, n_1-1, n_2-1} \right].$$

У протилежному випадку гіпотезу $H_0 : \sigma_1^2 = \sigma_2^2$ відхиляємо (рівень значущості α).

б) За альтернативи $H_1 : \sigma_1^2 > \sigma_2^2$ гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$ приймається при

$$\frac{\hat{S}_1^2}{\hat{S}_2^2} \leq F_{1-\alpha, n_1-1, n_2-1}.$$

У протилежному випадку гіпотезу $H_0 : \sigma_1^2 = \sigma_2^2$ відхиляємо (рівень значущості α).

в) За альтернативи $H_1 : \sigma_1^2 < \sigma_2^2$ гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$ приймається при

$$\frac{\hat{S}_1^2}{\hat{S}_2^2} \geq \frac{1}{F_{1-\alpha, n_2-1, n_1-1}}.$$

У протилежному випадку гіпотезу $H_0 : \sigma_1^2 = \sigma_2^2$ відхиляємо (рівень значущості α).

Приклад 19.5

Чи можна в умовах попереднього прикладу вважати, що точність вимірювань на обох фотосенсометрах однакова (рівень значущості 0.05)?

Розв'язок: $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \neq \sigma_2^2$.

$$\frac{\hat{S}_1^2}{\hat{S}_2^2} \approx \frac{3.38}{2.4} \approx 1.4,$$

$$F_{1-\frac{\alpha}{2}, n_2-1, n_1-1} = F_{1-\frac{0.05}{2}, 8-1, 10-1} = F_{0.975,7,9} = 4.197,$$

$$F_{1-\frac{\alpha}{2}, n_1-1, n_2-1} = F_{1-\frac{0.05}{2}, 10-1, 8-1} = F_{0.975,9,7} = 4.82,$$

$$\frac{\hat{S}_1^2}{\hat{S}_2^2} \approx 1.4 \in \left[\frac{1}{F_{1-\frac{\alpha}{2}, n_2-1, n_1-1}} ; F_{1-\frac{\alpha}{2}, n_1-1, n_2-1} \right] = \left[\frac{1}{4.197} ; 4.82 \right] = [0.238; 4.82],$$

а це означає, що дані не суперечать гіпотезі про те, що точність вимірювань на обох фотосенсометрах однакова, і ми приймаємо цю гіпотезу.

Приклад розв'язання цієї задачі за допомогою R див. на стор. 298.

Задача 19.1. У експериментальній групі людей, що випробовували новий препарат для схуднення, проводили зважування до та після курсу лікування. За результатами зважувань необхідно визначити, чи істотно змінювалась вага під впливом нових ліків. Рівень значущості обрати 0.05.

До лікування (кг): 120, 112, 116, 113, 119, 130, 122, 125, 127, 120.

Після лікування (кг): 131, 129, 91, 98, 122, 114, 108, 121, 128, 123.

Задача 19.2. За документацією дозатор наливає у кожну ємність в середньому 10 л розчину. Для перевірки роботи дозатора зробили тестові розливання. Перевірити з рівнем значущості 0.05 відповідність роботи дозатора документації.

Значення (л): 10.1, 10, 10.2, 9.9, 9.8, 9.9, 10, 10.1, 10.1, 10, 9.9, 10.

Задача 19.3. За значеннями генератора випадкових гауссівських величин перевірити з рівнем значущості 0,05 гіпотезу про одиничну дисперсію.

Значення: 4.4, 4.7, 5.5, 5.2, 5.4, 3.8, 3.9, 3.9, 4.6, 3.7.

Задача 19.4. Мікроскопічні нерівності однорідної поверхні спочатку вимірювали старим приладом, а потім – новим. Чи можна стверджувати, що новий прилад має вищу точність вимірювання (при рівні значущості 0.1), якщо виміри були такими:

Старий прилад (мкм): 4.56, 4.32, 4.53, 4.05, 3.31, 5.04, 4.88, 5.56, 7.19, 5.81.

Новий прилад (мкм): 4.42, 4.59, 5.45, 4.80, 5.07, 6.01, 4.44, 5, 5.89, 5.71.

Задача 19.5. Для перевірки гіпотези про те, що людина перестає рости у віці 19 років, вимірювали зріст у 100 осіб у віці 19 років та у 100 осіб у віці 20 років. За даними вимірів були розраховані наступні характеристики: середні значення зросту в першій та другій групах склали відповідно $\bar{\xi}^{(1)} = 171$см та $\bar{\xi}^{(2)} = 172$ см, а вибіркові дисперсії в групах склали $S_1^2 = 16$ (см^2) та $S_2^2 = 26$(см^2). Перевірити висунуту гіпотезу при рівні значущості 0.1. (Примітка: $t_{0.1,198} = -1.286$)

Задача 19.6. За нормативами дисперсія розсіювання ваги пакунку сухої шпаклівки складає 0.1 кг^2. З часом у фасувального автомату дисперсія розсіювання ваги зростає, що вимагає ремонту автомату. За даними контрольного зважування визначити, чи потрібно ремонтувати автомат (рівень значущості 0.1).

Результати зважувань (кг): 25.1, 25.09, 25.09, 25.07, 24.9, 25.01, 24.92, 25.07, 24.98, 24.94.

Задача 19.7. Середня вага дорослої людини складає 85 кг. Для того, щоб довести, що вживання гамбургерів та картоплі фрі не сприяє збільшенню ваги, контрольну групу у 100 дорослих людей годували гамбургерами та картоплею фрі тривалий час, а потім вимірювали їхню вагу. За результатами зважувань зробили такі розрахунки: середня вага $\bar{\xi} = 91$ кг, вибіркова дисперсія $S^2 = 12$ кг^2. Перевірити висунуту гіпотезу при рівні значущості 0.1, врахувавши, що $t_{0.9,99} = 1.29$.

Задача 19.8. Виміри зросту хлопців двох класів наведено нижче. Висунути та перевірити (при рівні значущості 0.1) гіпотезу про те, що середній зріст хлопців в обох класах відрізняється несуттєво.

Клас А (см): 163.9, 163.4, 163.6, 162.6, 155.8, 160.2, 156.9, 162.7, 159.2, 157.4.

Клас Б (см): 157.1, 157.9, 162.2, 159, 160.3, 165, 157.1, 160, 164.4, 163.5.

Задача 19.9. Нижче наведено дані вимірів діаметрів стандартних втулок двома різними штангенциркулями. Чи можна стверджувати (при рівні значущості 0.05), що обидва прилади мають однакову точність вимірів?

1 штангенциркуль (мм): 9.1, 9.8, 12.2, 11.4, 10.8, 12.1, 12.8, 12.3, 12, 11.2.

2 штангенциркуль (мм): 11.6, 11.3, 11.9, 13, 12.2, 12.5, 12.7, 12, 12.3, 12.2.

Задача 19.10. За документацією середній час спрацьовування запалу гранати Ф-6 складає 8 секунд. Зменшення цього часу тягне загрозу для життя. За наведеними нижче даними контрольних вимірів часу спрацьовування запалу висунути та перевірити відповідну гіпотезу із рівнем значущості 0.1.

Час спрацьовування: 7.28, 7.48, 8.56, 7.76, 8.09, 9.26, 7.29, 8, 9.11, 8.89, 5.1.

Задача 19.11. Вважається, що жінки живуть в середньому на 8 років довше за чоловіків. Для перевірки цього припущення були проаналізовані дати життя та смерті 50 чоловіків та 60 жінок. За результатами обробки статистичних даних було підраховано, що середня тривалість життя у чоловіків склала $\overline{\xi^{(1)}} = 56$ років, а у жінок – $\overline{\xi^{(2)}} = 65$років. При цьому вибіркова дисперсія у першій вибірці склала $S_1^2 = 3$ (роки^2), а за другою – $S_2^2 = 2.5$ (роки^2). Перевірити висунуту гіпотезу при рівні значущості 0.1. (Примітка: $t_{0.95,108} \approx 1.66$)

Задача 19.12. В результаті спостережень над випадковими величинами ξ та η отримали такі вибірки:

ξ : 45, 48, 53, 44, 59, 60, 41, 43, 57;

η : 51, 50, 42, 44, 39, 40, 48, 38, 59, 55, 51.

Чи можна вважати, що випадкові величини ξ та η мають однакові математичні сподівання? Похибка першого роду дорівнює 0.05. Припускається, що випадкові величини ξ та η

мають нормальний розподіл з рівними дисперсіями.

Задача 19.13. Нехай тепер

ξ : 2.50, 2.50, 2.60, 2.75, 2.80, 2.80, 2.95;

η : 2.50, 2.80, 2.85, 2.90, 2.90, 2.95, 3.40.

Чи можна вважати, що ξ та η мають однакові математичні сподівання? $\alpha = 0.05$.

Задача 19.14. В одному класі з 20 дітей навмання відібрали 10, яким щодня почали видавати апельсиновий сік. Інші 10 учнів щодня отримували молоко. Через деякий час зафіксували збільшення маси дітей в фунтах:

Сік: 4,0; 2,5; 3,5; 4.0; 1,5; 1,0; 3,5; 3,0; 2,5; 3,5.

Молоко: 1,5; 3,5; 2,5; 3,0; 2,5; 2,0; 2,0; 2,5; 1,5; 3,0.

Чи суттєво відрізняється середнє збільшення ваги дітей в групах?

Задача 19.15. З нормальної генеральної сукупності з $\sigma^2 = 25$ отримано дві вибірки об'ємом $n_1 = n_2 = 9$. Середнє першої вибірки $\bar{x} = 2$, другої – $\bar{y} = 3$. Чи можна пояснити цю розбіжність випадковими причинами при похибці першого роду 0.05?

Задача 19.16. Одним і тим самим приладом було зроблено дві серії вимірів: 1) 2.5; 3.2; 3.5; 3.8; 3.5; 2) 2.0; 2.7; 2.5; 2.9; 2.3; 2.6. а) Припускаючи, що виміри мають нормальний розподіл з однаковими дисперсіями, перевірити гіпотезу про рівність математичних сподівань при $\alpha = 0.05$. б) Перевірити гіпотезу про те, що дисперсії однакові для цих вимірів $\alpha = 0.05$. в) Перевірити гіпотезу про те, що дисперсії однакові для цих вимірів, якщо $a_{=}3, a_2 = 2.5$, $\alpha = 0.05$.

Задача 19.17. На станку-автоматі виготовляється один вид продукції. Критичним розміром виробів є зовнішній діаметр. Після налаштування станка відібрали 20 виробів. При цьому виявилось, що вибіркова дисперсія розміру зовнішнього діаметра складає 0,84 мм^2. Через деякий проміжок часу з метою контролю точності роботи станка та, за необхідності, його налаштування, відібрали 15 виробів. Вибіркова дисперсія, обчислена по них, дорівнює 1,07 мм^2. Чи свідчать наведені дані про зміну точності роботи станка? Взяти $\alpha = 0.05$.

Задача 19.18. Для зменшення дисперсії відбиваючої властивості фарби внесено зміни в технологію її виробництва. Щоб переконатись, що такі зміни насправді дають ефект, виготовили 10 пробних зразків та визначили відбиваючу властивість фарби, що виготовлена з використанням звичайної (А) та нової (В) технологій (в умовних одиницях). Отримано такі дані:

Технологія А: 40, 45, 195, 65, 145.

Технологія В: 110, 55, 120, 50, 80.

Чи свідчать ці дані про зміну дисперсії відбиваючої властивості фарби? Взяти $\alpha = 0.05$.

Задача 19.19. По двох вибірках з генеральних сукупностей об'єму $n_1 = 11$, $n_2 = 15$, підраховано $S_1^2 = \frac{1}{n_1}\sum_{k=1}^{n_1}(x_i - \bar{x})^2 = 0.76$ і $S_2^2 = \frac{1}{n_2}\sum_{k=1}^{n_2}(y_i - \bar{y})^2 = 0.76$. При $\alpha = 0.05$ перевірити гіпотезу про рівність дисперсій двох нормальних сукупностей.

Задача 19.20. 12 хлопчиків та 10 дівчаток пройшли тест на швидкочитання, у результаті якого виявилось, що середня швидкість читання хлопчиків дорівнює 74 слів за хвилину зі стандартним відхиленням 3, а у дівчаток - 79 зі стандартним відхиленням 4. Чи можна вважати, що здібності до читання хлопчиків та дівчаток відрізняються? Взяти $\alpha = 0.05$).

20 Базове моделювання в R

R — це програмно-статистичне середовище з інтерфейсом командного рядка та відкритим вихідним кодом. Все необхідне програмне забезпечення для встановлення R для різних операційних систем є на кількох сайтах, базовим є офіційний сайт http://www.r-project.org/. Як записано на офіційному сайті, R – це мова програмування та середовище для статистичної обробки даних та роботи з графікою (R is a language and environment for statistical computing and graphics).

R застосовується скрізь, де потрібна робота з даними, та дозволяє провести як базову попередню обробку даних (графіки, базові характеристики, таблиці спряженості), так і будувати різноманітні складні моделі, проводити їхній аналіз за допомогою величезного спектру класичних та найсучасніших методів аналізу даних. Фактично для всіх галузей науки та для всіх добре відомих (та не дуже відомих) статистичних методик в R є функції та пакети розширень, які їх підтримують.

На офіційній сторінці багато посилань, серед яких ті, які дозволяють завантажити R для різних операційних систем, та велика кількість документації. Для завантаження треба переходити за відповідними посиланнями та слідувати інструкціям.

Розроблено кілька оболонок, які мають за мету зробити роботу в R зручнішою. Вони також безкоштовні та можуть бути встановлені після того, як завантажено R. Найбільш популярними є

- RStudio: http://www.rstudio.com

- RCommander: http://www.rcommander.com/

Варто звернути увагу на посилання CRAN – Comprehensive R Archive Network – "CRAN is a network of ftp and web servers around the world that store identical, up-to-date, versions of code and documentation for R"https://cran.r-project.org/ – стабільна база пакетів для R та численні доповнення.

Детальніше про встановлення, основи програмування в R, графічні та статистичні можливості, можна знайти в численній літературі ([13], [16], [20], [21], тощо), вибравши, наприклад, звідси: https://cran.r-project.org/doc/contrib/.

20.1 Ймовірнісні розподіли в R

R призначений перш за все для статистичної обробки даних, методи та ідеологія якої ґрунтується на теорії ймовірностей. Тому R має багато можливостей для роботи з ймовірностями, випадковими величинами та їхніми розподілами: рахувати ймовірності, квантилі, значення функції розподілу та щільності розподілу, генерувати випадкові величини, зображувати розподіли, тощо.

20.1.1 Генерування випадкової вибірки

Багато ранніх робіт в галузі теорії ймовірностей були присвячені іграм та питанням перемішування, що базується на припущеннях симетрії. Тут базовим є припущення про випадкову вибірку: операції з добре перетасованою колодою карт або вибір пронумерованих кульок з добре перемішаної урни.

В R можна моделювати такі ситуації з функцією ‘sample()‘. Якщо треба, наприклад, вибрати 6 номерів з 49, треба записати:

```
sample(1:49,6)
[1]  1 29 28  3 27 36
```

Перший аргумент функції ‘sample()‘, ‘x‘, – це вектор значень, з яких вибираються значення, другий,‘size‘, – це розмір вибірки. Насправді запису ‘sample(49,6)‘ було б цілком достатньо в цьому випадку, оскільки окреме число в якості першого аргументу сприймається як таке, що представляє довжину послідовності натуральних чисел. ‘replace‘ – логічний аргумент, який вказує, чи повертаються вже вибрані числа назад. За замовчуванням ‘replace=FALCE‘, що означає, що раз обране число не повертається до базового набору чисел, а отже не може бути вибране повторно. Очевидно, що для цього випадку ‘size‘ не може бути меншим за бажану кількість елементів вибірки. Якщо є потреба формувати вибірку з повтореннями, треба вказати значення аргументу ‘replace=TRUE‘. Утворення вибірки з повторами зручне, наприклад, для моделювання багатократного підкидання монети. Наприклад, для моделювання 10 підкидань симетричної монети можна написати:

```
sample(c("Р","Г"), 10, replace=TRUE)
[1] "Г" "Р" "Г" "Г" "Р" "Р" "Г" "Р" "Р" "Р"
```

В даному випадку ймовірності випадіння герба та решки однакова. Проте можна змоделювати нерівноможливий випадок, використовуючи аргумент ‘prob‘, який задає вектор відповідних ймовірностей:

```
sample(c("succ", "fail"), 10, replace=T, prob=c(0.8, 0.2))
[1] "succ" "succ" "succ" "succ" "succ" "succ" "fail" "succ" "succ" "succ"
```

Це – згенерована послідовність незалежних випробувань Бернуллі з ймовірністю успіху 0.8. Саме таку послідовність можна простіше змоделювати, використовуючи функцію, що моделює біноміальний розподіл (нижче).

За допомогою функції ‘sample()‘ є можливість не лише згенерувати вибірку з певного датасету, але й випадковим чином перевпорядкувати датасет, утворити перестановку. Виклик функції ‘sample(x)‘, єдиним аргументом якої є датасет, є еквівалентним виклику

```
sample(x, size=length(x), replace=FALSE)
```

що означає "вибрати всі елементи датасету ‘x‘ у випадковому порядку, використовуючи кожен елемент лише один раз". Тобто створити випадкову перестановку елементів датасету:

```
sample(1:10)
[1]  9  8  2  5  6  3  1 10  4  7
```

20.1.2 Генерування комбінацій

Якщо є потреба згенерувати всі можливі k-елементні комбінації, що вибрані з n елементів, можна скористатися функцією ‘combn()‘

```
combn(x, m, FUN = NULL, simplify = TRUE, ...)
```

Аргументи цієї функції:
- ‘x‘ – вектор елементів (чисельних чи символьних), з яких вибираються комбінації (підмножини);
- ‘m‘ – кількість елементів, які вибираються.

```
combn(1:5, 2)
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
[1,]    1    1    1    1    2    2    2    3    3     4
[2,]    2    3    4    5    3    4    5    4    5     5
```

```
combn(c("Кривоніс", "Гуляйполе", "Убийвовк", "Непийпиво", "Вернидуб"), 4)
     [,1]        [,2]        [,3]        [,4]        [,5]       
[1,] "Кривоніс"  "Кривоніс"  "Кривоніс"  "Кривоніс"  "Гуляйполе"
[2,] "Гуляйполе" "Гуляйполе" "Гуляйполе" "Убийвовк"  "Убийвовк" 
[3,] "Убийвовк"  "Убийвовк"  "Непийпиво" "Непийпиво" "Непийпиво"
[4,] "Непийпиво" "Вернидуб"  "Вернидуб"  "Вернидуб"  "Вернидуб" 
```

20.2 Ймовірнісні розподіли

В базовій установці R реалізовані найчастіше вживані розподіли. Кожен з цих розподілів має скорочене ім'я, яке використовується для визначення відповідних функцій, що відповідають розподілу. Наприклад, ім'я нормального розподілу – ‘norm‘, і це ім'я використовується як корінь в імені функцій підрахунку значень, пов'язаних з цим розподілом. Для кожного розподілу є чотири базові функції, які можуть бути підраховані. Корінь назви функції вказує на вид розподілу. Перша літера – префікс – визначає, яка саме функція буде реалізована:

- ‘p‘ (‘probability‘) позначає функцію розподілу, перший аргумент (‘q‘) – точка (або вектор), в якій рахується значення функції розподілу;
- ‘q‘ (‘quantile‘) позначає квантильну функцію (обернену до функції розподілу), перший аргумент (‘p‘) – задана ймовірність (або вектор ймовірностей), для якої потрібне значення квантиля;
- ‘d‘ (‘density‘) позначає щільність розподілу, перший аргумент (‘x‘) – точка (або вектор), в якій рахується значення функції щільності розподілу (для дискретної випадкової величини під поняттям «щільність» розуміється ймовірність прийняття випадковою величиною конкретного значення);
- ‘r‘ (‘random‘) позначає функцію, яка генерує випадкову величину (величини), яка має вказаний розподіл, перший аргумент (‘n‘) – це розмір вибірки, яку треба згенерувати.

Якщо перший параметр функцій ‘d‘, ‘q‘, ‘p‘ є вектором, то функція порахує відповідні значення для всіх значень елементів вектора та поверне вектор ймовірностей (‘q‘ – квантилів).

Загальними параметрами для всіх розподілів є

– логічний параметр для ‘p‘- та ‘q‘- функцій ‘log.p‘ (‘log‘ для ‘d‘-функцій), який за замовчуванням має значення ‘FALSE‘, і рахує ймовірності ‘p‘ як ‘log(p)‘, якщо задати його значення як ‘TRUE‘.

– логічний параметр ‘lower.tail‘ для функцій з префіксами ‘p‘, ‘q‘. Його значення за замовчуванням – ‘FALSE‘. Якщо задати ‘lower.tail=TRUE‘, то ‘p‘-функція буде замість функції розподілу обчислювати так звану функцію виживання $P\{\xi > x\} = 1 - F_\xi(x)$, а ‘q‘-функція – верхній квантиль, тобто $z_{1-\alpha}$.

Інші параметри є параметрами, особливими для кожного окремого розподілу. Вони мають певну область визначення, отже, бажано ознайомитися з довідкою по конкретному розподілу в R, перш ніж почати роботу з ним. Довідка по розподілу викликається або повною його назвою, або назвою одної з функцій, пов'язаних з ним. Наприклад, команди

```
?Normal
?pnorm
```

повернуть сторінку довідки щодо нормального розподілу.

20.2.1 Дискретні розподіли

Наведемо параметри та імена деяких дискретних розподілів.

Табл. 1: Основні дискретні розподіли в R

Розподіл	Ім'я в R	Параметри
Біноміальний	binom	size – кількість випробувань Бернуллі prob – ймовірність 'успіху' в одному випробуванні
Геометричний	geom	prob – ймовірність 'успіху' в одному випробуванні
Гіпергеометричний	hyper	m – кількість білих куль в урні n – кількість чорних куль в урні k – кількість куль, які вийняли з урни
Негативний біноміальний	nbinom	size – задана кількість 'успіхів' prob – ймовірність 'успіху' в одному випробуванні
Поліноміальний	multinom	x – вектор з r невід'ємних цілих чисел, таких, що $x_1 + x_2 + ... + x_k = n$ size – загальна кількість результатів (n), за замовчуванням $size = sum(x)$ prob – вектор ймовірностей $p_1, ..., p_k$ довжини k
Пуассона	pois	lambda – параметр $\lambda > 0$ – середнє значення, інтенсивність потоку

Приклад 20.1

Знайдемо ймовірності отримати різну (від 0 до 10) кількість успіхів, якщо кількість випробувань Бернуллі – 10, а ймовірність успіху – 0.5 (ймовірності того, що випадкова величина, що має біноміальний розподіл з параметрами n=10, p=0.5, прийме значення i=1,. . . ,10).

```
dbinom(1:10,size=10,prob=0.5)
[1] 0.0097656250 0.0439453125 0.1171875000 0.2050781250 0.2460937500
[6] 0.2050781250 0.1171875000 0.0439453125 0.0097656250 0.0009765625
```

Ймовірність отримати від 3 до 7 успіхів:

```
sum(dbinom(3:7,size=10,prob=0.5))
[1] 0.890625
```

Значення функції розподілу в точках 1, 3, 6, 10:

```
pbinom(c(1, 3, 6, 10),size=10,prob=0.5)
[1] 0.01074219 0.17187500 0.82812500 1.00000000
```

Приклад 20.2

Побудуємо вибірку з 20 елементів, що мають геометричний розподіл з параметром p=0.1:

```
rgeom(20, prob = 0.1)
[1]  4 16 17  3  1 13 18  4  7 29 34 11  8  2  6 24 14 14  3 31
```

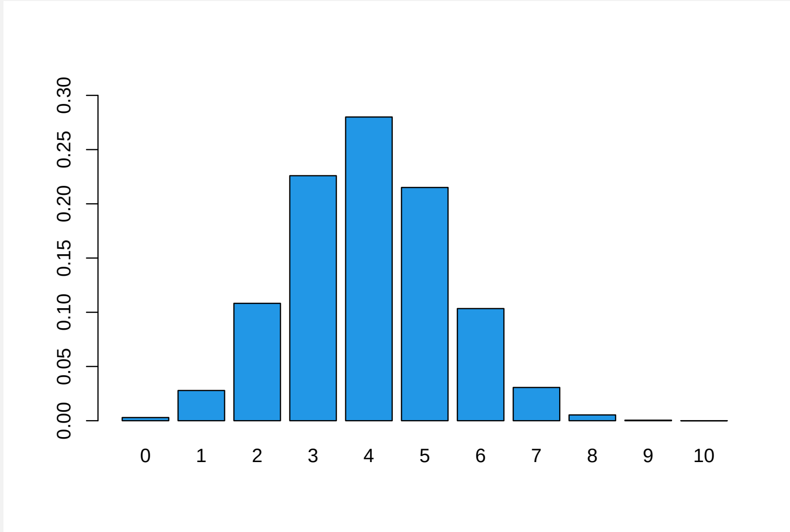
Квантилі цього розподілу (рівнів $\alpha = 0.1, ..., 1$, та $1 - \alpha$):

```
qgeom((1:10)/10, prob = 0.1)
[1]   0   2   3   4   6   8  11  15  21 Inf
qgeom((1:10)/10, prob = 0.1, lower.tail = FALSE)
[1] 21 15 11  8  6  4  3  2  0  0
```

Приклад 20.3 Гіпергеометричний розподіл

Нехай в урні 50 куль: 20 білих та 30 чорних. З урни, не повертаючи, виймають 10 куль. Побудуємо стовпчикову діаграму для ймовірностей значень випадкової величини, яка є кількістю вийнятих білих куль.

```
barplot(dhyper(0:10, 20, 30, 10), names=as.character(0:10),
ylim=c(0,0.3), col=12)
```



Бачимо, що найбільш ймовірно вийняти 4 білих кулі.

20.2.2 Неперервні розподіли

В таблиці 2 вказано деякі базові абсолютно неперервні розподіли випадкових величин. Зауважимо, що для деяких з них значення основних параметрів мають значення за замовчуванням. Наприклад, якщо скористатися однією з функцій нормального розподілу, не визначивши значення параметрів mean та sd, то функція буде застосована до стандартного нормального розподілу з параметрами, відповідно, 0 та 1.

Приклад 20.4 Рівномірний розподіл

Згенеруємо 50 псевдовипадкових чисел, що мають рівномірний розподіл на проміжку [0;10]:

Табл. 2: Основні неперервні розподіли в R

Розподіл	**Ім'я в R**	**Параметри**
Рівномірний	unif	min = 0 – початок проміжку max = 1 – кінець проміжку
Показниковий	exp	rate = 1 – параметр $\lambda = 1/\mathbf{M}\xi$
Нормальний	norm	mean=0 – математичне сподівання sd=1 – стандартне відхилення
Логнормальний	lnorm	meanlog=0 sdlog=1
Коші	cauchy	location = 0 scale = 1
Гамма	gamma	shape (параметр $\alpha \geq 0$) rate = 1 – параметр $\lambda > 0$ scale = 1/rate – альтернатива параметру rate $$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x},\ x \geq 0$$
Бета	beta	shape1, shape2 – параметри $a, b > 0$ $f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$, $x \in [0, 1]$
χ^2-розподіл Пірсона	chisq	df – кількість ступенів свободи
t-розподіл Стьюдента	t	df – кількість ступенів свободи
F-розподіл Фішера	f	df1 – кількість ступенів свободи чисельника df2 – кількість ступенів свободи знаменника

```
runif(50, min=0, max=10)
[1] 0.55105222 4.54865877 0.08103917 4.82421326 7.67126322 8.97944466
[7] 4.46019590 5.96269782 5.63057862 4.40542885 7.69685850 2.55730271
[13] 8.42158320 6.79192452 5.35514422 3.08432095 0.62589640 6.83660758
[19] 0.42493941 2.63578825 3.12393368 4.78995207 7.69145140 5.26840986
[25] 4.32960751 5.46814520 1.70446450 3.32909980 2.98892571 4.64556761
[31] 8.53846583 2.73931572 6.15658230 6.18415554 3.24368979 1.72226815
[37] 5.04359250 1.55049752 9.63805909 4.97190581 8.57391617 9.59205598
[43] 6.93559505 9.22914738 2.12727591 6.61891115 7.74917893 3.74282972
[49] 5.38442592 5.80301425
```

Приклад 20.5 Нормальний розподіл

Значення функції стандартного нормального розподілу в точці 1.96 можна знайти за допомогою функції

```
pnorm(1.96)
[1] 0.9750021
```

Значення функції розподілу нормального закону з параметрами 180, 20 в точці 196

```
pnorm(196, mean=180, sd=20)
[1] 0.7881446
```

Вектор децилів для стандартного нормального розподілу (вектор з послідовністю рівнів квантилів можна задавати в різний спосіб, наведений – не оптимальний, але інтуїтивно зрозумілий)

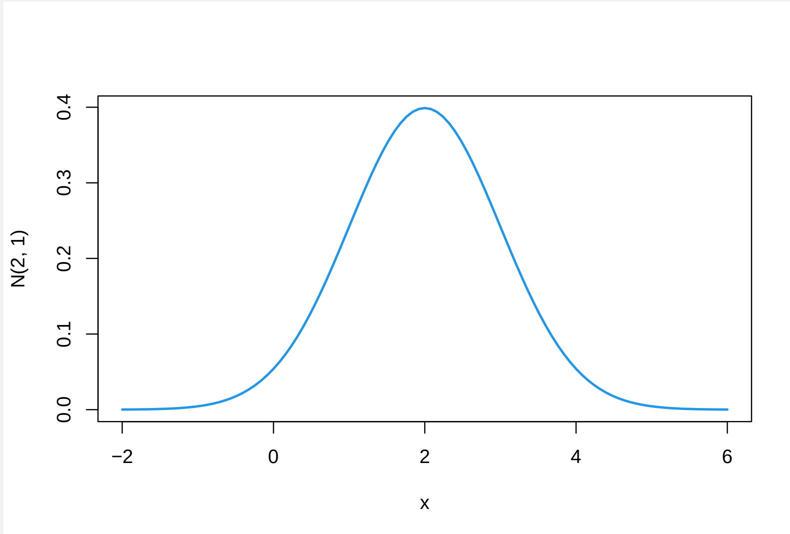
```
qnorm(c(0.1,0.2,0.3,0.4, 0.5, 0.6, 0.7, 0.8, 0.9))
[1] -1.2815516 -0.8416212 -0.5244005 -0.2533471 0.0000000 0.2533471 0.5244005
[8]  0.8416212 1.2815516
```

Згенеруємо 40 чисел, що мають стандартний нормальний розподіл:

```
rnorm(40)
[1] -1.162527534 -2.541593651  0.387676581  0.456161095  0.741387642
[6] -1.199063799 -0.256909586  0.843446876 -0.430361698 -0.040701509
[11] -0.883111447  1.006362241 -0.433081986  1.853119566  0.444346694
[16]  0.806669335  0.494931529  0.012523891  1.117890930 -0.613324728
[21] -0.107561402 -0.852084669 -1.667505370  0.863715269 -1.384934738
[26]  0.049469215 -0.007310359 -0.570362960  0.671287810 -0.737401975
[31]  1.413611011 -0.041488540 -1.056437509 -1.824710570 -2.582652175
[36]  0.306601664 -2.178124447  1.853522111  0.152857688 -0.316538886
```

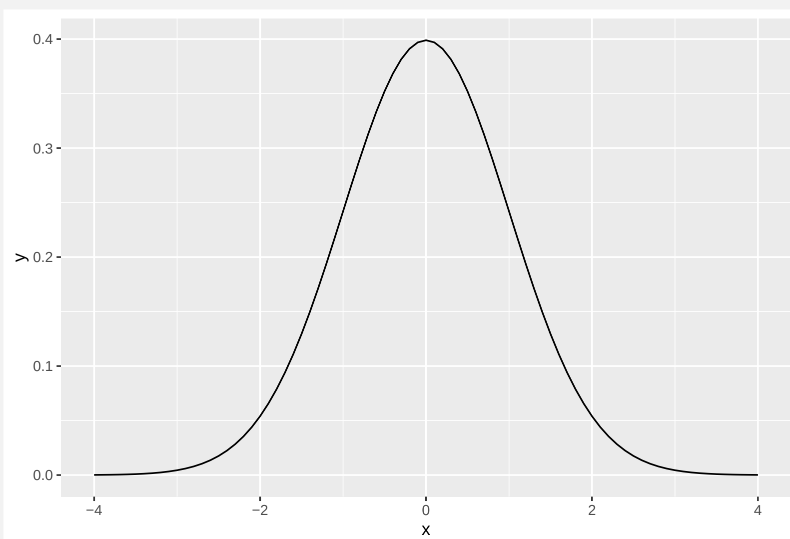
Графік щільності нормального розподілу з параметрами (2; 1):

```
curve(dnorm(x, 2, 1), from = -2, to = 6, ylab = paste(expression(N(2,1))),
col = 12, lwd = 2)
```



Можна зобразити подібні графіки за допомогою функцій пакету ggplot2, який має дещо іншу логіку використання, ніж графічні функції базових пакетів R. Для цього спочатку треба визначити вектор x – точки, в яких буде порахована щільність, застосувати функцію щільності в точках x, сформувати відповідний фрейм даних та зобразити результат (тут – графік щільності стандартного нормального розподілу):

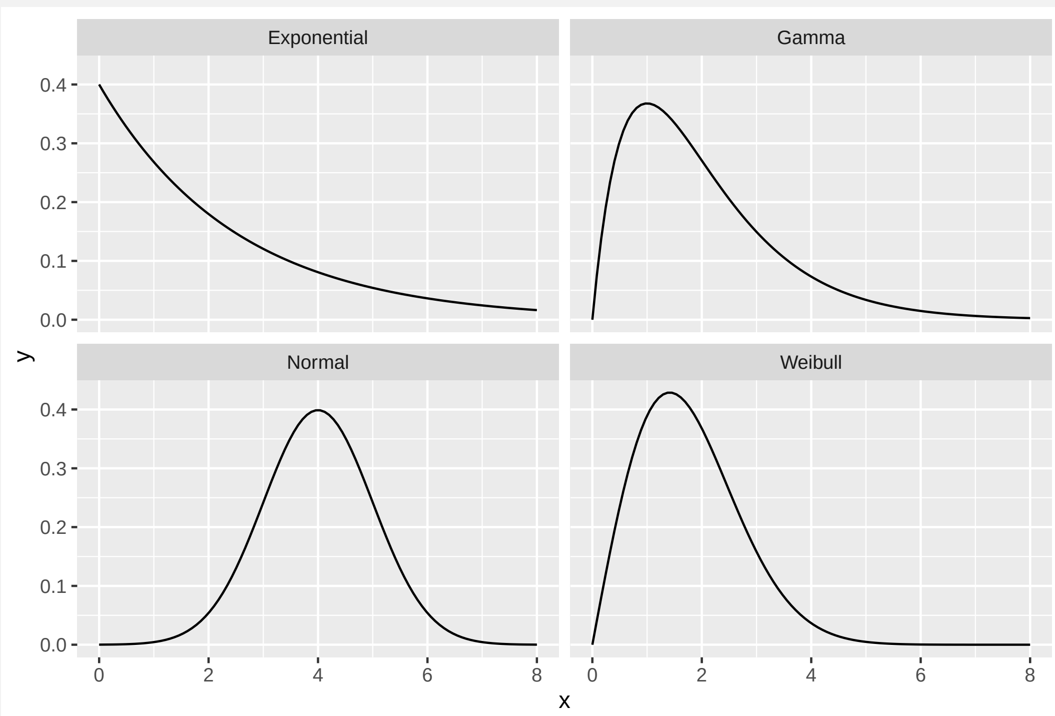
```
library(ggplot2)
x<-seq(-4, 4, .1)
dens<-data.frame(x=x, y=dnorm(x))
ggplot(dens, aes(x,y))+geom_line()
```



Приклад 20.6

Наступний код створює 2×2 графік для чотирьох щільностей:

```
x <- seq(from = 0, to = 8, length.out = 100)
#  задання області визначення щільності
# створюємо датафрейм з щільностями розподілів, вказуючи значення
параметрів:
df <- rbind(
 data.frame(x = x, dist_name = "Normal", y=dnorm(x, mean = 4, sd = 1)),
 data.frame(x = x, dist_name = "Exponential", y=dexp(x, rate  = 0.4)),
 data.frame(x = x, dist_name = "Weibull",y=dweibull(x,shape = 2,scale = 2)),
 data.frame(x = x, dist_name = "Gamma", y=dgamma(x, shape = 2, rate = 1)))
#створюємо графічне вікно, використовуючи facet_wrap, щоб об'єднати графіки:
ggplot(data = df, aes(x = x, y = y)) +
   geom_line() +
   facet_wrap(~dist_name) # "згортаємо" графіки за змінною dist_name.
```



Приклад 20.7 Моделювання закону великих чисел

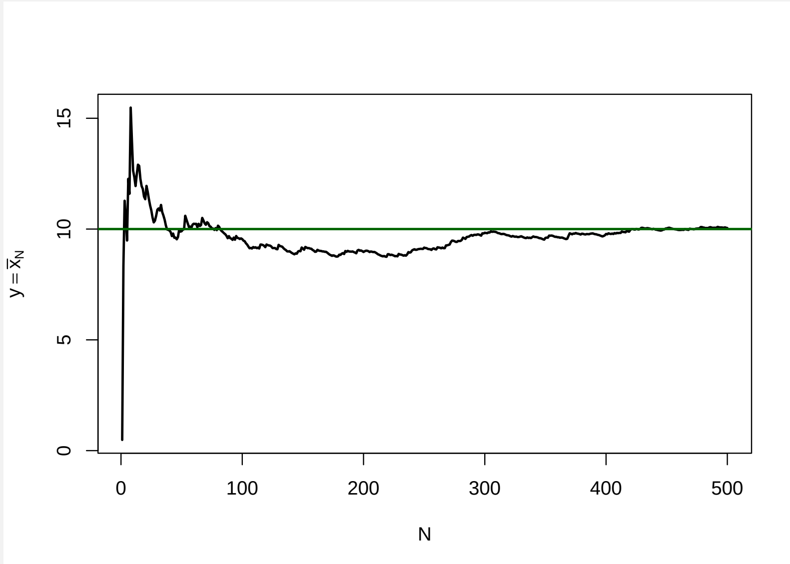
Згідно ЗВЧ, якщо ми маємо послідовність н.о.р.в.в. ξ_n, то за умови існування скінченного математичного сподівання цих випадкових величин їхнє середнє арифметичне $\bar{\xi}_n = \dfrac{1}{n}\sum_{i=1}^{n}\xi_i$ збігається за ймовірністю до $\mathbf{M}\xi_1$ зі збільшенням кількості доданків n.

Змоделюємо послідовність з $n = 500$ випадкових величин ξ_i, що мають показниковий розподіл з параметром $\lambda = 0.1$, і переконаємось, що послідовність середніх арифметичних $\bar{\xi}_N = \dfrac{1}{N}\sum_{i=1}^{N}\xi_i$, $N = 1, 2, ..., n$, дійсно наближається з ростом N до величини $\mathbf{M}\xi_1 = 1/\lambda = 10$.

```
set.seed(7)
# Задаємо параметри
n <- 500
lam <- 0.1
# Створюємо послідовність індексів та вектор з 500 показникових випадкових
# величин з заданим параметром 0.1
N <- 1:n
exp1 <- rexp(n, rate = lam)
# Створюємо вектор середніх значень змодельованих послідовностей
for (i in 1:500) {
  y[i] <- sum(exp1[1:i]/i)
}
```

Зобразимо отриману послідовність середніх на графіку

```
plot(N,y, type = "l", lwd=2, ylab = expression(y == bar(x)[N]))
abline(h = 1/lam, col = "darkgreen", lwd=2)
```



Бачимо, що з ростом кількості доданків N усереднена сума змодельованих реалізацій показникових випадкових величин дійсно стає близькою зо значення математичного сподівання 10.

Приклад 20.8 Моделювання центральної граничної теореми

Згідно ЦГТ, якщо у нас є послідовність н.о.р.в.в., то за відповідних умов випадкова величина, яка є їхнім середнім значенням, збігається при $n \to \infty$ за розподілом до нормально розподіленої випадкової величини:

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k \Longrightarrow N(\mathbf{M}\xi, \mathbf{D}\xi/n)$$

Змоделюємо 1000 рядків по 50 випадкових величин, які мають показниковий розподіл з параметром $\lambda = 0.1$, знайдемо середнє значення для кожного рядка (таким чином матимемо вибірку н.о.р. середніх значень), і подивимось на розподіл отриманої вибірки.

Для показникового розподілу $\mathbf{M}\xi = 1/\lambda$, $\mathbf{D}\xi = 1/\lambda^2$. Відповідно, згідно ЦГТ, середнє значення 50 таких доданків повинно мати приблизно нормальний розподіл з параметрами $\mu = 1/\lambda$ та $\sigma^2 = 1/(n\lambda^2)$.

Для $\lambda = 0.1$ маємо $\mu = 10$, $\sigma^2 = 100/50 = 2$.

Тепер змоделюємо всі потрібні величини. Створимо масив з 50000 експоненційних випадкових величин, сформуємо їх у матрицю exp, що має 1000 рядків, визначимо параметри:

```
set.seed(24)
lam <- 0.1
mu <- 1/lam
sigsq <- 1/lam^2/50
exp <- matrix(rexp(50000, rate = lam), 1000)
```

Створюємо вектор середніх значень для кожного з 1000 рядків (по 50 доданків) за допомогою функції apply пакету dplyr:

```
mexp <- apply(exp, 1, mean)
```

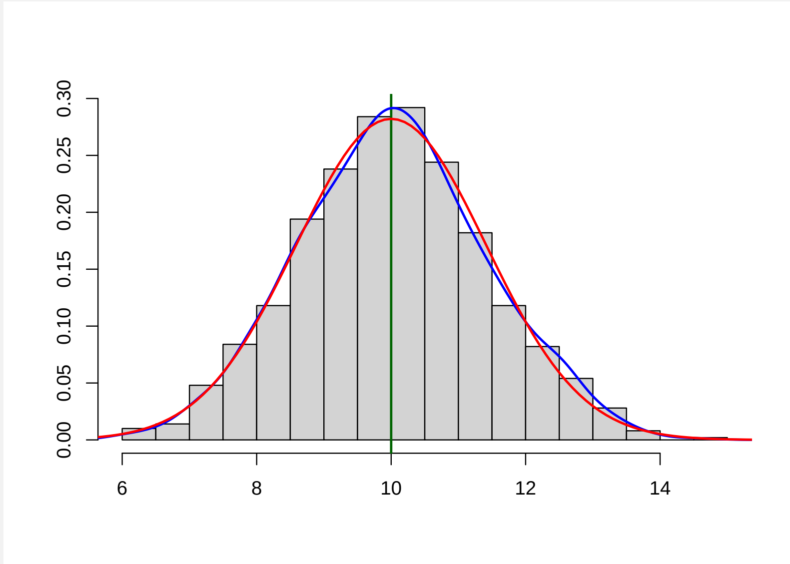
Обчислюємо оцінку вектору параметрів апроксимуючого нормального розподілу:

```
c(mean(mexp), var(mexp))
[1] 10.028530  1.944971
```

Бачимо, що отримані параметри дуже близькі до тих, які були обчислені теоретично (10 та 2 відповідно).

Щоб оцінити розподіл змодельованої вибірки середніх, побудуємо її гістограму, на якій зобразимо апроксимуючу функцію щільності змодельованого розподілу середніх значень (синя лінія) та теоретичну щільність граничного (згідно ЦГТ) нормального розподілу (червона лінія):

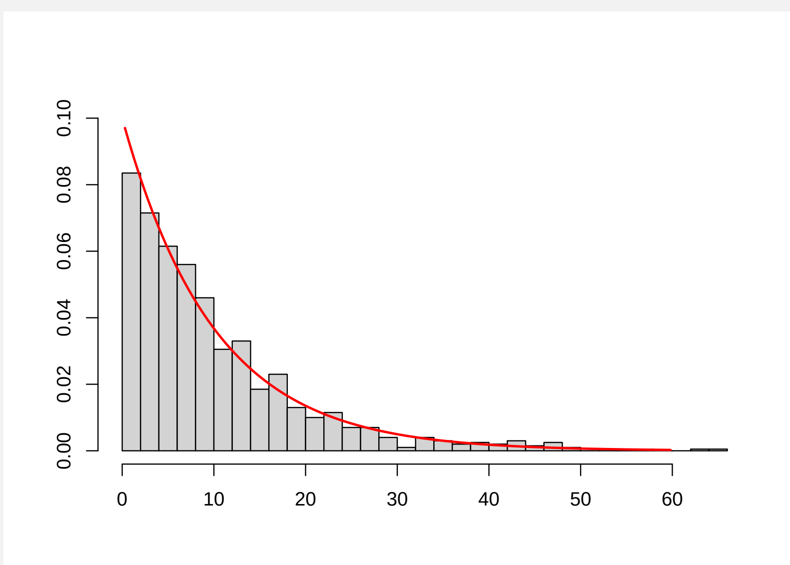
```
hist(mexp, breaks = "FD", probability = T, main = "", xlab = "", ylab = "")
abline(v = mu, col = "darkgreen", lwd=2)
lines(density(mexp), lwd=2, col="blue")
lines(seq(mu-3.5*sigsq, mu+3.5*sigsq, 0.1),
dnorm(seq(mu-3.5*sigsq, mu+3.5*sigsq, 0.1), mean=mu, sd=sqrt(sigsq)),
col="red", type = "l", lwd=2)
```



Бачимо, що змодельований розподіл дійсно дуже близький до теоретичного нормального.

Для порівняння подивимось на гістограму одного стовпчика нашого масиву – вибірки розміру 1000 з показникового розподілу з параметром $\lambda = 0.1$ (червоною лінією зображена щільність показникового розподілу з параметром 0.1):

```
hist(exp[,1], breaks = "FD", freq = F, main = " ", xlab = " ",
ylab = " ", ylim = c(0,0.1))
lines(seq(0.3,60,0.5), dexp(seq(0.3,60,0.5), rate = lam), col="red",
type = "l", lwd=2)
```



Отримана гістограма дійсно має форму показникового розподілу, і, вочевидь, відрізняється від розподілу середнього арифметичного 50 подібних доданків, яке має, згідно ЦГТ, нормальний розподіл, в чому ми маємо можливість переконатися, змоделювавши потрібні величини.

Зауважимо, що в різних пакетах R реалізовано багато інших розподілів окрім вказаних, їх можна подивитися на відповідній сторінці CRAN: https://cran.r-project.org/web/views/Distributions.html.

Окрім одно- та багатовимірних дискретних та неперервних розподілів представлено також різноманітні модифіковані версії розподілів та суміші розподілів.

Графічні можливості теж значно ширші за наведені приклади. Інформацію можна отримати в численних джерелах, присвячених опису функцій відповідних пакетів та прийомам роботи з ними. Доволі повна інформація по використанню функцій R, їхніх аргументів, використаних при обчисленнях методах міститься в довідці R. Для основних функцій довідка викликається знаком питання з назвою функції або ж стандартною функцією help():

```
?set.seed
help(hist)
```

20.3 Перевірка статистичних гіпотез

В R реалізовано багато критеріїв, як параметричних, так і непараметричних, перевірки статистичних гіпотез. У цьому підрозділі зосередимо свою увагу на тих, що описані у розділах 18 і 19 та проілюструємо розв'язання деяких прикладів з використанням R.

Приклад 20.9 Критерій згоди Колмогорова, приклад на стор. 245

Нехай маємо реалізацію вибірки x : $0.7; 2.3; 4.8; 9.7; 5.3; 6.8; 5.9; 8.7; 1.4; 3.2$. Треба перевірити гіпотезу про те, що величина, яка спостерігається, має рівномірний розподіл на $(0; 10)$ з рівнем значущості $\alpha = 0.05$:

$$H_0 : \qquad F_{\xi_0}(z) = \begin{cases} 0, & \text{якщо } z < 0, \\ \frac{z}{10}, & \text{якщо } z \in [0, 10], \\ 1, & \text{якщо } z > 10. \end{cases}$$

Розв'язок: для розв'язання цієї задачі використаємо критерій згоди Колмогорова. Аналітичний розв'язок можна знайти в прикладі на стор. 245. В R для виконання критерію згоди Колмогорова використовуємо функцію 'ks.test()'.

```
x<-c(0.7, 2.3, 4.8, 9.7, 5.3, 6.8, 5.9, 8.7, 1.4, 3)
ks.test(x, 'punif', min=0, max=10)

        Exact one-sample Kolmogorov-Smirnov test

data:  x
D = 0.12, p-value = 0.9949
alternative hypothesis: two-sided
```

Вихідними даними функції 'ks.test()' є найбільша відстань між емпіричною та гіпотетичною функціями розподілу (статистика D=0.12), що узгоджується з результатами прикладу на стор. 245, а також P-значення (p-value=0.9949). Статистичний висновок робиться таким чином: якщо P-значення більше за рівень значущості α, то нема підстав відхилити основну гіпотезу і H_0 приймається, інакше– H_0 відхиляється.

Для даного прикладу оскільки $0.9949 > 0.05$, то розподіл величини, що спостерігається, узгоджується з рівномірним розподілом на $(0, 10)$.

Функцію 'ks.test()' можна використовувати і для перевірки критерію Колмогорова-Смірнова на однорідність двох вибірок.

Приклад 20.10 Критерій Пірсона χ^2 про тип розподілу, приклад на стор. 247

При 50 підкиданнях монети герб з'явився 20 раз. Чи можна вважати, що монета симетрична? Прийняти $\alpha = 0.1$.

Розв'язок: спочатку в R створюємо реалізацію вибірки, яка складається з двадцяти '1', що відповідають гербам, та тридцяти '0', що відповідає решкам:

'c(rep(0,30),rep(1,20))'.

Для перевірки критерію χ^2 про тип розподілу використовуємо функцію 'chisq.test()'. Дані потрібно згрупувати, в R це можна зробити за допомогою функції 'table()'. Для перевірки симетричності монети основна гіпотеза полягає в тому, що ймовірності випадання герба ('1') та решки ('0') однакові,тобто розподіл є рівномірним на множині можливих значень:

$$H_0: \qquad p_0 = p_1 = \frac{1}{2}.$$

Для функції 'chisq.test()' за замовчуванням перевіряється саме рівномірний розподіл, тому як один з аргументів його не потрібно задавати.

```
x<-c(rep(0,30),rep(1,20))
chisq.test(table(x))


        Chi-squared test for given probabilities

data:  table(x)
X-squared = 2, df = 1, p-value = 0.1573
```

Статистика критерію дорівнює $X - squared = 2$, такий самий результат було і отримано аналітично у прикладі на стор. 247. У вихідних даних функції 'chisq.test()' також вказано ступінь сводоби $df = 1$ і Р-значення p-value $= 0.1573$. Оскільки Р-значення більше за рівень значущості α, $0.1573 > 0.1$, то нема підстав відхилити основну гіпотезу H_0. Отже, можна вважати, що монета симетрична, з рівнем значущості $\alpha = 0.1$.

Приклад 20.11 Критерій χ^2 Пірсона про тип розподілу, приклад на стор. 249

За час другої світової війни в південній частині Лондона впало 535 балістичних ракет. Уся ця територія була розділена на 576 ділянок площею по 0.25 км^2. Нижче наведені числа ділянок n_k на які впало k ракет

k	0	1	2	3	4	5
n_k	229	211	93	35	7	1

Чи узгоджуються ці дані з гіпотезою про те, що число ракет, які впали на кожну ділянку, мають розподіл Пуассона? Прийняти $\alpha = 0.05$.

Розв'язок: як і для попереднього випадку, використаємо функцію 'chisq.test()'. Для даного прикладу гіпотетичний розподіл відомий з точністю до параметра. Тому після формування вибірки знаходимо оцінку параметра пуассонівського розподілу як вибіркове середнє, обчислюємо вектор ймовірностей розподілу Пуассона із знайденим параметром, а потім вибіркові значення та вектор ймовірностей вводимо як аргументи 'chisq.test()'

```
 y<-c(rep(0,229),rep(1,211),rep(2,93),rep(3,35),rep(4,7),rep(5,1))
lambda<-mean(y)
p1<-c(dpois(0:4,lambda), 1-ppois(4,lambda))
chisq.test(table(y),p=p1)


        Chi-squared test for given probabilities

data:  table(y)
X-squared = 1.1724, df = 5, p-value = 0.9475

 qchisq(0.95,4)
[1] 9.487729
```

Статистика критерію дорівнює $X - squared = 1.1724$. Знайдемо квантиль χ^2-розподілу з 4 ступенями свободи порядку 0.95, використовуючи функцію 'qchisq(0.95,4)'. Оскільки статистика критерію менша за табличне значення $1.1724 < 9.487729$, то немає підстав відхиляти основну гіпотезу, отже, дані узгоджуються з гіпотезою про те, що число ракет, які впали на кожну ділянку, мають розподіл Пуассона.

Приклад 20.12 Критерій незалежності χ^2, приклад на стор. 255

Нижче наведені результати опитування 100 студентів перших трьох курсів, яким ставилося одне питання: "Чи вважаєте ви, що куріння заважає навчанню?"

З'ясувати чи підтверджують ці дані припущення про те, що відношення до куріння студентів на різних курсах різне? Прийняти $\alpha = 0.05$.

Відповідь/Курс	1	2	3	Сума
Так	-	30	25	55
Не знаю	8	5	7	20
Ні	15	10	-	25
Сума	23	45	32	100

Розв'язок: у даній задачі перевіряється гіпотеза незалежності двох змінних, а саме відношення до куріння та курс, на якому навчається студент. Основна гіпотеза

$$H_0 : p_{ij} = p_{i\bullet} \cdot p_{\bullet j} \quad \text{(відношення до куріння не залежить від курсу)},$$

альтернативна гіпотеза

$$H_1 : p_{ij} \neq p_{i\bullet} \cdot p_{\bullet j} \quad \text{(відношення до куріння на різних курсах різне, ці характеристики залежні)}.$$

Для перевірки незалежності в R використовуємо функцію 'chisq.test()'. У якості аргумента вставляємо матрицю частот без сумарних значень.

```
x<-rbind(c(0,30,25),c(8,5,7),c(15,10,0))
chisq.test(x)


       Pearson`s Chi-squared test

data:  x
X-squared = 44.241, df = 4, p-value = 5.716e-09
```

Статистика критерію дорівнює X-squared = 44.241, як і в прикладі на стор. 255. Оскільки Р-значення p-value = 5.716e-09 = $5.716 \cdot 10^{-9}$ значно менше за рівень значущості $\alpha = 0.05$, то основна гіпотеза відхиляється, а отже, відношення до куріння студентів на різних курсах різне.

Приклад 20.13 t-критерій Стьюдента для однієї вибірки, приклад на стор.266

Для перевірки твердження виробника про те, що генератор за зміну споживає в середньому 20 л пального, здійснили 10 випробувань. За 10 змін споживання генератора встановили:

$$19.1, 18.6, 19.1, 18.1, 16.6, 20.1, 19.8, 21.1, 24.4, 21.6.$$

Перевірити твердження виробника при рівні значущості 0.05.

Розв'язок: t-критерій Стьюдента для однієї вибірки в R реалізується за допомогою функції 't.test()'. Якщо як аргумент цієї функції задати тільки значення вибірки, то перевіряється гіпотеза про нульове середнє з двосторонньою альтернативою. Запис в аргументі $mu = 20$ означає, що основною гіпотезою є $H_0 : m = m_0 = 20$. Альтернативна гіпотеза $H_1 : m \neq m_0$ в R вибирається за замовчуванням. Для зміни альтернативи в функції 't.test()' варто для аргументу *alternative* вибрати одне із значень *"two.sided"*, *"less"*, *"greater"*.

```
x<-c(19.1, 18.6, 19.1, 18.1, 16.6, 20.1, 19.8, 21.1, 24.4, 21.6)
t.test(x,mu=20)


        One Sample t-test

data:  x
t = -0.22035, df = 9, p-value = 0.8305
alternative hypothesis: true mean is not equal to 20
95 percent confidence interval:
 18.31009 21.38991
sample estimates:
mean of x
    19.85
```

Вихідними даними є

- значення критерію $t = -0.22035$,
- ступінь свободи $df = 9$,
- Р-значення p-value = 0.8305,
- 95 % довірчий інтервал для середнього $(18.31009; 21.38991)$ та
- вибіркове середнє 19.85.

Оскільки Р-значення більше за рівень значущості $0.8305 > 0.05$, то нема підстав відхиляти основну гіпотезу. Отже, дані не суперечать твердженню виробника про те, що генератор за зміну споживає в середньому 20 л пального.

Приклад 20.14 t-критерій Стьюдента для двох вибірок, приклад на стор.269

Нижче наведено дані вимірювання рівня чутливості однорідного фотоматеріалу за допомогою двох фотосенсометрів. Чи можна вважати, що між показниками приладів немає систематичного розходження (рівень значущості 0.05)?

Перший прилад: 5.0, 5.1, 5.5, 6.0, 5.5, 3.6, 1.8, 7.8, 6.9, 2.7;

Другий прилад: 6.1, 5.4, 5.8, 2.7, 2.8, 2.8, 4.2, 2.4.

Розв'язок: як і в попередньому прикладі, використаємо функцію 't.test()' тільки вже для двох вибірок із зазначаенням того, що розглядається випадок для однакових дисперсій $var.equal = T$.

```
x<-c(5.0, 5.1, 5.5, 6.0, 5.5, 3.6, 1.8, 7.8, 6.9, 2.7)
y<-c(6.1, 5.4, 5.8, 2.7, 2.8, 2.8, 4.2, 2.4)
t.test(x,y,var.equal = T)

        Two Sample t-test

data:  x and y
t = 1.1842, df = 16, p-value = 0.2536
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.762544  2.692544
sample estimates:
mean of x mean of y
    4.990     4.025
```

Як результат отримаємо:

- значення критерію $t = 1.1842$,
- ступінь свободи $df = 16$,
- Р-значення p-value $= 0.2536$,
- 95 % довірчий інтервал для різниці середніх $(-0.762544, 2.692544)$ та
- пораховані вибіркові середні для кожної вибірки $4.990; 4.025$.

Оскільки Р-значення більше за рівень значущості $0.2536 > 0.05$, то нема підстав відхиляти основну гіпотезу. Отже, між показниками приладів нема систематичного розходження.

Приклад 20.15 Критерій перевірки рівності дисперсій, приклад на стор. 271

Чи можна в умовах попереднього прикладу вважати, що точність вимірювань на обох фотосенсометрах однакова (рівень значущості 0.05)?

Розв'язок: Для перевірки рівності дисперсій двох вибірок в R використовується 'var.test( )'.

```
x<-c(5.0, 5.1, 5.5, 6.0, 5.5, 3.6, 1.8, 7.8, 6.9, 2.7)
y<-c(6.1, 5.4, 5.8, 2.7, 2.8, 2.8, 4.2, 2.4)
var.test(x,y)


        F test to compare two variances

data:  x and y
F = 1.4118, num df = 9, denom df = 7, p-value = 0.664
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.2927044 5.9252930
sample estimates:
ratio of variances
          1.411777
```

Вихідними даними функції 'var.test( )' є

- статистика критерію $F = 1.4118$, що є відношенням вибіркових дисперсій,
- ступені свободи $numdf = 9$, $denomdf = 7$,
- Р-значення p-value $= 0.664$,
- 95 % довірчий інтервал для відношення дисперсій $(0.2927044; 5.9252930)$.

Оскільки Р-значення більше за рівень значущості $0.664 > 0.05$, то нема підстав відхиляти основну гіпотезу. Отже, точність вимірювань на обох фотосенсометрах однакова.

21 ЗАДАЧІ ДЛЯ ПРОГРАМІСТІВ

Задача 21.1. Наступний програмний код C++ 17 містить функцію shuffling, яка за задумом автора програми має тасувати елементи вектора, тобто, із усіх можливих перестановок елементів вектора навмання "обирати" якусь одну. Чи коректно працюватиме ця функція, якщо вважати, що умова (rand() $<$ RAND_MAX/2) насправді виконується з ймовірністю 0.5?

```
// C++ 17
#include <iostream>
#include <ctime>
#include<vector>
using namespace std;

void shuffling(vector<int> &v){
    for(int i=0; i<v.size()-1; i++){
        for(int j=i+1; j<v.size(); j++){
            if(rand() < RAND_MAX/2) swap(v[i], v[j]);
        }
    }
}

int main()
{
    srand(time(NULL));
    vector<int> v;
    for(int i=0; i<10; i++) v.push_back(i);
    shuffling(v);
    for(int i=0; i<v.size(); i++) cout << v[i] << " ";
    return 0;
}
```

Задача 21.2. Спираючись на властивості біноміальних коефіцієнтів, напишіть функцію, яка є рекурентною по двом своїм аргументам і обчислює значення біноміального коефіцієнта.

Задача 21.3. Дано функцію $f()$, яка повертає випадкове число з множини $\{1, 2, 3, 4, 5, 6, 7\}$ (кожне число має однакові шанси бути обраним). Напишіть функцію $g()$, яка буде повертати випадкове число з множини $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ (і теж кожне число повинно мати однакові шанси бути обраним), але так, щоб функція $g()$ могла із усіх можливих інших функцій викликати лише функцію $f()$. Як програмно перевірити правильність виконання цього завдання?

Задача 21.4. Дано функцію $f()$, яка повертає true з деякою невідомою ймовірністю p або false з ймовірністю $1-p$. Напишіть функцію $g()$, яка буде повертати true та false з ймовірністю 0.5, але так, щоб функція $g()$ могла із усіх можливих інших функцій викликати лише функцію $f()$. Як програмно перевірити правильність виконання цього завдання?

Задача 21.5.* Дано натуральне число n і масив *унікальних* цілих чисел **overleaf.** Розробіть алгоритм для вибору випадкового цілого числа в діапазоні $[0, n-1]$, якого немає в масиві **overleaf.** Будь-яке ціле число, яке знаходиться у згаданому діапазоні та не в масиві **overleaf,** повинно мати однакову ймовірність повернення. Оптимізуйте свій алгоритм таким чином, щоб він мінімізував кількість викликів вбудованої випадкової функції обраної мови програмування.

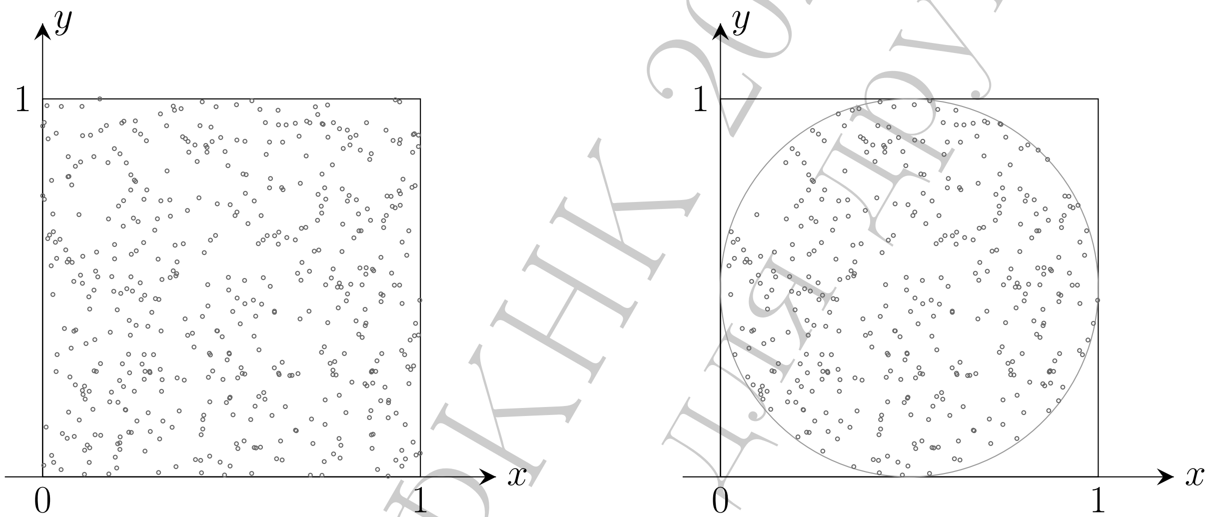
Задача 21.6. Допишіть функцію **randomizator** так, щоб вона повертала натуральне число з множини $\{1, 2...n\}$ і кожне число має однакові шанси бути обраним. Зверніть увагу, що не можна нічого писати до команди **return**.

```cpp
// C++ 17
#include <iostream>
#include <ctime>
using namespace std;

int randomizator(int n){
    return . . .
}

int main()
{
    srand(time(NULL));
    for(int i=0; i<10; i++) cout << randomizator(3) << " ";
    return 0;
}
```

Задача 21.7. Чи є цей цикл **while** скінченним? Відповідь обґрунтувати.

```cpp
// C++ 17
#include <iostream>
#include <ctime>
using namespace std;

int main()
{
```

```
    srand(time(NULL));
    int a = 2023, b;
    while(a){
        b = rand(); // випадкове число від 0 до RAND_MAX =  2147483647
        if(a>0) a-= b; else a+=b;
    }
    cout << "\n\nThe end.";
    return 0;
}
```



(а) (б)

Рис. 21.1: до задачі 21.8

Задача 21.8. Якщо випадкові величини ξ та η незалежні і мають рівномірний розподіл на відрізку $[0,1]$, то точка (ξ,η) має рівномірний розподіл у квадраті $S=[0,1]\times[0,1]$. Цей простий факт можна використати для генерування деякої кількості випадкових точок у квадраті S (див. рис. 21.1а). Але якщо потрібно, щоб точка (ξ,η) мала рівномірний розподіл у крузі $C=\left\{(x,y)|\ (x-0.5)^2+(y-0.5)^2\le 0.5^2\right\}$, який вписаний у квадрат K, то можна генерувати точки у квадраті S, а ті із них, що не потрапляють до круга C, відкидати (ігнорувати) (див. рис. 21.1б), але для відповідної програми це означатиме, що значну кількість точок ми генеруємо марно, витрачаючи на це час та пам'ять. Як обирати точки, щоб уникнути такого марнування? Напишіть відповідну програму. Подумайте, як перевірити, чи генерує вона насправді точки, які мають рівномірний розподіл у крузі C.
Вказівка: може допомогти перехід від декартової до полярної системи координат.

Задача 21.9. Продовження задачі 21.8 Розв'язати задачу 21.8, якщо замість круга буде а) деякий трикутник, б) деякий опуклий многокутник.

22 Додатки

22.1 Значення функції $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$

Значення функції $\Phi(x)$ для аргументу від 0 до 3 з кроком 0.01. Щоб знайти наприклад $\Phi(1.71)$, ми дивимося на рядок 1.7 і стовпець 0.01, звідки отримуємо $\Phi(1.71) = 0.9564$.

x	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.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	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.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	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	**.9564**	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

Табл. 3: Функція $\Phi(x)$

Якщо потрібно знайти табличне значення функції $\Phi(x)$ для від'ємного аргумента, то слід спочатку скористатись співвідношенням

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

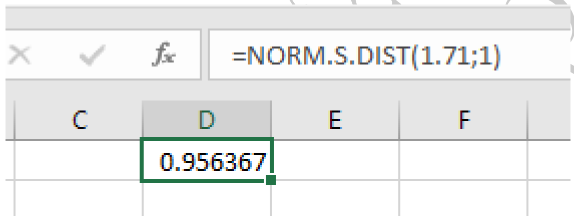
Для значень x, більших за 3, значення функції $\Phi(x)$ можна вважати рівними 1.

За необхідності значення функції $\Phi(x)$ можна обчислювати і іншими засобами. Наприклад:

- В електронних таблицях MS Excel значення $\Phi(x)$ обчислюємо за формулою

  =NORM.S.DIST(x;1)

  Наприклад, вираз =NORM.S.DIST(1.71;1) матиме наступне наближене значення:

- В пакеті MathCad значення $\Phi(x)$ обчислюємо за формулою

  pnorm(x, 0, 1)

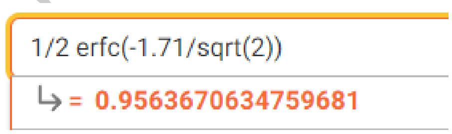
  Наприклад, вираз pnorm(1.71, 0, 1) матиме наступне наближене значення:

  $$\mathrm{pnorm}(1.71, 0, 1) = 0.956367$$

- На онлайн-платформі www.wolframalpha.com значення $\Phi(x)$ обчислюють за формулою

  1/2erfc(-x/sqrt(2))

  Наприклад, вираз 1/2erfc(-1.71/sqrt(2)) матиме наступне наближене значення:

- У мові програмування і програмному середовищі для статистичних обчислень R значення $\Phi(x)$ обчислюємо за формулою pnorm(x). Наприклад, вираз pnorm(1.71) матиме наступне наближене значення:

```
> pnorm(1.71)
[1] 0.9563671
```

22.2 Значення функції Колмогорова $K(x) = \sum\limits_{k=-\infty}^{\infty} (-1)^k e^{-2k^2x^2}$

x	$K(x)$	x	$K(x)$	x	$K(x)$	x	$K(x)$	x	$K(x)$
0.36	0.001	0.68	0.256	1	0.73	1.32	0.939	1.64	0.991
0.37	0.001	0.69	0.272	1.01	0.741	1.33	0.942	1.65	0.991
0.38	0.001	0.7	0.289	1.02	0.751	1.34	0.945	1.66	0.992
0.39	0.002	0.71	0.306	1.03	0.761	1.35	0.948	1.67	0.992
0.4	0.003	0.72	0.322	1.04	0.77	1.36	0.951	1.68	0.993
0.41	0.004	0.73	0.339	1.05	0.78	1.37	0.953	1.69	0.993
0.42	0.006	0.74	0.356	1.06	0.789	1.38	0.956	1.7	0.994
0.43	0.007	0.75	0.373	1.07	0.798	1.39	0.958	1.71	0.994
0.44	0.01	0.76	0.39	1.08	0.806	1.4	0.96	1.72	0.995
0.45	0.013	0.77	0.406	1.09	0.814	1.41	0.963	1.73	0.995
0.46	0.016	0.78	0.423	1.1	0.822	1.42	0.965	1.74	0.995
0.47	0.02	0.79	0.44	1.11	0.83	1.43	0.967	1.75	0.996
0.48	0.025	0.8	0.456	1.12	0.837	1.44	0.968	1.76	0.996
0.49	0.03	0.81	0.472	1.13	0.845	1.45	0.97	1.77	0.996
0.5	0.036	0.82	0.488	1.14	0.851	1.46	0.972	1.78	0.997
0.51	0.043	0.83	0.504	1.15	0.858	1.47	0.973	1.79	0.997
0.52	0.05	0.84	0.519	1.16	0.864	1.48	0.975	1.8	0.997
0.53	0.059	0.85	0.535	1.17	0.871	1.49	0.976	1.81	0.997
0.54	0.068	0.86	0.55	1.18	0.877	1.5	0.978	1.82	0.997
0.55	0.077	0.87	0.565	1.19	0.882	1.51	0.979	1.83	0.998
0.56	0.088	0.88	0.579	1.2	0.888	1.52	0.98	1.84	0.998
0.57	0.099	0.89	0.593	1.21	0.893	1.53	0.982	1.85	0.998
0.58	0.11	0.9	0.607	1.22	0.898	1.54	0.983	1.86	0.998
0.59	0.123	0.91	0.621	1.23	0.903	1.55	0.984	1.87	0.998
0.6	0.136	0.92	0.634	1.24	0.908	1.56	0.985	1.88	0.998
0.61	0.149	0.93	0.647	1.25	0.912	1.57	0.986	1.89	0.998
0.62	0.163	0.94	0.66	1.26	0.916	1.58	0.986	1.9	0.999
0.63	0.178	0.95	0.673	1.27	0.921	1.59	0.987	1.91	0.999
0.64	0.193	0.96	0.685	1.28	0.925	1.6	0.988	1.92	0.999
0.65	0.208	0.97	0.696	1.29	0.928	1.61	0.989	1.93	0.999
0.66	0.224	0.98	0.708	1.3	0.932	1.62	0.99	1.94	0.999
0.67	0.24	0.99	0.719	1.31	0.935	1.63	0.99	1.95	0.999

Табл. 4: Функція Колмогорова

22.3 Критичні значення λ_α для розподілу Колмогорова

$$P\{\lambda_n > \lambda_\alpha\} = \alpha$$

α	0.2	0.1	0.05	0.02	0.01	0.001
λ_α	1.073	1.224	1.358	1.520	1.627	1.950

Табл. 5: Критичні значення λ_α для розподілу Колмогорова

22.4 Квантилі розподілу Стьюдента

Квантилі порядку p розподілу Стьюдента з n степенями свободи.

$n \backslash p$	0.9	0.95	0.975	0.99	0.995
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750

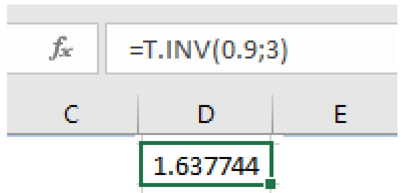
Табл. 6: Квантилі розподілу Стьюдента

За необхідності квантилів розподілу Стьюдента можна обчислювати і іншими засобами. Наприклад:

- В електронних таблицях MS Excel значення квантилів розподілу Стьюдента обчислюємо за формулою

  =T.INV(p;n)

  Наприклад, вираз =T.INV(0.9;3) матиме наступне наближене значення:



- В пакеті MathCad значення квантилів розподілу Стьюдента обчислюємо за формулою

  qt(p, n)

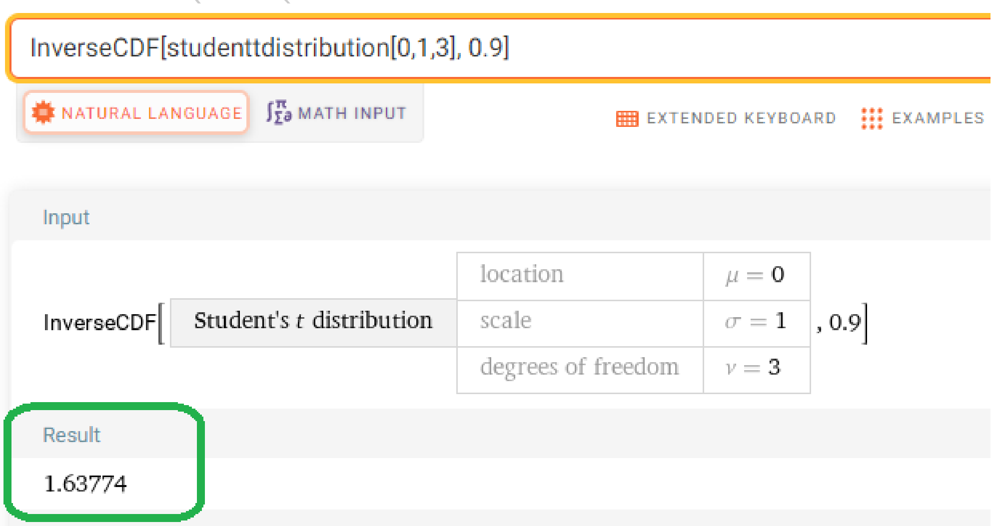
  Наприклад, вираз qt(0.9, 3) матиме наступне наближене значення:

  $$qt(0.9,3) = 1.638$$

- На онлайн-платформі www.wolframalpha.com значення квантилів розподілу Стьюдента обчислюють за формулою

  InverseCDF[studenttdistribution[0,1,n],p]

  Наприклад, вираз InverseeCDF[StudentsTDistribution[0,1,3],0.9] матиме наступне наближене значення:



22.5 Квантилі розподілу χ^2

Квантилі порядку p розподілу χ^2 з n степенями свободи.

n \ p	0.005	0.01	0.025	0.05	0.95	0.975	0.99	0.995
1	-	-	0.001	0.004	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672

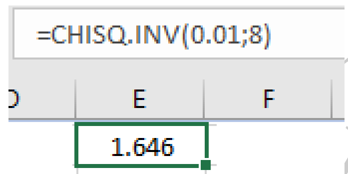
Табл. 7: Квантилі розподілу χ^2

За необхідності квантилів розподілу χ^2 можна обчислювати і іншими засобами. Наприклад:

- В електронних таблицях MS Excel значення квантилів розподілу χ^2 обчислюємо за формулою

  =CHISQ.INV(p; n)

  Наприклад, вираз =CHISQ.INV(0.01;8) матиме наступне наближене значення:

- В пакеті MathCad значення квантилів розподілу χ^2 обчислюємо за формулою

  qchisq(p, n)

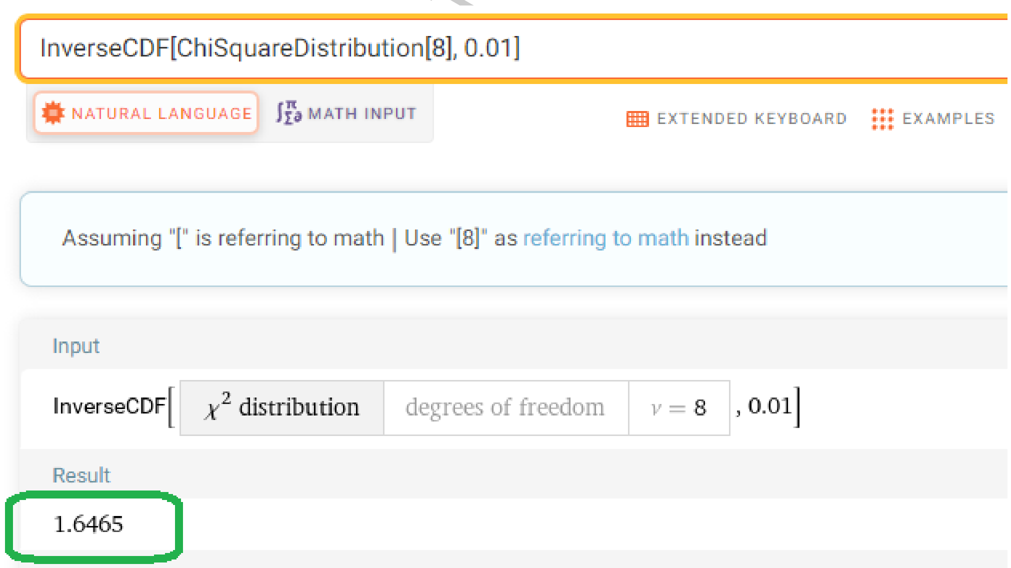
  Наприклад, вираз qchisq(0.01, 8) матиме наступне наближене значення:

  qchisq(0.01,8) = 1.646

- На онлайн-платформі www.wolframalpha.com значення квантилів розподілу χ^2 обчислюють за формулою

  InverseCDF[ChiSquareDistribution[n], p]

  Наприклад, вираз InverseCDF[ChiSquareDistribution[8], 0.01] матиме наступне наближене значення:

22.6 Квантилі розподілу Фішера

Квантилі порядку p розподілу Фішера з (n, m) степенями свободи.

$p = 0.95$

$m \backslash n$	1	2	3	4	5	6	7	8	10	12	20	40	100
1	161	200	216	225	230	234	237	239	242	244	248	251	253
2	19	19	19	19	19	19	19	19	19	19	19	19	19
3	10.1	9.6	9.3	9.1	9.0	8.9	8.9	8.8	8.8	8.7	8.7	8.6	8.6
4	7.7	6.9	6.6	6.4	6.3	6.2	6.1	6.0	6.0	5.9	5.8	5.7	5.7
5	6.6	5.8	5.4	5.2	5.1	5.0	4.9	4.8	4.7	4.7	4.6	4.5	4.4
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.06	4.00	3.87	3.77	3.71
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.64	3.57	3.44	3.34	3.27
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.35	3.28	3.15	3.04	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.14	3.07	2.94	2.83	2.76
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.98	2.91	2.77	2.66	2.59
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.85	2.79	2.65	2.53	2.46
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.75	2.69	2.54	2.43	2.35
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.67	2.60	2.46	2.34	2.26
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.60	2.53	2.39	2.27	2.19
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.54	2.48	2.33	2.20	2.12
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.49	2.42	2.28	2.15	2.07
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.45	2.38	2.23	2.10	2.02
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.41	2.34	2.19	2.06	1.98
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.38	2.31	2.16	2.03	1.94
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.35	2.28	2.12	1.99	1.91
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.32	2.25	2.10	1.96	1.88
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.30	2.23	2.07	1.94	1.85
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.27	2.20	2.05	1.91	1.82
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.25	2.18	2.03	1.89	1.80
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.24	2.16	2.01	1.87	1.78
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.22	2.15	1.99	1.85	1.76
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.20	2.13	1.97	1.84	1.74
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.19	2.12	1.96	1.82	1.73
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.18	2.10	1.94	1.81	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.16	2.09	1.93	1.79	1.70

Табл. 8: Квантилі розподілу Фішера, $p = 0.95$

Квантилі порядку p розподілу Фішера з (n, m) степенями свободи.

$$p = 0.975$$

$m \backslash n$	1	2	3	4	5	6	7	8	9	10	12	20	40	100
1	648	799	864	900	922	937	948	957	963	969	977	993	1006	1013
2	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5
3	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.3	14.2	14.0	14.0
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.56	8.41	8.32
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.33	6.18	6.08
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.17	5.01	4.92
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.47	4.31	4.21
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.00	3.84	3.74
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.67	3.51	3.40
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.42	3.26	3.15
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.23	3.06	2.96
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.07	2.91	2.80
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	2.95	2.78	2.67
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.84	2.67	2.56
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.76	2.59	2.47
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.68	2.51	2.40
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.62	2.44	2.33
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.56	2.38	2.27
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.51	2.33	2.22
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.46	2.29	2.17
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.42	2.25	2.13
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.39	2.21	2.09
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.36	2.18	2.06
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.33	2.15	2.02
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.30	2.12	2.00
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.28	2.09	1.97
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.25	2.07	1.94
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.23	2.05	1.92
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.21	2.03	1.90
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.20	2.01	1.88

Табл. 9: Квантилі розподілу Фішера, $p = 0.975$

Квантилі порядку p розподілу Фішера з (n, m) степенями свободи.

$$p = 0.99$$

$m \backslash n$	1	2	3	4	5	6	7	8	10	12	20	40	100
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.2	27.1	26.7	26.4	26.2
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.5	14.4	14.0	13.7	13.6
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.1	9.9	9.6	9.3	9.1
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.87	7.72	7.40	7.14	6.99
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.62	6.47	6.16	5.91	5.75
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.81	5.67	5.36	5.12	4.96
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.26	5.11	4.81	4.57	4.41
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.85	4.71	4.41	4.17	4.01
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.54	4.40	4.10	3.86	3.71
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.30	4.16	3.86	3.62	3.47
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.10	3.96	3.66	3.43	3.27
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	3.94	3.80	3.51	3.27	3.11
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.80	3.67	3.37	3.13	2.98
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.69	3.55	3.26	3.02	2.86
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.59	3.46	3.16	2.92	2.76
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.51	3.37	3.08	2.84	2.68
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.43	3.30	3.00	2.76	2.60
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.37	3.23	2.94	2.69	2.54
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.31	3.17	2.88	2.64	2.48
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.26	3.12	2.83	2.58	2.42
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.21	3.07	2.78	2.54	2.37
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.17	3.03	2.74	2.49	2.33
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.13	2.99	2.70	2.45	2.29
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.09	2.96	2.66	2.42	2.25
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.06	2.93	2.63	2.38	2.22
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.03	2.90	2.60	2.35	2.19
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.00	2.87	2.57	2.33	2.16
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	2.98	2.84	2.55	2.30	2.13

Табл. 10: Квантилі розподілу Фішера, $p = 0.99$

Квантилі порядку p розподілу Фішера з (n, m) степенями свободи.

$$p = 0.995$$

m \ n	1	2	3	4	5	6	7	8	10	12	20	40	100
2	199	199	199	199	199	199	199	199	199	199	199	199	199
3	55.6	49.8	47.5	46.2	45.4	44.8	44.4	44.1	43.7	43.4	42.8	42.3	42.0
4	31.3	26.3	24.3	23.2	22.5	22.0	21.6	21.4	21.0	20.7	20.2	19.8	19.5
5	22.8	18.3	16.5	15.6	14.9	14.5	14.2	14.0	13.6	13.4	12.9	12.5	12.3
6	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.25	10.03	9.59	9.24	9.03
7	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.38	8.18	7.75	7.42	7.22
8	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.21	7.01	6.61	6.29	6.09
9	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.42	6.23	5.83	5.52	5.32
10	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.85	5.66	5.27	4.97	4.77
11	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.42	5.24	4.86	4.55	4.36
12	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.09	4.91	4.53	4.23	4.04
13	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.82	4.64	4.27	3.97	3.78
14	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.60	4.43	4.06	3.76	3.57
15	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.42	4.25	3.88	3.58	3.39
16	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.27	4.10	3.73	3.44	3.25
17	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.14	3.97	3.61	3.31	3.12
18	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.03	3.86	3.50	3.20	3.01
19	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	3.93	3.76	3.40	3.11	2.91
20	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.85	3.68	3.32	3.02	2.83
21	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.77	3.60	3.24	2.95	2.75
22	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.70	3.54	3.18	2.88	2.69
23	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.64	3.47	3.12	2.82	2.62
24	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.59	3.42	3.06	2.77	2.57
25	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.54	3.37	3.01	2.72	2.52
26	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.49	3.33	2.97	2.67	2.47
27	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.45	3.28	2.93	2.63	2.43
28	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.41	3.25	2.89	2.59	2.39
29	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.38	3.21	2.86	2.56	2.36
30	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.34	3.18	2.82	2.52	2.32

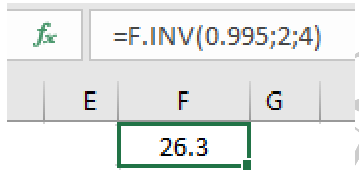
Табл. 11: Квантилі розподілу Фішера, $p = 0.995$

За необхідності квантилів розподілу Фішера можна обчислювати і іншими засобами. Наприклад:

- В електронних таблицях MS Excel значення квантилів розподілу Фішера обчислюємо за формулою

  =F.INV(p;n;m)

  Наприклад, вираз =F.INV(0.995;2;4) матиме наступне наближене значення:



- В пакеті MathCad значення квантилів розподілу Фішера обчислюємо за формулою

  qF(p, n, m)

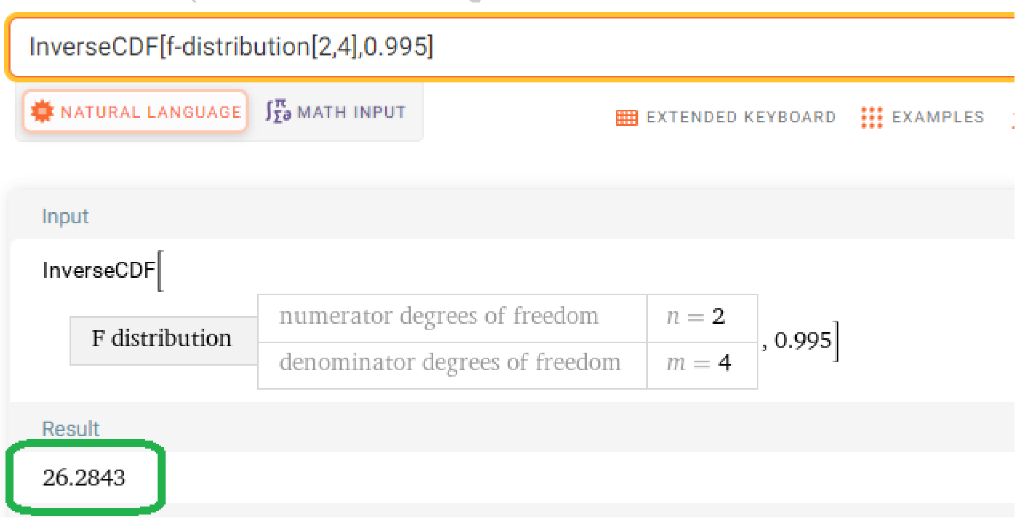
  Наприклад, вираз qF(0.995, 2, 4) матиме наступне наближене значення:

$$qF(0.995, 2, 4) = 26.284$$

- На онлайн-платформі www.wolframalpha.com значення квантилів розподілу Фішера обчислюють за формулою

  InverseCDF[f-distribution[n,m], p]

  Наприклад, вираз InverseCDF[f-distribution[2,4], 0.995] матиме наступне наближене значення:



22.7 Розподіл Пуассона

В таблиці наведено значення ймовірностей

$$\frac{\lambda^k e^{-\lambda}}{k!}$$

$k \setminus \lambda$	1	2	3	4	5	6	7	8
0	0.36788	0.13534	0.04979	0.01832	0.00674	0.00248	0.00091	0.00034
1	0.36788	0.27067	0.14936	0.07326	0.03369	0.01487	0.00638	0.00268
2	0.18394	0.27067	0.22404	0.14653	0.08422	0.04462	0.02234	0.01073
3	0.06131	0.18045	0.22404	0.19537	0.14037	0.08924	0.05213	0.02863
4	0.01533	0.09022	0.16803	0.19537	0.17547	0.13385	0.09123	0.05725
5	0.00307	0.03609	0.10082	0.15629	0.17547	0.16062	0.12772	0.09160
6	0.00051	0.01203	0.05041	0.10420	0.14622	0.16062	0.14900	0.12214
7	0.00007	0.00344	0.02160	0.05954	0.10444	0.13768	0.14900	0.13959
8	0.00001	0.00086	0.00810	0.02977	0.06528	0.10326	0.13038	0.13959
9	0.00000	0.00019	0.00270	0.01323	0.03627	0.06884	0.10140	0.12408
10	0.00000	0.00004	0.00081	0.00529	0.01813	0.04130	0.07098	0.09926
11	0.00000	0.00001	0.00022	0.00192	0.00824	0.02253	0.04517	0.07219
12	0.00000	0.00000	0.00006	0.00064	0.00343	0.01126	0.02635	0.04813
13	0.00000	0.00000	0.00001	0.00020	0.00132	0.00520	0.01419	0.02962
14	0.00000	0.00000	0.00000	0.00006	0.00047	0.00223	0.00709	0.01692
15	0.00000	0.00000	0.00000	0.00002	0.00016	0.00089	0.00331	0.00903

Табл. 12: Розподіл Пуассона

Відповіді

1.12 a) B, b) A, c) $\varnothing$, d) $A \cap C$. **1.15** $A^3 = A \times A \times A$, де $A = \{(a, b) \mid \ a, b = \overline{1, 6}\}$. **2.1** 6900 **2.2** 19958400, 3704778000 **2.3** а) 30, б) 120, в) 150. **2.4** 1140; 6840. **3.1** 2/3 для того, хто кидає першим, та 1/3 для того, хто кидає другим. **5.1** $\dfrac{2}{\pi}$. **6.1** Так, події A та B незалежні, бо нескладно показати, що $\mathrm{P}(AB) = \mathrm{P}(A)\mathrm{P}(B)$. **7.1** Ймовірності однакові.

Зміст

Предметний покажчик

Література

[1] Бобик О.І., Берегова Г.І., Копитко Б.І. Теорія ймовірностей і математична статистика – К.: Професіонал, 2007. – 560 с.

[2] Братійчук М.С., Чечельницький О.А. Математична статистика. – К.: ”Інтепроф”, 2009. – 244 с.

[3] Братійчук М.С., Чечельницький О.А. Лекції зі стохастики. Ймовірність. Статистика. Випадкові процеси.— Київ: електронна публікація на сайті факультету комп'ютерних наук та кібернетики, 2021.– 395 с. https://t.ly/TVfjI

[4] Василик О.І., Карташов М.В., Кнопова В.П. [та ін.] Методичні вказівки до лабораторних та самостійних робіт із дисципліни "Математична статистика". -– К. : ВПЦ "Київський університет", 2014. – 84 с.

[5] Голомозий В.В., Мішура Ю.С., Ральченко К.В. [та ін.] Збірник задач з теорії ймовірностей.– Київ:електронна публікація на сайті механіко-математичного факультету, 2023.– 214 с. https://t.ly/f6HVr

[6] Донченко В.С., Сидоров М.В.-С., Шарапов М.М. Теорія ймовірностей та математична статистика. – К.: Академія, 2009. – 286 с.

[7] Карташов М. В. Імовірність, процеси, статистика.– К. : ВПЦ Київський університет, 2006.– 494 с.

[8] Лебєдєв Є.О., Чечельницький О.А., Шарапов М.М., Братійчук М.С. Збірник задач з теорії ймовірностей. – К.: ВПЦ ”Київський університет”, 2006.

[9] Лебєдєв Є.О., Шарапов М.М. Курс лекцій з теорії ймовірностей. – К.: Норіта-плюс, 2007. – 168 с.

[10] Лебєдєв Є.О., Братійчук М.С., Чечельницький О.А., Шарапов М.М., Розора І.В. Збірник задач з прикладної статистики. – К.: ВПЦ "Київський університет 2010. – 116 с.

[11] Лебєдєв Є.О., Шарапов М.М. Вступ до теорії імовірностей. – К.: ВПЦ ”Київський університет”, 2010. – 151 с.

[12] Лебєдєв Є.О., Лівінська Г.В., Розора І.В., Шарапов М.М. Математична статистика: навч. посіб. – К.: ВПЦ ”Київський університет”, 2016. – 159 с.

[13] Майборода Р. Комп'ютерна статистика – професійний старт. – К.: ВПЦ ”Київський університет”, 2019.

[14] Теорія ймовірностей: Збірник задач; За ред.А.В. Скорохода. – К.: «Вища школа», 1976.

[15] Турчин В.М. Теорія ймовірностей і математична статистика. – Дніпропетровськ: Вид-во ДНУ, 2006.

[16] Dalgaard P. Introductory statistics with R. – Springer; 2nd edition, 2008. – 380 p. Dalgaard P. Introductory statistics with R. – Springer; 2nd edition, 2008. – 380 p.

[17] Dekking F.M., Kraaikamp C., Lopuhaä H.P., Meester L.E. A Modern Introduction to Probability and Statistics. Understanding Why and How. – Springer; 2005. – 486 p.

[18] Feller W. An Introduction to Probability Theory and Its Applications. – Wiley; 3rd edition, 1968. – Vol. 1. – 509 p.

[19] Feller W. An Introduction to Probability Theory and Its Applications. – Wiley; 2nd edition, 1971. – Vol. 2. – 669 p.

[20] Maindonald J.H. Using R for Data Analysis and Graphics. Introduction, Code and Commentary. – 96 p. https://t.ly/DqRRf

[21] Maindonald J.H., Braun W.J. Data Analysis and Graphics Using R – an Example-Based Approach. – Cambridge Univ. Press; 3rd edition, 2010. – 549 p.

[22] Mood A.M., Graybill F.A., Boes D.C. Introduction to the Theory of Statistics. – International Edition; 3rd Edition, 1974. – 480 p.

[23] Tijms H. Understanding Probability. Chance Rules in Everyday Life. – Cambridge Univ. Press, New York; 2007. – 442 p.

Навчальне видання

ШАРАПОВ Михайло Михайлович
РОЗОРА Ірина Василівна
ЛІВІНСЬКА Ганна Володимирівна
ПОНОМАРЬОВ Вадим Дмитрович

ЗБІРНИК ЗАДАЧ З ТЕОРІЇ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ

НАВЧАЛЬНИЙ ПОСІБНИК