І. П. Васильченко

ВИЩА МАТЕМАТИКА
ДЛЯ ЕКОНОМІСТІВ

(СПЕЦІАЛЬНІ РОЗДІЛИ)

Підручник

Затверджено
Міністерством освіти і науки України
як підручник для студентів
вищих навчальних закладів

Київ

КОНДОР

2007

УДК 51(075.8)
ББК 22.11я73
В19

Затверджено
Міністерством освіти і науки України
як підручник для студентів вищих навчальних закладів
(Лист МОНУ №14/18.2-1651 від 15.10.2003 р.)

Рецензенти:
Ф.Г. Гаращенко, *доктор фізико-математичних наук, професор*
(Київський національний університет імені Тараса Шевченка)
А.Ю. Дорошенко*, доктор фізико-математичних наук, професор*
(Київський славістичний університет)

Васильченко І.П. Вища математика для економістів (спеціальні
В19 **розділи).** – К.: Кондор, 2007.– 352 с.
ISBN 966-7982-59-9

Зміст спеціальних розділів охоплює такі розділи вищої математики: теорію ймовірностей, математичну статистику; елементи лінійної алгебри; лінійне програмування, що рекомендовані типовою навчальною програмою Міністерства освіти і науки України для економічних спеціальностей.

Теоретичній матеріал викладений доступно, а ілюстрація основних понять і теорем супроводжується великою кількістю прикладів і задач. До кожного параграфа наведені задачі для самостійної роботи.

Розраховано насамперед на студентів економічних спеціальностей, економістів-практиків, осіб, які займаються самоосвітою, може бути корисним також викладачам вищих навчальних закладів та коледжів.

ББК 22.11я73

ISBN 966-7982-59-9

ПРО АВТОРА

Іван Петрович Васильченко — професор кафедри загальної математики механіко-математичного факультету Київського національного університету імені Тараса Шевченка, доктор технічних наук, член Національного комітету України з теоретичної та прикладної механіки і математики, фахівець у галузі математичного моделювання.

Вищу освіту здобув на механіко-математич-ному факультеті Київського національного універ-ситету імені Тараса Шевченка, де і розпочав свою наукову і педагогічну діяльність, навчаючись спочатку в аспірантурі, а потім незмінно працюючи на кафедрі загальної математики, що сприяло набуттю досвіду викладання вищої математики студентам різних спе-ціальностей і різних форм навчання.

У 1986 р. І.П. Васильченко захистив докторську дисертацію. Через два роки йому було присвоєно вчене звання професора.

Автор понад 145 наукових праць з питань прикладної математики, зокрема трьох монографій і п’яти підручників. Серед них “Вища математика” (К., 1992), “Вища математика” (К., 1994), “Вища математика” (К., 1995), “Вища математика для економістів” (К., 2002). Одна із монографій опубліко-вана у видавництві Московського університету ім. М.В. Ломоносова (1990 р.).

ВСТУП

Пропонований підручник призначений для вивчення досить специфічних для засвоєння розділів вищої математики: теорії ймовірностей та математичної статистики; лінійного програмування, рекомендованих типовою навчальною програмою Міністерства освіти і науки України для економічних спеціальностей вищих навчальних закладів. Підручник є намаганням у доступній, простій і лаконічній формі викласти основний матеріал зазначених розділів вищої математики, не переобтяжуючи при цьому читача ретельністю та строгістю викладення. Намагаючись повідомити найважливіші означення і теореми та за можливістю навести їх доведення, все ж таки доведення деяких теорем у книзі випущено, оскільки, на наш погляд, вони ускладнюють засвоєння при початковому вивченні, тим більше якщо мова йде про нематематичні спеціальності.

Запропонований підручник є узагальненням багаторічного досвіду викладання автором вищої математики у Київському національному університеті імені Тараса Шевченка студентам різних спеціальностей та різних форм навчання. Його метою є забезпечення грунтовного засвоєння спеціальних розділів вищої математики, розкриття в доступному вигляді основних понять і теорем, сприяння формуванню навичок у застосуванні методів вищої математики, допомагати студентам при самостійному вивченні матеріалу та розв'язуванні задач. Відомо, що новий теоретичний матеріал засвоюється значно легше, якщо він супроводжується достатньою кількістю ілюструючих його прикладів. А тому у виданні зроблена спроба поєднати теоретичний матеріал з методичними рекомендаціями та розв'язуванням багатьох типових задач, зокрема економічного змісту. Оскільки в кожному параграфі є певна кількість детально розв'язаних прикладів і задач, що пояснюють теоретичний матеріал і сприяють більш глибокому його розумінню, вона може знайти застосування в педагогічній діяльності вищих навчальних закладів, як очної так і заочної форм навчання. Підібрані спеціальні вправи і задачі будуть сприяти повторенню і поглибленому вивченню відповідного параграфа та сприятимуть підвищеному інтересу читача до занять з математики, а також використанню математичних методів фахівцями економіки й менеджменту. Таким чином, кожний параграф містить у собі основні теоретичні відомості, приклади розв'язування типових задач та задачі для самостійної роботи, а тому книгу можна використовувати і як збірник задач.

Автор сподівається одержати від студентів та фахівців зауваження, рекомендації та побажання, спрямовані на поліпшення підручника.

Розділ I
ТЕОРІЯ ЙМОВІРНОСТЕЙ

§1. Основні поняття теорії ймовірностей

В економіці та інших сферах людської діяльності часто зустрічаються з явищами, коли при неодноразовому відтворенні одного і того ж експерименту в незмінних умовах, кожен раз відбуваються по іншому. Виконуючи експериментальні дослідження якого-небудь явища, ми переконуємося, що при достатньо великій кількості спостережень має місце розкиданість експериментальних результатів.

Будемо називати *подією* якісний або кількісний результат експерименту, який виконується за певних умов.

Подія називається *достовірною*, якщо вона обов'язково відбувається при заданих умовах і *неможливою*, якщо вона при цих умовах відбутися не може.

Подія називається *випадковою*, якщо вона при заданому комплексі умов може відбутися, а може й не відбутися. Об'єктивною математичною оцінкою можливості реалізації випадкової події є її *ймовірність*.

Математична наука, яка вивчає загальні закономірності випадкових явищ, незалежно від їх конкретної природи і дає методи кількісної оцінки впливу випадкових факторів на різні явища, *називається теорією ймовірностей.*

Класичне означення ймовірності. Випадкові події прийнято позначати великими літерами А, В, С тощо. Події A_1, A_2,..., A_n *називаються рівноможливими*, якщо при виконанні комплексу умов, кожна з яких має однакову можливість відбутися.

Приклад 1. В ящику лежать 3 однакових занумерованих кулі. Очевидно, є однакова можливість добути з ящика кулю з № 1 (подія A_1), з № 2 (подія A_2) або з № 3 (подія A_3), тобто A_1, A_2, A_3 — рівноможливі події.

Ймовірністю події А називається відношення числа т, сприяючих події А випадків, до загального числа п рівноможливих, єдино можливих і несумісних випадків.

$$P(A) = \frac{m}{n} \qquad (1)$$

Приклад 2. Знайти ймовірність того, що при підкиданні грального кубика випадає парне число очків.

Тут при m = 3, n = 6. Отже, $P(A)=\dfrac{3}{6}=\dfrac{1}{2}$.

Ймовірність будь-якої події A задовольняє нерівності:

$$0\le P\le 1 \qquad (2)$$

Дійсно, $P(A)=\dfrac{m}{n}$, де $0\le m\le n$, звідки $0\le\dfrac{m}{n}\le 1$, тобто $0\le P(A)\le 1$.

Безпосереднє обчислення ймовірностей. Основні формули комбінаторики. Для підрахунку m і n застосовуються формули комбінаторики: формули для числа розміщень, перестановок і сполучень, біном Ньютона.

Перестановка. Будь-який встановлений у скінченній множині порядок називається *перестановкою*.

Наприклад, нехай $\mathrm{X}=\{1;2;3\}$, тоді перестановки мають вигляд: $\{1;2;3\}$, $\{1;3;2\}$, $\{2;3;1\}$, $\{2;1;3\}$, $\{3;1;2\}$, $\{3;2;1\}$.

Число всіх перестановок з n елементів позначають P_n. Воно дорівнює добутку послідовних натуральних чисел від 1 до n включно:

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

Розміщення. Множина, в якій заданий порядок розміщення її елементів, називається *упорядкованою.* Нехай дано скінчену множину, яка складається з n елементів. Будь-яка її упорядкована m-елементна підмножина ($m\le n$) називається *розміщенням* з n елементів по m.

Розміщення відрізняється або власне елементами, або порядком їх розміщення. Нехай, наприклад, $\mathrm{X}=\{1;2;3\}$, m = 2, тоді всі можливі розміщення мають вигляд: $\{1;2\}$, $\{2;1\}$, $\{1;3\}$, $\{3;1\}$, $\{2;3\}$, $\{3;2\}$.

Число розміщень з n елементів по m позначають A_n^m. Це число знаходять за формулою:

$$A_n^m=n(n-1)(n-2)\ldots(n-m+1), \qquad (4)$$

яку можна записати у вигляді:

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

якщо m = n, то дістанемо формулу $A_n^n=P_n=n!$.

Сполучення. Нехай дано скінчену множину, яка складається з n елементів. Будь-яка її m-елементна підмножина ($m \le n$) називається *сполученням з n елементів по m.*

Наприклад, якщо $X = \{1;2;3\}$ і m = 2, тоді всі можливі сполучення мають вигляд: $\{1;2\}$, $\{2;3\}$, $\{1;3\}$. Число сполучень з n елементів по m позначають C_n^m.

Це число знаходять за формулою:

$$C_n^m = \frac{A_n^m}{P_m}, \tag{5}$$

яку можна записати

$$C_n^m = \frac{n!}{m!(n-m)!} \text{ \textit{або} } C_n^m = \frac{n(n-1)...(n-m+1)}{m!}.$$

Приклад 3. Є дві картки, на яких написані літери “О”, дві картки з літерою “К” та по одній картці з літерами “Е”, “Н”, “М”, “І”, “А”. Яка ймовірність того, що розміщуючи картки навмання одна за одною, можна одержати слово “ЕКОНОМІКА”?

Число способів розставити 9 карток одну за одною дорівнює $P_9 = 9!$. Нехай подія A полягає в тому, що одержане слово — “ЕКОНОМІКА”. Тоді A включає ті наслідки, за яких літера “Е” стоїть на першому місці, літера “К” — на 2-му та на 8-му місцях, “О” — на 3-му та 5-му місцях, “Н”, “М”, “І”, “А” — відповідно на 4-му, 6-му, 7-му, 9-му місцях. Оскільки картки з літерами “К” і “О” можна переставити, то число наслідків, які сприяють появі події A, дорівнює $m = 2!\cdot 2! = 4$. Отже,

$$P(A) = \frac{4}{9!} = \frac{1}{90720} \approx 11\cdot 10^{-6}.$$

Приклад 4. Набираючи номер телефону, абонент забув останні дві цифри, знаючи, що вони різні, набрав їх навмання. Знайти ймовірність того, що абонент вірно набрав номер телефону.

Очевидно, що в номері телефону грає роль порядок розміщення цифр. Загальне число способів набрати 2 цифри номера телефону дорівнює:

$$A_7^2 = \frac{7!}{(7-2)!} = \frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7}{1\cdot 2\cdot 3\cdot 4\cdot 5} = 42.$$

Нехай подія A полягає в тому, що набраний номер вірний. Тоді A включає один елементарний наслідок, отже,

$$P(A) = \frac{1}{42}.$$

Приклад 5. У вагоні 15 пасажирів. Поїзд зупиняється на 20 станціях. Яка ймовірність того, що два пасажири не зійдуть на одній і тій самій станції?

Кожний пасажир може вийти на будь-якій із 20 станцій. Загальне число наслідків дорівнює 15^{20}. Число наслідків, що сприяють, буде дорівнювати числу розміщень з 20 по 15: $A_{20}^{15} = 20 \cdot 19 \cdot 18 \cdot ... \cdot 6$. Із урахуванням порядку виходу пасажирів маємо:

$$P(A) = \frac{20 \cdot 19 \cdot 18 \cdot ... \cdot 6}{15^{20}}.$$

Приклад 6. Гра в лото: вгадати "*k* чисел із *n*", наприклад, спортлото "6 із 49". Яка ймовірність одержати головний виграш, вказавши *k* чисел вірно?

Маємо одну сприятливу подію, за якої відбувається виграш головного призу гри. Число всіх елементарних подій дорівнює числу можливих розміщень k чисел із n чисел без урахування порядку і без повторення, тобто дорівнює A_n^k. Таким чином, ймовірність головного виграшу в спорт лото дорівнює $\frac{1}{A_{49}^6} = \frac{1}{13983816}$.

При розв'язанні задач комбінаторики використовують наступні правила:

Правило суми. Якщо деякий об'єкт A може бути вибраний із сукупності m способами, а інший об'єкт B може бути вибраний n способами, то вибрати або A, або B можна $m+n$ способами.

Правило добутку. Якщо деякий об'єкт A можна вибрати із сукупності об'єктів m способами і після кожного такого вибору об'єкт B можна вибрати n способами, то пара об'єктів (A, B) у вказаному порядку може бути вибрана $m \cdot n$ способами.

Приклад 7. У партії із 10 деталей 7 стандартних. Знайти ймовірність того, що серед 6 взятих навмання деталей 4 стандартні.

Загальне число можливих елементарних наслідків випробувань рівне числу способів, якими можна добути 6 деталей із 10, тобто числу сполучень із 10 елементів по 6 елементів (C_{10}^6).

Визначимо число наслідків, сприятливих події A, що нас цікавить (серед 6 взятих деталей — 4 стандартні). 4 стандартні деталі можливо взяти із 7-ми стандартних деталей C_7^4 способами, причому останні $6 - 4 = 2$ деталі повинні бути нестандартними; вибрати ж 2 нестандартні деталі із $10 - 7 = 3$ нестандартних деталей можна C_3^2 способами. Отже, число сприятливих наслідків дорівнює $C_7^4 \cdot C_3^2$. Шукана ймовірність дорівнює відношенню числа наслідків, сприятливих події, до числа всіх елементарних наслідків:

$$P(A) = \frac{C_7^4 \cdot C_3^2}{C_{10}^6} = \frac{7!}{4!(7-4)!} \cdot \frac{3!}{2!(3-2)!} \cdot \frac{6!(10-6)!}{10!} = \frac{1}{2}.$$

Відносна частота. Статистична та геометрична ймовірність. *Відносною частотою події* називають відношення числа випробувань, в яких подія мала місце “m” до загального числа фактично проведених випробувань “n”.

$$W(A) = \frac{m}{n}.$$

Співставляючи означення ймовірності і відносної частоти, зауважимо: означення ймовірності не потребує, щоб випробування відбувалося в дійсності; означення ж відносної частоти припускає, щоб випробування були проведені фактично. Тобто, ймовірність обчислюють до експерименту, а відносну частоту — після експерименту.

Приклад 8. Відділ технічного контролю виявив 3 нестандартних деталі в партії із 80 випадково відібраних деталей. Відносна частота появи нестандартної деталі:

$$W(A) = \frac{3}{80}.$$

Якщо в однакових умовах проводити експерименти, в кожному з яких число випробувань достатньо велике, то відносна частота стає стійкою. Тобто, в різних експериментах відносна частота змінюється мало (тим менше, чим більше проведено випробувань), коливаючись біля деякого сталого числа. Зокрема, якщо дослідним шляхом встановлена відносна частота, то одержане число можна вважати за наближене значення ймовірності. За цією причиною поряд з класичним визначенням ймовірності, що припускає скінчене число елементарних наслідків випробувань, використовують і інші означення, зокрема: статистичне означення як *статистичну ймовірність* події приймають відносну частоту або число, близьке до неї.

Властивості ймовірності, які випливають з класичного означення, зберігаються і при статистичному означенні ймовірності.

Саме ці обставини дозволяють при вивчені випадкових подій застосовувати математичні методи, приписуючи кожній масовій випадковій події його ймовірність, за яку приймається те число, навколо якого коливається частота події.

Щоб позбутися недоліку класичного означення ймовірності, який полягає в тому, що воно неприйнятне для випробувань з нескінченим числом наслідків, вводять *геометричні ймовірності* — ймовірність попадання точки в область (відрізок, частину площини тощо).

Нехай відрізок l є частиною відрізка L. На відрізок L навмання поставлена точка. Якщо припустити, що ймовірність попадання точки на відрізок l пропорційна довжині цього відрізка і не залежить від його розташування відносно відрізка L, то ймовірність попадання точки на відрізок l визначається рівністю:

$$P = \frac{довжина\, l}{довжину\, L}.$$

Аналогічно можна говорити про плоску або просторову фігуру.

Приклад 9. На площині накреслені два концентричні кола, радіуси яких 15 і 20 см відповідно. Знайти ймовірність того, що точка кинута на вдачу у великий круг, попаде в кільце, утворене побудованими колами. Припускається, що ймовірність попадання точки в плоску фігуру пропорційна площі цієї фігури і не залежить від її розташування відносно великого круга.

Площа кільця:

$$s = \pi(20^2 - 15^2) = 175\pi.$$

Площа великого круга:

$$S = \pi \cdot 20^2 = 400\pi.$$

Шукана ймовірність:

$$P = \frac{175\pi}{400\pi} \approx 0,44.$$

ЗАДАЧІ

1.1. У партії із N деталей є n стандартних. На вдачу відібрані m деталей. Знайти ймовірність того, що серед відібраних деталей рівно k стандартних.

Відповідь: $P = \dfrac{C_n^k \cdot C_{N-n}^{m-k}}{C_N^m}$.

1.2. В упаковці є п'ять однакових костюмів, причому три з них чорного кольору. На вдачу відібрані два костюми. Знайти ймовірність того, що серед двох добутих костюмів будуть:

а) один чорний костюм;
б) два чорні костюми;
в) хоча б один чорний костюм.

Відповідь: а) $P = C_3^1 \cdot C_2^1 \cdot \dfrac{1}{C_5^2} = 0,6$;

б) $P = \dfrac{C_3^2}{C_5^2} = 0,3$;

в) $P = 0,9$.

1.3. Відділ технічного контролю виявив п'ять бракованих телевізорів у партії із випадково відібраних 100 телевізорів. Знайти відносну частоту появи бракованих телевізорів.

Відповідь: $W(A) = 0,05$.

1.4. На кожній із шести однакових карток надрукована одна із наступних літер: а, т, м, р, є, о. Картки ретельно перемішані. Знайти ймовірність того, що на чотирьох, взятих по одній і розташованих в одну лінію карток можливо буде прочитати слово “троє”.

Відповідь: $P = \dfrac{1}{360}$.

1.5. У круг радіуса R помістили менший круг радіуса r. Знайти ймовірність того, що точка навмання кинута у великий круг, попаде також і в малий круг. Вважати, що ймовірність попадання точки в круг пропорційна площі круга і не залежить від його розташування.

Відповідь: $P = \dfrac{r^2}{R^2}$.

1.6. На складі є 15 телевізорів, причому 10 із них виготовлені вітчизняним виробником. Знайти ймовірність того, що серед 5 взятих навмання телевізорів 3 виявилось вітчизняного виробництва.

Відповідь: $P = C_{10}^{3} \cdot C_{5}^{2} \cdot \dfrac{1}{C_{15}^{5}} \approx 0,4$.

1.7. При випробуванні партії приладів відносна частота придатних приладів виявилася рівною 0,8. Знайти число придатних приладів, якщо було перевірено 300 приладів.

Відповідь: 240 приладів.

§2. Основні теореми

Теореми додавання і множення ймовірностей. Умовні ймовірності

Означення 1. Події A_1, A_2,..., A_n називаються *сумісними*, якщо поява однієї із них не виключає можливості появи інших і *несумісними*, якщо поява однієї виключає можливість появи будь-якої іншої.

Наприклад. *Сумісні*: постріл по мішені із трьох гармат. Очевидно, що не виключається можливість попадання в ціль із всіх трьох гармат.

Несумісні: кидання монети — випадання герба виключає можливість випадання решки.

Означення 2. *Сумою двох сумісних подій* A_1 і A_2 називається подія B, яка містить у собі появу хоча б однієї з подій A_1 або A_2 або обох цих подій. Аналогічно для (A_1, A_2,..., A_n).

Сумою двох несумісних подій A_1 і A_2 називається подія B, яка полягає в появі події A_1 або події A_2.

Теорема 1. Ймовірність суми двох несумісних подій дорівнює сумі ймовірностей цих подій.

$$P(A_1 + A_2) = P(A_1) + P(A_2) . \tag{1}$$

Доведення. Нехай із “n” випадків m_1 випадків сприяють події A_1 і m_2 — події A_2. Тоді $P(A_1) = \dfrac{m_1}{n}$; $P(A_2) = \dfrac{m_2}{n}$. Очевидно, що події $A_1 + A_2$ сприяє $m_1 + m_2$ випадків, причому $m_1 + m_2 \le n$, оскільки, згідно з умовою теореми, події A_1 і A_2 — несумісні. Серед m_1 випадків немає ні одного, які сприяли б події A_2 і серед m_2 — які сприяли б події A_1. Отже, $P(A_1 + A_2) = \dfrac{m_1 + m_2}{n} = \dfrac{m_1}{n} + \dfrac{m_2}{n} = P(A_1) + P(A_2)$.

Теорема узагальнюється і на n подій:

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n). \qquad (2)$$

Наслідок. Сума ймовірностей подій, які утворюють повну систему подій дорівнює одиниці.

Якщо $A_1 + A_2 + A_3 + \ldots + A_n = U \to$ вірогідна подія, то $P(A_1 + A_2 + + A_3 + \ldots + A_n) = P(U) = 1$ або $P(A_1) + P(A_2) + P(A_3) + \ldots + P(A_n) = 1$.

Означення 3. Подія A_1 називається *незалежною* від події A_2, якщо ймовірність події A_1 не залежить від того відбулася чи ні подія A_2.

Наприклад. Два стрілка стріляють у ціль. Нехай "попадання" в ціль одним стрілком — A_1, а другим — A_2. Очевидно, що ймовірність попадання одного стрілка не залежить від попадання іншого.

Означення 4. Подія A_1 називається *залежною* від події A_2, якщо ймовірність події A_1 залежить від того, відбулася чи ні подія A_2.

Наприклад. В ящику 3 білі і 2 чорні кулі. Ймовірність дістати білу кулю (подія A_1) дорівнює $\dfrac{3}{5}$. Якщо дістати чорну кулю (подія A_2), то ймовірність події A_1 після події A_2 дорівнює вже $\dfrac{3}{4}$. Отже, події A_1 і A_2 — залежні.

Ймовірність настання події A_1 за умови, що подія A_2 вже відбулася називається *умовною* і позначається: $P(A_1/A_2)$.

Наприклад. $P(A_1) = \dfrac{3}{5}$; $P(A_1/A_2) = \dfrac{3}{4}$; із означення 3 випливає, що якщо A_1 не залежить від A_2, то $P(A_1/A_2) = P(A_1)$.

Означення 5. *Добутком двох подій* A_1 і A_2 називається подія B, яка полягає в сумісній появі цих подій (позначається A_1A_2).

Аналогічно *добутком кількох подій* A_1, A_2,..., A_n називається подія B, яка полягає в сумісній появі всіх цих подій.

Теорема 2. Ймовірність добутку двох *залежних* подій дорівнює добутку ймовірностей однієї з них на умовну ймовірність іншої, обчислену за умови, що перша подія вже відбулася.

$$P(A_1 A_2) = P(A_1) \cdot P(A_2/A_1) \qquad (3)$$

або

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

Теорема може бути узагальнена на будь-яке скінчене число залежних подій.

$$P(A_1 \cdot A_2 \cdot \ldots \cdot A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1A_2) \cdot \ldots \cdot P(A_n/A_1A_2\ldots A_{n-1}) \qquad (4)$$

Теорема 3. Ймовірність добутку двох *незалежних* подій дорівнює добутку ймовірностей цих подій.

$$P(A_1 \cdot A_2) = P(A_1) \cdot P(A_2) \qquad (5)$$

Теорема узагальнюється на будь-яке скінчене число незалежних подій.

Приклад 1. В ящику 5 білих і 2 чорні кулі. Із нього дістають навмання 3 кулі. Знайти ймовірність того, що при першому виборі з'явиться біла куля (подія A_1), при другому — знову біла куля (подія A_2), а при третьому — чорна (подія A_3).

Події A_1, A_2 і A_3 — залежні. Нас цікавить поява і події A_1 і A_2, і A_3. Згідно з формулою (4) маємо:

$$P(A_1 \cdot A_2 \cdot A_3) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1A_2) = \frac{5}{7} \cdot \frac{4}{6} \cdot \frac{2}{5} = 0{,}19.$$

Приклад 2. Нехай маємо 3 ящики, які містять по 20 деталей. У першому ящику 16, у другому 14 і у третьому 18 стандартних деталей. Із кожного ящика навмання виймають по одній деталі. Знайти ймовірність того, що всі три вийняті деталі виявляться стандартними.

Ймовірність того, що із першого ящика вийнята стандартна деталь (подія A_1), з другого (подія A_2), з третього (подія A_3) дорівнюють відповідно:

$$P(A_1) = \frac{16}{20} = 0{,}8; \quad P(A_2) = \frac{14}{20} = 0{,}7; \quad P(A_3) = \frac{18}{20} = 0{,}9.$$

Оскільки події A_1, A_2 і A_3 — незалежні в сукупності, то шукана ймовірність (за теоремою множення) дорівнює:

$$P(A_1 \cdot A_2 \cdot A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3) = 0{,}504\,.$$

Формула повної ймовірності. Формула Байєса. Нехай подія B може відбутися лише разом з однією із подій (гіпотез):

$$A_1, A_2, \ldots, A_n,$$

які є *несумісні* і утворюють повну систему подій. Відомі ймовірності цих подій:

$$P(A_1),\ P(A_2),\ \ldots,\ P(A_n)$$

і умовні ймовірності події B:

$$P(B/A_1),\ P(B/A_2),\ \ldots,\ P(B/A_n)\,.$$

Оскільки подія B може відбутися тільки з однією із *несумісних* подій, то відповідно до поняття добутку подій:

$$B = A_1B + A_2B + \ldots + A_nB\,,$$

де події A_1B , A_2B , ..., A_nB є також несумісні.

Отже, $P(B) = P(A_1B) + P(A_2B) + \ldots + P(A_nB)$,

звідси $P(B) = P(A_1)\cdot P(B/A_1) + P(A_2)\cdot P(B/A_2) + \ldots + P(A_n)\cdot P(B/A_n)$

або:

$$P(B) = \sum_{i=1}^{n} P(A_i)\cdot P(B/A_i). \tag{6}$$

Це формула повної ймовірності.

Нехай виконується рівність: $B = A_1B + A_2B + \ldots + A_nB$, тобто подія B може з'явитися лише з однією з несумісних подій A_1, A_2, ..., A_n. Знайдемо ймовірність будь-якої події A_i ($i = 1, 2, \ldots, n$), за умови, що подія B відбулася.

Згідно з теоремою 2 маємо:

$$P(A_iB) = P(B)\cdot P(A_i/B) = P(A_i)\cdot P(B/A_i)\,,$$

звідси $P(A_i/B) = \dfrac{P(A_i)\cdot P(B/A_i)}{P(B)}$;

згідно з формулою (6) маємо:

$$P(A_i/B) = \frac{P(A_i)\cdot P(B/A_i)}{\sum_{i=1}^{n} P(A_i)P(B/A_i)}. \qquad (7)$$

Це формула Байєса або формула ймовірності гіпотез. Вона дає можливість здійснити переоцінки ймовірностей подій (гіпотез) A_1, A_2, ..., A_n після того, як стало відомо, що подія B настала.

Приклад 3. На двох заводах виготовляють будівельну цеглу, причому перший завод випускає 70% загальної кількості цегли, а другий — 30%. Продукція першого заводу містить 80% цегли високої якості, а другого — 90%. На будівництво постачається продукція обох заводів.

а) яка ймовірність того, що одержана на будівництві цегла буде високої якості?

Позначення: B — подія, яка полягає в тому, що цегла буде високої якості, A_1– цегла виготовлена на першому заводі, A_2 — на другому.

Із умови задачі:

$$P(A_1) = 0{,}7, \quad P(A_2) = 0{,}3, \quad P(B/A_1) = 0{,}8, \quad P(B/A_2) = 0{,}9.$$

Згідно з формою повної ймовірності маємо :

$$P(B) = P(A_1)\cdot P(B/A_1) + P(A_2)\cdot P(B/A_2) = 0{,}7\cdot 0{,}8 + 0{,}3\cdot 0{,}9 = 0{,}83.$$

б) знайти ймовірність того, що цегла високої якості виготовлена на першому заводі?

Згідно з формулою Байєса:

$$P(A_1/B) = \frac{P(A_1)\cdot P(B/A_1)}{P(B)} = \frac{0{,}7\cdot 0{,}8}{0{,}83} = 0{,}67.$$

Приклад 4. Нехай задано: три урни I типу: в ній знаходиться 2 білі і 6 чорних куль; одна урна II типу: в ній знаходиться 1 біла і 8 чорних куль. Нехай навмання вибирається урна, а звідти — куля. Позначимо через B — подію “добута куля — біла”, її ймовірність $P(B)$. Позначимо через A_1 подію — “вибрана урна I типу”, A_2– “вибрана урна II типу”. Тоді $B = A_1B + A_2B$.

Отже, згідно з формулою (6):

$$P(B) = P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2),$$

але $P(A_1) = \frac{3}{4}$, $P(A_2) = \frac{1}{4}$; $P(B/A_1) = \frac{2}{8} = \frac{1}{4}$; $P(B/A_2) = \frac{1}{9}$.

Звідси:

$$P(B) = \frac{3}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{9} = \frac{31}{144}.$$

Приклад 5. Проводиться той же експеримент, що і в попередньому прикладі. Нехай добута біла куля. Яка ймовірність того, що вона добута із урни I-го типу?

$$P(A_1/B) = \frac{P(A_1)P(B/A_1)}{P(A_1)P(B/A_1) + P(A_2)P(B/A_2)} = \frac{\frac{3}{4} \cdot \frac{1}{4}}{\frac{3}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{9}} = \frac{27}{31}.$$

Приклад 6. Нехай є три партії деталей по 40 деталей у кожній. Число стандартних деталей у першій, другій і третій партіях відповідно дорівнюють 40, 25 і 15. Із навмання вибраної партії навдачу добута деталь, яка виявилася стандартною. Деталь повертається в партію і вдруге із тієї ж партії навмання виймають деталь, яка теж виявляється стандартною. Знайти ймовірність того, що деталі були добуті з третьої партії.

Позначимо через B подію в кожному з двох випробувань (з поверненням) була добута стандартна деталь. Зробимо три припущення (гіпотези): A_1 — деталі діставались із першої партії, A_2 — деталь добувалася з другої партії, A_3 — деталь добувалась з третьої партії.

Деталі добувалися із навмання взятої партії, а тому ймовірності гіпотез однакові:

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}.$$

Знайдемо умовну ймовірність $P(B/A_1)$, тобто ймовірність того, що з першої партії будуть послідовно добуті дві стандартні деталі. Ця подія достовірна, оскільки в першій партії всі деталі стандартні, тому: $P(B/A_1)=1$.

Знайдемо умовну ймовірність $P(B/A_2)$, тобто ймовірність того, що із другої партії будуть послідовно добуті (з поверненням) дві стандартні деталі:

$$P(B/A_2)=\frac{25}{40}\cdot\frac{25}{40}=\frac{25}{64}.$$

Знайдемо умовну ймовірність $P(B/A_3)$, тобто ймовірність того, що із третьої партії будуть послідовно добуті (з поверненням) дві стандартні деталі:

$$P(B/A_3)=\frac{15}{40}\cdot\frac{15}{40}=\frac{9}{64}.$$

Шукана ймовірність того, що обидві добуті стандартні деталі взяті із третьої партії, по формулі Байєса дорівнює:

$$P(A_3/B)=\frac{P(A_3)\cdot P(B/A_3)}{\sum_{i=1}^{3}P(A_i)P(B/A_i)}=\frac{\frac{1}{3}\cdot\frac{9}{64}}{\frac{1}{3}\cdot 1+\frac{1}{3}\cdot\frac{25}{64}+\frac{1}{3}\cdot\frac{9}{64}}=\frac{9}{98}.$$

ЗАДАЧІ

1.8. Робітник обслуговує чотири однотипні станки. Ймовірність того, що будь-який станок протягом години буде вимагати уваги робітника, дорівнює 0,6. Вважаючи неполадки на станках незалежними, знайти ймовірність того, що протягом години потребують уваги робітника: 1) всі чотири станки; 2) ні один станок; 3) хоча б один станок.

Відповідь: 1) 0,1296; 2) 0,0256; 3) 0,9744.

1.9. Нехай є два набори деталей. Ймовірність того, що деталь першого набору стандартна — 0,8, а другого — 0,9. Знайти ймовірність того, що вибрана навмання деталь (із навмання вибраного набору) — стандартна.

Відповідь: 0,85.

1.10. У першій коробці міститься 20 радіоламп, із них 18 стандартні, в другій коробці — 10 ламп, із них 9 — стандартні. Із другої коробки навмання вибрана лампа і перекладена в першу. Знайти ймовірність того, що лампа навмання взята із першої коробки, буде стандартна.

Відповідь: 0,9.

1.11. Перша фірма із кожних 100 ламп виробляє в середньому 90 стандартних, друга — 95, третя — 85, а продукція цих фірм містить відповідно 50, 30 і 20% всіх електроламп, які постачаються в магазини певної території. Знайти ймовірність придбання стандартної електролампи.

Відповідь: 0,905.

1.12. Фірмою відряджений автомобіль за різними матеріалами на чотири бази. Ймовірність наявності необхідного матеріалу на першій базі дорівнює 0,9, на другій — 0,95, на третій — 0,8, на четвертій — 0,6. Знайти ймовірність того, що тільки на одній базі не виявиться необхідного матеріалу.

Відповідь: 0,4434.

1.13. З першого станка на зборку поступає 40%, з другого — 30%, з третього — 20%, з четвертого — 10% всіх деталей. Серед деталей першого станка — 0,1% бракованих, другого — 0,2%, третього — 0,25%, четвертого — 0,5%. Знайти ймовірність того, що на зборку надійшла бракована деталь.

Відповідь: 0,002.

1.14. У магазині продаються 4 телевізори. Ймовірність того, що телевізори витримають гарантійний строк служби, відповідно рівна 0,8; 0,85; 0,9; 0,95. Знайти ймовірність того, що куплений навмання телевізор витримає гарантійний строк служби.

Відповідь: 0,875.

§3. Повторні випробування

Нехай необхідно обчислити ймовірність того, що при n випробуваннях подія A відбудеться рівно k разів, і, отже, не відбудеться $n - k$ разів. Зауважимо, що розглядатимемо лише незалежні події, і такі, які можуть мати лише одну і ту ж ймовірність, окрім того, не обов'язково, щоб подія A повторилася k разів у певній послідовності.

Формула Бернуллі. Ймовірність однієї складної події, яка полягає в тому, що в n випробуваннях подія A відбудеться k разів і не відбудеться $n - k$ разів, згідно з теоремою множення ймовірностей незалежних подій дорівнює $p^k q^{n-k}$, де p– ймовірність події A в кожному випробуванні одна і та ж, а $q = 1 - p$ – ймовірність того, що подія A не відбулась у кожному випробуванні також стала. Таких складних подій може бути стільки, скільки можна скласти сполучень із n елементів по k елементів, тобто C_n^k. Оскільки ці події несумісні, то за теоремою додавання ймовірностей несумісних подій шукана ймовірність дорівнює сумі ймовірностей всіх можливих складних подій. З урахуванням того, що ймовірності всіх цих подій однакові маємо:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k},$$

або

$$P_n(k) = \frac{n!}{k!(n-k)!} \cdot p^k \cdot q^{n-k} \tag{1}$$

Символ $P_n(k)$ означає ймовірність того, що в n випробуваннях подія з'явилася рівно k разів і, отже, не відбудеться $n - k$ разів. Одержану формулу називають *формулою Бернуллі*.

Приклад 1. Ймовірність того, що витрати електроенергії протягом однієї доби не перевищують встановлені норми, дорівнює 0,65. Знайти ймовірність того, що в найближчі 7 діб витрати електроенергії протягом 5 діб не перевищать норми.

Ймовірність нормальних витрат електроенергії протягом кожних із 7 діб стала і дорівнює p = 0,65. Отже, ймовірність перевитрат електроенергії на кожну добу також стала і дорівнює:

$$q = 1 - p = 1 - 0{,}65 = 0{,}35 .$$

Шукана ймовірність за формулою Бернуллі дорівнює:

$$P_7(5) = C_7^5 \cdot p^5 \cdot q^2 = \frac{7!}{5!(7-5)!} \cdot (0{,}65)^5 \cdot (0{,}35)^2 \approx 0{,}298 .$$

Локальна теорема Муавра-Лапласа. Формулою Бернуллі при достатньо великих значеннях n користуватися не дуже вигідно, оскільки вона вимагає дій над великими числами і в процесі обчислень може накопичуватися похибка, а сам результат може суттєво відрізнятися від дійсності. Наведемо лише формулювання локальної теореми Муавра-Лапласа (доведення її досить складне), яка дозволяє наближено знайти ймовірність появи події рівно k разів у n випробуваннях, якщо число випробувань достатньо велике:

Ймовірність того, що в n незалежних випробуваннях, в кожному із яких ймовірність появи події дорівнює p *(*$0 < p < 1$*), подія відбудеться рівно k разів, наближено дорівнює (тим точніше, чим більше n):*

$$P_n(k) \approx \frac{1}{\sqrt{npq}}\varphi(x) \tag{2}$$

де

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

Таблиця функцій $\varphi(x)$ для додатних значень x наведена в додатку 1; для від'ємних значень x користуються тією ж таблицею, оскільки функція $\varphi(x)$ — парна, тобто $\varphi(-x) = \varphi(x)$.

Приклад 2. Знайти ймовірність того, що подія A відбудеться рівно 50 разів у 200 експериментах, якщо ймовірність появи цієї події в кожному експерименті дорівнює 0,3.

Згідно з умовою, $n = 200$; $k = 50$; $p = 0,3$; $q = 0,7$. Скористаємося асимптотичною формулою Муавра-Лапласа:

$$P_{200}(50) \approx \frac{1}{\sqrt{200 \cdot 0,3 \cdot 0,7}} \cdot \varphi(x) = \frac{1}{10\sqrt{0,42}} \cdot \varphi(x).$$

Знайдемо значення x:

$$x = \frac{k-np}{\sqrt{npq}} = \frac{50 - 200 \cdot 0,3}{\sqrt{200 \cdot 0,3 \cdot 0,7}} = -\frac{10}{10\sqrt{0,42}} = -\frac{1}{0,65} = -1,54.$$

Функція $\varphi(x)$ парна, тому $\varphi(-1,54) = \varphi(1,54)$. Згідно з таблицею додатку 1 знайдемо $\varphi(1,54) = 0,1219$.

Шукана ймовірність:

$$P_{200}(50) \approx \frac{1}{6,5} \cdot 0,1219 = 0,0188 .$$

Формула Бернуллі наближено приведе до аналогічного результату (перевірити самостійно). Формула Лапласа приводить до прийнятного результату лише при достатньо великих значеннях *n*.

При малих значеннях *n* формули Бернуллі і Лапласа приведуть до розхожих відповідей.

Інтегральна теорема Лапласа. *Ймовірність того, що в n незалежних випробуваннях, у кожному із яких ймовірність появи події дорівнює p (* $0 < p < 1$*), подія відбудеться не менше* k_1 *разів і не більше* k_2 *разів наближено дорівнює:*

$$P_n(k_1, k_2) \approx \Phi(x'') - \Phi(x') \tag{3}$$

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

$$x' = \frac{k_1 - np}{\sqrt{npq}}; \quad x'' = \frac{k_2 - np}{\sqrt{npq}} .$$

Таблиця функцій Лапласа (оскільки невизначений інтеграл $\int_0^x e^{-(\frac{t^2}{2})} dt$ не виражається через елементарні функції) для додатних значень x ($0 \le x \le 5$) наведена в додатку 2; для значень $x > 5$ вважають $\Phi(x) = 0,5$. Для від'ємних значень *x* використовують ту ж таблицю, враховуючи, що функція Лапласа непарна, тобто $\Phi(-x) = -\Phi(x)$. Проілюструємо застосування наведеної теореми на прикладі.

Приклад 3. Ймовірність того, що деталь не пройшла перевірки ВТК, дорівнює *p* = 0,1. Знайти ймовірність того, що серед 500 випадково відібраних деталей випадуть неперевіреними від 60 до 80 деталей.

Згідно з умовою, $p = 0,1$; $q = 0,9$; $n = 500$; $k_1 = 60$; $k_2 = 80$.

Скористаємося інтегральною теоремою Лапласа:

$$P_{500}(60, 80) \approx \Phi(x'') - \Phi(x') .$$

Обчислимо нижню і верхню межу інтегрування x' і x'':

$$x' = \frac{k_1 - np}{\sqrt{npq}} = \frac{60 - 500 \cdot 0{,}1}{\sqrt{500 \cdot 0{,}1 \cdot 0{,}9}} = \frac{10}{\sqrt{45}} = 1{,}49,$$

$$x'' = \frac{k_2 - np}{\sqrt{npq}} = \frac{80 - 500 \cdot 0{,}1}{\sqrt{500 \cdot 0{,}1 \cdot 0{,}9}} = \frac{30}{\sqrt{45}} = 4{,}48,$$

Отже, $P_{500}(60,80) \approx \Phi(4{,}48) - \Phi(1{,}49)$.

Згідно з таблицями додатку маємо:

$$\Phi(4{,}48) = 0{,}499997\,; \quad \Phi(1{,}49) = 0{,}4319\,.$$

Шукана ймовірність:

$$P_{500}(60,80) \approx 0{,}5 - 0{,}4319 = 0{,}0681\,.$$

Зауваження. Інтегральну теорему Лапласа часто застосовують у вигляді:

$$P(x' \le \frac{m - np}{\sqrt{npq}} \le x'') \approx \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-\frac{t^2}{2}} dt \qquad (4)$$

де m — число появи події A при n незалежних випробуваннях змінюється від k_1 до k_2, а дріб $\frac{m - np}{\sqrt{npq}}$ змінюється від $x' = \frac{k_1 - np}{\sqrt{npq}}$ до $x'' = \frac{k_2 - np}{\sqrt{npq}}$.

Використовуючи вказану форму запису можна встановити, що ймовірність того, що в n незалежних випробуваннях, у кожному із яких ймовірність появи події дорівнює p ($0 < p < 1$), абсолютна величина відхилення відносної частоти появи події від ймовірності появи події не перевищить додатного числа ε, наближено дорівнює подвоєній функції Лапласа при $x = \varepsilon\sqrt{\frac{n}{pq}}$:

$$P(\left|\frac{m}{n} - p\right| \le \varepsilon) \approx 2\Phi(\varepsilon\sqrt{\frac{n}{pq}}). \qquad (5)$$

Вказівка. Замінивши нерівність $\left|\frac{m}{n}-p\right|\le\varepsilon$, їй рівнозначною $-\varepsilon\le\frac{m}{n}-p\le\varepsilon$, помножимо її на множник $\sqrt{\frac{n}{pq}}$ та скористаємося інтегральною теоремою Лапласа у формі $P(x'\le\frac{m-np}{\sqrt{npq}}\le x'')\approx\frac{1}{\sqrt{2\pi}}\int_{x'}^{x''}e^{-\frac{t^2}{2}}dt$ з урахуванням $x'=-\varepsilon\cdot\sqrt{\frac{n}{pq}}$ та $x''=\varepsilon\cdot\sqrt{\frac{n}{pq}}$.

Приклад 4. Ймовірність того, що деталь нестандартна p = 0,2. Знайти ймовірність того, що серед випадково відібраних 300 деталей відносна частота появи нестандартних деталей відхилиться від ймовірності p = 0,2 по абсолютній величині не більше ніж на 0,01.

За умовою: n = 300; p = 0,2; q = 0,8; ε = 0,01.

Потрібно знайти ймовірність:

$$P(\left|\frac{m}{300}-0,2\right|)\le 0,01.$$

Користуючись формулою (5), маємо:

$$P(\left|\frac{m}{300}-0,2\right|\le 0,01)\approx 2\Phi(0,01\sqrt{\frac{300}{0,2\cdot 0,8}})=2\Phi(0,43).$$

За таблицею додатку 2 знаходимо Ф(0,43) = 0,1664, отже, 2Ф(0,43) = 0,3328.

Зміст отриманого результату такий: якщо взято достатньо велике число проб по 300 деталей в кожній, то приблизно для 33,28% цих проб відхилення відносної частоти від сталої ймовірності p = 0,2 по абсолютній величині не перевищить 0,01.

ЗАДАЧІ

1.15. В урні 20 білих куль і 10 чорних куль. Добули підряд 4 кулі, причому кожна добута куля повертається в урну перед діставанням наступної і кулі в урні перемішуються. Яка ймовірність того, що із чотирьох добутих куль будуть дві білі?

Відповідь: $P_4(2)=\frac{8}{27}$.

1.16. У цеху 6 моторів. Для кожного мотора ймовірність того, що він у даний момент включений дорівнює 0,8. Знайти ймовірність того, що в даний момент:

а) включені 4 мотори;
б) включені всі мотори;
в) виключені всі мотори.

Відповідь: а) $P_6(4) = 0,246$;
б) $P_6(6) = 0,26$;
в) $P_6(0) = 0,000064$.

1.17. Знайти ймовірність того, що подія *A* настане рівно 70 разів у 243 випробуваннях, якщо ймовірність настання цієї події у кожному випробуванні дорівнює 0,25.

Відповідь: $P_{243}(70) = 0,0231$.

1.18. Знайти наближено ймовірність того, що при 400 випробуваннях подія настане рівно 104 рази, якщо ймовірність її появи в кожному експерименті дорівнює 0,2.

Відповідь: $P_{400}(104) = 0,0006$.

1.19. Ймовірність того, що деталь нестандартна $p = 0,1$. Знайти скільки деталей потрібно відібрати, щоб з ймовірністю, рівною 0,9544, можна було стверджувати, що відносна частота появи нестандартних деталей (серед відібраних) відхилилася від сталої ймовірності p по абсолютній величині не більше ніж на 0,03.

Відповідь: $n = 400$.

1.20. Ймовірність настання події у кожному із 100 незалежних випробувань стала і дорівнює $p = 0,8$. Знайти ймовірність того, що подія настане:

а) не менше 75 разів і не більше 90 разів;
б) не менше 75 разів ;
в) не більше 74 рази.

Відповідь: а) $P_{100}(75;90) = 0,8882$;
б) $P_{100}(75;100) = 0,8944$;
в) $P_{100}(0;74) = 0,1056$.

1.21. Ймовірність настання події у кожному із незалежних випробувань дорівнює 0,8. Скільки потрібно провести експериментів, щоб з ймовірністю 0,9 можна було сподіватися, що подія настане не менше 75 разів?

Відповідь: $n = 100$.

§4. Випадкові величини та їх розподіл

Закон розподілу дискретної випадкової величини. Одним із основних понять у теорії ймовірності є поняття випадкової величини.

Величина називається *випадковою*, якщо в результаті випробування вона приймає із множини можливих своїх значень одне і тільки одне наперед невідоме можливе значення, залежне від випадкових причин, які врахувати неможливо. Випадкові величини бувають *дискретні і неперервні*.

Приклад: а) *Дискретної величини*. Монету кидають один раз, при цьому герб може випасти або 0 раз або 1 раз, тобто випадкова величина (частота появи герба) може прийняти тільки одне із цих двох можливих значень.

б) *Неперервна випадкова величина*. Відстань польоту снаряду — випадкова величина, яка може прийняти будь-яке, але тільки одне значення на деякому проміжку, який є множиною можливих її значень.

Дискретна випадкова величина характеризується значеннями, які вона може приймати, і ймовірностями, з якими ці значення з'являються.

Множина всіх можливих значень дискретної випадкової величини з їх ймовірностями називається законом розподілу цієї випадкової величини.

Нехай нам відомі всі можливі значення $x_1, x_2, x_3, ..., x_n$ випадкової величини X (випадкові величини як правило позначаються великими буквами, а їх значення маленькими) із їх ймовірностями $p_1, p_2, p_3, ..., p_n$ — відповідно. Випадкова величина X при одному випробуванні приймає тільки одне із n можливих своїх значень із відповідною ймовірністю, тобто в результаті випробування обов'язково відбудеться одна із єдино можливих подій:

$$X = x_1;\ X = x_2;\ X = x_3;\ ...\ X = x_n.$$

Такі події утворюють повну систему подій і, отже, сума ймовірностей цих подій:

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

Запишемо *закон розподілу випадкової величини* формулою:

$$P(X = x_i) = p_i, \text{ де } (i = 1, 2,..., n)$$

або у вигляді таблиці:

X	x_1	x_2	x_3	...	x_n
$P(X = x_i)$	p_1	p_2	p_3	...	p_n

Приклад 1. У грошовій лотереї розігрується один виграш в 1000 гривень, 10 виграшів по 100 гривень і 100 виграшів по 1 гривні, при загальному числі білетів 10000. Знайти закон розподілу випадкового виграшу X для власників одного лотерейного білета.

Тут можливі значення для X є
$x_1 = 1000$, $x_2 = 100$, $x_3 = 1$, $x_4 = 0$.
Ймовірності їх відповідно будуть:
$p_1 = 0{,}0001$, $p_2 = 0{,}001$, $p_3 = 0{,}01$, $p_4 = 1 - (p_1 + p_2 + p_3) \approx 0{,}9889$.
Закон розподілу для виграшу X може бути заданий таблицею:

X	1000	100	1	0
P	0,0001	0,001	0,01	0,9889

Закон розподілу випадкової величини можливо задати графічно, для цього будують так званий полігон розподілу ймовірності.

Приклад 2. Ймовірність виготовлення нестандартної залізобетонної балки на заводі стала і дорівнює 0,05. Для перевірки якості контролер бере із партії не більше 4 виробів, при знаходженні нестандартної балки вся партія затримується. Скласти закон розподілу числа виробів, які перевіряє контролер і побудувати полігон розподілу цієї випадкової величини X.

Можливі значення шуканої випадкової величини X є числа 1, 2, 3, 4. Контролер перевіряє тільки одну балку $x_1 = 1$, якщо вона буде нестандартною, то партія затримується. Згідно з умовою: $P(X = 1) = 0{,}05$. Якщо перша балка стандартна, а друга — нестандартна, то контролер перевіряє два вироби: $x_2 = 2$. За теоремою множення ймовірностей $P(X = 2) = 0{,}95 \cdot 0{,}05 = 0{,}048$.

Контролер перевіряє три балки, якщо перші дві є стандартними, а третя нестандартна, тоді $P(X=3)=0,95\cdot 0,95\cdot 0,05=0,045$. Четверту балку перевіряють тоді, коли перші три стандартні, тоді $P(X=4)=(0,95)^3=0,857$.

Складемо таблицю закону розподілу.

X	1	2	3	4
P	0,05	0,048	0,045	0,857

Полігон розподілу випадкової величини X буде:

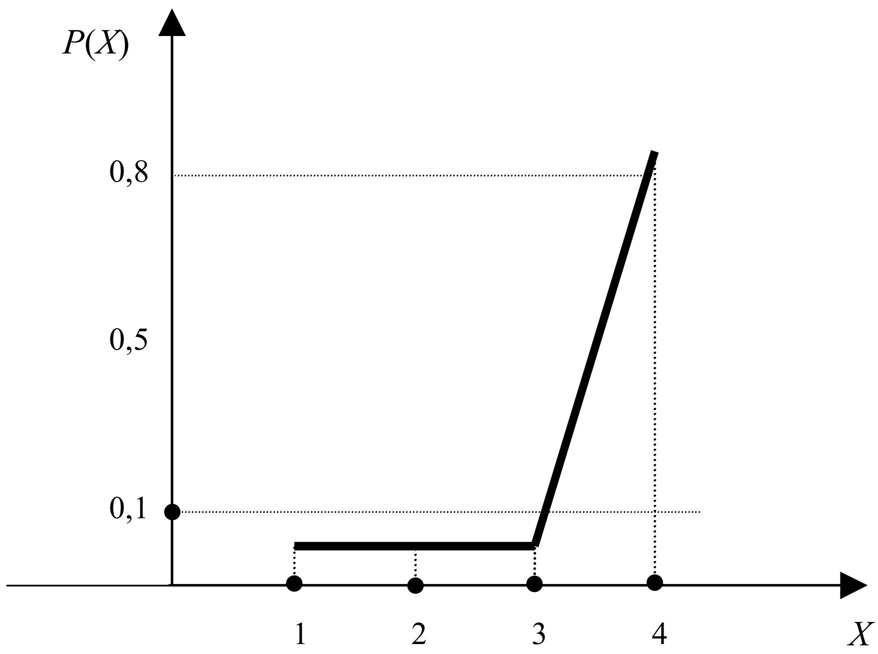


Рис. 1

Отже, дискретна випадкова величина характеризується значеннями $x_1, x_2, x_3,...$, які вона приймає, і ймовірностями $p_i, = P(X = x_i)$, з якими вона приймає ці значення і які повинні задовольняти умові $\sum_{i=1}^{n} p_i = 1$.

Функція розподілу випадкової величини та її властивості. Але попередній спосіб неприйнятний для неперервної випадкової величини X можливі значення якої суцільно заповнюють інтервал $(a; b)$. Таким чином, доцільно дати загальний спосіб задання будь-яких таких випадкових величин. З цією метою і вводять функцію розподілу ймовірності випадкових величин.

Нехай X — випадкова величина і x — довільне дійсне число. Функцією розподілу ймовірностей випадкової величини X називається ймовірність того, що випадкова величина X приймає значення менше ніж “x”:

$$F(x) = P(X < x) \tag{1}$$

або

$$F(x) = P(-\infty < X < x)\,. \tag{2}$$

Зрозуміло, що якщо x змінюється, то змінюється і $F(x)$, тобто $F(x)$ — функція від x.

Функція (1) іноді називається інтегральною функцією розподілу.

Введемо наступні визначення:

1. *Випадковою величиною* називається змінна величина, значення якої залежать від випадку і для якої визначена функція розподілу ймовірності.

2. *Випадкова величина неперервна*, якщо $F(x)$ — неперервно-диференційовна.

Розглянемо властивості функції розподілу $F(x)$, ураховуючи що “x” приймає будь-які числові значення на інтервалі $(-\infty;\infty)$.

Властивість 1. Згідно з визначенням (1) $F(x)$– ймовірність, отже,

$$0 \le F(x) \le 1 \tag{3}$$

Властивість 2. $F(x)$ — монотонно зростає в широкому змісті, тобто якщо $x_1 < x_2$, то $F(x_1) \le F(x_2)$.

Властивість 3. Справедливе співвідношення : $F(+\infty) = 1$

Властивість 4. Справедливе співвідношення: $F(-\infty) = 0$.

Властивість 5. $P(x_1 \le X \le x_2) = F(x_2) - F(x_1)$.

Властивість 6. Якщо всі значення дискретної або неперервної випадкової величини лежать на інтервалі $(a;\ b)$, то: $F(x) = 0$ при $x \le a$ (оскільки значення випадкової величини X, менші a неможливі) і $F(x) = 1$ при $x \ge b$ (оскільки подія $X < x$ при $x \ge b$ вірогідна).

Зазначені властивості дозволяють представити графік функції розподілу неперервної випадкової величини (рис. 2), (рис. 3).

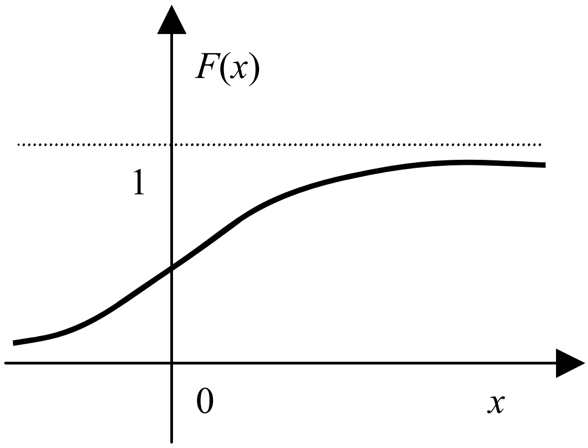


Рис. 2

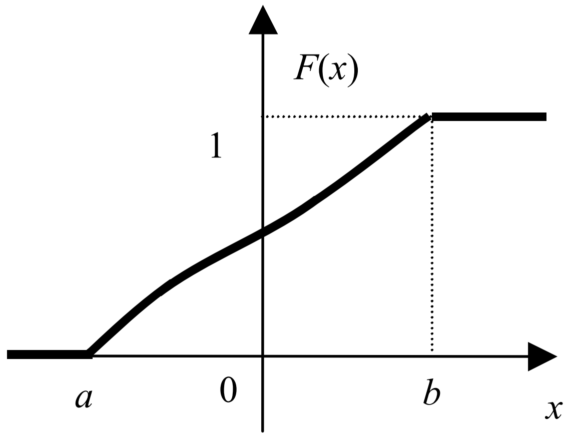


Рис. 3

Приклад 3. В ящику дві білі кулі і три чорні кулі. Із ящика дістають одну кулю. Знайти функцію розподілу числа появи білої кулі і побудувати її графік.

Випадкова величина X — число появи білої кулі, її закон розподілу можна представити у вигляді таблиці:

X	0	1
$P(X)$	0,6	0,4

Із таблиці і визначення функції розподілу випадкової величини X випливає:

1) якщо $x \le 0$, то $F(x) = 0$, оскільки лівіше нуля немає можливих значень X і тому подія неможлива.

2) якщо $0 < x \le 1$, то $F(x) = 0{,}6$, оскільки лівіше одиниці є можливе значення випадкової величини $X = 0$, ймовірність якої $P(X = 0) = 0{,}6$.

3) якщо $x > 1$, $F(x) = P(X = 0) + P(X = 1) = 0{,}6 + 0{,}4 = 1$, тобто при $x > 1$ подія $X < x$ вірогідна.

Отже, функція розподілу має вигляд:

$$F(x) = \begin{cases} 0, \ \textit{якщо} \ \ x \le 0 \\ 0{,}6, \ \textit{якщо} \ \ 0 < x \le 1, \\ 1, \ \textit{якщо} \ \ x > 1 \qquad . \end{cases}$$

Графік зображений на рис. 4.

Зауваження. $F(x)$ у точках $x = 0$, $x = 1$ має розрив 1-го роду, причому вона в цих точках неперервна зліва і має розрив справа.

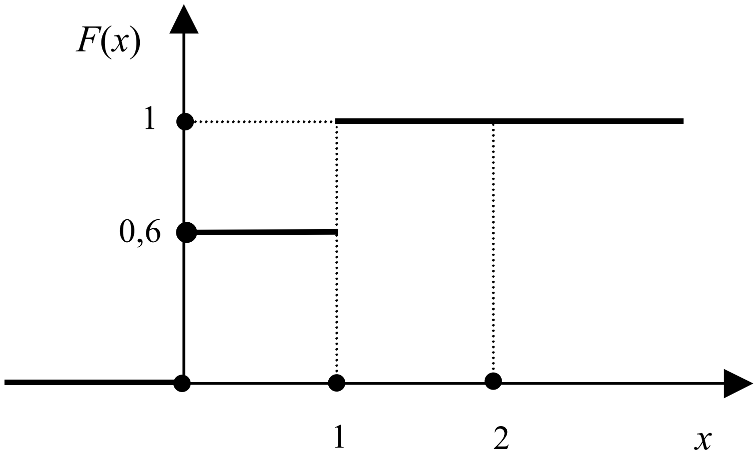


Рис. 4

Функція $F(x)$ існує як для неперервних, так і для дискретних випадкових величин. Графік функції розподілу дискретної випадкової величини має ступінчастий вигляд.

Однозначне відображення множини x_i на множину p_i розглядається як функція ймовірності дискретної випадкової величини. Для функції розподілу випадкової величини маємо:

$$F(x) = \sum_{x_i < x} p_i \ .$$

Підсумок проводиться за всіма i, для яких $x_i < x$. Також $F(x)$ є ступінчастою функцією зі стрибками висотою p_i у точках x_i (рис. 5).

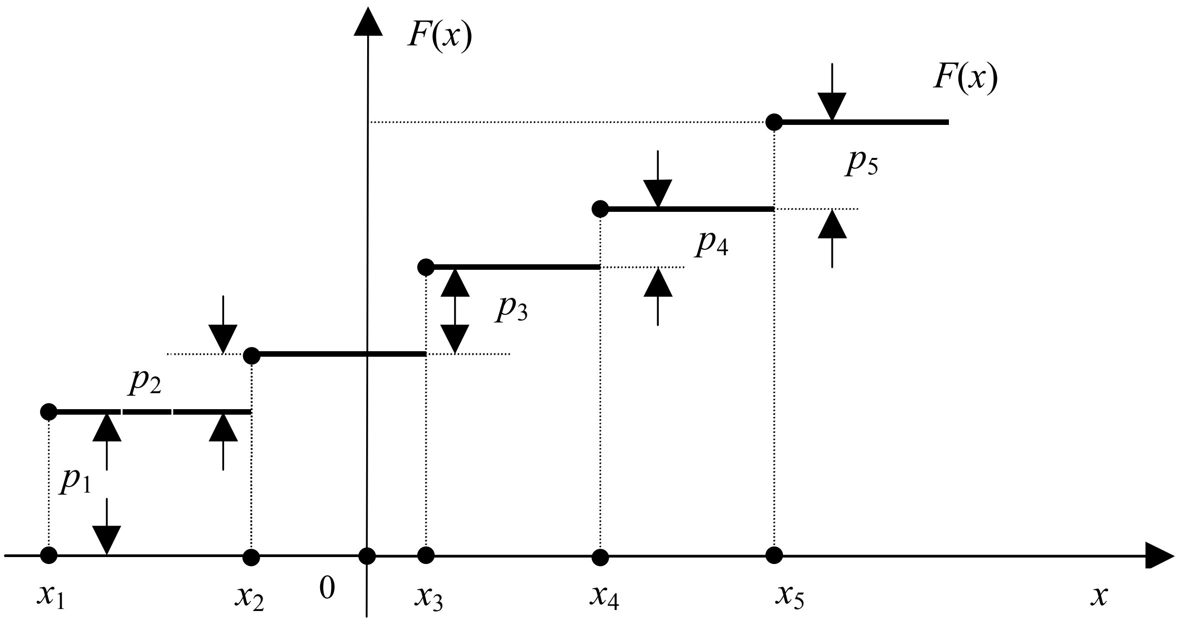


Рис. 5

Переконаємося в цьому на прикладі.

Приклад 4. Заданий ряд розподілу випадкової величини *X*:

x_i	10	20	30	40	50
p_i	0,15	0,25	0,45	0,1	0,05

Побудувати функцію розподілу ймовірності випадкової величини.
Якщо $x \le 10$, то $F(x) = P(X < x) = 0$;

$10 < x \le 20$, то $F(x) = P(X < x) = 0{,}15$

$20 < x \le 30$, то $F(x) = P(X < x) = 0{,}15 + 0{,}25 = 0{,}4$

$30 < x \le 40$, то $F(x) = P(X < x) = 0{,}15 + 0{,}25 + 0{,}45 = 0{,}85$

$40 < x \le 50$, то $F(x) = P(X < x) = 0{,}15 + 0{,}25 + 0{,}45 + 0{,}1 = 0{,}95$

$x > 50$, то $F(x) = 0{,}95 + 0{,}05 = 1$.

Отже, функція розподілу аналітично може бути записана так:

$$F(x) = \begin{cases} 0, \ \textit{якщо} \ \ x \le 10 \\ 0{,}15, \ \textit{якщо} \ \ 10 < x \le 20 \\ 0{,}4, \ \textit{якщо} \ \ 20 < x \le 30 \\ 0{,}85, \ \textit{якщо} \ \ 30 < x \le 40 \\ 0{,}95, \ \textit{якщо} \ \ 40 < x \le 50 \\ 1{,}00 \ \ \textit{якщо} \ \ x > 50 \end{cases}.$$

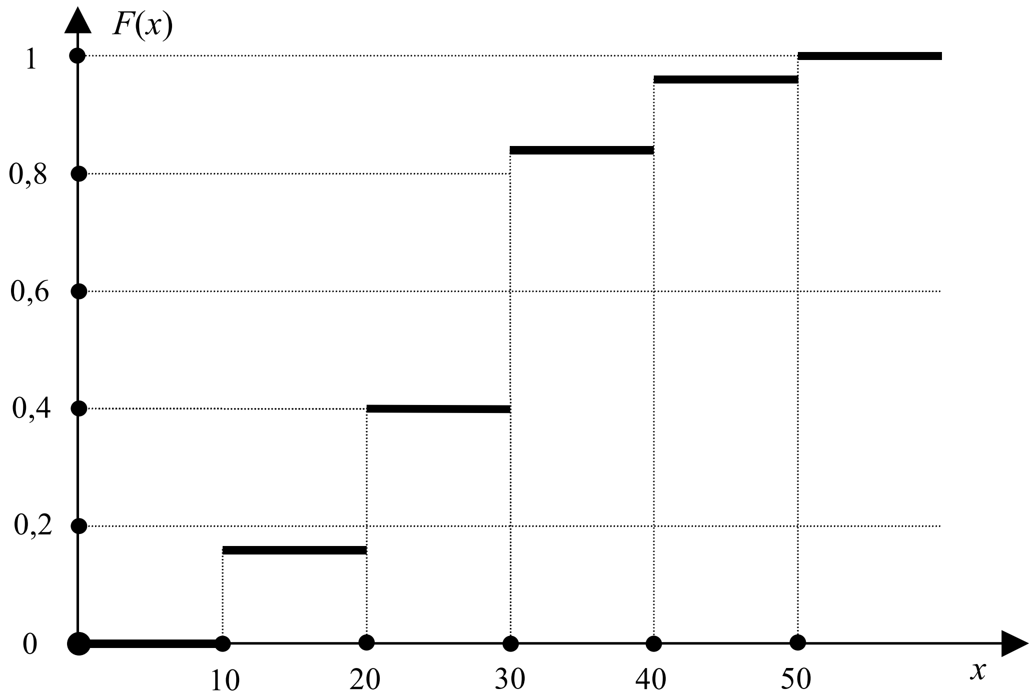


Рис. 6

Розглянемо деякі закони розподілу дискретної випадкової величини.

Біноміальний розподіл. Розглянемо як дискретну випадкову величину X число настання події A в n незалежних випробуваннях у кожному із яких подія A настане або не настане. Ймовірність настання події у всіх експериментах стала і дорівнює p, отже, ймовірність ненастання $p = 1 - q$.

Для знаходження закону розподілу величини X потрібно визначити можливі значення X та їх ймовірності. Очевидно подія A в експериментах може або не настати, або з’явитися 1 раз, або 2 рази, ..., або n раз. Таким чином, можливі значення X такі: $x_1 = 0$, $x_2 = 1$, ..., $x_{n+1} = n$. Залишається знайти ймовірність цих можливих значень, для цього достатньо скористатися формулою Бернуллі:

$$P_n(k) = C_n^k p^k q^{n-k}, \tag{4}$$

де $k = 0, 1, 2,..., n$.

Розподіл ймовірностей, визначених за формулою Бернуллі, називають *біноміальним*, оскільки праву частину рівності можна розглядати як загальний член розкладу біному Ньютона:

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

Отже, перший член розкладу p^n визначає ймовірність настання події n разів у n незалежних випробуваннях, другий член $np^{n-1}q$ визначає ймовірність настання події (n – 1) разів; ...; останній член q^n визначає ймовірність того, що подія не настане ні разу. Табличний запис біноміального закону:

X	n	$n-1$	....	k	....	0
P	p^n	$np^{n-1}q$	....	$C_n^k p^k q^{n-k}$	....	q^n

Приклад 5. Прилад складається із трьох незалежних працюючих елементів. Ймовірність відмови кожного елемента в одному експерименті дорівнює 0,2. Скласти закон розподілу числа елементів, які відмовили в одному експерименті.

Дискретна випадкова величина X (кількість елементів, що відмовили в одному експерименті) має наступні можливі значення: $x_1 = 0$ (ні один із елементів приладу не відмовив), $x_2 = 1$ (відмовив один елемент), $x_3 = 2$ (відмовили два елементи), $x_4 = 3$ (відмовили три елементи).

Відмови елементів незалежні один від одного, ймовірності відмови кожного елемента рівні між собою, а тому застосуємо формулу Бернуллі. Враховуючи, що за умовою $n = 3$, $p = 0{,}2$ (отже, $q = 1 - 0{,}2 = 0{,}8$,), одержимо $P_3(0) = q^3 = 0{,}8^3 = 0{,}512$, $P_3(1) = C_3^1 pq^2 = 3 \cdot 0{,}2 \cdot 0{,}8^2 = 0{,}384$, $P_3(2) = C_3^2 p^2 q = 3 \cdot 0{,}2^2 \cdot 0.8 = 0{,}096$, $P_3(3) = p^3 = 0{,}2^3 = 0{,}008$.

Перевірка: $0{,}512 + 0{,}384 + 0{,}096 + 0{,}008 = 1$.

Шуканий біноміальний закон розподілу X:

X	0	1	2	3
P	0,512	0,384	0,096	0,008

Розподіл Пуассона. Знайдемо ймовірність того, що при дуже великому числі n незалежних випробувань, у кожному з яких ймовірність події дуже мала, подія настане рівно k разів. Зауважимо, що при невеликому числі n користуються попередньою формулою Бернуллі. Якщо ж n велике, то вигідно скористатися асимптотичною формулою Лапласа. Але ця формула неприйнятна, якщо ймовірність події мала ($p \le 0{,}1$). У цих випадках (n велике, p мале) звертаються до асимптотичної формули Пуассона.

Скористаємося формулою Бернуллі для обчислення ймовірності, яка нас цікавить:

$$P_n(k) = \frac{n(n-1)(n-2)...(n-k+1)}{k!} p^k (1-p)^{n-k}.$$

Зробимо важливе припущення: добуток np зберігає стале значення, зокрема $np = \lambda$. У подальшому це буде означати, що середнє число настання події у різних серіях випробувань, тобто при різних значеннях n, залишається незмінним. Оскільки $np = \lambda$, то $p = \frac{\lambda}{n}$. Отже,

$$P_n(k) = \frac{n(n-1)(n-2)...(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

Спрямовуючи $n \to \infty$ ми при цьому знайдемо лише наближене значення шуканої ймовірності тому, що n хоча й велике, але в кожному випадку скінченне.

Зауважимо, що оскільки np зберігає стале значення, то при $n \to \infty$ ймовірність $p \to 0$.

Отже,

$$P_n(k) \approx \lim_{n\to\infty} \frac{n(n-1)(n-2)...(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty}\left[1\cdot\left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\cdot\ldots\cdot\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-k}\right] = \frac{\lambda^k}{k!}e^{-\lambda}\cdot 1.$$

Закон розподілу Пуассона ймовірностей масових (n велике) і рідкісних (p мале) подій виражається формулою:

$$P_n(k) = \frac{\lambda^k}{k!}e^{-\lambda}. \quad (5)$$

Зауваження. Для простоти запису знак наближено опущений, крім того існують спеціальні таблиці, користуючись якими можливо знаходити $P_n(k)$, знаючи k і λ.

Приклад 6. Підприємство відправило на базу 1000 якісних виробів. Ймовірність того, що по дорозі виріб пошкодиться, дорівнює 0,0001. Знайти ймовірність того, що у продаж надійдуть 5 неякісних виробів.

За умовою n = 1000, p = 0,0001, k = 5. Знайдемо λ: $\lambda = np = 1000 \cdot 0{,}0001$.

За формулою Пуассона шукана ймовірність наближено дорівнює:

$$P_{1000}(5) = \frac{\lambda^k}{k!}e^{-\lambda} = e^{-1}\cdot\frac{1}{5!} = \frac{1}{120e} \approx 0{,}003.$$

Геометричний розподіл. Нехай незалежні випробування в кожному із яких ймовірність настання події A рівна p ($0 < p < 1$) закінчуються як тільки настане подія A.

Таким чином, якщо подія A настала на k-ому експерименті, то на попередніх k – 1 випробуваннях вона не з'явилась, крім того, ймовірність її ненастання $q = 1 - p$. Позначимо через X дискретну випадкову величину — число випробувань, які необхідно провести до першого настання події A. Нехай у перших k – 1 експериментах подія A не настала , а в k-ому експерименті настала.

Згідно з теоремою множення ймовірностей незалежних подій ймовірність такої події дорівнює:

$$P(X=k)=q^{k-1}p\,. \qquad (6)$$

Поклавши $k = 1, 2, 3...$ одержимо геометричну прогресію з першим членом p і знаменником q $(0 < q < 1)$:

$$p\,,\ qp\,,\ q^2p\,,\ \ldots,\ q^{k-1}p\,,\ \ldots\,. \qquad (*)$$

З цієї причини розподіл (6) називається *геометричним*. Легко переконатися, що ряд (*) збігається і його сума дорівнює 1.

Дійсно,

$$S=\frac{p}{1-q}=\frac{p}{p}=1\,.$$

Приклад 7. Спортсмен виконує стрільбу по мішені до першого попадання в десятку. Ймовірність такого його попадання p = 0,7. Знайти ймовірність того, що попадання настане на четвертому пострілі.

За умовою p = 0,7, q = 0,3, k = 4.

Шукана ймовірність за формулою (6) $P=q^{k-1}p=0{,}3^3\cdot 0{,}7=0{,}019$.

Приклад 8. Скільки потрібно купити лотерейних білетів, щоб ймовірність виграшу була не менша ніж p?

Нехай загальна кількість лотерейних білетів дорівнює N і M — загальна кількість виграшів. Тоді ймовірність того, що куплений лотерейний білет виявиться із числа M виграшних білетів, дорівнює $\frac{M}{N}$. Придбання кожного окремо білета можливо розглядати як окреме випробування з ймовірністю "успіху" $p=\frac{M}{N}$ у серії із n незалежних випробувань (n — число куплених білетів). Якщо вважати, що ймовірність p мала, як це звичайно буває, а задана ймовірність P порівняно велика, то очевидно, що необхідно купити досить велику кількість лотерейних білетів, щоб ймовірність хоча б одного була б не менше P. А тому випадкове число виграшних білетів приблизно розподілено за законом Пуассона, тобто ймовірність того, що серед куплених n білетів виявиться рівно k виграшних, є

$$P(k)\approx\frac{a^k}{k!}e^{-a}\,,$$

де $a = n \cdot \dfrac{M}{N}$. Ймовірність того, що хоча б один із білетів буде виграшним, є $1 - P(0) = 1 - e^{-a}$, тим самим число n слід визначити як найменше ціле число для якого $e^{-a} \le 1 - P$.

Приклад 9. Нехай маємо певну кількість тіста *V*, із якого випікаються булочки з родзинками. Певну кількість родзинок *n* висипається в тісто, після чого все ретельно перемішується, а потім розрізається на рівні частини. Вважатимемо, що на окрему булку витрачається кількість тіста *v*, таким чином, всього випікається $N = \dfrac{V}{v}$ булок з родзинками. Очевидно, що хоча середні витрати родзинок на окрему булку складає цілком визначену величину $a = n \cdot \dfrac{v}{V}$, кількість родзинок в різних булках зовсім неоднакова. Яка ймовірність того, що в окремо взятій, випадково відібраній булочці виявиться хоча б одна родзинка?

Природно вважати, що кількість родзинок набагато менша кількості тіста, оскільки при перемішуванні тіста родзинки у решті-решт рухаються практично незалежно одна від одної, а тому незалежно одна від одної попадають чи не попадають у вибрану булку. Очевидно, після ретельного перемішування родзинки розподіляються в тісті приблизно рівномірно, так що попадання будь-якої із родзинок в будь-яку із булок одна і та ж, тобто $\dfrac{1}{N} = \dfrac{v}{V}$. Попадання окремої родзинки у певну булку можливо розглядати як “успіх” в окремому випробуванні, ймовірність якого є $p = \dfrac{v}{V}$. Незалежність руху родзинок при перемішуванні дозволяє вважати, що є *n* випробувань Бернуллі з ймовірністю “успіху” *p*, *n* — загальне число ізюминок, Така ймовірність порівняно мала, якщо булок випікається досить багато. У той же час число родзинок *n* порівняно велике. Отже, випадкове число родзинок в окремій булці рівне числу “успіхів” і приблизно розподілено за законом Пуассона : ймовірність *P*(*k*) того, що в булці виявиться рівно *k* родзинок, є

$$P(k) \approx \frac{a^k}{k!} \cdot e^{-a},$$

де $a = np = n \cdot \dfrac{v}{V}$ — середнє число родзинок, які припадають на одну булку. Ймовірність P того, що в булці виявиться хоча б одна родзинок, є

$$P = 1 - P(0) = 1 - e^{-a}.$$

Щільність розподілу ймовірностей. Закон розподілу неперервної випадкової величини зручно задавати за допомогою так званої функції щільності розподілу ймовірності $f(x)$.

Щільністю розподілу ймовірностей називається похідна від функції розподілу $F(x)$:

$$f(x) = F'(x) \tag{7}$$

Графік щільності розподілу ймовірностей $y = f(x)$ називається кривою розподілу ймовірностей випадкової величини X (рис. 7).

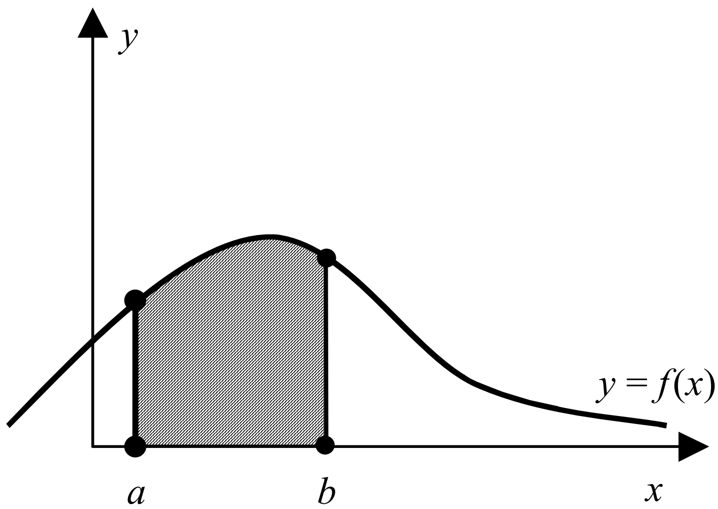


Рис. 7

Природно, що ймовірність того, що випадкова величина X прийме значення на інтервалі (a, b) дорівнює площі криволінійної трапеції, обмеженої кривою розподілу, віссю Ox і прямими $x = a$, $x = b$ (рис. 7), тобто

$$P(a < X < b) = \int_a^b f(x)dx. \tag{8}$$

Цей інтеграл називається інтегралом ймовірності. Згідно з означенням функції розподілу випадкової величини X: $F(x) = P(-\infty < X < x)$.

Отже, враховуючи (8) маємо:

$$F(x) = P(-\infty < X < x) = \int_{-\infty}^{x} f(x)dx\,, \tag{9}$$

тобто якщо відома функція розподілу $F(x)$ неперервної випадкової величини *X*, то $f(x) = F'(x)$ і, навпаки, якщо відома $f(x)$, то $F(x)$ обчислюється за формулою (9).

Властивості щільності ймовірностей.

1) Щільність розподілу ймовірностей неперервної випадкової величини X невід'ємна, тобто $f(x) \ge 0$ (оскільки $F(x)$ — монотонно зростає, то $F'(x) > 0$).

2) $\int\limits_{-\infty}^{\infty} f(x)dx = 1$.

Дійсно, інтеграл виражає ймовірність події, яка полягає в тому, що випадкова величина прийме значення, які належать інтервалу $(-\infty, \infty)$. Очевидно, така подія достовірна, отже, ймовірність її дорівнює 1. Геометрично це означає, що вся площа криволінійної трапеції, обмеженої віссю *Ox* і кривою розподілу, дорівнює одиниці.

Якщо всі значення випадкової величини X лежать на проміжку (a, b), то останню властивість можливо записати у вигляді $\int\limits_{a}^{b} f(x)dx = 1$.

Іноді функцію $f(x)$ *називають диференціальною функцією розподілу ймовірності, а функцію* $F(x)$ *— інтегральною функцією розподілу ймовірності. Зауважимо, що для описання розподілу ймовірностей дискретної величини, щільність розподілу незастосовна.*

Приклад 10. Випадкова величина X підпорядкована закону розподілу із щільністю $f(x)$, причому:

$$f(x) = \begin{cases} 0, & \textit{якщо} \quad x < 0 \\ k(2x - x^2), & \textit{якщо} \quad 0 \le x \le 2 \\ 0, & \textit{якщо} \quad x > 2 \end{cases} .$$

Необхідно: 1) знайти коефіцієнт k; 2) побудувати графік розподілу щільності $y = f(x)$; 3) знайти ймовірність попадання X у проміжок $(\frac{1}{2};\frac{3}{2})$.

1) Оскільки всі значення даної випадкової величини лежать на відрізку [0,2], то $\int\limits_0^2 k(2x - x^2)dx = 1$, звідси $k \cdot [x^2 - \frac{1}{3}x^3]_0^2 = 1$, або $k[4 - \frac{8}{3}] = 1$, тобто $k = \frac{3}{4}$.

2) Графіком функції $f(x)$ на інтервалі (0, 2) є парабола $y = \frac{3}{2}x - \frac{3}{4}x^2$, а поза цим інтервалом графіком є сама вісь абсцис (рис. 8).

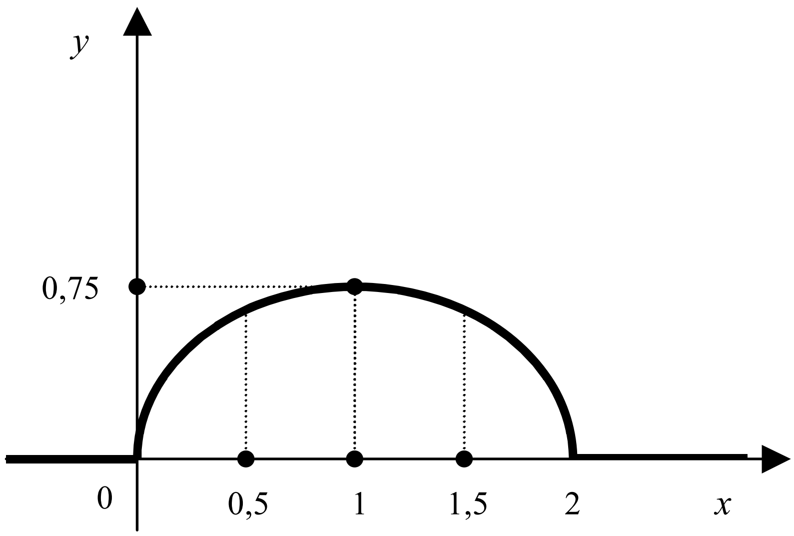


Рис. 8

3) Ймовірність попадання випадкової величини X на проміжок $(\frac{1}{2};\frac{3}{2})$ знаходиться із рівності:

$$P(\frac{1}{2} < X < \frac{3}{2}) = \int\limits_{\frac{1}{2}}^{\frac{3}{2}} (\frac{3}{2}x - \frac{3}{4}x^2)dx = [\frac{3}{4}x^2 - \frac{x^3}{4}]\Bigg|_{\frac{1}{2}}^{\frac{3}{2}} = \frac{27}{16} - \frac{27}{32} - \frac{3}{16} + \frac{1}{32} = \frac{11}{16}$$

.

Приклад 11. Нехай функція розподілу неперервної випадкової величини X:

$$F(x)=\begin{cases}0, \text{ якщо } x\le 0\\ \sin 2x, \text{ якщо } 0<x\le\dfrac{\pi}{4}\\ 1, \text{ якщо } x>\dfrac{\pi}{4}\end{cases}.$$

Знайти щільність розподілу $f(x)$.

Щільність розподілу дорівнює першій похідній від функції розподілу:

$$f(x)=F'(x)=\begin{cases}0, \text{ якщо } x<0\\ 2\cos 2x, \text{ якщо } 0<x<\dfrac{\pi}{4}\\ 0, \text{ якщо } x\ge\dfrac{\pi}{4}\end{cases}.$$

Зауважимо, що при $x = 0$ похідна $F'(x)$ не існує.

Приклад 12. Нехай щільність розподілу неперервної випадкової величини X:

$$f(x)=\begin{cases}0, \text{ якщо } x\le 0\\ \sin x, \text{ якщо } 0<x\le\dfrac{\pi}{2}\\ 0, \text{ якщо } x>\dfrac{\pi}{2}\end{cases}.$$

Знайти функцію розподілу *F(x)*.

Скористаємося формулою (9). Якщо $x\le 0$, то $f(x)=0$, отже, $F(x)=\int\limits_{-\infty}^{0} 0 \quad dx=0$.

Якщо $0<x\le\dfrac{\pi}{2}$, то

$$F(x)=\int\limits_{-\infty}^{0} 0 \quad dx+\int\limits_{0}^{x}\sin x \quad dx=(-\cos x)\Big|_0^x=1-\cos x.$$

Якщо $x > \frac{\pi}{2}$, то

$$F(x) = \int_{-\infty}^{0} 0 \quad dx + \int_{0}^{\frac{\pi}{2}} \sin x \quad dx + \int_{\frac{\pi}{2}}^{x} 0 \quad dx = (-\cos x)\Big|_0^{\frac{\pi}{2}} = 1.$$

Отже, шукана функція розподілу:

$$F(x) = \begin{cases} 0, \ \textit{якщо} \ \ x \le 0 \\ 1 - \cos x, \ \textit{якщо} \ \ 0 < x \le \dfrac{\pi}{2} \\ 1, \ \textit{якщо} \ \ x > \dfrac{\pi}{2} \end{cases}.$$

Приклад 13. Випадкова величина X задана функцією розподілу:

$$F(x) = \begin{cases} 0, \ \textit{якщо} \ \ x < 2 \\ (x-2)^2, \ \textit{якщо} \ \ 2 \le x \le 3 \\ 1, \ \textit{якщо} \ \ x > 3 \end{cases}.$$

Обчислити ймовірність попадання випадкової величини X в інтервали (1; 2,5) і (2,5; 3,5).

Ймовірність того, що X набуде значення із інтервалу (x_1, x_2), дорівнює приросту функції розподілу на цьому інтервалі: $P(x_1 < X < x_2) = F(x_2) - F(x_1)$.

Вважатимемо $x_1 = 1$, $x_2 = 2{,}5$, і $x_1 = 2{,}5$, $x_2 = 3{,}5$, одержимо

$$P_1 = F(2{,}5) - F(1) = (2{,}5 - 2)^2 - 0 = 0{,}25$$

$$P_2 = F(3{,}5) - F(2{,}5) = 1 - (2{,}5 - 2)^2 = 1 - 0{,}25 = 0{,}75 .$$

ЗАДАЧІ

1.22. Підручник виданий накладом 1000 екземплярів. Ймовірність того, що підручник переплетений неправильно рівна 0,01. Знайти ймовірність того, що тираж містить рівно три бракованих книжки.

Відповідь: $P_{1000}(3) = 0{,}007$.

1.23. Монета кинута 2 рази. Написати у вигляді таблиці закон розподілу випадкової величини X — числа випадання “герба”.

Відповідь:

X	2	1	0
P	0,25	0,5	0,25

1.24. У партії з N деталей M бракованих. Навмання беруть n деталей. Знайти ймовірність того, що з n деталей буде m бракованих.

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

1.25. Робітниця обслуговує 1000 веретен. Ймовірність обриву нитки на одному веретені протягом однієї хвилини дорівнює 0,004. Знайти ймовірність того, що протягом однієї хвилини обрив відбудеться на п’яти веретенах.

Відповідь: $P_{1000}(5)=0{,}1562$.

1.26. У партії із 10 деталей є 8 стандартних. Навмання відібрані 2 деталі. Скласти закон розподілу числа стандартних деталей серед відібраних.

Відповідь:

X	0	1	2
P	$\frac{1}{45}$	$\frac{16}{45}$	$\frac{28}{45}$

1.27. Випадкова величина X задана функцією розподілу (інтегральною функцією):

$$F(x)=\begin{cases}0, \text{ якщо } x<1\\ \dfrac{x-1}{2}, \text{ якщо } 1\le x\le 3\\ 1, \text{ якщо } x>3\end{cases} .$$

Обчисліть ймовірність попадання випадкової величини X в інтервали (1,5; 2,5) і (2.5; 3,5).

Відповідь: P_1=0,5; P_2=0,25.

1.28. Випадкова величина X має закон розподілу з щільністю:

$$f(x)=\begin{cases}\dfrac{a}{\sqrt{a^2-x^2}},\ \textit{якщо}\ \ |x|<a\\ 0,\ \textit{якщо}\ \ |x|\geq a\end{cases}.$$

Потрібно: 1) знайти коефіцієнт a; 2) знайти ймовірність попадання випадкової величини X на проміжок $(\frac{a}{2};a)$; 3) побудувати графік розподілу щільності ймовірності.

Відповідь: 1) $a=\frac{1}{\pi}$; 2) $P(\frac{a}{2}<X<a)=\frac{1}{3}$.

1.29. Випадкова величина задана щільністю розподілу:

$$f(x)=\begin{cases}0,\ \textit{якщо}\ \ x\leq 0\\ \dfrac{1}{2}\sin x,\ \textit{якщо}\ \ 0<x\leq\pi\\ 0,\ \textit{якщо}\ \ x>\pi\end{cases}.$$

Знайти: а) функцію розподілу; б) ймовірність того, що в результаті випробування випадкова величина набуде значення на інтервалі $(0;\ \frac{\pi}{4})$.

Відповідь:

$$F(x)=\begin{cases}0,\ \textit{якщо}\ \ x\leq 0\\ \dfrac{1}{2}(1-\cos x),\ \textit{якщо}\ \ 0<x\leq\pi\\ 1,\ \textit{якщо}\ \ x>\pi\end{cases}.$$

1.30. Встановити, що функція $f(x)=\dfrac{1}{x^2+\pi^2}$ є щільністю ймовірності деякої випадкової величини X та обчислити ймовірність попадання випадкової величини X на інтервал $(\pi\ ;\ \infty)$.

Відповідь: $P(\pi<X<\infty)=\frac{1}{4}$.

§ 5. Числові характеристики випадкових величин

Математичне сподівання дискретної випадкової величини. У багатьох питаннях повна характеристика випадкової величини не вимагається.

Випадкову величину залежно від вимог тієї чи іншої конкретної задачі можливо характеризувати деякими сталими числовими характеристиками. Однією з таких характеристик є *математичне сподівання дискретної випадкової величини* — суми добутків усіх її можливих значень на їх ймовірність.

Якщо випадкова величина X характеризується скінченим рядом розподілу:

x_i	x_1	x_2	x_3	...	x_n
p_i	p_1	p_2	p_3	...	p_n

то математичне сподівання $M(X)$ визначається за формулою:

$$M(X)=x_1p_1+x_2p_2+...+x_np_n\text{, або } M(X)=\sum_{i=1}^{n}x_ip_i\ . \qquad (1)$$

Оскільки $p_1+p_2+p_3+...+p_n=1$, то

$$M(X)=\frac{x_1p_1+x_2p_2+...+x_np_n}{p_1+p_2+p_3+...+p_n}\ .$$

Таким чином, $M(X)$ є зваженим середнім арифметичним значенням випадкової величини $x_1, x_2, ..., x_n$ за вагою $p_1, p_2, ..., p_n$.

Якщо $n=\infty$, то $M(X)=\sum\limits_{i=1}^{\infty}x_ip_i$ (за умови, що сума цього ряду скінчена).

Приклад 1. Стрілок вибиває 4, 8, 9, 10 очок з ймовірностями 0,1, 0,45, 0,3, 0,15. Знайти математичне сподівання числа очок X при одному пострілі.

$$M(X)=4\cdot 0{,}1+8\cdot 0{,}45+9\cdot 0{,}3+10\cdot 0{,}15=8{,}2\ .$$

Це середнє сподівання значень чисел вибитих очок при одному пострілі. Із (1) видно, що $M(X)$ — стале число.

Очевидно також, що у множині значень випадкової величини X завжди знайдуться значення більші або менші математичного сподівання, а тому математичне сподівання ще називають *центром розподілу випадкової величини.*

Приклад 2. Знайти математичне сподівання числа появи події A в одному випробуванні, якщо ймовірність події A дорівнює p.

Випадкова величина X — число появи події A в одному випробуванні — може приймати лише два значення: $x_1 = 1$ (подія A відбулася) з ймовірністю p і $x_2 = 0$ (подія A не відбулася) з ймовірністю $q = 1 - p$. Звідси математичне сподівання:

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

Отже, математичне сподівання числа появи події в одному випробуванні дорівнює ймовірності цієї події.

Легко переконатися, що у випадку n незалежних випробувань в кожному із яких ймовірність події A стала і дорівнює p математичне сподівання дорівнює:

$$M(X) = np .$$

Крім того, якщо в n випробуваннях випадкова величина X приймала m_1 раз значення x_1, m_2 раз значення x_2, ………….. , m_k раз значення x_k, причому $m_1 + m_2 +\ldots+ m_k = n$, тоді сума всіх значень, прийнятих X дорівнює:

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

Звідси середнє арифметичне $\overline{X}$ усіх значень, прийнятих випадковою величиною:

$$\overline{X} = \frac{x_1 m_1 + x_2 m_2 + \ldots + x_k m_k}{n} ;$$

$$\text{або } \overline{X} = x_1 W_1 + x_2 W_2 + \ldots + x_k W_k ;$$

де $W_i = \dfrac{m_i}{n} ; (i = 1, 2, \ldots, k)$ — відносна частота.

Якщо число випробувань досить велике, тоді відносна частота наближено дорівнює ймовірності появи події, тобто $W_i \approx p_i$, а тому

$$\overline{X} \approx x_1 p_1 + x_2 p_2 + \ldots + x_k p_k \text{ або } \overline{X} \approx M(X) .$$

Отже, *математичне сподівання наближено дорівнює (тим точніше, чим більше випробувань) середнє арифметичному спостережних значень випадкової величини.*

Властивості математичного сподівання.

1. Математичне сподівання сталої величини дорівнює тій же сталій:
$$M(C)=C\text{, де } C=const. \tag{2}$$

2. Математичне сподівання суми випадкових величин дорівнює сумі їх математичних сподівань:
$$M(X_1+X_2)=M(X_1)+M(X_2). \tag{3}$$

3. Сталий множник можливо виносити за знак математичного сподівання:
$$M(CX)=CM(X)\text{, де } C=const. \tag{4}$$

4. Математичне сподівання добутку незалежних випадкових величин X_1 і X_2 дорівнює добутку їх математичних сподівань:
$$M(X_1\cdot X_2)=M(X_1)\cdot M(X_2). \tag{5}$$

Формули (3), (5) дійсні для будь-якого скінченого числа незалежних випадкових величин.

5. Математичне сподівання різниці випадкових величин дорівнює різниці їх математичних сподівань
$$M(X_1-X_2)=M(X_1)-M(X_2).$$

6. Різниця $X – M(X)$ називається *відхиленням* випадкової величини X від її математичного сподівання. Математичне сподівання відхилення випадкової величини дорівнює нулю:
$$M[X-M(X)]=0 \tag{6}$$
дійсно $M[X-M(X)]=M(X)-M[M(X)]=M(X)-M(X)=0$, оскільки $M[M(X)]=M(X)$.

Дисперсія дискретної випадкової величини. Друга характеристика випадкової величини X — *дисперсія (“розсіювання”)* — дає можливість оцінити відстань (“розсіювання”) значень випадкової величини навколо її математичного сподівання.

Приклад 3. Нехай випадкові величини X_1 і X_2 мають такі закони розподілу:

X_1	–0,1	0	0,1	0,4
P	0,3	0,15	0,3	0,25

X_2	–10	0,5	10
P	0,4	0,2	0,4

Математичні сподівання:

$$M(X_1) = -0{,}1 \cdot 0{,}3 + 0 \cdot 0{,}15 + 0{,}1 \cdot 0{,}3 + 0{,}4 \cdot 0{,}25 = 0{,}1;$$
$$M(X_2) = -10 \cdot 0{,}4 + 0{,}5 \cdot 0{,}2 + 10 \cdot 0{,}4 = 0{,}1$$

У цьому прикладі математичні сподівання X_1 і X_2 однакові, а число можливих значень і самі значення різні, причому значення X_1 групуються поблизу її математичного сподівання, а значення X_2 — ні. Отже X_1 менше розсіюється по відношенню до свого математичного сподівання, ніж X_2.

Дисперсією випадкової величини називається математичне сподівання квадрата відхилення випадкової величини від її математичного сподівання.

$$D(X) = M[X - M(X)]^2. \qquad (7)$$

Середнім квадратичним відхиленням випадкової величини X від її математичного сподівання є квадратний корінь із дисперсії:

$$\sigma(x) = \sqrt{D(x)}\,. \qquad (8)$$

Середнє квадратичне відхилення також як і дисперсія, є мірою розсіювання випадкової величини відносно свого математичного сподівання.

Розмірності: D — розмірність квадрату випадкової величини X, σ — розмірність випадкової величини.

Дисперсію можливо обчислювати і іншим способом. Із (7) маємо:

$$D(X) = M[X - M(X)]^2 = M[X^2 - 2XM(X) + M^2(X)] =$$
$$= M(X^2) - 2M(X)M(X) + M^2(X) = M(X^2) - M^2(X)\,.$$

Отже,

$$D(X) = M(X^2) - M^2(X)\,.$$

Приклад 4. Знайти дисперсію випадкової величини X_1, задану таблицею із попереднього прикладу. Раніше обчислено, що $M(X_1) = 0{,}1$, отже,

$$M(X_1^2) = (-0{,}1)^2 \cdot 0{,}3 + 0^2 \cdot 0{,}15 + (0{,}1)^2 \cdot 0{,}3 + (0{,}4)^2 \cdot 0{,}25 = 0{,}046$$
$$D(X_1) = 0{,}046 - (0{,}1)^2 = 0{,}036\,.$$

На рис. 9 показано дискретне розсіювання з малим (*б*) і великим (*а*) розсіюванням.

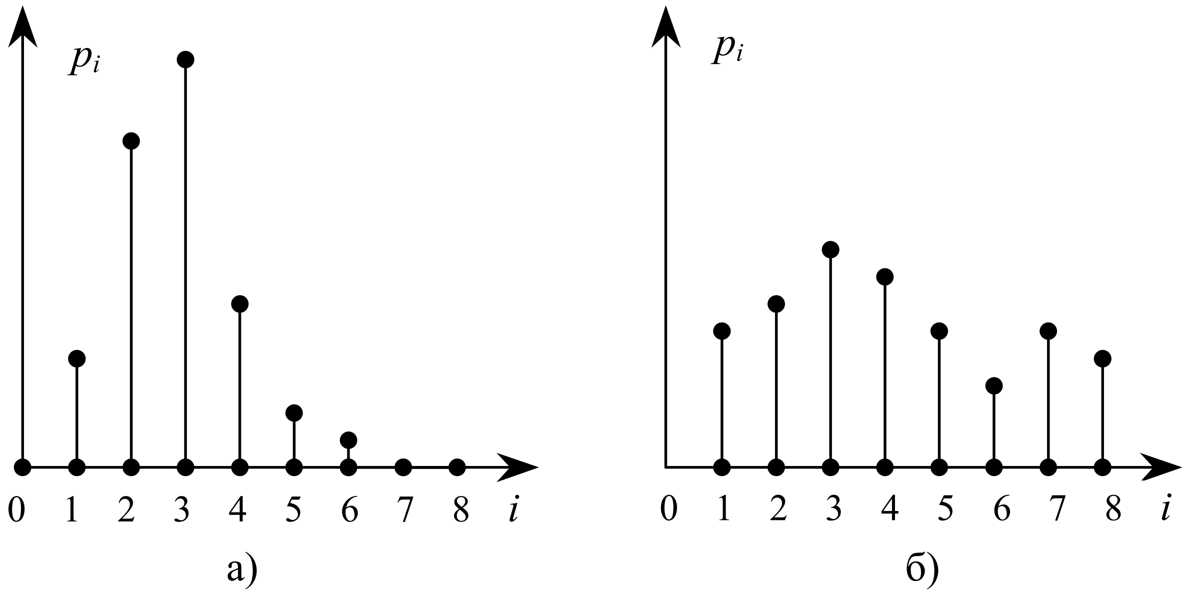


Рис. 9

Властивості дисперсії:

1. Дисперсія сталої величини дорівнює нулю:

$$D(C) = 0\,. \tag{9}$$

2. Дисперсія добутку сталої величини на випадкову величину дорівнює добутку квадрата сталої величини на дисперсію випадкової величини.

$$D(C \cdot X) = C^2 \cdot D(X)\,. \tag{10}$$

3. Дисперсія суми сталої C і випадкової величини дорівнює дисперсії випадкової величини:

$$D(C + X) = D(X)\,. \tag{11}$$

4. Дисперсія суми двох незалежних випадкових величин дорівнює сумі дисперсій цих величин:

$$D(X_1 + X_2) = D(X_1) + D(X_2)\,. \tag{12}$$

Ця властивість справедлива для будь-якого скінченого числа незалежних випадкових величин.

5. Дисперсія різниці двох незалежних випадкових величин дорівнює сумі їх дисперсій:

$$D(X - Y) = D(X) + D(Y)\,. \tag{13}$$

Дійсно,

$$D(X-Y)=D(X)+D(-Y)=D(X)+(-1)^2 D(Y)=D(X)+D(Y).$$

Зауваження. Здається, якщо X і Y мають однакові можливі значення і одне і теж математичне сподівання то і дисперсії цих величин рівні, оскільки можливі значення обох величин однаково розсіяні навколо своїх математичних сподівань. Проте в загальному випадку це не так, оскільки величина дисперсії визначається не тільки можливими значеннями, але і їх ймовірностями, які взагалі є різними. Наведемо ілюструючий приклад.

Приклад 5. Порівняти дисперсії випадкових величин заданих законами розподілу.

X	–1	1	2	3
P	0,48	0,01	0,09	0,42

Y	–1	1	2	3
P	0,19	0,51	0,25	0,05

Переконайтеся, що $M(X)=M(Y)=0{,}97$, $D(X)\approx 3{,}69$, $D(Y)\approx 1{,}21$.

Таким чином, можливі значення і математичні сподівання X і Y однакові, а дисперсії різні, причому $D(X)>D(Y)$.

Легко переконатися, що дисперсія настання події A в n незалежних випробуваннях, в кожному із яких ймовірність р настання події стала, дорівнює добутку числа n на ймовірності настання і ненастання події в одному випробуванні.

$$D(X)=npq.$$

Оскільки загальне число X появи події A в n випробуваннях дорівнює сумі настання події A в окремих випробуваннях:

$$X=X_1+X_2+\ldots+X_n$$

і $X_i (i=1,2,\ldots n)$ взаємно незалежні, то

$$D(X)=D(X_1)+D(X_2)+\ldots+D(X_n), \qquad (*)$$

де $D(X_i)=M(X_i^{\,2})-\left[M(X_i)\right]^2$, причому $M(X_i)=p$.

Математичне сподівання величини $X_1^{\,2}$, яка може приймати лише два значення, а саме: 1^2 із ймовірністю p і 0^2 із ймовірністю q дорівнює:

$$M(X_1^2)=1^2\cdot p+0^2\cdot q=p.$$

Тоді $D(X_1) = M(X_1^2) - \left[M(X_1)\right]^2 = p - p^2 = p(1-p) = pq$.

Тим самим дисперсія решти випадкових величин також рівна *pq*. Замінимо кожний доданок правої частини (*) через *pq*, одержимо $D(X) = npq$.

Справедливі наступні твердження:

1. Математичне сподівання середнє арифметичного однаково розподілених взаємно незалежних випадкових величин дорівнює математичному сподіванню a кожної із величин:

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

Дійсно:

$$M(\overline{X}) = M\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{M(X_1) + M(X_2) + \ldots + M(X_n)}{n} = \frac{na}{n} = a$$

.

2. Дисперсія середнє арифметичного n однаково розподілених взаємно незалежних випадкових величин у n раз менше дисперсії D кожної із величин:

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

Дійсно:

$$D(\overline{X}) = D\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{D(X_1) + D(X_2) + \ldots + D(X_n)}{n^2} = \frac{nD}{n^2} = \frac{D}{n}.$$

3. Середнє квадратичне відхилення середнє арифметичного n однаково розподілених взаємно незалежних випадкових величин у $\sqrt{n}$ *раз менше середнє квадратичного відхилення* σ *кожної із величин:*

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

Дійсно:

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

Висновок із (**) і (***): середнє арифметичне достатньо великої кількості взаємно незалежних випадкових величин має значно менше розсіювання, ніж кожна окрема величина.

Математичне сподівання і дисперсія неперервної випадкової величини. Згадаємо, що математичне сподівання дискретної випадкової величини X: $M(X)=\sum_{i=1}^{n} x_i p_i$. Нехай неперервна випадкова величина X приймає можливі значення на інтервалі (a, b) (рис.10) і відома щільність розподілу ймовірностей $f(x)$ — неперервна функція на цьому відрізку. Тоді *математичним сподіванням неперервної випадкової величини, можливі значення якої знаходяться на інтервалі (a, b) називається:*

$$M(X)=\int_a^b xf(x)dx . \tag{14}$$

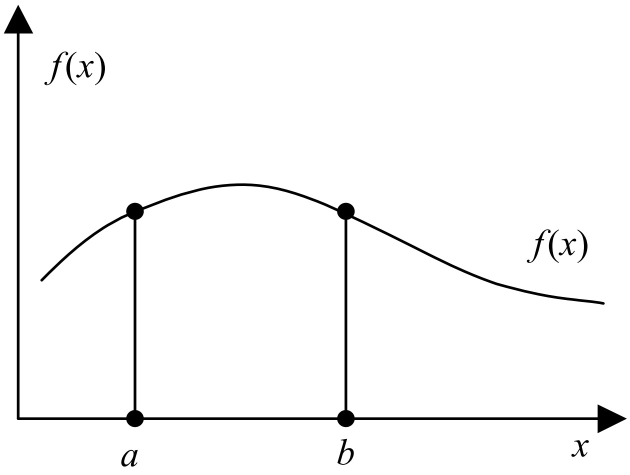


Рис. 10

Якщо можливі значення X знаходяться в інтервалі $(-\infty;\infty)$, то

$$M(X)=\int_{-\infty}^{\infty} xf(x)dx . \tag{15}$$

Дисперсією неперервної випадкової величини X називається математичне сподівання квадрата відхилення випадкової величини від її математичного сподівання (аналогічно визначається дисперсія дискретної випадкової величини)

$$D(X)=M(X-M(X))^2=\int_{-\infty}^{\infty}(x-M(X))^2 f(x)dx . \tag{16}$$

Якщо значення неперервної випадкової величини $X\in(a,b)$, то

$$D(X)=\int_a^b (x-M(X))^2 f(x)dx \tag{17}$$

Всі властивості математичного сподівання і дисперсії дискретних випадкових величин справедливі і для неперервних випадкових величин. Середнім квадратичним відхиленням неперервної випадкової величини X від її математичного сподівання є:

$$\sigma(X)=\sqrt{D(X)}\,. \tag{18}$$

Для дисперсії випадкової величини частіше буває прийнятна формула:

$$D(X)=M[(X-c)^2]-[M(X)-c]^2, \tag{19}$$

де c — довільне число.

Ця формула випливає із властивості лінійності математичного сподівання:

$$M[(X-c)^2]=M\left\{[(X-M(X))+(M(X)-c)]^2\right\}=M[(X-M(X))^2]+$$
$$(M(X)-c)^2+2(M(X)-c)M[X–M(X)]=M[(X–M(X))^2]+(M(X)–c)^2.$$

Оскільки $M[X-M(X)]=0$. Якщо c = 0, маємо $D(X)=M[X^2]-$ $-M^2[X]$.

Приклад 6. Випадкова величина X характеризується рядом розподілу:

x_i	0	1	2	3	4
p_i	0,3	0,5	0,2	0,09	0,03

Визначити математичне сподівання і дисперсію.

За формулою (1) знаходимо математичне сподівання:

$$M(X)=0\cdot 0,3+1\cdot 0,5+2\cdot 0,2+3\cdot 0,09+4\cdot 0,03=1,29\,.$$

Дисперсію знайдемо за формулою (19), вважаючи c = 2; звідси $M(X)$ – 2 = –0,71. Складемо таблицю:

x_i	0	1	2	3	4
x_i-c	–2	–1	0	1	2
$(x_i-c)^2$	4	1	0	1	4
p_i	0,3	0,5	0,2	0,09	0,03
$p_i(x_i-c)^2$	1,2	0,5	0	0,09	0,12

Звідси

$$M\left[(X-c)^2\right]=\sum_{i=0}^{4}p_i(x_i-c)^2=1,91;$$

$$D(X)=1,91-(-0,71)^2=1,91-0,50=1,41;$$

$$\sigma(X)=\sqrt{1,41}=1,19.$$

Зауваження. Легко переконатися, що для обчислення дисперсії неперервної випадкової величини зручно користуватися наступними формулами:

$$D(X)=\int_a^b x^2 f(x)\cdot dx-\left[M(X)\right]^2; \tag{20}$$

$$D(X)=\int_{-\infty}^{\infty} x^2 f(x)\cdot dx-\left[M(X)\right]^2. \tag{21}$$

Вказівка. Скористатися формулою (16) та рівностями

$\int_{-\infty}^{\infty} x\cdot f(x)\cdot dx=M(X)$ і $\int_{-\infty}^{\infty} f(x)\cdot dx=1$.

Приклад 7. Задана функція:

$$f(x)=\begin{cases}0, & \text{якщо} \quad x<0\\ \frac{1}{2}\sin x & \text{якщо} \quad 0\le x\le\pi\\ 0 & \text{якщо} \quad x>\pi\end{cases}.$$

Встановити, що $f(x)$ може бути щільністю ймовірності деякої випадкової величини *X*. Знайти математичне сподівання та дисперсію випадкової величини *X*.

Маємо:

$$\int_{-\infty}^{\infty} f(x)dx=\int_0^{\pi} f(x)dx=\frac{1}{2}\int_0^{\pi}\sin x dx=-\frac{1}{2}\cos x\Big|_0^{\pi}=1.$$

Оскільки функція $f(x) \geq 0$, то $f(x)$ може бути щільністю ймовірності деякої випадкової величини. Крім того, пряма $x = \pi/2$ є віссю симетрії відповідної дуги кривої $y = \frac{1}{2}\sin x$ (рис.11),

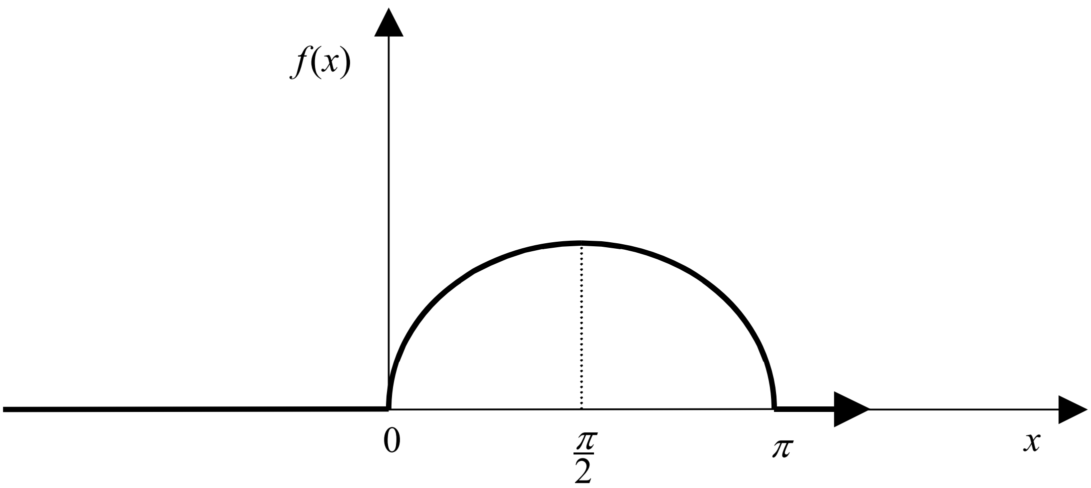


Рис. 11

а тому математичне сподівання випадкової величини X дорівнює $\pi/2$, тобто $M(X) = \frac{\pi}{2}$. Знайдемо дисперсію. Для цього у формулі (19) покладемо $c = 0$, $M(X) = \frac{\pi}{2}$, тоді залишається обчислити інтеграл, що визначає $M(X^2)$, тобто

$$M(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \frac{1}{2}\int_{0}^{\pi} x^2 \sin x dx = \frac{1}{2}(\pi^2 - 4),$$

отже,

$$D(X) = \frac{1}{2}(\pi^2 - 4) - (\frac{\pi}{2})^2 = \frac{\pi^2}{4} - 2,$$

$$\sigma = \sqrt{\frac{\pi^2}{4} - 2} \approx 0,69.$$

Мода і медіана. Якщо математичне сподівання $M(X)$ існує і крива розподілу симетрична відносно прямої $x = C$, то $M(X) = C$. Геометрично математичне сподівання як неперервної так і дискретної випадкової величини дорівнює абсцисі центру мас площі, обмеженої кривою розподілу і віссю абсцис. А тому при симетрії кривої розподілу відносно деякої прямої, паралельної осі ординат, математичне сподівання співпадає з абсцисою точки перетину цієї осі симетрії з віссю абсцис.

Модою дискретної випадкової величини X називається її найбільш ймовірностне значення.

Модою неперервної випадкової величини X називається те її значення, за якого щільність розподілу максимальна.

Моду будемо позначати символом $M_0(X)$.

Медіаною неперервної випадкової величини X називається таке її значення M_l, *для якого однаково ймовірно виявиться випадкова величина менше чи більше* $M_l(X)$, *тобто:*

$$P\left[X < M_l(X)\right] = P\left[X > M_l(X)\right] = \frac{1}{2}.$$

Геометрично мода є абсцисою тієї точки кривої розподілу, ордината якої максимальна. Ордината проведена в точці з абсцисою $x = M_l(X)$, ділить наполовину площу, обмежену кривою розподілу. Якщо пряма $x = a$ є віссю симетрії кривої розподілу $y = f(x)$, то

$$M_0 = M_l = M(X) = a.$$

Приклад 8. Задана щільність розподілу неперервної випадкової величини

$$f(x) = \begin{cases} 0, & \text{якщо} \quad x < 2 \\ a(x-2)(4-x), & \text{якщо} \quad 2 \le x \le 4. \\ 0, & \text{якщо} \quad x > 4 \end{cases}$$

Знайти a, моду, медіану і математичне сподівання величини X.

Оскільки всі значення даної випадкової величини лежать на відрізку $[2;4]$, то $\int_2^4 a(x-2)(4-x)\cdot dx = 1$, звідси: $-a\left(\frac{x^3}{3} - 3x^2 + 8x\right)\Big|_2^4 = 1$,

або $\frac{4}{3}a = 1$, тобто $a = \frac{3}{4}$.

Представимо щільність розподілу у вигляді:

$$f(x) = \frac{3}{4}(x-2)(4-x) = -\frac{3}{4}\left[(x-3)^2 - 1\right].$$

Оскільки при $x = 3$ щільність розподілу досягає максимуму, то $M_0(X) = 3$. (Зрозуміло, що дослідження на екстремум можна було б виконати методами диференціального числення).

Крива розподілу симетрична відносно прямої $x = 3$, а тому $M(X) = 3$ і $M_l(X) = 3$.

Приклад 9. Випадкова величина задана щільністю розподілу $f(x) = 2\cos 2x$ в інтервалі $(0; \frac{\pi}{4})$; за межами цього інтервалу $f(x) = 0$. Знайти: а) моду; б) медіану *X*.

а) Легко переконатися, що функція $f(x) = 2\cos 2x$ у відкритому інтервалі $(0; \frac{\pi}{4})$ немає максимуму, а тому X моди немає.

б) Знайдемо медіану $M_l(X) = m_l$, виходячи із визначення медіани $P(x < m_l) = P(x > m_l)$. Це теж саме, що і $P(x < m_l) = \frac{1}{2}$. Ураховуючи, що за умовою можливі значення X додатні запишемо цю рівність так:

$$P(0 < X < m_l) = \frac{1}{2}, \text{ або } 2\int_0^{m_l} \cos 2x \cdot dx = \sin 2m_l = \frac{1}{2}.$$

Звідси $2m_l = \arcsin\frac{1}{2} = \frac{\pi}{6}$; $m_l = \frac{\pi}{12}$.

ЗАДАЧІ

1.31. Знайти дисперсію і середнє квадратичне відхилення дискретної випадкової величини X, заданої законом розподілу:

X	–5	2	3	4
P	0,4	0,3	0,1	0,2

Відповідь: $D(X) = 15,21$; $\sigma(X) = 3,9$.

1.32. Знайти дисперсію дискретної випадкової величини X — числа появи події A в п’яти незалежних випробуваннях, якщо ймовірність настання події A в кожному випробуванні дорівнює 0,2.

Відповідь: $D(X) = 0{,}8$.

1.33. Дискретні незалежні випадкові величини, задані законами розподілу:

X_1	1	2
P	0,2	0,8

X_2	0,5	1
P	0,3	0,7

Знайти математичне сподівання добутку $X_1 \cdot X_2$ двома способами: а) скласти закон розподілу $X_1 \cdot X_2$; б) користуючись властивістю $M(X_1 \cdot X_2) = M(X_1) \cdot M(X_2)$.

Відповідь: 1,53.

1.34. Дискретні випадкові величини X_1 та X_2, задані законами розподілу, вказаними в задачі 1.33. Знайти математичне сподівання суми $X_1 + X_2$ двома способами: а) скласти закон розподілу $X_1 + X_2$; б) користуючись властивістю $M(X_1 + X_2) = M(X_1) + M(X_2)$.

Відповідь: 2,65.

1.35. Зайти дисперсію випадкової величини X, заданою функцією розподілу:

$$F(x) = \begin{cases} 0, & \text{якщо} \quad x \le -2 \\ \dfrac{x}{4} + \dfrac{1}{2}, & \text{якщо} \quad -2 < x \le 2 \\ 1, & \text{якщо} \quad x > 2 \end{cases}.$$

Відповідь: $D(X) = \dfrac{4}{3}$.

1.36. Випадкова величина X в інтервалі (3;5) задана щільністю розподілу $f(x) = -\dfrac{3}{4}x^2 + 6x - \dfrac{45}{4}$; за межами цього інтервалу $f(x) = 0$. Знайти моду, математичне сподівання і медіану X.

Відповідь: $M_0(X) = M(X) = M_l(X) = 4$.

1.37. Випадкова величина X задана щільністю розподілу $f(x) = c(x^2 + 2x)$ в інтервалі (0;1); за межами цього інтервалу $f(x) = 0$. Знайти: а) параметр с б) математичне сподівання величини *X*.

Відповідь: а) $c = \frac{3}{4}$; б) $M(X) = \frac{11}{16}$.

§ 6. Деякі закони розподілу неперервної випадкової величини

Рівномірний розподіл. Розподіл ймовірностей називається *рівномірним*, якщо на інтервалі, якому належать усі можливі значення випадкової величини, щільність розподілу зберігає стале значення.

Знайдемо щільність ймовірності рівномірного розподілу $f(x)$, враховуючи, що всі можливі значення випадкової величини містяться на інтервалі ($a;b$), на якому функція $f(x)$ зберігає стале значення *C*.

Згідно з умовою $f(x) = 0$ при $x < a$ і $x > b$. Знайдемо сталу *C*. Оскільки всі можливі значення випадкової величини належать інтервалу ($a;b$), то повинно виконуватися співвідношення:

$$\int_a^b f(x)dx = 1 \quad \text{або} \quad \int_a^b Cdx = 1, \tag{1}$$

звідси $C = \frac{1}{b-a}$.

Отже, щільність ймовірності рівномірного розподілу:

$$f(x) = \begin{cases} 0, & \textit{якщо} \quad x < a \\ \frac{1}{b-a}, & \textit{якщо} \quad a \le x \le b \\ 0, & \textit{якщо} \quad x > b \end{cases}.$$

Графік функції зображено на рис. 12.

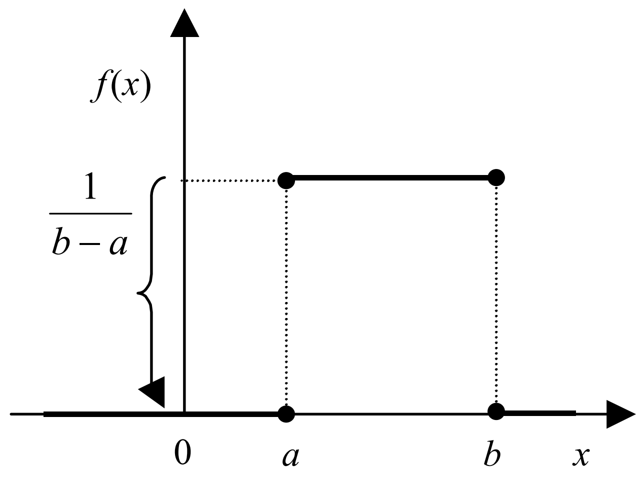


Рис. 12

Знайдемо функцію розподілу $F(x)$, математичне сподівання $M(X)$ та дисперсію $D(X)$ випадкової величини, рівномірно розподіленої на інтервалі ($a;b$):

$$F(x)=\int_a^x f(x)dx=\int_a^x \frac{dx}{b-a}=\frac{x-a}{b-a},$$

отже:

$$F(x)=\begin{cases} 0, & \text{якщо} \quad x<a \\ \dfrac{x-a}{b-a}, & \text{якщо} \quad a\le x\le b \\ 1, & \text{якщо} \quad x>b \end{cases}.$$

На рис.13 побудований графік функції розподілу $F(x)$.

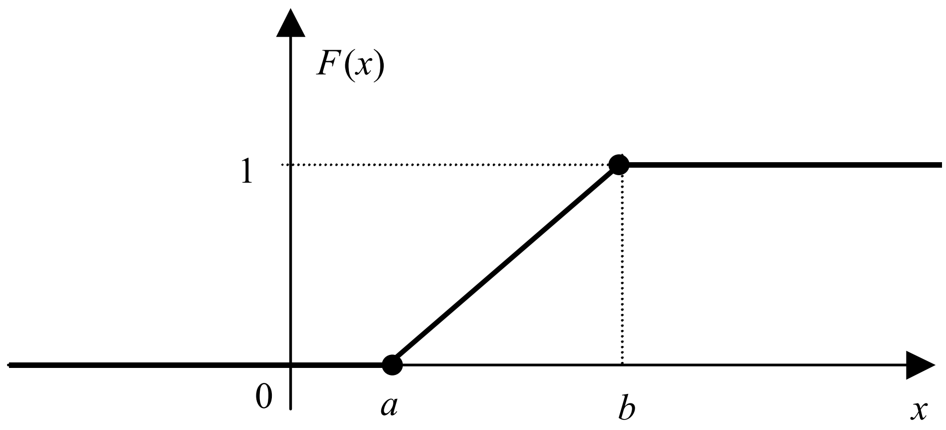


Рис. 13

$$M(X)=\int_a^b x\frac{dx}{b-a}=\frac{1}{b-a}\int_a^b xdx=\frac{b^2-a^2}{2(b-a)}=\frac{b+a}{2}. \qquad (*)$$

Для знаходження дисперсії випадкової величини X, рівномірно розподіленої на інтервалі (a; b), скористаємося формулою:

$$D(X)=M(X^2)-M^2(X).$$

Знайдемо

$$M(X^2)=\int_a^b x^2 f(x)dx=\int_a^b x^2 dF(x)=\int_a^b \frac{x^2}{b-a}dx=\frac{b^3-a^3}{3(b-a)}=\frac{b^2+ab+a^2}{3}.$$

Оскільки $M(X)=\dfrac{a+b}{2}$, то

$$D(X)=\frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2=\frac{(a-b)^2}{12}. \qquad (**)$$

$$\sigma(X)=\frac{b-a}{2\sqrt{3}} \qquad (***)$$

Середнє квадратичне відхилення випадкової величини X дорівнює квадратному кореню із дисперсії.

Приклад 1. Ціна поділу шкали амперметра дорівнює 0,2А. Показання закруглюють до найближчого цілого поділу. Знайти ймовірність того, що при відліку буде зроблена помилка, яка перевищує 0,02А.

Помилку заокруглення підрахунку можливо розглядати як випадкову величину X, що розподілена рівномірно в інтервалі між двома сусідніми цілими поділками. Щільність рівномірного розподілу $f(x)=\dfrac{1}{b-a}$, де $(b-a)$ — довжина інтервалу, в якому знаходяться можливі значення X; ззовні цього інтервалу $f(x)=0$. У заданій задачі довжина інтервалу, в якому знаходяться можливі значення X, дорівнює 0,2, а тому $f(x)=\dfrac{1}{0,2}=5$. Не важко переконатися, що помилка відліку перевищить 0,02, якщо вона буде в межах інтервалу (0,02; 0,18) .

За формулою $P(a<x<b)=\int\limits_a^b f(x)\cdot dx$ одержимо

$$P(0,02<X<0,18)=\int\limits_{0,02}^{0,18}5dx=5(0,18-0,02)=5\cdot 0,16=0,8.$$

Приклад 2. Знайти математичне сподівання, дисперсію та середнє квадратичне відхилення випадкової величини X, розподіленої рівномірно в інтервалі (3;9).

За формулами (*) — (***) математичне сподівання, дисперсія та середнє квадратичне відхилення відповідно дорівнюють:

$$M(X)=6;\ D(X)=3;\ \sigma(X)=\sqrt{3}.$$

Експоненціальний розподіл. *Експоненціальний (показниковий) розподіл* — це розподіл ймовірностей неперервної випадкової величини X з параметром $\lambda > 0$, заданий законом:

$$f(x) = \begin{cases} 0, & \text{якщо} \quad x < 0 \\ \lambda e^{-\lambda x}, & \text{якщо} \quad x \ge 0 \end{cases}. \tag{2}$$

Графік щільності експоненціального розподілу показаний на рис. 14.

Функція розподілу має вигляд:

$$F(x) = \int_{-\infty}^{x} f(z)dz = \int_{0}^{x} f(z)dz = \int_{0}^{x} \lambda e^{-\lambda z} dz = 1 - e^{-\lambda x} .$$

Графік $F(x)$ показаний на рис. 15.

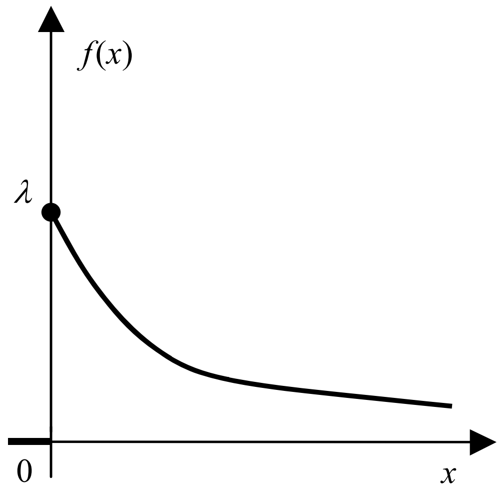


Рис. 14

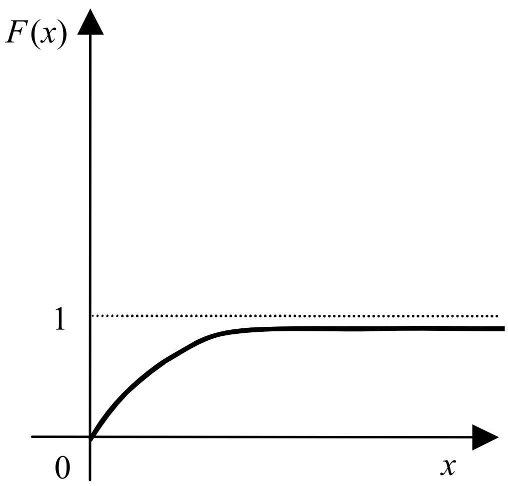


Рис. 15

Математичне сподівання:

$$M(X) = \int_{-\infty}^{\infty} x\lambda e^{-\lambda x} dx = \int_{0}^{x} x\lambda e^{-\lambda x} dx = \left| \begin{matrix} u = x; \ dv = e^{-\lambda x} dx \\ du = dx; \ v = -\dfrac{1}{\lambda} e^{-\lambda x} \end{matrix} \right| =$$

$$= -xe^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda}.$$

Отже: $M(X) = \dfrac{1}{\lambda}$.

Дисперсія:

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

Враховано, що $f(x) = 0$, якщо $x < 0$; $M(X)=\frac{1}{\lambda}$. Інтегрування проводиться за частинами (див. попередній випадок), але два рази.

Знайдемо середнє квадратичне відхилення, для чого слід добути квадратний корінь із дисперсії $D(X)=\frac{1}{\lambda^2}$, тобто $M(X)=\sigma(X)=\frac{1}{\lambda}$.

Ймовірність попадання в інтервал $(a;b)$ неперервної випадкової величини, яка розподілена за показниковим законом $F(x)=1-e^{-\lambda x}$ ($x\geq 0$) дорівнює:

$$P(a<X<b)=\int_a^b f(x)dx=\lambda\int_a^b e^{-\lambda x}dx=e^{-\lambda a}-e^{-\lambda b};$$

$$P(a<x<b)=e^{-\lambda a}-e^{-\lambda b}.$$

Значення функції e^{-x} потрібно знаходити за таблицями.

Експоненціальний розподіл використовується в теорії надійності. Функція надійності $R(t)$ визначає ймовірність безвідмовної роботи елемента за час t:

$$R(t)=e^{-\lambda t},$$

де λ — інтенсивність відмови.

Наприклад. Тривалість служби електролампи можна розглядати з хорошим наближенням, як експоненціальну розподілену величину.

Приклад 3. Неперервна випадкова величина X розподілена за показниковим законом із щільністю ймовірності $f(x)=2e^{-2x}$, якщо $x\geq 0$; $f(x)=0$, якщо $x<0$. Знайти ймовірність того, що в результаті випробування X попаде в інтервал (0,15; 0,4).

Скористаємося формулою:

$$P(a<X<b)=e^{-\lambda a}-e^{-\lambda b}.$$

Ураховуючи, що за умовою $a = 0{,}15$; $b = 0{,}4$; $\lambda = 2$ і скориставшись таблицею значень функції e^{-x}, одержимо:

$$P(0{,}15 < X < 0{,}4) = e^{-2\cdot 0{,}15} - e^{-2\cdot 0{,}4} = e^{-0{,}3} - e^{-0{,}8} = \frac{1}{1{,}3499} - \frac{1}{2{,}2255} = 0{,}292$$

Приклад 4. Студент пам'ятає, що щільність показникового розподілу має вигляд $f(x) = 0$, якщо $x<0$; $f(x) = Ce^{-\lambda x}$, якщо $x \ge 0$, але він забув чому дорівнює стала *C*. Потрібно знайти *C*.

Скориставшись властивістю щільності розподілу:

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

обчислимо:

$$\int_{0}^{\infty} Ce^{-\lambda x}dx = 1;\ -\frac{C}{\lambda}e^{-\lambda x}\Big|_{0}^{\infty} = 1;\ \frac{C}{\lambda} = 1;\ C = \lambda .$$

Приклад 5. Неперервна випадкова величина розподілена за експоненціальним законом $f(x) = 4e^{-4x}$, якщо $x \ge 0$; $f(x) = 0$, якщо $x < 0$. Знайти математичне сподівання, середнє квадратичне відхилення та дисперсію *X*.

За умовою $\lambda = 4$. Отже :

$$M(X) = \sigma(X) = \frac{1}{\lambda} = \frac{1}{4} = 0{,}25;\ D(X) = \frac{1}{\lambda^2} = \frac{1}{4^2} = 0{,}0625 .$$

Зауваження. Нехай є підстави вважати, що випадкова величина, яка вивчається на практиці має показниковий розподіл. Для того щоб перевірити цю гіпотезу знаходять математичне сподівання і середнє квадратичне відхилення. Математичне сподівання і середнє квадратичне відхилення експоненціального розподілу рівні між собою, а тому їх оцінки мають різнитися дуже мало. Якщо оцінки близькі одна до одної, то дані спостереження підтверджують гіпотезу про експоненціальний розподіл випадкової величини, яка вивчається; якщо оцінки відрізняються суттєво, то гіпотезу слід "відкинути".

Приклад 6. Час безвідмовної роботи елемента розподілено за показниковим законом $f(t) = 0{,}04e^{-0{,}04t}$, якщо $t \ge 0$ (t — години). Знайти ймовірність того, що елемент буде працювати безвідмовно 200 годин.

За умовою стала інтенсивність відмови $\lambda = 0,04$. Скористаємося формулою $R(t) = e^{-\lambda t}$.

$$R(200) = e^{-0,04 \cdot 200} = e^{-8} = 0,449 .$$

Шукана ймовірність того, що елемент пропрацює безвідмовно 200 годин, наближено дорівнює 0,45.

Вказівка. Оскільки функція розподілу $F(t) = P(T < t)$ визначає *ймовірність відмови* за час тривалістю *t*, то ймовірність безвідмовної роботи за цей час тривалістю *t* дорівнює $R(t) = P(T > t) = 1 - F(t)$. Як правило, тривалість часу безвідмовної роботи елемента має показниковий розподіл з функцією розподілу:

$$F(t) = 1 - e^{-\lambda t} \text{, а тому}$$

$$R(t) = 1 - F(t) = 1 - (1 - e^{-\lambda t}) = e^{-\lambda t} .$$

Нормальний закон розподілу. Серед усіх неперервних законів розподілу ймовірностей особливе місце посідає розподіл ймовірностей з щільністю:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} . \tag{3}$$

Такий розподіл називається *нормальним розподілом* з параметрами a, σ ($\sigma > 0$). Достатньо знати ці параметри, щоб задати нормальний закон. Про випадкову величину X з таким законом розподілу ймовірностей говорять, що вона розподілена нормально з параметрами a, σ. Множник вибраний так, щоб задовольнити умові:

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

Важливість нормального закону розподілу визначається тим, що до нього, як правило, приводять задачі, пов'язані з розподілом сум великого числа випадкових величин. Покажемо, що ймовірностний зміст параметрів такий: a є математичне сподівання, σ — середньо квадратичне відхилення.

Знайдемо $F(x)$:

$$F(x)=\int\limits_{-\infty}^{x} f(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left|\begin{array}{l}\dfrac{x-a}{\sigma}=t\\ x=a+\sigma t\\ dx=\sigma\cdot dt\end{array}\right|=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\sigma\cdot dt=$$

$$=\frac{1}{\sqrt{2\pi}}\left(\int\limits_{-\infty}^{0} e^{-\frac{t^2}{2}}dt+\int\limits_{0}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}dt\right).$$

Вираз $\int\limits_{-\infty}^{0} e^{-\frac{t^2}{2}}dt=\frac{1}{2}\int\limits_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt=\frac{\sqrt{2\pi}}{2}$ називають *інтегралом Пуассона*.

Як видно інтеграл від щільності $f(x)$ не виражається через елементарні функції. Для розрахунку ймовірностей випадкових величин з нормальним розподілом складені таблиці спеціальної функції:

$$\Phi(t)=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{t} e^{-\frac{x^2}{2}}dx, \tag{4}$$

яку називають інтегралом ймовірності (або функцією Лапласа). У таблиці 2 додатку наводяться значення $\Phi(t)$ тільки для додатних значень t; значення $\Phi(t)$ для відємних значень t обчислюються за допомогою властивості непарної функції:

$$\Phi(-t)=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{-t} e^{-\frac{x^2}{2}}dx=-\Phi(t). \tag{5}$$

Графік функції $2\Phi(t)$ наведений на рис. 16.

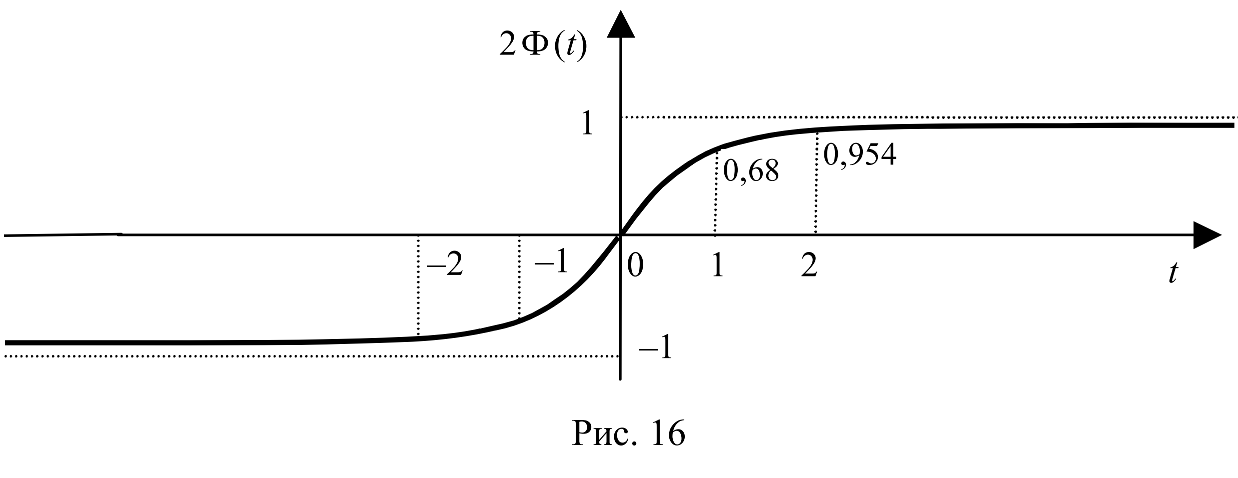


Рис. 16

При зміні t від 0 до $+\infty$ функція $2\Phi(t)$ зростає від значення $2\Phi(0)=0$ до значення $2\Phi(+\infty)=1$, при чому зростає дуже швидко, уже $2\Phi(3)=0.9973$, а $2\Phi(4)=0,999937$. Таким чином:

$$F(x)=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right) \tag{5'}$$

де:

$$\Phi\left(\frac{x-a}{\sigma}\right)=\frac{1}{\sqrt{2\pi}}\int\limits_0^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}dt \tag{6}$$

Знайдемо математичне сподівання нормальної випадкової величини X:

$$M(X)=\int\limits_{-\infty}^{\infty} xf(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}(a+\sigma t)e^{-\frac{t^2}{2}}\sigma dt=$$

$$=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} ae^{-\frac{t^2}{2}}dt+\frac{\sigma}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} te^{-\frac{t^2}{2}}dt=a+0=a$$

тобто математичне сподівання випадкової величини дорівнює параметру “a”.

Знайдемо дисперсію тієї ж величини X :

$$D(X)=\int\limits_{-\infty}^{\infty}(x-a)^2 f(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}(x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}}dx\,.$$

Введемо нову змінну $t=\dfrac{x-a}{\sigma}$. Звідси $x-a=\sigma t$, $dx=\sigma\cdot dt$, а межі інтегрування залишаються попередніми, одержимо:

$$D(X)=\frac{\sigma^2}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} t\cdot te^{-\frac{t^2}{2}}dt \;=\left|\begin{matrix} u=t & dv=te^{-\frac{t^2}{2}}dt \\ du=dt & v=-e^{-\frac{t^2}{2}}\end{matrix}\right|=$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\left(-te^{-\frac{t^2}{2}}\Bigg|_{-\infty}^{\infty}+\int\limits_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt\;\right)=\frac{\sigma^2}{\sqrt{2\pi}}(0+\sqrt{2\pi})=\sigma^2.$$

Інтегрування здійснювалося за частинами.

Підстановка: $\dfrac{(x-a)^2}{2\sigma^2}=t$, $(x-a)dx=\sigma^2 dt$ приводить останній інтеграл до гамма-функції Ейлера (властивості гамма-функції наводяться в більш детальних підручниках), тобто:

$$D(X)=\frac{2\sigma^2}{\sqrt{\pi}}\int_0^{\infty} t^{\frac{1}{2}}e^{-t}dt=\sigma^2\frac{2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right)=\sigma^2,$$

де $\Gamma\left(\dfrac{3}{2}\right)=\displaystyle\int_0^{\infty} t^{\frac{1}{2}}e^{-t}dt=\frac{\sqrt{\pi}}{2}$.

(оскільки гама-функція Ейлера має вигляд $\Gamma(\alpha)=\displaystyle\int_0^{\infty} t^{\alpha-1}e^{-t}dt$).

Зауваження. Значення $\Gamma(\alpha)$ можна знайти в таблицях гама-функції, які наводяться в багатьох довідниках.

Оскільки $D(X)=\sigma^2$, то $\sigma=\sqrt{D(X)}$.

Отже, параметр σ — є середнє квадратичне відхилення випадкової величини від свого математичного сподівання.

Властивості функції (3) зображено на рис. 17, де показано вплив параметра σ на криву нормального розподілу. Графік щільності нормального розподілу називають нормальною кривою (кривою Гауса). Досліджують функцію

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

методами диференціального числення:

1. Область визначення вся числова вісь $-\infty<x<\infty$.

2. За всіх значень x крива приймає лише додатні значення, тобто нормальна крива розташована над віссю OX.

3. Границя функції: $\lim\limits_{|x|\to\infty} y=0$, тобто вісь OX є горизонтальна асимптота графіка.

4. Знайдемо першу похідну:

$$y'=-\frac{x-a}{\sigma^3\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Легко переконатись, що $y' = 0$, якщо $x = a$; $y' > 0$, якщо $x < a$; $y' < 0$, якщо $x > a$.

Отже, якщо $x = a$ функція має максимум, рівний $\dfrac{1}{\sigma\sqrt{2\pi}}$.

5. Різниця $x - a$ у функцію входить у квадраті, а тому графік функції симетричний відносно прямої $x = a$.

6. Знайдемо другу похідну і прирівняємо її до нуля.

$$y'' = -\frac{1}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} \left[1 - \frac{(x-a)^2}{\sigma^2}\right].$$

Легко переконатись, що якщо $x = a + \sigma$ і $x = a - \sigma$, то $y'' = 0$, а при переході через ці точки вона змінює знак. Отже, точки графіка $(a - \sigma; \dfrac{1}{(\sigma\sqrt{2\pi})e})$ і $(a + \sigma; \dfrac{1}{(\sigma\sqrt{2\pi})e})$ є точки перегину.

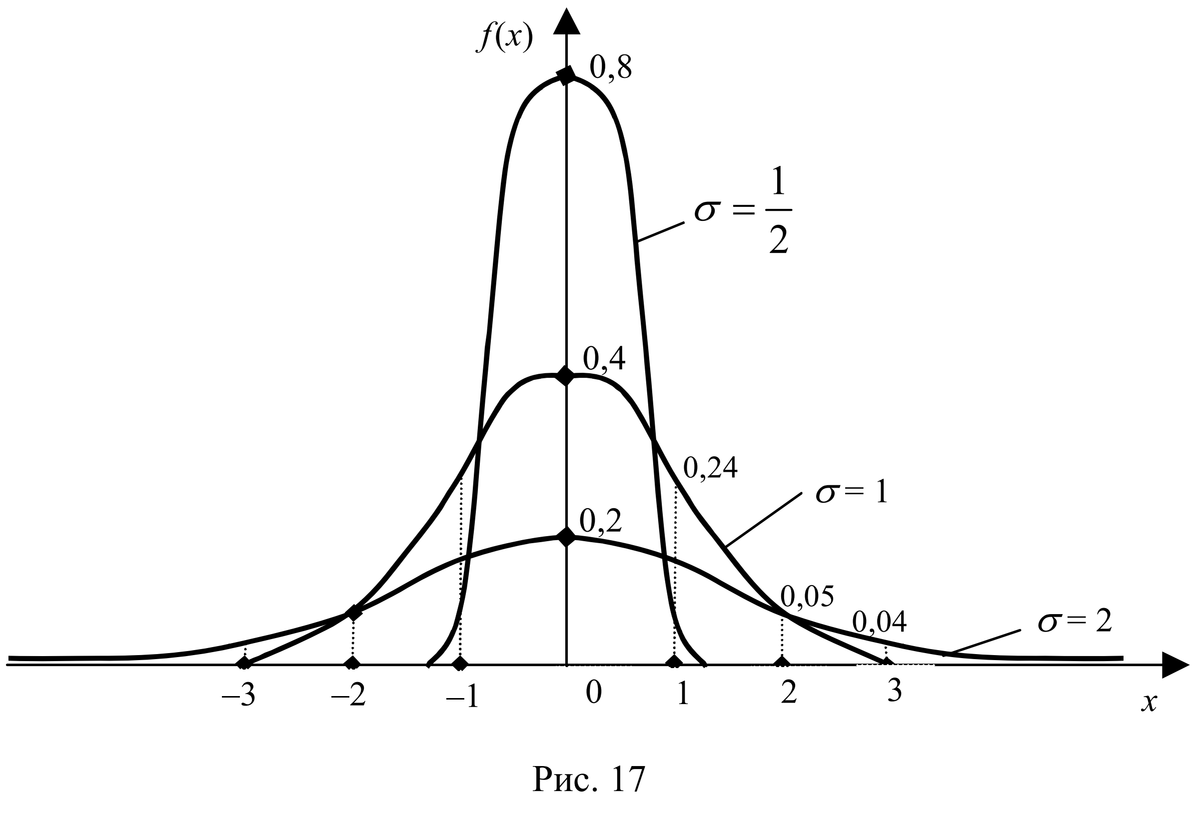


Рис. 17

На цьому рисунку наведені криві нормального розподілу при $a = 0$ і при значеннях $\sigma = \dfrac{1}{2}$, 1 і 2.

При $a = 0$ функція приймає вигляд:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{\sigma^2}}. \qquad (7)$$

Отже:

1) при $x = 0$, $f(0) = \frac{1}{\sigma\sqrt{2\pi}}$ функція (7) має максимум.

Зі зменшенням параметра σ крива розподілу стискається вздовж осі абсцис і витягується вздовж осі ординат, що призводить до збільшення ймовірності попадання випадкової величини X у будь-який фіксований окіл точки $x = 0$. Таким чином, параметр σ *характеризує розсіювання* випадкової величини X.

2) площа, обмежена кривою розподілу і віссю Ox:

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

3) якщо $a \neq 0$, то крива (3) нормального розподілу за однакових значень σ має ту саму форму, що і крива (7) і зі зростанням a буде зміщуватися зліва направо (рис. 18).

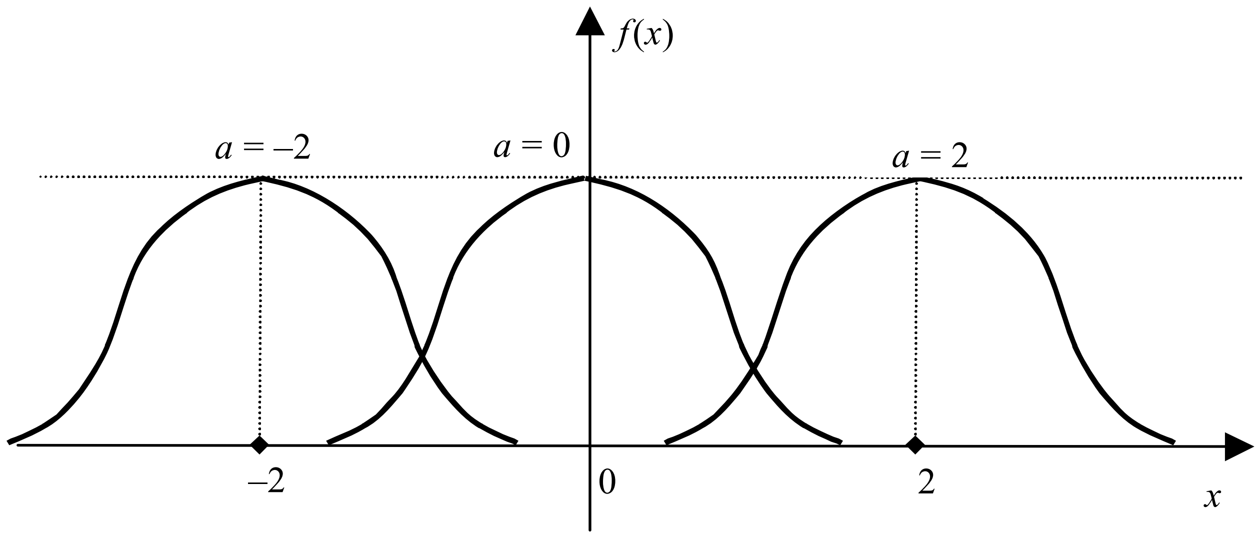


Рис. 18

Параметр "a" визначає зсув уздовж осі абсцис. За такого зсуву крива розподілу ймовірностей є симетричною відносно прямої $x = a$, так що значення "a" є центром симетрії розподілу.

Таким чином, у нормальному розподілі обидва параметри мають простий статистичний зміст: параметр “a” *є центром розподілу, параметр “σ^2” є дисперсією розподілу.*

За допомогою інтеграла ймовірності $\Phi(t)$ можливо обчислити ймовірність попадання випадкової величини X з нормальним розподілом у будь-який інтервал (x_1; x_2).

Відомо, що ймовірність попадання випадкової величини X в інтервал (x_1; x_2) у загальному випадку дорівнює :

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx \tag{8}$$

Для нормального розподілу маємо:

$$P(x_1 < X < x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left| \begin{matrix} t = \dfrac{x-a}{\sigma} \\ dx = \sigma dt \end{matrix} \right| = \frac{\sigma}{\sigma\sqrt{2\pi}} \int_{\frac{x_1-a}{\sigma}}^{\frac{x_2-a}{\sigma}} e^{-\frac{t^2}{2}} dt =$$

$$= \frac{1}{\sqrt{2\pi}} \left(\int_0^{\frac{x_2-a}{\sigma}} e^{-\frac{t^2}{2}} dt - \int_0^{\frac{x_1-a}{\sigma}} e^{-\frac{t^2}{2}} dt \right) = \Phi\left(\frac{x_2-a}{\sigma}\right) - \Phi\left(\frac{x_1-a}{\sigma}\right). \tag{9}$$

Виділимо *стандартний нормальний розподіл* з параметрами a = 0, $\sigma = 1$, з щільністю

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{10}$$

Крива стандартного нормального розподілу зображена на рис.17.

Вона симетрична відносно осі ординат (в силу парності функції $f(x)$), при x = 0 має єдиний максимум, рівний $\dfrac{1}{\sqrt{2\pi}} \approx 0,40$, і має дві точки перегину при $x = \pm 1$. При $x \to \pm\infty$ крива розподілу асимптотично наближається до осі абсцис, при чому дуже швидко (наприклад, при $f(3) = 0,0044$, $f(4) = 0,00013$). Враховуючи попередню формулу (10) при a = 0, σ = 1, маємо

$$P(x_1 < X < x_2) = \Phi(x_2) - \Phi(x_1), \tag{11}$$

У частинному випадку для симетричного інтервалу (–*t*; *t*) одержимо:

$$P(-t < X < t) = \Phi(t) - \Phi(-t) = 2\Phi(t)\,.$$

Останнє співвідношення дозволяє тлумачити функцію $\Phi(t)$ при $t > 0$ як ймовірність попадання в симетричний інтервал (–*t*; *t*) для випадкової величини зі стандартним нормальним законом розподілу.

Для загального нормального розподілу формулу $P(-t < X < t) = \Phi(t) - \Phi(-t) = 2\Phi(t)$ перепишемо у вигляді:

$$P(x_1 < X < x_2) = \Phi(t_2) - \Phi(t_1)\,, \tag{11'}$$

де $t_1 = \dfrac{x_1 - a}{\sigma}$, $t_2 = \dfrac{x_2 - a}{\sigma}$.

Якщо випадкова величина *X* розподілена за нормальним законом, то ймовірність того, що відхилення випадкової величини від її математичного сподівання по абсолютній величині не більше заданого числа $\varepsilon > 0$, обчислюється (з урахуванням (11')) за формулою:

$$P\left(\left|X - a\right| < \varepsilon\right) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right), \tag{12}$$

де $a = M(X)$, а $\sigma = \sqrt{D(X)}$.

У рівності (12) позначимо $\dfrac{\varepsilon}{\sigma} = t$, $\varepsilon = \sigma t$, тоді:

$$P\left(\left|X - a\right| < \sigma t\right) = P\left(a - t\sigma < X < a + t\sigma\right) = 2\Phi(t)\,.$$

Це співвідношення ще раз підкреслює роль стандарту σ як характеристики розсіювання випадкової величини X. Наприклад:

$$P\left(\left|X - a\right| < 3\sigma\right) = 2\Phi(3) = 0{,}9973 > 1 - 0{,}003\,; \tag{13}$$

$$P\left(\left|X - a\right| < 2\sigma\right) = 2\Phi(2) = 0{,}9545\,;$$

$$P\left(\left|X - a\right| < \sigma\right) = 2\Phi(1) = 0{,}6827\,.$$

Тобто відхилення випадкової величини *X* нормального закону розподілу від свого математичного сподівання за абсолютною величиною не перевищує σ, 2σ, 3σ з ймовірностями 0,6827; 0,9545; 0,9973 відповідно.

У застосуваннях формулу, як правило, трактують наступним чином: для нормально розподіленої випадкової величини практично достовірно, що її відхилення від центра (свого математичного сподівання) будуть менше потроєного стандартного відхилення (правило трьох сигм). Це означає, що тут нехтують можливістю появи випадків, які мають ймовірності менші, ніж 0,003.

Приклад 7. Нехай маса виловленої риби підкоряється нормальному закону з параметрами a = 375 *г*, σ =25 *г*. Знайти ймовірність того, що маса однієї зловленої риби буде: 1) від 300 до 425 *г*; 2) не більше 450 *г*; 3) більше 300 *г*.

За формулою (11'), якщо $x_1 = 300$, $x_2 = 425$, $a = 375$ і $\sigma = 25$ одержимо:

$$P(300 < X < 425) = \Phi(\frac{425-375}{25}) - \Phi(\frac{300-375}{25}) =$$
$$= \Phi(2) - \Phi(-3) = \Phi(2) + \Phi(3) = 0,9759$$

$$P(X < 450) = P(0 < X < 450) = \Phi(\frac{450-375}{25}) - \Phi(\frac{0-375}{25}) =$$
$$= \Phi(3) - \Phi(-15) = \Phi(3) + \Phi(15) = 0,9986$$
$$P(300 < X) = P(300 < X < +\infty) = \Phi(\frac{\infty-375}{25}) - \Phi(\frac{300-375}{25}) =$$
$$= \Phi(+\infty) - \Phi(-3) = \frac{1}{2} + \Phi(3) = 0,9986.$$

Формула (11') спрощується і стає формулою (12), якщо межі допустимих значень випадкової величини симетричні відносно математичного сподівання: одна менша, а інша більша від нього на одне і теж число $\varepsilon > 0$.

Приклад 8. Нехай діаметр виготовленої деталі є випадковою величиною, розподіленою за нормальним законом з параметрами a = 4,5 *см*, σ = 0,1 *см*. Знайти ймовірність того, що розмір діаметра взятої навмання деталі різниться від математичного сподівання не більше ніж на 2 *мм*.

За формулою (12), якщо ε = 0,2 *см* і σ = 0,1 *см*, одержимо, що шукана ймовірність

$$P(|X - 4,5| \le 0,2) = 2\Phi(\frac{0,2}{0,1}) = 2\Phi(2) = 0,9545 .$$

Приклад 9. Середнє квадратичне відхилення випадкової величини, розподілене за нормальним законом, дорівнює 3 *см*, а математичне сподівання дорівнює 18 *см*. Знайти межі, в яких з ймовірністю 0,97 слід очікувати значення випадкової величини.

За формулою (12), якщо σ=3 маємо $2\Phi(\frac{\varepsilon}{3}) = 0,97$. За таблицею значень функції $\Phi(x)$, знаходимо, що $2\Phi(x) = 0,97$, якщо x = 2,17. Отже, $\frac{\varepsilon}{3} = 2{,}17$, звідси $\varepsilon = 6,51$.

Таким чином, з ймовірністю 0,97 можливо очікувати, що відхилення випадкової величини від математичного сподівання не перевершить за абсолютною величиною 6,51 *см* ,тобто ця випадкова величина буде знаходитися в межах $18 \pm 6,51$, тобто від 11,49 *см* до 24,51 *см*.

Приклад 10. Вважається, що відхилення довжини деталі, яка виготовляється від стандарту є випадковою величиною, розподіленою за нормальним законом. Якщо стандартна довжина дорівнює a = 60 *см* і середнє квадратичне відхилення дорівнює σ = 0,6 *см*, то яку точність довжини виробу можна гарантувати з ймовірністю 0,9?

Потрібно знайти таке додатне число ε, для якого

$$P(|X - 60| < \varepsilon) = 0,9.$$

Оскільки

$$P(|X - 60| < \varepsilon) = 2\Phi(\frac{\varepsilon}{0,6}) = 2\Phi(1,67\varepsilon),$$

то задача зводиться до розв'язання нерівності $2\Phi(1,67\varepsilon) > 0,9$. За допомогою таблиці (додаток 2) встановимо, що $1,67\varepsilon > 1,65$. Залишається знайти найменше значення ε, яке задовольняє цій нерівності, звідси $\varepsilon = 0,99$.

Поняття про моменти розподілу, асиметрію випадкової величини. Введені раніше дві основні характеристики розподілу — центр розподілу $M(X) = a$ та дисперсія $D(X) = M\left(X - M(X)\right)^2$ — представляють собою частинні випадки розподілу, введені видатним математиком П.Л.Чебишевим для дослідження законів розподілу ймовірностей.

Розглянемо дискретну випадкову величину та її закон розподілу:

X	1	3	5	100
P	0,5	0,3	0,18	0,01

$$M(X)=1\cdot 0,5+3\cdot 0,3+5\cdot 0,18+100\cdot 0,01=3,3$$

Закон розподілу величини X^2 запишемо:

X^2	1	9	25	10000
P	0,5	0,3	0,18	0,01

$$M(X^2)=1\cdot 0,5+9\cdot 0,3+25\cdot 0,18+10000\cdot 0,01=107,7\ .$$

таким чином, у цьому прикладі перехід від $M(X)$ до $M(X^2)$ дозволяє повніше врахувати вплив на математичне сподівання того можливого значення, яке велике, а ймовірність має малу. А тому часто має сенс розглядати математичне сподівання цілого додатного степеня випадкової величини (не тільки дискретної, але й неперервної).

Початковим моментом порядка k випадкової величини X називають математичне сподівання величини X^k:

$$\nu_k=M(X^k)\,. \tag{14}$$

Зокрема, $\nu_1=M(X)$; $\nu_2=M(X^2)$ і т.д.

Користуючись цими моментами, дисперсію можна обчислити так:

$$D(X)=M(X^2)-\left[M(X)\right]^2=\nu_2-\nu_1^2\,. \tag{15}$$

Початковий момент першого порядку є центром розподілу випадкової величини X:

$$\nu_1=M(X)=a\,.$$

Центральним моментом порядку k випадкової величини X називається математичне сподівання величини $\left[X-M(X)\right]^k$:

$$\mu_k=M\left[\left(X-M(X)\right)^k\right]. \tag{16}$$

Зокрема: $\mu_1=M\left[X-M(x)\right]=0$

$$\mu_2=M\left[(X-M(x))^2\right]=D(X)\,. \tag{17}$$

Центральний момент другого порядку будь-якої випадкової величини дорівнює дисперсії, тобто $\mu_2 = D(X)$. Легко вводяться співвідношення, які зв'язують початкові і центральні моменти. Наприклад, порівнюючи (15) і (17), маємо $\mu_2 = \nu_2 - \nu_1^2$ і т.д.

Для розрахунку моментів часто корисними є наступні співвідношення між початковими і центральними моментами:

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3$$

$$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4 \text{ тощо.}$$

Ці співвідношення легко встановлюються за допомогою властивостей лінійності математичного сподівання, наприклад:

$$\mu_3 = M\left[\left(X - M(X)\right)^3\right] = M\left[X^3 - 3M(X)X^2 + 3M^2(X)X - M^3(X)\right] =$$
$$= M\left(X^3\right) - 3M(X)M\left(X^2\right) + 3M^2(X)M\left(X\right) - M^3(X) = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3 .$$

Легко переконатися, що якщо розподіл випадкової величини X симетричний відносно її математичного сподівання $M(X) = a$, тобто якщо крива розподілу ймовірностей симетрична відносно прямої $x = a$, то центральні моменти непарних моментів дорівнюють нулю, тобто $\mu_1 = \mu_3 = \mu_5 = ... = 0$. Зокрема, це справедливо для нормального і рівномірного розподілу. А тому відмінність центральних моментів непарних порядків від нуля свідчить про асиметрію розподілу. Як правило, для характеристики асиметрії розподілу використовують центральний момент третього порядку:

$$\mu_3 = M\left[\left(X - a\right)^3\right].$$

Найчастіше застосовується безрозмірний коефіцієнт асиметрії:

$$A_k = \frac{M\left[\left(X - a\right)^3\right]}{\sigma^3(X)} = \frac{\mu^3}{\left(\mu_2\right)^{3/2}} \tag{18}$$

відношення центрального моменту третього порядку до кубу середнього квадратичного відхилення σ.

Знак коефіцієнта вказує на правосторонню або лівосторонню асиметрію (рис. 19).

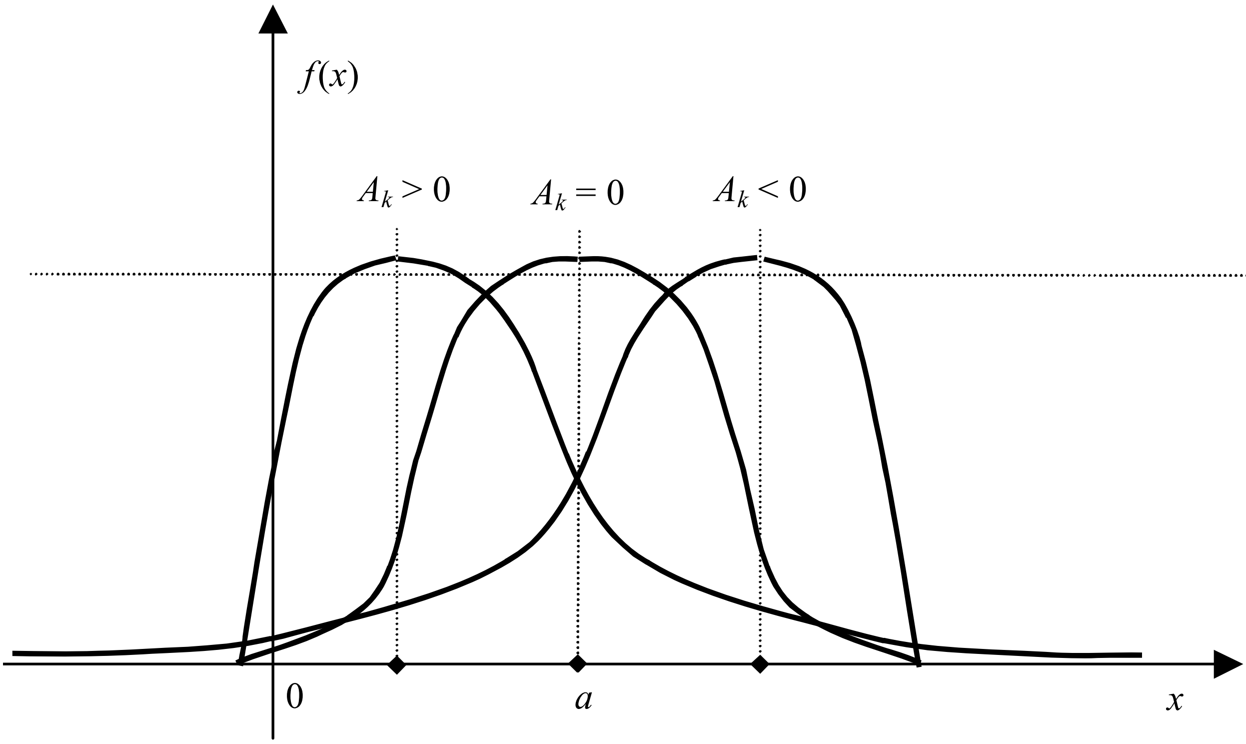


Рис. 19

Якщо розподіл симетричний відносно математичного сподівання, то $A_k = 0$. На рисунку зображена гістограма для $A_k > 0$ і $A_k < 0$.

Іноді застосовуються ще і моменти четвертого порядку $\mu_4 = M\left[\left(X-a\right)^4\right]$ для характеристики згладжування кривої розподілу відносно до кривої нормального розподілу. Відповідний безрозмірний коефіцієнт називається *ексцесом* і визначається за формулою:

$$\mathrm{E} = \frac{\mu_4}{\sigma^4(X)} - 3.$$

Для нормального розподілу центральний момент четвертого порядку дорівнює $\mu_4 = 3\sigma^4$, а тому ексцес дорівнює нулю.

Криві більш гостроверхі порівняно з нормальною мають додатній ексцес; для кривих більш плосковерхих, Е < 0 (рис. 20). Отже, якщо ексцес деякого розподілу відмінний від нуля, то крива цього розподілу різниться від нормальної кривої: якщо ексцес додатній, то крива має більш високу і гостру вершину ніж нормальна крива, якщо ексцес від'ємний, то порівняльна крива має більш низьку і плоску вершину ніж нормальна крива.

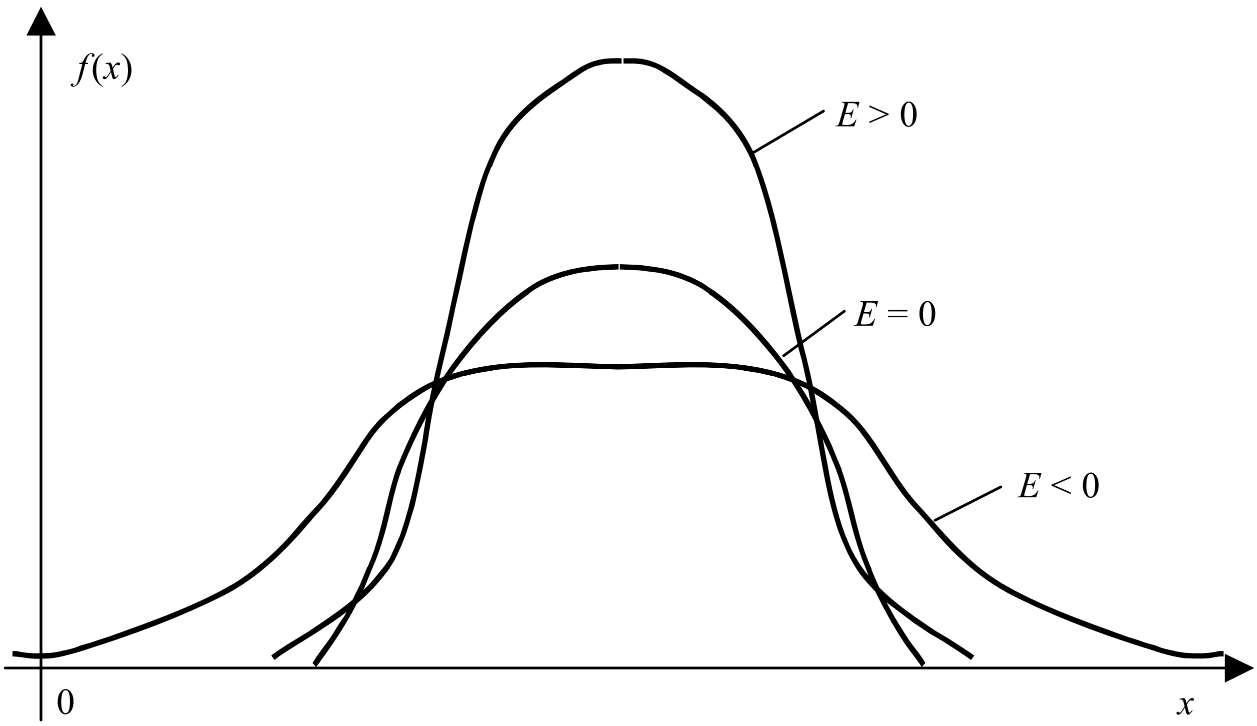


Рис. 20

Приклад 11. Випадкова величина X задана щільністю розподілу $f(x) = 2x$ в інтервалі (0;1); поза межами цього інтервалу $f(x) = 0$. Знайти початкові і центральні моменти першого, другого, третього і четвертого порядків.

За формулою (14)

$$\nu_k = \int_0^1 x^k f(x)dx$$

знайдемо початкові моменти:

$$\nu_1 = \int_0^1 x \cdot 2x \cdot dx = \frac{2}{3};\ \nu_2 = \int_0^1 x^2 \cdot 2x \cdot dx = \frac{1}{2};$$

$$\nu_3 = \int_0^1 x^3 \cdot 2x \cdot dx = \frac{2}{5};\ \nu_4 = \int_0^1 x^4 \cdot 2x \cdot dx = \frac{1}{3}.$$

Знайдемо центральні моменти. Центральний момент першого порядку будь-якої випадкової величини $\mu_1 = 0$. Скористаємося формулами, які виражають центральні моменти через початкові.

$$\mu_2 = \nu_2 - \nu_1^2;\ \mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3;\ \mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4.$$

Підставимо в ці формули, раніше знайдені початкові моменти, одержимо:

$$\mu_2 = \frac{1}{18};\ \mu_3 = -\frac{1}{135};\ \mu_4 = \frac{1}{135}.$$

ЗАДАЧІ

1.38. Ціна поділки шкали вимірювального приладу дорівнює 0,2. Показання приладу заокруглюють до найближчої ціни поділки. Знайти ймовірність того, що за відліком буде зроблена помилка: а) менша 0,04: б) більша 0,05.

Відповідь: а) $P(0 < X < 0{,}04) + P(0{,}16 < X < 20) = 0{,}4$;
б) $P(0{,}05 < X < 0{,}15) = 0{,}5$.

1.39. Неперервна випадкова величина X розподілена за показниковим законом, заданому при $x \geq 0$ щільністю розподілу $f(x) = 0{,}04 \cdot e^{-0{,}04x}$; якщо $x < 0$ $f(x) = 0$. Знайти ймовірність того, що в результаті випробування X попаде в інтервал (1;2).

Відповідь: $P(1 < X < 2) = 0{,}038$.

1.40. Знайти дисперсію і середнє квадратичне відхилення показникового розподілу, заданого при $x \geq 0$: а) щільністю $f(x) = 10 \cdot e^{-10x}$; б) функцією розподілу $F(x) = 1 - e^{-0{,}4x}$.

Відповідь: а) $D(X) = 0{,}01$; $\sigma(X) = 0{,}1$; б) $D(X) = 6{,}25$; $\sigma(X) = 2{,}5$.

1.41. Знайти: а) центральний момент третього порядку $\mu_3 = M\left[X - M(X)\right]^3$; б) асиметрію $A_k = \dfrac{\mu_3}{\sigma^3(X)}$; в) ексцес $E = \dfrac{\mu_4}{\sigma^4(X)} - 3$ показникового розподілу.

Відповідь: а) $\mu_3 = \dfrac{2}{\lambda^3}$; б) $A_k = 2$; в) E=6.

1.42. Функція $f(x) = \dfrac{(2a - x)}{2a^2}$ якщо $0 \leq x \leq 2a$; $f(x) = 0$ якщо $x < 0$ і $x > 2a$ є щільністю ймовірності випадкової величини X. Знайти її функцію розподілу, побудувати графік $f(x)$ і функції розподілу.

Відповідь: $F(x)=0$ якщо $x<0$; $F(x)=\frac{1}{a^2}(ax-\frac{x^2}{4})$ якщо $0\le x\le 2a$; $F(x)=1$ якщо $x>2a$.

1.43. Знайти математичне сподівання і дисперсію випадкової величини заданої в задачі 1.42.

Відповідь: $M(X)=\frac{2a}{3}$; $D(X)=\frac{2a^2}{9}$.

1.44. Розмір діаметру деталі, яка випускається цехом, розподіляється за нормальним законом з параметрами a = 5 *см*, $\sigma^2=0{,}81$. Знайти ймовірність того, що діаметр навмання вибраної деталі: а) складе від 4 до 7 *см*; б) різниться від математичного сподівання не більше ніж на 2 *см*.

Відповідь: а) 0,8536; б) 0,9737.

1.45. Результати замірів відстані між двома населеними пунктами розподілені за нормальним законом з параметрами a = 16 *км*, $\sigma=100$ *м*. Знайти ймовірність того, що відстань між цими пунктами: а) не менше 15,8 *км*; б) не більше 16,25 *км*; в) від 15,75 до 16,3 *км*.

Відповідь: а) 0,9772; б) 0,9938; в) 0,9924.

1.46. Діаметр деталі, що виготовляється цехом, є випадковою величиною, розподіленою за нормальним законом. Її дисперсія дорівнює 0,0001, а математичне сподівання дорівнює 2,5 *см*. Знайти межі, в яких з ймовірністю 0,9973 міститься діаметр навмання вибраної деталі.

Відповідь: від 2,47 до 2,53 *см*.

§7. Закон великих чисел і граничні теореми.

Для практики дуже важливо знати умови, за виконання яких сукупна дія багатьох випадкових причин призводить до результату, який майже не залежить від випадку, оскільки дозволяє завбачити хід явища. На ці умови вказують теореми, які мають загальну назву закону великих чисел. До них належать теореми Чебишова і Бернуллі.

Базою для доведення вказаних теорем є нерівність Чебишова, яка справедлива як для дискретних, так і для неперервних випадкових величин. Обмежимося доведенням цієї нерівності для дискретної величини X, заданої таблицею розподілу:

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

Нерівність Чебишова. *Яке б не було* $\varepsilon > 0$*, для будь-якої випадкової величини* X*, дисперсія якої скінченна, виконується нерівність:*

$$P\left(\left|X - M(X)\right| < \varepsilon\right) \geq 1 - \frac{D(X)}{\varepsilon^2}. \qquad (1)$$

Доведення. Оскільки події, які полягають у виконанні нерівностей $\left|X - M(X)\right| < \varepsilon$ і $\left|X - M(X)\right| \geq \varepsilon$ протилежні, то сума їх ймовірностей дорівнює одиниці, тобто:

$$P\left(\left|X - M(X)\right| < \varepsilon\right) + P\left(\left|X - M(X)\right| \geq \varepsilon\right) = 1.$$

Звідси

$$P\left(\left|X - M(X)\right| < \varepsilon\right) = 1 - P\left(\left|X - M(X)\right| \geq \varepsilon\right). \qquad (2)$$

Отже, задача зводиться до обчислення ймовірності:

$$P\left(\left|X - M(X)\right| \geq \varepsilon\right).$$

Для цього запишемо вираз дисперсії випадкової величини X:

$$D(X) = \left[x_1 - M(X)\right]^2 p_1 + \left[x_2 - M(X)\right]^2 p_2 + ... + \left[x_n - M(X)\right]^2 p_n$$

Очевидно всі доданки цієї суми невід'ємні. Доданки $\left|x_i - M(X)\right| < \varepsilon$ відкинемо і враховувати не будемо, а залишимо лише доданки $\left|x_j - M(X)\right| \geq \varepsilon$, внаслідок чого сума може лише зменшитися, таким чином:

$$D(X) \geq \left[x_{k+1} - M(X)\right]^2 p_{k+1} + \left[x_{k+2} - M(X)\right]^2 p_{k+2} + ... + \left[x_n - M(X)\right]^2 p_n.$$

(для визначеності відкинули k перших доданків).

Оскільки обидві частини нерівності $\left|x_j - M(X)\right| \geq \varepsilon$ (j = k+1, k+2,..., n) додатні, то підносячи їх до квадрата одержимо рівносильну нерівність $\left|x_j - M(X)\right|^2 \geq \varepsilon^2$. Замінимо в останній сумі кожний із доданків $\left|x_j - M(X)\right|^2$ числом ε^2 (при цьому нерівність лише підсилиться), одержимо:

$$D(X) \geq \varepsilon^2 \left(p_{k+1} + p_{k+2} + \ldots + p_n\right). \tag{3}$$

За теоремою додавання сума ймовірностей $\left(p_{k+1} + p_{k+2} + \ldots + p_n\right)$ є ймовірністю того, що X прийме лише одне, байдуже яке, із значень x_{k+1}, x_{k+2}, ..., x_n, а за будь-якого із них відхилення задовольняє нерівність $\left|x_j - M(X)\right| \geq \varepsilon$. Звідси сума $p_{k+1} + p_{k+2} + \ldots + p_n$ виражає ймовірність $P\left(\left|X - M(X)\right| \geq \varepsilon\right)$.

Це дозволяє переписати попередню нерівність (3) так:

$$D(X) \geq \varepsilon^2 P\left(\left|X - M(X)\right| \geq \varepsilon\right)$$

або

$$P\left(\left|X - M(X)\right| \geq \varepsilon\right) \leq \frac{D(X)}{\varepsilon^2} \tag{4}$$

Підставляючи (4) в (2), одержимо:

$$P\left(\left|X - M(X)\right| < \varepsilon\right) \geq 1 - \frac{D(X)}{\varepsilon^2},$$

що і потрібно було довести.

Зауваження. Нерівність Чебишова для практики не дуже застосовна, оскільки дає грубу, а іноді тривіальну оцінку, але її теоретичне значення дуже велике. Далі ми скористаємося цією нерівністю для запису теореми Чебишова.

Крім того, якщо випадкова величина X розподілена за нормальним законом, то, як відомо раніше, має місце точна формула:

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

Приклад 1. Дискретна випадкова величина X задана законом розподілу:

X	1	2
p	0,8	0,2

Скориставшись нерівністю Чебишова, оцініть ймовірність того, що абсолютна величина відхилення $X - M(X)$ буде менша одиниці (тобто $\varepsilon = 1$).

Знайдемо математичне сподівання X:

$$M(X) = 1 \cdot 0,8 + 2 \cdot 0,2 = 1,2 .$$

Запишемо закон розподілу X^2:

X^2	1	4
p	0,8	0,2

Знайдемо математичне сподівання X^2:

$$M(X^2) = 1 \cdot 0,8 + 4 \cdot 0,2 = 1,6$$

Знайдемо дисперсію $D(X)$:

$$D(X) = M(X^2) - \left[M(X)\right]^2 = 1,6 - \left(1,2\right)^2 = 0,16 .$$

Скористаємося нерівністю Чебишова:

$$P\left(\left|X - 1,2\right| < 1\right) \ge 1 - \frac{0,16}{1} \approx 0,84 .$$

Теорема Чебишова (закон великих чисел у формі Чебишова). Якщо випадкові величини $X_1, X_2, ..., X_n$ незалежні, мають математичне сподівання $a_1, a_2, ..., a_n$ та дисперсію $D(X_1)$, $D(X_2)$, ..., $D(X_n)$, кожна з яких обмежена одним і тим же числом C, то для будь-якого додатного числа ε виконується нерівність:

$$P\left(\left|\frac{X_1 + X_2 + ... + X_n}{n} - \frac{a_1 + a_2 + ... + a_n}{n}\right| < \varepsilon\right) \ge 1 - \frac{C}{n\varepsilon^2} . \qquad (5)$$

Доведення. Введемо позначення:

$$Y_n = \frac{X_1 + X_2 + ... + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}, \ A = \frac{a_1 + a_2 + ... + a_n}{n};$$

тоді з (5) маємо:

$$P\left(\left|Y_n - A\right| < \varepsilon\right) \geq 1 - \frac{C}{n\varepsilon^2}. \qquad (6)$$

це закон великих чисел у формі Чебишова.

Знайдемо математичне сподівання та дисперсію випадкової величини:

$$M\left(Y_n\right) = M\left(\frac{X_1 + X_2 + ... + X_n}{n}\right) = \frac{\sum_{i=1}^{n} M(X_i)}{n} = \frac{a_1 + a_2 + ... + a_n}{n} = A$$

$$D\left(Y_n\right) = D\left(\frac{X_1 + X_2 + ... + X_n}{n}\right) = \frac{1}{n^2}\left[D\left(X_1\right) + D\left(X_2\right) + ... + D\left(X_n\right)\right]$$

Згідно з умовою теореми $D\left(X_1\right) \leq C$, $D\left(X_2\right) \leq C$, ..., $D\left(X_n\right) \leq C$, то $D\left(Y_n\right) \leq \frac{nC}{n^2} = \frac{C}{n}$.

Застосуємо до величини Y_n нерівність Чебишова:

$$P\left(\left|Y_n - A\right| < \varepsilon\right) \geq 1 - \frac{D(Y_n)}{\varepsilon^2} = 1 - \frac{C}{n\varepsilon^2}.$$

Теорема Чебишова зберігає силу і для випадку, коли випадкові величини утворюють послідовність.

Наслідок 1. У нерівності (5) перейдемо до границі, при $n \to \infty$, одержимо:

$$\lim_{n\to\infty} P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} M(X_i)}{n}\right| < \varepsilon\right) \geq 1.$$

але, оскільки, ймовірність більше одиниці не буває, то

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} M(X_i)\right| < \varepsilon\right) = 1\,. \tag{7}$$

Отже, зі зростанням n середнє арифметичне випадкових величин $X_1, X_2, ..., X_n$ збігається за ймовірністю до середнього арифметичного їх математичних сподівань.

У термінах даного визначення теорему Чебишова можливо сформулювати наступним чином: при достатньо великій кількості незалежних експериментів середнє арифметичне значення випадкових величин збігається за ймовірностю до середнього арифметичного їх математичних сподівань.

Наслідок 2. Якщо математичне сподівання послідовності $\{X_n\}$ незалежних випадкових величин однакове $M\left(X_i\right) = a$, ($i = 1, 2, ..., n, ...$), дисперсія цих випадкових величин обмежена зверху одним і тим же числом C. Тоді для будь-якого додатного числа $\varepsilon > 0$ середнє арифметичне випадкових величин збігається за ймовірністю до сталого числа a, тобто:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| < \varepsilon\right) = 1\,. \tag{8}$$

Ця границя випливає із попередньої рівності, оскільки згідно з умовою при будь-якому n:

$$\frac{1}{n}\sum_{i=1}^{n} M(X_i) = \frac{n \cdot a}{n} = a\,.$$

Таким чином, середнє арифметичне розглянутих випадкових величин за достатньо великого n скільки завгодно мало різниться від сталої величини “a” з ймовірністю, близькою до 1.

Приклад 2. Дисперсія кожної із 1000 випадкових величин не перевищує 5. Оцінити знизу ймовірність того, що абсолютна величина відхилення середнє арифметичного цих випадкових величин від середнє арифметичного їх математичних сподівань не перевищує 0,2.

Згідно з умовою: $n = 1000$, $C = 5$, $\varepsilon = 0,2$. За формулою (5) маємо:

$$P\left(\left|\frac{1}{1000}\sum_{i=1}^{1000} X_i - \frac{1}{1000}\sum_{i=1}^{1000} M\left(X_i\right)\right| < 0,2\right) \ge 1 - \frac{5}{1000 \cdot \left(0,2\right)^2} \approx 0,875\,.$$

Теорема Чебишова є одним із тих законів великих чисел, які лежать в основі всіх практичних застосувань теорії ймовірностей. Іншим і до того ж найпростішим (і раніше всіх установленим) законом великих чисел є ***теорема Я.Бернуллі***: *Якщо в кожному із n незалежних експериментів ймовірність p настання події A стала, то ймовірність того, що відхилення відносної частоти від ймовірності p не перевищить по абсолютній величині додатне число ε більше ніж різниця* $1-\frac{pq}{n\varepsilon^2}$, тобто

$$P\left(\left|\frac{m}{n}-p\right|<\varepsilon\right)\geq 1-\frac{pq}{n\varepsilon^2}, \qquad (9)$$

де $q=1-p$.

Якщо $n\to\infty$, одержимо:

$$\lim_{n\to\infty}P\left(\left|\frac{m}{n}-p\right|<\varepsilon\right)=1. \qquad (10)$$

Приклад 3. Монету підкидають 500 разів. Оцінити зверху ймовірність відхилення частоти появи “герба” від ймовірності його появи менше ніж на 0,2.

Тут $n=500$, $p=q=\frac{1}{2}$, $\varepsilon=0,2$.

Використовуючи нерівність (9), одержимо:

$$P\left(\left|\frac{m}{500}-\frac{1}{2}\right|<0,2\right)\geq 1-\frac{\frac{1}{2}\cdot\frac{1}{2}}{500\cdot 0,04}=\frac{79}{80}.$$

Оскільки нерівність $\left|\frac{m}{500}-\frac{1}{2}\right|<0,2$ рівнозначна подвійній нерівності $150<m<350$, то можливо стверджувати, що ймовірність числа випадань “герба” в інтервалі $(150;350)$ більше $\frac{79}{80}$.

Теорема Ляпунова (центральна гранична теорема теорії ймовірностей). Ця теорема розв’язує головну задачу теорії ймовірностей: знайти і вивчити поведінку закону розподілу суми великого числа незалежних випадкових величин. Як пояснити, що нормально розподілені випадкові величини широко поширені при розв'язуванні практичних задач? Виявляється, що за певних умов сума незалежних випадкових величин розподілена асимптотично нормально. Теорема має винятково велике значення, а тому її називають центральною граничною теоремою:

Якщо X_1, X_2, ..., X_n *– незалежні випадкові величини, кожна із яких має один і той же закон розподілу з математичним сподіванням а та дисперсією* σ^2*, то при необмеженому зростанні числа* n *закон розподілу суми* $X = X_1 + X_2 + ... + X_n$ *необмежено наближається до нормального.*

Введемо позначення:

$$M(X_k) = a_k, \; D(X_k) = \sigma_k^2, S_n = X_1 + X_2 + ... + X_n, \; A_n = \sum_{k=1}^{n} a_k,$$

$$B_n^2 = \sum_{k=1}^{n} \sigma_k^2 .$$

Позначимо функцію розподілу нормованої суми через:

$$F_n(x) = P\left(\frac{S_n - A_n}{B_n} < x \right).$$

Вважається, що до послідовності X_1, X_2, … застосовна центральна гранична теорема, якщо при будь-якому x функція розподілу нормованої суми при $n \to \infty$ прямує до нормальної функції розподілу:

$$\lim_{n \to \infty} P\left(\frac{S_n - A_n}{B_n} < x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt .$$

Теорема справедлива для дискретних і неперервних випадкових величин і на практиці застосовна при числі $n \geq 10$, а іноді і менше 10.

Усі викладені в цьому параграфі теореми вказують умови, за виконання яких сукупні дії дуже багатьох випадкових причин призводять до результату, майже не залежному від випадку.

ЗАДАЧІ

1.47. У результаті 200 незалежних випробувань знайдені значення випадкової величини X: x_1, x_2, …, x_{200}, причому $M(X) = D(X) = 2$. Оцінити зверху ймовірність того, що абсолютна величина різниці між середнім арифметичним значень випадкової величини $\left(\sum_{i=1}^{200} x_i\right) \cdot \frac{1}{200}$ і математичним сподіванням буде менше $\frac{1}{5}$.

Відповідь: $\frac{3}{4}$.

1.48. Довести нерівність Чебишова у формі:

$$P\left(\left|X - M(X)\right| \geq \varepsilon\right) \leq D(X)\frac{1}{\varepsilon^2}.$$

Вказівка. Скористайтесь тим, що подіЇ $\left|X - M(X)\right| < \varepsilon$ і $\left|X - M(X)\right| \geq \varepsilon$ — протилежні.

1.49. Прилад складається із 20 незалежно працюючих елементів. Ймовірність відмови кожного елемента за час T дорівнює 0,02. За допомогою нерівності Чебишова оцінити ймовірність того, що абсолютна величина різниці між числом елементів, що відмовили, і середнім числом (математичним сподіванням) відмов за час T виявиться:

а) менше чотирьох;

б) більше чотирьох.

Відповідь: а) 0,9755
б) 0,0245.

1.50. Середня кількість споживання води у даному населеному пункті дорівнює 50000 л на день. Оцінити ймовірність того, що у цьому населеному пункті у зазначений день споживання води не перебільшить 150000 л.

Відповідь: $\geq \frac{2}{3}$.

1.51. Середнє споживання електроенергії за серпень місяць дорівнює 360 000 кВТ/год;

а) оцінити ймовірність того, що споживання електроенергії у серпні наступного року перебільшить 1 000 000 кВТ/год;

б) оцінити ту ж ймовірність, якщо відомо, що середнє квадратичне відхилення споживання електроенергії за серпень дорівнює 40 000 кВТ/год.

Відповідь: а) $\geq 0,36$

б) $\geq \dfrac{1}{256}$.

§ 8. Функції випадкових величин

Функції однієї випадкової величини. Нехай X — випадкова величина; її множину можливих значень позначимо через $\{x\}$; інша випадкова величина Y зв’язана з першою функціональною залежністю $Y = f(X)$, її множину можливих значень позначимо через $\{y\}$.

Наприклад. Якщо випадкова величина X є радіус кулі, що виготовляється, то об’єм цієї кулі $Y = \dfrac{4}{3}\pi X^3$ є також випадкова величина, що є функцією від X.

Розглянемо, як можна знайти закон розподілу Y, якщо відомий закон розподілу X, тобто $Y = f(X)$. Нехай аргумент X має наступний закон розподілу $P\left(X = x_i\right) = p_i$ ($i = 1,2,...,n$), тоді функція $Y = f(X)$ прийме відповідне значення y_i ($i = 1,2,...,n$) згідно з певним правилом $y = f(x)$ з тими ж ймовірностями, тобто:

$$P\left(Y = y_i\right) = P\left(X = x_i\right) = p_i, \qquad (i = 1,2,...,n).$$

1. *Нехай аргумент X — дискретна випадкова величина.*

а) Якщо різним можливим значенням аргументу X відповідають різні можливі значення функції Y, то ймовірності відповідних значень X та Y між собою рівні.

Приклад 1. Дискретна випадкова величина задана розподілом:

X	3	4
P	0,7	0,3

Знайти розподіл функції $Y = X^3$.

Знайдемо можливі значення *Y*: $y_1 = 3^3 = 27$; $y_2 = 4^3 = 64$. Напишемо шуканий розподіл *Y*:

Y	27	64
P	0,7	0,3

б) Якщо різним можливим значенням *X* відповідають значення *Y*, серед яких є рівні між собою, то необхідно додавати ймовірності значень *Y*, що повторюються.

Приклад 2. Дискретна випадкова величина *X* задана розподілом :

X	–3	3	4
P	0,3	0,6	0,1

Знайти розподіл функції $Y = X^2$.

Ймовірність можливого значення $y_1 = 9$ дорівнює сумі ймовірностей несумісних подій $X = -3$, $X = 3$, тобто 0,3 + 0,6 = 0,9. Ймовірність можливого значення $y_2 = 16$ дорівнює 0,1. Отже, шуканий розподіл *Y*:

Y	9	16
P	0,9	0,1

2. *Нехай аргумент X – неперервна випадкова функція*. Як знайти розподіл функції $Y = \varphi(X)$ при відомій щільності розподілу випадкового аргументу *X*? Встановлено, що якщо $y = \varphi(x)$ — диференційовна і строго монотонна функція, обернена функція якої є $x = \psi(y)$, то щільність розподілу $g(y)$ випадкової величини *Y* обчислюється за формулою:

$$g(y) = f\left[\psi(y)\right] \cdot \left|\psi'(y)\right|. \tag{1}$$

Приклад 3. Задана щільність розподілу $f(x)$ випадкової величини *X*, можливі значення якої містяться в інтервалі (*a;b*). Знайти щільність розподілу випадкової величини *Y* = 2*X*.

Оскільки функція $y = 2x$ диференційовна і строго зростає, то застосовна формула (1).

Знайдемо $\psi(y)$: $\psi(y) = x = \dfrac{y}{2}$.

Знайдемо $f[\psi(y)]$: $f[\psi(y)] = f\left(\dfrac{y}{2}\right)$.

Знайдемо похідну $\psi'(y)$: $\psi'(y) = \left(\dfrac{y}{2}\right)' = \dfrac{1}{2}$.

Очевидно, що $|\psi'(y)| = \dfrac{1}{2}$.

Згідно з формулою (1) маємо: $g(y) = \left(\dfrac{1}{2}\right)\cdot f\left(\dfrac{y}{2}\right)$. Оскільки x змінюється в інтервалі (a, b) і $y = 2x$, то $2a < y < 2b$.

Нехай задана функція $Y = \varphi(X)$ випадкового аргументу *X*.

Математичне сподівання цієї функції, знаючи закон розподілу аргументу, у випадку, якщо аргумент *X* — *дискретна випадкова величина*, визначається за формулою:

$$M[\varphi(X)] = \sum_{i=1}^{n} \varphi(x_i)\, p_i \,. \tag{2}$$

Якщо аргумент *X* — *неперервна випадкова величина*, то:

$$M[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(x) f(x) dx \,, \tag{3}$$

де $f(x)$ — задана щільність розподілу величини *X*.

У випадку, коли можливі значення *X* належать інтервалу (a, b), то:

$$M[\varphi(X)] = \int_{a}^{b} \varphi(x) f(x) dx \,. \tag{4}$$

Зауваження. Для знаходження математичного сподівання функції $Y = \varphi(X)$ можливо спочатку знайти щільність розподілу $g(y)$ величини *Y*, а потім скористатися формулою:

$$M(Y) = \int_{-\infty}^{\infty} y g(y) dy \,. \tag{5}$$

Приклад 4. Дискретна випадкова величина X задана розподілом:

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

Знайти математичне сподівання функції $Y = \varphi(X) = X^3 + 2$.

$$\varphi(2) = 2^3 + 2 = 10\,,\ \varphi(3) = 3^3 + 2 = 29\,,\ \varphi(6) = 6^3 + 2 = 110\,.$$

Шукане математичне сподівання дорівнює:

$$M\left[X^3 + 2\right] = 10 \cdot 0{,}3 + 29 \cdot 0{,}6 + 110 \cdot 0{,}1 = 31{,}4\,.$$

Приклад 5. Неперервна випадкова величина задана щільністю розподілу $f(x) = \frac{1}{2}\sin x$ в інтервалі $(0;\pi)$; ззовні цього інтервалу $f(x) = 0$. Знайти математичне сподівання випадкової величини $Y = \varphi(X) = X^2$.

Скористаємося формулою (4). За умовою $f(x) = \frac{1}{2}\sin x$, $\varphi(x) = x^2$, $a = 0$, $b = \pi$. Отже:

$$M\left[\varphi(x)\right] = \frac{1}{2}\int_0^{\pi} x^2 \sin x dx = \frac{\pi^2 - 4}{2}\,.$$

Вказівка: В останньому інтегралі слід двічі застосувати інтегрування за частинами.

Функція двох випадкових аргументів. Якщо кожній парі можливих значень випадкових величин X і Y відповідає одне можливе значення випадкової величини Z, то Z називають функцією двох випадкових аргументів X і Y:

$$Z = \varphi(X;Y)\,.$$

Якщо X і Y — *дискретні* незалежні випадкові величини, то для того щоб знайти розподіл функції $Z = X + Y$, необхідно знайти всі можливі значення Z, для цього достатньо скласти кожне можливе значення X зі всіма можливими значеннями Y; ймовірності знайдених можливих значень Z дорівнюють добуткам ймовірностей окремих значень X і Y.

Приклад 6. Дискретні незалежні випадкові величини X і Y задані розподілами:

X	1	4
P	0,3	0,7

Y	2	3
P	0,4	0,6

Знайти розподіл випадкової величини $Z = X + Y$.

Можливі значення Z є відповідні суми $z_1 = 1 + 2 = 3$, $z_2 = 1 + 3 = 4$, $z_3 = 4 + 2 = 6$, $z_4 = 4 + 3 = 7$.

Знайдемо ймовірності цих можливих значень. Для того, щоб $Z = 3$, достатньо X прийняти значення $x_1 = 1$ і Y — значення $y_1 = 2$. Ймовірності цих можливих значень відповідно рівні: 0,3 і 0,4. Оскільки аргументи X та Y незалежні, то події $X = 1$ і $Y = 2$ незалежні, а тому ймовірність їх сумісного настання (тобто ймовірність події $Z = 3$) згідно з теоремою множення дорівнює $0{,}3 \cdot 0{,}4 = 0{,}12$.

Аналогічно знайдемо:

$$P(Z = 1 + 3 = 4) = 0{,}3 \cdot 0{,}6 = 0{,}18\,;$$
$$P(Z = 4 + 2 = 6) = 0{,}7 \cdot 0{,}4 = 0{,}28\,;$$
$$P(Z = 4 + 3 = 7) = 0{,}7 \cdot 0{,}6 = 0{,}42\,;$$

Шуканий закон розподілу дорівнює:

Z	3	4	6	7
P	0,12	0,18	0,28	0,42

Перевірка: 0,12 + 0,18 + 0,28 + 0,42 = 1.

Якщо X та Y — *неперервні* незалежні випадкові величини, то щільність розподілу $g(z)$ суми $Z = X + Y$ (за умови, що щільність розподілу хоча б одного із аргументів задана в інтервалі $(-\infty, \infty)$ однією формулою) може бути задана за формулою:

$$g(z) = \int_{-\infty}^{\infty} f_1(x) f_2(z - x)\, dx\,. \tag{6}$$

або за рівносильною формулою:

$$g(z) = \int_{-\infty}^{\infty} f_1(z - y) f_2(y)\, dy\,, \tag{7}$$

де f_1 і f_2 — щільності розподілу аргументів.

Щільність розподілу суми незалежних випадкових величин називається композицією.

Приклад 7. Незалежні випадкові величини X та Y задані щільностями розподілу: $f_1(x)=\frac{1}{2}e^{-\frac{x}{2}}$ ($0\le x<\infty$), $f_2(y)=\frac{1}{3}e^{-\frac{y}{3}}$ ($0\le y<\infty$).

Знайти композицію цих законів, тобто щільність розподілу випадкової величини $Z = X + Y$.

Оскільки можливі значення аргументів невід'ємні, то застосовна формула:

$$g(z)=\int_0^z f_1(x)f_2(z-x)dx\,. \tag{8}$$

$$g(z)=\int_0^z\left[\frac{1}{2}e^{-\frac{x}{2}}\right]\cdot\left[\frac{1}{3}e^{-\frac{(z-x)}{3}}\right]dx=\frac{1}{6}e^{-\frac{z}{3}}\int_0^z e^{-\frac{x}{6}}dx=e^{-\frac{z}{3}}\left(1-e^{-\frac{z}{6}}\right).$$

Зокрема, тут $z\ge 0$, оскільки $Z = X + Y$ і згідно з умовою, можливі значення X і Y невід'ємні.

Таким чином:

$g(z)=e^{-\frac{z}{3}}\left(1-e^{-\frac{z}{6}}\right)$ в інтервалі ($0;\infty$), поза цим інтервалом $g(z) = 0$.

Рекомендуємо для перевірки переконатися, що $\int_0^z g(z)dz=1$.

ЗАДАЧІ

1.52. Дискретні випадкові величини X та Y задані розподілами:

а)

X	10	12	16
P	0,4	0,1	0,5

Y	1	2
P	0,2	0,8

б)

X	4	10
P	0,7	0,3

Y	1	7
P	0,8	0,2

Знайти розподіл випадкової величини $Z = X + Y$

Відповідь: а)

Z	11	12	13	14	17	18
P	0,08	0,32	0,02	0,08	0,10	0,40

б)

Z	5	11	17
P	0,56	0,38	0,06

1.53. Незалежні випадкові величини X та Y задані щільностями розподілу:

$$f_1(x) = \frac{1}{3}e^{-\frac{x}{3}} \ (0 \le x < \infty), \ f_2(y) = \frac{1}{5}e^{-\frac{y}{5}} \ (0 \le y < \infty).$$

Знайти композицію цих законів, тобто щільність розподілу випадкової величини $Z = X + Y$.

Відповідь: $$g(z) = \begin{cases} \frac{1}{2}e^{-\frac{z}{5}}\left(1 - e^{-\frac{2z}{15}}\right), & \textit{якщо} \quad z \ge 0 \\ 0, & \textit{якщо} \quad z < 0 \end{cases}.$$

§9. Системи двох випадкових величин

Часто результат експерименту описується не однією випадковою величиною X, а декількома випадковими величинами $X_1, X_2, ..., X_n$. У цьому випадку прийнято говорити, що вказані випадкові величини утворюють систему $(X_1, X_2, ..., X_n)$.

Будемо позначати через $(X; Y)$ двовимірну випадкову величину, можливі значення якої є пари чисел $(x; y)$. Складові X та Y, розглядувані одночасно, утворюють систему двох випадкових величин.

Двовимірну величину геометрично можна тлумачити як випадкову точку $M(X; Y)$ на площині XOY або як випадковий вектор $\overrightarrow{OM}$.

Приклад 1. Станок-автомат виготовляє облицювальну плитку. Якщо контрольованими розмірами є довжина X та ширина Y, то маємо двовимірну випадкову величину (X; Y); якщо контролюється і висота Z, то маємо трьохвимірну величину (X; Y; Z). Трьохвимірну випадкову величину геометрично можна тлумачити як точку $M(X$; Y; $Z)$ у трьохвимірному просторі або як вектор $\overrightarrow{OM}$.

Доцільно розрізняти дискретні (складові цих величин дискретні) і неперервні (складові-неперервні) багатовимірні випадкові величини.

Законом розподілу ймовірностей дискретної двовимірної випадкової величини називають відповідність між можливими значеннями (x_i, y_j) та їх ймовірностями $P(x_i, y_j)$, ($i = 1, 2,..., n$; $j = 1, 2,..., m$). Як правило, закон розподілу задають: а) у вигляді таблиці з подвійним входом, яка містить можливі значення та їх ймовірності;

б) аналітично, наприклад, у вигляді функцій розподілу.

Y \ X	x_1	x_2	...	x_i	...	x_n
y_1	p_{11}	p_{21}	...	p_{j1}	...	p_{n1}
...	...	...	...	...	...	...
y_j	p_{1j}	p_{2j}	...	p_{ij}	...	p_{nj}
...	...	...	...	...	...	...
y_m	p_{1m}	p_{2m}	...	p_{im}	...	p_{nm}

де p_{ij} — ймовірність події, яка полягає в одночасному виконанні рівностей ($X = x_i$, $Y = y_j$), ($i = 1, 2,..., n$; $j = 1, 2,..., m$), при цьому сума ймовірностей, які містяться у всіх клітинах таблиці дорівнює одиниці:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} = 1 . \tag{1}$$

Таблиця може мати нескінченну множину рядків і стовпців. Перший рядок таблиці містить усі можливі значення складової X, а перший стовпець — складової Y. На перетині стовпця x_i і рядка y_j зазначена ймовірність $p(x_i, y_j)$ того, що двовимірна випадкова величина прийме значення (x_i, y_j).

За розподілом двовимірної дискретної випадкової величини можливо знайти закони розподілу кожної із складових. Наприклад, події ($X = x_1$, $Y = y_1$), ($X = x_1$, $Y = y_2$), ..., ($X = x_1$, $Y = y_m$) несумісні, а тому ймовірність $P(x_1)$ того, що X прийме значення x_1, згідно з теоремою додавання така:

$$P(x_1) = p(x_1, y_1) + p(x_1, y_2) + ... p(x_1, y_m),$$

тобто дорівнює сумі ймовірностей стовпця x_1. У загальному випадку, щоб отримати $P(X = x_i)$, потрібно взяти суму ймовірності стовпця x_i. Аналогічно можливо отримати $P(Y = y_j)$, якщо взяти суму ймовірності рядка y_j.

Приклад 2. Знайти закон розподілу складових двовимірної випадкової величини, заданої законом розподілу:

$Y \backslash X$	x_1	x_2	x_3
y_1	0,15	0,25	0,20
y_2	0,08	0,18	0,14

Додамо ймовірності за стовпцями, отримаємо ймовірності можливих значень X: $P(x_1) = 0{,}23$, $P(x_2) = 0{,}43$, $P(x_3) = 0{,}34$. Отже, закон розподілу складової X:

X	x_1	x_2	x_3
P	0,23	0,43	0,34

Перевірка: $0{,}23 + 0{,}43 + 0{,}34 = 1$.

Просумувавши ймовірності за рядками, аналогічно знайдемо розподіл складової Y:

$P(y_1) = 0{,}60$, $P(y_2) = 0{,}40$.

Y	y_1	y_2
P	0,60	0,40

Перевірка: $0{,}60 + 0{,}40 = 1$.

Функцією розподілу двовимірної випадкової величини (байдуже, дискретної або неперервної) називають функцію $F(x, y)$*, яка визначає для кожної пари чисел x; y ймовірність того, що X прийме значення, менші x, і при цьому Y прийме значення, менше y:*

$$F(x, y) = P(X < x,\ Y < y).$$

Геометрично цю рівність можливо тлумачити так: $F(x, y)$ є ймовірність того, що випадкова точка (X, Y) попаде в нескінченний квадрат з вершиною (x, y), розташований лівіше і нижче цієї вершини (рис. 21).

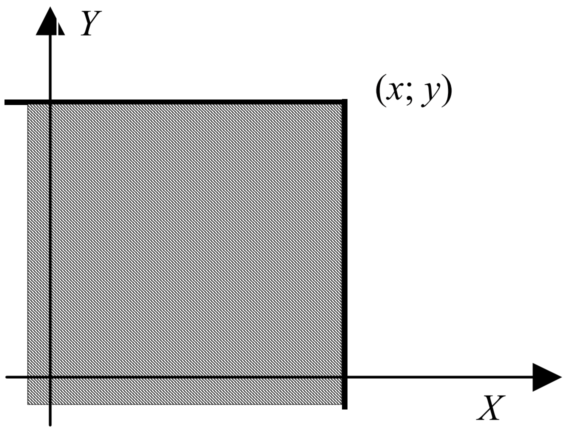


Рис. 21

Функція розподілу має такі властивості:

Властивість 1. Значення функції розподілу задовольняє подвійній нерівності $0 \le F(x, y) \le 1$.

Властивість 2. Функція розподілу є неспадна функція за кожним аргументом:

$$F(x_2; y) \ge F(x_1; y), \text{ якщо } x_2 > x_1,$$

$$F(x; y_2) \ge F(x; y_1), \text{ якщо } y_2 > y_1.$$

Властивість 3. Мають місце граничні співвідношення:

$$F(-\infty; y) = 0, \; F(x; -\infty) = 0, \; F(-\infty; -\infty) = 0, \; F(\infty; \infty) = 1.$$

Властивість 4. а) Якщо $y = \infty$ функція розподілу системи стає функцією розподілу складової *X*:

$$F(x; \infty) = F_1(x).$$

б) Якщо $x = \infty$ функція розподілу системи стає функцією розподілу складової *Y*:

$$F(\infty; y) = F_2(y).$$

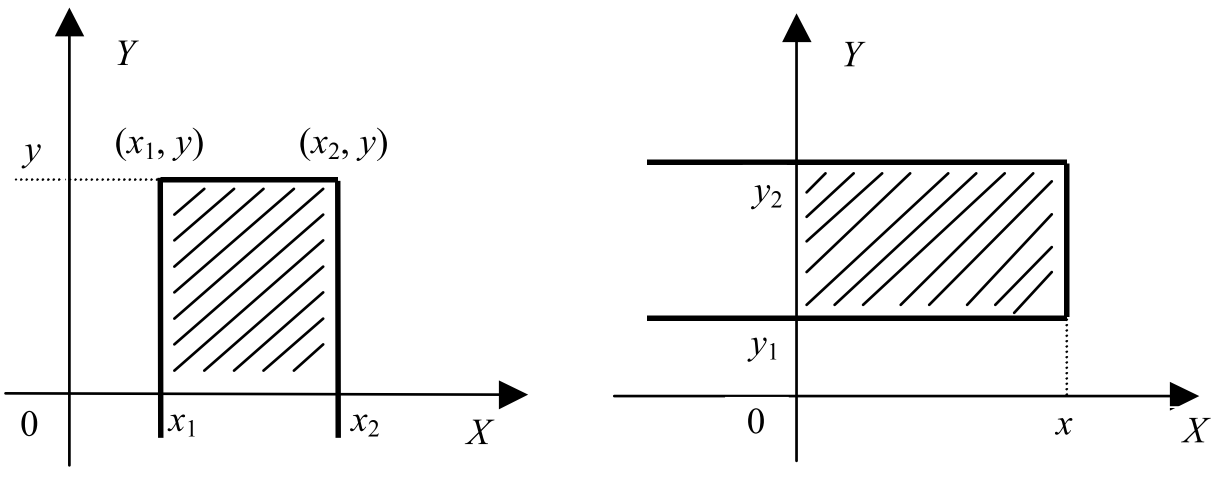
Використовуючи функцію розподілу, можливо знайти ймовірність попадання випадкової величини в напівсмугу (рис. 22а):

$$P(x_1 \le X < x_2, Y < y) = F(x_2, y) - F(x_1, y),$$

або (рис. 22б)

$$P(X < x, y_1 \le Y < y_2) = [F(x, y_2) - F(x, y_1)], \qquad (2)$$

що дорівнює приросту функції розподілу за одним із аргументів.



а) б)

Рис. 22

Аналогічно попадання випадкової величини в прямокутник $x_1 \le X < x_2$, $y_1 \le Y < y_2$ визначається за формулою (рис. 23):

$$P(x_1 \le X < x_2, y_1 \le Y < y_2) = [F(x_2, y_2) - F(x_1, y_2)] - [F(x_2, y_1) - F(x_1, y_1)]. \quad (3)$$

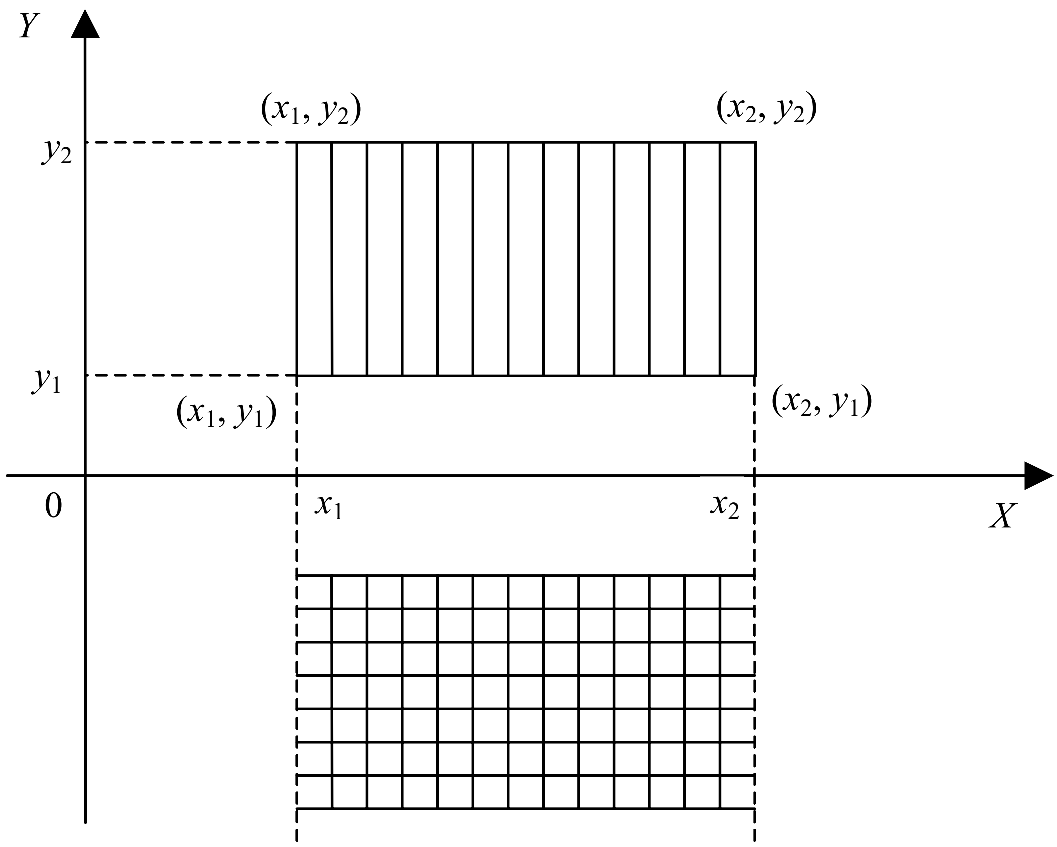


Рис. 23

Приклад 3. Знайти ймовірність попадання випадкової величини (X, Y) в прямокутник, обмежений прямими $x = 0$, $x=\frac{\pi}{6}, y=\frac{\pi}{3}, y=\frac{\pi}{2}$, якщо відома функція розподілу:

$$F(x,y)=\begin{cases}\sin x\cdot\sin y, & \textit{якщо}\quad 0\le x\le \pi/2,\quad 0\le y\le \pi/2\\ 0, & \textit{якщо}\quad x<0, \textit{або}\quad y<0\end{cases}.$$

Покладемо $x_1 = 0$, $x_2=\frac{\pi}{6}, y_1=\frac{\pi}{3}, y_2=\frac{\pi}{2}$, з формули (3) одержимо:

$$P\left(0<X<\frac{\pi}{6};\frac{\pi}{3}<Y<\frac{\pi}{2}\right)=\left[F\left(\frac{\pi}{6};\frac{\pi}{2}\right)-F\left(0;\frac{\pi}{2}\right)\right]-\left[F\left(\frac{\pi}{6};\frac{\pi}{3}\right)-F\left(0;\frac{\pi}{3}\right)\right]=$$

$$=\left[\sin\frac{\pi}{6}\cdot\sin\frac{\pi}{2}-\sin 0\cdot\sin\frac{\pi}{2}\right]-\left[\sin\frac{\pi}{6}\cdot\sin\frac{\pi}{3}-\sin 0\cdot\sin\frac{\pi}{3}\right]=$$

$$=\sin\frac{\pi}{6}\left(1-\sin\frac{\pi}{3}\right)=\frac{1}{2}\left(1-\frac{\sqrt{3}}{2}\right)=0,065.$$

Щільністю сумісного розподілу ймовірностей $f(x;y)$ неперервної двовимірної випадкової величини називають другу мішану похідну від функції розподілу:

$$f(x;y)=\frac{\partial^2 F(x;y)}{\partial x\partial y} \tag{4}$$

Щільність сумісного розподілу можливо розглядати як границю відношення ймовірності попадання випадкової точки в прямокутник зі сторонами Δx і Δy до площі цього прямокутника, якщо обидві його сторони прямують до нуля; геометрично її можна тлумачити як поверхню, яку називають *поверхнею розподілу*.

Знаючи щільність сумісного розподілу $f(x, y)$ можливо знайти функцію розподілу $F(x, y)$ за формулою:

$$F(x;y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f(x;y)\,dxdy. \tag{5}$$

Ймовірність попадання випадкової точки (X, Y) в область D визначається рівністю:

$$P\left[(X;Y)\subset D\right]=\iint\limits_{D} f(x;y)\,dxdy. \quad (6)$$

Функція щільності ймовірності має такі властивості:

Властивість 1. *Двовимірна щільність ймовірності невід'ємна*:

$$f(x;y)\geq 0\,.$$

Властивість 2. *Подвійний невласний інтеграл з нескінченними межами від двовимірної щільності ймовірності дорівнює одиниці:*

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x;y)\,dxdy=1. \quad (7)$$

Зокрема, якщо всі можливі значення (X, Y) належать скінченній області D, то остання умова матиме вигляд:

$$\iint\limits_{D} f\left(x;y\right)dxdy=1\,. \quad (8)$$

Приклад 4. Задана функція розподілу двовимірної випадкової величини:

$$F(x;y)=\begin{cases}(1-e^{-4x})\,(1-e^{-2y}), \text{якщо } x>0 \quad y>0\\ \quad 0, \qquad\qquad \text{якщо } x<0 \quad y<0\end{cases}\cdot$$

Знайти двовимірну щільність ймовірності системи (X, Y).

Використовуючи формулу (4), знайдемо частинні похідні:

$$\frac{\partial F}{\partial x}=4e^{-4x}\left(1-e^{-2y}\right),\ \frac{\partial^2 F}{\partial x\partial y}=8e^{-4x-2y}\,.$$

Отже, шукана двовимірна щільність ймовірності:

$$f\left(x;y\right)=\begin{cases}8e^{-4x-2y}, & \text{якщо} \quad x>0, \quad y>0\\ \quad 0, & \text{якщо} \quad x<0, \quad y<0\end{cases}.$$

Рекомендуємо читачу для перевірки переконатися, що

$$8\int\limits_{0}^{\infty}\int\limits_{0}^{\infty} e^{-4x-2y}dxdy=1\cdot$$

Приклад 5. Задана двовимірна щільність ймовірності системи випадкових величин (X, Y):

$$f(x;y)=\frac{20}{\pi^2\left(16+x^2\right)\left(25+y^2\right)}.$$

Знайти функцію розподілу системи.

Скористаємося формулою (5):

$$F(x;y)=\frac{20}{\pi^2}\int_{-\infty}^{y}\int_{-\infty}^{x}\frac{1}{\left(16+x^2\right)\left(25+y^2\right)}dxdy=\frac{20}{\pi^2}\int_{-\infty}^{y}\left(\frac{1}{25+y^2}\int_{-\infty}^{x}\frac{dx}{16+x^2}\right)dy=$$

$$=\frac{20}{\pi^2}\int_{-\infty}^{y}\frac{1}{25+y^2}\left(\frac{1}{4}arctg\frac{x}{4}+\frac{1}{4}\cdot\frac{\pi}{2}\right)dy=\frac{20}{\pi^2}\left(\frac{1}{4}arctg\frac{x}{4}+\frac{\pi}{8}\right)\int_{-\infty}^{y}\frac{dy}{25+y^2}=$$

$$=\frac{20}{\pi}\left(\frac{1}{4\pi}arctg\frac{x}{4}+\frac{1}{8}\right)\left(\frac{1}{5}arctg\frac{y}{5}+\frac{\pi}{10}\right)=\left(\frac{1}{\pi}arctg\frac{x}{4}+\frac{1}{2}\right)\left(\frac{1}{\pi}arctg\frac{y}{5}+\frac{1}{2}\right)$$

Приклад 6. Знайти ймовірність попадання випадкової точки в прямокутник (рис. 24) з вершинами $A(4;\ 0)$, $B(4\sqrt{3};0)$, $C(4;\ 5)$ і $D(4\sqrt{3};5)$, якщо щільність розподілу $f(x,\ y)$ задана як у попередньому прикладі.

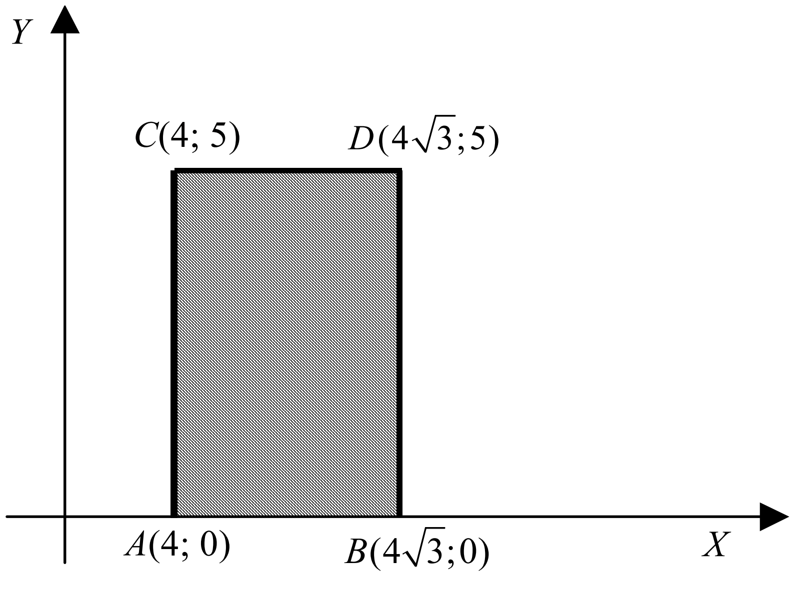


Рис. 24

Скориставшись формулою (6) маємо:

$$P[X;Y)\subset D]=\iint\limits_D \frac{20\,dx\,dy}{\pi^2(16+x^2)(25+y^2)}=\frac{20}{\pi^2}\int_0^5\left[\frac{1}{25+y^2}\int_4^{4\sqrt{3}}\frac{dx}{16+x^2}dy\right]=$$

$$=\frac{20}{\pi^2}\cdot\frac{1}{4}arctg\frac{x}{4}\Big|_4^{4\sqrt{3}}\int_0^5\frac{dy}{25+y^2}=\frac{20}{\pi^2}\cdot\frac{1}{4}\left(\frac{\pi}{3}-\frac{\pi}{4}\right)\cdot\frac{1}{5}arctg\frac{y}{5}\Big|_0^5=\frac{1}{\pi^2}\cdot\frac{\pi}{12}\cdot\frac{\pi}{4}=\frac{1}{48}$$

Приклад 7. Задана двовимірна щільність ймовірності $f(x,\ y)=$ $=\dfrac{C}{(9+x^2)(16+y^2)}$ системи (X, Y) двох випадкових величин. Знайти сталу C.

Скориставшись другою властивістю двовимірної щільності ймовірності (7), маємо:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{C}{\left(9+x^2\right)\left(16+y^2\right)}dxdy=1\ .$$

Звідси:

$$C\int_{-\infty}^{\infty}\left(\frac{1}{16+y^2}\int_{-\infty}^{\infty}\frac{dx}{9+x^2}\right)dy=1,\ C\cdot\frac{1}{3}arctg\frac{x}{3}\Big|_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{dy}{16+y^2}=1\,.$$

$$C\frac{1}{3}\left(\frac{\pi}{2}+\frac{\pi}{2}\right)\cdot\frac{1}{4}arctg\frac{y}{4}\Big|_{-\infty}^{\infty}=1;$$

$$C\cdot\frac{\pi}{3}\cdot\frac{1}{4}\left(\frac{\pi}{2}+\frac{\pi}{2}\right)=1;\ \ C=\frac{12}{\pi^2}\,.$$

Числові характеристики системи двох випадкових величин. Математичне сподівання дискретних випадкових величин X та Y, які входять в систему, визначаються за формулами:

$$m_x=M(X)=\sum_{i=1}^{n}\sum_{j=1}^{m}x_i p_{ij}\ ,\ m_y=M(Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}y_j p_{ij}\ , \tag{9}$$

а математичні сподівання неперервних випадкових величин за формулами:

$$m_x = M(X) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} xf(x;y)dxdy,\ m_y = M(Y) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} yf(x;y)dxdy \quad (10)$$

де: $f(x, y)$ — відома функція щільності ймовірності. Точка $(m_x; m_y)$ називається *центром розсіювання* системи випадкових величин (X, Y).

Математичне сподівання m_x та m_y можливо знайти простіше, якщо величини X та Y незалежні. У цьому випадку із законів розподілу цих випадкових величин X та Y можливо визначити математичні сподівання m_x та m_x за формулами, наведеними в § 5:

$$M(X) = \int\limits_{-\infty}^{\infty} xf_1(x)dx,\ M(Y) = \int\limits_{-\infty}^{\infty} yf_2(y)dy, \quad (11)$$

де: $f_1(x)$, $f_2(y)$ — щільності розподілу складових X та Y неперервної випадкової величини (X, Y).

Дисперсії дискретних випадкових величин X та Y знаходяться за формулами:

$$D(X) = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}(x_i - m_x)^2,\quad D(Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}(y_i - m_y)^2. \quad (12)$$

Дисперсії неперервних випадкових величин X та Y, які входять до системи, знаходяться за формулами:

$$\begin{gathered} D(X) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x - m_x)^2 f(x,y)dxdy = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x^2 f(x,y)dxdy - [M(X)]^2 \\ D(Y) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(y - m_y)^2 f(x,y)dxdy = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} y^2 f(x,y)dxdy - [M(Y)]^2 \end{gathered} \quad (13)$$

або

$$\begin{gathered} D(X) = \int\limits_{-\infty}^{\infty}[x - M(X)]^2 f_1(x)dx = \int\limits_{-\infty}^{\infty} x^2 f_1(x)dx - [M(X)]^2 \\ D(Y) = \int\limits_{-\infty}^{\infty}[y - M(Y)]^2 f_2(y)dy = \int\limits_{-\infty}^{\infty} y^2 f_2(y)dy - [M(Y)]^2 \end{gathered} \quad (14)$$

Подвійні інтеграли беруться за областю можливих значень системи.

Середнє квадратичне відхилення випадкових величин X та Y визначаються за формулами:

$$\sigma_x = \sqrt{D(X)}\,, \quad \sigma_y = \sqrt{D(Y)}\,. \qquad (15)$$

Приклад 8. Задана щільність сумісного розподілу двовимірної випадкової величини (X, Y):

$$f(x,y) = \begin{cases} xye^{-\left(x^2+y^2\right)}, & \left(x>0, \quad y>0\right) \\ 0, & \left(x<0, \text{або} \quad y<0\right) \end{cases}.$$

Знайти математичне сподівання і дисперсію складових X та Y.

Знайдемо спочатку щільність розподілу складової X:

$$f_1(x) = \int_0^\infty f(x,y)dy = \int_0^\infty xye^{-\left(x^2+y^2\right)}dy = xe^{-x^2}\int_0^\infty ye^{-y^2}dy = \frac{1}{2}xe^{-x^2} \quad (x>0).$$

Аналогічно отримаємо:

$$f_2(y) = \int_0^\infty f(xy)dx = \int_0^\infty xye^{-\left(x^2+y^2\right)}dx = \frac{1}{2}ye^{-y^2}\ , \ (y>0).$$

Математичне сподівання складової X:

$$M(X) = \int_0^\infty xf_1(x)dx = \int_0^\infty x\cdot\left(\frac{1}{2}xe^{-x^2}dx\right) = \begin{vmatrix} u = x & dv = xe^{-x^2}dx \\ du = dx & v = -\frac{1}{2}e^{-x^2} \end{vmatrix} =$$

$$= -\frac{1}{4}xe^{-x^2}\Big|_0^\infty + \frac{1}{4}\int_0^\infty e^{-x^2}dx = \frac{1}{4}\cdot\frac{\sqrt{\pi}}{2} = \frac{\sqrt{\pi}}{8},$$

оскільки інтеграл Пуассона $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$. Отже, $M(X) = \frac{\sqrt{\pi}}{8}$.

Очевидно, що $M(Y) = \frac{\sqrt{\pi}}{8}$.

Знайдемо дисперсію X:

$$D(X) = \int_0^\infty x^2 f_1(x)dx - [M(X)]^2 = \int_0^\infty x^2\left(\frac{1}{2}xe^{-x^2}\right)dx - \left(\frac{\sqrt{\pi}}{8}\right)^2 = \frac{1}{4} - \frac{\pi}{64}.$$

Отже, $D(X) = \frac{16-\pi}{64}$. Очевидно, що $D(Y) = \frac{16-\pi}{64}$.

Приклад 9. У двох банківських сейфах знаходяться по шість злитків золота. У першому сейфі: один злиток з №1, два — з №2, три– з №3; у другому сейфі: два злитка з №1, три — з №2, один — з №3. Нехай X та Y – номери злитків, які вийнято з першого та другого сейфа відповідно. З кожного сейфу вийняли по одному злитку. Записати закон розподілу випадкових величин (X, Y) та закони розподілу випадкових величин X та Y. Знайти математичне сподівання $M(X)$, $M(Y)$ та дисперсію $D(X)$, $D(Y)$.

Випадкова величина (X, Y) може приймати значення:

(1; 1) — 1×2 = 2 рази; (1; 2) — 1×3 = 3 рази; (1; 3) — 1×1 = 1 раз;

(2; 1) — 2×2 = 4 рази; (2; 2) — 2×3 = 6 разів; (2; 3) — 2×1 = 2 рази;

(3; 1) — 3×2 = 6 разів; (3; 2) — 3×3 = 9 разів; (3; 3) — 3×1 = 3 рази.

Всього таких значень 6×6 = 36. Позначимо $p_{ij} = P(X = x_i,\ Y = y_j)$, ($i = 1, 2, 3; j = 1, 2, 3;$).

Подія $\{X = 1,\ Y = 1\}$ відбувається тоді, коли з першого сейфа вийняли злиток №1, з другого — злиток №1, тоді $p_{11} = P(X = 1, Y = 1) =$ $= \frac{2}{36} = \frac{1}{18}$. Подія $\{X = 1,\ Y = 2\}$ відбувається тоді, коли з першого сейфа вийняли злиток №1, з другого — злиток №2, тоді $p_{12} = (X = 1,\ Y = 2) =$ $= \frac{3}{36} = \frac{1}{12}$. Аналогічно одержимо: $p_{13} = \frac{1}{36}$, $p_{21} = \frac{4}{36} = \frac{1}{9}$, $p_{22} = \frac{6}{36} = \frac{1}{6}$, $p_{23} = \frac{2}{36} = \frac{1}{6}$, $p_{31} = \frac{6}{36} = \frac{1}{6}$, $p_{32} = \frac{9}{36} = \frac{1}{4}$, $p_{33} = \frac{3}{36} = \frac{1}{12}$.

Закон розподілу випадкового вектора (X, Y) має вигляд:

X \ Y	1	2	3
1	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{36}$
2	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{1}{18}$
3	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{12}$

Використовуючи формули:

$$P(X = x_i) = \sum_{j=1}^{3} p_{ij} = p_i, \quad P(Y = y_i) = \sum_{j=1}^{3} p_{ij} = q_i,$$

можна визначити ймовірності:

$p_i = P(X = x_i)$:

$$p_1 = \frac{1}{18} + \frac{1}{12} + \frac{1}{36} = \frac{1}{6}, \quad p_2 = \frac{1}{9} + \frac{1}{6} + \frac{1}{18} = \frac{1}{3}, \quad p_3 = \frac{1}{6} + \frac{1}{4} + \frac{1}{12} = \frac{1}{2}$$

$q_j = P(Y = y_j)$:

$$q_1 = \frac{1}{18} + \frac{1}{9} + \frac{1}{6} = \frac{1}{3}, \quad q_2 = \frac{1}{12} + \frac{1}{6} + \frac{1}{4} = \frac{1}{2}, \quad q_3 = \frac{1}{36} + \frac{1}{18} + \frac{1}{12} = \frac{1}{6}.$$

Отже, закони розподілу випадкових величин X та Y мають вигляд:

X	1	2	3
p_i	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

Y	1	2	3
q_j	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

Математичні сподівання $M(X)$ та $M(Y)$ згідно з формулами (9) мають вигляд:

$$M(X) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i p_{ij} = 1\cdot\frac{1}{18} + 2\cdot\frac{1}{9} + 3\cdot\frac{1}{6} + 1\cdot\frac{1}{12} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{4} + 1\cdot\frac{1}{36} + 2\cdot\frac{1}{18} + 3\cdot\frac{1}{12} = \frac{7}{3};$$

$$M(Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} y_j p_{ij} = 1\cdot\frac{1}{18} + 2\cdot\frac{1}{12} + 3\cdot\frac{1}{36} + 1\cdot\frac{1}{9} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{18} + 1\cdot\frac{1}{6} + 2\cdot\frac{1}{4} + 3\cdot\frac{1}{12} = \frac{11}{6}$$

Точка $\left(\frac{7}{3}; \frac{11}{6}\right)$ — центр розсіювання для випадкового вектора (X, Y). Для знаходження дисперсії випадкових величин X та Y від системи величин (X, Y) перейдемо до системи величин $(X - m_x, Y - m_y)$ і складемо таблицю:

$X-\frac{3}{7}$ \ $Y-\frac{11}{6}$	$-\frac{5}{6}$	$\frac{1}{6}$	$\frac{7}{6}$
$-\frac{4}{3}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{36}$
$-\frac{1}{3}$	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{1}{18}$
$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{12}$

$$D(X)=\left(-\frac{4}{3}\right)^2\cdot\frac{1}{18}+\left(-\frac{1}{3}\right)^2\cdot\frac{1}{9}+\left(\frac{2}{3}\right)^2\cdot\frac{1}{6}+\left(-\frac{4}{3}\right)^2\cdot\frac{1}{12}+\left(-\frac{1}{3}\right)^2\cdot\frac{1}{6}+\left(\frac{2}{3}\right)^2\cdot\frac{1}{4}+$$
$$+\left(-\frac{4}{3}\right)^2\cdot\frac{1}{36}+\left(-\frac{1}{3}\right)^2\cdot\frac{1}{18}+\left(\frac{2}{3}\right)^2\cdot\frac{1}{12}=\frac{5}{9}.$$

$$D(Y)=\left(-\frac{5}{6}\right)^2\cdot\frac{1}{18}+\left(-\frac{5}{6}\right)^2\cdot\frac{1}{9}+\left(-\frac{5}{6}\right)^2\cdot\frac{1}{6}+\left(\frac{1}{6}\right)^2\cdot\frac{1}{12}+\left(\frac{1}{6}\right)^2\cdot\frac{1}{6}+\left(\frac{1}{6}\right)^2\cdot\frac{1}{4}+\left(\frac{7}{6}\right)^2\cdot\frac{1}{36}+$$
$$+\left(\frac{7}{6}\right)^2\cdot\frac{1}{18}+\left(\frac{7}{6}\right)^2\cdot\frac{1}{12}=\frac{17}{36}.$$

Звідси: $\sigma_x=\frac{\sqrt{5}}{3},\quad \sigma_y=\frac{\sqrt{17}}{6}.$

Кореляційний метод. Коефіцієнт кореляції. Важливу роль у теорії систем випадкових величин має так званий кореляційний момент, який вводиться для характеристики зв'язку між системою двох випадкових величин (X, Y).

Кореляційним моментом μ_{xy} випадкових величин X та Y називають математичне сподівання добутку відхилень цих величин:

$$\mu_{xy}=M\{[X-M(X)][Y-M(Y)]\}. \qquad (16)$$

Для дискретних випадкових величин кореляційний момент знаходиться за формулою:

$$\mu_{xy} = \sum_{i=1}^{n}\sum_{j=1}^{m} [x_i - M(X)][y_i - M(Y)]p(x_i, y_j), \tag{17}$$

а для неперервної - за формулою:

$$\mu_{xy} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[x - M(X)][y - M(Y)]f(x, y)dxdy. \tag{18}$$

Легко переконатися, що кореляційний момент можливо записати у вигляді:

$$\mu_{xy} = M(XY) - M(X)\cdot M(Y). \tag{19}$$

Тут

$$M(XY) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j p_{ij} \tag{20}$$

для дискретних випадкових величин X та Y і

$$M(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)dxdy \tag{21}$$

для неперервних величин.

Випадкові величини X та Y називаються *незалежними*, якщо ймовірність однієї із них набути значення, яке лежить у будь-якому проміжку області її значень, не залежить від того, які значення прийняла інша величина. У цьому випадку:

1. Функція розподілу системи (X, Y) дорівнює добутку функцій розподілу складових:

$$F(x, y) = F_1(x)\cdot F_2(y).$$

2. Щільність сумісного розподілу системи (X, Y) дорівнює добутку щільності розподілу складових:

$$f(x, y) = f_1(x)\cdot f_2(y).$$

3. Математичне сподівання системи (X, Y) дорівнює добутку математичних сподівань складових:

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

4. Згідно з (19) $\mu_{xy} = 0$ або з урахуванням (16) маємо:

$$\mu_{xy} = M[X - M(X)]\, M[Y - M(Y)] = 0.$$

Легко переконатися, що величина кореляційного моменту залежить від одиниць виміру випадкових величин, а тому це є недоліком цієї числової характеристики, оскільки при порівнянні кореляційних моментів різних систем випадкових величин виникають певні труднощі. А тому вводять нову безрозмірну числову величину — коефіцієнт кореляції.

Коефіцієнтом кореляції r_{xy} випадкових величин X та Y називають відношення кореляційного моменту до добутку середніх квадратичних відхилень цих величин:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y} \tag{22}$$

Оскільки розмірність μ_{xy} дорівнює добутку розмірностей величин X та Y, σ_x, σ_y мають розмірності відповідних величин X та Y, то r_{xy} — безрозмірна величина. У цьому полягає перевага коефіцієнту кореляції перед кореляційним моментом.

Очевидно коефіцієнт кореляції незалежних випадкових величин дорівнює нулю (оскільки $\mu_{xy} = 0$).

Отже, коефіцієнт кореляції — безрозмірна величина, причому $|r_{xy}| \le 1$, з іншого боку, абсолютна величина кореляційного моменту двох випадкових величин X та Y не перевищує середнє геометричного їх дисперсій:

$$|\mu_{xy}| \le \sqrt{D_x D_y}\ .$$

Коефіцієнт кореляції є оцінкою тісноти лінійного зв'язку між X та Y: чим ближче абсолютна величина коефіцієнта кореляції до одиниці, тим зв'язок сильніший; чим ближче абсолютна величина коефіцієнта кореляції до нуля, тим зв'язок слабкіший.

Корельованими називають дві випадкові величини, якщо їх $\mu_{xy} \ne 0$.

Некорельованими називають дві випадкові величини, якщо їх $\mu_{xy} = 0$.

Дві корельовані величини також і залежні; якщо дві величини залежні, то вони можуть бути як корельованими так і некорельованими. Із незалежності двох величин випливає їх некорельованість, але із некорельованості ще не можна зробити висновок про незалежність цих величин. Однак можна переконатися, що для нормально розподілених складових двовимірної випадкової величини поняття незалежності і некорельованості рівнозначні.

Неважко також переконатися, що для неперервних величин X та Y кореляційний момент можна також і обчислити за формулою:

$$\mu_{xy} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy - M(X)M(Y). \tag{23}$$

Приклад 10. Знайти коефіцієнт кореляції за умовою задачі 9.

Скористаємося таблицею розподілу системи ($X - m_x$, $Y - m_y$) випадкових величин.

За формулою (17) обчислимо спочатку кореляційний момент:

$$\mu_{xy} = \left(-\frac{4}{3}\right)\cdot\left(-\frac{5}{6}\right)\cdot\frac{1}{18} + \left(-\frac{4}{3}\right)\cdot\frac{1}{6}\cdot\frac{1}{12} + \left(-\frac{4}{3}\right)\cdot\frac{7}{6}\cdot\frac{1}{36} + \left(-\frac{1}{3}\right)\cdot\left(-\frac{5}{6}\right)\cdot\frac{1}{9} +$$

$$+\left(-\frac{1}{3}\right)\cdot\frac{1}{6}\cdot\frac{1}{6} + \left(-\frac{1}{3}\right)\cdot\frac{7}{6}\cdot\frac{1}{18} + \frac{2}{3}\left(-\frac{5}{6}\right)\cdot\frac{1}{6} + \frac{2}{3}\cdot\frac{1}{6}\cdot\frac{1}{4} + \frac{2}{3}\cdot\frac{7}{6}\cdot\frac{1}{12} = 0\,.$$

Оскільки $\mu_{xy} = 0$, то і коефіцієнт кореляції $r_{xy} = 0$.

ЗАДАЧІ

1.54. Задано розподіл ймовірностей дискретної двовимірної випадкової величини:

Y \ X	2	3	11	15
3,4	0,07	0,13	0,09	0,03
4,2	0,06	0,20	0,14	0,28

Знайти закон розподілу складових X та Y.

Відповідь:

X	2	3	11	15
P	0,13	0,33	0,23	0,31

Y	3,4	4,2
P	0,32	0,68

1.55. Знайти ймовірність того, що в результаті випробувань складова X двовимірної випадкової величини прийме значення $X < 2$ і при цьому складова Y прийме значення $Y < 3$, якщо відома функція розподілу системи.

$$F(x,y)=\left(\frac{1}{\pi}arctg\frac{x}{2}+\frac{1}{2}\right)\left(\frac{1}{\pi}arctg\frac{y}{3}+\frac{1}{2}\right).$$

Відповідь: $P(X<2,Y<3)=\dfrac{9}{16}$.

1.56. Задана функція розподілу двовимірної випадкової величини:

$$F(x,y)=\begin{cases}1-2^{-x}-2^{-y}+2^{-x-y} & якщо\ x\ge 0\ ,\ \ y\ge 0\\ 0 & якщо\ x<0\ або\ y<0\end{cases}$$

Знайти ймовірність попадання випадкової точки (X, Y) у прямокутник, обмежений лініями $x=1$, $x=2$, $y=3$, $y=5$ та двовимірну щільність ймовірності системи.

Відповідь: $P=\dfrac{3}{128}$.

$$f(x,y)=\begin{cases}\ln^2 2\cdot 2^{-x-y} & якщо\quad x\ge 0\quad ,\quad y\ge 0\\ 0 & якщо\quad x<0\quad або\quad y<0.\end{cases}$$

1.57. Щільність розподілу системи двох випадкових величин (X, Y) дорівнює:

$$f(x,y)=\frac{C}{\pi^2\left(3+x^2\right)\left(1+y^2\right)}.$$

Знайти:

а) величину C;

б) щільність розподілу складових X та Y, системи двох випадкових величин ($X;Y$);

в) функцію розподілу $F(x,y)$;

г) ймовірність попадання випадкової точки ($X;Y$) у квадрат, обмежений прямими $x=0$, $y=0$, $x=1$, $y=1$.

Відповідь: а) $\sqrt{3}$;

б) $f_1(x)=\sqrt{3}\cdot\dfrac{1}{\pi\left(3+x^2\right)}$, $f_2(y)=\dfrac{1}{\pi\left(3+y^2\right)}$;

в) $F(x,y)=\left(\dfrac{1}{\pi}arctg\dfrac{x}{\sqrt{3}}+\dfrac{1}{2}\right)\cdot\left(\dfrac{1}{\pi}arctgy+\dfrac{1}{2}\right)$;

г) $\dfrac{1}{24}$.

1.58 Задана щільність сумісного розподілу двовимірної випадкової величини (*X;Y*):

$$f(x,y)=\begin{cases}C\sin(x+y) & \text{в області А}\\ 0 & \text{за цією областю}\end{cases}.$$

Область А визначається нерівностями $0\le x\le\dfrac{\pi}{2}$, $0\le y\le\dfrac{\pi}{2}$.

Знайти: а) коефіцієнт *C;*
б) математичне сподівання *M*(*X*), *M*(*Y*);
в) середнє квадратичне відхилення σ_x та σ_y;
г) кореляційний момент μ_{xy};
д) коефіцієнт кореляції r_{xy}.

Відповідь: а) $C=\dfrac{1}{2}$;

б) $M(X)=M(X)=\dfrac{\pi}{4}$;

в) $\sigma_x^2=\sigma_y^2=\sigma_x\sigma_y=\dfrac{\pi^2}{16}+\dfrac{\pi}{2}-2;$;

г) $\mu_{xy}=\dfrac{8\pi-16-\pi^2}{16}$;

д) $r_{xy}\approx-0.2454$.

1.58. Задана таблиця, яка визначає закон розподілу системи двох випадкових величин (*X*, *Y*):

X \ *Y*	20	40	60
10	$\frac{3}{20}$	$\frac{1}{20}$	0
20	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{10}$
30	$\frac{1}{20}$	$\frac{1}{10}$	$\frac{1}{4}$

Знайти: а) математичне сподівання $M(X)$ та $M(Y)$;
б) дисперсії $D(X)$ та $D(Y)$;
в) коефіцієнт кореляції r_{xy}.

Відповідь: а) $M(X) = 22$, $M(Y) = 41$;
б) $D(X) = 56$, $D(Y) = 259$;
в) $r_{xy} = 0{,}56$.

1.58. Щільність розподілу системи двох випадкових величин (X, Y) дорівнює:

$$f(x,y)=\begin{cases}24xy & \text{в області А}\\ 0 & \text{за цією областю}\end{cases},$$

область А — трикутник, обмежений прямими: $x+y-1=0$, $x=0$, $y=0$.

Знайти: а) математичне сподівання $M(X)$ та $M(Y)$;
б) дисперсії $D(X)$ та $D(Y)$;
в) коефіцієнт кореляції r_{xy}.

Відповідь: а) $M(X)=M(Y)=\dfrac{2}{5}$;
б) $D(X)=D(Y)=\dfrac{1}{25}$;
в) $r_{xy}=-\dfrac{2}{3}$.

Розділ II
МАТЕМАТИЧНА СТАТИСТИКА

§1. Основні поняття і завдання математичної статистики. Вибірковий метод.

Встановлення закономірностей, яким підпорядковані масові випадкові явища, побудовано на вивченні статистичних даних результатів спостережень.

Перше завдання математичної статистики — вказати способи збору та групування (якщо даних дуже багато) статистичних даних.

Друге завдання математичної статистики — розробка методів аналізу статистичних даних, залежно від мети дослідження.

Отже, завдання математичної статистики полягає в розробці методів збору і обробки статистичних даних для одержання наукових та практичних висновків.

Наприклад, потрібно вивчити сукупність однорідних об'єктів щодо деякої якісної або кількісної ознаки, які характеризують ці об'єкти.

Іноді проводять суцільне обстеження, тобто обстежують кожний із об'єктів сукупності щодо ознаки, якою цікавляться. На практиці суцільне обстеження застосовується порівняно рідко, частіше випадково вибирають із усієї сукупності обмежене число об'єктів і піддають їх вивченню.

Вибірковою сукупністю, або *вибіркою* називають сукупність випадково відібраних об'єктів.

Генеральною сукупністю називається сукупність об'єктів, із яких проводиться вибірка.

Об'ємом сукупності (вибіркової або генеральної) називається число об'єктів цієї сукупності (наприклад, якщо із 100 деталей відібрано для обстеження 10 деталей, то об'єм генеральної сукупності $N = 100$, а об'єм вибірки $n = 10$).

При складанні вибірки можливо поступити двояко: після того, як об'єкт відібрано і над ним проведені обстеження, він може бути повернутий, або не повернутий в генеральну сукупність. Відповідно до цього вибірки поділяються на повторні і безповторні.

Повторною називають вибірку, за якою відібраний об'єкт (перед відбором наступного) повертається в генеральну сукупність.

Безповторною називають вибірку, за якою відібраний об'єкт у генеральну сукупність не повертається.

Вибірка повинна бути репрезентативною (представницькою) щоб об’єкти вибірки вірно її представляли. За законом великих чисел можливо стверджувати, що вибірка буде репрезентативною, якщо її здійснити випадково, тобто, таким чином, щоб будь-який об’єкт генеральної сукупності мав одну і ту ж ймовірність попасти у вибірку.

На практиці застосовуються різні способи відбору, які принципово можна розділити на два види:

1. Відбирання, що не потребує розчленування генеральної сукупності на частини:

а) простий випадковий безповторний відбір;

б) простий випадковий повторний відбір (простим випадковим називається такий відбір, за яким об’єкти здобувають по одному із всієї сукупності).

2. Відбирання, за яким генеральна сукупність розбивається на частини:

а) типовий відбір (об’єкти відбираються не із всієї генеральної сукупності, а із кожної її “типової” частини);

б) механічний відбір (генеральна сукупність “механічно” ділиться на стільки груп, скільки об’єктів має увійти у вибірку і із кожної групи відбирається один об’єкт);

в) серійний відбір (об’єкти “серіями” підкоряються суцільному обстеженню).

Статистичний розподіл вибірки. Нехай над випадковою кількісною ознакою X генеральної сукупності здійснено “n” незалежних спостережень і при цьому величина X прийняла n_1 раз значення x_1, n_2 раз — x_2,..., n_k раз — x_k причому $n_1 + n_2 + ... + n_k = n$ (одержана вибірка об’єму “n”). Складемо таблицю (2.1).

Таблиця 2.1

X	x_1	x_1	...	x_k
W	w_1	w_2	...	w_k

Перший рядок називається *варіаційним рядом* (значення записані по порядку їх зростання).

$w_i = \dfrac{n_i}{n}$ — відносна частота, x_i — варіанта, n_i — частота.

Ця таблиця називається статистичним розподілом вибірки випадкової величини.

До того ж

$$w_1 + w_2 + \ldots + w_k = \frac{n_1}{n} + \frac{n_2}{n} + \ldots + \frac{n_k}{n} = \frac{n_1 + n_2 + \ldots + n_k}{n} = \frac{n}{n} = 1 .$$

Приклад 1. Результати 30 спостережень над кількісною ознакою генеральної сукупності зведені в таблицю (2.2):

Таблиця 2.2

X	2	4	8	16	18
n_i	5	3	7	6	9

Знайти статистичний розподіл вибірки.

Знаходимо: $w_1 = \frac{5}{30}$; $w_2 = \frac{3}{30}$; $w_3 = \frac{7}{30}$; $w_4 = \frac{6}{30}$; $w_5 = \frac{9}{30}$.

Таблиця 2.3

X	2	4	8	16	18
W	$\frac{5}{30}$	$\frac{3}{30}$	$\frac{7}{30}$	$\frac{6}{30}$	$\frac{9}{30}$

Тут: $\sum W = 1$.

Зазначимо, що в теорії ймовірності під розподілом розуміють відповідність між можливими значеннями випадкової величини та їх ймовірностями, а в математичній статистиці — відповідність між спостережними варіантами та їх частотами, або відносними частотами.

Емпірична функція розподілу. Нехай відомий статистичний розподіл кількісної ознаки X. Введемо позначення: “n_x” — число спостережень, за яких спостерігались значення ознаки X, менше “x”; “n” — загальне число спостережень (об’єм вибірки).

Тоді: $\frac{n_x}{n}$ — відносна частота події $X < x$ (є функцією від x).

Емпіричною функцією розподілу називають функцію, яка визначає для кожного значення x відносну частоту події $X < x$, тобто:

$$F^*(x) = \frac{n_x}{n},$$

де n_x — число варіант менших за x.

Вона знаходиться емпіричним (дослідним) шляхом. Наприклад, щоб знайти $F^*(x_2)$ потрібно число спостережень, за яких спостерігалось $X < x_2$, поділити на загальне число спостережень:

$$F^*(x_2) = \frac{n_{x_2}}{n}.$$

Функція $F^*(x)$ — має всі властивості функції розподілу $F(x)$:

1) значення $F^*(x) \in [0;1]$;

2) $F^*(x)$ — не спадна функція;

3) якщо x_1– найменше значення ознаки, що спостерігалось, а x_k — найбільша варіанта, то

$$F^*(x) = 0, \text{ якщо } x \le x_1;$$
$$F^*(x) = 1, \text{ якщо } x > x_k.$$

Приклад 2. Побудувати емпіричну функцію за даними вибірки (об'єм вибірки $n = 60$):

Таблиця 2.4

X	2	6	10
n_i	12	18	30

Найменше спостережне значення ознаки дорівнює 2, отже, $F^*(x) = 0$, при $x \le 2$.

Значення $X < 6$, а значить $x_1 = 2$, спостерігалось 12 разів, отже, $F^*(x) = \dfrac{12}{60} = 0{,}2$ якщо $2 < x \le 6$.

Значення $X < 10$ спостерігалось 12 + 18 = 30 разів, отже, $F^*(x) = \dfrac{30}{60} = 0{,}5$ якщо $6 < x \le 10$.

Найбільше спостережне значення $x = 10$, отже, $F^*(x > 10) = 1$

Отже, $F^*(x) = \begin{cases} 0 \;\; \textit{якщо} \;\; x \le 2 \\ 0{,}2 \;\; \textit{якщо} \;\; 2 < x \le 6 \\ 0{,}5 \;\; \textit{якщо} \;\; 6 < x \le 10 \\ 1 \;\; \textit{якщо} \;\; x > 10 \end{cases}.$

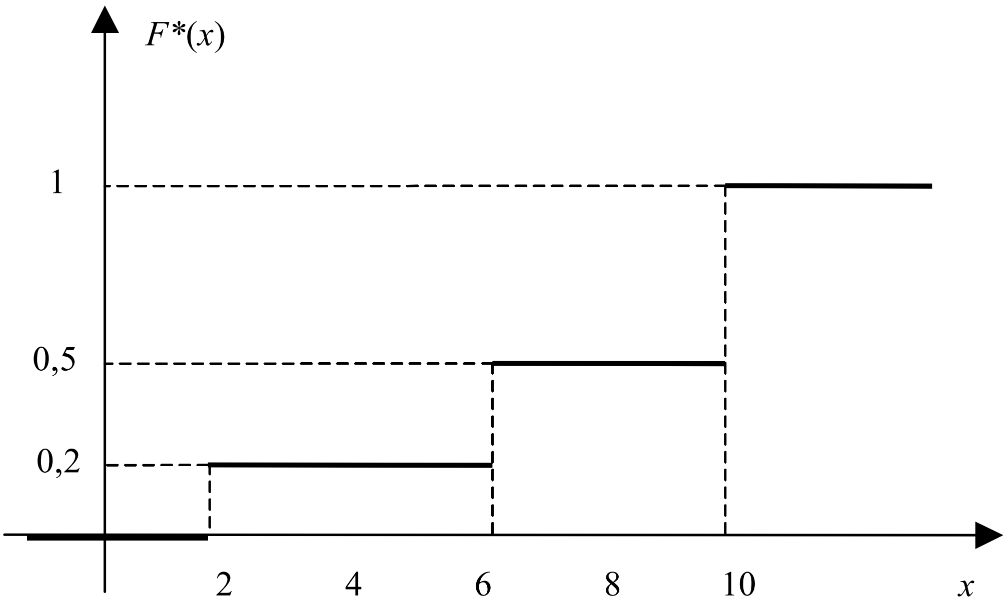


Рис. 25

З метою наочності користуються різними способами графічного зображення статистичного розподілу. До них належать полігон і гістограма.

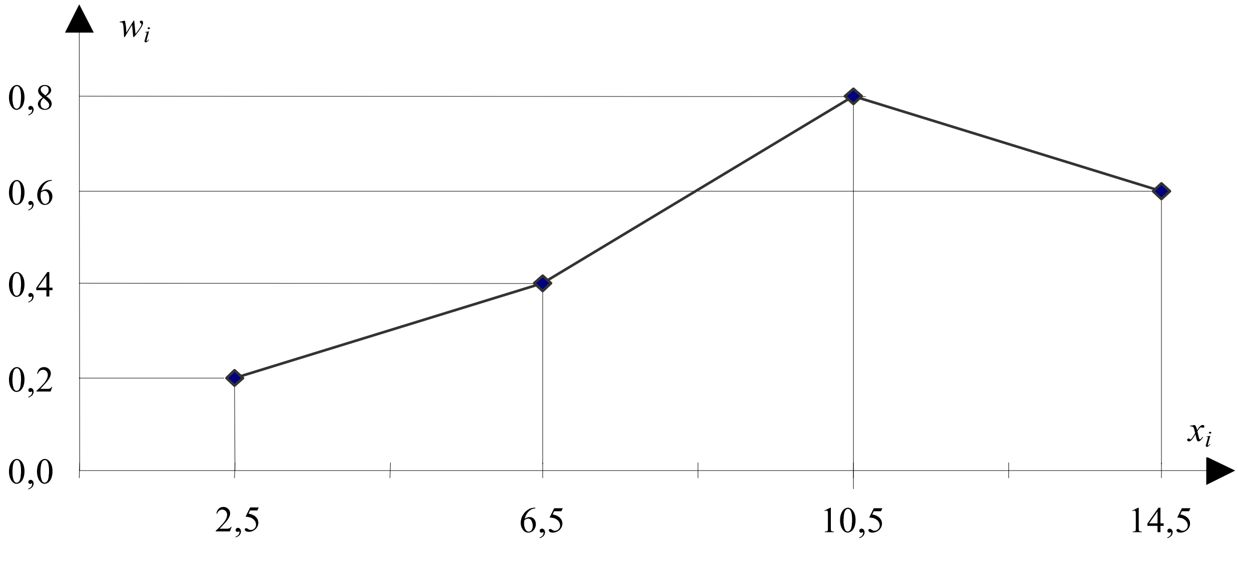


Рис. 26

Полігоном частот (відносних частот) називають ламану, відрізки якої сполучають точки $(x_1; n_1)$, $(x_2; n_2)$, ..., $(x_k; n_k)$, $((x_1; w_1)$, $(x_2; w_2)$, ..., $(x_k; w_k))$. Для побудови *полігону частот* (відносних частот) на осі абсцис відкладають варіанти x_i, а на осі ординат — відповідні їм частоти n_i (відносні частоти w_i). Точки $(x_i; n_i)$, $((x_i; w_i))$ сполучають відрізками прямих і одержують *полігон частот* (відносних частот).

На рис. 26 зображений полігон відносних частот наступного розподілу

Таблиця 2.5

X	2,5	6,5	10,5	14,5
W	0,2	0,4	0,8	0,6

При неперервному розподілі ознаки будують *гістограму* частот (відносних частот).

Гістограмою частот (відносних частот) називають ступінчасту фігуру, яка складається із прямокутників, основами яких є часткові інтервали довжиною h, а висоти рівні відношенню $\frac{n_i}{h}\left(\frac{w_i}{h}\right)$.

Для побудови гістограми частот (відносних частот) на осі абсцис відкладають часткові інтервали, а над ним проводять відрізки, паралельні осі абсцис на відстані $\frac{n_i}{h}\left(\frac{w_i}{h}\right)$. Площа i-го часткового прямокутника дорівнює $h \cdot \frac{n_i}{h} = n_i$ — сумі частот варіант i-го інтервалу ($h \cdot \frac{w_i}{h} = w_i$ — відносній частоті варіант, що попали в i-ий інтервал).

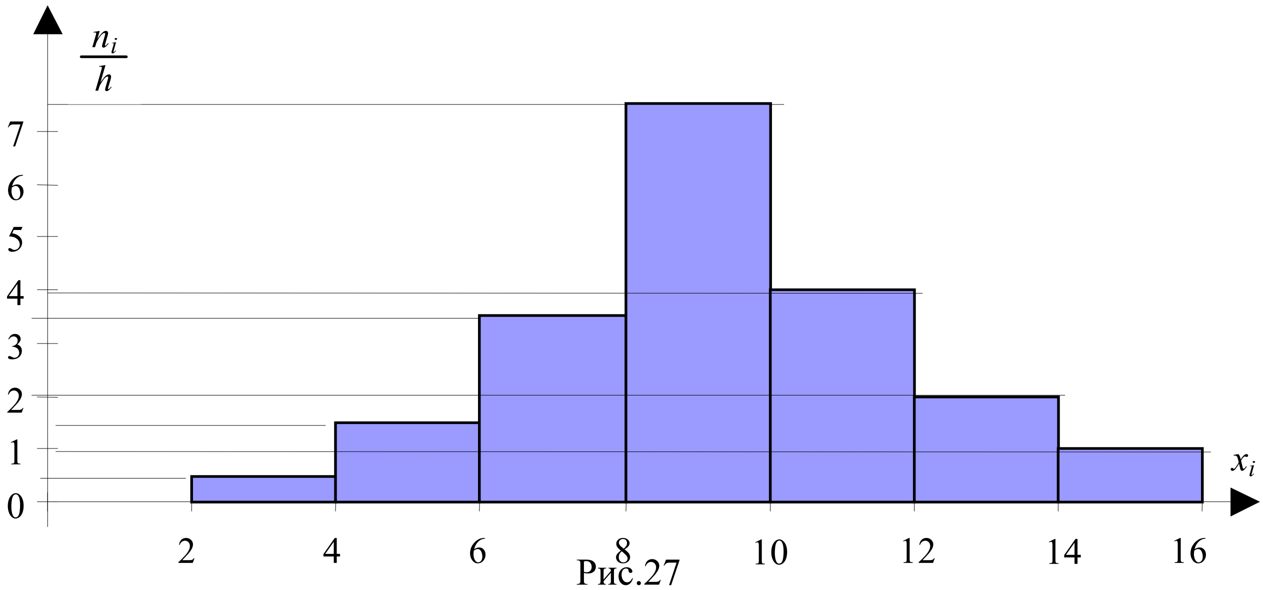


Рис.27

$F_i = h \cdot \frac{n_i}{h} = n_i$ (до рис. 27).

Отже, площа гістограми частот (відносних частот) дорівнює сумі всіх частот (відносних частот), тобто об'єму вибірки (одиниці).

На рис. 27 зображена гістограма частот розподілу об'єму $n = 40$, наведеного в таблиці (2.6):

Таблиця 2.6

Частковий інтервал $h = 2$	Сума частот варіант часткового інтервалу n_i	Щільність частот $\frac{n_i}{h}$
2 — 4	1	0,5
4 — 6	3	1,5
6 — 8	7	3,5
8 — 10	15	7,5
10 — 12	8	4,0
12 — 14	4	2,0
14 — 16	2	1

ЗАДАЧІ

2.1. Побудувати графік емпіричної функції розподілу.

Таблиця 2.7

x_i	7	9	13	16
n_i	4	5	12	9

2.2.Побудувати полігон частот і відносних частот розподілу.

Таблиця 2.8

x_i	2	4	6	8	10
n_i	15	20	40	13	12

2.3. Побудувати гістограму частот і відносних частот розподілу (в першому стовпці указаний частковий інтервал, у другому — сума частот варіант часткового інтервалу).

Таблиця 2.9

5 — 10	6
10 — 15	16
15 — 20	35
20 — 25	22

2.4. Побудувати гістограму відносних частот за заданим розподілом вибірки.

Таблиця 2.10

Номер інтервалу i	Частковий інтервал $x_i - x_{i+1}$	Сума частот варіант часткового інтервалу n_i
1	0 — 2	20
2	2 — 4	30
3	4 — 6	50

$$n = \sum n_i = 100$$

2.5. Побудувати гістограму частот за заданим розподілом вибірки об'єму $n = 100$.

Таблиця 2.11

Номер інтервалу i	Частковий інтервал $x_i - x_{i+1}$	Сума частот варіант інтервалу n_i	Щільність частоти n_i/h
1	1 — 5	10	2,5
2	5 — 9	20	5
3	9 — 13	50	12,5
4	13 — 17	12	3
5	17 — 21	8	2

§ 2.Числові характеристики статистичного розподілу

Основними числовими характеристиками випадкової величини в теорії ймовірності є математичне сподівання та дисперсія.

Аналогічні числові характеристики існують і для статистичних розподілів. Аналогією математичного сподівання випадкової величини X є середнє арифметичне спостережних значень випадкової величини, що визначається за формулою:
для генеральної сукупності

$$\bar{x}_Г = \frac{x_1 + x_2 + ...x_N}{N} \tag{1}$$

де $\bar{x}_Г$ – середнє генеральне; x_1, x_2 ,... x_N — варіанти (причому частота кожної варіанти дорівнює одиниці), N — об’єм генеральної сукупності.

Якщо значення ознаки x_1, x_2 ,... x_k мають відповідні частоти N_1, N_2,... N_k, причому $N_1 + N_2 +...+ N_k = N$, то

$$\bar{x}_Г = (x_1N_1 + x_2N_2 +...+ x_kN_k)\frac{1}{N}, \tag{2}$$

тобто генеральне середнє, є середня зважена значень ознаки з вагою, рівною відповідним частотам.

Якщо для вивчення генеральної сукупності здобута повторна вибірка об’єму *n,* то аналогічно:

$$\bar{x}_B = (x_1 + x_2 +...+ x_n)\frac{1}{n} \text{ або}$$

$$\bar{x}_B = \frac{x_1n_1 + x_2n_2 + ... + x_Kn_k}{n} = \frac{\sum_{i=1}^{k} x_in_i}{n}, \tag{3}$$

де n_1, n_2 ,..., n_k відповідні частоти значень x_1, x_2,..., x_k, причому $n_1 + n_2 +...+ n_k = n$.

Тобто вибіркова середня є середня зважена значень ознаки з вагою, рівною відповідним частотам.

Розглянемо $\bar{x}_в$ як випадкову величину $\overline{X_B}$ і x_1, x_2,..., x_n як незалежні і однаково розподілені випадкові величини X_1, X_2,..., X_n.

Оскільки ці величини однаково розподілені, то вони мають однакові числові характеристики (однакове математичне сподівання, яке позначимо через a), а також ураховуючи властивості математичного сподівання, запишемо:

$$M(\overline{X}_B) = M[(X_1 + X_2 + ... + X_n)\frac{1}{n}] = \frac{na}{n} = a. \tag{4}$$

Якщо вважатимемо, що кожна із величин X_1, X_2, ..., X_n має той самий розподіл, що і генеральна сукупність, яку також вважатимемо випадковою величиною, то числові характеристики цих величин і генеральної сукупності однакові. Зокрема, математичне сподівання ознаки X генеральної сукупності дорівнює математичному сподіванню a кожної із величин, тобто

$$M(X) = \overline{x}_Г = a. \tag{5}$$

З урахуванням (5) перепишемо формулу (4)

$$M(\overline{X_B}) = \overline{x}_Г \tag{6}$$

Отже, при збільшені об’єму вибірки n вибіркова середня прямує за ймовірністю до генеральної середньої, а це означає, що вибіркова середня є спроможною оцінкою генеральної середньої. Спроможною оцінкою називають статистичну оцінку, яка при $n \to \infty$ прямує за ймовірністю до параметра, що оцінюється.

Генеральною дисперсією $D_Г$ (вибірковою дисперсією D_B) називають середнє арифметичне квадратів відхилень спостережних значень ознаки генеральної (вибіркової) сукупності від їх середнього значення $\overline{x}_Г$ ($\overline{x}_B$).

$$D_Г = \frac{\sum_{i=1}^{N}(x_i - \overline{x}_Г)^2}{N} \text{ або } D_Г = \frac{\sum_{i=1}^{k}N_i(x_i - \overline{x}_Г)^2}{N}; \tag{7}$$

$$D_B = \frac{\sum_{i=1}^{n}(x_i - \overline{x}_B)^2}{n} \text{ або } D_B = \frac{\sum_{i=1}^{k}n_i(x_i - \overline{x}_B)^2}{n}, \tag{8}$$

де N_1, N_2 ,..., N_k (n_1, n_2 ,..., n_k) відповідні частоти значень ознаки x_1, x_2,..., x_k , причому $N_1 + N_2 +...+ N_k = N$ $(n_1 + n_2 +...+ n_k = n)$.

Крім дисперсії, для характеристики розсіювання значень кількісної ознаки X генеральної (вибіркової) сукупності навколо свого середнього значення користуються зведеною характеристикою — *генеральним (вибірковим) середнім квадратичним відхиленням:*

$$\sigma_Г = \sqrt{D_Г}\ , \text{або}\ \sigma_В = \sqrt{D_В}\ . \tag{9}$$

Приклад 1. Генеральна сукупність задана таблицею розподілу

x_i	1	3	7	10
N_i	6	7	9	4

Знайти генеральну дисперсію.

Знайдемо генеральну середню (2):

$$\overline{x}_Г = \frac{6\cdot1+7\cdot3+9\cdot7+4\cdot10}{6+7+9+4} = \frac{130}{26} = 5.$$

Знайдемо генеральну дисперсію (7)

$$D_Г = \frac{6(1-5)^2+7(3-5)^2+9(7-5)^2+4(10-5)^2}{26} = \frac{260}{26} = 10.$$

Приклад 2. Вибіркова сукупність задана таблицею розподілу

x_i	2	4	7	10
n_i	30	25	20	15

Знайти вибіркову дисперсію.

Знайдемо вибіркову середню (3):

$$\overline{x}_В = \frac{30\cdot2+25\cdot4+20\cdot7+15\cdot10}{30+25+20+15} = \frac{450}{90} = 5\ .$$

Знайдемо вибіркову дисперсію (8):

$$D_В = \frac{30\cdot(2-5)^2+25\cdot(4-5)^2+20\cdot(7-5)^2+15\cdot(10-5)^2}{90} = \frac{750}{90} \approx 8{,}3\ .$$

Зауваження. Обчислення дисперсії, байдуже — генеральної чи вибіркової — можна спростити шляхом перетворення, наприклад, формули (8). Підвівши до квадрата та вводячи відповідні позначення, отримаємо:

$$D = \overline{x^2} - \left[\overline{x}\right]^2. \tag{10}$$

Дисперсія дорівнює середньому квадратів значень ознаки мінус квадрат загальної середньої (середнє квадратів значень ознаки, яка належить всій сукупності),

де $$\overline{x} = \frac{\sum n_i x_i}{n}, \quad \overline{x^2} = \frac{\sum n_i x_i^2}{n}.$$

Приклад 3. Знайти дисперсію за заданим розподілом

x_i	2	4	7	10
n_i	30	25	20	15

Знайдемо загальну середню:

$$\overline{x} = \frac{30\cdot 2 + 25\cdot 4 + 20\cdot 7 + 15\cdot 10}{30+25+20+15} = \frac{450}{90} = 5.$$

Знайдемо середню квадратів значень ознаки:

$$\overline{x}^{\,2} = \frac{30\cdot 2^2 + 25\cdot 4^2 + 20\cdot 7^2 + 15\cdot 10^2}{30+25+20+15} = \frac{3000}{90} = 33{,}3.$$

Шукана дисперсія D = $\overline{x^2} - \left[\overline{x}\right]^2 \approx 33{,}3 - 5^2 = 8{,}3$.

Точкові оцінки параметрів. Нехай маємо випадкову величину X, закон розподілу якої містить невідомий параметр “a”. Потрібно на основі експериментальних даних знайти підходящу оцінку параметра a. Оцінка називається *точковою*, якщо вона визначається однією точкою на числовій осі, в якій має знаходиться значення невідомого параметра. Всі оцінки розглянуті раніше — точкові.

Позначимо через $X_1, X_2, ..., X_n$ спостережені значення випадкової величини в результаті проведення “n” незалежних дослідів (n різних вибірок одного об’єму). Нехай величина $\overline{a}$, обчислена на основі цих дослідів, є оцінкою параметра a, що означає, що $\overline{a}$ є функцією величин $X_1, X_2, ..., X_n$, тобто $\overline{a} = \overline{a}\left(X_1, X_2, ..., X_n\right)$.

Для того щоб оцінка $\overline{a}$ мала практичну цінність, вона повинна мати наступні властивості:

1. *Незсуненість оцінки*. Розрізняють оцінки зсунені та незсунені. Зсуненими називають оцінки, математичне сподівання яких не дорівнює оцінювальному параметру

$$M\left[\overline{a}\left(X_1, X_2, .., X_n\right)\right] \neq a\,. \tag{11}$$

Незсуненими називаються оцінки, для яких виконуються умови:

$$M\left[\overline{a}\left(X_1, X_2, .., X_n\right)\right] = a\,. \tag{12}$$

2. *Спроможність оцінки.* Оцінка $\overline{a}\left(X_1, X_2, ..., X_n\right)$ для параметра a називається *спроможною*, якщо вона збігається за ймовірністю до оцінювального параметра при необмеженому зростанні числа експериментів, тобто

$$\lim_{n\to\infty} P\left[\left|\overline{a}\left(X_1, X_2, .., X_n\right) - a\right| < \varepsilon\right] = 1 \tag{13}$$

або щоб дисперсія оцінки прямувала до нуля при $n \to \infty$:

$$\lim_{n\to\infty} D\left[\overline{a}\left(X_1, X_2, .., X_n\right)\right] = 0\,. \tag{14}$$

і, крім того, щоб оцінка була незсуненою.

Отже, спроможність оцінки означає, що при достатньо великій кількості дослідів "*n*" зі скільки завгодно великою вірогідністю, відхилення оцінки від дійсного значення параметра менше будь-якої наперед заданої величини.

3. *Ефективність оцінки.* Для того щоб оцінка була ефективною, необхідно, щоб дисперсія оцінки була мінімальною:

$$D\left[\overline{a}\left(X_1, X_2, .., X_n\right)\right] = D_{\min}\,. \tag{15}$$

Характеристики вибіркової сукупності є точковими оцінками відповідних характеристик генеральної сукупності, а тому розглянемо їх властивості:

а) середня вибірка $\overline{x_B}$ — є оцінкою середньої генеральної $\overline{x_\Gamma}$. Оцінка $\overline{x_B}$ — спроможна (оскільки зі збільшенням *n* вона збігається за ймовірністю до сталого числа — до середньої генеральної $\overline{x_\Gamma}$), незсуненим (оскільки $M\left(\overline{X_B}\right) = \overline{x_\Gamma}$) і ефективна (оскільки $D\left(\overline{X_B}\right) = \dfrac{D\left(\overline{X_\Gamma}\right)}{n}$).

б) дисперсія вибірки D_B є оцінкою дисперсії генеральної $D_Г$. Дисперсію вибірки запишемо у такому вигляді:

$$D_B = \overline{x_B^2} - \left(\overline{x_B}\right)^2.$$

Оскільки при $n \to \infty$ — $\overline{x_B^2}$ збігається за ймовірністю до $\overline{x_Г^2}$ і $\left(\overline{x_B}\right)^2$ збігається до $\left(\overline{x_Г}\right)^2$, то D_B — оцінка спроможна. Але вона є зсуненою, оскільки можна довести, що математичне сподівання вибіркової дисперсії не дорівнює оцінювальній генеральній дисперсії, а дорівнює:

$$M\left(D_B\right) = \frac{n-1}{n} D_Г. \tag{16}$$

Отже, якщо користуватися оцінкою D_B замість $D_Г$, то будемо здійснювати систематичну помилку у бік зменшення. Щоб усунути таке зсунення, вводять “виправлену” статистичну дисперсію, яку, як правило, позначають через s^2:

$$s^2 = \frac{n}{n-1} D_B \tag{17}$$

і вона вже не буде зсуненою:

$$M\left(s^2\right) = M\left[\frac{n}{n-1} D_B\right] = \frac{n}{n-1} M\left[D_B\right] = \frac{n}{n-1} \cdot \frac{n-1}{n} D_Г = D_Г.$$

Таким чином, середнє вибірки $\overline{x_B}$ і “виправлена” дисперсія s^2 є незсуненими і спроможними точковими оцінками генеральної середньої $\overline{x_Г}$ та дисперсії $D_Г$ відповідно.

Отже, із (17) маємо:

$$s^2 = \frac{n}{n-1} \frac{\sum_{i=1}^{n} n_i \left(x_i - \overline{x_B}\right)^2}{n} = \frac{\sum_{i=1}^{n} n_i \left(x_i - \overline{x_B}\right)^2}{n-1}. \tag{18}$$

Для оцінки середнього квадратичного відхилення генеральної сукупності використовують “виправлене” середнє квадратичне відхилення, що дорівнює квадратному кореню із виправленої дисперсії:

$$s = \sqrt{\frac{\sum_{i=1}^{n} n_i \left(x_i - \overline{x_B} \right)^2}{n-1}}\,. \qquad (19)$$

Наведені залежності дозволяють визначити наближені значення заміряної величини, тобто здійснити оцінку математичного сподівання випадкової величини X та її дисперсії.

Приклад 4. У результаті п’яти вимірювань деякої величини одним приладом (без систематичних помилок) одержані наступні результати (у мм): 93; 95; 104; 106; 107. Знайти:

а) вибірковий середній результат виміру;

б) вибіркову та “виправлену” дисперсії похибок приладу.

а) Знаходимо вибіркову середню:

$$\overline{x_B} = 93 + \frac{(0+2+11+13+14)}{5} = 93 + 8 = 101\,.$$

б) Знаходимо вибіркову дисперсію:

$$D_B = \frac{\sum_{i=1}^{n} \left(x_i - \overline{x_B} \right)^2}{n} =$$

$$= \frac{\left[(93-101)^2 + (95-101)^2 + (104-101)^2 + (106-101)^2 + (107-101)^2 \right]}{5} = 34.$$

Знайдемо “виправлену” дисперсію:

$$s^2 = \frac{n}{n-1} D_B = \frac{5}{4} \cdot 34 = 42{,}5\,.$$

Зауваження 1. Оскільки в цьому прикладі початкові варіанти x_i — великі числа, то для спрощення розрахунку доцільно відняти із кожної варіанти одне і те ж число C, тобто перейти до умовних варіант $u_i = x_i - C$. Тоді

$$\overline{x_B} = C + \frac{\sum_{i=1}^{n} n_i u_i}{n}\,. \qquad (20)$$

За число C вибирається будь-яке число, наприклад, у цьому прикладі $C = 93$, або число, близьке до вибіркової середньої, якщо воно відоме.

Звідси у формулі (20) легко переконатися, якщо записати $n_i u_i = n_i\left(x_i - C\right)$ і просумувати ліву і праву частини рівності за всіма значеннями i, одержимо:

$\sum n_i u_i = \sum n_i\left(x_i - C\right)$, або $\sum n_i u_i = \sum n_i x_i - C\sum n_i = \sum n_i x_i - Cn$,

тобто

$\sum n_i x_i = Cn + \sum n_i u_i$.

Отже, $\dfrac{\sum n_i x_i}{n} = C + \dfrac{\sum n_i u_i}{n}$, або $\overline{x_B} = C + \dfrac{\sum n_i u_i}{n}$.

Аналогічно зауваження можна зробити щодо обчислення інших параметрів розподілу, наприклад:

$$D_B\left(X\right) = D_B\left(U\right) = \overline{u^2} - \left[\overline{u}\right]^2 = \frac{\sum\limits_{i=1}^{n} n_i u_i^2}{n} - \left[\frac{\sum\limits_{i=1}^{n} n_i u_i}{n}\right]^2. \qquad (21)$$

Більш зручна формула, ніж (17) є:

$$s_x^2 = \frac{\sum n_i x_i^{\ 2} - \dfrac{\left[\sum n_i x_i\right]^2}{n}}{n-1} \qquad (22)$$

або в умовних варіантах вона має вигляд:

$$s_u^2 = \frac{\sum n_i u_i^{\ 2} - \dfrac{\left[\sum n_i u_i\right]^2}{n}}{n-1} \qquad (23)$$

причому, якщо $u_i = x_i - C$, то $s_x^2 = s_u^2$.

Приклад 5. Знайти вибіркову дисперсію та "виправлену" вибіркову дисперсію за заданим розподілом вибірки об'єму $n = 10$:

x_i	182	188	192
n_i	1	6	3

Варіанти порівняно великі числа, а тому перейдемо до умовних варіант $u_i = x_i - 187$ (ми вираховуємо із варіант число $C = 187$, наближене до вибіркової середньої). У результаті одержимо розподіл умовних варіант:

u_i	–5	1	+5
n_i	1	6	3

Знайдемо шукану вибіркову дисперсію за формулою (21):

$$D_B = \frac{1\cdot 5^2 + 6\cdot 1^2 + 3\cdot 5^2}{10} - \left[\frac{1\cdot(-5) + 6\cdot 1 + 3\cdot 5}{10}\right]^2 = 10{,}6 - 2{,}56 = 8{,}04 .$$

Згідно з формулою (23) знаходимо "виправлену" вибіркову дисперсію:

$$s_u^2 = \frac{1\cdot 5^2 + 6\cdot 1^2 + 3\cdot 5^2 - \frac{\left[1\cdot(-5) + 6\cdot 1 + 3\cdot 5\right]^2}{10}}{9} = \frac{106 - 25{,}6}{9} = 8{,}93 .$$

Усі початкові варіанти були зменшені на одне і те ж число $C = 187$, а тому дисперсія не змінилася, тобто шукана дисперсія дорівнює дисперсії в умовних варіантах:

$$s_x^2 = s_u^2 = 8{,}93 .$$

Зауваження 2. Якщо початкові варіанти є десятковими дробами з k десятковими знаками після коми, то щоб уникнути дій з дробами, перемножують початкові варіанти на стале число $C = 10^k$, тобто переходять до умовних варіант $u_i = Cx_i$. При цьому дисперсія збільшиться в C^2 разів. А тому для знаходження дисперсії потрібно дисперсію в умовних одиницях розділити на C^2:

$$D_B(X) = D_B(U)\cdot\frac{1}{C^2} \tag{24}$$

або

$$s_x^2 = \frac{s_u^2}{C^2} . \tag{25}$$

Приклад 6. Знайти вибіркову дисперсію та “виправлену” вибіркову дисперсію за заданим розподілом вибірки об’єму $n = 10$:

x_i	0,02	0,06	0,10
n_i	2	3	5

Щоб уникнути дій з дробами введемо умовні варіанти $u_i = 100x_i$, у результаті одержимо розподіл:

u_i	2	6	10
n_i	2	3	5

Знайдемо вибіркову дисперсію (21) умовних варіант:

$$D_B(U) = \frac{\sum_{i=1}^{n} n_i u_i^2}{n} - \left[\frac{\sum_{i=1}^{n} n_i u_i}{n}\right]^2 .$$

Підставимо в цю формулу умовні варіанти та їх частоти, одержимо:

$$D_B(U) = 61,6 - 51,84 = 9,76 .$$

Знайдемо шукану вибіркову дисперсію початкових варіант

$$D_B(X) = \frac{D_B(U)}{100^2} = \frac{9,76}{10000} = 0,000976 .$$

Знайдемо “виправлену” вибіркову дисперсію умовних варіант за формулою (23). Підставимо в цю формулу дані задачі, отримаємо:

$$s_u^2 = 10,844 .$$

Знайдемо шукану “виправлену” дисперсію початкових варіант:

$$s_x^2 = s_u^2 \cdot \frac{1}{100^2} = \frac{10,84}{1000} = 0,0010844 .$$

Зауваження 3. Порівнюючи формули (8) і (18) видно, що D_B і s^2 вирізняються лише знаменниками. Очевидно, що при достатньо великих значеннях n об’єму вибірки вибіркова та “виправлена” дисперсія різняться мало. На практиці користуються “виправленою” дисперсією, якщо приблизно $n < 30$. За великої кількості числових даних використовують метод добутків, або метод сум, які наводяться у більш детальних підручниках.

ЗАДАЧІ

2.6. Знайти вибіркове середнє за заданим розподілом вибірки об'єму $n = 10$:

Таблиця 2.12

x_i	1300	1350	1380
n_i	2	5	3

Вказівки: перейти до умовних варіант $u_i = x_i - 1350$.

Відповідь: $\overline{x_B} = 1349$.

2.7. Знайти вибіркову дисперсію за заданим розподілом вибірки об'єму $n = 60$:

Таблиця 2.13

x_i	310	330	345	350
n_i	10	40	6	4

Вказівки: перейти до умовних варіант $u_i = x_i - 330$.

Відповідь: $D_B(X) = D_B(U) = 115,58$.

2.8. Знайти вибіркову дисперсію за заданим розподілом вибірки об'єму $n = 70$:

Таблиця 2.14

x_i	0,2	0,6	0,7	0,9
n_i	10	20	25	15

Вказівки: перейти до умовних варіант $u_i = 10x_i$.

Відповідь: $D_B(X) = \dfrac{D_B(U)}{10^2} = \dfrac{4,40}{100} \approx 0,044$.

2.9. Знайти вибіркову дисперсію та "виправлену" вибіркову дисперсію за заданим розподілом вибірки об'єму $n = 50$:

Таблиця 2.15

x_i	18,5	18,9	19,4	19,7
n_i	5	10	20	15

Вказівки: перейти до умовних варіант $u_i = 10x_i - 195$.

Відповідь: $D_B\left(X\right)=\dfrac{D_B\left(U\right)}{10^2}=\dfrac{14,8}{100}\approx 0,148$.

$$s_x^2=s_u^2\cdot\frac{1}{100}=\frac{15,10}{100}\approx 0,1510.$$

2.10. Знайти “виправлену” вибіркову дисперсію за заданим розподілом вибірки об’єму n = 10:

Таблиця 2.16

x_i	23,6	26,2	28,3	30,5
n_i	2	3	4	1

Вказівки: перейти до умовних варіант $u_i = 10x_i - 268$.

Відповідь: $s_x^2=s_u^2\cdot\dfrac{1}{100}=489,17\cdot\dfrac{1}{100}=4,892$

§3. Інтервальні (надійні) оцінки параметрів

У деяких практичних задачах часто буває потрібно не тільки знайти для параметра a, придатне числове значення (дати точкову оцінку), але й оцінити його точність і надійність.

Такі задачі дуже важливі за малого числа спостережень, оскільки точкова оцінка $\overline{a}$ значною мірою є випадковою і наближена замінна a на $\overline{a}$ може призвести до серйозних помилок.

Для визначення точності оцінки $\overline{a}$ в математичній статистиці користуються надійними інтервалами, а для визначення надійності — надійними ймовірностями.

Нехай для параметра a одержана із експерименту незсувна оцінка $\overline{a}$. Потрібно оцінити можливу при цьому помилку.

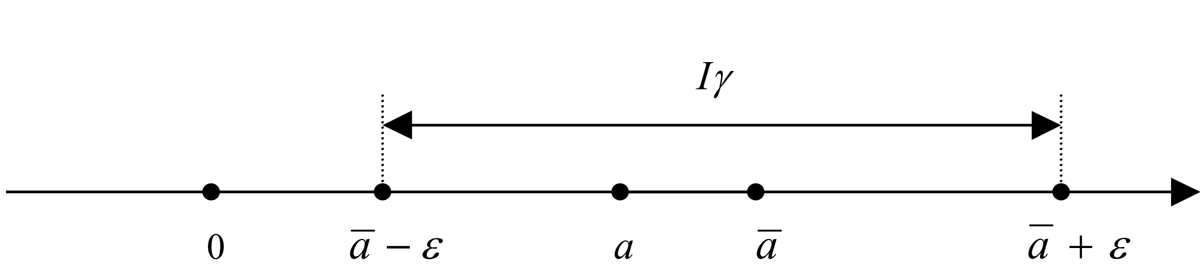


Рис. 28

Задамо деяку ймовірність γ (наприклад, $\gamma = 0{,}8$) і знаходимо таке значення $\varepsilon > 0$, для якого

$$P\left(\left|a - \overline{a}\right| < \varepsilon\right) = \gamma . \tag{1}$$

Представимо (1) у вигляді

$$P\left(\overline{a} - \varepsilon < a < \overline{a} + \varepsilon\right) = \gamma . \tag{2}$$

Рівність (2) означає, що невідоме значення параметра a з ймовірністю γ попадає в інтервал (рис. 28)

$$I_\gamma = \left(\overline{a} - \varepsilon,\ \overline{a} + \varepsilon\right). \tag{3}$$

Інтервал I_γ називається *надійним інтервалом*, $\left(\overline{a} - \varepsilon,\ \overline{a} + \varepsilon\right)$ — *надійними межами*, а ймовірність γ — *надійною ймовірністю* або надійністю, що відповідає даному надійному інтервалу і є ймовірністю того, що справжні значення параметра лежать в цьому інтервалі.

Розглянемо визначення надійного інтервалу для вибіркової середньої (математичного сподівання).

Нехай проведено n незалежних дослідів над випадковою величиною X з невідомою генеральною середньою $\overline{x}_Г$ та невідомою генеральною дисперсією. На основі експериментальних оцінок для цих параметрів визначена вибіркова середня $\overline{x}_В$. Потрібно побудувати надійний інтервал I_γ, який відповідає надійній ймовірності γ для величини $\overline{x}_Г$.

Відомо, що $M\left(\overline{X}_В\right) = \overline{x_Г},\ D\left(\overline{X}_В\right) = \dfrac{D\left(\overline{X}_Г\right)}{n}$.

Зауваження: Будемо розглядати вибіркову середню $\overline{x_В}$ як випадкову величину $\overline{X_В}$ ($\overline{X_В}$ змінюється від вибірки до вибірки) і вибіркові значення ознаки $x_1, x_2, \ldots, x_n$ — як однаково розподілені незалежні випадкові величини $X_1, X_2, \ldots, X_n$ (ці числа також змінюються від вибірки до вибірки).

Позначимо: ε_γ — точність оцінки (гранична похибка). Знайдемо тепер таку величину ε_γ, для якої (за формулою (1)):

$$P\left(\left|\overline{X}_В - \overline{X}_Г\right| < \varepsilon_\gamma\right) = \gamma . \tag{4}$$

Нехай закон розподілу випадкової величини є нормальним, а математичне сподівання кожної із них дорівнює “a” і середнє квадратичне відхилення — σ, то виразимо ймовірність через функцію Лапласа:

$$P\left(\left|\bar{X}_B - \bar{X}_{\tilde{A}}\right| < \varepsilon_\gamma\right) = 2\Phi\left(\frac{\varepsilon_\gamma}{\sigma}\right). \quad (5)$$

Замінивши $\bar{X}_Г$ на a і σ на $\sigma(\bar{X}_B) = \frac{\sigma}{\sqrt{n}}$, одержимо

$$P(\left|X_B - a\right|) < \varepsilon_\gamma = 2\Phi(\frac{\varepsilon_\gamma\sqrt{n}}{\sigma}) = 2\Phi(t),$$

де $t = \frac{\varepsilon_\gamma\sqrt{n}}{\sigma}$.

З урахуванням $\varepsilon_\gamma = \frac{t\sigma}{\sqrt{n}}$ можна написати $P(\left|X_B - a\right| < \frac{t\sigma}{\sqrt{n}} = 2\Phi(t)$.

Вважаючи, що ймовірність P відома і дорівнює γ, маємо робочу формулу, де вибіркову середню позначимо через $\bar{X}_B$

$$P(\bar{X}_B - \frac{t\sigma}{\sqrt{n}} < a < \bar{X}_B + \frac{t\sigma}{\sqrt{n}}) = \gamma. \quad (6)$$

Зміст отриманого співвідношення такий: з надійністю γ можна вважати, що надійний інтервал $(\bar{x}_B - \frac{t\sigma}{\sqrt{n}}, \bar{x}_B + \frac{t\sigma}{\sqrt{n}})$ покриває невідомий параметр a; точність оцінки $\varepsilon_\gamma = \frac{t\sigma}{\sqrt{n}}$.

Число t визначається із рівності $2\Phi(t) = \gamma$, або $\Phi(t) = \frac{\gamma}{2}$; за таблицями функції Лапласа (див. додаток 2) знаходять аргумент t, якому відповідає значення функції Лапласа рівне $\frac{\gamma}{2}$.

Надійні оцінки параметрів нормального розподілу. Нехай кількісна ознака X генеральної сукупності розподілена нормально, причому *середнє квадратичне відхилення σ цього розподілу невідоме*. Потрібно оцінити невідоме математичне сподівання a (генеральну середню) нормального розподілу ознаки за допомогою надійних інтервалів.

Виявляється, що за вибірковими даними, одержаними в незалежних спостереженнях, можливо побудувати таку випадкову величину T (її можливі значення будемо позначати через t), яка має розподіл, який не залежить від невідомого параметру σ:

$$T = \sqrt{n}\,\frac{\overline{X_B} - a}{S}, \tag{7}$$

де S — "виправлене" середнє квадратичне відхилення, $\overline{X_B}$ — вибіркове середнє, n — об'єм вибірки.

Для випадкової величини щільність ймовірності підкоряється закону:

$$S(t,n) = B_n\left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}}, \tag{8}$$

де B_n — коефіцієнт, який залежить від об'єму вибірки n:

$$B_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\cdot\Gamma\left(\frac{n-1}{2}\right)}.$$

Такий розподіл називається розподілом Стьюдента з $k = n-1$ ступенями волі і не залежить від невідомих параметрів a і σ, а визначається параметром n — об'ємом вибірки, що є важливим (див. Пояснення на стор. 138).

Оскільки $S(t,n)$ — парна функція від t, а тому, виходячи із формули (8, Розділ I, §4) для ймовірності попадання неперервної випадкової величини в заданий інтервал, ймовірність виконання нерівності:

$$|T| = \left|\sqrt{n}\,\frac{\overline{X}_B - a}{S}\right| < t_\gamma, \tag{9}$$

де t_γ — деяке стале число, дорівнює:

$$P\left(\left|\frac{\overline{X_B}-a}{S}\sqrt{n}\right|<t_\gamma\right)=2\int_0^{t_\gamma}S(t,n)dt=\gamma\ , \qquad (10)$$

де γ — називають надійністю.

Замінимо нерівність у круглих дужках рівносильною їй подвійною нерівністю, одержимо:

$$P\left(\overline{X_B}-t_\gamma\frac{S}{\sqrt{n}}<a<\overline{X_B}+t_\gamma\frac{S}{\sqrt{n}}\right)=\gamma\ . \qquad (11)$$

Отже, користуючись розподілом Стьюдента, ми знайшли надійний інтервал $\left(\overline{x}_B-t_\gamma\frac{s}{\sqrt{n}},\overline{x}_B+t_\gamma\frac{s}{\sqrt{n}}\right)$, що покриває невідомий параметр *a* з надійністю γ.

Тут випадкові величини $\overline{X_B}$ і S замінені невипадковими величинами $\overline{x_B}$ і s, що знайдені за вибіркою. Існують готові таблиці (див. додаток, таблиця 3), користуючись якими за заданими t_γ і n знаходять імовірність γ або за заданим γ і n можливо знайти t_γ.

Пояснення. Нехай X_i (i=1, 2, ..., n) — нормальні незалежні випадкові величини, причому математичне сподівання кожної із них дорівнює нулю, а середнє квадратичне відхилення — одиниці. Тоді сума квадратів цих величин

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

розділена за знаком χ^2 ("*хі квадрат*") з $k = n$ ступенями волі та щільністю розподілу

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

де $\Gamma(x)$ – гамма функція (див.: § 6, розділ I).

Очевидно, що розподіл "*хі квадрат*" визначається одним параметром — числом ступеней волі k, зі збільшенням якого розподіл повільно наближається до нормального.

Якщо Z — нормальна випадкова величина, причому $M(Z)=0$, $\sigma(Z)=1$, а V — незалежна від Z величина, яка розподілена за законом χ^2 з k ступенями волі, то величина

$$T=\frac{Z}{\sqrt{\frac{V}{k}}} \tag{12}$$

має розподіл, який називають *t*-розподілом, або розподілом Стьюдента з k ступенями волі.

Нехай кількістна ознака X генеральної сукупності розподілена нормально, причому $M(X)=a$, $\sigma(X)=\sigma$. Якщо із цієї сукупності здобувати вибірки об'єму n і за ними знаходити вибіркові середні, то не важко встановити, що вибіркова середня розподілена нормально, причому (див.: Розділ I, §5):

$$M\left(\overline{X_B}\right)=a\ ,\ \sigma\left(\overline{X_B}\right)=\frac{\sigma}{\sqrt{n}}\,.$$

Тоді випадкова величина

$$Z=\sqrt{n}\,\frac{\overline{X_B}-a}{\sigma} \tag{13}$$

також має нормальний розподіл як лінійна функція нормального аргументу $\overline{X_B}$ і $M(Z)=0$, $\sigma(Z)=1$. Можливо переконатися, що випадкова величина Z і

$$V=\frac{(n-1)\cdot S^2}{\sigma^2} \tag{14}$$

незалежні (S^2 — "виправлена" вибіркова дисперсія) і величина V розподілена за законом χ^2 з $k = n - 1$ ступенями волі. Якщо підставити (13) і (14) у (12) одержимо величину:

$$T=\frac{\left(\overline{x_B}-a\right)}{S}\cdot\sqrt{n} \tag{15}$$

яка розподілена за законом Стьюдента з $k = n - 1$ ступенями волі.

Для оцінки невідомого *генерального середнього квадратичного відхилення* σ за “виправленим” вибірковим середнім квадратичним відхиленням s та заданою надійністю γ, вимагатимемо виконання співвідношення:

$$P\left(\left|\sigma - s\right| < \varepsilon\right) = \gamma \text{ , або}$$
$$P\left(s - \varepsilon < \sigma < s + \varepsilon\right) = \gamma \text{ .}$$

Щоб можливо було користуватися готовою таблицею перетворимо подвійну нерівність

$$s - \varepsilon < \sigma < s + \varepsilon$$

у рівносильну нерівність

$$s(1 - \frac{\varepsilon}{s}) < \sigma < s(1 + \frac{\varepsilon}{s}) \text{ .}$$

Покладемо $q = \frac{\varepsilon}{s}$, одержимо

$$s(1 - q) < \sigma < s(1 + q) \text{ .}$$

Практично для відшукання q (див.: Зауваження 2) користуються таблицею Додатку 4. Обчисливши за вибіркою s і відшукавши за таблицею q, одержимо шуканий надійний інтервал, який покриває σ із заданою надійністю γ, тобто інтервал

$$s(1 - q) < \sigma < s(1 + q) \text{ .} \qquad (16)$$

Зауваження 1. Раніше вважалося, що $q < 1$**,** якщо $q > 1$, то записана нерівність прийме вигляд (ураховуючи, що $\sigma > 0$)

$$0 < \sigma < s(1 + q) \text{ .} \qquad (17)$$

Зауваження 2. Для відшукання q вводиться до розгляду випадкова величина “*хі* ”

$$\chi = \frac{S}{\sigma} \cdot \sqrt{n - 1} \text{ ,} \qquad (18)$$

де n — об’єм вибірки. Як було зазначено в поясненні співвідношення (14), величина $\frac{S^2}{\sigma^2}(n - 1)$ розподілена за законом χ^2 з $n - 1$ ступенями волі, а тому квадратний корінь із неї позначають через χ, а щільність розподілу χ має вигляд:

$$R(\chi,n)=\frac{\chi^{n-2}\cdot e^{-\frac{\chi^2}{2}}}{2^{\frac{n-3}{2}}\cdot\Gamma(\frac{n-1}{2})}, \qquad (19)$$

яка отримана з щільності розподілу випадкової величини, розподіленої за законом χ^2 з $k = n - 1$ ступенями волі

$$f(x)=\frac{x^{\frac{k}{2}-1}\cdot e^{-\frac{x}{2}}}{2^{\frac{k}{2}}\cdot\Gamma(\frac{k}{2})} \qquad (20)$$

після підстановки $k = n - 1$.

Розподіл $R(\chi,n)$ не залежить від оцінюваного параметра σ, а залежить лише від об'єму вибірки n.

Перетворення, за допомогою яких отримують $R(\chi,n)$, як і подальша оцінка нерівності (16), пов'язана з відшуканням q наводяться в більш детальних підручниках з теорії ймовірності та математичної статистики.

Приклад.1. Кількісна ознака X генеральної сукупності має нормальний розподіл. За вибіркою об'єму n = 16 знайдені вибіркове середнє $\overline{x}_B = 30,8$ та "виправлене" середнє квадратичне відхилення s = 0,6. Оцінити невідоме математичне сподівання за допомогою надійного інтервалу з надійністю γ = 0,95.

Знайдемо $t_\gamma = t(\gamma,n)$ за таблицею (див. додаток 3). За даними γ = 0,95, n = 16 знаходимо t_γ = 2,13.

Знайдемо надійні межі:

$$\overline{x_B}-t_\gamma\frac{s}{\sqrt{n}}=30,8-2,13\frac{0,6}{\sqrt{16}}=30,481$$
$$\overline{x_B}+t_\gamma\frac{s}{\sqrt{n}}=30,8+2,13\frac{0,6}{\sqrt{16}}=31,120.$$

Отже, з надійністю 0,95 невідомий параметр a міститься в надійному інтервалі: (30,481; 31,120).

Зауваження 3. Із граничних співвідношень

$$\lim_{n\to\infty} S(t,n) = \lim_{n\to\infty} B_n \cdot \left(1+\frac{t^2}{n-1}\right)^{-\frac{n}{2}} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} \tag{21}$$

випливає, що при необмеженому зростанні об'єму вибірки n розподіл Стьюдента прямує до нормального. А тому практично при $n > 30$ можливо замість розподілу Стьюдента користуватися нормальним розподілом. Проте важливо зауважити, що для малих значень n ($n < 30$) підміна розподілу нормальним приводить до грубих помилок, невиправдованого звуження надійного інтервалу, тобто до підвищення точності оцінки. Наприклад якщо $n = 5$ і $\gamma = 0,99$, то користуючись розподілом Стьюдента, знайдемо $t_\gamma = 4,6$, а використовуючи функцію Лапласа знайдемо $t_\gamma = 2,58$, тобто надійний інтервал стане більш вузьким, ніж знайдений за розподілом Стьюдента. Цей факт не засвідчує слабкість методу Стьюдента, а пояснюється тим, що мала вибірка містить малу інформацію про цікавлячу нас ознаку.

Приклад 2. За даними десяти незалежних рівноточних вимірів фізичної величини знайдені середнє арифметичне результатів окремих замірів $\overline{x_B} = 43,589$ і "виправлене" середнє квадратичне відхилення $s = 6,0$. Потрібно оцінити справжнє значення вимірюваної величини з надійністю $\gamma = 0,99$ за допомогою надійного інтервалу. Вважається, що результати замірів розподілені нормально.

Справжнє значення вимірюваної величини дорівнює її математичному сподіванню a.

А тому задача зводиться до оцінки математичного сподівання (за невідомим σ) за допомогою надійного інтервалу

$$\overline{x_B} - t_\gamma s \cdot \frac{1}{\sqrt{n}} < a < \overline{x_B} + t_\gamma s \cdot \frac{1}{\sqrt{n}},$$

який покриває a із заданою надійністю $\gamma = 0,99$. Користуючись таблицею Додаток 3, за $\gamma = 0,99$ і $n = 10$ знаходимо $t_\gamma = 3,25$.

Знайдемо точність оцінки: $t_\gamma s \cdot \frac{1}{\sqrt{n}} = 3,25 \cdot \frac{6}{\sqrt{10}} = \frac{19,5}{3,16} \approx 6,17$.

Знайдемо надійні межі: $\overline{x_B} - t_\gamma s \cdot \frac{1}{\sqrt{n}} = 43,589 - 6,17 = 37,419$.

$$\overline{x_B} + t_\gamma s \cdot \frac{1}{\sqrt{n}} = 43,589 + 6,17 = 49,759.$$

Отже, з надійністю 0,99 справжнє значення вимірюваної величини знаходиться в інтервалі

$$37{,}419 < a < 49{,}759.$$

Приклад 3. Із генеральної сукупності здобута вибірка об'єму $n = 15$.

Таблиця 2.17

варіанта x_i	1	1	2	4	5	6
частота n_i	3	2	3	3	2	2

Оцініть з надійністю 0,99 математичне сподівання a нормально розподіленої ознаки генеральної сукупності за вибірковою середньою за допомогою надійного інтервалу.

Вибіркову середню і "виправлене" середнє квадратичне відхилення знайдемо відповідно за формулами:

$$\overline{x_B} = \frac{\sum n_i x_i}{n}; \quad s = \sqrt{\frac{\sum n_i \left(x_i - \overline{x_B}\right)^2}{n-1}};$$

Підставимо в ці формули дані задачі, отримаємо $\overline{x_B} = 3$, $s = 1{,}93$. Користуючись таблицею Додатка 3 за $\gamma = 0{,}99$ і $n = 15$ знаходимо $t_\gamma = 2{,}98$.

Знайдемо шуканий надійний інтервал:

$$\overline{x_B} - t_\gamma s \cdot \frac{1}{\sqrt{n}} < a < \overline{x_B} + t_\gamma s \cdot \frac{1}{\sqrt{n}}.$$

Підставимо $\overline{x_B} = 3$, $t_\gamma = 2{,}98$, $s = 1{,}93$, $n = 15$, одержимо надійний інтервал

$$1{,}51 < a < 4{,}49,$$

що покриває невідоме математичне сподівання a з надійністю 0,99.

Приклад 4. Кількісна оцінка X генеральної сукупності розподілена нормально. За вибіркою об'єму $n = 20$ знайдено "виправлене" середнє квадратичне відхилення $s = 0{,}6$. Знайти надійний інтервал, який покриває генеральне середнє квадратичне відхилення σ із надійністю 0,99.

За таблицею Додатка 4 та даними $\gamma = 0,99$ і $n = 20$ знайдемо $q = 0,58$ ($q < 1$). Отже, надійний інтервал такий

$$0,6(1-0,58) < \sigma < 0,6(1+0,58) \text{ або}$$

$$0,252 < \sigma < 0,948 .$$

Приклад 5. Проведено вісім замірів одним приладом (без систематичної помилки) деякої фізичної величини, причому “виправлене” середнє квадратичне відхилення s випадкових помилок замірів дорівнює $s = 0,8$. Знайти точність приладу із надійністю 0,99. Вважатимемо, що результати замірів розподілені нормально.

Точність приладу характеризується середнім квадратичним відхиленням випадкових помилок замірів. А тому задача зводиться до відшукання надійного інтервалу, який покриває σ із заданою надійністю 0,99. За даними $\gamma = 0,99$ і $n = 8$ та таблицею Додатка 4 знайдемо $q = 1,38$ ($q > 1$).

Отже, надійний інтервал такий:

$$0 < \sigma < 0,8(1+1,38) \text{ або}$$

$$0 < \sigma < 1,9 .$$

Статистична перевірка гіпотез ймовірностей. При вивченні генеральної сукупності часто необхідно знати закон її розподілу. Якщо закон розподілу невідомий, то є підстава допустити, що він має певний вигляд, тобто висувають *гіпотезу*: генеральна сукупність розподілена по закону A.

Гіпотеза називається *статистичною*, якщо вона належить або до вигляду, або до параметрів випадкової величини і може бути перевірена *статистично*, тобто за допомогою результатів спостережень.

Приклад. Нехай у результаті спостережень деякої випадкової величини одержані емпіричні частоти.

Висуваємо статистичну гіпотезу про те, що генеральна сукупність розподілена нормально, тоді можна обчислити теоретичні або вирівнюючі частоти. Наприклад:

Таблиця 2.18

Емпіричні	7	12	36	76	104	83	28	8	5
(вирівнюючі)теоретичні	4	13	44	80	96	72	36	9	1

Як правило, емпіричні і теоретичні частоти різняться. Виникає необхідність встановити правило (критерій), яке дозволило б судити, чи є розбіжність між емпіричним і теоретичним розподілом випадковим чи значущим.

Критерієм згоди називається критерій, який дозволяє встановити, чи є розбіжність емпіричного і теоретичного розподілу випадковим чи значущим, тобто чи узгоджуються дані спостережень з висунутою статистичною гіпотезою чи ні.

При перевірці статистичної гіпотези можуть бути допущені помилки двох видів: вірна гіпотеза заперечується (*помилка першого роду*), хибна гіпотеза приймається (*помилка другого роду*).

Ймовірність допустити помилку першого роду називається *рівнем значущості критерія*. Чим менше рівень значущості, тим менша ймовірність відкинути вірну гіпотезу. Як рівень значущості приймається число $\alpha = 0{,}1;\ 0{,}05;\ 0{,}01;\ 0{,}001$ та інші. Якщо, наприклад, $\alpha = 0{,}01$ — то рівень значущості називається однопроцентним, це означає, що в одному випадку із ста ми ризикуємо зробити помилку першого роду (тобто відхилити правильну гіпотезу). Чим серйозніші наслідки помилки першого роду, тим менше має бути рівень значущості.

Нехай у результаті n спостережень одержано емпіричний розподіл частоти і у припущенні нормального розподілу генеральної сукупності теоретичні частоти n_i:

Таблиця 2.19

	x_1	x_2	...	x_s
Емпірична частота	n_1	n_2	...	n_s
Теоретична частота	n_1'	n_2'	...	n_s'

$(n = n_1 + n_1 + ... + n_s)$

Теоретичні частоти, взагалі кажучи, не співпадають з емпіричними. Як міру розходження частоти рядів емпіричного і теоретичного розподілу Пірсон запропонував формулу:

$$\chi^2 = \sum_{i=1}^{s} \frac{\left(n_i - n_i'\right)^2}{n_i'} \ . \qquad (22)$$

Це критерій згоди Пірсона (критерій “*хі квадрат*”).

Із (22) видно, що якщо всі відповідні теоретичні і емпіричні частоти співпадають, то $\chi^2 = 0$, тобто чим ближче χ^2 до нуля, тим краще узгоджуються емпіричні і теоретичні розподіли.

Величина χ^2 — випадкова і при достатньо великих n, вона має "*хі квадрат*" розподіл, тобто щільність розподілу її ймовірності має вигляд:

$$f(x) = \begin{cases} 0, \ \ \textit{якщо} \ \ x \le 0 \\ \dfrac{1}{2^{\frac{k}{2}} \cdot \Gamma\left(\dfrac{k}{2}\right)} e^{-\frac{x}{2}} \cdot x^{\frac{\kappa}{2}-1}, \ \ \textit{якщо} \ \ x > 0 \end{cases}, \tag{23}$$

де κ — параметр розподілу, який називають *числом ступенів волі*; $\Gamma\left(\dfrac{\kappa}{2}\right)$ — значення гамма-функції.

Число ступенів волі κ визначається як:

$$\kappa = S - r - 1, \tag{24}$$

де S — число груп емпіричного розподілу; r — сума числа параметрів теоретичного закону і числа додаткових співвідношень емпіричних частот (наприклад, для нормального закону $r = 3$, оскільки параметрів розподілу два (a і σ) і є одне додаткове співвідношення: $n = n_1 + n_1 + ... + n_s$).

Отже,

$$\kappa = s - 1 - r = s - 1 - 2 = s - 3.$$

Якщо, наприклад, припустити, що генеральна сукупність розподілена за законом Пуассона, то оцінюють один параметр λ, а тому $r = 1$ і $\kappa = s - 2$.

Сформулюємо правило (алгоритм) застосування критерію Пірсона:

1. Із генеральної сукупності утворюється випадкова вибірка, вибирається передбачуваний закон розподілу ознаки генеральної сукупності і знаходять параметри цього закону.
2. Обчислюють теоретичні частоти.
3. За формулою (22) знаходиться χ^2, позначаючи її значення χ_0^2.
4. За формулою (24) знаходиться величина κ.
5. Вибирається рівень значущості α.

6. За таблицями знаходиться $P\left(\chi^2 \geq \chi_0^2\right) = \beta$.

7. Якщо $\beta > \alpha$, то приймається гіпотеза про розподіл за предбачуваним законом (розходження між емпіричними і теоретичними частотами несуттєві), якщо $\beta \leq \alpha$, то гіпотезу потрібно відкинути або провести ще експеримент (якщо при повторному експерименті знову $\beta \leq \alpha$, то гіпотезу слід відкинути як неправдоподібну).

Крім критерію Пірсона є і інші: критерій Колмогорова, Романовського та інші.

Приклад 6. Використовуючи критерії Персона за рівнем значущості 0,025 перевірить чи узгоджується гіпотеза про нормальний розподіл генеральної сукупності X із емпіричним розподілом вибірки об'єму $n = 210$.

Таблиця 2.20

x_i	6	8	10	12	14	16	18	20	22
n_i	16	27	26	31	27	22	25	21	15

1) Вибіркове середнє $\overline{x_B} = 13,69$ і вибіркове середнє квадратичне відхилення $\sigma_B = 4,74$ знайдено за формулами:

$$\overline{x_B} = \frac{\sum n_i x_i}{n} \; ; \; \sigma_B = \sqrt{\frac{\sum n_i \left(x_i - \overline{x_i}\right)^2}{n}} \; .$$

2) Обчислюючи теоретичні частоти, враховуючи, що $n = 210$, $h = 2$, $\sigma_B = 4,74$ за формулою:

$$n_i' = \frac{nh}{\sigma_B} \cdot \varphi\left(u_i\right) = \frac{210 \cdot 2}{4,74} \cdot \varphi\left(u_i\right) = 88,61 \cdot \varphi\left(u_i\right),$$

де n — об'єм вибірки (сума всіх частот), h — крок (різниця між двома сусідніми варіантами),

$$u_i = \frac{x_i - \overline{x_B}}{\sigma_B} \; , \; \varphi(u) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}} \; .$$

Складемо розрахункову таблицю 2.21 (значення функції $\varphi(u)$ наведені в Додатку 1)

Таблиця 2.21

i	x_i	$u_i = \dfrac{x_i - \overline{x_B}}{\sigma_B}$	$\varphi(u_i)$	$n_i' = 88{,}61 \cdot \varphi(u_i)$
1	6	–1,62	0,1074	9,5
2	8	–1,20	0,1942	17,21
3	10	–0,78	0,2943	26,1
4	12	–0,36	0,3752	33,3
5	14	0,07	0,3980	35,3
6	16	0,49	0,3538	31,4
7	18	0,91	0,2637	23,4
8	20	1,33	0,1647	14,6
9	22	1,75	0,0863	7,6

Порівняємо емпіричні і теоретичні частоти за допомогою критерію Пірсона. Для цього:

а) складемо розрахункову таблицю 2.22, із якої знаходять спостережне значення критерію

$$\chi^2_{спост} = \sum \frac{\left(n_i - n_i'\right)^2}{n_i'} \quad ;$$

б) за таблицею критичних точок розподілу χ^2 (див.: Додаток 5), за рівнем значущості $\alpha = 0{,}025$ та числом ступенів волі $k = s - 3 = 9 - 3 = 6$ (s — число груп вибірки) знаходять критичну точку $\chi^2_{кр}(\alpha; k)$ правосторонньої критичної області

$$\chi^2_{кр}(0{,}025; 6) = 14{,}4 .$$

Оскільки $\chi^2_{спост} > \chi^2_{кр}$ — гіпотезу про нормальний розподіл генеральної сукупності відкидаємо. Іншими словами, емпіричні і теоретичні частоти різняться значно.

Таблиця 2.22

i	n_{i}	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$\frac{(n_i - n_i')^2}{n_i'}$
1	16	9,5	6,5	42,25	4,5
2	27	17,2	9,8	96,04	5,6
3	26	26,1	–0,1	0,01	0,0
4	31	33,3	–2,3	5,29	0,2
5	27	35,3	–8,3	68,89	2,0
6	22	31,4	–9,4	88,36	2,8
7	25	23,4	1,6	2,56	0,1
8	21	14,6	6,4	40,96	2,8
9	15	7,6	7,4	54,76	7,2
$\sum$	210				$\chi^2_{спост} = 25,2$

Якби $\chi^2_{спост} < \chi^2_{кр}$ — тоді не було б підстав відкидати гіпотезу про нормальний розподіл генеральної сукупності.

Приклад 7. Використовуючи критерії Пірсона за рівнем значущості 0,025 встановити випадкова чи значуща розбіжність між емпіричними частотами n_i та теоретичними частотами n_i' які обчислені, виходячи із гіпотези про нормальний розподіл генеральної сукупності X:

Таблиця 2.23

n_i	10	18	42	74	38	20	12
n_i'	8	20	38	78	41	20	9

Знайдемо спостережне значення критерію Пірсона $\chi^2_{спост} = \sum \frac{\left(n_i - n_i'\right)^2}{n_i'}$. Складемо розрахункову таблицю 2.24, із якої знайдемо спостережне значення критерію $\chi^2_{спост} = 2,55$.

За таблицею критичних точок розподілу χ^2 (див.: Додаток 5) за рівнем значущості 0,025 і числу ступенів волі $k = s - 3 = 7 - 3 = 4$ знаходимо критичну точку правосторонньої критичної області $\chi^2_{кр}(0,025;4) = 11,1$.

Оскільки $\chi^2_{спост} < \chi^2_{кр}$ — немає підстав відкинути гіпотезу про нормальний розподіл генеральної сукупності. Іншими словами, розбіжність між емпіричними та теоретичними частотами не значуща (випадкова).

Таблиця 2.24

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	10	8	2	4	0,5
2	18	20	-2	4	0,2
3	42	38	4	16	0,42
4	74	78	-4	16	0,21
5	38	41	-3	9	0,22
6	20	20	0	0	0
7	12	9	3	9	1
$\sum$	204				$\chi^2_{спост} = 2,55$

ЗАДАЧІ

2.11. Із генеральної сукупності здобута вибірка об'єму $n = 12$.

Таблиця 2.25

варіанта x_i	–0,6	–0,5	–0,3	0	0,3	0,7	0,9	1	1,2	1,4
частота n_i	1	2	1	1	1	1	1	1	2	1

Оцінити з надійністю 0,99 математичне сподівання a нормально розподіленої ознаки генеральної сукупності за допомогою надійного інтервалу.

Відповідь: $0{,}06 < a < 0{,}74$.

2.12. За даними 16 незалежних рівноточних вимірів деякої фізичної величини знайдені середні арифметичні результатів замірів $\overline{x_B} = 42{,}8$ і "виправлене" середнє квадратичне відхилення $s = 8$. Оцінити справжнє значення заміреної величини з надійністю $\gamma = 0{,}999$.

Відповідь: $34{,}66 < a < 50{,}94$.

2.13. Знайти точність приладу із надійністю 0,95, вважаючи що результати вимірювань розподілені нормально. З цією метою одним приладом проведено 10 замірів (без систематичних помилок) деякої фізичної величини, причому "виправлене" середнє квадратичне відхилення s випадкових помилок вимірювань дорівнює 0,8.

Відповідь: $0{,}28 < \sigma < 1{,}32$.

2.14. Знайти надійний інтервал, який покриває генеральне середнє квадратичне відхилення σ із надійністю 0,999, якщо за даними вибірки n із генеральної сукупності нормально розподіленої кількісної ознаки знайдено "виправлене" середнє квадратичне відхилення s:

а) $n = 10$, $s = 5{,}1$; б) $n = 50$, $s = 14$;

Відповідь: а) $0 < \sigma < 14{,}28$; б) $7{,}98 < \sigma < 20{,}02$.

2.15. Використовуючи критерій Пірсона за рівнем значущості 0,05 перевірте, чи узгоджується гіпотеза про нормальний розподіл генеральної сукупності X із емпіричним розподілом вибірки об'єму $n = 200$.

Таблиця 2.26

x_i	0,3	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9	2,1	2,3
n_i	6	9	26	25	30	26	21	24	20	8	5

Відповідь: $k = 8$; $\chi^2_{спост} = 7{,}71$; $\chi^2_{кр}(0{,}05;8) = 15{,}5$, немає підстав відкидати гіпотезу про нормальний розподіл генеральної сукупності.

2.16. Використовуючи критерій Пірсона за рівнем значущості 0,05 переконайтесь, випадкова чи значуща розбіжність між емпіричними частотами n_i і теоретичними частотами n_i', які обчислені виходячи із гіпотези про нормальний розподіл генеральної сукупності X.

Таблиця 2.27

а)

n_i	5	10	20	8	7
n_i'	6	14	18	7	5

Таблиця 2.28

б)

n_i	14	18	32	70	20	36	10
n_i'	10	24	34	80	18	22	12

Відповідь: а) випадкова; $k = 2$; $\chi^2_{спост} = 2,44$; $\chi^2_{кр}(0,05;2) = 6,0$; б) значуща; $k = 4$; $\chi^2_{спост} = 13,93$; $\chi^2_{кр}(0,05;4) = 9,5$.

§4. Елементи теорії кореляції

Функціональна, статистична та кореляційна залежності. Якщо на деякій множині X задана *функціональна* залежність $y = f(x)$ між змінними x та y, то кожному значенню $x \in X$ відповідає одне певне значення $y \in Y$. Але не будь-яка залежність між двома змінними є *функціональною*. Нехай Y — врожай зерна. X — кількість добрив. З однакових за площею ділянок землі при рівних кількостях внесених добрив збирають різний врожай, тобто Y не є функцією X. Наприклад, ріст людини та її вага: знаючи ріст людини не можна знайти її вагу і навпаки. Але, зокрема, залежність між ростом людини та її вагою існує, але вона не є *функціональною*. Об'єм валової продукції та собівартості продукції, що випускається, не є функціональною залежністю, а *статистичною*. Строга функціональна залежність реалізується рідко, оскільки обидві величини або одна з них підлягають ще дії випадкових чинників, у цьому випадку виникає *статистична залежність* за якою зміна однієї із величин тягне зміну розподілу іншої. Тобто залежність між змінними x та y статистична, якщо значенню x_1 відповідає один розподіл величини y:

Таблиця 2.29

Y	y_1	y_2	…	y_k
m	m_1	m_2	…	m_k

а значенню x_2 відповідає інший розподіл, відмінний від першого:

Таблиця 2.30

Y	y_1'	y_2'	…	y_k'
m	m_1'	m_2'	…	m_k'

Отже, розподіл змінюється разом зі зміною x. Зокрема, статистична залежність виявляється в тому, що при зміні однієї із величин змінюється середнє значення іншої, у цьому випадку статистичну залежність називають *кореляційною*.

Розглянемо дві таблиці:

Таблиця 2.31

x \ y	1	2	3	4
1	3	4	2	1
3	3	4	2	1
5	3	4	2	1

Таблиця 2.32

x \ y	1	2	5	7	8
1	1	2	3		
2		1	2	4	
4			3	5	3
6				2	3

Дані в цих таблицях можна розглядати як випадкові вибірки генеральних сукупностей за певними ознаками X та Y. Із таблиці 2.31 видно, що y приймає значення: 1; 2; 3; 4, а x — 1; 3; 5. У таблиці наведені частоти, наприклад, у першому рядку число показує, що пара (x, y) приймає значення (1; 1) три рази. Із таблиці 2.31 також видно, що статистична залежність між змінними x і y відсутня, оскільки кожному значенню x (x = 1; 3; 5) відповідає один і той же розподіл y:

Таблиця 2.33

y	1	2	3	4
m	3	4	2	1

У таблиці 2.32 статистичний зв'язок між x і y має місце, оскільки різним значення x відповідають різні розподіли змінної y:

Таблиця 2.34

Для $x = 1$

y	1	2	5
m	1	2	3

Для $x = 2$

y	2	5	7
m	1	2	4

Таблиці наведеного виду, в яких містяться статистичні дані системи двох випадкових величин x і y, називаються *кореляційними таблицями*.

При великій кількості спостережень, одне і теж значення x може зустрічатися m_x раз, одне і теж значення y — m_y раз, одна і та ж пара чисел (x, y) може спостерігатися m_{xy} раз. А тому дані спостережень групуються, тобто підраховуються частоти m_x, m_y, m_{xy}. Усі згруповані дані записують у вигляді таблиці, яку називають кореляційною.

Нехай після обробки дослідних даних та переходу до відповідних дискретних розподілів одержано розподіл 100 га орної землі за кількістю внесених добрив x (*ц* на 1 *га*) і за врожайністю y (*ц* з 1 *га*), наведено в таблиці 2.35.

Таблиця 2.35

y_j (*ц* з 1 *га*) / x_i (*ц* на 1 *га*)	20	22	24	26	28	30	Підсумок
20	12	6	1	–	–	–	19
40	3	10	8	4	–	–	25
60	–	3	7	12	5	–	27
90	–	–	1	10	14	4	29
Підсумок	15	19	17	26	19	4	100

У цій таблиці, наприклад, число 6, яке знаходиться на перетині 1-го рядка та 2-го стовпця, показує, що на 6 *га* із 100 було внесено по 20 *ц* добрив і при цьому одержана урожайність по 22 *ц* з *га*. Таблиці такого типу називають *кореляційними*. Кореляційна таблиця 2.35 дає можливість зробити висновок, що зі збільшенням кількості внесених добрив x урожайність y має тенденцію до підвищення. Але вигляд цієї залежності та її аналітичний вираз поки залишаються невідомими.

Характер такої залежності можна встановити по-різному. Наприклад, кожному значенню змінної x, поставимо у відповідність середнє арифметичне відповідного йому розподілу змінної y. Змінна x приймає чотири різні значення. Отже, вся сукупність розбивається ними на чотири групи, об'єми яких є 19; 25; 27; 29 *га*.

Хоча на кожний гектар кожної групи внесено однакову кількість добрив (по 20 *ц* на кожний гектар першої групи, по 40 *ц* — другої, по 60 *ц* — третьої, по 90 *ц* — четвертої), врожайність на них не стала однаковою. У першій групі по 20 *ц* на 12 *га*, по 22 *ц* на 6 га і 24*ц* на 1 *га*. У другій групі — по 20 *ц* на 3 *га*, по 22 *ц* на 10 *га*, по 24 *ц* на 8 *га*, по 26 *ц* на 4 *га*. Аналогічно картина має місце в третій і четвертій групах.

Розподіл за врожайністю тієї площі, на кожний гектар якої було внесено однакову кількість добрив, наведено в таблицях 2.36, 2.37, 2.38, 2.39.

Таблиця 2.36

Врожайність, *ц* з *га*	20	22	24	Σ
Площа, *га*	12	6	1	19

Таблиця 2.37

Врожайність, *ц* з *га*	20	22	24	26	Σ
Площа, *га*	3	10	8	4	25

Таблиця 2.38

Врожайність, *ц* з *га*	22	24	26	28	Σ
Площа, *га*	3	7	12	5	27

Таблиця 2.39

Врожайність, *ц* з *га*	24	26	28	30	Σ
Площа, *га*	1	10	14	4	29

Середні арифметичні цих розподілів є груповими середніми врожайності. Позначимо їх через $\overline{y}_1$, $\overline{y}_2$, $\overline{y}_3$, $\overline{y}_4$.

Групове середнє $\overline{y}_1$ виражає середню врожайність на тій площі, на кожний гектар якої внесено по 20 *ц* добрив, і є середнім арифметичним розподілу, що наведений у таблиці 2.36.

$$\overline{y}_1 = \frac{20 \cdot 12 + 22 \cdot 6 + 24 \cdot 1}{19} \approx 20{,}84 \text{ ц.}$$

Аналогічно групова середня $\overline{y}_2$ — середня врожайність на тій площі, на кожний гектар якої внесено 40 *ц* добрив:

$$\overline{y}_2 = \frac{20 \cdot 3 + 22 \cdot 10 + 24 \cdot 8 + 26 \cdot 4}{25} = 23{,}04 \text{ ц.}$$

і т.д. $\overline{y}_3 \approx 25{,}41$, $\overline{y}_4 \approx 27{,}45$.

Отже, зі зростанням x групові середні врожайності зростають, причому майже рівномірно.

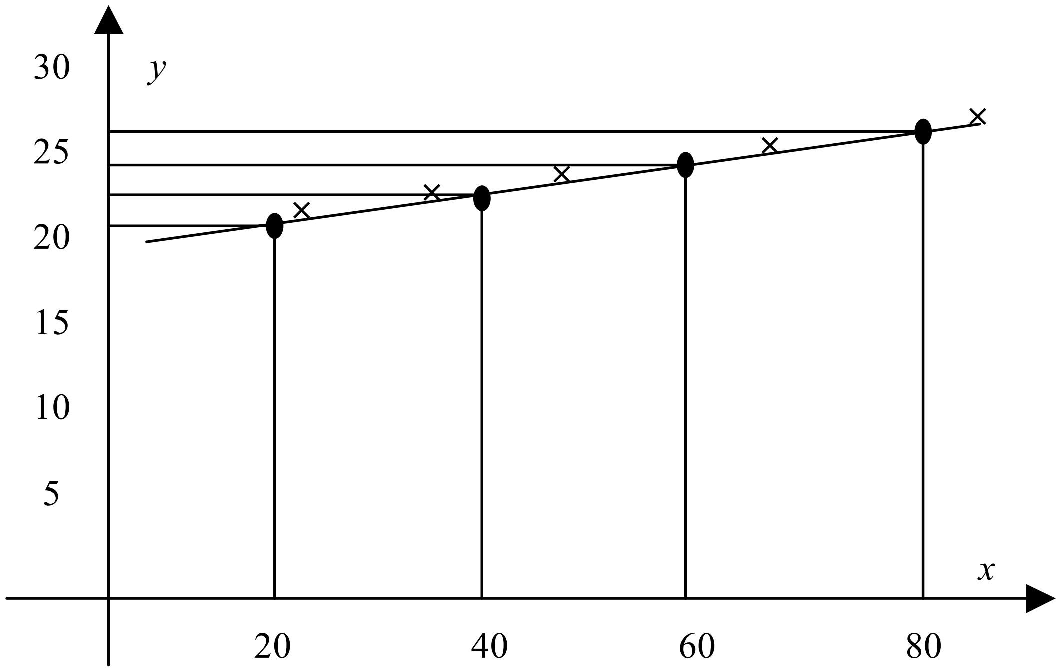


Рис. 29

На рис. 29 кружечками побудовані точки $(x_i; \overline{y}_i)$. Таким чином, залежність групових середніх врожайностей $\overline{y}_i$ від кількості внесених добрив x_i близька до лінійної, тобто

$$y = ax + b.$$

Можна поставити у відповідність кожному значенню врожайності середнє арифметичне відповідного йому розподілу площі за кількістю внесених добрив. За числом різних значень змінної y маємо шість груп з однаковою врожайністю. Але на ділянці кожної групи внесено неоднакову кількість добрив. Розподіли за кількістю внесених добрив тих ділянок, на яких одержана врожайність $y_1 = 20$ *ц* і $y_2 = 22$ *ц* наведені відповідно в таблицях 2.40, 2.41.

Таблиця 2.40

x_i *ц* на 1 *га*	Площа, *га*
20	12
40	3
Σ	15

Таблиця 2.41

x_i *ц* на 1 *га*	Площа, *га*
20	6
40	10
60	3
Σ	19

Рекомендуємо самостійно виписати розподіл відповідних значень y = 24; 26; 28; 30 та порахувати середні арифметичні цих розподілів, групові середні змінної x, позначивши їх через $\overline{x}_1,\ \overline{x}_2,...,\overline{x}_6$. Результати обчислень наведені в таблиці 2.42.

Таблиця 2.42

y_j, *ц* з *га*	$\overline{x}_j$ *ц* на *га*		y_j, *ц* з *га*	$\overline{x}_j$ *ц* на *га*	
20	24		26	68,5	
22	36,8		28	82,1	
24	50		30	90	

На тому ж рисунку хрестиками побудуємо точки $\left(\overline{x}_j; y_j\right)$. Можна вважати, що залежність групових середніх $\overline{x}_j$ від відповідних значень y наближена до лінійної

$$x = cy + d\ .$$

Отже, згідно з кореляційною таблицею 2.35 характер залежності між кількістю внесених добрив та врожайністю не був зрозумілий. Проте між значеннями однієї із цих величин і відповідними їм груповими середніми іншої існує майже функціональна залежність.

Розглянемо задачу в загальному вигляді. Нехай залежність між змінними x і y задана у вигляді кореляційної таблиці, тобто складемо кореляційну таблицю 2.43 загального вигляду і пояснимо її будову.

Таблиця 2.43

x_i \ y_j	y_1	y_2	…	y_j	…	y_l	n_i
x_1	m_{11}	m_{12}	…	m_{1j}	…	m_{1l}	n_1
x_2	m_{21}	m_{22}	…	m_{2j}	…	m_{2l}	n_2
…	…	…	…	…	…	…	…
x_i	m_{i1}	m_{i2}	…	m_{ij}	…	m_{il}	n_i
…	…	…	…	…	…	…	…
x_k	m_{k1}	m_{k2}	…	m_{kj}	…	m_{kl}	n_k
m_j	m_1	m_2	…	m_j	…	m_l	n

Тут m_{ij} — частота ($i = 1, 2, ..., k, j = 1, 2, ..., l$), з якою зустрічається пара (x_i, y_j); $n_i = \sum_{j=1}^{l} m_{ij}$; $m_j = \sum_{i=1}^{k} m_{ij}$; $n = \sum_{i=1}^{k} n_i = \sum_{j=1}^{l} m_j = \sum_{i=1}^{k}\sum_{j=1}^{l} m_{ij}$.

Із таблиці видно, що $x = x_1$ відповідає розподіл:

Таблиця 2.44

y_1	y_2	...	y_j	...	y_l
m_{11}	m_{12}	...	m_{1j}	...	m_{1l}

Знайдемо середнє значення y, що відповідає x_1:

$$\overline{y}_{x_1} = \frac{y_1 m_{11} + y_2 m_{12} + ... + y_l m_{1l}}{n_1}, \text{ де } n_1 = \sum_{j=1}^{l} m_{1j} .$$

Таке значення називають *умовним середнім значенням* величини y, якщо $x = x_1$.

Аналогічно знаходяться умовні середні значення $\overline{y}_x$ і для інших значень x, і тоді складемо таблицю 2.45.

Таблиця 2.45

x	x_1	x_2	...	x_i	...	x_k
$\overline{y}_x$	$\overline{y}_{x_1}$	$\overline{y}_{x_2}$	...	$\overline{y}_{x_i}$	...	$\overline{y}_{x_k}$
Частота	n_1	n_2	...	n_i	...	n_k

Аналогічно можна скласти таблицю 2.46 — таблицю умовних середніх $\overline{x}_y$ для y:

Таблиця 2.46

y	y_1	y_2	...	y_j	...	y_l
$\overline{x}_y$	$\overline{x}_{y_1}$	$\overline{x}_{y_2}$	...	$\overline{x}_{y_j}$	...	$\overline{x}_{y_l}$
Частота	m_1	m_2	...	m_j	...	m_l

Якщо кожному значенню x відповідає одне цілком певне умовне середнє значення $\overline{y}_x$, тобто якщо між змінними x та $\overline{y}_x$ існує функціональна залежність

$$\overline{y}_x = f(x) \qquad (1)$$

і $f(x) \neq const$ на множині значень *x*, то в цьому випадку статистична залежність між змінними називається кореляційною залежністю. Аналогічно, якщо існує функціональна залежність

$$\overline{x}_y = \varphi(y) \qquad (2)$$

і $\varphi(y) \neq const$ на множині значень *y*, то між змінними $\overline{x}_y$ та *y* також існує кореляційна залежність.

Рівняння (1), (2) називаються *кореляційними рівняннями* або *рівняннями регресії*. Відповідні графіки наведених залежностей називаються *кривими регресії*.

У теорії кореляції розв'язуються дві основі задачі. 1. За даними кореляційної таблиці визначають форму кореляційного зв'язку, тобто визначають вид функції $\overline{y}_x = f(x)$ або $\overline{x}_y = \varphi(y)$.

2. Оцінюють силу кореляційного зв'язку, тобто дають оцінку ступеню розсіювання значень *y* навколо умовного середнього $\overline{y}_x$ (або *x* навколо $\overline{x}_y$).

Лінійна кореляційна залежність та пряма регресія. Спочатку розглянемо першу задачу теорії кореляції: за даними таблиці 2.45 визначимо, яку дійсно функціональну залежність між *x* і $\overline{y}_x$ доцільно прийняти.

У прямокутній системі координат позначимо всі точки, які відповідають парам чисел $\left(x_i;\overline{y}_{x_i}\right)$ таблиці 2.45, тобто побудуємо *поле кореляції.* З'єднаємо ці точки ламаною лінією, яку називають *емпіричною лінією регресії y на x*.

Припустимо, що точки $M_i\left(x_i;\overline{y}_{x_i}\right)$ будуть групуватись навколо деякої прямої. Приймемо наближено залежність між *x* і $\overline{y}_x$ за лінійну:

$$\overline{Y}_x = ax + b. \qquad (3)$$

Шуканою прямою буде та, яка за змістом методу найменших квадратів найближче розташована до точок $M_1, M_2, ..., M_k$, причому точка M_1 ураховується n_1 раз, точка M_2 — n_2 рази і т.д. (стільки ж раз у розподілі зустрічаються відповідні значення x_i).

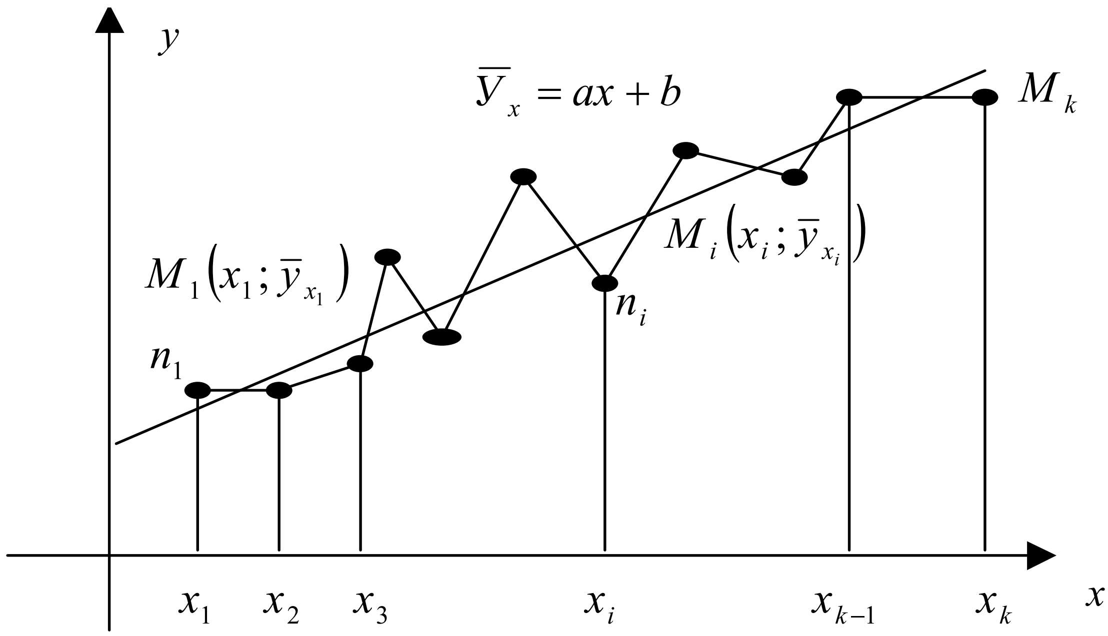


Рис.30

Невідомі коефіцієнти a і b визначимо найліпшим чином, щоб сума відповідних відхилень була мінімальною, тобто

$$n_1\left|\overline{Y}_{x_1}-\overline{y}_{x_1}\right|+n_2\left|\overline{Y}_{x_2}-\overline{y}_{x_2}\right|+\ldots+n_k\left|\overline{Y}_{x_k}-\overline{y}_{x_k}\right|\to\min$$

або $\quad n_1\left|ax_1+b-\overline{y}_{x_1}\right|+n_2\left|ax_2+b-\overline{y}_{x_2}\right|+\ldots+n_k\left|ax_k+b-\overline{y}_{x_k}\right|\to\min\,. \qquad (4)$

Різниці: $ax_i+b-\overline{y}_{x_i}\ \left(i=1,2,\ldots,k\right)$ називаються *похибками*. Застосуємо відомий спосіб найменших квадратів. Для цього складемо функцію

$$I\left(a;b\right)=n_1\left(ax_1+b-\overline{y}_{x_1}\right)^2+n_2\left(ax_2+b-\overline{y}_{x_2}\right)^2+\ldots+n_k\left(ax_k+b-\overline{y}_{x_k}\right)^2;$$

або $$I\left(a;b\right)=\sum_{i=1}^{k}n_i\left(ax_i+b-\overline{y}_{x_i}\right)^2. \qquad (5)$$

Отже, підберемо параметри a і b таким чином, щоб сума квадратів відхилень була мінімальною (у цьому полягає суть методу найменших квадратів). Оскільки кожне відхилення залежить від шуканих параметрів, то і сума квадратів відхилень (похибок) є функція цих параметрів. Для знаходження мінімуму знайдемо частинні похідні функції (5) по (a і b) і прирівняємо їх до нуля.

$$\frac{\partial I}{\partial a} = 2\sum_{i=1}^{k} n_i x_i \left(a x_i + b - \overline{y}_{x_i} \right) = 0$$

$$\frac{\partial I}{\partial b} = 2\sum_{i=1}^{k} n_i \left(a x_i + b - \overline{y}_{x_i} \right) = 0.$$

Виконавши елементарні перетворення, одержимо систему двох лінійних рівнянь щодо a і b.

$$\begin{cases} \left(\sum_{i=1}^{k} n_i x_i^2 \right) a + \left(\sum_{i=1}^{k} n_i x_i \right) b = \sum_{i=1}^{k} n_i x_i \, \overline{y}_{x_i}; \\ \left(\sum_{i=1}^{k} n_i x_i \right) a + \left(\sum_{i=1}^{k} n_i \right) b = \sum_{i=1}^{k} n_i \overline{y}_{x_i}. \end{cases} \tag{6}$$

Розділимо на n кожне рівняння системи (6) і введемо позначення:

$$\overline{x} = \frac{1}{n} \cdot \sum_{i=1}^{k} n_i x_i; \; \overline{x^2} = \frac{1}{n} \cdot \sum_{i=1}^{k} n_i x_i^{\,2}$$

$$\overline{y} = \frac{1}{n} \sum_{i=1}^{k} n_i \overline{y}_{x_i}; \; \overline{xy} = \frac{1}{n} \sum_{i=1}^{k} n_i x_i \overline{y}_{x_i}.$$

Лінійна система рівнянь прийме вигляд

$$\begin{cases} \overline{x^2} a + \overline{x} b = \overline{xy} \\ \overline{x} a + b = \overline{y}. \end{cases} \tag{7}$$

Розв'язавши цю систему, можемо знайти параметри a і b, а, отже, і шукане рівняння $\overline{y}_x = ax + b$.

Але більше доцільно ввести нову величину — вибірковий коефіцієнт кореляції; написати рівняння регресії в іншому вигляді. Зробимо це. Знайдемо b із другого рівняння (7):

$$b = \overline{y} - a\overline{x},$$

тоді

$$a = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \left(\overline{x} \right)^2} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x^2} = \rho_{yx}, \tag{8}$$

де $\sigma_x^2 = (\overline{x^2}) - \left(\overline{x} \right)^2$ — дисперсія. Підставимо знайдене значення b у рівняння (3):

$$\overline{y}_x = ax + \overline{y} - a\overline{x}$$

або $$\bar{Y}_x - \bar{y} = a\left(x - \bar{x}\right) = \rho_{yx}\left(x - \bar{x}\right). \tag{9}$$

Величина ρ_{yx} називається коефіцієнтом регресії y на x. Аналогічно за даними таблиці 2.46 можна знайти друге рівняння регресії виду:

$$\bar{X}_y = cy + d \quad \text{або} \quad \bar{X}_y - \bar{x} = \rho_{xy}\left(y - \bar{y}\right), \tag{10}$$

де параметр $\rho_{xy} = c = \dfrac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sigma_y^2}$ називається коефіцієнтом регресії x на y, а σ_y^2 — дисперсія змінної y щодо її загальної середньої $\bar{y}$.

Якщо коефіцієнти регресії ρ_{yx} і ρ_{xy} *додатні*, то кореляційна залежність між x і y — пряма (тому що зі зростання x зростають і умовні середні значення $\bar{y}_x$); якщо коефіцієнти регресії від'ємні, то кореляційна залежність між x і y — обернена (тобто зі зростанням x умовні середні значення $\bar{y}_x$ падають).

Коефіцієнт кореляції. Розглянемо другу задачу: оцінимо силу кореляційного зв'язку між змінним x і y. Така оцінка досягається за допомогою *коефіцієнта кореляції*.

Коефіцієнтом кореляції r змінних x і y називається число рівне середньому геометричному їх коефіцієнтів регресії і має їх знак, тобто

$$r = \pm\sqrt{\rho_{xy}\cdot\rho_{yx}}\,. \tag{11}$$

Отже, коефіцієнт кореляції додатній, якщо коефіцієнти регресії додатні, і від'ємний, якщо вони від'ємні. Підкореневий вираз у рівності (11) завжди додатний, оскільки коефіцієнти регресії, як відомо, мають однакові знаки. Помножимо обидві частини рівності (8) на дріб $\dfrac{\sigma_x}{\sigma_y}$:

$$\rho_{yx}\cdot\frac{\sigma_x}{\sigma_y} = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sigma_x\cdot\sigma_y}\,. \tag{12}$$

Позначимо праву частину рівності через r і назвемо її коефіцієнтом кореляції

$$r = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sigma_x\cdot\sigma_y}\,. \tag{13}$$

Підставимо r в (12) маємо

$$\rho_{yx} = r\frac{\sigma_y}{\sigma_x};$$

Аналогічно

$$\rho_{xy} = r\frac{\sigma_x}{\sigma_y}. \tag{14}$$

Перед дробом залишений тільки знак плюс, оскільки при такому записі знаки коефіцієнта кореляції і коефіцієнта регресії співпадають (див. формули 8 і 13), де знаки визначаються чисельниками, а вони тотожні в цих формулах.

Тепер рівняння регресії (9) і (10) можна записати так:

$$\overline{Y}_x - \overline{y} = r\frac{\sigma_y}{\sigma_x}\left(x - \overline{x}\right); \tag{15}$$

$$\overline{X}_y - \overline{x} = r\frac{\sigma_x}{\sigma_y}\left(y - \overline{y}\right). \tag{16}$$

Зауваження. За великої кількості спостережень одне і те ж число значень *x* можна зустріти n_x разів, одне і те ж значення $y - n_y$ разів, одна і та ж пара чисел $(x; y)$ може спостерігатися n_{xy} разів. Оцінкою коефіцієнта кореляції є вибірковий коефіцієнт кореляції r_B :

$$r_B = \frac{\sum n_{xy}xy - n\overline{x}\,\overline{y}}{n\tilde{\sigma}_x\tilde{\sigma}_y}, \qquad (17)^*$$

де *x*, *y* — варіанти (спостережні значення) ознак *X* і *У*; n_{xy} — частота пари варіант $(x; y)$; n — об’єм вибірки (сума всіх частот); $\tilde{\sigma}_x, \tilde{\sigma}_y$ — вибіркові середні квадратичні відхилення, $\overline{x}, \overline{y}$ — вибіркові середні.

---

[*] Для простоти запису замість $\sum_{i=1}^{k}$ будемо писати $\sum$.

Аналогічно знаходять вибіркові рівняння прямої лінії регресії:

$$\overline{Y}_x - \overline{y} = r_B \frac{\tilde{\sigma}_y}{\tilde{\sigma}_x}\left(x - \overline{x}\right); \tag{18}$$

$$\overline{X}_y - \overline{x} = r_B \frac{\tilde{\sigma}_x}{\tilde{\sigma}_y}\left(y - \overline{y}\right). \tag{19}$$

Якщо $r = 0$ із (15) і (16) випливає $\overline{Y}_x = \overline{y} = const$; $\overline{X}_y = \overline{x} = const$ (при $r = 0$ кореляційного зв'язку не існує, оскільки різним значенням x відповідають однакові умовні значення $\overline{y}_x$ і т.д.).

Відомо, що якщо величини $У$ і X незалежні, то коефіцієнт кореляції $r = 0$ (див.: Розділ I, §9), крім того, лінійного кореляційного зв'язку між x і y не існує, хоча нелінійний зв'язок може існувати. Завжди $-1 \le r \le 1$. Якщо $r = 1$ існує лінійна пряма функціональна залежність між x і y, а при $r = -1$ — обернена. При зростанні $|r|$ від 0 до 1 лінійна кореляційна залежність між x і y наближається до лінійної функціональної залежності, тобто поступово кореляційний зв'язок збільшується.

Коефіцієнт кореляції r вимірює силу (тісноту) *лінійного* зв'язку між $У$ і X. Вибірковий коефіцієнт кореляції r_B є оцінкою коефіцієнта кореляції r генеральної сукупності і тому також слугує для виміру лінійного зв'язку між величинами — кількісними ознаками $У$ і X.

Нелінійні кореляційні залежності. Між змінними x і y можуть існувати і нелінійні кореляційні залежності. Нехай залежність між змінними величинами x і y задані таблицею 2.43, обчислені групові середні $\overline{y}_{x_i}$, які відповідають кожному значенню x_i, а точки $M_i\left(x_i; \overline{y}_{x_i}\right)$ $\left(i = 1, 2, ..., k\right)$ розташовані приблизно на параболі другого порядку.

Рівняння параболи — параболічної регресії y на x будемо шукати у вигляді

$$У = a + bx + cx^2. \tag{20}$$

Із всіх парабол такого виду шукана найближче всього (згідно з принципом найменших квадратів) до точок M_i (рис. 31), причому точка M_i вибирається n_i разів (скільки разів зустрічається у розподілі значень x_i).

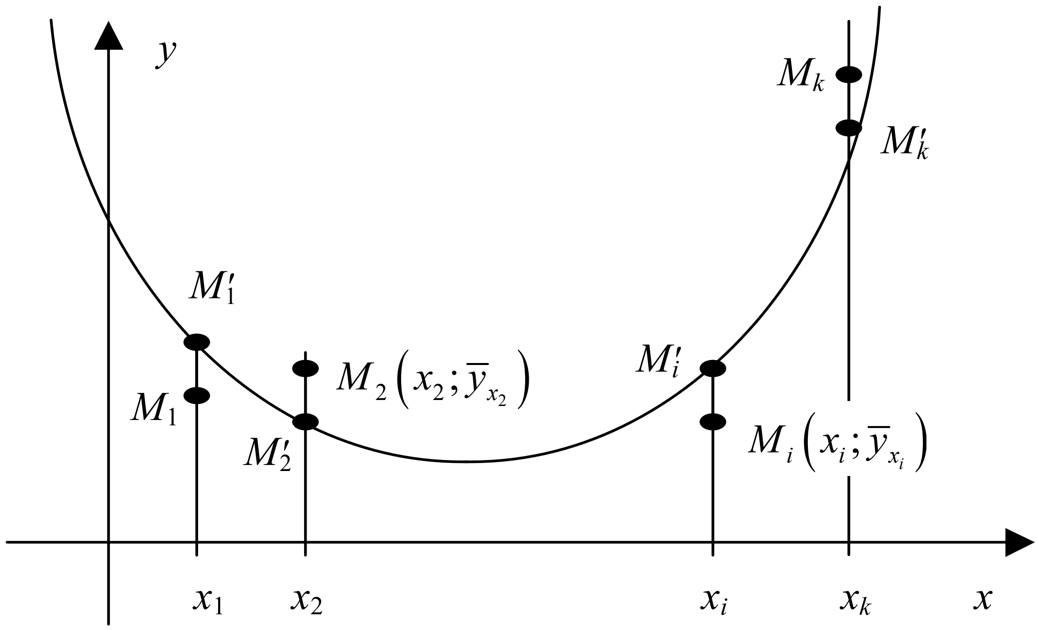


Рис. 31

Невідомі коефіцієнти *a*, *b*, *c* визначимо таким чином, щоб сума відповідних відхилень була мінімальною, тобто

$$n_i\left|a+bx_1+cx_1^2-\overline{y}_{x_1}\right|+$$

$$+n_2\left|a+bx_2+cx_2^2-\overline{y}_{x_2}\right|+\ldots+n_k\left|a+bx_k+cx_k^2-\overline{y}_{x_k}\right|\to\min \qquad (21)$$

Різниці $a+bx_i+cx_i^2-\overline{y}_{x_i}$ $\left(i=1,2,...,k\right)$ називаються похибкою. Застосуємо відомий спосіб найменших квадратів. Для цього складемо функцію:

$$I\left(a,b,c\right)=\sum_{i=1}^{k}n_i\left(a+bx_i+cx_i^2-\overline{y}_{x_i}\right)^2. \qquad (22)$$

Ця функція трьох незалежних змінних *a*, *b* і *c*. Необхідна умова екстремуму функції (рівність нулю частинних похідних за змінними *a*, *b* і *c*) дає три рівняння. Знаходження частинних похідних функції *I* і подальші алгебраїчні перетворення аналогічні виконаним раніше для лінійної функції. А тому наведемо лише кінцевий вигляд системи нормальних рівнянь щодо параметрів *a*, *b*, *c*.

$$\begin{cases} a\sum_{i=1}^{k} n_i + b\sum_{i=1}^{k} x_i n_i + c\sum_{i=1}^{k} x_i^2 n_i = \sum_{i=1}^{k} \overline{y}_{x_i} n_i \\ a\sum_{i=1}^{k} x_i n_i + b\sum_{i=1}^{k} x_i^2 n_i + c\sum_{i=1}^{k} x_i^3 n_i = \sum_{i=1}^{k} x_i \overline{y}_{x_i} n_i \\ a\sum_{i=1}^{k} x_i^2 n_i + b\sum_{i=1}^{k} x_i^3 n_i + c\sum_{i=1}^{k} x_i^4 n_i = \sum_{i=1}^{k} x_i^2 \overline{y}_{x_i} n_i . \end{cases} \tag{23}$$

Приклад 1. Залежність врожайності y (*ц* з 1 *га*) від глибини зрошення x (*см*) наводиться в таблиці 2.47.

Таблиця 2.47

x_i \ y_j	12	14	16	18	Σ
0	5	2	–	–	7
5	–	3	4	3	10
10	–	2	5	6	13
15	–	3	3	4	10
20	–	3	4	1	8
25	2	2	1		5
Σ	7	15	17	14	53

Встановити форму залежності між врожайністю та глибиною зрошення, знайти рівняння регресії.

Щоб з'ясувати форму залежності, обчислимо групові середні змінної y, які відповідають різним значенням x; маємо: $\overline{y}_{x_1} = (12\cdot 5 + 14\cdot 2):7 =$ $= 12{,}57$ *ц*, $\overline{y}_{x_2} = (14\cdot 3 + 16\cdot 4 + 18\cdot 3):10 = 16$ *ц*. Аналогічно $\overline{y}_{x_3} = 16{,}62$ *ц* $\overline{y}_{x_4} = 16{,}2$ *ц* ; $\overline{y}_{x_5} = 15{,}5$ *ц* ; $\overline{y}_{x_6} = 13{,}6$ *ц* .

Групові середні спочатку зростають, а потім спадають. Це дає підстави вважати, що між змінними x і y існує параболічна кореляційна залежність виду (20). Знайдемо параметри цієї функції. У таблиці 2.48 наведені результати допоміжних обчислень, виконаних для визначення коефіцієнтів системи нормальних рівнянь (23).

Звідси, система нормальних рівнянь має вигляд:

$$\begin{cases} 53a + 615b + 10125c = 818 \\ 615a + 10125b + 190125c = 9570 \\ 10125a + 190125b + 3875625c = 154156. \end{cases} \tag{24}$$

Таблиця 2.48

x_i	n_i	$\overline{y}_{x_i}$	$x_i n_i$	$x_i^2 n_i$	$x_i^3 n_i$	$x_i^4 n_i$	$\overline{y}_{x_i} n_i$	$x_i \overline{y}_{x_i} n_i$	$x_i^2 \overline{y}_{x_i} n_i$
0	7	12,57	0	0	0	0	87,93	0	0
5	10	16	50	250	1250	625	160	800	4000
10	13	16,62	130	1300	13000	130000	216,06	2160,6	21606
15	10	16,2	150	2250	33750	506250	162	2430	36450
20	8	15,5	160	3200	64000	1280000	124	2480	49600
25	5	13,6	125	3125	78125	1953125	68	1700	42500
Σ	**53**	—	**615**	**10125**	**190125**	**3875625**	**818**	**9570**	**154156**

Розв'язуючи її методом Гаусса знайдемо коефіцієнти *a, b, c*, які підставляють у рівняння (20). Рекомендуємо розв'язати одержану систему рівнянь самостійно (див.: Розділ III).

У випадку параболічної регресії *x* на *y* необхідно знайти функцію $x = a_1 + b_1 y + c_1 y^2$. У результаті одержуємо систему нормальних рівнянь щодо параметрів a_1; b_1:c_1, в якій порівняно з системою (20) *x* і *y* міняються місцями.

Припустимо, що аналіз залежності між змінними *x* і *y*, вираженої табл. 2.47 приводять до вибору форми кореляційної залежності *y* на *x* у вигляді рівняння гіперболи

$$y = \frac{a}{x} + b\,, \tag{25}$$

а у випадку регресії *x* на *y* — гіперболи

$$x = \frac{c}{y} + d\,. \tag{26}$$

Регресії такого типу називаються *гіперболічними.* Розглянемо спочатку гіперболічну регресію *y* на *x*. Введемо нову змінну $z = \frac{1}{x}$ та представимо рівняння (25) у вигляді $y = az + b$, тобто між *y* і *z* існує лінійна кореляційна залежність. Параметри *a* і *b* цього рівняння визначаються системою (7), в якій лише *x* замінимо на *z*. Повертаючись потім до змінної *x*, для рівняння (25) одержимо систему нормальних рівнянь щодо параметрів *a* і *b*.

$$\begin{cases} a\sum\limits_{i=1}^{k}\frac{1}{x_i}n_i + b\sum\limits_{i=1}^{k}n_i = \sum\limits_{i=1}^{k}\overline{y}_{x_i}n_i \\ a\sum\limits_{i=1}^{k}\frac{1}{x_i^2}n_i + b\sum\limits_{i=1}^{k}\frac{1}{x_i}n_i = \sum\limits_{i=1}^{k}\frac{1}{x_i}\overline{y}_{x_i}n_i. \end{cases} \tag{27}$$

У випадку гіперболічної регресії x на y система нормальних рівнянь для визначення параметрів c і d рівняння (26) знаходяться аналогічно.

Приклад 2. Розподіл однотипних підприємств за об'ємом виробленої за день продукції x та собівартість одиниці продукції y (грн.) наведено в таблиці 2.49.

Встановити форму та знайти рівняння кореляційної залежності між об'ємом продукції та собівартістю одного виробу.

Таблиця 2.49

x_i \ y_i	10	15	20	25	
25	—	—	1	2	3
50	—	2	2	—	4
75	—	5	3	1	9
100	1	3	—	—	4
125	3	1	1	—	5
$\sum n_i$	4	11	7	3	25

Таблиця 2.50

x_i	n_i	$\frac{1}{x_i}$	$\frac{1}{x_i}n_i$	$\frac{1}{x_i^2}n_i$	$\overline{y}_{x_i}n_i$	$\overline{y}_{x_i}\frac{1}{x_i}n_i$
25	3	0,04	0,12	0,0048	69,99	2,8
50	4	0,02	0,08	0,0016	70	1,4
75	9	0,013	0,117	0,0015	160,02	2,08
100	4	0,01	0,04	0,0004	71,12	0,55
125	5	0,008	0,04	0,00032	65	0,52
Σ	**25**	**—**	**0,397**	**0,00862**	**436,13**	**7,35**

Як і раніше, для встановлення форми кореляційної залежності знайдемо групові середні змінної *y*, що відповідає всім значенням змінної *x*; маємо: $\overline{y}_{x_1} = \left(20\cdot 1 + 25\cdot 2\right):3 = 23{,}33$; $\overline{y}_{x_2} = \left(15\cdot 2 + 20\cdot 2\right):4 = 17{,}5$; $\overline{y}_{x_3} = 17{,}78$; $\overline{y}_{x_4} = 13{,}75$; $\overline{y}_{x_5} = 13$.

Зі збільшенням *x* групові середні змінної *y* спадають, причому зменшення затухає. Це дає підстави вважати, що між змінними *x* та *y* існує гіперболічна кореляційна залежність *y* на *x*. Параметри функції (25) знайдемо із системи нормальних рівнянь (27). Результати попередніх обчислень для одержання її коефіцієнтів наведені в таблиці 2.50.

Отже, параметри *a* і *b* функції (25) визначаються системою:

$$\begin{cases} 0{,}397a + 25b = 436{,}13 \\ 0{,}00862a + 0{,}397b = 7{,}35. \end{cases}$$

Розв'язуючи його, знаходимо $a = 183{,}25$, $b = 14{,}54$, а шукане рівняння гіперболічної регресії *y* на *x* має вигляд $y = \dfrac{183{,}25}{x} + 14{,}54$.

Розглянемо нарешті випадок, якщо аналіз зв'язку між змінними *x* та *y*, заданих таблицею, приводить до вибору форми кореляційної залежності *y* на *x* у вигляді показникової функції

$$y = ba^x\,, \tag{28}$$

а при розгляді регресії *x* на *y* — показникової функції

$$x = dc^y\,. \tag{29}$$

Логарифмуючи обидві частини рівності (28) одержимо $\lg y = x\lg a + \lg b$. Отже, якщо між *x* та *y* існує кореляційна залежність *y* на *x* з параметрами *a* і *b*, то між $\lg y$ і *x* — лінійна кореляційна залежність з параметрами $\lg a$ і $\lg b$. Крім того, систему нормальних рівнянь для визначення $\lg a$ і $\lg b$ одержимо із системи (7) заміною *a* і *b* на $\lg a$ і $\lg b$, а $\overline{y}_{x_i}$ — на $\lg \overline{y}_{x_i}$, тобто

$$\begin{cases} \lg a\sum\limits_{i=1}^{k} n_i x_i + \lg b\sum\limits_{i=1}^{k} n_i = \sum\limits_{i=1}^{k} \lg \overline{y}_{x_i} n_i \\ \lg a\sum\limits_{i=1}^{k} n_i x_i^2 + \lg b\sum\limits_{i=1}^{k} n_i x_i = \sum\limits_{i=1}^{k} n_i x_i \lg \overline{y}_{x_i} \end{cases}. \tag{30}$$

Розв'язуючи її, знаходимо $\lg a$ і $\lg b$, а потім і параметри a і b показникової функції (28). Аналогічно можна одержати систему нормальних рівнянь для визначання логарифмів параметрів c і d рівняння (29).

ЗАДАЧІ

Вважаючи, що у всіх випадках між змінними x та y існує лінійна кореляційна залежність потрібно: а) обчислити коефіцієнт регресії; б) обчислити коефіцієнт регресії і розв'язати питання про тісноту зв'язку між вказаними змінними величинами; в) скласти рівняння прямих регресії.

Одержані наступні розподіли:

2.17. Облицьовувальна (будівельна) плитка довжиною x (*см*) та вагою y (*кг*) (таблиця 2.51).

Таблиця 2.51

x \ y	6	8	10	12	14	$\sum$
30	2	17	9	3	—	31
35	—	10	17	9	—	36
40	—	3	24	16	13	56
45	—	—	6	24	12	42
50	—	—	2	11	22	35
$\sum$	2	30	58	63	47	200

Відповідь: а) $\rho_{yx} \approx 0,2186$; $\rho_{xy} \approx 2,184$; б) $r \approx 0,691$;
в) $\overline{y}_x = 0,2186x + 2,4095$; $\overline{x}_y = 2,184y + 15,834$.

2.18. Рослин за масою кожної із них x і за масою насіння y (г) (таблиця 2.52).

Таблиця 2.52

$x \backslash y$	20	30	40	50	60	$\sum$
15	7	5	—	—	—	12
25	20	23	—	—	—	43
35	—	30	47	2	—	79
45	—	10	11	20	6	47
55	—	—	9	7	3	19
$\sum$	27	68	67	29	9	200

Відповідь: а) $\rho_{yx} \approx 0,7333$; $\rho_{xy} \approx 0,7318$; б) $r \approx 0,7325$;

в) $\overline{y}_x = 0,7333x + 9,735$; $\overline{x}_y = 0,7318y + 8,972$.

2.19. Фірм за основними фондами x та за готовою продукцією y (млн грн.) (таблиця 2.53).

Таблиця 2.53

$x \backslash y$	15	20	25	30	35	$\sum$
40	5	7	---	---	---	12
50	---	4	16	23	---	43
60	---	8	20	32	27	87
70	---	---	11	29	2	42
80	---	---	---	9	7	16
$\sum$	5	19	47	93	36	200

Відповідь: а) $\rho_{yx} \approx 0,2268$; $\rho_{xy} \approx 0,9725$; б) $r \approx 0,4697$;

в) $\overline{y}_x = 0,2268x + 4,7126$; $\overline{x}_y = 0,9725y + 32,7310$.

2.20. Підприємств за об'ємом продукції x і за її собівартістю y (грн.) (таблиця 2.54).

Таблиця 2.54

$x \backslash y$	2	2,5	3	3,5	4	$\sum$
1000	—	—	—	2	3	5
2000	—	—	3	6	2	11
3000	—	4	6	3	—	13
4000	1	6	4	1	—	12
5000	6	3	—	—	—	9
$\sum$	7	13	13	12	5	50

Відповідь: а) $\rho_{yx} \approx -0,00040$; $\rho_{xy} \approx -1713$; б) $r \approx -0,8277$; в) $\overline{y}_x = -0,00040x + 4,2220$; $\overline{x}_y = -1713y - 8253$.

Розділ III
ЕЛЕМЕНТИ ЛІНІЙНОЇ АЛГЕБРИ

§1. Основні поняття

Однією із важливих задач математики є дослідження та розв'язання систем рівнянь 1-ої степені. Нехай така система містить m невідомих $x_1, x_2, ..., x_m$, зв'язаних n рівняннями 1-ої степені (число рівнянь може і не співпадати з числом невідомих). У загальному вигляді цю систему можна записати так:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m = b_n \end{cases}. \tag{1}$$

У системі (1) $b_1, b_2, ..., b_n$ — задані вільні члени, a_{ik} (i = 1, 2, ..., n; k = 1, 2, ..., m) — задані коефіцієнти при невідомих. При цьому перший індекс у a_{ik} означає номер рівняння в системі, а другий — номер невідомого, при якому стоїть цей коефіцієнт.

Тобто, a_{ik} — коефіцієнт, який стоїть у i-му рівнянні системи при невідомому $x_к$. Аналогічно b_i означає вільний член в i-му рівнянні системи.

Розв'язком системи (1) називається така сукупність чисел $x_1, x_2, ..., x_m$, яка будучи підставлена у всі рівняння системи (1), обертає ці рівняння в числові рівності.

Запишемо таблицю, складену із коефіцієнтів при невідомих у системі (1):

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}. \tag{2}$$

Таку таблицю, розглядатимемо як єдине ціле і будемо називати *основною матрицею системи.*

Таблиця

$$B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{pmatrix}, \tag{3}$$

яка містить і стовпець, складений із вільних членів системи (1) називається *розширеною матрицею системи.*

Очевидно, що саме існування розв'язку системи (1), так і можливі числові значення елементів розв'язку повністю визначаються матрицями (2) і (3). А тому розглянемо деякі загальні властивості матриць.

Числовою матрицею, або просто матрицею, називається прямокутна таблиця чисел. Окремі числа цієї таблиці називаються елементом матриці. Елементи матриці A позначають символом a_{ik} де i — номер рядка, а k — номер стовпця, в якому стоїть вибраний елемент.

Якщо матриця містить n рядків і m стовпців, тоді говорять, що матриця має розмірність $n \times m$.

Особливо часто доводиться мати справу з матрицями, в яких кількість рядків дорівнює кількості стовпців. Такі матриці називаються *квадратними.*

Кількість рядків (а, отже, кількість стовпців) квадратної матриці називається *порядком* матриці.

§2. Визначники матриць другого порядку

Для квадратної матриці вводиться нове поняття — визначник матриці. Визначник квадратної матриці будемо позначати символом $\det A$ і визначатимемо його індуктивним методом.

Визначником матриці першого порядку (тобто матриці, яка складається із одного елемента, одного числа) називається число, яке утворює задану матрицю.

Визначником матриці другого порядку — $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ — називається число, обчислене за таким правилом: $\det A = a_{11}a_{22} - a_{12}a_{21}$.

Діагональ квадратної матриці, яка йде від лівого верхнього елемента таблиці до правого нижнього, називається *головною діагоналлю* матриці. Діагональ, яка іде від правого верхнього елемента, до лівого нижнього, називається побічною діагоналлю матриці.

Таким чином, для обчислення визначника матриці другого порядку потрібно від добутку елементів, які знаходяться на головній діагоналі матриці, відняти добуток елементів, які знаходяться на побічній діагоналі.

Для визначника матриці $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ вводиться символ $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.

Отже,

$$\det A=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}\,. \qquad (1)$$

Як видно із (1) визначник матриці другого порядку є алгебраїчною сумою двох доданків. Кожний із доданків є добутком двох елементів, причому до нього входить один елемент 1-го рядка та один елемент 2-го рядка, один елемент 1-го стовпця та один елемент 2-го стовпця заданої матриці. Зі знаком “+” береться добуток елементів головної діагоналі і зі знаком “–” — добуток елементів побічної діагоналі.

Із означення (1) легко отримати ряд властивостей визначників матриць.

Властивість 1. *При перестановці рядків матриці на місце стовпців і навпаки визначник матриці не змінюється.*

Нехай задана матриця $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, а матриця $A^*=\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$ отримана із A перестановкою рядків на місце стовпців (така матриця A^* називається часто транспонованою по відношенню до матриці A). Тоді

$$\det A^*=\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}=\det A\,.$$

Властивість 2. *При перестановці двох стовпців (або рядків) абсолютне значення визначника матриці не змінюється, а знак його змінюється на протилежний.*

Нехай задана матриця $A_1=\begin{pmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{pmatrix}$, отримана із A перестановкою стовпців. Тоді

$$\det A_1=\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix}=a_{12}a_{21}-a_{11}a_{22}=-(a_{11}a_{22}-a_{12}a_{21})=-\det A\,.$$

Властивість 3. *Якщо матриця має два однакові стовпці (рядки), тоді визначник матриці дорівнює нулю.*

Властивість 4. *Якщо всі елементи будь-якого стовпця (рядка) матриці помножити на одне і те саме число, тоді визначник матриці виявиться помноженим на те саме число.*

Властивість 5. *Якщо всі елементи будь-якого стовпця (рядка) матриці дорівнюють нулю, тоді визначник матриці дорівнює нулю.*

Властивість 6. *Нехай всі елементи будь-якого стовпця (рядка) матриці A є сумою двох елементів і відповідні стовпці матриць A_1 і A_2 складаються з таких елементів:*

$$A = \begin{pmatrix} a_{11} & + & a'_{11} & a_{12} \\ a_{21} & + & a'_{21} & a_{22} \end{pmatrix},$$

$$A_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},\ A_2 = \begin{pmatrix} a'_{11} & a_{12} \\ a'_{21} & a_{22} \end{pmatrix},$$

тоді

$$\det A = \det A_1 + \det A_2.$$

Таке твердження, як і твердження 3, 4, 5 доводиться безпосередньою перевіркою.

Властивість 7. *Визначник матриці не змінюється, якщо до елементів будь-якого стовпця (рядка) матриці додати величини, пропорційні елементам іншого стовпця (рядка).*

Нехай

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ і } A_1 = \begin{pmatrix} a_{11}+ & ka_{12} & a_{12} \\ a_{21}+ & ka_{22} & a_{22} \end{pmatrix},$$

тоді

$$\det A_1 = \begin{vmatrix} a_{11}+ & ka_{12} & a_{12} \\ a_{21}+ & ka_{22} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} ka_{12} & a_{12} \\ ka_{22} & a_{22} \end{vmatrix} =$$

$$= \det A + k \begin{vmatrix} a_{12} & a_{12} \\ a_{22} & a_{22} \end{vmatrix} = \det A.$$

§3. Визначники матриць третього порядку

Нехай задана будь-яка матриця порядку n

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Виберемо довільний елемент a_{ik} цієї матриці і викреслимо із матриці A той рядок і той стовпець, в яких міститься цей елемент (тобто викреслимо i-ий рядок та k-ий стовпець). Тоді дістанемо матрицю $(n-1)$-го порядку.

Називатимемо її субматрицею матриці A, що відповідає елементу a_{ik} та позначатимемо її символом D_{ik}.

Визначник субматриці D_{ik} назвемо мінором матриці A, що відповідає елементу a_{ik}, та позначимо символом M_{ik}.

Звідси,

$$M_{ik}=\det D_{ik}. \tag{1}$$

Нехай вихідна матриця A була матрицею третього порядку. Тоді дев'ять її можливих субматриць D_{11}, D_{12}, D_{13}, D_{21}, D_{22}, D_{23}, D_{31}, D_{32}, D_{33}, які відповідають різним елементам матриці A будуть матрицями другого порядку.

Означення. *Визначником матриці 3-го порядку*

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

називається число, визначене за таким правилом:

$$\det A=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=a_{11}M_{11}-a_{12}M_{12}+a_{13}M_{13}\,. \tag{2}$$

Оскільки для матриці A

$$D_{11}=\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix},\;\; D_{12}=\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}, D_{13}=\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

тоді

$$M_{11} = \det D_{11} = a_{22}a_{33} - a_{23}a_{32}$$
$$M_{12} = \det D_{12} = a_{21}a_{33} - a_{23}a_{31}$$
$$M_{13} = \det D_{13} = a_{21}a_{32} - a_{22}a_{31}\,.$$

Звідси

$$\det A = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}),$$
$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \quad (3)$$

Як видно із одержаної формули, визначник матриці третього порядку представляє собою алгебраїчну суму шести доданків. Кожний доданок є добутком трьох елементів по одному із кожного рядка і кожного стовпця. Перший з доданків, взятих зі знаком "+", є добутком елементів, розміщених на головній діагоналі матриці. Останні два містять елементи, розташовані у вершинах трикутників з основою, паралельною головній діагоналі.

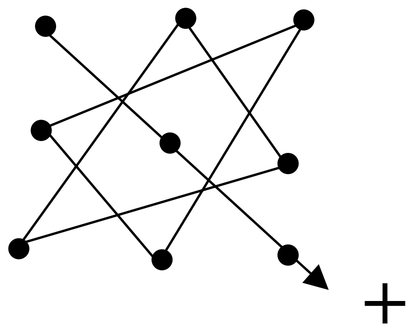


Рис. 1

Добутки, які входять зі знаком "–", містять елементи, розміщені на побічній діагоналі та елементи, розміщені у вершинах трикутників з основою, паралельною побічній діагоналі.

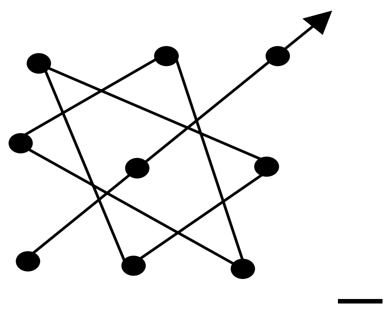


Рис. 2

Безпосередньою перевіркою можна пересвідчитись у тому, що всі властивості, встановлені для визначників матриць другого порядку, мають місце і для визначників третього порядку.

§4. Визначники матриць вищих порядків

Визначники матриць четвертого і більш високих порядків вводитимемо аналогічно визначникам матриць третього порядку. Нехай задана матриця

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Означення. *Визначником матриці A n-го порядку називається число, обчислене за таким правилом:*

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11}M_{11} - a_{12}M_{12} + \dots + (-1)^{n-1}a_{1n}M_{1n}. \quad (1)$$

Відповідно до сказаного вище в цій формулі M_{ik} є мінором матриці A, який відповідає елементу a_{ik}, тобто визначником субматриці D_{ik}, одержаним із матриці A викреслюванням i-того рядка і k-го стовпця.

Правило (1) є очевидним узагальненням правила (2, § 3). Воно дає можливість звести обчислення визначників матриць четвертого порядку до обчислень визначників третього порядку і т.ін.

Обчислення визначника матриці високого порядку за формулою (1) — операція достатньо трудомістка. А тому важливо, використовуючи властивості визначників матриць, скоротити обчислення.

Оскільки визначники матриць введені на основі принципу індукції, тоді природно сподіватися, що властивості визначників матриць вищих порядків будуть ті самі, що і властивості визначників другого порядку. Приймемо без доведення наступні два твердження про властивості визначників матриць n-го порядку (ці властивості були перевірені для визначників матриць другого порядку, а властивість 1 також і для матриць третього порядку).

Властивість 1. *При перестановці рядків матриці на місце стовпців визначник матриці не змінюється.*

Властивість 2. *При перестановці двох рядків або стовпців матриці абсолютне значення визначника матриці не змінюється, а знак змінюється на протилежний.*

Перша властивість дає можливість всі положення, встановлені для рядків матриці, переносити на її стовпці. Наприклад, визначник матриці можливо обчислювати і за форму-лою

$$\det A = a_{11}M_{11} - a_{21}M_{21} + ... + (-1)^{n-1}\, a_{n1}M_{n1} \tag{2}$$

тобто виконувати розклад визначника матриці не тільки за елементами 1-го рядка, але і за елементами 1-го стовпця. На основі тверджень виведемо основну формулу для розкладу визначника матриці за елементами будь-якого рядка або стовпця:

$$\begin{aligned}\det A &= (-1)^{i+1}a_{i1}M_{i1}+(-1)^{i+2}\, a_{i2}M_{i2} +...+ (-1)^{i+n}\, a_{in}M_{in} = \\ &= (-1)^{1+k}a_{1k}M_{1k} + (-1)^{2+k}a_{2k}M_{2k}+...+(-1)^{n+k}\, a_{nk}M_{nk}\end{aligned} \tag{3}$$

Перша частина формули є розкладом визначника за елементами i-го рядка, друга половина — розклад за елементами k-го стовпця.

Доведення. Нехай задана матриця n-го порядку

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{i-11} & a_{i-12} & ... & a_{i-1n} \\ a_{i1} & a_{i2} & ... & a_{in} \\ a_{i+11} & a_{i+12} & ... & a_{i+1n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix}.$$

Переставимо i-тий і (i – 1)-ий рядок матриці A, потім переставимо новий (i – 1)-ий рядок з (i – 2)-им рядком і т.д. до тих пір, доки вихідний i-ий рядок не стане 1-им рядком. Тоді дістанемо нову матрицю:

$$A' = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \\ a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i-11} & a_{i-12} & \dots & a_{i-1n} \\ a_{i+11} & a_{i+12} & \dots & a_{i+1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Матриця A' одержана із A у результаті $(i - 1)$ перестановок рядків. Кожна така перестановка у визначника матриці змінює тільки знак. Отже, $(-1)^{i-1}\det A' = \det A$.

Субматриця D'_{ik} матриці A', яка відповідає елементу a_{ik}, повністю співпадає з субматрицею D_{ik}, яка відповідає елементу a_{ik} матриці A.

Тому

$$\det D'_{ik} = \det D_{ik} = M_{ik}$$

і на основі означення (1) дістанемо

$$\det A' = a_{i1}M_{i1} - a_{i2}M_{i2} + \dots + (-1)^{n-1}a_{in}M_{in}$$

Таким чином,

$$\det A = (-1)^{i-1}\det A' = (-1)^{i+1}\det A' =$$
$$= (-1)^{i+1}a_{i1}M_{i1} + (-1)^{i+2}a_{i2}M_{i2} + \dots + (-1)^{i+n}a_{in}M_{in}$$

і перша половина формули (3) одержана.

Другу половину формули дістанемо, використовуючи можливість перестановки рядків на місце стовпців.

Для спрощення запису формул розкладу зручно ввести поняття “алгебраїчне доповнення, яке відповідає заданому елементу матриці”.

Означення. *Алгебраїчним доповненням елемента a_{ik} матриці A називається мінор M_{ik} цієї матриці, помножений на* $(-1)^{i+k}$.

Алгебраїчне доповнення елемента a_{ik} матриці A позначається, як правило, символом A_{ik}. Отже,

$$A_{ik} = (-1)^{i+k}M_{ik}. \tag{4}$$

Тепер формулу розкладу (3) можливо записати в такому вигляді:

$$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} = a_{1k}A_{1k} + a_{2k}A_{2k} + \dots + a_{nk}A_{nk}. \tag{5}$$

Визначник матриці дорівнює сумі добутків елементів будь-якого рядка або стовпця матриці на відповідне алгебраїчне доповнення.

Для скорочення запису часто вводиться спеціальне позначення суми ряду доданків. Суму $\alpha_1 + \alpha_2 + ... + \alpha_n$ будемо записувати символом $\sum_{k=1}^{n}\alpha_k$. Знак $\sum$ в такому записі є символом підсумовування, запис $k=1$ під знаком суми і n над знаком $\sum$ означає, що індекс к в окремих доданках набуває значення від 1 до n. Очевидно, що якщо p будь-яке число, тоді

$$\sum_{k=1}^{n} p\alpha_k = p\sum_{k=1}^{n}\alpha_k ,$$

тобто сталий множник можна винести за знак суми. Вкажемо ще, що

$$\sum_{k=1}^{n}(a_k + b_k)c_k = \sum_{k=1}^{n}a_k c_k + \sum_{k=1}^{n}b_k c_k .$$

Дійсно,

$$\sum_{k=1}^{n}(a_k + b_k)c_k = (a_1 + b_1)c_1 + (a_2 + b_2)c_2 + ... + (a_n + b_n)c_n =$$
$$= (a_1c_1 + b_1c_1 + a_2c_2 + b_2c_2 + ... + a_nc_n + b_nc_n) =$$
$$= (a_1c_1 + a_2c_2 + ... + a_nc_n) + (b_1c_1 + b_2c_2 + ... + b_nc_n) = \sum_{k=1}^{n}a_k c_k + \sum_{k=1}^{n}b_k c_k .$$

Використовуючи символ підсумовування формулу (5) записують у вигляді

$$\det A = \sum_{k=1}^{n}a_{ik}A_{ik} = \sum_{i=1}^{n}a_{ik}A_{ik} . \qquad (5')$$

За допомогою формули розкладу легко встановити інші властивості визначників матриць будь-якого порядку.

***Властивість* 3.** *Якщо матриця має два однакові рядки або стовпця, тоді визначник матриці дорівнює нулю.*

Доведення. Нехай у матриці A співпадають l-ий і s-ий рядок. Переставимо в цій матриці s-ий рядок на місце l-го і навпаки. Оскільки ці рядки однакові, тоді ні матриця, ні її визначник не зміняться. Але разом з тим при перестановці в матриці двох будь-яких рядків визначник матриці змінить свій знак, тобто det A= –det A, і тому det A = 0.

Властивість 4. *Якщо всі елементи будь-якого рядка або стовпця матриці помножити на одне і те саме число, тоді визначник матриці стане помноженим на те саме число.*

Доведення. Розглянемо матриці

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad A'=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ pa_{i1} & pa_{i2} & \dots & pa_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Розкладаючи визначник цих матриць за елементами i-го рядка, дістанемо:

$$\det A=\sum_{k=1}^{n} a_{ik}A_{ik}\ ; \quad \det A'=\sum_{k=1}^{n} pa_{ik}A_{ik}=p\sum_{k=1}^{n} a_{ik}A_{ik}\ .$$

Отже, $\det A'=p\det A$.

Властивість 5. *Якщо всі елементи будь-якого рядка або стовпця матриці дорівнюють нулю, тоді визначник матриці дорівнює нулю.*

Доведення. Розкладаючи визначник матриці за елементами рядка, всі елементи якого дорівнюють нулю, одержимо цю властивість.

Властивість 6. *Нехай всі елементи будь-якого рядка або стовпця матриці A є сумою двох доданків і нехай відповідні рядки матриць A_1 і A_2 складаються з цих доданків:*

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1}+\alpha_1 & a_{i2}+\alpha_2 & \dots & a_{in}+\alpha_n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_1 & \alpha^2 & \dots & \alpha^n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

тоді $\det A = \det A_1 + \det A_2$.

Доведення. Розкладаючи визначники матриць A_1, A_2, A за елементами i-го рядка одержимо

$$\det A = \sum_{k=1} (a_{ik} + \alpha_k) A_{ik} = \sum_{k=1} a_{ik} A_{ik} + \sum_{k=1} \alpha_k A_{ik} = \det A_1 + \det A_2$$

Властивість 7. *Визначник матриці не зміниться, якщо до елементів будь-якого рядка (або стовпця) матриці додати величини, пропорційні елементам іншого рядка (стовпця) тієї ж матриці.*

Доведення. Розглянемо матриці

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1} + pa_{l1} & a_{i2} + pa_{l2} & \dots & a_{in} + pa_{ln} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

і матрицю B, одержану заміною в матриці A i-го рядка l-им рядком ($i \neq l$)

Згідно доведеному раніше $\det A' = \det A + p\det B$. Але в матриці B рядок $a_{l1}, a_{l2}, \dots, a_{ln}$ входить двічі (на i-му і на l-му місцях) і тому $\det B = 0$, тобто $\det A' = \det A$.

Встановлені властивості дають практичний метод обчислення визначників матриць високих порядків. Покажемо на декількох прикладах зміст цього методу.

Приклад 1. Задана матриця:

$$A=\begin{pmatrix}1 & -2 & 3\\ 2 & 1 & 1\\ 3 & -2 & 2\end{pmatrix}.$$

Знайти $\det A$.

Розв'язок. Перетворимо задану матрицю так, щоб в одному рядку або в одному стовпці всі елементи, крім одного, стали рівними нулю. Перетворення при цьому проводимо такі, щоб визначник матриці не змінювався. Для цього до елементів 1-го і 3-го рядків додаємо подвоєні елементи 2-го рядка.

$$\det A=\begin{vmatrix}1 & -2 & 3\\ 2 & 1 & 1\\ 3 & -2 & 2\end{vmatrix}=\begin{vmatrix}5 & 0 & 5\\ 2 & 1 & 1\\ 7 & 0 & 4\end{vmatrix}.$$

Розкладаючи визначник за елементами 2-го стовпця за формулою (3) одержимо:

$$\det A=-0\begin{vmatrix}2 & 1\\ 7 & 4\end{vmatrix}+1\begin{vmatrix}5 & 5\\ 7 & 4\end{vmatrix}-0\begin{vmatrix}5 & 5\\ 2 & 1\end{vmatrix}=\begin{vmatrix}5 & 5\\ 7 & 4\end{vmatrix}=-15\,.$$

Приклад 2. Задана матриця:

$$A=\begin{pmatrix}2 & 3 & -1 & 2\\ 1 & 2 & -3 & 1\\ 4 & -2 & 1 & 0\\ 3 & 2 & 2 & 3\end{pmatrix}$$

Знайти $\det A$.

Розв'язок. Перетворимо матрицю так, щоб всі елементи 3-го рядка, крім одного, обертались на нуль. Для цього до елементів 1-го стовпця додамо елементи 3-го, помножені на –4, а до елементів 2-го стовпця додаємо елементи 3-го, помножені на 2. Тоді будемо мати:

$$\det A=\begin{vmatrix}2 & 3 & -1 & 2\\ 1 & 2 & -3 & 1\\ 4 & -2 & 1 & 0\\ 3 & 2 & 2 & 3\end{vmatrix}=\begin{vmatrix}6 & 1 & 1 & 2\\ 13 & -4 & -3 & 1\\ 0 & 0 & 1 & 0\\ -5 & 6 & 2 & 3\end{vmatrix}=\begin{vmatrix}6 & 1 & 2\\ 13 & -4 & 1\\ -5 & 6 & 3\end{vmatrix}.$$

Додаємо до елементів 1-го рядка елементи 2-го рядка, помножені на –2, і до елементів 3-го елементи 2-го, помножені на — 3. Тоді:

$$\det A = \begin{vmatrix} -20 & 9 & 0 \\ 13 & -4 & 1 \\ -44 & 18 & 0 \end{vmatrix} = -\begin{vmatrix} -20 & 9 \\ -44 & 18 \end{vmatrix} = 4 \cdot 9 \begin{vmatrix} 5 & 1 \\ 11 & 2 \end{vmatrix} = -36 .$$

Доведемо, на довершення, ще одну суттєву теорему, яку назвемо *теоремою про алгебраїчні доповнення.*

Теорема. *Сума добутків елементів будь-якого рядка (або стовпця) матриці на алгебраїчне доповнення, відповідні елементам іншого рядка (або стовпця) цієї ж матриці, дорівнює нулю.*

Доведення. Нехай задана матриця:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{l1} & a_{l2} & \dots & a_{ln} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Розглянемо допоміжну матрицю:

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ h_1 & h_2 & \dots & h_n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

одержану із матриці A заміною елементів l-го рядка числами $h_1, h_2, \dots h_n$.

Розкладаючи $\det A_1$ за елементами l-го рядка, дістанемо $\det A_1 = \sum_{k=1}^{n} h_k A_{lk}$, де A_{lk} — алгебраїчні доповнення, відповідні елементам l-го рядка як матриці A_1 так і матриці A. Якщо $h_k = a_{ik}$ одержимо матрицю, в якій два рядки повністю співпадають. Тому при $i \neq l$ визначник цієї матриці, тобто $\sum_{k=1}^{n} a_{ik} A_{ik_1}$ має дорівнювати нулю. Теорема доведена.

Об'єднуючи доведену теорему і формулу (5), можливо записати

$$\sum_{k=1}^{n} a_{ik} A_{lk_1} = \begin{cases} \det A & i = l \\ 0 & i \neq l \end{cases} \tag{6}$$

Аналогічно, проводячи розклад за елементами i-го стовпця, маємо

$$\sum_{i=1}^{n} a_{ik} A_{il} = \begin{cases} \det A & k = l \\ 0 & k \neq l \end{cases} \tag{7}$$

ЗАДАЧІ

3.1. Обчислить визначник другого порядку:

$$\begin{vmatrix} 5 & 3 \\ 6 & 4 \end{vmatrix}.$$

$$\begin{vmatrix} 5 & 3 \\ 6 & 4 \end{vmatrix} = 5 \cdot 4 - 6 \cdot 3 = 2$$

3.2. Обчислить визначник третього порядку:

$$\begin{vmatrix} 2 & 2 & 1 \\ 1 & 2 & 1 \\ 3 & 1 & 0 \end{vmatrix}$$

Згідно з правилом трикутника, маємо:

$$\begin{vmatrix} 2 & 2 & 1 \\ 1 & 2 & 1 \\ 3 & 1 & 0 \end{vmatrix} = 2{\cdot}2{\cdot}0 + 2{\cdot}1{\cdot}3 + 1{\cdot}1{\cdot}1 - 3{\cdot}2{\cdot}1 - 1{\cdot}1{\cdot}2 - 1{\cdot}2{\cdot}0 = 0 + 6 + 1 - 6 - 2 - 0 = -1.$$

3.3. Обчислить визначник третього порядку:

$$\begin{vmatrix} 5 & 3 & 2 \\ -1 & 2 & 4 \\ 7 & 3 & 6 \end{vmatrix}.$$

Розклавши визначник за елементами 1-го рядка, одержимо:

$$\begin{vmatrix} 5 & 3 & 2 \\ -1 & 2 & 4 \\ 7 & 3 & 6 \end{vmatrix} = 5\begin{vmatrix} 2 & 4 \\ 3 & 6 \end{vmatrix} - 3\begin{vmatrix} -1 & 4 \\ 7 & 6 \end{vmatrix} + 2\begin{vmatrix} -1 & 2 \\ 7 & 3 \end{vmatrix} = 5\cdot 0 - 3(-34) + 2(-17) = 68.$$

3.4. Обчислить попередній приклад інакше, використавши властивості визначників.

До елементів 1-го рядка додамо відповідні елементи 2-го рядка, помножені на 5, а до елементів 3-го рядка відповідні елементи 2-го рядка, помножені на 7.

$$\begin{vmatrix} 5 & 3 & 2 \\ -1 & 2 & 4 \\ 7 & 3 & 6 \end{vmatrix} = \begin{vmatrix} 0 & 13 & 22 \\ -1 & 2 & 4 \\ 7 & 3 & 6 \end{vmatrix} = \begin{vmatrix} 0 & 13 & 22 \\ -1 & 2 & 4 \\ 0 & 17 & 34 \end{vmatrix}.$$

Розклавши визначник за елементами 1-го стовпця, одержимо:

$$\begin{vmatrix} 0 & 13 & 22 \\ -1 & 2 & 4 \\ 0 & 17 & 34 \end{vmatrix} = 0\cdot\begin{vmatrix} 2 & 4 \\ 17 & 34 \end{vmatrix} - (-1)\begin{vmatrix} 13 & 22 \\ 17 & 34 \end{vmatrix} + 0\cdot\begin{vmatrix} 13 & 22 \\ 2 & 4 \end{vmatrix} = 13\cdot 34 - 17\cdot 22 = 68$$

.

3.5. Обчислити визначник

$$D = \begin{vmatrix} 3 & 5 & 7 & 2 \\ 1 & 2 & 3 & 4 \\ -2 & -3 & 3 & 2 \\ 1 & 3 & 5 & 4 \end{vmatrix}.$$

Виконаємо такі дії: 1) до елементів 1-го рядка додамо помножені на –3 відповідні елементи 2-го рядка; 2) до елементів 3-го рядка додамо подвоєні елементи 2-го рядка; 3) до елементів 4-го рядка додамо відповідні елементи 2-го рядка, помножені на –1. Тоді вихідний визначник перетвориться до вигляду:

$$D=\begin{vmatrix}0 & -1 & -2 & -10\\ 1 & 2 & 3 & 4\\ 0 & 1 & 9 & 10\\ 0 & 1 & 2 & 0\end{vmatrix}.$$

Розклавши цей визначник за елементами 1-го стовпця, маємо:

$$D=-\begin{vmatrix}-1 & -2 & -10\\ 1 & 9 & 10\\ 1 & 2 & 0\end{vmatrix}.$$

Додамо до елементів 1-го рядка елементи 3-го рядка і віднімемо від елементів 2-го рядка елементи 3-го рядка, одержимо:

$$D=-\begin{vmatrix}0 & 0 & -10\\ 0 & 7 & 10\\ 1 & 2 & 0\end{vmatrix}.$$

Розкладемо визначник за елементами 1-го стовпця.

$$D=-\begin{vmatrix}0 & -10\\ 7 & 10\end{vmatrix}=-70$$

3.6. Обчислити визначник п’ятого порядку

$$D=\begin{vmatrix}10 & -2 & -2 & -1 & 3\\ 0 & -1 & -4 & -7 & 2\\ -1 & 3 & 3 & 5 & -5\\ 6 & 2 & -2 & 1 & 2\\ -3 & 0 & -1 & 2 & 3\end{vmatrix}$$

Для перетворення в нуль всіх елементів (крім одного) будь-якого рядка чи стовпця, вибираємо той рядок або стовпчик, який складається з найменших чисел. У визначнику таким буде 2-й стовпчик. Залишимо в ньому без змін елемент $a_{22} = -1$, а решту перетворимо в нулі. Для цього виконаємо: ($I_р + (-2)\cdot II_р$, $III_р + 3\ II_р$, $IV_р + 2\ II_р$)[1].

Одержимо:

$$D = \begin{vmatrix} 10 & 0 & 6 & 13 & -1 \\ 0 & -1 & -4 & -7 & 2 \\ -1 & 0 & -9 & -16 & 1 \\ 6 & 0 & -10 & -13 & 6 \\ -3 & 0 & -1 & 2 & 3 \end{vmatrix}.$$

Розкладемо визначник за елементами 2-го стовпця.

$$D = (-1)\begin{vmatrix} 10 & 6 & 13 & -1 \\ -1 & -9 & -16 & 1 \\ 6 & -10 & -13 & 6 \\ -3 & -1 & 2 & 3 \end{vmatrix}.$$

В одержаному визначнику вже четвертого порядку з найменших елементів складається 4-ий рядок. Перетворимо в нулі всі його елементи, крім $a_{42} = -1$. Для цього виконаємо:

($I_{ст} + (-3)II_{ст}$, $III_{ст}+2\ II_{ст}$, $IV_{ст}+3\ II_{ст}$). У результаті одержимо:

$$D = -\begin{vmatrix} -8 & 6 & 25 & 17 \\ 26 & -9 & -34 & -26 \\ 36 & -10 & -33 & -24 \\ 0 & -1 & 0 & 0 \end{vmatrix}.$$

Розкладемо визначник за елементами 4-го рядка

$$D = -(-1)\begin{vmatrix} -8 & 25 & 17 \\ 26 & -34 & -26 \\ 36 & -33 & -24 \end{vmatrix} = 6\begin{vmatrix} -8 & 25 & 17 \\ 13 & -17 & -13 \\ 12 & -11 & -8 \end{vmatrix}.$$

(Ми винесли за знак визначника спільний множник з елементів 2-го рядка і спільний множник з елементів 3-го рядка).

---

[1] Тут і далі: р — рядок, ст — стовпець.

Для зменшення елементів цього визначника додамо 1-й стовпець до 2-го та 3-го:

$$D = 6\begin{vmatrix} -8 & 17 & 9 \\ 13 & -4 & 0 \\ 12 & 1 & 4 \end{vmatrix} = 6\left\{(-1)^{1+3}\cdot 9\begin{vmatrix} 13 & -4 \\ 12 & 1 \end{vmatrix} + (-1)^{3+3}\cdot 4\begin{vmatrix} -8 & 17 \\ 13 & -4 \end{vmatrix}\right\} = -1122$$

Останній визначник розклали за елементами 3-го стовпця.

3.7. Обчислити визначник n-го порядку, звівши його до трикутного вигляду:

$$D = \begin{vmatrix} 5 & 3 & 3 & \dots & 3 \\ 3 & 5 & 3 & \dots & 3 \\ 3 & 3 & 5 & \dots & 3 \\ \dots & \dots & \dots & \dots & \dots \\ 3 & 3 & 3 & \dots & 5 \end{vmatrix}.$$

Віднімемо 1-й рядок від усіх останніх.

$$D = \begin{vmatrix} 5 & 3 & 3 & \dots & 3 & 3 \\ -2 & 2 & 0 & \dots & 0 & 0 \\ -2 & 0 & 2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -2 & 0 & 0 & \dots & 0 & 2 \end{vmatrix}$$

До 1-го стовпця додамо суму всіх останніх.

$$D = \begin{vmatrix} 5+3(n-1) & 3 & 3 & 3 & \dots & 3 \\ 0 & 2 & 0 & 0 & \dots & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 2 \end{vmatrix} = (3n+2)2^{n-1}.$$

Обчислити визначники:

3.8. $\begin{vmatrix} -1 & 4 \\ -5 & 2 \end{vmatrix}$.

3.9. $\begin{vmatrix} a+b & a-b \\ a-b & a+b \end{vmatrix}$

3.10. $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$.

3.11. $\begin{vmatrix} 3 & 4 & -5 \\ 8 & 7 & -2 \\ 2 & -1 & 8 \end{vmatrix}$.

3.12. $\begin{vmatrix} 1 & -2 & 3 & 4 \\ 2 & 1 & -4 & 3 \\ 3 & -4 & -1 & -2 \\ 4 & 3 & 2 & -1 \end{vmatrix}$.

3.13. $\begin{vmatrix} \frac{3}{2} & -\frac{9}{2} & -\frac{3}{2} & -3 \\ \frac{5}{3} & -\frac{8}{3} & -\frac{2}{3} & -\frac{7}{3} \\ \frac{4}{3} & -\frac{5}{3} & -1 & -\frac{2}{3} \\ 7 & -8 & -4 & -5 \end{vmatrix}$.

3.14. $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$.

3.15. $\begin{vmatrix} 0 & 1 & 2 & 3 & 5 \\ 2 & -1 & 2 & -1 & 2 \\ 1 & 1 & 1 & -1 & 3 \\ 1 & 2 & 5 & 1 & 1 \\ -1 & 4 & 0 & 1 & 4 \end{vmatrix}$.

3.16. $\begin{vmatrix} 5 & 5 & 5 & \dots & 5 \\ 5 & 0 & 5 & \dots & 5 \\ 5 & 5 & 0 & \dots & 5 \\ \dots & \dots & \dots & \dots & \dots \\ 5 & 5 & 5 & \dots & 0 \end{vmatrix}$.

3.17. $\begin{vmatrix} 1 & 2 & 3 & \dots & n \\ -1 & 0 & 3 & \dots & n \\ -1 & -2 & 0 & \dots & n \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -2 & -3 & \dots & 0 \end{vmatrix}$.

3.18. $\begin{vmatrix} x & x & x & \dots & x & x & 1 \\ x & x & x & \dots & x & 2 & x \\ x & x & x & \dots & 3 & x & x \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x & n & x & \dots & x & x & x \\ x & x & x & \dots & x & x & x \end{vmatrix}$.

3.19. $\begin{vmatrix} -5 & 1 & -4 & 1 \\ 1 & 4 & -1 & 5 \\ -4 & 1 & -8 & -1 \\ 3 & 2 & 6 & 2 \end{vmatrix}$.

§5. Розв'язування систем *n* рівнянь із *n* невідомими

Правило Крамера. Встановивши основні властивості і способи обчислення визначників матриць будь-якого порядку, повернімося до основної задачі — розв'язуванню і дослідженню систем рівнянь першого степеня. Почнемо вивчення цього питання із розбору того основного випадку, коли кількість рівнянь співпадає з кількістю невідомих.

Нехай задана система:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}. \tag{1}$$

Приймемо, що система (1) має *який-небудь* розв'язок і потрібно тільки знайти його. Позначимо через A основну матрицю системи:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

Помножимо всі члени 1-го рівняння системи (1) на A_{11} — алгебраїчне доповнення елемента a_{11} матриці A, всі члени 2-го рівняння системи (1) на A_{21} — алгебраїчне доповнення елемента a_{21} матриці A, нарешті, всі члени n-го рівняння системи (1) на A_{n1} — алгебраїчне доповнення елементу a_{n1} матриці A. Тоді дістанемо систему:

$$\begin{cases} a_{11}A_{11}x_1 + a_{12}A_{11}x_2 + \ldots + a_{1n}A_{11}x_n = b_1A_{11} \\ a_{21}A_{21}x_1 + a_{22}A_{21}x_2 + \ldots + a_{2n}A_{21}x_n = b_2A_{21} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}A_{n1}x_1 + a_{n2}A_{n1}x_2 + \ldots + a_{nn}A_{n1}x_n = b_nA_{n1} \end{cases}. \tag{1'}$$

Додамо почленно всі рівняння системи, одержимо:

$$(\sum_{i=1}^{n} a_{i1}A_{i1})x_1 + (\sum_{i=1}^{n} a_{i2}A_{i2})x_2 + \ldots + (\sum_{i=1}^{n} a_{in}A_{in})x_n = \sum_{i=1}^{n} b_iA_{i1}\,.$$

Згідно з теоремою про алгебраїчні доповнення маємо:

$$\sum_{i=1}^{n} a_{i1}A_{i1} = \det A, \quad \sum_{i=1}^{n} a_{i2}A_{i1} = 0, \ldots, \quad \sum_{i=1}^{n} a_{in}A_{i1} = 0.$$

Тому одержане рівняння можливо переписати у вигляді

$$x_1 \cdot \det A = = \sum_{i=1}^{n} b_i A_{i1}.$$

Розглянемо матрицю:

$$B_1 = \begin{pmatrix} b_1 & a_{12} & \ldots & a_{1n} \\ b_2 & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ b_n & a_{n2} & \ldots & a_{nn} \end{pmatrix},$$

отриману із матриці A заміною елементів 1-го стовпця стовпцем, вільних членів рівнянь системи. Розкладаючи $\det B_1$ за елементами 1-го стовпця, одержимо $\det B_1 = \sum_{i=1}^{n} b_i A_{i1}.$, а тому $x_1 \cdot \det A = \det B_1$.

Аналогічно, помноживши рівняння системи (1) відповідно на A_{i2} ($i = 1, 2, \ldots, n$) і додаючи їх одержимо $x_2 \cdot \det A = \det B_2$,
де

$$B_2 = \begin{pmatrix} a_{11} & b_1 & \ldots & a_{1n} \\ a_{21} & b_2 & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & b_n & \ldots & a_{nn} \end{pmatrix}.$$

Роблячи так само і в подальшому, одержимо систему рівнянь:

$$\begin{cases} x_1 \cdot \det A = \det B_1 \\ x_2 \cdot \det A = \det B_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n \cdot \det A = \det B_n \end{cases}, \qquad (2)$$

де матрицю B_k одержано із A заміною k-го стовпця стовпцем вільних членів. Очевидно, будь-який розв’язок системи (1) є і розв’язком системи (2).

Покладемо тепер, що визначник основної матриці системи не дорівнює нулю. Тоді система (2) має єдиний розв’язок

$$x_i = \frac{\det B_i}{\det A} \quad i = 1, 2, ..., n . \tag{3}$$

Нагадаємо, що формулу (3) одержано, з припущенням, що система (1) має розв'язок. Безпосередньою підстановкою знайдених значень x_i у систему (1) можливо переконатися в тому, що вони є розв'язком системи (1) і, отже, в припущенні, що $\det A \neq 0$, система (1) має розв'язок і при тому єдиний.

Теорема (теорема Крамера): *Якщо визначник основної матриці системи n рівнянь 1-го степеня з n невідомими відмінний від нуля, тоді система має єдиний розв'язок. При цьому значення кожного із невідомих дорівнює частці від ділення визначників двох матриць: у знаменнику стоїть визначник основної матриці системи, а у чисельнику визначник матриці, одержаної із основної матриці системи заміною стовпця, що відповідає вибраному невідомому, стовпцем вільних членів.*

Із цієї теореми виходить, що якщо система рівнянь однорідна, тобто вільні члени у всіх рівняннях системи дорівнюють нулю, і якщо визначник основної матриці системи відмінний від нуля, тоді система має тільки нульовий розв'язок. Дійсно, в такому випадку, матриці, визначники яких стоять у чисельнику формули (3), містять стовпець, який включає у себе лише нулі, і, отже всі числа x_i дорівнюють нулю. Із доведеного випливає наступна теорема:

Якщо система n однорідних рівнянь 1-го степеня з n невідомими має хоча б один ненульовий розв'язок, тоді визначник основної матриці системи дорівнює нулю.

Дійсно, якби цей визначник був не рівний нулю, тоді система мала б тільки нульовий розв'язок, що суперечить умові.

Надалі ми доведемо, що рівність нулю визначника системи є не тільки обов'язковою та необхідною умовою існування ненульового розв'язку, але і умовою, достатньою для існування такого розв'язку. Інакше кажучи, якщо визначник системи однорідних рівнянь дорівнює нулю, тоді система має ненульовий розв'язок (і при цьому нескінчену множину таких розв'язків).

Розв'язування і дослідження систем рівнянь першого порядку методом повного виключення (Метод Гаусса). Формули Крамера дають можливість, використовуючи прийом обчислення визначників, знайти числові значення розв'язку системи рівнянь у випадку, коли визначник основної матриці системи відмінний від нуля. Але практичне застосування цих формул у багатьох випадках ускладнено. Перш за все, необхідно зазначити, що для знаходження розв'язків за формулами (3) слід обчислити n+1 визначник n-го порядку, що є досить трудомістким, навіть при використанні тих прийомів, які були описані в §4. Але, основним є те, що у випадку, коли коефіцієнти рівняння задані наближено (у реальних задачах це буває майже завжди), похибка розв'язку може бути досить велика. Це пояснюється тим, що доданки, які входять у кожний із визначників, через які визначається розв'язок системи, можуть бути досить великі (нагадаємо, вони є добутком n співмножників — різних коефіцієнтів розширеної матриці системи), а визначник, що є алгебраїчною сумою таких доданків, може бути малим. Навіть в тому випадку, коли коефіцієнти в системі вихідних рівнянь відомі точно, але самі обчислення ведуться з урахуванням лише заданого числа значущих цифр, в результаті ми з тих же причин зможемо одержать досить великі похибки. А тому при практичному розв'язуванні систем рівнянь переважно використовують не формули Крамера, а інші прийоми обчислень.

У нашому курсі ми розглянемо метод повного виключення щодо розв'язування систем рівнянь першого степеня також за умови, коли число рівнянь не співпадає з числом невідомих. Але викладення цього методу почнемо з основного випадку: якщо кількість рівнянь співпадає з числом невідомих.

Таким чином, нехай знову задана система n рівнянь з n невідомими:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}. \tag{4}$$

Оскільки хоча б один із коефіцієнтів a_{i1} відмінний від нуля (інакше до системи взагалі не входило б x_1) і рівняння в системі можливо міняти місцями, тоді без будь-якого обмеження загальності можна покласти, що $a_{11}\neq 0$. Поділимо перше рівняння системи на a_{11} і приведемо його до вигляду $x_1 + a_{12}^{(1)}x_2 + \ldots + a_{1n}^{(1)}x_n = b_1^{(1)}$,
де

$$a_{1k}^{(1)} = \frac{a_{1k}}{a_{11}}, \qquad b_1^{(1)} = \frac{b_1}{a_{11}},$$

Перемножуючи всі члени одержаного рівняння на a_{i1} і віднімаючи із i-го рівняння системи (4), дістанемо нову систему:

$$\begin{cases} 1\cdot x_1 + a_{12}^{(1)}x_2 + ... + a_{1n}^{(1)}x_n = b_1^{(1)} \\ 0\cdot x_1 + a_{22}^{(1)}x_2 + ... + a_{2n}^{(1)}x_n = b_2^{(1)} \\ .... \\ 0\cdot x_1 + a_{n2}^{(1)}x_2 + ... + a_{nn}^{(1)}x_n = b_n^{(1)} \end{cases}, \tag{5}$$

де

$$a_{ik}^{(1)} = a_{ik} - a_{i1}a_{1k}^{(1)}, \quad b_i^{(1)} = b_i - a_{i1}b_1^{(1)}$$

$i = 2, ..., n$; $k = 1, 2, ..., n$.

Оскільки рівняння системи (5) одержані як лінійні комбінації рівнянь системи (4), то будь-який розв'язок системи (4) є також і розв'язком системи (5). Разом з тим оскільки

$$a_{1k} = a_{1k}^{(1)}a_{11}, \quad b_1 = b_1^{(1)}a_{11}, \quad a_{ik} = a_{ik}^{(1)} + a_{i1}a_{1k}^{(1)}, \quad b_i = b_i^{(1)} + a_{i1}b_1^{(1)},$$

то рівняння системи (4) можуть бути одержані як лінійна комбінація рівнянь системи (5). Отже, будь-який розв'язок системи (5) є і розв'язком системи (4). Таким чином, системи (4) і (5) рівнозначні. (Лінійною комбінацією двох рівнянь $c_{11}x_1 + c_{12}x_2 + ... + c_{1n}x_n = d_1$ і, $c_{21}x_1 + c_{22}x_2 + ... + c_{2n}x_n = d_2$ будемо називати рівняння:

$\lambda_1(c_{11}x_1 + c_{12}x_2 + ... + c_{1n}x_n) + \lambda_2(c_{21}x_1 + c_{22}x_2 + ... + c_{2n}x_n) = \lambda_1 d_1 + \lambda_2 d_2$, де λ_1 та λ_2 — числа).

Порівняємо тепер визначники D_1 і D_2 основних матриць систем (4) і (5). Перший рядок основної матриці системи (5) одержаний із 1-го рядка основної матриці системи (4) діленням на a_{11}. Така операція відповідає діленню D_1 на a_{11}. Інші рядки одержані відніманням із відповідних рядків основної матриці системи (4) величин, пропорційних 1-му рядку. Ця операція не змінює величини визначника. Звідси виходить, що визначник D_2 основної матриці системи (5) дорівнює $\dfrac{D_1}{a_{11}}$. А тому $D_2 \neq 0$, якщо $D_1 \neq 0$ і $D_2 = 0$, якщо $D_1 = 0$. Зазначимо, нарешті, що обчислення ми проводили тільки з коефіцієнтами рівнянь системи (4), тому немає необхідності писати самі рівняння. Достатньо написати лише розширену матрицю системи і перетворити тільки елементи цієї матриці.

Будемо позначати перехід від однієї розширеної матриці до іншої, тобто фактично перехід від однієї системи рівнянь до системи, їй рівнозначної, символом $\Leftrightarrow$ або $\sim$. Тоді проведені операції можна записати так:

$$B_1 = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & a_{12}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} & b_2^{(1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} & b_n^{(1)} \end{pmatrix} = B_2 .$$

Приймемо спочатку, що визначник D_1 основної матриці системи (4) відмінний від нуля. Тоді, як зазначалося $D_2 \neq 0$, а тому в крайньому випадку одне з чисел $a_{i2}^{(1)}$ (i=1, 2, ..., n) відмінне від нуля, оскільки якби всі $a_{i2}^{(1)}$ дорівнювали нулю, дорівнював би нулю і визначник D_2 основної матриці системи (5).

Оскільки рівняння в системі (5) можна міняти місцями, тому, без обмеження, можливо вважати, що $a_{22}^{(1)} \neq 0$. Поділимо 2-е рівняння системи (5) на $a_{22}^{(1)}$, помножимо одержаний рядок на $a_{i2}^{(1)}$ (i = 1, 3, 4, ..., n) і віднімемо від і-го рядка.

Тоді будемо мати:

$$B_3 = \begin{pmatrix} 1 & 0 & a_{13}^{(2)} & \dots & a_{1n}^{(2)} & b_1^{(2)} \\ 0 & 1 & a_{23}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix}.$$

Система рівнянь, що відповідає матриці B_3, рівнозначна системі (5), а тому і вихідній системі (4). Визначник D_3 основної матриці цієї системи відмінний від нуля, оскільки відмінний від нуля визначник D_2. Звідси в крайньому випадку одне із чисел $a_{i3}^{(2)}$ (i = 3, ..., n) відмінне від нуля і можна знову провести ті самі операції, що і раніше. Продовжуючи аналогічні міркування, після n операцій одержимо матрицю:

$$B_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & b_1^{(n)} \\ 0 & 1 & 0 & \dots & 0 & b_2^{(n)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & b_n^{(n)} \end{pmatrix}.$$

Відповідна система рівнянь має вигляд:

$$\begin{cases} 1\cdot x_1 + 0\cdot x_2 + \ldots + 0\cdot x_n = b_1^{(n)} \\ 0\cdot x_1 + 1\cdot x_2 + \ldots + 0\cdot x_n = b_2^{(n)} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0\cdot x_1 + 0\cdot x_2 + \ldots + 1\cdot x_n = b_n^{(n)} \end{cases}. \tag{6}$$

Її єдиним розв'язком є:

$$x_1 = b_1^{(n)}, \quad x_2 = b_2^{(n)}, \quad \ldots., \quad x_n = b_n^{(n)}. \tag{7}$$

Оскільки система (6) рівнозначна системі (4), має єдиний розв'язок, то має єдиний розв'язок, що визначається формулами (7) і вихідна система(4).

Приклад 1. Розв'язати систему:

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 1 \\ 2x_1 + 3x_2 + 2x_3 - x_4 = -3 \\ x_1 + x_2 + x_3 + 2x_4 = 4 \\ x_1 - 3x_2 + 2x_3 - 2x_4 = 0 \end{cases}.$$

Проведемо перетворення:

$$B = \begin{pmatrix} 1 & 2 & -1 & 1 & 1 \\ 2 & 3 & 2 & -1 & -3 \\ 1 & 1 & 1 & 2 & 4 \\ 1 & -3 & 2 & -2 & 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & -1 & 1 & 1 \\ 0 & -1 & 4 & -3 & -5 \\ 0 & -1 & 2 & 1 & 3 \\ 0 & -5 & 3 & -3 & -1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 7 & -5 & -9 \\ 0 & 1 & -4 & 3 & 5 \\ 0 & 0 & -2 & 4 & 8 \\ 0 & 0 & -17 & 12 & 24 \end{pmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{pmatrix} 1 & 0 & 7 & -5 & -9 \\ 0 & 1 & -4 & 3 & 5 \\ 0 & 0 & 1 & -2 & -4 \\ 0 & 0 & -17 & 12 & 24 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 9 & 19 \\ 0 & 1 & 0 & -5 & -11 \\ 0 & 0 & 1 & -2 & -4 \\ 0 & 0 & 0 & -22 & -44 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

$x_1 = 1$; $x_2 = -1$; $x_3 = 0$; $x_4 = 2$.

Зазначимо, якщо система однорідна, тобто всі числа b_i (i=1, 2, ..., n) дорівнюють нулю, тоді дорівнюють нулю і всі числа $b_i^{(n)}$. Тому система (4) має в цьому випадку тільки нульовий розв'язок.

Нехай тепер визначник D_1 основної матриці системи (4) дорівнює нулю. Тоді вже не можна стверджувати, що серед чисел $a_{im}^{(m-1)}$ ($i = m$, m+1, ..., n), одержаних після (m–1)-го етапу перетворень буде хоча б одне, відмінне від нуля. Більше того, на якомусь етапі всі ці числа обов'язково стануть рівними нулю (інакше ми мали б розглянутий випадок). Таким чином, одержана матриця:

$$B_{m-1} = \begin{pmatrix} 1 & 0 & \dots & a_{1m}^{(m-1)} & a_{1(m+1)}^{(m-1)} & \dots & a_{1n}^{(m-1)} & b_1^{(m-1)} \\ 0 & 1 & \dots & a_{2m}^{(m-1)} & a_{2(m+1)}^{(m-1)} & \dots & a_{2n}^{(m-1)} & b_2^{(m-1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{m(m+1)}^{(m-1)} & \dots & a_{mn}^{(m-1)} & b_m^{(m-1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{n(m+1)}^{(m-1)} & \dots & a_{nn}^{(m-1)} & b_n^{(m-1)} \end{pmatrix}.$$

Переставимо m-ий стовпець матриці на місце n-го, а всі наступні за m-м стовпцем, крім стовпця вільних членів $b_i^{(m-1)}$, зсунемо на одне місце вліво (така операція, очевидно, означає перестановку невідомих у рівняннях системи або їх перенумерацію, що, звичайно, не змінює розв'язку системи). У результаті одержимо матрицю:

$$B_m = \begin{pmatrix} 1 & 0 & \dots & a_{1m}^{(m)} & \dots & a_{1n}^{(m)} & b_1^{(m)} \\ 0 & 1 & \dots & a_{2m}^{(m)} & \dots & a_{2n}^{(m)} & b_2^{(m)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{mm}^{(m)} & \dots & 0 & b_m^{(m)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nm}^{(m)} & \dots & 0 & b_n^{(m)} \end{pmatrix},$$

де

$$a_{in}^{(m)} = a_{im}^{(m-1)}, \quad b_i^{(m)} = b_i^{(m-1)}, \ i = 1, 2, ..., n;$$
$$a_{i(k+1)}^{(m-1)} = a_{ik}^{(m)}, \ k = m, m+1, ..., n.$$

Продовжуючи ті ж самі перетворення, що і раніше, одержимо в кінцевому рахунку матрицю:

$$
B_n = \begin{pmatrix}
1 & 0 & \dots & 0 & a_{1(k+1)}^{(n)} & \dots & a_{1n}^{(n)} & b_1^{(n)} \\
0 & 1 & \dots & 0 & a_{2(k+1)}^{(n)} & \dots & a_{2n}^{(n)} & b_2^{(n)} \\
\dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\
0 & 0 & \dots & 1 & a_{k(k+1)}^{(n)} & \dots & a_{kn}^{(n)} & b_k^{n)} \\
0 & 0 & \dots & 0 & 0 & \dots & 0 & b_{k+1}^{(n)} \\
\dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\
0 & 0 & \dots & 0 & 0 & \dots & 0 & b_n^{(n)}
\end{pmatrix}. \qquad (8)
$$

Матриці (8) відповідає система рівнянь:

$$
\begin{cases}
1\cdot x'_1 + 0\cdot x'_2 + \dots + 0\cdot x'_k + a_{1(k+1)}^{(n)}\cdot x'_{k+1} + \dots + a_{1n}^{(n)}\cdot x'_n = b_1^{(n)} \\
0\cdot x'_1 + 1\cdot x'_2 + \dots + 0\cdot x'_k + a_{2(k+1)}^{(n)}\cdot x'_{k+1} + \dots + a_{2n}^{(n)}\cdot x'_n = b_2^{(n)} \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
0\cdot x'_1 + 0\cdot x'_2 + \dots + 1\cdot x'_k + a_{k(k+1)}^{(n)}\cdot x'_{k+1} + \dots + a_{kn}^{(n)}\cdot x'_n = b_k^{(n)} \\
0\cdot x'_1 + 0\cdot x'_2 + \dots + 0\cdot x'_k + 0\cdot x'_{k+1} + \dots + 0\cdot x'_n = b_{k+1}^{(n)} \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
0\cdot x'_1 + 0\cdot x'_2 + \dots + 0\cdot x'_k + 0\cdot x'_{k+1} + \dots + 0\cdot x'_n = b_n^{(n)}
\end{cases}, \qquad (9)
$$

в якій невідомі x_i' різняться від невідомих x_i у системі (4) лише нумерацією. Оскільки система (9) рівнозначна системі (4), тоді висновок про розв’язуваність системи (4) рівносильний висновку про розв’язуваність системи (9).

Очевидно, що якщо хоча б одне із чисел $b_i^{(n)}$ ($i = k+1, \dots, n$) не дорівнює нулю, тоді рівняння системи (9), а тому і рівняння системи (4), несумісні. Якщо всі $b_i^{(n)}$ ($i = k + 1, \dots, n$) дорівнюють нулю, тоді рівняння сумісні. При цьому невідомим $x'_{k+1}, \dots\ x'_n$ можна надати будь-які значення, і система матиме такі розв’язки:

$$
\begin{array}{l}
x'_1 = b_1^{(n)} - (a_{1(k+1)}^{(n)} t_1 + \dots + a_{1n}^{(n)} t_e), \\
x'_2 = b_2^{(n)} - (a_{2(k+1)}^{(n)} t_1 + \dots + a_{2n}^{(n)} t_e), \\
-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,- \\
x'_k = b_k^{(n)} - (a_{k(k+1)}^{(n)} t_1 + \dots + a_{kn}^{(n)} t_e), \\
x'_{k+1} = t_1, \quad x'_{k+2} = t_2, \ \dots \ x'_n = t_e,
\end{array},
$$

де $t_1, t_2, \dots, t_l$ ($l = n - k$) — довільні.

Для того щоб було зручно повертатися до вихідної системи невідомих, корисно над стовпцями матриць, які одержують при проведенні перетворень, надписувати позначення відповідних невідомих. Зазначимо, крім того, що якщо вихідна система (4) однорідна, тоді всі числа $b_i^{(n)}$ ($i = 1, 2, ..., n$) дорівнюють нулю. Тому мають місце наступні два твердження.

1. Система однорідних рівнянь першого степеня завжди сумісна.

2. Якщо визначник системи однорідних рівнянь першого степеня дорівнює нулю, тоді система має нескінчену множину розв'язків.

Приклад 2.

$$\begin{cases} x_1 + 2x_2 - 2x_3 - 2x_4 = -1 \\ x_1 + 3x_2 - 2x_3 - 7x_4 = 2 \\ 2x_1 - x_2 + 2x_3 + x_4 = 3 \\ x_1 - x_2 + x_3 + 3x_4 = 0 \end{cases}.$$

Проведемо перетворення:

$$B = \begin{pmatrix} 1 & 2 & -2 & -2 & -1 \\ 1 & 3 & -2 & -7 & 2 \\ 2 & -1 & 2 & 1 & 3 \\ 1 & -1 & 1 & 3 & 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & -2 & -2 & -1 \\ 0 & 1 & 0 & -5 & 3 \\ 0 & -5 & 6 & 5 & 5 \\ 0 & -3 & 3 & 5 & 1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & -2 & 8 & -7 \\ 0 & 1 & 0 & -5 & 3 \\ 0 & 0 & 6 & -20 & 20 \\ 0 & 0 & 3 & -10 & 10 \end{pmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{pmatrix} 1 & 0 & -2 & 8 & -7 \\ 0 & 1 & 0 & -5 & 3 \\ 0 & 0 & 1 & -\frac{10}{3} & \frac{10}{3} \\ 0 & 0 & 3 & -10 & 10 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & \frac{4}{3} & -\frac{1}{3} \\ 0 & 1 & 0 & -5 & 3 \\ 0 & 0 & 1 & -\frac{10}{3} & \frac{10}{3} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Система рівнянь, що відповідає одержаній матриці, має вигляд:

$$\begin{cases} 1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + \frac{4}{3} x_4 = -\frac{1}{3} \\ 0 \cdot x_1 + 1 \cdot x_2 + 0 \cdot x_3 - 5x_4 = 3 \\ 0 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 - \frac{10}{3} x_4 = \frac{10}{3} \\ 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 0 \end{cases}.$$

Система сумісна, $x_4 = t$ — довільне число. Система має нескінчену множину розв'язків:

$$\begin{aligned} x_1 &= -\frac{1}{3}(1+4t) \\ x_2 &= 3+5t \\ x_3 &= \frac{10}{3}(1+t) \\ x_4 &= t \end{aligned} \quad .$$

причому t — довільне число.

Зазначимо, що якби вільні члени в рівняннях були іншими, ніж задано в умові, система могла б бути несумісною. Нехай, наприклад, $b_4 = 1$. Тоді перетворена матриця система буде:

$$\begin{pmatrix} 1 & 0 & 0 & \frac{4}{3} & -\frac{1}{3} \\ 0 & 1 & 0 & -5 & 3 \\ 0 & 0 & 1 & -\frac{10}{3} & \frac{10}{3} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

і останнє рівняння системи прийме такий вигляд $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 1$, що не має змісту.

Приклад 3.

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + 2x_3 - x_4 = -4 \\ x_1 + x_2 - 2x_3 - 3x_4 = 2 \\ 2x_1 + 2x_2 - 5x_3 - 6x_4 = 6 \end{cases} .$$

Проведемо перетворення:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 2 & -1 & -4 \\ 1 & 1 & -2 & -3 & 2 \\ 2 & 2 & -5 & -6 & 6 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -2 & -4 \\ 0 & 0 & -3 & -4 & 2 \\ 0 & 0 & -7 & -8 & 6 \end{pmatrix} \Leftrightarrow \begin{array}{c} \begin{matrix} x_1 & x_3 & x_4 & x_2 & \end{matrix} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 0 & -4 \\ 0 & -3 & -4 & 0 & 2 \\ 0 & -7 & -8 & 0 & 6 \end{pmatrix} \end{array} \Leftrightarrow$$

$$\Leftrightarrow \begin{array}{c} \begin{matrix} x_1 & x_3 & x_4 & x_2 & \end{matrix} \\ \begin{pmatrix} 1 & 0 & 3 & 1 & 4 \\ 0 & 1 & -2 & 0 & -4 \\ 0 & 0 & -10 & 0 & -10 \\ 0 & 0 & -22 & 0 & -22 \end{pmatrix} \end{array} \Leftrightarrow \begin{array}{c} \begin{matrix} x_1 & x_3 & x_4 & x_2 & \end{matrix} \\ \begin{pmatrix} 1 & 0 & 3 & 1 & 4 \\ 0 & 1 & -2 & 0 & -4 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & -22 & 0 & -22 \end{pmatrix} \end{array} \Leftrightarrow \begin{array}{c} \begin{matrix} x_1 & x_3 & x_4 & x_2 & \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Система сумісна, $x_2 = t$ — довільне число; $x_1 = 1 - t$, $x_2 = t$, $x_3 = -2$, $x_4 = 1$.

Розглянутий метод без будь-яких змін переноситься і на той випадок, якщо кількість невідомих не співпадає з кількістю рівнянь.

ЗАДАЧІ

3.20. Розв'язати систему:

$$\begin{cases} 3x_1 + 2x_2 = 7 \\ 4x_1 - 5x_2 = 40 \end{cases}.$$

Обчислимо визначник системи

$$\det A = \begin{vmatrix} 3 & 2 \\ 4 & -5 \end{vmatrix} = 3 \cdot (-5) - 4 \cdot 2 = -15 - 8 = -23.$$

Оскільки визначник системи відмінний від нуля, застосуємо правило Крамера. Для обчислення визначника $\det B_1$ замінимо стовпець $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ визначника системи стовпцем із вільних членів $\begin{pmatrix} 7 \\ 40 \end{pmatrix}$. Маємо:

$$\det B_1 = \begin{vmatrix} 7 & 2 \\ 40 & -5 \end{vmatrix} = 7(-5) - 40 \cdot 2 = -35 - 80 = -115.$$

Визначник $\det B_2$ дістанемо заміною стовпця $\begin{pmatrix} 2 \\ -5 \end{pmatrix}$ визначника системи стовпцем із вільних членів:

$$\det B_2 = \begin{vmatrix} 3 & 7 \\ 4 & 40 \end{vmatrix} = 92 .$$

Згідно з правилом Крамера знаходимо $x_1 = \dfrac{-115}{-23} = 5$; $x_2 = \dfrac{92}{-23} = -4$.

Сукупність чисел (5; –4) є єдиним розв'язком даної системи.

3.21. Розв'язати систему:

$$\begin{cases} 2x_1 + 5x_2 - 8x_3 = 8 \\ 4x_1 + 3x_2 - 9x_3 = 9 \\ 2x_1 + 3x_2 - 5x_3 = 7 \end{cases}.$$

Визначник з коефіцієнтів системи відмінний від нуля:

$$\det A = \begin{vmatrix} 2 & 5 & -8 \\ 4 & 3 & -9 \\ 2 & 3 & -5 \end{vmatrix} = 2\cdot3\cdot(-5)+5\cdot(-9)\cdot2+(-8)\cdot4\cdot3-(-8)\cdot3\cdot2-5\cdot4\cdot(-5)-2\cdot3\cdot(-9)=-14\neq0.$$

Тому можна застосувати правило Крамера:

$$\det B_1 = \begin{vmatrix} 8 & 5 & -8 \\ 9 & 3 & -9 \\ 7 & 3 & -5 \end{vmatrix} = -42; \quad \det B_2 = \begin{vmatrix} 2 & 8 & -8 \\ 4 & 9 & -9 \\ 2 & 7 & -5 \end{vmatrix} = -28 \quad \det B_3 = \begin{vmatrix} 2 & 5 & 8 \\ 4 & 3 & 9 \\ 2 & 3 & 7 \end{vmatrix} = -14 .$$

Звідси знаходимо $x_1 = \dfrac{-42}{-14} = 3$; $x_2 = \dfrac{-28}{-14} = 2$; $x_2 = \dfrac{-14}{-14} = 1$.

Сукупність чисел (3, 2, 1) являється єдиним розв'язком системи.

3.22. Розв'язати систему:

$$\begin{cases} 3x_1 + 4x_2 + 5x_3 + 7x_4 = 1 \\ 2x_1 + 6x_2 - 3x_3 + 4x_4 = 2 \\ 4x_1 + 2x_2 + 13x_3 + 10x_4 = 0 \\ 5x_1 \qquad + 21x_3 + 13x_4 = 3 \end{cases}.$$

Випишемо розширену матрицю системи:

$$\begin{pmatrix} 3 & 4 & 5 & 7 & 1 \\ 2 & 6 & -3 & 4 & 2 \\ 4 & 2 & 13 & 10 & 0 \\ 5 & 0 & 21 & 13 & 3 \end{pmatrix} \sim \begin{pmatrix} 3 & 4 & 5 & 7 & 1 \\ 2 & 6 & -3 & 4 & 2 \\ 4 & 2 & 13 & 10 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

/IVр + II – I – III/

Неважко побачити, що визначник з коефіцієнтів системи дорівнює нулю, оскільки 4-ий рядок його складається з нулів. Останній рядок розширеної матриці свідчить про те, що система несумісна.

3.23. Роз’язати систему:

$$\begin{cases} x_1 + 2x_2 + 3x_3 + x_4 = 0 \\ 2x_1 + 3x_2 + x_3 - 2x_4 = 1 \\ x_1 - 2x_2 - 3x_3 + x_4 = 2 \\ 3x_1 + x_2 - 5x_3 - x_4 = 0 \end{cases}.$$

Випишемо розширену матрицю системи:

$$\begin{pmatrix} 1 & 2 & 3 & 1 & .. & 0 \\ 2 & 3 & 1 & -2 & .. & 1 \\ 1 & -2 & -3 & 1 & .. & 2 \\ 3 & 1 & -5 & -1 & .. & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & 1 & .. & 0 \\ 0 & -1 & -5 & -4 & .. & 1 \\ 0 & -4 & -6 & 0 & .. & 2 \\ 0 & 0 & -3 & 0 & .. & -3 \end{pmatrix} \sim$$

/IIр. – 2·I, IIIр. – I, IVр. – II – III/

/поділимо IIIр. на (–2), IVр. на (–3)/

$$\sim \begin{pmatrix} 1 & 2 & 3 & 1 & .. & 0 \\ 0 & -1 & -5 & -4 & .. & 1 \\ 0 & 2 & 3 & 0 & .. & -1 \\ 0 & 0 & 1 & 0 & .. & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 & 3 & .. & 0 \\ 0 & -1 & -4 & -5 & .. & 1 \\ 0 & 0 & -8 & -7 & .. & 1 \\ 0 & 0 & 0 & 1 & .. & 1 \end{pmatrix}.$$

/IIIр. + 2·II/

У результаті всіх перетворень ця система лінійних рівнянь звелася до трикутного вигляду:

$$\begin{cases} x_1 + 2x_2 + x_4 + 3x_3 = 0 \\ -x_2 - 4x_4 - 5x_3 = 1 \\ -8x_4 - 7x_3 = 1 \\ x_3 = 1 \end{cases}.$$

Вона має єдиний розв'язок:
$x_3 = 1$ $x_4 = -1$ $x_2 = -2$ $x_1 = 2$

3.24. Розв'язати систему:

$$\begin{cases} x_1 + 2x_2 + 3x_3 + x_4 = 3 \\ x_1 + 4x_2 + 5x_3 + 2x_4 = 2 \\ 2x_1 + 9x_2 + 8x_3 + 3x_4 = 7 \\ 3x_1 + 7x_2 + 7x_3 + 2x_4 = 12 \\ 5x_1 + 7x_2 + 9x_3 + 2x_4 = 20 \end{cases}.$$

Проведемо перетворення:

$$\begin{pmatrix} 1 & 2 & 3 & 1 & 3 \\ 1 & 4 & 5 & 2 & 2 \\ 2 & 9 & 8 & 3 & 7 \\ 3 & 7 & 7 & 2 & 12 \\ 5 & 7 & 9 & 2 & 20 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & 3 & 1 & 3 \\ 0 & 2 & 2 & 1 & -1 \\ 0 & 5 & 2 & 1 & 1 \\ 0 & 1 & -2 & -1 & 3 \\ 0 & -3 & -6 & -3 & 5 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & 3 & 1 & 3 \\ 0 & 1 & -2 & -1 & 3 \\ 0 & 2 & 2 & 1 & -1 \\ 0 & 5 & 2 & 1 & 1 \\ 0 & -3 & -6 & -3 & 5 \end{pmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{pmatrix} 1 & 0 & 7 & 3 & -3 \\ 0 & 1 & -2 & -1 & 3 \\ 0 & 0 & 6 & 3 & -7 \\ 0 & 0 & 12 & 6 & -14 \\ 0 & 0 & -12 & -6 & 14 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & \frac{31}{6} \\ 0 & 1 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{7}{6} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Рівняння сумісні, $x_4 = t$ — довільне число, $x_1 = \frac{1}{2}t + \frac{31}{6}$, $x_2 = \frac{2}{3}$, $x_3 = -\frac{1}{2}t - \frac{7}{6}$, $x_4 = t$

3.25. Розв'язати систему:

$$\begin{cases} x_1 + 2x_2 + x_3 - x_4 = 1 \\ x_1 + 3x_2 - x_3 + 2x_4 = -4 \\ 2x_1 + x_2 - x_3 + x_4 = -1 \end{cases}$$

Проведемо перетворення:

$$\begin{pmatrix} 1 & 2 & 1 & -1 & 1 \\ 1 & 3 & -1 & 2 & -4 \\ 2 & 1 & -1 & 1 & -1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & 1 & -1 & 1 \\ 0 & 1 & -2 & 3 & -5 \\ 0 & -3 & -3 & 3 & -3 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 5 & -7 & 11 \\ 0 & 1 & -2 & 3 & -5 \\ 0 & 0 & -9 & 12 & -18 \end{pmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{pmatrix} 1 & 0 & 5 & -7 & 11 \\ 0 & 1 & -2 & 3 & -5 \\ 0 & 0 & 1 & -\frac{4}{3} & 2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{3} & 1 \\ 0 & 1 & 0 & \frac{1}{3} & -1 \\ 0 & 0 & 1 & -\frac{4}{3} & 2 \end{pmatrix}.$$

Система сумісна, $x_4 = t$ — довільне число,

$$x_1 = \frac{1}{3}t + 1, \quad x_2 = -\frac{1}{3}t - 1, \quad x_3 = \frac{4}{3}t + 2, \quad x_4 = t .$$

3.26. Роз'язати систему:

$$\begin{cases} x_1 + 2x_2 - x_3 + 3x_4 = 0 \\ 2x_1 - x_2 + 2x_3 - 4x_4 = 0 \\ x_1 + x_2 + 2x_3 + x_4 = 0 \\ x_1 + 5x_2 + x_3 + 9x_4 = 0 \\ x_1 - 5x_2 - 2x_3 - 11x_4 = 0 \end{cases}$$

Проведемо перетворення:

$$\begin{pmatrix} 1 & 2 & -1 & 3 & 0 \\ 2 & -1 & 2 & -4 & 0 \\ 1 & 1 & 2 & 1 & 0 \\ 1 & 5 & 1 & 9 & 0 \\ 1 & -5 & -2 & -11 & 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & -1 & 3 & 0 \\ 0 & -5 & 4 & -10 & 0 \\ 0 & -1 & 3 & -2 & 0 \\ 0 & 3 & 2 & 6 & 0 \\ 0 & -7 & -1 & -14 & 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & -1 & 3 & 0 \\ 0 & 1 & -3 & 2 & 0 \\ 0 & -5 & 4 & -10 & 0 \\ 0 & 3 & 2 & 6 & 0 \\ 0 & -7 & -1 & -14 & 0 \end{pmatrix} \Leftrightarrow$$

$$\Leftrightarrow\begin{pmatrix}1&0&5&-1&0\\0&1&-3&2&0\\0&0&-11&0&0\\0&0&11&0&0\\0&0&-22&0&0\end{pmatrix}\Leftrightarrow\begin{pmatrix}1&0&5&-1&0\\0&1&-3&2&0\\0&0&1&0&0\\0&0&11&0&0\\0&0&-22&0&0\end{pmatrix}\Leftrightarrow\begin{pmatrix}1&0&0&-1&0\\0&1&0&2&0\\0&0&1&0&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

Система сумісна, $x_4 = t$ — довільне число, $x_1 = t$, $x_2 = -2t$, $x_3 = 0$, $x_4 = t$.

Розв'язати системи за правилом Крамера.

3.27.

$$\begin{cases}3x_1 - 5x_2 = 13\\2x_1 + 7x_2 = 81\end{cases}.$$

3.28.

$$\begin{cases}3x_1 - 4x_2 = 1\\3x_1 + 4x_2 = 17\end{cases}.$$

3.29.

$$\begin{cases}7x_1 + 2x_2 + 3x_3 = 15\\5x_1 - 3x_2 + 2x_3 = 15\\10x_1 - 11x_2 + 5x_3 = 36\end{cases}.$$

3.30.

$$\begin{cases}x_1 + x_2 - 2x_3 = 6\\2x_1 + 3x_2 - 7x_3 = 16\\5x_1 + 2x_2 + x_3 = 16\end{cases}.$$

3.31

$$\begin{cases}4x_1 + 4x_2 + 5x_3 + 5x_4 = 0\\2x_1 + \phantom{x_2 +{}} 3x_3 - x_4 = 10\\x_1 + x_2 - 5x_3 = -10\\3x_2 + 2x_3 = 1\end{cases}.$$

3.32.

$$\begin{cases}2x_1 - x_2 + 3x_3 + 2x_4 = 4\\3x_1 + 3x_2 + 3x_3 + 2x_4 = 6\\3x_1 - x_2 - x_3 - 2x_4 = 6\\3x_1 - x_2 + 3x_3 - x_4 = 6\end{cases}.$$

Розв'язати системи методом Гаусса (метод послідовного виключення невідомих).

3.33.

$$\begin{cases}x_1 + 2x_2 + 3x_3 + x_4 = 0\\2x_1 + 3x_2 + x_3 - 2x_4 = 1\\x_1 - 2x_2 - 3x_3 + x_4 = 2\\3x_1 + x_2 - 5x_3 - x_4 = 0\end{cases}.$$

3.34.

$$\begin{cases}12x_1 + 14x_2 - 15x_3 + 23x_4 + 27x_5 = 5\\16x_1 + 18x_2 - 22x_3 + 29x_4 + 37x_5 = 8\\18x_1 + 20x_2 - 21x_3 + 32x_4 + 41x_5 = 9\\10x_1 + 12x_2 - 16x_3 + 20x_4 + 23x_5 = 4\end{cases}.$$

3.35.

$$\begin{cases} 3x_1+4x_2+5x_3+7x_4=1\\ 2x_1+6x_2-3x_3+4x_4=2\\ 4x_1+2x_2+13x_3+10x_4=0\\ 5x_1+21x_3+13x_4=3 \end{cases}.$$

3.36.

$$\begin{cases} x_1+3x_3-3x_4=-6,\\ 2x_1+x_2+x_3-5x_4=8,\\ 3x_1-5x_2-2x_3-5x_4=17,\\ 12x_2+7x_3-7x_4=-9 \end{cases}.$$

3.37.

$$\begin{cases} 2x_1+3x_2-x_3+x_4=1\\ 8x_1+12x_2-9x_3+8x_4=3\\ 4x_1+6x_2+3x_3-2x_4=3\\ 2x_1+3x_2+9x_3-7x_4=3 \end{cases}.$$

3.38.

$$\begin{cases} x_1-5x_2-8x_3+x_4=3\\ 3x_1+x_2-3x_3-5x_4=1\\ x_1-7x_3+2x_4=-5\\ 11x_2+20x_3-9x_4=2 \end{cases}.$$

§6. Ранг матриці, теорема про сумісність систем рівнянь першого степеня

Для дослідження багатьох питань, щодо розв'язання систем рівнянь першого степеня, часто вводять поняття *ранг матриці*.

Означення. *Рангом матриці називається найвищий порядок відмінного від нуля визначника квадратної субматриці, отриманої із заданої матриці викреслюванням деяких рядків і стовпців.*

Розглянемо, наприклад, матрицю:

$$\begin{pmatrix} 1 & 3 & 0 & 4\\ 3 & 2 & 0 & 1\\ 1 & 3 & 0 & 4 \end{pmatrix}.$$

Викресленням будь-якого числа рядків і стовпців неможливо із заданої матриці одержати квадратну матрицю порядку вище третього. Звідси ранг її не може бути більше трьох. Але, викреслюючи один із стовпців, ми будемо одержувати квадратні матриці, які мають два однакових рядки, а тому їх визначники дорівнюють нулю. Отже, ранг вихідної матриці менше 3-х. Викресливши, наприклад, 3-й і 4-й стовпці і 3-й рядок, отримаємо квадратну матрицю $\begin{pmatrix} 1 & 3\\ 3 & 2 \end{pmatrix}$, визначник якої не дорівнює нулю. Таким чином, всі визначники субматриці третього порядку дорівнюють нулю, але серед визначників матриць другого порядку є відмінний від нуля. Тим самим ранг вихідної матриці дорівнює двом.

Доведемо теорему: *ранг матриці не змінюється при лінійних операціях з її рядками.*

Дійсно, лінійні операції з рядками якої-небудь матриці приводять до тих самих лінійних операцій з рядками будь-якої субматриці. Але як зазначалося при лінійних операціях з рядками квадратних матриць визначники цих матриць одержуються один з одного множенням на число, відмінне від нуля. Звідси нульовий визначник залишається нульовим, а відмінний від нуля — відмінним від нуля, тобто не може змінитися найвищий порядок відмінного від нуля визначника субматриць. Не впливає, очевидно, на ранг матриці і перестановка стовпців, оскільки така перестановка може впливати лише на знак відповідних визначників.

Із доведеної теореми виходить, що розглянуті в минулому параграфі перетворені матриці мають той самий ранг, що і вихідні. Тому ранг основної матриці системи рівнянь першого степеня дорівнює числу одиниць на головній діагоналі перетвореної матриці.

Доведемо тепер теорему про сумісність систем рівнянь першого степеня (**теорема Кронекера-Капеллі**): *для того щоб система рівнянь першого степеня була сумісна, необхідно і достатньо, щоб ранг розширеної матриці співпадав з рангом основної матриці.*

Нехай ранг основної матриці системи дорівнює k. Якщо ранг розширеної матриці системи також дорівнює k, тоді це означає, що або система містить тільки k рівнянь, або — всі числа $b_i^{(n)}$ ($i = k+1, \ldots, k$) у перетвореній матриці дорівнюють нулю (у протилежному випадку ранг розширеної матриці перетвореної, а тому і вихідної системи був би $k+1$).

Нехай ранг перетвореної (а тому і вихідної) розширеної матриці системи більше k, тобто більше, ніж число одиниць на головній діагоналі перетвореної матриці. Тоді існує хоча б одна субматриця (k+1)-го порядку, визначник якої не дорівнює нулю. Така субматриця може бути одержана тільки додаванням до одиничної матриці порядку k (що знаходиться в лівому верхньому куту перетвореної матриці) якого-небудь одного рядка і стовпця, який складається із перших k вільних членів рівнянь перетвореної системи і будь-якого одного вільного члена із наступних $n–k$ рівнянь. Щоб визначник зазначеної субматриці був відмінний від нуля, відмінним від нуля має бути і цей останній доданий елемент, тобто число $b_i^{(n)}$ ($i = k+1, ..., k$). Але, як було доведено раніше, в цьому випадку (при $b_i^{(n)} \neq 0$) система несумісна. Отже, система сумісна тоді і тільки тоді, коли ранг основної матриці співпадає з рангом розширеної матриці.

ЗАДАЧІ

3.39. Обчислити ранг матриці за допомогою елементарних перетворень

$$A = \begin{pmatrix} 3 & 2 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ 0 & 4 & 5 & 1 \end{pmatrix}.$$

$$A \Rightarrow \begin{pmatrix} 3 & 2 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{pmatrix} \Rightarrow \begin{pmatrix} 3 & 1 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{pmatrix},$$

(III р –2·I)　　　(II ст : 2)　　　(I ст – 3·II, III ст – II, IV ст – 2·II)

де знак $\Rightarrow$ вказує, що з'єднані ним матриці одержують одну з одної елементарними перетвореннями, а тому мають один і той самий ранг.

Далі III р+3·II; I ст:2, додаючи його до III ст і віднімаючи із IV ст і переставивши, нарешті, місцями перші два стовпця, будемо мати:

$$A \Rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ранг матриці A дорівнює 2, тобто $r = 2$.

3.40. За допомогою елементарних перетворень обчислити ранг матриці:

$$\begin{pmatrix} 25 & 31 & 17 & 43 \\ 75 & 94 & 53 & 132 \\ 75 & 94 & 54 & 134 \\ 25 & 32 & 20 & 48 \end{pmatrix} \Rightarrow \begin{pmatrix} 25 & 31 & 17 & 43 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & 3 & 5 \end{pmatrix} \Rightarrow \begin{pmatrix} 25 & 31 & 17 & 43 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix},$$

$(\text{IIр-}3\cdot\text{I},\text{IIIр-II},\text{IVр-I})$

IVр відкинемо, оскільки $\text{IV} = \text{II} + \text{III}$ |

де $r = 3$, бо визначник трикутної матриці з перших трьох стовпців не дорівнює нулю.

Обчислення ранга матриці методом обрамлення. Вибираємо в даній матриці мінор другого порядку, відмінний від нуля. Потім обчислюємо мінори третього порядку, які обрамляють (містять у собі) вибраний, поки не знайдемо серед них відмінного від нуля. Далі обчислюємо мінори четвертого порядку, які обрамляють відмінний від нуля мінор третього порядку, поки не знайдемо серед них відмінний від нуля, і т.д. Якщо знайшли відмінний від нуля мінор r-го порядку, а всі мінори $(r + 1)$-го порядку, що його обрамляють, дорівнюють нулю або їх вже немає, то ранг матриці дорівнює r.

3.41. Обчислити ранг матриці:

$$\begin{pmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \\ 7 & 7 & 9 & 1 \end{pmatrix}.$$

Викреслимо IIIр., оскільки 2·IIр.+I є IIIр.
Отримуємо:

$$\begin{pmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 7 & 7 & 9 & 1 \end{pmatrix}.$$

Виберемо, наприклад, $\begin{vmatrix} 1 & 3 \\ 2 & -1 \end{vmatrix} \neq 0$.

Обчислимо мінори третього порядку, які обрамляють його

$$\begin{vmatrix} 1 & 3 & 5 \\ 2 & -1 & -3 \\ 7 & 7 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 2 & -7 & -13 \\ 7 & -14 & -26 \end{vmatrix} = \begin{vmatrix} 7 & 13 \\ 14 & 26 \end{vmatrix} = 0$$

$$\begin{vmatrix} 1 & 3 & -1 \\ 2 & -1 & 4 \\ 7 & 7 & 1 \end{vmatrix} \neq 0 .$$

IIст. – 3·I, IIIст. – 5·I, мінор III-го порядку відмінний від нуля.

Він міститься у визначнику четвертого порядку заданої матриці, який дорівнює нулю. Отже, $r = 3$.

3.42. Розв’язати системи рівнянь:

$$a)\begin{cases} x_1 + 2x_2 + 3x_3 = 2 \\ x_1 - x_2 + x_3 = 0 \\ x_1 + 3x_2 - x_3 = -2 \\ 3x_1 + 4x_2 + 3x_3 = 0 \end{cases} \qquad b)\begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 0 \\ x_1 - x_2 + x_3 + 2x_4 = 4 \\ x_1 + 5x_2 + 5x_3 - 4x_4 = -4 \\ x_1 + 8x_2 + 7x_3 - 7x_4 = -8 \end{cases}$$

$$\text{c)}\begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 0 \\ x_1 - x_2 + x_3 + 2x_4 = 4 \\ x_1 + 5x_2 + 5x_3 - 4x_4 = -4 \\ x_1 + 8x_2 + 7x_3 - 7x_4 = 6 \end{cases}.$$

а) Тут $r(A)=3$, $r(B)=3$; система сумісна, визначена.

Оскільки $\begin{vmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \\ 1 & 3 & -1 \end{vmatrix} = 14 \neq 0$,

тоді із перших трьох рівнянь системи, наприклад, згідно з формулами Крамера знаходимо

$x_1 = -1$, $x_2 = 0$, $x_3 = 1$;

b) Тут $r(A)=2$, $r(B)=2$; система сумісна, але не визначена.

Визначник $\begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix} = -3 \neq 0$

і з перших двох рівнянь системи

$$\begin{cases} x_1 + 2x_2 = -3x_3 + x_4 \\ x_1 - x_2 = 4 - x_3 - 2x_4 \end{cases}.$$

Знаходимо:

$$x_1 = \frac{8}{3} - \frac{5}{3}x_3 - x_4, \quad x_2 = -\frac{4}{3} - \frac{2}{3}x_3 + x_4,$$

де невідомим x_3 і x_4 можна надавати будь-які значення.

с) у цьому випадку $r(A) = 2$, $r(B) = 3$; і система несумісна.

3.43. Методом Гаусса (послідовного виключення невідомих) розв'язати однорідну систему рівнянь:

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 0 \\ 2x_1 + 3x_2 + 4x_3 + 5x_4 + x_5 = 0 \\ 3x_1 + 4x_2 + 5x_3 + x_4 + 2x_5 = 0 \\ x_1 + 3x_2 + 5x_3 + 12x_4 + 9x_5 = 0 \\ 4x_1 + 5x_2 + 6x_3 - 3x_4 + 3x_5 = 0 \end{cases}$$

і знайти її фундаментальну систему розв'язків.

Випишемо розширену матрицю системи (при цьому нульовий стовпець можна, звичайно, не писати). Після зрозумілих перетворень будемо мати:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 1 & 3 & 5 & 12 & 9 \\ 4 & 5 & 6 & -3 & 3 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -2 & -3 & -9 \\ 0 & -2 & -4 & -11 & -13 \\ 0 & 1 & 2 & 8 & 4 \\ 0 & -3 & -6 & -19 & -17 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -2 & -3 & -9 \\ 0 & 0 & 0 & -5 & 5 \\ 0 & 0 & 0 & 5 & -5 \\ 0 & 0 & 0 & -10 & 10 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 9 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

тобто задана система рівнозначна наступній:

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 0 \\ \quad x_2 + 2x_3 + 3x_4 + 9x_5 = 0 \\ \quad\quad -x_4 + x_5 = 0 \end{cases}.$$

Тут $r = 3$, і три невідомих можливо виразити через останні, наприклад, так:

$x_4 = x_5$

$x_2 = -2x_3 - 3x_4 - 9x_5 = -2x_3 - 12x_5$.

$x_1 = -2x_2 - 3x_3 - 4x_4 - 5x_5 = x_3 + 15x_5$

Фундаментальну систему можливо одержати, якщо вільним невідомим x_3, x_5 надавати значення x_3=1, x_5=0 (тоді x_1=1, x_2=-2, x_4=0) і значення $x_3 = 0$, $x_5 = 1$ (тоді $x_1 = 15$, $x_2 = -12$, $x_4 = 1$). Це дає фундаментальну систему розв'язків:

$e_1 = (1, -2, 1, 0, 0)$, $e_2 = (15, -12, 0, 1, 1)$.

З використанням фундаментальної системи часто записують загальний розв'язок у вигляді лінійної комбінації роз'язків e_1 та e_2, тобто:

$$e = \alpha_1 e_1 + \alpha_2 e_2 = (\alpha_1 + 15\alpha_2,\ -2\alpha_1 - 12\alpha_2,\ \alpha_1,\ \alpha_2,\ \alpha_2).$$

3.44. Знайти фундаментальну систему розв'язків системи лінійних рівнянь та записати її загальний розв'язок:

$$\begin{cases} 2x_1 + 3x_2 - 7x_3 + 3x_4 = 0 \\ 2x_1 - 3x_2 - 3x_3 + 5x_4 = 0 \\ 2x_1 + 9x_2 - 11x_3 + x_4 = 0 \\ 4x_1 + 6x_2 - 14x_3 + 6x_4 = 0 \end{cases}$$

$$\begin{pmatrix} 2 & 3 & -7 & 3 \\ 2 & -3 & -3 & 5 \\ 2 & 9 & -11 & 1 \\ 4 & 6 & -14 & 6 \end{pmatrix} \Rightarrow$$

IVр. відкинемо (пропорціон. до I), або IVр.=II+III, IIр. – I, IIIр. – I

$$\begin{pmatrix} 2 & 3 & -7 & 3 \\ 0 & -6 & 4 & 2 \\ 0 & 6 & -4 & -2 \end{pmatrix}$$

Третій рядок відкинемо. Система звелась до ступінчастої з головними невідомими x_1, x_2 і вільними x_3, x_4:

$$\begin{cases} 2x_1 + 3x_2 = 7x_3 - 3x_4 \\ \quad -6x_2 = -4x_3 - 2x_4 \end{cases}.$$

З останнього рівняння $x_2 = \frac{2}{3}x_3 + \frac{1}{3}x_4$. З першого $x_1 = \frac{1}{2}(5x_3 - 4x_4) =$ $= \frac{5}{2}x_3 - 2x_4$. Вільних невідомих 2. Тому беремо визначник другого порядку з одиничними елементами головної діагоналі і нульовими — побічної: $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$.

Покладемо $x_3 = 1$, $x_4 = 0$. Тоді $x_1 = \frac{5}{2}$, $x_2 = \frac{2}{3}$.

Одержимо вектор $e_1 = (\frac{5}{2}, \frac{2}{3}, 1, 0)$.

Далі покладемо $x_3 = 0$, $x_4 = 1$. Одержимо вектор $e_2 = (-2, \frac{1}{3}, 0, 1)$.

Вектори e_1 і e_2 являють собою фундаментальну систему розв'язків. Тепер загальний розв'язок можна записати у вигляді:

$$e = \alpha_1 e_1 + \alpha_2 e_2 = (\frac{5}{2}\alpha_1 - 2\alpha_2,\ \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2,\ \alpha_1,\ \alpha_2).$$

Надаючи коефіцієнтам α_1, α_2 будь-які (довільні) числові значення будемо одержувати різноманітні частинні розв'язки.

3.45. Знайти фундаментальну систему розв'язків системи лінійних рівнянь та записати її загальний розв'язок

$$\begin{cases} x_1 + x_2 - 3x_3 + x_4 + x_5 = 0 \\ x_1 - x_2 + x_3 + 3x_4 - 3x_5 = 0 \\ x_1 - 2x_2 + 3x_3 + 4x_4 - 5x_5 = 0 \\ x_1 \qquad\quad - x_3 + 2x_4 - x_5 = 0 \\ x_1 + 3x_2 - 7x_3 - x_4 + 5x_5 = 0 \end{cases}.$$

$$\begin{pmatrix} 1 & 1 & -3 & 1 & 1 \\ 1 & -1 & 1 & 3 & -3 \\ 1 & -2 & 3 & 4 & -5 \\ 1 & 0 & -1 & 2 & -1 \\ 1 & 3 & -7 & -1 & 5 \end{pmatrix} \underset{\Leftrightarrow}{\text{(від усіх рядків віднімемо четвертий)}} \begin{pmatrix} 0 & 1 & -2 & -1 & 2 \\ 0 & -1 & 2 & 1 & -2 \\ 0 & -2 & 4 & 2 & -4 \\ 1 & 0 & -1 & 2 & -1 \\ 0 & 3 & -6 & -3 & 6 \end{pmatrix}$$

Другий, третій та п’ятий рядки, які пропорціональні першому, викреслимо. В одержаній матриці переставимо перший і другий стовпці:

$$\begin{pmatrix} 1 & 0 & -2 & -1 & 2 \\ 0 & 1 & -1 & 2 & -1 \end{pmatrix}$$ Ранг матриці дорівнює 2.

Головні невідомі x_2 і x_1. Вільні — x_3, x_4, x_5. Система тепер має вигляд:

$$\begin{cases} x_2 = 2x_3 + x_4 - 2x_5 \\ x_1 = x_3 - 2x_4 + x_5 \end{cases}.$$

Надаючи вільним невідомим послідовно значення, рівні елементам стовпців визначника

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

($x_3 = 1$, $x_4 = 0$, $x_5 = 0$; 2) $x_3 = 0$, $x_4 = 1$, $x_5 = 0$; 3) $x_3 = 0$, $x_4 = 0$, $x_5 = 1$) одержимо:

1) $x_2 = 2$, $x_1 = 1$; 2) $x_2 = 1$, $x_1 = -2$; 3) $x_2 = -2$, $x_1 = 1$,
тобто вектори $e_1 = (1, 2, 1, 0, 0)$
$e_2 = (-2, 1, 0, 1, 0)$
$e_3 = (1, -2, 0, 0, 1)$
складають фундаментальну систему розв’язків. Загальний розв’язок системи тепер запишеться:

$$\begin{aligned} e &= \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = \\ &= (\alpha_1 - 2\alpha_2 + \alpha_3,\ 2\alpha_1 + \alpha_2 - 2\alpha_3,\ \alpha_1,\ \alpha_2,\ \alpha_3). \end{aligned}$$

3.46. Знайти фундаментальну систему розв’язків системи лінійних рівнянь та записати її загальний розв’язок

$$\begin{cases} 3x_1 + x_2 - 8x_3 + 2x_4 + x_5 = 0 \\ 2x_1 - 2x_2 - 3x_3 - 7x_4 + 2x_5 = 0 \\ x_1 + 11x_2 - 12x_3 + 34x_4 - 5x_5 = 0 \\ x_1 - 5x_2 + 2x_3 - 16x_4 + 3x_5 = 0 \end{cases}.$$

Матриця коефіцієнтів

$$A=\left(\begin{array}{cc|ccc}3 & 1 & -8 & 2 & 1\\ 2 & -2 & -3 & -7 & 2\\ \hline 1 & 11 & 12 & 34 & -5\\ 1 & -5 & 2 & -16 & 3\end{array}\right)$$ має ранг $r = 2$ (перевірте).

Виберемо за базисний мінор $M_2=\begin{vmatrix}3 & 1\\ 2 & -2\end{vmatrix}\neq 0$.

Тоді скорочена система має вигляд:

$$\begin{cases}3x_1+x_2=8x_3-2x_4-x_5\\ 2x_1-2x_2=3x_3+7x_4-2x_5\end{cases}.$$

Звідки, покладаючи x_3=c$_1$, x_4=c$_2$, x_5=c$_3$, знаходимо:

$$\begin{aligned}x_1&=\frac{19}{8}c_1+\frac{3}{8}c_2-\frac{1}{2}c_3\\ x_2&=\frac{7}{8}c_1-\frac{25}{8}c_2+\frac{1}{2}c_3\end{aligned}.$$

Загальний розв'язок системи

$$e=\begin{pmatrix}\frac{19}{8}c_1 & +\frac{3}{8}c_2 & -\frac{1}{2}c_3\\ \frac{7}{8}c_1 & -\frac{25}{8}c_2 & +\frac{1}{2}c_3\\ & c_1 & \\ & c_2 & \\ & c_3 & \end{pmatrix}.$$

Із загального розв'язку знаходимо фундаментальну систему розв'язків:

$$e_1(1,0,0)=\begin{pmatrix}\frac{19}{8}\\ \frac{7}{8}\\ 1\\ 0\\ 0\end{pmatrix},\quad e_2(0,1,0)=\begin{pmatrix}\frac{3}{8}\\ -\frac{25}{8}\\ 0\\ 1\\ 0\end{pmatrix},\quad e_3(0,0,1)=\begin{pmatrix}-\frac{1}{2}\\ \frac{1}{2}\\ 0\\ 0\\ 1\end{pmatrix}.$$

Із використанням фундаментальної системи загальний розв'язок може бути записаний:

$e = c_1e_1 + c_2e_2 + c_3e_3$.

Обчислити ранг матриць з допомогою елементарних перетворень:

3.47.

$$\begin{pmatrix} 3 & 5 & 7 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix}.$$

3.48.

$$\begin{pmatrix} 4 & 3 & 2 & 2 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & 3 \end{pmatrix}.$$

3.49.

$$\begin{pmatrix} -1 & 3 & 3 & -4 \\ 4 & -7 & -2 & 1 \\ -3 & 5 & 1 & 0 \\ -2 & 3 & 0 & 1 \end{pmatrix}.$$

3.50.

$$\begin{pmatrix} 24 & 19 & 36 & 72 & -38 \\ 49 & 40 & 73 & 147 & -80 \\ 73 & 59 & 98 & 219 & -118 \\ 47 & 36 & 71 & 141 & -72 \end{pmatrix}.$$

Обчислити ранг матриць методом обрамлення:

3.51.

$$\begin{pmatrix} 2 & -1 & 3 & -2 & 4 \\ 4 & -2 & 5 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{pmatrix}.$$

3.52.

$$\begin{pmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \\ 7 & 7 & 9 & 1 \end{pmatrix}.$$

3.53.

$$\begin{pmatrix} 3 & -1 & 3 & 2 & 5 \\ 5 & -3 & 2 & 3 & 4 \\ 1 & -3 & -5 & 0 & -7 \\ 7 & -5 & 1 & 4 & 1 \end{pmatrix}.$$

3.54.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}.$$

Розв'язати системи:

3.55.

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 0 \\ 2x_1 + 5x_2 + 3x_3 = 0 \\ 3x_1 + 4x_2 + 2x_3 = 0 \end{cases}.$$

3.56.

$$\begin{cases} x_1 + 2x_2 + 4x_3 - 3x_4 = 0 \\ 3x_1 + 5x_2 + 6x_3 - 4x_4 = 0 \\ 4x_1 + 5x_2 - 2x_3 + 3x_4 = 0 \\ 3x_1 + 8x_2 + 24x_3 - 19x_4 = 0 \end{cases}.$$

3.57.

$$\begin{cases} 2x_1 - 4x_2 + 5x_3 + 3x_4 = 0 \\ 3x_1 - 6x_2 + 4x_3 + 2x_4 = 0 \\ 4x_1 - 8x_2 + 17x_3 + 11x_4 = 0 \end{cases}.$$

3.58.

$$\begin{cases} 3x_1 + 2x_2 + x_3 + 3x_4 + 5x_5 = 0 \\ 6x_1 + 4x_2 + 3x_3 + 5x_4 + 7x_5 = 0 \\ 9x_1 + 6x_2 + 5x_3 + 7x_4 + 9x_5 = 0 \\ 3x_1 + 2x_2 + \qquad 4x_4 + 8x_5 = 0 \end{cases}.$$

3.59.

$$\begin{cases} 5x_1 + 6x_2 - 2x_3 + 7x_4 + 4x_5 = 0 \\ 2x_1 + 3x_2 - x_3 + 4x_4 + 2x_5 = 0 \\ 7x_1 + 9x_2 - 3x_3 + 5x_4 + 6x_5 = 0 \\ 5x_1 + 9x_2 - 3x_3 + x_4 + 6x_5 = 0 \end{cases}.$$

3.60.

$$\begin{cases} 3x_1 + 4x_2 + x_3 + 2x_4 + 3x_5 = 0 \\ 5x_1 + 7x_2 + x_3 + 3x_4 + 4x_5 = 0 \\ 4x_1 + 5x_2 + 2x_3 + x_4 + 5x_5 = 0 \\ 7x_1 + 10x_2 + x_3 + 6x_4 + 5x_5 = 0 \end{cases}.$$

§7. Основні операції з матрицями

У попередньому параграфі широко застосовувались лінійні операції з рядками і стовпцями різних матриць. Але в деяких питаннях лінійної алгебри доводиться розглядати операції з матрицями як з єдиним об'єктом.

В основі вивчення операцій з матрицями лежить поняття рівності матриць. Будемо виходити з наступного **означення**: *дві матриці однієї і тієї ж розмірності називаються рівним, якщо рівні всі їх відповідні елементи.*

Отже, матриці A і B однієї і тієї самої розмірності $n \times m$ рівні тоді і тільки тоді, коли $A_{ik} = B_{ik}$ $i = 1, 2,..., n$; $k = 1, 2,..., m$. При цьому ще раз підкреслимо, що порівнювати можна лише матриці однієї і тієї ж розмірності.

Сумою двох матриць A і B однієї і тієї самої розмірності $n\times m$ називається матриця C тієї самої розмірності така, що

$$(C)_{ik} = (A)_{ik} + (B)_{ik}\,. \qquad (1)$$

Отже, при додаванні матриць (додавати можливо тільки матриці однієї і тієї самої розмірності) треба складати всі їх відповідні елементи.

Оскільки додавання матриць зводиться до додавання чисел — елементів цих матриць, очевидно має місце комутативна і асоціативна властивість.

$$A + B = B + A;\;\; (A + B) + C = A + (B + C) \qquad (2)$$

Добутком матриці A на число λ (або числа λ на матрицю A) називається матриця B така, що

$$(B)_{ik} = \lambda(A)_{ik}\,, \qquad (3)$$

тобто при множенні матриці на число (або числа на матрицю) необхідно всі елементи матриці помножити на це число. Нагадаємо, що при множенні на число визначника матриці достатньо було помножити на це число лише елементи будь-якого рядка (або стовпця).

Легко перевірити, що при множенні матриці на число має місце розподільча властивість:

$$\lambda(A + B) = \lambda A + \lambda B;\;\; (\lambda + \eta)A = \lambda A + \eta A \qquad (4)$$

Визначимо тепер добуток двох матриць. Нехай дано матрицю A розмірності $n\times m$ і матриця B розмірності $m\times p$.

Означення. Добутком матриці A розмірності $n\times m$ на матрицю B розмірності $m\times p$ називається матриця C розмірності $n\times p$ така, що

$$(C)_{ik} = \sum_{i=1}^{m} (A)_{il}(B)_{lk} \qquad (5)$$

інакше кажучи, для одержання елемента, що знаходиться в i-ому рядку і в k-ому стовпці матриці добутку, потрібно обчислити суму добутків елементів i-го рядка першого множника і відповідних елементів k-го стовпця другого множника. Отже, для того щоб можливо було скласти вказану суму, потрібна рівність кількості стовпців у першій матриці (тобто кількість елементів у кожному рядку) кількості рядків у другій (тобто кількості елементів у кожному стовпці).

Приклад 1.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ -1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$$

Знайти $A \cdot B$

Матриця A має розмірність 3×2, матриця B — 2×2; добуток існує — це матриця розмірності 3×2.

$$\begin{pmatrix} 1 & 2 \\ 2 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1\cdot 2 + 2(-1) & 1\cdot 1 + 2\cdot 2 \\ 2\cdot 2 + 0(-1) & 2\cdot 1 + 0\cdot 2 \\ -1\cdot 2 + 1(-1) & (-1)\cdot 1 + 1\cdot 2 \end{pmatrix} = \begin{pmatrix} 0 & 5 \\ 4 & 2 \\ -3 & 1 \end{pmatrix}.$$

Добуток матриць не має переставної властивості: $A \cdot B$, взагалі кажучи, не дорівнює $B \cdot A$.

По-перше, з того, що можна обчислити $A \cdot B$, зовсім не виходить, що має зміст $B \cdot A$. Наприклад, у прикладі, що розглядали перестановка множників, тобто множення B на A неможливе оскільки не можна матрицю розмірності 2×2 помножити на матрицю розмірністю 3×2 — кількість стовпців першої матриці тут не дорівнює кількості рядків другої. Але навіть якщо добуток $B \cdot A$ існує, то нерідко $AB \neq BA$. Розглянемо приклад.

Нехай $A = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}$ *і* $B = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$.

Тоді

$$AB = \begin{pmatrix} 1\cdot 2 + 3(-1) & 1\cdot 1 + 3\cdot 0 \\ -1\cdot 2 + 2(-1) & -1\cdot 1 + 2\cdot 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -4 & -1 \end{pmatrix}$$

$$BA = \begin{pmatrix} 2\cdot 1 + 1(-1) & 2\cdot 3 + 1\cdot 2 \\ -1\cdot 1 + 0(-1) & -1\cdot 3 + 0\cdot 2 \end{pmatrix} = \begin{pmatrix} 1 & 8 \\ -1 & -3 \end{pmatrix}.$$

Отже, $AB \neq BA$

Разом з тим можливо довести (таке доведення ми рекомендуємо провести читачу), що

$$(A \cdot B) \cdot C = A(B \cdot C) \qquad (6)$$
$$A \cdot (B + C) = AB + AC.$$

(звичайно береться до уваги, що всі ці добутки мають зміст).

Відповідно до визначення добутку матриць завжди можливе множення квадратних матриць одного порядку, при цьому добуток буде матрицею того самого порядку. Зазначимо без доведення одну із властивостей добутку квадратних матриць одного порядку: ***визначник добутку двох матриць одного і того самого порядку дорівнює добутку визначників матриць, що перемножуються.***

Дуже часто доводиться розглядати добуток матриці розмірності nхm на матрицю розмірності $m\times1$, тобто на матрицю з одним стовпцем. Очевидно, ми повинні одержати в результаті матрицю розмірності $n\times1$, тобто також матрицю з одним стовпцем. Нехай, наприклад, необхідно помножити матрицю

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \text{ на матрицю } X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

У результаті одержимо матрицю $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$, елементи якої обчислюються за формулами:

$$b_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = \sum_{k=1}^{m} a_{1k}x_k$$
$$b_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m = \sum_{k=1}^{m} a_{2k}x_k \,.$$
....
$$b_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m = \sum_{k=1}^{m} a_{nk}x_k$$

Але це означає, що систему рівнянь першого порядку, що розглядалася в попередньому параграфі, можливо записати в дуже зручній матричній формі: $AX = B$.

Суттєву роль у різних застосуваннях матричної алгебри відіграє квадратна матриця, в якої всі діагональні елементи (тобто елементи, які знаходяться на головній діагоналі) дорівнюють 1, а всі інші елементи дорівнюють нулю. Така матриця називається *одиничною матрицею*. Очевидно, що визначник одиничної матриці

$$\det E = \det\begin{pmatrix} 1 & 0 & 0 & .. & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 0 \end{pmatrix} = \begin{vmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 \end{vmatrix} = 1.$$

Характерні такі властивості одиничної матриці: нехай задана квадратна матриця A порядку n і E — одинична матриця того самого порядку. Тоді $AE = EA = A$.

Дійсно $(AE)_{ik} = \sum_{l=1}^{n} (A)_{il} (E)_{lk}$, але $(E)_{lk} = \begin{cases} 0 & l \neq k \\ 1 & l = k \end{cases}$.

Тому в сумі $\sum_{l=1}^{n} (A)_{il} (E)_{lk}$ відмінні від нуля тільки ті складові, для яких $l = k$. Отже, $(AE)_{ik} = (A)_{ik}$ і звідси $AE = A$. Аналогічно одержуємо і для добутку EA.

ЗАДАЧІ

3.61. Знайти добуток $A{\cdot}B$ та $B{\cdot}A$ двох матриць

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 0 & 4 \\ 2 & 4 & -3 \\ -2 & 5 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 4 & 2 & 1 & 3 \\ -1 & 1 & 0 & 2 \end{pmatrix}.$$

Добуток $A\cdot B$ не існує, бо число стовпців матриці A не дорівнює числу рядків матриці B. Число стовпців матриці B дорівнює числу рядків матриці A. Отже існує добуток $B\cdot A$:

$$B\cdot A=\begin{pmatrix}4&2&1&3\\-1&1&0&2\end{pmatrix}\cdot\begin{pmatrix}1&-1&2\\3&0&4\\2&4&-3\\-2&5&0\end{pmatrix}=$$

$$=\begin{pmatrix}4\cdot1+2\cdot3+1\cdot3+3(-2) & 4(-1)+2\cdot0+1\cdot4+3\cdot5 & 4\cdot2+2\cdot4+1(-3)+3\cdot0\\(-1)\cdot1+1\cdot3+0\cdot2+2(-2) & (-1)(-1)+1\cdot0+0\cdot4+2\cdot5 & (-1)2+1\cdot4+0(-3)+2\cdot0\end{pmatrix}=$$

$$=\begin{pmatrix}6&15&13\\-2&11&2\end{pmatrix}.$$

3.62. Знайти матрицю $2A+5B$, якщо:

$$A=\begin{pmatrix}3&5&7\\2&-1&0\\4&3&2\end{pmatrix}\qquad B=\begin{pmatrix}1&2&4\\2&3&-2\\-1&0&1\end{pmatrix}$$

$$2A=\begin{pmatrix}6&10&14\\4&-2&0\\8&6&4\end{pmatrix},\qquad 5B=\begin{pmatrix}5&10&20\\10&15&-10\\-5&0&5\end{pmatrix}$$

$$2A+5B=\begin{pmatrix}11&20&34\\14&13&-10\\3&6&9\end{pmatrix}.$$

3.63. Знайти добуток $A\cdot B$ та добуток $B\cdot A$ двох матриць:

$$A=\begin{pmatrix}1&0&2&3\\2&1&0&3\\3&2&1&0\\0&1&-1&2\end{pmatrix},\qquad B=\begin{pmatrix}3&2&1&0\\0&3&4&1\\0&1&2&3\\1&0&1&-2\end{pmatrix}$$

3.64. Знайти добуток матриць $A \cdot B$:

$$A = \begin{pmatrix} -3 & 2 & -1 \\ 0 & -1 & 3 \\ 2 & 3 & -1 \end{pmatrix}; \qquad B = \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}$$

3.65. Знайти добуток матриць AB і BA, якщо:

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & 4 \\ 1 & 2 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 3 & 2 & 1 \end{pmatrix}$$

3.66. Знайти значення матричного многочлена:

$2A^2 + 3A + 5E$ при $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 1 \\ 4 & 1 & 1 \end{pmatrix}$, якщо E — одинична матриця третього порядку.

§8. Обернена матриця. Розв'язування матричних рівнянь

При розгляді дій з матрицями не вводиться операція ділення. Але можливо ввести поняття, яке дозволяє дати деякий еквівалент цієї дії.

Означення. *Квадратна матриця В називається оберненою квадратній матриці А, якщо добуток А·В є одинична матриця.*

Доведемо, що для будь-якої квадратної матриці А, визначник якої відмінний від нуля, існує одна і тільки одна обернена матриця, і наведемо спосіб її обчислення.

Нехай задана матриця

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \text{ і } \det A \neq 0.$$

Припустимо, що

$$B = \begin{pmatrix} b_{11} & b_{12} & ... & b_{1n} \\ b_{21} & b_{22} & ... & b_{2n} \\ ... & ... & ... & ... \\ b_{n1} & b_{n2} & ... & b_{nn} \end{pmatrix}$$

є шуканою матрицею і $E = \begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{pmatrix}$ — одиничною матрицею того самого порядку n.

Згідно з умовою $A \cdot B = E$, тому для визначення n^2 елементів b_{ik} матриці B маємо n систем рівнянь першого порядку, кожна з яких містить n рівнянь:

$$\begin{cases} a_{11}b_{11} + a_{12}b_{21} + ... + a_{1n}b_{n1} = 1 \\ a_{21}b_{11} + a_{22}b_{21} + ... + a_{2n}b_{n1} = 0 \\ .... \\ a_{n1}b_{11} + a_{n2}b_{21} + ... + a_{nn}b_{n1} = 0 \end{cases}$$

$$\begin{cases} a_{11}b_{12} + a_{12}b_{22} + ... + a_{1n}b_{n2} = 0 \\ a_{21}b_{12} + a_{22}b_{22} + ... + a_{2n}b_{n2} = 1 \\ .... \\ a_{n1}b_{12} + a_{n2}b_{22} + ... + a_{nn}b_{n2} = 0. \end{cases}$$

Такі системи мають одну і ту саму основну матрицю A.

Згідно з припущенням $\det A \neq 0$, тому кожна система має єдиний розв'язок, який можна обчислити за формулами Крамера. Оскільки в правій частині в кожній системі тільки один елемент дорівнює одиниці, а всі інші дорівнюють нулю, тоді

$$b_{11} = \frac{A_{11}}{\det A}, \quad b_{21} = \frac{A_{12}}{\det A}, \quad ......, \quad b_{n1} = \frac{A_{1n}}{\det A},$$

$$b_{12} = \frac{A_{21}}{\det A}, \quad b_{22} = \frac{A_{22}}{\det A}, \quad ......, \quad b_{n2} = \frac{A_{2n}}{\det A}$$

і взагалі $b_{ik} = \dfrac{A_{ki}}{\det A}$, $i,k=1,2,...,n$.

Отже, матриця B, обернена матриці A, яка позначається частіше символом A^{-1}, має вигляд

$$B = A^{-1} = \frac{1}{\det A}\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}. \tag{1}$$

Раніше було зазначено, що загалом, для довільних матриць A і B $A \cdot B \neq BA$. Але можливо довести, що $A^{-1} \cdot A = A \cdot A^{-1}$.

Дійсно $\left(A^{-1} \cdot A\right)_{ik} = \sum\limits_{l=1}^{n}(A^{-1})_{il} \cdot (A)_{lk} = \sum\limits_{l=1}^{n}\frac{A_{il}}{\det A}a_{lk} = \frac{1}{\det A}\sum\limits_{l=1}^{n}a_{lk}A_{li}$

Однак сума добутків елементів будь-якого рядка матриці на алгебраїчні доповнення відповідних елементів 2-го рядка, дорівнює нулю, а сума добутків елементів будь-якого рядка матриці на відповідні алгебраїчні доповнення елементів того самого рядка дорівнює самому визначнику.

Тому $(A^{-1} \cdot A)_{ik} = \begin{cases} 0 & i \neq k \\ 1 & i = k \end{cases}$

і, отже, $A^{-1} \cdot A = E = A \cdot A^{-1}$

Поняття "обернена матриця" може бути використане для розв'язання матричних рівнянь.

Нехай, наприклад, задане рівняння $AX = B$, де A і B — задані квадратні матриці порядку n, а X — шукана квадратна матриця того самого порядку. Нехай $\det A \neq 0$. Тоді обчислюємо матрицю A^{-1} і помножимо ліву і праву частини заданого рівняння зліва на A^{-1}:

$$A^{-1} \cdot (A \cdot X) = A^{-1} \cdot B$$

Оскільки

$$A^{-1} \cdot (A \cdot X) = (A^{-1} \cdot A) \cdot X$$

(згідно з асоціативною властивістю множення матриць), тоді

$$A^{-1} \cdot (A \cdot X) = E \cdot X = X$$

і одержуємо

$$X = A^{-1} \cdot B.$$

Для обчислення матриці A^{-1} оберненій матриці A, можливо, звичайно, використати формули (1). Проте, як правило, значно вигідніше використати для цього метод повного виключення. Це доцільно ще і тому, що всі n систем рівнянь, які служать для визначення стовпців матриці A^{-1}, різняться тільки правими частинами. Тому процес перетворення розширених матриць цих систем можна проводити одночасно для всіх матриць.

ЗАДАЧІ

3.67. Знайти матрицю, обернену матриці:

$$A=\begin{pmatrix}1 & 2 & -1\\ 0 & -1 & 1\\ 2 & 0 & -1\end{pmatrix}$$

Розглянемо матрицю

$$B^*=\left(\begin{array}{ccc|ccc}1 & 2 & -1 & 1 & 0 & 0\\ 0 & -1 & 1 & 0 & 1 & 0\\ 2 & 0 & -1 & 0 & 0 & 1\end{array}\right).$$

Перші три стовпці цієї матриці — стовпці заданої матриці A, наступні три стовпці, відділені рискою і складають разом одиничну матрицю, — стовпці вільних членів для систем рівнянь, які визначають елементи оберненої матриці.

Проводимо звичайні операції методом повного виключення:

$$B^*\Leftrightarrow\left(\begin{array}{ccc|ccc}1 & 2 & -1 & 1 & 0 & 0\\ 0 & 1 & -1 & 0 & -1 & 0\\ 2 & 0 & -1 & 0 & 0 & 1\end{array}\right)\Leftrightarrow\left(\begin{array}{ccc|ccc}1 & 2 & -1 & 1 & 0 & 0\\ 0 & 1 & -1 & 0 & -1 & 0\\ 0 & -4 & 1 & -2 & 0 & 1\end{array}\right)\Leftrightarrow\left(\begin{array}{ccc|ccc}1 & 0 & 1 & 1 & 2 & 0\\ 0 & 1 & -1 & 0 & -1 & 0\\ 0 & 0 & -3 & -2 & -4 & 1\end{array}\right)\Leftrightarrow$$

$$\Leftrightarrow\left(\begin{array}{ccc|ccc}1 & 0 & 1 & 1 & 2 & 0\\ 0 & 1 & -1 & 0 & -1 & 0\\ 0 & 0 & 1 & \frac{2}{3} & \frac{4}{3} & -\frac{1}{3}\end{array}\right)\Leftrightarrow\left(\begin{array}{ccc|ccc}1 & 0 & 0 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3}\\ 0 & 1 & 0 & \frac{2}{3} & \frac{1}{3} & -\frac{1}{3}\\ 0 & 0 & 1 & \frac{2}{3} & \frac{4}{3} & -\frac{1}{3}\end{array}\right).$$

$$A^{-1}=\begin{pmatrix} 1/3 & 2/3 & 1/3 \\ 2/3 & 1/3 & -1/3 \\ 2/3 & 4/3 & -1/3 \end{pmatrix}=\frac{1}{3}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & -1 \\ 2 & 4 & -1 \end{pmatrix}.$$

3.68. Знайти матрицю, обернену матриці:

$$A=\begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 1 & 1 & 2 \\ 1 & 2 & 3 & 0 \\ 2 & 2 & 2 & 5 \end{pmatrix}$$

Розглянемо матрицю:

$$B^*=\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 2 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 2 & 2 & 5 & 0 & 0 & 0 & 1 \end{array}\right)\Leftrightarrow\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -2 & -2 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right)\Leftrightarrow$$

$$\Leftrightarrow\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 & 2 & -1 & 0 & 0 \\ 0 & 1 & 2 & -2 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right)\Leftrightarrow\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & -4 & -3 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right)\Leftrightarrow$$

$$\Leftrightarrow\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 6 & 5 & -2 & -1 & 0 \\ 0 & 0 & 1 & -4 & -3 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right)\Leftrightarrow\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 17 & -2 & -1 & -6 \\ 0 & 0 & 1 & 0 & -11 & 1 & 1 & 4 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right)$$

$$A^{-1}=\begin{pmatrix} -1 & 1 & 0 & 0 \\ 17 & -2 & -1 & -6 \\ -11 & 1 & 1 & 4 \\ -2 & 0 & 0 & 1 \end{pmatrix}.$$

Другий спосіб знаходження оберненої матриці.

3.69. Знайти обернену A^{-1} матрицю до матриці A.

$$A = \begin{pmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{pmatrix}.$$

Обчислимо визначник матриці А:

$$\det A = \underset{Ip+(IIIp-IIp)}{\begin{vmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{vmatrix}} = \begin{vmatrix} 1 & 0 & 0 \\ 6 & 0 & 4 \\ 5 & -2 & -3 \end{vmatrix} = \begin{vmatrix} 3 & 4 \\ -2 & -3 \end{vmatrix} = -1 \neq 0.$$

Матриця A неособлива, оскільки $\det A = -1 \neq 0$.
Обчислимо алгебраїчні доповнення елементів цього визначника

$$A_{11} = \begin{vmatrix} 3 & -2 \\ 4 & -3 \end{vmatrix} = -1; \quad A_{12} = -\begin{vmatrix} 6 & 4 \\ 5 & -3 \end{vmatrix} = 38; \quad A_{13} = \begin{vmatrix} 6 & 3 \\ 5 & -2 \end{vmatrix} = -27$$

$$A_{21} = -\begin{vmatrix} 5 & 7 \\ -2 & -3 \end{vmatrix} = 1; \quad A_{22} = \begin{vmatrix} 2 & 7 \\ 5 & -3 \end{vmatrix} = -41; \quad A_{23} = -\begin{vmatrix} 2 & 5 \\ 5 & -2 \end{vmatrix} = 29$$

$$A_{31} = \begin{vmatrix} 5 & 7 \\ 3 & 4 \end{vmatrix} = -1; \quad A_{32} = -\begin{vmatrix} 2 & 7 \\ 6 & 4 \end{vmatrix} = 34; \quad A_{33} = \begin{vmatrix} 2 & 5 \\ 6 & 3 \end{vmatrix} = -24.$$

Згідно з формулою (1) записуємо A^{-1}

$$A^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ -38 & 41 & -34 \\ 27 & -29 & 24 \end{pmatrix}$$

3.70. Знайти матрицю обернену до даної:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 2 \\ 0 & 1 & -1 \end{pmatrix}.$$

Запишемо обернену матрицю у вигляді:

$$A^{-1}=\begin{pmatrix}\alpha_1 & \beta_1 & \gamma_1\\ \alpha_2 & \beta_2 & \gamma_2\\ \alpha_3 & \beta_3 & \gamma_3\end{pmatrix}.$$

Згідно з правилом множення матриць одержимо:

$$\begin{pmatrix}1 & 0 & 2\\ -1 & 1 & 2\\ 0 & 1 & -1\end{pmatrix}\cdot\begin{pmatrix}\alpha_1 & \beta_1 & \gamma_1\\ \alpha_2 & \beta_2 & \gamma_2\\ \alpha_3 & \beta_3 & \gamma_3\end{pmatrix}=$$

$$=\begin{pmatrix}\alpha_1+2\alpha_3 & \beta_1+2\beta_3 & \gamma_1+2\gamma_3\\ -\alpha_1+\alpha_2+2\alpha_3 & -\beta_1+\beta_2+2\beta_3 & -\gamma_1+\gamma_2+2\gamma_3\\ \alpha_2-\alpha_3 & \beta_2-\beta_3 & \gamma_2-\gamma_3\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Для знаходження елементів матриці A^{-1} запишемо системи:

$$\begin{cases}\alpha_1+2\alpha_3=1\\ -\alpha_1+\alpha_2+2\alpha_3=0\\ \alpha_2-\alpha_3=0\end{cases}\quad \begin{cases}\beta_1+2\beta_3=0\\ -\beta_1+\beta_2+2\beta_3=1\\ \beta_2-\beta_3=0\end{cases}\quad \begin{cases}\gamma_1+2\gamma_3=0\\ -\gamma_1+\gamma_2+2\gamma_3=0\\ \gamma_2-\gamma_3=1\end{cases}.$$

Розв'язки цих систем і дають нам елементи оберненої матриці:

$$A^{-1}=\begin{pmatrix}\frac{3}{5} & -\frac{2}{5} & \frac{2}{5}\\ \frac{1}{5} & \frac{1}{5} & \frac{4}{5}\\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5}\end{pmatrix}.$$

3.71. Знайти матрицю X з рівняння:

$$\begin{pmatrix}1 & 0 & 2\\ -1 & 1 & 2\\ 0 & 1 & -1\end{pmatrix}\cdot X=\begin{pmatrix}3 & -1 & 5\\ 2 & 2 & 3\\ 4 & 1 & 2\end{pmatrix}.$$

Помножимо обидві частини рівняння з лівого боку на матрицю, обернену до матриці: $A=\begin{pmatrix}1 & 0 & 2\\ -1 & 1 & 2\\ 0 & 1 & -1\end{pmatrix}$. Згідно з попереднім прикладом $A^{-1}=\frac{1}{5}\begin{pmatrix}3 & -2 & 2\\ 1 & 1 & 4\\ 1 & 1 & -1\end{pmatrix}$.

У лівій частині рівняння за асоціативним законом маємо:

$$A^{-1}\cdot(A\cdot X)=(A^{-1}\cdot A)\cdot X=EX=X$$

У правій частині буде:

$$\frac{1}{5}\begin{pmatrix}3 & -2 & 2\\ 1 & 1 & 4\\ 1 & 1 & -1\end{pmatrix}\cdot\begin{pmatrix}3 & -1 & 5\\ 2 & 2 & 3\\ 4 & 1 & 2\end{pmatrix}=\begin{pmatrix}\frac{13}{5} & -1 & \frac{13}{5}\\ \frac{21}{5} & 1 & \frac{16}{5}\\ \frac{1}{5} & 0 & \frac{6}{5}\end{pmatrix}=X.$$

Зауваження. Оскільки множення матриць некомутативне $A\cdot B\neq B\cdot A$, то в задачах такого типу потрібно уважно визначати, з якого боку слід множити обидві частини рівняння на матрицю, обернену до однієї з даних.

3.72. Розв’язати систему рівнянь:

$$\begin{cases}3x_1+2x_2+x_3=5\\ 2x_1-x_2+x_3=6\\ x_1+5x_2\qquad=-3\end{cases}$$

представивши її у вигляді матричного рівняння.

Перепишемо систему у вигляді $AX=B$, де

$$A=\begin{pmatrix}3 & 2 & 1\\ 2 & -1 & 1\\ 1 & 5 & 0\end{pmatrix},\quad X=\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix},\quad B=\begin{pmatrix}5\\ 6\\ -3\end{pmatrix}.$$

Розв’язок матричного рівняння має вигляд $X = A^{-1}B$. Знайдемо A^{-1}. Маємо

$$\det A = \begin{vmatrix} 3 & 2 & 1 \\ 2 & -1 & 1 \\ 1 & 5 & 0 \end{vmatrix} = 1\begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} - 5\begin{vmatrix} 3 & 1 \\ 2 & 1 \end{vmatrix} = 3 - 5 = -2 \neq 0 \ .$$

Обчислимо алгебраїчне доповнення елементів цього визначника.

$$A_{11} = \begin{vmatrix} -1 & 1 \\ 5 & 0 \end{vmatrix} = -5; \quad A_{12} = -\begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} = 1; \quad A_{13} = \begin{vmatrix} 2 & -1 \\ 1 & 5 \end{vmatrix} = 11$$

$$A_{21} = -\begin{vmatrix} 2 & 1 \\ 5 & 0 \end{vmatrix} = 5; \quad A_{22} = \begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} = -1; \quad A_{23} = -\begin{vmatrix} 3 & 2 \\ 1 & 5 \end{vmatrix} = -13$$

$$A_{31} = \begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} = 3; \quad A_{32} = -\begin{vmatrix} 3 & 1 \\ 2 & 1 \end{vmatrix} = -1; \quad A_{33} = \begin{vmatrix} 3 & 2 \\ 2 & -1 \end{vmatrix} = -7 \ .$$

Згідно з (1)

$$A^{-1} = -\frac{1}{2}\begin{pmatrix} -5 & 5 & 3 \\ 1 & -1 & -1 \\ 11 & -13 & -7 \end{pmatrix}.$$

Отже,

$$X = A^{-1} \cdot B = -\frac{1}{2}\begin{pmatrix} -5 & 5 & 3 \\ 1 & -1 & -1 \\ 11 & -13 & -7 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 6 \\ -3 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} -25 + 30 - 9 \\ 5 - 6 + 3 \\ 55 - 78 + 21 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} -4 \\ 2 \\ -2 \end{pmatrix}$$

$X = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$, тобто $x_1 = 2$, $x_2 = -1$, $x_3 = 1$.

§9. Модель багатогалузевої економіки

У макроекономіці виникає питання, пов'язане з ефективністю ведення багатогалузевого господарства: яким повинен бути обсяг виробництва кожної із *"n"* галузей, щоб задовольнити всі потреби в продукції цієї галузі? При цьому кожна галузь виступає, з одного боку, як виробник певної продукції, з іншого — як споживач продукції і своєї, і виробленої іншими галузями.

Розглянемо процес виробництва за проміжок часу (наприклад, рік). Введемо такі позначення:

x_i — загальний (валовий) обсяг продукції i-ої галузі ($i = 1, 2,\ldots, n$);

x_{ij} — обсяг продукції i-ої галузі, що споживається в j-ій галузі в процесі виробництва ($i, j = 1, 2,\ldots, n$);

y_i — обсяг кінцевої продукції i-ої галузі для невиробничого споживання.

Оскільки валовий обсяг продукції будь-якої i-ої галузі дорівнює сумарному обсягу продукції, що споживаються *"n"* галузями і кінцевої продукції, то

$$x_i = \sum_{j=1}^{n} x_{ij} + y_i, \qquad \left(i = 1,\ 2......n\right). \tag{1}$$

Рівняння (1) називається *співвідношенням балансу*. Будемо розглядати *вартісний міжгалузевий баланс*, коли всі величини, які входять в (1), мають вартісне вираження.

Введемо *коефіцієнти прямих витрат*

$$a_{ij} = \frac{x_{ij}}{x_j}, \quad \left(i,\ \ j = 1,\ 2.....n\right), \tag{2}$$

які відображають витрати продукції i-ої галузі на виробництво одиниці продукції j-ої галузі.

Можна вважати, що на деякому проміжку часу коефіцієнти a_{ij} будуть сталими і залежними від технології виробництва, що склалася. Отже це визначає лінійну залежність матеріальних витрат від валового випуску, тобто

$$x_{ij} = a_{ij} x_j, \quad \left(i, j = 1,\ 2,\ldots,n\right). \tag{3}$$

Внаслідок чого побудована на цій основі модель міжгалузевого балансу одержала назву лінійної.

Позначимо

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad A=\begin{pmatrix} a_{11} & a_{12}......a_{1n} \\ a_{21} & a_{22}......a_{2n} \\ .... & \\ a_{n1} & a_{n2}......a_{nn} \end{pmatrix}, \quad Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \tag{4}$$

де X — вектор валового випуску, Y — вектор кінцевого продукту, A — матриця прямих витрат. Тоді систему можна записати в матричному вигляді:

$$X = AX + Y. \tag{5}$$

Основне завдання міжгалузевого балансу полягає у відшуканні такого вектора валового випуску X, який при відомій матриці прямих витрат A забезпечує заданий вектор кінцевого продукту Y.

Перепишемо рівняння (5) у вигляді:

$$(E - A)X = Y. \tag{6}$$

Якщо матриця $(E - A)$) невироджена. Тобто $\left|E - A\right| \neq 0$, то

$$X = \left(E - A\right)^{-1} Y. \tag{7}$$

Матриця $S = \left(E - A\right)^{-1}$ називається матрицею повних витрат, кожний елемент S_{ij} якої є величиною валового випуску продукції i-ої галузі, необхідного для забезпечення випуску одиниці кінцевого продукту j-ої галузі y_j = (j = 1, 2,…, n).

Відповідно до економічного змісту задачі значення x_i повинні бути невід'ємними при невід'ємних значеннях $y_i \geq 0$ та $a_{ij} \geq 0$, де i, j = 1, 2,…, n.

Матриця $A \geq 0$ називається *продуктивною*, якщо для будь-якого вектора $Y \geq 0$ існує розв'язок $X \geq 0$ рівняння (2). У цьому випадку і модель називають *продуктивною*.

Існує кілька критеріїв продуктивності матриці A. Один із них полягає у тому, що матриця продуктивна, якщо $a_{ij} \geq 0$ для будь-яких i, j = 1, 2,…, n, і максимум сум елементів її стовпців не перевищує одиниці

$$\max_{j=1,\, 2....n} \sum_{i=1}^{n} a_{ij} \leq 1,$$

причому хоча б для одного із стовпців сума елементів строго менше одиниці

$$\sum_{i=1}^{n} a_{ij} \le 1 .$$

ЗАДАЧІ

3.73. Методом Гаусса розв'язати, у випадку сумісності, систему лінійних рівнянь. Указати вільні змінні, базисні змінні та базисний розв'язок, який їм відповідає. Перевірити цей розв'язок підстановкою.

$$\begin{cases} x_1 + x_2 - x_3 + x_4 = 2 \\ x_1 - 2x_2 + x_3 - x_4 = -3 \\ 2x_1 + x_2 \quad + 2x_4 = 5 \\ 5x_1 - 2x_2 + x_3 + x_4 = 1 \end{cases}.$$

Розрахунки за цим методом представимо у вигляді таблиці 3.1, до якої заносимо розширену матрицю системи.

Таблиця 3.1

№ ітерації	x_1	x_2	x_3	x_4	a_{ik}
0	1	1	–1	1	2
	1	–2	1	–1	–3
	2	1	0	2	5
	5	–2	1	1	1
1	2	–1	0	0	–1
	1	–2	1	–1	–3
	2	1	0	2	5
	4	0	0	2	4
2	4	0	0	2	4
	5	0	1	3	7
	2	1	0	2	5
	4	0	0	2	4
3	2	0	0	1	2
	–1	0	1	0	1
	–2	1	0	0	1
	0	0	0	0	0
3′	2	0	0	1	2
	–1	0	1	0	1
	–2	1	0	0	1

На кожній ітерації вибираються головний рядок та головний стовпчик, на перетині яких знаходиться головний елемент.

Для спрощення обчислень зручно за головний елемент вибирати елемент, рівний 1, та за головний стовпчик вибирати стовпчик, який містить якомога більше нулів.

Перша ітерація. Вибираємо третій головний стовпчик і другий головний рядок. На їх перетині стоїть головний елемент a_{23}, який ми виділимо рамкою. Тут і надалі через a_{qp} ми позначимо елемент, який стоїть на перетині рядка з номером q і стовпчика з номером p.

Далі перераховуємо елементи головного рядка за формулою:

$$a'_{qk} = \frac{a_{qk}}{a_{qp}}, \tag{1}$$

де $k = 0, 1, 2, ..., n$; q — номер головного рядка, p — номер головного стовпчика, a'_{qk} — елементи нової матриці, яка відповідає першій ітерації.

Оскільки для даного прикладу $a_{qp} = a_{23} = 1$, то, згідно з формулою (1), всі елементи головного рядка необхідно поділити на 1, а отже, переписати без зміни.

Елементи інших рядків обчислюються за формулою:

$$a'_{ik} = a_{ik} - \frac{a_{qk}}{a_{qp}} a_{ip}$$

або:

$$a'_{ik} = a_{ik} - a'_{qk} a_{ip} \tag{2}$$

де $i = 1, 2,..., n$; $k = 0, 1, 2,..., n$; q — номер в головного рядка ($q = 2$); p — номер ведучого стовпчика ($p = 3$).

Оскільки $a_{33} = 0$, то третій рядок, згідно з формулою (2), перепишеться без зміни. Для елементів першого рядка маємо:

$$a'_{11} = a_{11} - a'_{21}a_{13} = 1 - 1\cdot(-1) = 2;$$
$$a'_{12} = a_{12} - a'_{22}a_{13} = 1 - (-2)\cdot(-1) = -1;$$
$$a'_{13} = a_{13} - a'_{23}a_{13} = 0;$$
$$a'_{14} = a_{14} - a'_{24}a_{13} = 1 - (-1)\cdot(-1) = 0;$$
$$a'_{10} = a_{10} - a'_{20}a_{13} = 2 - (-3)\cdot(-1) = -1.$$

Елементи четвертого рядка обчислюються аналогічно: $a'_{41} = 4$, $a'_{42} = 0$, $a'_{43} = 0$, $a'_{44} = 2$, $a'_{40} = 4$.

Після цих обчислень головний стовпчик має перетворитись в одиничний.

Зауважимо, що на наступних ітераціях ні другий рядок, ні третій стовпчик вже не можуть бути вибраними за головні.

Друга ітерація. За головний вибираємо елемент $a'_{32}=1$ (головний рядок — третій, головний стовпчик — другий) Проводимо обчислення, аналогічні першій ітерації.

Третя ітерація. $a''_{14} = 2$ — головний елемент.

На третій ітерації з'явився нульовий рядок, який можна відкинути (крок 3').

Після останньої ітерації жоден рядок не може бути вибраний за головний, отже, розрахунки закінчені.

Оскільки ми отримали — три лінійно незалежних одиничних стовпчики при чотирьох змінних, то ця система рівнянь невизначена. Змінні, які відповідають лінійно незалежним одиничним стовпчикам, можуть бути вибрані за базисні. Для даного прикладу x_2, x_3, x_4 — базисні змінні. Усі інші (тобто x_1) — вільні.

Оскільки остання таблиця відповідає системі

$$\begin{cases} 2x_1 \qquad\qquad + x_4 = 2 \\ -x_1 \qquad + x_3 \qquad = 1 \\ -2x_1 + x_2 \qquad\qquad = 1, \end{cases}$$

то загальний розв'язок вихідної системи має вигляд (базисні змінні виражаються через вільні):

$$\begin{cases} x_2 = 1 + 2x_1 \\ x_3 = 1 + x_1 \\ x_4 = 2 - 2x_1. \end{cases}$$

Прирівнявши всі вільні змінні до нуля (тобто $x_1 = 0$), знайдемо базисний розв'язок:

$x_1 = 0$, $x_2 = 1$, $x_3 = 1$, $x_4 = 2$.

Перевіримо цей розв'язок підстановкою у вихідну систему:

$$\begin{cases} 0 + 1 - 1 + 2 = 2 \\ 0 - 2 \cdot 1 + 1 - 2 = -3 \\ 2 \cdot 0 + 1 + 2 \cdot 2 = 5 \\ 5 \cdot 0 - 2 \cdot 1 + 1 + 2 = 1. \end{cases}$$

Отже, ці значення є розв'язком.

3.74. Для виробництва продукції створено три фірми, кожна з яких випускає один вид продукції. У таблиці 3.2 задані:

— коефіцієнти прямих витрат a_{ik}, тобто кількість одиниць продукції i-ї фірми, яка використовується як проміжний продукт для випуску одиниці продукції k-ї фірми;

— кількість одиниць y_i продукції i-ї фірми, розрахованих на реалізацію (кінцевий продукт).

Визначити:

а) коефіцієнт повних витрат;

б) валовий випуск (план) для кожної фірми;

в) коефіцієнти непрямих витрат.

Таблиця 3.2.

Фірми	**Прямі витрати a_{ik}**			**Кінцевий продукт**
	I	**II**	**III**	y_i
I	0,1	0,2	0	10
II	0,1	0	0,2	30
III	0	0,2	0	20

Розв'язок. Нехай $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ — виробнича програма фірм, де x_i — валовий випуск продукції i-ї фірми (i=1,2,3).

Позначимо через $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ — план випуску товарної продукції (розрахованої на реалізацію). Матрицю коефіцієнтів прямих витрат позначимо через $A=\| a_{ik} \|$. Вектор Y та матриця A задано в таблиці.

Згідно з умовою задачі i-а фірма віддає рівно $a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3$ одиниць продукції на внутрішні потреби. Тоді виробничі зв'язки фірм можуть бути представлені за допомогою системи трьох рівнянь $x_i = y_i + a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3$, $i = 1, 2, 3$.

Іншими словами, валовий випуск продукції x_i складається з випуску товарної продукції y_i та випуску продукції для внутрішніх потреб.

У матричній формі цю рівність можна переписати:

$$Y + AX = X \text{ або } X - AX = Y.$$

Якщо E — одинична матриця третього порядку, то останнє рівняння перепишеться у вигляді $(E - A)X = Y$.

Його розв'язок у матричній формі має вигляд:

$$X = (E - A)^{-1} Y, \qquad (3)$$

де $(E - A)^{-1}$ — обернена матриця.

а) Елементи матриці $(E - A)^{-1}$ є ні що інше, як шукані коефіцієнти повних витрат. Позначимо $S = E - A$, тобто,

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0,1 & 0,2 & 0 \\ 0,1 & 0 & 0,2 \\ 0 & 0,2 & 0 \end{pmatrix} = \begin{pmatrix} 0,9 & -0,2 & 0 \\ -0,1 & 1 & -0,2 \\ 0 & -0,2 & 1 \end{pmatrix}.$$

Знайдемо матрицю S^{-1} методом Гаусса.

Розрахунки представимо у вигляді таблиці 3.3, у лівій частині якої записуємо вихідну матрицю S, справа — матрицю E.

Відповідні перетворення рядків таблиці проводимо так само, як і при розв'язуванні системи рівнянь (див. табл. 3.1), намагаючись отримати одиничні стовпчики (ітерації 1—3).

Таблиця 3.3

№ ітерації	матриця *S*			матриця *E*		
0	0,9	–0,2	0	1	0	0
	–0,1	1	–0,2	0	1	0
	0	–0,2	1	0	0	1
1	0,9	–0,2	0	1	0	0
	–0,1	0,96	0	0	1	0,2
	0	–0,2	1	0	0	1
2	0	8,44	0	1	9	1,8
	1	–9,6	0	0	–10	–2
	0	–0,2	1	0	0	1
3	0	1	0	0,12	1,07	0,21
	1	0	0	1,15	0,27	0,02
	0	0	1	0,02	0,21	1,04
3′	1	0	0	1,15	0,27	0,02
	0	1	0	0,12	1,07	0,21
	0	0	1	0,02	0,21	1,04

Якщо вихідна матриця невироджена, то після проведення n ітерацій (n-порядок системи) отримаємо n одиничних стовпчиків. Якщо вихідна матриця вироджена, то після деякої ітерації в лівій частині таблиці з'явиться нульовий рядок. Це буде свідчити про те, що оберненої матриці не існує.

Для нашого прикладу ми отримали три одиничних стовпчика. На останньому кроці ($3'$) шляхом перестановки рядків утворюємо в лівій частині таблиці одиничну матрицю. Тоді в правій частині таблиці буде записана обернена матриця.

Отже, матриця коефіцієнтів повних витрат має такий вигляд:

$$S^{-1} = \begin{pmatrix} 1,15 & 0,27 & 0,02 \\ 0,12 & 1,07 & 0,21 \\ 0,02 & 0,21 & 1,04 \end{pmatrix}.$$

Таким чином, наприклад, для випуску одиниці продукції першій, другій та третій фірм необхідно витратити, відповідно, 1,15; 0,27 та 0,02 одиниць продукції першої фірми.

б) Для визначення валового випуску продукції використовуємо рівність (3):

$$X = (E - A)^{-1} = \begin{pmatrix} 1,15 & 0,27 & 0,02 \\ 0,12 & 1,07 & 0,21 \\ 0,02 & 0,21 & 1,04 \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 30 \\ 20 \end{pmatrix} = \begin{pmatrix} 20 \\ 37,5 \\ 27,3 \end{pmatrix}.$$

Отже, $x_1 = 20$, $x_2 = 37,5$, $x_3 = 27,3$.

в) Коефіцієнти непрямих витрат знайдемо як різницю між повними витратами S^{-1} та прямими витратами A, або, в матричній формі:

$$S^{-1} - A = \begin{pmatrix} 1,05 & 0,07 & 0,02 \\ 0,02 & 1,07 & 0,01 \\ 0,02 & 0,01 & 1,04 \end{pmatrix}.$$

ЗАДАЧІ

Знайти обернені матриці:

3.75. $\begin{pmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{pmatrix}$.   **3.76**. $\begin{pmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & -1 \end{pmatrix}$.

3.77. $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$.

3.78. $\begin{pmatrix} 2 & 7 & 3 \\ 3 & 9 & 4 \\ 1 & 5 & 3 \end{pmatrix}$.

3.79. $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}$.

3.80. $\begin{pmatrix} 3 & 3 & -4 & -3 \\ 0 & 6 & 1 & 1 \\ 5 & 4 & 2 & 1 \\ 2 & 3 & 3 & 2 \end{pmatrix}$.

Розв'язати матричні рівняння:

3.81. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot X = \begin{pmatrix} 3 & 5 \\ 5 & 9 \end{pmatrix}$.

3.82. $X \cdot \begin{pmatrix} 3 & -2 \\ 5 & -4 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -5 & 6 \end{pmatrix}$.

3.83. $\begin{pmatrix} 1 & 2 & -3 \\ 3 & 2 & -4 \\ 2 & -1 & 0 \end{pmatrix} \cdot X = \begin{pmatrix} 1 & -3 & 0 \\ 10 & 2 & 7 \\ 10 & 7 & 8 \end{pmatrix}$.

3.84. $X \cdot \begin{pmatrix} 5 & 3 & 1 \\ 1 & -3 & -2 \\ -5 & 2 & 1 \end{pmatrix} = \begin{pmatrix} -8 & 3 & 0 \\ -5 & 9 & 0 \\ -2 & 15 & 0 \end{pmatrix}$

Розв'язати систему, представивши їх у вигляді матричних рівнянь:

3.85. $\begin{cases} 2x_1 + 3x_2 + 2x_3 = 9 \\ x_1 + 2x_2 - 3x_3 = 14 \\ 3x_1 + 4x_2 + x_3 = 16 \end{cases}$.

3.86. $\begin{cases} x_1 - x_2 + x_3 = 6 \\ 2x_1 + x_2 + x_3 = 3 \\ x_1 + x_2 + 2x_3 = 5 \end{cases}$.

3.87. Для виробництва промислової продукції створено три фірми, кожна з яких випускає один вид продукції. У таблиці задані:

коефіцієнти прямих витрат a_{ik}, тобто кількість одиниць продукції i-ї фірми, яка використовується як проміжний продукт для випуску одиниці продукції k-ї фірми;

кількість одиниць y_i продукції фірми, розрахованих на реалізацію (кінцевий продукт).

Визначити:

а) коефіцієнти повних витрат;

б) валовий випуск (план) для кожної фірми;

в) коефіцієнти непрямих витрат.

Таблиця 3.4

Варіант	**Фірми**	**Прямі витрати a_{ik}**			**Кінцевий продукт**
		I	**II**	**III**	y_i
1	I	0,1	0	0,1	30
	II	0,3	0	0,1	40
	III	0,1	0,2	0	50
2	I	0,2	0,1	0	50
	II	0	0,3	0,1	60
	III	0,2	0,1	0	20
3	I	0	0,2	0.2	40
	II	0,3	0,1	0	50
	III	0,1	0,2	0,1	60
4	I	0,1	0.2	0	100
	II	0,2	0,1	0,3	120
	III	0	0,1	0	110
5	I	0	0,1	0,3	70
	II	0,1	0,1	0	100
	III	0	0,2	0,2	120
6	I	0,2	0,1	0	180
	II	0,2	0	0.1	200
	III	0,1	0,1	0	210
7	I	0	0,3	0,1	30
	II	0,3	0,2	0,1	60
	III	0,1	0,2	0	100
8	I	0,2	0,1	0,1	400
	II	0,2	0,1	0	300
	III	0,1	0,1	0	200

ВІДПОВІДІ ДО РОЗДІЛУ III

3.8. 18. **3.9.** $4ab$ **3.10**. 0. **3.11.** 0.

3.12. 900. **3.13.** 1. **3.14.** 160. **3.15.** –98.

3.16. $(-1)^{n-1}5^n$ **3.17.** $n!$. **3.18.** $(-1)^{\frac{n(n-1)}{2}}\cdot x\cdot(x-1)\cdot(x-2)\cdot\ldots\cdot(x-n)$.

3.19. –264. **3.27.** $x_1 = 16$, $x_2 = 7$. **3.28.** $x_1 = 3$, $x_2 = 2$.

3.29. $x_1 = 2$, $x_2 = -1$, $x_3 = 1$. **3.30.** $x_1 = 3$, $x_2 = 1$, $x_3 = -1$.

3.31. $x_1 = 2$, $x_2 = -1$, $x_3 = 2$, $x_4 = -2$. **3.32.** $x_1 = 2$, $x_2 = x_3 = x_4 = 0$.

3.33. $x_1 = 2$, $x_2 = -2$, $x_3 = 1$, $x_4 = -1$.

3.34. $x_1 = a_4 - \frac{53}{18}a_5 + \frac{20}{9}$, $x_2 = -\frac{5}{2}a_4 + \frac{5}{6}a_5 - \frac{5}{3}$, $x_3 = \frac{2}{9}a_5 - \frac{1}{9}$, $x_4 = a_4$, $x_5 = a_5$, a_4, a_5 — довільні числа.

3.35. Не сумісна. **3.47.** $r = 2$. **3.48.** $r = 3$. **3.49.** $r = 2$.

3.50. $r = 3$. **3.51.** $r = 2$. **3.52.** $r = 3$. **3.53.** $r = 3$.

3.54. $r = 2$. **3.55.** система має тільки тривіальний розв'язок.

3.56. $c_1e_1 + c_2e_2$, $e_1 = (8, -6, 1, 0)$, $e_2 = (-7, 5, 0, 1)$.

3.57. $c_1e_1 + c_2e_2$, $e_1 = (1, 0, -\frac{5}{2}, \frac{7}{2})$, $e_2 = (0, 1, 5, -7)$.

3.58. $c_1e_1 + c_2e_2 + c_3e_3$, $e_1 = (1, 0, 0, -\frac{9}{4}, \frac{3}{4})$, $e_2 = (0, 1, 0, -\frac{3}{2}, \frac{1}{2})$, $e_3 = (0, 0, 1, -2, 1)$.

3.59. $c_1e_1 + c_2e_2$, $e_1 = (0, \frac{1}{3}, 1, 0, 0)$, $e_2 = (0, -\frac{2}{3}, 0, 0, 1)$.

3.60. $c_1e_1 + c_2e_2$, $e_1 = (-3, 2, 1, 0, 0)$, $e_2 = (-5, 3, 0, 0, 1)$.

3.63. $AB = \begin{pmatrix} 9 & 4 & 8 & 0 \\ 9 & 7 & 9 & -5 \\ 9 & 13 & 13 & 5 \\ 2 & 2 & 4 & -6 \end{pmatrix}$, $BA = \begin{pmatrix} 10 & 4 & 7 & 15 \\ 18 & 12 & 3 & 11 \\ 8 & 8 & -1 & 9 \\ 4 & 0 & 5 & -1 \end{pmatrix}$.

3.64. $\begin{pmatrix} -1 \\ 5 \\ 13 \end{pmatrix}$. **3.65.** $AB = \begin{pmatrix} 8 & 0 & 7 \\ 16 & 10 & 4 \\ 13 & 5 & 7 \end{pmatrix}$, $BA = \begin{pmatrix} 4 & 6 & 6 \\ 1 & 7 & 3 \\ 8 & 11 & 14 \end{pmatrix}$.

3.66. $2A^2 + 3A + 5E = \begin{pmatrix} 28 & 15 & 16 \\ 19 & 36 & 15 \\ 30 & 19 & 28 \end{pmatrix}$. **3.75.** $\begin{pmatrix} 1 & -1 & 1 \\ -38 & 41 & -34 \\ 27 & -29 & 24 \end{pmatrix}$.

3.76. $\begin{pmatrix} -8 & 29 & -11 \\ -5 & 18 & -7 \\ 1 & -3 & 1 \end{pmatrix}$. **3.77.** $\begin{pmatrix} 1/4 & 1/4 & 1/2 & 0 \\ 1/4 & 1/4 & -1/2 & 0 \\ 1/4 & -1/4 & 0 & 1/2 \\ 1/4 & -1/4 & 0 & -1/2 \end{pmatrix}$.

3.78. $\begin{pmatrix} -7/3 & 2 & -1/3 \\ 5/3 & -1 & -1/3 \\ 2 & 1 & 1 \end{pmatrix}$. **3.79.** $\begin{pmatrix} 1/9 & 2/9 & 2/9 \\ 2/9 & 1/9 & -2/9 \\ 2/9 & -2/9 & 1/9 \end{pmatrix}$.

3.80. $\begin{pmatrix} -7 & 5 & 12 & -19 \\ 3 & -2 & -5 & 8 \\ 41 & -30 & -69 & 111 \\ -59 & 43 & 99 & -159 \end{pmatrix}$. **3.81.** $\begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix}$.

3.82. $\begin{pmatrix} 3 & -2 \\ 5 & -4 \end{pmatrix}$. **3.83.** $\begin{pmatrix} 6 & 4 & 5 \\ 2 & 1 & 2 \\ 3 & 3 & 3 \end{pmatrix}$. **3.84.** $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

3.85. $x_1 = 2, x_2 = 3, x_3 = -2$.
3.86. $x_1 = 1, x_2 = -2, x_3 = 3$.

Розділ IV
ЛІНІЙНЕ ПРОГРАМУВАННЯ

§ 1. Теоретичні основи методів лінійного програмування

Математичне програмування — прикладна галузь математики, яка широко використовується в таких питаннях, як планування господарства, організація управління промисловістю, планування військової операцій і тощо. Задачі управління і планування, як правило, зводяться до вибору деякої системи числових параметрів або функцій (характеристики плану), що забезпечують найбільш ефективне досягнення поставленої мети (оптимальний план), з урахуванням обмеженої можливості ресурсів. Для оцінки ефективності плану вводиться так звана *цільова функція* (тобто показник якості плану), виражена через характеристики плану і яка приймає екстремальне значення (тобто найменше і найбільше значення) для оптимального плану.

Для великої кількості практично цікавих задач цільова функція виражається *лінійно* через характеристики плану, причому допустимі значення параметрів підпорядковані також лінійним рівностям або нерівностям. Знаходження при даних умовах абсолютного екстремуму цільової функції носить назву *лінійного програмування* (більш вдалим був би термін “лінійне планування”).

Задача про оптимальне використання ресурсів. Складемо економіко-математичну модель однієї із важливих задач лінійного програмування, а саме задачі про використання ресурсів, оскільки перш ніж застосовувати математичні методи для розв’язання задач економіки, необхідно виразити економічний зміст задачі через певні математичні залежності, тобто скласти так звану економіко-математичну модель задачі.

Вважаємо, що підприємство виготовляє n різних виробів. Для їх виробництва потрібно m різних видів ресурсів (сировини, допоміжних матеріалів, запасів машинного часу, людських резервів тощо.). Ці ресурси обмежені і складають у запланований період відповідно b_1, b_2,..., b_m умовних одиниць.

Відомі також технологічні коефіцієнти a_{ij}, які показують скільки одиниць i-го ресурсу потрібно для виробництва одиниці j-го виду виробу ($i = 1, 2, ..., m$; $j = 1, 2, ..., n$). Нехай прибуток, отриманий підприємством від реалізації одиниці виробу j-го виду, дорівнює c_j ($j = 1, 2, ..., n$). У запланований період усі показники a_{ij}, b_i та c_j вважатимемо "сталими".

Слід скласти такий план випуску продукції, за якого прибуток підприємства від її реалізації був би найбільшим.

Зведемо задані умови в таблицю.

Таблиця 4.1

Види ресурсів	**Запаси ресурсів**	**Технологічні коефіцієнти**					
		1	2	...	j	...	n
1	b_1	a_{11}	a_{12}	...	a_{1j}	...	a_{1n}
2	b_2	a_{21}	a_{22}	...	a_{2j}	...	a_{2n}
...	...	...	...	...	...	...	...
i	b_i	a_{i1}	a_{i2}	...	a_{ij}	...	a_{in}
...	...	...	...	...	...	...	...
m	b_m	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mn}
Прибуток		c_1	c_2	...	c_j	...	c_n

Вважатимемо, що підприємство буде випускати x_1 одиниць виробів 1-го виду, x_2 одиниць виробів 2-го виду, ..., x_n одиниць виробів n-го виду.

Очевидно, мають виконуватися обмеження

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \le b_1, \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \le b_2, \\ .... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \le b_m, \end{cases}$$

причому $x_1 \ge 0,\ x_2 \ge 0, ..., x_n \ge 0$, тобто змінні x_j ($j = 1, 2, ..., n$) повинні бути невід'ємні.

Система лінійних нерівностей разом з умовою невід'ємності змінних складає систему обмежень даної задачі із об'єму відповідного ресурсу, оскільки у ході виконання плану можливо використовувати або увесь запас цього ресурсу, або частину його.

Потрібно скласти оптимальний план роботи підприємства $X(x_1; x_2;...; x_n)$, тобто знайти такі невід'ємні значення $x_1; x_2;...; x_n$, які б задовольняли системі обмежень і за яких прибуток від реалізації всієї продукції $F = c_1x_1 + c_2x_2 +... + c_nx_n$ був максимальний.

Функція F виражає кінцеву мету оптимального планування в даному випуску — одержання найбільшого прибутку, а тому цю функцію називають *цільовою*.

Оскільки змінні $x_1; x_2;...; x_n$ входять і в цільову функцію і в систему обмежень у першому степені, а показники a_{ij}, b_i, c_j ($i = 1, 2, ..., m$; $j = 1, 2, ..., n$) є сталими, то поставлена задача є типовою задачею лінійного програмування.

Задача лінійного програмування. Задача лінійного програмування загалом зводиться до відшукання такого розв'язку $X(x_1; x_2;...; x_n)$ системи m лінійних рівнянь з n змінними (системи обмежень)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m, \\ x_j \geq 0 \ \left(j = 1, 2, ..., n\right), \end{cases} \tag{1}$$

за якого цільова функція (лінійна форма) $F = c_1x_1 + c_2x_2 +... + c_nx_n$ прийме оптимальне (максимальне або мінімальне) значення.

Розв'язок $X(x_1; x_2;...; x_n)$, за яким функція F обертається в оптимум (мінімум або максимум), називається *оптимальним розв'язком.*

Задача лінійного програмування і полягає у відшуканні цього оптимального розв'язку, виходячи з того, що оптимальний розв'язок — єдиний, хоча на практиці можуть зустрічатися задачі, де ця єдиність оптимального розв'язку порушується, що в подальшому буде вказано окремо.

У задачах лінійного програмування коефіцієнтам a_{ij}, b_i, c_j задані сталі величини, число рівнянь менше числа змінних, тобто $m < n$, а всі праві частини системи (1) без обмеження загальності, можна вважати невід'ємними, тобто $b_i \geq 0$ ($i = 1, 2, ..., m$. Якщо у деяких рівняннях системи (1) така умова порушена, то можна помножити обидві частини таких рівнянь на (–1).

Сформульовану задачу можливо записати в більш короткій формі, а саме: знайти розв'язок $X(x_1; x_2;...; x_n)$ системи обмежень

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \ \left(i = 1, 2, ..., m\right), \tag{2}$$

що приводить до оптимуму лінійну форму

$$F = \sum_{j=1}^{n} c_j x_j \ , \tag{3}$$

за умови, що $x_j \geq 0 \ \left(j = 1, 2, ..., n\right)$.

Система (2) може бути записана у векторно-матричній формі

$$AX = B, \tag{4}$$

де

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; \ B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \ A = \left(a_{ij}\right) = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix}$$

— відповідно вектор-стовпець змінних, вектор-стовпець вільних членів і матриця коефіцієнтів при змінних.

Матрицю A називають *матрицею умов задачі*, оскільки коефіцієнти a_{ij} представляють техніко-економічні або технологічні показники. Вектор $B = (b_1; b_2; ...; b_m)$ називають *вектором обмежень*, а задача лінійного програмування, система обмежень якої задана у вигляді системи рівнянь (1) має назву *канонічної*.

У багатьох задачах обмеження задаються не у вигляді системи рівнянь, а у вигляді системи лінійних нерівностей, причому можливі різні форми таких систем, наприклад:

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \ \left(i = 1, 2, ..., m\right) \tag{5}$$

або

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i \ \left(i = 1, 2, ..., m\right). \tag{6}$$

Крім того, система обмежень може бути змішаною: частина обмежень — нерівності типу (5), частина — типу (6), частина задана у вигляді рівнянь.

Насамкінець, будь-яку систему обмежень можливо привести до системи рівнянь виду (1). Для цього достатньо до лівої частини кожної нерівності (5) додати (відняти, якщо система задана у вигляді (6)) певне невід'ємне число — додаткову змінну, з тим, щоб кожна нерівність оберталась у рівняння.

Нехай система обмежень задана нерівностями виду (5). Додавши до лівої частини кожної нерівності відповідну додаткову невід'ємну змінну $x_{n+1}, x_{n+2}, ..., x_{n+m}$, де $x_{n+i} \geq 0$ $(i = 1, 2, ..., m)$, одержимо замість систем нерівностей (5) еквівалентну систему рівнянь:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n + x_{n+1} = b_1, \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n + x_{n+2} = b_2, \\ .... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n + x_{n+m} = b_m, \end{cases}$$

тобто систему обмежень, аналогічну системі (1).

Якщо в системі є нерівності виду (5) і (6), а частина задана у вигляді рівнянь (змішана система обмежень), число додаткових змінних менше m і дорівнює $m - t$, де m — загальне число обмежень системи, а t — число обмежень у вигляді рівнянь.

Таким чином, як би не були задані обмеження лінійного програмування, їх завжди можна звести до системи лінійних рівнянь, використовуючи для цього додаткові змінні, тобто звести задачу до канонічної.

Відомо, що система m лінійних рівнянь з n невідомими може бути як сумісною так і несумісною, а рівняння системи як залежними так і незалежними.

Несумісність системи означає, що її рівняння мають протиріччя. У цьому випадку система не має ні одного розв'язку, а отже, не має розв'язку і задача лінійного програмування. Будемо вважати, що усі m рівнянь системи (1) лінійно незалежні. У протилежному випадку із системи можливо виключити частину рівнянь так, щоб рівняння системи стали лінійно незалежними.

У випадку *векторної форми* запису системи обмежень, коли розглядається $n + 1$ векторів m-вимірного простору:

$$P_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \ P_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, ..., \ P_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}, \ P = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

систему (1) можна записати у вигляді одного векторного рівняння

$$P_1x_1 + P_2x_2 + ... + P_nx_n = P \tag{7}$$

і лінійна незалежність означає, що серед векторів P_1, P_2,..., P_n лише m векторів лінійно незалежні.

Як відомо, сумісна система m лінійних рівнянь з n змінними ($m < n$) має нескінченну множину всіляких розв'язків, зокрема і допустимих (які не мають від'ємних компонент), а кількість її базисних розв'язків не перевищує числа C_n^m .

Нагадаємо також, що допустимим базисним розв'язком є розв'язок, який містить m невід'ємних основних (базисних) змінних і $n - m$ неосновних (небазисних, або вільних) змінних. Неосновні змінні в базисному розв'язку дорівнюють нулю. Основні ж змінні, як правило, відмінні від нуля, тобто є додатними числами.

Якщо хоча б одна із основних змінних приймає нульове значення, то відповідний базисний розв'язок називається виродженим. За основні змінні можна вважати будь-які m змінних, якщо тільки визначник, складений із коефіцієнтів при них, відмінний від нуля або відповідні їм m векторів із загального числа n векторів P_1, P_2,..., P_n системи обмежень лінійно незалежні, тобто утворюють базис m-вимірного простору.

Оскільки компоненти оптимального розв'язку задачі лінійного програмування не можуть бути від'ємними, то пошук цього розв'язку потрібно обмежити тільки допустимими розв'язками системи обмежень. Для цього знадобиться ряд теорем, що належать до основних теорем лінійного програмування, доведення яких наводиться в більш детальних підручниках, присвячених як теорії лінійного програмування так і математичному програмуванню в цілому.

ТЕОРЕМА 1. *Множина всіх допустимих розв'язків системи обмежень задачі лінійного програмування є випуклою (див. §2).*

ТЕОРЕМА 2. *Якщо існує, і причому єдиний, оптимальний розв'язок задачі лінійного програмування, то він співпадає з однією з кутових точок множини допустимих розв'язків.*

Ця теорема показує, що пошук оптимального розв'язку можливо обмежити перебором скінченого числа кутових точок для відшукання яких потрібно побудувати області розв'язку системи обмежень.

Така побудова можлива тільки для дво- або трьохвимірного простору, а в загальному випадку задача залишається нерозв'язною. Отже, слід оперувати певним аналітичним методом, який дозволив би знаходити координати кутових точок. Для цього має місце наступна теорема.

ТЕОРЕМА 3. *Кожному допустимому базисному розв'язку задачі лінійного програмування відповідає кутова точка області допустимих розв'язків системи обмежень.*

ТЕОРЕМА 4 (обернена). *Кожній кутовій точці допустимих розв'язків системи обмежень відповідає допустимий базисний розв'язок.*

Із теореми 2 і 4 випливає наступне твердження.

НАСЛІДОК. *Якщо існує, і причому єдиний, оптимальний розв'язок задачі лінійного програмування, то він співпадає з одним із допустимих базисних розв'язків системи обмежень.*

Отже, оптимум цільової функції необхідно шукати серед скінченого числа допустимих базисних розв'язків. Але навіть у найпростіших задачах (за невеликих m і n) знаходження оптимального розв'язку шляхом розгляду всіх базисних розв'язків є трудомістким процесом, оскільки число базисних розв'язків може бути досить великим. А тому необхідна певна обчислювальна схема, яка дозволяє здійснювати перехід від одного допустимого базисного розв'язку до іншого, за яким цільова функція або наближається до оптимуму, або в крайньому випадку не змінить свого значення, що можливо, якщо інший базисний розв'язок — вироджений, або оптимальний розв'язок не єдиний. Такою обчислювальною схемою є, наприклад, *симплексний метод*, який розглянемо у параграфі §3.

§ 2. Геометричне представлення задачі лінійного програмування та її розв'язування

Системи лінійних нерівностей та область їх розв'язку. Нехай задана лінійна нерівність з двома змінними x_1 та x_2.

$$a_1x_1 + a_2x_2 + b \geq 0. \tag{1}$$

Якщо величини x_1 та x_2 вважати за координати точок площини, то сукупність точок площини, координати яких задовольняють нерівності (1) називаються *областю розв'язків* даної нерівності. Областю розв'язків нерівності (1) є напівплощина.

Щоб з'ясувати, яка із двох площин відповідає нерівності (1), достатньо привести цю нерівність до виду $x_2 \geq kx_1 + l$ або до виду $x_2 \leq kx_1 + l$. У першому випадку шукана площина лежить вище прямої $a_1x_1 + a_2x_2 + b = 0$, у другому — під нею. Якщо $a_2 = 0$, то нерівність приводиться до одного із видів $x_1 \geq h$, або $x_1 \leq h$, тобто напівплощина лежить справа або зліва від прямої $x_1 = h$.

У випадку, якщо задана система нерівностей

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + b_1 \geq 0, \\ a_{21}x_1 + a_{22}x_2 + b_2 \geq 0, \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + b_m \geq 0, \end{cases}, \tag{2}$$

де m — скінченне число, одержимо перетин скінченного числа напівплощини, що утворюють многокутну область D, яку називають *областю розв'язків* системи нерівностей (2). Така область не завжди буває обмеженою, вона може бути і необмеженою і, навіть, пустою. Останній випадок має місце тоді, коли система нерівностей (2) має протиріччя. Можливі випадки зайвих нерівностей, які входять до сумісної системи та визначають прямі, які не мають з областю D спільних точок. Такі нерівності можна виключити.

Область розв'язків має важливу властивість — вона є *випуклою*, тобто разом з будь-якими двома своїми точками містить і весь відрізок, що їх сполучає. Пряма, яка має з областю хоча б одну спільну точку, причому так, щоб вся область лежала по одну сторону від прямої, називається *опорною* щодо до цієї області.

На рис. 32а зображена *випукла* множина (випуклий многокутник), а на рис. 32б *невипуклий*, тобто існує хоча б одна пара точок множини, що відрізок, який сполучає ці точки, не належить повністю цій множині.

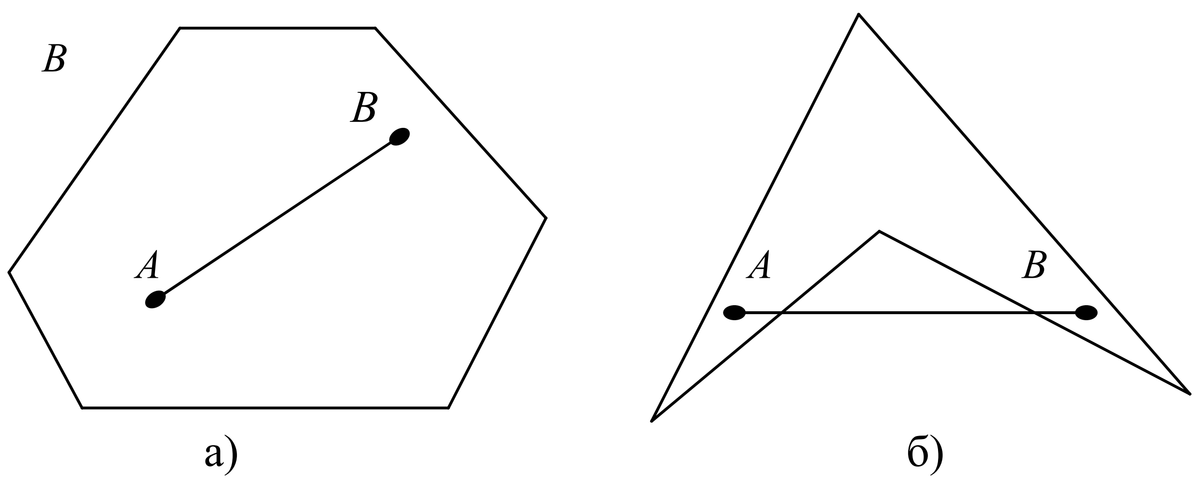


Рис. 32

Випуклі множини мають важливу властивість, яка визначається наступною теоремою.

ТЕОРЕМА. *Переріз (перетин) двох випуклих множин є також випуклою множиною.*

Користуючись методом математичної індукції, можна довести, що перетин скінченного числа випуклих множин також є випукла множина. Частинними випадками множин на площині є випуклі многокутники (трикутники, прямокутники, ромби, трапеції, правильні многокутники тощо).

Аналогічно слід тлумачити геометричний зміст і системи нерівностей з трьома змінними:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1 \ge 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2 \ge 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + b_m \ge 0. \end{cases} \tag{3}$$

Тут кожна із нерівностей виконується для одного із напівпросторів, на які розбиває весь простір відповідна площина. Система (3) є перетином напівпросторів, тобто многогранною областю розв'язків системи нерівностей.

Приклад 1. Знайти напівплощину, яка визначається рівнянням $3x_1 + 5x_2 - 15 \le 0$.

Замінимо знак нерівності на знак точної рівності, одержимо рівняння прямої $3x_1 + 5x_2 - 15 = 0$ або $x_2 = -\frac{3}{5}x_1 + 3$ (рис. 33). Проведемо дану нерівність до виду $x_2 \le -\frac{3}{5}x_1 + 3$. Отже, шукана площина розташована нижче прямої $x_2 = -\frac{3}{5}x_1 + 3$.

Приклад 2. Яка площина визначається нерівністю $3x_1 - 5x_2 \ge 0$?

Замінюючи знак нерівності на знак точної рівності, одержимо рівняння прямої $3x_1 - 5x_2 = 0$ або $x_2 = \frac{3}{5}x_1$, яка проходить через початок координат. Із нерівності $3x_1 - 5x_2 \ge 0$, тобто $x_2 \le \frac{3}{5}x_1$, випливає, що шукана півплощина розташована нижче прямої $x_2 = \frac{3}{5}x_1$ (рис. 34).

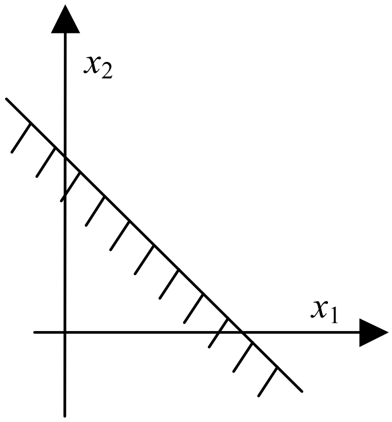


Рис. 33

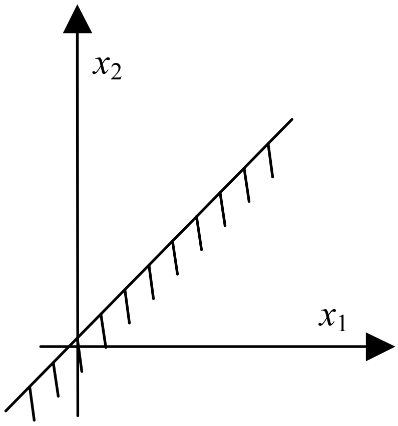


Рис. 34

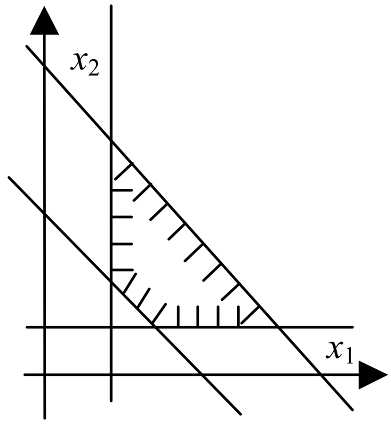


Рис. 35

Приклад 3. Знайти область розв'язків системи нерівностей:

$$\begin{cases} x_1 - 2 \geq 0 \\ x_2 - 2 \geq 0 \\ x_1 + x_2 - 6 \geq 0 \\ -7x_1 - 8x_2 + 56 \geq 0. \end{cases}$$

Замінюючи знаки нерівності на знак точних рівностей, одержимо рівняння чотирьох прямих: $x_1 - 2 = 0$; $x_2 - 2 = 0$; $x_1 + x_2 - 6 = 0$; $7x_1 + 8x_2 - 56 = 0$, зображених на рис. 35.

Приведемо дані нерівності до виду $x_1 \geq 2;\ x_2 \geq 2;\ x_2 \geq -x_1 + 6;$ $x_2 \leq -\frac{7}{8}x_1 + 7$. Штрихування показує ті із площини, що є областю розв'язку відповідних нерівностей. Отже, областю розв'язків системи нерівностей є випуклий чотирикутник.

Приклад 4. Знайти область розв'язку системи нерівностей:

$$\begin{cases} x_1 \geq 0 \\ x_1 + x_2 - 3 \geq 0 \\ x_1 - x_2 + 1 \leq 0 \\ x_1 \leq 3 \end{cases}.$$

Замінюємо знаки нерівностей на знаки точних рівностей, одержимо рівняння чотирьох прямих: $x_1 = 0;\ x_1 + x_2 - 3 = 0;\ x_1 - x_2 + 1 = 0;\ x_1 = 3$, зображених на рис. 36. Приведемо задані нерівності до виду $x_1 \geq 0;\ x_2 \geq -x_1 + 3;\ x_2 \geq x_1 + 1;\ x_1 \leq 3$. Областю розв'язків системи нерівностей є необмежена випукла фігура.

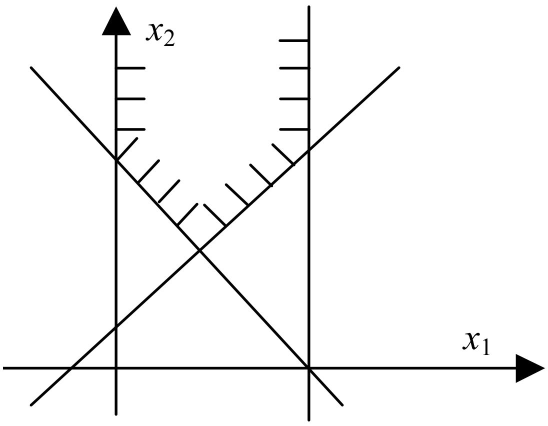


Рис. 36

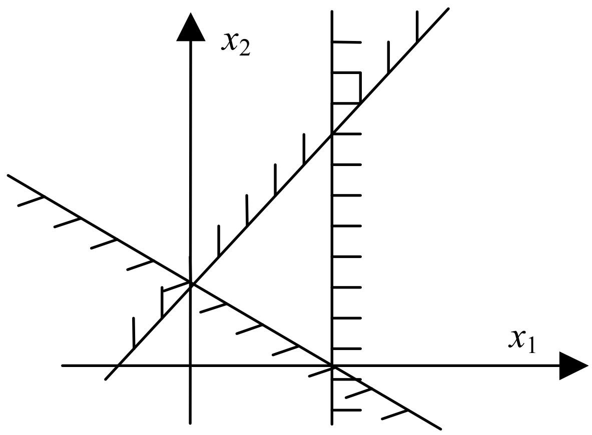


Рис. 37

Приклад 5. Знайти область розв'язків системи нерівностей:

$$\begin{cases} x_1 \ge 2 \\ x_1 + 2x_2 \le 2 \\ x_1 - 2x_2 \le -2. \end{cases}$$

Побудуємо відповідні прямі, звідки видно (рис. 37), що не існує ні однієї точки спільної для усіх трьох напівплощин. Це означає, що область розв'язків "пуста" і задана система нерівностей несумісна.

Приклад 6. Знайти область розв'язків системи нерівностей:

$$\begin{cases} 4x_1 - x_2 \ge -4 \\ x_1 - x_2 \ge -4 \\ x_1 \le 2 \\ 2x_1 - x_2 \ge 5 \\ x_2 \ge 0 \qquad . \end{cases}$$

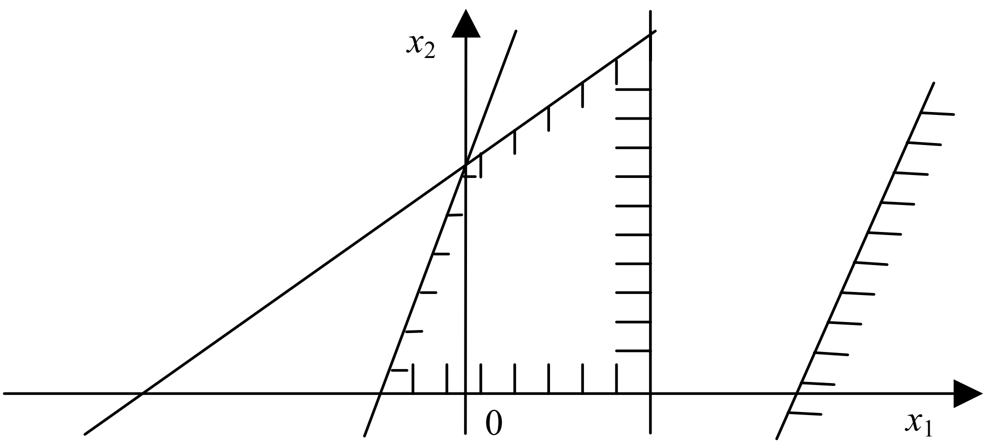


Рис. 38

Така система нерівностей не має розв'язків. Геометрично це означає, що не існує ні однієї точки координати, якої задовольняють усім нерівностям даної задачі (рис. 38).

Приклад 7. Знайти область розв'язків системи нерівностей:

$$\begin{cases} 4x_1 - x_2 \ge 0 \\ x_1 - 2x_2 \le 0 \\ 3x_1 + 4x_2 \le 15 \\ x_1 \le 3 \\ 3x_1 - x_2 \ge -6 \ . \end{cases}$$

П'яти заданим нерівностям відповідає множина точок площини, які утворюють трикутник *AOB* (рис. 39). Останні дві нерівності можуть бути виключені, оскільки остання нерівність визначає граничну пряму, яка не має з трикутником *AOB* спільних точок, а пряма, що визначається нерівністю $x_1 \le 3$, має спільну точку з трикутником і є опорною.

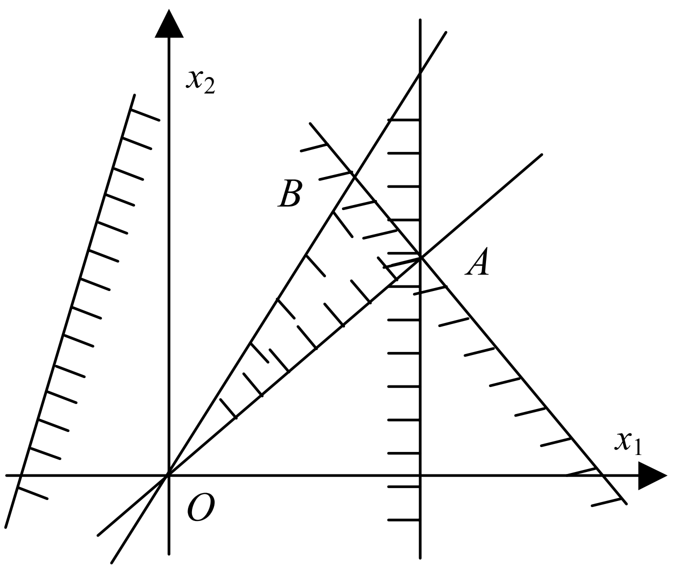


Рис. 39

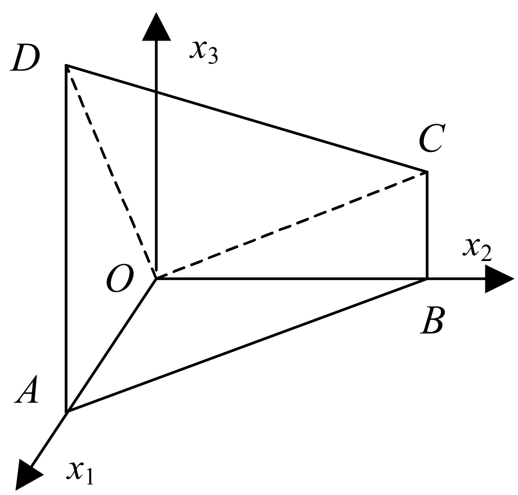


Рис. 40

Приклад 8. Знайти область розв'язків системи нерівностей $x_1 \ge 0$, $x_2 \ge 0$, $x_3 \ge 0$, $x_1 + x_2 - 2 \le 0$, $4x_1 + x_2 - 4x_3 \ge 0$.

Замінюючи знаки нерівностей на знаки точних рівностей, одержимо рівняння площини $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, $x_1 + x_2 - 2 = 0$, $4x_1 + x_2 - 4x_3 = 0$, які зображені на рис. 40. Областю розв'язків системи нерівностей буде випуклий чотирикутник *ABOCD*.

Геометричний розв'язок задач лінійного програмування. Проілюструємо, як відшукують оптимум цільової функції, використовуючи геометричне представлення системи обмежень і лінійної форми (цільової функції), тим самим мова буде йтися про геометричний метод розв'язування задач лінійного програмування.

Зауважимо, що геометричний метод має вузькі рамки застосування і говорити про нього як про особливий метод розв'язування задач лінійного програмування не можна, оскільки наглядність геометричного розв'язку досягається лише на площині та просторі і знайомство з геометричним представленням задачі відбувається лише в двовимірному та трьохвимірному просторі.

Покажемо, як розв'язується зазначена задача геометричним методом. Для цього обмежимося розглядом сумісної системи лінійних нерівностей з двома і трьома змінними. Нехай, крім цього, задана лінійна функція $F = c_1x_1 + c_2x_2 + c_0$. Знайдемо серед множини точок $(x_1; y_1)$ із області розв'язків сумісної системи нерівностей такі, які задають заданій лінійній функції найменше (найбільше) значення. Для кожної точки площини функція F приймає фіксоване значення $F = F_1$. Множина всіх таких точок є пряма $c_1x_1 + c_2x_2 + c_0 = F$, перпендикулярна вектору $\overline{n} = \{c_1; c_2\}$, який виходить із початку координат. Якщо таку пряму переміщувати паралельно самій собі в додатному напрямку вектора $\overline{n}$, то лінійна функція $F = c_1x_1 + c_2x_2 + c_0$ буде зростати, а в протилежному напрямку — спадати. Нехай при переміщенні прямої F в додатному напрямку вектора $\overline{n}$ вона спочатку зустрінеться з многокутником розв'язків у його вершині, тоді в цьому положенні F_1 пряма F стане *опорною*, і на цій прямій функція F приймає найменше значення.

При подальшому русі в тому ж напрямку (додатному) пряма F пройде через другу вершину многокутника розв'язків, виходячи із області розв'язків і стане також опорною прямою F_2, на якій F приймає найбільше значення серед усіх значень.

Отже, мінімізація та максимізація лінійної функції $F = c_1x_1 + c_2x_2 + c_0$ на многокутнику розв'язків досягається в точках перетину цього многокутника з опорними прямими, перпендикулярними вектору $\overline{n} = \{c_1; c_2\}$. Опорна пряма може мати з многокутником розв'язків або одну спільну точку (вершину многокутника), або нескінченну множину точок (ця множина є стороною многокутника).

Аналогічно лінійна функція трьох змінних $F = c_1x_1 + c_2x_2 + c_0$ приймає стале значення із площини, перпендикулярної вектору $\overline{n} = \{c_1; c_2; c_3\}$. Найменше і найбільше значення цієї функції на многограннику розв’язків досягається в точках перетину цього многогранника з опорними площинами, перпендикулярними вектору $\overline{n} = \{c_1; c_2; c_3\}$. Опорна площина може мати з многогранником розв’язків або одну спільну точку (вершина многогранника), або нескінченну множину точок (це є ребро або грань многогранника).

Разом з тим для задач лінійного програмування геометричний метод представляє певний інтерес, тим більше, що він дозволяє геометрично підтвердити справедливість основних теорем лінійного програмування.

Приклад 9. Знайти максимум функції $F = 2x_1 - x_2$ при обмеженнях:

$$\begin{cases} -2x_1 + x_2 \le 3 \\ 2x_1 - 3x_2 \le 6 \\ x_1 + x_2 \le 8 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases} \qquad (4)$$

Переконатися в справедливості теорем 1—4 § 1.

На рис. 41 зображена область допустимих розв’язків даної системи нерівностей. Це випуклий многокутник *OABCD*, який має п’ять кутових точок $O(0;0)$, $A(0;3)$, $B\left(\frac{5}{3};\frac{19}{3}\right)$, $D(3;0)$, $C(6;2)$.

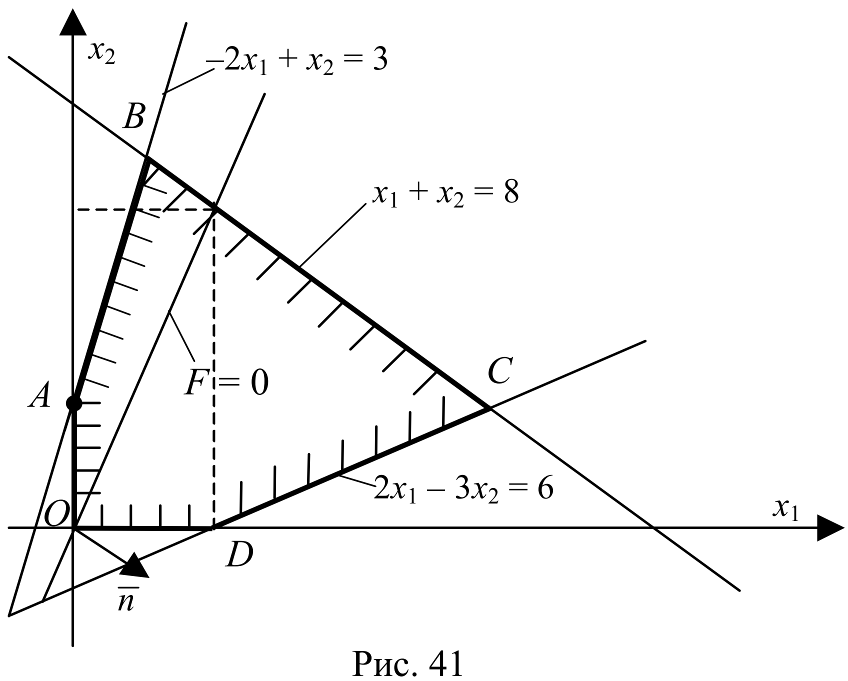


Рис. 41

Очевидно, що множина допустимих розв'язків в даному прикладі випукла, тобто справедлива теорема 1 §1. Потрібно знайти таку точку цього многокутника, яка б максимізувала цільову функцію $F = 2x_1 - x_2$, яка є лінійною функцією координат точок на площині. Розглянемо таке питання: як розташовуються на площині точки, в яких функція F приймає одне і те саме значення?

Для цього достатньо функцію F прирівняти до будь-якої сталої величини, тобто $F = const = a$. Це приводить до рівняння $2x_1 - x_2 = a$, яке є рівнянням прямої на площині. Отже, пряма $2x_1 - x_2 = a$ є множиною точок, в яких функція F приймає значення, рівне a. Змінюючи величину a, одержимо сімейство паралельних прямих. Кожну із прямих цього сімейства прийнято називати *лінією рівня.*

На рис.41 побудована лінія рівня $2x_1 - x_2 = 0$, яка відповідає значенню $F = 0$. При переході від однієї лінії рівня до іншої значення функції F змінюється. Якщо вихідну лінію рівня пересувати вправо, то значення F при цьому зростає. Необхідний напрямок руху вихідної лінії рівня можливо встановити, виходячи з того, що коефіцієнти при змінних у рівнянні прямої, є координатами вектора $\overline{n}$, перпендикулярного прямій. У даному випадку $\overline{n} = \{2;-1\}$. Із рис.41 видно, що значення функції F зростає при переміщенні вихідної лінії рівня в напрямку вектора $\overline{n}$. Якщо потрібно було б знайти мінімум функції F, то вихідну лінію рівня необхідно було б пересувати в сторону, протилежну вектору $\overline{n}$.

Переміщуючи вихідну лінію рівня $2x_1 - x_2 = 0$ у напрямку вектора $\overline{n}$, легко переконатися, що максимальне значення цільової функції на многокутнику розв'язків буде досягнуто в кутовій точці C, в якій лінія рівня при подальшому переміщенні вийде із цього многокутника. Для знаходження координат точок С необхідно розв'язати систему рівнянь

$$\begin{cases} 2x_1 - 3x_2 = 6 \\ x_1 + x_2 = 8 \end{cases},$$

звідки одержимо $C(6;2)$.

Підставимо координати точки C у вираз F, знайдемо максимальне значення функції: $F_{\max} = 2 \cdot 6 - 2 = 10$.

Таким чином, встановлено, що цільова функція досягла максимального значення в кутовій точці многокутника розв'язків, тобто геометрично підтвердили теорему 2 § 1.

Встановимо, що теореми 3, 4 § 1 також виконуються. Для цього систему обмежень (4) шляхом введення додаткових невід'ємних змінних x_3, x_4, x_5 зведемо до системи рівнянь

$$\begin{cases} -2x_1 + x_2 + x_3 = 3 \\ 2x_1 - 3x_2 + x_4 = 6 \\ x_1 + x_2 + x_5 = 8 \end{cases}. \tag{5}$$

Така система трьох рівнянь з п'ятьма змінними може мати не більше десяти базисних розв'язків ($C_5^3 = 10$). Знайдемо деякі із них.

Візьмемо як основні змінні x_3, x_4, x_5. Це можливо зробити оскільки коефіцієнти при них утворюють одиничну матрицю, визначник якої відмінний від нуля. Якщо прирівняти неосновні змінні x_1 і x_2 до нуля, одержимо базисний розв'язок (0; 0; 3; 6; 8).

Перші дві компоненти в цьому розв'язку є координатами x_1 і x_2 точки на площині, при цьому $x_1 = x_2 = 0$, що відповідає початку координат, тобто кутовій точці O. Можливо, наприклад, за основні змінні вибрати x_1, x_4 і x_5, оскільки визначник, складений із коефіцієнтів при них не дорівнює нулю, дійсно

$$D = \begin{vmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = 1 \neq 0 .$$

Змінні x_1 і x_3 тоді є неосновними, а тому вважатимемо $x_1 = x_3 = 0$, тоді з першого рівняння системи (5) знайдемо $x_2 = 3$, з другого — $x_4 = 6 + 3 \cdot 3 = 15$, з третього — $x_5 = 8 - 3 = 5$. Новому базисному розв'язку (0; 3; 0; 15; 5) відповідає точка $A(0; 3)$ многокутника $OABCD$.

Якщо за основні змінні вибрати x_1, x_2 та x_3, а це можливо, оскільки визначник складений із коефіцієнтів при цих змінних

$$D = \begin{vmatrix} -2 & 1 & 1 \\ 2 & -3 & 0 \\ 1 & 1 & 0 \end{vmatrix} = 5 \neq 0$$

відмінний від нуля, то базисний розв'язок з урахуванням $x_4 = x_5 = 0$ має вигляд (6; 2; 13; 0; 0), що відповідає кутовій точці $C(6; 2)$.

Якщо основні змінні є x_1, x_2, x_4, а $x_3 = x_5 = 0$, то аналогічно базисний розв'язок $\left(\frac{5}{3}; \frac{19}{3}; 0; \frac{65}{3}; 0\right)$ відповідає кутовій точці B.

Нарешті легко переконатися, що базисний розв’язок (3; 0; 9; 0; 5) відповідає точці D. Інші п’ять базисних розв’язків є недопустимими.

Зауважимо, що одержані висновки базуються на тому, що множина розв’язків задач лінійного програмування є замкнутим многокутником, система обмежень сумісна і лінійно незалежна (немає зайвих обмежень) і оптимальний розв’язок єдиний.

Розглянемо приклади, коли зазначені умови не виконуються.

Приклад 10. Знайти максимум функції $F = x_1 + 3x_2$ при обмеженнях:

$$\begin{cases} -3x_1 + 4x_2 \le 12 \\ 2x_1 - 3x_2 \le 6 \\ x_1 \ge 0 \\ x_2 \ge 0. \end{cases}$$

На рис. 42 зображено: необмежена многогранна область розв’язків заданої системи обмежень-нерівностей; вихідна лінія рівня $x_1 + 3x_2 = 9$; вектор $\overline{n} = \{1;3\}$, який вказує напрямок руху вихідної лінії рівня з метою знаходження $F_{\max}$.

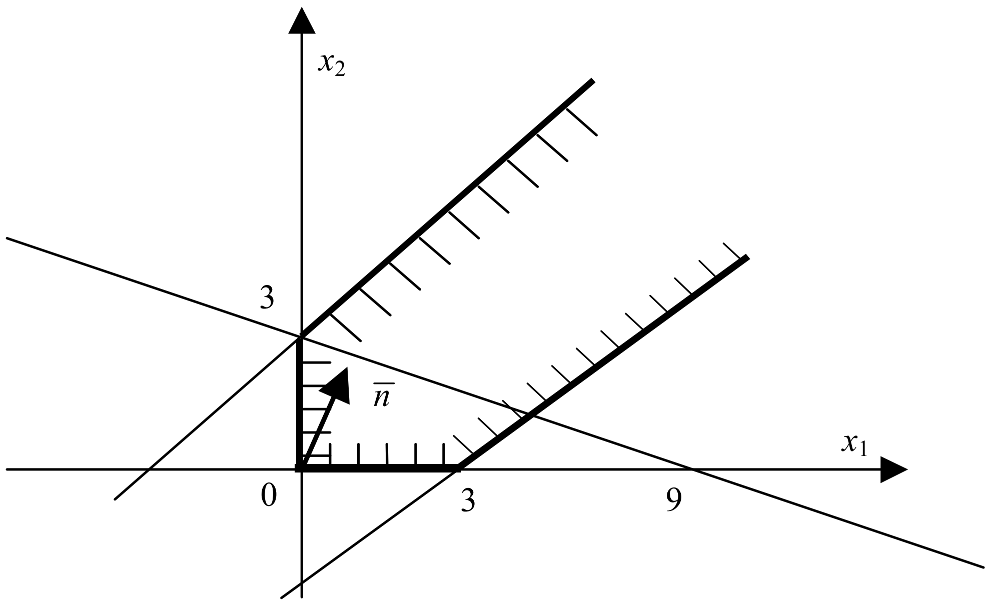


Рис. 42

Легко переконатися, що функція F може необмежено зростати при заданій системі обмежень, а тому умовно можна вважати, що $F_{\max} = \infty$.

Приклад 11. Знайти максимум функції $F = 2x_1 - x_2$ при обмеженнях:

$$\begin{cases} -2x_1 + x_2 \ge 3 \\ 2x_1 - 3x_2 \ge 6 \\ x_1 + x_2 \le 8 \\ x_1 \ge 0 \\ x_2 \ge 0. \end{cases}$$

Зображена на рис. 43 область не містить ні однієї загальної точки, яка б задовольняла всім нерівностям системи обмежень.

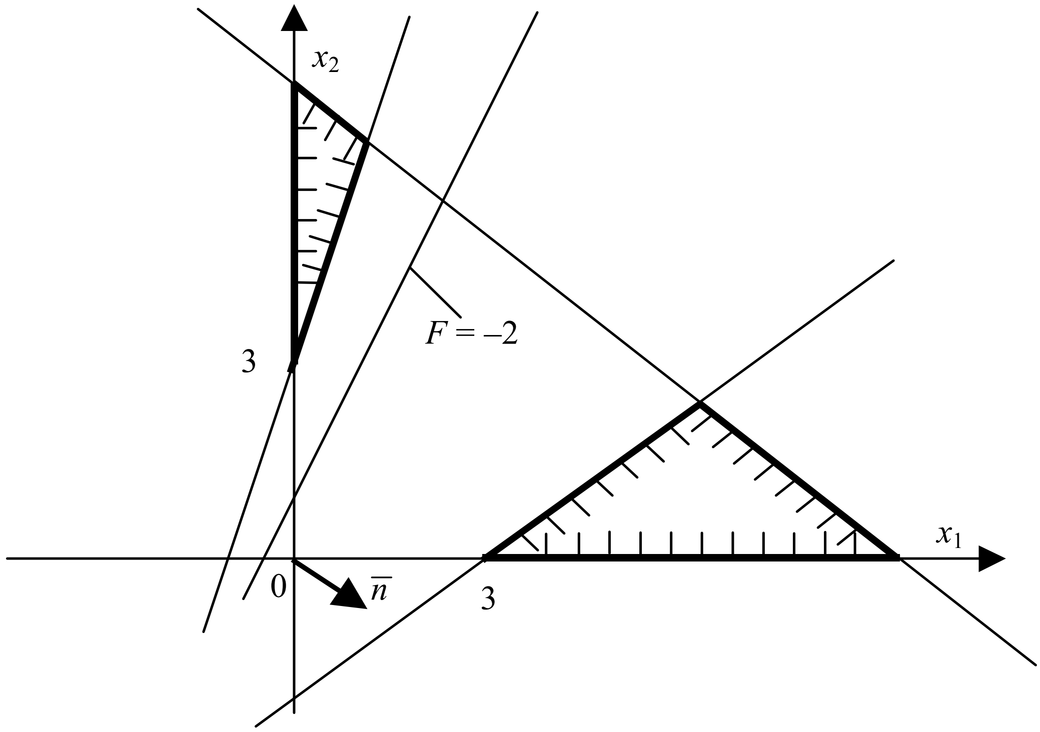


Рис. 43

Це вказує на те, що система обмежень містить протиріччя і не може мати ні одного розв’язку, в тому числі і оптимального.

Приклад 11. Знайти максимум функції $F = 2x_1 - x_2$ при обмеженнях:

$$\begin{cases} -2x_1 + x_2 \le 3 \\ 2x_1 - 3x_2 \ge 6 \\ x_1 + x_2 \le 8 \\ x_1 \ge 0 \\ x_2 \ge 0. \end{cases}$$

Областю розв'язків є точки трикутника *ABC*, які задовольняють усім нерівностям системи обмежень. За вихідну лінію рівня вибрали пряму $2x_1 - x_2 = 8$, з тим, щоб вона перетнула область розв'язків. Як видно з рис. 44, максимальне значення $F = 8$ досягається в точці $B(8;0)$. Зауважимо, що при побудові трикутника *ABC* ми не використовували пряму $-2x_1 + x_2 = 3$, яка відповідає першій нерівності, хоча всі точки цього трикутника задовольняють цій нерівності. Отже, нерівність $-2x_1 + x_2 \le 3$ зайва в системі обмежень (як і нерівність $x_1 \ge 0$).

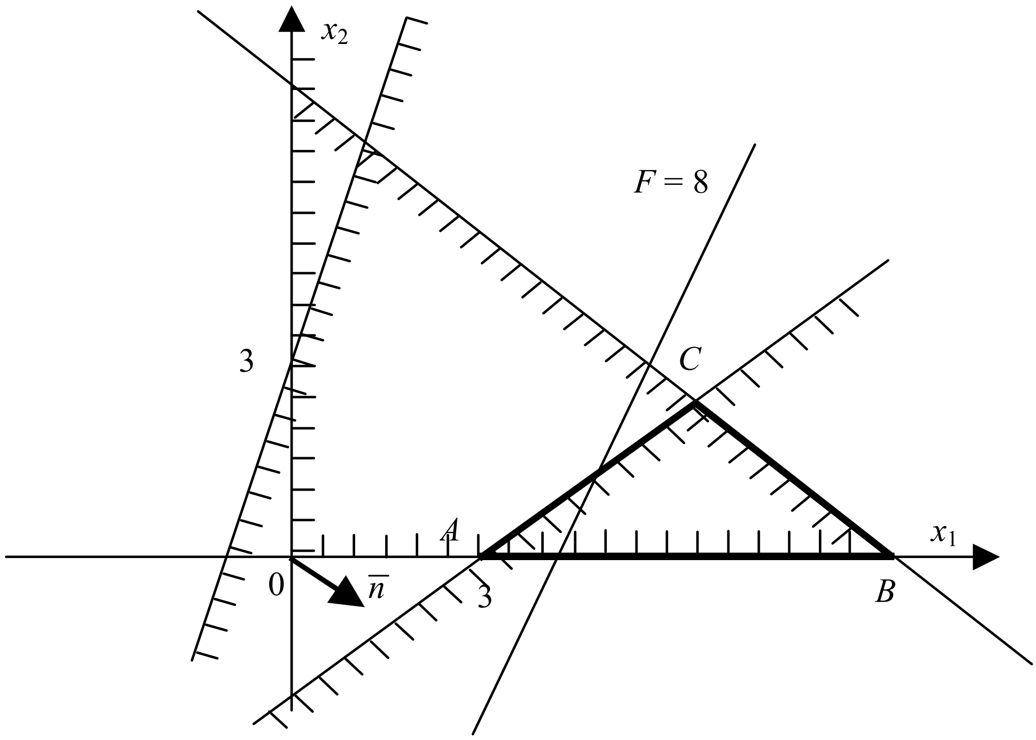


Рис. 44

Крім того, слід зазначити, що зайві нерівності в системі обмежень можливо встановити лише після побудови області розв'язків, що можливо лише для дуже обмеженого типу задач лінійного програмування. При розв'язуванні задач з будь-яким числом змінних аналітичним методом (наприклад, симплексним) питання про зайві обмеження просто не розглядаються, тобто всі нерівності системи обмежень враховуються.

Приклад 13. Знайти максимум функції $F = 2x_1 - x_2$ при обмеженнях:

$$\begin{cases} -2x_1 + x_2 \le 3 \\ 2x_1 - x_2 \le 6 \\ x_1 + x_2 \le 8 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}.$$

На рис.45 зображена область розв'язків системи обмежень і лінія рівня, що відповідає значенню $F = -2$. Якщо переміщати лінію рівня паралельно вихідній у напрямку вектора $\overline{n}$, то вона вийде із області розв'язків не в одній точці, як це було в попередніх прикладах, а зіллється з прямою CD, яка є граничною лінією області розв'язків. Усі точки відрізку CD дають одне і теж значення функції F, яке і є оптимальним значенням: $F_{\max} = 6$. Отже, існує не один, а нескінченна множина оптимальних розв'язків, що співпадають з точками відрізку CD, в окремому випадку з двома кутовими точками C і D.

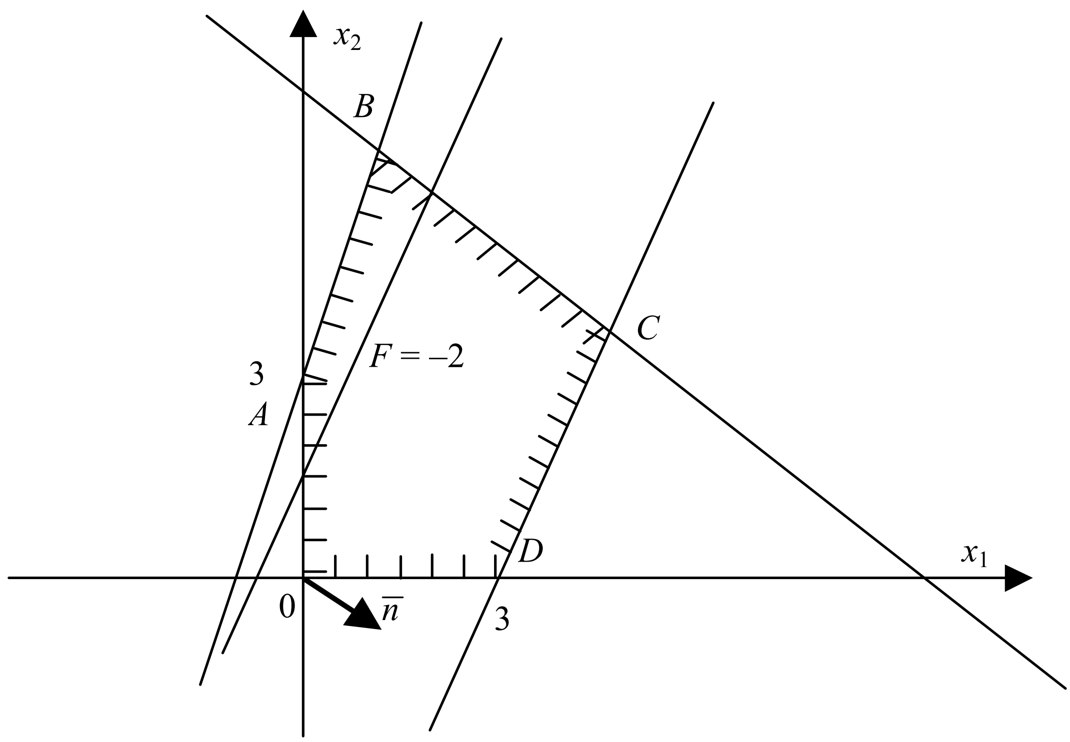


Рис.45

Приклад 14. Знайти мінімум функції $F = 2x_1 + x_2$ при обмеженнях

$$\begin{cases} -2x_1 + x_2 \le 3 \\ 2x_1 + 3x_2 \ge 16 \\ x_1 + x_2 \le 8 \\ x_1 \ge 0;\ x_2 \ge 0 \end{cases}$$

Областю розв'язків даної системи обмежень є трикутник ABC (рис. 46).

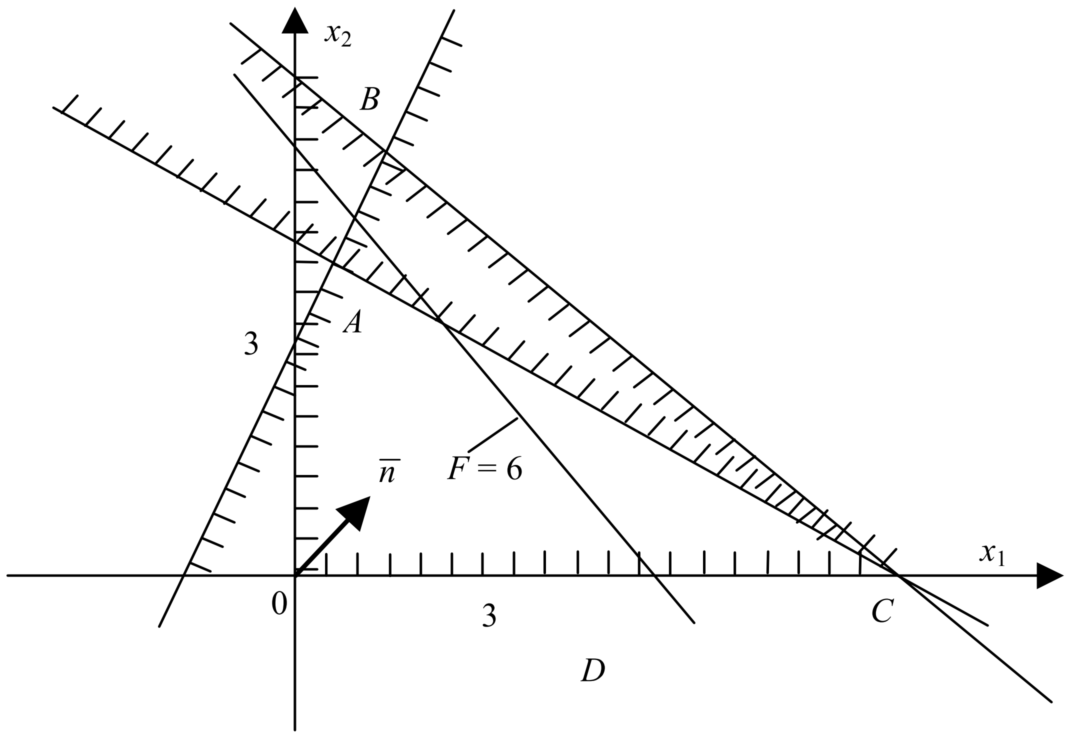


Рис. 46

На рисунку зображена також вихідна лінія рівня $2x_1 + x_2 = 6$ і вектор $\overline{n} = \{2;1\}$, який показує напрямок руху цієї лінії для досягнення максимуму функції F. Оскільки потрібно знайти мінімум цієї функції, то будемо переміщати вихідну лінію рівня в сторону, протилежну вектору $\overline{n}$. Як видно з рис. 46, мінімум функції F досягається в кутовій точці A, координати якої є розв'язком системи рівнянь:

$$\begin{cases} -2x_1 + x_2 = 3 \\ 2x_1 + 3x_2 = 16 \end{cases}.$$

Звідси $A\left(\frac{7}{8};\frac{19}{4}\right)$ і $F_{\min}=2\cdot\frac{7}{8}+\frac{19}{4}=\frac{13}{2}$.

ЗАДАЧІ

4.1. Як розташована напівплощина, координати точок якої задовольняють нерівності $x_1-x_2-5\geq 0$?

Знайти область розв'язку системи нерівностей та координати їх кутових точок:

4.2. $x_1+2x_2-8\geq 0,\ x_1-2x_2\leq 0,\ 3x_1+2x_2-32\leq 0,\ x_2\leq 7$.

4.3. $x_1+x_2\leq 6,\ 2x_1-x_2\leq 4,\ x_1+2x_2\geq 4,\ x_1\leq 3,\ x_2\leq 4$.

4.4. $5x_1-3x_2+15\geq 0,\ 0\leq x_2\leq 10,\ x_1+x_2-17\leq 0,\ 0\leq x_1\leq 11$.

4.5. $x_1\geq 3,\ x_2\geq 0,\ x_1+x_2\leq 0$.

4.6. $x_1-x_2+1\geq 0,\ 2x_1+x_2-7\geq 0,\ x_1-2x_2+4\geq 0$.

4.7. $x_2\geq 0,\ 4x_1-x_2\geq 0,\ x_2\leq 6,\ 4x_1+x_2\leq 40,\ x_1-x_2+8\geq 0$.

4.8. $5x_1-2x_2-10\leq 0,\ 3x_1+2x_2-10\geq 0,\ x_1+2x_2-6\geq 0,\ x_1\geq 0$.

4.9. $2x_1+x_2-2\geq 0,\ 3x_1-4x_2+12\geq 0,\ x_1\geq 0,\ x_2\geq 0$.

4.10. $3x_1+2x_2\geq 16,\ x_1+2x_2\geq 8,\ x_1\geq 1,\ x_2\geq 2$.

4.11. $x_1+2x_2\geq 12,\ 2x_1+x_2\leq 6,\ x_1\geq 2,\ x_2\leq 2$.

4.12. $x_1\geq 0,\ x_2\geq 1, x_3\geq 0,\ x_1+x_2+x_3-5\leq 0$.

Знайти максимум функції F за заданими обмеженнями:

4.13. $\begin{cases}3x_1-2x_2\geq -6,\\ 3x_1+x_2\geq 3,\\ x_1\leq 3,\end{cases}$ $F=2x_1+2x_2$.

4.14. $\begin{cases}x_1+x_2\geq 1,\\ -2x_1+x_2\leq 2,\\ x_1+x_2\leq 4,\\ x_1\leq 3,\\ x_1\geq 0, x_2\geq 0,\end{cases}$ $F=x_1+2x_2$.

4.15. $\begin{cases}3x_1+4x_2\geq 18,\\ 3x_1-x_2\geq 3,\\ x_2\leq 6,\\ 2x_1+x_2\leq 18,\\ 4x_1-x_2\leq 24,\end{cases}$ $F=8x_1-2x_2$.

4.16. $\begin{cases}x_1-2x_2\leq 4,\\ -x_1+2x_2\leq 4,\\ x_1+2x_2\leq 6,\\ x_1\geq 0, x_2\geq 0,\end{cases}$ $F=x_1+2x_2$.

4.17. $\begin{cases} 3x_1 + 2x_2 \ge 6, \\ 3x_1 - 2x_2 \ge -7, \\ 2x_1 - 4x_2 \le 8, \\ x_1 \ge 1, x_2 \ge 0, \end{cases}$.
$F = 3x_1 + 4x_2$

4.18. $\begin{cases} x_1 - x_2 \le -4, \\ x_1 + x_2 \le 2, \\ x_1 \le 1, \\ x_1 \ge 0, x_2 \ge 0, \end{cases}$.
$F = x_1 + 7x_2$

Знайти мінімум функції F за заданими обмеженнями:

4.19. $\begin{cases} 3 \le x_1 + x_2 \le 7, \\ 1 \le x_2 \le 4, \\ x_1 \le 4, \end{cases}$.
$F = x_1 - x_2$

4.20. $\begin{cases} x_1 - 2x_2 \le 2, \\ x_1 + 2x_2 \ge 3, \\ x_1 + 2x_2 \le 6, \\ x_1 - x_2 \ge 0, \\ x_1 \ge 0, x_2 \ge 0, \end{cases}$.
$F = x_1 - 3x_2$

4.21. $\begin{cases} x_1 - 2x_2 \ge 4, \\ x_1 + x_2 \ge 3, \\ x_1 + x_2 \le 7, \\ x_1 - 2 \le 0, \\ x_1 \ge 0, x_2 \ge 0, \end{cases}$.
$F = 2x_1 - x_2$

ВІДПОВІДІ

4.1. Нижче прямої $x_1 - x_2 - 5 = 0$.

4.2. Чотирикутник *ABCD* (рис. 47).

4.3. П’ятикутник *ABCDE*: *A*(-4;4), *B*(2;4), *C*(3;3), *D*(3;2), *E*(2,4;0,8).

4.4. Шестикутник *ABCDEO*: $A(0;5)$, $B(3;10)$, $C(7;10)$; $D(11;6)$, $E(11;0)$, $O(0;0)$.

4.5. Пуста область.

4.6. Точка (2;3).

4.7. Областю розв’язків є трапеція, останню нерівність можна виключити.

4.8. Необмежена область *ABCDE* (рис. 48).

4.9. Необмежена область *ABCDE* (рис. 49).

4.10. Необмежена область *ABCDE* (рис. 50).

4.11. Система несумісна.

4.12. Трикутна піраміда.

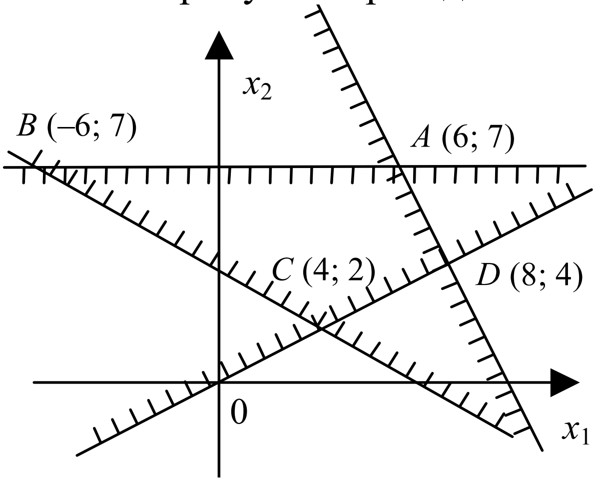


Рис. 47

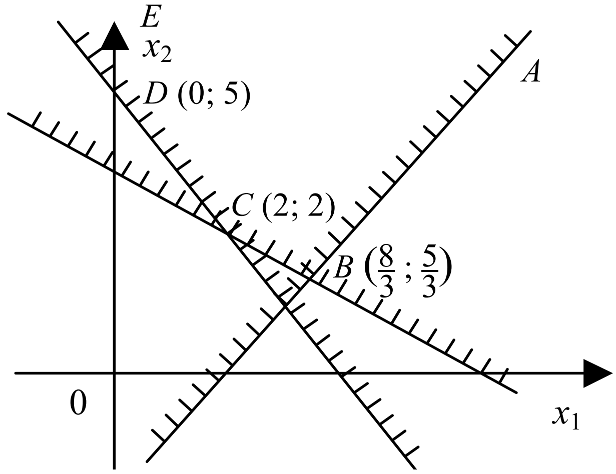


Рис. 48

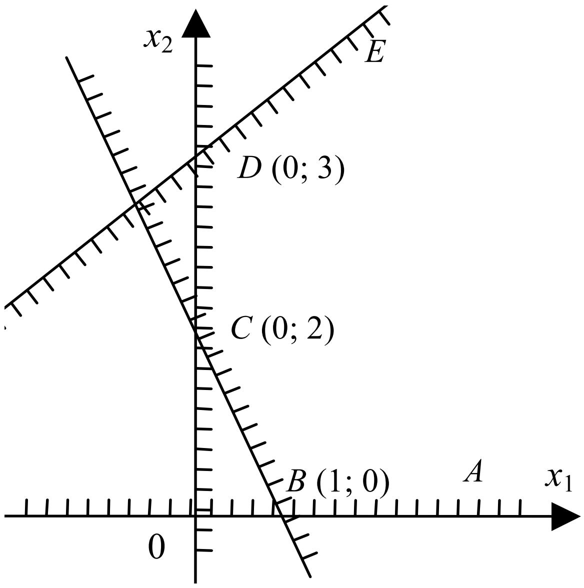


Рис. 49

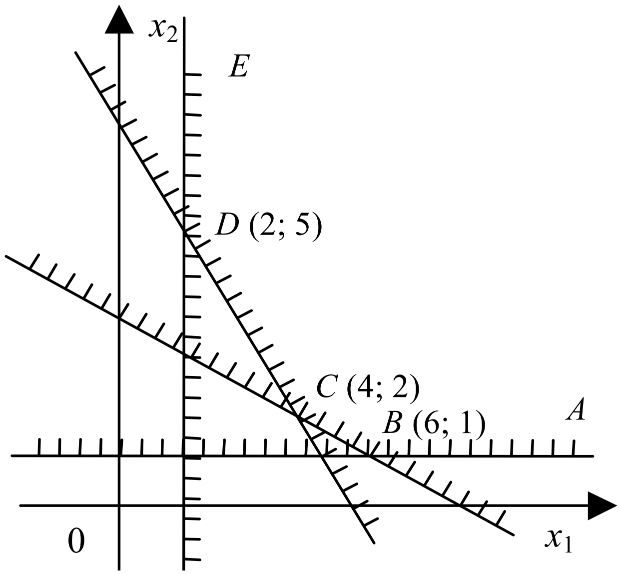


Рис. 50

4.13. $F_{\max} = 21$. **4.14.** $F_{\max} = \dfrac{22}{3}$.

4.15. $F_{\max} = 48$ в кожній точці відрізку AB, де $A(6;0)$, $B(7;4)$.

4.16. $F_{\max} = 6$ (оптимальний розв’язок не єдиний).

4.17. $F_{\max} = \infty$.

4.18. Розв’язку немає.

4.19. $F_{\min} = -4$ при $x_1 = 0$, $x_2 = 4$.

4.20. $F_{\min} = -4$.

4.21. Розв’язків немає.

§ 3. Симплексний метод

Розв'язання основної задачі лінійного програмування геометричним методом є наочним у випадку двох і навіть трьох змінних. Для випадку ж більшого числа змінних геометричний метод стає неможливим. Так званий *симплекс-метод* належить до числа аналітичних і є універсальним методом, яким можливо розв'язати будь-яку задачу лінійного програмування.

Ідея симплексного методу полягає у наступному. Використовуючи систему обмежень, приведену до загального вигляду, тобто до системи m лінійних рівнянь з n змінними ($m < n$), знаходять її будь-який базисний розв'язок, за можливістю найбільш простий. Якщо перший базисний розв'язок буде допустимим, то перевіряють його на оптимальність. Якщо він є неоптимальним, то переходять до іншого допустимого базисного розв'язку. Симплексний метод гарантує, що при цьому новому розв'язку цільова функція якщо і не досягне оптимуму, то наблизиться до нього. Продовжують таку дію, доти, доки не знайдуть розв'язок, який є оптимальним.

Якщо перший знайдений базисний розв'язок виявиться недопустимим, то за допомогою симплексного методу здійснюється перехід до інших базисних розв'язків, які дозволяють наблизитися до області допустимих розв'язків, доти, доки на якомусь кроці не одержимо допустимий базисний розв'язок. Після цього до нього застосовують механізм симплексного методу.

Таким чином, застосування симплексного методу розпадається на два етапи: 1) знаходження допустимого базисного розв'язку системи обмежень; 2) знаходження оптимального розв'язку.

При цьому кожний етап може містити кілька кроків, відповідних тому або іншому базисному розв'язку. Оскільки число базисних розв'язків завжди обмежено, то обмежено і число кроків симплексного методу.

Ідею методу прослідкуємо на конкретних прикладах.

Задача про використання сировини. Це задача є частинним випадком задачі про використання ресурсів, що сформульована в § 1.

Задача 1. Для виготовлення двох видів продукції (A і B) підприємство застосовує деревину чотирьох видів. Запаси деревини кожного виду обмежені і складають відповідно 100, 200, 100, 80 одиниць. Кількість одиниць деревини кожного виду, необхідна для виготовлення одного виробу *A* і одного виробу *B*, а також прибуток, отриманий підприємством від реалізації одиниці продукції, дані в таблиці 4.2.

Потрібно скласти такий план випуску продукції, який забезпечив би підприємству найбільший прибуток від реалізації всієї продукції.

Математичне формулювання задачі буде таким. Нехай x_1 і x_2 — відповідна кількість виробів *A* і *B*, запланованих до виробництва.

Таблиця 4.2

Вид деревини	**Запаси деревини**	**Кількість одиниць сировини, необхідної для виробництва одиниці продукції**	
		A	B
I	100	0	5
II	200	2	0
III	100	4	4
IV	80	2	2
Прибуток		4	5

Оскільки кількість деревини кожного виду обмежена, то мають виконуватися такі нерівності:

$$\begin{cases} 0 \cdot x_1 + 5x_2 \le 100 \\ 2x_1 + 0 \cdot x_2 \le 200 \\ 4x_1 + 4x_2 \le 100 \\ 2x_1 + 2x_2 \le 80 \end{cases}.$$

Така система нерівностей і є системою обмежень даної задачі. Цільова функція, яка виражає прибуток підприємства, має вигляд $F = 4x_1 + 5x_2$. Таким чином, задача зводиться до знаходження максимуму функції $F = 4x_1 + 5x_2$ при обмеженнях:

$$\begin{cases} 5x_2 \le 100 \\ 2x_1 \le 200 \\ 4x_1 + 4x_2 \le 100 \\ 2x_1 + 2x_2 \le 80 \end{cases}$$

$$x_1 \ge 0 \;,\; x_2 \ge 0 .$$

Додавання до лівої частини кожної нерівності додаткових невід'ємних змінних x_3, x_4, x_5, x_6 приведе систему обмежень-нерівностей до системи рівнянь.

$$\begin{cases} 5x_2 + x_3 = 100 \\ 2x_1 + x_4 = 200 \\ 4x_1 + 4x_2 + x_5 = 100 \\ 2x_1 + 2x_2 + x_6 = 80 \end{cases} \qquad (1)$$

$$x_j \geq 0 \ \left(j = 1, 2, ..., 6\right) .$$

В умовах цієї задачі додаткові змінні мають конкретний економічний зміст, а саме: виражають обсяг залишків деревини кожного виду після виконання плану випуску продукції.

Отже, необхідно знайти такі допустимі базисні розв'язки цієї системи обмежень, які б максимізували цільову функцію $F = 4x_1 + 5x_2$.

Оскільки система обмежень є системою чотирьох незалежних рівнянь із шістьма змінними, то число основних змінних повинно рівнятися чотирьом, а неосновних — двом. Для розв'язування задачі симплексним методом перш за все необхідно знайти будь-який базисний розв'язок. У даному випадку для цього достатньо вибрати за основні додаткові змінні x_3, x_4, x_5, x_6. Оскільки коефіцієнти при цих змінних утворюють одиничну матрицю, то відпадає необхідність обчислювати визначник. Ураховуючи неосновні змінні x_1 і x_2 рівними нулю, отримаємо базисний розв'язок (0; 0; 100; 200; 100; 80), який до того ж виявився допустимим (всі змінні невід'ємні).

А тому переходимо зразу до другого етапу, тобто пошуку оптимального розв'язку, оскільки відпадає необхідність у продовженні застосування першого етапу симплекс-методу.

I крок. *Основні змінні:* x_3, x_4, x_5, x_6; *неосновні:* x_1, x_2.

Виразимо основні змінні через неосновні. З урахуванням системи (1), маємо:

$$\begin{cases} \boxed{x_3 = 100 - 5x_2} \\ x_4 = 200 - 2x_1 \\ x_5 = 100 - 4x_1 - 4x_2 \\ x_6 = 80 - 2x_1 - 2x_2 \end{cases} \qquad (2)$$

$$F = 4x_1 + 5x_2 \quad .$$

Оскільки $x_1 = x_2 = 0$, маємо $x_3 = 100$; $x_4 = 200$; $x_5 = 100$; $x_6 = 80$, що дає базисний розв'язок (0; 0; 100; 200; 100; 80), який ми прийняли за вихідний. При цьому базисному розв'язку значення цільової функції $F = 4x_1 + 5x_2 = 0$. Вважаючи, що $x_1 = x_2 = 0$ (підприємство нічого не випускає), перша мета: знаходження будь-якого базисного розв'язку, досягнута. Тепер від отриманого першого початкового розв'язку необхідно перейти до іншого, за яким значення цільової функції збільшиться, що відбудеться зі збільшенням змінних x_1 і x_2. Іншими словами ці змінні невірно вважати неосновними, тобто рівними нулю, а тому їх необхідно перевести в число основних, що і буде означати переходом до нового базисного розв'язку.

Симплекс метод на кожному кроці передбачає перехід у число основних лише однієї із вільних змінних. Переведемо до числа основних змінну x_2, оскільки вона входить до виразу цільової функції з більшим коефіцієнтом. Як тільки одна із вільних змінних переходить до числа основних, одна із основних повинна бути переведена на її місце до числа неосновних. На запитання, яку із чотирьох основних змінних необхідно вивести, допоможуть відповісти наступні судження.

Значення x_2 необхідно зробити за можливістю великим, що відповідає кінцевій меті — максимізації цільової функції F, хоча збільшення x_2 може продовжуватися до певних меж, щоб не порушувати вимоги невід'ємності змінних. Наприклад, з першого рівняння системи (2) випливає, що змінна x_2 не повинна перевищувати числа $100/5$, тобто $x_2 \le 20$, оскільки при цих значеннях x_2 змінна x_3 залишається додатною (якщо $x_2 = 20$, то $x_3 = 0$; якщо $x_2 > 20$, то $x_3 < 0$). Аналогічно із третього рівняння системи (2) випливає, що $x_2 \le 100/4$, тобто $x_2 \le 25$, із четвертого — що $x_2 \le 80/2$, тобто $x_2 \le 40$ (у друге рівняння змінна x_2 не входить). Усім цим вимогам задовольняє $x_2 \le 20$, оскільки для відповіді на питання, яку змінну необхідно перевести до числа неосновних, необхідно прийняти $x_2 = \min\left\{\frac{100}{5}; \frac{100}{4}; \frac{80}{2}\right\} = \min\{20; 25; 40\} = 20$. Тоді $x_3 = 0$ і x_3 переходить до числа неосновних змінних, а x_4 і x_5 залишаються додатними.

II крок. *Основні змінні:* x_2, x_4, x_5, x_6; *неосновні:* x_1, x_3.

Виразимо основні змінні і цільову функцію через неосновні. У системі (2) вибираємо те рівняння, з якого одержане мінімальне значення відношення вільного члену до коефіцієнта при x_2. У даному випадку це перше рівняння виділене рамкою. Підставимо вираз $x_2 = 20 - 0{,}2x_3$, отриманий із першого рівняння, у всі інші рівняння системи (2) і в цільову функцію F одержимо:

$$\begin{cases} x_2 = 20 - 0{,}2x_3 \\ x_4 = 200 - 2x_1 \\ \boxed{x_5 = 20 - 4x_1 + 0{,}8x_3} \\ x_6 = 40 - 2x_1 + 0{,}4x_3 \end{cases} \tag{3}$$
$$F = 100 + 4x_1 - x_3 \quad .$$

При $x_1 = x_3 = 0$ маємо $F = 4x_1 + 5(20 - 0{,}2x_3) = 100$.

Це краще ніж на I кроці, але не шуканий максимум.

Наступне збільшення функції F можливе за рахунок уведення змінної x_1 у число основних. Оскільки ця змінна входить у вираз F з додатнім коефіцієнтом, а тому її збільшення приводить до збільшення лінійної форми F і її невигідно вважати неосновною, тобто рівною нулю. Для відповіді на запитання, яку змінну вивести із основної в неосновну, приймемо $x_1 = \min\left\{\frac{200}{2}; \frac{20}{4}; \frac{40}{2}\right\} = 5$. Тоді $x_5 = 0$ і x_5 переходить до числа неосновних змінних, а x_4 і x_6 залишаються при цьому додатними. Перше рівняння не використовується при знаходженні вказаного мінімуму, оскільки x_1 не входить в це рівняння.

III крок. *Основні змінні:* x_1, x_2, x_4, x_6; *неосновні:* x_3, x_5.

Продовжуючи аналіз, аналогічно попередньому, із третього рівняння системи (3) (воно виділено) маємо: $x_1 = 5 + 0{,}2x_3 - 0{,}25x_5$. Підставимо цей вираз в інші рівняння системи (3) і в лінійну форму, одержимо:

$$\begin{cases} x_1 = 5 + 0{,}2x_3 - 0{,}25x_5 \\ x_2 = 20 - 0{,}2x_3 \\ x_4 = 190 - 0{,}4x_3 + 0{,}5x_5 \\ x_6 = 30 + 0{,}5x_5 \end{cases} \tag{4}$$
$$F = 120 - 0{,}2x_3 - x_5 \quad .$$

Оскільки у виразі лінійної форми (цільової функції) змінні x_3 і x_5 входять з від'ємними коефіцієнтами, то ніяке збільшення F за рахунок цих змінних неможливе.

Відсутність на будь-якому кроці симплексного методу у виразі лінійної форми F, максимум якої шукається, неосновних змінних з додатними коефіцієнтами є критерієм оптимальності.

Отже, на III кроці критерій оптимальності досягнуто і задача розв'язана. Оптимальним є розв'язок (5; 20; 0; 190; 0; 30), , за яким $F_{\max} = 120$. Таким чином, для одержання найбільшого прибутку рівного 120 грошових од. із даних запасів деревини, підприємство повинно виготовити 5 одиниць продукції *A* і 20 одиниць продукції *B*. При цьому деревина I, III виявиться використаною повністю, а 190 од. і 30 од. деревини відповідно II і IV виду залишаться невикористаними.

Зауваження. Могло б статися, що один із коефіцієнтів лінійної форми F, наприклад, при x_3, опинився додатнім, з чого випливло б, що максимальне значення ще не одержане, оскільки можливе збільшення F за рахунок введення в основні змінні x_3.

Тоді можливо зустрітися з положенням, яке потребуює додаткового роз'яснення.

По-перше, хоча змінна x_3 і входить у вираз для x_1 (перше рівняння системи (4)), але маючи додатній коефіцієнт при будь-якому зростанні x_3 змінна x_1 не може стати від'ємною.

Іншими словами, тоді у першому рівнянні ніяких обмежень на зростання x_3 не накладається, а тому умовно можливо було б написати ∞, тобто $x_3 = \min\left\{\infty; \frac{20}{0,2}; \frac{190}{0,4}; \infty\right\} = 100$.

Домовимося в подальшому користуватися цим же позначенням, якщо змінна, яка знову вводиться в число основних не входить у якенебудь рівняння системи обмежень.

Знаходження допустимого базисного розв'язку. Не завжди у задачі, що розглядається перший одержаний базисний розв'язок є допустимим. Іноді він знаходиться не відразу, а через деяке число кроків. При цьому слід пам'ятати, що на першому етапі симплексного методу, тобто при знаходженні будь-якого допустимого розв'язку, лінійна форма в розрахунок не береться, а всі перетворення належать тільки до системи обмежень.

Нехай задача лінійного програмування задана в загальному вигляді, тобто системою m лінійних рівнянь з n невідомими ($m < n$) (або вона приведена до такого вигляду після введення додаткових невід'ємних змінних). Виберемо групу m основних змінних, які дозволяють знайти вихідний базисний розв'язок. Виразимо ці основні змінні через неосновні, маємо:

$$\begin{cases} x_1 = k_1 + b_{1,m+1}x_{m+1} + \ldots + b_{1n}x_n \\ x_2 = k_2 + b_{2,m+1}x_{m+1} + \ldots + b_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_i = k_i + b_{i,m+1}x_{m+1} + \ldots + b_{in}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_m = k_m + b_{m,m+1}x_{m+1} + \ldots + b_{mn}x_n \ . \end{cases} \qquad (5)$$

Отже, цьому способу розбиття на основні і неосновні відповідає базисний розв'язок $(k_1;k_2;...;k_i;...;k_m;0;0;...;0)$. Розглянемо загальний випадок, коли цей розв'язок є недопустимим.

Від отриманого базисного розв'язку необхідно спочатку перейти до будь-якого допустимого базисного розв'язку, причому не обов'язково, щоб цей перехід здійснювався відразу в один крок.

Завжди через скінченне число кроків буде здійснений перехід до допустимого базисного розв'язку, якщо система обмежень не має протиріч.

Оскільки згідно з припущенням, вихідний базисний розв'язок є недопустимим, серед вільних членів системи обмежень (5) є хоча б один від'ємний (число від'ємних вільних членів цієї системи співпадає з числом від'ємних компонент вихідного базисного розв'язку). Нехай ним є вільний член k_i i-го рівняння, тобто основна змінна x_i у відповідному базисному розв'язку від'ємна.

Для переходу до нового базисного розв'язку необхідно розв'язати два питання:

1) з'ясувати, яку з неосновних змінних слід перевести в основні; 2) підібрати змінну, яку із основних треба перевести в неосновну на місце тієї, що вибула в основні змінні.

Як правило, при переведенні неосновної змінної в основні вона зростає, оскільки замість нуля у вихідному базисному розв'язку вона має прийняти додатне значення в новому базисному розв'язку (виключення буде мати місце лише при вироджені). А тому для вирішення першого питання необхідно вміти знаходити неосновні змінні, при збільшенні яких зростає основна змінна, від'ємна у вихідному базисному розв'язку.

Повертаючись до i-го рівняння системи (5), в якому вільний член k_i від'ємний, змінна x_i зростає при зростанні тих неосновних змінних, коефіцієнти яких в цих рівняннях додатні. Отже, в основні можна переводити ті неосновні змінні, які в рівнянні системи (5) з від'ємним вільним членом мають додатні коефіцієнти.

Можливі три випадки:

1. Неосновних змінних з додатними коефіцієнтами в i-му рівнянні системи (5) не існує, тобто всі коефіцієнти b_{ij} від'ємні (як і вільний член k_i). У цьому випадку дана система обмежень несумісна — вона не має ні одного допустимого розв'язку, оскільки змінна x_i не може прийняти невід'ємних значень. Але, якщо немає жодного допустимого розв'язку системи обмежень, то немає і оптимального.

2. Якщо в i-му рівнянні є одна змінна x_{m+j}, коефіцієнт при якій додатній, то в цьому випадку саме ця змінна переводиться в основні.

3. Якщо в i-му рівнянні є кілька змінних з додатними коефіцієнтами, то в цьому випадку в основні можна перевести будь-яку із них. Для того щоб встановити, яка основна змінна повинна бути переведена в число неосновних, слід скористатися раніше встановленим правилом: знаходять відношення вільних членів до коефіцієнтів при змінній, яка переводиться в основні в усіх рівняннях, де знаки вільних членів і вказаних коефіцієнтів протилежні, а потім розглядають абсолютну величину цих відношень, вибираючи серед них найменшу (якщо в деяких рівняннях знаки вільних членів і коефіцієнтів співпадають, або змінна, яка переводиться в основні, відсутня, то відношенням вважають рівним ∞). Виділене рівняння із якого одержане найменше відношення, показує яка із основних змінних повинна бути переведена в неосновні.

Після цього виразимо нові основні змінні через неосновні, перейшовши до наступного базисного розв'язку. У результаті одержується система аналогічна (5), в якій число від'ємних вільних членів або співпадає з їх числом у системі (5), або на одиницю менше ніж у вихідному. Якщо у виділеному рівнянні вільний член від'ємний (додатній або рівний нулю), то в новому базисному розв'язку число від'ємних компонент стане на одиницю менше (залишиться таким же) як у вихідному рівнянні.

Отже, одержується поліпшений базисний розв'язок, який ближче до області допустимих розв'язків системи обмежень. У випадку, якщо воно є недопустимим, то до нього слід застосувати ту ж схему знову.

У результаті, через скінченне число кроків одержується допустимий базисний розв'язок.

Після знаходження допустимого розв'язку, переходять до другого етапу симплексного методу, суть якого була розглянута при розв'язуванні задачі 1.

Задача 2. Знайти максимум функції $F = x_1 + 4x_2$ при обмеженнях:

$$\begin{cases} -x_1 + 4x_2 \geq 4 \\ x_1 + x_2 \geq 6 \\ x_1 - x_2 \leq 4 \\ x_2 \leq 3 \end{cases}$$
$$x_1 \geq 0,\ x_2 \geq 0\ .$$

Введемо додаткові невід'ємні змінні x_3, x_4, x_5, x_6 і зводимо дану систему нерівностей до еквівалентної їй системи рівнянь:

$$\begin{cases} -x_1 + 4x_2 - x_3 = 4 \\ x_1 + x_2 - x_4 = 6 \\ x_1 - x_2 + x_5 = 4 \\ x_2 + x_6 = 3 \end{cases}$$
$$x_j \geq 0,\ \left(j = 1, 2, ..., 6\right).$$

Введені додаткові змінні вважатимемо за основні, оскільки в цьому випадку базисний розв'язок системи легко знаходиться. Тоді x_1 та x_2 — неосновні змінні.

І крок. *Основні змінні:* x_3, x_4, x_5, x_6; *неосновні:* x_1, x_2.

Виразимо основні змінні через неосновні, матимемо:

$$\begin{cases} \boxed{x_3 = -4 - x_1 + 4x_2} \\ x_4 = -6 + x_1 + x_2 \\ x_5 = 4 - x_1 + x_2 \\ x_6 = 3 - x_2 \end{cases} .$$

Базисний розв'язок (0; 0; –4; –6; 4; 3), який відповідає вибраному розбиттю змінних на основні і неосновні є недопустимим (дві змінні від'ємні), а тому він не є оптимальним. Отже, від цього базисного розв'язку перейдемо до поліпшеного.

Розглянемо будь-яке із двох рівнянь останньої системи з від'ємними вільними членами, наприклад друге. Воно показує, що в основні змінні можливо перевести x_1 і x_2, оскільки в цьому рівнянні вони мають додатні коефіцієнти. Отже, при їх збільшенні, що відбудется, якщо переведемо будь-яку із них в основні змінні, змінна x_4 збільшиться.

Щоб встановити, яку змінну слід перевести із основних в неосновні, знайдемо абсолютну величину найменшого відношення вільних членів системи до коефіцієнтів при вибраній змінній. Спробуємо перевести в основні змінну x_1, маємо $x_1 = \min\left\{\infty; \frac{6}{1}; \frac{4}{1}; \infty\right\} = 4$. Воно одержано із третього рівняння, яке показує, що в неосновні необхідно перевести змінну x_5, яка у вихідному базисному розв'язку додатна.

Крім того, одержаний базисний розв'язок, як і вихідний, містить дві від'ємні компоненти, тобто при переході до такого базисного розв'язку поліпшення не відбудеться. При переведенні в основні змінної x_2, абсолютна величина найменшого відношення вільних членів системи до коефіцієнтів при x_2 складе $x_2 = \min\left\{\frac{4}{4}; \frac{6}{1}; \infty; \frac{3}{1}\right\} = 1$.

Воно одержано із першого рівняння, в якому вільний член від'ємний. Отже, переводячи x_2 в основні, а x_3 в не основні змінні, отримаємо базисний розв'язок, в якому число від'ємних компонент на одиницю менше ніж у вихідному. А тому зупинимося на такій можливості: переведемо x_2 в основні, а x_3 в не основні змінні, тоді виділеним стане перше рівняння.

II крок. *Основні змінні:* x_2, x_4, x_5, x_6; *неосновні:* x_1, x_3.

Виразимо нові основні змінні через нові неосновні.

$$\begin{cases} x_2 = 1 + 0,25x_1 + 0,25x_3 \\ \hline \boxed{x_4 = -5 + 1,25x_1 + 0,25x_3} \\ \hline x_5 = 5 - 0,75x_1 + 0,25x_3 \\ x_6 = 2 - 0,25x_1 - 0,25x_3 \end{cases}.$$

Як ми і передбачали в новому базисному розв'язку (0; 1; 0; –5; 5; 2) лише одна змінна від'ємна (а саме x_4), отже, він також є недопустимим, а тому не оптимальним.

Щоб перейти від отриманого базисного розв'язку до іншого, розглянемо рівняння з від'ємним вільним членом, тобто друге, яке показує, що в основні змінні можна перевести x_1 та x_3.

Якщо перевести в основні змінну x_1, то найменше із абсолютних величин відношень вільних членів системи до коефіцієнтів при x_1, дорівнює $x_1 = \min\left\{\infty; \frac{5}{1,25}; \frac{5}{0,75}; \frac{2}{0,25}\right\} = 4$.

Отже, в неосновні змінні слід перевести x_4. Оскільки найменше відношення одержане із другого рівняння то його виділяємо. У новому базисному розв'язку вже не виявиться від'ємних компонент, тобто воно є допустимим.

III крок. *Основні змінні:* x_1, x_2, x_5, x_6; *неосновні змінні:* x_3, x_4.
Виразимо основні змінні через неосновні:

$$\begin{cases} x_1 = 4 - 0,2x_3 + 0,8x_4 \\ x_2 = 2 + 0,2x_3 + 0,2x_4 \\ \boxed{x_5 = 2 + 0,4x_3 - 0,6x_4} \\ x_6 = 1 - 0,2x_3 - 0,2x_4 \end{cases}.$$

Новий базисний розв'язок має вигляд (4; 2; 0; 0; 2; 1). З урахуванням отриманого допустимого базисного розв'язку виразимо лінійну форму через неосновні змінні, тобто $F = 12 + 0,6x_3 + 1,6x_4$. Зауважимо, що лінійну форму виписуємо лише тоді, якщо базисний розв'язок є допустимим. Оскільки ми шукаємо максимум лінійної форми то очевидно, що отриманий базисний розв'язок не є оптимальним.

Переводимо до числа основних змінних x_4, яка має більший додатний коефіцієнт. Знаходимо $x_4 = \min\{\infty\ ;\ \infty\ ;\ 2:0,6\ ;\ 1:0,2\} = \frac{10}{3}$. Це найменше відношення отримано із третього рівняння, яке і виділимо. Воно показує, що якщо $x_4 = \frac{10}{3}$, то $x_5 = 0$ і тому перейде до числа неосновних.

IV крок. *Основні змінні:* x_1, x_2, x_4, x_6; *неосновні змінні:* x_3, x_5.
Виразимо основні змінні через неосновні, одержимо:

$$\begin{cases} x_1 = \frac{20}{3} + \frac{1}{3}x_3 - \frac{4}{3}x_5 \\ x_2 = \frac{8}{3} + \frac{1}{3}x_3 - \frac{1}{3}x_5 \\ x_4 = \frac{10}{3} + \frac{2}{3}x_3 - \frac{5}{3}x_5 \\ \boxed{x_6 = \frac{1}{3} - \frac{1}{3}x_3 + \frac{1}{3}x_5} \end{cases}.$$

Виразимо цільову функцію через ті ж неосновні змінні; маємо $F = \frac{52}{3} + \frac{5}{3}x_3 - \frac{8}{3}x_5$. Отже, базисний розв'язок до якого ми прийшли, не є оптимальним: $(\frac{20}{3}; \frac{8}{3}; 0; \frac{10}{3}; 0; \frac{1}{3})$.

Збільшення цільової функції можливе при переході до нового базису, в якому змінна x_3 є основною. Знаходимо $x_3 = \min\left\{\infty;\infty;\infty;\frac{1}{3}:\frac{1}{3}\right\} = 1$. Таке найменше відношення, одержане із четвертого рівняння системи і вказує, що якщо $x_3 = 1$ то змінна $x_6 = 0$ і переходить до числа неосновних.

V крок. *Основні змінні:* x_1, x_2, x_3, x_4; *неосновні змінні:* x_5, x_6.
Виразимо основні змінні через неосновні, одержимо:

$$\begin{cases} x_1 = 7 - x_5 - x_6 \\ x_2 = 3 - x_6 \\ x_3 = 1 + x_5 - 3x_6 \\ x_4 = 4 - x_5 - 2x_6 \ . \end{cases}$$

Цільова функція, виражена через неосновні змінні нового базисного розв’язку, має вигляд $F = 19 - x_5 - 5x_6$. Критерій оптимальності для випадку максимізації цільової функції виконаний. Отже, базисний розв’язок (7; 3; 1; 4; 0; 0) є оптимальним, а максимум цільової функції $F_{\max} = 19$.

Задача про суміші. Розглянемо одну із економічних задач, в якій потрібно знайти мінімум лінійної форми, при цьому знаходження допустимого базисного розв’язку має певні труднощі. До групи задач про суміші відносять задачі про відшукання найбільш дешевого набору із відпущених вихідних матеріалів, які забезпечують одержання суміші із заданими властивостями. Іншими словами, одержувані суміші повинні містити у своєму складі n різних компонент у певних кількостях, а самі компоненти є складовими частинами m вихідних матеріалів.

Узагальнена таблиця задачі про суміші наведена в табл. 4.3.

Таблиця 4.3

Види матеріалів	**Ціна одиниці матеріалу**	**Кількість окремих компонент, які входять як складові матеріалу**					
		1	2	...	j	...	n
1	c_1	a_{11}	a_{21}	...	a_{1j}	...	a_{1n}
2	c_2	a_{21}	a_{22}	...	a_{2j}	...	a_{2n}
...	...	...	...	...	...	...	...

i	c_i	a_{i1}	a_{i2}	...	a_{ij}	...	a_{in}
...	...	...	...	...	...	...	...
m	c_m	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mn}
Необхідна кількість компонент у суміші		b_1	b_2	...	b_j	...	b_n

Коефіцієнти a_{ij} вказують на кількість j-ої компоненти в одиниці i-го матеріалу.

Потрібно одержати суміш із заданими властивостями при найменших затратах на придбання матеріалів.

Позначимо через x_i кількість матеріалу i-го виду, який входить в суміш. Тоді задача зводиться до відшукання мінімуму функції:

$$F = c_1x_1 + c_2x_2 + ... + c_mx_m$$

при обмеженнях

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + ... + a_{i1}x_i + ... + a_{m1}x_m \geq b_1, \\ a_{12}x_1 + a_{22}x_2 + ... + a_{i2}x_i + ... + a_{m2}x_m \geq b_2, \\ .... \\ a_{1j}x_1 + a_{2j}x_2 + ... + a_{ij}x_i + ... + a_{mj}x_m \geq b_j, \\ .... \\ a_{1n}x_1 + a_{2n}x_2 + ... + a_{in}x_i + ... + a_{mn}x_m \geq b_n, \end{cases}$$
$$x_i \geq 0 \ \ (i = 1, 2, ..., m) \quad .$$

Одним із частинних випадків загальної задачі про суміші є задача про дієту. Особливістю такої задачі є задоволення потреби індивідуума в поживних речовинах при найменшій загальній вартості продуктів, що використовуються.

Задача 3. На птахофермі в процесі годування птиць, кожна одиниця має щоденно одержать не менше 9 од. речовини A, 12 од. речовини B і 16 од. речовини C (речовини A, B, C можуть означати жири, білки і вуглеводи). Для годування птиць можна закупити три види кормів: I, II, III (наприклад, житні висівки, зерно, комбікорм). Вміст кожної речовини в різних видах корма і вартість одиниці кожного корму наведені в таблиці 4.4.

Таблиця 4.4

Вид корму	**Речовини**			**Вартість одиниці корму**
	A	*B*	*C*	
I	3	2	4	4
II	2	3	5	5
III	4	1	3	1,75

Потрібно забезпечити найбільш дешевий раціон відгодування.

Складемо економіко-математичну модель задачі.

Нехай x_1, x_2 і x_3 — кількість одиниць відповідно I, II і III видів кормів. Необхідно знайти мінімум цільової функції (лінійної форми) $F = 4x_1 + 5x_2 + 1,75x_3$ при наступних обмеженнях:

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 \geq 9 \\ 2x_1 + 3x_2 + x_3 \geq 12 \\ 4x_1 + 5x_2 + 3x_3 \geq 16 \\ x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0\ . \end{cases} \tag{6}$$

Зведемо систему нерівностей до системи рівнянь, вводячи додаткові невід'ємні змінні x_4, x_5 і x_6:

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 - x_4 = 9 \\ 2x_1 + 3x_2 + x_3 - x_5 = 12 \\ 4x_1 + 5x_2 + 3x_3 - x_6 = 16 \end{cases} \tag{7}$$
$$x_j \geq 0 \ \left(j = 1, 2, ..., 6\right) \quad .$$

Простіше всього одержати базисний розв'язок системи (7), якщо за основні вибрати додаткові змінні x_4, x_5 і x_6. Але базисний розв'язок (0; 0; 0; –9; –12; –16), в якому ці змінні є основними — недопустимий, а тому для переходу до будь-якого допустимого базисного розв'язку скористаємось симплексним методом.

I крок. *Основні змінні:* x_4, x_5, x_6; *неосновні:* x_1, x_2, x_3.

Згідно з системою (7) основні змінні виражаються через неосновні таким чином:

$$\begin{cases} \boxed{x_4 = -9 + 3x_1 + 2x_2 + 4x_3} \\ x_5 = -12 + 2x_1 + 3x_2 + x_3 \\ x_6 = -16 + 4x_1 + 5x_2 + 3x_3\ . \end{cases} \tag{8}$$

Переведемо в основні змінну x_1. Оскільки $x_1 = \min\left\{\frac{9}{3}; \frac{12}{2}; \frac{16}{4}\right\} = 3$, то при $x_1 = 3$ маємо: $x_4 = 0$ і x_4 переходить у неосновні змінні.

II крок. *Основні змінні:* x_1, x_5, x_6; *неосновні:* x_2, x_3, x_4.

Якщо виразити x_1 із першого (виділеного) рівняння системи (8) і підставити його в інші рівняння, одержимо таку систему:

$$\begin{cases} x_1 = 3 - \frac{2}{3}x_2 - \frac{4}{3}x_3 + \frac{1}{3}x_4 \\ x_5 = -6 + \frac{5}{3}x_2 - \frac{5}{3}x_3 + \frac{2}{3}x_4 \\ \boxed{x_6 = -4 + \frac{7}{3}x_2 - \frac{7}{3}x_3 + \frac{4}{3}x_4}. \end{cases} \qquad (9)$$

Базисний розв'язок (3; 0; 0; 0; –6; –4) хоча і є недопустимим, але все ж таки кращий ніж на I кроці, оскільки містить лише дві від'ємні компоненти.

Переведемо в основну змінну x_4. Оскільки $x_4 = \min\left\{\infty; 6:\frac{2}{3}; 4:\frac{4}{3}\right\} = 3$, то при $x_4 = 3$ маємо: $x_6 = 0$ і x_6 переходить у неосновні змінні.

III крок. *Основні змінні:* x_1, x_4, x_5; *неосновні:* x_2, x_3, x_6.

Якщо виразити x_4 із третього (виділеного) рівняння системи (9) і підставити його в інші рівняння, одержимо систему, що виражає основні змінні через неосновні:

$$\begin{cases} x_4 = 3 - \frac{7}{4}x_2 + \frac{7}{4}x_3 + \frac{3}{4}x_6 \\ x_1 = 4 - \frac{5}{4}x_2 - \frac{3}{4}x_3 + \frac{1}{4}x_6 \\ \boxed{x_5 = -4 + \frac{1}{2}x_2 - \frac{1}{2}x_3 + \frac{1}{2}x_6}. \end{cases} \qquad (10)$$

Переведемо в основну змінну x_6. Вважаємо $x_6 = \min\left\{\infty; \infty; 4:\frac{1}{2}\right\} = 8$.

При $x_6 = 8$ маємо: $x_5 = 0$ і x_5 переходить у неосновні змінні.

IV крок. *Основні змінні:* x_1, x_4, x_6; *неосновні:* x_2, x_3, x_5.

Виразимо із системи (10) основні змінні через неосновні:

$$\begin{cases} x_1 = 6 - \dfrac{3}{2}x_2 - \dfrac{1}{2}x_3 + \dfrac{1}{2}x_5 \\ \boxed{x_4 = 9 - \dfrac{5}{2}x_2 + \dfrac{5}{2}x_3 + \dfrac{3}{2}x_5} \\ x_6 = 8 - x_2 + x_3 + 2x_5 \end{cases}. \qquad (11)$$

Одержаний на цьому кроці базисний розв'язок (6; 0; 0; 9; 0; 8) є допустимим, а тому перший етап симплексного методу закінчений.

Переходимо до другого етапу, тобто будемо шукати оптимальний розв'язок. Спочатку перевіряємо, чи не є знайдений допустимий розв'язок оптимальним. Для цього виразимо цільову функцію $F = 4x_1 + 5x_2 + 1,75x_3$ через неосновні змінні. Підставимо замість x_1 її вираз із першого рівняння системи (11) і приводячи подібні члени, одержимо: $F = 24 - x_2 - 0,25x_3 + 2x_5$.

У задачах, які розв'язувались раніше, шукали максимум цільової функції F. У даній задачі мова йде про мінімізацію функції цілі, а тому вигідні ті змінні, які входять у вираз лінійної форми з від'ємними коефіцієнтами. У зазначеному випадку таких змінних дві: x_2 і x_3.

Переведемо в основну змінну x_2. Оскільки

$x_2 = \min\left\{\left(6 : \dfrac{3}{2}\right);\ \left(9 : \dfrac{5}{2}\right);\ (8 : 1)\right\} = \dfrac{18}{5}$, то при $x_2 = \dfrac{18}{5}$ маємо: $x_4 = 0$ і x_4 переходить в неосновні змінні.

V крок. *Основні змінні:* x_1, x_2, x_6; *неосновні:* x_3, x_4, x_5.

Виразимо із системи (11) основні змінні і цільову функцію через неосновні, маємо:

$$\begin{cases} \boxed{x_1 = \dfrac{3}{5} - 2x_3 + \dfrac{3}{5}x_4 - \dfrac{2}{5}x_5} \\ x_2 = \dfrac{18}{5} + x_3 - \dfrac{2}{5}x_4 + \dfrac{3}{5}x_5 \\ x_6 = \dfrac{22}{5} + \dfrac{2}{5}x_4 + \dfrac{7}{5}x_5 \end{cases} \qquad (12)$$

$$F = \frac{102}{5} - \frac{5}{4}x_3 + \frac{2}{5}x_4 + \frac{7}{5}x_5 .$$

Переведемо в основну змінну x_3. Вважаємо $x_3 = \min\left\{\left(\frac{3}{5}:2\right);\infty;\infty\right\} = \frac{3}{10}$. Якщо $x_3 = \frac{3}{10}$ маємо: $x_1 = 0$ і x_1 переходить в неосновні змінні.

VI крок. *Основні змінні:* x_2, x_3, x_6; *неосновні:* x_1, x_4, x_5.

Виразимо із системи (12) основні змінні і лінійну форму через неосновні:

$$\begin{cases} x_2 = \frac{39}{10} - \frac{1}{2}x_1 - \frac{1}{10}x_4 + \frac{2}{5}x_5 \\ x_3 = \frac{3}{10} - \frac{1}{2}x_1 + \frac{3}{10}x_4 - \frac{1}{5}x_5 \\ x_6 = \frac{22}{5} + \frac{2}{5}x_4 + \frac{7}{5}x_5 \\ F = \frac{801}{40} + \frac{5}{8}x_1 + \frac{1}{40}x_4 + \frac{33}{20}x_5 \end{cases}.$$

У вираз для цільової функції F входять всі змінні з додатними коефіцієнтами, а це означає, що критерій оптимальності при відшукуванні мінімуму цієї функції збережений.

Отже, оптимальним є розв'язок $\left(0;\ \frac{39}{10};\ \frac{3}{10};\ 0;\ 0;\ \frac{22}{5}\right)$, за якого $F_{\min} = \frac{801}{40} = 20{,}025$ гр.од.

Крім того, щоб забезпечити найбільш дешевий раціон харчування, необхідно закупити самий дорогий корм виду II, корму III, який самий дешевий, потрібно закупити менше, а корм виду I взагалі невигідний.

За цим оптимальним розв'язком будуть забезпечені норми речовин A і B, а речовини C виявиться на $\frac{22}{5} = 4{,}4$ од. більше норми.

Особливі випадки. У задачах, що розглядалися, система обмежень виявилася сумісною і мав місце скінчений оптимум, причому єдиний. Проілюструємо на прикладах, коли ці умови порушуються.

Задача 4. Серед розв'язків системи нерівностей:

$$\begin{cases} -3x_1 + 4x_2 \le 12 \\ 2x_1 - 3x_2 \le 3 \end{cases}$$

знайти невід'ємні x_1, x_2, для яких величина $F = 2x_1 + 3x_2$ — найбільша.

Введемо додаткові невід'ємні змінні і переведемо систему обмежень-нерівностей до системи рівнянь

$$\begin{cases} -3x_1 + 4x_2 + x_3 = 12 \\ 2x_1 - 3x_2 + x_4 = 3 \end{cases}$$
$$x_j \geq 0 \ \left(j = 1,2,3,4\right) .$$

Одержимо два рівняння з чотирма змінними. Якщо взяти за основні змінні x_3 і x_4, то вихідний базисний розв'язок (0; 0; 12; 3) — допустимий.

I крок. *Основні змінні:* x_3, x_4; *неосновні:* x_1, x_2.

Виразимо основні змінні і цільову функцію через неосновні:

$$\begin{cases} x_3 = 12 + 3x_1 - 4x_2 \\ x_4 = 3 - 2x_1 + 3x_2 \end{cases}$$
$$F = 2x_1 + 3x_2 \quad .$$

Переведемо x_2 в основні змінні, оскільки x_2 входить у вираз F з більшим додатним коефіцієнтом. Покладемо $x_2 = \min\left\{\frac{12}{4};\ \infty\right\} = 3$. При $x_2 = 3$ маємо: $x_3 = 0$ і x_3 переходить у неосновні змінні.

II крок. *Основні змінні:* x_2, x_4; *неосновні:* x_1, x_3.

Виразимо основні змінні і цільову функцію через неосновні, одержимо:

$$\begin{cases} x_2 = 3 + \frac{3}{4}x_1 - \frac{1}{4}x_3, \\ x_4 = 12 + \frac{1}{4}x_1 - \frac{3}{4}x_3 \end{cases}$$
$$F = 9 + \frac{17}{4}x_1 - \frac{3}{4}x_3 \quad .$$

Із виразу функції F випливає, що змінну x_1 необхідно перевести в основні. Покладемо $x_1 = \min\{\infty;\ \infty\}$, тоді змінна x_1 може зростати необмежено (ні змінна x_2, ні x_4 при цьому не стануть від'ємними). Отже, і функція F, максимум якої потрібно знайти, також може необмежено зростати. Звідси можна вважати, що $F_{\max} = \infty$. Геометричне розв'язання цього прикладу показало б, що дана система нерівностей має необмежену область розв'язків і лінійна форма при таких обмеженнях може приймати скільки завгодно великі значення.

Задача 5. Серед розв'язків системи нерівностей:

$$\begin{cases} -2x_2 + x_2 \ge 3 \\ 2x_1 - 3x_2 \ge 6 \\ x_1 + x_2 \le 8 \end{cases}$$

знайти невід'ємні x_1, x_2, для яких величина $F = 2x_1 - x_2$ — найбільша.

Зведемо систему обмежень-нерівностей до системи рівнянь, шляхом введення невід'ємних додаткових змінних:

$$\begin{cases} -2x_2 + x_2 - x_3 = 3 \\ 2x_1 - 3x_2 - x_4 = 6 \\ x_1 + x_2 + x_5 = 8 \end{cases} .$$

Змінні x_3, x_4, x_5 візьмемо за основні. Відповідний базисний розв'язок $(0;\ 0;\ -3;\ -6;\ 8)$ — недопустимий. А тому перш за все скористаємося симплексним методом для знаходження допустимого базисного розв'язку.

I крок. *Основні змінні:* x_3, x_4, x_5; *неосновні:* x_1, x_2.

Виразимо основні змінні через неосновні (цільову функцію поки не розглядаємо):

$$\begin{cases} x_3 = -3 - 2x_1 + x_2 \\ x_4 = -6 + 2x_1 - 3x_2 \\ x_5 = 8 - x_1 - x_2 \end{cases} .$$

Переведемо x_1 в основні змінні. Вважаємо $x_1 = \min\left\{\infty;\ \frac{6}{2};\ \frac{8}{1}\right\} = 3$.

При $x_1 = 3$ маємо: $x_4 = 0$ і x_4 переходить в неосновні змінні.

II крок. *Основні змінні:* x_1, x_3, x_5; *неосновні:* x_2, x_4.

Знову виразимо основні змінні через неосновні, переходимо до системи:

$$\begin{cases} x_1 = 3 + \frac{3}{2}x_2 + \frac{1}{2}x_4, \\ x_3 = -9 - 2x_2 - x_4 \\ x_5 = 5 - \frac{5}{2}x_2 - \frac{1}{2}x_4 \end{cases} .$$

У другому рівнянні одержаної системи і вільний член і всі коефіцієнти при неосновних змінних від'ємні. Це є ознакою того, що дана система несумісна. Вона не має ні одного розв'язку, в тому числі і оптимального (геометричний розв'язок було розглянуто у прикладі 11 §2).

Задача 6. Серед розв'язків системи нерівностей:

$$\begin{cases} -2x_1 + x_2 \le 3 \\ 2x_1 - x_2 \le 6 \\ x_1 + x_2 \le 8 \end{cases}$$

знайти невід'ємні x_1, x_2, для яких величина $F = 2x_1 - x_2$ — найбільша.

Зведемо систему нерівностей до системи рівнянь, шляхом введення додаткових змінних:

$$\begin{cases} -2x_1 + x_2 + x_3 = 3 \\ 2x_1 - x_2 + x_4 = 6 \\ x_1 + x_2 + x_5 = 8 \end{cases}$$
$$x_j \ge 0 \ \left(j = 1, 2, ..., 5\right).$$

Виберемо за основні, додаткові змінні, одержимо базисний розв'язок $\left(0;\ 0;\ 3;\ 6;\ 8\right)$, який допустимий, а тому його можна прийняти за вихідний на I кроці розв'язування.

I крок. *Основні змінні:* x_3, x_4, x_5; *неосновні:* x_1, x_2.

Виразимо основні змінні і лінійну форму через неосновні:

$$\begin{cases} x_3 = 3 + 2x_1 - x_2 \\ \boxed{x_4 = 6 - 2x_1 + x_2} \\ x_5 = 8 - x_1 - x_2 \end{cases}$$
$$F = 2x_1 - x_2 \qquad .$$

Переведемо x_1 в основні змінні. Знаходимо $x_1 = \min\left\{\infty;\ \frac{6}{2};\ \frac{8}{1}\right\} = 3$. При $x_1 = 3$ маємо: $x_4 = 0$ і x_4 переходить у неосновні змінні.

II крок. *Основні змінні:* x_1, x_3, x_5; *неосновні:* x_2, x_4.

Маємо:

$$\begin{cases} x_1 = 3 + \frac{1}{2}x_2 - \frac{1}{2}x_4, \\ x_3 = 9 - x_4 \\ x_5 = 5 - \frac{3}{2}x_2 + \frac{1}{2}x_4 \end{cases}$$
$$F = 6 + x_2 - x_4 - x_2 = 6 - x_4 .$$

У виразі лінійної форми F відсутня одна із неосновних змінних (x_2), а критерій оптимальності виконаний (можна вважати, що x_2 входить у вираз F з нульовим коефіцієнтом). А тому ця змінна не є ні вигідною ні невигідною.

Спробуємо все ж таки перевести x_2 в основні змінні. Вважатимемо $x_2 = \min\left\{\infty;\ \infty;\ 5 : \frac{3}{2}\right\} = \frac{10}{3}$. Заключаємо, що x_5 переходить в неосновні змінні.

III крок. *Основні змінні:* x_1, x_2, x_3; *неосновні:* x_4, x_5.

Маємо:

$$\begin{cases} x_1 = \frac{14}{3} - \frac{1}{3}x_4 - \frac{1}{3}x_5, \\ x_2 = \frac{10}{3} + \frac{1}{3}x_4 - \frac{2}{3}x_5 \\ x_3 = 9 - x_4 \end{cases}$$
$$F = 6 - x_4 .$$

Таким чином, і на цьому кроці одержується той же самий вираз цільової функції, а критерій оптимальності знову виконаний.

На II кроці оптимальний розв'язок мав компоненти $(3;\ 0;\ 9;\ 0;\ 5)$, а на III кроці — $\left(\frac{14}{3}; \frac{10}{3};\ 9;\ 0;\ 0\right)$.

Але обидва оптимальні розв'язки дають одне і теж максимальне значення: $F_{\max} = 6$.

Звідси випливає, що єдиність оптимального розв'язку може порушуватися. Це відбувається в тому випадку, коли на якомусь кроці розв'язування критерій оптимальності виконується, а у виразі лінійної форми F відсутня одна із неосновних змінних.

За геометричним розв'язуванням (див. подібний приклад 13, §2) було встановлено, що оптимальними розв'язками є всі точки відрізка CD (рис. 45).

ЗАДАЧІ

Використовуючи симплексний метод розв'язати наступні задачі:

Знайти максимум функції F за заданими обмеженнями:

4.22.

$$\begin{cases} x_1 + x_4 - 2x_5 = 1 \\ x_2 - 2x_4 + x_5 = 2 \\ x_3 + 3x_4 + x_5 = 3 \end{cases}$$
$$F = -x_4 + x_5 \quad .$$

4.23.

$$\begin{cases} x_1 - x_2 + x_3 = 1 \\ x_2 - 2x_3 + x_4 = 2 \end{cases}$$
$$F = x_1 + x_2 \quad .$$

4.24.

$$\begin{cases} x_1 - x_2 \ge -5 \\ 3x_1 - x_2 \ge -3 \\ 2x_1 + x_2 \le 14 \\ x_1 \le 5 \end{cases}$$
$$x_1 \ge 0 \ , \ \ x_2 \ge 0$$
$$F = x_1 + x_2 \quad .$$

4.25.

$$\begin{cases} x_1 + x_2 \le 1 \\ x_1 + x_2 \ge 4 \\ x_1 \le 3 \end{cases}$$
$$x_1 \ge 0 \ , \ \ x_2 \ge 0$$
$$F = x_1 + 2x_2 \quad .$$

4.26.

$$\begin{cases} x_1 + x_2 \ge 1 \\ -2x_1 + x_2 \le 2 \\ x_1 - x_2 \le 2 \end{cases}$$
$$x_1 \ge 0 \ , \ \ x_2 \ge 0$$
$$F = x_1 + 2x_2 \quad .$$

4.27.

$$\begin{cases} x_1 + x_2 \ge 1 \\ -2x_1 + x_2 \le 2 \\ x_1 + x_2 \le 4 \\ x_1 \le 3 \end{cases}$$
$$x_1 \ge 0 \ , \ \ x_2 \ge 0$$
$$F = x_1 + x_2 \quad .$$

Знайти мінімум функції F за заданими обмеженнями:

4.28.

$$\begin{cases} 2x_1 - x_2 \geq 2 \\ -x_1 + 2x_2 \geq 2 \\ x_1 - 2x_2 \geq -5 \\ -x_1 + x_2 \geq -1 \end{cases}$$

$$x_1 \geq 0 \ , \ \ x_2 \geq 0$$

$$F = x_1 + x_2 \quad .$$

4.29.

$$\begin{cases} x_1 + x_2 - 4 \geq 0 \\ -x_1 + 3x_2 + 6 \geq 0 \\ x_1 - 2 \geq 0 \end{cases}$$

$$x_1 \geq 0 \ , \ \ x_2 \geq 0$$

$$F = 2x_1 - 6x_2 \quad .$$

4.30.

$$\begin{cases} x_1 + x_2 + 2x_3 - x_4 = 3 \\ x_2 + 2x_4 = 1 \end{cases}$$

$$F = 5x_1 - x_2 \quad .$$

Відповіді:

4.22. $F_{\max} = \frac{11}{5}$, оптимальний розв'язок $(\frac{28}{5};\ 0;\ 0;\ 0;\ \frac{1}{5};\ \frac{12}{5})$.

4.23. $F_{\max} = +\infty$, оптимальний розв'язок не існує.

4.24. $F_{\max} = 11$.

4.25. Розв'язків не має.

4.26. $F_{\max} = \infty$.

4.27. $F_{\max} = 4$ (оптимальний розв'язок не єдиний).

4.28. $F_{\min} = 4$.

4.29. $F_{\max} = -\infty$.

4.30. $F_{\min} = -\frac{7}{4}$, оптимальний розв'язок $(0; 0;\ \frac{7}{4};\ \frac{1}{2})$.

§4. Двоїсті задачі

Складання двоїстої задачі та основні теореми. Кожній задачі лінійного програмування можна зіставити певним чином з нею зв'язану іншу задачу, яка називається двоїстою щодо до першої.

Якщо вихідна задача (I) лінійного програмування полягає у *мінімізації* лінійної функції $F = c_1x_1 + c_2x_2 + ... + c_nx_n$, при заданих обмеженнях у формі нерівностей

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \ge b_1, \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \ge b_2, \\ .... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \ge b_n, \end{cases}$$

за умовами невід'ємності x_j (j = 1, 2, ..., n), то з нею пов'язана двоїста задача (І′), яка полягає в тому, що потрібно максимізувати лінійну функцію $Z = b_1y_1 + b_2y_2 + ... + b_my_m$ за умови обмежень

$$\begin{cases} a_{11}y_1 + a_{21}y_2 + ... + a_{m1}y_m \le c_1, \\ a_{12}y_1 + a_{22}y_2 + ... + a_{m2}y_m \le c_2, \\ .... \\ a_{1n}y_1 + a_{2n}y_2 + ... + a_{mn}y_m \le c_n, \end{cases}$$

і невід'ємності $y_i \ge 0 \; \left(\left(i = 1, 2, ..., m \right) \right)$.

Відзначимо, що в задачі І і в двоїстій задачі І′ матриці

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} \text{ і } A' = \begin{pmatrix} a_{11} & a_{21} & ... & a_{m1} \\ a_{12} & a_{22} & ... & a_{m2} \\ ... & ... & ... & ... \\ a_{1n} & a_{2n} & ... & a_{mn} \end{pmatrix},$$

складені із коефіцієнтів при змінних, одержуються один із одного транспонування. У правих частинах системи обмежень кожної задачі стоять коефіцієнти лінійної функції, взятої із другої задачі. У системі обмежень задачі І (мінімізація) всі нерівності типу "$\ge$", а в системі обмежень задачі І′ (максимізація) всі нерівності типу "$\le$". Поняття двоїстості є взаємним, тобто якщо задачу І′ записати у формі, аналогічній задачі І, то двоїстою до неї буде вихідна задача І. А тому задачі І і І′ називаються симетричними *взаємно двоїстими* або взаємно спряженими. Ми будемо вивчати тільки симетричні двоїсті задачі, а тому будемо називати їх коротко *двоїстими задачами*.

Отже, кожній задачі лінійного програмування можливо поставити у відповідність двоїсту їй задачу. Початкову задачу будемо називати вихідною (або прямою). Пряма і двоїста задачі, взяті разом, утворюють пару взаємно двоїстих задач, причому будь-яку із них можна розглянути як вихідну, тоді інша буде двоїстою їй.

Звідси випливають наступні правила складання задачі, двоїстої щодо вихідної:

1. Зводять усі нерівності системи обмежень вихідної задачі до нерівностей одного змісту: якщо у вихідній задачі шукається максимум лінійної форми — до вигляду $\le$; якщо ж мінімум — до вигляду $\ge$. Для цього нерівності, в яких ця вимога не виконується, множать на (–1).

2. Виписують матрицю A коефіцієнтів при змінних вихідної задачі, одержаних після перетворення п. 1 і складають матрицю A', транспоновану щодо матриці A.

3. Складають систему обмежень двоїстої задачі, взявши як коефіцієнти при змінних елементи матриці A', а як вільні члени — коефіцієнти при змінних у лінійній формі вихідної задачі і записують нерівності протилежного змісту в порівнянні з нерівностями, одержаними у п.1.

4. Складають лінійну форму двоїстої задачі, прийнявши за коефіцієнти при змінних вільні члени системи обмежень вихідної задачі, одержаної у п. 1.

5. Вказують, що необхідно знайти при розв'язуванні двоїстої задачі; мінімум лінійної форми, якщо у вихідній задачі шукається максимум, і максимум, якщо у вихідній задачі шукається мінімум.

6. Записують умову невід'ємності змінних двоїстої задачі.

Приклад 1. Скласти задачу, двоїсту наступній: знайти максимум функції $F = 5x_1 + 3x_2$ при обмеженнях:

$$\begin{cases} 2x_1 - 3x_2 \le 3 \\ -2x_1 + 2x_2 \le 3 \\ x_1 + x_2 \ge 1 \\ 2x_1 + 2x_2 \le 5 \end{cases} . \qquad (1)$$
$$x_1 \ge 0,\ x_2 \ge 0$$

Третя нерівність системи не задовольняє п. 1 правила складання двоїстої задачі, а тому помножимо на (–1)

$$\begin{cases} 2x_1 - 3x_2 \le 3 \\ -2x_1 + 2x_2 \le 3 \\ -x_1 - x_2 \le -1 \\ 2x_1 + 2x_2 \le 5 \end{cases} .$$

Для полегшення складання двоїстої задачі краще користуватися розширеною матрицею B, в яку поряд з коефіцієнтами при змінних системи обмежень вихідної задачі запишемо вільні члени і коефіцієнти при змінних у лінійній формі, виділивши для цієї цілі додаткові стовпець і рядок. Матрицю B транспонуємо і, використовуючи транспоновану матрицю B', складаємо задачу, двоїсту вихідній. У даному випадку матриці B і B' мають вигляд:

$$B=\left(\begin{array}{cc|c} 2 & -3 & 3 \\ -2 & 2 & 3 \\ -1 & -1 & -1 \\ 2 & 2 & 5 \\ \hline 5 & 3 & F \end{array}\right),\quad B'=\left(\begin{array}{cccc|c} 2 & -2 & -1 & 2 & 5 \\ -3 & 2 & -1 & 2 & 3 \\ \hline 3 & 3 & -1 & 5 & Z \end{array}\right).$$

Двоїста задача зводиться до знаходження мінімуму функції $Z=3y_1+3y_2-y_3+5y_4$, при обмеженнях:

$$\begin{cases} 2y_1-2y_2-y_3+2y_4\geq 5 \\ -3y_1+2y_2-y_3+2y_4\geq 3 \end{cases} \tag{2}$$
$$y_1\geq 0,\ y_2\geq 0,\ y_3\geq 0,\ y_4\geq 0.$$

Приклад 2. Який геометричний зміст наступних взаємно двоїстих задач.

Вихідна задача (I): знайти невід'ємні значення (x_1,x_2) із умов:

$$\begin{cases} 2x_1+3x_2\geq 6 \\ x_1-2x_2\geq -2 \end{cases}$$

і мінімізації лінійної функції $F=4x_1+3x_2$.

Двоїста задача (I′): знайти невід'ємні значення (y_1,y_2) із умов:

$$\begin{cases} 2y_1+y_2\leq 4 \\ 3y_1-2y_2\leq 3 \end{cases}$$

і максимізації лінійної функції $Z=6y_1-2y_2$.

Побудуємо систему обмежень задач I і I′.

У точці $P\left(\frac{6}{7};\frac{10}{7}\right)$ досягається мінімум лінійної функції F, тобто $F_{\min}=4\cdot\frac{6}{7}+3\cdot\frac{10}{7}=\frac{54}{7}$, а в точці $P'\left(\frac{11}{7};\frac{6}{7}\right)$ — максимум лінійної функції Z, тобто $Z_{\max}=6\cdot\frac{11}{7}-2\cdot\frac{6}{7}=\frac{54}{7}$ (рис. 51). Звідси: $F_{\min}=Z_{\max}$.

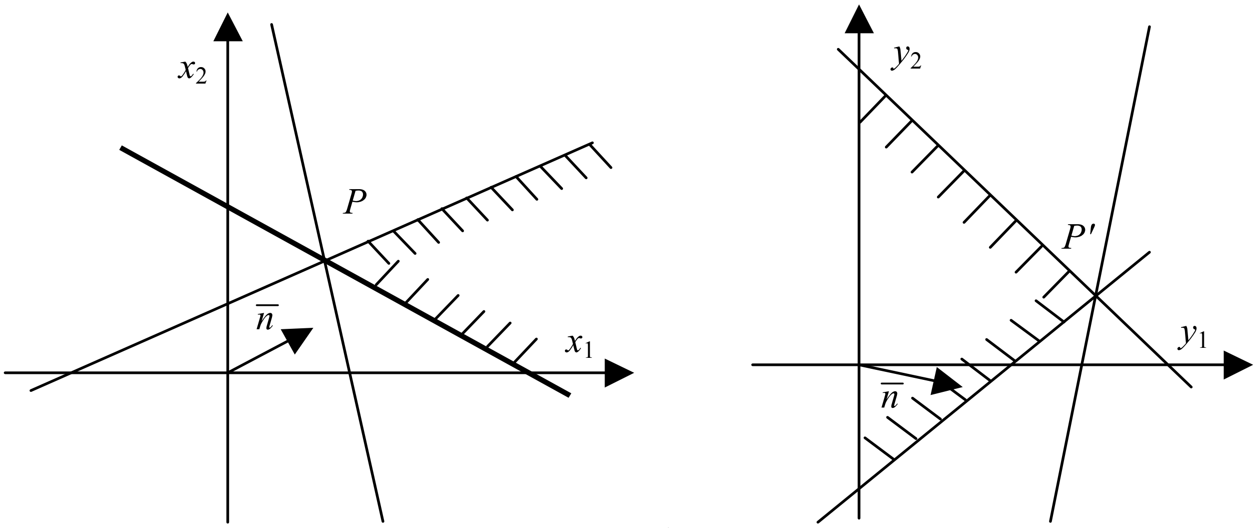


Рис. 51

Теорія двоїстості в лінійному програмуванні базується на таких основних теоремах.

ТЕОРЕМА 1. *Якщо одна із задач лінійного програмування має скінченний оптимум, то і двоїста до неї також має скінченний оптимум, причому оптимальне значення цільових функцій обох задач співпадають, тобто* $F_{\max} = Z_{\min}$*, або* $F_{\min} = Z_{\max}$*. Якщо ж цільова функція однієї із двоїстих задач не обмежена, то умови другої задачі несумісні.*

Приклад 2 підтверджує сформульоване твердження.

Перш ніж сформулювати наступну теорему, встановимо відповідність між змінними у вихідній і двоїстій задачах.

Відомо, що при розв'язуванні симплексним методом вихідної задачі для зведення системи нерівностей до еквівалентної їй системи рівнянь необхідно ввести m додаткових невід'ємних змінних (за числом нерівностей у системі обмежень) $x_{n+1}, x_{n+2}, ..., x_{n+i}, ..., x_{n+m}$, де $i = 1, 2, ..., m$ означає номер нерівності, в яке була введена додаткова змінна x_{n+i}.

Система обмежень двоїстої задачі складається із n нерівностей, які містять m змінних.

Якщо розв'язувати цю задачу симплексним методом, то слід ввести n додаткових невід'ємних змінних $y_{m+1}, y_{m+2}, ..., y_{m+j}, ..., y_{m+n}$, де $j = 1, 2, ..., n$ означає номер нерівності системи обмежень двоїстої задачі, в яке була введена додаткова змінна y_{m+j}. Відповідність між змінними у вихідній і двоїстій задачах буде така:

$$\begin{array}{cccccccccccc} x_1 & x_2 & \ldots & x_j & \ldots & x_n & x_{n+1} & x_{n+2} & \ldots & x_{n+i} & \ldots & x_{n+m} \\ \updownarrow & \updownarrow & \ldots & \updownarrow & \ldots & \updownarrow & \updownarrow & \updownarrow & \ldots & \updownarrow & \ldots & \updownarrow \\ y_{m+1} & y_{m+2} & \ldots & y_{m+j} & \ldots & y_{m+n} & y_1 & y_2 & \ldots & y_i & \ldots & y_m, \end{array}$$

де кожній початковій змінній вихідної задачі x_j ($j = 1, 2, ..., n$) ставиться у відповідність додаткова змінна y_{m+j}, введена в j-у нерівність двоїстої задачі, а кожній додатковій змінній x_{n+i} вихідної задачі ($i = 1, 2, ..., m$), введеної в i-у нерівність вихідної задачі — початкова змінна y_i двоїстої *задачі.*

ТЕОРЕМА 2. *Компоненти оптимального розв'язку однієї із задач (прямої або двоїстої) дорівнюють абсолютним величинам коефіцієнтів при відповідних змінних у виразі цільової функції другої задачі (двоїстої або прямої) при досягненні нею оптимуму і за умови, що одержаний оптимальний розв'язок не є виродженим.*

Із теорем 1 і 2 випливає, що якщо розв'язати одну із взаємно двоїстих задач, тобто знайти її оптимальний розв'язок і оптимум цільової функції, то можна записати оптимальний розв'язок і оптимум цільової функції другої задачі.

Оскільки доведення цих теорем не наводитися, переконаємося в їх справедливості на відповідних прикладах.

Приклад 3. Розв'язати симплексним методом пряму і двоїсту задачу, наведену в прикладі 1.

Розв'язування прямої задачі. Зведемо систему обмежень-нерівностей (1) до системи рівнянь, ввівши невід'ємні додатні змінні:

$$\begin{cases} 2x_1 - 3x_2 + x_3 = 3 \\ -2x_1 + 2x_2 + x_4 = 3 \\ x_1 + x_2 - x_5 = 1 \\ 2x_1 + 2x_2 + x_6 = 5 \end{cases}$$
$$x_j \ge 0 \ \left(j = 1, 2, ..., 6\right).$$

На 1-му кроці розв'язування за *основні* візьмемо додаткові змінні x_3, x_4, x_5, x_6; *неосновні змінні* x_1, x_2.

І крок. Виразимо основні змінні через неосновні (оскільки відповідний базисний розв'язок не є допустимим, цільову функцію на цьому кроці не розглядають).

$$\begin{cases} x_3 = 3 - 2x_1 + 3x_2 \\ x_4 = 3 + 2x_1 - 2x_2 \\ \boxed{x_5 = -1 + x_1 + x_2} \\ x_6 = 5 - 2x_1 - 2x_2\,. \end{cases}$$

Для одержання допустимого базисного розв’язку змінну x_1 переведемо в основні, при цьому змінна x_5 переходить в неосновні, оскільки $x_1 = \min\left\{\frac{3}{2};\infty;\frac{1}{1};\frac{5}{2}\right\} = 1$.

II крок. *Основні змінні* x_1, x_3, x_4, x_6; *неосновні змінні* x_2, x_5. Виразимо основні змінні і цільову функцію через неосновні:

$$\begin{cases} x_1 = 1 - x_2 + x_5 \\ \boxed{x_3 = 1 + 5x_2 - 2x_5} \\ x_4 = 5 - 4x_2 + 2x_5 \\ x_6 = 3 - 2x_5 \end{cases}$$
$$F = 5 - 2x_2 + 5x_5\ .$$

Переводячи в основні змінні x_5, яка $x_5 = \min\left\{\infty;\frac{1}{2};\infty;\frac{3}{2}\right\} = \frac{1}{2}$, переконуємося, що змінну x_3 необхідно перевести в неосновні.

III крок. *Основні змінні* x_1, x_4, x_5, x_6; *неосновні змінні* x_2, x_3. Маємо:

$$\begin{cases} x_1 = \frac{3}{2} + \frac{3}{2}x_2 - \frac{1}{2}x_3 \\ x_4 = 6 + x_2 - x_3 \\ x_5 = \frac{1}{2} + \frac{5}{2}x_2 - \frac{1}{2}x_3 \\ \boxed{x_6 = 2 - 5x_2 + x_3} \end{cases}$$
$$F = \frac{15}{2} + \frac{21}{2}x_2 - \frac{5}{2}x_3\ .$$

Змінну x_2 переведемо в основні, яка $x_2 = \min\left\{\infty;\infty;\infty;\frac{2}{5}\right\} = \frac{2}{5}$. Тоді як змінна x_6 переходить у неосновні.

IV крок. *Основні змінні* x_1, x_2, x_4, x_5; *неосновні змінні* x_3, x_6. Маємо:

$$\begin{cases} x_1 = \dfrac{21}{10} - \dfrac{1}{5}x_3 - \dfrac{3}{10}x_6 \\ x_2 = \dfrac{2}{5} + \dfrac{1}{5}x_3 - \dfrac{1}{5}x_6 \\ \boxed{x_4 = \dfrac{32}{5} - \dfrac{4}{5}x_3 - \dfrac{1}{5}x_6} \\ x_5 = \dfrac{3}{2} - \dfrac{1}{2}x_6 \end{cases}$$

$$F = \frac{117}{10} - \frac{2}{5}x_3 - \frac{21}{10}x_6.$$

Критерій оптимальності виконаний. Оптимальний розв’язок має вигляд $\left(\dfrac{21}{10}; \dfrac{2}{5}; 0; \dfrac{32}{5}; \dfrac{3}{2}; 0\right)$, при цьому розв’язуванні $F_{\max} = \dfrac{117}{10}$.

Розв’язування двоїстої задачі. Систему обмежень (2) двоїстої задачі зведемо до системи рівнянь:

$$\begin{cases} 2y_1 - 2y_2 - y_3 + 2y_4 - y_5 = 5 \\ -3y_1 + 2y_2 - y_3 + 2y_4 - y_6 = 3 \end{cases}$$
$$y_j \ge 0 \ \left(j = 1, 2, \ldots, 6\right).$$

На І-му кроці за основні приймаємо додаткові змінні y_5, y_6.

І крок. *Основні змінні* y_5, y_6; *неосновні змінні* $y_1; y_2, y_3, y_4$. Виразимо основні змінні через неосновні:

$$\begin{cases} y_5 = -5 + 2y_1 - 2y_2 - y_3 + 2y_4 \\ \boxed{y_6 = -3 - 3y_1 + 2y_2 - y_3 + 2y_4}. \end{cases}$$

Переведемо в основну змінну y_4, тоді y_6 переходить в неосновні, оскільки $y_4 = \min\left\{\dfrac{5}{2}; \dfrac{3}{2}\right\} = \dfrac{3}{2}$.

ІІ крок. *Основні змінні* y_4, y_5; *неосновні змінні* $y_1; y_2, y_3, y_6$. Маємо:

$$\begin{cases} y_4 = \dfrac{3}{2} + \dfrac{3}{2}y_1 - y_2 + \dfrac{1}{2}y_3 + \dfrac{1}{2}y_6 \\ \left| y_5 = -2 + 5y_1 - 4y_2 + y_6 \right| . \end{cases}$$

Змінну y_1 переводимо в основні. Оскільки $y_1 = \min\left\{\infty; \dfrac{2}{5}\right\} = \dfrac{2}{5}$, то змінна y_5 переходить в неосновні.

III крок. *Основні змінні* y_1, y_4; *неосновні змінні* $y_2; y_3, y_5, y_6$. Переконавшись, що відповідний базисний розв'язок задачі є допустимим, то виразимо через неосновні змінні не тільки основні , але і цільову функцію:

$$\begin{cases} y_1 = \dfrac{2}{5} + \dfrac{4}{5}y_2 + \dfrac{1}{5}y_5 - \dfrac{1}{5}y_6 \\ y_4 = \dfrac{21}{10} + \dfrac{1}{5}y_2 + \dfrac{1}{2}y_3 + \dfrac{3}{10}y_5 + \dfrac{1}{5}y_6 \end{cases}$$

$$Z = \frac{117}{10} + \frac{32}{5}y_2 + \frac{3}{2}y_3 + \frac{21}{10}y_5 + \frac{2}{5}y_6 .$$

Критерій оптимальності (для мінімізації цільової функції) виконаний. Оптимальний розв'язок має вигляд $\left(\dfrac{2}{5};\ 0;\ 0; \dfrac{21}{10};\ 0;\ 0\right)$, при цьому $Z_{\min} = \dfrac{117}{10}$.

Розв'язавши двоїсту задачу, переконалися в справедливості першої частини теореми 1: двоїста задача також має скінчений оптимум, причому $Z_{\min} = F_{\max} = \dfrac{117}{10}$.

Переконаємося, що справедливе і твердження теореми 2. Для цього випишемо змінні прямої і двоїстої задачі, дотримуючись їх відповідності:

$$\begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ y_5 & y_6 & y_1 & y_2 & y_3 & y_4 . \end{array}$$

Цільову функцію, отриману на останньому кроці розв'язуванні двоїстої задачі, виразимо через всі змінні цієї задачі:

$$Z = \frac{117}{10} + 0 \cdot y_1 + \frac{32}{5} y_2 + \frac{3}{2} y_3 + 0 \cdot y_4 + \frac{21}{10} y_5 + \frac{2}{5} y_6 \,.$$

Враховуючи відповідність коефіцієнтів при змінних y_j з коефіцієнтами при змінних x_i, одержимо розв'язок $\left(\frac{21}{10};\frac{2}{5};0;\frac{32}{5};\frac{3}{2};0\right)$, який співпадає з оптимальним розв'язком прямої задачі.

Зауваження. За розв'язком прямої задачі можна зразу одержати розв'язок двоїстої задачі. Якщо виразити цільову функцію *F*, одержану на IV кроці розв'язування прямої задачі, через всі змінні цієї задачі, то одержимо

$$F = \frac{117}{10} + 0 \cdot x_1 + 0 \cdot x_2 - \frac{2}{5} x_3 + 0 \cdot x_4 + 0 \cdot x_5 - \frac{21}{10} x_6 \,.$$

Згідно з теоремою 2, за відповідністю між змінними в прямій і двоїстих задачах і взявши абсолютну величину коефіцієнтів при змінних, знайдемо оптимальний розв'язок двоїстої задачі $\left(\frac{2}{5};\ 0;\ 0;\frac{21}{10};\ 0;\ 0\right)$.

При цьому $Z_{\min} = F_{\max} = \frac{117}{10}$.

Цим зручно користуватися, якщо розв'язування однієї із задач, наприклад прямої, пов'язано з певними труднощами.

Достатньо нагадати розв'язування задачі про суміші, де частина кроків була втрачена на пошуки допустимого базисного розв'язку. Згідно з теоремою 2, оптимальний розв'язок цієї задачі можливо отримати, розв'язавши двоїсту задачу. Переконаємося в цьому.

Приклад 4. Скласти і розв'язати симплексним методом задачу, двоїсту задачі про суміші (див. §3).

Оскільки всі нерівності системи (1) мають вигляд $\geq$, що відповідає знаходженню мінімуму цільової функції $F = 3x_1 + 4x_2 + 0,5x_3$, то п. 1 правила складання двоїстої задачі виконується.

Запишемо матрицю *B* прямої задачі і транспоновану матрицю *B′* двоїстої задачі:

$$B=\left(\begin{array}{ccc|c} 3 & 2 & 4 & 9 \\ 2 & 3 & 1 & 12 \\ 4 & 5 & 3 & 16 \\ \hline 4 & 5 & 1,75 & F \end{array}\right),\quad B'=\left(\begin{array}{ccc|c} 3 & 2 & 4 & 4 \\ 2 & 3 & 5 & 5 \\ 4 & 1 & 3 & 1,75 \\ \hline 9 & 12 & 16 & Z \end{array}\right).$$

Двоїста задача полягає у знаходженні максимуму функції $Z=9y_1+12y_2+16y_3$, за обмеженнями:

$$\begin{cases} 3y_1+2y_2+4y_3\le 4 \\ 2y_1+3y_2+5y_3\le 5 \\ 4y_1+y_2+3y_3\le 1,75 \end{cases}$$
$$y_1\ge 0,\ y_2\ge 0,\ y_3\ge 0.$$

Перетворимо систему нерівностей у систему рівнянь шляхом введення додаткових невід’ємних змінних:

$$\begin{cases} 3y_1+2y_2+4y_3+y_4=4 \\ 2y_1+3y_2+5y_3+y_5=5 \\ 4y_1+y_2+3y_3+y_6=1,75. \end{cases}$$
$$y_j\ge 0\ \left(j=1,2,...,6\right)$$

За основні приймемо додаткові змінні $y_4;y_5,y_6$. Неважко переконатися, що вихідний базисний розв’язок рівняння є допустимим.

І крок. *Основні змінні* y_4,y_5,y_6; *неосновні змінні* $y_1;y_2,y_3$. Виразимо основні змінні і цільову функцію Z через неосновні:

$$\begin{cases} y_4=4-3y_1-2y_2-4y_3 \\ y_5=5-2y_1-3y_2-5y_3 \\ \boxed{y_6=\dfrac{7}{4}-4y_1-y_2-3y_3} \end{cases}.$$
$$Z=9y_1+12y_2+16y_3$$

Змінну y_3 переводимо в основні. Оскільки $y_3=\min\left\{\frac{4}{4};\frac{5}{5};\frac{7}{12}\right\}=\frac{7}{12}$ переконуємось, що змінна y_6 переходить в неосновні.

II крок. *Основні змінні* $y_3; y_4, y_5$; *неосновні змінні* $y_1; y_2, y_6$.

$$\begin{cases} y_3 = \frac{7}{12} - \frac{4}{3}y_1 - \frac{1}{3}y_2 - \frac{1}{3}y_6 \\ y_4 = \frac{5}{3} + \frac{7}{3}y_1 - \frac{2}{3}y_2 + \frac{4}{3}y_6 \\ \boxed{y_5 = \frac{25}{12} + \frac{14}{3}y_1 - \frac{4}{3}y_2 + \frac{5}{3}y_6} \end{cases}$$

$$Z = \frac{28}{3} - \frac{37}{3}y_1 + \frac{20}{3}y_2 - \frac{16}{3}y_6.$$

Змінну y_2 переводимо в основні.

Оскільки $y_2 = \min\left\{\left(\frac{7}{12} : \frac{1}{3}\right); \left(\frac{5}{3} : \frac{2}{3}\right); \left(\frac{25}{12} : \frac{4}{3}\right)\right\} = \left\{\frac{21}{12}; \frac{5}{2}; \frac{25}{16}\right\} = \frac{25}{16}$, то в неосновні змінні переходить змінна y_5.

III крок. *Основні змінні* y_2, y_3, y_4; *неосновні змінні* y_1, y_5, y_6. Маємо:

$$\begin{cases} y_2 = \frac{25}{16} + \frac{7}{2}y_1 + \frac{5}{4}y_6 - \frac{3}{4}y_5 \\ \boxed{y_3 = \frac{1}{16} - \frac{5}{2}y_1 - \frac{3}{4}y_6 + \frac{1}{4}y_5} \\ y_4 = \frac{5}{8} + \frac{1}{2}y_5 - \frac{7}{6}y_6 \end{cases}$$

$$Z = \frac{79}{4} + 11y_1 - 5y_5 + 3y_6.$$

Змінну y_1 переведемо в основні.

Оскільки $y_1 = \min\left\{\infty, \frac{1}{16} : \frac{5}{2}; \infty\right\} = \frac{1}{40}$, то в неосновні змінні переходить y_3.

IV крок. Основні змінні y_1,y_2, y_4; неосновні змінні y_3, y_5, y_6.
Маємо:

$$\begin{cases} y_1 = \dfrac{1}{40} - \dfrac{2}{5}y_3 + \dfrac{1}{10}y_5 - \dfrac{3}{10}y_6 \\ y_2 = \dfrac{33}{20} - \dfrac{7}{5}y_3 - \dfrac{2}{5}y_5 + \dfrac{1}{5}y_6 \\ y_4 = \dfrac{5}{8} + \dfrac{1}{2}y_5 - \dfrac{7}{6}y_6 \end{cases}$$

$$Z = \frac{801}{40} - \frac{22}{5}y_3 - \frac{39}{10}y_5 - \frac{3}{10}y_6.$$

Критерій оптимальності досягнутий. Оптимальний розв'язок має вигляд $\left(\frac{1}{40}; \frac{33}{20}; \ 0; \frac{5}{8}; 0; \ 0\right)$. При цьому $Z_{maz} = \frac{801}{40} = 20{,}025$.

Запишемо співвідношення між змінними прямої і двоїстої задачі:

$$\begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ y_4 & y_5 & y_6 & y_1 & y_2 & y_3, \end{matrix}$$

Виразимо цільову функцію *Z*, одержану на IV кроці розв'язування через змінні y_j, маємо:

$$Z = \frac{801}{40} + 0 \cdot y_1 + 0 \cdot y_2 - \frac{22}{5}y_3 + 0 \cdot y_4 - \frac{39}{10}y_5 - \frac{3}{10}y_6.$$

Отже, оптимальний розв'язок прямої задачі має вигляд $\left(0; \ \frac{39}{10}; \frac{3}{10}; \ 0; \ 0; \ \frac{22}{5}\right)$. Ураховуючи, що $F_{\min} = Z_{\max}$, одержимо $F_{\min} = 20{,}025$, що співпадає з результатом одержаним у § 3.

Приклад 5. Скласти і розв'язати симплексним методом задачу, двоїсту задачі, наведеній у прикладі 10 §2.

Геометричне розв'язування прямої задачі показало, що при заданій системі обмежень:

$$\begin{cases} -3x_1 + 4x_2 \le 12 \\ 2x_1 - 3x_2 \le 6 \end{cases}$$
$$x_1 \ge 0, \ x_2 \ge 0$$

цільова функція $F = x_1 + 3x_2$ — необмежена, тобто $F_{\max} = \infty$.

Складемо задачу, двоїсту заданій. Запишемо матриці B і B'.

$$B = \left(\begin{array}{cc|c} -3 & 4 & 12 \\ 2 & -3 & 6 \\ \hline 1 & 3 & F \end{array}\right), \quad B' = \left(\begin{array}{cc|c} -3 & 2 & 1 \\ 4 & -3 & 3 \\ \hline 12 & 6 & Z \end{array}\right).$$

У двоїстій задачі будемо шукати мінімум функції $Z = 12y_1 + 6y_2$ при обмеженнях

$$\begin{cases} -3y_1 + 2y_2 \geq 1 \\ 4y_1 - 3y_2 \geq 3 \end{cases}$$
$$y_1 \geq 0,\ y_2 \geq 0.$$

Для зведення системи нерівностей до системи рівнянь введемо додаткові невід'ємні змінні y_3 і y_4 і вважатимемо їх основними на 1-му кроці розв'язування.

I крок. *Основні змінні* y_3, y_4; *неосновні змінні* $y_1; y_2$. Маємо:

$$\begin{cases} y_3 = -1 - 3y_1 + 2y_2 \\ \boxed{y_4 = -3 + 4y_1 - 3y_2\ .} \end{cases}$$

Змінну y_1 переведемо в основну. Оскільки $y_1 = \min\left\{\infty; \dfrac{3}{4}\right\} = \dfrac{3}{4}$, то змінна y_4 переходить в неосновні.

II крок. *Основні змінні* $y_1; y_3$; *неосновні змінні* y_2, y_4. Маємо:

$$\begin{cases} y_1 = \dfrac{3}{4} + \dfrac{3}{4}y_2 + \dfrac{1}{4}y_4 \\ y_3 = -\dfrac{13}{4} - \dfrac{1}{4}y_2 - \dfrac{3}{4}y_4\,. \end{cases}$$

У другому рівнянні одержаної системи від'ємні і коефіцієнти при всіх неосновних змінних і вільний член, що підтверджує несумісність системи обмежень двоїстої задачі. Це, у свою чергу, підтверджує справедливість другої частини теореми 1.

Особливі випадки (об'єктивно обумовлені оцінки). У попередньому пункті кожній додатковій змінній x_{n+i} прямої задачі ставиться у відповідність змінна y_i $\left(i=1,2,...,m\right)$ двоїстої задачі, а сам аналіз розглянутих прикладів дозволяє зробити такі висновки:

Якщо і-та компонента оптимального розв'язку двоїстої задачі обертається в нуль (додатна), то відповідна додаткова змінна в оптимальному розв'язку прямої задачі додатна (дорівнює нулю) і при підстановці в і-ту нерівність системи обмежень прямої задачі компонент оптимального розв'язку ця нерівність не обертається (обертається) в строгу рівність.

Зауваження. Якщо оптимальний розв'язок прямої задачі є виродженим, тоді вказаний порядок у відповідних змінних дещо порушується, а саме: одна із додаткових змінних прямої задачі дорівнює нулю згідно з виродженням, хоча відповідна їй змінна двоїстої задачі також дорівнює нулю. Перші m компоненти оптимального розв'язку двоїстої задачі називаються *об'єктивно обумовленими* оцінками. Така назва пов'язана з *економічним змістом* змінних двоїстої задачі.

Нехай, наприклад, прямою є задача про використання ресурсів, об'єми яких задані, то об'єктивно обумовлені оцінки двоїстої задачі відображають значимість кожного виду цих ресурсів для досягнення мети, поставленої у прямій задачі: одержання максимуму прибутку або продукції, що виробляється. Оцінки є своєрідним розрахунковим об'єктивно обумовленим оптимальним планом цін за одиницю кожного виду ресурсів.

Розглянемо задачу про використання сировини, розв'язану симплексним методом у §3. Приймемо її за вихідну і складемо двоїсту задачу.

Випишемо матрицю умов вихідної задачі B і матрицю B', транспоновану до неї.

$$B=\left(\begin{array}{cc|c} 0 & 5 & 100 \\ 2 & 0 & 200 \\ 4 & 4 & 100 \\ 2 & 2 & 80 \\ \hline 4 & 5 & F \end{array}\right),\quad B'=\left(\begin{array}{cccc|c} 0 & 2 & 4 & 2 & 4 \\ 5 & 0 & 4 & 2 & 5 \\ \hline 100 & 200 & 100 & 80 & Z \end{array}\right).$$

У двоїстій задачі потрібно знайти мінімум функції
$Z=100y_1+200y_2+100y_3+80y_4$ при обмеженнях:

$$\begin{cases}2y_2+4y_3+2y_4\ge 4\\ 5y_1+4y_3+2y_4\ge 5\end{cases}$$
$$y_i\ge 0\ \left(i=1,2,3,4\right).$$

Припустимо, що підприємство вирішило не виробляти продукцію, а продати свою сировину-деревину, маючи на увазі при цьому отримати виручку, не меншу максимального прибутку від реалізації самої продукції. За якими цінами в цьому випадку йому необхідно продавати кожний вид деревини?

Ці ціни визначаються оптимальним планом $\left(5;\ 20;\ 0;\ 190;\ 0;\ 30\right)$. Розглянемо розмірність математичних моделей вихідної і двоїстої задач.

Пряма задача:

$$\begin{pmatrix}\text{Витрати } i\text{-го виду}\\ \text{деревини на виробниц-}\\ \text{тво одного виробу } A\end{pmatrix}\cdot x_1+\begin{pmatrix}\text{Витрати } i\text{-го виду}\\ \text{деревини на виробниц-}\\ \text{тво одного виробу } B\end{pmatrix}\cdot x_2\le\begin{pmatrix}\text{Запаси } i\text{-го}\\ \text{виду деревини}\end{pmatrix}$$

$$\begin{matrix}\text{Сумарний}\\ \text{прибуток}\end{matrix}\quad F=\begin{pmatrix}\text{Прибуток від}\\ \text{реалізації одного}\\ \text{виробу } A\end{pmatrix}\cdot x_1+\begin{pmatrix}\text{Прибуток від}\\ \text{реалізації одного}\\ \text{виробу } B\end{pmatrix}\cdot x_2$$

Двоїста задача:

$$\sum_{i=1}^{4}\begin{pmatrix}\text{Витрати деревини}\\ i\text{-го виду на виробництво}\\ \text{одного виробу } A\end{pmatrix}\cdot y_i\ge\begin{pmatrix}\text{Прибуток від}\\ \text{реалізації одного}\\ \text{виробу } A\end{pmatrix}$$

$$\sum_{i=1}^{4}\begin{pmatrix}\text{Витрати деревини}\\ i\text{-го виду на виробництво}\\ \text{одного виробу } B\end{pmatrix}\cdot y_i\ge\begin{pmatrix}\text{Прибуток від}\\ \text{реалізації одного}\\ \text{виробу } B\end{pmatrix}$$

$$\text{Загальна вартість сировини}\ \ Z=\sum_{i=1}^{4}\begin{pmatrix}\text{Запаси деревини}\\ i\text{-го виду}\end{pmatrix}\cdot y_i$$

Із економічного змісту обох задач випливає, що y_i — умовна ціна одиниці деревини *i*-го виду $\left(i=1,2,3,4\right)$. Ці умовні ціни y_i — об'єктивно обумовлені оцінки — знаходять шляхом розв'язування двоїстої задачі, тобто вони є компонентами оптимального розв'язку двоїстої задачі.

Повернімося до попередньо сформульованої задачі і знайдемо її розв'язок. Шляхом введення додаткових змінних систему обмежень-нерівностей двоїстої задачі зведемо до системи рівнянь:

$$\begin{cases} 2y_2 + 4y_3 + 2y_4 - y_5 = 4 \\ 5y_1 + 4y_3 + 2y_4 - y_6 = 5 \end{cases}$$
$$y_i \geq 0 \ \left(i = 1, 2, ..., 6\right).$$

У даній задачі за основні змінні на І-му кроці доцільно вибрати не додаткові змінні, а змінні y_1 і y_2, оскільки базисний розв'язок І кроку опиниться у такому випадку допустимим.

І крок. *Основні змінні* y_1, y_2; *неосновні змінні* $y_3; y_4, y_5, y_6$. Маємо:

$$\begin{cases} y_1 = 1 - \frac{4}{5}y_3 - \frac{2}{5}y_4 + \frac{1}{5}y_6 \\ \boxed{y_2 = 2 - 2y_3 - y_4 + \frac{1}{2}y_5} \end{cases}$$
$$Z = 500 - 380y_3 - 160y_4 + 100y_5 + 20y_6.$$

Змінну y_3 переводимо в основну.

Оскільки $y_3 = \min\left\{\left(1 : \frac{4}{5}\right);\ 2 : 2\right\} = 1$, то змінна y_2 переходить в неосновні.

ІІ крок. *Основні змінні* $y_1; y_3$; *неосновні змінні* y_2, y_4, y_5, y_6.

$$\begin{cases} \boxed{y_1 = \frac{1}{5} + \frac{2}{5}y_2 - \frac{1}{5}y_5 + \frac{1}{5}y_6} \\ y_3 = 1 - \frac{1}{2}y_2 - \frac{1}{2}y_4 + \frac{1}{4}y_5 \end{cases}$$
$$Z = 120 + 190y_2 + 30y_4 + 5y_5 + 20y_6.$$

Критерій оптимальності виконаний. Оптимальний розв'язок $\left(\frac{1}{5};\ 0;\ 1;\ 0;\ 0;\ 0\right)$, $Z_{\min} = 120$.

Для більшого узагальнення спробуємо змінити умову попередньої задачі, переставивши місцем деякі цифри і проаналізуємо отриманий результат. Наприклад, матриця B вихідної задачі і матриця B', транспонованої задачі мають вигляд:

$$B = \left(\begin{array}{cc|c} 0 & 4 & 120 \\ 4 & 0 & 160 \\ 2 & 2 & 120 \\ 1 & 2 & 80 \\ \hline 2 & 3 & F \end{array}\right), \quad B' = \left(\begin{array}{cccc|c} 0 & 4 & 2 & 1 & 2 \\ 4 & 0 & 2 & 2 & 3 \\ \hline 120 & 160 & 120 & 80 & Z \end{array}\right).$$

Тобто вихідна задача зводиться до знаходження максимуму функції $F = 2x_1 + 3x_2$ за обмеженнями:

$$\begin{cases} 4x_2 \le 120 \\ 4x_1 \le 160 \\ 2x_1 + 2x_2 \le 120 \\ x_1 + 2x_2 \le 80 \end{cases}$$
$$x_1 \ge 0, \; x_2 \ge 0.$$

Неважко переконатися, що оптимальним розв’язком є (40; 20; 40; 0; 0; 0), за яким $F_{\max} = 140$.

У двоїстій задачі необхідно знайти мінімум функції $Z = 120y_1 + 160y_2 + 120y_3 + 80y_4$ за обмеженнями:

$$\begin{cases} 4y_2 + 2y_3 + y_4 \ge 2 \\ 4y_1 + 2y_3 + 2y_4 \ge 3 \end{cases}$$
$$y_i \ge 0 \; \left(i = 1,2,3,4\right).$$

Оптимальний розв’язок двоїстої задачі на III кроці прийме вигляд (перевірте самостійно)

$$\begin{cases} y_3 = \dfrac{1}{2} + 2y_1 - 4y_2 + y_5 - \dfrac{1}{2}y_6 \\ y_4 = 1 - 4y_1 + 4y_2 - y_5 + y_6 \end{cases}$$
$$Z = 140 + 40y_1 + 40y_5 + 20y_6.$$

Критерій оптимальності виконаний. Оптимальний розв’язок $\left(0;\ 0;\ \dfrac{1}{2};\ 1;\ 0;\ 0\right)$ і $Z_{\min} = 140$.

Характерним тут є те, що на останньому кроці розв'язування у виразі цільової функції відсутня змінна y_2. Отже, одержаний оптимальний розв'язок в двоїстій задачі не є єдиним.

Співвідношення між змінними в прямій і двоїстій задачах мають вигляд:

$$\begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ y_5 & y_6 & y_1 & y_2 & y_3 & y_4. \end{array}$$

Компоненти y_1, y_2, y_3, y_4 в оптимальному розв'язку двоїстої задачі оцінюють додаткові змінні x_3, x_4; x_5, x_6 в оптимальному розв'язку прямої задачі.

Рівність нулю змінної y_1 в одержаному оптимальному розв'язку означає, що в оптимальному розв'язку прямої (вихідної) задачі змінна x_3 додатна, тобто підстановка компонент оптимального розв'язку у першу нерівність системи обмежень прямої задачі не обертає її в строгу рівність.

Оскільки $y_2 = 0$, то змінна x_4 також повинна була б стати додатною. Але ця змінна в оптимальному розв'язку прямої задачі дорівнює нулю за рахунок *виродження* (змінна y_2 на останньому кроці у виразі цільової функції відсутня), на що було вказано у відповідному зауваженні. Тим самим підстановка компонент оптимального розв'язку в другу нерівність системи обмежень прямої задачі обертає його в тотожну рівність.

Оскільки змінні y_3 і y_4 додатні, а відповідні їм змінні оптимального розв'язку x_5 і x_6 прямої задачі дорівнюють нулю, то підстановка компонент оптимального розв'язку в третю і четверту нерівність прямої задачі також перетворюють ці нерівності в тотожні рівності.

Економічний зміст такого результату полягає в тому, що деревина I і II видів є з надлишками (надлишку деревини II виду немає за рахунок виродження) і ми не дуже дорожимо нею, встановлюємо для неї нульові ціни, а деревина III і IV видів є дефіцитною, вона повністю витрачається, а тому ми оцінюємо їх певними додатними оцінками $\left(y_3 = \frac{1}{2}; y_4 = 1 \right)$.

Зауваження. Якщо оптимальний розв'язок однієї із задач — вироджений, то в другій задачі порушується єдиність оптимального розв'язку.

ЗАДАЧІ

4.31. Скласти і розв'язати симплексним методом задачу, двоїсту наступній: мінімізувати функцію $F = 3x_1 + 2x_2$ за обмеженнями:

$$\begin{cases} 7x_1 + 2x_2 \ge 14 \\ 4x_1 + 5x_2 \ge 20 \end{cases}$$
$$x_1 \ge 0,\ x_2 \ge 0 \quad .$$

Відповідь: $F_{\min} = \dfrac{86}{9};\ \left(\dfrac{10}{9};\dfrac{28}{9}\right);\ Z_{\max} = \dfrac{86}{9};\ \left(\dfrac{7}{27};\dfrac{8}{27}\right).$

4.32. Симплексним методом знайти максимум функції $F = x_1 + x_2 + x_3$ за обмеженнями:

$$\begin{cases} 2x_1 + x_2 + 2x_3 \le 2 \\ 4x_1 + 2x_2 + x_3 \le 2 \end{cases}$$
$$x_1 \ge 0,\ x_2 \ge 0,\ x_3 \ge 0.$$

Скласти двоїсту задачу і розв'язати її симплексним методом.

Відповідь: $Z_{\min} = F_{\max} = \dfrac{4}{3}$.

4.33. Вихідна задача (I): знайти невід'ємні значення $\left(x_1, x_2\right)$, які максимізують лінійну функцію $F = 5x_1 + 4x_2$. За системою обмежень $4x_1 + 3x_2 \le 24,\ 3x_1 + 4x_2 \le 24$. Скласти двоїсту задачу і розв'язати її.

Відповідь: $F_{\max} = \dfrac{216}{7};\ \left(\dfrac{24}{7};\dfrac{24}{7}\right);\ Z_{\min} = \dfrac{216}{17};\ \left(\dfrac{8}{7};\dfrac{1}{7}\right).$

4.34. Знайти максимум функції $F = x_1 + x_2$ за обмеженнями:

$$\begin{cases} x_1 - x_2 \le 1 \\ -x_1 + 2x_2 \ge 1 \\ -2x_1 + x_2 \le 1 \\ x_2 \le 4 \end{cases}$$
$$x_1 \ge 0,\ x_2 \ge 0 \quad .$$

Розв'язати задачу симплексним методом, потім скласти двоїсту задачу і розв'язати її. Продемонструвати її геометричний розв'язок.

Відповідь: $F_{\max} = Z_{\min} = 9$.

4.35. Вихідна задача (I): знайти невід'ємні значення (x_1, x_2), мінімізуючі лінійну функцію $F = 3x_1 + 3x_2$. За системою обмежень $5x_1 - 4x_2 \ge -2,\ x_1 + 2x_2 \ge 6$. Скласти двоїсту задачу і розв'язати її.

Відповідь: $F_{\min} = \dfrac{78}{7};\ \left(\dfrac{10}{7};\dfrac{16}{7}\right);\ Z_{\max} = \dfrac{78}{7};\ \left(\dfrac{27}{14};\dfrac{3}{14}\right).$

4.36. Знайти мінімум функції $F = 2x_1 + 4x_2$ за обмеженнями:

$$\begin{cases} x_1 + 2x_2 \ge 10 \\ 2x_1 + x_2 \ge 10 \\ x_2 \ge 4 \end{cases}$$
$$x_1 \ge 0,\ x_2 \ge 0.$$

Розв'язати задачу геометрично, потім скласти двоїсту задачу і розв'язати її симплексним методом.

Відповідь: $F_{\min} = Z_{\max} = 22$.

4.37. Скласти і розв'язати симплексним методом задачу, двоїсту такій: максимізувати функцію $F = x_1 + 3x_2$ та за обмеженнями:

$$\begin{cases} x_1 + x_2 \ge 4 \\ -x_1 - x_2 \ge -2 \\ -x_1 \ge -2 \end{cases}$$
$$x_1 \ge 0,\ x_2 \ge 0\ \ .$$

Відповідь: $Z_{\min} = -\infty$.

§5. Транспортна задача

Однією з типових задач лінійного програмування є так звана *транспортна задача*. Специфіка економіко-математичної моделі транспортної задачі обумовлює виділення її окремим параграфом, оскільки вказана задача розв'язується не універсальним симплексним методом, а за допомогою так званого розподільчого методу та його різних модифікацій. Ця задача виникає при плануванні найбільш раціональних перевезень вантажу. В одних випадках це означає визначення такого плану перевезень, за яким вартість останніх була б мінімальною, а в інших більш важливим є виграш за часом. Іноді першу задачу називають *транспортною задачею за критерієм вартості*, а другу — *транспортною задачею за критерієм часу*. Перша задача є частинним випадком задачі лінійного програмування і може бути розв'язана симплексним методом, але в силу певних особливостей її можна — розв'язати простіше.

Економіко-математична модель транспортної задачі. Розглянемо задачу про перевезення деякого однорідного вантажу із пункту відправлення (від постачальника) у пункт призначення (до споживачів) при забезпеченні мінімальних витрат на перевезення.

Як правило, початкові умови таких задач записуються в таблицю. Наприклад, для n постачальників і m споживачів маємо таблицю 4.5.

У загальному плані розв'язування таких задач розбивається на два етапи:

1) визначення початкового (вихідного) опорного розв'язку, тобто знаходиться *початковий розподіл поставок.*

2) побудова послідовних ітерацій, тобто наближення до оптимального розв'язку, виконуючи так званий перерозподіл поставок.

Нехай ми маємо таблицю 4.5 вихідних даних задачі. Вихідний опорний розв'язок можна будувати або за *правилом урахування найменших витрат* або за правилом так званого *"північно-західного кута"*.

Таблиця 4.5

Постачальники	**Потужності постачальників**	**Споживачі та їх попит**					
		1	2	...	j	...	m
		S_1	S_2	...	S_j	...	S_m
1	P_1	a_{11} x_{11}	a_{12} x_{12}	...	a_{1j} x_{1j}	...	a_{1m} x_{1m}
2	P_2	a_{21} x_{21}	a_{22} x_{22}	...	a_{2j} x_{2j}	...	a_{2m} x_{2m}
...	...	...	...	...	...	...	...
i	P_i	a_{i1} x_{i1}	a_{i2} x_{i2}	...	a_{ij} x_{ij}	...	a_{im} x_{im}
...	...	...	...	...	...	...	...
n	P_n	a_{n1} x_{n1}	a_{n2} x_{n2}	...	a_{nj} x_{nj}	...	a_{nm} x_{nm}

Тут через a_{ij} позначаються витрати на перевезення одиниці вантажу від i-го постачальника ($i = 1, 2,..., n$) до j-го споживача ($j = 1, 2,..., m$), x_{ij} — кількість вантажу (поставка), яка планується для перевезень від i-го постачальника до j-го споживача, P_i — потужність i-го постачальника на запланований період, S_j — попит j-го споживача на цей же період.

Задача зводиться до знаходження мінімуму цільової функції, яка виражає сумарні витрати на перевезення всього вантажу, тобто функції

$$F = a_{11}x_{11} + a_{12}x_{12} + ... + a_{ij}x_{ij} + ... + a_{nm}x_{nm},$$

за такими обмеженнями:

$$\begin{cases} x_{11} + x_{12} + ... + x_{1m} = P_1 \\ x_{21} + x_{22} + ... + x_{2m} = P_2 \\ .... \\ x_{n1} + x_{n2} + ... + x_{nm} = P_n \\ x_{11} + x_{21} + ... + x_{n1} = S_1 \\ x_{12} + x_{22} + ... + x_{n2} = S_2 \\ .... \\ x_{1m} + x_{2m} + ... + x_{nm} = S_m. \end{cases} \quad (1)$$

Задача, в якій існують лише вказані обмеження, називається *відкритою моделлю*.

Якщо до обмежень (1) додати ще одне:

$$\sum_{i=1}^{n} P_i = \sum_{j=1}^{m} S_j, \quad (2)$$

тобто сумарна потужність постачальників дорівнює сумарному попиту споживачів, то відповідну модель задачі називають *закритою*.

Особливостями економіко-математичної моделі транспортної задачі порівняно з моделями задач лінійного програмування є:

1) оскільки система обмежень (1) має вигляд рівнянь, то відпадає необхідність вводити додаткові змінні.

2) у системі (1) матриця коефіцієнтів при змінних складається лише із одиниць і нулів.

3) система обмежень (1) містить n рівнянь, які зв'язують поставки i-го постачальника з потужністю P_i ($i = 1, 2,..., n$) цього постачальника і m рівнянь, які зв'язують поставки j-му споживачу з попитом

S_j (j = 1, 2,..., m) цього споживача, крім того, n — дорівнює числу рядків вихідної таблиці, а число m — числу стовпців.

4) число змінних x_{ij}, які входять в цільову функцію і в систему рівнянь (1) дорівнює добутку nm, тобто числу клітинок таблиці 4.5.

5) система обмежень (1) є система із $m + n$ рівнянь із mn змінними.

Специфічність економіко-математичної моделі транспортної задачі призвела до появи особливого методу її розв'язування — розподільчого методу. Існують також різні модифікації цього методу, проте всі теоретичні передумови, які лежать в основі симплексного методу збережені.

Будь-який розв'язок транспортної задачі $\left(x_{11};x_{12};...;x_{nm}\right)$ *називається розподілом поставок*. Оскільки поставки не можуть бути від'ємними, то мову можливо вести лише про допустимі базисні розв'язки системи обмежень (1).

Оптимальному розв'язку транспортної задачі, за якого цільова функція $F = a_{11}x_{11} + a_{12}x_{12} + ... + a_{nm}x_{nm}$ стає мінімальною, відповідає оптимальний розподіл поставки, який вважатимемо єдиним, хоча в деяких випадках єдиність оптимального розв'язку може порушуватись.

Якщо розв'язання задач симплексним методом означало перехід від одного базисного розв'язку до іншого, до тих пір поки не було знайдено оптимального розв'язку, або поки не було зроблено висновку про несумісність системи обмежень, то розв'язування транспортної задачі полягає у послідовному переході від одного розподілу поставок до іншого розподілу: від однієї розрахункової таблиці, яка відповідає певному розв'язку, до іншої. Оскільки кожна така таблиця містить певний розподіл поставок, які відповідають базисному розв'язку, то і клітинки таблиці мають відповідати основним (додатнім) і неосновним (рівними нулеві) змінним.

Випадок, коли відповідний базисний розв'язок стане виродженим, розглядається окремо.

Новий розподіл поставок має знижувати, або хоча б не збільшувати загальний кошторис витрат на перевезення. Перерозподіл поставок повинен здійснюватися, починаючи з вихідного (початкового), доки не буде знайдено оптимального розподілу поставок.

Зауваження. У клітинки, що відповідають основним змінним, записують поставки, а — неосновним змінним, тобто рівним нулю, залишають незаповненими (вільними). У верхній лівий кут запишемо показник витрат (у подальших таблицях — оцінка), а у нижньому правому — поставки.

Число заповнених клітинок визначається числом основних змінних системи обмежень (1), а останнє дорівнює числу лінійно незалежних рівнянь системи.

Початковий розподіл поставок. Розглянемо спочатку задачі, які мають закриту модель, тобто для них виконується обмеження (2): $\sum_{i=1}^{n} P_i = \sum_{j=1}^{m} S_j$. У цьому випадку система обмежень (1) містить $m + n$ рівнянь, але число її лінійно незалежних рівнянь на одиницю менше, оскільки, якщо скласти n перших рівнянь віднесених до постачальників і m рівнянь, віднесених до споживачів з урахуванням (2), одержимо рівність лівих і правих частин складених таким чином сум. Це свідчить про те, що система (1) у закритих моделях лінійно залежна. Якщо із системи виключити одне, будь-яке рівняння, то вона стане лінійно незалежною.

Отже, число лінійно незалежних рівнянь і число заповнюючих клітинок дорівнює $m + n - 1$. При цьому число незаповнених клітинок дорівнює $mn - (m + n - 1)$.

Як заповнити поставками $m + n - 1$ клітинок таблиці? Існують різні правила такого заповнення, а отже, і одержання початкового розподілу поставок. Розглянемо два з них: правило врахування найменших витрат і правило "північно-західного кута".

Нехай потрібно одержати початковий розподіл поставок у наступній транспортній задачі (таблиця 4.6).

Таблиця 4.6

Постачальники	**Потужність постачальників**	**Споживачі та їх попит**				
		1	2	3	4	5
		90	70	85	80	75
I	110	5	3	4	2	3
II	130	7	6	4	5	4
III	80	2	3	6	7	6
IV	80	7	5	6	3	4

Оскільки $\sum_{i=1}^{4} P_i = 400,\ \sum_{j=1}^{5} S_j = 400$, тобто $\sum_{i=1}^{4} P_i = \sum_{j=1}^{5} S_j$, то задача має закриту модель. Тут число постачальників $n = 4$; число споживачів $m = 5$; число всіх клітинок $4 \cdot 5 = 20$; число заповнених клітинок $4 + 5 - 1 = 8$; число вільних клітинок $20 - 8 = 12$.

Правило врахування найменших витрат. Дамо кожній клітині подвійний номер, співпадаючий з індексами відповідній цій клітині змінній (перший індекс співпадає з номером постачальника, другий — з номером споживача). Наприклад, клітина (3.4) стоїть на перетині 3-го рядка і 4-го стовпця. У таблиці 4.6 беремо клітини, які мають найменший показник витрат, такий показник дорівнює 2, який знаходиться в двох клітинах: (1.4) і (3.1). Дамо поставку в клітину (1.4). Щоб розв'язати питання про величину цієї поставки, зазначимо, що постачальник I має 110 од. вантажу, а споживачу 4 потрібно 80 од. вантажу. Розмір поставки визначається мінімальною із цих двох чисел, тобто в клітину (1.4) слід записати поставку, рівну 80 од. Споживач 4, одержав повністю 80 од. вантажу, може бути викреслений із таблиці 4.7 і на наступному етапі розподілу таблиця 4.7 вже має розмір 4×4, причому потужність постачальника I залишається 30 од. вантажу.

Таблиця 4.7

	90	**70**	**85**	**80**	**75**
110	5	3 30	4	2 80	3
130	7	6	4 85	5	4 45
80	2 80	3	6	7	6
80	7 10	5 40	6	3	4 30

Заповнюємо клітину (3.1). Враховуючи, що min{80; 90} = 80, даємо в цю клітину поставку, рівну 80 од. При цьому заповнені із таблиці викреслюється 3-ій рядок, а попит споживача I скорочується до 10 од.

Наступний за величиною показник витрат дорівнює 3, стоїть в клітинах (1.2) і (1.5) [клітини (3.2) і (4.4) у рахунок не беруться, оскільки вони стоять у викресленому рядку і стовпці].

Оскільки у постачальника I залишилось лише 30 од. вантажу, передаємо цей вантаж у клітину (1.2) (min{30; 70} = 30) і викреслимо 1-ий рядок. Клітини (2.3), (2.5), (4.5) мають показник витрат, рівний 4 од.

Дамо поставку в клітину (2.3). Розмір цієї поставки дорівнює 85 од. (min{85; 130} = 85). Із таблиці викреслюємо третій стовпець, а потужність постачальника II скоротилась до 45 одиниць, які передаємо в клітину (2.5) і викреслимо 2-ий рядок. Нестачу у споживача 5 30 одиниць вантажу візьмемо у постачальника IV і викреслимо 5-ий стовпець.

Залишок у постачальника IV 50 одиниць вантажу можливо передати споживачам 1 або 2. Забезпечимо спочатку споживача 2, оскільки в клітині (4.2) стоїть менший коефіцієнт витрат, ніж в клітині (4.1).

Але споживачу 2 потрібні не всі 50 одиниць. вантажу, а тільки 40, а тому в клітину (4.2) записуємо 40 одиниць і викреслюємо 2-ий стовпець, а залишок в 10 одиниць вантажу помістимо до клітини (4.1). Після останньої дії із таблиці викреслені одночасно і 1-ий стовпець і 4-ий рядок. Зауважимо, що тільки на останньому етапі відбувається викреслення і рядка і стовпця, а на попередніх етапах передача поставки в яку-небудь клітину супроводжувалась викресленням або одного стовпця, або одного рядка. Після одержання розподілу та підрахунку кількості заповнених клітин, виявилось, що їх 8, тобто стільки, скільки повинно бути. Можливо також переконатися, що на кожному рядку і по кожному стовпцю дотримується баланс, від кожного постачальника запланована відправка всього вантажу, який він має і кожному споживачу планується та кількість вантажу, яку він потребує.

Проведений початковий розподіл поставок не ставить за мету отримати оптимальний розподіл. Для цього існують наступні етапи розв'язування задачі, які мають за мету перехід до нових розподілів поставок, поки не буде знайдений їх оптимальний розподіл.

Перерозподіл поставок. У таблиці 4.7 поряд з вісьмома заповненими клітинами є 12 вільних клітин. Не замислюючись над питаннями вигідно це чи ні, дамо поставку в одну із вільних клітинок, наприклад, в клітину (1.1). Очевидно це порушить баланс в 1-му ряду і

1-му стовпці, який можливо відновити, якщо поставки в клітинах (1.2) і (1.4) зменшити на величину поставок, що подана в клітину (1.1). При цьому порушується баланс у 2-му стовпці, що можливо відновити, якщо додати цю поставку до поставки клітини (4.2).

Отже, передача поставки в одну із вільних клітин веде до зміни поставок у деяких заповнених клітинах. Надалі будемо говорити, що відбувається перерозподіл поставок у циклі.

Циклом називається замкнутий многокутник, сторони якого є горизонтальними і вертикальними відрізками, одна вершина якого співпадає з вільною клітиною, для якої утворюється цикл, а всі інші — з заповненими клітинами.

Якщо розподіл у таблиці такий, що заповнено рівно $m + n - 1$ клітини, то для кожної вільної клітини можливо побудувати цикл, причому єдиний.

Як приклад, на рис.52 зображені лише цикли для вільних клітин (1.1), (1.3), (3.4) і (4.4) таблиці 4.7. До номеру клітин циклу додамо ще послідовну нумерацію клітин, починаючи з тої, для якої утворений цей цикл із зазначеними поряд з цими номерами величинами поставок, взятих із таблиці 4.7, байдуже в якому напрямку (за чи проти годинникової стрілки).

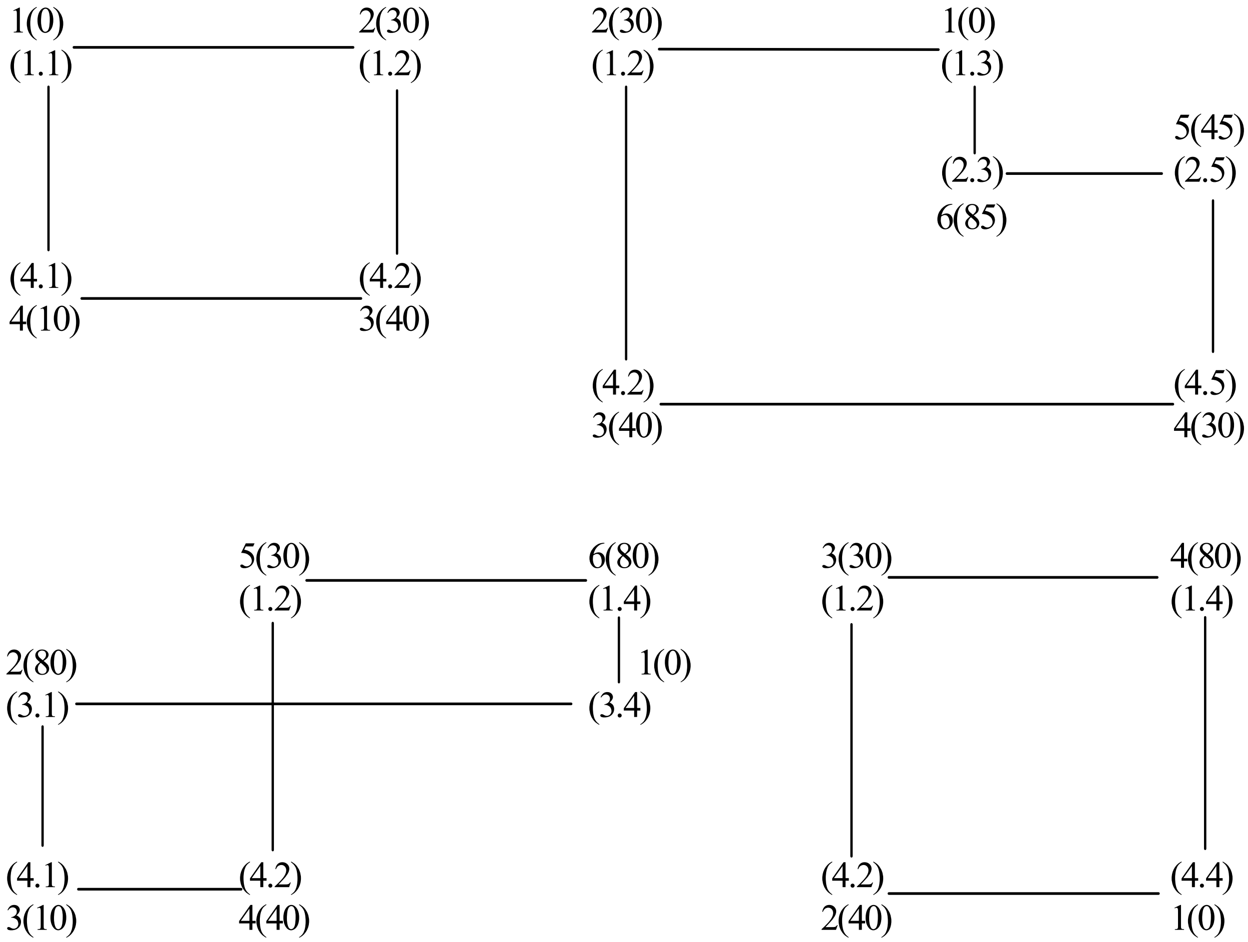


Рис. 52

При перерозподілі поставки по циклу вона додається до всіх поставок непарних клітинок (1, 3, 5 і т.д.) і вираховується із всіх поставок парних (2, 4, 6 і т.д.).

Оскільки число заповнених клітинок має зберігатися, то передача поставки у вільну клітину повинна супроводжуватися звільненням однієї із клітин циклу. Очевидно, що нею буде клітина з парним номером, яка має найменшу поставку. Це обумовлено тим, що всі поставки повинні бути невід'ємними.

У циклі, побудованому для клітини (1.1) мінімальна поставка парних клітин дорівнює 10 од. вантажу. Якщо цю поставку передати клітині (1.1), то в клітині (1.2) залишиться 20 од., в клітині (4.2) виявиться 50 од., а клітина (4.1) стане вільною.

Можна обчислити оцінку такого циклу t_{ij}, побудованого для будь-якої вільної клітини (i, j), яка дорівнює різниці між сумою показників витрат непарних та парних клітин циклу.

Наприклад:

$t_{11} = (5 + 5) - (3 + 7) = 0$ — оцінка циклу клітини (1.1);

$t_{13} = (4 + 4 + 5) - (4 + 4 + 3) = 2$ — оцінка циклу клітини (1.3);

$t_{44} = (3 + 3) - (2 + 5) = -1$ — оцінка циклу клітини (4.4);

$t_{34} = (7 + 7 + 3) - (2 + 5 + 2) = 8$ — оцінка циклу клітини (3.4).

Такі оцінки можуть бути як від'ємними t_{44}, додатними (t_{13}, t_{34}) так і дорівнювати нулю t_{11}.

Оскільки потрібно знайти мінімум функції F, то клітини, для яких оцінка циклу від'ємна є вигідними, а тому їх слід перевести в основні (заповнені). Передача поставки в таку клітину призведе до зменшення вартості витрат на величину, яка дорівнює добутку абсолютної величини оцінки на розмір поставки. Наприклад, перерозподіл поставок у циклі для клітини (4.4) знизило б витрати на $1{\cdot}40 = 40$ грош. од. Перерозподіл же поставок в циклі для клітини (1.3) збільшило б ці затрати на $2{\cdot}30 = 60$ грош. од. Передача поставки в клітину (1.1) не змінила б вартості витрат.

Отже, клітина, яка має від'ємну оцінку циклу є вигідною і її слід заповнювати, а додатню — невигідною. Про вільну клітину, яка має нульову оцінку, нічого сказати не можна.

Побудувавши цикли для всіх вільних клітин даного розподілу поставок і підрахувавши оцінки цих циклів, можна всі вільні клітини розбити на вигідні і невигідні. Серед вигідних клітин вибрати найвигіднішу і перерозподілити поставки в циклі для цієї клітини, у результаті перейти до нового розподілу поставок.

Зауваження. Найвигіднішою клітиною виявиться та, передача поставки в яку призведе до найбільшої економії, проте це не означає, що абсолютна величина від'ємної оцінки циклу цієї клітини повинна бути найбільшою, оскільки економія залежить також від розміру поставки, переміщуваної в циклі.

Зрозуміло, що така схема пошуку оптимального розв'язку досить громіздка, оскільки на кожному кроці розв'язування необхідно будувати цикли для всіх вільних клітин і підраховувати оцінки цих циклів, а тому з'явились різні модифікації розподільчого методу. Слід наголосити також, що цикли перерозподілу поставок у різних транспортних задачах можуть мати досить різну конфігурацію, а не обов'язково тому, як показано на рис. 52.

Знаходження оптимального розподілу поставок методом оцінки клітин. Ідея способу: кожній клітині ставиться у відповідність її оцінка рівна алгебраїчній сумі показника витрат клітини і відповідних чисел, записаних у додатковому рядку і стовпцю, поміщених знизу і справа від таблиці, які підбираються таким чином, щоб оцінки заповнених клітинок дорівнювали нулю. Оцінки вільних клітин можуть бути різних знаків і дорівнювати нулю. Встановлення оцінок для всіх заповнених і вільних клітин означає оцінку даного розподілу поставок.

Повернімося до розподілу поставок, наведених у таблиці 4.7, переписавши її з додатковим рядком і стовпчиком (таблиця 4.8).

Таблиця 4.8

5	3 30	4	2 80	3	–5
7	6	4 85	5	4 45	–7
2 80	3	6	7	6	–2
7 10	5 40	6	3	4 30	–7
0	+2	3	3	+3	

У 1-му стовпці додаткового рядка запишемо 0. Тоді для того, щоб оцінка в заповненій клітині (3.1) дорівнювала нулю, необхідно в 3-му рядку додаткового стовпця записати –2, оскільки в цьому випадку оцінка клітини (3.1) дорівнює (2 + 0 – 2) = 0. Щоб оцінку клітини (4.1) перетворити на нуль у 4-му рядку додаткового стовпця запишемо –7. У 5-му стовпці додаткового рядка запишемо число 3 для одержання нульової оцінки в клітині (4.5), в 2-му рядку додаткового стовпця для одержання нульової оцінки в клітині (2.5) запишемо число –7 і т. д., поки не заповнимо всі стовпці додаткового рядка і всі рядки додаткового стовпця.

Таблиця 4.9

0	0 70	2	0 40	1	–1
0	1	0 85	1	0 45	
0 80	3	7	8	7	
0 10	0	2	–1 40	0 30	
	+1		+1		

У нову таблицю 4.9 в кожну клітину замість показника витрат запишемо оцінку, яка для всіх заповнених клітин таблиці 4.9 дорівнює нулю. Всі вільні клітини, крім клітин (1.1), (2.1) і (4.4), мають додатні оцінки. Легко переконатися, що оцінка циклів, які були підраховані для циклів деяких клітин у попередньому параграфі збігається з оцінкою клітин, тим самим існує можливість обчислювати оцінки циклів, не будуючи самі цикли.

Про вигідність чи невигідність тієї чи іншої вільної клітини можливо судити за допомогою оцінки цієї клітини.

Оцінки, розташовані в таблиці 4.9, переконують, що початковий розподіл, проведений у таблиці 4.7 і 4.8, не є оптимальним, оскільки клітина (4.4) має від'ємну оцінку і її слід перевести в число заповнених. Цикл для цієї клітини був зображений на рис.52. Мінімальна поставка в парних клітинах цього циклу дорівнює 40 одиниць, яку слід перерозподілити таким чином: у клітину (4.4) запишемо цю поставку, із поставки (1.4) віднімемо, до поставки клітини (1.2) додамо, а клітина (4.2) залишиться вільною. В інших клітинах таблиці поставки не зміняться. Новий розподіл поставок запишемо в таблиці 4.9.

Надалі будемо добиватися, щоб в клітині (4.4) оцінка перетворилася на нуль. Очевидним є те, що до 4-го стовпця таблиці 4.9 додано 1, але ця дія псує нульову оцінку в клітині (1.4). Щоб цього не відбулося, віднімемо 1 із всіх оцінок 1-го рядка. Зберегти нульову оцінку в заповненій клітині (1.2) можливо шляхом додавання 1 до оцінок 2-го стовпця. Нові оцінки запишемо в таблиці 4.10.

Оцінки цієї таблиці такі, що одержаний розподіл не є оптимальним, але відмітимо, що вигідною на цьому кроці є клітина (1.1). Цикл розподілу для цієї клітини зображений на рис.53.

Таблиця 4.10

–1 10	0 70	1	0 30	0	
0	2	0 85	2	0 45	
0 80	4	7	9	7	–1
0	1	2	0 50	0 30	
1					

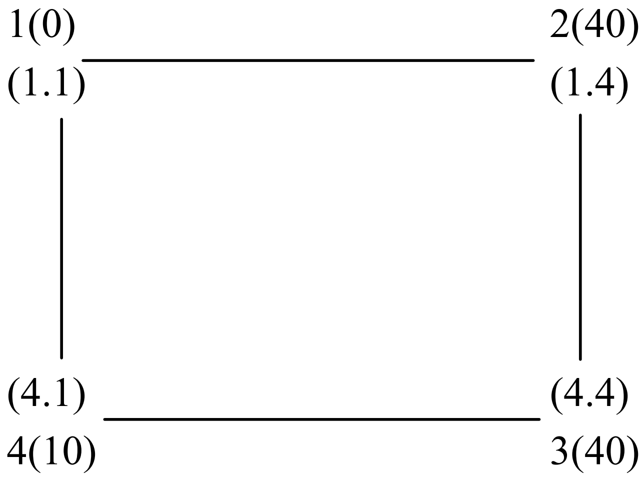


Рис. 53

Переконуємось, що в цикл для клітини (1.1) залучені інші клітини, ніж при початковому розподілі поставок.

Мінімальну поставку парних клітин, рівну 10 од., пересуваємо в циклі. Новий розподіл поставок запишемо в таблиці 4.10.

До оцінок 1-го стовпця таблиці 4.10 додамо 1, а із оцінок 3-го рядка віднімемо 1. Одержані оцінки розташуємо у таблиці 4.11.

Таблиця 4.11

0	0	1	0	0
1	2	0	2	0
0	3	6	8	6
1	1	2	0	0

Оскільки таблиця 4.11 вже не містить від'ємних оцінок, це є критерієм того, що розподіл поставок, одержаний у таблиці 4.10 є оптимальним.

Зауваження. Наявність нульової оцінки в незаповненій клітині (1.5) таблиці 4.11 є ознакою того, що одержаний оптимальний розподіл поставки не є єдиним. Наприклад, можна перерозподілити поставки в циклі для клітини (1.5) і одержати ще один оптимальний розподіл поставок, але вартість затрат при цьому не зменшиться.

На завершення слід обчислити мінімальні витрати на перевезення, для чого складемо підсумкову таблицю 4.12, в яку включимо показники витрат із таблиці 4.7 і оптимальний розподіл поставок із таблиці 4.10.

Таблиця 4.12

5 10	3 70	4	2 30	3
7	6	4 85	5	4 45
2 80	3	6	7	6
7	5	6	3 50	4 30

Дійсно, сумуючи добуток показника витрат на величину поставки для розподілу, приведеного в таблиці 4.12, одержимо:

$$F_{\min} = 50 + 210 + 60 + 340 + 180 + 160 + 150 + 120 = 1270 \text{ гр. одиниць.}$$

Порівняємо одержаний результат із витратами на перевезення при первісних розподілах поставок, наведених у таблиці 4.7.

$$F_1 = 90 + 160 + 340 + 180 + 160 + 70 + 200 + 120 = 1320 \text{ гр. одиниць.}$$

Отже, економія порівняно з початковим розподілом поставок складає $1320 - 1270 = 50$ гр. одиниць.

Аналогічний результат можна одержати і іншим шляхом.

На завершення вкажемо на друге правило початкового розподілу поставок.

Правило “північно-західного кута”. У цьому випадку не звертають увагу на показники витрат. Починаючи заповнення з клітини (1.1) — “північно-західного кута” таблиці, ступенями спускаються вниз до клітини (nm), уявно викреслюючи або один рядок, або один стовпець. На останньому кроці викреслюється останній (n-ий) рядок і останній (m-ий) стовпець. У таблиці 4.13 наведений початковий розподіл поставок за правилом “північно-західного кута” для задачі, що розглядається. Кількість заповнених клітин знову дорівнює 8. Легко переконатися, що початковий розподіл поставок у таблиці 4.13, виконаний за правилом “північно-західного кута”, значно гірший ніж розподіл поставок у таблиці 4.7, проведений з урахуванням найменших витрат.

Таблиця 4.13

	90	70	85	80	75
110	5 90	3 20	4	2	3
130	7	6 50	4 80	5	4
80	2	3	6 5	7 75	6
80	7	5	6	3 5	4 75

Дійсно:

$$F_2 = 450 + 60 + 300 + 320 + 30 + 525 + 15 + 300 = 2000 \text{ гр. од.},$$

тобто розподіл поставок у таблиці 4.13 значно далі знаходиться від оптимального ніж початковий розподіл поставок в таблиці 4.7 — 1320 гр. од.

Відкрита модель транспортної задачі. Випадок, коли сумарна потужність постачальників $\sum_{i=1}^{n} P_i$ не дорівнює сумарному попиту споживачів $\sum_{j=1}^{m} S_j$, тобто

$$\sum_{i=1}^{n} P_i \neq \sum_{j=1}^{m} S_j .$$

Можливі два випадки:

$$\text{а) } \sum_{i=1}^{n} P_i < \sum_{j=1}^{m} S_j$$

$$\text{б) } \sum_{i=1}^{n} P_i > \sum_{j=1}^{m} S_j \ .$$

В обох випадках модель замикають шляхом введення або фіктивного постачальника, або фіктивного споживача.

У випадку а) потужність фіктивного постачальника вважають рівною $\sum_{j=1}^{m} S_j - \sum_{i=1}^{n} P_i$. За показник витрат на перевезення від фіктивного постачальника можна вибрати будь-які числа, однакові по всьому рядку фіктивного постачальника. Але найпростіше всього вважати їх рівними нулю.

Після введення таким чином фіктивного постачальника модель задачі стає закритою і її можна розв'язувати способом оцінки клітинок.

Зауважимо, що в число заповнених клітин рівних $m+n-1$, в число постачальників необхідно включити і фіктивного постачальника. Після знаходження оптимального розподілу поставок, виявиться, що певна частина потреб одного або кількох споживачів не може бути задоволена за рахунок потужностей реальних постачальників.

У випадку б) вводиться фіктивний споживач, попит якого дорівнює $\sum_{i=1}^{n} P_i - \sum_{j=1}^{m} S_j$. Показниками витрат у стовпці фіктивного споживача можуть бути будь-які однакові числа, але краще такими показниками можуть бути нулі.

Приклад 1. Знайти оптимальний розподіл поставок згідно з таблицею 4.14.

Таблиця 4.14

Постачальники	**Потужність постачальників**	**Споживачі та їх попит**			
		1	**2**	**3**	**4**
		55	45	65	75
I	50	5	2	4	6
II	70	3	3	4	8
III	100	5	4	6	4

Тут $\sum_{i=1}^{3} P_i = 220$ одиниць, $\sum_{j=1}^{4} S_j = 240$ од.

Отже, попит споживачів на 20 од. перевищує обсяг потужностей постачальників. Введемо фіктивного постачальника, записуючи йому потужність, рівну 20 одиниць. Вихідну таблицю доповнюємо ще одним рядком — рядком фіктивного постачальника. У нову таблицю (4.15) відразу помістимо додатковий рядок і стовпець, щоб записати в них числа, які перетворюють оцінки заповнених клітин на нулі. У цій же таблиці проведений початковий розподіл поставок з урахуванням найменших витрат. Кількість заповнених клітин відповідає нормі: 4 + 4 – 1 = 7.

Таблиця 4.15

		55	45	65	75	
I	50	5	2 45	4 5	6	–3
II	70	3 10	3	4 60	8	–3
III	100	5 45	4	6	4 55	–5
Ф.	20	0	0	0	0 20	–1
		0	+1	–1	+1	

Таблиця 4.16

2	0 45	0 5	4	
0 30	1	0 40	6	
0 25	0	0	0 75	
–1	0	–2 20	0	+2

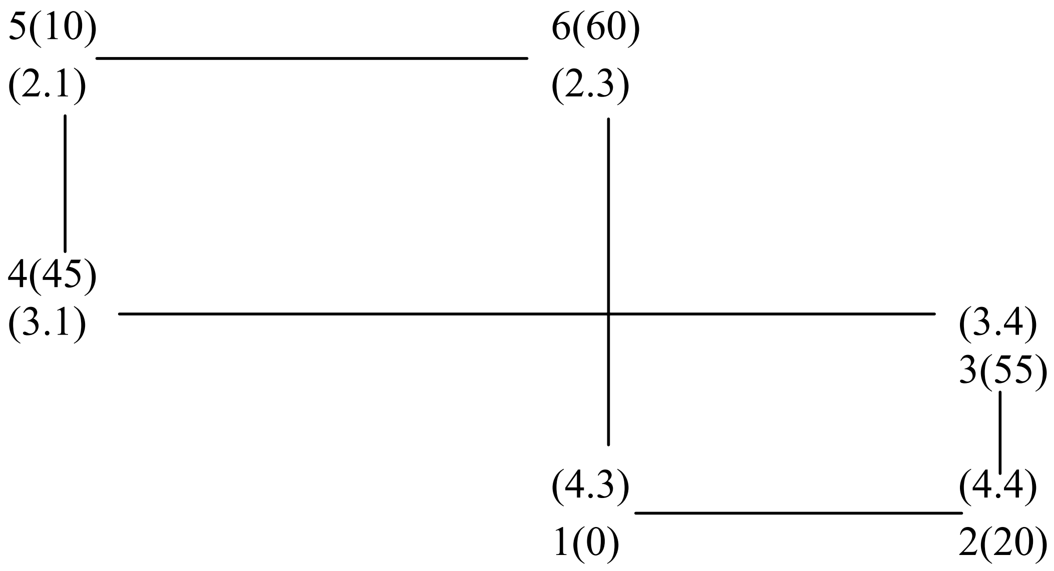


Рис.54

У таблиці 4.16 оцінюється початковий розподіл поставок. Згідно з таблицею 4.16 розподіл поставок у таблиці 4.15 не є оптимальним, оскільки дві клітини (4.1) і (4.3) мають від'ємні оцінки. Зазначені клітини є найвигіднішими і до них слід додати поставки. Візьмемо за правило надавати поставку насамперед в ту клітину, яка має найбільшу по абсолютній величині від'ємну оцінку (навіть якщо вона і не є самою вигідною). Такою клітиною є клітина (4.3), а тому будуємо для неї цикл (рис. 54). Мінімальна поставка парних клітин дорівнює 20 (вона знаходиться в клітині (4.4)). Переміщуючи цю поставку в 20 одиниць в циклі, одержимо новий розподіл поставок (таблиця 4.16). Оцінки нового розподілу запишемо в таблицю 4.17,

Таблиця 4.17

2	0	0	4
0	1	0	6
0	0	0	0
1	2	0	2

яка не містить від'ємних оцінок, а тому розподіл поставок, наведений у таблиці 4.16 є оптимальним, хоча не єдиним, оскільки нульові оцінки мають і дві вільні клітини.

За оптимальним розподілом поставок споживач 3 недоотримає 20 од. вантажу. Складемо підсумкову таблицю 4.18 та обчислимо мінімальні витрати на перевезення.

Таблиця 4.18

5	2 45	4 5	6
3 30	3	4 40	8
5 25	4	6	4 75

$$F_{\min} = 90 + 20 + 90 + 160 + 125 + 300 = 785 \text{ гр. одиниця.}$$

Приклад 2. Знайти оптимальний розподіл поставок, виконавши початковий розподіл поставок за правилом "північно-західного кута" (таблиця 4.19).

Таблиця 4.19

Поста-чаль-ники	**Потужність постачаль-ників**	**Споживачі та їх попит**			
		1	2	3	4
		65	55	135	65
I	160	3	5	7	8
II	80	5	5	6	3
III	90	6	10	3	3

Тут $\sum_{i=1}^{3} P_i = 330$ одиниць, $\sum_{j=1}^{4} S_j = 320$ одиниць, тобто $\sum_{i=1}^{3} P_i > \sum_{j=1}^{4} S_j$. Зайву потужність в 10 одиниць заплануємо фіктивному споживачу і побудуємо таблицю 4.20.

Таблиця 4.20

	1	2	3	4	Ф	
	65	55	135	65	10	
160	3 65	5 55	7 40	8	0	–3
80	5	5	6 80	3	0	–2
90	6	10	3 15	3 65	0 10	+1
	0	–2	–4	–4	–1	

Таблиця 4.21

0 65	0 55	0 30	1	–4 10
3	1	0 80	–3	–3
7	9	0 25	0 65	0
				+4

У таблиці 4.21 оцінюється початковий розподіл поставок, виконаний за правилом “північно-західного кута” згідно з таблицею 4.20.

Найбільша з абсолютною величиною від’ємна оцінка належить клітині (1.5). Будуємо для неї цикл і передаємо поставку в 10 од. у клітину (1.5), замінивши поставки в інших клітинах, які беруть участь у циклі.

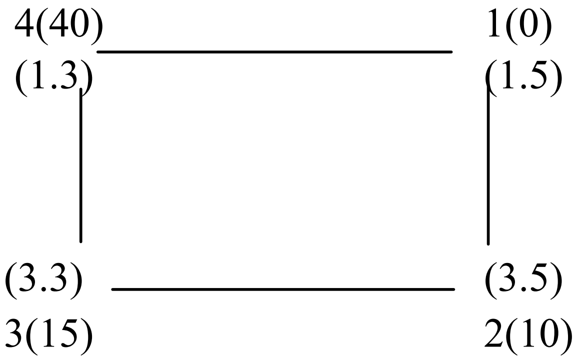


Рис. 55

Оцінимо розподіл поставок таблиці 4.21 і нові поставки запишемо в таблиці 4.22, яка включає і новий розподіл поставок, виконаний у циклі для клітини (2.4).

Таблиця 4.22

0 65	0 55	0 30	1	0 10
3	1	0 15	–3 65	1
7	9	0 90	0	4
			+3	

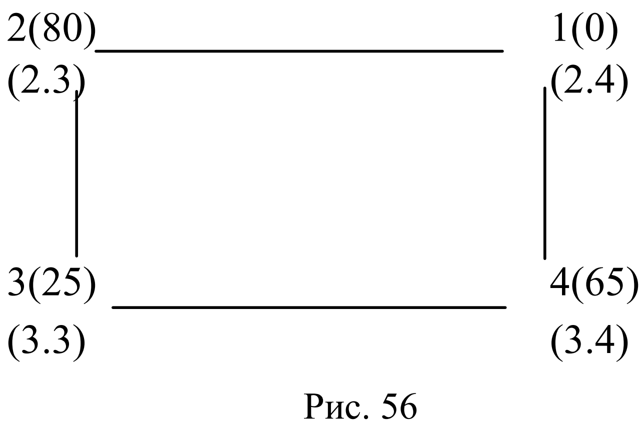


Рис. 56

У таблиці 4.23 оцінюється розподіл, наведений у таблиці 4.22.

Таблиця 4.23

0	0	0	4	0
3	1	0	0	1
7	9	0	3	4

Отриманий розподіл поставок у таблиці 4.22 є оптимальним, оскільки відсутні від’ємні оцінки в таблиці 4.23. При цьому розподілі у постачальника 1 залишиться 10 одиниць вантажу.

Щоб обчислити мінімальні витрати на перевезення складемо підсумкову таблицю 4.24.

Таблиця 4.24

3 65	5 55	7 30	8
5	5	6 15	3 65
6	10	3 90	3

Отже, $F_{\min} = 195 + 275 + 210 + 90 + 195 + 270 = 1235$ гр. одиниць.

Виродження в транспортних задачах. Виродженням в лінійному програмуванні означає обертання в нуль хоча б однієї основної змінної базисного розв’язку. Покажемо як це проявляється на прикладі конкретної задачі.

Знайти оптимальний розподіл поставок (таблиця 4.25).

Таблиця 4.25

Постачальники	**Потужність постачальників**	**Споживачі та їх попит**		
		1	2	3
		90	30	140
I	50	7	10	11
II	70	3	6	8
III	110	4	7	9
IV	30	8	1	5

Модель задачі закрита, оскільки $\sum_{i=1}^{4} P_i = \sum_{j=1}^{3} S_j$. Початковий розподіл поставок виконаємо за правилом “північно-західного кута” (таблиця 4.26).

Таблиця 4.26

	90	30	140
50	7 50	10	11
70	3 40	6 30	8
110	4	7	9 110
30	8	1	5 30

Початковий розподіл закінчений, баланс за рядками і стовпцями виконується, але у разі потреби побудувати цикли для клітин (1.3), (2.3), (3.1), (3.2), (4.1), (4.2), зробити не можливо. Отже, не виконується твердження, що для кожної вільної клітини можливо побудувати цикл.

Для з’ясування цього протиріччя перевіримо число заповнених клітин, які повинні відповідати даній задачі. Число споживачів n = 3, число постачальників n = 4; звідси число заповнених клітин має дорівнювати 4 + 3 – 1 = 6, а у таблиці заповнених клітин тільки 5, тобто менше необхідної кількості. Щоб зрівноважити цю кількість необхідно додати поставку ще в одну клітину. Оскільки баланс за рядками і за стовпцями вже встановлений, то розмір такої поставки встановлюємо як нульовий. Для з’ясування, в яку клітину записати нульову поставку, прослідкуємо, як виконувався початковий розподіл поставок. Нагадаємо, що записуючи в ту чи іншу клітину поставку, ми подумки викреслюємо або один рядок, або один стовпець, але коли записується остання поставка, одночасно викреслюються і рядок і стовпець. У нашому випадку спочатку викреслювався 1-ий рядок, а потім 1-ий стовпець. Коли в клітину (2.2) була задана поставка, рівна 30 одиницям, одночасно необхідно було викреслити і 2-ий рядок і 2-ий стовпець. Ми ж викреслимо лише, наприклад, тільки 2-ий рядок. У клітину (3.2) 2-го стовпця запишемо поставку, рівну нулю і вже після цього викреслимо цей стовпець. Зауважимо, що можна було спочатку викреслити 2-ий стовпець, а нульову поставку записати в клітину (2.3). Крім того, при розподілі поставок у таблиці відбулося порушення ступінчастості, яка характерна для розподілу поставок за правилом "північно-західного кута". Якщо записати нульову поставку в клітину (3.2) або в клітину (2.3), то така ступінчастість відновиться:

Таблиця 4.27

7 50	10	11	–7
3 40	6 30	8	–3
4	7 0	9 110	–4
8	1	5 30	0
0	–3	–5	

Зауваження. Не обов'язково число недостаючих заповнених клітин при початковому розподілі поставок має дорівнювати одиниці, як в даному випадку, а і двом, трьом і т.д.

А тому в таблиці можуть бути одна, дві, три нульові поставки і т.д.

Розподіл поставок, які містять нульові поставки в заповнених клітинах, називаються *виродженим розподілом.*

Воно відповідає виродженому базисному розв'язку (частина основних змінних дорівнює нулю).

У таблиці 4.27 заповнено необхідне число клітин — 6. Нульова поставка записана в клітину (3.2), оскільки у неї показник витрат менший, ніж у клітині (2.3).

У таблиці 4.27 також є додаткові рядок і стовпець, куди записують спеціально підібрані числа, які обертають оцінки заповнених клітин на нулі.

Оцінимо розподіл поставок, наведений у таблиці 4.27.

Таблиця 4.28

0 50	0	–1	
0 40	0 30	0	
0	0 0	0 110	+2
8	–2 0	0 30	+2
		–2	

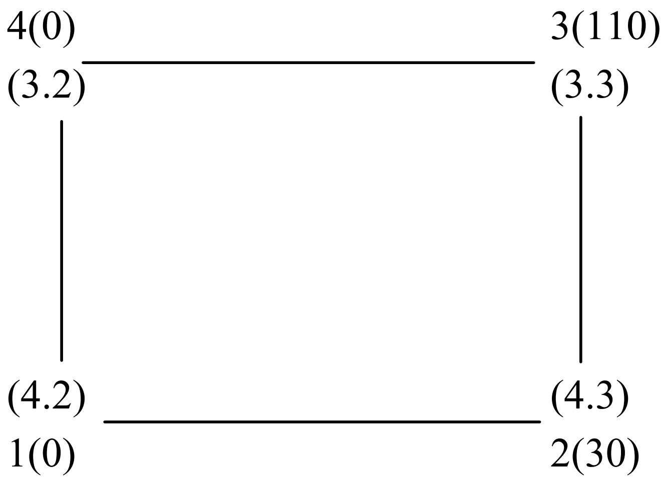


Рис. 57

У таблиці 4.28 нульові оцінки одержали клітини (1.2), (2.3), (3.1), хоча вони і не є заповненими, а тому при підрахунку оцінок не потрібно збереження цих нульових оцінок. Оскільки клітина (4.2) має найбільшу по абсолютній величині від'ємну оцінку, то будуємо для неї цикл (рис. 57). Мінімальна поставка в парних клітинах цього циклу дорівнює нулю. Щоб не порушувати прийнятий алгоритм розв'язування задачі, перемістимо в циклі для клітини (4.2) нульову поставку, після чого поставки в клітинах (3.3) і (4.3) збережуться, в клітину (4.2) перейде нульова поставка, а клітина (3.2) стане вільною.

Новий розподіл поставки запишемо в таблиці 4.28. До 4-го рядка таблиці 4.28 додаємо число 2, щоб в клітині (4.2) одержати нульову оцінку. Для збереження нульових оцінок у клітинах (3.3) і (4.3) із 3-го стовпця віднімемо 2, а до 3-го рядка додамо 2. Нові оцінки запишемо у таблицю 4.29.

Таблиця 4.29

0 20	0	–3 30	
0 70	0 0	–2	
2	2	0 110	–3
10	0 30	0	
		+3	

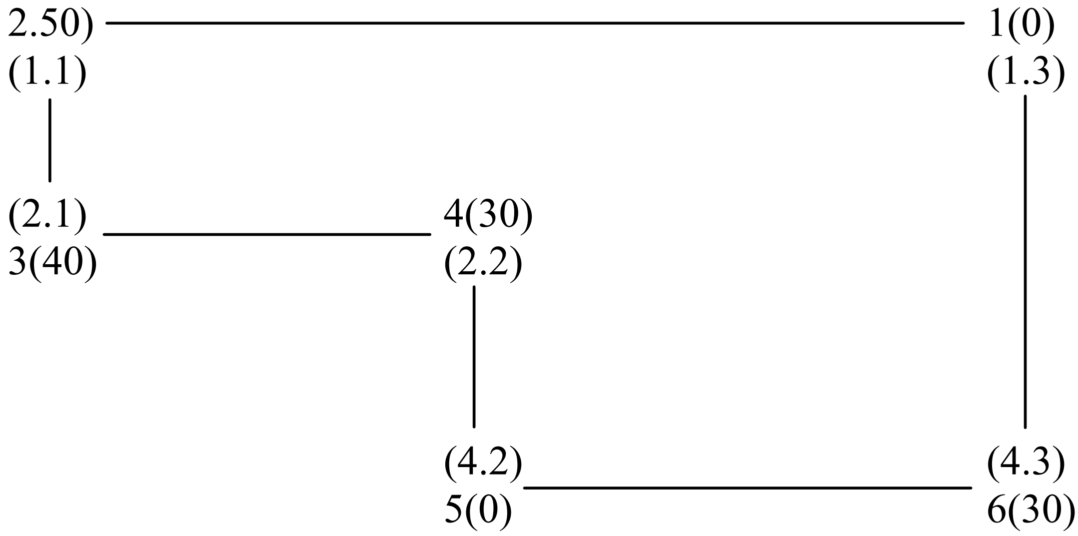


Рис. 58

У таблиці 4.29 є дві від'ємні оцінки, а тому побудуємо цикл для найбільшої від'ємної клітини (1.3) (рис. 58). В парних клітинах цього циклу стоять поставки, рівні 50, 30, 30 одиниць, причому мінімальна поставка знаходиться відразу в двох клітинах: (2.2) і (4.3). Передаючи поставку в 30 одиниць у клітину (1.3), необхідно звільнити лише одну клітину, а у нас звільнюється дві. Щоб число заповнених клітин залишилось без зміни, звільнимо лише одну, наприклад (4.3), а в другу (2.2) запишемо нульову поставку. Одержаний розподіл поставок, записаний у таблиці 4.29 є знову виродженим, крім того, в цьому випадку виродження відбулося в циклі.

Оцінимо розподіл поставок, наведений у таблиці 4.29 і випишемо це в таблицю 4.30.

Таблиця 4.30

0	0	0 50	+1
0 70	0 0	1	
–1 20	–1	0 90	+1
10	0 30	3	
		–1	

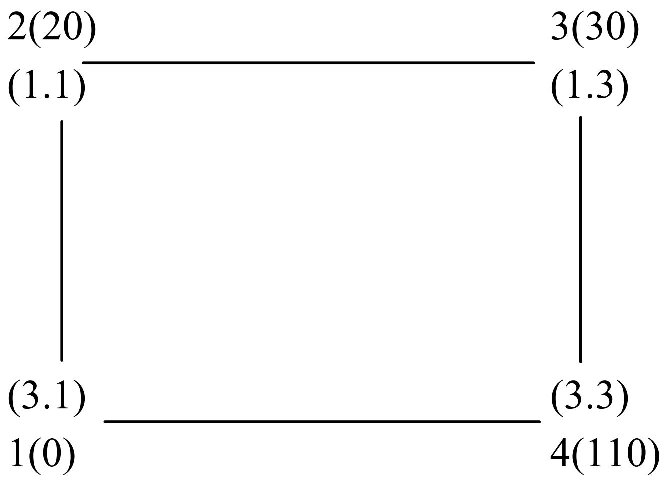


Рис. 59

У таблиці 4.30 клітини (3.1) і (3.2) мають однакові від'ємні оцінки. Побудуємо цикл для клітини (3.1) (рис.59), а одержаний розподіл поставок, проведений в цьому циклі внесемо у цю ж таблицю.

Оскільки таблиця 4.31 не містить від'ємних оцінок, то розподіл поставок у таблиці 4.30 є оптимальним.

Таблиця 4.31

1	1	0
0	0	0
0	0	0
10	0	2

ЗАДАЧІ

4.38. Виконати початковий розподіл поставок: а) за правилом урахування найменших витрат; б) за правилом "північно-західного кута" та підрахувати вартість витрат за цими розподілами поставок.

Постачальники	**Потужність постачальників**	**Споживачі та їх попит**			
		1	**2**	**3**	**4**
		60	40	90	60
I	120	4	4	7	5
II	80	2	3	6	8
III	50	5	1	5	9

4.39. Знайти оптимальний розподіл поставок, наведений у задачі 4.38.

4.40. На двох цукрових заводах А і В знаходиться відповідно 150 і 90 т цукру. У пункти 1,2,3 потрібно доставити відповідно 60, 70 і 110 т цукру. Вартість перевезення тони цукру із пункту А в пункт 1,2,3 складає відповідно 6, 10 і 4 гр. од., а із пункту В — 12, 2 і 8 гр. од. Скласти оптимальний план перевезень цукру, так щоб загальна сума транспортних витрат була найменшою.

4.41. У трьох автомобільних парках А, В і С знаходиться відповідно 60, 80 і 100 вантажних автомобілів. Скласти оптимальний план перегону цих автомобілів до чотирьох пунктів завантаження зерном, якщо пункту №1 необхідно 40 автомобілів, №2 — 60 автомобілів, №3 — 80 автомобілів і №4 — 60 автомобілів. Вартість перегону одного автомобіля із парку А у вказані пункти відповідно дорівнює 1, 2, 3, 4 гр. од., із парку В — 4, 3, 2, 0 гр. за одиницю, а із парку С — 0, 2, 2, 1. гр. од.

4.42. З трьох елеваторів А, В, С необхідно перевезти відповідно 10, 15, 25 тон сортового зерна в чотири пункти: пункт №1 — 5 т, №2 — 10 т, №3 — 20 т і №4 — 15 т. Вартість доставки однієї тони з елеватора А у зазначені пункти відповідно дорівнює 8, 3, 5, 2 гр. од., В — 4, 1, 6, 7 гр. од. і С — 1, 9, 4, 3 гр. од.

Скласти оптимальний план перевезення зерна в чотири пункти, мінімізуючи вартість перевезень.

Відповідь:

4.38. а) $F_1 = 1050$ гр. од.
б) $F_2 = 1490$ гр. од.

4.39. $F_{\min} = 1050$ гр. од.

4.40. $F_{\min} = 1020$ гр. од.

4.41. $F_{\min} = 280$ гр. од. Оптимальний план:
$x_{12} = x_{24} = x_{33} = 60,\ x_{23} = 20,\ x_{31} = 40$.

4.42. $F_{\min} = 140$ гр. од. Оптимальний план:
$x_{14} = x_{22} = 10,\ x_{23} = x_{31} = x_{34} = 5,\ x_{33} = 15$.

ДОДАТКИ

Додаток 1

Таблиця значень функції $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1893	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1,354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804

Продовження додатка 1

	0	1	2	3	4	5	6	7	8	9
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	00)9	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Додаток 2

Таблиця значень функції $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{z^2}{2}} dz$

x	Ф(*x*)	*x*	Ф(*x*)	*x*	Ф(*x*)	*x*	Ф(*x*)
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	03599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

Продовження додатка 2

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,52	0,4941
1,28	0,3997	1,61	0,4463	1,94	0,4738	2,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,50	0,499997
1,58	0,4429	1,91	0,4719	2,48	0,4934	5,00	0,499997

Додаток 3

Таблиця значень $t_{\gamma} = t(\gamma, n)$.

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0 , 999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,001	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Додаток 4

Таблиця значень $q = q(\gamma, n)$

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,48
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Додаток 5

Критичні точки розподілу χ^2

Число ступенів волі k	Рівень значущості					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Додаток 6

Критичні точки розподілу Стьюдента

Число ступенів волі k	Рівень значущості α (двохстороння критична область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,93
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	Рівень значущості α (одностороння критична область)					

ЗМІСТ

Про автора.. 3
Вступ.. 4

Розділ I

Теорія ймовірностей

§1. Основні поняття теорії ймовірностей.. 5
§2. Основні теореми.. 12
§3. Повторні випробування... 19
§4. Випадкові величини та їх розподіл.. 26
§5. Числові характеристики випадкових величин.......................... 45
§6. Деякі закони розподілу неперервної випадкової величини...................... 59
§7. Закон великих чисел і граничні теореми.................................. 80
§8. Функції випадкових величин.. 89
§9. Системи двох випадкових величин.. 95

Розділ II

Математична статистика

§1. Основні поняття математичної статистики. Вибірковий метод............... 115
§2. Числові характеристики статистичного розподілу................. 123
§3. Інтервальні (надійні) оцінки параметрів.................................. 134
§4. Елементи теорії кореляції.. 152

Розділ III

Елементи лінійної алгебри

§1. Основні поняття.. 173
§2. Визначники матриць другого порядку....................................... 174
§3. Визначники матриць третього порядку..................................... 177
§4. Визначники матриць вищих порядків....................................... 179
§5. Розв’язування систем n рівнянь із n невідомими..................... 193
§6. Ранг матриці, теорема про сумісність систем рівнянь першого степеня................ 210
§7. Основні операції з матрицями.. 221
§8. Обернена матриця. Розв’язування матричних рівнянь........... 227
§9. Модель багатогалузевої економіки.. 236

Розділ IV

Лінійне програмування

§1. Теоретичні осьови методів лінійного програмування.............. 248
§2. Геометричне представлення задачі лінійного програмування
та її розв’язування.. 254
§3. Симплексний метод... 272
§4. Двоїсті задачі... 294
§5. Транспортна задача... 314

Додатки.. 340

НАВЧАЛЬНЕ ВИДАННЯ

ВАСИЛЬЧЕНКО ІВАН ПЕТРОВИЧ

ВИЩА МАТЕМАТИКА ДЛЯ ЕКОНОМІСТІВ
(СПЕЦІАЛЬНІ РОЗДІЛИ)

Підручник

Редактор
Білокурський Сергій Петрович

Коректор
Медяник Наталія Василівна

Комп'ютерна верстка
Василенко Людмила Геннадіївна

Дизайн обкладинки
Сидоренко Марія Олексіївна

Підписано до друку 12.11.2006 р. Формат 60x84 1/16.
Друк офс. Папір офс. Гарнітура Таймс. Умов.друк.арк.22,1.
Обл..-вид.арк. 21,12. Наклад 500 прим. Зам. №

Видавництво „Кондор"
Свідоцтво про реєстрацію
ДК №1157 від 17.12.2002 р.