

МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ

ХАРКІВСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ МІСЬКОГО
ГОСПОДАРСТВА

ВИЩА ТА ПРИКЛАДНА МАТЕМАТИКА

Теорія ймовірностей та математична статистика Математичне програмування

*(Курс лекцій для студентів I курсу денної та заочної форм навчання
освітньо-кваліфікаційного рівня бакалавр
галузі знань 0306 – «Менеджмент і адміністрування»
напряму 6.030601 – «Менеджмент»)*

УДК 519.85 (042.4)

ББК 22.18

Курс лекцій з дисципліни «Вища та прикладна математика: Теорія ймовірностей та математична статистика. Математичне програмування» (для студентів 1 курсу денної форми навчання освітньо-кваліфікаційного рівня бакалавр галузі знань 0306 – «Менеджмент і адміністрування» напрям 6.030601 – «Менеджмент») [Текст]: М. І. Самойленко, Г. В. Білогурова, О. М. Штельма, В. П. Протопопова; Харк. нац. акад. міськ. госп-ва: ХНАМГ, 2012. – 144 с.

Автори: д.т.н., проф. М. І. Самойленко,
к.т.н., доцент Г. В. Білогурова,
О. М. Штельма,
В. П. Протопопова

Курс лекцій побудовано за вимогами кредитно-модульної системи організації навчального процесу.

Рецензент: зав. кафедри вищої математики Харківської національної академії міського господарства, д-р фіз.-мат. наук, професор О. В. Грицунов.

Рекомендовано кафедрою прикладної математики і інформаційних технологій, протокол №13 від 04.05.2012 р.

© Самойленко М. І., Білогурова Г. В., Шельма О. М., Протопопова В. П.,

ХНАМГ, 2012

Метою вивчення дисципліни є формування у студентів базових математичних знань для вирішення завдань у професійній діяльності, вмінь аналітичного мислення та математичного формулювання економічних задач, що виникають у процесі управління.

Предметом дисципліни є методи, математичні та цифрові моделі задач економіки та менеджменту прикладного напрямку, зокрема стохастичні задачі та задачі на пошук екстремуму.

МОДУЛЬ 1. ТЕОРІЯ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА. МАТЕМАТИЧНЕ ПРОГРАМУВАННЯ

Змістовий модуль 1. ТЕОРІЯ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА

Тема 1. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ

1.1. Необхідність і випадковість

Теорія про причинно-слідчі зв'язки стверджує, що в об'єктивній дійсності має місце причинно-слідчий ланцюжок подій. Кожна наступна подія обумовлена та визначена попередніми подіями. Тому в кожний момент часу відбувається та подія, і тільки та, що *повинна* відбутися. Подія, що відбувається, є наслідком попередніх подій у ланцюжку. Подія, що відбулася, сама стає причиною наступних подій.

Зі сказаного випливає, що в об'єктивній дійсності випадково ніщо не відбувається і не може відбутися.

На перший погляд, останнє твердження ставить під сумнів право на існування теорії ймовірностей, предметом дослідження якої є випадкові події, випадкові явища, випадкові процеси та їх закономірності.

Щоб примирити обидві теорії, розглянемо два експерименти з киданням монети.

Для першого експерименту створимо ідеальні умови. Нехай монета багаторазово кидається:

- з наданням їй однакового моменту імпульсу сили, прикладеного до однієї і тієї ж точки монети;
- у безповітряному просторі;
- з однакової висоти;
- з того самого вихідного положення монети;
- на горизонтальну поверхню однакової пружності в кожній її точці.

При багаторазовому повторенні описаного експерименту ми будемо одержувати той самий результат – або "орел", або "решку" (у залежності

від вихідного положення монети). Тобто результат експерименту не є випадковим.

Другий експеримент будемо проводити без дотримання будь-яких ідеальних умов: так, як це робиться в повсякденному побуті при киданні жереба. При багаторазовому повторенні такого експерименту в 50-и відсотках результатом буде "орел" і 50 відсотках – "решка", причому передбачити заздалегідь результат будь-якого експерименту неможливо. Тобто результат будь-якого нового експерименту є випадковим.

Висновок. Якщо ми не можемо врахувати або свідомо зневажаємо будь-які суттєві чинники, що впливають на протікання експерименту (процесу, явища), то він автоматично стає випадковим. При багаторазовому повторенні випадкового експерименту він може протікати по-різному, і передбачити точний його хід неможливо.

1.2. Стохастичний експеримент, елементарна та випадкова події, простір подій

Початковими для теорії ймовірностей є поняття стохастичного експерименту та елементарної події.



Означення 1.1. Стохастичним експериментом (випробуванням) називають дію або послідовність дій, яка може бути повторена необмежену кількість разів в однакових умовах, при цьому результат кожного експерименту неможливо визначити заздалегідь.

Стохастичний експеримент є основним об'єктом вивчення теорії ймовірностей. Він повинен мати стійку частотну поведінку стосовно появи тих або інших результатів.

Стохастичними експериментами можуть бути, наприклад, підкидання кубика, монетки, підрахунок кількості студентів, що прийшли на лекцію.

Стохастичний експеримент грає важливу роль при моделюванні соціально-економічних і природних процесів, виступаючи в якості дешевої та зрозумілої бази для вивчення закономірностей в випадкових процесах. Саме закономірності в випадкових явищах і процесах складають **предмет** теорії ймовірностей.

Зі стохастичним експериментом пов'язане поняття елементарної події.



Означення 1.2. Фіксований результат стохастичного експерименту, який не можна виразити через сукупність інших результатів, називають **елементарною подією**.

Елементарна подія (*атомарна подія*) – один з можливих наслідків стохастичного експерименту. Елементарну подію звичайно позначають ω ,

а множину всіх елементарних подій – Ω .



Означення 1.3. Множину Ω всіх елементарних подій називають **простором елементарних подій**.

Приклад 1.1. Припустимо, що монету підкидають один раз. Простір елементарних подій цього експерименту має вигляд $\Omega = \{Г, Р\}$, де Г означає появу герба, буква Р – появу решки.


Приклад 1.2. Монету підкидають двічі. Простором елементарних подій цього експерименту є множина $\Omega = \{ГГ, ГР, РГ, РР\}$. Тут ГР означає, наприклад, що при першому підкиданні з'явився герб, а при другому – решка.

Приклад 1.3. Підкидають шестигранний гральний кубик, на якому вибиті очки від 1 до 6. Нас цікавить число очок, яке випало. Простором елементарних подій тут може бути множина, яка складається з чисел 1,2,3,4,5,6, тобто $\Omega = \{1,2,3,4,5,6\}$.



Означення 1.4. Будь-яка підмножина A множини Ω елементарних подій називають **випадковою подією** $A \subset \Omega$.

Кажуть, що внаслідок експерименту *сталася* випадкова подія A , якщо наслідок експерименту (елементарна подія) належить до підмножини A ($A \subset \Omega$).

Примітка. У цьому підрозділі (в електронному варіанті підручника) теоретичний матеріал супроводжується стохастичним експериментом, що полягає в однократному киданні гральної кістки і спостереженні за числом очок, що випадає на її верхній грані. Кістка має шість граней, на кожній з яких намальовані крапки, що відповідають числу очок від одного до шести. Результатом такого експерименту може бути випадіння 1,2,3,4,5 або 6 очок. Експеримент можна повторювати необмежену кількість разів. Для виклику експерименту належить клацнути лівою кнопкою миші на позначці . Надалі будемо використовувати цю позначку як посилання на стохастичний експеримент з киданням монети. Доречі, для електронної реалізації стохастичного експерименту необхідно мати його цифрову модель. В цифрову модель вбудовується генератор випадкових чисел з потрібними характеристиками для забезпечення адекватності моделі реальному експерименту.

1.3. Математична модель стохастичного експерименту

Для застосування математичних методів управління різними процесами необхідно, перш за все, встановити співвідношення між величинами, що характеризують ці процеси. Кожне таке співвідношення являє собою математичну модель даного процесу.

У багатьох випадках математичну модель можна побудувати суто математичним шляхом на основі відомих законів механіки, фізики та інших дисциплін, які користуються кількісними співвідношеннями. Такі моделі називаються теоретичними.

Однак існують і такі процеси, для яких принципово неможливо побудувати моделі математичним шляхом. У таких випадках доводиться вдаватися до експериментального дослідження цих процесів і будувати відповідні моделі шляхом статистичної обробки отриманих даних. Такі моделі називаються статистичними.




Означення 1.5. Математична модель стохастичного експерименту – співвідношення між величинами, які характеризують цей експеримент і які отримані шляхом статистичної обробки експериментальних даних.

Як теоретичні моделі, що виводяться математично з первинних законів, так і статистичні моделі, отримані на основі статистичної обробки результатів спостережень, можуть бути детермінованими або стохастичними. Так, наприклад, залежність $Y = \varphi(x)$, що встановлюється за результатами спостережень випадкових величин Y і X методом найменших квадратів, є детермінованою моделлю. Якщо ж врахувати спостережувані в результаті дослідів випадкові відхилення експериментальних точок від кривої $y = \varphi(x)$ і написати залежність Y від X в вигляді: $Y = \varphi(X) + Z$, де Z – деяка випадкова величина, то вийде стохастична модель.

Методи побудови статистичних моделей становлять важливий розділ сучасної математичної статистики. Побудову статистичної моделі реального процесу зазвичай називають ідентифікацією цього процесу.

1.4. Алгебра випадкових подій

Прості випадкові події можуть утворювати складні події, а складні – ще більш складні. Наприклад, подія A_{2k} , що полягає у випадінні парного числа очок в експерименті , утворюється з трьох простих випадкових подій: A_2 – випадіння 2 очок, A_4 – випадіння 4 очок, A_6 – випадіння 6 очок. А більш складна подія A , що полягає у випадінні парного числа очок або числа очок, кратних трьом, складається з двох менш складних подій:

- A_{2k} – випадіння парного числа очок (події A_2, A_4, A_6);
- A_{3k} – випадіння числа очок, кратного трьом (події A_3, A_6).

Теорія ймовірностей, як це буде показано в наступному підрозділі, дає дослідникам можливість за відомими ймовірностями простих подій розраховувати ймовірності складних подій. Але щоб скористатися цією можливістю, необхідно володіти умінням розкладати складні події на прості або

складати з простих подій складні. Для розкладання або утворювання складних подій досить володіти двома операціями над подіями: додаванням і множенням.

1.4.1. Сума подій




Означення 1.6. **Сумою** двох подій A і B називають таку подію C , що відбувається тоді, коли відбувається або подія A , або подія B , або події A і B одночасно в одному експерименті.

Суму подій A і B прийнято записувати в такий спосіб: $A + B = C$.

Операція додавання має місце, коли прості події об'єднуються в складну з використанням сполучника "або".

Використання сполучника «або» еквівалентно використанню словосполучення "хоча б". Так, *означення 1.21* з використанням цього словосполучення звучало б у такий спосіб: «**Сумою** двох подій A і B називають таку подію C , що відбувається тоді, коли відбувається принаймні одна з подій A і B ».

Приклад 1.4. Нехай в умовах експерименту  подія $A = \{a_1, a_3, a_5\}$ визначається як випадіння непарного числа очок, а подія $B = \{a_3, a_6\}$ – як випадіння числа очок, кратного трьом. Тоді подія $C = \{a_1, a_3, a_5, a_6\}$, що полягає у випадінні непарного числа очок або числа очок, кратного трьом, буде сумою подій A і B .

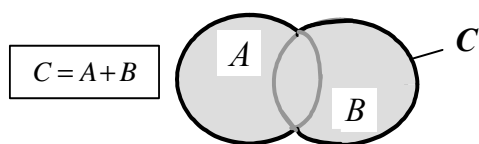


Рис.1.1

На рис.1.1 дана графічна інтерпретація операції додавання подій. Тут область A є множиною точок, які відповідають елементарним виходам експерименту, що сприяють події A ; область B – події B ; область C – принаймні одній з подій A і B .

Операція додавання може бути узагальнена на додавання декількох подій. У цьому випадку сумою декількох подій буде подія, що полягає у появі принаймні однієї з тих, що додаються.

Для операції додавання подій існує аналог операції над множинами – це об'єднання множин. Кількість елементів в об'єднаній множині обчислюються за формулами **включень та виключень**:

– для об'єднання двох множин за формулою

$$\text{Card}(A \cup B) = \text{Card } A + \text{Card } B - \text{Card}(A \cap B); \quad (1.1)$$

– для об'єднання трьох множин за формулою

$$\begin{aligned} \text{Card}(A \cup B \cup C) = & \text{Card } A + \text{Card } B + \text{Card } C - \text{Card}(A \cap B) - \\ & - \text{Card}(A \cap C) - \text{Card}(B \cap C) + \text{Card}(A \cap B \cap C); \end{aligned} \quad (1.2)$$

– для об'єднання чотирьох множин за формулою

$$\begin{aligned} \text{Card}(A \cup B \cup C \cup D) = & \text{Card } A + \text{Card } B + \text{Card } C + \text{Card } D - \\ & - \text{Card}(A \cap B) - \text{Card}(A \cap C) - \text{Card}(A \cap D) - \text{Card}(B \cap C) - \\ & - \text{Card}(B \cap D) - \text{Card}(C \cap D) + \text{Card}(A \cap B \cap C) + \text{Card}(A \cap B \cap D) + \\ & + \text{Card}(A \cap C \cap D) + \text{Card}(B \cap C \cap D) - \text{Card}(A \cap B \cap C \cap D); \quad (1.3) \end{aligned}$$


1.4.2. Добуток подій



Означення 1.7. *Добутком* двох подій A і B називають таку подію D , що відбувається тоді, коли відбувається і подія A , і подія B одночасно в одному експерименті.

Добуток подій A і B прийнято записувати в такий спосіб: $A * B = D$.

Операція множення має місце, коли прості події об'єднуються в складну з використанням сполучника "і".

Приклад 1.5. Нехай в умовах експерименту  подія $A = \{a_1, a_3, a_5\}$ визначається як випадіння непарного числа очок, а подія $B = \{a_3, a_6\}$ – як випадіння числа очок, кратного трьом. Тоді подія $D = \{a_6\}$, що полягає у випадінні непарного числа очок і одночасно кратного трьом, буде добутком подій A і B .

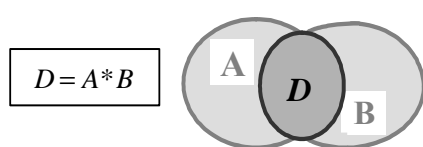


Рис.1.2

На рис.1.2 дана графічна інтерпретація операції множення подій. Тут область A є множиною точок, які відповідають елементарним виходам експерименту, що сприяють події A ; область B – події B ; область D є множиною точок, які відповідають елементарним виходам експерименту, що одночасно

сприяють і події A , і події B .

Операція множення може бути узагальнена на множення декількох подій. У цьому випадку добутком декількох подій буде подія, яка полягає в одночасній появі в одному експерименті всіх подій.

Для операції множення подій існує аналог операції над множинами – це перетин множин.

1.5. Означення ймовірності

1.5.1. Поняття ймовірності

Коли людина аналізує випадкову подію, його, насамперед, цікавить, як часто подія може відбутися або не відбутися в результаті експерименту. З цією метою вводиться спеціальна характеристика – *ймовірність події*.



Означення 1.8. Ймовірністю події називається чисельна міра ступеня об'єктивної можливості появи події в результаті нового експерименту.

У теорії ймовірностей прийнято події позначати заголовними латинськими літерами A , B і т.д., а ймовірності подій – відповідно $P(A)$, $P(B)$ і т.д.

Ймовірність довільної події (позначимо її через A) лежить у межах від нуля до одиниці: $0 \leq P(A) \leq 1$. Останнє співвідношення часто називають шкалою ймовірностей.

При вирішенні будь-яких задач необхідно стежити за тим, щоб ні в яких кінцевих або проміжних результатах не з'явилися значення ймовірності якоїсь події, що перевищує межі зазначеної шкали.

Усі події діляться на *достовірні*, *неможливі* і власне *випадкові*.



Означення 1.10. Достовірною називається подія, що у результаті експерименту неодмінно повинна відбутися (позначається: Ω).

Ймовірність достовірної події дорівнює одиниці: $P(\Omega) = 1$.



Означення 1.11. Неможливою називається подія, що у результаті експерименту не може відбутися (позначається: \emptyset).

Ймовірність неможливої події дорівнює нулю: $P(\emptyset) = 0$.



Означення 1.12. Випадковою називається подія, що при багаторазовому повторенні експерименту в одних виходах відбувається, а в інших – ні.

Ймовірність випадкової події (позначимо її через A) більше нуля і менше одиниці: $0 < P(A) < 1$.

У якості прикладів розглянемо експерименти з киданням гральної кіс-

тки. Так, при однократному киданні гральної кістки випадіння:


- не більш шести очок є достовірною подією;
- десятих очок – неможливою подією;
- трьох очок – випадковою подією.

Одним з ключових завдань розділу "Випадкові події" є чисельне визначення ймовірності випадкових подій. Перед тим, як перейти до визначення ймовірності найпростіших подій, введемо ще ряд означень.



Означення 1.13. Декілька подій в експерименті називаються **рівноможливими**, якщо за умовами симетрії експерименту немає підстави вважати появу якоїсь з них більш можливою, ніж появи інших.

Рівноможливі події мають рівний ступінь об'єктивної можливості відбутися в результаті експерименту.


В умовах експерименту  у якості прикладів *рівноможливих* подій можуть виступати події, що полягають у випадінні:

- 1, 2, 3, 4, 5, 6 очок;
- парного і непарного числа очок.

Прикладом *нерівноможливих* подій є випадіння двох очок і непарного числа очок.



Означення 1.14. Декілька подій називаються **несумісними**, якщо ніякі дві з них не можуть відбутися одночасно в одному експерименті.

Звертаючись знову до експерименту , у якості прикладів несумісних подій можна відзначити випадіння:

- 1, 2, 3, 4, 5, 6 очок;
- парного і непарного числа очок.

Прикладом сумісних подій є випадіння одного і непарного числа очок.



Означення 1.15. **Повною групою** подій називаються декілька попарно несумісних подій таких, що в результаті експерименту одна з них неодмінно повинна відбутися

В експерименті , повною групою подій є випадіння:

- 1, 2, 3, 4, 5, 6 очок;
- парного і непарного числа очок.




Означення 1.16. Якщо виходи експерименту утворюють повну групу несумісних рівноможливих подій, то вони називаються **випадками**

В умовах однократного експерименту  випадками є виходи, що полягають у випадінні:

- 1, 2, 3, 4, 5, 6 очок;
- парного і непарного числа очок.



Означення 1.11. Випадок називається **сприятливим** до події, якщо його поява тягне за собою появу цієї події.

Приклад 1.6. Нехай подія A полягає у випадінні непарного числа очок в експерименті . Тоді серед шести випадків, що полягають у випадінні 1, 2, 3, 4, 5 або 6 очок, сприятливими до події A будуть тільки три – випадіння 1, 2 або 3 очок.

1.5.2. Класичне означення ймовірності

Ймовірність події A як можливого виходу деякого експерименту визначається відношенням кількості випадків, що сприятливі до події A , до загальної кількості випадків у даному експерименті.

Таким чином, якщо m – кількість випадків, що сприятливі до події A , а n – загальна кількість випадків у даному експерименті, то ймовірність події A

$$P(A) = \frac{m}{n} . \quad (1.4)$$

Співвідношення (1.4) є класичною формулою обчислення ймовірності подій, які можуть виникати в результаті експерименту з виходами, що підпадають під означення 1.9. Обчислення ймовірності події за класичною формулою (1.4) обумовлює наступну послідовність дій (алгоритм):

1. Визначення загальної кількості випадків в експерименті, запропонованому за умовами задачі.
2. Визначення кількості випадків, що сприятливі до події, ймовірність якої необхідно знайти за умовами задачі.
3. Обчислення шуканої ймовірності події, як відношення чисел, знайдених у попередніх пунктах 1 і 2.

Наведену методику визначення ймовірності події називають схемою випадків.

Приклад 1.7. Нехай в умовах експерименту  слід визначити ймо-

вірність випадіння парного числа очок.

Розв’язання першим способом. У якості випадків при киданні гральної кістки розглядаємо групу подій: випадіння 1, 2, 3, 4, 5 і 6 очок. Тоді:

- загальна кількість випадків в експерименті – шість;
- кількість випадків, що сприятливі до події “парне число очок” – три: випадіння 2, 4 і 6 очок;
- шукана ймовірність – $3/6$, або $1/2$.

Розв’язання іншим способом. У якості випадків при киданні гральної кістки розглядаємо групу подій: випадіння “парного” і “непарного” числа очок.

- загальна кількість випадків в експерименті – два;
- кількість випадків, що сприятливі до події “парне число очок” – один (випадок “парне”);
- шукана ймовірність – $1/2$.

Як бачимо, результат вирішення не залежить від того, яку повну групу подій експерименту вважати випадками. Бажано тільки, щоб через випадки цієї групи можна було визначити шукану ймовірність події.

Слід зауважити, що не завжди всі можливі або сприятливі випадки в експерименті знаходяться так просто, як у *прикладі 1.4*. На підтвердження сказаного розглянемо ще два приклади.

Приклад 1.8. У ящику знаходиться 100 деталей, з них 10 бракованих. Навмання витягають 4 деталі. Знайти ймовірність події A – наявність рівно трьох стандартних деталей серед витягнутих.

Як бачимо, у прикладі усі можливі виходи експерименту являють собою комбінації 4 деталей, серед яких можуть бути групи, що складаються з деталей:

- тільки бракованих;
- трьох бракованих і однієї стандартної;
- двох бракованих і двох стандартних;
- однієї бракованої і трьох стандартних;
- тільки стандартних.

Виходи, що сприятливі, являють собою одну групу комбінацій, що складаються з трьох стандартних і однієї бракованої.

Цілком очевидно, що визначення кількості комбінацій у кожній групі методом прямого перебору являє собою досить складну задачу.

Приклад 1.9. З п'ятих літер розрізної абетки складене слово “КНИГА”. Дитина, що не вміє читати, розсипала літери, а потім зібрала їх

у довільному порядку. Знайти ймовірність того, що знову утворилося слово “КНИГА”.

У цьому прикладі загальна кількість випадків визначається кількістю можливих перестановок літер, з яких складається слово “КНИГА”, і ця кількість досить значна.

Задачами на відшукування кількості комбінацій елементів, подібними приведеним у *прикладів 1.8 – 1.9*, займається розділ математики, відомий як комбінаторика. В наступному підрозділі наведені найбільш важливі комбінаторні співвідношення. Після ознайомлення з цим підрозділом ми повернемося до розв’язання *прикладів 1.8 – 1.9*.

1.5.3. Статистичне означення ймовірності

Виникає необхідність в іншому визначенні ймовірності, застосовне тоді, коли експеримент не відповідає схемі випадків.

Нехай при багаторазовому повторенні стохастичного експерименту подія A наступила m раз в серії з n випробувань. Відносною частотою події $\varphi(A)$ називають відношення числа випробувань, в яких настала подія A , до числа всіх проведених випробувань: $\varphi(A) = \frac{m}{n}$.

Експериментально доведено, що частота має властивість стійкості: якщо число випробувань в серії досить велике, то відносні частоти події A в різних серіях одного і того ж експерименту мало відрізняються один від одного.

Статистичною ймовірністю події називають число, до якого прагнуть відносні частоти, якщо число дослідів необмежено зростає $P(A) = \lim_{n \rightarrow \infty} \frac{m}{n}$.

На відміну від апіорної (обчисленої до експерименту) класичної ймовірності статистична ймовірність є апостеріорної (отриманої після досвіду).

1.5.4. Основні поняття комбінаторного аналізу

Основні правила комбінаторного аналізу дозволяють визначати загальну кількість різних способів виконання певної роботи у залежності від способів виконання окремих її операцій і їх відношень між собою. До основних правил належать правило додавання та правило множення.

Правило додавання. Нехай деяку роботу можна виконати за допомогою k *взаємовиключних* операцій. При цьому перша операція може бути реалізована n_1 способами, друга – n_2 способами, ..., k -а – n_k способами. Тоді роботу можна виконати $n_1 + n_2 + \dots + n_k$ способами.

Правило множення. Нехай деяку роботу можна виконати за допомо-

гою k **послідовних** операцій. При цьому перша операція може бути реалізована n_1 способами, друга – n_2 способами, ... , k -а – n_k способами. Тоді всю роботу можна виконати $n_1 * n_2 * \dots * n_k$ способами.

1.6. Елементи комбінаторного аналізу

В комбінаториці розрізняють три види різних з'єднань (комбінацій) елементів довільної множини:

- перестановки;
- розміщення;
- сполучення.

При вирішенні багатьох задач теорії ймовірностей часто доводиться звертатися до формул комбінаторики, що дозволяють без здійснення повного перебору можливих з'єднань визначати їх загальну кількість.

Розглянемо послідовно усі види з'єднань і відповідні формули підрахунку їх кількості.

1.6.1. Перестановки



Означення 1.12. *Перестановками* з m елементів називають такі їх з'єднання, що відрізняються одне від одного порядком входження елементів.

Загальна кількість можливих перестановок з m елементів позначається P_m і визначається виразом

$$P_m = m! = 1 \cdot 2 \cdot \dots \cdot m. \quad (1.5)$$

Приклад 1.10. Скількома способами можна розташувати на полиці в ряд три різні книги.

Розв'язання. Загальна кількість можливих способів розташування книг визначається відповідно до виразу (1.5): $P_3 = 3! = 6$.

Легко помітити, що такий же результат можна одержати, застосовуючи правило множення. На перше місце на полицю можна поставити будь-яку з 3 книг (3 способи), після чого на друге місце – будь-яку з двох, що залишилися (2 способи), після чого на третьому місці буде стояти остання не розміщена книга (1 спосіб), що за правилом множення дає $3 \cdot 2 \cdot 1 = 6$ способів. Таким чином, правило множення можна вважати логічним обґрунтуванням формули (1.5).

1.6.2. Розміщення



Означення 1.13. Розміщеннями з n елементів по m називають такі з'єднання m елементів, що відрізняються одне від одного принаймні одним новим елементом або порядком їх входження ($m \leq n$).

Загальна кількість можливих розміщень з n елементів по m позначається A_n^m і визначається виразом

$$A_n^m = \frac{n!}{(n-m)!}. \quad (1.6)$$

Приклад 1.11. Скількома способами можна вибрати дві книги з трьох і розташувати їх у ряд на полиці.

Розв'язання. Загальна кількість можливих способів розташування книг визначається відповідно до виразу (1.6): $A_3^2 = \frac{3!}{(3-2)!} = 6$.

Знову ж, підставою для застосування формули (1.3) є правило множення. На перше місце на полицю можна поставити будь-яку з 3 книг (3 способи), після чого на друге – будь-яку з двох, що залишилися (2 способи). За правилом множення це складає $3 \cdot 2 = 6$ способів.

1.6.3. Сполучення



Означення 1.14. Сполученнями з n елементів по m називають такі з'єднання m елементів, що відрізняються одне від одного принаймні одним новим елементом ($m \leq n$).

Загальна кількість можливих сполучень з n елементів по m позначається C_n^m і визначається виразом

$$C_n^m = \frac{n!}{m!(n-m)!}. \quad (1.7)$$

Приклад 1.12. Скількома способами можна вибрати дві книги з трьох.

Розв'язання. Загальна кількість можливих способів вибору книг визначається відповідно до виразу (1.7): $C_3^2 = \frac{3!}{2!(3-2)!} = 3$.

Ця задача відрізняється від *прикладу 1.12* тим, що 2 книги, наприклад, AB і BA , є різними розміщеннями й однаковими сполученнями, тобто в сполученнях не враховується порядок елементів. Отже, кількість сполучень може бути отримана з числа розміщень шляхом його ділення на число перестановок, що вибираються. У даному випадку число розміщень 3 книг по 2 дорівнює 6; число перестановок з двох книг, що вибираються (або тих, що розміщуються у *прикладі 1.12*), дорівнює двом, так що вибрати 2 книги з 3 можна $6/2 = 3$ способами.

Загальний зв'язок між перестановками, розміщеннями і сполученнями показана нижче у формулі (1.10).

1.7. Приклади побудови ймовірнісного простору

Приклад 14. У коробці лежать 9 білих і 4 чорних кулі. Виймають намання дві кулі. Знайти ймовірність того, що вони:

- а)* білі;
- б)* чорні;
- в)* різнокольорові;
- г)* обидві білі або обидві чорні.

Розв'язання. Позначимо події:

A – кулі білі,

B – кулі чорні,

C – кулі різнокольорові,

D – кулі або білі, або чорні.

Для вирішення задачі використовуємо *класичну формулу ймовірності* (1.4)

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

що потребує обчислення загальної кількості n випадків в експерименті і кількості m сприятливих випадків для кожної з подій A, B, C, D , які ми позначимо відповідно m_A, m_B, m_C, m_D .

Загальна кількість випадків n в експерименті – це кількість комбінацій з $13=(9+4)$ куль по 2, тобто $n = C_{13}^2 = 78$.

а) Для події A кількість сприятливих випадків m_A – це кількість різних комбінацій з 9 білих куль по 2, тобто $m_A = C_9^2 = 36$.

Таким чином,
$$P(A) = \frac{m_A}{n} = \frac{36}{78} = \frac{6}{13}.$$

б) Для події B кількість сприятливих випадків m_B – це кількість різних

комбінацій з 4 чорних куль по 2, тобто $m_B = C_4^2 = 6$.

$$\text{Таким чином, } P(B) = \frac{m_B}{n} = \frac{6}{78} = \frac{1}{13}.$$

в) Для події C кількість сприятливих випадків m_C визначається відповідно до правила множення: $m_C = 4 \cdot 9 = 36$.

$$\text{Таким чином, } P(C) = \frac{m_C}{n} = \frac{36}{78} = \frac{6}{13}.$$

г) За результатами задач а) і б) даного прикладу дві білі кулі можна одержати $m_A = 36$ способами, а дві чорні – $m_B = 6$ способами. Тоді за правилом додавання $m_D = 36 + 6 = 42$.

$$\text{Таким чином, } P(D) = \frac{m_D}{n} = \frac{42}{78} = \frac{7}{13}.$$

Повернемося до розгляду *прикладів 1.8 і 1.9.*

Розв’язання прикладу 1.8. Загальна кількість можливих комбінацій по 4 деталі зі 100, що відрізняються принаймні одною новою деталлю, визначається як $n = C_{100}^4$. Усі вони утворюють повну групу несумісних рівноможливих випадків. Підрахуємо кількість сприятливих випадків до події A . Три стандартні деталі з 90 наявних у ящику можна витягти C_{90}^3 способами. З кожною отриманою вибіркою з 3 стандартних деталей може сполучитися одна нестандартна деталь з 10 наявних у ящику C_{10}^1 способами. Отже, за правилом множення $m = C_{90}^3 \cdot C_{10}^1$, а $P(A) = \frac{m}{n} = \frac{C_{90}^3 \cdot C_{10}^1}{C_{100}^4} = 0,3$.

Розв’язання прикладу 1.9. З п'ятих літер дитина може скласти різні з'єднання, що відрізняються одне від одного тільки порядком входження літер. Тому загальну кількість усіх можливих виходів експерименту знайдемо як кількість перестановок з 5 елементів: $n = P_5 = 5! = 120$. Усі виходи утворюють повну групу несумісних рівноможливих випадків, з котрих тільки один сприятливий до відновлення слова “КНИГА”. Отже, шукана ймовірність $P(A) = \frac{m}{n} = \frac{1}{120}$.

Тема 2. ОСНОВНІ ТЕОРЕМИ ТЕОРІЇ ЙМОВІРНОСТІ

2.1. Ймовірність суми подій

Основні теореми теорії ймовірностей дозволяють за відомими ймовірностями простих подій визначати ймовірності більш складних подій. Тобто передбачається, що ймовірності всіх подій, на які розкладається складна подія, відомі.

Основні теореми теорії ймовірностей включають дві теореми:

- теорему про ймовірність суми двох подій;
- теорему про ймовірність добутку двох подій.

Сформулюємо і доведемо першу з них.



Теорема 2.1. Ймовірність суми двох подій A і B дорівнює сумі їх імовірностей за відрахуванням ймовірності добутку цих же подій:

$$P(A + B) = P(A) + P(B) - P(A * B). \quad (2.1)$$

Доведення. Нехай результати досліду утворюють повну групу n несутісних рівноможливих подій (рис. 2.1). При цьому

- m з них сприятливі події A ;
- k з них сприятливі події B ;
- l з них сприятливі добутку подій $A * B$.

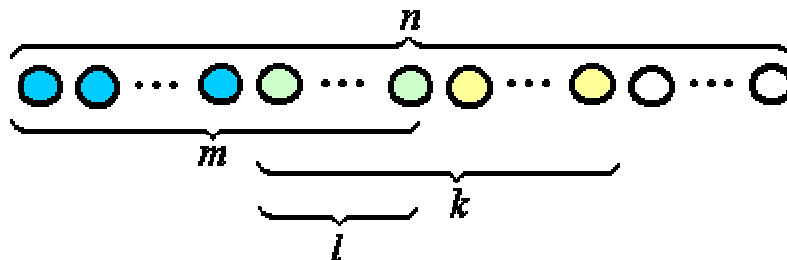


Рис. 2.1

Тоді відповідно до класичної формули визначення ймовірності

$$P(A) = \frac{m}{n}; \quad P(B) = \frac{k}{n}; \quad P(A * B) = \frac{l}{n}.$$

Відповідно до тієї ж формули ймовірність появи події A або B

$$P(A + B) = \frac{m + k - l}{n}.$$

Перетворимо останню рівність:

$$P(A + B) = \frac{m + k - l}{n} = \frac{m}{n} + \frac{k}{n} - \frac{l}{n} = P(A) + P(B) - P(A * B),$$

що й треба було довести.

Наслідок теореми 2.1. Ймовірність суми двох несумісних подій A і B дорівнює сумі їх імовірностей.

$$P(A + B) = P(A) + P(B). \quad (2.2)$$

Наслідок очевидний, оскільки ймовірність добутку несумісних подій дорівнює неможливій події, а ймовірність неможливої події дорівнює нулю: $A * B = \emptyset$ і $P(\emptyset) = 0$.

Наслідок легко узагальнюється на випадок кількох подій.

2.2. Повна група подій і протилежні події

Сума ймовірностей несумісних подій A_1, A_2, \dots, A_n , що складають **повну групу**, дорівнює одиниці:

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1. \quad (2.3)$$

Доведемо це твердження. Нехай несумісні події A_1, A_2, \dots, A_n утворюють повну групу в деякому просторі подій. Тоді, за визначенням повної групи, їх сума є достовірною подією:

$$A_1 + A_2 + \dots + A_n = U. \quad (2.4)$$

Події A_1, A_2, \dots, A_n несумісні, тому до лівої частини отриманої рівності можна застосувати узагальнений наслідок теореми 2.1, тобто маємо:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n). \quad (2.5)$$

Ймовірність достовірної події, що знаходиться в правій частині рівності (2.4), дорівнює одиниці: $P(U) = 1$. Співставляючи вирази (2.4) та (2.5), приходимо до формули для суми повної групи подій (2.3), що й треба було довести.



Означення 2.1. Дві події A і \bar{A} називаються **протилежними**, якщо вони утворюють повну групу несумісних подій.

Приклади. При стрільбі по мішені дві події, що полягають відповідно у влученні й промаху, є протилежними. При киданні монети події, що по-

лягають у випадінні "орла" і "решки", також є протилежними.

Сума ймовірностей протилежних подій дорівнює одиниці

$$P(A) + P(\bar{A}) = 1. \quad (2.6)$$

Згідно з означенням 2.1 події A і \bar{A} є несумісними і складають повну групу. Отже для них справедлива формула (2.3). Застосовуючи формулу (2.3) до подій A і \bar{A} , приходимо до виразу (2.6).

З (2.6) випливають рівності:


$$P(\bar{A}) = 1 - P(A); \quad (2.7)$$

$$P(A) = 1 - P(\bar{A}). \quad (2.8)$$

2.3. Залежні й незалежні події



Означення 2.2. Дві випадкові події називаються *незалежними*, якщо настання однієї з них не змінює ймовірності настання іншої.

Приклад 2.1. В умовах експерименту  випадіння 1, 2, 3, 4, 5 або 6 очок – це події взаємно незалежні. Скільки б гральну кістку не кидали, ймовірності випадіння 1, 2, 3, 4, 5 або 6 очок не зміняться.

Приклад 2.2. З урни, що містить 3 червоні й 2 сині кулі, послідовно витягують навмання дві. Поява червоної кулі при першому і другому вийманні – події залежні, оскільки ймовірність появи червоної кулі при другому вийманні буде залежати від того, яка куля була витягнута при першому вийманні. Так, якщо першою витягнутою кулею була червона, то ймовірність появи червоної кулі при другому вийманні дорівнює $2/4$; якщо – синя, то – $3/4$ (див. *приклад 2.3*).

Кілька подій називаються спільно незалежними або незалежними в сукупності, якщо кожна з них і будь-яка комбінація з них є подіями незалежними.

Попарна незалежність подій не означає їх незалежність в сукупності, проте незалежність подій в сукупності обумовлює їх попарну незалежність.

2.4. Умовна ймовірність



Означення 2.3. Для залежних подій A і B ймовірність події B , яка обчислена за умови, що подія A відбулася, називається **умовною** ймовірністю і позначається $P_A(B)$ або $P(B/A)$.

В умовах *прикладу 2.2* ймовірності появи червоної кулі при другому

вийманні $2/4$ і $3/4$ – це умовні ймовірності. Причому ймовірність $2/4$ обчислена за умови, що першою витягується червона куля, а ймовірність $3/4$ – що синя.

Приклад 2.3. З урни, що містить 3 червоні й 2 сині кулі, послідовно витягують навмання дві кулі. Визначити ймовірності появи кулі кожного кольору при першому вийманні, а також умовні ймовірності появи куль кожного кольору при другому вийманні.

Розв’язання. Введемо позначення:

n – загальне число куль в урні;

m – число червоних куль в урні;

k – число синіх куль в урні;

A_i – подія, яка полягає в тому, що навмання витягнута i -а куля опиниться червоною, $i = 1, 2$;

B_i – подія, яка полягає в тому, що навмання витягнута i -а куля опиниться синьою, $i = 1, 2$;

$P(A_1)$ – ймовірність події A_1 ;

$P(B_1)$ – ймовірність події B_1 ;

$P(A_2/A_1)$ – умовна ймовірність події A_2 за умови, що відбулася A_1 ;

$P(A_2/B_1)$ – умовна ймовірність події A_2 за умови, що відбулася B_1 ;

$P(B_2/A_1)$ – умовна ймовірність події B_2 за умови, що відбулася A_1 ;

$P(B_2/B_1)$ – умовна ймовірність події B_2 за умови, що відбулася B_1 .

Тоді відповідно до класичної формули визначення ймовірності (1.1) шукані ймовірності визначаються в такий спосіб:

$P(A_1)=3/5$, оскільки сприятливих виходів експерименту $m=3$ (число червоних куль в урні), а загальне число виходів $n=5$ (усього куль в урні);

$P(B_1)=2/5$, оскільки сприятливих виходів досліду $m=2$ (число синіх куль в урні), а загальне число виходів $n=5$ (усього куль в урні);

$P(A_2/A_1)=2/4$, оскільки сприятливих виходів експерименту $m=2$ (число червоних куль після виймання однієї червоної кулі), а загальне число виходів $n=4$ (усього куль в урні після першого виймання);

$P(A_2/B_1)=3/4$, оскільки сприятливих виходів експерименту $m=3$ (число червоних куль після виймання однієї синьої кулі), а загальне число виходів $n=4$ (усього куль в урні після першого виймання);

$P(B_2/A_1)=2/4$, оскільки сприятливих виходів експерименту $m=2$ (число синіх куль після виймання однієї червоної кулі), а загальне число виходів $n=4$ (усього куль в урні після першого виймання);

$P(B_2/B_1)=1/4$, оскільки сприятливих виходів експерименту $m=1$ (число синіх куль у результаті виймання однієї синьої кулі), а загальне число виходів $n=4$ (усього куль в урні після першого виймання).

2.5. Ймовірність добутку подій



Теорема 2.2. Ймовірність добутку двох подій A і B дорівнює ймовірності події A , помноженої на умовну ймовірність події B за умови, що подія A відбулася, або дорівнює ймовірності події B , помноженої на умовну ймовірність події A за умови, що подія B відбулася:

$$P(A*B) = P(A) * P_A(B) = P(B)*P_B(A) . \quad (2.9)$$

Доведення. Нехай результати експерименту складають повну групу несумісних рівноможливих подій (рис. 2.1). При цьому

m з них сприятливі події A ;

k з них сприятливі події B ;

l з них сприятливі добутку подій $A*B$.

Тоді відповідно до класичної формули

$$P(A) = \frac{m}{n}; \quad P(B) = \frac{k}{n}; \quad P_A(B) = \frac{l}{m}; \quad P_B(A) = \frac{l}{k} .$$

Відповідно до тієї ж формули ймовірність одночасної появи подій A і B

$$P(A*B) = \frac{l}{n} .$$

Перетворимо останню рівність:

$$P(A*B) = \frac{l}{n} = \begin{cases} \frac{l}{n} = \frac{l \cdot m}{n \cdot m} = \frac{m}{n} \cdot \frac{l}{m} = P(A)P_A(B); \\ \frac{l}{n} = \frac{l \cdot k}{n \cdot k} = \frac{k}{n} \cdot \frac{l}{k} = P(B)P_B(A) . \end{cases}$$

Таким чином, теорему доведено.

Наслідок теореми 2.2. Ймовірність добутку двох незалежних подій A і B дорівнює добутку їх імовірностей

$$P(A*B) = P(A) * P(B) . \quad (2.10)$$

Наслідок досить просто пояснюється, коли взяти до уваги, що для незалежних подій умовні ймовірності збігаються з безумовними:

$$P_B(A) = P(A); P_A(B) = P(B) .$$

Наслідок легко узагальнюється на випадок кількох подій.

Приклад 2.4. Стрілець робить три постріли по мішені. Ймовірність влучення в мішень при кожному пострілі однакова і дорівнює 0,9 . Знайти ймовірність того, що в мішені буде тільки дві пробоїни.

Розв’язання. Введемо позначення:

A_1 і \bar{A}_1 – відповідно влучення і промах при першому пострілі;

A_2 і \bar{A}_2 – відповідно влучення і промах при другому пострілі;

A_3 і \bar{A}_3 – відповідно влучення і промах при третьому пострілі.

Тоді подія A , яка полягає в тому, що після трьох пострілів у мішені буде тільки дві пробоїни, може наступити у випадку, якщо стрілець або влучить при першому і другому пострілах і промахнеться при третьому, або влучить при першому і третьому пострілах і промахнеться при другому, або влучить при другому і третьому пострілах і промахнеться при першому. З урахуванням введених позначень подію A можна розкласти через прості в такий спосіб:

$$A = A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 A_3 .$$

Оскільки доданки в наведеному розкладанні відповідають несумісним подіям, то ймовірність події A дорівнюватиме сумі ймовірностей цих подій (наслідок теореми 2.1):

$$P(A) = P(A_1 A_2 \bar{A}_3) + P(A_1 \bar{A}_2 A_3) + P(\bar{A}_1 A_2 A_3) .$$

Оскільки всі постріли є незалежними між собою, то кожний доданок в останньому виразі можна подати як добуток ймовірностей простих подій (наслідок теореми 2.2)

$$P(A) = P(A_1)P(A_2)P(\bar{A}_3) + P(A_1)P(\bar{A}_2)P(A_3) + P(\bar{A}_1)P(A_2)P(A_3) .$$

Ймовірності $P(A_1)$, $P(A_2)$ і $P(A_3)$ за умовою дорівнюють 0,9 . Невідомі ймовірності $P(\bar{A}_1)$, $P(\bar{A}_2)$ і $P(\bar{A}_3)$ легко визначаються як ймовірності протилежних подій: $P(\bar{A}_1) = 1 - P(A_1)$; $P(\bar{A}_2) = 1 - P(A_2)$; $P(\bar{A}_3) = 1 - P(A_3)$. Тобто усі вони дорівнюють 0,1 .

Підставляючи ймовірності простих подій в останнє розкладання ймовірності $P(A)$, одержимо шуканий результат $P(A) = 0,243$.

Тема 3. ЗАСТОСУВАННЯ ОСНОВНИХ ТЕОРЕМ

3.1. Формула повної ймовірності

Формулу **повної ймовірності** використовують для визначення **середньої** ймовірності події A , що може відбутися тільки з однією з повної групи несумісних подій $H_i, i = 1, 2, \dots, n$. При цьому відомі **апріорні** ймовірності $P(H_i)$ подій H_i і умовні ймовірності $P(A/H_i)$ настання події A за умови, що відбулася та або інша подія H_i .

Події H_i прийнято називати *гіпотезами*, тому що середня ймовірність події A визначається в момент, коли невідомо, яка з подій H_i відбудеться і спричинить настання події A .

Сума ймовірностей гіпотез повинна дорівнювати одиниці $\sum_{i=1}^n P(H_i = 1)$.



Теорема 3.1. Якщо деяка подія A може відбутися тільки з однією з повної групи несумісних подій (гіпотез) H_i ($i = 1, 2, \dots, n$) і відома апріорна ймовірність $P(H_i)$ кожної гіпотези й умовні ймовірності $P(A/H_i)$ події A за умови, що здійснилася та або інша гіпотеза, то повна, або середня ймовірність події A визначається за формулою

$$P(A) = \sum_{i=1}^n P(H_i)P(A/H_i). \quad (3.1)$$

Доведення. Подія A може відбутися або разом з подією H_1 , або з H_2 , ... , або з H_n , тобто складна подія A може бути розкладена в такий спосіб:

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

Покажемо, що з несумісності H_i випливає факт несумісності $H_i A$.

Якщо $H_i * H_j = \emptyset$, то $H_i * A * H_j * A = H_i * H_j * A * A = (H_i * H_j) * (A * A) = \emptyset * A = \emptyset$. Звідси ймовірність події A визначається відповідно до наслідку теорема 2.1, тобто

$$P(A) = P(H_1A) + P(H_2A) + \dots + P(H_nA) .$$

Застосовуючи до кожного доданка останнього виразу *теорему 2.2*, одержимо

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) + \dots + P(H_n)P(A/H_n) = \sum_{i=1}^n P(H_i)P(A/H_i),$$

що і потрібно було довести.

Безумовно, середня ймовірність $P(A)$ настання події A повинна бути більше самої маленької з умовних ймовірностей $P(A/H_i)$ і менше найбільшої.

Приклад 3.1. Завод випускає відеомагнітофони з гарантійним терміном експлуатації один рік. Відомо, що 20% продукції буде експлуатуватися в місцевості, яка знаходиться за полярним колом, 75% – у місцевості з помірним кліматом, 5% – у пустелі. Відомі також ймовірності безвідмовної роботи відеомагнітофонів у кожному регіоні протягом гарантійного терміну: 0,9 – за полярним колом; 0,99 – у місцевості з помірним кліматом; 0,8 – у пустелі.

Треба визначити, який відсоток виробів треба випустити додатково до плану для заміни відеомагнітофонів, що вийдуть з ладу в період гарантійного терміну. При цьому вважається, що при заміні виробів останні не відмовляють.

Розв’язання. Введемо позначення:

A – безвідмовна робота відеомагнітофона;

H_1 – гіпотеза, яка полягає в тому, що виріб буде експлуатуватися за полярним колом;

H_2 – гіпотеза, яка полягає в тому, що виріб буде експлуатуватися в місцевості з помірним кліматом;

H_3 – гіпотеза, яка полягає в тому, що виріб буде експлуатуватися в пустелі.

Тоді ймовірності здійснення гіпотез, виходячи з умов прикладу, складуть:

- для гіпотези H_1 величину $P(H_1) = 20\% / 100\% = 0,2$;
- для гіпотези H_2 величину $P(H_2) = 75\% / 100\% = 0,75$;
- для гіпотези H_3 величину $P(H_3) = 5\% / 100\% = 0,05$,

а відповідні умовні ймовірності події A набудуть значень: $P(A/H_1) = 0,9$; $P(A/H_2) = 0,99$; $P(A/H_3) = 0,8$.

Додатково до плану треба випустити стільки виробів, скільки їх вийде з ладу у всіх регіонах. Шуканий додатковий відсоток виробів – це середня ймовірність відмови виробів по всіх регіонах, помножена на 100%.

Обчислимо спочатку середню ймовірність безвідмовної роботи виробу:

$$P(A) = \sum_{i=1}^3 P(H_i)P(A/H_i) = 0,2*0,9 + 0,75*0,99 + 0,05*0,8 = 0,9625 .$$

Відповідно до формули (2.7), що зв'язує протилежні події, середня ймовірність відмови виробів по всіх регіонах визначиться як

$$P(\bar{A}) = 1 - P(A) = 1 - 0,9625 = 0,0375 .$$

Шукане рішення: $P(\bar{A}) * 100\% = 3,75\%$.

3.2. Формула Байєса

Формула Байєса використовується в тих же умовах, що і формула повної ймовірності. Єдина відмінність полягає в тому, що подія A вже відбулася.

Формула Байєса дозволяє визначати *апостеріорні* ймовірності гіпотез $P(H_j/A)$, $j=1, 2, \dots, n$, тобто умовні ймовірності гіпотез за умови, що подія A відбулася.



Теорема 3.2. Якщо деяка подія A може відбутися тільки з однією з повної групи несумісних подій (гіпотез) H_i ($i = 1, 2, \dots, n$) і відомі апріорні ймовірності гіпотез $P(H_i)$, умовні ймовірності $P(A/H_i)$ події A за умови, що здійснилася та або інша гіпотеза, а також відомо, що подія A відбулася, то апостеріорна ймовірність гіпотези H_j ($j \in \{1, 2, \dots, n\}$) визначається за формулою

$$P(H_j / A) = \frac{P(H_j)P(A / H_j)}{\sum_{i=1}^n P(H_i)P(A / H_i)} . \quad (3.2)$$

Доведення. На підставі *теорему 2.2* про ймовірність добутку двох подій визначимо ймовірність одночасної появи подій A і H_j ($j \in \{1, 2, \dots, n\}$) в одному експерименті:

$$P(A * H_j) = P(A) * P(H_j / A) = P(H_j) * P(A / H_j)$$

Другу частину отриманого співвідношення, тобто рівність

$$P(A) * P(H_j / A) = P(H_j) * P(A / H_j) ,$$

розв'яжемо відносно величини $P(H_j / A)$:

$$P(H_j / A) = \frac{P(H_j)P(A / H_j)}{P(A)} . \quad (3.3)$$

У знаменнику виразу (3.3) знаходиться повна ймовірність $P(A)$, що відповідно до *теорему 3.1* визначається сумою $\sum_{i=1}^n P(H_i)P(A / H_i)$.

Підставляючи в (3.3) названу суму, остаточно одержимо $P(H_j / A) = \frac{P(H_j)P(A / H_j)}{P(A)}$. Теорему доведено.

Прикладне значення формули Байєса досить значне. Вона знаходить застосування в:

- розпізнаванні образів для виявлення об'єктів за їх розмитим зображенням,
- технічній діагностиці для пошуку несправності,
- у медичній діагностиці для постановки діагнозу хворому,
- у радіолокаційній техніці для відділення сигналу від шуму і в багатьох інших випадках, коли необхідно виявити ймовірну причину (гіпотезу) появи події.

Формула Байєса, використовуючи інформацію про факт появи події, забезпечує корекцію апіорних імовірностей гіпотез, що дозволяє більш об'єктивно судити про причину, яка спонукала подію.

Приклад 3.2. На завод-виготовник надійшла рекламація на відеомагнітофон, що відмовив. Умови експлуатації магнітофона обговорені в прикладі 3.1. Треба визначити найімовірніший регіон, в якому він експлуатувався.

Розв'язання. За умовами задачі подією, що відбулася, є відмова відеомагнітофона. Якщо залишатися в рамках позначень, що мали місце при розгляданні прикладу 3.1, то ця подія позначається \bar{A} . Середня ймовірність цієї події $P(\bar{A})$ вже обчислена і складає 0,0375.

Умовні ймовірності події \bar{A} за умови, що відбулася та або інша гіпотеза, визначаються в такий спосіб:

$$\begin{aligned} P(\bar{A}/H_1) &= 1 - P(A/H_1) = 1 - 0,9 = 0,1 ; \\ P(\bar{A}/H_2) &= 1 - P(A/H_2) = 1 - 0,99 = 0,01 ; \\ P(\bar{A}/H_3) &= 1 - P(A/H_3) = 1 - 0,8 = 0,2 . \end{aligned}$$

Апостеріорні ймовірності гіпотез про регіон експлуатації відеомагнітофона, що відмовив, відповідно до формули Байєса, визначаються в такий спосіб:

для гіпотези H_1 (експлуатація за полярним колом)

$$P(H_1/\bar{A}) = \frac{P(H_1)P(\bar{A}/H_1)}{P(\bar{A})} = \frac{0,2 * 0,1}{0,0375} = 0,533;$$

для гіпотези H_2 (експлуатація в місцевості з помірним кліматом)

$$P(H_2/\bar{A}) = \frac{P(H_2)P(\bar{A}/H_2)}{P(\bar{A})} = \frac{0,75 * 0,01}{0,0375} = 0,2;$$

для гіпотези H_3 (експлуатація в пустелі)

$$P(H_3/\bar{A}) = \frac{P(H_3)P(\bar{A}/H_3)}{P(\bar{A})} = \frac{0,05 * 0,2}{0,0375} = 0,2667.$$

Таким чином, найбільш імовірним регіоном, з якого надійшла рекламація, є місцевість за полярним колом. Дана гіпотеза має найбільшу апостеріорну ймовірність – 0,533 .

Можна також уточнювати ймовірність гіпотези, враховуючи інші наявні дані (інші події, що відбулися). Для врахування кожної наступної події потрібно за апіорну ймовірність гіпотези брати її апостеріорну ймовірність з попереднього кроку.

Тема 4. ПОВТОРЕННЯ ЕКСПЕРИМЕНТУ

4.1. Повторні незалежні випробування. Схема Бернуллі

Нехай при незмінних умовах проводяться послідовні повторні випробування, в результаті яких можуть бути два можливі наслідки (дві елементарні події A і \bar{A}): A – подія відбувається, \bar{A} – подія не відбувається. Оскільки ймовірність події A в кожному випробуванні не залежить від наслідку (результату) інших випробувань, то такі випробування називають незалежними.



Означення 4.1. Випробування (експерименти) називаються **незалежними**, якщо ймовірність появи події A в кожному експерименті не залежить від того, з'явилася вона в попередніх експериментах чи ні.

Якщо в кожному незалежному випробуванні ймовірність настання події A одна й та сама і дорівнює p ($0 \leq p \leq 1$) і не залежить від номера випробування, то такі випробування називаються **схемою Бернуллі**.

У практиці економістів і менеджерів часто виникають задачі, безпосередньо або побічно зв'язані з обчисленням ймовірності складних подій при фіксованому числі повторення *незалежних експериментів* і відомої ймовірності настання деякої події A в одному експерименті. До таких задач насамперед відносяться:

- визначення ймовірності настання події A рівно k разів у n незалежних випробуваннях;
- визначення ймовірності настання події A не менше k_1 разів і не більш k_2 разів у n незалежних випробуваннях;
- визначення найімовірнішого числа настання події A в n незалежних випробуваннях.

Всі перелічені задачі можуть бути вирішені за допомогою основних теорем теорії ймовірностей. Як приклад розглянемо одну з них.

Приклад 4.1. Нехай необхідно визначити ймовірність ураження цілі

не менше двох разів при трьох пострілах, якщо ймовірність ураження цілі при одному пострілі дорівнює p .

Ця задача є окремим випадком другої задачі з тільки що перелічених. Її можна сформулювати іншим способом: визначити ймовірність ураження мішені не менше двох разів і не більше трьох разів при трьох пострілах, якщо ймовірність ураження мішені при одному пострілі дорівнює p .

Розв'язання прикладу 4.1. Складна, подія B , ймовірність якої потрібно визначити, може бути подана як сума менш складних:

$$B = C + D,$$

де C – подія, що полягає в ураженні мішені рівно 2 рази при трьох пострілах;

D – подія, що полягає в ураженні мішені рівно 3 рази при трьох пострілах.

У свою чергу, події C і D можуть бути розкладені на прості події:

$$C = A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 A_3 ;$$

$$D = A_1 A_2 A_3 ,$$

де A_i – подія, що полягає в ураженні мішені при i -му пострілі; $i = 1, 2, 3$;

\bar{A}_i – подія, протилежна події A_i , $i = 1, 2, 3$.

Відповідно до наслідку теореми 2.1 (C і D – несумісні події),

$$P(B) = P(C) + P(D) = P(\bar{A}_1 A_2 A_3) + P(A_1 \bar{A}_2 A_3) + P(A_1 A_2 \bar{A}_3) + P(A_1 A_2 A_3).$$

Оскільки виходи $\bar{A}_1 A_2 A_3$, $A_1 \bar{A}_2 A_3$ і $A_1 A_2 \bar{A}_3$ – рівноможливі події, то:

$$P(B) = 3 * P(A_1 A_2 \bar{A}_3) + 1 * P(A_1 A_2 A_3),$$

або

$$P(B) = C_3^2 * p^2 (1-p)^{3-2} + C_3^3 * p^3 (1-p)^{3-3}.$$

Відповідно до наслідку теореми 2.2 (A_i – події незалежні, $i = 1, 2, 3$) і рівності одиниці суми ймовірностей протилежних подій, остаточно одержуємо:

$$P(B) = C_3^2 p^2 (1-p)^{3-2} + C_3^3 p^3 (1-p)^{3-3} . \quad (4.1)$$

Примітка. Шуканий результат в останньому виразі поданий, на перший погляд, у незручній формі. Однак, саме ця форма найбільше підходить для порівняння рішення, отриманого за допомогою основних теорем теорії ймовірностей, з рішенням за формулою Бернуллі, яку буде розглянуто в наступному підрозділі.

Як було сказано раніше, всі задачі на повторення незалежних експериментів можуть бути вирішені за допомогою основних теорем теорії ймо-

вірностей. Але в умовах великого числа випробувань вирішення таких задач за допомогою основних теорем стає малоефективним через великі витрати часу на обчислювальні процедури. Щоб уникнути рутинних обчислень, у теорії ймовірності розроблені спеціальні математичні способи, які й складають предмет подальшого розгляду.

4.2. Формула Бернуллі

У вирішенні останньої задачі за допомогою основних теорем теорії ймовірностей при пошуку ймовірності настання події A рівно 2 рази в трьох експериментах, тобто ймовірності $P(B)$, ми змушені були вдатися до повного перебору можливих виходів, що сприяють події B . Така процедура повного перебору виправдує себе тільки при невеликому числі випробувань. У разі великого числа випробувань, набагато ефективніше використовувати формулу Бернуллі, що призначена саме для цієї цілі.



Теорема 4.2. Якщо робиться n незалежних випробувань, у кожному з яких подія A з'являється з однаковою ймовірністю p , то ймовірність того, що в цих випробуваннях подія A відбудеться рівно k разів (байдуже, в якій послідовності) визначається за формулою

$$P_n(k) = C_n^k p^k (1 - p)^{n-k} . \quad (4.2)$$

Теорема 4.2 приводиться без доказу, оскільки він примітивний і громіздкий. Читачеві пропонується довести теорему самостійно на основі використання комбінаторної формули обчислення числа сполучень.

Формула (4.2) відома як **формула Бернуллі**.

Рекомендується використовувати формулу Бернуллі при числі випробувань, що не перевищує числа 10.

Слід також пам'ятати, що формулу (4.2) може бути використано тільки в умовах *біноміального експерименту*, тобто при виконанні наступних вимог:

- експеримент повинен складатися з фіксованого числа випробувань (задане число n);
- кожне випробування приводить або до успіху, або до невдачі (до настання події A або \bar{A});
- ймовірність успіху (невдачі) в усіх випробуваннях повинна бути однаковою;
- усі випробування повинні бути незалежними одне від одного.

Вираз (4.2) є математичною моделлю багаторазового відтворення одного й того ж експерименту, наслідком якого може бути одна з двох подій - "успіх" або "невдача". При цьому ймовірність "успіху" не залежить від

часу (номера експерименту), умови експерименту забезпечують відсутність "після події", впливу наслідку даного експерименту на наслідок кожного з решти. Це – так звана *схема Бернуллі* проведення послідовних експериментів.

Набір чисел $P_n(m)$ ($m = 0, 1, 2, \dots, n$) називається **біномним розподілом**, а саму формулу (4.2) називають біномною, оскільки її права частина є загальним членом розкладу бінома Ньютона $(p + q)^n$. Зауважимо, що події $P_n(m)$ ($m = \overline{0, 1}$) є попарно несумісні, тому

$$\sum_{m=0}^n P_n(m) = P(\Omega) = 1,$$

тобто

$$\sum_{m=0}^n C_n^m p^m q^{n-m} = (p + q)^n = 1^n = 1.$$

4.3. Локальна теорема Муавра-Лапласа

Використання формули Бернуллі при $n \gg 1$, безумовно, краще прямого використання основних теорем теорії ймовірностей для тієї ж цілі. Однак наявність у формулі (4.2) числа сполучень робить її також незручною через трудомістке обчислення факторіалів. Зазначеного обчислення можна уникнути, якщо ймовірності $P_n(k)$ замінити їх оцінками.



Теорема 4.3 (локальна теорема Муавра-Лапласа). Якщо робиться n незалежних випробувань, у кожному з яких подія A з'являється з однаковою ймовірністю p , то ймовірність того, що в цих випробуваннях подія A відбудеться рівно k разів (байдуже, у якій послідовності) може бути оцінена (тим точніше, чим більше n) за формулою

$$P_n(k) \approx \frac{1}{\sqrt{n \cdot p \cdot q}} \varphi(x), \quad (4.3)$$

де $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ – функція Гаусса;

$$x = \frac{k - np}{\sqrt{npq}}; \quad q = 1 - p.$$

Функція Гаусса має відповідні таблиці. Це дозволяє уникнути складних обчислень за формулою (4.3), замінюючи їх простим обчисленням аргументу функції x . Функція Гаусса – парна функція, тобто $\varphi(-x) = \varphi(x)$. Тому таблиця складена тільки для невід'ємних значень аргументу. Причому для діапазону значень аргументу від 0 до 4. При $|x| > 4$ функція Гаусса приймається рівною 0, тобто подія з шуканою ймовірністю вважається практично неможливою.

Приклад 4.2. Нехай ймовірність відвідування студентом будь-якої лекції з курсу "Теорія ймовірностей" дорівнює $p = 0,9$. Визначити ймовірність відвідування студентом 12 лекцій з 16 запланованих у семестрі.

Розв'язання. Відповідно до *теорему 4.3* ймовірність відвідування студентом 12 лекцій з 16 при ймовірності відвідування кожної, рівної 0,9, визначається за формулою (4.3):

$$P_{16}(12) \approx \frac{1}{\sqrt{16 \cdot 0,9 \cdot (1 - 0,9)}} \varphi(x), \quad (4.4)$$

$$\text{де } x = \frac{12 - 16 \cdot 0,9}{\sqrt{16 \cdot 0,9 \cdot 0,1}} = \frac{-2,4}{1,2} = -2.$$

З погляду на парність функції Гаусса $\varphi(-2) = \varphi(2)$. За таблицею значень функції Гауса знаходимо: $\varphi(2) = 0,0540$. Підставляючи знайдене значення функції в (4.4), остаточно маємо

$$P_{16}(12) \approx \frac{1}{\sqrt{16 \cdot 0,9 \cdot (1 - 0,9)}} \cdot 0,054 = \frac{1}{1,2} 0,054 = 0,045.$$

Таким чином, шукана ймовірність $P_{16}(12) = 0,045$.

4.4. Інтегральна теорема Муавра-Лапласа

У *прикладі 4.1* було розглянуто задачу визначення появи події не менше двох разів і не більше трьох разів у трьох експериментах за допомогою основних теорем теорії ймовірностей. Як видно з результату (4.1), його можна отримати за допомогою формули Бернуллі (4.2), застосувавши її двічі. Крім того, задачу можна приблизно вирішити і за допомогою локальної теореми Муавра-Лапласа, також застосувавши її двічі. Вирішення задач розглянутого типу при $n \gg 1$ можна значно прискорити, якщо вдаватися до *інтегральної теореми Муавра-Лапласа*.



Теорема 4.4 (інтегральна теорема Муавра-Лапласа). Якщо робиться n незалежних випробувань, у кожному з яких подія A з'являється з однаковою ймовірністю p , то ймовірність того, що в цих випробуваннях подія A відбудеться не менше k_1 разів і не більш k_2 разів (байдуже, в якій послідовності) може бути оцінена (тим точніше, чим більше n) за формулою

$$P_n(k_1, k_2) \approx \int_{x_1}^{x_2} \phi(x) dx = \Phi(x_2) - \Phi(x_1), \quad (4.5)$$

$$\text{де } x_1 = \frac{k_1 - np}{\sqrt{npq}}; \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}; \quad q = 1 - p;$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt - \text{функція Лапласа.}$$

Функція Лапласа має відповідні таблиці. Це дозволяє уникнути інтегрування функції Гауса за формулою (4.5), замінюючи його обчисленням тільки аргументів функції x_1 і x_2 . Функція Лапласа – непарна функція, тобто $\Phi(-x) = -\Phi(x)$. Тому таблиця складена тільки для невід'ємних значень аргументу. Діапазон значень аргументу в таблиці змінюється від 0 до 5. При $x > 5$ функція Лапласа приймається рівною 0,5.

Приклад 4.3. Нехай ймовірність відвідування студентом будь-якої лекції з курсу "Теорія ймовірностей" дорівнює $p = 0,9$. Визначити ймовірність відвідування студентом не менше 12 лекцій з 16 запланованих у семестрі, якщо ймовірність його появи на будь-якій лекції складає 0,9.

Розв'язання. Згідно з *теоремою 4.4* ймовірність відвідування студентом не менше 12 лекцій з 16 при ймовірності відвідування будь-якої з них 0,9, визначається за формулою (4.5):

$$P_n(k_1, k_2) = P_{16}(12, 16) \approx \Phi(x_2) - \Phi(x_1), \quad (4.6)$$

$$\begin{aligned} \text{де } x_1 &= \frac{k_1 - np}{\sqrt{npq}} = \frac{12 - 16 \cdot 0,9}{\sqrt{16 \cdot 0,9 \cdot 0,1}} = \frac{-2,4}{1,2} = -2; \\ x_2 &= \frac{k_2 - np}{\sqrt{npq}} = \frac{16 - 16 \cdot 0,9}{\sqrt{16 \cdot 0,9 \cdot 0,1}} = \frac{1,6}{1,2} = 1,33. \end{aligned}$$

З погляду на непарність функції Лапласа $\Phi(-2) = -\Phi(2)$. За таблицею значень функції Лапласа знаходимо: $\Phi(1,33) = 0,4082$, $\Phi(2) = 0,4772$. Підставляючи знайдені значення функції в (4.6), остаточно одержуємо

$$P_{16}(12, 16) = \Phi(1,33) - \Phi(-2) = \Phi(1,33) + \Phi(2) = 0,4082 + 0,4772 = 0,8854.$$

Таким чином, шукана ймовірність $P_{16}(12,16) = 0,8854$.

4.5. Теорема Пуассона

Точність наближених формул Муавра-Лапласа погіршується з наближенням одного з чисел p чи q до нуля. Тому випадок, коли p чи q є малим числом, потребує окремого розгляду. Тобто нас цікавить оцінка для ймовірності числа подій, коли подія рідко відбувається («рідкісні» події).



Теорема 4.5 (теорема Пуассона). Нехай m фіксоване, а n і p міняються, причому так, що величина $\lambda = np$ є постійною. Тоді

$$\lim_{n \rightarrow \infty} P_n(m) = \frac{\lambda^m e^{-\lambda}}{m!}, \quad (4.7)$$

де число $\lambda = np$ – середнє число успіхів.

Сукупність значень $P_n(m)$ називається *розподілом Пуассона*, ($m = 0, 1, \dots$).

Приклад 4.4. Із статистичних обстежень визначено, що ймовірність захворіти грипом під час епідемії для кожної особи складає 0,08. Яка ймовірність того, що із 100 перевірених осіб хворими виявляться рівно 10 осіб?

Розв'язання. За умовою $n=100$, $m=10$, $p=0,08$. Тому

$$\lambda = np = 8.$$

Отже,

$$P_{100}(10) = \frac{8^{10} \cdot e^{-8}}{10!} = 0,0993.$$

4.6. Найімовірніше число настання подій

Нехай робиться n незалежних експериментів, у кожному з яких може наступити подія A з однаковою ймовірністю p . У проведених експериментах подія A може не наступити жодного разу ($k=0$), один раз ($k=1$), два рази ($k=2$) і т.д. За допомогою формули Бернуллі при невеликих n можна визначити (або за допомогою локальної теореми Муавра-Лапласа при великих n можна оцінити) ймовірність появи події A в n незалежних експериментах рівно k разів. У табл. 4.1 наведені можливі значення величини k і відповідні ймовірності $P_n(k)$.

Таблиця 4.1

k	0	1	...	k_0	...	n
$P_n(k)$	$P_n(0)$	$P_n(1)$...	$P_n(k_0)$...	$P_n(k)$

Серед множини чисел $k, k \in \{0, n\}$, є, принаймні, одне число k_0 , якому відповідає максимальна ймовірність $P_n(k)$.



Означення 4.2. Найімовірнішим числом настання події A в n незалежних експериментах при однаковій імовірності настання події A в кожному з них називається число k_0 , якому відповідає максимальна ймовірність $P_n(k)$, тобто число $k_0 = \arg \left(\max_{k=1, n} \{P(k)\} \right)$.

На практиці часто виникає необхідність визначити число k_0 . Щоб щоразу при визначенні цього числа не будувати таблицю, аналогічну табл. 4.1, слід користуватися спеціально для цього призначеною подвійною нерівністю

$$np - q \leq k_0 \leq np + p, \quad (4.8)$$

де p – ймовірність настання події A при одному експерименті; $q = (1 - p)$.

Число k_0 називають також *модую*.

4.7. Теорема Бернуллі (закон великих чисел Бернуллі)

До цього часу ми вважали, що ймовірність p в схемі Бернуллі відома. У той же час у реальних ситуаціях виникає обернена задача: як визначити невідому ймовірність деякої події?

Обґрунтований підхід для розв'язання цієї останньої задачі дається якраз законом великих чисел Бернуллі

Якщо проводиться n незалежних випробувань, у кожному з яких випадкова подія A з'являється з однаковою ймовірністю $P(A) = p$, то відносна частота $\frac{\mu}{n}$ появи події A (μ – число появ A) при великому n приблизно дорівнює ймовірності p :

$$\frac{\mu}{n} \approx p. \quad (4.9)$$

Наведене твердження є формальним змістом *теорему Бернуллі*, відомої як *закон великих чисел*.



Теорема 4.6 (теорема Бернуллі). З ймовірністю, як завгодно близькою до 1, очікується, що при досить великій кількості випробувань відносна частота появи події буде як завгодно мало відрізнятися від її ймовірності.

Зауважимо, що теорема не стверджує, що співвідношення (4.9) вірогідно, однак, якщо число n досить велике, то ймовірність виконання (4.9) близька до 1 (наприклад, 0.98 або 0.999), що **практично вірогідно**.

Іншими словами, якщо проводиться експеримент, що складається з досить великого числа n випробувань, то можна бути впевненим, що співвідношення (4.9) буде виконано.

Примітка. Автори рекомендують читачеві перевірити останнє твердження за допомогою експерименту з киданням монети (подія A – випадіння «орла») або киданням гральної кістки (подія A – випадіння парного числа).

Як користуватися подвійною нерівністю (4.8) ?

Насамперед, відзначимо особливості цієї нерівності:


- ◆ значення правої частини перевищує значення лівої рівно на одиницю;
- ◆ k_0 – ціле число;
- ◆ всередині діапазону значень $[np-q; np+p]$ може знаходитися тільки одне ціле число, або два цілих числа можуть знаходитися на його границях.

Визначення k_0 здійснюють у наступній послідовності:

- спочатку визначають величину np , якщо np – ціле число, то $k_0 = np$;
- потім визначають величину $np+p$,
 - якщо $(np+p)$ – ціле число, то існує два найімовірніших числа: $k_{01} = np+p$ і $k_{02} = k_{01}-1$;
 - якщо $(np+p)$ – неціле число, то k_0 – ціле число в діапазоні $[np-q; np+p]$.

Приклад 4.5. Яке найімовірніше число лекцій, що відвідав студент з $n=17$ запланованих у семестрі, якщо ймовірність відвідування кожної лекції $p=0,8$?

Розв’язання. Обчислюємо $np = 17 \cdot 0,8 = 13,6$. Оскільки np – неціле число, то відповідно до вище викладеної методики використання подвійної нерівності (4.8) обчислюємо $np+p = 13,6+0,8 = 14,4$. Оскільки $(np+p)$ – неціле число, то відповідно до тієї ж методики k_0 належить діапазону $[np-q; np+p]$, або дорівнює цілій частці величини $np+p$, тобто $k_0 = 14$.

Приклад 4.6. Яке найімовірніше число випадіння 6 очок в умовах експерименту  , якщо останній робиться 11 разів ?

Розв’язання. За умовою прикладу загальне число експериментів $n = 11$, ймовірність настання події в одному експерименті (випадіння шести очок) $p = \frac{1}{6}$. Як і в попередньому прикладі, вирішення здійснюємо відповідно до вище викладеної методики використання подвійної нерівності

(4.8). Тобто спочатку обчислюємо $np = 11 * \frac{1}{6} = \frac{11}{6}$. Далі, оскільки np – неціле число, обчислюємо $np+p = \frac{11}{6} + \frac{1}{6} = 2$. Оскільки $(np+p)$ – ціле число, то існують 2 найімовірніших числа: $k_{01} = np+p = 2$ і $k_{02} = np-q = k_{01}-1 = 1$.

4.8. Номер першого успішного випробування в схемі Бернуллі

Розглянемо схему Бернуллі з імовірністю успіху p в одному випробуванні. Випробування проводяться до появи першого успіху. Введемо величину τ , що дорівнює номеру першого успішного випробування.



Теорема 4.7. Можливість того, що перший успіх відбудеться у випробуванні з номером k ($k \in \mathbb{N} = \{1, 2, 3, \dots\}$), дорівнює

$$P(\tau = k) = pq^{k-1}, \quad (4.10)$$

де $q = 1 - p$.

Теорема 4.7 приводиться без доказу, оскільки він примітивний.

Набір чисел $\{pq^{k-1}, k = 1, 2, 3, \dots\}$ називається геометричним розподілом імовірностей.

Тема 5. ВИПАДКОВІ ВЕЛИЧИНИ

5.1. Основні означення

"Випадкові величини" – це традиційно друга велика тема теорії ймовірностей, що ще більше наближає нас до світу випадкових явищ і процесів



Означення 5.1. *Випадковою величиною* називають величину, яка в результаті експерименту приймає заздалегідь невідоме значення.

Приклади випадкових величин:

- кількість студентів, які присутні на лекції;
- кількість сонячних днів у році;
- вага осколка снаряда, що розірвався;

- час очікування громадського транспорту на зупинці;
- температура навколишнього середовища.

Випадкові величини за типом простору можливих значень діляться на *дискретні й неперервні*.



Означення 5.2. *Дискретною* називають випадкову величину, можливі значення якої належать до ліченої множини – скінченної або нескінченної.



Означення 5.3. *Неперервною* називають випадкову величину, можливі значення якої належать до неперервної множини – обмеженої або необмеженої.

Перші два приклади величин з наведених вище прикладів (кількість студентів і кількість сонячних днів) відносяться до дискретних випадкових величин, а три наступні – до неперервних. При цьому дискретні величини – скінчені й цілочислові: перша змінюється від нуля до загальної кількості студентів на потоці, друга від 0 до 366 у високосному році або від 0 до 365 у не високосному році. Перші дві неперервні величини (вага осколка і час очікування) обмежені з двох боків: вага осколка обмежена знизу нулем, а зверху вагою цілого снаряда, час очікування – знизу нулем, а зверху значенням максимального інтервалу часу, за яким рухається транспорт. Третя неперервна величина (температура навколишнього середовища) обмежена тільки знизу значенням абсолютного нуля $-273,2^{\circ}\text{C}$.

Для характеристики випадкової події досить знати ймовірність її настання в результаті експерименту. Для характеристики ж випадкової величини, що має більше двох можливих значень, цього недостатньо. Щоб мати вичерпну характеристику випадкової величини, треба знати її *закон розподілу*.



Означення 5.4. *Закон розподілу* випадкової величини – це співвідношення, що встановлює зв'язок між можливими значеннями випадкової величини і відповідними ймовірностями.

5.2. Форми задання дискретних випадкових величин

У теорії ймовірностей розрізняють декілька форм задання закону розподілу дискретної випадкової величини. На практиці використовують тільки дві найбільш корисні:

- ряд розподілу;
- інтегральна функція розподілу.

Розглянемо послідовно кожен форму задання закону розподілу дискретної випадкової величини.

5.2.1. Ряд розподілу

Ряд розподілу є найбільш простою і зрозумілою формою задання закону розподілу дискретної випадкової величини. Він являє собою таблицю, що складається з двох рядків. У першому рядку розташовуються в порядку зростання всі можливі значення дискретної випадкової величини, у другому – відповідні ймовірності.

Загальний вид ряду розподілу відповідає табл. 5.1.

Таблиця 5.1 – Закон розподілу у вигляді ряду розподілу

x_i	x_1	x_2	\dots	x_n
p_i	p_1	p_2	\dots	p_n

У результаті експерименту випадкова дискретна величина повинна прийняти одне з можливих значень. Оскільки всі можливі значення можна розглядати як повну групу несумісних подій, то сума відповідних ймовірностей обов'язково повинна дорівнювати одиниці (умова нормування)

$$\sum_{i=1}^n p_i = 1. \quad (5.1)$$

5.2.2. Інтегральна функція розподілу та її властивості



Означення 5.5. Інтегральна функція розподілу випадкової величини X – це функція $F(x)$, що при кожному значенні свого аргументу x чисельно дорівнює ймовірності того, що випадкова величина X опиниться менше, ніж значення аргументу x , тобто

$$F(x) = P \{ X < x \} .$$

Інтегральна функція має три *властивості*:

- *1-а властивість.* Інтегральна функція від аргументу "мінус нескінченність" дорівнює нулю: $F(-\infty) = 0$.
- *2-а властивість.* Інтегральна функція від аргументу "плюс нескінченність" дорівнює одиниці: $F(\infty) = 1$.
- *3-я властивість.* Інтегральна функція – функція, що не зменшується: якщо $x_2 > x_1$, то $F(x_2) \geq F(x_1)$.

5.2.3. Приклад побудови закону розподілу

Розглянемо побудову закону розподілу у вигляді ряду розподілу й у вигляді інтегральної функції на конкретному прикладі.

Приклад 5.1. Побудувати закон розподілу випадкової величини X – кількість будинків, які здано в експлуатацію в запланований термін, з 3, що будуються. Ймовірність здачі в запланований термін для кожного будинку однакова і дорівнює 0,9.

Розв’язання. Випадкова величина X – кількість будинків, зданих в експлуатацію в запланований термін, може приймати значення 0, 1, 2 або 3. За умовами прикладу загальна кількість будинків (кількість експериментів) $n = 3$, ймовірність побудувати кожний будинок у запланований термін (настання події A в одному експерименті) $p = 0,9$. Тоді ймовірності p_i ($i = 0, 1, 2, 3$) – це ймовірності того, що з 3 будинків у запланований термін буде здано рівно 0, 1, 2 або 3. Вони легко визначаються за допомогою формули Бернуллі:

$$p_0 = P(X = 0) = P_3(0) = C_3^0 p^0 (1-p)^{3-0} = 1 * 0,9^0 * (1-0,9)^{3-0} = 0,001 ;$$

$$p_1 = P(X = 1) = P_3(1) = C_3^1 p^1 (1-p)^{3-1} = 1 * 0,9^1 * (1-0,9)^{3-1} = 0,027 ;$$

$$p_2 = P(X = 2) = P_3(2) = C_3^2 p^2 (1-p)^{3-2} = 1 * 0,9^2 * (1-0,9)^{3-2} = 0,243 ;$$

$$p_3 = P(X = 3) = P_3(3) = C_3^3 p^3 (1-p)^{3-3} = 1 * 0,9^3 * (1-0,9)^{3-3} = 0,729 .$$

Отримані ймовірності дозволяють сформувати ряд розподілу (табл. 5.2) і побудувати інтегральну функцію випадкової величини X (рис. 5.1).

Таблиця 5.2 – Ряд розподілу випадкової величини X в умовах прикладу 5.1

x_i	0	1	2	3
p_i	0,001	0,027	0,243	0,729

Слід звернути увагу на рівність одиниці суми значень імовірностей у другому рядку ряду розподілу. Така рівність обов'язкова. Вона знаходиться в повній відповідності з формулою (5.1) і служить критерієм слушності побудови закону розподілу.

Ряд розподілу випадкової величини з погляду на її табличну природу не має наочності. Крім того, за рядом розподілу важко визначати і порівняти ймовірності влучення випадкової величини в заданий діапазон значень. Цих недоліків позбавлена інтегральна функція розподілу випадкової величини, яка подана у вигляді графіка.

Перед тим як перейти до побудови графіка інтегральної функції, слід спочатку одержати її аналітичний запис, що складається з часткових записів для кожного з $(n+1)$ -го діапазону, на які розбивається нескінченна числова ось можливими значеннями випадкової величини X , де n – загальна кількість можливих значень випадкової величини. В умовах прикладу $n = 4$ (можливі значення: 0, 1, 2 або 3). Отже кількість діапазонів дорівнює 5.

Аналітичний запис інтегральної функції випадкової величини X представимо у вигляді таблиці.

Таблиця 5.3 – таблично-аналітичне подання інтегральної функції розподілу випадкової величини X в умовах прикладу 5.1

Індекс діапазону i	Діапазон $x^{(i)}$	Значення інтегральної функції $F(x^{(i)})$
0	$x^{(0)} \leq 0$	$F(x^{(0)}) = P\{X < x^{(0)}\} = 0$
1	$0 \leq x^{(1)} \leq 1$	$F(x^{(1)}) = P\{X < x^{(1)}\} = P(X=0) = 0,001$
2	$1 \leq x^{(2)} < 2$	$F(x^{(2)}) = P\{X < x^{(2)}\} = P(X=0) + P(X=1) =$ $= 0,001 + 0,027 = 0,028$
3	$2 \leq x^{(3)} < 3$	$F(x^{(3)}) = P\{X < x^{(3)}\} = P(X=0) + P(X=1) + P(X=2) =$ $= 0,001 + 0,027 + 0,243 = 0,271$
4	$x^{(4)} \geq 3$	$F(x^{(4)}) = P\{X < x^{(4)}\} =$ $P(X=0) + P(X=1) + P(X=2) + P(X=3) =$ $= 0,001 + 0,027 + 0,243 + 0,729 = 1$

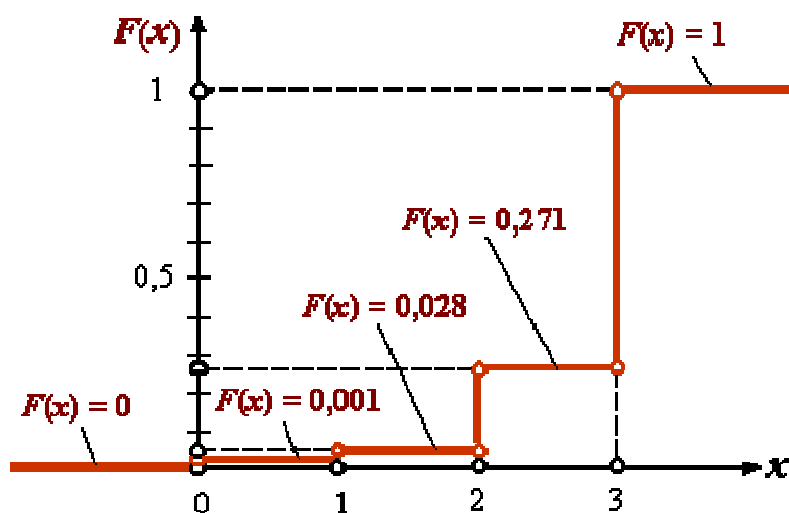


Рис. 5.1 – Графік інтегральної функції розподілу

На рис. 5.1 зображений графік інтегральної функції розподілу згідно з її таблично-аналітичним записом (табл. 5.3). Як видно з рисунку, графіком функції дискретної випадкової величини є сходи́ста неперервна лінія, визначена

на всій числовій осі x . Функція $F(x)$ змінюється від 0 до 1. Вона або зберігає своє значення в кожному i -му діапазоні зміни аргументу x , або стрибкоподібно збільшується в точках, що відповідають можливим значенням дискретної випадкової величини, тобто в точках, що розділяють діапазони.

5.2.4. Ймовірність влучення випадкової величини в заданий діапазон

На практиці при дослідженні випадкових величин часто виникає задача визначення ймовірності влучення значень деякої випадкової величини X у заданий діапазон $[a, b)$, тобто ймовірності $P\{a \leq X < b\}$. Така ймовірність легко визначається за допомогою інтегральної функції.

Введемо позначення:

A – подія, яка полягає в тому, що $X < a$;

B – подія, яка полягає в тому, що $X < b$;

C – подія, яка полягає в тому, що $a \leq X < b$.

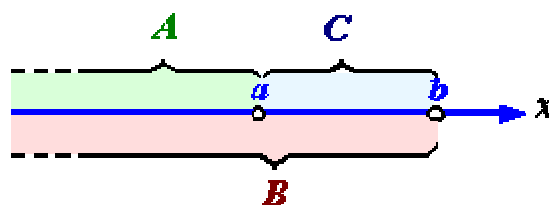


Рис. 5.2

Складна випадкова подія B являє собою суму подій A і C (див. рис. 5.2):

$$B = A + C.$$

Оскільки події A і C є несумісними, то

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

звідки

$$P(C) = P(B) - P(A) = P\{X < b\} - P\{X < a\}.$$

За означенням інтегральної функції $P\{X < b\} = F(b)$, $P\{X < a\} = F(a)$. Отже,

$$P(C) = F(b) - F(a).$$

Таким чином, ймовірність влучення випадкової величини в заданий діапазон визначається за формулою

$$P\{a \leq X < b\} = F(b) - F(a). \quad (5.2)$$

Приклад 5.2. В умовах прикладу 5.1 визначити ймовірність попадання випадкової величини X в діапазон $[2,5, 3,5)$, тобто ймовірність $P\{2,5 \leq X < 3,5\}$.

Розв'язання. У даному випадку ліва границя діапазону $a = 2,5$, а права $b = 3,5$. Підставляючи у формулу (5.2) значення аргументу інтегральної функції і обчислюючи значення інтегральної функції на границях заданого діапазону, одержуємо шуканий результат:

$$P\{2,5 \leq X < 3,5\} = F(b) - F(a) = F(3,5) - F(2,5) = 1 - 0,271 = 0,729.$$

Шукана ймовірність і значення інтегральної функції $F(x)$ в умовах прикладу легко визначаються за графіком інтегральної функції (див. рис. 5.3).

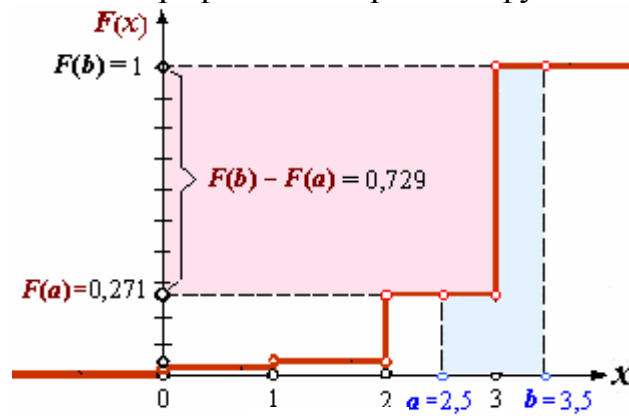


Рис. 5.3

5.3. Форми задання неперервної випадкової величини

У теорії ймовірностей розглядаються дві форми задання закону розподілу неперервної випадкової величини.

- інтегральна функція розподілу ймовірності;
- щільність розподілу ймовірності.

Обидві форми абсолютно рівноправні. Перша характеризує розподіл ймовірностей залежно від діапазону значень неперервної випадкової величини, а друга – від конкретних значень.

Розглянемо послідовно кожну форму задання неперервної випадкової величини.

5.3.1. Інтегральна функція розподілу

Інтегральна функція розподілу ймовірності – це універсальна форма задання випадкових величин. З її допомогою можна задати закон розподілу як дискретної випадкової величини, так і неперервної.

Інтегральну функцію неперервної випадкової величини легко представити як графік інтегральної функції довільної дискретної випадкової величини, в якій кількість дискретних значень прагне до нескінченності. На рис. 5.4 показано умовний процес перетворення інтегральної функції дискретної випадкової величини в інтегральну функцію неперервної величини при послідовному подрібненні діапазонів завдання функції навпіл, тобто при збільшенні кількості значень дискретної величини в 2 рази.

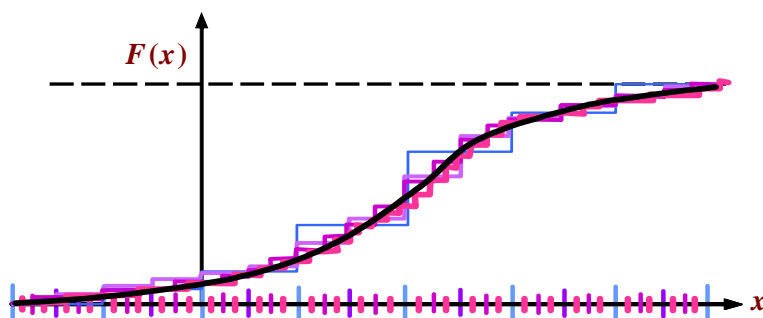


Рис. 5.4

Інтегральна функція неперервної випадкової величини зберігає всі властивості інтегральної функції дискретної випадкової величини. Нагадаємо їх:

1-а властивість. Інтегральна функція від аргументу "мінус нескінченність" дорівнює нулю: $F(-\infty) = 0$.

2-а властивість. Інтегральна функція від аргументу "плюс нескінченність" дорівнює одиниці: $F(\infty) = 1$.

3-я властивість. Інтегральна функція – функція, що не зменшується: якщо $x_2 > x_1$, то $F(x_2) \geq F(x_1)$.

Для інтегральної функції неперервної випадкової величини також справедлива формула (5.2), що дозволяє обчислювати ймовірність влучення значень випадкової величини в заданий діапазон $[a, b]$. На рис. 5.5 подана графічна інтерпретація процесу визначення ймовірності $P\{a \leq X < b\}$.

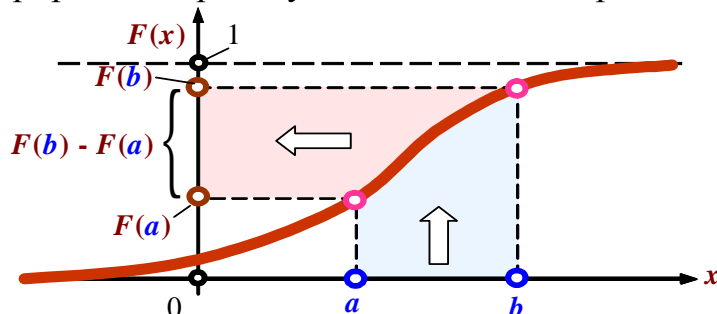


Рис. 5.5

5.3.2. Ймовірність конкретного значення неперервної випадкової величини

Ймовірність будь-якого конкретного значення a неперервної випадкової величини X можна встановити за допомогою формули (5.2) – як імовірність попадання X в деякий діапазон $[a, b]$ при правій границі b , що прямує до значення a :

$$P(X = a) = \lim_{b \rightarrow a} P\{a \leq X < b\} = \lim_{b \rightarrow a} [F(b) - F(a)] = F(a) - F(a) = 0,$$

тобто

$$P(X=a) = 0. \quad (5.3)$$

Рівність імовірності конкретного значення неперервної випадкової величини нулю (5.3) робить інтегральну функцію $F(x)$ безпорадною для характеристики окремих значень неперервної випадкової величини. Щоб дослідники могли порівнювати окремі значення неперервної випадкової величини з погляду їх імовірнісної появи в результаті експерименту, використовується інша форма задання закону розподілу цих величин – *щільність розподілу ймовірності*.

5.3.3. Щільність розподілу ймовірності

Інтегральна функція розподілу ймовірності за означенням 5.5 дорівнює ймовірності влучення значень випадкової величини X в діапазон значень від мінус нескінченності до аргументу x , тобто інтегральна функція характеризує нескінченний інтервал значень $(-\infty, x)$. Інтегральна функція, відповідно до формули (5.2), також дозволяє встановити ймовірність влучення значень випадкової величини в заданий діапазон $[a, b]$, довільно обраний на числовій осі. Однак поряд з функцією, що характеризує діапазони значень неперервної випадкової величини, становить інтерес і функція, що здатна характеризувати кожне значення неперервної величини. Такою функцією є щільність розподілу ймовірності.

Функція щільності розподілу ймовірності $f(x)$ являє собою відношення ймовірності влучення неперервної випадкової величини в малий діапазон $[x, x + \Delta x]$, до довжини цього діапазону Δx :

$$\begin{aligned} f(x) &= \lim_{\Delta x \rightarrow 0} \frac{P\{X \in [x, x + \Delta x]\}}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{P\{x \leq X < (x + \Delta x)\}}{\Delta x} = \\ &= \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{dF}{dx}. \end{aligned}$$



Означення 5.6. Щільністю розподілу ймовірності неперервної випадкової величини називається функція $f(x)$, що є першою похідною від інтегральної функції розподілу ймовірності $F(x)$:

$$f(x) = F'(x).$$

Вираз (5.4) дозволяє при відомій інтегральній функції розподілу $F(x)$ неперервної випадкової неперервної величини визначити функцію щільності розподілу $f(x)$. Не менш важливим є зворотне перетворення

$$F(x) = \int_{-\infty}^x f(t) dt, \quad (5.5)$$

яке дозволяє при відомій функції щільності розподілу $f(t)$ одержати інтегральну функцію розподілу $F(x)$.

5.3.4. Властивості щільності розподілу ймовірності

Щільність розподілу $f(x)$ неперервної випадкової величини успадковує усі властивості інтегральної функції розподілу $F(x)$. При цьому дві перші властивості інтегральної функції $F(x)$ трансформуються в одну властивість щільності розподілу.

1-а властивість. Інтеграл у нескінченних границях від щільності розподілу дорівнює одиниці (умова нормування):

$$\int_{-\infty}^{\infty} f(x)dx = 1. \quad (5.6)$$

Обґрунтування властивості (5.6) одержують шляхом узяття визначеного інтеграла від щільності розподілу:

$$\int_{-\infty}^{\infty} f(x)dx = F(x) \Big|_{-\infty}^{\infty} = F(\infty) - F(-\infty) = 1 - 0 = 1.$$

Тут $F(-\infty) = 0$ є перша властивість інтегральної функції, а $F(\infty) = 1$ – друга.

Геометричний зміст рівності (5.6) полягає в рівності одиниці площі, обмеженої графіком функції $f(x)$ і віссю абсцис.

2-а властивість. Щільність розподілу – функція невід’ємна:

$$f(x) \geq 0. \quad (5.7)$$

Така властивість щільності розподілу ймовірності впливає з третьої властивості інтегральної функції: похідна від функції, що не зменшується, не може бути від’ємною.

5.3.5. Ймовірність влучення неперервної випадкової величини в заданий діапазон

Ймовірність влучення неперервної випадкової величини в заданий діапазон $[a, b)$ може бути визначена за універсальною формулою (5.2):

$$P\{a \leq X < b\} = F(b) - F(a).$$

Альтернативну формулу для визначення цієї ж ймовірності можна одержати з (5.2) за допомогою зворотного перетворення (5.5):

$$\begin{aligned} P\{a \leq X < b\} &= F(b) - F(a) = \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx = \\ &= \int_{-\infty}^a f(x)dx + \int_a^b f(x)dx - \int_{-\infty}^a f(x)dx. \end{aligned}$$

Після скорочення подібних членів в останньому виразі остаточно одержуємо

$$P\{a \leq X < b\} = \int_a^b f(x)dx. \quad (5.8)$$

На рис. 5.6 подана графічна інтерпретація ймовірності влучення неперервної випадкової величини в діапазон $[a, b)$. Чисельно значення такої ймовірності дорівнює заштрихованій площі.

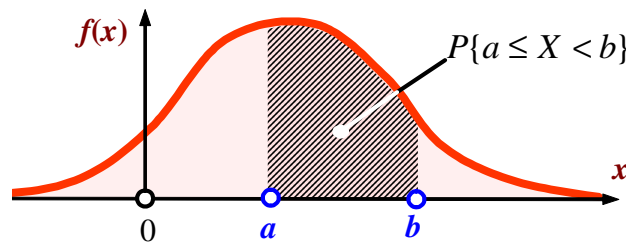


Рис. 5.6

Тема 6. ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВИХ ВЕЛИЧИН

Числові характеристики випадкових величин кількісно визначають різноманітні властивості випадкових величин. Вони дозволяють проводити порівняльний аналіз випадкових величин, давати оцінку очікуваним результатам експерименту, знаходити зв'язок і визначати залежність між різними випадковими величинами і багато чого іншого. Досить часто знання числових характеристик дає дослідникам можливість розв'язувати задачі з випадковими величинами, не знаючи закону їх розподілу. Більше того, для багатьох стохастичних задач метою їх розв'язання є визначення тієї або іншої числової характеристики.

Числові характеристики випадкових величин – це не випадкові величини. Кожна числова характеристика має тільки одне значення, що не залежить ні від результату конкретного експерименту, ні від кількості проведених експериментів.

Найбільш важливі числові характеристики є предметом розгляду цього підрозділу.

6.1. Характеристики положення випадкової величини на числовій осі

До числових характеристик положення випадкової величини на числовій осі відносяться:

- математичне сподівання;

- мода;
- медіана.

6.1.1. Математичне сподівання

Математичне сподівання випадкової величини є найбільш важливою її числовою характеристикою. Більша частина всіх числових характеристик випадкової величини безпосередньо зв'язана з її математичним сподіванням.

Математичне сподівання випадкової величини будемо позначати m_x або $M[X]$. Обидва позначення рівноправні. Надалі будемо користуватися обома позначеннями.



Означення 6.1. Математичне сподівання – це середньозважене за ймовірностями значення випадкової величини.

Математичне сподівання характеризує зміщення значень випадкової величини на числовій осі x відносно початку координат.

Математичне сподівання *дискретної* випадкової величини визначається за формулою

$$m_x = \sum_{i=1}^n x_i p_i, \quad (6.1)$$

де n – загальна кількість можливих значень випадкової величини; x_i і p_i – i -те можливе значення дискретної випадкової величини і відповідна ймовірність, $i = \overline{1, n}$.

Математичне сподівання *неперервної* випадкової величини визначається за формулою

$$m_x = \int_{-\infty}^{\infty} x \cdot f(x) dx, \quad (6.2)$$

де x – неперервна випадкова величина; $f(x)$ – щільність розподілу величини x .

На рис. 6.1 показані дві неперервні випадкові величини, задані у вигляді щільності розподілу з різними математичними сподіваннями ($m_{x2} > m_{x1}$).

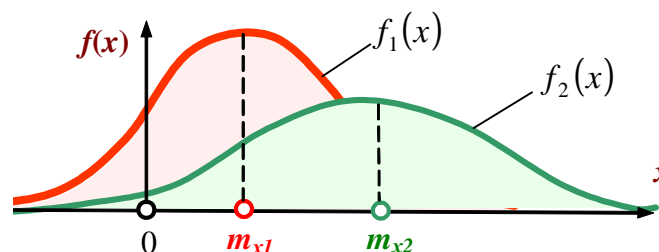


Рис. 6.1

6.1.2. Мода

Для позначення найбільш імовірного значення випадкової величини використовують так звану *моду*. Позначається мода випадкової величини символом t або M . Надалі будемо користуватися позначенням t .



Означення 6.2. Модою називають найбільш імовірне значення випадкової величини.

Мода t дискретної випадкової величини дорівнює такому її значенню x_m , якому відповідає максимальна ймовірність $p_m = \max_{i=1,n} \{p_i\}$.

Мода t неперервної випадкової величини дорівнює такому значенню аргументу x_m функції щільності розподілу $f(x)$, при якому $f(x_m) = \max_{x \in R^1} f(x)$.

На рис. 6.2 показана мода неперервної унімодальної (з однією модою) випадкової величини, заданої щільністю розподілу.

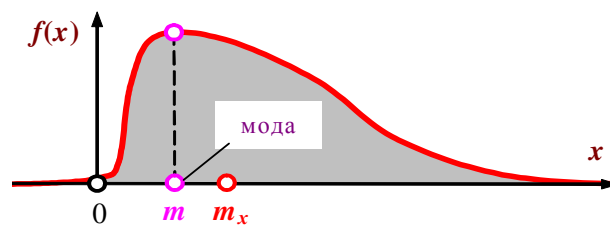


Рис. 6.2

Крім унімодальних розподілів випадкових величин, розрізняють *полімодальні* (рис. 6.3,а), *антимодальні* (рис. 6.3,б) й *безмодальні* (рис. 6.3,в).

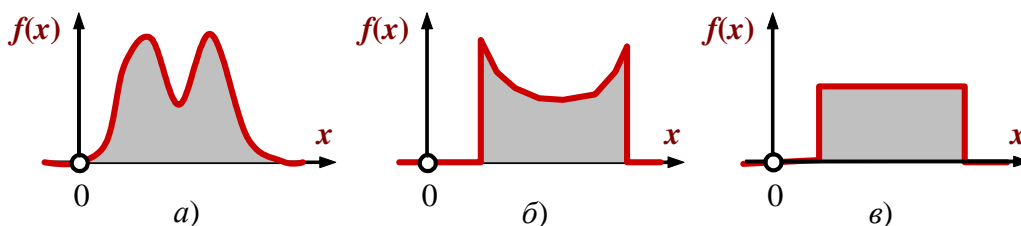


Рис. 6.3

6.1.3. Медіана

Неперервні випадкові величини, крім математичного сподівання m_x і моди t , мають ще одну характеристику положення на числовій осі – *медіану*. Ця характеристика позначається як Me .



Означення 6.3. Медіаною називають таке значення Me випадкової величини, для якого справедлива рівність $P\{X < Me\} = P\{X > Me\}$.

Перпендикуляр до числової осі, що проходить через медіану, поділяє площу, обмежену графіком щільності розподілу $f(x)$ і числовою віссю x , на дві рівні частини по 0,5 (рис. 6.4).

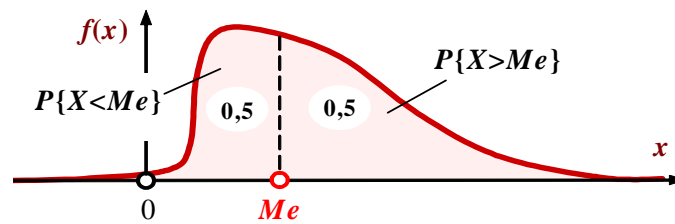


Рис. 6.4

Для симетричного унімодального закону розподілу випадкової величини значення математичного сподівання, моди і медіани збігаються.

6.2. Моменти випадкових величин

Для характеристики різних властивостей випадкових величин використовують початкові й центральні моменти

6.2.1. Початкові моменти



Означення 6.4. Початковим моментом k -го порядку α_k називають математичне сподівання k -го степеня випадкової величини $\alpha_k = M[X^k]$.

Для дискретної випадкової величини k -й початковий момент визначається за формулою

$$\alpha_k = \sum_{i=1}^n x_i^k p_i. \quad (6.3)$$

Для неперервної випадкової величини k -й початковий момент визначається за формулою

$$\alpha_k = \int_{-\infty}^{\infty} x^k f(x) dx. \quad (6.4)$$

6.2.2. Центральні моменти



Означення 6.5. Відхилення випадкової величини від математичного сподівання $(X - m_x)$ називають **центрованою випадковою величиною**.



Означення 6.6. *Центральним моментом s -го порядку μ_s* називають математичне сподівання s -го степеня центрованої випадкової величини

$$\mu_s = M[(X - m_x)^s].$$

Для дискретної випадкової величини s -й центральний момент визначається за формулою

$$\mu_s = \sum_{i=1}^n (x_i - m_x)^s p_i. \quad (6.5)$$

Для неперервної випадкової величини s -й центральний момент визначається за формулою

$$\mu_s = \int_{-\infty}^{\infty} (x - m_x)^s f(x) dx. \quad (6.6)$$

6.3. Властивості моментів випадкових величин

Особливої уваги заслуговують властивості початкових і центральних моментів першого і другого порядків. Розглянемо кожну з цих властивостей.

6.3.1. Перший початковий момент

Початковий момент 1-го порядку α_1 випадкової величини являє собою її математичне сподівання: $\alpha_1 = M[X^1] = m_x$.

6.3.2. Перший центральний момент

Центральний момент 1-го порядку μ_1 будь-якої випадкової величини завжди дорівнює нулю. Наступні перетворення першого початкового моменту дискретної випадкової величини підтверджують сказане:

$$\begin{aligned} \mu_1 &= M[(X - m_x)^1] = \sum_{i=1}^n (x_i - m_x) p_i = \sum_{i=1}^n x_i p_i - \sum_{i=1}^n m_x p_i = \\ &= m_x - m_x \sum_{i=1}^n p_i = m_x - m_x = 0. \end{aligned}$$

Аналогічний результат дають перетворення для неперервної випадкової величини:

$$\begin{aligned} \mu_1 &= M[(X - m_x)^1] = \int_{-\infty}^{\infty} (x - m_x) f(x) dx = \int_{-\infty}^{\infty} x f(x) dx - \int_{-\infty}^{\infty} m_x f(x) dx = \\ &= m_x - m_x \int_{-\infty}^{\infty} f(x) dx = m_x - m_x = 0. \end{aligned}$$

Перший центральний момент на практиці не використовується, оскільки нічого характеризувати не може.

6.3.3. Другий початковий момент

Початковий момент 2-го порядку α_2 випадковій величині характеризує ступінь відхилення випадкової величини навколо її математичного сподівання, а також зміщення випадкової величини на числовій осі відносно початку координат.

Другий початковий момент дискретної випадкової величини визначається за формулою

$$\alpha_2 = M[X^2] = \sum_{i=1}^n x_i^2 p_i . \quad (6.7)$$

Другий початковий момент неперервної випадкової величини визначається за формулою

$$\alpha_2 = M[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx . \quad (6.8)$$

З погляду на те, що другий початковий момент характеризує відразу дві властивості випадкової величини, він як самостійна числова характеристика не використовується. Проте, він має велике значення для визначення інших числових характеристик, про що мова йтиме в підрозділі 6.3.5.

6.3.4. Другий центральний момент

Центральний момент 2-го порядку μ_2 характеризує ступінь відхилення випадкової величини навколо її математичного сподівання. Ця числова характеристика має ще іншу назву – *дисперсія*. Дисперсія – більш поширений термін, що позначається як D_x .

Оскільки дисперсія характеризує тільки відхилення випадкової величини, вона в порівнянні з другим початковим моментом є більш важливою числовою характеристикою.

Дисперсія дискретної випадкової величини обчислюється за формулою

$$D_x = \mu_2 = M[(X - m_x)^2] = \sum_{i=1}^n (x_i - m_x)^2 p_i . \quad (6.9)$$

Дисперсія неперервної випадкової величини обчислюється за формулою

$$D_x = \mu_2 = M[(X - m_x)^2] = \int_{-\infty}^{\infty} (x - m_x)^2 f(x) dx . \quad (6.10)$$

На рис. 6.5 показані дві неперервні випадкові величини, задані у вигляді щільності розподілу з однаковими математичними сподіваннями ($m_{x2} = m_{x1} = m_x$) і різними дисперсіями ($D_{x2} > D_{x1}$).

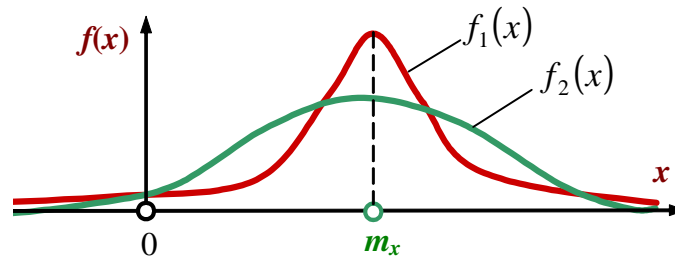


Рис. 6.5

6.3.5. Зв'язок дисперсії з початковими моментами

Визначення дисперсії D_x для неперервних випадкових величин зв'язано з трудомісткими обчисленнями визначених інтегралів. На практиці дисперсію обчислюють за допомогою другого початкового моменту α_2 і математичного сподівання (першого початкового моменту) m_x . Наступні математичні перетворення встановлюють зв'язок дисперсії з початковими моментами:

$$D_x = M[(X - m_x)^2] = \int_{-\infty}^{\infty} (x - m_x)^2 f(x) dx =$$

$$= \int_{-\infty}^{\infty} x^2 f(x) dx - 2m_x \int_{-\infty}^{\infty} x f(x) dx + m_x^2 \int_{-\infty}^{\infty} f(x) dx .$$

В отриманому виразі перший інтеграл дорівнює другому початковому моменту: $\int_{-\infty}^{\infty} x^2 f(x) dx = \alpha_2$; другий інтеграл – математичному сподіванню:

$\int_{-\infty}^{\infty} x f(x) dx = m_x$; третій – одиниці (1-ша властивість щільності розподілу):

$\int_{-\infty}^{\infty} f(x) dx = 1$. Замінюючи інтеграли зазначеними виразами, одержимо

$$D_x = \alpha_2 - 2m_x^2 + m_x^2 = \alpha_2 - m_x^2 .$$

Таким чином, для обчислення дисперсії випадкових величин використовують формулу

$$D_x = \alpha_2 - m_x^2 . \quad (6.11)$$

6.3.6. Середнє квадратичне відхилення

Дисперсія вимірюється у квадратних одиницях у порівнянні з одиницями виміру самої випадкової величини. Розбіжність в одиницях виміру може заподіяти незручності при аналізі дисперсійних властивостей випадкової величини. Це пов'язано з тим, що рядовому досліднику простіше порівнювати лінійні величини. Щоб одиниці виміру числової ха-

характеристики відхилення випадкової величини привести до лінійних, замість дисперсії використовують *середнє квадратичне відхилення*.



Означення 6.7. *Середнє квадратичне відхилення* являє собою квадратний корінь з дисперсії

$$\sigma_x = \sqrt{D_x} . \quad (6.12)$$

Помилка виміру є середнє квадратичне відхилення вимірюваної величини від її дійсного значення.

6.4. Моменти високих порядків

Для аналізу випадкових величин мають значення і моменти більш високих порядків. Особливої уваги заслуговують третій і четвертий центральні моменти.

6.4.1. Третій центральний момент і коефіцієнт асиметрії

Третій центральний момент характеризує ступінь відхилення випадкової величини навколо математичного сподівання, а також ступінь асиметрії її закону розподілу.

Третій центральний момент дискретної випадкової величини обчислюється за формулою

$$\mu_3 = M[(X - m_x)^3] = \sum_{i=1}^n (x_i - m_x)^3 p_i . \quad (6.13)$$

Третій центральний момент неперервної випадкової величини обчислюється за формулою

$$\mu_3 = M[(X - m_x)^3] = \int_{-\infty}^{\infty} (x - m_x)^3 f(x) dx . \quad (6.14)$$

У випадку симетричного закону розподілу $\mu_3 = 0$.

Для характеристики тільки ступеня асиметрії використовується коефіцієнт асиметрії $S = \frac{\mu_3}{\sigma_x^3}$. У випадку симетричного закону розподілу коефіцієнт асиметрії також дорівнює нулю.

На рис. 6.6 показані дві неперервні випадкові величини, задані у вигляді щільності розподілу, з різними за знаком коефіцієнтами асиметрії.

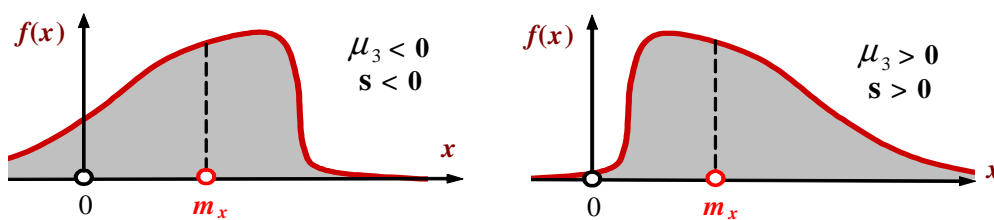


Рис. 6.6

6.4.2. Четвертий центральний момент і величина ексцес

Четвертий центральний момент характеризує ступінь відхилення випадкової величини навколо математичного сподівання, а також ступінь гостровершинності її закону розподілу.

Четвертий центральний момент дискретної випадкової величини обчислюється за формулою

$$\mu_4 = M[(X - m_x)^4] = \sum_{i=1}^n (x_i - m_x)^4 p_i. \quad (6.15)$$

Четвертий центральний момент неперервної випадкової величини обчислюється за формулою

$$\mu_4 = M[(X - m_x)^4] = \int_{-\infty}^{\infty} (x - m_x)^4 f(x) dx. \quad (6.16)$$

Для характеристики тільки ступеня гостровершинності закону розподілу використовується величина ексцес $E = \frac{\mu_4}{\sigma_x^4} - 3$. У разі нормального закону

розподілу випадкової величини ексцес дорівнює нулю ($E = 0$). Більш докладно нормальний закон розподілу буде розглянутий у підрозділі 7.2.3.

На рис. 6.7 показані три випадкові неперервні величини, задані у вигляді щільності розподілу, з різними величинами E . При цьому перша випадкова величина розподілена за нормальним законом з $E_1 = 0$, друга – з $E_2 > 0$, третя – з $E_3 < 0$.

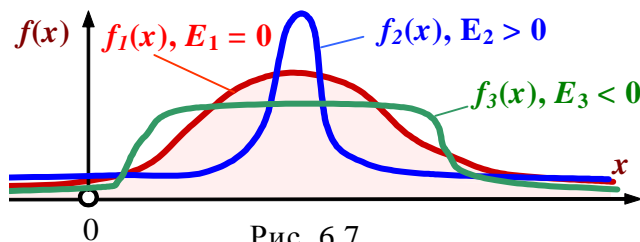


Рис. 6.7

Тема 7. ОКРЕМІ ЗАКОНИ РОЗПОДІЛУ

7.1. Закони розподілу дискретних випадкових величин

Скільки існує різних дискретних випадкових величин, стільки існує і законів їхнього розподілу. З усього різноманіття дискретних випадкових величин виділяють шість груп, що зустрічаються на практиці частіше за інші. Кожна група об'єднує випадкові величини, що мають закон розподілу, характерний тільки для цієї групи. Ймовірності конкретних значень таких випадкових величин обчислюються за однією і тією ж формулою. Відмінність випадкових величин, що входять в одну групу, визначається різними значеннями ключових компонент у визначальних формулах. Ключові компоненти формул називають параметрами закону розподілу.

Перелічимо закони розподілу для всіх згаданих груп дискретних величин:

1. Вироджений розподіл

$$P(X = a) = 1, \quad a = \text{const}.$$

2. Гіпергеометричний розподіл (N, M, n – натуральне число, $N \leq N, \quad n \leq n$),

$$P(X = m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}, \quad m = 0, 1, \dots, \min\{M, n\}.$$

3. Біноміальний розподіл (n – натуральне число, $0 < p < 1$)

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

4. Від'ємний Біноміальний розподіл (r – натуральне число, $0 < p < 1$)

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

5. Розподіл Пуассона з параметром $\lambda > 0$:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots$$

6. Геометричний розподіл ($0 < p < 1$)

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

З перелічених розподілів найбільш цікавими є біноміальний і пуассонівський закони розподілу. Тому розглянемо їх більш докладно.

7.1.1. Біноміальний закон розподілу

7.1.1.1. Загальна характеристика біноміальної випадкової величини

Нехай робиться n незалежних експериментів, у кожному з яких з однаковою ймовірністю p може відбутися деяка подія A . Подія A може мати саму різну природу. Випадкова величина X – число експериментів, у яких відбувається подія A – розподілена за біноміальним законом розподілу

$$P(X = m) = C_n^m p^m (1-p)^{n-m} \quad (7.1)$$

з рядом розподілу, що відповідає табл. 7.1.

Таблиця 7.1 – Ряд розподілу біноміальної випадкової величини

x_i	0	1	...	m	...	n
p_i	$(1-p)^n$	$np(1-p)^{n-1}$...	$C_n^m p^m (1-p)^{n-m}$...	p^n

Сума ймовірностей у другому рядку ряду розподілу (табл. 7.1) дорівнює одиниці, тобто $\sum_{i=0}^n C_n^i p^i (1-p)^{n-i} = 1$. Для доказу даного факту слід

суму $\sum_{i=0}^n C_n^i p^i (1-p)^{n-i}$ розглядати як розкладання біному Ньютона зі змінними p і $(1-p)$, тобто

$$\sum_{i=0}^n C_n^i p^i (1-p)^{n-i} = [p + (1-p)]^n = 1.$$

Біноміальний закон розподілу має два параметри:

p – ймовірність появи події A в одному експерименті;

n – загальне число експериментів (випробувань).

Ймовірність влучення дискретної випадкової величини, розподіленої за біноміальним законом, у заданий діапазон значень встановлюється за допомогою формули

$$P\{k_1 \leq X \leq k_2\} = \sum_{i=k_1}^{k_2} C_n^i p^i (1-p)^{n-i}.$$

7.1.1.2. Числові характеристики біноміальної випадкової величини

Математичне сподівання. Розглянемо попередньо випадкову величину X_i – число появ події A в i -му експерименті, $i = \overline{1, n}$. Ряд розподілу випадкової величини має вигляд:

x_{ij}	0	1
p_{ij}	$1-p$	p

. Математичне сподівання для X_i визначимо за формулою (6.1): $m_i = 0 \cdot (1-p) + 1 \cdot p = p$.

Біноміальна випадкова величина є сумою величин X_i . Тоді її математичне сподівання визначиться наступним перетворенням:

$$m_x = M[X] = M\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n M[X_i] = \sum_{i=1}^n p = np,$$

тобто

$$m_x = np. \quad (7.2)$$

Дисперсія. Визначимо попередньо другий початковий момент і дисперсію випадкової величини X_i – числа появ події A в i -му експерименті, $i = \overline{1, n}$. Ряд розподілу розглянутої величини наведено вище. Другий початковий момент випадкової величини X_i визначимо за формулою (6.7): $m_i = 0^2 \cdot (1-p) + 1^2 \cdot p = p$, дисперсію – за формулою зв'язку:

$$D_{xi} = \alpha_{2i} - m_{xi}^2 = p - p^2 = p(1-p).$$

Дисперсія біноміальної випадкової величини X (вона ж сума дисперсій незалежних випадкових величин X_i) визначиться за допомогою наступного перетворення:

$$D_x = D[X] = D\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n D[X_i] = \sum_{i=1}^n p(1-p) = np(1-p),$$

тобто

$$D_x = np(1-p). \quad (7.3)$$

Середнє квадратичне відхилення визначимо відповідно до формули (6.12):

$$\sigma_x = \sqrt{D_x} = \sqrt{np(1-p)}. \quad (7.4)$$

Приклад 7.1. Визначити математичне сподівання m_x , дисперсію D_x і середнє квадратичне відхилення випадкової величини X – числа появ “орла” при 10 киданнях монети.

Розв’язання. Кидання монети – це незалежні експерименти. Ймовірність появи “орла” при кожному киданні монети однакова і

дорівнює 0,5. Отже випадкова величина X розподілена по біноміальному закону. А це значить, що її математичне сподівання визначається за формулою (7.2):

$$m_x = np = 10 \cdot 0,5 = 5 ;$$

дисперсія – за формулою (7.3):

$$D_x = np(1-p) = 10 \cdot 0,5 \cdot (1-0,5) = 2,5 ;$$

середнє квадратичне відхилення – за формулою (7.4):

$$\sigma_x = \sqrt{D_x} = \sqrt{2,5} \approx 1,58 .$$

7.1.2. Закон розподілу Пуассона

Закон розподілу Пуассона пов'язаний з рідкими подіями, які складають найпростіший потік подій. Тому подальший виклад буде стосуватися матеріалу, що визначає і характеризує основні особливості випадкового потоку подій і його окремого випадку – найпростішого потоку подій.

7.1.2.1. Найпростіший потік подій



Означення 7.1. Випадковим потоком подій називають події, які відбуваються одна за одною у випадкові моменти часу.



Означення 7.2. Найпростішим потоком подій називається потік подій, який має наступні три властивості: *стаціонарність*; *ординарність*; *відсутність післядії*.

Прикладом найпростішого потоку подій може бути потік заявок на придбання квитків, що поступають по телефону в касу театру.

Дамо означення вище зазначеним властивостям найпростішого потоку.



Означення 7.3. Випадковий потік подій називають **стаціонарним**, якщо ймовірність влучення певного числа подій на заданий інтервал часу залежить тільки від довжини інтервалу T і не залежить від того, де на числовій осі t розташований цей інтервал.

Якщо інтервали часу T_1 і T_2 , які знаходяться на числовій осі t у різних місцях (рис. 7.1), рівні між собою, то рівні й імовірності появи визначеного числа подій m протягом цих інтервалів $P_1(X=m)$ і $P_2(X=m)$:

$$T_1 = T_2 \Rightarrow P_1(X=m) = P_2(X=m).$$



Рис. 7.1.



Означення 7.4. Випадковий потік подій називається **ординарним**, якщо ймовірність влучення двох і більше подій на нескінченно малий інтервал часу Δt занадто мала в порівнянні з ймовірністю влучення однієї події в цей інтервал.

Іншими словами, дві і більше події на одному нескінченно малому інтервалі часу відбутися не можуть, тобто має місце ліміт $\lim_{\Delta t \rightarrow 0} \frac{P\{X > 1\}}{P(X = 1)} = 0$.



Означення 7.5. Випадковий потік подій називається потоком **без післядії**, якщо ймовірність влучення певного числа подій на інтервал часу довжиною T не залежить від того, скільки подій потрапило на будь-який інший інтервал, що не перетинається з першим.

Ця властивість потоку говорить про те, що всі наступні події в потоці не залежать від попередніх.

7.1.2.2. Загальна характеристика пуассонівської випадкової величини

Для найпростішого потоку подій випадкова величина X – кількість подій, що потрапили на інтервал T , розподілена за законом Пуассона

$$P(X = m) = \frac{a^m}{m!} e^{-a}, \quad (7.5)$$

де $a = \lambda T = np$ – середнє число подій, що потрапляють на інтервал T (єдиний параметр закону розподілу);

λ – інтенсивність настання подій (кількість подій в одиницю часу);

T – певний період часу.

Ряд розподілу пуассонівської випадкової величини відповідає табл. 7.2.

Таблиця 7.2 – Ряд розподілу пуассонівської випадкової величини

x_i	0	1	...	m	...
p_i	e^{-a}	$a e^{-a}$...	$(a^m e^{-a})/m!$...

Доведення формули Пуассона (7.5). Введемо позначення:

λ – інтенсивність подій;

T – заданий інтервал часу.

Розіб'ємо інтервал часу довжиною T на ділянки Δt у кількості n . Причому $\Delta t = \frac{T}{n} \rightarrow 0$. З погляду на стаціонарність та ординарність потоку ймовірність того, що на ділянці Δt відбудеться одна подія, визначиться в такий спосіб:

$$p = \lambda \Delta t = \frac{\lambda T}{n},$$

а ймовірність того, що на ділянці Δt не відбудеться жодної події:

$$q = 1 - p = 1 - \lambda \Delta t = 1 - \frac{\lambda T}{n}.$$

За умови $n \rightarrow \infty$ ймовірність $p = \frac{\lambda T}{n} \rightarrow 0$. Ймовірність того, що за період часу T відбудеться рівно m подій, можна розглядати як ймовірність появи m подій в n незалежних випробуваннях при $n \rightarrow \infty$ і $p \rightarrow 0$, тобто обчислювати її за формулою Бернуллі:

$$P(X = m) = \lim_{n \rightarrow \infty} C_n^m p^m (1-p)^{n-m} = \lim_{n \rightarrow \infty} \left[\frac{\overbrace{n(n-1)\dots(n-m+1)}^m}{m!} \cdot p^m (1-p)^{n-m} \right] =$$

$$= \frac{n^m}{m!} \cdot p^m \lim_{n \rightarrow \infty} \frac{(1-p)^n}{(1-p)^m} = \frac{(np)^m}{m!} \cdot \frac{\lim_{n \rightarrow \infty} \left(1 - \frac{np}{n}\right)^n}{1^m} = \frac{(a)^m}{m!} \lim_{n \rightarrow \infty} \left(1 - \frac{a}{n}\right)^n.$$

Оскільки $\lim_{n \rightarrow \infty} \left(1 - \frac{a}{n}\right)^n = e^{-a}$ є визначним лімітом, то

$$P(X = m) = \frac{(a)^m}{m!} e^{-a}, \text{ що й треба було довести.}$$

Переконаємося, що сума ймовірностей у другому рядку ряду розподілу (табл. 7.2) дорівнює одиниці, тобто $\sum_{i=0}^{\infty} \frac{a^i}{i!} e^{-a} = 1$. Для доказу цієї рівності перетворимо її ліву частину: $\sum_{i=0}^{\infty} \frac{a^i}{i!} e^{-a} = e^{-a} \sum_{i=0}^{\infty} \frac{a^i}{i!}$. Тут сума $e^{-a} \sum_{i=0}^{\infty} \frac{a^i}{i!}$ є функціональний нескінченний ряд, що сходиться до функції e^a .

Здійснимо перетворення вихідної суми від початку до кінця:

$$\sum_{i=0}^{\infty} \frac{a^i}{i!} e^{-a} = e^{-a} \sum_{i=0}^{\infty} \frac{a^i}{i!} = e^{-a} \cdot e^a = 1.$$

Таким чином, вихідну рівність $\sum_{i=0}^{\infty} \frac{a^i}{i!} e^{-a} = 1$ доведено.

7.1.2.3. Числові характеристики пуассонівської випадкової величини

Математичне сподівання. Відповідно до формули (6.1) математичне сподівання дискретної випадкової величини визначиться в такий спосіб:

$$\begin{aligned} m_x = M[X] &= \sum_{i=0}^{\infty} x_i p_i = \sum_{i=0}^{\infty} i \cdot \frac{a^i}{i!} e^{-a} = e^{-a} \sum_{i=0}^{\infty} i \cdot \frac{a^i}{i!} = e^{-a} \cdot a \cdot \sum_{i=0}^{\infty} i \cdot \frac{a^{i-1}}{i!} = \\ &= e^{-a} \cdot a \cdot \sum_{i=1}^{\infty} \frac{a^{i-1}}{(i-1)!}. \end{aligned}$$

Остання сума являє собою функціональний ряд, що сходиться до функції e^a . Продовжимо перетворення:

$$m_x = e^{-a} \cdot a \cdot \sum_{i=1}^{\infty} \frac{a^{i-1}}{(i-1)!} = e^{-a} \cdot a \cdot e^a = a,$$

тобто математичне сподівання пуассонівської випадкової величини

$$m_x = a. \quad (7.6)$$

Дисперсія. Визначимо попередньо другий початковий момент відповідно до формули (6.7):

$$\alpha_2 = \sum_{i=0}^{\infty} x_i^2 p_i = \sum_{i=0}^{\infty} i^2 \cdot \frac{a^i}{i!} e^{-a} = e^{-a} \sum_{i=0}^{\infty} i^2 \frac{a^i}{i!} = e^{-a} a \sum_{i=1}^{\infty} i \frac{a^{i-1}}{(i-1)!} =$$

$$\begin{aligned}
&= e^{-a} a \sum_{i=1}^{\infty} (i-1+1) \frac{a^{i-1}}{(i-1)!} = e^{-a} a \left[\sum_{i=1}^{\infty} (i-1) \frac{a^{i-1}}{(i-1)!} + \overbrace{\sum_{i=1}^{\infty} \frac{a^{i-1}}{(i-1)!}}^{e^a} \right] = \\
&= e^{-a} a \left[a \sum_{i=2}^{\infty} \frac{a^{i-2}}{(i-2)!} + e^a \right] = e^{-a} a e^a (a+1) = a(a+1).
\end{aligned}$$

Дисперсію пуассонівської випадкової величини визначимо за формулою зв'язку:

$$D_x = \alpha_2 - m_x^2 = a(a+1) - a^2 = a,$$

тобто дисперсія

$$D_x = a. \quad (7.7)$$

Середнє квадратичне відхилення знайдемо відповідно до формули (6.12):

$$\sigma_x = \sqrt{D_x} = \sqrt{np(1-p)}. \quad (7.8)$$

7.1.2.4. Ймовірність влучення пуассонівської випадкової величини в заданий інтервал

Для пуассонівських випадкових величин існують дві спеціальні таблиці, що дозволяють вирішувати різні задачі, пов'язані з розподілом Пуассона, без обчислення факторіальних величин типу $m!$, степеневих величин типу a^m і показникових величин типу e^{-a} .

Перша таблиця дозволяє визначати ймовірність того, що пуассонівська випадкова величина приймає значення m , тобто ймовірність $P(X=m)$.

Друга таблиця дозволяє визначати ймовірність того, що пуассонівська випадкова величина приймає значення, яке менше або рівне m , тобто ймовірність $P\{X \leq m\}$.

Друга таблиця є більш універсальною, тому що дозволяє легко визначати ймовірності:

$P(X=m)$ як різницю $P\{X \leq m\} - P\{X \leq (m-1)\}$;

$P\{X \geq m\}$ як різницю $1 - P\{X \leq (m-1)\}$;

$P\{m_1 \leq X \leq m_2\}$ як різницю $P\{X \leq m_2\} - P\{X \leq (m_1-1)\}$.

7.2. Закони розподілу неперервних випадкових величин

Серед неперервних випадкових величин на особливу увагу заслуговують величини, що мають один з наступних законів розподілу: *рівномірний; показниковий; гамма-розподіл; нормальний..*

Розглянемо докладніше кожний з названих законів розподілу.

7.2.1. Рівномірний закон розподілу

7.2.1.1. Загальна характеристика



Означення 7.6. Неперервна випадкова величина розподілена за рівномірним законом, якщо її щільність розподілу має вигляд

$$f(x) = \begin{cases} 0, & x \notin [a, b]; \\ c, & x \in [a, b]. \end{cases} \quad (7.9)$$

Випадкову величину, що розподілена за законом (7.9) коротко позначають $X \sim R(a, b)$.

Графік щільності розподілу для рівномірно розподіленої випадкової величини має вигляд, показаний на рис. 7.1.

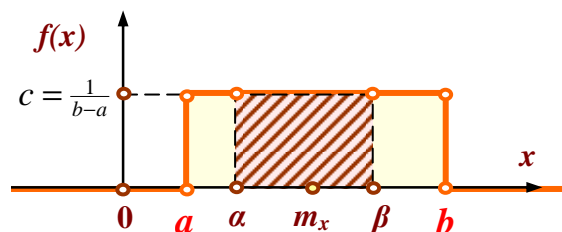


Рис. 7.1

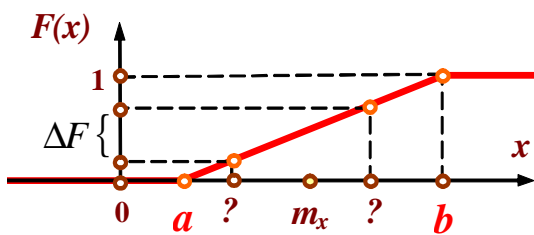


Рис. 7.2

Якщо випадкова величина розподілена на заданому інтервалі за рівномірним законом, то величина c у (7.9) має конкретне значення, що визначається за допомогою першої властивості щільності розподілу (5.6). Для цього не-

обхідно взяти визначений інтеграл у нескінченних границях від щільності розподілу (7.9) і дорівняти його одиниці:

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^a 0 dx + \int_a^b c dx + \int_b^{\infty} 0 dx = 0 + cx \Big|_a^b + 0 = c(b - a);$$

$$c(b - a) = 1.$$

Звідси $c = \frac{1}{b-a}$. Тоді

$$f(x) = \begin{cases} 0, & x \notin [a, b]; \\ \frac{1}{b-a}, & x \in [a, b]. \end{cases} \quad (7.10)$$

Рівномірний закон розподілу має два параметри: a і b .

Інтегральна функція розподілу, відповідно до зворотного перетворення (5.5), визначається в такий спосіб:

$$F(x) = \begin{cases} \int_{-\infty}^x 0 \cdot dt = 0, & x < a; \\ \int_{-\infty}^a 0 \cdot dt + \int_a^x \frac{1}{b-a} \cdot dt = \frac{x-a}{b-a}, & a \leq x \leq b; \\ \int_{-\infty}^a 0 \cdot dt + \int_a^b \frac{1}{b-a} \cdot dt + \int_b^x 0 \cdot dt = \frac{b-a}{b-a} = 1, & x > b. \end{cases}$$

Таким чином, інтегральна функція рівномірно розподіленої величини визначається як

$$F(x) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & x > b. \end{cases} \quad (7.11)$$

Графік інтегральної функції для рівномірно розподіленої випадкової величини має вигляд, показаний на рис. 7.2.

7.2.1.2. Числові характеристики

Математичне сподівання. Відповідно до формули (6.2) математичне сподівання неперервної випадкової величини визначиться в такий спосіб:

$$\begin{aligned} m_x &= \int_{-\infty}^{\infty} x f(x) dx = \\ &= \int_{-\infty}^a x \cdot 0 \cdot dx + \int_a^b x \frac{1}{b-a} \cdot dx + \int_b^{\infty} x \cdot 0 \cdot dx = \frac{1}{b-a} \cdot \frac{x^2}{2} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}, \end{aligned}$$

тобто

$$m_x = \frac{a+b}{2}. \quad (7.12)$$

Математичне сподівання (7.12) рівномірно розподіленої випадкової величини знаходиться в середині відрізка $[a, b]$ (див. рис. 7.1). Щільність розподілу (7.10) має осьову симетрію з віссю, що проходить через точку m_x паралельно осі ординат.

Дисперсія. Визначимо попередньо другий початковий момент відносно до формули (6.8):

$$\begin{aligned}\alpha_2 &= \int_{-\infty}^{\infty} x^2 f(x) dx = \\ &= \int_{-\infty}^a x^2 \cdot 0 \cdot dx + \int_a^b x^2 \frac{1}{b-a} dx + \int_b^{\infty} x^2 \cdot 0 \cdot dx = \frac{x^3}{3(b-a)} \Big|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3},\end{aligned}$$

тобто

$$\alpha_2 = \frac{a^2 + ab + b^2}{3}. \quad (7.13)$$

Дисперсію рівномірно розподіленої випадкової величини визначимо за формулою зв'язку:

$$D_x = \alpha_2 - m_x^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12},$$

тобто

$$D_x = \frac{(b-a)^2}{12}. \quad (7.14)$$

Середнє квадратичне відхилення знайдемо за формулою (6.12):

$$\sigma_x = \sqrt{D_x} = \frac{(b-a)\sqrt{3}}{6}. \quad (7.15)$$

7.2.1.3. Ймовірність влучення випадкової величини в заданий діапазон

Ймовірність влучення рівномірно розподіленої випадкової величини в заданий діапазон $P\{\alpha \leq X \leq \beta\}$, якщо діапазон $[\alpha, \beta]$ входить у діапазон $[a, b]$, можна визначити двома способами:

1-й спосіб. За формулою (5.2)

$$P\{\alpha \leq X \leq \beta\} = F(\beta) - F(\alpha) = \frac{\beta - a}{b - a} - \frac{\alpha - a}{b - a} = \frac{\beta - \alpha}{b - a}.$$

2-й спосіб. За формулою (5.8)

$$P\{\alpha \leq X \leq \beta\} = \int_{\alpha}^{\beta} f(x)dx = \int_{\alpha}^{\beta} \frac{1}{b-a} dx = \frac{x}{b-a} \Big|_{\alpha}^{\beta} = \frac{\beta - \alpha}{b-a}.$$

Таким чином,

$$[\alpha, \beta] \subset [a, b] \Rightarrow P\{\alpha \leq X \leq \beta\} = \frac{\beta - \alpha}{b - a}. \quad (7.16)$$

З геометричної точки зору ймовірності $P\{\alpha \leq X \leq \beta\}$ відповідає площа, виділена штрихуванням на рис. 7.1.

Якщо діапазон $[\alpha, \beta]$ не входить до діапазону $[a, b]$, то вираз (7.16) не є справедливим. У цьому випадку необхідно керуватися виразом

$$[\alpha, \beta] \not\subset [a, b] \Rightarrow P\{\alpha \leq X \leq \beta\} = \frac{(\beta - \alpha) \cap (b - a)}{b - a}, \quad (7.17)$$

де $(\beta - \alpha) \cap (b - a)$ – довжина діапазону на числовій осі, що є загальною (перетинанням) для діапазону $[\alpha, \beta]$ і діапазону $[a, b]$.

7.2.2. Показниковий закон розподілу

7.2.2.1. Загальна характеристика

У найпростішому потоці подій випадкова величина T – інтервал часу між двома послідовними подіями – розподілена за *показниковим законом*.



Означення 7.7. Неперервна випадкова величина розподілена за показниковим законом, якщо її щільність розподілу має вигляд

$$f(t) = \begin{cases} 0, & t < 0; \\ \lambda e^{-\lambda t}, & t \geq 0, \end{cases} \quad (7.18)$$

де λ – інтенсивність подій, тобто кількість подій в одиницю часу.

Випадкову величину, що розподілена за законом (7.18) коротко позначають $X \sim \text{Exp}(\lambda)$.

Графік щільності розподілу для випадкової величини, розподіленої за показниковим законом, має вигляд, показаний на рис. 7.3.

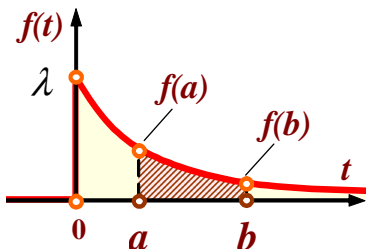


Рис. 7.3

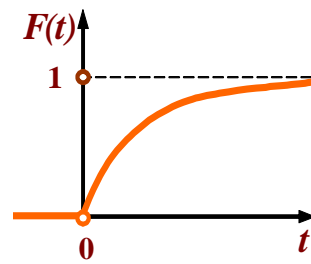


Рис. 7.4

Показниковий закон розподілу має тільки один параметр λ .

Інтегральна функція розподілу, відповідно до зворотного перетворення (5.5), визначається в такий спосіб:

$$F(t) = \begin{cases} \int_{-\infty}^t 0 \cdot dx = 0, & t < 0; \\ \int_{-\infty}^0 0 \cdot dx + \int_0^t \lambda e^{-\lambda x} dx = \lambda \left(-\frac{1}{\lambda} \right) e^{-\lambda x} \Big|_0^t = -e^{-\lambda t} + 1 = 1 - e^{-\lambda t}, & t \geq 0. \end{cases}$$

Таким чином, інтегральна функція випадкової величини, розподіленої за показниковим законом, визначається виразом

$$F(t) = \begin{cases} 0, & t < 0; \\ 1 - e^{-\lambda t}, & t \geq 0. \end{cases} \quad (7.19)$$

Графік інтегральної функції (7.19) зображений на рис. 7.4.

7.2.2.2. Числові характеристики

Математичне сподівання. Математичне сподівання випадкової величини, розподіленої за показниковим законом, визначається рівністю

$$m_x = \frac{1}{\lambda}. \quad (7.20)$$

Доведення. Відповідно до формули (6.2) математичне сподівання випадкової величини, розподіленої за показниковим законом, визначається в такий спосіб:

$$m_t = \int_{-\infty}^{\infty} t f(t) dx = \int_{-\infty}^0 t \cdot 0 \cdot dx + \int_0^{\infty} t \lambda e^{-\lambda t} dt = \int_0^{\infty} t \lambda e^{-\lambda t} dt.$$

Отриманий інтеграл інтегруємо частинами. Нагадаємо правило обчислення визначеного інтеграла частинами: $\int_a^b u dv = [uv]_a^b - \int_a^b v du$. Отже позна-

чимо: $u = -t$, $dv = -\lambda e^{-\lambda t} dt$, тоді $du = -dt$, $v = e^{-\lambda t}$. Інтегруємо частинами:

$$\begin{aligned} \int_0^{\infty} t \lambda e^{-\lambda t} dt &= \left[\underbrace{-t}_u \cdot \underbrace{e^{-\lambda t}}_v \right]_0^{\infty} - \left(\underbrace{-\int_0^{\infty} e^{-\lambda t} dt}_v \right) = \lim_{t \rightarrow \infty} \underbrace{\frac{-t}{e^{\lambda t}}}_0 - \lim_{t \rightarrow 0} \underbrace{(-te^{-\lambda t})}_0 - \frac{e^{-\lambda t}}{\lambda} \Big|_0^{\infty} = \\ &= -\lim_{t \rightarrow \infty} \underbrace{\frac{e^{-\lambda t}}{\lambda}}_0 + \lim_{t \rightarrow 0} \underbrace{\frac{e^{-\lambda t}}{\lambda}}_{\frac{1}{\lambda}} = \frac{1}{\lambda}, \end{aligned}$$

тобто

$$m_x = \frac{1}{\lambda}.$$

Примітка. Обчислення ліміту $\lim_{t \rightarrow \infty} \frac{-t}{e^{-\lambda t}}$ здійснюється за правилом Лопітала.

Дисперсія. Визначимо попередньо другий початковий момент відповідно до формули (6.8):

$$\alpha_2 = \int_{-\infty}^{\infty} t^2 f(t) dx = \int_{-\infty}^0 t^2 \cdot 0 \cdot dx + \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt.$$

Отриманий інтеграл інтегруємо частинами. Позначимо: $u = t^2$, $dv = \lambda e^{-\lambda t} dt$, тоді $du = 2t dt$, $v = -e^{-\lambda t}$. Інтегруємо частинами:

$$\alpha_2 = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt = \left[\underbrace{-t^2 e^{-\lambda t}}_{uv} \right]_0^{\infty} - \left(\underbrace{-2 \int_0^{\infty} t e^{-\lambda t} dt}_v \right) = \frac{2}{\lambda} \underbrace{\int_0^{\infty} t \lambda e^{-\lambda t} dt}_{m_t = \frac{1}{\lambda}} = \frac{2}{\lambda^2}.$$

Дисперсію визначимо за формулою зв'язку:

$$D_t = \alpha_2 - m_t^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2},$$

тобто

$$D_x = \frac{1}{\lambda^2}. \quad (7.21)$$

Середнє квадратичне відхилення визначимо відповідно до формули (6.12):

$$\sigma_x = \sqrt{D_x} = \sqrt{\frac{1}{\lambda^2}} = \frac{1}{\lambda}. \quad (7.22)$$

7.2.2.3. Ймовірність влучення випадкової величини в заданий діапазон

Ймовірність влучення випадкової величини в заданий діапазон $P\{a \leq T \leq b\}$, якщо діапазон $[a, b]$ входить у діапазон $[0, \infty]$, можна визначити двома способами:

1-й спосіб. За формулою (5.2)

$$P\{a \leq T \leq b\} = F(b) - F(a) = (1 - e^{-\lambda b}) - (1 - e^{-\lambda a}) = e^{-\lambda a} - e^{-\lambda b}.$$

2-й спосіб. За формулою (5.8)

$$P\{a \leq T \leq b\} = \int_a^b \lambda e^{-\lambda t} dt = -e^{-\lambda t} \Big|_a^b = -(e^{-\lambda b} - e^{-\lambda a}) = e^{-\lambda a} - e^{-\lambda b}.$$

Таким чином,

$$P\{a \leq T \leq b\} = e^{-\lambda a} - e^{-\lambda b}. \quad (7.23)$$

З геометричної точки зору ймовірності $P\{a \leq T \leq b\}$ відповідає площа, виділена штрихуванням на рис. 7.3.

7.2.3. Нормальний закон розподілу

7.2.3.1. Загальна характеристика

Найбільш простим законом, що досить точно відображує випадкові помилки вимірювань, є так званий нормальний закон розподілу.



Означення 7.8. Неперервна випадкова величина X розподілена за нормальним законом, якщо її щільність розподілу має вигляд

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (7.24)$$

де σ і m – параметри розподілу.

Випадкову величину, що розподілена за законом (7.24) коротко позначають $X \sim N(m, \sigma)$.

Графік щільності розподілу для випадкової величини, розподіленої за нормальним законом, має вигляд, показаний на рис. 7.5. Щільність розпо-

ділу (7.24) має осьову симетрію з віссю, що проходить через точку m_x паралельно осі ординат.

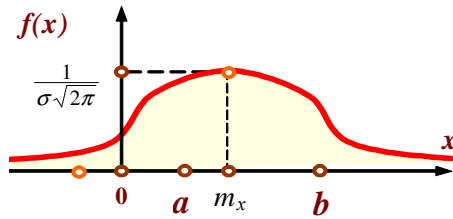


Рис. 7.5

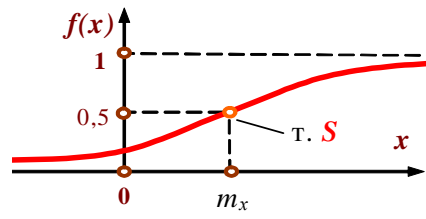


Рис. 7.6

Інтегральна функція розподілу, відповідно до зворотного перетворення (5.5), визначається в такий спосіб:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-m)^2}{2\sigma^2}} dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt .$$

Таким чином, інтегральна функція випадкової величини, розподіленої за нормальним законом, визначається інтегралом

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt . \quad (7.25)$$

Графік інтегральної функції (7.25) зображений на рис. 7.6. Крива інтегральної функції (7.25) має центральну симетрію відносно точки S .

7.2.3.2. Числові характеристики

Математичне сподівання. Математичне сподівання випадкової величини, розподіленої за нормальним законом, визначається виразом

$$m_x = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = m . \quad (7.26)$$

Дисперсія. Дисперсія випадкової величини, розподіленої за нормальним законом, визначається виразом

$$D_x = \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \sigma^2 . \quad (7.27)$$

Середнє квадратичне відхилення визначається відповідно до формули (6.12):

$$\sigma_x = \sqrt{D_x} = \sigma . \quad (7.28)$$

Центральні моменти. Центральні моменти будь-якого порядку нормально розподіленої випадкової величини визначаються рекурентним співвідношенням

$$\mu_s = (s-1)\mu_{s-2}\sigma^2. \quad (7.29)$$

Знаючи 1-й і 2-й центральні моменти, можна легко знайти будь-який інший.

Оскільки 1-й центральний момент для всіх випадкових величин дорівнює нулю, то всі непарні центральні моменти для нормально розподіленої випадкової величини також рівні нулю.

Оскільки 2-й центральний момент

$$\mu_2 = D_x = \sigma^2,$$

то всі парні центральні моменти нормально розподіленої випадкової величини легко визначаються за допомогою виразу (7.29):

$$\begin{aligned} \mu_4 &= (4-1)\mu_2\sigma^2 = 3 \cdot \sigma^2 \cdot \sigma^2 = 3\sigma^4; \\ \mu_6 &= (4-1)\mu_4\sigma^2 = 5 \cdot 3\sigma^2 \cdot \sigma^2 = 15\sigma^4; \\ \mu_8 &= (4-1)\mu_6\sigma^2 = 7 \cdot 15\sigma^2 \cdot \sigma^2 = 105\sigma^4; \\ &\dots \end{aligned}$$

Оскільки всі центральні непарні моменти для нормально розподіленої випадкової величини дорівнюють нулю, то *коефіцієнт асиметрії* також дорівнює нулю:

$$S = \frac{\mu_3}{\sigma^3} = \frac{0}{\sigma^3} = 0.$$

Коефіцієнт гостровершинності (величина ексцес) для нормально розподіленої випадкової величини також дорівнює нулю:

$$E = \frac{\mu_4}{\sigma^4} - 3 = \frac{3\sigma^4}{\sigma^4} - 3 = 0.$$

7.2.3.3. Ймовірність влучення випадкової величини в заданий діапазон

Ймовірність влучення нормально розподіленої випадкової величини в заданий діапазон $P\{a \leq X \leq b\}$ можна одержати за відомими формулами:

$$P\{a \leq X \leq b\} = \begin{cases} F(b) - F(a); \\ \int_a^b f(x)dx. \end{cases}$$

Однак у цьому разі інтегрування треба проводити чисельними методами з залученням обчислювальної техніки. Щоб уникнути необхідності інтегрувати інтеграл, елементарними методами, використовують окрему інтегральну функцію

$$F^*(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \quad (7.30)$$

тобто інтегральну функцію нормально розподіленої випадкової величини з параметрами: $m = 0$; $\sigma = 1$. Розподіл (7.30) називають *стандартним нормальним розподілом*.

Інтегральна функція $F(x)$ нормально розподіленої випадкової величини пов'язана зі стандартною інтегральною функцією співвідношенням

$$F(x) = F^*\left(\frac{x-m}{\sigma}\right).$$

Тоді ймовірність влучення випадкової величини в заданий діапазон

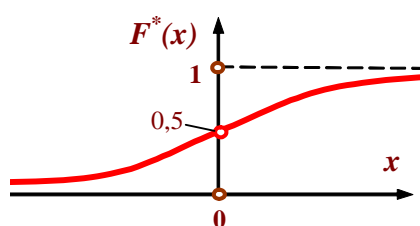


Рис.7.7

$$P\{a \leq X \leq b\} = F^*\left(\frac{b-m}{\sigma}\right) - F^*\left(\frac{a-m}{\sigma}\right). \quad (7.31)$$

На рис. 7.7 зображена інтегральна функція стандартного нормального розподілу (порівняй з рис. 7.6)

Розглянемо $F^*(x)$ від аргументу $x > 0$

$$\begin{aligned} F^*(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{t^2}{2}} dt}_{0,5} + \underbrace{\frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt}_{\hat{O}(\delta)} = \\ &= 0,5 + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt = \Phi(x) + 0,5 \end{aligned}$$

де $\Phi(x)$ – функція Лапласа.

Оскільки $F^*(x) = \hat{O}(\delta) + 0,5$, то (7.31) перепишеться як

$$P\{a \leq X \leq b\} = F^*\left(\frac{b-m}{\sigma}\right) - F^*\left(\frac{a-m}{\sigma}\right) = \Phi\left(\frac{b-m}{\sigma}\right) + 0,5 - \Phi\left(\frac{a-m}{\sigma}\right) - 0,5,$$

тобто

$$P\{a \leq X \leq b\} = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right). \quad (7.32)$$

Формула (7.32) має високу універсальність, оскільки дозволяє визначати ймовірність влучення на задану ділянку будь-якої нормально розподіленої випадкової величини незалежно від значень її параметрів m і σ .

7.2.3.4. Правило трьох сигм

Формула (7.32) може бути використана для обчислення ймовірності того, що відхилення випадкової величини X , розподіленої за нормальним законом, від її математичного сподівання за абсолютним значенням менше заданого числа δ . Часто такий розрахунок потрібний у практичних задачах, коли потрібно знайти ймовірність здійснення нерівності

$$|X - m| < \delta. \quad (7.33)$$

Перетворимо (7.33) у

$$m - \delta < X < m + \delta \quad (7.34)$$

і підставимо у формулу (7.32). Оскільки $\Phi(x)$ непарна функція, тобто $\Phi(-x) = -\Phi(x)$, маємо:

$$\begin{aligned} P(|X - m| < \delta) &= P(m - \delta < X < m + \delta) = \\ &= \Phi\left[\frac{(m + \delta) - m}{\sigma}\right] - \Phi\left[\frac{(m - \delta) - m}{\sigma}\right] = 2\Phi\left(\frac{\delta}{\sigma}\right), \end{aligned}$$

тобто ймовірність модуля відхилення випадкової величини, розподіленої за нормальним законом, можна обчислити за формулою

$$P(|X - m| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right). \quad (7.35)$$

Якщо вимірювати величину відхилення в одиницях σ , то можна вивести практично корисну закономірність, що відома як правило трьох сигм. Дійсно, покладемо у (7.35) $\delta = \sigma \cdot t$. Одержимо:

$$P(|X - m| < \sigma t) = 2\Phi(t).$$

Якщо $t=3$ і, отже, $\sigma \cdot t = 3\sigma$, то

$$P(|X - m| < 3\sigma) = 2\Phi(3) = 0,9973,$$

тобто ймовірність того, що відхилення за абсолютним значенням буде менше потроєного середньоквадратичного відхилення, дуже велика. Це означає, що ймовірність протилежної події, яка полягає в тому, що абсолютне відхилення перевищить потроєне σ , дуже мала, а саме дорівнює 0,0027. У цьому й полягає сутність правила трьох сигм.

Правило трьох сигм. Якщо випадкова величина розподілена нормально, то абсолютна величина її відхилення від математичного сподівання не перевершує потроєного середньоквадратичного відхилення.

7.3. Похідні розподіли від нормального розподілу

Розглянемо декілька розподілів, що зв'язані з нормальним розподілом і використовуються як інструмент для розв'язання багатьох задач математичної статистики.

7.3.1. Розподіл Пірсона

Розподіл Пірсона має ще іншу назву – χ -квадрат.

Нехай незалежні випадкові величини u_i розподілені за стандартним нормальним законом, тобто за нормальним законом з параметрами $m = 0$ і $\sigma = 1$. Тоді випадкова величина

$$\chi^2 = \sum_{i=1}^n u_i^2 \quad (7.36)$$

розподілена за законом χ -квадрат з числом ступенів свободи k , рівним n . Число ступенів свободи є абстрактним поняттям, що визначає в даному випадку умови незалежності величин u_i . Наявність будь-якої залежності між величинами u_i зменшує число ступенів свободи k на одиницю.

Зі збільшенням числа ступенів свободи k розподіл χ -квадрат наближається до стандартного нормального розподілу.

Для розподілу χ -квадрат складено таблицю ймовірності того, що величина χ^2 виявиться більшою за фіксоване значення χ_1^2 , тобто ймовірності $P\{\chi^2 > \chi_1^2\} = \beta$, де β – довірча ймовірність. Таблиця має два вхідних параметри: β і k .

7.3.2. Розподіл Ст'юдента

Нехай випадкова величина u розподілена за стандартним нормальним законом, а випадкова величина v розподілена за законом χ -квадрат з числом ступенів свободи k і не залежить від u . Тоді випадкова величина

$$t = \frac{u}{\sqrt{v/k}} \quad (7.37)$$

розподілена за законом Ст'юдента з числом ступенів свободи k .

Для розподілу Ст'юдента складено таблицю ймовірності того, що випадкова величина $|t|$ виявиться меншою за фіксоване значення t_1 , тобто ймовірності $P\{|t| < t_1\} = \beta$, де β – довірча ймовірність. Таблиця має два входи:

- рівень значущості $2\alpha = 1 - \beta$;
- число ступенів свободи k .

7.3.3. Розподіл Фішера

Нехай незалежні випадкові величини u і v розподілені за законом хі-квадрат відповідно зі ступенями свободи k_1 і k_2 . Тоді випадкова величина

$$F = \frac{u/k_1}{v/k_2} \quad (7.38)$$

розподілена за законом Фішера зі ступенями свободи k_1 і k_2 .

Для розподілу Фішера складено таблицю ймовірності того, що випадкова величина F виявиться більшою за фіксоване значення F_1 , тобто ймовірності $P\{F > F_1\} = \beta$, де β – довірча ймовірність. Таблиця має три входи:

- довірча ймовірність β ;
- число ступенів свободи k_1 ;
- число ступенів свободи k_2 .

7.4. Гамма-розподіл

7.4.1. Загальна характеристика

Гама-розподіл в теорії ймовірностей – це двопараметрична сім'я абсолютно неперервних розподілів. Гамма-розподіл складається з параметрів k і θ . Якщо k – ціле тоді розподіл показує суму k незалежних експоненціально розподілених випадкових величин, кожна з яких приймає значення θ . Якщо параметр k приймає ціле значення, то такий гамма-розподіл також називається розподілом Ерланга. Такий розподіл застосовується для визначення часу очікування в системах масового обслуговування.



Означення 7.9. Неперервна випадкова величина X розподілена за гамма-законом, якщо її щільність розподілу має вигляд

$$f(x) = \begin{cases} x^{k-1} \frac{e^{-x/\theta}}{\theta^k \Gamma(k)}, & x \geq 0 \\ 0, & x < 0 \end{cases},$$

де $\Gamma(k) = \int_0^{\infty} x^{k-1} e^{-x} dx$. При цьому мають місце наступні

властивості:

- $\Gamma(k) = (k-1) \cdot \Gamma(k-1)$;
- $\Gamma(0,5) = \sqrt{\pi}$;
- константи $k, \theta > 0$.

(7.39)

Якщо випадкова величина X відповідає означенню 7.39, то кажуть, що вона має гамма-розподіл з параметрами k і θ . Коротко позначають $X \sim \Gamma(k, \theta)$.

7.4.2. Числові характеристики

Математичне сподівання. Математичне сподівання випадкової величини, розподіленої за гамм-законом, визначається виразом

$$m_x = k\theta. \quad (7.26)$$

Дисперсія. Дисперсія випадкової величини, розподіленої за гамм-законом, визначається виразом

$$D_x = k\theta^2. \quad (7.27)$$

Середнє квадратичне відхилення визначається відповідно до формули (6.12):

$$\sigma_x = \sqrt{D_x} = \theta\sqrt{k}. \quad (7.28)$$

7.4.3. Особливості гамма-розподілу

• Якщо випадкової величини X_1, X_2, \dots, X_n розподілені за гамма-законом, тобто $X_i \sim \Gamma(k_i, \theta)$, $i = \overline{1, n}$, то

$$Y = \sum_{i=1}^n X_i \sim \Gamma\left(\sum_{i=1}^n k_i, \theta\right).$$

• Якщо $X \sim \Gamma(k, \theta)$ і a – довільна константа ($a > 0$), то $aX \sim \Gamma(k, a\theta)$.

7.4.2. Зв'язок з іншими розподілами

• Експоненціальний розподіл є частковим випадком гама-розподілу, в якому $k = 1$, $\theta = \lambda$.

• Якщо $\Gamma_1, \Gamma_2, \dots, \Gamma_k$ – незалежні експоненціальні випадкові величини з однаковим параметром λ , то $Y = \sum_{i=1}^k X_i \sim \Gamma(k, \theta)$, де $\theta = \lambda$.

• Розподіл хі-квадрат з числом ступенів свободи k є частковим випадком гамма-розподілу $\Gamma\left(\frac{k}{2}, 2\right)$

• При великих k гамма-розподіл $\Gamma(k, \theta)$ може бути наближеним нормальним розподілом з параметрами $m = k\theta$, $\sigma = k\theta^2$ при $k \rightarrow \infty$.

7.4. Перетворення випадкових величин

Нехай задана випадкова величина X з відомим законом розподілу. Вона піддається перетворенню

$$Y = \varphi(X),$$

результатом якого є випадкова величина Y . Потрібно визначити закон розподілу для перетвореної випадкової величини.

Нехай X – дискретна випадкова величина, тобто її можна задати за допомогою ряду розподілу (табл.7.2).

Таблиця 7.3 – Ряд розподілу випадкової величини X

X	x_1	x_2	\dots	x_k	\dots
P	$P(x_1)$	$P(x_2)$	\dots	$P(x_k)$	\dots

Визначимо інтегральну функцію розподілу для Y :

$$F(y) = P\{Y = \varphi(X) < y\} = \sum_{x_k: \varphi(x) < y} P(x_k),$$

де підсумовуються ймовірності тих величин x_k , для яких $\varphi(x) < y$.

Аналогічно записується функція розподілу для Y , якщо X - неперервна випадкова величина з щільністю розподілу $f(x)$

$$F(y) = P\{Y = \varphi(X) < y\} = \int_{x_k: \varphi(x) < y} P(x_k) dx.$$

Зокрема, якщо $\varphi(x)$ – строго монотонна функція, тоді

$$F(y) = P\{\varphi(X) < y\} = P\{X < \varphi^{-1}(y)\} = F(\varphi^{-1}(y)).$$

Якщо відомо, що Y є неперервною величиною, то можна визначити її щільність диференціюванням

$$f(y) = \frac{d}{dy} F(y) = P\{\varphi(X) < y\} = f(\varphi^{-1}(y)) \frac{d}{dy} \varphi^{-1}(y).$$

Тема 8. ВИПАДКОВІ ВЕКТОРИ І ФУНКЦІЇ ВИПАДКОВИХ АРГУМЕНТІВ

8.1. Випадкові вектори



Означення 8.1. Випадковим вектором називають вектор $\mathbf{X}=(X_1, X_2, \dots, X_n)$, компоненти якого є випадковими величинами.

Так само, як і для випадкової величини, для випадкового вектора вводяться поняття інтегральної функції розподілу й числові характеристики.

8.1.1. Інтегральна функція розподілу випадкового вектора



Означення 8.2. Інтегральна функція сумісного розподілу випадкового вектора – це така функція $F(x_1, x_2, \dots, x_n)$, яка при конкретних значеннях своїх аргументів чисельно дорівнює ймовірності того, що випадкові компоненти вектора виявляться менше за відповідні аргументи, тобто $F(x_1, x_2, \dots, x_n) = P\{X_1 < x_1, X_2 < x_2, \dots, X_n < x_n\}$.

Надалі, як правило, будуть розглядатися тільки двовимірні випадкові вектори $\mathbf{Z} = (X, Y)$, де X, Y – компоненти вектора. Однак усі наведені положення або в однаковій мірі справедливі і для багатовимірних векторів, або легко узагальнюються на випадок багатовимірних векторів.

У загальному випадку спільна інтегральна функція неперервного двовимірного випадкового вектора являє собою криволінійну поверхню $F(x, y)$, укладену між двома необмеженими площинами F_0 і F_1 , обумовленими відповідно рівностями $F=0$ і $F=1$ (рис. 8.1)

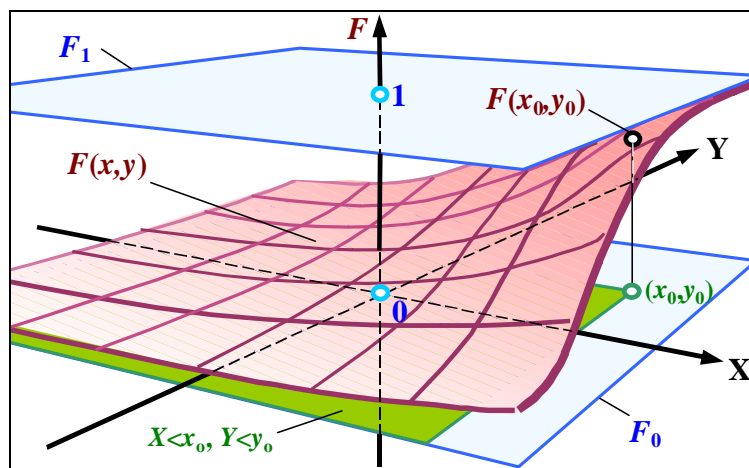


Рис. 8.1

Поверхня $F(x,y)$ асимптотично наближається до площини F_0 , коли або $x \rightarrow -\infty$, або $y \rightarrow -\infty$, або одночасно $x \rightarrow -\infty$ і $y \rightarrow -\infty$. При одночасному виконанні умов $x \rightarrow \infty$ і $y \rightarrow \infty$ поверхня $F(x,y)$ асимптотично наближається до площини F_1 . За означенням інтегральна функція двовимірного випадкового вектора $\mathbf{Z} = (X,Y)$ – це така функція $F(x,y)$, яка при кожних конкретних значеннях своїх аргументів x і y чисельно дорівнює ймовірності того, що випадкові компоненти вектора виявляться менше за відповідні аргументи, тобто $F(x,y) = P\{X < x, Y < y\}$. Іншими словами, інтегральна функція двовимірного випадкового вектора в конкретній точці (x_0, y_0) дорівнює ймовірності влучення випадкового вектора на затінену ділянку площини координат XOY (див. рис. 8.1 і 8.2).

Спільна інтегральна функція двовимірного випадкового вектора $\mathbf{Z} = (X,Y)$ має ряд властивостей, що формулюються в такий спосіб:

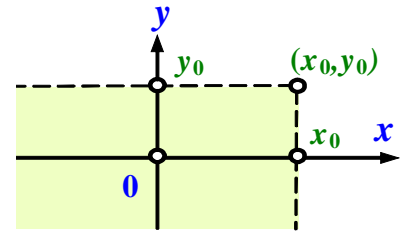


Рис. 8.2

1-а властивість

$$F(-\infty, -\infty) = 0 ; \quad (8.1)$$

$$F(x, -\infty) = 0 ; \quad (8.2)$$

$$F(-\infty, y) = 0 . \quad (8.3)$$

2-а властивість

$$F(\infty, \infty) = 1 . \quad (8.4)$$

3-я властивість

$$F(x, \infty) = P\{X < x, Y < \infty\} = P\{X < x\} = F_1(x) ; \quad (8.5)$$

$$F(\infty, y) = P\{X < \infty, Y < y\} = P\{Y < y\} = F_2(y) . \quad (8.6)$$

4-а властивість

$F(x,y)$ – функція, що монотонно зростає від обох своїх аргументів.

Разом із сумісним розподілом $F(x_1, x_2, \dots, x_n)$ існують розподіли ймовірностей кожної випадкової величини X_i , що входять до системи. На відміну від сумісного, ці розподіли називають **маргінальними**.

Урахування 3-ї властивості (8.5) – (8.6) дозволяє за відомою спільною інтегральною функцією $F(x,y)$ визначати маргінальні інтегральні функції розподілу компонент, тобто $F_1(x) = F(x, \infty)$; $F_2(y) = F(\infty, y)$.

Узагальнення 3-ї властивості на спільні інтегральні функції тривимірних векторів і функції більшої вимірності дозволяє одержати вирази для визначення часткових інтегральних функцій. Так, для тривимірної інтегральної функції $F(x_1, x_2, x_3)$ справедливі вирази: $F_1(x_1) = F(x_1, \infty, \infty)$; $F_2(x_2) = F(\infty, x_2, \infty)$, $F_{2,3}(x_2, x_3) = F(\infty, x_2, x_3)$ і т.і.

8.1.2. Ймовірність влучення випадкового вектора в заданий діапазон

Ймовірність влучення дискретного чи неперервного випадкового вектора в заданий діапазон (прямокутник) може бути визначена за допомогою однієї і тієї формули, заснованої на використанні інтегральної функції розподілу.

Нехай відома інтегральна функція $F(x,y)$ і задані параметри ділянки ΔZ , на яку з шуканою ймовірністю потрапляє випадковий вектор Z (див. рис. 8.3), тобто задані координати кутів прямокутника ΔZ . Тоді шукана ймовірність

$$P\{(X,Y) \in \Delta Z\} = F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - F(x + \Delta x, y) + F(x, y) . \quad (8.7)$$

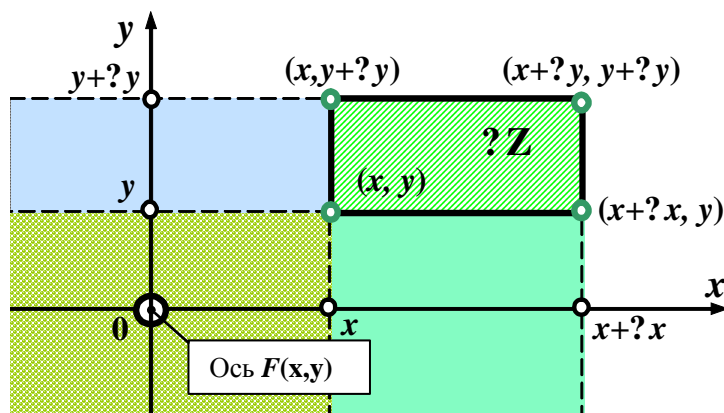


Рис. 8.3

8.1.3. Щільність сумісного розподілу випадкового вектора

Якщо компоненти випадкового вектора є неперервними величинами, то закон розподілу цього вектора може бути заданий у формі щільності розподілу (диференціальної функції розподілу).

Щільність розподілу випадкового вектора – це границя відношення ймовірності влучення випадкового вектора на нескінченно малу ділянку ΔZ до площі цієї ділянки:

$$\begin{aligned} f(x, y) &= \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P\{(X,Y) \in \Delta Z\}}{\Delta x \cdot \Delta y} = \\ &= \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - F(x + \Delta x, y) + F(x, y)}{\Delta x \cdot \Delta y}, \end{aligned}$$

тобто щільність розподілу двовимірного випадкового вектора являє собою другу часткову похідну від інтегральної функції

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \quad (8.8)$$

З (8.8) виходить, що при відомій щільності розподілу випадкового вектора інтегральна функція визначається за допомогою зворотного перетворення

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(t, \tau) dt d\tau. \quad (8.9)$$

Щільність розподілу випадкового вектора $f(x, y)$ успадковує всі властивості інтегральної функції $F(x, y)$. Так, наведені раніше 1-а і 2-а властивості інтегральної функції (8.1) – (8.4) трансформуються в **1-у властивість щільності розподілу**

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1; \quad (8.10)$$

3-я властивість (8.5) – (8.6) трансформується в **2-у властивість щільності розподілу**

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy; \quad (8.11)$$

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx; \quad (8.12)$$

4-та властивість – у **3-ю властивість щільності розподілу**

$$f(x, y) \geq 0. \quad (8.13)$$

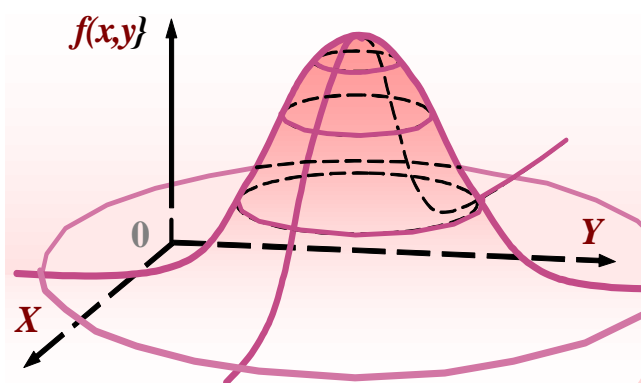


Рис. 8.4

З геометричної точки зору перша властивість щільності розподілу (8.11) означає, що об'єм, укладений між поверхнею $f(x, y)$ і координатною площиною XOY (рис. 8.4), дорівнює одиниці. Третя властивість (8.13) говорить про те, що поверхня $f(x, y)$ не може бути розташовувана нижче координатної площини XOY .

Для отримання маргінальної щільності за тією або іншою компонентою X_i треба проінтегрувати спільну щільність по всіх компонентах, крім обраної X_i .

8.1.4. Умовні закони розподілу



Означення 8.3. Умовний закон розподілу у формі $f(x/y)$ або $F(x/y)$ – це закон розподілу випадкової величини X , обчислений за умови, що випадкова величина Y прийняла конкретне значення.



Означення 8.4. Випадкові величини X і Y називаються **незалежними**, якщо закон розподілу випадкової величини X не залежить від того, яке значення прийняла випадкова величина Y . У протилежному разі величини X і Y називаються **залежними**.

Якщо випадкові величини X і Y є незалежними, то

$$f(x/y) = f(x) \quad \text{і} \quad f(y/x) = f(y) .$$

Якщо випадкові величини X і Y є залежними, то справедливо наступне співвідношення:

$$f(x,y) = f(x)*f(y/x) = f(y)*f(x/y) .$$

Звідки

$$f(y/x) = \frac{f(x,y)}{f(x)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dy} \quad \text{і} \quad f(x/y) = \frac{f(x,y)}{f(y)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dx} .$$

8.1.5. Числові характеристики випадкового вектора



Означення 8.5. Математичне сподівання випадкового вектора $\mathbf{X}=(X_1,X_2,\dots,X_n)$ є такий не випадковий вектор $\mathbf{m}=(m_1,m_2,\dots,m_n)$, компонентами якого є математичні сподівання відповідних компонент випадкового вектора \mathbf{X} .



Означення 8.6. Дисперсія випадкового вектора $\mathbf{X}=(X_1,X_2,\dots,X_n)$ є такий не випадковий вектор $\mathbf{D}=(D_1,D_2,\dots,D_n)$, компонентами якого є дисперсії відповідних компонент випадкового вектора \mathbf{X} .



Означення 8.7. Коваріацією k_{ij} випадкового вектора $\mathbf{X}=(X_1,X_2,\dots,X_n)$ називають другий змішаний центральний момент $M[(X_i - m_i)(X_j - m_j)]$.

Для $i=j$ коваріація співпадає з дисперсією відповідної компоненти випадкового вектора. Коваріацію k_{ij} також позначають як $Cov(X_i, X_j)$

Для двомірного випадкового вектора $\mathbf{Z}=(X,Y)$ коваріацією є другий змішаний центральний момент

$$k_{xy} = \mu_{11} = M[(X-m_x)(Y-m_y)] .$$

Для *дискретних* випадкових величин коваріація (або кореляційний момент) визначається за формулою

$$k_{xy} = \sum_{i=1}^n \sum_{j=1}^m (x_i - m_x)(y_j - m_y) p_{ij} , \quad (8.14)$$

де $p_{ij} = P(X=x_i, Y=y_j)$;

m_x – математичне сподівання компоненти X випадкового вектора \mathbf{Z} ;

m_y – математичне сподівання компоненти Y випадкового вектора \mathbf{Z} ;

n – кількість можливих значень компоненти X ;

m – кількість можливих значень компоненти Y .

Для *неперервних* випадкових величин коваріація визначається за формулою

$$k_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y) f(x, y) dx dy, \quad (8.15)$$

де $f(x,y)$ – щільність розподілу випадкового вектора \mathbf{Z} .

Кореляційний момент характеризує ступінь розкиду випадкових величин навколо їх математичного сподівання, а також ступінь лінійної залежності між випадковими величинами X і Y .

Для характеристики тільки ступеня лінійної залежності між випадковими величинами X і Y використовується **коефіцієнт кореляції**

$$r_{xy} = \frac{k_{xy}}{\sigma_x \sigma_y} . \quad (8.16)$$

Значення коефіцієнта кореляції r_{xy} знаходиться в діапазоні від -1 до $+1$. Якщо x і y є незалежними між собою величинами, то $r_{xy} = 0$.

Якщо x і y зв'язані лінійною залежністю $Y = aX + b$, то $r_{xy} = -1$ при $a < 0$ і $r_{xy} = 1$ при $a > 0$.

Для випадкового n -мірного вектора $\mathbf{X}=(X_1, X_2, \dots, X_n)$ задається n -мірна **кореляційна матриця**

$$\mathbf{K} = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1j} & \cdots & k_{1n} \\ k_{21} & k_{22} & \cdots & k_{2j} & \cdots & k_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ k_{i1} & k_{i2} & \cdots & k_{ij} & \cdots & k_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ k_{n1} & k_{n2} & \cdots & k_{nj} & \cdots & k_{nn} \end{bmatrix},$$

де $k_{ij} = M[(X_i - m_i)(X_j - m_j)]$;

$k_{ii} = M[(X_i - m_i)^2] = D_i$ – дисперсія i -го компонента випадкового вектора \mathbf{X} ;

$$k_{ij} = k_{ji}.$$

Для аналізу ступеня лінійної залежності між компонентами випадкового вектора \mathbf{X} використовується **нормована кореляційна матриця** \mathbf{R} , елементами якої є коефіцієнти кореляції r_{ij} відповідних компонент вектора \mathbf{X} ,

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2j} & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} & \cdots & r_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nj} & \cdots & r_{nn} \end{bmatrix},$$

$$\text{де } r_{ij} = \frac{k_{ij}}{\sigma_i \sigma_j}; \quad r_{ii} = \frac{D_i}{\sigma_i^2} = 1; \quad r_{ij} = r_{ji}.$$

8.2. Функції випадкових аргументів

Розв'язання багатьох прикладних задач вимагає знання законів розподілу або числових характеристик різних випадкових величин. У деяких випадках експеримент з виявлення закону розподілу вимагає постановки коштовних або тривалих за часом експериментів, а в деяких випадках і самий експеримент поставити неможливо. Часто цей бар'єр легко можна подолати. Якщо випадкова величина, що нас цікавить, є функцією випадкового аргументу, то її числові характеристики і закон розподілу можуть бути визначені за відомими характеристиками і законом розподілу випадкового аргументу, а також вигляду функціональної залежності.

8.2.1. Числові характеристики функції випадкових аргументів

Нехай випадкова величина Y є функцією випадкових аргументів $Y = \varphi(X_1, X_2, \dots, X_n) \dots$. Нехай відомий закон розподілу $g(y)$ функції випадкових аргументів. Тоді основні числові характеристики функції Y визначаються наступними виразами:

$$m_y = \int_{-\infty}^{\infty} y \cdot g(y) dy ; \quad (8.17)$$

$$D_y = \int_{-\infty}^{\infty} (y - m_y)^2 g(y) dy . \quad (8.18)$$

Однак, як уже наголошувалось, для визначення числових характеристик зовсім не обов'язково знати закон розподілу $g(y)$.

Нехай випадкова величина $Y = \varphi(X)$ є функцією випадкового дискретного аргументу X , для якого відомий закон розподілу у вигляді ряду розподілу

x_i	x_1	x_2	\dots	x_n
p_i	p_1	p_2	\dots	p_n

Тоді кожному значенню x_i можна поставити у відповідність значення $y_i = \varphi(x_i)$

x_i	x_1	x_2	\dots	x_n
p_i	p_1	p_2	\dots	p_n
$y_i = \varphi(x_i)$	$\varphi(x_1)$	$\varphi(x_2)$	\dots	$\varphi(x_n)$

} ряд розподілу X

} ряд розподілу Y

У загальному випадку для $y_i = \varphi(x_i)$ остання таблиця не є рядом розподілу (у точному розумінні цього терміна), однак усі необхідні для такого ряду значення випадкової функції і відповідні ймовірності в ній є. Таким чином,

$$m_y = M[\varphi(x)] = \sum_{i=1}^n \varphi(x_i) p_i . \quad (8.19)$$

Аналогічно для випадкової неперервної величини

$$m_y = M[\varphi(x)] = \int_{-\infty}^{\infty} \varphi(x) f(x) dx . \quad (8.20)$$

Для системи двох випадкових аргументів (8.19) і (8.20) матимуть відповідно вигляд:

$$m_z = M[\varphi(x, y)] = \sum_{i=1}^n \sum_{j=1}^n \varphi(x_i, y_j) p_{ij} ;$$

$$m_z = M[\varphi(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y) f(x, y) dx dy .$$

У загальному випадку (система з двох і більше випадкових аргументів) (8.19) і (8.20) матимуть відповідно вигляд:

$$m_y = M[\varphi(x_1, x_2, \dots, x_n)] = \sum_{i=1}^n \dots \sum_{j=1}^n \varphi(x_1, x_2, \dots, x_n) p_{i_1 i_2 \dots i_n} ;$$

$$m_y = M[\varphi(x_1, x_2, \dots, x_n)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \varphi(x_1, x_2, \dots, x_n) f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n .$$

Таким чином, математичне сподівання функції будь-якого числа випадкових аргументів може бути знайдене без знання закону розподілу $g(y)$.

Аналогічно можуть бути знайдені будь-які інші числові характеристики (моменти) функції випадкових аргументів. Наприклад, дисперсії

$$D_y = D[\varphi(x)] = \int_{-\infty}^{\infty} [\varphi(x) - m_y]^2 f(x) dx ,$$

$$D_z = D[\varphi(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\varphi(x, y) - m_y]^2 f(x, y) dx dy ,$$

$$D_y = D[\varphi(x_1, x_2, \dots, x_n)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [\varphi(x_1, x_2, \dots, x_n) - m_y]^2 f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n .$$

8.2.2. Теорема про числові характеристики функції випадкових аргументів

У багатьох випадках для знаходження числових характеристик функції випадкових аргументів не потрібно навіть знання закону розподілу випадкових аргументів. В основному це стосується лінійних і деяких елементарних нелінійних функцій.

Розглянемо обчислення математичного сподівання і дисперсії для найпростіших функцій випадкових аргументів.



Теорема 8.1. Математичне сподівання не випадкової величини c (константи) дорівнює самій не випадковій величині:

$$M[c] = c . \quad (8.21)$$

Доведення: $M[c] = \sum_{i=1}^n x_i p_i = c \cdot 1 = c$.



Теорема 8.2. Дисперсія не випадкової величини c дорівнює нулю:

$$D[c] = 0. \quad (8.22)$$

Доведення: $D[c] = M[(x_i - m_x)^2] = M[(c - c)^2] = M[0^2] = M[0] = 0$.



Теорема 8.3. Математичне сподівання добутку не випадкової величини c і випадкової величини X дорівнює добутку не випадкової величини c і математичного сподівання випадкової величини X :

$$M[cX] = cM[X]. \quad (8.23)$$

Доведення: $M[cX] = \sum_{i=1}^n cx_i = c \sum_{i=1}^n x_i = cM[X]$.



Теорема 8.4. Дисперсія добутку не випадкової величини c і випадкової величини X дорівнює добутку квадрата не випадкової величини c і дисперсії випадкової величини X :

$$D[cX] = c^2 D[X]. \quad (8.24)$$

Доведення:

$$\begin{aligned} D[cX] &= M[(cx - m_{cx})^2] = M[(cx - cm_x)^2] = M[c^2(x - m_x)^2] = \\ &= c^2 M[(x - m_x)^2] = c^2 D[X]. \end{aligned}$$

Наслідок теореми 8.4:

$$\sigma[cX] = \sqrt{D[cX]} = \sqrt{c^2 D[X]} = |c| \cdot \sigma_x. \quad (8.25)$$



Теорема 8.5. Математичне сподівання суми випадкових величин дорівнює сумі їх математичних сподівань:

$$M[X+Y] = M[X] + M[Y]. \quad (8.26)$$

Доведення:

$$\begin{aligned}
M[X + Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy = \\
&= \int_{-\infty}^{\infty} x \left\{ \int_{-\infty}^{\infty} f(x, y) dy \right\} dx + \int_{-\infty}^{\infty} y \left\{ \int_{-\infty}^{\infty} f(x, y) dx \right\} dy = \int_{-\infty}^{\infty} x \cdot f(x) dx + \int_{-\infty}^{\infty} y \cdot f(y) dy = \\
&= M[X] + M[Y].
\end{aligned}$$

Теорема 8.5 справедлива як для залежних, так і незалежних випадкових величин X і Y . Теорема 8.5 має також узагальнення на випадок суми декількох випадкових величин, тобто математичне сподівання суми n випадкових величин дорівнює сумі їх математичних сподівань:

$$M\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n M[X_i]. \quad (8.27)$$



Теорема 8.6. Математичне сподівання лінійної функції від n випадкових аргументів X_i ($i=1, 2, \dots, n$) дорівнює цій же лінійній функції від математичних сподівань випадкових величин X_i :

$$M[Y] = M\left[\sum_{i=1}^n a_i X_i + b\right] = \sum_{i=1}^n a_i M[X_i] + b. \quad (8.28)$$

Доведення. Твердження (8.28) очевидне з погляду на теореми 8.3, 8.5 і 8.1.



Теорема 8.7. Дисперсія суми випадкових величин X і Y дорівнює сумі їх дисперсій, збільшеній на подвоєний кореляційний момент цих же величин:

$$D[X + Y] = D[X] + D[Y] + 2k_{xy}. \quad (8.29)$$

Доведення:

$$\begin{aligned}
D[X + Y] &= M\left[\left((x + y) - m_{x+y}\right)^2\right] = M\left[\left(x + y - m_x - m_y\right)^2\right] = \\
&= M\left[\left(x - m_x\right)^2 + 2(x - m_x)(y - m_y) + (y - m_y)^2\right] = \\
&= M\left[\left(x - m_x\right)^2\right] + M\left[\left(y - m_y\right)^2\right] + 2M\left[(x - m_x)(y - m_y)\right] = D[X] + D[Y] + 2k_{xy}.
\end{aligned}$$

Наслідок теореми 8.7. Дисперсія суми незалежних величин дорівнює сумі їх дисперсій:

$$D[X+Y] = D[X] + D[Y] , \quad (8.30)$$

оскільки $k_{xy} = 0$.



Теорема 8.8. Дисперсія лінійної функції n випадкових незалежних аргументів X_i ($i=1,2,\dots,n$) визначається за формулою

$$D[Y] = D\left[\sum_{i=1}^n a_i X_i + b\right] = \sum_{i=1}^n a_i^2 D[X_i] . \quad (8.31)$$

Доведення. Формула (8.31) очевидна з погляду на теореми 8.4, 8.7 і 8.2.



Теорема 8.9. Математичне сподівання добутку випадкових величин X і Y визначається за формулою

$$M[XY] = M[X] * M[Y] + k_{xy} . \quad (8.32)$$

Доведення:

$$\begin{aligned} k_{xy} &= M[(X-m_x)(Y-m_y)] = M[XY] - m_x[Y] - m_y[X] + m_x m_y = \\ &= M[XY] - m_x m_y - m_x m_y + m_x m_y = M[XY] - m_x m_y . \end{aligned}$$

Звідки $M[XY] = M[X] * M[Y] + k_{xy}$.

Наслідок теореми 8.9. Математичне сподівання добутку незалежних випадкових величин дорівнює добутку їх математичних сподівань:

$$M[XY] = M[X] * M[Y] , \quad (8.33)$$

оскільки $k_{xy} = 0$.



Теорема 8.10. Дисперсія добутку незалежних випадкових величин X і Y визначається за формулою:

$$D[XY] = D[X] * D[Y] + m_y^2 D[X] + m_x^2 D[Y] . \quad (8.34)$$

Доведення:

$$\begin{aligned} D[XY] &= M[(xy - m_{xy})^2] = M[(xy - m_x m_y)^2] = M[x^2 y^2] - 2 m_x m_y M[xy] + m_x^2 m_y^2 = \\ &= M[x^2] * M[y^2] - 2 m_x m_y m_x m_y + m_x^2 m_y^2 = M[x^2] * M[y^2] - m_x^2 m_y^2 = \end{aligned}$$

$$= (D[X] + m_x^2) (D[Y] + m_y^2) - m_x^2 m_y^2 = D[X] * D[Y] + m_y^2 D[X] + m_x^2 D[Y] .$$

Наслідок теореми 8.10. При $m_x = 0$ і $m_y = 0$

$$D[XY] = D[X] * D[Y] . \quad (8.35)$$

8.2.3. Закон розподілу функції випадкових аргументів

В багатьох стохастичних задачах часто виникає необхідність визначити закон розподілу функції випадкового аргументу при відомому законі розподілу випадкового аргументу. Розглянемо таку задачу для монотонних функцій випадкового аргументу.

Нехай є випадкова неперервна величина X , яка розподілена в інтервалі (a, b) з щільністю розподілу $f(x)$. Нехай інша випадкова величина Y пов'язана з X функціональною залежністю $Y = \varphi(X)$. При цьому функція $\varphi(X)$ – функція монотонно **зростаюча** на інтервалі (a, b) , неперервна і має перші похідні (рис. 8.5). Потрібно знайти щільність розподілу $g(x)$ випадкової величини Y .

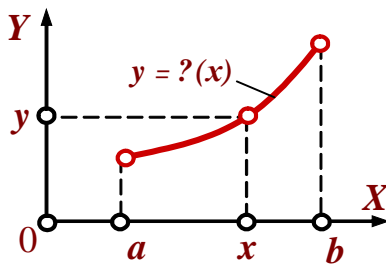


Рис. 8.5

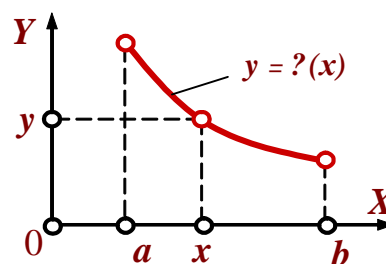


Рис. 8.6

Відповідно до означення 4.5 знайдемо інтегральну функцію випадкової величини Y

$$G(y) = P\{Y < y\} = P\{a < X < x\} = \int_a^x f(x) dx .$$

Визначимо x через y : $x = \varphi^{-1}(y)$, де φ^{-1} – функція, зворотна до функції φ . Тоді

$$G(y) = \int_a^{\varphi^{-1}(y)} f[\varphi^{-1}(y)] \cdot [\varphi^{-1}(y)]' dy .$$

Оскільки щільність розподілу $g(x)$ є похідною від інтегральної функції, то

$$g[y] = G'(y) = f[\varphi^{-1}(y)] \cdot [\varphi^{-1}(y)]' . \quad (8.36)$$

Нехай тепер функція $\varphi(X)$ – монотонно спадає на інтервалі (a,b) , неперервна і має перші похідні (рис. 8.6). Тоді

$$G(y) = P\{Y < y\} = P\{x < X < b\} = \int_x^b f(x)dx .$$

Визначимо x через y , тобто

$$G(y) = \int_{\varphi^{-1}(y)a}^b f[\varphi^{-1}(y)] \cdot [\varphi^{-1}(y)]' dy = - \int_b^{\varphi^{-1}(y)} f[\varphi^{-1}(y)] \cdot [\varphi^{-1}(y)]' dy ,$$

або

$$g[y] = G'(y) = -f[\varphi^{-1}(y)] \cdot [\varphi^{-1}(y)]' . \quad (8.37)$$

Враховуючи (8.36) і (8.37), узагальнена формула для щільності розподілу монотонної функції випадкового аргументу прийме остаточний вигляд:

$$g[y] = G'(y) = f[\varphi^{-1}(y)] \cdot |[\varphi^{-1}(y)]'| . \quad (8.38)$$

Дійсно, якщо $\varphi(x)$ – зростаюча функція, то похідна $\varphi'(x)$ невід'ємна, звідки похідна $[\varphi^{-1}(y)]'$ теж позитивна. Якщо $\varphi(x)$ – спадна функція, то похідна $\varphi'(x)$ негативна, звідки похідна $[\varphi^{-1}(y)]'$ теж від'ємна. Але знак «–» у (8.37) робить результат невід'ємним. Отже узагальнена формула щільності розподілу (8.38) справедлива в обох випадках.

Тема 9. ГРАНИЧНІ ТЕОРЕМИ

9.1. Закон великих чисел

9.1.1. Теорема Бернуллі

Якщо проводиться n незалежних випробувань, у кожному з яких випадкова подія A з'являється з імовірністю $P(A) = p$, то відносна частота μ/n появи події A (μ – число появ A) при великому n приблизно дорівнює ймовірності p :

$$\frac{\mu}{n} \approx p .$$

Наведене твердження можна уточнити в такий спосіб: $\frac{\mu}{n} \rightarrow p$ при $n \rightarrow \infty$, якщо для будь-якого $\varepsilon > 0$ і для досить великих n співвідношення

$$\left| \frac{\mu}{n} - p \right| < \varepsilon \quad (9.1)$$

виконується з імовірністю, що прямує до 1 з ростом n . Математичний запис даного твердження має вигляд

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} \rightarrow 1, \text{ якщо } n \rightarrow \infty. \quad (9.2)$$

Вираз (9.2) є формальним змістом *теорема Бернуллі*, відомої як *закон великих чисел*.



Теорема 9.1 (теорема Бернуллі). З імовірністю, як завжди близькою до 1, очікується, що при досить великій кількості випробувань відносна частота появи події буде як завжди мало відрізнятися від її ймовірності.

Зауважимо, що теорема не стверджує, що співвідношення (9.1) вірогідно, однак, якщо число n досить велике, то ймовірність виконання (9.1) близька до 1 (наприклад, 0.98 або 0.999), що **практично вірогідно**. Іншими словами, якщо проводиться експеримент, що складається з досить великого числа n випробувань, то можна бути впевненим, що співвідношення (9.1) буде виконано.

Примітка. Автори рекомендують читачеві перевірити останнє твердження за допомогою експерименту з киданням монети (подія A – випадіння «орла») або киданням гральної кістки (подія A – випадіння парного числа).

9.1.2. Закон великих чисел у формі Чебишова

9.1.2.1. Нерівність Чебишова

При будь-якому $\varepsilon > 0$

$$P(|\xi - M[\xi]| \geq \varepsilon) \leq \frac{D[\xi]}{\varepsilon^2}, \quad (9.3)$$

тобто абсолютне відхилення випадкової величини від її математичного сподівання більше або дорівнює ε з імовірністю, не більшою за відношення дисперсії цієї випадкової величини до квадрата ε .

З нерівності (9.3) виходить закон великих чисел у формі Чебишова.

9.1.2.2. Теорема Чебишова

Одне з основних тверджень закону великих чисел полягає в тому, що значення середньоарифметичного $\frac{1}{n} \sum_{i=1}^n \xi_i$ випадкових величин з рівними

математичними сподіваннями $M[\xi_i] = a$ при великому n виявляється приблизно рівним a :

$$\frac{1}{n} \sum_{i=1}^n \xi_i \approx a.$$

Надалі будемо говорити, що $\frac{1}{n} \sum_{i=1}^n \xi_i \rightarrow a$ при $n \rightarrow \infty$, якщо для будь-якого $\varepsilon > 0$ і досить великих n співвідношення

$$\left| \frac{1}{n} \sum_{i=1}^n \xi_i - a \right| < \varepsilon \quad (9.4)$$

виконується з імовірністю, що прямує до одиниці з ростом n . Дане висловлення записується в такий спосіб:

$$P \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - a \right| < \varepsilon \right\} \rightarrow 1, \text{ якщо } n \rightarrow \infty.$$

Це одне з тверджень закону великих чисел. Зауважимо, як і теорема Бернуллі, воно не означає, що співвідношення (9.4) вірогідне. Однак, якщо n досить велике, то ймовірність його виконання близька до 1, наприклад, 0.98 або 0.999, що означає його **практичну вірогідність**. Наведемо повне формулювання однієї з теорем закону великих чисел – *теорему Чебишова*.



Теорема 9.2 (теорема Чебишова). Якщо $\xi_1, \dots, \xi_n, \dots$ – послідовність попарно незалежних випадкових величин, що мають кінцеві дисперсії та обмежені однією і тією ж постійною:
 $D[\xi_1] < c, \quad D[\xi_2] < c, \quad \dots, \quad D[\xi_n] < c$, то для будь-якого $\varepsilon > 0$

$$P \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n M[\xi_i] \right| < \varepsilon \right\} \rightarrow 1, \text{ якщо } n \rightarrow \infty.$$

7.1.2.3. Перевірка закону великих чисел

Перевірка співвідношення (9.4) за допомогою різних експериментів, як правило, приводить до позитивного результату, тобто до його виконання. Однак слід звернути увагу на наявні випадки порушення закону великих чисел.

Розглянемо випадкову величину, яка розподілена за законом Коші з щільністю

$$p(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}. \quad (9.5)$$

Зауважимо, що щільність симетрична відносно нуля, однак 0 не є математичним сподіванням. Цей розподіл не має математичного сподівання.

Нагадаємо, що математичним сподіванням є $\int_{-\infty}^{\infty} xp(x)dx$, якщо

$\int_{-\infty}^{\infty} |x|p(x)dx < \infty$. Але остання нерівність для розподілу Коші не виконується.

Як наслідок, для послідовності незалежних випадкових величин, розподілених за законом Коші (9.5), закон великих чисел не виконується. Якби

середньоарифметичне $\bar{\xi}_n \equiv \frac{1}{n} \sum_{i=1}^n \xi_i$ сходилося з ростом n до якої-небудь

константи, то, з огляду на симетрію розподілу, такою константою міг бути тільки 0. Однак 0 не є точкою збіжності. Дійсно, можна показати, що при будь-якому $\varepsilon > 0$ і при будь-якому як завгодно великому n

$$\left| \frac{1}{n} \sum_{i=1}^n \xi_i \right| > \varepsilon \quad (9.6)$$

з імовірністю $P = 1 - \frac{2}{\pi} \arctg \varepsilon$.

Пояснимо сказане: можна показати, що середньоарифметичне $\bar{\xi}_n$ розподілене за законом (9.5), а функція розподілу для (9.5) є $\arctg x$. Ця ймовірність, як видно, не прямує до 0 з ростом n . Наприклад, якщо $\varepsilon = 0.03$, то ймовірність виконання (9.6) дорівнює приблизно $P \approx 0.98$, тобто подія (9.6) практично вірогідна, і можна впевнено очікувати її виконання з одного разу. Якщо $\varepsilon = 1$, то ймовірність (9.6) дорівнює 0.5, і виконання його хоча б один раз можна впевнено очікувати, зробивши 7 експериментів (тому що ймовірність невиконання жодного разу дорівнює $(0.5)^7 = 1/128$). І це при будь-якому фіксованому n , наприклад, $n = 200$. Експериментальна перевірка (рис. 9.1) підтверджує сказане.

Слід звернути увагу на те, що є рідкі спостереження, які відстоять дуже далеко від центру розподілу – точки 0.

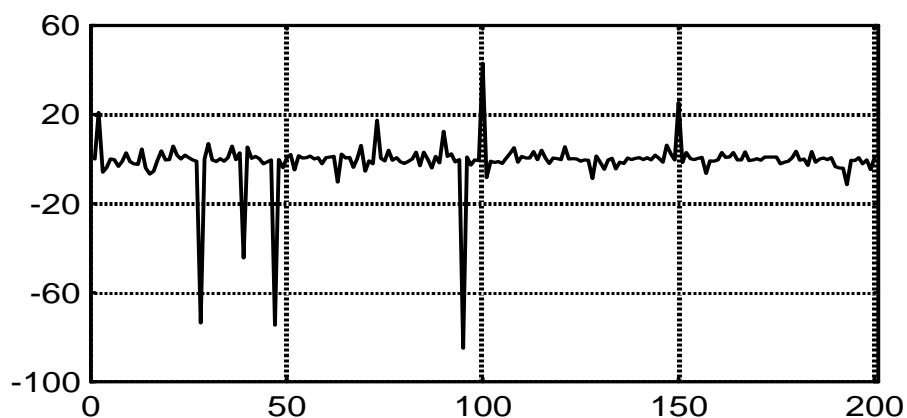


Рис. 9.1 – Вибірка спостережень, розподілених за законом Коші ($n = 200$)

7.1.2.4. Стиск розподілу з ростом числа доданків

Закон великих чисел у формі Чебишова означає, що розподіл випадкової величини

$$\bar{\xi}_n = \frac{1}{n} \sum_{i=1}^n \xi_i$$

стискується з ростом n . Якщо математичні сподівання однакові, тобто $M\xi_i = a$, то стиск відбувається навколо точки a .

Перекоонатися в стиску можна, спостерігаючи гістограми при різних значеннях n (наприклад, для $n = 10, 40, 160, 640$). Згенеруємо k разів (наприклад, $k = 20$) випадкову величину $\bar{\xi}_n \equiv \bar{\xi} : \bar{x}_1, \dots, \bar{x}_k$ і побудуємо для кожної такої вибірки середніх гістограму. Порівнюючи гістограми для різних n , можна легко помітити стиск (табл. 9.1 і рис. 9.2).

Таблиця 9.1 – Розсіяння середніх

n	$\bar{x} \min$	$\bar{x} \max$	w
10	0.371	0.687	0.32
40	0.418	0.606	0.19
160	0.472	0.550	0.08
320	0.523	0.469	0.05

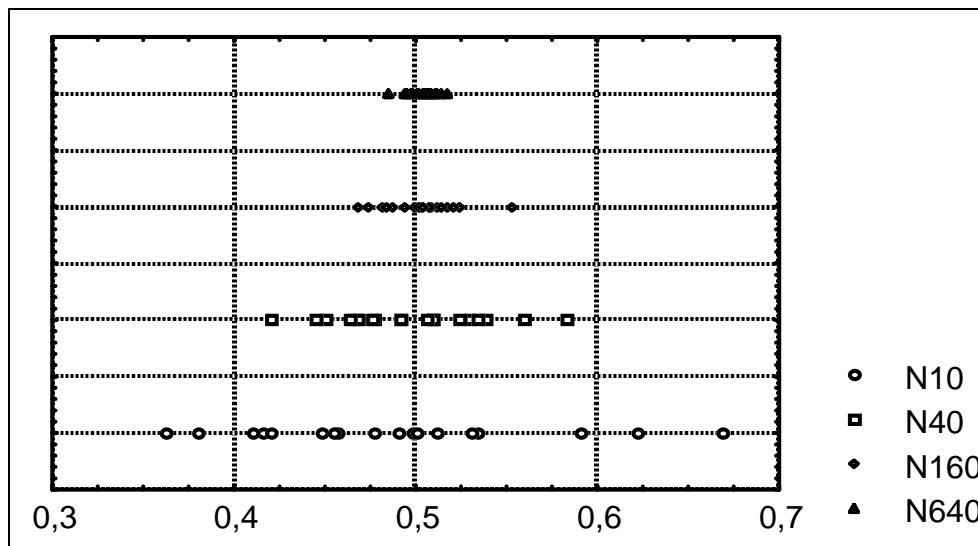


Рис. 9.2 – Гістограми розсіяння середніх при різних n

9.2. Посилений закон великих чисел.

9.2.1. Теорема Бореля



Теорема 9.3 (теорема Бореля). Відносна частота $f_n \equiv \frac{\mu_n}{n}$ появи випадкової події A з ростом числа n незалежних випробувань прямує до дійсної ймовірності p події A

$$\frac{\mu_n}{n} \rightarrow p \quad (9.7)$$

з імовірністю 1.

Іншими словами, при будь-якому експерименті з нескінченним числом випробувань має місце збіжність послідовності f_n до p . У справедливості сказаного можна переконатися за допомогою експерименту з кидання монети або гральної кістки. В останньому випадку для спрощення і прискорення експерименту слід розглядати подію A_1 (поява непарного числа очок) або A_2 (поява парного числа очок).

На рис. 9.3 наведені три графіки відносних частот f_n випадіння герба при киданні монети, отриманих у результаті натурних випробувань при $n=50$; на рис. 9.4 – при $n=500$.

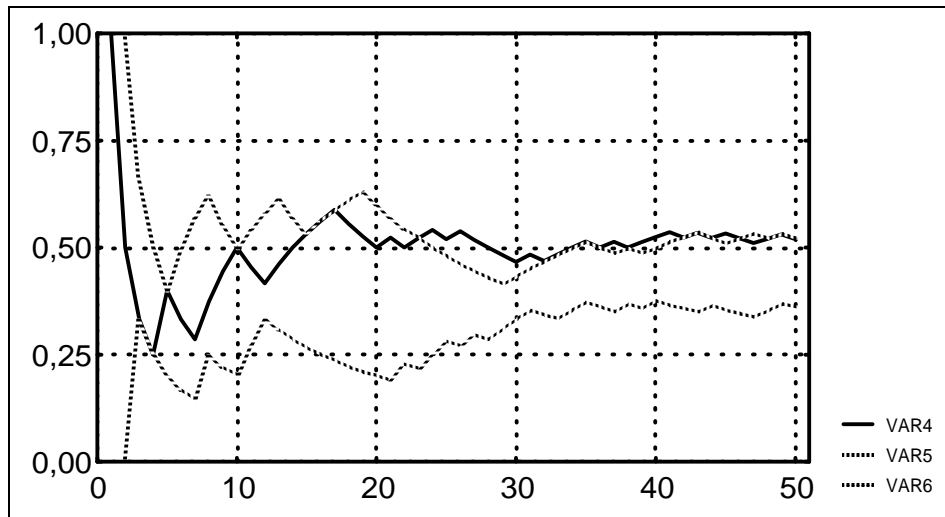


Рис. 9.3 – Відносна частота випадіння герба при зміні n від 1 до 50

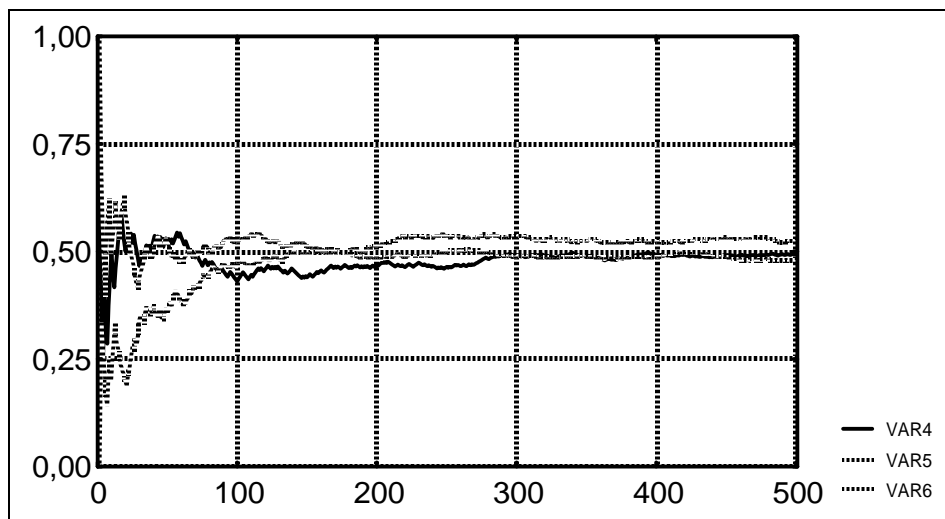


Рис. 9.4 – Відносна частота випадіння герба при зміні n від 1 до 500

Графіки переконливо свідчать про прямування відносних частот до дійсної ймовірності 0,5. Причому чим більша кількість випробувань, тим більш очевидне наближення f_n до p .

Будемо говорити, що послідовність випадкових величин $\xi_1, \xi_2, \dots, \xi_n$ підкоряється посиленому закону великих чисел, якщо за умову $n \rightarrow \infty$

$$\frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n M[\xi_i] \rightarrow 0 \quad (9.8)$$

з імовірністю 1.

В окремому випадку, при рівних математичних сподіваннях, $M[\xi_i]=a$, це означає

$$\frac{1}{n} \sum_{i=1}^n \xi_i \rightarrow a, \text{ якщо } n \rightarrow \infty \quad (9.9)$$

з імовірністю 1.

На рис. 9.5 наведені графіки результатів трьох експериментів з відстеження послідовностей середніх арифметичних випадкових величин ξ_i , рівномірно розподілених в інтервалі (0;1), коли i змінюється від 1 до 50; на рис. 9.6 – від 1 до 500. Графіки переконливо демонструють прямування середніх арифметичних до математичного сподівання 0.5 з ростом n .

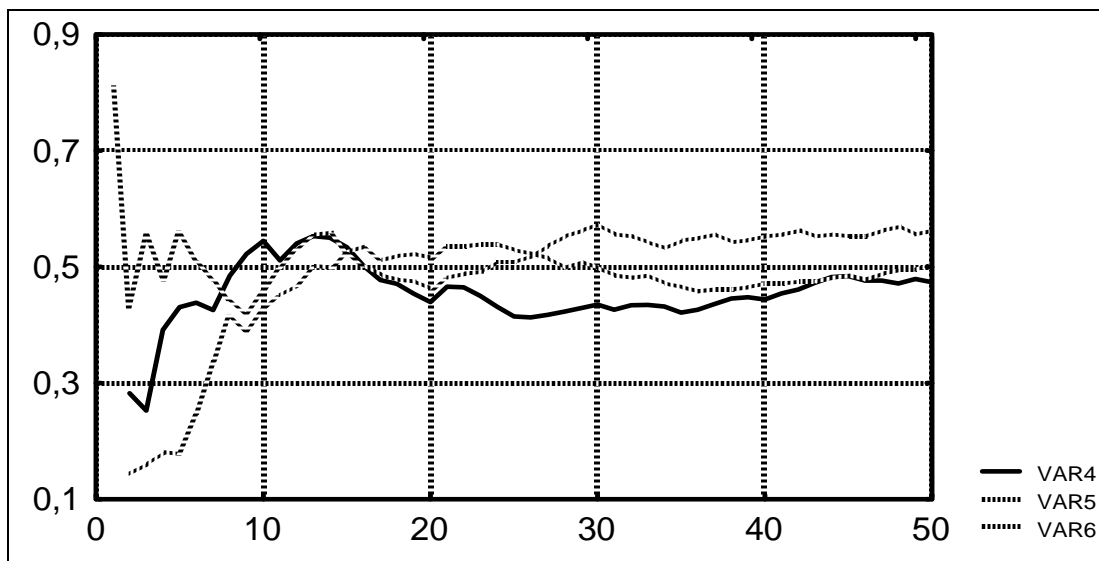


Рис. 9.5 – Графіки середніх арифметичних для рівномірно розподілених випадкових величин на інтервалі (0; 1) залежно від їх кількості ($i = \overline{1,50}$)

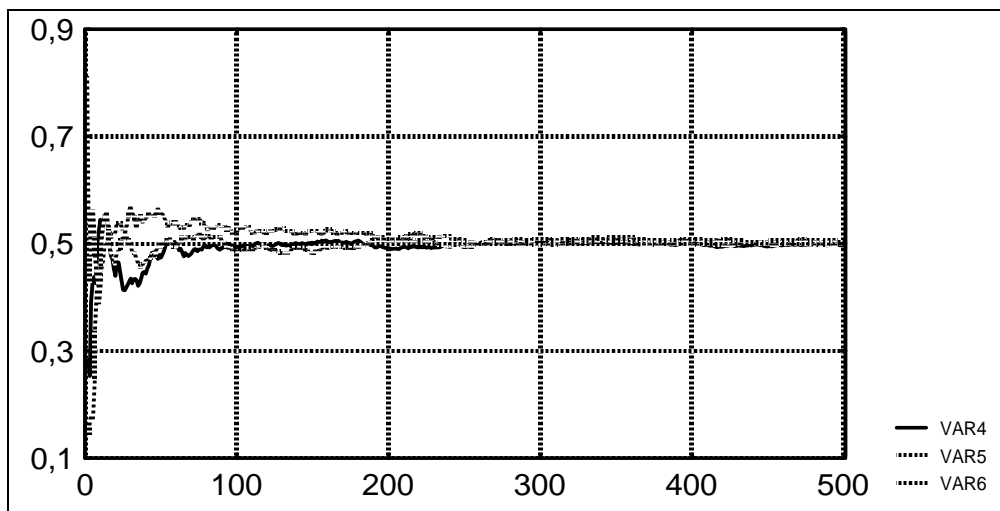


Рис. 9.6. – Графіки середніх арифметичних для рівномірно розподілених випадкових величин на інтервалі (0; 1) залежно від їх кількості ($i = \overline{1,500}$)

Достатню умову виконання (9.8) дає *теорема Колмогорова*.

9.2.2. Теорема Колмогорова



Теорема 9.4 (теорема Колмогорова). Якщо послідовність взаємно незалежних випадкових величин $\xi_1, \dots, \xi_n, \dots$ задовольняє умові

$$\frac{1}{n^2} \sum_{n=0}^{\infty} D[\xi_n] < \infty,$$

то вона підкоряється посиленому закону великих чисел.

Для незалежних і однаково розподілених випадкових величин теорема Колмогорова трансформується в більш просту теорему.



Теорема 9.5 (теорема Колмогорова у спрощеному трактуванні). Необхідною і достатньою умовою для застосовності посиленого закону великих чисел до послідовності незалежних величин є існування математичного сподівання.

9.2.3. Основна теорема статистики

Нехай x_1, x_2, \dots, x_n – вибірка з n незалежних спостережень над випадковою величиною X з теоретичною (дійсною, справжньою) функцією розподілу $F(x)$. Розташуємо спостереження в порядку зростання; одержимо варіаційний ряд $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$.

Визначимо функцію емпіричного розподілу

$$F_n^*(x) \equiv F_n^*(x : x_1, x_2, \dots, x_n) = \frac{\mu_n(x)}{n},$$

де $\mu_n(x)$ – число тих спостережень, для яких $x_i < x$. Ясно, що $F_n^*(x)$ – східчаста функція; що утворюється, якщо значенням x_1, \dots, x_n надати ймовірності, рівні $1/n$. До того ж, $F_n^*(x)$ – функція випадкова, тому що залежить від спостережень x_1, \dots, x_n .



Теорема 9.6 (теорема Глiвенко – основна теорема статистики). З ростом n максимальне абсолютне відхилення емпіричної функції розподілу від теоретичної (дійсної) прямує до нуля з імовірністю 1:

$$P\left(\sup_x |F_n^*(x) - F(x)| \xrightarrow{n \rightarrow \infty} 0\right) = 1.$$

Проілюструємо цю теорему на прикладах спостережень випадкової величини X , розподіленої за рівномірним законом на інтервалі $(0; 1)$, при числі випробувань $n=10$ (рис. 9.7), $n=40$ (рис. 9.8) і $n=160$ (рис. 9.9) .

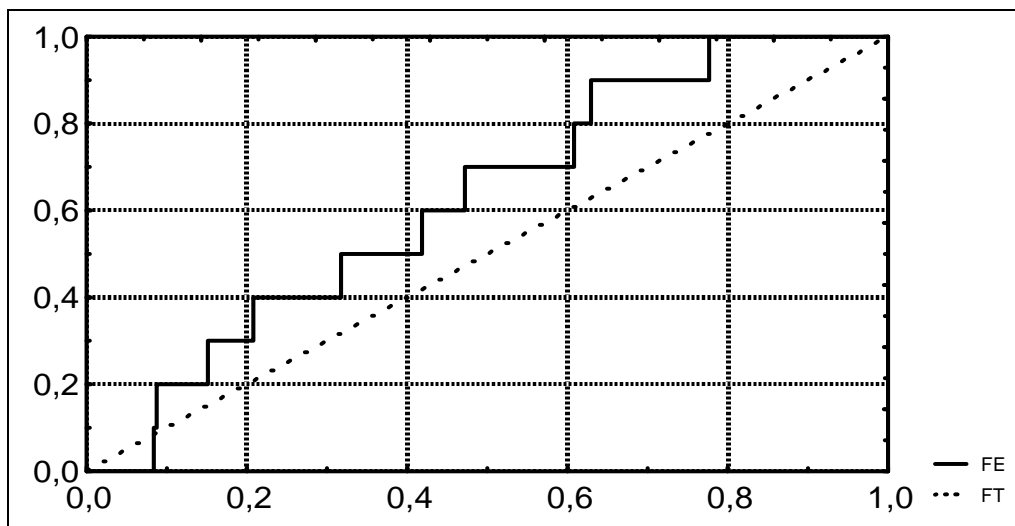


Рис. 9.7 – Функції емпіричного FE і теоретичного FT розподілів рівномірно розподіленої випадкової величини X при числі спостережень $n=10$

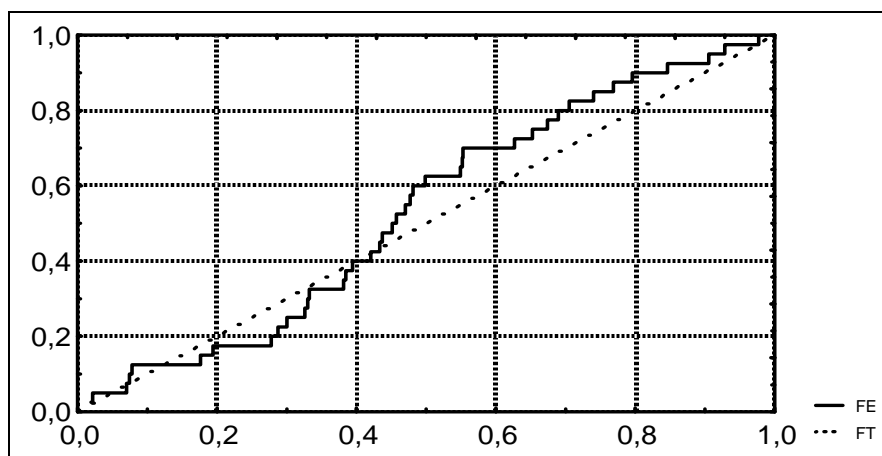


Рис. 9.8 – Функції емпіричного FE і теоретичного FT розподілів рівномірно розподіленої випадкової величини X при числі спостережень $n=40$

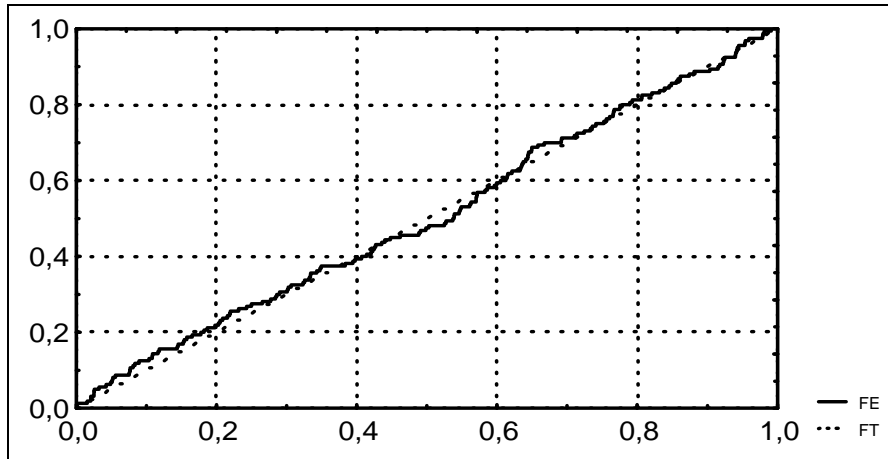


Рис. 9.9 – Функції емпіричного FE і теоретичного FT розподілів рівномірно розподіленої випадкової величини X при числі спостережень $n=160$

7.3. Центральна гранична теорема

9.3.1. Зміст центральної граничної теореми

Закон великих чисел стверджує, що при $n \rightarrow \infty$ середньоарифметичне випадкових величин ξ_i з рівними математичними сподіваннями прямує до їх математичного сподівання

$$\frac{1}{n} \sum_{i=1}^n \xi_i \rightarrow a ,$$

де $a = M[\xi_i]$.

Центральна гранична теорема стверджує дещо більше, а саме, що при $n \rightarrow \infty$ розподіл середньоарифметичного випадкових величин (за багаторазовим підсумовуванням середньоарифметичне є випадкова величина) наближається до нормального з параметрами a (математичне сподівання) і σ^2/n (дисперсія) :

$$\frac{1}{n} \sum_{i=1}^n \xi_i \sim N(a, \frac{\sigma^2}{n}) , \quad (9.10)$$

де $\sigma^2 = D[\xi_i]$.

При нормуванні суми гранична теорема записується в такий спосіб:

$$\frac{\sum_{i=1}^n \xi_i - na}{\sigma \sqrt{n}} \sim N(0,1) .$$

Наведемо формулювання центральної граничної теореми у формі Ліндеберга.

9.3.2. Теорема Ліндеберга



Теорема 9.7 (теорема Ліндеберга). Якщо послідовність взаємно незалежних випадкових величин $\xi_1, \xi_2, \dots, \xi_n, \dots$ при будь-якому постійному $\tau > 0$ задовольняє умові Ліндеберга

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{|x-a_k| > \tau B_n} (x-a_k)^2 dF_k(x) = 0, \quad (9.11)$$

де $a_k = M[\xi_k]$, $B_n^2 = D\left[\sum_{k=1}^n \xi_k\right]$, то при $n \rightarrow \infty$

$$P\left\{\frac{\sum_{k=1}^n (\xi_k - a_k)}{\sqrt{D\left[\sum_{k=1}^n \xi_k\right]}} < x\right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz, \quad (9.12)$$

тобто нормована сума випадкових величин $\xi_1, \xi_2, \dots, \xi_n$ розподілена за нормальним законом з параметрами 0 (математичне сподівання) і 1 (дисперсія).

9.3.3. Теорема Ляпунова

Умова Ліндеберга (9.11) досить універсальна, але незручна при практичній перевірці. Замість неї доцільно використовувати умову Ляпунова: при деякому $\delta > 0$

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^{2+\delta}} \sum_{k=1}^n M\left[|\xi_k - a_k|^{2+\delta}\right] = 0. \quad (9.13)$$

Для нормованих величин умова Ляпунова має вигляд

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n M\left[|\alpha_k|^{2+\delta}\right] = 0. \quad (9.14)$$



Теорема 9.8 (теорема Ляпунова – центральна гранична теорема у формі Ляпунова). Якщо $s_n = \xi_1 + \xi_2 + \dots + \xi_n$ – сума незалежних випадкових величин, $A_n = M[s_n]$, $B_n^2 = D[s_n]$ і виконується умова (9.13), то розподіл випадкової величини s_n наближається до нормального з параметрами A_n і B_n^2 .

На практиці частіше використовується умова Ляпунова

а) при $\delta=1$:

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^3} \sum_{k=1}^n M[|\xi_k - a_k|^3] = 0;$$

б) при $\delta=2$:

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^4} \sum_{k=1}^n M[|\xi_k - a_k|^4] = 0.$$

9.3.4. Сума однаково розподілених доданків

Центральним граничним теоремам підпорядковані послідовності випадкових величин з різними законами розподілу. На практиці частіше має місце більш простий випадок – послідовності випадкових величин з однаковими законами розподілу або послідовності реалізацій однієї і тієї ж випадкової величини.

Наслідок центральної граничної теореми. Якщо незалежні випадкові величини $\xi_1, \xi_2, \dots, \xi_n$ однаково розподілені і мають кінцеву відмінну від нуля дисперсію, то виконується (9.12) і (9.13). При цьому розподіл суми $s_n = \xi_1 + \xi_2 + \dots + \xi_n$ з ростом n наближається до нормального з параметрами $A_n = M[s_n]$, $B_n^2 = D[s_n]$.

Переконаємося статистично в тому, що сума декількох випадкових величин розподілена приблизно за нормальним законом. Зробимо це на прикладі суми

$$S = \sum_{k=1}^m x_k \quad (9.15)$$

шести ($m = 6$) незалежних випадкових величин, що мають *beta*-розподіл з параметрами $a=b=0.5$, тобто з щільністю розподілу

$$p(x | a, b) = \frac{x^{a-1} (1-x)^{b-1}}{B(a, b)} = \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}}, \quad (9.16)$$

де $B(a, b) = \int_0^1 z^{a-1} (1-z)^{b-1} dz$ – beta-функція.

Щільність розподілу доданків при обраних значеннях параметрів має *U-образний* вигляд, дуже далекий від нормального. Переконаємося в цьому, побудувавши графік щільності (рис. 9.10).

Щоб **статистично** оцінити закон розподілу для суми S , слід багаторазово (N разів, наприклад, $N=500$), промоделювати підсумовування: одержимо S_1, S_2, \dots, S_N – вибірку для суми. Для цієї вибірки побудуємо гістограму і порівняємо її візуально з нормальною щільністю (рис. 9.11 – 9.13).

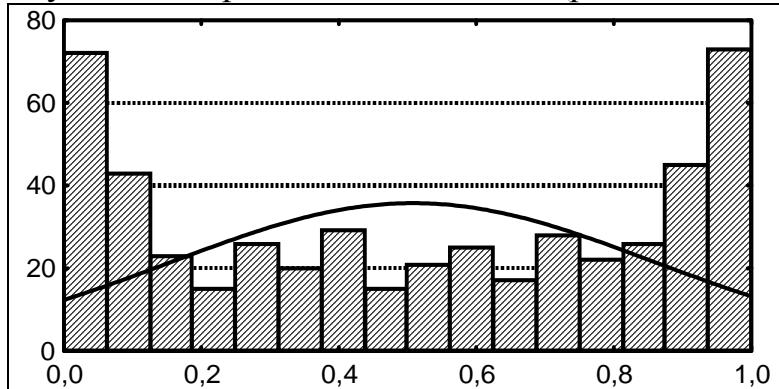


Рис. 9.10 – Гістограма одного доданка

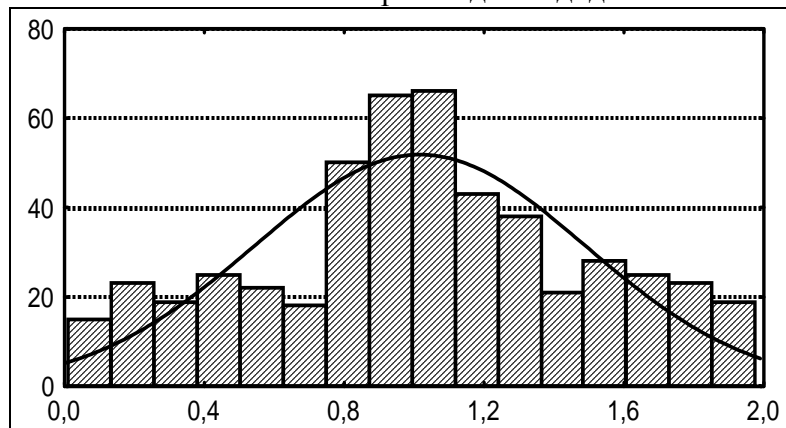


Рис. 9.11 – Гістограма суми двох доданків

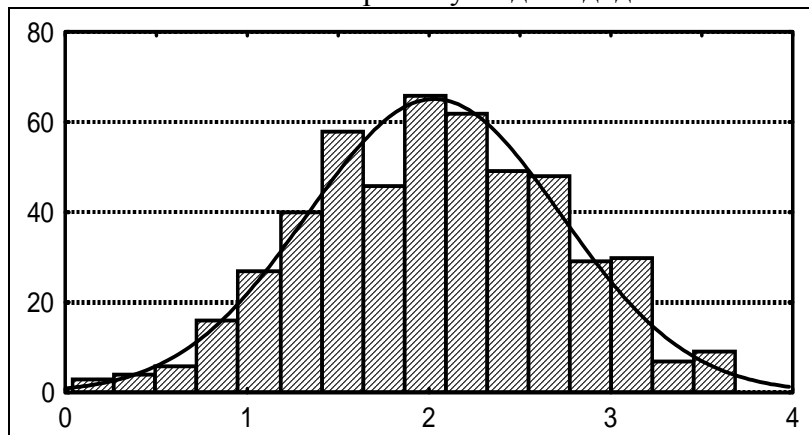


Рис. 9.12 – Гістограма суми чотирьох доданків

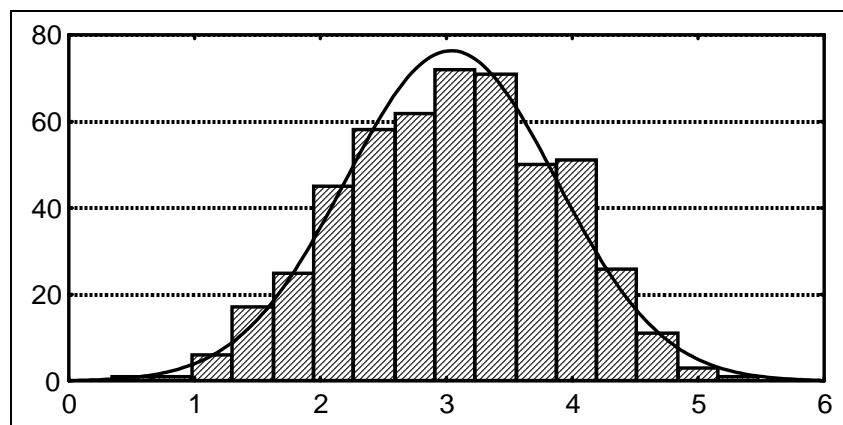


Рис. 9.13 – Гістограма суми шести доданків

На закінчення нагадаємо, що відповідно до центральної граничної теореми розподіл суми випадкових величин збігається до нормального і в тому випадку, коли доданки розподілені за *різними* законами розподілу.

Тема 10. ОСНОВНІ ПОНЯТТЯ МАТЕМАТИЧНОЇ СТАТИСТИКИ: ВИБІРКОВІ СПОСТЕРЕЖЕННЯ ТА ВИБІРКОВІ ОЦІНКИ

Математична статистика займається встановленням закономірностей, яким підпорядковані масові випадкові явища. Встановлення закономірностей ґрунтується на вивченні статистичних даних – результатів спостережень, які можна розглядати як реалізації деякої випадкової величини X .

Множина всіх значень випадкової величини x_1, x_2, \dots, x_N , що отримують в наслідок спостережень, утворює *генеральну сукупність*. Кількість спостережень N називають *обсягом* генеральної сукупності.

10.1. Основні положення вибіркового методу

Вибірковий метод – це система наукових принципів випадкового відбору певної частини генеральної сукупності з метою її дослідження та встановлення (оцінки) невідомих параметрів генеральної сукупності.

Вибірковий метод дозволяє суцільне обстеження генеральної сукупності звести до вибіркового обстеження. В основі теорії вибіркового спостереження лежать теореми закону великих чисел.

Сукупність одиниць, відібраних на основі науково розроблених принципів, називається *вибірковою сукупністю* або просто *вибіркою*. Завдання вибірки - дати вірне уявлення про характер генеральної сукупності, оцінити значення її параметрів.

Основні положення вибіркового методу передбачають, що вибіркова сукупність має досить великий обсяг.

Щоб властивості вибірки достатньо відображували властивості генеральної сукупності, вона повинна бути представничою – репрезентативною. Неточності несущільного спостереження називають *помилками репрезентативності*.

Оскільки вибіркова сукупність не точно відтворює склад генеральної, то й вибіркова оцінка не збігається з відповідними характеристиками генеральної. Розбіжність між ними називають *похибкою репрезентативності*. Розрізняють *випадкові* та *систематичні* помилки. Випадкові – наслідок випадковості відбору елементів для дослідження і пов'язаними з цим відмінностями структур вибіркової і генеральної сукупності щодо ознак, які вивчають. Систематичні мають місце тоді, коли при формуванні вибіркової сукупності порушується принцип випадковості, тобто здійснюється упереджений відбір елементів, або недосконала основа відбору їх називають помилками зміщення. В організації вибіркового обстеження важливо уникнути систематичних помилок. Випадкових помилок уникнути неможливо, але теорія вибіркового методу спостереження дає математичну основу для обчислення кількості та визначення напрямів їх зменшення.

Вибірковий метод відрізняється від інших методів несущільного спостереження (монографічного, основного масиву, анкетування) тим, що заздалегідь встановлюється кількість одиниць обстежень генеральної сукупності та їх порядок відбору, при якому вибіркова сукупність у достатній мірі репрезентувала б генеральну сукупність.

10.2. Способи відбору даних

Вибірка може бути *повторної* і *безповторною*. Повторною називають вибірку, при якій об'єкт (перед відбором наступного) повертається в генеральну сукупність. Безповторною називають вибірку, при якій відібраний об'єкт у генеральну сукупність не повертається.

Вибірка репрезентативна (представницька), якщо імовірнісні властивості вибірки збігаються з імовірнісними властивостями генеральної сукупності. Представницьку вибірку можна отримати, якщо вибирати об'єкти для дослідження випадково, тобто всі об'єкти генеральної сукупності з однаковою ймовірністю можуть піддаватися дослідженню (аналізу).

На практиці застосовують такі способи відбору (повторні або неповторні):

- *Простий* випадковий відбір – це класичний спосіб формування вибіркової сукупності. При цьому відбір одиниць у вибіркову сукупність провадиться випадково або з використанням таблиці випадкових чисел.

При цьому способі відбору для кожної одиниці сукупності створюються однакові умови, щоб вона могла потрапити у вибірку сукупності.

- *Механічний* відбір – це вибірка при якій генеральна сукупність. Заздалегідь поділяється на певне число рівних за кількістю одиниць груп, після чого із кожної групи відбирається для вибіркової сукупності тільки одна одиниця. Відбір одиниць у вибірку сукупності здійснюється через рівні інтервали. Механічний відбір має переваги перед простим випадковим відбором, оскільки при його проведенні до вибіркової сукупності обов'язково потраплять різні частини генеральної і результати є більш репрезентативними.

- *Типовий* відбір – це спосіб формування вибіркової сукупності залежно від складу генеральної сукупності. Типовою називається така вибірка, при якій генеральна сукупність заздалегідь поділяється на типові, якісно однорідні групи за істотними ознаками. Дуже часто цей спосіб відбору має назву пропорційного представництва, або методу кварт. Він найчастіше застосовується при вивченні громадської думки.

- *Серійний* – іноді доцільно проводити відбір до вибіркової сукупності не окремих одиниць, а їх груп, щоб в їх межах спостерігати усі без винятку одиниці сукупності. При серійному відборі основою вибірки є серія одиниць, яка розглядається і вивчається як одне ціле. Серійний відбір значно легше організувати і провести, але треба мати на увазі що він дає значно більшу похибку, і щоб забезпечити точність як і при інших способах відбору слід збільшити чисельність вибіркової сукупності.

10.3. Вибірковий розподіл. Емпіричний ряд розподілу. Гістограма та її властивості. Емпірична функція розподілу та її властивості

Вибірковий розподіл – це розподіл, який відповідає певній виборці з генеральної сукупності. Скільки існує вибірок – стільки маємо вибіркових розподілів. На практиці оперують лише з однією вибіркою, при цьому прагнуть отримати характеристики, які притаманні всій генеральній сукупності. Такі характеристики вдається отримати для репрезентативних вибірок.

Нехай з генеральної сукупності витягнута вибірка обсягом n , причому значення (*варіанта*) x_1 спостерігалася m_1 раз, x_2 – m_2 раз, ..., x_k – m_k раз, і

$n = \sum_{i=1}^k m_i$ (сума частот дорівнює обсягу вибірки). Послідовність варіант x_i ,

записаних у зростаючому порядку, сумісно з відповідною послідовністю

частот m_i утворюють *емпіричний ряд розподілу*. Величину $\frac{m_i}{n} = P_i^*$ назива-

ють *відносною частотою* або *емпіричною ймовірністю*. Ряд розподілу випадкової величини, що є якісною ознакою (наприклад ознака зі значення-

ми: відмінний, добрий, посередній), називають *атрибутивним рядом*. А ряд розподілу, що утворений зі значень числової випадкової величини, називають *варіаційним рядом*.

Таблиця 10.1 – Варіаційний ряд

Варіанта x_i	x_1	x_2	...	x_k	Обсяг вибірки
Частота m_i	m_1	m_2	...	m_k	$n = \sum_{i=1}^k m_i$

Побудова емпіричних графіків і діаграм дозволяє установити на першому етапі дослідження, до якого типу теоретичного розподілу скоріше за все відноситься отримане емпіричне розподіл.

Варіаційний ряд можна зобразити графічно у вигляді полігона частот або гістограми. Полігоном частот називають ламану, що сполучає точки з координатами (x_1, m_1) , (x_2, m_2) , ..., (x_k, m_k) . Полігон частот нагадує багатокутник розподілу.

Полігоном відносних частот називають ламану, відрізки якої сполучають точки з координатами (x_1, p_1^*) , (x_2, p_2^*) , ..., (x_k, p_k^*) (рис.10.1).

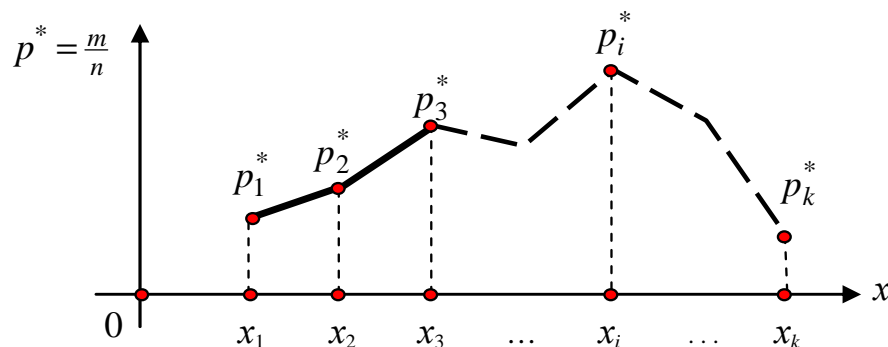


Рис.10.1 – Полігон відносних частот

Варіаційний ряд для неперервної випадкової величини (ознаки) частіше зображують у вигляді гістограми відносних частот.

Гістограмою відносних частот називають ступінчасту фігуру, що складається з прямокутників. В основі кожного прямокутника лежить інтервал Δx , а висота прямокутника h_i визначається відношенням $\frac{p_i^*}{\Delta x}$ (щільність відносної частоти).

Властивості гістограм. Площа i -го прямокутника гістограми дорівнює відносній частоті i -ї варіанти: $h_i \cdot \Delta x = p_i^*$. Площа гістограми відносних частот дорівнює сумі всіх відносних частот, тобто одиниці.

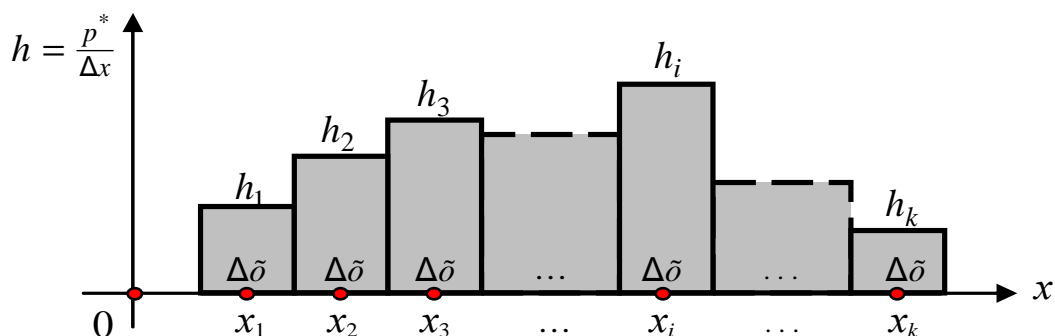


Рис.10.2 – Гістограма розподілу

Емпіричною функцією розподілу (комулятою) називають функцію $F^*(x)$, що визначає для кожного аргументу x відносну частоту події $X < x$. Комулята визначається співвідношенням $F^*(x) = \frac{n_x}{n}$, де $n_x = \sum_{i=1}^x n_i$ – число варіант, менших за x , n – обсяг вибірки (табл.10.2).

Таблиця 10.2 – Емпірична функція розподілу

x_i	x_1	x_2	...	x_k	
m_i	m_1	m_2	...	m_k	$\sum_{i=1}^k m_i = n$
p_i^*	$p_1^* = \frac{m_1}{n}$	$p_2^* = \frac{m_2}{n}$...	$p_k^* = \frac{m_k}{n}$	$\sum_{i=1}^k p_i^* = 1$
$F^*(x)$	p_1	$p_1 + p_2$...	$\sum_{i=1}^k p_i = 1$	

Інтегральну функцію розподілу генеральної сукупності називають *теоретичною* функцією розподілу. Різниця між емпіричною $F^*(x)$ та теоретичною $F(x)$ функціями розподілу полягає в тому, що теоретична функція визначає ймовірність події $X < x$, а емпірична функція визначає відносну частоту цієї ж події.

Графік емпіричної функції розподілу має східчастий вигляд. Із збільшенням кількості спостережень він стає більш гладким, а емпірична функція розподілу наближається до теоретичної функції розподілу генеральної сукупності.

Властивості емпіричної функції розподілу. Емпірична функція розподілу $F^*(x)$ володіє усіма властивостями інтегральної функції розподілу $F(x)$ випадкової величин, а саме:

- значення $F^*(x)$ належать відрітку $[0, 1]$;

- $F^*(x)$ – неспадна функція;
- Якщо x_1 – найменша варіанта, то $F^*(x) = 0$ при $x \leq x_1$;
- Якщо x_k – найбільша варіанта, то $F^*(x) = 1$ при $x > x_k$.

10.3. Статистичні оцінки (вибіркові моменти) та їх властивості. Збіжність статистичних оцінок

Вибіркові моменти в математичній статистиці – це оцінка теоретичних моментів розподілу на основі вибірки. Вибіркові моменти поділяють на початкові та центральні.

Нехай x_1, x_2, \dots, x_n – довільна вибірка обсягу n з певної генеральної сукупності. Тоді *початковим вибіровим моментом порядку k* є випадкова

величина $\alpha_k = \frac{1}{n} \sum_{i=1}^n x_i^k$. При $k=1$ початковий вибіровий момент являє собою

вибірове середнє $\tilde{x} = \alpha_1 = \frac{1}{n} \sum_{i=1}^n x_i$.

Центральний вибіровий момент порядку s – це випадкова величина

$$\mu_s = \frac{1}{n} \sum_{i=1}^n (x_i - \tilde{x})^s.$$

Генеральна сукупність має певний тип розподілу, який характеризується певними параметрами розподілу. За допомогою вибірок можна знайти оцінки цих параметрів. Для знаходження оцінок використовують статистики – функції від вибірових значень. Прикладами статистик є:

вибірове середнє $\tilde{x} = \frac{1}{n} \sum_{i=1}^n x_i$;

вибірова дисперсія $\tilde{\mu}_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \tilde{x})^2$;

вибіровий k -й початковий момент $\tilde{\alpha}_k = \frac{1}{n} \sum_{i=1}^n x_i^k$;

вибіровий k -й центральний момент $\tilde{\mu}_k = \frac{1}{n} \sum_{i=1}^n (x_i - \tilde{x})^k$ та ін.

Оскільки результати випробувань випадкові, будь-яка статистика являє собою випадкову величину. Щоб статистика могла служити оцінкою параметра θ , необхідно, щоб розподіл статистики був зосереджений в достатній близькості від невідомого значення параметра θ , тобто

ймовірність великих відхилень цієї статистики від θ була досить мала. Бажано також, щоб точність оцінювання збільшувалася при збільшенні обсягу вибірки.

Основними властивостями статистичних оцінок є спроможність, незміщеність і ефективність:

- оцінка повинна бути незміщеною $M(\hat{\theta}_n) = \theta$, тобто не мати систематичної похибки;

- оцінка має бути ефективною, тобто такою, що при заданому обсязі вибірки n має найменшу можливу дисперсію;

- оцінка θ_n повинна бути спроможною, тобто при постійному збільшенні обсягу вибірки n вона наближається до значення параметра θ , який оцінює.

Потрібно відзначити, якщо деяка оцінка є спроможною, то вона асимптотично незміщена, проте зворотне твердження не є вірним, тобто умова спроможності є сильнішою, ніж умова незміщеності.

Для статистичного оцінювання параметрів генеральної сукупності бажано використовувати оцінки, які задовольняють одночасно вимогам спроможності, незміщеності й ефективності. Саме такі статистичні оцінки (емпіричні моменти) збігаються до відповідних теоретичних моментів розподілу.

Вибіркове середнє \tilde{x} є незміщеною оцінкою, проте вибіркова дисперсія $\tilde{\mu}_2$ є зміщеною оцінкою. Тому її треба скорегувати. Так, вираз

$\frac{n}{n-1} \tilde{\mu}_2$ приводить зміщену оцінку до незміщеної.

Тема 11. МЕТОДИ ТОЧКОВОГО ТА ІНТЕРВАЛЬНОГО ОЦІНЮВАННЯ ПАРАМЕТРІВ

11.1. Точкові оцінки параметрів розподілу. Метод моментів

Точковою називають оцінку, яка визначається одним числом.

Найвідоміші методи знаходження оцінок: метод моментів, метод функції правдоподібності, метод найменших квадратів. Розглянемо детальніше метод моментів.

Метод моментів точкової оцінки невідомих параметрів заданого розподілу полягає в прирівнюванні теоретичних моментів розподілу емпіричним моментам цього ж розподілу.

Якщо розподіл визначається одним параметром, то для отримання точкової оцінки параметра прирівнюють один теоретичний момент одному-

емпіричному моменту. Наприклад, для показникового розподілу має місце одне рівняння $\frac{1}{\lambda} = \tilde{x}$. Звідки точковою оцінкою параметра λ буде дріб $\frac{1}{\tilde{x}}$.

Якщо розподіл визначається двома параметрами, то необхідно скласти систему двох рівнянь. Наприклад, для рівномірного розподілу має місце система рівнянь

$$\begin{cases} \frac{b-a}{2} = \tilde{x}; \\ \frac{(b-a)^2}{12} = \tilde{\mu}_2. \end{cases}$$

Звідки точковими оцінками параметрів будуть розв'язки системи відносно невідомих a та b .

11.2. Інтервальні оцінки параметрів розподілу. Метод моментів

На відміну від точкової оцінки, інтервальна оцінка визначається двома числами.

Точкові оцінки мають той недолік, що за ним не можна судити про точність одержуваних оцінок. Тому виникає необхідність визначення на підставі вибірових значень такого інтервалу (θ_1, θ_2) , який покривав би невідоме значення параметра θ із заданою ймовірністю.

Нехай $P(\theta_1 \leq \theta \leq \theta_2) = \alpha$, де випадковий інтервал (θ_1, θ_2) , який називається довірчим інтервалом, із заданою ймовірністю містить оцінюваний параметр θ . Величину α називають довірчим рівнем або надійністю. Величина $\delta = (\theta_2 - \theta_1)/2$ характеризує точність інтервальної оцінки. Зазвичай величину α беруть рівною 0,95, 0,99 чи 0,999. Величину $1 - \alpha$ називають рівнем значущості відхилення оцінки. Кінці довірчого інтервалу θ_1 і θ_2 називають довірчими границями.

Один з поширених методів побудови довірчих інтервалів полягає в наступному. За обраним значенням обчислюється незміщена точкова оцінка θ_n параметра θ . Будь-яким чином обчислюється дисперсія статистики θ_n або її оцінка σ_n^2 . Потім будують довірчий інтервал вигляду $(\theta_n - k_1 \sigma_n, \theta_n + k_2 \sigma_n)$, де k_1 і k_2 – коефіцієнти, значення яких визначають обраний довірчий рівень і апріорні припущення про розподіл генеральної сукупності (наприклад, нормальність або симетричність розподілу). Але оскільки такий інтервал визначається не однозначно, накладаються додаткові умови, щоб цей інтервал мав мінімальну довжину. Якщо розподіл статистики θ_n симетричний (або близький до симетричного), то довірчий інтервал мінімальної довжини виходить при $k_1 = k_2$. На такий

основі будується відомий критерій Стюдента для нормально розподілених сукупностей.

У самому загальному випадку (при мінімальних припущеннях щодо розподілу генеральної сукупності) довірчі інтервали можна побудувати на підставі нерівності Чебишова або інших подібних нерівностей. Однак такі інтервальні оцінки мають невелику точність.

11.3. Інтервальні оцінки за даними вибірки

Обмежуємося інтервальною оцінкою математичного сподівання.

Нехай для кількісної ознаки X з невідомим законом розподілу отримано вибірку обсягом n , тобто маємо вибіркиму середню $\bar{x} = \sum_{i=1}^n x_i$ та скор-

еговане квадратичне відхилення $s = \frac{n}{n-1} \tilde{s}_2 = \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$. Треба з ви-

значити довірчий інтервал, що покриває істинне значення ознаки X (математичне сподівання m_x) з довірчим рівнем (надійністю) γ .

Довірчий інтервал визначається подвійною нерівністю

$$\bar{x} - t_\gamma \frac{s}{\sqrt{n}} < m_x < \bar{x} + t_\gamma \frac{s}{\sqrt{n}},$$

де t_γ знаходять зі спеціальної таблиці по відомим n та γ .

Тема 12. ПЕРЕВІРКИ СТАТИСТИЧНИХ ГІПОТЕЗ

12.1. Загальний алгоритм перевірки гіпотез. Типи помилок

У прикладних задачах часто потрібно за емпіричними даними перевірити те чи інше припущення – гіпотезу. Якщо гіпотеза стосується виду невідомого розподілу або параметрів відомих розподілів, то її називає *статистичною*.

Як правило, висувається дві гіпотези – основна гіпотеза, наприклад, генеральна сукупність розподілена за рівномірним законом розподілу, і – конкуруюча (альтернативна) гіпотеза, наприклад, генеральна сукупність не розподілена за рівномірним законом розподілу.

Висунута гіпотеза може бути правильною або неправильною, тому виникає необхідність її перевірки. Оскільки перевірку проводять статистичними методами, її називають *статистичною*.

При перевірці можуть бути допущені помилки двох родів:

- помилка першого роду – відкинута правильна гіпотеза;

- помилка другого роду – прийнята неправильна гіпотеза.

Ймовірність зробити помилку першого роду позначають α , а ймовірність помилки другого роду – β . Будь-яку статистичну гіпотезу необхідно статистично перевірити.

Статистичні гіпотези перевіряються за допомогою статистичних критеріїв. Для перевірки основної гіпотези використовують спеціально підібрану випадкову величину K (статистику), точне або наближене значення якої відомо.

Критерієм згоди називають критерій перевірки гіпотези про належність розподілу генеральної сукупності конкретного класу розподілів. Є кілька критеріїв згоди: χ^2 (« χ -квадрат» або критерій Персона), критерій Колмогорова-Смирнова (застосовується тільки для безперервних розподілів) та інші.

Після вибору певного критерію, множину всіх його можливих значень розбивають на дві непересічні підмножини: одна з них містить значення критерію, при яких основна гіпотеза відкидається – *критична область*, а інша – при якому вона приймається, тобто *область прийняття гіпотези*.

Основний принцип перевірки статистичних гіпотез: якщо K належить критичній області – основну гіпотезу відкидають, в іншому випадку основну гіпотезу приймають.

Критичною точкою називають точку $k_{кр}$, яка відокремлює критичну область від області прийняття гіпотези. Розрізняють односторонню (правосторонню, лівосторонню) і двосторонню критичну області (рис.12.1).

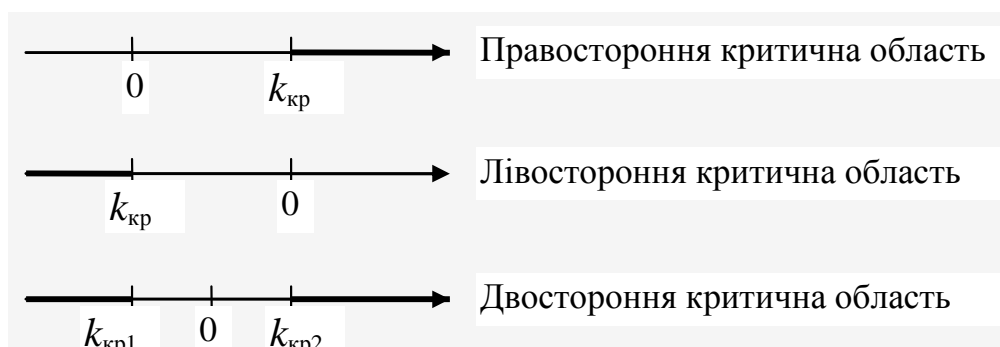


Рис.12.1 – Критичні області

Для перевірки статистичної гіпотези достатньо відшукати критичну точку $k_{кр}$ і порівняти її зі статистикою k

Для правосторонньої критичної області точку $k_{кр}$ шукають виходячи з вимоги, щоб ймовірність перебільшення статистики K значення $k_{кр}$ дорівнювала прийнятому рівню значущості, тобто $P[K > k_{кр}^2] = \alpha$.

12.2. Критерій згоди Пірсона

Перевагою критерію Пірсона є його універсальність – застосовується однаково при різних розподілах.

Пірсоном доведено, що при $n \rightarrow \infty$ закон розподілу випадкової величини $\chi^2 = \sum_{i=1}^n \frac{(m_i - m_i^T)^2}{m_i^T}$, незалежно від того, якому закону розподілу під-

порядкована генеральна сукупність, прагне до закону розподілу $\chi^2 = \sum_{i=1}^n \delta_i^2$

з $k = n$ ступенями свободи, де x_i ($i = \overline{1, n}$) – нормальні незалежні випадкові величини, причому математичне сподівання кожної з них дорівнює нулю, а середнє квадратичне відхилення – одиниці. Диференціальна функція розподілу χ^2 має такий вигляд:

$$f(x) = \begin{cases} 0, & x \leq 0, \\ \frac{1}{2^{\frac{k}{2}} \cdot \Gamma(\frac{k}{2})} e^{-\frac{x}{2}} \cdot x^{\frac{k}{2}-1}, & x > 0, \end{cases}$$

де $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ – гамма-функція; $\Gamma(n+1) = n!$

Зі збільшенням числа ступенів свободи розподіл χ^2 повільно наближається до нормального.

Ідея критерію Пірсона полягає в наступному: спостережувані (емпіричні) частоти m_i^T порівнюються з теоретичними m_i^T (обчислені в припущенні певного виду закону розподілу, що перевіряється). Зазвичай емпіричні та теоретичні частоти відрізняються. Можливо, що розбіжність частот випадкова (незначна) і пояснюється або малим числом спостережень, або способом їх угруповання, або іншими причинами. Можливо, розбіжність частот не випадкова (значна) і пояснюється тим, що теоретичні частоти обчислені виходячи з невірної гіпотези. Критерій Пірсона встановлює, узгоджуються або не узгоджуються дані спостережень з висунутою гіпотезою на прийнятому рівні значущості α (досить малої ймовірності помилки першого роду).

Критерій Пірсона має правосторонню критичну область. Ймовірність попадання критерію χ^2 в критичну область, в припущенні справедливості нульової гіпотези, дорівнює прийнятому рівню значущості:

$$P[\chi^2 > \chi_{кр}^2(\alpha, k)] = \alpha$$

Таким чином, правобічна критична область визначається нерівністю $\chi^2 > \chi_{кр}^2(\alpha, k)$, а область прийняття нульової гіпотези – нерівністю $\chi^2 \leq \chi_{кр}^2(\alpha, k)$.

Для того щоб при заданому рівні значущості, перевірити нульову гіпотезу H_0 – генеральна сукупність розподілена за певним законом треба:

1. У припущенні, що справедлива гіпотеза H_0 , обчислити теоретичні частоти m_i^T .

2. Обчислити спостережуване значення критерію $\chi^2 = \sum_{i=1}^n \frac{(m_i - m_i^T)^2}{m_i^T}$.

3. По заданому рівню значущості α , і числу ступенів свободи k ($k = s - 1 - r$, де s – число груп або часткових інтервалів вибірки, r – число параметрів передбачуваного розподілу) за таблицею розподілу критичних точок розподілу χ^2 знайти критичну точку $\chi_{кр}^2(\alpha, k)$.

4. Якщо $\chi^2 \leq \chi_{кр}^2$ – нема підстав відкидати нульову гіпотезу. Якщо $\chi^2 > \chi_{кр}^2$ – нульову гіпотезу треба відкинути.

Зауваження 1. Обсяг вибірки повинен бути досить великий, не менше 50. Кожна група повинна містити не менше 5-8 варіант, нечисленні групи слід об'єднати в одну, підсумовуючи частоти.

Зауваження 2. Оскільки можливі помилки першого і другого роду, особливо, якщо узгодження теоретичних та емпіричних частот «занадто добре», слід проявляти обережність. Наприклад, можна повторити обстеження, збільшити число спостережень, скористатися іншими критеріями, побудувати графік розподілу, обчислити асиметрію та ексцес.

Зауваження 3. З метою контролю обчислень, спостережуване значення критерію обчислюють за $\chi^2 = \sum \frac{m_i^2}{m_i^T} - n$.

Змістовий модуль 2. МАТЕМАТИЧНЕ ПРОГРАМУВАННЯ

Тема 13. ПРЕДМЕТ МАТЕМАТИЧНОГО ПРОГРАМУВАННЯ

Математичне програмування як прикладний розділ вищої математики відіграє винятково важливу роль у підготовці фахівців управлінського та економічного профілю. Використання математичних методів в професійній діяльності дозволяє вирішувати оптимальним способом багато менеджерських та економічних задач. Іншими словами, майбутній

фахівець стає власником надійного інструменту для одержання найвищого економічного ефекту в конкретних виробничих умовах.

Прикладами можливих, що наочно ілюструють корисність і необхідність знання запропонованої дисципліни, можуть бути наступні задачі (надаються в змістовній постановці):

- одержання максимального випуску продукції або максимального прибутку при заданих матеріальних, трудових або тимчасових витратах;
- забезпечення планових показників підприємства при мінімальних фінансових витратах;
- досягнення максимально короткого терміну виготовлення продукції, будівництва об'єкта, товарообігу, виробничого циклу і т.п. при існуючих або заданих виробничих ресурсах (матеріальних, трудових, енергетичних, тимчасових та ін.).



У наведених прикладах максимальний випуск продукції, максимальний прибуток, мінімальні фінансові вкладення, максимально короткий термін є шукані *оптимуми* (*максимуми* або *мінімуми*).

У математиці максимум і мінімум мають ще одну назву – *екстремум*, а задачі пошуку екстремуму називають екстремальними задачами.



У наведених прикладах умови, що накладаються на вирішення задачі (задані матеріальні, трудові й тимчасові витрати; планові показники; виробничі ресурси), називають *обмеженнями* задачі.

Обмеження задач визначають *область припустимих рішень*.



Ті припустимі рішення, при яких досягається оптимум, називають *оптимальними* або *екстремальними рішеннями*.

У загальному випадку екстремальна задача може мати одне, декілька, безліч, нескінчену безліч або жодне оптимальне рішення.



У практиці менеджера або економіста оптимальне рішення прийнято називати *оптимальним планом*.

Змістова постановка задачі повинна дозволяти переходити до строгої *математичної моделі*. У протилежному випадку необхідно пройти досить трудомісткі й копіткі процеси математичного моделювання та ідентифікації, але останні тут не розглядаються.



У загальному вигляді екстремальна задача формулюється наступним чином: визначити найбільше (максимальне) або найменше (мінімальне) значення деякої функції $Y(x_1, x_2, \dots, x_n)$ при умовах $f_i(x_1, x_2, \dots, x_n) \leq b_i$ ($i = \overline{1, m}$), де y і f_i – задані функції, а b_i – дійсні числа.

Наведене формулювання є узагальненням постановок ряду приватних задач математичного програмування, що можуть розрізнятися між собою як видом функцій y і f_i , так і характером перемінних (дискретний, безупинний).



Функцію $Y(x_1, x_2, \dots, x_n)$, яку оптимізують (мінімізують або максимізують), називають *цільовою функцією*.

Залежно від особливостей функцій y і f_i математичне програмування можна розділити на ряд самостійних дисциплін, що вивчають і розробляють методи вирішення окремих класів задач.

Насамперед, задачі математичного програмування поділяють на задачі *лінійного* і *нелінійного програмування*. При цьому, якщо всі функції y і f_i є лінійними, то відповідні задачі відносяться до класу задач лінійного програмування. Якщо ж хоча б одна із зазначених функцій є нелінійною, то відповідна задача відноситься до класу задач нелінійного програмування.

Найбільш вивченим розділом математичного програмування є лінійне програмування. Для вирішення задач лінійного програмування розроблено цілий ряд ефективних методів, алгоритмів і програм.

Серед задач нелінійного програмування найбільш глибоко вивчені задачі *опуклого програмування*. Це задачі, в результаті вирішення яких визначається екстремум опуклої функції, заданої на опуклій замкнутій множині.

У свою чергу, серед задач опуклого програмування більш докладно досліджені задачі *квадратичного програмування*. У результаті вирішення таких задач потрібно в загальному випадку знайти екстремум квадратичної функції при обмеженнях на змінні у вигляді системи лінійних рівнянь або лінійних нерівностей.

Визначеними класами задач математичного програмування є задачі *цілочислового*, *параметричного* і *дрібно-лінійного програмування*.

У задачах *цілочислового*, або *дискретного* програмування невідомі можуть приймати тільки цілочислові значення.

У задачах параметричного програмування цільова функція, або функції, що визначають область можливих змін перемінних, або то і інше залежать від деяких параметрів.

У задачах дрібно-лінійного програмування цільова функція являє собою відношення двох лінійних функцій, а функції, що визначають область припустимих рішень, також є лінійними.

Особливі класи становлять задачі *стохастичного* і *динамічного програмування*. Якщо в цільовій функції або функціях-обмеженнях, містяться випадкові величини, то така задача відноситься до класу задач стохастичного програмування. Задача, процес знаходження якої є багатоетапним, належить до класу задач динамічного програмування.

Таким чином, *математичне програмування* є математичною дисципліною, що досліджує екстремальні задачі і розробляє методи їх вирішення. Математичне програмування як наука знаходиться у процесі постійного розвитку. Вченими всього світу розроблено багато методів для вирішення різних класів задач математичного програмування. Разом з тим багато задач ще не мають ефективних методів розв'язання і чекають своїх дослідників.

На рис.13.1 подана графічна ілюстрація нелінійної оптимізаційної задачі однієї змінної щодо пошуку максимально можливої відстані польоту снаряда s_{\max} при відомій початковій швидкості.

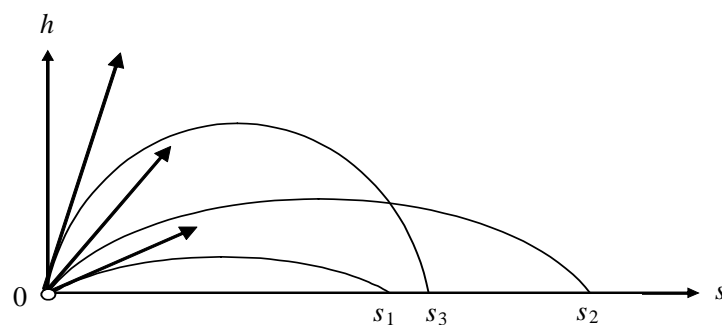


Рис.13.1 – Приклад екстремальної задачі з однією змінною

На рис. 13.2 подана графічна ілюстрація нелінійної оптимізаційної задачі двох змінних щодо пошуку Місяцехідом найнижчої точки на поверхні Місяця та графічна ілюстрація простішої стратегії її розв'язання (методу покоординатного спуску). Тут третя вісь Y_0 прямує на читача, при цьому значення еквіпотенціальних рівнів постійно зменшуються, тобто $y_1 > y_2 > y_3 > \dots$.

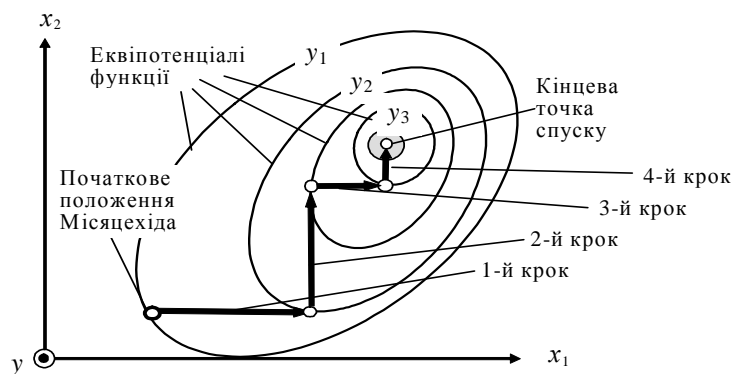


Рис.13.1 – Приклад екстремальної задачі з двома змінними

Тема 14. ЛІНІЙНЕ ПРОГРАМУВАННЯ

Лінійне програмування – найбільш вивчений розділ математичного програмування. Задача лінійного програмування є частинним випадком загальної задачі математичного програмування. Така задача досить часто зустрічається в інженерно-економічній і фінансовій діяльності. Диференціальний алгоритм вирішення загальної задачі математичного програмування може бути з успіхом використаний для розв’язання задач лінійного програмування. Через значне спрощення алгоритму оптимізації внаслідок лінійності цільової функції і обмежень він викликає особливий інтерес і заслуговує окремого детального розгляду.

14.1. Постановка задачі лінійного програмування

Загальна задача лінійного програмування формулюється таким чином: *знайти оптимум лінійної функції $y(\mathbf{x})$, якщо на змінні задачі накладені лінійні обмеження у вигляді рівностей і нерівностей, а також обмеження у вигляді гіперпаралелепіпеда.*

Аналітичний запис цієї задачі такий:

$$y(\mathbf{x}) = \mathbf{c}^T \mathbf{x} + c_0 \rightarrow \underset{\mathbf{x} \in \Omega \subset \mathbf{R}^n}{\text{opt}}, \quad (14.1)$$

$$\Omega: \mathbf{A}_1 \mathbf{x} + \mathbf{b}_1 \leq 0; \quad (14.2)$$

$$\mathbf{A}_2 \mathbf{x} + \mathbf{b}_2 = 0; \quad (14.3)$$

$$\mathbf{A}_3 \mathbf{x} + \mathbf{b}_3 \geq 0; \quad (14.4)$$

$$\mathbf{x} \geq 0, \quad (14.5)$$

де \mathbf{x} – n -вимірний вектор дійсних змінних; \mathbf{c} – n -вимірний вектор коефіцієнтів функції цілі; c_0 – вільний член функції цілі; $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ – матриці коефіцієнтів лінійних систем розмірності $m_1 \times n, m_2 \times n, m_3 \times n$ відповідно, $m_2 < n$; $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ – вектори вільних членів обмежень розмірності $m_1 \times 1, m_2 \times 1, m_3 \times 1$ відповідно.

Частинні задачі лінійного програмування можуть не містити однієї або двох систем обмежень типу (14.2) – (14.5), однаково яких. Крім того, замість умови невід’ємності (14.5) може мати місце двостороння або одностороння обмеженість змінних.

Задачу, складену з (14.1), (14.2) і (14.5), називають *стандартною задачею лінійного програмування*.

Канонічна, або основна задача лінійного програмування, має вигляд:

$$y(\mathbf{x}) = \mathbf{c}^T \mathbf{x} + c_0 \rightarrow \underset{\mathbf{x} \in \Omega \subset \mathbf{R}^n}{\max}, \quad (14.6)$$

$$\Omega: \mathbf{A} \mathbf{x} + \mathbf{b} = 0; \quad (14.7)$$

$$\mathbf{x} \geq 0, \quad (14.8)$$

де \mathbf{A} – матриця коефіцієнтів розмірності $m \times n$, $m < n$; \mathbf{b} – вектор вільних членів розмірності $m \times 1$.

Очевидно, що обмеження-нерівність типу " \leq " можна перетворити на обмеження-рівність додаванням до його лівої частини додаткової невід'ємної змінної, а кожне обмеження-нерівність типу " \geq " – в обмеження-рівність відніманням з його лівої частини додаткової невід'ємної змінної. Задачу мінімізації лінійної функції y множенням останньої на -1 можна звести до задачі максимізації. Таким чином, задачу лінійної оптимізації (14.1) – (14.5) завжди можна перетворити на задачу (14.6) – (14.8) і навпаки.

Укладання математичної моделі загальної задачі математичного програмування або її канонічної форми вимагає певних зусиль. Особливий інтерес являє собою векторне-матричне подання канонічної задачі лінійного програмування, що дозволяє легко переходити до таблиці диференціального алгоритму.

Розглянемо декілька прикладів одержання математичних моделей задач лінійного програмування поки без їхнього остаточного вирішення.

Приклад 14.1. Скласти математичну модель задачі лінійного програмування такого змісту:

Для виготовлення 3-х видів виробів A , B , C використовується токарне, фрезерне, зварювальне і шліфувальне устаткування. Витрати часу на обробку одного виробу для кожного з типів устаткування зазначені в табл.14.1. У ній же поданий загальний фонд робочого часу кожного з типів використовуваного устаткування, а також прибуток від реалізації одного виробу кожного виду. Треба визначити, скільки виробів і якого виду слід виготовити підприємству, щоб прибуток від їхньої реалізації був максимальним.

Таблиця 14.1

Тип Устаткування	Витрати часу на обробку одного виробу, верстата-год.			Загальний фонд робо- чого часу устаткуван- ня, год.
	A	B	C	
Фрезерне	2	4	5	120
Токарське	1	8	6	280
Зварювальне	7	4	5	240
Шліфувальне	4	6	7	360
Прибуток, грн.	10	14	12	–

Розв'язання. Припустимо, що буде виготовлено x_1 одиниць виробів виду A , x_2 – виду B і x_3 – виду C . Тоді для виробництва такої кількості виробів потрібно затратити $2x_1 + 4x_2 + 5x_3$ верстата-годин фрезерного устаткування. Оскільки загальний фонд робочого часу верстатів даного типу не

може перевищувати 120, повинна виконуватися нерівність $2x_1 + 4x_2 + 5x_3 \leq 120$.

Аналогічні міркування щодо можливого використання токарського, зварювального і шліфувального устаткування приведуть до таких нерівностей:

$$x_1 + 8x_2 + 6x_3 \leq 280 ;$$

$$7x_1 + 4x_2 + 5x_3 \leq 240 ;$$

$$4x_1 + 6x_2 + 7x_3 \leq 360 .$$

При цьому кількість виготовлених виробів не може бути від'ємною:

$$x_1 \geq 0; \quad x_2 \geq 0; \quad x_3 \geq 0 . \quad (14.9)$$

Далі, якщо буде виготовлено x_1 одиниць виробів виду A , x_2 – виду B і \tilde{o}_3 – виду C , то прибуток становитиме

$$y = 10x_1 + 14x_2 + 12x_3 . \quad (14.10)$$

Підприємство зацікавлене одержувати максимальний прибуток від реалізації своєї продукції. Отже, необхідно серед усіх невід'ємних рішень системи нерівностей

$$\begin{aligned} 2x_1 + 4x_2 + 5x_3 &\leq 120; \\ x_1 + 8x_2 + 6x_3 &\leq 280; \\ 7x_1 + 4x_2 + 5x_3 &\leq 240; \\ 4x_1 + 6x_2 + 7x_3 &\leq 360 \end{aligned} \quad (14.11)$$

знайти таке, при якому функція (14.10) набуває максимального значення.

Оскільки функція (14.10) лінійна і система (14.11) містить тільки лінійні нерівності, розглянута задача є задача лінійного програмування, яка в алгебраїчній формі запишеться таким чином:

$$o = 10\tilde{o}_1 + 14\tilde{o}_2 + 12\tilde{o}_3 \rightarrow \max_{\tilde{o}_1, \tilde{o}_2, \tilde{o}_3 \in \Omega}$$

$$2x_1 + 4x_2 + 5x_3 \leq 120;$$

$$x_1 + 8x_2 + 6x_3 \leq 280;$$

$$7x_1 + 4x_2 + 5x_3 \leq 240;$$

$$4x_1 + 6x_2 + 7x_3 \leq 360;$$

$$\tilde{o}_1, \tilde{o}_2, \tilde{o}_3 \geq 0 .$$

14.2. Графічне розв'язання задачі лінійного програмування

Припустима множина рішень задачі лінійного програмування Ω утворює опуклий гіпербагатогранник, на границі якого знаходиться оптимум функції цілі $y(x)$.

Якщо задача лінійного програмування, записана у формі стандартної (14.1), (14.2) і (14.5), містить не більше двох змінних, або задача, записана в канонічній формі (14.6) – (14.8), містить не більше $m + 2$ змінних, де m –

число лінійно незалежних обмежень-рівностей, то оптимальну граничну точку \mathbf{x}^* із Ω легко знайти графічним способом. Якщо задача задана в канонічній формі, її треба попередньо перетворити в стандартну, тобто в задачу лінійного програмування такого вигляду:

$$y(\mathbf{x}) = [c_1 \quad c_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \underset{\mathbf{x} \in \Omega \subset \mathbf{R}^2}{\text{opt}}, \quad (14.13)$$

$$\Omega: \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{12} & \alpha_{22} \\ \dots & \dots \\ \alpha_{m1} & \alpha_{m2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix} \leq 0; \quad (14.14)$$

$$\mathbf{x} \geq 0. \quad (14.15)$$

Через двовимірний простір припустимих рішень гіпербагатогранник Ω задачі (14.13) – (14.15) трансформується у багато. Сторони багатокутника утворюються прямими $a_{i1}x_1 + a_{i2}x_2 + b_i = 0$, $i = \overline{1, m}$; $x_1 = 0$ і $x_2 = 0$, відповідні рівняння для яких утворюються з початкової системи обмежень (14.14) – (14.15) шляхом заміни знаків нерівностей « \leq » і « \geq » на знак точної рівності « $=$ ».

Оскільки оптимальне рішення є граничною точкою багатокутника рішень, графічне вирішення задачі лінійного програмування полягає у графічній побудові багатокутника рішень і відшукування на його границі точки \mathbf{x}^* , що оптимізує функцію цілі. Останнє досягається побудовою лінії рівня (еквіпотенціалі) $c_1x_1 + c_2x_2 = y_0$ (y_0 – константа), що проходить через багатокутник рішень, і її переміщенням паралельно самої собі в напрямку вектора $\mathbf{c}^T = [c_1 \quad c_2]$ доти, поки вона не досягне крайньої граничної точки перетину з багатокутником. У досягнутій точці цільова функція має максимальне значення. При переміщенні лінії рівня у протилежному напрямку координати крайньої граничної точки перетину визначають точку мінімуму функції y . При цьому можливі різні випадки, зображені на рис.14.1 – 14.4. Область припустимих рішень на кожному рисунку виділена сірим фоном.

На рис.14.2 показано випадок, коли функція приймає максимальне значення в єдиній точці, на рис.14.3 – у будь-якій точці на відрізку AB . На рис.14.4 зображено випадок, коли функція, що максимізується, не обмежена зверху на області припустимих рішень, а на рис.14.5 – випадок, коли система обмежень задачі несутісна.

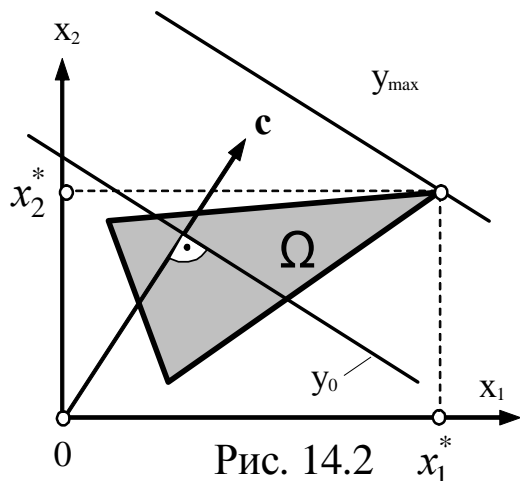


Рис. 14.2

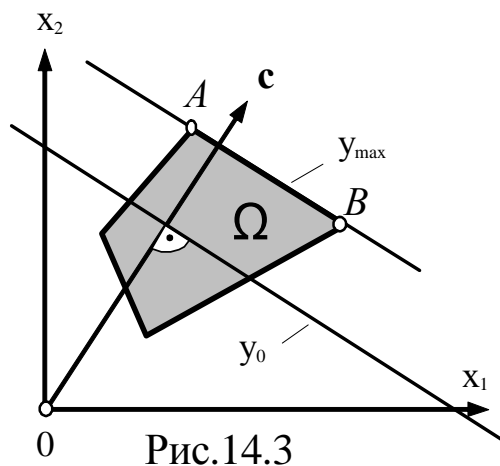


Рис. 14.3

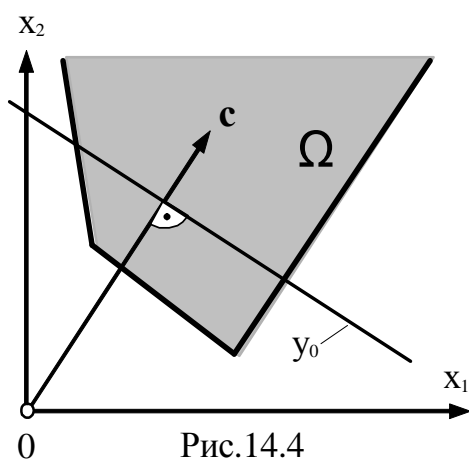


Рис. 14.4

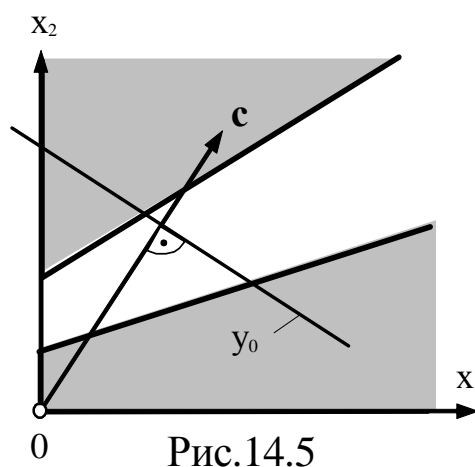


Рис. 14.5

На рисунках розглянуті можливі випадки при пошуку максимуму цільової функції. Аналогічні ситуації можуть виникнути і при пошуку мінімуму.

Послідовність розв'язання задачі лінійного програмування графічним способом, як правило, включає такі етапи.

1. Приведення математичної моделі задачі до вигляду (14.13) – (14.15).

2. Побудова прямих, визначених рівняннями

$$a_{i1}x_1 + a_{i2}x_2 + b_i = 0, \quad i = \overline{1, m}; \quad x_1 = 0 \text{ і } x_2 = 0.$$

3. Знаходження напівплощин, визначених кожним з обмежень задачі.

4. Виділення багатокутника рішень (області припустимих рішень Ω).

5. Побудова прямої $y_0 = c_1x_1 + c_2x_2$, що проходить через багатокутник рішень.

6. Побудова вектора $c^T = [c_1 \ c_2]$.

7. Переміщення прямої $y_0 = c_1x_1 + c_2x_2$ в напрямку (у зворотному напрямку) вектора c до границі області Ω .

8. Визначення координат граничної точки й обчислення значення цільової функції в цій точці, що є максимальною (мінімальною).

Приклад 14.2. Для виробництва двох видів виробів A і B підприємство використовує три види сировини. Норми витрати сировини кожного виду наведені в табл.14.2. У ній же зазначено прибуток від реалізації одного виробу кожного виду і загальна кількість сировини, яка може бути використана підприємством.

Таблиця 14.2

Вид сировини	Норми витрати сировини на один виріб, кг		Загальна кількість сировини, кг
	A	B	
I	12	4	300
II	4	4	120
III	3	12	252
Прибуток від реалізації одного виробу, <i>гр.</i>	30	40	–

Вироби A і B можуть виготовлятися в будь-яких співвідношеннях (збут забезпечений). Треба скласти такий план їхнього випуску, при якому прибуток підприємства від реалізації усіх виробів буде максимальним.

Розв’язання. Припустимо, що буде виготовлено x_1 одиниць виробів виду A , x_2 – виду B . Оскільки виробництво продукції обмежено наявністю у розпорядженні підприємства сировиною кожного виду і кількість виготовлених виробів не може бути від’ємною, повинні виконуватися нерівності

$$12x_1 + 4x_2 \leq 300 ;$$

$$4x_1 + 4x_2 \leq 120 ;$$

$$3x_1 + 12x_2 \leq 252 ;$$

$$x_{1,2} \geq 0 .$$

Загальний прибуток від реалізації x_1 виробів виду A і x_2 – виду B складе $y = 30x_1 + 40x_2$.

Таким чином, приходимо до такої математичної задачі: серед усіх невід’ємних рішень складеної системи лінійних нерівностей треба знайти таке, при якому функція y набуває максимального значення.

Вирішимо сформульовану задачу лінійного програмування графічним способом. Оскільки задача подана в стандартній формі, відпадає необхідність у першому етапі вирішення. Почнемо розв’язання задачі з побудови багатокутника рішень, для чого всі знаки нерівностей в системі обмежень замінимо точними рівностями і побудуємо відповідні прямі:

$$12x_1 + 4x_2 = 300 ;$$

$$4x_1 + 4x_2 = 120 ;$$

$$3x_1 + 12x_2 = 252 ;$$

$$x_1 = 0 ;$$

$$x_2 = 0 .$$

Ці прямі зображені на рис.14.6. Кожна з побудованих прямих поділяє площину на дві напівплощини. Координати точок однієї напівплощини задовольняють початковій нерівності, а іншою – ні. Щоб визначити шукану напівплощину, потрібно взяти точку, що належить однієї з площин, і перевірити, чи задовольняють її координати цій нерівності. Якщо координати точки задовольняють даній нерівності, шуканою є та напівплощина, якій належить ця точка, у протилежному разі – інша напівплощина.

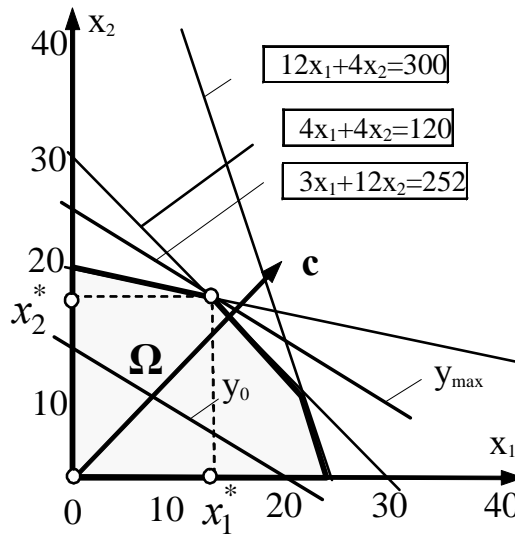


Рис.14.6

Знайдемо, наприклад, напівплощину, визначену нерівністю $12x_1 + 4x_2 \leq 300$. Для цього, побудувавши пряму $12x_1 + 4x_2 = 300$, візьмемо будь-яку точку, що належить однієї з двох отриманих напівплощин. Нехай це буде точка початку координат $x_0^T = [0 \ 0]$. Координати цій точці задовольняють нерівності $12 \cdot 0 + 4 \cdot 0 < 300$. Виходить, напівплощина, якій належить точка початку координат x_0 , відповідає нерівності $12x_1 + 4x_2 \leq 300$.

Аналогічно робимо при пошуку інших напівплощин. Переріз знайдених напівплощин визначає опуклий багатокутник припустимих рішень. На рисунку багатокутник виділений сірим фоном. Будь-яка точка з області задовольняє системі обмежень задачі.

Поставлена задача буде вирішена, якщо ми зможемо знайти таку точку з області Ω , в якій функція y приймає максимальне значення. Щоб знайти зазначену точку, побудуємо пряму $y_0 = 3x_1 + 4x_2$ (нехай $y_0 = 600$) і вектор $c^T = [30 \ 40]$.

Якщо тепер взяти будь-яку точку, що належить побудованій прямої і багатокутнику Ω , то її координати визначають такий план виготовлення виробів A і B , при якому прибуток від їхньої реалізації складе 600 *гр.* Далі, перемішуючи побудовану пряму в напрямку вектора c паралельно самій собі, будемо збільшувати праву частину рівняння прямої. Останньою спільною точкою прямої і багатокутника є точка M . Координати цієї точки і визначають оптимальний план випуску виробів A і B , при якому прибуток від їхньої реалізації буде максимальним.

Знайдемо координати точки M як точки перетину двох прямих $4x_1 + 4x_2 = 120$, $3x_1 + 12x_2 = 252$. Шукані координати є рішенням системи цих двох рівнянь. Вирішивши систему, одержимо $x_1^* = 12$, $x_2^* = 18$. Отже, якщо підприємство виготовить 12 виробів виду A і 18 виробів виду B , воно одержить максимальний прибуток $y^* = 30 \cdot 12 + 40 \cdot 18 = 1080$ *грн.*

Тема 15. ДВОЇСТІСТЬ У ЛІНІЙНОМУ ПРОГРАМУВАННІ

Кожній задачі лінійного програмування можна певним чином поставити у відповідність іншу задачу лінійного програмування, що називають *двоїстою* стосовно вихідної (початкової) задачі. Вихідна та двоїста задачі тісно зв'язані між собою й утворюють єдину пару двоїстих задач, причому задача, двоїста стосовно двоїстої задачі, збігається з вихідною. Залежно від структури моделі вихідної задачі розрізняють *симетричні, несиметричні й змішані двоїсті задачі*.

Економічний зміст двоїстої задачі

Розглянемо задачу *виробничого планування*. У розпорядженні підприємства є m видів ресурсів в кількостях, що відповідно дорівнюють b_1, b_2, \dots, b_m . Ці ресурси повинні бути використані для виробництва n видів продукції, вартість одиниці якої відома й дорівнює c_j ($j=1, 2, \dots, n$). Крім того, відомі норми споживання a_{ij} кожного з ресурсів на виготовлення одиниці всіх видів продукції. План виробництва $\bar{x}^T = (x_1, x_2, \dots, x_n)$ варто скласти з умови максимізації загальної вартості продукції (прибутку)

$$y(\bar{x}) = \sum_{j=1}^n c_j x_j \rightarrow \max, \text{ при обмеженнях на використання ресурсів:}$$

$$\begin{cases} \sum_{j=1}^n a_{ij} x_j \leq b_i, i = \overline{1, m}; \\ x_j \geq 0, j = \overline{1, n}. \end{cases}$$

За вихідним даними цієї задачі сформулюємо іншу економічну задачу. Для цього припустимо, що підприємству дозволено на його розсуд реалізовувати (продавати) всі зазначені ресурси. У зв'язку із цим виникає необхідність установити оптимальні ціни z_1, z_2, \dots, z_m на ці ресурси, користуючись наступними міркуваннями:

1) покупець ресурсів прагне мінімізувати їх загальну вартість

$$d(\bar{z}) = \sum_{i=1}^m b_i z_i \rightarrow \min;$$

2) з іншого боку, підприємство за кожен вид ресурсів бажає виручити суму, не меншу тієї, котру воно може одержати в результаті використання цих ресурсів $\sum_{i=1}^m a_{ji} z_i \geq c_j$. Це пояснюється тим, що в протилежному випадку йому вигідніше організувати перепродаж саме існуючих ресурсів.

Наведене формулювання дозволяє одержати наступну математичну модель: мінімізувати загальну вартість всіх ресурсів $d(\bar{z}) = \sum_{i=1}^m b_i z_i$ при умовах:

$$\sum_{i=1}^m a_{ji} z_i \geq c_j, j = \overline{1, n}, \text{ де } z_i \geq 0, i = \overline{1, m}.$$

Сформульована задача є двоїстою до вихідної задачі. Ми отримали симетричну пару двоїстих задач.

Вихідна задача	Двоїста задача
$y(\bar{x}) = \sum_{j=1}^n c_j x_j \rightarrow \max$ $\begin{cases} \sum_{j=1}^n a_{ij} x_j \leq b_i, i = \overline{1, m}; \\ x_j \geq 0, j = \overline{1, n}. \end{cases}$	$d(\bar{z}) = \sum_{i=1}^m b_i z_i \rightarrow \min$ $\begin{cases} \sum_{i=1}^m a_{ji} z_i \geq c_j, j = \overline{1, n}; \\ z_i \geq 0, i = \overline{1, m}. \end{cases}$

Загальні правила побудови двоїстих пар задач

1. Спочатку задачу лінійного програмування необхідно привести у відповідність до вимоги: при максимізації цільової функції обмеження-нерівності повинні бути записані зі знаком відношення " \leq ", а при мінімізації – зі знаком " \geq ".

2. Кількість змінних двоїстої задачі дорівнює кількості обмежень вихідної задачі (кожному i -му обмеженню вихідної задачі відповідає змінна z_i двоїстої задачі), а кількість обмежень двоїстої задачі дорівнює кількості змінних вихідної задачі (кожній j -й змінній вихідної задачі відповідає j -е обмеження двоїстої задачі).

3. Кожному i -му обмеженню-нерівності вихідної задачі відповідає у двоїстій задачі умова невід'ємності змінної ($z_i \geq 0$), а рівності – змінна z_i без обмежень на знак (будь-якого знака). Навпаки, позитивній змінній $x_j \geq 0$ відповідає у двоїстій задачі j -е обмеження-нерівність, а змінній довільного знака – рівність.

4. Матриці систем обмежень двоїстої пари задач взаємно транспоновані. Отже, рядок коефіцієнтів a_{ij} в j -м обмеженні двоїстої задачі є стовпець коефіцієнтів при x_j в обмеженнях вихідної задачі й навпаки.

5. Вільні члени обмежень однієї із задач є коефіцієнтами при відповідних змінних у цільовій функції іншої задачі. При цьому максимум цільової функції вихідної задачі переходить у мінімум цільової функції двоїстої задачі і навпаки, мінімум цільової функції вихідної задачі – у максимум цільової функції двоїстої задачі.

Приклади побудови двоїстих задач

Приклад 15.1. Побудувати двоїсту задачу до вихідної задачі:

$$y(\bar{x}) = 2x_1 + 3x_2 + x_3 \rightarrow \max_{x \in \Omega};$$

$$\Omega: \begin{cases} f_1 = x_1 + 2x_2 - x_3 \leq 0; \\ f_2 = 2x_1 + x_2 + 3x_3 \leq 10; \\ x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0. \end{cases}$$

Розв'язання. Двоїста задача згідно з правилами побудови двоїстих задач має вигляд має таку математичну модель:

$$d(\bar{z}) = 6z_1 + 10z_2 \rightarrow \min_{z \in P};$$

$$P: \begin{cases} g_1 = z_1 + 2z_2 \geq 2; \\ g_2 = 2z_1 + z_2 \geq 3; \\ g_3 = -z_1 + 3z_2 \geq 1; \\ z_1 \geq 0, \quad z_2 \geq 0. \end{cases}$$

Ця задача є симетричною до вихідної задачі:

Вихідна задача		Двоїста задача	
1.	$y(\bar{x}) = \sum_{j=1}^n c_j x_j \rightarrow \max$	1.	$d(\bar{z}) = \sum_{i=1}^m b_i z_i \rightarrow \min$
2.	$\sum_{j=1}^n a_{ij} x_j = b_i, \quad i = \overline{1, m};$	2.	$z_i \begin{matrix} \geq \\ \leq \end{matrix} 0, \quad i = \overline{1, m};$
3.	$x_j \geq 0, \quad j = \overline{1, n}.$	3.	$\sum_{i=1}^m a_{ji} z_i \geq c_j, \quad j = \overline{1, n};$

Приклад 15.2. Побудувати двоїсту задачу до вихідної задачі:

$$y(\bar{x}) = x_1 + 4x_2 + 3x_3 + 2x_4 \rightarrow \max;$$

$$\Omega: \begin{cases} 3x_1 + x_2 + 2x_3 + x_4 = 12; \\ 2x_1 + 2x_2 - x_3 + 2x_4 \leq 11; \\ 4x_1 + x_2 + 3x_3 - x_4 \geq 9; \\ x_1 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0. \end{cases}$$

Розв'язання. Оскільки вихідна задача є задачею максимізації, то згідно з правилами побудови двоїстих задач третю нерівність обмежень треба привести до нерівності типу « \leq », для чого помножимо її на -1 (мінус одиницю). Одержимо: $-4x_1 - x_2 - 3x_3 + x_4 \leq -9$. Двоїста задача має таку математичну модель:

$$d(\bar{z}) = 12z_1 + 11z_2 - 9z_3 \rightarrow \min;$$

$$\mathcal{D}: \begin{cases} 3z_1 + 2z_2 - 4z_3 \geq 1; \\ z_1 + 2z_2 - z_3 = 4; \\ 2z_1 - z_2 - 3z_3 \geq 3; \\ z_1 + 2z_2 + z_3 \geq 2; \\ z_2 \geq 0, \quad z_3 \geq 0. \end{cases}$$

Друге обмеження двоїстої задачі записано у вигляді рівності, тому що відповідна їй змінна у вихідній задачі x_2 може набувати будь-якого значення. На змінну z_1 не накладається обмеження невід'ємності, тому що відповідне їй перше обмеження у вихідній задачі має вигляд строгої рівності.

Тема 16. МЕТОДИКА РОЗВ'ЯЗУВАННЯ ТРАНСПОРТНОЇ ЗАДАЧІ

16.1. Змістовна постановка задачі

Однорідний продукт, зосереджений у m пунктах відправлення в кількостях a_1, a_2, \dots, a_m одиниць відповідно, необхідно доставити в кожний із n пунктів призначення в кількості b_1, b_2, \dots, b_n одиниць відповідно. Вартість (відстань) перевезення одиниці продукту з i -го пункту відправлення в j -й пункт призначення дорівнює c_{ij} і відома для кожного маршруту. Нехай x_{ij} – кількість продукту, перевезеного з i -го пункту відправлення в j -й пункт призначення. Задача полягає у визначенні таких розмірів x_{ij} для всіх маршрутів, при яких сумарна вартість (відстань) перевезень була б мінімальною.

16.2. Математична модель задачі

Позначимо:

c_{ij} – тарифи (вартість, час, відстань) перевезення одиниці вантажу з i -го пункту відправлення в j -й пункт призначення;

a_j – запаси вантажу в i -му пункті відправлення;

b_i – потреба у вантажі в j -му пункті призначення;

x_{ij} – кількість одиниць вантажу, перевезеного з i -го пункту відправлення в j -й пункт призначення.

Тоді математична модель транспортної задачі про планування перевезень має такий вигляд:

$$y = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min_{x_{ij} \in \Omega} ; \quad (16.1)$$

$$\Omega: f_j = \sum_{i=1}^m x_{ij} = b_j, \quad j = \overline{1, n} ; \quad (16.2)$$

$$f_{n+i} = \sum_{j=1}^n x_{ij} = a_i, \quad i = \overline{1, m} ; \quad (16.3)$$

$$x_{ij} \geq 0; \quad i = \overline{1, m}; \quad i = \overline{1, m} . \quad (16.4)$$

Тут (16.1) – цільова функція, що визначає вартість перевезень усього вантажу. Саме екстремальне (мінімальне) значення цієї функції необхідно знайти в задачі. Причому значення змінних x_{ij} , при яких цільова функція досягає свого мінімуму, повинні належати області припустимих рішень Ω .

Вирази (16.2) – (16.4) визначають область припустимих рішень Ω . При цьому вираз (16.2) відбиває потреби у вантажі в пунктах призначення, вираз (16.3) визначає запаси вантажів у пунктах відправлення, а вираз (16.4) відокремлює невід'ємну область значень x_{ij} , в яку дані змінні не можуть потрапляти за своїм фізичним змістом.

Вирази (16.2) – (16.4) називаються обмеженнями задачі лінійного програмування. Вирішення задачі (частковий набір значень змінних x_i) називається припустимим, якщо воно одночасно задовольняє всім обмеженням задачі. Вирішення задачі називається оптимальним, якщо воно забезпечує оптимум (у даному випадку мінімум) функції цілі.

Вважатимемо, що функції y, f_1, f_2, \dots, f_n – неперервні лінійні функції, задані на невід'ємному ортанті евклідового простору \mathbf{R}^n . Дані функції мають місце, коли перевезений вантаж є рідиною, сипкою речовиною, дрібними заготівлями або дрібною неспакованою продукцією. Такий вантаж характеризується параметрами, що являють собою вагу, довжину (погонні метри), площу (квадратні метри), об'єм і т.п.

Якщо загальна потреба у вантажі в пунктах призначення дорівнює запасу вантажу в пунктах відправлення, тобто

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j, \quad (16.5)$$

то модель такої транспортної задачі називається закритою. У протилежному випадку – відкритою.

Для можливості розв'язання транспортної задачі необхідно і достатньо, щоб запаси вантажу в пунктах відправлення були рівні потребам у вантажі в пунктах призначення, тобто щоб виконувалася рівність (16.13).

У випадку перевищення запасу над потребою, тобто $\sum_{i=1}^m a_i > \sum_{j=1}^n b_j$, вводиться фіктивний $(n+1)$ -й пункт призначення з потребою $b_{n+1} = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j$.

При цьому відповідні тарифи вважаються рівними нулю: $c_{i,n+1} = 0$ ($i = \overline{1, m}$). Така задача буде вже транспортною задачею, для якої умова (16.13) виконується.

Аналогічно, якщо $\sum_{i=1}^m a_i < \sum_{j=1}^n b_j$, вводиться фіктивний $(m+1)$ -й пункт відправлення з запасом вантажу $a_{m+1} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i$. При цьому відповідні тарифи вважаються рівними нулю: $c_{m+1,j} = 0$ ($j = \overline{1, n}$). Така задача буде вже транспортною задачею, для якої умова (16.13) виконується.

Далі будемо розглядати закриту модель транспортної задачі. Якщо ж модель конкретної задачі є відкритою, то, виходячи зі сказаного вище, її необхідно перетворити так, щоб виконувалася рівність (16.13).

У відкритій моделі область припустимих значень (за інших рівних умов) значно ширше, тому цільова функція досягає кращих значень або, принаймні, не гірше.

16.3. Розв'язування транспортної задачі на ПОЕМ

Вирішення будь-якої задачі математичного програмування традиційним ручним способом потребує від економістів і менеджерів великих затрат сил і часу для здійснення ітераційних процесів наближення до оптимуму. Ситуація різко змінюється, якщо для вирішення задач математичного програмування залучити сучасні інформаційні технології. Існує ряд потужних інформаційних систем, що значно знижують ризик одержання помилкового результату і на декілька порядків скорочують час вирішення задач.

Для вирішення задач математичного програмування з економічним ухилом найбільш удалим є використання сучасної інформаційної системи *Microsoft Excel* версії 7.0 і вище. Пояснюється це, насамперед, тим, що дана система є програмним інструментом для вирішення інших (не зв'язаних

із пошуком екстремуму) задач економіки. Великою перевагою системи є її універсальність. Практично будь-які типи задач математичного програмування можуть бути успішно вирішені за допомогою *Microsoft Excel*, зокрема різні типи транспортних задач.

Інформаційна система *Microsoft Excel* має вбудовану програму *Solver* (Пошук рішення), що являє собою потужний допоміжний інструмент для виконання складних обчислень, у тому числі вирішення більшості задач математичного програмування. Розглянемо використання програми *Solver* на прикладі конкретної закритої транспортної задачі.

Приклад 16.1. Три заводи, що виготовляють бетонні конструкції, постачаються цементом з чотирьох складів. Попит заводів b_j відповідно дорівнює 280, 90 і 180 тис.т/міс. Пропускна здатність складів a_i відповідно становить 200, 150, 80 і 120 тис.т/міс. Відстані перевезень (у

км) із i -го складу на j -й завод подані в матриці $C = [c_{ij}] = \begin{bmatrix} 1 & 5 & 3 \\ 6 & 8 & 9 \\ 2 & 7 & 4 \\ 4 & 1 & 11 \end{bmatrix}$.

Потрібно скласти план перевезень цементу зі складів на заводи, що задовольняв би пропускним спроможностям складів і потребам заводу, а сумарний пробіг вантажного транспорту був би мінімальним.

Технологія вирішення транспортної задачі за допомогою інформаційної системи Microsoft Excel

Вихідні дані для програми *Solver* повинні бути подані у вигляді електронної таблиці, що містить чотири типи областей:

- область змінних задачі;
- область заданих параметрів задачі;
- область проміжних результатів;
- область цільової функції.

Область змінних задачі – це обов'язкова область, що за своєю конфігурацією відповідає формі матриці змінних X . Кожний осередок області відповідає одному елементу x_{ij} матриці X . Змінні можуть мати початкові значення, але не обов'язково. У разі їхньої відсутності програма сама їх введе. Осередки змінних не повинні містити формул.

Область вихідних даних задачі – це обов'язкова область, що містить константи, задані умовою задачі. Для транспортної задачі ця область має три складові:

- підобласть для матриці відстаней (тарифів) $C=[c_{ij}]$;
- підобласть для вектора запасів вантажу в пунктах відправлення $a=[a_i]$;
- підобласть для вектора потреб у вантажі в пунктах призначення $b=[b_j]$.

Осередки всіх підобластей не повинні містити формул. Усі початкові

дані мають бути введені в ці підобласті до початку вирішення задачі.

Область проміжних результатів містять формули, що відображають залежності між даними таблиці. Для транспортної задачі ця область розпадається на три підобласті:

- підобласть $C \times X$ для добутків елементів матриці X на відповідні елементи матриці C . Необов'язкова область. При її наявності кожний осередок містить формулу, що визначає добуток $x_{ij}c_{ij}$.

- підобласть функцій обмежень типу (16.2), що визначають запаси в пунктах відправлення. Це обов'язкова область, кожний осередок якої містить формулу для визначення запасу вантажу у відповідному пункті від-

правлення $\sum_{j=1}^n x_{ij}$;

- підобласть функцій обмежень типу (1.11), що визначають потребу у вантажі в пунктах призначення. Це обов'язкова область, кожний осередок якої містить формулу для визначення потреби відповідного пункту при-

значення $\sum_{i=1}^m x_{ij}$.

Область цільової функції складається з одного (і тільки одного) осередку, в якому записана формула для визначення критерію (16.1), тобто

формула подвійної суми $\sum_{i=1}^m \sum_{j=1}^n x_{ij}$ (при наявності області $C \times X$) або спеціальна функція “СУММПРОИЗВ” (при відсутності області $C \times X$).

Питання використовувати або не використовувати область $C \times X$ для вирішення транспортної задачі вирішує сам користувач. У випадку її наявності при малій розмірності транспортної задачі підсилюється наочність її вирішення, при великій – погіршується.

Слід зазначити, якщо транспортна задача має відкриту модель, то при її вирішенні немає необхідності приводити модель до закритої, як це має місце при ручному обчисленні. Крім того, немає також необхідності в процедурі пошуку початкового опорного плану (початкового приближеного розв'язку).

16.4. Приклад вирішення транспортної задачі за допомогою інформаційної системи Microsoft Excel

На рис.16.1 показано розподіл областей електронної таблиці і їхнє заповнення вихідними даними в умовах Прикладу 16.1. Як видно з таблиці, область змінних задачі заповнюється нулями, а підобласті вихідних даних заповнюються даними, взятими з умови задачі.

Область проміжних результатів заповнюється в такий спосіб:

- в осередку $BF\$3$ записують формулу $=СУМ(B3:D3)$ і копіюють в осередки $BF\$4:BF\6 ;

– в осередку \$Y\$8 записують формулу $=\text{СУМ}(B3:B6)$ і копіюють в осередки \$C\$8: \$Y\$8;

– в осередку \$F\$12 записують формулу $=B3*B12$ і копіюють в осередки \$G\$12: \$H\$12, а потім осередки \$F\$12:\$H\$12 копіюють в осередки \$F\$13: \$H\$15.

Нарешті, в осередок цільової функції \$H\$9 при наявності області $C \times X$ записують формулу $=\text{СУМ}(F12:H15)$, а при відсутності – формулу $=\text{СУММПРОИЗВ}(B12:D15;B3:D6)$.

Microsoft Excel - Трансп-з

Файл Правка Вид Вставка Формат Сервис Данные Окно ?

Courier New 10 Ж К Ч [Icons] % , +0,00 -0,00 [Icons] [Icons] [Icons] [Icons]

J23 =

	A	B	C	D	E	F	G	H	I
1	Вирішення закритої транспортної задачі про планування перевезень								
2								$a_i (i=1,2,3,4)$	
3		0	0	0	=	0	=	200	
4	$[x_{ij}] =$	0	0	0	=	0	=	150	
5		0	0	0	=	0	=	80	
6		0	0	0	=	0	=	120	
7									
8		0	0	0					
9								0	y_{min}
10	$b_j (j=1,2,3)$	280	90	180					
11									
12		1	5	3		0	0	0	
13	$[c_{ij}] =$	6	8	9	$[c_{ij}x_{ij}] =$	0	0	0	
14		2	7	4		0	0	0	
15		4	1	11		0	0	0	
16									
17									
18									
19									

Область CxX

Готово

Пуск Проводник - ИсслОперУ... Microsoft Word Microsoft Excel - Тра... En 11:08

Рис.16.1

Вся інша текстова інформація, подана в електронній таблиці (рис.16.1), не є обов'язковою. Її відсутність не вплине на вирішення задачі.

Подальша підготовка до запуску процесу вирішення задачі пов'язана безпосередньо з програмою *Solver*, яка ініціалізується командою *Сервіс / Пошук рішення*. При цьому на екрані з'являється діалогове вікно програми *Solver*, яке вимагає встановити параметри вирішення задачі. На рис.16.2 показано діалогове вікно програми з необхідними установками.

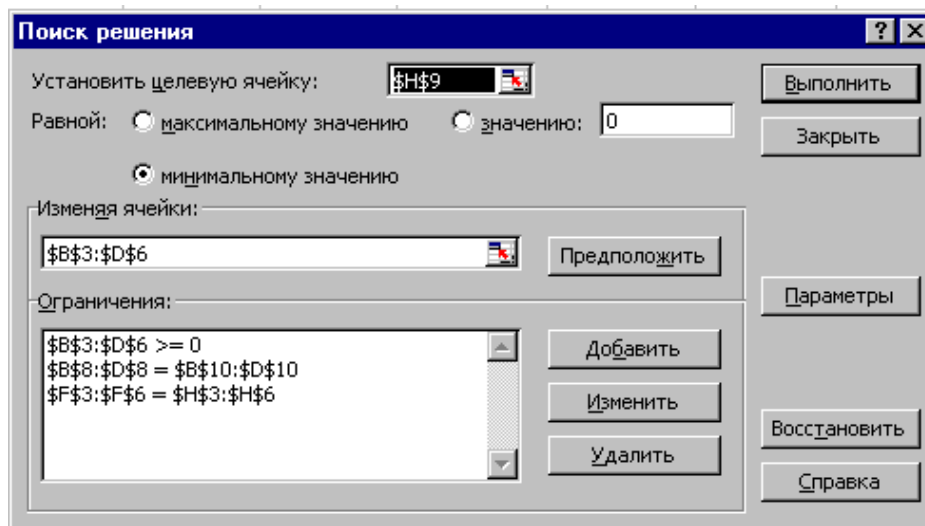


Рис.16.2 – Діалогове вікно програми Solver

Для транспортної задачі з Прикладу 16.1 установки діалогового вікна повинні бути наступними:

- як цільовий осередок вказують осередок \$H\$9;
- вибирають селекторну кнопку «мінімальне значення»;
- у вікні «Змінюючи осередки» вказують діапазон кліток \$B\$3:\$D\$6;
- у вікні «Обмеження» послідовно вказують обмеження задачі: \$B\$3:\$D\$6 >= 0; \$B\$8:\$D\$8 = \$B\$10:\$D\$10; \$F\$3:\$F\$6 = \$H\$3:\$H\$6.

Далі необхідно запустити процес обчислення натисканням кнопки «Виконати» в діалоговому вікні. Окремі кроки процесу відображаються в рядку стану. Після завершення пошуку рішення нові значення будуть вставлені в електронну таблицю, а на екрані з'явиться нове діалогове вікно, що містить інформацію про завершення процесу пошуку рішення. У цьому вікні необхідно вибрати опцію «Зберегти знайдене рішення». У результаті вибору нові значення залишаться в таблиці. У протилежному разі програма відновить значення, що були в таблиці до натискання кнопки «Виконати».

Для транспортної задачі з Прикладу 16.1 остаточний вигляд електронної таблиці подано на рис.16.3. Тут у цільовому осередку знаходиться оптимальне значення критерію оптимальності – число 1830, а в області змінних (осередки \$B\$3:\$D\$6) – шукані значення x_{ij}^* .

Microsoft Excel - Транспорт-3

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка

Введите вопрос

Courier New 10 Ж К У

H9 =СУММПРОИЗВ(B12:D15;B3:D6)

	A	B	C	D	E	F	G	H	I
1		Вирішення закритої транспортної задачі про планування перевезень							
2								$a_i (i=1,2,3,4)$	
3		100	0	100	=	200	=	200	
4	$[x_{ij}] =$	150	0	0	=	150	=	150	
5		0	0	80	=	80	=	80	
6		30	90	0	=	120	=	120	
7									
8		280	90	180					
9								1830	Y_{min}
10	$b_j (j=1,2,3)$	280	90	180					
11									
12		1	5	3		100	0	300	
13	$[c_{ij}] =$	6	8	9	$[c_{ij}x_{ij}] =$	900	0	0	
14		3	7	4		0	0	320	
15		4	1	11		120	90	0	
16									
17									

Готово NUM

Пуск Обзор - Исслед-опер ТранЗад-2у - Microsoft ... Microsoft Excel - Тр... 17:28

Рис.16.3

Тема 17. ЦІЛОЧИСЛОВЕ ПРОГРАМУВАННЯ

17.1. Сутність і класифікація задач цілочислового програмування

Цілочислове програмування є окреим випадком дискретного програмування. До останньої відносяться такі відомі задачі як: *задача розподілу ресурсів; задача сіткового планування й керування; задача календарного планування* та ін.

За структурою математичної моделі задачі *дискретного програмування* розділяють на такі класи:

- задачі з неподільностями (невідомі приймають тільки цілі значення);
- екстремальні комбінаторні задачі (цільова функція задана на кінцевій множині, елементами якого є перестановки з n об'єктів);
- задачі на незв'язаних і неопуклих областях;
- задачі з розривними цільовими функціями.

Методи розв'язання задач дискретного програмування за принципом підходу до проблеми розділяють на такі класи:

- методи відтинань (відсікань);
- метод гілок і границь;
- метод випадкового пошуку;

– наближені евристичні методи.

Метод відтинань. Спочатку умова цілочисельності тимчасово відкидається, і задача вирішується традиційним методом. Якщо отримане рішення не задовольняє умові цілочисельності, то вводять додаткове обмеження, якому свідомо задовольняє будь-який цілочисловий план і не задовольняє знайдений оптимальний нецілочисловий план, і вирішують нову задачу. Процедура повторюється доти, поки не буде знайдено цілочислове рішення або не буде встановлена несумісність обмежень задачі. Додаткове обмеження являє собою гіперплощину, яка відсікає від області припустимих рішень отримане нецілочислове рішення. Найвідоміший метод відтинань – метод Гоморі.

Метод гілок і границь. Спочатку умова цілочисельності відкидається, і задача вирішується традиційними методами оптимізації. В результаті одержують нижню (при пошуку мінімуму) оцінку цільової функції. Потім вихідну множину цілочислових планів ділять на кінцеве число підмножин. Для кожної з яких знову визначається оцінка. Підмножина цілочислових планів, для яких оцінка виявиться найближчою до вихідної задачі, залишається для подальшого розв’язання, а решта підмножин відкидається. Для підмножини, що залишилися, повторюють процедуру розбиття на підмножини. Процес розв’язання припиняється, коли у черговій множині, що залишається після оцінювання, міститься тільки один цілочисловий план.

Якщо в задачі цілочислового лінійного програмування є тільки дві змінні, її можна вирішити графічним методом.

Приклад 17.1. Знайти оптимальне рішення задачі лінійного програмування:

$$y(\bar{x}) = x_1 + 2x_2 \rightarrow \max_{\bar{x} \in \Omega},$$
$$\Omega : \begin{cases} f_1 = 4x_1 - 2x_2 \leq 11; \\ f_2 = -x_1 + 3x_2 \leq 9; \\ f_3 = x_1 + 2x_2 \geq 2; \\ x_j \geq 0, j = 1, 2; \\ x_j = \text{int}, j = 1, 2;. \end{cases}$$

Розв’язання. Спочатку вирішимо задачу без умови цілочисельності. Необхідні побудови показано на рис.17.1.

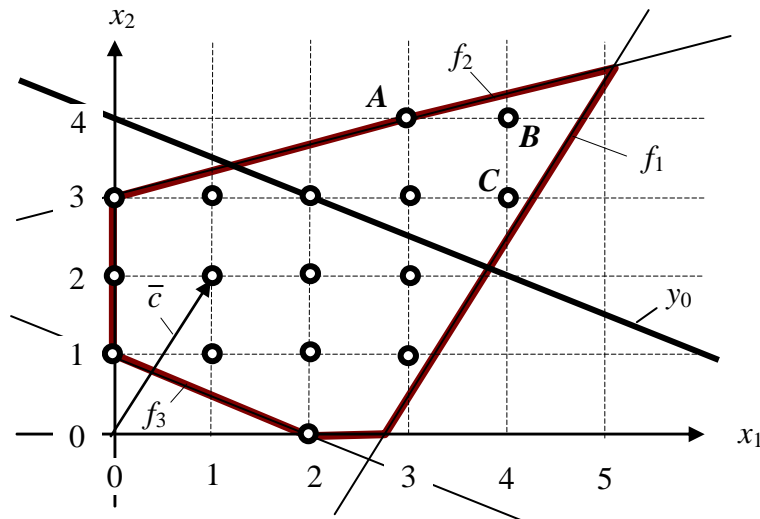


Рис. 17.1

Нецілочислове рішення лежить на перетині прямих f_1 й f_2 (рис.17.1):

$$\begin{cases} 4x_1 - 2x_2 = 11; \\ -x_1 + 3x_2 = 9. \end{cases}$$

Звідки $\bar{x}^* = \begin{bmatrix} 5,1 \\ 4,7 \end{bmatrix}$, $y^* = 5,1 + 2 \cdot 4,7 = 14,5$.

Умові цілочисельності змінних задовольняють координати 16 жирних точок на рис. 17.1.

Якщо пряму цільової функції y_0 зміщати в напрямку вектора \bar{c} , то знайдемо цілочисловий максимум $y^* = 12$ в точці В з координатами (4, 4). Для порівняння оцінимо значення цільової функції в трьох точках А, В і С:

$$y_A = 3 + 2 \cdot 4 = 11;$$

$$y_B = 4 + 2 \cdot 4 = 12;$$

$$y_C = 4 + 2 \cdot 3 = 10.$$

Відповідь: цілочисловий максимум досягається в точці В: $\bar{x}^* = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$.

17.2. Задача про призначення (кадрова задача)

Змістовна постановка задачі. Нехай є n робіт і n кандидатів для виконання цих робіт. Призначення кандидата i на роботу j пов'язане з витратами c_{ij} , $i, j = \overline{1, n}$. Потрібно знайти призначення кандидатів на всі роботи, що дає мінімальні сумарні витрати. При цьому кожного кандидата можна призначити тільки на одну роботу і кожна робота може бути зайнята тільки одним кандидатом.

Рішенням цієї задачі є перестановки $(p_1, p_2, \dots, p_j, \dots, p_n)$ чисел $(1, 2, \dots, i, \dots, n)$. Кожне із зроблених призначень описується відповідністю $i \rightarrow p_j, \quad i = \overline{1, n}$.

Кінцева множина, на якій шукається рішення, являє собою множину всіх перестановок чисел $(1, 2, \dots, i, \dots, n)$. Як відомо, кожна така перестановка може бути описана точкою в n^2 -вимірному евклідовім просторі. Цю точку зручно представляти у вигляді матриці $\mathbf{X} = [x_{ij}]$ розмірності $n \times n$, де

$$x_{ij} = \begin{cases} 1 & \text{якщо } i\text{-й кандидат призначається на } j\text{-у роботу,} \\ 0 & \text{в протилежному разі.} \end{cases}$$

Математична постановка цієї задачі як задачі цілочислового лінійного програмування має вигляд:

$$\sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min_{x_{ij} \in \Omega}, \quad (17.1)$$

$$\Omega: \sum_{j=1}^n x_{ij} = 1, \quad i = \overline{1, n}, \quad (17.2)$$

$$\sum_{i=1}^n x_{ij} = 1, \quad j = \overline{1, n} \quad (17.3)$$

17.3. Задача про інвестиції (фінансування проектів)

Змістовна постановка задачі. Нехай протягом T років можливе здійснення n дослідних проектів. Очікуваний ефект проекту j , виражений в «сьогоднішніх» одиницях корисності, становить c_j . Витрати на рік i на здійснення проекту j становлять a_{ij} , а загальний ліміт капіталовкладень на дослідження в році i дорівнює b_i . Потрібно вказати максимально ефективний набір проектів, що не виводить за ліміти вкладень.

Математична модель задачі. Введемо змінні

$$x_j = \begin{cases} 1 & \text{якщо проект } j \text{ здійснюється,} \\ 0 & \text{в протилежному разі.} \end{cases}$$

Тоді задача про фінансування проектів зведеться до задачі цілочислового лінійного програмування:

$$\sum_{j=1}^n c_j x_j \rightarrow \max_{x_j \in \Omega}, \quad (17.4)$$

$$\Omega: \sum_{j=1}^n p_j x_j \leq b_i, \quad i = \overline{1, T}, \quad (17.5)$$

$$x_j \in \{0,1\}, \quad j = \overline{1,n}. \quad (17.6)$$

17.4. Задача про розподіл обладнання (парку машин)

Змістовна постановка задачі. Є n типів сільськогосподарських машин і m видів робіт, що підлягають виконанню в обсягах b_i , $i = \overline{1,m}$ (будемо вважати, що всі обсяги виражені в гектарах). Задана продуктивність j -ї машини на i -й роботі a_{ij} , а також собівартість d_{ij} обробки 1 га i -ї роботи j -ю машиною. Собівартість самих машин (їх оренда) складає c_j . Слід знайти оптимальний машинний парк для даного комплексу робіт і вказати його розподіл по роботах, щоб виконати завдання і добитися мінімальної сумарної вартості робіт та оренди машин.

Математична модель задачі. Введемо змінні: x_j - кількість машин j -го типу; y_{ij} - кількість гектарів, що обробляються j -машиною на i -й роботі.

Тоді задача про розподіл парку машин зведеться до задачі цілочислового лінійного програмування:

$$\sum_{j=1}^n c_j x_j + \sum_{j=1}^n \sum_{i=1}^m d_{ij} y_{ij} \rightarrow \min_{x_j, y_{ij} \in \Omega}, \quad (17.7)$$

$$\Omega: \sum_{j=1}^n a_{ij} y_{ij} = b_i, \quad i = \overline{1,m}, \quad (17.8)$$

$$x_j - \sum_{i=1}^m y_{ij} \geq 0, \quad j = \overline{1,n}, \quad (17.9)$$

$$x_j \geq 0; \quad x_j = \text{int}, \quad j = \overline{1,n}, \quad (17.10)$$

$$y_{ij} \geq 0; \quad y_{ij} = \text{int}, \quad j = \overline{1,n}, \quad i = \overline{1,m}. \quad (17.11)$$

НАВЧАЛЬНЕ ВИДАННЯ

САМОЙЛЕНКО Микола Іванович,
БІЛОГУРОВА Ганна Вікторівна,
ШТЕЛЬМА Ольга Миколаївна,
ПРОТОПОПОВА Валентина Петрівна

ВИЩА ТА ПРИКЛАДНА МАТЕМАТИКА
Теорія ймовірностей та математична статистика.
Математичне програмування

(Курс лекцій для студентів 1 курсу денної та заочної форм навчання
освітньо-кваліфікаційного рівня бакалавр
галузі знань 0306 – «Менеджмент і адміністрування»
напрямку 6.030601 – «Менеджмент»)

Відповідальний за випуск *М. І. Самойленко*
За авторською редакцією
Комп'ютерне верстання *О. А. Балашова*

План 2012, поз. 158Л

Підп. до друку 10.05.2012	Формат 60х84/16
Друк на ризографі	Ум.-друк. арк. 8,41
Тираж 50 пр.	Зам. №

Видавець і виготовлювач:
Харківська національна академія міського господарства,
вул. Революції, 12, Харків, 61002
Електронна адреса: rectorat@ksame.kharkov.ua
Свідоцтво суб'єкта видавничої справи:
ДК №4064 від 12.05.2011