

І.І. Горбань

Теорія ймовірностей
і
математична статистика

для наукових працівників та інженерів

**Національна академія наук України
Інститут проблем математичних машин і систем**

519.2

І.І. Горбань

**Теорія ймовірностей і математична
статистика**
для наукових працівників та інженерів

**Київ
2003**

УДК 519.2+600.1

Теорія ймовірностей і математична статистика для наукових працівників та інженерів / Горбань І.І. – Київ, 2003. – С. 244: іл. – Бібліогр.: 114 назв.

Книга охоплює основні питання теорії ймовірностей і математичної статистики, з якими часто зустрічаються на практиці наукові працівники та інженери різних спеціальностей.

У книзі викладені основи теорії випадкових величин і випадкових функцій, кореляційної теорії, марковських випадкових процесів, випадкових течій, вибіркової теорії, оцінки параметрів і законів розподілу випадкових величин, регресійного аналізу, теорії перевірки статистичних гіпотез, оцінки характеристик випадкових функцій, а також питання дисперсійного аналізу, кластерного аналізу і робастних статистичних методів. Матеріал представлено у конспективному стилі: на базі упорядкованих визначень, теорем, положень, наслідків, зауважень, попереджень, приміток та прикладів.

Для читача достатньо знань вищої математики в об'ємі типових курсів вищих технічних навчальних закладів.

Рецензенти:

А.О. Морозов, чл.-кор. НАН України, доктор технічних наук, професор

М.С. Сівов, заслужений діяч науки і техніки України, доктор технічних наук, професор

С.Я. Жук, доктор технічних наук, професор

Рекомендована до видання Вченою Радою Інституту проблем математичних машин і систем НАН України

ISBN 966-02-2664-0

© Горбань І. І., Київ, 2003

Вступ

Монографія, що пропонується, присвячена прикладним питанням теорії ймовірностей та математичної статистики. Вона розрахована на наукових працівників та інженерів, які використовують ймовірнісні методи для розв'язання різноманітних прикладних завдань.

Орієнтація на таку специфічну категорію читачів обумовила стиль викладання матеріалу:

- 1) книга складається з упорядкованого набору визначень, теорем, положень, наслідків, зауважень, приміток, прикладів та значної кількості систематизуючих таблиць;
- 2) у книзі практично відсутні доведення теорем і положень. У деяких випадках замість них наведені коментарі, що дають уявлення про ідею доведення;
- 3) у книгу не увійшли теоретичні та вузько спеціалізовані питання;
- 4) матеріал викладено з урахуванням знань, що отримують студенти вищих технічних навчальних закладів при вивченні математичного аналізу, теорії матриць і теорії функцій комплексної змінної.

Книга складається з одинадцяти розділів та чотирнадцяти додатків. Перші п'ять розділів містять основи теорії ймовірностей, а останні сім – основи математичної статистики.

Перший розділ присвячений викладанню основних положень теорії випадкових величин, другий – загальним положенням теорії випадкових функцій, третій – питанням кореляційної теорії, що описує випадкові функції у рамках моментів двох перших порядків. До четвертого розділу увійшли питання теорії марковських випадкових процесів, яка базується на припущенні, що одновимірні ймовірнісні характеристики у фіксований момент часу залежать лише від значення випадкової функції у попередній момент часу й не залежать від того, що було раніше. У п'ятому розділі описуються основи теорії випадкових течій, випадковість яких проявляється у випадкових часових моментах появи подій. Шостий розділ присвячений загальним положенням вибіркової теорії, що є базою для усіх питань математичної статистики, сьомий – методології оцінки параметрів та законів розподілу вибірових величин, а восьмий – регресійному аналізу, який забезпечує опис одних випадкових величин та функцій за допомогою інших величин і функцій. У дев'ятому розділі розглянуті питання теорії перевірки статистичних гіпотез, яка є теоретичною базою для синтезу оптимальних алгоритмів прийняття рішень за результатами спостережень випадкових величин і випадкових функцій та для оцінки характеристик випадкових величин. Десятий розділ присвячений різним оцінкам характеристик випадкових функцій: математичного сподівання, дисперсії та інших.

В останньому розділі книги наведені відомості про деякі сучасні напрямки розвитку математичної статистики, а саме: дисперсійний аналіз, що є ефективним інструментом для виявлення факторів, від яких залежать випадкові величини, кластерний аналіз, що дозволяє виявити особливості структури ймовірнісних характеристик випадкових величин, а також методи отримання регуляризованих та робастних оцінок параметрів випадкових величин, що стійкі до похибок.

У додатки винесені окремі теоретичні питання, що не вписуються у загальну канву викладання основного матеріалу, опис деяких спеціальних функцій,

випадкових величин, а також розрахункові математичні таблиці.

До списку літератури включені деякі видання з теорії випадкових функцій, математичної статистики та прикладних питань. Кожен розділ завершується таблицею, яка має допомогти читачеві у підборі додаткової літератури відповідно до його вимог.

Автор висловлює глибоку вдячність офіційним рецензентам – чл.-кор. НАН України д.т.н. професору Морозову А.О., заслуженому діячу науки і техніки України д.т.н професору Сівову М.С. і д.т.н професору Жуку С.Я., а також неофіційним рецензентам – академіку НАН України д.т.н та д.ф.-м.н. професору Коваленко І.М. і д.ф.-м.н. професору Кнопову П.С., які знайшли час уважно ознайомитися з рукописом і внесли цінні зауваження та рекомендації щодо його покращення.

* * *

Книга, що пропонується, є відкоригованим варіантом монографії 2003 р., в якій внесено виправлення, що були виявлені після друку.

З відгукami та зауваженнями можна звертатися до автора за адресою: Проспект Глушкова, 42, Київ, 03187, Україна, Інститут проблем математичних машин і систем НАН України, або електронною поштою: igor.gorban@yahoo.com.

1. Основні положення теорії ймовірностей

1.1. Математичне поняття ймовірності

Існує декілька математичних визначень поняття ймовірності. Найбільш відомі класичне, статистичне і аксіоматичне визначення. Розглянемо їх.

1.1.1. Класична та статистична ймовірності

Визначення 1. *Випадковим* називається явище, що залежить від випадку.

Зауваження. Характерною рисою випадкового явища є те, що результати його спостереження неможливо передбачити заздалегідь абсолютно точно.

Визначення 2. *Однаковими умовами* при спостереженні випадкового явища називають фіксовані стани всіх факторів (зв'язків, величин, характеристик, параметрів та ін.), що контролюються.

Зауваження 1. При спостереженні випадкового явища в однакових умовах стани всіх факторів, що не контролюються, вважають *нефіксованими*, тобто такими, що можуть довільно змінюватися.

Зауваження 2. Наявність нефіксованих факторів та можливість їх зміни обумовлює випадковість випадкового явища.

Визначення 3. Спостереження випадкового явища називається *дослідом*, а дослід в однакових умовах – *іспитами*.

Визначення 4. *Елементарною подією* ω називається будь-який можливий результат (наслідок) дослідів. Сукупність Ω елементарних подій утворює *простір елементарних подій* або *простір вибірок*. Кожна точка цього простору – елементарна подія.

Зауваження. Простір елементарних подій може бути довільного типу. У найпростішому випадку – це числова множина, множина символів або слів. У загальному ж випадку – це множина, що складається з об'єктів будь-якого типу.

Приклад 1. Множина оцінок, що може одержати студент на іспиті, утворює простір елементарних подій. У даному випадку його можна розглядати як множину, що складається з чотирьох слів (незадовільно, задовільно, добре, відмінно), або як числову множину, елементи якої – чотири числа (2, 3, 4, 5).

Визначення 5. *Випадковою подією (подією)* називається довільний факт, який може статися у результаті дослідів.

Зауваження 1. Випадкові події відбуваються, коли відбуваються елементарні події. Кожна подія A пов'язана з настанням елементарних подій ω з деякої сукупності Ω_A . Про елементарні події ω , що входять до сукупності Ω_A , говорять, що вони *сприяють* події A .

Зауваження 2. В загальному випадку сукупності Ω_A і Ω_B елементарних подій ω , що сприяють різним подіям A і B , можуть перетинатися.

Положення 1. В ряді випадків елементарні події можна вважати несумісними та рівноможливими.

Несумісність означає, що у результаті дослідів може статися одна і тільки одна елементарна подія, а *рівноможливість* – що можливість настання елементарних подій однакова.

Особливості несумісності та рівноможливості покладені до основи так званого *класичного визначення ймовірності*, що визначається для *скінченного* числа елементарних подій.

Визначення 6. У класичному розумінні *ймовірністю* $p(A)$ події A називається відношення числа L_A несумісних рівноможливих елементарних подій ω , що сприяють події A , до загального числа L елементарних подій:

$$p(A) = \frac{L_A}{L}. \quad (1.1.1)$$

Приклад 2. Ймовірність $p(A)$ того, що при киданні гральної кості випаде число від одиниці до чотирьох включно, є $\frac{4}{6} = \frac{2}{3}$.

Ймовірність може бути визначена і на базі результатів спостережень. Цей підхід називається *статистичним*.

Визначення 7. Частотою $p_N(A)$ події A в N іспитах називається відношення числа випадків N_A , в яких сталася подія A , до загального числа іспитів:

$$p_N(A) = \frac{N_A}{N}. \quad (1.1.2)$$

Положення 2. В багатьох випадках (але не завжди) зі збільшенням кількості іспитів N частота події $p_N(A)$ стабілізується, тобто прямує до деякої фіксованої величини $p(A)$.

Визначення 8. *Статистичною ймовірністю* називається величина $p(A)$, до якої прямує частота $p_N(A)$ події A при $N \rightarrow \infty$:

$$p(A) = \lim_{N \rightarrow \infty} p_N(A). \quad (1.1.3)$$

Зауваження. На практиці кількість іспитів завжди обмежена, тому розрахувати на базі експериментів статистичну ймовірність принципово неможливо. Можливо лише отримати більш чи менш точну оцінку цієї ймовірності.

Положення 3. Найбільш коректне з математичної точки зору є аксіоматичне визначення ймовірності, що було запропоновано у 30-х роках ХХ століття А.М. Колмогоровим. Цей підхід базується на сучасних теоріях множини і міри. Тому розглянемо спочатку основні положення цих теорій.

1.1.2. Основні положення теорії множин

Визначення 1. *Множиною* називають сукупність об'єктів довільної природи. Ці об'єкти називаються *елементами множини*.

Зауваження. Наведене визначення не є математично строгим. У математиці це початкове поняття, що вводиться аксіоматичним чином.

Визначення 2. *Підмножиною* A множини Ω називається така множина, усі елементи якої є елементами множини Ω . При цьому пишуть $A \subset \Omega$ або $\Omega \supset A$.

Зауваження. По визначенню пуста множина ϕ є підмножиною кожної множини.

Визначення 3. Дві множини A і B називаються *рівними*, якщо вони мають однакові елементи ($A \subset B$, $A \supset B$).

Визначення 4. *Скінченною множиною* називається множина, кількість елементів якої скінченна; *нескінченною множиною* називається множина, кількість елементів якої нескінченна; *зліченною множиною* називається множина, елементи якої можна перелічити, *незліченною множиною* називається множина, елементи якої перелічити не можна.

Дискретною множиною називається множина, що складається з кінцевої або зліченної кількості елементів.

Приклади 1. Сукупність цифри одиниця, числа π , знаків ∇ і $*$ утворюють скінченну множину, що має чотири елементи. Сукупності цілих додатних чисел та просто цілих чисел утворюють злічені множини. Сукупності точок на прямій, на відрізку або у просторі утворюють незлічені (*континуальні*) множини.

Положення. Для множин визначають операції *об'єднання* \cup (*логічної суми* $+$), *перетину* \cap (*логічного добутку* \times (\cdot)) і *доповнення* \setminus ($-$).

Об'єднання $A = A_1 \cup A_2$ означає, що множина A створена з усіх елементів, які входять тільки до множини A_1 , або тільки до множини A_2 , або одночасно і до множини A_1 , і до множини A_2 .

Перетин $A = A_1 \cap A_2$ означає, що множина A створена з усіх елементів, які входять одночасно і до множини A_1 , і до множини A_2 .

Доповнення $B = \bar{A} = \Omega \setminus A$, де $A \subset \Omega$, означає, що множина B створена з усіх елементів множини Ω , які не входять до множини A .

Зауваження. Операції об'єднання і перетину узагальнюють на випадок скінченної та нескінченної (зліченної й незліченної) множин.

Нехай $s \in S$, де S – скінченна, зліченна або незліченна множина, A_s – множина, що характеризується індексом s . Тоді об'єднання $\bigcup_{s \in S} A_s$ є множина, яка створюється з усіх елементів, що входять хоча би до однієї множини A_s , а перетин $\bigcap_{s \in S} A_s$ – є множина, що створюється з усіх елементів, що входять одночасно до всіх множин A_s .

Приклад 2. Нехай A_s – множина точок у прямокутнику $s \leq x \leq s+1$, $0 \leq y \leq 1$, де $0 \leq s \leq 0,5$ (тобто $s \in S = [0, 0,5]$) (рис. 1.1.1). Тоді $\bigcup_{s \in S} A_s$ – множина точок, що входять до прямокутника $0 \leq x \leq 1,5$, $0 \leq y \leq 1$, а $\bigcap_{s \in S} A_s$ – множина точок, що входять до прямокутника $0,5 \leq x \leq 1$, $0 \leq y \leq 1$.

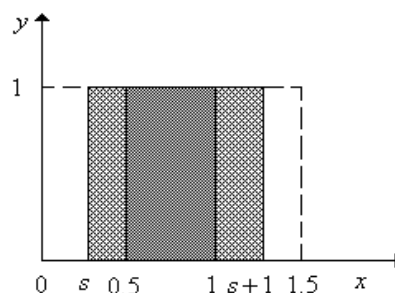


Рис. 1.1.1. Ілюстрація приклада

Операції об'єднання, перетину і доповнення є базовими для формування алгебри множин.

Визначення 5. *Алгеброю множин* називають клас підмножин Ξ множини Ω , що задовольняє наступним умовам:

- 1) $\Omega \in \Xi$;

- 2) $\phi \in \Xi$;
- 3) якщо $A \in \Xi$, то $\bar{A} \in \Xi$;
- 4) якщо $A_1 \in \Xi$ і $A_2 \in \Xi$, то $A_1 \cup A_2 \in \Xi$ і $A_1 \cap A_2 \in \Xi$.

1.1.3. Основи теорії міри

Визначення 1. Мірою називається така функція $\mu(A)$ підмножин A не пустої множини Ω , що для всіх A , для котрих визначена міра, $\mu(A) \geq 0$, $\mu(\phi) = 0$ та для підмножин A_n ($n = 1, 2, \dots$), що попарно не перетинаються,

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n). \quad (1.1.4)$$

Зауваження 1. Міра є математичним поняттям, що узагальнює поняття довжини, площі, об'єму, маси, енергії та ін. величин, які кількісно характеризують підмножини деякої множини.

Зауваження 2. Умова (1.1.4) називається аксіомою зліченної адитивності.

Визначення 2. Міра $\mu(\Omega)$ множини Ω називається зосередженою на підмножині $A \subset \Omega$, якщо $\mu(\Omega \setminus A) = \mu(\bar{A}) = 0$.

Визначення 3. Міра $\mu(\Omega)$ множини Ω називається дискретною, якщо вона зосереджена на скінченній або зліченній множині точок $\omega_n \in \Omega$ ($n = 1, 2, \dots, N$ або $n = 1, 2, \dots$) і неперервною (недискретною), якщо вона зосереджена на континуальній множині точок.

Зауваження 1. Недискретна міра на скінченній множині точок дорівнює нулю.

Приклад 1. Площа (тобто міра) точки на площині дорівнює нулю, але площа континуальної множини точок відмінна від нуля.

Зауваження 2. Для довільної міри μ вимагають, щоб вона була визначена для всіх A деякого класу підмножин множини Ω , що створюють так звану σ -алгебру (сігму-алгебру).

Визначення 4. σ -алгеброю називають клас \mathfrak{F} підмножин множини Ω з визначеними для них операціями об'єднання, перетину і доповнення, що задовольняють таким вимогам:

$$1) \quad \Omega \in \mathfrak{F}; \quad (1.1.5)$$

$$2) \quad \text{при } A \in \mathfrak{F} \Rightarrow \bar{A} = \Omega \setminus A \in \mathfrak{F}; \quad (1.1.6)$$

$$3) \quad \text{при } A_n \in \mathfrak{F}, n = 1, 2, \dots \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{F}. \quad (1.1.7)$$

Зауваження 1. Потужність множини \mathfrak{F} може бути більшою ніж потужність зліченної множини.

Зауваження 2. З аксіом (1.1.5) – (1.1.7) випливає, що

$$\text{при } A_n \in \mathfrak{F}, n = 1, 2, \dots \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathfrak{F}. \quad (1.1.8)$$

Зауваження 3. Пуста множина входить до \mathfrak{F} (це випливає з виразів (1.1.5), (1.1.6)).

Зауваження 4. Вирази (1.1.5) – (1.1.8) відображають той факт, що при будь-яких об'єднаннях, перетинах та доповненнях підмножин $A_n \in \mathfrak{F}$ ($n = 1, 2, \dots$) отримуються множини, що входять до множини \mathfrak{F} , тобто операції об'єднання, перетину та доповнення довільних підмножин множини \mathfrak{F} не виводять за границі множини \mathfrak{F} .

Зауваження 5. Множина Ω може породжувати множину різних σ -алгебр.

Зауваження 6. До σ -алгебри \mathfrak{F} множини Ω можуть входити, а можуть і не входити елементи множини Ω .

Зауваження 7. Серед різних σ -алгебр множини Ω завжди існує *найменша* σ -алгебра.

Визначення 5. *Вимірним простором* називається множина, яка задається не пустою множиною Ω та σ -алгеброю його підмножин \mathfrak{F} (тобто парою (Ω, \mathfrak{F})). При цьому підмножини множини \mathfrak{F} називаються *вимірними множинами*.

Зауваження. Міра задається для всіх елементів вимірних множин \mathfrak{F} , тобто для таких, що створюють σ -алгебру.

Визначення 6. *Простором з мірою (полем)* називається вимірний простір (Ω, \mathfrak{F}) , для якого задана міра μ . Таким чином, простір з мірою задається трійкою $(\Omega, \mathfrak{F}, \mu)$.

Визначення 7. Міра μ називається *скінченною*, якщо $\mu(\Omega) < \infty$, і *нормованою*, якщо $\mu(\Omega) = 1$.

Зауваження. Для скінченної міри $\mu(A) \leq \mu(\Omega)$ для всіх $A \in \mathfrak{F}$.

Визначення 8. Нехай Ω – M -вимірний евклідовий простір R^M , K – клас напіввідкритих паралелепіпедів $\prod_{m=1}^M [a_m, b_m)$, $-\infty < a_m < b_m < \infty$ ($m = \overline{1, M}$). Тоді σ -алгеброю *борелівських множин в R^M (борелівським полем)* називається найменша σ -алгебра, що містить клас K .

1.1.4. Аксиоматичне визначення ймовірності

Положення. При аксіоматичному визначенні ймовірності

- 1) задається *простір елементарних подій Ω з елементарними подіями $\omega \in \Omega$* ;
- 2) задається *найменша σ -алгебра \mathfrak{F} (борелівське поле) підмножин, які називаються подіями*;
- 3) кожній події A ставиться відповідна нормована міра $p(A)$, що називається *ймовірнісною мірою (розподілом ймовірностей)* або, скорочено, *ймовірністю*.

Зауваження. Простір з ймовірнісною мірою називається *ймовірнісним простором*. Ймовірнісний простір, як і любий простір з мірою, задається трійкою $(\Omega, \mathfrak{F}, p)$.

Визначення 1. *Достовірною подією I простору Ω називають подію, що має місце при всіх елементарних подіях $\omega \in \Omega$.*

Визначення 2. *Неможливою подією називається подія, що не настає ні при якому $\omega \in \Omega$.*

Визначення 3. Події A_1 і A_2 називаються *несумісними*, якщо $A_1 \cap A_2 = \phi$.

Визначення 4. Події A_1, A_2, \dots, A_I називаються *попарно несумісними*, якщо $A_i \cap A_j = \phi$ при довільних $i \neq j$ ($i, j = \overline{1, I}$).

Аксиоми. Як нормована міра ймовірність $p(A)$ задається такими трьома аксіомами:

- 1) $p(A) \geq 0$;
- 2) $p(\bigcup_n A_n) = \sum_n p(A_n)$ для *попарно несумісних* подій (як скінченних, так і зліченних);

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

Зауваження 1. З наведених аксіом випливає, що $0 \leq p(A) \leq 1$ і $p(\phi) = 0$.

Зауваження 2. З рівності $p(A) = 1$ не випливає, що A – достовірна подія, а з рівності $p(B) = 0$ не випливає, що B – неможлива подія.

Зауваження 3. У разі сумісних подій A_1 і A_2 ймовірність $p(A_1 \cap A_2)$ події $A_1 \cap A_2$ визначається (четверта аксіома) таким чином з ймовірністю $p(A_1) \neq 0$ події A_1 і умовною ймовірністю $p(A_2 / A_1)$ події A_2 при умові, що настала подія A_1 :

$$p(A_1 \cap A_2) = p(A_1)p(A_2 / A_1). \quad (1.1.9)$$

Примітка. Ймовірність $p(A_2 / A_1)$ не визначена, якщо $p(A_1) = 0$. У подальшому всюди, де зустрічається умовна ймовірність, вважається, що ця умова виконується

Зауваження 4. З аксіом випливають два положення, що називаються аксіомами неперервності:

а) якщо $A_n \subset A_{n+1}$, $n \geq 1$, то $p(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} p(A_n)$;

б) якщо $A_{n+1} \subset A_n$, $n \geq 1$, то $p(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} p(A_n)$.

Визначення 5. Події A_1 і A_2 називаються незалежними (незалежними по ймовірності), якщо $p(A_1 \cap A_2) = p(A_1)p(A_2)$. При цьому $p(A_2 / A_1) = p(A_2)$, якщо $p(A_1) \neq 0$, і $p(A_1 / A_2) = p(A_1)$, якщо $p(A_2) \neq 0$.

Зауваження. З наведеного визначення випливає, що події A_1 і A_2 незалежні, якщо поява однієї з них не викликає зміни ймовірності появи іншої події.

Визначення 6. Події A_1, A_2, \dots, A_M називаються незалежними у сукупності, якщо $p(A_1 \cap A_2 \cap \dots \cap A_M) = p(A_1)p(A_2)\dots p(A_M)$.

Зауваження. З незалежності подій у сукупності випливає попарна незалежність подій. Зворотне твердження у загальному випадку невірно.

Визначення 7. Повною групою несумісних подій називається набір несумісних подій A_1, A_2, \dots, A_M , для яких $A_1 \cup A_2 \cup \dots \cup A_M = \Omega$.

1.1.5. Основні теореми і формули теорії ймовірностей

Теорема 1. Нехай A_1, A_2, \dots, A_M – повна група несумісних подій. Тоді сума їх ймовірностей дорівнює одиниці:

$$\sum_{m=1}^M p(A_m) = 1. \quad (1.1.10)$$

Зауваження 1. Теорема узагальнюється на нескінченний випадок ($M \rightarrow \infty$).

Зауваження 2. Оскільки протилежні події несумісні, сума їх ймовірностей дорівнює одиниці:

$$p(A) + p(\bar{A}) = 1. \quad (1.1.11)$$

Теорема 2 (додавання). Нехай A і B – довільні події (в загальному випадку сумісні). Тоді ймовірність суми цих подій

$$p(A \cup B) = p(A) + p(B) - p(A \cap B). \quad (1.1.12)$$

Теорема 3. Нехай A_1, A_2, \dots, A_M – події, що сумісні в загальному випадку. Тоді

а) ймовірність суми цих подій

$$p\left(\bigcup_{m=1}^M A_m\right) = \sum_{m=1}^M p(A_m) - \sum_{m,n=1}^M p(A_m \cap A_n) + \dots + (-1)^{M-1} p(A_1 \cap \dots \cap A_M); \quad (1.1.13)$$

б) ймовірність добутку цих подій

$$p\left(\bigcap_{m=1}^M A_m\right) = \sum_{m=1}^M p(A_m) - \sum_{m,n=1}^M p(A_m \cup A_n) + \dots + (-1)^{M-1} p(A_1 \cup \dots \cup A_M). \quad (1.1.14)$$

Теорема 4 (добутку). Нехай A_1, A_2, \dots, A_M – залежні в загальному випадку M подій. Тоді ймовірність добутку цих подій

$$p\left(\bigcap_{m=1}^M A_m\right) = p(A_1)p(A_2 / A_1)\dots p(A_M / A_1 \cap \dots \cap A_{M-1}). \quad (1.1.15)$$

Зауваження. Якщо події незалежні в сукупності, то

$$p\left(\bigcap_{m=1}^M A_m\right) = \prod_{m=1}^M p(A_m),$$

(дивись визначення в підрозділі 1.1.4).

Теорема 5 (формула повної ймовірності). Нехай подія A може статися сумісно з однією і тільки однією подією з числа H_1, H_2, \dots, H_M , що утворюють повну групу несумісних подій (гіпотез). Тоді ймовірність $p(A)$ події A може бути розрахована як

$$p(A) = \sum_{m=1}^M p(H_m) p(A / H_m). \quad (1.1.16)$$

Теорема 6 (Байеса або гіпотез). Нехай H_1, H_2, \dots – послідовність попарно несумісних подій (гіпотез), що утворюють повну групу. Тоді для кожної пари подій H_m, A має місце формула Байеса:

$$p(H_m / A) = \frac{p(H_m \cap A)}{p(A)} = \frac{p(H_m)p(A / H_m)}{\sum_{n=1}^{\infty} p(H_n)p(A / H_n)}. \quad (1.1.17)$$

Визначення. Досліди називаються *незалежними*, якщо ймовірність того чи іншого результату досліду не залежить від результатів інших дослідів.

Теорема 7 (часткова теорема про повторення дослідів). Проводиться N незалежних дослідів в однакових умовах (N незалежних іспитів). В результаті іспиту може статися або не статися подія A . Ймовірність появи події A в кожному іспиті дорівнює p , а ймовірність не появи – $q = 1 - p$. Тоді ймовірність того, що в N іспитах подія A станеться n раз, дорівнює

$$P_N(n) = C_N^n p^n q^{N-n}, \quad (1.1.18)$$

де C_N^n – число об'єднань із N по n : $C_N^n = \frac{N!}{n!(N-n)!}$.

Примітка. Розподіл ймовірностей (1.1.18) називається *біномним*.

Теорема 8 (загальна теорема про повторення дослідів). Проводиться N незалежних дослідів, у кожному з котрих може статися або не статися подія A . Ймовірність появи події A в m -му досліді дорівнює p_m , а ймовірність не появи – $q_m = 1 - p_m$ ($m = \overline{1, N}$). Тоді ймовірність того, що в результаті N дослідів подія A станеться n раз, може бути знайдена з такого рівняння

$$\sum_{n=0}^N P_N(n) z^n = \prod_{m=1}^N (q_m + p_m z).$$

1.2. Скалярні випадкові величини

1.2.1. Основні визначення

Визначення 1. Випадковою величиною X називається довільна детермінована числова вимірна функція, що визначена на просторі Ω елементарних подій ω . При цьому значення x випадкової величини X може бути представлено у вигляді деякої функції $x = \psi(\omega)$, де $\omega \in \Omega$. Множина значень випадкової величини утворює простір значень випадкової величини.

Зауваження 1. У теорії випадкових величин вважають, що до завершення досліду невідомо яка саме елементарна подія станеться, а тому і якого значення x набуде випадкова величина X .

Зауваження 2. Випадкову величину звичайно задають не тільки простором значень випадкової величини, а і характеристиками, які характеризують ймовірність появи тих чи інших значень з цього простору.

Коли для випадкової величини заданий як простір значень випадкової величини, так і ймовірності появи цих значень говорять, що випадкова величина ймовірнісно визначена.

Зауваження 3. В окремому випадку функція $x = \psi(\omega)$ може мати вигляд $x = \omega$. При цьому випадковою величиною X є така змінна, значення якої утворюють простір Ω елементарних подій ω , а простір значень випадкової величини співпадає з простором елементарних подій.

Зауваження 4. Надалі випадкові величини переважно будемо позначати прописними літерами, а значення, що вони приймають, – рядковими.

Положення 1. Випадкові величини класифікують за різними ознаками. Одними з поширених ознак класифікації є тип простору значень випадкової величини (дійсний простір, комплексний, неперервний, дискретний і т.д.) і вимірність цього простору.

Зауваження. У подальшому всюди, де не оговорено протилежне, вважається, що простір значень випадкової величини – дійсний.

Визначення 3. Якщо простір значень випадкової величини має один вимір, випадкова величина називається скалярною, а якщо багато вимірів, випадкова величина називається векторною.

Зауваження. Надалі векторні випадкові величини і їхні значення будемо позначати літерами зі стрілкою над ними.

Положення 2. Випадкові величини описують за допомогою різних характеристик. Найбільш повний опис дають ймовірнісні характеристики, менш повне – неймовірнісні (числові) характеристики. Познайомимося з обома типами характеристик на прикладі скалярних випадкових величин.

1.2.2. Ймовірнісні характеристики випадкової величини

Положення 1. У скалярному (одновимірному) випадку випадкову величину X найбільш повно і точно описують такі три ймовірнісні характеристики:

1) функція розподілу ймовірностей

$$F(x) = P\{X < x\}; \quad (1.2.1)$$

2) щільність розподілу ймовірностей

$$f(x) = \frac{dF(x)}{dx}; \quad (1.2.2)$$

3) характеристична функція, що є математичним сподіванням комплексної випадкової величини $e^{j\omega x}$, або при іншому трактуванні – перетворенням Фур'є щільності розподілу ймовірностей:

$$Q(j\omega) = \int_{-\infty}^{\infty} f(x)e^{j\omega x} dx, \quad (1.2.3)$$

де $P\{*\}$ – ймовірність умови; j – уявна одиниця.

Примітка. Визначення поняття математичного сподівання дано у наступному підрозділі.

Зауваження 1. Функція розподілу ймовірностей $F(x)$ і характеристична функція $Q(j\omega)$ – функції, що не мають розмірності, а щільність розподілу ймовірностей $f(x)$ має розмірність, яка обернена розмірності випадкової величини.

Зауваження 2. Кожна випадкова величина X однозначно визначає функцію розподілу $F(x)$, але кожна функція розподілу $F(x)$ *неоднозначно* визначає випадкову величину.

Приклад. Нехай випадкова величина X приймає три значення: $-1, 0, 1$ з ймовірністю $1/3$. Функція розподілу такої випадкової величини збігається з функціями розподілу випадкових величин $Y_n = X^{2n+1}$, де $n = 0, 1, \dots$

Зауваження 3. Перераховані ймовірнісні характеристики застосовують для опису як неперервних, так і дискретних випадкових величин. У разі, коли випадкова величина X дискретна, щільність розподілу ймовірностей прийнятих значень x_i ($i = 1, 2, \dots$) може бути записана за допомогою δ -функції (додаток 1):

$$f(x) = p_i \delta(x - x_i), \quad (1.2.4)$$

де p_i – ймовірність значення x_i . При цьому множину ймовірностей p_i ($i = 1, 2, \dots$) називають *законом розподілу ймовірностей*.

Однаковий опис неперервних і дискретних випадкових величин забезпечується також застосуванням інтеграла Стілт'еса, що вироджується у звичайний інтеграл Рімана, коли функція $f(x)$ неперервна, і у суму значень функції, коли функція дискретна.

Зауваження 4. Всі три перераховані ймовірнісні характеристики випадкової величини пов'язані між собою, причому однозначно. Це означає, що задання випадкової величини у формі однієї з них дозволяє розрахувати інші. Наприклад, функція розподілу ймовірностей може бути обчислена з щільності розподілу ймовірностей шляхом інтегрування останньої:

$$F(x) = \int_{-\infty}^x f(x) dx, \quad (1.2.5)$$

а щільність розподілу ймовірностей – із характеристичної функції за допомогою оберненого перетворення Фур'є:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Q(j\omega) e^{-j\omega x} d\omega. \quad (1.2.6)$$

Зауваження 5. Відзначимо, що поряд із зазначеними ймовірнісними характеристиками іноді застосовують й інші ймовірнісні характеристики, наприклад, *твірну функцію моментів* і *твірну функцію факторіальних моментів*, що описуються відповідно співвідношеннями

$$M_x(s) = \int_{-\infty}^{\infty} e^{sx} dF(x) \quad (1.2.7)$$

$$\gamma_x(s) = \int_{-\infty}^{\infty} s^x dF(x). \quad (1.2.8)$$

Положення 2. Не всяка функція може виступати в ролі функції розподілу, щільності ймовірностей або характеристичної функції. Ті і тільки ті функції, що мають властивості, наведені у табл. 1.2.1, 1.2.2 і 1.2.3, є відповідно функціями розподілу, щільностями ймовірностей і характеристичними функціями деяких випадкових величин.

Таблиця 1.2.1

Властивості функції розподілу випадкової величини

Назва властивості	Математичне трактування
Невід'ємність	$F(x) \geq 0$
Обмеженість	$0 \leq F(x) \leq 1$
Монотонність	$F(x)$ – неспадна і неперервна зліва

Таблиця 1.2.2

Властивості щільності ймовірностей випадкової величини

Назва властивості	Математичне трактування
Невід'ємність	$f(x) \geq 0$
Нормованість до одиниці	$\int_{-\infty}^{\infty} f(x) dx = 1$
Адитивність	$P\{x_1 \leq X < x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx$

Таблиця 1.2.3

Властивості характеристичної функції випадкової величини

Назва властивості	Математичне трактування
Обмеженість	$ Q(j\omega) \leq Q(0) = 1$
Спряженість (для дійсних X)	$Q(-j\omega) = Q^*(j\omega)$

Приклади деяких випадкових величин, що часто зустрічаються на практиці, наведено у табл. 1.2.4 і 1.2.5. Кілька сот інших типів випадкових величин наведено в роботах [25] та [69].

Таблиця 1.2.4

Розподіли деяких дискретних випадкових величин

Назва розподілу	Закон розподілу ймовірностей
Дискретний рівномірний	$p_i = \frac{1}{I}, \quad i = \overline{0, I-1}$
Біномний	$p_i = C_I^i \vartheta^i (1-\vartheta)^{I-i}, \quad i = \overline{0, I}, \quad 0 \leq \vartheta \leq 1$
Пуассона	$p_i = \frac{\lambda^i}{i!} \exp(-\lambda), \quad i = 0, 1, \dots$
Геометричний	$p_i = \vartheta(1-\vartheta)^i, \quad i = 0, 1, \dots, \quad 0 \leq \vartheta \leq 1$

Гіпергеометричний	$p_i = \frac{C_I^i C_{N-I}^{n-i}}{C_N^n}, \quad i = \overline{0, n}, \quad N \geq n \geq 0, \quad N \geq I \geq 0$
Бернуллі	$p_i = ip + (1-i)q, \quad i = \overline{0, 1}, \quad q = 1-p$
Паскаля	$p_i = C_{m+i-1}^i g^m (1-g)^i, \quad i = 0, 1, \dots, \quad m = 1, 2, \dots, \quad 0 \leq g \leq 1$

Примітка. C_N^n – біномний коефіцієнт: $C_N^n = \frac{N!}{n!(N-n)!}$.

Таблиця 1.2.5

Розподіли деяких неперервних випадкових величин

Назва розподілу	Щільність ймовірностей
Рівномірний	$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$
Гауссівський (нормальний)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right], \quad x \in (-\infty, \infty), \quad \sigma > 0$
Логнормальний	$f(x) = \frac{1}{\sqrt{2\pi}x\sigma} \exp\left[-\frac{(\ln(x/m))^2}{2\sigma^2}\right], \quad x > 0, \quad m > 0, \quad \sigma > 0$
Експоненціальний	$f(x) = \lambda \exp(-\lambda x), \quad \lambda > 0, \quad x > 0$
Хі-квадрат	$f(x) = \frac{x^{(\nu-2)/2} \exp(-x/2)}{2^{\nu/2} \Gamma(\nu/2)}, \quad x \geq 0, \quad \nu = 1, 2, \dots$
Гамма	$f(x) = \frac{(x/b)^{c-1} \exp(-x/b)}{b\Gamma(c)}, \quad x \geq 0,$ де b – параметр масштабу ($b > 0$), c – параметр форми ($c > 0$)
Бета	$f(x) = \frac{x^{\nu-1} (1-x)^{\mu-1} \Gamma(\nu+\mu)}{\Gamma(\nu)\Gamma(\mu)}, \quad x \in [0, 1], \quad \nu > 0, \quad \mu > 0$
Вейбула	$f(x) = \frac{cx^{c-1}}{b^c} \exp(-(x/b)^c), \quad x \geq 0, \quad c > 0, \quad b > 0$
Накагамі	$f(x) = \frac{2}{\Gamma(m)} \left(\frac{m}{\sigma^2}\right)^m x^{2m-1} \exp\left(-\frac{m}{\sigma^2} x^2\right), \quad x > 0,$ де m – параметр масштабу ($m > 0$), σ – параметр форми ($\sigma > 0$)
Райса	$f(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + \alpha^2}{2\sigma^2}\right) I_0\left(\frac{\alpha}{\sigma^2} x\right), \quad x > 0,$ де α – параметр «нецентральності» ($\alpha > 0$), σ – параметр масштабу ($\sigma > 0$)

Трикутний (Сімпсона)	$f(x) = \begin{cases} 4(x-a)/(b-a)^2, & x \in (a, (a+b)/2), \\ 4(b-x)/(b-a)^2, & x \in ((a+b)/2, b), \\ 0, & x \notin (a, b) \end{cases}$
Коші	$f(x) = \frac{1}{\pi} \frac{h}{h^2 + (x-x_0)^2}, \quad x \in (-\infty, \infty), \quad h > 0$
Пірсона	$\frac{df(x)}{dx} = \frac{x-a}{b_0 + b_1x + b_2x^2} f(x),$ де a, b_0, b_1, b_2 – параметри розподілу
Стюдента	$f(x) = \frac{\Gamma((m+1)/2)}{\Gamma(m/2)\sqrt{m\pi}} \left(1 + \frac{x^2}{m}\right)^{-(m+1)/2},$ $x \in [0, \infty), \nu > 0, \mu > 0$
Лапласа	$f(x) = \frac{1}{2\beta} e^{-\frac{ x-m }{\beta}}, \quad x \in (-\infty, \infty),$ де m, β – параметри
Парето	$f(x) = \frac{\alpha}{m} \left(\frac{m}{x}\right)^{\alpha+1}, \quad x \geq m, \quad \alpha > 0$

Примітка. $\Gamma(m)$ – гамма-функція (додаток 2).

1.2.3. Неймовірнісні (числові) характеристики випадкової величини

Поряд із ймовірнісними характеристиками, що вичерпно описують випадкову величину, часто використовують різноманітні числові характеристики, які дозволяють уявити загалом характерні риси розподілу і при необхідності провести апроксимацію його за допомогою інших розподілів. Як такі використовують моменти, кумулянти, параметри розподілу Пірсона, ентропію й ін. Почнемо розгляд із моментів.

Визначення 1. Математичним сподіванням $M[\varphi(X)]$ функції $\varphi(X)$ випадкової величини X з щільністю ймовірностей $f(x)$ називається інтеграл

$$M[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(x) f(x) dx. \quad (1.2.9)$$

Зауваження 1. Математичне сподівання існує не завжди: тільки, коли інтеграл (1.2.9) існує у сенсі абсолютної збіжності.

Зауваження 2. Математичне сподівання $M[\varphi(X)]$ являє собою детерміновану величину, що характеризує середнє значення функції $\varphi(X)$ з урахуванням щільності ймовірностей $f(x)$.

Визначення 2. Математичним сподіванням m_x випадкової величини X з щільністю ймовірностей $f(x)$ називається математичне сподівання функції $\varphi(X) = X$:

$$m_x = M[X] = \int_{-\infty}^{\infty} xf(x)dx. \quad (1.2.10)$$

Визначення 3. Дисперсією $D_x = D[X]$ дійсної випадкової величини X з щільністю ймовірностей $f(x)$ називається математичне сподівання функції $\varphi(X) = (X - m_x)^2$:

$$D_x = M[(X - m_x)^2] = \int_{-\infty}^{\infty} (x - m_x)^2 f(x)dx. \quad (1.2.11)$$

Визначення 4. Середньоквадратичним відхиленням (СКВ) σ_x дійсної випадкової величини X називається корінь із дисперсії: $\sigma_x = \sqrt{D_x}$.

Зауваження 1. Математичне сподівання m_x випадкової величини X характеризує середнє її значення з урахуванням щільності ймовірностей $f(x)$, а дисперсія D_x і СКВ σ_x – розсіювання цієї випадкової величини відносно математичного сподівання.

Зауваження 2. Математичне сподівання m_x випадкової величини X і СКВ σ_x має таку ж розмірність, як і випадкова величина; розмірність дисперсії D_x дорівнює квадрату розмірності випадкової величини.

Зауваження 3. Математичне сподівання не випадкової (детермінованої) величини (константи) дорівнює цій величині, а дисперсія і СКВ – нулю.

Визначення 5. Коефіцієнтом варіації називається відношення математичного сподівання до СКВ випадкової величини: $\frac{m_x}{\sigma_x}$.

Визначення 6. Медіаною розподілу x_e називається значення випадкової величини X , що поділяє площу під щільністю ймовірностей $f(x)$ навпіл, тобто є розв'язком рівняння $F(x) = 0,5$.

Визначення 7. Квантилем x_p порядку p називається корінь рівняння $F(x) = p$, де p – фіксоване число ($0 < p < 1$).

Зауваження 1. Медіана являє собою квантиль порядку 0,5: $x_e = x_{0,5}$.

Зауваження 2. Квантилі $x_{0,25}$, $x_{0,5}$, $x_{0,75}$ іноді називають *квартілями*, квантилі $x_{0,1}$, $x_{0,2}, \dots, x_{0,9}$ – *децилями*, а квантилі $x_{0,01}$, $x_{0,02}, \dots, x_{0,09}$ – *процентилями*.

Зауваження 3. Квантилі (у тому числі і медіана) можуть бути визначені неоднозначно. Неоднозначність спостерігається, коли рівняння $F(x) = p$ має більш одного кореня.

Визначення 8. Модою x_m розподілу випадкової величини X називається точка максимуму щільності ймовірностей $f(x)$.

Зауваження 1. Мода визначається як для неперервних, так і для дискретних випадкових величин.

Зауваження 2. Розподіл може мати одну моду (*уніmodalьний*), дві моди (*біmodalьний*), багато мод (*мультиmodalьний*) або взагалі не мати мод.

Зауваження 3. Якщо розподіл уніmodalьний і симетричний відносно математичного сподівання, мода і медіана збігаються з математичним сподіванням. Якщо розподіл уніmodalьний і асиметричний, то мода може

розташовуватися як зліва від медіани (у випадку додатної асиметрії), так і праворуч від неї (у випадку від'ємної асиметрії).

Визначення 9. Початковим моментом m_ν , ν -го порядку випадкової величини X називається математичне сподівання функції $\varphi(X) = X^\nu$, абсолютним початковим моментом ν -го порядку – математичне сподівання функції $\varphi(X) = |X|^\nu$, центральним моментом μ_ν , ν -го порядку – математичне сподівання функції $\varphi(X) = (X - m_x)^\nu$, а абсолютним центральним моментом ν -го порядку – математичне сподівання функції $\varphi(X) = |X - m_x|^\nu$, де ν – ціле додатне число.

Зауваження 1. Абсолютними моментами користуються рідко. Звичайно обходяться початковими і центральними моментами.

Зауваження 2. Початкові і центральні моменти пов'язані між собою, зокрема:

$$\begin{aligned} \mu_1 &= 0; & \mu_2 &= D_x = m_2 - m_x^2; \\ \mu_3 &= m_3 - 3m_x m_2 + 2m_x^2; & \mu_4 &= m_4 - 4m_x m_3 + 6m_x^2 m_2 - 3m_x^4. \end{aligned}$$

Зауваження 3. Центральні непарні моменти гауссівського розподілу дорівнюють нулю, а парні – виражаються через дисперсію D_x :

$$\mu_\nu = \begin{cases} 1 \cdot 3 \cdot 5 \dots (\nu - 1) D_x^{\nu/2} & \text{при } \nu - \text{парному}; \\ 0 & \text{при } \nu - \text{непарному}. \end{cases}$$

Початкові моменти гауссівського розподілу цілком визначаються математичним сподіванням m_x і дисперсією D_x , зокрема:

$$\begin{aligned} m_1 &= m_x; & m_2 &= D_x + m_x^2; \\ m_3 &= 3m_x D_x + m_x^3; & m_4 &= 3D_x^2 + 6m_x^2 D_x + m_x^4. \end{aligned}$$

Зауваження 4. Не всі розподіли мають моменти довільного порядку. Розподіл Коші, наприклад, не має моментів порядку $\nu \geq 2$.

Положення. Множина всіх моментів (за умови, що вони існують і скінченні) однозначно визначає ймовірнісні характеристики розподілу випадкової величини. При цьому характеристична функція

$$Q(j\omega) = 1 + \sum_{\nu=1}^{\infty} \frac{(j\omega)^\nu}{\nu!} m_\nu, \quad (1.2.12)$$

де

$$m_\nu = \left. \frac{d^\nu Q(j\omega)}{j^\nu d\omega^\nu} \right|_{\omega=0}. \quad (1.2.13)$$

Доведення цього положення базується на розкладенні експоненти виразу (1.2.3) у ряд Маклорена й обчисленні математичних сподівань степінних функцій.

Зауваження 1. Вираз (1.2.12) являє собою розкладення характеристичної функції в ряд Маклорена.

Зауваження 2. Часто з'ясовується, що для спрощення розрахунків зручніше розкласти в ряд Маклорена не характеристичну функцію, а логарифм характеристичної функції:

$$\ln Q(j\omega) = \sum_{\nu=1}^{\infty} \frac{\kappa_\nu}{\nu!} (j\omega)^\nu. \quad (1.2.14)$$

Визначення 10. Кумулянтами або семіінваріантами називаються коефіцієнти розкладення логарифма характеристичної функції (1.2.14) в ряд Маклорена. Ці коефіцієнти визначаються таким чином:

$$\kappa_\nu = \left. \frac{d^\nu \ln Q(j\omega)}{j^\nu d\omega^\nu} \right|_{\omega=0}. \quad (1.2.15)$$

Зауваження 1. З співвідношення (1.2.14) випливає, що характеристична функція пов'язана з кумулянтами таким співвідношенням:

$$Q(j\omega) = \exp \left[\sum_{\nu=1}^{\infty} \frac{\kappa_\nu}{\nu!} (j\omega)^\nu \right]. \quad (1.2.16)$$

Зауваження 2. Кумулянт κ_ν являє собою поліном від моментів m_1, \dots, m_ν і, навпаки, момент m_ν являє собою поліном від кумулянтів $\kappa_1, \dots, \kappa_\nu$. Зокрема:

$$\begin{aligned} \kappa_1 &= m_1 = m_x; \\ \kappa_2 &= m_2 - m_x^2 = \mu_2 = D_x; \\ \kappa_3 &= m_3 - 3m_2 m_x + 2m_x^3 = \mu_3; \\ \kappa_4 &= m_4 - 3m_2^2 - 4m_3 m_x + 12m_2 m_x^2 - 6m_x^4 = \mu_4 - 3\mu_2^3. \end{aligned}$$

Зауваження 3. Моменти і кумулянти випадкової величини просто виражаються через твірні функції моментів:

$$m_\nu = \left. \frac{d^\nu M_x(s)}{ds^\nu} \right|_{s=0}; \quad \kappa_\nu = \left. \frac{d^\nu \ln M_x(s)}{ds^\nu} \right|_{s=0}; \quad (1.2.17)$$

і, навпаки,

$$M_x(s) = \sum_{\nu=0}^{\infty} \frac{m_\nu}{\nu!} s^\nu; \quad \ln M_x(s) = \sum_{\nu=1}^{\infty} \frac{\kappa_\nu}{\nu!} s^\nu. \quad (1.2.18)$$

Визначення 11. Коефіцієнтами асиметрії й ексцесу називаються відповідно величини

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}}; \quad (1.2.19)$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\kappa_4}{\kappa_2^2}. \quad (1.2.20)$$

Зауваження 1. Коефіцієнт асиметрії характеризує ступінь асиметричності розподілу. При симетричному розподілі $\gamma_1 = 0$. Коефіцієнт ексцесу характеризує ступінь «приплюснутості» вершини щільності ймовірностей. При гауссівському розподілі $\gamma_2 = 0$.

Зауваження 2. Замість величин γ_1 і γ_2 іноді використовують величини γ_1^2 й $\gamma_2 + 3$ або $(\gamma_2 + 3)/2$.

Визначення 12. Розмахом називається різниця між максимальним і мінімальним значеннями випадкової величини.

Визначення 13. Ентропією безрозмірної (нормованої) випадкової величини X з щільністю ймовірностей $f(x)$ називається величина

$$H_x = -M[\log_2 f(x)] = - \int_{-\infty}^{\infty} f(x) \log_2 f(x) dx. \quad (1.2.21)$$

Зауваження 1. Ентропія є мірою априорної непевності значення, що прийме величина X після проведення досліду.

Зауваження 2. У випадку, коли величина X – детермінована, ентропія $H_x = 0$. Якщо X – випадкова величина, що визначена на скінченному інтервалі $[a, b]$, то ентропія максимальна, коли розподіл рівномірний на інтервалі.

Зауваження 3. При даній дисперсії σ_x^2 найбільшу ентропію серед неперервних випадкових величин має випадкова величина X з гауссівським законом розподілу. При цьому ентропія $H_x = \log_2 \sqrt{2\pi e} \sigma_x$.

Визначення 14. Початковим факторіальним моментом порядку ν випадкової величини X (відносно $x = 0$) називається величина

$$m_{[\nu]} = M[X^{[\nu]}] = M[X(X-1)\dots(X-\nu+1)]. \quad (1.2.22)$$

Зауваження. Початковий факторіальний момент $m_{[\nu]}$ випадкової величини X пов'язаний із твірною функцією факторіальних моментів $\gamma_x(s)$ такими співвідношеннями:

$$m_{[\nu]} = \gamma_x^{(\nu)}(1);$$

$$\gamma_x(s+1) = \sum_{\nu=0}^{\infty} m_{[\nu]} \frac{s^{\nu}}{\nu!}.$$

Визначення 15. Центральним факторіальним моментом порядку ν випадкової величини X називається величина

$$\mu_{[\nu]} = M[(X - m_x)^{[\nu]}]. \quad (1.2.23)$$

1.3. Векторні випадкові величини

1.3.1. Ймовірнісні характеристики векторної випадкової величини

Положення 1. У векторному випадку N -вимірну випадкову величину \vec{X} найбільш повно описують такі ймовірнісні характеристики:

1) N -вимірна функція розподілу ймовірностей

$$F_N(\vec{x}) = P\{X_1 < x_1, \dots, X_N < x_N\} \equiv P\{\vec{X} < \vec{x}\}; \quad (1.3.1)$$

2) N -вимірна щільність розподілу ймовірностей, що визначається за формулою

$$f_N(\vec{x}) = \frac{\partial^N F_N(\vec{x})}{\partial x_1 \partial x_2 \dots \partial x_N}; \quad (1.3.2)$$

3) N -вимірна характеристична функція

$$Q_N(j\vec{\omega}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_N(\vec{x}) e^{j\vec{\omega}\vec{x}} d\vec{x}. \quad (1.3.3)$$

Ці характеристики пов'язані одна з одною (див. табл. 1.3.1).

Властивості. Властивості ймовірнісних характеристик векторних випадкових величин наведено у табл. 1.3.2 – 1.3.4.

Положення 2. Скалярну комплексну випадкову величину $\dot{Z} = X + jY$ можна розглядати як векторну двовимірну дійсну випадкову величину $\vec{Z} = (X, Y)$, а векторну N -вимірну комплексну випадкову величину $\dot{\vec{Z}} = \vec{X} + j\vec{Y}$ – як векторну $2N$ -вимірну дійсну випадкову величину $\vec{Z} = (\vec{X}, \vec{Y})$.

Зауваження. Представлення комплексних випадкових величин за допомогою векторних випадкових величин в деяких випадках значно полегшує викладки. Цим прийомом будемо користуватися і далі.

Таблиця 1.3.1

Основні ймовірнісні характеристики векторної випадкової величини

Назва характеристики	Математичне трактування
Функція розподілу ймовірностей	$F_N(\vec{x}) = P\{X_1 < x_1, \dots, X_N < x_N\};$ $F_N(\vec{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_N} f_N(\vec{x}) d\vec{x}$
Щільність розподілу ймовірностей	$f_N(\vec{x}) = \frac{\partial^N F_N(\vec{x})}{\partial x_1 \dots \partial x_N};$ $f_N(\vec{x}) = \frac{1}{(2\pi)^N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} Q_N(j\vec{\omega}) e^{-j\vec{\omega}\vec{x}} d\vec{\omega}$
Характеристична функція	$Q_N(j\vec{\omega}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_N(\vec{x}) e^{j\vec{\omega}\vec{x}} d\vec{x}$

Таблиця 1.3.2

Основні властивості функції розподілу ймовірностей

Властивість	Математичне трактування
Невід'ємність	$F_N(\vec{x}) \geq 0$
Обмеженість	$0 \leq F_N(\vec{x}) \leq 1$
Монотонність	$F_N(\vec{x})$ неспадна і неперервна зліва по кожній координаті x_n
Симетричність	$F_N(\vec{x})$ не змінюється при будь-яких перестановках компонент вектора \vec{x}
Узгодженість	$F_M(x_1, \dots, x_M) = F_N(x_1, \dots, x_M, \infty, \dots, \infty)$ при $M < N$

Таблиця 1.3.3

Основні властивості щільності ймовірностей

Властивість	Математичне трактування
Невід'ємність	$f_N(\vec{x}) \geq 0$
Нормованість до одиниці	$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_N(\vec{x}) d\vec{x} = 1$
Симетричність	$f_N(\vec{x})$ не змінюється при будь-яких перестановках аргументів x_n
Узгодженість	$f_M(x_1, \dots, x_M) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_N(x_1, \dots, x_N) dx_{M+1} \dots dx_N$ при $M < N$

Таблиця 1.3.4

Основні властивості характеристичної функції

Властивість	Математичне трактування
Обмеженість	$ Q_N(j\vec{\omega}) \leq 1$
Симетричність	$Q(j\vec{\omega})$ не змінюється при будь-яких перестановках аргументів $j\omega_n$
Узгодженість	$Q_M(j\omega_1, \dots, j\omega_M) = Q_M(j\omega_1, \dots, j\omega_M, 0, \dots, 0)$ при $M < N$

1.3.2. Умовні ймовірнісні характеристики векторної випадкової величини

При розв'язанні ряду практичних задач досить часто використовується апарат умовних ймовірнісних характеристик випадкових величин.

Розглянемо N -вимірну випадкову величину $\vec{Z} = (\vec{X}, \vec{Y})$, що складається з M -вимірної випадкової величини \vec{X} і $(N-M)$ -вимірної випадкової величини \vec{Y} .

Визначення 1. Умовною функцією розподілу $(N-M)$ -вимірної випадкової величини \vec{Y} називається $(N-M)$ -вимірна функція розподілу $F_{N-M}(\vec{y}/\vec{x})$, визначена за умови, що M -вимірна випадкова величина \vec{X} прийняла конкретне значення \vec{x} .

Визначення 2. Умовною щільністю ймовірностей $(N-M)$ -вимірної випадкової величини \vec{Y} називається $(N-M)$ -вимірна щільність ймовірностей $f_{N-M}(\vec{y}/\vec{x})$, визначена за умови, що випадкова величина \vec{X} прийняла конкретне значення \vec{x} .

Визначення 3. Умовною характеристичною функцією $(N-M)$ -вимірної випадкової величини \vec{Y} називається $(N-M)$ -вимірна характеристична функція $Q_{N-M}(j\omega_1, \dots, j\omega_{N-M}/\vec{x})$, визначена за умови, що випадкова величина \vec{X} прийняла конкретне значення \vec{x} .

Зауваження. Умовна функція розподілу $F_{N-M}(\vec{y}/\vec{x})$, умовна щільність ймовірностей $f_{N-M}(\vec{y}/\vec{x})$ й умовна характеристична функція $Q_{N-M}(j\omega_1, \dots, j\omega_{N-M}/\vec{x})$ пов'язані між собою такими співвідношеннями:

$$F_{N-M}(\vec{y}/\vec{x}) = \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_{N-M}} f_{N-M}(\vec{u}/\vec{x}) du_1 \dots du_{N-M}; \quad (1.3.4)$$

$$f_{N-M}(\vec{y}/\vec{x}) = \frac{\partial^{N-M} F_{N-M}(\vec{y}/\vec{x})}{\partial y_1 \dots \partial y_{N-M}}; \quad (1.3.5)$$

$$Q_{N-M}(j\omega_1, \dots, j\omega_{N-M}/\vec{x}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{N-M}(\vec{y}/\vec{x}) e^{j\vec{\omega}\vec{y}} d\vec{y}. \quad (1.3.6)$$

Положення 1. Спільна щільність ймовірностей $f_N(\vec{x}, \vec{y})$ системи випадкових величин $\vec{Z} = (\vec{X}, \vec{Y})$ пов'язана з умовною щільністю ймовірностей $f_{N-M}(\vec{y}/\vec{x})$ випадкової величини \vec{Y} і щільністю ймовірностей $f_M(\vec{x})$ випадкової величини \vec{X} співвідношенням

$$f_N(\vec{x}, \vec{y}) = f_{N-M}(\vec{y}/\vec{x}) f_M(\vec{x}), \quad (1.3.7)$$

що впливає із теореми множення ймовірностей.

Визначення 4. Випадкові величини X і Y називаються *незалежними*, якщо їхня спільна щільність ймовірностей $f_2(x, y)$ дорівнює добутку одновимірних щільностей ймовірностей $f(x)$ і $f(y)$ відповідно величин X і Y :

$$f_2(x, y) = f(x) f(y). \quad (1.3.8)$$

Зауваження. Для незалежних і тільки незалежних випадкових величин X і Y

$$F_2(x, y) = F(x) F(y); \quad (1.3.9)$$

$$Q_2(j\omega_x, j\omega_y) = Q(j\omega_x) Q(j\omega_y). \quad (1.3.10)$$

Визначення 5. Випадкові величини X_1, \dots, X_N називаються *незалежними в сукупності (взаємно незалежними)*, якщо їхня спільна щільність ймовірностей $f_N(x_1, \dots, x_N)$ дорівнює добутку одновимірних щільностей ймовірностей $f(x_1), \dots, f(x_N)$:

$$f_N(x_1, \dots, x_N) = f(x_1) \dots f(x_N). \quad (1.3.11)$$

Зауваження. Для незалежних і тільки незалежних у сукупності випадкових величин X_1, \dots, X_N

$$F_N(x_1, \dots, x_N) = F(x_1) \dots F(x_N); \quad (1.3.12)$$

$$Q_N(j\omega_1, \dots, j\omega_N) = Q(j\omega_1) \dots Q(j\omega_N). \quad (1.3.13)$$

Визначення 6. Випадкові величини X_1, \dots, X_N називаються *попарно незалежними*, якщо довільна пара випадкових величин X_i, X_j , ($i, j = \overline{1, N}$) є незалежною.

Зауваження 1. Якщо випадкові величини X_1, \dots, X_N взаємно незалежні, то вони попарно незалежні. Обернене твердження *невірне*.

Приклад. Нехай для випадкових величин X_1, X_2, X_3 справедливі рівності $f_2(x_1, x_2) = f(x_1) f(x_2)$, $f_2(x_1, x_3) = f(x_1) f(x_3)$, $f_2(x_2, x_3) = f(x_2) f(x_3)$. З цих рівностей не випливає, що $f_3(x_1, x_2, x_3) = f(x_1) f(x_2) f(x_3)$. Дійсно, відповідно до теореми множення теорії ймовірностей $f_3(x_1, x_2, x_3) = f_2(x_1, x_2) f(x_3 / x_1, x_2) = f(x_1) f(x_2) f(x_3 / x_1, x_2)$. Звідси очевидно, що лише в тому випадку, коли умовна щільність ймовірностей випадкової величини X_3 дорівнює її безумовній щільності ймовірностей ($f(x_3 / x_1, x_2) = f(x_3)$), можна стверджувати, що випадкові величини X_1, X_2, X_3 незалежні у сукупності.

Зауваження 2. Іноді залежні випадкові величини можна розділити на незалежні групи. При цьому підгрупи випадкових величин із різних груп незалежні.

Положення 2. Щільність ймовірностей $f_N(x_1, \dots, x_N)$ N -вимірної випадкової величини (X_1, \dots, X_N) описується виразом

$$f_N(x_1, \dots, x_N) = f_1(x_N / x_1, \dots, x_{N-1}) \dots f_1(x_2 / x_1) f_1(x_1), \quad (1.3.14)$$

тобто цілком визначається одновимірними умовними щільностями ймовірностей $f_1(x_n / x_1, \dots, x_{n-1})$ ($n = \overline{2, N}$) й одновимірною безумовною щільністю ймовірностей $f_1(x_1)$.

Доведення цього співвідношення може бути проведено методом математичної індукції з застосуванням співвідношення (1.3.7).

Зауваження. Якщо компоненти вектора (X_1, \dots, X_N) незалежні, то з виразу (1.3.14) випливає вираз (1.3.11).

1.3.3. Неймовірнісні (числові) характеристики векторної випадкової величини

Положення 1. Більшість із введених у підрозділі 1.3.3 понять узагальнюються на багатовимірний випадок, причому не тільки для дійсних, але й для комплексних випадкові величин.

Зауваження. Для опису комплексних випадкових величин можна застосовувати підхід, який наведено у положенні 2 підрозділу 1.3.1.

Визначення 1. Математичним сподіванням M -вимірної векторної функції $\vec{\varphi}(\vec{X})$ N -вимірної випадкової величини $\vec{X} = (X_1, \dots, X_N)$ з щільністю ймовірностей $f_N(x_1, \dots, x_N)$ називається M -вимірний вектор, що описується інтегралом

$$M[\vec{\varphi}(\vec{X})] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \vec{\varphi}(x_1, \dots, x_N) f_N(x_1, \dots, x_N) dx_1 \dots dx_N, \quad (1.3.15)$$

якщо останній існує.

Зауваження 1. Математичне сподівання $M[\vec{\varphi}(\vec{X})]$ векторної функції $\vec{\varphi}(\vec{X})$ векторної випадкової величини \vec{X} являє собою невипадковий (детермінований) вектор, що розраховується з урахуванням щільності ймовірностей $f_N(\vec{x})$ цієї випадкової величини.

Зауваження 2. У загальному випадку вимірність N випадкової величини \vec{X} і вимірність M математичного сподівання функції $\vec{\varphi}(\vec{X})$ відрізняються одна від одної.

Визначення 2. Математичним сподіванням \vec{m}_x N -вимірної векторної випадкової величини \vec{X} називається N -вимірний вектор, що являє собою математичне сподівання функції $\vec{\varphi}(\vec{X}) = \vec{X}$:

$$\vec{m}_x = M[\vec{X}]. \quad (1.3.16)$$

Зауваження. Для векторної комплексної випадкової величини $\vec{Z} = \vec{X} + j\vec{Y}$ математичним сподіванням називається вектор

$$\vec{m}_z = \vec{m}_x + j\vec{m}_y.$$

Визначення 3. Дисперсією $\vec{D}_x = D[\vec{X}]$ N -вимірної векторної дійсної випадкової величини \vec{X} називається N -вимірний вектор, що являє собою математичне сподівання функції

$$\vec{\varphi}(\vec{X}) = ((X_n - m_{x_n})^2, \quad n = \overline{1, N}). \quad (1.3.17)$$

Зауваження. Дисперсією \vec{D}_z N -вимірної векторної комплексної випадкової величини $\vec{Z} = \vec{X} + j\vec{Y}$ називається N -вимірний комплексний вектор, що являє собою математичне сподівання вектора

$$((X_n - m_{x_n})^2 + j(Y_n - m_{y_n})^2, \quad n = \overline{1, N}).$$

Визначення 4. Середньоквадратичним відхиленням (СКВ) $\vec{\sigma}_x$ N -вимірної векторної випадкової величини \vec{X} називається N -вимірний вектор, компоненти якого дорівнюють кореню з компонент вектора дисперсії \vec{D}_x .

Зауваження 1. Компоненти вектора математичного сподівання \vec{m}_x характеризують середні значення відповідних компонент вектора \vec{X} , а компоненти вектора дисперсії \vec{D}_x і СКВ $\vec{\sigma}_x$ – їхнє розсіювання щодо відповідних компонент вектора математичного сподівання.

Зауваження 2. Компоненти вектора математичного сподівання \vec{m}_x векторної випадкової величини \vec{X} і компоненти вектора СКВ $\vec{\sigma}_x$ мають таку ж розмірність, як і відповідні компоненти випадкової величини \vec{X} , а розмірності компонент вектора дисперсії \vec{D}_x дорівнюють квадратам розмірностей відповідних компонент \vec{X} .

Приклад. Тиск масла, швидкість вихлопних газів і рівень шумовипромінювання двигуна літака складають тривимірний випадковий вектор \vec{X} , компоненти якого вимірюються відповідно в Па, м/с і дБ. У даному випадку одиницями виміру першої, другої і третьої компонент векторів \vec{m}_x і $\vec{\sigma}_x$ є відповідно Па, м/с і дБ, а компоненти вектора \vec{D}_x – відповідно Па², (м/с)² і (дБ)².

Зауваження 3. Середньоквадратичним відхиленням $\vec{\sigma}_z$ N -вимірної векторної комплексної випадкової величини $\vec{Z} = \vec{X} + j\vec{Y}$ називається N -вимірний комплексний вектор, дійсні компоненти якого дорівнюють кореню з відповідних дійсних компонент дисперсії \vec{D}_z , а уявні компоненти – з відповідних уявних компонент дисперсії \vec{D}_z .

Визначення 5. Початковим моментом $m_{v_1 \dots v_N}$ порядку $v = v_1 + \dots + v_N$ компонент N -вимірної векторної дійсної випадкової величини \vec{X} називається скалярна величина, що описується виразом

$$m_{v_1 \dots v_N} = M[X_1^{v_1} \dots X_N^{v_N}], \quad (1.3.18)$$

де v_n – ціле додатне число ($n = 1, N$).

Зауваження 1. Моменти $m_{v_1 \dots v_N}$ являють собою коефіцієнти розкладення характеристичної функції $Q_N(j\omega_1, \dots, j\omega_N)$ в ряд Маклорена:

$$Q_N(j\omega_1, \dots, j\omega_N) = \sum_{v_1, \dots, v_N=0}^{\infty} \frac{m_{v_1 \dots v_N}}{v_1! \dots v_N!} (j\omega_1)^{v_1} \dots (j\omega_N)^{v_N}. \quad (1.3.19)$$

Звідси випливає, що

$$m_{v_1 \dots v_N} = (-j)^{v_1 + \dots + v_N} \left. \frac{\partial^{v_1 + \dots + v_N} Q_N(j\omega_1, \dots, j\omega_N)}{\partial \omega_1^{v_1} \dots \partial \omega_N^{v_N}} \right|_{\omega_1 = \dots = \omega_N = 0}. \quad (1.3.20)$$

Зауваження 2. У загальному випадку моменти того ж самого v -го порядку, але з різною комбінацією параметрів v_1, \dots, v_N , відрізняються один від одного.

Визначення 6. Кумулянт $\kappa_{v_1 \dots v_N}$ порядку $v = v_1 + \dots + v_N$ компонент N -вимірної векторної випадкової величини \vec{X} називається коефіцієнт розкладення логарифма характеристичної функції в степінний ряд:

$$\ln Q_N(j\omega_1, \dots, j\omega_N) = \sum_{v_1, \dots, v_N=0}^{\infty} \frac{\kappa_{v_1 \dots v_N}}{v_1! \dots v_N!} (j\omega_1)^{v_1} \dots (j\omega_N)^{v_N}. \quad (1.3.21)$$

При цьому

$$\kappa_{v_1 \dots v_N} = (-j)^{v_1 + \dots + v_N} \left. \frac{\partial^{v_1 + \dots + v_N} \ln Q_N(j\omega_1, \dots, j\omega_N)}{\partial \omega_1^{v_1} \dots \partial \omega_N^{v_N}} \right|_{\omega_1 = \dots = \omega_N = 0}. \quad (1.3.22)$$

Визначення 7. Центральним моментом $\mu_{v_1 \dots v_N}$ порядку $v = v_1 + \dots + v_N$ компонент N -вимірної векторної випадкової величини \vec{X} називається скалярна величина, що описується виразом

$$\mu_{v_1 \dots v_N} = M[X_1^{0 \ v_1} \dots X_N^{0 \ v_N}], \quad (1.3.23)$$

де $\vec{X} = \vec{X} - \vec{m}_x$ – центрований випадковий вектор.

Визначення 8. Змішаний центральний момент другого порядку μ_{11} випадкових величин X_1 і X_2 називається *кореляційним моментом*, змішаний початковий момент другого порядку m_{11} – *коваріаційним моментом*, а змішаний центральний момент другого порядку, нормований на відповідні СКВ σ_{x_1} і σ_{x_2} , – *коефіцієнтом кореляції*

$$r = \frac{\mu_{11}}{\sigma_{x_1} \sigma_{x_2}}. \quad (1.3.24)$$

Зауваження 1. Кореляційний і коваріаційний моменти випадкових величин X_1 і X_2 мають розмірність, рівну добутку розмірностей цих величин. Коефіцієнт кореляції – безрозмірна величина.

Зауваження 2. Кореляційний момент μ_{11} випадкових величин X_1 і X_2 пов'язаний із коваріаційним моментом m_{11} і математичними сподіваннями m_{x_1} і m_{x_2} цих величин співвідношенням

$$\mu_{11} = m_{11} - m_{x_1} m_{x_2}, \quad (1.3.25)$$

яке впливає із визначення кореляційного і коваріаційного моментів.

Зауваження 3. Кореляційний момент μ_{11} і коефіцієнт кореляції r характеризують *лінійний і тільки лінійний зв'язок* між випадковими величинами, що розглядаються. При відсутності лінійного зв'язку $\mu_{11} = r = 0$. При наявності тільки лінійного зв'язку модуль коефіцієнта кореляції $|r| = 1$.

Зауваження 4. Поняття кореляційного моменту і коваріаційного моменту, що сформульовані вище для дійсних випадкових величин, узагальнюються на випадок комплексних випадкових величин.

Визначення 9. Коваріаційним моментом m_{11} комплексної випадкової величини $Z = X + jY$ називається математичне сподівання добутку її дійсної та уявної частин:

$$m_{11} = M[XY],$$

а кореляційним моментом μ_{11} – математичне сподівання добутку її центрованих дійсної та уявної частин:

$$\mu_{11} = M[(X - m_x)(Y - m_y)].$$

Визначення 10. Матрицею коваріаційних моментів K двох комплексних випадкових величин $\dot{Z}_1 = X_1 + jY_1$, $\dot{Z}_2 = X_2 + jY_2$ називається матриця

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix},$$

де елементи матриці $K_{nm} = M[\dot{Z}_n \dot{Z}_m^*]$ ($n, m = \overline{1, 2}$), а матрицею кореляційних моментів R – матриця

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix},$$

де елементи матриці $R_{nm} = M[(\dot{Z}_n - \dot{m}_{z_n})(Z_m^* - m_{z_m}^*)]$ ($n, m = \overline{1, 2}$).

Зауваження. Діагональні елементи R_{11} , R_{22} матриці кореляційних моментів R являють собою модулі дисперсій випадкових величин \dot{Z}_1 і \dot{Z}_2 :

$$R_{11} = |\dot{D}[\dot{Z}_1]| = M[(X_1 - m_{x_1})^2 + (Y_1 - m_{y_1})^2];$$

$$R_{22} = |\dot{D}[\dot{Z}_2]| = M[(X_2 - m_{x_2})^2 + (Y_2 - m_{y_2})^2].$$

Визначення 11. Випадкові величини X_1 і X_2 називаються *некорельованими* (лінійно незалежними), якщо їхній кореляційний момент $\mu_{11} = 0$, тобто коваріаційний момент

$$m_{11} = m_{x_1} m_{x_2}. \quad (1.3.26)$$

Зауваження. Некорельованість і незалежність випадкових величин – різні поняття. Причому з незалежності випадкових величин випливає некорельованість, але в загальному випадку з некорельованості не випливає незалежність. Взаємоднозначний зв'язок цих понять має місце лише в деяких окремих випадках, наприклад, коли випадкові величини спільно гауссівські (визначення поняття спільно гауссівських випадкових величин наведено у підрозділі 1.4).

Визначення 12. Дійсні випадкові величини X_1 і X_2 називаються *ортогональними*, якщо коваріаційний момент $m_{11} = 0$.

Зауваження 1. Поняття некорельованості й ортогональності – різні поняття. Якщо хоча б одне з математичних сподівань m_{x_1} , m_{x_2} дорівнює нулю, з ортогональності випадкових величин випливає їх некорельованість, а з некорельованості – ортогональність.

Зауваження 2. Поняття некорельованості й ортогональності узагальнюються на випадок N випадкових величин, причому не обов'язково дійсних.

Визначення 13. Комплексні випадкові величини X_1, \dots, X_N називаються *некорельованими* (попарно), якщо

$$M[X_n X_m^*] = M[X_n] M[X_m^*], \quad n \neq m, \quad n, m = \overline{1, N}, \quad (1.3.27)$$

де зірочкою позначена процедура комплексного спряження.

Визначення 14. Комплексні випадкові величини X_1, \dots, X_N називаються *ортогональними* (попарно), якщо

$$M[X_n X_m^*] = 0, \quad n \neq m, \quad n, m = \overline{1, N}. \quad (1.3.28)$$

Визначення 15. Умовною ентропією безрозмірної (нормованої) випадкової величини Y за умови, що безрозмірна випадкова величина X прийняла значення x , називається величина

$$H_{y/x} = -M[\log_2 f(y/x)]. \quad (1.3.29)$$

Зауваження 1. Ентропія безрозмірної випадкової величини $Z = (X, Y)$

$$H_z = -M[\log_2 f(x, y)]$$

описується співвідношенням $H_z = H_x + H_{y/x}$.

Доведення цього положення випливає з визначення поняття ентропії і теореми множення теорії ймовірностей.

Зауваження 2. Ентропія випадкової величини $Z = (X, Y)$ не більше ентропії складових випадкових величин:

$$H_z \leq H_x + H_y, \quad (1.3.30)$$

причому знак рівності має місце, коли X і Y незалежні.

Зауваження 3. Із зауваження 2 випливає, що при проведенні вимірів випадкової величини X з точністю ΔX , що не залежить від X , ентропія випадкової величини $(X, \Delta X)$ дорівнює сумі ентропії випадкової величини X й ентропії точності виміру ΔX цієї величини.

Зауваження 4. Мірою залежності безрозмірних випадкових величин X і Y є величина

$$J(x, y) = H_x + H_y - H_z. \quad (1.3.31)$$

Зауваження 5. Ентропія векторної випадкової величини $\vec{X} = (X_1, \dots, X_N)$ описується виразом

$$H_{\vec{x}} = H_{x_1} + H_{x_2/x_1} + \dots + H_{x_N/x_1, \dots, x_{N-1}}, \quad (1.3.32)$$

що випливає із співвідношення (1.3.14). У випадку, коли компоненти вектора \vec{X} незалежні,

$$H_{\vec{x}} = H_{x_1} + H_{x_2} + \dots + H_{x_N}. \quad (1.3.33)$$

Визначення 16. Кривою сталої щільності ймовірностей називається розв'язок рівняння $f_N(x_1, \dots, x_N) = C$, де C – константа.

Зауваження. Криві сталої щільності ймовірностей часто використовуються для наочного уявлення щільностей ймовірностей двох випадкових величин.

Положення. Використовуючи поняття умовної щільності ймовірностей й умовної характеристичної функції, за аналогією з безумовними моментами і кумулянтами, вводять поняття умовних моментів і умовних кумулянтів.

Зауваження 1. Варто пам'ятати, що всі умовні моменти і кумулянти залежать від умов, тобто є функціями умов.

Зауваження 2. На практиці часто використовуються умовне математичне сподівання випадкової величини й умовна дисперсія, що визначаються так:

$$\bar{m}_{y/x} = M[\bar{Y}/\bar{x}];$$

$$\bar{D}_{y/x} = D[\bar{Y}/\bar{x}].$$

1.4. Перетворення випадкових величин

1.4.1. Поняття про математичні перетворення

Визначення 1. набір правил, що ставлять кожному об'єкту x класу A у відповідність деякий об'єкт y класу B , називається *перетворенням* (відображенням) класу A у клас B .

Зауваження 1. Якщо класи A і B являють собою числові множини, то y називається *функцією аргументу x* .

Приклади 1. Приклади функцій: $y = \lg x, y = x^5$.

Зауваження 2. Якщо клас A являє собою множину функцій, а клас B – числову множину, то y називається *функціоналом x* .

Приклади 2. Приклади функціоналів: 1) $y = \int_{-\infty}^{\infty} x(t)dt$; 2) $y = \max x(t)$.

Зауваження 3. Якщо обидва класи A і B являють собою множини функцій, то перетворення називається *оператором*.

Приклади 3. Приклади операторів: $y(\tau) = \int_{-\infty}^{\tau} x(t)dt, y(t) = \lg x(t)$.

Зауваження 4. Найбільш загальним є поняття оператора. Поняття функції і функціонала входять у поняття оператора як окремі випадки.

Зауваження 5. У даному розділі розглянуті перетворення, що являють собою функції випадкових величин. При цьому аргументом виступають значення випадкової величини. У розділі 2 розглянуті перетворення операторного типу.

Положення. Функції класифікуються за різноманітними ознаками. Деякі поширені класи функцій випадкових величин наведено в табл. 1.4.1.

Таблиця 1.4.1

Класи функцій випадкових величин

Ознака класифікації	Вид функції	
Стохастичні особливості параметрів функції	Детермінована (з детермінованими параметрами)	Функція з випадковими параметрами
Тип функції	Лінійна	Нелінійна

Приклади 4. Функції $y = |x|$ і $y = x^2$, що описують роботу лінійного і квадратичного детекторів, являють собою нелінійні детерміновані функції, а функція $y = \sin(\omega x + \phi)$, де ω – детермінований параметр, ϕ – випадковий параметр, x – аргумент, являє собою нелінійну функцію з випадковим параметром ϕ .

Зауваження. У цьому розділі розглядаються детерміновані лінійні і нелінійні функції випадкових величин.

При перетворенні випадкових величин формуються нові випадкові величини. Ймовірнісні і неймовірнісні характеристики нових випадкових величин у загальному випадку відрізняються від відповідних характеристик початкових випадкових величин. Знаючи вид функції і характеристики початкових випадкових величин, можна розрахувати характеристики випадкових величин, що одержуються у результаті перетворення. Розглянемо методику такого розрахунку на прикладі скалярної випадкової величини.

1.4.2. Перетворення скалярної випадкової величини

Теорема 1. Нехай випадкова величина X з щільністю ймовірностей $f_1^x(x)$ підлягає перетворенню $y = \varphi(x)$, що має *однозначну* обернену диференційовану функцію $x = \eta(y)$. У цьому випадку щільність ймовірностей $f_1^y(y)$ випадкової величини Y описується співвідношенням

$$f_1^y(y) = f_1^x(\eta(y)) \left| \frac{d\eta(y)}{dy} \right|. \quad (1.4.1)$$

Доведення. Оскільки випадкові величини X і Y пов'язані однозначно детермінованою залежністю, то ймовірність перебування випадкової величини X в інтервалі $x \leq X \leq x + dx$ дорівнює ймовірності перебування випадкової величини Y в інтервалі $y \leq Y \leq y + dy$, тобто $f_1^y(y)|dy| = f_1^x(x)|dx|$. Звідси випливає співвідношення (1.4.1).

Теорема 2. Нехай випадкова величина X з щільністю ймовірностей $f_1^x(x)$ підлягає перетворенню $y = \varphi(x)$, обернена функція якого має Q гілок і описується диференційованими функціями $x = \eta_q(y)$, $q = \overline{1, Q}$. У цьому випадку щільність ймовірностей $f_1^y(y)$ випадкової величини Y виражається через щільність ймовірностей $f_1^x(x)$ випадкової величини X таким чином:

$$f_1^y(y) = \sum_{q=1}^Q f_1^x(\eta_q(y)) \left| \frac{d\eta_q(y)}{dy} \right|. \quad (1.4.2)$$

Співвідношення (1.4.2) доводиться, як і співвідношення (1.4.1).

Положення 1. Нехай випадкова величина X з щільністю ймовірностей $f_1^x(x)$ підлягає перетворенню $y = \varphi(x)$. Тоді початкова моментна функція m_ν^y ν -го порядку описується виразом

$$m_\nu^y = M[Y^\nu] = \int_{-\infty}^{\infty} \varphi^\nu(x) f_1^x(x) dx, \quad (1.4.3)$$

який випливає з визначення математичного сподівання функції.

Зауваження 1. З співвідношення (1.4.3) випливає, що математичне сподівання випадкової величини Y описується співвідношенням

$$m_y = \int_{-\infty}^{\infty} \varphi(x) f_1^x(x) dx. \quad (1.4.4)$$

Зауваження 2. З співвідношення (1.4.4) випливає, що, якщо функція $y = \varphi(x)$ – лінійна ($y = ax + b$), то математичне сподівання m_y випадкової величини Y зв'язано з математичним сподіванням m_x випадкової величини X співвідношенням

$$m_y = am_x + b. \quad (1.4.5)$$

Положення 2. Нехай випадкова величина X з щільністю ймовірностей $f_1^x(x)$ підлягає перетворенню $y = \varphi(x)$. Тоді центральна моментна функція μ_ν^y ν -го порядку описується виразом

$$\mu_\nu^y = M[(Y - m_y)^\nu] = \int_{-\infty}^{\infty} [\varphi(x) - m_y]^\nu f_1^x(x) dx. \quad (1.4.6)$$

Зауваження 1. З співвідношення (1.4.6) випливає, що дисперсія випадкової величини Y описується співвідношенням

$$D_y = M[(Y - m_y)^2] = \int_{-\infty}^{\infty} [\varphi(x) - m_y]^2 f_1^x(x) dx. \quad (1.4.7)$$

Зауваження 2. З співвідношення (1.4.7) випливає, що коли функція $y = \varphi(x)$ – лінійна ($y = ax + b$), то дисперсія D_y випадкової величини Y пов'язана з дисперсією D_x випадкової величини X співвідношенням

$$D_y = a^2 D_x. \quad (1.4.8)$$

Зміна закону розподілу випадкової величини при нелінійному перетворенні широко використовується для одержання випадкових величин із заданим законом розподілу.

Розглянемо методику розв'язання цієї задачі на найпростішому прикладі.

Приклад. Є випадкова величина X з рівномірним законом розподілу на інтервалі $(a, b]$ ($f_1^x(x) = 1/(b-a)$). Потрібно знайти перетворення $y = \varphi(x)$, що перетворює цю випадкову величину в неперервну випадкову величину Y з заданою функцією розподілу $F_1^y(y)$. При цьому функція $y = \varphi(x)$ монотонно зростаюча і має однозначну обернену функцію $x = \eta(y)$.

Розв'язок. На підставі виразу (1.4.1) щільність ймовірностей випадкової величини Y

$$f_1^y(y) = \frac{1}{b-a} \frac{dx}{dy}.$$

Інтегруючи цей вираз, маємо $x = c + (b-a)F_1^y(y)$, де c – константа. Її можна визначити з умов на границях ($F_1^y(-\infty) = 0$, $F_1^y(\infty) = 1$) як $c = a$. Звідси

$$F_1^y(y) = \frac{x-a}{b-a}$$

і

$$y = \varphi(x) = F_1^{-y}\left(\frac{x-a}{b-a}\right), \quad (1.4.9)$$

де $F_1^{-y}(y)$ – функція, обернена $F_1^y(y)$.

Зауваження 1. При комп'ютерному моделюванні часто необхідно мати масив чисел із заданим нерівномірним законом розподілу $F_1^y(y)$. Такий масив можна одержати шляхом перетворення масиву чисел, що генерується датчиком випадкових чисел із рівномірним на інтервалі $(0,1]$ законом розподілу, нелінійною функцією, що описується виразом (1.4.9).

Зауваження 2. Можна показати, що за допомогою перетворення

$$y = a + (b-a)F_1^x(x) \quad (1.4.10)$$

неперервна випадкова величина X з функцією розподілу $F_1^x(x)$ трансформується у випадкову величину Y з рівномірним законом розподілу на інтервалі $(a, b]$.

1.4.3. Перетворення векторної випадкової величини

Розглянемо більш складний випадок, коли векторна випадкова величина \vec{X} перетворюється у векторну випадкову величину \vec{Y} .

Теорема 1. Нехай H -вимірна випадкова величина \vec{X} з щільністю ймовірностей $f_H^x(\vec{x})$ підлягає нелінійному в загальному випадку перетворенню

випадкової величини $\vec{X} = (X_1, X_2)$, з двовимірною щільністю ймовірностей $f_2^x(x_1, x_2)$

Функція	Щільність ймовірностей $f_1^y(y)$
$y = x_1 + x_2$	$f_1^y(y) = \int_{-\infty}^{\infty} f_2^x(y - x_2, x_2) dx_2$
$y = x_1 - x_2$	$f_1^y(y) = \int_{-\infty}^{\infty} f_2^x(y + x_2, x_2) dx_2$
$y = x_1 \cdot x_2$	$f_1^y(y) = \int_{-\infty}^{\infty} f_2^x\left(\frac{y}{x_2}, x_2\right) \frac{dx_2}{ x_2 }$
$y = \frac{x_1}{x_2}$	$f_1^y(y) = \int_{-\infty}^{\infty} f_2^x(yx_2, x_2) x_2 dx_2$

Розглянемо моменти випадкової величини. У ряді випадків вони можуть бути досить просто розраховані, минаючи етап обчислення щільності ймовірностей.

Приклад. Нехай є лінійне перетворення $y = \sum_{q=1}^Q \alpha_q x_q$, де α_q – константи. Для цього перетворення неважко розрахувати моменти випадкової величини. Результати розрахунків наведено в табл. 1.4.3.

Таблиця 1.4.3

Залежність математичного сподівання m_y і дисперсії D_y випадкової величини Y від математичних сподівань m_{x_q} і дисперсій D_{x_q} випадкових величин X_q для перетворення $y = \sum_{q=1}^Q \alpha_q x_q$

Назва моментної функції	Співвідношення
Математичне сподівання	$m_y = \sum_{q=1}^Q \alpha_q m_{x_q}$
Дисперсія	$D_y = \sum_{q=1}^Q \alpha_q^2 D_{x_q} + \sum_{q \neq h} \alpha_q \alpha_h R_{x_q x_h}$

Примітка. $R_{x_q x_h}$ – змішаний центральний момент другого порядку.

Зауваження 1. При розрахунку суми випадкових величин X_1, X_2 математичне сподівання результату Y дорівнює сумі математичних сподівань додатків, а при розрахунку різниці випадкових величин X_1, X_2 – їхньої різниці.

Зауваження 2. При відсутності кореляції між випадковими величинами X_q ($q = \overline{1, Q}$) дисперсія випадкової величини Y істотно спрощується:

$$D_y = \sum_{q=1}^Q \alpha_q^2 D_{x_q}.$$

Зауваження 3. При розрахунку суми або різниці некорельованих випадкових величин дисперсія результату дорівнює сумі дисперсій.

1.5. Приклади випадкових величин

1.5.1. Гауссівська випадкова величина

На практиці дуже часто доводиться мати справу з гауссівськими випадковими величинами, як скалярними, так і векторними. Зупинимося на них більш докладно. Почнемо зі скалярного випадку.

Положення 1. Формули, що описують основні ймовірнісні характеристики скалярної гауссівської випадкової величини, наведено в табл. 1.5.1. Щільність ймовірностей гауссівської випадкової величини при різних СКВ зображено на рис. 1.5.1.

Таблиця 1.5.1

Ймовірнісні характеристики скалярної гауссівської випадкової величини

Назва характеристики	Аналітичний вираз
Щільність ймовірностей	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right], \quad x \in (-\infty, \infty), \quad \sigma > 0$
Функція розподілу	$F(x) = \Phi\left(\frac{x-m}{\sigma}\right),$ де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$ – інтеграл ймовірності
Характеристична функція	$Q(j\omega) = \exp\left(jm\omega - \frac{\sigma^2\omega^2}{2}\right)$

Примітки. 1. У поданих формулах m – математичне сподівання, а σ – СКВ.

2. Таблиця інтегралів ймовірності наведено у додатку 7.

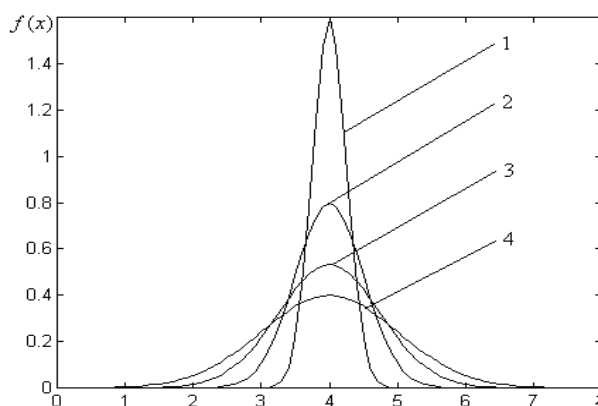


Рис. 1.5.1. Щільність ймовірностей гауссівського розподілу. Математичне сподівання $m = 4$, СКВ $\sigma = 0,25; 0,5; 0,75; 1$ (відповідно криві 1–4)

Зауваження 1. Криві щільності ймовірностей гауссівської випадкової величини мають одну моду, яка збігається з математичним сподіванням m , а СКВ дорівнює параметру σ . На відстані $\pm 3\sigma$ від математичного сподівання m щільність ймовірностей падає до величини 0,4% від максимуму. При цьому площа під кривою $f(x)$ на інтервалі $m \pm 3\sigma$ дорівнює 99,7%. Таким чином, випадкова величина, що має гауссівський закон розподілу, з ймовірністю 0,997 знаходиться в інтервалі $(m - 3\sigma, m + 3\sigma)$.

Зауваження 2. При нульовому математичному сподіванні ($m = 0$) і СКВ $\sigma \rightarrow \infty$ крива щільності ймовірностей зливається з віссю абсцис, а при $\sigma \rightarrow 0$ вироджується у дельта-функцію.

Зауваження 3. Для гауссівської скалярної випадкової величини X справедливі співвідношення:

$$\frac{d^n M[g(X)]}{dD^n} = 2^{-n} M\left[\frac{d^{2n} g(X)}{dX^{2n}}\right];$$

$$M[Xg(X)] = mM[g(X)] + DM\left[\frac{dg(X)}{dX}\right],$$

де $g(X)$ – довільна функція випадкової величини X , а D – дисперсія цієї величини.

Зауваження 4. За допомогою наведених формул ітераційно можна розрахувати момент будь-якого порядку. Для цього треба покласти $g(X) = X^\nu$.

Визначення 1. Дві випадкові величини X_1 і X_2 називаються *спільно гауссівськими*, якщо їхня спільна щільність ймовірностей має такий вигляд:

$$f_2(x_1, x_2) = C \exp[-P_2(x_1, x_2)], \quad (1.5.1)$$

де C – коефіцієнт, що нормує, ($C > 0$), $P_2(x_1, x_2)$ – квадратична форма, що додатньо напіввизначена для довільних x_1 і x_2 , тобто поліном другого степеня, що при будь-яких x_1 і x_2 приймає невід'ємні значення ($P_2(x_1, x_2) \geq 0$).

Положення 1. Для спільно гауссівських випадкових величин X_1, X_2 з урахуванням властивостей щільності ймовірностей багатовимірних випадкових величин вираз (1.5.1) може бути записаний так:

$$f_2(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \times \exp\left[-\frac{\sigma_2^2(x_1 - m_1)^2 - 2r\sigma_1\sigma_2(x_1 - m_1)(x_2 - m_2) + \sigma_1^2(x_2 - m_2)^2}{2\sigma_1^2\sigma_2^2(1-r^2)}\right]. \quad (1.5.2)$$

При цьому характеристична функція має вигляд

$$Q_2(j\omega_1, j\omega_2) = \exp\left(jm_1\omega_1 + jm_2\omega_2 - \frac{1}{2}(\sigma_1^2\omega_1^2 + 2r\sigma_1\sigma_2\omega_1\omega_2 + \sigma_2^2\omega_2^2)\right), \quad (1.5.3)$$

де m_1 і m_2 – математичні сподівання випадкових величин X_1 і X_2 , σ_1 і σ_2 – відповідні СКВ, r – коефіцієнт кореляції цих величин.

Зауваження 1. Використовуючи матричну символіку, вираз (1.5.2) можна записати в більш компактному вигляді:

$$f_2(x_1, x_2) = \frac{1}{2\pi\sqrt{|R|}} \exp\left[-\frac{1}{2}(\bar{x} - \bar{m})^T R^{-1}(\bar{x} - \bar{m})\right], \quad (1.5.4)$$

де $\bar{x} = (x_1, x_2)^T$ – вектор-стовпець, що описує координати випадкових величин X_1 і X_2 , $\bar{m} = (m_1, m_2)^T$ – вектор-стовпець, що описує математичні сподівання випадкових величин X_1 і X_2 , T – оператор транспонування, R^{-1} – матриця, обернена кореляційної матриці

$$R = \begin{vmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{vmatrix},$$

а $|R|$ – визначник матриці R .

Зауваження 2. Характеристична функція двох спільно гауссівських випадкових величин задовольняє диференційному рівнянню

$$\frac{\partial^n Q_2(j\omega_1, j\omega_2)}{\partial r^n} = (-\sigma_1\sigma_2\omega_1\omega_2)^n Q_2(j\omega_1, j\omega_2).$$

Дві спільно гауссівські випадкові величини мають властивості, що наведено нижче.

Властивість 1. Якщо спільно гауссівські випадкові величини X_1 і X_2 некорельовані, то вони незалежні.

Зауваження. Для спільно гауссівських випадкових величин із незалежності впливає їхня некорельованість, а з некорельованості – їхня незалежність (див. зауваження до визначення 9 підрозділу 1.2).

Властивість 2. Для спільно гауссівських випадкових величин X_1 і X_2 кривої сталої щільності ймовірностей є еліпс (еліпс розсіювання) (рис. 1.5.2), що описується рівнянням

$$\frac{(x_1 - m_1)^2}{\sigma_1^2} - 2r \frac{(x_1 - m_1)(x_2 - m_2)}{\sigma_1\sigma_2} + \frac{(x_2 - m_2)^2}{\sigma_2^2} = (1 - r^2)C^2, \quad C > 0 \quad (1.5.4)$$

або в матричній символіці

$$(\bar{x} - \bar{m})^T R^{-1} (\bar{x} - \bar{m}) = C^2, \quad C > 0. \quad (1.5.5)$$

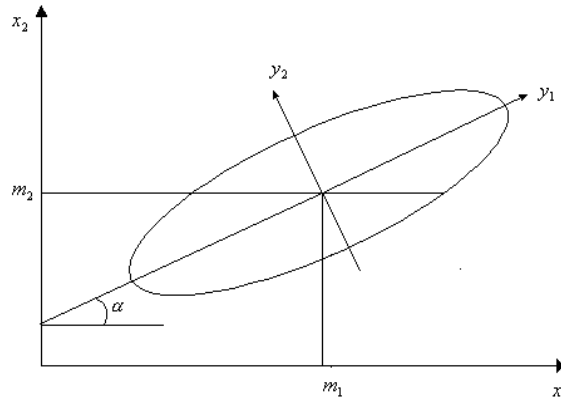


Рис. 1.5.2. Еліпс розсіювання для гауссівських випадкових величин X_1 і X_2

Зауваження 1. Центр еліпса розсіювання розташований у точці (m_1, m_2) , що описується математичними сподіваннями m_1 , m_2 , а розміри осей еліпса визначаються СКВ σ_1 , σ_2 , коефіцієнтом кореляції r і величиною параметра C . Нахил осей еліпса щодо осей координат залежить від СКВ σ_1 , σ_2 і коефіцієнта кореляції r . Кут нахилу α однієї з осей еліпса описується виразом

$$\alpha = \begin{cases} \frac{1}{2} \arctg\left(\frac{2r\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}\right) & \text{при } \sigma_1 \neq \sigma_2, \\ \pi/4 & \text{при } \sigma_1 = \sigma_2. \end{cases} \quad (1.5.6)$$

Оскільки осі еліпса взаємно ортогональні, кут нахилу другої осі відрізняється від α на 90° .

Зауваження 2. З виразів (1.5.4) – (1.5.6) видно, що при різних значеннях параметра C утворюються різні за розміром еліпси розсіювання, але центри і орієнтація осей усіх цих еліпсів однакові.

Зауваження 3. З виразів (1.5.4), (1.5.5) випливає, що коли гауссівські випадкові величини мають нульові математичні сподівання, то центр еліпсів розсіювання знаходиться у початку координат, а якщо ці величини некорельовані ($r = 0$), то осі еліпса орієнтовані уздовж осей координат.

Зауваження 4. З виразу (1.5.6) видно, що при рівних СКВ випадкових величин ($\sigma_1 = \sigma_2$) незалежно від рівня кореляції між ними кут повороту осей еліпсів розсіювання однаковий: $\alpha = \pi/4$. Якщо СКВ рівні ($\sigma_1 = \sigma_2$) і кореляція між випадковими величинами відсутня ($r = 0$), то еліпси вироджуються в коло радіуса σ .

Властивість 3. Дві корельовані (залежні) спільно гауссівські випадкові величини X_1 і X_2 приводяться до некорельованих (незалежних) спільно гауссівських випадкових величин Y_1 , Y_2 із нульовими математичними сподіваннями за допомогою такого лінійного перетворення:

$$\begin{aligned} Y_1 &= (X_1 - m_1) \cos \alpha + (X_2 - m_2) \sin \alpha, \\ Y_2 &= -(X_1 - m_1) \sin \alpha + (X_2 - m_2) \cos \alpha, \end{aligned} \quad (1.5.7)$$

де α визначається співвідношенням (1.5.6).

Властивість 4. При будь-яких лінійних перетвореннях спільно гауссівських випадкових величин утворюються також спільно гауссівські випадкові величини.

Властивість 5. Якщо дві випадкові величини спільно гауссівські, то кожна з випадкових величин гауссівська, проте якщо обидві випадкові величини гауссівські, то не обов'язково вони спільно гауссівські.

Зауваження. Якщо обидві випадкові величини гауссівські, то вони спільно гауссівські тільки тоді, коли вони незалежні.

Властивість 6. Якщо випадкові величини X_1 і X_2 – спільно гауссівські, то умовна щільність ймовірностей однієї з цих величин при фіксованому значенні іншої величини являє собою гауссівську щільність ймовірностей, зокрема:

$$\begin{aligned} f_1(x_2 / X_1 = x_1) &= f_2(x_1, x_2) / f_1(x_1) = \\ &= \frac{1}{\sigma_2 \sqrt{2\pi(1-r^2)}} \exp \left\{ -\frac{1}{2\sigma_2^2(1-r^2)} \left[x_2 - m_2 - r \frac{\sigma_2}{\sigma_1} (x_1 - m_1) \right]^2 \right\}. \end{aligned} \quad (1.5.8)$$

Зауваження. З виразу (1.5.8) випливає, що умовне математичне сподівання m_{x_2/x_1} залежить від значення випадкової величини X_1 , а умовна дисперсія D_{x_2/x_1} – ні:

$$m_{x_2/x_1} = m_2 + r \frac{\sigma_2}{\sigma_1} (x_1 - m_1), \quad D_{x_2/x_1} = \sigma_2^2 (1-r)^2.$$

Положення 2. Наведені співвідношення, що стосуються двовимірних спільно гауссівських випадкових величин, узагальнюються на багатовимірні випадкові величини.

Визначення 2. Випадкові величини X_1, \dots, X_N називаються *спільно гауссівськими*, якщо їхня спільна щільність ймовірностей має такий вигляд:

$$f_N(x_1, \dots, x_N) = \frac{1}{(2\pi)^{N/2} |R|^{1/2}} \exp \left[-\frac{1}{2} (\bar{x} - \bar{m})^T R^{-1} (\bar{x} - \bar{m}) \right], \quad (1.5.9)$$

де $\vec{x} = (x_1, \dots, x_N)^T$ – вектор-стовпець, що описує координати випадкових величин X_1, \dots, X_N , $\vec{m} = (m_1, \dots, m_N)^T$ – вектор-стовпець, що описує математичні сподівання цих величин, R – кореляційна матриця цих величин:

$$R = \begin{pmatrix} R_{11} & \dots & R_{1N} \\ \dots & \dots & \dots \\ R_{N1} & \dots & R_{NN} \end{pmatrix}. \quad (1.5.10)$$

1.5.2. Розподіли, пов'язані з гауссівськими

Ряд розподілів, що часто зустрічаються на практиці, тісно пов'язані з гауссівськими розподілами. З деякими з них, а саме розподілами Релея і Райса, познайомимося у цьому підрозділі. Деякі інші розподіли, що утворюються з гауссівських при різноманітних нелінійних перетвореннях, будуть розглянуті нижче.

Визначення 1. Релеївським розподілом називається розподіл, щільність ймовірностей якого описується таким виразом:

$$f_1(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x > 0. \quad (1.5.11)$$

Зауваження 1. Функція розподілу, що відповідає щільності ймовірностей (1.5.11), має вигляд

$$F_1(x) = 1 - \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x > 0. \quad (1.5.12)$$

Зауваження 2. Математичне сподівання релеївського розподілу $m_x = \sigma \sqrt{\frac{\pi}{2}}$, дисперсія $D_x = 2\sigma^2 - m_x^2$, максимум щільності ймовірностей знаходиться в точці $x = \sigma$, а початковий момент третього порядку $m_3 = 3\sigma^3 \sqrt{\frac{\pi}{2}}$.

Зауваження 3. Уявлення про щільність ймовірностей розподілу Релея дає крива 1 на рис. 1.5.3.

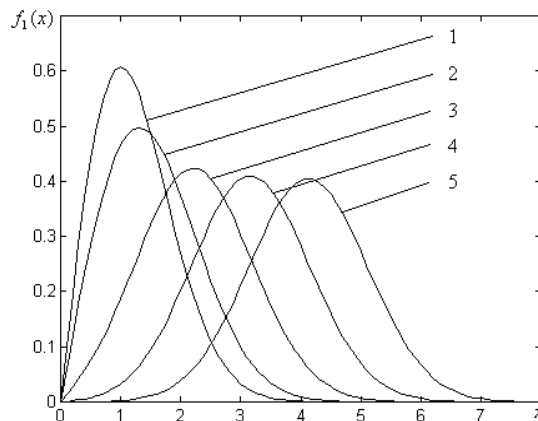


Рис. 1.5.3. Щільності ймовірностей розподілів Релея (1) і Райса (2 – 5). Криві 1 – 5 побудовані для $\sigma = 1$ і значень параметра $\alpha = \overline{0,4}$ відповідно

Теорема 1. Нехай у декартовій системі координат компоненти двовимірного випадкового вектора $\vec{X} = (X_1, X_2)$ є незалежними гауссівськими величинами з

нульовими математичними сподіваннями ($m_{x_1} = m_{x_2} = 0$) і однаковими дисперсіями σ^2 . Тоді довжина X вектора \vec{X} описується розподілом Релея, а кут повороту Φ цього вектора – рівномірним законом на інтервалі $(0, 2\pi]$.

Доведення. Нехай виконуються умови теореми. Тоді в декартовій системі координат двовимірна щільність ймовірностей

$$f_2(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x_1^2 + x_2^2}{2\sigma^2}\right). \quad (1.5.13)$$

Якобіан перетворення декартової системи координат у полярну систему координат

$$J_2(x, \varphi) = \begin{vmatrix} \cos \varphi & -x \sin \varphi \\ \sin \varphi & x \cos \varphi \end{vmatrix} = x.$$

Тому щільність ймовірностей координат вектора \vec{X} в полярній системі координат (x, φ) має вигляд

$$f_2(x, \varphi) = \frac{x}{2\pi\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (1.5.14)$$

При інтегруванні двовимірної щільності ймовірностей (1.5.14) по куту в границях $(0, 2\pi]$ і по x в границях $[0, \infty)$ утворюються шукані щільності ймовірностей довжини і кута повороту вектора \vec{X} у вигляді виразів (1.5.11) і $f_1(\varphi) = \frac{1}{2\pi}$ відповідно.

Зауваження. Неважко показати, що, коли дві незалежні випадкові величини X_1 й X_2 описуються гауссівськими законами розподілу з однаковими дисперсіями σ^2 і необов'язково нульовими математичними сподіваннями m_{x_1} і m_{x_2} , щільність ймовірностей перебування кінця радіус-вектора $\vec{X} = (X_1, X_2)$ на колі радіусом R із центром у точці (m_{x_1}, m_{x_2}) описується виразом (1.5.11).

Визначення 2. Узагальненим релеївським розподілом (розподілом Райса) називається розподіл, щільність ймовірностей якого

$$f_1(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + \alpha^2}{2\sigma^2}\right) I_0\left(\frac{\alpha x}{\sigma^2}\right), \quad x > 0, \quad (1.5.15)$$

де α – параметр «нецентральності», σ^2 – параметр масштабу ($\sigma > 0$), $I_0(*)$ – функція Бесселя нульового порядку уявного аргументу.

Зауваження 1. Функція розподілу, що відповідає щільності ймовірностей (1.5.15), описується інтегралом

$$F_1(x) = \int_0^x \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + \alpha^2}{2\sigma^2}\right) I_0\left(\frac{\alpha x}{\sigma^2}\right) dx, \quad (1.5.16)$$

який не виражається через табульовані функції.

Зауваження 2. Щільність ймовірностей розподілу Райса при малому відношенні $\frac{\alpha}{\sigma}$ добре апроксимується формулою

$$f_1(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + \alpha^2}{2\sigma^2}\right) \left(1 + \frac{\alpha^2 x^2}{4\sigma^4}\right), \quad x > 0,$$

що утворюється при використанні перших двох членів розкладення функції Бесселя в степінний ряд, а при $\alpha \gg \sigma$ – асимптотичною формулою

$$f_1(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\alpha)^2}{2\sigma^2}\right) \left(1 + \frac{\sigma^2}{8\alpha x}\right) \sqrt{\frac{x}{\alpha}}, \quad x > 0.$$

Зауваження 3. Математичне сподівання розподілу Райса

$$m_x = \sigma\sqrt{\frac{\pi}{2}} \left[\left(1 + \frac{\alpha^2}{2\sigma^2}\right) I_0\left(\frac{\alpha^2}{4\sigma^2}\right) + \frac{\alpha^2}{2\sigma^2} I_1\left(\frac{\alpha^2}{4\sigma^2}\right) \right] \exp\left(-\frac{\alpha^2}{4\sigma^2}\right), \quad (1.5.17)$$

дисперсія

$$D_x = 2\sigma^2 + \alpha^2 - m_x^2, \quad (1.5.18)$$

де $I_1(*)$ – функція Бесселя першого порядку уявного аргументу.

Зауваження 4. При $\alpha \gg \sigma$ розподіл Райса наближається до гауссівського з математичним сподіванням $m_x \approx \alpha$ і дисперсією $D_x \approx \sigma^2$. Уявлення про щільність ймовірностей розподілу Райса дає рис. 1.5.3 (криві 2 – 5).

Теорема 2. Нехай у декартовій системі координат компоненти двовимірного випадкового вектора $\vec{X} = (X_1, X_2)$ є незалежними гауссівськими величинами з математичними сподіваннями m_{x_1} , m_{x_2} і однаковими дисперсіями σ^2 . Тоді довжина X вектора \vec{X} описується розподілом Райса з параметрами $\alpha = \sqrt{m_{x_1}^2 + m_{x_2}^2}$, σ^2 , а кут повороту Φ цього вектора – розподілом із щільністю ймовірностей

$$f_1(\phi) = \frac{1}{2\pi} \exp\left(-\frac{\alpha^2}{2\sigma^2}\right) + \frac{\alpha \cos(\phi - \phi_0)}{\sigma\sqrt{2\pi}} F\left[\frac{\alpha}{\sigma} \cos(\phi - \phi_0)\right] \times \\ \times \exp\left[-\frac{\alpha^2}{2\sigma^2} \sin^2(\phi - \phi_0)\right], \quad |\phi - \phi_0| \leq \pi,$$

де $\phi_0 = \arctg \frac{m_{x_2}}{m_{x_1}}$, $F[*]$ – інтеграл ймовірності.

Відомості про книги з тематики розділу 1 наведено в табл. 1.5.2.

Таблиця 1.5.2

Література до розділу 1

Характер книг	Номери книг у списку літератури
Довідники	24, 40, 47, 49
Книги навчального плану для інженерів	3, 16, 25, 32, 69, 74, 75, 84, 100
Книги з математичним ухилом	5, 10, 15, 20, 22, 61, 88, 89, 105
Книги прикладного характеру	55, 56, 66, 75, 82, 91, 96, 97

2. Основні положення теорії випадкових функцій

2.1. Уявлення про випадкові функції

2.1.1. Визначення і класифікація

Визначення 1. Випадковою функцією $X(t)$ називається числова функція незалежного аргументу, значення якої при будь-якому фіксованому значенні аргументу $t \in T$ (де T – область визначення аргументу) є випадковою величиною. Ця випадкова величина називається *перерізом* випадкової функції. Множина S значень усіх перерізів випадкової функції утворює *простір станів* (фазовий простір).

Визначення 2. i -ю *реалізацією* випадкової функції $X(t)$ (*вибірковою функцією*) називається не випадкова (детермінована) функція $x_i(t)$ ($i = 1, 2, \dots$), що для кожного i -го досліду ставить у відповідність кожному $t \in T$ одне із значень $x \in S$.

Зауваження 1. Випадкова функція поєднує в собі риси як випадкової величини, так і звичайної не випадкової функції. При фіксації значення аргументу вона перетворюється у випадкову величину,; при фіксації досліду – в реалізацію випадкової функції.

Зауваження 2. Будь-яка випадкова функція може бути представлена множиною (ансамблем) нескінченної кількості реалізацій, причому навіть у тому випадку, коли за умовами задачі кількість реалізацій обмежена. Перехід від обмеженого числа реалізацій до необмеженого забезпечується припущенням, що реалізації можуть повторюватися нескінченне число разів.

Зауваження 3. При нескінченній кількості різних реалізацій зобразити їх усі на папері неможливо. Графічно можна представити лише окремі реалізації з ансамблю. Приклад такого графічного опису випадкової функції наведено на рис. 2.1.1.

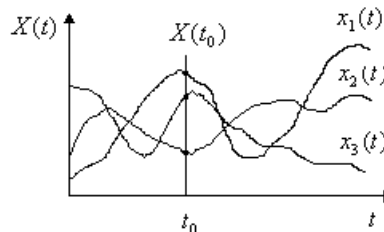


Рис. 2.1.1. Реалізації випадкової функції $X(t)$

Приклад. Проводиться стрільба ракетами по цілі. Відхилення траєкторії польоту ракети від розрахункової траєкторії можна розглядати як випадкову функцію $X(t)$. При цьому відхилення траєкторії $x_i(t)$ кожної i -ї ракети являє собою i -у реалізацію випадкової функції $X(t)$. В результаті відслідковування польоту кількох ракет можна одержати сімейство реалізацій $x_i(t)$ ($i = 1, 2, \dots$) випадкової функції $X(t)$. Величина відхилення траєкторії ракети у фіксований момент часу t_0 являє собою випадкову величину (переріз $X(t_0)$ випадкової функції). Величина відхилення траєкторії $x_i(t_0)$ кожної i -ї ракети у момент часу t_0 є i -м значенням перерізу $X(t_0)$.

Положення. Складність та різноманітність випадкових функцій робить доцільним їх упорядкування. Ознаки, за якими звичайно класифікують випадкові функції, пов'язані з особливостями простору станів S , області T визначення аргументу, а також характеру зміни реалізацій. У табл. 2.1.1 і 2.1.2

наведені деякі класи випадкових функцій (ВФ), що використовуються на практиці і мають спеціальні назви.

Таблиця 2.1.1

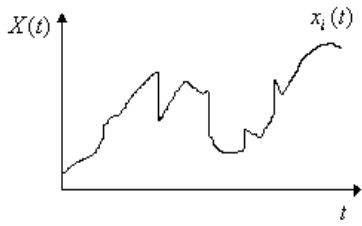
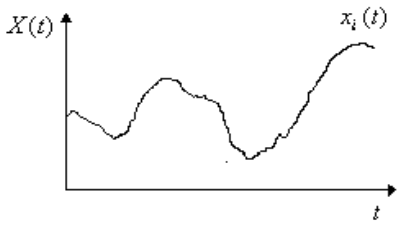
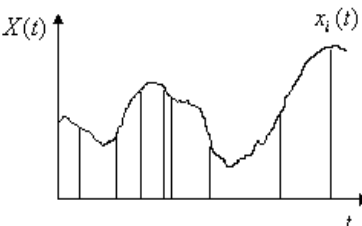
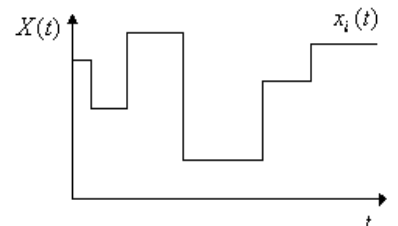
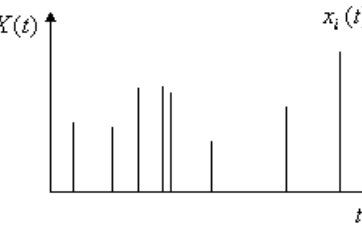
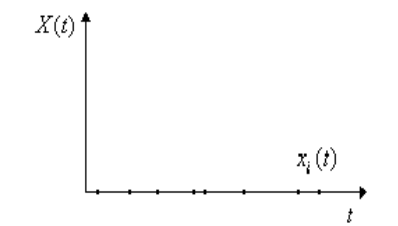
Класифікація випадкових функцій (варіант 1)

Ознака класифікації	Випадкова функція	
Тип простору станів S	Дійсна ($S \in R^N$)	Комплексна ($S \in Z^N$)
Вимірність N простору станів S	Скалярна ($S \in R$ або $S \in Z$)	Векторна ($S \in R^N$ або $S \in Z^N$)
Вимірність області визначення аргументу T	Випадковий (стохастичний) процес ($T \in R$)	Випадкове поле ($T \in R^N$)

Примітка. R – множина дійсних чисел, Z – множина комплексних чисел.

Таблиця 2.1.2

Класифікація випадкових функцій (варіант 2)

 <p>1. Неперервнозначна ВФ (з розривами 1-го роду)</p>	 <p>2. Неперервна ВФ в області станів S (без розривів)</p>
 <p>3. Випадкова послідовність (S – неперервна, T – дискретна)</p>	 <p>4. Дискретна ВФ (S – дискретна, T – неперервна)</p>
 <p>5. Дискретна випадкова послідовність (S і T дискретні)</p>	 <p>6. Точкова ВФ (випадкова течія)</p>

Особливу увагу належить звернути на скалярні дійсні випадкові (стохастичні) процеси, які будуть детально розглянуті в подальшому.

Графічний засіб опису випадкових функцій, незважаючи на наочність, дає лише поверхове уявлення про випадкову функцію. Більш повний і точний опис дозволяє одержати аналітичний підхід. Зупинимося на ньому більш докладно.

2.1.2. Ймовірнісні характеристики випадкових функцій

Положення. Випадкову функцію можна розглядати як N -вимірну векторну випадкову величину. При цьому для випадкової функції можна визначити поняття *функції розподілу ймовірності* $F_N(\vec{x}; \vec{t})$, *щільності розподілу ймовірності* $f_N(\vec{x}; \vec{t})$ і *характеристичної функції* $Q_N(j\vec{\omega}; \vec{t})$ так, як це зроблено в табл. 2.1.3.

Таблиця 2.1.3

Основні ймовірнісні характеристики випадкової функції

Назва характеристики	Математичне трактування
Функція розподілу ймовірності	$F_N(\vec{x}; \vec{t}) = P\{X(t_1) < x_1, \dots, X(t_N) < x_N\};$ $F_N(\vec{x}; \vec{t}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_N} f_N(\vec{x}; \vec{t}) d\vec{x}$
Щільність розподілу ймовірності	$f(\vec{x}; \vec{t}) = \frac{\partial^N F(\vec{x}; \vec{t})}{\partial x_1 \dots \partial x_N};$ $f_N(\vec{x}; \vec{t}) = \frac{1}{(2\pi)^N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} Q_N(j\vec{\omega}; \vec{t}) e^{-j\vec{\omega}\vec{x}} d\vec{\omega}$
Характеристична функція	$Q_N(j\vec{\omega}; \vec{t}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_N(\vec{x}; \vec{t}) e^{j\vec{\omega}\vec{x}} d\vec{x}$

Зауваження. Математична коректність такого представлення випадкової функції забезпечується при спрямованості N до нескінченності.

Властивості. Ймовірнісні характеристики випадкових функцій мають ті самі властивості, що і відповідні характеристики векторних випадкових величин. Основні властивості ймовірнісних характеристик випадкових функцій наведено в табл. 2.1.4 – 2.1.6.

Таблиця 2.1.4

Основні властивості функції розподілу ймовірності

Властивість	Математичне трактування
Невід'ємність	$F_N(\vec{x}; \vec{t}) \geq 0$
Обмеженість	$0 \leq F_N(\vec{x}; \vec{t}) \leq 1$
Монотонність	$F_N(\vec{x}; \vec{t})$ неспадна і неперервна зліва по кожній координаті x_n
Симетричність	$F_N(\vec{x}; \vec{t})$ не змінюється при будь-яких перестановках пар (x_n, t_n)
Узгодженість	$F_M(x_1, \dots, x_M; t_1, \dots, t_M) = F_N(x_1, \dots, x_M, \infty, \dots, \infty; t_1, \dots, t_N)$ при $M < N$

Таблиця 2.1.5

Основні властивості щільності ймовірностей

Властивість	Математичне трактування
Невід'ємність	$f_N(\vec{x}; \vec{t}) \geq 0$
Нормованість до одиниці	$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_N(\vec{x}; \vec{t}) d\vec{x} = 1$
Симетричність	$f_N(\vec{x}; \vec{t})$ не змінюється при будь-яких перестановках пар (x_n, t_n)
Узгодженість	$f_M(x_1, \dots, x_M; t_1, \dots, t_M) =$ $= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_N(x_1, \dots, x_N; t_1, \dots, t_N) dx_{M+1} \dots dx_N$ при $M < N$

Таблиця 2.1.6

Основні властивості характеристичної функції

Властивість	Математичне трактування
Обмеженість	$ Q_N(j\vec{\omega}; \vec{t}) \leq 1, \quad Q_N(0; \vec{t}) = 1$
Симетричність	$Q_N(j\vec{\omega}; \vec{t})$ не змінюється при будь-яких перестановках пар $(j\omega_n, t_n)$
Узгодженість	$Q_M(j\omega_1, \dots, j\omega_M; t_1, \dots, t_M) =$ $= Q_N(j\omega_1, \dots, j\omega_M, 0, \dots, 0; t_1, \dots, t_N)$ при $M < N$

2.1.3. Умовні ймовірнісні характеристики випадкових функцій

Положення. Уявимо випадкову функцію $X(t)$ сукупністю її перерізів $X(t_1), \dots, X(t_N)$ (рис. 2.1.2). Розділимо ці перерізи на дві групи. До першої групи віднесемо будь-які M перерізів, а до другої – $(N - M)$ перерізів, що залишилися, (на рисунку для простоти до першої групи віднесені M перших перерізів, а до другої – $(N - M)$ останніх). Тоді випадкова функція $X(t)$ може бути описана безумовною щільністю ймовірності $f_N(\vec{x}; \vec{t})$, що подається у вигляді добутку безумовної щільності ймовірностей $f_M(x_1, \dots, x_M; t_1, \dots, t_M)$ перерізів першої групи й умовної щільності ймовірностей $f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M)$ перерізів другої групи:

$$f_N(x_1, \dots, x_N; t_1, \dots, t_N) = f_M(x_1, \dots, x_M; t_1, \dots, t_M) \times \\ \times f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M).$$

Визначення 1. Перерізи t_1, t_2, \dots, t_N випадкової функції називаються *незалежними*, якщо випадкові величини, що відповідають цим перерізам, незалежні.

Визначення 2. Перерізи t_1, t_2, \dots, t_N випадкової функції називаються *незалежними в сукупності (взаємно незалежними)*, якщо випадкові величини, що відповідають цим перерізам, незалежні у сукупності, тобто

$$f_N(\vec{x}; \vec{t}) = \prod_{n=1}^N f_1(x_n; t_n), \quad F_N(\vec{x}; \vec{t}) = \prod_{n=1}^N F_1(x_n; t_n) \quad \text{або} \quad Q_N(j\vec{\omega}; \vec{t}) = \prod_{n=1}^N Q_1(j\omega_n; t_n).$$

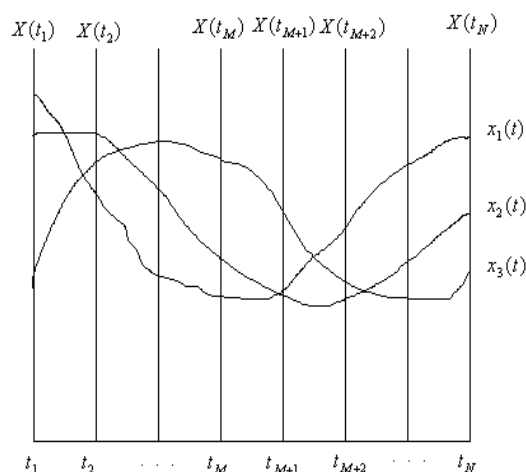


Рис. 2.1.2. Перерізи випадкової функції $X(t)$

Таким чином, вичерпною характеристикою випадкової функції з незалежними у сукупності перерізами є одновимірні функції розподілу, одновимірні щільності розподілу або одновимірні характеристичні функції.

Зауваження 1. Говорячи про незалежність перерізів випадкової функції, звичайно розуміють їхню незалежність у сукупності.

Зауваження 2. Якщо перерізи випадкової функції незалежні, то

$$f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) = f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N)$$

і

$$f_N(x_1, \dots, x_N; t_1, \dots, t_N) = f_M(x_1, \dots, x_M; t_1, \dots, t_M) f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N).$$

Властивості. Властивості умовних ймовірнісних характеристик наведено в табл. 2.1.7 – 2.1.9.

Таблиця 2.1.7

Основні властивості умовної функції розподілу

Властивість	Математичне трактування
Невід'ємність	$F_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) \geq 0$
Обмеженість	$0 \leq F_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) \leq 1$
Монотонність	$F_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M)$ неспадна і неперервна зліва по кожній координаті x_n ($n = \overline{M+1, N}$)
Симетричність	$F_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M)$ не змінюється при будь-яких перестановках пар аргументів (x_n, t_n) у межах лівої чи правої сторони від знака умови *)
Узгодженість	$F_{N-M-K}(x_{M+K+1}, \dots, x_N; t_{M+K+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) = F_{N-M}(\infty, \dots, \infty, x_{M+K+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M)$ при $K + M < N$

*) **Примітка.** При перестановці пар аргументів (x_n, t_n) із лівої сторони у праву і навпаки, з правої сторони у ліву, умовна функція розподілу змінюється.

Таблиця 2.1.8

Основні властивості умовної щільності ймовірностей

Властивість	Математичне трактування
Невід'ємність	$f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) \geq 0$
Нормованість до одиниці	$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) \times dx_{M+1} \dots dx_N = 1$
Симетричність	$f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M)$ не змінюється при будь-яких перестановках пар аргументів (x_n, t_n) у межах лівої чи правої сторони від знака умови *)
Узгодженість	$f_{N-M-K}(x_{M+K+1}, \dots, x_N; t_{M+K+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) \times dx_{M+1} \dots dx_{M+K}$ при $K + M < N$

*) *Примітка.* При перестановці пар аргументів (x_n, t_n) із лівої сторони у праву і навпаки, з правої сторони у ліву, умовна щільність змінюється.

Таблиця 2.1.9

Основні властивості умовної характеристичної функції

Властивість	Математичне трактування
Обмеженість	$ \mathcal{Q}_{N-M}(j\omega_{M+1}, \dots, j\omega_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) \leq \mathcal{Q}_{N-M}(0, \dots, 0; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) = 1$
Симетричність	$\mathcal{Q}_{N-M}(j\omega_{M+1}, \dots, j\omega_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M)$ не змінюється при перестановках пар $(j\omega_n, t_n)$ у межах лівої сторони від знака умови чи при перестановках пар (x_n, t_n) у межах правої сторони від знака умови
Узгодженість	$\mathcal{Q}_{N-M-K}(j\omega_{M+K+1}, \dots, j\omega_N; t_{M+K+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) = \mathcal{Q}_{N-M}(0, \dots, 0, j\omega_{M+K+1}, \dots, j\omega_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M)$ при $K + M < N$

На практиці зручно користуватися такими правилами.

Правило 1. Частина аргументів умовної щільності ймовірностей випадкової функції, що розташовані ліворуч знака умови, можна виключати шляхом інтегрування по «зайвих» аргументах у нескінченних границях.

Зауваження. Це правило виражає властивість узгодженості умовної щільності ймовірностей.

Правило 2. Частина аргументів умовної щільності ймовірностей, що розташовані праворуч знака умови, можна виключати шляхом інтегрування по «зайвих» аргументах у нескінченних границях із вагою, що дорівнює умовній щільності ймовірностей «зайвих» аргументів при фіксованому значенні параметрів, що залишаються праворуч знака умови:

$$f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_K; t_1, \dots, t_K) =$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M \times \quad (2.1.1)$$

$$\times f_{M-K}(x_{K+1}, \dots, x_M; t_{K+1}, \dots, t_M / x_1, \dots, x_K; t_1, \dots, t_K) dx_{K+1} \dots dx_M$$

$$(0 < K < M < N).$$

Доведення цього правила базується на властивості узгодженості щільності ймовірностей випадкової функції.

Визначення 3. Випадкова функція називається умовною, якщо на значення цієї функції в деяких перерізах накладені обмеження (умови). Умовні випадкові функції позначаються таким чином: $X(t / X(t_1) = x(t_1), \dots, X(t_M) = x(t_M))$, $X(t / X(t_m) = x(t_m), m = \overline{1, M})$, $X(t / x_1, \dots, x_M; t_1, \dots, t_M)$ або $X(t / x(t_m), m = \overline{1, M})$.

Зауваження 1. Умовна випадкова функція $X(t / X(t_m) = x(t_m), m = \overline{1, M})$ може розглядатися як функція, яка побудована з безумовної випадкової функції $X(t)$ при одержанні даних про значення випадкової функції в точках t_1, \dots, t_M . У цьому випадку щільність ймовірностей $f_N(x_1, \dots, x_N; t_1, \dots, t_N)$ випадкової функції $X(t)$ називається *апостеріорною*, а умовна щільність ймовірностей

$$f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M)$$

умовної випадкової функції $X(t / X(t_m) = x(t_m), m = \overline{1, M})$ – *апостеріорною*.

Зауваження 2. Основною областю застосування апарата умовних ймовірнісних характеристик є опис випадкових функцій із залежними перерізами.

2.2. Неймовірнісні характеристики випадкових функцій

2.2.1. Моментні і кумулянтні функції випадкових функцій

Визначення 1. Початковою N -вимірною моментною функцією порядку $\nu = \nu_1 + \dots + \nu_N$ випадкової функції $X(t)$ називається функція

$$m_{\nu_1 \dots \nu_N}(t_1, \dots, t_N) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_1^{\nu_1} \dots x_N^{\nu_N} f_N(x_1, \dots, x_N; t_1, \dots, t_N) dx_1 \dots dx_N =$$

$$= M[X^{\nu_1}(t_1) \dots X^{\nu_N}(t_N)], \quad (2.2.1)$$

де ν_n – ціле додатне число ($n = \overline{1, N}$).

Визначення 2. Математичним сподіванням випадкової функції $X(t)$ називається одновимірна початкова моментна функція першого порядку

$$m_x(t) = m_1(t) = M[X(t)] = \int_{-\infty}^{\infty} x f_1(x; t) dx. \quad (2.2.2)$$

Визначення 3. Центрованою випадковою функцією $\overset{0}{X}(t)$ називається випадкова функція

$$\overset{0}{X}(t) = X(t) - m_x(t). \quad (2.2.3)$$

Визначення 4. Центральною N -вимірною моментною функцією порядку $\nu = \nu_1 + \dots + \nu_N$ випадкової функції $X(t)$ називається початкова N -вимірна моментна функція порядку $\nu = \nu_1 + \dots + \nu_N$ відповідної центрованої випадкової функції $\overset{0}{X}(t)$:

$$\mu_{\nu_1 \dots \nu_N}(t_1, \dots, t_N) = M[\overset{0}{X}^{\nu_1}(t_1) \dots \overset{0}{X}^{\nu_N}(t_N)]. \quad (2.2.4)$$

Зауваження. Поряд із початковими і центральними моментними функціями іноді використовують *абсолютні початкові і центральні моментні функції*, які визначаються відповідно як

$$M[|X(t_1)|^{v_1} \dots |X(t_N)|^{v_N}] \text{ і } M[|X(t_1)|^{v_1} \dots |X(t_N)|^{v_N}].$$

Моментні функції мають важливі властивості, що часто використовуються на практиці.

Властивість 1. Початкові і центральні моментні функції є не випадковими функціями аргументів t_1, \dots, t_N .

Властивість 2. Початкові і центральні моментні функції симетричні щодо аргументів.

Властивість 3. Будь-яка ймовірнісна функція (функція розподілу, щільність ймовірностей або характеристична функція) вимірності N цілком визначається нескінченною послідовністю усіх моментних функцій вимірності до N включно. Зокрема, одновимірна характеристична функція $Q_1(j\omega; t)$ цілком визначається нескінченною послідовністю усіх одновимірних моментних функцій. При цьому

$$Q_1(j\omega; t) = 1 + \sum_{v=1}^{\infty} \frac{(j\omega)^v}{v!} m_v(t). \quad (2.2.5)$$

Доведення. За визначенням характеристичної функції

$$Q_1(j\omega, t) = \int_{-\infty}^{\infty} e^{j\omega x} f_1(x; t) dx. \quad (2.2.6)$$

Розкладемо експоненту в степінний ряд:

$$e^z = 1 + \sum_{v=1}^{\infty} \frac{z^v}{v!}. \quad (2.2.7)$$

Підставивши вираз (2.2.7) у формулу (2.2.6), маємо

$$Q_1(j\omega; t) = \int_{-\infty}^{\infty} \left[1 + \sum_{v=1}^{\infty} \frac{(j\omega x)^v}{v!} \right] f_1(x; t) dx.$$

Інтегрування дає формулу (2.2.5).

Аналогічно доводиться більш загальне твердження, що стосується багатовимірного випадку:

$$Q_N(j\omega_1, \dots, j\omega_N; t_1, \dots, t_N) = \sum_{v_1, \dots, v_N=0}^{\infty} \frac{m_{v_1, \dots, v_N}(t_1, \dots, t_N)}{v_1! \dots v_N!} (j\omega_1)^{v_1} \dots (j\omega_N)^{v_N}. \quad (2.2.8)$$

Властивість 4. Моментні функції вимірності N цілком визначаються будь-якою ймовірнісною характеристикою вимірності N . Зокрема, початкова одновимірна моментна функція $m_v(t)$ будь-якого порядку v цілком визначається одновимірною характеристичною функцією. При цьому

$$m_v(t) = \frac{\partial^v Q_1(j\omega; t)}{j^v \partial \omega^v} \Big|_{\omega=0}. \quad (2.2.9)$$

Доведення. Розкладаючи характеристичну функцію $Q_1(j\omega; t)$ в ряд Маклорена по аргументу $j\omega$ в області нуля, маємо:

$$Q_1(j\omega; t) = 1 + \sum_{v=1}^{\infty} \frac{(j\omega)^v \partial^v Q_1(j\omega; t)}{v! j^v \partial \omega^v} \Big|_{\omega=0}. \quad (2.2.10)$$

Порівняння виразів (2.2.5) і (2.2.10) приводить до виразу (2.2.9).

Зауваження. Так само доводиться співвідношення

$$m_{v_1 \dots v_N}(t_1, \dots, t_N) = (-j)^{v_1 + \dots + v_N} \frac{\partial^{v_1 + \dots + v_N} Q_N(j\omega_1, \dots, j\omega_N; t_1, \dots, t_N)}{\partial \omega_1^{v_1} \dots \partial \omega_N^{v_N}} \Big|_{\omega_1 = \dots = \omega_N = 0}, \quad (2.2.11)$$

аналогічне (1.3.20).

Для опису випадкових функцій використовують також кумулянтні функції.

Визначення 5. Кумулянтними функціями називаються коефіцієнти розкладу $\kappa_{v_1 \dots v_N}(t_1, \dots, t_N)$ в ряд Маклорена логарифма характеристичної функції:

$$\ln Q_N(j\omega_1, \dots, j\omega_N; t_1, \dots, t_N) = \sum_{v_1, \dots, v_N=0}^{\infty} \frac{\kappa_{v_1 \dots v_N}(t_1, \dots, t_N)}{v_1! \dots v_N!} (j\omega_1)^{v_1} \dots (j\omega_N)^{v_N}, \quad (2.2.12)$$

де

$$\kappa_{v_1 \dots v_N}(t_1, \dots, t_N) = (-j)^{v_1 + \dots + v_N} \frac{\partial^{v_1 + \dots + v_N} \ln Q_N(j\omega_1, \dots, j\omega_N; t_1, \dots, t_N)}{\partial \omega_1^{v_1} \dots \partial \omega_N^{v_N}} \Big|_{\omega_1 = \dots = \omega_N = 0}. \quad (2.2.13)$$

Зауваження 1. На практиці часто представляють інтерес лише кумулянтні функції нижчого порядку. При цьому кумулянтні функції з $v_i = 1$ і $v_n = 0$ ($n = \overline{1, N}$, $n \neq i$) позначають $R_1(t_i)$, з $v_i = v_j = 1$ і $v_n = 0$ ($n = \overline{1, N}$, $n \neq i$, $n \neq j$) – $R_2(t_i, t_j)$, з $v_i = v_j = v_k = 1$ і $v_n = 0$ ($n = \overline{1, N}$, $n \neq i$, $n \neq j$, $n \neq k$) – $R_3(t_i, t_j, t_k)$ і т.д.

Примітка. Іноді всі ці функції називають *кореляційними функціями* випадкової функції, хоча частіше під *кореляційною функцією* випадкової функції розуміють тільки кумулянтну функцію $R_2(t_i, t_j)$.

Зауваження 2. Моментні функції однозначно виражаються через кумулянтні функції і, навпаки, кумулянтні функції – через моментні функції. Зокрема справедливі такі співвідношення:

$$m_1(t) = R_1(t);$$

$$m_{11}(t_1, t_2) = R_2(t_1, t_2) + R_1(t_1)R_1(t_2);$$

$$m_{111}(t_1, t_2, t_3) = R_3(t_1, t_2, t_3) + [R_1(t_1)R_2(t_2, t_3) + R_1(t_2)R_2(t_1, t_3) + R_1(t_3)R_2(t_1, t_2)] + R_1(t_1)R_1(t_2)R_1(t_3);$$

Зауваження 3. Моментна і кумулянтна функції будь-якого порядку вимірності N несуть менше інформації про випадкову функцію, ніж характеристична функція, функція розподілу або щільність ймовірностей тієї ж вимірності. При цьому ймовірнісні характеристики вимірності N в загальному випадку визначаються нескінченним числом моментних або кумулянтних функцій. Моментна ж функція будь-якого порядку ν вимірності N та кумулянтна (або кореляційна) функція будь-якого порядку ν вимірності N цілком визначаються будь-якою N -вимірною ймовірнісною характеристикою випадкової функції.

2.2.2. Одновимірні і двовимірні моментні функції випадкових функцій

Основними одновимірними характеристиками випадкової функції є математичне сподівання $m_x(t)$, дисперсія $D_x(t)$ і відхилення $\sigma_x(t)$ (СКВ).

Визначення математичного сподівання дано вище.

Визначення 1. Дисперсією $D_x(t)$ випадкової функції $X(t)$ називається центральна одновимірна моментна функція другого порядку:

$$D_x(t) = \int_{-\infty}^{\infty} [x - m_x(t)]^2 f_1(x; t) dx. \quad (2.2.14)$$

Визначення 2. Середньоквадратичним відхиленням (СКВ) $\sigma_x(t)$ випадкової функції $X(t)$ називається корінь із дисперсії: $\sigma_x(t) = \sqrt{D_x(t)}$.

Зауваження. Розмірність математичного сподівання і СКВ така ж, як і випадкової функції. Розмірність дисперсії – квадрат розмірності випадкової величини.

Властивості одновимірних моментних функцій перших двох порядків наведено в табл. 2.2.1.

Приклад. Напругу електричної мережі можна розглядати як випадкову функцію $X(t)$. У цьому випадку $X(t)$ має розмірність напруги, що вимірюється у вольтах. Математичне сподівання $m_x(t)$ і СКВ $\sigma_x(t)$ мають ту ж розмірність і вимірюються теж у вольтах. Дисперсія має розмірність напруги у квадраті і пропорційна потужності флюктуаційної складової струму.

Таблиця 2.2.1

Властивості одновимірних моментних функцій

Назва	Розмірність	Особливості
Математичне сподівання	як у ВФ	$M[X(t) \pm \varphi(t)] = m_x(t) \pm \varphi(t)$ $M[A(t)X(t)] = A(t) m_x(t)$
Дисперсія	квадрат розмірності ВФ	$D_x(t) = m_2(t) - m_x^2(t)$ $D[X(t) \pm \varphi(t)] = D_x(t)$ $D[A(t)X(t)] = A^2(t)D_x(t)$
СКВ	як у ВФ	$\sigma_x(t) = \sqrt{D_x(t)}$ $\sigma[X(t) \pm \varphi(t)] = \sigma_x(t)$ $\sigma[A(t)X(t)] = A(t)\sigma_x(t)$

Примітка. Функції $\varphi(t)$ і $A(t)$ – детерміновані; D – оператор дисперсії; σ – оператор середньоквадратичного відхилення.

Визначення 3. Кореляційною функцією (КФ) $R_x(t, t')$ випадкової функції $X(t)$ (у вузькому сенсі, див. примітку до зауваження 1 визначення 5 підрозділу 2.2.1) називається двовимірна центральна моментна функція другого порядку:

$$\begin{aligned} R_x(t, t') &= M[(X(t) - m_x(t))(X(t') - m_x(t'))] = \\ &= \int \int_{-\infty-\infty}^{\infty \infty} [x - m_x(t)][x' - m_x(t')] f_2(x, x'; t, t') dx dx'. \end{aligned} \quad (2.2.15)$$

Визначення 4. Коваріаційною функцією $K_x(t, t')$ випадкової функції $X(t)$ називається двовимірна початкова функція другого порядку:

$$K_x(t, t') = M[X(t)X(t')] = \int \int_{-\infty-\infty}^{\infty \infty} xx' f_2(x, x'; t, t') dx dx'. \quad (2.2.16)$$

Зауваження. Кореляційна і коваріаційна функції пов'язані між собою:

$$R_x(t, t') = K_x(t, t') - m_x(t)m_x(t'). \quad (2.2.17)$$

Визначення 5. Нормованою кореляційною функцією (НКФ) називається функція

$$r_x(t, t') = \frac{R_x(t, t')}{\sigma_x(t)\sigma_x(t')}. \quad (2.2.18)$$

Зауваження 1. Поняття кореляційної функції, коваріаційної функції і нормованої кореляційної функції випадкової функції аналогічні поняттям кореляційного моменту, коваріаційного моменту і нормованого кореляційного моменту (коефіцієнта кореляції) векторної випадкової величини.

Зауваження 2. Розмірність кореляційної і коваріаційної функцій – квадрат розмірності випадкової функції, тобто такий же, як і дисперсії. Нормована кореляційна функція – безрозмірна функція.

Основні властивості кореляційної функції і нормованої кореляційної функції наведено в табл. 2.2.2.

Таблиця 2.2.2

Властивості кореляційної функції і нормованої кореляційної функції

№ з/п	Властивість	Математичне трактування
1	У точці $t' = t$ 1) значення КФ дорівнює дисперсії; 2) значення НКФ дорівнює одиниці	$R_x(t, t) = D_x(t),$ $r_x(t, t) = 1$
2	Симетричність щодо аргументів	$R_x(t, t') = R_x(t', t),$ $r_x(t, t') = r_x(t', t)$
3	Обмеженість	$ R_x(t, t') \leq \sigma_x(t)\sigma_x(t'),$ $r_x(t, t') \leq 1$
4	При додатку до ВФ $X(t)$ детермінованої функції $\varphi(t)$ КФ і НКФ не змінюються	$Y(t) = X(t) + \varphi(t),$ $R_y(t, t') = R_x(t, t'),$ $r_y(t, t') = r_x(t, t')$
5	При множенні ВФ $X(t)$ на детерміновану функцію $A(t)$ 1) КФ множиться на значення детермінованої функції при відповідних аргументах; 2) НКФ не змінюється	$Y(t) = A(t)X(t),$ $R_y(t, t') = A(t)A(t')R_x(t, t'),$ $r_y(t, t') = r_x(t, t')$

Властивості 1 – 3 достатньо очевидні. Властивість 1 випливає з визначення поняття дисперсії випадкової функції, властивість 2 – із властивості 2 моментних функцій, а властивість 3 – із *нерівності Коші – Шварца* (іноді її називають *нерівністю Коші – Буняковського*):

$$\left| \int_x f_1(x)f_2(x)dx \right| \leq \sqrt{\int_x |f_1(x)|^2 dx \cdot \int_x |f_2(x)|^2 dx}, \quad (2.2.20)$$

де $f_1(x), f_2(x)$ – довільні функції, що інтегруються в області X .

Доведення властивостей 4 і 5 будується на основі безпосереднього розрахунку кореляційної функції і нормованої кореляційної функції для випадкових функцій $Y(t) = X(t) + \varphi(t)$ і $Y(t) = A(t)X(t)$.

Визначення 6. Перерізи t, t' випадкової функції називаються некорельованими, якщо для цих перерізів кореляційна функція $R_x(t, t')$ дорівнює нулю.

Зауваження 1. Некорельованість і незалежність – різні поняття. Можливі різноманітні їхні поєднання (рис. 2.2.1). Якщо перерізи некорельовані, то це не означає, що вони обов'язково незалежні. Проте якщо перерізи незалежні, то вони обов'язково некорельовані.

Зауваження 2. Кореляційна функція є мірою лінійного (і тільки лінійного) зв'язку між перерізами випадкової функції.

Зауваження 3. Якщо перерізи лінійно залежні: $X(t') = aX(t) + b$, де a і b – деякі константи, то кореляційна функція $R_x(t, t')$ приймає максимально можливе значення, рівне добутку СКВ $\sigma_x(t)\sigma_x(t')$. Нормована кореляційна функція при цьому дорівнює одиниці.

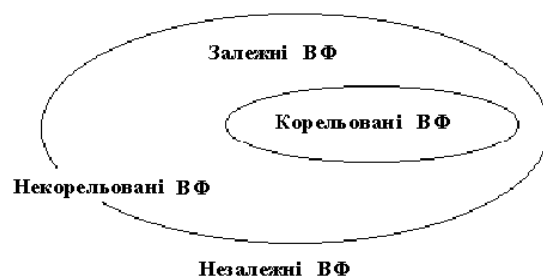


Рис. 2.2.1. Співвідношення між залежними, незалежними, корельованими і некорельованими випадковими функціями

Визначення 7. Перерізи t, t' випадкової функції називаються ортогональними, якщо для цих перерізів коваріаційна функція $K_x(t, t')$ дорівнює нулю.

Зауваження 1. З виразу (2.2.17) і зауваження 1 до визначення 6 випливає, що перерізи випадкової функції можуть бути ортогональними і при цьому як некорельованими, так і корельованими, як залежними, так і незалежними.

Зауваження 2. Якщо хоча б одне з математичних сподівань $m_x(t), m_x(t')$ дорівнює нулю, то з ортогональності перерізів випадкової функції випливає їхня некорельованість, а з некорельованості – їхня ортогональність.

2.2.3. Умовні моментні і кумулянтні функції випадкових функцій

Для опису умовних випадкових функцій використовують умовні моментні функції.

Визначення 1. Умовною початковою моментною функцією вимірності $(N - M)$ порядку $\nu = \nu_{M+1} + \dots + \nu_N$ випадкової функції $X(t)$ називається математичне сподівання

$$\begin{aligned} m_{\nu_{M+1} \dots \nu_N}(t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) = \\ = M[X^{\nu_{M+1}}(t_{M+1}) \dots X^{\nu_N}(t_N) / x_1, \dots, x_M; t_1, \dots, t_M], \end{aligned} \quad (2.2.21)$$

яке обчислюється за умови, що в перерізах t_1, \dots, t_M випадкова функція $X(t)$ прийняла значення x_m , ($m = \overline{1, M}$).

Визначення 2. Умовним математичним сподіванням випадкової функції (апостеріорним математичним сподіванням) $X(t)$ називається умовна одновимірна моментна функція першого порядку

$$m_x(t/x_1, \dots, x_M; t_1, \dots, t_M) = M_x[X(t)/x_1, \dots, x_M; t_1, \dots, t_M] = \int_{-\infty}^{\infty} x f_1(x; t/x_1, \dots, x_M; t_1, \dots, t_M) dx. \quad (2.2.22)$$

Зауваження. Безумовне (апріорне) математичне сподівання пов'язане з умовним співвідношенням

$$m_x(t) = M_{x_1, \dots, x_M}[m_x(t/X_1, \dots, X_M; t_1, \dots, t_M)],$$

де M_{x_1, \dots, x_M} – оператор математичного сподівання, що діє на X_1, \dots, X_M .

Визначення 3. Дисперсією умовних середніх називається функція, обчислена для умовного математичного сподівання $m_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$:

$$D_m(t) = D_{x_1, \dots, x_M}[m_x(t/X_1, \dots, X_M; t_1, \dots, t_M)], \quad (2.2.23)$$

де D_{x_1, \dots, x_M} – оператор дисперсії, що діє на X_1, \dots, X_M .

Визначення 4. Умовною дисперсією $D_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ називається функція, що визначається для випадкової функції $X(t)$ таким чином:

$$D_x(t/x_1, \dots, x_M, t_1, \dots, t_M) = D_x[X(t)/x_1, \dots, x_M; t_1, \dots, t_M]. \quad (2.2.24)$$

Визначення 5. Середньою умовною дисперсією $D_D(t)$ називається функція, що визначається для умовної дисперсії $D_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ таким чином:

$$D_D(t) = M_{x_1, \dots, x_M}[D_x(t/X_1, \dots, X_M; t_1, \dots, t_M)], \quad (2.2.25)$$

де математичне сподівання обчислюється по X_1, \dots, X_M .

Уявлення про різноманітні моментні функції дає рис. 2.2.2.

Зауваження 1. Умовне математичне сподівання $m_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ відрізняється від безумовного математичного сподівання $m_x(t)$. Виключенням є випадок, коли перерізи випадкових функцій $X(t)$ незалежні.

Зауваження 2. Для значень t_1, \dots, t_M умовне математичне сподівання $m_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ збігається з реалізацією $x(t_m)$, ($m = \overline{1, M}$), що спостерігається, а умовна дисперсія $D_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ дорівнює нулю.

Зауваження 3. Для значень $t > t_M$ умовне математичне сподівання $m_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ характеризує поведінку умовної випадкової функції $X(t/x(t_m), m = \overline{1, M})$ в середньому, а умовна дисперсія $D_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ – розсіювання значень цієї випадкової функції відносно умовного математичного сподівання.

Зауваження 4. Дисперсія умовних середніх $D_m(t)$ характеризує розсіювання умовного математичного сподівання $m_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ при зміні умов, а середня умовна дисперсія $D_D(t)$ – середню величину умовної дисперсії $D_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ при зміні умов.

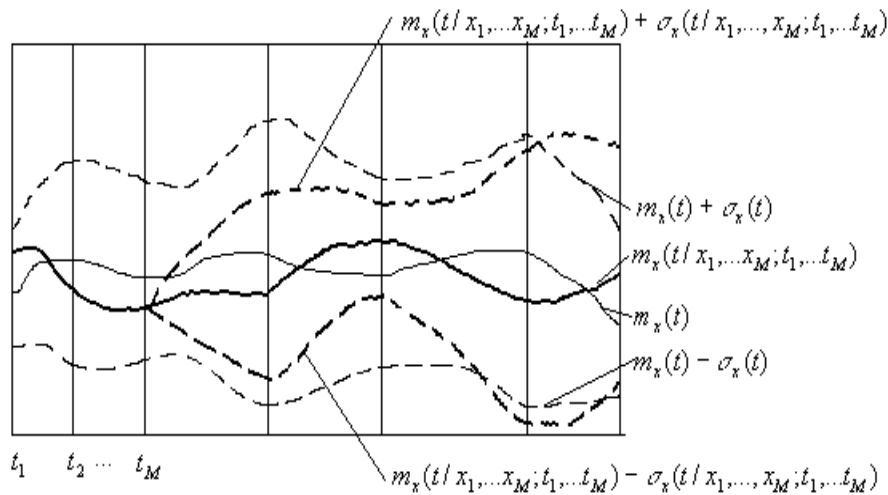


Рис. 2.2.2. Моментні функції випадкової функції: *априорне* математичне сподівання $m_x(t)$ безумовної випадкової функції $X(t)$; границі $m_x(t) \pm \sigma_x(t)$ односигмового відхилення від математичного сподівання $m_x(t)$ безумовної випадкової функції $X(t)$; *апостеріорне* математичне сподівання $m_x(t/x_1, \dots, x_M, t_1, \dots, t_M)$ випадкової функції $X(t)$; границі $m_x(t/x_1, \dots, x_M; t_1, \dots, t_M) \pm \sigma_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ односигмового відхилення від апостеріорного математичного сподівання $m_x(t/x_1, \dots, x_M; t_1, \dots, t_M)$ випадкової функції $X(t)$

Теорема. Безумовна дисперсія $D_x(t)$ випадкової величини $X(t)$ дорівнює сумі дисперсії умовних середніх $D_m(t)$ і середньої умовної дисперсії $D_D(t)$:

$$D_x(t) = D_m(t) + D_D(t). \quad (2.2.26)$$

Для характеристики *лінійного* зв'язку між перерізами умовної випадкової функції $X(t/x(t_m), m = \overline{1, M})$ використовується умовна кореляційна функція.

Визначення 6. Умовною кореляційною функцією $R_x(t, t'/x_1, \dots, x_M; t_1, \dots, t_M)$ називається двовимірна умовна центральна моментна функція другого порядку, що визначається так:

$$\begin{aligned} R_x(t, t'/x_1, \dots, x_M; t_1, \dots, t_M) &= \\ &= M_x^0 [X(t/x_1, \dots, x_M; t_1, \dots, t_M) X(t'/x_1, \dots, x_M; t_1, \dots, t_M)], \end{aligned} \quad (2.2.27)$$

де $X^0(t/x_1, \dots, x_M; t_1, \dots, t_M)$ – центрована умовна випадкова функція:

$$\begin{aligned} X^0(t/x_1, \dots, x_M; t_1, \dots, t_M) &= \\ &= X(t/x_1, \dots, x_M; t_1, \dots, t_M) - m_x(t/x_1, \dots, x_M; t_1, \dots, t_M). \end{aligned} \quad (2.2.28)$$

Основні умовні моментні функції перелічені в табл. 2.2.3.

Зауваження. Подібно кумулянтній функції $\kappa_{v_1 \dots v_N}(t_1, \dots, t_N)$ можна визначити умовну кумулянтну функцію $\kappa_{v_{M+1} \dots v_N}(t_{M+1}, \dots, t_N/x_1, \dots, x_M; t_1, \dots, t_M)$. У даному випадку у формулі (2.2.13) замість характеристичної функції $Q_N(j\omega_1, \dots, j\omega_N; t_1, \dots, t_N)$ треба використовувати умовну характеристичну функцію $Q_{N-M}(j\omega_{M+1}, \dots, j\omega_N; t_{M+1}, \dots, t_N/x_1, \dots, x_M; t_1, \dots, t_M)$. Умовні кумулянтні

функції однозначно виражаються через умовні моментні функції і, навпаки, умовні моментні функції – через умовні кумулянтні функції. Залежність між цими функціями така ж, як і між безумовними кумулянтними і моментними функціями.

Таблиця 2.2.3

Основні умовні моментні функції випадкової функції $X(t)$

Назва	Математичне визначення
Умовне математичне сподівання	$m_x(t/x_1, \dots, x_M; t_1, \dots, t_M) = M_x[X(t)/x_1, \dots, x_M; t_1, \dots, t_M]$
Дисперсія умовних середніх	$D_m(t) = D_{x_1, \dots, x_M}[m_x(t/X_1, \dots, X_M; t_1, \dots, t_M)]$
Умовна дисперсія	$D_x(t/x_1, \dots, x_M; t_1, \dots, t_M) = D_x[X(t)/x_1, \dots, x_M; t_1, \dots, t_M]$
Середня умовна дисперсія	$D_D(t) = M_{x_1, \dots, x_M}[D_x(t/X_1, \dots, X_M; t_1, \dots, t_M)]$
Умовна кореляційна функція	$R_x(t, t'/x_1, \dots, x_M; t_1, \dots, t_M) = M_x^0[X(t/x_1, \dots, x_M; t_1, \dots, t_M)X(t'/x_1, \dots, x_M; t_1, \dots, t_M)]$

2.3. Векторні випадкові функції

Визначення 1. Векторною випадковою функцією називається векторна функція $\vec{X}(t)$, компоненти якої являють собою звичайні (скалярні) випадкові функції: $\vec{X}(t) = (X_1(t), \dots, X_H(t))$.

Приклад. Векторною випадковою функцією описуються випадкові відхилення показань відразу декількох авіаційних приладів (наприклад, висотоміру, дальноміру, датчиків тиску, витрати палива та ін.).

Зауваження. Найбільш повний опис векторної випадкової функції дають ймовірнісні характеристики спільного закону розподілу (багатовимірні щільність ймовірностей, функція розподілу, характеристична функція).

Визначення 2. Випадкові функції $X_1(t)$ і $X_2(t)$ називаються незалежними, якщо сукупність значень функції $X_1(t)$ не залежить від сукупності значень функції $X_2(t)$.

Зауваження. Необхідною і достатньою умовами незалежності випадкових функцій $X_1(t)$ і $X_2(t)$ є можливість представлення спільної щільності ймовірностей $f_{2N}(\vec{x}_1, \vec{x}_2; \vec{t}_1, \vec{t}_2)$ у вигляді добутку щільностей ймовірностей $f_N(\vec{x}_1; \vec{t}_1)$ і $f_N(\vec{x}_2; \vec{t}_2)$.

Визначення 3. Компоненти векторної випадкової функції $\vec{X}(t) = (X_1(t), \dots, X_H(t))$ називаються незалежними в сукупності, якщо щільність ймовірностей спільного закону розподілу $f_{NH}(\vec{x}_1, \dots, \vec{x}_n; \vec{t}_1, \dots, \vec{t}_n)$ може бути подана у вигляді

$$f_{NH}(\vec{x}_1, \dots, \vec{x}_n; \vec{t}_1, \dots, \vec{t}_n) = \prod_{h=1}^H f_N(\vec{x}_h; \vec{t}_h), \quad (2.3.1)$$

де N – вимірність компонент векторної випадкової функції (для спрощення вони вважаються однакової вимірності), H – кількість компонент.

Зауваження 1. Якщо компоненти векторної випадкової функції незалежні, то задача опису векторної випадкової функції зводиться до незалежного опису сукупності скалярних випадкових функцій-компонент. Якщо ж компоненти залежні, то задача ускладнюється. У цьому випадку необхідне урахування зв'язків між компонентами.

Зауваження 2. Якщо компоненти векторної випадкової функції $\vec{X}(t) = (X_1(t), \dots, X_H(t))$ попарно незалежні, то зовсім не обов'язково, що компоненти незалежні в сукупності.

Визначення 4. Математичним сподіванням $\vec{m}_x(t)$ векторної випадкової функції $\vec{X}(t) = (X_1(t), \dots, X_H(t))$ називається *невипадкова* векторна функція, компонентами якої є математичні сподівання відповідних компонент векторної випадкової функції:

$$\vec{m}_x(t) = M[\vec{X}(t)]. \quad (2.3.2)$$

Визначення 5. Дисперсією $\vec{D}_x(t)$ векторної випадкової функції називається *невипадкова* векторна функція, компонентами якої є дисперсії відповідних компонент векторної випадкової функції $\vec{X}(t)$:

$$\vec{D}_x(t) = D[\vec{X}(t)]. \quad (2.3.3)$$

Визначення 6. Кореляційною функцією $R_x(t, t')$ векторної випадкової функції $\vec{X}(t)$ називається *невипадкова* квадратна матриця розміром $H \times H$ з елементами

$$R_{hl}(t, t') = M[X_h^0(t) X_l^0(t')] \quad (2.3.4)$$

$$(h, l = \overline{1, H}).$$

Зауваження. Діагональні елементи матриці $R_{hl}(t, t')$ являють собою кореляційні функції компонент. Позадіагональні елементи описують кореляційні зв'язки між компонентами. Вони називаються *взаємними кореляційними функціями*.

Визначення 7. Нормованою кореляційною функцією називається *невипадкова* функція

$$r_{hl}(t, t') = \frac{R_{hl}(t, t')}{\sqrt{R_{hh}(t, t') R_{ll}(t, t')}} \quad (h, l = \overline{1, H}). \quad (2.3.5)$$

Основні властивості кореляційної функції і нормованої кореляційної функції векторної випадкової функції наведено в табл. 2.3.1.

Зауваження. Нелінійні стохастичні (випадкові) зв'язки між перерізами випадкової функції характеризуються не кореляційними моментами, а моментними функціями, порядок котрих більше двох.

Таблиця 2.3.1

Властивості кореляційної функції (КФ) і нормованої кореляційної функції (НКФ) векторної випадкової функції

Властивість	Математичне трактування
Значення КФ дорівнює дисперсії, коли збігаються індекси КФ і аргументи КФ	$R_{hh}(t, t) = D_h(t), h = \overline{1, H}$
Симетричність щодо системи індекс-аргумент	$R_{hl}(t, t') = R_{lh}(t', t)$
1) Модуль елементів КФ не перевищує добутку відповідних СКВ 2) Модуль елементів НКФ не перевищує одиниці	$ R_{hl}(t, t') \leq \sigma_h(t)\sigma_l(t')$ $ r_{hl}(t, t') \leq 1$

Примітка. Якщо індекси не збігаються, то значення, рівне одиниці, НКФ приймає необов'язково при $t = t'$.

2.4. Перетворення випадкових функцій

2.4.1. Класифікація операторів

Положення. Функції можуть перетворюватися різноманітними операторами. Будь-який оператор D умовно може бути представлений у вигляді деякого функціонального блока (системи) (рис. 2.4.1), що перетворює функцію $x(t)$, яка називається *впливом*, у функцію $y(t)$, що називається *реакцією* або *відгуком* на вплив.

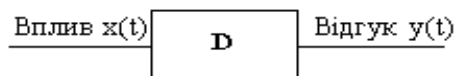


Рис. 2.4.1. Умовне зображення оператора

Оператори можуть бути різного типу. Класифікація операторів за критеріями, що використовуються найчастіше, наведено в табл. 2.4.1.

Таблиця 2.4.1

Класифікація операторів перетворення випадкових функцій

Критерій класифікації	Види операторів		
	Характер оператора	Детермінований	Випадковий
Тип оператора	Лінійний однорідний	Лінійний неоднорідний	Нелінійний
Інерційність	Безінерційний		Інерційний

Визначення 1. *Детермінованим оператором* називається оператор, усі параметри якого детерміновані, а *випадковим* – оператор, параметри якого випадкові.

Визначення 2. Оператор L називається *лінійним однорідним*, якщо він має властивість *суперпозиції*:

- 1) $L[x_1(t) + x_2(t)] = L[x_1(t)] + L[x_2(t)]$;
- 2) $L[\alpha x(t)] = \alpha L[x(t)]$.

Приклади 1. Лінійними операторами є оператори диференціювання й інтегрування функцій.

Визначення 3. Оператор D називається *лінійним неоднорідним*, якщо його можна представити у вигляді суми лінійного оператора L і деякої детермінованої функції $\varphi(t)$:

$$D[x(t)] = L[x(t)] + \varphi(t).$$

Визначення 4. Оператори, що не відносяться до лінійних, називаються *нелінійними*.

Визначення 5. *Безінерційним* (функціональним) називається оператор, для якого значення відгуку $y(t)$ при значенні аргументу t визначається значенням функції $x(t)$ тільки при тому ж значенні аргументу t .

Визначення 6. *Інерційним* оператором називається оператор, для якого значення відгуку $y(t)$ при значенні t визначається значеннями функції $x(t)$ як при тому ж значенні t , так й при інших її значеннях.

Приклади 2. Наведемо декілька прикладів пристроїв, робота яких описується різноманітними операторами. Лінійний підсилювач сигналу з рівномірною амплітудно-частотною і лінійною фазовою характеристиками в широкому діапазоні частот приблизно описується лінійним однорідним безінерційним оператором. Підсилювач сигналу з нерівномірною амплітудно-частотною характеристикою (фільтр) описується лінійним однорідним інерційним оператором, а детектор сигналу описується нелінійним безінерційним оператором.

Зауваження 1. Якщо вплив являє собою випадкову функцію, то і відгук є випадковою функцією.

Зауваження 2. Одними з основних задач теорії операторів є знаходження по ймовірнісних характеристиках і моментних функціях впливу $X(t)$ відповідних ймовірнісних характеристик і моментних функцій відгуку $Y(t)$.

2.4.2. Перетворення випадкових функцій безінерційними операторами

Положення. При безінерційному перетворенні однієї випадкової функції $X(t)$ в іншу випадкову функцію $Y(t)$ кожний переріз впливу $X(t)$ породжує відповідний переріз відгуку $Y(t)$. При цьому відбувається перетворення однієї випадкової величини в іншу. Ця особливість безінерційного перетворення дозволяє узагальнити усі відомі результати, що стосуються перетворення випадкових величин, на випадок безінерційних перетворень випадкових функцій.

Єдина особливість, яку треба враховувати при узагальненні результатів, це залежність від аргументу характеристик випадкових функцій (як ймовірнісних, так і неймовірнісних характеристик).

Зупинимося спочатку на щільності ймовірностей.

Теорема 1. Нехай безінерційний оператор описується деякою нелінійною в загальному випадку функцією $y = \varphi(x)$, що диференціюється і має *однозначну* обернену функцію $x = \eta(y)$. У цьому випадку одновимірна щільність ймовірностей $f_1^y(y;t)$ відгуку $Y(t)$ на вплив $X(t)$ описується співвідношенням

$$f_1^y(y;t) = f_1^x(\eta(y);t) \left| \frac{d\eta(y)}{dy} \right|, \quad (2.4.1)$$

де $f_1^x(\eta(y);t)$ – одновимірна щільність ймовірностей впливу $X(t)$.

Теорема 2. Нехай безінерційний оператор описується деякою диференційованою в загальному випадку нелінійною функцією $y = \varphi(x)$, обернена функція якої має Q гілок, що описуються функціями $x = \eta_q(y)$, $q = \overline{1, Q}$.

Тоді одновимірна щільність ймовірностей $f_1^y(y, t)$ відгуку $Y(t)$ на вплив $X(t)$ описується співвідношенням

$$f_1^y(y; t) = \sum_{q=1}^Q f_1^x(\eta_q(y); t) \left| \frac{d\eta_q(y)}{dy} \right|. \quad (2.4.2)$$

Співвідношення (2.4.1) і (2.4.2) доводяться так само, як і співвідношення (1.3.1) і (1.3.2).

Розглянемо векторний випадок, коли вплив описується векторною випадковою функцією $\vec{X}(t)$, а відгук – векторною випадковою функцією $\vec{Y}(t)$. Тоді справедлива така теорема.

Теорема 3. Нехай безінерційний оператор описується деякою диференційованою в загальному випадку нелінійною функцією $\vec{y} = \vec{\varphi}(\vec{x})$, що має однозначні обернені функції:

$$\begin{aligned} x_1 &= \eta_1(y_1, \dots, y_H), \\ &\dots\dots\dots \\ x_H &= \eta_H(y_1, \dots, y_H), \end{aligned}$$

де H – вимірність векторів $\vec{x} = (x_1, \dots, x_H)$ і $\vec{y} = (y_1, \dots, y_H)$. Тоді H -вимірна щільність ймовірностей $f_H^y(\vec{y}; t)$ відгуку $\vec{Y}(t)$ на вплив $\vec{X}(t)$ описується співвідношенням

$$f_H^y(\vec{y}; t) = f_H^x(\eta_1(\vec{y}), \dots, \eta_H(\vec{y}); t) |J_H(\vec{y})|, \quad (2.4.3)$$

де $f_H^x(\vec{x}; t)$ – H -вимірна щільність ймовірностей випадкової функції $\vec{X}(t)$, $J_H(\vec{y})$ – яacobіан перетворення змінних:

$$J_H(\vec{y}) = \begin{vmatrix} \frac{\partial \eta_1}{\partial y_1} & \dots & \frac{\partial \eta_1}{\partial y_H} \\ \dots & \dots & \dots \\ \frac{\partial \eta_H}{\partial y_1} & \dots & \frac{\partial \eta_H}{\partial y_H} \end{vmatrix}. \quad (2.4.4)$$

Зауваження. Якщо обернені функції η_1, \dots, η_H неоднозначні, то в правій частині співвідношення (2.4.3) треба брати суму по всіх підобластях.

Теорема 4. Нехай виконуються умови теореми 1. Тоді N -вимірна щільність ймовірностей $f_N^y(y_1, \dots, y_N; t_1, \dots, t_N)$ скалярного відгуку $Y(t)$ на скалярний вплив $X(t)$ описується співвідношенням

$$f_N^y(y_1, \dots, y_N; t_1, \dots, t_N) = f_N^x(\eta(y_1), \dots, \eta(y_N); t_1, \dots, t_N) \left| \prod_{n=1}^N \frac{\partial \eta(y_n)}{\partial y_n} \right|, \quad (2.4.5)$$

де $f_N^x(x_1, \dots, x_N; t_1, \dots, t_N)$ – N -вимірна щільність ймовірностей впливу $X(t)$.

Доведення. Уявимо випадкові скалярні функції $X(t)$ і $Y(t)$ їхніми перерізами в N точках t_1, \dots, t_N . Сукупність цих перерізів можна розглядати як векторні випадкові функції $\vec{X}(t)$ і $\vec{Y}(t)$, кожну компоненту яких визначено

лише в одній точці (для компонент $X_n(t)$ і $Y_n(t)$ в точках $t_n, n = \overline{1, N}$). Звернемо увагу на те, що N -вимірні щільності ймовірностей векторних випадкових функцій $\vec{X}(t)$, $\vec{Y}(t)$ і відповідні N -вимірні щільності ймовірностей скалярних випадкових функцій $X(t)$, $Y(t)$ збігаються. На цій підставі для розрахунку N -вимірної щільності ймовірностей $f_N^y(y_1, \dots, y_N; t_1, \dots, t_N)$ відгуку $Y(t)$ на вплив $X(t)$ можна скористатися формулою (2.4.3), що зв'язує N -вимірну щільність ймовірностей $\vec{Y}(t)$ з N -вимірною щільністю ймовірності впливу $\vec{X}(t)$. Врахуємо, що в даному випадку всі обернені функції однакові: $\eta_1 = \dots = \eta_N = \eta$, а в якобіані перетворення (2.4.3) усі позадіагональні елементи дорівнюють нулю. При цьому формула (2.4.3) трансформується у формулу (2.4.5).

Теорема 5. Нехай безінерційний оператор описується парою функцій $y_1 = \varphi(x_1, x_2)$, $y_2 = x_2$, що мають однозначні обернені функції:

$$\begin{aligned} x_1 &= \eta_1(y_1, y_2), \\ x_2 &= y_2. \end{aligned}$$

Тоді одновимірна щільність ймовірностей $f_1^y(y_1; t)$ відгуку $Y_1(t)$ на вплив $\vec{X}(t) = (X_1(t), X_2(t))$ описується таким співвідношенням:

$$f_1^y(y_1; t) = \int_{-\infty}^{\infty} f_2^x(\eta(y_1, y_2), y_2; t) \left| \frac{\partial \eta_1(y_1, y_2)}{\partial y_1} \right| dy_2, \quad (2.4.6)$$

де $f_2^x(x_1, x_2; t)$ – двовимірна щільність ймовірностей впливу $\vec{X}(t)$.

Доведення. Скористаємося теоремою 3. У даному випадку якобіан перетворення має вигляд

$$J_n(\vec{y}) = \begin{vmatrix} \frac{\partial \eta_1}{\partial y_1} & \frac{\partial \eta_1}{\partial y_2} \\ 0 & 1 \end{vmatrix} = \frac{\partial \eta_1}{\partial y_1}.$$

Тоді на підставі формули (2.4.3) маємо

$$f_2^y(y_1, y_2; t) = f_2^x(\eta(y_1, y_2), y_2; t) \left| \frac{\partial \eta_1}{\partial y_1} \right|.$$

Переходячи від двовимірної щільності ймовірностей $f_2^y(y_1, y_2; t)$ до одновимірної $f_1^y(y_1; t)$, одержуємо співвідношення (2.4.6).

Зауваження. Формула (2.4.6) дозволяє знаходити одновимірні щільності ймовірностей випадкових функцій $Y(t)$, які одержані в результаті додавання, віднімання, множення і ділення двох випадкових функцій $X_1(t)$ і $X_2(t)$ (табл. 2.4.2).

Зупинимося коротко на моментних функціях. У ряді випадків вони можуть бути досить просто розраховані, минаючи етап обчислення щільності ймовірностей відгуку $Y(t)$.

Приклад 1. Нехай безінерційний оператор описується функцією $y = \varphi(x)$. Тоді одновимірну початкову моментну функцію $m_\nu^y(t)$ ν -го порядку можна розрахувати за формулою

$$m_\nu^y(t) = M[Y^\nu(t)] = \int_{-\infty}^{\infty} \varphi^\nu(x) f_1^x(x; t) dx. \quad (2.4.7)$$

Таблиця 2.4.2

Зв'язок одновимірної щільності ймовірностей $f_1^y(y;t)$ відгуку $Y(t)$ з двовимірною щільністю ймовірності $f_2^x(x_1, x_2; t)$ впливу $\vec{X}(t) = (X_1(t), X_2(t))$ для різних операторів

Оператор	Щільність ймовірностей $f_1^y(y;t)$ відгуку $Y(t)$
$y = x_1 + x_2$	$f_1^y(y;t) = \int_{-\infty}^{\infty} f_2^x(y - x_2, x_2; t) dx_2$
$y = x_1 - x_2$	$f_1^y(y;t) = \int_{-\infty}^{\infty} f_2^x(y + x_2, x_2; t) dx_2$
$y = x_1 \cdot x_2$	$f_1^y(y;t) = \int_{-\infty}^{\infty} f_2^x\left(\frac{y}{x_2}, x_2; t\right) \frac{dx_2}{ x_2 }$
$y = \frac{x_1}{x_2}$	$f_1^y(y;t) = \int_{-\infty}^{\infty} f_2^x(yx_2, x_2; t) x_2 dx_2$

Зауваження. З співвідношення (2.4.7) випливає, що математичне сподівання відгуку $Y(t)$ описується співвідношенням

$$m_y(t) = \int_{-\infty}^{\infty} \varphi(x) f_1^x(x;t) dx.$$

Приклад 2. Нехай безінерційний оператор описується функцією $y = \varphi(x)$. Знайдемо одновимірну центральну моментну функцію $\mu_\nu^y(t)$ ν -го порядку. Цю моментну функцію можна представити у вигляді

$$\mu_\nu^y(t) = M\left[(Y(t) - m_y(t))^\nu\right] = \int_{-\infty}^{\infty} [\varphi(x) - m_y(t)]^\nu f_1^x(x;t) dx.$$

Приклад 3. Нехай є лінійний безінерційний оператор виду $\sum_{q=1}^Q \alpha_q x_q$, (де α_q – константа, $q = \overline{1, Q}$). Для цього оператора також можна безпосередньо розрахувати моментні функції. Результати розрахунків наведено в табл. 2.4.3.

Зауваження 1. При додаванні випадкових функцій $X_1(t)$, $X_2(t)$ математичне сподівання суми $Y(t)$ дорівнює сумі математичних сподівань доданків.

Зауваження 2. При відсутності кореляції між функціями $X_q(t)$, $q = \overline{1, Q}$ кореляційна функція і дисперсія випадкової функції $Y(t)$ істотно спрощуються:

$$R_y(t, t') = \sum_{q=1}^Q \alpha_q^2 R_{x_q}(t, t'),$$

$$D_y(t) = \sum_{q=1}^Q \alpha_q^2 D_{x_q}(t).$$

Зауваження 3. При додаванні некорельованих випадкових функцій дисперсія суми дорівнює сумі дисперсій, а при відніманні двох некорельованих

випадкових функцій $X_1(t)$ і $X_2(t)$ дисперсія різниці дорівнює сумі дисперсії зменшеного і від'ємника.

Таблиця 2.4.3

Залежність математичного сподівання $m_y(t)$, кореляційної функції $R_y(t, t')$ і дисперсії $D_y(t, t')$ відгуку $Y(t)$ від математичних сподівань $m_{x_q}(t)$, кореляційних функцій $R_{x_q x_h}(t)$ і дисперсій $D_{x_q}(t)$ впливів $X_q(t)$

для безінерційного оператора вигляду $y = \sum_{q=1}^Q \alpha_q x_q$

Назва моментної функції	Співвідношення
Математичне сподівання	$m_y(t) = \sum_{q=1}^Q \alpha_q m_{x_q}(t)$
Кореляційна функція	$R_y(t, t') = \sum_{q=1}^Q \alpha_q^2 R_{x_q x_q}(t, t') + \sum_{q \neq h} \alpha_q \alpha_h R_{x_q x_h}(t, t')$
Дисперсія	$D_y(t) = \sum_{q=1}^Q \alpha_q^2 D_{x_q}(t) + \sum_{q \neq h} \alpha_q \alpha_h R_{x_q x_h}(t, t)$

2.4.3. Перетворення випадкових функцій лінійними інерційними операторами

Інерційність оператора суттєво ускладнює задачу знаходження ймовірнісних характеристик відгуку $Y(t)$ по ймовірнісних характеристиках впливу $X(t)$.

Якщо при безінерційному перетворенні для знаходження одновимірної щільності ймовірностей відгуку $Y(t)$ потрібно знання одновимірної щільності ймовірностей впливу $X(t)$, то при інерційному перетворенні (навіть лінійному) цих знань вже недостатньо.

При відсутності будь-яких припущень, які б спрощували задачу, потрібні дані про нескінченновимірну щільність ймовірностей впливу $X(t)$, що призводить до неподоланих труднощів точного розв'язування задачі в загальному вигляді. Тому доводиться вдаватися до методів неповного опису випадкових функцій. До таких методів належить метод опису за допомогою моментних функцій.

Розглянемо цей метод відповідно до лінійного інерційного перетворення випадкової функції $X(t)$ у випадкову функцію $Y(\tau)$. Аналітично таке перетворення може бути описано таким чином:

$$Y(\tau) = L_t[X(t)], \quad (2.4.8)$$

де L_t – лінійний оператор, що впливає на аргумент t .

Визначення 1. Оператори D_1 і D_2 називаються *переміщувальними (комутативними)*, якщо

$$D_1[D_2[x]] = D_2[D_1[x]].$$

Теорема 1. Лінійні оператори є переміщувальними.

Теорема 2. Оператор математичного сподівання M є лінійним.

На основі цих двох теорем неважко встановити залежність між моментними функціями відгуку $Y(\tau)$ і впливу $X(t)$ при лінійному перетворенні (табл. 2.4.4).

Таблиця 2.4.4

Залежність між моментними функціями відгуку і впливу при лінійному перетворенні

Назва моментної функції	Залежність між моментними функціями відгуку і впливу
Математичне сподівання	$m_y(\tau) = L_t[m_x(t)]$
Кореляційна функція	$R_y(\tau, \tau') = L_t L_{t'}[R_x(t, t')]$
Дисперсія	$D_y(\tau) = R_y(\tau, \tau') _{\tau=\tau'}$

Найпростішими лінійними операторами є оператори диференціювання та інтегрування:

$$y(t) = \frac{d}{dt} x(t),$$

$$y(\tau) = \int_{-\infty}^{\tau} x(t) dt.$$

Математичне визначення цих операцій для випадкової функції наведено у додатку 3. Для таких операторів в табл. 2.4.5 – 2.4.6 наведені залежності між моментними функціями відгуку і впливу, які часто використовуються на практиці.

Таблиця 2.4.5

Залежність між моментними функціями відгуку і впливу для операторів диференціювання

Назва моментної функції	Залежність між моментними функціями відгуку і впливу
Математичне сподівання	$m_y(t) = \frac{d}{dt} m_x(t)$
Кореляційна функція	$R_y(t, t') = \frac{\partial^2 R_x(t, t')}{\partial t \partial t'}$
Дисперсія	$D_y(t) = \frac{\partial^2 R_x(t, t')}{\partial t \partial t'} _{t=t'}$

Таблиця 2.4.6

Залежність між моментними функціями відгуку і впливу для операторів інтегрування

Назва моментної функції	Залежність між моментними функціями відгуку і впливу
Математичне сподівання	$m_y(\tau) = \int_{-\infty}^{\tau} m_x(t) dt$
Кореляційна функція	$R_y(\tau, \tau') = \int_{-\infty}^{\tau} \int_{-\infty}^{\tau'} R_x(t, t') dt dt'$
Дисперсія	$D_y(\tau) = \int_{-\infty}^{\tau} \int_{-\infty}^{\tau} R_x(t, t') dt dt'$

Відомості про книги з тематики розділу 2 наведено в табл. 2.4.7.

Таблиця 2.4.7

Література до розділу 2

Характер книг	Номери книг у списку літератури
Довідники	24, 40, 47, 49
Книги навчального плану для інженерів	3, 16, 25, 32, 40, 69, 73, 75, 84
Книги з математичним ухилом	5, 10, 15, 20, 22, 29, 61, 88, 89, 105
Книги прикладного характеру	38, 55, 56, 66, 74, 82, 84, 86, 92, 94

3. Кореляційна теорія випадкових функцій

Розділ теорії випадкових функцій, в якому випадкові функції описуються із залученням лише моментних функцій першого і другого порядку, називається *кореляційною теорією випадкових функцій*.

Опис випадкових функцій моментними функціями перших двох порядків зручний і тому часто використовується для розв'язання практичних задач. Однак слід мати на увазі, що в загальному випадку такий опис не є повним. Виняток складають гауссівські (нормальні) випадкові функції.

3.1. Гауссівські випадкові функції

Визначення 1. Випадкова функція $X(t)$ називається *гауссівською* (нормальною), якщо для будь-якої скінченної множини її перерізів $X(t_1), \dots, X(t_N)$ спільна щільність ймовірностей перерізів описується співвідношенням

$$f_N(\vec{x}; \vec{t}) = \frac{1}{\sqrt{(2\pi)^N |R_x|}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{m}_x)^T R_x^{-1} (\vec{x} - \vec{m}_x)\right), \quad (3.1.1)$$

де $\vec{x} = (x(t_1), \dots, x(t_N))^T$ – N -вимірний вектор-стовпець, n -а компонента якого являє собою значення випадкової величини $X(t_n)$; $\vec{m}_x = (m_x(t_1), \dots, m_x(t_N))^T$ – N -вимірний вектор-стовпець математичного сподівання, кожна n -а компонента якого являє собою математичне сподівання випадкової величини $X(t_n)$; $R_x = \{R_x(t_n, t_m)\}$ – кореляційна матриця розміром $N \times N$, елементи якої являють собою центральні кореляційні моменти випадкових величин $X(t_n)$, $X(t_m)$ ($n, m = \overline{1, N}$); R_x^{-1} – матриця, обернена кореляційній матриці R_x ; T – оператор транспонування.

Гауссівські випадкові функції мають цілий ряд важливих властивостей, що відрізняють їх від інших випадкових функцій.

Властивість 1. Гауссівська випадкова функція вичерпно визначається математичним сподіванням $m_x(t)$ і кореляційною функцією $R_x(t, t')$. Звідси випливає, що кореляційна теорія дає вичерпний її опис. Слід підкреслити, що усі моменти вищих порядків гауссівської випадкової функції виражаються через моменти перших двох її порядків.

Властивість 2. Для гауссівських випадкових функцій некорельованість перерізів тотожна їх незалежності, а незалежність – їх некорельованості.

Доведення. Якщо перерізи гауссівської функції некорельовані, то $R_x(t, t') = 0$ при $t \neq t'$ і тоді N -вимірна щільність ймовірностей $f_N(\vec{x}; \vec{t})$ факторизується (представляється у вигляді добутку одновимірних щільностей ймовірностей):

$$f_N(\vec{x}; \vec{t}) = f_1(x_1; t_1) \dots f_1(x_N; t_N). \quad (3.1.2)$$

Навпаки, якщо перерізи незалежні, то щільність ймовірностей $f_N(\vec{x}; \vec{t})$ представляється формулою (3.1.2). При цьому кореляційна функція $R_x(t, t') = 0$, якщо $t \neq t'$.

Властивість 3. Якщо випадкова функція $X(t)$ – гауссівська, то умовна випадкова функція $X(t / X(t_m) = x(t_m), (m = \overline{1, M}))$ також гауссівська.

Доведення. Для будь-якої випадкової функції при $N > M$ справедливе співвідношення

$$f_{N-M}(x_{M+1}, \dots, x_N; t_{M+1}, \dots, t_N / x_1, \dots, x_M; t_1, \dots, t_M) = \frac{f_N(x_1, \dots, x_N; t_1, \dots, t_N)}{f_M(x_1, \dots, x_M; t_1, \dots, t_M)}. \quad (3.1.3)$$

Якщо випадкова функція гауссівська, то обидві щільності ймовірностей в правій частині формули (3.1.3) мають вигляд виразу (3.1.1). Ділення щільностей ймовірностей приводить до виразу, який має такий же вигляд.

Таким чином, умовна щільність ймовірностей виявляється гауссівською.

Властивість 4. Будь-яка гауссівська випадкова функція $X(t)$ з корельованими (залежними) перерізами може бути зведена до гауссівської випадкової функції з некорельованими (незалежними) перерізами.

Доведення. Відмітимо, що коли кореляційна матриця є діагональною, тобто $R_x(t_n, t_m) = 0$ при $t_n \neq t_m$, то перерізи відповідної гауссівської випадкової функції є некорельованими. Тому задача зведення перерізів гауссівської випадкової функції $X(t)$ до сукупності некорельованих гауссівських випадкових величин зводиться до діагоналізації кореляційної матриці. Діагоналізація досягається шляхом обчислення власних векторів кореляційної матриці, а потім перетворенням координатних осей. Координати зміщують на величину математичних сподівань і повертають так, щоб вони збігалися з власними векторами.

Властивість 5. У результаті лінійного перетворення гауссівської випадкової функції лінійним оператором одержується гауссівська випадкова функція. Ця властивість називається властивістю *сталості гауссівських випадкових функцій* по відношенню до лінійних перетворень.

Властивість 6. При нелінійному перетворенні гауссівської випадкової функції властивість гауссівості втрачається.

Властивість 7. Характеристична функція, що відповідає щільності ймовірностей (3.1.1) має такий вигляд:

$$Q_N(j\vec{\omega}; \vec{t}) = \exp\left(j\vec{\omega}^T \vec{m}_x - \frac{1}{2} \vec{\omega}^T R_x \vec{\omega}\right).$$

Властивість 8. При фіксованій дисперсії гауссівська випадкова функція має максимальну ентропію.

3.2. Стаціонарні та нестаціонарні випадкові функції

3.2.1. Стаціонарні випадкові функції та їх опис

Положення 1. У кореляційній теорії особливу роль відіграють так звані стаціонарні випадкові функції. У різних випадках під цим терміном розуміють різні поняття. Розглянемо найбільш поширені його визначення.

Визначення 1. Випадкова функція $X(t)$ називається *стаціонарною у вузькому розумінні (строго)*, якщо її N -вимірні ймовірнісні характеристики при будь-якому N залежать тільки від довжини інтервалів $t_2 - t_1, \dots, t_N - t_1$ і не залежать від положення цих інтервалів на осі t , де t_1, \dots, t_N – значення аргументу функції $X(t)$.

Випадкові функції, що не відносяться до класу стаціонарних, називаються *нестаціонарними*.

Властивість 1. Основною властивістю стаціонарної функції є незалежність її ймовірнісних характеристик будь-якої N -ї вимірності від зсуву значень аргументу t , зокрема N -вимірна щільність ймовірностей

$$f_N(x_1, \dots, x_N; t_1, \dots, t_N) = f_N(x_1, \dots, x_N; t_1 + \tau, \dots, t_N + \tau), \quad (3.2.1)$$

де τ – довільне число.

Властивість 2. Одновимірні ймовірнісні характеристики стаціонарної випадкової функції $X(t)$ не залежать від аргументу t , зокрема щільність ймовірностей $f_1(x; t) = f_1(x)$.

Властивість 3. Двовимірні ймовірнісні характеристики стаціонарної випадкової функції $X(t)$ залежать лише від різниці $\tau = t_2 - t_1$ значень аргументу t , зокрема щільність ймовірностей $f_2(x_1, x_2; t_1, t_2) = f_2(x_1, x_2; \tau)$.

Моментні функції стаціонарної випадкової функції мають специфічні особливості, що наведені в табл. 3.2.1.

Таблиця 3.2.1

Особливості моментних функцій стаціонарних випадкових функцій

Особливість
$m_x(t) = m_x = \text{const}$
$D_x(t) = D_x = \text{const}$
$R_x(t_1, t_2) = R_x(t_1 - t_2) = R_x(\tau)$
$r_x(\tau) = \frac{R_x(\tau)}{D_x}$

Зауваження 1. При розв'язанні задач у рамках кореляційної теорії будь-які особливості ймовірнісних характеристик вимірності $N > 2$ не відіграють суттєвої ролі, зокрема характерна для стаціонарної у вузькому розумінні випадкової функції особливість залежності ймовірнісних характеристик вимірності $N > 2$ лише від різниці значень аргументу t .

Зауваження 2. Іноді розглядають інші стаціонарні у вузькому розумінні випадкові функції (див. визначення 1а – 1г).

Визначення 1а. Випадкова функція $X(t)$ називається *стаціонарною у вузькому розумінні порядку K* , якщо її ймовірнісні характеристики порядку $N \leq K$ інваріантні до зсуву вздовж осі t , зокрема рівність (3.2.1) має місце при $N \leq K$.

Примітка. Якщо рівність (3.2.1) слухна при $N = K$, то, як впливає з властивості узгодженості щільності ймовірностей, ця рівність слухна й при $N < K$. Тому визначення 1а можна переформулювати, змінивши умову $N \leq K$ на умову $N = K$.

Визначення 1б. Випадкова функція $X(t)$ називається *асимптотично стаціонарною у вузькому розумінні*, якщо її N -вимірні ймовірнісні характеристики при усіх N мають границі при спрямованості до нескінченності зсуву уздовж осі t , зокрема існує границя щільності ймовірностей:

$$\lim_{\tau \rightarrow \infty} f_N(x_1, \dots, x_N; t_1 + \tau, \dots, t_N + \tau).$$

Визначення 1в. Випадкова функція $X(t)$ називається *стаціонарною у вузькому розумінні на інтервалі T* , якщо вона відповідає визначенню 1 для усіх $t \in T$.

Визначення 1г. Випадкова функція $X(t)$ називається *періодично стаціонарною (циклостаціонарною) у вузькому розумінні*, якщо її N -вимірні ймовірнісні характеристики при усіх N періодичні з періодом T_0 , зокрема

$$f_N(x_1, \dots, x_N; t_1, \dots, t_N) = f_N(x_1, \dots, x_N; t_1 + kT_0, \dots, t_N + kT_0),$$

де $k = 0, \pm 1, \dots$

Поняття стаціонарності у вузькому розумінні та усі його модифікації узагальнюються на випадок двох та декількох випадкових функцій. Зупинимось на основній модифікації цього поняття.

Визначення 2. Дві випадкові функції $X(t)$ і $Y(t)$ називаються *сумісно стаціонарно пов'язаними у вузькому розумінні*, якщо їх сумісні $(N + M)$ -вимірні ймовірнісні характеристики при довільних N і M залежать лише від довжини інтервалів $t_2 - t_1, \dots, t_N - t_1$; $t'_1 - t_1, \dots, t'_M - t_1$, де t_1, \dots, t_N – значення аргументу функції $X(t)$, а t'_1, \dots, t'_M – значення аргументу функції $Y(t)$.

Зауваження 1. З наведеного визначення випливає, що, зокрема, сумісна щільність ймовірностей

$$f_{N+M}(x_1, \dots, x_N, y_1, \dots, y_M; t_1, \dots, t_N, t'_1, \dots, t'_M) = f_{N+M}(x_1, \dots, x_N, y_1, \dots, y_M; t_1 + \tau, \dots, t_N + \tau, t'_1 + \tau, \dots, t'_M + \tau).$$

Зауваження 2. З стаціонарності випадкових функцій $X(t)$ і $Y(t)$ не випливає, що вони сумісно стаціонарно пов'язані у вузькому розумінні.

На практиці використовують ще інше поняття стаціонарності.

Визначення 3. Випадкова функція $X(t)$ називається *стаціонарною в широкому розумінні*, якщо її математичне сподівання стає ($m_x(t) = m_x = \text{const}$), а коваріаційна функція залежить тільки від різниці значень аргументу t : $K_x(t_1, t_2) = K_x(t_1 - t_2)$.

Зауваження 1. Поняття стаціонарності у вузькому і широкому розумінні не тотожні. Випадкові функції, які стаціонарні у вузькому розумінні, стаціонарні і в широкому розумінні. Зворотне твердження в загальному випадку невірне (рис. 3.2.1).

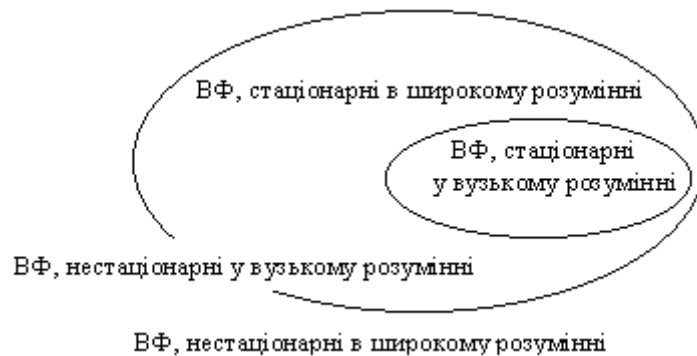


Рис. 3.2.1. Співвідношення між стаціонарними і нестационарними випадковими функціями різних типів

Зауваження 2. Для гауссівських випадкових функцій поняття стаціонарності у вузькому і широкому розумінні *тотожні*.

Зауваження 3. Інколи використовують поняття стаціонарних у широкому розумінні випадкових функцій, що аналогічні розглянутим вище модифікаціям стаціонарних у вузькому розумінні випадкових функцій.

На практиці використовують також поняття *сумісно стаціонарно пов'язаних у широкому розумінні випадкових функцій*, що аналогічне поняттю сумісно стаціонарно пов'язаних у вузькому розумінні випадкових функцій.

Визначення 4. Дві випадкові функції $X(t)$ і $Y(t)$ називаються *сумісно стаціонарно пов'язаними у широкому розумінні*, якщо їх математичні сподівання стали, а взаємна коваріаційна функція інваріантна до зсуву вздовж осі t :

$$K_{xy}(t_1, t_2) = M_{xy}[X(t_1)Y(t_2)] = K_{xy}(\tau), \quad \tau = t_2 - t_1. \quad (3.2.2)$$

Зауваження. Як і у випадку стаціонарних у вузькому розумінні випадкових функцій, стаціонарність у широкому розумінні випадкових функцій *не гарантує* їх стаціонарну пов'язаність у широкому розумінні.

Положення 2. В подальшому під стаціонарною випадковою функцією будемо розуміти функцію, *стаціонарну в широкому розумінні*.

Властивості. Стаціонарні випадкові функції мають ряд специфічних властивостей, що наведені у таблиці 3.2.2.

Таблиця 3.2.2

Властивості кореляційних функцій (КФ) і нормованих кореляційних функцій (НКФ) дійсних стаціонарних випадкових функцій

Властивість
$ R_x(\tau) \leq D_x; r_x(\tau) \leq 1$
Максимум КФ і НКФ знаходиться при $\tau = 0$: $R_x(0) = D_x; r_x(0) = 1$
КФ і НКФ – парні: $R_x(\tau) = R_x(-\tau); r_x(\tau) = r_x(-\tau)$
Взаємні КФ і НКФ не є парними, однак $R_{xy}(\tau) = R_{yx}(-\tau); r_{xy}(\tau) = r_{yx}(-\tau)$
Якщо КФ неперервна при $\tau = 0$, то вона неперервна при будь-якому τ

Зауваження. Стаціонарні випадкові функції, що зустрічаються на практиці, як правило, мають кореляційну функцію, що неперервно спадає зі збільшенням $|\tau|$, або спадає з деякою осциляцією (рис. 3.2.2).

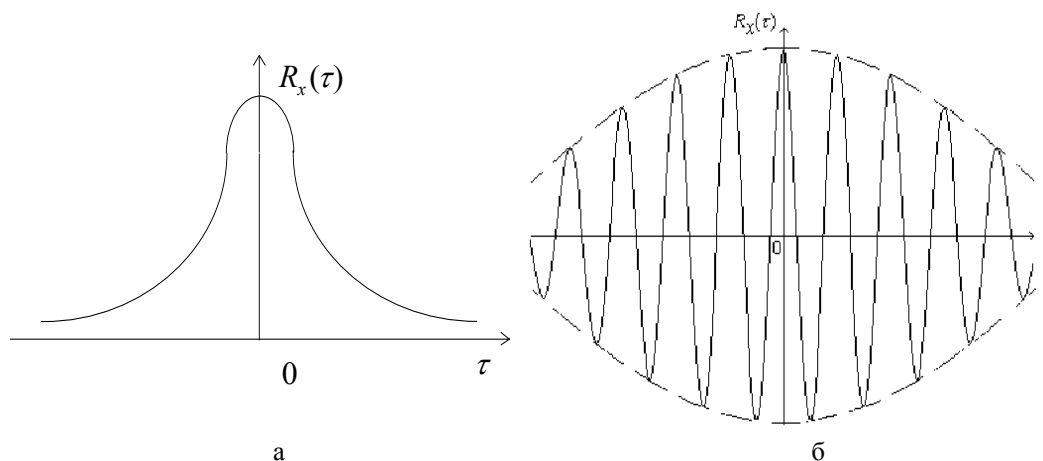


Рис. 3.2.2. Кореляційна функція широкопasmової (а) і вузькопasmової (б) випадкових функцій. Визначення понять широкопasmовості і вузькопasmовості випадкової функції дано в розділі 3.5.3. На рисунку 3.2.2,б пунктиром зображені огинаючі функції. Це поняття конкретизовано у наступному підрозділі

Положення 3. Термін “стаціонарність” найчастіше пов’язують з часом. Для стаціонарних випадкових функцій одного аргументу, фізичний зміст якого не є час, використовують термін “однорідність”. У випадку функцій декількох аргументів, серед яких є час, застосовують обидва терміни.

Приклад. Стаціонарне однорідне випадкове поле – поле, що є стаціонарним у часі та однорідним по інших аргументах.

Зауваження. Випадкове поле може бути стаціонарним, але неоднорідним, може бути однорідним по деяких аргументах та неоднорідним по інших і т.д.

3.2.2. Опис нестационарних випадкових функцій

Положення 1. Деякі нестационарні випадкові функції можуть бути зведені до стаціонарних.

Приклад. Нестационарна випадкова функція $Y(t) = \varphi(t)X(t) + g(t)$, де $\varphi(t), g(t)$ – неслучайні функції, $X(t)$ – випадкова стаціонарна функція, зводиться до стаціонарної центруванням на $g(t)$ і нормуванням на $\varphi(t)$.

Для такої функції справедливі такі співвідношення:

$$m_y(t) = \varphi(t)m_x + g(t);$$

$$R_y(t_1, t_2) = \varphi(t_1)\varphi(t_2)R_x(t_2 - t_1);$$

$$D_y(t) = \varphi^2(t)D_x;$$

$$r_y(t_1, t_2) = r_x(t_2 - t_1).$$

Положення 2. Для опису нестационарних випадкових функцій іноді застосовують характеристики, що є аналогом кореляційних моментів. Розглянемо одну з них, що використовують для так званих функцій зі стаціонарними першими приростами.

Визначення 1. Структурною функцією випадкової функції $X(t)$ називається функція

$$B_x(t_1, t_2) = M_x[(X(t_1) - X(t_2))^2]. \quad (3.2.3)$$

Визначення 2. Функцією зі стаціонарними першими приростами називається випадкова функція $X(t)$, приріст якої $\Delta X_\theta(t) = X(t + \theta) - X(t)$ за інтервал θ являє собою стаціонарну у широкому розумінні випадкову функцію, тобто для якої

$$m_\theta = M_\theta[\Delta X_\theta(t)] = const, \quad (3.2.4)$$

$$R_\theta(\tau) = M_\theta[\Delta X_\theta(t)\Delta X_\theta(t + \tau)] - m_\theta^2. \quad (3.2.5)$$

Зауваження 1. Клас випадкових функцій зі стаціонарними першими приростами включає клас усіх стаціонарних випадкових функцій.

Зауваження 2. Для стаціонарних випадкових функцій

$$m_\theta = 0; \quad (3.2.6)$$

$$R_\theta(\tau) = 2R_x(\tau) - R_x(\tau - \theta) - R_x(\tau + \theta); \quad (3.2.7)$$

$$B_x(\tau) = 2[R_x(0) - R_x(\tau)], \quad (3.2.8)$$

що випливає з формул (3.2.3) – (3.2.5).

Зауваження 3. Іноді використовують функції зі стаціонарними приростами другого порядку, що визначаються аналогічним чином.

3.2.3. Інтервали кореляції

Перейдемо до розгляду параметрів, що характеризують кореляцію, але спочатку зупинимося на терміні “огинаюча”, яка буде використана при цьому.

Визначення 1. Огинаючою $y = f(x)$ однопараметричного (відносно параметра λ) сімейства кривих $\varphi(x, y, \lambda) = 0$ називається крива, яка або торкається кожною своєю точкою однієї з кривих сімейства, або водночас належить всім кривим цього сімейства.

Зауваження 1. Рівняння огинаючої одержується з системи рівнянь

$$\begin{cases} \varphi(x, y, \lambda) = 0, \\ \frac{\partial \varphi(x, y, \lambda)}{\partial \lambda} = 0 \end{cases} \quad (3.2.9)$$

виключенням параметра λ .

Зауваження 2. Сімейство кривих може не мати огинаючої, а може мати декілька огинаючих.

Зауваження 3. Система рівнянь (3.2.9) визначає не тільки точки огинаючої, але також і особливі точки.

Зауваження 4. Виключення параметра λ , можливо, і приводить до огинаючої, якщо в області простору x, y і λ , що розглядається, якобіан

$$\begin{vmatrix} \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi'}{\partial x} \\ \frac{\partial \varphi}{\partial y} & \frac{\partial \varphi'}{\partial y} \end{vmatrix} \neq 0, \quad \varphi''_{\lambda\lambda} \neq 0. \quad (3.2.10)$$

Приклад. Огинаючою сімейства кривих $y = A(t) \cos 2\pi ft$ (де f – частота, t – час) є криві $y = A(t)$ і $y = -A(t)$, які відповідають параметрам $f = \frac{k}{t}$ і $f = \frac{2k+1}{2t}$ ($k = 0, \pm 1, \dots$).

Положення 1. Якщо сімейство однопараметричних кривих описується відносно параметра λ виразом $y = A(t) \cos \Phi(t, \lambda)$, де $A(t)$ – неперервна функція аргументу t , $\Phi(t, \lambda)$ – неперервна функція аргументів t, λ , то серед огинаючих цього сімейства є огинаючі $y = A(t)$ і $y = -A(t)$.

Зауваження. Якщо деякий сигнал представлено виразом $y = A(t) \cos \Phi(t)$, то $A(t)$ називається *огинаючою сигналу*, а $\Phi(t)$ – його *фазою*.

Визначення 2. *Інтервалом кореляції* τ_k випадкової функції називається таке значення аргументу τ кореляційної функції $R_x(\tau)$, починаючи з якого значення кореляційної функції (або огинаючої) стає настільки малим, що ним можна знехтувати.

Зауваження. Наведене визначення якісного характеру може бути конкретизоване одним з таких способів:

1) *інтервал кореляції* – це значення аргументу кореляційної функції, починаючи з якого нормована кореляційна функція (або її огинаюча) стає меншою за деяку величину (наприклад 0,1);

2) *інтервал кореляції* – це значення аргументу кореляційної функції, що дорівнює площі під кривою модуля нормованої кореляційної функції $r_x(\tau)$ або її огинаючої $\rho_x(\tau)$:

$$\tau_k = \int_0^{\infty} |r_x(\tau)| d\tau, \quad (3.2.11)$$

$$\tau_k = \int_0^{\infty} |\rho_x(\tau)| d\tau. \quad (3.2.12)$$

Все викладене досить очевидним чином узагальнюється на випадок векторних випадкових функцій.

3.2.4. Стаціонарні векторні випадкові функції

Визначення 1. Векторна функція $\vec{X}(t) = (X(t_1), \dots, X(t_N))$ називається *стаціонарною в широкому розумінні*, якщо математичне сподівання її стале, а всі елементи коваріаційної функції $K_{\vec{x}}(t_1, t_2)$ (включаючи взаємні) залежать тільки від різниці аргументів $\tau = t_2 - t_1$.

Деякі властивості векторної стаціонарної випадкової функції збігаються з властивостями скалярної стаціонарної випадкової функції. Основні властивості наведені в табл. 3.2.3.

Таблиця 3.2.3

Властивості векторних стаціонарних випадкових функцій

Властивості
$R_{hl}(-\tau) = R_{lh}(\tau)$
$ R_{hl}(\tau) \leq \sqrt{D_h D_l}$
$ r_{hl}(\tau) \leq 1$

3.3. Ергодичні випадкові функції

Для деяких стаціонарних (однорідних) випадкових функцій розрахунок наймовірніших характеристик може бути проведено не шляхом усереднення по ансамблю реалізацій, а усередненням даних лише однієї реалізації. Такі випадкові функції називаються ергодичними. Відомо декілька визначень поняття ергодичності випадкових функцій. Розглянемо їх.

Визначення 1. Випадкова функція (процес) називається *ергодичною*, якщо будь-яка її наймовірнісна характеристика, отримана усередненням по множині можливих реалізацій, з ймовірністю скільки завгодно близькою до одиниці, дорівнює часовому середньому, отриманому з однієї-єдиної реалізації випадкової функції шляхом усереднення за нескінченний інтервал часу.

Теорема 1. Необхідною умовою ергодичності випадкової функції є її стаціонарність у вузькому розумінні.

Визначення 2. Випадкова функція (процес) називається *ергодичною по відношенню до математичного сподівання*, якщо (дивись додаток 3)

$$\text{l.i.m.}_{T \rightarrow \infty} \frac{1}{T} \int_{\frac{T}{2}}^{\frac{T}{2}} X(t) dt = m_x \quad (3.3.1)$$

Теорема 2. Випадкова функція $X(t)$ ергодична по відношенню до математичного сподівання тоді і тільки тоді, коли математичне сподівання m_x випадкової функції стале, а її кореляційна функція $R_x(t_1, t_2)$ задовольняє співвідношення

$$\lim_{T \rightarrow \infty} \frac{1}{T^2} \int_{\frac{T}{2}}^{\frac{T}{2}} \int_{\frac{T}{2}}^{\frac{T}{2}} R_x(t_1, t_2) dt_1 dt_2 = 0 \quad (3.3.2)$$

Ця теорема носить назву *ергодичної теореми*.

Зауваження 1. Достатньою умовою справедливості рівності (3.3.2) є

$$\lim_{|t_2 - t_1| \rightarrow \infty} |R_x(t_1, t_2)| = 0 \quad (3.3.3)$$

тобто кореляційна функція $R_x(t_1, t_2)$ прямує до нуля при необмеженому збільшенні модуля різності аргументів $|t_2 - t_1|$.

Перевірка умови (3.3.3) звичайно викликає менше труднощів, ніж умови (3.3.2).

Зауваження 2. Для стаціонарних випадкових функцій умова (3.3.2) може бути записана таким чином:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(1 - \frac{|\tau|}{T}\right) R_x(\tau) d\tau = 0.$$

Достатньою умовою при цьому є $\lim_{|\tau| \rightarrow \infty} |R_x(\tau)| = 0$.

Зауваження 3. Теорема 2 має виключну цінність. Справа в тому, що розрахунок математичного сподівання шляхом усереднення реалізацій по великому ансамблю звичайно дуже важко реалізувати на практиці. Зробити ж розрахунок математичного сподівання шляхом усереднення за часом всього лише однієї реалізації або, у крайньому випадку, декількох значно простіше.

Визначення 3. Стаціонарна випадкова функція $X(t)$ називається *ергодичною по відношенню до кореляційної функції* $R_x(\tau)$, якщо

$$R_x(\tau) = \text{l.i.m.}_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t + \tau) X^*(t) dt.$$

Теорема 3. Необхідною і достатньою умовами ергодичності стаціонарної випадкової функції по відношенню до кореляційної функції $R_x(\tau)$ є рівність

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(1 - \frac{\tau}{t}\right) [R_x^2(\tau) + R_x(\tau + \tau_0) R_x(\tau - \tau_0)] d\tau = 0$$

при будь-якому фіксованому τ_0 .

3.4. Кореляційні та коваріаційні функції комплексних випадкових функцій

На практиці важливу роль відіграють кореляційні та коваріаційні функції комплексних випадкових функцій. Зупинимося на них.

Визначення 1. Коваріаційною функцією $\dot{K}_x(t_1, t_2)$ комплексної випадкової функції $\dot{X}(t)$ називається двовимірна початкова моментна функція другого порядку:

$$\dot{K}_x(t_1, t_2) = M[\dot{X}(t_1) X^*(t_2)]. \quad (3.4.1)$$

Визначення 2. Кореляційною функцією $\dot{R}_x(t_1, t_2)$ комплексної випадкової функції $\dot{X}(t)$ називається двовимірна центральна моментна функція другого порядку:

$$\dot{R}_x(t_1, t_2) = M[(\dot{X}(t_1) - \dot{m}_x(t_1))(X^*(t_2) - m_x^*(t_2))]. \quad (3.4.2)$$

Зауваження 1. Коваріаційна функція зв'язана з кореляційною функцією таким чином:

$$\dot{K}_x(t_1, t_2) = \dot{R}_x(t_1, t_2) + \dot{m}_x(t_1) m_x^*(t_2). \quad (3.4.3)$$

Зауваження 2. Якщо математичне сподівання дорівнює нулеві, то коваріаційна і кореляційна функції збігаються: $\dot{K}_x(t_1, t_2) = \dot{R}_x(t_1, t_2)$.

Визначення 3. Значення комплексної випадкової функції $\dot{X}(t)$ у моменти часу t_1 і t_2 називаються *некорельованими*, якщо кореляційна функція $\dot{R}_x(t_1, t_2) = 0$, тобто

$$\dot{K}_x(t_1, t_2) = \dot{m}_x(t_1)m_x^*(t_2). \quad (3.4.4)$$

Визначення 4. Значення комплексної випадкової функції $\dot{X}(t)$ у моменти часу t_1 і t_2 називаються *ортгональними*, якщо коваріаційна функція $\dot{K}_x(t_1, t_2) = 0$, тобто

$$\dot{R}_x(t_1, t_2) = -\dot{m}_x(t_1)m_x^*(t_2). \quad (3.4.5)$$

Визначення 5. *Взаємною коваріаційною функцією* $\dot{K}_{xy}(t_1, t_2)$ комплексних випадкових функцій $\dot{X}(t)$ і $\dot{Y}(t)$ називається двовимірна початкова взаємна моментна функція другого порядку:

$$\dot{K}_{xy}(t_1, t_2) = M[\dot{X}(t_1)Y^*(t_2)]. \quad (3.4.6)$$

Визначення 6. *Взаємною кореляційною функцією* $\dot{R}_{xy}(t_1, t_2)$ комплексних випадкових функцій $\dot{X}(t)$ і $\dot{Y}(t)$ називається двовимірна центральна взаємна моментна функція другого порядку:

$$\dot{R}_{xy}(t_1, t_2) = M[(\dot{X}(t_1) - \dot{m}_x(t_1))(Y^*(t_2) - m_y^*(t_2))]. \quad (3.4.7)$$

Положення 1. Коваріаційні й кореляційні зв'язки між перерізами комплексних випадкових функцій $\dot{X}(t)$ і $\dot{Y}(t)$ у моменти часу t_1 і t_2 описуються коваріаційною і кореляційною матрицями:

$$\dot{K}(t_1, t_2) = \begin{vmatrix} \dot{K}_x(t_1, t_2) & \dot{K}_{xy}(t_1, t_2) \\ \dot{K}_{yx}(t_1, t_2) & \dot{K}_y(t_1, t_2) \end{vmatrix}, \quad (3.4.8)$$

$$\dot{R}(t_1, t_2) = \begin{vmatrix} \dot{R}_x(t_1, t_2) & \dot{R}_{xy}(t_1, t_2) \\ \dot{R}_{yx}(t_1, t_2) & \dot{R}_y(t_1, t_2) \end{vmatrix}. \quad (3.4.9)$$

Визначення 7. Комплексні випадкові функції $\dot{X}(t)$ і $\dot{Y}(t)$ називаються *некорельованими*, якщо для двох довільних моментів часу t_1 і t_2 взаємна кореляційна функція $\dot{R}_{xy}(t_1, t_2)$ дорівнює нулеві: $\dot{R}_{xy}(t_1, t_2) = 0$, тобто матриця (3.4.9) – діагональна.

Визначення 8. Комплексні випадкові функції $\dot{X}(t)$ і $\dot{Y}(t)$ називаються *ортгональними*, якщо для двох довільних моментів часу t_1 і t_2 взаємна коваріаційна функція $\dot{K}_{xy}(t_1, t_2)$ дорівнює нулеві: $\dot{K}_{xy}(t_1, t_2) = 0$, тобто матриця (3.4.8) – діагональна.

Визначення 9. *Нормованими кореляційними функціями* називаються функції

$$\dot{r}_x(t_1, t_2) = \frac{\dot{R}_x(t_1, t_2)}{\sqrt{D_x(t_1)D_x(t_2)}}, \quad (3.4.10)$$

$$\dot{r}_{xy}(t_1, t_2) = \frac{\dot{R}_{xy}(t_1, t_2)}{\sqrt{D_x(t_1)D_y(t_2)}}. \quad (3.4.11)$$

Зауваження. Кореляційні і нормовані кореляційні функції характеризують *лінійні і тільки лінійні зв'язки* між різними перерізами однієї чи двох функцій.

Властивості. Кореляційні функції комплексних випадкових функцій мають властивості, що широко використовуються на практиці. Деякі з них наведено в табл. 3.4.1.

Таблиця 3.4.1

Властивості кореляційних функцій $\dot{R}_{xy}(t_1, t_2)$, $\dot{R}_x(t_1, t_2)$ і нормованих кореляційних функцій $\dot{r}_{xy}(t_1, t_2)$, $\dot{r}_x(t_1, t_2)$ комплексних випадкових функцій $\dot{X}(t)$, $\dot{Y}(t)$

№ з/п	Назва	Математичне трактування
1	Ермітовість	$\dot{R}_{xy}(t_1, t_2) = R_{yx}^*(t_2, t_1)$; $\dot{r}_{xy}(t_1, t_2) = r_{yx}^*(t_2, t_1)$
2	Обмеженість	$ \dot{R}_{xy}(t_1, t_2) ^2 \leq D_x(t_1)D_y(t_2)$; $ \dot{r}_{xy}(t_1, t_2) \leq 1$ (знак рівності має місце при $\dot{Y}(t_1) = a\dot{X}(t_1) + b$, де a і b – константи)
3	Невід’ємна визначеність	Для довільних t_1, \dots, t_N і комплексних $\dot{z}_1, \dots, \dot{z}_N$ $\sum_{m,n=1}^N \dot{R}_x(t_m, t_n) \dot{z}_m \dot{z}_n^*$ дійсна і невід’ємна

Примітка. Доведення властивостей 1, 3 будується на визначеннях взаємної кореляційної і кореляційної функцій, а властивості 2 – на нерівності Коші – Шварца.

3.5. Спектральний аналіз стаціонарних випадкових функцій

3.5.1. Основи спектрального аналізу

Теорема 1. Якщо періодична з періодом T дійсна функція $x_T(t)$ абсолютно інтегровна, то вона може бути представлена у вигляді збіжного ряду Фур’є:

$$x_T(t) = \sum_{\nu=-\infty}^{\infty} \dot{A}_\nu \exp(j \frac{2\pi\nu}{T} t), \quad (3.5.1)$$

де

$$\dot{A}_\nu = \frac{1}{T} \int_T x_T(t) e^{-j \frac{2\pi\nu}{T} t} dt. \quad (3.5.2)$$

Вирази (3.5.1) і (3.5.2) описують розкладення періодичної функції по гармонічних коливаннях (гармоніках) з круговими частотами $\frac{2\pi\nu}{T} = \omega$ і комплексними амплітудами \dot{A}_ν . Модуль ν -ї комплексної амплітуди $A_\nu = |\dot{A}_\nu|$ називається ν -ю амплітудою, а аргумент $\theta_\nu = \arg(\dot{A}_\nu)$ – ν -ю початковою фазою.

Визначення 1. Сукупність амплітуд $|\dot{A}_\nu|$, які розглядаються як функція ν , називається математичним амплітудним спектром; сукупність фаз θ_ν – математичним фазовим спектром; а сукупність комплексних амплітуд \dot{A}_ν – математичним амплітудно-фазовим комплексним спектром.

Теорема 2. Якщо неперіодична дійсна функція $x(t)$ абсолютно інтегровна, то вона може бути представлена у вигляді інтеграла Фур’є:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{A}(\omega) e^{j\omega t} d\omega, \quad (3.5.3)$$

де

$$\dot{A}(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt. \quad (3.5.4)$$

Зауваження. Ряд Фур'є або інтеграл Фур'є, породжений абсолютно інтегрованою дійсною функцією $x(t)$, рівномірно збігається до $x(t)$ на кожному закритому інтервалі, в якому функція $x(t)$ неперервна та збігається до $\frac{1}{2}[x(t-0) + x(t+0)]$ на кожному відкритому інтервалі, де функція $x(t)$ та її похідна $x'(t)$ кусково-неперервні або де функція $x(t)$ має обмежену варіацію (де функція обмежена та має скінченну кількість екстремумів і точок розриву першого роду).

Спектри мають ряд важливих властивостей. Наведемо деякі з них.

Властивість 1. Математичні спектри періодичної функції $x_T(t)$ дискретні, а неперіодичної $x(t)$ – неперервні (рис. 3.5.1).

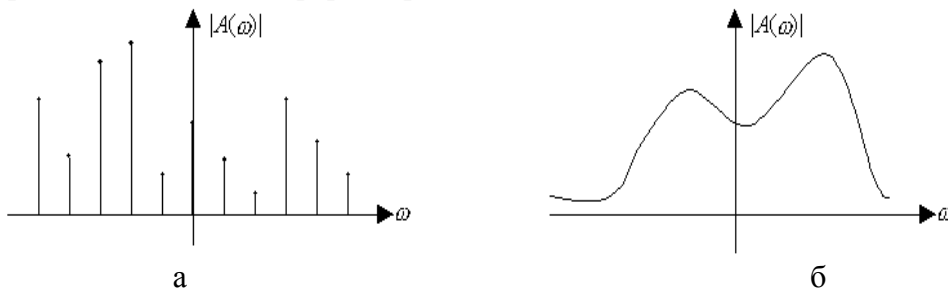


Рис. 3.5.1. Амплітудні спектри періодичної (а) і неперіодичної (б) функцій

Властивість 2. Якщо функція дійсна, то її математичний амплітудний спектр $A(\omega)$ – парний, а математичний фазовий спектр $\theta(\omega)$ – непарний, тобто $\dot{A}(\omega) = A^*(-\omega)$.

Зауваження. У формулах (3.5.3) і (3.5.4) кругова частота $\omega = 2\pi f$ (f – частота) приймає як додатні, так і від'ємні значення. Використання поняття від'ємних кругових частот зручно з математичної точки зору, однак не має фізичного сенсу.

Для дійсних функцій $x(t)$ від поняття математичного спектра $\dot{A}(\omega)$ можна перейти до поняття фізичного спектра $\dot{A}^+(\omega)$

Визначення 2. Фізичним спектром $\dot{A}^+(\omega)$ дійсної функції $x(t)$ називається функція, яка визначається для невід'ємних кругових частот ω як

$$\dot{A}^+(\omega) = \begin{cases} \dot{A}(\omega) + \dot{A}(-\omega) = 2\dot{A}(\omega) & \text{при } \omega > 0, \\ \dot{A}(0) & \text{при } \omega = 0. \end{cases} \quad (3.5.5)$$

Зауваження 1. Абсолютно інтегровна неперіодична дійсна функція $x(t)$ може бути представлена таким чином:

$$x(t) = \frac{1}{2\pi} \int_0^{\infty} A^+(\omega) \cos[\omega t + \theta^+(\omega)] d\omega,$$

де $A^+(\omega)$; $\theta^+(\omega)$ – відповідно амплітуда і фаза фізичного спектру:

$$A^+(\omega) = |\dot{A}^+(\omega)|, \quad \theta^+(\omega) = \arg(\dot{A}^+(\omega)),$$

$$\dot{A}^+(\omega) = A^+(\omega) e^{j\theta^+(\omega)} = \begin{cases} 2 \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt & \text{при } \omega > 0, \\ \int_{-\infty}^{\infty} x(t) dt & \text{при } \omega = 0. \end{cases}$$

Зауваження 2. Аналогічно абсолютно інтегровна періодична функція $x_T(t)$ з періодом T може бути представлена у вигляді збіжного ряду Фур'є:

$$x(t) = \sum_{\nu=0}^{\infty} A_{\nu}^+ \exp(j \frac{2\pi\nu}{T} t + \theta_{\nu}^+),$$

де A_{ν}^+ і θ_{ν}^+ – відповідно амплітуда і фаза ν -ї гармоніки фізичного спектру:

$$A_{\nu}^+ = |\dot{A}_{\nu}^+|, \quad \theta_{\nu}^+ = \arg(\dot{A}_{\nu}^+),$$

$$\dot{A}_{\nu}^+ = A_{\nu}^+ e^{j\theta_{\nu}^+} = \begin{cases} \frac{2}{T} \int_T x_T(t) e^{-j\frac{2\pi\nu}{T} t} dt & \text{при } \nu > 0, \\ \frac{1}{T} \int_T x_T(t) dt & \text{при } \nu = 0. \end{cases}$$

3.5.2. Спектральне представлення випадкових функцій

Апарат спектрального аналізу ефективно використовується для опису випадкових функцій. Методику його застосування розробили у 30-х роках американський вчений Н. Вінер і радянський вчений А.Я. Хінчин.

Визначення 1. Миттєвим (випадковим) спектром $\dot{A}_x(f)$ випадкової функції $X(t)$ називається комплексна випадкова функція $\dot{A}_x(f)$, що пов'язана з $X(t)$ перетворенням Фур'є:

$$\dot{A}_x(f) = \int_{-\infty}^{\infty} X(t) e^{-j2\pi f t} dt.$$

Положення 1. Миттєвий спектр стаціонарної у широкому розумінні випадкової функції має ряд специфічних властивостей, які впливають з його визначення та властивостей стаціонарних функцій.

Властивість 1. Математичне сподівання $m_a(f)$ миттєвого спектра стаціонарної випадкової функції $X(t)$ пов'язано з математичним сподіванням m_x самої випадкової функції виразом $m_a(f) = m_x \delta(f)$, де $\delta(f)$ – дельта-функція.

Властивість 2. Коваріаційна функція миттєвого спектра стаціонарної випадкової функції

$$K_a(f_1, f_2) = S_x(f_1) \delta(f_2 - f_1), \quad (3.5.6)$$

де $S_x(f) = \int_{-\infty}^{\infty} K_x(\tau) e^{-j2\pi f \tau} d\tau$ – функція, що називається спектральною щільністю потужності (визначення цієї функції та її властивості наведено далі).

Властивість 3. З виразу (3.5.6) випливає, що миттєвий спектр стаціонарної випадкової функції не є стаціонарною випадковою функцією.

Властивість 4. Відліки миттєвого спектра, що відповідають різним частотам, ортогональні.

Властивість 5. Якщо математичне сподівання стаціонарної випадкової функції дорівнює нулеві, то відліки миттєвого спектра, що відповідають різним частотам, не тільки ортогональні, але й *некорельовані*.

Визначення 2. *Спектральною щільністю потужності*, просто *спектральною щільністю* або *енергетичним спектром* стаціонарної випадкової функції $X(t)$ називається не випадкова функція $S_x(f)$, що визначається через коваріаційну функцію $K_x(\tau)$ і зв'язана з нею парою перетворень Фур'є:

$$S_x(f) = \int_{-\infty}^{\infty} K_x(\tau) e^{-j2\pi f\tau} d\tau, \quad (3.5.7)$$

$$K_x(\tau) = \int_{-\infty}^{\infty} S_x(f) e^{j2\pi f\tau} df. \quad (3.5.8)$$

Перетворення (3.5.7), (3.5.8) називаються *перетвореннями Вінера – Хінчина*. Розглянемо особливості спектральної щільності потужності $S_x(f)$.

Особливість 1. Представлення в формулі (3.5.7) коваріаційної функції $K_x(\tau)$ через кореляційну функцію і інтегрування отриманого виразу приводить до висновку, що спектральна щільність потужності $S_x(f)$ стаціонарної випадкової функції $X(t)$ пов'язана зі спектральною щільністю потужності $S_x^0(f)$

центрованої випадкової функції $\overset{0}{X}(t)$ та математичним сподіванням m_x випадкової функції $X(t)$ таким чином:

$$S_x(f) = S_x^0(f) + m_x^2 \delta(f), \quad (3.5.9)$$

тобто спектральні щільності потужності $S_x(f)$ і $S_x^0(f)$ відрізняються лише на нульовій частоті.

Особливість 2. З виразу (3.5.8) випливає, що дисперсія центрованої стаціонарної випадкової функції

$$D_x = R_x(0) = \int_{-\infty}^{\infty} S_x^0(f) df, \quad (3.5.10)$$

тобто дисперсія дорівнює площі під кривою спектральної щільності потужності $S_x^0(f)$.

Зауваження. З виразу (3.5.10) видно, що енергетичний спектр $S_x^0(f)$ характеризує розподіл дисперсії випадкової функції за частотою.

Особливість 3. На відміну від миттєвого спектра $\dot{A}_x(f)$ енергетичний спектр $S_x(f)$ є характеристикою не якоїсь однієї реалізації, а *всієї множини реалізацій*. При цьому він не містить інформації про фазову структуру реалізації і характеризує в середньому амплітудний спектр всіх реалізацій.

Особливість 4. Можна показати, що енергетичний спектр $S_x^0(f)$ пов'язаний з миттєвим спектром $\dot{A}_x(f)$ центрованої випадкової функції $\overset{0}{X}(t)$ співвідношенням

$$S_x^0(f) = \lim_{T \rightarrow \infty} \frac{1}{T} M[\dot{A}_x(f) \dot{A}_x^*(f)], \quad (3.5.11)$$

де T – інтервал визначення випадкової функції $\overset{0}{X}(t)$.

Основні властивості енергетичного спектра випадкової функції наведені далі.

Властивість 1. Енергетичний спектр стаціонарної випадкової функції $\overset{0}{X}(t)$ (навіть комплексної) дійсний і невід’ємно визначений, тобто $S_x(f) \geq 0$. Це випливає з виразу (3.5.6) (при $f_1 = f_2 = f$ коваріаційна функція $K_a(f, f) \geq 0$).

Властивість 2. Оскільки коваріаційна функція $K_x(\tau)$ дійсної стаціонарної випадкової функції $\overset{0}{X}(t)$ – парна, то, як випливає з виразу (3.5.7), енергетичний спектр такої функції теж парний, тобто $S_x(f) = S_x(-f)$.

Властивість 3. Для дійсної випадкової функції $\overset{0}{X}(t)$ пару перетворень (3.5.7), (3.5.8) можна записати з застосуванням енергетичного спектра $S_x^0(f)$ випадкової функції $\overset{0}{X}(t)$ таким чином:

$$S_x^0(f) = 2 \int_0^{\infty} R_x(\tau) \cos 2\pi f \tau d\tau, \quad (3.5.12)$$

$$R_x(\tau) = 2 \int_0^{\infty} S_x^0(f) \cos 2\pi f \tau df. \quad (3.5.13)$$

Властивість 4. Враховуючи формули (3.5.12), (3.5.13), а також вираз (3.5.5), який пов’язує фізичний спектр дійсної функції з її математичним спектром, можна визначити фізичну спектральну щільність потужності $S_x^+(f)$ за допомогою пари перетворень:

$$S_x^+(f) = \begin{cases} 4 \int_0^{\infty} K_x(\tau) \cos 2\pi f \tau d\tau & \text{при } f > 0, \\ 2 \int_0^{\infty} K_x(\tau) \tau d\tau & \text{при } f = 0, \end{cases} \quad (3.5.14)$$

$$K_x(\tau) = \int_0^{\infty} S_x^+(f) \cos 2\pi f \tau df. \quad (3.5.15)$$

Зауваження. Формули (3.5.14), (3.5.15) зручно використовувати на практиці при проведенні фізичних експериментів, а еквівалентні формули (3.5.7), (3.5.8), або (3.5.12) (3.5.13) – при математичних розрахунках.

Визначення 3. Нормованою спектральною щільністю потужності стаціонарної у широкому розумінні центрованої випадкової функції $\overset{0}{X}(t)$ називається невід’ємна функція $s_x^0(f)$, що пов’язана зі спектральною щільністю потужності $S_x^0(f)$ таким чином:

$$s_x^0(f) = \frac{S_x^0(f)}{D_x^0}, \quad (3.5.16)$$

де D_x^0 – дисперсія випадкової функції.

Зауваження 1. Застосовуючи формули (3.5.7), (3.5.8), неважко отримати вирази

$$s_x(f) = \int_{-\infty}^{\infty} r_x(\tau) e^{-j2\pi f\tau} d\tau, \quad (3.5.17)$$

$$r_x(\tau) = \int_{-\infty}^{\infty} s_x(f) e^{j2\pi f\tau} df, \quad (3.5.18)$$

що пов'язують нормовану спектральну щільність потужності $s_x(f)$ центрованої випадкової функції $X(t)$ з її нормованою кореляційною функцією $r_x(\tau)$.

Зауваження 2. Нормована спектральна щільність потужності $s_x(f)$ відповідає вимогам щільності ймовірностей:

$$s_x(f) \geq 0, \quad \int_{-\infty}^{\infty} s_x(f) df = 1,$$

а нормована кореляційна функція $r_x(\tau)$ – вимогам характеристичної функції. Тому нормовану спектральну щільність потужності можна розглядати як *аналог* щільності ймовірностей, а нормовану кореляційну функцію – як *аналог* характеристичної функції.

Зауваження 3. Нормована спектральна щільність потужності, як і нормована кореляційна функція, на відміну від щільності ймовірностей і характеристичної функції не вичерпно характеризують випадкову функцію. Тому аналогія, що підкреслена у зауваженні 2, залишається лише аналогією.

Визначення 3. *Ефективною шириною спектра (ефективною смугою) Δf_e* називається величина, що визначається подібно інтервалу кореляції:

$$\Delta f_e = \frac{1}{S_0} \int_0^{\infty} S_x(f) df, \quad (3.5.19)$$

де $S_0 = \max_f S_x(f)$ – максимальне значення спектральної щільності потужності.

Властивість 5. Ефективна ширина спектра Δf_e пов'язана з інтервалом кореляції τ_k за допомогою оберненої функції:

$$\Delta f_e = \frac{1}{\tau_k}. \quad (3.5.20)$$

3.5.3. Вузькосмугові і широкосмугові випадкові функції. Білий шум

Визначення 1. Випадкова функція називається *вузькосмуговою*, якщо

$$\Delta f_e \ll f_0, \quad (3.5.21)$$

де f_0 – частота, що відповідає максимальній спектральній щільності потужності (іноді називається несучою). Якщо співвідношення (3.5.21) не виконується, то випадкова функція називається *широкосмуговою*.

Зауваження 1. Енергетичний спектр вузькосмугової випадкової функції зосереджений в районі частот $\pm f_0$ (кругових частот $\pm \omega_0$) (рис. 3.5.2).

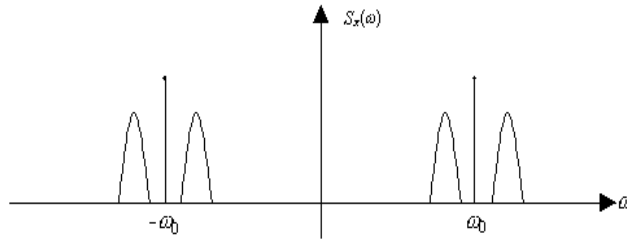


Рис. 3.5.2. Енергетичний спектр амплітудно-модульованого сигналу

Зауваження 2. Якщо енергетичний спектр вузькосмугової випадкової функції приблизно симетричний відносно несучої частоти f_0 (а це спостерігається в цілому ряді випадків, зокрема при амплітудній модуляції сигналу, див. рис. 3.5.2), то кореляційна функція $R_x(\tau)$ може бути представлена у вигляді добутку огинаючої $Q_c(\tau)$ та швидкоосцилюючої функції $\cos 2\pi f_0 \tau$:

$$R_x(\tau) = Q_c(\tau) \cos 2\pi f_0 \tau$$

(рис. 3.2.2, б).

Визначення 2. Білим шумом називається стаціонарна випадкова функція $N(t)$ з нульовим математичним сподіванням, спектральна щільність потужності якої

$$S_x(f) = \frac{N_0}{2} = \text{const}. \quad (3.5.22)$$

Основні властивості білого шуму наведені далі.

Властивість 1. Коваріаційна функція білого шуму являє собою δ -функцію, точніше

$$K_x(\tau) = \frac{N_0}{2} \delta(\tau). \quad (3.5.23)$$

Це прямо випливає з виразу (3.5.6).

Зауваження 1. Оскільки математичне сподівання білого шуму дорівнює нулеві, за допомогою того ж самого виразу (3.5.23) описується кореляційна функція $R_x(\tau)$.

Зауваження 2. Співвідношення (3.5.23) і зауваження 1 означають, що перерізи, які відстоять як завгодно мало один від одного, між собою некорельовані, тобто інтервал кореляції $\tau_k = 0$.

Властивість 2. Ефективна смуга частот білого шуму дорівнює нескінченності. Це випливає з виразів (3.5.19), (3.5.22).

Властивість 3. Дисперсія білого шуму дорівнює нескінченності. Це випливає з виразу (3.5.23).

Зауваження. Факт нескінченно великої дисперсії (потужності) білого шуму підкреслює ту обставину, що білий шум фізично не можна реалізувати. Це лише зручна математична абстракція, яка наближено описує випадкові процеси з малими змінами спектральної щільності в широкому діапазоні частот.

Визначення 3. Нестационарним білим шумом називається нестационарна випадкова функція $Y(t) = X(t)N(t)$, де $X(t)$ – деяка випадкова чи детермінована функція, а $N(t)$ – білий шум.

Зауваження 1. Кореляційна функція такого нестационарного білого шуму

$$R_y(t_1, t_2) = M[|X(t_1)|^2] \frac{N_0}{2} \delta(t_2 - t_1).$$

Зауваження 2. Інколи нестационарний білий шум визначають за допомогою виразу $Y(t) = X(t) + N(t)$ або $Y(t) = X_1(t) + X_2(t)N(t)$, де $X(t)$, $X_1(t)$ та $X_2(t)$ – деякі випадкові чи детерміновані функції.

3.5.4. Взаємна спектральна щільність потужності

Метод спектрального опису однієї випадкової функції, що розглянутий вище, узагальнюється на випадок двох стаціонарно пов'язаних випадкових функцій.

Визначення 1. Взаємною спектральною щільністю (потужності) двох стаціонарно пов'язаних випадкових функцій $X(t)$ і $Y(t)$ називається невідмінна функція $\dot{S}_{xy}(f)$, що визначається через коваріаційну функцію $K_{xy}(\tau)$ і зв'язана з нею парою перетворень Фур'є:

$$\dot{S}_{xy}(f) = \int_{-\infty}^{\infty} K_{xy}(\tau) e^{-j2\pi f\tau} d\tau, \quad (3.5.24)$$

$$K_{xy}(\tau) = \int_{-\infty}^{\infty} \dot{S}_{xy}(f) e^{j2\pi f\tau} df. \quad (3.5.25)$$

Зауваження 1. На відміну від спектральної щільності потужності $S_x(f)$ однієї випадкової функції $X(t)$, взаємна спектральна щільність $\dot{S}_{xy}(f)$ у загальному випадку не є дійсною функцією.

Зауваження 2. На відміну від $S_x(f)$, взаємна спектральна щільність $\dot{S}_{xy}(f)$ не є парною. При цьому

$$\dot{S}_{xy}(f) = S_{yx}^*(f). \quad (3.5.26)$$

Ця рівність доводиться із застосуванням виразу (3.5.24).

Зауваження 3. Можна довести, що

$$|\dot{S}_{xy}(f)|^2 \leq S_x(f)S_y(f). \quad (3.5.27)$$

Визначення 2. Функцією частотної когерентності двох стаціонарно пов'язаних випадкових функцій $X(t)$ і $Y(t)$ називається невідмінна функція

$$\gamma_{xy}^2(f) = \frac{|\dot{S}_{xy}(f)|^2}{S_x(f)S_y(f)}. \quad (3.5.28)$$

Зауваження 1. З виразу (3.5.27) випливає, що функція частотної когерентності $\gamma_{xy}^2(f)$ може знаходитися у межах $[0, 1]$. Якщо випадкові функції $X(t)$, $Y(t)$ некорельовані, то для усіх $f \neq 0$ $\gamma_{xy}^2(f) = 0$, а якщо вони лінійно пов'язані, то $\gamma_{xy}^2(f) = 1$.

Зауваження 2. Функція частотної когерентності $\gamma_{xy}^2(f)$ є аналогом квадрата нормованої взаємної кореляційної функції $r_{xy}^2(\tau)$. Як і $r_{xy}^2(\tau)$ вона характеризує зв'язок між випадковими функціями $X(t)$ і $Y(t)$. Однак, на відміну від $r_{xy}^2(\tau)$, вона характеризує не тільки лінійний, а й нелінійний зв'язок.

Зауваження 3. Якщо випадкова функція $Y(t)$ залежить не тільки від випадкової функції $X(t)$, а і від інших детермінованих чи випадкових функцій, функція частотної когерентності $\gamma_{xy}^2(f)$ залежить й від цих інших функцій.

Коли $Y(t)$ залежить від $X(t)$ лінійно і при цьому існує залежність $Y(t)$ від інших функцій, функція $\eta_{xy}(f) = 1 - \gamma_{xy}^2(f)$ є мірою зв'язку між функцією $Y(t)$ та цих інших функцій.

3.6. Перетворення стаціонарних випадкових функцій лінійними операторами

3.6.1. Математичний опис лінійних інерційних операторів

Визначення 1. Лінійним інерційним інтегральним оператором називається оператор, який описується таким співвідношенням:

$$y(t) = \int_{-\infty}^{\infty} x(t_1)h(t, t_1)dt_1, \quad (3.6.1)$$

де $x(t)$, $y(t)$ – відповідно вхідний і вихідний процеси, $h(t, t_1)$ – вагова функція, що називається імпульсною перехідною характеристикою лінійного оператора.

Зауваження 1. Імпульсна перехідна характеристика $h(t, t_1)$ лінійного оператора являє собою реакцію (відгук) системи на δ -вплив:

$$h(t, t_2) = \int_{-\infty}^{\infty} \delta(t_2 - t_1)h(t, t_1)dt_1. \quad (3.6.2)$$

Ця властивість імпульсної перехідної характеристики іноді береться як її визначення.

Зауваження 2. Будь-який лінійний оператор може бути представлений у вигляді інтегрального оператора (3.6.1).

Зауваження 3. Будь-який лінійний оператор має фільтруючі властивості. Тому його можна розглядати як лінійний фільтр. Вичерпною характеристикою такого фільтра є імпульсна перехідна характеристика $h(t, t_1)$.

Визначення 2. Оператор називається *стаціонарним (інваріантним до зсуву)*, якщо зсув вхідного процесу на деяку фіксовану величину приводить до такого ж зсуву вихідного процесу. В протилежному випадку оператор називається *нестаціонарним*.

Зауваження 1. Усі безінерційні оператори є стаціонарними. Серед лінійних інерційних інтегральних операторів зустрічаються як стаціонарні, так і нестаціонарні.

Зауваження 2. Для лінійного інерційного інтегрального стаціонарного оператора імпульсна перехідна характеристика $h(t, t_1)$ залежить лише від різниці аргументів t, t_1 , тобто $h(t, t_1) = h(\tau)$, де $\tau = t - t_1$. В цьому випадку вираз (3.6.1) має вигляд згортки:

$$y(t) = \int_{-\infty}^{\infty} x(t_1)h(t - t_1)dt_1 = \int_{-\infty}^{\infty} x(t - \tau)h(\tau)d\tau. \quad (3.6.3)$$

Визначення 3. Лінійний оператор називається *дійсним*, якщо його імпульсна перехідна характеристика дійсна.

Визначення 4. Стаціонарний лінійний оператор називається *сталим*, якщо при будь-якому обмеженому впливі його відгук обмежений.

Зауваження. Оскільки

$$|y(t)| = \left| \int_{-\infty}^{\infty} x(t - \tau)h(\tau)d\tau \right| \leq \int_{-\infty}^{\infty} |x(t - \tau)| |h(\tau)| d\tau, \quad (3.6.4)$$

імпульсна перехідна характеристика $h(\tau)$ сталого оператора повинна бути

абсолютно інтегрованою:

$$\int_{-\infty}^{\infty} |h(\tau)| d\tau < \infty.$$

Визначення 5. *Лінійним фільтром, що фізично реалізується, називається лінійний оператор, імпульсна перехідна характеристика якого $h(t)$ при $t < 0$ тотожна нулю. Лінійні фільтри, які не відносяться до класу фільтрів, що фізично реалізуються, називаються фільтрами, що фізично не реалізуються.*

Зауваження 1. Фільтр, що фізично реалізується, – це такий фільтр, значення відгуку $y(t)$ якого для кожного поточного значення $t = t_0$ визначається лише поточним $x(t_0)$ і попередніми значеннями вхідного впливу $x(t)$ (тобто значеннями, які відповідають нерівності $t \leq t_0$).

Зауваження 2. Для фільтра, що фізично реалізується, формули (3.6.3) в загальному випадку мають вигляд

$$y(t) = \int_{-\infty}^t x(t_1)h(t-t_1)dt_1 = \int_0^{\infty} x(t-\tau)h(\tau)d\tau. \quad (3.6.5)$$

Якщо вхідний вплив $x(t)$ починається у нульовий момент часу t , то

$$y(t) = \int_0^t x(t_1)h(t-t_1)dt_1 = \int_0^t x(t-\tau)h(\tau)d\tau. \quad (3.6.6)$$

Визначення 6. *Передаточною характеристикою (функцією) лінійного оператора $\dot{K}(\omega) = K(\omega) e^{-j\varphi(\omega)}$ (частотною характеристикою оператора) називається спектр його імпульсної перехідної характеристики $h(t)$:*

$$\dot{K}(\omega) = \int_{-\infty}^{\infty} h(\tau)e^{-j\omega\tau} d\tau. \quad (3.6.7)$$

Визначення 7. Модуль $K(\omega)$ передаточної функції $\dot{K}(\omega)$ називається амплітудно-частотною характеристикою оператора, а аргумент $\varphi(\omega)$ – його фазо-частотною характеристикою.

Зауваження 1. Імпульсна перехідна характеристика лінійного оператора пов'язана з передаточною функцією оберненим перетворенням Фур'є:

$$h(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{K}(\omega)e^{j\omega\tau} d\omega. \quad (3.6.8)$$

Зауваження 2. Миттєві спектри $\dot{A}_x(\omega)$ і $\dot{A}_y(\omega)$ відгуку і впливу пов'язані співвідношенням

$$\dot{A}_y(\omega) = \dot{K}(\omega) \dot{A}_x(\omega). \quad (3.6.9)$$

Зауваження 3. Спектральні щільності потужності $S_y(\omega)$ і $S_x(\omega)$ відгуку і впливу пов'язані з передаточною функцією $\dot{K}(\omega)$ співвідношенням

$$S_y(\omega) = |\dot{K}(\omega)|^2 S_x(\omega). \quad (3.6.10)$$

Положення 1. Якщо лінійний оператор з передаточною функцією $\dot{K}(\omega)$ та імпульсною перехідною характеристикою $h(t)$ складається з двох *попередовно* діючих лінійних операторів з передаточними функціями $\dot{K}_1(\omega)$, $\dot{K}_2(\omega)$ та імпульсними перехідними характеристиками $h_1(t)$, $h_2(t)$, тоді

$$\dot{K}(\omega) = \dot{K}_1(\omega) \dot{K}_2(\omega), \quad (3.6.11)$$

$$h(t) = \int_{-\infty}^{\infty} h_1(t_1) h_2(t-t_1) dt_1.$$

Положення 2. Якщо лінійний оператор з передаточною функцією $\dot{K}(\omega)$ складається з двох *паралельно* діючих лінійних операторів з передаточними функціями $\dot{K}_1(\omega)$, $\dot{K}_2(\omega)$, причому оператор з передаточною функцією $\dot{K}_2(\omega)$ забезпечує *обернений зв'язок*,

$$\dot{K}(\omega) = \frac{\dot{K}_1(\omega)}{1 + \dot{K}_1(\omega) \dot{K}_2(\omega)}. \quad (3.6.12)$$

Положення 3. Для опису лінійних операторів іноді замість імпульсної перехідної характеристики $h(\tau)$ застосовують перехідну характеристику $g(\tau)$.

Визначення 5. *Перехідною характеристикою* $g(\tau)$ лінійного оператора називається функція, що описує відгук системи на одиничний вплив.

Зауваження. Перехідна характеристика $g(\tau)$ пов'язана з імпульсною перехідною характеристикою $h(\tau)$:

$$g(t) = \int_0^t h(\tau) d\tau, \quad h(t) = \frac{dg(t)}{dt}. \quad (3.6.13)$$

Положення 4. Для опису стаціонарних лінійних операторів іноді замість передаточної характеристики $\dot{K}(\omega)$ застосовують *передаточну функцію* $\dot{H}(p)$, яка визначається як перетворення Лапласа імпульсної перехідної характеристики:

$$\dot{H}(p) = \int_0^{\infty} h(\tau) e^{-p\tau} d\tau, \quad p = \alpha + j\omega.$$

Зауваження 1. Для лінійного оператора, що фізично реалізується, передаточна характеристика $\dot{K}(\omega)$ є окремим випадком передаточної функції $\dot{H}(p)$ (при $p = j\omega$).

Зауваження 2. Оператор сталий, якщо $\dot{H}(p)$ не має полюсів при $\alpha \geq 0$. У противному разі оператор несталий.

Положення 5. Лінійні оператори можуть бути описані також за допомогою лінійних диференціальних рівнянь типу

$$a_M \frac{d^M y(t)}{dt^M} + \dots + a_0 y(t) = b_N \frac{d^N x(t)}{dt^N} + \dots + b_0 x(t),$$

де a_m , b_n ($m = \overline{1, M}$, $n = \overline{1, N}$) – сталі коефіцієнти, а також початкових умов, що накладаються на вихідний процес: $y(0)$, $y'(0)$, ..., $y^{(M-1)}(0)$.

Зауваження до підрозділу 3.6.1. Для опису роботи лінійного пристрою можна використовувати будь-який з розглянутих вище способів: операторний,

за допомогою імпульсної перехідної характеристики, передаточної характеристики, передаточної функції, перехідної характеристики або диференційного рівняння.

Всі ці способи еквівалентні. У подальшому ми будемо переважно використовувати підхід, що заснований на імпульсній перехідній характеристиці $h(t)$ та передаточної характеристиці $\dot{K}(\omega)$.

Результати розрахунку амплітудно-частотної характеристики $\dot{K}(\omega)$, імпульсної перехідної характеристики $h(t)$ і ефективної смуги пропускання $\Delta\omega_e$ різноманітних фільтрів наведено в табл. 3.6.1.

Таблиця 3.6.1

Характеристики фільтрів

Назва фільтра	Передаточна функція $\dot{K}(\omega)$	Імпульсна перехідна характеристика $h(t)$	Ефективна смуга пропускання
Ідеальний фільтр низьких частот	K_0 при $ \omega \leq \Delta\Omega$ 0 при $ \omega > \Delta\Omega$	$K_0 \frac{\Delta\Omega}{\pi} \times$ $\times \frac{\sin(\Delta\Omega t)}{\Delta\Omega t}$	$\Delta\Omega$
Ідеальний смуговий фільтр	K_0 при $\left\{ \begin{array}{l} \omega - \frac{\Delta\omega}{2} \leq \omega \leq \omega_0 + \frac{\Delta\omega}{2} \\ -\omega_0 - \frac{\Delta\omega}{2} \leq \omega \leq -\omega_0 + \frac{\Delta\omega}{2} \end{array} \right.$ 0 – в інших випадках	$K_0 \frac{\Delta\omega}{\pi} \times$ $\times \frac{\sin(\frac{\Delta\omega t}{2})}{\frac{\Delta\omega t}{2}} \cos \omega_0 t$	$\Delta\omega$
RC-ланцюг	$\frac{\alpha}{\alpha + j\omega}$	$1(t)\alpha e^{-\alpha t}$	$\frac{\alpha\pi}{2}$
Окремий резонансний контур	$\frac{\alpha}{\alpha + j(\omega - \omega_0)}$	$1(t)\alpha e^{-\alpha t} \times$ $\times \cos \omega_0 t$	$\alpha\pi$
Колоколоподібний фільтр низьких частот	$\frac{1}{2} e^{-\frac{\omega^2}{2\beta^2}}$	$\frac{1(t)\beta}{\sqrt{2\pi}} e^{-\frac{t^2\beta^2}{2}}$	$2\sqrt{2\pi}\beta$
Колоколоподібний смуговий фільтр	$\frac{1}{2} e^{-\frac{(\omega-\omega_0)^2}{2\beta^2}}$	$\frac{1(t)\beta}{\sqrt{2\pi}} e^{-\frac{t^2\beta^2}{2}} \times$ $\times \cos \omega_0 t$	$4\sqrt{2\pi}\beta$

Примітки: 1. $1(t) = \begin{cases} 1 & \text{при } t \geq 0; \\ 0 & \text{при } t < 0. \end{cases}$

2. Колоколоподібний фільтр низьких частот можна розглядати як фільтр, складений з нескінченного числа RC -ланцюгів, включених послідовно, а колоколоподібний смуговий фільтр – як фільтр, що складається з нескінченного числа окремих резонансних контурів.

3.6.2. Перетворення випадкових функцій лінійними інерційними операторами

Моментні функції і спектральні щільності потужності відгуку $Y(t)$ і впливу $X(t)$ пов'язані між собою. Формули, які описують ці зв'язки для нестационарних і стаціонарних випадкових функцій, наведені відповідно в табл. 3.6.2 і 3.6.3 (в обох випадках вважається, що вплив починається в нульовий момент часу t).

Таблиця 3.6.2

Перетворення нестационарних випадкових функцій для стаціонарного фільтра, який фізично реалізується

Моментні функції	
	$m_y(t) = \int_0^t m_x(t - \tau)h(\tau)d\tau$
	$K_y(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} K_x(t_1 - \tau_1, t_2 - \tau_2)h(\tau_1)h(\tau_2)d\tau_1d\tau_2$
	$K_{xy}(t_1, t_2) = \int_0^{t_2} K_x(t_1, t_2 - \tau)h(\tau)d\tau$
	$R_y(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} R_x(t_1 - \tau_1, t_2 - \tau_2)h(\tau_1)h(\tau_2)d\tau_1d\tau_2$
	$R_{xy}(t_1, t_2) = \int_0^{t_2} R_x(t_1, t_2 - \tau)h(\tau)d\tau$

Таблиця 3.6.3

Перетворення стаціонарних випадкових функцій для стаціонарного фільтра, що фізично реалізується

№ з/п	Моментні функції і спектральні щільності
1	$m_y(t) = m_x \int_0^t h(\tau)d\tau$
2	$R_y(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} R_x(t_2 - \tau_2 - (t_1 - \tau_1))h(\tau_1)h(\tau_2)d\tau_1d\tau_2$
3	$R_{xy}(t_1, t_2) = \int_0^{t_2} R_x(t_2 - \tau - t_1)h(\tau)d\tau$
4	$S_y(\omega) = K^2(\omega)S_x(\omega) \quad (t_0 > T)$

Зауваження 1. Із формул табл. 3.6.3 випливає, що відгук на вплив стаціонарної випадкової функції, в загальному випадку, являє собою нестационарну випадкову функцію. Однак при цьому треба мати на увазі, що всі

лінійні фільтри, які фізично реалізуються, на практиці мають імпульсну перехідну характеристику $h(t)$ кінцевої тривалості T (вона наближається до нуля при необмеженому збільшенні t). Тому, коли відгук на вплив розглядається в момент t_0 , що відстає від початку більш ніж за тривалість T , формули 1, 2, 3 табл. 3.6.3 набувають спрощеного вигляду:

$$m_y = m_x \int_0^T h(\tau) d\tau,$$

$$R_y(\tau) = \int_0^T \int_0^T R_x(\tau - (\tau_2 - \tau_1)) h(\tau_1) h(\tau_2) d\tau_1 d\tau_2,$$

$$R_{xy}(\tau) = \int_0^T R_x(\tau - \tau_1) h(\tau_1) d\tau_1.$$

Звідси випливає, що при $t_0 > T$ відгук являє собою стаціонарну випадкову функцію, а вплив і відгук – стаціонарно пов'язані.

Зауваження 2. Формула 4 табл. 3.6.3 відповідає випадку $t_0 > T$, коли відгук можна вважати стаціонарною випадковою функцією.

3.6.3. Перетворення білого шуму лінійними операторами

Враховуючи співвідношення (3.5.23), неважко одержати формули, які описують реакцію фільтра на вплив білого шуму:

$$R_y(\tau) = \frac{N_0}{2} \int_0^\infty h(\tau_1 + \tau) h(\tau_1) d\tau_1, \quad S_y(\omega) = \frac{N_0}{2} K^2(\omega),$$

$$R_{xy}(\tau) = \frac{N_0}{2} h(\tau).$$

Кореляційні функції $R_y(\tau)$, дисперсії D_y і спектральні щільності потужності $S_y(\omega)$ відгуку різноманітних фільтрів на вплив білого шуму наведені в табл. 3.6.4.

Таблиця 3.6.4

Кореляційні функції $R_y(\tau)$, дисперсії D_y і спектральні щільності потужності $S_y(\omega)$ відгуку різноманітних фільтрів на вплив білого шуму

Назва фільтра	Кореляційна функція $R_y(\tau)$	Дисперсія D_y	Спектральна щільність потужності $S_y(\omega)$
Ідеальний фільтр низьких частот	$D_y \frac{\sin(\Delta\Omega\tau)}{\Delta\Omega\tau}$	$\frac{K_0^2 N_0 \Delta\Omega}{2\pi}$	$\frac{K_0^2 N_0}{2}$ при $ \omega \leq \Delta\Omega$ 0 при $ \omega > \Delta\Omega$
Ідеальний смуговий фільтр	$D_y \frac{\sin \frac{\Delta\omega\tau}{2}}{\frac{\Delta\omega\tau}{2}} \cos \omega_0 \tau$	$\frac{K_0^2 N_0 \Delta\omega}{2\pi}$	$\frac{K_0^2 N_0}{2}$ при $\left\{ \begin{array}{l} \omega_0 - \frac{\Delta\omega}{2} \leq \omega \leq \omega_0 + \frac{\Delta\omega}{2} \\ -\omega_0 - \frac{\Delta\omega}{2} \leq \omega \leq -\omega_0 + \frac{\Delta\omega}{2} \end{array} \right.$ 0 в інших випадках
RC-ланцюг	$D_y e^{-\alpha \tau }$	$\frac{\alpha N_0}{4}$	$\frac{N_0 \alpha^2}{2(\alpha^2 + \omega^2)}$

Окремий резонансний контур	$D_y e^{-\alpha \tau } \cos \omega_0 \tau$	$\frac{\alpha N_0}{2}$	$\frac{N_0 \alpha^2}{2(\alpha^2 + (\omega_0 - \omega)^2)}$
Колоколоподібний низькочастотний фільтр	$D_y e^{-\frac{\tau^2 \beta^2}{2}}$	$\frac{\beta N_0}{4\sqrt{2\pi}}$	$\frac{N_0}{4} e^{-\frac{\omega^2}{2\beta^2}}$
Колоколоподібний смуговий фільтр	$D_y e^{-\frac{\tau^2 \beta^2}{2}} \cos \omega_0 \tau$	$\frac{\beta N_0}{2\sqrt{2\pi}}$	$\frac{N_0}{4} e^{-\frac{(\omega - \omega_0)^2}{2\beta^2}}$

Примітка. Зростання в два рази дисперсії окремого резонансного контуру у порівнянні з RC-ланцюговим пов'язане з розширенням в два рази смуги пропускання фільтра. З тим же пов'язане і зростання дисперсії колоколоподібного низькочастотного фільтра у порівнянні з аналогічним смуговим.

Нарешті зазначимо, що у додатку 4 наведені деякі корисні відомості, які стосуються огинаючої та спектральних перетворень випадкових функцій, що пов'язані з нею.

Відомості про книги з тематики розділу 3 наведено в табл. 3.6.5.

Таблиця 3.6.5

Література до розділу 3

Характер книг	Номери книг у списку літератури
Довідники	3, 24, 40, 47, 49
Книги навчального плану для інженерів	16, 32, 42, 73, 75, 84, 100
Книги з математичним ухилом	5, 10, 15, 20, 23, 61, 88, 89, 105
Книги прикладного характеру	27, 33, 35, 37, 38, 54, 55, 56, 66, 72, 74, 82, 85, 86, 92, 94, 97, 104

4. Марковські випадкові процеси

4.1. Початкові уявлення про марковські процеси і їх класифікація

Визначення 1. Випадковий процес $X(t)$, $t \in T$ називається *марковським*, якщо для будь-яких значень $t_1 < t_2 < \dots < t_N$, що належать T , одновимірною умовною функцією розподілу

$$F_1(x_N; t_N / x_1, \dots, x_{N-1}; t_1, \dots, t_{N-1}) = F_1(x_N; t_N / x_{N-1}; t_{N-1}). \quad (4.1.1)$$

Зауваження 1. З виразу (4.1.1) випливає, що якщо стан випадкового процесу $X(t)$ в поточний момент часу t_N залежить лише від стану в момент часу t_{N-1} , процес є марковським.

Зауваження 2. З виразу (4.1.1) випливає, що для марковського випадкового процесу $X(t)$ N -вимірною функцією розподілу:

$$F_N(x_1, \dots, x_N; t_1, \dots, t_N) = F_1(x_1; t_1) \prod_{n=1}^{N-1} F_1(x_{n+1}; t_{n+1} / x_n; t_n), \quad (4.1.2)$$

тобто N -вимірною функцією розподілу може бути знайдена, якщо відома одновимірною функцією розподілу при $t = t_1$ і послідовність умовних функцій розподілу для моментів часу $t_n > t_1$, ($n = \overline{2, N}$).

Поширеними критеріями класифікації марковських випадкових процесів є тип області T визначення аргументу і тип простору станів S .

Звичайно розрізняють 4 основних класи марковських випадкових процесів: ланцюг Маркова; марковські послідовності; дискретні марковські процеси; неперервні марковські процеси.

Початкові уявлення про класи цієї класифікації дає табл. 4.1.1.

Таблиця 4.1.1

Класифікація марковських процесів

Область T	Область S	
	Дискретна	Неперервна
Дискретна	Ланцюг Маркова	Марковська послідовність
Неперервна	Дискретний марковський процес	Неперервний марковський процес

4.2. Ланцюг Маркова

Розглянемо випадковий процес $X(t)$, у якого область визначення T – дискретна множина точок $t_0 < t_1 < \dots$, а простір станів являє собою дискретну множину точок $\{\theta_l, l = \overline{1, L}\} = S$ (рис. 4.2.1).

В моменти часу t_0, t_1, \dots можливі зміни стану (тобто перехід з одного стану в інший). Такий випадковий процес можна розглядати як послідовність дискретних випадкових величин $X_N = X(t_N)$, $N = 0, 1, \dots$, яка називається *дискретною випадковою послідовністю*.

Визначення 1. *Простим ланцюгом Маркова* називається дискретна випадкова послідовність, для якої одновимірною умовною функцією розподілу

$$F_1(x_N / x_0, \dots, x_{N-1}) = F_1(x_N / x_{N-1}), \quad N = 1, 2, \dots \quad (4.2.1)$$

Зауваження. З виразу (4.2.1) випливає, що ймовірність $P\{x_N / x_0, \dots, x_{N-1}\}$

того, що X_N прийме значення x_N за умови, що випадкові величини X_0, X_1, \dots, X_{N-1} приймали значення x_0, x_1, \dots, x_{N-1} , визначається рівністю

$$P\{x_N / x_0, \dots, x_{N-1}\} = P\{x_N / x_{N-1}\}.$$

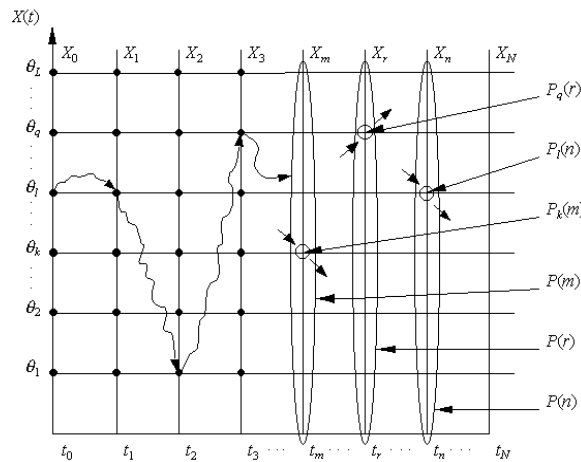


Рис. 4.2.1. Умовне зображення ланцюга Маркова

Визначення 2. Складним ланцюгом Маркова порядку M називається дискретна випадкова послідовність, для якої одновимірна умовна функція розподілу

$$F_1(x_N / x_0, \dots, x_{N-1}) = F_1(x_N / x_{N-M}, \dots, x_{N-1}), \quad N = M, M+1, \dots$$

Визначення 3. Простим ланцюгом Маркова для M -вимірного вектора \vec{X}_N називається дискретна векторна випадкова послідовність $\vec{X}_0, \vec{X}_1, \dots$, для якої одновимірна умовна функція розподілу

$$F_M(\vec{x}_N / \vec{x}_0, \dots, \vec{x}_{N-1}) = F_M(\vec{x}_N / \vec{x}_{N-1}).$$

Теорема 1. Складний ланцюг Маркова порядку M може бути зведений до простого ланцюга Маркова для M -вимірного вектора $\vec{X}_N = (X_{N-M+1}, \dots, X_N)$.

Ця теорема прямо випливає з визначень 2 і 3.

Зауваження. Можливість зведення складного ланцюга Маркова до простого ланцюга Маркова для вектора дозволяє обмежитися розглядом лише простих марковських ланцюгів.

Теорема 2. Простий марковський ланцюг вичерпно описується N -вимірною сумісною ймовірністю

$$P\{X_0 = \theta_{l_0}, \dots, X_N = \theta_{l_N}\} = P\{X_0 = \theta_{l_0}\} \prod_{n=1}^N P\{X_n = \theta_{l_n} / X_{n-1} = \theta_{l_{n-1}}\}.$$

Зауваження 1. Доведення теореми базується на визначенні простого марковського ланцюга.

Визначення 4. Ймовірністю переходу з стану θ_k в момент часу t_m в стан θ_l в момент часу t_n за $(n - m)$ кроків називається умовна ймовірність

$$\pi_{kl}(m, n) = P\{X_n = \theta_l / X_m = \theta_k\},$$

де $\theta_k, \theta_l \in S$.

Упорядкована сукупність ймовірностей переходу з стану θ_k в стан θ_l ($\theta_k, \theta_l \in S$; $k, l = \overline{1, L}$) для заданих моментів часу t_m, t_n утворює квадратну матрицю $\Pi(m, n)$ розміром $L \times L$.

Визначення 5. Матриця називається *стохастичною*, якщо:

- 1) всі її елементи невід'ємні;
- 2) сума елементів будь-якого її рядка дорівнює одиниці.

Теорема 3. Матриця ймовірностей переходу є стохастичною.

Теорема 4. Нехай для простого ланцюга Маркова відомі безумовні ймовірності $P_k(m)$ перебування в станах θ_k на m -му кроці ($k = \overline{1, L}$) і ймовірності переходу $\pi_{kl}(m, n)$ із цих станів в деякий стан θ_l на n -му кроці. Тоді безумовна ймовірність $P_l(n)$ перебування в стані θ_l на n -му кроці визначається співвідношенням

$$P_l(n) = \sum_{k=1}^L P_k(m) \pi_{kl}(m, n), \quad (l = \overline{1, L}; \quad 0 \leq m \leq n). \quad (4.2.2)$$

Матрична форма рівняння (4.2.2) має вигляд

$$P(n) = P(m) \Pi(m, n),$$

де $P(n)$ – вектор-рядок розміром L з елементами $P_l(n)$, $l = \overline{1, L}$, який описує ймовірності перебування у всіх можливих станах на n -му кроці.

Теорема 5. Нехай ймовірність переходу із стану k на m -му кроці в стан l на n -му кроці визначається ймовірністю $\pi_{kl}(m, n)$, ймовірність переходу з стану k на m -му кроці в проміжний стан q на r -му кроці ($0 \leq m < r < n$) – ймовірністю переходу $\pi_{kq}(m, r)$, а ймовірність переходу із проміжного стану q на r -му кроці в стан l на n -му кроці – ймовірністю переходу $\pi_{ql}(r, n)$. Тоді

$$\pi_{kl}(m, n) = \sum_{q=1}^L \pi_{kq}(m, r) \pi_{ql}(r, n), \quad (k, l = \overline{1, L}; \quad 0 \leq m \leq n),$$

або у матричній формі

$$\Pi(m, n) = \Pi(m, r) \Pi(r, n). \quad (4.2.3)$$

Рівняння (4.2.3) називається *рівнянням Маркова для ймовірностей переходу*.

Зауваження. Рівняння (4.2.3) є необхідною умовою того, що ланцюг марковський, однак ця умова не є достатньою (тобто ще існують випадкові послідовності, які не є ланцюгами Маркова, однак для них справедливе співвідношення (4.2.3)).

Теорема 6. Нехай відомі ймовірності перебування у всіх можливих станах в початковий момент часу (тобто відомий вектор $P(0)$) і матриці ймовірностей однокрокового переходу $\Pi(n-1, n)$ ($n = \overline{1, N}$). Тоді вектор $P(N)$, що описує ймовірності перебування у всіх можливих станах в момент часу t_N , може бути представлений таким чином:

$$P(N) = P(0) \prod_{n=1}^N \Pi(n-1, n).$$

Визначення 5. Простий ланцюг Маркова називається *однорідним*, якщо всі матриці ймовірностей переходу $\Pi(m, n)$ не залежать від моментів часу t_m, t_n , для яких вони визначені, а залежать тільки від різниці $(n-m)$, тобто

$$\Pi(m, n) = \Pi(n-m). \quad (4.2.4)$$

Ланцюги Маркова, які не задовольняють співвідношення (4.2.4), називаються *неоднорідними*.

Зауваження до теореми 6. Для однорідного ланцюга Маркова

$$P(N) = P(0) \Pi^N, \quad (4.2.5)$$

де Π – матриця однокрокових переходів.

Визначення 6. Вектором фінальних ймовірностей називається вектор

$$P = \lim_{N \rightarrow \infty} P(N).$$

Зауваження. Не всі ланцюги Маркова мають фінальні ймовірності.

Визначення 7. Марковські ланцюги називаються *ергодичними*, якщо:

- 1) існує вектор фінальної ймовірності;
- 2) вектор фінальної ймовірності не залежить від стану, в якому знаходився ланцюг в початковий момент часу (при t_0).

Зауваження. Поняття ергодичності для ланцюга Маркова відрізняється від аналогічного поняття для випадкового процесу.

Визначення 8. *Стаціонарним або рівноважним ланцюгом Маркова* називається дискретна однорідна випадкова послідовність, для якої вектор ймовірностей $P(n)$ перебування в різноманітних станах в момент часу t_n не залежить від моменту часу t_n , тобто $P(n) = P$ для всіх n .

Теорема 8. Однорідний ланцюг Маркова стаціонарний тоді і тільки тоді, коли

$$P = P\Pi. \quad (4.2.6)$$

Зауваження 1. Теорема випливає з співвідношення (4.2.5).

Зауваження 2. З співвідношення (4.2.6) виходить, що у випадку стаціонарного ланцюга Маркова вектор P ймовірностей перебування в різноманітних станах є власним вектором матриці ймовірностей переходу Π .

Зауваження 3. Стаціонарний однорідний ланцюг Маркова ергодичний.

Розглянуті типи ланцюгів Маркова і їхні основні властивості зведені в табл. 4.2.1.

Таблиця 4.2.1

Типи ланцюгів Маркова

Тип ланцюга	Математичне визначення	Основні властивості
Однорідний	$\Pi(m, n) = \Pi(n - m)$	Для однорідного ланцюга $P(N) = P(0)\Pi^N$
Стаціонарний	$P(n) = P$ для всіх n	Для однорідного і стаціонарного ланцюга $P = P\Pi$
Ергодичний	1) $P = \lim_{N \rightarrow \infty} P(N)$ 2) P не залежить від початкового стану	Стаціонарний однорідний ланцюг є ергодичним

Приклади. Характерні приклади залежності ймовірності перебування в l -му стані для різноманітних типів ланцюгів Маркова наведені на рис. 4.2.2.

Наявність певних типів станів часто докорінним чином змінює властивості ланцюгів Маркова. Розглянемо деякі з таких станів.

Визначення 9. Стан θ_l називається *незворотним*, якщо ймовірність переходу з цього стану в будь-який інший стан більша за нуль, а ймовірність повернення з будь-якого іншого стану в θ_l -й стан менша за одиницю.

Стан θ_l називається *зворотним*, якщо ймовірність переходу з цього стану в

будь-який інший стан більша за нуль, а ймовірність повернення з будь-якого іншого стану – дорівнює одиниці.

Визначення 10. Зворотний стан θ_l називається *зворотним нульовим*, якщо математичне сподівання кількості кроків, що необхідно для повернення у стан θ_l , – нескінченна величина, і *зворотним додатним*, якщо математичне сподівання є скінченною величиною.

Визначення 11. Зворотні стани θ_l, θ_k називаються *сполученими*, коли ймовірності переходу з стану θ_l в стан θ_k і навпаки, з стану θ_k в θ_l , більші за нуль хоча б для двох моментів часу.

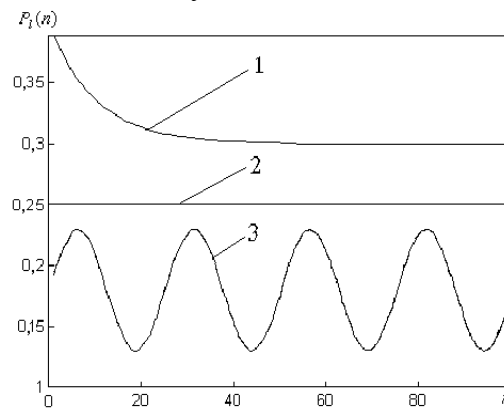


Рис. 4.2.2. Залежність ймовірності перебування в l -му стані в різні моменти часу для різних типів ланцюгів Маркова: 1 – ергодичного; 2 – стаціонарного; 3 – неергодичного.

Визначення 12. Ланцюг Маркова називається *незвідним*, якщо множина станів цього ланцюга утворює клас сполучених станів.

Визначення 13. Нехай r – найбільший спільний дільник усіх номерів кроків n_1, n_2, \dots , для яких ймовірність повернення у сполучений стан θ_l є додатна величина. Тоді стан θ_l називається *періодичним* з періодом r , якщо $r > 1$, та *аперіодичним*, якщо $r = 1$.

Визначення 14. Стан θ_l називається *поглинаючим* станом, якщо ймовірність переходу у цей стан більша за нуль, а ймовірність виходу з нього дорівнює нулю для всіх моментів часу. Ланцюги Маркова з поглинаючими станами називаються *поглинаючими*.

Визначення 15. Стан називається *відзеркалюючим (відображаючим)*, якщо ймовірність переходу в нього дорівнює нулю.

Положення. *Марковські послідовності* (випадкові процеси Маркова, задані в дискретні моменти часу при неперервному розподілі значень процесу) наближено можна розглядати як ланцюги Маркова з нескінченним числом станів ($L \rightarrow \infty$).

4.3. Неперервний марковський процес

Визначення 1. Нехай $X_0 = X(t_0), \dots, X_N = X(t_N)$ значення неперервного процесу $X(t)$ у довільні моменти часу $t_0 < t_1 < \dots < t_N$. Процес $X(t)$ називається *марковським*, якщо для будь-якого моменту часу t_N одновимірною умовною щільністю ймовірностей

$$f_1(x_N; t_N / x_0, \dots, x_{N-1}; t_0, \dots, t_{N-1}) = f_1(x_N; t_N / x_{N-1}; t_{N-1}). \quad (4.3.1)$$

Зауваження 1. З співвідношення (4.3.1) випливає, що безумовна щільність ймовірностей марковського випадкового процесу $X(t)$ може бути представлена таким чином:

$$f_N(x_0, \dots, x_N; t_0, \dots, t_N) = f_1(x_0; t_0) \prod_{n=1}^N \Pi(x_n; t_n / x_{n-1}; t_{n-1}), \quad (4.3.2)$$

де $\Pi(x_n; t_n / x_{n-1}; t_{n-1}) = f_1(x_n; t_n / x_{n-1}; t_{n-1})$ – щільність ймовірностей переходу з стану x_{n-1} в момент часу t_{n-1} в стан x_n в момент часу t_n .

Зауваження 2. З співвідношення (4.3.1) випливає, що коли значення випадкового процесу в будь-які незбіжні моменти часу незалежні, то процес – марковський.

Властивості щільності ймовірностей переходу наведені в табл. 4.3.1.

Таблиця 4.3.1

Основні властивості щільності ймовірностей переходу для неперервного марковського процесу

Властивість	Математичне трактування
Невід'ємність	$\Pi(x; t / x'; t') \geq 0$
Нормованість	$\int_{-\infty}^{\infty} \Pi(x; t / x'; t') dx = 1$
Сингулярність	$\lim_{t \rightarrow t'} \Pi(x; t / x'; t') = \delta(x - x')$
Задовольняє узагальненому рівнянню Маркова	$\Pi(x; t / x_0; t_0) = \int_{-\infty}^{\infty} \Pi(x; t / x'; t') \Pi(x'; t' / x_0; t_0) dx'$

Примітка. Узагальнене рівняння Маркова іноді називають *рівнянням Смолуховського*.

Теорема 1. Для марковського випадкового процесу $X(t)$ щільність ймовірностей переходу $\Pi(x; t / x_0; t_0)$ з стану x_0 в момент часу t_0 в стан x в момент часу t визначається рівнянням

$$-\frac{\partial}{\partial t_0} \Pi(x; t / x_0; t_0) = \sum_{\nu=1}^{\infty} \frac{A_{\nu}(x_0; t_0)}{\nu!} \frac{\partial^{\nu} \Pi(x; t / x_0; t_0)}{\partial x_0^{\nu}}, \quad (4.3.3)$$

де $A_{\nu}(x_0; t_0) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M[(X(t_0 + \Delta t / x_0; t_0) - X(t_0 / x_0; t_0))^{\nu}]$.

Зауваження. Випадкова величина $X(t_0 + \Delta t / x_0; t_0) - X(t_0 / x_0; t_0)$ являє собою прирощення стану, яке трапилося за час Δt . Тому коефіцієнти $A_{\nu}(x_0; t_0)$ можна трактувати як локальні швидкості зміни початкових моментів ν -го порядку прирощення стану.

Доведення теореми 1. Розглянемо три перерізи: $X(t_0)$; $X(t_0 + \Delta t)$; $X(t)$. Для цих перерізів рівняння Смолуховського має вигляд

$$\Pi(x; t / x_0; t_0) = \int_{-\infty}^{\infty} \Pi(x; t / x'; t_0 + \Delta t) \Pi(x'; t_0 + \Delta t / x_0; t_0) dx'. \quad (4.3.4)$$

Використаємо очевидну тотожність

$$\Pi(x; t / x_0; t_0 + \Delta t) = \int_{-\infty}^{\infty} \Pi(x; t / x_0; t_0 + \Delta t) \Pi(x'; t_0 + \Delta t / x_0; t_0) dx' \quad (4.3.5)$$

(тут перший співмножник під інтегралом не залежить від змінної інтегрування, а інтеграл від другого співмножника дорівнює одиниці). Віднімаючи від рівняння (4.3.4) тотожність (4.3.5), маємо:

$$\begin{aligned} & \Pi(x; t / x_0; t_0) - \Pi(x; t / x_0; t_0 + \Delta t) = \\ & = \int_{-\infty}^{\infty} [\Pi(x; t / x'; t_0 + \Delta t) - \Pi(x; t / x_0; t_0 + \Delta t)] \Pi(x'; t_0 + \Delta t / x_0; t_0) dx'. \end{aligned} \quad (4.3.6)$$

Розкладемо вираз в квадратних дужках рівняння (4.3.6) в ряд Тейлора:

$$\begin{aligned} & \Pi(x; t / x'; t_0 + \Delta t) - \Pi(x; t / x_0; t_0 + \Delta t) = \\ & = \sum_{\nu=1}^{\infty} \frac{(x' - x_0)^\nu}{\nu!} \frac{\partial^\nu \Pi(x; t / x_0; t_0 + \Delta t)}{\partial x_0^\nu}. \end{aligned} \quad (4.3.7)$$

Підставимо рівність (4.3.7) у вираз (4.3.6), розділимо його на Δt і зробимо граничний перехід ($\Delta t \rightarrow 0$). В результаті отримаємо рівняння (4.3.3).

Визначення 2. Рівняння (4.3.3) з коефіцієнтами, що дорівнюють нулю для всіх $\nu \geq 3$, називається *дифузійним або першим (зворотним) рівнянням Колмогорова*:

$$\begin{aligned} & -\frac{\partial}{\partial t_0} \Pi(x; t / x_0; t_0) = \\ & = a(x_0; t_0) \frac{\partial}{\partial x_0} \Pi(x; t / x_0; t_0) + \frac{1}{2} b(x_0; t_0) \frac{\partial^2}{\partial x_0^2} \Pi(x; t / x_0; t_0), \end{aligned} \quad (4.3.8)$$

де $a(x_0; t_0) = A_1(x_0; t_0)$ називається *коефіцієнтом зносу*, а коефіцієнт $b(x_0; t_0) = A_2(x_0; t_0)$ – *коефіцієнтом дифузії*.

Визначення 3. Рівняння, аналогічне (4.3.8), вигляду

$$\frac{\partial}{\partial t} \Pi(x; t / x_0; t_0) = -\frac{\partial}{\partial x} [a(x; t) \Pi(x; t / x_0; t_0)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [b(x; t) \Pi(x; t / x_0; t_0)] \quad (4.3.9)$$

називається *рівнянням Фоккера – Планка – Колмогорова або прямим рівнянням Колмогорова*.

Зауваження 1. Рівняння (4.3.8) і (4.3.9) є залежними.

Зауваження 2. З рівняння (4.3.9) випливає рівняння для одновимірної щільності ймовірностей

$$\frac{\partial}{\partial t} f_1(x; t) = -\frac{\partial}{\partial x} [a(x; t) f_1(x; t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [b(x; t) f_1(x; t)]. \quad (4.3.10)$$

Доведення. Для будь-якого $t > t_0$

$$f_1(x; t) = \int_{-\infty}^{\infty} f_1(x_0; t_0) \Pi(x; t / x_0; t_0) dx_0.$$

Тоді, помножуючи (4.3.9) на $f_1(x_0; t_0)$ і інтегруючи обидві частини по x_0 , з врахуванням останнього співвідношення, отримуємо рівняння (4.3.10).

Визначення 4. Марковський процес, що описується рівняннями (4.3.8), (4.3.9), називається *дифузійним*.

Визначення 5. Дифузійний марковський процес називається *однорідним у часі*, якщо щільність ймовірностей переходу $\Pi(x; t / x_0; t_0)$ не залежить прямо від часу t і t_0 , а визначається лише різницею $\tau = t - t_0$: $\Pi(x; t / x_0; t_0) = \Pi(x / x_0; \tau)$.

Зауваження 1. В однорідному за часом дифузійному марковському процесі коефіцієнти a і b не залежать від часу.

Зауваження 2. З співвідношення (4.3.10) випливає, що для стаціонарного у вузькому розумінні (см. розділ 3.2) дифузійного однорідного марковського процесу

$$\frac{d}{dx}[b(x)f_1(x)] = 2a(x)f_1(x) + C, \quad (4.3.11)$$

де C – константа, що визначається з умови нормування.

Теорема 2. Якщо неперервний марковський процес стаціонарний у вузькому розумінні, то він однорідний. Зворотнє твердження невірне.

Доведення. Двовимірна щільність ймовірностей марковського процесу

$$f_2(x_1, x_2; t_1, t_2) = f_1(x_1; t_1)P(x_2; t_2 / x_1; t_1).$$

Для стаціонарного у вузькому розумінні процесу $f_2(x_1, x_2; t_1, t_2) = f_2(x_1, x_2; \tau)$, $f_1(x_1; t_1) = f_1(x_1)$, де $\tau = t_2 - t_1$. Звідси випливає, що $P(x_2; t_2 / x_1; t_1) = P(x_2 / x_1; \tau)$, тобто процес однорідний.

Визначення 6. Дифузійний однорідний марковський процес називається *ергодичним*, якщо:

- 1) існує фінальна одновимірна щільність ймовірностей $f_{1f}(x) = \lim_{t \rightarrow \infty} f_1(x; t)$;
- 2) фінальна ймовірність не залежить від стану процесу в первісний момент часу.

Зауваження 1. Поняття ергодичності для неперервного марковського процесу аналогічно поняттю ергодичності ланцюга Маркова і відрізняється від загального поняття ергодичності для випадкового процесу.

Зауваження 2. Стаціонарний дифузійний марковський процес є ергодичним.

4.4. Приклади дифузійних марковських процесів

4.4.1. Чисто дифузійний (вінерівський) процес

Визначення 1. *Стохастичним диференціальним рівнянням* називається рівняння

$$\frac{dx}{dt} = h(x, t) + g(x, t)n(t),$$

де $h(x, t)$ і $g(x, t)$ – детерміновані функції аргументів, що задовольняють умовам *Ліпшиця*

$$|h(x, t) - h(y, t)| + |g(x, t) - g(y, t)| \leq L|x - y| \quad (L = \text{const} > 0),$$

$n(t)$ – гауссівський білий шум.

Розглянемо найпростіше стохастичне диференціальне рівняння, що описує чисто дифузійний процес.

Нехай в газовій рідині знаходиться частка (рис. 4.4.1) масою m . Молекули середовища хаотично рухаються. При зіткненні молекул з часткою спостерігається переміщення частки. Розглянемо переміщення вздовж фіксованого напрямку. В кожний момент часу t відбувається велика кількість незалежних зіткнень. При цьому сила поштовхів $n(t)$, яка викликає переміщення вздовж вибраного напрямку, може інтерпретуватися як випадкова функція, яка має гауссівський розподіл з нульовим математичним сподіванням і δ -подібною кореляційною функцією $R(t, t')$, тобто $n(t)$ може розглядатися як гауссівський білий шум, у якого

$$M[n(t)] = 0; \quad (4.4.1)$$

$$R(t, t') = \frac{1}{2} N_0 \delta(t' - t). \quad (4.4.2)$$

У відповідності до другого закону Ньютона швидкість $v(t)$ переміщення частки може бути визначена з рівняння

$$m \frac{dv(t)}{dt} = n(t). \quad (4.4.3)$$

Вважаючи, що у початковий момент часу $t = 0$ частка мала нульову швидкість, а її маса дорівнює умовній одиниці, рівняння (4.4.3) можна записати у вигляді вінерівського випадкового процесу, що визначається нижче.

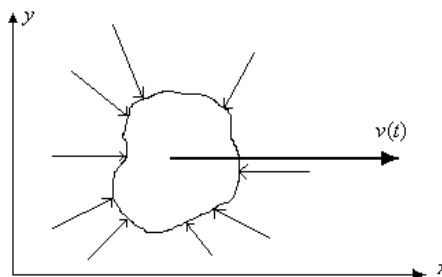


Рис. 4.4.1. Частка, що переміщується в просторі під впливом хаотичного руху молекул

Визначення 2. Вінерівським або чисто дифузійним називається випадковий процес $v(t)$, що описується рівнянням

$$\frac{dv(t)}{dt} = n(t), \quad v(0) = 0, \quad (4.4.4)$$

де $n(t)$ – гауссівський білий шум.

Зауваження 1. З рівняння (4.4.4) випливає, що

$$v(t) = \int_0^t n(t_1) dt_1. \quad (4.4.5)$$

Зауваження 2. Властивості вінерівського процесу наведені в табл. 4.4.1.

Примітка. Визначення процесу з незалежними приростами та мартингалу, що фігурують у таблиці, наведено нижче.

Визначення 3. Випадковий процес $X(t)$, що визначений на інтервалі T , називається процесом з незалежними приростами, якщо для усіх $t_0 < t_1 < \dots < t_N$ з T випадкові величини

$$X(t_0), X(t_1) - X(t_0), \dots, X(t_N) - X(t_{N-1})$$

незалежні у сукупності.

Визначення 4. Випадковий процес $X(t)$, що визначений на інтервалі T , називається мартингалом, якщо для усіх $t \in T$ математичне сподівання $M[X(t)] < \infty$ і для усіх $t_0 < t_1 < \dots < t_N \in T$ математичне сподівання

$$M[X(t_N) / x(t_0), \dots, x(t_{N-1})] = x(t_{N-1}).$$

Зауваження 3. Для вінерівського процесу пряме рівняння Колмогорова з урахуванням наведених властивостей має вигляд

$$\frac{\partial f_1(v; t)}{\partial t} = \frac{1}{4} N_0 \frac{\partial^2 f_1(v; t)}{\partial v^2}.$$

Примітка. Аналітичний розв'язок цього рівняння наведений у рядку 3 таблиці 4.4.1.

Таблиця 4.4.1

Властивості вінерівського процесу

№ з/п	Властивості	Математичне трактування
1	Центрованість	$M[V(t)] = 0$
2	Лінійне зростання дисперсії	$\sigma_v^2(t) = \int_0^t \int_0^t M[n(\tau_1)n(\tau_2)]d\tau_1d\tau_2 = \frac{N_0 t}{2}$
3	Процес гауссівський	$f_1(v;t) = \frac{1}{\sqrt{\pi N_0 t}} \exp\left(-\frac{v^2}{N_0 t}\right)$
4	Процес марковський	$v(t_3) = v(t_2) + \int_{t_2}^{t_3} n(\tau)d\tau$
5	Нульовий коефіцієнт зносу	$a(v;t) = 0$
6	Сталий коефіцієнт дифузії	$b(v;t) = \frac{N_0}{2}$
7	Процес з незалежними приростами	$f_N(v_0, \dots, v_N - v_{N-1}; t_0, \dots, t_N - t_{N-1}) = f_1(v_0; t_0) \dots f_1(v_N - v_{N-1}; t_N - t_{N-1})$
8	Мартингал	$M[V(t_N / v(t_0), \dots, v(t_{N-1}))] = v(t_{N-1})$

- Примітки:**
1. Властивість 3 випливає з того, що $n(t)$ – гауссівський процес, а $v(t)$ і $n(t)$ зв'язані між собою лінійним рівнянням.
 2. Властивість 4 випливає з того, що стан випадкового процесу в момент часу t_3 залежить лише від стану в момент часу t_2 і не залежить від станів в моменти часу, що передують t_2 .
 3. Властивість 5 випливає з виразу (4.4.5) і співвідношення (4.4.1), а властивість 6 – з виразу (4.4.5) і співвідношення (4.4.2).
 4. Властивість 7 випливає з того, що прирости вінерівського процесу – некорельовані і гауссівські.
 5. Властивість 8 випливає з властивості 4 та властивостей білого шуму.

4.4.2. Гауссівський марковський процес

Теорема. Випадковий процес $X(t)$, який задовольняє стохастичному диференційному рівнянню

$$\frac{dx(t)}{dt} + \alpha x(t) = \gamma n(t), \quad (4.4.6)$$

де α, γ – сталі коефіцієнти, а $n(t)$ – гауссівський білий шум, є гауссівським марковським процесом.

Доведення. Той факт, що випадковий процес $X(t)$ є гауссівським випливає з того, що $n(t)$ – гауссівський білий шум, а рівняння (4.4.6) – лінійне.

Те, що випадковий процес $X(t)$ є марковським, встановлюється таким чином.

Загальним розв'язком однорідного рівняння, що відповідає рівнянню (4.4.6), є $x(t) = Ce^{-\alpha t}$, де C – константа. Окремим розв'язком рівняння (4.4.6) є

$$x(t) = \gamma e^{-\alpha t} \int_0^t e^{\alpha \tau} n(\tau) d\tau$$

(в цьому можна переконатися безпосередньою підстановкою розв'язку в рівняння

(4.4.6). Тоді загальним розв'язком рівняння (4.4.6) при початковій умові $x(0) = x_0$ є

$$x(t) = x_0 e^{-\alpha t} + \gamma e^{-\alpha t} \int_0^t e^{\alpha \tau} n(\tau) d\tau. \quad (4.4.7)$$

Розглянемо три послідовні моменти часу $t_3 > t_2 > t_1, t_1 = 0$. З формули (4.4.7) випливає, що

$$x(t_3) = x(t_2) e^{-\alpha(t_3-t_2)} + \gamma e^{-\alpha t_3} \int_{t_2}^{t_3} e^{\alpha \tau} n(\tau) d\tau. \quad (4.4.8)$$

З виразу (4.4.8) видно, що значення функції $x(t)$ в точці t_3 не залежить від моменту часу t_1 , а повністю визначається значенням в момент часу t_2 , тобто процес марковський.

Основні характеристики гауссівського марковського процесу наведені в табл. 4.4.2.

Теорема 2. Стаціонарний гауссівський випадковий процес з кореляційною функцією $R(\tau) = \sigma^2 e^{-\alpha|\tau|}$ є марковським.

Таблиця 4.4.2

Характеристики гауссівського марковського процесу

Назва характеристики	Математичне трактування
Математичне сподівання	$m(t) = x_0 e^{-\alpha t}$
Дисперсія	$\sigma^2(t) = \sigma^2 [1 - e^{-2\alpha t}]$, де $\sigma^2 = \frac{\gamma^2 N_0}{4\alpha}$
Кореляційна функція	$R(t_1, t_2) = \sigma^2 e^{-\alpha t_2-t_1 } (1 - e^{-2\alpha t})$, $\tau = t_2 - t_1; t = \min(t_1, t_2) \geq 0$
Кореляційна функція у стаціонарному стані ($t \rightarrow \infty$)	$R(\tau) = \sigma^2 e^{-\alpha \tau }$
Коефіцієнт зносу	$a(x; t) = -\alpha x(t)$
Коефіцієнт дифузії	$b(x; t) = \gamma^2 \frac{N_0}{2}$

Примітка. Функції, що описують математичне сподівання, дисперсію і у стаціонарному стані кореляційну функцію гауссівського марковського процесу, зображені на рис. 4.4.2.

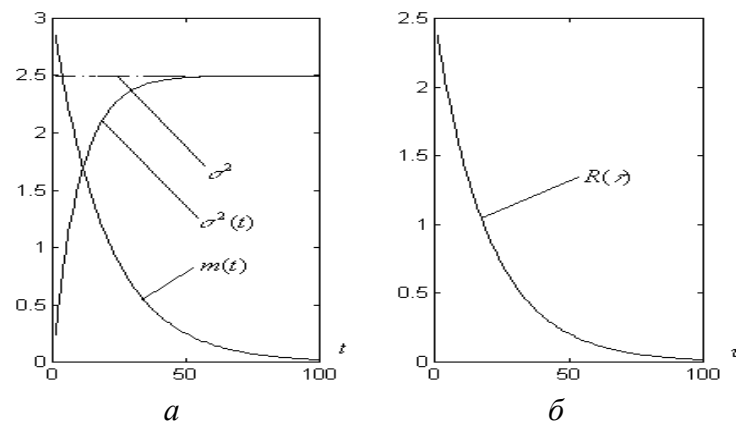


Рис. 4.4.2. Поведінка функцій, що описують математичне сподівання $m(t)$, дисперсію $\sigma^2(t)$ (а) і у стаціонарному стані кореляційну функцію $R(\tau)$ (б) гауссівського марковського процесу

Положення 1. Пряме рівняння Колмогорова для щільності ймовірностей гауссівського випадкового процесу має вигляд

$$\frac{\partial f_1(x;t)}{\partial t} = \alpha \frac{\partial}{\partial x} [x f_1(x;t)] + \frac{\gamma^2 N_0}{2} \frac{\partial^2 f_1(x;t)}{\partial x^2}.$$

Розв'язок цього рівняння описується гауссівською функцією

$$f_1(x;t) = \frac{1}{\sqrt{2\pi\sigma(t)}} \exp\left(-\frac{(x-m(t))^2}{2\sigma^2(t)}\right).$$

Положення 2. Гауссівський марковський процес є ергодичним в розумінні теорії марковських процесів. Фінальна щільність ймовірностей такого процесу визначається співвідношенням

$$f_1(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

4.5. Методи розв'язання рівнянь Колмогорова

4.5.1. Аналітичні методи розв'язання рівнянь Колмогорова

Рівняння Колмогорова (4.3.8) – (4.3.10) відносяться до класу *лінійних диференціальних рівнянь в окремих похідних параболічного типу*.

Розв'язки цих рівнянь повинні задовольняти певним *початковим і граничним* умовам.

Положення 1. Як початкові умови звичайно задають для первісного моменту часу або щільності ймовірностей переходу $\Pi(x;t/x_0;t_0)$, або безумовні щільності ймовірностей $f_1(x;t)$ і їх похідні по x .

Положення 2. Граничні умови звичайно задають для течії ймовірності.

Визначення 1. Функція

$$G(x;t) = a(x;t)f_1(x;t) - \frac{1}{2} \frac{\partial}{\partial x} [b(x;t)f_1(x;t)] \quad (4.5.1)$$

називається *течією ймовірності*.

Зауваження 1. Поняття течії ймовірності пов'язане з трактуванням величини t як змінної часу, а величини x – як просторової змінної. При цьому щільність ймовірностей $f_1(x;t)$ розглядається як щільність деякої субстанції (як концентрація часток) в точці x в момент часу t . Течія ймовірності же являє собою просторово-часову функцію, що залежить від щільності субстанції і швидкості її зміни в просторі. Частина течії $a(x;t)f_1(x;t)$, що визначається щільністю субстанції, називається *систематичною течією*, а частина $\frac{1}{2} \frac{\partial}{\partial x} [b(x;t)f_1(x;t)]$ – *випадковою дифузійною течією*.

Зауваження 2. Звичайно граничні умови задають одним з таких рівнянь:

- 1) $G(x_1;t) = G(x_2;t) = 0$. Ця умова відповідає ідеальним віддзеркалюючим границям x_1 і x_2 ;
- 2) $f_1(x_1;t) = f_1(x_2;t) = 0$. Ця умова відповідає ідеальним поглинаючим границям x_1 і x_2 .

Визначення 2. Рівнянням *неперервності* називається рівняння

$$\frac{\partial}{\partial t} f_1(x;t) + \frac{\partial}{\partial x} G(x;t) = 0. \quad (4.5.2)$$

Зауваження 1. Рівняння неперервності являє собою пряме рівняння Колмогорова (4.3.10), записане відносно течії ймовірності.

Зауваження 2. Фізичний сенс рівняння неперервності полягає в тому, що часові зміни щільності субстанції $f_1(x;t)$ супроводжуються просторовими змінами течії ймовірності $G(x;t)$, причому змінами, при яких субстанція не зникає і не додається, а лише перерозподіляється в просторі.

Положення 3. Для стаціонарного дифузійного марковського процесу з рівняння (4.5.2) випливає, що $G(x;t) = const$. Тоді рівняння (4.5.1) може бути представлене у вигляді

$$-\frac{d}{dx} [b(x)f_1(x)] - 2a(x)f_1(x) = -2G. \quad (4.5.3)$$

Це рівняння являє собою лінійне неоднорідне диференціальне рівняння першого порядку.

Загальний розв'язок рівняння (4.5.3) має вигляд

$$f_1(x) = \frac{C}{b(x)} \exp \left[2 \int_{x_1}^x \frac{a(y)}{b(y)} dy \right] - \frac{2G}{b(x)} \int_{x_1}^x \exp \left\{ 2 \int_{x_1}^z \frac{a(y)}{b(y)} dy \right\} dz.$$

Величину течії ймовірності G визначають з граничних умов, а сукупність параметрів C і x_1 – з умов нормування $f_1(x)$, невід'ємності $f_1(x)$, а також заданих початкових умов.

Зауваження. Описаний підхід розв'язання рівнянь Колмогорова не є єдиним. Відомі також інші методи, наприклад, методи з розподілом змінних, перетворення Лапласа, характеристичної функції, заміни незалежних змінних, гауссівського наближення, а також різноманітні чисельні методи. Деякі з цих методів розроблені не тільки для стаціонарних, а й для нестаціонарних випадків.

4.5.2. Чисельні методи розв'язання рівнянь Колмогорова

Розглянемо один з найпростіших чисельних методів розв'язання диференціальних рівнянь, який називається методом *явних різницевих схем*.

Наведемо схему розв'язання прямого нестаціонарного рівняння Колмогорова (4.3.10) з початковими умовами $f_1(x_0,0)$, $\frac{\partial f_1(x_0,0)}{\partial x_0}$, $\frac{\partial^2 f_1(x_0,0)}{\partial x_0^2}$ і граничними

умовами $f_1(c,t) = f_1(d,t) = 0$.

Область існування нестаціонарного розв'язку обмежена областю $t \geq 0$, $c \leq x \leq d$. Накладемо на цю область сітку з прямокутними комірками розміром $\Delta t \times \Delta x$. При цьому вузли сітки будуть знаходитися в точках $t_n = \Delta t n$;

$x_i = \Delta x i + c$, $n = 0, 1, \dots$; $i = \overline{0, I}$; $I = \left[\frac{d-c}{\Delta x} \right]$, де $[*]$ – оператор виділення цілої частини.

Задача полягає в знаходженні значень функції $f_1(x;t)$ у вузлах сітки.

Користуючись правилами диференціювання, запишемо рівняння (4.3.10) у такому вигляді:

$$\begin{aligned} \frac{\partial f_1(x;t)}{\partial t} &= \frac{1}{2}b(x;t) \frac{\partial^2}{\partial x^2} f_1(x;t) + \left[\frac{\partial}{\partial x} b(x;t) - a(x;t) \right] \frac{\partial}{\partial x} f_1(x;t) + \\ &+ \frac{1}{2} \left[\frac{\partial^2}{\partial x^2} b(x;t) - \frac{\partial}{\partial x} a(x;t) \right] f_1(x;t). \end{aligned} \quad (4.5.4)$$

Представимо похідні функції $f_1(x;t)$ різницевиими відношеннями:

$$\begin{aligned} \left. \frac{\partial f_1(x;t)}{\partial t} \right|_{x_i, t_n} &\approx \frac{f_{i,n+1} - f_{i,n}}{\Delta t}, \\ \left. \frac{\partial f_1(x;t)}{\partial x} \right|_{x_i, t_n} &\approx \frac{f_{i+1,n} - f_{i-1,n}}{2\Delta x}, \\ \left. \frac{\partial^2 f_1(x;t)}{\partial x^2} \right|_{x_i, t_n} &\approx \frac{f_{i+1,n} - 2f_{i,n} + f_{i-1,n}}{\Delta x^2}, \end{aligned}$$

де $f_{i,n} = f_1(x_i, t_n)$.

Якщо коефіцієнти $a(x;t)$, $b(x;t)$ і їх похідні відомі, то рівняння (4.5.4) може бути представлено такою різницевою схемою:

$$f_{i,n+1} = \alpha_{i,n} f_{i+1,n} + \beta_{i,n} f_{i,n} + \gamma_{i,n} f_{i-1,n}, \quad n \geq 0, \quad (4.5.5)$$

де $\alpha_{i,n}, \beta_{i,n}, \gamma_{i,n}$ – коефіцієнти, що визначаються коефіцієнтами $a(x;t)$, $b(x;t)$.

Рекурентна формула (4.5.5) дозволяє з використанням початкових і граничних умов розрахувати $f_1(x;t)$ у всіх вузлах сітки.

Серйозним недоліком розглянутої схеми розв'язання рівняння є необхідність вибору дуже дрібного кроку сітки Δx і Δt для забезпечення усталеності розв'язку. Це зв'язане з тим, що при дрібному кроці похибки рахування, які утворилися на попередніх кроках, зменшуються у часі, в результаті чого ці похибки практично не впливають на розв'язок. Однак при не дуже дрібному кроці похибки не зменшуються, а накопичуються у часі. При цьому розв'язок одержується невірним (не стійким).

Теорема. Явна різницева схема стійка, якщо $b(x;t) \geq 0$ і $\Delta t < \frac{\Delta x^2}{b(x;t)}$.

Зауваження. Окрім розглянутого чисельного способу розв'язання рівнянь Колмогорова відомі і інші чисельні способи. Деякі з них мають низку переваг у порівнянні з явною різницевою схемою. Наприклад, так звана *неявна різницева схема* менш чутлива до похибок обчислень.

4.6. Дискретний марковський процес

Розглянемо випадковий процес $X(t)$, у якого область T визначення аргументу – неперервна множина точок $t \in T$, а простір станів S – дискретна множина точок $\theta_l \in S$, $l = \overline{1, L}$. В будь-які випадкові моменти часу $t_0 < t_1 < \dots$ можливі зміни стану. Такий процес являє собою *дискретну випадкову функцію*.

Визначення 1. *Дискретним марковським процесом* називається дискретна випадкова функція, для якої одновимірна функція розподілу

$$F_1(x_N; t_N / x_0, \dots, x_{N-1}; t_0, \dots, t_{N-1}) = F_1(x_N; t_N / x_{N-1}; t_{N-1}).$$

Зауваження. З наведеної рівності випливає, що ймовірність $P\{x_N; t_N / x_0, \dots, x_{N-1}; t_0, \dots, t_{N-1}\}$ того, що випадкова величина $X(t_N)$ прийме

значення $x_N \in S$, за умови, що випадкові величини $X(t_0), \dots, X(t_{N-1})$ приймали значення $x_0, \dots, x_{N-1} \in S$, визначається рівністю

$$P\{x_N; t_N / x_0, \dots, x_{N-1}; t_0, \dots, t_{N-1}\} = P\{x_N; t_N / x_{N-1}; t_{N-1}\}.$$

Теорема 1. Нехай для дискретного марковського процесу відомі безумовні ймовірності $P_k(t) = P\{X(t) = \theta_k\}$ перебування у станах θ_k у момент часу t ($k = \overline{1, L}$) і ймовірність переходу $\pi_{kl}(t, t')$ з цих станів у деякий стан θ_l у момент часу t' . Тоді безумовна ймовірність $P_l(t') = P\{X(t') = \theta_l\}$ перебування у стані θ_l у момент часу t' визначається таким співвідношенням:

$$P_l(t') = \sum_{k=1}^L P_k(t) \pi_{kl}(t, t'). \quad (4.6.1)$$

Основні властивості ймовірності переходу наведені у табл. 4.6.1.

Зауваження 1. Властивості 1 – 4 ймовірності переходу дискретного марковського процесу аналогічні властивостям щільності ймовірностей неперервного марковського процесу, а властивості 1, 2 і 4 – відповідним властивостям ланцюга Маркова.

Зауваження 2. Важливою особливістю дискретного марковського процесу є властивість сингулярності, яка означає, в даному випадку, що ймовірність переходу в будь-який новий стан за короткий проміжок часу Δt істотно менше ймовірності того, що стан залишиться незмінним.

Таблиця 4.6.1

Основні властивості ймовірності переходу для дискретного марковського процесу

№ з/п	Властивість	Математичне трактування
1	Невід'ємність	$\pi_{kl}(t, t') \geq 0$
2	Нормованість	$\sum_{l=1}^L \pi_{kl}(t, t') = 1, k = \overline{1, L}$
3	Сингулярність	$\lim_{\Delta t \rightarrow 0} \pi_{kl}(t_0, t_0 + \Delta t) = \delta_{kl}^*$
4	Задовольняє рівняння Колмогорова – Чепмена	$\pi_{kl}(t_0, t') = \sum_{q=1}^L \pi_{kq}(t_0, t) \pi_{ql}(t, t'),$ $t_0 < t < t'$

*) **Примітка.** Функція δ_{kl} називається *символом Кронекера* і визначається таким чином:

$$\delta_{kl} = \begin{cases} 1 & \text{при } k = l; \\ 0 & \text{при } k \neq l. \end{cases}$$

Теорема 2. Дискретний марковський процес повністю описується системою лінійних диференціальних рівнянь, що називаються *прямими рівняннями Колмогорова*:

$$\frac{\partial}{\partial t} \pi_{kl}(t_0, t) = \sum_{q=1}^L \pi_{kq}(t_0, t) a_{ql}(t), \quad k, l = \overline{1, L}, \quad (4.6.2)$$

де $a_{ql}(t) = \frac{\partial}{\partial t_1} \pi_{ql}(t, t_1) \Big|_{t_1=t}$.

Визначення 2. Коефіцієнти $a_{ql}(t)$, що визначають систему рівнянь Колмогорова (4.6.2), називаються *інфінітезимальними (породжувачими) ймовірностями*.

Зауваження до теореми 2. Інфінітезимальні ймовірності зв'язані між собою співвідношенням

$$a_{ll}(t) = -\sum_{\substack{q=1 \\ q \neq l}}^L a_{ql}(t). \quad (4.6.3)$$

Доведення теореми 2 і співвідношення (4.6.3). Розкладемо ймовірності переходу $\pi_{ql}(t, t')$ ($q, l = \overline{1, L}$) по t' у ряд Тейлора у районі точки t . Враховуючи властивість сингулярності, можна записати для $t' = t + \Delta t$

$$\pi_{ll}(t, t + \Delta t) \approx 1 + a_{ll}(t)\Delta t; \quad (4.6.4)$$

$$\pi_{ql}(t, t + \Delta t) \approx a_{ql}(t)\Delta t \quad (q \neq l). \quad (4.6.5)$$

Підставляючи співвідношення (4.6.4), (4.6.5) у рівняння Колмогорова – Чепмена, виділяючи в лівій частині прирощення ймовірності переходу, поділяючи обидві частини рівняння на Δt і, нарешті, роблячи граничний перехід, одержуємо вираз (4.6.2). Справедливість співвідношення (4.6.3) випливає з виразів (4.6.4), (4.6.5) і властивості нормованості ймовірності переходу.

Теорема 3. Дискретний марковський процес повністю описується також системою зворотних рівнянь Колмогорова:

$$\frac{\partial}{\partial t_0} \pi_{kl}(t_0, t) = -\sum_{q=1}^L a_{kq}(t_0) \pi_{ql}(t_0, t), \quad k, l = \overline{1, L}, \quad t > t_0. \quad (4.6.6)$$

Доведення теореми 3 аналогічно доведенню теореми 2.

Теорема 4. Безумовна ймовірність $P_l(t)$ перебування в стані l у момент часу t зв'язана з безумовними ймовірностями перебування у станах $q = \overline{1, L}$ у той же момент часу t таким рівнянням:

$$\frac{d}{dt} P_l(t) = \sum_{q=1}^L P_q(t) a_{ql}(t), \quad l = \overline{1, L}. \quad (4.6.7)$$

Доведення. Помножимо обидві частини рівняння (4.6.2) на $P_k(t_0)$ і просумуємо по всіх k . Враховуючи рівність (4.6.1), отримаємо співвідношення (4.6.7).

Зауваження. Вираз (4.6.7) можна розглядати як систему L рівнянь, що зв'язують безумовні ймовірності $P_l(t)$ з інфінітезимальними ймовірностями $a_{ql}(t)$.

Визначення 3. Дискретний марковський процес називається *однорідним*, якщо ймовірності переходу $\pi_{kl}(t, t')$ залежать тільки від різниці аргументів t, t' :

$$\pi_{kl}(t, t') = \pi_{kl}(\tau),$$

де $\tau = t' - t$.

Зауваження 1. З рівнянь (4.6.4), (4.6.5) випливає, що для однорідного дискретного марковського процесу інфінітезимальні ймовірності не залежать від часу: $a_{kl} = \text{const}$.

Зауваження 2. Для однорідного дискретного марковського процесу рівняння (4.6.2), (4.6.6) можуть бути представлені в такому вигляді:

$$\frac{\partial}{\partial \tau} \pi_{kl}(\tau) = \sum_{q=1}^L \pi_{kq}(\tau) a_{ql}, \quad k, l = \overline{1, L},$$

$$\frac{\partial}{\partial \tau} \pi_{kl}(\tau) = -\sum_{q=1}^L a_{kq} \pi_{ql}(\tau), \quad k, l = \overline{1, L}.$$

Визначення 4. Дискретний марковський процес називається *ергодичним*, якщо для вектора безумовних ймовірностей $P(t) = \{P_l(t), l = \overline{1, L}\}$, існує вектор фінальної ймовірності $P = \lim_{t \rightarrow \infty} P(t)$ і він не залежить від стану, у якому знаходився процес у певний момент часу.

Зауваження. Поняття ергодичності для дискретного марковського процесу аналогічне тому ж поняттю для ланцюгів Маркова і неперервного марковського процесу і відрізняється від поняття ергодичності для випадкової функції.

Визначення 5. Однорідний дискретний марковський процес називається *стаціонарним*, якщо його вектор безумовних ймовірностей $P(t)$ не залежить від часу: $P(t) = P$.

Зауваження 1. Для однорідного дискретного марковського процесу введено поняття стаціонарності тотожно поняттю стаціонарності у вузькому розумінні для випадкового процесу.

Зауваження 2. Стаціонарний дискретний марковський процес є ергодичним.

4.7. Застосування теорії марковських процесів

Питання про доцільність використання теорії марковських процесів для розв'язання тієї або іншої практичної задачі визначається, у першу чергу, її змістом і можливістю побудови для неї марковської моделі, з одного боку, не дуже складної, а з іншого – такої, що адекватно відображає закономірності, що властиві задачі. Коректне обґрунтування принципової можливості використання марковської моделі – перший і дуже відповідальний етап розв'язання задачі.

Другий етап, не менш відповідальний, – розв'язання питання про те, який конкретно тип марковської моделі використати і з якими параметрами. Звичайно можна по-різному сформулювати задачу – і як неперервну, і як дискретну.

Моделі, основані на ланцюгах Маркова, простіші і зрозуміліші, ніж моделі, що використовують дискретні або неперервні марковські процеси. Крім того, вони легше моделюються з використанням комп'ютерної техніки. Тому, якщо немає вагомих причин використати інші моделі, слід віддавати перевагу марковським ланцюгам.

Для полегшення прийняття рішення про використання марковських моделей у табл. 4.7.1 перелічені деякі типи випадкових процесів, що мають властивість марковості.

Таблиця 4.7.1

Випадкові процеси, що мають властивість марковості

Тип процесу	Умова марковості
Довільний процес	Перерізи ВП незалежні
Простий ланцюг Маркова	$F_1(x_N / x_0, \dots, x_{N-1}) = F_1(x_N / x_{N-1}), N = 1, 2, \dots$
Складний ланцюг Маркова	$F_1(x_N / x_0, \dots, x_{N-1}) = F_1(x_N / x_{N-M}, \dots, x_{N-1}), N = 1, 2, \dots$

Дискретний марковський процес	$F_1(x_N; t_N / x_0, \dots, x_{N-1}; t_0, \dots, t_{N-1}) = F_1(x_N; t_N / x_{N-1}; t_{N-1})$
Неперервний марковський процес	$F_1(x_N; t_N / x_1, \dots, x_{N-1}; t_1, \dots, t_{N-1}) = F_1(x_N; t_N / x_{N-1}; t_{N-1})$
Стационарний неперервний гауссівський процес	$m = x_0,$ $R(\tau) = \sigma^2 e^{-\alpha \tau }$
Нестационарний неперервний гауссівський процес	$m(t) = x_0 e^{-\alpha t},$ $R(t_1, t_2) = \sigma^2 e^{-\alpha t_1 - t_2 } (1 - e^{-2\alpha t}), \tau = t_2 - t_1; t = \min(t_1, t_2) \geq 0$
Чисто дифузійний (вінерівський) процес	Гауссівський неперервний ВП з кореляційною функцією у вигляді δ -функції і нульовим математичним сподіванням

4.8. Напівмарковський випадковий процес

Розглянемо дискретний випадковий процес, $X(t)$ у якому область визначення T аргументу – неперервна множина $t \in T$, простір станів S – дискретна множина $\theta_l \in S (l = \overline{1, L})$, а зміна станів відбувається у випадкові моменти часу $t_0 < t_1 < \dots$

Визначення 1. *Напівмарковським* називається дискретний випадковий процес $X(t)$, однокрокові переходи якого зі стану $\theta_j (j = \overline{1, L})$ у стан $\theta_k (k = \overline{1, L})$ описуються матрицею Π ймовірностей однокрокових переходів з елементами π_{jk} , а час T_{jk} перебування у стані θ_j до переходу у стан θ_k матрицею $F(t)$ щільності ймовірностей з елементами $f_{jk}(t)$.

Зауваження 1. Для повного ймовірнісного опису напівмарковського процесу крім матриць Π і $F(t)$ треба задати початкові умови, а саме стан θ_i у момент часу t_0 .

Зауваження 2. Характерною рисою напівмарковського процесу є те, що матриці Π і $F(t)$ не залежать від поведінки процесу за межами кроків, що розглядаються.

Зауваження 3. Якщо термін перебування у станах – стала величина T_c (при цьому $f_{jk}(t) = \delta(t - T_c)$), то відліки напівмарковського процесу у моменти часу $t = 0, T_c, 2T_c, \dots$ являють собою однорідний ланцюг Маркова, який називають *вкладеним ланцюгом Маркова*.

Зауваження 4. У загальному випадку напівмарковський процес не є марковським, але доведено, що в окремому випадку, коли щільності ймовірностей $f_{jk}(t) (j, k = \overline{1, L})$ не залежать від станів θ_j і θ_k та описуються експоненційною функцією, тобто $f_{jk}(t) = \nu e^{-\nu t}$, де ν – константа, напівмарковський процес є дискретним марковським процесом.

Теорема 1. Нехай θ_i – стан напівмарковського процесу в початковий момент часу t_0 , Π – матриця ймовірностей однокрокових переходів, а $F(t)$ – матриця

щільності ймовірностей часу перебування у станах $(j, k = \overline{1, L})$. Тоді безумовна щільність ймовірностей $w_i(t)$ часу очікування у стані θ_i ($i = \overline{1, L}$) описується виразом

$$w_i(t) = \sum_{k=1}^L \pi_{ik} f_{ik}(t), \quad (4.8.1)$$

Примітка. Теорема доводиться за допомогою теорем додавання та множення ймовірностей.

Зауваження. З формули (4.8.1) випливає, що математичне сподівання m_{T_i} часу очікування у стані θ_i до переходу у стани θ_k ($k = \overline{1, L}$) описується формулою

$$m_{T_i} = \sum_{k=1}^L \pi_{ik} m_{T_{ik}}, \quad (4.8.2)$$

де $m_{T_{ik}} = \int_0^{\infty} t f_{ik}(t) dt$.

Теорема 2. Нехай для напівмарковського процесу задані ймовірності π_{jk} однокрокових переходів і щільності $f_{jk}(t)$ ймовірностей часу перебування у станах $(j, k = \overline{1, L})$. Тоді умовна ймовірність $\Phi_{ij}(t)$ того, що в момент часу t процес опиниться у стані θ_j при умові, що у початковий момент часу $t_0 = 0$ він знаходиться у стані θ_i , описується рівнянням

$$\Phi_{ij}(t) = \delta_{ij} \Psi_i(t) + \sum_{k=1}^L \pi_{ik} \int_0^t f_{ik}(\tau) \Phi_{kj}(t-\tau) d\tau, \quad (1 \leq i, j \leq L), \quad (4.8.3)$$

де $\Psi_i(t) = P\{T_i > t\}$ – ймовірність того, що до моменту t процес залишиться у стані θ_i :

$$\Psi_i(t) = \int_t^{\infty} w_i(t) dt,$$

$w_i(t)$ – безумовна щільність ймовірностей часу очікування у стані θ_i .

Примітка. Ймовірність $\Phi_{ij}(t)$ описує перехід зі стану у стан на фіксованому інтервалі. Тому її називають *інтервально-перехідною ймовірністю*.

Зауваження 1. У правій частині рівняння (4.8.3) перший доданок описує ситуацію, коли до моменту часу t первинний стан залишається незмінним ($\theta_j = \theta_i$), а другий доданок – ситуацію, коли мають місце зміни станів, причому за час $\tau < t$ процес здійснює однокроковий перехід зі стану θ_i у стан θ_k (можливо навіть у стан θ_i), а потім на інтервалі часу $(\tau, t]$ переходить за декілька кроків у стан θ_j .

Зауваження 2. Розв'язання рівняння (4.8.3) – непроста задача. Якщо вкладений ланцюг Маркова – стаціонарний, то при $t \rightarrow \infty$ розв'язок рівняння (4.8.3) має вигляд

$$\Phi_{ij}(t) = \frac{P_j m_{T_j}}{\sum_{k=1}^L P_k m_{T_k}}, \quad (4.8.4)$$

де $P_k (k = \overline{1, L})$ – фінальна ймовірність перебування у стані θ_k , m_{T_j} – математичне сподівання часу очікування у стані θ_k , що визначається за формулою (4.8.2).

Примітка. Розв'язок (4.8.4) не залежить від часу t та стану θ_i , в якому знаходився процес у первинний момент часу t_0 .

Відомості про книги до розділу 4 наведені в табл. 4.8.1.

Таблиця 4.8.1

Література до розділу 4

Характер книг	Номери книг у списку літератури
Довідники	24, 40, 47, 49
Книги навчального плану для інженерів	32, 41, 42, 73, 77, 78, 99
Книги з математичним ухилом	7, 10, 15, 20, 22, 30, 61, 87, 88, 89, 105
Книги прикладного характеру	6, 38, 57, 82, 85, 86, 92, 94, 95, 98, 104

5. Випадкові течії

5.1. Класифікація і основні характеристики випадкових течій

Визначення 1. *Течією подій* називається послідовність подій, що розгортаються у часі та просторі.

Зауваження. Течією подій можна розглядати як точкову функцію.

Приклад 1. Течіями подій є: послідовність відмов різноманітних вузлів літака за період його експлуатації; послідовність звернень до даних інформаційної системи; послідовність рисок на лінійці і т.д.

Визначення 2. Течії з детермінованим законом появи подій (наприклад, малі риси на лінійці) називаються *регулярними течіями*, а течії з випадковим законом появи подій – *випадковими течіями*.

Визначення 3. Події, що відрізняються фізичним змістом, називаються *неоднорідними*, а течії, що ними утворюються, – *неоднорідними течіями*.

Визначення 4. *Однорідною течією* називається течія подій, фізичний зміст яких або однаковий, або несуттєвий в рамках задачі, яка розглядається.

Приклад 2. Течію відмов комп'ютерів різноманітних типів за період їх тестування можна віднести до однорідних течій, якщо ставити задачу виявлення найбільш надійного типу комп'ютера, і до неоднорідних течій, якщо ставити задачу виявлення найбільш надійного вузла в кожному комп'ютері.

Зауваження. Неоднорідні течії подій часто можуть бути зведені до сукупності однорідних течій.

Положення 1. Найбільш повно випадкову течію описують її ймовірнісні характеристики (рис. 5.1.1):

- спільна ймовірність $P_{k_1 \dots k_M}(t_1, t_1 + \Delta t_1; \dots; t_M, t_M + \Delta t_M)$ складної події, яка полягає в тому, що на інтервалах $[t_1, t_1 + \Delta t_1], \dots, [t_M, t_M + \Delta t_M]$, які не перетинаються, відбудеться, відповідно, k_1, \dots, k_M подій;
- N -вимірна щільність ймовірностей $f_N(t_1, \dots, t_N)$ моментів t_1, \dots, t_N настання подій;
- N -вимірна щільність ймовірностей $f_N(\tau_1, \dots, \tau_N)$ тривалості інтервалів $\tau_1 = t_1 - 0; \tau_2 = t_2 - t_1; \dots; \tau_N = t_N - t_{N-1}$ між сусідніми подіями.

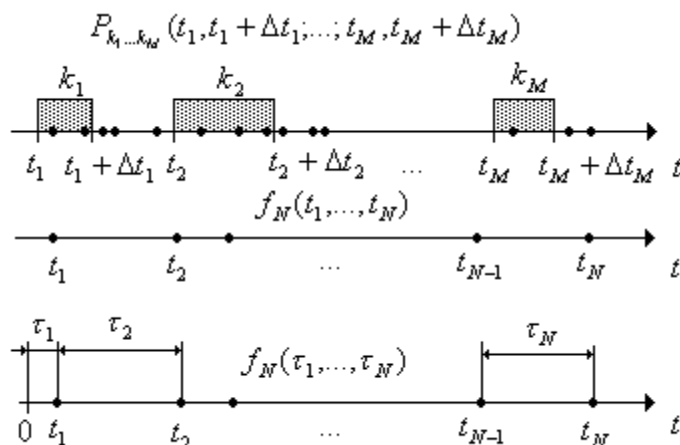


Рис. 5.1.1. Основні ймовірнісні характеристики випадкових течій

Визначення 5. Випадкова течія називається *стаціонарною*, якщо ймовірність $P_{k_1 \dots k_M}(t_1 + \theta, t_1 + \Delta t_1 + \theta; \dots; t_M + \theta, t_M + \Delta t_M + \theta)$ спостереження $k_1 \dots k_M$ подій

відповідно на інтервалах часу $[t_1 + \theta, t_1 + \Delta t_1 + \theta; \dots; t_M + \theta, t_M + \Delta t_M + \theta]$ для довільних M не залежить від зсуву θ , тобто

$$P_{k_1 \dots k_M}(t_1 + \theta, t_1 + \Delta t_1 + \theta; \dots; t_M + \theta, t_M + \Delta t_M + \theta) = P_{k_1 \dots k_M}(t_i, t_i + \Delta t_i; \dots; t_M, t_M + \Delta t_M).$$

Визначення 6. Випадкова течія називається *ординарною*, якщо ймовірність настання більш однієї події $P_k(t, t + \Delta t)$ на довільному інтервалі $[t, t + \Delta t]$ ($k \geq 2$) мала у порівнянні з ймовірністю настання однієї події $P_1(t, t + \Delta t)$ на тому ж інтервалі: $P_1(t, t + \Delta t) \gg P_k(t, t + \Delta t)$, $k = 2, 3, \dots$

Зауваження. Для ординарної випадкової течії факт одночасного настання двох і більше подій практично не може мати місця.

Визначення 7. Випадкова течія подій називається *течією без післядії* (або *течією з незалежними приростами*), якщо для будь-яких інтервалів часу, які не перетинаються, число подій, що спостерігаються на одному з них, не залежить від числа подій, що спостерігаються на іншому, тобто

$$P_{k_1 \dots k_M}(t_1, t_1 + \Delta t_1; \dots; t_M, t_M + \Delta t_M) = \prod_{m=1}^M P_{k_m}(t_m, t_m + \Delta t_m).$$

Ймовірнісні характеристики для ординарної течії без післядії наведено на рис. 5.1.2.

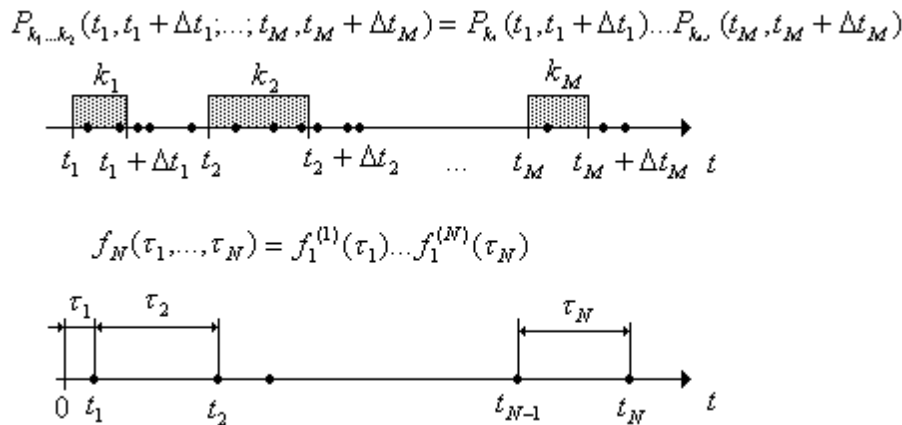


Рис. 5.1.2. Ймовірнісні характеристики для ординарної течії без післядії

Зауваження 1. Для ординарної течії без післядії тривалості інтервалів між сусідніми подіями незалежні:

$$f_N(\tau_1, \dots, \tau_N) = \prod_{n=1}^N f_1^{(n)}(\tau_n).$$

Ця властивість притаманна не тільки течії без післядії, але і деяким іншим течіям (див. підрозділ 5.4).

Зауваження 2. У загальному випадку регулярна течія не є течією без післядії.

Приклади 3.

1. Течія пасажирів, що входять на станцію метро, можна вважати течією без післядії, якщо прийняти, що причини появи пасажирів на станції незалежні.

2. Течія же пасажирів, що входять в метро після закінчення, наприклад, футбольного матчу, не можна вважати течією без післядії, бо причина появи на станції більшості пасажирів викликана однією і тією ж причиною.

3. Течія пасажирів, що виходять з поїзда метро, не можна вважати течією без післядії, бо причина їхньої появи на пероні одна і та ж сама: прибуття поїзда.

Типові задачі. При розгляді випадкових течій звичайно потрібно розв'язувати два класи задач. Перший клас зв'язаний з визначенням числа подій на деякому заданому інтервалі T , другий клас – з визначенням величини інтервалу між сусідніми подіями або між довільною точкою відліку і моментом настання чергової події.

Положення 2. Для опису випадкових течій поряд з ймовірнісними характеристиками, що найбільш повно характеризують випадкову течію, використовують і неймовірнісні числові характеристики, зокрема (рис. 5.1.3):

- математичне сподівання $m(t_1, t_2)$ і дисперсію $D(t_1, t_2)$ числа подій на інтервалі $[t_1, t_2]$;
- математичне сподівання m_{τ_n} і дисперсію D_{τ_n} інтервалу між подіями з номерами $n-1$ і n ;
- миттєву інтенсивність течії $\mu(t)$, що визначається нижче;
- миттєву щільність течії $\lambda(t)$, що теж визначається нижче, та ін.

Зауваження. Для стаціонарної випадкової течії моменти $m(t, t+T)$ і $D(t, t+T)$ не залежать від часу t , тобто $m(t, t+T) = m(T)$, $D(t, t+T) = D(T)$.

Визначення 8. Миттєвою інтенсивністю течії $\mu(t)$ називається похідна по аргументу t від математичного сподівання $m(t) = m(0, t)$ числа подій на інтервалі $[0, t]$:

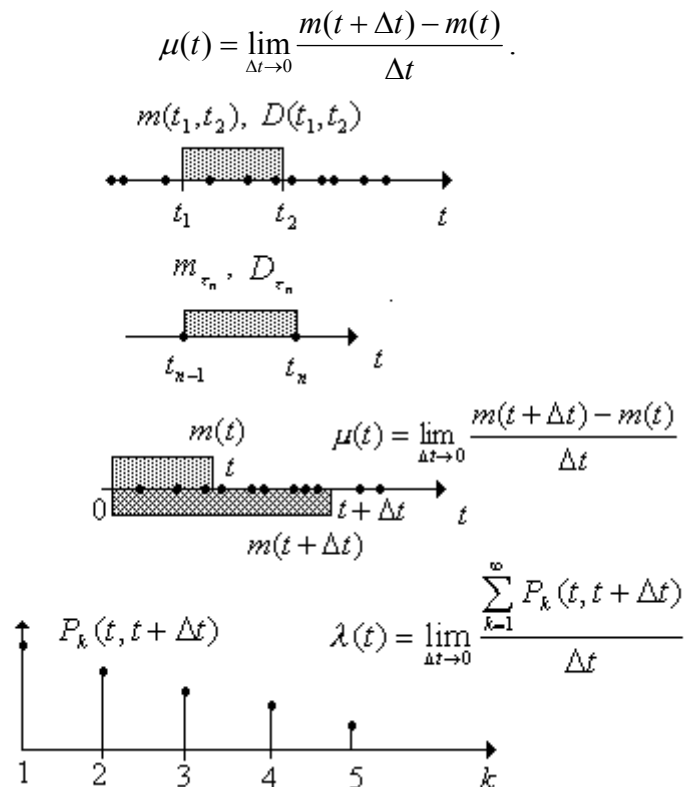


Рис. 5.1.3. Основні неймовірнісні (числові) характеристики випадкової течії

Визначення 9. Миттєвою щільністю течії (миттєвим значенням параметра течії) $\lambda(t)$ називається границя відношення ймовірності $\sum_{k=1}^{\infty} P_k(t, t + \Delta t)$ настання хоча б однієї події за інтервал часу Δt до величини цього інтервалу при $\Delta t \rightarrow 0$:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{\sum_{k=1}^{\infty} P_k(t, t + \Delta t)}{\Delta t}.$$

Зауваження 1. Якщо течія стаціонарна, то миттєва інтенсивність течії і миттєва щільність течії не залежать від часу, тобто $\mu(t) = \mu$; $\lambda(t) = \lambda$.

Зауваження 2. Якщо течія ординарна, то миттєва інтенсивність течії $\mu(t)$ дорівнює миттєвій щільності $\lambda(t)$.

Зауваження 3. Деякі автори називають функцію $\mu(t)$ щільністю течії, а $\lambda(t)$ – інтенсивністю течії.

5.2. Найпростіша течія

Визначення 1. Найпростішою течією називається однорідна стаціонарна ординарна течія без післядій.

Теорема 1. Нехай найпростіша течія має щільність λ . Тоді ймовірність $P_k(T)$ того, що на інтервалі $[0, T]$ відбудеться рівно k подій (рис. 5.2.1), описується законом Пуассона:

$$P_k(T) = \frac{(\lambda T)^k}{k!} e^{-\lambda T}.$$

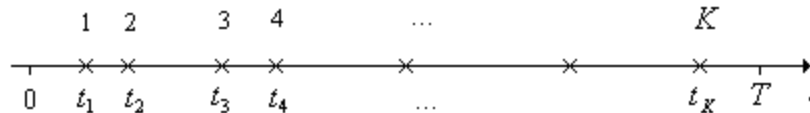


Рис. 5.2.1. K подій в інтервалі $[0, T]$

Теорема 2. Нехай щільність найпростішої течії дорівнює λ . Тоді (рис. 5.2.2):

- 1) тривалість інтервалу τ між двома сусідніми подіями;
- 2) тривалість τ_0 між довільною точкою відліку і моментом появи чергової події;

3) залишок часу $\Delta\tau_n = t_{n+1} - (t_n + \theta_n)$ до моменту настання чергової події t_{n+1} , за умови, що попередня подія відбулася в момент t_n , а в інтервалі $[t_n, t_n + \theta_n]$ подій не було, описуються однаковими на вигляд експоненціальними щільностями ймовірності $f_1(\tau)$, $f_1(\tau_0)$, $f_1(\Delta\tau_n)$ з параметром λ :

$$f_1(\tau) = \lambda e^{-\lambda\tau}.$$

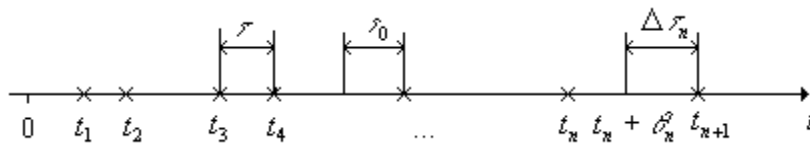


Рис. 5.2.2. Тривалості інтервалів: τ – тривалість інтервалу між двома будь-якими сусідніми подіями; τ_0 – тривалість інтервалу між довільним моментом часу і моментом настання чергової події; $\Delta\tau_n$ – залишок часу до настання чергової події

Основні формули, що описують моменти найпростішої течії, наведені в табл. 5.2.1.

Таблиця 5.2.1

Моменти найпростішої течії

Назва моменту	Співвідношення параметрів
Математичне сподівання числа подій на інтервалі T	$m(T) = \lambda T$
Дисперсія числа подій на інтервалі T	$D(T) = \lambda T$
Математичні сподівання тривалостей $\tau, \tau_0, \Delta\tau_n$	$m_\tau = m_{\tau_0} = m_{\Delta\tau_n} = \frac{1}{\lambda}$
Дисперсії тривалостей $\tau, \tau_0, \Delta\tau_n$	$D_\tau = D_{\tau_0} = D_{\Delta\tau_n} = \frac{1}{\lambda^2}$

Теорема 3. Інтенсивність найпростішої течії μ дорівнює його щільності λ .

Зауваження. Має місце і більш загальне твердження: якщо течія стаціонарна і ординарний, то інтенсивність течії дорівнює його щільності.

5.3. Узагальнення найпростішої течії

Теорема 1. Течія, що утворена накладенням M незалежних найпростіших течій з щільностями λ_m ($m = \overline{1, M}$), є найпростіша течія з щільністю $\lambda = \sum_{m=1}^M \lambda_m$.

Зауваження 1. Доведена і більш загальна теорема, що аналогічна центральній граничній теоремі теорії ймовірностей. Ця теорема стверджує, що у достатньо загальних умовах течія, яка утворена накладенням M стаціонарних ординарних течій з довільним законом розподілу інтервалів між подіями, при $M \rightarrow \infty$ збігається до найпростішої течії.

Зауваження 2. На практиці течію можна вважати найпростішою, якщо вона утворена більш ніж 5 – 10 незалежними стаціонарними ординарними течіями.

Визначення 1. *Нестаціонарною пуассонівською* течією називається нестаціонарна ординарна течія однорідних подій без післядії.

Теорема 2. Нехай нестаціонарна пуассонівська течія має щільність течії $\lambda(t)$. Тоді ймовірність $P_k(t_1, t_2)$ того, що в інтервалі $[t_1, t_2]$ відбудеться рівно k подій, описується законом Пуассона:

$$P_k(t_1, t_2) = \frac{a^k(t_1, t_2)}{k!} e^{-a(t_1, t_2)},$$

де $a(t_1, t_2) = \int_{t_1}^{t_2} \lambda(t) dt$.

Теорема 3. Нехай на інтервалі $[t_1, t_2]$ нестаціонарна пуассонівська течія практично стаціонарна (тобто $\lambda(t)$ мало змінюється на інтервалі $[t_1, t_2]$). Тоді на інтервалі стаціонарності $t \in [t_1, t_2]$ величина інтервалу τ між сусідніми подіями описується експоненціальною щільністю ймовірності:

$$f_1(\tau) = \lambda_{cp} e^{-\tau \lambda_{cp}},$$

де λ_{cp} – середнє значення $\lambda(t)$ на інтервалі стаціонарності, який розглядається.

Зауваження. В загальному випадку для двох різних інтервалів стаціонарності щільності ймовірностей між сусідніми подіями $f_1(\tau)$ відрізняються одна від одної.

Моменти для нестационарної пуассонівської течії наведені в табл. 5.3.1.

Таблиця 5.3.1

Моменти для числа подій на інтервалі $[t_1, t_2]$ і для величини інтервалу τ між сусідніми подіями (течія нестационарна пуассонівська)

Назва моменту	Співвідношення параметрів
Математичне сподівання числа подій на інтервалі $[t_1, t_2]$	$m(t_1, t_2) = a(t_1, t_2)$
Дисперсія числа подій на інтервалі $[t_1, t_2]$	$D(t_1, t_2) = a(t_1, t_2)$
Математичне сподівання величини інтервалу τ між сусідніми подіями	$m_\tau = \frac{1}{\lambda_{cp}}$
Дисперсії величини інтервалу τ між сусідніми подіями	$D_\tau = \frac{1}{\lambda_{cp}^2}$

5.4. Течії з обмеженою післядією

Визначення. Течією з обмеженою післядією називається течія, у якій тривалості інтервалів τ_n ($n=1,2,\dots$) між сусідніми подіями є незалежними випадковими величинами, тобто

$$f_N(\tau_1, \dots, \tau_N) = \prod_{n=1}^N f_1^{(n)}(\tau_n).$$

Приклад. Нехай деякий блок працює безперервно до своєї відмови, після чого він миттєво замінюється новим. Для усунення відмов використовуються блоки різноманітної модифікації з різним гарантійним строком експлуатації. Строк безвідмовної роботи блоку випадковий. Блоки виходять з ладу незалежно один від одного. За цих умов течія відмов є течією з обмеженою післядією.

5.4.1. Течії Пальма

Визначення 1. Течіями Пальма називаються однорідні ординарні течії з обмеженою післядією, у яких закони розподілу інтервалів між моментами появи сусідніх подій однакові, тобто $f_1^{(2)}(\tau) = \dots = f_1^{(N)}(\tau) = f_1(\tau)$.

Зауваження 1. Найпростіша течія є окремим випадком течії Пальма: в ній інтервали τ_1, \dots, τ_N між послідовними подіями являють собою незалежні випадкові величини, які розподілені за показовим законом.

Зауваження 2. Нестационарна пуассонівська течія не є течією Пальма (див. зауваження 1 до теореми 3 в підрозділі 5.3).

Приклад 1. Якщо в прикладі підрозділу 5.4 для усунення відмов використовуються блоки однакової модифікації з однаковим гарантійним строком експлуатації, то течія відмов являє собою течію Пальма. Якщо до того ж строк роботи блоків описується показовим законом, то течія буде найпростіша.

Приклад 2. Група з N літаків летить в бойовому порядку один за одним. Кожний літак, крім ведучого, намагається утриматись на заданій відстані L від

того, який іде попереду. Внаслідок похибок навігаційних систем відстані між літаками L_1, \dots, L_{N-1} різні. Моменти подолання літаками заданого рубежу утворюють течія Пальма.

Зауваження 3. Для течії Пальма спільна щільність ймовірностей тривалості будь-якого числа N інтервалів може бути наведена у вигляді

$$f_N(\tau_1, \dots, \tau_N) = f_1^{(1)}(\tau_1) \prod_{n=2}^N f_1(\tau_n),$$

де $f_1^{(1)}(\tau_1)$ – щільність ймовірностей тривалості першого інтервалу (інтервалу між початковим моментом часу $t=0$, в якому події не було, і моментом t_1 настання першої події), $f_1(\tau)$ – щільність ймовірностей тривалості τ інтервалу між двома сусідніми подіями (рис. 5.4.1).

Зауваження 4. Щільності $f_1^{(1)}(\tau_1)$ і $f_1(\tau)$ в загальному випадку відрізняються одна від одної хоча, наприклад, в окремому випадку найпростішої течії вони збігаються.

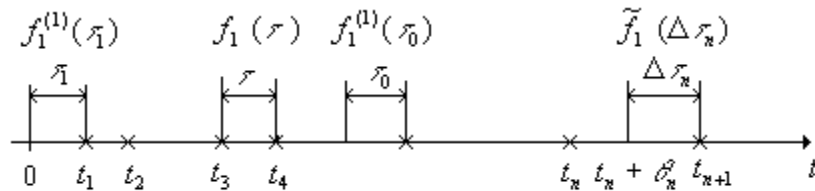


Рис. 5.4.1. Тривалість інтервалів: τ_1 – тривалість інтервалу між початковим моментом часу $t=0$, в якому події не було, і моментом t_1 настання першої події; τ – тривалість інтервалу між будь-якими двома сусідніми подіями; τ_0 – тривалість інтервалу між довільним моментом часу і моментом настання чергової події; $\Delta\tau_n$ – залишок часу до настання чергової події.

Теорема 1. Щільності ймовірностей $f_1^{(1)}(\tau_1)$ і $f_1(\tau)$ тривалості першого інтервалу і наступних інтервалів зв'язані між собою формулою Пальма:

$$f_1^{(1)}(\tau_1) = \lambda \left[1 - \int_0^{\tau_1} f_1(\tau) d\tau \right],$$

де $\lambda = \mu = \frac{1}{m_\tau}$.

Теорема 2. Щільність ймовірностей інтервалу τ_0 між довільним моментом часу і моментом настання чергової події описується тією ж щільністю ймовірності $f_1^{(1)}(\tau_0)$, що і щільність ймовірностей тривалості першого інтервалу (див. рис. 5.4.1).

Зауваження 1. Умовна щільність ймовірностей залишку часу $\Delta\tau_n$ до моменту t_{n+1} настання чергової події, за умови, що вона не трапилась на інтервалі $[t_n, t_n + \theta_n]$ (див. рис. 5.4.1), описується щільністю ймовірності $\tilde{f}_1(\Delta\tau_n)$. В загальному випадку вона відрізняється як від $f_1^{(1)}(\tau_1)$, так і від $f_1(\tau)$.

Положення. Течії на виході систем масового обслуговування часто добре описуються течіями Пальма.

Теорема 3 (Пальма). Нехай на багатоканальну систему масового обслуговування поступає течія заявок, що описуються течією Пальма. Час обслуговування заявок випадковий з показниковим законом. Якщо усі канали обслуговування зайняті, заявка одержує відмову (тобто не обслуговується). Тоді течія заявок, що одержали відмову описуються течією Пальма.

Зауваження. Якщо вхідна течія найпростіша, течія заявок, що одержали відмову, не обов'язково найпростіша.

5.4.2. Течія Ерланга

Визначення 1а. Течією Ерланга k -го порядку є течія, що одержується з найпростішої течії шляхом відкидання всіх подій, номери яких без одиниці не кратні k (рис. 5.4.2).

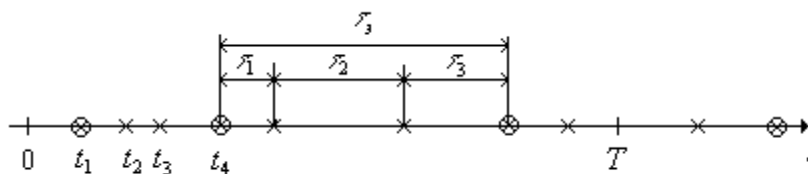


Рис. 5.4.2. Тривалості інтервалів для течії Ерланга (на прикладі течії Ерланга 3-го порядку): τ_e – тривалість інтервалу між двома сусідніми подіями течії Ерланга; τ_n – тривалість інтервалу між двома сусідніми подіями породжуючої найпростішої течії

Визначення 1б. Течією Ерланга порядку k називається течія, в якій тривалість інтервалу між двома сусідніми подіями являє собою випадкову величину τ_e , яка дорівнює сумі k незалежних випадкових величин, розподілених за експоненційним законом з параметром щільності λ :

$$\tau_e = \sum_{n=1}^k \tau_n ;$$

$$f_1(\tau_n) = \lambda e^{-\lambda \tau_n}, \quad \tau_n \geq 0 .$$

Зауваження 1. Визначення 1а, 1б течії Ерланга еквівалентні.

Зауваження 2. Течія Ерланга є окремим випадком течії Пальма.

Зауваження 3. Порядок k течії Ерланга характеризує ступінь післядії: чим більше k , тим більше ступінь післядії. При $k = 1$ течія Ерланга вироджується в найпростішу течію. Доведено, що при $k \rightarrow \infty$ течія Ерланга прямує до регулярної течії.

Теорема 1. В течії Ерланга k -го порядку тривалість інтервалів τ_e між моментами появи сусідніх подій описується щільністю ймовірності

$$f_k(\tau_e) = \lambda \frac{(\lambda \tau_e)^{k-1}}{(k-1)!} e^{-\lambda \tau_e}, \quad \tau_e \geq 0,$$

де λ – щільність породжуючої найпростішої течії.

Теорема 2. В течії Ерланга k -го порядку ймовірність $P_{l,k}(T)$ спостереження l подій на інтервалі часу $[0, T]$ описується формулою

$$P_{l,k}(T) = \sum_{n=k}^{(l+1)k-1} \frac{(\lambda T)^n}{n!} e^{-\lambda T} .$$

Основні параметри течії Ерланга наведені в табл. 5.4.1.

Таблиця 5.4.1

Основні параметри течії Ерланга

Назва параметра	Співвідношення параметрів для течії Ерланга	Співвідношення параметрів для узагальненого течії Ерланга
Математичне сподівання інтервалу між подіями	$m_{\tau_e, k} = \frac{k}{\lambda}$	$m_{\tau_{yz}, k} = \sum_{n=1}^k \frac{1}{\lambda_n}$
Дисперсія інтервалу між подіями	$D_{\tau_e, k} = \frac{k}{\lambda^2}$	$D_{\tau_{yz}, k} = \sum_{n=1}^k \frac{1}{\lambda_n^2}$
Миттєва щільність течії Ерланга	$\lambda_e = \frac{\lambda}{k}$	$\lambda_{yz} = \left(\sum_{n=1}^k \frac{1}{\lambda_n} \right)^{-1}$

Положення. Течія Ерланга часто використовується для моделювання реальних течій подій з урахуванням їх реальної післядії. При цьому реальна течія моделюється течією Ерланга такого порядку k і з таким параметром щільності λ_e , щоб для реальної течії і моделюючої був забезпечений збіг математичних сподівань і дисперсій. Вказана умова погодження досягається, якщо

$$k = \frac{m_{\tau_p}^2}{D_{\tau_p}}; \quad \lambda_e = \frac{1}{m_{\tau_p}},$$

де m_{τ_p} і D_{τ_p} – математичне сподівання і дисперсія реальної течії.

5.4.3. Узагальнена течія Ерланга

Визначення 1. Узагальненою течією Ерланга порядку k називається течія, в якій тривалість інтервалів τ_{yz} між двома сусідніми подіями являє собою суму k випадкових величин, розподілених за експоненційним законом з параметрами λ_n , $n = \overline{1, k}$, тобто

$$f_1(\tau_n) = \lambda_n e^{-\lambda_n \tau_n}.$$

Зауваження 1. Узагальнена течія Ерланга порядку k вироджується у звичайну течію Ерланга порядку k , коли параметри $\lambda_n = \lambda = const$.

Теорема 1. В узагальненій течії Ерланга порядку k тривалість інтервалу між двома сусідніми моментами часу описується формулою

$$f_k(\tau_{yz}) = \prod_{n=1}^k \lambda_n \sum_{n=1}^k \left(\frac{e^{-\lambda_n \tau_{yz}}}{\prod_{\substack{n,i=1 \\ n \neq i}}^k (\lambda_i - \lambda_n)} \right).$$

Параметри узагальненої течії Ерланга наведені в табл. 5.4.1.

5.4.4. Транспортна течія

Визначення. Транспортною течією називається течія Пальма, в якій тривалість інтервалу τ між двома сусідніми подіями являє собою суму сталої величини τ_0 і випадкової величини τ' : $\tau = \tau_0 + \tau'$ (рис. 5.4.3).

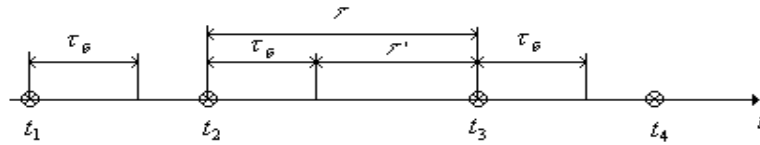


Рис. 5.4.3. Тривалість інтервалу для транспортної течії

Зауваження 1. Транспортна течія часто зустрічається при розв'язанні задач, зв'язаних з безпекою руху. В цих задачах величина τ_0 трактується як інтервал безпеки, в середині якого знаходження двох транспортних засобів неприпустимо. Назва «транспортна течія» не є загальноживаною.

Зауваження 2. Звичайно розглядають два класи транспортних течій:

1) з експоненційним розподілом інтервалу τ' :

$$f_1(\tau) = \begin{cases} 0, & \tau < \tau_0; \\ \lambda e^{-\lambda(\tau-\tau_0)}, & \tau \geq \tau_0, \end{cases}$$

2) з рівномірним розподілом інтервалу τ' в інтервалі $[0, a]$:

$$f_1(\tau) = \begin{cases} 0, & \tau < \tau_0, \tau > a + \tau_0; \\ \frac{1}{a}, & \tau_0 \leq \tau \leq \tau_0 + a. \end{cases}$$

У табл. 5.4.2 наведені відповідні цим щільностям параметри.

Таблиця 5.4.2

Параметри транспортних течій

Назва параметра	Співвідношення параметрів для експоненційного розподілу	Співвідношення параметрів для рівномірного розподілу
Математичне сподівання інтервалу між подіями	$m_\tau = \tau_0 + \frac{1}{\lambda}$	$m_\tau = \tau_0 + \frac{a}{2}$
Дисперсія інтервалу між подіями	$D_\tau = \frac{1}{\lambda^2}$	$D_\tau = \frac{a^2}{12}$
Миттєва щільність течії	$\lambda_T = \frac{\lambda}{1 + \lambda\tau_0}$	$\lambda_T = \frac{2}{a + 2\tau_0}$

Система класифікації, котра упорядковує розглянуті класи випадкових течій наведена на рис. 5.4.4.

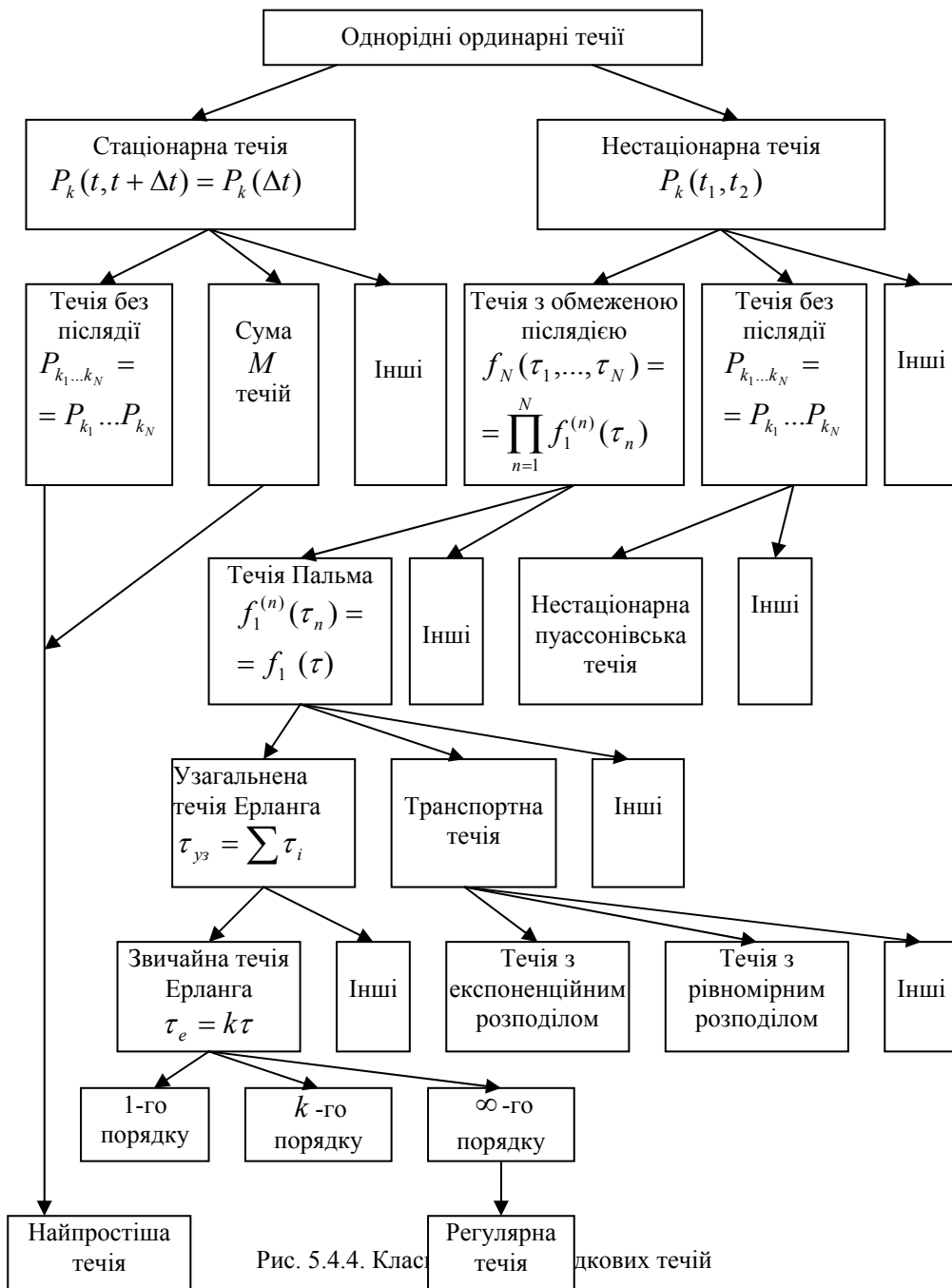


Рис. 5.4.4. Класифікація однорідних течій

5.5. Використання теорії випадкових течій

Випадкові течії широко використовуються при розв'язанні різних задач, у першу чергу задач теорії масового обслуговування.

Теорія масового обслуговування описує роботу систем, що обробляють масові заявки. Прикладами таких систем є телефонні станції, банки, комп'ютерні мережі, мережа INTERNET, різні довідкові служби, служба швидкої допомоги, ремонтні майстерні, підприємства торгівлі, підприємства побутового обслуговування, служби забезпечення зльоту та посадки літаків, склади та ін.

Заявки, які надходять у систему на обслуговування, і заявки, які залишають систему після обслуговування, утворюють дві течії: вхідна і вихідна. Обидві

течії випадкові. Час обслуговування кожної заявки випадковий. Тому вихідна течія відрізняється своїми статистичними характеристиками від вхідної течії.

Друга галузь використання випадкових течій – теорія надійності. У рамках цієї теорії за допомогою випадкових течій описуються порушення працездатності і відкази виробів.

Третя галузь використання випадкових течій зв'язана з описом різних фізичних уявлень, наприклад, ядерного розпаду, старіння матеріалу під впливом ударних дій та ін.

Випадкові течії використовуються також у медицині, біології та різних галузях техніки.

Відомості про книги до розділу 5 наведені в табл. 5.5.1.

Таблиця 5.5.1

Література до розділу 5

Характер книг	Номери книг у списку літератури
Довідники	24, 40, 47, 49
Книги навчального плану для інженерів	16, 32
Книги з математичним ухилом	20, 23, 88, 89, 105, 109, 113
Книги прикладного характеру	38, 79, 83

6. Основні положення вибіркової теорії

Визначення 1. *Статистичною інформацією* називається будь-яка інформація про випадкові явища чи процеси (функції).

Визначення 2. *Математична статистика* – розділ математики, присвячений встановленню закономірностей випадкових явищ або процесів на основі реєстрації, систематизації й обробки результатів спостережень або вимірів.

6.1. Способи представлення статистичної інформації про випадкову величину й властивості цієї інформації

6.1.1. Генеральна й вибірка сукупності

Визначення 1. *Генеральною сукупністю* називається нескінченна множина всіх реалізацій випадкового явища чи процесу (функції).

Примітка 1. Множина реалізацій може бути як зліченою, так і незліченою. На практиці часто вважають (і у подальшому ми будемо вважати) її зліченою множиною.

Зауваження. Генеральною сукупністю випадкової величини X є вся множина її реалізацій x_n , а випадкової функції $X(t)$ – уся множина реалізацій $x_n(t)$, де $n = 1, 2, \dots$.

Примітка 2. Генеральна сукупність визначається не тільки для скалярних, але і для векторних випадкових явищ, причому як неперервних, так і дискретних. В цьому розділі з метою спрощення викладення будемо розглядати тільки випадкові явища, що описуються скалярною випадковою величиною.

Визначення 2. *Членами генеральної сукупності* називаються її елементи.

Визначення 3. Скінченна множина членів генеральної сукупності x_1, \dots, x_N , що одержана при кінцевому числі N дослідів, називається *вбіркою з генеральної сукупності* або просто *вбіркою*.

Визначення 4. Елементи вибірки називаються *вбірковими значеннями* або *реалізаціями*.

Визначення 5. *Обсягом вибірки* називається число N вибіркових значень, що складають вибірку.

Визначення 6. Кажуть, що *вбірка* x_1, \dots, x_N *належить розподілу* $F_1(x)$, якщо вона отримана з генеральної сукупності, що описується функцією розподілу $F_1(x)$.

Визначення 7. Множина різноманітних вибірок обсягу N , сформованих з однієї генеральної сукупності, являє собою N -вимірний випадковий вектор

$$\vec{X} = \begin{pmatrix} x_1^{(1)} & \dots & x_N^{(1)} \\ x_1^{(2)} & \dots & x_N^{(2)} \\ \dots & \dots & \dots \end{pmatrix} = (X_1, \dots, X_N),$$

який називається *випадковою вибіркою (вбірковою сукупністю)*.

Зауваження 1. Необхідно відрізнити випадкову вибірку, що являє собою випадковий вектор \vec{X} , від просто вибірки \vec{x} , що являє собою реалізацію випадкової вибірки \vec{X} і не є випадковим вектором.

Зауваження 2. Компоненти випадкової вибірки (X_1, \dots, X_N) є взаємно незалежні випадкові величини X_n ($n = \overline{1, N}$) з однаковим законом розподілу $F_1(x_n)$, що співпадає з законом розподілу елементів генеральної сукупності.

Зауваження 3. N -вимірна функція розподілу $F_N(\vec{x})$ випадкової вибірки \vec{X} дорівнює добутку одновимірних функцій розподілу $F_1(x_n)$ компонент x_n випадкової вибірки:

$$F_N(\vec{x}) = \prod_{n=1}^N F_1(x_n).$$

Визначення 8. Функцією вірогідності $L_{\vec{x}}$ називається багатовимірна щільність ймовірностей випадкової вибірки \vec{X} .

Зауваження. Якщо елементи вибірки незалежні, функція вірогідності $L_{\vec{x}}$ дорівнює добутку одновимірних щільностей ймовірностей $f_1(x_n)$ компонент X_n випадкової вибірки \vec{X} :

$$L_{\vec{x}} = \prod_{n=1}^N f_1(x_n). \quad (6.1.1)$$

При цьому треба мати на увазі, що щільності ймовірностей $f_1(x_n)$ компонент ймовірності X_n співпадають з щільністю ймовірності генеральної сукупності.

Примітка. Крім наведеної широко розповсюдженої моделі вибірки іноді використовують інші моделі. В деяких з них компоненти випадкової вибірки \vec{X} мають різний закон розподілу та залежні між собою. Далі головним чином розглядається розповсюджена модель

Визначення 9. Статистикою називається довільна функція $Y = Y(X_1, \dots, X_N)$ випадкової вибірки \vec{X} .

Зауваження. Статистика $Y(X_1, \dots, X_N)$ являє собою випадкову величину. Розподіл цієї випадкової величини однозначно визначається функцією вірогідності $L_{\vec{x}}$ випадкової вибірки ймовірності або, що теж саме, щільністю ймовірності $f_1(x)$ генеральної сукупності.

Визначення 10. Вибірка називається *представницькою* (репрезентативною), якщо вона дозволяє з необхідною точністю описати властивості випадкової величини.

Зауваження 1. Для одержання репрезентативної вибірки її обсяг має бути достатньо великим, а її члени повинні вибиратися з генеральної сукупності випадково й незалежно.

Зауваження 2. По генеральній сукупності можна обчислити ймовірнісні і числові характеристики, наприклад, функцію розподілу $F_1(x)$, математичне сподівання m , дисперсію D . По конкретних реалізаціях можна обчислити оцінки тих самих характеристик, наприклад, оцінку функції розподілу $F_1^*(x)$, оцінку математичного сподівання m^* , оцінку дисперсії D^* .

Зауваження 3. Характеристики, що отримані по генеральній сукупності і що отримані по конкретних реалізаціях, відрізняються. Якщо вибірка є репрезентативною, то розбіжності в характеристиках невеликі.

6.1.2. Варіаційний ряд. Статистичний розподіл вибірки. Статистична функція розподілу

Визначення 1. Варіаційним (статистичним) рядом називається вибірка, що упорядкована за зростанням або спаданням вибірових значень x_n ($n = \overline{1, N}$).

Визначення 2. Частотою вибірових значень ω_j^* називається відношення числа N_j однакових значень x_j , що зустрічаються у виборці до обсягу N вибірки:

$$\omega_j^* = \frac{N_j}{N}.$$

Зауваження 1. Частоту вибірових значень ω_j^* можна розглядати як не випадкову величину, якщо йдеться про конкретну вибірку, і як випадкову величину (статистику випадкової вибірки \bar{X}), якщо мова йде про множину різних вибірок.

Зауваження 2. В обох випадках частота вибірових значень є аналогом ймовірності p_j того, що випадкова величина X прийме значення x_j .

Зауваження 3. Частота вибірових значень ω_j^* має такі властивості:

1) $\sum_j \omega_j^* = 1$;

2) при $N \rightarrow \infty$ частота $\omega_j^* \rightarrow p_j$.

Визначення 3. Статистичним розподілом вибірки називається перелік вибірових значень x_j і відповідних їм частот ω_j^* : (x_j, ω_j^*) , $j = \overline{1, J}$.

Визначення 4. Полігоном частот називається ламана, що з'єднує послідовний ряд точок (x_j, ω_j^*) , $j = \overline{1, J}$ для конкретної вибірки.

Приклад 1. Розглянемо вибірку (6; 3; 5; 2; 3; 2; 5; 2; 5; 5). Варіаційний ряд, що відповідає вибірці, має такий вигляд: (2; 2; 2; 3; 3; 5; 5; 5; 5; 6).

Статистичним розподілом вибірки є пари чисел $(2; \frac{3}{10})$, $(3; \frac{1}{5})$, $(5; \frac{4}{10})$, $(6; \frac{1}{10})$.

Відповідний полігон частот зображений на рис. 6.1.1.

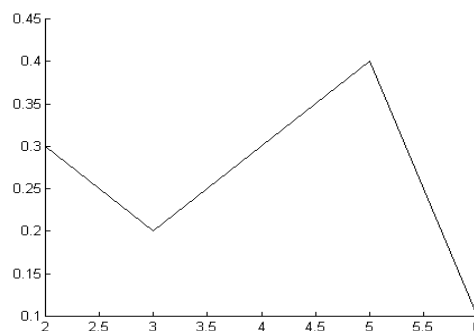


Рис. 6.1.1. Приклад полігона частот

Зауваження. Будувати статистичний розподіл вибірки має сенс лише для вибірки дискретних випадкових величин, що приймають невелике число J значень x_j . Для неперервних випадкових величин, а також для дискретних випадкових величин, які приймають велику кількість значень, використовують інтервальний статистичний розподіл вибірки, що базується на понятті розряду.

Визначення 5. Розрядами називаються піддіапазони значень випадкової величини X , що не перекриваються.

Зауваження 1. При розподілі діапазону значень випадкової величини X на розряди слід враховувати, що кількість розрядів не повинна бути занадто великою. Практика показує, що в більшості випадків доцільно вибрати число розрядів порядку 10 – 20.

Зауваження 2. Розряди можуть бути різної довжини, однак для спрощення розрахунків часто рекомендується використовувати розряди однакової довжини.

У тих випадках, коли випадкова величина розподілена дуже нерівномірно, доцільно вибрати розряди неоднакової довжини: робити їх більш вузькими в області найбільшої щільності і менш вузькими в області найменшої щільності.

Зауваження 3. Щоб уникнути непорозумінь, треба чітко визначитись, до якого розряду відносити значення, що потрапили на стик розрядів.

Визначення 6. Частотою вибіркового розподілу p_i^* , відповідного i -му розряду, називають відношення кількості N_i значень вибірки, що потрапили в

i -й розряд, до обсягу N вибірки: $p_i^* = \frac{N_i}{N}$.

Визначення 7. Інтервальним статистичним розподілом вибірки називається перелік розрядів і частот, відповідних цим розрядам.

Визначення 8. Щільністю частоти f_i^* вибіркового розподілу називають

відношення частоти p_i^* i -го розряду до його довжини Δx_i : $f_i^* = \frac{p_i^*}{\Delta x_i}$.

Зауваження 1. Як і частота вибірових значень, частота і щільність частоти вибіркового розподілу можуть розглядатися як невинпадкові величини, якщо йдеться про конкретну вибірку, і як випадкові величини, якщо мова йде про випадкову вибірку.

Зауваження 2. Щільність частоти f_i^* являє собою оцінку щільності ймовірностей генеральної сукупності.

Визначення 9. Гістограмою розподілу називається стовпчаста діаграма, яка використовується для графічного зображення інтервальних розподілів конкретної вибірки. Основи стовпців представляють розряди, а висота кожного i -го стовпця дорівнює відповідній частоті p_i^* .

Зауваження. Для зображення інтервальних статистичних розподілів вибірки звичайно використовують гістограми. Значно рідше з цією метою використовують полігони.

Приклад 2. При випробуванні висотоміра проведено 100 вимірів величини помилки. Результати вимірів оформлені у вигляді статистичного ряду (табл. 6.1.1), гістограми розподілу (рис. 6.1.2) і полігона (рис. 6.1.3).

Таблиця 6.1.1

Приклад інтервального статистичного розподілу

Розряд	[-4,5:- 3,5]	(-3,5:- 2,5]	(-2,5:- 1,5]	(-1,5:- 0,5]	(-0,5: 0,5]	(0,5: 1,5]	(1,5: 2,5]	(2,5: 3,5]
p_j^*	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020

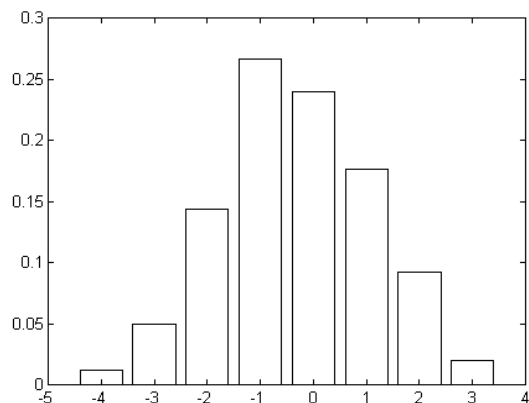


Рис. 6.1.2. Приклад гістограми розподілу

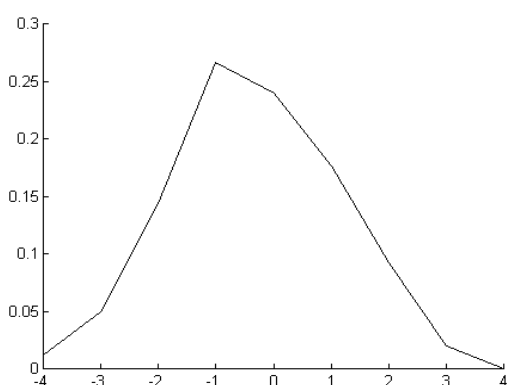


Рис. 6.1.3. Приклад полігона інтервального розподілу

Визначення 10. *Статистичною функцією розподілу називається випадкова функція $F^*(x)$, яка визначає для кожного значення x випадкової величини X частоту події $X \leq x$.*

Зауваження 1. Статистична функція розподілу $F^*(x)$ є аналогом функції розподілу $F(x)$, що в теорії математичної статистики звичайно називається *істинною функцією розподілу або теоретичною*.

Зауваження 2. Статистична функція розподілу $F^*(x)$ для конкретної вибірки графічно може бути представлена у вигляді неспадаючої ступінчастої функції аргументу x .

Зауваження 3. Властивості статистичної функції розподілу $F^*(x)$ такі самі, як і теоретичної функції розподілу.

Визначення 11. *Інтервальною статистичною функцією розподілу називається випадкова функція $F^*(x)$, яка визначає для кожного розряду частоту події $X \leq x$.*

Приклад графічного зображення реалізації інтервальної статистичної функції розподілу показано на рис. 6.1.4.

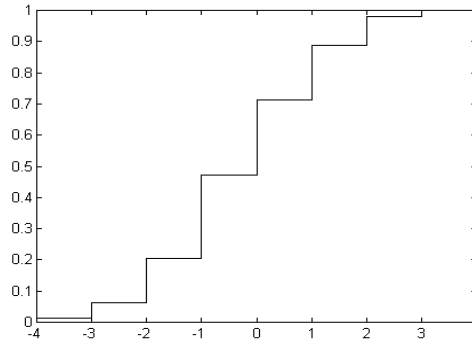


Рис. 6.1.4. Приклад інтервальної статистичної функції розподілу для конкретної вибірки (дані відповідають прикладу 2)

6.1.3. Моменти вибіркового розподілу

Визначення 1. *Вибірковим початковим моментом порядку ν випадкової вибірки $\vec{X} = (X_1, \dots, X_N)$ обсягу N називається випадкова величина m_ν^* (статистика), яка визначається за формулою*

$$m_\nu^* = \frac{1}{N} \sum_{n=1}^N X_n^\nu. \quad (6.1.2)$$

Визначення 2. *Вибірковим середнім значенням випадкової вибірки $\vec{X} = (X_1, \dots, X_N)$ (вибірковим математичним сподіванням) називається випадкова величина*

$$m_x^* = \frac{1}{N} \sum_{n=1}^N X_n. \quad (6.1.3)$$

Визначення 3. *Вибірковим центральним моментом порядку ν випадкової вибірки $\vec{X} = (X_1, \dots, X_N)$ обсягу N називається випадкова величина*

$$\mu_\nu^* = \frac{1}{N} \sum_{n=1}^N (X_n - m_x^*)^\nu. \quad (6.1.4)$$

Визначення 4. *Вибірковою дисперсією випадкової вибірки $\vec{X} = (X_1, \dots, X_N)$ називається випадкова величина*

$$D_x^* = \frac{1}{N} \sum_{n=1}^N (X_n - m_x^*)^2. \quad (6.1.5)$$

6.2. Закон великих чисел

Ствердження. При необмеженому зростанні обсягу вибірки N статистичні аналоги ймовірнісних і моментних характеристик випадкових величин і випадкових функцій збігаються за ймовірністю до відповідних точних характеристик. Це твердження є теоретичним обґрунтуванням правомірності використання вибіркового методу в математичній статистиці. Воно ґрунтується на законі великих чисел, який розглядається далі.

На даний час відомі кілька форм цього закону. Зупинимося на деяких із них.

6.2.1. Нерівність Чебишева

Доведемо лему, що називається *нерівністю Чебишева*.

Лема. Нехай випадкова величина X має математичне сподівання $m_x < \infty$ і дисперсію $D_x < \infty$. Тоді ймовірність того, що абсолютна величина відхилення $|X - m_x|$ буде більшою за будь-яке наперед задане додатне число ε , не більше $\frac{D_x}{\varepsilon^2}$, тобто

$$P(|X - m_x| > \varepsilon) \leq \frac{D_x}{\varepsilon^2}. \quad (6.2.1)$$

Доведення. Нехай випадкова величина X описується функцією розподілу $F(x)$. Тоді її дисперсія

$$D_x = \int_{-\infty}^{\infty} (x - m_x)^2 dF(x).$$

Виключимо з області інтегрування область, яка відповідає нерівності $|x - m_x| < \varepsilon$. Оскільки підінтегральний вираз додатний, маємо

$$D_x \geq \int_{|x - m_x| > \varepsilon} (x - m_x)^2 dF(x).$$

В області інтегрування $(x - m_x)^2 > \varepsilon^2$. Тому

$$D_x \geq \varepsilon^2 \int_{|x - m_x| > \varepsilon} dF(x) = \varepsilon^2 P(|x - m_x| > \varepsilon),$$

звідки випливає співвідношення (6.2.1).

6.2.2. Теорема Чебишева

Теорема. Нехай X_1, \dots, X_N – попарно незалежні випадкові величини з математичними сподіваннями m_{x_1}, \dots, m_{x_N} й обмеженими дисперсіями D_{x_1}, \dots, D_{x_N} ($D_{x_n} \leq C < \infty$, $n = \overline{1, N}$). Тоді для будь-якого як завгодно малого додатного ε середньоарифметичне Y_N випадкових величин X_1, \dots, X_N збігається за ймовірністю до середньоарифметичного m_{y_N} їх математичних сподівань m_{x_1}, \dots, m_{x_N} :

$$\lim_{N \rightarrow \infty} P(|Y_N - m_{y_N}| > \varepsilon) = 0. \quad (6.2.2)$$

Доведення. Нехай виконуються умови теореми. Тоді математичне сподівання m_{y_N} і дисперсія D_{y_N} середньоарифметичного $Y_N = \frac{1}{N} \sum_{n=1}^N X_n$ можуть бути записані таким чином:

$$m_{y_N} = \frac{1}{N} \sum_{n=1}^N m_{x_n};$$

$$D_{y_N} = \frac{1}{N^2} \sum_{n=1}^N D_{x_n}.$$

Оскільки $D_{x_n} \leq C$, то $D_{y_N} \leq \frac{C}{N}$.

Використовуючи до випадкової величини Y_N нерівність Чебишева, маємо

$$P\left(|Y_N - m_{y_N}| > \varepsilon\right) \leq \frac{D_{y_N}}{\varepsilon^2} \leq \frac{C}{N\varepsilon^2}.$$

При $N \rightarrow \infty$ з останньої нерівності випливає вираз (6.2.2).

Наслідок. Середньоарифметичне вибірових значень випадкової величини X збігається за ймовірністю до її математичного сподівання.

Доведення. Якщо X_1, \dots, X_N є вибіровими значеннями випадкової величини X , то математичні сподівання цих величин рівні. Тоді вираз (6.2.2) може бути записаний таким чином:

$$\lim_{N \rightarrow \infty} P\left(\left|\frac{1}{N} \sum_{n=1}^N X_n - m_x\right| > \varepsilon\right) = 0.$$

Зауваження. Аналогічним чином доводиться збіжність за ймовірністю інших вибірових моментів.

6.2.3. Теорема Бернуллі

Теорема. Нехай проводиться серія незалежних випробувань, в кожному з яких може або з'явитися або не з'явитися деяка подія A . Ймовірність появи події A стала і дорівнює p_a . Тоді при необмеженому збільшенні числа випробувань N частота $\frac{N_a}{N}$ появи події A збігається за ймовірністю до її ймовірності p_a :

$$\lim_{N \rightarrow \infty} P\left(\left|\frac{N_a}{N} - p_a\right| > \varepsilon\right) = 0. \quad (6.2.3)$$

Доведення. Розглянемо незалежні випадкові величини X_1, \dots, X_N , які представляють число появи події в кожному випробуванні. Ці величини можуть приймати два значення: 1 – з ймовірністю p_a і 0 – з ймовірністю $q_a = (1 - p_a)$.

Математичне сподівання кожної з величин X_n дорівнює $m_{x_n} = p_a$, а дисперсія

$$D_{x_n} = (1 - p_a)^2 p_a + (0 - p_a)^2 q_a = p_a q_a < \infty.$$

Середньоарифметичне випадкових величин X_1, \dots, X_N являє собою частоту $\frac{N_a}{N}$ появи події A . Тоді, на підставі теореми Чебишева, справедлива рівність (6.2.3).

Відомості про книги до розділу 6 наведені в табл. 6.2.1.

Таблиця 6.2.1

Література до розділу 6

Характер книг	Номери книг у списку літератури
Довідники	3, 11, 24, 47, 49, 70, 93, 108
Книги навчального плану для інженерів	13, 16, 32, 42, 80, 91
Книги з математичним ухилом	5, 21, 22, 45, 46, 50, 59, 61, 90, 101, 105
Книги прикладного характеру	1, 28, 31, 52, 55, 56, 62, 65, 66, 71, 84, 97, 106, 107

7. Оцінка параметрів і законів розподілу випадкової величини

7.1. Поняття про точкову оцінку параметрів розподілу

У попередньому розділі розглядалися питання визначення ймовірнісних характеристик випадкової величини, коли апіорна інформація про її властивості взагалі відсутня. На практиці часто доводиться стикатися з іншою ситуацією: є деякі відомості (або припущення) про властивості випадкової величини і треба, використовуючи дані вибірки, конкретизувати (уточнити) ці відомості. Класична задача такого роду наведена нижче.

Постановка завдання. Є випадкова вибірка $X_1, \dots, X_N = \vec{X}$ обсягу N випадкової величини X з відомою структурою (виглядом) розподілу $F(x/\theta_1, \dots, \theta_K)$ або $F(x; \theta_1, \dots, \theta_K)$ і невідомими у конкретній реалізації вибірки значеннями параметрів $\theta_1, \dots, \theta_K$. Необхідно, використовуючи реалізацію x_1, \dots, x_N вибірки X_1, \dots, X_N , визначити значення невідомих параметрів $\theta_1, \dots, \theta_K$.

Зауваження 1. Задача з розподілом $F(x/\theta_1, \dots, \theta_K)$ відповідає випадку, коли відсутня апіорна інформація щодо параметрів $\theta_1, \dots, \theta_K$, а з розподілом $F(x; \theta_1, \dots, \theta_K)$ – випадку, коли така інформація існує.

Зауваження 2. Параметри, для яких відсутня апіорна інформація, іноді розглядають як випадкові параметри з рівномірним законом розподілу.

Визначення. Точковою оцінкою або просто оцінкою $\vec{\Theta}^*$ параметра $\vec{\theta} = (\theta_1, \dots, \theta_K)$ називається статистика (тобто функція вибірки), значення $\vec{\theta}^*$ якої розглядаються як приблизне значення параметра $\vec{\theta}$.

Сформульована задача зводиться до знаходження оцінок $\Theta_1^*, \dots, \Theta_K^*$, які при заданій структурі розподілу $F(x/\theta_1, \dots, \theta_K)$ або $F(x; \theta_1, \dots, \theta_K)$ є функціями $\alpha_1, \dots, \alpha_K$ вибірки X_1, \dots, X_N :

$$\begin{aligned}\Theta_1^* &= \alpha_1(X_1, \dots, X_N); \\ &\dots\dots\dots \\ \Theta_K^* &= \alpha_K(X_1, \dots, X_N).\end{aligned}$$

Зауваження 1. Якщо розглядається випадкова вибірка \vec{X} , то результат оцінки $\Theta_1^*, \dots, \Theta_K^*$ – випадкові величини; якщо розглядається реалізація \vec{x} випадкової вибірки, то результат оцінки $\theta_1^*, \dots, \theta_K^*$ – не випадкові величини.

Зауваження 2. Для одержання оцінок можуть бути використані різні статистики. Оцінки, одержані при використанні різних статистик, відрізняються одна від одної.

Зауваження 3. Поряд із точковими оцінками часто використовують також інтервальні оцінки, що розглянуті в розділі 7.5.

7.2. Властивості оцінок параметрів розподілу

7.2.1. Зміщені й незміщені оцінки

Визначення 1. Оцінка $\vec{\Theta}^*$ фіксованого параметра $\vec{\theta}$ називається *незміщеною*, якщо математичне сподівання умовної випадкової величини $\vec{\Theta}^*/\vec{\theta}$ за сукупністю вибірок будь-якого кінцевого обсягу дорівнює параметру, що оцінюється:

$$M[\bar{\Theta}^* / \bar{\theta}] = \bar{\theta}. \quad (7.2.1)$$

У протилежному випадку оцінка називається *зміщеною*. *Зміщення* (*систематична похибка*) оцінки визначається за формулою

$$\bar{\varepsilon}_0 = M[\bar{\Theta}^* / \bar{\theta}] - \bar{\theta}.$$

Примітка. У загальному випадку систематична похибка залежить від параметра $\bar{\theta}$, що оцінюється, і не залежить від випадкової вибірки \bar{X} .

Зауваження 1. Точність оцінки параметра $\bar{\theta}$ часто характеризують за допомогою величини середнього квадрата похибки

$$\Delta^2 = M[(\bar{\Theta}^* - \bar{\theta})^2 / \bar{\theta}].$$

У скалярному випадку цю величину можна представити таким чином:

$$\Delta^2 = \sigma^2 + \varepsilon_0^2, \quad (7.2.2)$$

де $\sigma^2 = M[(\Theta^* - M[\Theta^* / \theta])^2 / \theta]$ – дисперсія оцінки.

Звідси видно, що величина Δ^2 складається з двох компонент – дисперсії оцінки σ^2 , що характеризує випадкові властивості оцінки, та квадрата систематичної похибки ε_0^2 .

Зауваження 2. У загальному випадку факт зміщення оцінки $\bar{\Theta}^*$ залежить від закону розподілу випадкової величини X . При одному законі оцінка може бути зміщеною, при іншому – незміщеною.

Зауваження 3. Факт зміщення оцінки залежить також від того, чи відомі априорно інші параметри розподілу, наприклад, математичне сподівання, дисперсія. Оцінки, незміщені при наявності таких априорних відомостей, можуть виявитись зміщеними за відсутності таких відомостей, і навпаки.

Деякі приклади незміщених оцінок скалярних параметрів наведені в табл. 7.2.1.

Таблиця 7.2.1

Приклади незміщених оцінок параметрів розподілу

Параметр	Незміщена оцінка для випадку, коли математичне сподівання	
	невідомо	відомо
m_x	$m_x^* = \frac{1}{N} \sum_{n=1}^N X_n$	m_x
D_x	$D_x^* = \frac{1}{N-1} \sum_{n=1}^N (X_n - m_x^*)^2$	$D_x^* = \frac{1}{N} \sum_{n=1}^N (X_n - m_x)^2$
σ_x (при гауссівському розподілі X)	$\sigma_x^* = k_N \sqrt{\frac{1}{N-1} \sum_{n=1}^N (X_n - m_x^*)^2}$	$\sigma_x^* = k_{N+1} \sqrt{\frac{1}{N} \sum_{n=1}^N (X_n - m_x)^2}$
R_{xy}	$R_{xy}^* = \frac{1}{N-1} \sum_{n=1}^N (X_n - m_x^*)(Y_n - m_y^*)$	$R_{xy}^* = \frac{1}{N} \sum_{n=1}^N (X_n - m_x)(Y_n - m_y)$

Примітка. Коефіцієнти k_N близькі до одиниці. Вони визначаються співвідношенням

$$k_N = \sqrt{\frac{N-1}{2}} \frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)}, \text{ де } \Gamma(*) \text{ – гама-функція (додаток 2). Значення коефіцієнтів } k_N \text{ наведені}$$

в табл. 7.2.2.

Таблиця 7.2.2

Значення коефіцієнтів k_N

N	k_N	N	k_N	N	k_N	N	k_N	N	k_N
3	1,1284	6	1,0506	12	1,0230	25	1,0104	40	1,0064
4	1,0853	7	1,0423	15	1,0181	30	1,0087	45	1,0056
5	1,0640	10	1,0280	20	1,0134	35	1,0072	50	1,0051

Визначення 2. Оцінка $\bar{\Theta}^*$ параметра $\bar{\theta}$ називається *асимптотично незміщеною*, якщо математичне сподівання умовної випадкової величини $\bar{\Theta}^* / \bar{\theta}$ за сукупністю вибірок обсягу $N \rightarrow \infty$ дорівнює параметру, що оцінюється:

$$\lim_{N \rightarrow \infty} M[\bar{\Theta}^* / \bar{\theta}] = \bar{\theta}. \quad (7.2.3)$$

7.2.2. Спроможні оцінки

Визначення. Оцінка $\bar{\Theta}^*$ фіксованого параметра $\bar{\theta}$ називається *спроможною*, якщо вона збігається за ймовірністю до цього параметра:

$$\lim_{N \rightarrow \infty} P\{|\bar{\Theta}^* - \bar{\theta}| > \varepsilon / \bar{\theta}\} = 0, \quad (7.2.4)$$

де N – об'єм вибірки, $\varepsilon > 0$ – завгодно мале число.

Зауваження 1. Спроможність оцінки означає, що вона відповідає закону великих чисел (теоремі Чебишева). Достатньою умовою виконання співвідношення (7.2.4) є $\lim_{N \rightarrow \infty} M\left[\frac{|\bar{\Theta}^* - \bar{\theta}|^2}{\bar{\theta}}\right] = 0$.

Зауваження 2. Спроможність є асимптотичною характеристикою оцінки, яка спостерігається при $N \rightarrow \infty$.

7.2.3. Ефективні оцінки

Важливою характеристикою оцінки є її ефективність. Для фіксованого та випадкового параметрів поняття ефективності оцінки визначається по-різному.

Розглянемо спочатку поняття ефективності оцінки для скалярного фіксованого параметра.

Визначення 1. Оцінка Θ_e^* називається *ефективною оцінкою* скалярного фіксованого параметра θ , якщо математичне сподівання квадрата відхилення оцінки Θ_e^* від параметра θ за сукупністю вибірок заданого обсягу N менше, ніж для будь-яких інших оцінок θ_i^* :

$$M[(\Theta_e^* - \theta)^2 / \theta] < M[(\Theta_i^* - \theta)^2 / \theta], \quad i=1, 2, \dots \quad (7.2.5)$$

Зауваження 1. Величина $M[(\Theta^* - \theta)^2 / \theta]$ в загальному випадку не є дисперсією оцінки $D[\Theta^* / \theta]$. Вона дорівнює дисперсії лише для незміщених оцінок.

Зауваження 2. Для незміщених оцінок умова ефективності може бути записана у вигляді

$$D[\Theta_e^* / \theta] < D[\Theta_i^* / \theta], \quad i=1, 2, \dots \quad (7.2.6)$$

Зауваження 3. Ефективність оцінки, як і незміщеність, залежить від наявності апріорних даних про розподіли випадкової величини X і від вигляду цього розподілу. Тому, говорячи про ефективність оцінки, треба зазначити умови, коли вона ефективна.

Визначення 2. Відносною ефективністю оцінки називається параметр l , який визначається як відношення математичного сподівання квадрата відхилення ефективної оцінки Θ_e^* за умови θ до математичного сподівання квадрата відхилення оцінки Θ^* за умови θ :

$$l = \frac{M[(\Theta_e^* - \theta)^2 / \theta]}{M[(\Theta^* - \theta)^2 / \theta]} \quad (7.2.7)$$

Зауваження. Параметр відносної ефективності належить інтервалу $[0,1]$. У випадку, коли оцінка ефективна, $l = 1$.

Визначення 3. Оцінка Θ^* називається *асимптотично ефективною*, якщо при нескінченному зростанні обсягу вибірки N параметр її відносної ефективності прямує до одиниці.

Визначення 4. Відносною ефективністю $l'(\theta)$ оцінки Θ_1^* щодо оцінки Θ_2^* називається відношення математичних сподівань квадратів їх відхилень за умови θ :

$$l'(\theta) = \frac{M[(\Theta_1^* - \theta)^2 / \theta]}{M[(\Theta_2^* - \theta)^2 / \theta]} \quad (7.2.8)$$

Зауваження. Параметр $l'(\theta)$ відносної ефективності оцінок дозволяє порівнювати ефективність різних оцінок, уникаючи знаходження ефективних оцінок.

Нижня межа дисперсії $D[\Theta^* / \theta]$ оцінки Θ^* визначається за допомогою теореми, що наведена нижче.

Теорема 1. Нехай по вибірці \vec{x} обсягу N випадкового вектора \vec{X} оцінюється скалярний фіксований параметр θ . При цьому межі області визначення умовної щільності ймовірностей $f_N(\vec{x} / \theta)$ не залежать від θ , ця щільність ймовірностей абсолютно інтегровна по \vec{x} і подвійно диференційована по θ й для оцінки Θ^* / θ існують перші два моменти. Тоді нижня межа дисперсії задовольняє нерівності

$$D[\Theta^* / \theta] = M[(\Theta^* - M[\Theta^* / \theta])^2 / \theta] \geq \frac{\left(1 + \frac{\partial \varepsilon_0}{\partial \theta}\right)^2}{M\left[\left(\frac{\partial \ln f_N(\vec{X} / \theta)}{\partial \theta}\right)^2 / \theta\right]}, \quad (7.2.9)$$

що називається *нерівністю Рао – Крамера* (Крамера – Рао), де D – оператор дисперсії; M – оператор математичного сподівання, який у лівій частині нерівності діє на параметр Θ^* , а у правій частині – на вектор \vec{X} ; ε_0 – систематична похибка оцінки.

Доведення. Розглянемо математичне сподівання

$$M[\Theta^* / \theta] = \int_{-\infty}^{\infty} \theta^*(\bar{x}) f_N(\bar{x} / \theta) d\bar{x} . \quad (7.2.10)$$

У загальному випадку зміщеної оцінки

$$\int_{-\infty}^{\infty} \theta^*(\bar{x}) f_N(\bar{x} / \theta) d\bar{x} = \theta + \varepsilon_0 . \quad (7.2.11)$$

Продиференціюймо вираз (7.2.11) по θ :

$$\int_{-\infty}^{\infty} \theta^*(\bar{x}) \frac{\partial f_N(\bar{x} / \theta)}{\partial \theta} d\bar{x} = 1 + \frac{\partial \varepsilon_0}{\partial \theta} . \quad (7.2.12)$$

Зауважимо що

$$\frac{\partial f_N(\bar{x} / \theta)}{\partial \theta} = f_N(\bar{x} / \theta) \frac{\partial \ln f_N(\bar{x} / \theta)}{\partial \theta} . \quad (7.2.13)$$

Тоді вираз (7.2.12) може бути записаний так:

$$\int_{-\infty}^{\infty} \theta^*(\bar{x}) \frac{\partial \ln f_N(\bar{x} / \theta)}{\partial \theta} f_N(\bar{x} / \theta) d\bar{x} = 1 + \frac{\partial \varepsilon_0}{\partial \theta}$$

або як

$$M \left[\Theta^* \frac{\partial \ln f_N(\vec{X} / \theta)}{\partial \theta} / \theta \right] = 1 + \frac{\partial \varepsilon_0}{\partial \theta} . \quad (7.2.14)$$

Розглянемо тотожність

$$\int_{-\infty}^{\infty} f_N(\bar{x} / \theta) d\bar{x} = 1 . \quad (7.2.15)$$

Продиференціюймо вираз (7.2.15) по θ :

$$\int_{-\infty}^{\infty} \frac{\partial f_N(\bar{x} / \theta)}{\partial \theta} d\bar{x} = 0 . \quad (7.2.16)$$

Враховуючи рівність (7.2.13), вираз (7.2.16) набуває вигляду

$$\int_{-\infty}^{\infty} \frac{\partial \ln f_N(\bar{x} / \theta)}{\partial \theta} f_N(\bar{x} / \theta) d\bar{x} = 0 ,$$

тобто

$$M \left[\frac{\partial \ln f_N(\vec{X} / \theta)}{\partial \theta} / \theta \right] = 0 . \quad (7.2.17)$$

Домножимо рівність (7.2.17) на $M[\Theta^* / \theta]$:

$$M \left[M[\Theta^* / \theta] \frac{\partial \ln f_N(\vec{X} / \theta)}{\partial \theta} / \theta \right] = 0 . \quad (7.2.18)$$

Відніmemo з виразу (7.2.14) вираз (7.2.18):

$$M\left[\left(\Theta^* - M[\Theta^* / \theta]\right) \frac{\partial \ln f_N(\bar{X} / \theta)}{\partial \theta} / \theta\right] = 1 + \frac{\partial \varepsilon_0}{\partial \theta}. \quad (7.2.19)$$

Приймаючи до уваги нерівність Коши – Шварца, після тривіальних перетворень отримуємо формулу (7.2.9).

Визначення 5. Інформацією за Фішером називається величина

$$J_N = M\left[\left(\frac{\partial \ln f_N(\bar{X} / \theta)}{\partial \theta}\right)^2 / \theta\right]. \quad (7.2.20)$$

Зауваження 1. Інформацію за Фішером можна представити іншим чином:

$$J_N = -M\left[\frac{\partial^2 \ln f_N(\bar{X} / \theta)}{\partial \theta^2} / \theta\right]. \quad (7.2.21)$$

В цьому неважко переконатися. Дійсно, продиференціюємо вираз (7.2.13):

$$\frac{\partial^2 f_N(\bar{x} / \theta)}{\partial \theta^2} = \frac{\partial f_N(\bar{x} / \theta)}{\partial \theta} \frac{\partial \ln f_N(\bar{x} / \theta)}{\partial \theta} + f_N(\bar{x} / \theta) \frac{\partial^2 \ln f_N(\bar{x} / \theta)}{\partial \theta^2}.$$

Скористаємося рівністю (7.2.13):

$$\frac{\partial^2 f_N(\bar{x} / \theta)}{\partial \theta^2} = f_N(\bar{x} / \theta) \left(\frac{\partial \ln f_N(\bar{x} / \theta)}{\partial \theta}\right)^2 + f_N(\bar{x} / \theta) \frac{\partial^2 \ln f_N(\bar{x} / \theta)}{\partial \theta^2}.$$

Проінтегруймо цей вираз по \bar{x} . Враховуючи рівність (7.2.16), отримуємо

$$\int_{-\infty}^{\infty} \left(\frac{\partial \ln f_N(\bar{x} / \theta)}{\partial \theta}\right)^2 f_N(\bar{x} / \theta) d\bar{x} = - \int_{-\infty}^{\infty} \frac{\partial^2 \ln f_N(\bar{x} / \theta)}{\partial \theta^2} f_N(\bar{x} / \theta) d\bar{x},$$

тобто

$$M\left[\left(\frac{\partial \ln f_N(\bar{x} / \theta)}{\partial \theta}\right)^2\right] = -M\left[\frac{\partial^2 \ln f_N(\bar{x} / \theta)}{\partial \theta^2}\right].$$

Враховуючи рівність (7.2.20), маємо вираз (7.2.21).

Зауваження 2. З урахуванням виразу (7.2.20) нерівність Рао – Крамера можна записати таким чином:

$$D[\Theta^* / \theta] \geq \frac{\left(1 + \frac{\partial \varepsilon_0}{\partial \theta}\right)^2}{J_N}. \quad (7.2.22)$$

Зауваження 3. Для незміщеної оцінки

$$D[\Theta^* / \theta] \geq \frac{1}{J_N}. \quad (7.2.23)$$

Зауваження 4. Оскільки мінімум середнього квадрата похибки $M\left[(\Theta^* - \theta)^2 / \theta\right]$ має місце при $M[\Theta^* / \theta] = \theta$, то нерівність Рао – Крамера може бути записана у вигляді подвійної нерівності:

$$M\left[(\Theta^* - \theta)^2 / \theta\right] \geq D[\Theta^* / \theta] \geq \frac{\left(1 + \frac{\partial \varepsilon_0}{\partial \theta}\right)^2}{J_N}. \quad (7.2.24)$$

Як бачимо, нижня межа квадрата відхилення оцінки Θ^* від параметра θ й дисперсії оцінки Θ^* співпадають.

Зауваження 5. З виразу (7.2.24) видно, що для забезпечення нульової дисперсії величина $\frac{\partial \varepsilon_0}{\partial \theta}$ повинна дорівнювати (-1). Звідси випливає, що неможливо одноразово забезпечити нульове зміщення та нульову дисперсію.

Зауваження 6. Для однорідної незалежної вибірки нерівність Рао – Крамера має вигляд:

$$M\left[(\Theta^* - \theta)^2 / \theta\right] \geq D[\Theta^* / \theta] \geq \frac{\left(1 + \frac{\partial \varepsilon_0}{\partial \theta}\right)^2}{N J_1}, \quad (7.2.25)$$

де J_1 – інформація за Фішером:

$$J_1 = -M\left[\frac{\partial^2 \ln f_1(X/\theta)}{\partial \theta^2} / \theta\right].$$

Нерівність (7.2.25) слідує з виразів (7.2.21), (7.2.24).

Зауваження 7. Замість наведеного вище визначення ефективної оцінки (див. визначення 1) часто використовують інше визначення, що базується на нерівності Рао – Крамера.

Визначення 6а. Ефективною оцінкою Θ_e^* називають оцінку, для якої математичне сподівання

$$M\left[(\Theta_e^* - \theta)^2 / \theta\right] = \frac{\left(1 + \frac{\partial \varepsilon_0}{\partial \theta}\right)^2}{J_N}. \quad (7.2.26)$$

Теорема 2. Оцінка θ_e^* ефективна у розумінні визначення 6а тоді і тільки тоді, коли вона не зміщена та має наступну структуру:

$$\theta_e^* = \theta + \frac{1}{J_N} \frac{\partial \ln f_N(\bar{x}/\theta)}{\partial \theta}. \quad (7.2.27)$$

Враховуючи теорему 2, можна дати інше визначення ефективності оцінки, що еквівалентно визначенню 6а.

Визначення 6б. Ефективною оцінкою Θ_e^* називають оцінку, що задовольняє системі рівнянь

$$\begin{cases} M[\Theta_e^* / \theta] = \theta, \\ M\left[(\Theta_e^* - \theta)^2 / \theta\right] = D[\Theta_e^* / \theta] = \frac{1}{J_N} \end{cases} \quad (7.2.28)$$

Зауваження 1. У загальному випадку визначення 1 та 6 не еквівалентні.

Зауваження 2. Якщо ефективна оцінка за визначенням 6 завжди задовольняє нерівності (7.2.24), то ефективна оцінка за визначенням 1 не завжди відповідає рівності (7.2.26).

Зауваження 3. Коли не існує ефективної за визначенням 6 оцінки, вираз (7.2.24) характеризує не потенційну точність оцінки, а нижню межу точності оцінки.

Зауваження 4. З визначення 6б випливає, що ніяка зміщена оцінка не є ефективною, але треба пам'ятати, що коли існує ефективна оцінка, то існують зміщені оцінки, в яких дисперсія менша, ніж у ефективної оцінки.

Зауваження 5. Не зважаючи на те, що зміщена оцінка може мати меншу дисперсію, при наявності ефективної оцінки її майже не використовують.

Причина в тому, що зменшення дисперсії похибки σ^2 супроводжується збільшенням систематичної складової похибки ε_0 , при цьому (див. формулу (7.2.2)) середній квадрат похибки Δ^2 виявляється не меншим середнього квадрата похибки ефективної оцінки.

Зауваження 6. Ефективна за визначенням оцінка існує тоді і тільки тоді, коли функція $\frac{\partial \ln f_N(\bar{x}/\theta)}{\partial \theta}$ задовольняє рівності (7.2.27), тобто коли умовна щільність ймовірностей $f_N(\bar{x}/\theta)$ належить до експоненційного сімейства функцій:

$$f_N(\bar{x}/\theta) = \exp[k_1(\theta)\theta_e^* + k_2(\theta)]h(\bar{x}),$$

де $k_1(\theta)$ і $k_2(\theta)$ – деякі функції параметра θ , а $h(\bar{x})$ – деяка функція вектора \bar{x} .

Зауваження 7. Для неефективних оцінок відомі межі, які більш точні, ніж ті, що дає нерівність Рао – Крамера. Такими межами є, наприклад, межа Баттачари, що використовує похідні високого порядку, а також межа Баранкіна. Розрахунок цих та інших подібних меж викликає значні труднощі. Тому їх використовують дуже рідко.

Поняття ефективності узагальнюються на векторні параметри. Розглянемо оцінку $\bar{\theta}^*$ K -вимірною векторного фіксованого параметра $\bar{\theta}$.

Визначення 7. Інформаційною матрицею Фішера називається матриця I з елементами

$$I_{ij} = M \left[\frac{\partial \ln f_N(\bar{X}/\bar{\theta})}{\partial \theta_i} \frac{\partial \ln f_N(\bar{X}/\bar{\theta})}{\partial \theta_j} \right] / \bar{\theta} \quad (i, j = \overline{1, K}), \quad (7.2.29)$$

що визначаються для K -вимірною параметра $\bar{\theta}$ за допомогою N -вимірної умовної щільності ймовірностей $f_N(\bar{X}/\bar{\theta})$ N -вимірною випадкового вектора \bar{X} .

Зауваження 1. Можна показати, що

$$\begin{aligned} & M \left[\frac{\partial \ln f_N(\bar{X}/\bar{\theta})}{\partial \theta_i} \frac{\partial \ln f_N(\bar{X}/\bar{\theta})}{\partial \theta_j} \right] / \bar{\theta} = \\ & = -M \left[\frac{\partial^2 \ln f_N(\bar{X}/\bar{\theta})}{\partial \theta_i \partial \theta_j} \right] / \bar{\theta} \quad (i, j = \overline{1, K}). \end{aligned}$$

Тому вираз (7.2.29) може бути записаний інакше:

$$I_{ij} = -M \left[\frac{\partial^2 \ln f_N(\bar{X}/\bar{\theta})}{\partial \theta_i \partial \theta_j} \right] / \bar{\theta} \quad (i, j = \overline{1, K}). \quad (7.2.30)$$

Зауваження 2. Якщо N -вимірний випадковий вектор \bar{X} описується гауссівським розподілом із математичними сподіваннями $\bar{S}(\bar{\theta})$ і кореляційною матрицею R , то елементи інформаційної матриці Фішера

$$I_{ij} = \frac{\partial \bar{S}^T(\bar{\theta})}{\partial \theta_i} R^{-1} \frac{\partial \bar{S}(\bar{\theta})}{\partial \theta_j} \quad (i, j = \overline{1, K}). \quad (7.2.31)$$

Нижня межа кореляційної матриці помилок

$$R_{\bar{\theta}^*} = M[(\bar{\Theta}^* - \bar{\theta})(\bar{\Theta}^* - \bar{\theta})^T / \bar{\theta}] \quad (7.2.32)$$

визначається теоремою, що узагальнює теорему 1.

Теорема 3. Нехай по вибірці \vec{x} випадкового вектора \vec{X} оцінюється векторний параметр $\vec{\theta} = (\theta_1, \dots, \theta_K)$. При цьому існує обернена матриця D розміром $K \times K$ з елементами

$$d_{ij} = 1 + \frac{\partial}{\partial \theta_j} \varepsilon_{0i}(\vec{\theta}) \quad (i, j = \overline{1, K}),$$

де $\varepsilon_{0i}(\vec{\theta})$ ($i = \overline{1, K}$) – компоненти вектора зміщення $\vec{\varepsilon}_0(\vec{\theta})$. Тоді

$$U^T R_{\vec{\theta}^*}^{-1} U \leq U^T D I D^{-1} U, \quad (7.2.33)$$

де U – K -вимірний допоміжний вектор.

Зауваження 1. Нерівність (7.2.33) – одна з форм узагальненої нерівності Рао – Крамера.

Зауваження 2. Для незміщеної оцінки нерівність (7.2.33) набуває вигляду

$$U^T R_{\vec{\theta}^*}^{-1} U \leq U^T I U. \quad (7.2.34)$$

Зауваження 3. Нерівність (7.2.34) аналітично виражає той факт, що нижня межа еліпсоїда розсіювання $U^T R_{\vec{\theta}^*}^{-1} U = 1$ довільної незміщеної оцінки $\vec{\Theta}^*$ цілком знаходиться у еліпсоїді розсіювання $U^T I U = 1$, що визначається інформаційною матрицею Фішера.

Визначення 8а. Незміщенні оцінки $\Theta_{1e}^*, \dots, \Theta_{Ke}^*$, що складають вектор $\vec{\Theta}_e^*$, називають *спільно ефективними*, якщо кореляційна матриця помилок $R_{\vec{\theta}_e^*}^{-1}$ задовольняє рівності

$$U^T R_{\vec{\theta}_e^*}^{-1} U = U^T I U. \quad (7.2.35)$$

Зауваження 1. У разі, коли рівність (7.2.35) має місце тільки при обсязі вибірки $N \rightarrow \infty$, оцінки $\Theta_{1e}^*, \dots, \Theta_{Ke}^*$ називають *спільно асимптотично ефективними*.

Зауваження 2. Спільно ефективні оцінки, що задовольняють рівності (7.2.35), існують не завжди. Необхідною та достатньою умовами їх існування є можливість представлення вектора $\vec{\theta}_e^* = (\theta_{1e}^*, \dots, \theta_{Ke}^*)$ спільно ефективних оцінок таким чином:

$$\vec{\theta}_e^* = \vec{\theta} + I^{-1} b^T, \quad (7.2.36)$$

де b – вектор із компонентами $\frac{\partial \ln f_N(\vec{x} / \vec{\theta})}{\partial \theta_i}$ ($i = \overline{1, K}$).

Зауваження 3. Ефективна оцінка залишається ефективною при довільних лінійних перетвореннях. При нелінійних перетвореннях ефективність оцінки не зберігається.

Зауваження 4. У разі, коли не існують спільно ефективні оцінки, що задовольняють рівності (7.2.35), іноді використовують інше поняття спільно ефективних оцінок.

Визначення 8б. Оцінки $\Theta_{1e}^*, \dots, \Theta_{Ke}^*$, що складають вектор $\vec{\Theta}_e^*$ називають *спільно ефективними*, якщо при тому же обсязі вибірки N для всіх інших оцінок $\Theta_{1i}^*, \dots, \Theta_{Ki}^*$ вектора $\vec{\Theta}_i^*$ має місце така нерівність

$$U^T R_{\vec{\theta}_e^*}^{-1} U > U^T R_{\vec{\theta}_i^*}^{-1} U \quad (i = 1, 2, \dots). \quad (7.2.37)$$

Перейдемо до розгляду поняття ефективності оцінки для випадкового параметра.

Теорема 4. Нехай по вибірці \vec{x} обсягу N оцінюється скалярний випадковий параметр θ , який описується щільністю ймовірності $f_1(\theta)$. При цьому $(N + 1)$ –

вимірна щільність ймовірностей $f_{N+1}(\bar{x}, \theta)$ подвійно диференційована по θ , $\frac{\partial f_{N+1}(\bar{x}; \theta)}{\partial \theta}$ та $\frac{\partial^2 f_{N+1}(\bar{x}; \theta)}{\partial \theta^2}$ абсолютно інтегровні по \bar{x} і θ , а

$$\lim_{\theta \rightarrow \pm\infty} \int_{-\infty}^{\infty} (\theta^* - \theta) f_{N+1}(\bar{x}; \theta) d\bar{x} = 0 .$$

Тоді

$$M[(\Theta^* - \Theta)^2] \geq J_N^{-1}, \quad (7.2.38)$$

де J_N – інформація за Фішером, що визначається виразом:

$$\begin{aligned} J_N &= M \left[\left(\frac{\partial \ln f_{N+1}(\bar{X}; \Theta)}{\partial \Theta} \right)^2 \right] = \\ &= -M \left[\frac{\partial^2 \ln f_{N+1}(\bar{X}; \Theta)}{\partial \Theta^2} \right], \end{aligned} \quad (7.2.39)$$

у лівій частині нерівності (7.2.38) оператор математичного сподівання M діє на оцінку Θ^* та параметр Θ , а у виразі (7.2.39) – на \bar{X} та Θ .

Зауваження 1. Нерівність (7.2.38) являє собою нерівність Рао – Крамера, що описує нижню межу середнього квадрата помилки оцінки Θ^* випадкового параметра Θ .

Зауваження 2. Формули (7.2.38), (7.2.39) узагальнюються на випадок K – вимірного векторного випадкового параметра $\vec{\Theta}$. В цьому разі нерівність Рао – Крамера набуває вигляду

$$U^T R_{\vec{\theta}^*}^{-1} U \leq U^T I U, \quad (7.2.40)$$

де $R_{\vec{\theta}^*}$ – кореляційна матриця помилок: розміром $K \times K$

$$R_{\vec{\theta}^*} = M[(\vec{\Theta}^* - \vec{\Theta})(\vec{\Theta}^* - \vec{\Theta})^T], \quad (7.2.41)$$

I – інформаційна матриця Фішера з елементами

$$I_{ij} = M \left[\frac{\partial \ln f_{N+K}(\bar{X}; \vec{\Theta})}{\partial \Theta_i} \frac{\partial \ln f_{N+K}(\bar{X}; \vec{\Theta})}{\partial \Theta_j} \right], \quad (7.2.42)$$

де у виразі (7.2.41) оператор математичного сподівання діє на оцінку $\vec{\Theta}^*$ та параметр $\vec{\Theta}$, а у виразі (7.2.42) – на \bar{X} та $\vec{\Theta}$.

Зауваження 3. Для векторних випадкових параметрів також використовують поняття *спільно ефективних та спільно асимптотично ефективних оцінок* (див. визначення 8а, зауваження 1 до визначення 8а та визначення 8б).

Зауваження 4. Необхідною та достатньою умовами існування ефективної оцінки випадкового параметра $\vec{\Theta}$ за визначенням 8а є гауссівський розподіл апостеріорної щільності ймовірностей $f_K(\vec{\theta}; \bar{x})$.

Зауваження 5. При лінійних перетвореннях ефективність оцінки зберігається, а при нелінійних – ні.

7.2.4. Достатні оцінки

Визначення. Оцінка $\bar{\theta}^*$ фіксованого параметра $\bar{\theta}$ називається *достатньою*, якщо N -вимірна умовна щільність ймовірностей $f_N(x_1, \dots, x_N / \bar{\theta}^*)$ вибірки випадкової величини X не залежить від параметра $\bar{\theta}$.

Зауваження 1. З наведеного визначення випливає, що у разі достатньої оцінки, щільність ймовірностей $f_N(\bar{x} / \bar{\theta}^*)$ не містить ніякої інформації про параметр $\bar{\theta}$, тобто достатня оцінка несе всю корисну інформацію про параметр $\bar{\theta}$, яка зосереджена у вибірці.

Приклад. Оцінка математичного сподівання $m_x^* = \frac{1}{N} \sum_{n=1}^N x_n$ для вибірки обсягом N достатня, а оцінка $m_x^* = \frac{1}{N-1} \sum_{n=1}^{N-1} x_n$ для тієї ж вибірки недостатня.

Зауваження 2. Якщо оцінка ефективна, то вона обов'язково достатня. Зворотнє твердження невірно.

Зауваження 3. Достатні оцінки існують не завжди.

Зауваження 4. Якщо достатня статистика існує, то вона визначається неоднозначно. Довільна статистика, що пов'язана з достатньою статистикою взаємно однозначно залежністю, теж достатня.

Теорема. Необхідною й достатньою умовами достатності оцінки $\bar{\theta}^*$ параметра $\bar{\theta} = (\theta_1, \dots, \theta_K)$ є можливість представлення умовної щільності ймовірностей $f_N(x_1, \dots, x_N / \bar{\theta})$ у наступному факторизованому вигляді:

$$f_N(x_1, \dots, x_N / \theta) = f_1(\theta^* / \theta) h(x_1, \dots, x_N), \quad (7.2.43)$$

де $f_1(\theta^* / \theta)$ – щільність ймовірностей оцінки θ^* , яка одержана за вибіркою \bar{x} , $h(x_1, \dots, x_N)$ – довільна функція вибірки, яка не залежить від параметра θ .

Зауваження. З теореми випливає наступне твердження: якщо $\bar{\theta}^*$ – достатня оцінка параметра $\bar{\theta}$, то апостеріорна щільність ймовірностей $f_K(\bar{\theta} / \bar{x})$ залежить не від самих вибіркових значень \bar{x} , а від оцінки $\bar{\theta}^*(\bar{x})$.

Властивості деяких оцінок параметрів розподілу випадкової величини наведені в табл. 7.2.3.

Таблиця 7.2.3

Властивості деяких оцінок параметрів розподілу випадкової величини X при відсутності апіорної інформації про її математичне сподівання

Параметр, що оцінюється	Статистика, що використовується для оцінки параметра	Властивості оцінки			
		незміщенність	спроможність	ефективність (при гауссівському розподілі)	достатність
Математичне сподівання m_x	$m_x^* = \frac{1}{N} \sum_{n=1}^N X_n$	+	+	+	+
Дисперсія D_x	$D_x^* = \frac{1}{N} \sum_{n=1}^N (X_n - m_x^*)^2$	-	+	$-(l_1)$	+
	$D_x^* = \frac{1}{N-1} \sum_{n=1}^N (X_n - m_x^*)^2$	+	+	$-(l_2 < l_1)$	+

Ймовірність випадкової події A	Частота $p^* = \frac{N_A}{N}$	+	+	+	+
----------------------------------	----------------------------------	---	---	---	---

Примітка. При цьому більш ефективною є зміщена оцінка ($l_2 < l_1$).

7.3. Методи отримання оцінок параметрів розподілу випадкових величин

7.3.1. Метод моментів

Ідея методу. Нехай щільність ймовірностей $f_1(x; \theta)$ деякої випадкової величини X залежить від одного невідомого параметра θ , що підлягає оцінці. Будемо вважати, що параметр θ не є моментом випадкової величини X . Визначимо для випадкової величини, що досліджується, початковий момент ν -го порядку:

$$m_\nu = \int_{-\infty}^{\infty} x^\nu f(x; \theta) dx.$$

Цей момент можна розглядати як детерміновану функцію ψ_ν параметра θ :

$$m_\nu = \psi_\nu(\theta). \quad (7.3.1)$$

Розв'язавши рівняння (7.3.1) відносно θ , знайдемо

$$\theta = \psi_\nu^{-1}(m_\nu), \quad (7.3.2)$$

де ψ_ν^{-1} – функція, що є оберненою до ψ_ν .

Використаємо як оцінку m_ν відповідний вибірковий момент і підставимо його у вираз (7.3.2). Тоді оцінка Θ^* параметра розподілу може бути визначена так:

$$\Theta^* = \psi_\nu^{-1}\left(\frac{1}{N} \sum_{n=1}^N X_n^\nu\right).$$

Зауваження 1. Для знаходження оцінки параметра розподілу випадкової величини можна використовувати аналогічним чином і центральні моменти.

Зауваження 2. Описаний метод знаходження оцінки Θ^* невідомого параметра θ випадкової величини X легко узагальнюється на багатовимірний випадок, коли необхідно одержати оцінки $\Theta_1^*, \dots, \Theta_K^*$ одночасно K параметрів $\theta_1, \dots, \theta_K$. В цьому випадку моменти m_1, \dots, m_K K різних порядків можуть бути записані у вигляді системи рівнянь

$$\begin{cases} m_1 = \psi_1(\theta_1, \dots, \theta_K); \\ \dots \\ m_K = \psi_K(\theta_1, \dots, \theta_K). \end{cases} \quad (7.3.3)$$

Розв'язавши систему рівнянь відносно параметрів, що оцінюються, і, замінивши моменти m_1, \dots, m_K відповідними оцінками m_1^*, \dots, m_K^* , маємо

$$\begin{cases} \Theta_1^* = \phi_1(m_1^*, \dots, m_K^*); \\ \dots \\ \Theta_K^* = \phi_K(m_1^*, \dots, m_K^*). \end{cases}$$

Зауваження 3. Оцінки, які отримуються з використанням описаного методу, залежать від вибору моментів. Доведено, що всі ці оцінки справжні, коли

моменти існують та функції Φ_1, \dots, Φ_k – раціональні і приймають скінченні значення.

7.3.2. Метод максимуму вірогідності

Ідея методу. Нехай відома N -вимірна щільність ймовірностей $L_{\bar{X}}(\vec{\theta})$ (функція вірогідності) випадкової вибірки X_1, \dots, X_N , яка залежить від вектора параметрів $\vec{\theta}$, що оцінюються. В якості оцінки $\vec{\theta}^*$ параметра $\vec{\theta}$ можна вибрати таке значення $\vec{\theta}$, при якому функція вірогідності набуває максимального значення.

Визначення. Оцінкою максимальної вірогідності називається така оцінка $\vec{\theta}^*$ параметра $\vec{\theta}$ розподілу випадкової величини X , при якій функція вірогідності приймає максимальне значення.

Розглянемо особливості обчислення максимально вірогідної оцінки. Врахуємо, що через незалежність компонент випадкової вибірки функція вірогідності $L_{\bar{X}}(\vec{\theta})$ може бути представлена у вигляді добутку щільностей ймовірностей компонент:

$$L_{\bar{X}}(\vec{\theta}) = \prod_{n=1}^N f_1(x_n / \vec{\theta}). \quad (7.3.4)$$

Прологарифмуємо обидві частини виразу (7.3.4). В результаті маємо

$$\ln L_{\bar{X}}(\vec{\theta}) = \sum_{n=1}^N \ln f_1(x_n / \vec{\theta}). \quad (7.3.5)$$

Натуральний логарифм – монотонно зростаюча функція. Тому точки екстремумів функції (7.3.4) і функції (7.3.5) збігаються.

Тоді шукана оцінка $\vec{\theta}^*$ максимуму функції вірогідності може бути визначена із системи рівнянь

$$\begin{cases} \sum_{n=1}^N \frac{\partial \ln f_1(x_n / \vec{\theta})}{\partial \theta_1} = 0; \\ \dots \\ \sum_{n=1}^N \frac{\partial \ln f_1(x_n / \vec{\theta})}{\partial \theta_k} = 0. \end{cases} \quad (7.3.6)$$

Зауваження 1. Система рівнянь (7.3.6) може мати кілька розв'язків. Причому деяким розв'язкам відповідають максимуми, а деяким – мінімуми функції вірогідності. Серед усієї цієї множини розв'язків необхідно відібрати розв'язок, який би відповідав максимум-максимуму функції вірогідності (максимальному з усіх максимальних її значень).

Зауваження 2. Метод максимальної вірогідності може бути використаний і для знаходження оцінок невідомих параметрів розподілу дискретної випадкової величини. В цьому випадку система рівнянь (7.3.6) записується відносно не щільності ймовірності, а ймовірності.

Властивості. Оцінки максимальної вірогідності мають такі властивості:

- 1) оцінка максимальної вірогідності завжди спроможна;
- 2) якщо для заданого обсягу вибірки N існує ефективна оцінка, то – це оцінка максимальної вірогідності;
- 3) оцінка максимальної вірогідності завжди є асимптотично ефективною й асимптотично незміщеною;

1) коли апіорна щільність ймовірностей фактично не додає інформації про параметр, що оцінюється (наприклад, коли при оцінці математичного сподівання випадкової величини X апостеріорна щільність ймовірностей математичного сподівання задається у вигляді величини, яка рівномірно розподілена на інтервалі, що розглядається);

2) коли інформація, яка використовується, є невірною (наприклад, зроблено припущення, що закон розподілу гауссівський, а в дійсності він істотно відрізняється від такого).

7.4. Закони розподілу оцінок параметрів гауссівського розподілу

Одним з основних параметрів, що характеризують оцінку, є її точність.

Визначення. В одновимірному випадку як *кількісну міру точності* точкової оцінки звичайно використовують або величину $M[(\Theta^* - \theta)^2]$ (для незміщеної оцінки – дисперсію $D[\Theta^*]$), або ймовірність $P(|\theta^* - \theta| \leq \varepsilon)$ того, що абсолютне відхилення оцінки Θ^* від істинного значення θ не перевищує деяку задану величину ε .

У багатовимірному випадку звичайно використовують матрицю похибок R_{θ^*} (для незміщеної оцінки – кореляційну матрицю похибок).

Зауваження. У будь-якому випадку для розрахунку точності необхідно знати закон розподілу оцінки.

Розглянемо методику визначення закону розподілу оцінки і точнісних характеристик на найпростішому прикладі оцінки математичного сподівання гауссівської випадкової величини в умовах, коли дисперсія випадкової величини X відома точно.

Приклад. Нехай випадкова величина X має гауссівський розподіл із параметрами m_x , σ_x^2 . Дисперсія σ_x^2 апіорно відома, а математичне сподівання m_x – ні. Величина математичного сподівання оцінюється по вибірці обсягу N таким чином:

$$m_x^* = \frac{1}{N} \sum_{n=1}^N X_n.$$

Треба визначити: щільність ймовірностей оцінки $f_1(m_x^*)$, дисперсію оцінки $D[m_x^*]$ та ймовірність $P(|m_x^* - m_x| \leq \varepsilon)$.

Розв'язання. Випадкова величина X має гауссівський розподіл. Тому оцінка m_x^* математичного сподівання, яка являє собою середнє арифметичне вибірових значень X_n , теж має гауссівський розподіл. Математичне сподівання цього розподілу $M[m_x^*] = m_x$. Враховуючи, що X_n незалежні, неважко розрахувати дисперсію оцінки: $D[m_x^*] = \frac{D_x}{N}$.

Для гауссівського розподілу ймовірність $P(|m_x^* - m_x| \leq \varepsilon)$ може бути отримана в такому вигляді:

$$P(|m_x^* - m_x| \leq \varepsilon) = 2F\left(\varepsilon \sqrt{\frac{N}{D_x}}\right) - 1,$$

де $F(x)$ – табулірована гауссівська функція розподілу (додаток 7).

Деякі результати, що стосуються оцінок різних параметрів при різному обсязі апіорної інформації, наведені в табл. 7.4.1. В цій таблиці є посилання на χ^2 -розподіл із N ступенями свободи й розподіл Стюдента (t -розподіл). Короткі описи цих розподілів наведені у додатках 5 і 6. Відповідні таблиці ймовірностей

$P_{\chi^2}(\chi^2 > \chi^2_\gamma)$ і $P(T > t_\beta) = \int_{t_\beta}^{\infty} S(t, N) dt$, корисні при розрахунках, дані в додатках 8 і 9.

Зауваження до підрозділу. Щоб визначити точність точкової оцінки Θ^* , як правило, необхідно знати істинне значення параметра θ , який оцінюється. Якщо ця інформація відсутня, не можна розрахувати ні середній квадрат похибки $M[(\Theta^* - \theta)^2]$, ні ймовірність $P(|\theta^* - \theta| \leq \varepsilon)$. Неможливість розрахунку точності точкових оцінок в умовах, коли значення параметрів невідомі, є істотним недоліком таких оцінок. Позбавлені цього недоліку інтервальні оцінки, які розглядаються в наступному підрозділі.

Таблиця 7.4.1

Оцінки параметрів гауссівського розподілу по вибірці випадкової величини X обсягу N

Оцінка	Умови		Формули, що описують оцінку, тип розподілу оцінки й параметри розподілу оцінки
	відомі	невідомі	
$m_x^* \sqrt{\frac{N}{D_x}}$	D_x	m_x	1) $m_x^* = \frac{1}{N} \sum_{n=1}^N X_n$ 2) величина $m_x^* \sqrt{\frac{N}{D_x}}$ має гауссівський розподіл 3) $M\left[m_x^* \sqrt{\frac{N}{D_x}}\right] = m_x \sqrt{\frac{N}{D_x}}$ 4) $D\left[m_x^* \sqrt{\frac{N}{D_x}}\right] = 1$ 5) $P\left(\left m_x^* - m_x\right \sqrt{\frac{N}{D_x}} \leq \varepsilon\right) = 2F(\varepsilon) - 1$
$\frac{D_x^* N}{D_x}$	m_x	D_x	1) $D_x^* = \frac{1}{N} \sum_{n=1}^N (X_n - m_x)^2$ 2) величина $\frac{D_x^* N}{D_x}$ має χ^2 -розподіл із N ступенями свободи 3) $M\left[\frac{D_x^* N}{D_x}\right] = N$ 4) $D\left[\frac{D_x^* N}{D_x}\right] = 2N$ 5) $P\left(\left \frac{D_x^* N}{D_x} - N\right \leq \varepsilon\right) = P_{\chi^2}(\chi^2 > \chi^2_1) - P_{\chi^2}(\chi^2 > \chi^2_2)$, де $\chi^2_{1,2} = (N \mp \varepsilon)$

$\frac{D_x^* N}{D_x}$	-	m_x, D_x	<p>1) $D_x^* = \frac{1}{N-1} \sum_{n=1}^N (X_n - m_x^*)^2$, де</p> $m_x^* = \frac{1}{N} \sum_{n=1}^N X_n$ <p>2) величина $\frac{D_x^* N}{D_x}$ має χ^2-розподіл з $(N-1)$ ступенями свободи</p> <p>3) $M\left[\frac{D_x^* N}{D_x}\right] = N-1$</p> <p>4) $D\left[\frac{D_x^* N}{D_x}\right] = 2(N-1)$</p>
-----------------------	---	------------	---

Закінчення табл. 7.4.1

			<p>5) $P\left(\left \frac{D_x^* N}{D_x} - (N-1)\right \leq \varepsilon\right) = P_{\chi^2}(\chi^2 > \chi_1^2) - P_{\chi^2}(\chi^2 > \chi_2^2)$, де $\chi_{1,2}^2 = ((N-1) \mp \varepsilon)$</p>
$t^* = \frac{m_x^* - m_x}{\sqrt{D_x^*}} \sqrt{N}$	-	m_x, D_x	<p>1) $m_x^* = \frac{1}{N} \sum_{n=1}^N X_n$</p> $D_x^* = \frac{1}{N-1} \sum_{n=1}^N (X_n - m_x^*)^2$ <p>2) величина t^* має розподіл Стьюдента (t-розподіл) з $(N-1)$ ступенями свободи</p> <p>3) $M[t^*] = 0$</p> <p>4) $D[t^*] = \frac{N-1}{N-3}$</p> <p>5) $P(t^* \leq \varepsilon) = 1 - 2 \int_{\varepsilon}^{\infty} S(t, N-1) dt$</p>

7.5. Інтервальне оцінювання

7.5.1. Загальні поняття про інтервальне оцінювання

Постановка задачі. Нехай для параметра θ є незміщена оцінка Θ^* . Позначимо γ ймовірність $P(|\theta^* - \theta| \leq \varepsilon)$ того, що абсолютне відхилення оцінки Θ^* від істинного значення параметра θ буде не більше деякої заданої величини ε . Уявимо величину γ таким чином:

$$\gamma = P(\theta^* - \varepsilon \leq \theta \leq \theta^* + \varepsilon). \quad (7.5.1)$$

Цей вираз визначає ймовірність знаходження істинного значення параметра θ в інтервалі $I_\gamma = (\Theta^* - \varepsilon, \Theta^* + \varepsilon)$. Середня точка цього інтервалу визначається випадковою величиною – оцінкою Θ^* . Тому місцезнаходження інтервалу I_γ на осі θ є випадковою величиною. Іншими словами, інтервал I_γ –

випадковий. При цьому ймовірність γ можна трактувати як ймовірність того, що випадковий інтервал I_γ накріє точку θ (рис. 7.5.1).

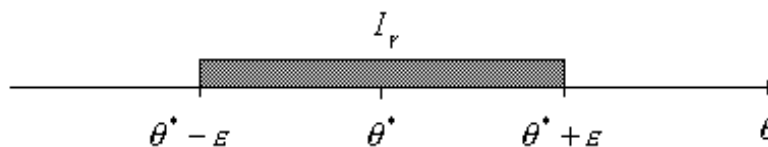


Рис. 7.5.1. Довірчий інтервал

Визначення 1. Довірчим інтервалом (інтервальною оцінкою) параметра θ , що оцінюється, називається випадковий інтервал I_γ із фіксованою тривалістю 2ϵ і середньою точкою, яка визначається випадковою оцінкою Θ^* :

$$I_\gamma = [\Theta^* - \epsilon, \Theta^* + \epsilon]. \quad (7.5.2)$$

Зауваження. Випадковою величиною може бути не тільки місцезнаходження довірчого інтервалу на осі параметрів, але також і його довжина, тобто величина 2ϵ . Приклади таких ситуацій розглянуті в розділі 7.5.3.

Визначення 2. Межі довірчого інтервалу називаються довірчими межами (нижньою й верхньою довірчими межами).

Визначення 3. Довірчою ймовірністю (коефіцієнтом довіри) називається ймовірність γ того, що довірчий інтервал (7.5.2) накріє параметр θ , що оцінюється, тобто значення параметра θ попаде до інтервалу I_γ :

$$\gamma = P(\theta \in I_\gamma).$$

7.5.2. Наближений метод розрахунку

меж довірчого інтервалу за заданою довірчою ймовірністю

Постановка завдання. Нехай для параметра θ є незміщена оцінка Θ^* . Якщо закон розподілу Θ^* відомий, то для знаходження по заданій довірчій ймовірності відповідних меж довірчого інтервалу I_γ необхідно розв'язати рівняння

$$P(|\theta^* - \theta| \leq \epsilon) = \gamma \quad (7.5.3)$$

відносно ϵ .

Ідея методу. Складність розв'язання завдання полягає в тому, що часто закон розподілу Θ^* залежить від ряду невідомих параметрів, в тому числі і від самого параметра θ , що оцінюється. Для подолання цих труднощів інколи замінюють точні значення невідомих параметрів їх точковими оцінками.

Зауваження. Заміна точних значень невідомих параметрів їх точковими оцінками, як правило, дає добрий результат, але тільки в тому випадку, коли кількість дослідів достатньо велика (не менша за кілька десятків). Коли ж число дослідів мало (до десяти – двадцяти), використання цього прийому небажано.

Приклад. Розглянемо методику застосування приблизного методу розрахунку меж довірчого інтервалу за заданою довірчою ймовірністю γ на прикладі математичного сподівання m_x випадкової величини X , яка має довільний закон розподілу.

Як оцінку m_x^* математичного сподівання m_x будемо використовувати середнє арифметичне вибірових значень:

$$m_x^* = \frac{1}{N} \sum_{n=1}^N X_n.$$

Вважаючи, що обсяг вибірки N досить великий, можна вважати, що оцінка m_x^* має гауссівський розподіл. Математичне сподівання цієї оцінки $M[m_x^*] = m_x$, а дисперсія $D[m_x^*] = \frac{D_x}{N}$.

Замінімо невідому дисперсію D_x випадкової величини X її оцінкою:

$$D_x^* = \frac{1}{N-1} \sum_{n=1}^N (X_n - m_x^*)^2.$$

Тоді, використовуючи функцію гауссівського розподілу $F(x)$, рівняння (7.5.3) можна записати у вигляді

$$2F\left(\varepsilon \sqrt{\frac{N}{D_x^*}}\right) - 1 = \gamma,$$

звідки

$$\varepsilon = \sqrt{\frac{D_x^*}{N}} \arg F\left(\frac{1+\gamma}{2}\right),$$

де $\arg F(x)$ – функція, обернена до $F(x)$.

Використовуючи величину $t_\gamma = \arg F\left(\frac{1+\gamma}{2}\right)$, довірчий інтервал можна записати таким чином:

$$I_\gamma = \left[m_x^* - t_\gamma \sqrt{\frac{D_x^*}{N}}; m_x^* + t_\gamma \sqrt{\frac{D_x^*}{N}} \right].$$

Аналогічно, хоча дещо складніше, розраховуються межі довірчого інтервалу для дисперсії D_x^* випадкової величини X , яка має гауссівський розподіл, і для коефіцієнта кореляції r_{xy} двох гауссівських випадкових величин X і Y . Результати розрахунків наведені в табл. 7.5.1.

Таблиця 7.5.1

Величини параметра ε , який визначає межі довірчого інтервалу (приблизний метод розрахунку меж)

Параметр, що оцінюється	Параметр ε
Математичне сподівання m_x	$\varepsilon = t_\gamma \sqrt{\frac{D_x^*}{N}}$
Дисперсія D_x	$\varepsilon = t_\gamma \sqrt{\frac{0.8N + 1.2}{N(N-1)}} D_x^*$
Коефіцієнт кореляції r_{xy}	$\varepsilon = t_\gamma \sqrt{\frac{1+r_{xy}^*}{N-1}}$

Примітка. Параметр t_γ відповідає гауссівському розподілу.

7.5.3. Точні методи побудови довірчих інтервалів для параметрів випадкової величини

Розглянутий у попередньому розділі наближений метод побудови довірчих інтервалів базується на заміні істинних невідомих параметрів розподілу випадкової величини X їх приблизними значеннями, які були отримані в результаті оцінювання.

Точні методи побудови довірчих інтервалів використовують інший підхід.

Ідея методів. Нехай відхилення $(\Theta^* - \theta)$ оцінки Θ^* від параметра θ , який оцінюється, залежить від невідомих параметрів, але при цьому відомо перетворення величини $(\Theta^* - \theta)$ в іншу величину $(\Xi^* - \xi)$, функція розподілу якої не залежить від невідомих параметрів. Тоді, визначивши для нової випадкової величини $(\Xi^* - \xi)$ межі довірчого інтервалу, за допомогою оберненого перетворення можна знайти межі довірчого інтервалу для цього параметра θ .

Розглянемо застосування цієї методики на конкретному прикладі.

Приклад. Нехай потрібно по заданій довірчій ймовірності γ розрахувати точні межі довірчого інтервалу для математичного сподівання m_x гауссівської випадкової величини X у випадку, коли оцінка математичного сподівання

$$m_x^* = \frac{1}{N} \sum_{n=1}^N X_n,$$

а дисперсія D_x випадкової величини X невідома.

Розв'язання. Величина відхилення $(m_x^* - m_x)$ залежить від невідомої дисперсії D_x і тому безпосередньо розрахувати межі довірчого інтервалу неможливо.

Перейдемо від випадкової величини $(m_x^* - m_x)$ до нормованої випадкової величини

$$(\Xi^* - \xi) = (m_x^* - m_x) \sqrt{\frac{N}{D_x^*}},$$

де D_x^* – оцінка дисперсії:

$$D_x^* = \frac{1}{N-1} \sum_{n=1}^N (X_n - m_x^*)^2.$$

Як бачимо з табл. 7.4.1, величина $(\Xi^* - \xi)$ описується розподілом Стьюдента із $N-1$ ступенями свободи. Тоді

$$P(|\xi^* - \xi| \leq t_\gamma) = 1 - 2 \int_{t_\gamma}^{\infty} S(t, N-1) dt = \gamma,$$

де $S(t, N-1)$ – щільність розподілу ймовірності Стьюдента.

Використовуючи таблицю $\beta = \int_{t_\beta}^{\infty} S(t, N-1) dt$ (додаток 9), неважко оберненим інтерполюванням обчислити величину t_γ для $\gamma = 1 - 2\beta$ і далі визначити межі інтервалу у вигляді

$$m_x^* - t_\gamma \sqrt{\frac{D_x^*}{N}} \leq m_x \leq m_x^* + t_\gamma \sqrt{\frac{D_x^*}{N}}. \quad (7.5.4)$$

Зауваження. З виразу (7.5.4) видно, що у даному випадку не тільки розташування інтервалу на осі значень параметра, але і *ширина* довірчого інтервалу є випадковою величиною.

Аналогічно обчислюються межі довірчого інтервалу для дисперсії гауссівського розподілу.

Дещо складніше з коефіцієнтом кореляції r_{xy} . Для отримання довірчого інтервалу в цьому випадку можна скористатися перетворенням Фішера, яке стосовно до оцінки коефіцієнта кореляції r_{xy}^* має такий вигляд:

$$Z^* = \frac{1}{2} \ln \frac{1+r_{xy}^*}{1-r_{xy}^*},$$

$$r_{xy}^* = \text{th} Z^*.$$

Особливість цього перетворення полягає в тому, що вже при невеликому обсязі N вибірки випадкова величина Z^* має розподіл, що близький до гауссівського розподілу з математичним сподіванням

$$m_z = \frac{1}{2} \ln \frac{1+r_{xy}}{1-r_{xy}} + \frac{r_{xy}}{2(N-1)}$$

і дисперсією $D_z = \frac{1}{N-3}$, що не залежить від r_{xy} .

Формули для розрахунку меж довірчих інтервалів різних параметрів гауссівського розподілу наведені в табл. 7.5.2.

Таблиця 7.5.2

Довірчі інтервали для різних параметрів гауссівської випадкової величини (точні методи розрахунку)

№ з/п	Параметр, який оцінюється	Довірчий інтервал
1	Математичне сподівання m_x	$I_\gamma = \left[m_x^* - t_\gamma \sqrt{\frac{D_x^*}{N}}, m_x^* + t_\gamma \sqrt{\frac{D_x^*}{N}} \right]$
2	Дисперсія D_x при відомому математичному сподіванні m_x	$I_\gamma = \left[\frac{D_x^* N}{\chi_2^2}, \frac{D_x^* N}{\chi_1^2} \right]$
3	Дисперсія D_x при невідомому математичному сподіванні m_x	$I_\gamma = \left[\frac{D_x^* (N-1)}{\chi_2^2}, \frac{D_x^* (N-1)}{\chi_1^2} \right]$
4	Коефіцієнт кореляції при невідомих математичних сподіваннях m_x і m_y	$I_\gamma = \left[\text{th} \left(Z^* - \frac{t_\gamma}{\sqrt{N-3}} \right), \text{th} \left(Z^* + \frac{t_\gamma}{\sqrt{N-3}} \right) \right]$

Примітки: 1. У графі 1 величина t_γ характеризує довірчу ймовірність γ для розподілу Стьюдента (при використанні таблиці з додатка 9 належить брати параметр $\beta = \frac{1-\gamma}{2}$).

2. У графі 2 величини $\chi_{1,2}^2$ – табличні значення для χ^2 -розподілу з N ступенями свободи, які відповідають ймовірностям $\frac{1-\gamma}{2}$ і $\frac{1+\gamma}{2}$ відповідно.

3. У графі 3 величини $\chi_{1,2}^2$ – табличні значення для χ^2 -розподілу з $N-1$ ступенями свободи, які відповідають ймовірностям $\frac{1-\gamma}{2}$ і $\frac{1+\gamma}{2}$ відповідно.
4. У графі 4 величина t_γ характеризує довірчу ймовірність γ для гауссівського розподілу.

7.6. Оцінка ймовірності події по частоті її появи

Постановка задачі. Є випадкова величина X , що може приймати тільки два значення: 1 – із невідомою ймовірністю p і 0 – із невідомою ймовірністю $q = 1 - p$. Проводять N незалежних дослідів. Результатом x_n кожного n -го дослідів ($n = \overline{1, N}$) є одне з двох згаданих значень випадкової величини. Значення випадкової величини в різноманітних дослідах не залежать одне від одного. Потрібно за отриманими N значеннями x_n оцінити ймовірність появи одиниці (ймовірність p).

Розв'язання 1 (знаходження точкової оцінки). Як точкову оцінку P^* ймовірності p виберемо частоту появи одиниці:

$$P^* = \frac{1}{N} \sum_{n=1}^N X_n.$$

Частота P^* є незміщеною, спроможною і ефективною оцінкою ймовірності p . Математичне сподівання й дисперсія цієї оцінки відповідно такі:

$$m_{P^*} = p, \quad (7.6.1)$$

$$\sigma_{P^*}^2 = \frac{p(1-p)}{N}. \quad (7.6.2)$$

Розв'язання 2 (знаходження довірчого інтервалу). Нехай при проведенні дослідів обидва можливих значення випадкової величини зареєстровані багато разів. Тоді частота P^* має розподіл, близький до гауссівського, із математичним сподіванням і дисперсією, що визначаються відповідно до формул (7.6.1) і (7.6.2). Тоді довірчий інтервал

$$I_\gamma = [P_1, P_2] \quad (7.6.3)$$

з межами $P_1 = P^* - \varepsilon$, $P_2 = P^* + \varepsilon$, який відповідає довірчій ймовірності γ , визначається таким виразом:

$$2F\left(\frac{\varepsilon}{\sigma_{P^*}}\right) - 1 = \gamma, \quad (7.6.4)$$

де $F(x)$ – гауссівська функція розподілу (див. додаток 7).

З виразу (7.6.4) випливає, що

$$\varepsilon = \sigma_{P^*} t_\gamma, \quad (7.6.5)$$

де $t_\gamma = \arg F\left(\frac{1+\gamma}{2}\right)$ – функція, обернена гауссівській функції розподілу $F\left(\frac{1+\gamma}{2}\right)$.

Тоді з урахуванням виразів (7.6.1) – (7.6.3), (7.6.5) можна записати нерівність

$$|P^* - p| \leq t_\gamma \sqrt{\frac{p(1-p)}{N}}. \quad (7.6.6)$$

Зводячи обидві частини нерівності (7.6.6) у квадрат і, вирішуючи їх відносно p , неважко отримати вираз

$$P_{1,2} = \frac{P^* + \frac{t_\gamma^2}{2N} \mp t_\gamma \sqrt{\frac{P^*(1-P^*)}{N} + \frac{t_\gamma^2}{4N^2}}}{1 + \frac{t_\gamma^2}{N}}, \quad (7.6.7)$$

що описує межі довірчого інтервалу I_γ .

Уявлення про характер зміни меж довірчого інтервалу дають графіки (рис. 7.6.1).

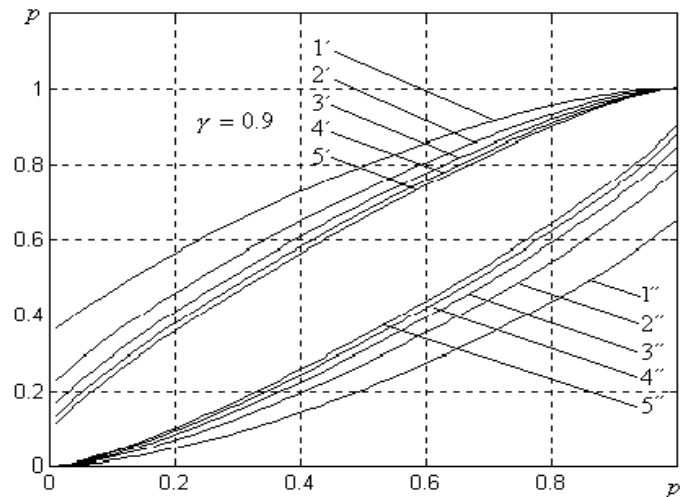


Рис. 7.6.1. Верхні (1'–5') і нижні (1''–5'') межі (P_2, P_1) довірчого інтервалу для $N = 5$ (1', 1''), $N = 10$ (2', 2''), $N = 15$ (3', 3''), $N = 20$ (4', 4'') і $N = 25$ (5', 5'')

Зауваження 1. Припущення про гауссівський характер розподілу частоти P^* обмежує можливості коректного застосування формули (7.6.7). Рекомендується використовувати її у тому випадку, коли обидва значення випадкової величини X зустрічаються у досліді не менш чотирьох разів.

Зауваження 2. Якщо обидва значення випадкової величини зустрічаються у досліді не менш десяти разів, то для розрахунку меж довірчого інтервалу можна використовувати спрощену формулу

$$P_{1,2} = P^* \mp t_\gamma \sqrt{\frac{P^*(1-P^*)}{N}}, \quad (7.6.8)$$

яка випливає з формули (7.6.7) при малих $\frac{t_\gamma^2}{N}$.

Зауваження 3. Якщо одне зі значень випадкової величини зустрічається в досліді дуже рідко (менш чотирьох разів) або кількість дослідів дуже мала, то формула (7.6.7) дає великі помилки. В цьому випадку можна застосовувати точний спосіб розрахунку [16] меж довірчого інтервалу.

Відомості про книги до розділу 7 наведені в табл. 7.6.1.

Таблиця 7.6.1

Література до розділу 7

Характер книг	Номери книг у списку літератури
Довідники	11, 24, 40, 47, 49, 81, 93, 108

Книги навчального плану для інженерів	13, 16, 32, 42, 75, 80, 84, 91
Книги з математичним ухилом	5, 21, 22, 45, 46, 50, 59, 90, 101
Книги прикладного характеру	2, 28, 43, 52, 53, 55, 56, 62, 63, 65, 76, 85, 97, 100, 102, 106, 107

8. Основи регресійного аналізу

8.1. Регресійні залежності

Реальні об'єкти, що досліджуються, пов'язані один з одним складними зв'язками. Самі об'єкти та зв'язки між ними можуть бути описані за допомогою математичних моделей, які дають деяке наближене уявлення про реальний стан об'єктів та реальні процеси, що протікають. Для такого приблизного опису застосовують різні моделі: детерміновані, випадкові та змішані (що включають як детерміновані, так і випадкові елементи).

Положення 1. Розділ математичної статистики, який присвячений теорії наближеного опису одних *випадкових об'єктів* за допомогою інших *детермінованих* та *випадкових* об'єктів, носить назву *регресійного аналізу*.

Зауваження 1. У регресійному аналізі є два основних напрямки. Перший пов'язаний із приблизним описом випадкових величин за допомогою інших випадкових величин, другий – із приблизним описом випадкових функцій за допомогою інших випадкових та детермінованих функцій.

Зауваження 2. Регресійний аналіз випадкових величин часто використовується у тому випадку, коли випадкова величина, що досліджується, чи зовсім не доступна до безпосереднього виміру, чи хоча й доступна, але її вимірювання складно реалізувати на практиці. При цьому, однак, може бути проведений відносно простий вимір інших випадкових величин, пов'язаних якимось чином із величиною, що досліджується.

Визначення 1. У регресійному аналізі (як випадкових величин, так й випадкових функцій) величини, що вимірюються, називаються *факторами*.

Примітка до визначення 1. У залежності від конкретної задачі та контексту фактор може бути незалежною змінною, випадковою величиною або значенням випадкової величини. У подальшому, для позначення факторів, які є випадковими величинами, будемо використовувати великі літери, а для позначання інших видів факторів – малі літери.

Приклад. Знос кілець двигуна й потужність його шумовипромінювання являють собою випадкові величини, складним чином пов'язані одна з одною. Наявність такого зв'язку дозволяє судити про стан кілець по результатах виміру потужності шумовипромінювання (у даному випадку цей параметр виступає як фактор).

Зауваження 3. Регресійний аналіз випадкових величин може бути представлено як двоетапну процедуру. Перший етап – *побудова (чи уточнення) регресійної моделі*, яка приблизно описує функціональну залежність випадкової величини, що досліджується, від факторів: випадкових величин, що безпосередньо вимірюються. Другий етап – *використання регресійної моделі* для прогнозу значень випадкової величини, що досліджується, за отриманими значеннями факторів.

Зауваження 4. Часто етапи побудови регресійної моделі й прогнозування значень розділені й незалежні. Однак іноді, коли у процесі прогнозування отримується нова інформація, вони можуть перемешуватися.

Визначення 2. Залежність між двома випадковими величинами будемо називати *стохастичною (випадковою)*, якщо *закон розподілу* однієї випадкової величини залежить від значення, яке набула інша випадкова величина.

Визначення 3. Для двох стохастично залежних випадкових величин X , Y регресією Y на X називається будь-яка функція $\hat{y}(X)$, що приблизно описує

стохастичну залежність Y від X (рис. 8.1.1). При цьому випадкова величина Y представляється у вигляді суми двох випадкових величин $\hat{y}(X)$ і $h(X, Y)$:

$$Y = \hat{y}(X) + h(X, Y), \quad (8.1.1)$$

де $\hat{y}(X)$ – функція регресії, $h(X, Y)$ – поправковий член.

Примітка. Функція регресії $\hat{y}(X)$ і поправковий член $h(X, Y)$ розглядаються як детерміновані функції, але оскільки аргументи в них випадкові, то $\hat{y}(X)$ і $h(X, Y)$ – випадкові величини.

Визначення 4. Значення функції $\hat{y}(x)$ регресії y на x називається прогнозом або оцінкою значення y .

Визначення 5. Множина пар значень (x_n, y_n) , $n = \overline{1, N}$, стохастично залежних випадкових величин X, Y називається полем регресії, а залежність $y = \hat{y}(x)$ – кривою регресії (див. рис. 8.1.1).

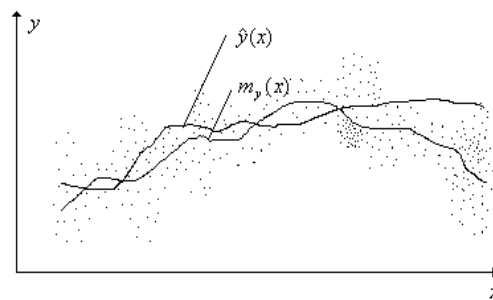


Рис. 8.1.1. Поле регресії (x_n, y_n) , $n = \overline{1, N}$, крива регресії $y = \hat{y}(x)$ і математичне сподівання $m_y(x)$ умовної випадкової величини Y/x

Зауваження 1. Поняття оцінки в теорії оцінки параметрів розподілу відрізняється від цього ж поняття в теорії регресійного аналізу. У першому випадку оцінка являє собою наближене значення параметра, яке близьке до істинного його значення. При цьому помилка наближення пов'язана з обсягом вибірки. Із збільшенням обсягу помилка зменшується аж до нуля. У теорії ж регресійного аналізу наближення пов'язане із суттю задачі. В цьому випадку отримати точну оцінку (прогноз) значення Y принципово неможливо ні за яких умов.

Зауваження 2. Щоб уникнути неоднозначності термінології, відміченої у зауваженні 1, в подальшому термін “оцінка” будемо використовувати тільки в сенсі, прийнятому в теорії оцінки параметрів розподілу. Для задач регресії будемо застосовувати термін “прогноз”.

Положення 2. На практиці використовують дві *регресійні моделі*: *детерміновану* й *стохастичну (випадкову)*. Зміст їх такий. Нехай $\hat{y}(x)$ є функція регресії $\hat{y}(x)$, що представляє залежність випадкової величини Y від випадкової величини X . В результаті дослідження отримано значення x_n ($n = \overline{1, N}$) випадкової величини X . У межах детермінованої регресійної моделі можна прогнозувати, що випадкова величина Y при цьому x_n буде приймати значення $\hat{y}(x_n)$ ($n = \overline{1, N}$). У межах стохастичної моделі можна прогнозувати, що випадкова величина Y при цьому x_n буде приймати випадкове значення $\hat{y}(x_n) + Z(x_n)$, де $Z(x_n)$ – випадкова величина з деяким законом розподілу.

Зауваження 1. Закон розподілу випадкової величини $Z(x_n)$ в ідеалі повинен дорівнювати закону розподілу умовної випадкової величини $h(x_n, Y)$. Отримати обґрунтовану інформацію про цей закон не завжди можливо. На практиці, як правило, вважають $Z(x_n)$ випадковою величиною, яка має чи гауссівський чи рівномірний закон розподілу. Слід відзначити, що для багатьох задач заміна реального закону розподілу на гауссівський чи рівномірний закон є допустимою.

Зауваження 2. На практиці найчастіше використовують детерміновану регресійну модель. Тому у подальшому розгляді будемо приділяти їй більше уваги.

Положення 3. Точність прогнозування звичайно характеризується середнім по y квадратом помилки при даному x :

$$\varepsilon(x) = M_y[(Y/x - \hat{y}(x))^2] \quad (8.1.2)$$

(тобто осередненим по y квадратом поправкового члена $h(x, Y)$), де Y/x – умовна (при даному x) випадкова величина; $M_y[*]$ – оператор обчислення математичного сподівання по параметру Y .

Примітка. У випадку стохастичної моделі для характеристики точності прогнозу іноді використовують метрики, що встановлюють ступінь різниці законів розподілу випадкових величин $\hat{y}(x_n) + Z(x_n)$ і Y/x (див. п. 9.5).

Теорема. По критерію мінімуму середнього квадрата помилки найкращий (оптимальний) прогноз забезпечується в тому випадку, коли прогнозуєме значення $\hat{y}(x)$ дорівнює математичному сподіванню $m_y(x)$ умовної випадкової величини Y/x , тобто коли $\hat{y}(x) = m_y(x)$.

Справедливість цієї теореми впливає з виразу (8.1.2), що може бути зведений до вигляду:

$$\varepsilon(x) = D_y(x) + (m_y(x) - \hat{y}(x))^2, \quad (8.1.3)$$

де $D_y(x) = D[Y/x]$ і $m_y(x) = M[Y/x]$ – відповідно умовні дисперсія і математичне сподівання умовної випадкової величини Y/x (рис. 8.1.2).

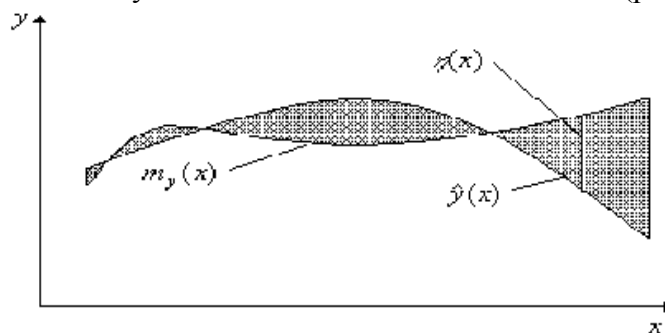


Рис. 8.1.2. Математичне сподівання $m_y(x)$ умовної випадкової величини Y/x , значення $\hat{y}(x)$, яке прогнозується, і розбіжність $\eta(x)$ між математичним сподіванням $m_y(x)$ і прогнозованим значенням $\hat{y}(x)$

Положення 4. Якщо умовне математичне сподівання $m_y(x)$ невідомо, то задача прогнозування полягає у визначенні функції регресії $\hat{y}(x)$ по обмеженій статистичній інформації про випадкові величини X і Y .

Положення 5. Звичайно функцію регресії $\hat{y}(x)$ визначають по вибірках (x_1, \dots, x_N) і (y_1, \dots, y_N) випадкових величин X і Y . Для спрощення задачі розв'язок шукають у заздалегідь обумовленому класі функцій ψ . Розв'язок $\hat{y}(x)$, який при цьому одержується, в загальному випадку не є оптимальним, бо не гарантує виконання умови $\hat{y}(x) = m_y(x)$.

Зауваження 1. Пошук розв'язку у класі заздалегідь обумовлених функцій обумовлено не тільки прагненням спростити задачу, але також тим, що часто обсяг вибірки N малий. При цьому задовільні оцінки параметрів функції регресії $\hat{y}(x)$ можуть бути отримані лише при умові, що число цих параметрів, як мінімум у декілька разів менше за обсяг вибірки N .

Зауваження 2. Підбір придатного класу функцій – складна справа. Для отримання добрих результатів звичайно розглядають декілька класів функцій: ψ_1, \dots, ψ_K . Для кожного k -го класу знаходять оптимальні параметри функції регресії $\hat{y}(x)$, розраховують помилки, що виникають при використанні цієї функції, й серед усіх отриманих таким чином функцій $\hat{y}(x)$ ($k = \overline{1, K}$) відбирають функцію $\hat{y}_{k0}(x)$, яка забезпечує мінімальну помилку.

Ясно, що така методика пошуку функції регресії не гарантує отримання оптимального розв'язку у класі довільних функцій.

Зауваження 3. Клас функцій ψ , у котрому може знаходитись прийнятний розв'язок, залежить від характеру зв'язку між випадковими величинами X і Y . Іноді, проводячи аналіз поля регресії, можливо швидко підібрати потрібний клас функцій, іноді зробити це складно. Слід відзначити, що як правило, розв'язок шукають у класі неперервних чи кусково-неперервних функцій.

Визначення 6. Середньою квадратичною регресією Y на X класу ψ називається така функція $\hat{y}(x) = \hat{y}_0(x)$ класу ψ , що забезпечує мінімум середньої помилки

$$\begin{aligned} \varepsilon &= M_x[\varepsilon(X)] = M_x[M_y[(Y/X - \hat{y}(X))^2]] = \\ &= M_x[D_y(X) + (m_y(X) - \hat{y}(X))^2], \end{aligned} \quad (8.1.4)$$

тобто мінімум осередненого по X і Y квадрата поправкового члена $h(X, Y)$.

Визначення 7. Залишковою дисперсією називається середня помилка ε_0 середньої квадратичної регресії Y на X класу ψ (тобто мінімальна середня помилка для даного класу ψ).

Зауваження 1. Залишкова дисперсія складається з двох частин: середньої умовної дисперсії $M_x[D_y(X)]$ і середнього квадрата розбіжності $M_x[\eta^2(X)]$ між умовним математичним сподіванням $m_y(X)$ і середньою квадратичною регресією $\hat{y}_0(X)$.

Зауваження 2. Середня квадратична регресія $\hat{y}_0(x)$ класу функцій ψ є найкращим у класі ψ середньоквадратичним наближенням до умовного математичного сподівання $m_y(x)$ (це прямо впливає із співвідношення (8.1.4)).

Зауваження 3. Якщо $m_y(x)$ належить класу функцій ψ , то $\hat{y}_0(x) = m_y(x)$, тобто середньоквадратична регресія є оптимальною в класі довільних функцій.

При цьому $M_x[\eta^2(X)] = 0$ і залишкова дисперсія ε_0 дорівнює середній умовній дисперсії:

$$\varepsilon_0 = M_x[D_y(X)].$$

Зауваження 4. Залишкові дисперсії для регресії Y на X і для регресії X на Y в загальному випадку не співпадають.

8.2. Лінійна середньоквадратична регресія

Визначення 1. Лінійною середньоквадратичною регресією $\hat{y}_0(x)$ називається середньоквадратична регресія класу лінійних функцій, тобто така функція

$$\hat{y}_0(x) = a_0 + b_0x, \quad (8.2.1)$$

параметри a_0, b_0 якої визначаються з умови мінімуму середньої помилки ε .

Розрахунок параметрів. Використовуючи співвідношення (8.1.4), можна записати середню помилку $\varepsilon(a, b)$ як функцію параметрів a, b у вигляді

$$\varepsilon(a, b) = M_x[D_y(X) + (m_y(X) - a - bX)^2].$$

Обчислюючи окремі похідні по a і b від цього виразу й, прирівнюючи їх до нуля, одержуємо систему двох рівнянь. Розв'язавши цю систему відносно a і b , можна отримати

$$a_0 = m_y - r_{xy} \frac{\sigma_y}{\sigma_x} m_x; \quad b_0 = r_{xy} \frac{\sigma_y}{\sigma_x},$$

де $\sigma_x = \sqrt{D_x}$; $\sigma_y = \sqrt{D_y}$; $r_{xy} = \frac{R_{xy}}{\sigma_x \sigma_y}$.

Підставивши ці значення у вираз (8.2.1), отримаємо рівняння лінійної середньоквадратичної регресії

$$\hat{y}_0(x) = m_y + r_{xy} \frac{\sigma_y}{\sigma_x} (x - m_x). \quad (8.2.2)$$

Залишкова дисперсія ε_0 лінійної середньоквадратичної регресії може бути отримана у вигляді

$$\varepsilon_0 = \sigma_y^2 (1 - r_{xy}^2). \quad (8.2.3)$$

Визначення 2. Коефіцієнт $b_0 = r_{xy} \frac{\sigma_y}{\sigma_x}$ називається *коефіцієнтом регресії*, а пряма, що визначається рівнянням (8.2.2), – *прямою середньоквадратичної регресії Y на X* .

Зауваження 1. З виразів (8.2.2), (8.2.3) видно, що в загальному випадку прямі регресії і остаточні дисперсії відрізняються для лінійних середньоквадратичних регресій Y на X і X на Y .

Зауваження 2. З виразу (8.2.3) випливає, що величина залишкової дисперсії ε_0 лінійної середньоквадратичної регресії залежить від рівня кореляції випадкових величин X і Y . При зростанні кореляції залишкова дисперсія зменшується до нуля (коли $r_{xy} = 1$). Тому можна говорити про те, що кореляція характеризує ступінь лінійної залежності між випадковими величинами.

Теорема. Для гауссівських випадкових величин X і Y оптимальною середньоквадратичною регресією Y на X класу довільних функцій є лінійна середньоквадратична регресія. При цьому залишкова дисперсія ε_0 дорівнює умовній дисперсії $D_y(x)$.

Доведення. Доведення теореми засноване на аналізі умовної щільності ймовірностей $f_1(y/x)$. Ця щільність ймовірностей може бути представлена таким чином:

$$f_1(y/x) = \frac{1}{\sigma_y \sqrt{2\pi(1-r_{xy}^2)}} \exp \left\{ -\frac{1}{2\sigma_y^2(1-r_{xy}^2)} \left[y - m_y - r_{xy} \frac{\sigma_y}{\sigma_x} (x - m_x) \right]^2 \right\}.$$

Математичне сподівання цього умовного розподілу

$$m_y(x) = m_y + r_{xy} \frac{\sigma_y}{\sigma_x} (x - m_x), \quad (8.2.4)$$

а дисперсія

$$D_y(x) = \sigma_y^2 (1 - r_{xy}^2). \quad (8.2.5)$$

Зіставляючи вирази (8.2.4), (8.2.2) і (8.2.5), (8.2.3), приходимо до рівняння $\hat{y}_0(x) = m_y(x)$, $\varepsilon_0 = D_y(x)$.

Зауваження. Оскільки для гауссівських випадкових величин X і Y умовна дисперсія $D_y(x)$ не залежить від x (однакова для будь-яких значень x), то і середній по Y квадрат помилки $\varepsilon(x)$ лінійного прогнозу при даному x (див. співвідношення (8.1.3)) не залежить від x , тобто точність прогнозу виявляється однаковою при будь-яких x .

8.3. Оцінка параметрів лінійної середньоквадратичної регресії

Параметри лінійної середньоквадратичної регресії виражаються через перші й другі моменти випадкових величин. Якщо моменти m_x , m_y , D_x , D_y , R_{xy} випадкових величин, що розглядаються, невідомі, а є лише вибірки (x_1, \dots, x_L) , (y_1, \dots, y_L) обсягу L , то параметри a_0^* і b_0^* лінійної середньоквадратичної регресії можна оцінити за допомогою відповідних моментів: вибірових математичних сподівань m_x^* , m_y^* , вибірових дисперсій $D_x = \sigma_x^{*2}$, $D_y = \sigma_y^{*2}$ і вибірової кореляції $R_{xy}^* = r_{xy}^* \sigma_x^* \sigma_y^*$. При такому наближенні вибірові параметри регресії a_0^* і b_0^* мають вигляд

$$a_0^* = m_y^* - r_{xy}^* \frac{\sigma_y^*}{\sigma_x^*} m_x^*; \quad (8.3.1)$$

$$b_0^* = r_{xy}^* \frac{\sigma_y^*}{\sigma_x^*}. \quad (8.3.2)$$

Незміщеними спроможними оцінками моментів випадкових величин є

$$m_x^* = \frac{1}{L} \sum_{l=1}^L x_l; \quad m_y^* = \frac{1}{L} \sum_{l=1}^L y_l; \quad (8.3.3)$$

$$D_x^* = \frac{1}{L-1} \sum_{l=1}^L (x_l - m_x^*)^2; \quad D_y^* = \frac{1}{L-1} \sum_{l=1}^L (y_l - m_y^*)^2; \quad (8.3.4)$$

$$R_{xy}^* = \frac{1}{L-1} \sum_{l=1}^L (x_l - m_x^*)(y_l - m_y^*). \quad (8.3.5)$$

Теорема. Якщо вибірові параметри регресії a_0^* і b_0^* обчислюються за допомогою вибірових моментів (8.3.3) – (8.3.5), то вибірові параметри є спроможними і незміщеними.

Справедливість першої частини теореми впливає із спроможності вибіркових моментів, що використовуються для оцінки параметрів регресії, а справедливість другої частини – із незміщеності цих моментів.

Зауваження. Для оцінки точності лінійного прогнозу звичайно використовують оцінку залишкової дисперсії

$$\varepsilon^*(a_0^*, b_0^*) = D_y^*(1 - r_{xy}^{*2}).$$

Ця оцінка також є незміщеною і спроможною.

8.4. Множинна регресія

Визначення 1. Множинною регресією X_1 на X_2, \dots, X_N називається будь-яка функція $\hat{y}(X_2, \dots, X_N)$, яка приблизно описує статистичну залежність $Y = X_1$ від X_2, \dots, X_N . При цьому випадкова величина X_1 представляється у вигляді суми двох випадкових величин $\hat{y}(X_2, \dots, X_N)$ і $h_1(X_1, X_2, \dots, X_N)$:

$$X_1 = \hat{y}(X_2, \dots, X_N) + h_1(X_1, X_2, \dots, X_N),$$

де $h_1(X_1, X_2, \dots, X_N)$ – поправковий член.

Положення 1. Для множинної регресії мірою точності прогнозування є середній квадрат помилки

$$\varepsilon(x_2, \dots, x_N) = M_{x_1} [(X_1 / x_2, \dots, x_N - \hat{y}(x_2, \dots, x_N))^2].$$

Теорема 1. По критерію мінімуму середнього квадрата помилки найкращий (оптимальний) прогноз забезпечується множинною регресією в тому випадку, коли значення $\hat{y}(x_2, \dots, x_N)$, що прогнозується, дорівнює математичному сподіванню $m_{x_1}(x_2, \dots, x_N)$ умовної випадкової величини $X_1 / x_2, \dots, x_N$, тобто коли

$$\hat{y}(x_2, \dots, x_N) = m_{x_1}(x_2, \dots, x_N). \quad (8.4.1)$$

Визначення 2. Середньою квадратичною регресією X_1 на X_2, \dots, X_N класу функцій ψ називається така функція $\hat{y}_0(X_2, \dots, X_N)$ з цього класу, що забезпечує мінімум середньої помилки

$$\begin{aligned} \varepsilon_0 &= M_{x_2, \dots, x_N} [\varepsilon(X_2, \dots, X_N)] = \\ &= M_{x_2, \dots, x_N} [D_{x_1}(X_2, \dots, X_N) + (m_{x_1}(X_2, \dots, X_N) - \hat{y}_0(X_2, \dots, X_N))^2], \end{aligned} \quad (8.4.2)$$

де $D_{x_1}(x_2, \dots, x_N)$ – умовна дисперсія випадкової величини X_1 .

Визначення 3. Середня помилка $\varepsilon_0 = M_{x_2, \dots, x_N} [\varepsilon(X_2, \dots, X_N)]$ середньої квадратичної регресії X_1 на X_2, \dots, X_N класу ψ називається *залишковою дисперсією множинної регресії*.

Зауваження 1. З виразу (8.4.2) випливає, що як й у одновимірному випадку, середня квадратична регресія $\hat{y}_0(x_2, \dots, x_N)$ класу функцій ψ є в класі ψ найкращим наближенням до математичного сподівання $m_{x_1}(x_2, \dots, x_N)$.

Зауваження 2. З виразів (8.4.1) і (8.4.2) випливає, що коли математичне сподівання належить до класу функцій ψ , то середня квадратична регресія $\hat{y}_0(x_2, \dots, x_N)$ є найкращою у класі довільних функцій. При цьому залишкова дисперсія множинної регресії $\varepsilon_0 = M_{x_2, \dots, x_N} [D_{x_1}(X_2, \dots, X_N)]$.

Визначення 4. Лінійна функція, яка апроксимує статистичну залежність X_1 від X_2, \dots, X_N , називається *лінійною регресією X_1 на X_2, \dots, X_N* .

Визначення 5. Коефіцієнтами лінійної регресії X_1 на X_2, \dots, X_N називаються функції

$$\beta_{1n} = -\frac{\Lambda_{1n}}{\Lambda_{11}} \quad (n = \overline{2, N}),$$

де Λ_{in} – елементи матриці, що обернена до матриці центральних кореляційних моментів $R(x_i, x_n)$ випадкових величин X_i і X_n ($i, n = \overline{1, N}$).

Теорема 2. Середньою квадратичною регресією X_1 на X_2, \dots, X_N класу лінійних функцій є функція

$$y = \hat{y}_0(x_2, \dots, x_N) = m_{x_1} + \sum_{n=2}^N \beta_{1n}(x_n - m_{x_n}).$$

Визначення 6. Звідним (повним) коефіцієнтом кореляції лінійної множинної регресії називається величина

$$r(x_1, \hat{y}_0) = \sqrt{1 - \frac{1}{\sigma_{x_1}^2 \Lambda_{11}}}.$$

Зауваження. Звідний коефіцієнт кореляції $r(x_1, \hat{y}_0)$ є мірою кореляції між X_1 і X_2, \dots, X_N . При відсутності кореляції цей коефіцієнт дорівнює нулю, а при повній кореляції між X_1 і X_2, \dots, X_N (лінійної функціональної залежності) – одиниці.

Теорема 3. Залишкова дисперсія лінійної множинної регресії визначається виразом

$$\varepsilon_0 = \sigma_{x_1}^2 (1 - r^2(x_1, \hat{y}_0)).$$

Положення 2. Для кількісної оцінки ступеня лінійного впливу різноманітних випадкових величин X_2, \dots, X_N на прогноз X_1 при множинній регресії використовують так звані окремі коефіцієнти кореляції.

Визначення 7. Окремим коефіцієнтом кореляції величин X і Z відносно виключеного Y називається коефіцієнт кореляції $R_{xz(y)} = M[\Delta X \Delta Z]$ залишків $\Delta X = X - \hat{X}_0(Y)$; $\Delta Z = Z - \hat{Z}_0(Y)$, де $\hat{X}_0(Y)$, $\hat{Z}_0(Y)$ – прогнози X і Z по Y .

Зауваження 1. Окремий коефіцієнт кореляції несе інформацію про ступінь лінійного зв'язку між величинами X і Z за умови, що лінійний зв'язок з іншими змінними усунений.

Зауваження 2. Нормований окремих коефіцієнт кореляції величин X і Z відносно виключеного Y виражається через звичайні коефіцієнти кореляції r_{xy}, r_{yz}, r_{xz} :

$$r_{xz(y)} = \frac{r_{xz} - r_{xy} r_{yz}}{\sqrt{(1 - r_{xy}^2)(1 - r_{yz}^2)}}.$$

Зауваження 3. Окремі коефіцієнти кореляції, розраховані для різноманітних пар випадкових величин, дозволяють оцінити значимість лінійних зв'язків між величинами й обгрунтовано прийняти рішення про виключення з регресійної моделі тих величин, що практично не впливають на прогноз.

8.5. Нелінійна середньоквадратична регресія

Положення 1. В тому випадку, коли залежність між випадковими величинами носить явно нелінійний характер (див., наприклад, рис. 8.1.1),

використання лінійної регресії недоречно. В цьому випадку необхідно використовувати нелінійну регресію.

Зауваження 1. У ряді випадків пошук придатної функції регресії має сенс обмежити класом *гладких* функцій. Такі функції, як відомо, можуть бути розкладені у ряд Тейлора (у багатовимірний, якщо число факторів більше одиниці). При обмеженні ряду скінченним числом членів задача пошуку зводиться до знаходження коефіцієнтів відповідного поліному.

Розглянемо схему розв'язання цієї задачі на прикладі одновимірного випадку.

Приклад. Якщо клас нелінійних функцій являє собою клас поліномів ступеня K , то регресія Y на X являє собою функцію

$$\hat{y}(x) = \sum_{k=0}^K a_k x^k,$$

де a_k – параметри, що підлягають визначенню.

Оптимальне значення цих параметрів визначається з умови мінімуму середнього квадрата помилки

$$\varepsilon(a_0, \dots, a_K) = M_x \left[M_y \left[\left(Y / X - \sum_{k=0}^K a_k X^k \right)^2 \right] \right].$$

Методика знаходження цих параметрів не відрізняється від розглянутої раніше для лінійної регресії.

Зауваження 2. Опис функції регресії за допомогою поліномів має сенс у випадку, коли функція регресії *неперіодична*. Для *періодичних* функцій регресії з періодом X використовують розклад у ряд Фур'є:

$$\hat{y}(x) = \sum_{k=-\infty}^{\infty} a_k \exp\left(j \frac{2\pi}{X} kx\right).$$

Примітка 1. При розв'язанні конкретних задач число членів ряду обмежують.

Примітка 2. При застосуванні ряду Фур'є клас функцій ψ , у якому ведеться пошук оптимального розв'язку, не обмежений гладкими функціями.

Положення 2. Ступінь регресійного зв'язку між двома випадковими величинами характеризує кореляційне відношення – поняття, введене К. Пірсоном.

Визначення. *Кореляційним відношенням* називається величина

$$\eta_{Y/X} = 1 - \frac{\varepsilon}{\sigma_y^2},$$

де ε – середній квадрат помилки передбачення випадкової величини Y по випадковій величині X , σ_y^2 – дисперсія випадкової величини Y .

Зауваження 1. У випадку лінійної регресії кореляційне відношення дорівнює квадрату коефіцієнта кореляції r_{xy} і характеризує лінійну залежність між випадковими величинами X і Y . У випадку нелінійної середньоквадратичної регресії кореляційне відношення характеризує не тільки лінійну, але і *нелінійну* залежність між випадковими величинами.

Зауваження 2. Кореляційні відношення часто використовуються при виборі найбільш придатного для прогнозування класу функцій. Той клас, для якого

кореляційне відношення більше, більш придатний для прогнозування, ніж той, для якого це відношення менше.

Положення 3. У випадку множинної регресії знаходження придатного класу функцій ψ являє собою складну задачу. Відомо декілька методик пошуку таких класів. Але жодний з них не дає повністю формалізованого розв'язку задачі. Усі ці методики містять елементи волюнтаризму.

Положення 4. Один із розповсюджених підходів до пошуку придатного класу функцій базується на обмеженні області пошуку функціями, що допускають представлення залежності $\hat{y}(X_2, \dots, X_N)$ у вигляді лінійної комбінації M заданих функцій $\varphi_m(X_2, \dots, X_N)$ ($m = \overline{1, M}$). Ці функції називають *регресорами*. Вони являють собою детерміновані функції випадкових аргументів (факторів) X_2, \dots, X_N :

$$\hat{y}(X_2, \dots, X_N) = \sum_{m=1}^M \alpha_m \varphi_m(X_2, \dots, X_N), \quad (8.5.1)$$

де α_m – константа, що відображає “вагу” m -го регресора.

Зауваження 1. Вираз (8.5.1) описує приблизну залежність випадкової величини X_1 від випадкових величин X_2, \dots, X_N за допомогою розкладу X_1 у ряд по заданих функціях $\varphi_m(X_2, \dots, X_N)$ ($m = \overline{1, M}$).

Зауваження 2. Залежність (8.5.1) є лінійною відносно регресорів φ_m . При нелінійних залежностях $\varphi_m(X_2, \dots, X_N)$ вона є *нелінійною* відносно випадкових функцій (факторів) X_2, \dots, X_N .

Зауваження 3. Якщо регресори є поліномами факторів X_2, \dots, X_N , то функція регресії $\hat{y}(X_2, \dots, X_N)$ являє собою також поліном.

Зауваження 4. При застосуванні моделі (8.5.1) задача пошуку функції множинної *нелінійної* регресії зводиться до розглянутої у підрозділі 8.4 процедури знаходження оптимальних коефіцієнтів α_m ($m = \overline{1, M}$) множинної *лінійної* регресії.

Зауваження 5. У підході, що описано, елементи волюнтаризму, про який ішла мова раніше, проявляються в заданні функцій $\varphi_m(X_2, \dots, X_N)$ ($m = \overline{1, M}$). Ці функції звичайно підбирають з урахуванням існуючих відомостей про характер залежностей, які шукаються, й аналізу поля регресії.

8.6. Коефіцієнт рангової кореляції

Визначення 1. Процедура побудови варіаційного ряду по зменшенню елементів називається *ранжируванням*.

Визначення 2. *Рангом елемента* x_n ($n = \overline{1, N}$) називається номер j , що має елемент x_n у варіаційному ряді після ранжирування елементів.

Про залежність між двома випадковими величинами X і Y можна судити по залежності між рангами A_1, \dots, A_N , B_1, \dots, B_N , розрахованими для випадкових вибірок X_1, \dots, X_N , Y_1, \dots, Y_N цих випадкових величин.

Кількісною мірою залежності між цими випадковими величинами є кореляція рангів.

Визначення 3. Коефіцієнтом рангової кореляції Спірмена випадкових величин X , Y називається коефіцієнт кореляції, розрахований для рангів цих величин:

$$\rho_s = \frac{R_{ab}}{\sqrt{D_a D_b}}, \quad (8.6.1)$$

де R_{ab} – кореляційний момент рангів A і B ; D_a, D_b – дисперсії рангів A і B .

Зауваження. Коефіцієнт рангової кореляції характеризує як лінійний, так і нелінійний зв'язок між випадковими величинами.

Теорема 1. Нехай $d_n = a_n - b_n$ – різниця рангів, отриманих по вибірках x_1, \dots, x_N , y_1, \dots, y_N випадкових величин X , Y ($n = \overline{1, N}$), що розглядаються. Тоді оцінка ρ_s^* коефіцієнта рангової кореляції Спірмена

$$\rho_s^* = 1 - \frac{6 \sum_{n=1}^N d_n^2}{N(N^2 - 1)}. \quad (8.6.2)$$

Теорема 2. Нехай випадкові величини X , Y , які розглядаються, мають гауссівські розподіли. Тоді оцінка коефіцієнта рангової кореляції Спірмена ρ_s^* пов'язана з оцінкою коефіцієнта кореляції r_{xy}^* співвідношенням

$$\rho_s^* = \frac{6(N-2)}{\pi(N+1)} \arcsin \frac{r_{xy}^*}{2} + \frac{6}{\pi(N+1)} \arcsin r_{xy}^*.$$

Примітка. При великому обсязі вибірки ($N > 10$) справедливі співвідношення

$$\rho_s^* \approx \frac{6}{\pi} \arcsin \frac{r_{xy}^*}{2};$$

$$r_{xy}^* \approx 2 \sin\left(\frac{\pi}{6} \rho_s^*\right).$$

Зауваження. Поряд із коефіцієнтом рангової кореляції Спірмена інколи використовують інші параметри, які кількісно характеризують зв'язок між випадковими величинами. До їх числа відноситься, наприклад, *вибірковий коефіцієнт рангової кореляції Кендала*.

Приклад. Нехай є вибірки двох випадкових величин X , Y : $\vec{x} = (3; 5; 1; 2; 1024)$; $\vec{y} = (1; 9; -33; 5; 8)$. Необхідно розрахувати оцінку коефіцієнта рангової кореляції Спірмена.

Розв'язання. Вектори рангів \vec{a} , \vec{b} , відповідні векторам \vec{x} , \vec{y} , мають вигляд $\vec{a} = (3; 2; 5; 4; 1)$, $\vec{b} = (4; 1; 5; 3; 2)$. Тоді $\vec{d} = (-1; 1; 0; 1; -1)$, $\rho_s^* = 0,8$.

8.7. Регресійний аналіз випадкових функцій

Перейдемо до розгляду другого напрямку регресійного аналізу, який пов'язаний з наближеним описом випадкових функцій за допомогою інших випадкових та детермінованих функцій. Спочатку розглянемо приклад.

Приклад. Реальне просторово-часове коливання $U(t, \vec{x})$, що приймається й обробляється радіолокаційною станцією, являє собою випадкове поле, де t – час, \vec{x} – просторовий вектор. Наближено це поле можна описати різними математичними моделями, які з більшою чи меншою помилкою дозволяють

прогнозувати коливання для усіх визначених t і \bar{x} . У найпростішому випадку математичною моделлю коливання може виступати детермінований корисний сигнал $s(t, \bar{x})$ (тоді $U(t, \bar{x}) = s(t, \bar{x})$), у більш складному випадку – сума детермінованого корисного сигналу $s(t, \bar{x})$ і випадкової завади $N(t, \bar{x})$ (тоді $U(t, \bar{x}) = s(t, \bar{x}) + N(t, \bar{x})$). У першому випадку, як бачимо, випадкова функція наближено описується детермінованою функцією, у другому – сумою детермінованої й випадкової функцій.

Розглянемо деякі розповсюджені регресійні моделі.

Модель 1. Випадкова функція $Y(x_1, \dots, x_N)$, що залежить від *невипадкових величин* x_1, \dots, x_N , представляється у вигляді суми детермінованої функції $\hat{y}(x_1, \dots, x_N)$ (функції регресії) й випадкової функції $H(x_1, \dots, x_N)$ (поправкового члена):

$$Y(x_1, \dots, x_N) = \hat{y}(x_1, \dots, x_N) + H(x_1, \dots, x_N). \quad (8.7.1)$$

Примітка. В одновимірному випадку, коли величина x являє собою час, функція регресії іноді називається *трендом*.

Зауваження 1. Слід підкреслити, що у даній моделі вважається, що випадковість випадкової функції $Y(x_1, \dots, x_N)$ обумовлена лише випадковістю поправкового члена $H(x_1, \dots, x_N)$. Функція регресії $\hat{y}(x_1, \dots, x_N)$, яка використовується для прогнозу $Y(x_1, \dots, x_N)$ по значеннях x_1, \dots, x_N , – не випадкова функція.

Зауваження 2. Відповідно до теореми 1 підрозділу 8.4 по критерію мінімуму квадрата помилки оптимальний прогноз забезпечується, коли функція $\hat{y}(x_1, \dots, x_N)$ дорівнює математичному сподіванню $m_y(x_1, \dots, x_N)$ випадкової функції $Y(x_1, \dots, x_N)$ ($\hat{y}(x_1, \dots, x_N) = m_y(x_1, \dots, x_N)$).

Зауваження 3. Найкращій прогноз у класі функцій ψ забезпечує така функція $\hat{y}(x_1, \dots, x_N) = \hat{y}_0(x_1, \dots, x_N)$ з цього класу, при якій досягається мінімум середньої помилки

$$\begin{aligned} \varepsilon &= \frac{1}{V_{\bar{x}}} \int_{\bar{X}} M_y [(Y(x_1, \dots, x_N) - \hat{y}(x_1, \dots, x_N))^2] dx_1 \dots dx_N = \\ &= \frac{1}{V_{\bar{x}}} \int_{\bar{X}} [D_y(x_1, \dots, x_N) + (m_y(x_1, \dots, x_N) - \hat{y}(x_1, \dots, x_N))^2] dx_1 \dots dx_N, \end{aligned} \quad (8.7.2)$$

де $V_{\bar{x}}$ – об'єм області \bar{X} інтегрування, $D_y(x_1, \dots, x_N)$ – дисперсія випадкової величини $Y(x_1, \dots, x_N)$.

Якщо математичне сподівання $m_y(x_1, \dots, x_N)$ входить до класу функцій ψ , що розглядаються, то $\hat{y}_0(x_1, \dots, x_N) = m_y(x_1, \dots, x_N)$ і прогноз є оптимальним у класі довільних функцій. При цьому, як видно з виразу (8.7.2), мінімальна середня помилка (залишкова дисперсія)

$$\varepsilon_0 = \frac{1}{V_{\bar{x}}} \int_{\bar{X}} D_y(x_1, \dots, x_N) dx_1 \dots dx_N.$$

Зауваження 4. З наведених даних випливає, що є лише одна суттєва різниця між моделлю, що розглядається у цьому підрозділі, і моделлю, яка описує множинну регресію випадкової величини, – це спосіб розрахунку середнього. Тому результати, що були отримані для випадкових величин, можуть бути

використані й для регресійного аналізу випадкових функцій, але з відповідною корекцією. Це стосується як лінійних, так і нелінійних регресійних моделей.

Зауваження 5. На практиці використовують дві *регресійні моделі: детерміновану й стохастичну (випадкову)*. Суть цих моделей така, як й аналогічних регресійних моделей випадкових величин. Нехай \hat{y} є функція регресії $\hat{y}(x_1, \dots, x_N)$, що представляє залежність випадкової функції $Y(x_1, \dots, x_N)$ від аргументів x_1, \dots, x_N . У межах детермінованої регресійної моделі можна прогнозувати, що при конкретних $(x_1, \dots, x_N) = (a_1, \dots, a_N)$ випадкова функція буде приймати значення $\hat{y}(a_1, \dots, a_N)$. У межах стохастичної моделі можна прогнозувати, що випадкова функція $Y(x_1, \dots, x_N)$ при конкретних a_1, \dots, a_N буде мати вигляд $\hat{y}(a_1, \dots, a_N) + Z(a_1, \dots, a_N)$, де $Z(a_1, \dots, a_N)$ – випадкова функція з деяким законом розподілу.

Зауваження 6. Закон розподілу випадкової функції $Z(x_1, \dots, x_N)$ в ідеалі повинен дорівнювати закону розподілу поправкового члена $H(x_1, \dots, x_N)$. Часто отримати інформацію про цей закон нелегко, а іноді й неможливо. На практиці, як правило, вважають $Z(x_1, \dots, x_N)$ гауссівською випадковою функцією, причому часто з некорельованими перерізами. Як і у випадку регресійних моделей випадкових величин, для багатьох задач допустима заміна реального закону розподілу на гауссівський закон.

Зауваження 7. На практиці найчастіше використовують детерміновану регресійну модель. Тому у подальшому розгляді орієнтація зроблена на неї.

Зауваження 8. У рамках детермінованої моделі регресія випадкової функції являє собою різновид фільтрації, у результаті якої флуктуаційні компоненти подавляються, а регулярна компонента (математичне сподівання) підкреслюється.

Модель 2. При використанні підходу, що описаний у підрозділі 8.5 (див. положення 4), можна представити модель випадкової функції $Y(x_1, \dots, x_N)$ такою лінійною (відносно $\varphi_m(x_1, \dots, x_N)$) моделлю:

$$Y(x_1, \dots, x_N) = \sum_{m=1}^M \alpha_m \varphi_m(x_1, \dots, x_N) + H(x_1, \dots, x_N), \quad (8.7.3)$$

де α_m – константа, що відображає “вагу” m -го регресору, $\varphi_m(x_1, \dots, x_N)$ – m -й детермінований регресор, який є детермінованою функцією N незалежних аргументів x_1, \dots, x_N , $H(x_1, \dots, x_N)$ – поправковий член, що являє собою випадкову функцію цих аргументів.

Зауваження 1. Методика отримання коефіцієнтів α_m при необмеженій кількості експериментальних даних (чи при дуже великій їх кількості) у цілому не відрізняється від тої, що була розглянута у підрозділі 8.5. Методика отримання цих коефіцієнтів при обмеженій кількості експериментальних даних розглянута у підрозділі 8.8.

Модель 3. Часто зустрічаються задачі, в яких значення отримуються (чи можуть бути розглянуті) у дискретній послідовності точок x_n із фіксованою відстанню Δx одна від одної. Причому дані отримуються послідовно. Тоді для опису функції регресії (тренду) $\hat{y}_n = \hat{y}(x_n)$ може бути використана *авторегресійна модель*, яка аналітично записується так:

$$\hat{y}_n = -\beta_1 \hat{y}_{n-1} - \beta_2 \hat{y}_{n-2} - \dots - \beta_r \hat{y}_{n-r} + \gamma_1 x_{n-1} + \gamma_2 x_{n-2} + \dots + \gamma_r x_{n-p}, \quad (8.7.4)$$

де r, p – параметри, що вказують на глибину залежності n -го значення функції регресії \hat{y}_n від попередніх значень $\hat{y}_{n-1}, \hat{y}_{n-2}, \dots, \hat{y}_{n-r}$ цієї функції та попередніх значень $x_{n-1}, x_{n-2}, \dots, x_{n-p}$ фактора x .

Примітка 1. Модель (8.7.4) іноді називають моделлю *розподіленого лага*. При цьому оставляють назву “авторегресійна модель” для моделі, у якої фактор x явно не визначений:

$$\hat{y}_n = -\beta_1 \hat{y}_{n-1} - \beta_2 \hat{y}_{n-2} - \dots - \beta_r \hat{y}_{n-r}. \quad (8.7.5)$$

Примітка 2. Часто розглядають моделі, у яких попередні значення функції регресії $\hat{y}_{n-1}, \hat{y}_{n-2}, \dots, \hat{y}_{n-r}$ явно не визначені:

$$\hat{y}_n = \gamma_1 x_{n-1} + \gamma_2 x_{n-2} + \dots + \gamma_r x_{n-p}. \quad (8.7.6)$$

Зауваження 1. Окремі моделі (8.7.5), (8.7.6) та загальна модель (8.7.4) є різновидами лінійної моделі (8.7.3)

Зауваження 2. Особливістю багатьох задач, що описуються за допомогою моделей (8.7.4) – (8.7.6), є знаходження у реальному масштабі часу прогнозу чергового значення випадкової функції y_n на базі даних, отриманих на попередніх кроках.

8.8. Оцінка коефіцієнтів лінійної регресійної моделі випадкової функції

На практиці для наближеного опису випадкових функцій широко використовується лінійна регресійна модель (8.7.3). Зупинимося на питанні оцінки коефіцієнтів α_m цієї моделі при обмеженому обсязі експериментальних даних.

Для оцінки цих коефіцієнтів звичайно використовують *метод найменших квадратів (МНК)*.

8.8.1. Метод найменших квадратів

Постановка задачі. Розглядається лінійна регресійна модель випадкової функції $Y(x_1, \dots, x_N)$:

$$Y(x_1, \dots, x_N) = \sum_{m=1}^M \alpha_m \varphi_m(x_1, \dots, x_N) + H(x_1, \dots, x_N), \quad (8.8.1)$$

де x_1, \dots, x_N – незалежні аргументи; $\varphi_m(x_1, \dots, x_N)$ – задані функції (регресори); $H(x_1, \dots, x_N)$ – поправковий член (випадкова функція).

Є результати спостережень $y_l = y(x_{1l}, \dots, x_{Nl})$ випадкової функції $Y(x_1, \dots, x_N)$ у L різних точках (x_{1l}, \dots, x_{Nl}) ($l = \overline{1, L}$). Треба по результатах цих спостережень (експериментів) оцінити коефіцієнти α_m функції регресії

$$\hat{y}(x_1, \dots, x_N) = \sum_{m=1}^M \alpha_m \varphi_m(x_1, \dots, x_N). \quad (8.8.2)$$

Розв’язання. Застосуємо матричний опис величин. Нехай $\vec{y} = (y_1, \dots, y_L)^T$ – L -вимірний вектор-стовпець значень випадкової величини у L спостереженнях; $\vec{\alpha} = (\alpha_1, \dots, \alpha_M)^T$ – M -вимірний вектор-стовпець коефіцієнтів,

які повинні бути оцінені; Φ – матриця розміром $L \times M$ рангом M , строки якої описують значення регресорів $\varphi_m(x_1, \dots, x_N)$ у різних спостереженнях (у різних точках (x_{1l}, \dots, x_{Nl}) ($l = \overline{1, L}$)); $\vec{h} = (h_1, \dots, h_L)^T$ – L -вимірний вектор-стовпець значень поправкового члена $H(x_1, \dots, x_N)$ (значень помилки прогнозу) у L спостереженнях. Тоді вирази (8.8.1), (8.8.2) можна записати у вигляді

$$\vec{y} = \Phi \vec{\alpha} + \vec{h}. \quad (8.8.3)$$

Розглянемо далі дві моделі: канонічну й неканонічну. В *канонічній моделі* перерізи випадкової функції $Y(x_1, \dots, x_N)$ незалежні і мають однакові дисперсії σ_y^2 , а у *неканонічній моделі* перерізи залежні і (або) мають неоднакові дисперсії σ_y^2 .

Випадок 1 (канонічна модель). Якщо перерізи випадкової функції $Y(x_1, \dots, x_N)$ незалежні і мають однакові дисперсії σ_y^2 , помилка прогнозу може бути охарактеризована величиною $\varepsilon = \vec{h}^T \vec{h}$. Тоді з виразу (8.8.3) маємо

$$\varepsilon = (\vec{y} - \Phi \vec{\alpha})^T (\vec{y} - \Phi \vec{\alpha}). \quad (8.8.4)$$

Мінімум помилки буде мати місце, коли похідні помилки по параметрах $\alpha_1, \dots, \alpha_M$ дорівнюють нулеві. Проводячи диференціювання виразу (8.8.4) і, прирівнюючи похідні до нуля, після тривіальних математичних перетворень можна отримати систему рівнянь

$$\Phi^T \Phi \vec{\alpha} = \Phi^T \vec{y}, \quad (8.8.5)$$

розв'язок якої дає найкращу по критерію МНК оцінку $\vec{\alpha}^*$ вектора $\vec{\alpha}$:

$$\vec{\alpha}^* = (\Phi^T \Phi)^{-1} \Phi^T \vec{y}. \quad (8.8.6)$$

Зауваження 1. Матрицю $G = \Phi^T \Phi$ (розміром $M \times M$) часто називають *інформаційною матрицею*, а обернену матрицю $C = (\Phi^T \Phi)^{-1}$ – *матрицею помилок*.

Зауваження 2. Оскільки оцінка (8.8.6) базується на випадкових спостереженнях, вона є випадковою векторною величиною. Доведено, що у разі, коли математичне сподівання поправкового члена дорівнює нулю, така оцінка $\vec{\alpha}^*$ є незміщеною, спроможною й ефективною серед усіх *незміщених лінійних* оцінок (якщо випадкова величина має гауссівський закон розподілу, то МНК-оцінка є ефективною не тільки серед усіх незміщених лінійних, але й нелінійних оцінок).

Зауваження 3. Доведено, що у цьому випадку кореляційна матриця оцінки $\vec{\alpha}^*$

$$R_{\vec{\alpha}^*} = M[(\vec{\alpha}^* - \vec{\alpha})(\vec{\alpha}^* - \vec{\alpha})^T] = (\Phi^T \Phi)^{-1} \sigma_y^2 = C \sigma_y^2, \quad (8.8.7)$$

а дисперсія оцінки функції регресії \hat{y}^* (яка відповідає отриманим оцінкам $\vec{\alpha}^*$)

$$D[\hat{y}^*] = \vec{\phi}^T R_{\vec{\alpha}^*} \vec{\phi}, \quad (8.8.8)$$

де $\vec{\phi} = (\varphi_1, \dots, \varphi_M)^T$ – M -вимірний вектор-стовпець регресорів $\varphi_m(x_1, \dots, x_N)$ ($m = \overline{1, M}$).

Зауваження 4. Якщо випадкова функція $Y(x_1, \dots, x_N)$ – гауссівська з математичним сподіванням $\hat{y}(x_1, \dots, x_N) = \vec{\alpha}^T \vec{\phi}(x_1, \dots, x_N)$, то вектор оцінок регресійних коефіцієнтів $\vec{\alpha}^*$ має гауссівський закон розподілу з математичним

сподіванням $\vec{\alpha}$ і кореляційної матрицею $R_{\vec{\alpha}^*}$, тобто M -вимірна щільність ймовірностей вектора коефіцієнтів $\vec{\alpha}^*$ має вигляд

$$f_M(\vec{\alpha}^*) = \frac{1}{(2\pi)^{\frac{M}{2}} |R_{\vec{\alpha}^*}|^{1/2}} \exp\left(-\frac{(\vec{\alpha}^* - \vec{\alpha})^T R_{\vec{\alpha}^*}^{-1} (\vec{\alpha}^* - \vec{\alpha})}{2}\right).$$

Доведення цього положення засноване на властивостях лінійного перетворення гауссівських випадкових величин.

Випадок 2 (неканонічна модель). Якщо перерізи випадкової функції $Y(x_1, \dots, x_N)$ залежні і (або) мають неоднакові дисперсії σ_y^2 , вирази (8.8.7), (8.8.8) прямо не придатні для розрахунку кореляційної матриці оцінки коефіцієнтів $\vec{\alpha}^*$ і дисперсії оцінки функції регресії \hat{y}^* . Щоб можна було використовувати ці вирази, спочатку слід перейти від випадкової функції $Y(x_1, \dots, x_N)$ до випадкової функції $Z(x_1, \dots, x_N)$, яка має некорельовані значення з однаковою дисперсією σ_z^2 . Такий перехід здійснюється лінійним перетворенням, що реалізується множенням випадкової функції $Y(x_1, \dots, x_N)$ на спеціальну не вироджену матрицю P^{-1} розміром $L \times L$. Для отриманої таким чином випадкової функції $Z(x_1, \dots, x_N)$ вже справедливі вирази (8.8.7), (8.8.8).

Розглядаючи спочатку регресійну залежність для випадкової функції $Z(x_1, \dots, x_N)$, а потім переходячи до випадкової функції $Y(x_1, \dots, x_N)$, можна довести, що по критерію МНК найкраща оцінка $\vec{\alpha}^*$ вектора $\vec{\alpha}$ описується виразом

$$\vec{\alpha}^* = (\Phi^T \Omega^{-1} \Phi)^{-1} \Phi^T \Omega^{-1} \vec{y}, \quad (8.8.9)$$

де $\Omega = PP^T$.

При цьому кореляційна матриця оцінки $\vec{\alpha}^*$

$$R_{\vec{\alpha}^*} = (\Phi^T \Omega^{-1} \Phi)^{-1} \sigma_z^2. \quad (8.8.10)$$

Зауваження 1. Для неканонічної моделі інформаційна матриця $G = \Phi^T \Omega^{-1} \Phi$.

Зауваження 2. Оцінка коефіцієнтів (8.8.9) є незміщеною і ефективною у класі незміщених лінійних оцінок. Доведено, крім того, що при деяких обмеженнях, щодо матриць Φ і Ω , ця оцінка є й спроможною.

Зауваження 3. З наведених зауважень, щодо ефективності оцінок (8.8.6) і (8.8.9), не впливає, що оцінки не можуть бути кращими. Це можливо, наприклад, при спеціальному плануванні експериментів та у класі зміщених оцінок. Такі оцінки розглядаються у підрозділах 8.8.2.

8.8.2. Планування експериментів

Положення 1. У підрозділі 8.8.1 відзначалося, що для лінійної регресійної моделі при зафіксованих результатах спостережень $y_l = y(x_{l1}, \dots, x_{lN})$ випадкової функції $Y(x_1, \dots, x_N)$ у точках (x_{l1}, \dots, x_{lN}) ($l = \overline{1, L}$) МНК забезпечує отримання серед деякого класу оцінок ефективною оцінкою вектора коефіцієнтів $\vec{\alpha}$ функції регресії $\mathfrak{f}(x_1, \dots, x_N) = \sum_{m=1}^M \alpha_m \varphi_m(x_1, \dots, x_N)$, тобто отримання найкращої оцінки у цьому класі.

Поліпшення цієї оцінки може бути можливим, якщо ведеться *планування експериментів*. З виразу (8.8.7) видно, що зміна матриці Φ призводить до зміни

кореляційної матриці оцінки R_{α^*} . З цього випливає, що при фіксованому обсязі експериментів відповідним підбором (плануванням) експериментів (тобто зміною матриці Φ) можна підвищити якість оцінки α^* й навіть отримати оптимальну оцінку по деякому заздалегідь вибраному критерію. При розв'язанні такої задачі використовують різні критерії оптимальності. Найбільш відомі критерії оптимальності наведено в табл.8.8.1.

Таблиця 8.8.1

Назва критерію	Математичне формулювання критерію
D -оптимальність	$ \Phi_0^T \Phi_0 ^{-1} = \min_{\Phi} \Phi^T \Phi ^{-1}$
A -оптимальність	$\text{tr}(\Phi_0^T \Phi_0)^{-1} = \min_{\Phi} \text{tr}(\Phi^T \Phi)^{-1}$
G -оптимальність	$\max_{\bar{x}} D[\bar{f}(\bar{x}, \Phi_0)] = \min_{\Phi} \max_{\bar{x}} D[\bar{f}(\bar{x}, \Phi)]$

Примітка. Наведені у таблиці формули відповідають канонічній лінійній регресійній моделі. У випадку неканонічної моделі можна рекомендувати таку схему оптимізації плану. Спочатку за допомогою матриці P слід перейти до канонічної моделі, знайти для неї оптимальний по деякому критерію план експериментів, а потім за допомогою оберненого перетворення отримати шуканий план Φ .

Зауваження. D -оптимальний план мінімізує узагальнену дисперсію, A -оптимальний план мінімізує середню дисперсію, а G -оптимальний план мінімізує дисперсію прогнозу у точці факторного простору, де ця дисперсія максимальна.

Положення 2. При плануванні експериментів особливе значення має ортогональність плану.

Визначення. План називається *ортогональним*, якщо відповідна інформаційна матриця G – діагональна.

Зауваження 1. Ортогональність плану означає, що усі коефіцієнти регресійної моделі оцінюються незалежно. При цьому кореляційні матриці (8.8.7), (8.8.10) – діагональні і точність оцінки кожного коефіцієнту не залежить від кількості коефіцієнтів та точності оцінки інших коефіцієнтів.

Зауваження 2. Ортогональність плану експериментів дуже спрощує розв'язання задачі знаходження оцінки коефіцієнтів α_m регресійної моделі та розрахунку точності їх оцінки. Тому, якщо можливо, має сенс використовувати ортогональні плани.

На кінець відзначимо, що розглянуті підходи, моделі та методи розв'язання задач регресійного аналізу ніяк не претендують на повноту. Є дуже багато важливих результатів, які заслуговують детального викладення, але обмежені рамки книги не дозволяють цього зробити. Наприклад, не розглянуті питання регресії векторних випадкових величин і функцій, питання оптимізації планів та багато інших.

Відомості про книги до розділу 8 наведено в табл. 8.8.2.

Таблиця 8.8.2

Література до розділу 8

Характер книг	Номери книг у списку літератури
Довідники	11, 24, 40, 47, 49
Книги навчального плану для інженерів	17, 32, 42, 67, 80
Книги з математичним ухилом	5, 21, 50, 60, 101

Книги прикладного характера	1, 100, 106
-----------------------------	-------------

9. Основи теорії перевірки статистичних гіпотез

9.1. Основні поняття теорії перевірки статистичних гіпотез

Визначення 1. Будь-яке припущення про властивості ймовірнісних об'єктів називається *статистичною гіпотезою*.

Визначення 2. Гіпотеза називається *простою*, якщо їй відповідає лише один повністю визначений закон розподілу ймовірнісного об'єкта, що досліджується, і *складною*, якщо цій гіпотезі відповідає декілька можливих законів розподілу.

Зауваження. При опису ймовірнісного об'єкта за допомогою функції вірогідності, що являє собою однопараметричне сімейство функцій $L(\bar{x}/\theta)$, параметр θ у випадку простої гіпотези дорівнює конкретному значенню θ_0 , а у випадку складної гіпотези – одному із значень деякої множини Ω ($\theta \in \Omega$).

Приклад 1. Досліджується наробка на відмову деякого технічного пристрою, що складається з різних блоків. Закон розподілу блоків відрізняється один від одного. При цьому гіпотеза про те, що пристрій за один рік роботи не вийде з ладу, є простою, а гіпотеза про те, що він вийде з ладу, – складною, бо закон розподілу відмови пристрою визначається законами розподілу відмов окремих блоків. Гіпотези ж про відмову окремих блоків є простими.

Основна задача перевірки статистичних гіпотез формулюється таким чином.

Формулювання основної задачі. Є деякий ймовірнісний об'єкт X , який може знаходитися в одному із K станів з апіорною ймовірністю P_k знаходження у стані S_k ($k=\overline{1, K}$). Множина можливих станів S_k ($k=\overline{1, K}$) утворює повну групу, тобто $\sum_{k=1}^K P_k = 1$. Про фактичний стан об'єкта висловлюється K гіпотез (*альтернатив*): H_0, \dots, H_{K-1} .

Вимагається за наявними реалізаціями x_1, \dots, x_N ймовірнісного об'єкта X прийняти рішення про ту з гіпотез, яка найбільшою мірою відповідає фактичному стану об'єкта. При цьому кількість можливих рішень завжди дорівнює кількості гіпотез H_k , тобто дорівнює K .

Визначення 3. Функцією вірогідності $L(\bar{x}/H_k)$ гіпотези H_k називається умовна щільність ймовірностей випадкової вибірки $\vec{X} = (X_1, \dots, X_N)$ ймовірнісного об'єкта X , яка відповідає гіпотезі H_k .

Визначення 4. Множина всіх значень, що може приймати випадкова вибірка \vec{X} , утворює *простір вибірок* G (*простір спостережень*).

Визначення 5. Множина всіх можливих рішень про приналежність вибірки до однієї з висунутих гіпотез утворює *простір рішень* γ .

Розглянемо варіант розв'язання основної задачі в найпростішому випадку, коли можливих гіпотез усього дві: H_0 і H_1 .

Визначення 6. У двоальтернативному випадку одна з гіпотез (H_0) вважається *основною* (*нульовою*), а друга (H_1) – *конкуруючою* (*альтернативною*).

Нехай проведено експеримент, в результаті якого отримана вибірка x_1, \dots, x_N випадкової величини X . Закон розподілу цієї випадкової величини невідомий. Однак вважається (існує гіпотеза), що цей закон може описуватися або функцією вірогідності $L(\vec{x}/H_0)$ (основна гіпотеза H_0), або функцією вірогідності $L(\vec{x}/H_1)$ (конкуруюча гіпотеза H_1).

Визначення 7. Область G_0 простору спостережень G , що відповідає рішенням γ_0 про справедливість основної гіпотези H_0 , називається *допустимою* областю, а область G_1 цього ж простору, що відповідає рішенням γ_1 про справедливість конкуруючої гіпотези, називається *критичною* областю.

Зауваження. Допустима й критична області не обов'язково повинні бути однозв'язними.

Визначення 8. Правило, у відповідності з яким проводиться розподіл простору спостережень G на допустиму G_0 і критичну G_1 області, називається *правилом прийняття рішення (критерієм прийняття рішення)*.

Зауваження 1. Звичайно функції вірогідності $L(\vec{x}/H_0)$ і $L(\vec{x}/H_1)$ *перекриваються*. Це породжує серйозну проблему визначення правила прийняття рішення.

Зауваження 2. Завдання перевірки гіпотез полягає в тому, щоб з'ясувати, до якої саме з цих двох функцій вірогідності відноситься конкретна вибірка \vec{x} , яка розглядається. У такій постановці завдання зводиться до розподілу простору вибірок (простору спостережень) G на дві області G_0 і G_1 , що не перекриваються. Якщо вектор \vec{x} потрапляє в область G_0 , приймається рішення γ_0 , що вірна гіпотеза H_0 , а якщо він потрапляє в область G_1 – рішення γ_1 , що вірна гіпотеза H_1 (рис. 9.1.1).



Рис. 9.1.1. Простір спостережень і простір рішень

Приклад 2. Радіолокаційна станція веде виявлення об'єктів. Рішення про наявність або відсутність об'єкта приймається по енергії прийнятого сигналу. Розглядаються дві гіпотези: ціль є (основна гіпотеза H_0) і цілі немає (конкуруюча гіпотеза H_1). В даному випадку функції вірогідності $L(\vec{x}/H_0)$ і $L(\vec{x}/H_1)$ перекриваються (рис. 9.1.2).

Наведений приклад ілюструє той факт, що при перекритті функцій вірогідності можливі помилки в прийнятті рішень.

У випадку розв'язання двоальтернативної задачі можуть спостерігатися чотири ситуації:

1) висувається основна гіпотеза H_0 і при цьому вибірка \vec{x} належить допустимій області G_0 . Тоді правильно приймається основна гіпотеза H_0 ;

2) висувається основна гіпотеза H_0 і при цьому вибірка \vec{x} належить критичній області G_1 . Тоді помилково відхиляється основна гіпотеза H_0 ;

3) висувається конкуруюча гіпотеза H_1 і при цьому вибірка \vec{x} належить допустимій області G_0 . Тоді помилково приймається основна гіпотеза H_0 ;

4) висувається конкуруюча гіпотеза H_1 і при цьому вибірка \vec{x} належить критичній області G_1 . Тоді правильно відхиляється основна гіпотеза H_0 (правильно приймається конкуруюча гіпотеза H_1).

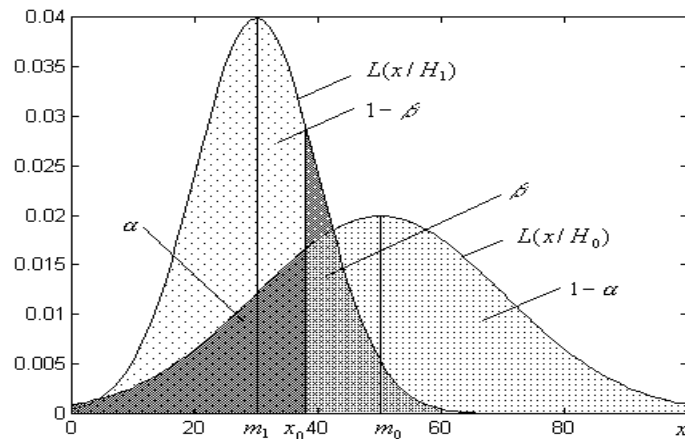


Рис. 9.1.2. Функції вірогідності та ймовірні характеристики системи виявлення

Визначення 9. Помилкове відхилення основної гіпотези H_0 називається помилкою першого роду (у різноманітних додатках – пропуском цілі, ризиком виробника та ін.).

Визначення 10. Помилкове прийняття основної гіпотези H_0 називається помилкою другого роду (у різноманітних додатках – хибною тривоною, помилкою верифікації та ін.).

Ймовірності, відповідні описаним вище різноманітним ситуаціям, наведені в табл. 9.1.1 і зображені на рис. 9.1.2 у вигляді площ під функціями вірогідності $L(\vec{x}/H_0)$ і $L(\vec{x}/H_1)$.

Визначення 11. Ймовірність помилки першого роду α називається рівнем значимості.

Визначення 12. Ймовірність $(1 - \beta)$ правильного відхилення основної гіпотези H_0 називається потужністю критерію прийняття рішення.

Зауваження 1. Потужність $(1 - \beta)$ критерію є кількісною мірою якості критерію. При фіксованому рівні значимості α критерій, який дає найбільшу потужність, забезпечує найкраще прийняття рішення.

Зауваження 2. Коли простір спостережень G є одновимірним, часто межі розподілу областей, що відповідають найбільш потужному критерію, являють собою одну точку (поріг). В цьому випадку обидві області G_0 , G_1 однозв'язні. Рішення приймається по перевищенню або неперевищенню виставленого порога x_0 (рис. 9.1.2). Слід зазначити, що можливі ситуації, коли межі розділу являють собою не одну точку, а цілу множину точок. Це відбувається у випадку багатозв'язних областей G_0 , G_1 .

Таблиця 9.1.1

Ймовірності різноманітних подій при розв'язанні двоальтернативної задачі

Гіпотеза, що висувається	Рішення γ_0 (потрапляння в допустиму область G_0)	Рішення γ_1 (потрапляння в критичну область G_1)
Основна гіпотеза H_0	$1 - \alpha = P(\vec{x} \in G_0 / H_0) =$ $= \int_{G_0} L(\vec{x} / H_0) d\vec{x}$	$\alpha = P(\vec{x} \in G_1 / H_0) =$ $= \int_{G_1} L(\vec{x} / H_0) d\vec{x}$
	Ймовірність правильного прийняття гіпотези H_0	Ймовірність помилки першого роду (рівень значимості)
Конкуруюча гіпотеза H_1	$\beta = P(\vec{x} \in G_0 / H_1) =$ $= \int_{G_0} L(\vec{x} / H_1) d\vec{x}$	$1 - \beta = P(\vec{x} \in G_1 / H_1) =$ $= \int_{G_1} L(\vec{x} / H_1) d\vec{x}$
	Ймовірність помилки другого роду	Ймовірність правильного прийняття гіпотези H_1 (потужність критерію)

9.2. Критерій якості прийняття рішення

Значимість того або іншого вірного або невірного рішення різна в різноманітних ситуаціях. Важливість прийняття рішень характеризують *матрицею втрат*

$$P = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix},$$

де P_{00} – плата за правильне прийняття основної гіпотези H_0 ; P_{01} – плата за помилку першого роду; P_{10} – плата за помилку другого роду; P_{11} – плата за правильне прийняття конкуруючої гіпотези H_1 .

Приклад. Людина збирається вийти на вулицю. Виглянувши у вікно, вона приймає одну з двох гіпотез: дощу не буде (гіпотеза H_0) або буде дощ (гіпотеза H_1). Якщо вона вирішує, що буде дощ, то бере із собою парасольку, в іншому випадку – ні. Для молодої здорової людини мати парасольку, коли йде дощ, і не мати парасольку, коли дощу немає, практично однаково важливо, а промокнути під дощем майже так само неприємно, як цілий день носити із собою парасольку в ясну погоду. Цілком інша ситуація, якщо людина літня та хвора. Для хворої людини знаходитися на вулиці в дощову погоду не дуже бажано навіть у випадку, якщо у неї є парасолька, а опинитися без парасольки, коли йде дощ, вкрай небезпечно. В той самий час носити із собою парасольку в ясну погоду достатньо важко.

На підставі вищезгаданого для здорової людини співвідношення між елементами матриці втрат можуть мати такий вигляд:

$$P_{01} \approx P_{10} \gg P_{00} \approx P_{11}.$$

Для хворої ж літньої людини – інакше:

$$P_{01} \gg \gg P_{11} > P_{10} \gg P_{00}.$$

Визначення 1. Умовними середніми втратами r_0 , r_1 для основної й конкуруючої гіпотез називаються такі величини, що визначаються елементами матриці втрат Π і помилками α , β першого й другого роду:

$$\begin{aligned} r_0 &= \Pi_{00}(1-\alpha) + \Pi_{01}\alpha; \\ r_1 &= \Pi_{10}\beta + \Pi_{11}(1-\beta). \end{aligned}$$

Зауваження. Якщо апріорно відомо, що ймовірність справедливості гіпотези H_0 дорівнює p , а справедливість гіпотези H_1 дорівнює $q=1-p$, то повні середні втрати

$$R = pr_0 + qr_1 = p\Pi_{00} + q\Pi_{11} + p\alpha(\Pi_{01} - \Pi_{00}) + q\beta(\Pi_{10} - \Pi_{11}). \quad (9.2.1)$$

Визначення 2. Середнім ризиком називаються повні середні втрати, що визначаються співвідношенням (9.2.1).

На практиці використовують цілий ряд різноманітних критеріїв прийняття рішень. Розглянемо деякі з них.

Критерій 1. Правило прийняття рішення, яке мінімізує середній ризик, називається *байєсівським правилом або критерієм Байєса*.

Визначення 3. Середній ризик, що відповідає критерію Байєса, називається *байєсівською функцією втрат*.

Для знаходження правила Байєса скористаємось формулами, наведеними у табл. 9.1.1 для ймовірностей α і $(1-\beta)$. Підставляючи ці формули у вираз (9.2.1), маємо

$$R = p\Pi_{00} + q\Pi_{10} - \int_{G_1} [q(\Pi_{10} - \Pi_{11})L(\vec{x}/H_1) - p(\Pi_{01} - \Pi_{00})L(\vec{x}/H_0)]d\vec{x}. \quad (9.2.2)$$

Мінімум середнього ризику забезпечується в тому випадку, коли значення інтеграла максимальне. Максимум же значення інтеграла досягається тоді, коли область інтегрування G_1 вибрана таким чином, що в неї входять тільки ті точки, в яких підінтегральний вираз приймає додатні значення, тобто

$$q(\Pi_{10} - \Pi_{11})L(\vec{x}/H_1) - p(\Pi_{01} - \Pi_{00})L(\vec{x}/H_0) > 0. \quad (9.2.3)$$

З виразу (9.2.3) випливає

$$l(\vec{x}) = \frac{L(\vec{x}/H_1)}{L(\vec{x}/H_0)} > c, \quad (9.2.4)$$

де $c = \frac{p\Pi_{01} - \Pi_{00}}{q\Pi_{10} - \Pi_{11}}$ – константа, що визначається умовою задачі.

Визначення 4. Функція $l(\vec{x}) = \frac{L(\vec{x}/H_1)}{L(\vec{x}/H_0)}$, що являє собою відношення функцій вірогідності, називається *відношенням вірогідності*.

Зауваження. З виразу (9.2.4) випливає, що правило Байєса передбачає:

1) розрахунок функцій вірогідності $L(\vec{x}/H_1)$ і $L(\vec{x}/H_0)$ для конкретної вибірки $\vec{x} = (x_1, \dots, x_N)$;

2) обчислення відношення вірогідності $l(\vec{x})$;

3) порівняння відношення вірогідності $l(\vec{x})$ з порогом. В тому випадку, коли нерівність (9.2.4) справедлива, приймається рішення, що вибірка \vec{x} належить критичній області G_1 (основна гіпотеза H_0 відхиляється). В тому випадку, коли

нерівність (9.2.4) не справедлива, приймається рішення, що вибірка \vec{x} належить допустимій області G_0 (основна гіпотеза H_0 приймається).

Критерій 2. Правило прийняття рішення $l(\vec{x}) > \frac{p}{q}$, що припускає порівняння відношення вірогідності $l(\vec{x})$ з відношенням апіорних ймовірностей p, q гіпотез H_0, H_1 , називається *критерієм ідеального спостерігача*.

Зауваження 1. Критерій ідеального спостерігача – окремий випадок критерію Байєса. Він впливає з критерію Байєса при $\Pi_{00} = \Pi_{11} = 0$; $\Pi_{01} = \Pi_{10} = \Pi$.

Зауваження 2. Для критерію ідеального спостерігача байєсівська функція втрат має вигляд $R = \Pi(p\alpha + q\beta)$. Звідси випливає, що в даному випадку мінімізація середнього ризику R зводиться до мінімізації апіорної ймовірності помилкового рішення $(p\alpha + q\beta)$.

Визначення 5. Умовна ймовірність $P(H_k / \vec{x})$ справедливості гіпотези H_k за умови, що отримана вибірка \vec{x} , називається *апостеріорною ймовірністю* гіпотези H_k .

Критерій 3. Правило прийняття рішення $P(H_0 / \vec{x}) < P(H_1 / \vec{x})$, що припускає порівняння апостеріорних ймовірностей $P(H_0 / \vec{x})$ і $P(H_1 / \vec{x})$, називається *критерієм максимуму апостеріорної ймовірності*.

Теорема. Критерій максимуму апостеріорної ймовірності збігається з критерієм ідеального спостерігача.

Доведення теореми засновано на формулі повної ймовірності.

Критерій 4. Критерій ідеального спостерігача для випадку, коли апіорні ймовірності p, q гіпотез H_0, H_1 однакові ($l(\vec{x}) > 1$), називається *критерієм максимальної вірогідності*.

Критерій 5. Критерій, який мінімізує помилку другого роду β за умови, що помилка першого роду α не перевищує задану величину α_{\max} , називається *критерієм Неймана – Пірсона*.

Критерій 6. Правило, яке мінімізує максимальне значення умовних середніх втрат r_0, r_1 , називається *мінімаксним правилом*.

Зауваження 1. Якщо є цілий ряд вирішальних правил із своїми умовними середніми втратами r_{0j}, r_{1j} ($j = 1, 2, \dots$), то мінімаксне правило припускає вибір серед них правила з номером j , що задовольняє умові $j^* = \arg \min_j \max(r_{0j}, r_{1j})$.

Зауваження 2. Мінімаксне правило є дуже "обережним" і застосовується достатньо рідко.

Різноманітні критерії прийняття рішення показані на рис. 9.2.1.

9.3. Перевірка гіпотез про значення математичного сподівання гауссівської випадкової величини

Застосування описаних у попередньому розділі критеріїв проілюструємо таким прикладом.

Постановка задачі. Нехай спостерігається вибірка \vec{X} гауссівської випадкової величини X з відомою дисперсією σ_x^2 . Математичне сподівання m_x

цієї величини точно не відомо, але припускається, що воно може дорівнювати або m_0 (гіпотеза H_0), або m_1 (гіпотеза H_1).



Рис. 9.2.1. Різноманітні критерії прийняття рішення

Потрібно перевірити гіпотези H_0 , H_1 і прийняти рішення про приналежність вибірки випадковій величині з математичним сподіванням m_0 або m_1 .

Розв’язання. Зважаючи на те, що обидва розподіли гауссівські, байєсівське вирішальне правило можна записати таким чином:

$$l(\vec{x}) = \exp\left(\frac{1}{2\sigma_x^2} \left(\sum_{n=1}^N (x_n - m_0)^2 - \sum_{n=1}^N (x_n - m_1)^2\right)\right) > c. \quad (9.3.1)$$

Припустимо, що $m_0 > m_1$. Тоді, логарифмуючи вираз (9.3.1), після нескладних алгебраїчних перетворень одержуємо

$$\frac{1}{N} \sum_{n=1}^N x_n < \frac{m_0 + m_1}{2} + \frac{\sigma_x^2 \ln c}{N(m_1 - m_0)} = K. \quad (9.3.2)$$

З виразу (9.3.2) видно, що байєсівське вирішальне правило зводиться до знаходження точкової оцінки математичного сподівання $m_x^* = \frac{1}{N} \sum_{n=1}^N x_n$ й порівняння її з порогом K . Якщо $m_x^* < K$, то приймається гіпотеза H_0 , у протилежному ж випадку – гіпотеза H_1 .

Поріг K залежить від параметра c . Для критерію максимальної вірогідності $c = 1$ і $K = \frac{m_0 + m_1}{2}$.

Ймовірності помилки першого α й другого β роду можуть бути розраховані таким чином (див. рис. 9.1.2):

$$\alpha = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_0} e^{-\frac{x^2}{2}} dx = F(z_0);$$

$$\beta = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{\infty} e^{-\frac{x^2}{2}} dx = 1 - F(z_1),$$

де $z_0 = \frac{(K - m_0)\sqrt{N}}{\sigma_x}$; $z_1 = \frac{(K - m_1)\sqrt{N}}{\sigma_x}$; $F(x)$ – гауссівська функція розподілу.

9.4. Послідовний аналіз

Розглянуті в розділі 9.2 критерії прийняття рішення не є оптимальними з точки зору обсягу N вибірки, яка необхідна для прийняття рішення із заданим рівнем помилок.

Приклад. Нехай відомо, що випадкова величина X має рівномірний розподіл. Для основної гіпотези H_0 її щільність ймовірностей

$$f_1(x/H_0) = \begin{cases} \frac{1}{b-a} & \text{при } a \leq x \leq b, \\ 0 & \text{- в інших випадках,} \end{cases}$$

а для альтернативної гіпотези H_1 –

$$f_1(x/H_1) = \begin{cases} \frac{1}{d-c} & \text{при } c \leq x \leq d, \\ 0 & \text{- в інших випадках.} \end{cases}$$

Щільності ймовірностей перекриваються, тобто $a < c < b < d$ (рис. 9.4.1).

Необхідно прийняти рішення з ймовірністю одиниця.

Розв'язання. Нехай перша вибірка потрапила до I або III інтервалу ($[a, c]$ або $(b, d]$). Тоді вже при $N = 1$ з ймовірністю одиниця може бути прийнята гіпотеза H_0 або H_1 . Якщо ж перша вибірка влучила в другий інтервал ($[c, b]$), то достовірно прийняти рішення не можна. Необхідно отримати другу вибірку і проаналізувати її. Якщо вона потрапила до I або III інтервалу, то можна прийняти відповідне рішення, у противному ж разі треба здійснити чергову вибірку. І так до тих пір, поки чергова вибірка не потрапить до I або III інтервалу.

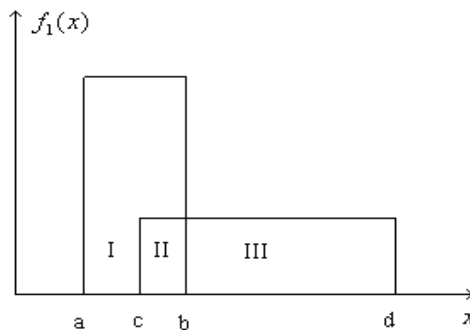


Рис. 9.4.1. Щільності двох рівномірних розподілів

Наведений приклад вказує на те, що в загальному випадку вірогідність прийняття рішення залежить від тих значень, що приймають вибірки. Можливі ситуації, коли навіть по одній вибірці можна прийняти рішення з дуже малим рівнем помилки.

Визначення 1. *Послідовним аналізом* називається процедура прийняття рішення, що передбачає послідовне отримання вибірових значень, їхній аналіз і прийняття одного з трьох рішень:

- 1) справедлива гіпотеза H_0 ;
- 2) справедлива гіпотеза H_1 ;
- 3) обсяг інформації не достатній для обґрунтованого прийняття однієї з гіпотез H_0 , H_1 , для прийняття рішення необхідно збільшити обсяг N вибірки.

Положення 1. Обсяг вибірки N , що гарантується при послідовному аналізі прийняття рішення з необхідною вірогідністю, є випадковою величиною.

Положення 2. При послідовному аналізі простір вибірок G розбивається на три області: G_0 – допустиму область, G_1 – критичну область і G_{\sim} – проміжну область.

Положення 3. Межі областей G_0 , G_1 , G_{\sim} можуть змінюватися зі зміною обсягу вибірки N .

Визначення 2. При послідовному аналізі оптимальною процедурою розділу простору спостережень G називається таке правило, що мінімізує середній обсяг N вибірки, необхідний для прийняття рішення із заданими помилками першого α й другого β роду.

Теорема 1. Оптимальна процедура розділу простору спостережень не залежить від закону розподілу випадкової величини X .

Теорема 2. Для оптимальної процедури розділу простору спостережень області G_0 , G_1 , G_{\sim} визначаються нерівностями, наведеними в табл. 9.4.1. Така процедура називається послідовним критерієм Вальда.

Таблиця 9.4.1

Нерівності, які визначають для послідовного критерію Вальда області G_0 , G_1 , G_{\sim} простору спостережень G

Назва області	Нерівності
Допустима G_0	$l(x_1, \dots, x_N) \leq c_0, \quad c_0 < l(x_1, \dots, x_{N-1}) < c_1$
Критична G_1	$l(x_1, \dots, x_N) \geq c_1, \quad c_0 < l(x_1, \dots, x_{N-1}) < c_1$
Проміжна G_{\sim}	$c_0 < l(x_1, \dots, x_N) < c_1$

Наведені в табл. 9.4.1 пороги c_0 , c_1 , описуються такими виразами:

$$c_0 = \frac{\beta}{(1-\alpha)},$$

$$c_1 = \frac{(1-\beta)}{\alpha}.$$

Теорема 3. У критерії Вальда середній обсяг вибірки, необхідний для прийняття гіпотези H_0 із заданими помилками першого α й другого β роду, визначається співвідношенням

$$N_0 = \frac{1}{a_0} \left[(1-\alpha) \ln \frac{\beta}{1-\alpha} + \alpha \ln \frac{1-\beta}{\alpha} \right],$$

а для прийняття гіпотези H_1 – співвідношенням

$$N_1 = \frac{1}{a_1} \left[\beta \ln \frac{\beta}{1-\alpha} + (1-\beta) \ln \frac{1-\beta}{\alpha} \right],$$

де

$$a_0 = \int_{-\infty}^{\infty} f_1(x/H_0) \ln \frac{f_1(x/H_1)}{f_1(x/H_0)} dx,$$

$$a_1 = \int_{-\infty}^{\infty} f_1(x/H_1) \ln \frac{f_1(x/H_1)}{f_1(x/H_0)} dx.$$

Зауваження 1. При одних і тих самих помилках середній обсяг вибірки, що забезпечує прийняття однієї з гіпотез по критерію Вальда, менш обсягу,

необхідного для прийняття рішення по будь-якому критерію фіксованого обсягу (зокрема, по критерію Байєса).

Зауваження 2. Незважаючи на факт, відмічений у зауваженні 1, внаслідок випадковості обсягу вибірки конкретний обсяг вибірки при використанні критерію Вальда може перевищувати величину, необхідну для прийняття рішення з використанням інших критеріїв.

Зауваження 3. Враховуючи зауваження 2, доцільно вживати таку послідовність застосування критеріїв:

- визначити обсяг N вибірки, необхідний для критерію фіксованого обсягу;
- застосувати критерій Вальда;
- якщо при використанні критерію Вальда не вдається прийняти рішення про справедливість однієї з гіпотез H_0 , H_1 при досягненні обсягу вибірки N , приймати рішення по критерію фіксованого обсягу.

9.5. Перевірка непараметричних статистичних гіпотез

9.5.1. Поняття про критерії згоди

Визначення 1. Статистична гіпотеза називається *параметричною*, якщо вона висловлюється відносно деякого параметра закону розподілу ймовірного об'єкта, і *непараметричною*, якщо вона висловлюється відносно самого закону розподілу.

Приклад 1. Розглянуті вище критерії Байєса й Вальда використовують параметричні гіпотези.

Приклад 2. Гіпотеза про те, що випадкова величина X має певний тип закону розподілу $F_h(x)$, є непараметричною. Такою ж є і гіпотеза про те, що ця випадкова величина має закон розподілу, відмінний від $F_h(x)$.

Зауваження 1. Непараметрична гіпотеза може бути як простою, так і складною.

Приклад 3. Гіпотеза $H_0: F_1(x) = F_h(x)$ – проста, а гіпотеза $H_1: F_1(x) \neq F_h(x)$ – складна.

Зауваження 2. Як правило, для непараметричних складних гіпотез записати функцію вірогідності неможливо. Тому для перевірки непараметричних гіпотез застосовують методи, що не використовують відношення вірогідності. Розглянемо ці методи. Спочатку зупинимось на визначенні одного з найфундаментальніших понять сучасної математики – понятті метричного простору.

Визначення 2. *Метричним* простором називається множина S , для довільних елементів x, y якої визначена дійсна функція $\mu(x, y)$, що задовольняє таким постулатам Лінденбаума:

- 1) $\mu(x, y) = 0$ тоді і тільки тоді, коли $x = y$;
- 2) $\mu(x, y) \leq \mu(z, x) + \mu(z, y)$,

де x, y, z – довільні елементи множини S .

Функція $\mu(x, y)$ називається *метрикою* або *відстанню*.

Приклад 4. Метрикою є евклідова відстань $\mu(\vec{x}, \vec{y})$ між точками (векторами) $\vec{x} = (x_1, \dots, x_N)$ і $\vec{y} = (y_1, \dots, y_N)$ N -вимірного простору R^N , що визначається співвідношенням

$$\mu(\vec{x}, \vec{y}) = \sqrt{\sum_{n=1}^N (x_n - y_n)^2}.$$

Зауваження 1. Неметричні простори на практиці зустрічаються дуже часто.

Приклади 5. Температурно-часовий простір, у якому описується залежність температури від часу, є неметричним простором. Неметричним простором є множина точок, яка задається за допомогою базису, орти якого відповідають різним фізичним величинам (наприклад, масі, прискоренню й силі).

Властивості. Метрика має такі властивості:

- 1) $\mu(x, y) \geq 0$ (властивість невід'ємності);
- 2) $\mu(x, y) = \mu(y, x)$ (властивість комутативності).

Зауваження 2. На одній і тій самій множині S за допомогою різних метрик можна утворити різноманітні метричні простори.

Приклад 6. Нехай S –множина неперервних дійснозначних функцій $f(x)$, визначених на числовій осі. На множині цих функцій за допомогою метрик

$$\mu(f, g) = \max_x |f(x) - g(x)|,$$

$$\mu(f, g) = \int_{-\infty}^{\infty} |f(x) - g(x)| dx$$

утворюються два різних метричних простори.

Визначення 3. Для простої непараметричної гіпотези H_0 критерієм згоди D називається метрика μ , яка характеризує ступінь відмінності гіпотетичної функції розподілу $F_h(x)$ гіпотези H_0 від статистичної функції розподілу $F^*(x)$:

$$D = \mu[F_h(x), F^*(x)]. \quad (9.5.1)$$

Зауваження. Метрика (9.5.1) визначена на множині двох функцій, одна з яких ($F^*(x)$) випадкова. Тому критерій згоди D є випадковим функціоналом.

Ідея засобів перевірки непараметричних гіпотез. Нехай відомий закон розподілу випадкової величини D для простої основної гіпотези H_0 . Тоді, задавши певний рівень значимості α , можна визначити поріг d_α , що відповідає цьому значенню. Правило прийняття рішення таке (рис. 9.5.1): якщо у конкретному експерименті значення випадкової величини D менше порога d_α , то приймається рішення про справедливість основної гіпотези H_0 , у противному разі – про її несправедливість (гіпотеза H_0 відхиляється).

Зауваження 1. Описаний метод перевірки непараметричних гіпотез забезпечує прийняття рішення із заданою помилкою першого роду (рівнем значимості α), однак при цьому він не дозволяє проконтролювати величину помилки другого роду β . Це може приводити до ситуацій, коли із-за великої помилки другого роду рішення про справедливість гіпотези H_0 виявляється цілком необґрунтованим.

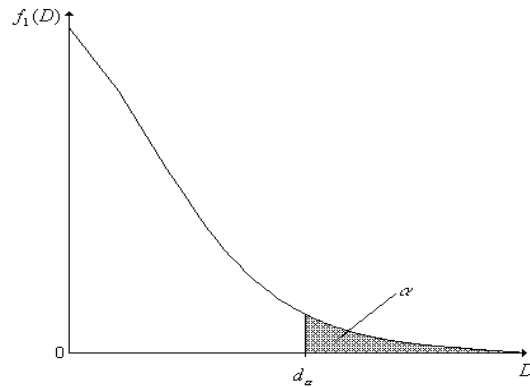


Рис. 9.5.1. Поріг d_α , відповідний рівню значимості α

Зауваження 2. Для зниження небезпеки прийняття необґрунтованого рішення звичайно рекомендують перевіряти рішення по декількох різноманітних критеріях, що ґрунтуються на використанні різних метрик.

Зауваження 3. Описаний метод перевірки простих непараметричних гіпотез природно узагальнюється на випадок складних гіпотез H_0 , які складаються з декількох простих гіпотез H_{0i} ($i = \overline{1, I}$). При цьому перевірка гіпотези H_0 зводиться до перевірки окремих гіпотез H_{0i} .

Зауваження 4. Основними недоліками методів перевірки непараметричних гіпотез є:

- 1) свавілля у виборі критерію;
- 2) свавілля у встановленні рівня значимості α і неконтрольованість рівня помилки другого роду β ;
- 3) можливість отримання прийнятних результатів лише при дуже великому обсязі вибірки, істотно більшому, ніж звичайно необхідно для отримання аналогічних результатів при перевірці параметричних гіпотез.

9.5.2. Критерій згоди Колмогорова

Визначення 1. Функція розподілу випадкової величини Z вигляду

$$K(z) = 1 - 2 \sum_{\nu=1}^{\infty} (-1)^{\nu-1} e^{-2\nu^2 z^2}, \quad z > 0 \quad (9.5.2)$$

називається *розподілом Колмогорова*.

Теорема. Нехай для обсягу вибірки N максимальне абсолютне відхилення статистичної функції розподілу $F^*(x)$ від гіпотетичної функції розподілу $F_h(x)$ дорівнює D :

$$D = \max_{x \in (-\infty, \infty)} |F^*(x) - F_h(x)|. \quad (9.5.3)$$

Тоді при $N \rightarrow \infty$ випадкова величина $Z = \sqrt{ND}$ незалежно від вигляду закону розподілу $F_h(x)$ має розподіл Колмогорова.

Визначення 2. Критерій згоди, що описується формулою (9.5.3), носить назву *критерію Колмогорова*.

Для розподілу Колмогорова є таблиці (додаток 10). При наближених розрахунках використовують той факт, що ряд (9.5.1) швидко збігається. Обмежуючись першим членом ряду, маємо

$$K(z) = 1 - 2e^{-2z^2}. \quad (9.5.4)$$

Тоді, задаючи величину допустимого рівня значимості $\alpha = 1 - K(z)$, легко з урахуванням співвідношення (9.5.4) отримати значення порога

$$d_\alpha = \sqrt{\frac{1}{2N} \ln \frac{2}{\alpha}}, \quad (9.5.5)$$

що відповідає цьому рівню.

Порядок використання критерію згоди Колмогорова такий:

1) по заданому рівню значимості α й обсягу вибірки N за допомогою таблиць розподілу або по формулі (9.5.5) визначають поріг d_α ;

2) по наявній вибірці x_1, \dots, x_N будують статистичну функцію розподілу $F^*(x)$;

3) для гіпотетичного розподілу $F_h(x)$, відповідного гіпотезі H_0 , і отриманого статистичного розподілу $F^*(x)$, визначають значення $d = \max_x |F^*(x) - F_h(x)|$;

4) проводять порівняння отриманої величини d з порогом d_α і по результатах порівняння приймають рішення про справедливість (при $d < d_\alpha$) або несправедливість (при $d \geq d_\alpha$) гіпотези H_0 .

Зауваження 1. Недоліком критерію Колмогорова є те, що величина D розбіжності між статистичною $F^*(x)$ і гіпотетичною $F_h(x)$ функціями розподілу нечутлива до невеликих, але дуже значимих розбіжностей в області «хвостів».

Зауваження 2. Недоліком критерію Колмогорова є також те, що він не враховує різну ймовірність появи у вибірці тих або інших значень.

Зауваження 3. Міжнародний стандарт ГОСТ 27.005-97 відзначає, що при використанні критерію Колмогорова, обсяг вибірки N повинен бути не менш 100, а число розрядів при розрахунку статистичної функції розподілу – не менш 30.

9.5.3. Критерій згоди омега-квадрат

Визначення. Критерієм згоди омега-квадрат називається критерій, що визначається співвідношенням

$$\Omega^2 = \int_{-\infty}^{\infty} (F^*(x) - F_h(x))^2 \varphi(F_h(x)) dF_h(x), \quad (9.5.6)$$

де $F^*(x)$ і $F_h(x)$ – відповідно статистична й гіпотетична функції розподілу, $\varphi(*)$ – вагова функція.

Теорема 1. Для одиничної вагової функції при спрямуванні обсягу вибірки N до нескінченності розподіл випадкової величини Ω^2 , що визначається

співвідношенням (9.5.6), незалежно від вигляду гіпотетичного розподілу $F_h(x)$ прямує до розподілу, який називається *омега-квадрат*.

Теорема 2. Для вагової функції $\varphi(F_h(x)) = [F_h(x)(1 - F_h(x))]^{-1}$ при спрямуванні обсягу вибірки N до нескінченності розподіл випадкової величини Ω^2 , що визначається співвідношенням (9.5.6), незалежно від вигляду гіпотетичного розподілу $F_h(x)$ прямує до розподілу, який називається *зваженим розподілом омега-квадрат*.

Зауваження 1. Порядок застосування цих критеріїв такий самий, як і критерію Колмогорова. Для полегшення розрахунків використовують такі наближені формули:

$$\Omega^2 = \frac{1}{12N^2} + \frac{1}{N} \sum_{n=1}^N \left(F_h(x_n) - \frac{2n-1}{2N} \right)^2, \quad (9.5.7)$$

$$\Omega^2 = -N - 2 \sum_{n=1}^N \left(\frac{2n-1}{2N} \ln F_h(x_n) + \left(1 + \frac{2n-1}{2N} \right) \ln(1 - F_h(x_n)) \right) \quad (9.5.8)$$

відповідно для розподілу омега-квадрат і зваженого розподілу омега-квадрат.

Таблиці розподілу омега-квадрат і зваженого розподілу омега-квадрат наведено відповідно у додатках 11 і 12.

Зауваження 2. ГОСТ 27.005-97 рекомендує використовувати критерій омега-квадрат вигляду (9.5.8). При цьому відзначається, що обсяг вибірки повинен бути не менш $N = 50$.

9.5.4. Критерій згоди хі-квадрат

Теорема 1. Нехай область значень випадкової величини X поділена на I інтервалів не обов'язково однакової довжини (рис. 9.5.2). І нехай ймовірність влучення випадкової величини X в i -й інтервал дорівнює P_i , а число елементів вибірки обсягу N , які потрапили в i -й інтервал, дорівнює l_i . Тоді при наближенні обсягу вибірки N до нескінченності незалежно від розподілу P_i функція

$$\chi^2 = \sum_{i=1}^I \frac{(l_i - NP_i)^2}{NP_i} \quad (9.5.9)$$

має розподіл хі-квадрат із $(I-1)$ ступенями свободи.

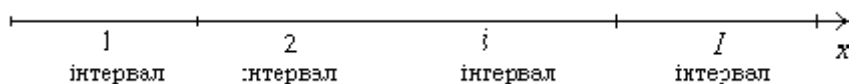


Рис. 9.5.2. Розподіл області значень випадкової величини X на I інтервалів

Теорема 2. Нехай виконуються всі умови теореми 1, але замість ймовірності P_i потрапляння випадкової величини X в інтервали використовуються оцінки $P_i(\theta_1^*, \dots, \theta_K^*)$, які являють собою функції випадкових точкових оцінок K параметрів $\theta_1^*, \dots, \theta_K^*$, котрі мають асимптотично гауссівський розподіл. Тоді при спрямованості обсягу вибірки N до нескінченності функція

$$\chi^2 = \sum_{i=1}^I \frac{[l_i - NP_i(\theta_1^*, \dots, \theta_K^*)]^2}{NP_i(\theta_1^*, \dots, \theta_K^*)}, \quad (9.5.10)$$

має розподіл хі-квадрат із $(I - K - 1)$ ступенями свободи.

Визначення 1. Критерії згоди, що описуються співвідношеннями (9.5.9) і (9.5.10), називаються *критеріями згоди хі-квадрат (критеріями Пірсона)*.

Зауваження 1. При використанні критерію згоди хі-квадрат слід дотримуватись таких рекомендацій ГОСТ 27.005-97:

- 1) обсяг вибірки N повинен бути не менш 100;
- 2) кількість інтервалів I треба вибирати від 7 до 25 (практика показує, що доцільно брати $I = 15 \div 18$ при $100 < N < 200$; $I = 18 \div 20$ при $N = 200$; $I = 25 \div 30$ при $N = 400$; $I = 35 \div 40$ при $N = 1000$);
- 3) вибирати величини інтервалів за умови, що $NP_i \geq 5$, де $i = \overline{1, I}$.

Зауваження 2. Порогове значення d_α , відповідне заданому рівню значимості α , визначається для критерію (9.5.9) із рівняння

$$\alpha = \int_{d_\alpha}^{\infty} f(\chi^2, I - 1) d\chi^2 \quad (9.5.11)$$

і для критерію (9.5.10) – із рівняння

$$\alpha = \int_{d_\alpha}^{\infty} f(\chi^2, I - K - 1) d\chi^2. \quad (9.5.12)$$

9.6. Перевірка гіпотези про незалежність випадкових величин

Дослідження системи випадкових величин (X_1, \dots, X_M) істотно спрощується, коли вони незалежні. Тому важливим завданням є перевірка гіпотези про незалежність випадкових величин, що досліджуються. При перевірці цієї гіпотези використовують той факт, що у випадку незалежності випадкових величин X_m ($m = \overline{1, M}$) спільна функція розподілу $F_M(x_1, \dots, x_M)$ факторизується:

$$F_M(x_1, \dots, x_M) = \prod_{m=1}^M F_1(x_m).$$

Розглянемо методику розв'язання цієї задачі на прикладі перевірки гіпотези про незалежність двох випадкових величин X і Y .

Постановка завдання. Нехай для системи випадкових величин (X, Y) із невідомою функцією розподілу $F(x, y)$ отримана вибірка $(x_1, y_1), \dots, (x_N, y_N)$ обсягом N .

Потрібно перевірити гіпотезу H_0 про незалежність випадкових величин X і Y , тобто перевірити, що $F(x, y) = F(x)F(y)$.

Розв'язання 1. Скористаємось критерієм хі-квадрат для функціонала $F^*(x, y) = F^*(x)F^*(y)$. Розіб'ємо області значень величини X і Y відповідно на I і J інтервалів (не обов'язково однакової довжини).

Вираз, аналогічний співвідношенню (9.5.9), має вигляд

$$\chi^2 = N \sum_{i=1}^I \sum_{j=1}^J \frac{\left(l_{ij} - \frac{l_i l_j}{N} \right)^2}{l_i l_j}, \quad (9.6.1)$$

де l_{ij} – кількість вибірових значень системи випадкових величин (X, Y) , які влучили в ij -й інтервал (i -й інтервал випадкової величини X і j -й інтервал випадкової величини Y); $l_i = l_{ix} = \sum_j l_{ij}$ – кількість вибірових значень випадкової величини X , які влучили в i -й інтервал; $l_j = l_{jy} = \sum_i l_{ij}$ – кількість вибірових значень випадкової величини Y , які влучили в j -й інтервал.

Доведено, що випадкова величина (9.6.1) при спрямованості обсягу вибірки N до нескінченності має хі-квадрат розподіл із $(I-1)(J-1)$ ступенями свободи.

Для зручності розрахунків звичайно використовують таблицю, яка називається *таблицею спряженості двох ознак* (табл. 9.6.1).

Таблиця 9.6.1

Таблиця спряженості двох ознак

Y	X				
	Δx_1	Δx_2	...	Δx_I	l_j
Δy_1	l_{11}	l_{12}	...	l_{1I}	l_{1y}
Δy_2	l_{21}	l_{22}	...	l_{2I}	l_{2y}
\vdots
Δy_J	l_{J1}	l_{J2}	...	l_{JI}	l_{Jy}
l_i	l_{ix}	l_{2x}	...	l_{Ix}	N

Зауваження. Порогове значення d_α , яке відповідає заданому рівню значимості α , визначається рівнянням

$$\alpha = \int_{d_\alpha}^{\infty} f(\chi^2, (I-1)(J-1)) d\chi^2.$$

Розв'язання 2. Використаємо критерій омега-квадрат із метрикою

$$\Omega^2 = \int_{-\infty}^{\infty} (F^*(x, y) - F^*(x)F^*(y))^2 dF^*(x, y). \quad (9.6.2)$$

Інтегруючи вираз (9.6.2), можна отримати

$$\Omega^2 = \frac{1}{N^5} \sum_{i=1}^N (l_{1i} l_{3i} - l_{2i} l_{4i})^2, \quad (9.6.3)$$

де $l_{1i}, l_{2i}, l_{3i}, l_{4i}$ – кількість вибірових значень, які лежать відповідно в першому, другому, третьому й четвертому квадрантах відносно точки (x_i, y_j) (рис. 9.6.1).

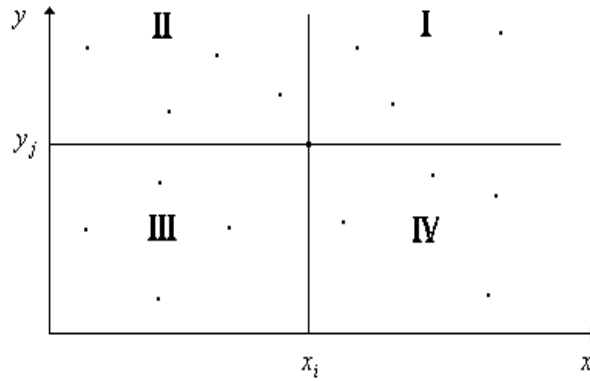


Рис. 9.6.1. Квадранти відносно точки (x_i, y_j)

Функція розподілу $F(b)$ випадкової величини $B = \frac{\pi^4}{2} N \Omega^2$ табульована. Відповідні таблиці наведені в додатку 13.

9.7. Перевірка гіпотези про приналежність вибірок одному розподілу

Визначення 1. Гіпотезою однорідності даних називається припущення про те, що M вибірок $(x_1^1, \dots, x_{N_1}^1), (x_1^2, \dots, x_{N_2}^2), \dots, (x_1^M, \dots, x_{N_M}^M)$ обсягом N_1, N_2, \dots, N_M випадкових величин X^1, X^2, \dots, X^M належать одному і тому ж розподілу $F(x)$ ($F(x) = F_1(x) = F_2(x) = \dots = F_M(x)$), тобто відповідають одній і тій самій випадковій величині X ($X = X^1 = X^2 = \dots = X^M$).

Для перевірки гіпотези однорідності даних використовують різноманітні критерії.

Розв'язання 1. Скористаємось критерієм хі-квадрат для функціонала $\sum_{m=1}^M (F^*(x^m) - F^*(x))$. Розіб'ємо область значень випадкових величин X^1, X^2, \dots, X^M на I інтервалів (рис. 9.7.1).

Вираз, аналогічний співвідношенню (9.5.10), має вигляд

$$\chi^2 = N \sum_{m=1}^M \sum_{i=1}^I \frac{\left(l_i^m - \frac{N_m l_i}{N} \right)^2}{N_m l_i}, \quad (9.7.1)$$

де l_i^m – кількість вибірових значень випадкової величини X^m , що потрапили в i -й інтервал; l_i – сумарна кількість вибірових значень випадкових величин X^m ($m = \overline{1, M}$), що потрапили в i -й інтервал; N – сумарний обсяг всіх вибірок:

$$N = \sum_{m=1}^M N_m.$$

Випадкова величина, що визначається формулою (9.7.1), має хі-квадрат розподіл із $(M-1)(I-1)$ ступенями свободи.

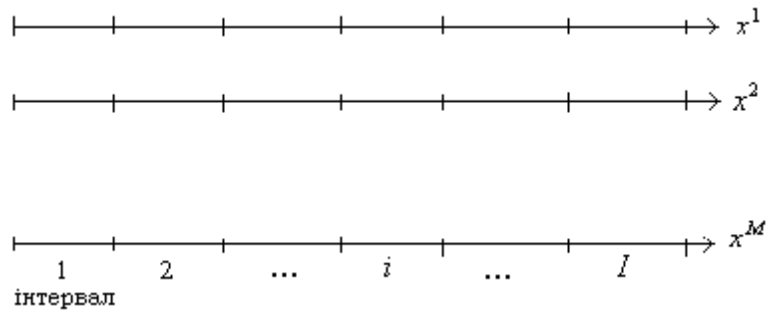


Рис. 9.7.1. Розбиття області значень випадкових величин X^1, X^2, \dots, X^M на I інтервалів

Розв'язання 2. Для перевірки однорідності двох вибірок ($M = 2$) використаємо критерій Смирнова для функціонала

$$D = \max_x |F_1^*(x) - F_2^*(x)|.$$

Доведено, що при спрямованості обсягу вибірок N_1 і N_2 до нескінченності величина $z = D \sqrt{\frac{N_1 + N_2}{N_1 N_2}}$ має розподіл Колмогорова (9.5.2). Приймаючи наближення (9.5.5), маємо

$$d_\alpha = \sqrt{\frac{N_1 N_2}{2(N_1 + N_2)}} \ln \frac{2}{\alpha}.$$

Розглянуті критерії згоди зведені в табл. 9.7.1.

Зауваження. Поряд із розглянутими критеріями використовують і інші, зокрема рангові критерії (Вілкоксона, Кендала, Спірмена), знакові критерії, критерії інверсій та ін..

Таблиця 9.7.1

Критерії згоди й відповідні пороги для заданого рівня значимості α й обсягу вибірки N

Назва критерію	Математичне формулювання критерію	Поріг
Критерій Колмогорова для простої гіпотези	$D = \max_x F^*(x) - F_h(x) $	$d_\alpha = \sqrt{\frac{1}{2N} \ln \frac{2}{\alpha}}$ таблиця (додаток 10)
Критерій омега-квадрат (Смирнова) для простої гіпотези	$\Omega^2 = \int_{-\infty}^{\infty} (F^*(x) - F_h(x))^2 dF_h(x)$ $\Omega^2 = \frac{1}{12N^2} + \frac{1}{N} \sum_{n=1}^N \left(F_h(x_n) - \frac{2n-1}{2N} \right)^2$	таблиця (додаток 11)
Критерій зваженого розподілу омега-квадрат (Смирнова) для простої гіпотези	$\Omega^2 = \int_{-\infty}^{\infty} (F^*(x) - F_h(x))^2 \varphi(F_h(x)) dF_h(x)$ $\Omega^2 = -1 - \frac{2}{N} \sum_{n=1}^N \left(\frac{2n-1}{2N} \ln F_h(x_n) + \left(1 - \frac{2n-1}{2N} \right) \ln(1 - F_h(x_n)) \right)$	таблиця (додаток 12)

Критерій хі-квадрат (Пірсона) для простої гіпотези	$\chi^2 = \sum_{i=1}^I \frac{(l_i - NP_i)^2}{NP_i}$ (розподіл із $(I - 1)$ ступенями свободи)	таблиця (додаток 8)
Критерій хі-квадрат (Фішера) для складної гіпотези	$\chi^2 = \sum_{i=1}^I \frac{[l_i - NP_i(\theta_1^*, \dots, \theta_K^*)]^2}{NP_i(\theta_1^*, \dots, \theta_K^*)}$ (розподіл із $(I - K - 1)$ ступенями свободи)	таблиця (додаток 8)
Критерій хі-квадрат для гіпотези про незалежність випадкових величин	$\chi^2 = N \sum_{i=1}^I \sum_{j=1}^J \frac{\left(l_{ij} - \frac{l_i l_j}{N} \right)^2}{l_i l_j}$ (розподіл із $(I - 1)(J - 1)$ ступенями свободи)	таблиця (додаток 8)
Критерій омега-квадрат для гіпотези про незалежність випадкових величин	$\Omega^2 = \int_{-\infty}^{\infty} (F^*(x, y) - F^*(x)F^*(y))^2 dF^*(x, y)$ $\Omega^2 = \frac{1}{N^5} \sum_{i=1}^N (l_{1i}l_{3i} - l_{2i}l_{4i})^2$	таблиця для $B = \frac{\pi^4}{2} N \Omega^2$ (додаток 13)
Критерій хі-квадрат для гіпотези про однорідність даних	$\chi^2 = N \sum_{m=1}^M \sum_{i=1}^I \frac{\left(l_i^m - \frac{N_m l_i}{N} \right)^2}{N_m l_i}$ (розподіл із $(M - 1)(I - 1)$ ступенями свободи)	таблиця (додаток 8)
Критерій Смирнова для гіпотези про однорідність даних	$D = \max_x F_1^*(x) - F_2^*(x) $	$d_\alpha = \sqrt{\frac{N_1 N_2}{2(N_1 + N_2)} \ln \frac{2}{\alpha}}$ (додаток 10)

Відомості про книги до розділу 9 наведені в табл. 9.7.2.

Таблиця 9.7.2

Література до розділу 9

Характер книг	Номери книг у списку літератури
Довідники	9, 11, 24, 47, 49, 68, 81, 93, 108
Книги навчального плану для інженерів	4, 13, 32, 42, 69, 75, 84, 91, 100
Книги з математичним ухилом	12, 18, 21, 44, 50, 58, 90, 101
Книги прикладного характеру	2, 8, 14, 19, 26, 27, 34, 36, 51, 52, 53, 55, 56, 57, 63, 64, 66, 76, 85, 92, 96, 97, 102, 103, 106

10. Оцінка характеристик випадкових функцій

10.1. Особливості застосування вибіркового методу при визначенні характеристик випадкових функцій

Визначення 1. Генеральною сукупністю випадкової функції $X(t)$ називається нескінченна множина реалізацій $x_n(t)$ ($n = 1, 2, \dots$) цієї випадкової функції.

Примітка 1. Множина реалізацій може бути як зліченою, так і незліченою. На практиці часто вважають (і у подальшому ми будемо вважати) її зліченою множиною.

Визначення 2. Елементами (членами) генеральної сукупності $X(t)$ називаються її реалізації $x_n(t)$ ($n = 1, 2, \dots$).

Визначення 3. Скінченна множина членів генеральної сукупності $x_n(t)$ ($n = \overline{1, N}$), отримана при кінцевому числі N дослідів, тобто вектор $\vec{x}(t) = (x_1(t), \dots, x_N(t))$, називається вибіркою з генеральної сукупності або просто вибіркою.

Зауваження. На практиці тривалість експериментів обмежена. Тому при статистичних дослідженнях випадкової функції $X(t)$ вибірка обсягом N являє собою сукупність відрізків реалізацій $x_n(t)$ тривалістю T .

Визначення 4. Множина різноманітних вибірок обсягу N , сформованих з однієї генеральної сукупності, являє собою N -вимірну векторну випадкову функцію $\vec{X}(t) = (X_1(t), \dots, X_N(t))$, яка називається *випадковою вибіркою*.

Зауваження 1. Слід відрізняти випадкову вибірку $\vec{X}(t)$ від простої вибірки $\vec{x}(t)$, що є реалізацією випадкової вибірки $\vec{X}(t)$.

Зауваження 2. Оскільки наслідки дослідів заздалегідь невідомі, то елементи випадкової вибірки можна вважати незалежними.

Положення. Для оцінок характеристик випадкових функцій, як і для випадкових величин, використовують *точкове і інтервальне оцінювання*.

Зауваження. Для точкових оцінок існують поняття незміщеності, спроможності, ефективності й достатності, які аналогічні розглянутим раніше для випадкових величин. Інтервальні оцінки характеризують розміром довірчого інтервалу і довірчою ймовірністю.

Отримання оцінки. Звичайно алгоритм отримання оцінки характеристики випадкової функції складається з таких етапів:

1. На основі наявних апріорних даних і даних вибірки $x_1(t), \dots, x_N(t)$ будують модель випадкової функції $X(t)$, що досліджується, (наприклад, приймають припущення, що випадкова функція стаціонарна).

2. Будують дискретний аналог випадкової функції $X(t)$. Для цього на інтервалі спостереження T задають моменти часу t_1, \dots, t_l ($t_1 = 0$; $t_l = T$) з рівномірним або нерівномірним кроком і випадкову функцію $X(t)$ ототожнюють з сукупністю I в загальному випадку залежних випадкових величин

$$X(t_1), X(t_2), \dots, X(t_l). \quad (10.1.1)$$

3. Перевіряють, чи не суперечить модель дослідним даним. Якщо модель визнається такою, що не суперечить ним, то вона приймається. У противному

випадку її уточнюють, і знов прийняті припущення перевіряють для дискретного наближення (10.1.1).

4. На основі оцінки характеристик дискретного наближення випадкової функції будують наближену оцінку шуканої характеристики $\Theta^*(t_1, \dots, t_M)$.

10.2. Оцінка математичного сподівання випадкової функції

10.2.1. Оцінка математичного сподівання випадкової функції по ансамблю реалізацій

Нехай у результаті досліду отримана вибірка $x_n(t)$ ($n = \overline{1, N}$) з N реалізацій довільної випадкової функції $X(t)$ (рис. 10.2.1). Розглянемо оцінку $m_x^*(t)$ математичного сподівання цієї випадкової функції по ансамблю реалізацій $x_n(t)$:

$$m_x^*(t) = \frac{1}{N} \sum_{n=1}^N x_n(t). \quad (10.2.1)$$

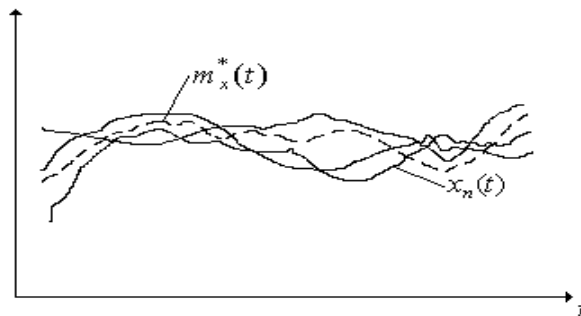


Рис. 10.2.1. Реалізації $x_n(t)$ ($n = \overline{1, N}$) і оцінка $m_x^*(t)$ математичного сподівання нестационарної випадкової функції $X(t)$

Якщо існують перші два моменти випадкової функції $X(t)$, то оцінка (10.2.1) має такі властивості.

Властивість 1. Оцінка (10.2.1) – незміщена.

Справедливість цього твердження випливає з того, що

$$M[m_x^*(t)] = \frac{1}{N} \sum_{n=1}^N M[X_n(t)] = m_x(t),$$

де $m_x(t)$ – математичне сподівання випадкової функції $X(t)$.

Властивість 2. Оцінка (10.2.1) – спроможна.

Справедливість цього твердження випливає з того, що

$$1) \quad D[m_x^*(t)] = \frac{1}{N} D_x(t), \text{ де } D_x(t) \text{ – дисперсія випадкової величини } X(t);$$

$$2) \quad \lim_{N \rightarrow \infty} D[m_x^*(t)] = 0 \text{ (це означає, що } m_x^*(t) \text{ збігається з } m_x(t) \text{ у середньоквадратичному);}$$

3) збіжність у середньоквадратичному гарантує збіжність за ймовірністю:

$$\lim_{N \rightarrow \infty} P\{|m_x^*(t) - m_x(t)| > \varepsilon\} = 0.$$

Властивість 3. Оцінка (10.2.1) – не обов'язково ефективна.

Приклад. Оцінка неефективна, якщо відомо, що випадкова функція $X(t)$ стаціонарна. Тоді дисперсія оцінки $D[m_x^*(t)]$ може бути зменшена за рахунок усереднення по аргументу t .

10.2.2. Оцінка математичного сподівання випадкової функції по одній реалізації

Нехай у результаті досліду отримана вибірка $x(t)$ (рис. 10.2.2) стаціонарної випадкової функції, що має властивість ергодичності ($t \in [0, T]$).

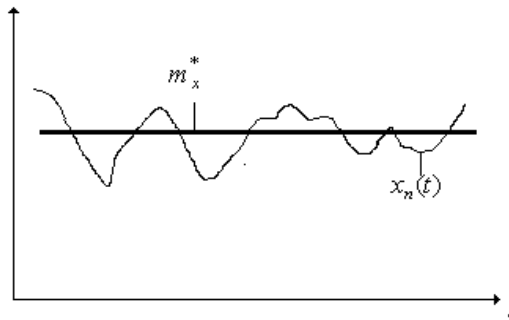


Рис. 10.2.2. Реалізація $x(t)$ і оцінка m_x^* математичного сподівання стаціонарної випадкової функції $X(t)$

Розглянемо оцінку m_x^* математичного сподівання цієї випадкової функції по єдиній реалізації:

$$m_x^* = \frac{1}{T} \int_0^T x(t) dt. \quad (10.2.2)$$

Якщо існують перші два моменти випадкової функції $X(t)$ та кореляційна функція $R_x(t_1, t_2) = R_x(t_2 - t_1)$ інтегровна, то оцінка має такі властивості.

Властивість 1. Оцінка (10.2.2) – незміщена. Справедливість цього твердження випливає з того, що $M[m_x^*] = \frac{1}{T} \int_0^T M[X(t)] dt = m_x$, де m_x – математичне сподівання випадкової функції $X(t)$.

Властивість 2. Оцінка (10.2.2) – спроможна. Справедливість цього твердження випливає з того, що

$$1) D[m_x^*] = \frac{1}{T^2} \int_0^T \int_0^T R_x(t_1, t_2) dt_1 dt_2 = \frac{2}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) R_x(\tau) d\tau, \quad (10.2.3)$$

де $R_x(\tau)$ – кореляційна функція випадкової функції $X(t)$;

$$2) \lim_{T \rightarrow \infty} \frac{2}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) R_x(\tau) d\tau = 0 \text{ для ергодичних випадкових функцій;}$$

3) збіжність у середньоквадратичному гарантує збіжність за ймовірністю:

$$\lim_{N \rightarrow \infty} P\{|m_x^*(t) - m_x(t)| > \varepsilon\} = 0.$$

Властивість 3. Оцінка (10.2.2) – ефективна в класі лінійних оцінок, тобто серед оцінок вигляду

$$m_x^* = \frac{1}{T} \int_0^T g(t)x(t) dt.$$

Оцінка (10.2.2) має мінімальну дисперсію.

Властивість 4. Дисперсія оцінки (10.2.2) описується співвідношенням

$$D[m_x^*] = \frac{2D_x}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) r(\tau) d\tau,$$

де $r(\tau)$ – коефіцієнт кореляції. Звідси видно, що при збільшенні тривалості T дисперсія оцінки зменшується (рис. 10.2.3).

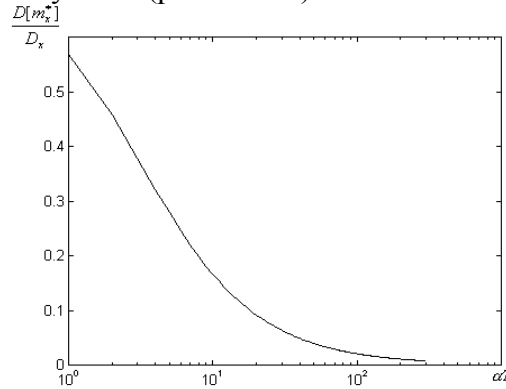


Рис. 10.2.3. Залежність нормованої дисперсії $\frac{D[m_x^*]}{D_x}$ оцінки m_x^* математичного сподівання від тривалості реалізації T (коефіцієнт кореляції $r(\tau) = e^{-\alpha|\tau|}$)

10.2.3. Похибка дискретизації при отриманні оцінки математичного сподівання стаціонарної випадкової функції

Нехай для однієї реалізації стаціонарної випадкової функції $X(t)$, що має властивість ергодичності, в результаті проведення дослідження отримані в дискретній множині точок $t_1 = 0; t_2 = \Delta t, \dots, t_I = (I-1)\Delta t$ значення $x(t_i)$ ($i = \overline{1, I}$) (рис. 10.2.4).

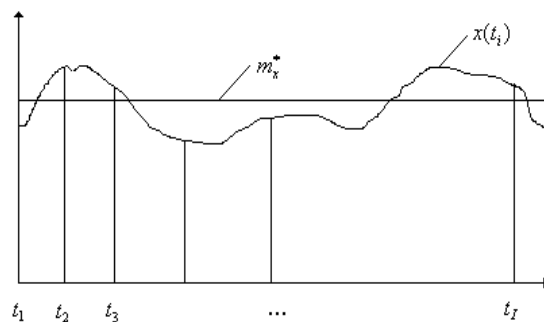


Рис. 10.2.4. Множина значень $x(t_i)$ реалізації стаціонарної випадкової функції $X(t)$ у дискретній множині точок і оцінка m_x^* її математичного сподівання

Розглянемо оцінку m_x^* математичного сподівання випадкової функції, обчислену по множині значень $x(t_i)$:

$$m_x^* = \frac{1}{I} \sum_{i=1}^I x(t_i). \quad (10.2.4)$$

Примітка. Завдання такого роду часто виникає при комп'ютерній обробці експериментальних даних.

Якщо існують перші два моменти випадкової функції $X(t)$ та кореляційна функція $R_x(t_2 - t_1)$ інтегровна, то оцінка (10.2.4) має такі властивості.

Властивість 1. Оцінка (10.2.4) – незміщена, тому що

$$M[m_x^*] = \frac{1}{I} \sum_{i=1}^I M[X(t_i)] = m_x.$$

Властивість 2. Оцінка (10.2.4) – спроможна.

Ця властивість впливає з таких причин:

1) у дискретному випадку вираз (10.2.3) може бути записаний так:

$$D[m_x^*] = \frac{2}{I} \int_0^T \left(1 - \frac{\tau}{T}\right) R_x(\tau) d\tau,$$

де $R_x(\tau) = D_x r_x(\tau) \sum_{i=1}^I \delta(t_i - \tau)$, D_x – дисперсія випадкової функції $X(t)$, $r_x(\tau)$ – кореляційна функція випадкової функції $X(t)$, $\delta(t)$ – дельта-функція, $R_x(\tau)$ – кореляційна функція випадкової функції $X(t)$, звідки

$$D[m_x^*] = \frac{D_x}{I} \left[1 + 2 \sum_{i=1}^{I-1} \left(1 - \frac{i}{I}\right) r_x(i\Delta t) \right];$$

2) $\lim_{I \rightarrow \infty} D[m_x^*] = 0$ для ергодичної випадкової функції;

3) збіжність у середньоквадратичному гарантує збіжність за ймовірністю.

Властивість 3. Для конкретних умов (конкретної кореляційної функції) при фіксованій тривалості реалізації $T = (I-1)\Delta t$ дисперсія $D[m_x^*(t)]$ оцінки зменшується зі збільшенням кількості відліків I .

Зауваження. Розрахунки показують (рис. 10.2.5), що при зростанні I на інтервалі (10, 100) швидкість зменшення дисперсії помилки в ряді випадків невелика. Тому є сенс вибирати I не більш $15 \div 20$.

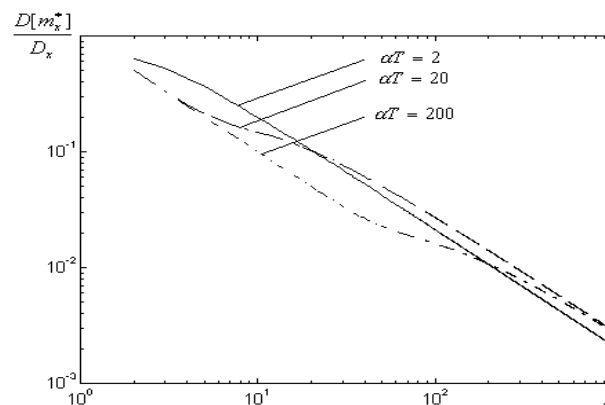


Рис. 10.2.5. Залежність нормованої дисперсії $\frac{D[m_x^*]}{D_x}$ оцінки математичного сподівання від кількості значень реалізації I , використаних при обчисленні оцінки (коефіцієнт кореляції $r(\tau) = e^{-\alpha|\tau|}$)

10.2.4. Оцінка математичного сподівання стаціонарної випадкової функції по аргументу й множині реалізацій

Нехай у результаті досліду для стаціонарної випадкової функції $X(t)$, що має властивість ергодичності, отримана вибірка $x_n(t)$ ($t \in [0, T]$, $n = \overline{1, N}$) з N реалізацій (рис. 10.2.6).

Розглянемо оцінку m_x^* математичного сподівання цієї випадкової функції $X(t)$ по аргументу t і множині реалізацій:

$$m_x^* = \frac{1}{N} \sum_{n=1}^N m_{xn}^* \quad (10.2.5)$$

$$\text{де } m_{xn}^* = \frac{1}{T} \int_0^T x_n(t) dt.$$

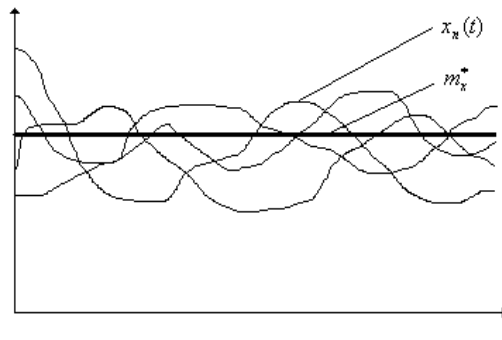


Рис. 10.2.6. Реалізації $x_n(t)$ ($n = \overline{1, N}$) і оцінка математичного сподівання m_x^* стаціонарної випадкової функції $X(t)$

Якщо існують перші два моменти випадкової функції $X(t)$ та кореляційна функція $R_x(t_2 - t_1)$ інтегровна, то оцінка (10.2.5) має такі властивості.

Властивість 1. Оцінка (10.2.5) – незміщена.

Ця властивість випливає з того, що

$$M[m_x^*] = \frac{1}{NT} \sum_{n=1}^N \int_0^T M[X(t)] dt = m_x,$$

де m_x – математичне сподівання випадкової функції $X(t)$.

Властивість 2. Оцінка (10.2.5) – спроможна.

Ця властивість випливає з таких причин:

$$1) \quad D[m_x^*] = \frac{2}{NT} \int_0^T \left(1 - \frac{\tau}{T}\right) R_x(\tau) d\tau, \quad \text{де } R_x(\tau) \text{ – кореляційна функція}$$

випадкової функції $X(t)$;

$$2) \quad \lim_{T \rightarrow \infty} D[m_x^*] = 0;$$

3) збіжність у середньоквадратичному гарантує збіжність за ймовірністю.

Властивість 3. Оцінка (10.2.5) – ефективна в класі лінійних оцінок.

Зауваження. В тому випадку, коли тривалості реалізацій T_n різні, доцільно використовувати оцінку

$$m_x^* = \frac{1}{N} \frac{\sum_{n=1}^N \alpha_n m_{xn}^*}{\sum_{n=1}^N \alpha_n},$$

де $\alpha_n = \frac{1}{D_{x_n}}$, $D_{x_n} = D[m_{x_n}^*]$ – дисперсія оцінки $m_{x_n}^*$.

Дисперсія цієї оцінки

$$D[m_x^*] = \frac{1}{\sum_{n=1}^N \alpha_n}.$$

10.2.5. Оцінка математичного сподівання нестационарної випадкової функції по одній реалізації

Нехай досліджується нестационарна випадкова функція $X(t)$, математичне сподівання якої змінюється повільно, причому на довільному інтервалі $2T_0$ практично за лінійним законом:

$$m_x(\tau) \approx m_x(t) + m'_x(t)(\tau - t), \quad t - T_0 < \tau < t + T_0. \quad (10.2.6)$$

В результаті досліду такої випадкової функції на інтервалі $T > 2T_0$ отримана одна реалізація $x(t)$.

Розглянемо оцінку $m_x^*(t)$ математичного сподівання, що являє собою усереднену на ковзному інтервалі $2T_0$ реалізацію $x(t)$ випадкової функції $X(t)$:

$$m_x^*(t) = \frac{1}{2T_0} \int_{t-T_0}^{t+T_0} x(\tau) d\tau. \quad (10.2.7)$$

Примітка. Метод обчислення такої оцінки називається *методом ковзного середнього*. Цей метод був розроблений В. С. Пугачовим.

Властивість. Якщо існують перші два моменти випадкової функції $X(t)$ та кореляційна функція $R_x(t_1, t_2)$ подвійно інтегровна, то дисперсія $D[m_x^*(t)]$ оцінки $m_x^*(t)$ дорівнює

$$D[m_x^*(t)] = M[(m_x^*(t) - m(t))^2] = \frac{1}{4T_0^2} \int_{t-T_0}^{t+T_0} \int_{t-T_0}^{t+T_0} R_x(\tau, \tau') d\tau d\tau'. \quad (10.2.8)$$

Таблиця 10.2.1

Різноманітні оцінки математичного сподівання й властивості цих оцінок

Назва оцінки	Математичне формулювання	Властивість оцінки	Примітка
Оцінка математичного сподівання по ансамблю	$m_x^* = \frac{1}{N} \sum_{n=1}^N x_n(t)$	1) незміщена 2) спроможна 3) не обов'язково ефективна	для будь-яких ВФ
Оцінка математичного сподівання по одній реалізації	$m_x^* = \frac{1}{T} \int_0^T x(t) dt$	1) незміщена 2) спроможна 3) ефективна в класі лінійних оцінок	для ергодичних ВФ
Оцінка математичного сподівання по одній реалізації (дискретний випадок)	$m_x^* = \frac{1}{I} \sum_{i=1}^I x(t_i)$	1) незміщена 2) спроможна 3) збільшення I понад 15 – 20 інколи недоцільно	для ергодичних ВФ

Оцінка математичного сподівання по аргументу й множині реалізацій (неперервний випадок)	$m_x^* = \frac{1}{N} \sum_{n=1}^N m_{x_n}^*$ де $m_{x_n}^* = \frac{1}{T} \int_0^T x_n(t) dt$	1) незміщена 2) спроможна	для ергодичних ВФ
Оцінка математичного сподівання по одній реалізації (метод ковзного середнього)	$m_x^*(t) = \frac{1}{2T_0} \int_{t-T_0}^{t+T_0} x(\tau) d\tau$	$D[m_x^*(t)] = \frac{1}{4T_0^2} \int_{t-T_0}^{t+T_0} \int_{t-T_0}^{t+T_0} R_x(\tau, \tau') d\tau d\tau'$	для нестационарної ВФ, для якої $m_x(\tau) \approx m_x(t) + m_x'(t)(\tau - t)$, $t - T_0 < \tau < t + T_0$

Різноманітні оцінки математичного сподівання й властивості цих оцінок зведені в табл. 10.2.1.

10.3. Оцінка дисперсії випадкової функції

10.3.1. Оцінка дисперсії випадкової функції по ансамблю реалізацій

Нехай є вибірка з N реалізацій $x_n(t)$ довільної випадкової функції $X(t)$ при невідомому математичному сподіванні $m_x(t)$. По такій вибірці можна отримати дві оцінки:

$$D_x^*(t) = \frac{1}{N} \sum_{n=1}^N [x_n(t) - m_x^*(t)]^2, \quad D_x^*(t) = \frac{1}{N-1} \sum_{n=1}^N [x_n(t) - m_x^*(t)]^2,$$

де $m_x^*(t)$ – оцінка математичного сподівання.

Зауваження 1. Якщо існують перші два моменти випадкової функції, то перша оцінка зміщена, друга – ні. При цьому обидві оцінки спроможні.

Зауваження 2. Для гауссівської випадкової функції дисперсії цих двох оцінок відповідно такі:

$$D[D_x^*(t)] = \frac{2}{N} D_x^2(t), \quad D[D_x^*(t)] = \frac{2}{N-1} D_x^2(t),$$

де $D_x(t)$ – дисперсія випадкової функції $X(t)$.

10.3.2. Оцінка дисперсії випадкової функції по одній реалізації

Нехай є одна реалізація $x(t)$ ергодичної випадкової функції $X(t)$. По цій реалізації при невідомому математичному сподіванні m_x можна обчислити оцінку дисперсії:

$$D_x^* = \frac{1}{T} \int_0^T (x(t) - m_x^*)^2 dt, \quad (10.3.1)$$

де m_x^* – оцінка математичного сподівання.

Якщо існують перші два моменти випадкової функції $X(t)$ та кореляційна функція інтегровна, то оцінка має такі властивості.

Властивість 1. Оцінка (10.3.1) – зміщена на величину $D[m_x^*]$.

У справедливості цього твердження можна переконатися шляхом перетворення формули (10.3.1) до вигляду

$$D_x^* = \frac{1}{T} \int_0^T (x(t) - m_x)^2 dt - (m_x^* - m_x)^2, \quad (10.3.2)$$

а після цього – обчислення математичного сподівання від обох частин виразу (10.3.2).

Властивість 2. Оцінка (10.3.1) спроможна для гауссівського розподілу $X(t)$.

Властивість 3. При гауссівському розподілі $X(t)$ дисперсія оцінки описується формулою

$$D[D_x^*] = \frac{4}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) R_x^2(\tau) d\tau. \quad (10.3.3)$$

Звідси видно, що при збільшенні тривалості інтервалу T дисперсія $D[D_x^*]$ зменшується (рис. 10.3.1).

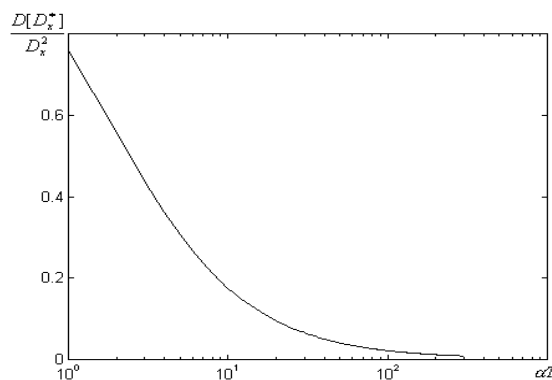


Рис. 10.3.1. Залежність відносної дисперсії $\frac{D[D_x^*]}{D_x^2}$ оцінки дисперсії D_x^* від тривалості

реалізації (коефіцієнт кореляції $r(\tau) = e^{-\alpha|\tau|}$)

10.3.3. Похибки дискретизації при оцінці дисперсії

Нехай для однієї реалізації стаціонарної випадкової функції $X(t)$, що має властивість ергодичності, в результаті проведення дослідів отримані значення $x(t_i)$ в дискретній множині точок $t_1 = 0; t_2 = \Delta t, \dots, t_l = (l-1)\Delta t$ (рис. 10.2.4).

По цих значеннях можна отримати оцінку дисперсії:

$$D_x^* = \frac{1}{l} \sum_{i=1}^l [x(t_i) - m_x]^2, \quad (10.3.4)$$

де m_x – значення математичного сподівання.

Якщо існують перші два моменти випадкової функції та кореляційна функція інтегровна, то оцінка має такі властивості.

Властивість 1. Оцінка (10.3.4) – незміщена.

Зауваження. Аналогічна оцінка дисперсії, що використовує оцінку математичного сподівання m_x^* , є зміщеною.

Властивість 2. Оцінка (10.3.4) – спроможна.

Властивість 3. Дисперсія оцінки (10.3.4) визначається співвідношенням

$$D[D_x^*] = \frac{2D_x^2}{I} \left[1 + 2 \sum_{i=1}^{I-1} \left(1 - \frac{i}{I} \right) r_x^2(i\Delta t) \right],$$

аналогічним співвідношенню (10.3.3).

Зауваження 1. Для конкретних умов (конкретної кореляційної функції і фіксованої тривалості реалізації $T = (I-1)\Delta t$) дисперсія $D[D_x^*]$ оцінки зменшується зі збільшенням кількості відліків I (рис. 10.3.2).

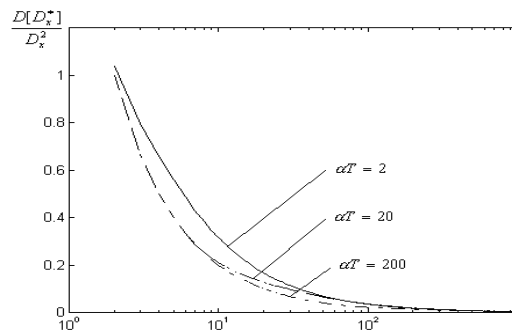


Рис. 10.3.2. Залежність відносної дисперсії $\frac{D[D_x^*]}{D_x^2}$ оцінки дисперсії D_x^* від кількості відліків

I , які використовуються для обчислення оцінки (коефіцієнт кореляції $r(\tau) = e^{-\alpha|\tau|}$)

Зауваження 2. Розрахунки показують, що на інтервалі (10, 100), як і у випадку оцінки математичного сподівання, при зростанні I інколи відбувається повільне зменшення дисперсії. Тому збільшення кількості відліків I понад 20 часто не має сенсу.

Різноманітні оцінки дисперсії випадкової функції при невідомому математичному сподіванні зведені в табл. 10.3.1.

Таблиця 10.3.1

Різноманітні оцінки дисперсії випадкової функції при невідомому математичному сподіванні

Назва оцінки	Математичне формулювання	Властивість оцінки	Примітка
Оцінка дисперсії по ансамблю	$D_x^*(t) = \frac{1}{N} \sum_{n=1}^N [x_n(t) - m_x^*(t)]^2$	1) зміщена 2) спроможна	для будь-якої ВФ
	$D_x^*(t) = \frac{1}{N-1} \sum_{n=1}^N [x_n(t) - m_x^*(t)]^2$	1) незміщена 2) спроможна	для будь-якої ВФ
Оцінка дисперсії по одній реалізації (неперервний випадок)	$D_x^* = \frac{1}{T} \int_0^T (x(t) - m_x^*)^2 dt$	1) зміщена на $D[m_x^*]$ 2) спроможна при гауссівському розподілі	для ергодичної ВФ
Оцінка дисперсії по одній реалізації (дискретний випадок)	$D_x^* = \frac{1}{I} \sum_{i=1}^I (x(t_i) - m_x^*)^2$	1) зміщена на $D[m_x^*]$ 2) спроможна при гауссівському розподілі 3) збільшення I понад 20 інколи не має сенсу	для ергодичної ВФ

10.4. Оцінка кореляційної функції

10.4.1. Оцінка кореляційної функції по ансамблю реалізацій

Нехай є вибірка з N реалізацій $x_n(t)$ довільної випадкової функції $X(t)$. При відомому математичному сподіванні $m_x(t)$ оцінка кореляційної функції $R_x^*(t_1, t_2)$ може бути обчислена так:

$$R_x^*(t_1, t_2) = \frac{1}{N} \sum_{n=1}^N (x_n(t_1) - m_x(t_1))(x_n(t_2) - m_x(t_2)),$$

а при невідомому математичному сподіванні $m_x(t)$ –

$$R_x^*(t_1, t_2) = \frac{1}{N-1} \sum_{n=1}^N (x_n(t_1) - m_x^*(t_1))(x_n(t_2) - m_x^*(t_2)),$$

де $m_x^*(t)$ – оцінка математичного сподівання.

Зауваження. Обидві оцінки є незміщеними і спроможними.

10.4.2. Оцінка кореляційної функції стаціонарної випадкової функції

Нехай є одна реалізація $x(t)$ ергодичної випадкової функції $X(t)$ ($t \in [0, T]$). Потрібно оцінити кореляційну функцію цієї випадкової функції. Розглянемо два випадки.

Випадок 1. Математичне сподівання m_x відомо.

Тоді як оцінки можуть бути використані два вирази:

$$R_{x_1}^*(\tau) = \frac{1}{T} \int_0^{T-|\tau|} x(t) x(t+|\tau|) dt, \quad 0 \leq |\tau| \leq T, \quad (10.4.1)$$

$$R_{x_2}^*(\tau) = \frac{1}{T-|\tau|} \int_0^{T-|\tau|} x(t) x(t+|\tau|) dt, \quad 0 \leq |\tau| \leq T, \quad (10.4.2)$$

де $\overset{0}{x}(t) = x(t) - m_x$.

Визначення. Оцінки вигляду (10.4.1) і (10.4.2) називаються *вибірковими кореляційними функціями*.

Якщо існують перші два моменти випадкової функції $X(t)$ та кореляційна функція $R_x(\tau)$ інтегровна, то оцінки (10.4.1) і (10.4.2) мають такі властивості.

Властивість 1. Оцінка (10.4.1) – зміщена на величину $\frac{|\tau|}{T} R_x(\tau)$. При $T \rightarrow \infty$ вона стає незміщеною (тобто оцінка (10.4.1) асимптотично незміщена). Оцінка (10.4.2) – незміщена.

Властивість 2. При гауссівському розподілі випадкової функції $X(t)$ обидві оцінки є спроможними.

Властивість 3. При гауссівському розподілі випадкової функції $X(t)$ дисперсії оцінок (10.4.1) і (10.4.2) описуються, відповідно, виразами

$$D[R_{x_1}^*(\tau)] = \frac{1}{T} \int_{-(T-\tau)}^{T-\tau} \left(1 - \frac{\tau + |t|}{T}\right) [R_x^2(t) + R_x(t+\tau)R_x(t-\tau)] dt, \quad (10.4.3)$$

$$D[R_{x_2}^*(\tau)] = \frac{1}{T(1-T^{-2}|\tau|^2)} \int_{-(T-\tau)}^{T-\tau} \left(1 - \frac{\tau + |t|}{T}\right) [R_x^2(t) + R_x(t+\tau)R_x(t-\tau)] dt. \quad (10.4.4)$$

Зауваження. При великих T дисперсії оцінок $D[R_{x_1}^*(\tau)]$ і $D[R_{x_2}^*(\tau)]$ обернено пропорційні T .

Властивість 4. При $\tau \rightarrow T$ дисперсія зміщеної оцінки $D[R_{x_1}^*(\tau)]$ наближається до нуля, а дисперсія незміщеної оцінки $D[R_{x_2}^*(\tau)]$ – до нескінченності (рис. 10.4.1).

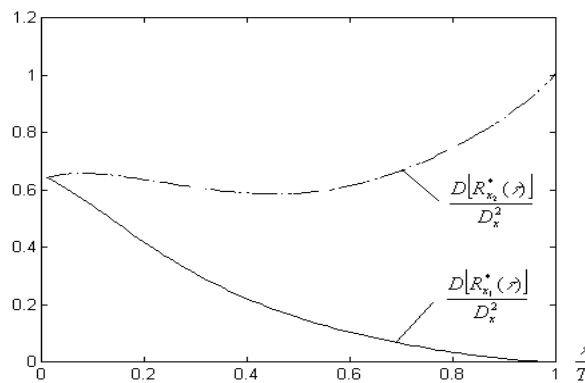


Рис. 10.4.1. Залежність відносної дисперсії зміщеної $R_{x_1}^*(\tau)$ і незміщеної $R_{x_2}^*(\tau)$ оцінок кореляційної функції $R_x(\tau)$ від величини τ (коефіцієнт кореляції $r_x(\tau) = e^{-\alpha|\tau|}$, $\alpha T = 5/2$)

Зауваження. Властивість 4 часто робить незміщену оцінку (10.4.2) непридатною для застосування. Тому звичайно віддають перевагу зміщеній оцінці (10.4.1).

Випадок 2. Математичне сподівання m_x невідомо.

В цьому випадку як оцінку кореляційної функції доцільно брати

$$R_x^*(\tau) = \frac{1}{T} \int_0^{T-|\tau|} (x(t) - m_x^*) (x(t+|\tau|) - m_x^*) dt, \quad (10.4.5)$$

де m_x^* – оцінка математичного сподівання.

Властивість. Якщо існують перші два моменти випадкової функції $X(t)$ та кореляційна функція інтегровна, то оцінка (10.4.5) – зміщена. Причому величина зміщення виявляється більшою, ніж у випадку оцінки (10.4.1).

Зауваження. При гауссівському розподілі оцінка (10.4.5) спроможна.

10.4.3. Оцінка кореляційної функції стаціонарної випадкової функції по дискретних відліках

Нехай для однієї реалізації стаціонарної випадкової функції $X(t)$, що має властивість ергодичності, в результаті проведення досліду отримані значення $x(t_i)$ ($t_i = (i-1)\Delta t$, $i = \overline{1, I}$).

При апріорно невідомому математичному сподіванні m_x як оцінку доцільно брати

$$R_x^*(l\Delta t) = \frac{1}{I} \sum_{i=1}^{I-l} (x((i-1)\Delta t) - m_x^*)(x((i+l-1)\Delta t) - m_x^*), \quad (l = 0, I-1), \quad (10.4.6)$$

де m_x^* – оцінка математичного сподівання.

Якщо існують перші два моменти випадкової функції $X(t)$ та кореляційна функція інтегровна, то оцінка (10.4.6) має такі властивості.

Властивість 1. Оцінка (10.4.6) – зміщена.

Властивість 2. Оцінка (10.4.6) – спроможна.

Властивість 3. Дисперсія оцінки (10.4.6) при гауссівському розподілі $X(t)$ і великій кількості відліків I має вигляд

$$D[R_x^*(l\Delta t)] \approx \frac{1}{I} \sum_{i=-\infty}^{+\infty} [R_x^{*2}((i-1)\Delta t) + R_x^*((i+l-1)\Delta t)R_x^*((i-l-1)\Delta t)]. \quad (10.4.7)$$

З виразу (10.4.7) видно, що при збільшенні кількості відліків I дисперсія оцінки зменшується.

Різні оцінки кореляційної функції випадкової функції при невідомому математичному сподіванні m_x зведені в табл. 10.4.1, де $x^*(t) = x(t) - m_x^*$.

Таблиця 10.4.1

Різні оцінки кореляційної функції випадкової функції при невідомому математичному сподіванні

Назва оцінки	Математичне формулювання	Властивість оцінки	Примітка
Оцінка кореляційної функції по ансамблю	$R_x^*(t_1, t_2) = \frac{1}{N} \sum_{n=1}^N x_n^*(t_1) x_n^*(t_2)$	1) зміщена 2) спроможна	для будь-якої ВФ
	$R_x^*(t_1, t_2) = \frac{1}{N-1} \sum_{n=1}^N x_n^*(t_1) x_n^*(t_2)$	1) незміщена 2) спроможна	для будь-якої ВФ
Оцінка кореляційної функції по одній реалізації (неперервний випадок)	$R_x^*(t_1, t_2) = \frac{1}{T} \int_0^{T-\tau_0} x^*(t) x^*(t+\tau) d\tau$	1) зміщена 2) спроможна при гауссівському розподілі	для ергодичних ВФ
Оцінка кореляційної функції по одній реалізації (дискретний випадок)	$R_x^*(l\Delta t) = \frac{1}{I} \sum_{i=1}^{I-l} x_n^*((i-1)\Delta t) x_n^*((i+l-1)\Delta t)$	1) зміщена 2) спроможна при гауссівському розподілі	для ергодичних ВФ

10.5. Оцінка спектральної щільності стаціонарної випадкової функції

Нехай є одна реалізація $x(t)$ стаціонарної ергодичної випадкової функції $X(t)$ ($t \in [0, T]$) із нульовим математичним сподіванням.

Миттєвий математичний спектр цієї реалізації має вигляд

$$\dot{A}(\omega) = \int_0^T x(t)e^{-j\omega t} dt.$$

Визначення. Вибірковою спектральною щільністю або періодограмою $S_T(\omega)$ називається миттєвий спектр потужності реалізації $x(t)$, нормований на величину інтервалу спостереження T :

$$S_T(\omega) = \frac{1}{T} |\dot{A}(\omega)|^2.$$

Зауваження. Періодограма $S_T(\omega)$ зв'язана з вибірковою кореляційною функцією $R_x^*(\tau)$ перетворенням Фур'є:

$$S_T(\omega) = \int_{-T}^T R_x^*(\tau) e^{-j\omega\tau} d\tau. \quad (10.5.1)$$

Теорема 1. Нехай $\Theta^*(t)$ – спроможна оцінка деякої статистичної характеристики $\theta(t)$. Тоді перетворення Фур'є цієї оцінки $F[\Theta^*(t)]$ не є спроможною оцінкою функції $F[\theta(t)]$, яка являє собою перетворення Фур'є цієї характеристики.

Зауваження. Оцінка (10.5.1) не є спроможною оцінкою спектральної щільності потужності $S_x(\omega)$. В результаті цього при будь-якому фіксованому ω значення оцінки $S_T(\omega)$ сильно коливається від реалізації до реалізації, причому збільшення часу спостереження T не приводить до зменшення дисперсії оцінки.

Положення. Спроможні оцінки $S_x^*(\omega)$ спектральної щільності потужності $S_x(\omega)$ формуються шляхом усереднення неспроможних оцінок. Усереднення може проводитися:

1) *по ансамблю.* Тоді інтервал спостереження розбивається на N інтервалів тривалістю T_n . На кожному n -му інтервалі обчислюється своя періодограма $S_{T_n}(\omega)$ і спроможна оцінка формується шляхом усереднення періодограм:

$$S_x^*(\omega) = \frac{1}{N} \sum_{n=1}^N S_{T_n}(\omega); \quad (10.5.2)$$

2) *по частоті.* В цьому випадку використовується одна періодограма $S_T(\omega)$. Вона усереднюється з вагою:

$$S_x^*(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega_1) S_T(\omega - \omega_1) d\omega_1, \quad (10.5.3)$$

де $W(\omega)$ – спектральне вікно, що задовольняє таким умовам:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega) d\omega = 1, \quad W(\omega) = W(-\omega); \quad (10.5.4)$$

3) у часовій області. При цьому використовується оцінка кореляційної функції $R_x^*(\tau)$. Ця величина помножується на часове вікно $w(\tau)$, після чого виконується перетворення Фур'є:

$$S_x^*(\omega) = \int_{-T}^T w(\tau) R_x^*(\tau) e^{-j\omega\tau} d\tau. \quad (10.5.5)$$

До часового вікна ставлять такі вимоги:

$$w(0) = 1, \quad w(\tau) = w(-\tau);$$

$$w(\tau) = 0 \quad \text{при} \quad |\tau| \geq \frac{T}{N}.$$

Зауваження. Часове вікно $w(\tau)$ зв'язане зі спектральним вікном $W(\omega)$ парою перетворень Фур'є:

$$w(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega) e^{j\omega\tau} d\omega, \quad W(\omega) = \int_{-\infty}^{\infty} w(\tau) e^{-j\omega\tau} d\tau.$$

Якщо існують перші два моменти випадкової функції $X(t)$, кореляційна функція $R_x(\tau)$ інтегровна, а спектральна щільність потужності $S_x(\omega)$ подвійно диференційована, то усереднені оцінки мають такі властивості.

Властивість 1. Оцінка (10.5.2) – незміщена, а оцінки (10.5.3) і (10.5.5) – зміщені.

Величина їх зміщення

$$\Delta S_x(\omega) \approx K_{\Delta} S_x''(\omega), \quad (10.5.6)$$

де

$$K_{\Delta} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \omega^2 W(\omega) d\omega. \quad (10.5.7)$$

Властивість 2. Дисперсія оцінок (10.5.3) і (10.5.5) при гауссівському розподілі $X(t)$ і великому T визначається співвідношенням

$$D[S_x^*(\omega)] \approx K_D S_x^2(\omega), \quad (10.5.8)$$

де

$$K_D = \frac{1}{T} \int_{-\infty}^{\infty} w^2(\tau) d\tau = \frac{1}{2\pi T} \int_{-\infty}^{\infty} W^2(\omega) d\omega. \quad (10.5.9)$$

Зауваження. З метою зменшення зміщення $\Delta S_x(\omega)$ належить брати якомога більш вузьке спектральне вікно $W(\omega)$ (див. формулу (10.5.7)), а для зменшення дисперсії оцінки $D[S_x^*(\omega)]$, навпаки, – більш широке вікно $W(\omega)$ (див. формули (10.5.4), (10.5.8), (10.5.9)).

Відомості про книги до розділу 10 наведені в табл. 10.5.1.

Таблиця 10.5.1

Література до розділу 10

Характер книг	Номери книг у списку літератури
Довідники	24, 47, 49
Книги навчального плану для інженерів	32, 91
Книги з математичним ухилом	50, 101
Книги прикладного характеру	2, 14, 26, 51, 52, 53, 54, 55, 66, 76, 84, 85, 86, 92, 96, 97, 100, 106

11. Додаткові питання

11.1. Дисперсійний аналіз

11.1.1. Основи дисперсійного аналізу

Загальна постановка задачі. Часто виникає потреба з'ясувати, як впливають ті чи інші фактори (чи умови проведення дослідів) на результат статистичних спостережень. Завдання такого роду зводяться до порівняння вибірок. У розділі 9 були розглянуті приклади таких завдань. При цьому вважалось, що закони розподілу кожної вибірки невідомі.

Інколи можна припустити, що вибірки отримуються з гауссівських генеральних сукупностей. Тоді завдання зводиться до зіставлення математичних сподівань та дисперсій різних вибірок. Розв'язання цієї задачі будується на базі розрахунку оцінок дисперсій. Тому метод порівняння гауссівських сукупностей носить назву *дисперсійного аналізу*.

Формулювання основної задачі. Нехай спостерігаються M вибірок $(x_1^1, \dots, x_{N_1}^1), (x_1^2, \dots, x_{N_2}^2), \dots, (x_1^M, \dots, x_{N_M}^M)$ обсягом N_1, N_2, \dots, N_M при можливо різних значеннях деякого фактора Φ (рис. 11.1.1). Відомо, що вибірки відповідають незалежним гауссівським випадковим величинам X^1, X^2, \dots, X^M із невідомими або частково невідомими параметрами: математичними сподіваннями m_1, m_2, \dots, m_M та дисперсіями D_1, D_2, \dots, D_M .

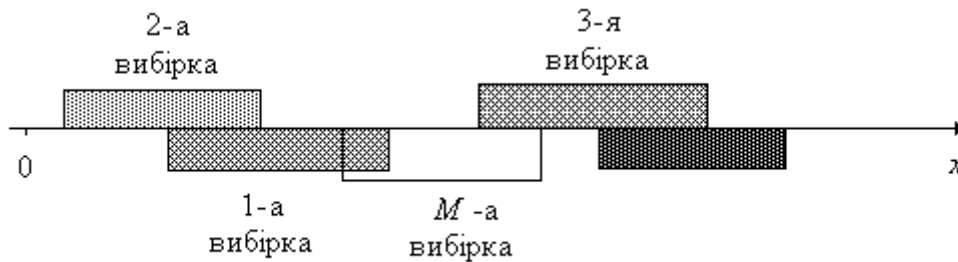


Рис. 11.1.1. Розмахи вибірок, що отримані при різних значеннях одного фактора Φ

Потрібно по результатах аналізу даних прийняти рішення про їх однорідність (їх залежність чи незалежність від фактора).

Розв'язок. По результатах спостережень кожної m -ї вибірки можна обчислити оцінки математичного сподівання m_m^* та дисперсії D_m^* :

$$m_m^* = \frac{1}{N_m} \sum_{n=1}^{N_m} x_n^m; \quad (11.1.1)$$

$$D_m^* = \frac{1}{N_m - 1} \sum_{n=1}^{N_m} (x_n^m - m_m^*)^2. \quad (11.1.2)$$

По результатах спостереження всіх M вибірок обсягом $N = \prod_{m=1}^M N_m$ можна обчислити оцінки математичного сподівання m^* та дисперсії D^* *об'єднаної вибірки*:

$$m^* = \frac{1}{N} \sum_{m=1}^M \sum_{n=1}^{N_m} x_n^m = \frac{1}{N} \sum_{m=1}^M N_m m_m^*; \quad (11.1.3)$$

$$D^* = \frac{1}{N-1} \sum_{m=1}^M \sum_{n=1}^{N_m} (x_n^m - m^*)^2 = \frac{1}{N-1} [(N-M)D_\Sigma^* + (M-1)D_0^*], \quad (11.1.4)$$

де D_Σ^* – параметр, що характеризує середнє розсіяннє у межах вибірки:

$$D_\Sigma^* = \frac{1}{N-M} \sum_{m=1}^M (N_m - 1)D_m^*, \quad (11.1.5)$$

а D_0^* – параметр, що характеризує міжвибіркове розсіяннє (розсіяннє оцінок математичних сподівань m_m^*):

$$D_0^* = \frac{1}{M-1} \sum_{m=1}^M N_m (m_m^* - m^*)^2. \quad (11.1.6)$$

Зауваження 1. Розподіли статистик, що часто використовуються при аналізі, наведено в табл. 11.1.1.

Таблиця 11.1.1

Розподіли статистик, що використовуються в дисперсійному аналізі

Статистика	Назва розподілу
$\frac{N-M}{D} D_\Sigma^*$	χ^2 - розподіл із $(N-M)$ ступенями свободи
$\frac{M-1}{D} D_0^*$	χ^2 - розподіл із $(M-1)$ ступенями свободи
$\frac{(m_n^* - m_k^*) - (m_n - m_k)}{\sqrt{D}} \sqrt{\frac{N_n N_k}{N_n + N_k}}$	стандартизований гауссівський розподіл
$\frac{(m_n^* - m_k^*) - (m_n - m_k)}{\sqrt{D_\Sigma^*}} \sqrt{\frac{N_n N_k}{N_n + N_k}}$	t - розподіл із $N_n + N_k - 2$ ступенями свободи
$\frac{D_n^*}{D_k^*}$	v^2 - розподіл із $m_1 = N_n - 1$, $m_2 = N_k - 1$
$\ln \sqrt{\frac{D_n^*}{D_k^*}}$	Z - розподіл із $m_1 = N_n - 1$, $m_2 = N_k - 1$ (Фішера)
$\frac{D_0^*}{D_\Sigma^*}$	v^2 - розподіл із $m_1 = M - 1$, $m_2 = N - M$
$\ln \sqrt{\frac{D_0^*}{D_\Sigma^*}}$	Z - розподіл із $m_1 = M - 1$, $m_2 = N - M$ (Фішера)

Примітка. Характеристики та таблиці вказаних розподілів наведені у додатках 5 – 9, 14.

Зауваження 2. На основі різноманітних статистик, насамперед тих, що входять до табл. 11.1.1, будують різні критерії слухності. Деякі з них наведено в табл. 11.1.2.

Зауваження 3. Методика аналізу даних, що представлена для скалярних величин, узагальнюється на випадок векторних величин.

Зауваження 4. Наведена методика виявлення залежності даних від одного фактора Φ узагальнюється на випадок кількох факторів (векторного фактора $\vec{\Phi} = (\Phi_1, \dots, \Phi_L)$). Відповідний аналіз називається *багатофакторним аналізом*. Відомо декілька його модифікацій. З деякими з них познайомимося в наступному підрозділі.

Примітка. Існує розділ математичної статистики, який носить назву *факторного аналізу*. Поняття фактора у факторному аналізі і в багатофакторному дисперсійному аналізі відрізняються.

Таблиця 11.1.2

Критерії слушності

Гіпотеза, що перевіряється	Критична область, в якій гіпотеза відхиляється при рівні значимості, що менша за α
$D_n = D_k$	$\left \ln \sqrt{\frac{D_n^*}{D_k^*}} \right > z_{1-\frac{\alpha}{2}}$ (при $m_1 = N_n - 1, m_2 = N_k - 1$)
$m_n = m_k$ (при $D_n = D_k$)	$\left \frac{m_n^* - m_k^*}{\sqrt{D_\Sigma^*}} \sqrt{\frac{N_n N_k}{N_n + N_k}} \right > t_{1-\frac{\alpha}{2}}$ (при $N_n + N_k - 2$ ступенях свободи)
$m_1 = m_2 = \dots = m_M$ (при $D_1 = D_2 = \dots = D_M$)	$\left \ln \sqrt{\frac{D_0^*}{D_\Sigma^*}} \right > z_{1-\frac{\alpha}{2}}$ (при $m_1 = N - 1, m_2 = N - M$)

11.1.2. Багатофакторний дисперсійний аналіз

При багатофакторному аналізі обсяг даних, що отримується для різних значень фактора $\vec{\Phi} = (\Phi_1, \dots, \Phi_L)$, як правило, дуже обмежений. Для деяких значень вектора $\vec{\Phi}$ іноді не вдається отримати жодного результату.

Визначення. При багатофакторному аналізі схема спостережень, яка охоплює всі можливі комбінації значень факторів Φ_1, \dots, Φ_L , носить назву *повного плану експериментів*. Скорочена схема спостережень, яка охоплює не всі, а лише частину можливих комбінацій цих значень, називається *неповним планом*.

Розглянемо особливості багатофакторного аналізу на прикладі двофакторного аналізу.

Постановка задачі. Ведеться аналіз залежності статистичних даних від двох факторів Φ_1 та Φ_2 . Нехай фактор Φ_1 приймає M_1 значень, а фактор Φ_2 – відповідно M_2 значень. В результаті спостережень даних при різних комбінаціях значень факторів Φ_1 та Φ_2 отримана вибірка

$$\begin{aligned}
 & x_{11}, x_{12}, \dots, x_{1M_2}, \\
 & x_{21}, x_{22}, \dots, x_{2M_2}, \\
 & \dots\dots\dots \\
 & x_{M_1 1}, x_{M_1 2}, \dots, x_{M_1 M_2}
 \end{aligned}
 \tag{11.1.7}$$

обсягом $M = M_1 M_2$ (рис. 11.1.2).

Потрібно по результатах спостережень виявити залежність чи незалежність вибірки від факторів Φ_1 та Φ_2 .

Розв'язок. За результатами спостережень можна обчислити оцінки математичного сподівання $m_{\bullet m_2}^*$ і $m_{m_1 \bullet}^*$ по стовпцях і рядках масиву (11.1.7) та загальну оцінку математичного сподівання m^* :

$$m_{\bullet m_2}^* = \frac{1}{M_1} \sum_{m_1=1}^{M_1} x_{m_1 m_2}; \tag{11.1.8}$$

$$m_{m_1 \bullet}^* = \frac{1}{M_2} \sum_{m_2=1}^{M_2} x_{m_1 m_2}; \tag{11.1.9}$$

$$m^* = \frac{1}{M_1 M_2} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} x_{m_1 m_2}. \tag{11.1.10}$$

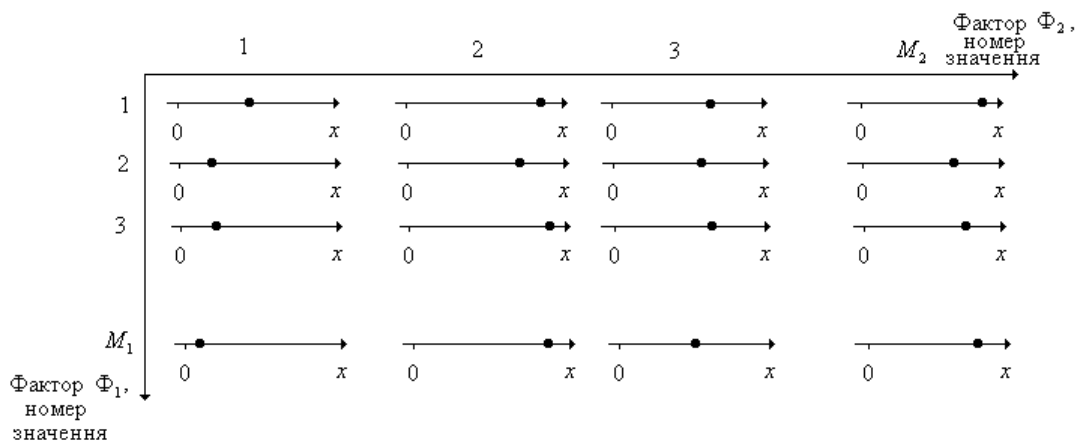


Рис. 11.1.2. Вибірка, що отримана при різних значеннях факторів Φ_1 та Φ_2

Крім того, можна розрахувати загальну оцінку дисперсії

$$\begin{aligned}
 D^* &= \frac{1}{M_1 M_2 - 1} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} (x_{m_1 m_2} - m^*)^2 = \\
 &= \frac{1}{M_1 M_2 - 1} [(M_1 - 1)(M_2 - 1)D_{\Sigma}^* + (M_1 - 1)M_2 D_1^* + M_1(M_2 - 1)D_2^*],
 \end{aligned}
 \tag{11.1.11}$$

яка складається з трьох часткових оцінок: оцінки дисперсії

$$D_{\Sigma}^* = \frac{1}{(M_1 - 1)(M_2 - 1)} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} (x_{m_1 m_2} + m^* - m_{m_1 \bullet}^* - m_{\bullet m_2}^*)^2, \tag{11.1.12}$$

що обумовлена флуктуаціями даних у стовпцях та рядках, оцінки дисперсії

$$D_1^* = \frac{1}{M_1 - 1} \sum_{m_1=1}^{M_1} (m_{m_1 \cdot}^* - m^*)^2, \quad (11.1.13)$$

що характеризує флуктуації даних між рядками, та оцінки дисперсії

$$D_2^* = \frac{1}{M_2 - 1} \sum_{m_2=1}^{M_2} (m_{\cdot m_2}^* - m^*)^2, \quad (11.1.14)$$

що характеризує флуктуації даних між стовпцями.

Зауваження 1. Якщо дані мають гауссівський закон розподілу, то оцінки дисперсій D_Σ^* , D_1^* , D_2^* взаємно незалежні й мають хі-квадрат розподіл відповідно з $(M_1 - 1)(M_2 - 1)$, $(M_1 - 1)$ і $(M_2 - 1)$ ступенями свободи.

Зауваження 2. Статистика $\frac{D_1^*}{D_\Sigma^*}$ має ν^2 -розподіл з параметрами $m_1 = M_1 - 1$,

$m_2 = (M_1 - 1)(M_2 - 1)$, а $\ln \sqrt{\frac{D_1^*}{D_\Sigma^*}}$ – розподіл Фішера з такими ж параметрами. Ці

статистики характеризують відносну флуктуацію даних по рядках і використовуються для перевірки гіпотези про рівність рядкових математичних сподівань (див. табл. 11.1.1 і 11.1.2).

Зауваження 3. Статистика $\frac{D_2^*}{D_\Sigma^*}$ має ν^2 -розподіл з параметрами $m_1 = M_2 - 1$,

$m_2 = (M_1 - 1)(M_2 - 1)$, а $\ln \sqrt{\frac{D_2^*}{D_\Sigma^*}}$ – розподіл Фішера з такими ж параметрами. Ці

статистики характеризують відносну флуктуацію даних по стовпцях і використовуються для перевірки гіпотези про рівність математичних сподівань, що розраховані по стовпцях.

Положення. При значній кількості факторів L та великому числі їх значень M для реалізації багатофакторного плану необхідно проведення дуже великої кількості експериментів, число яких не менше за M^L . На практиці реалізувати таку кількість спостережень часто неможливо. В цьому випадку шукають різні шляхи для спрощення задачі. До їх числа слід віднести використання неповних факторних планів та виявлення детермінованих зв'язків між факторами.

Приклад 1. Найпростішим неповним трьохфакторним планом експериментів є так званий *план латинського квадрата*. Цей план побудований таким чином, що для різних фіксованих комбінацій значень двох факторів, спостереження проводять тільки для одного значення третього фактора. При цьому загальна кількість спостережень дорівнює всього M^2 , а не M^3 , як у випадку повного плану.

Приклад 2. Якщо вдається встановити, що між деякими l факторами Φ_1, \dots, Φ_L існують детерміновані зв'язки, та знайти аналітичний їх опис, задачу багатофакторного аналізу можна значно спростити: замість дослідження залежностей від усіх цих l факторів можна проводити дослідження залежності лише від одного з них (або деякого віртуального фактора, що зв'язаний з ними

відомими функціями), а потім перерахувати результати для усіх l факторів. При цьому вимірність задачі знижується від L до $L - l + 1$.

11.2. Основи кластерного аналізу

Іноді спостерігається групування експериментальних даних у деяких областях простору спостереження (рис. 11.2.1).

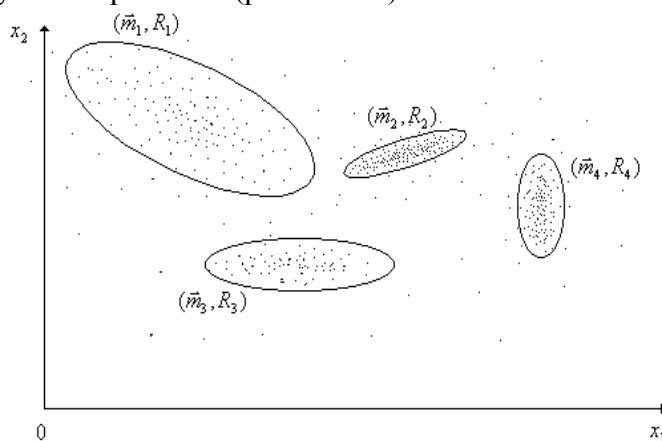


Рис. 11.2.1. Кластеризація даних

Визначення. *Кластерами* називаються області концентрації даних у просторі спостереження.

Зауваження 1. Кластери являють собою області найімовірніших значень, які можуть бути отримані при проведенні експериментів.

Зауваження 2. Кластери характеризують розміщення мод багатомодового розподілу.

Зауваження 3. При наявності групування даних часто виникає потреба в з'ясуванні кількості кластерів K (мод розподілу) та обчисленні числових їх параметрів, насамперед математичних сподівань \bar{m}_k , дисперсій \bar{D}_k та матриць кореляційних моментів R_k ($k = \overline{1, K}$).

Зауваження 4. По експериментальних даних з'ясувати точно кількість кластерів, що відповідають усім модам реального розподілу даних, принципово неможливо. Неможливо й абсолютно точно знайти параметри кластерів. Можна лише дати оцінку кількості кластерів K^* та обчислити оцінки їх числових параметрів (наприклад, математичних сподівань \bar{m}_k^* , дисперсій \bar{D}_k^* і матриць кореляційних моментів R_k^* , де $k = \overline{1, K^*}$).

Основними задачами кластерного аналізу є розрахунок числа кластерів та їх параметрів. Зараз розроблено декілька підходів до розв'язання цих завдань. Розглянемо один із найпоширеніших, що використовується для розрахунку параметрів кластерів при відомій кількості кластерів (мод розподілу).

Математичне формулювання задачі. Існує деякий невідомий розподіл ймовірностей $f(\bar{x})$ N -вимірної випадкової величини \bar{X} . Є інформація, що кількість мод цього розподілу (тобто кількість можливих кластерів) дорівнює $K = 2^L$, де L – ціле додатне число. Отримано I експериментальних даних.

Потрібно розрахувати параметри кластерів: оцінки математичного сподівання \bar{m}_k^* , дисперсії \bar{D}_k^* і матриці кореляційних моментів R_k^* ($k = \overline{1, K}$).

Схема розв'язку. Задачу вирішують в L етапів. На першому етапі вводять припущення, що випадкова величина \bar{X} має гауссівський закон розподілу. Використовуючи це припущення, розраховують оцінки математичного сподівання \bar{a}_1^* та матриці кореляційного моменту R_{11}^* . По матриці R_{11}^* знаходять N -вимірні власні вектори \bar{b}_{1n}^* ($n = \overline{1, N}$) та N -вимірний вектор власних чисел $\bar{\lambda}_1^*$. Таким чином, на першому етапі фактично розраховують параметри еліпсоїда розсіювання: його центр, що описується математичним сподіванням \bar{a}_1^* , нахил осей, який характеризується векторами \bar{b}_{1n}^* , та розмір цих осей, що характеризується власними числами (компонентами вектора $\bar{\lambda}_1^*$).

Далі знаходять власний вектор \bar{b}_{1m}^* , який відповідає найбільшому власному числу λ_{1m}^* . Через центр еліпсоїда розсіювання \bar{a}_1^* перпендикулярно до знайденого власного вектора \bar{b}_{1m}^* проводять площину, що ділить еліпсоїд на дві частини.

На другому та подальших етапах для кожної множини, що отримана на попередньому етапі, проводиться та ж сама процедура. При цьому на l -му етапі кількість повторів процедури дорівнює 2^{l-1} .

Після завершення L етапів усі I відліків виявляються розділеними на $K = 2^L$ множин (кластерів).

Зауваження. Алгоритм, що описаний, забезпечує ефективний розподіл даних по кластерах, але потребує при цьому значних обчислювальних витрат. У спеціальній літературі можна знайти й інші алгоритми, більш економічні в обчислювальному плані.

11.3. Робастні статистичні методи

11.3.1. Робастні оцінки

Визначення 1. Оцінка параметра чи характеристики випадкової величини називається *непараметричною*, якщо при її розрахунку не використовуються дані про розподіл випадкової величини, і *параметричною*, якщо вона будується за допомогою таких даних.

Примітка. Параметрична оцінка може залежати як від виду, так і від параметрів розподілу випадкової величини.

Зауваження 1. Непараметричними оцінками є вибіркові оцінки числових та нечислових характеристик випадкової величини (вбіркові моменти, статистичний розподіл (емпірична функція розподілу), емпірична щільність розподілу та ін.).

Приклади 1. Непараметричними оцінками є вибіркове середнє значення

$$m_x^* = \frac{1}{N} \sum_{n=1}^N x_n \quad (11.3.1)$$

випадкової величини X , що отримана по конкретній вибірці (x_1, x_2, \dots, x_N) , а також вибіркова дисперсія

$$D_x^* = \frac{1}{N} \sum_{n=1}^N (x_n - m_x^*)^2, \quad (11.3.2)$$

емпірична функція розподілу

$$F^*(x) = \frac{1}{N} \sum_{x_n < x} 1 \quad (11.3.3)$$

(додавання проводиться по всіх n , для котрих $x_n < x$) і емпірична щільність розподілу

$$f^*(x) = \frac{1}{N} \sum_{n=1}^N \delta(x - x_n). \quad (11.3.4)$$

Приклади 2. Параметричними оцінками є оцінки

$$m_x^* = \frac{1}{k} \sum_{n=1}^N x_n f(x_n), \quad (11.3.5)$$

$$D_x^* = \frac{1}{k} \sum_{n=1}^N (x_n - m_x^*)^2 f(x_n) \quad (11.3.6)$$

математичного сподівання та дисперсії випадкової величини X , що отримані по конкретній вибірці (x_1, x_2, \dots, x_N) з урахуванням щільності ймовірностей $f(x)$

цієї випадкової величини, де $k = \sum_{n=1}^N f(x_n)$.

Приклад 3. Нехай щільність ймовірностей випадкової величини $X \in f(x, \theta)$, де θ – деякий невідомий параметр. Оцінка θ^* параметра θ , що знаходиться за методом максимуму вірогідності, є параметричною оцінкою.

Положення 1. Непараметричні оцінки (зокрема, вибіркові характеристики), як правило (хоча й не завжди), чутливі до результатів спостереження (відліків), що значно вирізняються (знаходяться від інших на великій відстані). Поява у вибірці таких відліків приводить до суттєвих змін оцінок. Найбільш виразно це проявляється при невеликому обсязі вибірки.

Приклад 4. На рис. 11.3.1 наведено результати спостереження деякої випадкової величини X . Сім її відліків згруповані поблизу одиниці, а восьмий відлік знаходиться поблизу дев'яти.

Оцінка вибіркового середнього без врахування відліку, що значно вирізняється, приблизно дорівнює 1, а оцінка вибіркової СКВ – 0,2. З урахуванням відліку, що значно вирізняється, ці оцінки становлять відповідно 2 та 2,7. Як бачимо, додатне урахування відліку, що значно вирізняється, змінює оцінку середнього у два рази, а оцінку СКВ у тринадцять разів.

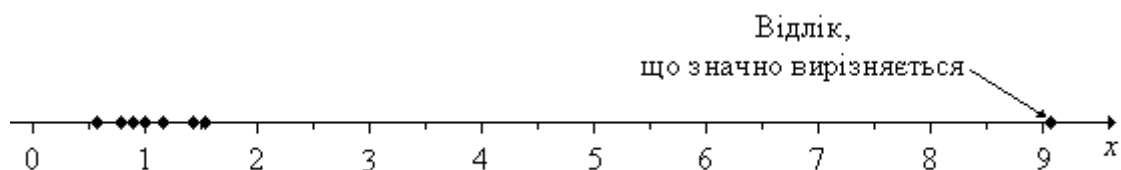


Рис. 11.3.1. Результати спостереження випадкової величини з відліком, що значно вирізняється від інших

Положення 2. Непараметричні оцінки об'єктивно характеризують вибірку. При цьому причини, що приводять до появи відліків, які значно вирізняються, залишаються поза увагою.

Зауваження 1. Причини появи відліків, які значно вирізняються, можуть бути різні, наприклад:

- можуть бути отримані у вибірці *багатомодової випадкової величини*. Найімовірніше, що такі відліки несуть інформацію про другорядні моди;
- можуть знаходитися у вибірці *одномодової (унімодальної) випадкової величини* в результаті випадкового характеру вибірки. Ймовірність появи таких відліків мала, але вона не дорівнює нулеві;
- можуть бути обумовлені грубими помилками (промахами) експериментів;
- можуть бути викликані зміною умов при формуванні вибірки.

Можливі також інші причини.

Зауваження 2. З'ясування причин появи відліків, що значно вирізняються, є складною задачею.

Простий і достатньо ефективний засіб, що дає можливість прояснити причини появи таких відліків – це повтор експериментів у тих самих умовах (формування додатних вибірок) і зіставлення отриманих вибірок.

При появі у всіх додатних вибірках відліків, що значно вирізняються, причому у одних й тих самих областях, можна припустити, що такі відліки обумовлені багатомодовим характером розподілу. При відсутності в нових вибірках відліків, що значно вирізняються, можна вважати зі значною мірою упевненості, що відліки першої вибірки, які досліджуються, були обумовлені випадковим характером вибірки. При цьому випадкова величина має унімодальний розподіл. При появі у нових вибірках відліків, що значно вирізняються, причому постійно у різних областях, можна припустити або постійні промахи в експериментах, або недотримання одних й тих самих умов проведення іспитів.

Примітка. Наведені варіанти прийняття рішень про причини появи відліків, що значно вирізняються, носять не більш як рекомендаційний характер. На практиці можуть виникати дуже складні ситуації, які не описуються переліченими схемами.

Зауваження 3. Якщо відліки, що значно вирізняються, обумовлені багатомодовим характером розподілу випадкової величини, то ефективність непараметричних оцінок висока, якщо ж вони викликані іншими причинами, то ефективність оцінок низька.

З метою забезпечення високої ефективності оцінок слід рекомендувати завжди проводити аналіз відліків і враховувати результати аналізу при розрахунку оцінок.

Якщо відліки, що значно вирізняються, підвищують ефективність непараметричних оцінок, то такі оцінки з успіхом можна використовувати, якщо ж ні, то використання непараметричних оцінок проблематичне. В останньому випадку має сенс звернутися до спеціальних параметричних оцінок.

Визначення 2. Оцінка T^* параметра t , яка отримується по вибірці випадкової величини, називається *стійкою*, якщо при значних змінах значень елементів вибірки вона мало змінюється.

Зауваження 1. Поняття стійкості оцінки часто ув'язують із чутливістю до деяких змін елементів вибірки. При цьому оцінка може бути стійкою до одних змін і нестійкою до інших.

Приклад 5. Непараметрична оцінка математичного сподівання

$$\bar{x}_c = \frac{1}{N} \sum_n y_n, \quad (11.3.7)$$

де $y_n = \begin{cases} x_n, & \text{якщо } |x_n| < c, \\ c, & \text{якщо } |x_n| \geq c, \end{cases}$ c – константа, що характеризує ступінь стійкості,

зовсім нечутлива до змін елементів, значення яких по модулю більше за c . У той же час вона залежить від елементів, значення яких менше за c . Причому в останньому випадку суттєву роль відіграє характер змін цих елементів. Наприклад, при зміщенні усіх елементів на фіксовану величину оцінка зміщується більше, ніж при зміщенні різних елементів по модулю на ту ж саму величину.

Зауваження 2. Під стійкою оцінкою часто розуміють оцінку, що нечутлива до відліків, що значно вирізняються.

Визначення 3а. Оцінка називається *робастною*, якщо вона нечутлива до малих відхилень від розрахункових умов.

Примітка. Наведене якісне визначення поняття робастності оцінки може бути конкретизовано таким чином.

Визначення 3б. Оцінка T^* параметра t , яка отримується по вибірці випадкової величини X , називається *робастною*, якщо при спрямованості до нуля відстані між реальною функцією розподілу $F(x)$ і розрахунковою функцією розподілу $F_0(x)$ ($\mu(F(x), F_0(x)) \rightarrow 0$) відстань між відповідними функціями розподілу оцінок $\mathfrak{F}_{F(x)}(t^*)$ і $\mathfrak{F}_{F_0(x)}(t^*)$ теж наближається до нуля ($\mu(\mathfrak{F}_{F(x)}(t^*) \mathfrak{F}_{F_0(x)}(t^*)) \rightarrow 0$).

Зауваження 1. Поняття робастності має змістовний сенс тільки для параметричних оцінок, оскільки непараметричні оцінки не залежать від розподілу і тому $\mu(\mathfrak{F}_{F(x)}(t^*) \mathfrak{F}_{F_0(x)}(t^*)) = 0$.

Зауваження 2. Під словосполученням “робастна оцінка” часто розуміють оцінку, що нечутлива до малих відхилень функції розподілу (щільності ймовірностей) у районі “хвостів” (рис. 11.3.2).

Зауваження 3. Робастність і стійкість – різні характеристики оцінки. Однак на практиці в багатьох випадках робастні оцінки є стійкими, і навпаки, стійкі оцінки є робастними. Тому нерідко ці поняття розглядають як синоніми.

Приклади 6. Робастними оцінками математичного сподівання (зсуву) випадкової величини X є:

– медіана розподілу

$$m_x^* = \text{med}(x_1, \dots, x_N), \quad (11.3.8)$$

– α -урізане середнє значення

$$\bar{x}_\alpha = \frac{1}{N - 2N_\alpha} \sum_n x_n, \quad (11.3.9)$$

де додавання проводиться по усіх елементах урізаної вибірки, що формується шляхом відкидання N_α її мінімальних і N_α її максимальних елементів.

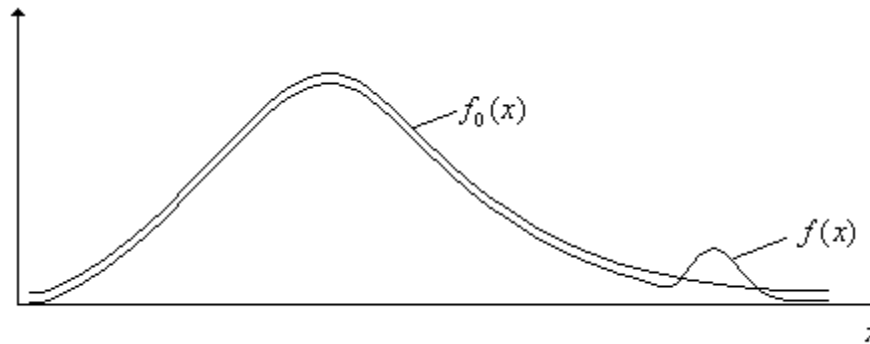


Рис. 11.3.2. Щільності ймовірностей випадкової величини: розрахункова щільність ймовірностей $f_0(x)$ та реальна $f(x)$

Приклад 7. Робастною оцінкою масштабу, що є аналогом СКВ випадкової величини X , є абсолютне медіанне відхилення від оцінки зсуву:

$$s_x^* = \text{med}(|x_1 - m_x^*|, \dots, |x_N - m_x^*|). \quad (11.3.10)$$

Зауваження 1. Теорія робастних оцінок почала інтенсивно розвиватися лише в останні роки. За цей час розроблено ряд теоретичних положень, і на їх основі отримані важливі практичні результати. Крім робастних оцінок, наведених у прикладах 6, 7, запропоновано багато інших. Причому не тільки оцінок зсуву й масштабу. В першу чергу це робастні оцінки кореляційних і регресійних залежностей.

Зауваження 2. Як правило, робастні оцінки засновані на сортуванні даних, що потребує значних обчислювальних витрат. Тому перехід від неробастних оцінок до робастних пов'язаний з підвищенням вимог до швидкодії обчислювальних засобів.

11.3.2. Перевірка гіпотез з елементами робастної обробки даних

Постановка питання. В підрозділі 9.2 розглядалася задача перевірки основної H_0 й конкуруючої H_1 гіпотез по вибірці $\vec{x} = (x_1, \dots, x_N)$. Було відмічено, що за різними критеріями оптимальні правила прийняття рішення

базуються на розрахунку відношення функцій вірогідності $l(\vec{x}) = \frac{L(\vec{x}/H_1)}{L(\vec{x}/H_0)}$ і на

порівнянні логарифма цього відношення $\ln l(\vec{x})$ з порогом c :

$$\ln l(\vec{x}) > c. \quad (11.3.11)$$

Якщо нерівність (11.3.11) не задовольняється, приймається рішення, що справедлива гіпотеза H_0 . У противному разі приймається рішення, що справедлива гіпотеза H_1 .

У випадку, коли елементи вибірки незалежні і мають однаковий закон розподілу, цей алгоритм може бути записаний у вигляді

$$\ln l(x) = \sum_{n=1}^N \ln l_1(x_n) > c, \quad (11.3.12)$$

де $l_1(x_n) = \frac{L(x_n / H_1)}{L(x_n / H_0)}$ – відношення функцій вірогідності при гіпотезах H_1 і H_0 ,

що розраховані для елемента x_n .

Якщо функції вірогідності $L(x_n / H_0)$ і $L(x_n / H_1)$ відомі абсолютно точно, то проблем не виникає. Але коли вони невідомі, побудувати відношення функцій вірогідності неможливо. Заміна невідомих функцій вірогідності деякими наближеними, наприклад, гауссівськими, не завжди дає позитивний результат. Більш доцільно переформулювати задачу інакше.

Формулювання задачі. По результатах спостереження вибірки $\vec{x} = (x_1, \dots, x_N)$ з незалежними однаково розподіленими елементами перевіряються дві гіпотези: H_0 і H_1 . Гіпотеза H_0 полягає в тому, що елементи вибірки \vec{x} описуються функцією вірогідності $\tilde{L}(x_n / H_0)$, яка відноситься до деякого класу моделей \mathfrak{Z}_0 , а гіпотеза H_1 – в тому, що її елементи описуються функцією вірогідності $\tilde{L}(x_n / H_1)$, яка відноситься до іншого класу моделей \mathfrak{Z}_1 . При цьому класи \mathfrak{Z}_0 і \mathfrak{Z}_1 не перетинаються.

Найпоширенішими є однопараметричні моделі ε -забруднення:

$$\tilde{L}(x_n / H_0) = (1 - \varepsilon_0)L(x_n / H_0) + \varepsilon_0 w_0(x_n), \quad (11.3.13)$$

$$\tilde{L}(x_n / H_1) = (1 - \varepsilon_1)L(x_n / H_1) + \varepsilon_1 w_1(x_n), \quad (11.3.14)$$

де $\varepsilon_0, \varepsilon_1$ – параметри, що характеризують рівень забруднення (відхилення від вихідних функцій вірогідності $L(x_n / H_0)$ і $L(x_n / H_1)$), $w_0(x_n)$ і $w_1(x_n)$ – довільні щільності ймовірностей.

Розв'язок. Доведено, що існують найменш сприятливі щільності ймовірностей $q_0(x)$ і $q_1(x)$, які описуються такими виразами:

$$q_0(x) = \begin{cases} (1 - \varepsilon_0)L(x / H_0), & l_1(x) < c'', \\ (1 - \varepsilon_0) \frac{L(x / H_1)}{c''}, & l_1(x) \geq c''; \end{cases}$$

$$q_1(x) = \begin{cases} (1 - \varepsilon_1)L(x / H_1), & l_1(x) > c', \\ (1 - \varepsilon_1)c' L(x / H_0), & l_1(x) \leq c', \end{cases}$$

де c' і c'' – невід'ємні числа ($c' < c''$), що визначаються з умови забезпечення властивостей щільностей ймовірностей $q_0(x)$ і $q_1(x)$.

При цьому для $q_0(x)$ і $q_1(x)$ відношення вірогідностей $l_q(x)$ може бути отримано у такому вигляді:

$$l_q(x) = \begin{cases} bc'', & l_1(x) \geq c'', \\ bl_1(x), & c' < l_1(x) < c'', \\ bc', & l_1(x) \leq c', \end{cases} \quad (11.3.15)$$

$$\text{де } b = \frac{(1 - \varepsilon_1)}{(1 - \varepsilon_0)}.$$

Розрахувавши ці відношення вірогідностей $l_q(x_n)$ для різних x_n , можна використати їх у виразі (11.3.12) замість $l_1(x_n)$.

Зауваження 1. Наведена модернізована процедура перевірки гіпотез має робастні властивості.

Зауваження 2. Зараз розроблені робастні процедури також для інших моделей \mathfrak{T}_0 і \mathfrak{T}_1 , опис яких можна знайти у спеціальній літературі.

11.3.3. Регуляризація оцінок

Постановка питання. Проблема забезпечення стійкості оцінок до змін розрахункових умов (тобто проблема робастності) турбувала дослідників давно. У 60-х роках минулого століття ця проблема була сформульована А.Н. Тихоновим як проблема *регуляризації (стабілізації) оцінок*.

Ним було підмічено, що в деяких випадках дуже висока розрахункова ефективність оцінок на практиці не підтверджується. Дослідження виявили, що ця ситуація, зокрема, спостерігається тоді, коли при розрахунку оцінки проводиться обернення погано обумовленої матриці, тобто матриці з великим перепадом значень власних чисел.

Розв'язок. Для стабілізації оцінок в умовах використання погано обумовленої матриці A було запропоновано А.Н. Тихоновим використовувати замість цієї матриці іншу матрицю \tilde{A} , що відрізняється від матриці A на величину $\Delta A = rI$, де r – параметр, який називається *параметром регуляризації*, I – одинична матриця.

При вдалому виборі параметра регуляризації отримуються оцінки, що нечутливі до похибок умов.

Зауваження 1. Обґрунтований вибір величини параметра регуляризації r – непроста задача. Існують різні підходи до вибору цієї величини. На практиці добрі результати дає r , при якій власні числа матриці \tilde{A} відрізняються не більш як в 10 – 100 разів.

Зауваження 2. Відомі також інші методи регуляризації оцінок, опис яких можна знайти в спеціальній літературі.

Відомості про книги з тематики розділу 11 наведені в табл. 11.3.1.

Таблиця 11.3.1

Література до розділу 11

Характер книг	Номери книг у списку літератури
Довідники	24, 40, 47, 49, 81, 93, 108
Книги навчального плану для інженерів	16, 17, 32, 75, 84
Книги з математичним ухилом	21, 46, 50, 90, 101, 110
Книги прикладного характеру	43, 52, 53, 55, 56, 62, 63, 65, 76, 85, 97, 100, 102

Додатки

Додаток 1

Дельта – функція

Визначення. Дельта-функцією називається функція, яка визначається таким чином:

$$\delta(t-t_0) = \begin{cases} 0, & \text{при } t \neq t_0; \\ \infty, & \text{при } t = t_0. \end{cases}$$

Основні властивості дельта-функції зведені у табл. Д1.

Таблиця Д1

Основні властивості δ -функції

Властивості
$\delta(\tau) = \int_{-\infty}^{\infty} e^{j2\pi f\tau} df$
$\delta(t-t_0) = \lim_{\sigma \rightarrow 0} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-t_0)^2}{2\sigma^2}}$
$\delta(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} d\omega$
$\int_a^b x(t)\delta(t-t_0)dt = \begin{cases} 0 & \text{при } t_0 < a; t_0 > b (a < b); \\ \frac{1}{2}x(t_0) & \text{при } t_0 = a \quad \text{або} \quad t_0 = b; \\ x(t_0) & \text{при } a < t_0 < b, \end{cases}$ <p style="text-align: center;">де $x(t)$ – функція, неперервна в точці $t = t_0$</p>
$\int_{t_0-\varepsilon}^{t_0+\varepsilon} \delta(t-t_0)dt = 1, \quad \varepsilon > 0$
Спектр $A_{\delta(t-t_0)}(\omega) = e^{-j\omega t_0}$
Похідна $\delta^{(r)}(t) = (-1)^r r! \frac{\delta(t)}{t^r}$, де $r=0, 1, 2, \dots$

Додаток 2

Гама-функція

Гама-функція $\Gamma(z)$ звичайно визначається у вигляді інтеграла Ейлера другого роду

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt \quad (\operatorname{Re} z > 0),$$

або як розв'язок функціонального рівняння $\Gamma(z+1) = z\Gamma(z)$, де $\Gamma(1) = 1$.

Гама-функція для дійсних значень z зображена на рис. Д 2.1.

Для цілих значень аргументу $z = N$ вона може бути обчислена за такою формулою:

$$\Gamma(N) = (N-1)! \quad (N = 0, 1, \dots).$$

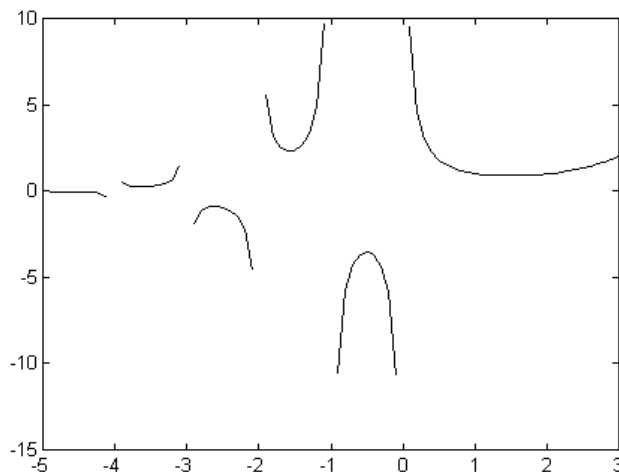


Рис. Д 2.1. Вигляд гама-функції для дійсних аргументів

Відзначимо, що $\Gamma(0,5) = \sqrt{\pi}$,

$$\Gamma(N + 0,5) = \sqrt{\pi} \frac{\prod_{n=1}^N (2n-1)}{2^N},$$

$$\Gamma(-N + 0,5) = \sqrt{\pi} \frac{(-2^N)}{\prod_{n=1}^N (2n-1)}.$$

Додаток 3

Неперервність, диференційованість та інтегровність випадкових функцій

Математиками використовується чотири різних типи збіжності.

Визначення 1. Нехай ϵ послідовність випадкових величин $X = \{X_1, \dots, X_N\}$ і випадкова величина X . Для всіх X_1, \dots, X_N і X визначені функції розподілу $F_1(x), \dots, F_N(x)$ і $F(x)$. Тоді послідовність X :

1) збігається до X за розподілом, якщо в кожній точці x , де $F(x)$ неперервна, $F_N(x) \rightarrow F(x)$ при $N \rightarrow \infty$;

2) збігається до X за ймовірністю, якщо $P\{|X_N - X| > \epsilon\} \rightarrow 0$ при будь-якому $\epsilon > 0$ і $N \rightarrow \infty$;

3) збігається до X у середньоквадратичному, якщо $M[|X_N - X|^2] \rightarrow 0$ при $N \rightarrow \infty$. При такій збіжності пишуть $\text{l.i.m.}_{N \rightarrow \infty} X_N = X$;

4) збігається до X з ймовірністю одиниця (майже обов'язково), якщо $P\{X_N \rightarrow X\} = 1$ при $N \rightarrow \infty$. При такій збіжності пишуть $\lim_{N \rightarrow \infty} X_N = X$.

Зауваження 1. Найбільш слабка збіжність – за розподілом. Більш сильна збіжність – за ймовірністю. Ще більш сильна збіжність – збіжність у середньоквадратичному і з ймовірністю одиниця. При цьому слід мати на увазі, що деякі послідовності збігаються в середньоквадратичному, але не збігаються з ймовірністю одиниця; інші ж – збігаються з ймовірністю одиниця, але не збігаються в середньоквадратичному. Співвідношення між різними типами збіжності умовно зображені на рис. Д 3.1.

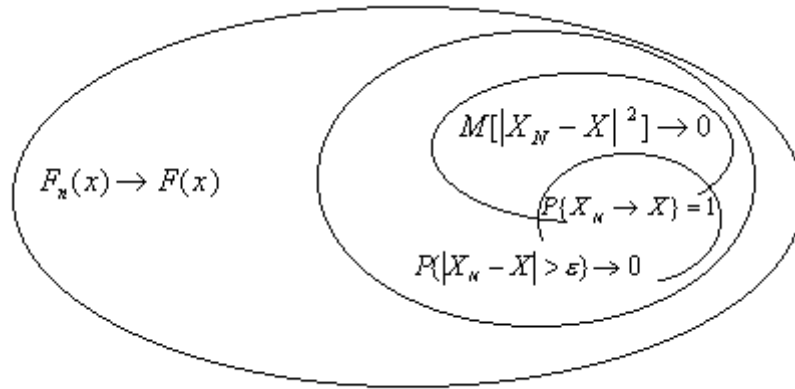


Рис. Д 3.1. Співвідношення між різними типами збіжності

Зауваження 2. На практиці найбільш часто використовується збіжність у середньоквадратичному.

Зауваження 3. Поняття збіжності випадкової послідовності X поширюється на випадкові функції $X(t)$ ($t \in T$).

Визначення 2. Випадковою функцією (процесом) другого порядку називають таку випадкову функцію $X(t)$ ($t \in T$), математичне сподівання квадрата якої обмежено для усіх $t \in T$: $M[X^2(t)] < \infty$.

Зауваження. Для випадкових функцій другого порядку існують поняття неперервності, диференційованості та інтегровності.

Визначення 3. Випадкова функція другого порядку називається неперервною в середньоквадратичному в точці t , якщо

$$\text{l.i.m.}_{\Delta t \rightarrow 0} X(t + \Delta t) = X(t),$$

тобто $\lim_{\Delta t \rightarrow 0} M[|X(t + \Delta t) - X(t)|^2] = 0$.

Теорема 1. Випадкова функція $X(t)$ другого порядку неперервна в середньоквадратичному в точці t тоді і тільки тоді, коли її математичне сподівання $m_x(t)$ неперервно в точці t і кореляційна функція $R_x(t_1, t_2)$ неперервна в точці $t = t_1 = t_2$.

Визначення 4. Випадкова функція $X(t)$ другого порядку називається диференційованою в середньоквадратичному в точці t , якщо існує функція

$$X'(t) = \text{l.i.m.}_{\Delta t \rightarrow 0} \frac{X(t + \Delta t) - X(t)}{\Delta t}.$$

При цьому випадкова функція $X'(t)$ називається похідною випадкової функції $X(t)$.

Теорема 2. Випадкова функція $X(t)$ другого порядку диференційована в середньоквадратичному в точці t тоді і тільки тоді, коли її математичне сподівання $m_x(t)$ диференційоване в точці t і в точці $t_1 = t_2 = t$ існує змішана похідна другого порядку

$$\frac{\partial^2}{\partial t_1 \partial t_2} R_x(t_1, t_2)$$

від кореляційної функції $R_x(t_1, t_2)$.

Визначення 5. Випадкова функція $X(t)$ другого порядку називається *інтегрованою* на інтервалі $T(\tau)$, якщо при довільному розподілі інтервалу $T(\tau)$ на N інтервалів $\Delta t_n = t_n - t_{n-1}$ незалежно від вибору точок t_n існує функція

$$Y(\tau) = \text{l.i.m.}_{\max \Delta t_n \rightarrow 0} \sum_n X(t_n) \Delta t_n = \int_{T(\tau)} X(t) dt.$$

При цьому випадкова функція $Y(\tau)$ називається *інтегралом* випадкової функції $X(t)$.

Зауваження. Доведено, що випадкова функція $X(t)$ другого порядку з математичним сподіванням $m_x(t)$ і кореляційною функцією $R_x(t_1, t_2)$ інтегровна по Ріману, якщо існують інтеграли $\int_{T(\tau)} m_x(t) dt$, $\int_{T(\tau)} \int_{T(\tau)} R_x(t_1, t_2) dt_1 dt_2$.

При цьому

$$M_x \int_{T(\tau)} X(t) dt = \int_{T(\tau)} m_x(t) dt,$$

$$M_x \int_{T(\tau)} \int_{T(\tau)} X(t_1) X(t_2) dt_1 dt_2 = \int_{T(\tau)} \int_{T(\tau)} R_x(t_1, t_2) dt_1 dt_2 + \left[\int_{T(\tau)} m_x(t) dt \right]^2.$$

Додаток 4

Перетворення Гільберта й аналітичний сигнал

Визначення 1. Функції $s(t)$ і $s_{\perp}(t)$ називаються *спряженими за Гільбертом*, якщо вони зв'язані між собою співвідношеннями

$$s_{\perp}(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{\tau - t} d\tau = H[s(t)],$$

$$s(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s_{\perp}(\tau)}{\tau - t} d\tau = -H[s_{\perp}(t)].$$

Перетворення Гільберта H та функції, що спряжені за Гільбертом, мають ряд важливих особливостей.

Особливість 1. Перетворення Гільберта – лінійне перетворення.

Особливість 2. Функції, спряжені за Гільбертом, ортогональні, тобто

$$\int_{-\infty}^{\infty} s(t) s_{\perp}(t) dt = 0.$$

Особливість 3. Перетворення Гільберта – *однорідне (стаціонарне) по аргументу t* , тобто при затримці функції $s(t)$ на фіксовану величину t_0 (при цьому вона набуває вигляду $s(t - t_0)$) спряжена за Гільбертом функція має вигляд $s_{\perp}(t - t_0)$. Ця особливість випливає з особливості 1.

Особливість 4. Спектри $\dot{F}_s(\omega)$ і $\dot{F}_{s_{\perp}}(\omega)$ спряжених за Гільбертом функцій $s(t)$ і $s_{\perp}(t)$ зв'язані між собою співвідношенням

$$\dot{F}_s(\omega) = -j \dot{F}_{s_{\perp}}(\omega) \text{sign}(\omega),$$

де $\text{sign}(\omega) = \begin{cases} 1 & \text{при } \omega \geq 0, \\ -1 & \text{при } \omega < 0. \end{cases}$

Зауваження 1. Модулі спектрів $|\dot{F}_s(\omega)|$ і $|\dot{F}_{s_{\perp}}(\omega)|$ спряжених за Гільбертом функцій $s(t)$ і $s_{\perp}(t)$ однакові:

$$|\dot{F}_s(\omega)| = |\dot{F}_{s\perp}(\omega)|.$$

Зауваження 2. Аргументи спектрів $\arg \dot{F}_s(\omega)$ і $\arg \dot{F}_{s\perp}(\omega)$ спряжених за Гільбертом функцій $s(t)$ і $s_{\perp}(t)$ зв'язані співвідношенням

$$\arg \dot{F}_s(\omega) = \arg \dot{F}_{s\perp}(\omega) - \frac{\pi}{2} \text{sign}(\omega).$$

Особливість 5. Перетворення Гільберта – *однорідне по частоті та фазі*. Це означає, що заміна у спектральному розкладанні функції $s(t)$ усіх кругових частот ω на $\omega + \Omega$ або усіх фаз φ на $\varphi + \Phi$ приводить до спряженої функції, спектральне розкладання якої відрізняється від спектрального розкладання функції $s_{\perp}(t)$ зсувом відповідно кругових частот на Ω , або початкових фаз на Φ .

Особливість 6. Перетворення Гільберта від гармонічного коливання $s(t) = \cos(\omega_0 t + \varphi)$ з круговою частотою ω_0 і початковою фазою φ дає спряжену за Гільбертом функцію $s_{\perp}(t) = \sin(\omega_0 t + \varphi)$.

Особливість 7. Перетворення Гільберта можна розглядати як фільтрацію функції $s(t)$ за допомогою фільтра з імпульсною перехідною характеристикою

$$h(t) = \frac{1}{\pi t}, \text{ якій відповідає частотна характеристика } K(\omega) = \begin{cases} -j, & \omega \geq 0, \\ j, & \omega < 0 \end{cases}.$$

Зауваження. Перетворення Гільберта забезпечує поворот фаз усіх частотних компонент на кут $-\pi/2$.

Приклади. Перетворення Гільберта деяких функцій наведено в табл. Д 4.1.

Таблиця Д 4.1

Перетворення Гільберта

Функція $s(t)$	Спряжена функція $s_{\perp}(t)$
$\delta(t)$	$\frac{1}{\pi t}$
$\cos t$	$\sin t$
$\sin t$	$-\cos t$
$\frac{\sin t}{t}$	$\frac{1 - \cos t}{t}$
$\frac{1}{1+t^2}$	$\frac{t}{1+t^2}$
$\begin{cases} 1, & t \leq T, \\ 0, & t > T \end{cases}$	$-\frac{1}{\pi} \ln \left \frac{t-T}{t+T} \right $

Визначення 2. Аналітичним комплексним сигналом називається комплексна функція

$$\dot{s}(t) = s(t) + js_{\perp}(t).$$

Визначення 3. Огинаючою аналітичного сигналу називається функція $S(t) = |\dot{s}(t)|$, а фазою – функція $\Phi(t) = \arg \dot{s}(t)$.

Зауваження 1. Огинаюча $S(t)$ і фаза $\Phi(t)$ аналітичного сигналу $\dot{s}(t)$ однозначно пов'язані з вхідними спряженими функціями $s(t)$ і $s_{\perp}(t)$:

$$S(t) = \sqrt{s^2(t) + s_{\perp}^2(t)},$$

$$\Phi(t) = \operatorname{arctg} \frac{s_{\perp}(t)}{s(t)}.$$

Зауваження 2. Аналітичний сигнал $\dot{s}(t)$ пов'язаний з огинаючою $S(t)$ і фазою $\Phi(t)$ співвідношенням

$$\dot{s}(t) = S(t)e^{j\Phi(t)}.$$

Положення 1. Аналітичний сигнал $\dot{s}(t)$ вузькосмугового сигналу зі смугою $\Delta\omega_e \ll \omega_0$ можна представити таким чином:

$$\dot{s}(t) = S(t)e^{j(\omega_0 t + \Phi(t))}.$$

Визначення 4. *Комплексною огинаючою* вузькосмугового сигналу називається функція

$$\dot{S}(t) = S(t)e^{j\Phi(t)}.$$

Положення 2. Перетворення Гільберта широко використовують при аналізі як детермінованих, так і випадкових функцій.

Розглянемо деякі особливості, що характерні для випадкових функцій $X(t)$ і $X_{\perp}(t)$, які пов'язані перетворенням Гільберта ($X_{\perp}(t) = H[X(t)]$).

Особливість 1. Математичне сподівання $m_{x\perp}(t)$ випадкової функції $X_{\perp}(t)$ є перетворенням Гільберта математичного сподівання $m_x(t)$ випадкової функції $X(t)$:

$$m_{x\perp}(t) = H[m_x(t)].$$

Особливість 2. Кореляційні функції $R_x(t_1, t_2)$, $R_{x\perp}(t_1, t_2)$ та взаємні кореляційні функції $R_{xx\perp}(t_1, t_2)$, $R_{x\perp x}(t_1, t_2)$ центрованих випадкових функцій $X(t)$, $X_{\perp}(t)$ пов'язані між собою перетворенням Гільберта:

$$R_{x\perp}(t_1, t_2) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R_{xx\perp}(\tau, t_2)}{\tau - t_1} d\tau,$$

$$R_x(t_1, t_2) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R_{xx\perp}(t_2, \tau)}{\tau - t_1} d\tau,$$

$$R_{xx\perp}(t_1, t_2) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R_{x\perp}(t_2, \tau)}{\tau - t_1} d\tau,$$

$$R_{x\perp x}(t_1, t_2) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R_x(t_2, \tau)}{\tau - t_1} d\tau.$$

Особливість 3. Якщо центрована випадкова функція $X(t)$ стаціонарна у широкому розумінні, то випадкова функція $X_{\perp}(t)$ теж стаціонарна у широкому розумінні.

Особливість 4. Взаємна кореляційна функція $R_{xx\perp}(\tau)$ стаціонарної у широкому розумінні центрованої випадкової функції $X(t)$ непарна, а кореляційні функції випадкових функцій $X(t)$, $X_{\perp}(t)$ збігаються і парні:

$$R_{xx\perp}(-\tau) = -R_{xx\perp}(\tau). \\ R_{x\perp}(\tau) = R_x(\tau) = R_x(-\tau).$$

Особливість 5. З наведеної особливості 4 та особливості 7 перетворення Гільберта впливає, що спектральні щільності потужності $S_x(f)$, $S_{x\perp}(f)$ випадкових функцій $X(t)$, $X_{\perp}(t)$ збігаються ($S_x(f) = S_{x\perp}(f)$), а взаємна

$$\text{спектральна щільність потужності } \dot{S}_{xx\perp}(f) = \begin{cases} -jS_x(f), & f \geq 0, \\ jS_x(f), & f < 0. \end{cases}$$

Особливість 6. З наведеної особливості 4 впливає, що у однакові моменти t центровані випадкові функції $X(t)$, $X_{\perp}(t)$ некорельовані.

Особливість 7. Кореляційна функція $R_x(\tau)$ центрованого стаціонарного випадкового аналітичного сигналу $\dot{X}(t) = X(t) + jX_{\perp}(t)$ пов'язана з кореляційними функціями $R_x(\tau)$ і $R_{x\perp}(\tau)$ таким чином:

$$R_{\dot{X}}(\tau) = 2R_x(\tau) + 2jR_{x\perp}(\tau).$$

Особливість 8. Спектральна щільність аналітичного сигналу $\dot{S}_{\dot{X}}(f)$ центрованого стаціонарного випадкового аналітичного сигналу $\dot{X}(t)$ пов'язана зі спектральною щільністю $S_x(f)$ функції $X(t)$ і взаємною спектральною щільністю $\dot{S}_{x\perp}(f)$ функцій $X_{\perp}(t)$, $X(t)$ таким чином:

$$\dot{S}_{\dot{X}}(f) = 2S_x(f) + j\dot{S}_{x\perp}(f).$$

Додаток 5

Хі-квадрат розподіл

Визначення. Нехай є m взаємно незалежних гауссівських стандартизованих (тобто з нульовим математичним сподіванням і одиничною дисперсією) випадкових величин X_n ($n = \overline{1, m}$). Тоді сума квадратів цих величин

$$\chi^2(x; m) = \sum_{n=1}^m X_n^2$$

являє собою χ^2 -розподіл із m ступенями свободи.

Щільність ймовірностей цього розподілу має вигляд

$$f(x; m) = \begin{cases} 0, & x \leq 0; \\ \frac{1}{\Gamma\left(\frac{m}{2}\right)2^{\frac{m}{2}}} x^{\frac{m-2}{2}} \exp\left(-\frac{x}{2}\right), & x > 0. \end{cases}$$

Вигляд щільності ймовірностей χ^2 -розподілу при різноманітних значеннях m показаний на рис. Д 5.

Математичне сподівання $m_{\chi^2} = m$, дисперсія $D_{\chi^2} = 2m$, а мода знаходиться в точці $m - 2$ ($m > 2$).

При $m \rightarrow \infty$ χ^2 -розподіл наближається до гауссівського з математичним сподіванням m і дисперсією $2m$.

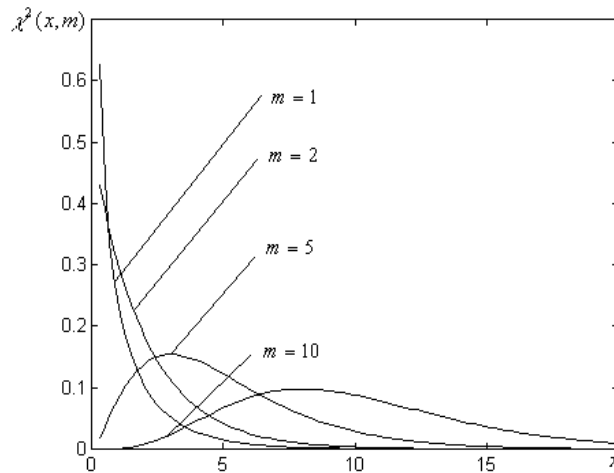


Рис. Д 5. Вигляд щільності ймовірностей χ^2 -розподілу при різноманітних значеннях m

Додаток 6

Розподіл Стюдента (t -розподіл)

Визначення. Нехай є $m+1$ взаємно незалежних гауссівських стандартизованих випадкових величин X_n ($n = \overline{0, m}$). Тоді величина

$$t = \frac{X_0}{\sqrt{\frac{1}{m}(X_1^2 + X_2^2 + \dots + X_m^2)}}$$

являє собою розподіл Стюдента (t -розподіл) із m ступенями свободи.

Щільність ймовірностей цього розподілу описується співвідношенням

$$S(t; m) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\sqrt{m\pi}} \left(1 + \frac{t^2}{m}\right)^{-\frac{m+1}{2}}.$$

Вигляд щільності ймовірностей розподілу Стюдента показаний на рис. Д 6.

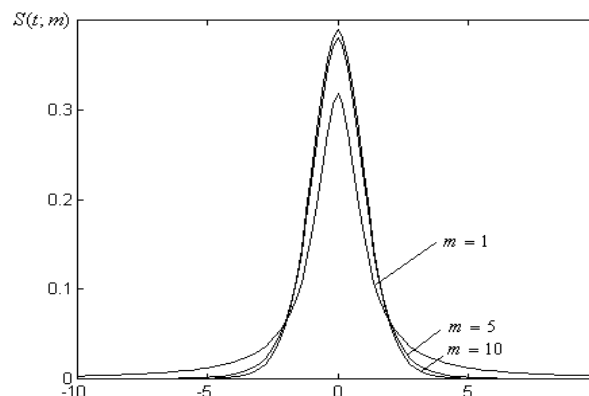


Рис. Д 6. Щільність ймовірностей розподілу Стюдента

Математичне сподівання $m_t = 0$, дисперсія $D_t = \frac{m}{m-2}$ ($m > 2$), а мода знаходиться в точці 0. Функція $S(t; m)$ симетрична відносно осі ординат.

При $m \rightarrow \infty$ розподіл Стюдента наближається до гауссівського із математичним сподіванням 0 і дисперсією 1.

Додаток 7

Таблиці гауссівської функції розподілу
Значення гауссівської функції розподілу

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

x	F(x)	x	F(x)	x	F(x)	x	F(x)
-0,0	0,5000	-2,0	0,0228	0,0	0,5000	2,0	0,9772
-0,1	0,4602	-2,1	0,0179	0,1	0,5398	2,1	0,9821
-0,2	0,4207	-2,2	0,0139	0,2	0,5793	2,2	0,9861
-0,3	0,3821	-2,3	0,0107	0,3	0,6179	2,3	0,9893
-0,4	0,3446	-2,4	0,0082	0,4	0,6554	2,4	0,9918
-0,5	0,3085	-2,5	0,0062	0,5	0,6915	2,5	0,9938
-0,6	0,2743	-2,6	0,0047	0,6	0,7257	2,6	0,9953
-0,7	0,2420	-2,7	0,0035	0,7	0,7580	2,7	0,9965
-0,8	0,2119	-2,8	0,0026	0,8	0,7881	2,8	0,9974
-0,9	0,1841	-2,9	0,0019	0,9	0,8159	2,9	0,9981
-1,0	0,1587	-3,0	0,0014	1,0	0,8413	3,0	0,9986
-1,1	0,1357	-3,1	0,0010	1,1	0,8643	3,1	0,9990
-1,2	0,1151	-3,2	0,0007	1,2	0,8849	3,2	0,9993
-1,3	0,0968	-3,3	0,0005	1,3	0,9032	3,3	0,9995
-1,4	0,0808	-3,4	0,0003	1,4	0,9192	3,4	0,9997
-1,5	0,0668	-3,5	0,0002	1,5	0,9332	3,5	0,9998
-1,6	0,0548	-3,6	0,0002	1,6	0,9452	3,6	0,9998
-1,7	0,0446	-3,7	0,0001	1,7	0,9554	3,7	0,9999
-1,8	0,0359	-3,8	0,0001	1,8	0,9641	3,8	0,9999
-1,9	0,0288	-3,9	0,0000	1,9	0,9713	3,9	1,0000

Значення функції $t_\gamma = \arg F\left(\frac{1+\gamma}{2}\right)$

γ	t_γ	γ	t_γ	γ	t_γ	γ	t_γ
0,80	1,282	0,86	1,475	0,91	1,694	0,97	2,169
0,81	1,310	0,87	1,513	0,92	1,750	0,98	2,325
0,82	1,340	0,88	1,554	0,93	1,810	0,99	2,576
0,83	1,371	0,89	1,597	0,94	1,880	0,9973	3,000
0,84	1,404	0,90	1,643	0,95	1,960	0,999	3,290
0,85	1,439			0,96	2,053		

Додаток 8

Таблиця χ^2 -розподілу

$$P(\chi^2 > \chi_\gamma^2) = \gamma = \int_{\chi_\gamma^2}^{\infty} \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \chi^{n-2} e^{-\frac{\chi^2}{2}} d\chi^2, \quad \chi^2 > 0$$

<i>n</i>	<i>γ</i>									
	0,999	0,995	0,99	0,98	0,975	0,95	0,9	0,8	0,75	0,70
1	0,0 ⁵ 157	0,0 ⁴ 393	0,0 ³ 157	0,0 ³ 628	0,0 ³ 982	0,0039	0,0158	0,0642	0,102	0,148
2	0,002	0,0100	0,0201	0,0404	0,0506	0,103	0,211	0,446	0,575	0,713
3	0,0243	0,0717	0,115	0,185	0,216	0,352	0,584	1,005	1,213	1,424
4	0,0908	0,207	0,297	0,429	0,484	0,711	1,064	1,649	1,923	2,195
5	0,210	0,412	0,554	0,752	0,831	1,145	1,610	2,343	2,675	3,000
6	0,381	0,676	0,872	1,134	1,237	1,635	2,204	3,070	3,455	3,828
7	0,598	0,989	1,239	1,564	1,690	2,167	2,833	3,822	4,255	4,671
8	0,857	1,344	1,646	2,032	2,180	2,733	3,490	4,594	5,071	5,527
9	1,152	1,735	2,088	2,532	2,700	3,325	4,168	5,380	6,899	6,393
10	1,479	2,156	2,558	3,059	3,247	3,940	4,865	6,179	6,737	7,267
11	1,834	2,603	3,053	3,609	3,816	4,575	5,578	6,989	7,584	8,148
12	2,214	3,074	3,571	4,178	4,404	5,226	6,304	7,807	8,438	9,034
13	2,617	3,565	4,107	4,765	5,009	5,892	7,042	8,634	9,299	9,926
14	3,041	4,075	4,660	5,368	5,629	6,571	7,790	9,467	10,165	10,821
15	3,483	4,501	5,229	5,985	6,262	7,261	8,547	10,307	11,036	11,721
16	3,942	5,142	5,812	6,614	6,908	7,962	9,312	11,152	11,912	12,624
17	4,416	5,697	6,408	7,255	7,564	8,672	10,085	12,002	12,792	13,531
18	4,905	6,265	7,015	7,906	8,231	9,390	10,865	12,857	13,675	14,440
19	5,407	6,844	7,633	8,567	8,907	10,117	11,651	13,716	14,562	15,352
20	5,921	7,434	8,260	9,237	9,591	10,851	12,443	14,578	15,452	16,266
21	6,447	8,034	8,897	9,915	10,283	11,591	13,240	15,445	16,344	17,128
22	6,983	8,643	9,542	10,600	10,982	12,338	14,041	16,314	17,240	18,101
23	7,529	9,260	10,196	11,293	11,688	13,091	14,848	17,187	18,137	19,021
24	8,085	9,886	10,856	11,992	12,401	13,848	15,659	18,062	19,037	19,943
25	8,649	10,520	11,524	12,697	13,120	14,611	16,473	18,940	19,939	20,867
26	9,222	11,160	12,198	13,409	13,844	15,379	17,292	19,820	20,843	21,792
27	9,803	11,808	12,879	14,125	14,573	16,151	18,114	20,703	21,749	22,719
28	10,391	12,461	13,565	14,847	15,308	16,928	18,939	21,588	22,675	23,647
29	10,986	13,121	14,256	15,574	16,047	17,708	19,768	22,475	23,567	24,577
30	11,588	13,787	14,953	16,306	16,791	18,493	20,599	23,364	24,478	25,508

Продовження табл.

<i>n</i>	<i>γ</i>									
	0,50	0,30	0,20	0,10	0,05	0,025	0,02	0,01	0,005	0,001
1	0,455	1,074	1,642	2,706	3,841	5,024	5,412	6,635	7,879	10,827
2	1,386	2,408	3,219	4,605	5,991	7,378	7,824	9,210	10,597	13,815
3	2,366	3,665	4,642	6,251	7,815	9,348	9,837	11,345	12,838	16,268
4	3,357	4,878	5,989	7,779	9,488	11,143	11,668	13,277	14,860	18,465
5	4,351	6,064	7,289	9,236	11,070	12,832	13,388	15,086	16,750	20,517
6	5,348	7,231	8,558	10,645	12,592	14,449	15,033	16,812	18,548	22,457
7	6,346	8,383	9,803	12,017	14,067	16,013	16,622	18,475	20,278	24,322
8	7,344	9,524	11,030	13,362	15,507	17,535	18,168	20,090	21,955	26,125
9	8,343	10,656	12,242	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	9,342	11,781	13,442	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	10,341	12,899	14,631	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	11,340	14,011	15,812	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	12,340	15,119	16,985	19,812	22,362	24,736	25,472	27,688	29,819	34,528
14	13,339	16,222	18,151	21,064	23,685	26,119	26,873	29,141	31,319	36,123
15	14,338	17,322	19,311	22,307	24,996	27,488	28,259	30,578	32,801	37,697
16	15,338	18,418	20,465	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	16,338	19,511	21,615	24,769	27,587	30,191	30,995	33,409	35,718	40,790
18	17,338	20,601	22,760	25,989	28,869	31,526	32,346	34,805	37,156	42,321
19	18,338	21,689	23,900	27,204	30,141	32,852	33,687	36,191	38,582	43,820
20	19,337	22,775	25,038	28,412	31,410	34,170	35,020	37,566	39,997	45,315

Закінчення табл.

<i>n</i>	γ									
	0,50	0,30	0,20	0,10	0,05	0,025	0,02	0,01	0,005	0,001
21	20,337	23,858	26,171	29,615	32,671	35,479	36,343	38,932	41,401	46,697
22	21,337	24,939	27,301	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	22,337	26,018	28,429	32,007	35,172	38,076	38,968	41,638	44,181	49,728
24	23,337	27,096	29,553	33,196	36,415	39,364	40,270	42,980	45,558	51,179
25	24,337	28,172	30,675	34,382	37,652	40,646	41,566	44,314	46,928	52,620
26	25,336	29,246	31,795	35,563	38,885	41,923	42,856	45,642	48,290	54,052
27	26,336	30,319	32,912	36,741	40,113	43,194	44,140	46,963	49,645	55,476
28	27,336	31,391	34,027	37,916	41,337	44,161	45,419	48,278	50,993	56,893
29	28,336	32,461	35,139	39,087	42,557	45,722	46,693	49,588	52,336	58,302
30	29,336	33,530	36,250	40,256	43,773	46,979	47,962	50,892	53,672	59,703

Додаток 9

Таблиця розподілу Стюдента (*t*-розподілу)

$$P(T > t_{\beta}) = \beta = \int_{t_{\beta}}^{\infty} \frac{\Gamma(\frac{N+1}{2})}{\sqrt{N\pi} \Gamma(N/2)} (1 + \frac{t^2}{N})^{-\frac{N+1}{2}} dt$$

<i>N</i>	β						
	0,45	0,4	0,35	0,3	0,25	0,2	0,15
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963
2	0,142	0,289	0,445	0,617	0,816	1,061	1,336
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093
11	0,128	0,259	0,395	0,539	0,695	0,873	1,083
12	0,128	0,258	0,393	0,537	0,692	0,868	1,076
13	0,128	0,258	0,392	0,535	0,690	0,865	1,071
14	0,127	0,257	0,393	0,534	0,688	0,862	1,067
15	0,127	0,257	0,391	0,533	0,687	0,860	1,064
16	0,127	0,256	0,390	0,531	0,684	0,856	1,058
17	0,126	0,256	0,389	0,530	0,683	0,854	1,055

Продовження табл.

<i>N</i>	β					
	0,1	0,05	0,025	0,01	0,005	0,0005
1	3,078	6,314	12,706	31,821	63,657	636,619
2	1,886	2,920	4,303	6,965	9,965	31,598
3	1,638	2,353	3,181	4,541	5,841	12,940
4	1,533	2,132	2,776	3,747	4,604	8,610
5	1,476	2,015	2,571	3,365	4,032	6,869
6	1,440	1,943	2,447	3,143	3,707	5,959
7	1,415	1,895	2,365	2,998	3,499	5,405
8	1,397	1,860	2,306	2,896	3,355	5,041
9	1,383	1,833	2,262	2,821	3,250	4,781
10	1,372	1,812	2,228	2,764	3,169	4,587

Закінчення табл.

12	1,356	1,782	2,179	2,681	3,055	4,318
14	1,345	1,761	2,145	2,624	2,977	4,140
16	1,337	1,746	2,120	2,583	2,921	4,015
18	1,330	1,734	2,101	2,552	2,878	3,922
20	1,325	1,725	2,086	2,528	2,845	3,850
25	1,316	1,708	2,060	2,485	2,787	3,725
30	1,310	1,697	2,042	2,457	2,750	3,646

Додаток 10

Таблиця розподілу Колмогорова

$$P(Z < z) = K(z) = 1 - 2 \sum_{\nu=1}^{\infty} (-1)^{\nu-1} e^{-2\nu^2 z^2}, \quad z > 0$$

z	$K(z)$	z	$K(z)$	z	$K(z)$	z	$K(z)$
0,1	0,000	0,6	0,136	1,1	0,822	1,6	0,988
0,2	0,000	0,7	0,289	1,2	0,888	1,7	0,994
0,3	0,000	0,8	0,456	1,3	0,932	1,8	0,997
0,4	0,003	0,9	0,607	1,4	0,960	1,9	0,998
0,5	0,036	1,00	0,730	1,5	0,978	2,0	0,999

Додаток 11

Таблиця розподілу омега-квадрат

$$P(\Omega^2 < \omega^2) = F(\omega^2)$$

ω^2	$F(\omega^2)$	ω^2	$F(\omega^2)$	ω^2	$F(\omega^2)$	ω^2	$F(\omega^2)$	ω^2	$F(\omega^2)$
0,00	0,000	0,20	0,7325	0,40	0,9277	0,60	0,9776	0,80	0,9927
0,01	0,001	0,21	0,7511	0,41	0,9320	0,61	0,9788	0,81	0,9931
0,02	0,003	0,22	0,7681	0,42	0,9360	0,62	0,9800	0,82	0,9934
0,03	0,025	0,23	0,7838	0,43	0,9397	0,63	0,9811	0,83	0,9938
0,04	0,067	0,24	0,7983	0,44	0,9432	0,64	0,9821	0,84	0,9941
0,05	0,1237	0,25	0,8116	0,45	0,9465	0,65	0,9831	0,85	0,9944
0,06	0,1860	0,26	0,8239	0,46	0,9496	0,66	0,9840	0,86	0,9947
0,07	0,2484	0,27	0,8353	0,47	0,9525	0,67	0,9845	0,87	0,9950
0,08	0,3081	0,28	0,8459	0,48	0,9552	0,68	0,9857	0,88	0,9952
0,09	0,3638	0,29	0,8557	0,49	0,9577	0,69	0,9865	0,89	0,9955
0,10	0,4151	0,30	0,8648	0,50	0,9602	0,70	0,9872	0,90	0,9957
0,11	0,4619	0,31	0,8733	0,51	0,9624	0,71	0,9879	0,91	0,9959
0,12	0,5045	0,32	0,8811	0,52	0,9645	0,72	0,9886	0,92	0,9962
0,13	0,5432	0,33	0,8884	0,53	0,9665	0,73	0,9892	0,93	0,9964
0,14	0,5784	0,34	0,8953	0,54	0,9684	0,74	0,9898	0,94	0,9966
0,15	0,6104	0,35	0,9016	0,55	0,9702	0,75	0,9904	0,95	0,9967
0,16	0,6395	0,36	0,9076	0,56	0,9719	0,76	0,9909	0,96	0,9969
0,17	0,6660	0,37	0,9131	0,57	0,9734	0,77	0,9914	0,97	0,9971
0,18	0,6902	0,38	0,9183	0,58	0,9749	0,78	0,9918	0,98	0,9972
0,19	0,7123	0,39	0,9232	0,59	0,9763	0,79	0,9923	0,99	0,9974

Додаток 12

Таблиця зваженого розподілу омега-квадрат

$$P(\Omega^2 < \omega^2) = F(\omega^2) \text{ (для вагової функції } \varphi(F_h(x)) = [F_h(x)(1 - F_h(x))]^{-1} \text{)}$$

ω^2	$F(\omega^2)$	ω^2	$F(\omega^2)$	ω^2	$F(\omega^2)$	ω^2	$F(\omega^2)$	ω^2	$F(\omega^2)$
0,00	0,0000	1,0	0,6427	2,0	0,9081	3,0	0,9726	4	0,9913
0,1	0,0000	1,1	0,6912	2,1	0,9190	3,1	0,9756	5	0,9971
0,2	0,0096	1,2	0,7324	2,2	0,9285	3,2	0,9783	6	0,9990

Закінчення табл.

0,3	0,0618	1,3	0,7676	2,3	0,9368	3,3	0,9806	7	0,9996
0,4	0,1512	1,4	0,7977	2,4	0,9440	3,4	0,9828	8	0,9999
0,5	0,2531	1,5	0,8235	2,5	0,9505	3,5	0,9846	9	0,9999
0,6	0,3520	1,6	0,8457	2,6	0,9560	3,6	0,9863		
0,7	0,4411	1,7	0,8648	2,7	0,9610	3,7	0,9877		
0,8	0,5189	1,8	0,8813	2,8	0,9654	3,8	0,9891		
0,9	0,5857	1,9	0,8957	2,9	0,9692	3,9	0,9902		

Додаток 13

Таблиця B -розподілу

$$P\left(\frac{\pi^4}{2} N \Omega^2 < b\right) = F(b)$$

b	$F(b)$	b	$F(b)$	b	$F(b)$	b	$F(b)$	b	$F(b)$
0,30	0,00000	1,3	0,60697	2,3	0,90185	3,3	0,97094	4,3	0,99075
0,40	0,00086	1,4	0,66131	2,4	0,91357	3,4	0,97414	4,4	0,99174
0,50	0,01158	1,5	0,70763	2,5	0,92377	3,5	0,97698	4,5	0,99261
0,60	0,4867	1,6	0,74704	2,6	0,93268	3,6	0,97949	4,6	0,99339
0,70	0,11594	1,7	0,78060	2,7	0,94047	3,7	0,98172	4,7	0,99409
0,80	0,20293	1,8	0,80922	2,8	0,94730	3,8	0,98370	4,8	0,99471
0,90	0,29652	1,9	0,83369	2,9	0,95329	3,9	0,98515	4,9	0,99627
1,00	0,38730	2,0	0,85469	3,0	0,95857	4,0	0,98702	5,0	0,99576
1,1	0,47027	2,1	0,87275	3,1	0,96322	4,1	0,98841	6,0	0,99857
1,2	0,54354	2,2	0,88335	3,2	0,96732	4,2	0,98965	7,0	0,99952

Додаток 14

ν^2 -розподіл і розподіл Фішера

Визначення. Нехай випадкові величини X_1 і X_2 мають χ^2 -розподіл відповідно з N_1 і N_2 ступенями свободи. Тоді випадкова величина

$$X = \nu^2(N_1, N_2) = \frac{X_1}{X_2} \text{ має } \nu^2\text{-розподіл з } N_1 \text{ і } N_2 \text{ ступенями свободи.}$$

Щільність ймовірностей цього розподілу

$$f(x; N_1, N_2) = \begin{cases} 0, & x \leq 0, \\ \left(\frac{N_1}{N_2}\right)^{\frac{N_1}{2}} \frac{\Gamma\left(\frac{N_1 + N_2}{2}\right)}{\Gamma\left(\frac{N_1}{2}\right)\Gamma\left(\frac{N_2}{2}\right)} x^{\frac{N_1-2}{2}} \left(1 + \frac{N_1}{N_2}x\right)^{-\frac{N_1+N_2}{2}}, & x > 0. \end{cases}$$

Властивості. Математичне сподівання ν^2 -розподілу $m_x = \frac{N_2}{N_2 - 2} (N_2 > 2)$,

дисперсія $D_x = \frac{2N_2^2(N_1 + N_2 - 2)}{N_1(N_2 - 2)^2(N_2 - 4)} (N_2 > 4)$, мода знаходиться в точці

$$\frac{N_2(N_1 - 2)}{N_1(N_2 + 2)}.$$

Сума квадратів m незалежних гауссівських стандартизованих випадкових величин має χ^2 -розподіл із m ступенями свободи, тому ν^2 -розподіл можна

розглядати як розподіл, що породжений відношенням двох нормованих сум N_1 і N_2 квадратів незалежних гауссівських стандартизованих випадкових величин:

$$X = v^2(N_1, N_2) = \frac{\frac{1}{N_1} \sum_{n=1}^{N_1} X_{1n}^2}{\frac{1}{N_2} \sum_{n=1}^{N_2} X_{2n}^2}.$$

Положення. v^2 -розподіл пов'язаний з розподілом Фішера виразом $Z = \frac{1}{2} \ln v^2(N_1, N_2)$. Щільність ймовірностей розподілу Фішера

$$f(z; N_1, N_2) = 2 \left(\frac{N_1}{N_2} \right)^{\frac{N_1}{2}} \frac{\Gamma\left(\frac{N_1 + N_2}{2}\right)}{\Gamma\left(\frac{N_1}{2}\right)\Gamma\left(\frac{N_2}{2}\right)} e^{N_1 z} \left(1 + \frac{N_1}{N_2} e^{2z}\right)^{-\frac{N_1 + N_2}{2}}.$$

Предметний покажчик

А

- аксіома ймовірності 1.1.4
- неперервності 1.1.4
- зліченної адитивності 1.1.3
- алгебра множин 1.1.2
- σ (сигма) 1.1.3
- амплітуда 3.5.1
- аналіз багатофакторний 11.1.1
- дисперсійний 11.1.1
- послідовний 9.4
- регресійний 8.1
- факторний 11.1.1

В

- вектор фінальних ймовірностей 4.2
- величина випадкова 1.2.1
- векторна 1.2.1
- гауссівська (нормальна) 1.2.2
- , дисперсія 1.2.3, 1.3.3
- , ентропія 1.2.3
- ймовірно визначена 1.2.1
- , кумулянт (семіінваріант) 1.2.3, 1.3.3
- , коефіцієнт асиметрії 1.2.3
- варіації 1.2.3
- ексцесу 1.2.3
- комплексна 1.3.3
- , математичне сподівання 1.2.3, 1.3.3
- , медіана 1.2.3
- , мода 1.2.3
- , момент 1.2.3
- коваріаційний 1.2.3, 1.3.3
- кореляційний 1.2.3, 1.3.3
- початковий 1.2.3, 1.3.3
- факторіальний 1.2.3
- центральний 1.2.3, 1.3.3
- , розмах 1.2.3
- , розподіл (див. розподіл)
- , середньоквадратичне відхилення (СКВ) 1.2.3, 1.3.3
- скалярна 1.2.1
- величини випадкові
- незалежні 1.3.2
- незалежні попарно 1.3.2
- незалежні у сукупності (взаємно незалежні) 1.3.2
- некорельовані (лінійно незалежні) 1.3.3
- некорельовані (попарно) 1.3.3
- ортогональні 1.3.3
- спільно гауссівські 1.5.1
- вибірка 6.1.1
- випадкова 6.1.1
- із генеральної сукупності 6.1.1, 10.1
- представницька 6.1.1
- репрезентативна 6.1.1
- , обсяг 6.1.1
- відгук 2.4.1
- відношення кореляційне 8.5
- вірогідності 9.2

- відображення 1.4.1
- відстань 9.5.1
- вплив 2.4.1
- втрати
- умовні середні 9.2
- , байєсівська функція 9.2

Г

- гіпотеза альтернативна 9.1
- конкуруюча 9.1
- непараметрична 9.5.1
- нульова 9.1
- однорідності даних 9.7
- основна 9.1
- параметрична 9.5.1
- проста 9.1
- статистична 9.1
- гістограма 6.1.2
- група повна 1.1.4
- несумісних подій 1.1.5

Д

- дослід 1.1.1
- незалежний 1.1.5
- дисперсія 1.2.3, 1.3.3, 2.2.2
- вибіркова 6.1.3
- залишкова 8.1, 8.4
- умовна 2.2.3
- середня 2.2.3
- умовних середніх 2.2

Е

- елемент (член) генеральної сукупності 6.1.1, 10.1
- ентропія 1.2.3, 1.3.3
- умовна 1.3.3

З

- залежність стохастична (випадкова) 8.1
- зв'язок лінійний 3.4
- обернений 3.6.1
- значення вибіркоче 6.1.1
- середнє 6.1.3
- некорельовані 3.4
- ортогональні 3.4

І

- інтеграл випадкової функції д.3
- інтервал довірчий 7.5.1
- іспит 1.1.1

Й

- ймовірність, аксіоматичне визначення 1.1.4
- довірча 7.5.1
- інтервально-перехідна 4.8
- інфінітезимальна (породжуюча) 4.6
- , класичне визначення 1.1.1
- переходу 4.2

- події 1.1.1
- статистична 1.1.1

К

- квантиль 1.2.3
- квартиль 1.2.3
- кластер 11.2
- коефіцієнт асиметрії 1.2.3
- варіації 1.2.3
- довіри 7.5.1
- дифузії 4.3
- ексцесу 1.2.3
- зносу 4.3
- кореляції 1.3.3
- – зведений 8.4
- – окремий 8.4
- рангової кореляції Спірмена 8.6
- регресії 8.2, 8.4
- критерій Байєса 9.2
- згоди 9.5.1
- – омега-квадрат 9.5.3
- – хі-квадрат 9.5.4
- ідеального спостерігача 9.2
- Колмогорова 9.5.2
- максимальної вірогідності 9.2
- максимуму апостеріорної ймовірності 9.2
- Неймана – Пірсона 9.2
- послідовний Вальда 9.4
- прийняття рішення 9.1
- потужності 9.1
- Смирнова 9.7
- кумулянт 1.2.3, 1.3.3

Л

- ланцюг Маркова (марковський) 4.2
- – вкладений 4.8
- – ергодичний 4.2
- – незвідний 4.2
- – неоднорідний 4.2
- – однорідний 4.2
- – простий 4.2
- – рівноважний 4.2
- – складний 4.2
- – стаціонарний 4.2

М

- мартингал 4.4.1
- матриця інформаційна 8.8.1
- – Фішера 7.2.3
- коваріаційних моментів 1.3.3
- помилок 8.8.1
- стохастична 4.2
- медіана розподілу 1.2.3
- межа довірча 7.5.1
- метод ковзного середнього 10.2.5
- метрика 9.5.1
- міра 1.1.3
- апріорної непевності 1.2.3
- дискретна 1.1.3
- зосереджена на підмножині 1.1.3

- ймовірнісна 1.1.4
- кінцева 1.1.3
- нормована 1.1.3
- множина 1.1.2
- вимірна 1.1.3
- дискретна 1.1.2
- , елемент 1.1.2
- кінцева 1.1.2
- континуальна 1.1.2
- незліченна (континуальна) 1.1.2
- неkontинуальна 1.1.2
- нескінченна 1.1.2
- перелічена 1.1.2
- скінченна 1.1.2
- упорядкована 9.5.1
- мода (розподілу) 1.2.3
- модель авторегресійна 8.7
- розподіленого лагу 8.7
- момент вибіркового 6.1.3
- коваріаційний 1.3.3
- кореляційний 1.3.3
- початковий 1.2.3, 6.1.3
- – абсолютний 1.2.3
- – факторіальний 1.2.3
- центральний 1.2.3, 1.3.3, 6.1.3

Н

- непевність апріорна 1.2.3
- нерівність Коші – Шварца (Коші – Буняковського) 2.2.2
- Рао – Крамера 7.2.3
- Чебишева 6.2.1
- несумісність 1.1.1

О

- область допустима 9.1
- критична 9.1
- огинаюча сигналу 3.2.3, д.4
- сімейства кривих 3.2.3
- оператор 1.4.1
- безінерційний 2.4.1
- випадковий 2.4.1
- детермінований 2.4.1
- інваріантний (до зсуву) 3.6.1
- інерційний 2.4.1
- комутативний 2.4.3
- лінійний інерційний інтегральний 3.6.1
- лінійний однорідний 2.4.1
- лінійний неоднорідний 2.4.1
- нестационарний 3.6.1
- переміщувальний 2.4.3
- стаціонарний 3.6.1
- , частотна характеристика 3.6.1
- операція доповнення 1.1.2
- об'єднання 1.1.2
- перетину 1.1.2
- оцінка 6.1.1
- апостеріорна 7.3.3
- апріорна 7.3.3
- асимптотично незміщена 7.2.1

- асимптотично ефективна 7.2.3
- – спільно ефективна 7.2.3
- , відносна ефективність 7.2.3
- достатня 7.2.4
- ефективна 7.2.3
- зміщена 7.2.1
- інтервальна 7.5.1
- максимальної вірогідності 7.3.2
- непараметрична 11.3.1
- параметрична 11.3.1
- робастна 11.3.1
- спільно ефективна 7.2.3
- спроможна 7.2.2
- стійка 11.3.1
- точкова 7.1

П

- параметр регуляризації 11.3.3
- перетворення 1.4.1
- Вінера – Хінчина 3.5.2
- перетини
 - незалежні 2.1.3
 - незалежні в сукупності 2.1.3
 - некорельовані 2.2.2
 - ортогональні 2.2.2
- періодограма 10.2.5
- план латинського квадрата 11.2
 - неповний 11.1.2
 - повний 11.1.2
 - ортогональний 8.8.2
- події незалежні 1.1.4
 - – у сукупності 1.1.4
 - несумісні 1.1.4
 - – попарно 1.1.4
- подія 1.1.4
 - випадкова 1.1.1
 - достовірна 1.1.4
 - неможлива 1.1.4
 - , повна група 1.1.4
 - елементарна 1.1.1, 1.1.4
- полігон частот 6.1.2
- помилка верифікації 9.1
 - другого роду (хибної тривоги) 9.1
 - першого роду (пропуску цілі) 9.1
- послідовність дискретна випадкова 4.2
- похідна випадкової функції д.3
- похибка систематична 7.2.1
- правило Байєса 9.2
 - мінімаксне 9.2
- прийняття рішення 9.1
- прогноз 8.1
- пропуск цілі 9.1
- простір
 - вибірок 1.1.1
 - вимірний 1.1.3
 - дійсний 1.2.1
 - дискретний 1.2.1
 - елементарних подій 1.1.1
 - із мірою 1.1.3
 - ймовірнісний 1.1.4

- комплексний 1.2.1
- метричний 9.5.1
- неперервний 1.2.1
- процес (див. ще функція) вінерівський 4.4.1
 - із незалежними приростами 4.4.1
 - марковський 4.1, 4.3
 - – дискретний 4.6
 - – дифузійний 4.3
 - – однорідний 4.6
 - – стаціонарний 4.6
 - – ергодичний 4.3, 4.6
 - – – однорідний у часі 4.3
 - напівмарковський 4.8
 - чисто дифузійний 4.4.1
- пряма середньоквадратичної регресії 8.2

Р

- ранг елемента 8.6
- ранжирування 8.6
- реакція 2.4.1
- реалізація 2.1.1, 6.1.1
- регресія 8.2
 - множинна 8.4
 - , поле 8.1
 - середньо квадратична 8.1, 8.4
 - – – лінійна 8.2, 8.4
- регресори 8.5
- ризик виробника 9.1
 - середній 9.2
- рівноможливість 1.1.1
- рівняння дифузійне 4.3
 - Маркова 4.2, 4.3
 - Колмогорова 4.3, 4.6
 - – зворотне 4.6
 - – пряме 4.6
 - лінійні диференційні 4.5.1
 - неперервності 4.5.1
 - Смолуховського 4.3
 - стохастичне диференційне 4.4.1
 - Фоккера – Планка – Колмогорова 4.3
- розмах випадкової величини 1.2.3
- розподіл Бернуллі 1.2.2
 - бета 1.2.2
 - бімодальний 1.2.3
 - біномний 1.1.5, 1.2.2
 - Вейбула 1.2.2
 - вибіркового 6.1.2
 - випадкової величини 1.2.2
 - – функцій 1.3.1
 - гамма 1.2.2
 - Гаусса (гауссівський або нормальний) 1.2.2
 - геометричний 1.2.2
 - гіпергеометричний 1.2.2
 - дискретний рівномірний 1.2.2
 - експоненціальний 1.2.2
 - інтервальний статистичний 6.1.2
 - ймовірності 1.1.4
 - Колмогорова 9.5.2
 - Коші 1.2.2

- Лапласа 1.2.2
- логнормальний 1.2.2
- мультимодальний 1.2.3
- Накагамі 1.2.2
- омега-квадрат 9.5.3
- зважений 9.5.3
- Парето 1.2.2
- Пірсона 1.2.2
- Пуасона 1.2.2
- Райса 1.2.2, 1.5.2
- Релея 1.5.2
- узагальнений (Райса) 1.5.2
- рівномірний 1.2.2
- Сімпсона (трикутний) 1.2.2
- Стьюдента д.6
- унімодальний 1.2.3
- χ^2 -квадрат 1.2.2
- щільність 1.2.2, 1.3.1
- розряд 6.1.2
- ряд варіаційний 6.1.2
- статистичний 6.1.2

С

- семіінваріант 1.2.3
- символ Кронекера 4.6
- спектр амплітудний 3.5.1
- амплітудно-фазовий комплексний 3.5.1
- дискретний 3.5.1
- енергетичний 3.5.2
- , ефективна смуга 3.5.2
- , ефективна ширина 3.5.2
- миттєвий 3.5.2
- неперервний 3.5.1
- фазовий 3.5.1
- фізичний 3.5.1
- сподівання математичне 1.2.3, 1.3.3, 2.2.1
- апостеріорне 2.2.3
- векторної функції 1.3.3
- випадкової величини 1.2.3
- – – векторної 1.3.3
- – – функції 2.2.1
- – умовне 2.2.3
- стан аперіодичний 4.2
- відзеркалюючий 4.2
- відображаючий 4.2
- ергодичний 4.2
- зворотний 4.2
- – додатний 4.2
- – нульовий 4.2
- незворотний 4.2
- періодичний 4.2
- поглинаючий 4.2
- сполучений 4.2
- статистика 6.1.1
- сукупність 6.1.1
- вибіркова 6.1.1
- генеральна 6.1.1
- схема явна 4.5.2
- неявна 4.5.2
- сигнал

- аналітичний (комплексний) д.4
- – –, амплітуда д.4
- – –, огибаюча д.4
- – –, фаза д.4

Т

- таблиця спряженості ознак 9.6
- течія випадкова 5.1
- без післядії 5.1
- дифузійна 4.5.1
- Ерланга
- – порядку k 5.4.2
- – – узагальнена 5.4.3
- із незалежними приростами 5.1
- з обмеженою післядією 5.4
- ймовірності 4.5.1
- , миттєва інтенсивність 5.1
- , миттєва щільність 5.1
- найпростіша 5.2
- неоднорідна 5.1
- однорідна 5.1
- ординарна 5.1
- Пальма 5.4.1
- подій 5.1
- Пуасона (пуасонівська) 5.2
- – нестационарна 5.3
- регулярна 5.1
- систематична 4.5.1
- стаціонарна 5.1
- – транспортна 5.4.4
- тренд 8.7

У

- умови однакові 1.1.1

Ф

- фаза початкова 3.5.1
- сигналу д.4
- фактори 8.1
- фільтр лінійний, що фізично реалізується 3.6.1
- – що фізично не реалізується 3.6.1
- функціонал 1.4.1
- функція аргументу 1.4.1
- асимптотично стаціонарна у вузькому розумінні 3.2.1
- байєсівська (втрат) 9.2
- вірогідності 6.1.1
- вузькосмугова 3.5.3
- вибіркова 2.1.1
- – кореляційна 10.4.2
- випадкова 2.1.1
- – векторна 2.3
- – – стаціонарна у широкому розумінні 3.2.4
- – дискретна 4.6
- –, дисперсія 2.2.2
- – другого порядку д.3
- – ергодична 3.3

- по відношенню до кореляційної функції 3.3
- по відношенню до математичного сподівання 3.3
- із стаціонарними першими приростами 3.2.2
- інтегровна на інтервалі д.3
- , коваріаційна функція 2.2.2
- комплексна 3.4
- , кореляційна функція 2.2.2
- нормована 2.2.2
- , математичне сподівання 2.2.1
- незалежна 2.3
- незалежна в сукупності 2.3
- неперервна д.3
- нестационарна 3.2.1
- періодично стаціонарна у вузькому розумінні 3.2.1
- , перетин 2.1.1
- , похідна д.3
- , простір станів 2.1.1
- , реалізація 2.1.1
- , середньоквадратичне відхилення 2.2.2
- стаціонарна у вузькому розумінні (строго) 3.2.1
- стаціонарна у вузькому розумінні порядку K 3.2.1
- стаціонарна у вузькому розумінні на інтервалі 3.2.1
- стаціонарна у широкому розумінні 3.2.1, 3.2.4
- сумісно стаціонарно пов'язана у вузькому розумінні 3.2.1
- сумісно стаціонарно пов'язана у широкому розумінні 3.2.1
- умовна 1.3.2, 2.1.2
- , гама д.2
- , дельта д.1
- гауссівська (нормальна) 3.1
- , інтервал кореляції 3.2.1, 3.2.3
- кореляційна 2.2.1, 2.2.2, 2.3, 3.4
- взаємна 2.3, 3.4
- нормована 2.2.2, 2.3, 3.4
- умовна 2.2.3
- коваріаційна 2.2.2, 3.4
- взаємна 3.4
- кумулянтна 2.2.1
- моментна 2.2.1
- некорельована 3.4
- однозначна обернена 1.4.2
- ортогональна 3.4
- регресії 8.1

- розподілу 1.2.2, 2.1.2
- істинна 6.1.2
- статистична 6.1.2
- інтервальна 6.1.2
- теоретична 6.1.2
- , частота 6.1.2
- спряжена за Гільбертом д.4
- структурна 3.2.2
- твірна 1.2.2
- моментів 1.2.2
- факторіальних 1.2.2
- характеристична 1.2.2, 1.3.1, 1.3.2, 2.1.2
- центральна моментна 2.2.1
- центрована 2.2.1
- циклостационарна 3.2.1
- частотної когерентності 3.5.4
- широкопasmового 3.5.3

Х

- характеристика амплітудно-частотна 3.6.1
- імпульсна перехідна 3.6.1
- ймовірнісна 1.2.2, 1.3.1
- лінійного оператора 3.6.1
- передаточна 3.6.1
- перехідна 3.6.1
- фазо-частотна 3.6.1

Ч

- частота вибіркового розподілу 6.1.2
- вибіркових значень 6.1.2
- події 1.1.1

Ш

- шум білий 3.5.3
- нестационарний 3.5.3

Щ

- щільність
- ймовірностей 1.2.2
- апостеріорна 2.1.3
- апріорна 2.1.3
- вибіркова спектральна 10.2.5
- умовна 1.3.2
- переходу 4.2
- потужності 3.5.2
- спектральна 3.5.2
- взаємна 3.5.4
- нормована 3.5.2
- розподілу 1.2.2, 2.1.2

Я

- явище випадкове 1.1.1

Список літератури

1. Айвазян С. А. и др. Прикладная статистика: исследование зависимостей. – М.: Финансы и статистика, 1985. – 487 с.
2. Амиантов И. Н. Избранные вопросы статистической теории связи. – М.: Сов. радио, 1971. – 416 с.
3. Анго А. Математика для электро- и радиоинженеров. – М.: Наука, 1967. – 779 с.
4. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физматгиз, 1963. – 500 с.
5. Андерсон Т. Статистический анализ временных рядов. – М.: Мир, 1976. – 755 с.
6. Барлоу Р., Прошан Ф. Статистическая теория надежности и испытания на безотказность. – М.: Наука, 1984. – 328 с.
7. Баруча-Рид А.Т. Элементы теории марковских процессов и их приложение. – М.: Наука, 1969. – 511 с.
8. Башаринов А. Е., Флейшман Б. С. Методы статистического последовательного анализа и их приложения. – М.: Сов. радио, 1962. – 350 с.
9. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. – М.: Наука, 1983. – 416 с.
10. Боровков А. А. Теория вероятностей. – М.: Наука, 1976. – 352 с.
11. Бронштейн И. Н., Семендяев К. А. Справочник по математике для инженеров и учащихся втузов. – М.: Наука, 1981. – 719 с.
12. Вальд А. Последовательный анализ. – М.: Физматгиз, 1960. – 215 с.
13. Ван дер Варден Б. Л. Математическая статистика. – М.: ИЛ, 1960. – 434 с.
14. Ван Трис Г. Теория обнаружения, оценок и модуляции. – М.: Сов. радио, 1972. – Т. 1. – 743 с.; 1975. – Т. 2. – 343 с.; 1977. – Т. 3. – 662 с.
15. Вентцель А. Д. Курс теории случайных процессов. – М.: Наука, 1975. – 319 с.
16. Вентцель Е. С. Теория вероятностей. – М.: Физматгиз, 1969. – 576 с.
17. Вучков И., Бояджиева Л., Солаков Е. Прикладной линейный регрессионный анализ. – М.: Финансы и статистика, 1987. – 239 с.
18. Гаек Я., Шидак Э. Теория ранговых критериев. – М.: Наука, 1971. – 375 с.
19. Гаткин Н. Г. и др. Помехоустойчивость типового тракта обнаружения сигналов. – К.: Техніка, 1971. – 204 с.
20. Гихман И. И., Скороход А. В. Введение в теорию случайных процессов. – М.: Наука, 1977. – 567 с.
21. Гихман И. И., Скороход А. В., Ядренко М. И. Теория вероятностей и математическая статистика. – К.: Вища школа, 1988. – 439 с.
22. Гнеденко Б. В. Курс теории вероятностей. – М.: Физматгиз, 1961. – 406 с.
23. Гнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания. – М.: Наука, 1966. – 431 с.
24. Горбань И.И. Справочник по теории случайных функций и математической статистике для научных работников и инженеров. – К.: Институт кибернетики им. В.М. Глушкова НАН Украины, 1998. – 150 с.
25. Губарев В.В. Таблицы характеристик случайных величин и векторов. – Новосибирск, Новосибирск. электротехн. ин-т, 1981. – 225 с. (Рук. деп. в ВИНТИ, №3146-81).
26. Гуткин Л. С. Теория оптимальных методов радиоприема при флуктуационных помехах. – М.: Сов. радио, 1972. – 447 с.
27. Давенпорт В. Б., Рут В. Л. Введение в теорию случайных сигналов и шумов. – М.: ИЛ, 1960. – 468 с.
28. Длин А. М. Математическая статистика в технике. М.: Сов. наука, 1958. – 465 с.
29. Дуб Д. Вероятностные процессы. – М.: ИЛ, 1956. – 606 с.
30. Дынкин Е. Б. Марковские процессы. – М.: Физматгиз, 1963. – 859 с.
31. Дюге Д. Теоретическая и прикладная статистика. – М.: Наука, 1972. – 384 с.
32. Дятлов Г. И., Кудрицкий В. Д. Вероятностные основы моделирования сложных систем. – К.: КВВАИУ, 1992. – 529 с.
33. Евсиков Ю.А., Чапурский В.В. Преобразование случайных процессов в радиотехнических устройствах. – М.: Высшая школа, 1977. – 264 с.
34. Зельнер А. Байесовские методы в эконометрии. – М.: Статистика, 1980. – 438 с.
35. Зиновьев А. Л., Филиппов Л. И. Введение в теорию сигналов и цепей. – М.: Высшая школа, 1975. – 261 с.

36. Ильичев В. И. и др. Статистическая теория обнаружения гидроакустических сигналов. – М.: Наука, 1992. – 414 с.
37. Казаков И. Е. Статистическая теория систем управления в пространстве состояний. – М.: Наука, 1975. – 432 с.
38. Карлин С. Основы теории случайных процессов. – М.: Мир, 1971. – 536 с.
39. Кендел М., Стюарт А. Статистические выводы и связи. – М.: Наука, 1973. – 899 с.
40. Коваленко И. Н., Кузнецов Н.Ю., Шуренков В.М. Случайные процессы. Справочник. – К.: Наукова думка, 1983. – 366 с.
41. Коваленко И. Н., Сарманов О. В. Краткий курс теории случайных процессов. – К.: Вища школа, 1978. – 262 с.
42. Коваленко И. Н., Филиппова А. А. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1973. – 368 с.
43. Кнопов П. С. Оптимальные оценки параметров стохастических систем. – К.: Наукова думка, 1981. – 151 с.
44. Козин И. В. Элементы теории оптимального обнаружения и приема сигналов. – Л.: ЛГУ, 1974. – 124 с.
45. Козлов М. В., Прохоров А. В. Введение в математическую статистику. – М.: МГУ, 1987. – 364 с.
46. Колмогоров А. Н. Теория вероятностей и математическая статистика. – М.: Наука, 1986. – 535 с.
47. Корн Г., Корн Т. Справочник по математике. – М.: Наука, 1973. – 832 с.
48. Королюк В. С. Стохастичні моделі систем. – К.: Либідь, 1993. – 136 с.
49. Королюк В. С. и др. Справочник по теории вероятностей и математической статистике. – М.: Наука, 1985. – 640 с.
50. Крамер Г. Математические методы статистики. – М.: Мир, 1976. – 648.
51. Кремер И. Я. и др. Пространственно-временная обработка сигналов. – М.: Радио и связь, 1984. – 222 с.
52. Куликов Е. И., Трифонов А. П. Оценка параметров сигналов на фоне помех. – М.: Сов. радио, 1978. – 295 с.
53. Кульбак С. Теория информации и статистика. – М.: Наука, 1967. – 408 с.
54. Купер Дж., Макгиллем К. Вероятностные методы анализа сигналов и систем. – М.: Мир, 1989. – 376 с.
55. Левин Б. Р. Теоретические основы статистической радиотехники. – М.: Сов. радио, 1968. – Т. 1. – 748 с.; 1969. – Т. 2. – 504 с.; 1976. – Т. 3. – 285 с.
56. Левин Б. Р. Теоретические основы статистической радиотехники. – М.: Радио и связь, 1989. – 653 с.
57. Левин Б. Р., Шварц В. Вероятностные модели и методы в системах связи и управления. – М.: Радио и связь, 1985. – 312 с.
58. Леман И. Проверка статистических гипотез. – М.: Наука, 1979. – 408 с.
59. Линник Ю. В. Избранные труды. Математическая статистика. – Л.: Наука, 1982. – 284 с.
60. Лоули Д., Максвелл А. Факторный анализ как статистический метод. – М.: Мир, 1967. – 144 с.
61. Лозв М. Теория вероятностей. – М.: ИЛ, 1962. – 719 с.
62. Марпл-мл. С. Л. Цифровой спектральный анализ и его приложения. – М.: Мир, 1990. – 584 с.
63. Маркел Дж. Д., Грэй А. Х. Линейное предсказание речи. – М.: Связь, 1980. – 308 с.
64. Мартынов Г. В. Критерии омега-квадрат. – М.: Наука, 1978. – 80 с.
65. Мельник М. Основы прикладной статистики. – М.: Энергоатомиздат, 1983. – 416.
66. Миддтон Д. Введение в статистическую теорию связи. – М.: Сов. радио, 1961. – Т. 1. – 782 с.; 1962. – Т. 2. – 829 с.
67. Мостеллер Ф., Тьюки Д. Анализ данных и регрессия. – М.: Финансы и статистика, 1982. – 239 с.
68. Мюллер П. и др. Таблицы по математической статистике. – М.: Финансы и статистика, 1982. – 278 с.
69. Переверзев Е.С. и др. Случайные сигналы в задачах оценки состояния технических средств. – Киев: Наукова думка, 1992. – 250 с.
70. Поллард Дж. Справочник по вычислительным методам статистики. – М.: Финансы и статистика, 1982. – 344 с.

71. Прикладная статистика: классификация и снижение размерности/ Под ред. С. А. Айвазяна. – М.: Финансы и статистика, 1982. – 198 с.
72. Применение цифровой обработки сигналов/ Под ред. Э. Оппенгейма. – М.: Мир, 1980. – 552 с.
73. Прохоров Ю. В., Розанов Ю. А. Теория вероятностей: основные понятия, предельные теоремы, случайные процессы. – М.: Наука, 1967. – 495 с.
74. Пугачев В. С. Теория случайных функций и ее применение к задачам автоматического управления. – М.: Физматгиз, 1968. – 883 с.
75. Пугачев В. С. Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.
76. Репин В. Г., Тартаковский Г. П. Статистический синтез при априорной неопределенности и адаптация информационных систем. – М.: Сов. радио, 1977. – 432 с.
77. Рихтмайер Р., Мортон К. Разностные методы решения краевых задач. – М.: Мир, 1972. – 344 с.
78. Розанов Ю. А. Лекции по теории вероятностей. – М.: Наука, 1968. – 120 с.
79. Розенберг В. Я., Прохоров А. И. Что такое теория массового обслуживания. – М.: Сов. радио, 1965. – 255 с.
80. Румшиский Л. З. Элементы теории вероятностей. – М.: Физматгиз, 1960. – 155 с.
81. Рунион Р. Справочник по непараметрической статистике. Современный подход. – М.: Финансы и статистика, 1982. – 198 с.
82. Рытов С. М. Введение в статистическую радиофизику. – М.: Наука, 1966. – 404 с.
83. Саати Т. Элементы теории массового обслуживания и ее применения. – М.: Сов. радио, 1965. – 509 с.
84. Сборник задач по теории вероятностей, математической статистике и теории случайных функций/ Под ред. А. А. Свешникова. – М.: Наука, 1965. – 632 с.
85. Сейдж Э., Мелс Дж. Теория оценивания и ее применение в связи и управлении. – М.: Связь, 1976. – 495 с.
86. Свешников А. А. Прикладные методы теории случайных функций. – М.: Наука, 1968. – 463 с.
87. Скороход А. В. Случайные процессы с независимыми приращениями. – М.: Наука, 1964. – 277 с.
88. Скороход А. В. Елементи теорії ймовірностей та випадкових процесів. – К.: Вища школа, 1975. – 295 с.
89. Скороход А. В. Лекції з теорії випадкових процесів. – К.: Либідь, 1990. – 168 с.
90. Смирнов Н. В. Теория вероятностей и математическая статистика. – М.: Наука, 1970. – 288 с.
91. Смирнов Н. В., Дунин-Барковский И. В. Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1965. – 511 с.
92. Сосулин Ю. Г. Теория обнаружения и оценивания стохастических сигналов. – М.: Сов. радио, 1978. – 320 с.
93. Справочник по прикладной статистике. – М.: Финансы и статистика, 1989. – Т. 1. – 510 с.; 1990. – Т. 2. – 526 с.
94. Стратанович Р. Л. Избранные вопросы теории флуктуации в радиотехнике. – М.: Сов. радио, 1961. – 558 с.
95. Стратанович Р. Л. Условные марковские процессы и их применение к теории оптимального управления. – М.: МГУ, 1966. – 352 с.
96. Теория обнаружения сигналов/ Под ред. П. А. Бакута. – М.: Радио и связь, 1984. – 440 с.
97. Тихонов В.И. Статистическая радиотехника. – М.: Сов. Радио, 1982. – 624 с.
98. Тихонов В. И., Кульман Н. К. Нелинейная фильтрация и квазикогерентный прием сигналов. – М.: Сов. радио, 1975. – 704 с.
99. Тихонов В.И., Миронов М.А. Марковские процессы. – М.: Сов. Радио, 1977. – 486 с.
100. Тихонов В. И., Харисов В.Н. Статистический анализ и синтез радиотехнических устройств и систем. – М.: Радио и связь, 1991. – 608 с.
101. Уилкс С. Математическая статистика. – М.: Наука, 1967. – 632 с.
102. Фалькович С. Е. Оценка параметров сигнала. – М.: Сов. радио, 1975. – 334 с.
103. Фалькович С. Е., Пономарев В. И., Шкварко Ю. В. Оптимальный прием пространственно-временных сигналов в радиоканалах с рассеиванием. – М.: Радио и связь, 1989. – 295 с.
104. Фано Р. Передача информации. Статистическая теория связи. – М.: Мир, 1965. – 438 с.

105. Феллер В. Введение в теорию вероятностей и ее приложения. – М.: Мир, 1967. – Т. 1. – 498 с.; Т. 2. – 752 с.
106. Фомин В.Н. Рекуррентное оценивание и адаптивная фильтрация. – М.: Наука, 1984. – 286 с.
107. Хан Г., Шапиро С. Статистические модели в инженерных задачах. – М.: Мир, 1969. – 390 с.
108. Хастингс Н., Пикок Дж. Справочник по статистическим распределениям. – М.: Статистика, 1980. – 95 с.
109. Хинчин А.Я. Работы по математической теории массового обслуживания. – М.: Физматгиз, 1963. – 235 с.
110. Хьюбер П. Робастность в статистике. М.: Мир, 1984. – 303 с.
111. Шефтель З. Г. Теорія ймовірностей. – К.: Вища школа, 1994. – 192 с.
112. Шишлянникова В. Н., Шишлянникова С. Н. Численные и графические методы. – Рига: РИИГВФ, 1963. – 314 с.
113. Шор Я. Б. Статистические методы анализа и контроля качества и надежности. – М.: Сов. радио, 1962. – 361 с.
114. Янке Е., Эмде Ф., Леш Ф. Специальные функции. Формулы, графики, таблицы. – М.: Наука, 1968. – 344 с.

Зміст

Вступ	3
1. Основні положення теорії ймовірностей	5
1.1. Математичне поняття ймовірності	5
1.1.1. Класична та статистична ймовірності	5
1.1.2. Основні положення теорії множин	6
1.1.3. Основи теорії міри	8
1.1.4. Аксиоматичне визначення ймовірності	9
1.1.5. Основні теореми і формули теорії ймовірностей	10
1.2. Скалярні випадкові величини	12
1.2.1. Основні визначення	12
1.2.2. Ймовірнісні характеристики випадкової величини	12
1.2.3. Неймовірнісні (числові) характеристики випадкової величини	16
1.3. Векторні випадкові величини	20
1.3.1. Ймовірнісні характеристики векторної випадкової величини	20
1.3.2. Умовні ймовірнісні характеристики векторної випадкової величини	22
1.3.3. Неймовірнісні (числові) характеристики векторної випадкової величини	24
1.4. Перетворення випадкових величин	28
1.4.1. Поняття про математичні перетворення	28
1.4.2. Перетворення скалярної випадкової величини	30
1.4.3. Перетворення векторної випадкової величини	31
1.5. Приклади випадкових величин	34
1.5.1. Гауссівська випадкова величина	34
1.5.2. Розподіли, пов'язані з гауссівськими	38
2. Основні положення теорії випадкових функцій	41
2.1. Уявлення про випадкові функції	41
2.1.1. Визначення і класифікація	41
2.1.2. Ймовірнісні характеристики випадкових функцій	43
2.1.3. Умовні ймовірнісні характеристики випадкових функцій	44
2.2. Неймовірнісні характеристики випадкових функцій	47
2.2.1. Моментні і кумулянтні функції випадкових функцій	47
2.2.2. Одновимірні і двовимірні моментні функції випадкових функцій	49
2.2.3. Умовні моментні і кумулянтні функції випадкових функцій	52
2.3. Векторні випадкові функції	55
2.4. Перетворення випадкових функцій	57
2.4.1. Класифікація операторів	57
2.4.2. Перетворення випадкових функцій безінерційними операторами	58
2.4.3. Перетворення випадкових функцій лінійними інерційними операторами	62
3. Кореляційна теорія випадкових функцій	65
3.1. Гауссівські випадкові функції	65
3.2. Стаціонарні та нестаціонарні випадкові функції	66
3.2.1. Стаціонарні випадкові функції та їх опис	66
3.2.2. Опис нестаціонарних випадкових функцій	70

3.2.3. Інтервали кореляції	70
3.2.4. Стаціонарні векторні випадкові функції	72
3.3. Ергодичні випадкові функції	72
3.4. Кореляційні та коваріаційні функції комплексних випадкових функцій	73
3.5. Спектральний аналіз стаціонарних випадкових функцій	75
3.5.1. Основи спектрального аналізу	75
3.5.2. Спектральне представлення випадкових функцій	77
3.5.3. Вузькосмугові і широкосмугові випадкові функції. Білий шум	80
3.5.4. Взаємна спектральна щільність потужності	81
3.6. Перетворення стаціонарних випадкових функцій лінійними операторами	82
3.6.1. Математичний опис лінійних інерційних операторів	82
3.6.2. Перетворення випадкових функцій лінійними інерційними операторами	86
3.6.3. Перетворення білого шуму лінійними операторами	88
4. Марковські випадкові процеси	90
4.1. Початкові уявлення про марковські процеси та їх класифікація	90
4.2. Ланцюг Маркова	90
4.3. Неперервний марковський процес	94
4.4. Приклади дифузійних марковський процесів	97
4.4.1. Чисто дифузійний (вінерівський) процес	97
4.4.2. Гауссівський марковський процес	99
4.5. Методи розв'язання рівнянь Колмогорова	101
4.5.1. Аналітичні методи розв'язання рівнянь Колмогорова	101
4.5.2. Чисельні методи розв'язання рівнянь Колмогорова	102
4.6. Дискретний марковський процес	103
4.7. Застосування теорії марковський процесів	106
4.8. Напівмарковський випадковий процес	107
5. Випадкові течії	110
5.1. Класифікація і основні характеристики випадкових течій	110
5.2. Найпростіша течія	113
5.3. Узагальнення найпростішої течії	114
5.4. Течії з обмеженою післядією	115
5.4.1. Течії Пальма	115
5.4.2. Течії Ерланга	117
5.4.3. Узагальнена течія Ерланга	118
5.4.4. Транспортна течія	119
5.5. Використання теорії випадкових течій	120
6. Основні положення вибіркової теорії	122
6.1. Способи представлення статистичної інформації про випадкову величину і властивості цієї інформації	122
6.1.1. Генеральна і вибірка сукупності	122
6.1.2. Варіаційний ряд. Статистичний розподіл вибірки. Статистична функція розподілу	124
6.1.3. Моменти вибіркового розподілу	127
6.2. Закон великих чисел	127
6.2.1. Нерівність Чебишева	128
6.2.2. Теорема Чебишева	128

6.2.3. Теорема Бернуллі	129
7. Оцінка параметрів і законів розподілу випадкової величини	130
7.1. Поняття про точкову оцінку параметрів розподілу	130
7.2. Властивості оцінок параметрів розподілу	130
7.2.1. Зміщені і незміщені оцінки	130
7.2.2. Спроможні оцінки	132
7.2.3. Ефективні оцінки	132
7.2.4. Достатні оцінки	139
7.3. Методи отримання оцінок параметрів розподілу випадкових..... величин	140
7.3.1. Метод моментів	140
7.3.2. Метод максимуму вірогідності	141
7.3.3. Метод максимуму апостеріорної щільності ймовірності	142
7.4. Закони розподілу оцінок параметрів гауссівського розподілу	143
7.5. Інтервальне оцінювання	146
7.5.1. Загальні поняття про інтервальне оцінювання	146
7.5.2. Наближений метод розрахунку меж довірчого інтервалу за заданою довірчою ймовірністю.....	147
7.5.3. Точні методи побудови довірчих інтервалів для параметрів випадкової величини	148
7.6. Оцінка ймовірності події по частоті її появи	150
8. Основи регресійного аналізу	153
8.1. Регресійні залежності	153
8.2. Лінійна середньоквадратична регресія	157
8.3. Оцінка параметрів лінійної середньоквадратичної регресії	158
8.4. Множинна регресія	159
8.5. Нелінійна середньоквадратична регресія	160
8.6. Коефіцієнт рангової кореляції.	162
8.7. Регресійний аналіз випадкових функцій	163
8.8. Оцінка коефіцієнтів лінійної регресійної моделі випадкової функції	166
8.8.1. Метод найменших квадратів	166
8.8.2. Планування експериментів	168
9. Основи теорії перевірки статистичних гіпотез	170
9.1. Основні поняття теорії перевірки статистичних гіпотез	170
9.2. Критерій якості прийняття рішення	173
9.3. Перевірка гіпотез про значення математичного сподівання гауссівської випадкової величини	175
9.4. Послідовний аналіз	177
9.5. Перевірка непараметричних статистичних гіпотез	179
9.5.1. Поняття про критерії згоди	179
9.5.2. Критерій згоди Колмогорова	181
9.5.3. Критерій згоди омега-квадрат	182
9.5.4. Критерій згоди хі-квадрат	183
9.6. Перевірка гіпотези про незалежність випадкових величин	184
9.7. Перевірка гіпотези про приналежність вибірок одному розподілу .	186
10. Оцінка характеристик випадкових функцій	189
10.1. Особливості застосування вибіркового методу при визначенні характеристик випадкових функцій	189
10.2. Оцінка математичного сподівання випадкової функції	190

10.2.1. Оцінка математичного сподівання випадкової функції по ансамблю реалізацій	190
10.2.2. Оцінка математичного сподівання випадкової функції по одній реалізації	191
10.2.3. Похибка дискретизації при отриманні оцінки математичного сподівання стаціонарної випадкової функції	192
10.2.4. Оцінка математичного сподівання стаціонарної випадкової функції по аргументу й множині реалізацій	194
10.2.5. Оцінка математичного сподівання нестаціонарної випадкової функції по одній реалізації	195
.....	196
10.3. Оцінка дисперсії випадкової функції	196
10.3.1. Оцінка дисперсії випадкової функції по ансамблю реалізацій	196
10.3.2. Оцінка дисперсії випадкової функції по одній реалізації	197
10.3.3. Похибки дискретизації при оцінці дисперсії	199
10.4. Оцінка кореляційної функції	199
10.4.1. Оцінка кореляційної функції по ансамблю реалізацій	199
10.4.2. Оцінка кореляційної функції стаціонарної випадкової функції	201
10.4.3. Оцінка кореляційної функції стаціонарної випадкової функції по дискретних відліках	202
10.5. Оцінка спектральної щільності стаціонарної випадкової функції	204
11. Додаткові питання	204
11.1. Дисперсійний аналіз	204
11.1.1. Основи дисперсійного аналізу	206
11.1.2. Багатофакторний дисперсійний аналіз	209
11.2. Основи кластерного аналізу	210
11.3. Робастні статистичні методи	210
11.3.1. Робастні оцінки	214
11.3.2. Перевірка гіпотез з елементами робастної обробки даних	216
11.3.3. Регуляризація оцінок	217
Додатки	217
Додаток 1. Дельта-функція	217
Додаток 2. Гама-функція	218
Додаток 3. Неперервність, диференційованість та інтегровність..	220
випадкових функцій	223
Додаток 4. Перетворення Гільберта й аналітичний сигнал	224
Додаток 5. Хі-квадрат розподіл	225
Додаток 6. Розподіл Стьюдента (t -розподіл)	225
Додаток 7. Таблиці гауссівської функції розподілу	227
Додаток 8. Таблиця χ^2 -розподілу	228
Додаток 9. Таблиця розподілу Стьюдента (t -розподілу)	228
Додаток 10. Таблиця розподілу Колмогорова	228
Додаток 11. Таблиця розподілу омега-квадрат	229
Додаток 12. Таблиця зваженого розподілу омега-квадрат	229
Додаток 13. Таблиця F -розподілу	231
Додаток 14. ν^2 -розподіл і розподіл Фішера	236
Предметний показчик	236

Список літератури.....