

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

**Ю. В. Куц,
Ю. Ю. Лисенко**

СПЕЦІАЛЬНІ РОЗДІЛИ МАТЕМАТИКИ КУРС ЛЕКЦІЙ

Навчальний посібник

Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня бакалавра
за освітньою програмою «Комп'ютерно-інтегровані системи та технології в
приладобудуванні»
спеціальності 151 Автоматизація та комп'ютерно-інтегровані технології

Електронне мережне навчальне видання

Київ
КПІ ім. Ігоря Сікорського
2022

Рецензент *Петрик В.Ф.*, к.т.н., доцент,
КПІ ім. Ігоря Сікорського
Єременко В.С., д.т.н., професор,
КПІ ім. Ігоря Сікорського

Відповідальний редактор *Баженов В.Г.*, к.т.н., доцент

*Гриф надано Методичною радою КПІ ім. Ігоря Сікорського
(протокол № 6 від 24.06.2022 р.)
за поданням Вченої ради приладобудівного факультету
(протокол № 5/22 від 30.05.2022 р.)*

Сучасний науково-технічний прогрес вимагає розроблення нових та адаптацію існуючих методів та алгоритмічно-програмних засобів вимірювань та контролю, засобів автоматизації виробничих процесів, постійного удосконалення методів оброблення та представлення результатів вимірювань. Все це вимагає знань та досвіду практичного застосування методів та критеріїв математичної статистики. Навчальний посібник складається з чотирьох розділів і містить матеріали лекцій з загальних питань, зокрема висвітлено питання перетворення Лапласа, теорії ймовірностей, математичної статистики та випадкових процесів. Кожний розділ супроводжується контрольними запитаннями з метою перевірки засвоєння знань, самоконтролю та самостійної роботи. Приведено довідковий матеріал для ознайомлення та розуміння процесів.

Даний посібник призначений для здобувачів ступеня бакалавра спеціальності «Автоматизація і комп'ютерно-інтегровані технології» та споріднених спеціальностей, а також наукових та інженерно-технічних працівників, фахівців, які займаються розробленням комп'ютерно-інтегрованих систем та технологій у галузі приладобудування.

Реєстр. № НП 21/22-670. Обсяг 8,2 авт. арк.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
проспект Перемоги, 37, м. Київ, 03056
<https://kpi.ua>

Свідоцтво про внесення до Державного реєстру видавців, виготовлювачів
і розповсюджувачів видавничої продукції ДК № 5354 від 25.05.2017 р.

© Ю. В. Куц, Ю. Ю. Лисенко
© КПІ ім. Ігоря Сікорського, 2022

ЗМІСТ

ВСТУП	5
Розділ 1. ОПЕРАЦІЙНЕ ЧИСЛЕННЯ	6
Лекція 1. ПЕРЕТВОРЕННЯ ЛАПЛАСА	6
Лекція 2. ОСНОВНІ ВЛАСТИВОСТІ ПЕРЕТВОРЕННЯ ЛАПЛАСА.....	14
Лекція 3. ВИЗНАЧЕННЯ ОРИГІНАЛІВ ЗА ЇХ ЗОБРАЖЕННЯМИ	25
Лекція 4. ЗАСТОСУНКИ ОПЕРАЦІЙНОГО ЧИСЛЕННЯ.....	31
Питання для самоконтролю	40
Розділ 2. ОСНОВИ ТЕОРІЇ ЙМОВІРНОСТЕЙ	42
Лекція 5. ОСНОВНІ ПОНЯТТЯ ТА ВИЗНАЧЕННЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ.....	42
Лекція 6. ЙМОВІРНІСНИЙ ПРОСТІР. ТЕОРЕМИ ДОДАВАННЯ І МНОЖЕННЯ ЙМОВІРНОСТЕЙ ПОДІЙ.....	51
Лекція 7. ВИПАДКОВА ВЕЛИЧИНА. ТЕОРЕМИ ПРО ДИСПЕРСІЮ ТА МАТЕМАТИЧНЕ СПОДІВАННЯ ДИСКРЕТНИХ ВИПАДКОВИХ ВЕЛИЧИН	58
Лекція 8. ПОВТОРЮВАННЯ ВИПРОБУВАНЬ. ЗАКОНИ ВЕЛИКИХ ЧИСЕЛ	66
Лекція 9. НЕПЕРЕРВНІ ВИПАДКОВІ ВЕЛИЧИНИ.....	74
Лекція 10. ФУНКЦІОНАЛЬНІ ПЕРЕТВОРЕННЯ ВИПАДКОВИХ ВЕЛИЧИН	86
Лекція 11. СИСТЕМА ДВОХ ВИПАДКОВИХ ВЕЛИЧИН	94
Питання для самоконтролю	104
Розділ 3. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ.....	107
Лекція 12. ОСНОВИ СТАТИСТИЧНОГО ОПРАЦЮВАННЯ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ.....	107
Лекція 13. МЕТОДИ ВИЗНАЧЕННЯ ТОЧКОВИХ ТА ІНТЕРВАЛЬНИХ ОЦІНОК РЯДУ СПОСТЕРЕЖЕНЬ	118
Лекція 14. МЕТОДИ ОЦІНЮВАННЯ ПАРАМЕТРІВ ЗАКОНУ РОЗПОДІЛУ ДАНИХ.....	127
Лекція 15. МЕТОДИ ВИЗНАЧЕННЯ ЗАКОНА РОЗПОДІЛУ ДАНИХ.....	135

Лекція 16. ЗАГАЛЬНІ ВІДОМОСТІ ПРО КОРЕЛЯЦІЙНИЙ ТА РЕГРЕСІЙНИЙ АНАЛІЗ.....	145
Питання для самоконтролю	162
Розділ 4. ВИБРАНІ ПИТАННЯ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ.....	164
Лекція 17. ОСНОВНІ ПОЛОЖЕННЯ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ.....	164
Питання для самоконтролю	172
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	174
ДОДАТОК 1	176
ДОДАТОК 2	177
ДОДАТОК 3	178
ДОДАТОК 4	179

ВСТУП

Конспект лекцій складено на основі сілабусу дисципліни «Спеціальні розділи математики» для студентів спеціальності 151«Автоматизація та комп'ютерно-інтегровані технології», освітньої програми «Комп'ютерно-інтегровані системи та технології в приладобудуванні». Конспект лекцій складається з чотирьох розділів і включає питання операційного числення, основ теорії ймовірностей та математичної статистики, базових відомостей про випадкові процеси.

Актуальність ґрунтовної математичної підготовки фахівців, що займаються розробленням та удосконаленням комп'ютерно-інтегрованих технологій в галузі приладобудування постійно зростає. Це пояснюється наступним. Науково-технічний прогрес постійно вимагає, по-перше, освоєння нових методів та алгоритмічно-програмних засобів вимірювань та контролю, засобів автоматизації виробничих процесів; по-друге, обґрунтування необхідної точності вимірювань на основі коректного застосування методів теорії ймовірностей та математичної статистики; по-третє, постійне удосконалення форм і методів оброблення та представлення результатів вимірювань, що знаходить відбиття в нових стандартах і нормативних документах, вимагає знання та досвіду практичного застосування методів та критеріїв математичної статистики; по-четверте, застосування різних методів вимірювань, контролю та діагностики постійно поширюється на нові сфери діяльності людини, в яких об'єкти досліджень мають яскраво виражену ймовірнісну природу. Це стосується, до прикладу, вимірювань та контролю в області нанотехнологій.

Мета навчального видання полягає в ознайомленні студентів з основними поняттями і теоретичними положеннями перетворення Лапласа, теорії ймовірностей та математичної статистики, базових понять теорії випадкових процесів, сприянні розвитку логічного та аналітичного мислення студентів.

Розділ 1. ОПЕРАЦІЙНЕ ЧИСЛЕННЯ

Лекція 1.

ПЕРЕТВОРЕННЯ ЛАПЛАСА

1.1. Стисла історична довідка

Операційне або символічне числення є ефективним методом дослідження теоретичних і прикладних задач у різних областях науки і техніки, зокрема, у фізиці, математиці, механіці, теорії автоматичного регулювання, електротехніки, радіотехніки та ін. Особливо значну роль відіграє операційне числення як ефективний інструмент розв'язання лінійних диференціальних рівнянь, обчисленні інтегралів, розв'язання інтегральних рівнянь типу згортки.

Методи операційного числення базуються на заміні операцій диференціювання та інтегрування заданої функції алгебраїчними операціями. У результаті цього багато математичних дій, виконання яких у диференціальному та інтегральному численні є достатньо складними, зводяться до простіших алгебраїчних дій.

Основи операційного числення були закладені ще у XVIII ст., але строге математичне обґрунтування з'явилося лише у перші роки минулого століття на основі теорії функції комплексної змінної. Значний внесок у розвиток цього розділу математики зробили видатні вчені Г. Лейбніц (1646-1716), Л. Ейлер (1707-1783), Ж. Лагранж (1736-1813), Ж. Фур'є (1768-1830), О. Коші (1789-1857), П. Лаплас (1749-1827). Саме П. Лаплас запропонував інтегральне перетворення для обчислення зображень, яке носить його ім'я.

В Україні вперше теоретичні основи операційного числення були обґрунтовані у 1862 році в монографії «Символьне числення та його застосування до інтегрування лінійних диференціальних рівнянь» професора Київського університету М.Є. Ващенко-Захарченка.

Англійський інженер-електрик О. Хевісайд (1850-1925) ввів у символічне числення оператор $p = d/dt$ та розглянув правила його застосування у

розрахунках. Крім того він запропонував одиничну функцію названу його ім'ям, використання якої дає змогу формально обмежити час існування функції напіввіссю $t \geq 0$ і задовольнити умову застосування перетворення Лапласа.

Строге математичне обґрунтування операційного числення було дано пізніше, у 20-х роках ХХ ст. Бромвичем і Карсоном, які пов'язали цей метод із методом інтегральних перетворень.

1.2. Загальні відомості про метод інтегральних перетворень

В цілому *інтегральним перетворенням* функції $f(t)$ (*функції-оригіналу*), яка належить певному класу функцій, називається перетворення виду

$$F(p) = \int_a^b k(t, p) \cdot f(t) dt \quad (1.1)$$

де $k(t, p)$ – *ядро перетворення*. Вид ядра перетворення визначає його властивості та залежності від характеру задач, на розв'язання яких воно орієнтовано. Ядро може бути як дійсним, так і комплекснозначним. До прикладу, якщо $k(t, p) = e^{-pt}$ маємо перетворення Лапласа, якщо $k(t, p) = 1/\pi t$ – перетворення Гільберта, а якщо $k(t, p) = e^{-i\omega t}$ – перетворення Фур'є.

Функцію $F(p)$ називають *зображенням* функції-оригіналу $f(t)$. Той факт, що функція $F(p)$ є зображенням оригіналу $f(t)$ символічно позначатимемо наступним чином: $f(t) \rightarrow F(p)$.

Більшість перетворень виду (1.1) є зворотними. Тому для певного інтегрального перетворення функція $f(t)$ може бути отримана з функції $F(p)$ за допомогою відповідної інтегральної *функції обернення*, тобто $F(p) \rightarrow f(t)$.

Загалом оператор (1.1) діє в просторі, елементами якого є функції $f(t)$. Згідно з (1.1) функція-оригінал $f(t)$ змінної t (найчастіше часу) перетворюється у відповідну функцію-зображення $F(p)$ іншої змінної p . Таке перетворення

переводить дослідження $f(t)$ в область їх зображень $F(p)$. Це дає змогу замінити математичні дії над функціями часу, тобто дії над оригіналами, більш простими діями над їхніми зображеннями. Після розв'язку цих рівнянь здійснюють зворотний перехід до функції дійсної змінної t . Зазначене вище ілюструє рис. 1.1.

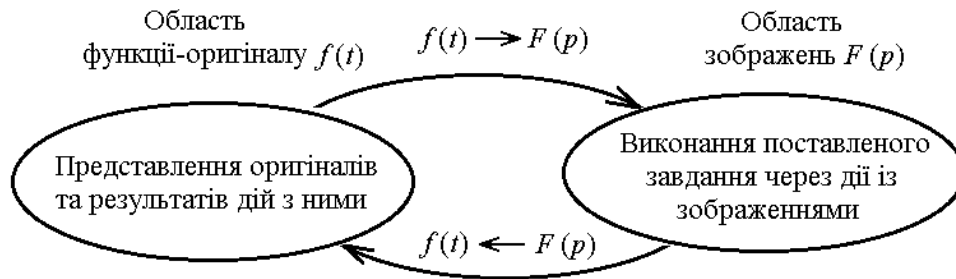


Рис. 1.1. Умовне графічне представлення загальної методології застосування інтегральних перетворень

Математичні операції з зображеннями виконуються значно простіше, а пошук та аналіз інформаційних ознак сигналів у значній кількості практичних завдань легше здійснити для функцій-зображень. До прикладу, перетворення Фур'є дає змогу перейти від дослідження неперервних в часі сигналів до їх образів – спектрів, що представлені набором чисел в частотній області.

Метод інтегральних перетворень виявився ефективним інструментом для розв'язання інтегральних рівнянь, лінійних диференціальних рівнянь за заданих початкових умов, систем лінійних диференціальних рівнянь, дослідження перехідних процесів в динамічних системах, аналізу роботи електричних кіл.

Опанування інтегрального перетворення Лапласа і розвинутого на його основі операторного метода передбачає вивчення:

- інтегрального перетворення Лапласа, його властивостей та застосування для визначення зображень;
- правил виконання операцій із зображеннями;
- способів визначення оригіналів за їх зображеннями.

1.3. Комплексні числа та функції

Комплексні числа вводять в математиці за допомогою уявної одиниці i , як числа, що дає в квадраті « -1 », тобто $i = \sqrt{-1}$. Загальна форма *комплексного числа* в *алгебраїчній формі* дається наступним виразом

$$a = \alpha + i\beta, \quad (1.2)$$

де α, β дійсні числа, які мають назви відповідно дійсної та уявної частини числа a і позначаються наступним чином: $\alpha = \operatorname{Re} a$, $\beta = \operatorname{Im} a$ (від французьких слів “*reel*” – дійсний та “*imaginaire*” – уявний).

Комплексні числа зображуються на площині векторами чи точками: число $a = \alpha + i\beta$ зображується точкою з абсцисою α і ординатою β або вектором, що починається в точці 0 (полюсі) і закінчується в точці a (рис. 1.2).

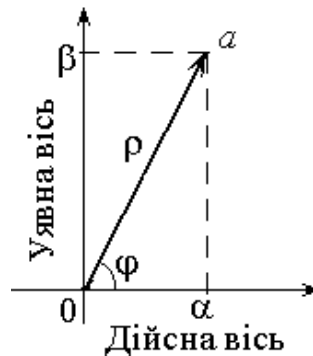


Рис. 1.2. Графічне представлення комплексного числа вектором на площині

Якщо ввести замість декартових координат точки, яка відповідає числу a , полярні координати, то отримаємо другу, *тригонометричну форму* запису комплексного числа

$$a = \rho(\cos \varphi + i \sin \varphi), \quad (1.3)$$

де ρ – довжина радіус-вектора відповідної точки, яка називається ще *модулем* або *абсолютною величиною* комплексного числа, φ – плоский кут (в радіанах), який називають *аргументом* комплексного числа. Для модуля і аргументу комплексного числа ще використовують такі позначення

$$\rho = |a|, \quad \varphi = \arg a. \quad (1.4)$$

Зв'язок між парами координат ρ , φ та α , β встановлюють рівняння

$$\alpha = \rho \cos \varphi, \quad \beta = \rho \sin \varphi, \quad (1.5)$$

$$\rho = \sqrt{\alpha^2 + \beta^2}, \quad \varphi = \operatorname{Arctg} \frac{\beta}{\alpha}. \quad (1.6)$$

В цих перетвореннях $0 \leq \rho < \infty$, а значення кута φ має нескінченну множину значень, які відрізняються одне від одного на $2\pi k$ (k – ціле число). *Головне значення аргументу* міститься в проміжку $-\pi/2 < \varphi \leq \pi/2$.

Для комплексних чисел має місце формула Ейлера

$$\cos kx + i \cdot \sin kx = e^{ikx}, \quad (1.7)$$

з якої витікають наступні рівняння, що дають змогу виразити тригонометричні функції через комплексні експоненти

$$\cos kx = \frac{e^{ikx} + e^{-ikx}}{2}, \quad \sin kx = \frac{e^{ikx} - e^{-ikx}}{2i}. \quad (1.8)$$

Використовуючи формулу Ейлера (1.7) отримують третю, *показникову*, форму запису комплексного числа

$$a = \rho e^{i\varphi} \quad (1.9)$$

Всі три форми запису комплексного числа (1.2), (1.3) та (1.9) еквівалентні.

Для комплексних чисел існує поняття спряжених чисел. Два комплексних числа називаються *взаємно спряженими* (позначаються a та \bar{a}), якщо їх дійсні частини рівні, а уявні частини відрізняються тільки знаком. Добуток комплексно спряжених чисел дає дійсне число, до прикладу

$$a \cdot \bar{a} = (\alpha + i\beta) \cdot (\alpha - i\beta) = \alpha^2 - i\alpha\beta + i\alpha\beta - i^2\beta^2 = \alpha^2 + \beta^2.$$

Два комплексних числа є рівними, якщо рівні їх дійсні та уявні частини.

Над комплексними числами виконують арифметичні операції додавання (дійсні частини додаються до дійсної, а уявні – до уявної), віднімання, множення та ділення.

Зведення комплексного числа в n -ю степінь виконують за *формулою Муавра*

$$a^n = [\rho(\cos \varphi + i \sin \varphi)]^n = \rho^n (\cos n\varphi + i \sin n\varphi). \quad (1.10)$$

Здобуття кореня n -ї степені комплексного числа як дія, обернена до зведення

комплексного числа в n -ю степiнь, виконують за формулою Муавра для дробного показника. Якщо $a = \rho(\cos \varphi + i \sin \varphi)$, то маємо

$$\sqrt[n]{a} = a^{\frac{1}{n}} = \sqrt[n]{\rho} \left(\cos \frac{\varphi + 2\pi k}{n} + i \sin \frac{\varphi + 2\pi k}{n} \right). \quad (1.11)$$

Додавання, віднімання, множення, ділення та зведення в цілу степiнь – дії однозначні; здобуття кореня n -ї степені комплексного числа завжди дає n різних значень.

Якщо величина $z = x + iy$ – комплексна змінна, то результат проведення над z певних алгебраїчних дій є алгебраїчною функцією комплексної змінної, що позначатимемо наступним чином: $w = f(z)$.

Комплексна функція від комплексного z (або дійсного t) аргументу може бути представлена у виді

$$f(z) = u(z) + iv(z), \quad (1.12)$$

де $u(z)$, $v(z)$ – дійсні функції комплексного аргументу z , які називаються відповідно дійсною і уявною частинами функції $f(z)$ (або $u(t)$, $v(t)$ – дійсні функції дійсного аргументу t функції $f(z)$.)

Тепер розглянемо поняття *аналітичної функції*.

Означення. Якщо функція $w = f(z)$ диференційована у всіх точках деякого кола з центром s_0 , то вона називається *аналітичною в точці s_0* ; функція називається *аналітичною у зв'язаній області* (тобто у зв'язаній області декількох змінних), якщо вона аналітична у всіх точках цієї області.

Необхідною і достатньою умовою того, щоби комплексна функція $u + iv = f(x + iy)$ була аналітичною є виконання умови Коші-Рімана:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \quad (1.13)$$

1.4. Оригінал та зображення

Сутність операторного методу, який ґрунтується на перетворенні Лапласа, полягає у відображенні функції дійсної змінної t (часової області) функцією комплексної змінної – оператора $p = s + i\sigma$ (область комплексних величин) таким чином, щоб замість інтегрально-диференціальних рівнянь функції дійсної змінної t отримати алгебраїчні рівняння для функції аргументу p .

Визначення оригіналу. Оригіналом називається комплексна функція $f(t) = u(t) + iv(t)$ дійсної змінної t , яка задовольняє умови:

- 1) $f(t)$ – однозначна неперервна або кусково-неперервна функція разом зі своїми похідними n -го порядку в інтервалі $(-\infty, \infty)$;
- 2) $f(t) = 0$ при $t < 0$;
- 3) існують такі числа $M > 0$ і $s < 0$, що для всіх $t < 0$ виконується нерівність

$$|f(t)| < Me^{st}. \quad (1.14)$$

Число $s_0 \geq 0$, для якого нерівність (1.13) виконується за будь-якого $s = s_0 + \varepsilon$, ($\varepsilon > 0$) і не виконується за значення $s = s_0 - \varepsilon$ (s_0 – точна нижня межа чисел s), називається показником росту функції $f(t)$.

Перша умова означає, що деякий фізичний процес, який описує функція $f(t)$, починається у момент часу $t = 0$ і розвиток цього процесу до початкового моменту (тобто при $t < 0$) не має значення. Третій умові задовольняють всі обмежені функції: функції-оригінали $f(t)$ при $t \rightarrow \infty$ є або обмеженими, або збігаючими до нескінченності, але не швидше ніж показникова функція $e^{s_0 t}$. Їх називають функціями експоненціального типу.

Означення. Зображенням функції-оригіналу $f(t)$ називається функція $F(p)$ комплексного змінного $p = s + i\sigma$, $s = \operatorname{Re} p$, $\sigma = \operatorname{Im} p$, яка визначається інтегралом Лапласа

$$F(p) = \int_0^{\infty} e^{-pt} f(t) dt . \quad (1.15)$$

Область існування зображення визначається наступною теоремою

Теорема 1.1. Якщо $f(t)$ – оригінал з показником росту s_0 , то функція $F(p)$ збігається у напівплощині $\operatorname{Re} p > s_0$ і є на ній аналітичною функцією.

Графічне представлення області існування зображення на площині (s, σ) подано на рис. 1.3.

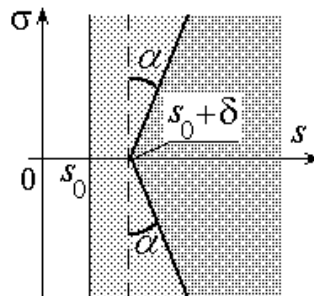


Рис. 1.3. Графічне представлення області існування зображення

1.5. Відшукання зображень деяких функцій

Одинична функція Хевісайда. Функція-оригінал дається виразом

$$\eta(t) = \begin{cases} 1, & t > 0, \\ 0, & t < 0 \end{cases} \quad (1.16)$$

Для значення $t = 0$ маємо: $\eta(0) = 0,5$, $\lim_{t \rightarrow 0+0} \eta(t) = 1$.

Отримаємо зображення функції Хевісайда за формулою (1.15)

$$F(p) = \int_0^{\infty} e^{-pt} \eta(t) dt = \lim_{a \rightarrow \infty} \int_0^a e^{-pt} dt = \lim_{a \rightarrow \infty} \frac{e^{-pt}}{-p} \Big|_0^a = \lim_{a \rightarrow \infty} \left(\frac{1}{p} - \frac{1}{p} e^{-pa} \right). \quad (1.17)$$

Якщо $\operatorname{Re} p > 0$, то $\lim_{a \rightarrow \infty} e^{-pa} = 0$. Отже маємо: $\eta(t) \rightarrow 1/p$, $\operatorname{Re} p > 0$.

Експоненціальна функція $e^{\alpha t}$. За формулою (1.15) визначаємо

$$F(p) = \int_0^{\infty} e^{-pt} e^{\alpha t} dt = \lim_{a \rightarrow \infty} \int_0^a e^{-(p-\alpha)t} dt = -\lim_{a \rightarrow \infty} \frac{e^{-(p-\alpha)t}}{p-\alpha} \Big|_0^a = -\lim_{a \rightarrow \infty} \left(\frac{1}{p-\alpha} - \frac{e^{-(p-\alpha)a}}{p-\alpha} \right). \quad (1.18)$$

Якщо $\operatorname{Re}(p - \alpha) > 0$, то $\lim_{a \rightarrow \infty} e^{-(p-\alpha)a} = 0$. Отже маємо:

$$e^{at} \rightarrow \frac{1}{p - \alpha}, \quad \operatorname{Re} p > \operatorname{Re} \alpha.$$

Гармонічна функція $\sin t$. Представимо гармонічну функцію з використанням формули Ейлера (1.8) та скористаємось виразом (1.15)

$$\begin{aligned} F(p) &= \int_0^{\infty} e^{-pt} \sin t dt = \lim_{a \rightarrow \infty} \int_0^a e^{-pt} \frac{e^{it} - e^{-it}}{2i} dt = \frac{1}{2i} \lim_{a \rightarrow \infty} \left[\int_0^a e^{-(p-i)t} dt - \int_0^a e^{-(p+i)t} dt \right] = \\ &= \frac{1}{2i} \lim_{a \rightarrow \infty} \left[-\frac{e^{-(p-i)t}}{p-i} \Big|_0^a + \frac{e^{-(p+i)t}}{p+i} \Big|_0^a \right] = \frac{1}{2i} \left[\frac{1}{p-i} - \frac{1}{p+i} \right] = \frac{1}{2i} \cdot \frac{2i}{p^2 - i^2} = \frac{1}{p^2 + 1}. \end{aligned}$$

Відшукування зображень безпосередньо за формулою (1.15) пов'язане з обчисленням невласних інтегралів і може бути достатньо складною задачею. У багатьох випадках для цього набагато зручнішим є використання властивостей перетворення Лапласа, які розглядатимуться в наступній лекції.

Зведення формул перетворення Лапласа наведено у Додатку 1.

Лекція 2.

ОСНОВНІ ВЛАСТИВОСТІ ПЕРЕТВОРЕННЯ ЛАПЛАСА

Визначення зображень оригіналів, які уявляють собою складні функції дійсного аргументу через інтеграл Лапласа – важка задача. Значно простіше вона виконується за допомогою певних правил, які складають суть операційного числення. Ці правила сформульовані на основі доведення ряду теорем операційного числення.

2.1. Теорема додавання, подібності, запізнювання, випередження, зміщення

Перетворення Лапласа має властивість лінійності, що доводить теорем 1.

Теорема 2.1. Якщо $f_1(t) \rightarrow F_1(p)$, $f_2(t) \rightarrow F_2(p)$, ..., $f_n(t) \rightarrow F_n(p)$, і

$\alpha_1, \alpha_2, \dots, \alpha_n$ – числа, то

$$\alpha_1 f_1(t) + \alpha_2 f_2(t) + \dots + \alpha_n f_n(t) \rightarrow \alpha_1 F_1(p) + \alpha_2 F_2(p) + \dots + \alpha_n F_n(p). \quad (2.1)$$

Доведення. З формули інтеграла Лапласа витікає, що $\alpha_1 f_1(t) \rightarrow \alpha_1 F_1(p)$ та $\alpha_1 f_1(t) + \alpha_2 f_2(t) \rightarrow \alpha_1 F_1(p) + \alpha_2 F_2(p)$ оскільки інтегрування є лінійною операцією. Продовжуючи додавати до сум (2.1) інші пари оригіналів та зображень з індексами 3, 4, ...n доводимо справедливість цієї теореми, тобто доводить властивість лінійності перетворення Лапласа.

Теорема 2.2. Якщо $f(t) \rightarrow F(p)$ і α – число, то

$$f(\alpha t) \rightarrow \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right). \quad (2.2)$$

Доведення. Нехай $\alpha > 0$. Виконуючи заміну змінної $\alpha t = \tau$ маємо

$$f(\alpha t) \rightarrow \int_0^{\infty} e^{-pt} f(\alpha t) dt = \frac{1}{\alpha} \int_0^{\infty} e^{-\frac{p}{\alpha} \tau} f(\tau) d\tau, \quad (2.3)$$

або $f(\alpha t) \rightarrow \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right)$, що і доводить теорему подібності.

Поклавши в рівнянні (2.2) $\alpha = 1/\beta$ як наслідок отримуємо таку властивість

$$F(\beta p) \rightarrow \frac{1}{\beta} f\left(\frac{t}{\beta}\right). \quad (2.4)$$

Теорема подібності справедлива і у випадку, коли α – комплексне число.

Теорема 2.3. Якщо $f(t) \rightarrow F(p)$ і $t_0 > 0$, то

$$f(t-t_0) \rightarrow e^{-t_0 p} F(p), \quad (2.5)$$

Доведення. За умовою теореми маємо

$$f(t-t_0) = \begin{cases} 0, & t \leq 0, \\ f(t-t_0), & t > t_0. \end{cases} \quad (2.6)$$

Зсуваючи графік функції $f(t)$ (рис. 2.1а) праворуч на величину t_0 отримаємо графік функції $f(t-t_0)$ (рис. 2.1б).

Зображення оригіналу функції $f(t-t_0)$ визначається наступним чином

$$f(t-t_0) \rightarrow \int_0^{\infty} e^{-pt} f(t-t_0) dt = \int_{t_0}^{\infty} e^{-pt} f(t-t_0) dt \quad (2.7)$$

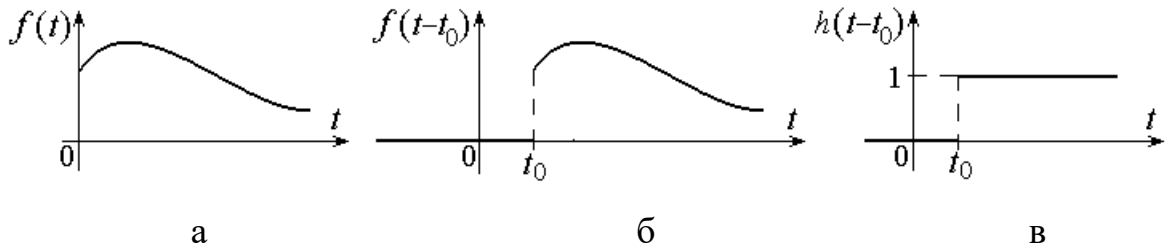


Рис. 2.1. Графічна ілюстрація до теореми запізнювання

Виконуючи заміну змінної $t-t_0 = \tau$ отримаємо

$$\int_{t_0}^{\infty} e^{-pt} f(t-t_0) dt = \int_0^{\infty} e^{-p(\tau+t_0)} f(\tau) d\tau = e^{-pt_0} \int_0^{\infty} e^{-p\tau} f(\tau) d\tau. \quad (2.8)$$

Маємо: $f(t-t_0) \rightarrow e^{-pt_0} F(p)$, $t > t_0$, що доводить теорему запізнювання.

Множення зображення $F(p)$ на e^{-pt_0} зсуває графік його оригіналу $f(t)$ праворуч на t_0 , тобто приводить до запізнювання оригіналу на час t_0 .

Застосовуючи теореми подібності та запізнювання можна визначити зображення для оригіналу виду $f(\alpha t - t_0)$, $t_0 > 0$, де α – число. Нехай $f(t) \rightarrow F(p)$. Тоді згідно з теоремою подібності $f(\alpha t) \rightarrow \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right)$. За теоремою запізнювання визначаємо

$$f(\alpha t - t_0) = f\left[\alpha\left(t - \frac{t_0}{\alpha}\right)\right] \rightarrow \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right) e^{-\frac{t_0}{\alpha} p}. \quad (2.9)$$

Функція

$$h(t-t_0) = \begin{cases} 1, & t > t_0, \\ 0, & t < t_0 \end{cases} \quad (2.10)$$

називається *узагальненою одиничною функцією* (рис. 2.1в). Вона має зображення

$$h(t-t_0) \rightarrow \frac{e^{-pt_0}}{p}. \quad (2.11)$$

Теорема 2.4. Якщо $f(t) \rightarrow F(p)$ і $t_0 > 0$, то

$$f(t+t_0) \rightarrow e^{-t_0 p} \left(F(p) - \int_0^{\infty} e^{-pt} f(t) dt \right), \quad (2.12)$$

Доведення. Графік функції $f(t+t_0)$ (рис. 2.2а) отримують шляхом зміщення графіка $f(t)$ (рис. 2.2б) ліворуч на величину t_0

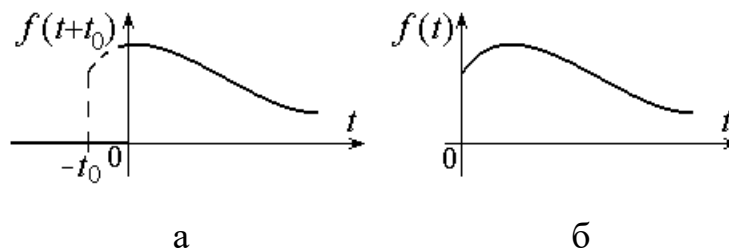


Рис. 2.2. Графічна ілюстрація до теореми випередження

Зміщена частина графіку $f(t)$ в інтервалі $t_0 < t < 0$ вироджується у відрізок осі t , а усічена частина для $t > 0$ є графіком функції $f(t+t_0)$ (рис. 2.2а). Маємо

$$f(t+t_0) \rightarrow \int_0^{\infty} e^{-pt} f(t+t_0) dt \quad (2.13)$$

Зробивши заміну змінної $t+t_0 = u$, $t_0 \leq u < \infty$ отримаємо

$$\begin{aligned} \int_0^{\infty} e^{-pt} f(t+t_0) dt &= \int_0^{\infty} e^{-p(u-t_0)} f(u) du = e^{pt_0} \int_{t_0}^{\infty} e^{-pu} f(u) du = \\ &= e^{pt_0} \left(F(p) - \int_0^{\infty} e^{-pu} f(u) du \right), \end{aligned} \quad (2.14)$$

що і доводить теорему.

Теорема 2.5. Якщо $f(t) \rightarrow F(p)$ і p_0 – число, то

$$e^{-p_0 t} f(t) \rightarrow F(p+p_0). \quad (2.15)$$

Доведення. Згідно визначення зображення функції маємо

$$e^{-p_0 t} f(t) \rightarrow \int_0^{\infty} e^{-pt} e^{-p_0 t} f(t) dt = \int_0^{\infty} e^{-(p+p_0)t} f(t) dt, \quad (2.16)$$

або $e^{-p_0 t} f(t) \rightarrow F(p + p_0)$, що і доводить теорему зміщення.

Таким чином, у випадку множення оригіналу на $e^{\pm p_0 t}$ в площині p відбувається зміщення на вектор $\mp p_0$ (рис. 2.3).

Теорему зміщення застосовують при розгляді фізичних явищ, пов'язаних із загасанням коливань.

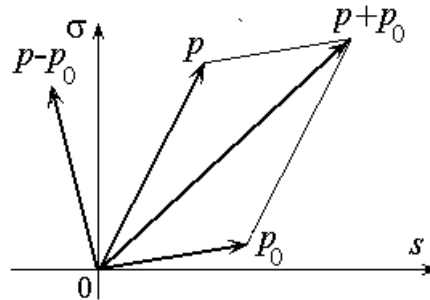


Рис. 2.3. Графічна ілюстрація до теореми зміщення

2.2. Теореми про диференціювання оригіналу та диференціювання зображення

Наступні дві теореми стосуються диференціювання оригіналу та диференціювання зображення.

Теорема 2.6. Якщо $f(t) \rightarrow F(p)$ і p_0 – число, і функції $f'(t), f''(t), \dots, f^{(n)}(t)$ є оригіналами, то

$$\begin{aligned} f'(t) &\rightarrow pF(p) - f(0), \\ f''(t) &\rightarrow p^2 F(p) - pf(0) - f'(0), \\ &\dots\dots\dots \\ f^{(n)}(t) &\rightarrow p^n F(p) - p^{n-1} f'(0) - \dots - f^{(n-1)}(0), \end{aligned} \quad (2.17)$$

де $f^{(k)}(0) = \lim_{t \rightarrow 0+0} f^{(k)}(t)$, $k = 0, 1, 2, \dots, n-1$

Доведення. За визначенням зображення маємо

$$f'(t) \rightarrow \int_0^{\infty} e^{-pt} f'(t) dt. \quad (2.18)$$

Інтегруючи праву частину за частинами отримаємо

$$\int_0^{\infty} e^{-pt} f'(t) dt = e^{-pt} f(t) \Big|_0^{\infty} + p \int_0^{\infty} e^{-pt} f(t) dt. \quad (2.19)$$

Оскільки $\operatorname{Re}(p) = s > s_0$, то $\operatorname{Re}(p) = s > s_0 \left| e^{-pt} f(t) \right| < M e^{-(s-s_0)t}$, тому $\lim_{t \rightarrow \infty} e^{-pt} f(t) = 0$. Якщо оригінал $f(t)$ в точці $t=0$ є неперервною функцією, то $f(0) = 0$, оскільки $f(t) = 0$ при $t < 0$; якщо при $t=0$ функція $f(t)$ має розрив першого роду, то при $t \rightarrow 0$ функція $f(t)$ має границю справа, тобто $\lim_{t \rightarrow 0+0} f^{(k)}(t) = f(0)$, тобто $e^{-pt} f(t) \Big|_0^{\infty} = -f(0)$.

Таким чином, $f'(t) \rightarrow pF(p) - f(0)$, $\operatorname{Re}(p) > s_0$.

Якщо при $t=0$ функція $f(t)$ неперервна, то $f'(t) \rightarrow pF(p)$. Це означає, що диференціюванню оригіналу $f(t)$ відповідає у просторі зображень множення на p функції $F(p)$.

Інтегруючи двічі по частинам інтеграл $\int_0^{\infty} e^{-pt} f''(t) dt$ отримаємо:

$$f''(t) \rightarrow p^2 F(p) - pf(0) - f'(0).$$

Використовуючи метод математичної індукції та припускаючи, що вираз

$$f^{(n)}(t) \rightarrow p^n F(p) - p^{n-1} f'(0) - \dots - f^{(n-1)}(0) = \Phi(p)$$

справедливий для n , покажемо, він є вірним і для $n+1$.

З доведення теореми для $n=1$ маємо $\left[f^{(n)}(t) \right]' \rightarrow p\Phi(p) - f^{(n)}(0)$ або

$$f^{(n+1)}(t) \rightarrow p^{n+1} F(p) - p^n f(0) - \dots - pf^{(n-1)}(0) - f^n(0). \quad (2.20)$$

Формула є вірною і для будь-якого натурального числа n . Якщо функція $f^{(n)}(t)$ має зображення, то і $f^{(n-1)}(t)$ має зображення; зворотне взагалі не можливе.

Якщо при $t=0$ функція $f(t)$ та її похідна $f^{(k)}(t)$, $k=1, 2, \dots, n$ є неперервними, то $f^{(n)}(t) \rightarrow p^n F(p)$, отже n -кратному диференціюванню в просторі оригіналів відповідає множення на p^n функції $F(p)$ в просторі зображень. В цьому частинному випадку величину p розглядають як оператор.

Наведемо без доведення наступну теорему.

Теорема 2.7. Якщо $F(p) \rightarrow f(t)$, $\text{Re } p > s_0$, то

$$\begin{aligned} F'(p) &\rightarrow -tf(t), \\ F''(p) &\rightarrow (-1)^2 t^2 f(t), \end{aligned} \tag{2.21}$$

.....

$$F^{(n)}(p) \rightarrow (-1)^n t^n f(t), \text{Re } p_1 > s_0 .$$

Отже диференціюванню зображення $F(p)$ відповідає у просторі оригіналів операція множення на $-t$ функції $f(t)$.

2.3. Теорема про інтегрування оригіналу та інтегрування зображення

Спочатку доведемо теорему про інтегрування оригіналу.

Теорема 2.8. Якщо $f(t) \rightarrow F(p)$, $\text{Re } p > s_0$, то

$$\int_0^t f(\tau) d\tau \rightarrow \frac{F(p)}{p}, \text{Re } p > s_0 \tag{2.22}$$

Доведення. Функція $\varphi(t) = \int_0^t f(\tau) d\tau$, вочевидь, задовольняє умовам 1) та 2)

функції-оригіналу, а також умові 3), оскільки

$$\left| \int_0^t f(\tau) d\tau \right| < \int_0^t M e^{s_0 \tau} d\tau = \frac{M}{s_0} (e^{s_0 t} - 1) < M_1 e^{s_0 t}, \tag{2.23}$$

де $M_1 = M/s_0$. Отже функції $f(t)$ є оригіналом з показником росту s_0 .

Нехай $\varphi(t) \rightarrow \Phi(p)$, $\text{Re } p > s_0$, тоді згідно з теоремою диференціювання

оригіналу маємо $\varphi'(t) \rightarrow p\Phi(p)$, $\varphi(0)=0$, оскільки $\varphi'(t) = f(t)$, то $F(p) = p\Phi(p)$. Звідси отримаємо

$$\int_0^t f(\tau) d\tau \rightarrow \frac{F(p)}{p}.$$

Розглянемо тепер інтегрування зображень

Теорема 2.9. Якщо $f(t) \rightarrow F(p)$, $\operatorname{Re} p > s_0$ та інтеграл $\varphi(t) = \int_p^\infty F(p) dp$

збігається у напівплощині $\operatorname{Re} p > s_1 > s_0$, то

$$\int_p^\infty F(p) dp \rightarrow \frac{f(t)}{t}, \operatorname{Re} p > s_1 > s_0. \quad (2.24)$$

Доведення. В напівплощині $\operatorname{Re} p > s_0 + \delta$ (рис. 2.1) інтеграл Лапласа $F(p) = \int_0^\infty e^{-pt} f(t) dt$ збігається рівномірно відносно p . Тому в цій напівплощині його інтегрують за параметром p , і за контур інтегрування можна обрати будь-який промінь, що виходить з точки p і утворює з дійсною віссю гострий кут. Маємо

$$\int_p^\infty F(p) dp = \int_p^\infty dp \int_0^\infty e^{-pt} f(t) dt. \quad (2.25)$$

Оскільки за умовою інтеграл $\int_p^\infty F(p) dp$ збігається у напівплощині

$\operatorname{Re} p > s_1 > s_0$, тому

$$\int_p^\infty dp \int_0^\infty e^{-pt} f(t) dt = \int_0^\infty f(t) dt \int_p^\infty e^{-pt} dp. \quad (2.26)$$

або

$$\int_p^\infty F(p) dp = \int_0^\infty e^{-pt} \frac{f(t)}{t} dt. \quad (2.27)$$

Отже функція $f(t)/t$ є оригіналом з показником росту $s_1 > s_0$, а інтегрування $F(p)$ зводиться в просторі оригіналу до ділення на t функції $f(t)$.

2.4. Згортка функцій та її властивості

Означення. Згортокою неперервних функцій $\varphi(t)$ та $f(t)$, $0 \leq t < \infty$ (позначається як $\varphi(t) * f(t)$) називають інтеграл

$$\varphi * f = \int_0^t \varphi(t-\tau) f(\tau) d\tau . \quad (2.28)$$

Згортокою є дія, яка парі функцій з певної множини ставить у відповідність іншу певну функцію з цієї множини. Дія отримання згортки називається *згортанням функцій*.

Згортка функцій має наступні властивості.

1. Комутативність: $\varphi * f = f * \varphi$. Ця властивість легко доводиться з (2.28) шляхом заміни $t - \tau = u$.

2. Асоціативність: $(\varphi * f) * \psi = \varphi * (f * \psi)$.

3. Дистрибутивність відносно додавання: $\varphi * (f + \psi) = \varphi * f + \varphi * \psi$.

4. Абсолютна величина згортки: $|\varphi * f| \leq |\varphi| * |f|$.

5. Якщо функції $\varphi(t)$ та $f(t)$ – неперервні функції при $t \geq 0$, то їх згортка $\varphi * f$ є також неперервною функцією при $t \geq 0$.

6. Якщо функції $f(t)$ та $\varphi(t)$ неперервні при $t \geq 0$, і їх згортка $f * \varphi = 0$, то принаймні одна з них при $t \geq 0$ дорівнює нулю (*теорема Тітчмарша*).

7. Якщо функції $f(t)$ та $\varphi(t)$ оригінали, то і їх згортка $f * \varphi$ також є оригіналом.

2.5. Множення зображень

В операційному численні важливе місце відіграє теорема про множення зображень (теорема Бореля) та її наслідок – інтеграл Дюамеля.

Теорема 2.9 (Е. Бореля). Якщо $f(t) \rightarrow F(p)$, $\operatorname{Re} p > s_0$ та $\varphi(t) \rightarrow \Phi(p)$ $\operatorname{Re} p > s_1$, то $f * \varphi \rightarrow F(p)\Phi(p)$.

Доведення. Маємо

$$\varphi * f = \int_0^t e^{-pt} f * \varphi dt = \int_0^\infty e^{-pt} dt \int_0^t f(t-\tau)\varphi(\tau) d\tau. \quad (2.29)$$

В цьому виразі двократний інтеграл по області $D: 0 \leq t < \infty, 0 \leq \tau \leq t$ (рис. 2.4) збігається абсолютно, оскільки зображення оригіналу $f * \varphi$ визначено в напівплощині $\operatorname{Re} p > s_0$ при $s_0 > s_1$. Тому у двократному інтегралі можна змінити порядок інтегрування ($0 \leq \tau < \infty, \tau \leq t < \infty$):

$$f * \varphi = \int_0^\infty \varphi(\tau) d\tau \int_\tau^\infty e^{-pt} f(t-\tau) dt. \quad (2.30)$$

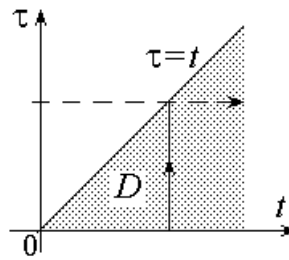


Рис. 2.4. Графічне зображення області D

Позначивши $t - \tau = u$ отримаємо

$$f * \varphi = \int_0^\infty e^{-pt} \varphi(\tau) d\tau \int_\tau^\infty e^{-pu} f(u) du \rightarrow F(p)\Phi(p). \quad (2.31)$$

Отже згортці в просторі оригіналу відповідає множення функцій в просторі зображень.

Тепер розглянемо інтеграл Дюамеля, якій має винятково важливе значення в аналізі лінійних систем.

Теорема 2.10. Якщо $f * \varphi \rightarrow F(p)\Phi(p)$, то

$$\int_0^t f(\tau)\varphi'_t(t-\tau) d\tau + f(t)\varphi(0) \rightarrow pF(p)\Phi(p), \quad (2.32)$$

або

$$\int_0^t \varphi(\tau) f'_i(t-\tau) d\tau + \varphi(t) f(0) \rightarrow pF(p)\Phi(p). \quad (2.33)$$

Доведення. Згідно з властивістю диференціювання оригіналу маємо

$$\frac{d}{dt} \int_0^t f(\tau) \varphi(t-\tau) d\tau \rightarrow pF(p)\Phi(p),$$

або

$$\frac{d}{dt} \int_0^t \varphi(\tau) f(t-\tau) d\tau \rightarrow pF(p)\Phi(p),$$

оскільки, у відповідності до властивості комутативності згортки, $\varphi * f = f * \varphi$.

З курсу математичного аналізу відомо наступне. Якщо для інтегралу

$\int_{\alpha(t)}^{\beta(t)} F(t, \tau) d\tau$, що залежить від параметра t , виконуються умови:

- 1) $\alpha(t), \beta(t)$ та $\alpha'(t), \beta'(t)$ – неперервні функції на відрізку $[t_0, t_1]$;
- 2) $\frac{dF(t, \tau)}{dt}$ – неперервна функція в області, яка обмежена лініями $t = t_0, t = t_1$

, і $\tau = \alpha(t), \tau = \beta(t)$, то похідна по аргументу $t_0 \leq t \leq t_1$ від інтегралу дорівнює

$$\frac{d}{dt} \int_{\alpha(t)}^{\beta(t)} F(t, \tau) d\tau = \frac{d}{dt} \int_{\alpha(t)}^{\beta(t)} \frac{\partial F(t, \tau)}{\partial t} d\tau + F(t, \beta(t))\beta'(t) - F(t, \alpha(t))\alpha'(t). \quad (2.34)$$

Користуючись формулою (2.34) отримаємо:

$$\frac{d}{dt} \int_0^t f(\tau) \varphi(t-\tau) d\tau = \int_0^t f(\tau) \varphi'(t-\tau) d\tau + f(t) \varphi(0) \rightarrow pF(p)\Phi(p),$$

що і доводить цю теорему.

Висновок: якщо одна з функцій $f(t)$ чи $\varphi(t)$ диференційована, а інша неперервна, то згортка $f * \varphi$ цих функцій є диференційованою.

Вирази (2.32), (2.33) вперше у 1853 р. застосував Жан Марі Констан Дюамель до розрахунку динаміки механічних систем. Пізніше цей метод почав використовуватись і для розрахунку електротехнічних ланцюгів.

Для використання інтеграла Дюамеля в електротехнічних розрахунках для пошуку реакції ланцюга на певний електричний сигнал необхідно попередньо визначити *перехідну функцію* ланцюга яка є відгуком ланцюга на ступінчатий одиничний вхідний сигнал. Для лінійного ланцюга перехідною функцією може бути аперіодичний, коливальний, загасаючий коливальний процеси або їх комбінація.

Якщо вхідний сигнал ланцюга заданий функцією де незалежна змінна, реакція ланцюга на цей сигнал виражається формулою

$$U_{\text{вих}}(t) = \int_0^t U'(\tau)h(t-\tau)d\tau + U(0)h(t) ,$$

де $U'(t)$ – похідна вхідного впливу в часі.

Властивості перетворення Лапласа зведені в таблицю в Додатку 2.

Лекція 3.

ВИЗНАЧЕННЯ ОРИГІНАЛІВ ЗА ЇХ ЗОБРАЖЕННЯМИ

Перетворення Лапласа кожному оригіналу $f(t)$ ставить у відповідність одне і тільки одне зображення $F(p)$. Обернена задача відновлення оригіналу за зображенням теж розв'язується однозначно. Ця процедура здійснюється за допомогою оберненого перетворення Лапласа, яке виконується згідно з теоремою обернення.

3.1. Теорема обернення

Теорема 3.1. Якщо функція $f(t)$ є оригіналом з показником зростання s_0 і $F(p)$ – зображення цієї функції, то в будь-якій точці неперервності $f(t)$ виражається через $F(p)$ за формулою

$$f(t) = \frac{1}{2\pi i} \int_{s-i\infty}^{s+i\infty} e^{pt} F(p) dp , \quad (3.1)$$

де інтеграл береться по будь-якій прямій $\operatorname{Re} p = s > s_0$ і розуміється як головне значення, тобто

$$\int_{s-i\infty}^{s+i\infty} e^{pt} F(p) dp = \lim_{R \rightarrow \infty} \int_{s-iR}^{s+iR} e^{pt} F(p) dp . \quad (3.2)$$

У точках розриву маємо:

$$\int_{s-iR}^{s+iR} e^{pt} F(p) dp = \frac{1}{2} [f(t-0) + f(t+0)] .$$

Формула (3.1) ще має назву *формули обернення Рімана-Мелліна* і визначає *обернене перетворення Лапласа*.

З формули (3.1) витікає таке: якщо два оригінала $f_1(t)$ та $f_2(t)$ мають одне і те саме зображення $F(p)$, то в точках неперервності вони є рівними, оскільки виражаються через $F(p)$ за допомогою одного і того самого інтеграла. Це власне доводить *теорему одного єдино можливого оригіналу*:

якщо функції $f_1(t)$ та $f_2(t)$ мають одне і те саме зображення $F(p)$, то вони є рівними в точках неперервності.

Відшукування оригіналів безпосередньо за формулою (3.1) є складною математичною задачею. Тому на практиці користуються простішим способом, використовуючи *наслідки* з теореми обернення – так звані *теореми розкладання*.

3.2. Перша та друга теореми розкладання

Теорема 3.2. Якщо функція $F(p)$ аналітична в нескінченно віддаленій точці, $F(\infty) = 0$ і в околі цієї точки її розвинення в ряд за степенями $1/p$ має вид

$$F(p) = \sum_{n=0}^{\infty} \frac{c_n}{p^{n+1}} , \quad (3.3)$$

то функція $f(t) = \sum_{n=0}^{\infty} c_n \frac{t^n}{n!}$ є оригіналом, що відповідає зображенню $F(p)$.

Теорема 3.3. Якщо зображення $F(p)$ є однозначною функцією й має лише

скінченне число полюсів $p_1, p_2, \dots, p_k, \dots, p_n$, що лежать у скінченній площині, то відповідний оригінал $f(t)$ дорівнює сумі лишків функції $e^{pt}F(p)$, що обчислені в усіх полюсах p_k функції $F(p)$, тобто

$$f(t) = \sum_{k=1}^n \text{res} \{ e^{pt} F(p) \} \quad (3.4)$$

В теорії комплексних функцій залишком в ізольованій особливій точці p_0 (яка належить області комплексних чисел) називається інтеграл виду

$$\frac{1}{2\pi i} \oint_{\gamma} F(p) dp$$

де γ – контур, що належить околу точки p_0 і охоплює її. Обхід контуру – позитивний, тобто проти ходу стрілки годинника. Залишки позначаються як res (від французького *residu* – віднімати):

$$\text{res}_{p_0} F(p) = \frac{1}{2\pi i} \oint_{\gamma} F(p) dp .$$

До прикладу, залишок $F(p)$ в простому полюсі a визначається як:

$$\text{res}_{p_0} F(p) = \lim_{p \rightarrow a} [F(p) \cdot (p - a)] .$$

Для розв'язання практичних задач важливе значення має обернене перетворення дробово-раціональних функцій виду $F(p) = \frac{P_m(p)}{Q_n(p)}$, де $P_m(p), Q_n(p)$ – многочлени степенів відповідно m і n ($m < n$), p_1, p_2, \dots, p_r – корені многочлена $Q_n(p)$ з кратностями l_1, l_2, \dots, l_r ($l_1 + l_2 + \dots + l_r = n$).

Нагадаємо загальну властивість алгебраїчних рівнянь, які представляються в такому загальному виді

$$Q_n(p) = p^n + a_1 p^{n-1} + \dots + a_n = 0 \quad (3.5)$$

Корень рівняння (3.5) називається *коренем многочлена* $Q_n(p)$. Якщо p_1 – корень многочлена $Q_n(p) = 0$, то $Q_n(p)$ ділиться на $(p - p_1)$ без залишку. Якщо

$Q_n(p)$ ділиться на $(p - p_1)^{l_1}$, але вже не ділиться на $(p - p_1)^{l_1+1}$, то p_1 називається l_1 -кратним коренем рівняння $Q_n(p) = 0$. В цьому випадку p_1 є спільним коренем многочлена $Q_n(p)$ та його похідних до $(l_1 - 1)$ порядку включно. Однократний корінь рівняння (3.5) називають ще *простим коренем*.

Основна теорема алгебри стверджує: будь-яке рівняння n -того степеня, коефіцієнти якого дійсні чи комплексні числа, має n коренів дійсних чи комплексних, якщо k -кратні корені рахувати за k коренів. Отже многочлен $Q_n(p)$ можна представити у вигляді

$$Q_n(p) = (p - p_1)^{l_1} \cdot (p - p_2)^{l_2} \cdot \dots \quad (3.6)$$

З курсу алгебри відома ціла низка способів розкладання многочлена $Q_n(p)$ (3.5) на множники. Нагадаємо ці способи.

1. Спосіб винесення спільного множника за скобки:

$$p^3 + 2p^2 + p = p(p^2 + 2p + 1).$$

2. Розкладання способом групування:

$$ap^3 - bp + bp^3 - ap = p^3(a + b) - p(a + b) = (p^3 - p)(a + b) = p(p^2 - 1)(a + b).$$

3. Розкладання за формулами множення:

$$(p^2 \pm 2p + a^2) = (p \pm a)^2,$$

$$(p^2 - a^2) = (p - a)(p + a),$$

$$(p^3 \pm 3p^2a + 3pa^2 \pm a^3) = (p \pm a)^3,$$

$$(p^3 \pm a^3) = (p \pm a)(p^2 - pa + a^2).$$

4. Заміна одного з членів рівною йому сумою двох членів

$$p^2 + 7p + 12 = p^2 + 3p + 4p + 12 = p(p + 3) + 4(p + 3) = (p + 3)(p + 4).$$

5. Додавання до многочлена двох протилежних членів

$$p^2 + 6p + 8 = p^2 + 6p + 9 - 9 + 8 = (p + 3)^2 - 1 = (p + 2)(p + 4).$$

6. Розклад квадратного тричлена на лінійні множники, тобто представлення у вигляді добутку двочленів першого ступеня. Для рівнянь виду $ap^2 + bp + c = 0$

корені знаходять як $p_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. Тоді квадратний тричлен представляється добутком

$$ap^2 + bp + c = (p - p_1)(p - p_2).$$

Можливе застосування перелічених способів розкладання многочлена в комбінації один з одним.

Повернімось до виразу (3.4). На підставі відомої загальної формули для обчислення лишку

$$\text{res}\varphi(z_0) = \frac{1}{(n-1)!} \lim_{z \rightarrow z_0} \frac{d^{n-1}}{dz^{n-1}} \left\{ (z - z_0)^n \varphi(z) \right\} \quad (3.7)$$

запишемо (3.4) у такому виді

$$f(t) = \sum_{k=1}^r \frac{1}{(l_k - 1)!} \lim_{p \rightarrow p_k} \frac{d^{l_k - 1}}{dp^{l_k - 1}} \left\{ (p - p_k)^{l_k} e^{pt} F(p) \right\}, \quad (3.8)$$

де підсумовування виконується по всіх полюсах p_k зображення $F(p)$, тобто по усіх нулях многочлена $Q_n(p)$.

У випадку, коли усі полюси p_k прості, формула (3.8) спрощується:

$$f(t) = \sum_{k=1}^r \lim_{p \rightarrow p_k} \left\{ (p - p_k) e^{pt} F(p) \right\}. \quad (3.9)$$

Ще більше спростити визначення $f(t)$ можна, якщо усі корені p_1, p_2, \dots, p_r многочлена $Q_n(p)$ прості. Тоді є можливість скористатись формулою

$$f(t) = \sum_{k=1}^r \frac{P_m(p_k)}{Q'_n(p_k)} e^{p_k t}. \quad (3.10)$$

На практиці при відшуванні оригіналу не дуже складного дробово-раціонального зображення достатньо розкласти заданий дріб у суму

де $f^{(k)}(0) = \lim_{t \rightarrow 0+0} f^{(k)}(t)$, $k = 0, 1, 2, \dots, n-1$

У відповідності до теореми лінійності отримаємо

$$x^{(n)}(t) + a_1 x^{(n-1)}(t) + a_2 x^{(n-2)}(t) + \dots + a_n x(t) \rightarrow p^n X(p) - p^{n-1} c_0 - \dots - c_{n-1} + a_{n-2} [p^2 X(p) - p c_0 - c_1] + a_{n-1} [p X(p) - c_0] + a_n X(p). \quad (3.14)$$

Оскільки $f(t) \rightarrow F(p)$ згідно з теоремою про єдність оригіналу у просторі зображень отримаємо алгебраїчне рівняння

$$p^n X(p) - p^{n-1} c_0 - \dots - c_{n-1} + \dots + a_{n-1} [p X(p) - c_0] + a_n X(p) = F(p), \quad (3.15)$$

або

$$(p^n + \dots + a_{n-1} p + a_n) X(p) = F(p) + p^{n-1} c_0 + \dots + c_{n-1} + \dots + a_{n-1} c_0. \quad (3.16)$$

Рівняння (3.16) називається операторним рівнянням диференціального рівняння (3.11) з початковими умовами (3.12). Це рівняння є алгебраїчним відносно функції $X(p)$. З рівняння (3.16) отримаємо зображення

$$X(p) = \frac{F(p) + p^{n-1} c_0 + \dots + c_{n-1} + \dots + a_{n-1} c_0}{(p^n + \dots + a_{n-1} p + a_n)}. \quad (3.17)$$

Далі, згідно з теоремою розкладання та користуючись властивостями перетворення Лапласа знаходимо для зображення $X(p)$ оригінал $x(t)$. Неперервна функція $x(t)$, згідно теореми про єдність оригіналу, є рішенням рівняння (3.11), яке задовольняє початкові умови (3.12).

Лекція 4.

ЗАСТОСУНКИ ОПЕРАЦІЙНОГО ЧИСЛЕННЯ

Перетворення Лапласа і розвинене на його основі операційне числення є зручним інструментом розв'язування різних задач в математиці, фізиці, електротехніці, радіотехніці, теорії автоматичного регулювання тощо. Такий

спосіб розв'язування технічних задач отримав назву *операційного методу*. Розглянемо деякі за стосунки цього методу.

4.1. Застосування операційного числення до розв'язання звичайних диференціальних рівнянь

Постановка задачі. Нехай маємо диференціальне рівняння

$$x^{(n)} + a_1 x^{(n-1)} + \dots + a_n x = f(t) \quad (4.1)$$

де a_1, a_2, \dots, a_n – коефіцієнти рівняння (задані числа), $f(t)$ – задана функція, яка є оригіналом, $x = x(t)$ – невідома функція. Треба знайти розв'язок рівняння (4.1), який задовольняє початковим умовам:

$$x(0) = x_0, x'(0) = x'_0, \dots, x^{(n-1)}(0) = x_0^{(n-1)}, \quad (4.2)$$

де $x_0, x'_0, \dots, x_0^{(n-1)}$ – задані числа.

Розв'язання задачі. Методика пошуку розв'язок рівняння операційним методом полягає у наступному.

1. Здійснюється перехід від функцій $x(t)$ та $f(t)$ до їх зображень : перехід $x(t) \rightarrow X(p)$ здійснюється формально, а $f(t) \rightarrow F(p)$ – згідно з фактичним виразом функції $f(t)$.

2. Застосовуються теореми про диференціювання оригіналу і лінійність і здійснюється перехід від рівняння (4.1) до рівняння виду

$$L(p) \cdot X(p) + Q(p) = F(p). \quad (4.3)$$

яке називається *зображуючим* або *операторним рівнянням*. Многочлен

$L(p) = p^n + a_1 p^{n-1} + \dots + a_n$ називається *характеристичним многочленом* рівняння

(4.1), $Q(p)$ – многочлен степеня $(n-1)$, коефіцієнти якого залежать від початкових

даних $x_0, x'_0, \dots, x_0^{(n-1)}$ На відміну від диференціального рівняння (4.1), рівняння

(4.3) є лінійним алгебраїчним відносно $X(p)$.

3. Рівняння (4.3) розв'язується відносно $X(p)$

$$X(p) = \frac{F(p) - Q(p)}{L(p)}. \quad (4.4)$$

За нульових початкових умов $x_0 = x'_0 = \dots = x_0^{(n-1)} = 0$ маємо $Q(p) = 0$ і вираз (4.4) спрощується

$$X(p) = F(p)/L(p). \quad (4.4)$$

4. Здійснюється зворотний перехід від зображення $X(p) \rightarrow x(t)$. Функція $x(t)$ і є розв'язком рівняння (4.1).

Функція $Z(p) = 1/L(p)$ називається *передавальною функцією*. Якщо розуміти $F(p)$ як зображення вхідного сигналу, $X(p)$ – як зображення реакції системи на це збурення, то зображення відгуку отримують шляхом множення з

$$X(p) = Z(p)F(p). \quad (4.5)$$

Проілюструємо використання наведеної методики наступним прикладом.

Приклад 4.1. Знайти розв'язок диференціального рівняння

$$x'' + 4x' + 4x = e^{-2t} (\cos t + 2 \sin t),$$

що задовольняє початковим умовам $x(0) = -1$, $x'(0) = 1$.

Розв'язання. Нехай $x(t) \rightarrow X(p)$. Тоді згідно з теоремою про диференціювання оригіналу та з урахуванням початкових умов маємо

$$x'(t) \rightarrow pX(p) + 1, \quad x''(t) \rightarrow p^2X(p) + p - 1.$$

За таблицею відповідності оригіналів і зображень і теоремою лінійності знаходимо

$$(\cos t + 2 \sin t) \rightarrow \frac{p}{p^2 + 1} + 2 \frac{1}{p^2 + 1} = \frac{p + 2}{p^2 + 1}.$$

Далі, за теоремою зсунення маємо

$$e^{-2t} (\cos t + 2 \sin t) \rightarrow \frac{p + 4}{(p + 2)^2 + 1}.$$

З урахуванням отриманих зображень представимо операторне рівняння як

$$p^2 X(p) + p - 1 + 4pX(p) + 4 + 4X(p) = \frac{p+4}{(p+2)^2 + 1}$$

і після перетворень

$$(p^2 + 4p + 4)X(p) + p + 4 = \frac{p+4}{(p+2)^2 + 1}.$$

З останнього рівняння знаходимо зображення розв'язку

$$X(p) = -\frac{p^3 + 7p^2 + 16p + 11}{\left[(p+2)^2 + 1\right](p+2)^2}.$$

За результатами розкладання дробу у суму найпростіших дроби матимемо

$$X(p) = -\frac{p+2}{(p+2)^2 + 1} - 2\frac{1}{(p+2)^2 + 1} + \frac{1}{(p+2)^2}.$$

З урахуванням того, що $\cos t \rightarrow \frac{p}{p^2 + 1}$, $\sin t \rightarrow \frac{1}{p^2 + 1}$, $t \rightarrow \frac{1}{p^2}$ отримаємо

такий розв'язок

$$x(t) = e^{-2t} (t - \cos t - 2 \sin t).$$

4.2. Розв'язання систем звичайних лінійних диференціальних рівнянь зі сталими коефіцієнтами

Розв'язування систем звичайних лінійних диференціальних рівнянь принципово не відрізняється від розв'язування одного рівняння. Відмінність лише в тому, що замість одного операторного рівняння отримуємо систему операторних рівнянь. Кожне з цих рівнянь є лінійним алгебраїчним відносно зображення оригіналу. Розв'язавши його отримуємо зображення оригіналу, після чого переходимо до оригіналу. Проілюструємо сказане наступним прикладом

Приклад 4.2. Знайти розв'язок системи звичайних лінійних диференціальних рівнянь

$$\begin{cases} x' - x - 2y + 9t = 0, \\ y' - 2x - y = 4e^t \end{cases}$$

з початковими умовами $x(0) = 0$, $y(0) = 2$.

Розв'язання. Введемо позначення: $x(t) \rightarrow X(p)$, $y(t) \rightarrow Y(p)$. Оскільки за таблицею відповідності оригіналів та зображень маємо $t \rightarrow \frac{1}{p^2}$, $e^t \rightarrow \frac{1}{p-1}$, то з урахуванням початкових умов та теоремою про диференціювання і теоремою лінійності отримаємо таку систему операторних рівнянь відносно зображень $X(p), Y(p)$

$$\begin{cases} (p-1)X(p) - 2Y(p) = 1 - \frac{9}{p}, \\ -2X(p) + (p-1)Y(p) = 2 + \frac{4}{p-1}. \end{cases}$$

Визначимо $X(p), Y(p)$ за формулою Крамера: $X(p) = \frac{\Delta x}{\Delta}$, $Y(p) = \frac{\Delta y}{\Delta}$, де детермінант другого порядку обчислюється як

$$\Delta = \begin{vmatrix} p-1 & -2 \\ -2 & p-1 \end{vmatrix} = p^2 - 2p - 3 = (p-3)(p+1).$$

Тоді маємо

$$\Delta x = \begin{vmatrix} \frac{p^2-9}{p^2} & -2 \\ \frac{2p+2}{p-1} & p-1 \end{vmatrix} = \frac{p^4 + 2p^3 - 4p^2 + 18p - 9}{p^2(p-1)},$$

$$\Delta y = \begin{vmatrix} p-1 & \frac{p^2-9}{p^2} \\ -2 & \frac{2p+2}{p-1} \end{vmatrix} = \frac{2p^3 + 4p^2 - 18}{p^2}.$$

Отже зображення оригіналів визначаємо як

$$X(p) = \frac{\Delta x}{\Delta} = \frac{p^4 + 2p^3 - 4p^2 + 18p - 9}{p^2(p-1)(p-3)(p+1)},$$

$$Y(p) = \frac{\Delta y}{\Delta} = \frac{2p^3 + 4p^2 - 18}{p^2(p-3)(p+1)}.$$

Розкладемо зображення у суму найпростіших дробів, користуючись методом невизначених коефіцієнтів. Для $X(p)$ матимемо

$$X(p) = \frac{p^4 + 2p^3 - 4p^2 + 18p - 9}{p^2(p-1)(p-3)(p+1)} = \frac{A}{p^2} + \frac{B}{p} + \frac{C}{p-1} + \frac{D}{p-3} + \frac{E}{p+1}.$$

З цього рівняння отримаємо наступну систему:

$$\begin{cases} B + C + D + E = 1, \\ A - 3B - 2C - 4E = 2, \\ -3A - B - 3C - D + 3E = -4, \\ -A + 3B = 18, \\ 3A = -9. \end{cases}$$

Розв'язання цієї системи дає такі результати:

$$A = -3, B = 5, C = -2, D = 2, E = -4.$$

З урахуванням визначених коефіцієнтів запишемо зображення $X(p)$ у такій формі

$$X(p) = -\frac{3}{p^2} + \frac{5}{p} - \frac{2}{p-1} + \frac{2}{p-3} - \frac{4}{p+1}.$$

В аналогічний спосіб визначимо зображення $Y(p)$

$$Y(p) = \frac{6}{p^2} - \frac{4}{p} + \frac{2}{p-3} + \frac{4}{p+1}.$$

Використовуючи таблицю відповідності зображень та оригіналів та теорему лінійності запишемо розв'язок системи диференціальних рівнянь

$$x(t) = -3t + 5 - 2e^t + 2e^{3t} - 4e^{-t}, \quad y(t) = 6t - 4 + 2e^{3t} + 4e^{-t}.$$

4.3. Застосування операційного числення для розв'язання задач електротехніки та теорії керування

Методи операційного числення для аналізу процесів, що відбуваються в електричних колах, виявились дуже зручними та ефективними. Вперше їх запропонував відомий англійський фізик О. Хевісайд.

Застосування перетворення Лапласа для розрахунку перехідних процесів в електричних контурах розглянемо на прикладі підключення послідовно з'єднаних елементів резистора R , котушки індуктивності L та конденсатора C до джерела напруги E (рис. 4.1). Початкові умови: струм в колі $i(0) = 0$, заряд на обкладинках конденсатора $Q(0) = 0$, прикладена електрорушійна сила $e(t)$.

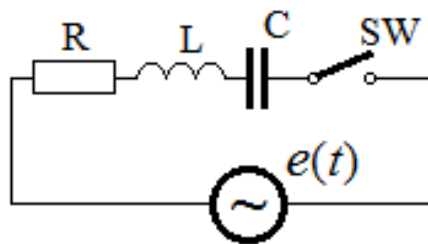


Рис. 4.1. Дослідження перехідних процесів в послідовному RLC контурі

Диференціальне рівняння, що описує зміну сили струму в електричному колі має наступний вид

$$L \frac{di(t)}{dt} + Ri(t) + \frac{Q(t)}{C} = e(t). \quad (4.6)$$

Зв'язок заряду та електричного струму подають наступні формули

$$\frac{dQ(t)}{dt} = i(t), \quad Q(0) = 0 \Rightarrow Q(t) = \int_0^t i(\tau) d\tau. \quad (4.7)$$

З урахуванням (4.7) рівняння (4.6) набуває такого вигляду

$$L \frac{di(t)}{dt} + Ri(t) + \frac{1}{C} \int_0^t i(\tau) d\tau = e(t). \quad (4.8)$$

В (4.7) перша складова – це падіння напруги на індуктивності, друга – падіння напруги на опорі, третя складова – падіння напруги на ємності.

Введемо позначення зображень: $e(t) \rightarrow E(p)$, $i(t) \rightarrow I(p)$. Тоді операторне рівняння електричного кола запишеться у такому виді

$$LpI(p) + RI(p) + \frac{1}{Cp}I(p) = E(p). \quad (4.9)$$

З (4.9) визначимо вираз для зображення сили струму в електричному колі

$$I(p) = \frac{E(p)}{R + Lp + \frac{1}{Cp}}. \quad (4.10)$$

Означення. *Операторний опір* $Z(p)$ електричного кола – це відношення операторного зображення напруги до операторного зображення струму, що діють на елементі або ділянці кола. Має розмірність Ом. Для виразу (4.10) це

$$Z(p) = \frac{E(p)}{I(p)} = R + Lp + \frac{1}{Cp}. \quad (4.11)$$

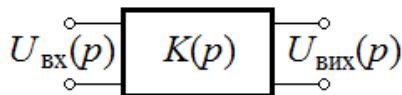
Означення. Величина, обернена операторному опору, називається *операторною провідністю* (має розмірність См – сіменс)

$$Y(p) = 1/Z(p). \quad (4.12)$$

Тоді закон Ома в операторній формі запишеться як

$$I(p) = E(p)/Z(p). \quad (4.13)$$

Означення. *Передавальною функцією* $K(p)$ електричного чотирьополосника називається відношення зображення напруги на його виході до зображення напруги на вході (рис. 4.2)



$$K(p) = \frac{U_{\text{вих}}(p)}{U_{\text{вх}}(p)}. \quad (4.14)$$

Рис. 4.2. Визначення передавальної функції

Користуючись теоремою обернення або її наслідками з (4.10) можна знайти силу струму як функцію часу $i(t)$.

Приклад 4.3. Виконати аналіз перехідного процесу в RC ланцюжку (рис. 4.3а), на який подається одиничний імпульс напруги (рис. 4.3б).

Необхідно знайти закон зміни напруги на конденсаторі $u_C(t)$ за умови, що напруга на ньому в момент подачі імпульсу дорівнює 0, тобто $u_C(0) = 0$.

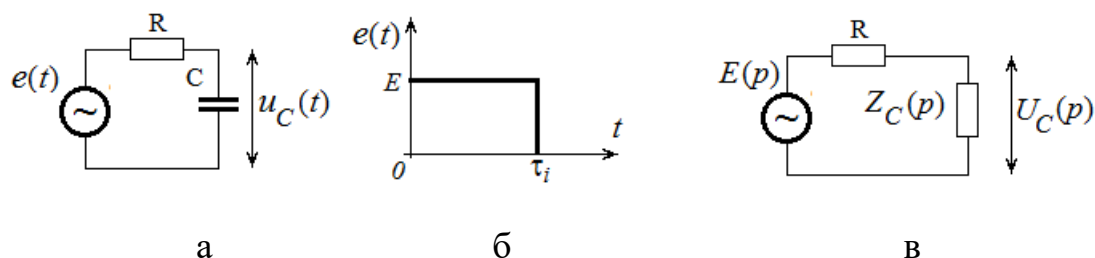


Рис. 4.3. Електрична ланка (а), одиничний імпульс напруги (б) та схема заміщення (в)

Розв'язання. Будемо шукати розв'язок за допомогою передавальної функції ланцюга. Скористаємось схемою заміщення ланцюга, яка представлена на рис. 4.3в. Передавальна функція визначається виразом

$$K(p) = \frac{U_C(p)}{E(p)} = \frac{(pC)^{-1}}{R + (pC)^{-1}} = \frac{1}{pRC + 1} = \frac{1}{p\tau_0 + 1}.$$

Для аналітичного запису вхідного сигналу (рис. 4.3б) скористаємось принципом суперпозиції: представимо $e(t)$ сумою двох більш простих сигналів $e(t) = E \cdot 1(t) - E \cdot 1(t - \tau_i)$ після чого знайдемо відгук на кожний з них і складемо отримані результати.

1. Знайдемо зображення по Лапласу першої складової вхідного сигналу

$$e_1(t) = E \cdot 1(t) \rightarrow E_1(p) = E/p.$$

2. Визначимо зображення 1-го елементарного відгуку

$$U_{c1}(p) = E_1(p) \cdot K(p) = E/p(p\tau_0 + 1).$$

3. Визначимо оригінал відгуку від першої складової вхідного сигналу.

Для цього знайдемо залишки по двом полюсам: $p_1 = 0$, $p_2 = -1/\tau_0$.

$$u_{C1}(t) = \lim_{p \rightarrow 0} \frac{Epe^{pt}}{p(p\tau_0 + 1)} + \lim_{p \rightarrow -\frac{1}{\tau_0}} \frac{E(p + 1/\tau_0)e^{pt}}{p(p\tau_0 + 1)} = E \cdot 1(t) + \frac{E}{\tau_0} \cdot \left(\frac{1}{-1/\tau_0} \right) \cdot e^{\frac{t}{\tau_0}} =$$

$$= E \cdot 1(t) - Ee^{\frac{t}{\tau_0}} \cdot 1(t) = E \cdot \left(1 - e^{\frac{t}{\tau_0}} \right) \cdot 1(t).$$

4. В аналогічний спосіб визначимо оригінал відгуку від другої складової вхідного сигналу

$$u_{C2}(t) = E \cdot \left(1 - e^{\frac{t-\tau_0}{\tau_0}} \right) \cdot 1(t - \tau_0).$$

5. Визначимо відгук на прямокутний імпульс як суму відгуків від двох складових

$$u_C(t) = u_{C1}(t) + u_{C2}(t) = E \cdot \left(1 - e^{\frac{t}{\tau_0}} \right) \cdot 1(t) + E \cdot \left(1 - e^{\frac{t-\tau_0}{\tau_0}} \right) \cdot 1(t - \tau_0).$$

З отриманого результату витікає, що на інтервалі часу $[0, \tau_0)$ конденсатор заряджається і напруга на ньому збільшується за експоненціальною функцією, а на інтервалі $[\tau_0, \infty)$ – розряджається, напруга експоненціально спадає.

Питання для самоконтролю

1. Які форми представлення комплексних чисел вам відомі?
2. Які функції можуть бути оригіналами для перетворення Лапласа?
3. Дайте визначення зображення функції-оригіналу перетворення Лапласа.
4. Отримайте зображення по Лапласу функції Хевісайда.
5. Згортка функцій та її властивості.
6. Наведіть теореми додавання, подібності, запізнювання для перетворення Лапласа та поясніть зміст всіх складових (функцій та аргументів).
7. Наведіть теореми випередження та зміщення для перетворення Лапласа та поясніть зміст всіх складових (функцій та аргументів).

8. Інтеграл Дюамеля та теорема про його зображення по Лапласу. Використання інтеграла Дюамеля в електротехнічних розрахунках.
9. Теорема обернення. Її зміст та використання для визначення функції-оригіналу за її зображенням.
10. Перша теорема розкладання, її зміст та використання для визначення зображень.
11. Друга теорема розкладання, її зміст та використання для визначення зображень.
12. Способи розкладання многочлена $Q_n(p)$ на множники.
13. Наведіть формули для визначення функції-оригіналу за полюсами зображення та розшифруйте зміст всіх складових.
14. Як розкласти зображення по Лапласу на прості дроби?
15. З якою метою проводиться розкладання зображень по Лапласу на прості дроби?
16. Операторний метод розв'язання звичайних диференціальних рівнянь.
17. Викладіть алгоритм розв'язання систем звичайних лінійних диференціальних рівнянь зі сталими коефіцієнтами.
18. Застосування операційного числення для розв'язання задач електротехніки та теорії управління.
19. Означення та зміст понять операторний опір та операторна провідність.
20. Передавальна функція та її використання в електротехнічних розрахунках.
21. Викладіть алгоритм аналізу перехідних процесів операторним методом.

Розділ 2. ОСНОВИ ТЕОРІЇ ЙМОВІРНОСТЕЙ

Лекція 5.

ОСНОВНІ ПОНЯТТЯ ТА ВИЗНАЧЕННЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ

5.1. Загальні поняття та визначення

Теорію ймовірностей називається математична наука, що вивчає закономірності у випадкових явищах.

Випадкове явище – це таке явище, яке за умови багаторазового відтворення одного й того ж досліду (випробування, експерименту) перебігає кожного разу дещо в інший спосіб. Для них характерна значна степінь невизначеностей, непередбачуваностей.

Приклади випадкових явищ:

1. Повторне зважування одного і того ж тіла (або будь-яке інше вимірювання);
2. Випробування на визначення для серії однотипних виробів часу безвідмовної роботи;
3. Стрільба з однієї гармати по одній цілі;
4. Відхилення центра маси літака від заданої траєкторії під час польоту;
5. Досліди з підкиданням монети чи гральних кісток;
6. Контроль партії виробів одним з методів НК.

Первинним поняттям теорії ймовірностей є поняття множини.

Множина – набір, сукупність, збірка будь-яких об'єктів, що називаються його елементами, і які володіють спільною для всіх елементів характеристичною властивістю.

Георг Кантор (1845 –1918), німецький математик, який створив основи теорії множин, дав таке означення: «Множество есть многое, мыслимое как единое» (Множина є багато чого, що мислиться як єдине).

Множини можуть мати найрізноманітнішу природу.

Відомо 3 види множин: скінченні, нескінченні і пуста множина, тобто множина, що не містить жодного елемента, позначається як \emptyset .

Для порівняння множин вводиться поняття потужності множини:

Потужність множини – це узагальнене поняття кількості її елементів.

Позначення множин точок на прямій:

$[\alpha, \beta]$ – відрізок (включає межові точки);

(α, β) – інтервал;

$(\alpha, \beta]$ – проміжок;

$[\alpha, +\infty)$ – проміжок.

Числові множини:

N – множина натуральних чисел;

R – множина раціональних чисел.

Порівнювати множини можна за їх потужностями.

Г. Кантор визначав кардинальне число множини A як його властивість, яка лишається після абстрагування від якості елементів множини і їх порядку у множині. Позначення кардинального числа множини A : \bar{A} або $card A$. Якщо A – скінченна множина з « n » елементів, то $card A = n$.

Потужність всіх дійсних чисел дорівнює потужності континуума (від лат. *continuum* – безперервне, суцільне), що позначається так: $card A = c$, де c – потужність континуум.

Дві множини A і B називаються *рівнопотужними*, якщо існує взаємно однозначна функція $f: A \rightarrow B$ з областю визначення A і множиною значень B (тобто коли кожному елементу з A можна поставити у відповідність елементи з B). Цей факт записує так:

$$card A = card B$$

5.2. Випадкові події

Будь-яка наука, що розвиває загальну теорію певного кола явищ, ґрунтується на певному колі понять та визначеностей. Це стосується і теорії ймовірностей.

Дослід (експеримент, випробування) – деяка відтворювана сукупність умов, за

виконання яких спостерігається те чи інше явище, фіксується той чи інший результат. Дослід може проводитись як людиною, так і відбуватись незалежно від дій людини.

Якщо результати досліду варіюються за його повторювання, кажуть про дослід з *випадковим наслідком*. Саме такі досліди є *предметом вивчення теорії ймовірностей*.

Результат проведеного досліду називається *подією* (елементарною подією). За результатами досліду подія може відбутись чи не відбутись. Елементарні події позначатимемо ω_i .

Сукупність всіх елементарних подій: $\Omega = \{\omega_1, \dots, \omega_n\}$ утворює простір елементарних подій. Запис $\omega_i \in \Omega$ читається як: « ω_i є елементом простору Ω ». Математичний об'єкт Ω можна розглядати як достовірну подію.

Множину можна задати перелічивши всі її елементи, або визначивши правило, за яким визначається приналежність елемента до множини.

Приклад 5.1. Під час однократного підкидання монети, простір елементарних подій задається наступним чином:

$$\Omega = \{\Gamma, P\}, \quad \text{де } \Gamma - \text{ герб, } P - \text{ решітка.}$$

Приклад 5.2. У випадку n -кратного підкидання монети маємо

$$\Omega = \{\omega: \omega = (a_1, \dots, a_n), a_i = \Gamma \text{ чи } P\}$$

Читається як « Ω – це простір елементарних подій, утворених сукупністю наслідків (a_1, \dots, a_n) , кожен з яких може набути значення Γ чи P ». В цьому випадку загальне число наслідків дорівнює 2^n .

Поряд з поняттям простору елементарних подій введемо більш складне поняття – поняття події. Необхідність його введення пов'язане з тим, що експериментатора зазвичай цікавить не те, який конкретно наслідок має місце за результатами випробування, а те, чи належить наслідок певній підмножині (тобто частині) множини Ω . Позначення $A \subset \Omega$ читається як: «подія A є підмножиною Ω (частиною Ω)», або « A міститься у Ω ». Знак \subset – знак приналежності, знак «включення».

5.3. Дії з множинами та їх позначення

$\omega_i \in A$ – елемент ω_i належить підмножині (події) A ;

$\omega_i \notin A$ – елемент ω_i не належить A ;

$A = \emptyset$ – A є пустою множиною, тобто множиною, що не має жодного елемента (це аналог задання 0 на числовій осі);

$A = B$ – рівність множин, має місце коли $A \subset B$ і $B \subset A$ (тобто A є підмножиною B і навпаки), тобто A і B складаються з одних і тих самих елементів;

$A \cup B$ – об'єднання множин – це множина, що складається з елементів, що належать хоча б одній з множин A і B ;

$A \cap B$ – переріз (пересечение) множин – це множина, що складається з елементів, що належать як A , так і B .

Операції об'єднання і перерізу комутативні, асоціативні та взаємодистрибутивні.

$\bar{A} = B \setminus A = \{\omega_i \in B : \omega_i \notin A\}$ – доповнення множини A ; пуста множина: $\emptyset = A \setminus A$.

$A^c = \Omega \setminus A$ – доповнення підмножини A до Ω

Операції об'єднання, перерізу і доповнення пов'язані законами де Моргана, наприклад: $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

Для наочного зображення множин і операцій з ними застосовують круги Ейлера (діаграми Венна). Ймовірнісне трактування теоретико-множинних операцій представлено та їх графічне зображення за допомогою діаграм Венна в Таблиці 1.1.

Розглянемо ряд означень, пов'язаних з поняттям «подія».

1. *Ймовірністю події* називається міра степені об'єктивної можливості цієї події. Позначатимемо $P(A)$ – ймовірність події A ; ймовірність довільної події A належить інтервалу $[0, 1]$, тобто $0 \leq P(A) \leq 1$.

Таблиця 1.1. Ймовірнісне трактування теоретико-множинних операцій

Теоретико-множинні об'єкти та операції	Ймовірнісне трактування	Геометричне представлення
Ω – множина	простір елементарних подій, достовірна подія	
ω_i – елемент Ω , $\omega_i \in \Omega$	елементарна подія	
A – підмножина множини Ω	подія	
\emptyset – пуста множина	неможлива подія	
$A \subset B$ – підмножина A є частиною підмножини B	подія A тягне подію B	
A^c – доповнення підмножини A до Ω	подія A не відбулась	
$A \cup B$ – об'єднання підмножин A і B	відбулась принаймні одна з подій – A чи B	
$A \cap B$ – переріз підмножин A і B	відбулись одночасно дві події – A і B	
$B \setminus A$ – різниця підмножин B і A	відбулась подія B в той час як подія A не відбулась	
$A \cap B = \emptyset$ – множини A і B не мають спільних елементів	події A і B не сумісні	

2. *Вірогідною* називається подія U , яка за результатами досліду неодмінно повинна відбутись: $P(U)=1$;

3. *Неможливою* називається подія V , яка за результати досліду не може відбутись: $P(V)=0$;

4. Дві події називаються *несумісними*, якщо в ході випробувань поява однієї

з них виключає можливість появи іншої. Наприклад, попадання з одного пострілу у 2 цілі.

5. Дві події називаються *сумісними*, якщо в ході випробувань поява однієї з них не виключає можливість появи іншої. До прикладу, попадання артилерійським залпом (10 пострілів) в 2,3,...цілі.

6. Події називаються *єдино можливими*, коли в ході випробувань неминуче станеться хоча б одна з цих подій. Приклад: події за наслідком кидання монети, кубиків.

7. Дві єдино можливі несумісні події називаються *протилежними*. Приклад: події за наслідком кидання монети. Позначаються: A і \bar{A} .

8. Якщо за наслідками ряду випробувань можуть з'явитись декілька можливих подій і немає підстав припускати, що поява одних подій є більш можливою ніж інших. Такі події називаються *рівно можливими*. Приклад: кидання грального кубіка, монети.

9. *Повна* група подій. Декілька подій в певному досліді утворюють повну групу подій, якщо за результатами досліді неминуче повинна з'явитись хоча б одна з них.

Класичне визначення ймовірності деякої події A дається формулою

$$P(A) = m/n, \quad (5.1)$$

де m – кількість елементарних наслідків досліді, що сприяють появі події A , n – число всіх можливих елементів простору Ω . Отже дослідження теоретико-ймовірнісних схем із скінченим числом рівноймовірних наслідків зводиться до розв'язання задач комбінаторики.

Існує і інше, статистичне, означення ймовірності події A , яке ґрунтується на результатах ряду повторюваних випробувань. В його основі лежить поняття відносної частоти події:

$$W(A) = m(n)/n, \quad (5.2)$$

де $m(n)$ – число випробувань, в яких настала подія A за загального числа випробувань n .

За *статистичну ймовірність* приймається відносна частота події A за $n \rightarrow \infty$:

$$P(A) = \lim_{n \rightarrow \infty} W(A). \quad (5.3)$$

Класичне визначення ймовірності використовують для аналізу теоретичних моделей, а статистичне – для опрацювання експериментальних даних. До прикладу, при оцінюванні випадання решітки при підкиданні монети (за припущення про рівні ймовірності подій) $P(A) = 1/2$; якщо при підкиданні монети в 40 дослідах отримали 19 решіток, то маємо $W(A) = 19/40$, отже приймаємо $P(A) \approx 19/40$.

5.4. Елементи комбінаторики

Комбінаторика – це розділ математики, який вивчає найпростіші «поєднання» елементів множин.

1. **Перестановками** з « n » елементів називаються такі їх поєднання, які відрізняються одне від одного тільки порядком елементів. Число перестановок

$$P_n = n!. \quad (5.4)$$

Факторіали великих чисел можуть бути наближено обчислені за *формулою Стірлінга*:

$$n! \approx \sqrt{2\pi n} \times n^n \times e^{-n}, \quad n \gg 1. \quad (5.5)$$

2. **Розміщенням** (вибір без повернення) з « n » елементів по « m » називають такі їх поєднання, які відрізняються один від одного самими елементами, або їх порядком.

До прикладу, розміщенням з 3 елементів a, b, c по 2: ab, ac, bc, ba, ca, cb . Подібні комбінації утворюються, наприклад, у випадку послідовного вибору без повернення елементів a_{i1}, \dots, a_{im} з деякої загальної сукупності a_1, a_2, \dots, a_n .

Число всіх розміщень визначається формулою

$$A_n^m = \frac{n!}{(n-m)!}. \quad (5.6)$$

3. **Сполучення** з « n » елементів по « m » називають їх поєднання, що

відрізняються один від одного самими елементами. Тобто в комбінаціях a_{i_1}, \dots, a_{i_m} , складених із загальної сукупності a_1, \dots, a_n обсягу n не враховується порядок елементів, так що комбінації з одними і тими самими елементами вважаються рівними. До прикладу, сполучення з 3 елементів a, b, c по 2: $ab; ac; bc$;

Число всіх сполучень визначається формулою

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n!}{m!(n-m)!} . \quad (5.7)$$

Властивість: $C_n^m = C_n^{n-m}$.

4. Вибір елементів з поверненням. Нехай з певної сукупності n елементів a_1, a_2, \dots, a_n виконується вибір, за якого послідовно вибирається один з елементів a_i , який щоразу повертається у загальну сукупність. За r кроків реєструється вибірка $a_{i_1}, a_{i_2}, \dots, a_{i_r}$. Число можливих комбінацій (вбірок) дорівнює

$$K = n^r. \quad (5.8)$$

Цей результат легко перевірити на прикладі кидання монети.

5.5. «Геометричні» ймовірності

«Геометричні» ймовірності застосовують для обчислення ймовірності на множинах нескінченних обсягів. Такі множини є рівнопотужними з множинами точок прямої або деякої фігури чи об'єму. Це дає змогу обчислювати ймовірності через відношення довжин відрізків, площ фігур або об'ємів тривимірних тіл.

Приклад 5.3. Нехай на відрізок довжиною L деякої прямої навмання «кидається» точка. Якою є вірогідність того, що вона впаде не далі, ніж на відстань l від середини вказаного відрізка? (див. рис. 5.1).

Розв'язання. Подія A – «точка знаходиться від середини на відстані не більше l » – настає в результаті попадання в будь-яку точку відрізка $2l$. Доля таких точок x у всьому відрізку може бути визначена як відношення довжин відрізків

$$L(A)/L = 2l/L .$$

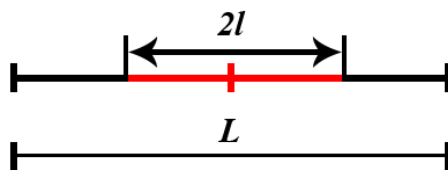


Рис. 5.1. Визначення геометричної ймовірності за довжиною відрізків

Для всіх можливих випадків шукана ймовірність визначається як

$$P(A) = \frac{L(A)}{L} = \begin{cases} \frac{2l}{L}, & \text{якщо } 2l < L, \\ 1, & \text{якщо } 2l \geq L. \end{cases}$$

Приклад 5.4. Нехай на відрізок довжиною L «кидається» навмання і незалежно одна від одної 2 різні точки ξ_1 і ξ_2 . Якою є ймовірність того, що відстань між ними буде не більше l ?

Розв'язання. Розглянемо наступну модель. Координату т. ξ_1 , відкладемо на відрізьку $[0, L]$ осі x_1 , а координату ξ_2 – на $[0, L]$ осі x_2 (рис. 5.2).

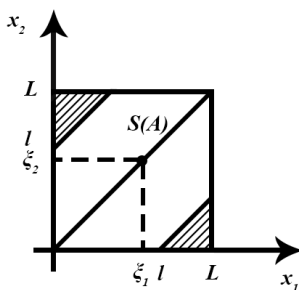


Рис. 5.2. Визначення геометричної ймовірності за площею фігур

Можна вважати, що точка (ξ_1, ξ_2) кидається навмання в квадрат зі стороною L .

Шукана ймовірність співпадає з ймовірністю події A , яка полягає в тому, що випадково кинута точка (ξ_1, ξ_2) попадає в область квадрата, обмежену прямими з рівняннями $x_2 = x_1 \pm l$ (на рис. 5.2 це незашифрована область). Тому ймовірність події A може бути визначена як відношення площ:

$$P(A) = \frac{S(A)}{S} = \frac{S(A)}{L \cdot L} = \frac{L^2 - (L-l)^2}{L^2} = 1 - \left(1 - \frac{l}{L}\right)^2.$$

Лекція 6.

ЙМОВІРНІСНИЙ ПРОСТІР.

ТЕОРЕМИ ДОДАВАННЯ І МНОЖЕННЯ ЙМОВІРНОСТЕЙ ПОДІЙ

6.1. Ймовірнісна модель стохастичних експериментів

Трійка об'єктів $\{\Omega, \mathcal{F}, P\}$, де $\Omega = \{\omega_1, \dots, \omega_n\}$ – простір подій, \mathcal{F} – алгебра множин, тобто непорожня сукупність підмножин з Ω , $\mathcal{F} = \{A: A \subset \Omega\}$, а $P = \{P(A); A \in \mathcal{F}\}$ – ймовірність будь-якої події $A \in \mathcal{F}$ називається ймовірнісною моделлю чи ймовірнісним простором.

Події $A \in \mathcal{F}$ утворюються з елементарних подій за допомогою теоретично-множинних операцій (об'єднання, перерізу, доповнення). Алгебри подій будуються за наступним загальним принципом. Говоритимемо, що система множин $D = \{D_1, \dots, D_n\}$ утворює *розбиття* множин Ω , а $D_i \in$ атомами розбиття, якщо D_i непорожні, попарно несумісні, а їх сума $\sum D_i = \Omega$ (рис. 6.1).

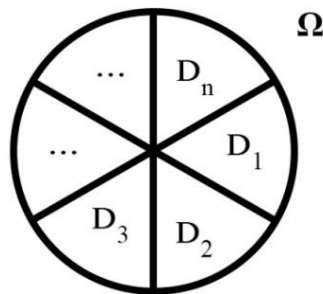


Рис. 6.1. Розбиття множини Ω

Всі можливі об'єднання атомів з D_i разом з пустою множиною утворюють алгебру, породжену розбиттям D . Цей факт позначається таким чином: $\mathcal{F} = \alpha(D)$.

Задана на алгебрі множин ймовірнісна міра (або просто ймовірності) має наступні властивості:

1. $0 \leq P(A) \leq 1$ (невід'ємність).
2. $P(\emptyset) = 0$.
3. $P(\Omega) = 1$.

$$4. P(D_1) + P(D_2) + \dots + P(D_n) = 1 \text{ (нормованість)}$$

$$5. P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

В частинному випадку, якщо $A \cap B = \emptyset$, то:

$$P(A \cup B) = P(A) + P(B); \quad (6.1)$$

$$P(\bar{A}) = P(\Omega \setminus A) = P(\Omega) - P(A) = 1 - P(A). \quad (6.2)$$

6.2. Аксиоми сучасної теорії ймовірностей (аксиоми Колмогорова)

Узагальнюючи наведені вище дані викладемо аксиоми сучасної теорії ймовірності, які покладені в основу побудови ймовірнісного простору.

1. \mathcal{F} є алгеброю множин.

Система \mathcal{F} підмножин множини Ω називається алгеброю, якщо $\Omega \in \mathcal{F}$, об'єднання, перетин і різниця двох множин системи також належить цій системі, тобто якщо $A \in \mathcal{F}$, $B \in \mathcal{F}$ то $A \cap B \in \mathcal{F}$, $A \cup B \in \mathcal{F}$, $A \setminus B \in \mathcal{F}$.

2. Кожній множині $A \in \mathcal{F}$ поставлене у відповідність деяке невід'ємне число $P(A)$, що називається ймовірністю події A .

$$3. P(\Omega) = 1.$$

4. $P(\emptyset) = 0$, де \emptyset – порожня множина, тобто множина, яка не містить жодного елемента.

5. Якщо події A і B несумісні, тобто $A \cap B = \emptyset$, то

$$P(A \cup B) = P(A + B) = P(A) + P(B). \quad (6.3)$$

Сукупність об'єктів $\{\Omega, \mathcal{F}, P\}$, що задовольняє аксиоми 1–5, називається *ймовірнісним простором* і повністю визначає ймовірнісну модель експерименту зі скінченним числом наслідків.

6.3. Умовна ймовірність

Для довільних подій A та B з ймовірністю $P(B) \neq 0$ умовна ймовірність події A за умови, що B відбулася визначається як відношення $P(AB)/P(B)$ і позначається:

$$P(A|B) = \frac{P(AB)}{P(B)}. \quad (6.4)$$

Це визначення можна пояснити такими міркуваннями. Нехай простір подій Ω складається з N елементів, подія A – з n , подія B – з m , а подія AB – з p елементарних подій. Графічну ілюстрацію простору подій подібно на рис. 6.2.

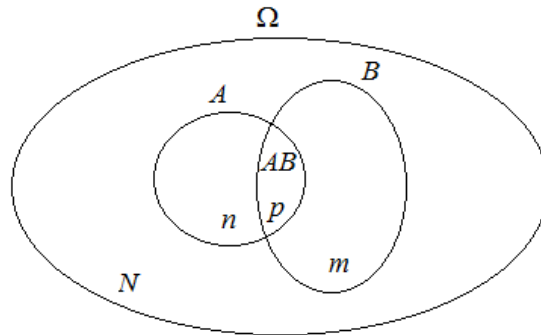


Рис. 6.2. Представлення простору події кругами Ейлера

Всі елементарні події рівноймовірні.

До проведення експерименту маємо: $P(B) = m/N$, $P(A) = n/N$.

Нехай визначається $P(A)$ і в ході експерименту відбулася подія B . Це приводить до того, що початковий простір Ω трансформувався (звужився) до події B . Отже, новий (апостеріорний) простір подій $\Omega_1 = B \cap \Omega = B$. З цього простору тільки подія AB пов'язана зі здійсненням події A . Отже

$$P(A|B) = \frac{p}{m} = \frac{p}{N} \times \frac{1}{m/N} = \frac{P(AB)}{P(B)},$$

$$P(AB) = P(A|B) \times P(B) = P(B|A) \times P(A). \quad (6.5)$$

Означення. Подія A не залежить від події B , якщо має місце рівняння:

$$P(A|B) = P(A). \quad (6.6)$$

Якщо A не залежить від B , то і B не залежить від A . Тому для незалежних подій:

$$P(AB) = P(A) \cdot P(B). \quad (6.7)$$

Поняття незалежних подій можна розширити на довільну кількість подій. Якщо події B_1, B_2, \dots, B_n незалежні, то $P(B_1 B_2 \dots B_n) = P(B_1) P(B_2) \dots P(B_n)$.

6.4. Теорема множення ймовірностей

Означення. Добутком двох подій A і B називають подію AB , яка полягає у спільній появі цих подій.

Теорема 6.1. *Ймовірність спільної появи двох залежних подій дорівнює добутку ймовірності першої події A на умовну ймовірність події B , визначену за припущення, що подія A відбулася:*

$$P(AB) = P(A) \cdot P(B|A). \quad (6.8)$$

Для незалежних подій A і B маємо: $P(B|A) = P(B)$

В цьому частинному випадку має місце наступна теорема.

Теорема 6.2: *Ймовірність сумісної появи двох незалежних подій дорівнює добутку ймовірностей цих подій.*

$$P(AB) = P(A) \cdot P(B). \quad (6.8)$$

Ці теореми поширюються на довільне число залежних чи незалежних подій.

Теорема 6.3. *Ймовірність появи хоча б однієї з подій A_1, A_2, \dots, A_n незалежних в сукупності, дорівнює різниці між одиницею і добутком ймовірностей протилежних подій A_1, A_2, \dots, A_n .*

$$P(A) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot P(\bar{A}_n). \quad (6.9)$$

Доведення. Позначимо як A подію появи хоча б однієї з подій A_1, A_2, \dots, A_n . Події A і $\bar{A}_1 \cdot \bar{A}_2 \dots \bar{A}_n$ (жодна з подій не наступила) протилежні, отже сума їх ймовірностей дорівнює одиниці:

$$P(A) + P(\bar{A}_1 \cdot \bar{A}_2 \dots \bar{A}_n) = 1. \quad (6.10)$$

Користуючись теоремою множення маємо:

$$P(A) = 1 - P(\bar{A}_1 \cdot \bar{A}_2 \dots \bar{A}_n) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \dots \cdot P(\bar{A}_n). \quad (6.11)$$

6.5. Теорема додавання ймовірностей

Означення. Сумою $A + B$ двох подій A і B називають подію, що полягає у появі A чи B , чи обох цих подій.

Означення. Дві події називаються сумісними, якщо поява однієї з них не виключає появи іншої в тому самому випробуванні.

Теорема 6.4: Якщо події A_1, A_2, \dots, A_r є такими, що кожні дві з них є несумісними, то ймовірність появи хоча б однієї з них дорівнює сумі ймовірностей:

$$P(A_1 + A_2 + \dots + A_r) = P(A_1) + \dots + P(A_r). \quad (6.12)$$

Цей вираз відбиває властивість адитивності ймовірнісної міри.

Довести це можна в такий спосіб. Нехай з n рівноймовірних випадків m_1 випадків сприятливі для настання події A (рис. 6.3). Тоді:

$$P(A) = \frac{m_1}{n}$$

Нехай також m_2 випадків сприятливі для появи події B , тому: $P(B) = \frac{m_2}{n}$.

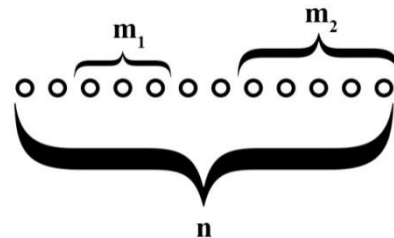


Рис. 6.3. Ілюстрація до теореми 6.4.

Число випадків сприятливих подій або A або B дорівнює $m_1 + m_2$, тому:

$$P(A + B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A) + P(B). \quad (6.13)$$

Цей висновок можна поширити на довільну сукупність попарно несумісних подій.

Приклад 6.1. Стрілок стріляє по мішені, яка розділена на три частини. Ймовірність попадання у першу область дорівнює 0,25, у другу – 0,4. Знайти ймовірність того, що стрілок за одного пострілу поцілить або в першу, або у другу область.

Роз'яання. Позначимо: подія A – «стрілок поцілив у першу область» і подія B – «стрілок поцілив у другу область». Ці події несумісні, бо попадання у одну область виключає попадання в другу, тому застосуємо теорему додавання ймовірностей: $P(A + B) = P(A) + P(B) = 0,25 + 0,4 = 0,65$.

Теорема 6.5. *Ймовірність появи хоча б однієї з двох сумісних подій дорівнює сумі їх ймовірностей без ймовірності спільної появи цих подій.*

$$P(A + B) = P(A) + P(B) - P(A) \cdot P(B|A). \quad (6.14)$$

У випадку незалежних подій:

$$P(A + B) = P(A) + P(B) - P(A) \cdot P(B). \quad (6.15)$$

Приклад 6.2. Ймовірність попадання у ціль під час стрільби першої та другої гармати дорівнює відповідно 0,6 та 0,5. Знайти ймовірність попадання в ціль в одному залпі (з обох гармат) хоча б з одної гармати.

Роз'яання. Позначимо: подія A – «перша гармата поцілила у ціль» і подія B – «друга гармата поцілила у ціль». Ймовірність попадання в ціль кожною з гармат не залежить від результату стрільби другої гармати, тому події A і B є незалежними.

Ймовірність події AB (обидві гармати поцілили у ціль)

$$P(AB) = P(A) \cdot P(B) = 0,6 \cdot 0,5 = 0,3.$$

Тоді шукана ймовірність

$$P(A + B) = P(A) + P(B) - P(A) \cdot P(B) = 0,6 + 0,5 - 0,3 = 0,8.$$

Означення. Повною групою називають сукупність єдино можливих подій випробування.

Приклад 6.3. Стрілець виконує два постріли по мішені. Подія A_1 – одне попадання, подія A_2 – два попадання, і подія A_3 – промах, утворюють повну групу подій.

Теорема 6.6. *Сума ймовірностей подій A_1, A_2, \dots, A_r , які утворюють повну групу, дорівнює одиниці:*

$$P(A_1) + P(A_2) + \dots + P(A_r) = 1. \quad (6.16)$$

Протилежними називаються дві єдино можливі події, що утворюють повну групу.

Приклад 6.4. Попадання в ціль – подія A , промах – \bar{A} .

Теорема 6.7. Сума ймовірностей протилежних подій дорівнює одиниці

$$P(A) + P(\bar{A}) = 1. \quad (6.17)$$

Приклад 6.5. Скористаємось умовою прикладу 2.2. Визначимо для події $A+B$ протилежну подію – жодна з гармат не влучила у ціль. Її ймовірність, враховуючи незалежність подій можна визначити як добуток ймовірностей протилежних подій : $(1 - P(A)) \cdot (1 - P(B)) = 0,4 \cdot 0,5 = 0,2$. Оскільки події A і B незалежні, то $P(A + B) = 1 - P(\bar{A} \cdot \bar{B})$. Тоді шукана ймовірність

$$P(A + B) = 1 - (1 - P(A)) \cdot (1 - P(B)) = 1 - 0,2 = 0,8.$$

6.6. Формула повної ймовірності

Нехай подія A може наступити за умови однієї з несумісних подій B_1, B_2, \dots, B_n , які утворюють повну групу (рис. 6.4). І нехай відомі ймовірності цих подій і умовні ймовірності виду $P(A|B)$. Як знайти ймовірність $P(A)$?

Відповідь на це питання дає наступна теорема.

Теорема 6.8. Ймовірність події A , яка може наступити лише за умови появи події з несумісних подій B_1, B_2, \dots, B_n , які утворюють повну групу, дорівнює сумі добутків ймовірностей кожної з цих подій на відповідну умовну ймовірність події

$$P(A) = P(B_1) \cdot P(A|B_1) + \dots + P(B_n) \cdot P(A|B_n) = \sum_{i=1}^n P(B_i) \cdot P(A|B_i). \quad (6.18)$$

6.7. Поняття гіпотез. Формула Бейеса

Нехай подія A може наступити за умови появи однієї з несумісних подій B_1, B_2, \dots, B_n , що утворюють повну групу. Оскільки наперед невідомо, яка з подій

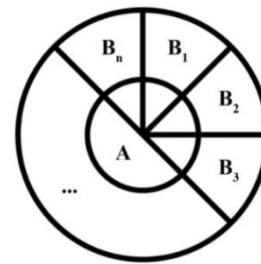


Рис. 6.4. Представлення події A з несумісними подіями B_1, B_2, \dots, B_n ,

наступить, їх називають *гіпотезами* (науковими припущеннями). Ймовірність появи події A визначається формулою повної ймовірності

$$P(A) = P(B_1) \cdot P(A|B_1) + \dots + P(B_n) \cdot P(A|B_n). \quad (6.19)$$

Припустимо, що проведено випробування, за результатами якого мала місце подія A . Як трансформувалась (у зв'язку з тим, що подія A вже настала) ймовірність гіпотез? Простими словами будемо шукати умовні ймовірності $P(B_1|A)$, $P(B_2|A)$, ... $P(B_n|A)$.

З теореми 1 множень ймовірностей маємо:

$$P(AB_i) = P(A) \cdot P(B_i|A); \quad P(AB_i) = P(B_i) \cdot P(A|B_i).$$

З цих рівнянь отримаємо апостеріорну ймовірність i -тої гіпотези :

$$P(B_i|A) = \frac{P(B_i) \cdot P(A|B_i)}{P(A)} = \frac{P(B_i) \cdot P(A|B_i)}{\sum_{i=1}^n P(B_i) \cdot P(A|B_i)}. \quad (6.20)$$

Ця формула називається формулою **Бейеса** (за прізвищем видатного англійського математика, пресвітеріанського священника Т. Bayes (1701-1761) вперше опублікована у 1763 р.). Вона дає змогу переоцінити ймовірність гіпотез після того, як стає відомим результат випробування, в якому з'явилась подія A .

Лекція 7.

ВИПАДКОВА ВЕЛИЧИНА.

ТЕОРЕМИ ПРО ДИСПЕРСІЮ ТА МАТЕМАТИЧНЕ СПОДІВАННЯ ДИСКРЕТНИХ ВИПАДКОВИХ ВЕЛИЧИН

7.1. Визначення випадкової величини

В ході проведених експериментів мають справу з подіями різного фізичного змісту. Аналізувати та досліджувати безпосередньо самі події дуже важко, оскільки вони визначаються не числами, а вербальними (від латинського *verbalis*

– словесний) конструкціями, певними логічними умовами. Тому від випадкових подій переходять до випадкових величин. Це значно спрощує дослідження самих подій і дає змогу математично формалізувати цей процес.

Означення. Дійсною випадковою величиною називають числову функцію $\xi = \xi(\omega)$ з областю визначення $\Omega = \{\omega_i\}$ та областю значень $x \subseteq R$.

Випадкова величина (ВВ) – це числова функція, задана на просторі подій. Таким чином ВВ – це правило (функція), яка ставить у відповідність елементарній події $\omega \in \Omega$ число $x \subseteq R$.

Випадкові величини зазвичай позначають літерами грецького алфавіту: ξ, η, ζ .

Якщо ВВ ξ набуває кінцеве або зліченне число значень, така ВВ називається *дискретною*, у протилежному випадку ВВ називається *неперервною*.

Ілюстрацію задання дискретної випадкової величини показано на рис. 7.1.

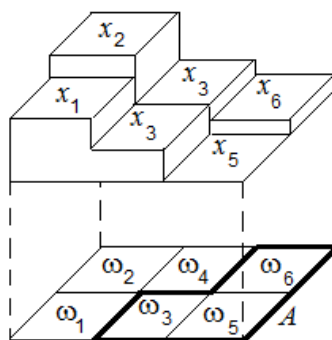


Рис. 7.1. Графічне представлення задання дискретної ВВ

Означення. Законом розподілу дискретної ВВ називається співвідношення, яке пов'язує її можливі значення з відповідними ймовірностями, з якими ці значення приймаються.

Форми представлення закону розподілу: таблична, графічна, аналітична – тобто формулою, яка **відтворює** правило обчислення ймовірностей. Найпростішою формою закону розподілу дискретної ВВ $\xi(\omega)$ є таблиця.

Таблиця 7.1. Таблиця із законом розподілу дискретної ВВ

$\xi(\omega)$	x_1	x_2	x_3	...	x_{n-1}	x_n
---------------	-------	-------	-------	-----	-----------	-------

P	P_1	P_2	P_3	\dots	x_{n-1}	P_n
-----	-------	-------	-------	---------	-----------	-------

Варто пам'ятати, що сума ймовірностей у цьому ряді завжди дорівнює одиниці: $P_1 + P_2 + \dots + P_n = 1$. Таку таблицю або її графічне представлення називають *розподілом ймовірності* ВВ ξ . Крім табличного способу розподіли можуть задаватись графічно чи аналітично, тобто формулою, яка відтворює правило обчислення ймовірностей. Прикладами такого завдання є розподіли Бернуллі та Пуассона, які розглядатимуться в лекції 4.

Та обставина, що для різних ω ВВ ξ може набувати одне й те саме значення є суттєвою: множина $\{x_1 \ x_2 \ \dots x_n \ \dots\}$ можливих значень ВВ може бути значно спрощена відносно множини Ω .

В прикладних задачах теорії ймовірностей частіше розглядають не самі ВВ, а їх розподіли. Це пов'язано з тим, що зазвичай результатами експериментів є не випадкові події ω_i , а ВВ, тобто певні числові значення, які реєструються різними приладами чи іншими технічними засоба.

7.2. Числові характеристики дискретних ВВ

7.2.1. Математичне сподівання дискретної ВВ

Математичне сподівання дискретної ВВ – це сума добутків всіх її можливих значень та їх ймовірностей

$$M\xi = \sum_{i=1}^n x_i P_i . \quad (7.1)$$

Це не випадкова величина – число. Ймовірносний сенс математичного сподівання дискретної ВВ полягає у тому, що воно наближено дорівнює середньому арифметичному спостережуваних значень ВВ. Дійсно, спостережувані значення рівноімовірні, тобто $P_i = 1/n$, то

$$M\xi = \sum_{i=1}^n x_i P_i = \frac{1}{n} \sum_{i=1}^n x_i . \quad (7.2)$$

Математичне сподівання характеризує розміщення розподілу ВВ на осі x , тому його ще називають *центром розподілу*.

Математичне сподівання дискретної ВВ має наступні **властивості**.

1. $MC = C$, тобто математичне сподівання константи $C = const$ дорівнює константі.
2. $M(C \cdot \xi) = C \cdot M\xi$.
3. Математичне сподівання добутку двох незалежних ВВ дорівнює добутку їх математичних сподівань

$$M(\xi \cdot \eta) = M\xi \cdot M\eta. \quad (7.3)$$

4. Математичне сподівання суми двох ВВ (залежних чи незалежних) дорівнює сумі їх математичних сподівань

$$M(\xi + \eta) = M\xi + M\eta. \quad (7.4)$$

5. Математичне сподівання відхилення (тобто різниці) між ВВ та її математичним сподіванням дорівнює нулю: $M(\xi - M\xi) = 0$.

7.2.2. Дисперсія дискретної ВВ

Означення. *Дисперсія* (від латинського *dispersus* – розсіювання) дискретної ВВ – це математичне сподівання квадрату відхилення ВВ від її математичного сподівання

$$D\xi = M(\xi - M\xi)^2 = \sum_{i=1}^n (x_i - M\xi)^2 \cdot P_i . \quad (7.5)$$

Це не випадкова величина. Вона характеризує розсіювану ВВ в околі її математичного сподівання.

Дисперсія дискретної ВВ має наступні **властивості**.

1. $D\xi = M(\xi^2) - (M\xi)^2$.

2. $\mathbf{DC} = 0$.

3. $\mathbf{D}(C\xi) = C^2 \cdot \mathbf{D}\xi$.

4. Дисперсія суми двох незалежних ВВ дорівнює сумі їх дисперсій

$$\mathbf{D}(\xi + \eta) = \mathbf{D}\xi + \mathbf{D}\eta. \quad (7.6)$$

5. Дисперсія різниці двох незалежних ВВ дорівнює сумі їх дисперсій

$$\mathbf{D}(\xi - \eta) = \mathbf{D}\xi + \mathbf{D}\eta. \quad (7.7)$$

6. $\mathbf{D}\xi \geq 0$.

7.2.3. Середнє квадратичне відхилення (СКВ) ВВ

Середнє квадратичне відхилення ВВ – це додатне значення кореня квадратного з дисперсії

$$\sigma_\xi = +\sqrt{\mathbf{D}\xi}. \quad (7.8)$$

Зручність використання цієї характеристики у випадку досліджень фізичних величин особливо відчутна: розмірність σ_ξ співпадає з розмірністю ξ , тобто $\dim(\sigma_\xi) = \dim(\xi)$.

Для середнього квадратичного відхилення (СКВ) ВВ доведена наступна теорема.

Теорема 7.1. СКВ суми скінченного числа взаємно незалежних ВВ дорівнює квадратному кореню з суми квадратів СКВ цих величин

$$\sigma_{(\xi_1 + \xi_2 + \dots + \xi_n)} = \sqrt{(\sigma_{\xi_1})^2 + (\sigma_{\xi_2})^2 + \dots + (\sigma_{\xi_n})^2}. \quad (7.9)$$

Ця теорема широко використовується в теорії похибок для визначення СКВ випадкової складової результату непрямого вимірювання фізичних величин.

7.3. Дисперсія числа появи подій в незалежних випробуваннях

Нехай здійснюється n незалежних випробувань, в кожному з яких ймовірність події A постійна. Як визначити дисперсію числа появи події A ?

Відповідь на це питання дає наступна теорема.

Теорема 7.2. Дисперсія числа появи події A в n незалежних випробуваннях, в кожному з яких ймовірність $P(A)$ появи події A постійна, дорівнює добутку числа випробувань на ймовірність появи і не появи події в одному випробуванні

$$D\xi = npq. \quad (7.10)$$

Доведення: Розглянемо випадкову величину $\mu(\omega)$ – число появи події A в n незалежних випробуваннях. Вочевидь, загальне число появи події в цих випробуваннях дорівнює сумі появ події в окремих випробуваннях

$$\mu(\omega) = \mu_1(\omega) + \mu_2(\omega) + \dots + \mu_n(\omega), \quad (7.11)$$

де $\mu_i(\omega)$ – число настання події в i -тому випробуванні.

ВВ $\mu_i(\omega)$ взаємно незалежні, оскільки наслідок кожного випробування не залежить від наслідків решти випробувань, тому (див. властивості дисперсії)

$$D\mu = D\mu_1 + D\mu_2 + \dots + D\mu_n. \quad (7.12)$$

Обчислимо $D\mu_i$ за формулою

$$D\mu_i = M\mu_i^2 - (M\mu_i)^2. \quad (7.13)$$

За умовою теореми $M\mu_i = p$.

Визначимо математичне сподівання $M\mu_i^2$. Якщо $\mu_i(\omega)$ набуває в кожному окремому випробуванні тільки 2 значення (поява події A і не поява події \bar{A} – протилежні події), а саме «1» з ймовірністю p і «0» з ймовірністю $q=1-p$

$$M\mu_i^2 = 1^2 p + 0^2 q = p. \quad (7.14)$$

Тоді $D\mu_i = p - p^2 = p(1-p) = pq$.

Вочевидь, дисперсія кожної з решти ВВ також дорівнює pq . Тоді $D\mu = npq$.

7.4. Однаково розподілені взаємно незалежні випадкові величини

Така модель сукупності ВВ є базовою для опрацювання результатів багаторазових вимірювань числового значення певної фізичної величини.

Результат кожного окремого вимірювання залежить від багатьох випадкових факторів: варіації напруг джерел живлення, механічних вібрацій, коливань метеопараметрів, наявності зовнішніх електромагнітних завад тощо. Такі фактори не можуть бути заздалегідь враховані. Це дає підстави розглядати можливі результати n окремих вимірювань однієї фізичної величини, тобто сукупність значень x_1, x_2, \dots, x_n , як такі, що з'явилися в результаті випробувань сукупності n взаємно незалежних ВВ $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$. Останні мають однакові розподіли ймовірності (за умови незмінності числового значення фізичної величини, методики та засобу вимірювання тощо), отже і однакові числові характеристики – математичне сподівання та дисперсію (у разі опрацювання результатів фізичних експериментів значення $\mathbf{M}\xi, \mathbf{D}\xi, \sigma_\xi$ – розмірні величини).

Розглянемо n взаємно незалежних ВВ $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$, які мають однакові розподіли, отже і однакові числові характеристики. Така ймовірнісна модель використовується у випадку опрацювання даних повторюваних вимірювань фізичної величини незмінного числового значення. Найбільший інтерес для практичного застосування представляє вивчення середнього арифметичного ВВ

$$\bar{\xi}(\omega) = \frac{\xi_1(\omega) + \xi_2(\omega) + \dots + \xi_n(\omega)}{n}. \quad (7.15)$$

Для $\bar{\xi}(\omega)$ мають місце наступні теореми.

Теорема 7.3: Математичне сподівання середнього арифметичного однаково розподілених взаємно незалежних ВВ дорівнює математичному сподіванню « a » кожної з цих величин

$$\mathbf{M}\bar{\xi}(\omega) = \mathbf{M}\xi_i(\omega) = a. \quad (7.16)$$

Теорема 7.4: Дисперсія середнього арифметичного n однаково розподілених взаємно незалежних ВВ в n разів менша за дисперсію складових

$$\mathbf{D}\bar{\xi}(\omega) = \mathbf{D}\left(\frac{\xi_1(\omega) + \xi_2(\omega) + \dots + \xi_n(\omega)}{n}\right) = \frac{\mathbf{D}\xi}{n}. \quad (7.17)$$

Теорема 7.5. Середнє квадратичне відхилення середнього арифметичного n однаково розподілених взаємно незалежних випадкових величин в \sqrt{n} разів менше середнього квадратичного відхилення σ_ξ кожної з них

$$\sigma_{\bar{\xi}} = \sigma_\xi / \sqrt{n}. \quad (7.18)$$

7.5. Поняття про моменти розподілу випадкової величини вищих порядків

Розглянемо дискретну ВВ $\xi(\omega)$, задану наступним розподілом:

$\xi(\omega)$	1	2	5	100
p	0,6	0,2	0,19	0,01

Її математичне сподівання:

$$\mathbf{M}\xi = 1 \cdot 0,6 + 2 \cdot 0,2 + 5 \cdot 0,19 + 100 \cdot 0,01 = 2,95.$$

Тепер розглядаємо закон розподілу $\xi^2(\omega)$ та визначимо $\mathbf{M}\xi^2(\omega)$

$\xi^2(\omega)$	1	4	25	10000
p	0,6	0,2	0,19	0,01

$$\mathbf{M}\xi^2(\omega) = 1 \cdot 0,6 + 4 \cdot 0,2 + 25 \cdot 0,19 + 10000 \cdot 0,01 = 106,15.$$

З цього прикладу бачимо, що $\mathbf{M}\xi^2(\omega) \gg \mathbf{M}\xi$. Це пояснюється тим, що після зведення в квадрат, внесок великих, але малоїмовірних значень ВВ в математичне сподівання суттєво збільшилось. Ось чому виявляється доцільним розглядати математичне сподівання цілої додатної степені ВВ більшої за одиницю (тобто моменти розподілу ВВ вищих порядків)!

Означення. Початковим моментом порядку k випадкової величини $\xi(\omega)$ називають математичне сподівання величини $\xi^k(\omega)$

$$\nu_k = \mathbf{M}\xi^k(\omega). \quad (7.19)$$

Означення. Центральним моментом порядку k випадкової величини $\xi(\omega)$ називають математичне сподівання величини $(\xi - \mathbf{M}\xi)^k$

$$\mu_k = \mathbf{M}(\xi - \mathbf{M}\xi)^k. \quad (7.20)$$

Використовуючи властивості математичного сподівання можна отримати:

$$\mu_2 = v_2 - v_1^2$$

$$\mu_3 = v_3 - 3v_1 v_2 + 2 v_1^3$$

$$\mu_4 = v_4 - 4 v_3 v_1 + 6 v_2 v_1^2 - 3 v_1^4$$

Моменти більш високих порядків використовують вкрай рідко.

Лекція 8.

ПОВТОРЮВАННЯ ВИПРОБУВАНЬ. ЗАКОНИ ВЕЛИКИХ ЧИСЕЛ

8.1. Схема Бернуллі. Розподіл Бернуллі

Схема Бернуллі – це ймовірнісна модель, що відповідає n незалежним випробуванням з двома наслідками. Така модель виникає, до прикладу, під час контрольних випробувань приладів чи виробів (справний/несправний, норма/брак), або у випадку підкидання монет (герб/решітка).

Два наслідка кожного випробування в схемі Бернуллі позначають символами 1 чи 0, або називають успіхом та неуспіхом, а відповідні їм ймовірності – літерами p і $q=1-p$. Якщо позначити число успіхів (число одиниць) в n випробуваннях як μ_n , то

$$P(\mu_n = m) = C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m}, \quad m = \overline{0, n}, \dots (8.1)$$

тобто це ймовірність складної події типу $\underbrace{AA \dots A}_m \underbrace{\bar{A}\bar{A} \dots \bar{A}}_{n-m}$ – досягнення m успіхів та $(n-m)$ неуспіхів в n випробуваннях.

Виведення наведеної вище формули ґрунтується на таких міркуваннях:

1) неважливо в якому порядку відбуваються успіхи;

- 2) імовірність однієї складної події (m успіхів в n випробуваннях) визначається як $p^m q^{n-m}$;
- 3) загальна кількість складних подій дорівнює числу сполучень з n елементів по m – C_n^m .

Розподіл ВВ μ_n з ймовірностями $P(\mu_n = m)$ називається розподілом *Бернуллі* або *біноміальним*. Назва «біноміальний» пояснюється тим, що $P(\mu_n = m)$ можна розглядати як загальний член розкладу бінома Ньютона

$$(p + q)^n = C_n^n p^n + C_n^{n-1} p^{n-1} q + \dots + C_n^m p^m q^{n-m} \dots + C_n^0 q^n . \quad (8.2)$$

За великих n для обчислення ймовірності того, що число успіхів в n випробуваннях схеми Бернуллі знаходиться в межах між значеннями m_1 та m_2 , тобто $m_1 \leq \mu_n \leq m_2$, використовують *локальну та інтегральну теорему Муавра-Лапласа*. Асимптотична формула обчислення ймовірності такої події для $p=0,5$ отримана видатним англійським математиком А. Муавром у 1730 р., а в 1783 р.; відомий французький математик П. Лаплас узагальнив її для довільного $p \in [0, 1]$.

8.2. Локальна теорема Лапласа

Теорема 8.1. Якщо ймовірність p появи події A в кожному досліді постійна і відмінна від нуля і одиниці, то ймовірність $P_n(m)$ того, що подія A в n дослідях з'явиться m разів наближено дорівнює (чим більше n тим точніше) значенню функції

$$P_n(m) = y(x) = \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{npq}} \cdot \varphi(x), \quad (8.3)$$

$$\text{де } x = \frac{m - np}{\sqrt{npq}} .$$

Для функції $\varphi(x)$ складені таблиці. Ця функція парна: $\varphi(x) = \varphi(-x)$.

Задовільна точність обчислення функції $y(x)$ досягається, якщо $npq \geq 10$.

8.3. Інтегральна теорема Муавра-Лапласа

Теорема 8.2. Якщо ймовірність успіху в кожному випробуванні дорівнює p , $0 < p < 1$, постійна, то за умови $n \rightarrow \infty$ для будь-яких a і b має місце співвідношення

$$P\left\{a \leq \frac{\mu_n - np}{\sqrt{npq}} \leq b\right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx = \Phi(b) - \Phi(a).$$

В останній формулі маємо: np – математичне сподівання ВВ μ_n ($M\mu_n = np$), npq – дисперсія ВВ μ_n ($D\mu_n = npq$), $\Phi(\cdot)$ – функція Лапласа

$$\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_0^a e^{-x^2/2} dx. \quad (8.4)$$

Функція Лапласа непарна: $\Phi(-x) = -\Phi(x)$. Значення цієї функції табульовані. На практиці найчастіше використовуються в розрахунках такі значення: $\Phi(0) = 0$; $\Phi(1) = 0,3413$; $\Phi(2) = 0,4772$; $\Phi(3) = 0,4987$.

8.4. Розподіл Пуассона

Інший приклад аналітичного завдання дискретної ВВ – розподіл Пуассона.

Нехай виконується n незалежних випробувань, в кожному з яких ймовірність настання події A дорівнює $P(A) = p$. Для визначення ймовірності появи k разів події A в таких дослідах використовують формулу Бернуллі. Однак вона не придатна у разі, якщо $p \leq 0,1$, а $n \gg 1$. В таких випадках вдаються до формули Пуассона.

Зробимо важливе припущення: добуток np зберігає постійне значення: $np = \lambda$. Це означає, що середнє число появи подій A в різних серіях випробувань (за різних n) лишається незмінним.

Скористаємось формулою Бернуллі з підстановкою $p = \lambda/n$:

$$P_n(k) = C_n^k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} \cdot \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k}. \quad (8.5)$$

Визначимо границю $\lim_{n \rightarrow \infty} P_n(k)$

$$\lim_{n \rightarrow \infty} P_n(k) = \frac{\lambda^k}{k!} \cdot \lim_{n \rightarrow \infty} \frac{n(n-1) \cdot \dots \cdot (n-(k-1))}{n^k} \cdot \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \cdot \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-k}. \quad (8.6)$$

Перша і остання границі прямують до 1. Другий множник, згідно з виразом для другої чудової границі, дорівнює

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = \lim_{n \rightarrow \infty} \left[\left(1 - \frac{\lambda}{n}\right)^{-n/\lambda} \right]^{-\lambda} = e^{-\lambda}. \quad (8.7)$$

Тому маємо :

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

Ця формула виражає закон *розподілу Пуассона ймовірностей масових* (n велике) *рідких* (p мале) подій.

Закон Пуассона є однопараметровим, тобто залежить від одного параметра λ , зміст якого полягає в наступному: він є одночасно і математичним сподіванням, і дисперсією ВВ з розподілом Пуассона.

8.5. Закони великих чисел

На перший погляд здається, що у випадку, коли про кожну ВВ з їх певної сукупності маються достатньо скромні (неповні) відомості, навряд чи можна встановити якісь закономірності для суми значного числа ВВ. Але це не так! Виявляється, що за певних доволі загальних умов сумарна поведінка достатньо великої кількості ВВ майже втрачає випадковий характер і стає закономірною. Ці умови визначають ряд теорем, які мають загальну назву – закони великих чисел. До них належать, в першу чергу, теореми Чебишова та Бернуллі. Для їх кращого розуміння, спочатку розглянемо нерівність Чебишова.

8.5.1. Нерівність Чебишова (або нерівність Бьєнеме – Чебишова)

Ця нерівність отримана французьким статистиком І.-Ж Бьєнеме у 1853 р, а пізніше – відомним російським математиком П.Л. Чебишовим (1821 –1894).

Лема. Якщо ВВ ξ може набувати лише невід'ємних значень, тобто $\xi \geq 0$, і має скінченне математичне сподівання, то $P(\xi \geq 1) \leq M\xi$.

Доведення. Наведемо доведення для дискретної ВВ $\xi \geq 0$, значення якої мають ймовірності $p_i, i = 1, 2, \dots$. Для всіх значень ВВ $\xi \geq 1$ можна записати

$$P(\xi \geq 1) = \sum_{i: x_i \geq 1} p_i \leq \sum_{i: x_i \geq 1} x_i p_i \leq \sum_i x_i p_i = M\xi, \quad (8.8)$$

що і доводить цю лему. Для випадку неперервної ВВ доведення виконується за такою ж схемою. Нерівність (8.8) називають ще першою формою нерівності Чебишова.

Друга форма нерівності Чебишова дається у такому формулюванні: якщо ВВ ξ має скінченні математичне сподівання $M\xi$ та дисперсію $D\xi$, то для будь-якого числа $\varepsilon > 0$ виконуються нерівності

$$P(|\xi - M\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2} \quad (8.9)$$

$$P(|\xi - M\xi| < \varepsilon) \geq 1 - \frac{D\xi}{\varepsilon^2}. \quad (8.10)$$

Доведення. Для будь-якого числа $\varepsilon > 0$ розглянемо випадкову величину $\eta = \frac{(\xi - M\xi)^2}{\varepsilon^2}$, яка може набувати лише невід'ємних значень. З урахуванням леми

можна скласти нерівність $P(\eta \geq 1) \leq M\eta$, або

$$P\left(\eta = \frac{(\xi - M\xi)^2}{\varepsilon^2} \geq 1\right) \leq M\left(\eta = \frac{(\xi - M\xi)^2}{\varepsilon^2}\right) = \frac{1}{\varepsilon^2} M(\xi - M\xi)^2 = \frac{D\xi}{\varepsilon^2}. \quad (8.11)$$

Враховуючи, що нерівності $\frac{(\xi - M\xi)^2}{\varepsilon^2} \geq 1$ і $|\xi - M\xi| \geq \varepsilon$ еквівалентні, нерівність (8.11) доводить нерівність (8.9).

Оскільки події $|\xi - \mathbf{M}\xi| < \varepsilon$ та $|\xi - \mathbf{M}\xi| \geq \varepsilon$ протилежні, можна записати

$$P(|\xi - \mathbf{M}\xi| < \varepsilon) + P(|\xi - \mathbf{M}\xi| \geq \varepsilon) = 1. \quad (8.12).$$

Використовуючи (8.92) та (8.8) можна отримати іншу форму представлення нерівності Чебишова

$$P(|\xi - \mathbf{M}\xi| < \varepsilon) = 1 - P(|\xi - \mathbf{M}\xi| \geq \varepsilon) \geq 1 - \frac{\mathbf{D}\xi}{\varepsilon^2}, \quad (8.13)$$

що і доводить нерівність (8.10).

8.5.2. Теорема Чебишова

Теорема 8.3. Якщо $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$ попарно незалежні ВВ, а їх дисперсії рівномірно обмежені (не перевищують постійне число C), то яким би малим не було додатне число ε , ймовірність нерівності

$$\left| \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n} - \frac{\mathbf{M}\xi_1(\omega) + \mathbf{M}\xi_2(\omega) + \dots + \mathbf{M}\xi_n(\omega)}{n} \right| < \varepsilon \quad (8.14)$$

буде як завгодно близькою до одиниці, якщо число ВВ ε достатньо великим.

Іншими словами

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n} - \frac{\mathbf{M}\xi_1(\omega) + \mathbf{M}\xi_2(\omega) + \dots + \mathbf{M}\xi_n(\omega)}{n} \right| < \varepsilon \right) = 1. \quad (8.15)$$

Отже теорема Чебишова стверджує: якщо розглядається достатньо велике число незалежних ВВ, які мають обмежені дисперсії, то майже вірогідною можна вважати подію, яка полягає в тому, що відхилення середнього арифметичного ВВ від середнього арифметичного їх математичних сподівань буде за абсолютною величиною скільки завгодно малим.

В цій теоремі не робилось ніяких припущень щодо рівності математичних сподівань випадкових величин.

Сутність теореми Чебишева полягає у наступному. Хоча окремі незалежні ВВ можуть набувати значення суттєво віддалені від своїх математичних сподівань, середнє арифметичне достатньо великої кількості ВВ з великою

ймовірністю набуває значень близьких до певного постійного числа, а саме до числа $\frac{M\xi_1(\omega) + M\xi_2(\omega) + \dots + M\xi_n(\omega)}{n}$. Іншими словами окремі ВВ можуть мати значний розкид, а їх середнє арифметичне розсіяно мало. Отже, не можна впевнено передбачити яке з можливих значень прийме кожна з ВВ, але можна передбачити, яке значення матиме їх середнє арифметичне.

Висновок: середнє арифметичне достатньо великої кількості незалежних ВВ (дисперсії яких рівномірно обмежені) втрачає характер випадкової величини. Це пояснюється тим, що відхилення $\xi_i - M\xi_i$ мають знаковмінний характер, а в середньому арифметичному вони компенсуються.

8.5.3. Практичне значення теореми Чебишова

Для зменшення випадкової складової похибки вимірювання значення фізичної величини виконують декілька її вимірювань, а в якості її числового значення приймають арифметичне середнє (туту і далі, згідно «Міжнародний словника метрології: основні та загальні поняття і відповідні терміни» використовуємо такі терміни, до прикладу, фізична величина – довжина, значення фізичної величини – 4,5 м, числове значення фізичної величини – 4,5). За яких умов такий спосіб визначення числового значення фізичної величини можна вважати правильним? Відповідь на це питання дає теорема Чебишова.

Розглянемо числові значення кожного вимірювання як набір невинпадкових чисел, які виникають в експерименті з рядом ВВ $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$. До цих ВВ можна застосувати теорему Чебишова, якщо вони попарно незалежні, мають однакове математичне сподівання і їх дисперсії рівномірно обмежені.

Перша умова виконується, якщо результати кожного вимірювання не залежать від решти результатів.

Друга вимога виконується, якщо вимірювання виконані без систематичних похибок, а вимірювана величина є незмінною впродовж процесу вимірювання.

Третя вимога виконується, якщо прилад забезпечує певну точність

вимірювань, а в ряду спостережень відсутні значення, що містять грубу похибку вимірювання (промах), тобто мають обмежену дисперсію.

За виконання цих умов і достатньо великому n ймовірність виконання нерівності $|\bar{\xi} - a| < \varepsilon$ як завгодно близька до одиниці. Тобто достовірним є висновок про те, що середнє арифметичне як завгодно мало відрізняється від істинного (але не відомого) значення вимірюваної величини. Однак внаслідок наближеного виконання цих умов годі сподіватись, що за рахунок збільшення $n \rightarrow \infty$ можна як завгодно зменшувати похибку визначення середнього.

На теоремі Чебишова ґрунтується широко вживаний в статистиці **вибірковий метод**, сутність якого полягає у тому, що за порівняно невеликою випадковою вибіркою роблять висновок про характеристики всієї генеральної сукупності.

8.5.4. Теорема Бернуллі

Ця теорема є найпростішим видом закону великих чисел. Нехай виконується n незалежних випробувань, в кожному з яких ймовірність появи події A дорівнює p . Чи можна передбачити якою буде відносна частота появи події A ?

Відповідь на це запитання дає доведена швейцарським математиком Яковом Бернуллі (1713 р) теорема, яка поклала початок теорії ймовірностей як науки.

Теорема 8.4 (Бернуллі). Якщо в кожному з n незалежних випробувань ймовірність p появи події A постійна, то як завгодно близькою до 1 є ймовірність того, що відхилення її відносної частоти від ймовірності p за абсолютною величиною, буде скільки завгодно малим, якщо число випробувань є достатньо великим:

$$\lim_{n \rightarrow \infty} P(|m/n - p| < \varepsilon) = 1. \quad (8.16)$$

Але збіжність m/n до p не означає, що $\lim_{n \rightarrow \infty} m/n = p$.

Відносна частота за достатньо великих значень n має властивість стійкості і обґрунтовує дане раніше статистичне визначення ймовірності.

Лекція 9.

НЕПЕРЕРВНІ ВИПАДКОВІ ВЕЛИЧИНИ

Означення. *Неперервною* називають ВВ, яка може набувати всі значення з певного скінченного чи нескінченного проміжку. Число можливих значень неперервної ВВ безмежне.

9.1. Функція розподілу ймовірностей випадкової величини

Загальною формою закону розподілу ВВ ξ (дискретної чи неперервної) є *інтегральна функція розподілу*, або просто *функція розподілу*.

Функція розподілу $F(x)$, $x \in R$ ВВ ξ визначається виразом:

$$F(x) = P(\xi < x), \quad -\infty < x < \infty. \quad (9.1)$$

Формула (9.1) тлумачиться наступним чином: значення функції розподілу за будь-яких x дорівнює ймовірності того, що ВВ набуває значень менших за $x \in (-\infty; \infty)$. Функція розподілу має такі властивості:

- 1) вона монотонна (якщо $x_1 < x_2$, то $F(x_1) \leq F(x_2)$);
- 2) вона неперервна зліва ($\lim_{x \rightarrow x_0} F(x) = F(x_0)$);
- 3) $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$, $F(\infty) = \lim_{x \uparrow \infty} F(x) = 1$.

Для $a < b$ $a \leq \xi < b$ маємо

$$\Delta F(x) = F(b) - F(a) = P(\xi < b) - P(\xi \leq a) = P(a \leq \xi < b). \quad (9.2)$$

Для довільно фіксованого числа C маємо:

$$P(\xi = C) = F(C + 0) - F(C). \quad (9.3)$$

Отже, якщо функція $F(x)$ в точці "C" безперервна, то ймовірність того, що $\xi = C$ дорівнює 0, а якщо в т. C є стрибок (розрив), то $F(C + 0) - F(C) = \Delta F(C) > 0$, тобто величина стрибка $\Delta F(C)$ дорівнює ймовірності того, що ВВ ξ набуває значення C.

9.2. Щільність розподілу ймовірності

Означення. Якщо існує така функція $p(x)$, що для всіх x має місце співвідношення

$$F(x) = \int_{-\infty}^x p(y)dy, \quad (9.4)$$

то $p(x)$ називається *щільністю розподілу ймовірності* випадкової величини, а ВВ, які мають цю властивість, називаються *неперервними*. Отже, якщо $F(x)$ неперервна, то

$$p(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x \leq \xi < x + \Delta x)}{\Delta x} = \frac{dF(x)}{dx}, \quad x \in (-\infty, \infty) \quad (9.5)$$

9.2.1. Ймовірнісний сенс щільності розподілу

За визначенням щільності розподілу маємо

$$p(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}. \quad (9.6)$$

З диференціального числення відомо, що приріст функції наближено дорівнює диференціалу функції: $F(x + \Delta x) - F(x) \approx dF(x) \approx F'(x)dx = p(x)dx$. Цей вираз має наступний ймовірнісний сенс: ймовірність того, що ВВ набуде значень з інтервалу $(x, x + \Delta x)$ тобто $\xi(\omega) \in (x, x + \Delta x)$, наближено дорівнює (з точністю до нескінченно малих вищого порядку відносно Δx) добутку щільності ймовірності в точці x на довжину інтервалу Δx .

Цей результат трактується наступним чином: ймовірність того, що ВВ ξ набуде значень з інтервалу $(x, x + \Delta x)$, наближено дорівнює площі прямокутника з основою Δx і висотою $p(x)$. Геометрично це положення пояснює рис. 9.1, на якому зображено: на епюрі (а) – функції розподілу ймовірностей неперервної ВВ, на епюрі (б) – щільність розподілу ймовірності.

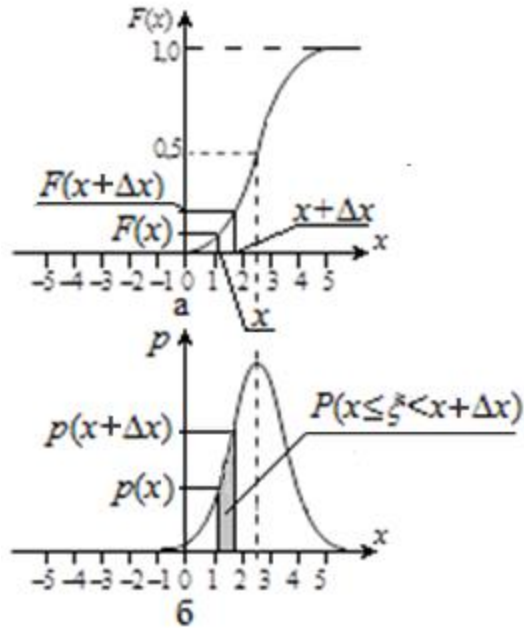


Рис. 9.1. Геометрична інтерпретація ймовірнісного сенсу щільності розподілу

9.2.2. Властивості щільності розподілу

1) $p(x) \geq 0$;

2) Умова нормування ймовірнісної міри: $\int_{-\infty}^{\infty} p(x)dx = 1$;

3) для довільних a, b має місце співвідношення

$$P(a \leq \xi < b) = F(b) - F(a) = \int_a^b p(x)dx . \quad (9.7)$$

Функція $p(x)$ може мати розриви першого роду. Функцію $F(x)$ ще називають *інтегральною*, а $p(x)$ – *диференціальною* функціями розподілу.

Значення $p(x)$ можуть бути більшими за 1. Крім того, якщо ВВ має розмірність, то функція $p(x)$ має розмірність зворотну до розмірності ВВ.

9.3. Приклади розподілів неперервних ВВ

9.3.1. Рівномірний розподіл імовірностей:

Щільність рівномірного розподілу ймовірностей (рис. 9.2,а) визначається як

$$p(x) = \begin{cases} 0, & x \notin [a, b], \\ 1/(b-a), & x \in [a, b] \end{cases} \quad -\infty < a < b < \infty, \quad (9.8)$$

де a – ліва межа, b – права межа розподілу.

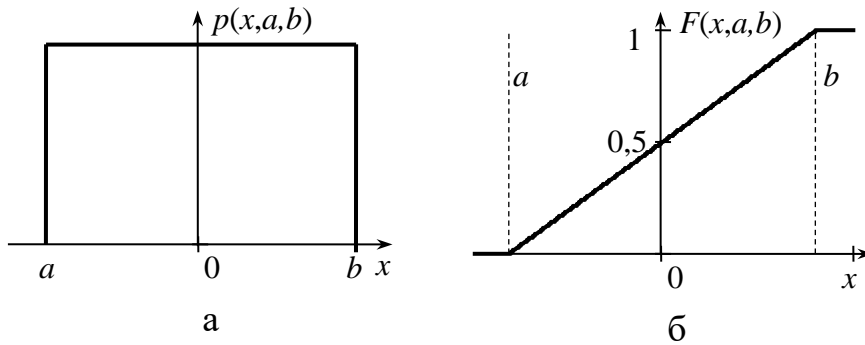


Рис. 9.2. Диференціальна (а) та інтегральна (б) функції рівномірного розподілу

Функція розподілу (інтегральна, рис. 9.2,б) визначається як

$$F(x; a; b) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b. \\ 1, & x > b \end{cases} \quad (9.9)$$

Математичне сподівання ВВ ξ з рівномірним розподілом дорівнює $\mathbf{M}\xi = \frac{a+b}{2}$, а дисперсія – $\mathbf{D}\xi = \frac{(b-a)^2}{12}$. Такий розподіл має, до прикладу, похибка квантування аналого-цифрового перетворювача.

9.3.2. Трикутний розподіл імовірностей (розподіл Сімпсона)

Щільність трикутного розподілу ймовірностей (рис. 9.3,а) визначається як

$$p(x, a, b) = \begin{cases} 0, & a > x > b, \\ \frac{4(x-a)}{(b-a)^2}, & a \leq x \leq \frac{b+a}{2}, \\ \frac{4(b-x)}{(b-a)^2}, & \frac{b+a}{2} \leq x \leq b, \end{cases} \quad (9.10)$$

де a – ліва межа, b – права межа розподілу.

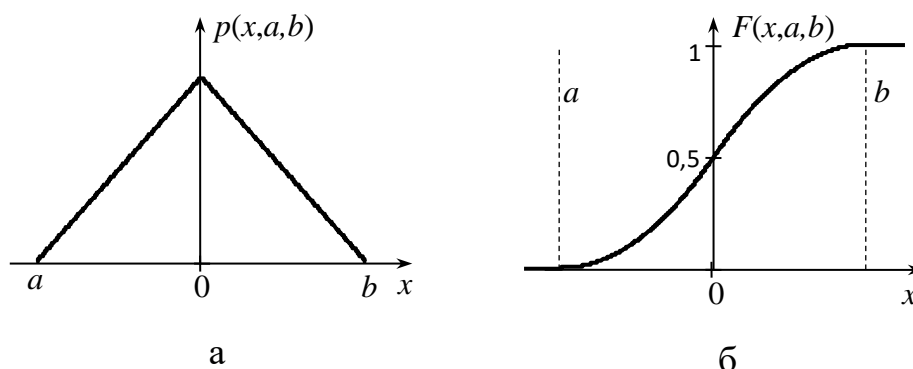


Рис. 9.3. Диференціальна (а) та інтегральна (б) функції розподілу Сімпсона

Функція розподілу (інтегральна, рис. 9.3,б) визначається виразом

$$F(x;a;b) = \begin{cases} 0, & x < a, \\ \frac{2(x-a)^2}{(b-a)^2}, & a \leq x \leq \frac{b+a}{2}, \\ 1 - \frac{2(b-x)^2}{(b-a)^2}, & \frac{b+a}{2} \leq x \leq b, \\ 1, & x > b. \end{cases} \quad (9.11)$$

Математичне сподівання ВВ ξ з розподілом Сімпсона дорівнює $\mathbf{M}\xi = (a+b)/2$, а дисперсія – $\mathbf{D}\xi = (b-a)^2/6$. Такий розподіл має, до прикладу, похибка квантування часового інтервалу, яка виникає двічі – на початку і в кінці часового інтервалу.

9.3.3. Гаусовий розподіл імовірностей

Розподіл гаусової ВВ відіграє важливу роль у теорії похибок. Цей закон вперше отримав А. Муар в 1738 р., а К. Гаусс в 1809 р. використав його в теорії похибок. Щільність гауссового розподілу (рис. 9.4,а) дається формулою

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty, \quad \sigma > 0, \quad (9.12)$$

де μ – параметр розташування; σ – параметр масштабу. Математичне сподівання

гауссової ВВ ξ дорівнює $\mathbf{M}\xi = \mu$, а дисперсія – $\mathbf{D}\xi = \sigma^2$.

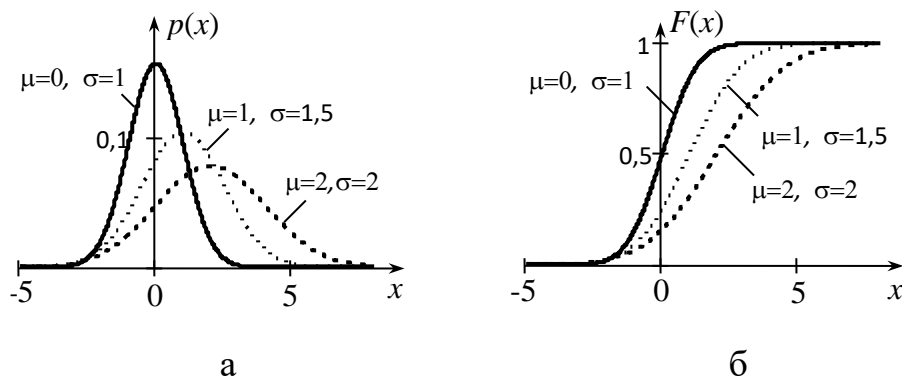


Рис. 9.4. Диференціальна (а) та інтегральна (б) функції розподілу ймовірності Гаусса

Функція $p(x)$ симетрична відносно μ , парна.

Стандартний вид цього розподілу має місце для параметрів $\sigma = 1, \mu = 0$

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} . \quad (9.13)$$

Графік щільності ймовірності стандартного розподілу Гаусса з позначенням ймовірності попадання значень ВВ в різні інтервали наведено на рис. 9.5.

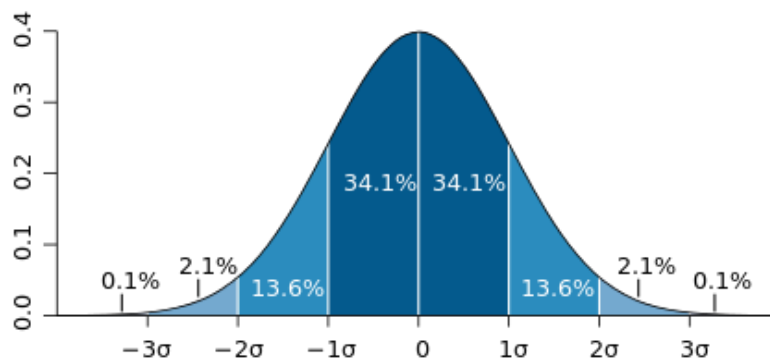


Рис. 9.5. Графік щільності ймовірності стандартного розподілу Гаусса

Функція розподілу (інтегральна, рис. 9.4,б) не виражається через елементарні функції, натомість обчислюється через *інтеграл ймовірності*, або *інтеграл помилок (функцію Лапласа)*, значення якого табульовані

$$erf(x) = 2 \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du . \quad (9.14)$$

Отже, $erf(x)$ – це вірогідність того, що ВВ зі стандартним Гауссовим (нормальним) розподілом належить інтервалу $\pm x$, тобто $\xi \in [-x, x]$.

В теорії імовірностей частіше використовують інтеграл імовірності Гаусса (функцію стандартного розподілу нормованих ухилень ВВ $\frac{\xi - a}{\sigma}$)

$$\Phi(x) = \frac{1}{2\pi} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = \frac{1}{2} [1 + erf(x)], \quad x \in (-\infty, \infty). \quad (9.15)$$

Застосовуючи цю функцію можна обчислювати ймовірності подій з гауссовим розподілом за довільних параметрів розподілу

$$\begin{aligned} P(x_1 < \xi < x_2) &= P(x_1 - \mu < \xi - \mu < x_2 - \mu) = P\left(\frac{x_1 - \mu}{\sigma} < \frac{\xi - \mu}{\sigma} < \frac{x_2 - \mu}{\sigma}\right) = \\ &= \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right) = \frac{1}{2} erf\left(\frac{x_2 - \mu}{\sigma}\right) - \frac{1}{2} erf\left(\frac{x_1 - \mu}{\sigma}\right). \end{aligned} \quad (9.16)$$

В таблицях приведені значення функції для додатних значень аргументів $u = \frac{x - \mu}{\sigma} > 0$. Значення функції $\Phi(u)$ для негативних значень аргументу розраховується як

$$\Phi(-u) = 1 - \Phi(u). \quad (9.17)$$

9.3.4. Розподіл χ^2 (розподіл Пірсона)

Цей розподіл пов'язаний з ім'ям видатного англійського математика, статистика, біолога та філософа К. Пірсона (1857 – 1936).

Нехай $\xi_i (i = 1, 2, 3 \dots n)$ – ВВ з розподілом Гаусса, і нехай $\mathbf{M}\xi_i = 0$, а $\mathbf{D}\xi_i = 1$.

Тоді сума квадратів цих величин $\chi_n^2 = \sum_{i=1}^n \xi_i^2$ розподілена за законом χ^2 («хи квадрат») з $k = n$ степенями свободи. Якщо ці величини зв'язані одним лінійним співвідношенням, то число степенів свободи зменшується на одиницю – $k = n - 1$.

Диференціальна функція цього розподілу дається формулою

$$p(x) = \begin{cases} 0, & x \leq 0, \\ \frac{1}{2^{0,5k} \Gamma(0,5k)} \cdot e^{-x/2} \cdot x^{(k/2-1)}, & x > 0, \end{cases} \quad (9.18)$$

де $\Gamma(0,5k)$ – гамма- функція від параметру $0,5k$, $\Gamma(z) = \int_0^{\infty} t^{z-1} \cdot e^{-t} dt$.

Функція $\Gamma(z)$ (інтеграл Ейлера) поширює значення факторіала $z!$ на випадок довільного комплексного числа, зокрема якщо z – ціле і $z > 0$, то $\Gamma(z) = z!$.

Отже розподіл χ^2 визначається одним параметром – числом степенів свободи k . Для ВВ з розподілом χ^2 маємо: $\mathbf{M}\chi^2 = k$, $\mathbf{D}\chi^2 = 2k$.

Багато розподілів визначаються через розподіл χ^2 , до прикладу, розподіл ВВ $\sqrt{\chi^2}$ – довжини випадкового вектора (ξ_1, \dots, ξ_n) з незалежними нормально розподільними компонентами.

Графіки щільності ймовірності та функції розподілу χ^2 (Пірсона) наведено відповідно на рис. 9.6а,б.

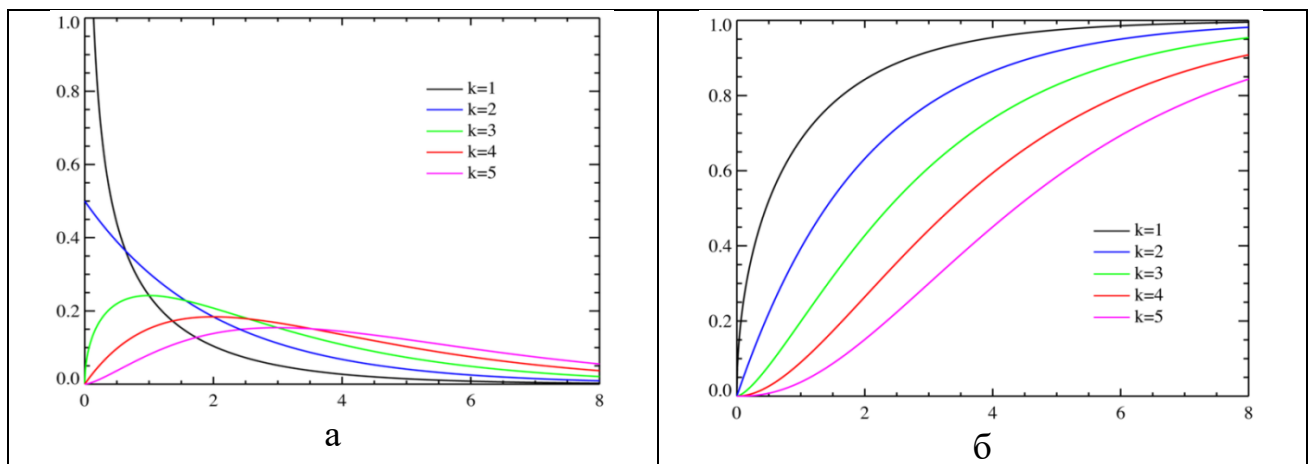


Рис. 9.6. Графіки щільності ймовірності (а) та функції розподілу (б) χ^2

За великих n використовують апроксимацію цього розподілу нормальним: розподіл нормованої величини $\frac{\chi^2 - k}{\sqrt{2k}}$ збігається зі стандартним гауссовим розподілом.

9.3.5. Розподіл Стьюдента

Нехай ξ – ВВ з гаусовим розподілом, $\mathbf{M}\xi = 0$, $\mathbf{D}\xi = 1$, а η – незалежна від ξ ВВ з розподілом χ^2 і k степенями свободи. Тоді нова ВВ

$$T = \frac{\xi}{\sqrt{\eta/k}} \quad (9.19)$$

має розподіл, який називається t -розподілом, або розподілом *Стьюдента* (за псевдонімом англійського статистика В. Госсета) з k степенями свободи.

Зі збільшенням k цей розподіл швидко наближається до гауссового.

9.3.6. Показниковий розподіл.

ВВ ξ розподілена за показниковим (експоненційним) законом з параметром $\lambda > 0$, якщо її щільність розподілу має такий аналітичний вираз

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (9.20)$$

Цей розподіл однопараметровий. Відповідна функція розподілу (інтегральна) має вид:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0. \end{cases} \quad (9.21)$$

Графіки щільності і функції показникового розподілу для різних значень параметру λ наведено на рис. 9.7а,б.

Математичне сподівання і дисперсія дорівнюють відповідно

$$\mathbf{M}\xi = \int_0^{\infty} x \cdot \lambda \cdot e^{-\lambda x} dx = \frac{1}{\lambda}, \quad \mathbf{D}\xi = \frac{1}{\lambda^2}. \quad (9.22)$$

Приклад використання цього розподілу: у найпростішому потоці подій ВВ τ – інтервал часу між двома послідовними подіями – розподілена за показниковим законом. В теорії надійності цей розподіл використовують для оцінювання тривалості безвідмовної роботи технічних пристроїв. В цих випадках змінна x в

(9.20), (9.21) набуває значення часу.

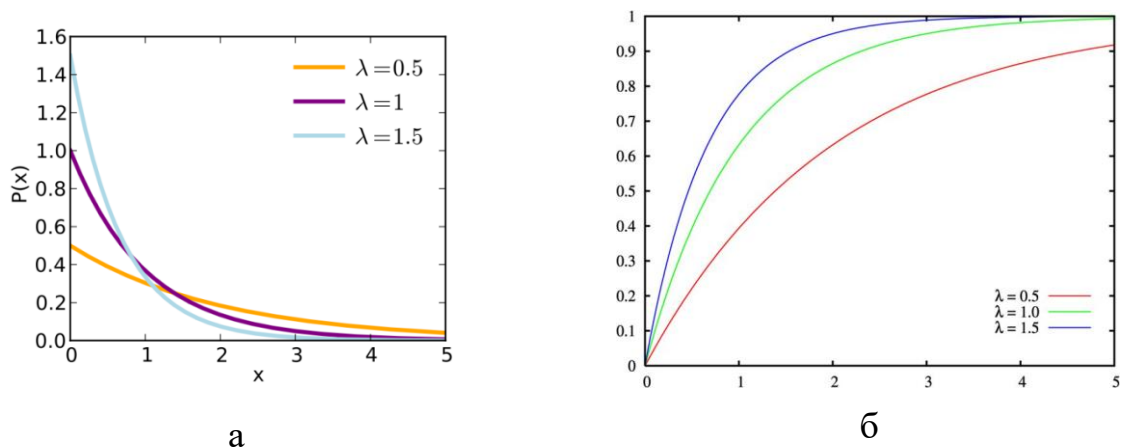


Рис. 9.7. Графіки щільності (а) та функції (б) показникового розподілу

9.4. Характеристики неперервних ВВ

9.4.1. Математичне сподівання

Вперше це поняття застосував Христіан Гюйгенс в 1657 році в роботі «Про розрахунки в азартній грі».

Означення. Математичним сподіванням (або першим початковим моментом) неперервної ВВ ξ з функцією розподілу $F(x)$, яка не має стрибків, називається число, яке визначається інтегралом

$$\mathbf{M}\xi = \int_{-\infty}^{\infty} xp(x)dx . \quad (9.23)$$

В загальному випадку, якщо функція $F(x)$ має стрибки, математичне сподівання дається виразом

$$\mathbf{M}\xi = \int_{-\infty}^{\infty} xdF(x) = \sum_j \int_{x_j}^{x_{j+1}} xp(x)dx + \sum_j p_j x_j , \quad (9.24)$$

де $x_j, j = 0, \pm 1, \pm 2, \dots$ – точки розриву функції $F(x)$, $p(x) = F'(x)$, p_j – стрибки $F(x)$ в цих точках. Для фізичних величин розмірності ВВ та їх математичного сподівання співпадають.

Властивості математичного сподівання:

1. $\mathbf{M}\left(\sum_{i=1}^n \xi_i\right) = \sum_{i=1}^n \mathbf{M}\xi_i$.
2. $\mathbf{M}c = c$, де c – постійна величина, або не випадкова числова функція.
3. $\mathbf{M}(c\xi) = c\mathbf{M}\xi$.
4. Математичне сподівання функції $\varphi(\xi)$ від ξ , яка має розподіл $F(x)$, визначається виразом

$$\mathbf{M}[\varphi(\xi)] = \int_{-\infty}^{\infty} \varphi(x) dF(x) = \int_{-\infty}^{\infty} \varphi(x) p(x) dx . \quad (9.25)$$
5. Якщо ξ_1 та ξ_2 незалежні, то $\mathbf{M}(\xi_1 \cdot \xi_2) = \mathbf{M}\xi_1 \cdot \mathbf{M}\xi_2$.

9.4.2. Дисперсія

Означення. Дисперсією ВВ ξ називається математичне сподівання квадрата відхилення значень ξ від її математичного сподівання

$$\mathbf{D}\xi = \mathbf{M}(\xi - \mathbf{M}\xi)^2 = \int_{-\infty}^{\infty} (x - \mathbf{M}\xi)^2 dF(x) . \quad (9.26)$$

Ця числова характеристика визначає розкид значень ВВ в околі її математичного сподівання.

Дисперсія безперервної ВВ визначається формулою

$$\mathbf{D}\xi = \int_{-\infty}^{\infty} (x - \mathbf{M}\xi)^2 p(x) dx . \quad (9.27)$$

Якщо значення ВВ є відліками сигналу (до прикладу, струму чи напруги), то дисперсія має сенс середньої потужності сигналу. Дисперсія має розмірність квадрата розмірності ξ . Арифметичне значення корня квадратного з дисперсії – це середнє квадратичне відхилення випадкової величини, $\sigma_{\xi} = +\sqrt{\mathbf{D}\xi}$.

Означення. ВВ $(\xi - \mathbf{M}\xi)$ називається *середнім ухиленням*, а ВВ $(\xi - \mathbf{M}\xi) / \sqrt{\mathbf{D}\xi}$ ($\mathbf{D}\xi \neq 0$) називається *нормованим ухиленням*.

Дисперсія неперервної ВВ має такі самі властивості як і дисперсія дискретної ВВ. Нагадаємо їх.

1. Дисперсія постійної c чи числової функції $c(t)$ дорівнює нулю:

$$Dc = Dc(t) = 0.$$

2. $D(c \cdot \xi) = c^2 D\xi$, $D(c(t) \cdot \xi) = c^2(t) \cdot D\xi$.

3. Для суми двох незалежних ВВ ξ_1 та ξ_2 маємо

$$D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2. \quad (9.28)$$

За аналогією до випадкових подій дві ВВ називаються *незалежними*, якщо значення однієї з них не впливає на розподіл значень іншої.

Якщо ВВ ξ_1 та ξ_2 залежні, то:

$$D(\xi_1 + \xi_2) = M(\xi_1 + \xi_2)^2 - (M\xi_1 + M\xi_2)^2 = [M\xi_1^2 - M\xi_1] + [M\xi_2^2 - M\xi_2] + 2[M(\xi_1 \cdot \xi_2) - M\xi_1 \cdot M\xi_2] = D\xi_1 + D\xi_2 + 2\text{cov}(\xi_1, \xi_2), \quad (9.29)$$

де $\text{cov}(\xi_1, \xi_2)$ – коваріація випадкових величин ξ_1 та ξ_2 .

9.4.3. Моменти вищих порядків

Початковий момент k -того порядку в загальному випадку визначається як

$$M\xi^k = \int_{-\infty}^{\infty} x^k dF(x), \quad k = 1, 2, 3, \dots \quad (9.30)$$

Центральний момент k -того порядку в загальному випадку визначається як

$$M(x - M\xi)^k = \int_{-\infty}^{\infty} (x - M\xi)^k dF(x). \quad (9.31)$$

В теорії похибок велике значення відіграє *змішаний центральний момент другого порядку*, який називається *коваріацією* (або *кореляційним моментом*).

Для двох ВВ ξ та η він в загальному виді визначається як

$$K_{\xi\eta} = M\{(x - M\xi) \cdot (y - M\eta)\} \quad (9.32)$$

Кореляційний момент, нормований за допомогою середніх квадратичних відхилень, називається *коефіцієнтом кореляції* ВВ ξ та η

$$\rho_{\xi\eta} = \frac{K_{\xi\eta}}{\sqrt{D\xi \cdot D\eta}} = \frac{K_{\xi\eta}}{\sigma_{\xi} \cdot \sigma_{\eta}} . \quad (9.33)$$

Коефіцієнт $\rho_{\xi\eta}$ задовольняє нерівність $-1 \leq \rho_{\xi\eta} \leq 1$ і характеризує ступінь лінійного зв'язку між ξ та η . Якщо $\rho_{\xi\eta} = 0$, то ВВ називають *некорельованими*.

Лекція 10.

ФУНКЦІОНАЛЬНІ ПЕРЕТВОРЕННЯ ВИПАДКОВИХ ВЕЛИЧИН

10.1. Функція одного випадкового аргументу та її розподіл

Якщо кожному можливному значенню ВВ $\xi(\omega)$ поставлено у відповідність одне можливе значення ВВ $\eta(\omega)$, то $\eta(\omega)$ називається *функцією випадкового аргументу $\xi(\omega)$*

$$\eta(\omega) = \varphi(\xi(\omega)) . \quad (10.1)$$

Розглянемо як можна знайти розподіл функції за відомим розподілом дискретного та неперервного аргументів (далі аргумент ω для спрощення аналітичних виразів опускатимемо).

Нехай ξ – дискретна ВВ. Можливі два випадки.

1. Якщо різним можливим значенням аргументу ξ відповідають різні можливі значення функції η , то ймовірності відповідних значень ξ та η є рівними між собою. До прикладу, якщо дискретна ВВ ξ задана розподілом

ξ	2	5
P	0,7	0,3

то розподіл функції $\eta = \xi^2$ визначається таблицею

η	4	25
P	0,7	0,3

2. Якщо різним можливим значенням аргументу ξ відповідають значення функції η , серед яких є рівні між собою, то слід додавати ймовірності повторюваних значень η . До прикладу, якщо дискретна ВВ ξ задана розподілом

ξ	-2	2	3
P	0,4	0,3	0,3

і треба знайти розподіл функції $\eta = \xi^2$ можна побачити, що значення -2 і 2 дають однакові значення ВВ η , тому їх імовірності в розподілі η повинні додаватись. З урахуванням цього розподіл ВВ η набуває такого вигляду

η	4	9
P	0,7	0,3

Тепер розглянемо випадок, коли ξ – неперервна ВВ і відома її щільність розподілу $p(x)$. Як визначити розподіл функції $\eta = \varphi(\xi)$?

Доведене наступне твердження: якщо функція $y = \varphi(x)$ диференційована, строго зростаюча або строго спадна і має обернену функцію $x = \psi(y)$, то щільність розподілу $g(y)$ для ВВ η знаходять за виразом

$$g(y) = p(\psi(y)) \cdot |\psi'(y)|. \quad (10.2)$$

Приклад 10.1. Неперервна ВВ ξ має гауссовий розподіл, а $\mathbf{M}\xi = 0$. Необхідно знайти щільність розподілу ВВ $\eta = \xi^3$.

Розв'язання. Оскільки функція $y = x^3$ є диференційованою та строго зростаючою, то можна застосувати (10.2). Спочатку знайдемо обернену функцію $x = \psi(y) = y^{1/3}$

За умовою задачі $p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}$, тому маємо $p(\psi(y)) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y^{2/3}}{2\sigma^2}}$.

Визначимо похідну від оберненої функції по y : $\psi'(y) = \frac{d\psi(y)}{dy} = \frac{1}{3} y^{-2/3}$.

Тоді шукана щільність розподілу визначається як

$$g(y) = \frac{1}{3y^{\frac{2}{3}} \sqrt{2\pi}\sigma} e^{-\frac{y^{2/3}}{2\sigma^2}}.$$

Користуючись формулою (10.2) можна довести, що лінійна функція виду $\eta = A\xi + B$ аргументу ξ з гауссовим розподілом і параметрами $\mathbf{M}\xi = a$, $\sigma = \sqrt{\mathbf{D}\xi}$ також має гауссовий розподіл з параметрами

$$\mathbf{M}\eta = Aa + B, \quad \sigma_\eta = |A| \cdot \sigma. \quad (10.3)$$

Приклад 10.2. Знайти щільність розподілу лінійної функції $\eta = 3\xi + 1$, якщо її аргумент має гауссовий розподіл з параметрами $\mathbf{M}\xi = 2$, $\sigma = 0,5$.

Роз'язання. Знайдемо математичне сподівання та середнє квадратичне відхилення ВВ η

$$\mathbf{M}\eta = 3 \cdot 2 + 1 = 7, \quad \sigma_\eta = 3 \cdot 0,5 = 1,5.$$

Після цього можна записати аналітичний вираз щільності розподілу

$$p(y) = \frac{1}{1,5 \cdot \sqrt{2\pi}} e^{-\frac{(y-7)^2}{4,5}}.$$

10.2. Математичне сподівання функції одного випадкового аргументу

Розглянемо перший випадок. Нехай аргумент ξ – дискретна ВВ. І нехай задана функція $\eta = \varphi(\xi)$ аргументу ξ . Необхідно знайти математичне сподівання цієї функції, якщо розподіл аргументу ξ заданий таблицею

ξ	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Вочевидь η також дискретна ВВ зі значеннями $y_1 = \varphi(x_1)$, $y_2 = \varphi(x_2)$, ..., $y_n = \varphi(x_n)$ та тими самими ймовірностями. Розподіл нової ВВ задається таблицею

η	y_1	y_2	...	y_n
P	p_1	p_2	...	p_n

Отже математичне сподівання дискретної ВВ η визначається як

$$\mathbf{M}[\varphi(\xi)] = \sum_{i=1}^n \varphi(x_i) \cdot p_i = \sum_{i=1}^n y_i \cdot p_i. \quad (10.4)$$

Розглянемо другий випадок. Нехай аргумент ξ – неперервна ВВ. І нехай аргумент ξ задано щільністю розподілу $p(x)$. Для визначення математичного сподівання функції $\eta = \varphi(\xi)$ знайдемо спочатку щільністю розподілу $g(y)$ ВВ η , після чого отримаємо

$$\mathbf{M}\eta = \int_{-\infty}^{\infty} yg(y)dy. \quad (10.5)$$

Якщо визначення $g(y)$ ускладнене, можна скористатись іншою формулою:

$$\mathbf{M}[\varphi(\xi)] = \int_{-\infty}^{\infty} \varphi(x) \cdot p(x)dx. \quad (10.6)$$

10.3. Функція двох випадкових аргументів. Розподіл суми незалежних доданків

Якщо кожній парі можливих значень ВВ ξ та η відповідає одне можливе значення ВВ τ , то τ називають *функцією двох випадкових аргументів* ξ та η : $\tau = \varphi(\xi, \eta)$.

Характерний приклад функції двох випадкових аргументів – $\tau = \xi + \eta$. Задача визначення розподілу ВВ τ за розподілами складових часто виникає в інженерній практиці. До прикладу, якщо ξ – похибка показань вимірювального приладу (з гауссовим розподілом), а η – похибка заокруглення показань до найближчої поділки шкали (розподілена рівномірно), то виникає задача пошуку закону розподілу суми похибок $\tau = \xi + \eta$. Розглянемо два випадки.

1. Нехай ξ та η – дискретні ВВ. Для того, щоби скласти закон розподілу функції $\tau = \xi + \eta$ необхідно знайти всі можливі значення τ та їх імовірності. Розглянемо таку задачу на прикладі.

Приклад 10.3. Дискретні незалежні ВВ задані розподілами

ξ	1	2
P	0,4	0,6

η	3	4
P	0,2	0,8

Необхідно скласти розподіл ВВ $\tau = \xi + \eta$.

Розв'язання. Можливі значення τ – це суми кожного можливого значення ξ зі всіма можливими значеннями η

$$z_1 = 1 + 3 = 4; z_2 = 1 + 4 = 5; z_3 = 2 + 3 = 5; z_4 = 2 + 4 = 6.$$

Визначимо ймовірності цих значень нової ВВ. Для того, щоби $\tau = 4$ достатньо аби тільки ξ набула значення 1, а $\eta = 3$. Ймовірності цих значень отримуємо з розподілів – це відповідно 0,4 та 0,2. Оскільки аргументи ξ та η незалежні, то події $\xi = 1$ та $\eta = 3$ також незалежні, тому ймовірність їх сумісної появи згідно з теоремою множення ймовірностей дорівнює $0,4 \cdot 0,2 = 0,08$. Аналогічно знайдемо:

$$P(\tau = 1 + 4 = 5) = 0,4 \cdot 0,8 = 0,32;$$

$$P(\tau = 2 + 3 = 5) = 0,6 \cdot 0,2 = 0,12;$$

$$P(\tau = 2 + 4 = 6) = 0,6 \cdot 0,8 = 0,48.$$

Залишилось записати розподіл у виді таблиці:

τ	4	5	6
P	0,08	0,32+0,12=0,44	0,48

Перевіряємо суму ймовірностей повної групи подій : $0,08 + 0,44 + 0,48 = 1$.

2. Нехай ξ та η – неперервні ВВ. Доведено наступне твердження: якщо ξ та η незалежні і мають щільності розподілу відповідно $p_1(x)$ та $p_2(x)$ то щільність розподілу $g(z)$ суми $\tau = \xi + \eta$ (за умови, що щільність розподілу хоча би одного з аргументів задана на інтервалі $(-\infty, \infty)$ однією формулою) визначається згорткою цих розподілів

$$g(z) = \int_{-\infty}^{\infty} p_1(x)p_2(z-x)dx \quad (10.7)$$

або за рівносильним виразом

$$g(z) = \int_{-\infty}^{\infty} p_1(z-y)p_2(y)dy . \quad (10.8)$$

Означення. Диференціальну функцію (щільність розподілу) суми незалежних ВВ називають *композицією*.

Закон розподілу ймовірностей називають *стійким*, якщо композиція таких законів і самі закони складових відрізняються лише параметрами. Гауссовий розподіл має властивість стійкості: композиція гауссових розподілів також є гауссовим розподілом, математичне сподівання та дисперсія якого дорівнюють відповідно сумам математичних сподівань та дисперсій складових.

10.4. Функція багатьох випадкових аргументів. Метод лінеаризації функцій ВВ

В інженерній практиці часто зустрічаються ймовірнісні задачі, які можуть бути наближено розв'язані без знання законів розподілу ймовірностей ВВ. Для їх розв'язання достатньо визначити числові характеристики ВВ – математичне сподівання, дисперсію та коефіцієнти кореляції. До кола таких завдань належит і оцінювання похибок результатів непрямих вимірювань. Рівняння вимірювання в цьому випадку представляється функціональною залежністю результатів x_1, x_2, \dots, x_n прямих вимірювань

$$y = f(x_1, x_2, \dots, x_n) . \quad (10.9)$$

Прикладами таких задач є наступні: 1) необхідно визначити дисперсію вимірювання швидкості руху об'єкта за формулою $v=s/t$, якщо шлях s , який пройшов об'єкт та відповідний час t отримано з певними середніми квадратичними похибками; 2) необхідно визначити дисперсію значення об'єма правильної призми з прямокутником в основі $V=ab c$, сторони призми a, b, c отримано з певними середніми квадратичними похибками. Величини s, t, a, b, c в таких задачах розглядаються як ВВ.

Передбачається, що функція (10.9) є неперервною.

Розглядаючи результати вимірювань як реалізації ВВ $\xi_1, \xi_2, \dots, \xi_n$, а значення y – як реалізацію ВВ η , ймовірнісну модель вимірювання представимо у виді

$$\eta = f(\xi_1, \xi_2, \dots, \xi_n). \quad (10.10)$$

Значна частина нелінійних функціональних залежностей, які зустрічаються на практиці, в обмеженому діапазоні значень можуть бути наближено представлені лінійними, до прикладу в околі математичних сподівань їх випадкових аргументів. На цьому припущенні власне і ґрунтується метод лінеаризації функцій ВВ.

Означення. *Лінеаризацією* функції називається наближена заміна нелінійної функції лінійною. Замінивши нелінійну функцію випадкових аргументів лінійною отримуємо можливість знаходити числові характеристики ВВ η за числовими характеристиками випадкових аргументів $\xi_1, \xi_2, \dots, \xi_n$, застосовуючи для цього властивості математичного сподівання і дисперсії. До прикладу, для лінійної функції одного випадкового аргументу виду $\eta = k\xi + a$, маємо $\mathbf{M}\eta = k\mathbf{M}\xi + a$, $\mathbf{D}\eta = k^2\mathbf{D}\xi$.

Обґрунтуємо метод лінеаризації функцій ВВ для опрацювання результатів непрямих вимірювань наступним чином.

Розглянемо спочатку задачу лінеаризації функції одного випадкового аргументу $\eta = f(\xi)$; передбачається, що η, ξ – неперервні ВВ.

З курсу вищої математики відомо, що будь-яка неперервна диференційована функція $y = f(x)$ може бути розкладена в ряд Тейлора в околі певної точки a

$$f(x) = f(a) + f'(a) \cdot (x - a) + f''(a) \cdot \frac{(x - a)^2}{2} + \dots, \quad (10.11)$$

$$\text{де } f'(a) = \left. \frac{\partial f(x)}{\partial x} \right|_{x=a}; \quad f''(a) = \left. \frac{\partial^2 f(x)}{\partial x^2} \right|_{x=a}.$$

З математичної точки зору лінеаризація функції одного випадкового аргументу є наближене представлення цієї функції першими двома членами ряду

Тейлора. Геометрично лінеаризована функція уявляється дотичною до функції $y = f(x)$ в точці a .

Перейдемо від ВВ η, ξ до їх не випадкових значень y, x , оскільки, строго кажучи, диференціювати по випадковим аргументам не можна. Розглянемо ряд Тейлора в околі точки $\mathbf{M}\xi$ і обмежимося двома першими членами розкладу

$$y \approx f(\mathbf{M}\xi) + f'(\mathbf{M}\xi) \cdot (x - \mathbf{M}\xi) . \quad (10.12)$$

Можна очікувати, що таке саме лінійне співвідношення пов'язує ВВ η, ξ

$$\eta \approx f(\mathbf{M}\xi) + f'(\mathbf{M}\xi) \cdot (\xi - \mathbf{M}\xi) . \quad (10.13)$$

Числові характеристики для цієї лінеаризованої функції визначаються як

$$\mathbf{M}\eta = \mathbf{M} [f(\mathbf{M}\xi) + f'(\mathbf{M}\xi) \cdot (\xi - \mathbf{M}\xi)] = f(\mathbf{M}\xi) , \quad (10.14)$$

$$\mathbf{D}\eta = [f'(\mathbf{M}\xi)]^2 \cdot \mathbf{D}\xi . \quad (10.15)$$

Поширюючи цей результат на випадок "n" незалежних аргументів маємо

$$\mathbf{M}\eta = f(\mathbf{M}\xi_1, \mathbf{M}\xi_2, \dots, \mathbf{M}\xi_n) , \quad (10.16)$$

$$\mathbf{D}\eta = \sum_{j=1}^n [f'(\mathbf{M}\xi_j)]^2 \cdot \mathbf{D}\xi = \sum_{j=1}^n \left(\left. \frac{\partial f(x)}{\partial x_j} \right|_{x=\mathbf{M}\xi_j} \right)^2 \cdot \sigma_{\xi_j}^2 . \quad (10.17)$$

Похідні $\left. \frac{\partial f(x)}{\partial x_j} \right|_{x=\mathbf{M}\xi}$ в рівнянні (10.17) називають *коефіцієнтами впливу*.

Якщо аргументи $\xi_1, \xi_2, \dots, \xi_n$ корельовані, їх зв'язки необхідно врахувати у формулі (10.17), яка трансформується до такого виду

$$\mathbf{D}\eta = \sum_{j=1}^n \left(\left. \frac{\partial f}{\partial x_j} \right|_{x=\mathbf{M}\xi_j} \right)^2 \cdot \sigma_{\xi_j}^2 + 2 \sum_{i < j} \left(\left. \frac{\partial f}{\partial x_i} \right|_{x=\mathbf{M}\xi_i} \right) \cdot \left(\left. \frac{\partial f}{\partial x_j} \right|_{x=\mathbf{M}\xi_j} \right) \cdot r_{ij} \cdot \sigma_{\xi_i} \cdot \sigma_{\xi_j} , \quad (10.18)$$

де r_{ij} – коефіцієнт кореляції ВВ ξ_i, ξ_j

На завершення розгляду цього питання наведемо наступне означення.

Означення. Функція $f(\xi_1, \xi_2, \dots, \xi_n)$ декількох випадкових аргументів

$\xi_1, \xi_2, \dots, \xi_n$ називається *майже лінійною*, якщо у всьому діапазоні практично можливих значень аргументів вона може бути з достатньою для практики точністю замінена лінійною функцією

$$f(\xi_1, \xi_2, \dots, \xi_n) \approx f(\mathbf{M}_{\xi_1}, \mathbf{M}_{\xi_2}, \dots, \mathbf{M}_{\xi_n}) + \sum_{j=1}^n \left(\left. \frac{\partial f}{\partial x_j} \right|_{x=\mathbf{M}_{\xi_j}} \right) \cdot (\xi_j - \mathbf{M}_{\xi_j}) \quad (10.19)$$

Нижче наведені формули диференціювання, які найчастіше використовуються для визначення коефіцієнтів впливу:

$$(x_1 x_2)' = x_1' x_2 + x_1 x_2'; \quad \left(\frac{x_1}{x_2} \right)' = \frac{x_1' x_2 - x_1 x_2'}{x_2^2};$$

$$(x_1^{x_2})' = x_2 x_1^{x_2-1} x_1' + x_1^{x_2} \ln x_1 \cdot x_2';$$

$$(\alpha x)' = \alpha; \quad (x^\alpha)' = \alpha x^{\alpha-1}; \quad (\sqrt{x})' = \frac{1}{2\sqrt{x}};$$

$$(e^x)' = e^x; \quad (a^x)' = a^x \ln a;$$

$$(\ln x)' = \frac{1}{x}; \quad (\sin x)' = \cos x; \quad (\cos x)' = -\sin x;$$

$$(\arctg x)' = \frac{1}{1+x^2}; \quad (tg x)' = \frac{1}{\cos^2 x}.$$

Лекція 11.

СИСТЕМА ДВОХ ВИПАДКОВИХ ВЕЛИЧИН

11.1. Загальне поняття про систему декількох ВВ

Крім одновимірних ВВ, значення яких визначаються одним числом, в теорії ймовірностей досліджуються ВВ, можливі значення яких визначаються двома, трьома, ... n числами. Такі ВВ називають відповідно *двовимірними*, *тривимірними*, ... *n -вимірними*.

Найбільш поширеним в практиці інженерних розрахунків є двовимірні ВВ. Двовимірну ВВ будемо надалі позначати як (ξ, η) . Кожну з величин ξ та η

називають *складовою (компонентою) ВВ*. Обидві величини ξ та η , що розглядаються одночасно, створюють систему двох ВВ (але функціонально вони не пов'язані). Двовимірний ВВ має місце, до прикладу, під час вимірювання двох параметрів деталі – внутрішнього та зовнішнього діаметрів труби, визначення площі дефекту на поверхні деталі за двома декартовими координатами тощо.

Геометрично двовимірний ВВ (ξ, η) можна інтерпретувати як випадкову точку на площині, тобто як точку з випадковими координатами.

Розрізняють дискретні та неперервні багатовимірні ВВ.

11.2. Закон розподілу ймовірностей дискретної двовимірної ВВ

Означення. *Законом розподілу дискретної двовимірної ВВ* називають перелік можливих значень компонент випадкової величини (ξ, η) (тобто пар чисел (x_i, y_j)) та їх ймовірностей $P(x_i, y_j)$, $i=1, 2, \dots, n, j=1, 2, \dots, m$. Такий розподіл зручно представляти у вигляді таблиці (див. Таблицю 11.1).

Табл. 11.1. Загальний вид розподілу дискретної двовимірної ВВ

η	ξ	x_1	...	x_i	...	x_n	$\sum_{j=1}^m P(x_i, y_j)$
y_1		$P(x_1, y_1)$...	$P(x_i, y_1)$...	$P(x_n, y_1)$	$P(y_1)$
...		
y_j		$P(x_1, y_j)$...	$P(x_i, y_j)$...	$P(x_n, y_j)$	$P(y_j)$
...		
y_m		$P(x_1, y_m)$...	$P(x_i, y_m)$...	$P(x_n, y_m)$	$P(y_m)$
	$\sum_{j=1}^m P(x_i, y_j)$	$P(x_1)$		$P(x_i)$		$P(x_n)$	-

Оскільки представлені в Таблиці 11.1 події утворюють повну групу подій, то сума їх ймовірностей дорівнює одиниці. Знаючи розподіл двовимірної дискретної ВВ можна знайти розподіл кожної компоненти. До прикладу, події $(\xi=x_1, \eta=y_1)$, $(\xi=x_1, \eta=y_2), \dots, (\xi=x_1, \eta=y_m)$ несумісні, тому ймовірність $P(x_1)=P(\xi=x_1)$ згідно з теоремою додавання ймовірностей (крайній лівий стовпчик Таблиці 11.1) дорівнює

$$P(x_1) = P(x_1, y_1) + P(x_1, y_2) + \dots + P(x_1, y_m) = \sum_{j=1}^m P(x_1, y_j). \quad (11.1)$$

Аналогічно склавши «рядок y_j » отримаємо ймовірність $P(y_j) = P(\eta = y_j)$.

11.3. Інтегральна функція двовимірної ВВ та її властивості

Означення. Інтегральною функцією двовимірної ВВ (ξ, η) називають функцію $F(x, y)$, що визначає для кожної пари чисел x, y ймовірність того, що ξ набуде значення меншого за x і одночасно η матиме значення менше за y

$$F(x, y) = P(\xi < x, \eta < y). \quad (11.2)$$

Геометрично це визначення можна трактувати наступним чином: функція $F(x, y)$ – це ймовірність того, що ВВ (ξ, η) потрапить на площині з декартовою системою координат XOY у нескінченний квадрант з вершиною в точці (x, y) , розміщений ліворуч та нижче цієї вершини (рис. 11.1).

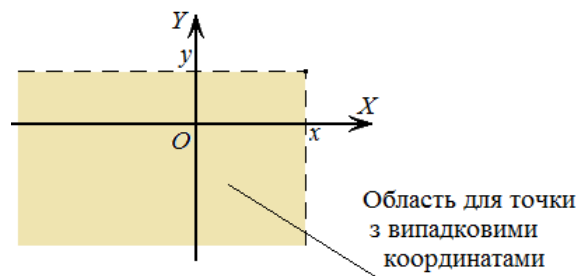


Рис. 11.1. Геометрична інтерпретація змісту інтегральної функції двовимірної ВВ

Функція $F(x, y)$ має наступні властивості.

1. Значення інтегральної функції задовольняє подвійній нерівності

$$0 \leq F(x, y) \leq 1;$$

2. $F(x, y)$ – неспадна функція, тобто

$$F(x_2, y) \geq F(x_1, y) \quad \text{якщо } x_2 > x_1,$$

$$F(x, y_2) \geq F(x, y_1) \quad \text{якщо } y_2 > y_1.$$

3. Мають місце граничні співвідношення:

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

$$F(-\infty, -\infty) = 0; \quad F(\infty, \infty) = 1.$$

4. Якщо $x_1 \leq \xi < x_2, \eta < y$, то $P(x_1 \leq \xi < x_2, \eta < y) = F(x_2, y) - F(x_1, y)$.
5. За умови $y \rightarrow \infty$ інтегральна функція системи ВВ перетворюється на інтегральну функцію компоненти ξ , тобто $F(x, \infty) = F_1(x)$; за умови $x \rightarrow \infty$ інтегральна функція системи перетворюється на інтегральну функцію компоненти η , тобто $F(\infty, y) = F_2(y)$.

11.4. Диференціальна функція неперервної двовимірної ВВ та її властивості

Якщо $F(x, y)$ всюди неперервна і всюди має неперервну змішану частинну похідну другого порядку, то має сенс говорити про диференціальну функцію двовимірної ВВ.

Означення. Диференціальною функцією розподілу $p(x, y)$ двовимірної неперервної ВВ (ξ, η) називають другу змішану частинну похідну від інтегральної функції

$$p(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} . \quad (11.3)$$

Геометрично функція $p(x, y)$ трактується як поверхня, яку називають *поверхнею розподілу*.

Ймовірнісний сенс диференціальної функції неперервної двовимірної ВВ трактується наступним чином: функцію $p(x, y)$ можна розглядати як межу відношення ймовірності попадання випадкової точки на площині з випадковими координатами (ξ, η) в прямокутник зі сторонами $\Delta x, \Delta y$ до площі цього прямокутника, якщо $\Delta x, \Delta y \rightarrow 0$, тобто

$$p(x, y) = \lim_{\Delta x, \Delta y \rightarrow 0} \frac{P(x - 0,5\Delta x < \xi \leq x + 0,5\Delta x, y - 0,5\Delta y < \eta \leq y + 0,5\Delta y)}{\Delta x \Delta y} . \quad (11.4)$$

Диференціальна функція двовимірної ВВ має властивості:

1. Диференціальна функція неперервної двовимірної ВВ невід'ємна

$$p(x, y) \geq 0 .$$

2. Двократний невластний інтеграл від диференціальної функції неперервної двовимірної ВВ з нескінченними межами дорівнює одиниці:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = 1 .$$

3. Диференціальні функції компонент неперервної двовимірної ВВ визначаються наступним чином

$$p_1(x) = \int_{-\infty}^{\infty} p(x, y) dy, \quad p_2(y) = \int_{-\infty}^{\infty} p(x, y) dx. \quad (11.5)$$

11.5. Числові характеристики системи двох ВВ: кореляційний момент та коефіцієнт кореляції

Для опису системи двох ВВ крім математичних сподівань та дисперсій її компонент користуються й іншими характеристиками, до кола яких належать кореляційний момент і коефіцієнт кореляції.

Означення. Кореляційним моментом $\mu_{x,y}$ ВВ ξ та η називають математичне сподівання добутку відхилень цих ВВ від їх математичних сподівань

$$\mu_{x,y} = \mathbf{M}[(\xi - \mathbf{M}\xi) \cdot (\eta - \mathbf{M}\eta)]. \quad (11.6)$$

Кореляційний момент дискретних ВВ обчислюють за формулою

$$\mu_{x,y} = \sum_{i=1}^n \sum_{j=1}^m (x_i - \mathbf{M}\xi) \cdot (y_j - \mathbf{M}\eta) \cdot P(x_i, y_j), \quad (11.7)$$

а для неперервних ВВ –

$$\mu_{x,y} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mathbf{M}\xi) \cdot (y - \mathbf{M}\eta) \cdot p(x, y) dx dy . \quad (11.8)$$

Характеристика $\mu_{x,y}$ дає уявлення про зв'язок між ξ та η . Для незалежних ВВ маємо $\mu_{x,y} = 0$. Дві ВВ називають *корельованими*, якщо $\mu_{x,y} \neq 0$.

Дві корельовані ВВ також є залежними. Обернене твердження не завжди має місце: залежні величини можуть виявитись як корельованими, так і

некорельованими.

Означення. Коефіцієнтом кореляції $r_{x,y}$ ВВ ξ та η називають відношення

$$r_{x,y} = \frac{\mu_{x,y}}{\sigma_x \sigma_y}. \quad (11.9)$$

У випадку, якщо ξ та η розмірні величини, розмірність $\mu_{x,y}$ дорівнює добутку розмірностей цих ВВ. Натомість $r_{x,y}$ – безрозмірна величина і змінюється в зручному для аналізу обмеженому інтервалі значень $-1 \leq r_{x,y} \leq 1$.

11.6. Умовні закони розподілу складових системи ВВ

Раніше було встановлено: якщо події A і B залежні, то умовна ймовірність події B відмінна від її безумовної ймовірності і визначається як $P(B|A) = \frac{P(AB)}{P(A)}$.

Аналогічне положення має місце і для ВВ. Нехай (ξ, η) – двовимірна ВВ, і ВВ ξ та η залежні між собою. Умовний розподіл характеризує залежність ВВ.

Означення. Умовним розподілом однієї з ВВ називається розподіл цієї величини, визначений за умови, що інша ВВ прийняла певне значення (або потрапила в певний інтервал значень).

Умовні розподіли мають місце як для дискретних, так і для неперервних ВВ. Розглянемо ці два випадки.

11.6.1. Умовні закони розподілу складових системи дискретних ВВ

Розглянемо дискретну двовимірну ВВ (ξ, η) , розподіл якої визначає набір ймовірностей $P(\xi = x_i, \eta = y_j)$, $i = 1, \dots, n$, $j = 1, \dots, m$.

Нехай можливими значенням компонент ВВ є такі значення

$$x_1, x_2, \dots, x_n; \quad y_1, y_2, \dots, y_m.$$

Припустимо, що в результаті випробування ВВ η набула значення $\eta = y_1$; в цьому випадку ξ дорівнюватиме одному зі своїх можливих значень –

x_1, x_2, \dots чи x_n (див. тал. 11.1). Позначимо умовну ймовірність того, що ξ набуде, до прикладу, значення x_1 за умови $\eta = y_1$, через $P(x_1 | y_1)$. В загальному випадку умовні ймовірності позначатимемо як $P(x_i | y_j)$, ($i = \overline{1, n}; j = \overline{1, m}$).

Зазначимо, що у загальному випадку $P(x_i | y_j) \neq P(x_i)$.

Означення. Умовним розподілом дискретної компоненти ξ за умови $\eta = y_j$ називають сукупність умовних ймовірностей $P(x_1 | y_j), P(x_2 | y_j), \dots, P(x_n | y_j)$, обчислених за припущення, що подія $\eta = y_j$ вже відбулась.

В аналогічний спосіб визначається умовний розподіл компоненти η .

Якщо відомий закон розподілу двовимірної дискретної ВВ, можна обчислити, за аналогією до умовної імовірності подій, умовні закони розподілу її компонент. До прикладу, умовний закон розподілу ξ за умови, що подія $\eta = y_1$ вже відбулась, визначається за формулою

$$P(x_i | y_1) = \frac{P(x_i, y_1)}{P(y_1)}, \quad i = \overline{1, n}. \quad (7.10)$$

Аналогічно визначаються умовні закони розподілу складової η

$$P(y_j | x_i) = \frac{P(x_i, y_j)}{P(x_i)}, \quad i = \overline{1, n}; \quad j = \overline{1, m}. \quad (7.11)$$

Важливою властивістю умовного закону розподілу дискретної ВВ є те, що сума ймовірностей умовного розподілу дорівнює одиниці. Ця властивість умовних розподілів використовується для контролю правильності обчислень.

11.6.2. Умовні закони розподілу складових системи неперервних ВВ

Нехай (ξ, η) – неперервна двовимірна ВВ, і $p_1(x)$ та $p_2(y)$ – щільності розподілу ймовірностей випадкових величин ξ і η .

Означення. Умовною диференціальною функцією $\varphi(x|y)$ складової ξ за заданого значення $\eta = y$ називають відношення диференціальної функції $p(x, y)$ системи ВВ до диференціальної функції $p_2(y)$ складової η :

$$\varphi(x|y) = \frac{p(x, y)}{p_2(y)}. \quad (11.12)$$

Відмінність функції $\varphi(x|y)$ від безумовної диференціальної функції $p_1(x)$ полягає в тому, що $\varphi(x|y)$ дає розподіл ξ за умови, що складова η набула значення $\eta = y$, натомість функція $p_1(x)$ дає розподіл ВВ ξ незалежно від значення ВВ η .

Аналогічно визначається умовна диференціальна функція $\psi(y|x)$ складової η за даного значення $\xi = x$:

$$\psi(y|x) = \frac{p(x, y)}{p_1(x)}. \quad (11.13)$$

Якщо $p(x, y)$ відома, то умовні диференціальні функції можуть бути визначені згідно з такими виразами:

$$\varphi(x|y) = \frac{p(x, y)}{\int_{-\infty}^{+\infty} p(x, y) dx}, \quad \psi(y|x) = \frac{p(x, y)}{\int_{-\infty}^{+\infty} p(x, y) dy}. \quad (11.14)$$

Запишемо формули у (11.14) в іншому виді

$$p(x, y) = p_2(y) \cdot \varphi(x|y) = p_1(x) \cdot \psi(y|x). \quad (7.15)$$

З (11.15) випливає наступний висновок: множенням закону розподілу однієї складової на умовний закон розподілу іншої знаходимо закон розподілу системи неперервних ВВ.

Властивості умовної диференціальної функції:

1. $\psi(y|x) \geq 0$.
2. $\int_{-\infty}^{+\infty} \psi(y|x) dy = 1$.

11.6.3. Критерій незалежності ВВ

Випадкові величини ξ і η є незалежними тоді і тільки тоді, коли умовний розподіл ВВ ξ за умови $\eta = y$ збігається з безумовним розподілом ВВ ξ .

Зокрема, дискретні величини ξ і η є незалежними тоді і тільки тоді, коли всі умовні ймовірності збігаються з безумовними, тобто $\forall i, j: p(x_i | y_j) = p(x_i)$.

11.7. Умовні числові характеристики

Оскільки умовний розподіл має всі властивості безумовного розподілу, то за ним можна визначити числові характеристики, які називаються *умовними числовими характеристиками*.

11.7.1. Умовне математичне сподівання.

Важливою числовою характеристикою умовного розподілу ймовірностей є умовне математичне сподівання.

Означення. Умовне математичне сподівання дискретної ВВ η , якщо $\xi = x$ (де x – одне з можливих значень ξ) – це сума добутків можливих значень η на їх умовні ймовірності

$$\mathbf{M}(\eta | \xi = x) = \sum_{j=1}^m y_j P(y_j | x). \quad (11.16)$$

Умовне математичне сподівання неперервної випадкової величини η – це інтеграл виду

$$\mathbf{M}(\eta | \xi = x) = \int_{-\infty}^{\infty} y \cdot \psi(y | x) dy. \quad (11.17)$$

В такий самий спосіб визначається і $M(\xi | \eta = y)$.

Функція $g(\xi) = M(\eta | \xi)$ називається *функцією регресії* ВВ η на ВВ ξ , а її графік – *лінією регресії*. Лінія регресії графічно зображає залежність «в середньому» випадкової величини η від значень ξ .

Властивості умовного математичного сподівання:

1. $\mathbf{M}(c|\eta) \equiv \text{const}, c = \text{const}.$

2. $\mathbf{M}((a\xi + b)|\eta) = a\mathbf{M}(\xi|\eta) + b.$

3. $\mathbf{M}((\xi_1 + \xi_2)|\eta) = \mathbf{M}(\xi_1|\eta) + \mathbf{M}(\xi_2|\eta).$

4. Нехай ξ_1, ξ_2 – незалежні ВВ. За умови, що випадкова величини η набула будь-яке конкретне значення, маємо

$$\mathbf{M}((\xi_1 \cdot \xi_2)|\eta) = \mathbf{M}(\xi_1|\eta) \cdot \mathbf{M}(\xi_2|\eta). \quad (11.18)$$

5. $\mathbf{M}(\mathbf{M}(\xi|\eta)) = \mathbf{M}\xi.$

6. Нехай $u(\xi)$ та $v(\eta)$ – функції ВВ ξ и η , тоді

$$\mathbf{M}(u(\xi)v(\eta)|\eta) = v(\eta)\mathbf{M}(u(\xi)|\eta). \quad (11.19)$$

7. Якщо ξ та η – незалежні випадкові величини, то

$$\mathbf{M}(\xi|\eta) \equiv \mathbf{M}\xi.$$

11.7.2. Умовна дисперсія

Означення. Умовною дисперсією $\mathbf{D}(\xi|\eta)$ ВВ ξ відносно ВВ η називають

$$\mathbf{D}(\xi|\eta) = \mathbf{M}\left(\left[\xi - \mathbf{M}(\xi|\eta)\right]^2|\eta\right). \quad (11.20)$$

Формула (11.20) для дискретних та неперервних ВВ набуває відповідно такого вигляду

$$\mathbf{D}(\xi|y_j) = \sum_{i=1}^n \left[x_i - \mathbf{M}(\xi|y_j) \right]^2 P(x_i|y_j), \quad (11.21)$$

$$\mathbf{D}(\xi|y) = \int_{-\infty}^{+\infty} \left[x - \mathbf{M}(\xi|y) \right]^2 p(x|y) dx. \quad (11.22)$$

Властивості умовної дисперсії:

1. $\mathbf{D}(\text{const}|\eta) \equiv 0.$

2. $\mathbf{D}((a\xi + b)|\eta) = a^2\mathbf{D}(\xi|\eta).$

$$3. \mathbf{D}(\xi|\eta) = \mathbf{M}(\xi^2|\eta) - \mathbf{M}^2(\xi|\eta).$$

4. Нехай ξ_1, ξ_2 – незалежні ВВ за умови, що ВВ η набула будь-яке конкретне значення, тоді

$$\mathbf{D}(\xi_1 + \xi_2|\eta) = \mathbf{D}(\xi_1|\eta) + \mathbf{D}(\xi_2|\eta). \quad (11.23)$$

5. Нехай $u(\xi)$ та $v(\eta)$ – функції ВВ ξ та η , тоді

$$\mathbf{D}(u(\xi)v(\eta)|\eta) = v^2(\eta)\mathbf{D}(u(\xi)|\eta). \quad (11.24)$$

6. Якщо ξ та η – незалежні ВВ, то

$$\mathbf{D}(\xi|\eta) \equiv \mathbf{D}\xi.$$

$$7. \mathbf{D}\xi = \mathbf{M}(\mathbf{D}(\xi|\eta)) + \mathbf{M}(\mathbf{M}(\xi|\eta) - \mathbf{M}\xi)^2. \quad (11.25)$$

Питання для самоконтролю

1. Поняття множини. Дії з множинами. Приклади числових множин та множин точок на прямій.

2. Потужності множини та їх використання для порівняння множин.

3. Імовірнісна модель стохастичних експериментів з кінцевим числом наслідків та аксіоми, покладені в основу її побудови.

4. Властивості імовірності випадкових подій.

5. Основні терміни та визначення теорії ймовірностей: випробування, випадкова подія, імовірність події, вірогідні та неможливі, сумісні та несумісні події, протилежні події, повна група подій. Запишіть формули для найпростіші «поєднань» – перестановок, розміщень, сполучень.

6. Назвіть основні дії з множинами та поясніть їх зміст.

7. Наведіть основні формули комбінаторики.

8. Для яких випадків розрахунку ймовірностей застосовують поняття

«геометрична» ймовірність?

9. Наведіть приклади застосування розрахунку «геометричної» ймовірності.
10. Ймовірнісний простір та його компоненти.
11. Аксиоми теорії ймовірності.
12. Поняття умовної ймовірності подій.
13. Визначення та зміст понять залежні та незалежні випадкові події.
14. Теореми множення ймовірностей випадкових подій.
15. Теореми додавання ймовірностей випадкових подій.
16. Формула повної ймовірності. Формула Бейеса
17. Випадкова величина та її види.
18. Розподіл дискретної ВВ.
19. Визначення математичного сподівання дискретної ВВ та його властивості.
20. Визначення дисперсії дискретної ВВ та наведіть її властивості.
21. Сформулюйте теорему, яка визначає дисперсію числа появи події A в n незалежних випробуваннях.
22. Сформулюйте теорему про математичне сподівання і дисперсію середнього арифметичного.
23. Дайте визначення початкових і центральних моментів ВВ.
24. Розкрийте зміст схеми Бернуллі.
25. Дайте пояснення локальної теореми Лапласа
26. Розкрийте зміст інтегральної теореми Муавра-Лапласа
27. Розподіл Пуассона та зміст його параметра.
28. Нерівність Чебишова.
29. Теорема Чебишова та її практичне значення.
30. Теорема Бернуллі.
31. Рівномірний розподіл та його характеристики.
32. Розподіл Сімпсона та його характеристики.
33. Розподіл Гаусса та його характеристики.
34. Функція розподілу імовірностей ВВ та її властивості.

35. Щільність розподілу імовірностей неперервної ВВ та її властивості.
36. Розподіл χ^2 (розподіл Пірсона).
37. Визначення та формула математичного сподівання неперервної ВВ.
38. Сформулюйте властивості математичного сподівання.
39. Визначення та формула неперервної ВВ.
40. Як визначається розподіл функції дискретного випадкового аргументу?
41. Як визначається розподіл функції неперервної випадкового аргументу?
42. Як визначається математичне сподівання функції одного дискретного випадкового аргументу?
43. Як визначається математичне сподівання функції одного неперервного випадкового аргументу?
44. Наведіть методику визначення розподілу суми двох дискретних ВВ.
45. Наведіть формули композиції щільностей розподілу двох ВВ та необхідні умови для їх застосування.
46. Розкрийте зміст поняття «стійкий закон розподілу ймовірностей».
47. Дайте визначення та розкрийте зміст умовних розподілів ВВ.
48. Запишіть вираз для умовної щільності розподілу.
49. Сформулюйте критерій незалежності ВВ, використовуючи поняття умовної ймовірності.
50. Дайте визначення умовного математичного сподівання дискретної і неперервної ВВ.
51. Властивості умовного математичного сподівання.
52. Дайте визначення умовної дисперсії дискретної і неперервної ВВ.
53. Властивості умовної дисперсії.

Розділ 3. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

Лекція 12.

ОСНОВИ СТАТИСТИЧНОГО ОПРАЦЮВАННЯ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ

Відомо, що повну інформацію про ВВ містить її закон розподілу ймовірності. Встановлення закону розподілу ВВ за результатами експериментів – основне завдання математичної статистики. Можна сказати, що теорія ймовірності і математична статистика вирішують взаємно зворотні завдання. Теорія ймовірності за відомими законами розподілу передбачає результати випадкових експериментів (наскільки можливий той або інший результат, чи має місце залежність величин, наскільки вона сильна і т. п.). Математична статистика за результатами експерименту намагається встановити закони розподілу, залежності ВВ, рівність або відмінності розподілів та ін.

12.1. Основні терміни та визначення

В ході дослідження явищ навколишнього середовища, вимірювання фізичних величин, контролю якості продукції і т.і. завжди мають справу зі спостереженнями (результатами вимірювань), отриманими на об'єктах різної природи. Як спостереження можуть розглядатись такі процеси як вимірювання, випробування, тестування тощо. В цьому контексті *спостереження* розуміють не як процес отримання інформації про властивість об'єкта, а як певне число яке кількісно характеризує цю властивість і є власне *результатом спостереження*. Відповідно *статистичні дані* – це результати спостережень, набори іменованих чи не іменованих чисел.

Результат спостереження – зафіксоване значення величини або числа за показуючим пристроєм засобу вимірювальної техніки в заданий момент часу.

Результат спостереження можна прийняти як результат вимірювання у випадку вимірювань з однократними спостереженнями. Якщо вимірювання

проводиться з багаторазовими спостереженнями, то результат вимірювання отримують шляхом опрацювання ряду спостережень, яке ґрунтується на статистичних методах.

Завдання математичної статистики полягає в тому, щоб на підставі знання деяких властивостей підмножини елементів, взятих із деякої множини, висловити деякі твердження про властивості цієї множини – генеральної сукупності. Сукупність, яка об'єднує всі елементи, що мають деяку спільну ознаку, називається *генеральною*. Наприклад, якщо ознака людська (національність, стать, освіта, коефіцієнт інтелекту IQ), то генеральна сукупність – все населення Землі. Це дуже велика сукупність, тобто число елементів в сукупності n велике. Число елементів називається *обсягом сукупності*. Таким чином, уся множина об'єктів, що підлягає дослідженню, є *генеральною сукупністю* (рис. 12.1). Сукупності можуть бути кінцевими і нескінченними. Генеральна сукупність «всі люди» хоч і дуже велика, проте звичайно кінцева. Генеральна сукупність «усі зірки» – нескінченна.

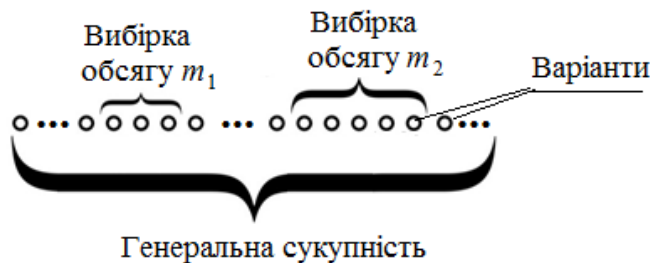


Рис. 12.1 Формування вибірок з генеральної сукупності

Якщо дослідник проводить вимірювання деякої неперервної ВВ ξ , то кожен результат вимірювання x можна вважати елементом деякої необмеженої генеральної сукупності. У цій сукупності незліченна кількість результатів розподілена за ймовірністю під впливом похибок в приладах, неуважності експериментатора, випадкових завад в самому явищі і т.і.

Якщо проведемо n повторних вимірювань, тобто отримаємо n конкретних різних числових значень x_1, x_2, \dots, x_n , то цей результат експерименту в

ймовірнісному сенсі можна розглядати як вибірку *обсягу* n з гіпотетичної генеральної сукупності результатів одиничних вимірювань сукупності n випадкової величини $\xi_1, \xi_2, \dots, \xi_n$, які мають однакові розподіли.

Вибірка – це одна або декілька вибірових одиниць, узятих з генеральної сукупності і призначених для отримання інформації про неї. Елементи вибірки називають **варіантами**. Вибірковою одиницею може бути один з елементів, з яких складається генеральна сукупність, або певна кількість продукції, матеріалу, які створюють сукупність і узяті з одного місця, в один час для формування частини вибірки. До прикладу, для задач контролю та діагностики одиницею є результат одного спостереження під час вимірювання, а вибіркою – сукупність результатів спостережень при проведенні вимірювань з багаторазовими спостереженнями.

Вибірка повинна достатньо повно відображати особливості всіх об'єктів генеральної сукупності, щоб отримувані оцінки були вірогідними, тобто вона має бути репрезентативною (представницькою). Вибірка буде **репрезентативною**, якщо відбір елементів у вибірку проводиться випадково. Це означає, що всі елементи генеральної сукупності мають однакову ймовірність потрапити у вибірку. Якщо вибірка виявиться не репрезентативною (її називають **зміщеною**), то з ростом її обсягу точність певних статистичних оцінок може навіть зменшитися або можуть бути зроблені помилкові висновки.

Функції результатів спостережень $f(x_1, x_2, \dots, x_i \dots x_n)$, що використовуються для оцінювання параметрів розподілів або для перевірки статистичних гіпотез, називають **статистиками** або **статистичними оцінками**. Якщо в імовірнісній моделі числові значення x_1, x_2, \dots, x_n , розглядаються як результати спостережень ВВ $\xi_1, \xi_2, \dots, \xi_n$, то статистики, як функції ВВ, тобто $f(\xi_1, \xi_2, \dots, \xi_n)$, самі є випадковими величинами. Статистики, що є вибіровими аналогами характеристик ВВ (математичного сподівання, медіани, дисперсії, моментів) і використовувані для оцінювання цих характеристик, називають **статистичними характеристиками**.

У ймовірнісній моделі вибірки спостережуваних значень розглядають як реалізацію незалежних однаково розподілених ВВ $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$, $\omega \in \Omega$. При цьому вважають, що отримані в ході спостережень конкретні значення x_1, x_2, \dots, x_n відповідають певній елементарній події $\omega = \omega_0$, тобто

$$x_1 = \xi_1(\omega_0), x_2 = \xi_2(\omega_0), \dots, x_n = \xi_n(\omega_0), \quad \omega_0 \in \Omega. \quad (12.1)$$

У випадку повторних спостережень будуть отримані інші значення, які відповідатимуть іншій елементарній події $\omega = \omega_1$. Мета опрацювання статистичних даних полягає в тому, щоб за наслідками спостережень, які відповідають елементарній події $\omega = \omega_0$, зробити висновки про ймовірнісну міру P і результати спостережень при різних можливих ω .

Варіаційним рядом $x_{[1]}, x_{[2]}, \dots, x_{[n]}$ називається вибірка, записана в порядку зростання її варіант $x_{[1]} \leq x_{[2]} \leq \dots \leq x_{[n]}$.

Елементи варіаційного ряду позначаються як $x_{[i]}$.

У випадку дослідження неперервної (або з обсягом варіант $n > 100$) ВВ варіаційний ряд розбивається на класи (інтервали) і отримується інтервальний варіаційний ряд. Ширину інтервалів Δx визначають формулою

$$\Delta x = (x_{[n]} - x_{[1]}) / k, \quad (12.2)$$

де k – кількість класів.

Оптимальна кількість класів залежить від обсягу вибірки. Для визначення кількості класів можна користуватися Таблицею 12.1.

Таблиця 12.1. Визначення кількості класів в залежності від обсягу вибірки

Обсяг вибірки n	25÷40	40÷60	60÷80	100÷200	200÷1000
Кількість класів k	5÷6	6÷8	7÷10	8÷12	10÷15

Припустимо, що зроблено n вимірювань випадкової дискретної величини ξ і отримано k різних значень x_1, x_2, \dots, x_k . При цьому значення x_1 спостерігалось m_1 разів, x_2 – m_2 разів, \dots , x_k – m_k разів. Отже з нескінченної гіпотетичної генеральної

сукупності результатів вимірювань зроблено вибірку обсягу $n = m_1 + m_2 + \dots + m_k$. Числа m_1, m_2, \dots, m_k називають **частотами** спостережуваних значень x_1, x_2, \dots, x_k . Величини $\tilde{m}_1 = m_1 / n, \tilde{m}_2 = m_2 / n, \dots, \tilde{m}_k = m_k / n$ називають **відносними частотами** варіант x_i . Ясно, що $\tilde{m}_1 + \tilde{m}_2 + \dots + \tilde{m}_k = 1$. Отримані результати зручно представити у табличному вигляді (Таблиця 12.2). Вважатимемо, що варіанти x_1, x_2, \dots, x_k розташовані в порядку зростання, тобто побудовано варіаційний ряд (перший рядок таблиці). Другий рядок таблиці являє собою варіаційний ряд для частот, третій – для відносних частот, четвертий – для кумулятивних (накопичених) відносних частот.

Таблиця 12.2. Структура таблиці для запису результатів спостережень

x_i	x_1	x_2	...	x_i	...	x_k
m_i	m_1	m_2	...	m_i	...	m_k
\tilde{m}_i	m_1 / n	m_2 / n	...	m_i / n	...	m_k / n
$F_i = \sum_{j=1}^i \tilde{m}_j$	m_1 / n	$\frac{m_1 + m_2}{n}$...	$\sum_{j=1}^i m_j / n$...	$\sum_{j=1}^k m_j / n = 1$

Якщо кількість варіант k не дуже велика, то для того, щоб отримати більш наочного уявлення про розподіл ВВ ξ , будують *полігони частот*. Для цього на осі абсцис відкладають значення варіант x_1, x_2, \dots, x_k , а на осі ординат відповідні значення частот m_1, m_2, \dots, m_k , або відносних частот, або кумулятивних відносних частот F_i . Очевидно, що полігон відносних частот дає уявлення про розподіл ймовірностей, а графік кумулятивних відносних частот називають **емпіричною функцією розподілу ймовірностей**. Ця функція виражає залежність між значеннями кількісної ознаки і накопиченою частотою. Отже, *емпіричною функцією розподілу* $F_n(x)$ називається частка елементів вибірки, менших x . Ця функція містить всю інформацію про результати спостережень. Вона визначена на всій числовій осі. Ясно, що $F(x) = 0$ для всіх $x < x_1$ і $F(x) = 1$ для всіх $x > x_k$. На інтервалі $x_1 < x < x_k$ функція $F(x)$ має вид ступінчастої монотонно зростаючої від

0 до 1 функції, такої, що $F(a) = P(\xi < a)$ (рис. 12.2).

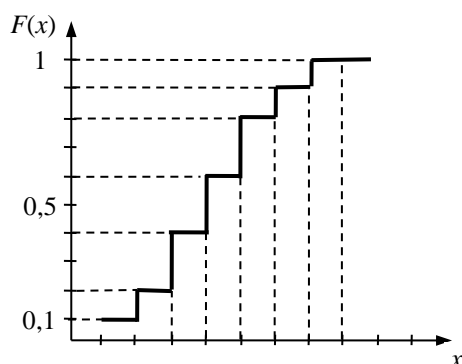


Рис. 12.2. Приклад емпіричної функції розподілу імовірностей

У випадку інтервального варіаційного ряду частота дорівнює загальній кількості варіант в даному класі. Всі класи, крім останнього, є напіввідкритими справа інтервалами (наприклад, $[a_i; a_{i+1})$), а останній закритий $[a_{k-1}; a_k]$.

Для запису емпіричної функції розподілу у вигляді формули зручно зкористатись функцію $c(x, y)$ двох змінних

$$c(x, y) = \begin{cases} 0, & x \leq y, \\ 1, & x > y. \end{cases} \quad (12.3)$$

Випадкові величини, що формують результати спостережень, позначимо $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$, $\omega \in \Omega$. Тоді емпірична функція розподілу $F_n(x)$ має вигляд

$$F_n(x) = F_n(x, \omega) = \frac{1}{n} \sum_{i=1}^n c(x, \xi_i(\omega)). \quad (12.4)$$

Згідно із законом великих чисел для кожного дійсного числа x емпірична функція розподілу $F_n(x)$, за умови $n \rightarrow \infty$, сходиться до функції розподілу $F(x)$ результатів спостережень, тобто

$$F_n(x) \rightarrow F(x). \quad (12.5)$$

Для інтервального варіаційного ряду також можна скласти таблицю (Таблиця 12.3), в якій a_i – границя класових інтервалів.

Якщо на осі абсцис відкласти класові інтервали і над ними побудувати прямокутники з висотами, рівними відповідним щільностям \tilde{m}_i відносної частоти,

то площа кожного прямокутника дорівнюватиме відносній частоті $S_i = \Delta x p_i = \Delta x \tilde{m}_i / \Delta x = \tilde{m}_i$. Отримана в такий спосіб ступінчаста фігура називається **гістограмою** (рис. 12.3).

Таблиця 12.3. Варіант таблиці для інтервального варіаційного ряду

Номер класу	1	2	...	k
Класовий інтервал	$[a_0; a_1)$	$[a_1; a_2)$...	$[a_{k1}; a_k)$
Частота	m_1	m_2	...	m_k
Відносна частота	$\tilde{m}_1 = m_1 / n$	\tilde{m}_2	...	\tilde{m}_k
Щільність відносної частоти	$p_1 = \tilde{m}_1 / \Delta X$	p_2	...	p_k

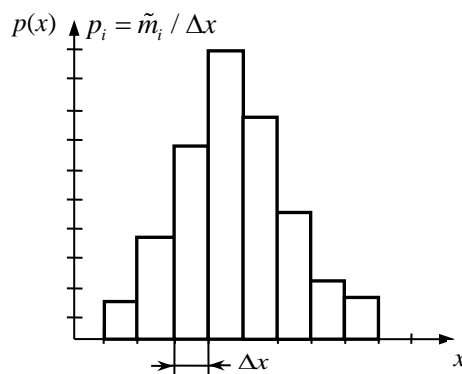


Рис. 12.3. Приклад емпіричної функції щільності розподілу ймовірностей

Площа під гістограмою дорівнює одиниці, оскільки дорівнює сумі площ прямокутників $S = \sum_{i=1}^k S_i = \sum_{i=1}^k \tilde{m}_i = 1$. Ламана лінія, яка проходить по осі абсцис, обводить гістограму і знову проходить по осі абсцис, є графіком **емпіричної функції щільності ймовірностей**.

За умови нескінченно великої кількості інтервалів, частота \tilde{m}_i є наближенням ймовірності потрапляння ВВ в i -й інтервал

$$\int_{x_i}^{x_{i+1}} p(x) dx \approx p_i. \quad (12.6)$$

Цю властивість використовують для порівняння теоретичного і емпіричного розподілів.

Вважатимемо емпіричну щільність ймовірностей неперервної ВВ постійною всередині кожного інтервалу, а емпіричну функцію розподілу на кожному інтервалі – як лінійно зростаючою від початкового до кінцевого інтервального значення.

12.2. Точкові оцінки характеристик ряду спостережень

Методологічною основою статистичного опрацювання результатів вимірювання є *вибірковий метод*, сутність якого полягає у перенесенні результатів дослідження вибірки на генеральну сукупність. Така можливість обґрунтовується наступним чином. Нехай генеральна сукупність має певний генеральний параметр Q . Проводиться серія з n незалежних вимірювань ВВ ξ з певним невідомим законом розподілу ймовірностей. Результатом вимірювання є вибірка $(x_1, \dots, x_i, \dots, x_n)$ обсягу n . Необхідно знайти оцінку генерального параметра Q

$$\hat{Q} = \varphi(x_1, \dots, x_i, \dots, x_n). \quad (12.6)$$

Тут і далі оцінки параметрів і характеристик позначатимемо символом $\hat{}$.

Параметром Q може бути, до прикладу, математичне сподівання, дисперсія, інші моменти ВВ вищих порядків, медіана та мода розподілів ВВ.

Вважається, що x_i отримують за незмінних умов, і всі результати вимірювання незалежні. Кожне з x_i , $j = \overline{1, n}$ є реалізацією генеральної сукупності ξ . Нагадаємо, що генеральна сукупність – це сукупність всіх можливих значень досліджуваної ВВ, тобто власне сама ВВ. Кожне значення x_i розглядається як реалізація ВВ ξ_i , а вся вибірка $(x_1, \dots, x_i, \dots, x_n)$ – як реалізація випадкового вектора $(\xi_1, \dots, \xi_i, \dots, \xi_n)$, де всі ξ_i , $j = \overline{1, n}$ незалежні і підпорядковані тому самому розподілу ймовірності, що й ξ . Такий перехід дає можливість вважати вибірковий параметр \hat{Q} функцією ВВ

$$\hat{Q} = \varphi(\xi_1, \dots, \xi_i, \dots, \xi_n), \quad (12.7)$$

що дає змогу не тільки знаходити аналітичні розв'язки (12.7), але й отримувати вирази для розподілів ймовірностей випадкових векторів $(\xi_1, \dots, \xi_i, \dots, \xi_n)$ та випадкових величин Q .

Означення. *Оцінювання* – це визначення наближеного значення невідомої характеристики або параметра розподілу (генеральній сукупності), іншою оцінюваною складовою математичної моделі реального (фізичного, технічного, економічного, тощо.) явища або процесу за результатами спостережень. Параметром генеральної сукупності у випадку вимірювань може бути число, набір чисел (вектор), або функція. Оцінювання проводять за допомогою оцінок – *статистик*.

Означення. *Точкове оцінювання* – спосіб оцінювання, що полягає в тому, що безпосередньо отримане значення оцінки приймають в якості невідомого значення параметра розподілу. Відповідно *точкова оцінка* параметра Q – це статистична оцінка, яка визначається одним числом \hat{Q} (12.6) за вибіркою $x_1, \dots, x_i, \dots, x_n$. Така функція вибірки власне і називається точковою оцінкою \hat{Q} , оскільки вона визначає одну точку на числовій осі. Наприклад, як оцінки положення центру розподілу використовуються вибіркове середнє, медіана, мода, середнє арифметичне максимального та мінімального елементів вибірки, геометричне середнє, гармонічне середнє, тощо.

Наявність декількох методів оцінювання одних і тих же параметрів призводить до необхідності вибору кращого між ними. Степінь відповідності Q його оцінці \hat{Q} залежить не тільки від обсягу n вибірки, але й від виду функції φ . Остання з вимоги найкращого наближення \hat{Q} до Q повинна задовольняти таким показникам якості як незсуненість, обґрунтованість, ефективність.

Незсуненість оцінки передбачає, що зі збільшенням кількості спостережень середнє значення оцінки прямує до значення оцінюваного параметра:

$$\lim_{n \rightarrow \infty} (M(\hat{Q})) = Q. \quad (12.8)$$

Якщо ця умова не виконується, оцінку називають *зсуненою*, а величину зсуву визначають як різницю $M(\hat{Q}) - Q$. Цей зсув може зумовлюватися як властивостями самої функції φ , так і похибками вимірювання, калібрування, невідповідним характером отриманої вибірки, чи комбінацією цих чинників.

Якщо зсув не залежить від обсягу вибірки він називається *систематичною похибкою*. Наявність систематичних похибок унеможлиблює оцінювання істинного значення вимірюваного параметра.

Оцінка називається **обґрунтованою** коли зі збільшенням обсягу вибірки вона збігається з відповідним параметром генеральної сукупності, тобто для будь якого як завгодно малого додатного числа ε має місце рівняння

$$\lim_{n \rightarrow \infty} P(|Q - \hat{Q}| \leq \varepsilon) = 1, \quad (12.9)$$

тобто зі зростанням n оцінки \hat{Q} все щільніше концентруються в околі Q .

Вимога **ефективності** оцінки полягає в наступному – для вибірок рівного обсягу ефективна оцінка повинна мати мінімальне розсіювання (вибіркову дисперсію). Якщо маємо $\mathbf{D}(\hat{Q}_1) \leq \mathbf{D}(\hat{Q}_2)$, то більш ефективною є оцінка \hat{Q}_1 .

Оцінкою математичного сподівання генеральної сукупності є вибіркове середнє

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (12.10)$$

Ця оцінка є незсуненою, обґрунтованою та ефективною.

Якщо вибіркові дані згруповані у варіаційний ряд, то оцінку математичного сподівання визначають за формулою

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k m_i x_i, \quad (12.11)$$

де x_i – значення варіанти для варіаційного ряду або середина класового інтервалу для інтервального варіаційного ряду; m_i – частота варіанти або класова частота.

Вибіркова медіана Me також може слугувати оцінкою математичного сподівання генеральної сукупності (особливо у випадку симетричного закону розподілу генеральної сукупності). Для дисперсій вибіркової медіани та вибіркового середнього справедливе співвідношення:

$$\lim_{n \rightarrow \infty} (\mathbf{D}(\hat{Me})) = \frac{\pi}{2} \mathbf{D}(\bar{x}). \quad (12.12)$$

З (12.12) випливає, що дисперсія оцінки медіани перевищує дисперсію вибіркового середнього \bar{x} в $\pi/2$ разів, тобто оцінка вибіркового середнього є більш ефективною.

Вибіркова дисперсія \hat{s}^2 є оцінкою дисперсії генеральної сукупності:

$$\hat{s}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mathbf{M}\xi)^2. \quad (12.13)$$

Ця оцінка є незсуненою, але якщо значення математичного сподівання $\mathbf{M}\xi$ невідоме, і використовується вибіркоче середнє, то отримана оцінка дисперсії

$$\hat{s}_*^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (12.14)$$

є зсуненою: $\hat{s}_*^2 = \frac{n-1}{n} \hat{s}^2$.

Для отримання незсуненої оцінки дисперсії використовують формулу

$$\hat{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (12.15)$$

У цій формулі знаменник $(n-1) = \nu$ називається *кількістю степенів свободи*. В загальному випадку значення ν дорівнює різниці між обсягом вибірки, з якої обчислено цю оцінку, мінус кількість констант, необхідних для обчислення оцінки і які визначені на підставі тих самих значень. Наприклад, у рівнянні (12.15) втрачено один степінь свободи під час визначенні \bar{x} на основі n спостережень.

У випадку, коли вибіркочі дані представлені інтервальним варіаційним рядом, для достатньо великого обсягу n і малої кількості класів, оцінка вибіркової дисперсії є завищеною (зсуненою) на величину $\Delta x^2 / 12$, де Δx – ширина класового інтервалу. Для розрахунку цієї оцінки вводять *поправку Шеннарда*:

$$\hat{s}'^2 = \hat{s}^2 - \Delta x^2 / 12. \quad (12.16)$$

Точкова оцінка стандартного відхилення вибіркового середнього \bar{x} в \sqrt{n} разів менша за оцінку стандартного відхилення одного результату спостереження

$$\hat{s}_{\bar{x}} = \frac{\hat{s}}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2}. \quad (12.17)$$

Отже статистика є характеристикою сукупності вимірювань, тоді як \hat{s} є характеристикою окремого вимірювання.

Лекція 13.

МЕТОДИ ВИЗНАЧЕННЯ ТОЧКОВИХ ТА ІНТЕРВАЛЬНИХ ОЦІНОК РЯДУ СПОСТЕРЕЖЕНЬ

Для знаходження точкових оцінок найчастіше використовують метод максимальної правдоподібності, метод моментів та метод квантилів.

13.1. Метод максимальної правдоподібності

Цей метод запропонований у 1912 році видатним англійським статистиком Р.А. Фішером (1890 –1962 р.). Метод базується на принципі, згідно якого найкращою оцінкою параметра Q є та оцінка, яка найбільш ймовірна за результатами проведення експерименту. В якості найбільш правдоподібного значення параметра Q береться та його оцінка \hat{Q} , для якої ймовірність отримати в n експериментах дану вибірку (x_1, x_2, \dots, x_n) є максимально великою. Кожна з величин x_i має щільність розподілу ймовірності $f(x_i, Q)$. Функція правдоподібності визначається наступним рівнянням:

$$L(x_1, x_2, \dots, x_n | Q) = f(x_1 | Q) \cdot f(x_2 | Q) \cdot \dots \cdot f(x_n | Q), \quad (13.1)$$

тобто ця функція дорівнює добутку умовних ймовірностей n випадкових величин, параметри розподілу яких відповідають максимуму ймовірності появи параметра Q . Вона має максимум при $\hat{Q} = Q$, де \hat{Q} є розв'язком рівняння

$$\frac{\partial L(x_1, x_2, \dots, x_n | Q)}{\partial Q} = 0, \quad (13.2)$$

або рівнозначного рівняння

$$\frac{\partial \ln L(x_1, x_2, \dots, x_n | Q)}{\partial Q} = 0. \quad (13.3)$$

Якщо потрібно оцінити не один, а декілька невідомих параметрів Q_1, Q_2, \dots, Q_k , то оцінки максимуму функцій правдоподібності для цих параметрів знаходять вирішуючи систему з k рівнянь:

$$\begin{cases} \frac{\partial \ln L(x_1, x_2, \dots, x_n | Q_1, \dots, Q_k)}{\partial Q_1} = 0, \\ \dots, \\ \frac{\partial \ln L(x_1, x_2, \dots, x_n | Q_1, \dots, Q_k)}{\partial Q_k} = 0. \end{cases} \quad (13.4)$$

Застосування методу максимальної правдоподібності розглянемо на прикладі оцінювання параметрів положення $\mathbf{M}\xi = \nu$ та розсіювання $+\sqrt{\mathbf{D}\xi} = \sigma$ ВВ, розподіленої за гауссовим законом, за вибіркою (x_1, x_2, \dots, x_n) .

Функція правдоподібності визначається рівнянням

$$L(x_1, x_2, \dots, x_n | \nu, \sigma) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left(-\frac{\sum_{i=1}^n (x_i - \nu)^2}{2\sigma^2}\right). \quad (13.5)$$

Якщо цю функцію домножити на $(2\pi)^{0,5n}$ і взяти її логарифм, отримаємо:

$$\ln L(x_1, x_2, \dots, x_n | \nu, \sigma) = -n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \nu)^2. \quad (13.6)$$

Диференціюємо отримане рівняння за параметрами ν та σ :

$$\begin{cases} \frac{\partial L(x_1, x_2, \dots, x_n | \nu, \sigma)}{\partial \nu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \nu) = 0, \\ \frac{\partial L(x_1, x_2, \dots, x_n | \nu, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \nu)^2 = 0. \end{cases} \quad (13.7)$$

Перше рівняння системи дорівнюватиме 0 у випадку, якщо

$$\sum_{i=1}^n (x_i - \nu) = \sum_{i=1}^n x_i - n\nu = 0, \quad (13.8)$$

звідки, з урахуванням $\hat{\nu} = \nu$ отримаємо

$$\hat{\nu} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}. \quad (13.9)$$

Отриману оцінку математичного сподівання підставимо у друге рівняння системи (13.7), матимемо

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (13.10)$$

Отримані оцінки (13.9) та (13.10) збігаються з оцінками математичного сподівання (12.10) та середньо квадратичного відхилення (12.14), отже ці оцінки є обґрунтованими та ефективними. Оцінка математичного сподівання \bar{x} є незсуненою і має гауссовий розподіл, а оцінка дисперсії $\hat{\sigma}^2 = \hat{s}^2$ зсунута і має розподіл χ^2 , який за умови $n \rightarrow \infty$ наближається до гауссового.

Отже, оцінки, отримані за методом максимальної правдоподібності є: 1) обґрунтованими; 2) асимптотично ефективними; 3) мають асимптотично гауссовий розподіл; 4) відповідають умові: якщо для параметра Q існує ефективна оцінка, то рівняння правдоподібності має єдиний розв'язок, який збігається з цією оцінкою.

13.2. Метод моментів

Цей метод розроблений відомим англійським математиком та статистиком К. Пірсоном (1857 – 1936) і є історично першим методом оцінювання параметрів. Він полягає у прирівнюванні певного числа вибірових моментів до відповідних моментів розподілу, функціонально зв'язаних з невідомими параметрами.

$$\begin{cases} m_1(Q_1, \dots, Q_k) = m_1^*(x_1, \dots, x_n), \\ \dots, \\ m_j(Q_1, \dots, Q_k) = m_j^*(x_1, \dots, x_n), \\ \dots, \\ m_k(Q_1, \dots, Q_k) = m_k^*(x_1, \dots, x_n). \end{cases} \quad (13.11)$$

де $m_j(Q)$ і $m_j^*(x_1, \dots, x_n)$ відповідно j -й момент розподілу (початковий чи центральний) з невідомими параметрами та j -й вибіркового моменту (початковий чи центральний).

За аналогією до моментів ВВ ξ (7.19) та (7.20) для вибірки (x_1, x_2, \dots, x_n) вводяться вибіркові моменти, які визначаються з цієї вибірки:

- початковий вибіркового моменту порядку j – це випадкова величина

$$v_j^* = \frac{1}{n} \sum_{i=1}^n x_i^j ; \quad (13.12)$$

- центральний вибіркового моменту порядку j – це випадкова величина

$$\mu_j^* = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^j . \quad (13.13)$$

Оцінки параметра Q_j розраховують шляхом розв'язання отриманих рівнянь (13.11) відносно цих параметрів.

Оцінки, отримані за методом моментів є асимптотично незміщеними та асимптотично нормальними, але не є асимптотично ефективними, тобто навіть за великих обсягів вибірки вони не мають найменшої можливої дисперсії.

13.3. Метод квантилів

Квантиль – це одна з числових характеристик ВВ. Розглянемо це поняття на прикладі неперервної ВВ ξ , яка має розподіл ймовірностей $F(x)$. Якщо зафіксовано дійсне число $\alpha \in [0, 1]$, α -квантилем (або квантилем рівня α) розподілу $F(x)$ називається число x_α таке, що

$$F(x_\alpha) = \alpha . \quad (13.14)$$

Графічну ілюстрація визначення поняття квантиля подано на рис. 13.1.

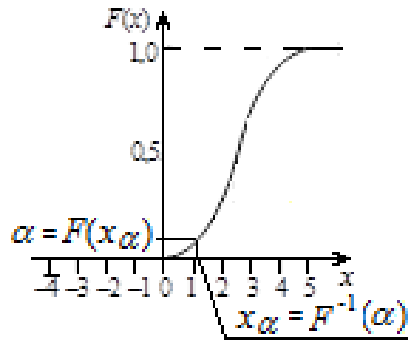


Рис. 13.1. Визначення квантиля рівня α

Очевидно, що для неперервних розподілів справедлива рівність

$$P\left(x_{\frac{1-\alpha}{2}} \leq \xi \leq x_{\frac{1+\alpha}{2}}\right) = \alpha. \quad (13.15)$$

В Таблиця 13.1 наведено квантилі стандартного гауссового розподілу, які часто вживаються в практиці статистичних розрахунків.

Таблиця 13.1. Найбільш вживані квантилів стандартного гауссового розподілу

Ймовірність, %	99,99	99,90	99,00	97,72	97,50	95,00	90,00	84,13	50,00
Квантиль	3,715	3,090	2,326	2,000	1,960	1,645	1,282	1,000	0,000

Вибіркові квантилі відсікають в межах варіаційного ряду певну частину його членів. Квантиль 0,25 (або 25-процентий квантиль або *нижній квантиль*) змінної – це таке значення $(x_{0,25})$, що 25 % значень змінної потрапляють нижче даного значення. Відповідно квантиль 0,75 називають *верхнім квантилем*.

Перейдемо тепер безпосередньо до методу квантилів. Цей метод подібний до методу моментів. При використанні цього методу теоретичні квантилі, як функції невідомих параметрів закону розподілу, прирівнюються до вибірових квантилів. У результаті отримуємо систему рівнянь

Довірча ймовірність визначається як величина $p = 1 - \alpha$, де α називається **рівнем значущості**, яка дорівнює ймовірності того, що надійний інтервал не накриває значення параметра.

Припустимо, що необхідно визначити довірчий інтервал для параметра Q генеральної сукупності, використовуючи незміщену та обґрунтовану вибірку оцінку \hat{Q} . Ця оцінка має вибіркового розподілу з середнім $M\hat{Q} = Q$ і стандартним відхиленням $\sigma_{\hat{Q}} \rightarrow 0$ за умови $n \rightarrow \infty$. Можна вказати ймовірність $P_d(|\hat{Q} - Q| \leq \varepsilon) = 1 - \alpha$, з якою відхилення оцінки від її математичного сподівання не перевищуватиме деякого додатного числа ε

$$P_d(\hat{Q} - \varepsilon \leq Q \leq \hat{Q} + \varepsilon) = 1 - \alpha, \quad (13.17)$$

де $\hat{Q} - \varepsilon$, $\hat{Q} + \varepsilon$ – нижня та верхня границі довірчого інтервалу. Ці границі є випадковими величинами, оскільки отримана за вибіркового спостереженнями оцінка \hat{Q} є випадковою величиною.

На рис.13.2 дано графічне представлення визначення довірчого інтервалу та довірчої ймовірності в задачі оцінювання математичного сподівання ВВ.

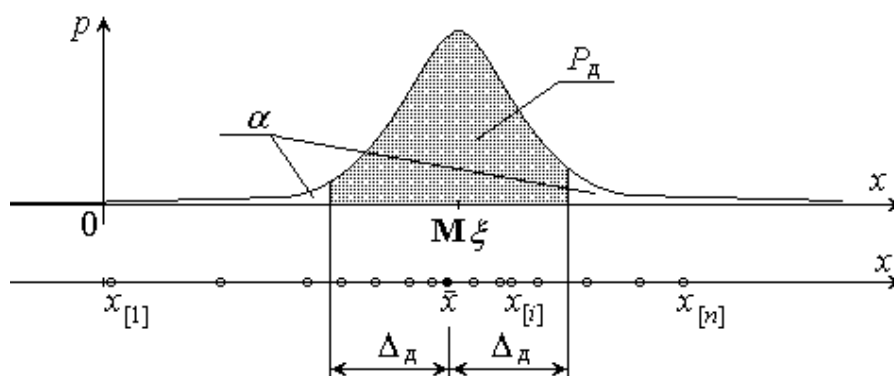


Рис. 13.2. Графічна ілюстрація процесу побудови точкової та інтервальної оцінок параметру розподілу ВВ за вибіркою $(x_1, \dots, x_i, \dots, x_n)$

13.5. Порядкові статистики ряду спостережень

Розглянуті вище методи точкового і інтервального оцінювання передбачають обґрунтування законів розподілу ВВ, для яких визначаються їх параметри.

Зазвичай припускають, що результати спостережень формуються за сукупним впливом багатьох малих факторів, і в силу центральної граничної теореми теорії ймовірності ця величина добре наближається (за розподілом) до гауссової ВВ. Таке твердження є справедливим, якщо фактори, що визначають похибки діють адитивно і незалежно один від одного. На припущенні про гауссовий закон розподілу результатів спостережень побудовано більшість класичних моделей і методів математичної статистики.

Проте таке теоретичне припущення не завжди можна підтвердити експериментальними дослідженнями, що обмежує можливість використання гауссового розподілу і, відповідно, параметричних методів статистики.

Альтернативою параметричним методом статистики є так званий непараметричний підхід, за яким наближення функції розподілу знаходиться безпосередньо з отриманих вибіркового даних. Переваги непараметричних методів полягають у тому, що їх можна застосовувати:

- 1) для перевірки гіпотез про параметри генеральної сукупності, коли випадкові величини не розподілені за гауссовим законом;
- 2) для номінальних і порядкових даних;
- 3) для перевірки гіпотез, які не пов'язані з параметрами генеральної сукупності;

Крім того у більшості випадків для непараметричних методів обчислення простіші, ніж для параметричних, а самі методи зрозуміліші.

Недоліки непараметричних методів полягають у тому, що вони менш точні, чим відповідні параметричні критерії, менш інформативні та менш ефективні.

Одним з типів методів, які входять в групу непараметричних є рангові, в яких використовуються не самі вибіркові значення, а їх ранги. Такі методи дають змогу оцінити властивості ВВ незалежно від їх функцій розподілу.

Рангом числа x_i у вибірці $(x_1, \dots, x_i, \dots, x_n)$ називають той номер r_i , який воно отримує при упорядкуванні всієї групи за зростанням (спаданням). Так, ранг $r_1 = 1$ отримує найменше (найбільше) число у вибірці, ранг 2 – найменше (найбільше) із залишившихся і т.д. Ранг $r_n = n$ отримає найбільше (найменше) з чисел у вибірці.

Якщо допускається, що результати спостережень $(x_1, \dots, x_i, \dots, x_n)$, розподілені неперервно, то рівність серед чисел $(x_1, \dots, x_i, \dots, x_n)$ теоретично має нульову ймовірність, але на практиці величини x_i вимірюються з обмеженою точністю і тому їх рівність є можливою. Якщо серед чисел $(x_1, \dots, x_i, \dots, x_n)$ є однакові, вони отримують спільний *середній ранг*; якщо однакові величини займають третє, четверте та п'яте місця у вибірці, кожне з них отримує середній ранг $r_3 = r_4 = r_5 = (3+4+5)/3$. Такі ранги називають “зв’язаними”. Правильність визначення рангів здійснюється за перевіркою виконання рівності:
$$\sum_i r_i = \frac{n(n+1)}{2}.$$

Вибірка, елементи якої розміщені на місцях, номери яких відповідають їх рангам називаються **ранжованою** або **варіаційним рядом**. Елементи варіаційного ряду називаються **порядковими статистиками**.

При заміні чисельних значень спостережень їх рангами неминуче виникає втрата інформації. Іноді така втрата може бути повною. Нехай, наприклад, $(x_1, \dots, x_i, \dots, x_n)$, результати незалежних вимірювань невідомої величини А. Якщо замінити дану вибірку послідовністю рангів отримуємо деяку перестановку чисел $1, 2, \dots, n$. Ця перестановка не містить інформації про значення величини А, яке може мати розмірність, а ранги – безрозмірні.

Попри цей недолік, методи рангової статистики дають змогу вирішувати всі ті задачі оброблення експериментальних даних, що й параметричні методи, зокрема ті, що ґрунтуються на припущенні про гауссовість, тому їх широко застосовують для розв’язання різних статистичних задач.

Лекція 14.

МЕТОДИ ОЦІНЮВАННЯ ПАРАМЕТРІВ ЗАКОНУ РОЗПОДІЛУ ДАНИХ

14.1. Оцінювання параметрів за відомого закону розподілу

Загальна постановка задачі наступна. Відомим є закон розподілу ймовірностей ВВ $\xi(\omega)$, $\omega \in \Omega$ (його аналітичний запис), але невідомі параметри розподілу. За отриманої вибірки $(x_1, \dots, x_i, \dots, x_n)$ необхідно визначити точкові та інтервальні оцінки (довірчий інтервал $\pm \Delta x_d$ та довірчу ймовірність P_d) центру розподілу. Отже оцінюється параметр $M\xi$. Довірчий інтервал будується відносно оцінки \bar{x} .

14.1.1. Оцінювання параметрів розподілу Гаусса

Нагадаємо, що щільність розподілу Гаусса (диференціальна функція розподілу) (рис. 14.1a) та функція розподілу (інтегральна) (рис. 14.1б) визначаються відповідно формулами:

$$p(x; \mu; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (14.1)$$

$$F(x; \mu; \sigma) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt, \quad (14.2)$$

де μ – параметр положення (зсуву), σ – параметр розсіювання.

Для розв'язання практичних задач часто застосовується функція нормованого або стандартного розподілу Гаусса ($\mu = 0$, $\sigma = 1$):

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt. \quad (14.3)$$

Розглянемо спочатку **точкові оцінки** параметра μ , які отримують за рядом спостережень x_i , $i = \overline{1, n}$

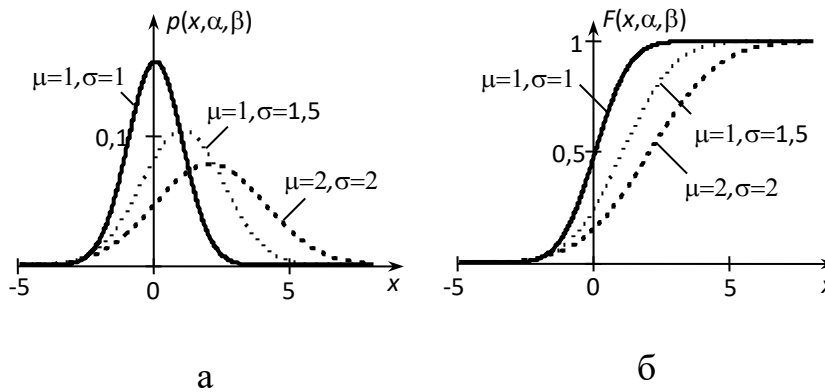


Рис. 14.1. Диференціальна (а) та інтегральна (б) функції розподілу імовірності Гаусса

Для розв'язання практичних задач часто застосовується функція нормованого або стандартного розподілу Гаусса ($\mu = 0, \sigma = 1$):

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt. \quad (14.3)$$

Розглянемо спочатку **точкові оцінки** параметра μ , які отримують за рядом спостережень $x_i, i = \overline{1, n}$

1. *Оцінка найбільшої правдоподібності* параметра μ визначається як

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (14.4)$$

Оцінка є обґрунтованою, незміщеною, ефективною. Оцінка розподілена так само як ВВ з математичним сподіванням $\mathbf{M}(\hat{\mu}) = \mu$ та дисперсією $\mathbf{D}(\hat{\mu}) = \sigma^2 / n$.

2. *Оцінка за значенням медіани.* У якості оцінки параметра μ можна використовувати вибірккову медіану:

$$\hat{\mu} = \mathbf{Med}(x), \quad (14.5)$$

де $\hat{\mu} = \mathbf{Med}(x)$ – оператор вибіркової медіани.

В статистиці *медіана* – це величина ознаки, що розташована посередині ранжованого ряду, тобто це величина, що розташована в середині ряду величин, переставлених у зростаючому або спадному порядку (нагадаємо, що в теорії ймовірностей *медіаною функції розподілу* $F(x)$ називається таке число x_{med} , для

якого $F(x_{med}) = 0,5$). Медіана ділить ряд значень ознаки на дві рівні частини, по обидві сторони від неї розміщується однакова кількість одиниць сукупності. Медіана є квантилем порядку $1/2$ і обчислюється як

$$\mathbf{Med}(x) = \begin{cases} x_{[(n+1)/2]}, & \text{якщо } n \text{ непарне,} \\ 0,5(x_{[n/2]} + x_{[(n+1)/2]}), & \text{якщо } n \text{ парне.} \end{cases} \quad (14.6)$$

Ефективність оцінки при $n \rightarrow \infty$ прямує до $2/\pi = 0,637$, тобто для досягнення оцінкою ефективності оцінки максимальної правдоподібності необхідний в $\pi/2$ більший обсяг спостережень.

3. Оцінювання за порядковими статистиками (квантилями). Діксон запропонував прості оцінки μ :

– середнє з двох *оптимальних* порядкових статистик:

$$\hat{\mu} = 0,5 \cdot (x_{[i]} + x_{[j]}). \quad (14.7)$$

– середнє за варіаційним рядом з усіх спостережень крім двох крайніх:

$$\hat{\mu} = \frac{1}{n-2} \sum_{i=2}^{n-1} x_{[i]}. \quad (14.8)$$

Оптимальні порядкові статистики для оцінки (14.7) наведені у Таблиці 14.1. Якщо $n > 20$, можна скористатись залежністю $i=[0,27n]$, $j=[0,73n]$, де $[\cdot]$ – оператор заокруглення до найближчого цілого.

Ефективність оцінки (14.6) прямує до 0,81 при $n \rightarrow \infty$, а ефективність оцінки (14.7) не поступається оцінці (14.4).

Таблиця 14.1. Оптимальні порядкові статистики для оцінки Діксона

<i>n</i>	2	3	4	5	6	7	8	9	10	11
<i>i</i>	1	1	2	2	2	2	3	3	3	3
<i>j</i>	2	3	3	4	5	6	6	7	8	9
<i>n</i>	12	13	14	15	16	17	18	19	20	∞
<i>i</i>	4	4	4	4	5	5	5	6	6	$0,27n$
<i>j</i>	9	10	11	12	12	13	14	14	15	$0,73n$

Ефективність оцінки (14.6) прямує до 0,81 при $n \rightarrow \infty$, а ефективність оцінки (14.7) не поступається оцінці (14.4).

4. Швидкі оцінки Кенуя. До таких оцінок належать наступні.

– Середньоквартильна оцінка:

$$\hat{\mu} = 0,5 \cdot (x_{[0,25n]} + x_{[0,75n]}). \quad (14.9)$$

Ефективність оцінки $E \approx 1,21/n$. Метод рекомендований лише для швидкого наближеного оцінювання;

– швидка оцінка за трьома квантилями:

$$\hat{\mu} = 0,2x_{[n/16]} + 0,6x_{[n/2]} + 0,2x_{[15n/16]}. \quad (14.10)$$

Ефективність оцінки $\approx 0,83$. Метод достатньо стійкий до відхилення розподілу спостережень від розподілу Гаусса.

– Оцінка за п'ятьма квантилями:

$$\hat{\mu} = \frac{1}{6}(x_{[n/16]} + x_{[n/4]} + 2x_{[n/2]} + x_{[3n/4]} + x_{[15n/16]}). \quad (14.11)$$

Ефективність оцінки $\approx 0,93$. Метод стійкий до відхилення розподілу спостережень від розподілу Гаусса.

5. Стійка оцінка Ходжеса-Лемана за середніми Уолша. Ця оцінка застосовується у випадку відхилення розподілу спостережень від розподілу Гаусса та наявності аномальних значень. За рядом спостережень $x_i, i = \overline{1, n}$ визначають множину $(n(n+1)/2)$ середніх Уолша:

$$z_{i,j} = 0,5(x_i + x_j), \quad i \leq j. \quad (14.12)$$

Оцінка Ходжеса-Лемана визначається як медіана отриманої множини середніх Уолша: $\hat{\mu} = \mathbf{Med}(z)$.

Перейдемо до розгляду **інтервальних оцінок** параметра μ

6. Інтервальна оцінка μ за відомого значення σ

Для визначення меж довірчого інтервалу користуються γ -квантилями стандартного розподілу Гаусса, наведеними в Таблиці 14.2.

Нижню і верхню межі інтервальної оцінки параметра μ (14.4), (14.8) для довірчої ймовірності P розраховують за формулами:

$$\mu_{\text{н}}(n, P_{\text{д}}) = \hat{\mu} - u_{\gamma} \sigma / \sqrt{n}, \quad \mu_{\text{в}}(n, P_{\text{д}}) = \hat{\mu} + u_{\gamma} \sigma / \sqrt{n}, \quad (14.13)$$

де u_{γ} – γ -квантиль стандартного розподілу Гаусса (Таблиця 14.2), $\gamma = (1 + P) / 2$ – для двосторонньої оцінки і $\gamma = P$ для односторонньої оцінки.

Таблиця 14.2. Значення квантилів стандартного розподілу Гаусса

γ	u_{γ}	γ	u_{γ}	γ	u_{γ}	γ	u_{γ}
0,50	0,00000	0,76	0,70630	0,950000	1,64485	0,999200	3,15591
0,52	0,05015	0,78	0,77219	0,960000	1,75069	0,999400	3,23888
0,54	0,10043	0,80	0,84162	0,970000	1,88079	0,999500	3,29053
0,56	0,15097	0,81	0,87790	0,975000	1,95996	0,999600	3,35279
0,58	0,20189	0,8200	0,91536	0,980000	2,05375	0,999700	3,43161
0,60	0,25335	0,8400	0,99446	0,990000	2,32635	0,999800	3,54008
0,62	0,30548	0,8600	1,08032	0,992000	2,40891	0,999900	3,71902
0,64	0,35846	0,8800	1,17499	0,994000	2,51214	0,999950	3,89059
0,66	0,41246	0,9000	1,28155	0,995000	2,57583	0,999990	4,26489
0,68	0,46770	0,9100	1,34075	0,996000	2,65207	0,999995	4,41717
0,70	0,52440	0,9200	1,40507	0,997000	2,74778	0,999999	4,75342
0,72	0,58284	0,9300	1,47579	0,998000	2,87816		
0,74	0,64334	0,9400	1,55477	0,999000	3,09023		

7. *Інтервальна оцінка μ за невідомого значення σ .* Нижня і верхня границі інтервальної оцінки параметра μ (14.4), (14.8) з довірчою ймовірністю P розраховують за формулами:

$$\mu_{\text{н}}(n, P_{\text{д}}) = \hat{\mu} - t_{\gamma}(v) \frac{s}{\sqrt{n}}, \quad \mu_{\text{в}}(n, P_{\text{д}}) = \hat{\mu} + t_{\gamma}(v) \frac{s}{\sqrt{n}}, \quad (14.14)$$

де $t_{\gamma}(v)$ – γ -квантиль розподілу Стюдента з $\nu = n - 1$ степенями **свободи** (таблиця Додатка 3), s – оцінка СКВ, $\gamma = (1 + P) / 2$ – для двосторонньої оцінки і $\gamma = P$ для односторонньої оцінки.

Для практичного використання часто обирають $P = 0,95$, тоді можна застосовувати апроксимацію $\gamma = \frac{1 + P}{2} = 0,975$:

$$t_{0,975}(v) = 1,96 + \frac{2,5}{v - 0,8}. \quad (14.15)$$

Перейдемо до розгляду **точкових оцінок** параметра σ , які отримують за рядом спостережень x_i , $i = \overline{1, n}$.

1. Оцінка найбільшої правдоподібності параметра σ :

$$\hat{\sigma} = \sqrt{s^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2}. \quad (14.16)$$

2. Оцінка σ за вибіркоvim розмахом R . За обсягу спостережень $n \leq 20$ рекомендується застосовувати наступну формулу для спрощеного розрахунку:

$$\hat{\sigma} = R \cdot d(n), \quad (14.17)$$

де $R = x_{[n]} - x_{[1]}$, $d(n)$ – параметр який пов’язує значення СКВ та розмах для розподілу Гаусса (Таблиця 14.3).

Таблиця 14.3. Коефіцієнт d для визначення σ за вибіркоvim розмахом

n	d	n	d	n	d	n	d
2	0,8862	7	0,3698	12	0,3069	17	0,2787
3	0,5908	8	0,3562	13	0,2998	18	0,2747
4	0,4857	9	0,3367	14	0,2935	19	0,2711
5	0,4299	10	0,3249	15	0,2880	20	0,2677
6	0,3946	11	0,3152	16	0,2831		

3. Оцінка σ Даутона за порядковими статистиками. Простою з огляду на обчислювальну складність та достатньо ефективна ($E \approx 0,94$) є така оцінка СКВ:

$$\hat{\sigma} = \frac{1,77245}{n(n-1)} \sum_{i=1}^n (x_{[i]} \cdot (2 \cdot i - n - 1)). \quad (14.18)$$

4. Інтервальні оцінки параметра σ . Методи визначення інтервальних оцінок параметра σ залежать від методів визначення точкових оцінок. Так нижню і верхню межі інтервальної оцінки параметра σ (14.16) при довірчій ймовірності P_d визначають за формулами:

$$\sigma_{\text{н}}(n, P_d) = \sqrt{\frac{1}{\chi_{\gamma'}^2} \sum_{i=1}^n (x_i - \hat{\mu})^2}, \quad \sigma_{\text{в}}(n, P_d) = \sqrt{\frac{1}{\chi_{\gamma''}^2} \sum_{i=1}^n (x_i - \hat{\mu})^2}. \quad (14.19)$$

де $\hat{\mu}$ – вибіркова оцінка параметра μ , χ^2_{γ} – γ -квантиль розподілу Пірсона з $\nu = n - 1$ степенями свободи (якщо значення $\hat{\mu}$ відоме, то $\nu = n$) (таблиця Додатка 4), $\gamma' = (1 + P_d) / 2$, $\gamma'' = (1 - P_d) / 2$ для двосторонньої оцінки і $\gamma' = P_d$, $\gamma'' = 1 - P_d$ – для односторонньої.

14.1.2. Оцінювання параметрів рівномірного розподілу

Щільність і функція **рівномірного** розподілу дається формулами (9.8) та (9.9). Їх графіки подані відповідно на рис. 9.2а та рис. 9.2б. Розглянемо точкові оцінки параметрів рівномірного закону розподілу. За достатнього обсягу вибірки $n > 50$ для оцінювання параметрів a та b запропоновано прості оцінки:

$$a = x_{[1]}, b = x_{[n]}. \quad (14.20)$$

Ці оцінки не є стійкими до наявності спостережень з надмірною похибкою та значною мірою залежать від обсягу даних.

За обсягу даних $n \leq 50$ більш доцільним є застосування наступних оцінок:

$$a = \bar{x} - s\sqrt{3}, b = \bar{x} + s\sqrt{3}, \quad (14.21)$$

де \bar{x} – оцінка математичного сподівання, отримана за методом максимальної правдоподібності, s – оцінка СКВ.

Для забезпечення стійкості до наявності у ряді спостережень даних з надмірною похибкою у виразі (14.21) замість оцінок \bar{x} та s застосовують відповідні усічені оцінки:

$$\bar{x} = \mathbf{Med}(x), s = 1,15 \cdot \mathbf{Med}(\Delta x), \quad (14.22)$$

де Δx – множина модулів різниць між результатами спостережень та медіани ряду спостережень $\Delta x_i = |x_i - \mathbf{Med}(x)|$.

14.1.3. Оцінювання параметрів трикутного розподілу

Щільність і функція **трикутного** розподілу дається формулами (9.10) та (9.11). Їх графіки подані відповідно на рис. 9.3а та рис. 9.3б. За достатнього обсягу даних

$n > 50$ для оцінювання параметрів a та b беруть прості оцінки (14.20): $a = x_{[1]}$, $b = x_{[n]}$. Так само як і для рівномірного розподілу ці оцінки не є стійкими до наявності спостережень з надмірною похибкою та значною мірою залежать від обсягу даних.

За обсягу даних $n \leq 50$ більш доцільним є застосування наступних оцінок:

$$a = \bar{x} - s\sqrt{6}, \quad b = \bar{x} + s\sqrt{6}. \quad (14.23)$$

Для забезпечення стійкості до наявності у ряді спостережень даних з надмірною похибкою у виразі (14.23) замість оцінок \bar{x} та s застосовують відповідні стійкі або усічені оцінки, наприклад:

$$\bar{x} = \mathbf{Med}(x), \quad s = 1,75 \cdot \mathbf{Med}(\Delta x). \quad (14.24)$$

14.2. Оцінювання за невідомого закону розподілу ймовірностей

14.2.1. Оцінка для центру розподілу

Як первинні (грубі) оцінки центру групування значень ВВ за невідомого закону розподілу ймовірностей можуть бути використані різні граничні нерівності, які дають змогу отримати інтервальні оцінки центру розподілу.

1. *Нерівність Чебишова.* Виконаємо підстановку $\varepsilon = k\sigma$ у нерівність Чебишова (8.9)

$$P(|x - \mu| \geq k\sigma) < 1/k^2, \quad (14.25)$$

де μ і σ – відповідно математичне сподівання та стандартне відхилення розподілу імовірності.

З нерівності (14.25) з урахуванням того, що $P(|x - \mu| \geq k\sigma) = 1 - P_d$ впливає

$$\xi - \frac{\sigma}{\sqrt{1 - P_d}} \leq \mu \leq \xi + \frac{\sigma}{\sqrt{1 - P_d}}, \quad (14.26)$$

де P_d – довірча ймовірність.

Якщо замість значення випадкової величини ξ використовується вибіркове середнє $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, то виконується нерівність

$$\bar{x} - \frac{\sigma}{\sqrt{n(1-P_d)}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n(1-P_d)}} . \quad (14.27)$$

Якщо відомо, що розподіл симетричний відносно центра μ , то відповідний довірчий інтервал визначається як:

$$\bar{x} - \frac{2\sigma}{3\sqrt{n(1-P_d)}} \leq \mu \leq \bar{x} + \frac{2\sigma}{3\sqrt{n(1-P_d)}} . \quad (14.28)$$

Отже тільки знання того факту, що розподіл ВВ симетричний, дає змогу побудувати більш вузький довірчий інтервал для центру розподілу.

Лекція 15.

МЕТОДИ ВИЗНАЧЕННЯ ЗАКОНА РОЗПОДІЛУ ДАНИХ

Вибір методу та правильне оцінювання закону розподілу даних є важливим етапом статистичного аналізу експериментальних даних, оскільки від виду закону залежить коректність визначення точкових та інтервальних оцінок параметрів розподілу та інших статистичних характеристик.

Визначення законів розподілу ймовірності полягає у висуненні та перевірці гіпотез про закони розподілу статистичних даних.

15.1. Загальні поняття про статистичні гіпотези та їх перевірку

Нехай ξ – досліджувана неперервна ВВ зі щільністю ймовірності $p_\xi(x)$, θ – невідомий інформативний параметр, $X = (x_1, \dots, x_n)$ – реалізація обсягу N випадкової вибірки $\xi = (\xi_1, \dots, \xi_n)$, $\hat{\theta}$ – оцінка параметра θ .

Означення. *Статистичною перевіркою гіпотез* називають перевірку будь-якого припущення щодо ймовірнісних характеристик ВВ, які спостерігаються в експерименті. Припущення в статистиці називається *гіпотезою*, позначається літерою H і може стосуватись вигляду розподілу досліджуваної ВВ, її параметрів, числових характеристик, наявності/відсутності дефектів ОК.

Означення. Нульовою гіпотезою H_0 називається основна (висунута) гіпотеза, яка полягає в порівнянні параметра θ з його оцінкою $\hat{\theta}$. Нульова гіпотеза зазвичай стверджує, що відмінність між порівнюваними величинами відсутня, а наявні відхилення обумовлені випадковим характером оцінки $\hat{\theta}$.

Означення. Альтернативною (конкуруючою) називається гіпотеза H_1 , яка приймається в тому випадку, коли нульова гіпотеза H_0 відхиляється. Альтернативних гіпотез може бути одна або більше, наприклад,

$$H_0: \theta = \theta_0; H_1: \theta = \theta_1 \text{ або } \theta \neq \theta_0.$$

Вибір альтернативної гіпотези визначається постановкою задачі.

Означення. Гіпотеза називається *простою*, якщо вона припускає, що параметр θ приймає тільки одне значення, наприклад, $H_0: \theta = \theta_0; H_1: \theta = \theta_1$.

Означення. Гіпотеза називається *складною*, якщо вона припускає, що параметр θ приймає скінченну чи зліченну кількість значень, наприклад, $H_0: \theta = \theta_0; H_1: \theta \neq \theta_0$ або $H_1: \theta > \theta_0$. Найчастіше перевіряють прості гіпотези.

Означення. *Параметричною* називається гіпотеза, якщо закон розподілу ВВ ξ передбачається відомим і перевіряються припущення щодо значень його параметрів.

Означення. *Непараметричною* називається гіпотеза, якщо закон розподілу ВВ ξ невідомий, а висновки, зроблені на основі непараметричної гіпотези, не залежать від невідомого теоретичного розподілу ВВ ξ чи його параметрів.

Прийняття рішення на підставі статистичної перевірки гіпотез завжди супроводжується помилками двох видів.

Означення. *Помилкою першого роду* називається помилка відхилення вірної нульової гіпотези. Імовірність помилки першого роду позначається $P(H_1|H_0) = \alpha$.

Означення. *Помилкою другого роду* називається помилка неправильного прийняття нульової гіпотези. Імовірність помилки другого роду позначається $P(H_0|H_1) = \beta$. Якщо необмежено збільшувати обсягу вибірки N , є принципова можливість одержання як завгодно малих помилок α і β .

Означення. *Статистичним критерієм* називається правило, згідно з яким гіпотеза H_0 приймається або відхиляється. Статистичний критерій встановлює, для яких результатів випадкової вибірки гіпотеза H_0 приймається, а для яких – відхиляється. Для побудови статистичного критерію використовується *статистика критерію* $T(\xi_1, \dots, \xi_n)$ – функція від вектора спостережень, за значеннями якої приймають або відхиляють гіпотезу H_0 . Зокрема, статистикою критерію є оцінка $\hat{\theta} = \theta(\xi_1, \dots, \xi_n)$ параметру θ . Значення статистики критерію, за яких гіпотеза H_0 відхиляється, утворюють *критичну область* гіпотези, що перевіряється.

В процесі побудови критерію прагнуть зменшити до мінімуму ймовірності помилок першого і другого роду. Характеристиками якості статистичного критерію перевірки гіпотез служать рівень значущості і потужність статистичного критерію.

Означення. *Рівень значущості α* – це ймовірність відхилення вірної гіпотези H_0 , тобто ймовірність помилки першого роду. Значення α зазвичай вибирається з ряду 0,005; 0,01; 0,05; 0,1; в технічних застосунках, як правило, $\alpha = 0,05$.

Означення. *Потужністю критерія* називається ймовірність $M = 1 - \beta$ недопущення помилки другого роду. Із двох критеріїв, що характеризуються однією і тією ж імовірністю α , обирають той, що має більшу потужність (меншу помилку β).

Для статистичної перевірки гіпотез необхідно виконати такі дії.

1. Сформулювати гіпотези H_0 і H_1 .
2. Визначити статистику критерію – оцінку $\hat{\theta} = \theta(\xi_1, \dots, \xi_n)$ параметру θ .
3. Задати рівень значущості статистичного критерію α і визначити критичну область для гіпотези H_0 , що перевіряється.
4. Одержати реалізацію (x_1, \dots, x_n) випадкової вибірки $\xi = (\xi_1, \dots, \xi_n)$.
5. Підставити у функцію $\hat{\theta} = \theta(\xi_1, \dots, \xi_n)$ реалізацію випадкової вибірки.
6. Якщо значення $\hat{\theta}$ потрапляють у критичну область гіпотеза H_0 відхиляється.

15.2 Попереднє оцінювання закону розподілу даних за діаграмою β_1, β_2

На рис. 15.1 показані області в площині (β_1, β_2) для різних розподілів – розподілу Гаусса, бета-розподілу (окремий випадок – експоненціальний розподіл) і логарифмічно нормального, де β_1 – корінь квадратний з нормованого показника асиметрії, а β_2 – нормований показник ексцесу.

$$\beta_1 = \sqrt{\frac{\mu_3}{(\mu_2)^{3/2}}}; \quad \beta_2 = \frac{\mu_4}{\mu_2^2}, \quad (15.1)$$

де μ_j – центральні моменти ВВ j -того порядку. Сюди входить і t -розподіл Стюдента – симетричний розподіл, що сходиться до нормального, коли його параметр (число ступенів свободи) довільно збільшується.

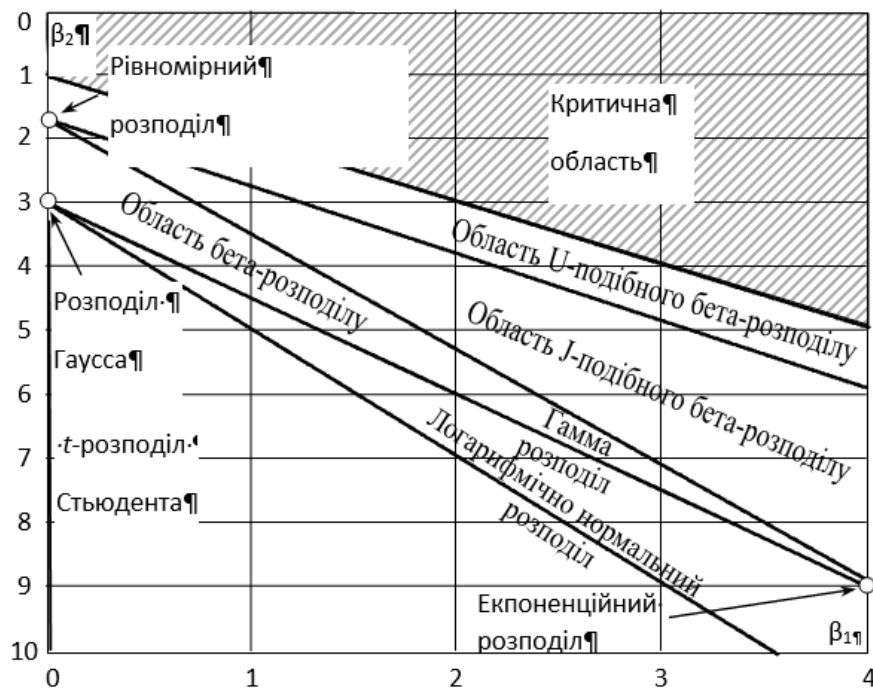


Рис. 15.1. Області у площині (β_1, β_2) для різних розподілів

Для гауссового розподілу $\beta_1 = 0$ і $\beta_2 = 3$. Тому на рис. 15.1 цей розподіл представлений однією точкою, так само як експоненціальний ($\beta_1 = 4,0$ і $\beta_2 = 9,0$) і рівномірний ($\beta_1 = 0$ і $\beta_2 = 1,8$) розподіли. Це пояснюється тим, що у цих розподілів немає параметра форми і тому вони мають єдину форму. Гамма-

розподіл, логарифмічно нормальний розподіл і t -розподіл Стьюдента представлені кривими. Таким чином, гамма- розподіл можна підібрати для усіх значень β_1 і β_2 , що лежать поблизу середньої кривої.

Велика область значень β_1 і β_2 не охоплена жодним з розглянутих раніше розподілів. Такі розподіли можуть бути описані спеціальними сім'ями розподілів, наприклад Джонсона або Пірсона.

Для застосування графіків з рис. 15.1 необхідно замість невідомих значення β_1 і β_2 використати їх оцінки $\hat{\beta}_1$ і $\hat{\beta}_2$, що визначаються з отриманої вибірки

$$\hat{\beta}_1 = \sqrt{\frac{\hat{\mu}_3}{(\hat{\mu}_2)^{3/2}}}; \quad \hat{\beta}_2 = \frac{\hat{\mu}_4}{\hat{\mu}_2^2}, \quad (15.2)$$

де $\hat{\mu}_j$ – оцінки центральних моментів ВВ j -того порядку.

Оцінки $\hat{\mu}_j$ розраховуються за вибірками ВВ $(x_1, \dots, x_i, \dots, x_n)$ обсягу n

$$\hat{\mu}_j = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^j. \quad (15.3)$$

Точка з координатами $(\hat{\beta}_1, \hat{\beta}_2)$ наноситься на рис. 15.1. Якщо ця точка лежатиме досить близько від точки, кривої або області, що відповідає одному з названих розподілів, то останній використовують для опису емпіричних даних та висування гіпотези про вид розподілу, значущість якого кількісно оцінюють за допомогою різних критеріїв. Після підтвердження гіпотези можна приступити до знаходження оцінок параметрів розподілу.

Застосовуючи цей метод необхідно враховувати важливе обмеження: для будь-якої множини даних обчислені значення $\hat{\beta}_1$ і $\hat{\beta}_2$ є лише оцінками для β_1 і β_2 , що схильні до коливань від вибірки до вибірки.

15.3. Методи оцінювання закону розподілу з групуванням даних

Розглянуті тут методи застосовують загальні критерії згоди, які полягають у порівнянні вибіркових групових оцінок емпіричних даних (гістограми, полігону

частот тощо) з теоретичними, заздалегідь відомими функціями розподілу. Всі відомі загальні критерії згоди можна поділити на три групи:

- критерії, що ґрунтуються на вивченні різниці між теоретичною щільністю розподілу і емпіричною гистограмою;
- критерії, що ґрунтуються на відстані між теоретичною і емпіричною функціями розподілу імовірності;
- кореляційно-регресійні критерії, які засновані на дослідженні кореляційних та регресійних зв'язків між емпіричними та теоретичними порядковими статистиками.

15.3.1. Аналіз варіаційних рядів

В лекції 12 було дано початкове уявлення про такі поняття як полігон частот, гистограма та емпірична функція розподілу. Тут уточнимо ці поняття та наведемо методику побудови відповідних графіків.

Означення. *Полігоном* частот для неперервного розподілу називають ламану лінію, відрізки якої з'єднують точки $(\bar{x}_1, \hat{m}_1), (\bar{x}_2, \hat{m}_2), \dots, (\bar{x}_j, \hat{m}_j)$, де \bar{x}_j – середини інтервалів, на які розбито область значень ВВ ξ , \hat{m}_j – відповідні частоти інтервалів, $j = \overline{1, l}$, l – кількість інтервалів. Якщо немає іншої інформації, областю значень ξ є інтервал від мінімального до максимального значення ряду спостережень. Частоти інтервалів визначають як кількість значень ряду, що потрапляють у межі j -го інтервалу.

Означення. *Полігон відносних частот* – це ламана лінія, відрізки якої з'єднують точки $(\bar{x}_1, \hat{m}_1 / n), (\bar{x}_2, \hat{m}_2 / n), \dots, (\bar{x}_j, \hat{m}_j / n)$, де n – обсяг вибірки.

Означення. *Гистограма* – ступінчаста фігура, яка складається з прямокутників, основи яких відповідають довжині інтервалів h , а висоти рівні відношенню $\hat{m}_j / (n \cdot h)$. Площа гистограми завжди дорівнює одиниці.

Означення. *Емпірична функція розподілу* (функція розподілу вибірки) – функція $F(x)$, яка визначає для кожного значення x відносну частоту події $\xi < x$:

$$F(x) = n_x / n, \quad (15.4)$$

де n_x – кількість варіант, менших за x , n – обсяг вибірки.

Емпірична функція розподілу має наступні властивості:

1. Область значень функції $[0,1]$.
2. Функція $F(x)$ – неспадна.
3. Якщо x_1 – найменша варіанта, x_k – найбільша, то $F(x) = 0$, якщо $x < x_1$ та $F(x) = 1$, якщо $x > x_k$.

Методика побудови полігону частот та гістограми

1. Визначається мінімальне x_{\min} та максимальне x_{\max} вибірккові значення.
2. Експериментальні дані розбиваються на l інтервалів. У випадку $n < 100$ кількість інтервалів визначають як найближче ціле від кореня квадратного з обсягу ряду спостережень n ($l = [\sqrt{n}]$). Якщо обсяг даних $n > 100$, то за формулою *Старджесса* отримують $l = 1 + 3,31 \lg(n)$. Рекомендується обирати непарне значення l .
3. Визначається ширина інтервалів гістограми $h = (x_{\max} - x_{\min}) / l$.
4. Визначаються границі інтервалів $int_j = x_{\min} + j \cdot h, j = 0 \dots l$ та середини інтервалів $\bar{x}_j = (int_j + int_{j-1}) / 2, j = 1 \dots l$.
5. Розраховуються емпіричні частоти \hat{m}_j – кількість значень ряду, що потрапили в j -й інтервал ($\hat{m}_j = \hat{m}_j + 1$ якщо $int_{j-1} \leq x_i < int_j, i = 1 \dots n, j = 1 \dots l$).
6. Будується полігон частот (рис. 15.2).

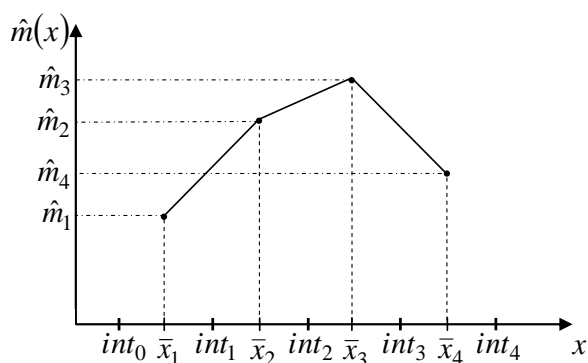


Рис. 15.2. Приклад полігона частот

7. Розраховуються оцінки гістограми $p_j = \hat{m}_j / (n \cdot h)$, якщо будь-яке значення $p_j = 0$, то необхідно зменшити кількість інтервалів l , або об'єднати сусідні інтервали.

8. Будується гістограма за виглядом якої виноситься гіпотеза про тип закону розподілу (рис. 15.3).

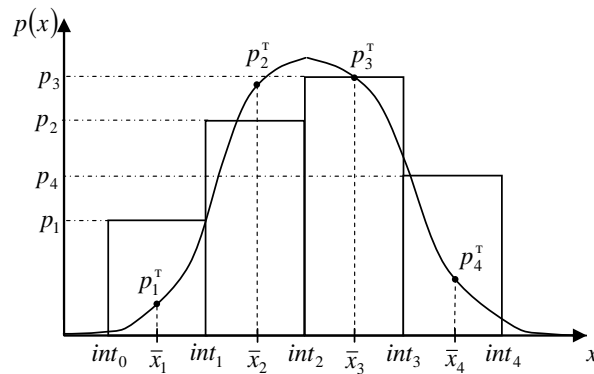


Рис. 15.3. Приклад гістограми та відповідної щільності розподілу

Методика побудови оцінки інтегральної функції розподілу

1. Визначається мінімальне x_{\min} та максимальне x_{\max} вибіркові значення.
2. Експериментальні дані розбиваються на l інтервалів. Кількість інтервалів визначають як найближче ціле від кореня квадратного з обсягу ряду спостережень n ($l = \lceil \sqrt{n} \rceil$), якщо $n < 100$. Якщо обсяг даних $n > 100$, то за формулою Старджесса $l = 1 + 3,31 \lg(n)$.

3. Визначається ширина інтервалів $h = (x_{\max} - x_{\min}) / l$.

4. Визначаються границі інтервалів $int_j = x_{\min} + j \cdot h, j = 0 \dots l$ та середини інтервалів $\bar{x}_j = (int_j + int_{j-1}) / 2, j = 1 \dots l$.

5. Розраховуються емпіричні частоти \hat{m}_j – кількість значень ряду, що потрапили в j -й інтервал:

$$(\hat{m}_j = \hat{m}_j + 1 \text{ якщо } int_{j-1} \leq x_i < int_j, i = 1 \dots n, j = 1 \dots l).$$

6. Розраховуються відносні частоти $w_j = \hat{m}_j / n$.

7. Розраховуються оцінки інтегральної функції : $F_j = \sum_{i=1}^j w_i, j = 1 \dots l$.

8. Будується оцінка інтегральної функції розподілу за виглядом якої виноситься гіпотеза про тип закону розподілу (рис. 15.4).

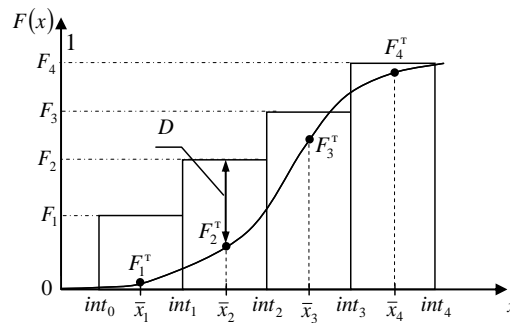


Рис. 15.4. Приклад емпіричної і теоретичної функцій розподілу

15.3.2. Перевірка гіпотез про згоду оцінки закону розподілу з теоретичною моделлю

Критерій χ^2 Пірсона

1. За досліджуваною вибіркою розраховуються параметри для обраної функції щільності $p(x)$. Наприклад для гауссового закону – це оцінки математичного сподівання \bar{x} та СКВ $s(x)$.

2. За обраною функцією щільності $p(x)$ та розрахованими значеннями інтервалів гістограми \bar{x}_j визначаються відповідні теоретичні значення $p_j^T = p(\bar{x}_j)$ (рис. 15.3).

3. Розраховуються теоретичні частоти для всіх інтервалів: $m_j^T = p_j^T n h$.

4. Розраховується статистика χ_p^2 :

$$\chi_p^2 = \sum_{j=1}^l \frac{(\hat{m}_j - m_j^T)^2}{m_j^T} = n h \sum_{j=1}^l \frac{(p_j - p_j^T)^2}{p_j^T}. \quad (154.5)$$

Якщо $\chi_p^2 < \chi_\alpha^2(v)$ – приймається гіпотеза про обраний закон розподілу.

Значення $\chi_\alpha^2(v)$ знаходять за таблицями χ^2 -розподілу (Таблиця 15.1), значення $v = l - k - 1$, – кількість параметрів обраного закону розподілу.

Таблиця 15.1. Граничні значення статистики $\chi^2_\alpha(\nu)$

ν	Рівень значимості α			ν	Рівень значимості α		
	0,1	0,05	0,01		0,1	0,05	0,01
1	2,71	3,84	6,64	11	17,3	19,7	24,7
2	4,61	5,99	9,21	12	18,5	21,0	26,2
3	6,25	7,81	11,3	13	19,8	22,4	27,7
4	7,78	9,49	13,3	14	21,1	23,7	29,1
5	9,24	11,1	15,1	15	22,3	25	30,6
6	10,6	12,6	16,8	16	23,5	26,3	32,0
7	12,0	14,1	18,5	17	24,8	27,6	33,4
8	13,4	15,5	20,1	18	26,0	28,9	34,8
9	14,7	16,9	21,7	19	27,2	30,1	36,2
10	16,0	18,3	23,2	20	28,4	31,4	37,6

Критерій Колмогорова-Смірнова

Мірою розбіжності між емпіричним і теоретичним законами розподілу в критерії Колмогорова-Смірнова є максимальне значення D – модуля різниці між емпіричною $F(x)$ і вибраною теоретичною $F^T(x)$ функціями розподілу:

$$D = \max |F^T(x) - F(x)|. \quad (15.6)$$

Колмогоровим доведено, що незалежно від вигляду передбачуваної функції розподілу неперервної ВВ ξ , у разі необмеженого збільшення числа незалежних вимірювань n , ймовірність нерівності $D\sqrt{n} \leq \lambda$ наближається до межі ймовірності $P(\lambda)$, рівній:

$$P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k \cdot e^{-2k^2\lambda^2}. \quad (15.7)$$

Послідовність застосування критерію наступна:

1. Будується оцінка інтегральної функції розподілу $F(x)$ за виглядом якої висувається гіпотеза про тип закону розподілу, і відповідна інтегральна функція $F^T(x)$ наноситься на один графік з емпіричною функцією розподілу (рис. 15.4).
2. Знаходиться максимальне значення D (15.6), (рис. 15.4).
3. Знаходиться величина λ , яка дорівнює

$$\lambda = D\sqrt{n}. \quad (15.8)$$

4. Рішення щодо гіпотези про закон розподілу приймається двома способами.

4.1. Якщо для заданої ймовірності P , $\lambda < \lambda(P)$, то приймається гіпотеза про те, що закон розподілу $F(x)$ відповідає обраному $F^T(x)$. Значення $\lambda(P)$ приведені в Таблиця 15.2.

Таблиця 15.2. Граничні значення $\lambda(P)$

P	0,99	0,95	0,9
$\lambda(P)$	0,44	0,52	0,57

4.2. За обчисленим значенням λ за формулою (15.7) або Таблиця 15.3 визначається ймовірність $P(\lambda)$ як ймовірність того, що за рахунок випадкових причин максимальна розбіжність між емпіричною і теоретичною функціями розподілу буде не менше, ніж отримана за результатами вимірювань. Отже, якщо ймовірність $P(\lambda)$ достатньо велика, то приймається гіпотеза про відповідність експериментального розподілу $F(x)$ теоретичному $F^T(x)$.

Таблиця 15.3. Значення імовірності $P(\lambda)$

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
0,0	1,000	0,7	0,711	1,2	0,112
0,3	1,000	0,8	0,544	1,3	0,068
0,4	0,997	0,9	0,393	1,4	0,040
0,5	0,964	1	0,270		
0,6	0,864	1,1	0,178		

Лекція 16.

ЗАГАЛЬНІ ВІДОМОСТІ ПРО КОРЕЛЯЦІЙНИЙ ТА РЕГРЕСІЙНИЙ АНАЛІЗ

16.1. Загальні відомості про поняття кореляції

Кореляція в широкому розумінні слова означає зв'язок, співвідношення між об'єктивно існуючими явищами та процесами. В математичній статистиці – це

імовірна або систематична залежність між ознаками різних явищ і процесів, яка на відміну від функціональної, обтяжена дією випадкових факторів.

Функціональний і кореляційний зв'язок – два основні типи зв'язку, що визначають співвідношення між явищами і процесами. Будь-який причинний вплив виражається функціональним або кореляційним зв'язком. Проте не кожна функція або не кожна кореляція відповідає причинній залежності між явищами.

Для ефективного вивчення зв'язків необхідно використовувати сукупності, однорідні відносно тих ознак, зв'язок яких вивчається. Якщо визначають час, витрачений працівником на вироблення одиниці виробу на підприємствах, що розрізняються між собою тільки технічним рівнем виробництва, то слід очікувати, що в цьому випадку матиме місце дуже тісний зв'язок між цими ознаками. Чим тісніший зв'язок між явищами, тим, більше виключається дія другорядних причин і тим менше позначаються (впливають) випадкові чинники. В результаті кореляційний зв'язок наближається до функціонального. Тому функціональний зв'язок може розглядатися як граничний випадок кореляції.

Іноді істинний функціональний зв'язок важко виявити внаслідок похибок вимірювання, зміни умов реалізації, помилкового або формального розгляду причинних стосунків. І тоді невинні змінні, що знаходяться у функціональній залежності, перетворюються у випадкові, а зв'язок починає набувати стохастичного характеру. Наприклад, закон вільного падіння виконується точно тільки в безповітряному просторі. У випадку відхилення від цієї умови закон проявляється у вигляді кореляції.

Причинний вплив може бути виражений у вигляді функціонального або кореляційного зв'язку. Але звідси зовсім не витікає зворотне твердження, що за будь-яким кореляційним або функціональним зв'язком ховається причинна залежність. Завдання ж наукового дослідження полягає в пошуку причинних залежностей. Тільки знання істинних причин явищ дозволяє правильно тлумачити спостережувані закономірності. Проте кореляція, як формально-статистичне поняття, сама по собі не розкриває причинного характеру зв'язку. За

допомогою кореляційного аналізу не можна вказати, яке з явищ брати за причину, а яке за наслідок. Кореляція лише дає оцінку сили, або тісноти зв'язку.

На рис. 16.1 наведено класифікацію видів кореляції.

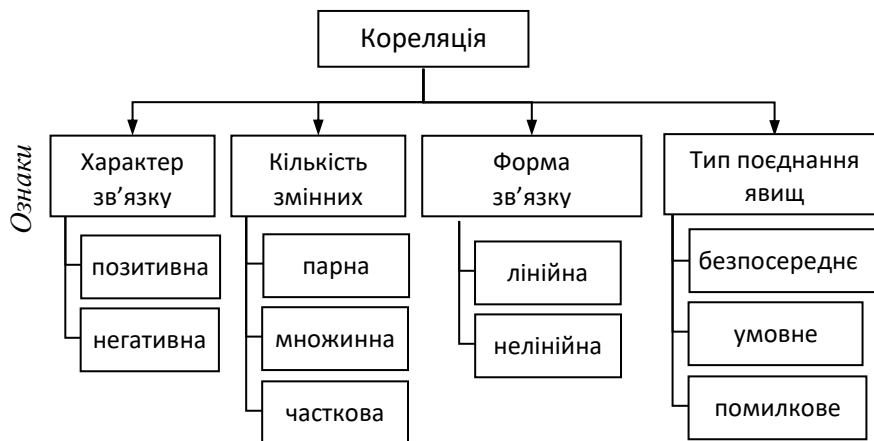


Рис. 16.1. Класифікаційні ознаки та види кореляції

1. За характером кореляції розрізняють:

а) позитивну кореляцію. Вона має місце, якщо зі збільшенням або зменшенням значень однієї змінної значення іншої відповідно збільшуються або зменшуються (рис. 16.2, *а*). Позитивна кореляція існує, наприклад, між продуктивністю праці і заробітною платою, між ростом і вагою людини. Вона називається ще рівно напрямленою (чи прямою) кореляцією;

б) негативну (або обернену) кореляцію. При цьому виді кореляції із збільшенням або зменшенням значень однієї змінної значення іншої відповідно зменшуються або збільшуються (рис. 16.2, *б*). Негативна кореляція існує, наприклад, між продуктивністю праці і вартістю виробу.

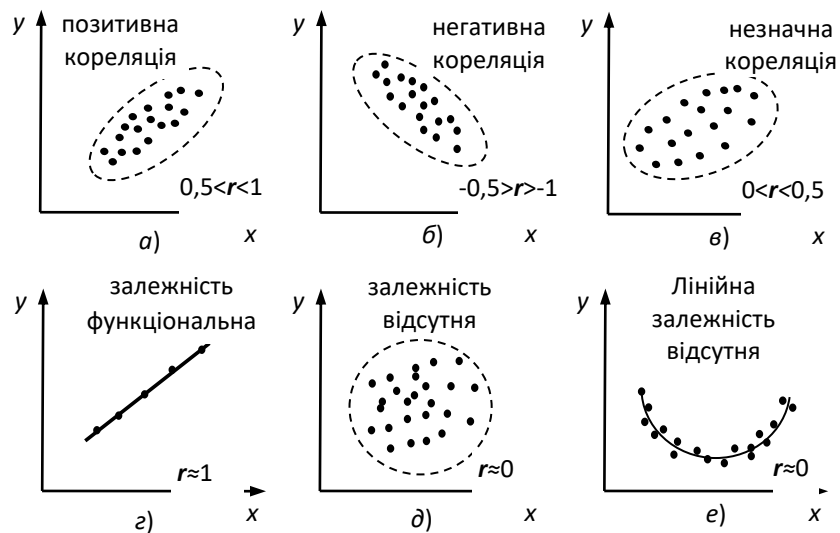


Рис. 16.2. Графічне зображення функціонального та стохастичного зв'язків між величинами x та y у вигляді діаграм розсіювання

2. За кількістю змінних розрізняють:

а) просту, або парну, кореляцію. Це кореляція між двома змінними;

б) множинну кореляцію. Це кореляція між більш ніж двома змінними. За її допомогою охоплюється увесь причинно-наслідковий комплекс, в якому окремі явища, є наслідком не однієї, а декількох причин. Множинна кореляція служить віддзеркаленням об'єктивно існуючих множинних зв'язків. Встановлення цих зв'язків разом з їх поясненням, розкриває механізм явищ;

в) часткову кореляцію. Це кореляція між двома змінними при "фіксованому" впливі інших змінних, що включені в аналіз. За її допомогою досліджується причинно-наслідковий комплекс і розкривається внутрішня структура співвідношень між ознаками досліджуваних явищ. Однак якщо визначати кореляцію між залежною змінною (наслідок) і кожною пояснюючою змінною (причиною) окремо, то вплив інших змінних позначатиметься на мірі зв'язності виділених змінних. Це може привести до помилкових висновків.

3. За формою зв'язку розрізняють:

а) лінійну кореляцію. Для цього виду кореляції між досліджуваними змінними існують лінійні співвідношення;

б) нелінійну кореляцію. При цьому виді кореляції між досліджуваними

змінними існують нелінійні співвідношення.

4. За типом поєднання явищ розрізняють:

а) безпосередню кореляцію. В цьому випадку досліджувані явища сполучені між собою безпосередньо. Визначальна змінна чинить прямий вплив на залежну змінну. Безпосередня кореляція існує, якщо з одного явища логічно витікає інше, для пояснення цієї кореляції не треба залучати інші явища;

б) непряму кореляцію. Про непряму кореляцію говорять у випадку, коли досліджувані змінні не мають безпосереднього причинно-наслідкового зв'язку, а визначаються загальною для них причиною. Логічно обґрунтувати такий зв'язок можна за допомогою інших явищ. Але у цьому разі існує небезпека переходу на формальний шлях дослідження, що може привести до помилкової кореляції.

Приклад непрямої кореляції дає статистика дореволюційної Росії. Було встановлено тісну кореляцію між числом пожеж і розмірами врожаю. Очевидно, погані врожаї ніяк не можна вважати причиною пожеж в будівлях, а боротися з пожежами не можливо за допомогою агротехнічних заходів. В дійсності в цьому прикладі між змінними існує лише непрямий зв'язок: і розміри врожаю, і число пожеж суттєво залежать від третього явища – метеорологічних умов: сильна посуха призводить до поганого врожаю і до виникнення пожеж.

в) помилкова кореляція. Під помилковою кореляцією (нонсенс-кореляцією) розуміють формальний зв'язок між явищами, що не знаходить логічного пояснення та оснований лише на кількісному співвідношенні між ними.

В статистиці відомий приклад помилкової кореляції між числом лелек, що звили гнізда в районах Східної Пруссії, і народжуваністю дітей. Виконані заради жарту обчислення дали позитивну кореляцію між цими явищами.

Завдання кореляційного аналізу:

1) Визначення ступеня зв'язку (тісноти, сили, строгості, інтенсивності) двох і більше явищ. Загальні знання про об'єктивно існуючі причинні зв'язки повинні доповнюватися науково пояснюваними знаннями про міру залежності між ними. Тут розмова йде, в основному, про верифікацію вже відомих зв'язків.

2) Відбір факторів, що чинять найбільш суттєвий вплив на результативну ознаку, на основі вимірювання ступеня зв'язності між явищами. Відібрані фактори використовують у подальшому аналізі. Найважливішими факторами вважаються ті, які корелюють найсильніше з досліджуваними явищами.

3) Виявлення невідомих причинних зв'язків. При вирішенні цієї задачі необхідно враховувати своєрідні взаємозв'язки в причинно-наслідковому комплексі і особливості науково-методологічних правил статистичного дослідження, що спираються на кількісні зв'язки між явищами.

16.2. Коефіцієнт парної кореляції Пірсона

Ступінь стохастичного лінійного зв'язку між ВВ ξ і η із гауссовим розподілом незалежно від їхнього роду оцінюється коефіцієнтом кореляції:

$$r = \frac{\mathbf{M}\{(\xi - \mathbf{M}\{\xi\})(\eta - \mathbf{M}\{\eta\})\}}{\sqrt{\mathbf{D}(\xi)\mathbf{D}(\eta)}} = \frac{\text{cov}(\xi, \eta)}{\sigma_{\xi}\sigma_{\eta}}, \quad (16.1)$$

де $\text{cov}(\xi, \eta)$ – оператор коваріації ВВ ξ і η .

Властивості коефіцієнта кореляції:

1. $|r| \leq 1$;
2. Якщо $r = 0$, то ξ і η – незалежні ВВ;
3. Якщо $|r| = 1$, то між ξ і η має місце функціональний зв'язок;
4. Якщо $0 < |r| < 1$, то між ξ і η має місце стохастичний зв'язок.

Для визначення коефіцієнта кореляції між ξ і η проводять ряди незалежних випробувань, результатом яких є пари величин (x_i, y_i) – реалізації (ξ, η) . На кореляційному полі – координатній площині у вигляді точок з координатами (x_i, y_i) , будується діаграма розсіювання (рис. 16.2), за її аналізом висувається гіпотеза про вид залежності між ξ і η (лінійна/нелінійна). У разі гіпотези про лінійність визначають статистичну оцінку коефіцієнта кореляції:

$$\hat{r} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (16.2)$$

де $\bar{x} = n^{-1} \sum_{i=1}^n x_i$, $\bar{y} = n^{-1} \sum_{i=1}^n y_i$.

За малого обсягу вибірки n оцінка \hat{r} є зміщеною відносно генерального параметра r . Для усунення зміщення оцінку корегують:

$$\hat{r}^* = \hat{r} \left[1 + \frac{1-r^2}{2(n-3)} \right]. \quad (16.3)$$

Значення оцінок коефіцієнта кореляції завжди є відмінним від нуля. Тому виникає задача перевірки значущості коефіцієнта кореляції. Гіпотеза H_0 про випадковість відхилень від нуля вибіркового коефіцієнта кореляції у випадку генеральної сукупності з параметром $r = 0$, перевіряється за t -розподілом з $(n-2)$ степенями свободи (таблиця Додатку 3) статистики:

$$\hat{t} = \frac{\hat{r} \sqrt{n-2}}{\sqrt{1-\hat{r}^2}}. \quad (16.4)$$

Якщо $|\hat{t}| \geq t_{n-2, \alpha}$ – гіпотеза $H_0 : r = 0$ відхиляється, і \hat{r} визнається значущим.

16.3. Кореляційний аналіз, що ґрунтується на порядкових статистиках.
Рангова кореляція

У випадку, коли визначається взаємозв'язок між рядами спостережень ВВ, які мають довільні закони розподілів, а також коли аналізуються ознаки, що не піддаються безпосередньому кількісному оцінюванню, однак мають ряд якісних градацій застосовують коефіцієнти рангової кореляції.

16.3.1. Коефіцієнт рангової кореляції Спірмена

Для розрахунку коефіцієнта рангової кореляції за обома рядами спостережень x_i та y_i формують ряди рангів R_x та R_y . За вибіркою будуються варіаційні ряди, потім кожному значенню вихідної вибірки ставиться ранг у відповідності його номера у варіаційному ряду. Якщо у вибірці є однакові значення, то для них розраховується середній ранг, який приписується кожному з чисел. Такі ранги називаються зв'язаними, а їх група – зв'язною.

Коефіцієнт рангової кореляції Спірмена r_s :

$$r_s = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n D_i^2, \quad (16.5)$$

де n – кількість пар спостережень, $D_i = R_{x_i} - R_{y_i}$ – різниці пар рангів у досліджуваних вибірках.

Для перевірки правильності розрахунків сума різниць рангів D_i повинна дорівнювати нулю.

Коефіцієнт рангової кореляції Спірмена має наступні властивості:

1. $|r_s| \leq 1$;
2. $r_s = 1$ – випадок рівності рангів у обох рядах, тобто різниці $D_i = 0$;
3. $r_s = -1$, має місце протилежне впорядкування послідовностей рангів;
4. $r_s = 0$, кореляція відсутня.

У випадку, коли коефіцієнт рангової кореляції Спірмена визначений для кількісних показників, розподілених за гауссовим законом, для визначення коефіцієнта кореляції Пірсона має місце рівняння: $r = 2 \sin(\pi/6) r_s$.

16.3.2. Коефіцієнт рангової кореляції Кенделла

Іншу рангову міру зв'язку запропонував М. Дж. Кенделл. Вона базується на визначенні узгодженості пар рангів у досліджуваних вибірках. Якщо пари (x_i, y_i)

узгоджені ($1 \leq i < j \leq n$), то повинні виконуватися нерівності $x_i < y_i$ та $y_i < y_j$ або $x_i > x_j$ та $y_i > y_j$ (тобто $\text{sign}(x_j - x_i) \cdot \text{sign}(y_j - y_i) = 1$). Нехай S – кількість узгоджених пар, а Q – кількість неузгоджених пар. Тоді міра перевищення узгодженості над неузгодженістю є статистика:

$$T = S - Q = \sum_{i < j} \text{sign}(x_j - x_i) \cdot \text{sign}(y_j - y_i);$$

$$T \in [-0,5 \cdot n(n-1); 0,5 \cdot n(n-1)]$$
(16.6)

Значення $\max T = 0,5 \cdot n(n-1)$ визначається через абсолютно узгодженні порядки x_1, \dots, x_n та y_1, \dots, y_n . Для визначення ступеня узгодженості вибірок застосовується коефіцієнт рангової кореляції Кенделла:

$$\hat{r}_k = \frac{T}{T_{\max}} = \frac{2(S - Q)}{n(n-1)} = 1 - \frac{4}{n(n-1)} Q.$$
(16.7)

Він має такі властивості:

1. Якщо між послідовностями порядкових статистик існує повна відповідність, тобто кожний елемент займає одне й те саме місце в обох рангах, то $\hat{r}_k = 1$.

2. Якщо має місце взаємообернена залежність тобто у другій послідовності ранги розміщені в зворотньому порядку порівняно з першою, то $\hat{r}_k = -1$.

3. У всіх інших випадках $-1 < \hat{r}_k < 1$.

Методика оцінювання кореляції двох вибірок $X = (x_1, \dots, x_n)$ і $Y = (y_1, \dots, y_n)$ за способом, який запропонував Кенделл, полягає у наступному:

1. Впорядковуємо вибірку X за зростанням, а значення вибірки Y перерозподіляємо за значеннями впорядкованої вибірки X , зберігаючи взаємну відповідність значень x_i та y_i як у початковій вибірці (отримуємо вибірку Y^*).

2. Визначаємо ранги отриманої вибірки $Y^* - R^*$.

3. За отриманою послідовністю рангів R^* визначаємо кількість інверсій. *Інверсією* Q_i називають пару рангів R_i^* та R_j^* ($i < j$, $i = 1, \dots, n-1$, $j = i+1, \dots, n-1$), якщо $R_i^* > R_j^*$. Отриману сумарну кількість інверсій використовують як оцінку Q .

4. За формулою (16.7) визначаємо оцінку коефіцієнта кореляції Кенделла.

16.4. Регресійний аналіз

Якщо взаємозв'язок ВВ η та ξ виявлено, наступним етапом є встановлення виду залежності між η та ξ . Якщо залежність в загальному випадку виражається як $\mathbf{M}(\eta) = g(\xi)$, тобто шуканою є залежність математичного сподівання ВВ η від ξ , то така залежність називається *регресійною*, а визначення такої залежності та оцінювання її статистичних властивостей є змістом *регресійного аналізу*.

Залежно від моделі шуканої функції $g(\xi)$ розрізняють *лінійну, нелінійну та множинну регресії*, а відповідно до методів оцінювання параметрів моделі – *параметричну та непараметричну регресії*.

Якщо досліджуються вибірки $(x_i, i = \overline{1, n})$ та $(y_i, i = \overline{1, n})$, загальна схема регресійного аналізу містить наступні етапи:

1. Знаходження вибіркової оцінки регресії.

1.1. Вибір моделі та оцінювання її параметрів.

1.2. Оцінювання статистичної значущості отриманих параметрів.

2. Оцінювання адекватності обраної моделі, тобто оцінювання статистичної значущості вибіркової регресії у порівнянні із розсіюванням значень y_i , яке характеризується СКВ s_y .

3. Визначення довірчих областей, які із заданою імовірністю містять середні або індивідуальні значення η , за якими визначають межу похибки (розширеної невизначеності) значення η для заданого значення ξ .

Найбільш повно розроблений апарат регресійного аналізу який передбачає,

що вибіркові значення y_i статистично незалежні та мають гауссовий розподіл. У випадку порушення цієї умови застосовують методи непараметричної регресії.

16.4.1. Лінійна регресія

У випадку лінійної регресії розглядається модель залежності двох ВВ η та ξ в наступному загальному виді

$$\eta(\xi) = k\xi + b. \quad (16.8)$$

Початковими даними (рис. 16.3) можуть бути n пар вибіркових значень (x_i, y_i) , $i = \overline{1, n}$, або у випадку значень $y_{i,k}$ величини η з багаторазовими спостереженнями обсягом m , тобто якщо $k = \overline{1, m}$ – n пар x_i та середніх значень \bar{y}_i . Далі будемо використовувати лише позначення y_i .

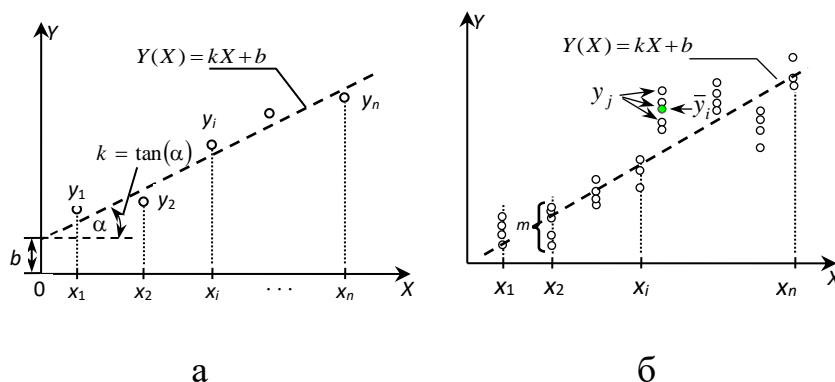


Рис. 16.3. Графічне зображення початкових даних для визначення регресії: значення η отримані з одноразових (а) чи багаторазових (б) вимірювань

Існує декілька методів оцінювання коефіцієнтів регресії k і b – спрощені, метод найменших квадратів (МНК) та метод найбільшої правдоподібності. Розглянемо декілька методів оцінювання коефіцієнтів лінійної регресії – *метод Бартлетта-Кенуя та метод Керріча*, які належать до кола спрощених методів.

Метод Бартлетта – Кенуя. Пари спостережень (x_i, y_i) упорядковуються по x і розбиваються на 3 приблизно рівні групи обсягом M (перша і остання групи повинні бути обов'язково рівного обсягу). У кожній групі знаходяться суми $\sum y_i$ і $\sum x_i$ (позначимо їх відповідно Y_1, Y_2, Y_3 та X_1, X_2, X_3). Тоді коефіцієнти

регресії оцінюються за допомогою співвідношень:

$$k = \frac{Y_3 - Y_1}{X_3 - X_1}, \quad b = \bar{y} - k\bar{x} \quad \text{або} \quad b = \frac{Y_2}{M} - k \frac{X_2}{M}. \quad (16.9)$$

СКВ коефіцієнта k : $S_k = \frac{0,8s\sqrt{n}}{X_3 - X_1}$, де $s = \frac{8}{9} \sum_{i=1}^{n-1} \frac{|y_i - y_{i+1}|}{n}$.

У виразі (16.9) замість значень сум або середніх значень можна використовувати відповідні значення медіан, що забезпечує більшу стійкість до наявності спостережень з надмірною похибкою.

Якщо n пар спостережень розбивають на чотири групи, які містять $1/6, 1/3, 1/3$ та $1/6$ частин спостережень, то коефіцієнт k та його СКВ визначаються як

$$k = \frac{3Y_1 + Y_2 - Y_3 - 3Y_4}{3X_1 + X_2 - X_3 - X_4}, \quad s_k = \frac{1,9s\sqrt{n}}{3X_1 + X_2 - X_3 - X_4}. \quad (16.10)$$

Ці оцінки застосовні для великих вибірок обсягу $n \geq 100$.

Метод Керріча. Для окремого випадку лінійної залежності виду $y = bx$ Керріч запропонував простий метод оцінювання. Для множини $\{x_i, y_i, i = \overline{1, n}\}$ обчислюємо різниці d_i , оцінки їх математичного сподівання та СКВ

$$d_i = \lg y_i - \lg x_i; \quad \bar{d} = \frac{1}{n} \sum_{i=1}^n d_i, \quad S_{\bar{d}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (d_i - \bar{d})^2}. \quad (16.11)$$

Оскільки будь-яке відношення y_i / x_i є оцінкою коефіцієнта k , то будь-яке значення d_i є оцінкою $\lg(k)$. Коли $S_{\bar{d}} / \bar{d} \ll 1$, оцінкою $\lg(k)$ є величина \bar{d} , тобто $\lg(k) = \bar{d}$. Отже, оцінка буде дорівнювати $\tilde{b} = 10^{\bar{d}}$.

Метод найменших квадратів. Значення коефіцієнтів регресії k і b знаходять із умови мінімуму суми квадратів відхилень початкових значень y_i і значень отриманих за побудованою функцією $Y(X) = y_{\text{перг.}}$:

$$\min \left(\sum_{i=1}^n (y_{\text{перг.}} - y_i)^2 \right). \quad (16.12)$$

З цієї умови, значення коефіцієнтів регресії знаходять за формулами:

$$b = \frac{\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n (x_i \cdot y_i)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}, \quad k = \frac{n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}. \quad (16.13)$$

Після визначення коефіцієнтів k і b перевіряють їх статистичну значущість.

Перевірка значущості коефіцієнтів k і b

Для оцінювання значущості використовується критерій Стюдента для перевірки 2-х гіпотез:

H_0 – оцінка коефіцієнта статистично незначуща, дорівнює 0;

H_1 – оцінка коефіцієнта статистично значуща не дорівнює 0;

Розраховується t -статистика:

$$t_k = |k|/S_k, \quad t_b = |b|/S_b, \quad (16.14)$$

де S_k, S_b – оцінки СКВ коефіцієнтів k і b , отримані згідно з формулами:

$$S_k = \frac{S_y}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}, \quad S_b = S_k \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}, \quad S_y = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - y_{\text{рег},i})^2}, \quad (16.15)$$

де \bar{x} – середнє арифметичне вибірових значень ряду $\{x_i, i = \overline{1, n}\}$.

Якщо $t_k < t_\alpha(v)$, або $t_b < t_\alpha(v)$ приймається гіпотеза H_0 , тобто для обсягу вибірки n коефіцієнт k або b є статистично незначущим.

Якщо $t_k \geq t_\alpha(v)$ приймається гіпотеза H_1 , відповідний коефіцієнт є статистично значущим. Значення $t_\alpha(v)$ знаходиться як $(1-a) \cdot 100\%$ односторонній довірчий інтервал з таблиці розподілу Стюдента для ступенів свободи $v = n - 2$ та рівня значущості $\alpha = 1 - P$.

Перевірка адекватності отриманої моделі

Загалом проводиться перевірка гіпотези, що обрана модель виду $Y = g(X)$, яка побудована за результатами парних вимірювань $x_i, y_i, i = 1 \dots n$, задовільно погоджується з цими даними. Як правило приймають рівень значущості $\alpha = 0,05$. Для перевірки зазначеної гіпотези застосовують критерій Фішера, за умови, що похибка значень y має гауссовий закон розподілу.

Критерій Фішера. Перевірка адекватності моделі зводиться до перевірки двох гіпотез:

H_0 – модель не адекватна; H_1 – модель адекватна (лінійна);

Ідея перевірки полягає в тому, що значення спостережень y_i (\bar{y}_i) повинні лежати приблизно на прямій $\eta(\xi)$, тобто їх відхилення від прямої регресії не мають бути занадто великими по відношенню до відхилення значень в групі спостережень $y_{i,k}$ від їх середнього \bar{y}_i .

Статистика Фішера розраховується за формулами:

$$F = S_y^2 / S_{гр}^2, \quad (16.16)$$

де S_y – СКВ розсіювання значень y_i навколо регресії (6.12),

$S_{гр}$ – СКВ розсіювання значень в групі спостережень навколо \bar{y}_i , у випадку багаторазових спостережень Y :

$$S_{гр}^2 = \frac{1}{n(m-1)} \sum_{i=1}^n \sum_{k=1}^m (y_{i,k} - \bar{y}_i)^2, \quad (16.17)$$

у випадку однократних вимірювань ВВ η значення $S_{гр}^2$ має бути відомим до проведення експерименту.

Якщо $F < F_\alpha(v_1, v_2)$ – модель адекватна, якщо $F \geq F_\alpha(v_1, v_2)$ – модель не адекватна. Значення $F_\alpha(v_1, v_2)$ знаходиться з таблиці розподілу Фішера для

ступенів свободи $v_1 = n - 2, v_2 = m - 1$ та рівня значущості $\alpha = 1 - P$. У випадку однократних вимірювань Y значення $F_\alpha(v_1, \infty)$ знаходяться з Таблиці 16.1.

Таблиця 16.1. Значення коефіцієнтів Фішера $F_\alpha(v_1, \infty)$

v_1	2	3	4	5	6	7	8	9	10	11	12
$\alpha=0,01$	21,31	2,09	1,95	1,85	1,78	1,72	1,68	1,64	1,61	1,58	1,55
$\alpha=0,05$	3,01	2,61	2,38	2,22	2,11	2,02	1,95	1,89	1,84	1,80	1,76

Визначення довірчих інтервалів значень у розрахованих за регресією

Оцінюється СКВ випадкової складової похибки середнього значення \bar{y} та індивідуального значення y , визначених за отриманою залежністю:

$$s_{\bar{y}}(x) = u_{\bar{y}}(x) = S_y \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}, \quad (16.18)$$

$$s_y(x) = u_y(x) = S_y \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}. \quad (16.19)$$

За отриманими значеннями s оцінюють довірчий інтервал $\Delta(\cdot)$ випадкової похибки середнього та індивідуального значення y :

$$\Delta_{\bar{y}}(x) = t_{\alpha/2}(v) \cdot s_{\bar{y}}(x), \quad \Delta_y(x) = t_{\alpha/2}(v) \cdot s_y(x), \quad (16.20)$$

де $t_{\alpha/2}(v)$ – квантиль розподілу Стьюдента, для довірчої імовірності $P = 1 - \alpha$ та $v = n - 2$ ступеней свободи.

Отже, як видно з формул (16.18)-(16.20) довірчі області, які із заданою імовірністю P містять середні або індивідуальні значення y залежать від відстані значення x (за яким визначають y) від середнього \bar{x} (див рис. 16.4).

16.4.2 Нелінійна регресія

Загальна методика визначення параметрів нелінійної регресії передбачає:

- визначення загальної нелінійної моделі регресії;
- виконання нелінійних перетворень експериментальних даних з метою зведення нелінійної моделі до лінійної;
- визначення параметрів для трансформованої лінійної моделі регресії;
- визначення параметрів нелінійної моделі регресії;
- оцінювання границі похибки лінеаризованої функції;
- оцінювання границі похибки шуканої функції.

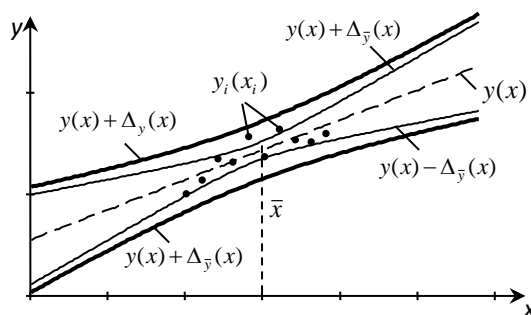


Рис. 16.4. Графік довірчих інтервалів регресії

Розглянемо ряд характерних прикладів нелінійної регресії.

Степенева функція (геометрична регресія). Розглянемо степеневу функцію, яка в найбільш загальному вигляді задається наступним виразом:

$$F(x, k, m) = kx^m. \quad (16.21)$$

За умови, що $F(x, k, m) > 0$, $x > 0$, прологарифмуємо (16.21):

$$\ln F(x, k, m) = \ln k + m \ln x. \quad (16.22)$$

Якщо функція $F(\cdot)$ є наближаючою для деякої функції $f(\cdot)$ то функція $\ln F(\cdot)$ буде наближаючою для функції $\ln f(\cdot)$. Якщо ввести змінну $x' = \ln x$, тоді, як виходить з (16.22), $\ln F(\cdot)$ буде лінійною функцією від x' : $\Phi(x')$. Позначимо:

$$m = A, \quad \ln k = B, \quad (16.23)$$

тоді (16.23) приймає вигляд:

$$\Phi(x', A, B) = Ax' + B. \quad (16.24)$$

Отже задача побудови регресії зводиться до знаходження наближаючої функції у вигляді лінійної. В результаті для знаходження регресії у вигляді степеневій використовують наступний алгоритм:

1. Логарифмують вихідні значення x і y .
2. Згідно (16.13) визначають параметри A і B як, відповідно, k і b , наближаючої функції вигляду (16.24).
3. Знаходять згідно з (16.23) значення параметрів k і m : $k = e^B$, $m = A$.
4. Оцінюють границі похибки лінеаризованої функції $\Phi(x, A, B)$: $\Delta_{\text{лін},\bar{y}}(x)$ та $\Delta_{\text{лін},y}(x)$ (6.20).
5. Оцінюють межі похибки шуканої функції $y(x)$:

$$y(x) \pm \Delta_{\bar{y}}(x) = e^{(\Phi(\ln(x), A, B) \pm \Delta_{\text{лін},\bar{y}}(\ln(x)))}, \quad y(x) \pm \Delta_y(x) = e^{(\Phi(\ln(x), A, B) \pm \Delta_{\text{лін},y}(\ln(x)))}.$$

Показникова функція. Ця функція задається виразом:

$$F(x, k, m) = ke^{mx}, k > 0 \quad (16.25)$$

Прологарифмувавши вираз (16.26), за умови $F(x, k, m) > 0$, $x > 0$, маємо:

$$\ln F(x, k, m) = \ln k + mx. \quad (16.26)$$

Позначимо:

$$x' = x, \quad m = A, \ln k = B, \quad (16.27)$$

тоді (16.26) набуває вигляду (16.24) і задача побудови регресії зводиться до попередньої. Отже для знаходження регресії у вигляді показникової використовують наступний алгоритм:

1. Логарифмують вихідні значення тільки y .
2. Згідно (16.13) визначають параметри A і B як, відповідно, k і b наближаючої функції вигляду (16.25).
3. Знаходять згідно з (16.26) значення параметрів k і m : $k = e^B$, $m = A$.
4. Оцінюють границі похибки лінеаризованої функції $\Phi(x, A, B)$: $\Delta_{\text{лін},\bar{y}}(x)$ та $\Delta_{\text{лін},y}(x)$ (16.20).

5. Оцінюють границі похибки шуканої функції $y(x)$:

$$y(x) \pm \Delta_y(x) = e^{(\Phi(x,A,B) \pm \Delta_{\text{лін},\bar{y}}(x))}, y(x) \pm \Delta_y(x) = e^{(\Phi(x,A,B) \pm \Delta_{\text{лін},y}(x))}.$$

Узагальнені відомості для деяких видів нелінійної регресії експериментальних даних наведено у Таблиці 16.2.

Таблиця 16.2. Нелінійні функції та оцінки їх параметрів за лінійною регресією.

Назва	Функція	Проміжне перетворення		Значення	
		для x	для y	k	m
Степенева	$y = kx^m$	$\ln(x)$	$\ln(y)$	e^B	A
Показникова	$y = ke^{mx}$	—	$\ln(y)$	e^B	A
Дробово-лінійна	$y = \frac{1}{kx+m}$	—	$\frac{1}{y}$	A	B
Логарифмічна	$y = k \ln x + m$	$\ln(x)$	—	A	B
Гіперболічна	$y = k \frac{1}{x} + m$	$\frac{1}{x}$	—	A	B

Питання для самоконтролю

1. Поясніть терміни “вибірка”, “генеральна сукупність”, “порядкова статистика”, “статистична оцінка”.
2. Які показники якості статистичних оцінок вам відомі?
3. Наведіть означення незсуненої, обґрунтованої та ефективної точкової оцінки та поясніть їх зміст.
4. Наведіть незсунену, обґрунтовану та ефективну оцінку математичного сподівання генеральної сукупності.
5. Які оцінки називають точковими?
6. Які оцінки називають інтервальними?
7. Що таке ранг елемента вибірки та які правила його визначення?
8. Розкрийте зміст визначення точкових оцінок за методом максимальної правдоподібності.
9. Як визначаються вибіркові моменти за рядом спостережень?
10. Розкрийте зміст визначення точкових оцінок за методами моментів та

квантилів.

11. Що таке квантиль розподілу?
12. Наведіть відомі вам методи отримання точкових оцінок параметрів закону розподілу Гаусса.
13. Наведіть відомі вам методи отримання інтервальних оцінок параметрів закону розподілу Гаусса.
14. Розкрийте зміст поняття «статистична гіпотеза» та наведіть приклади аналітичного запису основної і альтернативної гіпотез.
15. Наведіть визначення помилок першого і другого роду.
16. Для якої мети застосовують діаграму $\beta_1\beta_2$?
17. Наведіть алгоритм побудови полігону частот, гістограми та інтегральної функції розподілу? Як перевірити їх правильність?
18. Поясніть суть терміну “кореляція”. Як оцінюють ступінь кореляційного зв’язку?
19. Наведіть класифікацію видів кореляції за різними ознаками.
20. Проілюструйте графічно різні види парної кореляції.
21. Поясніть зміст парної, множинної та часткової видів кореляції.
22. Наведіть формулу статистичної оцінки коефіцієнта парної кореляції Пірсона.
23. В яких випадках використовують кореляційний аналіз, що ґрунтується на порядкових статистиках?
24. Які передумови застосування регресійного аналізу ви знаєте?
25. Які методи оцінювання коефіцієнтів регресійної моделі вам відомі?
26. Наведіть загальну послідовність застосування регресійного аналізу.
27. Які передумови застосування методу найменших квадратів?
28. Розкрийте сутність довірчих границь рівняння регресії.

Розділ 4. ВИБРАНІ ПИТАННЯ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ

Лекція 17.

ОСНОВНІ ПОЛОЖЕННЯ ТЕОРІЇ ВИПАДКОВИХ ПРОЦЕСІВ

Спостерігаючи та вивчаючи фізичні явища навколишнього світу, діагностуючи різні технічні та біологічні об'єкти дослідник часто зустрічається з процесами, перебіг яких в часі неможливо наперед передбачити і представити певною детермінованою функцією. Ця невизначеність викликана різними випадковими чинниками, які впливають на процес і не можуть бути заздалегідь виявлені, усунені чи враховані. Характерним прикладом є виникнення теплового шуму в провідниках. Це явище привносить до того, що в електронних схемах напруга теплових шумів додається до інформаційних сигналів і, за низьких рівнів останніх, може створити суттєві перешкоди у розв'язанні задач вимірювання інформативних параметрів сигналів. Крім того самі досліджувані явища можуть мати випадкову природу. Такі явища породжують в технічних системах сигнали, аргументами яких є і час, і випадкові події. Методів класичної теорії ймовірностей недостатньо для створення моделей таких сигналів та дослідження таких явищ. Для їх аналізу створено спеціальний розділ теорії ймовірностей – *теорію випадкових процесів*, що вивчає закономірності динаміки розвитку випадкових явищ.

17.1. Визначення випадкового процесу

У широкому розумінні процес – це послідовна в часі зміна явищ, станів, стадій у розвитку будь-чого (від латинського *procesus* – просування).

Означення. *Випадковим процесом* називається будь-який процес, що протікає в часі t і залежить від елементарних подій ω з області подій Ω .

Означення. Параметричне сім'я (сукупність) дійсних ВВ називається *дійсним випадковим процесом*. Випадковий процес позначається як параметрична множина $\{\xi(\omega, t), \omega \in \Omega, t \in T\}$, де T – область визначення процесу $\xi(t)$.

Означення. Багатовимірні сумісні функції розподілу ймовірності

випадкових процесів $\xi(\omega, t_1), \xi(\omega, t_2), \dots, \xi(\omega, t_n)$ для всіх наборів $\{t_j, j = \overline{1, n}\}$ і $n = 1, 2, 3, \dots$

$$F_{(n)}(x_1, \dots, x_n; t_1, \dots, t_n) = P\{\xi(t_1) < x_1, \dots, \xi(t_n) < x_n\} \quad (17.1)$$

називаються *скінченномірними функціями розподілу* випадкового процесу $\xi(t)$.

Випадковий процес вважається повністю заданим якщо задана послідовність його скінченномірних функцій розподілу

$$\begin{aligned} & F_{(1)}(x; t), \\ & F_{(2)}(x_1, x_2; t_1, t_2), \\ & \dots \quad \dots \quad \dots \quad \dots \\ & F_{(n)}(x_1, \dots, x_n; t_1, \dots, t_n), \\ & t_j \in T, \quad j = \overline{1, n}, \quad n = 1, 2, 3, \dots \end{aligned} \quad (17.2)$$

Означення. Процеси, які можна задати послідовністю (17.2) називаються *сепарабельними*.

Це поняття пояснимо наступним прикладом. Сепарабельність функції, у випадку функції декількох змінних – це можливість відокремлення впливу аргументів на загальний результат. Наприклад, функція виду $f(x_1, \dots, x_n) = \sum f_i(x_i), i = \overline{1, n}$ – сепарабельна, оскільки кожна складова є функцією одного аргументу, натомість функція $f(x_1, \dots, x_n) = f(x_1x_2 + x_2x_3 + \dots + x_{n-1}x_n)$ не сепарабельна.

17.2. Реалізації випадкового процесу

Означення. Функція, що з'являється в результаті спостереження за випадковим процесом, називається його *реалізацією* або *вибірковою функцією процесу*.

Випадковий процес має подвійний характер. Функція $\xi(t_1, \omega)$ для кожної фіксованої елементарної події $\omega \in \Omega$ є звичайною не випадковою функцією дійсного параметра t . З іншого боку, для кожного фіксованого $t = t_1$ функція

$\xi(t_1, \omega) \in \text{ВВ}$ і має визначений розподіл ймовірності

$$F(x, t_1) = P\{\xi(t_1) < x\}. \quad (17.3)$$

Ілюстрацію цього положення подано на рис. 17.1.

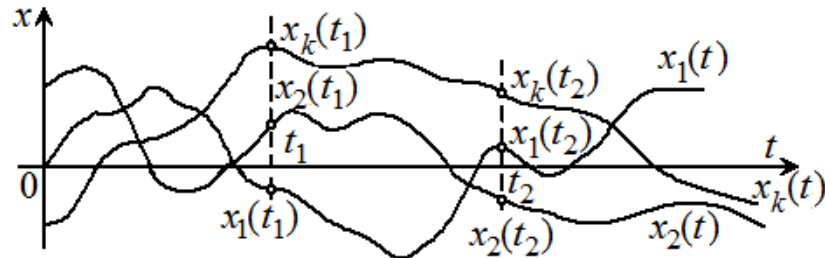


Рис. 17.1. Приклад реалізацій випадкового процесу

На рис. 17.1 показано сукупність трьох реалізацій випадкового процесу $\xi(\omega, t)$. Така сукупність називається ще *ансамблем реалізацій*.

Якщо зафіксувати два моменти часу t_1 і t_2 , то в кожному експерименті отримаємо два значення $x(t_1)$ і $x(t_2)$ однієї реалізації. В цілому сумісний розгляд цих значень приводить до системи двох ВВ $(\xi(\omega_g, t_1), \xi(\omega_g, t_2))$.

Оскільки випадкові процеси, які розглядаються у фіксовані моменти часу, є випадковими величинами, то можна ставити питання про визначення математичного сподівання (середнього значення) та дисперсії випадкового процесу:

$$v_1(t) = \mathbf{M}\{\xi(\omega, t)\}, \quad D(t) = \mathbf{M}\left\{\left(\xi(\omega, t) - v_1(t)\right)^2\right\}. \quad (17.4)$$

Так само, як і у випадку ВВ дисперсія характеризує розкид значень випадкового процесу відносно середнього значення $M(t)$. Зі збільшенням $D(t)$ збільшується ймовірність появи дуже великих за абсолютною величиною значень ВВ. Як і для ВВ для випадкового процесу зручною характеристикою є середнє квадратичне відхилення (СКВ) $\sigma(t) = +\sqrt{D(t)}$. У випадку досліджень фізичних процесів значення СКВ має таку саму розмірність як і ВВ.

Далі в цьому розділі, з метою спрощення аналітичних викладок, в загальному записі випадкових процесів опускаємо аргумент ω .

17.3. Неперервний і дискретний параметр часу

Параметр t може набувати різних значень із деякого інтервалу $[T_1, T_2]$ дійсної прямої. Існує два найбільш важливих випадки:

1. Послідовність значень параметра $t_j \in [T_1, T_2]$, $j = \dots, -2, -1, 0, 1, 2, \dots$ або $j = 0, 1, 2, \dots$ є скінченною або зліченною, тобто область зміни параметра складається з ізольованих точок, як правило цілих чисел (але для часу – іменованих), що являють собою арифметичну прогресію.

2. Другий випадок, коли $t_j \in [T_1, T_2]$, де $[T_1, T_2]$ – деякий скінчений інтервал на дійсній прямій.

Першому випадку відповідають випадкові процеси з дискретним часом – *випадкові послідовності*. Реалізація випадкової послідовності називається *часовим рядом*. Другому випадку відповідають випадкові процеси з неперервним часом – власне *випадкові процеси*.

17.4. Кореляційна теорія та її застосування в технічних системах

Багато тверджень, що використовуються в теорії випадкових процесів, можуть бути сформульовані подвійно – в широкому і строгому (вузькому) розумінні. Про формулювання тих або інших тверджень, властивостей об'єктів дослідження в широкому сенсі говорять у випадках, коли обмежуються дослідженням тільки перших двох моментів (моментних функцій) випадкового процесу $\xi(t)$ – середнього значення і кореляційної функції.

Для реальних (дійсних) процесів використовується модель гільбертового випадкового процесу.

Означення. *Гільбертовим* називається комплекснозначний випадковий процес $\{\xi(t), t \in T\}$, у якого для всіх t виконується умова:

$$\mathbf{M}|\xi(t)|^2 < \infty, \quad t \in T. \quad (17.5)$$

Означення. *Комплекснозначним* називається випадковий процес

$\{\xi(t), t \in T\}$ у якого дійсна та уявна частина суть дійсні процеси:
 $\xi(t) = \xi_1(t) + i\xi_2(t)$, де $i = \sqrt{-1}$.

Означення. Коваріацією $B(t_1, t_2)$, $t_1, t_2 \in T \times T$ гільбертового випадкового процесу $\xi(t)$ називається функція, що визначається наступним чином:

$$B(t_1, t_2) = \mathbf{M}\xi(t_1)\overline{\xi(t_2)}. \quad (17.6)$$

Означення. Кореляційною функцією $R(t_1, t_2)$, $t_1, t_2 \in T \times T$ називається функція, що визначається наступним чином:

$$R(t_1, t_2) = \mathbf{M}[\xi(t_1) - \mathbf{M}\xi(t_1)][\overline{\xi(t_2) - \mathbf{M}\xi(t_2)}]. \quad (17.7)$$

Кореляційну функцію $R(t_1, t_2)$ випадкового процесу $\xi(t)$ часто називають автокореляційною функцією процесу $\xi(t)$.

Характерний вид функції $R(t_1, t_2)$ наведено на рис. 17.2.

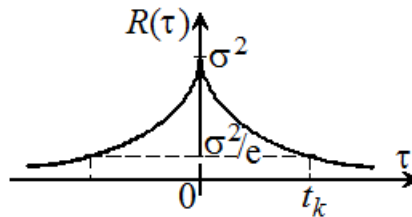


Рис. 17.2. Визначення інтервалу кореляції

Означення. Інтервал часу t_k , на якому кореляційна функція, яка характеризує зв'язок між значеннями випадкового процесу, зменшується в M разів, називається інтервалом (часом) кореляції випадкового процесу.

Ця характеристика визначається доволі суб'єктивно (рис. 17.2). Зазвичай обирають $M = 10$ або $M = e$. Проте вона дає змогу оцінити зв'язок між двома рознесеними в часі значеннями випадкового процесу: на інтервалі кореляції більшому за t_k значення випадкового процесу слабо зв'язані один з одним. Отже знання кореляційної функції дають змогу судити про швидкість зміни випадкового процесу.

Зв'язок між коваріацією та кореляційною функцією дається формулою:

$$B(t_1, t_2) = R(t_1, t_2) + \mathbf{M}\xi(t_1)\mathbf{M}\xi(t_2) = R(t_1, t_2) + v_1(t_1)v_1(t_2). \quad (17.8)$$

Якщо моменти часу співпадають, тобто $t_1 = t_2 = t$, вираз (17.8) набуває виду:

$$B(t, t) = R(t, t) + \nu_1^2(t). \quad (17.9)$$

Всі три елемента в останньому виразі мають самостійний фізичний зміст і характеризують потужність компонент і самого процесу $\xi(t)$ в момент часу t :

$B(t, t)$ – повна потужність процесу,

$R(t, t)$ – потужність флуктацій (змінна складова процесу),

ν_1^2 – потужність постійної складової.

Для двох випадкових процесів $\xi(t)$ і $\eta(t)$ вводиться поняття взаємної кореляційної функції $R_{\xi\eta}(t_1, t_2)$.

Означення. Взаємно кореляційною функцією $R_{\xi\eta}(t_1, t_2)$, $t_1, t_2 \in T_1 \times T_2$ випадкових процесів $\{\xi(t), t \in T_1\}$ і $\{\eta(t), t \in T_2\}$ називається функція, що визначається виразом:

$$R_{\xi\eta}(t_1, t_2) = \mathbf{M}[\xi(t_1) - \mathbf{M}\xi(t_1)][\eta(t_2) - \mathbf{M}\eta(t_2)]. \quad (17.10)$$

17.5. Стаціонарні випадкові процеси

Важливим класом випадкових процесів є *стаціонарні випадкові процеси*, тобто такі процеси які поводять себе більш або менш однаково при зсуві початку відліку в часі. Розрізняють випадкові процеси стаціонарні в широкому і вузькому сенсі.

Означення. Випадковий процес називають *стаціонарним у вузькому сенсі*, коли послідовність його скінченновимірних функцій розподілу не залежить від початку відліку процесу в часі, тобто коли для довільних t_1, \dots, t_n , таких, що $t_k + \tau \in T$, $k = 1, 2, \dots, n$ незалежно від τ , виконується умова:

$$F_{(n)}(x_1, \dots, x_n; t_1 + \tau, \dots, t_n + \tau) = F_{(n)}(x_1, \dots, x_n; t_1, \dots, t_n), \quad n = 1, 2, 3, \dots \quad (17.11)$$

Для такого процесу будь яка характеристика, що визначається сумісним розподілом, наведеним вище, не залежить від τ . Це відноситься до всіх моментів даного розподілу, якщо вони існують.

Означення. *Стационарним в широкому сенсі* називають гільбертів процес $\{\xi(t), t \in T\}$, у якого перші два моменти не залежать від початку відліку, тобто якщо

$$\mathbf{M}\xi(t) = \nu_1 = \text{const};$$

$$\mathbf{M}[\xi(t_1) - \mathbf{M}\xi(t_1)][\xi(t_2) - \mathbf{M}\xi(t_2)] = R(t_2 - t_1) = R(\tau), \quad (17.12)$$

де $R(\tau)$ – неперервна функція по τ .

Кореляційна функція стаціонарного процесу є парною функцією

$$R(\tau) = R(-\tau) \quad (17.13)$$

і задовольняє нерівності

$$R(0) \geq |R(\tau)| \quad (17.14)$$

Із стаціонарності у вузькому сенсі випливає стаціонарність у широкому сенсі. Зворотнє твердження не завжди вірне.

Важливий в практичних застосуваннях *гауссовий стаціонарний в широкому сенсі процес є стаціонарним і у вузькому сенсі.*

Для двох стаціонарних випадкових процесів вводиться поняття стаціонарної зв'язаності.

Означення. Випадкові процеси $\xi(t)$ і $\eta(t)$ називаються *стаціонарно зв'язаними* у вузькому сенсі, якщо сумісна функція розподілу $n + m$ випадкових величин

$$\xi(t_1 + \tau), \dots, \xi(t_n + \tau);$$

$$\eta(t_1 + \tau), \dots, \eta(t_m + \tau)$$

для будь яких $n, m = 1, 2, 3, \dots$ не залежить від τ .

Означення. Стаціонарні в широкому сенсі випадкові процеси $\xi(t)$ і $\eta(t)$ називаються *стаціонарно зв'язаними в широкому сенсі*, якщо їхня взаємна кореляційна функція

$$R_{\xi\eta}(t_1, t_2) = \mathbf{M}[\xi(t_1) - \mathbf{M}\xi(t_1)][\eta(t_2) - \mathbf{M}\eta(t_2)] = R_{\xi\eta}(t_2 - t_1) = R_{\xi\eta}(\tau) \quad (17.15)$$

залежить тільки від різниці аргументів $\tau = t_2 - t_1$.

На відміну від автокореляційної функції стаціонарного процесу взаємна кореляційна функція двох стаціонарно зв'язаних процесів в загальному випадку є непарною і задовольняє нерівності

$$R_{\xi\eta}(-\tau) = R_{\eta\xi}(\tau) \quad (17.16)$$

$$R_{\xi\eta}(0) \geq |R_{\xi\eta}(\tau)|. \quad (17.17)$$

17.6. Енергетичний спектр випадкового процесу

Важливою характеристикою випадкового процесу є його енергетичний спектр. Він визначається як перетворення Фур'є від кореляційної функції

$$G(f) = \int_{-\infty}^{\infty} R(\tau) e^{-i2\pi f\tau} d\tau, \quad (17.18)$$

де f – частота.

Справедливим є і обернене перетворення Фур'є від енергетичного спектру

$$R(\tau) = \int_{-\infty}^{\infty} G(f) e^{i2\pi f\tau} df \quad (17.19)$$

Енергетичний спектр показує розподіл потужності випадкового процесу (до прикладу шумової складової сигналу) по осі частот.

Означення. Випадковий процес з постійним енергетичним спектром, $G(f) = const$, називається *білим шумом*.

Модель випадкового процесу у вигляді білого шуму використовують в аналізі систем управління, засобів вимірювання, контролю та діагностики.

17.7. Ергодичні випадкові процеси

В деяких випадках, наприклад, в апаратурному аналізі випадкових процесів виникає необхідність визначення характеристик досліджуваних явищ усередненням їх реалізацій в часі. Це можливо в аналізі стаціонарних випадкових

процесів, що задовольняють ергодичній гіпотезі.

Оскільки випадковий процес є функцією двох змінних $t \in (-\infty, \infty)$ і $\omega \in \Omega$, то маємо можливість для визначення характеристик процесів, крім усереднення по множині реалізацій (за ймовірнісною мірою, див. рис.17.1), проводити усереднення в часі. Необхідну і достатню умову, за виконання якої можна замінити усереднення по множині реалізацій усередненням в часі, дає ергодична теорема, згідно з якою для виконання рівності (процес $\xi(t)$ –стаціонарний в широкому сенсі)

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) dt = \mathbf{M} \xi(t)^2 \quad (17.20)$$

необхідно і достатньо щоб

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T}^T R(\tau) \left(1 - \frac{\tau}{T}\right) d\tau = 0. \quad (17.21)$$

Процеси, що задовольняють наведеній умові, називаються ергодичними відносно математичного сподівання. Якщо замість процесу $\xi(t)$ розглядати процес $\eta(t, \tau) = \xi(t) \times \xi(t - \tau)$, то у випадку застосування до нього ергодичної теореми для всіх t і τ говорять, що процес $\xi(t)$ ергодичний відносно кореляційної функції.

У виразах (17.20) та (17.21) позначено *l.i.m.* – збіжність у середньому (від англійського limit in mean). Це один з видів збіжності послідовності ВВ, яка вводиться наступним чином: послідовність випадкових величин ξ_1, ξ_2, \dots є такою, що збігається у середньому порядку p , $0 < p < \infty$ до ВВ ξ , якщо

$$\mathbf{M} |\xi_n - \xi|^p \rightarrow 0, \quad n \rightarrow \infty.$$

Для $p=2$ таку збіжність називають збіжністю у середньому квадратичному.

Питання для самоконтролю

1. Дайте визначення дійсного випадкового процесу.
2. Які випадкові процеси називаються сепарабельними?

3. Чим відрізняються випадкові процеси від випадкових послідовностей?
4. Дайте визначення гільбертова комплекснозначного випадкового процесу.
5. Покажіть зв'язок між ковариацією та кореляційною функцією.
6. Запишіть аналітичний вираз для взаємно кореляційної функції двох випадкових процесів.
7. Дайте визначення стаціонарного у вузькому і стаціонарного у широкому сенсі випадкового процесу.
8. В чому полягає властивість ергодичності випадкового процесу і як її можна використати для оброблення сигналів?

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Операційне числення: практикум [Електронний ресурс]: навч. осіб. / В.П. Легеза, Л.М. Олещенко; КПІ ім. Ігоря Сікорського.– Електронні текстові дані (1 файл: 1,4 Мбайт). – К.: КПІ ім. Ігоря Сікорського, 2018. – 70 с.
2. Теоретичні основи інформаційно-вимірювальних систем: Підручник / В.П. Бабак, С.В. Бабак, В.С. Єременко та ін.; під ред. чл.-кор. НАН України В.П. Бабака / – К.: Ун-т новітніх технологій, 2017. – 496 с.
3. Теорія ймовірностей, випадкові процеси та математична статистика / В.П. Бабак, Б.Г. Марченко, М.Є. Фриз. – К.: Техніка, 2004.- 288 с.
4. Статистичний аналіз даних вимірювань: навч. посіб / Єременко В.С., Куц Ю.В., Мокійчук В.М., Самойліченко О.В. –К.:НАУ, 2015.– 321 с.
5. ДСТУ ISO 3534-1:2008 Статистика. Словник термінів і позначки. Частина 1. Загальні статистичні терміни та терміни теорії ймовірностей (ISO 3534-1:2006, IDT) чин. з 01.01.2010. К. : Держспоживстандарт (в електронному вигляді), 42 с. – (Національний стандарт України).
6. Статистика. Словник термінів і позначень. Частина 2. Прикладна статистика (ISO 3534-2: 2006, IDT): ДСТУ ISO 3534-2:2008. – [Чинний від 2010-01-01]. — К. : Держспоживстандарт України, 2008. – 43 с. – (Національний стандарт України).
7. Мартыненко В.С. Операционное исчисление. – К.: Вища школа, 1973.–268 с.
8. Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и 1. Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и математическая статистика. – К.: Вища школа, 1979. – 408 с.
9. Жерновий Ю.В. Збірник задач з теорії ймовірностей та математичної статистики: Для студентів нематематичних спеціальностей. – Львів, 2014. – 18 с.
10. Теорія ймовірностей у задачах автоматизації виробництва: Навчально-методичний посібник з курсу “Спеціальні розділи математики” для студ. спец. „Автоматизоване управління технологічними процесами” / Уклад.: А.І. Жученко, В.В. Миленський, Л.Д. Ярошук. – К.:НТУУ «КПІ», 2008. — 70 с.

11. Жильцов О.Б. Теорія ймовірностей та математична статистика у прикладах і задачах: навч. посіб. для студ. вищ. навч. закл. / О.Б. Жильцов; за ред. Г.О. Михаліна. — К.: Київ. ун-т ім. Б. Грінченка, 2015. — 336 с.

12. Дорожовець М. Опрацювання результатів вимірювань: навч. посібник / М. Дорожовець. — Львів : Видавництво Національного університету “Львівська політехніка”, 2007. — 624 с.

13. Марченко Б.Г., Приймак М.В., Щербак Л.М. Теоретичні основи аналізу стохастичних сигналів і шумів. Навчальний посібник.— Тернопіль: ТДТУ ім. І. Пулюя, 2001. —179 с.

14. Куц Ю.В., Щербак Л.М. Статистична фазометрія.— Тернопіль: ТДТУ ім. І. Пулюя. —2009. —383 с.

15. Forbes C. Statistical Distribution / C. Forbes, M. Evans, N. Hastings, B. Peacock // John Wiley and Sons Inc, New Jersey, 2008. — 202 p.

16. Walk Ch. Hand–book on Statistical Distributions for experimentalist / Ch. Walk // Internal Report, SUF-PFY/ 96-01, Stockholm, 2007.

17. Donnelly R. The Complete Idiot's Guide to Statistics / R.A. Donnelly (Jr.) // Alpha, — 2007, — pp. 395.

18. Технології електромагнітного неруйнівного контролю: Лабораторний практикум. Навчальний посібник / Ю. В. Куц, Ю. Ю. Лисенко; КПІ ім. Ігоря Сікорського. — Київ: КПІ ім. Ігоря Сікорського, 2021. — 63 с.

19. Технології теплового неруйнівного контролю: підручник / А. Г. Протасов, Ю. Ю. Лисенко; КПІ ім. Ігоря Сікорського. — Київ : КПІ ім. Ігоря Сікорського, 2021. — 133 с.

20. Mei, Z.; Kuts, Y.; Kochan, O.; Lysenko, I.; Levchenko, O.; Vlah-Vyhrynovska, H. Using Signal Phase in Computerized Systems of Non-destructive Testing. Meas. Sci. Rev. 2022, 22, 32–43.

21. Застосування фазових характеристик сигналу в автоматизованій вихрострумівій дефектокопії / М. О. Редька, Ю. В. Куц, Є. В. Шаповалов, В. М. Учанін, Ю. Ю. Лисенко, О. Д. Близнюк // Технічна діагностика і неруйнівний контроль. — 2022. — №1. — С. 45-53.

Формули перетворення Лапласа

Оригінал	Зображення
$\eta(t)$	$\frac{1}{p}$
$\eta(t-t_0)$	$\frac{1}{p} e^{-t_0 p}$
$e^{\alpha t}$	$\frac{1}{p-\alpha}$
t^n	$\frac{n!}{p^{n+1}}$
$\sin \alpha t$	$\frac{\alpha}{p^2 + \alpha^2}$
$\cos \alpha t$	$\frac{p}{p^2 + \alpha^2}$
$\sin^2 \alpha t$	$\frac{2\alpha^2}{p(p^2 + 4\alpha^2)}$
$\cos^2 \alpha t$	$\frac{p^2 + 2\alpha^2}{p(p^2 + 4\alpha^2)}$
$\sin(\omega t - \varphi_0)$	$e^{\frac{-\varphi_0}{\omega}} \frac{\omega}{p^2 + \omega^2}$
$\cos(\omega t - \varphi_0)$	$e^{\frac{-\varphi_0}{\omega}} \frac{p}{p^2 + \omega^2}$
$e^{-\alpha t} \sin \omega t$	$\frac{\omega}{(p+\alpha)^2 + \omega^2}$
$e^{-\alpha t} \cos \omega t$	$\frac{\omega}{(p+\alpha)^2 + \omega^2}$
$\delta(t)$	1
$\delta'(t)$	p

Оригінал	Зображення
A	$\frac{A}{p}$
$te^{-\alpha t}$	$\frac{1}{(p+\alpha)^2}$
$t \sin \alpha t$	$\frac{2\alpha p}{(p^2 + \alpha^2)^2}$
$\sin(\omega t + \psi)$	$\frac{p \sin \psi + \omega \cos \psi}{p^2 + \omega^2}$
$e^{i(\omega t + \psi)}$	$\frac{e^{i\psi}}{p - i\omega}$
$1 - e^{-\alpha t}$	$\frac{\alpha}{p(p+\alpha)}$
$(1-t)e^{-\alpha t}$	$\frac{p}{(p+\alpha)^2}$
$f(t-\tau)$	$e^{-p\tau} F(p)$
$\frac{\sin \alpha t}{t}$	$\frac{\pi}{2} - \operatorname{arctg} \frac{p}{\alpha}$
$\sin \alpha n$	$\frac{e^p \sin \alpha}{e^{2p} - 2e^p \cos \alpha + 1}$
$\cos \alpha n$	$\frac{e^p (e^p - \cos \alpha)}{e^{2p} - 2e^p \cos \alpha + 1}$
$a^{\alpha n}$	$\frac{e^p}{e^p - a^\alpha}$

Властивості перетворення Лапласа

Властивість	Оригінал	Зображення
Лінійність	$\alpha f_1(t) + \beta f_2(t)$	$\alpha F_1(p) + \beta F_2(p)$
Подібність	$f(\alpha t)$	$\frac{1}{\alpha} F\left(\frac{p}{\alpha}\right)$
Зсуення	$e^{\alpha t} f(t)$	$F(p - \alpha)$
Запізнювання	$f(t - \tau)$	$e^{-p\tau} F(p)$
Диференціювання оригіналу	$f^{(n)}(t)$	$p^n F(p) - p^{n-1} f(0) - \dots - f^{(n-1)}(0)$
Диференціювання зображення	$(-1)^n t^n f(t)$	$F^{(n)}(p)$
Інтегрування оригіналу	$\int_0^t f(\tau) d\tau$	$\frac{F(p)}{p}$
Інтегрування зображення	$\frac{f(t)}{t}$	$\int_p^{+\infty} F(p) dp$

Значення квантилів розподілу Стюдента $t_p(v)$ з v ступенями свободи

v	довірча імовірність P ($P=1-\alpha$, де α – рівень значущості)									
	0,6	0,65	0,7	0,75	0,8	0,85	0,9	0,95	0,975	0,99
1	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,71	31,82
2	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965
3	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541
4	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747
5	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365
6	0,265	0,404	0,543	0,718	0,906	1,134	1,440	1,943	2,447	3,143
7	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998
8	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896
9	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821
10	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764
11	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718
12	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681
13	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650
14	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624
15	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602
16	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583
17	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567
18	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552
19	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539
20	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528
25	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485
30	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457
40	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423
60	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390
120	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358
∞	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326

Примітки. Квантилі рівня $\alpha = 0,5$ при будь-якому v рівні нулю.

Квантилі рівня $P < 0,5$ визначаються за формулою:

$$t_P(v) = -t_{1-P}(v). \quad (\text{ДЗ.1})$$

Для проміжних значень P , що лежать між двома сусідніми табличними значеннями P_1 і P_2 , тобто $P_1 < P < P_2$, значення квантиля $t_P(v)$ може бути визначено методом лінійної інтерполяції:

$$t_P(v) = (P - P_1) \frac{t_{P_2} - t_{P_1}}{P_2 - P_1} + t_{P_1}. \quad (\text{ДЗ.2})$$

ДОДАТОК 4

Значення квантилів розподілу χ^2 Пірсона
для ν ступенів свободи та рівня імовірності P

ν	P									
	0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995
1	$3,9 \cdot 10^{-5}$	$1,5 \cdot 10^{-4}$	0,001	0,004	0,016	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,042	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	10,085	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	10,865	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	11,651	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	13,240	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	14,041	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	14,848	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	15,659	33,196	36,415	39,364	42,980	45,559
25	10,520	11,524	13,120	14,611	16,473	34,382	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	35,563	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	18,114	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	37,916	41,337	44,461	48,278	50,993
29	13,121	14,256	16,047	17,708	19,768	39,087	42,557	45,722	49,588	52,336
30	13,787	14,953	16,791	18,493	20,599	40,256	43,773	46,979	50,892	53,672
31	14,458	15,655	17,539	19,281	21,434	41,422	44,985	48,232	52,191	55,003
32	15,134	16,362	18,291	20,072	22,271	42,585	46,194	49,480	53,486	56,328
33	15,815	17,074	19,047	20,867	23,110	43,745	47,400	50,725	54,776	57,648

продовження таблиці Додатка 4

<i>v</i>	<i>P</i>									
	0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995
34	16,501	17,789	19,806	21,664	23,952	44,903	48,602	51,966	56,061	58,964
35	17,192	18,509	20,569	22,465	24,797	46,059	49,802	53,203	57,342	60,275
36	17,887	19,233	21,336	23,269	25,643	47,212	50,998	54,437	58,619	61,581
37	18,586	19,960	22,106	24,075	26,492	48,363	52,192	55,668	59,892	62,883
38	19,289	20,691	22,878	24,884	27,343	49,513	53,384	56,896	61,162	64,181
39	19,996	21,426	23,654	25,695	28,196	50,660	54,572	58,120	62,428	65,476
40	20,707	22,164	24,433	26,509	29,051	51,805	55,758	59,342	63,691	66,766
41	21,421	22,906	25,215	27,326	29,907	52,949	56,942	60,561	64,950	68,053
42	22,138	23,650	25,999	28,144	30,765	54,090	58,124	61,777	66,206	69,336
43	22,859	24,398	26,785	28,965	31,625	55,230	59,304	62,990	67,459	70,616
44	23,584	25,148	27,575	29,787	32,487	56,369	60,481	64,201	68,710	71,893
45	24,311	25,901	28,366	30,612	33,350	57,505	61,656	65,410	69,957	73,166
46	25,041	26,657	29,160	31,439	34,215	58,641	62,830	66,617	71,201	74,437
47	25,775	27,416	29,956	32,268	35,081	59,774	64,001	67,821	72,443	75,704
48	26,511	28,177	30,755	33,098	35,949	60,907	65,171	69,023	73,683	76,969
49	27,249	28,941	31,555	33,930	36,818	62,038	66,339	70,222	74,919	78,231
50	27,991	29,707	32,357	34,764	37,689	63,167	67,505	71,420	76,154	79,490
100	67,328	70,06	74,22	77,93	82,358	118,49	124,34	129,56	135,80	140,16