МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРНІВЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ЮРІЯ ФЕДЬКОВИЧА

Опорний конспект лекцій з навчальної дисципліни «**Вища математика**»

для студентів спеціальностей:
«Телекомунікації та радіотехніка»,
«Видавництво та поліграфія»,
«Метрологія та інформаційно-вимірювальна техніка»,
«Трудове навчання та технології»,
«Професійна освіта (машинобудування)»,
навчально-наукового
Інституту фізико-технічних та комп'ютерних наук

Чернівці
ЧНУ ім. Юрія Федьковича
2021

УДК 51(075.8)
В558

Рекомендовано Вченою радою
навчально-наукового Інституту фізико-технічних та комп'ютерних наук
Чернівецького національного університету імені Юрія Федьковича
(*Протокол № 11 від 26.11.2021*)

Укладач: Івашко Віктор Вікторович, канд. фіз.-мат. наук, асистент.

В558 Івашко В.В. Конспект лекцій з навчальної дисципліни «Вища математика». Чернівці : Чернівецький національний університет імені Юрія Федьковича 2021. – 201 с.

Опорний конспект лекцій з навчальної дисципліни «Вища математика» для студентів всіх форм навчання спеціальностей: «Телекомунікації та радіотехніка», «Видавництво та поліграфія», «Метрологія та інформаційно-вимірювальна техніка», «Трудове навчання та технології», «Професійна освіта (машинобудування)».

УДК 51(075.8)

ЗМІСТ

ВСТУП

Описано курс лекцій зі «Вищої математики», який включає такі розділи, як «Теорія ймовірностей», «Математична статистика», «Функція комплексної змінної» та «Елементи дискретної математики».

Курс покликаний забезпечити оволодіння студентами теоретичних основ навчальної дисципліни та формування практичних умінь застосовувати одержані знання при розв'язуванні прикладних задач зі спеціальностей: «Телекомунікації та радіотехніка»; «Метрологія та інформаційно-вимірювальна техніка»; «Видавництво та поліграфія», «Трудове навчання та технології» та «Професійна освіта (машинобудування)» або у процесі вивчення природничих наук, що передбачає застосування сучасних методів вищої математики.

Особливу увагу приділено математичній статистиці та функції комплексної змінної з метою їх застосування для дослідження та розв’язування фізичних, математичних та статистичних задач.

У процесі вивчення курсу також закладаються вміння й навички щодо застосування понять і фактів математики у фізиці, інформатиці, метрології, статистиці та інших природничих науках.

Основна мета курсу – формування у студентів базових уявлень про сучасні методи аналізу масових, стаціонарних і однорідних випадкових процесів та подій, освоєння методів функцій комплексної змінної для дослідження природних явищ, законів та методів математичної логіки.

Матеріал викладено за принципом, де значна увага акцентується на розв'язку задач. Такий підхід забезпечує добре засвоєння теоретичного матеріалу і сприяє розвитку творчих здібностей у майбутніх фахівців.

РОЗДІЛ 1. ТЕОРІЯ ЙМОВІРНОСТЕЙ

Теорія ймовірностей зародилася в середині XVII ст. При спробі створити теорію азартних ігор. Перші праці, в яких згадувались основні поняття теорії ймовірностей, належать таким ученим свого часу, як Християн Гюйгенс, П'єр Ферма, Блез Паскаль та інші.

Вагомий внесок у розвиток теорії ймовірності належить знаменитому французькому вченому П'єр-Симону Лапласу. Він першим систематично виклав основи теорії ймовірностей, довів одну з форм центральної граничної теореми і показав практичне застосування теорії ймовірностей до конкретних задач, зокрема для аналізу помилок спостережень при виконанні вимірювань.

Нині теорію ймовірностей будують на так званій аксіоматичній основі, яку запропонував видатний радянський вчений Андрій Миколайович Колмогоров.

Теорія ймовірностей – це розділ вищої математики, який вивчає закономірності випадкових явищ, тобто випадкові події, величини, їх функції, властивості та операції над ними.

Предмет теорії ймовірностей полягає у вивченні математичних моделей реальних випадкових подій (явищ), які називають імовірнісними моделями. Такі моделі дають можливість зрозуміти математичну сутність реальних випадкових явищ та спрогнозувати перебіг досліджуваних випадкових подій. Наприклад, можна здійснити прогнозування зростання або падіння курсу валют, темпів виробництва, попиту на товари та послуги, результатів виборів, спортивних змагань та інші.

Лекція 1. Випадкові події. Означення ймовірності

Тема 1.1. Випадкові події. Класифікація подій

Основні поняття теорії ймовірностей випливають із її призначення як науки. Теорія ймовірностей займається, по суті, виявленням закономірностей в масових випадкових явищах (процесах) природи. Тому основний матеріал для роботи в теорії ймовірностей – це події.

Експериментом називають певна сукупність умов, що забезпечують спостереження за певним реальним явищем або *випадковою подією*.

Кожне проведення експерименту (забезпечення певних умов) називають *випробуванням*, а відповідний результат – його наслідком, або *елементарною подією*.

Сукупність усіх елементарних подій, які пов’язані з певним експериментом, позначають Ω і називають *множиною* (або *простором*) *елементарних подій*.

Кожну реальну випадкову подію, пов’язану з експериментом, ототожнюють з математичною моделлю, тобто з певною сукупністю

результатів цього експерименту. Такі моделі називають *подіями*. Події позначають великими латинськими літерами A, B, C, або A_1, A_2, …, A_n.

Подія A, яка під час виконання певної сукупності умов може відбутися або не відбутися, називається *випадковою*.

Подія, яка обов'язково відбувається під час кожного виконання комплексу умов, називається *вірогідною* (або *достовірною*). Якщо подія A найімовірніше, не може відбутися при виконанні певної сукупності умов, то вона називається *неможливою*. Вірогідну подію позначають через U, а неможливу – V.

Події A, B, C ... називаються несумісними, якщо поява однієї з них унеможливлює появу іншої події в одному і тому ж випробуванні. Якщо поява події не виключає появи іншої, в одному і тому ж випробуванні, то такі події називають сумісними.

Події називаються *протилежними*, якщо в умовах випробування вони, по суті, є єдиними його наслідками та несумісні. Подію, протилежну до події A, позначають через $\bar{A}$.

Події A_1, A_2, …, A_n *рівноможливі*, якщо жодна з них не має «переваги» над іншою.

Події A_1, A_2, …, A_n, попарно несумісні та рівноможливі, складають *повну групу подій*, з яких хоча б одна подія неминуче відбудеться, тобто сума цих подій – подія достовірна.

Тема 1.2. Класичне означення ймовірності

Якщо подія складається з m частинних випадків, що входять до повної групи з n попарно несумісних і рівноможливих випадків, тоді ймовірність такої події дорівнює

$$P(A)=\frac{m}{n}, \tag{1.1}$$

тобто ймовірність $P(A)$ події A дорівнює відношенню кількості результатів випробувань, які сприятливі події A, до загальної кількості всіх можливих результатів випробувань.

Імовірність характеризується такими властивостями:

1. Імовірність достовірної події дорівнює одиниці:

$$P(U)=1. \tag{1.2}$$

2. Імовірність неможливої події дорівнює нулю:

$$P(V)=0. \tag{1.3}$$

3. Імовірність випадкової події задовольняє нерівність:

$$0\le P(A)\le 1. \tag{1.4}$$

Задача 1.1. В ящику 7 однакових за розміром кульок: 1 червона, 2 сині, 4 білих. Знайдіть імовірність появи синьої кулі, якщо з ящика навмання беруть тільки одну кульку.

Розв'язок

Нехай подія A – це навмання взята синя куля.

З ящика можна взяти будь-яку кульку із семи, тобто всіх можливих випадків (наслідків) $n = 7$.

Появі синьої кулі сприяли лише 2 кулі, тому $m = 2$. За формулою (1) отримаємо:

$$P(A)=\frac{m}{n}=\frac{2}{7}\approx 0{,}29.$$

Задача 1.2. Знайти ймовірність того, що вибране випадковим способом двозначне число ділиться на 5.

Розв'язок

Цей експеримент полягає в тому, що випадково вибирається двозначне число. Наслідком такого випробування є одне з чисел від 10 – 99. Оскільки двозначних чисел 90, то $n = 90$, при цьому вибір кожного числа можна вважати рівноможливим.

Нехай подія A – це вибране двозначне число, яке ділиться на 5.

Загальна кількість наслідків експерименту $n = 90$. Кількість чисел, які можна поділити на 5, 18, тому $m = 18$.

$$P(A)=\frac{m}{n}=\frac{18}{90}=\frac{1}{35}\approx 0{,}03.$$

Тема 1.3. Елементи комбінаторики

Комбінаторика – розділ математики, який займається вивченням розміщення об'єктів відповідно до спеціальних правил та методів підрахунку кількості всіх можливих способів, за якими ці розміщення можуть бути виконанні.

Перестановки – це комбінації, які складаються з одних і тих самих n різних елементів, які відрізняються виключно порядком їх розкладу.

Кількість усіх перестановок з n елементів позначають P_n. Воно дорівнює добутку послідовності натуральних чисел від 1 до n (факторіал).

$$P_n=n!=1\cdot 2\cdot 3\cdot\ldots\cdot(n-1)\cdot n. \qquad (1.5)$$

Задача 1.3. Голова компанії планує відвідати 7 філій своєї фірми, які знаходяться в різних містах. Скільки існує різних маршрутів поїздок?

Розв'язок

Кількість різних маршрутів дорівнює кількістю перестановок із 7-ми елементів.

Тобто, згідно зі формулою (1.5), одержимо:

$$P_n=7!=1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7=5040.$$

Розміщення – це комбінації, які складені з n різних елементів по k елементах, які відрізняються між собою своїм порядком розташування або складом. Нехай задано множину, яка складається з n елементів. Усяка її упорядкована k елементна підмножина ($k < n$) називається розміщенням з n елементів по k.

Кількість розміщень з n елементів по k знаходять за допомогою формули:

$$A_n^k = \frac{n!}{(n-k)!}. \tag{1.6}$$

Задача 1.4. Керівництво фірми, яке складається з голови, заступника, та головного бухгалтера, обирають за конкурсом з поміж 10 претендентів. Знайти кількість варіантів вибору керівництва компанії.

Розв'язок

Оскільки порядок вибору членів керівництва є важливим, то кількість варіантів вибору буде розміщення з 10 по 3.

Згідно з формулою (1.6), отримаємо:

$$A_{10}^3 = \frac{10!}{(10-3)!} = 8 \cdot 9 \cdot 10 = 720.$$

Сполучення – це комбінації, які складені зі n різних елементів по k елементах, які відрізняються між собою хоча б одним елементом. Нехай задано множину, яка складається з n елементів. Усяка її k елементна підмножина ($k < n$) називається сполученням з n елементів по k.

Кількість сполучень з n елементів по k знаходять за допомоги формули:

$$C_n^k = \frac{n!}{k!(n-k)!}. \tag{1.7}$$

Задача 1.5. Приватне підприємство має ліцензію на проведення десяти видів діяльності. На початку роботи підприємство планує займатися тільки п'ятьма видами діяльності. Скільки існує способів вибору цих п'яти видів діяльності?

Розв'язок

Оскільки порядок вибору видів діяльності неважливий, то кількість способів вибору п'яти видів діяльності буде дорівнювати кількості комбінацій з 10 елементів по 5.

Згідно з формулою (1.7), одержимо:

$$C_{10}^5 = \frac{10!}{5!(10-5)!} = \frac{6 \cdot 7 \cdot 8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 252.$$

Задача 1.6. Чотирнадцять варіантів контрольної роботи у формі окремих екзаменаційних карток випадково розподіляються серед 12 здобувачів освіти, які сидять в одному ряду. Кожний студент отримує одну картку. Знайдіть імовірність того, що варіанти 1 і 2 не будуть використані взагалі.

Розв'язок

Маємо справу з експериментом по розподілу 14 карток серед 12 студентів. Результатами є впорядковані (за студентами) набори розданих 12 з 14 варіантів контрольних робіт.

У цьому разі елементарні події відрізняються одна від одної не лише номерами варіантів, які розподіляються серед студентів, а й порядком розподілу.

Такі елементарні події є розміщеннями, а кількість усіх елементарних подій (розміщень) знаходимо за формулою (1.6):

$$n = A_{14}^{12} = \frac{14!}{(14-12)!}.$$

Вважаємо, що всі елементарні події рівноможливі.

Нехай подія A – варіанти 1 і 2 залишаться нерозподіленими. Тоді інші 12 карток розподіляться серед 12 здобувачів освіти.

Такі розподіли, по суті, є перестановками, а їх кількість розраховується за формулою (1.5):

$$m = P_{12} = 12!,$$

де m – кількість елементарних подій, які сприяють події A.

Згідно з формулою (1.1), одержимо:

$$P(A) = \frac{m}{n} = \frac{12!}{A_{14}^{12}} = \frac{12!\cdot(14-12)!}{14!} = \frac{1\cdot 2}{13\cdot 14} = \frac{1}{91} \approx 0{,}01.$$

Задача 1.7. Група складається з 10 хлопців і 5 дівчат, серед яких вибирають двох учнів, для участі у конференції. Яка ймовірність того, що:

А) виберуть двох хлопців;

Б) виберуть хлопця та дівчину?

Розв'язок

У групі 15 осіб, з них 10 хлопців і 5 дівчат. Вибір двох людей із 15 визначається так:

$$C_{15}^{2} = \frac{15!}{2!(15-2)!} = \frac{14\cdot 15}{1\cdot 2} = 105.$$

Загальна кількість випадків $n = 105$.

А) нехай подія A – вибрали двох хлопців.

Кількість випадків, які сприяють події A, визначається кількістю (числом) вибору 2 хлопців із 10:

$$m = C_{10}^{2} = \frac{10!}{2!(10-2)!} = \frac{9\cdot 10}{1\cdot 2} = 45.$$

Тоді, за формулою (1.1), одержимо:

$$P(A) = \frac{m}{n} = \frac{45}{105} = 0{,}43.$$

Б) нехай подія B – вибрали хлопця та дівчину.

Кількість випадків, що сприяють події B, визначається кількістю вибору 1 хлопця із 10 та 1 дівчини з 5.

Це можна зробити за формулою:

$$m = C_{10}^{1}\cdot C_{5}^{1} = \frac{10!}{1!(10-1)!}\cdot\frac{5!}{1!(5-1)!} = 10\cdot 5 = 50.$$

Отже,

$$P(B) = \frac{m}{n} = \frac{50}{105} = 0{,}48.$$

Задача 1.8. У партії з N деталей – M непридатних. Навмання виймають n деталей. Знайдіть імовірність того, що з n деталей буде m непридатних.

Розв'язок

Елементарним наслідком є вибір будь-яких деталей із загальної кількості *N*. Кількість усіх таких наслідків дорівнює кількості комбінацій з *N* по *n*, тобто C_N^n.

Подія *A* – це виймання *n* деталей, з яких *m* непридатних. Наслідком, який сприяє події *A*, є поява групи з *n* деталей, в яких *n*-*m* якісних деталей і *m* непридатних.

Кількість таких груп дорівнює

$$C_{N-M}^{n-m} \cdot C_M^m,$$

бо групу з *m* бракованих деталей можна утворити

$$C_M^m,$$

а групу *n*-*m* якісних деталей

$$C_{N-M}^{n-m}.$$

Водночас будь-яка група придатних (якісних) деталей може комбінуватись з будь-якою групою непридатних деталей.

Шукана ймовірність події діорівнює відношенню кількості наслідків, які сприяють цій події, до кількості всіх можливих елементарних наслідків:

$$P(A) = \frac{C_{N-M}^{n-m} \cdot C_M^m}{C_N^n}.$$

Тема 1.4. Статичне визначення ймовірності

Використання класичного означення до задач природничого або наукового характеру не завжди можливо з різного роду причин.

Оскільки часто неможливо подати результат експерименту як сукупність подій, які можна вважати рівноможливими.

Наприклад, із міркувань симетрії, на яких, по суті, базуються міркування про рівноймовірність подій, розрахувати ймовірність того, що новонароджена дитина була хлопчиком, неможливо.

З цієї причини поряд із класичним означенням імовірності використовуються також *статистичне означенням ймовірності*, відповідно до кого ймовірністю події вважають її відносну частоту.

Наприклад, якщо через μ позначити кількість появ події в *n* – незалежних випробуваннях, тоді відношення кількості появ події до загального числа випробувань, що відбулись, називають *відносною частотою появи події*, тобто

$$W = \frac{\mu}{n}. \quad (1.8)$$

Класичне означення ймовірності не ставить вимогу, щоби випробування відбувалось в дійсності, проте означення відносної частоти вимагає реального проведення випробувань. Тобто ймовірність визначають до випробування, а відносну частоту після.

Задача 1.9. Відділ технічного контролю знайшов у партії з 90 деталей 3 непридатні деталі. Знайдіть відносну частоту непридатних деталей.

Розв'язок

Нехай A – подія появи непридатної деталі. Загальна кількість деталей $n = 90$. Непридатних деталей є $\mu = 3$.

Тоді, згідно з формулою (1.8), одержимо

$$W(A)=\frac{\mu}{n}=\frac{3}{90}\approx 0{,}03.$$

Якщо відносна частота була встановлена дослідним шляхом, то отримане число можна вважати наближеним значенням шуканої ймовірності.

Наприклад, за даними статистики, відносна частота народження дівчат на 1000 дітей характеризується такими числами за кожен місяць:

Таблиця 1.1

Місяць	1	2	3	4	5	6	7	8	9	10	11	12
W_i	0,486	0,489	0,49	0,471	0,478	0,482	0,462	0,484	0,485	0,491	0,482	0,473

Тема 1.5. Геометричне означення ймовірності

Класичне означення ймовірності використовується лише для експериментів з обмеженою кількістю рівномірних елементарних подій, тобто коли простір елементарних подій обмежений. Для математичного опису експерименту з нескінченною кількістю рівноможливих результатів застосовують так зване геометричне означення ймовірності.

Імовірність випадкової події A дорівнює

$$P(A)=\frac{m(A)}{m(\Omega)}, \tag{1.9}$$

де $m(A)$ – довжина або площина, або об'єм підмножини A; $m(\Omega)$ – довжина, або площина, або об'єм області Ω.

Це означення ймовірності має, по суті, просту геометричну інтерпретацію: це частина, яку складає довжина, площина або об'єм множини A від довжини, площі або об'єму всієї області Ω. Випадкові події в цьому випадку реалізуються як певні геометричні об'єкти.

Задача 1.10. На відрізок OA довжиною L осі Ox навмання поставлена точка B (рис. 1.1). Знайти ймовірність того, що менший із відрізків, що утворились (OB та BA), має довжину, більшу за L/3. Припускається, що ймовірність попадання точки пропорційна довжині відрізку й не залежить від його розміщення на осі.

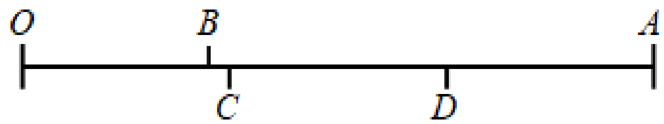


Рис. 1.1

Розв'язок

Розіб'ємо відрізок *OA* точками *C* та *D* на 3 рівні частини. Умова задачі буде виконана, якщо точка *B* попаде на відрізок *CD* довжиною *L*/3.

Шукана ймовірність:

$$P=\frac{L/3}{L}=\frac{1}{3}.$$

Задача 1.11. На площі накреслено два кола з радіусами 5 та 10 см відповідно (рис. 1.2). Знайдіть імовірність того, що точка, яка кинута навмання у більше коло, попаде в кільце, утворене побудованими колами. Припускається, що ймовірність попадання точки в кільце пропорційна його площі та не залежить від її розміщення відносно великого кола.

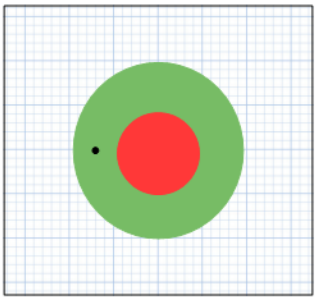


Рис.1.2

Розв'язок

Площа кільця (фігури *g*):

$$S_g=\pi\,10^2-\pi\,5^2=\pi\left(10^2-5^{12}\right)=75\pi\,.$$

Площа великого кола (фігури *G*):

$$S_G=\pi\,10^2=100\pi\,.$$

Шукана ймовірність:

$$P=\frac{S_g}{S_G}=\frac{75\pi}{100\pi}=0{,}75.$$

Задача 1.12. У сигналізатор надходять сигнали від двох датчиків руху, причому надходження кожного з них рівноймовірне у будь-який момент проміжку часу тривалістю *T*. Моменти надходження сигналів не залежать один від одного. Сигналізатор спрацьовує, якщо різниця між моментами надходження сигналів менше деякого часу *t* (*t* < *T*). Знайти ймовірність того, що сигналізатор спрацює за період часу *T*, якщо кожен з датчиків надішле по одному сигналу.

Розв'язок

Нехай *x* та *y* – моменти надходження сигналів від першого та другого датчиків відповідно.

Згідно з умовою, повинні виконуватися дві нерівності:

$$0\le x\le T\,,$$
$$0\le y\le T\,.$$

Розглянемо прямокутну систему координат *xOy*. В цій системі зазначені нерівності задовольняють координати будь-якої точки певного квадрата *OTAT*. Отже, квадрат *OTAT* можна розглядати як геометричну фігуру *G*, координати точок якої являють собою всі можливі значення моментів надходження сигналів.

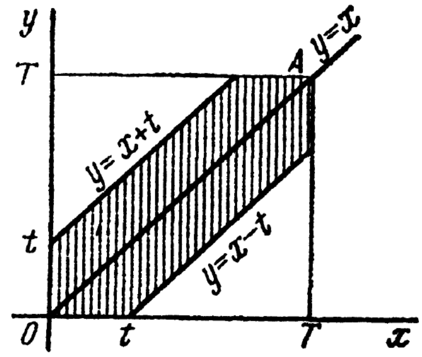


Рис. 1.3

Сигналізатор спрацьовує, якщо різниця між моментами надходження сигналів менше t, тобто якщо $y - x < t$ при $y > x$ та $x - y < t$ при $x > y$, або це то саме, що

$$\begin{gathered} y<x+t \text{ при } y>x, \\ y>x-t \text{ при } y<x. \end{gathered} \tag{1.10}$$

Перша нерівність (1.10) справедлива для координат тих точок фігури G, які лежать вище прямої $y = x$ і нижче $y = x + t$; друга нерівність (1.10) справедлива для координат тих точок фігури G, які лежать ниже $y = x$ і вище $y = x - t$.

З рис. 1.3 видно, що всі точки, координати яких задовольняють нерівності (1.10), належать заштрихованій області. Отже, цю область (шестикутник) можна розглядати як фігуру g, координати точок якої сприяють спрацюванню сигналізатора у відповідні моменти часу (x та y).

$$P=\frac{S_g}{S_G}=\frac{T^2-2\cdot(T-t)^2/2}{T^2}=\frac{T^2-(T-t)^2}{T^2}=\frac{T^2-(T^2-2Tt+t^2)}{T^2}=\frac{T^2-T^2+2Tt-t^2}{T^2}=\frac{t(2T-t)}{T^2}.$$

Лекція 2. Основні теореми теорії ймовірностей

Тема 2.1. Умовна ймовірність. Залежні та незалежні події

Якщо при розрахунку ймовірності події A не вказуються ніякі умови, крім комплексу (сукупності) умов, то таку ймовірність називають *безумовною* її позначають $P(A)$.

Імовірність настання події B, розрахована в припущені, що подія A уже відбулась, називається *умовною ймовірністю* події B при умові A і позначається:

$$P(B/A) \text{ або } P_A(B).$$

Задача 1.13. В ящику міститься 5 червоних і 7 білих кульок. Із ящика два рази виймають по одній кульці. Знайти ймовірність того, що за другою спробою (випробуванням) буде вийнята біла кулька (подія B), якщо першою була вийнята також біла кулька (подія A).

Розв'язок

В ящику є 12 (5 + 7 = 12) кульок. Після першого випробування в ньому залишилось 11 кульок, із них 6 білих. Відповідно, шукана ймовірність, на основі класичного визначення, дорівнює:

$$P_A(B)=\frac{6}{11}\approx 0{,}54.$$

Цей самий результат можна одержати іншим способом за допомоги формули:

$$P_A(B)=\frac{P(AB)}{P(A)},\quad \text{за умови} \quad P(A)>0.$$

Імовірність появи білої кульки при першій спробі:

$$P(A)=\frac{7}{12}.$$

$P(AB)$ – це ймовірність того, що при першій та другій спробі з'явиться біла кулька. Загальна кількість результатів спільної появи обох кульок (неважливо якого кольору) дорівнює кількості розміщень:

$$A_{12}^{2}=\frac{12!}{(12-2)!}=11\cdot 12=132.$$

З цього числа події AB сприяють 6 * 7 випадків, тобто

$$P(AB)=\frac{6\cdot 7}{132},$$

$$P_A(B)=\frac{P(AB)}{P(A)}=\frac{6\cdot 7}{132}\cdot\frac{12}{7}=\frac{6\cdot 12}{132}=\frac{12}{22}=\frac{6}{11}\approx 0{,}54.$$

Дві події A і B називаються *незалежними* тоді та тільки тоді, якщо ймовірність появи однієї із них не залежить від появи або непояви іншої події.

Події A_1, A_2, … A_n називаються *незалежними в сукупності* тоді і тільки тоді, якщо ймовірність кожної з них не змінюється через настання або ненастання інших подій окремо або в комбінації (будь-якій).

Тема 2.2. Теорема додавання ймовірностей несумісних подій

Сумою двох подій ($A + B$) називають таку подію (C), яка полягає у появі події A або події B, або обох цих подій. Наприклад, зроблено два постріли, A – подія, що означає попадання при першому пострілі, а B – при другому, тоді $A + B$ означає влучання при першому або другому, або в обох випадках.

Зокрема, якщо дві події A і B несумісні, то $A + B$ – це подія, яка полягає в появі однієї із цих подій, тобто або A, або B.

Сумою $\Sigma_i A_i$ подій A_1, A_2, … A_n називають таку подію, поява якої рівносильна появі принаймні однієї з подій A_i, $i = 1, 2, \dots , n$.

Теорема: ймовірність появи однієї з кількох несумісних подій, не важливо якої, дірівнює сумі ймовірностей цих подій.

$$P(A+B)=P(A)+P(B), \tag{1.11}$$

$$P\left(\sum_{i=1}^{n} A_i\right)=P(A_1+A_2+\ldots+A_n)=P(A_1)+P(A_2)+\ldots P(A_n). \tag{1.12}$$

Доведення. Введемо позначення: n – це загальна кількість можливих елементарних наслідків випробування; m_1 – кількість наслідків, які сприяють

події A; m_2 – кількість наслідків, які сприяють події B. Кількість елементарних наслідків, які сприяють події A або події B, дорівнює сумі: $m_1 + m_2$. Отже,

$$P(A+B)=\frac{m_1+m_2}{n}=\frac{m_1}{n}+\frac{m_2}{n},$$

Узявши до уваги, що $m_1/n = P(A)$ і $m_2/n = P(B)$, отримаємо:

$$P(A+B)=P(A)+P(B),$$

Наслідок. Імовірність появи однієї із декількох попарно несумісних подій, неважливо якої, дорівнює сумі ймовірностей цих подій.

Задача 1.14. В ящику 30 кульок: 10 червоних, 5 синіх і 15 білих. Знайти ймовірність появи кольорової кульки.

Розв'язок

Імовірність появи червоної та синьої кульок:

$$P(A)=10/30=1/3, \qquad P(B)=5/30=1/6.$$

Події A та B несумісні, тобто поява кулі одного кольору виключає появу кулі іншого кольору, відповідно:

$$P(A+B)=P(A)+P(B)=1/3+1/6=1/2.$$

Задача 1.15. В коропці знаходиться 6 білих і 4 чорних кульок. З неї виймають навмання 2 кульки. Яка ймовірність того, що вони одного кольору?

Розв'язок

Нехай подія A – це поява двох білих кульок, а подія B – це поява двох чорних кульок. Події A та B несумісні. Кількість всіх єдино можливих, рівноможливих та несумісних випадків дорівнює кількості пар, які можна утворити з десяти різних кульок:

$$n=C_{10}^2=\frac{10!}{2!(10-2)!}=\frac{9\cdot 10}{2}=45$$

n – кількість загальних випадків. Кількість випадків, сприятливих для появи двох білих кульок, дорівнює:

$$m_1=C_6^2=15\text{, а для двох чорних } m_2=C_4^2=6.$$

Тоді, згідно з формулою (1.12), отримаємо:

$$P(A+B)=P(A)+P(B)=\frac{15}{45}+\frac{6}{45}=\frac{7}{15}\approx 0{,}47.$$

Тема 2.3. Теорема множення залежних і незалежних подій

Добутком двох подій $A \times B$ називають таку подію C, яка полягає в одночасній появі обох цих подій.

Добутком подій $A_1, A_2, \ldots A_n$ називають таку подію, поява якої рівносильна появі кожної з подій (A_i , $i = 1, 2, \ldots , n$).

Теорема: якщо події A і B незалежні одна від одної, тоді ймовірність добутку цих подій дорівнює добутку ймовірностей цих подій, тобто

$$P(A\cdot B)=P(A)\cdot P(B). \qquad (1.13)$$

Імовірність появи кількох подій, незалежних у сукупності, обчислюється згідно з формулою

$$P\left(\prod_{i=1}^{n} A_i\right)=P(A_1\cdot A_2\cdot\ldots\cdot A_n)=P(A_1)\cdot P(A_2)\cdot\ldots P(A_n). \qquad (1.14)$$

Задача 1.16. У трьох ящиках знаходиться по 10 деталей. У першому ящику – 8, в другому – 7 і в третьому – 9 стандартних деталей. Із кожного ящика навмання виймають по одній деталі. Знайдіть імовірність того, що всі три вийняті деталі будуть стандартними.

Розв'язок

Імовірність того, що з першого ящика вийнята стандартна деталь (подія *A*), дорівнює:

$$P(A)=\frac{8}{10}=0{,}8.$$

Імовірність того, що з другого ящика вийнята стандартна деталь (подія *B*), дорівнює:

$$P(B)=\frac{7}{10}=0{,}7.$$

Імовірність того, що із третього ящика вийнята стандартна деталь (подія *C*), дорівнює:

$$P(C)=\frac{9}{10}=0{,}9.$$

Оскільки *A*, *B*, *C* незалежні, то шукана ймовірність (згідно з теоремою добутку – формула (1.14)) визначається через формулу

$$P(A\cdot B\cdot C)=P(A)\cdot P(B)\cdot P(C)=0{,}8\cdot 0{,}7\cdot 0{,}9=0{,}504.$$

Якщо події *A* і *B* залежні між собою, причому відомі як $P(A)$, так і $P_A(B)$, то ймовірність добутку цих подій шукається за допомогою нижченаведеної теореми.

Теорема: ймовірність добутку двох залежних подій дорівнює добутку ймовірності однієї із цих подій на умовну ймовірність іншої, розраховану за умови, що перша подія вже відбулась.

$$P(A\cdot B)=P(A)\cdot P_A(B). \qquad (1.15)$$

Нехай, ми маємо три події *A*, *B*, *C*, які залежні між собою, тоді теорема добутку ймовірностей матиме такий вигляд:

$$P(A\cdot B\cdot C)=P(A)\cdot P_A(B)\cdot P_{AB}(C). \qquad (1.16)$$

Оскільки A * B = B * A, то і для добутку ймовірностей має місце закон комутативності:

$$P(A\cdot B)=P(A)\cdot P_A(B)=P(B)\cdot P_B(A). \qquad (1.17)$$

Задача 1.17. Студент знає 20 із 25 питань програми. Знайдіть ймовірність того, що студент знає три питання, які будуть в екзаменаційному білеті.

Розв'язок

Імовірність того, що студент знає відповідь на перше питання (подія *A*), дорівнює:

$$P(A)=\frac{20}{25}.$$

Імовірність того, що студент знає відповідь на друге питання (подія *B*), така

$$P(B)=\frac{19}{24}$$

(оскільки події *A*, *B* залежні між собою).

Імовірність того, студент знає відповідь на третє питання (подія *C*), дорівнює:

$$P(C)=\frac{18}{23}$$

(оскільки події *A*, *B*, *C* залежні між собою). Згідно з формулою (1.16), маємо:

$$P(A\cdot B\cdot C)=\frac{20}{25}\cdot\frac{19}{24}\cdot\frac{18}{23}\approx 0{,}5.$$

Тема 2.4. Теорема додавання ймовірностей сумісних подій

Суму двох сумісних подій можна подати у вигляді:

$$A+B=A\cdot\bar{B}+B\cdot\bar{A}+A\cdot B$$

Геометрично (у формі кіл Ейлера) суму двох сумісних подій *A* і *B* продемонстровано на нижченаведеному рисунку

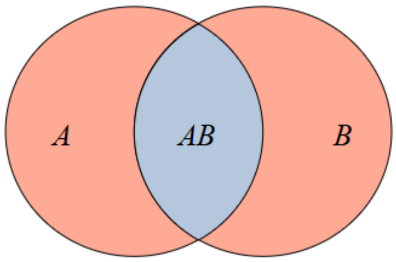


Рис. 1.4

Теорема: ймовірність появи хоча б однієї із двох сумісних подій дорівнює сумі ймовірностей цих подій без імовірності добутку цих подій:

$$P(A+B)=P(A)+P(B)-P(A\cdot B). \qquad (1.18)$$

Зауваження! При використанні формули (1.8) слід звернути увагу на те, що події *A* і *B* можуть бути як залежними, так і незалежними.

Для *сумісних і незалежних подій* теорема додавання запишеться формулою

$$P(A+B)=P(A)+P(B)-P(A)\cdot P(B). \qquad (1.19)$$

Для *сумісних і залежних подій* теорема додавання запишеться формулою

$$P(A+B)=P(A)+P(B)-P(A)\cdot P_A(B). \qquad (1.20)$$

Задача 1.18. Для виконання завдання власник фірми звернувся до двох виконавців. Імовірність того, що перший виконавець надасть згоду дорівнює 0,8, а другий – 0,9. Знайдіть імовірність того, що замовлення буде прийнято принаймні одним з виконавців.

Розв'язок

A – завдання прийнято для виконання;
A_1 – перший виконавець дав згоду виконати завдання;
A_2 – другий виконавець дав згоду виконати завдання.

Тоді $A = A_1 + A_2$. Оскільки згода на виконання замовлення першим виконавцем не виключає згоди на виконання замовлення іншим – події сумісні. Тоді

$$P(A)=P(A_1)+P(A_2)-P(A_1\cdot A_2).$$

Події A_1 та A_2 незалежні, оскільки ймовірність згоди кожного з виконавців не залежить від того, чи дав згоду інший виконавець. Тоді

$$P(A_1\cdot A_2)=P(A_1)\cdot P(A_2).$$

Отже,

$$P(A)=0{,}8+0{,}9-0{,}8\cdot 0{,}9=0{,}98.$$

Тема 2.5. Протилежні події. Ймовірність протилежної події

Протилежними називають дві єдино можливі події, які утворюють повну групу подій. Подію, протилежну до події A (ненастання події A), позначають через $\bar{A}$.

Наприклад: влучання і промах при пострілі в мішень – протилежні події. Скласти іспит та не скласти – протилежні події. Якщо подія A – влучання в мішень, то $\bar{A}$ – промах. Якщо A – подія складання іспиту, то $\bar{A}$ – подія його не складання.

Теорема: ймовірність суми протилежних подій дорівнює одиниці, тобто:

$$P(A)+P(\bar{A})=1. \qquad (1.21)$$

Задача 1.19. Імовірність того, що банк дасть кредит, дорівнює 0,7. Знайдіть імовірність того, що банк відмовить у кредитуванні.

Розв'язок

Події A – кредит буде надано та $\bar{A}$ – кредиті не надано – протилежні. Використовуючи формулу (1.21), ймовірність буде такою:

$$P(\bar{A})=1-P(A)=1-0{,}7=0{,}3.$$

Задача 1.20. В коробці в довільному порядку розкладено 20 деталей, причому 5 з них стандартні. Навмання беруть три деталі. Знайдіть ймовірність того, що принаймні одна з них буде стандартною.

Розв'язок

Задачу можна розв'язати двома способами.

Перший спосіб

Принаймні одна з взятих деталей буде стандартною, якщо відбудеться одна з трьох несумісних подій:
A_1 – одна деталь стандартна і дві нестандартні,
A_2 – дві деталь стандартна і одна нестандартні,
A_3 – три деталі стандартні.

Подію A можна записати у вигляді суми цих трьох подій:

$$A=A_1+A_2+A_3.$$

За теоремою додавання – формула (1.12) – маємо:

$$P(A)=P(A_1)+P(A_2)+P(A_3).$$

Знайдемо ймовірність кожної з цих подій (на основі задачі 1.8 з попередньої лекції):

$$P(A_1)=\frac{C_5^1\cdot C_{15}^2}{C_{20}^3}=\frac{5}{1}\cdot\frac{14\cdot 15}{1\cdot 2}\cdot\frac{1\cdot 2\cdot 3}{18\cdot 19\cdot 20}=\frac{35}{76};$$

$$P(A_2)=\frac{C_5^2\cdot C_{15}^1}{C_{20}^3}=\frac{4\cdot 5}{1\cdot 2}\cdot\frac{15}{1}\cdot\frac{1\cdot 2\cdot 3}{18\cdot 19\cdot 20}=\frac{5}{38};$$

$$P(A_3)=\frac{C_5^3}{C_{20}^3}=\frac{3\cdot 4\cdot 5}{1\cdot 2\cdot 3}\cdot\frac{1\cdot 2\cdot 3}{18\cdot 19\cdot 20}=\frac{1}{144}.$$

Додавши знайдені величини, отримаємо:

$$P(A)=\frac{35}{76}+\frac{5}{38}+\frac{1}{144}=\frac{137}{228}\approx 0{,}60.$$

Другий спосіб

Події A – принаймні одна з взятих деталей буде стандартною і $\bar{A}$ – жодна зі взятих деталей не стандартна є протилежними, тому, згідно з формулою

$$P(A)+P(\bar{A})=1, \qquad P(A)=1-P(\bar{A}).$$

Імовірність появи події $\bar{A}$ дорівнює:

$$P(\bar{A})=\frac{C_{15}^3}{C_{20}^3}=\frac{13\cdot 14\cdot 15}{1\cdot 2\cdot 3}\cdot\frac{1\cdot 2\cdot 3}{18\cdot 19\cdot 20}=\frac{91}{228}.$$

Отже, шукана ймовірність:

$$P(A)=1-P(\bar{A})=1-\frac{91}{228}=\frac{137}{228}\approx 0{,}60.$$

Тема 2.6. Ймовірність появи хоча б однієї події

Нехай в результаті випробовування може виникнути n подій, причому відомі їх ймовірності, тоді ймовірність того, що виникне хоча б одна із цих подій, шукається за наступною теоремою.

Теорема: ймовірність появи хоча б однієї з подій A_1, A_2, … A_n, незалежних у сукупності, дорівнює різниці між одиницею і добутком ймовірностей протилежних подій $\bar{A}_1$, $\bar{A}_2$, ... $\bar{A}_n$.

$$P(A)=1-q_1\cdot q_2\cdot\ldots\cdot q_n, \tag{1.22}$$

де $q_i = 1 - p_i$ – ймовірності протилежних подій.

Задача 1.21. Ймовірність влучання в мішень із трьох рушниць, дорівнює: $p_1 = 0{,}7$, $p_2 = 0{,}8$ та $p_3 = 0{,}9$. Знайдіть імовірність влучення при одному залпі із трьох рушниць.

Розв'язок

Події A_1, A_2, A_3 – попадання в ціль із трьох рушниць, які незалежні сукупно. Ймовірності промахів такі:

$$q_1=1-p_1=1-0{,}7=0{,}3=P(\bar{A}_1);$$

$$q_2=1-p_2=1-0{,}8=0{,}2=P(\bar{A}_2);$$

$$q_3 = 1 - p_1 = 1 - 0{,}9 = 0{,}1 = P(\bar{A}_3).$$

Шукана ймовірність за формулою (1.22) буде дорівнювати:

$$P(A) = 1 - 0{,}3 \cdot 0{,}2 \cdot 0{,}1 = 0{,}994.$$

Якщо події A_1, A_2, … A_n мають однакову ймовірність, яка дорівнює p, тобто $P(A_1) = P(A_2) = ... = P(A_n) = p$, то ймовірність появи хоч би однієї із цих подій визначається за формулою

$$P(A) = 1 - q^n. \tag{1.23}$$

Задача 1.22. Підписано 4 однотипних договори страхування терміном на 1 рік. Імовірність того, що за будь-яким із договорів протягом року надійде запит на відшкодування, дорівнює 0,1. Припускаючи, що надходження запитів незалежні між собою, знайдіть імовірність того, що:

А) запити надійдуть за чотирма договорами;

Б) запити не надійдуть за жодним договором;

В) запит надійде принаймні за одним договором.

Розв'язок

Нехай A_i, $i = 1, 2, 3, 4$ – це подія, яка полягає в тому, що запит надійде за i-тим договором. Тоді $\bar{A}_i$ – це протилежна подія до події A_i. Тоді:

$$P(A_i) = p = 0{,}1; \qquad P(\bar{A}_i) = q = 1 - p = 1 - 0{,}1 = 0{,}9.$$

А) оскільки події незалежні, то, згідно з формулою для незалежних подій,

$$P(A_1 \cdot A_2 \cdot A_3 \cdot A_4) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4) = 0{,}1^4 = 0{,}0001;$$

Б) $$P(\bar{A}_1 \cdot \bar{A}_2 \cdot \bar{A}_3 \cdot \bar{A}_4) = P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot P(\bar{A}_3) \cdot P(\bar{A}_4) = 0{,}9^4 = 0{,}6561;$$

В) нехай подія A – це запит, який надійде принаймні за одним зі договором, є протилежною подією до події $\bar{A}$ – запит не надійде за жодним договором. Використовуючи формулу (1.23), одержимо: $P(A) = 1 - 0{,}9^4 = 0{,}3439.$

Тема 2.7. Формула повної ймовірності

Нехай подія A може відбутися з однією і лише однією із n-попарно несумісних подій H_1, H_2, … H_n, які утворюють повну групу. Тобто

$$A = A \cdot H_1 + A \cdot H_2 + ... + A \cdot H_n.$$

Потрібно знайти ймовірність події A. На основі теореми додавання несумісних подій маємо:

$$P(A) = P(A \cdot H_1) + P(A \cdot H_2) + ... + P(A \cdot H_n).$$

Далі, використавши теорему множення залежних подій, отримаємо:

$$P(A) = P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A) + ... + P(H_n) \cdot P_{H_n}(A),$$

або

$$P(A) = \sum_{i=1}^{n} P(H_i) P_{H_i}(A). \tag{1.24}$$

Формулу (1.24) називають ***формулою повної ймовірності.***

$$P(H) = P(H_1) + P(H_2) + ... + P(H_n) = 1. \tag{1.25}$$

Теорема: ймовірність події A, яка може наступити тільки при умовній ймовірності однієї з декількох подій H_1, H_2, … H_n , які утворюють повну групу,

дорівнює сумі добутків ймовірності кожної з цих подій на умовну ймовірність події A.

Задача 1.23. Маємо дві коробки деталей. Імовірність того, що в першій коробці деталь стандартна, дорівнює 0,8, а в другій – 0,9. Знайти ймовірність того, що взята навмання із будь-якої коробки деталь буде стандартна.

Розв'язок

Нехай подія A – вийнята стандартна деталь. H_1 – деталь може бути вийнята з першої коробки, H_2 – деталь може бути вийнята з другої коробки, тобто:

$$A = A \cdot H_1 + A \cdot H_2.$$

На основі формули повної імовірності (1.24), маємо:

$$P(A) = P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A).$$

Імовірність вибору коробки (формула 1.25):

$$P(H_1) = \frac{1}{2}, \qquad P(H_2) = \frac{1}{2}.$$

Імовірність стандартної деталі з відповідної коробки:

$$P_{H_1}(A) = 0{,}8, \qquad P_{H_2}(A) = 0{,}9.$$

Підставивши значення, одержимо:

$$P(A) = 0{,}5 \cdot 0{,}8 + 0{,}5 \cdot 0{,}9 = 0{,}85.$$

Тема 2.8. Ймовірність гіпотез. Формула Байєса

Нехай A – це подія, яка може відбутися з однією і лише з однією із n попарно несумісних подій H_1, H_2, … H_n, тобто:

$$A = A \cdot H_1 + A \cdot H_2 + \ldots + A \cdot H_n.$$

Якщо наперед невідомо, яка з цих подій відбудеться, то події H_1, H_2, … H_n прийнято називати *гіпотезами*.

Нехай подія A відбулася. Потрібно знайти ймовірність події H_i, з якою відбулася подія A, тобто $P_A(H_i)$. На основі теореми добутку ймовірностей отримаємо:

$$P(A H_i) = P(A) \cdot P_A(H_i) = P(H_i) \cdot P_{H_i}(A),$$

звідки

$$P_A(H_i) = \frac{P(H_i) \cdot P_{H_i}(A)}{P(A)}.$$

Використавши формулу повної імовірності (1.24), отримаємо:

$$P_A(H_i) = \frac{P(H_i) \cdot P_{H_i}(A)}{\sum_{i=1}^{n} P(H_i) P_{H_i}(A)}. \qquad (1.26)$$

Дану формулу (1.26) називають формулою Байєса (за ім'ям англійського математика, який опублікував в 1764 р. цю формулу) або формулою гіпотез.

Задача 1.24. В нас є 6 коробок: 2 коробки зі складу H_1, в яких є 3 білі та 2 червоні кульки, 1 коробка із складу H_2, в якій є 4 білі та 1 червона кульки, 3

коробки зі складу H_3, в яких є 2 білі і 3 червоні кульки. Навмання вибирають ящик та кульку. Знайдіть ймовірність того, що

А) біла кулька вийнята зі складу H_1;

Б) біла кулька вийнята зі складу H_2;

В) біла кулька вийнята зі складу H_3.

Розв'язок

Імовірність вибору ящиків з першого складу: $P(H_1)=2/6$.

Імовірність вибору ящиків з другого складу: $P(H_2)=1/6$.

Імовірність вибору ящиків з третього складу: $P(H_3)=3/6$.

Подія *A* – вийнята біла кулька. Ймовірність вибору білої кульки при умові вибору першого складу: $P_{H_1}(A)=3/5$. Ймовірність вибору білої кульки при умові вибору другого складу: $P_{H_2}(A)=4/5$. Ймовірність вибору білої кульки при умові вибору третього складу: $P_{H_3}(A)=2/5$.

Знайдемо, за формулою повної ймовірності (1.24),

$$P(A)=\sum_{i=1}^{3} P(H_i)\,P_{H_i}(A)=\frac{2}{6}\cdot\frac{3}{5}+\frac{1}{6}\cdot\frac{4}{5}+\frac{3}{6}\cdot\frac{2}{5}=\frac{16}{30}=\frac{8}{15}.$$

За формулою Байєса (1.26), одержимо:

А) $$P_A(H_1)=\frac{P(H_1)\cdot P_{H_1}(A)}{\sum_{i=1}^{3} P(H_i)P_{H_i}(A)}=\frac{\frac{2}{6}\cdot\frac{3}{5}}{\frac{16}{30}}=\frac{3}{8}\approx 0{,}38;$$

Б) $$P_A(H_2)=\frac{P(H_2)\cdot P_{H_2}(A)}{\sum_{i=1}^{3} P(H_i)P_{H_i}(A)}=\frac{\frac{1}{6}\cdot\frac{4}{5}}{\frac{16}{30}}=\frac{4}{16}\approx 0{,}25;$$

В) $$P_A(H_3)=\frac{P(H_3)\cdot P_{H_3}(A)}{\sum_{i=1}^{3} P(H_i)P_{H_i}(A)}=\frac{\frac{3}{6}\cdot\frac{2}{5}}{\frac{16}{30}}=\frac{3}{8}\approx 0{,}38.$$

Порівнюючи отримані ймовірності, можна зробити висновок, що біла кулька була вийнята із коробок 1-го або 3-го складу, оскільки імовірність цих подій найбільша.

Лекція 3. Повторні випробування

Тема 3.1. Послідовність незалежних випробувань. Формула Бернуллі

На практиці часто трапляються експерименти, які полягають в тому, що одне й те ж випробування проводять декілька разів. Проте водночас у кожному випробуванні деякі події можуть як відбуватись, так і не відбуватись.

Наприклад, підкидання кубика або монети кілька разів, перевірка деталей на якість, взятих по одній із певної кількості.

Послідовні випробування називаються незалежними, якщо здійснення (настання) будь-якого результату в n-му випробуванні не залежить від результатів у попередніх випробуваннях.

Нехай здійснено n послідовних випробувань, імовірність події в кожному випробуванні однакова, а самі випробування незалежні, тобто поява події в i-му випробуванні не залежить від її появи в інших випробуваннях, то таку схему (серію) випробувань називають схемою Бернуллі.

Якщо випадкова подія відбувається в кожному випробуванні з певною ймовірністю $P(A) = p$, тоді вона не відбувається з ймовірністю $q = 1 - p$. Ймовірність того, що при n випробуваннях згідно зі схемою Бернуллі подія A відбудеться k разів, знаходиться за так званою *формулою Бернуллі*:

$$P_n(k)=C_n^k p^k q^{n-k}. \qquad (1.27)$$

Задача 1.25. Підписано чотири однотипних договори страхування терміном на один рік. Імовірність того, що за будь-яким із договорів протягом року надійде запит на відшкодування, дорівнює 0,1. Знайдіть ймовірність того, що надійде рівно два запити.

Розв'язок

Згідно з умовою задачі, $n = 4$, $k = 2$, $p = 0{,}1$, $q = 1 – 0{,}1 = 0{,}9$.

Скориставшись формулою Бернуллі (1.27), матимемо

$$P_4(2)=C_4^2\cdot 0{,}1^2\cdot 0{,}9^{4-2}=0{,}0486.$$

Задача 1.26. Математично встановлено, що на кожну тисячу дітей, що народились, в середньому припадає 485 дівчат і 515 хлопчиків. У сім’ї є 5 дітей. Знайти ймовірність того, що серед цих дітей:

А) 3 дівчини;

Б) не більше трьох дівчат.

Розв'язок

А) згідно з умовою, ймовірність народження дівчинки $p = 0{,}485$; а ймовірність народження хлопчика $q = 0{,}515$; $n = 5$; $k = 3$.

З користаємось формулою Бернуллі:

$$P_5(3)=C_5^3\cdot(0{,}485)^3\cdot(0{,}515)^2=0{,}31.$$

Б) умова, щоб в сім’ї було не більше трьох дівчат, буде виконуватися тоді, коли серед 5-ти дітей в сім’ї:

Немає ні однієї дівчинки – $P_5(0)$,

$$P_5(0)=C_5^0\cdot(0{,}485)^5\cdot(0{,}515)^0=0{,}02;$$

є одна дівчинка і 4 хлопчики – $P_5(1)$,

$$P_5(1)=C_5^1\cdot(0{,}485)^1\cdot(0{,}515)^4=0{,}17;$$

дві дівчинки і 3 хлопчики – $P_5(2)$,

$$P_5(2)=C_5^2\cdot(0{,}485)^2\cdot(0{,}515)^3=0{,}33;$$

три дівчинки і 2 хлопчики – $P_5(3)$,

$$P_5(3)=C_5^3\cdot(0{,}485)^3\cdot(0{,}515)^2=0{,}31.$$

Усі події несумісні, тому шукана ймовірність, згідно з теоремою додавання несумісних подій, дорівнює сумі ймовірностей кожної із цих подій: $P_5(k\le 3)=P_5(0)+P_5(1)+P_5(2)+P_5(3)=0{,}83.$

Під час розрахунку ймовірностей для різних значень k будуть різні значення ймовірностей.

Тема 3.2. Найімовірніше число настання події

Число k_0 (появи події в n незалежних випробуваннях) називається *найімовірнішим*, якщо ймовірність найбільша.

Для розрахунку найімовірнішого числа появи використовуємо нерівність:

$$np-q\le k_0\le np+p. \qquad (1.28)$$

Можливі такі випадки:

1. Якщо $np-q$ – дробове число, то існує єдине найімовірніше число k_0;
2. Якщо $np-q$ – ціле число, то існує два найімовірніших числа k_0 і k_0+1;
3. Якщо np – ціле число, то найімовірніше число $k_0=np$.

Задача 1.27. Підписано чотири однотипних договори страхування терміном на один рік. Імовірність того, що за будь-яким зі договорів протягом року надійде запит на відшкодування, дорівнює 0,1. Знайдіть найімовірнішу кількість запитів.

Розв'язок

Згідно, з умовою задачі, $n = 4$, $p = 0{,}1$, тоді

$$q=1-0{,}1=0{,}9.$$

Скориставшись формулою (1.28), матимемо:

$$4\cdot 0{,}1-0{,}9\le k_0\le 4\cdot 0{,}1+0{,}1,$$

або $-0{,}5\le k_0\le 0{,}5.$

Отже, найімовірніше число $k_0 = 0$.

Тема 3.3. Локальна теорема Муавра-Лапласа

Формула Бернуллі дає нам можливість обчислити ймовірність того, що подія A з’явиться k разів у n незалежних випробуваннях, якщо ймовірність появи події в кожному випробуванні відповідно стала.

Але використання формули Бернуллі при великих значеннях n призводить до труднощів, пов’язаних із громіздкими розрахунками.

Локальна теорема Муавра-Лапласа визначає так звану асимптотичну формулу, за якою можна наближено розрахувати ймовірність $P_n(k)$, якщо кількість випробувань велика.

В математиці *асимптотичний аналіз* є методом опису граничної поведінки. Найпростіша задача – розгляд функції $f(x)$, при описі її властивостей, коли x стає занадто великим. Наприклад, якщо $f(x) = x^2+3x$, елемент $3x$ стає незначним в порівнянні з x^2, при занадто великих x. Тоді кажуть, що функція $f(x)$ є асимптотично еквівалентна x^2, при $x \to \infty$ й символічно записують як $f(x) \sim x^2$.

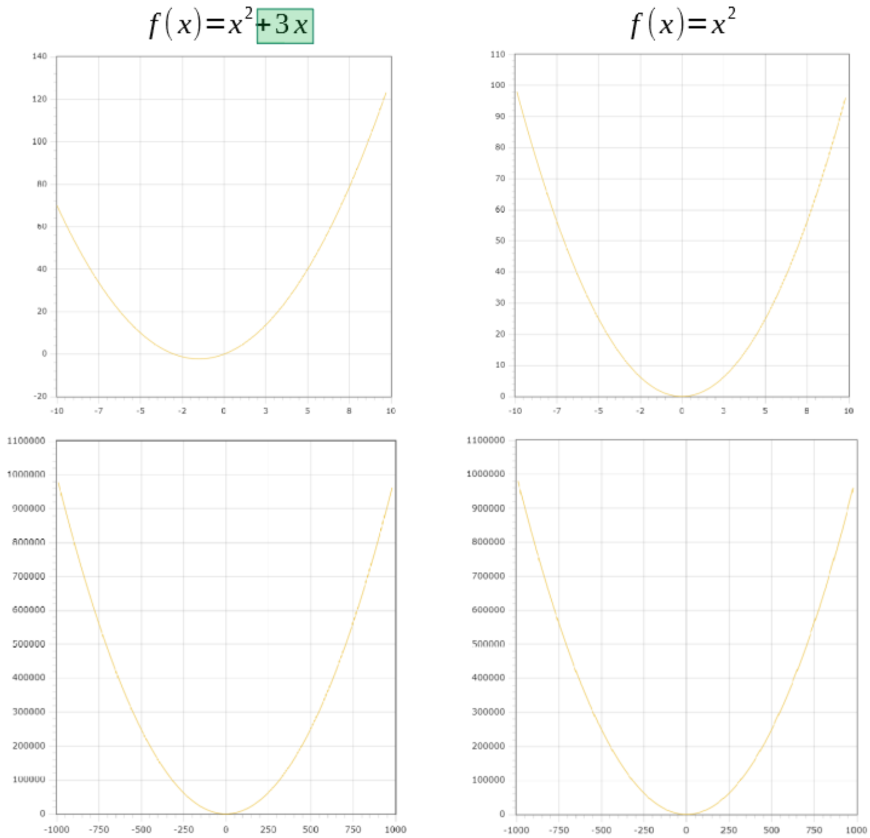


Рис. 1.5. Задача асимптотичного аналізу

Зверніть увагу! Для частинного випадку, коли $p = 0{,}5$, асимптотична формула була виведена ще в 1730 році Муавром.

1783 році Лаплас узагальнив цю формулу для довільних значень p, $0 < p < 1$. Тому цю теорему, ще називають теоремою Муавра-Лапласа.

Теорема: якщо ймовірність p настання події A в n незалежних випробуваннях стала nf задовольняє нерівність: $0 < p < 1$, то ймовірність того, що подія з’явиться k разів у n незалежних випробуваннях дорівнює:

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi(x), \qquad (1.29)$$

де

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x = \frac{k-np}{\sqrt{npq}}.$$

Функція $\varphi(x)$ – парна, тобто виконується рівність

$$\varphi(-x) = \varphi(x).$$

Для функції $\varphi(x)$ складена спеціальна таблиця, в якій за заданими значеннями x знаходять відповідне значення функції.

Таблиця 1.2

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3988	3986	3989	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3952	3845	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3725	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0794	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0011	0011	0010	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001
4,0	0001	0001	0001	0001	0001	0001	0001	0001	0001	0001
4,1	0001	0001	0001	0001	0001	0001	0001	0001	0001	0001
4,2	0001	0001	0001	0001	0000	0000	0000	0000	0000	0000

Задача 1.28. Записи страхової компанії показали, що 30% власників договорів по страхуванню подали запити на відшкодування. Для перевірки було відібрано 15 чоловік (які мають відповідні договори). Знайдіть імовірність того, що протягом року може надійти п’ять запитів.

Розв'язок

Згідно з умовою задачі, $n = 15$, $k = 5$, $p = 0,3$, $q = 1 - 0,3 = 0,7$.

Скористаємось локальною теоремою Лапласа, тому спочатку розрахуємо:

$$x = \frac{5 - 15 \cdot 0,3}{\sqrt{15 \cdot 0,3 \cdot 0,7}} \approx 0,28.$$

За табл. 1.2 знаходимо значення функції $\varphi(x)$ при $x = 0{,}28$;

$$\varphi(0{,}28)=0{,}3836.$$

Тоді за формулою (1.29), отримаємо:

$$P_{15}(5)=\frac{0{,}3836}{\sqrt{15\cdot 0{,}3\cdot 0{,}7}}=\frac{0{,}3836}{1{,}77}=0{,}217.$$

Тема 3.4. Інтегральна теорема Лапласа

Часто виникає задача, коли потрібно знайти, при n незалежних випробувань, в кожному з яких ймовірність появи події A стала і дорівнює p ($0 < p < 1$), ймовірність того, що подія A відбудеться не менше ніж k_1 і не більше ніж k_2 разів. Цю ймовірність позначають: $P_n(k_1, k_2)$.

Отримати достатньо точний результат, у цьому випадку, можна за допомогою інтегральної теореми Лапласа.

Теорема: якщо ймовірність p настання події A в кожному з випробуваннях стала ($0 < p < 1$), то ймовірність $P_n(k_1, k_2)$ того, що подія A відбудеться в n випробуваннях не менш ніж k_1 і не більш ніж k_2 разів, розраховується наближено за такою формулою:

$$P_n(k_1,k_2)\simeq\frac{1}{\sqrt{2\pi}}\int_{x_1}^{x_2} \mathrm{e}^{-\frac{z^2}{2}}dz, \tag{1.30}$$

де

$$x_1=\frac{k_1-np}{\sqrt{npq}}, \qquad x_2=\frac{k_2-np}{\sqrt{npq}}.$$

$$\frac{1}{\sqrt{2\pi}}\int_{x_1}^{x_2} \mathrm{e}^{-\frac{z^2}{2}}dz = \frac{1}{\sqrt{2\pi}}\int_{x_1}^{0} \mathrm{e}^{-\frac{z^2}{2}}dz+\frac{1}{\sqrt{2\pi}}\int_{0}^{x_2} \mathrm{e}^{-\frac{z^2}{2}}dz=\frac{1}{\sqrt{2\pi}}\int_{0}^{x_2} \mathrm{e}^{-\frac{z^2}{2}}dz-\frac{1}{\sqrt{2\pi}}\int_{0}^{x_1} \mathrm{e}^{-\frac{z^2}{2}}dz.$$

Якщо ввести позначення

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{0}^{x} \mathrm{e}^{-\frac{z^2}{2}}dz,$$

то формула (1.30) набуде вигляду

$$P_n(k_1,k_2)=\Phi(x_2)-\Phi(x_1). \tag{1.31}$$

Функція $\Phi(x)$ називається *функцією Лапласа*. $\Phi(x)$ – непарна функція, тобто

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

Функція $\Phi(x)$ також табульована. В табл. 1.3 наведені значення лише до $x = 5$, оскільки для $x > 5$ можна вважати, що $\Phi(x) = 0{,}5$.

Таблиця 1.3

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,50	0,1915	1,00	0,3413	1,50	0,4332	2,00	0,4772	2,50	0,4938
0,01	0,0040	0,51	0,1950	1,01	0,3438	1,51	0,4345	2,01	0,4778	2,52	0,4941
0,02	0,0080	0,52	0,1985	1,02	0,3461	1,52	0,4357	2,02	0,4783	2,54	0,4945
0,03	0,0120	0,53	0,2019	1,03	0,3485	1,53	0,4370	2,03	0,4788	2,56	0,4948
0,04	0,0160	0,54	0,2054	1,04	0,3508	1,54	0,4382	2,04	0,4793	2,58	0,4951
0,05	0,0199	0,55	0,2088	1,05	0,3531	1,55	0,4394	2,05	0,4798	2,60	0,4953
0,06	0,0239	0,56	0,2123	1,06	0,3554	1,56	0,4406	2,06	0,4803	2,62	0,4956
0,07	0,0279	0,57	0,2157	1,07	0,3577	1,57	0,4418	2,07	0,4808	2,64	0,4959
0,08	0,0319	0,58	0,2190	1,08	0,3599	1,58	0,4429	2,08	0,4812	2,66	0,4961
0,09	0,0359	0,59	0,2224	1,09	0,3621	1,59	0,4441	2,09	0,4817	2,68	0,4963
0,10	0,0398	0,60	0,2257	1,10	0,3643	1,60	0,4452	2,10	0,4821	2,70	0,4965
0,11	0,0438	0,61	0,2291	1,11	0,3665	1,61	0,4463	2,11	0,4826	2,72	0,4967
0,12	0,0478	0,62	0,2324	1,12	0,3686	1,62	0,4474	2,12	0,4830	2,74	0,4969
0,13	0,0517	0,63	0,2357	1,13	0,3708	1,63	0,4484	2,13	0,4834	2,76	0,4971
0,14	0,0557	0,64	0,2389	1,14	0,3729	1,64	0,4495	2,14	0,4838	2,78	0,4973
0,15	0,0596	0,65	0,2422	1,15	0,3749	1,65	0,4505	2,15	0,4842	2,80	0,4974
0,16	0,0636	0,66	0,2454	1,16	0,3770	1,66	0,4515	2,16	0,4846	2,82	0,4976
0,17	0,0675	0,67	0,2486	1,17	0,3790	1,67	0,4525	2,17	0,4850	2,84	0,4977
0,18	0,0714	0,68	0,2517	1,18	0,3810	1,68	0,4535	2,18	0,4854	2,86	0,4979
0,19	0,0753	0,69	0,2549	1,19	0,3830	1,69	0,4545	2,19	0,4857	2,88	0,4980
0,20	0,0793	0,70	0,2580	1,20	0,3849	1,70	0,4554	2,20	0,4861	2,90	0,4981
0,21	0,0832	0,71	0,2611	1,21	0,3869	1,71	0,4564	2,21	0,4864	2,92	0,4982
0,22	0,0871	0,72	0,2642	1,22	0,3888	1,72	0,4573	2,22	0,4868	2,94	0,4984
0,23	0,0910	0,73	0,2673	1,23	0,3907	1,73	0,4582	2,23	0,4871	2,96	0,4985
0,24	0,0948	0,74	0,2704	1,24	0,3925	1,74	0,4591	2,24	0,4875	2,98	0,4986
0,25	0,0987	0,75	0,2734	1,25	0,3944	1,75	0,4599	2,25	0,4878	3,00	0,4987
0,26	0,1026	0,76	0,2764	1,26	0,3962	1,76	0,4608	2,26	0,4881	3,05	0,4989
0,27	0,1064	0,77	0,2794	1,27	0,3980	1,77	0,4616	2,27	0,4884	3,10	0,49903
0,28	0,1103	0,78	0,2823	1,28	0,3997	1,78	0,4625	2,28	0,4887	3,15	0,49918
0,29	0,1141	0,79	0,2852	1,29	0,4015	1,79	0,4633	2,29	0,4890	3,20	0,49931
0,30	0,1179	0,80	0,2881	1,30	0,4032	1,80	0,4641	2,30	0,4893	3,25	0,49942
0,31	0,1217	0,81	0,2910	1,31	0,4049	1,81	0,4649	2,31	0,4896	3,30	0,49952
0,32	0,1255	0,82	0,2939	1,32	0,4066	1,82	0,4656	2,32	0,4898	3,35	0,49960
0,33	0,1293	0,83	0,2967	1,33	0,4082	1,83	0,4664	2,33	0,4901	3,40	0,49966
0,34	0,1331	0,84	0,2995	1,34	0,4099	1,84	0,4671	2,34	0,4904	3,50	0,49977
0,35	0,1368	0,85	0,3023	1,35	0,4115	1,85	0,4678	2,35	0,4906	3,60	0,49984
0,36	0,1406	0,86	0,3051	1,36	0,4131	1,86	0,4686	2,36	0,4909	3,70	0,49989
0,37	0,1443	0,87	0,3078	1,37	0,4147	1,87	0,4693	2,37	0,4911	3,80	0,499928
0,38	0,1480	0,88	0,3106	1,38	0,4162	1,88	0,4699	2,38	0,4913	3,90	0,499952
0,39	0,1517	0,89	0,3133	1,39	0,4177	1,89	0,4706	2,39	0,4916	4,00	0,499968
0,40	0,1554	0,90	0,3159	1,40	0,4192	1,90	0,4713	2,40	0,4918	4,20	0,499987
0,41	0,1591	0,91	0,3186	1,41	0,4207	1,91	0,4719	2,41	0,4920	4,40	0,4999946
0,42	0,1628	0,92	0,3212	1,42	0,4222	1,92	0,4726	2,42	0,4922	4,60	0,4999979
0,43	0,1664	0,93	0,3238	1,43	0,4236	1,93	0,4732	2,43	0,4925	4,80	0,4999992
0,44	0,1700	0,94	0,3264	1,44	0,4251	1,94	0,4738	2,44	0,4927	5,00	0,4999997
0,45	0,1736	0,95	0,3289	1,45	0,4265	1,95	0,4744	2,45	0,4929		
0,46	0,1772	0,96	0,3315	1,46	0,4279	1,96	0,4750	2,46	0,4931		
0,47	0,1808	0,97	0,3340	1,47	0,4292	1,97	0,4756	2,47	0,4932		
0,48	0,1844	0,98	0,3365	1,48	0,4306	1,98	0,4761	2,48	0,4934		
0,49	0,1879	0,99	0,3389	1,49	0,4319	1,99	0,4767	2,49	0,4936		

Задача 1.29. Ймовірність того, що деталь не пройшла перевірку, складає 0,2. Знайти імовірність того, що серед 400 випадково відібраних деталей, не перевірених буде від 70 до 100.

Розв'язок

Згідно з умовою, $p = 0{,}2$, $n = 400$, $k_1 = 70$, $k_2 = 100$, $q = 1 - 0{,}2 = 0{,}8$.

Скористаємось інтегральною теоремою Лапласа:

$$P_{400}(70,100)=\Phi(x_2)-\Phi(x_1).$$

Знайдемо верхню та нижню межі інтегрування:

$$x_1=\frac{70-400\cdot 0{,}2}{\sqrt{70\cdot 0{,}2\cdot 0{,}8}}=-1{,}25, \qquad x_2=\frac{100-400\cdot 0{,}2}{\sqrt{100\cdot 0{,}2\cdot 0{,}8}}=2{,}5.$$

Отже,

$$P_{400}(70,100)=\Phi(2{,}5)-\Phi(-1{,}25)=\Phi(2{,}5)+\Phi(1{,}25).$$

За таблицею 1.3 знаходимо, що

$$\Phi(2{,}5)=0{,}4938, \qquad \Phi(1{,}25)=0{,}3944.$$

Тоді

$$P_{400}(70,100)=0{,}4938+0{,}3944=0{,}8882.$$

Із теореми Лапласа випливає теорема, яка характеризує ймовірність відхилення відносної частоти від постійної ймовірності в незалежних випробуваннях.

Теорема: якщо p – стала ($0 < p < 1$), при $n \to \infty$, то ймовірність того, що відхилення відносної частоти появи події (m/n) від постійної її ймовірності p по абсолютному значенню не буде перевищувати задане число ε, визначається згідно з формулою

$$P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)\to 2\Phi\left(\varepsilon\cdot\sqrt{\frac{n}{pq}}\right). \tag{1.32}$$

Теорема Бернуллі при достатньо великих n, p, q дає змогу оцінити ймовірність того, що в n незалежних випробуваннях абсолютне значення відхилення відносної частоти появи події A від її ймовірності p не перевищить заданого додатного числа ε:

$$P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)\approx 2\Phi\left(\varepsilon\cdot\sqrt{\frac{n}{pq}}\right). \tag{1.33}$$

Задача 1.30. Імовірність того, що деталь бракована, становить 0,1. Знайдіть імовірність того, що серед 400 випадково відібраних деталей, відносна частота появи бракованих деталей відхилиться від постійної ймовірності 0,1 за абсолютним значенням не більш ніж на 0,03.

Розв'язок

Згідно з умовою, $p = 0{,}1$, $n = 400$, $\varepsilon = 0{,}03$, $q = 1 - 0{,}1 = 0{,}9$.

Користуючись формулою (1.33), знаходимо:

$$P\left(\left|\frac{m}{400}-0{,}1\right|\le 0{,}03\right)\approx 2\Phi\left(0{,}03\cdot\sqrt{\frac{400}{0{,}1\cdot 0{,}9}}\right)=2\Phi(2).$$

За таблицею 1.3 знаходимо:

$$\Phi(2)=0{,}4772, \qquad 2\Phi(2)=0{,}9544.$$

Отже, шукана ймовірність дорівнює:

$$P\left(\left|\frac{m}{400}-0{,}1\right|\le 0{,}03\right)=0{,}9544.$$

Зміст отриманого результату такий: якщо здійснити достатньо велику кількість перевірок по 400 деталей у кожній, то в 95,44% цих перевірок (контролів) відхилення відносної частоти від постійної ймовірності (0,1) за абсолютним значенням не перевищуватиме 0,03.

Задача 1.31. Імовірність того, що деталь нестандартна становить 0,1. Знайти, скільки необхідно вибрати деталей, щоби з ймовірністю 0,9544 можна було сказати, що відносна частота появи нестандартних деталей відхилиться від постійної ймовірності по абсолютному значенні не більш ніж на 0,03.

Розв'язок

Згідно з умовою, $p = 0{,}1$, $P = 0{,}9544$, $\varepsilon = 0{,}03$, $q = 1 - 0{,}1 = 0{,}9$.

Скористаємось формулою:

$$P\left(\left|\frac{m}{n}-0{,}1\right|\le 0{,}03\right)\approx 2\Phi\left(0{,}03\cdot\sqrt{\frac{n}{0{,}1\cdot 0{,}9}}\right).$$

Звідси отримаємо

$$2\Phi\left(0{,}03\cdot\sqrt{\frac{n}{0{,}1\cdot 0{,}9}}\right)=2\Phi\left(0{,}03\cdot\frac{\sqrt{n}}{\sqrt{0{,}09}}\right)=2\Phi\left(\frac{0{,}03}{0{,}3}\cdot\sqrt{(n)}\right)=2\Phi\left(0{,}1\cdot\sqrt{n}\right).$$

$$2\Phi\left(0{,}1\cdot\sqrt{n}\right)=0{,}9544\Rightarrow\Phi\left(0{,}1\cdot\sqrt{n}\right)=0{,}4772.$$

З іншого боку

$$0{,}1\cdot\sqrt{n}=x.$$

З таблиці 1.3 для $\Phi(x) = 0{,}4772$ знаходимо, що $x = 2$.

Тоді,

$$0{,}1\cdot\sqrt{n}=2\Rightarrow n=400.$$

Зміст отриманого результату такий: якщо здійснити достатньо велику кількість перевірок по 400 деталей у кожній, то в 95,44% цих перевірок відхилення відносної частоти появи бракованої деталі буде відрізнятись від постійної ймовірності 0,1 за абсолютним значенням не більше ніж на 0,03.

Тобто відносна частота знаходиться в межах від 0,07 до 0,13.

Інакше кажучи, кількість бракованих деталей в 95,44% перевірок буде коливатись в межах від 28 (7% від 400) до 52 (13% від 400) штук.

Тема 3.5. Теорема Пуассона

Згідно з теоремою Пуассона, вираховують імовірності у випадках, коли ймовірність настання події стала і дуже мала, а кількість випробувань досить велика.

Теорема: якщо ймовірність настання події досить мала, то для достатньо великої кількості випробувань імовірність того, що подія відбудеться k раз, асимптотично обчислюється за формулою

$$P_n(k)=\frac{a^k\cdot e^{-a}}{k!}, \tag{1.34}$$

де $a = n\cdot p$.

Задача 1.32. Завод по виготовленню моніторів відвантажив на базу 5000 якісних кольорових моніторів. Імовірність того, що під час перевезення

моніторів якість знизиться, дорівнює 0,0002. Знайдіть імовірність того, що на базу потрапить три бракованих телевізори.

Розв'язок

Згідно з умовою задачі, n = 5000, p = 0,0002, k = 3, a = 5000·0,0002 = 1.

Шукана ймовірність, за теоремою Пуассона, тобто, за формулою (1.34), дорівнює:

$$P_{5000}(3)=\frac{1^3\cdot e^{-1}}{3!}=\frac{2{,}73^{-1}}{1\cdot 2\cdot 3}=\frac{\frac{1}{2{,}73}}{6}=\frac{0{,}3663}{6}\approx 0{,}06.$$

Потік подій – це послідовність подій, які настають у випадковий момент часу. Прикладом потоку подій може бути надходження викликів на пункт швидкої медичної допомоги, прибуття ремонтної бригади та інше.

Посеред властивостей, якими володіють потоки, окремо виділяють властивість стаціонарності, відсутності післядії та ординарності.

Властивість стаціонарності характеризується тим, що ймовірність появи k подій за довільний проміжок часу залежить лише від числа k та t (довжини проміжку часу) і не залежить від початку його відліку; водночас вважається, що різні проміжки не перетинаються.

Отже, якщо потік володіє властивістю стаціонарності, то ймовірність появи k подій за проміжок часу t є функцією, яка залежить від k та t.

Властивість відсутності післядії характеризується тим, що ймовірність появи k подій за довільний часовий проміжок не залежить від того, відбувались чи не відбувались події у моменти часу, які передували початку проміжку, що розглядається.

Отже, якщо потік володіє властивістю післядії, то наявна взаємна незалежність появи тієї чи іншої кількості подій в проміжках, які не перетинаються.

Властивість ординарності характеризується тим, що поява двох чи більше подій за невеликий проміжок часу практично неможлива. Тобто ймовірність появи більш ніж однієї події за невеликий проміжок часу, настільки мала у порівнянні з ймовірністю появи лише однієї події, що нею нехтують.

Отже, якщо потік володіє властивістю ординарності, то за нескінченно малого проміжку часу може відбутись не більше однієї події. Надалі будемо розглядати такі потоки, які володіють властивістю стаціонарності.

Інтенсивністю потоку λ називають середню кількість подій, які виникли за одиницю часу.

Якщо постійна інтенсивність потоку відома, то ймовірність появи k подій простого потоку за час t визначається згідно, з формулою Пуассона:

$$P_t(k)=\frac{(\lambda t)^k\cdot e^{-\lambda t}}{k!}. \tag{1.35}$$

Задача 1.33. Середня кількість викликів, яка надходять на АТС за одну хвилину, дорівнює двом. Знайдіть імовірність того, що за 5 хв на АТС надійде 2 виклики.

Розв'язок

Згідно з умовою задачі, $\lambda = 2$, $t = 5$, $k = 2$.

Використовуючи формулу Пуассона (1.35), отримаємо:

$$P_5(2)=\frac{(2\cdot 5)^2\cdot e^{-2\cdot 5}}{2!}=\frac{100\cdot 0{,}000045}{2}\approx 0{,}000025.$$

Ця подія неможлива, оскільки ймовірність її настання дуже мала.

Лекція 4. Випадкові величини і функції розподілу

Тема 4.1. Дискретна випадкова величина

Випадковою (X) називають величину, яка в результаті випробування набуде тільки одне з можливих значень, яке наперед невідоме та залежне від випадкових причин, які попередньо не можуть бути враховані.

Законом розподілу випадкової величини називають таке співвідношення, яке встановлює зв’язок між можливими її значеннями та відповідними їм ймовірностями.

Дискретною випадковою величиною називають величину, яка набуває ізольовані (дискретні) числові значення з усіх можливих.

Для будь-якої випадкової величини відомості про їх можливі значення недостатні. Потрібно також знати ще ймовірність появи цих значень. Перелік всіх випадкових значень дискретної випадкової величини та їх ймовірностей відповідно носить назву *закону розподілу дискретних випадкових величин*.

Причому повинна обов'язково виконуватись рівність

$$p_1+p_2+p_3+...+p_n=1.$$

Закон розподілу дискретних випадкових величин задається нижчеописаними способами.

Табличний спосіб. Цей спосіб полягає в тому, що надається таблиця, в якій записані окремі значення випадкової величини X та їх ймовірності P відповідно.

X	x_1	x_2	x_3	...	x_n
P	p_1	p_2	p_3	...	p_n

*Графічний спосіб***.** Згідно із цим способом функцію задають графіком у декартовій системі координат. На осі абсцис відкладають значення випадкової величини, а на осі ординат – значення їх імовірностей.

Якщо на площині XOY з'єднати послідовно точки (x_1, p_1), (x_2, p_2), … , (x_n, p_n), то дістанемо криву, яку називають *багатокутником розподілу випадкової величини X*. Прикладом такого способу задання закону розподілу є так званий полігон розподілу ймовірностей.

*Аналітичний спосіб***.** Згідно із цим способом, надається математична формула, за якою можна розрахувати ймовірності того, що випадкова величина набуде те чи інше значення.

Найважливіші такі види дискретних розподілів:

1. Рівномірний.
2. Біноміальний.
3. Показниковий.
4. Геометричний.
5. Гіпергеометричний.

Рівномірний розподіл – розподіл випадкової величини, при якому випадкова величина набуває n різних значень з однаковими ймовірностями.

$$P_k = \frac{1}{n}, \quad \text{де } k = 1, 2, 3, ..., n. \tag{1.36}$$

Кількість очок, які випадають на верхній грані грального куба, є дискретною випадковою величиною, яка має наступний рівномірний закон розподілу:

X	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

Задача 1.34. Тільки один з п’яти ключів відмикає замок. Скласти закон розподілу кількості випробувань відмикання замка, якщо випробуваний ключ у подальшому не використовується.

Розв'язок

Імовірність відмикання замка кожним ключем

$$p = \frac{1}{5}.$$

У цьому випадку випадкова величина – це кількість випробуваних ключів під час відмикання замка, яка набуває значення: 1, 2, 3, 4, 5, (оскільки маємо п’ять ключів).

При $X = 1$, $P(1) = 0{,}2$ (ймовірність відкрити 1 ключем з п’яти).

При $X = 2$, $P(2) = 0{,}2$ (ймовірність відкрити 2 ключем з п’яти).

При $X = 3$, $P(3) = 0{,}2$ (ймовірність відкрити 3 ключем з п’яти).

При $X = 4$, $P(4) = 0{,}2$ (ймовірність відкрити 4 ключем з п’яти).

При $X = 5$, $P(5) = 0{,}2$ (ймовірність відкрити 5 ключем з п’яти).

Закон розподілу цієї випадкової величини буде мати вигляд:

X	1	2	3	4	5
P	0,2	0,2	0,2	0,2	0,2

Здійснимо контроль: $0{,}2+0{,}2+0{,}2+0{,}2+0{,}2=1$.

Біноміальний розподіл (розподіл Бернуллі) – розподіл випадкової величини зі значеннями $k = 0, 1, 2, 3, \ldots, n$, імовірності якої знаходяться з наступної формули

$$P_k = C_n^k p^k q^{n-k}. \quad (1.37)$$

Біноміальним розподілом характеризується випадкова величина, яка дорівнює кількістю «успіхів» у серії з *n* незалежних випробувань, у кожному з яких «успіх» відбувається з певною ймовірністю *p*.

Задача 1.35. Імовірність скласти іспит на «відмінно» для кожного з шести студентів дорівнює 0,4. Наведіть алгоритм розподілу кількості п'ятірок («*A*» або відмінно), які студенти отримали на іспиті.

Розв'язок

Позначимо випадкову величину через число *X* – кількість п'ятірок, отриманих студентами на екзамені. Вона відповідно набуває значення: 0, 1, 2, 3, 4, 5, 6. Згідно, з умовою, $p = 0{,}4$; $q = 0{,}6$. Знайдемо імовірності для відповідних значень *X*.

При $X = 0$ (на іспиті жоден студент не отримав «*A*»)

$$P_0 = C_6^0 (0{,}4)^0 (0{,}6)^6 \approx 0{,}047.$$

При $X = 1$ (на іспиті 1 студент отримав «*A*»)

$$P_1 = C_6^1 (0{,}4)^1 (0{,}6)^5 \approx 0{,}187.$$

При $X = 2$ (на іспиті 2 студенти отримали «*A*»)

$$P_2 = C_6^2 (0{,}4)^2 (0{,}6)^4 \approx 0{,}311.$$

При $X = 3$ (на іспиті 3 студенти отримали «*A*»)

$$P_3 = C_6^3 (0{,}4)^3 (0{,}6)^3 \approx 0{,}276.$$

При $X = 4$ (на іспиті 4 студенти отримали «*A*»)

$$P_4 = C_6^4 (0{,}4)^4 (0{,}6)^2 \approx 0{,}138.$$

При $X = 5$ (на іспиті 5 студентів отримали «*A*»)

$$P_5 = C_6^5 (0{,}4)^5 (0{,}6)^1 \approx 0{,}037.$$

При $X = 6$ (на іспиті всі студенти отримали «*A*»)

$$P_6 = C_6^6 (0{,}4)^6 (0{,}6)^0 \approx 0{,}004.$$

Закон розподілу для цьої випадкової величини буде мати вигляд:

X	0	1	2	3	4	5	6
P	0,047	0,187	0,311	0,276	0,138	0,037	0,004

Відповідний багатокутник розподілу для цієї випадкової величини наведений на рис. 1.6.

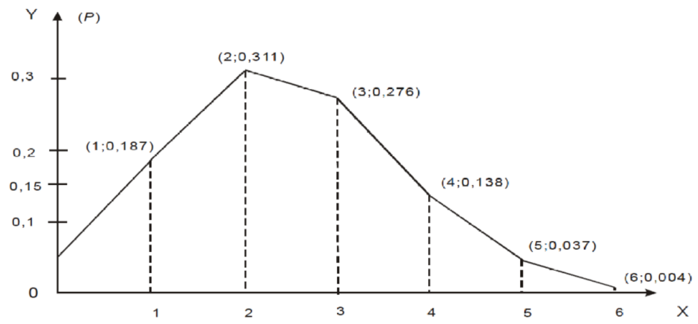


Рис. 1.6

Показниковий розподіл (так званий розподіл Пуассона) – розподіл випадкової величини зі значеннями k = 1, 2, 3, …, n, ймовірності якої знаходяться з допомогою формули

$$P_k = \frac{a^k \cdot e^{-a}}{k!}, \quad \text{де } a = n \cdot p. \tag{1.38}$$

Зверніть увагу! З допомогою цього розподілу будуються математичні моделі роботи різного роду об’єктів по обслуговуванню населення.

Задача 1.36. У магазин надійшло 100 пляшок мінеральної води. Ймовірність того, що під час перевезення пляшка розіб’ється, становить 0,01. Знайдіть ряд розподілу випадкової величини (X), яка характеризує кількість розбитих пляшок ($X \in [1, 5]$).

Розв'язок

Імовірності появи окремих значень випадкової величини розраховується за формулою

$$P_k = \frac{a^k \cdot e^{-a}}{k!}.$$

де, згідно з умовою, $a = n \cdot p = 100 \cdot 0{,}01 = 1$.

При X = 1 (тобто, при перевезенні розбилась одна пляшка мінеральної води):

$$P_1 = \frac{1^1 \cdot e^{-1}}{1!} \approx 0{,}367884.$$

При X = 2 (тобто при перевезенні розбилось дві пляшки мінеральної води):

$$P_2 = \frac{1^2 \cdot e^{-1}}{2!} \approx 0{,}18394.$$

При X = 3 (тобто при перевезенні розбилось три пляшки мінеральної води):

$$P_3 = \frac{1^3 \cdot e^{-1}}{3!} \approx 0{,}06131.$$

При X = 4 (тобто при перевезенні розбилось чотири пляшки мінеральної води):

$$P_4 = \frac{1^4 \cdot e^{-1}}{4!} \approx 0{,}01533.$$

При X = 5 (тобто при перевезенні розбилось п'ять пляшок мінеральної води):

$$P_5 = \frac{1^5 \cdot e^{-1}}{5!} \approx 0{,}00306.$$

Закон розподілу для цієї випадкової величини має вигляд:

X	1	2	3	4	5
P	0,36788	0,18394	0,06131	0,01533	0,00306

Відповідний багатокутник розподілу для цієї випадкової величини такий (рис. 1.7):

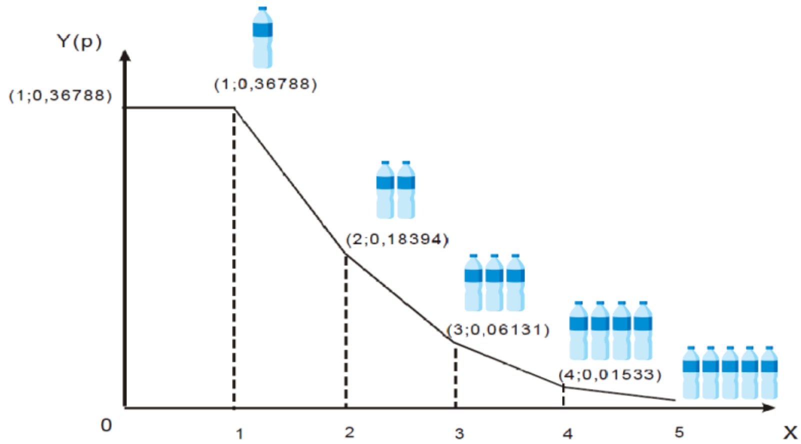


Рис. 1.7

Геометричний розподіл – розподіл випадкової величини, яка набуває значень $k \in \mathbb{N}$ з ймовірностями:

$$P_k = q^{k-1} \cdot p, \tag{1.39}$$

де k = 1, 2, 3, …, n, p – число фіксоване, $q = 1 - p$.

Зверніть увагу! Геометричним розподілом характеризується випадкова величина, яка дорівнює кількості спроб до появи першого «успіху» в серії незалежних випробувань, у кожному з яких імовірність успіху p дорівнює певному значенню.

Задача 1.37. Із рушниці відбувається стрільба по мішені до першого влучення. Ймовірність влучання в ціль становить 0,6. Знайдіть ймовірність того, що влучення відбудеться при третьому пострілі.

Розв'язок

Згідно з умовою задачі, p = 0,6, відповідно $q = 1 - p = 1 - 0{,}6 = 0{,}4$, а k = 3. Тоді

$$p_3 = q^{k-1} \cdot p = 0{,}4^{3-1} \cdot 0{,}6 = 0{,}4^2 \cdot 0{,}6 = 0{,}096.$$

Гіпергеометричний розподіл – розподіл випадкової величини, яка набуває значень k = 1, 2, 3, …, n, а ймовірності знаходяться з допомогою формули

$$P_k = \frac{C_n^k \cdot C_{N-n}^{m-k}}{C_N^k}, \tag{1.40}$$

де $n \geq m$, а $N \geq n$.

Зверніть увагу! Цей розподіл знаходить широке використання. Він виникає під час перевірки виготовленої продукції, вивчення думки громадськості та інше.

Задача 1.38. Серед 50 деталей 20 – якісних. Знайдіть ймовірність того, що серед 5 випадково відібраних деталей 3 виявляться бракованими.

Розв'язок

Згідно з умовою задачі, $N = 50$, $n = 20$, $m = 5$, $k = 3$. Тоді

$$P_3 = \frac{C_{20}^{3} \cdot C_{50-20}^{5-3}}{C_{50}^{3}} = \frac{C_{20}^{3} \cdot C_{30}^{2}}{C_{50}^{3}} = 0{,}234.$$

Тема 4.2. Неперервна випадкова величина

Неперервною випадковою величиною називають величину, яка набуває всі свої можливі значення з деякого інтервалу (проміжку) значень.

На відміну від розподілу дискретної випадкової величини, розподіл неперервної випадкової величини неможливо задати шляхом зазначення значень, яких вона набуває, та відповідних їм ймовірностей.

Множина можливих значень неперервної випадкової величини незліченна. Відповідно виникає питання про інший спосіб задання випадкової величини. Він полягає у заданні функції розподілу.

Інтегральною функцією розподілу випадкової величини (CDF – cumulative distribution function) або *функцією розподілу ймовірностей* називають функцію $F(x)$, яка визначає для кожного значення x ймовірність того, що випадкова величина X набуває значення, меншою від дійсного числа x, тобто

$$F(x) = P(X < x). \qquad (1.41)$$

Геометрично це виглядає так: $F(x)$ є ймовірністю того, що випадкова величина набуде значення, яке зображується точкою, що знаходиться лівіше значення x. Випадкову величину будемо вважати неперервною, якщо її інтегральна функція розподілу $F(x)$ неперервно диференційована.

Властивості інтегральної функції розподілу:

1. Інтегральна функція розподілу $F(x)$ задовольняє нерівність

$$0 \leq F(x) \leq 1.$$

Ця властивість випливає з означення інтегральної функції як ймовірності події.

2. Інтегральна функція $F(x)$ неспадна, тобто $F(x_2) \geq F(x_1)$, якщо $x_2 > x_1$.

Звідси випливає такий наслідок: імовірність того, що випадкова величина X набуде значення, що належить проміжку (a, b), дорівнює приросту інтегральної функції:

$$P(a < X < b) = F(b) - F(a). \qquad (1.42)$$

3. Якщо можливі значення випадкової величини X належать проміжку (a, b) то: $F(x) = 0$, при $x \leq a$; $F(x) = 1$, при $x \geq b$.

Задача 1.39. Дискретна випадкова величина X задана табличним способом:

X	1	4	8
P	0,3	0,1	0,6

Знайдіть інтегральну функцію і побудуйте її графік.

Розв'язок

Випадкова величина не може бути меншою за одиницю. Отже, якщо для всіх $x \le 1$ події неможливі, то $F(x) = 0$ (третя властивість).

Якщо $1 < x \le 4$, то $F(x) = 0{,}3$.

Нехай, наприклад, $x = 1{,}4$. Тоді $F(1{,}4)$ є ймовірністю події $X < 1{,}4$. Але ця випадкова величина тільки в одному випадку набуває значень, меншим від 1,4, а саме з ймовірністю 0,3.

Якщо $4 < x \le 8$, то $F(x) = 0{,}3 + 0{,}1 = 0{,}4$. Ці дві події несумісні, то згідно з теоремою додавання ймовірностей, вони дорівнюють сумі ймовірностей.

Якщо $x > 8$, то $F(x) = 0{,}4 + 0{,}6 = 1$.

Отже, інтегральна функція буде мати такий вигляд

$$F(x)=\begin{cases} 0, \text{ при } x \le 1 \\ 0{,}3, \text{ при } 1 < x \le 4 \\ 0{,}4, \text{ при } 4 < x \le 8 \\ 1, \text{ при } x > 8 \end{cases}$$

Зобразимо графік $F(x)$:

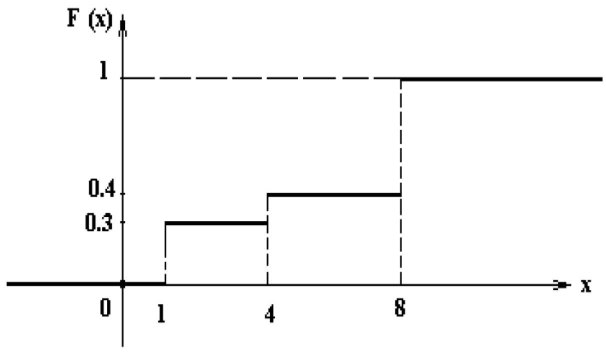


Рис. 1.8

Задача 1.40. Випадкова величина X задана інтегральною функцією розподілу

$$F(x)=\begin{cases} 0, \text{ при } x \le 0; \\ x-4, \text{ при } 0 < x \le 1; \\ 1, \text{ при } x > 1. \end{cases}$$

Знайдіть імовірність того, що в результаті випробування випадкова величина набуде значення, що належить інтервалу (-1\2, +1\2).

Розв'язок

Відповідно до формули (1.42), знаходимо:

$$P\left(-\frac{1}{2} < X < \frac{1}{2}\right) = F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right) = \left(\frac{1}{2} - 4\right) - 0 = -\frac{7}{2} = -3{,}5.$$

Диференціальною функцією розподілу (PDF – probability density function), або *густиною ймовірності* (*щільність*) $f(x)$ неперервної випадкової величини X, називають першу похідну від інтегральної функції:

$$f(x)=F'(x). \tag{1.43}$$

З цього випливає, що інтегральна функція розподілу неперервної випадкової величини X є, по суті, первісною для диференціальної функції.

Знаючи інтегральну функцію розподілу $F(x)$, можна легко знайти диференціальну функцію $f(x)$, бо, згідно з означенням: $f(x) = F'(x)$.

Знаючи диференціальну функцію $f(x)$, інтегральна функція $F(x)$ знаходиться за формулою

$$F(X)=\int_{-\infty}^{+\infty} f(x)\,dx. \tag{1.44}$$

Властивості диференціальної функції розподілу:

1. Диференціальна функція додатна – $f(x) \geq 0$. Ця властивість зумовлена тим, що інтегральна функція неспадна: $f(x) = F'(x) \geq 0$.
2. Невластивий інтеграл від диференціальної функції дорівнює одиниці, тобто:

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

Ця властивість зумовлена тим, що ймовірність випадкової величини потрапляє в інтервал (+∞, -∞), що є подією достовірною. Зокрема, якщо всі можливі значення випадкової величини належать проміжку (a, b), то

$$\int_{a}^{b} f(x)\,dx=1.$$

Знаючи диференціальну функцію, можна розрахувати ймовірність того, що неперервна випадкова величина набуде значення, що належить інтервалу (a, b).

3. Ймовірність того, що неперервна випадкова величина набуде значення, що належить проміжку (a, b), дорівнює визначеному інтегралу від диференціальної функції, в межах від a до b:

$$P(a<X<b)=\int_{a}^{b} f(x)\,dx. \tag{1.45}$$

Задача 1.41. Заданий рівномірний закон розподілу неперервної випадкової величини X:

$$f(x)=\begin{cases} 0, \text{при } x\leq a \\ \dfrac{1}{b-a}, \text{при } a<x\leq b \\ 0, \text{при } x>b \end{cases}$$

Знайти інтегральну функцію рівномірного закону розподілу.

Розв'язок

Скористаємось формулою (1.44):

$$F(X)=\int_{-\infty}^{+\infty} f(x)\,dx.$$

Якщо $x \le a$, то $f(x) = 0$, відповідно $F(x) = 0$.

Якщо $a < x \le b$, $f(x) = 1/(b - a)$, тоді

$$F(X)=\int_{-\infty}^{+\infty} f(x)\,dx=\int_{-\infty}^{a} 0\,dx+\int_{a}^{x} \frac{1}{b-a}\,dx=\frac{x-a}{b-a}.$$

Якщо $x > b$, то

$$F(X)=\int_{-\infty}^{a} 0\,dx+\int_{a}^{b} \frac{1}{b-a}\,dx+\int_{b}^{x} 0\,dx=\frac{b-a}{b-a}=1.$$

Отже, шукана інтегральна функція рівномірного заданого закону розподілу матиме вигляд:

$$F(x)=\begin{cases} 0, при\ x\le a \\ \dfrac{x-a}{b-a}, при\ a<x\le b \\ 1, при\ x>b \end{cases}$$

Задача 1.42. Неперервна випадкова величина X задана інтегральною функцією розподілу:

$$F(x)=\begin{cases} 0, при\ x\le 0 \\ x^2, при\ 0<x\le 1 \\ 1, при\ x>1 \end{cases}$$

Знайти:

А) диференціальну функцію розподілу;

Б) ймовірність того, що випадкова величина *X* набуде значення з інтервалу (0,1/2).

Побудувати графіки інтегральної та диференціальної функції розподілу.

Розв'язок

А) за означенням диференціальної функції, маємо:

$$f(x)=F'(x)=\begin{cases} 0, при\ x\le 0 \\ 2x, при\ 0<x\le 1 \\ 0, при\ x>1 \end{cases}$$

Б) ймовірність того, що випадкова величина *X* набуде значення з інтервалу (*a*, *b*), визначається формулою (1.45):

$$P\left(0<X<\frac{1}{2}\right)=\int_{0}^{\frac{1}{2}} 2x\,dx=\frac{1}{4}.$$

Схематичні графіки функцій $F(x)$ і $f(x)$ мають вигляд:

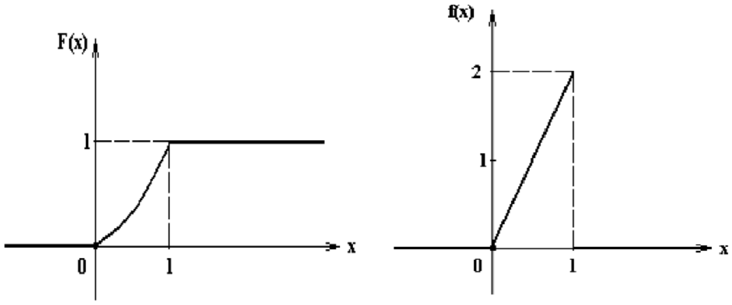


Рис. 1.9

Часто диференціальну функцію розподілу неперервної випадкової величини ще називають *густиною ймовірності в точці x*.

4.2.1. Нормальний закон розподілу

Нормальний закон розподілу (який, часто називають *законом Гауса*) виконує важливу роль в теорії ймовірностей і має серед інших законів розподілу особливе значення. Цей закон часто трапляється на практиці. Важливою особливістю цього закону є те, що він є граничним законом, до якого наближаються решта законів розподілу, які часто трапляються в типових умовах.

Нормальним законом розподілу випадкової величини X називають закон, диференціальна функція розподілу якого задається формулою:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-(x-a)^2}{2\sigma^2}} \qquad (1.46)$$

Нормальний закон розподілу визначається двома параметрами: a і σ. Достатньо задати ці два параметри, щоб задати нормальний закон розподілу.

Графік диференціальної функції нормального закону розподілу називають *нормальною кривою* або *кривою Гауса*.

Дослідивши диференціальну функції $f(x)$ (1.46), шляхом методу диференціального числення, одержимо:

1. Функція $f(x)$ визначена для всіх значень x.
2. При всіх значеннях x функція $f(x)$ набуває додатних значень, тобто графік нормальної кривої розміщений над віссю OX.
3. Границя функції за необмеженого зростання x (за абсолютним значенням) дорівнює нулю, тобто:

$$\lim_{x\to\pm\infty}\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-(x-a)^2}{2\sigma^2}}=0.$$

Це означає, що вісь OX є горизонтальною асимптотою графіка функції.

Величина a носить назву *математичного сподівання*, а *σ – дисперсії X*.

4. Дослідивши функцію на екстремум, знаходимо функцію, яка має максимум, що рівний:

$$f(a)=\frac{1}{\sigma\sqrt{2\pi}}.$$

5. Функція $f(x)$ має дві точки перегину:

$$\left(a-\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right), \qquad \left(a+\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right).$$

Схематичний графік нормальної кривої такий:

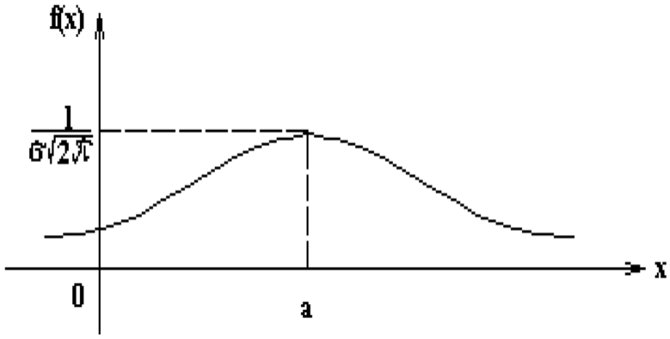


Рис. 1.10

При $a = 0$ і $\sigma = 1$ нормальну криву називають нормованою і позначають:

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

Ймовірність попадання в заданий інтервал величини, яка розподілена за нормальним законом, визначається за формулою

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \qquad (1.47)$$

Задача 1.43. Випадкова величина X розподілена за нормальним законом. Дано $a = 30$ та $\sigma = 10$. Знайти ймовірність того, що X набуде значення, що належить інтервалу (10, 50).

Розв'язок

Згідно з умовою задачі, $a = 30$ і $\sigma = 10$, $\alpha = 10$ і $\beta = 50$, тому за формулою (1.47) отримаємо:

$$P(10 < X < 50) = \Phi\left(\frac{50-30}{10}\right) - \Phi\left(\frac{10-30}{10}\right) = 2\Phi(2).$$

За допомогою таблиці значень функції Лапласа (див. попередню лекцію) знаходимо, що

$$\Phi(2) = 0{,}4772.$$

Звідси:

$$P(10 < X < 50) = 2 \cdot 0{,}4772 = 0{,}9544.$$

Правило трьох сигм: якщо випадкова величина X розподілена відповідно до нормально закону, то $P(|X - a| \geq 3\sigma) \to 0$. Тобто імовірність того, якщо абсолютна величина відхилення X від її математичного сподівання a більше 3σ, прямує до 0, то це означає, що $|X - a| < 3\sigma$ – практично достовірна подія.

На практиці це правило використовують так: «якщо закон розподілу випадкової величини X невідомий, але $|X - a| < 3\sigma$, тоді можна припустити, що X розподілена нормально».

Нормальний закон розподілу випадкової величини широко застосовується на практиці. Це пов’язано з тим, що коли на випадкову величину впливає сума багатьох факторів, кожний із яких впливає як завгодно

мало, то випадкова величина має розподіл, близький до нормального. На практиці розглядають ще інші закони розподілу, такі, наприклад, як розподіл χ^2, розподіл Стьюдента, розподіл Фішера, Гамма-розподіл та інші. Розглянемо основні з них.

Розподіл χ^2 («*хі-квадрат*»). Нехай X_i (i =1, 2, n) – нормальні, незалежні випадкові величини. Тоді сума квадратів цих величин

$$\chi^2 = \sum_{i=1}^{n} X_i^2,$$

розподілена за законом χ^2 з $k = n$ степенями вільності. Якщо величини X_i зв’язані одним лінійним співвідношенням, наприклад $\Sigma X_i = n * X$, то кількість степенів вільності буде $k = n - 1$. Диференціальна функція розподілу має вигляд

$$f(x) = \begin{cases} 0 \; при \; x \le 0 \\ \dfrac{1}{2^{\frac{k}{2}} \cdot \Gamma\left(\frac{k}{2}\right)} \cdot e^{-\frac{x}{2}} \cdot x^{\frac{k}{2}-1} \; при \; x > 0 \end{cases},$$

де

$$\Gamma(x) = \int_0^{\infty} t^{x-1} \cdot e^{-t} dt.$$

Розподіл χ^2 визначається параметром – кількістю степенів вільності k. Коли k зростає, розподіл χ^2 «повільно» прямує до нормального розподілу.

Розподіл Стьюдента. Нехай X – нормальна випадкова величина ($a = 0$, $\sigma = 1$), Y – незалежна від X величина, яка розподілена за законом «*хі-квадрат»* із k степенями вільності.

Тоді величина:

$$T = \frac{X}{\sqrt{\frac{Y}{k}}}$$

має розподіл, який називають *t-розподілом* або *розподілом Стьюдента* з k степенями вільності.

При зростанні k розподіл Стьюдента «швидко» наближається до нормального розподілу.

F-розподіл або *розподіл Фішера*. Нехай X та Y незалежні випадкові величини, розподілені за законом χ^2 зі степенями вільності k_1 та k_2 відповідно. Тоді величина

$$F = \frac{X / k_1}{Y / k_2}$$

має розподіл, який називається F-розподілом зі степенями вільності k_1 та k_2 відповідно (інколи його позначають через V^2).

Густина ймовірності цього розподілу така:

$$f(x)=\begin{cases} 0 \; при \; x\le 0; \\ C_0 \dfrac{x^{(k_1-2)/2}}{(k_2+k_1x)^{(k_1+k_2)/2}} \; при \; x>0, \end{cases} \qquad C_0=\frac{\Gamma\left(\frac{(k_1+k_2)}{2}\right)\cdot k_1^{k_1/2}\cdot k_2^{k_2/2}}{\Gamma(k_1/2)\cdot\Gamma(k_2/2)}.$$

Видно, що F-розподіл визначається двома параметрами – кількостями степенів вільності (k_1, k_2).

4.2.2 Показниковий закон розподілу

Показниковим (*експоненціальним*) *законом розподілу* випадкової величини X називають розподіл ймовірностей, який описується диференціальною функцією:

$$f(x)=\begin{cases} 0 \; при \; x<0 \\ \lambda e^{-\lambda x} \; при \; x\ge 0 \end{cases}, \tag{1.48}$$

де λ – стала величина $\lambda > 0$.

Прикладом неперервної випадкової величини, яка характеризує показниковий закон розподілу, може бути випадковий час між двома послідовними подіями простого потоку (простий потік Пуассона).

Скориставшись формулою (1.44), можна знайти інтегральну функцію розподілу показникового закону:

$$F(X)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{0} 0\cdot dx+\lambda\int_{0}^{x} e^{-\lambda x}dx=1-e^{-\lambda x},$$

тобто

$$F(X)=1-e^{-\lambda x}. \tag{1.49}$$

Графіки інтегральної та диференціальної функцій показникового розподілу мають такий вигляд:

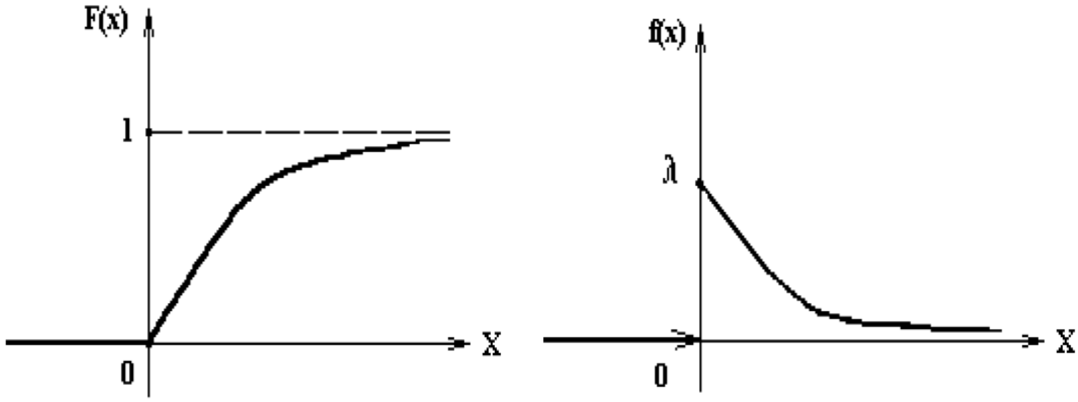


Рис. 1.11

Задача 1.44. Диференціальна функція має вигляд

$$f(x)=\begin{cases} 0 \; при \; x<0; \\ 2e^{-2x} \; при \; x\ge 0. \end{cases}$$

Знайдіть імовірність того, що у результаті випробування випадкова величина X набуде значення з проміжку (0,3; 1).

Розв'язок

Імовірність того, що у результаті випробування випадкова величина X набуде значення з проміжку (0,3; 1), будемо шукати за допомогою формули $P(a < X < b) = F(b) - \mathrm{F}(a)$. Скористаємось при цьому формулою (1.49). Згідно з умовою, $\lambda = 2$. Тоді

$$P(0{,}3<X<1)=F(1)-F(0{,}3)=\left(1-\mathrm{e}^{-2\cdot 1}\right)-\left(1-\mathrm{e}^{-2\cdot 0{,}3}\right) =$$
$$= \left(1-0{,}13534\right)-\left(1-0{,}54881\right)=0{,}86466-0{,}45119=0{,}41347.$$

Тема 4.3. Багатовимірні випадкові величини

Окрім одновимірних випадкових величин, існують випадкові величини, можливі значення яких визначаються за допомогою двох, трьох, … , n числами. Такі величини називаються відповідно *двовимірними*, *трьохвимірними*, …, *n-вимірними випадковими величинами*.

Позначимо через (X, Y) двовимірну величину. Обидві величини X і Y розглядаються одночасно, утворюючи так звану систему двох випадкових величин. Двовимірну випадкову величину (X, Y) геометрично можна інтерпретувати як деяку випадкову точку $M(X, Y)$ на площині (тобто як точку з випадковими координатами).

Наприклад, верстат-автомат штампує дерев'яні плитки. Якщо, для контролю, ми маємо такі розміри, як довжина X та ширина Y, то отримаємо двовимірну випадкову величину.

Законом розподілу дискретної двовимірної величини (X і Y) називають перелік всіх можливих значень цієї величини (тобто пар чисел) x_i, y_j та їх ймовірностей $p(x_i, y_j)$ $i = 1, 2, …, n, j = 1, 2, …, m$. Закон розподілу подають у вигляді таблиці з подвійним входом. Перший рядок таблиці містить усі можливі значення величини X, а перший стовпчик – всі можливі значення величини Y. В комірках на перетині стовпчика x_i і рядка y_j вказані ймовірності $p(x_i, y_j)$ того, що двовимірна випадкова величина набуде значення (x_i, y_j).

Таблиця 1.4

X/Y	x_1	x_2	…	x_i
y_1	$p(x_1, y_1)$	$p(x_2\ y_1)$	…	$p(x_i, y_1)$
y_1	$p(x_1, y_2)$	$p(x_2, y_2)$	…	$p(x_i, y_2)$
…	…	…	…	…
y_j	$p(x_1, y_j)$	$p(x_2, y_j)$	…	$p(x_i, y_j)$

Оскільки події X, Y утворюють повну групу подій, то сума ймовірностей, розміщених в усіх комірках таблиці, дорівнює 1. Знаючи закон розподілу 2D випадкової величини, можна знайти закон розподілу кожної з її складових.

Задача 1.45. Знайти закон розподілу складових двовимірної випадкової величини, заданої законом розподілу:

X/Y	x_1	x_2	x_3
y_1	0,3	0,1	0,2
y_2	0,18	0,16	0,06

Розв'язок

Розклавши ймовірності за стовпчиками, одержимо ймовірності можливих значень

$$P(x_1)=0{,}30+0{,}18=0{,}48;$$
$$P(x_2)=0{,}10+0{,}16=0{,}26;$$
$$P(x_3)=0{,}20+0{,}06=0{,}26.$$

X	x_1	x_2	x_3
P	0,48	0,26	0,26

Контроль: 0,48 + 0,26 + 0,26 = 1.

Аналогічно, склавши ймовірності по відповідних рядках, одержимо ймовірності можливих значень:

$$P(y_1)=0{,}30+0{,}10+0{,}20=0{,}60;$$
$$P(y_2)=0{,}18+0{,}16+0{,}06=0{,}40.$$

Закон розподілу складової випадкової величини Y матиме вигляд:

X	y_1	y_2
P	0,6	0,4

Контроль: 0,6 + 0,4 = 1.

Лекція 5. Числові характеристики випадкових величин

Тема 5.1. Числові характеристики дискретних випадкових величин

До числових характеристик, які описують розподіл випадкової величини, належать:

1. Математичне сподівання;
2. Дисперсія;
3. Середньоквадратичне відхилення;
4. Мода;
5. Медіана.

Математичне сподівання дискретної випадкової величини – це сума добутків всіх можливих значень цієї величини на відповідні ймовірності, яку позначають через $M(X)$.

Нехай випадкова величина X задана табличним законом розподілу,

X	x_1	x_2	x_3	...	x_n
P	p_1	p_2	p_3	...	p_n

Тоді математичне сподівання визначається за формулою

$$M(X)=x_1\cdot p_1+x_2\cdot p_2+...+x_n\cdot p_n, \quad \text{або} \quad M(X)=\sum_{i=1}^{n} x_i\cdot p_i. \qquad (1.50)$$

Задача 1.46. Знайдіть математичне сподівання випадкової величини X, знаючи закон її розподілу.

X	3	5	2
P	0,1	0,6	0,3

Розв'язок

Шукане математичне сподівання випадкової величини X дорівнює сумі добутків усіх можливих значень цієї величини на відповідні ймовірності:

$$M(X)=3\cdot 0{,}1+5\cdot 0{,}6+2\cdot 0{,}3=3{,}9.$$

Математичне сподівання має таку ж розмірність, що й випадкова величина.

Задача 1.47. Знайти математичне сподівання кількості появи події A в одному випробуванні, якщо ймовірність події A дорівнює p.

Розв'язок

Випадкова величина X – це кількість появи події A в одному випробуванні, яке може набувати тільки два значення: $x_1 = 1$ (подія відбулась) з ймовірністю p і $x_2 = 0$ (подія не відбулась) з ймовірністю $q = 1 - p$.

Шукане математичне сподівання:

$$M(X)=1\cdot p+0\cdot q=p.$$

Отже, математичне сподівання кількості появи події в одному випробуванні дорівнює імовірності цієї події.

Ймовірнісний зміст математичного сподівання полягає в тому, що математичне сподівання приблизно дорівнює середньому арифметичному значень випадкової величини.

Нехай здійснено n випробувань, в яких випадкова величина X набуває m_1 разів значення x_1, m_2 разів значення x_2, ..., m_k разів значення x_k, причому $m_1 + m_2 + ... + m_k = n$. Відповідно сума всіх значень, набутих X, дорівнює:

$$x_1\cdot m_1+x_2\cdot m_2+...+x_k\cdot m_k.$$

Знайдемо середнє арифметичне значень, набутих випадковою величиною, для чого розділимо знайдену суму на кількість випробувань:

$$\overline{X}=\frac{(x_1\cdot m_1+x_2\cdot m_2+...+x_k\cdot m_k)}{n}, \quad \text{або} \quad \overline{X}=x_1\left(\frac{m_1}{n}\right)+x_2\left(\frac{m_2}{n}\right)+...+x_k\left(\frac{m_k}{n}\right). \qquad (*)$$

Зауважимо, що відношення m_1/n – це відносна частота W_1 значення x_1, m_2/n – це відносна частота W_2 значення x_2, і так далі. Запишемо відношення (*) наступним чином:

$$\overline{X}=x_1\cdot W_1+x_2\cdot W_2+...+x_k\cdot W_k. \qquad (**)$$

Припустимо, що кількість випробувань достатньо велике. Тоді відносна частота приблизно дорівнює імовірності появи події:

$$W_1 \simeq p_1, \quad W_2 \simeq p_2, \quad ..., \quad W_k \simeq p_k.$$

Замінивши у відношенні (**) відносні частоти відповідними ймовірностями, отримаємо:

$$\overline{X} \simeq x_1 \cdot p_1 + x_2 \cdot p_2 + ... + x_k \cdot p_k.$$

Перша частина отриманого виразу є $M(X)$. Отже,

$$\overline{X} \simeq M(X).$$

Властивості математичного сподівання:

1. Математичне сподівання сталої величини дорівнює сталій величині:

$$M(C) = C.$$

2. Сталу величину можна виносити за знак математичного сподівання:

$$M(C \cdot X) = C \cdot M(X).$$

3 Математичне сподівання добутку кількох незалежних випадкових величин дорівнює добутку їх математичних сподівань:

$$M(X \cdot Y) = M(X) \cdot M(Y).$$

4 Математичне сподівання суми двох випадкових величин дорівнює сумі їх математичних сподівань:

$$M(X + Y) = M(X) + M(Y).$$

5. Математичне сподівання кількості появи події в n незалежних випробуваннях дорівнює добутку кількості випробувань на ймовірність події у кожному випробуванні:

$$M(X) = n \cdot p.$$

Задача 1.48. Знайдіть математичне сподівання суми кількості очок, які можуть випасти при киданні двох гральних кубиків.

Розв'язок

Нехай X – це кількість очок, що можуть випасти на першому кубику, Y – на другому. Вони набуватимуть однакових значень 1, 2, 3, 4, 5, 6 з імовірностями p = 1/6.

$$M(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = \frac{7}{2},$$

$$M(Y) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = \frac{7}{2}.$$

Математичне сподівання суми кількості очок дорівнює:

$$M(X + Y) = M(X) + M(Y) = \frac{7}{2} + \frac{7}{2} = \frac{14}{2} = 7.$$

Задача 1.49. Ймовірність влучання в мішень при одному пострілі дорівнює p = 0,7. Знайдіть математичне сподівання кількості попадань при десяти пострілах.

Розв'язок

Попадання в мішень при кожному пострілі не залежить від результату інших пострілів, тому вказані події незалежні й відповідно шукане математичне сподівання дорівнює:

$$M(X)=n\cdot p=10\cdot 0{,}7=7.$$

Математичне сподівання не дає повної картини характеристики випадкової величини, бо для різних розподілів математичні сподівання можуть бути однакові.

Для розрахунку відхилення випадкової величини від її математичного сподівання використовують таку числову характеристику, як дисперсія.

Дисперсією дискретної випадкової величини називають математичне сподівання квадрата відхилення випадкової величини від її математичного сподівання та позначають через $D(X)$:

$$D(X)=M[X-M(X)]^2. \qquad (1.51)$$

Для розрахунку дисперсії застосовують формулу, яку отримують, розписавши (1.51).

$$D(X)=M[X-M(X)]^2 =$$

$$= M[X^2-2\cdot X\cdot M(X)+M^2(X)] = M(X^2)-2\cdot M(X)\cdot M(X)+M^2(X) =$$

$$= M(X^2)-2\cdot M^2(X)+M^2(X) = M(X^2)-M^2(X).$$

Отже,

$$D(X)=M(X^2)-[M(X)]^2. \qquad (1.52)$$

Задача 1.50. Знайти дисперсію випадкової величини, яка має такий закон розподілу:

X	2	3	5
P	0,1	0,6	0,3

Розв'язок

Знайдемо математичне сподівання $M(X)$:

$$M(X)=2\cdot 0{,}1+3\cdot 0{,}6+5\cdot 0{,}3=3{,}5.$$

$$M(X^2)=2^2\cdot 0{,}1+3^2\cdot 0{,}6+5^2\cdot 0{,}3=13{,}3.$$

Дисперсію шукаємо за формулою (1.52):

$$D(X)=M(X^2)-[M(X)]^2=13{,}3-3{,}5^2=1{,}05.$$

Властивості дисперсії:

1. Дисперсія сталої величини C дорівнює нулю:

$$D(C)=0.$$

2 Сталий множник можна виносити за знак дисперсії, піднісши його до квадратуа

$$D(C\cdot X)=C^2\cdot D(X).$$

3. Дисперсія суми двох незалежних випадкових величин дорівнює сумі дисперсії цих величин:

$$D(X+Y)=D(X)+D(Y).$$

4. Дисперсія різниці двох незалежних випадкових величин дорівнює сумі дисперсій:

$$D(X-Y)=D(X)+D(Y).$$

5. Дисперсія кількості появи події A в n незалежних випробуваннях, в кожному з яких ймовірність p появи події стала, дорівнює добутку кількості випробувань на імовірність появи та непояви події в одному випробуванні:

$$D(X)=n\cdot p\cdot q.$$

6. Дисперсія добутку двох незалежних випадкових величин X, Y дорівнює:

$$D(X\cdot Y)=D(X)\cdot D(Y)+n^2\cdot D(X)+m^2\cdot D(Y),$$

де

$$m=M(X), \qquad n=M(Y).$$

Задача 1.51. Кубик підкидається 10 разів. Знайдіть математичне сподівання та дисперсію появи 6 очок на грані, що випала.

Розв'язок

З умови задачі маємо n = 10, p = 1/6, q = 1 - p = 1 - 1/6 = 5/6.

Щоби знайти математичне сподівання, скористаємося п'ятою властивістю математичного сподівання:

$$M(X)=n\cdot p=10\cdot\frac{1}{6}=\frac{5}{3}.$$

Для знаходження дисперсії, скористаємося її п'ятою властивістю:

$$D(X)=n\cdot p\cdot q=\frac{5}{3}\cdot\frac{5}{6}=\frac{25}{18}.$$

Середньоквадратичне відхилення дискретної випадкової величини X – це корінь квадратний її дисперсії:

$$\sigma(X)=\sqrt{D(X)}. \qquad (1.53)$$

Середньоквадратичне відхилення має ту саму розмірність, що й випадкова величина. Тому в тих випадках, коли оцінка розсіювання повинна мати ту ж розмірність, що як і випадкова величина, замість дисперсії. розглядають таку величину, як середньоквадратичне відхилення.

Властивості середньоквадратичного відхилення:

1. Середньоквадратичне відхилення сталої величини С дорівнює нулю:

$$\sigma(C)=0.$$

2. Сталий множник можна виносити за знак середнього квадратичного відхилення, попередньо взявши по модулю:

$$\sigma(C\cdot X)=|C|\cdot\sigma(X).$$

3. Середньоквадратичне відхилення суми n незалежних випадкових величин дорівнює:

$$\sigma(X)=\sqrt{\sigma^2(X_1)+\sigma^2(X_2)+...+\sigma^2(X_n)}.$$

Мода Mo(X) *дискретної випадкової величини* – це, найімовірніше, її значення в деякому околі цього значення.

Розподіл називається *унімодальним*, якщо він має єдину моду (наприклад, біноміальний, нормальний, показниковий тощо).

Розподіл називається *полімодальним*, якщо він має більше однієї моди (наприклад, рівномірний тощо).

Медіаною Me(*X*) *дискретної випадкової величини* називається таке число, для якого справедлива умова

$$P(X<Me(X))=P(X>Me(X)).$$

Медіана неперервної випадкової величини *X* завжди існує, а якщо *X* має дискретний розподіл, то медіани може і не бути.

Задача 1.52. Випадкова величина *X* має ряд розподілу, який поданий у таблиці. Знайти числові характеристики випадкової величини *X*.

X	1	2	3
P	0,2	0,6	0,2

Розв'язок

Математичне сподівання випадкової величини *X* розрахуємо за формулою (1.50):

$$M(X)=\sum_{i=1}^{3} x_i\cdot p_i=1\cdot 0{,}2+2\cdot 0{,}6+3\cdot 0{,}2=2.$$

Дисперсію випадкової величини *X* знайдемо за формулою (1.52):

$$D(X)=M(X^2)-[M(X)]^2=1^2\cdot 0{,}2+2^2\cdot 0{,}6+3^2\cdot 0{,}2-2^2=4{,}4-4=0{,}4.$$

Середньоквадратичне відхилення випадкової величини X вирахуємо за формулою (1.53):

$$\sigma(X)=\sqrt{D(X)}=\sqrt{0{,}4}\approx 0{,}633.$$

Мода і медіана дорівнює:

$$Mo(X)=Me(X)=2.$$

Наведемо числові характеристики основних законів розподілу дискретних випадкових величин.

1. Рівномірний розподіл із параметром *n*:

$$M(X)=\frac{n+1}{2},\quad D(X)=\frac{n^2-1}{12},\quad \sigma(X)=\sqrt{\frac{n^2-1}{12}}.$$

2. Біномінальний розподіл з параметром *n* і *p*:

$$M(X)=n\cdot p,\quad D(X)=n\cdot p\cdot q,\quad \sigma(X)=\sqrt{n\cdot p\cdot q}.$$

3. Показниковий із параметром λ:

$$M(X)=\lambda,\quad D(X)=\lambda,\quad \sigma(X)=\sqrt{\lambda}.$$

4. Геометричний із параметром *p*:

$$M(X)=\frac{1}{p},\quad D(X)=\frac{q}{p^2},\quad \sigma(X)=\frac{\sqrt{q}}{p}.$$

Задача 1.53. Імовірність попадання у мішень, що рухається, з рогатки дорівнює 0,001. Відбувається 2000 пострілів. Знайти числові характеристики випадкової величини *X*, яка дорівнює кількості попадань у мішень.

Розв'язок

Випадкова величини X має показниковий розподіл з параметром λ, який дорівнює $\lambda = n \cdot p$. Згідно умови задачі, n = 2000, p = 0,001, тому

$$\lambda = 2000 \cdot 0{,}001 = 2.$$

Зкориставшсь формулами для знаходження числових характеристик основних законів розподілу дискретних випадкових величин при показниковому розподілі отримаємо:

$$M(X)=2, \quad D(X)=2, \quad \sigma(X)=\sqrt{2}\approx 1{,}41.$$

Тема 5.2. Числові характеристики середнього арифметичного однаково розподілених взаємно незалежних випадкових величин

Нехай, X_1, X_2, ..., X_n – це взаємно незалежні випадкові величини, які мають однаковий розподіл, а отже, однакові числові характеристики (математичне сподівання, дисперсію).

Середнє арифметичне цих величин:

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Математичне сподівання середнього арифметичного однаково розподілених взаємонезалежних випадкових величин дорівнює математичному сподіванню кожної з цих величин:

$$M(\overline{X}) = a.$$

Дисперсія середнього арифметичного однаково розподілених взаємно незалежних випадкових величин в n-разів менша від дисперсії D – кожної із цих величин:

$$D(\overline{X}) = \frac{D}{n}.$$

Середньоквадратичне відхилення середнього арифметичного n однаково розподілених взаємонезалежних випадкових величин в $\sqrt{n}$ раз менше середнього квадратичного відхилення кожної із цих величин:

$$\sigma(\overline{X}) = \frac{\sigma}{\sqrt{n}}.$$

Наприклад, якщо середньоквадратичне відхилення окремого виміру рівне σ = 4 м, а всього зроблено n = 16 вимірів, то середньоквадратичне відхилення середнього арифметичного тих вимірів буде дорівнює лише 1 м.

Середнє арифметичне вимірів виявилось більш близьким до істинного значення величини, яку вимірюють, ніж результат окремого виміру.

При $n \to \infty$ математичне сподівання середнього арифметичного n незалежних однаково розподілених випадкових величин не змінюється і дорівнює середньому значенню випадкових величин, а дисперсія прямує до нуля. Це означає, що середнє арифметичне значення випадкових величин при $n \to \infty$ перестає бути випадковою величиною і прямує до сталого значення.

Для встановлення певної ознаки деякої величини проводять n вимірювань $X_1, X_2, ..., X_n$, які залежать від багатьох випадкових факторів, мають одинаків розподілі та є взаємонезалежними.

За характеристику досліджуваної ознаки доцільніше обирати середнє арифметичне значення отриманих результатів вимірювання: $X_1, X_2, ..., X_n$, оскільки розсіювання (дисперсія) їх середнього арифметичного значення є меншим, ніж розсіювання кожного результату вимірювання, причому воно прямує до нуля при $n \to \infty$.

Значення середнього арифметичного результатів вимірювань ознаки випадкової величини більш надійніше і ближче до справжньої характеристики ознаки, ніж окремий результат.

Тема 5.3. Числові характеристики неперервних випадкових величин

Математичним сподіванням неперервної випадкової величини X, можливі значення якої належать проміжку $[a, b]$, називають визначений інтеграл:

$$M(X)=\int_a^b x\cdot f(x)dx. \tag{1.54}$$

Якщо можливі значення випадкової величини X належать всій дійсній осі, $x\in(-\infty,+\infty)$ то

$$M(X)=\int_{-\infty}^{+\infty} x\cdot f(x)dx. \tag{1.55}$$

Дисперсія неперервної випадкової величини X, якщо можливі значення належать інтервалу $[a, b]$, визначається за формулою

$$D(X)=\int_a^b [x-M(X)]^2\cdot f(x)dx, \tag{1.56}$$

де

$$M(X)=\int_a^b x\cdot f(x)dx.$$

Якщо можливі значення випадкової величини X належать всій дійсній осі, $x\in(-\infty,+\infty)$, то дисперсія визначається за формулою

$$D(X)=\int_{-\infty}^{+\infty} [X-M(x)]^2\cdot f(x)dx, \tag{1.57}$$

де

$$M(X)=\int_{-\infty}^{+\infty} x\cdot f(x)dx.$$

Формули (1.56) та (1.57), як і у випадку дискретної випадкової величини, можна записати у вигляді

$$D(X)=\int_a^b x^2\cdot f(x)dx-[M(X)]^2, \quad \text{або} \quad D(X)=\int_{-\infty}^{+\infty} x^2\cdot f(x)dx-[M(X)]^2. \tag{1.58}$$

Середньоквадратичне відхилення, як і для дискретної випадкової величини, визначається за тією ж формулою:

$$\sigma(X)=\sqrt{D(X)}.$$

Задача 1.54. Знайдіть математичне сподівання, дисперсію та середньоквадратичне відхилення випадкової величини *X*, якщо

$$f(x)=\begin{cases} 0\; при\; x\le 0 \\ 2x\; при\; 0<x\le 1 \\ 0\; при\; x>1 \end{cases}$$

Розв'язок

Математичне сподівання знаходимо за формулою (1.54):

$$M(X)=\int_a^b x\cdot f(x)dx.$$

Підставивши значення, одержимо:

$$M(X)=\int_0^1 x\cdot 2x\,dx = 2\int_0^1 x^2 dx = 2\frac{x^3}{3}\Big|_0^1 = \frac{2}{3}.$$

Дисперсію визначаємо згідно з формулою (1.58). Підставивши значення, одержимо:

$$D(X)=\int_0^1 x^2\cdot 2x dx-\left(\frac{2}{3}\right)^2=\frac{1}{18}. \qquad \sigma(X)=\sqrt{\frac{1}{18}}.$$

Тема 5.4. Числові характеристики нормального закону розподілу

Нехай неперервна випадкова величина *X* розподілена за нормальним законом:

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

За означенням математичного сподівання для неперервної випадкової величини маємо:

$$M(X)=\int_{-\infty}^{+\infty} x\cdot f(x)dx = \int_{-\infty}^{+\infty} x\cdot\frac{1}{\sigma\sqrt{2\pi}}\cdot e^{\frac{-(x-a)^2}{2\sigma^2}}dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} x\cdot e^{\frac{-(x-a)^2}{2\sigma^2}}dx =$$

Узявши до уваги, що $z = (x\text{-}a)/\sigma$, звідки $x = \sigma z + a$, $dx = \sigma dz$,

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} x\cdot e^{\frac{-(x-a)^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma z e^{\frac{-z^2}{2}}dz+\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{\frac{-z^2}{2}}dz = a.$$

Під знаком інтегралу непарна функція, тобто перший інтеграл суми, дорівнює 0, а інтеграл Пуассона, тобто другий інтеграл у цій сумі, доірвнює $\sqrt{2\pi}$.

Отже, *математичне сподівання нормального закону розподілу* дорівнює параметру *a* диференціальної функції розподілу:

$$M(X)=a. \qquad (1.59)$$

Згідно з означенням дисперсії неперервної випадкової величини, з урахуванням того, що $M(X) = a$, маємо:

$$D(X)=\int_{-\infty}^{+\infty} x^2 \cdot f(x)\,dx-[M(X)]^2 = \int_{-\infty}^{+\infty} x^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(x-a)^2}{2\sigma^2}}\,dx-(a)^2 =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-a)^2 \cdot e^{\frac{-(x-a)^2}{2\sigma^2}}\,dx.$$

Інтеграл шукається за допомогою заміни $z = (x\text{-}a)/\sigma$, звідки $x = \sigma z + a$, $dx = \sigma dz$.

$$D(X)=\sigma^2, \qquad \sigma=\sqrt{D(X)}. \tag{1.60}$$

Імовірність відхилення випадкової величини, розподіленої за нормальним законом, від свого математичного сподівання знаходиться за формулою

$$P(|X-a|<\alpha)=2\Phi\left(\frac{\alpha}{\sigma}\right). \tag{1.61}$$

Тема 5.5. Числові характеристики показникового закону розподілу

Нехай неперервна випадкова величина розподілена за показниковим законом:

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

де λ – стала величина, $\lambda > 0$.

Використовуючи формулу (1.55), одержимо:

$$M(X)=\int_0^{+\infty} x \cdot f(x)\,dx = \int_0^{+\infty} x \cdot \lambda e^{-\lambda x}\,dx = \lambda \int_0^{+\infty} x \cdot e^{-\lambda x}\,dx.$$

Інтегруючи частинами,

$$M(X)=\frac{1}{\lambda}. \tag{1.62}$$

Отже, математичне сподівання показникового розподілу дорівнює оберненій величині параметра λ.

Використовуючи формулу (1.58) для розрахунку дисперсії, матимемо

$$D(X)=\int_0^{+\infty} x^2 \cdot f(x)\,dx-[M(X)]^2 = \lambda \int_0^{+\infty} x^2 \cdot e^{-\lambda x}\,dx-\frac{1}{\lambda^2}.$$

Інтегруючи частинами, отримаємо:

$$D(X)=\frac{1}{\lambda^2}. \tag{1.63}$$

Звідси середньоквадратичне відхилення, як корінь квадратний дисперсії, дорівнює:

$$\sigma(X)=\sqrt{D(X)}=\frac{1}{\lambda}. \tag{1.64}$$

Відтак математичне сподівання та середньоквадратичне відхилення показникового закону розподілу рівні між собою:

$$M(X)=\sigma(X)=\frac{1}{\lambda}.$$

Задача 1.55. Знайдіть числові характеристики випадкової величини, розподіленої згідно із законом:

$$f(x)=\begin{cases} 0 \, при \, x<0\,; \\ \dfrac{2}{e^{2x}} при \, x\ge 0. \end{cases}$$

Розв'язок

Запишемо випадкову величину X у формі

$$f(x)=\begin{cases} 0 \, при \, x<0\,; \\ 2e^{-2x} при \, x\ge 0. \end{cases}$$

Отримаємо випадкову величину, розподілену за показниковим законом з параметром $\lambda = 2$. Відповідно до формул (1.62 – 1.64), маємо:

$$M(X)=\frac{1}{\lambda}=\frac{1}{2}, \quad D(X)=\frac{1}{\lambda^2}=\frac{1}{4}, \quad \sigma(X)=\sqrt{\frac{1}{\lambda^2}}=\frac{1}{2.}$$

Тема 5.6. Кореляційний момент та коефіцієнт кореляції

Для характеристики зв'язку між випадковими величинами, наприклад X і Y, використовується так званий кореляційний момент.

Кореляційний момент випадкових величин X і Y – це математичне сподівання добутку відхилень цих величин від своїх математичних сподівань:

$$\mu_{x,y}=M\{[X-M(X)]\cdot[Y-M(Y)]\}. \qquad (1.65)$$

Для розрахунку кореляційних моментів дискретних випадкових величин використовують формулу:

$$\mu_{x,y}=\sum_{i=1}^{n}\sum_{j=1}^{m}[x_i-M(X)]\cdot[y_j-M(Y)]\cdot p(x_i,y_j), \qquad (1.66)$$

а для неперервної випадкової величини –

$$\mu_{x,y}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-M(x)]\cdot[y-M(Y)]\cdot f(x,y). \qquad (1.67)$$

Із означення кореляційного моменту випливає, що він має розмірність, що дорівнює добутку розмірностей випадкових величин X і Y. Щоби мати однакову розмірність, або безрозмірну величину, яка характеризує випадкові величини X і Y, вводять поняття коефіцієнта кореляції.

Коефіцієнт кореляції випадкових величин X і Y ($r_{x,y}$) – це відношення кореляційного моменту до добутку середніх квадратичних відхилень цих величин:

$$r_{x,y}=\frac{\mu_{x,y}}{\sigma_x\sigma_y}. \qquad (1.68)$$

Властивості коефіцієнта кореляції:

1. $|r_{x,y}| \le 1$;
2. Якщо X та Y незалежні, то $r_{x,y} = 0$;
3. Якщо між X та Y є лінійна залежність $Y = aX + b$, де a та b – постійні, то $r_{x,y} = 1$.

Дві випадкові величини X і Y називаються *корельованими*, якщо їх кореляційний момент (або те саме – коефіцієнт кореляції) відмінний від нуля. Дві випадкові величини X і Y називаються *некорельованими* величинами, якщо їх кореляційний момент дорівнює нулю.

Зверніть увагу! Із означення корельованості величин X і Y випливає їх залежність, але із їх залежності ще не випливає їх корельованість. Із незалежності двох випадкових величин випливає їх некорильованість, проте в разі їх некорильованості ще не можна говорити про незалежність цих величин.

Слід зазначити, що з некореляваності нормально розподілених величин випливає їх незалежність.

Лекція 6. Закон великих чисел. Центральна гранична теорема

Тема 6.1. Нерівність Чебишова

Теорія ймовірностей вивчає закономірності масових випадкових подій (явищ). Закономірності можна виявити при великій кількості випадкових явищ, що відбуваються за однакових умов.

Тобто характеристики випадкових подій і випадкових величин у цих умовах стають *стійкими*: середній їх результат (частота події, середні значення випадкових величин) перестає бути випадковою величиною і може бути передбачений з великою імовірністю.

В понятті стійкості середніх величин і закладений фізичний зміст «*закону великих чисел*»: за дуже великої кількості випадкових експериментів середній їх результат перестає бути випадковим і може бути передбаченим практично достовірна. У теорії ймовірностей під законом великих чисел розуміють певну групу математичних теорем, у кожній з яких за тих чи інших умов визначається факт наближення середніх характеристик великої кількості випадкових експериментів до невипадкових сталих.

Закон великих чисел – це зв’язок між теорією ймовірностей як математичною наукою та закономірностями випадкових явищ, у випадку масових спостережень над ними.

Другу групу математичних (граничних) теорем об’єднують під назвою – центральна гранична терема. По суті, центральна гранична теорема визначає умови виникнення нормального розподілу випадкової величини. На основі тверджень центральної граничної теореми випливає, що нормальний розподіл виникає тоді, коли відбувається підсумовування багатьох незалежних випадкових величин, порівняльних щодо порядку свого впливу на розсіювання суми.

Різні форми закону великих чисел і центральної граничної теореми різняться умовами, які накладаються на випадкові величини та їх суми.

Під час розв’язування задач та при доведенні граничних теорем зокрема важливу роль виконує нерівність Чебишова.

Розглянемо закон розподілу дискретної величини X:

X	x_1	x_2	x_3	...	x_n
P	p_1	p_2	p_3	...	p_n

Знайдемо ймовірність відхилення випадкової величини від його математичного сподівання.

Перша нерівність Чебишова: для довільної випадкової величини, яка набуває додатні значення та має скінченне математичне сподівання, справедлива нерівність:

$$P(X \ge 1) \le M(X). \tag{1.69}$$

Якщо X – дискретна випадкова величина, то

$$P(X \ge 1) = \sum_i p(x_i) \le \sum_i x_i \cdot p(x_i) = M(X).$$

Якщо X – неперервна випадкова величина, $f(x)$ – її густина ймовірності, то

$$P(X \ge 1) = \int_1^\infty f(x)dx \le \int_1^\infty x \cdot f(x)dx \le \int_0^\infty x \cdot f(x)dx = M(X).$$

Якщо X набуває лише невід'ємні значення

$$M(X) < \infty, \qquad \alpha > 0,$$

то

$$P(X \ge \alpha) \le \frac{M(X)}{\alpha}. \tag{1.70}$$

Нерівність (1.70) називають нерівністю Маркова.

Друга нерівність Чебишова: ймовірність того, що відхилення випадкової величини від його математичного сподівання (за абсолютним значенням менше деякого додатного числа) задовольняє нерівність

$$P(|X - M(X)| < \varepsilon) \ge 1 - \frac{D(X)}{\varepsilon^2}. \tag{1.71}$$

Задача 1.56. Випадкова величина X має математичне сподівання $M(X) = 1$ і середньоквадратичне відхилення $\sigma = 0{,}2$. Скориставшись нерівністю Чебишова, оцінити ймовірності таких подій:

A) $A = (0{,}5 < X < 1{,}5)$;

B) $B = (0{,}75 < X < 1{,}35)$;

C) $C = (X < 2)$.

Розв'язок

A) використовуючи формулу (1.71), одержимо:

$$P(A) = P(0{,}5 < X < 1{,}5) \;=\; P(|X - 1| < 0{,}5) \;=$$

$$= \; 1 - P(|X - 1| \ge 0{,}5) \;\ge\; 1 - \frac{0{,}2^2}{0{,}5^2} = 0{,}84; \quad P(A) \ge 0{,}84.$$

B) використаємо те, що звуження інтервалу не може призвести до збільшення ймовірності:

$$P(B) = P(0{,}75 < X < 1{,}35) \;\ge\; P(0{,}75 < X < 1{,}25) \;=$$

$$= \quad P(|X-1|<0{,}25) \;=\; 1-P(|X-1|\geq 0{,}25) \;\geq\; 1-\frac{0{,}2^2}{0{,}25^2} = 0{,}36$$

$$P(B)\geq 0{,}36.$$

C) $P(C)=P(X<2) \;=\; P(|X-1|<1) \;=\; 1-P(|X-1|\geq 1) \;\geq\; 1-\frac{0{,}2^2}{1^2} = 0{,}96$

$$P(C)\geq 0{,}96.$$

Інформація про розподіл випадкової величини дає змогу уточнити відповідну оцінку ймовірності на основі нерівності Чебишова.

Задача 1.57. Дискретна випадкова величина задана законом розподілу.

Знайдіть ймовірність того, що $|X - M(X)| < 3$ та оцініть її,

X	-2	-1	2	3	4
P	0,1	0,2	0,3	0,15	0,25

використовуючи нерівністю Чебишова.

Розв'язок

Знайдемо математичне сподівання та дисперсію:

$$M(X)=(-2)\cdot 0{,}1+(-1)\cdot 0{,}2+2\cdot 0{,}3+3\cdot 0{,}15+4\cdot 0{,}25=1{,}65.$$

$$D(X)=(-2)^2\cdot 0{,}1+(-1)^2\cdot 0{,}2+2^2\cdot 0{,}3+3^2\cdot 0{,}15+4^2\cdot 0{,}25-1{,}65^2=4{,}4275.$$

Тоді нерівність $|X - M(X)| < 3$ розпишеться:

$$|X-1{,}65|<3 \Rightarrow -3+1{,}65<X<3+1{,}65 \Rightarrow -1{,}35<X<4{,}65.$$

Оцінимо цю ймовірність за допомогою нерівності Чебишова:

$$P(|X-1{,}65|<3)\geq 1-\frac{4{,}4275}{3^2}\approx 0{,}51.$$

Якщо $|X - M(X)| \geq \varepsilon$, це еквівалентне $|X - M(X)|^2 \geq \varepsilon^2$. Застосувавши першу нерівність Чебишова,

$$P(|X-M(X)|\geq \varepsilon) \;=\; P\left(\frac{1}{\varepsilon^2}|X-M(X)|^2\geq 1\right)\leq \frac{M(X-M(X)^2)}{\varepsilon^2},$$

$$P(|X-M(X)|\geq \varepsilon)\leq \frac{D(X)}{\varepsilon^2}. \qquad (1.72)$$

Задача 1.58. Дисперсія випадкової величини X рівна 0,001. Знайдіть імовірність того, що випадкова величина X відрізняється від її математичного сподівання більше ніж на 0,1.

Розв'язок

Для знаходження ймовірності використаємо нерівність (1.72)

$$P(|X-M(X)|\geq 0{,}1)\leq \frac{0{,}001}{0{,}1^2}=0{,}01.$$

Тема 6.2. Теорема Чебишова (закон великих чисел)

Теорема: якщо X_1, X_2, …, X_n – попарно незалежні випадкові величини, причому дисперсії їх рівномірно обмежені, $D(X_1) \le C$, $D(X_2) \le C$, …, $D(X_n) \le C$, де C = const, то для довільного $\varepsilon > 0$ справедлива рівність:

$$\lim_{n\to\infty} P\left(\left|\overline{X} - M\left(\overline{X}\right)\right| < \varepsilon\right) = 1, \tag{1.73}$$

де

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Зміст теореми Чебишова (так званого закону великих чисел) полягає в тому, що середньоарифметичні значення доволі великої кількості незалежних випадкових величин (з рівномірно обмеженими дисперсіями) перестають бути випадковими.

Це пояснюється тим, що при вимірюванні чи проведенні експерименту за істинну величину беруть середнє значення цих величин.

Середнє арифметичне значення великої кількості незалежних випадкових величин втрачає випадковий характер і набуває *властивість стійкості*.

На теоремі Чебишова ґрунтується вибірковий метод, який часто використовують у статистиці. Зміст методу полягає в тому, що на підставі вивчення певної ознаки для достатньо великої вибірки об’єктів (випадкової) роблять висновок стосовно генеральної сукупності.

Тема 6.3. Теорема Бернуллі та теорема Пуассона (граничні теореми)

Теореми Чебишова – це, по суті, найпростіше доведення наступної граничної теореми Бернуллі.

Теорема Бернуллі: якщо в кожному із n незалежних випробувань ймовірність p виникнення події A стала, то як завгодно близькою до одиниці буде ймовірність того, що відхилення відносної частоти m/n (m – число появ події A в n випробуваннях) від ймовірності p буде як завгодно малим, якщо кількість випробувань досить велика, тобто:

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1. \tag{1.74}$$

Англійський статистик Карл Пірсон підкинув монету 12000 разів і при цьому герб випав 6019 разів, тобто частота випадання герба 0,5016. Іншого разу він підкинув монету 24000 разів і 12012 разів спостерігав випадіння герба. Частота випадання герба 0,5005.

Зверніть увагу! Правильність теореми Бернуллі випливає з нерівності Чебишова (1.71), яка застосовується до випадкової величини, тобто:

$$P\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) = 1 - \frac{p \cdot q}{n \cdot \varepsilon^2} \tag{1.75}$$

Задача 1.59. Ймовірність того, що автомат, який продає квитки, при опусканні монети працюватиме безпомилково, дорівнює 0,96. Визначте за допомогою теореми Бернуллі, скільки потрібно зробити випробувань

правильності роботи автомату, щоб із імовірністю не меншою ніж 0,99 можна було стверджувати, що відхилення частоти правильної роботи автомата від імовірності правильної роботи за абсолютним значенням не перевищуватиме 0,03.

Розв'язок

Використаємо теорему Бернуллі:

$$P\left(\left|\frac{m}{n}-p\right|\leq\varepsilon\right)=1-\frac{p\cdot q}{n\cdot\varepsilon^2}$$

Згідно з умовою, $\varepsilon = 0{,}03$, $p = 0{,}96$. Отже, $q = 1 - p = 1 - 0{,}96 = 0{,}04$. Вимога задачі $P \geq 0{,}99$ виконуватиметься тоді, коли

$$1-\frac{p\cdot q}{n\cdot\varepsilon^2}\geq 0{,}99.$$

Або при

$$\frac{p\cdot q}{n\cdot\varepsilon^2}\leq 0{,}01.$$

Звідси:

$$\frac{p\cdot q}{0{,}01\cdot\varepsilon^2}\leq n\Rightarrow n\geq\frac{p\cdot q}{0{,}01\cdot\varepsilon^2},\quad n\geq\frac{0{,}96\cdot 0{,}04}{0{,}01\cdot 0{,}03^2}=\frac{0{,}0384}{0{,}000009}\approx 42\,67.$$

n може бути лише цілим числом, заокруглення треба виконувати в подібних випадках лише у бік збільшення n.

Отже, $n \geq 4267$. При таких значеннях n ймовірність P буде не менша за 0,99. Узагальненням теореми Бернуллі є теорема Пуассона, тобто в цьому випадку ймовірності подій у кожному випробуванні різні.

Теорема Пуассона: частота події A в n незалежних випробуваннях, в кожному з яких вона може настати відповідно з ймовірностями p_1, p_2, … p_n, при необмеженому збільшенні кількості випробувань, збігається за ймовірністю до середнього арифметичного ймовірностей подій в окремих дослідах, тобто

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n}-\overline{p}\right|<\varepsilon\right)=1, \tag{1.76}$$

де

$$\overline{p}=\frac{p_1+p_2+\ldots+p_n}{n}.$$

Окрім законів великих чисел, описаних у вище розглянутих теоремах, спостерігається ще одне доволі цікаве явище, яке полягає в тому, що за великої кількості випадкових доданків, кожний з яких робить лише невеликий внесок у загальну суму, розподіл кожного з цих доданків не впливає на сумарний результат.

Тема 6.4. Центральна гранична теорема (теорема Ляпунова)

Центральну граничну теорему використовують, щоб дати відповідь на такі питання:

1. Коли сума багатьох випадкових величин не надто відрізняється від сталої величини, тобто перестає бути випадковою величиною, і тому її поведінка може прогнозуватись зі значною імовірністю.
2. За яких умов можна зі значною імовірністю прогнозувати кількість появ деякої випадкової події при великій кількості незалежних випробувань?
3. При яких обмеженнях сума багатьох випадкових величин стане розподіленою за нормальним законом?

Теорема: Нехай X_i – це незалежні однаково розподілені випадкові величини з математичним сподіванням: $M(X_i) = a$ та дисперсіями $D(X_i) = \sigma^2$. Тоді при великому n розподіл суми: $X_1 + X_2 +\ldots+ X_n= X$ стає близький до нормального розподілу.

При цьому:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - na}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{\alpha - na}{\sigma\sqrt{n}}\right). \qquad (1.77)$$

Отже,

$$P(|X - na| \le \delta) = P(na - \delta < X < na + \delta) \approx 2\Phi\left(\frac{\delta}{\sigma\sqrt{n}}\right). \qquad (1.78)$$

З останніх рівностей випливає те, що ймовірність p події можна оцінити, використовуючи експериментальні результати.

Дійсно, якщо в кожному випробуванні подія A відбувається з імовірністю p і в серії з n незалежних випробувань досліджувана подія A відбувається k разів, тоді статистична частота появи події A дорівнювати k/n.

Ця частота – випадкова величина, яка в свою чергу є сумою незалежних випадкових величин X_k, які набувають значень $1/n$ з ймовірністю p, коли подія A відбувається у k-му випробуванні, і 0 з ймовірністю $q = 1 - p$ в протилежному випадку, при цьому $k = 1, 2, \ldots, n$.

Для таких випадкових величин X_k математичне сподівання та дисперсія визначаються за формулами:

$$M(X_k) = \frac{p}{n} + 0\cdot(1-p) = \frac{p}{n}, \qquad \sigma^2 = D(X_k) = \frac{p\cdot q}{n^2}, \qquad k = 1, 2, 3, \ldots, n.$$

Тоді при досить великих значеннях n ($n \ge 30$) виконується рівність

$$P\left(\left|\frac{k}{n} - p\right| \le \delta\right) = 2\Phi\left(\frac{\delta}{\sigma\sqrt{n}}\right) = 2\Phi\left(\delta\sqrt{\frac{n}{p\cdot q}}\right). \qquad (1.79)$$

Наслідок: при $n \ge 30$ розподіл суми однаково розподілених випадкових величин мало відрізняється від нормального розподілу.

Задача 1.60. Імовірність того, що деталь неякісна, дорівнює 0,1. Знайти ймовірність того, що серед виготовлених 400 деталей відносна частота появи неякісної деталі відхиляється від імовірності 0,1 не більше ніж на 0,03.

Розв’язок

Згідно з умовою задачі, $n = 400$, $p = 0{,}1$, $q = 1 - p = 0{,}9$, $\delta = 0{,}03$. Для знаходження ймовірності застосуємо формулу (1.79):

$$P\left(\left|\frac{k}{400} - 0{,}1\right| \le 0{,}03\right) = 2\,\Phi\left(0{,}03 \cdot \sqrt{\frac{400}{0{,}1 \cdot 0{,}9}}\right).$$

$$P\left(\left|\frac{k}{400} - 0{,}1\right| \le 0{,}03\right) = 2\,\Phi(2).$$

За допомогою таблиці значень функції Лапласа знаходимо необхідне нам значення $\Phi(2) = 0{,}4772$.

$$P\left(\left|\frac{k}{400} - 0{,}1\right| \le 0{,}03\right) = 2 \cdot 0{,}4772 = 0{,}9544.$$

Контрольні запитання до розділу 1

1. Що таке «випадкова подія»? Наведіть класифікацію випадкових подій.
2. Напишіть основні формули комбінаторики. Опишіть їх.
3. В чому зміст класичного, статичного та геометричного означення ймовірності?
4. Розкрийте зміст теореми додавання ймовірностей несумісних подій.
5. Який зміст теореми множення залежних і незалежних подій?
6. Поясніть зміст теореми додавання ймовірностей сумісних подій.
7. Що таке «протилежні події»? Напишіть формулу ймовірність протилежної події, ймовірності появи хоча б однієї події.
8. Напишіть формулу повної ймовірності та формулу Байєса. Опишіть їх.
9. Що ви розумієте під поняттям «послідовність незалежних випробувань»? Запишіть формулу Бернуллі.
10. Що таке «найімовірніше число настання події під час повторних випробувань»?
11. Розкрийте зміст локальної теореми Муавра-Лапласа.
12. Розкрийте зміст інтегральної теорема Лапласа.
13. Що таке «інтегральна функція розподілу неперервної випадкової величини». Опишіть її властивості.
14. Що таке «диференціальна функція розподілу неперервної випадкової величини»? Які її властивості?
15. Розкрийте зміст нормального та показникового законів розподілу випадкової величини.
16. Які ви знаєте числові характеристики дискретної та неперервної випадкової величини? Напишіть їх формули.
17. Що таке «кореляційний момент» та «коефіцієнт кореляції»?
18. Що таке «нерівність Чебишова»? Напишіть її формулу.
19. В чому полягає зміст теореми Чебишова (закону великих чисел)?
20. Розкрийте зміст теореми Бернуллі та теореми Пуассона (як граничних теорем).
21. Який зміст центральної граничної теореми?

РОЗДІЛ 2. МАТЕМАТИЧНА СТАТИСТИКА

Статистика (від італ. *stato* – держава) – це наука, яка вивчає кількісну сторону суспільних явищ і процесів у нерозривному зв'язку з їх якісним змістом.

Статистику поділяють на два типи: описову та пояснювальну. Прикладом описової статистики є книга рекордів Гіннеса. Пояснювальна статистика здійснює формулювання висновків про ті чи інші явища, робить прогноз.

Математична статистика – це розділ математики, в якому на основі дослідних (вихідних) даних вивчаються ймовірнісні закономірності масових явищ.

Основні *завдання математичної статистики* – це статистична перевірка гіпотез, оцінка розподілу статистичних імовірностей та його параметрів, вивчення статистичної залежності, визначення основних числових характеристик випадкових вибірок, якими є: вибіркове середнє, вибіркові дисперсії, стандартне відхилення.

Предмет математичної статистики полягає у розробці методів збору та обробки статистичних даних з метою отримання наукових та практичних висновків.

Математична статистика виникла (XVII ст.) та почала розвиватись одночасно з теорією ймовірностей. Подальшим розвитком (кінець XIX – початок XX ст.) математичної статистики займалися А. Марков, П. Чебишов, О. Ляпунов, а також Ф. Гальтон, К. Гаусс, К. Пірсон та інші.

У XX ст. найбільший внесок у математичну статистику зробили Е. Слуцький, В. Романовський, А. Колмогоров, Стьюдент (псевдонім У. Госсета), Ю. Нейман, Е. Пірсон, А. Вальд, А. Скороход, В. Королюк та інші вчені.

Лекція 7. Вибірковий метод

Тема 7.1. Генеральна та вибіркова сукупність

Основні поняття математичної статистики:

1. *Генеральна сукупність* – це вся сукупність об'єктів, які досліджуються.
2. *Вибіркова сукупність* (*вибірка*) – це об'єкти, довільно або випадково відібрані з генеральної сукупності для дослідження.

3 *Обсяг* (*об'єм*) *сукупності* – це кількість об'єктів цієї сукупності.

Наприклад, плоди дерева (200 шт.) обстежують на наявність специфічного смаку. Для цього відбирають 10 екземплярів. Тут 200 – це обсяг генеральної сукупності, 10 – обсяг вибірки.

Вибірка повинна бути представницькою (репрезентативною), тобто правильно відображати властивості генеральної сукупності, які вивчаються. Досліджувані явища мають бути масовими. Лише тоді статистичні дані будуть достовірними.

Вибірки бувають повторні та безповторні. *Повторною* називають вибірку, коли відібраний об’єкт повертається до генеральної сукупності перед відбором іншого об’єкта. Вибірка називається *безповторною*, якщо взятий об’єкт назад до генеральної сукупності не повертається. Найчастіше використовують безповторні вибірки.

План проведення статистичних досліджень:

1. Формулювання завдання дослідження та визначення обсягу вибірки (мета, об’єкти вивчення, їх кількість, певні ознаки, характеристики).
2. Збір необхідних даних, вибір доцільної форми їх подання для подальшого дослідження (водночас застосовують методи: спостереження, порівняння, усне та письмове анкетування).
3. Проводять остаточну обробку статистичного матеріалу та його вивчення.
4. За результатами формулюють певні висновки.

Зручною та наочною формою подання даних, одержаних у результаті статистичного дослідження, є таблиці, графіки, кругові діаграми тощо.

Тема 7.2. Способи відбору статистичного матеріалу

На практиці використовують різноманітні способи відбору об’єктів із генеральної сукупності.

Простий (*випадковий*) *відбір*. Навмання вибирається певна кількість об’єктів з усієї генеральної сукупності. Тобто, вибираємо m об’єктів із n об’єктів генеральної сукупності. Для цього, наприклад, нумеруємо картки від 1 до n на об’єктах і вибираємо по одній картці і так далі. Випадковий відбір може бути повторним або безповторним. Для великого об'єму вибірки використовують готові таблиці «випадкових чисел».

Типовий відбір. Він полягає у відборі об'єктів не з усієї сукупності, а з певної її типової частини. Наприклад, соціологи вивчають різні групи людей. Типовими ознаками можуть бути місце роботи, вік, професія...

Механічний відбір. Полягає в тому, що сукупність поділяють на стільки частин, скільки планується об’єктів у вибірку. З кожної частини вибирають один об’єкт. Наприклад, якщо обстежується 20% телевізорів з партії, то беруть кожен п’ятий, якщо 10% – кожен десятий. Щоб механічний відбір був репрезентативним, враховують специфіку технологічного процесу.

Серійний відбір. Об’єкти з генеральної сукупності вибирають серіями або партіями. Серійний відбір використовують тоді, коли ознака, яку досліджують, практично не змінюється в різних серіях. На практиці розглянуті способи відбору можуть поєднувати та комбінувати.

Тема 7.3. Статистичний розподіл вибірки

Статистичний розподіл вибірки

Об’єкти обстежують за їх певними характеристиками або ознаками (місце проживання, вік, професія, …), які називають *варіантами*.

Сукупність значень ознаки, записаних у порядку їх зростання, називають *варіаційним рядом*.

Нехай X – статистична змінна, x_i ($i = 1, 2, 3, \ldots n$) – її значення. Тоді

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

– варіаційний ряд, якщо

$$x_{i+1}' \geq x_i'.$$

Якщо кількість варіантів велика, то сукупність значень для зручності розбивають на інтервали або проміжки. Кожен інтервал характеризується одним числом – його серединою. Величину інтервалу ще називають інтервальною різницею. Частинні інтервали варіантів, які розміщені у зростаючий послідовності, називають *інтервальним варіаційним рядом*.

Важлива характеристика – частота появи цих варіантів.

Додатне число n_i, яке вказує на те, скільки разів варіант x_i ($i = 1, 2, 3, \ldots m$) трапляється в таблиці даних, називається *частотою варіанта*.

Ряд n_1, n_2, ..., n_m називається *рядом частот*. Сума усіх частот повинна дорівнювати обсягу вибірки:

$$\sum_{n=1}^{m} n_i = n. \tag{2.1}$$

Статистичний розподіл вибірки встановлює зв’язок між рядом варіанта, що зростає або спадає, і відповідними частотами. Він може бути зображений у вигляді таблиці:

X	x_1	x_2	…	x_m
n_i	n_1	n_2	…	n_m

Відношення частоти варіанта n_i до обсягу вибірки (n) називають *відносною частотою* та позначають

$$W_i = \frac{n_i}{n}. \tag{2.2}$$

Сума усіх відносних частот

$$\sum_{n=1}^{m} W_i = 1. \tag{2.3}$$

Залежність між значеннями варіаційного ряду і відповідними їм відносними частотами називають *статистичним розподілом відносних частот*:

X	x_1	x_2	x_3	…	x_m
W_i	n_1/n	n_2/n	n_3/n	…	n_m/n

Задача 2.1. Під час дослідження кількісної ознаки з генеральної сукупності було отримано вибірку:

4; 3; 6; 4; 7; 2; 5; 1; 2; 5; 4; 4; 3; 5; 6; 3; 4; 1; 3; 4.

Знайдіть: обсяг вибірки, варіаційний ряд, статистичний розподіл вибірки та статистичний розподіл відносних її частот.

Розв'язок

Оскільки вибірка містить 20 значень, то об'єм (обсяг) вибірки n = 20. Будуємо варіаційний ряд вибірки – записуємо всі її значення у порядку зростання:

1; 1; 2; 2; 3; 3; 3; 3; 4; 4; 4; 4; 4; 4; 5; 5; 5; 6; 6; 7.

У цій вибірці лише сім різних значень, тобто сім варіантів:

1; 2; 3; 4; 5; 6; 7.

Знаходимо їх частоти:

$n_1 = 2$; $n_2 = 2$; $n_3 = 4$; $n_4 = 6$; $n_5 = 3$; $n_6 = 2$; $n_7 = 1$.

Запишемо шуканий статистичний розподіл частот вибірки:

X	1	2	3	4	5	6	7
n_i	2	2	4	6	3	2	1

Знайдемо відносні частоти варіантів вибірки:

$$W_1=\frac{2}{20}=0{,}1; \quad W_2=\frac{2}{20}=0{,}1; \quad W_3=\frac{4}{20}=0{,}2; \quad W_4=\frac{6}{20}=0{,}3;$$

$$W_5=\frac{3}{20}=0{,}15; \quad W_6=\frac{2}{20}=0{,}1; \quad W_7=\frac{1}{20}=0{,}05.$$

Отже, шуканий розподіл відносних частот має такий вигляд:

X	1	2	3	4	5	6	7
W_i	0,1	0,1	0,2	0,3	0,15	0,1	0,05

Тема 7.4. Емпірична функція розподілу та її властивості

За даними статистичного розподілу вибірки будується емпірична функція розподілу.

Емпіричною функцією розподілу, або *функцією розподілу вибірки*, називають функцію $F^*(x)$, яка визначає для кожного значення x відносну частоту події $X < x$.

Математично це записується так:

$$F^*=\frac{n_x}{n}, \tag{2.4}$$

де n_x – кількість варіантів, які менші від x, а n – об'єм вибірки.

Емпірична функція розподілу має такі властивості:

1. $0 \le F^*(x) \le 1$.
2. $F^*(x)$ – функція зростаюча.
3. $F^*(x)=0$, при $x \le x_1$ та $F^*(x)=1$, при $x > x_2$, де x_1 – найменше значення варіантів, x_2 – найбільше значення варіантів.

Задача 2.2. Нехай маємо розподіл частот:

X	3	5	7	10	15
n_i	2	4	7	4	3

Знайти емпіричну функцію розподілу та побудувати її графік.

Розв'язок

Об'єм вибірки:

$$n=\sum_{n=1}^{m} n_i=2+4+7+4+3=20.$$

Оскільки найменше значення варіантів $x_1 = 3$, то

$$F^*(x)=0,$$

для всіх $x \le 3$.

Значення $x < 5$, а саме: $x_1 = 3$, спостерігається двічі, тому

$$F^*(x)=2/20=0{,}1,$$

при $3 < x \le 5$.

Значення $x < 7$, а саме: $x_1 = 3$, $x_2 = 5$ спостерігається $2 + 4 = 6$ разів, тому

$$F^*(x)=6/20=0{,}3,$$

при $5 < x \le 7$.

Значення $x < 10$, а саме: $x_1 = 3$, $x_2 = 5$, $x_3 = 7$, спостерігається $2 + 4 + 7 = = 13$ разів, тому

$$F^*(x)=13/20=0{,}65,$$

при $7 < x \le 10$.

Значення $x < 15$, а саме: $x_1 = 3$, $x_2 = 5$, $x_3 = 7$, $x_4 = 10$, спостерігається $2 + 4 + 7 + 4 = 17$ разів, тому:

$$F^*(x)=17/20=0{,}85,$$

при $10 < x \le 15$.

Оскільки найбільше значення варіантів $x_5 = 15$, то

$$F^*(x)=1,$$

при $x > 15$.

Отже, запишемо шукану емпіричну функцію розподілу:

$$F^*(x)=\begin{cases} 0, \text{при } x \le 3 \\ 0{,}1, \text{при } 3 < x \le 5 \\ 0{,}3, \text{при } 5 < x \le 7 \\ 0{,}65, \text{при } 7 < x \le 10 \\ 0{,}85, \text{при } 10 < x \le 15 \\ 1, \text{при } x > 15 \end{cases}$$

Графік цієї функції зображено на рис. 2.1.

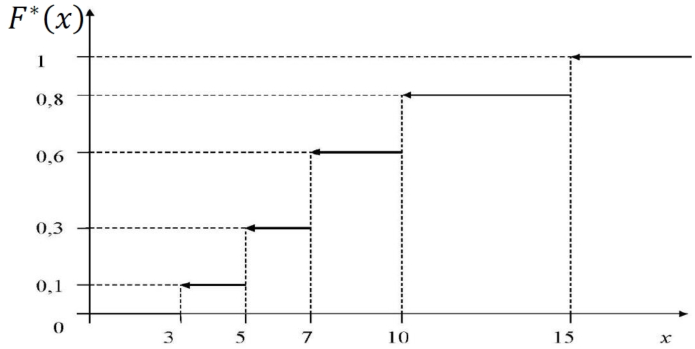


Рис. 2.1. Графік емпіричної функції розподілу

Тема 7.5. Згруповані розподіли вибірки

Часто для спрощення статистичного дослідження використовують згрупований розподіл вибірки.

Загальна схема побудови згрупованого розподілу частот:

1. Визначити найбільше x_{max} та найменше x_{min} значення варіантів і розрахувати їх розмах: $R = x_{max} - x_{min}$.
2. Задати непарне число класів k. При загальній кількості замірів $n \geq 100$ доцільно обрати $9 \leq k \leq 15$, а при $n \leq 100$ можна вибрати $k = 7$.
3. Визначити ширину класу $h = R/k$.
4. Встановити границі класів і підрахувати кількість варіант у кожному із них. При обчислені кількості варіантів значення x_i, що знаходиться на границі класів, слід відносити завжди до одного й того ж класу, тобто там де ця кількість трапилась вперше. Відтак воно стає нижньою границею класу.
5. Визначити частоту для кожного класу і записати ряд розподілу.

Згрупований розподіл накопиченої частоти та відносної частоти. Часто поряд з розподілом частот варіантів необхідно мати розподіл їх накопичених частот.

Розподіл накопиченої частоти отримують послідовним додаванням частот чергового інтервалу, починаючи з першого і закінчуючи останнім. Розподіл накопиченої частоти (позначається через F_i) дозволяє відповісти на питання: скільки існує варіантів, які менші, ніж варіант x_i.

Розподіл накопиченої відносної частоти отримують послідовним додаванням відносних частот чергового інтервалу, починаючи з першого і закінчуючи останнім. Розподіл накопиченої відносної частоти (позначається через F_i/n) дозволяє відповісти на питання: яка частина варіантів менша ніж варіант x_i.

Згрупований розподіл щільності частот та щільності відносних частот. Якщо поділити всі частоти на ширину інтервалу h, то одержимо *розподіл щільності частот вибірки*:

$$n_i/h.$$

Якщо поділити всі відносні частоти на ширину інтервалу h, то одержимо *розподіл щільності відносних частот вибірки*:

$$W_i / h.$$

Задача 2.3. У таблиці 2.1 наведена вибірка середньомісячної платні 100 співробітників підприємства. Впорядкуйте цю вибірку. Запишіть розподіл частот та відносних частот; згрупований розподіл частот та відносних частот; згрупований розподіл накопичених частот та накопичених відносних частот; згрупований розподіл щільності частот та щільності відносних частот середньомісячної заробітної плати співробітників підприємства.

Таблиця 2.1

Середньомісячна заробітна плата 100 співробітників									
368	292	304	332	334	314	324	304	342	323
336	304	323	338	326	304	320	321	322	321
313	302	336	324	312	312	364	362	356	302
322	310	348	334	362	381	304	366	298	304
381	304	368	298	338	326	316	340	328	322
302	292	314	342	321	322	332	290	298	296
296	324	298	338	352	290	304	318	332	322
360	331	312	331	304	316	282	332	342	338
342	324	322	325	302	328	330	354	316	324
314	334	350	324	314	340	314	324	326	308

Розв'язок

Запишемо розподіл частот (та відносних частот) цієї вибірки:

Таблиця 2.2

Розподіл частот (та відносних частот) зарплати 100 співробітників											
x_i	n_i	W_i	x_i	n_i	W_i	x_i	n_i	W_i	x_i	n_i	W_i
282	1	0,01	314	5	0,05	328	2	0,02	350	1	0,01
290	2	0,02	316	3	0,03	330	1	0,01	352	1	0,01
292	2	0,02	318	1	0,01	331	2	0,02	354	1	0,01
296	2	0,02	320	1	0,01	332	4	0,04	356	1	0,01
298	4	0,04	321	3	0,03	334	3	0,03	360	1	0,01
302	4	0,04	322	6	0,06	336	2	0,02	362	2	0,02
304	9	0,09	323	2	0,02	338	4	0,04	364	1	0,01
308	1	0,01	324	7	0,07	340	2	0,02	366	1	0,01
310	1	0,01	325	1	0,01	342	4	0,04	368	2	0,02
312	4	0,04	326	3	0,03	348	1	0,01	381	2	0,02

Наступний крок, який веде до суттєвого спрощення досліджень, – це групування. Знаходимо максимальне та мінімальне значення варіантів вибірки:

$$x_{max}=381, \qquad x_{min}=282.$$

Розмах: $R=x_{max}-x_{min}=381-282=99.$

Задамо кількість класів $k = 11$. Визначимо ширину класу: $h = R/k = 99/11 = 9$.

Додавши частоти для кожного класу інтервалів, значення x_i, що знаходиться на їх границі, заносимо до того класу, де ця кількість трапилась вперше. Результати записуємо у вигляді таблиці.

Таблиця 2.3

Згрупований розподіл середньомісячної зарплати 100 співробітників		
Інтервал	n_i	
282 – 291	3	= 1 + 2
291 – 300	8	= 2 + 2 +4
300 – 309	14	= 4 + 9 + 1
309 – 318	14	= 1 + 4 + 5 + 3 + 1
318 – 327	23	= 1 + 3 + 6 + 2 +7 + 1+ 3
327 – 336	14	2 + 1 + 2 + 4+ 3+ 2
336 – 345	10	= 4 + 2 + 4
345 – 354	4	= 1 + 1 + 1 + 1
354 – 363	4	= 1 + 1 + 2
363 – 372	4	= 1 + 1 + 2
372 – 381	2	= 2
Розмах	100	

Запишемо розподіл накопичених частот, який отримується шляхом послідовного додаванням частот чергового інтервалу.

Таблиця 2.4

Згрупований розподіл накопичених частот зарплати 100 співробітників		
Зарплата	F_i	
≤ 291	3	= 3
≤ 300	11	= 3 + 8
≤ 309	25	= 11 + 14
≤ 318	39	= 25 + 14
≤ 327	62	= 39 + 23
≤ 336	76	= 62 + 14
≤ 345	86	= 76 + 10
≤ 354	90	= 86 + 4
≤ 363	94	= 90 + 4
≤ 372	98	= 94 + 4
≤ 381	100	= 98 + 2

Якщо поділити частоти та накопичені частоти на обсяг вибірки $n = 100$, то отримаємо розподіл відносних частот W_i і накопичених відносних частот F_i/n.

Запишемо розподіл відносних частот W_i та накопичених відносних частот F_i/n.

Таблиця 2.5

Згрупований розподіл відносних та накопичених відносних частот середньомісячної зарплати 100 співробітників			
Інтервал	W_i	Зарплата	F_i/n
282 – 291	0,03	≤ 291	0,03
291 – 300	0,08	≤ 300	0,11
300 – 309	0,14	≤ 309	0,25
309 – 318	0,14	≤ 318	0,39
318 – 327	0,23	≤ 327	0,62
327 – 336	0,14	≤ 336	0,76
336 – 345	0,10	≤ 345	0,86
345 – 354	0,04	≤ 354	0,90
354 – 363	0,04	≤ 363	0,94
363 – 372	0,04	≤ 372	0,98
372 – 381	0,02	≤ 381	1

Запишемо згрупований розподіл щільності частот та відносних частот. Для цього поділимо частоти та відносні частоти на ширину інтервалу $h = 9$.

Щоб підбити підсумки, одержані результати зводимо в одну таблицю (див. таблиця 2.6).

Таблиця 2.6

Згрупований розподіл відносних та накопичених відносних частот середньомісячної зарплати 100 співробітників							
Інтервал	n_i	W_i	n_i/h	W_i/h	Зарплата	F_i	F_i/n
282 – 291	3	0,03	1/3	1/300	≤ 291	3	0,03
291 – 300	8	0,08	8/9	2/225	≤ 300	11	0,11
300 – 309	14	0,14	14/9	7/450	≤ 309	25	0,25
309 – 318	14	0,14	14/9	7/450	≤ 318	39	0,39
318 – 327	23	0,23	23/9	23/900	≤ 327	62	0,62
327 – 336	14	0,14	14/9	7/450	≤ 336	76	0,76
336 – 345	10	0,10	10/9	1/90	≤ 345	86	0,86
345 – 354	4	0,04	4/9	1/225	≤ 354	90	0,90
354 – 363	4	0,04	4/9	1/225	≤ 363	94	0,94
363 – 372	4	0,04	4/9	1/225	≤ 372	98	0,98
372 – 381	2	0,02	2/9	1/450	≤ 381	100	1

Тема 7.6. Полігон частот

Статистичні розподіли вибірки можна подати також і графічно. Завдяки цьому можна побачити характерні зміни у ряду розподілу, не користуючись аналізом цифрових даних.

Якщо в результаті вибірки ми отримали статистичний розподіл ознаки X, яку потрібно дослідити, то матимемо перелік варіантів ознаки:

$$x_1, x_2, x_3, \ldots, x_m.$$

та відповідних їм частот:

$$n_1, n_2, n_3, \ldots, n_m,$$

або відносних частот:

$$W_1, W_2, W_3, \ldots, W_m,$$

Полігоном частот – це ламана, відрізки якої з’єднують точки з координатами:

$$K_1(x_1, n_1), K_2(x_2, n_2), K_3(x_3, n_3), \ldots, K_m(x_m, n_m).$$

Полігоном відносних частот – це ламана, відрізки якої з’єднують точки з координатами:

$$L_1(x_1, W_1), L_2(x_2, W_2), L_3(x_3, W_3), \ldots, L_m(x_m, W_m).$$

Для побудови полігону частот на осі абсцис відкладають варіанти x_i ознаки, а на осі ординат – відповідні їм частоти n_i ($i = 1, 2, 3, \ldots, m$). Точки $K_i(x_i, n_i)$ з’єднують прямими відрізками і одержують полігон частот.

Для побудови полігону відносних частот на осі абсцис відкладають варіанти x_i ознаки, на осі ординат – відповідні їм відносні частоти W_i ($i = 1, 2, 3, \ldots, m$), а потім точки $L_i(x_i, W_i)$ з’єднують відрізками прямих.

При неперервному розподілі ознаки X у разі великої кількості спостережень, увесь проміжок, в якому розміщені спостережні значення ознаки, як правило, розбивають на кілька так званих частинних інтервалів однакової довжини h та знаходять суму частот варіантів, що потрапили в і-й інтервал. Для вибору оптимальної величини інтервалу рекомендовано використовувати формулу

$$h \approx \frac{x_{max} - x_{min}}{1 + 3{,}2 \cdot \ln(n)},$$

де x_{max} та x_{min}– це найбільше та найменше значення вибірки, n – об’єм.

Якщо задано інтервальний розподіл вибірки, то для побудови полігону частот за даними вибірки з’єднують точки, абсцисами яких є значеннями середин частинних інтервалів, а ординатами – відповідні їм значення частот.

Задача 2.4. Вибірку задано інтервальним розподілом частот:

$(x_i; x_{i+1})$	(1; 3)	(3; 5)	(5; 7)	(7; 9)	(9; 11)
n_i	13	9	5	16	7

Побудувати полігон відносних частот.

Розв'язок

Знайдемо об'єм вибірки:

$$n = 13 + 9 + 5 + 16 + 7 = 50.$$

Для побудови інтервального статистичного розподілу відносних частот визначимо відносні частоти:

$$W_1 = \frac{13}{50} = 0{,}26; \qquad W_2 = \frac{9}{50} = 0{,}18;$$

$$W_3 = \frac{5}{50} = 0{,}1; \qquad W_4 = \frac{16}{50} = 0{,}32;$$

$$W_5 = \frac{7}{50} = 0{,}14;$$

Побудуємо інтервальний статистичний розподіл відносних частот:

$(x_i; x_{i+1})$	(1; 3)	(3; 5)	(5; 7)	(7; 9)	(9; 11)
W_i	0,26	0,18	0,1	0,32	0,14

Знайдемо середини частинних інтервалів:

$$x_1^{сер} = \frac{1+3}{2} = 2; \qquad x_2^{сер} = \frac{3+5}{2} = 4;$$

$$x_3^{сер} = \frac{5+7}{2} = 6; \qquad x_4^{сер} = \frac{7+9}{2} = 8;$$

$$x_5^{сер} = \frac{9+11}{2} = 10.$$

Відкладемо на осі абсцис значення середин частинних інтервалів, а на осі ординат значення відносних частот. Послідовно з'єднуючи між собою точки L_i відрізками, одержимо полігон відносних частот (рис. 2.2).

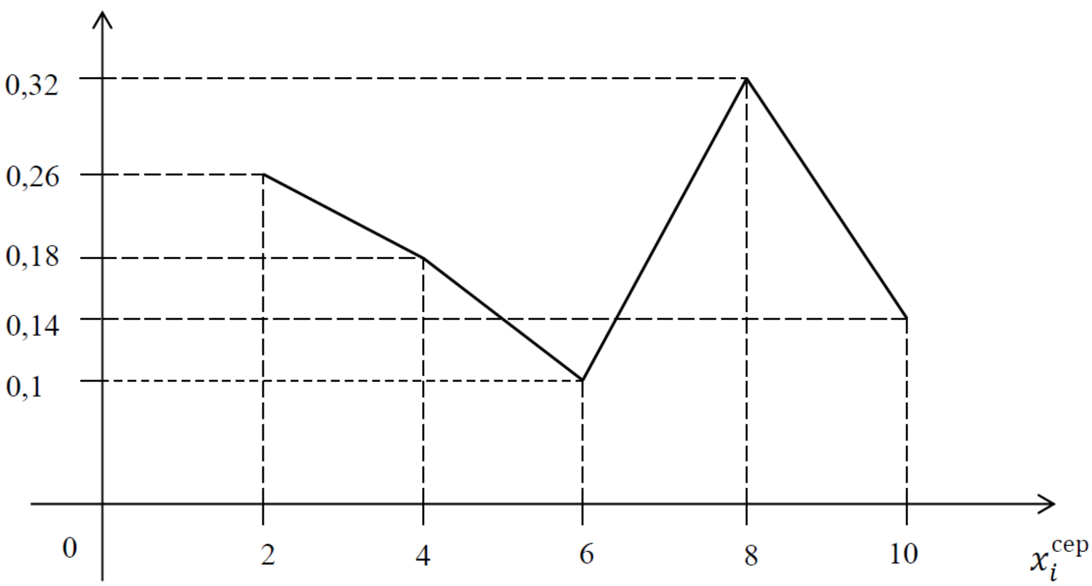


Рис. 2.2. Графік побудови полігону відносних частот

Тема 7.7. Гістограма частот

Гістограма частот – це ступінчаста фігура, яка складається з прямокутників, основами яких є частинні інтервали варіантів довжиною $h = x_i - x_{i-1}$, а висоти – n_i/h (щільність частоти).

Гістограма відносних частот – це ступінчаста фігура, яка складається з прямокутників, основами яких є частинні інтервали варіантів довжиною $h = xi - x_{i-1}$, а висоти – W_i/h (щільність відносної частоти).

Площа гістограми частот дорівнює обсягу вибірки, а площа гістограми відносних частот – одиниці.

Для побудови гістограми частот (або відносних частот) інтервал $[x_{min}; x_{max}]$, від найменшого значення x_{min}, що спостерігали, до найбільшого значення x_{max}, розбивають на декілька відрізків рівної довжини h. Потім підраховують суму частот (відносних частот) значень варіантів, які належать кожному з отриманих відрізків.

Якщо в i-у інтервалі ($i = 1, 2, 3, ..., m$) кількість спостережуваних варіантів спостерігали, з врахуванням їх частот дорівнює n_i, то будують прямокутник $П_i$, основою якого буде i-й інтервал довжиною h, а висотою – відрізок довжиною n_i/h (для відносних частот висота – Wi/h).

Площа такого прямокутника дорівнює $h*n_i/h = n_i$ (у випадку відносних частот – $h*W_i/h = W_i$). Тому площа усіх прямокутників дорівнює сумі $n_i(W_i)$, тобто обсягу вибірки n.

$$S = \sum_{n=1}^{m} n_i = n.$$

У випадку гістограми відносних частот площа прямокутників буде дорівнювати сумі усіх відносних частот

$$S = \sum_{n=1}^{m} W_i = 1.$$

Задача 2.5. Вибірку задано розподілом частот:

x_i	2	4	6	8	10	12	14
n_i	2	3	5	1	4	2	3

Побудувати гістограму відносних частот.

Розв'язок

Щоб побудувати гістограму відносних частот на основі цього статистичного розподілу, складемо спочатку інтервальний статистичний розподіл відносних частот і знайдемо їх щільність.

Обсяг вибірки:

$$n = 2+3+5+1+4+2+3 = 20.$$

Відносні частоти:

$$W_1 = \frac{2}{20} = 0{,}1; \qquad W_2 = \frac{3}{20} = 0{,}15;$$

$$W_3=\frac{5}{20}=0{,}25; \qquad W_4=\frac{1}{20}=0{,}05;$$

$$W_5=\frac{4}{20}=0{,}2; \qquad W_6=\frac{2}{20}=0{,}1;$$

$$W_7=\frac{3}{20}=0{,}15.$$

Знайдемо частинні інтервали, їх довжину і щільність відносних частот. Частинні інтервали визначимо з умови, що варіанти, які задані в дискретному статистичному розподілі, мають бути серединами частинних інтервалів.

Отже, довжина частинних інтервалів:

$$h=x_i-x_{i-1}=2.$$

Шуканий інтервальний статистичний розподіл відносних частот буде мати вигляд:

$(x_i;\, x_{i+1})$	(1; 3)	(3; 5)	(5; 7)	(7; 9)	(9; 11)	(11; 13)	(13; 15)
W_i	0,1	0,15	0,25	0,05	0,2	0,1	0,15

Щільності відносних частот:

$$\frac{W_1}{h}=\frac{0{,}1}{2}=0{,}05; \qquad \frac{W_2}{h}=\frac{0{,}15}{2}=0{,}075;$$

$$\frac{W_3}{h}=\frac{0{,}25}{2}=0{,}125; \qquad \frac{W_4}{h}=\frac{0{,}05}{2}=0{,}025;$$

$$\frac{W_5}{h}=\frac{0{,}2}{2}=0{,}1; \qquad \frac{W_6}{h}=\frac{0{,}1}{2}=0{,}05;$$

$$\frac{W_7}{h}=\frac{0{,}15}{2}=0{,}075.$$

Побудуємо гістограму відносних частот (рис. 2.3).

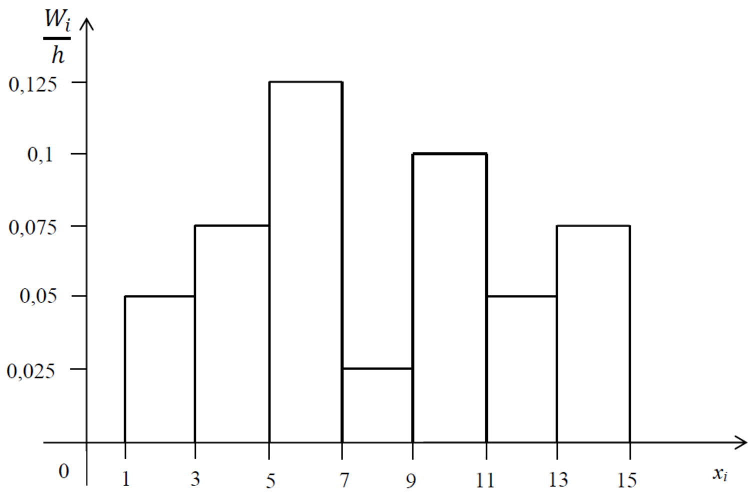


Рис. 2.3. Гістограма відносних частот

Лекція 8. Числові характеристики статистичного матеріалу

Тема 8.1. Числові характеристики вибірки

Доповненням до табличних та графічних методів представлення даних, по суті важливим засобом обробки даних, є розрахунок їх числових характеристик.

Найважливіші з них:

1. Вибіркове середнє.
2. Вибіркова дисперсія.
3. Середньоквадратичне відхилення.

Ці характеристики можуть бути розраховані за даними, що знаходяться у вибірці, або за даними, що входять у кінцеву сукупність (генеральну).

Числові характеристики, розраховані за вибіркою, або ті, що використовуються для опису даних вибірки, називають *статистиками*.

Вибіркове середнє статистичного матеріалу – це сума всіх значень вибірки, поділена на об'єм вибірки n:

$$\overline{x_B} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{2.5}$$

Якщо ми маємо справу із частотною таблицею, то вибіркове середнє вираховуємо згідно з формулою

$$\overline{x_B} = \frac{1}{n}\sum_{i=1}^{m} n_i x_i, \tag{2.6}$$

де n – об’єм вибірки, m – кількість різних варіантів, n_1, n_2, ..., n_m – частоти варіантів x_i – значення і-го варіанта.

Вибіркове середнє значення для інтервального статистичного ряду знаходиться за формулою:

$$\overline{x_B} = \frac{1}{n}\sum_{i=1}^{s} n_i z_i, \tag{2.7}$$

де n – об’єм вибірки, s – кількість інтервалів, z_i – середина і-го інтервалу, n_1, n_2, ..., n_s – кількість елементів вибірки, які потрапили в і-й інтервал.

Вибіркове середнє – це аналог математичного сподівання, який використовується досить часто. Воно може набувати різні числові значення при різних вибірках однакового об’єму.

Тому можна розглядати розподіли (теоретичний та емпіричний) вибіркового середнього разом із числовими характеристики цього розподілу (такий розподіл називають вибірковим).

Основні властивості вибіркового середнього:

1. При добутку усіх варіантів вибірки на однаковий множник вибіркове середнє також множиться на цей множник:

$$\frac{1}{n}\sum_{i=1}^{m} n_i (c x_i) = \frac{c}{n}\sum_{i=1}^{m} n_i x_i = c\,\overline{x}.$$

2. Якщо додати (або відняти) до всіх варіантів вибірки однакове число, то вибіркове середнє зростає або зменшується на це число:

$$\frac{1}{n}\sum_{i=1}^{m} n_i(x_i \pm c) = \frac{1}{n}\sum_{i=1}^{m} n_i x_i \pm \frac{c}{n}\sum_{i=1}^{m} n_i = \overline{x} \pm c.$$

Степеневим середнім вибірки називають середнє значення, яке знаходять згідно з формулою

$$\overline{x_c} = \left(\sum_{i=1}^{m} \frac{n_i}{n} x^{\alpha}\right)^{\frac{1}{\alpha}}. \qquad (2.8)$$

При $\alpha = 1$ отримаємо вибіркове середнє, при $\alpha = 2$ – середньоквадратичне значення вибірки:

$$\overline{x_2} = \left(\sum_{i=1}^{m} \frac{n_i}{n} x^2\right)^{\frac{1}{2}}.$$

При $\alpha = -1$ матимемо середнє гармонічне:

$$\overline{x_{-1}} = \overline{x_{гарм}} = \frac{n}{\sum_{i=1}^{m} \frac{n_i}{x_i}}.$$

Середнім геометричним вибірки називають середнє значення, яке обчислюється за формулою

$$\overline{x_{Г}} = \sqrt[n]{x_1^{n_1} \cdot x_2^{n_2} \cdot \ldots \cdot x_m^{n_m}}. \qquad (2.9)$$

Це середнє обчислюється лише за умови, що всі значення варіантів додатні, тобто $x_i > 0$, $i = 1, 2, ..., m$.

Середнє геометричне застосовується у статистиці для визначення темпу зростання під час дослідження змін ознаки з часом.

Зверніть увагу! Вибір того чи іншого середнього для характеристики розподілу пов'язане з якісним аналізом цього розподілу. Окрім вказаних степеневих середніх, у статистиці застосовують ще структурні середні, які не залежать від значень варіантів, розташованих на краях розподілу та пов'язаних із рядом частот.

До структурних середніх відносять моду та медіану.

Мода M_0 – це елемент, який найчастіше трапляється у вибірці.

Для дискретних статистичних рядів:

$$M_o = x_i,\ якщо\ n_i = max\, n_i. \qquad (2.10)$$

Для інтервальних статистичних рядів:

$$M_o = x_i + h \cdot \frac{n_i - n_{i-1}}{2n_i - n_{i-1} - n_{i+1}}, \qquad (2.11)$$

де x_i – початок інтервалу з найбільшою частотою, n_i – частота i-го інтервалу.

Медіаною M_e називають варіант, який ділить варіаційний ряд на дві частини, які однакові за кількістю варіантів.

Для дискретних статистичних рядів:

$$M_e=\begin{cases} x_m \, при \, n=2m-1 \\ \dfrac{x_m+x_{m+1}}{2}, при \, n=2m \end{cases} \quad (2.12)$$

де x_m – середина варіаційного ряду.

Для інтервальних статистичних рядів:

$$M_e=x_i+h\cdot\frac{\frac{n}{2}-\sum_{j=1}^{i-1}n_j}{n_i}, \quad (2.13)$$

де x_i – початок медіанного інтервалу, тобто такого, якому відповідає перша з нагромаджених частот, що перевищує половину всіх спостережень, h – довжина i-го інтервалу, n_i – частота медіанного інтервалу.

Вибіркова дисперсія – це середнє значення квадратів відхилення варіантів від вибіркового середнього з урахуванням відповідних частот.

$$D_B=\frac{1}{n}\sum_{i=1}^{m}n_i(x_i-\overline{x_B})^2 \quad \text{або} \quad D_B=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x_B})^2. \quad (2.14)$$

Зверніть увагу! Розрахунок вибіркової дисперсії спрощується, якщо її знаходити за формулою

$$D_B=\frac{1}{n}\sum_{i=1}^{m}n_i x_i^2-\left(\frac{1}{n}\sum_{i=1}^{m}n_i x_i\right)^2=\frac{1}{n}\sum_{i=1}^{m}n_i x_i^2-(\overline{x_B})^2. \quad (2.15)$$

Виправлена вибірковою дисперсія (*варіанса*) – це сума квадратів відхилень елементів від вибіркового середнього, поділена на (n – 1):

$$S^2=\frac{n}{n-1}D_B. \quad (2.16)$$

Вибіркову дисперсію та виправлену вибіркову дисперсію для інтервального розподілу знаходять, відповідно, за формулами

$$D_B=\frac{1}{n}\sum_{i=1}^{s}n_i(z_i-\overline{x_B})^2=\frac{1}{n}\sum_{i=1}^{s}n_i z_i^2-(\overline{x_B})^2, \quad (2.17)$$

$$S^2=\frac{1}{n-1}\sum_{i=1}^{s}n_i(z_i-\overline{x_B})^2=\frac{1}{n-1}\sum_{i=1}^{s}n_i z_i^2-(\overline{x_B})^2, \quad (2.18)$$

$$S^2=\frac{n}{n-1}D_B. \quad (2.19)$$

Вибіркове середньоквадратичне відхилення – це квадратний корінь із вибіркової дисперсії:

$$\sigma_B=\sqrt{D_B}. \quad (2.20)$$

Стандарт – це арифметичний квадратний корінь із виправленої вибіркової дисперсії:

$$S=\sqrt{S^2}. \quad (2.21)$$

Розмах – це різниця між найбільшим і найменшим значенням варіаційного ряду (між крайніми елементами), що позначається

$$R = x_{max} - x_{min}. \tag{2.22}$$

Задача 2.6. Розрахувати вибіркове середнє, дисперсію, середньоквадратичне відхилення, варіансу, стандарт, моду, медіану та розмах вибірки, заданої статистичним розподілом частот:

x_i	-1	2	5	8	10
n_i	5	10	20	5	10

Розв'язок

Об'єм вибірки:

$$n = 5+10+20+5+10 = 50.$$

Знайдемо вибіркове середнє, вибіркову дисперсію, вибіркове середньоквадратичне відхилення, варіансу, стандарт, моду, медіану та розмах вибірки:

$$\overline{x_B} = \frac{1}{n}\sum_{i=1}^{m} n_i x_i = \frac{1}{50}\left(5\cdot(-1)+10\cdot 2+20\cdot 5+5\cdot 8+10\cdot 10\right) = \frac{255}{50} = 5{,}1;$$

$$D_B = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^2 - (\overline{x})^2 = \frac{1}{50}\left(5\cdot 1+10\cdot 4+20\cdot 25+5\cdot 64+10\cdot 100\right) - (5{,}1)^2 =$$

$$= \frac{1}{50}\cdot 1865 - 26{,}01 = 37{,}3 - 26{,}01 = 11{,}29;$$

$$\sigma_B = \sqrt{D_B} = \sqrt{11{,}29} \approx 3{,}36;$$

$$S^2 = \frac{n}{n-1}D_B = \frac{50}{49}\cdot 11{,}29 = 11{,}52;$$

$$S = \sqrt{S^2} = \sqrt{11{,}52} \approx 3{,}39;$$

$$M_o = M_e = 5;$$

$$R = x_{max} - x_{min} = 10 - (-1) = 11.$$

Задача 2.7. Знайти вибіркове середнє, дисперсію, середньоквадратичне відхилення, варіансу, стандарт, моду, медіану та розмах вибірки, заданої інтервальним розподілом:

$(x_i; x_{i+1})$	[2; 4)	[4; 6)	[6; 8)	[8; 10)	[10; 12)
n_i	2	8	35	40	15

Розв'язок

Об'єм вибірки:

$$n = 2+8+35+40+15 = 100.$$

Перетворимо цей інтервальний розподіл вибірки на точковий, знайшовши середини частинних інтервалів:

$$x_1^{сер} = \frac{2+4}{2} = 3; \qquad x_2^{сер} = \frac{4+6}{2} = 5;$$

$$x_3^{сер}=\frac{6+8}{2}=7; \qquad x_4^{сер}=\frac{8+10}{2}=9;$$

$$x_5^{сер}=\frac{10+12}{2}=11.$$

Отже, в результаті ми отримали такий статистичний розподіл вибірки:

$x_i^{сер}$	3	5	7	9	11
n_i	2	8	35	40	15

Знайдемо вибіркове середнє, вибіркову дисперсію, вибіркове середньоквадратичне відхилення, варіансу, стандарт, моду, медіану та розмах вибірки:

$$\overline{x_B}=\frac{1}{n}\sum_{i=1}^{s} n_i z_i=\frac{1}{100}(2\cdot 3+8\cdot 5+35\cdot 7+40\cdot 9+15\cdot 11)=\frac{816}{100}=8,16;$$

$$D_B=\frac{1}{n}\sum_{i=1}^{s} n_i z_i^2-(\overline{x})^2=\frac{1}{100}(2\cdot 9+8\cdot 25+35\cdot 49+40\cdot 81+15\cdot 121)-(8,16)^2 =$$

$$=\frac{1}{100}\cdot 6988-66,5856=69,88-66,5856=3,29;$$

$$\sigma_B=\sqrt{D_B}=\sqrt{3,29}\approx 1,81;$$

$$S^2=\frac{n}{n-1}D_B=\frac{100}{99}\cdot 3,29=3,32;$$

$$S=\sqrt{S^2}=\sqrt{3,32}\approx 1,82;$$

$$M_o=8+2\cdot\frac{40-35}{2\cdot 40-35-15}=8+2\cdot\frac{5}{30}=8,33;$$

$$M_e=8+2\cdot\frac{\frac{100}{2}-(8+2+35)}{40}=8+2\cdot\frac{50-45}{40}=8,25;$$

$$R=x_{max}-x_{min}=12-2=10.$$

Тема 8.2. Метод добутків обчислення вибіркового середнього та дисперсії

Нехай вибірку задано у вигляді розподілу рівновіддалених або нерівновіддалених варіантів та відповідних їм частот. У цьому випадку вибіркове середнє і дисперсію зручно знаходити методом добутків.

Алгоритм методу добутків:

1. У перший стовпчик таблиці записують рівновіддалені варіанти вибірки, розміщуючи їх у порядку зростання.
2. Другий стовпчик таблиці записують частоти n_i відповідних варіантів. Суму елементів даного стовпчика (обсяг вибірки n) записують в останню його комірку.

3. Третій стовпчик містить умовні варіанти вибірки. Для знаходження умовних варіантів вибірки необхідно:
3.1. Значення варіанта вибірки (C), яке знаходиться приблизно посередині варіаційного ряду, обрати за так званий «умовний нуль».
3.2. Знайти різницю h між будь-якими двома сусідніми варіантами.
3.3. Обчислити умовні варіанти:

$$u_i = \frac{x_i - C}{h}.$$

Зауважимо, що умовні варіанти завжди повинні бути цілими числами;
4. У четвертий стовпчик записують добутки частот та відповідні їм умовні варіанти ($n_i * u_i$). Суму елементів стовпчика записують в його останню комірку:

$$\sum_{i=1}^{m} n_i \cdot u_i.$$

5. Знаходять добутки частот на квадрати відповідних умовних варіантів $n_i * (u_i)^2$ та записують їх у п'ятий стовпчик. Суму елементів стовпчика записують в його останню комірку:

$$\sum_{i=1}^{m} n_i \cdot u_i^2.$$

6. Знаходять добутки частот на квадрати умовних варіантів, збільшених на одиницю $n_i * (u_i + 1)^2$, та записують їх у шостий стовпчик. Суму елементів стовпчика записують в його останню комірку:

$$\sum_{i=1}^{m} n_i \cdot (u_i + 1)^2.$$

7. Перевіряють розрахунок: сума елементів шостого стовпчика повинна задовольняти тотожність

$$\sum_{i=1}^{m} n_i \cdot (u_i + 1)^2 = \sum_{i=1}^{m} n_i \cdot u_i^2 + 2 \sum_{i=1}^{m} n_i \cdot u_i + n.$$

8. Розраховують умовні моменти за формулами:

$$M_1^* = \frac{1}{n} \sum_{i=1}^{m} n_i \cdot u_i, \quad M_2^* = \frac{1}{n} \sum_{i=1}^{m} n_i \cdot u_i^2.$$

9. Розраховують вибіркове середнє та вибіркову дисперсію:

$$\overline{x_B} = M_1^* \cdot h + C, \quad D_B = (M_2^* - (M_1^*)^2) \cdot h^2.$$

Зверніть увагу! Якщо варіанти вибірки не рівновіддалені, то інтервал (проміжок,) на якому розміщені всі варіанти вибірки, ділять на кілька однакових за довжиною частинних інтервалів, кожен з яких повинен містити від 8 до 10 варіантів. Потім знаходять середини частинних інтервалів, які й утворюють послідовність рівновіддалених варіантів. За частоту кожної середини частинного інтервалу обирають суму частот варіантів, які потрапили у відповідний частинний інтервал.

Під час розрахунку вибіркової дисперсії з метою зменшення похибки, зумовленої групуванням (особливо при малій кількості інтервалів), враховують поправку Шеппарда, а саме: обчислюють дисперсію за формулою

$$D_B^{'}=D_B-\frac{1}{12}\cdot h^2. \qquad (2.23)$$

Задача 2.8. Знайти методом добутків вибіркове середнє та вибіркову дисперсію для статистичного розподілу вибірки об’єму n = 100.

x_i	2	8	9	13	15	18	20	21	24	27
n_i	8	1	15	13	9	3	4	10	22	15

Розв’язок

Розіб’ємо інтервал [2, 27] на п’ять частинних інтервалів довжиною h = 5:

(2; 7] (7; 12] (12; 17] (17; 22] (22; 27].

Новими варіантами будуть середини цих частинних інтервалів:

$$y_1=\frac{2+7}{2}=4{,}5; \qquad y_2=\frac{7+12}{2}=9{,}5; \qquad y_3=\frac{12+17}{2}=14{,}5;$$

$$y_4=\frac{17+22}{2}=19{,}5; \qquad y_5=\frac{22+27}{2}=24{,}5.$$

За частоти m_i варіантів y_i візьмемо суму частот варіантів, які потрапили у відповідний i-й інтервал:

$$m_1=8, \qquad m_2=1+15=16, \qquad m_3=13+9=22,$$

$$m_4=3+4+10=17, \qquad m_5=22+15=37.$$

Отже, ми отримали такий статистичний розподіл рівновіддалених варіантів:

y_i	4,5	9,5	14,5	19,5	24,5
m_i	8	16	22	17	37

Складемо розрахункову таблицю:

y_i	m_i	u_i	m_iu_i	$m_iu_i^2$	$m_i(u_i+1)^2$
4,5	8	-2	-16	32	8
9,5	16	-1	-16	16	0
14,5	22	0	0	0	22
19,5	17	1	17	17	68
24,5	37	2	74	148	333
	$n=100$		$\sum_{i=1}^{5} m_iu_i=59$	$\sum_{i=1}^{5} m_iu_i^2 = 213$	$\sum_{i=1}^{5} m_i(u_i+1)^2 = 431$

Перевірка:

$$\sum_{i=1}^{5} m_i \cdot (u_i+1)^2 = 431,$$

$$\sum_{i=1}^{5} m_i \cdot u_i^2 + 2\sum_{i=1}^{5} m_i \cdot u_i + n = 213 + 2 \cdot 59 + 100 = 431.$$

Обчислимо умовні моменти першого і другого порядків:

$$M_1^* = \frac{1}{n}\sum_{i=1}^{m} n_i \cdot u_i = \frac{59}{100} = 0{,}59, \quad M_2^* = \frac{1}{n}\sum_{i=1}^{m} n_i \cdot u_i^2 = \frac{213}{100} = 2{,}13.$$

Методом добутків знайдемо вибіркове середнє й дисперсію, враховуючи, що «умовний нуль» $C = 14{,}5$:

$$\overline{x_B} = M_1^* \cdot h + C = 0{,}59 \cdot 5 + 14{,}5 = 17{,}45;$$

$$D_B = (M_2^* - (M_1^*)^2) \cdot h^2 = (2{,}13 - 0{,}59^2) \cdot 5^2 = 44{,}5475.$$

Оскільки кількість частинних інтервалів мала, то скористаємося поправкою Шеппарда:

$$D_B^{'} = D_B - \frac{1}{12} \cdot h^2 = 44{,}5475 - \frac{1}{12} \cdot 5^2 \approx 42{,}46.$$

Тема 8.3. Метод сум обчислення вибіркового середнього та дисперсії

Нехай вибірку задано у вигляді розподілу рівновіддалених варіантів і відповідних їм частот. Тоді вибіркове середнє та вибіркову дисперсію зручно знаходити методом сум.

Алгоритм методу сум:

1. У перший стовпчик таблиці записують рівновіддалені варіанти x_i вибірки, розміщуючи їх у порядку зростання.
2. У другий стовпчик таблиці записують відповідні частоти n_i варіантів. Суму елементів цього стовпчика (обсяг вибірки n) записують в його останню комірку.
3. Третій стовпчик містить умовні варіанти вибірки. Для знаходження умовних варіант вибірки необхідно:

3.1. Значення варіанта вибірки з найбільшою частотою (C) обрати за «умовний нуль», це значення ще називають модою.

3.2. У порожні комірки третього стовпця, які розміщені над нулем (крім верхньої), зверху вниз записують послідовно нагромаджені частоти. Додавши всі нагромаджені частоти зверху до нуля, отримують відповідно число b_1, яке записують у верхню комірку.

3.3. У порожні комірки третього стовпця, які розміщені під нулем (крім нижньої), знизу вверх записують послідовно нагромаджені частоти. Додавши всі нагромаджені частоти знизу до нуля, отримують відповідно число a_1, яке записують у його нижню комірку.

4. У четвертому стовпці знову записують «умовний нуль», над і під яким записують ще по одному нулю. Аналогічно до третього, заповнюють четвертий стовпець, при цьому нагромаджуються частоти третього стовпця. Суму

нагромаджених частот, розміщених над нулем, позначають b_2 і записують у верхній комірці четвертого стовпця. Суму нагромаджених частот, розміщених під нулем, позначають a_2 і записують у нижню клітинку четвертого стовпця.

5. Обчислюють:

$$d_1 = a_1 - b_1, \quad s_1 = a_1 + b_1, \quad s_2 = a_2 + b_2.$$

6. Обчислюють умовні моменти за формулами

$$M_1^* = \frac{d_1}{n}, \quad M_2^* = \frac{s_1 + 2s_2}{n}.$$

7. Обчислюють вибіркове середнє та вибіркову дисперсію за формулами

$$\overline{x_B} = M_1^* \cdot h + C \qquad D_B = (M_2^* - (M_1^*)^2) \cdot h^2.$$

Задача 2.9. Знайдіть методом сум вибіркове середнє та дисперсію для статистичного розподілу вибірки об'єму n = 100.

x_i	52	58	64	70	76	82	88	94	100	106
n_i	4	8	10	13	18	21	14	9	2	1

Розв'язок

Складемо розрахункову таблицю:

1	2	3	4
x_i	n_i	$b_1 = 126$	$b_2 = 131$
52	4	4	4
58	8	12	16
64	10	22	38
70	13	35	73
76	18	53	0
82	21	0	0
88	14	26	0
94	9	12	16
100	2	3	4
106	1	1	1
		$a_1 = 42$	$a_2 = 21$

Знайдемо d_1, s_1, s_2:

$$d_1 = a_1 - b_1 = 42 - 126 = -84,$$

$$s_1 = a_1 + b_1 = 42 + 126 = 168,$$

$$s_2 = a_2 + b_2 = 21 + 131 = 152.$$

Обчислимо умовні моменти першого і другого порядку:

$$M_1^* = \frac{d_1}{n} = \frac{-84}{100} = -0{,}84,$$

$$M_2^* = \frac{s_1 + 2s_2}{n} = \frac{168 + 2\cdot 152}{100} = 4{,}72;$$

Знайдемо вибіркове середнє та дисперсію. Врахувавши, що «умовний нуль» $C = 82$, а крок (відстань між двома сусідніми варіантами) $h = 6$, одержимо:

$$\overline{x_B} = M_1^* \cdot h + C = -0{,}84 \cdot 6 + 82 = 76{,}96;$$

$$D_B = \left(M_2^* - \left(M_1^*\right)^2\right) \cdot h^2 = \left(4{,}72 - \left(-0{,}84\right)^2\right) \cdot 6^2 = 144{,}5184.$$

Лекція 9. Статистичні оцінки параметрів розподілу

Тема 9.1. Точкові оцінки

Статистичною оцінкою Θ^* невідомого параметра Θ теоретичного розподілу називають таку функцію $f(X_1, X_2, ..., X_n)$, де $X_1, X_2, ..., X_n$ – спостережувані випадкові величини.

Точкова оцінка – це статистична оцінка, яка визначається одним числом $\Theta^* = f(x_1, x_2, ..., x_n)$, де $x_1, x_2, ..., x_n$ – результати n спостережень над кількісною ознакою випадкової величини.

Наприклад, вибіркове середнє, дисперсія та вибіркове середньоквадратичне відхилення – точкові оцінки відповідних числових характеристик генеральної сукупності.

Точкові оцінки параметрів розподілу – випадкові величини, їх можна вважати первинними результатами обробки вибірки, тому що невідомо з якою точністю кожна з них оцінює відповідну числову характеристику генеральної сукупності.

Якщо обсяг вибірки доволі великий, то точкові оцінки задовольняють потреби точності.

Щоб статистичні оцінки давали, якмога кращі наближення, вони повинні задовольняти певним вимогам. Розглянемо ці вимоги.

Нехай, Θ^* – це статистична оцінка невідомого параметра Θ деякого теоретичного розподілу. Припустимо, що за вибіркою обсягом n знайдена оцінка Θ_1^*. Для інших вибірок такого ж об'єму одержано оцінки $\Theta_2^*, \Theta_3^*, ..., \Theta_m^*$. Оцінку Θ^* можна розглядати як випадкову величину, а числа $\Theta_2^*, \Theta_3^*, ..., \Theta_m^*$ – як її можливі значення відповідно.

Якщо Θ_k^* ($k = 1, 2, 3, ..., m$) будуть більші від Θ, то оцінка Θ^* дає наближене значення Θ з так званим *надлишком*. У цьому випадку математичне сподівання випадкової величини Θ^* є більшим від Θ: $M(\Theta^*) > \Theta$.

Якщо Θ^* дає оцінку з *нестачею*, то математичне сподівання випадкової величини Θ^* менше від Θ: $M(\Theta^*) < \Theta$.

Отже, використання статистичної оцінки, математичне сподівання якої не дорівнює параметру Θ, спричиняє систематичні похибки. Умова $M(\Theta^*) = \Theta$ застерігає (застраховує) від таких похибок.

Незміщеною називають точкову оцінку Θ^*, математичне сподівання якої дорівнює оцінюваному параметру Θ для будь-якого об'єму вибірки:

$$M(\Theta^*)=\Theta .$$

Зміщеною називають точкову оцінку Θ^*, математичне сподівання якої відрізняється від оцінюваного параметра Θ.

Якщо $M(\Theta^*) = \Theta + b(\Theta)$, то $b(\Theta)$ називають *зсувом оцінки*.

Вимога або умова про незміщеність оцінки Θ^* недостатня, оскільки можливі значення Θ^* можуть бути дуже відхилятись від середнього значення, а дисперсія $D(\Theta^*)$ може бути великою. Тоді, знайдена за даними однієї вибірки оцінка може значно відрізнятися від середнього значення Θ^* та від параметра Θ.

Якщо дисперсія $D(\Theta^*)$ незначна, то можливість допустити помилку виключена. Тому для статистичної оцінки виникає вимога стосовно її ефективності.

Ефективною називають таку оцінку Θ^*, яка при заданому обсягу n має найменшу можливу дисперсію.

Обґрунтованою називають статистичну оцінку, яка при $n \to \infty$ прямує за ймовірністю до параметра, що оцінюється.

Послідовність оцінок Θ_k^* (k = 1, 2, 3, ..., m) параметра Θ називається *спроможною*, якщо $\Theta_k^* \to \Theta$, $n \to \infty$.

Послідовність оцінок Θ_k^* (k = 1, 2, 3, ..., m) параметра Θ називається *сильно спроможною*, якщо $\Theta_k^* \to \Theta$ з ймовірністю 1 при $n \to \infty$.

Незміщеною оцінкою генерального середнього (або математичного сподівання) є вибіркове середнє:

$$\overline{x_B}=\frac{1}{n}\sum_{i=1}^{m} n_i \cdot x_i ,$$

де n_i – частота варіантів x_i; x_i – варіанти вибірки; $n = n_1 + n_2 + ... + n_m$ – об’єм вибірки.

Зверніть увагу! Якщо варіанти вибірки x_i – досить великі або дуже малі (близькі до нуля) числа, то для спрощення розрахунків доцільно відняти (у разі великих від’ємних чисел – додати) від кожного варіанта одне й те саме число C (як C можна вибрати будь-яке число, розміщене приблизно посередині варіаційного ряду), потім поділити (у разі близьких до нуля чисел – помножити) на одне й те саме число b (як b можна вибрати найбільший спільний дільник – НСД), а відтак до умовних варіантів:

$$u_i=\frac{x_i \pm C}{b} \quad \text{або} \quad u_i=x_i \cdot b \tag{2.24}$$

Тоді

$$\overline{x_B}=\frac{b}{n}\sum_{i=1}^{m} n_i u_i \mp C = b\overline{u_b} \mp C \quad \text{або} \quad \overline{x_B}=\frac{\sum_{i=1}^{m} n_i u_i}{bn}=\frac{\overline{u_b}}{b}. \tag{2.25}$$

Зміщеною оцінкою генеральної дисперсії є вибіркова дисперсія:

$$D_B=\frac{1}{n}\sum_{i=1}^{m} n_i(x_i-\overline{x_B})^2.$$

Ця оцінка зміщена, оскільки

$$M(D_B)=\frac{n-1}{n}D_Г.$$

Для обчислення вибіркової дисперсії можна скористатися зручнішою формулою:

$$D_B=\frac{1}{n}\sum_{i=1}^{m} n_i x_i^2-\left(\frac{1}{n}\sum_{i=1}^{m} n_i x_i\right)^2=\overline{x^2}-(\overline{x_B})^2.$$

Зверніть увагу! Якщо варіанти вибірки – досить великі або дуже малі (близькі до нуля) числа, то для спрощення розрахунків доцільно відняти (у разі великих від'ємних чисел – додати) від кожної варіанти одне й те саме число *C* (як *C* можна вибрати будь-яке число, розміщене приблизно посередині варіаційного ряду), потім поділити (у разі близьких до нуля чисел – помножити) на одне й те саме число *b* (як *b* можна вибрати найбільший спільний дільник – НСД), далі перейти до умовних варіантів:

$$u_i=\frac{x_i\pm C}{b} \quad \text{або} \quad u_i=x_i\cdot b.$$

Тоді

$$D_B=b^2\left(\frac{1}{n}\sum_{i=1}^{m} n_i u_i^2-\left(\frac{1}{n}\sum_{i=1}^{m} n_i u_i\right)^2\right)=b^2\left(\overline{u^2}-(\bar{u})^2\right), \tag{2.26}$$

або

$$D_B=\left(\frac{1}{b^2 n}\sum_{i=1}^{m} n_i u_i^2-\left(\frac{1}{bn}\sum_{i=1}^{m} n_i u_i\right)^2\right)=\frac{\left(\overline{u^2}-(\bar{u})^2\right)}{b^2}. \tag{2.27}$$

Незміщеною оцінкою генеральної дисперсії є виправлена вибіркова дисперсія:

$$S^2=\frac{n}{n-1}D_B.$$

Поправка $n/(n - 1)$ – це так звана *поправка Бесселя*.

Задача 2.10. З генеральної сукупності отримано деяку вибірку обсягу n = 100:

xi	0,002	0,004	0,006	0,008	0,01	0,012	0,014
ni	7	29	35	12	9	5	3

Знайти зміщену оцінку генеральної дисперсії.

Розв'язок

Зміщеною оцінкою генеральної дисперсії $D_Г$ є вибіркова дисперсія $D_В$. Варіанти вибірки є малими, близькими до нуля числами, відповідно перейдемо до умовних варіантів:

$$u_i = x_i \cdot b = x_i \cdot 1000.$$

Як b вибрано число 1000, оскільки в такому разі ми отримаємо цілі числа:

$$u_1 = x_1 \cdot 1000 = 0{,}002 \cdot 1000 = 2;$$

$$\ldots$$

$$u_7 = x_7 \cdot 1000 = 0{,}014 \cdot 1000 = 14.$$

Тепер можна знайти вибіркову дисперсію:

$$D_B = \left(\frac{1}{b^2 n} \sum_{i=1}^{m} n_i u_i^2 - \left(\frac{1}{bn} \sum_{i=1}^{m} n_i u_i \right)^2 \right) =$$

$$= \frac{7 \cdot 2^2 + 29 \cdot 4^2 + 35 \cdot 6^2 + 12 \cdot 8^2 + 9 \cdot 10^2 + 5 \cdot 12^2 + 3 \cdot 14^2}{1000^2 \cdot 100} -$$

$$- \left[\frac{7 \cdot 2 + 29 \cdot 4 + 35 \cdot 6 + 12 \cdot 8 + 9 \cdot 10 + 5 \cdot 12 + 3 \cdot 14}{1000 \cdot 100} \right]^2 =$$

$$= 4728 \cdot 10^{-8} - 394384 \cdot 10^{-10} = 784{,}16 \cdot 10^{-8}.$$

Задача 2.11. У результаті статистичних досліджень випадкової величини отримано вибірку:

47, 45, 46, 46, 46, 45, 47, 44, 46, 45, 45, 46, 46, 44, 46, 48, 46,
46, 45, 46, 44, 46, 45, 47, 46, 46, 47, 46, 46, 48, 44, 46, 45, 46,
45, 46, 44, 47, 46, 46, 45, 47, 48, 44, 46, 46, 45, 46, 47, 45.

Знайдіть незміщені оцінки генерального середнього та генеральної дисперсії.

Розв'язок

Побудуємо статистичний розподіл вибірки:

xi	44	45	46	47	48
ni	6	11	23	7	3

Об'єм вибірки n = 50.

Незміщеною оцінкою генерального середнього є вибіркове середнє.

Варіанти вибірки є великими числами, відповідно переходимо до умовних варіантів:

$$u_i = \frac{x_i \pm C}{b} = \frac{x_i - 46}{1}.$$

Як C вибрано число 46, b (НСД), у цьому випадку воно дорівнює 1.

$$u_1 = x_1 - 46 = 44 - 46 = -2;$$

$$\ldots$$

$$u_5 = x_5 - 46 = 48 - 46 = 2.$$

Тепер можна знайти вибіркове середнє:

$$\overline{x_B}=\frac{b}{n}\sum_{i=1}^{m}n_i u_i \mp C=\frac{6\cdot(-2)+11\cdot(-1)+23\cdot 0+7\cdot 1+3\cdot 2}{50}+46=-\frac{10}{50}+46=45{,}8.$$

Щоб визначити незміщену оцінку генеральної дисперсії (виправлену вибіркову дисперсію), знайдемо вибіркову дисперсію й помножимо її на поправку Бесселя:

$$D_B=\left(\frac{1}{b^2 n}\sum_{i=1}^{m}n_i u_i^2-\left(\frac{1}{bn}\sum_{i=1}^{m}n_i u_i\right)^2\right)=$$

$$=\frac{6\cdot(-2)^2+11\cdot(-1)^2+23\cdot 0^2+7\cdot 1^2+3\cdot 2^2}{1^2\cdot 50}-\left[\frac{6\cdot(-2)+11\cdot(-1)+23\cdot 0+7\cdot 1+3\cdot 2}{1\cdot 50}\right]^2=$$

$$=\frac{54}{50}-\left[\frac{(-10)}{50}\right]^2=1{,}04; \quad S^2=\frac{n}{n-1}D_B=\frac{50}{49}\cdot 1{,}04\approx 1{,}06.$$

Тема 9.2. Методи визначення точкових статистичних оцінок

9.2.1. Методом моментів

Методом моментів знаходження точкових оцінок називають метод, за допомогою якого обчислення невідомих параметрів заданого розподілу здійснюють шляхом прирівнювання теоретичних та емпіричних моментів.

Якщо розподіл визначається одним параметром, то для знаходження оцінки здійснюється прирівняння математичного сподівання до вибіркового середнього:

$$M(X)=\overline{x_B},$$

а потім із цього рівняння визначають шукану точкову оцінку невідомого параметра.

Якщо розподіл визначається двома параметрами, то і точкові оцінки знаходять за допомогою системи рівнянь:

$$\begin{cases} M(X)=\overline{x_B}; \\ D(X)=D_B. \end{cases} \tag{2.28}$$

Лівими частинами цих рівнянь є математичне сподівання та дисперсія, які дорівнюють вибірковому середньму та вибірковій дисперсії.

Задача 2.12**.** Випадкова величина X – це зріст дорослої людини, яка розподілена за нормальним законом з параметрами a, σ. У результаті досліджень отримано такий статистичний розподіл зросту дорослих людей:

Ріст (см)	(145; 155]	(155; 165]	(165; 175]	(175; 185]	(185; 195]	(195; 205]	(205; 215]
Число осіб	24	112	263	322	202	66	11

Знайдіть за допомогою методу моментів точкову оцінку невідомих параметрів a, σ нормального розподілу.

Розв'язок

Перетворимо цей інтервальний розподіл на точковий, вибравши як варіанти середини частинних інтервалів:

Ріст (см)	150	160	170	180	190	200	210
Число осіб	24	112	263	322	202	66	11

Оскільки нормальний закон розподілу залежить від двох параметрів, то відповідно треба знайти вибіркове середнє та вибіркову дисперсію, а потім прирівняти їх до математичного сподівання й дисперсії.

Перейдемо до умовних варіантів:

$$u_i=\frac{x_i-180}{10};$$

$$u_1=\frac{150-180}{10}=-3; \quad \ldots \quad u_7=\frac{210-180}{10}=3.$$

Знайдемо вибіркове середнє та вибіркову дисперсію:

$$\overline{x_B}=\frac{10}{n}\sum_{i=1}^{m} n_i u_i+180 \ =$$

$$=\ \frac{10\left(24\cdot(-3)+112\cdot(-2)+263\cdot(-1)+322\cdot 0+202\cdot 1+66\cdot 2+11\cdot 3\right)}{1000}+180\ =$$

$$=\ 0{,}01\cdot(-192)+180=178{,}08.$$

$$D_B=b^2\left(\frac{1}{n}\sum_{i=1}^{m} n_i u_i^2-\left(\frac{1}{n}\sum_{i=1}^{m} n_i u_i\right)^2\right)\ =$$

$$=\ 10^2\left(\frac{24\cdot(-3)^2+112\cdot(-2)^2+263\cdot(-1)^2+322\cdot 0^2+202\cdot 1^2+66\cdot 2^2+11\cdot 3^2}{1000}\right)\ -$$

$$-\ 10^2\cdot\left(\frac{24\cdot(-3)+112\cdot(-2)+263\cdot(-1)+322\cdot 0+202\cdot 1+66\cdot 2+11\cdot 3}{1000}\right)^2\ =$$

$$=\ 100\cdot\left(\frac{1492}{1000}-\left(\frac{-192}{1000}\right)^2\right)=145{,}5136.$$

Оскільки параметр а нормального закону розподілу є математичним сподіванням, *а* параметр σ – середньоквадратичним відхиленням, тоді оцінками цих параметрів є:

$$a^*=\overline{x_B}=178{,}08; \qquad \sigma^*=\sqrt{D_B}=\sqrt{145{,}5136}\approx 12{,}0629.$$

9.2.2. Методом найменших квадратів

Згідно з *методом найменших квадратів*, статистичні оцінки визначаються з умови мінімізації суми квадратів відхилень варіантів вибірки від статистичної оцінки Θ^*

Отже, використовуючи метод найменших квадратів, можна, наприклад, визначити статистичну оцінку для

$$\overline{X_Г}=M(X).$$

Для цього скористаємося функцією

$$u=\sum_{i=1}^{n} n_i(x_i-\Theta^*)^2.$$

Використовуючи умову екстремуму, отримаємо:

$$\frac{du}{d\Theta^*}=-2\sum_{i=1}^{n} n_i(x_i-\Theta^*)=0 \quad \rightarrow \quad \sum_{i=1}^{n} n_i x_i-\sum_{i=1}^{n} n_i\Theta^*=0 \quad \rightarrow$$

$$\rightarrow \quad \Theta^*=\frac{\sum_{i=1}^{n} n_i x_i}{n}=\overline{x_B}.$$

Звідси для $\Theta = X_Г$ точковою статистичною оцінкою Θ^* буде $\overline{x_B}$ – вибіркове середнє.

9.2.3. Метод максимальної правдоподібності

Метод максимальної правдоподібності займає головне місце в теорії статистичного оцінювання параметрів. Розробив його Рональд Фішер та свого часу доповнив Карл Фрідріх Гаусс.

Цей метод полягає у знаходженні максимуму функції одного або кількох оцінюваних параметрів.

Нехай X – дискретна випадкова величина з відомим законом розподілу, який визначається невідомим параметром Θ. За даними вибірки $x_1, x_2, ..., x_n$ одержаної в результаті спостережень над випадковою величиною. Необхідно знайти точкову оцінку $\Theta^*(x_1, x_2, ..., x_n)$ параметра Θ.

Функція правдоподібності L дискретної випадкової величини – це функція аргументу Θ.

$$L(x_1,x_2,...,x_n;\Theta)=p(x_1;\Theta)\cdot p(x_2;\Theta)\cdot...\cdot p(x_n;\Theta),$$

де $p(x_i; \Theta)$ – імовірність того, що в результаті випробування випадкова величина набуде значення x_i.

Оцінкою максимальної правдоподібності параметра Θ називають таке значення Θ^*, за допомогою якого функція правдоподібності досягає свого максимуму.

Логарифмічна функція правдоподібності – це функція ln L. Оскільки функції L і ln L досягають свого максимуму при одному і тому самому значенні Θ, то зручніше знаходити максимум функції ln L, а не L.

Якщо випадкова величина (X) неперервна, то відомою вважається густина ймовірності $f(x)$, а невідомим – параметр, від якого залежить ця густина.

Функцією правдоподібності L неперервної випадкової величини називають таку функцію аргументу Θ:

$$L(x_1, x_2, \ldots, x_n; \Theta) = f(x_1; \Theta) \cdot f(x_2; \Theta) \cdot \ldots \cdot f(x_n; \Theta). \qquad (2.29)$$

Оцінку максимальної правдоподібності невідомого параметра розподілу неперервної випадкової величини знаходять так само, як і у випадку дискретної випадкової величини, тобто:

1. Визначають похідну:

$$\frac{d \ln L}{d\Theta} \quad \text{або} \quad \frac{dL}{d\Theta}.$$

2. Знаходять корені рівняння:

$$\frac{d \ln L}{d\Theta} = 0 \quad \text{або} \quad \frac{dL}{d\Theta} = 0.$$

Ці рівняння називають *рівняннями правдоподібності*;

3. Визначають другу похідну:

$$\frac{d^2 \ln L}{d\Theta^2} \quad \text{або} \quad \frac{d^2 L}{d\Theta^2}.$$

Якщо параметр Θ двовимірний, тобто $\Theta = (\Theta_1, \Theta_2)$, то для пошуку максимуму функції правдоподібності складають і розв'язують таку систему рівнянь:

$$\begin{cases} \dfrac{d \ln L}{d\Theta_1} = 0; \\ \dfrac{d \ln L}{d\Theta_2} = 0. \end{cases} \quad \text{або} \quad \begin{cases} \dfrac{dL}{d\Theta_1} = 0; \\ \dfrac{dL}{d\Theta_2} = 0. \end{cases}$$

Наприклад, якщо ознака генеральної сукупності має нормальний закон розподілу, то функція максимальної правдоподібності набере наступного вигляду:

$$L(x_1, x_2, \ldots, x_n; \Theta_1^*, \Theta_2^*) = \frac{1}{(2\pi\Theta_2^*)^{\frac{n}{2}}} \cdot \mathrm{e}^{-\frac{\sum_{i=1}^{n}(x_i - \Theta_1^*)^2}{2\Theta_2^*}}.$$

Водночас як статистичні оцінки Θ_1^*, Θ_2^* вибирають ті значення, за яких задана вибірка буде найімовірнішою, тобто функція $L(x_1, x_2, \ldots, x_n; \Theta_1^*, \Theta_2^*)$ набуває максимуму.

На практиці зручно від функції $L(x_1, x_2, \ldots, x_n; \Theta_1^*, \Theta_2^*)$ перейти до її логарифма, тобто

$$\ln L(x_1, x_2, \ldots, x_n; \Theta_1^*, \Theta_2^*) = -\frac{n}{2}(\ln \pi + \ln \Theta_2^*) - \frac{\sum_{i=1}^{n}(x_i - \Theta_1^*)^2}{2\Theta_2^*}.$$

Згідно з умовою екстремуму, для цієї функції отримаємо:

$$\begin{cases} \dfrac{d\ln L}{d\Theta_1^*} = -\dfrac{\sum\limits_{i=1}^{n}(x_i - \Theta_1^*)}{\Theta_2^*} = 0; \\ \dfrac{d\ln L}{d\Theta_2^*} = -\dfrac{n}{2\Theta_2^*} + \dfrac{\sum\limits_{i=1}^{n}(x_i - \Theta_1^*)^2}{2(\Theta_2^*)^2} = 0. \end{cases}$$

З першого рівняння системи одержимо:

$$\Theta_1^* = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x_B}.$$

З другого рівняння системи:

$$\Theta_2^* = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x_B})^2 = D_B.$$

Отже, для $\overline{X_Г}$, $= M(X)$ точковою статистичною оцінкою є $\overline{x_B}$, а для $D_Г$ – це $D_В$.

Тема 9.3. Інтервальні оцінки

Якщо обсяг вибірки малий, то точкові оцінки можуть давати суттєві похибки, тому питання точності оцінок у даному випадку важливе. Тому, використовують так звані інтервальні оцінки.

Інтервальною називають статистичну оцінку, яка визначається двома числами – кінцями інтервалів.

Нехай знайдена за даними вибірки статистична оцінка Θ^* буде оцінкою невідомого параметра Θ.

Різниця між статистичною оцінкою Θ^* та її оцінювальним параметром Θ, взята за абсолютною величиною, називається точністю оцінки, а саме:

$$|\Theta - \Theta^*| < \delta, \tag{2.30}$$

де δ є точністю оцінки.

Оскільки Θ^* є випадковою величиною, то і δ буде випадковою, тому нерівність (2.30) справджується з деякою ймовірністю.

Імовірність, з якою береться нерівність (2.30), наступна:

$$P(|\Theta - \Theta^*| < \delta) = \gamma \tag{2.31}$$

її називають *надійністю*.

Рівність (2.31) також можна розписати як

$$P(\Theta^* - \delta < \Theta < \Theta^* + \delta) = \gamma \tag{2.32}$$

Найчастіше надійність задається наперед і, залежно від обставин, воно набуває значення: 0,95 або 0,99 або 0,999. Інтервал [Θ^* - δ, Θ^* + δ], який покриває оцінюваний параметр Θ генеральної сукупності із заданою надійністю γ, називають *довірчим*.

9.3.1. Довірчий інтервал для оцінки математичного сподівання нормального розподілу при відомому середньоквадратичному відхиленні

Нехай ознака X генеральної сукупності має нормальний закон розподілу. Побудуємо довірчий інтервал для $\overline{X}_{Г}$ генеральної сукупності (або X), знаючи числове значення середньоквадратичного відхилення сукупності σ; із заданою надійністю γ. Оскільки X, розподілена нормально, то і $\overline{X}$ теж буде розподілено нормально. Тоді

$$M(\overline{X})=a, \qquad \sigma(\overline{X})=\frac{\sigma}{\sqrt{n}}.$$

Скориставшись (2.32), одержимо:

$$P(|\overline{X}-a|<\delta)=\gamma. \tag{2.33}$$

Рівність (2.33) можна записати, позначивши:

$$\delta=x\sigma/\sqrt{n},$$

$$P(|X-a|<\delta)=2\Phi(\delta/\sigma). \tag{2.34}$$

Замінивши X на $\overline{X}$ та σ на $\sigma(\overline{X})=\sigma/\sqrt{n}$, одержимо:

$$P(|\overline{X}-a|<\delta)=2\Phi(\delta\sqrt{n}/\sigma)=2\Phi(x).$$

Враховуючи, що $\delta = x\sigma_{Г}/\sqrt{n}$, остання формула запишеться так:

$$P(|\overline{X}-a|<\frac{x\sigma}{\sqrt{n}})=2\Phi(x).$$

для (2.34) вона набирає вигляду:

$$P\left(\left|\frac{\overline{X}-a}{\frac{\sigma}{\sqrt{n}}}\right|<x\right)=2\Phi(x)=\gamma. \tag{2.35}$$

З рівності (2.35) знаходимо аргументи x, а саме:

$$2\Phi(x)=\gamma\rightarrow\Phi(x)=\frac{\gamma}{2}.$$

Аргумент x можна знайти за значенням функції Лапласа за допомоги відповідної таблиці (див. лекцію № 3).

Для того, щоби отримати робочу формулу, позначимо у формулі (2.35) середнє вибіркове через $\overline{x}_{В}$, тоді:

$$P(\overline{x_B}-\frac{x\cdot\sigma}{\sqrt{n}}<a<\overline{x_B}+\frac{x\cdot\sigma}{\sqrt{n}})=2\Phi(x)=\gamma.$$

Отже, довірчий інтервал дорівнюватиме:

$$[\overline{x_B}-\frac{x\cdot\sigma}{\sqrt{n}};\overline{x_B}+\frac{x\cdot\sigma}{\sqrt{n}}] \tag{2.36}$$

Величина $x\sigma_{Г}/\sqrt{n}$ називається *точністю оцінки* або *похибкою вибірки*.

Задача 2.13. Вимірявши 40 випадково відібраних після виготовлення деталей, знайшли вибіркове середнє, яке дорівнює 15 см. З надійністю $\gamma = 0{,}99$ побудувати довірчий інтервал для середньої величини всієї партії деталей, якщо генеральна дисперсія дорівнює 0,99 см^2.

Розв'язок

Для побудови довірчого інтервалу необхідно знати: $\overline{x}$, σ, n, x.
Згідно умови задачі:

$$\overline{x_B}=15\,см, \qquad \sigma=\sqrt{D}=\sqrt{0{,}99}=0{,}9949\,см, \qquad n=40.$$

Величина x обчислюється з рівняння:

$$\Phi(x)=\frac{\gamma}{2}=\frac{0{,}99}{2}=0{,}495\,;$$

$\Phi(x) = 0{,}495 \rightarrow x = 2{,}58$ (*таблиця значень функції Лапласа*).
Знайдемо числові значення кінців довірчого інтервалу:

$$\overline{x_B}-\frac{x\cdot\sigma}{\sqrt{n}}=15-\frac{2{,}56\cdot 0{,}9949}{\sqrt{40}}=15-0{,}4027=14{,}6\,см.$$

$$\overline{x_B}+\frac{x\cdot\sigma}{\sqrt{n}}=15+\frac{2{,}56\cdot 0{,}9949}{\sqrt{40}}=15+0{,}4027=15{,}4\,см.$$

Відтак маємо:

$$14{,}6<\overline{X}<15{,}4$$

Отже, з надійністю 0,99 (99% гарантії) оцінюваний параметр $\overline{X}$ перебуває всередині інтервалу [14,6; 15,4].

9.3.2. Довірчий інтервал для оцінки математичного сподівання нормального розподілу при невідомому середньоквадратичному відхиленні

Для малих вибірок, з якими ми стикаємося, досліджуючи різні ознаки наприклад в фізиці, математиці, економіці, техніці чи сільському господарстві, для оцінювання $\overline{X}_Г = a$ при невідомому значенні σ сукупності неможливо скористатися нормальним законом розподілу.

В такому випадку для побудови довірчого інтервалу застосовується наступна випадкова величина:

$$t=\frac{\overline{x_B}-a}{\frac{S}{\sqrt{n}}} \tag{2.37}$$

яка має розподіл Стьюдента з $k = n - 1$ степенями вільності.

Тоді (2.38) набуває наступного вигляду:

$$P\left(\left|\frac{\overline{x_B}-a}{\frac{S}{\sqrt{n}}}\right|<t_\gamma\right)=P\left(\overline{x_B}-\frac{t_\gamma\cdot S}{\sqrt{n}}<a<\overline{x_B}+\frac{t_\gamma\cdot S}{\sqrt{n}}\right)=2\int_0^{t_\gamma} f(t)\,dt=\gamma\,,$$

тут $f(t)$ – щільність розподілу Стьюдента, яка є парною функцією.

Розрахувавши за цим статистичним розподілом $\overline{x}_В$, S (виправлене вибіркове середньоквадратичне відхилення – стандарт) та визначивши за таблицею розподілу Стьюдента значення t_γ, будуємо довірчий інтервал:

Тут $t_\gamma(\gamma; n)$ – знаходимо згідно таблицею значень розподілу Стьюдента:

Таблиця 2.7

$n \backslash \gamma$	0,95	0,99	0,999	$n \backslash \gamma$	0,95	0,99	0,999
5	2,78	4,60	8,61	**20**	2,093	2,861	3,883
6	2,57	4,03	6,86	**25**	2,064	2,797	3,745
7	2,45	3,71	5,96	**30**	2,045	2,756	3,659
8	2,37	3,50	5,41	**35**	2,032	2,720	3,600
9	2,31	3,36	5.04	**40**	2,023	2,708	3,558
10	2,26	3,25	4,78	**45**	2,016	2,692	3,527
11	2,23	3,17	4,59	**50**	2,009	2,679	3,502
12	2,20	3,11	4,44	**60**	2,001	2,662	3,464
13	2,18	3,06	4,32	**70**	1,996	2,649	3,439
14	2,16	3,01	4,22	**80**	1,991	2,640	3,418
15	2,15	2,98	4,14	**90**	1,987	2,633	3,403
16	2,13	2,95	4,07	**100**	1,984	2,627	3,392
17	2,12	2,92	4,02	**120**	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Задача 2.14. Випадково вибрана партія з двадцяти приладів була випробувана, щодо терміну безвідмовної роботи (t_i, в годинах). Результати випробувань наведено у вигляді дискретного статистичного розподілу:

t_i	100	170	240	310	380
n_i	2	5	10	2	1

З надійністю γ = 0,99 побудувати довірчий інтервал для a (середнього часу безвідмовної роботи приладу).

Розв'язок

Для побудови довірчого інтервалу необхідно знайти середнє вибіркове і виправлене середньоквадратичне відхилення.

Обчислимо $\overline{x_B}$:

$$\overline{x_B}=\frac{1}{n}\sum_{i=1}^{m} n_i t=\frac{2\cdot 100+5\cdot 170+10\cdot 240+2\cdot 310+1\cdot 380}{20}=\frac{4450}{20}=222{,}5.$$

Визначимо D_B:

$$D_B=\frac{1}{n}\sum_{i=1}^{m} n_i t_i^2-(\overline{x})^2=\frac{2\cdot 100^2+5\cdot 170^2+10\cdot 240^2+2\cdot 310^2+1\cdot 380^2}{20}-222{,}5^2 =$$

$$= 53855-49506{,}25=4348{,}75.$$

Виправлене середньоквадратичне відхилення (стандарт) дорівнюватиме:

$$S=\sqrt{\frac{n}{n-1}D_B}=\sqrt{\frac{20}{19}\cdot 4348{,}75}\approx 67{,}66.$$

$$2\int_0^{t_\gamma} f(t)dt = \gamma = 0{,}99.$$

За таблицею значень розподілу Стьюдента, згідно заданої надійності γ = 0,99 та об'єму вибірки n = 20 знаходимо значення t_γ = 2,861.

Обчислимо кінці довірчого інтервалу:

$$\overline{x_B} - \frac{t_\gamma \cdot S}{\sqrt{n}} = 222{,}5 - \frac{2{,}861 \cdot 67{,}66}{\sqrt{20}} = 222{,}5 - 43{,}2847 = 179{,}2153\, год.$$

$$\overline{x_B} + \frac{t_\gamma \cdot S}{\sqrt{n}} = 222{,}5 + \frac{2{,}861 \cdot 67{,}66}{\sqrt{20}} = 222{,}5 + 43{,}2847 = 265{,}7847\, год.$$

Отже, з надійністю γ = 0,99 можна стверджувати, що $\overline{X}_Г = a$ буде міститися в інтервалі:

$$179{,}21 < a < 265{,}78.$$

При великих обсягах вибірки, а саме: n > 30, на підставі центральної граничної теореми теорії ймовірностей (теореми Ляпунова), розподіл Стьюдента наближається до нормального закону. У цьому разі t_γ знаходиться за таблицею значень функції Лапласа.

9.3.3. Довірчий інтервал для оцінки середньоквадратичного відхилення нормального розподілу

Нехай кількісна ознака X генеральної сукупності має нормальний закон розподілу. Необхідно оцінити невідоме генеральне середньоквадратичне відхилення σ по виправленому вибірковому середньоквадратичному відхиленню S (стандарту). Поставимо перед собою задачу знаходження довірчого інтервалу, який охоплюватиме шукане σ із заданою надійністю γ.

Нехай справедливі такі відношення:

$$P(|\sigma - S| < \delta) = \gamma, \qquad P(S - \delta < \sigma < S + \delta) = \gamma.$$

Для того, щоби можна було скористатись відповідною таблицею перетворимо подвійну нерівність

$$S - \delta < \sigma < S + \delta$$

в рівносильну нерівність

$$S(1 - \delta / S) < \sigma < S(1 + \delta / S).$$

Зробивши заміну $\delta/S = q$, одержимо:

$$S(1 - q) < \sigma < S(1 + q). \qquad (2.38)$$

Необхідно знайти q, із цією метою введемо в розподіл випадкову величину «хі»:

$$\chi = (S/\sigma)\sqrt{n-1},$$

де n – об'єм вибірки.

Величина $S^2(n - 1)/\sigma^2$ розподілена за законом χ^2 із $n - 1$ степенями вільності, тому корінь квадратний цієї величини позначають через χ.

Щільність розподілу χ має вигляд

$$R(\chi, n) = \frac{\chi^{n-2} e^{-\chi^2/2}}{2^{(n-3)/2}\Gamma\left(\frac{n-1}{2}\right)}. \quad (2.39)$$

Цей розподіл не залежить від оцінювального параметра σ, а тільки від об'єму вибірки n.

Перетворимо нерівність (2.38) так, щоби вона набуло виду: $\chi_1 < \chi < \chi_2$. Імовірність цієї нерівності рівна заданій точності, тобто:

$$\int_{\chi_1}^{\chi_2} R(\chi, n)\, d\chi = \gamma.$$

Припускаючи, що $q < 1$, перепишемо нерівність (2.38) наступним чином:

$$\frac{1}{S(1+q)} < \frac{1}{\sigma} < \frac{1}{S(1-q)}.$$

Помноживши всі члени нерівності на $S\sqrt{n}$-1, одержимо:

$$\frac{\sqrt{(n-1)}}{(1+q)} < \frac{S\sqrt{(n-1)}}{\sigma} < \frac{\sqrt{(n-1)}}{(1-q)}, \qquad \frac{\sqrt{(n-1)}}{(1+q)} < \chi < \frac{\sqrt{(n-1)}}{(1-q)}.$$

Імовірність того, що ця нерівність і, відповідно, рівносильна їй нерівність (2.38) буде виконуватись, дорівнює:

$$\int_{\sqrt{(n-1)}/(1+q)}^{\sqrt{(n-1)}/(1-q)} R(\chi, n)\, d\chi = \gamma.$$

За допомоги наведеного рівняння по відомих γ та n можна знайти q. На практиці для знаходження $q(\gamma, n)$, користуються такою таблицею:

Таблиця 2.8

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Розрахувавши по вибірці S та знайшовши по таблиці q, отримаємо шуканий довірчий інтервал:

$$S(1-q)<\sigma<S(1+q), \text{ якщо } q<1 \text{ та } 0<\sigma<S(1+q), \text{ якщо } q>1.$$

Задача 2.15. Кількісна ознака X розподілена за нормальним законом. Відповідно до вибірки об'ємом n = 25, знайдено виправлене середньоквадратичне відхилення S = 0,8 (стандарт). Знайти довірчий інтервал середнього квадратичного відхилення генеральної сукупності, з надійністю γ = 0,95.

Розв'язок

Згідно з умовою задачі, n = 25; S = 0,8; γ = 0,95. З відповідної таблиці знаходимо q = 0,32 (q < 1).

Шуканий довірчий інтервал такий

$$0{,}8(1-0{,}32)<\sigma<0{,}8(1+0{,}32), \quad 0{,}544<\sigma<1{,}056.$$

Лекція 10. Елементи теорії кореляції

Тема 10.1. Функціональна, статистична та кореляційна залежності

На практиці ми зустрічаємося з різними типами зв'язку між випадковими величинами. Одним з таких типів є *функціональна залежність*.

Така залежність виникає тоді, коли зв'язок між величинами настільки тісний, що, знаючи можливі значення однієї з величин, можна точна вказати можливі значення іншої. Функціональну залежність величини Y від величини X можна задати формулою

$$Y=f(X).$$

Нехай, наприклад, $Y = X^2$, можливі значення величини X : 1, 2, 3. Тоді можливі значення величини Y : 1, 4, 9.

До величин, що пов'язані функціональними залежностями, належить значна кількість тих, що зустрічаються у фізиці. Наприклад, тиск P, температура T та об'єм газу V пов'язані функціональною залежністю $P = RT/V$, де R – універсальна газова стала.

Водночас разом з цим точна функціональна залежність між величинами реалізується досить рідко, оскільки величини зазнають впливу багатьох випадкових факторів. Серед цих факторів можуть бути спільні для обох величин. У таких випадках виникають статистичні залежності.

Наприклад, якщо величина Y залежить від випадкових факторів Z_1, Z_2, Z_3, Z_4, а величина X – від випадкових факторів Z_3, Z_4, Z_5, то між Y та X існує статистична залежність – серед випадкових факторів, що діють на величини Y та X , є спільні – Z_3, Z_4.

Статистичною залежністю між величинами називають таку залежність, коли зміна однієї з величин викликає зміну розподілу іншої. Зокрема, якщо при зміні однієї з величин змінюється середнє значення іншої, то така залежність називається *кореляційною*.

Статистичні та кореляційні залежності знаходять численні застосування в усіх сферах природознавства, зокрема в геології. Наприклад, існує кореляційна залежність між органічними речовинами, наявних у ґрунті, з неорганічними мінеральними компонентами порід. Такого типу залежності лежать в основі багатьох методів підрахунку природничих ресурсів, на них спираються прогностичні побудови, розрахункові схеми.

Кореляційні зв'язки є окремим випадком статистичних. Термін «кореляція» походить від англійського слова correlation – співвідношення, відповідність. При таких залежностях певному значенню однієї з величин може відповідати одразу кілька значень іншої.

Для дослідження кореляційного зв'язку між величинами часто використовується математична модель, яка називається рівнянням регресії. При цьому дослідження складається з двох етапів:

1. Виявлення на підставі великої кількості спостережень того, як змінюється в середньому величина Y в залежності від величини X, тобто знаходження рівняння зв'язку між X та Y.
2. Виявлення ступеня взаємозв'язку явищ, що досліджуються.

Тема 10.2. Умовні середні та вибіркове рівняння регресії

Умовними математичними сподіваннями вважають умовні середні, які знаходять за даними спостереження.

Умовним середнім $\bar{y}_x$ називають середнє арифметичне значень Y, які відповідають $X = x$. Наприклад, якщо при $x_1 = 2$ величина Y набуває значення $y_1 = 5$, $y_2 = 6$, $y_3 = 10$, то умовне середнє $\bar{y}_x = (5+6+10)/3 = 7$.

Аналогічно визначається умовне середнє $\bar{x}_y$.

Умовним середнім $\bar{x}_y$ називають середнє арифметичне значень X, які відповідають $Y = y$.

Рівняння регресії Y на X та X на Y:

$$M(Y|x) = f(x), \quad M(X|y) = \varphi(y).$$

Умовне математичне сподівання $M(Y|x)$ – функція x, відповідно його оцінка, тобто умовне середнє $\bar{y}_x$ також є функцією від x; позначивши цю функцію через $f^*(x)$, одержимо рівняння:

$$\bar{y}_x = f^*(x).$$

Це рівняння називають *вибірковим рівнянням регресії* Y на X; функцію $f^*(x)$ – *вибірковою регресією* Y на X, а її графік – *вибірковою лінією регресії* Y на X. Аналогічно рівняння

$$\bar{x}_y = \varphi^*(y).$$

називають вибірковим рівнянням регресії X на Y ; функцію $\varphi^*(y)$ – вибірковою регресією X на Y, а її графік – вибірковою лінією регресії X на Y.

Регресія (англ. regression) – це форма зв'язку між випадковими величинами, закон зміни математичного сподівання однієї випадкової величини залежно від значень іншої.

Тема 10.3. Пошук параметрів вибіркового рівняння прямої лінії регресії за незгрупованими даними

Нехай проводиться вивчення системи якісних ознак (X, Y). У результаті n незалежних випробувань отримані n пар чисел: (x_1; y_1), (x_2; y_2), ..., (x_n; y_n).

Знайдемо за даними спостереженнями вибіркове рівняння прямої лінії середньоквадратичної регресії. Будемо шукати рівняння вигляду

$$\bar{y}_x = kx + b.$$

регресії Y на X.

Умовний коефіцієнт прямої лінії регресії Y на X називають вибірковим коефіцієнтом регресії Y на X та позначають через ρ_{yx}; він по суті являє собою оцінку коефіцієнта регресії (β).

Отже, будемо шукати вибіркове рівняння прямої лінії регресії Y на X вигляду:

$$Y = \rho_{yx} x + b. \tag{2.40}$$

Підберемо параметри ρ_{yx} та b так, щоби точки (x_1; y_1), (x_2; y_2), ..., (x_n; y_n), отримані за даними спостереженнями, на площині xOy лежали якмога ближче до прямої (2.40).

Уточнимо зміст цієї вимоги. Назвемо відхиленням таку різницю:

$$Y_i - y_i, \quad (i = 1, 2, \ldots, n),$$

де Y_i – розрахована згідно з рівняння (2.40) ордината, яка відповідає спостережуваному значенню x_i; y_i – спостережувана ордината, яка відповідає x_i.

Підберемо параметри ρ_{yx} та b так, щоби сума квадратів відхилень була мінімальною (в цьому і полягає суть методу найменших квадратів). Оскільки кожне відхилення залежить від шуканих параметрів, то і сума квадратів відхилень є функцією F цих параметрів (тимчасово замість ρ_{yx} будемо писати ρ):

$$F(\rho, b) = \sum_{i=1}^{n} (Y_i - y_i)^2, \quad \text{або} \quad F(\rho, b) = \sum_{i=1}^{n} (\rho x_i + b - y_i)^2.$$

Для знаходження мінімуму прирівняємо до нуля відповідні частини рівностей:

$$\frac{dF}{d\rho} = 2\sum_{i=1}^{n} (\rho x_i + b - y_i) x_i = 0;$$

$$\frac{dF}{db} = 2\sum_{i=1}^{n} (\rho x_i + b - y_i) = 0.$$

Виконуючи елементарні перетворення, одержимо систему двох лінійних рівнянь відносно ρ та b:

$$\left(\sum x^2\right)\rho + \left(\sum x\right)b = \sum xy; \quad \left(\sum x\right)\rho + nb = \sum y. \qquad (2.41)$$

Розв'язавши цю систему, знайдемо шукані параметри:

$$\rho_{yx} = \left(n\sum xy - \sum x \cdot \sum y\right) / \left(n\sum x^2 - \left(\sum x\right)^2\right); \qquad (2.42)$$

$$b = \left(\sum x^2 \cdot \sum y - \sum x \cdot \sum xy\right) / \left(n\sum x^2 - \left(\sum x\right)^2\right). \qquad (2.43)$$

Аналогічно можна отримати вибіркове рівняння прямої лінії регресії *X* на *Y*:

$$\bar{x}_y = \rho_{xy} x + C. \qquad (2.44)$$

де ρ_{xy} – вибірковий коефіцієнт регресії *X* на *Y*.

Задача 2.16. Знайдіть вибіркове рівняння прямої лінії регресії *Y* на *X* за такими даними спостережень:

x	1	1,5	3	4,5	5
y	1,25	1,4	1,5	1,75	2,25

Розв'язок

Знайдемо шукані параметри, для чого підставимо розраховані за таблицею суми у формули (2.42, 2.43):

$$\rho_{yx} = (5 \cdot 26{,}975 - 15 \cdot 8{,}15) / (5 \cdot 57{,}5 - 15^2) = 0{,}202;$$

$$b = (57{,}5 \cdot 8{,}15 - 15 \cdot 26{,}975) / 62{,}5 = 1{,}024.$$

Запишемо шукане рівняння регресії:

$$Y = 0{,}202\,x + 1{,}024.$$

Для того, щоби отримати уявлення про те, наскільки добре розраховані за даним рівнянням значення, Y_i узгоджується зі спостережуваними величинами y_i. Знайдемо Y_i - y_i. Результаті розрахунків наведемо в наступній таблиці:

x_i	Y_i	y_i	Y_i - y_i
1,00	1,226	1,25	-0,024
1,50	1,327	1,40	-0,073
3,00	1,630	1,50	0,130
4,50	1,933	1,75	0,183
5,00	2,034	2,25	-0,216

Як видно з таблиці, не всі відхилення достатньо малі. Це зумовлено малим об'ємом вибірки.

Тема 10.4. Кореляційна таблиця

При великій кількості спостережень одне й те саме значення x може зустрітися n_x раз, одне і те ж значення y – n_y разів, одна і та сама пара чисел (x, y) – n_{xy} разів. Тому дані спостережень групують, тобто, розраховують частоти n_x, n_y, n_{xy}. Згруповані типи даних записують у виді таблиці, яку називають *кореляційною*.

Таблиця 2.9

Y	X				
	10	20	30	40	n_y
0,4	5	–	7	14	26
0,6	–	2	6	4	12
0,8	3	19	–	–	22
n_x	8	21	13	18	n = 60

У першому рядку таблиці вказані спостережувані значення (10; 20; 30; 40) ознаки X, а в першому стовпчику – спостережувані значення (0,4; 0,6; 0,8) ознаки Y. На перетині рядків зі стовпчиками знаходяться частоти n_{xy}. Наприклад, частота 5 вказує на те, що пара чисел (10; 0,4) зустрічаються 5 разів. Усі частоти розташовані в прямокутнику, сторони якого виділені жирною лінією. Тире «–» означають, що відповідна пара чисел, наприклад (20; 0,4), не зустрічаються взагалі.

В останньому стовпчику записані суми частот рядків. Наприклад, сума частот першого рядку виділеного прямокутника дорівнює $n_y = 5 + 7 + 14 = 26$; це число вказує на те, що значення ознаки Y, дорівнює 0,4 (у співвідношенні до різних значень ознаки X) та зустрічається 26 разів.

В останньому рядку записані суми частот стовпчиків. Наприклад, число 8 вказує на те, що значення ознаки X дорівнює 10 (у співвідношенні до різних значень ознаки Y) та зустрічається 8 разів. В комірці, яка розташована у нижньому правому куті таблиці, розміщена сума всіх частот (загальна кількість усіх спостережень n). Очевидно,

$$\sum n_x = \sum n_y = n.$$

У нашій задачі:

$$\sum n_x = 8+21+13+18=60 \quad \text{або} \quad \sum n_y = 26+12+22=60.$$

Тема 10.5. Пошук параметрів вибіркового рівняння прямої лінії регресії за згрупованими даними

Для отримання параметрів рівняння прямої лінії регресії Y на X була одеражана система рівнянь:

$$\left.\begin{aligned} \left(\sum x^2\right)\rho_{yx} + \left(\sum x\right)b = \sum xy\,, \\ \left(\sum x\right)\rho_{yx} + nb = \sum y\,. \end{aligned}\right\} \qquad (2.45)$$

Припускалося, що значення X і відповідні їм значення Y спостерігались по одному разу. Нехай отримана більша кількість даних (для задоволення оцінки шуканих параметрів потрібно, щоби було хоча б 50 спостережень), серед яких і ті, що повторюються, які згруповані у вигляді кореляційної таблиці. Запишемо систему (2.45), так щоби вона відображала дані кореляційної таблиці. Скористаємося тотожностями:

$$\sum x = n\bar{x} \Rightarrow \bar{x} = \sum x/n;$$
$$\sum y = n\bar{y} \Rightarrow \bar{y} = \sum y/n;$$
$$\sum x^2 = n\overline{x^2} \Rightarrow \overline{x^2} = \sum x^2/n;$$
$$\sum xy = \sum n_{xy}xy.$$

Підставивши праві частини рівнянь в систему (2.45) та скоротивши обидві частини другого рівняння на n, отримаємо:

$$\left.\begin{array}{c} \left(n\overline{x^2}\right)\rho_{yx} + (n\bar{x})b = \sum n_{xy}xy, \\ (\bar{x})\rho_{yx} + b = \bar{y}. \end{array}\right\} \qquad (2.46)$$

Розв'язавши цю систему, знайдемо параметри ρ_{yx} та b, а отже, шукане рівняння:

$$\bar{y}_x = \rho_{yx}x + b. \qquad (2.47)$$

Однак більш доцільно ввести нову величину – вибірковий коефіцієнт кореляції, тобто записати рівняння регресії в іншому вигляді. Знайдемо з другого рівняння (2.46) параметр b:

$$b = \bar{y} - \rho_{yx}\bar{x}.$$

Підставивши праву частину цього рівняння у формулу (2.47), отримаємо:

$$\bar{y}_x - \bar{y} = \rho_{yx}(x - \bar{x}). \qquad (2.48)$$

Знайдемо коефіцієнт регресі, і враховуючи при цьому, що $\overline{x^2} - (\bar{x})^2 = \widetilde{\sigma}_x^2$:

$$\rho_{yx} = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n[\overline{x^2} - (\bar{x})^2]} = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n\widetilde{\sigma}_x^2}.$$

Помножимо обидві частини рівняння на дріб $\widetilde{\sigma}_x / \widetilde{\sigma}_y$:

$$\rho_{yx}\cdot\frac{\widetilde{\sigma}_x}{\widetilde{\sigma}_y} = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n\widetilde{\sigma}_x\widetilde{\sigma}_y}. \qquad (2.49)$$

Позначимо праву частину рівняння через r_B та назвемо її *вибірковим коефіцієнтом кореляції*.

$$r_B = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n\widetilde{\sigma}_x\widetilde{\sigma}_y}.$$

Підставивши r_B у формулу (2.49), одержимо:

$$\rho_{yx}\cdot\frac{\widetilde{\sigma}_x}{\widetilde{\sigma}_y} = r_B, \qquad \rho_{yx} = r_B\cdot\frac{\widetilde{\sigma}_y}{\widetilde{\sigma}_x}.$$

Підставивши праву частину даного рівняння в (2.48), остаточно отримаємо вибіркове рівняння прямої лінії регресії *Y* на *X*:

$$\bar{y}_x - \bar{y} = r_B \frac{\tilde{\sigma}_y}{\tilde{\sigma}_x}(x - \bar{x}).$$

Аналогічно знаходять вибіркове рівняння прямої лінії регресії *X* на *Y*:

$$\bar{x}_y - \bar{x} = r_B \frac{\tilde{\sigma}_x}{\tilde{\sigma}_y}(y - \bar{y}).$$

Рівняння вибіркових прямих регресії можна записати в більш симетричній формі:

$$\frac{\bar{y}_x - \bar{y}}{\tilde{\sigma}_y} = r_B \frac{x - \bar{x}}{\tilde{\sigma}_x}, \qquad \frac{\bar{x}_y - \bar{x}}{\tilde{\sigma}_x} = r_B \frac{y - \bar{y}}{\tilde{\sigma}_y}. \tag{2.50}$$

Зверніть увагу! Вибірковий коефіцієнт кореляції є оцінкою коефіцієнта кореляції:

$$r = \frac{\mu_{xy}}{\sigma_x \sigma_y} = \frac{M(XY) - M(X)M(Y)}{\sigma_x \sigma_y}. \tag{2.51}$$

Тема 10.6. Вибірковий коефіцієнт кореляції

Як було встановлено, вибірковий коефіцієнт кореляції визначається рівнянням

$$r_B = \frac{\sum n_{xy} xy - n\bar{x}\,\bar{y}}{n\tilde{\sigma}_x \tilde{\sigma}_y}, \tag{2.52}$$

де x, y – варіанти ознак *X* та *Y* відповідно; n_{xy} – частота пар варіантів (x, y); n – об'єм вибірки; $\tilde{\sigma}_x, \tilde{\sigma}_y$ – вибіркові середньоквадратичні відхилення, $\bar{x}, \bar{y}$ – вибіркові середні.

Відомо, що якщо величини *X* та *Y* незалежні, то коефіцієнт кореляції $r = 0$. Якщо $r = \pm 1$, то *X* та *Y* пов'язані між собою лінійною функціональною залежністю. Звідки випливає, що коефіцієнт кореляції r вимірює по суті, силу (або щільність) лінійного зв'язку між *Y* та *X*.

Вибірковий коефіцієнт кореляції r_B являє собою оцінку коефіцієнта кореляції r генеральної сукупності, і тому він також використовується для виміру сили лінійного зв'язку між величинами – кількісними ознаками *Y* і *X*.

Припустимо, що вибірковий коефіцієнт кореляцій по вибірці був знайдений зі значенням відмінним від нуля. Оскільки вибірка утворена у випадковий спосіб, то ще не можна зробити висновок, що коефіцієнт кореляції генеральної сукупності також відмінний від нуля.

Виникає необхідність перевірити гіпотезу про значущість вибіркового коефіцієнта кореляції. Якщо гіпотеза про рівність нулю генерального коефіцієнта кореляції буде відкинута, то вибірковий коефіцієнт кореляції буде значущим, а величини *X* та *Y* – корельованими; якщо гіпотеза буде прийнята, то вибірковий коефіцієнт кореляції незначущий, а величини *X* та *Y* некорельовані.

Якщо вибірка має достатньо великий об'єм і добре представляє генеральну сукупність (репрезентативна), то висновок стосовно щільності

лінійної залежності між ознаками, отриманий за даними вибірки, певною мірою може поширюватись на генеральну сукупність. Наприклад, для оцінки коефіцієнта кореляції $r_Г$ нормально розподіленої генеральної сукупності (при $n \geq 50$) можна скористатись формулою

$$r_B - 3\frac{1-r_B^2}{\sqrt{n}} \leq r_Г \leq r_B + 3\frac{1+r_B^2}{\sqrt{n}}. \qquad (2.53)$$

Зверніть увагу! Знак вибіркового коефіцієнта кореляції збігається зі знаками вибіркових коефіцієнтів регресії, що випливає зі формул

$$\rho_{yx} = r_B \frac{\tilde{\sigma}_y}{\tilde{\sigma}_x}, \qquad \rho_{xy} = r_B \frac{\tilde{\sigma}_x}{\tilde{\sigma}_y}.$$

Вибірковий коефіцієнт кореляції дорівнює середньому геометричному вибіркових коефіцієнтів регресії:

$$\rho_{yx}\rho_{xy} = r_B^2, \qquad r_B = \pm\sqrt{\rho_{xy}\rho_{yx}}. \qquad (2.54)$$

Знак при радикалі збігається зі знаком коефіцієнтів регресії.

Тема 10.7. Вибіркове кореляційне відношення та його властивості

Для оцінки щільності лінійного кореляційного зв'язку між ознаками у вибірці використовують вибірковий коефіцієнт кореляції. Для оцінки щільності нелінійного кореляційного зав'язку вводять такі характеристики:

$$\eta_{yx} - \textit{вибіркове клреляційне відношення } Y \textit{ на } X\,;$$

$$\eta_{xy} - \textit{вибіркове клреляційне відношення } X \textit{ на } Y\,;$$

Вибірковим кореляційним відношенням Y на X називають відношення міжгрупового середньоквадратичного відхилення до загального середньоквадратичного відхилення ознаки *Y*:

$$\eta_{yx} = \sigma_{міжгр} / \sigma_{заг},$$

або в інших позначеннях

$$\eta_{yx} = \sigma_{\bar{y}_x} / \tilde{\sigma}_y. \qquad (2.55)$$

Звідси:

$$\sigma_{\bar{y}_x} = \sqrt{D_{міжгр}} = \sqrt{\left(\sum n_x (\bar{y}_x - \bar{y})^2\right)/n};$$

$$\tilde{\sigma}_y = \sqrt{D_{заг}} = \sqrt{\left(\sum n_y (y - \bar{y})^2\right)/n},$$

де n – об'єм вибірки; n_x – частота значення x ознаки X; n_y – частота значення y ознаки Y; $\bar{y}$ – загальна середня ознака Y; $\bar{y}_x$ – умовна середня ознака Y.

Аналогічно можна визначити вибіркове кореляційне відношення X на Y:

$$\eta_{xy} = \sigma_{\bar{x}_y} / \tilde{\sigma}_x. \qquad (2.56)$$

Задача 2.17. Знайти η_{yx} за даними кореляційної таблиці:

Y	*X*			
	10	20	30	n_y
15 25	4 6	28 –	6 6	38 12
n_x	10	28	12	$n = 50$
$\bar{y}_x$	21	15	20	

Розв'язок

Знайдемо загальне середнє: $\bar{y}=(\sum n_{yx})/n=(38\cdot15+12\cdot25)/50=17{,}4.$

Знайдемо загальне середньоквадратичне відхилення:

$$\tilde{\sigma}_y=\sqrt{(\sum n_y(y-\bar{y})^2)/n}=\sqrt{(38(15-17{,}4)^2+12(25-17{,}4)^2)/50}=4{,}27.$$

Знайдемо міжгрупове середньоквадратичне відхиленням:

$$\sigma_{\bar{y}_x}=\sqrt{(\sum n_x(\bar{y}_x-\bar{y})^2)/n}\ =$$

$$=\ \sqrt{(10(21-17{,}4)^2+28(15-17{,}4)^2+12(20-17{,}4)^2)/50}=2{,}73.$$

Шукане кореляційне відношення: $\eta_{yx}=\sigma_{\bar{y}_x}/\tilde{\sigma}_y=2{,}73/4{,}27=0{,}64.$

Властивості вибіркового кореляційного відношення:

1. Кореляційне відношення задовольняє нерівність

$$0\le\eta\le1.$$

2. Якщо $\eta = 0$, то ознака Y з ознакою X кореляційною залежністю не пов'язанні.
3. Якщо $\eta = 1$, то ознака Y пов'язана з ознакою X функціональною залежністю.
4. Вибіркове кореляційне відношення не менше абсолютної величини вибіркового коефіцієнта кореляції:

$$\eta\ge|r_B|.$$

5. Якщо вибіркове кореляційне відношення дорівнює за абсолютною величиною вибірковому коефіцієнту кореляції, то має місце точна лінійна кореляційна залежність.

Тема 10.8. Задача криволінійної кореляції

Криволінійний – це зв'язок, при якому рівномірна зміна ознаки однієї величини призводить до нерівномірної зміни ознаки іншої.

При дослідженні криволінійних зв'язків принципове значення має вибір форми та рівняння зв'язку, яке якмога точніше буде відтворювати наявний зв'язок. Для вирішення даного завдання використовуються ті ж самі принципи, що й при лінійному зв'язку.

Якщо графік регресії

$$\bar{y}_x=f(x)\quad \text{або}\quad \bar{x}_y=\varphi(y)$$

зображається кривою лінією, то кореляцію називають *криволінійною*.

Наприклад, функції регресії Y на X можуть мати вигляд

$$\bar{y}_x = ax^2 + bx + c\,(\textit{параболічна кореляція}\ 2-\textit{го порядку});$$

$$\bar{y}_x = ax^3 + bx^2 + cx + d\,(\textit{параболічна кореляція}\ 3-\textit{го порядку});$$

Для визначення виду функції регресії будують точки $(x;\bar{y}_x)$, за розміщенням яких роблять висновок стосовно приблизного виду функції регресії; при кінцевому розв'язку беруть до уваги особливості, які випливають зі змісту розв'язуваної задачі.

Теорія криволінійної кореляції розв'язує ті ж задачі, що й теорія лінійної кореляції (встановлення форми та щільності кореляційного зв'язку). Невідомі параметри рівняння регресії шукають методом найменших квадратів. Для оцінки щільності криволінійної кореляції використовують вибіркове кореляційне відношення.

Розглянемо задачу параболічної кореляції 2-го порядку, припустивши що дані *n* спостережень (вибірки) дозволяють нам стверджувати, що має місце така кореляція. В такому випадку вибіркове рівняння регресії *Y* на *X* має вигляд

$$\bar{y}_x = Ax^2 + Bx + C, \tag{2.57}$$

де *A*, *B* та *C* – невідомі параметри.

Користуючись методом найменших квадратів, одержимо систему лінійних рівнянь відносно невідомих параметрів:

$$\left.\begin{array}{c}(\sum n_x x^4)A + (\sum n_x x^3)B + (\sum n_x x^2)C = \sum n_x \bar{y}_x x^2;\\ (\sum n_x x^3)A + (\sum n_x x^2)B + (\sum n_x x)C = \sum n_x \bar{y}_x x;\\ (\sum n_x x^2)A + (\sum n_x x)B + nC = \sum n_x \bar{y}_x.\end{array}\right\} \tag{2.58}$$

Знайдені із цієї системи рівнянь параметри *A*, *B* та *C* підставляють у рівняння (2.57) та отримують шукане рівняння регресії.

Задача 2.18. Знайти вибіркове рівняння регресії *Y* на *X* вигляду

$$\bar{y}_x = Ax^2 + Bx + C.$$

За даними кореляційної таблиці:

Y	*X*			
	1	1,1	1,2	n_y
6	8	2	–	10
7	–	30	–	30
7,5	–	1	9	10
n_x	8	33	9	$n = 50$
$\bar{y}_x$	6	6,73	7,5	

Розв'язок

Складемо розрахункову таблицю:

x	n_x	$\bar{y}_x$	$n_x x$	$n_x x^2$	$n_x x^3$	$n_x x^4$	$n_x \bar{y}_x$	$n_x \bar{y}_x x$	$n_x \bar{y}_x x^2$
1	8	6	8	8	8	8	48	48	48
1,1	33	6,73	36,3	39,93	43,93	48,32	222,09	244,30	268,73
1,2	9	7,5	10,8	12,96	15,55	18,66	67,5	81	97,2
Σ	50	–	55,1	60,89	67,48	74,98	337,59	373,3	413,93

Підставивши числа (суми) нижнього рядка розрахункової таблиці, одержимо таку систему рівнянь:

$$\left.\begin{array}{c} 74{,}98\,A+67{,}48\,B+60{,}89\,C=413{,}93\,; \\ 67{,}48\,A+60{,}89\,B+55{,}10\,C=373{,}30\,; \\ 60{,}89\,A+55{,}10\,B+50\,C=337{,}59. \end{array}\right\}$$

Розв'язавши цю систему, знаходимо: A = 1,94; B = 2,98; C = 1,10. Записуємо кінцеве рівняння регресії: $\bar{y}_x=1{,}94\,x^2+2{,}98\,x+1{,}10.$

Тема 10.9. Множинна кореляція

Досі ми розглядали кореляційний зв'язок між двома ознаками. Якщо ж використовується зв'язок між кількома ознаками, то таку кореляцію називають *множинною*.

У найпростішому випадку кількість ознак дорівнює трьом і зв'язок між ними – лінійний:

$$z = ax + by + c.$$

В такому випадку виникають задачі:

1. Знайти за даними спостереження вибіркове рівняння зв'язку вигляду:

$$z = Ax + By + C. \qquad (2.59)$$

2. Оцінити силу зв'язку між Z та X, Y.
3. Оцінити силу зв'язку між Z та X (при сталій ознаці Y) і між Z та Y (при сталій ознаці X).

Перша задача розв'язується за допомогою методу найменших квадратів, причому замість рівняння (2.59) зручніше шукати рівняння зв'язку вигляду:

$$z - \bar{z} = A\left(x - \bar{x}\right) + B\left(y - \bar{y}\right), \qquad (2.60)$$

де

$$A = \frac{r_{xz} - r_{yz} r_{xy}}{1 - r_{xy}^2} \cdot \frac{\sigma_z}{\sigma_x}; \qquad B = \frac{r_{yz} - r_{xz} r_{xy}}{1 - r_{xy}^2} \cdot \frac{\sigma_z}{\sigma_y}.$$

Тут r_{xz}, r_{yz}, r_{xy} – коефіцієнти кореляції між ознаками X і Z, Y і Z та X і Y відповідно; σ_x, σ_y, σ_z – середньоквадратичні відхилення.

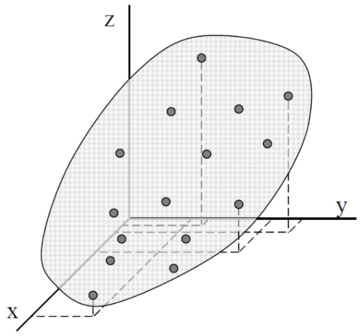


Рис. 2.1. Площина множинної регресії

Сила зв'язку Z з ознаками X, Y оцінюється *вибірковим сукупним коефіцієнтом кореляції*:

$$R = \sqrt{\frac{r_{xz}^2 - 2 r_{xy} r_{xz} r_{yz} + r_{yz}^2}{1 - r_{xy}^2}}, \qquad (2.61)$$

причому $0 \le R \le 1$.

Сила зв'язку між Z і X, (при сталому Y) та між Z і Y (при сталому X), оцінюється відповідними *частинними вибірковими коефіцієнтами кореляції*:

$$r_{xz(y)} = \frac{r_{xz} - r_{xy} r_{yz}}{\sqrt{\left(1 - r_{xy}^2\right)\left(1 - r_{yz}^2\right)}}; \qquad r_{yz(x)} = \frac{r_{yz} - r_{xy} r_{xz}}{\sqrt{\left(1 - r_{xy}^2\right)\left(1 - r_{xz}^2\right)}}. \qquad (2.62)$$

Ці коефіцієнти мають ті ж властивості та зміст, що й звичайний вибірковий коефіцієнт кореляції, тобто служать для оцінки лінійного зв'язку між ознаками.

Лекція 11. Метод Монте-Карло

Тема 11.1. Предмет методу Монте-Карло

Датою народження методу Монте-Карло прийнято вважати 1949 р., коли американські вчені Ніколас Метрополіс та Станіслав Улям опублікували статтю: «Метод Монте-Карло», в якій систематично його описали.

Назва методу пов'язана з назвою міста Монте-Карло (Монако), де в так званих гральних будинках (казино) грають у рулетку – один із найпростіших пристроїв отримання випадкових чисел, на використанні яких і заснований цей метод.

ЕОМ дозволяють з легкістю отримати так званні псевдовипадкові числа (при розв'язку задач їх вважають випадковими числами); що у свою чергу привело до широкої інтеграції даного методу в багатьох галузях науки та техніки (статистична фізика, теорія масового обслуговування, теорія ігор та інші).

Метод Монте-Карло використовується для розв'язку інтегралів, особливо багатовимірних, для розв'язку систем алгебраїчних рівнянь високого порядку, при дослідженні різного роду складних систем (автоматичного керування, економічних, біологічних та інших).

Предмет (*зміст*) *методу Монте-Карло* полягає в такому: необхідно знайти значення (число) a, деякої досліджуваної величини. Для цього випадкову величину X обирають так, щоби її математичне сподівання дорівнювало значенню a, тобто

$$M(X)=a. \tag{2.63}$$

На практиці діють таким чином: здійснюють n випробувань, в результаті яких отримують n можливих значень величини X; розраховують середнє арифметичне:

$$\bar{x}=\sum_{i=1}^{n}\frac{x_i}{n}, \tag{2.64}$$

та вважають отримане значення оцінкою (наближеним значенням) a^* шуканого числа a:

$$a^*\sim a=\bar{x}. \tag{2.65}$$

Оскільки метод Монте-Карло потребує виконання значного числа випробувань, його ще часто називають *методом статистичних випробувань*.

Теорія цього методу вказує, як доцільніше вибрати випадкову величину X та як знайти її можливі значення. Зокрема, розробляють методи зменшення дисперсії використовуваних випадкових величин, в результаті чого зменшується похибка, яка виникає при заміні шуканого математичного сподівання a на його оцінку a^*.

Пошук можливих значень випадкової величини X (моделюванням) називається «*розіграш випадкової величини*». Розглянемо деякі способи

(методи) розіграшу випадкових величин і з'ясуємо, як при цьому оцінити допустиму похибку.

Тема 11.2. Оцінка похибки методу Монте-Карло

Нехай для одержання оцінки a^* математичного сподівання а випадкової величини X було здійснено n незалежних випробувань (розіграно n можливих значень X) і за ними ж було знайдено вибіркове середнє x, яке взято як шукана оцінка $a^* = \overline{x}$.

Зрозуміло, що якщо повторити дослід, то будуть отримані інші значення X, відповідне інше середнє та оцінка. Звідси можна зробити висновок, що отримати точну оцінку математичного сподівання неможливо. Природно, виникає питання стосовно величини допустимої похибки.

Обмежимося пошуком тільки верхньої межі допустимої похибки δ із заданою ймовірністю (надійністю) γ:

$$P\left(\left|\overline{X} - a\right| \le \delta\right) = \gamma .$$

Шукана верхня границя похибки δ є не чим іншим, як «точністю оцінки» математичного сподівання по вибірковому середньому за допомогою довірчих інтервалів, про які мова йшла на попередніх лекціях (див. лекція № 9). Тому скористаємось отриманими раніше результатами та розглянемо три можливих випадки.

1. *Випадкова величина X розподілена нормально і її середньоквадратичне відхилення σ відоме.* В даному випадку з надійністю γ верхню границю похибки δ можна знайти за формулою:

$$\delta = \frac{t \cdot \sigma}{\sqrt{n}}, \tag{2.66}$$

де n – кількість випробувань (розіграних значень X); t – значення аргументу функції Лапласа, при якому $\Phi(t) = \gamma/2$; σ – відоме середньоквадратичне відхилення X.

Задача 2.19. З надійністю $\gamma = 0{,}95$ знайдіть верхню границю похибки δ, якщо для оцінки математичного сподівання нормальної величини X із відомим середньоквадратичним відхиленням $\sigma = 0{,}5$ було розіграно 100 можливих її значень.

Розв'язок

Згідно з умовою задачі, $n = 100$, $\sigma = 0{,}5$, $\Phi(t) = 0{,}95/2 = 0{,}475$. Використовуючи таблицю значень функції Лапласа (див. лекцію № 3), знаходимо: $t = 1{,}96$. Тоді

$$\delta = \frac{t \cdot \sigma}{\sqrt{n}} = \frac{1{,}96 \cdot 0{,}5}{\sqrt{100}} = 0{,}098.$$

2. *Випадкова величина X розподілена нормально, але її середньоквадратичне відхилення σ невідоме.* В цьому випадку з надійністю γ верхню границю похибки δ знаходять за формулою

$$\delta = \frac{t_\gamma \cdot s}{\sqrt{n}}. \tag{2.67}$$

де n – кількість випробувань (розіграних значень X); s – «виправлене» середньоквадратичне відхилення; t_γ – значення функції, яке можна знайти за допомогою таблиці розподілу Стьюдента (див. лекцію № 9).

Задача 2.20. З надійністю $\gamma = 0,95$, знайдіть верхню границю похибки δ, якщо для оцінки математичного сподівання нормальної величини X, було розіграно 100 її можливих значень, і по ним розраховано «виправлене» середньоквадратичне відхилення $s = 0,5$.

Розв'язок

Згідно з умовою задачі, $n = 100$, $s = 0,5$. Використовуючи таблицю значень функції t_γ (див. лекцію № 9), для $\gamma = 0,95$ знаходимо $t_\gamma = 1,984$. Тоді

$$\delta = \frac{t_\gamma \cdot s}{\sqrt{n}} = \frac{1,984 \cdot 0,5}{\sqrt{100}} = 0,099.$$

3. *Випадкова величина X розподілена за законом відмінним від нормального*.

У такому випадку, при досить великій кількості випробувань ($n > 30$), з надійністю, яка приблизно дорівнює γ, верхня границя похибки може бути розрахована за формулою (2.66), якщо середньоквадратичне відхилення σ величини X відоме, якщо ж σ невідоме, то можна підставити у формулу (2.66) його оцінку s («виправлене» середньоквадратичне відхилення) або скористатись формулою (2.67).

Зауважимо, що чим більше число n, тим менше відмінностей між результатами, які дають обидві формули. Це пояснюється тим, що при $n \to \infty$ розподіл Стьюдента наближається до нормального.

Зверніть увагу! Для того щоби знайти найменшу кількість випробувань, які забезпечать наперед задану верхню границю похибки δ, необхідно виразити n за допомогою формул (2.66) та (2.67), тобто

$$n = \frac{t^2 \cdot \sigma^2}{\delta^2}, \qquad n = \frac{t_\gamma^2 \cdot s^2}{\delta^2}. \qquad (2.68)$$

Тема 11.3. Випадкові числа

Раніше було сказано, що метод Монте-Карло заснований на застосуванні або використані випадкових чисел. Наведемо визначення цих чисел. Позначимо через R неперервну випадкову величину, яка розподілена рівномірно на проміжку (0, 1).

Випадковими числами називають можливі значення r неперервної випадкової величини R, розподіленої рівномірно на інтервалі (0, 1).

Насправді використовують не рівномірно розподілену величину R, можливі значення якої, загалом можуть мати нескінченну кількість десяткових знаків, а *квазірівномірну випадкову величину* R^*, можливі значення якої мають кінцеву кількість десяткових знаків.

У результаті заміни R на R^* модельована величина має не точний, а наближений заданий розподіл. Нижче наведені сукупність випадкових чисел, запозичена зі книги: [Большев Л. Н., Смирнов Н. В. Таблицы математической статистики : Наука, 1965, 428 с.]

10 09 73 25 33	76 52 01 35 86	34 67 35 48 76	80 95 90 91 17
37 54 20 48 05	64 89 47 42 96	24 80 52 40 37	20 63 61 04 02
08 42 26 89 53	19 64 50 93 03	23 20 90 25 60	15 95 33 47 64
99 01 90 25 29	09 37 67 07 15	38 31 13 11 65	88 67 67 43 97
12 80 79 99 70	80 15 73 61 47	64 03 23 66 53	98 95 11 68 77
66 06 57 47 17	34 07 27 68 50	36 69 73 61 70	65 81 33 98 85
31 06 01 08 05	45 57 18 24 06	35 30 34 26 14	86 79 90 74 39
85 26 97 76 02	02 05 16 56 92	68 66 57 48 18	73 05 38 52 47
63 57 33 21 35	05 32 54 70 48	90 55 35 75 48	28 46 82 87 09
73 79 64 57 53	03 52 96 47 78	35 80 83 42 82	60 93 52 03 44
98 52 01 77 67	14 90 56 86 07	22 10 94 05 58	60 97 09 34 33
11 80 50 54 31	39 80 82 77 32	50 72 56 82 48	29 40 52 42 01
83 45 29 96 34	06 28 89 80 83	13 74 67 00 78	18 47 54 06 10
88 68 54 02 00	86 50 75 84 01	36 76 66 79 51	90 36 47 64 93
99 59 46 73 48	87 51 76 49 69	91 82 60 89 28	93 78 56 13 68
65 48 11 76 74	17 46 85 09 50	58 04 77 69 74	73 03 95 71 86
80 12 43 56 35	17 72 70 80 15	45 31 82 23 74	21 11 57 82 53
74 35 09 98 17	77 40 27 72 14	43 23 60 02 10	45 52 16 42 37
69 91 62 68 03	66 25 22 91 48	36 93 68 72 03	76 62 11 39 90
09 89 32 05 05	14 22 56 85 14	46 42 75 67 88	96 29 77 88 22
91 49 91 45 23	68 47 92 76 86	46 16 28 35 54	94 75 08 99 23
80 33 69 45 98	26 94 03 68 58	70 29 73 41 35	53 14 03 33 40
44 10 48 19 49	85 15 74 79 54	32 97 92 65 75	57 60 04 08 81
12 55 07 37 42	11 10 00 20 40	12 86 07 46 97	96 64 48 94 39
63 60 64 93 29	16 50 53 44 84	40 21 95 25 63	43 65 17 70 82

61 19 69 04 46	26 45 74 77 74	51 92 43 37 29	65 39 45 95 93
15 47 44 52 66	95 27 07 99 53	59 36 78 38 48	82 39 61 01 18
94 55 72 85 73	67 89 75 43 87	54 62 24 44 31	91 19 04 25 92
42 48 11 62 13	97 34 40 87 21	16 86 84 87 67	03 07 11 20 59
23 52 37 83 17	73 20 88 98 37	68 93 59 14 16	26 25 22 96 63
04 49 35 24 94	75 24 63 38 24	45 86 25 10 25	61 96 27 93 35
00 54 99 76 54	64 05 18 81 59	96 11 96 38 96	54 69 28 23 91
35 96 31 53 07	26 89 80 93 54	33 35 13 54 62	77 97 45 00 24
59 80 80 83 91	45 42 72 68 42	83 60 94 97 00	13 02 12 48 92
46 05 88 52 36	01 39 09 22 86	77 28 14 40 77	93 91 08 36 47
32 17 90 05 97	87 37 92 52 41	05 56 70 70 07	86 74 31 71 57
69 23 46 14 06	20 11 74 52 04	15 95 66 00 00	18 74 39 24 23
19 56 54 14 30	01 75 87 53 79	40 41 92 15 85	66 67 43 68 06
45 15 51 49 38	19 47 60 72 46	43 66 79 45 43	59 04 79 00 33
94 86 43 19 94	36 16 81 08 51	34 88 88 15 53	01 54 03 54 56
98 08 62 48 26	45 24 02 84 04	44 99 90 88 96	39 09 47 34 07
33 18 51 62 32	41 94 15 09 49	89 43 54 85 81	88 69 54 19 94
80 95 10 04 06	96 38 27 07 74	20 15 12 33 87	25 01 62 52 98
79 75 24 91 40	71 96 12 82 96	69 86 10 25 91	74 85 22 05 39
18 63 33 25 37	98 14 50 65 71	31 01 02 46 74	05 45 56 14 27
74 02 94 39 02	77 55 73 22 70	97 79 01 71 19	52 52 75 80 21
54 17 84 56 11	80 99 33 71 43	05 33 51 29 69	56 12 71 92 55
11 66 44 98 83	52 07 98 48 27	59 38 17 15 39	09 97 33 34 40
48 32 47 79 28	31 24 96 47 10	02 29 53 68 70	32 30 75 75 46
69 07 49 41 38	87 63 79 19 76	35 58 40 44 01	10 51 82 16 15

Рис. 2.2

Тема 11.4. Розігрування дискретної випадкової величини

Нехай потрібно змоделювати (розіграти) дискретну випадкову величину X, тобто отримати послідовність її можливих значень x_i ($i = 1, 2, ..., n$), знаючи при цьому закон розподілу X:

X	x_1	x_2	x_3	...	x_n
p	p_1	p_2	p_3	...	p_n

Позначимо через R неперервну випадкову величину, яка розподілена рівномірно на інтервалі (0, 1), а через r_j ($j = 1, 2, ...$) – її можливі значення, тобто випадкові числа.

Розіб'ємо інтервал $0 \le R < 1$ на осі Or точками з координатами p_1, $p_1 + p_2$, $p_1 + p_2 + p_3$, ..., $p_1 + p_2 + p_3 + ... + p_{n-1}$ на n окремих інтервалів Δ_1, Δ_2, ..., Δ_n, тобто

$$\Delta_1 = p_1 - 0 = p_1, \quad \Delta_2 = (p_1 + p_2) - p_1 = p_2, \quad ... ,$$

$$\Delta_n = 1 - (p_1 + p_2 + ... + p_{n-1}) = p_n.$$

Як видно, довжина частинного інтервалу з індексом дорівнює ймовірності з тим самим індексом:

$$\Delta_i = p_i. \tag{2.69}$$

Теорема: Якщо кожному випадковому числу r_j ($0 \le r_j < 1$), яке попало в інтервал Δ_i, поставити у відповідність можливі значення x_i, то розігрувана випадкова величина буде мати заданий закон розподілу:

X	x_1	x_2	...	x_n
p	p_1	p_2	...	p_n

Доведення

Оскільки при попаданні випадкового числа r_j в частинний інтервал Δ_i, величина, що розігрується, набуває можливі значення x_i, а таких інтервалів всього n, то величина, що розігрується, має ті ж самі можливі значення, що й X, а точніше $x_1, x_2, ..., x_n$.

Ймовірність потрапляння випадкової величини R в інтервал Δ_i дорівнює його довжині, в наслідок формули (2.69). Таким чином, імовірність попадання випадкової величини R в інтервал Δ_i дорівнює p_i.

Відповідно, ймовірність того, що розігрувана величина прийме можливі значення x_i, також дорівнює p_i (оскільки ми домовились, що у випадку попадання випадкового числа r_j в інтервал Δ_i вважати, що розігрувана величина набула можливе значення x_i). Отже, розігрувана випадкова величина має заданий закон розподілу.

Правило 2.1. Для того щоби розіграти (з моделювати) дискретну випадкову величину, яка задана законом розподілу

X	x_1	x_2	x_3	...	x_n
p	p_1	p_2	p_3	...	p_n

Необхідно:

1. Розбити інтервал (0, 1) осі Or на n частинних інтервалів: $\Delta_1 - (0, p_1)$, $\Delta_2 - (p_1; p_1 + p_2)$, ..., $\Delta_n - (p_1 + p_2 + ... + p_{n-1}; 1)$.
2. Вибрати (наприклад, із таблиці випадкових чисел) випадкове число r_j.
3. Якщо r_j попало в частинний інтервал Δ_i, то розігрувана дискретна випадкова величина набула можливе значення x_i.

Задача 2.21. Розіграйте 8 можливих значень дискретної випадкової величини X, із наступним законом розподілу:

X	3	11	24
p	0,25	0,16	0,59

Розв'язок

Розіб'ємо інтервал (0, 1) осі Or на 3 частинних інтервали: (0; 0,25), (0,25; 0,41) та (0,41; 1).

Виберемо довільно з таблиці випадкових чисел 8 випадкових чисел, наприклад: 0,10; 0,37; 0,08; 0,99; 0,12; 0,66; 0,31; 0,85.

Випадкове число $r_1 = 0{,}10$ належить частинному інтервалу Δ_1, тому розігрувана дискретна випадкова величина набула можливе значення $x_1 = 3$.

Випадкове число $r_2 = 0{,}37$ належить частинному інтервалу Δ_2, тому розігрувана дискретна випадкова величина набула можливе значення $x_2 = 11$. Аналогічно одержимо решту можливих значень.

Отже, можливі розігруванні значення X такі:

3, 11, 3, 24, 3, 24, 11, 24.

Зверніть увагу! В подальшому буде показано, що розігрування *подій* можна звести до розігрування значень *дискретної випадкової величини*. Спочатку розглянемо повну групу подій, яка складається з двох подій, а потім з n подій. Зрозуміло, що група з двох подій є окремим випадком повної групи n подій.

Тема 11.5. Розігрування протилежних подій

Нехай потрібно розіграти випробування, в кожному з яких подія A з'являється з відомою ймовірністю p, і відповідно, не з'являється з ймовірністю $q = 1 - p$.

Уведемо дискретну випадкову величину X з двома можливими значеннями ($x_1 = 1$, $x_2 = 0$) і відповідні їм ймовірності $p_1 = p$, $p_2 = q$. Домовимось, що якщо у випробуваннях величина X набула можливе значення $x_1 = 1$, то подія A відбулась, якщо $X = x_2 = 0$, то подія A не відбулась, тобто відбулась протилежна подія $\bar{A}$.

Отже, розіграш протилежних подій A і $\bar{A}$ зведено до розігрування дискретної випадкової величини X із заданим законом розподілу:

X	1	0
p	p	q

Для розіграшу X потрібно інтервал (0, 1) розбити точкою р на два частинних інтервали: Δ_1 – (0; p) та Δ_2 – (p; 1). Потім вибрати випадкове число r_j. Якщо r_j попадає в інтервал Δ_1, то $X = x_1$, (подія A відбулась); якщо r_j попадає в інтервал Δ_2, то $X = x_2 = 0$ (подія A не відбулась).

Задача 2.22. Розіграйте 6 випробувань, в кожному з яких подія A з'являється з ймовірністю $p = 0{,}35$.

Розв'язок

Виберемо з таблиці випадкових чисел такі числа: 0,10; 0,36: 0,08; 0,99; 012; 0,06.

Вважаючи, що при $r_j < 0{,}35$ подія A відбулась, а при $r_j \geq 0{,}35$ відбулась протилежна подія $\bar{A}$, отримаємо шукану послідовність подій:

$$A, \bar{A}, A, \bar{A}, A, A.$$

Тема 11.6. Розігрування повної групи подій

Розігрування повної групи n ($n > 2$) несумісних подій A_1, A_2, ..., A_n, ймовірності яких p_1, p_2, ..., p_n відомі, можна звести до розігрування дискретної випадкової величини X із таким законом розподілу:

X	1	2	3	...	n
p	p_1	p_2	p_3	...	p_n

Достатня умова є: якщо у випробуваннях величина X набула значення $x_i = i$ ($i = 1, 2, ..., n$), то відбулась подія A_i. Справедливість цього твердження

випливає із того, що кількість n можливих значень X дорівнює числу подій повної групи, а ймовірності можливих значень x_i та відповідних їм подій A_i однакові: $P(X = x_i) = P(A_i) = p_i$.

Отже, поява у випробуванні події A рівносильна події, в якій дискретна випадкова величина X набула можливе значення x_i.

Задача 2.23. Події A і B незалежні та сумісні. Розіграйте 6 випробувань, в кожному з яких імовірність появи події A дорівнює 0,6, а події B – 0,2.

Розв'язок

Можливі 4 результати випробувань:

$$A_1 = AB \quad \text{при чому} \quad P(AB) = P(A) \cdot P(B) = 0{,}6 \cdot 0{,}2 = 0{,}12;$$

$$A_2 = A\overline{B} \quad \text{при чому} \quad P(A\overline{B}) = P(A) \cdot P(\overline{B}) = 0{,}6 \cdot 0{,}8 = 0{,}48;$$

$$A_3 = \overline{A}B \quad \text{при чому} \quad P(\overline{A}B) = P(\overline{A}) \cdot P(B) = 0{,}4 \cdot 0{,}2 = 0{,}08;$$

$$A_4 = \overline{AB} \quad \text{при чому} \quad P(\overline{AB}) = P(\overline{A}) \cdot P(\overline{B}) = 0{,}4 \cdot 0{,}8 = 0{,}32.$$

Ця задача зводиться до розіграшу дискретної випадкової величини X за такий законом розподілу:

X	1	2	3	4
p	0,12	0,48	0,08	0,32

Виберемо 6 випадкових чисел, наприклад: 0,45; 0,65; 0,06; 0,59; 0,33; 0,70.

Побудуємо частинні інтервали: Δ_1 – (0; 0,12), Δ_2 – (0,12; 0,60), Δ_3 – (0,60; 0,68), Δ_4 – (0,68; 1).

Випадкове число r_1 = 0,45 належить інтервалу Δ_2, тому відповідно відбудеться подія $A_2 = \overline{A}B$. Аналогічно знайдемо результати решти випробувань.

Отже, шукана послідовність результатів розіграшу випробувань така:

$$A\overline{B}, \overline{A}B, AB, A\overline{B}, A\overline{B}, \overline{AB}.$$

Тема 11.7. Розігрування неперервної випадкової величини

Нехай потрібно розіграти (змоделювати) неперервну випадкову величину X, тобто отримати послідовність її можливих значень x_i (i = 1, 2, ...), знаючи при цьому функцію розподілу $F(x)$.

Теорема: Якщо r_i – випадкове число, то можливі значення x_i неперервної випадкової величини X, що розігрується, із заданою функцією розподілу $F(x)$, яка відповідає r_i, є коренями рівняння

$$F(x_i) = r_i. \tag{2.70}$$

Дійсно, поза як функція $F(x)$ – монотонно зростаюча в інтервалі всіх можливих значень X, то в цьому інтервалі більшим значенням аргументів відповідають більші значення функцій, і навпаки. Тому якщо $c < x_i < d$, то $F(c) < r_i < F(d)$, і навпаки.

З даних нерівностей випливає, що якщо випадкова величина ξ замкнена в інтервалі

$$c<\xi<d, \tag{2.71}$$

то і випадкова величина R буде замкнутою в інтервалі

$$F(c)<R<F(d), \tag{2.72}$$

та навпаки. Таким чином, нерівності (2.71) і (2.72) рівносильні, а отже і рівноймовірні:

$$P(c<\xi<d)=P[F(c)<R<F(d)]. \tag{2.73}$$

Оскільки величина R розподілена рівномірно в інтервалі (0, 1), то ймовірність попадання R в деякий інтервал, який належить інтервалу (0, 1), дорівнює його довжині. Зокрема,

$$P(c<\xi<d)=F(d)-F(c). \tag{2.74}$$

Отже, ймовірність попадання ξ в інтервал (c, d) дорівнює приросту функції розподілу $F(x)$ на даному інтервалі, а це означає, що $\xi = X$. Іншими словами, числа x_i, які визначається за допомогою формули (2.70), є можливими значеннями величини X із заданою функцією розподілу $F(x)$, що й потрібно було довести.

Правило 2.2. Для того, щоби знайти можливі значення x_i неперервної випадкової величини X, знаючи її функцію розподілу $F(x)$, необхідно вибрати випадкові числа r_i, прирівняти функцію розподілу $F(x)$ до цих випадкових чисел r_i і розв'язати відносно x_i отримані рівняння:

$$F(x_i)=r_i.$$

Зверніть увагу! Якщо розв'язати це рівняння в явному виді не вдається, то використовують графічні або числові методи розв'язку.

Задача 2.24. Розіграйте 3 можливих значень неперервної випадкової величини X, яка розподілена рівномірно на інтервалі (2, 10).

Розв'язок

Запишемо функцію розподілу величини X. Оскільки розподіл рівномірний на інтервалі (a; b), то

$$F(x)=(x-a)/(b-a)$$

Згідно з умовою задачі, $a = 2$, $b = 10$, відповідно:

$$F(x)=(x-2)/(8).$$

За правилом, зазначеним вище, запишемо рівняння для знаходження можливих значень x_i, для чого прирівняємо функцію розподілу до випадкового числа:

$$(x_i-2)/8=r_i.$$

Звідси знаходимо:

$$x_i=8r_i+2.$$

Виберемо (довільно або навмання) три випадкових числа, наприклад: r_1 = 0,11; r_2 = 0,17; r_3 = 0,66.

Підставимо ці числа у рівняння відносно x_i. В результаті отримаємо відповідні можливі значення X:

$$x_1=8\cdot 0{,}11+2=2{,}88; \quad x_2=8\cdot 0{,}17+2=3{,}36; \quad x_3=8\cdot 0{,}66+2=7{,}28.$$

Задача 2.25. Неперервна випадкова величина X розподілена згідно показникового закону, який заданий відповідною функцією розподілу (параметр λ відомий, причому $\lambda > 0$).

$$F(x)=1-e^{-\lambda x}, \qquad x>0.$$

Знайдіть, явну формулу для розіграшу можливих значень X.

Розв'язок

У відповідності до правила 2.2. запишемо:

$$1-e^{-\lambda x_i}=r_i.$$

Розв'яжемо дане рівняння відносно x_i.

$$e^{-\lambda x_i}=1-r_i \quad \text{або} \quad -\lambda x_i=\ln(1-r_i), \quad \rightarrow \quad x_i=-\frac{1}{\lambda}\ln(1-r_i).$$

Випадкове число r_i знаходиться в інтервалі (0, 1); відповідно число 1 - r_i також знаходитиметься в інтервалі (0, 1). Тому, для розіграшу можливих значень X можна скористатись простішою формою знайденої формули:

$$x_i=-\frac{1}{\lambda}\ln(r_i).$$

Зверніть увагу! Відомо, що

$$F(x)=\int_{-\infty}^{x} f(x)dx, \quad \text{зокрема} \quad F(x_i)=\int_{-\infty}^{x_i} f(x)dx.$$

Звідси випливає, що якщо відома густина ймовірності $f(x)$, то для розіграшу X можна замість рівнянь $F(x_i) = r_i$ розв'язати відносно x_i наступне рівняння

$$\int_{-\infty}^{x_i} f(x)dx=r_i.$$

Правило 2.3. Для того, щоби знайти можливі значення x_i неперервної випадкової величини X, знаючи її густина ймовірності $f(x)$, необхідно вибрати випадкове число r_i і розв'язати відносно x_i рівняння:

$$\int_{-\infty}^{x_i} f(x)dx=r_i, \quad \text{або} \quad \int_{a}^{x_i} f(x)dx=r_i,$$

де a – найменше кінцеве можливе значення X.

Задача 2.26. Задано густину ймовірності неперервної випадкової величини X, $f(x) = \lambda(1 — \lambda x/2)$ на інтервалі $(0, 2/\lambda)$; за межами даного інтервалу $f(x) = 0$. Знайдіть явну формулу для розігрування (моделювання) можливих значень X.

Розв'язок

Запишемо відповідно до правила, зазначеного вище, рівняння

$$\lambda\int_{0}^{x_i} (1-\lambda x/2)dx=r_i.$$

Виконавши інтегрування і розв'язавши отримане квадратне рівняння відносно x_i, отримаємо:

$$x_i = 2\left(1-\sqrt{1-r_i}\right)/\lambda .$$

Тема 11.8. Метод суперпозиції

Нехай функція розподілу випадкової величини X, що розігрується, може бути представлена у вигляді лінійної комбінації двох функцій розподілу:

$$F(x) = C_1 F_1(x) + C_2 F_2(x), \quad (C_1 > 0, C_2 > 0)$$

При $x \to \infty$ кожна із функцій розподілу прямує до 1, тому $C_1 + C_2 = 1$.

Уведемо допоміжну дискретну випадкову величину Z, із таким законом розподілу:

Z	1	2
p	C_1	C_2

Бачимо, що

$$P(Z=1) = C_1, \quad P(Z=2) = C_2. \tag{2.75}$$

Виберемо два незалежних випадкових числа r_1 і r_2. За допомогою числа r_1 будемо розігрувати можливі значення Z. Якщо виявиться, що $Z = 1$, то шукане можливе значення X знаходять із рівняння $F_1(x) = r_2$, а якщо $Z = 2$, то з рівняння $F_2(x) = r_2$.

Доведемо, що функція розподілу величини, що розігрується, дорівнює заданій функції розподілу. Для цього скористаємося формулою повної ймовірності (див. лекцію № 2):

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A).$$

Позначимо через A подію $X < x$, тоді

$$P(A) = P(X<x) = F(x). \tag{2.76}$$

Розглянемо гіпотези B_1: $Z = 1$ та B_2: $Z = 2$. Імовірність цих гіпотез, за (2.76), дорівнює:

$$P(B_1) = P(Z=1) = C_1, \quad P(B_2) = P(Z=2) = C_2. \tag{2.77}$$

Умовні ймовірності появи події А відповідно дорівнюють:

$$P_{B_1}(A) = P_{B_1}(X<x) = F_1(x), \quad P_{B_2}(A) = P_{B_2}(X<x) = F_2(x). \tag{2.78}$$

Підставивши формули (2.76-2.78) у формулу повної ймовірності, остаточно отримаємо:

$$F(x) = C_1 F_1(x) + C_2 F_2(x),$$

що й потрібно було довести.

Зверніть увагу! Метод суперпозиції узагальнюється на суму з n функцій розподілу.

Правило 2.4. Для того, щоби розіграти можливі значення випадкової величини X, функція розподілу якої

$$F(x)=C_1F_1(x)+C_2F_2(x),$$

де $C_1 > 0$, $C_2 > 0$, причому $C_1 + C_2 = 1$, потрібно вибрати два незалежних випадкових числа r_1 та r_2 і по випадковому числу r_1 розіграти можливі значення допоміжної дискретної випадкової величини Z:

Z	1	2
p	C_1	C_2

Якщо виявиться, що $Z = 1$, то розв'язують відносно x рівняння $F_1(x) = r_2$; якщо $Z = 2$, то відповідно $F_2(x) = r_2$.

Задача 2.27. Знайти явну формулу для розіграшу неперервної випадкової величини X, яка задана функцією розподілу

$$F(x)=1-0{,}25\left(\mathrm{e}^{-2x}+3\mathrm{e}^{-x}\right),\quad 0<x<\infty.$$

Розв'язок

Скористаємось методом суперпозиції. Для цього представимо задану функцію у вигляді

$$F(x)=0{,}25\left(1-\mathrm{e}^{-2x}\right)+0{,}75\left(1-\mathrm{e}^{-x}\right).$$

Отже, можна прийняти:

$$F_1(x)=0{,}25\left(1-\mathrm{e}^{-2x}\right),\quad F_2(x)=0{,}75\left(1-\mathrm{e}^{-x}\right),$$

$$C_1=0{,}25,\quad C_2=0{,}75.$$

Уведемо допоміжну дискретну випадкову величину Z, закон розподілу якої

Z	1	2
p	0,25	0,75

Виберемо незалежні один від одного випадкові числа r_1 та r_2. Розіграємо Z по випадковому числу r_1, для чого, згідно з правилом 2.1, побудуємо частинні інтервали: Δ_1 – (0; 0,25), та Δ_2 – (0,25; 1). Якщо $r_1 < 0{,}25$ то $Z = 1$, якщо $r_1 \geq 0,\ 25$, то $Z = 2$.

Отже, можливі значення X знаходять шляхом розв'язування відносно x таких рівнянь:

$$1-\mathrm{e}^{-2x}=r_2,\quad \text{якщо}\quad r_1<0{,}25;$$

$$1-\mathrm{e}^{-x}=r_2,\quad \text{якщо}\quad r_1\geq 0{,}25.$$

Використовуючи розв'язок задачі 2.25, в якому була знайдена явна формула: $x = -(1/\lambda)\ln(r)$ для розігрування можливих значень показникового розподілу з заданим параметром λ, то в кінцевому випадку одержимо:

$$x=-(1/2)\ln r_2,\quad \text{якщо}\quad r_1<0{,}25;$$

$$x=-\ln r_2,\quad \text{якщо}\quad r_1\geq 0{,}25.$$

Тема 11.9. Наближене розігрування нормальної випадкової величини

Нагадаємо, що якщо випадкова величина R розподілена рівномірно на інтервалі (0, 1), то її математичне сподівання і дисперсія відповідно дорівнюють

$$M(R)=1/2, \tag{2.79}$$

$$D(R)=1/12. \tag{2.80}$$

Складемо суму з n незалежних, рівномірно розподілених на інтервалі (0, 1) випадкових величин R_j (j = 1, 2, ..., n):

$$\sum_{j=1}^{n} R_j. \tag{2.81}$$

Для того, щоби пронормувати цю суму, знайдемо спочатку її математичне сподівання та дисперсію.

Відомо, що математичне сподівання суми випадкових величин дорівнює сумі математичних сподівань випадкових величин, що підсумовується. Тобто

$$M\left[\sum_{j=1}^{n} R_j\right]=n/2.$$

Аналогічно дисперсія суми випадкових випадкових величин дорівнює сумі дисперсій випадкових величин, що підсумовується. Тобто

$$D\left[\sum_{j=1}^{n} R_j\right]=n/12.$$

Звідси середньоквадратичне відхилення суми дорівнюватиме:

$$\sigma_{\sum}=\sqrt{n/12}.$$

Пронормуємо суму, що розглядається, для чого віднімемо від неї її математичне сподівання та поділимо на середньоквадратичне відхилення:

$$\frac{\sum_{j=1}^{n} R_j-(n/2)}{\sqrt{n/12}}.$$

Згідно із центральною граничною теоремою, при $n \to \infty$ розподіл цієї нормованої випадкової величини прямує до нормального з параметрами a = 0, σ = 1. *При кінцевому n розподіл наближено нормальний*. Зокрема, при n = 12 отримаємо досить добре і зручне для розрахунків наближення:

$$\sum_{j=1}^{12} R_j-6.$$

Правило 2.5. Для того, щоби розіграти можливі значення хі нормальної випадкової величини X, з параметрами a = 0, σ = 1, необхідно просумувати 12 незалежних один від одного випадковий чисел і відняти від отриманої суми 6:

$$x_i=\sum_{j=1}^{12} r_j-6=S_i-6$$

Задача 2.28. *А*) Розіграти 100 можливих значень нормальної величини X з $a = 0$, $\sigma = 1$; *Б*) оцінити параметри розіграної величини.

Розв'язок

А) виберемо 12 випадкових чисел зі першого рядку таблиці випадкових чисел, просумуємо їх і від отриманої суми віднімемо число 6, в результаті одержимо:

$$x_1 = (0{,}10 + 0{,}09 + ... + 0{,}67) - 6 = -0{,}99.$$

Аналогічним способом зайдемо решту можливих значень випадкової величини X, вибираючи з кожного наступного рядка перші 12 випадкових чисел,

Б) зробимо необхідні розрахунки та отримаємо шукані оцінки:

$$a^* = \overline{x_B} \approx -0{,}05; \qquad \sigma^* = \sqrt{D_B} \approx 1{,}04.$$

Одержані оцінки задовільні, оскільки a^* близький до нуля, а σ^* мало відрізняється від одиниці.

Зверніть увагу! Якщо потрібно розіграти можливі значення z_i нормальної випадкової величини Z із математичним сподіванням a та середньоквадратичним відхиленням σ, то спочатку необхідно розіграти, згідно з правилом 5, можливі значення x_i, а вже потім знаходити можливі значення z_i згідно з формулою

$$z_i = \sigma x_i + a. \tag{2.82}$$

Ця формула отримана зі наступної рівності:

$$(z_i - a)/\sigma = x_i.$$

Лекція 12. Відомості про ланцюги Маркова

Тема 12.1. Поняття ланцюга Маркова

Ланцюгом Маркова називається така послідовність випробувань, в якій в кожному з випробувань відбулася тільки одна з k несумісних подій $A_1, A_2, ..., A_k$ повної групи, причому умовна ймовірність $p_{ij}(s)$ того, що у випробуванні s настане подія $A_j (j = 1, 2, ..., k)$, при умові, що в s-1 настала подія $A_i (i = 1, 2, ..., k)$, не залежить від результатів попередніх подій.

Наприклад, якщо послідовність випробувань утворює ланцюг Маркова і повна група складається з чотирьох несумісних подій: A_1, A_2, A_3, A_4, при цьому відомо, що в шостому випробувані з'явилась подія A_2, то умовна ймовірність того, що в сьомому (наступному) випробувані настане подія A_4, не залежить від того, які події відбулись в першому, другому, ..., п'ятому випробувані.

Зазначимо, що незалежні випробування є окремим випадком ланцюга Маркова. Дійсно, якщо випробування не залежать від появи деякої визначеної події в будь-якому випробувані, то вони не залежать і від результатів попередніх випробувань. Звідси і випливає, що ланцюг Маркова є доповненням поняття незалежних випробувань.

Надалі будемо використовувати термінологію, яка прийнята при опису ланцюгів Маркова. Нехай дано систему, яка в кожен момент часу знаходиться в одному із k станів: першому, другому, ..., k-му. У визначені моменти часу в результаті випробування стан системи змінюється, наприклад із i в j. Зокрема, система може залишитись в тому ж стані (тобто «перейти» зі стану i в стан $j = i$).

Отже, події називають *станами системи*, а випробування – *зміни її станів*.

Дамо означення ланцюга Маркова.

Ланцюгом Маркова називається послідовність випробувань, в кожному з яких система має тільки один із k станів повної групи, при чому умовна ймовірність $p_{ij}(s)$ того, що s-му випробуванні система буде знаходитися в стані j, за умови, що після s - 1 випробування вона знаходилась в стані i, не залежить від результатів раніше проведених випробувань.

Ланцюгом Маркова з дискретним часом називається ланцюг, зміна стану якого відбувається у визначенні фіксовані моменти часу.

Ланцюгом Маркова з неперервним часом називається ланцюг, зміна стану якого відбувається в довільні, тобто випадкові, можливі моменти часу.

Тема 12.2. Однорідний ланцюг Маркова. Перехідні ймовірності. Матриця переходу

Однорідним називається ланцюг Маркова, якщо умовна ймовірність $p_{ij}(s)$ (перехід зі стану i в стан j) не залежить від номера випробування. Тому замість $p_{ij}(s)$ використовують позначення p_{ij}.

Задача 2.29. Випадкове блукання.

Нехай на прямій Ox в точці з цілочисловими координатами $x = n$ знаходиться матеріальна точка. У визначені моменти часу t_1, t_2, t_3, ... з точкою відбуваються «поштовхи». Під дією поштовху точка з ймовірністю p зміщується на одиницю вправо та з ймовірністю $1 - p$ на одиницю вліво. Зрозуміло, що положення (координата) точки після поштовху залежить від того, де вона знаходилась після безпосереднього попереднього поштовху, та не залежить від того, як вона рухалась під дією решти попередніх поштовхів.

Отже, *випадкове блукання* – це задача однорідного ланцюга Маркова з дискретним часом.

Обмежимось елементами теорії кінцевих однорідних ланцюгів Маркова.

Перехідною ймовірністю p_{ij} називають умовну ймовірність того, що із стану i (в якому система виявилась в результаті деяких випробувань, неважливо якого номера) в результаті наступного випробування система перейде в стан j.

Таким чином, в позначенні p_{ij} перший індекс вказує на номер попереднього стану, а другий – на номер наступного стану. Наприклад, p_{11} – ймовірність переходу з першого стану в перший; p_{23} — ймовірність переходу із другого стану в третій.

Нехай кількість станів кінцева і дорівнює k.

Матрицею переходу системи називають матрицю, яка складається з перехідних імовірностей цієї системи.

$$\wp = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{pmatrix}$$

Оскільки, в кожному рядку матриці розміщені ймовірності подій (переходу зі стану i в стан j), які утворюють повну групу, то сума ймовірностей цих подій дорівнює одиниці. Іншими словами, сума перехідних ймовірностей кожного рядка матриці переходу дорвнює одиниці:

$$\sum_{i=1}^{k} p_{ij} = 1, \qquad (1 = 1, 2, \ldots, k).$$

Наведемо задачу матриці переходу системи, яка може знаходитись в трьох станах:

$$\wp = \begin{pmatrix} 0{,}5 & 0{,}2 & 0{,}3 \\ 0{,}4 & 0{,}5 & 0{,}1 \\ 0{,}6 & 0{,}3 & 0{,}1 \end{pmatrix}$$

Тут $p_{11} = 0{,}5$ – ймовірність переходу зі стану $i = 1$ в той самий стан $j = 1$; $p_{21} = 0{,}4$ – ймовірність переходу зі стану $i = 2$ в стан $j = 1$. Аналогічний зміст мають і решта елементів матриці.

Тема 12.3. Рівність Маркова

Позначимо через $P_{ij}(n)$ ймовірність того, що в результаті n кроків (випробувань) система перейде зі стану i в стан j. Наприклад, $P_{25}(10)$ – ймовірність переходу через 10 кроків із другого стану в п'ятий.

Підкреслимо, що при $n = 1$ одержимо перехідні ймовірності:

$$P_{ij}(1) = p_{ij}.$$

Поставимо перед собою задачу: знаючи перехідні ймовірності pij, знайти ймовірності $P_{ij}(n)$ переходу системи зі стану i в стан j після n кроків. З цією метою введемо проміжний (між i та j) стан r. Тобто будемо вважати, що з вихідного стану i після m кроків система перейде в проміжний стан r з ймовірністю $P_{ir}(m)$, після чого за $n - m$ кроків, що залишились, – із проміжного стану r у кінцевий стан j із ймовірністю $P_{rj}(n - m)$.

Згідно формули повної ймовірності, отримаємо:

$$P_{ij}(n) = \sum_{r=1}^{k} P_{ir}(m) \cdot P_{rj}(n - m). \qquad (2.83)$$

Цю формулу називають *рівністю Маркова*.

Уведемо позначення: A – шукана подія (після n кроків система перейде з вихідного стану i в кінцевий стан j), відповідно, $P(A) = P_{ij}(n)$; $B_r(r = 1, 2, ..., k)$ – гіпотези (після m кроків система перейде з вихідного стану i в проміжний стан r), відповідно, $P(B_r) = P_{ir}(m)$; $PB_r(A)$ – умовна ймовірність настання події A, за

умови, що має місце гіпотеза B_r (після n - m кроків система перейде із проміжного стану r у кінцевий стан j), відповідно, $PB_r(A) = P_{rj}(n - m)$.

Знову, використовуючи формулу повної ймовірності, одержимо:

$$P(A)=\sum_{r=1}^{k} P(B_r)\cdot P_{B_r}(A),$$

або, відповідно до прийнятих нами позначень,

$$P_{ij}(n)=\sum_{r=1}^{k} P_{ir}(m)\cdot P_{rj}(n-m),$$

що збігається з формулою (2.83) Маркова.

Покажемо, що, знаючи всі перехідні ймовірності $p_{ij} = P_{ij}(1)$, тобто знаючи матрицю переходу зі стану в стан за один крок, можна знайти ймовірності $P_{ij}(2)$ переходу зі стану в стан за два кроки, а отже і саму матрицю переходу; по відомій матриці можна знайти матрицю переходу зі стану в стан за три кроки, і т. д.

Для цього підставимо $n = 2$, $m = 1$ в рівність Маркова:

$$P_{ij}(n)=\sum_{r=1}^{k} P_{ir}(m)\cdot P_{rj}(n-m).$$

У результаті одержимо:

$$P_{ij}(2)=\sum_{r=1}^{k} P_{ir}(1)\cdot P_{rj}(2-1),$$

або

$$P_{ij}(2)=\sum_{r=1}^{k} p_{ir}\cdot p_{rj}. \qquad (2.84)$$

Отже, за формулою (2.84) можна знайти всі ймовірності $P_{ij}(2)$, і саму матрицю.

Оскільки безпосереднє використання формули (2.84) є досить втомлюючим процесом, а матричне обчислення приводить, радше, до ланцюга Маркова, запишемо співвідношення, що випливає із формули (2.84), в матричній формі:

$$\wp_2=\wp_1\cdot\wp_1=\wp_1^2.$$

Підставивши $n = 3$, $m = 2$ в (2.84) аналогічно отримаємо:

$$\wp_3=\wp_1\cdot\wp_2=\wp_1\cdot\wp_1^2=\wp_1^3.$$

У загальному випадку:

$$\wp_n=\wp_1^n.$$

Задача 2.30. Задана матриця переходу:

$$\wp_1=\begin{pmatrix}0{,}4 & 0{,}6\\ 0{,}3 & 0{,}7\end{pmatrix}$$

Знайти матрицю переходу вигляду:

$$\wp_2=\begin{pmatrix}P_{11}(2) & P_{12}(2)\\ P_{21}(2) & P_{22}(2)\end{pmatrix}$$

Розв'язок

Скористаємося формулою

$$\wp_n=\wp_1^n.$$

Помноживши матриці, в кінцевому випадку одержимо:

$$\wp_2=\begin{pmatrix}0{,}4 & 0{,}6\\ 0{,}3 & 0{,}7\end{pmatrix}\times\begin{pmatrix}0{,}4 & 0{,}6\\ 0{,}3 & 0{,}7\end{pmatrix}.$$

$$\wp_2=\begin{pmatrix}0{,}34 & 0{,}66\\ 0{,}33 & 0{,}67\end{pmatrix}.$$

Контрольні запитання до розділу 2

1. Розкрийте зміст понять «генеральна» та «вибіркова сукупність». Наведіть способи відбору статистичного матеріалу.
2. Що таке емпірична функція розподілу? Які її властивості?
3. Що таке згруповані розподіли вибірки?
4. Що ви розумієте під поняттями полігону та гістограми частот?
5. Які числові характеристики вибірки ви знаєте?
6. Розкрийте зміст методу добутків та методу сум обчислення вибіркового середнього та вибіркової дисперсії.
7. Що таке «точкова оцінка»?
8. Які існують методи визначення точкових статистичних оцінок? Опишіть їх.
9. Що таке «інтервальна оцінка» та «надійність»?
10. Опишіть алгоритм побудови довірчого інтервалу для оцінки математичного сподівання нормального розподілу при відомому середньоквадратичному відхиленні.
11. Опишіть алгоритм побудови довірчого інтервалу для оцінки математичного сподівання нормального розподілу при невідомому середньоквадратичному відхиленні.
12. Опишіть алгоритм побудови довірчого інтервалу, для оцінки середнього квадратичного відхилення нормального розподілу.
13. Розкрийте зміст понять: функціональна, статистична та кореляційна залежності.
14. Що таке умовні середні та вибіркове рівняння регресії?

15. Що таке кореляційна таблиця?
16. Наведіть визначення вибіркового коефіцієнта кореляції.
17. Що таке вибіркове кореляційне відношення, та які його властивості?
18. Наведіть та опишіть задачу криволінійної кореляції.
19. Наведіть та опишіть задачу множинної кореляції.
20. Розкрийте зміст методу Монте-Карло.
21. Опишіть оцінку похибок за допомогою методу Монте-Карло.
22. Опишіть розігрування дискретної випадкової величини методом Монте-Карло.
23. Опишіть розігрування неперервної випадкової величини методом Монте-Карло.
24. Опишіть наближене розігрування нормальної випадкової величини методом Монте-Карло.
25. Що таке «ланцюг Маркова», «перехідна ймовірність» та «матриця переходу»?
26. Запишіть рівність Маркова. В чому полягає її зміст?

РОЗДІЛ 3. ФУНКЦІЯ КОМПЛЕКСНОЇ ЗМІННОЇ

Теорія функцій комплексної змінної (або скорочено ФКЗ) – розділ математичного аналізу, в якому розглядаються і вивчаються функції комплексного аргументу.

Важливу роль у ФКЗ відіграє операційне числення, яке знаходить широке застосування в сучасній автоматиці, телемеханіці, теорії автоматичного регулювання. Методи операційного числення дозволяють спростити алгоритм розв'язку диференціальних рівнянь та систем диференціальних рівнянь шляхом переходу і подальшого розв'язку більш простих алгебраїчних рівнянь.

Лекція 13. Функції комплексної змінної та дії над ними

Тема 13.1. Комплексні числа та дії над ними

Комплексним числом z називається впорядкована пара дійсних чисел (x, y): $z = (x, y)$. Дійсні числа x і y називаються *дійсною і уявною частинами* комплексного числа z відповідно і позначаються $x = \mathrm{Re}\, z$ (від франц. Reelle «дійсний»), $y = \mathrm{Im}\, z$ (від франц. Imaginaire – «уявний»).

Два комплексних числа $z_1 = \{x_1, y_1)$ і $z_2 = \{x_2, y_2)$ називаються *рівними* тоді і тільки тоді, коли $x_1 = x_2$ та $y_1 = y_2$, тобто

$$z_1 = z_2 \Leftrightarrow \begin{cases} x_1 = x_2; \\ y_1 = y_2. \end{cases} \tag{3.1}$$

Сумою комплексних чисел $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$ називається комплексне число z, яке визначається рівністю

$$z = z_1 + z_2 = (x_1 + x_2, y_1 + y_2). \tag{3.2}$$

Різницею комплексних чисел $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$ називається комплексне число z, що визначається рівністю

$$z = z_1 - z_2 = (x_1 - x_2, y_1 - y_2). \tag{3.3}$$

Добутком комплексних чисел $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$ називається комплексне число z, яке визначається рівністю

$$z = z_1 \cdot z_2 = (x_1 \cdot x_2 - y_1 \cdot y_2, x_1 \cdot y_2 + x_2 \cdot y_1). \tag{3.4}$$

Справедливі такі властивості операцій над комплексними числами:

1. $z_1 + z_2 = z_2 + z_1$;
2. $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$;
3. $z_1 \cdot z_2 = z_2 \cdot z_1$;
4. $z_1 \cdot (z_2 \cdot z_3) = (z_1 \cdot z_2) \cdot z_3$;
5. $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$.

Геометрично комплексне число $z = (x, v)$ можна зобразити точкою $M(x, y)$ на декартовій площині R^2 або вектором $\vec{z} = (x, y)$, що йде з початку координат в точку $M(x, y)$, тобто як радіус-вектором точки M.

Площина C, на якій комплексні числа зображуються як точки, називається *комплексною площиною*. Вісь Ox називається *дійсною віссю* комплексної площини, а вісь Oy – *уявною віссю*. Комплексна площину C, що

доповнюється нескінченно віддаленою точкою $z = \infty$, називають *розширеною комплексною площиною* і позначають $\overline{C}$.

Масштабна одиниця осі Ox, тобто комплексне число z = (l, 0), є матеріальною одиницею. Масштабна одиниця осі Oy, тобто число z = (0, 1), називається уявною одиницею і позначається (0, 1) = i.

За правилом множення комплексних чисел, одержимо:

$$(0,1)\cdot(0,1)=(0\cdot 0-1\cdot 1, 0\cdot 1+1\cdot 0)=(-1,0)=-1,$$

тобто $i^2 = -1$. Аналогічно $i^3 = -i$, $i^4 = 1$, і т. д.

Два комплексних числа $z=(x,y)$ і $\bar{z}=(x,-y)$, які відрізняються знаком уявної частини, називають *комплексно спряженими* числами (рис. 3.1).

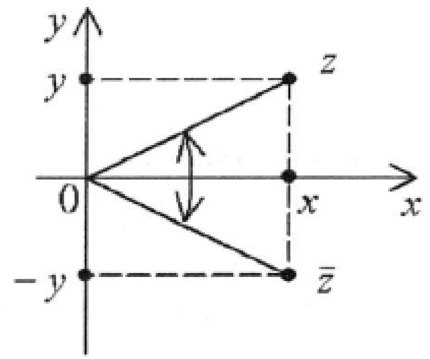


Рис. 3.1

Дійсне невід'ємне число

$$|z|=\sqrt{x^2+y^2}=\sqrt{(\Re z)^2+(\Im z)^2} \tag{3.5}$$

називають *модулем комплексного числа* $z = (x, y)$.

Геометрично модуль комплексного числа – це відстань від точки, яка зображує число z, до початку координат (або довжина радіус-вектора точки).

Кут φ між додатним напрямом осі Ox і вектором $\vec{z}=(x,y)$ називають *аргументом комплексного числа* $z = (x, y)$ (позначення $Arg\ z$). Цей кут називають *загальним значенням аргументу*, він визначений неоднозначно, з точністю до $2k\pi$, $k \in \mathbb{Z}$. Аргумент комплексного числа z = 0 не визначено.

Головним значенням аргументу комплексного числа називають значення кута φ, укладеного в проміжку довжиною 2π (позначення $arg\ z$). Для визначеності можна вважати, що $-\pi < arg\ z < \pi$.

Загальне і головне значення аргументу пов'язані співвідношенням:

$$Arg\, z=arg\, z+2k\pi, k\in Z. \tag{3.6}$$

З означення модуля і аргументу комплексного числа $z = (x, v)$ випливає:

$$x=|z|\cos(\phi), \quad y=|z|\sin(\phi). \tag{3.7}$$

Оскільки, $-\pi < arg\ z < \pi$, то з рівності tg φ = y/x матимемо

$$arg\, z=\begin{cases} arctg\frac{y}{x}, x>0; \\ arctg\frac{y}{x}+\pi, x<0, y\geq 0; \\ arctg\frac{y}{x}-\pi, x<0, y<0; \\ \pi/2, x=0, y>0; \\ -\pi/2, x=0, y<0. \end{cases} \tag{3.8}$$

Справедливі такі властивості модуля і аргументу комплексного числа:

1. $z\cdot\overline{z}=(x,y)\cdot(x,-y)=(x^2+y^2,0)=|z|^2;$
2. $|z_1\cdot z_2|=|z_1|\cdot|z_2|,\qquad Arg(z_1\cdot z_2)=Arg(z_1)\cdot Arg(z_2);$
3. $\left|\frac{z_1}{z_2}\right|=\frac{|z_1|}{|z_2|},\qquad Arg\left(\frac{z_1}{z_2}\right)=Arg(z_1)-Arg(z_2).$

Існують кілька форм запису комплексних чисел.

Алгебраїчною формою запису комплексного числа $z = (x, y)$ називається зображення його у вигляді

$$z=x+iy \tag{3.9}$$

(оскільки $z = (x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1)\,(y, 0) = x + iy$).

Комплексні числа, записані в алгебраїчній формі можна додавати, віднімати і множити як звичайні двочлени, враховуючи, що $i^2 = -1$, тобто для комплексних чисел $z_1 = x_1 + iy_1$ і $z_2 = x_2 + iy_2$ справедливі рівності

$$z_1+z_2=(x_1+x_2)+i(y_1+y_2),$$

$$z_1-z_2=(x_1-x_2)+i(y_1-y_2),$$

$$z_1 z_2=(x_1 x_2-y_1 y_2)+i(x_1 y_2+x_2 y_1).$$

Операцію ділення комплексних чисел, записаних в алгебраїчній формі, можна визначити за допомогою операції множення. Тобто, щоб обчислити значення z_1/z_2, треба чисельник і знаменник дробу помножити на число, комплексно спряжене зі знаменником:

$$\frac{z_1}{z_2}=\frac{z_1\overline{z_2}}{z_2\overline{z_2}}=\frac{x_1x_2+y_1y_2}{x_2^2+y_2^2}+i\,\frac{y_1x_2-x_1y_2}{x_2^2+y_2^2}. \tag{3.10}$$

Тригонометричною формою запису комплексного числа $z = (x, y)$ називається зображення його у вигляді

$$z=|z|(\cos\phi+i\sin\phi), \tag{3.11}$$

(бо $z = (x, y) = x + iy = |\,z\,| \cdot \cos\varphi + |\,z\,| \cdot \sin\varphi = |\,z\,|\,(\cos\varphi + i \sin\varphi)$).

Числа, записані в тригонометричної формі, зручно множити і ділити, використовуючи властивості модуля і аргументу, а саме для комплексних чисел $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ і $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2,)$ справедливі рівності

$$z_1\cdot z_2=r_1\cdot r_2(\cos(\phi_1+\phi_2)+i\sin(\phi_1+\phi_2)),$$

$$\frac{z_1}{z_2}=\frac{r_1}{r_2}(\cos(\phi_1-\phi_2)+i\sin(\phi_1-\phi_2)).$$

Показниковою формою запису комплексного числа $z = (x, y)$ називається зображення його у вигляді

$$z=|z|e^{i\phi}, \tag{3.12}$$

(позаяк за формулою Ейлера $\cos\varphi + i\sin\varphi = e^{i\varphi}$).

Числа, записані в показниковій формі, зручно множити і ділити, використовуючи властивості показникової функції, а саме для комплексних чисел $z_1=r_1e^{i\phi_1}$ і $z_2=r_2e^{i\phi_2}$, справедливі рівності

$$z_1\cdot z_2=r_1\cdot r_2e^{i\phi_1}\cdot e^{i\phi_2}=r_1\cdot r_2e^{i(\phi_1+\phi_2)},$$

$$\frac{z_1}{z_2}=\frac{r_1 e^{i\phi_1}}{r_2 e^{i\phi_2}}=\frac{r_1}{r_2}e^{i(\phi_1-\phi_2)}.$$

У наслідок зазначених властивостей модуля і аргументу комплексного числа, операції множення і ділення комплексних чисел зручніше виконувати, якщо ці числа записані в тригонометричній або показниковій формі.

Властивості модуля і аргументу комплексного числа дозволяють отримати формулу зведення комплексного числа до цілого додатного степеня:

$$z^n=\left(|z|(\cos\phi+i\sin\phi)\right)^n=|z|^n(\cos n\phi+i\sin n\phi), \qquad (3.13)$$

або в показниковій формі:

$$z^n=\left(|z|\cdot e^{i\phi}\right)^n=|z|^n\cdot e^{in\phi}. \qquad (3.14)$$

Цю формулу називають *формулою Муавра*.

Легко перевірити, що ця формула залишається справедливою для n = 0 і для цілих від'ємних степенів.

Коренем n-го степеня з комплексного числа z називається таке комплексне число $\alpha=\sqrt[n]{z}$, , для якого $\alpha^n=z$.

Згідно означенням та формулою Муавра модулем шуканого кореня $\sqrt[n]{|z|}$ буде аргумент $\frac{\phi+2\pi k}{n}$, де $k\in\mathbb{Z}$ Таким чином,

$$\sqrt[n]{z}=\sqrt[n]{|z|(\cos\phi+i\sin\phi)}=\sqrt[n]{|z|}\left(\cos\frac{\phi+2\pi k}{n}+i\sin\frac{\phi+2\pi k}{n}\right), \qquad (3.15)$$

для k = 0, 1, ... (n-1), або

$$\sqrt[n]{z}=\sqrt[n]{\left(|z|\cdot e^{i\phi}\right)}=\sqrt[n]{|z|}\cdot e^{i\frac{\phi+2\pi k}{n}}. \qquad (3.16)$$

для k = 0, 1, ... (n-1).

Корінь n-го степеня з комплексного числа z має n різних значень, модулі яких однакові й дорівнюють $\sqrt[n]{|z|}$, а аргументи двох послідовних значень відрізняються на кут $2\pi/n$ (рис. 3.2). Отже, всі значення кореня лежать на окружності з центром у початку координат радіусом $\sqrt[n]{|z|}$.

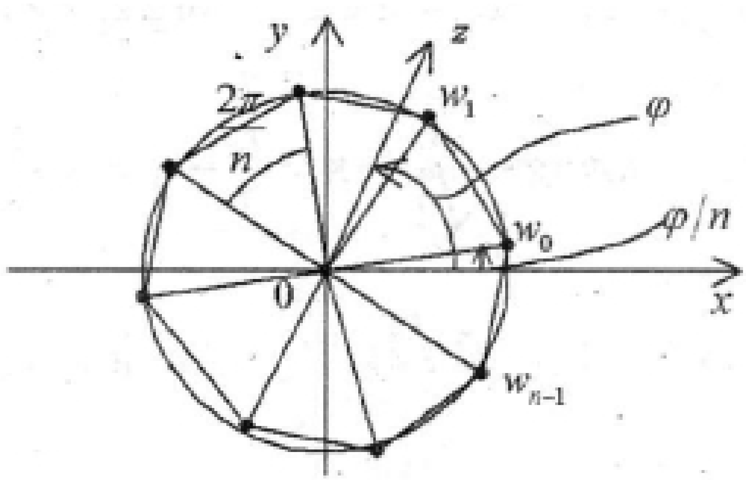


Рис. 3.2

Тема 13.2. Функції комплексної змінної

Величина $z = x + iy$ називається *комплексною змінною*, якщо x і y – змінні величини.

При зміні x і y комплексна змінна $z = x + y$ буде пробігати деяку множину точок комплексної площині C, яку називають також *площиною комплексної змінної* z (рис. 3.3).

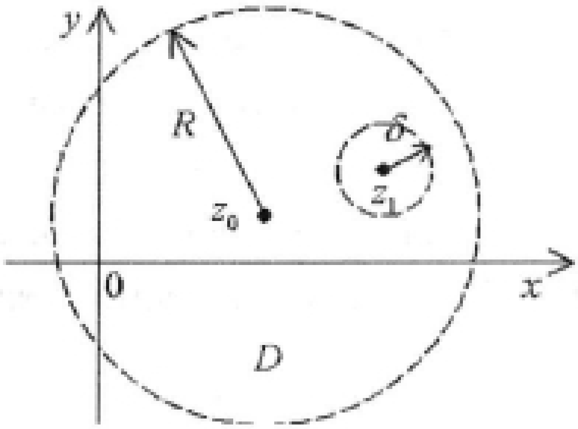


Рис. 3.3

Оскільки відстань між двома точками $z_1 = x_1 + iy_1$ та $z_2 = x_2 + iy_2$ знаходиться за формулою

$$\rho(z_1, z_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = |z_1 - z_2|, \tag{3.17}$$

то рівняння кола радіуса R із центром в точці z_0 має вигляд $|z - z_0| = R$, а нерівність $|z - z_0| < R$ задає безліч точок z, що лежать всередині кола радіусом R з центром в z_0.

δ-околом точки z_1 називається внутрішня частина кола з центром z_0 і довільним радіусом δ, тобто множина $|z - z_1| < \delta$.

Границею нескінченно віддаленої точки $z = \infty$ називається множина, яка складається з точок $z \in C$, для яких $|z| > R$, і самої точки $z = \infty$.

Точка множини D називається *внутрішньою*, якщо існує окіл цієї точки, який міститься в D.

Множина D називається *відкритою*, якщо кожна її точка є внутрішньою.

Множина D називається *зв'язаною*, якщо будь-які дві точки D можна з'єднати між собою ламаною, яка міститься в D.

Точка z називається *граничною точкою* множини D, якщо в будь-якому околі цієї точки знайдуться точки, які як належать D, так і не належать D.

Множина граничних точок називається *границею множини* D.

Множину, що містить усі свої граничні точки, називають *замкнутою*.

Область – це відкрита зв'язана множина.

Область D називається *замкнутою областю*, якщо вона містить усі свої граничні точки (позначення: $\overline{D}$).

Область називається *однозв'язною*, якщо її межа складається з однієї неперервної кривої без самоперетинів (можливо, замкнутої). Всі інші області називаються *багатозв'язними*.

Якщо кожному числу n з натурального ряду 1, 2, ... по певному правилу поставлено у відповідність комплексне число z_n, то кажуть, що задана *послідовність* $\{z_n\}$.

Оскільки $z_n = x_n + iy_n$, то задання послідовності комплексних чисел $\{z_n\}$ рівносильне завданням двох послідовностей дійсних чисел $\{x_n\}$ і $\{y_n\}$ відповідно.

Комплексне число $A = a + ib$ називається *границею* послідовності $\{z_n\}$ при $n \to \infty$, якщо для будь-якого $\varepsilon > 0$ існує додатне число $N(\varepsilon)$, що при всіх $n > N(\varepsilon)$ виконується нерівність $|z_n - A| < \varepsilon$ (позначення: $\lim\limits_{n\to\infty} z_n = A$).

Послідовність, яка має границю, називається *збіжною*.

Теорема: співвідношення $\lim\limits_{n\to\infty} x_n + i\,y_n = a + ib$ еквівалентне двом співвідношенням:

$$\lim_{n\to\infty} x_n = a, \quad \text{та} \quad \lim_{n\to\infty} y_n = b.$$

З цього твердження, зокрема, випливають, такі співвідношення:

$$\lim_{n\to\infty}(z_n + w_n) = \lim_{n\to\infty} z_n + \lim_{n\to\infty} w_n;$$

$$\lim_{n\to\infty} z_n \cdot w_n = \lim_{n\to\infty} z_n \cdot \lim_{n\to\infty} w_n;$$

$$\lim_{n\to\infty} \frac{z_n}{w_n} = \frac{\lim\limits_{n\to\infty} z_n}{\lim\limits_{n\to\infty} w_n}, \quad \text{якщо} \quad \lim_{n\to\infty} w_n \neq 0.$$

Нескінченно віддалена точка $A = \infty$ називається границею послідовності $\{z_n\}$, якщо для будь-якого як завгодно великого числа $R > 0$ знайдеться такий номер $N(R)$, що при всіх $n > N(R)$ виконується нерівність $|z_n| > R$ (позначення: $\lim\limits_{n\to\infty} z_n = \infty$).

Поняття функції комплексної змінної є окремим випадком загального математичного поняття функції.

Нехай D і E – деякі множини точок на комплексній площині, а комплексне число z може набувати будь-яке значення з D (z – комплексна змінна, D – областю її зміни).

Величина w називається *функцією незалежної змінної* z, якщо кожному значенню z за деяким правилом відповідає одне або кілька комплексних значень $w \in E$ (позначення: $w = f(z)$).

Якщо кожному $z \in D$ відповідає єдине значення $w \in E$, то кажуть, що на множині D задана *однозначна функція* $w = f(z)$.

Якщо кожному $z \in D$ відповідає декілька значень $w \in E$, то кажуть, що на множині D задана *багатозначна функція* $w = f(z)$.

Множина D називається областю визначення функції $w = f(z)$, а множина $E_1 = f(D) \subseteq E$ – областю значень функції $w = f(z)$.

Запишемо комплексні числа z і w в алгебраїчній формі:

$$z = x + iy, \qquad w = u + iv, \tag{3.18}$$

зауваживши, що $u = u(x, y)$, $v = v(x, y)$.

Отже, задання функції комплексної змінної:

$$w = f(z) = u(x, y) + iv(x, y), \tag{3.19}$$

еквівалентне заданню двох дійсних функцій від двох дійсних змінних $u = u(x, y)$, $v = v(x, y)$.

Функція $u = u(x, y)$ називають *дійсною частиною* функції $w = f(z)$, а функція $v = v(x, y)$ – *уявною частиною* функції $w = f(z)$ ($\mathrm{Re}\, f(z) = u(x, y)$, $\mathrm{Im}\, f(z) = v(x, y)$).

Число A називається границею функції $f(z)$ при $z \to z_0$, $z_0 \in C$, якщо для будь-якого $\varepsilon > 0$ знайдеться таке $\delta > 0$, що $|f(z) - A| < \varepsilon$, як тільки $0 < |z - z_0| < \delta$, $z \in D$ (позначення: $\lim\limits_{z \to z_0} f(z) = A$).

Нескладно показати, що співвідношення $\lim\limits_{z \to z_0} f(z) = A$, де $f(z) = u(x, y) + iv(x, y)$, а $A = B + iC$, $z_0 = x_0 + iy_0$, еквівалентне двом дійсним співвідношенням:

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} u(x, y) = B, \qquad \lim_{\substack{x \to x_0 \\ y \to y_0}} v(x, y) = C. \tag{3.20}$$

Функція $f(z)$ називається *неперервною в точці* z_0, якщо вона визначена в деякому околі цієї точки та

$$\lim_{z \to z_0} f(z) = f(z_0). \tag{3.21}$$

Якщо $f(z)$ визначена на множині D та неперервна в кожній точці цієї множини, то кажуть, що вона *неперервна на множині* D.

Знову можна легко показати, що умова неперервності функції $f(z)$ в точці z_0 еквівалентна двом співвідношенням:

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} u(x, y) = u(x_0, y_0), \qquad \lim_{\substack{x \to x_0 \\ y \to y_0}} v(x, y) = v(x_0, y_0).$$

Отже, функція комплексної змінної неперервна в точці z_0 тоді і тільки тоді, коли її дійсна і уявна частини, які розглядаються як функції дійсних змінних x та y, неперервні у відповідній точці (x_0, y_0).

Тема 13.3. Диференціювання функції комплексної змінної

Нехай задана однозначна функція $w = f(z)$ на області D (відкрита зв'язана множина) комплексної площини C.

Похідною функції $w = f(z)$ в точці $z \in D$ називається границя відношення приросту функції до приросту аргументу:

$$\lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} = f'(z) = \frac{dw}{dz}. \tag{3.22}$$

Якщо ця границя існує, то функція $f(z)$ називається *диференційованою* в точці z.

Функція $w = f(z)$ називається *аналітичною* в точці z_0, якщо існує такий окіл цієї точки, в кожній точці якого у функції $f(z)$ існує похідна.

Функція $w = f(z)$ називається *аналітичною* (*регулярною*, *голоморфною*) *функцією* в області D, якщо вона аналітична в кожній точці області D.

Оскільки означення похідної функції комплексної змінної повністю аналогічне означенню похідної функції дійсної змінної, то, в разі диференціювання $f(z)$, усі відомі правила диференціювання функції дійсної залишаються в силі, в тому числі *правила диференціювання складених та обернених функцій*, тобто:

1. $(f_1(z)+f_2(z))'=f_1'(z)+f_2'(z);$

2. $(f_1(z)-f_2(z))'=f_1'(z)-f_2'(z);$

3. $(f_1(z)\cdot f_2(z))'=f_1'(z)\cdot f_2(z)+f_1(z)\cdot f_2'(z);$

4. $\left(\dfrac{f_1(z)}{f_2(z)}\right)'=\dfrac{f_1'(z)\cdot f_2(z)-f_1(z)\cdot f_2'(z)}{f_2^2(z)};$

5. $(c\cdot f_1(z))'=c f_1'(z), c\in C;$

6. Якщо функція $\varphi(z)$ диференційована в точці z, а функція $f(w)$ диференційована у відповідній точці $w = \varphi(z)$, то функція $f(\varphi(z))$ диференційовна в точці z і справедливі та рівність:

$$(f(\varphi(z)))'=f'(\varphi)\cdot\varphi'(z);$$

7. Якщо функція $f(z)$ диференційована в точці z, причому $f'(z) \neq 0$, і існує зворотна до неї функція $f^{-1}(w)$, тоді $f^{-1}(w)$ диференційована у відповідній точці $w = f(z)$, і справедлива така рівність:

$$(f^{-1}(w))'=\frac{1}{f'(z)}.$$

Умови Коші-Рімана (або *Ейлера-Даламбера*): для того щоб функція $f(z) = u(x, y) + iv(x, y)$, яка визначена в деякій області D, була диференційована в точці z цієї області, як функція комплексної змінної, необхідно і достатньо, щоб функції $u(x, y)$ та $v(x, y)$ були диференційованими у відповідній точці (x, y) (як функції дійсних змінних), і виконувалися такі умови:

$$\frac{du}{dx}=\frac{dv}{dy}, \qquad \frac{du}{dy}=-\frac{dv}{dx}. \tag{3.23}$$

При виконанні умов цієї теореми похідна функції $f(z)$ може бути представлена у вигляді:

$$f'(z_0)=\frac{du}{dx}+i\frac{dv}{dx}=\frac{dv}{dy}-i\frac{du}{dy}=\frac{du}{dx}-i\frac{du}{dy}=\frac{dv}{dy}+i\frac{dv}{dx}. \tag{3.24}$$

Похідні елементарних функцій $w = z^n$, $w = e^z$, $w = \ln z$, $w = \sin z$, $w = \cos z$, $w = \operatorname{sh} z$, $w = \operatorname{ch} z$, $\arcsin z$, $\arccos z$ та $\operatorname{arctg} z$ знаходяться за тими ж правилами, що й для дійсного аргументу:

Таблиця 3.1

$(z^n)'=nz^{n-1}$	$(\arcsin z)'=\dfrac{1}{\sqrt{1-z^2}}$
$(e^z)'=e^z$	$(\arccos z)'=\dfrac{-1}{\sqrt{1-z^2}}$
$(\ln z)'=\dfrac{1}{z}$	$(\arctan z)'=\dfrac{1}{1+z^2}$
$(\sin z)'=\cos z$	$(\sinh z)'=\cosh z$
$(\cos z)'=-\sin z$	$(\cosh z)'=\sinh z$

Отже, основні елементарні функції комплексної змінної є аналітичними функціями. Будь-яка функція комплексного змінного, що є композицією кінцевого числа основних елементарних функцій, буде аналітичної або диференційованою в області свого визначення.

Дійсна функція $u = u(x, y)$ називається *гармонійною* в області D, якщо вона володіє в D неперервними частинними похідними другого порядку включно і задовольняє в кожній точці D рівняння Лапласа:

$$\frac{d^2u}{dx^2}+\frac{d^2u}{dy^2}=0. \tag{3.25}$$

Теорема: дійсна і уявна частини функції комплексної змінної $w = f(z) = u(x, y) + iv(x, y)$, аналітичної в деякій області D, є гармонійними функціями в тій же області D (зворотне твердження хибне).

Гармонійні функції u та v, які пов'язані між собою умовами Коші-Рімана, називаються *спряженими*.

Теорема: для того щоб дві гармонійні функції $u = u(x, y)$ і $v = v(x, y)$ утворювали аналітичну функцію $w = f(z) = u(x, y) + iv(x, y)$, необхідно і достатньо, щоб вони були спряженими.

Наслідок: аналітична функція $w = f(z) = u(x, y) + iv(x, y)$, з точністю до постійного доданка, може бути задана своєю дійсною або уявною частиною.

Дійсно, якщо функція $f(z) = u + iv$ аналітична в деякій області D, то її дійсна і уявна частини пов'язані умовами Коші – Рімана:

$$\frac{du}{dx}=\frac{dv}{dy}, \qquad \frac{du}{dy}=-\frac{dv}{dx}.$$

Нехай відома одна з частин аналітичної функції, наприклад $u(x, y)$. З умови $du/dx=dv/dy$ можна знайти

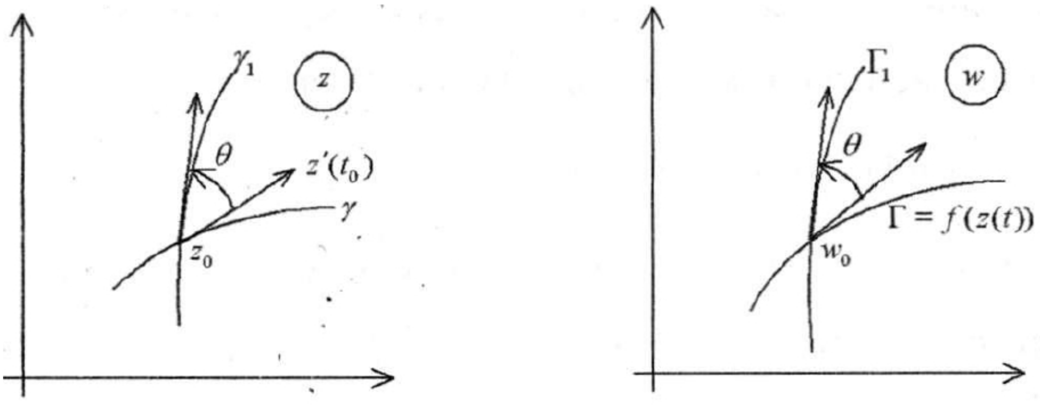
$$v(x,y)=\int\frac{du}{dx}dy+\phi(x)$$

(з точністю до невідомої функції $\varphi(x)$). Цю функцію $\varphi(x)$, з точністю до постійного доданка, знаходимо з другої умови: $du/dy=-dv/dx$.

А точніше:

$$\frac{d}{dx}\left(\int \frac{du}{dx}dy+\phi(x)\right)=-\frac{du}{dy} \quad \text{або} \quad \phi(x)=-\int\left[\left(\int \frac{du}{dx}dy\right)_x'+\frac{du}{dy}\right]dx.$$

Нехай $f(z)$ – функція комплексної змінної, яка визначена й неперервна в деякій області D, і така, що в точці $z_0 \in D$ існує похідна $f'(z_0) \neq 0$. Проведемо через точку z_0 деяку криву γ: $z(t) = x(t) + iy(t)$ ($z(t_0) = z_0$), для якої існує похідна $z'(t_0) \neq 0$ (рис. 3.4а). Ця крива має дотичну в точці z_0 з кутом нахилу рівним $Arg\ z'(t_0)$.



а б

Рис. 3.4

Крива γ перетвориться за допомогою відображення $w = f(z)$ в криву Г, яка розташована в площині w (рис. 3.4б); рівняння кривої Г матиме вигляд $w = f(z(t))$ ($w_0 = f(z_0)$). Оскільки

$$arg f'(z_0) \approx arg(f(z) - f(z_0)) - arg(z - z_0), \tag{3.26}$$

то величині $|f'(z_0)|$ можна дати таку інтерпретацію.

Геометричний зміст величини $Arg\ f'(z_0)$: аргумент похідної $Arg\ f'(z_0)$ дорівнює куту повороту кожної з гладких ліній, що проходять через точку z_0 при відображенні $w = f(z)$ (якщо $f'(z_0) \neq 0$, то відображення $w = f(z)$ зберігає кути між кривими). Оскільки

$$|f'(z_0)| = \lim_{z \to z_0} \frac{|f(z) - f(z_0)|}{|z - z_0|}, \tag{3.27}$$

де числа $|z - z_0|$ і $|f(z) - f'(z_0)|$ являють собою відповідні відстані між точками z і z_0 площині z та між їх образами $f(z)$ і $f'(z_0)$ в площині w, тоді величині $|f'(z_0)|$ можна дати таку інтерпретацію.

Геометричний зміст величини $|f'(z_0)|$: модуль похідної $|f'(z_0)|$ можна розглядати як розтягнення в точці z_0 при відображенні за допомогою функції $w = f(z)$ (якщо $|f'(z_0)| > 1$, то відбувається розтягнення; якщо $|f'(z_0)| < 1$, то відбувається стиснення).

Відображення за допомогою неперервної функції, що зберігає кути між кривими, які проходять через точку z_0, називається *конформним* в цій точці.

Якщо відображення зберігає не тільки величину кутів, але і напрямки їх відліку, то воно називається *конформним відображенням першого роду*; якщо ж напрямки відліку кутів змінюються на протилежні, то введуть мову про *конформне відображення другого роду*.

Отже, відображення за допомогою аналітичної в деякої області G функції $w = f(z)$ комплексно змінної є конформним відображенням першого роду в усіх точках, в яких її похідна відмінна від нуля. При цьому $Arg\, f'(z_0)$ означає кут повороту, а $|f'(z_0)|$ – коефіцієнт розтягування при даному відображенні.

Тема 13.4. Інтегрування функції комплексної змінної

Дуга називається *гладкою*, якщо вона може бути задана неперервно-диференційованою функцією. Крива, що складається з кінцевого числа гладких дуг, називається *кусково-гладкою*.

Нехай на комплексній площині задана орієнтована кусково-гладка крива Г, уздовж якої визначена функція $f(z)$ (рис. 3.5).

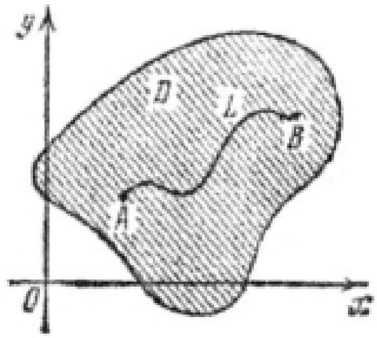


Рис. 3.5

Розіб'ємо цю криву на n частин $[z_{k-1}, z_k]$ точками $z_0, z_1, ..., z_n$, пронумерованими в напрямку від z_0 – початкової точки кривої Г, до z_n – кінцевої точки Г. Позначимо через $\Delta z_k = z_k - z_{k-1}$, $(k = 1, 2, ..., n)$. У кожній частині виберемо довільно точку z_k $(k = 1, 2, ..., n)$ і складемо суму:

$$\sigma = \sum_{k=1}^{n} f(z_k)\Delta z_k, \qquad (3.28)$$

яку назвемо *інтегральною сумою*. Спрямуємо n в нескінченність так, щоб

$$\lambda = \max_k \{|\Delta z_k|\} \to 0.$$

Якщо існує межа (границя) інтегральної суми σ за умови довільного способу розбиття кривої Г на частини, довільного вибору точок z_k $(k = 1, 2, ..., n)$ і за умови

$$\lambda = \max_k \{|\Delta z_k|\} \to 0,$$

то ця границя називається *інтегралом від функції $f(z)$ по дузі (контуру)* Г і позначається:

$$\int_\Gamma f(z)dz = \lim_{\lambda_n \to 0} \sum_{k=1}^{n} f(z_k)\Delta z_k. \qquad (3.29)$$

З визначення інтеграла випливає, що його значення залежать не тільки від функції $f(z)$, але і від шляху інтегрування Г. Якщо крива Г замкнута (точки z_0 і z_n збігаються), то інтеграл позначають символом

$$\oint f(z)dz.$$

Якщо функція $f(z)$ неперервна вздовж кривої Г, то інтеграл (3.29) існує. Нехай $z = x + iy$ і $f(z) = u(x, y) + iv(x, y)$, тоді

$$\int_{\Gamma} f(z)dz=\int_{\Gamma}(u+iv)d(x+iy)=\int_{\Gamma}(u+iv)(dx+idy)=\int_{\Gamma} udx-vdy+i\int_{\Gamma} vdx+udy.$$

Отже, інтеграл від функції комплексної змінної дорівнює сумі двох криволінійних інтегралів 2-го роду від дійсних функцій:

$$\int_{\Gamma} f(z)dz=\int_{\Gamma} u(x,y)dx-v(x,y)dy+i\int_{\Gamma} v(x,y)dx+u(x,y)dy. \qquad (3.30)$$

Якщо крива Г задана параметричних рівнянням: $z = z(t)$, $t_1 \le t \le t_2$ (або $x = x(t)$, $y = y(t)$, $t_1 \le t \le t_2$), тоді

$$\int_{\Gamma} f(z)dz=\int_{t_1}^{t_2} f(z(t))z'(t)dt=\int_{t_1}^{t_2}[u(x(t),y(t))x_t'-v(x(t),y(t))y_t']dt \; +$$

$$+ \; i\int_{t_1}^{t_2}[v(x(t),y(t))x_t'+u(x(t),y(t))y_t']dt.$$

(часто як параметр t вибирають кут: $\varphi = arg\, z$).

Оскільки інтеграл від функції комплексної змінної зводиться до криволінійного інтегралу другого роду від дійсних функцій, то його властивості будуть аналогічні відомим нам властивостям криволінійних інтегралів.

Властивості інтеграла від функції комплексної змінної:

1. $\int_{\Gamma^+} f(z)dz=-\int_{\Gamma^-} f(z)dz;$

де Γ^+ н Γ^- – один і той самий контур, прохідний в додатному і від'ємному напрямках (додатним напрямком обходу замкнутого контуру вважається напрямок, при якому внутрішня область, обмежена даним замкнутим контуром, залишається зліва від напрямку руху);

2. $\int_{\Gamma} a\cdot f(z)dz=a\cdot\int_{\Gamma} f(z)dz, a=const;$

3. $\int_{\Gamma} f(z)dz=\int_{\Gamma_1} f(z)dz+...+\int_{\Gamma_n} f(z)dz;$

4. $\int_{\Gamma}(f_1(z)+f_2(z))dz=\int_{\Gamma} f_1(z)dz+\int_{\Gamma} f_2(z)dz;$

5. Якщо вздовж кривої Г $|f(z)| \le M$ і $|\Gamma|$ - довжина кривої Г, то

$$\left|\int_{\Gamma} f(z)dz\right|\le M\cdot|\Gamma|;$$

6. $\left|\int_{\Gamma} f(z)dz\right|\le\int_{\Gamma}|f(z)|\cdot|dz|=\int_{\Gamma}|f(z)|ds,$

де ds – диференціал довжини дуги.

Задача 3.1. Визначити інтеграл:

$$\int_{\Gamma}\Im zdz,$$

де Г – дуга параболи $y = 2x^2$ від точки 0 до точки $1 + 2i$.

Розв'язок

Позаяк для всіх точок Г має вигляд $y = 2x^2$, то, відповідно,

$$\int_{\Gamma} \Im z dz = \left| \begin{array}{c} \Gamma: y=2x^2, 0 \le x \le 1, \Im z = y = 2x^2 \\ dz = d(x+iy) = d(x+i2x^2) = (1+4ix)dx \end{array} \right| = \int_0^1 2x^2(1+4ix)dx =$$

$$\frac{2x^3}{3}\Big|_0^1 + \frac{8x^4}{4}\Big|_0^1 = \frac{2}{3} + 2i.$$

Теорема Коші: якщо D – однозв'язна область комплексної площини і $f(z)$ – однозначна аналітична в цій області функція, то для будь-якої замкнутої кривої Г, лежачої в області D, інтеграл від $f(z)$ вздовж Г дорівнює нулю, тобто

$$\oint_{\Gamma} f(z)dz = 0.$$

Відзначимо, що теорема Коші залишається справедливою і для багатозв'язної області.

Теорема Коші для n-зв'язної області: якщо D – замкнута n-зв'язна область комплексної площини і $f(z)$ – аналітична в цій області функція, то інтеграл від $f(z)$ по межі області $\overline{D}$ дорівнює нулю (при цьому передбачається, що обхід граничних кривих проводиться в такому напрямку, щоб область $\overline{D}$ залишалася ліворуч).

$$\oint_{\Gamma} f(z)dz = \int_{\Gamma_1^+} f(z)dz + \int_{\Gamma_2^-} f(z)dz = \int_{\Gamma_1^+} f(z)dz - \int_{\Gamma_2^+} f(z)dz = 0,$$

справедлива рівність

$$\int_{\Gamma_1} f(z)dz = \int_{\Gamma_2} f(z)dz.$$

Це означає, що інтеграл від функції $f(z)$, аналітичної в однозв'язній області D, не залежить від кривої Г (від шляху інтегрування), а залежить тільки від початкової і кінцевої точок цієї кривої. Тому для інтеграла вздовж довільної кривої Г, яка з'єднує точки z_0 і z, можна користуватися позначенням

$$\int_{z_0}^{z} f(z)dz.$$

Нехай функція $f(z)$ аналітична в області D. Функція $F(z)$ називається *первісною функції* $f(z)$, якщо $F'(z) = f(z)$ для всіх точок із D.

Легко показати, що функція $F(z)$ також буде аналітичною в області D.

Теорема: якщо функція $F(z)$ є деякою первісної функції $f(z)$, то сукупність усіх первісних функції $f(z)$ задається формулою

$$F(z) + C, C = const.$$

Невизначеним інтегралом від функції $f(z)$ називається сукупність всіх первісних функції $f(z)$ і позначається символом

$$\int f(z)dz.$$

Отже,

$$\int f(z)dz = F(z) + C.$$

Формула Ньютона-Лейбніца: якщо $f(z)$ – аналітична функція в однозв'язній області D, то

$$\int_{z_1}^{z_2} f(z)\,dz = F(z_2) - F(z_1). \qquad (3.31)$$

де $F(z)$ – будь-яка первісна функції $f(z)$, z_0 і z – будь-які дві точки з області D, інтегрування ведеться за довільним напрямком, який лежить в області D.

Цей результат дозволяє зводити обчислення інтеграла від аналітичної функції $f(z)$ до відшукання будь-якої первісної функції $F(z)$ і, відповідно, використовувати відому таблицю первісних і методи інтегрування функцій дійсної змінної.

Зокрема, якщо функції $f(z)$ і $g(z)$ аналітичні, то буде справедлива *формула інтегрування частинами*:

$$\int_{z_1}^{z_2} f(z)\,d(g(z)) = f(z)\cdot g(z)\Big|_{z_1}^{z_2} - \int_{z_1}^{z_2} g(z)\,d(f(z)). \qquad (3.32)$$

Задача 3.2. Знайти інтеграл:

$$\int_0^{3i} z^2\,dz,$$

використовуючи формулу Ньютона-Лейбніца.

Розв'язок

Застосуємо формулу Ньютона-Лейбніца:

$$\int_0^{3i} z^2\,dz = \frac{z^3}{3}\Big|_0^{3i} = \frac{(3i)^3}{3} - 0 = 9i.$$

Тема 13.5. Інтегральна формула Коші

Інтегральна формула Коші виражає фундаментальну властивість аналітичних функцій.

Інтегральна формула Коші: нехай D – однозв'язна область, обмежена замкнутим контуром Г, а функція $f(z)$ – аналітична в області D і неперервна на $D \cup Г$. Тоді для будь-якої точки $z_0 \in D$ справедлива формула Коші:

$$f(z_0) = \frac{1}{2\pi i} \oint_{Г} \frac{f(z)}{z - z_0}\,dz. \qquad (3.33)$$

Теорема про існування похідних аналітичної функції: якщо функція $f(z)$ аналітична в області D, то в кожній точці $z_0 \in D$ у функції $f(z)$ існують всі похідні вищих порядків, які можуть бути знайдені за допомогою формули

$$f(z_0) = \frac{1}{2\pi i} \oint_{Г} \frac{f(z)}{z - z_0}\,dz. \qquad (3.34)$$

Теорема про середнє: якщо функція $f(z)$ неперервна в замкнутому колі $|z-z_0| < R$ і аналітична всередині цього кола, то її значення в центрі кола дорівнює середньому значенню на колі, тобто

$$f(z_0)=\frac{1}{2\pi}\int_0^{2\pi} f(z_0+R\cdot e^{i\phi})\,d\phi. \qquad (3.35)$$

Нерівності Коші: якщо функція $f(z)$ аналітична в колі $|z - z_0| < R$, то в точці z_0 виконуються нерівності

$$\left|f^{(n)}(z_0)\right|\le\frac{Mn!}{R^n},\ n=0,1,2,\ldots, \qquad (3.36)$$

де M – максимум модуля функції $f(z)$ на колі $|z - z_0| = R$.

Теорема Ліувілля: якщо функція $f(z)$ аналітична і обмежена на всій комплексній площині C, то $f(z)$ є константою.

Теорема Морера: нехай D – однозв'язна область, а функція $f(z)$ неперервна в D, і інтеграл від $f(z)$ по будь-якому замкнутому контуру, який цілком лежить в D, дорівнює нулю. Тоді для функції $f(z)$ існує первісна в області D.

Інтегральна формула Коші і формула для похідних вищих порядків дають зручний спосіб для обчислення інтегралів по замкнутому контуру, що охоплює особливі точки функції $f(z)$.

Задача 3.3. Обчислити інтеграл

$$\oint_{|z-2|=3}\frac{e^{z^2}}{z^2-6z}\,dz,$$

використовуючи інтегральну формулу Коші.

Розв'язок

Розглянемо підінтегральну функцію

$$\phi(z)=\frac{e^{z^2}}{z^2-6z}.$$

Аналітичність цієї функції порушується в точках: $z = 0$, $z = 6$ (точки, в яких знаменник дорівнює 0).

Контур $|z-2| = 3$, по якому обчислюється інтеграл, є колом із центром в точці $C = 2$ і радіусом 3. Усередині цього контуру лежить точка $z_0 = 0$, тому всередині контуру підінтегральна функція не аналітична.

Запишемо цю функцію у вигляді

$$\phi(z)=\frac{\frac{e^{z^2}}{z-6}}{z}=\frac{f(z)}{z},\ f(z)=\frac{e^{z^2}}{z-6}.$$

Тоді функція $f(z)$ аналітична всередині замкнутого контуру Скористаємося інтегральною формулою Коші:

$$\oint_{|z-2|=3}\frac{e^{z^2}}{z^2-6z}\,dz=\oint_{|z-2|=3}\frac{\frac{e^{z^2}}{z-6}}{z}\,dz=\left|f(z)=\frac{e^{z^2}}{z-6},\ z_0=0\right| =$$

$$= 2\pi i f(0)=2\pi i\frac{e^0}{-6}=-\frac{\pi i}{3}.$$

Лекція 14. Ряди аналітичних функцій

Тема 14.1. Числові ряди. Ознаки збіжності

Нехай $\{z_n = x_n + iy_n\}$ $(n = 1, 2, ...)$ – послідовність комплексних чисел. *Числовим рядом* називається вираз вигляду:

$$z_1+z_2+\ldots+z_n=\sum_{n=1}^{\infty} z_n, \qquad (3.37)$$

числа z_1, z_2 ... – його члени (z_n – загальний член ряду).

Суми типу

$$S_n=z_1+z_2+\ldots+z_n=\sum_{k=1}^{n} z_k, \qquad (n=1,2,\ldots),$$

називаються *частковими сумами ряду*.

Числовий ряд називається *збіжним*, якщо існує кінцева границя S послідовності часткових сум ряду, тобто

$$\lim_{n\to\infty} S_n=S.$$

В цьому випадку число S називається *сумою ряду* і позначається

$$S=\sum_{n=1}^{\infty} z_n.$$

Якщо послідовність часткових сум ряду розбігається, тобто границя послідовності часткових сум ряду дорівнює нескінченності або не існує, то ряд називається *розбіжним*.

Числовий ряд називається *абсолютно збіжним*, якщо збіжний ряд, складений із модулів його членів, тобто ряд

$$\sum_{n=1}^{\infty} |z_n|=|z_1|+|z_2|+\ldots+|z_n|+\ldots. \qquad (3.38)$$

Збіжний ряд називається *умовно збіжним*, якщо ряд складений з модулів його членів, розбігається.

Ряд

$$\sum_{k=n+1}^{\infty} z_k,$$

отриманий з ряду

$$\sum_{k=1}^{\infty} z_k,$$

шляхом відкидання перших n членів, називається *n-м залишком ряду*.

Помітивши, що

$$S_n=\sum_{k=1}^{n} x_k+i\sum_{k=1}^{n} y_k,$$

і співвідношення

$$\lim_{n\to\infty} S_n=S,$$

еквівалентне двом співвідношенням

$$\lim_{n\to\infty}\sum_{k=1}^{n} x_k=\sigma\,,\quad \lim_{n\to\infty}\sum_{k=1}^{n} y_k=\tau\,,$$

робимо висновок, що збіжність ряду з комплексними елементами еквівалентна одночасній збіжності рядів, складених з дійсних і уявних частин даного ряду.

Завдяки цьому, багато важливих властивостей рядів з дійсними членами переносяться на випадок рядів з комплексними членами, а саме:

1. *Необхідна умова* (або *ознака*) *збіжності*. Якщо ряд

$$\sum_{n=1}^{\infty} z_n\,,$$

збігається, то

$$\lim_{n\to\infty} z_n=0.$$

2. Сума (різниця) двох збіжних рядів є збіжним рядом, тобто

$$\sum_{n=1}^{\infty} z_n\pm\sum_{n=1}^{\infty} w_n=\sum_{n=1}^{\infty}(z_n\pm w_n)=S+\sigma\,. \tag{3.39}$$

3. Постійний множник можна винести за знак ряду, тобто якщо збігається ряд

$$\sum_{n=1}^{\infty} z_n=S\,,$$

то для будь-якого комплексного числа λ ряд

$$\sum_{n=1}^{\infty}(\lambda z_n)\,,$$

теж збігається і справедлива рівність

$$\sum_{n=1}^{\infty}(\lambda z_n)=\lambda S\,. \tag{3.40}$$

4. Якщо до збіжної ряду додати (відняти) кінцеве число членів, то вийде також збіжний ряд.

5. *Критерій збіжності Коші для рядів.* Ряд вигляду

$$\sum_{n=1}^{\infty} z_n\,,$$

збігається тоді і тільки тоді, коли для будь-якого $\varepsilon > 0$ існує таке $N(\varepsilon)$, що при всіх $n > N(\varepsilon)$ і при будь-якому натуральному p виконується нерівність:

$$\left|\sum_{k=n+1}^{n+p} z_k\right|<\varepsilon\,. \tag{3.41}$$

6. Якщо збігається ряд

$$\sum_{n=1}^{\infty}|z_n|\,,$$

то ряд

$$\sum_{n=1}^{\infty} z_n$$

також збігається (зворотне твердження хибне).

Дослідити питання про абсолютну збіжність ряду

$$\sum_{n=1}^{\infty} z_n$$

та збіжність ряду

$$\sum_{n=1}^{\infty} |z_n|$$

можна на підставі будь-якої ознаки збіжності дійсних рядів з невід'ємними членами, наприклад, за ознакою порівняння, ознакою Даламбера, ознакою Коші та інші.

Ознака порівняння. Нехай ряди

$$\sum_{n=1}^{\infty} |z_n| \quad \text{та} \quad \sum_{n=1}^{\infty} |w_n|$$

такі, що їх члени $|z_n|$ і $|w_n|$, починаючи з деякого номера N, задовольняють нерівність $|z_n| \le |w_n|$. Тоді

1. Якщо збігається ряд $\sum_{n=1}^{\infty} |w_n|$, то і ряд $\sum_{n=1}^{\infty} |z_n|$ теж збігається.
2. Якщо ряд $\sum_{n=1}^{\infty} |z_n|$ розбіжний, то і ряд $\sum_{n=1}^{\infty} |w_n|$ – розбіжний.

Ознака д'Аламбера. Нехай ряд

$$\sum_{n=1}^{\infty} |z_n|$$

такий, що існує кінцева границя

$$\lim_{n \to \infty} \frac{|z_{n+1}|}{|z_n|} = D.$$

Тоді

1. Якщо $D < 1$, то $\sum_{n=1}^{\infty} |z_n|$ збіжний;
2. Якщо $D > 1$, то $\sum_{n=1}^{\infty} |z_n|$ розбіжний;
3. Якщо $D = 1$, то ознака не дає відповіді на питання про збіжність ряду.

Радикальна ознака Коші. Нехай ряд

$$\sum_{n=1}^{\infty} |z_n|$$

такий, що існує кінцева границя

$$\lim_{n \to \infty} \sqrt[n]{|z_n|} = K.$$

Тоді

1. Якщо $K < 1$, то $\sum_{n=1}^{\infty} |z_n|$ збіжний;
2. Якщо $K > 1$, то $\sum_{n=1}^{\infty} |z_n|$ розбіжний;;
3. Якщо $K = 1$, то ознака не дає відповіді на питання про збіжність ряду.

Задача 3.4. Дослідити на збіжність ряд

$$\sum_{n=1}^{\infty}\left(\frac{1-3n}{2n+1}+i\,\frac{n+1}{2n+1}\right).$$

Розв'язок

Перевіримо ряд на необхідну умову (ознаку) збіжності. Оскільки

$$\lim_{n=1}^{\infty}\left[\frac{1-3n}{2n+1}+i\,\frac{n+1}{2n+1}\right] = \lim_{n\to\infty}\left[\frac{1-3n}{2n+1}\right] + i\lim_{n\to\infty}\left[\frac{n+1}{2n+1}\right]$$

$$\lim_{n\to\infty}\left(\frac{1-3n}{2n+1}\right)^{\frac{\infty}{\infty}} = \lim_{n\to\infty}\left(\frac{\frac{1}{n}-3}{2+\frac{1}{n}}\right)=-\frac{3}{2}, \quad \lim_{n\to\infty}\left(\frac{n+1}{2n+1}\right)^{\frac{\infty}{\infty}} = \lim_{n\to\infty}\left(\frac{1+\frac{1}{n}}{2+\frac{1}{n}}\right)=\frac{1}{2},$$

$$\lim_{n\to\infty}\left[\frac{1-3n}{2n+1}+i\,\frac{n+1}{2n+1}\right]=-\frac{3}{2}+i\,\frac{1}{2}\neq 0.$$

Отже, необхідна умова збіжності порушена. Тому ряд розбіжний.

Тема 14.2. Функціональні ряди

Функціональним рядом називається вираз вигляду

$$f_1(z)+f_2(z)+\ldots+f_n(z)+\ldots=\sum_{n=1}^{\infty}f_n(z), \qquad (3.42)$$

де $f_1(z)$, $f_2(z)$, ... – функції, визначені в деякій області D (загальна для всіх функцій $f_n(z)$).

При кожному фіксованому $z \in D$ розглянутий функціональний ряд:

$$\sum_{n=1}^{\infty}f_n(z)$$

перетворюється у відповідний йому числовий ряд, який при одних значеннях $z \in D$ збігається, а при інших – розбігається. Якщо в точці z відповідний числовий ряд збігається (розбігається), то z називається *точкою збіжності* (*точкою розбіжності*) функціонального ряду

$$\sum_{n=1}^{\infty}f_n(z).$$

Множиною збіжності функціонального ряду називається множина всіх точок $z \in D$, в яких даний ряд збігається.

На множині збіжності функціонального ряду визначена функція $S(z)$, яка є в кожної точки z множини збіжності функціонального ряду сумою відповідного збіжного числового ряду.

Ряд називається *рівномірно збіжним* в області D до функції $S(z)$, якщо для будь-якого $\varepsilon > 0$ знайдеться такий $N(\varepsilon)$, що для всіх $n > N(\varepsilon)$ і всіх точок $z \in D$ виконано нерівність $|S(z) - S_n(z)| < \varepsilon$, де $S_n(z)$ – часткова сума ряду

$$\sum_{n=1}^{\infty} f_n(z).$$

Зверніть увагу! Зі збіжності ряду в D не випливає його рівномірна збіжність в області D.

Ознака рівномірної збіжності Веєрштрасса: нехай дано функціональний ряд

$$\sum_{n=1}^{\infty} f_n(z).$$

Якщо ряд

$$\sum_{n=1}^{\infty} a_n$$

з постійними додатними членами збігається і $|f_n(z)| \le a_n$, при всіх $n > N$ у всіх точках z з області D, то ряд

$$\sum_{n=1}^{\infty} f_n(z)$$

абсолютно і рівномірно збігається в D.

Ряд

$$\sum_{n=1}^{\infty} a_n \tag{3.43}$$

називається *мажоруючим* для ряду $\sum_{n=1}^{\infty} f_n(z)$.

Справедливі такі *властивості рівномірно збіжних рядів*:

1. Якщо функції $f_n(z)$ неперервні в області D і ряд $\sum_{n=1}^{\infty} f_n(z)$ рівномірно збігається в цій області, то сума ряду $S(z)$ також неперервна в D.
2. Якщо ряд $\sum_{n=1}^{\infty} f_n(z)$, членами якого є неперервні в області D функції $f_n(z)$, рівномірно збігається в цій області до функції $S(z)$, то його можна почленно інтегрувати уздовж будь-якої кривої Г, яка цілком лежить в області D, та справедлива рівність

$$\int_{\Gamma} S(z)dz = \int_{\Gamma} f_1(z)dz + \int_{\Gamma} f_2(z)dz + \ldots = \sum_{n=1}^{\infty} \int_{\Gamma} f_n(z)dz.$$

3. Якщо функція $g(z)$ обмежена в області D і ряд $\sum_{n=1}^{\infty} f_n(z)$ рівномірно збігається в D до функції $S(z)$, то ряд $\sum_{n=1}^{\infty} g(z) f_n(z)$ також рівномірно збігається в цій області до функції $g(z)S(z)$.

Теорема Веєрштрасса: нехай ряд $\sum_{n=1}^{\infty} f_n(z)$, членами якого є однозначні і аналітичні в області D функції $f_n(z)$, рівномірно збігаються в кожному замкнутому колі, що міститься в D. Тоді:

1. Сума ряду $S(z)$ є аналітичною в області D.

2. Ряд $\sum_{n=1}^{\infty} f_n(z)$ можна почленно диференціювати будь-яку кількість разів, і справедливою буде рівність

$$\sum_{n=1}^{\infty} f_n^{(k)}(z) = S^{(k)}(z), \qquad k \in \mathbb{N}.$$

3. Всі ряди $\sum_{n=1}^{\infty} f_n^{(k)}(z)$, отримані в результаті диференціювання ряду $\sum_{n=1}^{\infty} f_n(z)$, рівномірно збігаються в кожному замкнутому колі, що міститься в D.

Задача 3.5. Дослідити на рівномірну збіжність ряд вигляду

$$1 + z + z^2 + \ldots = \sum_{k=0}^{\infty} z^k.$$

Розв'язок

Оскільки члени ряду – це геометрична прогресія з $q = z$, то

$$S_n(z) = \sum_{k=0}^{n-1} z^k = \frac{1 - z^n}{1 - z}.$$

При $|z| < 1$ існує

$$\lim_{n \to \infty} S_n(z) = \lim_{n \to \infty} \frac{1 - z^n}{1 - z} = \frac{1}{1 - z}.$$

Тому цей ряд збігається в середині кругу $|z| < 1$ (його сума: $S_n(z) = 1/(1 - z)$).

Згідно з означенням про рівномірну збіжність, в колі $|z| < 1$: для будь-якого $\varepsilon > 0$ знайдеться таке $N(\varepsilon)$, що при $n >$ N: $|S(z) - S_n(z)| < \varepsilon$ для всіх z з кола $|z| < 1$. В даному випадку, оскільки

$$|S(z) - S_n(z)| = \left| \frac{1}{1 - z} - \frac{1 - z^n}{1 - z} \right| = \left| \frac{z^n}{1 - z} \right|,$$

повинна також виконуватись нерівність

$$\left| \frac{z^n}{1 - z} \right| < \varepsilon, \qquad |z| < 1.$$

Однак, яким би не було значення n, $|S(z) - S_n(z)| < \varepsilon$ не виконуватиметься для всіх z, взятих з одиничного кола, і, відповідно, рівномірної збіжності в цьому випадку (колі) немає. Проте в будь-якому колі $|z| < r$ ($r < 1$) буде мати місце рівномірна збіжність, тому що

$$|S(z) - S_n(z)| = \left| \frac{z^n}{1 - z} \right| \le \frac{|z|^n}{1 - |z|} < \frac{r^n}{1 - r}, \qquad |z| < r.$$

Тема 14.3. Степеневі ряди

Степеневі ряди, будучи окремим випадком функціональних рядів, відіграють винятково важливу роль у теорії функцій.

Степеневим рядом називається ряд вигляду

$$\sum_{n=0}^{\infty} c_n (z-z_0)^n = c_0 + c_1(z-z_0) + c_2(z-z_0)^2 + ... + c_n(z-z_0)^n + ..., \qquad (3.44)$$

де z_0, c_0, c_1, ... – фіксовані комплексні числа (c_0, c_1, ... – коефіцієнти ряду, z_0 – його центр), а z – комплексна зміна.

Теорема Коші-Адамара: нехай

$$R = \frac{1}{\overline{\lim\limits_{n\to\infty}} \sqrt[n]{c_n}}.$$

Тоді степеневий ряд

$$\sum_{n=0}^{\infty} c_n (z-z_0)^n:$$

1. При $R = 0$ збігається тільки в точці z_0.
2. При $R = \infty$ абсолютно збігається на всій комплексній площині C.
3. При $0 < R < \infty$ збігається абсолютно в точках кола $|z - z_0| < R$ і розбігається за межами цього кола ($|z - z_0| > R$).

У разі, коли $0 < R < \infty$, число R називається *радіусом збіжності степеневого ряду*, а коло $|z - z_0| > R$ – колом збіжності степеневого ряду.

Застосовуючи ознаки збіжності рядів д'Аламбера та Коші, можна знайти радіус збіжності степеневого ряду. Наприклад, застосувавши до ряду

$$\sum_{n=0}^{\infty} c_n (z-z_0)^n$$

ознаку д'Аламбера (в припущенні, що відповідна границя існує), отримаємо

$$\lim_{n\to\infty} \left| \frac{c_{n+1}(z-z_0)^{n+1}}{c_n(z-z_0)^n} \right| = \lim_{n\to\infty} \left| \frac{c_{n+1}(z-z_0)}{c_n} \right| = |z-z_0| \lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|} = D.$$

Згідно з ознакою д'Аламбера, ряд буде збігатись, якщо $D < 1$, і розбігатись, якщо $D > 1$. Тому радіус збіжності можна знайти за формулою

$$R = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|}. \qquad (3.45)$$

Аналогічно, застосовуючи ознаку Коші, отримаємо:

$$R = \frac{1}{\overline{\lim\limits_{n\to\infty}} \sqrt[n]{c_n}}. \qquad (3.46)$$

Теорема Абеля:

1. Якщо ряд $\sum\limits_{n=0}^{\infty} c_n (z-z_0)^n$ збігається в точці $z_1 \neq z_0$, то він абсолютно збігається в колі $|z - z_0| < |z_1 - z_0|$.

2. Якщо ряд $\sum\limits_{n=0}^{\infty} c_n (z-z_0)^n$ розбігається в деякій точці z_2, то він розбігається поза колом $|z - z_0| > |z_2 - z_0|$ (рис. 3.6).

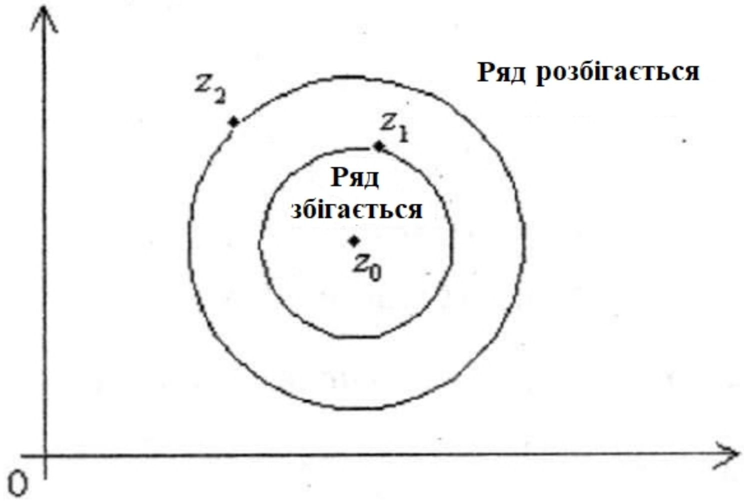


Рис. 3.6

Справедливі такі *властивості степеневих рядів*:

1. Ряд $\sum\limits_{n=0}^{\infty} c_n (z-z_0)^n$ збігається абсолютно і рівномірно в будь-якому колі $|z-z_0| < r < R$, що лежить всередині кола збіжності.
2. Сума $S(z)$ степеневого ряду є аналітичної функцією всередині кола збіжності.
3. Степеневий ряд $\sum\limits_{n=0}^{\infty} c_n (z-z_0)^n = S(z)$ можна почленно інтегрувати всередині кола збіжності, тобто

$$\int_{z_0}^{z'} S(z)\,dz = c_0 (z'-z_0) + \frac{c_1}{2}(z'-z_0)^2 + \frac{c_2}{3}(z'-z_0)^3 + \ldots + \frac{c_n}{n+1}(z'-z_0)^{n+1} + \ldots =$$

$$= \sum_{n=0}^{\infty} \frac{c_n}{n+1}(z'-z_0)^{n+1},$$

де z' – довільна точка всередині кола збіжності вихідного ряду.
4. У кожній внутрішній точці кола збіжності сума степеневого ряду

$$S(z) = \sum_{n=0}^{\infty} c_n (z-z_0)^n = c_0 + c_1(z-z_0) + c_2(z-z_0)^2 + \ldots + c_n(z-z_0)^n + \ldots,$$

має похідні будь-якого порядку, причому їх можна знайти шляхом почленного диференціюванням ряду відповідну кількість разів.
5. Отримані при інтегруванні та диференціюванні степеневі ряди мають те саме коло збіжності, що й вихідний ряд.

Зверніть увагу! При почленній інтеграції та диференціюванні степеневого ряду збіжність (або розбіжність) у граничних точках кола збіжності може не зберігатися.

Задача 3.6. Знайти радіус збіжності степеневого ряду

$$\sum_{n=0}^{\infty} (\sin i\, n)\, z^n.$$

Розв'язок

Щоб скористатися формулою Коші-Адамара, обчислимо границю:

$$\rho=\overline{\lim_{n\to\infty}}\sqrt[n]{|\sin i\,n|}=\overline{\lim_{n\to\infty}}\sqrt[n]{\frac{1}{2}\left(\mathrm{e}^{n}-\mathrm{e}^{-n}\right)}=\lim_{n\to\infty}\left(\frac{1}{2}\right)^{\frac{1}{n}}\cdot\lim_{n\to\infty}\mathrm{e}\left(1-\mathrm{e}^{-2n}\right)^{\frac{1}{n}}=e.$$

Отже, $R = \rho^{-1} = e^{-1}$.

Тема 14.4. Ряд Тейлора

Якщо функцію $f(z)$ можна розкласти в степеневий ряд в області D і якщо цю функцію можна представити у вигляді суми збіжного в області D степеневого ряду, то існують числа z_0, c_0, c_1, ..., такі, що справедлива рівність

$$f(z)=\sum_{n=0}^{\infty}c_n(z-z_0)^n=c_0+c_1(z-z_0)+c_2(z-z_0)^2+\ldots+c_n(z-z_0)^n+\ldots,\qquad (3.47)$$

Теорема про єдність розкладу функції в степеневий ряд: нехай в колі $U = \{z\colon |z\text{-}z_0| < R\}$ функцію $f(z)$ можна розкласти в степеневий ряд. Тоді $f(z)$ аналітична в колі U функцій і коефіцієнти c_0, c_1, ... однозначно визначаються за формулою Тейлора:

$$c_0=\frac{f(z_0)}{0!};\quad c_1=\frac{f'(z_0)}{1!};\quad c_2=\frac{f''(z_0)}{2!};\quad \ldots;\quad c_n=\frac{f^{(n)}(z_0)}{n!}.\qquad (3.48)$$

Рядом Тейлора аналітичної функції $f(z)$ називається степеневий ряд $\sum\limits_{n=0}^{\infty}c_n(z-z_0)^n$, коефіцієнти якого визначаються за формулою Тейлора.

З цих тверджень, зокрема, випливає, що однозначна функція $f(z)$ буде аналітичною в області D тоді і тільки тоді, коли в межах кожної точки $z_0 \in D$ її можна розкласти в ряд за степенями z - z_0.

Формули Тейлора і формули для похідних дозволяють легко отримати такі Тейлоровскі розклади елементарних функцій ($z_0 = 0$):

$$e^z=1+z+\frac{z^2}{2!}+\ldots+\frac{z^n}{n!}+\ldots=\sum_{n=0}^{\infty}\frac{z^n}{n!};\qquad (3.49)$$

$$\sin z=z-\frac{z^3}{3!}+\frac{z^5}{5!}-\ldots+(-1)^n\frac{z^{2n+1}}{(2n+1)!}+\ldots=\sum_{n=0}^{\infty}(-1)^n\frac{z^{2n+1}}{(2n+1)!};\qquad (3.50)$$

$$\cos z=1-\frac{z^2}{2!}+\frac{z^4}{4!}-\ldots+(-1)^n\frac{z^{2n}}{(2n)!}+\ldots=\sum_{n=0}^{\infty}(-1)^n\frac{z^{2n}}{(2n)!}.\qquad (3.51)$$

Ці ряди збігаються на всій комплексній площині C, тому що функції e^z, $\sin z$ і $\cos z$ аналітичні на всій площині C. Їх можна вважати визначенням функцій e^z, $\sin z$ та $\cos z$.

Отримані формули з використанням методів підстановки, почленного інтегрування і диференціювання рядів дозволяють отримувати розкладання в ряд Тейлора багатьох інших функцій.

Тема 14.5. Властивість єдиності аналітичної функції

Нехай функція $f(z)$ аналітична в області D і $A \in C$. Точка $z_0 \in D$ називається *А-точкою*, якщо $f(z_0) = A$. Якщо А = 0, то точка z_0 називається *0-точкою* (або *нулем*) функції $f(z)$.

Нехай z_0 – нуль аналітичної функції $f(z)$. Можливі два випадки:

1. $f(z) \equiv 0$ на D (тривіальний випадок).
2. Знайдеться така точка $z_1 \in D$, що $f(z_1) \neq 0$.

У другому випадку функція $f(z)$ розкладається в ряд Тейлора в деякому околі U точки z_0, Позаяк z_0 – нуль аналітичної функції $f(z)$, то $f(z_0) = 0$, а отже, $c_0 = 0$. Оскільки $f(z) \neq 0$, то серед інших коефіцієнтів c знайдеться хоча б один ненульовий. Номер n першого відмінного від нуля коефіцієнта c_n називається *кратністю* (або *порядком*) *нуля* z_0. Зокрема, якщо $n = 1$, то нуль z_0 називається *простим* (або *однократним*).

Якщо z_0 – n-кратний нуль аналітичної функції $f(z)$, то розклад функції в ряд Тейлора має вигляд

$$f(z)=c_n(z-z_0)^n+c_{n+1}(z-z_0)^{n+1}+\ldots=\sum_{k=n}^{\infty} c_k(z-z_0)^k =$$

$$(z-z_0)^n\{c_n+c_{n+1}(z-z_0)+\ldots\}=(z-z_0)^n\phi(z),$$

де $\varphi(z) = c_n + c_{n+1}(z - z_0) + \ldots$ – аналітична функція і $\varphi(z_0) = c_n \neq 0$.

Очевидно, що порядок n нуля z_0 дорівнює порядку першої похідної $f^{(n)}(z)$, відмінної від нуля в точці z_0.

Теорема: якщо функція $f(z)$, що не дорівнює тотожно нулю, аналітична в деякому околі точки z_0 і $f(z_0) = 0$, то знайдеться окіл точки z_0, в якій у функції $f(z)$ немає інших нулів, окрім точки z_0.

Точка z_0 називається *ізольований нулем* функції $f(z)$, якщо існує такий окіл точки z_0, в якому у функції $f(z)$ немає інших нулів, крім точки z_0.

Теорема єдиності аналітичної функції: нехай дана область D і $M \subset D$, причому існує гранична точка множини M, яка належить D. Тоді існує не більше однієї аналітичної в області D функції $f(z)$, яка на множині M набуває вказані значення.

Принцип максимуму модуля: нехай $f(z) \neq const$ – функція, однозначна і аналітична в усіх точках області D розширеної площині. Тоді в жодній точці $z_0 \in D$ модуль $|f(z)|$ не може мати максимуму.

Наслідок: якщо функції $f(z)$ неперервна в замкненій області $\overline{D}$ і аналітична всередині цієї області, то $|f(z)|$ досягає максимального значення на границі області.

Тема 14.6. Ряд Лорана

Узагальненим степеневим рядом (або *рядом Лорана*) називається ряд вигляду

$$f(z)=\sum_{-\infty}^{\infty} c_n(z-z_0)^n. \tag{3.52}$$

Оскільки

$$\sum_{-\infty}^{\infty} c_n (z-z_0)^n = \sum_{-\infty}^{-1} c_n (z-z_0)^n + \sum_{0}^{+\infty} c_n (z-z_0)^n = \sum_{1}^{+\infty} \frac{c_{-\nu}}{(z-z_0)^{\nu}} + \sum_{0}^{+\infty} c_n (z-z_0)^n,$$

то ряд $f(z)$ збігається, якщо збігаються обидва ряди:

$$\sum_{n=0}^{\infty} c_n (z-z_0)^n, \qquad \sum_{n=-\infty}^{-1} c_n (z-z_0)^n.$$

Ряд

$$\sum_{n=-\infty}^{-1} c_n (z-z_0)^n,$$

називають *головною частиною ряду Лорана*,

Ряд $\sum_{n=0}^{\infty} c_n (z-z_0)^n$ *– правильною частиною ряду Лорана.*

Отже, ряд $f(z)$ є сумою двох степеневих рядів, один з яких (правильна частина ряду) збігається в середині деякого кола з центром z_0 і радіусом R, а інший – поза колом меншого радіуса r із тим самим центром.

Теорема Лорана про існування розкладу в ряд функції, аналітичної в колі: нехай функція $f(z)$ аналітична в круговому кільці $V = \{r < |z\text{-}z_0| < R\}$. Тоді в кожній точці z цього кільця функцію $f(z)$ можна зобразити у вигляді суми узагальненого степеневого ряду:

$$f(z) = \sum_{-\infty}^{\infty} c_n (z-z_0)^n.$$

Наслідок: коефіцієнти ряду Лорана визначаються за формулами

$$c_n = \frac{1}{2\pi i} \int_{|\zeta - z_0| = \rho} \frac{f(\zeta)\, d\zeta}{(\zeta - z_0)^{n+1}}, \qquad n = 0, \pm 1, \pm 2, \ldots, \qquad r < \rho < R. \tag{3.53}$$

Теорема єдиності розкладання функції в ряд Лорана: якщо в деякому круговому кільці $V = \{r < |z\text{-}z_0| < R\}$ функцію $f(z)$ можна розкласти в ряд Лорана, то це можна зробити єдиним сros обом.

Справедливі такі *властивості*:

1. Ряд Лорана $f(z)$ збігається в круговому кільці $V = \{r < |z\text{-}z_0| < R\}$ зі можливим додаванням деяких (або всіх) точок границі, при цьому не виключено, що $r = 0$ і $R = \infty$.
2. Сума $S(z)$ ряду Лорана $f(z)$ є аналітичною функцією всередині кругового кільця V.
3. Ряд Лорана можна почленно інтегрувати й диференціювати всередині кільця V будь-яку кількість разів. Отримані при цьому ряди мають те саме кільце збіжності (V), що й вихідний ряд; збіжність у граничних точках може не зберігатися.
4. Якщо V є кільцем збіжності ряду Лорана функції $f(z)$ і $0 < r < R$, то на внутрішній і на зовнішній межах кільця V лежать особливі точки функції $f(z)$.

Лекція 15. Теорія лишків та її застосування

Тема 15.1. Ізольовані особливі точки аналітичної функції. Їх класифікація

Точка z_0 називається *особливою точкою функції* $f(z)$, якщо $f(z)$ не аналітична в цій точці.

Особлива точка $z_0 \in \overline{C}$ функції $f(z)$ називається *ізольованою особливою точкою* $f(z)$, якщо існує окіл $0 < |z - z_0| < R$ точки z_0, у всіх точках якої функція $f(z)$ аналітична.

Розрізняють три типи ізольованих особливих точок функції $f(z)$ в залежності від поведінки $f(z)$ в їх околі:

1. Точка z_0 називається *усувною особливою точкою* (або правильною особливою точкою) функції $f(z)$, якщо існує кінцева границя:

$$\lim_{z \to z_0} f(z).$$

2. Точка z_0 називаєтеся *полюсом функції* $f(z)$, якщо

$$\lim_{z \to z_0} f(z) = \infty.$$

3. Точка z_0 називається *істотно особливою точкою* функції $f(z)$, якщо $\lim\limits_{z \to z_0} f(z)$ не існує.

Тип ізольованої особливої точки z_0 функції $f(z)$ залежить від виду головної частини ряду Лорана.

Теорема: для того щоб точка z_0 була усувною особливою точкою функції $f(z)$, необхідно і достатньо, щоб ряд Лорана функції $f(z)$ в околі $0 < |z - z_0| < R$ цієї точки не містив головної частини, тобто мав вигляд степеневого ряду.

Отже, в околі $0 < |z - z_0| < R$ усувної особливої точки z_0 функція $f(z)$ збігається з функцією $S(z)$, яка аналітична у всьому околі точки z_0 (включаючи точку z_0). Визначивши $f(z_0) = S(z_0)$, функцію $f(z)$ можна зробити аналітичною вже в околі $|z - z_0| < R$ точки z_0 і в такий спосіб «усунути» особливість у точці z_0.

Кратність (порядок) нуля z_0 функції $g(z) = 1/f(z)$ називається *порядком полюса* z_0 функції $f(z)$.

Теорема: для того щоб точка z_0 була n-кратним полюсом функції $f(z)$, необхідно і достатньо, щоб головна частина ряду Лорана функції $f(z)$ в околі $0 < |z - z_0| < R$ точки z_0 містила лише кінцеву кількість нульових членів, тобто ряд Лорана мав вигляд

$$f(z) = \sum_{k=-n}^{\infty} c_k (z - z_0)^k, \qquad n > 0.$$

Таким чином, порядок полюса z_0 функції $f(z)$ дорівнює номеру старшого ненульового коефіцієнта головної частини ряду Лорана функції $f(z)$ в околі $0 < |z - z_0| < R$ точки z_0.

Наслідок: точка z_0 є полюсом порядку n функції $f(z)$ тоді і тільки тоді, коли функція $f(z)$ подана у вигляді

$$f(z)=\frac{h(z)}{(z-z_0)^n},$$

де h(z) – аналітична функція в околі $|z - z_0| < R$ точки z_0 і $h(z_0) \neq 0$.

Теорема: для того щоб точка z_0 була істотно особливою точкою функції $f(z)$ необхідно і достатньо, щоб головна частина ряду Лорана функції $f(z)$ в околі $0 < |z - z_0| < R$ точки z_0 містила нескінченну кількість нульових членів, тобто ряд Лорана мав вигляд

$$f(z)=\sum_{-\infty}^{\infty} c_n (z-z_0)^n.$$

Теорема Сохоцкого: нехай $z_0 \in \overline{C}$ – істотно особлива точка функції $f(z)$. Тоді для будь-якого $A \in \overline{C}$ знайдеться послідовність точок $\{z_n\}_{n=1}^{\infty}$, така, що $z_n \to z_0$ і

$$\lim_{n\to\infty} f(z_n)=A.$$

Зверніть увагу! Для дослідження ізольованих особливих точок часто буває корисними використовувати вже відомі тейлоровські розклади основних елементарних функцій.

Нескінченність як особлива точка. Розглянемо однозначну функцію $f(z)$, аналітичну у всіх точках множини $|z| > R$. Виконуючи перетворення $w = 1/z$, ми зведемо дослідження такої функції до вивчення функції $f^*(\mathrm{w}) = f(1/w)$, аналітичної в усіх точках околі початку координат, крім, можливо, самого початку координат. При цьому $w = 0$ буде служити образом нескінченно віддаленої точки $z = \infty$. Залежно від того, чи буде точка $w = 0$ усувною, полюсом або істотно особливою для $f^*(w) = f(1/w)$, називатимемо точку $z = \infty$ усувною, полюсом або істотно особливою відповідно.

Таким чином, зв'язки між характером точки по відношенню до функції і відповідним розкладам в ряд Лорана виходять такі ж, як і у випадку кінцевої точки, тільки ролі членів з додатними і від'ємними степенями змінюються між собою.

Теорема: для того щоб точка $z_0 = \infty$ була усувною особливою точкою функції $f(z)$ необхідно і достатньо, щоб ряд Лорана в «проколотому» околі цієї точки мав вигляд

$$f(z)=\sum_{n=0}^{-\infty} c_n z^n = c_0+\frac{c_{-1}}{z}+\frac{c_{-2}}{z^2}+\ldots$$

Теорема: для того, щоб точка $z_0 = \infty$ була полюсом функції $f(z)$ необхідно і достатньо, щоб її ряд Лорана мав вигляд

$$f(z)=\sum_{n=0}^{-\infty} c_n z^n+(c_1 z+c_2 z^2+\ldots+c_N z^N), \qquad N>0.$$

Теорема: для того щоб точка $z_0 = \infty$ була істотно особливою точкою функції $f(z)$, необхідно і достатньо, щоб її ряд Лорана мав вигляд

$$f(z)=\sum_{n=-\infty}^{+\infty} c_n z^n.$$

Задача 3.7. Знайти ізольовані особливі точки функції:

$$f(z)=\frac{z-1}{(z^2+1)(z+3)^2}$$

та визначити їх тип.

Розв'язок

Функція має три особливі точки $z_1 = i$, $z_2 = -i$, $z_3 = -3$, в яких знаменник функції перетворюється в нуль.

Для $z_1 = i$ одержимо:

$$f(z)=\frac{z-1}{(z-i)(z+i)(z+3)^2}=\frac{1}{(z-i)}\cdot\frac{z-1}{(z+i)(z+3)^2}=\frac{h_1(z)}{z-i},\quad h_1(z)=\frac{z-1}{(z+i)(z+3)^2}.$$

Оскільки $h_1(i) \neq 0$, то z_1 – полюс 1-го порядку.

Аналогічно доводиться, що $z_2 = -i$ теж є полюсом 1-го порядку.

Для $z_3 = -3$ одержимо:

$$f(z)=\frac{z-1}{(z-i)(z+i)(z+3)^2}=\frac{h_3(z)}{(z+3)^2},\quad h_3(z)=\frac{z-1}{(z-i)(z+i)}.$$

Позаяк $h_3(-3) \neq 0$, то z_3 – полюс 2-го порядку.

Тема 15.2. Лишки

Нехай z_0 – ізольована особлива точка функції $f(z)$. Тоді $f(z)$ аналітична в деякому проколотому околі $V = \{0 < |z - z_0| < R\}$ точки z_0, і ряд Лорана має вигляд

$$f(z)=\sum_{n=-\infty}^{\infty} c_n (z-z_0)^n.$$

Лишком функції $f(z)$ в точці z_0 називається коефіцієнт c_{-1}. Позначення:

$$res_{z_0} f = c_{-1}. \tag{3.54}$$

Нехай $z_0 = \infty$ – ізольована особлива точка функції $f(z)$, і ряд Лорана функції $f(z)$ в околі має вигляд

$$f(z)=\sum_{n=-\infty}^{+\infty} c_n z^n.$$

Лишком функції $f(z)$ в точці $z_0 = \infty$ називається число $-c_{-1}$. Позначення:

$$res_{\infty} f = -c_{-1}. \tag{3.55}$$

Теорема: якщо $z_0 \in C$ – ізольована особлива точка функції $f(z)$, і якщо $f(z)$ аналітична на замкнутій кривій γ та у всіх точках площини, обмежених цією кривою, за винятком точки z_0, то

$$res_{z_0} f = \frac{1}{2\pi i}\oint_{\gamma} f(z)\,dz, \tag{3.56}$$

де γ взято в додатному напрямку (проти годинникової стрілки).

Як γ можна взяти коло $|z - z_0| = \rho$ з центром z_0 і досить малим радіусом $\rho < R$.

Теорема: нехай $z_0 = \infty$ – ізольована особлива точка функції $f(z)$ і

$$f(z)=\sum_{n=-\infty}^{+\infty} c_n z^n,\quad (|z|>R).$$

Нехай також γ – замкнута кусково-гладка крива, яка лежить в області $|z| > R$. Тоді

$$res_{z_0} f = -c_{-1} = \frac{1}{2\pi i}\oint_{\gamma} f(z)\,dz, \qquad (3.57)$$

де γ взято у від'ємному напрямку (за годинниковою стрілкою).

Основна теорема про лишки: нехай функція $f(z)$ аналітична в області D і на її границі Γ, за винятком кінцевої кількості точок, що лежать усередині D. Тоді

$$\oint_{\Gamma} f(z)dz = 2\pi i \sum_{k=1}^{n} res_{z_k} f, \qquad (3.58)$$

де контур Γ узятий в додатному напрямку.

В усувній особливій точці $z_0 \in C$ лишок дорівнює нулю, тому що ряд Лорана в проколотому околі цієї точки складається тільки з правильної частини.

Теорема: нехай $z_0 \in C$ – простий полюс функції $f(z)$. Тоді

$$res_{z_0} f = \lim_{z \to z_0} (z - z_0) f(z). \qquad (3.59)$$

Теорема: нехай проколотому околі точки z_0 функція $f(z)$ має вигляд $f(z) = \varphi(z)/\psi(z)$, де функції $\varphi(z)$ і $\psi(z)$ аналітичні в точці z_0, та $\varphi(z_0) \neq 0$, $\psi(z_0) = 0$ і $\psi'(z_0) \neq 0$. Тоді

$$res_{z_0} f = \frac{\phi(z_0)}{\psi'(z_0)}. \qquad (3.60)$$

Теорема: нехай z_0 – n-кратний полюс ($n > 1$) функції $f(z)$. Тоді

$$res_{z_0} f = \frac{1}{(n-1)!} \lim_{z \to z_0} \left((z - z_0)^n f(z) \right)^{(n-1)}. \qquad (3.61)$$

Для обчислення лишків в істотно особливих точках аналогічних формул не існує і треба знаходити головні частини розкладу Лорана.

Теорема про суму лишків: нехай у функції $f(z)$ на площині існує лише кінцева кількість особливих точок $z_1, z_2, ..., z_n$. Тоді

$$\sum_{k=1}^{n} res_{z_k} f + res_{\infty} f = 0. \qquad (3.62)$$

Задача 3.8. Знайти лишки функції:

$$f(z) = \frac{z-1}{(z^2+1)(z+3)^2},$$

в її особливих точках і в точці ∞.

Розв'язок

У попередній задачі (задача 3.7) було встановлено, що $f(z)$ має три особливі точки $z_1 = i$, $z_2 = -i$ та $z_3 = -3$. Причому перші дві з них – полюси першого порядку, а третя – полюс другого порядку.

Для $z_1 = i$ одержимо

$$res_i f = \lim_{z \to i} \frac{(z-i)(z-1)}{(z^2+1)(z+3)^2} = \lim_{z \to i} \frac{(z-i)(z-1)}{(z-i)(z+i)(z+3)^2} = \lim_{z \to i} \frac{(z-1)}{(z+i)(z+3)^2} =$$

$$= \frac{i-1}{2i(i+3)^2} = \frac{i-1}{2i(-1+6i+9)} = \frac{i-1}{4(-3+4i)} = \frac{(i-1)(-3-4i)}{4((-3)^2+4^2)} = \frac{7+i}{100}.$$

Для $z_2 = -i$ одержимо

$$res_{-i} f = \lim_{z \to -i} \frac{(z+i)(z-1)}{(z-i)(z+i)(z+3)^2} = \lim_{z \to -i} \frac{(z-1)}{(z-i)(z+3)^2} = \frac{(-i-1)}{-2i(-i+3)^2} =$$

$$= \frac{i+1}{2i(-1-6i+9)} = \frac{i+1}{4(3+4i)} = \frac{(i+1)(3-4i)}{4((3)^2+(4i)^2)} = \frac{7-i}{100}.$$

Для $z_3 = -3$ одержимо

$$(z-z_0)^2 f(z) = \frac{(z+3)^2(z-1)}{(z^2+1)(z+3)^2} = \frac{z-1}{z^2+3},$$

$$\left((z-z_0)^2 f(z)\right)' = \left(\frac{z-1}{z^2+3}\right)' = \frac{z^2+1-2z(z-1)}{(z^2+1)^2} = \frac{-z^2+2z+1}{(z^2+1)^2},$$

$$res_{-3} f = \frac{1}{1!} \lim_{z \to -3} \frac{-z^2+2z+1}{(z^2+1)^2} = \frac{-9-6+1}{(9+1)^2} = -\frac{14}{100}.$$

Для $z = \infty$ із рівності (теорема про суму лишків)

$$res_{\infty} f + \frac{7+i}{100} + \frac{7-i}{100} - \frac{14}{100} = 0$$

отримаємо:

$$res_{\infty} f = -\frac{(7+i)}{100} - \frac{(7-i)}{100} + \frac{14}{100} = 0.01(-7-i-7+i+14) = 0.$$

Тема 15.3. Розрахунок деяких класів інтегралів за допомогою лишків

Обчислення інтегралів по замкнутому контуру

Нехай функція $f(z)$ має усередині замкнутого контуру Г тільки ізольовані особливі точки z_k (k = 1, 2, ... n). Тоді інтеграл від функції $f(z)$ по додатно орієнтованому контуру Г можна знайти, застосовуючи основну теорему про лишки:

$$\oint_{\Gamma} f(z) dz = 2\pi i \sum_{k=1}^{n} res_{z_k} f.$$

Якщо функція $f(z)$ має в розширеній комплексній площині $\overline{C}$ тільки ізольовані особливі точки, то, згідно з теоремою про суму лишків, для обчислення інтеграла по замкненому контуру досить знайти лишок у нескінченно віддаленій точці.

Обчислення інтегралів вигляду $\int_0^{2\pi} R(\cos\phi, \sin\phi) d\phi$,

де R – раціональна функція від косинуса та синуса

Такі інтеграли за допомогою заміни:

$$z = e^{i\phi}, \quad d\phi = \frac{dz}{iz} = -\frac{idz}{z},$$

$$\cos\phi = \frac{1}{2}(e^{i\phi} + e^{-i\phi}) = \frac{1}{2}\left(z + \frac{1}{z}\right), \quad \sin\phi = \frac{1}{2i}(e^{i\phi} - e^{-i\phi}) = \frac{1}{2i}\left(z - \frac{1}{z}\right)$$

(при зміні φ від 0 до 2π точка z описує коло $|z| = 1$) зводяться до інтегралів по замкнутому контуру.

Обчислення невласних інтегралів

Інтеграл, визначений рівністю:

$$\int_{-\infty}^{\infty} f(x)dx = \lim_{R\to\infty} \int_{-R}^{R} f(x)dx,$$

де $f(x)$ – функція, задана на всій осі Ox, називатися *невласних інтегралом функції f(x) у розумінні головного значення*.

Якщо існує кінцева границя $\lim\limits_{R\to\infty} \int_{-R}^{R} f(x)dx$, то невласний інтеграл $\int_{-\infty}^{\infty} f(x)dx$ в розумінні головного значення називається *збіжним*; якщо межа не існує або дорівнює нескінченності, то невласний інтеграл $\int_{-\infty}^{\infty} f(x)dx$ у розумінні головного значення називається *розбіжним*.

Теорема: якщо збігається кожен з інтегралів $\int_{-\infty}^{0} f(x)dx = \lim\limits_{R\to\infty} \int_{-R}^{0} f(x)dx$ та $\int_{0}^{\infty} f(x)dx = \lim\limits_{R\to\infty} \int_{0}^{R} f(x)dx$, то невласний інтеграл $\int_{-\infty}^{\infty} f(x)dx$ у розумінні головного значення також збігається і дорівнює сумі цих інтегралів.

Теорема: нехай функція $f(x)$, $x \in (-\infty; +\infty)$ та задовольняє умови:

1. Функція $f(z)$ має лише ізольовані особливі точки в комплексній площині C, причому жодна з них не лежить на осі Ox.
2. Якщо $\gamma(R)$ – півколо радіуса R із центром у початку координат, що лежить у верхній (або в нижній) півплощині, то границя:

$$\lim_{R\to\infty} \int_{\gamma(R)} f(z)dz = 0.$$

Тоді невласний інтеграл $\int_{-\infty}^{\infty} f(x)dx$ у розумінні головного значення дорівнює сумі лишків функції $f(z)$ в особливих точках, що лежать у верхній півплощині, помноженої на $2\pi i$ (відповідно, дорівнює сумі лишків в особливих точках з нижньої півплощини, помноженої на $-2\pi i$).

Лемма Жордана: нехай функція $F(z)$ аналітична в півплощині Im $z \geq -a$, за винятком кінцевої кількості ізольованих особливих точок, і $\lim\limits_{z\to\infty} F(z) = 0$. Якщо, $F(z)$ – дуга кола $|z| = R$, яка розташована в півплощині Im $z \geq -a$, то

$$\lim_{R\to\infty}\int_{\gamma(R)} e^{itz}F(z)\,dz=0, \quad \text{для всіх} \quad t>0.$$

(для випадку $t < 0$ справедливе аналогічне твердження, якщо як $\gamma(R)$ взяти дугу кола $|z| = R$, що лежить в півплощини $\operatorname{Im} z \le -a$).

Задача 3.9. Обчислити інтеграл

$$\int_{|z|=2}\frac{z-1}{(z^2+1)(z+3)^2}\,dz.$$

Розв'язок

У попередній задачі (задача 3.8) були знайдені особливі точки підінтегральної функції $z_1 = i$, $z_2 = -i$, $z_3 = -3$ та, відповідно, лишки в цих точках:

$$res_i f=\frac{7+i}{100}, \quad res_{-i} f=\frac{7-i}{100}, \quad res_{-3} f=-\frac{14}{100}.$$

Усередині кола $|z| = 2$ знаходяться особливі точки $z_1 = i$, $z_2 = -i$, а точка $z_3 = -3$ лежить поза межами кола. Отже,

$$\int_{|z|=2}\frac{z-1}{(z^2+1)(z+3)^2}\,dz=2\pi i(0{,}07+0{,}01i+0{,}07-0.01i)=0{,}28\pi i.$$

Тема 15.4. Логарифмічні лишки. Принцип аргументу

Нехай функція $f(z)$ аналітична в області D і на її границі Γ, за винятком кінцевої кількості точок – полюсів, що лежать всередині D. Крім того, функція $f(z)$ на границі не має нулів, а всередині області D має кінцеву кількість нулів.

Логарифмічним лишком функції $f(z)$ відносно кривої Γ називається число

$$\frac{1}{2\pi i}\int_{\Gamma}\frac{f'z}{f(z)}\,dz, \tag{3.63}$$

де Γ обходиться в додатному напрямку.

Теорема: якщо z_0 – нуль кратності n аналітичної функції $f(z)$, то логарифмічний лишок функції $f(z)$ в точці z_0 дорівнює n; якщо z_0 – полюс порядку p, то логарифмічний лишок дорівнює $-p$.

Теорема про логарифмічний лишок: нехай Γ – замкнутий контур, що лежить в області аналітичності функції $f(z)$, а функція $f(z)$ аналітична в усіх точках всередині Γ, за винятком кінцевої кількості полюсів, і не має на Γ ні нулів, ні полюсів. Тоді різниця між кількістю нулів N і полюсів P (з урахуванням кратності) функції $f(z)$ всередині контуру Γ дорівнює логарифмічному лишку відносно цього контуру, тобто

$$\frac{1}{2\pi i}\int_{\Gamma}\frac{f'z}{f(z)}\,dz=N-P. \tag{3.64}$$

Принцип аргументу: нехай Γ – замкнутий контур, що лежить в області аналітичності функції $f(z)$, а функція $f(z)$ аналітична в усіх точках усередині Γ, за винятком кінцевої кількості полюсів, і не має на контурі Γ, ні нулів, ні полюсів. Тоді різниця між кількістю нулів N і полюсів P (з урахуванням кратності) функції $f(z)$ всередині контуру Γ дорівнює зміні $arg\, f(z)$ при обході точкою z контуру Γ в додатному напрямку, поділеному на 2π:

$$N-P=(\Delta_\Gamma arg f)/2\pi . \quad (3.65)$$

Теорема Руше: якщо функції $f(z)$ і $g(z)$ однозначні й аналітичні у всіх точках замкнутого контуру Г і всередині нього, причому в точках цього контуру справедлива нерівність $|f(z)| > |g(z)|$, то всередині контуру Г функція $f(z) + g(z)$ має стільки ж нулів, скільки і функція $f(z)$.

Лекція 16. Перетворення Лапласа

Тема 16.1. Означення перетворення Лапласа

Перетворення Лапласа є однією з можливих реалізацій так званого принципу відповідності, коли можна встановити відповідність між елементами двох різних класів об'єктів, наприклад між функціями та операторами, між функціями та функціями тощо.

В основі перетворення Лапласа лежить ідея інтегрального перетворення. При цьому, наприклад, розв'язку деякої вихідної задачі (звичайного диференціального рівняння зі сталими коефіцієнтами), що являє собою функцію дійсної змінної $f(t)$, ставиться у відповідність функція комплексної змінної $F(p)$, яка задовольняє деяке алгебраїчне рівняння; або рівнянню в частинних похідних ставиться у відповідність звичайне диференціальне рівняння.

Перетворення Лапласа ставить у відповідність функції $f(t)$ дійсної змінної t функцію $F(p)$ комплексної змінної p за допомогою інтегрального перетворення:

$$F(p)=\int_0^\infty e^{-pt} f(t)\,dt. \quad (3.66)$$

Природно, що цей інтеграл існує не для довільної функції $f(t)$. Тому необхідно вказати клас функцій (клас оригіналів), для яких це перетворення існує.

Розглянемо функцію $f(t)$, яка визначена для всіх значень дійсної змінної $-\infty < t < +\infty$ та задовольняє такі умови:

1. При $t < 0$ функція $f(t) \equiv 0$.
2. При $t \geq 0$ функція $f(t)$ на довільній обмеженій ділянці осі має не більш ніж скінченну кількість точок розриву першого роду.
3. При $t \to \infty$ функція $f(t)$ має обмежену швидкість зростання, тобто для кожної функції цього класу існують такі дійсні сталі $M \geq 0$ та a, що для всіх $t > 0$ виконується умова:

$$|f(t)| \leq M e^{at}.$$

Зазначимо, що інтеграл (3.66) є невласним інтегралом, який залежить від змінної p як від параметра. Із цього випливає, що цей інтеграл збігається не за всіх значень параметра p. Наприклад, якщо функція $f(t)$ при $t \to \infty$ прямує до граничного значення, що відрізняється від нуля, а $Re\ p < 0$, то інтеграл розбігається.

Теорема: інтеграл (3.66) збігається в області $Re\ p > a$, де a – показник зростання функції $f(t)$, причому для довільного $x_0 > a$ інтеграл (3.66) при $Re\ p \geq x_0 > a$ збігається рівномірно.

Для довільного $p = x + iy$ при $x > a$ можна вказати таке $\varepsilon > 0$, що $x > a_1 = a + \varepsilon$, причому

$$|f(t)| < M\,\mathrm{e}^{a_1 t},$$

тоді

$$|F(p)| = \left|\int_0^\infty \mathrm{e}^{-pt} f(t)\,dt\right| \leq M\int_0^\infty \mathrm{e}^{-xt}\mathrm{e}^{a_1 t}\,dt = \frac{M}{x - a_1},\ x > a_1, \tag{3.67}$$

тобто інтеграл збігається при $x > a$. Якщо $x \geq x_0 > a$, то аналогічна оцінка дає

$$|F(p)| \leq M\int_0^\infty \mathrm{e}^{-(x_0 - a_1)t}\,dt = \frac{M}{x_0 - a_1}, \tag{3.68}$$

що доводить рівномірну збіжність.

Функцію $F(p)$, визначену за допомогою (3.66), називають *зображенням Лапласа функції* $f(t)$. Функцію $f(t)$ називають *оригіналом функції* $F(p)$. Зв'язок функцій $f(t)$ та $F(p)$ символічно позначають так:

$$f(t) \risingdotseq F(P) \quad (f(t) \to F(P)) \quad \text{або} \quad F(p) \risingdotseq f(t) \quad (F(p) \to f(t)).$$

Нагадаємо, що серед функцій комплексної змінної найбільш важливий клас – аналітичні функції. З'ясуємо, чи є аналітичною функцією зображення Лапласа $F(p)$.

Теорема: зображення Лапласа $F(p)$ функції $f(t)$ є аналітичною функцією комплексної змінної p в області $Re\ p > a$, де a – показник зростання функції $f(t)$. Доведення не наводимо.

Необхідною ознакою зображення є:

$$\lim_{|p| \to \infty} F(p) = 0, \tag{3.69}$$

Тема 16.2. Зображення Лапласа елементарних функцій

Одинична функція Хевісайда: нехай

$$f(t) = \sigma_0(t) = \begin{cases} 0, t < 0, \\ 1, t \geq 0. \end{cases} \tag{3.70}$$

Тоді

$$f(t) \ \risingdotseq\ F(p) = \int_0^\infty \mathrm{e}^{-pt}\,dt = \frac{1}{p},$$

а функція $F(p)$ визначена в області $Re\ a > 0$. Отже,

$$f(t) = \sigma_0(t) = \begin{cases} 0, t < 0, \\ 1, t \geq 0. \end{cases} \risingdotseq \frac{1}{p},\ \ Re\, p > 0. \tag{3.71}$$

Домовимося надалі під функцією $f(t)$ розуміти добуток $f(\mathrm{t}) \cdot \sigma_0(t)$, тобто функцію, що тотожно дорівнює нулю при $t < 0$.

Показникова функція:

$$f(t)=e^{\alpha t}. \tag{3.72}$$

Обчислимо інтеграл (1) і дістанемо:

$$F(p)=\int_0^\infty e^{-pt}e^{\alpha t}dt=\frac{1}{p-\alpha} \ \risingdotseq\ e^{\alpha t}, \quad Re\, p > Re\, a. \tag{3.73}$$

Степенева функція:

$$f(t)=t^\nu, \qquad \nu>-1. \tag{3.74}$$

У цьому випадку інтеграл (3.63) має вигляд

$$F(p)=\int_0^\infty e^{-pt}t^\nu dt, \quad Re\, p > 0.$$

Почнемо з випадку, коли змінна p набуває дійсних значень $p = x > 0$. Тоді

$$F(x)=\int_0^\infty e^{-xt}t^\nu dt=\frac{1}{x^{\nu+1}}\int_0^\infty e^{-s}s^\nu ds=\frac{\Gamma(\nu+1)}{x^{\nu+1}},$$

де $\Gamma(\nu+1)$ – гамма-функція Ейлера. Унаслідок єдиності аналітичного продовження функція $F(p)$ в області $Re\, p > 0$ має вигляд

$$F(p)=\frac{\Gamma(\nu+1)}{p^{\nu+1}}. \tag{3.75}$$

При цьому у випадку дробових або ірраціональних значень v необхідно обирати ту гілку багатозначної функції $1/p^{\nu+1}$, яка є безпосереднім аналітичним продовженням в область $Re\, p > 0$ дійсної функції $1/x^{\nu+1}$ дійсної змінної $x > 0$.

Таким чином,

$$t^\nu \ \risingdotseq\ \frac{\Gamma(\nu+1)}{p^{\nu+1}}, \qquad \nu>-1, \ Re\, p > 0. \tag{3.76}$$

Зокрема, для цілих $\nu = n$ можна записати:

$$t^\nu \ \risingdotseq\ \frac{\Gamma(n+1)}{p^{n+1}}=\frac{n!}{p^{n+1}}, \ Re\, p > 0. \tag{3.77}$$

Тема 16.3. Властивості зображень Лапласа

Лінійність зображення. Унаслідок властивостей визначених інтегралів справедливе твердження:

якщо $F_i(p) \ \risingdotseq\ f_i(t), \quad Re\, p > a_i$, ($i = 1, ..., n$), то

$$F(p)=\sum_{i=1}^n \alpha_i F_i(p) \ \risingdotseq\ \sum_{i=1}^n \alpha_i f_i(t), \quad Re\, p > max\, a_i \tag{3.78}$$

де α_i – задані сталі, a_i – показник зростання функції $f_i(t)$.

Властивість подібності. Нехай

$$F(p) \ \risingdotseq\ f(t), \quad Re\, p > a,$$

тоді

$$\frac{1}{\alpha}F\left(\frac{p}{\alpha}\right) \ \risingdotseq\ f(\alpha t), \qquad \alpha>0, \quad Re\, p > a, \tag{3.79}$$

$$\int_0^\infty e^{-pt}f(\alpha t)dt=\frac{1}{\alpha}\int_0^\infty e^{-\frac{p}{\alpha}\tau}f(\tau)d\tau=\frac{1}{\alpha}F\left(\frac{p}{\alpha}\right).$$

Теорема запізнювання. Нехай

$$F(p) \risingdotseq f(t), \quad Re\, p > a,$$

та задано функцію

$$f_\tau(t)=\begin{cases} 0, t<\tau, \tau>0, \\ f(t-\tau), t\geq\tau. \end{cases} \qquad (3.80)$$

Тоді:

$$f_\tau(t) \risingdotseq F_\tau(p)=e^{-p\tau}F(p), \quad Re\, p > a, \qquad (3.81)$$

$$F_\tau(p)=\int_0^\infty e^{-pt}f_\tau(t)dt=\int_\tau^\infty e^{-pt}f(t-\tau)dt = \int_0^\infty e^{-p(t'+\tau)}f(t')dt'=e^{-p\tau}F(p).$$

Отже, множення зображення на $e^{-p\tau}$ зміщує графік оригіналу вправо на τ, відповідає явищу, яке у фізиці називають *запізненням*.

Зображення похідної. Якщо $f'(t)$ задовольняє умови існування зображення при $Re\, p > a$ та

$$f(t) \risingdotseq F(p), \quad Re\, p > a,$$

то

$$f'(t) \risingdotseq pF(p)-f(0), \quad Re\, p > a. \qquad (3.82)$$

Це одна з основних властивостей зображення: операції диференціювання в просторі зображень ставиться у відповідність операція множення на змінну p, якщо початкова умова для оригіналу − нульова ($f(0) = 0$).

Зображення похідної вищих порядків. Якщо $f^{(n)}(t)$ задовольняє умови існування зображення при $Re\, p > a$ та

$$f(t) \risingdotseq F(p), \quad Re\, p > a,$$

то

$$f^{(n)}(t) \risingdotseq p^n\left\{F(p)-\frac{f(0)}{p}-\frac{f'(0)}{p^2}-\ldots-\frac{f^{(n-1)}(0)}{p^n}\right\}, \quad Re\, p > a. \qquad (3.83)$$

Формула (3.80) особливо спрощується у випадку, коли

$$f(0)=f'(0)=\ldots=f^{(n-1)}(0)=0, \quad f^{(n)}(t) \risingdotseq p^{(n)}F(p). \qquad (3.84)$$

Зображення інтеграла. Нехай

$$f(t) \risingdotseq F(p), \quad Re\, p > a.$$

Тоді

$$\varphi(t)=\int_0^t f(\tau)d\tau \risingdotseq \frac{1}{p}F(p), \quad Re\, p > a. \qquad (3.85)$$

Зображенням інтеграла $\int_0^t f(\tau)d\tau$ є повторний інтеграл $\int_0^\infty e^{-pt}dt\int_0^t f(\tau)d\tau$, що виражає подвійний інтеграл по області, яка задається нерівностями $0 \leq t \leq T$, $0 \leq \tau \leq t$, де $T \to \infty$ (рис. 3.7). Цю область можна задати по іншому ($0 \leq \tau \leq T$, $\tau \leq t \leq T$, де $T \to \infty$) і змінити порядок інтегрування:

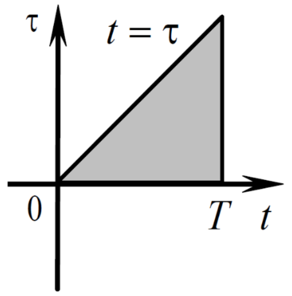


Рис. 3.7

$$\int_0^\infty e^{-pt}dt\int_0^t f(\tau)d\tau=\int_0^t f(\tau)d\tau\int_\tau^\infty e^{-pt}dt=\frac{1}{p}\int_0^\infty e^{-p\tau}f(\tau)d\tau=\frac{1}{p}F(p).$$

Зображення багатократного інтеграла. Нехай

$$f(t)\risingdotseq F(p),\quad Re\,p>a.$$

Тоді

$$\varphi(t)=\int_0^t dt_1\int_0^{t_1}dt_2...\int_0^{t_{n-1}}dt_n f(t_n)\risingdotseq\frac{1}{p^n}F(p),\quad Re\,p>a. \qquad (3.86)$$

Цю властивість доводять аналогічно попередній.

Згорткою $f_1 * f_2$ функцій $f_1(t)$ і $f_2(t)$ називається функція $\varphi(t)$, що визначається співвідношенням:

$$\varphi(t)=\int_0^t f_1(\tau)f_2(t-\tau)d\tau=\int_0^t f_1(t-\tau)f_2(\tau)d\tau. \qquad (3.87)$$

Зображення згортки. Якщо

$$f_1(t)\risingdotseq F_1(p),\quad Re\,p>a_1,\quad f_2(t)\risingdotseq F_2(p),\quad Re\,p>a_2.$$

Тоді

$$\varphi(t)=\int_0^t f_1(\tau)f_2(t-\tau)d\tau\risingdotseq F_1(p)\cdot F_2(p),\quad Re\,p>\max\{a_1,a_2\}. \qquad (3.88)$$

Отже, згортці у просторі оригіналів відповідає добуток у просторі зображень.

Диференціювання зображення. Нехай

$$F(p)\risingdotseq f(t),\quad Re\,p>a,$$

Тоді

$$F'(p)\risingdotseq -tf(t),\quad Re\,p>a. \qquad (3.89)$$

$$F'(p)=\frac{d}{dp}\int_0^\infty e^{-pt}f(t)dt=-\int_0^\infty e^{-pt}tf(t)dt\risingdotseq -tf(t),$$

Багатократне диференціювання зображення. Нехай

$$F(p)\risingdotseq f(t),\quad Re\,p>a.$$

Тоді

$$F^{(n)}(p)\risingdotseq(-1)^{(n)}t^{(n)}f(t),\quad Re\,p>a. \qquad (3.90)$$

Інтегрування зображення. Нехай

$$F(p)\risingdotseq f(t),\quad Re\,p>a,$$

функція $f(t)/t$ задовольняє умови існування зображення. Тоді

$$\frac{f(t)}{t} \risingdotseq \int_0^{\infty} e^{-pt}\frac{f(t)}{t}dt=\int_p^{\infty} F(q)dq\,, \quad Re\, p > a. \tag{3.91}$$

Теорема зміщення зображення. Нехай

$$f(t) \risingdotseq F(p), \quad Re\, p > a.$$

Тоді для довільного комплексного числа λ справедливе співвідношення

$$F(p+\lambda) \risingdotseq e^{-\lambda t}f(t), \quad Re\, p > a - Re\,\lambda, \tag{3.92}$$

$$F(p+\lambda)=\int_0^{\infty} e^{-(p+\lambda)t}f(t)dt=\int_0^{\infty} e^{-pt}e^{-\lambda t}f(t)dt \risingdotseq e^{-\lambda t}f(t).$$

Теорема зміщення застосовується у фізичних явищах, пов'язаних із затухаючими коливаннями.

Гранична теорема 1. Нехай

$$f(t) \risingdotseq F(p), \quad Re\, p > a,$$

функція $f'(t)$ належить до класу оригіналів та існує граничне значення $f(t)$, коли $t \to \infty$. Тоді

$$\lim_{t\to\infty} f(t)=\lim_{p\to 0} pF(p). \tag{3.93}$$

Властивість (3.93) безпосередньо випливає, якщо здійснити граничний перехід у формулі для зображення похідної функції:

$$\lim_{p\to 0}\int_0^{\infty} e^{-pt}f'(t)dt=\lim_{p\to 0}(pF(p)-f(0)),$$

звідки

$$\lim_{t\to\infty} f(t)-f(0)=\lim_{p\to 0} pF(p)-f(0).$$

Гранична теорема 2. Нехай

$$f(t) \risingdotseq F(p), \quad Re\, p > a,$$

і функція $f'(t)$ належить до класу оригіналів. Тоді

$$\lim_{t\to +0} f(t)=\lim_{p\to\infty} pF(p). \tag{3.94}$$

Властивість (3.94) безпосередньо випливає з необхідної умови існування зображення похідної функції:

$$\lim_{p\to\infty}(pF(p)-f(0))=0\,,$$

звідки

$$\lim_{p\to\infty} pF(p)=f(0).$$

Диференціювання за параметром. Нехай

$$f(t,\lambda) \risingdotseq F(p,\lambda), \quad Re\, p > a,$$

і при $\lambda_1 < \lambda < \lambda_2$ та $t > 0$ існує $df/d\lambda$. Тоді для всіх λ із вказаного інтервалу справедливе співвідношення:

$$\frac{df(t,\lambda)}{d\lambda} \risingdotseq \frac{dF(p,\lambda)}{d\lambda} \tag{3.95}$$

Ця властивість використовується при розв'язуванні диференціальних рівнянь у частинних похідних.

Наведені вище властивості допомагають знайти зображення багатьох функцій, деякі з яких наведені нижче ($t > 0$).

1) $1 \fallingdotseq 1/p$, $\operatorname{Re} p > 0$;

2) $t^{\nu} \fallingdotseq \dfrac{\Gamma(\nu+1)}{p^{\nu+1}}$, $\nu > -1$, $\operatorname{Re} p > 0$;

3) $t^{n} \fallingdotseq \dfrac{n!}{p^{n+1}}$, n – натуральне, $\operatorname{Re} p > 0$;

4) $e^{\alpha t} \fallingdotseq \dfrac{1}{p-\alpha}$, $\operatorname{Re} p > \operatorname{Re}\alpha$;

5) $\sin\omega t \fallingdotseq \dfrac{\omega}{p^2+\omega^2}$, $\operatorname{Re} p > |\operatorname{Im}\omega|$;

6) $\cos\omega t \fallingdotseq \dfrac{p}{p^2+\omega^2}$, $\operatorname{Re} p > |\operatorname{Im}\omega|$;

7) $\operatorname{sh}\lambda t \fallingdotseq \dfrac{\lambda}{p^2-\lambda^2}$, $\operatorname{Re} p > |\operatorname{Re}\lambda|$;

8) $\operatorname{ch}\lambda t \fallingdotseq \dfrac{p}{p^2-\lambda^2}$, $\operatorname{Re} p > |\operatorname{Re}\lambda|$;

9) $t^{n}e^{\alpha t} \fallingdotseq \dfrac{n!}{(p-\alpha)^{n+1}}$, $\operatorname{Re} p > \operatorname{Re}\alpha$;

10) $t\sin\omega t \fallingdotseq \dfrac{2p\omega}{(p^2+\omega^2)^2}$, $\operatorname{Re} p > |\operatorname{Im}\omega|$;

11) $t\cos\omega t \fallingdotseq \dfrac{p^2-\omega^2}{(p^2+\omega^2)^2}$, $\operatorname{Re} p > |\operatorname{Im}\omega|$;

12) $e^{\lambda t}\sin\omega t \fallingdotseq \dfrac{\omega}{(p-\lambda)^2+\omega^2}$, $\operatorname{Re} p > \operatorname{Re}\lambda + |\operatorname{Im}\omega|$

13) $e^{\lambda t}\cos\omega t \fallingdotseq \dfrac{p-\lambda}{(p-\lambda)^2+\omega^2}$, $\operatorname{Re} p > \operatorname{Re}\lambda + |\operatorname{Im}\omega|$;

14) $\dfrac{\sin\omega t}{t} \fallingdotseq \dfrac{\pi}{2} - \operatorname{arctg}\dfrac{p}{\omega}$, $\operatorname{Re} p > |\operatorname{Im}\omega|$;

15) $|\sin\omega t| \fallingdotseq \dfrac{\omega}{p^2+\omega^2}\operatorname{cth}\dfrac{p\pi}{2\omega}$, $\operatorname{Re} p > |\operatorname{Im}\omega|$;

16) $\operatorname{si} t \fallingdotseq \dfrac{1}{p}\left(\dfrac{\pi}{2} - \operatorname{arctg} p\right)$, $\operatorname{Re} p > 0$.

Тема 16.4. Обернене перетворення Лапласа. Формула Мелліна

Нехай відомо, що функція комплексної змінної $F(p)$ є зображенням функції $f(t)$ з обмеженим показником зростання $|f(t)| < Me^{at}$, і значення сталої а задано. Необхідно за даною функцією $F(p)$ відновити оригінал $f(t)$. Розв'язати цю задачу допоможе наступна теорема.

Теорема (*формула Мелліна*): Нехай відомо, що функція $F(p)$ в області $Re\, p > a$ є зображенням кусково-гладкої функції $f(t)$ дійсної змінної t, що має показник зростання a. Тоді

$$f(t) = \frac{1}{2\pi i}\int_{x-i\infty}^{x+i\infty} \mathrm{e}^{pt}F(p)\,dp, x > a. \qquad (3.96)$$

Розглянемо допоміжну функцію $\varphi(t) = e^{-xt} f(t)$, $x > a$. Ця функція є кусково-гладкою, тобто на довільній обмеженій ділянці осі t має скінченну кількість точок розриву першого роду й експоненціально прямує до нуля при $t \to \infty$.

Її можна зобразити за допомогою *інтеграла Фур'є*:

$$\varphi(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\xi \int_{-\infty}^{\infty} \varphi(\eta)\mathrm{e}^{i\xi(t-\eta)}d\eta. \qquad (3.97)$$

Використаємо тепер визначення функції φ(t)

$$\mathrm{e}^{-xt}f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\xi \int_{-\infty}^{\infty} \mathrm{e}^{-x\eta}f(\eta)\mathrm{e}^{i\xi(t-\eta)}d\eta \;=$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{i\xi t}d\xi \int_{0}^{\infty} \mathrm{e}^{-(x+i\xi)\eta}f(\eta)\,d\eta. \qquad (3.98)$$

Оскільки $f(\eta) = 0$ при $\eta < 0$, множимо обидві частини (3.98) на e^{xt} і, здійснивши перехід до нової змінної $p = x + i\xi$ у результаті одержимо:

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{(x+i\xi)t}d\xi\int_{0}^{\infty} e^{-(x+i\xi)\eta}f(\eta)d\eta=\frac{1}{2\pi i}\int_{x-i\infty}^{x+i\infty} e^{pt}F(p)dp. \qquad (3.99)$$

Зазначимо, що інтегрування у формулі Мелліна здійснюється на комплексній площині по прямій, яка паралельна уявній осі та проходить справа від прямої $Re\ p = a$. Далі покажемо, що значення інтеграла (3.96) не залежить від x за умови, що пряма інтегрування лежить справа від прямої $Re\ p = a$.

Формула Мелліна в певному змісті є оберненим перетворенням Лапласа. Оскільки в процесі виведення її від невідомої функції $f(t)$ ми прийшли до інтеграла Фур'є, який збігається до $f(t)$ лише в точках неперервності цієї функції, то й (3.96) визначає $f(t)$ лише в точках її неперервності.

Розглянемо застосування формули Мелліна (3.96) для знаходження зображення добутку, якщо відомі зображення співмножників.

Теорема: нехай

$$f_1(t)\ \doteqdot\ F_1(p),\quad Re\,p > a_1,\quad f_2(t)\ \doteqdot\ F_2(p),\quad Re\,p > a_2.$$

Тоді

$$f(t)=f_1(t)\cdot f_2(t)\ \doteqdot\ F(p)=\frac{1}{2\pi i}\int_{x-i\infty}^{x+i\infty} F_1(q)F_2(p-q)dq\ =$$

$$=\ \frac{1}{2\pi i}\int_{x-i\infty}^{x+i\infty} F_1(p-q)F_2(q)dq,$$

причому функція $F(p)$ визначена та аналітична всюди в області $Re\ p > a_1 + a_2$, а інтегрування проводиться по довільній прямій, що паралельна до уявної осі й задовольняє у першому випадку умову

$$a_1 < Re\ q < Re\ p - a_2,$$

а в другому – умову

$$a_2 < Re\ q < Re\ p - a_1.$$

Оскільки функція f(t) задовольняє всі умови існування зображення, то для неї існує перетворення Лапласа

$$f(t)\ \doteqdot\ F(p)=\int_0^{\infty} e^{-pt}f_1(t)f_2(t)dt=\frac{1}{2\pi i}\int_0^{\infty} e^{-pt}f_2(t)dt\int_{x-i\infty}^{x+i\infty} e^{qt}F_1(q)dq\ =$$

$$=\ \frac{1}{2\pi i}\int_{x-i\infty}^{x+i\infty} F_1(q)dq\int_0^{\infty} e^{-(p-q)t}f_2(t)dt\ =\ \frac{1}{2\pi i}\int_{x-i\infty}^{x+i\infty} F_1(q)F_2(p-q)dq.$$

Зазначимо, що $Re\ q = x > a_1$, а функція $F_2(p - q)$ визначена для $Re\ (p - q) > a_2$, звідки $Re\ p > a_1 + a_2$. Крім того, змінювати порядок інтегрування можна в наслідок рівномірної збіжності даних невласних інтегралів, залежних від параметрів. *Ця теорема у деякому розумінні обернена до властивості* 8.

Розглянемо тепер достатні умови, при виконанні яких задана функція $F(p)$ комплексної змінної p є зображенням функції $f(t)$ дійсної змінної t, та покажемо, як знайти останню.

Теорема: нехай функція $F(p)$ комплексної змінної p задовольняє наступні умови:
1. $F(p)$ – аналітична в області $Re\, p > a$.
2. В області $Re\, p > a$ функція $F(p)$ прямує до 0 при $|p| \to \infty$ рівномірно відносно $arg\, p$.
3. Для всіх $Re\, p = x > a$ збігається інтеграл:

$$\int_{-\infty}^{\infty} |F(p)|\, dy < M, \quad p = x + iy, \quad x > a.$$

Тоді функція $F(p)$ при $Re\, p > a$ є зображенням функції $f(t)$ дійсної змінної t, яка визначається виразом

$$f(t) = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} e^{pt} F(p)\, dp, x > a. \qquad (3.100)$$

Тема 16.5. Розв'язання диференціальних та інтегральних рівнянь за допомогою перетворення Лапласа

Розглянемо лінійне диференціальне рівняння зі сталими коефіцієнтами для функції $f(t)$:

$$\hat{L} f(t) = \varphi(t),$$

де $\hat{L}$ – диференціальний оператор, що містить лінійну комбінацію похідних різного порядку по t. Наприклад,

$$\frac{d^2 f}{dt^2} + \gamma \frac{df}{dt} + \omega_0^2 f = \varphi(t), \quad f(0) = 0, \quad f'(0) = 0.$$

У просторі зображень такому диференціальному рівнянню для $f(t)$ відповідає алгебраїчне рівняння для F(p),

$$F(p) \risingdotseq f(t),$$

розв'язавши яке, можна знову повернутися до функції f(t), здійснивши обернене перетворення Лапласа.

Розглянемо лінійне диференціальне рівняння n-го порядку загального вигляду зі сталими коефіцієнтами:

$$L[x] = a_0 \frac{d^n x}{dt^n} + a_1 \frac{d^{n-1} x}{dt^{n-1}} + \ldots + a_{n-1} \frac{dx}{dt} + a_n x = f(t), \qquad (3.101)$$

доповнене початковими умовами:

$$x(0) = x_0, \quad x'(0) = x_1, \quad \ldots \quad x^{(n-1)}(0) = x_{n-1}.$$

Нехай функція $x(t)$, усі її похідні та $f(t)$ задовольняють умови існування зображення,

$$X(p) \risingdotseq x(t), \quad F(p) \risingdotseq f(t),$$

Тоді домножимо (3.101) на e^{-pt} і проінтегруємо за t в межах від 0 до ∞. У результаті отримаємо алгебраїчне рівняння:

$$(a_0 p^n + a_1 p^{n-1} + \ldots + a_{n-1} p + a_n) X(p) \; =$$

$$= F(p)+x_0(a_0 p^{n-1}+a_1 p^{n-2}+\ldots+a_{n-1}) +$$

$$+ x_1(a_0 p^{n-2}+a_1 p^{n-3}+\ldots+a_{n-2})+\ldots+x_{n-1}a_0,$$

з якого знаходимо X(p). За відомого зображення $X(p)$ можна здійснити обернене перетворення Лапласа і в такий спосіб знайти розв'язок (3.101), що задовольняє задані початкові умови.

Метод Лапласа є ефективним методом розв'язку лінійних інтегральних рівнянь Вольтерра типу згортки.

Задача 3.10. Знайти розв'язок інтегрального рівняння:

$$1-\cos(t) \risingdotseq \int_0^t sh(t-\tau)f(\tau)d\tau .$$

Розв'язок

Перейдемо до зображень за формулами

$$1-\cos(t) \risingdotseq \frac{1}{p(p^2+1)}, \quad sh(t)*f(t) \risingdotseq \frac{1}{p^2-1}F(p),$$

у просторі зображень одержимо алгебраїчне рівняння

$$\frac{1}{p(p^2+1)}=\frac{1}{p^2-1}F(p),$$

з якого знаходимо

$$F(p)=\frac{p^2-1}{p(p^2+1)}=\frac{2p}{p^2+1}-\frac{1}{p},$$

і відповідно відновлюємо оригінал:

$$f(t)=2\cos(t)-1.$$

Контрольні запитання до розділу 3

1. Що таке «комплексні числа»? Які дії над ними ви знаєте? Опишіть їх.
2. Що таке «функція комплексної змінної»? Яка функція від комплексної змінної називається неперервною в точці, а яка – неперервною на множині?
3. Що таке «похідна від функції комплексної змінної»? Розкрийте зміст умов Коші-Рімана.
4. Що називається інтегралом від функції комплексної змінної по дузі (контуру)? Розкрийте зміст теореми Коші.
5. Що називається невизначеним інтегралом від функції комплексної змінної? Запишіть формулу Ньютона-Лейбніца.
6. Запишіть інтегральну формулу Коші. В чому полягає зміст нерівності Коші?
7. Розкрийте зміст теореми Ліувілля та теореми Морера.

8. Що таке «числові ряди»? Які ознаки збіжності числових рядів ви знаєте? Опишіть їх.
9. Що таке «функціональні ряди»? Розкрийте зміст ознаки рівномірної збіжності (рядів) Веєрштрасса.
10. Розкрийте зміст теореми Веєрштрасса.
11. Що таке «степеневі ряди»? Розкрийте зміст теорема Коші-Адамара.
12. Розкрийте зміст теореми Абеля.
13. Що таке «ряд Тейлора»? Наведіть приклади розкладу в ряд Тейлора елементарних функцій.
14. Розкрийте зміст теореми про єдиність аналітичної функції. Що таке «принцип максимуму»?
15. Що таке «ряд Лорана»? Розкрийте зміст теорема Лорана про існування розкладу в ряд функції, аналітичної в колі.
16. Що таке «ізольовані особливі точки аналітичної функції»? Наведіть їх класифікацію.
17. Що таке «лишки»? Розкрийте зміст основної теореми про лишки.
18. Які класи інтегралів зручно розраховувати допомогою лишків? Наведіть приклади.
19. Що таке «логарифмічний лишок»? Розкрийте зміст принципу аргументу.
20. Що таке «перетворення Лапласа»? Що називається зображенням Лапласа, а що – оригіналом?
21. Наведіть приклади зображень Лапласа елементарних функцій.
22. Які ви знаєте властивості зображень Лапласа? Опишіть їх.
23. Запишіть формулу Мелліна. Розкрийте зміст оберненого перетворення Лапласа.
24. Наведіть приклад розв'язку диференціальних рівнянь за допомогою перетворення Лапласа.
25. Наведіть приклад розв'язку інтегральних рівнянь за допомогою перетворення Лапласа.

РОЗДІЛ 4. ЕЛЕМЕНТИ ДИСКРЕТНОЇ МАТЕМАТИКИ

Дискретна математика – галузь математики, що вивчає властивості будь-яких дискретних структур. Як синонім іноді вживається термін «дискретний аналіз», що вивчає властивості структур скінченного характеру.

Дискретна математика по суті – динамічний розділ вищої математики. Основною сферою її практичного застосовування є комп'ютерні технології. Це, насамперед, пов'язано з необхідністю проєктування, розробки та експлуатації електронних обчислювальних машин, засобів збору, передачі та обробки інформації, систем автоматизованого управління та проектування.

На перетині дискретної математики з програмуванням з'являються нові дисципліни, такі як розробка та аналіз обчислювальних та комбінаторних алгоритмів, алгоритмізація процесів автоматизації тощо. Основу дискретної математики складають такі галузі математичних знань, як теорія множин, математична логіка (булева алгебра) та теорія графів.

Лекція 17. Множини. Відображення множин

Тема 17.1. Поняття множини

Поняття множини належить до фундаментальних понять не тільки математики, а й будь-якого абстрактного мислення. Чітких визначень поняття «множина», «елемент множини», «приналежність до множини», і так далі не існує. Однак вони і не потрібні, оскільки кожен, це розуміє апріорно.

Наприклад, можна говорити про такі множини, як: множина вищих навчальних закладів країни; студентів навчального закладу; підприємств міста; гравців та вболівальників на стадіоні; літер українського алфавіту; слів у словнику і так далі. У математиці постійно доводиться працювати з множинами різної природи, наприклад: множиною дійсних чисел, множиною прямокутних трикутників, множиною векторів у тривимірному просторі, тощо.

Водночас інтуїтивне розуміння множини та способів її задання призводить до відомих парадоксів і логічних проблем типу: «множена всіх множин», і так далі. Для розв'язання цих проблем використовують конструкції, які базуються на введенні так званої аксіоми вибору або еквівалентному їй певному твердженні. В цьому курсі ми будемо мати справу з кінцевими або рахунковими множинами, для яких такого роду проблеми не виникають.

Отже, *множина* – це сукупність певних об'єктів, що називаються елементами множини, які можна добре відрізнити один від одного.

Множина, яка не містить жодного елемента, називається *порожньою множиною* та позначається через символ ø або $\{\}$.

Традиційно множини позначають через великими літерами латинського алфавіту: A, B, C, ..., а їх елементи – маленькими літерами (a, b, c, ...). Якщо елемент множини a належить множині A, то пишуть:

$$a \in A, \text{ в протилежному випадку } a \notin A,$$

Кажуть, що дві множини A і B *рівні*:

$$A = B, \tag{4.1}$$

якщо обидві складаються з однакових елементів. У цьому визначенні є один нюанс. Чи рівні наступні множини

$$A = \{1,2,3,4,5\}, \quad B = \{5,4,3,2,1\}?$$

Так, рівні, оскільки множини можуть бути і неупорядкованими. А чи рівні множини

F_1 = {яблуко, груша, апельсин}, F_2 = {яблуко, яблуко, груша, апельсин}?

Якщо, як елементи ми розглядаємо одні і ті самі яблука, груші, апельсини, то так, ці множини рівні.

Кажуть, що множина A є *підмножиною* множини B, якщо кожен елемент множини A є елементом множини B. Для цього використовують таке стандартне позначення:

$$A \subset B \tag{4.2}$$

Проте, це відношення не строге, якщо $A = B$, то $A \subset B$ і відповідно $B \subset A$. Тобто, справедливе також таке позначення:

$$A \subseteq B, \quad \text{(підмножина або дорівнює).} \tag{4.3}$$

Тема 17.2. Потужність множини

Зафіксуємо деякі позначення для множин, з якими ми постійно маємо справу. По-перше, це множина натуральних чисел:

$$\mathbb{N} = \{0,1,2,\ldots\}.$$

Будемо вважати, що нуль є натуральним числом. У цьому питанні багато суперечок і кожен по – своєму правий. Цей варіант зручний тим, що якщо буде потрібним ряд із додатних цілих чисел, то естетичніше виключити нуль, ніж його додавати в іншому випадку.

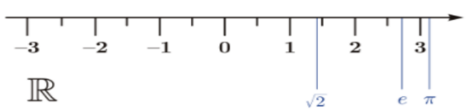
По-друге, множина цілих чисел:

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

По-третє, множина раціональних чисел – $\mathbb{Q}$.

$$\mathbb{Q} = \left\{ \frac{m}{n}, n \in \mathbb{N}, m \in \mathbb{Z} \right\}.$$

По-четверте, найголовніша множина – це множина дійсних чисел $\mathbb{R}$.



І, нарешті, множина комплексних чисел $\mathbb{C}$.

$$z = x + iy.$$

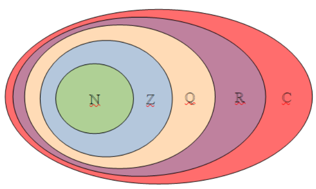


Рис. 4.1

Розглянуті вище множини є множинами чисел. А що таке «число»? По суті число виражає деяку кількісну характеристику, необхідну для розрахунку.

У цьому розумінні тільки комплексні числа вибиваються з ряду. Їх поява зумовлена операцією вилучення квадратного кореня.

Проте для нас важливим буде визначення чисел як *алгебраїчного поля*. У загальній алгебрі *полем* називається множина M, яка є комутативною групою щодо операції додавання «+» з нульовим елементом 0, а також з операцією множення «*» над ненульовими елементами цієї множини. При цьому ці операції повинні задовольняти властивість дистрибутивності множення щодо додавання.

Множини $\mathbb{Q}$, $\mathbb{R}$ і $\mathbb{C}$ є полями зі звичайними операціями додавання і множення. Часто вживаються такі стійкі вираження, як «поле комплексних чисел» над «полем дійсних чисел».

Тепер ми можемо розглянути таке поняття в теорії множин як *потужність множини*.

У побуті ми часто зустрічаємося з такими множинами як: множина гостей, множина кімнат у квартирі, множина цукерок в тарілці, і так далі. Нас перш за все цікавить кількість елементів у цій множині. Тобто *для кінцевих множин поняття потужності збігається з кількістю елементів*.

Використовують різні позначення для потужності множин, зокрема:

$$|A|, \#A \text{ та інші.} \tag{4.4}$$

Потужності множин поділяються на два типи.

Множини $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ називаються *зліченними*, тобто кожному елементу цих множин можна присвоювати певний номер.

Множини $\mathbb{R}$, $\mathbb{C}$ незліченні, тобто їхні елементи не можна перерахувати, відповідно потужність цих множин називається *континуум*.

З кінцевими множинами більш-менш все зрозуміло. Звернемо увагу на нескінченні множини.

Згідно з означенням, множина натуральній чисел $\mathbb{N}$ нескінченна. Тоді множина A називається *нескінченною*, якщо існує така підмножина $B \subset A$, що між елементами множин $\mathbb{N}$ і B існує взаємно однозначна відповідність. Можна дати ще одне еквівалентне визначення.

Множина називається *нескінченною*, якщо в неї існує власна підмножина (тобто, яка не збігається із самою множиною), для елементів якої можна встановити взаємно однозначну відповідність.

Розглянемо власну підмножину $\mathbb{N}_1 = \mathbb{N} \setminus \{0\}$. Тоді будь-якому елементові $a \in \mathbb{N}_1$ відповідає єдиний елемент $b = a - 1$, який належить множині $\mathbb{N}$. Потужність множини $\mathbb{N}$ позначається $\aleph_0$ (алеф).

Отже, множини, які мають таку саму потужність, що й множина натуральних чисел ($\aleph_0$), тобто елементам якої можна присвоїти номери, (інакше кажучи перенумерувати), називаються *зліченними*.

Потужність $\aleph_0$ є найменшою потужністю нескінченних множин. Добре відомо, що $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$. Нескінченною множиною, яка має велику потужність,

є множина $\mathbb{R}$. Потужність цієї множини називається *потужністю континууму* і позначається через $\aleph_1$ або c.

Теорема Георга Кантора: множина дійсних чисел $\mathbb{R}$, що міститься у відрізку [0,1], незлічення.

Добре відомо, що $\aleph_0 < \aleph_1$. Виникає питання про існування множини, що має потужність a, таку що

$$\aleph_0 < a < \aleph_1. \tag{4.5}$$

Континуум-гіпотеза була сформульована Георгом Кантором в 1877 році й стверджує, що таких множин не існує. У 1940 році К. Гедель довів, що *заперечення* континуум-гіпотези недоказове при аксіомі вибору.

Тема 17.3. Операції над множинами

Для наочної ілюстрації операцій над множинами використовуються так звані *круги Ейлера* або *діаграми Венна*. Множини зображуються на площині у вигляді кругів або інших плоских фігур, незалежно від їхньої потужності. Результат виконання операцій позначається відповідно кольором або заштриховкою.

Об'єднанням двох множин A та *B* називається множина *C*, яка складається тільки з елементів множини *A* та елементів множини *B*, взятих по одному разу. Позначення: $C = A \cup B$ (рис. 4.2).

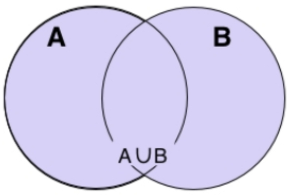


Рис. 4.2

Перетином двох множин A та *B* називається множина *C*, елементами якої є ті і тільки ті елементи, що належать одночасно множинам *A* і *B*. Позначення: $C = A \cap B$ (рис. 4.3).

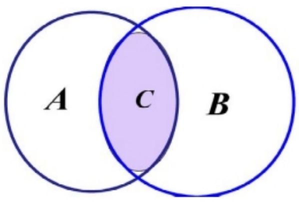


Рис. 4.3

Різницею двох множин A та *B* називається множина *C*, елементами якої є ті і тільки ті елементи множини *A*, що не є елементами множини *B*. Позначення: $C = A \setminus B$ (рис. 4.4).

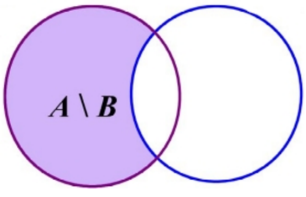


Рис. 4.4

Симетричною різницею двох множин A та *B* називається множина *C*, елементами якої є ті і тільки ті елементи множин *A* та *B*, що не належать цим множинам одночасно. Позначення $C = A \Delta B$ (рис. 4.5).

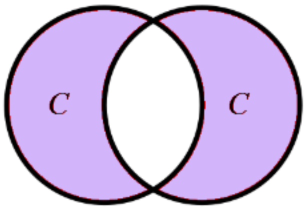


Рис. 4.5

За означення симетричної різниці також приймають очевидну рівність:

$$A \Delta B = (A \setminus B) \cup (B \setminus A). \tag{4.6}$$

Якщо $A \subseteq B$, різниця $B \setminus A$ називається *доповненням* множини *B* у множині *A* і позначається $\overline{B_A}$. Як правило, доповнення розглядається в універсумі, тобто $\overline{A} = U \setminus A$ (рис. 4.6).

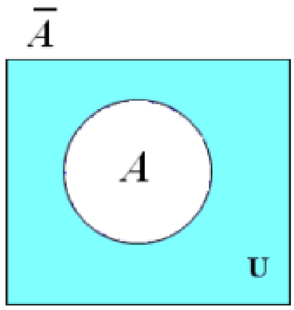


Рис. 4.6

Універсальна множина (*універсум*) – в теорії множин така множина *U*, для якої перетин цієї множини з будь-якою множиною *X* збігається з цією множиною *X*. Універсальна множина єдина.

Тема 17.4. Відображення множин

Ще одне поняття, яке розуміється інтуїтивно, стосується відображення множин.

Нехай задано дві множини *X* і *Y*. Ми будемо говорити, що *A* є *відображенням* з *X* в *Y*, якщо задано правило, яке кожному елементу $x \in X$ ставить у однозначну відповідність елемент $y \in Y$. При цьому пишуть:

$$A : X \to Y. \tag{4.7}$$

Ми будемо завжди говорити про однозначні відображення. У випадках, коли необхідно розглядати багатозначні відображення, ми будемо говорити про однозначні відображення в підмножини.

Замість терміна «відображення» будемо також використовувати і інші синоніми, наприклад *функція*. Якщо множини *X* і *Y* є просторами, то часто вживають термін *оператор*. Якщо множина *Y* являє собою множину чисел (дійсних або комплексних), то говорять про *функціонал*.

Іноді відображення може бути визначено не на всьому *X*, а лише на деякій підмножині $X_1 \subset X$, ця підмножина називається *областю визначення відображення A* і позначається $D(A)$. Підмножина $Y_1 \subset Y$, яка складається з $y \in$

Y, але таких, для яких існує $x \in D(A)$, $y = A[x]$, називається *областю значення відображення Y*.

Для будь-якого $x \in D(A)$ елемент $y = A[x]$ називається *образом x*. А для фіксованого $y \in Y$ множини таких $x \in D(A)$, що $A[x] = y$ називається *прообразом y*. Поняття образу та прообразу легко узагальнюються на поняття образу і прообразу для множин.

Якщо областю значень відображення є всі Y, то такі відображення називається *сюр'єктивним*.

Якщо для двох різних x_1, $x_2 \in X$ їх образи різні, то таке відображення називається *ін'єктивним*.

Відображення, яке є одночасно сюр'єктивним і ін'єктивним називаються *бієктивним* (*бієкцією*). Бієктивне відображення, або бієкція, – це взаємно однозначне відображення (рис. 4.7 а).

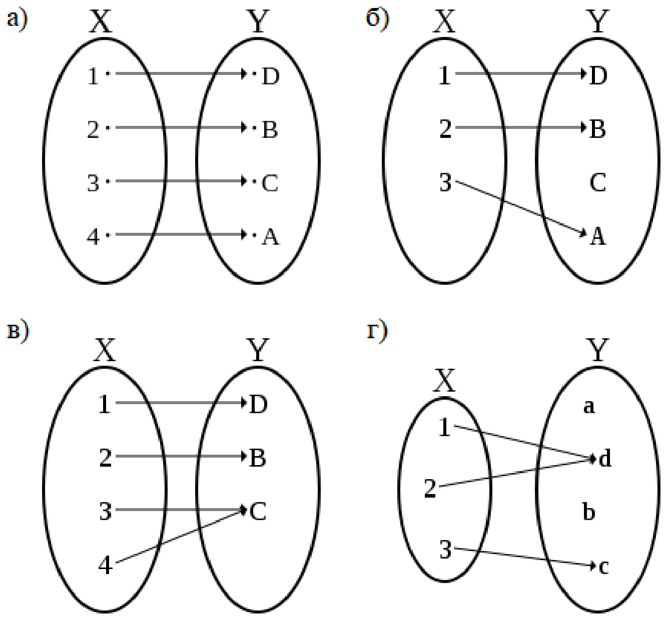


Рис. 4.7. Типи відображень множин: *а* – бієктивне відображення;
б – ін'єктивне, але не сюр'єктивне відображення;
в – сюр'єктивне, але не ін'єктивне відображення;
г – несюр'єктивне і неін'єктивне відображення

Уведемо поняття прямого добутку двох множин X і Y.

Прямим добутком, або *декартовим добутком*, двох множин $X \times Y$, називається множина, що складається з усіх упорядкованих пар (x, y), де $x \in X$, а $y \in Y$ (рис. 4.8).

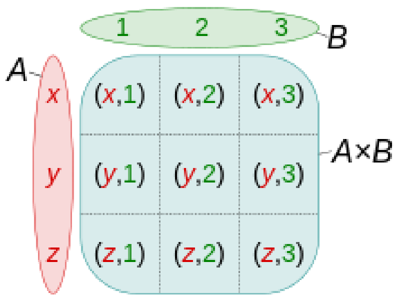


Рис. 4.8. Декартів добуток $A \times B$ множин $A = \{x, y, z\}$ та $B = \{1, 2, 3\}$

Наприклад, декартовим добутком $\mathbb{R} \times \mathbb{R}$ є множина точок – $\mathbb{R}^2$. Аналогічно можна ввести поняття і декартового добутку n множин.

Теорема: нехай скінченні множини A_i ($i = 1, ..., n$), потужності яких дорівнюють m_i відповідно. Тоді:

$$\left| \underset{i=1}{\overset{n}{\times}} A_i \right| = m_1 \cdot m_2 \cdot \ldots \cdot m_n = \prod_{i=1}^{n} m_i .$$

Звідси випливає, що

$$|A^n| = |A|^n.$$

Розглянемо деяку непорожню множину X. Для деяких упорядкованих пар із цієї множини введемо поняття *відношення*. Кажуть, що x_1 *знаходиться у відношенні* зі x_2 та пишуть:

$$x_1 \phi\, x_2. \tag{4.8}$$

Відношення мають такі властивості:

1. Відношення називається *рефлексивним*, якщо $x \varphi x$ для всіх $x \in X$.
2. Відношення називається *симетричним*, якщо з $x_1 \varphi x_2$ слідує, що $x_2 \varphi x_1$.
3. Відношення називається *транзитивним*, якщо з $x_1 \varphi x_2$ і $x_2 \varphi x_3$ випливає, що $x_1 \varphi x_3$.

Приклади рефлексивних відношень: $=, \leq, \geq$ та $\subseteq$. Приклали відношень, що не є рефлексивними : $\neq, <, >$ та $\subset$.

Приклади симетричних відношень:

1. Рівність множини цілих дійсних чисел.
2. Подібність трикутників.
3. Рівність по модулю n.
4. Властивості зворотного відношення.

Приклади транзитивних відношень:

1. Строга нерівність: $(a<b),(b<c) \Rightarrow (a<c)$.
2. Нестрога нерівність: $(a \leq b),(b \leq c) \Rightarrow (a \leq c)$.
3. Строга підмножина: $(A \subset B; \wedge B \subset C) \Rightarrow (A \subset C)$.
4. Нестрога підмножина: $(A \subseteq B; \wedge B \subseteq C) \Rightarrow (A \subseteq C)$.
5. Подільність: $(a|b),(b|c) \Rightarrow (a|c)$.
6. Рівність: $(a=b),(b=c) \Rightarrow (a=c)$.
7. Еквівалентність: $(a \Leftrightarrow b),(b=c) \Rightarrow (a=c)$.
8. Імплікація: $(a \Rightarrow b),(b \Rightarrow c) \Rightarrow (a \Rightarrow c)$.
9. Паралельність: $(a||b),(b||c) \Rightarrow (a||c)$.
10. Відношення подібності геометричних фігур.

Приклади нетранзитивних відношень

1. Харчовий ланцюжок: це відношення не завжди транзитивне (задача: вовки їдять оленів, олені їдять траву, але вовки не їдять траву).
2. Бути переважніше ніж. Якщо ми хочемо яблуко замість апельсина, а замість яблука ми б хотіли кавун, то це не означає, що ми віддамо перевагу кавуну.
3. Бути другом.

4. Бути колегою по роботі.
5. Бути підлеглим. Наприклад, у часи феодального ладу в Західній Європі була в ходу приказка: «васал мого васала – не мій васал».
6. Бути схожим на іншу людину.

Відношення, що володіє всіма трьома властивостями: рефлективністю, симетричністю та транзитивністю, називається *відношенням еквівалентності*. Якщо на множині X задано відношення еквівалентності, то цю множину можна розбити на класи еквівалентності. Будемо говорити, що підмножина $K \subset X$ є *класом еквівалентності* для відношення еквівалентності, якщо для всіх x_1, $x_2 \in K$ має місце $x_1\ \varphi\ x_2$ і в $X \setminus K$ не існує елементів, що знаходяться у відношеннях із елементами з K.

Для будь-якого елементу $x \in X$ клас еквівалентності визначається так:

$$K_x = \{x' \in X : x_1 \phi\, x'\}. \tag{4.9}$$

Для будь-яких x і y, які не перебувають у відношенні еквівалентності, відповідні класи еквівалентності не перетинаються. Тому всю множина X можна розбити на класи еквівалентності:

$$X = \bigcup_{\alpha} K_\alpha.$$

При цьому кожен клас еквівалентності можна ототожнити з деяким новим елементом і розглядати вихідну множину як множину класів еквівалентності.

Лекція 18. Елементи математичної логіки

Тема 18.1. Булева алгебра

Логіка (від грец. *logos* – слово, думка, мова, розум) – це сукупність наук про закони та форми мислення. Можна казати про традиційну логіку, яка є початковою сходинкою логіки та вивчає загальні форми висновків та способи зв'язку висловлювань у висновках (Аристотель IV століття до н.е.).

Криза традиційної логіки підштовхнула дослідників до розроблення інших підходів формалізації висновків (Джордж Буль, Огюст де Морган – середина XIX ст.) Внаслідок цього виникла математична логіка (теоретична логіка, символічна логіка) та закладено основи алгебри логік – булевої алгебри.

Алгеброю називається сукупність множини M, що є носієм алгебри, зі заданою на ній сукупністю операцій $S = \{f_{11}, \dots f_{1n}, f_{21}, \dots f_{2m}, \dots, f_{k1}, \dots f_{kp}\}$, яка називається сигнатурою алгебри і позначається: $A = \langle M, S \rangle$.

Алгебраїчна система, або *алгебраїчна структура*, – множина символів деякого алфавіту (носій) з заданим на ньому набором операцій і відношень (сигнатура) цих символів, який задовольняє деяку *систему аксіом*. Результат цих операцій також має належати множині.

Множина – вироджена алгебраїчна система з порожнім набором операцій та відношень. *Групоїд* – множина з однієї бінарної операцією (зазвичай множення). *Кільце* – структура з двома бінарними операціями.

Булева алгебра – це алгебраїчна структура, з двома бінарними операціями: $\wedge$ (булеве множення) – узагальнення кон'юнкції, $\vee$ (булеве додавання) – узагальнення диз'юнкції, та унарною операцією: $\neg$ або $\overline{x}$ (булеве доповнення) – узагальнення заперечення; які задовольняють такі аксіоми:

1. $x_1 \vee x_2 = x_2 \vee x_1$, $x_1 \wedge x_2 = x_2 \wedge x_1$ (комутативність).
2. $x_1 \vee (x_2 \vee x_3) = (x_1 \vee x_2) \vee x_3$, $x_1 \wedge (x_2 \wedge x_3) = (x_1 \wedge x_2) \wedge x_3$ (асоціативність).
3. $x_1 \vee (x_1 \wedge x_2) = x_1$, $x_1 \wedge (x_1 \vee x_2) = x_1$ (закон поглинання).
4. $x_1 \vee (x_2 \wedge x_3) = (x_1 \vee x_2) \wedge (x_1 \vee x_3)$, $x_1 \wedge (x_2 \vee x_3) = (x_1 \wedge x_2) \vee (x_1 \wedge x_3)$ (дистрибутивність).
5. $(x_1 \vee \overline{x_1}) \wedge x_2 = x_2$, $(x_1 \wedge \overline{x_1}) \vee x_2 = x_2$ (доповнення).

Аксіоми 1,2,3 визначають *ґратку*. Аксіоми 1,2,3,4 визначають *дистрибутивну ґратку*. Аксіоми 1,2,3,5 визначають *доповнену ґратку*.

Отже, *булева алгебра* – це алгебраїчна структура, що є доповненою дистрибутивною ґраткою, та частина математики, яка вивчає подібні структури.

Логічними (*булевими*) *змінними* називають змінні, які набувають значення 0 (хибне) або 1 (істина). *Логічною* називають функцію $f(x_1, x_2, \ldots, x_n)$, яка задана на множині логічних змінних $\{x_1, x_2, \ldots, x_n\}$ і набуває значення 0 або 1. З комбінаторики відомо, що для n-місної функції кількість різних наборів змінних дорівнює 2^n. Логічні функції можна задати формулами або таблицями, які мають назву *таблиць істинності*. Область визначення логічної функції – це множина наборів, на яких вона визначена.

Тема 18.2. Основні елементарні логічні функції

Інверсія (логічне «НЕ», заперечення, доповнення) – логічна функція однієї змінної, яка задається таблицею істинності:

x	$\overline{x}$
0	1
1	0

Легко побачити, що $\overline{\overline{x}} = x$.

Кон'юнкція (логічне «І», логічний добуток) – це бінарна логічна функція, яка задається формулою

$$f(x_1, x_2) = x_1 \wedge x_2 = x_1 \cdot x_2. \qquad (4.10)$$

Таблиця істинності кон'юнкції має вигляд

x_1	x_2	$x_1 \wedge x_2$
0	0	0
0	1	0
1	0	0
1	1	1

x_1 та x_2 – логічні співмножники (змінні).

Диз'юнкція (логічне «АБО», математичне додавання) – це бінарна логічна функція, яка задається формулою

$$f(x_1, x_2) = x_1 \vee x_2. \tag{4.11}$$

Її таблиця істинності має вигляд

x_1	x_2	$x_1 \vee x_2$
0	0	0
0	1	1
1	0	1
1	1	1

x_1 та x_2 – логічні доданки (змінні).

Операції кон'юнкції та диз'юнкції можна узагальнити на n логічних співмножників і n логічних доданків.

Тема 18.3. Основні тотожності алгебри Буля

Диз'юнкція та кон'юнкція комутативні:

$$x_1 \vee x_2 = x_2 \vee x_1, \qquad x_1 \wedge x_2 = x_2 \wedge x_1.$$

Диз'юнкція та кон'юнкція асоціативні:

$$x_1 \vee (x_2 \vee x_3) = (x_1 \vee x_2) \vee x_3, \qquad x_1 \wedge (x_2 \wedge x_3) = (x_1 \wedge x_2) \wedge x_3.$$

Диз'юнкція та кон'юнкція взаємно дистрибутивні:

$$x_1 \vee (x_2 \wedge x_3) = (x_1 \vee x_2) \wedge (x_1 \vee x_3), \qquad x_1 \wedge (x_2 \vee x_3) = (x_1 \wedge x_2) \vee (x_1 \wedge x_3).$$

Правила де Моргана (закони де Моргана):

$$\overline{x_1 \vee x_2} = \overline{x_1} \wedge \overline{x_2}, \qquad \overline{x_1 \wedge x_2} = \overline{x_1} \vee \overline{x_2},$$

$$\overline{x_1 \vee x_2 \vee \ldots \vee x_n} = \overline{x_1} \wedge \overline{x_2} \wedge \ldots \wedge \overline{x_n}, \qquad \overline{x_1 \wedge x_2 \wedge \ldots \wedge x_n} = \overline{x_1} \vee \overline{x_2} \vee \ldots \vee \overline{x_n}.$$

Тотожності над логічними змінними:

$$x \vee x = x, \qquad x \wedge x = x,$$

$$x \vee \overline{x} = 1, \qquad x \wedge \overline{x} = 0,$$

$$x \vee 0 = x, \qquad x \wedge 0 = 0,$$

$$x \vee 1 = 1, \qquad x \wedge 1 = x.$$

Закон поглинання:

$$x_1 \vee (x_1 \wedge x_2) = x_1 \wedge (x_1 \vee x_2) = x_1,$$

$$(x_1 \vee x_2) \wedge x_2 = x_2, \qquad (x_1 \wedge x_2) \vee x_2 = x_2.$$

Закон склеювання:

$$(x_1 \vee x_2) \wedge (x_1 \vee \overline{x_2}) = x_1, \qquad (x_1 \wedge x_2) \vee (x_1 \wedge \overline{x_2}) = x_1.$$

Якщо в логічному виразі відсутні дужки, то спочатку виконуються операції заперечення, потім кон'юнкції, далі – диз'юнкції.

Інверсія, кон'юнкція, диз'юнкція – основні функції (операції) математичної логіки, що дозволяють визначити інші логічні функції (операції), наприклад такі, як: імплікація, еквівалентність, сума за модулем 2 штрих Шефера та стрілка Пірса.

Імплікація – бінарна функція. Позначається через: $x_1 \to x_2$. Висловлювання $x_1 \to x_2$ хибне тоді і лише тоді, коли x_1 – істинне, а x_2 – хибне. Таблиця істинності:

x_1	x_2	$x_1 \to x_2$
0	0	1
0	1	1
1	0	0
1	1	1

Для операції імплікації мають місце такі закони:

1. $x_1 \to x_2 = \overline{x_1} \to \overline{x_2}$ (закон контрпропозиції).

2. $(x_1 \to x_2) \wedge (x_2 \to x_3) = x_1 \to x_3$ (закон силогізму).

Еквівалентність – бінарна функція. Позначається через: $x_1 \leftrightarrow x_2$. (або $x_1 \equiv x_2$). Висловлювання $x_1 \leftrightarrow x_2$ істинне тоді і лише тоді, коли x_1 і x_2 – істинні, або x_1 і x_2 – хибні. Має місце формула

$$x_1 \leftrightarrow x_2 = (x_1 \to x_2) \wedge (x_2 \to x_1). \tag{4.12}$$

Таблиця істинності має вигляд

x_1	x_2	$x_1 \leftrightarrow x_2$
0	0	1
0	1	0
1	0	0
1	1	1

Сума за модулем 2 (*строга диз'юнкція*, *нееквівалентність*) – бінарна операція. Позначається через $x_1 \oplus x_2$. Висловлювання $x_1 \oplus x_2$ хибне тоді і лише тоді, коли x_1 і x_2 – істинні або x_1 і x_2 – хибні, тобто має місце рівність $x_1 \oplus x_2 = x_1 \leftrightarrow x_2$, яка означає, що логічне додавання є запереченням еквівалентності. Також має місце рівність:

$$x_1 \oplus x_2 = x_1 \wedge \overline{x_2} \vee \overline{x_1} \wedge x_2. \tag{4.13}$$

Таблиця істинності має вигляд

x_1	x_2	$x_1 \oplus x_2$
0	0	0
0	1	1
1	0	1
1	1	0

Властивості:

1. $x \oplus x = 0$.
2. $x \oplus 0 = x$.
3. $x_1 \oplus x_2 = x_2 \oplus x_1$ – (комутативність).
4. $x_1 \oplus (x_2 \oplus x_3) = (x_1 \oplus x_2) \oplus x_3$ – (асоціативність).
5. $x_1 \wedge (x_2 \oplus x_3) = x_1 \wedge x_2 \oplus x_1 \wedge x_3$ – (дистрибутивність).

Штрих Шефера – бінарна операція. Позначається через $x_1|x_2$. Висловлювання $x_1|x_2$ істинне, тоді і лише тоді, коли хоча б одне з висловлень x_1 і x_2 хибне. Тобто операція штрих Шефера є запереченням операції кон'юнкції. Має місце формула

$$x_1|x_2 = \overline{x_1 \wedge x_2} = \overline{x_1} \vee \overline{x_2}. \tag{4.14}$$

Таблиця істинності:

x_1	x_2	$x_1|x_2$
0	0	1
0	1	1
1	0	1
1	1	0

Стрілка Пірса або *функція Вебба*. Позначається через $x_1 \downarrow x_2$. Висловлювання $x_1 \downarrow x_2$ істинне тоді і лише тоді, коли x_1 і x_2 – хибні. Тобто операція стрілка Пірса є запереченням операції диз'юнкції.

Має місце формула

$$x_1 \downarrow x_2 = \overline{x_1 \vee x_2} = \overline{x_1} \wedge \overline{x_2}. \tag{4.15}$$

Стрілка Пірса визначається таблицею істинності:

x_1	x_2	$x_1 \downarrow x_2$
0	0	1
0	1	0
1	0	0
1	1	0

Тема 18.4. Булеві функції

Булева функція (функція алгебри логіки, логічна функція) – в дискретній математиці відображення: $B^n \to B$, де $B = \{0, 1\}$ – булева множина, B^n – множина всіх можливих послідовностей з 0 та 1 довжини n.

Булевою функцією n змінних називається відображення вигляду

$$\{0{,}1\}^n \ni \langle x_1, \ldots, x_n \rangle \xrightarrow{f} f(x_1, \ldots, x_n) \in \{0{,}1\},$$

яке набуває значень 0 та 1, аргументи якого є вектори, складені з 0 та 1. Булеву функцію n змінних називають ще n-арною операцією на множині $\{0, 1\}$.

Булевих функцій двох змінних усього 16, кожна з яких може набувати всього по чотири значення. Розташуємо їх у таблиці, де в рядках наведено значення функції для фіксованого набору значень аргументів x_1, x_2, а в стовпчиках – усі значення функцій.

Наступна таблиця є фактично таблицею істинності для всіх булевих функцій двох змінних. Серед цих функцій є ті, що вже були визначені вище.

Таблиця 4.1

x_1	x_2	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
0	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

Дійсно:

$f_1(x_1, x_2)$ – кон'юнкція;

$f_6(x_1, x_2)$ – логічне додавання;

$f_7(x_1, x_2)$ – диз'юнкція;

$f_8(x_1, x_2)$ – стрілка Пірса;

$f_9(x_1, x_2)$ – еквівалентність;

$f_{13}(x_1, x_2)$ – імплікація;

$f_{14}(x_1, x_2)$ – штрих Шеффера.

Функції:

$$f_{10}(x_1, x_2) = \bar{x}_2 \quad \text{та} \quad f_{12}(x_1, x_2) = \bar{x}_1 \quad \text{– заперечення.}$$

Вони мають так звані фіктивні змінні. Змінна x_i, називається *фіктивною*, якщо

$$f(x_1, \ldots x_{i-1}, 0, x_{i+1}, \ldots x_n) = f(x_1, \ldots x_{i-1}, 1, x_{i+1}, \ldots x_n)$$

при довільних значеннях решти змінних.

Функції сталі:

$f_0(x_1, x_2) = 0$ – константа 0;

$f_{15}(x_1, x_2) = 1$ – константа 1.

Крім того, функції присвоєння значення однієї зі змінних (також з фіктивними змінними):

$f_3(x_1, x_2) = x_1 \bar{x}_2 \vee x_1 x_2 = x_1,$

$f_5(x_1, x_2) = \bar{x}_1 x_2 \vee x_1 x_2 = x_2.$

І, нарешті, функції, що являють собою різновиди імплікацій:

$f_2(x_1, x_2) = \overline{x_1 \rightarrow x_2}$ – ліва коімплікація;

$f_4(x_1, x_2) = \overline{x_1 \leftarrow x_2}$ – права коімплікація;

$f_{11}(x_1, x_2) = x_1 \leftarrow x_2$ – права імплікація.

Логічна функція *n* змінних задається таблицею істинності, де всім можливим набором значень аргументу зіставляються відповідні значення функції.

Наприклад:

$$f(x_1, x_2, x_3) = \overline{x_1} \wedge x_2 \vee x_3,$$

має таку таблицю істинності:

x_1	x_2	x_3	$\bar{x}_1 \wedge x_2 \vee x_3$
1	1	1	1
0	1	1	1
1	0	1	0
0	0	1	1
1	1	0	0
0	1	0	1
1	0	0	0
0	0	0	1

Тема 18.5. Диз'юнктивна та кон'юнктивна нормальна форма булевих функцій

Елементарною кон'юнкцією називається кон'юнкція попарно різних змінних або їх заперечень, наприклад:

$$x_1 \wedge x_2 \wedge \overline{x_3} \quad \text{або} \quad x_1 \wedge \overline{x_2} \wedge x_3.$$

Елементарною диз'юнкцією називається диз'юнкція попарно різних змінних або їх заперечень, наприклад:

$$x_1 \vee x_2 \vee \overline{x_3}$$

Диз'юнктивною нормальною формою (ДНФ) булевої функції називається диз'юнкція будь-якої скінченної множини попарно різних елементарних кон'юнкцій, наприклад:

$$f(x_1,x_2,x_3)=(x_2\wedge x_3)\vee(x_1\wedge\overline{x_2}\wedge x_3)\vee(x_1\wedge x_3).$$

Кон’юнктивною нормальною формою (КНФ) називається кон’юнкція будь-якої скінченної множини попарно різних диз’юнкцій, наприклад:

$$f(x_1,x_2,x_3)=(x_1\vee x_3)\wedge(x_2\vee\overline{x_3})\wedge(\overline{x_1}\vee x_2\vee x_3).$$

Досконалою ДНФ (досконалою КНФ) булевої функції називається така її ДНФ (КНФ), в якої будь-яка елементарна кон’юнкція (диз’юнкція) містить в собі в прямому або інверсному вигляді всі істотні змінні.

Будь-яка формула алгебри Буля може бути зведена до ДНФ (або КНФ) так:

1. За допомогою законів де Моргана позбутися у формулі від значень заперечення відносно декількох змінних.
2. За допомогою закону дистрибутивності (3-тя тотожність).
3. Спростити отриманий вираз згідно з тотожностями пунктів 5, 6, 7 (див. основні тотожності алгебри Буля).

Перетворення ДНФ до досконалої ДНФ (ДДНФ) можна виконати розщепленням кон’юнкцій, які містять не всі істотні змінні, шляхом множення цих кон’юнкцій на диз’юнкцію вигляду $(x_i\vee\overline{x_i})$, де x_i – недостатня змінна. Далі вся формула зводиться до ДНФ.

Аналогічно для перетворення КНФ до досконалої КНФ (ДКНФ) треба розщепити диз'юнкцію, яка містить не всі істотні змінні (наприклад, не містить змінну x_i) шляхом додавання до цієї диз'юнкції доданка $(x_i\wedge\overline{x_i})$. Далі формула зводиться до КНФ.

КНФ можна отримати з ДНФ шляхом перетворення:

$$f=x_1\vee x_2\vee\ldots\vee x_n=\overline{\overline{f}}=\overline{\overline{x_1\vee x_2\vee\ldots\vee x_n}}=\overline{x_1\wedge x_2\wedge\ldots\wedge x_n}. \qquad (4.16)$$

Перетворимо вираз під знаком інверсії до ДНФ, потім застосуємо закони де Моргана та отримаємо КНФ. Аналогічно з КНФ можна отримати ДНФ.

Лекція 19. Елементи теорії графів

Тема 19.1. Основні поняття теорії графів

Теорія графів – це розділ дискретної математики, що вивчає властивості графів. Ця теорія знаходить застосування та корисна в різноманітних сферах людської діяльності: фізика, електротехніка, теорія зв’язку, проєктування ЕОМ, архітектура тощо.

Неформально *граф* – це набір точок і ліній, які їх з'єднують. Формальних визначень поняття граф багато, всі вони задають, певною мірою споріднені, але водночас розбіжні поняття.

У строгому визначенні *графом* називається така пара множин $G = (V, E)$, де V є підмножина будь-якої зліченної *множини вершин*, а E – підмножина декартового добутку $V \times V$ (*множина ребер*).

Кількість вершин графа $n(G) = |V(G)|$, а кількість ребер $m(G) = |E(G)|$. Кількість вершин $n(G)$ графа називають його *порядком*.

Інцидентність – поняття, що використовується для ребра і вершини: якщо v_1, v_2 – вершини, а $e_1 = (v_1, v_2)$ – ребро, що їх з'єднує, тоді вершина v_1 і ребро e_1 інцидентні, вершина v_2 і ребро e_1 також інцидентні. Математично це записується так: $e_1 = (v_1, v_2) \in E(G)$.

Два ребра, інцидентні одній вершині, називаються *суміжними*; дві вершини, інцидентні одному ребру, також називаються *суміжними*.

Зв'язний граф – це граф, що містить рівно одну компоненту зв'язності. Це означає, що між будь-якою парою вершин цього графа існує шлях.

Шлях – це ланцюг, усі ребра якого орієнтовані в напряму руху від початкової до кінцевої вершини ланцюга.

Маршрут у теорії графів – це скінченна або нескінченна послідовність ребер $S = \{\ldots, e, e_1, \ldots, e_n, \ldots\}$, в якій кожні два сусідні ребра $e_{i\text{-}1}$ і e_{i} мають спільну вершину. Маршрут називають *скінченним*, якщо він має початкову і кінцеву вершини. Якщо маршрут має початкову вершину, але не має кінцевої (або навпаки), то його називають *дносторонньо-нескінченним*. Якщо маршрут не має початкової і кінцевої вершин, то його називають *двосторонньо-нескінченними*.

Довжина маршруту дорівнює кількості ребер у ньому, причому кожне ребро вказується стільки разів, скільки воно зустрічається в даному маршруті.

Цикл у теорії графів – це ланцюг $v_1e_1v_2e_2\ldots v_ne_nv_I$, в якому перша та остання вершини збігаються з початковою. Якщо відсутні інші збіжності вершин, то такий цикл називається *простим*. Цикл, який містить усі ребра графа називається *ейлеровим*, а простий цикл, який містить усі вершини графа – *гамільтоновим*.

Множина ребер E може бути порожньою (рис. 4.9 *а*). Такий граф називається *нульовим* графом. Якщо ж множина вершин V – порожня, то порожня також множина E. Такий граф називається *порожнім*.

Лінії, що зображують ребра графа, можуть перетинатися, але точки перетину не є вершинами (рис. 4.9 *б*). Ребро може з'єднувати деяку вершину саму із собою (рис. 4.9 *в*), таке ребро називається *петлею*.

Цей випадок відповідає наявності в множині E пар вигляду (v, v). Різні ребра можуть бути інцидентними одній і тій самій парі вершин (тобто одну й ту саму пару вершин з'єднує більше ніж одне ребро), такі ребра називаються *кратними* (рис. 4.9 *г*).

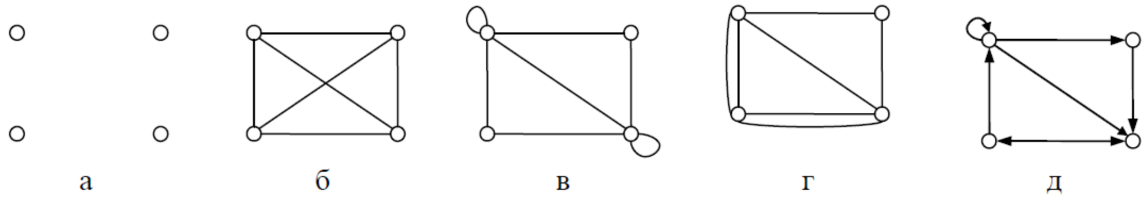


Рис. 4.9. Приклади графів

Граф називається *простим*, якщо кожну пару вершин з'єднує не більше ніж одне ребро. Граф називається *мультиграфом*, якщо він має кратні ребра. Граф називається *псевдографом*, якщо він має петлі та кратні ребра.

Розглядають також орієнтовані графи (до цього були розглянуті *неорієнтовані графи*).

Орієнтованим графом називається граф $D = (V, E)$, де V – множина вершин, $E \subseteq V \times V$ – множина орієнтованих ребер або дуг.

При зображенні орієнтованих графів напрямки ребер позначаються стрілками (рис. 4.9 *д*). Орієнтований граф може мати кратні ребра, петлі, а також петлі, що з'єднують одні й ті самі вершини, але у зворотних напрямках.

Кожному неорієнтованому графу можна поставити у відповідність орієнтований граф із тією самою множиною вершин, в якій кожне ребро замінено двома орієнтованими ребрами, що є інцидентними тим самим вершинам і мають зворотні напрямки.

Розглянемо декілька прикладів графів. *Задача про Кенігсбергзські мости* була сформульована в XVIII ст Леонардо Ейлером та поклала початок теорії графів.

Суть задачі така. Ейлер, живучи в Кенігсберзі, любив прогулюватися понад річкою і через мости (рис. 4. 10).

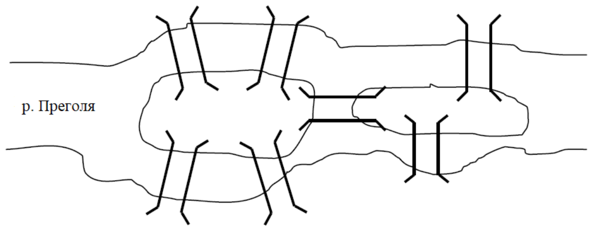


Рис. 4.10

В нього виникало питання: чи існує такий шлях, щоб вийти з дому, обійти всі мости і повернутися додому, але обійти мости так, щоб кожен міст пройти рівно один раз. Ейлер сам цю задачу сформулював і сам її розв'язав, тим самим поклавши початок використанню графів у математиці. Більше того у цьому разі суша – це вершини графа, мости – ребра. Тоді схема мостів у задачі Ейлера приводить нас до графа, який показаний на рис. 4.11.

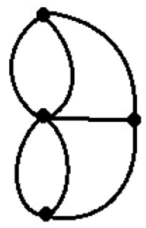


Рис. 4.11

Як видно з рис. 4.11, з вершини до вершини є кілька ребер. Такий граф називають неорієнтованим мультиграфом. Як виявилось, ця задача не має розв'язків.

Прикладом іншої проблеми, яку можна промоделювати на основі теорії графів, є *задача про три будинки та три колодязі*. Є три будинки та три колодязі, які якимось способом розташовані на площині. Потрібно провести від кожного будинку до кожного колодязя стежку так, щоб стежки не перетинались (рис. 4.12).

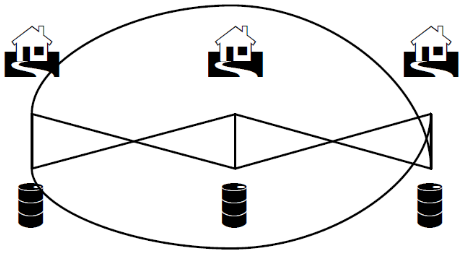


Рис. 4.12

Казимир Куратовський (відомий польський математик) у 1930 році показав, що ця задача також не має розв'язків.

Розглянемо *структуру певного Web-сайту*, де вершинами будуть сайти, а ребрами – гіперлінки. У цьому разі зображення графа може мати вигляд, який наведений на рис. 4.13.

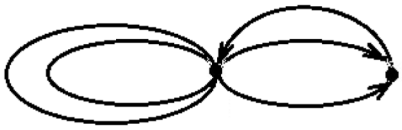


Рис. 4.13

Зліва на рис. 4.13 показано гіперлінк сторінки на саму себе. Як видно з рисунка, петлі можуть бути кратні. Граф, що містить лише петлі (x, x), називається *псевдографом*. У цілому, граф на рис. 4.13 є орієнтованим, мульти-, псевдографом.

Тема 19.2. Способи задання графів

Задати граф означає задати множини його вершин і ребер, а також відношення інцидентності. Коли граф G – скінченний, для опису його вершин та ребер досить їх занумерувати. Нехай v_1, v_2, …, v_n – вершини графа G; e_1, e_2, …, e_m – його ребра.

Відношення інцидентності можна означити матрицею ε, яка має n рядків та m стовпців. Рядки відповідають вершинам графа, а стовпці – його ребрам. Якщо ребро e_j є інцидентним вершині v_i, то $\varepsilon_{ij} = 1$, в іншому випадку $\varepsilon_{ij} = 0$. Це *матриця інцидентності* звичайного графа G, яка є одним зі способів його визначення.

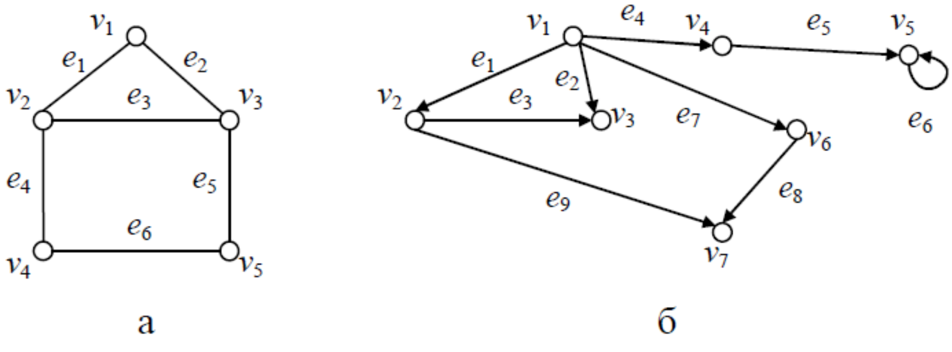


Рис. 4.14

Наприклад, для графа, який зображено на рис. 4.14 *а*, матриця інцидентності має вигляд таблиця 4.2.

Таблиця 4.2

	e_1	e_2	e_3	e_4	e_5	e_6
v_1	1	1	0	0	0	0
v_2	1	0	1	1	0	0
v_3	0	1	1	0	1	0
v_4	0	0	0	1	0	1
v_5	0	0	0	0	1	1

Таблиця 4.3

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9
v_1	-1	-1	0	-1	0	0	-1	0	0
v_2	1	0	-1	0	0	0	0	0	-1
v_3	0	1	1	0	0	0	0	0	0
v_4	0	0	0	1	-1	0	0	0	0
v_5	0	0	0	0	1	2	0	0	0
v_6	0	0	0	0	0	0	1	-1	0
v_7	0	0	0	0	0	0	0	1	1

У матриці інцидентності ε орієнтованого графа D, якщо вершина v_i – початок дуги e_j, то ε_{ij} = -1, якщо v_i – кінець e_j, то ε_{ij} = 1; якщо e_j – петля, а v_i – інцидентна їй вершина, то ε_{ij} = 2. Наприклад, для орієнтованого графа, зображеного на рис. 4.14 *б*, матриця інцидентності наведена у таблиці 4.3.

У кожному рядку матриці інцидентності для неорієнтованого або орієнтованого графа тільки два елементи відмінні від 0 (або один, якщо ребро є петлею). Тому такий спосіб задання графа не досить економний. Відношення інцидентності можна задати ще *списком ребер* графа.

Кожний рядок цього списку відповідає ребру, в ньому записано номери вершин, інцидентних йому. Для неорієнтованого графу порядок цих вершин у рядку довільний, для орієнтованого першим записується номер або інше найменування початку ребра, а другим – його кінця.

У таблицях 4.4 та 4.5 наведено списки ребер відповідно для графів з рис. 4.14.

Таблиця 4.4

Ребро	Вершини
e_1	v_1, v_2
e_2	v_1, v_3
e_3	v_2, v_3
e_4	v_2, v_4
e_5	v_3, v_5
e_6	v_4, v_5

Таблиця 4.5

Ребро	Вершини
e_1	v_1, v_2
e_2	v_1, v_3
e_3	v_2, v_3
e_4	v_1, v_4
e_5	v_4, v_5
e_6	v_5, v_5
e_7	v_1, v_6
e_8	v_6, v_7
e_9	v_2, v_7

За списком ребер графа можна легко визначити матрицю інцидентності. Справді, кожний рядок цього списку відповідає стовпцю матриці з тим самим номером. Для неорієнтованого графа в рядку списку записуються номери елементів стовпця матриці інцидентності, що дорівнюють 1, а для орієнтованого графу в цьому рядку першим зазначається номер елемента стовпця матриці, який дорівнює -1, другим – номер елемента, що дорівнює 1.

Поняття матриці інцидентності та списку ребер можна легко узагальнити на випадок мультиграфа.

Розглянемо третій спосіб задання графів за допомогою *матриці суміжності* – квадратна матриця δ, стовпцям і рядкам якої відповідають вершини графа. Для неорієнтованого графа δ_{ij} дорівнює кількості ребер, інцидентних i-й та j-й вершинам, для орієнтованого – цей елемент матриці відповідає кількості ребер з початком в i-й вершині та кінцем у j-й вершині.

Таблиця 4.6

	v_1	v_2	v_3	v_4	v_5
v_1	0	1	1	0	0
v_2	1	0	1	1	0
v_3	1	1	0	0	1
v_4	0	1	0	0	1
v_5	0	0	1	1	0

Таблиця 4.7

	v_1	v_2	v_3	v_4	v_5	v_6	v_7
v_1	0	1	1	1	0	1	0
v_2	0	0	1	0	0	0	1
v_3	0	0	0	0	0	0	0
v_4	0	0	0	0	1	0	0
v_5	0	0	0	0	1	0	0
v_6	0	0	0	0	0	0	1
v_7	0	0	0	0	0	0	0

Отже, матриця суміжності неорієнтованого графу симетрична ($\delta_{ij} = \delta_{ji}$), а орієнтованого – необов'язково. Якщо вона все ж симетрична, то для кожного ребра орієнтованого графа існує ребро, яке з'єднує ті самі вершини, але йде у зворотному напрямку. Очевидно, що орієнтований граф із симетричною матрицею суміжності відповідає неорієнтованому графу, який має ту саму матрицю суміжності.

Тема 19.3. Ізоморфізм графів та основні операції над ними

Отже, граф можна задати різними способами. Він може бути зображений на кресленні (рисунку), заданий матрицею інцидентності, списком ребер або матрицею суміжності. Вигляд креслення залежить від форми ліній і взаємного розташування вершин. Іноді не так легко зрозуміти, чи однакові графи, зображені різними кресленнями, як, наприклад, на рис. 4.15. Вигляд матриць та списку ребер залежить від нумерації вершин і ребер графа. Строго кажучи, граф вважається повністю заданим, якщо нумерацію його вершин зафіксовано.

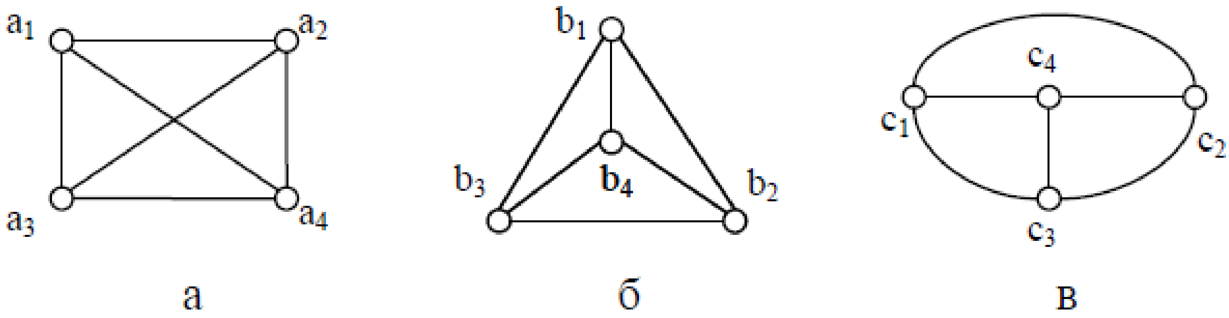


Рис. 4.15

Нехай існує бієкція φ, яка діє з множини вершин графа G на множину вершин графа H так, що для будь-яких вершин v_1 та v_2 графа G їх образи $\varphi(v_1)$ і $\varphi(v_2)$ є суміжними в H тоді й тільки тоді, коли v_1 та v_2 – суміжні в G. Така бієкція називається *ізоморфізмом* графа G на граф H, а графи G і H є *ізоморфними*.

Іншими словами, графи G і H ізоморфні, якщо шляхом перестановки рядків і стовпців матриці суміжности графа G можна отримати матрицю суміжності графа H. Таким чином, графи, зображені на рис. 4.15.

Перенумерація вершин графа задається рядком $a_1, \ldots, a_n$ нових номерів вершин, розташованих у початковому порядку. Нова матриця суміжності утворюється з початкової шляхом переміщення кожного елемента δ_{ij} в a_i-й рядок та a_j-й стовпець, тобто внаслідок перестановки $(a_1, \ldots, a_n)$ рядків і стовпців початкової матриці.

Тому, щоб дізнатися, чи зображують дві матриці суміжності ізоморфні графи, можна, наприклад, здійснити усі перестановки рядків та стовпців першої матриці. Якщо після однієї із цих перестановок виникне матриця, що тотожно збігається з другою, графи, які зображаються цими матрицями, ізоморфні. Проте, щоб пересвідчитися таким способом у тому, що графи неізоморфні, доведеться виконати всі $n!$ перестановок рядків і стовпців, а це – досить трудомістка операція.

Матриця інцидентності графа та список його ребер залежать від нумерації ребер і вершин. Перехід від однієї пари нумерації до іншої визначається перестановками $(v_1, \ldots, v_n)$ вершин та $(e_1, \ldots, e_m)$ ребер графу, який розглядається. Матриця інцидентності утворюється з початкової внаслідок перестановки стовпців (j-й на e_j-те місце) і рядків (i-й на v_i-те). Рядки списку ребер переставляються так само, як і рядки матриці інцидентності, причому кожний номер j в рядках списку замінюється номером v_j.

Розглянемо граф G. Граф H називається *підграфом* графа G, якщо виконані такі включення:

$$V_H \subset V_G, \qquad E_H \subset E_G.$$

Якщо H – підграф графа G ($H \subset G$) і $V_H = V_G$ то підграф H називається *фактором*.

Нехай G деякий граф, що містить більше однієї вершини. Тоді для будь-якої вершини $v \in V_G$ можна розглянути *операцію видалення вершини* з графа G. Будемо говорити, що $H = G - v$ граф в результаті видалення вершини v з графа G, якщо $V_H = V_G \setminus v$ і зі множини ребер вихідного графа видалені всі ребра, які починалися або закінчувалися у вершині v. Можна розглянути і *операцію додавання вершини* в граф, коли до множини вершин додається новий елемент – нова вершина. Можна визначити *операцію додавання нового ребра* в існуючому графі.

Розглянемо два графи G і H. Граф F ми назвемо *об'єднанням графів* G і H, якщо

$$V_F = V_H \cup V_G, \qquad E_F = E_H \cup E_G.$$

Трохи більш складною операцією над графами є операція декартового добутку графів.

Нехай G_1 і G_2 – це два графи. *Декартовим добутком двох графів* називається граф

$$G = G_1 \times G_2$$

множина вершин якого складається з декартового добутку множин вершин графів G_1 і G_2:

$$V_G = V_{G_1} \times V_{G_2},$$

а множина ребер E_G визначається:

$$((u_1, u_2)(v_1, v_2)) \in E_G,$$

тоді і тільки тоді, коли або

$$u_1 = v_1 \quad \text{і} \quad (u_2, v_2) \in E_{G_2}, \quad \text{або} \quad u_2 = v_2 \quad \text{і} \quad (u_1, v_1) \in E_{G_1}.$$

Тема 19.4. Дерева в теорії графів

Серед графів вагоме значення мають графи зі спеціальною структурою, які називаються деревами.

Дерево – це зв’язний ациклічний (без циклів) граф. З того, що граф зв’язний, випливає, що між будь-якими його вершинами є маршрут типу $v_1e_1v_2e_2\ldots e_nv_n$. Однак, оскільки граф ациклічний, то він завжди простий (рис. 4.16).

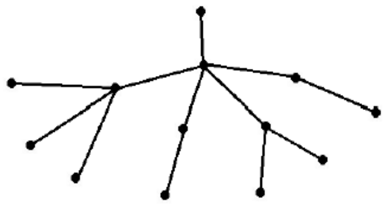


Рис. 4.16

Деревами можуть бути як орієнтовані, так і неорієнтовані графи, але їх визначення трохи різні. Неорієнтований граф називається *неорієнтованим деревом*, якщо він зв'язний і не має циклів (рис. 4.17).

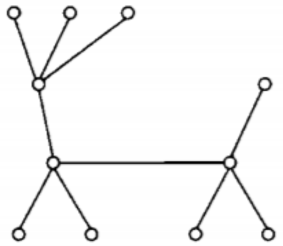
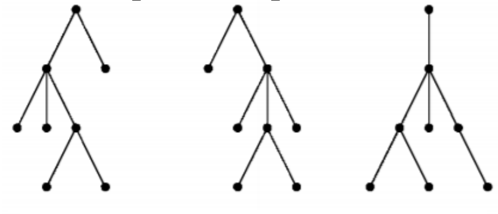


Рис. 4.17. Неорієнтоване дерево

Орієнтований граф $G = (V, E)$ називається *орієнтованим деревом*, якщо він зв'язний, у ньому існує тільки одна вершина $v_0 \in V$, до якої не входять ніякі ребра, і в кожну вершину $v \in V \setminus v_0$ входить рівно одне ребро (рис. 4.18). При цьому вершина v_0 називається *коренем дерева*.



а б в

Рис. 4.18. Приклади орієнтованих дерев: *а* та *б* – ізоморфні, але *б* і *в* та *а* і *в* неізоморфні

Розглянемо більш докладно спочатку неорієнтовані дерева. Ці графи мають ряд характерних властивостей, кожна з яких може бути покладена у визначення дерева. Будемо розглядати дерево $G = (V, E)$.

1. У дереві для будь-яких двох вершин існує єдиний маршрут, що з'єднує ці вершини.
2. Кількість ребер $|E|$ рівно на 1 менше кількості вершин $|V|$.
3. Якщо у дерева видалити будь-яке ребро, то цей граф стане незв'язаним.
4. При додаванні нового ребра в дерево це дерево стане графом, який має рівно один цикл.

Вершина в неорієнтованому графі називається «*висячою*», якщо в неї рівно одне ребро. Якщо з дерева видалити одну висячу вершину і ребро, яке має цю вершина, то ми знову отримаємо неорієнтоване дерево.

Орієнтовані дерева мають власну термінологію. Наприклад, вершина називається «*листком*», якщо з неї не виходить жодне ребро. Лист є коренем тільки в тому випадку, коли дерево складається з однієї вершини. Шлях з кореня до будь-якого листа називається «*гілкою*» *дерева*.

Максимальна довжина гілки називається *висотою дерева*. *Глибиною вершини* називається довжина шляху з кореня до цієї вершини. Якщо з вершини u йде ребро в v, вершина u називається батьківською по відношенню до v, а вершина v називається *дочірньою* по відношенню до u.

Орієнтоване дерево задає частковий порядок на множині вершин. Нагадаємо, що *порядок* називається *частковим*, якщо для деяких елементів множини задано відношення порядку, тобто рефлексивне, антисиметричне і транзитивне. Елементи, які пов'язані відношенням порядку, називаються *порівнювальними*.

Для орієнтованого дерева дві вершини u і v будуть порівнянними, якщо вони належать одній із гілок дерева.

Для порівнювальних вершин пишуть $u \geq v$, якщо існує шлях зі u у v та $u \leq v$, якщо існує шлях зі v в u.

Так само і для неорієнтованого дерева, якщо в ньому видалити вершину, що є листом, і ребро, що входить в цю вершину, ми отримаємо знову орієнтоване дерево.

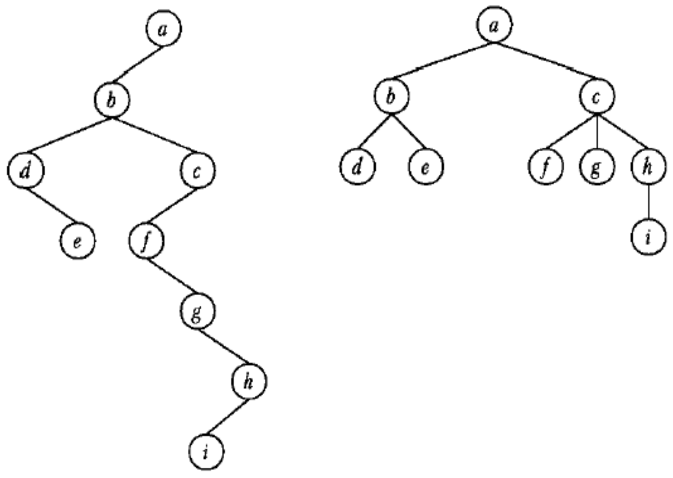
Означення *бінарного дерева*:

1. Є одна вершина – корінь дерева.
2. Усі інші вершини належать одному з піддерев (лівому або правому), які не перетинаються.

За допомогою бінарних дерев можна зобразити довільне *упорядковане дерево*.

При переході до бінарних дерев для кожної вершини ліве ребро з’єднує її зі старшим сином, праве ребро – з наступним (молодшим) братом у початковому дереві.

На рис. 4.19 зображено бінарне дерево *а* та відповідне йому упорядковане дерево *б*.



а б

Рис. 4.19

Контрольні запитання до розділу 4

1. Що таке «множина»? Які множини називають рівними? Наведіть приклади множин.
2. Що таке «потужність множин»? Які множини називають нескінченними?
3. Які операції над множинами ви знаєте? Опишіть їх.
4. Що таке «відображення множини»? Які типи відображення множин вам відомі?
5. Що таке «булева алгебра»?
6. Які ви знаєте основні елементарні логічні функції? Запишіть їх таблиці істинності.
7. Запишіть основні тотожності алгебри Буля.
8. Що таке «булева функція»? Наведіть класифікацію булевих функцій.
9. Що таке «диз'юнктивна» та «кон'юнктивна нормальні форми булевих функцій»? Наведіть приклади.
10. Що таке «графи» в дискретній математиці? Наведіть приклади графів.
11. Які способи задання графів ви знаєте? Опишіть їх.
12. Розкрийте зміст поняття ізоморфізму графів.
13. Які основні операції над графами ви знаєте?
14. Що таке «дерева» в теорії графів? Наведіть приклади дерев.

СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Огірко О. І., Галайко Н. В. Теорія ймовірностей та математична статистика : навч. посіб. Львів : ЛьвДУВС, 2017. 292 с.
2. Коноваленко О. Є., Ткачук М. А., Грабовський А. В. Дискретна математика: навч.-метод. посібник. Харків : НТУ «ХПІ», 2016. 84 с.
3. Щоголев С. А. Основи теорії ймовірностей та математичної статистики : навч.-метод. посіб. Одеса : Одеський національний університет імені І. І. Мечникова, 2015. 206 с.
4. Конечная Н. Н., Сафонова Т. А., Троицкая О. Н. Теория функций комплексного переменного : учебное пособие. Архангельск : КИРА, 2015. 111 с.
5. Олійник Л. О. Дискретна математика : навч. посіб. Дніпродзержинськ : ДДТУ, 2015. 256 с.
6. Швець В. Т. Вища математика: теорія функцій комплексної змінної : навч. посіб. Одеса : Видавництво БМВ, 2014. 284 с.
7. Єжов С. М., Разумова М. А. Теорія функцій комплексної змінної : навч. посіб. Київ : Видавничо-поліграфічний центр «Київський університет», 2012. 191 с.
8. Барковський В. В., Барковська Н. В., Лопатін О. К. Теорія ймовірностей та математична статистика : навч. посіб. 5-те вид. Київ : Центр учбової літератури, 2010. 424 с.
9. Медведєв М. Г., Пащенко І. О. Теорія ймовірностей та математична статистика. Підручник. Київ : Вид-во «Ліра-К», 2008. 536 с.
10. Гмурман В. Е. Теория вероятностей и математическая статистика : учеб. пособ. 9-е изд., стер. Москва : Высшая школа, 2003. 479 с.
11. Захаров В. К., Севастьянов Б. А., Чистяков В. П. Теория вероятностей. Москва : Наука. Главная редакция физико-математической литературы, 1983. 160 с.
12. Математическая статистика : учебник для техникумов / Иванова В. М., Калинина В. Н., Нешумова Л. А., Решетникова И. О. Москва : Высшая школа, 1975. 398 с.