

В.В. Михайленко

**ТЕОРІЯ ЙМОВІРНОСТЕЙ,
МАТЕМАТИЧНА СТАТИСТИКА
ТА ВИПАДКОВІ ФУНКЦІЇ**

КУРС ЛЕКЦІЙ

Рекомендовано Міністерством освіти і
науки України як навчальний посібник для
студентів нематематичних спеціальностей
вищих навчальних закладів



2003

УДК 519.21
М 69

В.В. Михайленко

Теорія ймовірностей, математична статистика та випадкові функції. Курс лекцій: Навчальний посібник. – Житомир: ЖІТІ, 2003. – 292 с.

ISBN 966-683-063-9

Дано елементарний виклад у вигляді курсу лекцій традиційних тем теорії ймовірностей, математичної статистики та кореляційної теорії випадкових функцій.

Для студентів нематематичних спеціальностей вищих навчальних закладів, що вивчають дисципліну “Теорія ймовірностей та математична статистика”.

Іл.: 66. – Табл.: 36. – Бібліогр.: 21 назв.

Рецензенти:

доктор фіз.-мат. наук, професор Г.Л. Кулініч
доктор фіз.-мат. наук, професор Р.Є. Майборода
доктор техн. наук, професор І.Г. Грабар

УДК 519.21

ISBN 966-683-063-9

© В.В. Михайленко, 2003

ВІД АВТОРА

В основу посібника покладено курс лекцій, який читається автором в Житомирському державному технологічному університеті. Передусім посібник написано для студентів нематематичних спеціальностей вищих навчальних закладів, які вивчають дисципліну “Теорія ймовірностей та математична статистика”. Виклад матеріалу у вигляді лекцій не є традиційним. Кожна лекція розрахована орієнтовно на аудиторне прослуховування протягом двох академічних годин. Тому в межах однієї лекції може закінчуватись одна і розпочинатись інша тема. В залежності від спеціальності та математичної підготовки студентів окремі теми лекцій, а також окремі лекції можуть виключатись з курсу, зберігатись у запропонованому в посібнику обсязі або розширяватись з використанням літератури, наведеної в кінці посібника.

Посібник може використовуватись для початкового самостійного вивчення традиційних тем теорії ймовірностей, математичної статистики та кореляційної теорії випадкових функцій за умови математичної підготовки читача на рівні вимог вищого технічного навчального закладу. Для читача, що не володіє елементарною математичною культурою, робота з посібником навряд чи принесе користь.

Виклад теорії ймовірностей проводиться на найпростішому рівні, проте аксіоматична основа теорії зберігається. Розглядаються скінченні та зчисленні ймовірнісні простори, а також ймовірнісні простори з геометричними ймовірностями та ймовірностями, що вводяться за допомогою звичайних рімановських інтегралів від невід'ємних функцій. Основна увага у другому розділі приділяється статистичній перевірці статистичних гіпотез. Вивчення третього розділу передбачає наявність певних навичок комплексного аналізу та знайомство з перетворенням Фур’є.

При вивчені будь-якої математичної дисципліни смисл її математичних залежностей значно легше зрозуміти з математичного аналізу і на прикладах розв’язування простих модельних задач, ніж на реальних фізичних прикладах. Тому основні положення і методи теорії ймовірностей та інших розділів даного посібника ілюструються в основному на простих модельних задачах, а не на конкретних задачах з техніки, економіки тощо. Звичайно, в розвитку інтуїції спеціаліста ці реальні задачі відіграють надзвичайно важливу роль. Проте автор рекомендує студентам та іншим читачам переходити до опрацювання існуючих і до побудови своїх власних математичних моделей реальних ймовірнісних об’єктів та процесів лише після набуття необхідних навичок саме на простих модельних прикладах.

Наведені у посібнику розв’язки певної кількості задач не звільняють від необхідності самостійного розв’язування задач. Для цього можуть використовуватися задачники зі списку літератури.

Автор висловлює подяку В.О. Ковалю за змістовні консультації під час написання посібника та рецензентам Г.Л. Кулінічу, Р.Є. Майдороді, І.Г. Грабару за критичні зауваження, які сприяли поліпшенню книги.

ВСТУП

Вивчаючи різні дисципліни до знайомства з теорією ймовірностей, ви мали справу з такими математичними моделями явищ або експериментів, в яких умови однозначно визначають результат. Такі моделі прийнято називати **детермінованими**. Класичним прикладом є теоретична механіка. Якщо, наприклад, у деякий момент часу задати положення і швидкість матеріальної точки, то її подальший рух під дією сили тяжіння визначається однозначно, оскільки відповідне диференціальне рівняння має єдиний розв'язок. Проте застосувати цю модель до реальних фізичних тіл, наприклад, до польоту кулі, не завжди вдається. Початкова швидкість кулі наперед точно невідома; при різних пострілах вона приймає різні значення, а тому траєкторія однозначно визначатись уже не буде. Отже, невизначеність тут вноситься випадковістю початкової швидкості. Якщо коливання значень початкової швидкості невеликі (наприклад, менші за похибку при чисельному інтегруванні рівнянь руху), то можна використовувати детерміновану модель, в якій рух однозначно визначається початковими умовами.

Невизначеність результату при збереженні основних умов експерименту спостерігається для широкого кола явищ. При підкиданні монети ми не можемо передбачити результат: випаде монета “гербом” доверху чи ні. Результати окремих вимірювань, отриманих одним і тим же приладом в одних і тих же умовах, різні. Деталі, виготовлені на одному і тому ж верстаті, що працює в одному і тому ж режимі, отримуються не зовсім однаковими. Звичайно, є багато причин, які впливають на таку невизначеність результату. В останньому прикладі це може бути неоднорідність матеріалу, вібрація інструменту, коливання електроенергії тощо. Врахувати вплив всіх цих причин на результат неможливо, тому що кількість їх надто велика і закони їхньої дії невідомі.

Отже, спільна дія дуже великої кількості різноманітних причин, кожна з яких окремо не може вплинути на результат експерименту, призводить до того, що результат експерименту не визначається однозначно. Кажуть, що результат такого експерименту є **випадковим**.

Результати окремих експериментів (підкидань монети, вимірювань тощо) поводять себе дуже “неправильно”. І теорія ймовірностей не ставить перед собою задачу передбачити результат такого окремого експерименту. Будь-яка наука безсила це зробити. Проте при спостереженні за результатами великої кількості тотожних експериментів проявляються певні закономірності. Саме дослідження таких закономірностей на математичних моделях випадкових явищ і є предметом теорії ймовірностей як науки.

Коротка історична довідка. Народження теорії ймовірностей відноситься до середини 17 ст. і пов'язане з вивченням і намаганням створити теорію азартних ігор. Перші праці, в яких зароджувались основні поняття теорії ймовірностей, належать таким видатним вченим свого часу, як П. Ферма (1601–1665), Б. Паскаль (1623–1662), Х. Гюйгенсу (1629–1695) та іншим вченим.

Наступний етап розвитку теорії ймовірностей пов'язаний з іменами Я. Бернуллі (1654–1705), А. Муавра (1667–1754), Т. Байеса (1702–1761), П. Лапласа (1749–1827), К. Гаусса (1777–1855), С. Пуассона (1781–1840),

О. Коші (1789–1857) та інших вчених. Цей етап можна охарактеризувати як етап формування основ теорії ймовірностей. Саме на цьому етапі в науку було введено поняття класичної ймовірності (Я. Бернуллі, А. Муавр), сформовано теореми про суму (Т. Байес) та добуток (А. Муавр) ймовірностей, введено початкове поняття випадкової величини (С. Пуассон) тощо. Майже два століття основним полем застосування теорії ймовірностей були азартні ігри, страхування та демографія. І все ж до середини 19 ст. теорія ймовірностей ще не сформувалась в математичну науку. Її застосування до вивчення явищ природи були інколи непереконливими і слабко обґрунтованими, а невизначеність і нечіткість основних понять призводили до парадоксів, в яких нелегко було розібратися. В результаті до теорії ймовірностей багато вчених відносилися не серйозно.

Початок перетворенню теорії ймовірностей у самостійну математичну науку зі специфічною тематикою й методами поклав видатний російський вчений П.Л. Чебишов (1821–1894). Чебишов написав з теорії ймовірностей протягом досить великого інтервалу часу (більше сорока років) всього чотири статті. Проте вони були настільки багаті ідеями, що саме в руслі ідей Чебишова теорія ймовірностей будувалась протягом століття.

Наприкінці 19 та на початку 20 століття з'явилися серйозні вимоги з боку природознавства (задачі теорії похибок, теорії стрільби, теорії статистики та ін.), які сприяли подальшому розвитку теорії ймовірностей. Цей, найбільш плідний період розвитку теорії ймовірностей, пов'язаний з іменами учнів і послідовників Чебишова – А.А. Маркова (1856–1922) і О.М. Ляпунова (1857–1918) та інших вчених.

Будь-яка змістовна математична наука будується за такою схемою: спочатку йде накопичення фактів, потім їх узагальнення і далі побудова науки на чіткій логічній основі. Саме на початку 20 ст. ряд вчених – Е. Борель (1871–1956), С.Н. Бернштейн (1880–1968), Р. Мізес (1883–1953) та інші доклали серйозних зусиль до критичного перегляду вихідних посилань і понять теорії ймовірностей. З цього часу теорія ймовірностей почала будуватись на формально-логічній основі, коли основні факти теорії отримуються не на інтуїтивному рівні, а як чіткі логічні наслідки з вихідних означенень і аксіом. Перша аксіоматична побудова теорії ймовірностей належить С.Н. Бернштейну і відноситься до 1917 р. Проте система аксіом Бернштейна проіснувала недовго. У 30-х роках минулого століття інший російський вчений А.М. Колмогоров (1903–1987) ввів систему аксіом, яка до цього часу використовується для побудови теорії ймовірностей у всьому світі. Колмогоровим і його науковою школою закінчено дослідження з класичної проблематики теорії ймовірностей, що йде від Чебишова. Йому належать основоположні праці з теорії випадкових процесів, математичної статистики та їх застосувань у механіці, геофізиці, біології, техніці. Не можна також не згадати таких видатних вчених 20-го століття, як П. Леві (1886–1971), О.Я. Хінчина (1894–1959), В. Феллера (1906–1970), Н. Вінера (1894–1956), Р. Фішера (1890–1962); з вчених, які працювали на Україні – Й.І. Гіхмана (1918–1985), Ю.В. Лінника (1915–1972), М.Ф. Кравчука (1892–1942), Б.В. Гнеденка (1912–1997) та інших.

Докладно про історію розвитку теорії ймовірностей можна дізнатися з підручника І.М. Коваленка, Б.В. Гнеденка “Теория вероятностей” (Київ, 1990).

Розділ 1. ЕЛЕМЕНТИ ТЕОРИЇ ЙМОВІРНОСТЕЙ

Лекція № 1

Поняття випадкового експерименту і випадкової події.

Схематизація випадкових подій. Операції над подіями. Відносна частота випадкової події. Ймовірність події

Теорія ймовірностей перш за все є дисципліною математичною. Тому, як і в інших математичних дисциплінах, в основі ймовірнісних моделей лежать формально-логічні визначення ряду основних понять, таких, наприклад, як “подія”, “ймовірність”, “випадкова величина” та ін. і певна система аксіом.

Вихідними поняттями теорії ймовірностей є поняття *випадкового* (стохастичного) *експерименту*, *випадкової події* і *ймовірності випадкової події*.

Під випадковими (стохастичними) експериментами будемо розуміти експерименти, результати яких не можна передбачити заздалегідь. Всякий випадковий експеримент проводиться при деяких основних умовах. Позначимо ці умови через S . *Будемо цікавитись лише такими випадковими експериментами, які можна повторювати при збереженні умов S будь-яку кількість разів (принаймні теоретично).*

Наведемо декілька простих прикладів випадкових експериментів.

- 1) Підкидається монета. Не можна точно сказати, якою стороною вона впаде, “гербом” чи “цифрою”.
- 2) Підкидається шестигранний гральний кубик. Не можна заздалегідь передбачити, яка кількість очок при цьому випаде.
- 3) З урни, що містить певну кількість червоних та білих кульок, виймається одна кулька. Не можна наперед сказати, якого кольору кульку буде вийнято.

Можна навести безліч прикладів випадкових експериментів у техніці, фізиці, економіці та ін. Але ми і далі досить часто будемо звертатись саме до наведених вище прикладів, тому що на них найпростіше прослідкувати за основними поняттями і законами теорії ймовірностей без додаткової концентрації уваги на самій суті експерименту.

Предметом спостереження у тому чи іншому випадковому експерименті може бути деякий процес, фізичне явище тощо. Для

реально відтворюваного експерименту поняття результату, як такого, що спостерігається, означає, що існує принципова можливість реєстрації даного результату експерименту за допомогою того чи іншого пристроя (у найпростішому випадку, візуально). Будь-який результат, що спостерігається при проведенні випадкового експерименту, інтерпретується як **випадкова подія**. Отже, **випадкова подія** – це така подія, яка є результатом випадкового експерименту, тобто подія, яка при виконанні комплексу умов S даного експерименту може відбутися або не відбутися.

Оскільки наперед не можна сказати, відбудеться випадкова подія чи ні, то, на перший погляд, здається, що ніяких закономірностей в явищах природи і суспільства, в яких існують випадкові події, бути не може. Але насправді це не так. Численні спостереження над випадковими подіями привели до відкриття нового типу закономірностей, які називаються **ймовірнісними** або **стохастичними** (в протилежність детермінованим). Найпростішим прикладом стохастичного закону є такий закон: при достатньо великому числі підкидань монети одним і тим же способом (тобто, при виконанні одного і того ж комплексу умов S) майже у половині випадків випадає герб. Наведемо результати трьох експериментів:

Експериментатор	Кількість підкидань	Випав “герб”	Відношення числа випадань “герба” до числа підкидань монети
Ж. Бюффон	4040	2048	0,5069
К. Пірсон	12000	6019	0,5016
К. Пірсон	24000	12012	0,5005

Наведені дані підтверджують, що не дивлячись на те, що при однократному проведенні випадкового експерименту появляється випадкової події непередбачувана (може як відбутися, так і не відбутися), при багатократному повторенні цього експерименту спостерігаються певні закономірності в результатах.

В основі прийнятого на сьогоднішній день теоретико-множинного методу побудови теорії ймовірностей, якого ми будемо надалі притримуватись, лежить припущення, що кожному випадковому експерименту можна поставити у відповідність деяку множину Ω . Елементи цієї множини повинні давати найбільш повну інформацію про можливі результати даного експерименту. Множина Ω

називається **множиною елементарних подій**, а її елементи – **елементарними подіями**.

Приклад 1. Проводиться експеримент: один раз підкидається монета. Множина елементарних подій цього експерименту має вигляд $\Omega = \{\text{Г}, \text{Ц}\}$, де Г – означає випадання “герба”, Ц – випадання “цифри”. Отже, під елементарними подіями розуміють результати експерименту, які одночасно реалізовуватись не можуть.

Приклад 2. Монету підкидають два рази. Множиною елементарних подій для цього експерименту є $\Omega = \{\text{ГГ}, \text{ГЦ}, \text{ЦГ}, \text{ЦЦ}\}$, де, наприклад, ГЦ означає, що при першому підкиданні з'явився “герб”, а при другому – “цифра”.

Приклад 3. Підкидають гральний кубик, на якому вибито очки від 1 до 6. Множиною елементарних подій є множина, яка складається з чисел 1, 2, 3, 4, 5, 6, тобто $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Приклад 4. Гральний кубик підкидають m разів. Як множину елементарних подій можна розглядати множину всіх m -вимірних векторів i_1, i_2, \dots, i_m , де кожна компонента i_k приймає значення 1, 2, 3, 4, 5, 6, тобто $\Omega = \{i_1, i_2, \dots, i_m\}$, $i_k = \{1, 2, 3, 4, 5, 6\}$, $1 \leq k \leq m$.

У розглянутих прикладах множини елементарних подій є скінченими множинами, так що кількість елементів цих множин можна порахувати. У багатьох задачах теорії ймовірностей ми будемо мати справу з експериментами, які мають нескінченну кількість елементарних подій.

Приклад 5. Монета підкидається до першого випадання “герба”. Множина елементарних подій має вигляд: $\Omega = \{\omega_1, \omega_2, \omega_3, \dots, \omega_n, \dots\}$, де $\omega_n = \underbrace{\text{ЦЦ} \dots \text{ЦГ}}_{n-1}$ означає, що “герб” випав при n -му підкиданні. У

даному прикладі Ω – нескінчена, але зчисленна множина (див. додаток 1).

Приклад 6. У коло з радіусом R і центром у початку координат **навмання** кидають точку. За множину елементарних подій приймається множина точок $\Omega = \{(x, y) : x^2 + y^2 \leq R^2\}$. Множина Ω у даному випадку є неперервною множиною. Тут і далі слово “навмання” означає, що шанси відбудутися для кожної з елементарних подій однакові, тобто припускається, що жодна з елементарних подій не має переваг перед іншими.

Отже, кожному випадковому експерименту ставиться у відповідність множина елементарних подій Ω . Разом з тим, для кожного з таких експериментів можуть розглядатись події, які не є елементарними, але про кожну з яких можна сказати, відбулась вона в даному експерименті чи ні. Кажуть, що ці події спостерігаються в даному експерименті. Нехай A – довільна подія, яка спостерігається в експерименті. Оскільки кожна елементарна подія ω дає повну інформацію про результат експерименту, то знаючи, що результат експерименту описується елементом ω , завжди можна сказати, відбулась подія A чи ні. По відношенню до події A всю множину елементарних подій Ω можна розбити на дві підмножини A' і A'' наступним чином: якщо результат експерименту описується елементом ω з множини A' , то подія A в цьому експерименті відбулася, якщо ж $\omega \in A''$ – то – не відбулася. Елементи ω з множини A' називаються елементарними подіями, сприятливими для події A . Кажуть, що множина A' є відображенням або інтерпретацією події A на множині Ω . Подія A ототожнюється з множиною A' , $A = A'$. Отже, подія A – це деяка підмножина множини Ω , яка складається з усіх тих елементарних подій ω , які сприяють події A . Якщо результат експерименту описується елементом $\omega \in \Omega$ і $\omega \in A$, то кажуть, що подія A відбулася. Якщо ж $\omega \notin A$, то подія A в цьому експерименті не відбулася.

Приклад 7. Монету підкидають два рази. Нехай A – подія, яка полягає в тому, що хоча б один раз з'явиться герб. Тоді

$$\Omega = \text{ГГ}, \text{ГЦ}, \text{ЦГ}, \text{ЦЦ}, \quad A = \text{ГГ}, \text{ГЦ}, \text{ЦГ} .$$

Приклад 8. Нехай один раз підкидається гральний кубик і A – подія, яка полягає в тому, що число очок, яке випало, ділиться на три. Тоді $\Omega = 1, 2, 3, 4, 5, 6$, $A = 3, 6$.

Зauważення 1. Побудову множини Ω (якщо її не задано при описанні експерименту) проводять, виходячи з вимоги, щоб всі результати, які цікавлять дослідника у даному експерименті, однозначно описувались на основі цієї множини. Іншими словами, якщо нас цікавлять події A, B, C , які спостерігаються у даному експерименті, то множина Ω повинна будуватися так, щоб існували підмножини цієї множини, рівносильні подіям A, B, C . Оскільки поняття “елементарної події” строго не визначено, то ця задача не має единого розв’язку і залежить від набору подій, які нас цікавлять. Якщо ж вимагати, щоб правило однозначного описання виконувалось для всіх подій, що спостерігаються в експерименті, то в цьому випадку поняття “елементарної події” стає більш визначенім, а саме, в сукупності всіх підмножин множини Ω , які відповідають всім

подіям, що спостерігаються, елементарні події є **одноелементними** підмножинами.

Вище поняття “подій” зведено до поняття множини, яка є підмножиною множини елементарних подій Ω . Над множинами можна виконувати певні операції, аналогічні операціям над числами в алгебрі (див. додаток 1). А це означає, що і над подіями можна виконувати такі ж операції. Розглянемо їх.

1) Перш за все введемо поняття **вірогідної** (достовірної) події.

Вірогідна (достовірна) подія – це подія, яка співпадає з усією множиною Ω , тобто подія, яка відбувається при кожному проведенні експерименту.

2) Порожня множина \emptyset , тобто множина, яка не містить жодного елемента з Ω , ототожнюється з подією, яку називають **неможливою**, тобто такою, яка у даному експерименті відбутися не може.

3) Кажуть, що з події A **випливає** подія B , якщо $A \subset B$. Тобто, якщо подія A відбулася, то і подія B також відбулася. За означенням приймається, що $\emptyset \subset A$.

4) Події A і B називаються **рівними** (рівносильними), якщо рівні відповідні множини A і B , тобто $A = B$ – $A \subset B$ і $B \subset A$.

5) **Сумою** (або об’єднанням) подій A та B називається подія $A + B$, якій відповідає множина $A \cup B$, тобто подія, яка відбувається тоді і тільки тоді, коли відбувається хоча б одна з подій A або B .

Приклад 9. Кидають точку в два круги, що перетинаються. Тоді подією $A + B$ є попадання в один із кругів (в тому числі і в їх спільну частину (рис. 1)).

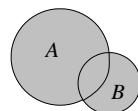


Рис.1

6) **Добутком** (або перерізом) двох подій A та B називається подія AB , якій відповідає множина $A \cap B$, тобто подія, яка відбувається тоді і тільки тоді, коли одночасно відбувається і подія A і подія B .

У прикладі 9 подія AB – це попадання у спільну частину двох кругів.

7) Події A і B називаються **несумісними**, якщо їх добуток є подія неможлива, тобто $AB = \emptyset$. У противному разі події A та B називаються **сумісними**. Отже, події A та B несумісні у даному експерименті, якщо вони не можуть відбутися одночасно.

8) Якщо A – випадкова подія, то подія \bar{A} називається подією, *протилежною* до події A , тобто \bar{A} відбувається тоді і тільки тоді, коли не відбувається A . Отже, $A + \bar{A} = \Omega$, $A\bar{A} = \emptyset$.

9) *Різницєю* подій A та B називається подія $A - B$, якій відповідає множина $A \setminus B$, тобто подія, яка відбувається тоді і тільки тоді, коли відбувається подія A , але не відбувається подія B . Очевидно, що $\bar{A} = \Omega - A$.

10) Поняття суми та добутку подій переноситься на зчисленну кількість подій. Подія

$$A_1 + A_2 + \dots + A_n + \dots = \bigcup_{n=1}^{\infty} A_n$$

відбувається тоді і тільки тоді, коли відбувається хоча б одна з подій A_n , $n = 1, 2, \dots$

Подія

$$A_1 A_2 \dots A_n \dots = \bigcap_{n=1}^{\infty} A_n$$

відбувається тоді і тільки тоді, коли відбуваються одночасно всі події A_n , $n = 1, 2, \dots$

Приклад 10. Стрільба з гармати ведеться до першого попадання в ціль. Введемо випадкову подію A_n – попадання при n -тому пострілі, $n \geq 1$. Тоді випадкову подію A – зроблено парну кількість пострілів, можна подати $A = A_2 + A_4 + \dots + A_{2n} + \dots = \bigcup_{n=1}^{\infty} A_{2n}$.

Всі властивості операцій над множинами, розглянуті в додатку 1, автоматично переносяться і на операції над подіями.

Означення. Будемо говорити, що події $A_1, A_2, \dots, A_n, \dots$

$A_k \neq \emptyset$, $k \geq 1$ утворюють *повну групу подій*, якщо вони попарно несумісні і сума цих подій є подія достовірна (вірогідна), тобто

$$A_i A_j = \begin{cases} \emptyset, & i \neq j \\ A_i, & i = j \end{cases} \text{ і } A_1 + A_2 + \dots + A_n + \dots = \Omega.$$

Очевидно, що елементарні події утворюють повну групу подій.

Розглянемо деякий випадковий експеримент і подію A , яка спостерігається в цьому експерименті. Повторимо експеримент n разів. Нехай $k_n A$ – число експериментів, в яких відбулась подія A . Це число називають абсолютною частотою події A або просто частотою.

Відношення $P_n^* A = \frac{k_n A}{n}$ називають **відносною частотою** події

A у проведений серії з n експериментів.

Відносна частота події має такі властивості:

- 1) $0 \leq P_n^* A \leq 1$;

- 2) $P_n^* \Omega = 1$, де Ω – вірогідна (достовірна) подія;

- 3) якщо A і B – дві події, які спостерігаються у даному експерименті і є **несумісними** $AB = \emptyset$, то $P_n^* A+B = P_n^* A + P_n^* B$.

Перша властивість є наслідком того, що $k_n A \leq n$. Друга властивість випливає з того, що достовірна подія відбувається при кожному експерименті з даної серії. Третя властивість також очевидна, оскільки для несумісних подій A та B $k_n A+B = k_n A + k_n B$.

Відносна частота події може бути знайдена тільки в результаті проведення серії з n експериментів і, взагалі кажучи, змінюється, якщо проводиться інша серія з n експериментів або якщо змінюється n . Проте практика підтверджує, що при достатньо великих n для більшості таких серій експериментів відносна частота зберігає майже постійну величину, причому відхилення спостерігається тим рідше, чим більше n . Наведена вище таблиця показує, що відносна частота появи “герба” у довгих серіях експериментів мало відрізняється від $\frac{1}{2}$.

Якщо при великих n відносна частота $P_n^* A$ події A мало відрізняється від деякого фіксованого числа p , то кажуть, що подія A **статистично стійка**, а число p є **ймовірністю** події A . Таке означення ймовірності отримало назву **частотного** або **статистичного**.

Зauważення 2. Статистична стійкість буде нашою основною вимогою до випадкової події. Теорія ймовірностей застосовна не до всіх подій, які залежать від випадку, тобто можуть відбутися або не відбутися, а застосовна лише до статистично стійких подій. Саме нерозуміння цього основного твердження є причиною багатьох помилкових висновків, які нібито обґрунтуються методами теорії ймовірностей. Підкреслимо, що статистична стійкість випадкової події передбачає виконання нашої основної вимоги до випадкового експерименту, а саме його багатократну повторюваність при незмінних умовах, теоретично – скільки завгодно разів.

Як бачимо, поняття ймовірності тісно пов'язане з властивістю стійкості відносної частоти. Наведене емпіричне визначення ймовірності події характеризує природно-науковий зміст поняття ймовірності, але не є його формальним визначенням. Щоб прийти до формального визначення ймовірності, потрібно аксіомувати властивості ймовірності, які випливають з того природного змісту цього поняття, який ми хочемо йому приписати. При формалізації поняття ймовірності ми повинні наділити це поняття властивостями відносної частоти, тобто притримуватись точки зору, що ймовірність події A є число, близьке до відносної частоти появи події A в довгих серіях тотожних експериментів.

Лекція № 2

Аксіоматичне означення ймовірності події. Ймовірнісний простір. Класична ймовірнісна схема

У минулій лекції ми формалізували поняття випадкової події, звівши його до поняття підмножини. Наступним нашим кроком буде введення формального поняття ймовірності випадкової події. Виходимо з того, що ймовірність об'єктивно існує (як об'єктивно існує, наприклад, маса тіла) і на мові математики “ймовірність випадкової події A дорівнює деякому числу P ”. Іншими словами, ймовірність – це числовая функція $P(A)$, аргументом якої є випадкова подія A . Областю визначення функції дійсної змінної не завжди є всі дійсні числа. Ймовірність, як функцію випадкової події, також не завжди можна визначити для всіх випадкових подій (підмножин множини Ω). Доводиться обмежуватись деякою системою (класом) підмножин із Ω .

Нехай Ω – довільна множина елементарних подій, а F – деяка система підмножин множини Ω .

Означення 1. Система F називається алгеброю подій, якщо

- 1) $\Omega \in F$, тобто достовірна подія належить до системи F ;
- 2) якщо події A та B належать до системи F , то події $A+B$, AB , $A-B$ також належать до F .

Як наслідок, маємо $\emptyset = \Omega - \Omega \in F$, $\bar{A} = \Omega - A \in F$, тобто до F належить як неможлива подія, так і подія, протилежна до будь-якої події A з F .

Про систему подій F , яка є алгеброю, кажуть, що вона замкнена відносно введених операцій над подіями.

Очевидно, що найменшою системою подій, яка утворює алгебру, є система $F = \emptyset, \Omega$. Алгеброю буде також система $F = \emptyset, A, \bar{A}, \Omega$, де $A \subset \Omega$. Переконайтесь у цьому самостійно. Легко довести, що система F всіх підмножин множини Ω , включаючи \emptyset , є також алгеброю. Зокрема, для множини Ω , що містить n елементів, система всіх підмножин F є алгеброю з 2^n елементами (див. додаток 1).

Приклад 1. 1) Підкидається монета. $\Omega = \{\Gamma, \Gamma\}$. Алгеброю подій є $F = \{\emptyset; \Gamma; \Gamma; \{\Gamma, \Gamma\} = \Omega\}$, $N(F) = 2^2 = 4$.

Приклад 2. Підкидається гральний кубик. Тоді $\Omega = \omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6$,

$$F = \emptyset, \omega_1, \dots, \omega_6; \omega_1, \omega_2, \dots; \omega_1, \omega_2, \omega_3, \dots; \omega_1, \omega_2, \omega_3, \omega_4, \dots;$$

$$\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \dots; \omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6 = \Omega, N F = 2^6 = 64.$$

Наведемо ще два приклади алгебри подій.

Приклад 3. На півінтервал a, b навмання кидають точку. Множиною елементарних подій цього експерименту є $\Omega = x : a \leq x < b$. Отже, елементарні події – це точки півінтервалу. Нехай F – система півінтервалів c, d , $a \leq c < d \leq b$ та їх скінченних сум. Легко переконатись, що F – алгебра. Розглянемо окрему точку $c \subset \Omega$ та множину раціональних точок $Q = x : a \leq x < b$, x – раціональне число, $Q \subset \Omega$. Тоді c та Q не належать до алгебри F .

Приклад 4. В одиничний квадрат навмання кидають точку. Тоді $\Omega = x, y : 0 \leq x \leq 1, 0 \leq y \leq 1$. Отже, елементарними подіями є точки квадрата. Розглянемо систему F квадрованих підмножин множини Ω , тобто підмножин, які мають площину. Відомо, що об'єднання, переріз і різниця квадрованих множин є також квадрованою множиною. Тому система F – алгебра. Розглянемо множину $Q = x, y : 0 \leq x \leq 1, 0 \leq y \leq 1; x, y$ – раціональні числа. Очевидно, що Q – підмножина множини Ω , $Q \subset \Omega$. Проте Q не належить до алгебри F , оскільки не має площини.

Для експериментів з нескінченною множиною елементарних подій Ω досить часто доводиться розглядати складні випадкові події, які є сумами або добутками зчисленної кількості інших випадкових подій (див. лекцію 1). У таких випадках потрібно вимагати, щоб алгебра

подій була одночасно і так званою σ - алгеброю (“сигма”- алгеброю), а саме, замкненою відносно операцій над зчисленною кількістю подій.

Означення 2. Алгебра подій F називається σ - алгеброю, якщо з того, що $A_n \in F$, $n = 1, 2, \dots$ випливає, що $\bigcup_{n=1}^{\infty} A_n \in F$ і $\bigcap_{n=1}^{\infty} A_n \in F$.

Приклад 5 (приклад σ -алгебри). Нехай $\Omega = R = x : -\infty < x < \infty$ – числовий вісь. Розглянемо систему \tilde{F} множин цієї осі, які можна утворити з множин вигляду $x : x < a$ (далі будемо використовувати скорочений запис $x < a$) застосуванням скінченної або зчисленної кількості операцій об'єднання і перерізу множин та операції різниці множин, причому можна покласти $a = +\infty$, тобто вважаємо, що $R \in \tilde{F}$. До системи \tilde{F} входять множини вигляду $x \geq a$ (різниця $R \setminus x < a$), множини $a \leq x < b$ (переріз $x \geq a \cap x < b$), окремі точки $x = a$ (переріз $\bigcap_{n=1}^{\infty} \left\{ a \leq x < a + \frac{1}{n} \right\}$), множини $a \leq x \leq b$ (об'єднання $a \leq x < b \cup x = b$), множини $a < x < b$ (різниця $a \leq x < b \setminus x = a$), множини $a < x \leq b$ (об'єднання $a < x < b \cup x = b$) та інші множини. Оскільки множина Q раціональних чисел є зчисленною, то їх можна пронумерувати: $Q = r_1, r_2, \dots, r_n, \dots$. Тоді $Q = \bigcup_{n=1}^{\infty} x = r_n$.

Отже, множина Q також входить до \tilde{F} . До \tilde{F} входить і множина ірраціональних чисел $R \setminus Q$. Практично всі множини, які зустрічаються у застосуваннях, належать до \tilde{F} . Множини з \tilde{F} називаються **борелівськими**. Очевидно, що система множин \tilde{F} є σ -алгеброю. Її називають σ -алгеброю борелівських множин.

Аналогічно, можна ввести σ -алгебру борелівських множин на площині, використовуючи множини $x, y : x < a, y < b$. Множина Q з прикладу 4 є борелівською.

Зauważення 1. Для кращого розуміння того, навіщо вводиться поняття σ -алгебри подій, доцільно нагадати ситуацію, що в свій час мала місце в звичайній алгебрі чисел, де операція знаходження кореня не виконувалась, якщо залишатись в межах не тільки раціональних (наприклад, $\sqrt{2}$ – ірраціональне, а не раціональне число), а й дійсних чисел (наприклад, \sqrt{a} ,

де $a < 0$, не існує). В результаті раціональні числа були доповнені ірраціональними, а потім і комплексними числами. Тим самим було побудовано поле (алгебру) комплексних чисел, в якому операція знаходження кореня завжди виконується.

Після введення операцій над подіями також виникає природне питання про виконання цих операцій, тобто, чи не може трапитись так, що в результаті введених операцій над подіями утвориться деяка “ірраціональна” подія, непередбачена теорією і така, що не може здійснитись при жодному комплексі умов. В σ -алгебрі подій подібна ситуація виключається. Очевидно, що сказане вище стосується тільки тих моделей, для яких множина елементарних подій Ω – незчисленна.

Оскільки при введенні поняття ймовірності доводиться обмежуватись системою підмножин множини Ω , які утворюють σ -алгебру F , уточнимо поняття події, як такої, що спостерігається в експерименті. За означенням подія A **спостерігається**, якщо $A \subset \Omega$ і $A \in F$. Якщо ж $A \subset \Omega$, але $A \notin F$, то будемо говорити, що подія A **не спостерігається** у даному експерименті. Отже, в загальному випадку клас подій, що спостерігається, вужчий, ніж система всіх підмножин множини Ω .

Зauważення 2. Якщо в прикладах 3, 4 залишатись у межах алгебр, то окрім точки множин Ω не можуть інтерпретуватись як події, що спостерігаються. З переходом до відповідних σ -алгебр борелівських множин окрім точки стають такими, що спостерігаються.

Той чи інший вибір σ -алгебри F підмножин множини Ω залежить від природи Ω . Простою є ситуація у випадку, коли Ω – скінчenna або зчисленна множина, $\Omega = \omega_1, \omega_2, \dots, \omega_n, \dots$. Таку множину елементарних подій будемо називати далі **дискретною**. Очевидно, що система F всіх підмножин дискретної множини Ω є σ -алгеброю. Важливо, що у цьому випадку довільну випадкову подію (підмножину) $A \subset \Omega$ можна отримати з елементарних подій ω_k шляхом скінченої або зчисленної кількості операцій над ними. Тому для дискретної множини Ω в якості σ -алгебри завжди вибирається система всіх підмножин цієї множини і, отже, кожна підмножина $A \subset \Omega$ є подією, що спостерігається.

Складніша ситуація у випадку неперервних (незчисленних) множин елементарних подій з прикладів 3, 4. У цих випадках система всіх підмножин множини Ω також є σ -алгеброю. Але отримати, наприклад, півінтервал c, d , який природно вважати подією, шляхом зчисленної кількості операцій над елементарними подіями (точками відрізка) неможливо. Проте зчисленної кількості операцій цілком

достатньо, якщо їх застосувати до інтервалів. Тому для неперервних множин Ω , як у прикладах 3, 4, роль σ -алгебри відіграють відповідні системи борелівських множин. Як видно з прикладу 5, це досить широкий клас множин. В усікому разі борелівських множин цілком достатньо для потреб теорії ймовірностей.

Перейдемо до поняття ймовірності, яке формалізується системою аксіом:

Аксіома 1. Кожній випадковій події A з σ -алгебри подій F можна поставити у відповідність невід'ємне число $P(A)$, яке називається ймовірністю цієї події:

$$P(A) \geq 0 \quad \forall A \in F.$$

Аксіома 2. Ймовірність достовірної події дорівнює одиниці:

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

Аксіома 3. Якщо події A та B з σ -алгебри подій F несумісні, то ймовірність їх суми дорівнює сумі ймовірностей:

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

Відмітимо, що аксіоми 1–3 повністю узгоджуються з властивостями частоти подій (див. лекцію 1).

Наведемо деякі **властивості** ймовірностей, які є безпосередніми наслідками аксіом:

- 1) $P(\bar{A}) = 1 - P(A);$
- 2) $P(\emptyset) = 0;$
- 3) $P(A+B) = P(A) + P(B) - P(AB);$
- 4) $P(A+B) \leq P(A) + P(B);$
- 5) якщо $A \subset B$, то $P(B \setminus A) = P(B) - P(A);$
- 6) якщо $A \subset B$, то $P(A) \leq P(B);$
- 7) $0 \leq P(A) \leq 1.$

Доведення. 1) Оскільки $A + \bar{A} = \Omega$ і A та \bar{A} – несумісні, то $P(\bar{A}) + P(A) = 1 \Rightarrow P(\bar{A}) = 1 - P(A);$

- 2) якщо $A = \Omega$, то $\bar{A} = \emptyset \Rightarrow P(\emptyset) = 1 - P(\Omega) = 0;$

3) події $A+B$ та B можна записати у вигляді суми несумісних подій: $A+B = A+B\Omega = A+B(A+\bar{A}) = A+BA+B\bar{A} = A+B\bar{A};$

$B = B\Omega = BA + B\bar{A}.$ Тому $P(A+B) = P(A) + P(B\bar{A})$ і $P(B) = P(BA) + P(B\bar{A}).$ Звідси отримуємо рівність 3).

Властивості 4)-7) пропонується довести **самостійно**.

Трійку (Ω, F, P) , де Ω – множина елементарних подій, F – σ -алгебра подій, P – числові функції на F , яка задовольняє аксіомам 1-3, називають **ймовірністю простором** у широкому розумінні. Для побудови змістової математичної теорії аксіоми 1-3 слід доповнити розширеною аксіомою додавання.

Аксіома 3'. Якщо події $A_1, A_2, \dots, A_k, \dots$ з σ -алгебри подій F попарно несумісні, то

$$P\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P A_k .$$

Використовуючи аксіому 3', можна довести таку теорему.

Теорема (про неперервність ймовірності). Якщо послідовність подій $A_1, A_2, \dots, A_n, \dots$ така, що $A_{n+1} \subset A_n$, $n = 1, 2, \dots$ і $\bigcap_{n=1}^{\infty} A_n = A$, то

$$\lim_{n \rightarrow \infty} P A_n = P A .$$

Доведення. Скористаємося рівністю $A_n = A + \sum_{k=n}^{\infty} A_k \bar{A}_{k+1}$, $n = 1, 2, \dots$

Всі доданки цієї рівності попарно несумісні. Тому за аксіомою 3' отримуємо: $P A_n - P A = \sum_{k=n}^{\infty} P A_k \bar{A}_{k+1} .$

Зокрема, при $n = 1$ маємо рівність $P A_1 - P A = \sum_{k=1}^{\infty} P A_k \bar{A}_{k+1}$, де ряд у правій частині збігається. Різниця $P A_n - P A$ є залишком цього ряду. Тому $\lim_{n \rightarrow \infty} P A_n = P A .$

Кажуть, що побудовано ймовірнісну модель експерименту, якщо побудовано ймовірнісний простір (Ω, F, P) , тобто вказано множину елементарних подій Ω , σ -алгебру подій F і введено ймовірність P . Теорія ймовірностей саме і “починається” після того, як побудовано **ймовірнісний простір**.

Нехай події $H_i \in F$ ($i = 1, 2, \dots, n$) утворюють повну групу подій. Тоді згідно з аксіомами 2) і 3) $P(H_1) + P(H_2) + \dots + P(H_n) = 1$.

Це означає, що однічна ймовірність достовірної події розподілена на множині несумісних подій, які утворюють повну групу. У загальному випадку відповідність між подіями деякої σ -алгебри подій F і ймовірностями цих подій називається **розподілом ймовірностей**.

Побудова ймовірнісного простору є основним етапом математичної формалізації того чи іншого випадкового експерименту. Сама задача моделювання реальних фізичних експериментів, взагалі кажучи, виходить за межі теорії ймовірностей. Найбільш складним при побудові ймовірнісного простору є задання розподілу ймовірностей. Наведені аксіоми не дають ніяких вказівок про числові значення ймовірностей тих подій, які нас цікавлять, а визначають лише загальні властивості, які повинна мати ймовірність як чисрова функція. Тому, залишаючись в межах аксіоматичної теорії, задачу про розподіл ймовірностей розв'язати не можна. Вибір тих чи інших значень для ймовірності здійснюється на основі додаткових міркувань, що випливають з практики.

Приклад 6. Підкидається гральний кубик. Побудувати ймовірнісний простір, який відповідає цьому експерименту.

$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ – множина елементарних подій.
 $F = \{\text{система всіх підмножин множини } \Omega, \text{ включаючи } \emptyset\}.$

Введемо розподіл ймовірностей формулою

$$P(A) = \frac{N_A}{6}, \quad (1)$$

де N_A – кількість елементарних подій ω_i , які входять до події A (наприклад, якщо $A = \{\text{випала парна кількість очок}\}$, то $A = \{\omega_2, \omega_4, \omega_6\}$ і $N_A = 3$). Аксіоми 1-3 виконуються. Дійсно, $P(A) \geq 0$, оскільки $N_A \geq 0$; $P(\Omega) = 1$, оскільки $N_\Omega = 6$. Нехай A та B – довільні несумісні події, тобто $AB = \emptyset$ і $N_A = k$, $N_B = m$. Тоді $N_{A+B} = N_A + N_B$ і $P(A+B) = P(A) + P(B)$. Отже, функція (1) є ймовірністю і простір Ω, F, P побудовано.

Розподіл ймовірностей для даного експерименту можна задати й іншим способом. Нехай p_1, p_2, \dots, p_6 – шість довільних невід'ємних чисел, які задовольняють умові $p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1$. Введемо ймовірність елементарних подій як $P(\omega_k) = P_{\omega_k} = p_k$, а ймовірність довільної події $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}\}$, де i_1, i_2, \dots, i_k – довільні числа від 1 до 6, як

$$P(A) = P_{\omega_{i_1}} + P_{\omega_{i_2}} + \dots + P_{\omega_{i_k}}. \quad (2)$$

Легко показати, що для даного розподілу ймовірностей аксіоми 1-3 також виконуються. Попередній випадок отримуємо, якщо покладемо

$p_1 = p_2 = \dots = p_6 = \frac{1}{6}$. Отже, наведений приклад показує, що

ймовірність не визначається однозначно системою аксіом. Вибір розподілу ймовірностей згідно з формулою (1) пов'язаний з міркуваннями “симетрії” грального кубика, коли немає ніяких підстав надавати перевагу випаданню однієї з граней кубика в порівнянні з іншою гранню. Якщо ж кубик не “симетричний”, то модель розподілу ймовірностей потрібно шукати серед (2), коли є p_i відмінні від $\frac{1}{6}$.

Питання про те, які значення ймовірності приписувати тим чи іншим подіям в реальних експериментах, розв'язуються **методами математичної статистики**. Наведене у попередній лекції означення відносної частоти може використовуватись для наближеної оцінки ймовірності події, точність якої тим більша, чим більше число експериментів проведено. З іншого боку, знання ймовірності події A , що нас цікавить, дозволяє передбачити з певною точністю відносну частоту даної події при проведенні достатньо великого числа реальних експериментів. Подібна задача наукового прогнозування є особливо актуальною в тих випадках, коли оцінка ймовірності за частотою пов'язана з проведенням дорогоцінних або важких експериментів. *Задачі, які розв'язуються в теорії ймовірностей, саме і полягають у тому, щоб за заданими ймовірностями деяких простих подій, відомими з практики, визначати ймовірності більш складних подій, які нас цікавлять.*

У багатьох задачах ймовірнісний простір будується на основі проведення аналогій між даним експериментом і деякою вже вивченою моделлю випадкового експерименту з відомим розподілом ймовірностей. До таких експериментів відносяться експерименти, що зводяться до так званої **класичної схеми**. **Класичний метод співставлення випадковій події A певної ймовірності P_A** задовольняє наведеним вище аксіомам і полягає в наступному:

1) будується повна група G рівноможливих попарно несумісних подій E_1, E_2, \dots, E_n , кожній з яких ставиться у відповідність

ймовірність $\frac{1}{n}$;

2) знаходиться число m таких подій $E_{j_1}, E_{j_2}, \dots, E_{j_m}$ цієї групи, які в сумі дають дану подію A , $A = E_{j_1} + E_{j_2} + \dots + E_{j_m}$;

3) ймовірність події A визначається відношенням

$$P(A) = \frac{m}{n}. \quad (3)$$

Кожний експеримент, для якого множина Ω є скінченною множиною рівноможливих елементарних подій

$\left(P(\omega_1) = P(\omega_2) = \dots = P(\omega_n) = \frac{1}{n} \right)$, називається **класичною схемою**.

Класична схема є математичною формалізацією експериментів, в яких елементарні події мають певну симетрію по відношенню до умов проведення експерименту, так що немає ніяких підстав вважати одну елементарну подію більш ймовірною, ніж іншу. Таку властивість, наприклад, мають експерименти з діставання навмання певної кількості кульок з урні, в якій лежить задана кількість кульок, що не розрізняються на дотик. Тому класичну схему називають ще **схемою урн**. Означення ймовірності за формулою (3) отримало назву класичного означення ймовірності. Словесне **означення класичної ймовірності**: ймовірність дорівнює відношенню числа m елементарних подій, які сприяють даній події A , до загального числа n всіх рівноможливих елементарних подій. У сучасному викладі теорії ймовірностей формулу (3) слід розглядати не як означення ймовірності, а як спосіб її знаходження для подій, які зводяться до схеми урн. Цей спосіб отримав назву **безпосереднього підрахунку ймовірностей**.

Лекція № 3

Комбінаторний метод знаходження ймовірності в класичній схемі

Для підрахунку числа елементарних подій, з яких складаються події в класичній схемі, досить часто використовуються відомі формули **комбінаторики**.

Сформулюємо і приймемо без доведення **основний принцип комбінаторики** (правило добутку): нехай потрібно виконати одну за одною k деяких дій. Якщо першу дію можна виконати n_1 способами, після цього другу дію – n_2 способами, після першої і другої дій третю дію – n_3 способами і так до k -ї дії, яку можна виконати n_k способами, то всі k дій разом можна виконати $n_1 \cdot n_2 \cdot n_3 \cdot \dots \cdot n_k$ способами.

Приклад 1. Скільки чотирьохзначних чисел можна скласти з цифр 0, 1, 2, 3, 4, 5, якщо жодна з цифр не повинна повторюватись більше одного разу?

Розв'язання. Першою цифрою може бути одна з цифр 1, 2, 3, 4, 5, тобто її можна вибрати 5-ма способами (першу дію можна виконати 5-ма способами). Якщо першу цифру вибрано, то другу цифру можна вибрати також 5-ма способами. Отже, першу та другу цифри можна вибрати $5 \cdot 5 = 25$ способами. Якщо перші дві цифри вибрані, то для вибору третьої цифри є 4 способи. Після вибору перших трьох цифр четверту цифру можна вибрати 3 способами. Згідно з правилом добутку загальне число способів вибору числа дорівнює $5 \cdot 5 \cdot 4 \cdot 3 = 300$, тобто можна скласти 300 чисел.

Приклад 2. Скільки можна скласти шестизначних телефонних номерів?

Розв'язання. Для вибору кожної цифри є 10 можливостей, оскільки цифри можуть повторюватись. Тому шестизначний телефонний номер можна вибрати $10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 10^6$ способами, тобто можна скласти 10^6 номерів.

Розглянемо деяку множину E , що містить n елементів: $E = x_1, x_2, \dots, x_n$. Довільний набір k елементів цієї множини, тобто $x_{i_1}, x_{i_2}, \dots, x_{i_k}$, де $x_{i_m} \in E$, $m = 1, 2, \dots, k$ домовимось називати **рядком** довжиною k . Рядки довжиною k можуть відрізнятись один від одного як складом елементів, так і порядком розташування елементів у рядку. Не виключається повторення елементів рядка. Рядки довжиною k назовемо **розміщеннями з n елементів по k** , причому будемо розрізняти розміщення з повтореннями і без повторень елементів.

Розміщення без повторень. Будемо розглядати всі можливі рядки довжиною k , тобто $x_{i_1}, x_{i_2}, \dots, x_{i_k}$, які задовольняють умові: всі елементи $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ різні. Іншими словами, маємо справу з розміщеннями з n елементів по k (коротко: з n по k) без повторень (як правило, слова “без повторень” опускають і кажуть просто про розміщення з n по k). Позначимо число таких розміщень через A_n^k . Перший елемент розміщення з n по k можна вибрати n способами. Якщо перший елемент вибрано, то для вибору другого елемента є $n - 1$ можливість, третього – $n - 2$ можливості і так далі. Вибір k -го елемента можна здійснити $n - \underline{\underline{k-1}}$ способами. Згідно з правилом

добутку кількість розміщень (без повторень) з n елементів по k дорівнює

$$A_n^k = n(n-1)(n-2)\dots(n-(k-1)).$$

Позначимо через $n!$ добуток n перших натуральних чисел, тобто

$$n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot \dots \cdot (n-1) \cdot n$$

і назовемо цей добуток терміном “ n -факторіал”. За означенням приймемо, що $0!=1$. Тоді число A_n^k можна записати у вигляді

$$A_n^k = \frac{n!}{(n-k)!}. \quad (1)$$

Важливим є випадок, коли $n = k$, тобто коли в рядку містяться всі елементи множини E (кожен по одному разу). Такі рядки називають *перестановками* з n елементів. Число P_n всіх перестановок дорівнює

$$P_n = n!. \quad (2)$$

Приклад 3. Скільки існує способів розсадити 5 чоловік на 5 стільців?

Розв'язання. $P_5 = 5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$ (способів).

Про розміщення без повторень можна говорити як про *упорядковані підмножини* множини E , наприклад $\langle x_1, x_2, x_3 \rangle$ і $\langle x_2, x_1, x_3 \rangle$ – дві різні підмножини, оскільки порядок елементів в них різний.

Розміщення з повтореннями. Задача про відшукання числа рядків довжиною k , елементи яких можуть повторюватись, допускає таку інтерпретацію. Нехай всі n елементів множини E занумеровані і знаходяться в урні. Будемо виймати ці елементи один за одним, записувати їх номери і повертати назад в урну. Після того, як ми виймемо k елементів, отримаємо рядок довжиною k , причому елементи в рядку можуть повторюватись. На основі тих же міркувань, що й у попередньому випадку, число $A_n^{k(\text{повт.})}$ всіх можливих розміщень з повтореннями дорівнює

$$A_n^{k(\text{повт.})} = n^k.$$

Приклад 4. Відповідь на запитання з прикладу 2 можна записати у вигляді $A_{10}^{6(\text{повт.})} = 10^6$.

Рекомендується надати відповідну інтерпретацію з урною розміщенням без повторень.

Сполучення. Якщо порядок елементів у рядку довжиною k не є важливим і рядки з однаковим набором елементів, але різним

порядком, вважаються однаковими, то говорять про **сполучення** k елементів із n (сполучення з n по k). Сполучення також можуть бути без повторень і з повтореннями.

Сполучення без повторень. Під сполученням (як правило, слова “без повторень” опускають і кажуть просто “сполучення”) k елементів із n елементів множини E розуміють довільну k -елементну підмножину множини E (кажуть: сполучення з n по k). Нехай число C_n^k всіх k -елементних підмножин n -елементної множини E відоме. Кожну з цих підмножин можна упорядкувати $k!$ (k -факторіал) способами. Отримаємо $C_n^k k!$ всіх упорядкованих k -елементних підмножин множини E . З іншого боку, число таких підмножин дорівнює A_n^k . Отже,

$$C_n^k k! = A_n^k \text{ або } C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}. \quad (3)$$

Формула (3) є однією з основних формул комбінаторики.

Приклад 5. Скільки існує способів розсадити 5 чоловіків і 5 жінок на 12 стільців?

Розв’язання. Стільці для жінок можна вибрати C_{12}^5 способами. Після цього залишаються 7 стільців, на які можна C_7^5 способами розсадити чоловіків. Всього за правилом добутку отримуємо

$$C_{12}^5 \cdot C_7^5 = \frac{12!}{5! 7!} \cdot \frac{7!}{5! 2!} = \frac{12!}{2 \cdot 5! \cdot 5!} \text{ (способів)}.$$

Числа C_n^k мають ряд цікавих властивостей. Наведемо їх.

$$1) \quad C_n^k = C_n^{n-k}; \quad 2) \quad C_n^k = C_{n-1}^k + C_{n-1}^{k-1}, \quad 1 \leq k \leq n; \quad 3) \quad C_n^0 + C_n^1 + \dots + C_n^n = 2^n.$$

Властивість 1) є очевидною і безпосередньо випливає з формули (3).

$$\begin{aligned} \text{Доведемо властивість 2): } & C_{n-1}^k + C_{n-1}^{k-1} = \frac{n-1!}{k! n-1-k!} + \frac{n-1!}{(k-1)! n-k!} = \\ & = \frac{n-1! n-k}{k! n-k!} + \frac{n-1! k}{k! n-k!} = \frac{n-1! n-k+k}{k! n-k!} = \frac{n!}{k! n-k!} = C_n^k. \end{aligned}$$

Ліва частина в 3) дорівнює числу всіх підмножин множини E (\emptyset , одноелементних підмножин, двохелементних, трьохелементних і так далі). У додатку 1 доводиться, що це число дорівнює 2^n .

Сполучення з повтореннями. Розглянемо рядки довжиною k , елементи в яких можуть повторюватись, зокрема k може бути більшим за n . Як і в попередньому випадку, розрізняємо рядки тільки за складом елементів, наприклад, рядки $\langle x_1, x_2, x_2 \rangle$ і $\langle x_2, x_1, x_2 \rangle$ вважаємо однаковими. Кожному складу рядка довжиною k можна поставити у відповідність один і тільки один числовий рядок $\langle N_1, N_2, N_3, \dots, N_n \rangle$, де число N_i показує, що елемент x_i множини E зустрічається в рядку N_i разів. Зрозуміло, що

$$N_1 + N_2 + N_3 + \dots + N_n = k.$$

Приклад 6. $E = \langle x_1, x_2, x_3, x_4 \rangle$. Для рядка $\langle x_1, x_2, x_2, x_3, x_3 \rangle$ відповідний числовий рядок має вигляд $\langle 2, 2, 0, 2 \rangle$.

Різні склади рядка довжиною k , компоненти якого належать до n -елементної множини E , називаються **сполученнями з повтореннями** з n елементів по k . Позначимо їх число $C_n^{k(\text{повт.})}$. Отже, $C_n^{k(\text{повт.})}$ – це число всіх можливих числових рядків $\langle N_1, N_2, N_3, \dots, N_n \rangle$.

Запишемо кожен з таких рядків у вигляді рядка з одиниць і нулів

$$\left(\underbrace{11\dots1}_{N_1} \underbrace{011\dots1}_{N_2} \underbrace{011\dots1}_{N_3} \dots \underbrace{11\dots1}_{N_{n-1}} \underbrace{011\dots1}_{N_n} \right), \quad (4)$$

де кожне число замінюється на відповідне число одиниць і ставиться нуль післяожної групи одиниць, крім останньої. Якщо $N_i = 0$, то набір одиниць для даного числа відсутній. Наприклад, для рядка $\langle 0, 3, 0, 1 \rangle$ можна записати $\langle 100111001 \rangle$, а для рядка $\langle 0, 0, 3, 1 \rangle$ – $\langle 100011101 \rangle$. Очевидно, що число одиниць, які містяться в рядку (4), дорівнює $N_1 + N_2 + \dots + N_n = k$, а число нулів дорівнює $n - 1$. Знайдемо число можливих рядків вигляду (4). Оскільки k одиниць можна розмістити C_{k+n-1}^k способами, а $n - 1$ нулів на решту $n - 1$ місць можна розмістити $C_{n-1}^{n-1} = 1$ способом, то число можливих рядків (4) дорівнює C_{k+n-1}^k . З іншого боку, це число дорівнює $C_n^{k \text{ повт.}}$. Отже, для числа сполучень з повтореннями з n елементів по k маємо:

$$C_n^{k(\text{повт.})} = C_{k+n-1}^k = \frac{(n+k-1)!}{k!(n-1)!}. \quad (5)$$

Приклад 7. Скількома способами можна скласти набір із 3 троянд двох сортів?

Розв'язання. $C_2^{3(\text{новт})} = C_4^3 = 4$ (способами).

Приклад 8. Скількома способами можна скласти набір із 2 троянд трьох сортів?

Розв'язання. $C_3^{2(\text{новт})} = C_4^2 = 6$ (способами).

З викладеного вище випливає, що всі основні положення комбінаторики є простими наслідками **основного принципу комбінаторики** (правила добутку).

При розв'язанні ймовірнісних задач важливо виділити експерименти, де можна використовувати ті чи інші комбінаторні формули. Кожна з цих формул визначає загальне число елементарних подій у деякій схемі урн, тобто в іdealізованому експерименті з вибору навмання k елементів із n елементів вихідної множини $E = \{x_1, x_2, \dots, x_n\}$. При цьому в постановці кожного такого експерименту повинно бути чітко обумовлено, яким способом проводиться вибір і що розуміється під різними вибірками. Існують **два принципово різні схеми вибору**: у *першій схемі* вибір здійснюється без повернення елементів (це означає, що відбираються або зразу всі k елементів, або послідовно по одному елементу, причому кожного разу відібраний елемент вилучається з вихідної множини E). У *другій схемі* вибір здійснюється **поелементно** з обов'язковим поверненням відібраного елемента на кожному кроці і ретельним “перемішуванням” вихідної множини перед наступним вибором. Після того, як вибір тим чи іншим способом здійснено, відіbrane елементи можуть бути або упорядковані (тобто розміщені в послідовний ланцюжок), або ні. В результаті маємо такі чотири різні постановки експерименту з вибору навмання k елементів із загального числа n елементів множини E .

A. Схема вибору, яка призводить до сполучень

Якщо експеримент полягає у виборі k елементів без повернення і без упорядкування, то різними результатами слід вважати k -елементні підмножини множини E , які мають різний склад. Комбінації елементів, які при цьому отримуються (елементарні події), є сполученнями з n елементів по k , а їх загальне число $N(\Omega)$ підраховується за формулою

$$N(\Omega) = C_n^k = \frac{n!}{k!(n-k)!}.$$

Б. Схема вибору, яка призводить до розміщень

Якщо експеримент полягає у виборі k елементів без повернення, але з упорядкуванням цих елементів в процесі вибору в послідовний ланцюжок, то різними результатами даного експерименту будуть упорядковані k -елементні підмножини множини E , які розрізняються або набором елементів, або порядком їх розміщення в ланцюжку. Комбінації елементів, які отримуються в даному випадку, тобто елементарні події, є розміщеннями з n елементів по k , а їх загальне число визначається формулою

$$N(\Omega) = A_n^k = \frac{n!}{(n-k)!}.$$

У випадку, коли $k = n$, експеримент полягає в довільному упорядкуванні множини E , тобто зводиться до випадкової перестановки елементів всієї множини. При цьому $N(\Omega) = n!$.

В. Схема вибору, яка призводить до сполучень з повтореннями

Якщо експеримент полягає в поелементному виборі з поверненням k елементів множини $E = \{x_1, x_2, \dots, x_n\}$, але без наступного упорядкування, то різними результатами такого експерименту будуть довільні k -елементні набори, які відрізняються складом елементів. При цьому окремі набори можуть містити елементи, що повторюються. Наприклад, при $k = 4$ набори $\omega_1 = x_1, x_1, x_2, x_1$ і $\omega_2 = x_2, x_1, x_1, x_1$ не розрізняються для даного експерименту, а набір $\omega_3 = x_1, x_1, x_3, x_1$ відрізняється від кожного з попередніх. Комбінації елементів, які отримуються в даному експерименті, є сполученнями з повтореннями, а їх загальне число визначається формулою

$$N(\Omega) = C_{n+k-1}^k = \frac{n+k-1!}{k! (n-1)!}.$$

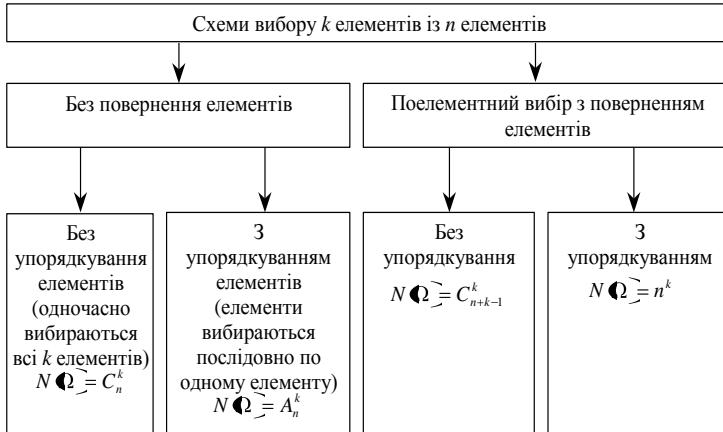
Г. Схема вибору, яка призводить до розміщень з повтореннями

Якщо поелементний вибір k елементів із множини $E = \{x_1, x_2, \dots, x_n\}$ здійснюється з поверненням і з упорядкуванням їх в послідовний ланцюжок, то різними результатами будуть довільні k -елементні набори (взагалі кажучи, з повтореннями), які відрізняються або складом, або порядком розміщення елементів. Наприклад, при $k = 4$ множини $\omega_1 = x_1, x_1, x_2, x_1$, $\omega_2 = x_2, x_1, x_1, x_1$, і $\omega_3 = x_1, x_1, x_3, x_1$ є різними елементарними подіями для даного

експерименту. Комбінації, які отримуються в цьому експерименті, є розміщеннями з повтореннями, а їх загальне число знаходиться за формuloю

$$N \Omega = n^k.$$

Зведемо основні висновки даної лекції до схеми:



Приклад 9. З колоди в 36 карт навмання вибираються 4 карти. Знайти ймовірності таких подій: $A = \{\text{серед вибраних карт всі карти бубнової масти}\}$, $B = \{\text{серед вибраних карт буде хоча б один туз}\}$.

Розв'язання. Ω – всі можливі комбінації з 36 карт по чотири (сполучення з 36 по 4); $N \Omega = C_{36}^4$, $N A = C_9^4$;

$$P(A) = \frac{N(A)}{N(\Omega)} = \frac{C_9^4}{C_{36}^4} = \frac{9!}{4!28!} \cdot \frac{4!32!}{36!} = \frac{6 \cdot 7 \cdot 8 \cdot 9}{33 \cdot 34 \cdot 35 \cdot 36} = \frac{2}{935} \approx 0,0021.$$

$$\text{Нехай } \bar{B} = \{\text{не буде жодного туза}\}; N(\bar{B}) = C_{36-4}^4 = C_{32}^4;$$

$$P(\bar{B}) = \frac{C_{32}^4}{C_{36}^4} = \frac{32!}{4!28!} \cdot \frac{4!32!}{36!} = \frac{29 \cdot 30 \cdot 31 \cdot 32}{33 \cdot 34 \cdot 35 \cdot 36} \approx 0,61;$$

$$P(B) = 1 - P(\bar{B}) \approx 0,39.$$

Приклад 10. Група з 8 чоловік займає всі 8 місць: а) за круглим столом, б) з одного боку прямокутного столу. Знайти ймовірність того, що дві певні особи сидітимуть поряд.

Розв'язання. Ω – перестановки з 8 елементів; $N \Omega = 8!$

а) Одна з двох даних осіб може зайняти місце за круглим столом одним із 8 способів; після цього друга особа може зайняти місце поруч

одним із 2 способів. Інші 6 чоловік можуть зайняти 6 вільних місць $6!$ способами. Отже, згідно з основним принципом комбінаторики

$$N(A) = 8 \cdot 2 \cdot 6!; \quad P(A) = \frac{8 \cdot 2 \cdot 6!}{8!} = \frac{8 \cdot 2}{7 \cdot 8} = \frac{2}{7}.$$

б) Рекомендується показати **самостійно**, що у цьому випадку

$$P(A) = \frac{1}{4}.$$

Приклад 11. У кондитерській є 7 видів тістечок. Покупець вибив чек на 4 тістечка. Вважаємо, що всі набори тістечок, які замовляються, є рівноможливими. Знайти ймовірності таких подій: $A = \{\text{тістечка одного виду}\}$, $B = \{\text{тістечка різних видів}\}$.

Розв'язання. Ω – сполучення з повтореннями з 7 по 4;

$$N(\Omega) = C_{7+4-1}^4 = C_{10}^4;$$

$$N(A) = C_7^1 = 7; \quad P(A) = \frac{7}{C_{10}^4} = \frac{7 \cdot 4! \cdot 6!}{10!} = \frac{1}{30}.$$

$$N(B) = \frac{C_7^4}{C_{10}^4}.$$

Приклад 12. Шість пасажирів піднімаються на ліфті семиповерхового будинку. Вважаємо, що рух ліфту розпочинається з нульового поверху. Знайти ймовірності таких подій:

$A = \{\text{на перших трьох поверхах не вийде жоден пасажир}\}$,

$B = \{\text{всі пасажири вийдуть на одному поверсі}\}$.

Розв'язання. $\Omega = i_1, i_2, i_3, i_4, i_5, i_6$, $1 \leq i_k \leq 7$;

i_k – номер поверху, на якому виходить k -пасажир.

$$N(\Omega) = 7^6; \quad A = i_1, i_2, i_3, i_4, i_5, i_6, \quad 4 \leq i_k \leq 7; \quad N(A) = 4^6;$$

$$P(A) = \frac{4^6}{7^6} = \left(\frac{4}{7}\right)^6.$$

$$B = 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, \dots, 7, 7, 7, 7, 7, 7; \quad ;$$

$$N(B) = C_7^1 = 7; \quad P(B) = \frac{7}{7^6} = \frac{1}{7^5}.$$

Лекція № 4

Дискретний ймовірнісний простір. Геометричні ймовірності.

Неперервний ймовірнісний простір. Умовні ймовірності та незалежність подій

Дискретний ймовірнісний простір. Розглянемо випадковий експеримент з дискретною множиною елементарних подій $\Omega = \omega_1, \omega_2, \dots, \omega_k, \dots$. Як ми вже знаємо, у цьому випадку довільна підмножина $A \subset \Omega$ інтерпретується як подія, що спостерігається в даному експерименті. Ймовірнісну модель експерименту Ω, F, P , де F – система всіх підмножин множини Ω , можна подати у спрощеному вигляді Ω, P .

Припустимо, що кожній елементарній події ω_k поставлено у відповідність число $P(\omega_k) = p_k$, причому

$$p_k \geq 0, k = 1, 2, \dots; \quad \sum_{k=1}^{\infty} p_k = 1. \quad (1)$$

Числа p_k , $k = 1, 2, \dots$ назовемо ймовірностями елементарних подій. Ймовірність довільної події $A \subset \Omega$ задамо як суму ймовірностей тих елементарних подій, з яких складається A :

$$P(A) = \sum_{\substack{k \\ \omega_k \in A}} p_k. \quad (2)$$

Неважко перевірити, що при такому розподілі ймовірностей всі аксіоми з лекції 2 виконуються. Збіжність ряду (2) гарантована збіжністю ряду (1).

Якщо множина Ω – скінчenna, $\Omega = \omega_1, \omega_2, \dots, \omega_k$, отримуємо модель випадкового експерименту зі скінченним числом *нерівноможливих* елементарних подій (див. приклад 6 з лекції 2). Зокрема, якщо $p_1 = p_2 = \dots = p_n = 1/n$, прийдемо до класичної схеми.

Питання вибору конкретних значень p_k в тій чи іншій моделі виходить за межі теорії ймовірностей.

Геометричні ймовірності. Припустимо, що елементарні події випадкового експерименту є рівноможливими і утворюють нескінченну неперервну сукупність, яку можна зобразити точками деякої квадрованої області Ω на площині, а довільну випадкову подію A у цьому експерименті – точками квадрованої підмножини $A \subset \Omega$. Позначимо через $S(\Omega)$, $S(A)$ відповідні площини і введемо розподіл

ймовірностей за формулою

$$P(A) = \frac{S(A)}{S(\Omega)}. \quad (3)$$

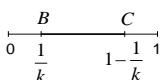
Легко перевірити, що аксіоми теорії ймовірностей для такого розподілу задовільняються.

Аналогічні розподіли ймовірностей можна ввести для випадків, коли рівноможливі результати експерименту зображаються неперервними областями Ω на прямій чи в просторі, а події A – відповідними підмножинами $A \subset \Omega$ з довжиною $l(A)$ чи об'ємом $V(A)$:

$$P(A) = \frac{l(A)}{l(\Omega)}; \quad P(A) = \frac{V(A)}{V(\Omega)}. \quad (4)$$

Формули (3), (4) носять називу **формул геометричної ймовірності**. Всім задачам на геометричну ймовірність можна дати таке тлумачення: в області Ω навмання вибирається точка. Яка ймовірність, що ця точка належить A ?

Приклад 1. На відрізку одиничної довжини навмання вибирається точка. Знайти ймовірність того, що відстань від цієї точки до кожного з кінців відрізка буде більшою, ніж $\frac{1}{k}$, де $k > 2$.



Розв'язання. Множина елементарних подій – це точки відрізка $[0, 1]$. Нехай подія A – відстань від точки до кожного з кінців відрізка більша, ніж $\frac{1}{k}$.

Цій події відповідають точки інтервалу $(\frac{1}{k}, 1 - \frac{1}{k})$. Оскільки $l(\Omega) = 1$,

$$l(A) = 1 - \frac{2}{k} = \frac{k-2}{k}, \text{ то}$$

$$P(A) = \frac{l(A)}{l(\Omega)} = \frac{k-2}{k}.$$

Зauważення 1. Формули геометричних ймовірностей (3), (4) введені нами в межах алгебри подій F (квадрованих підмножин, підмножин, що мають довжину чи об'єм). У багатьох практичних задачах цього буває досить. Відповідними σ -алгебрами є системи борелівських множин на площині, прямій та в просторі. Ми не будемо обчислювати ймовірності нових (у порівнянні з алгеброю F) подій із σ -алгебри. Розглянемо лише подію $A = c \subset \Omega$, що містить окрему точку та подію $Q = r_1, r_2, \dots, r_n, \dots$,

що складається з раціональних точок, $Q \subset \Omega$ (див. приклад 5 лекції 2) (нагадаємо, що c та Q не можуть інтерпретуватись як події, що спостерігаються, якщо залишатись в межах алгебри подій). Оскільки $\emptyset \subset c \subset c - 1/n, c + 1/n$, то згідно з властивістю ймовірності 6) з лекції 2 $0 \leq P(c) \leq \frac{2}{n \cdot l(\Omega)}$. Звідси при $n \rightarrow \infty$ знаходимо, що

$$P(c) = 0. \text{ Далі, } Q = \bigcup_{n=1}^{\infty} r_n. \text{ Тому згідно з розширеною аксіомою додавання } P(Q) = \sum_{n=1}^{\infty} P(r_n) = 0.$$

Отже, якщо в дискретному ймовірнісному просторі нульову ймовірність мають лише неможливі події (з того, що $P(A) = 0$, випливає $A = \emptyset$), то в загальному випадку нульову ймовірність мають не тільки неможливі події, але і деякі з можливих подій, наприклад попадання голки в окрему точку чи в множину раціональних точок.

Неперервний ймовірнісний простір. Позначимо через Ω множину точок всієї площини, а через $f(x; y)$ – деяку невід’ємну функцію, для якої існує інтеграл Рімана $\iint_A f(x, y) dx dy$ по довільній квадрованій області A . Допустимо, що існує невласний інтеграл від функції $f(x; y)$ на всій площині і $\iint_{\Omega} f(x, y) dx dy = 1$. Нехай F – алгебра квадрованих множин із Ω . Задамо розподіл ймовірностей для кожної множини $A \in F$:

$$P(A) = \iint_A f(x, y) dx dy. \quad (3')$$

Рекомендується переконатись **самостійно**, що аксіоми теорії ймовірностей для такого розподілу виконуються. Простір $(\Omega, F, P(A))$ назовемо неперервним ймовірнісним простором. Зокрема, якщо Ω – квадрована фігура скінчених розмірів, то покладаючи $f(x, y) = \frac{1}{S(\Omega)}$,

коли $(x, y) \in \Omega$ і $f(x, y) = 0$, коли $(x, y) \notin \Omega$, прийдемо до формули геометричної ймовірності (3). Поняття неперервного ймовірнісного простору можна розглядати як узагальнення поняття ймовірнісного простору з геометричною ймовірністю на випадок “нерівноможливих” елементарних подій.

Аналогічно будуються неперервні ймовірнісні простори на прямій та в просторі довільної розмірності.

Умовні ймовірності та незалежність подій. В лекції 2 введено поняття ймовірності як числової функції множин, що задовольняє трьом аксіомам. Таку ймовірність називають безумовною ймовірністю, підкреслюючи цим, що вона не залежить ні від яких інших додаткових умов, крім фіксованого комплексу умов S , яким характеризується експеримент.

Досить часто доводиться розглядати ймовірність випадкової події A за додаткової умови, що відбулася деяка випадкова подія B . Таку ймовірність будемо називати **умовною ймовірністю** і позначати символом $P(A|B)$ (ймовірність A за умови B).

Розглянемо класичну схему. Нехай можливі результати експерименту зводяться до n випадків. Припустимо, що події A сприяють m випадків, а події B – k випадків. Оскільки ніяких припущення щодо сумісності чи несумісності подій A та B не робиться, то, взагалі кажучи, існують випадки, які сприяють і події A , і події B одночасно. Позначимо число таких випадків через l . Тоді

$$P(A) = \frac{m}{n}; P(B) = \frac{k}{n}; P(AB) = \frac{l}{n}. \text{ Знайдемо } P(A|B), \text{ тобто умовну}$$

ймовірність події A за умови, що подія B відбулася. Якщо відомо, що подія B відбулася, то з раніше можливих n випадків залишаються можливими тільки ті k , які сприяють події B . При цьому подія A може відбутися лише за рахунок перерізу AB , якому сприяють l

елементарних подій. Отже, $P(A|B) = \frac{l}{k}$. Оскільки $\frac{l}{k} = \frac{\frac{l}{n}}{\frac{k}{n}}$, то

$$P(A|B) = \frac{P(AB)}{P(B)}. \quad (5)$$

Зрозуміло, що наші міркування мають смисл лише тоді, коли $P(B) \neq 0$.

Рівність (5) дозволяє дати загальне означення умовної ймовірності.

Означення 1. Нехай (Ω, F, P) – довільний ймовірнісний простір.

Якщо $A, B \in F$ і $P(B) \neq 0$, то умовою ймовірністю події A за умови, що відбулася подія B , називається число, яке обчислюється за формулою (5).

Легко перевірити, що для умовних ймовірностей виконуються всі аксіоми безумовних ймовірностей:

$$A.1. \quad P(A/B) = \frac{P(AB)}{P(B)} \geq 0, \quad P(B) \neq 0;$$

$$A.2. \quad P(\Omega/B) = \frac{P(\Omega B)}{P(B)} = \frac{P(B)}{P(B)} = 1;$$

A.3. Нехай A_1, A_2 – несумісні $A_1 A_2 = \emptyset$ події. Тоді

$$\begin{aligned} P[(A_1 + A_2)/B] &= \frac{P[(A_1 + A_2)B]}{P(B)} = \\ &= \frac{P(A_1 B) + P(A_2 B)}{P(B)} = P(A_1/B) + P(A_2/B). \end{aligned}$$

Якщо A_1, A_2 – сумісні події $A_1 A_2 \neq \emptyset$, то має місце така формула суми для умовних ймовірностей:

$$P[A_1 + A_2/B] = P[A_1/B] + P[A_2/B] - P[A_1 A_2/B], \quad P[B] \neq 0.$$

Доведемо ІІ:

$$\begin{aligned} P[A_1 + A_2/B] &= \frac{P[A_1 + A_2 B]}{P[B]} = \frac{P[A_1 B + A_2 B]}{P[B]} = \\ &= \frac{P(A_1 B) + P(A_2 B) - P(A_1 A_2 B)}{P[B]} = P(A_1/B) + P(A_2/B) - P(A_1 A_2/B). \end{aligned}$$

Рекомендується показати **самостійно**, що

$$P(B/A) = 1, \quad P[AB/C] = P[A/C]/P[B/AC].$$

Очевидно, що у випадку $P[A] \neq 0$ події A і B в (5) можна помінити місцями. З формули (5) безпосередньо випливає так звана “теорема множення”: якщо $P[A] \neq 0$, $P[B] \neq 0$, то

$$P[AB] = P[A]P[B/A] = P[B]P[A/B]. \quad (6)$$

“Теорему множення” можна узагальнити на **випадок довільної скінченної кількості подій**, а саме:

$$P[A_1 A_2 A_3 \dots A_n] = P[A_1] P[A_2/A_1] P[A_3/A_1 A_2] \dots P[A_n/A_1 A_2 \dots A_{n-1}]. \quad (7)$$

Доведення (за індукцією) рекомендується провести **самостійно** (див. додаток 1).

Зauważення 2. Формули (6), (7) дозволяють знаходити ймовірності одночасної появи подій AB та $A_1 A_2 \dots A_n$, якщо **умовні** ймовірності відомі.

Для знаходження останніх досить часто застосовують так званий **метод допоміжного експерименту**, суть якого полягає в наступному: нехай у

деякому експерименті, в якому нас цікавить подія B , стало відомо, що відбулася подія A . Така ситуація призводить до необхідності зміни початкової множини елементарних подій і перегляду ймовірностей, які спочатку приписувались цим подіям. Іншими словами, для знаходження ймовірності $P(B/A)$ ми формуємо новий (“допоміжний”) експеримент, який є деякою модифікацією початкового експерименту. **Безумовна** ймовірність події B у цьому новому експерименті і приймається за **умовну** ймовірність $P(B/A)$. Методом допоміжного експерименту ми скористалися вище, отримуючи формулу (5) в класичній схемі.

Приклад 2. З урни, в якій містяться 4 білих і 6 чорних кульок, навмання послідовно і без повернення виймаються дві кульки. Нехай A – перша кулька біла, B – друга кулька біла. Знайдемо $P(B/A)$.

Допоміжний експеримент: з урни, в якій 3 білих і 6 чорних кульок (згідно з умовою A одну білу кульку вже вийнято) навмання виймається одна кулька. Ймовірність того, що вона біла дорівнює

$$\frac{3}{3+6} = \frac{1}{3}. \text{ Отже, } P(B/A) = \frac{1}{3}. \text{ Знайдемо тепер } P(B/A) \text{ безпосередньо}$$

за формулою $P(B/A) = \frac{P(BA)}{P(A)}$. Маємо $N(B/A) = A_{10}^2 = \frac{10!}{8!} = 90$;

$$N(A) = 4 \cdot 9 = 36; P(A) = \frac{36}{90} = \frac{2}{5}; N(AB) = 4 \cdot 3 = 12;$$

$$P(BA) = P(AB) = \frac{12}{90} = \frac{2}{15}.$$

Отже, $P(B/A) = \frac{1}{3}$. Отримали ту ж саму ймовірність. Відмітимо, що

безумовна ймовірність події B дорівнює $\frac{2}{5}$. Дійсно,

$$P(B) = P[(A + \bar{A})B] = P(AB) + P(\bar{A}B) = \frac{12}{90} + \frac{24}{90} = \frac{36}{90} = \frac{2}{5}, \text{ оскільки}$$

$$N(\bar{A}B) = 6 \cdot 4 = 24.$$

Означення 2. Випадкові події A і B називаються **незалежними**, якщо

$$P(AB) = P(A)P(B). \quad (8)$$

Порівнявши формулу (6) з означенням незалежності (8), для незалежних подій A та B $P(A) \neq 0, P(B) \neq 0$ отримуємо

$$P(B/A) = P(B), P(A/B) = P(A). \quad (9)$$

Отже, настання однієї з двох незалежних подій не змінює ймовірності іншої. Очевидно, що кожна з рівностей (9) є гарантією незалежності подій A та B .

Зауваження 3. Формула (8) дає можливість виділити незалежні події, якщо модель сформовано, тобто ймовірнісний простір побудовано. Однак на практиці далеко не завжди можна скористатись цією формулою як критерієм незалежності подій. На практиці частіше застосовується гіпотеза про “фізичну” незалежність подій: вважаються незалежними події, які не зв’язані причинно. Наприклад, постріли з декількох гармат в одну і ту ж мішень – незалежні.

Якщо математичну модель, що описує деякий експеримент, побудовано правильно, то “фізично” незалежним подіям реального експерименту відповідають події моделі, які незалежні і в розумінні формулі (8).

Приклад 3. Експеримент полягає в підкиданні двох симетричних монет. Тоді $\Omega = \{\text{ГГ}, \text{ЦЦ}, \text{ЦГ}, \text{ГЦ}\}$. Нехай $A = \{\text{ГГ}, \text{ГЦ}\}$ – перша монета впаде “гербом” доверху; $B = \{\text{ЦГ}, \text{ГГ}\}$ – друга монета впаде “гербом” доверху; $AB = \{\text{ГГ}\}$ – обидві монети впадуть “гербом” доверху. Тоді

$$P(A) = P(B) = \frac{1}{2}, \quad P(AB) = \frac{1}{4}. \quad \text{Отже, } P(AB) = P(A)P(B). \quad \text{Події } A \text{ та } B \text{ незалежні як “фізично”, так і у моделі.}$$

Теорема. Якщо події A та B незалежні, то незалежні і події A та \bar{B} .

$$\begin{aligned} \text{Доведення.} \quad & P(A) = P(A\bar{B} + \bar{A}) = P(AB + A\bar{B}) = \\ & = P(AB) + P(A\bar{B}), \text{ то } P(A\bar{B}) = P(A) - P(AB) = P(A) - P(A)P(B) = \\ & = P(A)(1 - P(B)) = P(A)P(\bar{B}). \end{aligned}$$

З теореми маємо такий наслідок.

Наслідок. Якщо події A та B незалежні, то незалежні і події \bar{A} та B , A та \bar{B} .

Поняття незалежності можна поширити на випадок довільної кількості подій.

Означення 3. Події A_1, A_2, \dots, A_n називаються **незалежними в сукупності**, якщо для довільного набору індексів $1 \leq i_1 < i_2 < i_3 < \dots < i_k \leq n$ $2 \leq k \leq n$ справедлива рівність

$$P(A_{i_1} A_{i_2} \dots A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \dots P(A_{i_k}). \quad (10)$$

Зокрема, для незалежних в сукупності подій формула (7) приймає вигляд

$$P(A_1 A_2 \dots A_n) = P(A_1) P(A_2) \dots P(A_n). \quad (11)$$

З незалежності подій A_1, A_2, \dots, A_n в сукупності випливає їх попарна незалежність, тобто

$$P(A_{i_1} A_{i_2}) = P(A_{i_1}) P(A_{i_2}), \quad 1 \leq i_1 < i_2 \leq n.$$

Обернене твердження, взагалі кажучи, не вірне, тобто попарно незалежні події можуть не бути незалежними в сукупності.

Стосовно формул (11) можна зробити те ж саме зауваження 3, що й стосовно формул (8).

Приклад 4. Знайти ймовірність того, що при шести підкиданнях монети принаймні один раз випаде “герб”.

Розв'язання. Позначимо нашу подію через A . Тоді протилежна подія \bar{A} – при шести підкиданнях монети всі шість разів випаде “цифра”. Нехай подія A_i – при i -му підкиданні випаде “цифра”. Тоді

$P(A_i) = \frac{1}{2}, \quad i = 1, 2, \dots, 6$. Зрозуміло, що події A_i – незалежні в сукупності. Тому за формулою (11) $P(\bar{A}) = P(A_1 A_2 \dots A_6) =$

$$= P(A_1) P(A_2) \dots P(A_6) = \frac{1}{2^6}. \text{ Отже, } P(A) = 1 - P(\bar{A}) = 1 - \frac{1}{2^6} = \frac{63}{64}.$$

Лекція № 5

Формула повної ймовірності. Формула Байєса. Послідовності незалежних випробувань. Формула Бернуллі

Введемо поняття складної події.

Означення. Складною подією будемо називати подію A , яка виражається через інші події, що спостерігаються, за допомогою алгебраїчних операцій над подіями.

Для знаходження ймовірностей таких подій досить часто використовуються вже відомі нам формулі:

$$P(A+B) = P(A) + P(B) - P(AB);$$

$$P(AB) = P(A)P(B/A) = P(B)P(A/B);$$

$$P(A_1 A_2 \dots A_n) = P(A_1) P(A_2/A_1) \dots P(A_n/A_1 A_2 \dots A_{n-1}).$$

Справедлива формула для суми трьох подій:

$$P(A+B+C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

Рекомендується довести цю формулу **самостійно**.

Ймовірність суми трьох і більше подій зручно знаходити, переходячи до протилежної події:

$$P(A_1 + A_2 + \dots + A_n) = 1 - P(\overline{A_1 + A_2 + \dots + A_n}) = 1 - P(\overline{A_1} \overline{A_2} \dots \overline{A_n}).$$

Якщо події A_1, A_2, \dots, A_n – незалежні в сукупності, то $P(\overline{A_1} \overline{A_2} \dots \overline{A_n}) = P(\overline{A_1})P(\overline{A_2}) \dots P(\overline{A_n})$ і ймовірність настання хоча б однієї з подій A_1, A_2, \dots, A_n дорівнює

$$P(A_1 + A_2 + \dots + A_n) = 1 - P(\overline{A_1})P(\overline{A_2}) \dots P(\overline{A_n}) = 1 - q_1 q_2 \dots q_n,$$

де q_k – ймовірність настання події $\overline{A_k}$, $1 \leq k \leq n$.

Якщо незалежні в сукупності події A_1, A_2, \dots, A_n мають однакову ймовірність, наприклад, p , то ймовірність настання хоча б однієї з них дорівнює $P(A_1 + A_2 + \dots + A_n) = 1 - q^k$, де $q = 1 - p$.

Розглянемо події H_1, H_2, \dots, H_n , які утворюють **повну групу подій** (див. лекцію 1), тобто

$$H_i \neq \emptyset, \quad H_i H_j = \emptyset \quad (i \neq j), \quad H_1 + H_2 + \dots + H_n = \Omega.$$

Нехай потрібно знайти ймовірність деякої події A . Звичайно, ця подія може відбутися разом з однією із подій повної групи H_1, H_2, \dots, H_n . По відношенню до події A події H_1, H_2, \dots, H_n будемо називати **гіпотезами**. Позначимо через $P(H_i)$ ймовірності гіпотез. Ці ймовірності ще часто називають **апріорними**.

Твердження. Ймовірність події A дорівнює сумі добутків ймовірностей кожної з гіпотез на умовні ймовірності події A при цих гіпотезах, тобто

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

Формулу (1) називають **формулою повної ймовірності**.

Доведення. Оскільки події H_1, H_2, \dots, H_n попарно несумісні, то і події AH_1, AH_2, \dots, AH_n також попарно несумісні. Дійсно, $(AH_i)(AH_j) = (AA)(H_i H_j) = A\emptyset = \emptyset$. Подію A можна записати у вигляді $A = A\Omega = A(H_1 + H_2 + \dots + H_n) = AH_1 + AH_2 + \dots + AH_n$, а її

ймовірність – у вигляді

$$P(A) = P(AH_1) + P(AH_2) + \dots + P(AH_n) \geq \sum_{i=1}^n P(H_i A).$$

Оскільки згідно з “теоремою множення” (див. лекцію 4)

$$P(H_i A) = P(H_i) P(A|H_i), \text{ тоді } P(A) = \sum_{i=1}^n P(H_i) P(A|H_i).$$

Формула повної ймовірності (1) дає відповідь на питання про ймовірність події A , якщо відомі апріорні ймовірності гіпотез.

Сформулюємо тепер таку **задачу**. Нехай є повна група гіпотез H_1, H_2, \dots, H_n . Ймовірності цих гіпотез до проведення експерименту відомі і дорівнюють $P(H_1), P(H_2), \dots, P(H_n)$. Проводиться експеримент, в результаті якого відбувається деяка подія A . Поставимо запитання, як змінюються ймовірності гіпотез після проведення експерименту, тобто після того, як стає відомо, що відбулася подія A . Отже, необхідно знайти умовну ймовірність $P(H_i|A)$ для кожної гіпотези. Ці умовні ймовірності називають **апостеріорними** ймовірностями гіпотез. Відповідь на поставлене запитання дає так звана **формула Байєса**. Отримаємо її. Згідно з теоремою про добуток ймовірностей

$$P(AH_i) = P(A)P(H_i|A) = P(H_i)P(A|H_i).$$

Звідси

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)}.$$

Підставимо в останню рівність замість ймовірності $P(A)$ її значення з формули повної ймовірності (1). В результаті будемо мати

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{k=1}^n P(H_k)P(A|H_k)}. \quad (2)$$

Це і є **формула Байєса**. Формула Байєса (2) дає можливість зробити "переоцінку" ймовірності кожної гіпотези після того, як стає відома деяка нова інформація про виконання тих чи інших подій. Особливе значення ця формула має для таких експериментів, в яких події H_i безпосередньо не спостерігаються, хоча апріорні ймовірності $P(H_i)$ і умовні ймовірності $P(A|H_i)$ відомі з додаткових експериментів. Така ситуація має місце, наприклад, коли відсутній пристрій, який дозволяє реєструвати факт настання гіпотез H_i , або ж,

коли застосовується пристрій для реєстрації гіпотез, який призводить до руйнування предмету спостереження (в техніці це називають руйнівним контролем). У таких випадках переоцінку ймовірностей гіпотез після проведення експерименту можна зробити на основі події A , що відбулася в результаті цього експерименту.

Приклад 1 (на застосування формули повної ймовірності).

У цеху на автоматичних верстатах трьох типів виготовляють одні і ті ж деталі. Продуктивність верстатів однаакова, але якість роботи різна. Відомо, що верстати першого типу виробляють 94% деталей відмінної якості, верстати другого типу – 90%, верстати третього типу – 85%. Всі виготовлені в цеху за зміну деталі в нерозсортованому вигляді занесені до складу. Знайти ймовірність того, що навмання вибрана деталь буде відмінної якості, якщо відомо, що верстатів першого типу 5, другого – 3 і третього – 2.

Розв'язання. Нехай подія A – навмання вибрано деталь відмінної якості. Введемо три гіпотези:

H_1 – вибрано деталь, яку виготовлено на верстаті першого типу;

H_2 – вибрано деталь, яку виготовлено на верстаті другого типу;

H_3 – вибрано деталь, яку виготовлено на верстаті третього типу.

Оскільки продуктивність верстатів однаакова, то для гіпотез H_1, H_2, H_3 маємо такі ймовірності :

$$P(H_1) = \frac{5}{10} = \frac{1}{2}; \quad P(H_2) = \frac{3}{10}; \quad P(H_3) = \frac{2}{10} = \frac{1}{5}.$$

Умовні ймовірності події A дорівнюють

$$P(A/H_1) = 0,94; \quad P(A/H_2) = 0,9; \quad P(A/H_3) = 0,85.$$

Тому

$$P(A) = \sum_{i=1}^3 P(H_i)P(A/H_i) = \frac{1}{2}0,94 + \frac{3}{10}0,9 + \frac{1}{5}0,85 = 0,91.$$

Приклад 2 (на застосування формули Байєса). В умовах попередньої задачі знайти ймовірність того, що навмання вибрану деталь виготовлено на верстаті першого типу, якщо відомо, що ця деталь відмінної якості.

Розв'язання. За формuloю (2)

$$P(H_1/A) = \frac{P(H_1)P(A/H_1)}{\sum_{i=1}^3 P(H_i)P(A/H_i)} = \frac{\frac{1}{2} \cdot 0,94}{0,91} \approx 0,52.$$

Приклад 3. З урни, яка містить 10 білих і 20 чорних кульок, одну кульку невідомого кольору загублено. Яка ймовірність витягнути з урни білу кульку?

Розв'язання. Нехай H_1 – загублено білу кульку ; H_2 – загублено чорну кульку; подія A – з кульок, що залишилися, витягнуто білу кульку. Тоді

$$P(H_1) = \frac{10}{30} = \frac{1}{3}; \quad P(H_2) = \frac{20}{30} = \frac{2}{3};$$

$$P(A/H_1) = \frac{10-1}{30-1} = \frac{9}{29}; \quad P(A/H_2) = \frac{10}{30-1} = \frac{10}{29};$$

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) = \frac{1}{3} \cdot \frac{9}{29} + \frac{2}{3} \cdot \frac{10}{29} = \frac{9+20}{3 \cdot 29} = \frac{1}{3}.$$

Відмітимо, що ймовірність витягнути з урни білу кульку до того, як загубили кульку, також дорівнює $\frac{1}{3}$.

Приклад 4. В умовах попередньої задачі знайти ймовірності того, що загублено білу та чорну кульки, якщо відомо, що витягнуто білу кульку.

Розв'язання.

$$P(H_1/A) = \frac{\frac{1}{3} \cdot \frac{9}{29}}{\frac{1}{3}} = \frac{9}{29}; \quad P(H_2/A) = \frac{\frac{2}{3} \cdot \frac{10}{29}}{\frac{1}{3}} = \frac{20}{29}.$$

Перейдемо до **поняття послідовності незалежних випробувань**.

На практиці досить часто зустрічаються задачі, в яких один і той же експеримент або аналогічні експерименти повторюються велику кількість разів. У результаті кожного експерименту може відбутися чи ні деяка подія A , причому цікавляється не результатом кожного окремого експерименту, а загальним числом наставань події A у серії експериментів. Послідовні експерименти (будемо називати їх послідовними випробуваннями) називаються незалежними відносно події A , якщо ймовірність цієї події у кожному експерименті (випробуванні) не залежить від того, відбулася ця подія чи ні в інших експериментах (випробуваннях). Наприклад, декілька послідовних підкидань монети є незалежними.

Незалежні випробування можуть проводитись в однакових або різних умовах. У першому випадку ймовірність події A у всіх випробуваннях однакова, у другому випадку – різна. Далі розглядається перший випадок. Розв'яжемо таку задачу: яка

ймовірність P_n того, що у серії з n незалежних випробувань подія A настане рівно m ($m=0, 1, \dots, n$) разів? Очевидно, що m наставань події A у n випробуваннях є випадковою подією. Позначимо цю подію через B_m . Нехай ймовірність події A у кожному випробуванні дорівнює p . Введемо події

$$A_i = \text{при } i\text{-му випробуванні настала подія } A \quad (i=1,2,3,\dots,n).$$

Зрозуміло, що $P(A_i) = p$, $P(\bar{A}_i) = 1 - p = q$. Однією з елементарних випадкових подій, які сприяють події B_m , є такий результат серії випробувань, коли подія A відбувається m разів підряд, а потім $(n-m)$ разів підряд не відбувається

$$A_1 A_2 \dots A_m \bar{A}_{m+1} \bar{A}_{m+2} \dots \bar{A}_n.$$

Оскільки випробування незалежні, то

$$P(A_1 A_2 \dots A_m \bar{A}_{m+1} \bar{A}_{m+2} \dots \bar{A}_n) = P(A_1) \dots P(A_m) P(\bar{A}_{m+1}) \dots P(\bar{A}_n) = p^m q^{n-m}. \quad (3)$$

Кожен інший варіант настання події B_m також складається з m подій A_i і $(n-m)$ подій \bar{A}_j (ще кажуть: з m “успіхів” і $(n-m)$ “невдач”). Кількість таких варіантів дорівнює кількості способів, якими можна з n випробувань вибрати m , в яких подія A відбувається, тобто дорівнює C_n^m . Ймовірність кожного окремого варіанту, як і в (3), дорівнює $p^m q^{n-m}$. Оскільки всі варіанти настання події B_m між собою несумісні, то

$$P(B_m) = P_n(n) = C_n^m p^m q^{n-m}. \quad (4)$$

Отже, ми приходимо до такого результату: якщо проводяться n незалежних випробувань, у кожному з яких подія A може відбутися з ймовірністю p , то ймовірність настання події A рівно m разів визначається за формулою (4). Цю формулу називають **формулою Бернуллі**.

Приклад 5. Монета підкидається п'ять разів. Знайти ймовірність того, що: а) три рази випаде “герб”; б) “герб” випаде не менше трьох разів.

Розв'язання. За формулою Бернуллі:

$$\text{а) } P_5(3) = C_5^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^{5-3} = \frac{C_5^3}{2^5} = \frac{5}{16};$$

$$\text{б) } P_5(3 \leq m \leq 5) = P_5(3) + P_5(4) + P_5(5) = \frac{1}{2}, \text{ оскільки}$$

$$P_5 \cdot 3 = C_5^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{10}{32}; \quad P_5 \cdot 4 = C_5^4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^1 = \frac{5}{32};$$

$$P_5 \cdot 5 = C_5^5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^0 = \frac{1}{32}.$$

Зауваження. Якщо розглядати ймовірність $P_n m$ як функцію від m , то $P_n m$ при фіксованому n спочатку збільшується зі збільшенням m від 0 до деякого m_0 , а потім зменшується зі збільшенням m від m_0 до n (число m_0 називають **найбільш ймовірним числом появи події A**).

Справді, з відношення

$$\frac{P_n m+1}{P_n m} = \frac{C_n^{m+1} p^{m+1} q^{n-m+1}}{C_n^m p^m q^{n-m}} = \frac{n-m}{m+1} \cdot \frac{p}{q}$$

маємо:

- 1) $P_n m+1 > P_n m$, якщо $n-m > m+1$ q , тобто $m < n+1-p-1$;
- 2) $P_n m+1 = P_n m$, якщо $m = n+1-p-1$;
- 3) $P_n m+1 < P_n m$, якщо $m > n+1-p-1$. Звідси, якщо $n+1-p$ є цілим числом, то найбільш ймовірними будуть числа $m_0 = n+1-p$ та m_0-1 ; якщо ж $n+1-p$ не є цілим числом, то $m_0 = \lfloor n+1-p \rfloor$, тобто є цілою частиною числа $n+1-p$.

Лекція № 6

Формули наближених обчислень у схемі Бернуллі

У лекції 5 введено поняття послідовності незалежних випробувань і отримано формулу Бернуллі

$$P(B_m = P_n m) = C_n^m p^m q^{n-m}. \quad (1)$$

Відмітимо, що події B_m ($m = 0, 1, 2, \dots, n$) попарно несумісні і в сумі дають достовірну подію Ω . Тому повинно бути

$$\sum_{m=1}^n P(B_m) = \sum_{m=1}^n P_n m = P(\Omega) = 1. \quad (2)$$

Саме це маємо за формулою бінома Ньютона:

$$\sum_{m=1}^n P(B_m) = \sum_{m=1}^n C_n^m p^m q^{n-m} = p + q^n = 1, \text{ оскільки } p + q = 1.$$

Формула Бернуллі описує розподіл ймовірностей між можливими значеннями числа наставань події A у n випробуваннях. Оскільки ймовірність $P_n m$ є членом розкладу бінома $p+q^n$, розподіл ймовірностей такого типу називають ***біномним розподілом***.

Схему незалежних випробувань, в кожному з яких подія A може відбутися з однаковою ймовірністю p , ще називають ***схемою Бернуллі***.

При великих n обчислення за формулою Бернуллі стають громіздкими. У той самий час для потреб практики необхідно обчислювати біномні ймовірності $P_n m$ або $P_n m_1 \leq m \leq m_2$ (ймовірність того, що число “успіхів” лежить в межах від m_1 до m_2) саме для великих n . Ускладнення обчислювального характеру виникають також при малих p або q . У таких випадках часто користуються різними наближеними формулами для ймовірностей $P_n m$ і $P_n m_1 \leq m \leq m_2$, до розгляду яких ми і переходимо.

Теорема Пуассона. Нехай в схемі Бернуллі ймовірність “успіху” в кожному з n випробувань дорівнює $p = \frac{\lambda}{n}$, де $\lambda = const$, $0 < \lambda < n$.

Тоді

$$\lim_{\substack{n \rightarrow \infty \\ (p \rightarrow 0)}} P_n m = \frac{\lambda^m}{m!} e^{-\lambda} \text{ для будь-якого } m = 0, 1, 2, \dots$$

Доведення. За формулою Бернуллі

$$\begin{aligned} P_n m &= \frac{n(n-1) \dots (n-m+1)}{m!} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} = \\ &= \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{m-1}{n}\right). \end{aligned}$$

Звідси при $n \rightarrow \infty$ отримуємо твердження теореми, оскільки

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

На практиці теоремою Пуассона користуються у формі наближеної рівності

$$P_n m \approx \frac{\lambda^m}{m!} e^{-\lambda}, \quad \lambda = np. \quad (3)$$

Цією формулою зручно користуватись при великих n і значеннях p , близьких до нуля, тобто для тих подій, що трапляються рідко. Для

порівняння точних (формула 1) і наближених (формула 3) значень ймовірностей наведемо таку таблицю

	m	0	1	2	3	4	5
(1)	$P_{10} \Phi(n) = 1/5$	0,1074	0,2684	0,3020	0,2013	0,0880	0,0264
	$P_{100} \Phi(n) = 1/50$	0,1326	0,2707	0,2734	0,1823	0,0902	0,0353
(3)	$\lambda^m e^{-\lambda} / m!$ $\lambda = np = 2$	0,1353	0,2707	0,2707	0,1805	0,0902	0,0361

Бачимо, що при великих n і малих p похибка може бути досить малою. Формула (3) часто використовується при $n > 100$, $p < 0,1$ і $np < 10$. Якщо малим є значення $q = 1 - p$, то пуассонівським наближенням користуються для числа “невдач”.

У випадку, коли обидва параметри p і q помітно відрізняються від нуля, використовуються **локальна та інтегральна теореми Муавра-Лапласа**. Ми не зупиняємося тут на точному формулюванні і доведеннях названих теорем, а наведемо лише ті наближені рівності, які є їх наслідками (обґрунтування цих рівностей буде проведено в кінці курсу теорії ймовірностей).

Введемо функції

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}; \quad \Phi(x) = \int_{-\infty}^x \Phi(u) du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du. \quad (4)$$

Для цих функцій складені спеціальні таблиці (див. Додаток 2).

Локальна теорема Муавра-Лапласа використовується на практиці для достатньо великих n у вигляді наближеної рівності

$$P_n(n) \approx \frac{1}{\sqrt{npq}} \Phi\left(\frac{n-m}{\sqrt{npq}}\right), \quad \text{де } x_m = \frac{m-np}{\sqrt{npq}}. \quad (5)$$

Якщо значення p або q близьке до нуля, формула (5) може давати надто грубі наближення. У цьому випадку слід користуватись формулами Пуассона (3). Достатньо точні результати формула (5) дає при $p = q = \frac{1}{2}$. Слід зазначити також, що m і n в (5) не повинні відрізнятись одне від одного надто сильно. Наприклад, для $P_n(n)$ формула (5) дає погане наближення.

Наближена рівність, у формі якої застосовують на практиці інтегральну теорему Муавра-Лапласа, має вигляд

$$P_n \left(a \leq \frac{m-np}{\sqrt{npq}} \leq b \right) \approx \Phi(b) - \Phi(a). \quad (6)$$

Оскільки з нерівностей $m_1 \leq m \leq m_2$ випливають нерівності $\frac{m_1-np}{\sqrt{npq}} \leq \frac{m-np}{\sqrt{npq}} \leq \frac{m_2-np}{\sqrt{npq}}$, то формулу (6) можна записати у вигляді, зручному при обчисленні ймовірності події $m_1 \leq m \leq m_2$, а саме

$$P_n(m_1 \leq m \leq m_2) \approx \Phi(x_2) - \Phi(x_1),$$

де $x_1 = \frac{m_1-np}{\sqrt{npq}}, \quad x_2 = \frac{m_2-np}{\sqrt{npq}}.$ (7)

Якщо n не менше кількох сотень, а p не дуже близьке до 0 або 1, то наближення (6), (7) є цілком задовільними.

Значення функцій $\Phi(x)$ і $\Phi(-x)$, як вже говорилось, табулювані. При користуванні таблицями цих функцій слід мати на увазі, що

$$\Phi(x) = \varphi(x), \quad \Phi(-x) = 1 - \varphi(x). \quad (8)$$

Доведемо другу з цих рівностей:

$$\begin{aligned} \Phi(-x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du - \frac{1}{\sqrt{2\pi}} \int_{-x}^{\infty} e^{-\frac{u^2}{2}} du = \\ &= 1 - \frac{1}{\sqrt{2\pi}} \int_{-x}^{\infty} e^{-\frac{u^2}{2}} du = \left| \begin{array}{l} \text{заміна} \\ u = -t \\ du = -dt \end{array} \right| = 1 + \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{t^2}{2}} dt = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \\ &= 1 - \varphi(x). \end{aligned}$$

Тут ми скористалися інтегралом Пуассона

$$\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = \sqrt{2\pi}. \quad (*)$$

Формула (6) дозволяє дати оцінку близькості відносної частоти та ймовірності події. Нехай p – ймовірність настання події A в схемі

Бернуллі, а m – загальне число “успіхів”. Тоді $\frac{m}{n}$ є відносною частотою події A . Знайдемо ймовірність події $\left\{ \left| \frac{m}{n} - p \right| \leq \varepsilon \right\}$, де $\varepsilon > 0$ – довільне число.

Якщо n достатньо велике, то за формулою (6)

$$\begin{aligned} P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) &= P\left(-\varepsilon\sqrt{\frac{n}{pq}} \leq \frac{m-np}{\sqrt{npq}} \leq \varepsilon\sqrt{\frac{n}{pq}}\right) \approx \\ &\approx \Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) - \Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right) = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) - 1. \end{aligned} \quad (9)$$

Якщо перейти до границі при $n \rightarrow \infty$, то можна записати

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) = 2\Phi(+\infty) - 1 = 1. \quad (10)$$

Тут ми знову скористались інтегралом Пуассона (*).

На основі рівності (10) можна **стверджувати**, що яким би малим не було наперед задане число $\varepsilon > 0$, при достатньо великому числі випробувань з ймовірністю, як завгодно близькою до 1, відносна частота $\frac{m}{n}$ події відрізняється від ймовірності p цієї події не більше ніж на ε .

Сформульований результат носить назив **теореми Бернуллі**.

Теорема Бернуллі є математичним підтвердженням відомого з практики факту, що відносна частота події при великому числі випробувань є близькою до ймовірності цієї події:

$$\frac{m}{n} \approx p. \quad (11)$$

Однак, тут необхідно дати певні пояснення. Насамперед зазначимо, що теорема Бернуллі не дає підстав стверджувати існування граници $\lim_{n \rightarrow \infty} \frac{m}{n} = p$ в звичайному для нас розумінні, тому що

не можна, наприклад, категорично заперечувати факт настання при всіх випробуваннях тільки “успіху” (в цьому випадку $m=n$) і $\lim_{n \rightarrow \infty} \frac{m}{n} = 1$) або тільки “невдачі” (тут $m=0$ і $\lim_{n \rightarrow \infty} \frac{m}{n} = 0$). А якщо це так,

то тоді виникає природне питання, який же математичний смисл наближеної рівності (11)? Відповідь на поставлене питання саме і дає теорема Бернуллі, підkreślуючи, насамперед, ймовірнісний характер цієї рівності і стверджуючи, що ймовірність її виконання з будь-якою наперед заданою точністю ε зі збільшенням числа випробувань прямує до 1. Іншими словами, при невеликому числі випробувань відносна частота події є випадковою; при збільшенні

числа випробувань вона стабілізується і наближається до невипадкового числа – ймовірності події.

До теореми Бернуллі ми ще повернемося в кінці курсу теорії ймовірностей.

Наведемо ряд прикладів на використання наближених формул (5), (7) та (9).

Приклад 1. Знайти ймовірність того, що подія A відбудеться рівно 80 разів в 400 випробуваннях, якщо ймовірність її настання в кожному випробуванні дорівнює 0,2.

Розв'язання. За умовою $n = 400$; $m = 80$; $p = 0,2$; $q = 0,8$. Скористаємося формулою (5):

$$P_{400} \approx \frac{1}{\sqrt{400 \cdot 0,2 \cdot 0,8}} \cdot \varphi \left(\frac{1}{8} \right) \text{де } x = \frac{m-np}{\sqrt{npq}} = \frac{80-400 \cdot 0,2}{8} = 0.$$

За таблицею значень функції φ (див. Додаток 2) знаходимо $\varphi \approx 0,3989$.

Отже, $P_{400} \approx \frac{1}{8} \cdot 0,3989 = 0,04986$. Формула Бернуллі (2) дає майже

такий же результат: $P_{400} \approx 0,0498$.

Приклад 2. Ймовірність випадання “герба” при кожному підкиданні “неправильної” монети дорівнює $p = 0,75$. Знайти ймовірність того, що при 10 підкиданнях монети “герб” випаде рівно 8 разів.

Розв'язання. За умовою $n = 10$; $m = 8$; $p = 0,75$; $q = 0,25$. За наближеною формулою (5)

$$P_{10} \approx \frac{1}{\sqrt{10 \cdot 0,75 \cdot 0,25}} \varphi(x) = 0,7301 \varphi(x) \text{ де } x = \frac{m-np}{\sqrt{npq}} =$$

$$= \frac{8-10 \cdot 0,75}{\sqrt{10 \cdot 0,75 \cdot 0,25}} \approx 0,36. \text{ За таблицею } \varphi(0,36) \approx 0,3739. \text{ Отже,}$$

$P_{10} \approx 0,7301 \cdot 0,3739 \approx 0,273$. Формула Бернуллі (2) приводить до $P_{10} \approx 0,282$. Розбіжність результатів пояснюється тим, що в даному прикладі n мале, а формула (5) дає хороші результати лише при достатньо великих n .

Приклад 3. Ймовірність того, що деталь не буде перевірено, дорівнює 0,2. Знайти ймовірність того, що серед 400 випадково відібраних деталей неперевірених буде від 70 до 100 деталей.

Розв'язання. За умовою $n = 400$; $p = 0,2$; $q = 0,8$; $m_1 = 70$; $m_2 = 100$. Скористаємося формулою (7):

$$x_1 = \frac{m_1 - np}{\sqrt{npq}} = \frac{70 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = -1,25; \quad x_2 = \frac{m_2 - np}{\sqrt{npq}} = \frac{100 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = 2,5.$$

Отже, $P_{400} \{0 \leq m \leq 100\} \approx \Phi(2,5) - \Phi(-1,25) = \Phi(2,5) - \Phi(-1)$.

За таблицею (див. Додаток 2) $\Phi(2,5) \approx 0,9938$, $\Phi(-1) \approx 0,8944$. Тому $P_{400} \{0 \leq m \leq 100\} \approx 0,8882$.

Приклад 4. Ймовірність того, що деталь нестандартна, дорівнює 0,1. Знайти ймовірність того, що для випадково відібраних 400 деталей відносна частота нестандартних деталей буде відрізнятись від ймовірності $p = 0,1$ не більше ніж на 0,03.

Розв'язання. За умовою $n = 400$; $p = 0,1$; $q = 0,9$; $\epsilon = 0,03$. Скористаємося наближеною формулою (9):

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

За таблицею $\Phi(2) \approx 0,9772$. Отже, $2\Phi(2) - 1 = 0,9544$. Отриманий результат можна сформулювати в такій формі: якщо взяти достатньо велике число проб по 400 деталей в кожній, то майже в 95,44% цих проб відхилення відносної частоти від постійної ймовірності $p = 0,1$ не перевищить 0,03.

Лекція № 7

Поняття випадкової величини

Елементарні події випадкових експериментів можуть бути об'єктами довільної природи, наприклад, точками мішені, послідовностями “успіхів” і “невдач” в схемі Бернуллі, неперервними функціями тощо. Проте при проведенні більшості випадкових експериментів нас цікавить (і, як правило, реєструється) не сама елементарна подія, а деяка величина X (або ряд величин), що характеризує результат, наприклад, число “успіхів” в тій же схемі Бернуллі, а не сама послідовність з “успіхів” і “невдач”. Ця величина в результаті експерименту може приймати те або інше значення (але тільки одне), причому наперед, до експерименту, не відомо, яке саме. Подібні величини носять назву **випадкових**.

Оскільки у будь-якому разі результат випадкового експерименту визначається елементарною подією, під випадковою величиною X природно розуміти деяку числову функцію елементарної події:

$$X = X(\omega), \omega \in \Omega. \quad (1)$$

Випадкові величини низче будемо позначати великими літерами латинського алфавіту, тоді як їх можливі значення – відповідними малими буквами, наприклад X, Y – випадкові величини, а x_i, y_i – їх можливі значення.

Розглянемо ряд прикладів.

Приклад 1. Підкидається гральний кубик. Множиною елементарних подій є $\Omega = \omega_1, \omega_2, \dots, \omega_6$, де ω_i – випадання грані в i очок. Нехай $X(\omega_i) = i$, тобто кожному підкиданню кубика ставиться у відповідність число очок, що випало при цьому підкиданні.

Приклад 2. Два рази підкидається монета. Множиною елементарних подій є $\Omega = \text{ГГ}, \text{ГЦ}, \text{ЦГ}, \text{ЦЦ}$. Позначимо через X випадкову величину, яка дорівнює числу випадань “герба”. Тоді таблиця значень функції $X(\omega)$ має такий вигляд:

ω	ГГ	ГЦ	ЦГ	ЦЦ
$X(\omega)$	2	1	1	0

Приклад 3. Монета підкидається до першого випадання “герба”. Тоді $\Omega = \omega_1, \omega_2, \omega_3, \dots, \omega_n, \dots$, де $\omega_n = \underbrace{\text{ЦЦ} \dots \text{ЦГ}}_{n-1}$ означає, що “герб”

випав при n -му підкиданні. Нехай $X(\omega_n) = n$ – число підкидань монети до першого випадання “герба”. Можливі значення X – це всі натуральні числа.

Приклад 4. Проводиться n незалежних випробувань (схема Бернуллі). Елементарна подія – це деяка послідовність “успіхів” і “невдач” в n випробуваннях $\underbrace{\text{УУНННУ} \dots \text{У}}_n$. Нехай X_n – число “успіхів” в серії з n випробувань. Можливими значеннями величини X_n є $0, 1, 2, \dots, n$.

Приклад 5. У коло радіуса R з центром в початку координат навмисно кидають точку; $\Omega = \omega = x, y : x^2 + y^2 \leq R^2$. Нехай $X(\omega) = \sqrt{x^2 + y^2}$ – відстань від точки попадання ω до центра кола. Можливі значення X – це точки відрізка $[0, R]$.

У наведених прикладах випадкова величина описана як функція елементарної події. Але задання лише самої функції ще недостатньо для повної характеристики випадкової величини. Необхідно ще мати змогу визначити ймовірності, з якими випадкова величина приймає свої значення.

Якщо множина елементарних подій Ω – **дискретна**, як у прикладах 1-4, то завжди можна сказати, з якою ймовірністю функція X набуває своїх значень. Так, у першому прикладі $P(X=i) = 1/6$, $i = 1, 2, \dots, 6$. У другому прикладі $P(X=2) = 1/4$; $P(X=1) = 1/2$; $P(X=0) = 1/4$. Для третього прикладу $P(X=n) = \frac{1}{2^n}$, $n = 1, 2, 3, \dots$. Для четвертого прикладу $P(X_n=m) = C_n^m p^m q^{n-m}$, $m = 0, 1, 2, \dots, n$.

Зокрема, у випадку **дискретних множин** Ω завжди можна визначити ймовірність того, що випадкова величина X приймає значення з деякого скінченного чи нескінченного числового інтервалу, наприклад, ймовірність $P(X < x) = P(\omega : X(\omega) < x)$, де x – довільне дійсне число. Так, у другому прикладі

$$P(X < 2) = P(X=0) + P(X=1) = 3/4.$$

Тому **випадковою величиною в дискретних ймовірнісних просторах** Ω, P можна назвати довільну числову функцію елементарної події (1). Очевидно, що множина можливих значень $x_1, x_2, \dots, x_n, \dots$ такої функції – дискретна (скінчена або зчисленна). Подібні випадкові величини називаються **дискретними** і вважаються повністю описаними з ймовірнісної точки зору, якщо задані ймовірності їх можливих значень

$$P(X=x_i) = p_i = \sum_{\substack{k \\ \omega_k : X(\omega_k) = x_i}} P(\omega_k), \quad i = 1, 2, \dots. \quad (2)$$

Оскільки події $X=x_i$, $i = 1, 2, \dots$ попарно несумісні і утворюють повну групу подій, то $\sum_{i=1}^{\infty} p_i = 1$ (якщо X набуває скінченної кількості значень, ця сума містить скінченну кількість доданків).

Розподіл ймовірностей (2) називається **законом розподілу дискретної випадкової величини (ДВВ)**.

У випадку, коли множина елементарних подій Ω випадкового експерименту неперервна (як у прикладі 5) і значення функції (1) неперервно заповнюють скінчений або нескінчений числовий проміжок, на випадкову величину X (функцію (1)) накладаються додаткові умови. Перш за все відмітимо, що в цьому випадку цікавляться не ймовірністю окремого значення випадкової величини, а ймовірностями того, що це значення попадає в певний числовий проміжок, наприклад, ймовірностями $P \omega : x_1 \leq X \omega < x_2$ або

$P \omega : X \omega < x$, де x_1, x_2, x – дійсні числа. Як вже говорилося в лекції 2, у випадку неперервної множини елементарних подій Ω розглядаються не всі її підмножини, а деякий клас підмножин, що утворює σ -алгебру подій F . Для подій з F і тільки для них визначається ймовірність. Тому, наприклад, ймовірність $P \omega : X \omega < x$ визначено лише тоді, коли множина таких $\omega \in \Omega$, для яких $X \omega < x$, належить до σ -алгебри подій F . Саме визначеність ймовірності події $\omega : X \omega < x$ для довільного дійсного числа x покладається в основу означення випадкової величини в загальному випадку.

Означення 1. *Випадковою величиною* X називається дійсна числова функція $X = X \omega$, визначена на множині елементарних подій Ω і така, що для довільного дійсного числа x множина таких ω , для яких $X \omega < x$, належить до σ -алгебри подій F даного експерименту:

$$\omega : X \omega < x \in F \quad \forall x \in -\infty, \infty . \quad (3)$$

Зауваження 1. Якщо умова (3) виконується, то для довільної борелівської множини B на прямій подія $\omega : X \omega \in B$ буде такою, що спостерігається. Дійсно, цю подію можна одержати застосуванням скінченної або зчисленної кількості операцій до подій вигляду $\omega : X \omega < x$ (див. приклад 5 лекції 2).

З умови (3), зокрема, випливає, що

$$\omega : X \omega \geq x = \overline{\omega : X \omega < x} \in F ;$$

$$\omega : x_1 \leq X \omega < x_2 = \omega : X \omega < x_2 \setminus \omega : X \omega < x_1 \in F ; \quad (4)$$

$$\omega : X \omega = x = \bigcap_{n=1}^{\infty} \left\{ \omega : x \leq X \omega < x + \frac{1}{n} \right\} \in F .$$

Приклад 6. Повернемось до експерименту з прикладу 5. За ймовірнісну модель даного експерименту приймемо Ω, F_s, P . Тут Ω – множина точок круга, F_s – алгебра квадрованих підмножин множини Ω , а ймовірність задано формулою геометричних ймовірностей $P(A) = \frac{S(A)}{S(\Omega)}$ (див. лекцію 4). Нехай F – σ -алгебра борелівських підмножин множини Ω . Очевидно, що вона включає в себе всі множини з алгебри F_s . Покажемо, що подія $\omega: X \rightarrow \omega < x$ належить до F_s , а отже, і до σ -алгебри F для довільного $x \in (-\infty, \infty)$. Дійсно,

$$\omega: X \rightarrow \omega < x = \begin{cases} \emptyset & , x \leq 0 \\ u, v : u^2 + v^2 < x^2 & , 0 < x \leq R \\ \Omega & , x > R \end{cases}.$$

Очевидно, що для дискретної випадкової величини X умова (3) завжди виконується, оскільки будь-яка підмножина множини Ω в цьому випадку є подією, що спостерігається.

Згідно з умовою (3) для всіх $x \in (-\infty, \infty)$ визначена ймовірність $P(\omega < x)$. Нижче також будемо використовувати скорочений запис $\omega: X \rightarrow \omega < x = X < x$. Введемо позначення

$$F_x(x) = P(X < x). \quad (5)$$

Означення 2. Функція (5) називається *функцією розподілу* випадкової величини X .

Вона визначена на всій числовій осі. Нижче замість $F_x(x)$ будемо писати просто $F(x)$. Сформулюємо загальні *властивості функції розподілу*:

Властивість 1. Значення функції розподілу належать відрізку $[0, 1] : 0 \leq F(x) \leq 1$.

Ця властивість випливає з означення функції розподілу і аксіоматичного означення ймовірності.

Властивість 2. $F(x)$ – неспадна функція, тобто $F(x_2) \geq F(x_1)$ при $x_2 > x_1$.

Дійсно, нехай $x_2 > x_1$. Розглянемо подію $C = X < x_2$. Цю подію можна записати у вигляді суми двох несумісних подій, а саме $C = A + B$, де $A = X < x_1$, $B = x_1 \leq X < x_2$. Тому

$$P(C) = P(A) + P(B) \quad \text{або} \quad P(X < x_2) - P(X < x_1) = P(x_1 \leq X < x_2).$$

Згідно з означенням функції розподілу і аксіоматичним означенням ймовірності $F(x_2) - F(x_1) = P(x_1 \leq X < x_2) \geq 0$.

Отже, $F(x_2) \geq F(x_1)$, що й потрібно було показати.

Властивість 3. Ймовірність попадання випадкової величини у півінтервал $[x_1, x_2]$ дорівнює приросту її функції розподілу на цьому півінтервалі:

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1).$$

Ця властивість є наслідком доведення попередньої властивості.

Властивість 4. Якщо можливі значення випадкової величини належать інтервалу $[x_1, x_2]$, то $F(x) = 0$ при $x \leq x_1$ і $F(x) = 1$ при $x \geq x_2$.

Дійсно, якщо $x \leq x_1$, то подія $X < x$ є неможливою і тому $F(x) = P(X < x) = 0$. Якщо ж $x \geq x_2$, то подія $X < x$ є вірогідною (достовірною) і тому $F(x) = P(X < x) = 1$.

Очевидним наслідком даної властивості є наступна властивість.

Властивість 5. $F(-\infty) = 0$, $F(+\infty) = 1$.

З теореми про неперервність ймовірності (див. лекцію 2) і властивості 3 випливає, що

$$\begin{aligned} P(X=x) &= \lim_{n \rightarrow \infty} P\left\{x \leq X < x + \frac{1}{n}\right\} = \\ &= \lim_{n \rightarrow \infty} F\left(x + \frac{1}{n}\right) - F(x) = F(x+0) - F(x), \end{aligned} \tag{6}$$

де $F(x+0)$ означає границю справа. Останньою формулою можна користуватись для знаходження розподілу ймовірностей (2) дискретної випадкової величини, якщо відома її функція розподілу $F(x)$.

Властивість 6. $F(x)$ неперервна зліва, тобто $F(x-0) = F(x)$ для будь-якого x , де $F(x-0)$ – границя зліва.

Доведення. Розглянемо послідовність чисел $x_0 < x_1 < x_2 < \dots < x_n < \dots < x$ таку, що $x_n \rightarrow x$ при $n \rightarrow \infty$. Тоді

$$X < x = X < x_0 + \sum_{k=1}^{\infty} x_{k-1} \leq X < x_k , \quad \text{де події } X < x_0 ,$$

$x_{k-1} \leq X < x_k$, $k = 1, 2, \dots$ – попарно несумісні. Згідно з розшириною аксіомою додавання (див. лекцію 2) і властивістю 3) функції $F(x)$

$$\begin{aligned} \text{отримуємо } F(x) &= P(X < x) = P(X < x_0) + \sum_{k=1}^{\infty} P(x_{k-1} \leq X < x_k) = \\ &= F(x_0) + \sum_{k=1}^{\infty} [F(x_k) - F(x_{k-1})] = F(x_0) + \lim_{n \rightarrow \infty} \sum_{k=1}^n [F(x_k) - F(x_{k-1})] = \\ &= F(x_0) + \lim_{n \rightarrow \infty} [F(x_n) - F(x_0)] = \lim_{n \rightarrow \infty} F(x_n) . \end{aligned}$$

Звідси, $F(x) = F(x) - 0$.

Отже, ми показали, що функція розподілу $F(x)$ випадкової величини є неспадною, неперервною зліва і задовольняє умовам $F(-\infty) = 0$, $F(+\infty) = 1$.

Виявляється, що справедливим є й обернене твердження, тобто будь-яка неспадна, неперервна зліва функція $F(x)$, яка задовольняє умовам $F(-\infty) = 0$, $F(+\infty) = 1$, є функцією розподілу деякої випадкової величини.

Наведемо приклад такої випадкової величини. Для цього розглянемо ймовірнісний простір $\tilde{\Omega}, \tilde{F}, \tilde{P}$, в якому $\tilde{\Omega} = R$, тобто елементарними подіями $\tilde{\omega} \in \tilde{\Omega}$ є точки всієї числовової осі, \tilde{F} – система борелівських множин на прямій (див. приклад 5 лекції 2), а ймовірність \tilde{P} введено таким чином, що для борелівських множин $\tilde{\omega} : \tilde{\omega} < x$, $x \in R$ вона дорівнює

$$\tilde{P}(\tilde{\omega} : \tilde{\omega} < x) = F(x), \quad (8)$$

де $F(x)$ – довільна неспадна, неперервна зліва функція, яка задовольняє умовам $F(-\infty) = 0$, $F(+\infty) = 1$.

Можна показати (ми цього робити не будемо), що розподілом ймовірностей (8) однозначно визначається ймовірність для будь-якої

борелівської множини з \tilde{F} , зокрема,

$$\tilde{P} \tilde{\omega}: x_1 \leq \tilde{\omega} < x_2 = F(x_2) - F(x_1). \quad (9)$$

Розглянемо в просторі $\tilde{\Omega}, \tilde{F}, \tilde{P}$ випадкову величину

$$\tilde{X} = \tilde{X}(\tilde{\omega}) = \tilde{\omega}. \quad (10)$$

Тоді згідно з (8)

$$\tilde{P} \tilde{\omega}: \tilde{X}(\tilde{\omega}) < x = \tilde{P} \tilde{X} < x = F(x), \quad (11)$$

тобто функція $F(x)$ є функцією розподілу випадкової величини (10).

Зauważення 2. Випадкова величина X є повністю описаною з ймовірнісної точки зору, якщо її функція розподілу $F(x)$ відома. Зручність використання функції розподілу $F(x)$ для описання випадкової величини пояснюється тим, що при проведенні випадкового експерименту реєструється, як правило, не елементарна подія, якою закінчується експеримент, а значення випадкової величини. Тобто, реєструється функція, а не значення аргумента. Вивчення окремо взятої випадкової величини X за допомогою функції розподілу $F(x)$ зв'язане по суті з переходом від вихідного ймовірнісного простору Ω, F, P до ймовірнісного простору $\tilde{\Omega}, \tilde{F}, \tilde{P}$ і від випадкової величини X до випадкової величини (10) \tilde{X} , яка має таку ж функцію розподілу $F(x)$, тобто

$$P X < x = \tilde{P} \tilde{X} < x = F(x). \quad (12)$$

Якщо задано функцію розподілу $F(x)$ випадкової величини X , то кажуть, що задано закон розподілу випадкової величини у формі функції розподілу. У деяких випадках закон розподілу випадкової величини можна задати в більш зручній формі, наприклад, у формі розподілу ймовірностей (2) для дискретної випадкової величини. У наступній лекції ми познайомимося ще з одним типом випадкової величини, що допускає більш зручну форму закону розподілу.

Лекція № 8

Дискретні та неперервні випадкові величини

На практиці зустрічаються випадкові величини в основному двох типів – дискретні та неперервні. З дискретними випадковими величинами ми частково ознайомились у минулій лекції.

Закон розподілу дискретної випадкової величини зручно подавати у вигляді таблиці (див. лекцію 7)

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

яку ще називають *рядом розподілу*.

Якщо ряд розподілу дискретної випадкової величини відомий, то завжди можна побудувати функцію розподілу і навпаки, за заданою функцією розподілу завжди можна записати ряд розподілу (див. формулу (6) лекції 7). Нехай є ряд розподілу деякої дискретної випадкової величини. Тоді для функції розподілу можна записати $F(x) = P(X < x) = \sum_{x_i < x} p_i$, де сумування поширюється на ті

індекси i , для яких $x_i < x$. Функцію $F(x)$ можна також подати у вигляді

$$F(x) = \begin{cases} 0, & x \leq x_1, \\ p_1, & x_1 < x \leq x_2, \\ p_1 + p_2, & x_2 < x \leq x_3, \\ \dots & \dots \\ p_1 + p_2 + \dots + p_{n-1}, & x_{n-1} < x \leq x_n, \\ 1, & x > x_n. \end{cases}$$

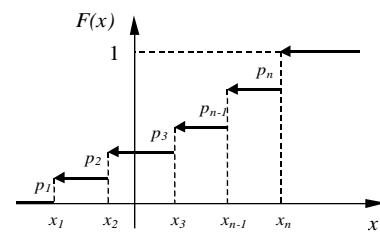


Рис. 1

Дійсно, нехай значення випадкової величини пронумеровані в порядку їх зростання. Оскільки не існує значень X , менших за x_1 , то $F(x) = 0$ при $x \leq x_1$ (нерівність $X < x$ для $x \leq x_1$ є неможливою подією, тому ймовірність цієї події дорівнює нулю). Якщо $x_1 < x \leq x_2$, то $F(x) = p_1$, оскільки існує лише одне значення X , яке задовільняє нерівності $X < x$. Це значення дорівнює x_1 і приймається з ймовірністю p_1 . Якщо $x_2 < x \leq x_3$, то нерівність $X < x$ виконується

тільки тоді, коли $X = x_1$ або $X = x_2$. Ймовірність цієї події дорівнює $p_1 + p_2$. Отже, при переході через значення x_1 функція $F(x)$ робить стрибок, який дорівнює p_1 ; при переході через значення x_2 – стрибок p_2 і т. д. При переході через значення x_n функція $F(x)$ робить останній стрибок p_n . Для всіх $x > x_n$ $F(x) = 1$. Графік функції $F(x)$ показано на рис. 1.

Приклад 1. Проводиться три незалежних постріли в ціль. Ймовірність влучення в ціль при кожному пострілі дорівнює 0,4. Побудувати функцію розподілу числа влучень в ціль.

Розв'язання. Нехай X – число влучень. Можливими значеннями випадкової величини X будуть:

$$x_1 = 0, \quad x_2 = 1, \quad x_3 = 2, \quad x_4 = 3.$$

Ймовірності можливих значень знаходимо за формулою Бернуллі:

$$P(X = x_i) = C_n^{x_i} p^{x_i} q^{n-x_i}, \text{ де } n = 3;$$

$$P(X = 0) = C_3^0 p^0 q^{3-0} = 1 \cdot 1 \cdot (1 - 0,4)^3 = 0,216;$$

$$P(X = 1) = C_3^1 p^1 q^2 = 3 \cdot 0,4 \cdot 0,6^2 = 0,432;$$

$$P(X = 2) = C_3^2 p^2 q^1 = 3 \cdot 0,4^2 \cdot 0,6 = 0,288;$$

$$P(X = 3) = C_3^3 p^3 q^0 = 1 \cdot 0,4^3 \cdot 1 = 0,064.$$

Отже, маємо ряд розподілу випадкової величини:

x_i	0	1	2	3
$P(X = x_i)$	0,216	0,432	0,288	0,064

Побудуємо функцію розподілу $F(x) = \sum_{x_i \leq x} P(X = x_i)$.

$$\text{При } x \leq 0: F(x) = \sum_{x_i < 0} P(X = x_i) = 0.$$

$$\text{При } 0 < x \leq 1: F(x) = \sum_{x_i < 1} P(X = x_i) = P(X = 0) = 0,216.$$

При $1 < x \leq 2$:

$$F(x) = \sum_{x_i < 2} P(X = x_i) = P(X = 0) + P(X = 1) = 0,216 + 0,432 = 0,648.$$

При $1 < x \leq 3$:

$$F(x) = \sum_{x_i < 3} P(X = x_i) = P(X = 0) + P(X = 1) + P(X = 2) = 0,936.$$

При $x > 3$: $F(x) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 1$.

Графік функції розподілу показано на рис.2.

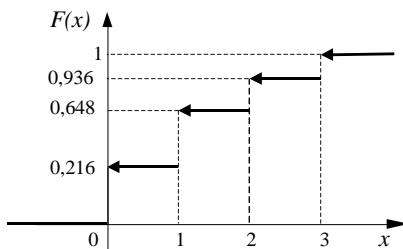


Рис. 2

Перейдемо до **неперервної випадкової величини**, під якою будемо розуміти таку випадкову величину, **функція розподілу $F(x)$ якої є неперервною скрізь і диференційовною за виключенням, можливо, скінченного числа точок.**

З цього означення випливає, що ймовірність всякого окремого значення неперервної випадкової величини дорівнює нулю, тобто якщо X – неперервна випадкова величина, то $P(X = x) = 0$.

Дійсно, для неперервної випадкової величини $P(X = x) = P(x + 0) - P(x) = 0$ (див. формулу (6) лекції 7).

Звідси, зокрема, випливає, що для неперервних випадкових величин $P(x_1 \leq X < x_2) = P(x_1 < X < x_2) = P(x_1 < X \leq x_2) =$

$= P(x_1 \leq X \leq x_2)$, тобто ймовірність попадання неперервної випадкової величини в проміжок не залежить від того, відкритий цей проміжок, замкнений чи напіввідкритий.

Відмітимо, що рівність нулю ймовірності, з якою **неперервна** випадкова величина приймає свої значення, зовсім не означає, що ці події є неможливими. Навпаки, саме ці події спостерігаються в експерименті. Якщо подія неможлива, то її ймовірність дорівнює нулю, це вірно. Обернене твердження, взагалі кажучи, неправильне, тобто з того, що ймовірність деякої події дорівнює нулю, ще не випливає, що ця подія є неможливою (див. зауваження 1 з лекції 4). Неперервна випадкова величина, приймаючи конкретні значення при кожному повторенні експерименту, саме і реалізує події, ймовірності яких дорівнюють нулю і які у той же час є можливими. Ніякого парадоксу в цьому немає, якщо цікавиться не ймовірностями елементарних подій, тобто ймовірностями окремих значень неперервної випадкової величини, а ймовірностями її попадання в

окрім інтервалі. При формуванні ймовірнісної моделі задаються саме ці останні ймовірності, а не ймовірності “попасті в точку”.

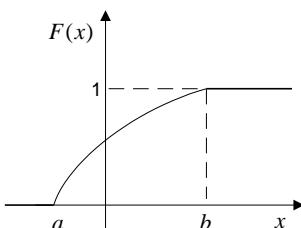


Рис. 3

Враховуючи наведені у минулій лекції властивості, функцію розподілу неперервної випадкової величини можна схематично зобразити у вигляді рис. 3.

Приклад 2. В одиничний квадрат $\Omega = \alpha, \beta : 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1$ навмання кидається точка; Ω – множина елементарних подій (точки квадрата) (рис. 4). Алгебра подій складається з усіх квадрованих підмножин множини Ω . Ймовірність вводиться за формулою геометричних ймовірностей.

Нехай випадкова величина X – абсциса точки попадання. Знайдемо її функцію розподілу. Маємо:

$$\begin{aligned} x \leq 0: \quad F(x) &= P(\alpha, \beta : X = \alpha < x) = 0; \\ 0 < x \leq 1: \quad F(x) &= P(\alpha, \beta : X = \alpha < x) = x; \\ 1 < x: \quad F(x) &= 1. \end{aligned}$$

Графік функції розподілу показано на рис.5.

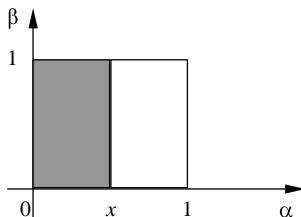


Рис. 4

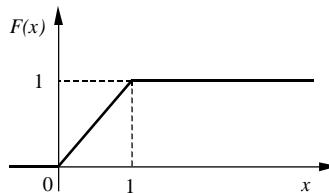


Рис. 5

Розглянемо неперервну випадкову величину з функцією розподілу $F(x)$, яку будемо вважати не тільки неперервною, але й диференційованою у точці x . Знайдемо ймовірність попадання величини X в інтервал $x, x + \Delta x$:

$$P(x < X < x + \Delta x) = F(x + \Delta x) - F(x).$$

Складемо відношення цієї ймовірності до довжини інтервалу Δx , тобто вираз для середньої ймовірності на інтервалі, і перейдемо до границі при $\Delta x \rightarrow 0$. У результаті будемо мати

$$\lim_{\Delta x \rightarrow 0} \frac{P(x < X < x + \Delta x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = f'(x). \quad (1)$$

Позначимо $f(x) = F'(x)$. Функцію $f(x)$ називають **щільністю розподілу ймовірностей** неперервної випадкової величини X . Її ще називають **диференціальною функцією розподілу** або **диференціальним законом розподілу**.

Означення. Границя відношення ймовірності попадання неперервної випадкової величини в елементарний інтервал $[x, x + \Delta x]$ до довжини цього інтервалу Δx , коли Δx прямує до нуля, називається **щільністю розподілу ймовірностей випадкової величини у точці x** .

Щільність $f(x)$ є однією з найважливіших розрахункових функцій в теорії ймовірностей. Якщо проінтегрувати рівність $F'(x) = f(x)$ у межах від $-\infty$ до x , то будемо мати

$$F(x) = \int_{-\infty}^x f(u) du. \quad (2)$$

Для ймовірності попадання випадкової величини у заданий інтервал $[x_1, x_2]$ можна записати

$$P(x_1 < X < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(u) du. \quad (3)$$

З означення функції $f(x)$ випливають такі властивості:

Властивість 1. $f(x) \geq 0$.

Ця властивість є наслідком того, що функція розподілу $F(x)$ є неспадкою функцією.

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

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

Властивість 2 має назву “умова нормування функції $f(x)$ ”, і є узагальненням властивості $\sum_{i=1}^n p_i = 1$ для дискретного розподілу

ймовірностей. Геометрично вона означає, що площа фігури, обмеженої графіком функції $f(x)$ і віссю Ox , дорівнює 1.

Властивість 3. Існує хоча б одне значення $\xi \in [x, x + \Delta x]$, для якого виконується рівність

$$P[x \leq X < x + \Delta x] = f(\xi) \Delta x. \quad (4)$$

Ця властивість випливає з теореми “про середнє значення” для інтеграла $\Delta F(x) = \int_x^{x+\Delta x} f(u) du$.

Якщо покласти $\Delta x = dx \rightarrow 0$, то тоді $\xi \rightarrow x$. З рівності (4) отримуємо вираз для ймовірності попадання випадкової величини в нескінченно малий інтервал:

$$P[x \leq X < x + dx] = f(x) dx.$$

Цей вираз називається *елементом ймовірності*.

Приклад 3. Випадкова величина називається рівномірно розподіленою на відрізку $[a, b]$, якщо щільність розподілу $f(x)$ постійна на цьому відрізку, а за його межами дорівнює нулю. Знайти функції $f(x)$ та $F(x)$ для такої випадкової величини.

З умови маємо, що $f(x) = 0$ для $-\infty < x < a$, $b < x < \infty$ і $f(x) = C = \text{const}$ для $a \leq x \leq b$. Величину C знаходимо з умови нормування $f(x)$:

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^a f(x) dx + \int_a^b f(x) dx + \int_b^{\infty} f(x) dx = C(b-a).$$

$$\text{Звідси } C = \frac{1}{b-a}.$$

Отже, для щільності *рівномірного розподілу* ймовірностей маємо:

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

Скориставшись співвідношенням $F(x) = \int_{-\infty}^x f(u) du$, знаходимо функцію рівномірного розподілу (інтегральну функцію рівномірного розподілу)

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

Графіки функцій $f(x)$ і $F(x)$ для рівномірного розподілу показано на рис. 6.

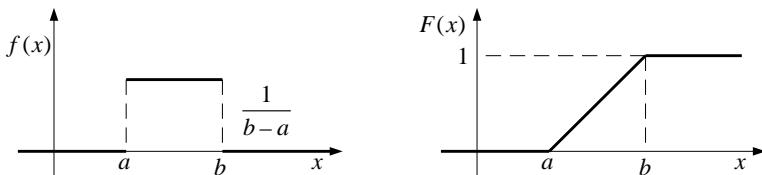


Рис. 6

Приклад 4. Випадкова величина X розподілена за законом

$$f(x) = \begin{cases} a \sin x, & 0 \leq x \leq \pi \\ 0, & x < 0 \text{ або } x > \pi. \end{cases}$$

Необхідно: 1) знайти коефіцієнт a ; 2) побудувати графік щільності розподілу; 3) знайти ймовірність попадання випадкової величини в інтервал $\left(0; \frac{\pi}{4}\right)$.

1) Коефіцієнт a знаходимо з умови нормування:

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^0 0 dx + \int_0^{\pi} a \sin x dx + \int_{\pi}^{+\infty} 0 dx = -a \cos x \Big|_0^{\pi} = 2a = 1. \text{ Отже, } a = \frac{1}{2}.$$

2) Графік щільності розподілу показано на рис. 7.

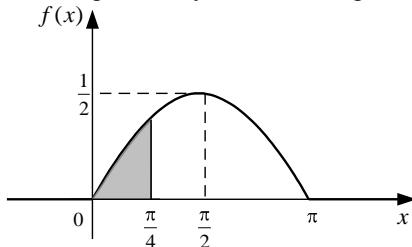


Рис. 7

3) Ймовірність попадання випадкової величини в інтервал $\left(0; \frac{\pi}{4}\right)$:

$$P\left\{0 < X < \frac{\pi}{4}\right\} = \int_0^{\frac{\pi}{4}} \frac{1}{2} \sin x \, dx = -\frac{1}{2} \cos x \Big|_0^{\frac{\pi}{4}} = -\frac{1}{2} \left(\frac{\sqrt{2}}{2} - 1\right) \approx 0,15.$$

Лекція № 9

Числові характеристики випадкових величин

При розв'язуванні багатьох практичних задач теорії ймовірностей буває достатньо знати лише деякі суттєві риси розподілу випадкової величини, наприклад, деяке середнє значення, навколо якого групуються можливі значення випадкової величини; деяке число, яке характеризує рівень "розсіювання" цих значень навколо середнього значення тощо. У таких випадках користуються числовими характеристиками випадкової величини.

Зупинимося на таких важливих поняттях, як *початкові та центральні моменти*. Саме поняття "момент" запозичене з механіки, де воно застосовується при описуванні розподілу маси.

Початковий момент α_r *порядку r* обраховується за формулою

$$\alpha_r = \begin{cases} \sum_i x_i^r p_i & \text{— для дискретної випадкової величини,} \\ \int_{-\infty}^{+\infty} x^r f(x) dx & \text{— для неперервної випадкової величини.} \end{cases} \quad (1)$$

Величина r може приймати значення $0, 1, 2, \dots$. Підсумування у випадку дискретної випадкової величини проводиться за всіма її значеннями. Як видно, початковий момент будь-якого порядку повністю визначається *рядом розподілу* x_i, p_i для дискретної випадкової величини і щільністю розподілу $f(x)$ — для неперервної випадкової величини, якщо сума та інтеграл в (1) збігаються абсолютно. Остання умова може і не виконуватись. Тоді кажуть, що початковий момент не існує.

Особливо важливе значення має початковий момент першого порядку, який називається *математичним сподіванням* і позначається $M X$ або m_x :

$$m_x = M[X] = \alpha_1 = \begin{cases} \sum_i x_i p_i & \text{для дискретної випадкової величини,} \\ \int_{-\infty}^{\infty} x f(x) dx & \text{для неперервної випадкової величини.} \end{cases} \quad (2)$$

Зauważення 1. Очевидно, що будь-який степінь X^r випадкової величини X сам є випадковою величиною. Нехай X – дискретна випадкова величина з рядом розподілу x_i, p_i . Набір пар x_i^r, p_i , взагалі кажучи, не є рядом розподілу дискретної випадкової величини X^r . Наприклад, для парних r різним можливим значенням x_i та $x_j = -x_i$ випадкової величини X відповідає одне і те ж значення $x_i^r = x_j^r$ випадкової величини X^r , а тому ймовірності p_i та p_j повинні додаватись. Проте це ніяк не впливає на суму $\sum_i x_i^r p_i$ з (1), яка є нічим іншим, як математичним сподіванням випадкової величини X^r , тобто сумою добутків можливих значень цієї величини (з врахуванням їх кратності) на ймовірності цих значень. Теж саме можна сказати і відносно інтеграла з (1), якщо замість p_i розглянути елемент ймовірності $f(x) dx$, а суму замінити на інтеграл. Отже, початковий момент порядку r випадкової величини X можна ще означити як математичне сподівання випадкової величини X^r , тобто $\alpha_r X = M[X^r]$.

Математичне сподівання випадкової величини X зв'язане своєрідною залежністю з **середнім арифметичним** значенням цієї величини при великій кількості експериментів. Ця залежність такого ж типу, як і залежність між частотою та ймовірністю, а саме, зі збільшенням числа експериментів середнє арифметичне значення випадкової величини “наближається до її математичного сподівання”.

Дійсно, нехай є ряд розподілу випадкової величини X :

$$P[X = x_i] = p_i \quad i = 1, 2, \dots, n.$$

Нехай проведено N незалежних експериментів, в кожному з яких величина X приймає певне значення. Припустимо, що значення x_1 спостерігалось m_1 разів, значення $x_2 - m_2$ разів, ..., значення $x_n - m_n$ разів. Очевидно, що $\sum_{i=1}^n m_i = N$. Знайдемо середнє арифметичне значення, яке позначимо $M^* X$:

$$M^* X = \frac{x_1 m_1 + x_2 m_2 + \dots + x_n m_n}{N} = \sum_{i=1}^n x_i \frac{m_i}{N}.$$

В останньому виразі $\frac{m_i}{N}$ є відносною частотою події $X = x_i$.

Позначимо її через p_i^* . Тоді $M^* X = \sum_{i=1}^n x_i p_i^*$. Як ми вже знаємо (див. лекцію 6), зі збільшенням числа експериментів N відносна частота $p_i^* = \frac{m_i}{N}$ події $X = x_i$ стає близькою до ймовірності p_i цієї події.

Тому і середнє арифметичне значення випадкової величини зі збільшенням числа N буде наближатись до її математичного сподівання. Для достатньо великих N має місце наближена рівність

$$M^* X \approx M X .$$

Точний математичний зміст розглянутого зв'язку між математичним сподіванням випадкової величини та середнім арифметичним її значень, що спостерігались, а також строге математичне доведення цього зв'язку будуть предметом вивчення в кінці курсу теорії ймовірностей.

Відмітимо, що математичне сподівання існує не для всіх випадкових величин; для його існування необхідно, щоб ряд або інтеграл в (2) збігались, причому абсолютно.

Крім математичного сподівання на практиці застосовуються й інші параметри, що характеризують **положення** випадкової величини на числовій прямій, зокрема **мода** та **медіана**. **Мода** випадкової величини дискретного типу – це таке її значення, яке має найбільшу ймовірність (якщо, звичайно, це значення єдине), тобто x_m – мода, якщо

$$P(X = x_m) = \max_k P(X = x_k) .$$

Будемо позначати моду буквою d_x (рис. 1). Геометричне зображення ряду розподілу дискретної випадкової величини у вигляді рис. 1 ще називають многокутником розподілу.

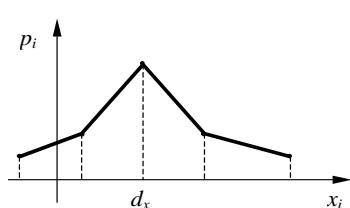


Рис. 1

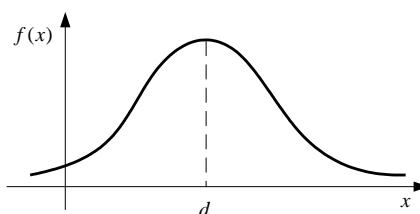


Рис. 2

Модою d_x випадкової величини X неперервного типу називається точка максимуму щільності розподілу $f(x)$ (рис. 2).

Мода може не існувати, мати єдине значення (унімодальний розподіл) або мати цілу множину значень (мультимодальний розподіл).

У загальному випадку мода та математичне сподівання не співпадають. При **симетричному** ($f(x)$ симетрична відносно деякої осі) **унімодальному** розподілі і існуванні математичного сподівання **мода та математичне сподівання співпадають.**

Часто використовується ще одна характеристика положення – **медіана**. Цією характеристикою, як правило, користуються для неперервної випадкової величини.

Медіаною випадкової величини X неперервного типу називається таке її значення h_x , яке задоволяє рівнянню

$$P(X < h_x) = P(X > h_x),$$

тобто однаково ймовірно, прийме випадкова величина менше чи більше значення за h_x . Геометрично медіана – це абсциса, в якій площа, обмежена кривою щільності розподілу, ділиться пополам. Площа кожної половини дорівнює 0,5, оскільки вся площа, обмежена кривою $f(x)$, дорівнює 1. Тому можна ще записати: $F(h_x) = P(X < h_x) = 0,5$, тобто медіана – це точка, в якій функція розподілу $F(x)$ дорівнює 0,5 (рис. 3).

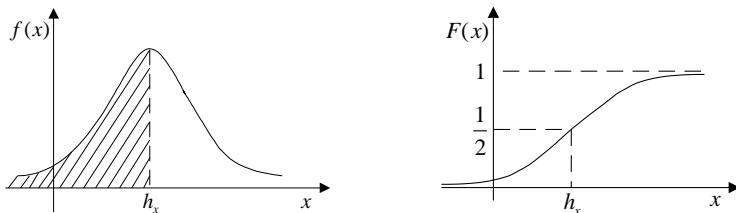


Рис. 3

Іншим типом числових характеристик випадкової величини є **центральні моменти**.

Центральним моментом порядку r називається число β_r , яке обирається за формулою

$$\beta_r = \begin{cases} \sum_i x_i - m_x^r p_i & \text{для дискретної випадкової величини,} \\ \int_{-\infty}^{+\infty} x - m_x^r f(x) dx & \text{для неперервної випадкової величини.} \end{cases}$$

Найбільше застосування має центральний момент другого порядку, який називається **дисперсією**. Для дисперсії вводиться спеціальне позначення $D X$:

$$D X = \beta_2 = \begin{cases} \sum_i x_i - m_x^2 p_i & \text{для дискретної випадкової величини,} \\ \int_{-\infty}^{+\infty} x - m_x^2 f(x) dx & \text{для неперервної випадкової величини.} \end{cases} \quad (3)$$

Дисперсія випадкової величини є мірою “розсіювання” цієї величини навколо її математичного сподівання. Дійсно, якщо випадкова величина постійно дорівнює своєму математичному сподіванню (взагалі кажучи, в цьому випадку величина не є випадковою), то $D X = 0$, оскільки $x_i = m_x$. Якщо ж випадкова величина не завжди дорівнює математичному сподіванню m_x , то вираз $\sum_i x_i - m_x^2 p_i$ відмінний від нуля. Ця відмінність при незмінних p_i , очевидно, тим більша, чим більше відхилення x_i від m_x , тобто, чим далі розміщаються одна від одної точки на числовій прямій, яким відповідають значення x_i та m_x . Отже, зі збільшенням відхилення значень x_i від m_x дисперсія зростає. Дисперсія випадкової величини має розмірність квадрата випадкової величини. Для більшої наочності характеристики розсіювання зручніше користуватись величиною, розмірність якої співпадає з розмірністю випадкової величини. Для цього з дисперсії добувають квадратний корінь. Величину, що отримується, називають **середнім квадратичним** або **стандартним відхиленням** випадкової величини X від її математичного сподівання. Позначають цю величину σ_x :

$$\sigma_x = \sqrt{D X}. \quad (4)$$

Введемо випадкову величину $\overset{\circ}{X} = X - m_x$. Вона називається або **централою випадковою величиною**, або **відхиленням X від математичного сподівання**.

Має місце співвідношення

$$\beta_r X = \alpha_r \begin{bmatrix} \overset{\circ}{X} \end{bmatrix},$$

тобто центральний момент випадкової величини X дорівнює початковому моменту того ж порядку центрованої випадкової величини. Зокрема, $D X = M \begin{bmatrix} \overset{\circ}{X^2} \end{bmatrix}$, тобто дисперсія випадкової величини X дорівнює математичному сподіванню квадрата відповідної центрованої випадкової величини.

Доведено, що вся сукупність початкових (або центральних) моментів повністю визначає функцію розподілу випадкової величини. Тому вся сукупність моментів повністю описує і всі властивості випадкової величини, тоді як кожен окремий момент характеризує яку-небудь одну властивість цієї величини.

Отримаємо формулу, яка зв'язує значення дисперсії і математичного сподівання. Для цього введемо додаткову випадкову величину X^2 . Нехай X – неперервна випадкова величина. Тоді $D X = \int_{-\infty}^{\infty} x - m_x^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 f(x) dx - 2m_x \int_{-\infty}^{\infty} xf(x) dx + m_x^2 \int_{-\infty}^{\infty} f(x) dx = M[X^2] - 2M[X] \cdot M[X] + M^2[X] = M[X^2] - M^2[X]$. Отже,

$$D X = M[X^2] - M^2[X]. \quad (5)$$

Рівність (5) справедлива і для дискретної випадкової величини.

Математичне сподівання та дисперсія – числові характеристики випадкової величини, які найчастіше використовуються. Вони характеризують найважливіші риси розподілу: його положення та рівень розсіювання. Для більш грунтовного описування розподілу застосовуються моменти вищих порядків. Третій центральний момент

$$\beta_3 = \int_{-\infty}^{\infty} x - m_x^3 f(x) dx \quad \text{використовується для характеристики асиметрії}$$

(або “скісності”) розподілу. Величина $\alpha_x = \frac{\beta_3}{\sigma_x^3}$ називається **коєфіцієнтом асиметрії** або **“скісності” розподілу**. Зрозуміло, що при симетричному розподілі $\alpha_x = 0$. На рис. 4 показано два асиметричних розподіли; один з них (крива 1) має додатну асиметрію $\alpha_x > 0$, інший (крива 2) – від’ємну асиметрію $\alpha_x < 0$.

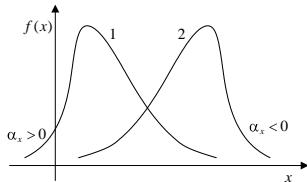


Рис. 4

Коефіцієнт ексцесу обраховують за формулою $e_x = \frac{\beta_4}{\sigma_x^4} - 3$.

Число 3 віднімається від $\frac{\beta_4}{\sigma_x^4}$ тому, що для одного з найважливіших розподілів, а саме нормального розподілу випадкової величини, з яким ми познайомимось у наступній лекції, $\frac{\beta_4}{\sigma_x^4} = 3$.

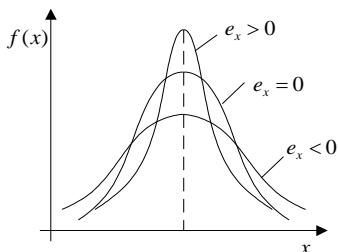


Рис. 5

На рис. 5 показані криві з різним ексцесом.

Отже, для нормального розподілу ексцес дорівнює нулю. Криві розподілу, які більш гостровершинні в порівнянні з кривою нормального розподілу, мають додатній ексцес; більш пологі при вершині криві – від’ємний ексцес. Характеристика e_x використовується, як правило, для симетричних розподілів. На

Математичне сподівання, мода, медіана, початкові та центральні моменти, зокрема, дисперсія, середнє квадратичне відхилення, асиметрія, ексцес – числові характеристики випадкових величин, які найчастіше використовуються. У багатьох практичних задачах повна характеристика випадкової величини – закон розподілу – або невідома, або не може бути знайдена. У таких випадках обмежуються наближеним описуванням випадкової величини за допомогою числових характеристик, кожна з яких виражає яку-небудь характеристичну властивість розподілу.

Приклад. Випадкову величину задано щільністю розподілу

$$f(x) = \begin{cases} \frac{\cos x}{2}, & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 0, & |x| > \frac{\pi}{2}. \end{cases}$$

Знайти дисперсію та середнє квадратичне відхилення.

Розв'язання. Знайдемо математичне сподівання:

$$\begin{aligned}
 m_x &= \int_{-\infty}^{\infty} xf(x) dx = \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x \cos x dx = \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} xd \sin x = \frac{1}{2} \left[x \sin x \right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} - \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin x dx = \\
 &= \frac{1}{2} \left[x \sin x \right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} + \cos x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{1}{2} \left[\frac{\pi}{2} - \frac{\pi}{2} \right] = 0; \\
 D(X) &= \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x^2 \cos x dx = \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x^2 d \sin x = \frac{1}{2} \left[x^2 \sin x \right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} - \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 2x \sin x dx = \\
 &= \frac{\pi^2}{4} - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x \sin x dx = \frac{\pi^2}{4} - 2, \quad \text{оскільки} \quad \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x \sin x dx = - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} xd \cos x = \\
 &= - \left[x \cos x \right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos x dx = \sin x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = 2; \\
 \sigma_x &= \sqrt{\frac{\pi^2}{4} - 2} = \frac{1}{2} \sqrt{\pi^2 - 8}.
 \end{aligned}$$

Лекція № 10

Нормальний закон розподілу випадкової величини

Нормальний закон розподілу (НЗР), який ще називають розподілом Гаусса, відіграє надзвичайно важливу роль у теорії ймовірностей і займає серед інших законів розподілу особливе місце. НЗР – це закон розподілу, який найчастіше зустрічається на практиці. Основна особливість, що виділяє НЗР серед інших законів розподілу, полягає в тому, що він є граничним законом, до якого при виконанні ряду типових умов наближаються інші закони розподілу. Далі ми познайомимося з теоремами, що встановлюють властивість нормального закону як граничного для суми незалежних випадкових доданків (при деяких обмеженнях на ці доданки, які на практиці досить часто виконуються).

Щільність розподілу ймовірностей для НЗР залежить від двох параметрів m , σ і має вигляд

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

Легко перевірити, що умова нормування виконується:

$$\int_{-\infty}^{\infty} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \left| \begin{array}{l} \text{Зробимо заміну:} \\ \frac{x-m}{\sigma} = t; \quad dx = \sigma dt \end{array} \right. * = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = 1. \quad \text{Tут ми скористалися інтегралом Пуассона} \\ \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}. \quad (2)$$

Крива розподілу за нормальним законом має симетричний “горбоподібний” вигляд (крива Гаусса) (рис.1). Цю криву ще називають нормальнюю кривою.

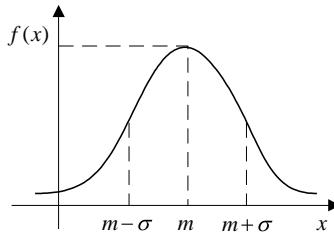


Рис. 1

Максимальна ордината, яка дорівнює $\frac{1}{\sigma\sqrt{2\pi}}$, відповідає точці

$x = m$; в міру віддалення від точки $x = m$ щільність розподілу спадає і при $x \rightarrow \pm\infty$ крива асимптотично наближається до осі абсцис.

Рекомендується показати **самостійно**, що точками перегину кривої Гаусса є точки $x = m \pm \sigma$.

Покажемо, що математичне сподівання для НЗР дорівнює m :

$$M(X) = m_x = \int_{-\infty}^{\infty} xf(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-\frac{(x-m)^2}{2\sigma^2}} dx = \left| \begin{array}{l} \text{Зробимо заміну:} \\ x = \sigma t + m \end{array} \right. * =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma t + m e^{-\frac{t^2}{2}} \sigma dt = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-\frac{t^2}{2}} dt + \frac{m}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = m, \text{ оскільки}$$

перший інтеграл дорівнює нулю, бо підінтегральна функція непарна, а другий інтеграл є інтегралом Пуассона (2). Мода та медіана для НЗР співпадають з математичним сподіванням.

Знайдемо дисперсію для НЗР:

$$\begin{aligned} D(X) &= \sigma_x^2 = \int_{-\infty}^{\infty} |x - m|^2 f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} |x - m|^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \\ &= \left| \begin{array}{l} \text{Зробимо заміну} \\ * \end{array} \right| = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^2 t^2 e^{-\frac{t^2}{2}} \sigma dt = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt = \\ &= \left| \begin{array}{l} \text{Інтегрування частинами:} \\ t = u, \quad dt = du \\ te^{-\frac{t^2}{2}} = dv, \quad v = -e^{-\frac{t^2}{2}} \end{array} \right| = \frac{\sigma^2}{\sqrt{2\pi}} \left[-te^{-\frac{t^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \right] = \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sigma^2. \text{ Тут ми скористалися тим, що } \lim_{t \rightarrow \pm\infty} te^{-\frac{t^2}{2}} = 0. \text{ Реко-} \end{aligned}$$

мендується показати **самостійно**, що це дійсно так.

Отже, дисперсія випадкової величини X , розподіленої за нормальним законом з параметрами m та σ , дорівнює σ^2 , тобто параметр σ є середнім квадратичним відхиленням випадкової величини X :

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

Дослідимо вплив на форму та розміщення кривої Гаусса значень параметрів m та σ . Як відомо, графіки функцій $f(x)$ і $f(x-a)$ мають однакову форму і отримуються один з одного зсувом вздовж осі Ox . Тому зміна параметра m не змінює форму кривої Гаусса, а призводить лише до зсуву цієї кривої вздовж осі Ox : вправо, якщо m збільшується, і вліво, якщо m зменшується (рис.2). Оскільки

$\max f(x) = \frac{1}{\sigma\sqrt{2\pi}}$, то при зростанні параметра σ максимальна

ордината кривої Гаусса спадає, а сама крива стискається до осі Ox . Зменшення параметра σ призводить до того, що крива Гаусса розтягується у додатному напрямку осі Oy (стає більш

“гостровершинною”) (рис.3; $m_x = m = 0$).

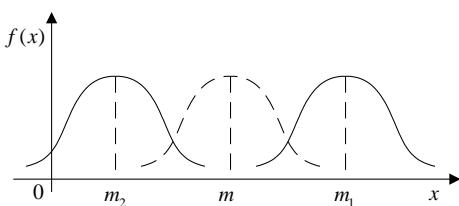


Рис. 2

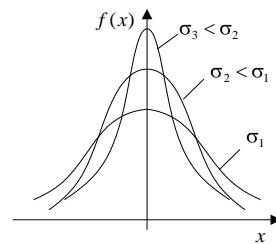


Рис. 3

Отже, зміна параметра σ призводить до зміни форми кривої (змінюється масштаб кривої як у напрямку осі Ox , так і в напрямку осі Oy).

Якщо $m = 0$, $\sigma = 1$, то НЗР називається **нормованим** і позначається $N(0, 1)$. Щільність розподілу (1) при $m = 0$, $\sigma = 1$ називають нормованою і позначають через $\varphi(x)$ (див. лекцію 6):

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (3)$$

Для цієї функції складені таблиці (див. Додаток 2, табл. 2). У загальному випадку нормальній розподіл з параметрами m , σ позначається $N(m, \sigma)$.

Знайдемо для НЗР **центральні моменти** довільного порядку:

$$\begin{aligned} \beta_s X &= \int_{-\infty}^{\infty} (x - m)^s f(x) dx = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} (x - m)^s e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \left| \begin{array}{l} \text{Заміна} \\ * \end{array} \right| = \\ &= \frac{\sigma^s}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^s e^{-\frac{t^2}{2}} dt. \end{aligned}$$

Очевидно, що для непарних s всі моменти $\beta_s = 0$,

оскільки функція під інтегралом непарна, а межі інтегрування симетричні. Для парних s маємо:

$$\begin{aligned} \beta_s X &= \frac{\sigma^s}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t \cdot t^{s-1} e^{-\frac{t^2}{2}} dt = \left| \begin{array}{l} \text{Інтегрування} \\ \text{частинами} \end{array} \right| = \\ &= \frac{\sigma^s}{\sqrt{2\pi}} \left[-t^{s-1} e^{-\frac{t^2}{2}} \Big|_{-\infty}^{\infty} + s-1 \int_{-\infty}^{\infty} t^{s-2} e^{-\frac{t^2}{2}} dt \right] = \end{aligned} \quad (4)$$

$$= s-1 \frac{\sigma^s}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^{s-2} e^{-\frac{t^2}{2}} dt \quad (\text{оскільки } \lim_{t \rightarrow \pm\infty} t^{s-1} \cdot e^{-\frac{t^2}{2}} = 0).$$

З іншого боку, за означенням

$$\beta_{s-2} X = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} x - m)^{s-2} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{\sigma^{s-2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^{s-2} e^{-\frac{t^2}{2}} dt.$$

Порівнюючи останній вираз з (4), отримуємо рекурентну формулу

$$\beta_s X = s-1 \sigma^2 \beta_{s-2} X, \quad (5)$$

яка дає змогу виразити центральні моменти вищих порядків через центральні моменти нижчих порядків. Оскільки $\beta_0 X = 1$, то

$$\beta_2 X = \sigma^2, \beta_4 X = 3\sigma^4, \beta_6 X = 3 \cdot 5\sigma^6, \dots$$

Оскільки $\beta_3 X = 0$, то **асиметрія** НЗР дорівнює нулю. Оскільки

$$\beta_4 X = 3\sigma^4, \text{ то для } \text{експесу} \text{ НЗР маємо } e_x = \frac{\beta_4 X}{\sigma^4} - 3 = 3 - 3 = 0$$

(див. лекцію 9).

Розглянемо **нормований** розподіл $N(0,1)$. Функція розподілу для нього має вигляд:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt. \quad (6)$$

Для функції $\Phi(x)$ складені таблиці (див. Додаток 2, табл. 1).

Через функцію (6) можна виразити **функцію розподілу** для НЗР $N(m, \sigma)$:

$$F(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \left| \begin{array}{l} \text{Заміна} \\ (*) \end{array} \right| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{t^2}{2}} dt = \Phi\left(\frac{x-m}{\sigma}\right).$$

Знайдемо ймовірність попадання випадкової величини X , розподіленої за нормальним законом $N(m, \sigma)$, в інтервал $\alpha < x < \beta$:

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \Phi\left(\frac{\beta-m}{\sigma}\right) - \Phi\left(\frac{\alpha-m}{\sigma}\right). \quad (7)$$

Функція $\Phi(x)$ (6) має всі властивості функції розподілу, які доведені в лекції 7. Нагадаємо лише, що

$$\Phi(-x) = 1 - \Phi(x) \quad (\text{див. лекцію 6}). \quad (8)$$

Знайдемо ймовірність попадання нормально розподіленої

випадкової величини в інтервал, симетричний відносно математичного сподівання:

$$\begin{aligned} P |X - m| < \varepsilon &= P m - \varepsilon < X < m + \varepsilon = \Phi\left(\frac{m + \varepsilon - m}{\sigma}\right) - \Phi\left(\frac{m - \varepsilon - m}{\sigma}\right) = \\ &= \Phi\left(\frac{\varepsilon}{\sigma}\right) - \Phi\left(-\frac{\varepsilon}{\sigma}\right) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right) - 1. \text{ Отже,} \\ P |X - m| < \varepsilon &= 2\Phi\left(\frac{\varepsilon}{\sigma}\right) - 1. \end{aligned} \quad (9)$$

Якщо $\varepsilon = 3\sigma$, то $P |X - m| < 3\sigma = 2\Phi(3) - 1 = 0,997$. Звідси випливає, що подія $m - 3\sigma < X < m + 3\sigma$ є практично вірогідною (достовірною), тобто можливі значення випадкової величини X , розподіленої за законом $N(m, \sigma)$, розміщаються майже вірогідно на відрізку $m - 3\sigma, m + 3\sigma$. Отриманий факт має назву “правило трьох сигм”. Це правило дає можливість за відомим середнім квадратичним відхиленням і математичним сподіванням випадкової величини, розподіленої за законом $N(m, \sigma)$, вказати інтервал її практично можливих значень. З правила трьох сигм випливає також наближений спосіб визначення середнього квадратичного відхилення σ : беруть максимальне і практично можливе відхилення випадкової величини від її середнього значення і ділять це відхилення на 3. Зрозуміло, що цей грубий спосіб може використовуватись лише тоді, коли немає інших, більш точних способів визначення σ .

У деяких підручниках замість функції $\Phi(x)$ (6) використовується функція Лапласа (інтеграл ймовірностей)

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt. \quad (10)$$

Рекомендується показати **самостійно**, що мають місце такі співвідношення:

$$\begin{aligned} \Phi(x) &= \frac{1}{2} + \Phi_0(x); F(x) = \frac{1}{2} + \Phi_0\left(\frac{x-m}{\sigma}\right); \\ P \alpha < X < \beta &= \Phi_0\left(\frac{\beta-m}{\sigma}\right) - \Phi_0\left(\frac{\alpha-m}{\sigma}\right); \\ \Phi_0(-x) &= -\Phi_0(x); P |X - m| < \varepsilon = 2\Phi_0\left(\frac{\varepsilon}{\sigma}\right). \end{aligned} \quad (11)$$

Крім того, $\Phi_0 +\infty = \frac{1}{2}$; $\Phi_0 -\infty = -\frac{1}{2}$; $\Phi_0 0 = 0$.

Розглянемо декілька прикладів.

Приклад 1. Прилад для вимірювання деякої величини працює без систематичних похибок (це означає, що математичне сподівання випадкової величини – похибки вимірювання – дорівнює нулю).

Відомо, що похибка вимірювання розподілена за нормальним законом. Нехай її середнє квадратичне відхилення (ще кажуть, стандартне відхилення) дорівнює $\sigma = 10$ од. Знайти ймовірність того, що вимірювання буде проведено з похибкою, яка не перевищить за абсолютною величиною а) 10 од.; б) 20 од.; в) 30 од.

За формулою (9):

$$\text{а)} P |X| < 10 = 2\Phi\left(\frac{10}{10}\right) - 1 = 2\Phi 1 - 1 = 0,6826, \text{ оскільки за таблицею } \Phi 1 = 0,8413;$$

$$\text{б)} P |X| < 20 = 2\Phi\left(\frac{20}{10}\right) - 1 = 2\Phi 2 - 1 = 0,9544; \Phi 2 = 0,9772;$$

$$\text{в)} P |X| < 30 = 2\Phi\left(\frac{30}{10}\right) - 1 = 2\Phi 3 - 1 = 0,9974; \Phi 3 = 0,9987.$$

Приклад 2. Деталь виготовляється на верстаті. Довжина X цієї деталі є випадковою величиною, розподіленою за нормальним законом із середнім значенням 20 см і середнім квадратичним відхиленням 0,2 см. Знайти ймовірність того, що відхилення довжини деталі в ту чи іншу сторону не перевищить 0,3 см.

Скористаємося формулою (9). У нашому випадку $m = 20$, $\sigma = 0,2$, $\varepsilon = 0,3$. Тому

$$P |X - 20| < 0,3 = 2\Phi\left(\frac{0,3}{0,2}\right) - 1 = 2\Phi 1,5 - 1 = 0,8664, \text{ оскільки за таблицею } \Phi 1,5 = 0,9332.$$

Отже, приблизно 87% всіх деталей, виготовлених на верстаті, будуть мати довжину між 19,7 см і 20,3 см. Інші 13% деталей матимуть більше відхилення довжини від середнього значення.

Приклад 3. В умовах попередньої задачі знайти, яку точність довжини деталі можна гарантувати з ймовірністю 0,95.

$$\text{Розв'яжемо рівняння } P |X - 20| < \varepsilon = 2\Phi\left(\frac{\varepsilon}{0,2}\right) - 1 = 0,95 \text{ або}$$

$\Phi\left(\frac{\varepsilon}{0,2}\right) = 0,975$ відносно ε . За таблицею знаходимо таке значення аргумента $\frac{\varepsilon}{0,2}$, для якого функція Φ приймає значення 0,975. Маємо

$$\frac{\varepsilon}{0,2} = 1,96 \text{ або } \varepsilon = 0,392. \text{ Отже, з ймовірністю, навіть більшою ніж}$$

0,95 можна гарантувати, що відхилення довжини деталі від номіналу не перевищують 0,4 см.

Як вже було сказано, НЗР – це закон розподілу, який найчастіше зустрічається на практиці. Нормальний розподіл мають, наприклад, такі випадкові величини, як відхилення розмірів, ваги предметів від номінальних значень; відхилення при стрільбі; ріст жінок чи чоловіків певного віку, що проживають на даній території; похибки зважувань і, взагалі, похибки вимірювань будь-якого роду в різних сферах людської діяльності. Про умови виникнення нормального розподілу і про причини його широкого поширення у випадкових явищах природи мова йтиме у наступних лекціях.

Лекція № 11

Деякі інші важливі для практики розподіли випадкових величин

Розглянемо ще ряд розподілів, що моделюють ті чи інші випадкові явища. З одним із них (*рівномірним розподілом*) ми вже познайомилися в лекції 8. Випадкова величина розподілена рівномірно, якщо її значення неперервно заповнюють деякий скінчений інтервал, причому щільність розподілу ймовірностей на цьому інтервалі є постійною величиною.

Рекомендується показати **самостійно**, що для математичного сподівання та середнього квадратичного відхилення рівномірно розподіленої на відрізку a, b випадкової величини X справедливі формули

$$M X = \frac{b+a}{2}, \quad \sigma_x = \sqrt{D X} = \frac{b-a}{2\sqrt{3}}, \quad (1)$$

а для ймовірності попадання X в інтервал $\alpha, \beta \subset a, b$ – формула

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \frac{\beta - \alpha}{b - a}. \quad (2)$$

Вираз для функції $F(x)$ наведено в лекції 8.

Випадковою величиною з рівномірним розподілом є, наприклад, положення хвилинної стрілки в момент зупинки годинника. З рівномірним розподілом мають справу у вимірювальній практиці при заокругленні результатів виміру до цілих поділок шкали вимірювального приладу. Похибка в результаті округлення результату виміру до найближчої цілої поділки є випадковою величиною, що рівномірно розподілена на відрізку $-0,5; 0,5$. Очевидно, що рівномірний розподіл мають похибки, що виникають при заокругленні даних в розрахунках.

Іншим прикладом розподілу неперервної випадкової величини є **показниковий розподіл**. Щільність розподілу ймовірностей для нього має вигляд

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{якщо } x \geq 0, \\ 0, & \text{якщо } x < 0. \end{cases} \quad (3)$$

Графік функції $f(x)$ показано на рис.1.

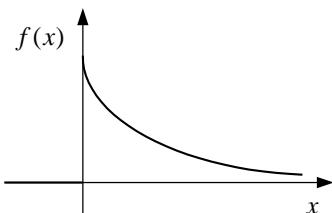


Рис. 1

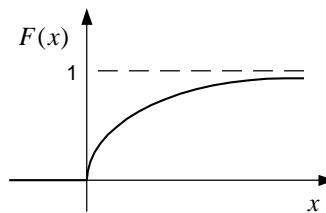


Рис. 2

Показниковий розподіл залежить лише від одного параметра λ , що вигідно відрізняє його від тих розподілів, які залежать від декількох параметрів. Як правило, параметри розподілів невідомі і необхідно знаходити їх оцінки (наближені значення). Зрозуміло, що простіше оцінити один параметр, ніж декілька.

Знайдемо інтегральну функцію показникового розподілу:

$$F(x) = \int_{-\infty}^x f(x) dx = \int_{-\infty}^0 0 \cdot dx + \lambda \int_0^x e^{-\lambda x} dx = 1 - e^{-\lambda x}. \text{ Отже,}$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{якщо } x \geq 0, \\ 0, & \text{якщо } x < 0. \end{cases} \quad (4)$$

Графік функції $F(x)$ показано на рис. 2.

Ймовірність попадання в заданий інтервал визначається формулою

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = e^{-\alpha\lambda} - e^{-\beta\lambda}. \quad (5)$$

Знайдемо математичне сподівання показникового розподілу:

$$M(X) = m_x = \int_0^{\infty} x \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} x e^{-\lambda x} dx = \lambda \left(-\frac{1}{\lambda} x e^{-\lambda x} \Big|_0^{\infty} + \frac{1}{\lambda} \int_0^{\infty} e^{-\lambda x} dx \right) = \frac{1}{\lambda}.$$

Для знаходження дисперсії скористаємося вже відомою нам формулою

$$D(X) = M[X^2] - M^2(X). \quad \text{Тому} \quad D(X) = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Рекомендується провести необхідні викладки **самостійно**.

Отже, для показникового розподілу математичне сподівання дорівнює середньому квадратичному відхиленню

$$m_x = \sigma_x = \frac{1}{\lambda}. \quad (6)$$

Приклад 1. Середній час безперебійної роботи пристрою 15 год. Припустивши, що час безперебійної роботи пристрою має показниковий розподіл, знайти ймовірність того, що пристрій працюватиме без перебою не менше 20 год.

За умовою $M(X) = \frac{1}{\lambda} = 15$. Звідси $\lambda = \frac{1}{15}$. Скористаємося

функцією розподілу (4):

$$P(X \geq 20) = 1 - P(X < 20) = 1 - F(20) = 1 - \left(1 - e^{-\frac{20}{15}}\right) = e^{-\frac{20}{15}} \approx 0,264.$$

До показникового розподілу ми ще повернемося після розгляду розподілу Пуассона.

Як приклади розподілів дискретної випадкової величини розглянемо **біномний розподіл та розподіл Пуассона**.

Як ми вже знаємо (див. лекції 5-8), випадкова величина X має біномний розподіл, якщо її можливими значеннями є числа $0, 1, 2, \dots, n$, а відповідні ймовірності p_m знаходяться за формулою Бернуллі

$$p_m = P(X = m) = C_n^m p^m q^{n-m}, \quad m = 0, 1, 2, \dots, n, \quad 0 < p < 1, \quad q = 1 - p. \quad (7)$$

Умови виникнення біномного розподілу нам вже відомі. Знайдемо основні числові характеристики цього розподілу:

$$M[X] = m_x = \sum_{m=0}^n m C_n^m p^m q^{n-m}; D[X] = M[X^2] - m_x^2, \text{де}$$

$$M[X^2] = \sum_{m=0}^n m^2 C_n^m p^m q^{n-m}. \quad (8)$$

Для цього скористаємося виразом

$$p+q^n = \sum_{m=0}^n C_n^m p^m q^{n-m}. \quad (9)$$

Продиференціюємо ліву та праву частини (9) по p , а потім домножимо на p :

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

$$p+q=1, \text{ то } m_x = np.$$

Продиференціюємо (9) двічі по p і домножимо на p^2 :

$$\begin{aligned} n(n-1)(p+q^{n-2}) \cdot p^2 &= p^2 \cdot \sum_{m=0}^n m(m-1) C_n^m p^{m-2} q^{n-m} = \sum_{m=0}^n m(m-1) C_n^m p^m q^{n-m} = \\ &= \sum_{m=0}^n m^2 C_n^m p^m q^{n-m} - \sum_{m=0}^n m C_n^m p^m q^{n-m} = M[X^2] - m_x. \end{aligned}$$

$$\text{Звідси, } M[X^2] = m_x + n(n-1)p^2 = np + n^2 p^2 - np^2.$$

Для дисперсії маємо:

$$D[X] = np + n^2 p^2 - np^2 - np^2 = np - np^2 = np(1-p) = npq.$$

Отже, для випадкової величини, розподіленої за біномним законом,

$$m_x = np, \quad D_x = npq, \quad \sigma_x = \sqrt{npq}. \quad (10)$$

Формули (10) розкривають ймовірнісний смисл аргумента $x = \frac{m-np}{\sqrt{npq}}$ функцій $\phi(x)$ та $\Phi(x)$, що фігурують в наближених

рівностях Муавра-Лапласа (див. лекцію 6). А саме, x – це відхилення числа “успіхів” m в n незалежних випробуваннях від середнього значення, віднесене до середнього квадратичного (стандартного) відхилення.

Розподіл Пуассона. Нехай дискретна випадкова величина X приймає значення $0, 1, 2, \dots, m, \dots$. Якщо ймовірності, з якими ці значення приймаються, визначаються формулою

$$p_m = P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda} \quad m = 0, 1, 2, \dots , \quad (11)$$

то кажуть, що має місце розподіл Пуассона з параметром $\lambda > 0$.

Задання закону розподілу формулою (11) є коректним, тобто

$$\sum_{m=0}^{\infty} p_m = 1. \text{ Дійсно, } \sum_{m=0}^{\infty} p_m = \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = e^{-\lambda} \cdot e^{\lambda} = 1.$$

Знайдемо математичне сподівання розподілу Пуассона:

$$\begin{aligned} M(X) &= \sum_{m=0}^{\infty} mp_m = \sum_{m=0}^{\infty} m \frac{\lambda^m}{m!} e^{-\lambda} = \sum_{m=1}^{\infty} m \frac{\lambda^m}{m!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{m=1}^{\infty} \frac{m \lambda^{m-1}}{m!} = \\ &= \lambda e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda, \text{ де } k = m-1. \end{aligned}$$

Знаходимо дисперсію $D(X) = M(X^2) - M^2(X)$:

$$\begin{aligned} M(X^2) &= \sum_{m=0}^{\infty} m^2 \frac{\lambda^m}{m!} e^{-\lambda} = \lambda \sum_{m=1}^{\infty} m \frac{\lambda^{m-1}}{(m-1)!} e^{-\lambda} = \lambda \sum_{m=1}^{\infty} [m-1+1] \frac{\lambda^{m-1}}{(m-1)!} e^{-\lambda} = \\ &= \lambda \left[\sum_{m=1}^{\infty} m-1 \frac{\lambda^{m-1}}{(m-1)!} e^{-\lambda} + \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} e^{-\lambda} \right] = \lambda(\lambda+1) = \lambda^2 + \lambda; \end{aligned}$$

$D(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$. Отже, параметр λ розподілу Пуассона є одночасно математичним сподіванням і дисперсією цього розподілу:

$$M(X) = D(X) = \lambda. \quad (12)$$

Ця властивість пуассонівського розподілу часто використовується на практиці для перевірки правдоподібності гіпотези про те, що випадкова величина X розподілена за законом Пуассона. Для цього на основі експериментальних даних визначають статистичні характеристики – математичне сподівання та дисперсію – випадкової величини. Якщо їх значення є близькими, то це свідчить на користь гіпотези про пуассонівський розподіл. Різка відмінність цих характеристик, навпаки, свідчить проти гіпотези. (Аналогічна перевірка можлива і для показникового розподілу, де експериментально визначаються математичне сподівання та середнє квадратичне відхилення).

Знайдемо для випадкової величини X , розподіленої за законом Пуассона, ймовірність того, що вона прийме значення, не менші за деяке число k :

$$P(X \geq k) = p_k + p_{k+1} + \dots = 1 - p_0 + p_1 + \dots + p_{k-1} = 1 - \sum_{m=0}^{k-1} p_m. \quad \text{Зокрема,}$$

ймовірність того, що X прийме ненульове значення, визначається формулою $P(X > 0) = 1 - p_0 = 1 - e^{-\lambda}$.

У лекції 6 було доведено теорему Пуассона, з якої випливає, що розподіл Пуассона є граничним випадком біномного розподілу при необмеженому збільшенні числа випробувань n і необмеженому зменшенні ймовірності p “успіху” в кожному випробуванні, причому так, що величина $\lambda = np$ залишається постійною. Тому біномний розподіл, для якого математичне сподівання мало відрізняється від дисперсії, тобто $np \approx npq$ (див. формули (10)), може бути замінений на розподіл Пуассона. Оскільки ймовірність “успіху” p в цьому випадку є малою, то розподіл Пуассона ще називають законом рідких явищ.

Крім наближення біномного розподілу, закон розподілу Пуассона застосовується при вивчені так званого *простого потоку подій*.

Означення. Простий потік подій – це послідовність подій A_k $k = 1, 2, \dots$, які відбуваються у випадкові моменти часу $t_1, t_2, \dots, t_k, \dots$, тобто на осі абсцис (осі часу) випадково розподіляються точки. Припускається, що випадковий розподіл точок на осі задовільняє третьом умовам:

- 1) Умові *стаціонарності*, яка полягає в тому, що ймовірність попадання того чи іншого числа подій в інтервал часу довжиною t залежить лише від довжини інтервалу t і не залежить від місця знаходження цього інтервалу на осі. Отже, середнє число подій, що відбуваються за одиницю часу, є постійним. Позначимо його через μ і назовемо *цільністю потоку*.
- 2) Умові *відсутності післядії*, яка означає, що для інтервалів часу, які не перетинаються, число подій, що попадають в один з цих інтервалів, не залежить від числа подій, що попадають в інший інтервал.
- 3) Умові *одинарності*, яка означає, що ймовірність попадання в елементарний інтервал Δt двох чи більше подій дуже мала в порівнянні з ймовірністю попадання однієї події, так що цією ймовірністю можна знехвати.

Відомо, що за умов 1), 2), 3) число X точок (подій), які попадають в довільний інтервал довжиною t , розподіляється за законом Пуассона з математичним сподіванням $M X = \lambda = \mu t$. Ймовірність того, що за час t відбудеться рівно m подій, дорівнює

$$P_m t = \frac{\mu t^m}{m!} e^{-\mu t}. \quad (13)$$

Зокрема, ймовірність того, що за час t не відбудеться жодної події, дорівнює $P_0 t = e^{-\mu t}$.

При розумній ідеалізації простим потоком подій можуть бути виклики на телефонній станції; реєстрація радіоактивних частинок; збої на автоматичній лінії; послідовність замовлень на обслуговування тощо.

Приклад 2. На автоматичну телефонну станцію надходить простий потік викликів, інтенсивність (щільність) якого $\mu = 0,8$ (викликів/хв.). Знайти ймовірність того, що за дві хвилини:

- не надійде жодного виклику;
- надійде рівно один виклик;
- надійде хоча б один виклик.

Розв'язання. X – випадкова величина (число викликів за 2 хвилини), яка розподілена за законом Пуассона з параметром $\mu t = 0,8 \cdot 2 = 1,6$. Маємо:

a) $P_0 2 = \frac{\mu t^0}{0!} e^{-\mu t} = e^{-\mu t} = e^{-1,6} \approx 0,202$; б) $P_1 2 = \frac{1,6^1}{1!} e^{-1,6} \approx 0,323$;

в) $P X \geq 1 = 1 - P X = 0 = 1 - P_0 2 \approx 0,798$.

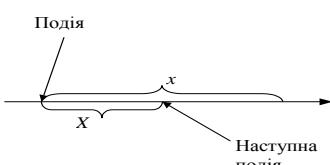


Рис. 3

Нехай випадкова величина X – інтервал часу між двома сусідніми подіями у простому потоці подій (рис. 3). Покажемо, що X має показниковий розподіл з параметром, що дорівнює щільності потоку μ . Знайдемо

функцію розподілу $F(x) = P X < x$, $x > 0$. Для виконання нерівності $X < x$ необхідно, щоб відбулася хоча б одна подія на проміжку часу $t = x$. За формулою (13) $F(x) = P X < x = 1 - P_0 x = 1 - e^{-\mu x}$, $x > 0$. Згідно з (4), маємо показниковий розподіл. Для щільності розподілу ймовірностей можна записати

$$f(x) = F'(x) = \mu e^{-\mu x}, \quad x > 0.$$

Лекція № 12

Спільний розподіл декількох випадкових величин (випадкові вектори)

В теорії ймовірностей та її застосуваннях досить часто зустрічаються задачі, в яких результат експерименту описується не однією випадковою величиною, а двома або більше випадковими величинами, що утворюють систему. Система випадкових величин – це сукупність випадкових величин, які з певних причин розглядаються одночасно. Позначимо, наприклад, через X_1, X_2, \dots, X_n – оцінки (чи бали) випускника школи. Маємо систему n випадкових величин, які в сукупності характеризують успішність.

Нехай є система випадкових величин X_1, X_2, \dots, X_n . Розглядаючи кожен набір значень x_1, x_2, \dots, x_n , які приймаються відповідно випадковими величинами X_1, X_2, \dots, X_n , як координати точки в n -вимірному просторі, будемо називати систему величин X_1, X_2, \dots, X_n випадковою точкою в n -вимірному просторі або **n -вимірним випадковим вектором**. Як і окрема випадкова величина, випадковий вектор є функцією елементарних подій.

Нижче розглядаються лише двовимірні випадкові вектори X, Y . У цьому випадку кожній елементарній події ω ставиться у відповідність пара дійсних чисел x та y , які є значеннями випадкових величин X та Y . Як і для окремо взятої випадкової величини, повною характеристикою **системи** випадкових величин є її **закон розподілу** (закон розподілу випадкового вектора). Цей закон розподілу може мати різну форму (функція розподілу, щільність розподілу, таблиця ймовірностей окремих значень випадкового вектора тощо).

Нехай X та Y – дискретні випадкові величини з можливими значеннями x_i та y_j , $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$. Розподіл системи цих випадкових величин (розподіл дискретного випадкового вектора X, Y) задано, якщо вказано ймовірності $p_{ij} = P(X = x_i, Y = y_j)$ того, що одночасно випадкова величина X приймає значення x_i , а випадкова величина Y – значення y_j , тобто ймовірності добутку подій $X = x_i$ та $Y = y_j$ для довільних i, j . Ці ймовірності заносяться до таблиці.

Табл.1

$x_i \backslash y_j$	x_1	x_2	\dots	x_n
y_1	p_{11}	p_{21}	\dots	p_{n1}
y_2	p_{12}	p_{22}	\dots	p_{n2}
\dots	\dots	\dots	\dots	\dots
y_m	p_{1m}	p_{2m}	\dots	p_{nm}

Таблиця 1 – таблиця розподілу двовимірного дискретного випадкового вектора зі скінченим числом можливих значень (таблиця може мати і нескінчені розміри, якщо кількість можливих значень випадкового вектора є нескінченною (але **зчисленною**)).

Всі можливі події $X = x_i, Y = y_j \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m$ утворюють повну групу несумісних подій. Тому $\sum_{i,j} p_{ij} = 1$. Проводячи підсумовування лише по одному з індексів i або j , отримуємо

$$\begin{aligned} \sum_j p_{ij} &= \sum_j P(X = x_i, Y = y_j) = p_{i1} + p_{i2} + \dots + p_{im} = P(X = x_i) = p_{x_i}; \\ \sum_i p_{ij} &= \sum_i P(X = x_i, Y = y_j) = p_{1j} + p_{2j} + \dots + p_{nj} = P(Y = y_j) = p_{y_j}. \end{aligned} \quad (1)$$

Отже, за відомим розподілом випадкового вектора можна знайти розподілі його компонент.

Введемо інтегральну функцію розподілу (коротко: **функцією розподілу**) двовимірного випадкового вектора або ще кажуть: **функцією спільногорозподілу випадкових величин** X та Y . Під цією функцією розуміють функцію двох дійсних змінних x та y , що визначається як ймовірність одночасного виконання двох нерівностей $X < x$ та $Y < y$:

$$F(x, y) = P(X < x, Y < y). \quad (2)$$

Геометрична інтерпретація: величина $F(x, y)$ виражає ймовірність попадання випадкової точки у внутрішню частину прямого кута з вершиною в точці (x, y) і зі сторонами, паралельними координатним осям (рис.1).

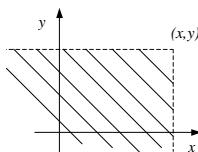


Рис. 1

Функція (2) визначається однаково як для системи дискретних, так і системи неперервних випадкових величин.

Властивості функції розподілу $F(x, y)$:

Властивість 1. $0 \leq F(x, y) \leq 1$.

Ця властивість випливає з самого означення функції як ймовірності.

Властивість 2. $F(x, y)$ – неспадна функція по кожному з аргументів:

$$F(x_2, y) \geq F(x_1, y), \text{ якщо } x_2 > x_1;$$

$$F(x, y_2) \geq F(x, y_1), \text{ якщо } y_2 > y_1.$$

Ця властивість випливає з геометричної інтерпретації функції розподілу. Якщо збільшити x , тобто змістити праву границю кута на рис.1 вправо, то цим самим не можна зменшити ймовірність попадання точки в новий квадрант. Аналогічно, для змінної y .

Властивість 3. $F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0$.

Ці властивості також стають очевидними, якщо звернутись до геометричної інтерпретації функції розподілу.

Властивість 4. $F(+\infty, +\infty) = 1$. Дійсно, квадрант на рис.1 перетворюється у всю площину, попадання в яку є достовірною подією.

Властивість 5. При $y \rightarrow +\infty$ функція розподілу випадкового вектора переходить у функцію розподілу випадкової величини X , а при $x \rightarrow +\infty$ – у функцію розподілу випадкової величини Y :

$$F(x, +\infty) = F_x(x); \quad F(+\infty, y) = F_y(y).$$

Дійсно, при $y \rightarrow +\infty$ квадрант перетворюється у півплощину. При цьому подія $Y < +\infty$ є достовірною, а $F(x, +\infty)$ визначається ймовірністю події $X < x$, тобто $F(x, +\infty) = P(X < x) = F_x(x)$ – функція розподілу складової X випадкового вектора.

Властивість 6. Ймовірність попадання випадкової точки (випадкового вектора) X, Y у довільний прямокутник зі сторонами, паралельними координатним осям (рис.2), знаходиться за формулою

$$P(a \leq X < b, c \leq Y < d) = F(b, d) - F(a, d) - F(b, c) + F(a, c). \quad (3)$$

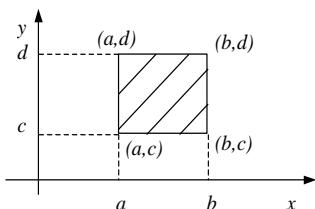


Рис. 2

Доведення. Розглянемо події:

$$A = a \leq X < b, c \leq Y < d,$$

$$B = a \leq X < b, Y < c,$$

$$C = X < a, c \leq Y < d,$$

$$D = X < a, Y < c, E = X < b, Y < d.$$

Тоді $E = A + B + C + D$, а $P(E) =$

$$= P(A) + P(B) + P(C) + P(D), \text{ оскільки події } A, B, C, D \text{ – несумісні.}$$

Звідси $P(A) = P(E) - P(B) - P(C) - P(D)$. Крім того, $P(E) = F(b, d)$;

$$P(B) = F(b, c) - F(a, c); \quad P(C) = F(a, d) - F(a, c); \quad P(D) = F(a, c).$$

Отже, $P(A) = F(b, d) - F(b, c) + F(a, c) - F(a, d) + F(a, c) - F(a, c) = F(b, d) - F(a, d) - F(b, c) + F(a, c)$, що й потрібно було довести.

Розглянемо випадковий вектор X, Y , складові X та Y якого є неперервними випадковими величинами. Якщо існує скінчenna границя

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P(x \leq X < x + \Delta x, y \leq Y < y + \Delta y)}{\Delta x \Delta y} = f(x, y), \quad (4)$$

тобто границя відношення ймовірності попадання випадкової точки (вектора) X, Y в елементарний прямокутник, що прилягає до точки

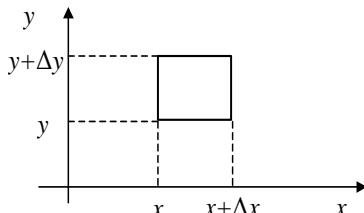


Рис. 3

x, y (рис. 3), до площині цього прямокутника при прямуванні його розмірів до нуля, то така границя називається **щільністю розподілу випадкового вектора** (ще кажуть: **спільною щільністю розподілу**).

Скориставшись формулою (3), перепишемо границю (4) у вигляді

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - F(x + \Delta x, y) + F(x, y)}{\Delta x \Delta y} = f(x, y). \quad (5)$$

Будемо вважати, що функція розподілу $F(x, y)$ є неперервною і диференційованою по кожному з аргументів і існує $\frac{\partial^2 F}{\partial x \partial y}$ (саме ці умови приймаємо за означення **неперервного** випадкового вектора). Тоді формулу (5) можна переписати у вигляді

$$\frac{\partial^2 F}{\partial x \partial y}(x, y) = f(x, y). \quad (6)$$

Отже, щільність розподілу випадкового вектора є другою мішаною частинною похідною від функції розподілу цього вектора.

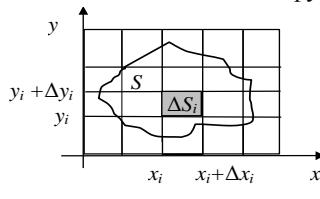


Рис. 4

Розглянемо довільну область S на площині (рис. 4). Розіб'ємо цю область на елементарні підобласті сіткою координатних ліній. Для кожного елементарного прямокутника ΔS_i , який має спільні точки з областю S , згідно з

означенням (4), можна записати $P\{X, Y \in \Delta S_i\} \approx f(x_i, y_i)\Delta x_i \Delta y_i$. Тоді

$$P\{X, Y \in S\} \approx P\{(X, Y) \in \sum_i \Delta S_i\} = \sum_i P\{(X, Y) \in \Delta S_i\} \approx \sum_i f(x_i, y_i) \Delta x_i \Delta y_i. \quad (7)$$

Тут ми скористалися тим, що елементарні прямокутники не перетинаються і що ймовірність суми несумісних подій дорівнює сумі ймовірностей. У граничному випадку, коли $\Delta x_i \rightarrow 0$, $\Delta y_i \rightarrow 0$, з (7) отримуємо

$$P\{X, Y \in S\} = \iint_S f(x, y) dx dy. \quad (8)$$

Якщо скористатись формулою (8) для області на рис.1, то отримаємо ще один зв'язок між функцією розподілу і щільністю розподілу випадкового вектора, а саме

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy. \quad (9)$$

Формули (6) і (9) рівносильні. З (9) і **власливості 5** для функції $F(x, y)$ випливають рівності

$$F_x(x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dx dy; \quad F_y(y) = \int_{-\infty}^{\infty} \int_{-\infty}^y f(x, y) dx dy, \quad (10)$$

а з останніх – рівності

$$f_x(x) = \frac{dF_x(x)}{dx} = \int_{-\infty}^{\infty} f(x, y) dy; \quad f_y(y) = \frac{dF_y(y)}{dy} = \int_{-\infty}^{\infty} f(x, y) dx. \quad (11)$$

Формули (11) є неперервними аналогами формул (1) для дискретного випадкового вектора і встановлюють зв'язок між законом спільного розподілу випадкових величин X та Y і законами розподілу кожної з цих величин окремо.

Відмітимо, що з формули (9) і **власливості 4** для $F(x, y)$ випливає, що

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1. \quad (12)$$

Крім того, з самого означення спільної (двовимірної) щільності маємо, що $f(x, y) \geq 0$.

З формул (1), (10), (11) випливає, що за відомим законом розподілу системи випадкових величин X, Y (таблиця 1, функція $F(x, y)$ або $f(x, y)$) завжди можна знайти закони розподілу окремо взятих величин X та Y , що входять до системи. Виникає запитання: чи не

можна за законами розподілу окремих випадкових величин, які входять до системи, відновити закон розподілу системи? Це можна зробити у випадку незалежних випадкових величин.

Незалежні випадкові величини

Означення. Випадкові величини X та Y називаються **незалежними**, якщо функцію їх спільного розподілу $F(x, y)$ можна подати у вигляді

$$F(x, y) = F(x, +\infty) F(y, +\infty) = F_x(x) F_y(y) \quad (13)$$

для довільних дійсних x, y .

Зauważення. Умову незалежності (13) можна переписати ще так:

$$P(X < x, Y < y) = P(X < x) P(Y < y). \quad (14)$$

Отже, незалежність випадкових величин X та Y означає незалежність зв'язаних з ними випадкових подій $X < x$ та $Y < y$ (див. означення незалежності випадкових подій в лекції 7).

Розглянемо події $a \leq X < b$ і $c \leq Y < d$, де $a < b$ і $c < d$ – довільні числа. Скориставшись умовою (13) в рівності (3), отримуємо

$$\begin{aligned} P(a \leq X < b, c \leq Y < d) &= F_x(b) F_y(d) - F_x(a) F_y(d) - F_x(b) F_y(c) + \\ &+ F_x(a) F_y(c) = [F_x(b) - F_x(a)][F_y(d) - F_y(c)]. \end{aligned}$$

Звідси, врахувавши властивість 3 функції розподілу з лекції 7, можна записати

$$P(a \leq X < b, c \leq Y < d) = P(a \leq X < b) P(c \leq Y < d). \quad (15)$$

Отже, якщо умова (13) виконується, то довільні випадкові події $a \leq X < b$ та $c \leq Y < d$ є незалежними.

Справедливе загальне **тврдження**: з рівності (13) випливає незалежність довільних випадкових подій, пов'язаних з випадковими величинами X та Y , тобто

$$P(X \in B_1, Y \in B_2) = P(X \in B_1) P(Y \in B_2) \quad (16)$$

для довільних борелівських множин B_1 та B_2 (див. приклад 5 лекції 2).

Доведення цього твердження ми наводити не будемо.

Навпаки якщо рівність (16) виконується для довільних борелівських множин, то виконується і (14), а отже, і (13). Тому умова (16) є еквівалентною формою означення незалежності випадкових величин (13).

Означення (13) (або його еквівалентна форма (16)) носить загальний характер. Для неперервних і дискретних випадкових величин зручніше користуватись такими рівносильними умовами незалежності.

Неперервні випадкові величини зі щільностями $f_x(x)$, $f_y(y)$ незалежні тоді і тільки тоді, коли

$$f(x,y) = f_x(x)f_y(y) \quad (17)$$

для всіх дійсних x, y .

Дійсно, диференціюючи рівність (13) по x і по y , дістанемо (17):

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial^2}{\partial x \partial y} (F_x(x)F_y(y)) = \frac{dF_x(x)}{dx} \cdot \frac{dF_y(y)}{dy} = f_x(x)f_y(y).$$

З іншого боку, скориставшись умовою (17) у співвідношенні (9), отримаємо (13):

$$\begin{aligned} F(x,y) &= \int_{-\infty}^x \int_{-\infty}^y f(x,y) dx dy = \int_{-\infty}^x \int_{-\infty}^y f_x(x)f_y(y) dx dy = \int_{-\infty}^x f_x(x) dx \cdot \int_{-\infty}^y f_y(y) dy = \\ &= F_x(x)F_y(y). \end{aligned}$$

Якщо X та Y – дискретні випадкові величини, то їх незалежність рівносильна умові

$$p_{ij} = P\{X = x_i, Y = y_j\} = P\{X = x_i\}P\{Y = y_j\} = p_{x_i}p_{y_j} \quad (18)$$

для всіх i, j .

Приклад. Два студенти незалежно один від одного повинні відповісти на одні і ті ж два запитання. Ймовірність правильної відповіді на кожне з питань для першого студента $p_1 = 0,7$, для другого – $p_2 = 0,4$. Випадкова величина X – число правильних відповідей першого студента, Y – число правильних відповідей другого студента. Побудувати таблицю розподілу $\|p_{ij}\|$ системи випадкових величин X, Y (випадкового вектора).

Розв'язання. Можливі значення випадкових величин X та Y : $x_1 = 0, x_2 = 1, x_3 = 2$; $y_1 = 0, y_2 = 1, y_3 = 2$. Можливі пари значень випадкового вектора X, Y : $(0,0), (0,1), (0,2), (1,0), (1,1), (1,2), (2,0), (2,1), (2,2)$. Знайдемо відповідні цим парам ймовірності. Згідно з умовою задачі події $\{X = x_i\}$ і $\{Y = y_j\}$ слід розглядати як незалежні. Тому $p_{ij} = P\{X = x_i, Y = y_j\} = P\{X = x_i\}P\{Y = y_j\}$.

Знайдемо спочатку ряди розподілів випадкових величин X та Y , скориставшись формулою Бернуллі:

x_i	0	1	2
p_{x_i}	0,09	0,42	0,49

y_j	0	1	2
p_{y_j}	0,36	0,48	0,16

Ймовірності p_{ij} знаходимо за формулою $p_{ij} = P_{x_i} P_{y_j}$:

$$\begin{aligned} p_{11} &= P_{x_1} P_{y_1} = 0,0324; \quad p_{12} = P_{x_1} P_{y_2} = 0,0432; \quad p_{13} = P_{x_1} P_{y_3} = 0,0144; \\ p_{21} &= P_{x_2} P_{y_1} = 0,1512; \quad p_{22} = P_{x_2} P_{y_2} = 0,2016; \quad p_{23} = P_{x_2} P_{y_3} = 0,0672; \\ p_{31} &= P_{x_3} P_{y_1} = 0,1764; \quad p_{32} = P_{x_3} P_{y_2} = 0,2352; \quad p_{33} = P_{x_3} P_{y_3} = 0,0784. \end{aligned}$$

Заносимо знайдені значення p_{ij} до таблиці.

$y_i \backslash x_i$	0	1	2
0	0,0324	0,1512	0,1764
1	0,0432	0,2016	0,2352
2	0,0144	0,0672	0,0784

У розглянутому прикладі закон розподілу випадкового вектора знайдений нами на основі розподілів його компонент, оскільки ці компоненти є незалежними. У загальному випадку, коли X та Y залежні, цього зробити не можна. Необхідно ще знати **умовний** закон розподілу (умовний ряд розподілу, якщо X та Y – дискретні).

Умовний закон розподілу

Для того щоб дати повну характеристику системі випадкових величин, недостатньо знати розподіли кожної з цих величин, необхідно ще мати інформацію про залежність між величинами, що входять до системи. Ця залежність характеризується так званими **умовними законами розподілу**. **Умовним законом розподілу** випадкової величини X , що входить до системи X, Y , називається її закон розподілу, знайдений за умови, що інша випадкова величина (величина Y) приймає певне значення y . Analogічно, для випадкової величини Y . Умовний закон розподілу задають або у вигляді умовної функції розподілу, або умовного ряду розподілу, або умовної щільності розподілу.

Введемо формально умовні щільності $f(x/y)$ і $f(y/x)$ за допомогою формул

$$f(x/y) = \frac{f(x, y)}{f_y(y)}, \quad f(y/x) = \frac{f(x, y)}{f_x(x)} \quad (19)$$

і вияснимо ймовірнісний зміст введених таким чином величин. Для цього перепишемо (19) у вигляді

$$f(x, y) = f_x(x)f(y/x) = f_y(y)f(x/y)$$

і утворимо рівності

$$f(x, y)\Delta x\Delta y = f_x(x)\Delta x \cdot f(y/x)\Delta y = f_y(y)\Delta y \cdot f(x/y)\Delta x. \quad (20)$$

Нехай подія A – попадання випадкової точки X, Y у прямокутник зі сторонами $\Delta x, \Delta y$ (див. рис.3); подія B – попадання у горизонтальну смугу $y \leq Y < y + \Delta y$; подія C – попадання у вертикальну смугу $x \leq X < x + \Delta x$. Тоді $A = BC$. Ми знаємо, що $P(A) = P(BC) = P(B)P(C/B) = P(C)P(B/C)$. Введемо в ці рівності щільності:

$$\begin{aligned} f(x, y)\Delta x\Delta y &\approx f_y(y)\Delta y \cdot P(C/B), \\ f(x, y)\Delta x\Delta y &\approx f_x(x)\Delta x \cdot P(B/C), \end{aligned} \quad (21)$$

оскільки $P(A) \approx f(x, y)\Delta x\Delta y$; $P(B) \approx f_y(y)\Delta y$; $P(C) \approx f_x(x)\Delta x$.

Порівнюючи (20) і (21), тобто $f(x/y)\Delta x \approx P(C/B)$; $f(y/x)\Delta y \approx P(B/C)$, бачимо, що $f(x/y)$ і $f(y/x)$ – умовні щільності (читається: $f(x/y)$ – умовна щільність випадкової величини X за умови, що випадкова величина Y прийняла значення y ; аналогічно, для $f(y/x)$).

Аналогами формул (19) для дискретного випадкового вектора є формули

$$p_{y_j/x_i} = \frac{P_{ij}}{P_{x_i}}, \quad p_{x_i/y_j} = \frac{P_{ij}}{P_{y_j}} \quad (\text{див. формули (1)}). \quad (22)$$

Розшифруємо ці формули:

$$\begin{aligned} P\{Y = y_j / X = x_i\} &= \frac{P\{X = x_i; Y = y_j\}}{P\{X = x_i\}}, \\ P\{X_i = x_i / Y = y_j\} &= \frac{P\{X = x_i; Y = y_j\}}{P\{Y = y_j\}}. \end{aligned} \quad (23)$$

Набір ймовірностей $P\{Y = y_j / X = x_i\} = p_{y_j/x_i}$ є **умовним рядом розподілу** випадкової величини Y за умови $X = x_i$. Аналогічно, $P\{X = x_i / Y = y_j\} = p_{x_i/y_j}$. Легко переконатись, що умовний ряд розподілу має властивості звичайного ряду розподілу:

$$\sum_{j=1}^m p_{y_j/x_i} = \sum_{j=1}^m \frac{p_{ij}}{P_{x_i}} = \frac{1}{P_{x_i}} \sum_{j=1}^m p_{ij} = \frac{1}{P_{x_i}} \cdot p_{x_i} = 1. \quad \text{Аналогічно, } \sum_{i=1}^n p_{x_i/y_j} = 1.$$

Очевидно, що у випадку незалежних випадкових величин X та Y

$$f_x(x/y) = f_x(x), \quad f_y(y/x) = f_y(y), \quad p_{y_j/x_i} = p_{y_j}, \quad p_{x_i/y_j} = p_{x_i}. \quad (24)$$

Лекція № 13

Числові характеристики системи випадкових величин (X,Y).

Умовні числові характеристики. Регресія.

Нормальний розподіл на площині

Як ми вже знаємо (див. лекцію 9), математичне сподівання випадкової величини X^r повністю визначається законом розподілу випадкової величини X . Можна ввести поняття добутку двох випадкових величин XY як випадкової величини з можливими значеннями xy , а також добутку будь-яких степенів цих величин $X^r Y^s$, можливими значеннями якого є числа $x^r y^s$. Можна довести, хоча це і здається очевидним, що математичне сподівання випадкової величини $X^r Y^s$ повністю визначається таблицею спільного розподілу $\|p_{ij}\|$ випадкового вектора X, Y , якщо X та Y – дискретні випадкові величини, або ж щільністю спільного розподілу $f(x, y)$, якщо X та Y – неперервні випадкові величини. Випадкові величини X^r , $X^r Y^s$ – це приклади функцій випадкової величини чи випадкового вектора. Загальне поняття функцій випадкових величин розглянатиметься в наступній лекції.

Перейдемо до числових характеристик системи випадкових величин X, Y , якими є початкові та центральні моменти різних порядків.

Початкові моменти α_{rs} визначаються формулою

$$\alpha_{rs} = M[X^r Y^s] = \begin{cases} \sum_i \sum_j x_i^r y_j^s p_{ij}, \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^r y^s f(x, y) dx dy, \end{cases} \quad (1)$$

а **центральні моменти** β_{rs} – формулою

$$\beta_{rs} = M[\overset{\circ}{X}^r \overset{\circ}{Y}^s] = \begin{cases} \sum_i \sum_j (x_i - m_x)^r (y_j - m_y)^s p_{ij}, \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)^r (y - m_y)^s f(x, y) dx dy. \end{cases} \quad (2)$$

Сума $r + s = r, s = 0, 1, 2, \dots$ називається порядком моменту.

Нагадаємо, що $\overset{\circ}{X} = X - m_x$, $\overset{\circ}{Y} = Y - m_y$ – центровані випадкові

величини; $p_{ij} = P\{X = x_i; Y = y_j\}$ – ймовірність того, що X приймає значення x_i , а Y – значення y_j .

Найширше застосування мають моменти першого та другого порядку.

Момент α_{10} є математичним сподіванням випадкової величини X :

$$\alpha_{10} = \sum_i \sum_j x_i^1 y_j^0 p_{ij} = \sum_i x_i (\sum_j p_{ij}) = \sum_i x_i p_{x_i} = M[X] = m_x$$

$$\left(\alpha_{10} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^1 y^0 f(x, y) dx dy = \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f(x, y) dy \right) dx = \int_{-\infty}^{\infty} x f_x(x) dx = M[X] = m_x \right).$$

Аналогічно, початковий момент α_{01} є математичним сподіванням випадкової величини Y :

$$\alpha_{01} = M[X^0 Y^1] = M[Y] = m_y.$$

Точку m_x, m_y на площині xOy називають центром розсіювання випадкового вектора X, Y .

З (2) випливає, що центральні моменти 1-го порядку дорівнюють нулю: $\beta_{10} = \beta_{01} = 0$, крім того $\beta_{00} = 1$. Рекомендується показати це самостійно.

Для початкових моментів 2-го порядку можна записати:

$$\alpha_{20} = M[X^2 Y^0] = M[X^2]; \quad \alpha_{02} = M[X^0 Y^2] = M[Y^2];$$

$$\alpha_{11} = M[X^1 Y^1] = M[XY].$$

Центральні моменти другого порядку β_{20} і β_{02} є дисперсіями:

$$\begin{aligned} \beta_{20} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)^2 f(x, y) dx dy = \int_{-\infty}^{\infty} (x - m_x)^2 \left(\int_{-\infty}^{\infty} f(x, y) dy \right) dx = \\ &= \int_{-\infty}^{\infty} (x - m_x)^2 f_x(x) dx = D[X] = D_x. \end{aligned}$$

Аналогічно,

$$\beta_{02} = M[\overset{\circ}{X^0} \overset{\circ}{Y^2}] = M[\overset{\circ}{Y^2}] = M[(Y - m_y)^2] = D[Y] = D_y.$$

Особливу роль відіграє мішаний центральний момент другого порядку β_{11} , який називається **коваріацією** (або кореляційним моментом) випадкових величин X та Y і позначається K_{xy} :

$$K_{xy} = \beta_{11} = M[\overset{\circ}{X} \overset{\circ}{Y}] = \begin{cases} \sum_i \sum_j x_i - m_x \quad y_j - m_y \quad p_{ij}, \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x - m_x \quad y - m_y \quad f(x, y) \quad dxdy. \end{cases} \quad (3)$$

Очевидно, що $K_{xy} = K_{yx}$. Крім того, $K_{xx} = M[\overset{\circ}{X} \overset{\circ}{X}] = M[\overset{\circ}{X}^2] = M[X - m_x]^2 = D_x$, тобто дисперсія є коваріацією випадкової величини X самої з собою. Аналогічно, $D_y = K_{yy}$.

Твердження. Для незалежних випадкових величин X та Y коваріація дорівнює нулю.

Доведення:

$$\begin{aligned} K_{xy} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y) f(x, y) dxdy = \begin{cases} \text{якщо } X, Y \text{ – незалежні,} \\ \text{то } f(x, y) = f_x(x)f_y(y) \end{cases} = \\ &= \int_{-\infty}^{\infty} (x - m_x) f_x(x) dx \int_{-\infty}^{\infty} (y - m_y) f_y(y) dy = \left(\int_{-\infty}^{\infty} xf(x) dx - m_x \int_{-\infty}^{\infty} f_x(x) dx \right) \times \\ &\times \left(\int_{-\infty}^{\infty} yf_y(y) dy - m_y \int_{-\infty}^{\infty} f_y(y) dy \right) = (m_x - m_x)(m_y - m_y) = 0. \end{aligned}$$

Доведення твердження для дискретних випадкових величин рекомендується провести **самостійно**.

Означення 1. Якщо коваріація K_{xy} випадкових величин X та Y дорівнює нулю, то ці величини називаються **некорельованими**, у протилежному разі – **корельованими**.

Отже, якщо дві випадкові величини незалежні, то вони некорельовані.

Коваріація K_{xy} характеризує як рівень залежності випадкових величин X та Y , так і їх розсіювання навколо точки m_x, m_y . Для характеристики лише залежності, а не розсіювання, вводять нормований кореляційний момент

$$k_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} = \frac{K_{xy}}{\sqrt{D_x D_y}}, \quad (4)$$

який називають **коєфіцієнтом кореляції** випадкових величин X та Y .

Відмітимо, що коєфіцієнт кореляції у більшості книг прийнято позначати літерою r . Проте в подальшому викладенні матеріалу нам буде зручніше користуватись саме позначенням (4).

До коефіцієнта кореляції ми ще повернемося в наступній лекції.

Вище ми встановили, що з незалежності двох випадкових величин випливає їх некорельованість. Чи справедливе обернене твердження? Взагалі кажучи, ні. З некорельованості не випливає незалежність величин, тобто умова незалежності більш жорстка, ніж умова некорельованості.

Приклад. Розглянемо систему випадкових величин X, Y , рівномірно розподілену в кругу $S = \{(x, y) : x^2 + y^2 \leq r^2\}$ (рис. 1).

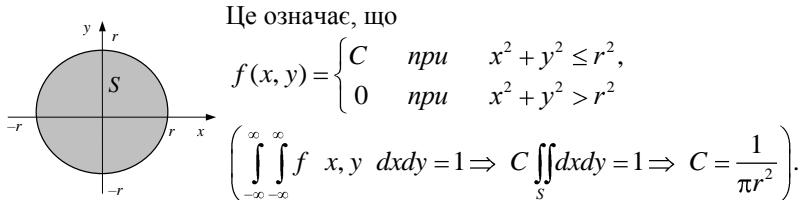


Рис. 1

Легко переконатись, що величини X та Y залежні. Дійсно, якщо величина X приймає значення 0, то величина Y може прийняти довільне значення від $-r$ до $+r$; якщо ж X прийняла значення r , то величина Y може прийняти лише одне значення – нуль. З міркувань симетрії випливає, що $M[X] = M[Y] = m_x = m_y = 0$ (рекомендується підтвердити це розрахунками).

Знайдемо K_{xy} :

$$K_{xy} = \iint_S xy f(x, y) dx dy = \frac{1}{\pi r^2} \iint_S xy dx dy = \frac{1}{\pi r^2} \int_{-r}^r x \left[\int_{-\sqrt{r^2 - x^2}}^{\sqrt{r^2 - x^2}} y dy \right] dx =$$

$$= \frac{1}{\pi r^2} \int_{-r}^r x \left[\frac{y^2}{2} \Big|_{-\sqrt{r^2 - x^2}}^{\sqrt{r^2 - x^2}} \right] dx = 0.$$

Отже, випадкові величини X та Y некорельовані. Однак, вони є залежними.

Вище відмічалось, що до основних числових характеристик системи випадкових величин X, Y належать: математичні сподівання m_x і m_y ; дисперсії D_x і D_y (або середні квадратичні відхилення $\sigma_x = \sqrt{D_x}$ і $\sigma_y = \sqrt{D_y}$); коваріація K_{xy} (або коефіцієнт кореляції k_{xy}).

Поряд з цими характеристиками використовуються так звані **умовні числові характеристики** системи випадкових величин X, Y .

Означення 2. Умовним математичним сподіванням випадкової величини X , яка входить до системи X, Y , називається її математичне сподівання, знайдене за умови, що інша випадкова величина (величина Y) прийняла певне значення y . Analogічне означення можна дати і для умовного математичного сподівання випадкової величини Y .

Умовні математичні сподівання знаходяться на основі умовних законів розподілу (див. лекцію 12):

$$M[X|y_j] = \sum_i x_i p_{x_i/y_j}; \quad M[Y|x_i] = \sum_j y_j p_{y_j/x_i} \quad (5)$$

(читається: $M[X|y_j]$ – умовне математичне сподівання випадкової величини X за умови, що випадкова величина Y прийняла значення y_j ; аналогічно, для $M[Y|x_i]$). Нагадаємо, що $p_{x_i/y_j} = P[X=x_i | Y=y_j]; p_{y_j/x_i} = P[Y=y_j | X=x_i]$ – умовні ймовірності значень випадкових величин X та Y відповідно за умови, що інша величина прийняла певне значення.

Використовуючи умовні щільності розподілу (див. лекцію 12), можна записати вирази для умовних математичних сподівань у випадку неперервних випадкових величин X та Y :

$$M[X|y] = \int_{-\infty}^{\infty} xf(x/y)dx; \quad M[Y|x] = \int_{-\infty}^{\infty} yf(y/x)dy. \quad (6)$$

Означення 3. Умовне математичне сподівання випадкової величини Y при заданому $X=x$, тобто, $M[Y|x]=m_{y|x}$ називається **регресією** Y на x . Analogічно, $M[X|y]=m_{x|y}$ – регресія X на y . Графік залежності $M[Y|x]$ від x називається **лінією регресії** (або кривою регресії) Y на x . Analogічно, графік залежності $M[X|y]$ від y називається лінією регресії X на y (рис. 2).

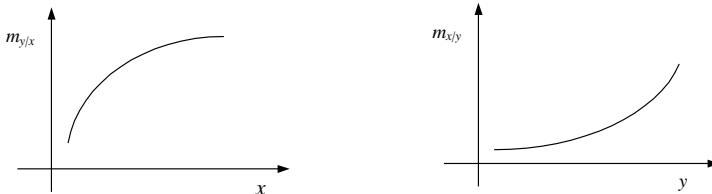


Рис. 2

Зрозуміло, що для незалежних випадкових величин X та Y ліній регресії паралельні до координатних осей, оскільки математичне сподівання кожної з цих величин не залежить від того, яке значення приймає інша.

Лінії регресії можуть бути паралельними до координатних осей і у випадку залежних випадкових величин, зокрема, співпадати з координатними осями, що має місце у розглянутому вище прикладі 1.

Рекомендується **самостійно** знайти за щільністю $f(x, y)$ з прикладу 1 умовні щільності $f(y/x)$, $f(x/y)$ і показати, що умовні математичні сподівання (6) дорівнюють нулю.

Аналогічно умовним математичним сподіванням можна вводити умовні дисперсії $D[Y|x]$ та $D[X|y]$, умовні початкові та центральні моменти тощо.

Раніше ми познайомилися з різними розподілами для окрім взятої випадкової величини. Аналогічно, існують різні розподіли системи двох випадкових величин (двохимірні розподіли), наприклад, рівномірний розподіл в деякій області площини, як у прикладі 1. Серед двовимірних розподілів найчастіше зустрічається двовимірний нормальний розподіл або ще кажуть **нормальний розподіл на площині**.

Означення 4. Кажуть, що система неперервних випадкових величин X, Y розподілена за нормальним законом, якщо спільна щільність розподілу має вигляд

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-k_{xy}^2}} e^{-\frac{1}{2(1-k_{xy}^2)} \left[\frac{(x-m_x)^2}{\sigma_x^2} + \frac{(y-m_y)^2}{\sigma_y^2} - \frac{2k_{xy}(x-m_x)(y-m_y)}{\sigma_x\sigma_y} \right]}. \quad (7)$$

Проінтегрувавши функцію $f(x, y)$ по y в нескінченних межах, отримаємо щільність розподілу випадкової величини X

$$f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{x-m_x^2}{2\sigma_x^2}}. \quad (8)$$

Рекомендується провести викладки **самостійно**. Аналогічно, можна отримати щільність розподілу випадкової величини Y

$$f_y(y) = \int_{-\infty}^{\infty} f(x, y) dx = \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{y-m_y^2}{2\sigma_y^2}}. \quad (9)$$

З наведених формул видно, що параметри $m_x, m_y, \sigma_x, \sigma_y$ в (7) є математичними сподіваннями та середніми квадратичними

відхиленнями випадкових величин X та Y . Можна показати (ми цього робити не будемо), що параметр k_{xy} є коефіцієнтом кореляції випадкових величин X та Y .

Отже, двовимірний нормальний розподіл повністю визначається заданням його характеристик $m_x, m_y, \sigma_x, \sigma_y, k_{xy}$.

Для нормального розподілу системи X, Y терміни “незалежність випадкових величин” та “некорельованість випадкових величин” еквівалентні. Дійсно, з (7) при $k_{xy} = 0$ випливає рівність

$$f(x, y) = f_x(x) f_y(y),$$

де $f_x(x)$, $f_y(y)$ – щільності (8), (9). Отже, якщо дві нормально розподілені випадкові величини некорельовані $k_{xy} = 0$, то вони і незалежні. Справедливість оберненого **тврдження** встановлена нами вище для довільних випадкових величин X та Y .

Знайдемо умовну щільність розподілу $f(y/x) = f(x, y) / f_x(x)$:

$$\begin{aligned} f(y/x) &= \frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-k_{xy}^2}} e^{-\frac{1}{2(1-k_{xy}^2)\left[\frac{x-m_x}{\sigma_x^2} + \frac{y-m_y}{\sigma_y^2} - \frac{2k_{xy}(x-m_x)(y-m_y)}{\sigma_x\sigma_y}\right] + \frac{x-m_x}{2\sigma_x^2}}} = \\ &= \frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-k_{xy}^2}} e^{-\frac{1}{2\sigma_y^2(1-k_{xy}^2)\left[y - \left(m_y + k_{xy}\frac{\sigma_y}{\sigma_x}(x-m_x)\right)\right]^2}}. \end{aligned}$$

Видно, що умовна щільність $f(y/x)$ є щільністю нормального закону з умовним математичним сподіванням $m_{y|x}$ та умовним середнім квадратичним відхиленням $\sigma_{y|x}$:

$$m_{y|x} = m_y + k_{xy} \frac{\sigma_y}{\sigma_x} (x - m_x), \quad \sigma_{y|x} = \sigma_y \sqrt{1 - k_{xy}^2}. \quad (10)$$

Отже, лінія регресії $M(Y|x) = m_{y|x}$ є прямою лінією. Аналогічно, можна знайти лінію регресії $M(X|y) = m_{x|y}$:

$$m_{x|y} = m_x + k_{xy} \frac{\sigma_x}{\sigma_y} (y - m_y), \quad (11)$$

яка також є прямою лінією. Отже, для нормального розподілу системи X, Y регресія завжди лінійна.

Багатовимірні випадкові вектори

Розглянуті у попередній та даній лекціях питання про закони спільного розподілу і числові характеристики двох випадкових величин узагальнюються на системи випадкових величин довільної розмірності $n > 2$:

$F(x_1, x_2, \dots, x_n) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$ – функція розподілу;

$$f(x_1, x_2, \dots, x_n) = \lim_{\substack{\Delta x_i \rightarrow 0 \\ \Delta x_n \rightarrow 0}} \frac{P(x_1 < X_1 < x_1 + \Delta x_1, \dots, x_n < X_n < x_n + \Delta x_n)}{\Delta x_1 \dots \Delta x_n} -$$

щільність спільного розподілу;

$$P(X_1, X_2, \dots, X_n \in S) = \int_S \int \dots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n;$$

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n;$$

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n};$$

$$m_{x_i} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_i f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n - \text{математичне сподівання}$$

випадкової величини X_i ;

$$D_{x_i} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - m_{x_i})^2 f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n - \text{дисперсія}$$

випадкової величини X_i ;

$$K_{x_i x_j} = K_{ij} = M \begin{bmatrix} X_i - m_{x_i} & X_j - m_{x_j} \end{bmatrix} - \text{коваріація випадкових величин } X_i \text{ та } X_j;$$

$$k_{x_i x_j} = k_{ij} = \frac{K_{ij}}{\sigma_{x_i} \sigma_{x_j}} - \text{коєфіцієнт кореляції випадкових величин } X_i \text{ та } X_j.$$

Коваріації K_{ij} утворюють так звану коваріаційну матрицю

$$\|K_{ij}\| = \begin{vmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ K_{21} & K_{22} & \dots & K_{2n} \\ \dots & \dots & \dots & \dots \\ K_{n1} & K_{n2} & \dots & K_{nn} \end{vmatrix}.$$

Слід мати на увазі, що $K_{ii} = D X_i = D_{x_i}$ і $K_{ij} = K_{ji}$, тобто коваріаційна матриця симетрична відносно головної діагоналі.

Часто використовується матриця коефіцієнтів кореляції

$$\|k_{ij}\| = \begin{vmatrix} 1 & k_{12} & \dots & k_{1n} \\ k_{21} & 1 & \dots & k_{2n} \\ \dots & \dots & \dots & \dots \\ k_{n1} & k_{n2} & \dots & 1 \end{vmatrix},$$

що також симетрична відносно головної діагоналі, на якій стоять одиниці.

Випадкові величини X_1, X_2, \dots, X_n називаються незалежними, якщо $F(x_1, x_2, \dots, x_n) = F_{x_1}(x_1) F_{x_2}(x_2) \dots F_{x_n}(x_n)$, де $F_{x_i}(x_i)$ – функція розподілу випадкової величини X_i , $i=1, 2, \dots, n$.

Для неперервних випадкових величин X_1, X_2, \dots, X_n умова незалежності приймає вигляд

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n).$$

Якщо незалежними є дискретні величини, то

$$P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\} = P\{X_1 = x_1\} P\{X_2 = x_2\} \dots P\{X_n = x_n\},$$

для довільних значень x_1, x_2, \dots, x_n .

Лекція № 14

Числові характеристики функцій випадкових величин

Нехай на множині можливих значень x випадкової величини X задано невипадкову функцію $y = \varphi(x)$. В результаті будемо мати випадкову величину Y з можливими значеннями y . Кажуть, що випадкова величина Y є функцією випадкової величини X і пишуть $Y = \varphi(X)$. Аналогічно, можна говорити про випадкову величину Y як функцію декількох випадкових величин (декількох випадкових аргументів): $Y = \varphi(X_1, X_2, \dots, X_n)$.

Знаючи закон розподілу випадкового аргумента X , можна знайти числові характеристики його функції (випадкової величини Y). Наприклад, основні з них, математичне сподівання та дисперсія, знаходяться за допомогою формул:

$$m_y = M Y = \begin{cases} \sum_i \varphi(x_i) p_i, \\ \int_{-\infty}^{\infty} \varphi(x) f(x) dx, \end{cases} \quad D Y = \begin{cases} \sum_i [\varphi(x_i) - m_y]^2 p_i, \\ \int_{-\infty}^{\infty} [\varphi(x) - m_y]^2 f(x) dx. \end{cases} \quad (1)$$

Якщо $Z = \varphi(X, Y)$ – функція двох випадкових аргументів X і Y , то

$$\begin{aligned} m_z &= M Z = \begin{cases} \sum_i \sum_j \varphi(x_i, y_j) p_{ij}, \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y) f(x, y) dxdy, \end{cases} \\ D Z &= \begin{cases} \sum_i \sum_j [\varphi(x_i, y_j) - m_z]^2 p_{ij}, \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\varphi(x, y) - m_z]^2 f(x, y) dxdy, \end{cases} \end{aligned} \quad (2)$$

де $f(x, y)$ – щільність спільного розподілу аргументів X та Y , а $p_{ij} = P(X = x_i, Y = y_j)$.

Формули (2) можна узагальнити на випадок довільної кількості випадкових аргументів.

У багатьох випадках числові характеристики функцій випадкових величин можна знайти без використання законів розподілу аргументів. Для цього достатньо знати лише числові характеристики аргументів. При знаходженні числових характеристик функції використовується ряд **власивостей** цих характеристик, до розгляду яких ми переходимо.

Властивість 1. Математичне сподівання сталої дорівнює цій сталій, тобто $M c = c$, де $c = const$.

Дійсно, стала можна розглядати як випадкову величину, що приймає значення $c = const$ з ймовірністю 1. А тому $M c = c \cdot 1 = c$.

Властивість 2. Невипадковий множник можна виносити за знак математичного сподівання, а саме, $M cX = cM X$.

$$\begin{aligned} \text{Дійсно, } \quad M cX &= \sum_i cx_i p_i = c \sum_i x_i p_i = cM X \quad \text{або} \\ M cX &= \int_{-\infty}^{\infty} cx f(x) dx = c \int_{-\infty}^{\infty} xf(x) dx = cM X. \end{aligned}$$

Властивість 3. Дисперсія сталої дорівнює нулю, тобто $D c = 0$. Ця властивість випливає з самого означення дисперсії. Проте можна скористатись відомою нам формулою

$$D X = M[X^2] - M^2 X . \quad (3)$$

Розглядаючи сталу c як випадкову, тобто $X = c$, будемо мати $D c = M[c^2] - M^2 c = c^2 - c^2 = 0$.

Властивість 4. Невипадковий множник виносиється за знак дисперсії з піднесенням до квадрату, тобто $D cX = c^2 D X$.

Знову скористаємося формулою (3):

$$D cX = M[c^2 X^2] - M^2 cX = c^2 M[X^2] - c^2 M^2 X = c^2 D X .$$

Властивість 5. Математичне сподівання суми двох випадкових величин дорівнює сумі їх математичних сподівань

$$M X_1 + X_2 = M X_1 + M X_2 . \quad (4)$$

Дійсно, для неперервних випадкових величин X_1 і X_2 маємо:

$$\begin{aligned} M X_1 + X_2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 + x_2 f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 f(x_1, x_2) dx_1 dx_2 + \\ &+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2 f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{\infty} x_1 f(x_1) dx_1 + \int_{-\infty}^{\infty} x_2 f(x_2) dx_2 = M X_1 + M X_2 . \end{aligned}$$

Для дискретних випадкових величин X_1 і X_2 доведення властивості 5 пропонується провести **самостійно**.

Використовуючи метод математичної індукції, властивість 5 можна узагальнити на довільне число випадкових величин X_i :

$$M \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n M X_i . \quad (5)$$

Відмітимо, що формулі (4) і (5) справедливі для довільних випадкових величин: залежних і незалежних, корельованих і некорельованих.

Властивість 6. Математичне сподівання добутку двох випадкових величин дорівнює сумі добутку їх математичних сподівань і коваріації

$$M X_1 X_2 = M X_1 M X_2 + K_{x_1 x_2} . \quad (6)$$

Для доведення позначимо, як і раніше, $M X_1 = m_{x_1}$, $M X_2 = m_{x_2}$ та скористаємося означенням коваріації і вже відомими нам

властивостями 2, 5:

$$K_{x_1 x_2} = M \begin{bmatrix} X_1 - m_{x_1} & X_2 - m_{x_2} \end{bmatrix} = M \begin{bmatrix} X_1 X_2 - m_{x_1} X_2 - m_{x_2} X_1 + m_{x_1} m_{x_2} \end{bmatrix} = \\ = M X_1 X_2 - m_{x_1} M X_2 - m_{x_2} M X_1 + m_{x_1} m_{x_2} = M X_1 X_2 - m_{x_1} m_{x_2} .$$

Звідси отримуємо (6).

Відмітимо, що формулою (6) іноді зручно користуватись для знаходження коваріації випадкових величин X_1 і X_2 .

Зauważення. Якщо в (6) покласти $X_1 = X_2 = X$ і врахувати, що

$$K_{x_1 x_2} = K_{xx} = D X , \text{ то отримаємо вже відому нам формулу (3).}$$

З (6) випливає, що для **некорельованих** випадкових величин X_1 і X_2

$$M X_1 X_2 = M X_1 M X_2 . \quad (7)$$

Формула (7) узагальнюється на добуток довільного числа випадкових величин, але в цьому випадку необхідно накладати деякі додаткові умови на ці величини. Проте всі ці умови виконуються, якщо випадкові величини **незалежні**.

Отже, для **незалежних** випадкових величин X_1, X_2, \dots, X_n математичне сподівання добутку дорівнює добутку математичних сподівань

$$M \left[\prod_{i=1}^n X_i \right] = \prod_{i=1}^n M X_i . \quad (8)$$

Властивість 7. Дисперсія суми двох випадкових величин дорівнює сумі їх дисперсій і подвоєної коваріації

$$D X_1 + X_2 = D X_1 + D X_2 + 2K_{x_1 x_2} . \quad (9)$$

$$\begin{aligned} \text{Дійсно, } D X_1 + X_2 &= M \left[X_1 + X_2 - M [X_1 + X_2]^2 \right] = \\ &= M \left[X_1 - M X_1 + X_2 - M X_2 \right]^2 = M \left[X_1 - M X_1 \right]^2 + X_2 - M X_2 \right]^2 + \\ &+ 2 \left[X_1 - M X_1 \right] \left[X_2 - M X_2 \right] = M \left[X_1 - M X_1 \right]^2 + M \left[X_2 - M X_2 \right]^2 + \\ &+ 2M \left[X_1 - M X_1 \right] \left[X_2 - M X_2 \right] = D X_1 + D X_2 + 2K_{x_1 x_2} . \end{aligned}$$

Якщо один з доданків є сталою, наприклад, $X_2 = c = const$, то $D X_1 + c = D X_1$, тобто невипадковий доданок не змінює дисперсії.

Якщо врахувати, що $D X_1 = K_{x_1 x_1}$, $D X_2 = K_{x_2 x_2}$ і $K_{x_1 x_2} = K_{x_2 x_1}$, то

формулу (9) можна записати у вигляді

$$D(X_1 + X_2) = K_{x_1 x_1} + K_{x_2 x_2} + K_{x_1 x_2} + K_{x_2 x_1} = \sum_{i=1}^2 \sum_{j=1}^2 K_{x_i x_j}. \quad (9')$$

Має місце більш загальне твердження, а саме, дисперсія суми довільної кількості випадкових величин дорівнює сумі всіх елементів їх коваріаційної матриці $\|K_{x_i x_j}\|$

$$D\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \sum_{j=1}^n K_{x_i x_j}. \quad (10)$$

Якщо в останній формулі перейти до дисперсій $D(X_i) = K_{x_i x_i}$ і врахувати симетричність матриці $\|K_{x_i x_j}\|$ відносно головної діагоналі, тобто $K_{x_i x_j} = K_{x_j x_i}$, то ця формула набуває вигляду, який узагальнює формулу (9) на довільну кількість доданків

$$D\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n D(X_i) + 2 \sum_{i < j} K_{x_i x_j}. \quad (10')$$

Тут $\sum_{i < j} K_{x_i x_j}$ означає суму тільки тих елементів коваріаційної матриці, які розміщені над головною діагоналлю.

Відмітимо, що доведення формул (10') (або (10)) аналогічне доведенню формули (9). Пропонується провести його **самостійно**.

Знайдемо дисперсію лінійної функції випадкових величин з невипадковими коефіцієнтами c_0, c_1, \dots, c_n :

$$\begin{aligned} D\left[c_0 + \sum_{i=1}^n c_i X_i\right] &= D\left[\sum_{i=1}^n c_i X_i\right] = \sum_{i=1}^n D(c_i X_i) + 2 \sum_{i < j} M\left[c_i X_i - M(c_i X_i) \times \right. \\ &\times \left. c_j X_j - M(c_j X_j)\right] = \sum_{i=1}^n c_i^2 D(X_i) + 2 \sum_{i < j} c_i c_j K_{x_i x_j}. \end{aligned}$$

Отже, справедлива формула

$$D\left[c_0 + \sum_{i=1}^n c_i X_i\right] = \sum_{i=1}^n c_i^2 D(X_i) + 2 \sum_{i < j} c_i c_j K_{x_i x_j}. \quad (11)$$

Зокрема, при $n = 2$

$$D(c_0 + c_1 X_1 + c_2 X_2) = c_1^2 D(X_1) + c_2^2 D(X_2) + 2c_1 c_2 K_{x_1 x_2}. \quad (12)$$

Для **некорельованих** (зокрема, незалежних) випадкових величин наведені вище формули спрощуються. З (10') одержуємо, що дисперсія суми **некорельованих** випадкових величин дорівнює сумі

дисперсій:

$$D \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n D X_i . \quad (13)$$

Формули (11), (12) для **некорельованих** випадкових величин приймають вигляд:

$$D \left[c_o + \sum_{i=1}^n c_i X_i \right] = \sum_{i=1}^n c_i^2 D X_i , \quad (14)$$

$$D c_o + c_1 X_1 + c_2 X_2 = c_1^2 D X_1 + c_2^2 D X_2 .$$

Властивість 8. Дисперсія добутку двох незалежних випадкових величин X_1 та X_2 може бути знайдена за формулою

$$D X_1 X_2 = M[X_1^2] M[X_2^2] - M^2 X_1 M^2 X_2 . \quad (15)$$

Дійсно, якщо X_1 та X_2 – незалежні випадкові величини, то такими будуть і величини X_1^2 та X_2^2 . Тому, скориставшись формулою (3), отримуємо

$$D X_1 X_2 = M[X_1^2 X_2^2] - M^2 X_1 X_2 = M[X_1^2] M[X_2^2] - M^2 X_1 M^2 X_2 .$$

Розглянемо ряд прикладів на використання наведених вище властивостей.

Приклад 1. Розглянемо послідовність незалежних випробувань, в кожному з яких може відбутися чи ні деяка подія A . Нехай ймовірність настання події A в i -му випробуванні дорівнює p_i , $i = 1, 2, \dots, n$. Позначимо через X загальне число наставань події A в усіх n випробуваннях. Потрібно знайти математичне сподівання та дисперсію випадкової величини X .

Введемо для кожного i , $i = 1, 2, \dots, n$ випадкову величину X_i , яка приймає значення 1, якщо подія A в i -му випробуванні відбувається, та 0, якщо – не відбувається. Зрозуміло, що ці значення приймаються з ймовірностями відповідно p_i та $q_i = 1 - p_i$. Для випадкової величини X_i маємо $M X_i = 1 p_i + 0 q_i = p_i$, $D X_i = 1 - p_i^2 p_i + 0 - p_i^2 q_i = p_i q_i$.

Очевидно, що випадкову величину X можна записати у вигляді суми $X = \sum_{i=1}^n X_i$. Оскільки випробування незалежні, то й випадкові величини X_i також незалежні. Скориставшись формулами (5) і (13), отримуємо

$$M X = \sum_{i=1}^n M X_i = \sum_{i=1}^n p_i, \quad D X = \sum_{i=1}^n D X_i = \sum_{i=1}^n p_i q_i .$$

Якщо ймовірність настання події A в кожному випробуванні однакова, тобто $p_i = p$ для всіх $i = 1, 2, \dots, n$ (схема випробувань Бернуллі), то

$$M(X) = \sum_{i=1}^n p_i = np, \quad D(X) = \sum_{i=1}^n p_i q_i = npq.$$

Це вже відомі нам формули для біномного розподілу (див. лекцію 11, формули (10)).

Оскільки формула (5) справедлива для довільних випадкових величин, то вираз для математичного сподівання $M(X) = \sum_i p_i$ не змінюється й тоді, коли розглядається послідовність залежних випробувань, тобто коли ймовірність настання події A в кожному випробуванні залежить від того, відбувалась чи ні ця подія в інших випробуваннях.

Приклад 2. Знайдемо коефіцієнт кореляції випадкових величин X та Y , якщо $Y = aX + b$, де a і b – невипадкові величини, причому $a \neq 0$.

Формула для коефіцієнта кореляції має вигляд $k_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y}$

(див. лекцію 13). Знаходимо коваріацію

$$\begin{aligned} K_{xy} &= M[X - M(X)(Y - M(Y))] = M[X - m_x(aX + b) - am_x - b] = \\ &= M[a(X - m_x)^2] = aD(X) = a\sigma_x^2. \end{aligned}$$

Далі, $D(Y) = \sigma_y^2 = a^2 D(X)$. Звідси $\sigma_y = |a| \sigma_x$. Отже,

$$k_{xy} = \frac{a\sigma_x^2}{\sigma_x |a| \sigma_x} = \frac{a}{|a|} = \begin{cases} -1, & \text{якщо } a < 0, \\ 1, & \text{якщо } a > 0. \end{cases}$$

Маємо такий результат: коефіцієнт кореляції випадкових величин X та Y , зв'язаних лінійною залежністю $Y = aX + b$, дорівнює 1, якщо $a > 0$ та -1 , якщо $a < 0$.

Приклад 3. Покажемо, що для довільних випадкових величин X та Y коефіцієнт кореляції по модулю не перевищує 1; $|k_{xy}| \leq 1$. Для цього знайдемо дисперсію випадкової величини $Z = \sigma_y X \pm \sigma_x Y$, де $\sigma_y = \sqrt{D(Y)}$, $\sigma_x = \sqrt{D(X)}$. З (12) маємо: $D(Z) = \sigma_y^2 D(X) + \sigma_x^2 D(Y) \pm 2\sigma_x \sigma_y K_{xy} = 2\sigma_x^2 \sigma_y^2 \pm 2\sigma_x \sigma_y K_{xy} = 2\sigma_x \sigma_y |\sigma_x \sigma_y \pm K_{xy}|$.

Оскільки дисперсія невід'ємна, то $\sigma_x \sigma_y \pm K_{xy} \geq 0$. Отже, $|K_{xy}| \leq \sigma_x \sigma_y$ і

$$|k_{xy}| = \left| \frac{K_{xy}}{\sigma_x \sigma_y} \right| \leq 1.$$

З розглянутих прикладів випливає, що коефіцієнт кореляції k_{xy} характеризує рівень лінійної залежності між випадковими величинами X і Y . Якщо $k_{xy} = 0$, то такої залежності не існує. Якщо між X і Y існує функціональна залежність $Y = aX + b$, то $k_{xy} = 1$ або $k_{xy} = -1$. При $0 < k_{xy} < 1$ кажуть, що між X і Y має місце додатна кореляція. У цьому випадку при зростанні однієї випадкової величини інша має тенденцію також зростати. При $-1 < k_{xy} < 0$ між випадковими величинами має місце від'ємна кореляція, коли при зростанні однієї з них інша має тенденцію спадати.

Слід відмітити, що при $k_{xy} = 0$ відсутня лінійна залежність між випадковими величинами. При цьому ці випадкові величини можуть бути зв'язані іншим типом залежності (нелінійно!).

На закінчення отримаємо **формулу повного математичного сподівання**, яка є простим наслідком формули повної ймовірності (див. лекцію 5)

$$P(A) = \sum_{i=1}^n P(H_i) P(A|H_i), \quad (16)$$

де H_1, H_2, \dots, H_n – повна група подій. Нехай X – деяка неперервна випадкова величина. Введемо подію $A = \{x < X < x + \Delta x\}$. Тоді, поділивши ліву і праву частини рівняння (16) на Δx , будемо мати

$$\frac{P\{x < X < x + \Delta x\}}{\Delta x} = \sum_{i=1}^n P(H_i) \frac{P\{x < X < x + \Delta x\}/H_i}{\Delta x}. \quad (17)$$

Якщо в (17) перейти до границі при $\Delta x \rightarrow 0$, то, пригадавши означення щільності розподілу (див. лекцію 8), можна записати

$$f(x) = \sum_{i=1}^n P(H_i) f(x)/H_i, \quad (18)$$

де за означенням

$$f(x)/H_i = \lim_{\Delta x \rightarrow 0} \frac{P\{x < X < x + \Delta x\}/H_i}{\Delta x} - \quad (19)$$

умовна щільність розподілу випадкової величини X відносно події H_i . Підставляючи (18) в означення математичного сподівання, отримуємо

$$M X = \int_{-\infty}^{\infty} x f(x) dx = \sum_{i=1}^n P(H_i) \int_{-\infty}^{\infty} x f(x/H_i) dx. \quad (20)$$

Якщо ввести за означенням $M(X|H_i) = \int_{-\infty}^{\infty} x f(x/H_i) dx$ – умовне

математичне сподівання випадкової величини X відносно події H_i , то формула (20) набуде вигляду

$$M X = \sum_{i=1}^n P(H_i) M(X|H_i). \quad (21)$$

Це і є формула повного математичного сподівання.

Нехай тепер X – дискретна випадкова величина з законом розподілу $x_k, p_k = P\{X = x_k\}$, $k = 1, 2, \dots$. Тоді

$$\begin{aligned} M X &= \sum_k x_k P\{X = x_k\} = \sum_k x_k \left[\sum_{i=1}^n P(H_i) P\{X = x_k\}/H_i \right] = \\ &= \sum_{i=1}^n P(H_i) \left[\sum_k x_k P\{X = x_k\}/H_i \right] = \sum_{i=1}^n P(H_i) M(X|H_i), \end{aligned}$$

де $M(X|H_i) = \left[\sum_k x_k P\{X = x_k\}/H_i \right]$ – умовне математичне сподівання випадкової величини X відносно події H_i , $i = 1, 2, \dots, n$. Отже, знову прийшли до формулі (21). Зауважимо, що випадкова величина X у формулі (21) може бути функцією випадкових аргументів: $X = \varphi(X_1, X_2, \dots, X_m)$, де φ – невипадкова функція.

Лекція № 15

Закони розподілу функцій випадкових величин. Задача композиції

Вище ми мали справу з числовими характеристиками функцій випадкової величини $Y = \varphi(X)$ чи випадкового вектора $Y = \varphi(X_1, X_2, \dots, X_n)$. Для знаходження цих характеристик достатньо знати закон розподілу аргументів (випадкової величини X чи випадкового вектора (X_1, X_2, \dots, X_n)). У деяких випадках числові

характеристики функції вдається знайти з використанням лише числових характеристик аргументів.

Іноді потрібно знати закон розподілу функції випадкової величини чи декількох випадкових величин. Наприклад, для знаходження ймовірності попадання функції випадкових величин в ті чи інші області своїх можливих значень.

Почнемо з функції одного випадкового аргументу. Нехай щільність розподілу $f(x)$ неперервної випадкової величини X відома.

Потрібно знайти щільність розподілу $g(y)$ функції $Y = \varphi(X)$.

Припустимо, що функція $y = \varphi(x)$ є строго монотонною і диференційованою на інтервалі a, b можливих значень випадкової величини X (це означає, що існує строго монотонна і диференційовна обернена функція $x = \psi(y)$). Тоді кожному інтервалу $[y, y + \Delta y]$ на осі Oy відповідає повністю визначений інтервал $[l]$ на осі Ox : $l = (x, x + \Delta x)$, якщо функція $y = \varphi(x)$ зростає (рис.1), та $l = (x + \Delta x, x)$, якщо вона спадає (рис.2).

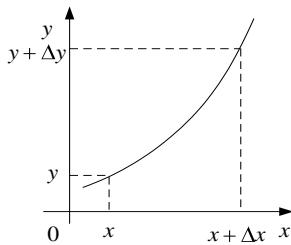


Рис. 1

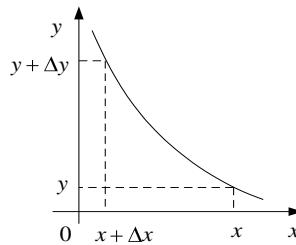


Рис. 2

Об'єднуючи обидва випадки, довжину Δl інтервалу l запишемо у вигляді $\Delta l = |\Delta x|$. Очевидно, що $P[y < Y < y + \Delta y] = P[X \in l]$.

А тому

$$g(y) = \lim_{\Delta y \rightarrow 0} \frac{P[y < Y < y + \Delta y]}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{P[X \in l]}{|\Delta x|} \left| \frac{\Delta x}{\Delta y} \right| = f(x) \left| \frac{dx}{dy} \right| = f(\psi(y)) \left| \frac{d\psi}{dy} \right|.$$

Отже, якщо $y = \varphi(x)$ – строго монотонна і диференційовна функція, то

$$g(y) = f(\psi(y)) \left| \frac{d\psi}{dy} \right|. \quad (1)$$

У випадку, коли $y = \varphi(x)$ є немонотонною функцією, фіксованому значенню y може відповісти декілька значень x_1, x_2, \dots, x_n (рис. 3)

аргумента x , тобто обернена функція $x = \psi(y)$ є неоднозначною: $x_1 = \psi_1(y), x_2 = \psi_2(y), \dots, x_n = \psi_n(y)$. Інтервалу $y, y + \Delta y$ відповідають на осі Ox декілька інтервалів l_i довжиною $|\Delta x_i|$.

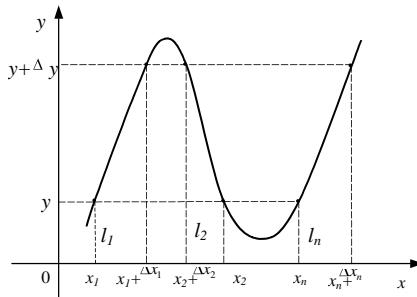


Рис.3

Очевидно, що

$$P\{y < Y < y + \Delta y\} = P\{X \in l_1\} + P\{X \in l_2\} + \dots + P\{X \in l_n\}. \text{ Тому}$$

$$\begin{aligned} g(y) &= \lim_{\Delta y \rightarrow 0} \frac{P\{y < Y < y + \Delta y\}}{\Delta y} = \lim_{\substack{\Delta y \rightarrow 0 \\ \Delta x_i \rightarrow 0}} \frac{\sum_{i=1}^n P\{X \in l_i\}}{\Delta y} = \sum_{i=1}^n \lim_{\substack{\Delta y \rightarrow 0 \\ \Delta x_i \rightarrow 0}} \frac{P\{X \in l_i\}}{|\Delta x_i|} \left| \frac{\Delta x_i}{\Delta y} \right| = \\ &= \sum_{i=1}^n f(x_i) \left| \frac{dx_i}{dy} \right| = \sum_{i=1}^n f(\psi_i(y)) \left| \frac{d\psi_i}{dy} \right|. \text{ Отже,} \\ g(y) &= \sum_{i=1}^n f(\psi_i(y)) \left| \frac{d\psi_i}{dy} \right|. \end{aligned} \quad (2)$$

Розглянемо декілька прикладів на використання формул (1) і (2).

Приклад 1. Знайти щільність розподілу функції $Y = aX + b$, якщо щільність розподілу $f(x)$ випадкової величини X відома, а величини a та b – невипадкові.

Функція $y = ax + b$ – монотонна; $x = \psi(y) = \frac{y - b}{a}$ – обернена функція; $\frac{d\psi}{dy} = \frac{1}{a}$. Тому згідно з формулою (1) $g(y) = f\left(\frac{y - b}{a}\right) \frac{1}{|a|}$.

Отже, лінійне перетворення випадкової величини X призводить до розтягу (або стиску) графіка її щільності розподілу вздовж координатних осей та зміщення цього графіка вздовж осі ординат.

Тип кривої $f(x)$, а отже, і тип закону розподілу випадкової величини X при цьому не змінюється.

Приклад 2. Нехай X – нормальню розподілена випадкова величина з параметрами m_x, σ_x . Показати самостійно, не користуючись результатами прикладу 1, що нормальню розподіленою буде і випадкова величина $Y = aX + b$, де a та b – невипадкові числа, причому $m_y = am_x + b$, $\sigma_y = |a|\sigma_x$.

Приклад 3. Нехай випадкова величина X має нормальню розподіл з параметрами $m_x = 0, \sigma_x = 1$. Знайдемо розподіл випадкової величини $Y = X^2$.

У даному випадку $x_1 = \psi_1(y) = +\sqrt{y}, x_2 = \psi_2(y) = -\sqrt{y}$,

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \text{ Тому за формулою (2)}$$

$$\begin{aligned} g(y) &= f(\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right| + f(-\sqrt{y}) \left| \frac{1}{-2\sqrt{y}} \right| = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{2\sqrt{y}} = \\ &= \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \quad y > 0. \end{aligned}$$

Зauważення 1. Якщо X – дискретна випадкова величина з рядом розподілу x_i, p_i , $i = 1, 2, \dots, n$, то функція $Y = \varphi(X)$ також буде дискретною випадковою величиною з можливими значеннями $y_1 = \varphi(x_1), y_2 = \varphi(x_2), \dots, y_n = \varphi(x_n)$. Складемо таблицю

y_1	y_2	\dots	y_n
p_1	p_2	\dots	p_n

Якщо всі числа y_1, y_2, \dots, y_n – різні, то ця таблиця буде рядом розподілу випадкової величини Y . Якщо ж серед вказаних чисел є однакові, то відповідні стовпчики таблиці потрібно об'єднати в один, а відповідні ймовірності додати.

Розглянемо тепер **таку задачу**: нехай системи випадкових величин (Y_1, Y_2) та (X_1, X_2) зв'язані функціонально

$$Y_1 = \varphi_1(X_1, X_2), \quad Y_2 = \varphi_2(X_1, X_2). \quad (3)$$

Вважаємо, що перетворення (3) є взаємно однозначним і відомі формулі оберненого перетворення

$$X_1 = \psi_1(Y_1, Y_2), \quad X_2 = \psi_2(Y_1, Y_2). \quad (4)$$

Припустимо, що всі функції $\varphi_1, \varphi_2, \psi_1, \psi_2$ є диференційовними. Потрібно, знаючи щільність розподілу $f(x_1, x_2)$ системи (X_1, X_2) ,

знати щільність розподілу $g(y_1, y_2)$ системи (Y_1, Y_2) . При зроблених припущеннях кожній точці $Q(x_1, x_2)$ елементарної області ΔS у площині x_1Ox_2 відповідає лише одна точка $Q_1(y_1, y_2)$ відповідної елементарної області ΔS_1 у площині y_1Oy_2 (рис. 4).

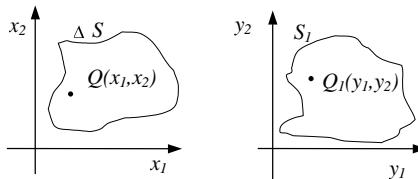


Рис.4

Тому, якщо відбувається одна з подій $X_1, X_2 \in \Delta S$ або

$Y_1, Y_2 \in \Delta S_1$, то обов'язково відбувається й інша. Отже,

$$P(Y_1, Y_2 \in \Delta S_1) = P(X_1, X_2 \in \Delta S).$$

За означенням щільності розподілу

$$\begin{aligned} g(y_1, y_2) &= \lim_{\Delta S_1 \rightarrow Q_1} \frac{P(Y_1, Y_2 \in \Delta S_1)}{\Delta S_1} = \lim_{\substack{\Delta S_1 \rightarrow Q_1 \\ \Delta S \rightarrow Q}} \frac{P(X_1, X_2 \in \Delta S)}{\Delta S} \frac{\Delta S}{\Delta S_1} = \\ &= f(x_1, x_2) \lim_{\substack{\Delta S \rightarrow Q \\ \Delta S_1 \rightarrow Q_1}} \frac{\Delta S}{\Delta S_1} = f(x_1, x_2) \frac{dS}{dS_1} = f(x_1, x_2)|I|, \end{aligned}$$

де $x_1 = \psi_1(y_1, y_2)$, $x_2 = \psi_2(y_1, y_2)$, а величина I , як відомо з математичного аналізу, дорівнює

$$I = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \frac{\partial \psi_1}{\partial y_1} \frac{\partial \psi_2}{\partial y_2} - \frac{\partial \psi_1}{\partial y_2} \frac{\partial \psi_2}{\partial y_1} \text{ — якобіан. Отже,}$$

$$g(y_1, y_2) = f(\psi_1(y_1, y_2); \psi_2(y_1, y_2)) |I|. \quad (5)$$

Формула (5) є узагальненням формули (1).

Приклад 4. Нехай X_1, X_2 — нормальню розподілений випадковий вектор (див. лекцію 13) з параметрами $m_1, m_2, \sigma_1, \sigma_2, k_{12}$. Тоді випадковий вектор Y_1, Y_2 , де $Y_1 = \frac{X_1 - m_1}{\sigma_1}$ і $Y_2 = \frac{X_2 - m_2}{\sigma_2}$, буде

мати нормальній розподіл з параметрами $m_{y_1} = m_{y_2} = 0$, $\sigma_{y_1} = \sigma_{y_2} = 1$ і коефіцієнтом кореляції k_{12} .

Рекомендується показати це **самостійно**, скориставшись формулou (5).

Поставимо тепер **таку задачу**: нехай задано випадкову величину $Y = \phi(X_1, X_2)$, причому існує однозначний розв'язок $X_2 = \psi(Y, X_1)$ і функції ϕ та ψ – диференційовні. Потрібно, знаючи щільність розподілу $f(x_1, x_2)$ системи випадкових величин (X_1, X_2) , знайти щільність розподілу випадкової величини Y .

Для цього розглянемо перетворення

$$Y = \phi(X_1, X_2), \quad Z = X_1, \quad (6)$$

яке є частинним випадком перетворення (3). За умовою існує обернене перетворення

$$X_1 = Z, \quad X_2 = \psi(Y, Z). \quad (7)$$

Застосуємо формулу (5), в позначеннях якої $y_1 = y; y_2 = z$;
 $\psi_1(y_1, y_2) = z = y_2$; $\psi_2(y_1, y_2) = \psi(y, z)$. Тому $g(y, z) = f|z; \psi(y, z)|I|$,

де $I = -\frac{\partial \psi}{\partial y}$.

Оскільки $z = x_1$, то

$$g(y, x_1) = f|x_1; \psi(y, x_1)| \left| \frac{\partial \psi(y, x_1)}{\partial y} \right|. \quad (8)$$

Щоб отримати щільність розподілу g_y у випадкової величини Y , необхідно проінтегрувати вираз (8) по x_1 в нескінченних межах. Отже,

$$g_y(y) = \int_{-\infty}^{\infty} f|x_1; \psi(y, x_1)| \left| \frac{\partial \psi(y, x_1)}{\partial y} \right| dx_1. \quad (9)$$

Зauważення 2. Якщо в (6), (7) покласти $Y = \phi(X_1, X_2) = \phi(X_2)$, $X_2 = \psi(Y)$, то з (9) можна отримати формулу (1). Дійсно,

$$g_y(y) = \int_{-\infty}^{\infty} f|x_1; \psi(y)| \left| \frac{\partial \psi}{\partial y} \right| dx_1 = f_2|\psi(y)| \left| \frac{\partial \psi}{\partial y} \right|, \quad \text{де } f_2|x| \text{ – щільність}$$

розподілу випадкової величини X_2 .

Зауваження 3. Якщо функція $Y = \varphi(X_1, X_2)$ така, що існує однозначний розв'язок $X_1 = \tilde{\psi}(Y, X_2)$, то, покладаючи $Z = X_2$ і проводячи аналогічні міркування, для щільності випадкової величини Y отримуємо

$$g_y(y) = \int_{-\infty}^{\infty} f \left. \frac{\partial \tilde{\psi}(y, x_2)}{\partial y} \right| dx_2. \quad (9')$$

Рекомендується провести викладки **самостійно**.

Розподіл лінійної функції двох випадкових величин. Нехай $Y = a_1 X_1 + a_2 X_2 + b$, де a_1, a_2, b – невипадкові величини.

$$\text{Тоді } x_2 = \psi(y, x_1) = \frac{1}{a_2} y - a_1 x_1 - b; \quad x_1 = \tilde{\psi}(y, x_2) = \frac{1}{a_1} y - a_2 x_2 - b;$$

$$\frac{\partial \psi}{\partial y} = \frac{1}{a_2}; \quad \frac{\partial \tilde{\psi}}{\partial y} = \frac{1}{a_1}.$$

За формулами (9) і (9') маємо

$$g(y) = \int_{-\infty}^{\infty} f \left(x_1; \frac{1}{a_2} y - a_1 x_1 - b \right) \frac{1}{|a_2|} dx_1 = \int_{-\infty}^{\infty} f \left(\frac{1}{a_1} y - a_2 x_2 - b; x_2 \right) \frac{1}{|a_1|} dx_2 \quad (10)$$

або

$$g(y) = \int_{-\infty}^{\infty} f \left(x; \frac{1}{a_2} y - a_1 x - b \right) \frac{1}{|a_2|} dx = \int_{-\infty}^{\infty} f \left(\frac{1}{a_1} y - a_2 x - b; x \right) \frac{1}{|a_1|} dx. \quad (10')$$

Якщо в формулах (10') покласти $b = 0$, $a_1 = a_2 = 1$, то матимемо **розподіл суми двох випадкових величин** $Y = X_1 + X_2$

$$g(y) = \int_{-\infty}^{\infty} f(x; y - x) dx = \int_{-\infty}^{\infty} f(y - x; x) dx. \quad (11)$$

Задача композиції. Потрібно знайти щільність розподілу суми $Y = X_1 + X_2$ незалежних випадкових величин X_1 і X_2 .

За умовою щільність розподілу системи X_1, X_2 дорівнює добутку щільностей розподілу випадкових величин X_1 та X_2 : $f(x_1, x_2) = f_1(x_1)f_2(x_2)$.

Згідно з формулами (11)

$$g(y) = \int_{-\infty}^{\infty} f_1(x)f_2(y - x) dx = \int_{-\infty}^{\infty} f_1(y - x)f_2(x) dx. \quad (12)$$

Розподіл суми незалежних випадкових величин називають **композицією** розподілів цих величин, а задачу знаходження розподілу – **задачою композиції**.

Приклад 5. Нехай X_1, X_2 – незалежні випадкові величини, рівномірно розподілені на відрізку $[0,1]$. Знайти щільність розподілу їх суми $Y = X_1 + X_2$.

За умовою $f_1(x) = f_2(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & x \notin [0,1] \end{cases}$. 3 (12) маємо

$$g(y) = \int_{-\infty}^{\infty} f_1(x)f_2(y-x)dx = \int_0^1 f_2(y-x)dx = \left| \begin{array}{l} \text{заміна} \\ y-x=z \end{array} \right| = - \int_y^{y-1} f_2(z)dz =$$

$$= \int_{y-1}^y f_2(z)dz = \int_{y-1}^y f_2(x)dx.$$

Розглянемо можливі випадки:

1) Якщо $y < 0$, то $g(y) = 0$. 2) Якщо $y > 2$, то $g(y) = 0$. 3) Якщо

$$0 \leq y \leq 1, \text{ то } g(y) = \int_0^y dx = y. 4) \text{ Якщо } 1 \leq y \leq 2, \text{ то } g(y) = \int_{y-1}^1 dx = 2 - y.$$

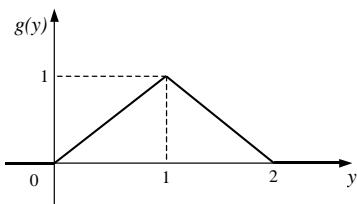


Рис. 5

Об'єднуючи розглянуті випадки, отримуємо

$$g(y) = \begin{cases} 0, & x \notin [0,2] \\ y, & x \in [0,1] \\ 2-y, & x \in (1,2]. \end{cases}$$

Графік щільності розподілу $g(y)$ показано на рис. 5. Одержаній

розподіл називається розподілом Симпсона або трикутним розподілом.

Приклад 6. Нехай $\left(\overset{*}{X}_1, \overset{*}{X}_2 \right)$ – нормальні розподілені випадковий вектор, причому $m_1 = m_2 = 0$, $\sigma_1 = \sigma_2 = 1$. Покажемо, що випадкова величина $Y = a_1 \overset{*}{X}_1 + a_2 \overset{*}{X}_2$, де a_1 і a_2 – довільні невипадкові величини, має нормальній закон розподілу з параметрами $m_y = 0$, $\sigma_y^2 = a_1^2 + a_2^2 + 2k_{12}a_1a_2$, де k_{12} – коефіцієнт кореляції випадкових величин $\overset{*}{X}_1$ та $\overset{*}{X}_2$.

Для цього скористаємося, наприклад, першою з формул (10').

Щільність спільного розподілу системи випадкових величин

$$\left(\begin{array}{c} * \\ X_1, X_2 \end{array} \right) \quad \text{має вигляд} \quad f(x_1, x_2) = \frac{1}{2\pi\sqrt{1-k_{12}^2}} e^{-\frac{1}{2(1-k_{12}^2)}[x_1^2+x_2^2-2k_{12}x_1x_2]}$$

(див. лекцію 13). Тому

$$\begin{aligned} g(y) &= \frac{1}{2\pi\sqrt{1-k_{12}^2}|a_2|} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-k_{12}^2)}[x^2+\frac{1}{a_2^2}(y-a_1x)^2-2k_{12}x\frac{1}{a_2}(y-a_1x)]} dx = \\ &= \frac{1}{2\pi\sqrt{1-k_{12}^2}|a_2|} \int_{-\infty}^{\infty} e^{-Ax^2+2Bx+C} dx, \end{aligned}$$

$$\text{де } A = \frac{1}{2(1-k_{12}^2)a_2^2}(a_1^2 + a_2^2 + 2k_{12}a_1a_2) > 0; \quad B = \frac{1}{2(1-k_{12}^2)a_2^2}(a_1 + k_{12}a_2)y;$$

$$C = -\frac{1}{2(1-k_{12}^2)a_2^2}y^2.$$

Скориставшись інтегралом Пуассона, можна показати, що

$$\int_{-\infty}^{\infty} e^{-Ax^2+2Bx+C} dx = \sqrt{\frac{\pi}{A}} e^{-\frac{B^2+AC}{A}}. \quad \text{А тому } g(y) = \frac{1}{2\pi\sqrt{1-k_{12}^2}|a_2|} \sqrt{\frac{\pi}{A}} e^{-\frac{B^2+AC}{A}}.$$

Підставивши в останню рівність вирази для величин A, B, C і

$$\text{спростивши, отримуємо } g(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{y^2}{2\sigma_y^2}}, \quad \sigma_y^2 = a_1^2 + a_2^2 + 2k_{12}a_1a_2.$$

Отже, випадкова величина $Y = a_1 \overset{*}{X}_1 + a_2 \overset{*}{X}_2$ розподілена за нормальним законом $N(0, \sigma_y)$.

Рекомендується викладки даного прикладу провести більш грунтовно.

Приклад 7. Нехай тепер X_1, X_2 – довільний нормально розподілений випадковий вектор з характеристиками $m_1, m_2, \sigma_1, \sigma_2, k_{12}$. Покажемо, що сума $Y = X_1 + X_2$ буде розподілена за нормально.

Дійсно, цю суму можна подати у вигляді $Y = X_1 + X_2 = \sigma_1 \overset{*}{X}_1 + \sigma_2 \overset{*}{X}_2 + m_1 + m_2$, де $\overset{*}{X}_1 = \frac{X_1 - m_1}{\sigma_1}, \quad \overset{*}{X}_2 = \frac{X_2 - m_2}{\sigma_2}$.

Оскільки сума $\sigma_1^* X_1 + \sigma_2^* X_2^*$ розподілена нормально (див. приклади 4, 6), то такою буде і сума $Y = X_1 + X_2$ (див. приклад 2).

Очевидно, що для нормально розподіленого випадкового вектора X_1, X_2 нормально розподіленою буде будь-яка лінійна комбінація його компонент $Y = a_1 X_1 + a_2 X_2 + b$, причому $m_y = a_1 m_1 + a_2 m_2$, $\sigma_y^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + 2k_{12} \sigma_1 \sigma_2$.

Якщо випадкові величини X_1 і X_2 мають нормальній розподіл і незалежні, то випадковий вектор X_1, X_2 нормальній розподілений. Тому будь-яка лінійна комбінація таких випадкових величин розподілена нормально.

Має місце більш загальне твердження, а саме: **закон розподілу лінійної функції** $Y = \sum_{i=1}^n a_i X_i + b$, де $X_i \quad i=1, 2, \dots, n$ – **незалежні** і **нормально розподілені випадкові величини, є нормальним.**

Доведення цього факту можна провести методом математичної індукції.

Стійкість нормального закону по відношенню до лінійного перетворення широко використовується на практиці. Зокрема, реакція Y багатьох технічних систем на випадкові збурення X_1, X_2, \dots, X_n (які часто є нормальними і незалежними випадковими величинами) вибирається лінійною.

Лекція № 16

Закон великих чисел. Нерівність Чебишова. Теореми Чебишова та їх наслідки

Раніше вже відмічалось, що математичні закони теорії ймовірностей є відображенням в абстрактній формі реальних статистичних закономірностей, що проявляються в **масовості** випадкових явищ. Це можуть бути як закономірності в масі результатів великого числа однорідних випадкових експериментів, так і закономірності, які проявляються внаслідок дії великого числа різного роду випадкових впливів (суть останніх закономірностей полягає в тому, що велике число **випадкових впливів** може породжувати випадкову величину з **конкретним** законом розподілу, що є предметом вивчення наступної лекції). З деякими

закономірностями, що мають місце при великому числі незалежних випадкових експериментів, ми вже знайомі. Це властивість *стійкості* відносної частоти подій (лекція 6) та середнього арифметичного значень випадкової величини, що спостерігались (лекція 9). У загальному випадку властивість стійкості масових випадкових явищ проявляється в тому, що конкретні особливості кожного окремого випадкового явища майже не впливають на середній результат маси таких явищ. Саме в цій стійкості “середніх” і полягає “фізичний” зміст “закону великих чисел”: при дуже великому числі випадкових експериментів середній їх результат практично перестає бути випадковим і може бути передбаченим майже вірогідно. В теорії ймовірностей під законом великих чисел мають на увазі ряд математичних теорем, в кожній з яких за тих чи інших умов встановлюється факт наближення середніх характеристик великого числа випадкових експериментів до деяких невипадкових постійних.

Закон великих чисел – це свого роду зв’язок між теорією ймовірностей як математичною науковою і закономірностями випадкових явищ при масових спостереженнях над ними.

Доведення теорем, що відносяться до “закону великих чисел”, ґрунтуються на нерівності Чебишова, до розгляду якої ми переходимо.

Нехай ϵ випадкова величина X з математичним сподіванням $M X = m$ і дисперсією $D X = D$.

Нерівність Чебишова. Яке б не було додатне число ϵ , ймовірність того, що величина X відхиляється від свого математичного сподівання не менше ніж на ϵ , обмежена зверху величиною $\frac{D}{\epsilon^2}$:

$$P |X - m| \geq \epsilon \leq \frac{D}{\epsilon^2}. \quad (1)$$

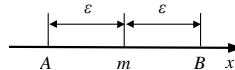
Доведення. 1) Нехай X – дискретна випадкова величина з рядом розподілу $P X = x_i = p_i$. Запишемо вираз для дисперсії: $D X = D = = (x_1 - m)^2 p_1 + (x_2 - m)^2 p_2 + \dots + (x_n - m)^2 p_n + \dots$. Відкінемо у правій частині цього виразу ті доданки, для яких $|x_i - m| < \epsilon$. Нехай це перші k доданків. У доданках, що залишаються, замінимо $(x_j - m)^2$ на ϵ^2 . Очевидно, що тоді можна записати $D \geq \epsilon^2 (p_{k+1} + p_{k+2} + \dots + p_n + \dots)$. З іншого боку, $p_{k+1} + p_{k+2} + \dots + p_n + \dots = P |X - m| \geq \epsilon$, оскільки події $|X - m| < \epsilon$ і $|X - m| \geq \epsilon$ утворюють повну групу несумісних подій і

$$P |X - m| \geq \varepsilon = 1 - P |X - m| < \varepsilon .$$

$$\text{Отже, } D \geq \varepsilon^2 P |X - m| \geq \varepsilon \text{ або } P |X - m| \geq \varepsilon \leq \frac{D}{\varepsilon^2}.$$

2) Нехай тепер X – неперервна випадкова величина зі щільністю розподілу $f(x)$.

Тоді $D = \int_{-\infty}^{\infty} (x - m)^2 f(x) dx$. Проведемо



очевидні оцінки:

$$\begin{aligned} D &\geq \int_{-\infty}^A (x - m)^2 f(x) dx + \int_B^{\infty} (x - m)^2 f(x) dx \geq \varepsilon^2 \left(\int_{-\infty}^A f(x) dx + \int_B^{\infty} f(x) dx \right) = \\ &= \varepsilon^2 \int_{|x-m| \geq \varepsilon} f(x) dx = \varepsilon^2 P |X - m| \geq \varepsilon . \end{aligned}$$

Звідси $P |X - m| \geq \varepsilon \leq \frac{D}{\varepsilon^2}$. Отже, нерівність (1) доведено.

Приклад. Оцінити зверху ймовірність того, що випадкова величина X з математичним сподіванням m і дисперсією σ^2 відхиляться від m не менше, ніж на 3σ .

Розв'язання. Покладаючи в нерівності Чебишова $\varepsilon = 3\sigma$, отримуємо $P |X - m| \geq 3\sigma \leq \frac{D}{9\sigma^2} = \frac{1}{9}$. Нерівність Чебишова дає

верхню оцінку ймовірності даного відхилення для будь-якого закону розподілу випадкової величини X . На практиці в більшості випадків ймовірність того, що X прийме значення поза інтервалом

$m - 3\sigma, m + 3\sigma$, значно менша, ніж $\frac{1}{9}$. Наприклад, для нормального

закону ця ймовірність наближено дорівнює 0,003 (див. лекцію 10). Тому нерівність Чебишова має обмежене практичне застосування, тоді як теоретичне значення цієї нерівності надто велике.

Перейдемо до однієї з форм закону великих чисел – теореми Чебишова. Попередньо розглянемо таку допоміжну задачу. Нехай X – випадкова величина з математичним сподіванням m і дисперсією D . Проводяться n незалежних експериментів над цією величиною і знаходиться середнє арифметичне всіх значень величини X , що спостерігаються у цій серії експериментів. Це середнє значення змінюється від серії до серії з n експериментів, а тому є випадковою величиною. Необхідно знайти математичне сподівання та дисперсію

середнього арифметичного і вияснити, як вони змінюються зі збільшенням n .

Нехай X_1 – значення, яке може прийняти випадкова величина X в першому експерименті, X_2 – в другому експерименті, …, X_n – в n -му експерименті. Очевидно, що сукупність величин X_1, X_2, \dots, X_n є сукупністю n **незалежних** випадкових величин, кожна з яких розподілена за тим же законом, що й сама величина X . Розглянемо середнє арифметичне

$$Y_n = \frac{1}{n} \sum_{i=1}^n X_i. \quad (2)$$

Випадкова величина Y_n є лінійною функцією **незалежних** випадкових величин X_1, X_2, \dots, X_n . Використовуючи результати лекції 14 і враховуючи, що $M X_i = m$, $D X_i = D$, знаходимо

$$M Y_n = \frac{1}{n} M \left(\sum_{i=1}^n X_i \right) = \frac{1}{n} \sum_{i=1}^n M X_i = \frac{1}{n} nm = m,$$

$$D Y_n = \frac{1}{n^2} \left(\sum_{i=1}^n D X_i \right) = \frac{1}{n^2} \sum_{i=1}^n D X_i = \frac{1}{n^2} nD = \frac{D}{n}.$$

Отже, математичне сподівання випадкової величини Y_n (середнього арифметичного) не залежить від числа експериментів n і дорівнює математичному сподіванню випадкової величини X ; дисперсія величини Y_n необмежено спадає зі збільшенням числа експериментів і при достатньо великому n може бути як завгодно малою. Тому середнє арифметичне при великому числі експериментів веде себе майже як не випадкова величина. Теорема Чебишова саме і встановлює в точній кількісній формі цю властивість стійкості середнього арифметичного.

Теорема Чебишова. Нехай випадкова величина X має скінченну дисперсію. Тоді при необмеженому збільшенні числа незалежних експериментів над цією величиною середнє арифметичне значень, що спостерігаються, збігається за ймовірністю до математичного сподівання випадкової величини.

Доведення. Перш за все зупинимося на понятті “**збіжність за ймовірністю**”. Кажуть, що послідовність випадкових величин X_1, X_2, \dots, X_n збігається за ймовірністю до випадкової чи невипадкової величини a , якщо зі збільшенням n ймовірність того, що X_n і a будуть як завгодно близькі, необмежено наближається до 1. Це означає, що для достатньо великого n

$$P |X_n - a| < \varepsilon > 1 - \delta, \quad (3)$$

де ε, δ – як завгодно малі додатні числа.

Теорема Чебишова стверджує, що зі збільшенням n середнє арифметичне $\frac{1}{n} \sum_{i=1}^n X_i$ збігається за ймовірністю до m , тобто для достатньо великого n

$$P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - m \right| < \varepsilon \right\} > 1 - \delta. \quad (4)$$

Застосуємо нерівність Чебишова (1) до випадкової величини $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$: $P |Y_n - M Y_n| \geq \varepsilon \leq \frac{D Y_n}{\varepsilon^2}$. Раніше ми показали, що $M Y_n = m$, а $D Y_n = \frac{D X}{n} = \frac{D}{n}$. Тому $P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - m \right| \geq \varepsilon \right\} \leq \frac{D}{n \varepsilon^2}$.

Яким би малим не було число ε , завжди існує таке N , що для всіх $n > N$ виконується нерівність $\frac{D}{n \varepsilon^2} < \delta$, де δ – будь-яке мале число. Отже, для достатньо великого n $P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - m \right| \geq \varepsilon \right\} < \delta$ або

$$P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - m \right| < \varepsilon \right\} > 1 - \delta, \text{ що й потрібно було довести.}$$

Теорема Чебишова обґрунтуете правило середнього арифметичного, яким широко користуються в практиці вимірювань. При цьому смисл рівностей $M X_i = m$ $i = 1, 2, \dots, n$ полягає в тому, що всі n вимірювань позбавлені систематичної похибки, а вимога скінченної дисперсії $D X_i < \infty$ $i = 1, 2, \dots, n$ означає, що всі вимірювання здійснюються з деякою гарантованою точністю.

Наслідком теореми Чебишова є вже відома нам (див. лекцію 6) **теорема Бернуллі** про зв'язок між відносною частотою подій та її ймовірністю. Нехай проводяться n незалежних експериментів, в кожному з яких може відбутися (з ймовірністю p) чи не відбутися (з ймовірністю $q = 1 - p$) деяка подія A . Позначимо через m загальне число наставань події A в серії з n експериментів. Тоді $p_n^* = \frac{m}{n} -$

відносна частота цієї події в n експериментах. Оскільки m в кожній серії з n експериментів може бути різним, то p_n^* – випадкова величина.

Теорема Бернуллі. При необмеженому збільшенні числа незалежних експериментів n (зі збереженням постійних умов) відносна частота p_n^* події A збігається за ймовірністю до ймовірності p цієї події в окремому експерименті, тобто для достатньо великого n

$$P |p_n^* - p| < \varepsilon > 1 - \delta, \quad (5)$$

де ε, δ – як завгодно малі додатні числа.

Зauważення. Формула (10) в лекції 6

$$\lim_{n \rightarrow \infty} P |p_n^* - p| < \varepsilon = 1, \quad (5')$$

яка отримана як наслідок інтегральної теореми Муавра-Лапласа, є однією з різновидностей запису збіжності за ймовірністю.

Доведення. Як і в прикладі 1 лекції 14, введемо незалежні випадкові величини X_1, X_2, \dots, X_n . Кожна з величин X_i приймає значення 1, якщо подія A в i -му експерименті відбувається, та 0, якщо – не відбувається, причому

$$P X_i = 1 = p; \quad P X_i = 0 = q; \quad M X_i = p; \quad D X_i = pq \quad i = 1, 2, \dots, n.$$

Очевидно, що відносна частота події p_n^* – це середнє арифметичне величин X_1, X_2, \dots, X_n :

$$p_n^* = \frac{1}{n} \sum_{i=1}^n X_i \quad \left(M [p_n^*] = \frac{1}{n} \sum_{i=1}^n M X_i = \frac{np}{n} = p \right).$$

Згідно з теоремою Чебишова (4) отримуємо співвідношення (5). Теорему доведено.

Теорема Чебишова можна узагальнити на випадок, коли характеристики випадкової величини, що спостерігається, змінюються від експерименту до експерименту.

Узагальнена теорема Чебишова. Якщо X_1, X_2, \dots, X_n – незалежні випадкові величини з математичними сподіваннями m_1, m_2, \dots, m_n і дисперсіями $D X_1, D X_2, \dots, D X_n$ і якщо всі дисперсії обмежені зверху одним і тим же числом L : $D X_i < L$ ($i = 1, 2, \dots, n$), то зі зростанням n середнє арифметичне величин X_1, X_2, \dots, X_n збігається за ймовірністю до середнього арифметичного їх математичних сподівань, тобто для достатньо великого n

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n m_i\right| < \varepsilon\right\} > 1 - \delta, \quad (6)$$

де ε, δ – як завгодно малі додатні числа.

Доведення. Розглянемо величину $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$. Знайдемо її математичне сподівання та дисперсію:

$$M Y_n = \frac{1}{n} \sum_{i=1}^n m_i; \quad D Y_n = \frac{1}{n^2} \sum_{i=1}^n D X_i. \quad \text{Застосуємо до величини } Y_n$$

нерівність Чебишова: $P |Y_n - M Y_n| \geq \varepsilon \leq \frac{D Y_n}{\varepsilon^2}$ або

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n m_i\right| \geq \varepsilon\right\} \leq \frac{\sum_{i=1}^n D X_i}{n^2 \varepsilon^2} < \frac{nL}{n^2 \varepsilon^2} = \frac{L}{n \varepsilon^2}. \quad \text{Яким би малим не}$$

було ε , можна вибрати n настільки великим, що буде виконуватись нерівність $\frac{L}{n \varepsilon^2} < \delta$. Тоді $P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n m_i\right| \geq \varepsilon\right\} < \delta$. Перейдемо до протилежної події: $P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n m_i\right| < \varepsilon\right\} > 1 - \delta$. Теорему доведено.

Теорема Бернуллі встановлює стійкість відносної частоти події при постійних умовах проведення серії експериментів. Але ця властивість частоти має місце і в більш загальному випадку, коли умови проведення експерименту змінюються від експерименту до експерименту, так що ймовірності p_1, p_2, \dots, p_n появі події A в цих експериментах різні. Цей факт встановлює теорема Пуассона, яка є наслідком узагальненої теореми Чебишова.

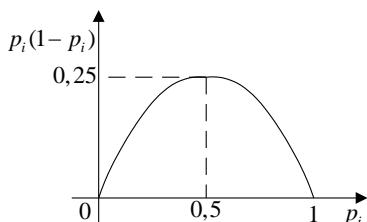
Теорема Пуассона. Якщо проводиться n незалежних експериментів і ймовірність настання події A в i -му експерименті дорівнює p_i , то зі збільшенням n відносна частота події A збігається за ймовірністю до середнього арифметичного ймовірностей p_i , тобто для достатньо великого n

$$P\left\{\left|p_n^* - \frac{1}{n} \sum_{i=1}^n p_i\right| < \varepsilon\right\} > 1 - \delta \quad (7)$$

де ε, δ – як завгодно малі додатні числа.

Доведення. Як і при доведенні теореми Бернуллі, введемо незалежні випадкові величини X_i , $i=1, 2, \dots, n$ (випадкову величину X_i з можливими значеннями 0, 1 ще називають індикатором події A в i -му експерименті):

$P X_i = 1 = p_i$; $P X_i = 0 = q_i = 1 - p_i$; $M X_i = p_i$; $D X_i = p_i q_i$ $i = 1, 2, \dots, n$ (див. приклад 1 лекції 14). Всі дисперсії $D X_i$ обмежені зверху числом 0,25, що є максимальним значенням величини $p_i q_i = p_i(1 - p_i)$.



Для відносної частоти подій

$$p_n^* \text{ маємо } p_n^* = \frac{1}{n} \sum_{i=1}^n X_i;$$

$$M[p_n^*] = \frac{1}{n} \sum_{i=1}^n M X_i = \frac{1}{n} \sum_{i=1}^n p_i.$$

Використовуючи узагальнену теорему Чебишова, отримуємо (7).

Узагальнення закону великих чисел на випадок залежних випадкових величин належить А. Маркову.

Теорема Маркова. Якщо дисперсії залежних випадкових величин в послідовності X_1, X_2, \dots, X_n задовільняють умові

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} D \left[\sum_{i=1}^n X_i \right] = 0, \quad (8)$$

то для всіх $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n M X_i \right| \geq \varepsilon \right\} = 0. \quad (9)$$

Тут ми знову скористалися іншим записом збіжності за ймовірністю.

Доведення. Застосуємо нерівність Чебишова до величини $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$. Як і раніше, $M Y_n = \frac{1}{n} \sum_{i=1}^n M X_i$. Для залежних випадкових величин дисперсія суми не дорівнює сумі дисперсій. Позначимо $D Y_n = \frac{1}{n^2} D \left[\sum_{i=1}^n X_i \right] = D_n$. Отже,

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n M(X_i)\right| \geq \varepsilon\right\} \leq \frac{D_n}{\varepsilon^2}. \text{ Оскільки } \lim_{n \rightarrow \infty} D_n = 0, \text{ то яким би}$$

малим не було $\varepsilon > 0$, можна вибрати n настільки великим, що для довільного малого $\delta > 0$ буде виконуватись нерівність

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n M(X_i)\right| \geq \varepsilon\right\} < \delta, \quad (10)$$

з якої випливає твердження теореми.

Зазначимо, що узагальнена теорема Чебишова випливає з теореми Маркова як окремий випадок. Дійсно, якщо випадкові величини X_1, X_2, \dots, X_n незалежні і їх дисперсії обмежені зверху одним і тим же числом, то умова (8) у теоремі Маркова виконується автоматично.

Закон великих чисел у всіх його формах має широке практичне застосування, в чому ви переконаєтесь при вивченні курсу математичної статистики.

Лекція № 17

Характеристичні функції. Центральна гранична теорема

У попередній лекції ми розглянули різні форми закону великих чисел. Всі вони стверджують одне: факт збіжності за ймовірністю тих чи інших послідовностей випадкових величин до невипадкових чисел. Ні в одній із форм закону великих чисел не йшла мова про закони розподілу випадкових величин. Граничні закони розподілу є предметом іншої групи теорем, що об'єднуються під єдиною назвою центральної граничної теореми.

Всі форми центральної граничної теореми присвячені встановленню умов, за яких виникає нормальний закон розподілу. Оскільки ці умови на практиці досить часто виконуються, нормальний закон є найпоширенішим законом розподілу. Він виникає у всіх випадках, коли випадкову величину можна подати у вигляді суми достатньо великого числа незалежних (або слабкозалежних) елементарних доданків, кожен з яких окремо порівняно мало впливає на суму.

Одна з найбільш загальних форм центральної граничної теореми доведена А.Ляпуновим в 1900 р. Для доведення цієї теореми Ляпунов розробив спеціальний метод характеристичних функцій.

Означення. Характеристичною функцією випадкової величини X називається функція дійсного аргументу t $-\infty < t < \infty$

$$g(t) = M[e^{itX}], \text{ де } i = \sqrt{-1} - \text{ уявна одиниця.}$$

При розв'язуванні багатьох задач теорії ймовірностей буває зручніше користуватись характеристичними функціями випадкових величин, ніж їх законами розподілу. Знаючи закон розподілу випадкової величини, легко знайти її характеристичну функцію.

Нехай X – дискретна випадкова величина з рядом розподілу $P(X = x_k) = p_k$. Тоді

$$g(t) = \sum_k e^{itx_k} p_k. \quad (1)$$

Якщо X – неперервна випадкова величина зі щільністю розподілу $f(x)$, то

$$g(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx. \quad (2)$$

Інтеграл (2) називається перетворенням Фур'є функції $f(x)$. Отже, характеристична функція $g(t)$ неперервної випадкової величини є перетворенням Фур'є щільності розподілу $f(x)$ цієї випадкової величини. Якщо відома характеристична функція $g(t)$, то щільність розподілу $f(x)$ знаходиться за допомогою інтеграла

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} g(t) dt, \quad (3)$$

який називається оберненим перетворенням Фур'є.

З перетворенням Фур'є можна познайомитись у спеціальних курсах математичного аналізу (наприклад, Л.Д.Кудрявцев “Математичний аналіз”, т. 2). Для нас зараз важливим є те, що щільність $f(x)$ і характеристична функція $g(t)$ зв'язані взаємно-однозначно і кожну з цих функцій можна знайти, якщо відома інша.

Приклад 1. Знайти характеристичну функцію випадкової величини, що має нормований нормальній розподіл $N(0,1)$.

Розв'язання. Щільність розподілу за умовою має вигляд

$$f(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \text{ Згідно з (2) одержуємо:}$$

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

Зробимо очевидні перетворення:

$$\begin{aligned} 2itx - x^2 &= 2itx - x^2 + (it)^2 - (it)^2 = -(x^2 - 2itx + (it)^2) + (it)^2 = \\ &= -(x - it)^2 - t^2; \quad (i^2 = -1). \text{ Отже,} \end{aligned}$$

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

Тут ми скористалися тим, що

$$\int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}} dx = \left| \frac{x-it}{dx} = dz \right| = \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}. \quad \text{Отже, для випадкової}$$

величини з розподілом $N(0,1)$ характеристична функція має вигляд

$$g(t) = e^{-\frac{t^2}{2}}.$$

Якщо підставити $g(t) = e^{-\frac{t^2}{2}}$ у формулу (3) і провести аналогічні викладки, то отримаємо функцію $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ – щільність розподілу $N(0,1)$. Рекомендується провести ці викладки **самостійно**.

Основні властивості характеристичних функцій

1. Характеристична функція невипадкової величини a дорівнює $g(t) = e^{ita}$. Дійсно, $M[e^{ita}] = e^{ita}$.

2. Якщо випадкові величини X і Y зв'язані співвідношенням $Y = aX$, де a – невипадковий множник, то їх характеристичні функції зв'язані співвідношенням $g_y(t) = g_x(at)$.

Дійсно, $g_y(t) = M[e^{itY}] = M[e^{itaX}] = M[e^{i(at)X}] = g_x(at)$.

3. Якщо існує початковий момент випадкової величини порядку k , то він зв'язаний з характеристичною функцією співвідношенням

$$\alpha_k = M[X^k] = g^{(k)}(0)(i)^{-k}, \quad g^{(k)}(0) = \frac{d^k g(t)}{dt^k} \Big|_{t=0}.$$

$$\begin{aligned} \text{Дійсно, } g^{(k)}(t) &= \frac{d^k}{dt^k} M[e^{itX}] = i^k M[X^k e^{itX}]. \quad \text{Якщо } t \rightarrow 0, \quad \text{то} \\ g^{(k)}(0) &= i^k M[X^k]. \end{aligned}$$

Зокрема, для математичного сподівання випадкової величини маємо:

$$M[X] = g'(0) \frac{1}{i} = -ig'(0).$$

4. Характеристична функція суми **незалежних** випадкових величин дорівнює добутку характеристичних функцій доданків.

Доведення. Нехай $Y = \sum_{k=1}^n X_k$ і $g_k(t)$ – характеристична функція випадкової величини X_k ; випадкові величини X_1, X_2, \dots, X_n за умовою незалежні. Потрібно показати, що $g_y(t) = \prod_{k=1}^n g_k(t)$. Дійсно,

$$g_y(t) = M[e^{itY}] = M\left[e^{it\sum_{k=1}^n X_k}\right] = M\left[\prod_{k=1}^n e^{itX_k}\right]. \text{ Оскільки } X_k \text{ – незалежні,}$$

то незалежні і величини e^{itX_k} . Математичне сподівання добутку незалежних випадкових величин дорівнює добутку їх математичних сподівань (лекція 14). Отже, $g_y(t) = \prod_{k=1}^n M[e^{itX_k}] = \prod_{k=1}^n g_k(t)$, що й потрібно було довести.

Розглянемо незалежні випадкові величини X_1, X_2, \dots, X_n , що мають один і той же закон розподілу з математичним сподіванням m і дисперсією σ^2 . Введемо випадкову величину

$$Y_n = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i \quad (4)$$

з характеристиками

$M[Y_n] = \sum_{i=1}^n M[X_i] = nm$, $D[Y_n] = \sum_{i=1}^n D[X_i] = n\sigma^2$ і відповідну їй нормовану випадкову величину (нормовану суму (4))

$$\overset{*}{Y}_n = \frac{\overset{*}{Y}_n - M[Y_n]}{\sqrt{D[Y_n]}} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\overset{*}{X}_i - m}{\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \overset{*}{X}_i, \quad (5)$$

де $\overset{*}{X}_i = \frac{\overset{*}{X}_i - m}{\sigma}$ – нормовані випадкові величини $i = 1, 2, \dots, n$. Очевидно, що

$$M\left[\overset{*}{Y}_n\right] = 0, \quad D\left[\overset{*}{Y}_n\right] = 1, \quad M\left[\overset{*}{X}_i\right] = 0, \quad D\left[\overset{*}{X}_i\right] = 1 \quad i = 1, 2, \dots, n. \quad (6)$$

Випадкові величини $\overset{*}{X}_i$ мають один і той же закон розподілу, а тому одні і ту ж характеристичну функцію $g(t)$. Розкладемо цю функцію в ряд Маклорена

$$g(t) = g(0) + \frac{g'(0)}{1!}t + \frac{g''(0)}{2!}t^2 + \alpha(t)t^2 = 1 - \frac{t^2}{2} + \alpha(t)t^2, \quad (7)$$

де $\alpha(t) \rightarrow 0$ при $t \rightarrow 0$. Крім того, на основі (6) враховано, що

$$g(0) = 1, \quad g'(0) = iM\left[\overset{*}{X}_i\right] = 0, \quad g''(0) = i^2M\left[\overset{*}{X}_i^2\right] = -D\left[\overset{*}{X}_i\right] = -1$$

(див. властивість 3 характеристичних функцій).

Враховуючи 2-гу та 4-ту властивості характеристичних функцій, для характеристичної функції випадкової величини $\overset{*}{Y}_n$ отримуємо

$$g_n(t) = \left[g\left(\frac{t}{\sqrt{n}}\right) \right]^n = \left[1 - \frac{t^2}{2n} + \alpha\left(\frac{t}{\sqrt{n}}\right) \frac{t^2}{2n} \right]^n. \quad (8)$$

Знайдемо границю виразу (8) при $n \rightarrow \infty$. Для цього спочатку прологарифмуємо цей вираз:

$$\begin{aligned} \lim_{n \rightarrow \infty} \ln g_n(t) &= \lim_{n \rightarrow \infty} n \ln \left[1 - \frac{t^2}{2n} + \alpha\left(\frac{t}{\sqrt{n}}\right) \frac{t^2}{2n} \right] = \lim_{n \rightarrow \infty} n \left[-\frac{t^2}{2n} + \alpha\left(\frac{t}{\sqrt{n}}\right) \frac{t^2}{2n} \right] = \\ &= \lim_{n \rightarrow \infty} \left[-\frac{t^2}{2} + \alpha\left(\frac{t}{\sqrt{n}}\right) \frac{t^2}{2} \right] = -\frac{t^2}{2}, \quad \text{оскільки } \alpha\left(\frac{t}{\sqrt{n}}\right) \rightarrow 0 \text{ при } n \rightarrow \infty. \end{aligned}$$

Тут мискористались еквівалентністю нескінченно малих x та $\ln(1+x)$ при $x \rightarrow 0$. Отже,

$$\lim_{n \rightarrow \infty} g_n(t) = e^{-\frac{t^2}{2}}, \quad (9)$$

а це означає, що послідовність характеристичних функцій $g_n(t)$ нормованих сум $\overset{*}{Y}_n$ збігається при $n \rightarrow \infty$ до характеристичної функції нормованого нормальногорозподілу $N(0,1)$ (див. приклад 1).

В теорії ймовірностей **доведено**, що зі збіжності характеристичних функцій випливає збіжність законів розподілу. **Тому закон розподілу нормованої суми $\overset{*}{Y}_n$ зі збільшенням n наближається до нормованогоального закону.** Одержаний результат є однією з самих простих форм центральної граничної теореми (випадок однаково розподілених доданків).

Оскільки лінійне перетворення не змінює вигляду закону розподілу (див. лекцію 15), а з (5) маємо

$$Y_n = \sigma\sqrt{n}\overset{*}{Y}_n + mn, \quad (10)$$

то зі збільшенням n закон розподілу суми (4) $\overset{*}{Y}_n$ також наближається до нормальногозакону. Тому сформулюємо одержаний результат у такій формі.

Теорема 1. (*центральна гранична теорема для однаково розподілених доданків*). Якщо X_1, X_2, \dots, X_n – незалежні випадкові величини, що мають один і той же закон розподілу з математичним сподіванням m і дисперсією σ^2 , то при необмеженому збільшенні n закон розподілу суми $Y_n = \sum_{i=1}^n X_i$ необмежено наближається до нормального.

Приклад 2. Нехай $Y_n = m$ – число “успіхів” у схемі Бернуллі з n випробувань. Як ми вже знаємо, випадкова величина Y_n має біномний розподіл з характеристиками $M Y_n = np$, $D Y_n = npq$, де p – ймовірність “успіху” в кожному випробуванні, $q = 1 - p$. Крім того,

$$Y_n = \sum_{i=1}^n X_i, \quad (11)$$

де X_i – індикатор “успіху” в i -му випробуванні (див. лекцію 14, приклад 1). Випадкові величини X_i ($i = 1, 2, \dots, n$) є незалежними і однаково розподіленими з математичним сподіванням $M X_i = p$ і дисперсією $D X_i = pq$. Отже, зі збільшенням n розподіл нормованої

суми (11) $\overset{*}{Y}_n = \frac{Y_n - np}{\sqrt{npq}} = \frac{m - np}{\sqrt{npq}}$ наближається до нормованого нормального розподілу $N(0,1)$. Тому для достатньо великих n справедлива наблизена рівність (див. лекцію 10)

$$P\left(a \leq \frac{m - np}{\sqrt{npq}} \leq b\right) \approx \Phi(b) - \Phi(a), \quad (12)$$

де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$ – функція розподілу закону $N(0,1)$.

Наблизена рівність (12) – це вже відома нам рівність (див. лекцію 6, формулу (6)), що виражає собою суть інтегральної теореми Муавра-Лапласа. Отже, інтегральна теорема Муавра-Лапласа є простим наслідком центральної граничної теореми для однаково розподілених доданків.

На рис. 1 показано, як веде себе біномний розподіл при $p = 0,25$ і $n = 5, 10, 20$ (p_m – ймовірність настання рівно m “успіхів”). Видно, що зі зростанням n “графіки” біномного розподілу “розповзаються” у

ширину, симетризуються і при великих n “наближаються” до кривої Гаусса

$$f(x) = \frac{1}{\sqrt{2\pi} \sqrt{npq}} e^{-\frac{(x-np)^2}{2npq}}. \quad (13)$$

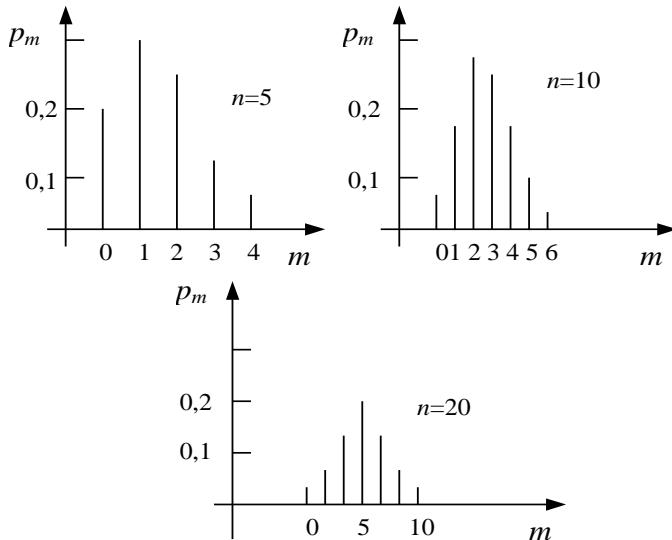


Рис.1

Для достатньо великих n ймовірність p_m може бути наблизено знайдена як $p_m \approx f(x=m)$ (див формулу (5) лекції 6).

Центральна гранична теорема при достатньо широкому комплексі умов справедлива і для неоднаково розподілених випадкових величин. Наприклад, має місце більш загальна теорема.

Теорема 2. (Теорема Ляпунова) Нехай X_1, X_2, \dots, X_n – незалежні випадкові величини з довільними розподілами, кожна з яких має скінченні математичне сподівання, дисперсію і третій центральний момент. Якщо при $n \rightarrow \infty$ виконується умова

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n M \left[|X_i - M|^{3+} \right]}{\left(\sum_{i=1}^n D(X_i) \right)^{\frac{3}{2}}} = 0, \quad (14)$$

то розподіл випадкової величини $\bar{Y}_n^* = \frac{\bar{Y}_n - M_{\bar{Y}_n}}{\sqrt{D_{\bar{Y}_n}}}$, де $\bar{Y}_n = \sum_{i=1}^n X_i$, зі

збільшенням n необмежено наближається до нормальногорозподілу з нульовим математичним сподіванням і одиничною дисперсією, тобто

до нормального закону зі щільністю $\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$. Очевидно, що розподіл

суми $\bar{Y}_n = \sum_{i=1}^n X_i$ також наближається до нормального.

У практичних задачах центральну граничну теорему часто застосовують для знаходження ймовірності того, що сума декількох випадкових величин попаде в заданий інтервал. Нехай X_1, X_2, \dots, X_n – незалежні випадкові величини з математичними сподіваннями m_1, m_2, \dots, m_n і дисперсіями D_1, D_2, \dots, D_n . Нехай умови центральної граничної теореми виконуються і число доданків n є достатнім для того, щоб закон розподілу величини $\bar{Y} = \sum_{i=1}^n X_i$ можна було вважати наближено нормальним. Тоді ймовірність того, що випадкова величина \bar{Y} потрапить в інтервал (α, β) можна наближено знайти за формулою

$$P(\alpha < \bar{Y} < \beta) = \Phi\left(\frac{\beta - m_y}{\sigma_y}\right) - \Phi\left(\frac{\alpha - m_y}{\sigma_y}\right), \quad (15)$$

де m_y, σ_y – математичне сподівання та середнє квадратичне відхилення величини \bar{Y} , а $\Phi(x)$ – функція розподілу нормального закону $N(0,1)$ (див. лекцію 10). Величини m_y, σ_y знаходяться за відомими m_i, D_i :

$$m_y = \sum_{k=1}^n m_k, \quad \sigma_y = \sqrt{D_{\bar{Y}}} = \sqrt{\sum_{k=1}^n D_k} = \sqrt{\sum_{k=1}^n D_k}.$$

Отже, для того щоб наближено знайти ймовірність потрапляння суми великого числа випадкових величин у заданий інтервал, не обов'язково знати закони розподілу цих величин, достатньо знати тільки їх числові характеристики. Звичайно, мається на увазі, що виконується основна умова центральної граничної теореми – **рівномірно малий вплив доданків на розсювання суми** (у цьому полягає фізичний зміст умови (14) в теоремі Ляпунова).

Приклад 3. Похибки при будь-якому вимірюванні можуть бути як систематичними, так і випадковими. Якщо від систематичних похибок

можна завжди позбавитись, то від випадкової похибки – ніколи. Випадкова похибка при вимірюванні – це сума великого числа “елементарних” похибок, кожна з яких породжується тим чи іншим випадковим фактором, що сам по собі мало впливає на результат вимірювання. А тому вклад кожної “елементарної” випадкової похибки в сумарну випадкову похибку, як правило, незначний. Отже, за теоремою Ляпунова випадкова похибка вимірювання повинна мати наблизено нормальній закон розподілу. Як показує практика, це дійсно так. Якщо вимірюється деяка величина a , то результат вимірювання X є випадковою величиною $X = a + \Delta X$. При відсутності систематичної похибки вимірювань $M X = a$, а тому ΔX – випадкова похибка і $M \Delta X = 0$. Отже, випадкова похибка вимірювань має

щільність розподілу $P_{\Delta x}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$, де σ – середнє квадратичне відхилення випадкової величини ΔX (середня квадратична похибка вимірювань).

Розділ 2. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

Лекція № 18

Поняття вибірки. Методи описування вибірки. Характеристики вибірки

Як ми вже знаємо, методи теорії ймовірностей працюють з того моменту, як побудовано ймовірнісний простір того чи іншого випадкового явища, тобто по суті з'ясовано ймовірнісну природу цього явища. Наприклад, відомі ймовірності елементарних подій випадкового експерименту, функція розподілу випадкової величини тощо. Тоді, застосовуючи методи теорії ймовірностей, можна з'ясувати, як будуть вести себе ті або інші характеристики випадкового явища в експериментах, тобто спрогнозувати поведінку самого явища.

Але у більшості випадків, що зустрічаються на практиці, ані ймовірності окремих випадкових подій, ані точний вираз функції розподілу випадкової величини нам невідомі. Тому постає питання про їхнє експериментальне визначення.

Методи, які дають змогу за результатами випробувань робити певні ймовірнісні висновки, є предметом вивчення **математичної статистики**. Отже, вихідними у математичній статистиці є експериментальні дані. Це дає підстави розглядати її як самостійний розділ науки про випадкові явища. Разом з тим, як ми побачимо далі, кожна задача математичної статистики є по суті своєрідною задачею теорії ймовірностей. Тому вивчати математичну статистику, не знаючи основ теорії ймовірностей, неможливо.

Подамо деякі найпростіші і в той же час найтиповіші задачі математичної статистики:

1. *Оцінка на основі експериментальних даних невідомої ймовірності р випадкової події A.* На практиці ця задача зустрічається досить часто. Нехай, наприклад, потрібно оцінити число бракованих виробів у досить великий партії з N виробів. Якщо партія містить M бракованих виробів, то $\frac{M}{N}$ – це ймовірність того, що

навмання вибраний виріб буде бракованим. Зрозуміло, що задача оцінки числа бракованих виробів зводиться до задачі оцінки невідомої ймовірності.

2. *Оцінка на основі результатів вимірювань невідомої функції розподілу випадкової величини.* Задача ставиться так: у результаті незалежних випробувань над випадковою величиною X одержані такі її значення: x_1, x_2, \dots, x_n . Необхідно наблизено оцінити невідому функцію розподілу $F(x)$ випадкової величини X .

3. *Оцінка невідомих параметрів розподілу.* Задача ставиться так: припускається, що випадкова величина X має функцію розподілу певного вигляду, яка залежить від k параметрів. Значення цих параметрів невідомі (про вигляд функції розподілу часто можна зробити досить точні висновки на основі загальнотеоретичних міркувань, наприклад, на основі центральної граничної теореми). Необхідно на основі дослідних (експериментальних) даних оцінити значення невідомих параметрів функції розподілу.

4. *Статистична перевірка гіпотез.* Одна з основних задач статистичної перевірки гіпотез ставиться так: на основі деяких міркувань можна вважати, що функцією розподілу випадкової величини X є функція $F(x)$. Запитується: чи сумісні дослідні дані з гіпотезою, що випадкова величина X має розподіл $F(x)$? Зокрема, якщо закон розподілу випадкової величини X не викликає сумніву, а перевірки потребують значення деяких параметрів, що характеризують розподіл, то в задачі запитується: чи не спростовують дослідні дані гіпотезу, що параметри закону розподілу мають певні значення?

В основі математичної статистики лежить ряд вихідних понять, без попереднього знайомства з якими неможливе вивчення сучасних методів обробки експериментальних даних. Це, насамперед, поняття *генеральної сукупності* та *вибірки*.

Нехай потрібно дослідити деяку ознаку, що властива великій кількості однотипних предметів (наприклад, розміри деталей даного типу тощо). Сукупність значень цієї ознаки для всіх N предметів даного типу називається *генеральною сукупністю* (можна й так: генеральною сукупністю називається сукупність всіх об'єктів, що підлягають дослідженням відносно деякої ознаки). Як правило, число N у генеральній сукупності досить велике. На практиці суцільне обстеження всіх об'єктів генеральної сукупності застосовується рідко, а іноді це зробити просто неможливо, наприклад, коли генеральна сукупність містить надто велике число об'єктів. Якщо дослідження об'єкту пов'язане з його знищеннем або потребує великих

матеріальних затрат, то суцільне дослідження всіх об'єктів генеральної сукупності з практичної точки зору втрачає будь-який смисл. У таких випадках з генеральної сукупності випадково відбирають обмежену кількість об'єктів (виробів) і піддають їх вивченню відносно тієї чи іншої ознаки. Вибіркою сукупністю або просто **вибіркою** обсягу n ($n < N$) називають сукупність n випадково відібраних об'єктів з генеральної сукупності. Кожний об'єкт вибірки відібраний **випадково**, якщо всі об'єкти генеральної сукупності мали однакову ймовірність попасті до вибірки.

Нехай вивчається деяка випадкова величина X , закон розподілу якої невідомий. З цією метою над випадковою величиною X проводиться ряд незалежних дослідів. Множину всіх можливих значень випадкової величини X також назовемо генеральною сукупністю, а ті її значення, що спостерігались у результаті дослідів – вибіркою. Результати дослідів (вибірку) подають у вигляді таблиці з двох рядків, перший з яких вказує на номер досліду i , а другий – на результат досліду x_i .

i	1	2	3	...	n
x_i	x_1	x_2	x_3	...	x_n

Серед величин x_i можуть бути однакові числа. Таблицю, в якій містяться номера і

результати вимірювань, назовемо **статистичним рядом**. Статистичний ряд є первинною формою запису вибірки і підлягає подальшій обробці.

У першу чергу, зі статистичного ряду вписують значення x_1 і число спостережень n_1 цього значення, потім x_2 і число спостережень n_2 і так далі, аж до x_k і числа спостережень n_k . Очевидно, що $\sum_{i=1}^k n_i = n$, де n – обсяг вибірки. Числа x_i , $i = 1, 2, \dots, k$ називають **варіантами**. Послідовність варіант, записаних у зростаючому порядку, називають **варіаційним рядом**. Різницю між найбільшим та найменшим значеннями варіаційного ряду називають **розмахом** вибірки: $R = x_k - x_1$ – розмах вибірки. Число спостережень n_i називають **частотою**, а його відношення до обсягу вибірки n , тобто $\frac{n_i}{n} = w_i$ – **відносною частотою**. Очевидно, що $\sum_{i=1}^k w_i = 1$.

Статистичним розподілом вибірки називають перелік варіант, записаних у зростаючому порядку, і відповідних їм частот або відносних частот.

За відомим статистичним розподілом вибірки (статистичним розподілом частот) будується так звана **статистична функція розподілу**, яку ще називають **емпіричною функцією розподілу**.

Емпірична функція розподілу $F^*(x)$ визначається за відомими відносними частотами співвідношенням

$$F^*(x) = \sum_{x_i < x} \frac{n_i}{n}, \quad (1)$$

де підсумовування поширюється на всі x_i , які менші за x .

Очевидно, що $F^*(x)$ – неспадна функція, яка набуває значень з відрізка $[0;1]$. Якщо $x \leq x_1$, де x_1 – найменше значення вибірки, то $F^*(x) = 0$. При $x > x_k$, де x_k – найбільше значення вибірки, $F^*(x) = 1$.

Приклад 1. Знайти емпіричну функцію $F^*(x)$ за даним статистичним розподілом вибірки. Побудувати графік цієї функції.

x_i	1	3	4	5
n_i	6	8	6	5

Розв'язання: Обсяг вибірки $n = 6 + 8 + 6 + 5 = 25$. Найменша варіанта $x_1 = 1$, тому $F^*(x) = 0$ при $x \leq 1$. Якщо $1 < x \leq 3$, то

$$F^*(x) = \frac{6}{25} = 0,24. \text{ При } 3 < x \leq 4 \text{ згідно з (1) маємо}$$

$$F^*(x) = \frac{6}{25} + \frac{8}{25} = 0,56. \text{ Аналогічно, при } 4 < x \leq 5$$

$$F^*(x) = \frac{6}{25} + \frac{8}{25} + \frac{6}{25} = 0,80. \text{ При } x > 5 \quad F^*(x) = \frac{6}{25} + \frac{8}{25} + \frac{6}{25} + \frac{5}{25} = 1.$$

Отже,

$$F^*(x) = \begin{cases} 0 & \text{при } x \leq 1, \\ 0,24 & \text{при } 1 < x \leq 3, \\ 0,56 & \text{при } 3 < x \leq 4, \\ 0,80 & \text{при } 4 < x \leq 5, \\ 1 & \text{при } x > 5. \end{cases}$$

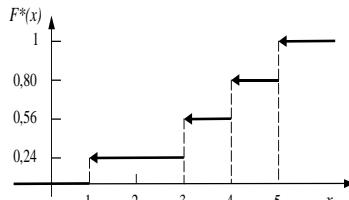


Рис. 1

Графік функції $F^*(x)$ показано на рис. 1

Емпірична функція розподілу будь-якої випадкової величини (дискретної чи неперервної) завжди є розривною східчастою функцією, стрибки якої відповідають значенням випадкової величини,

що спостерігались, і по величині дорівнюють відносним частотам цих значень.

Емпіричну функцію розподілу $F^*(x)$ можна ще визначити як закон зміни відносних частот події $X < x$ на даному статистичному матеріалі: $F^*(x) = P^* X < x$. Згідно з теоремою Бернуллі при необмеженому збільшенні числа дослідів (експериментів) n відносна частота події $X < x$, тобто $P^* X < x$, збігається за ймовірністю до ймовірності цієї події. Це означає, що емпірична функція розподілу $F^*(x)$ зі збільшенням n збігається за ймовірністю до справжньої функції розподілу $F(x)$ випадкової величини X .

З метою більшої наочності будують різні графічні зображення вибірки, зокрема **полігон** та **гістограму** частот.

Полігоном частот називають ламану з вершинами в точках (x_i, n_i) .

Полігоном відносних частот називають ламану з вершинами в точках (x_i, w_i) , де $w_i = \frac{n_i}{n}$.

Приклад 2. Побудувати полігон відносних частот за заданим розподілом вибірки:

x_i	1	2	4	6
n_i	4	7	6	3

Розв'язання. Знайдемо відносні частоти. Оскільки обсяг вибірки $n = 4 + 7 + 6 + 3 = 20$, то $w_1 = \frac{4}{20} = 0,2$; $w_2 = \frac{7}{20} = 0,35$; $w_3 = \frac{6}{20} = 0,3$;

$w_4 = \frac{3}{20} = 0,15$. Отже, розподіл відносних частот має вигляд

x_i	1	2	4	6
w_i	0,2	0,35	0,3	0,15

Будуємо точки $(1; 0,2)$; $(2; 0,35)$; $(4; 0,3)$; $(6; 0,15)$ і з'єднуємо їх відрізками прямих.

Отримуємо шуканий полігон відносних частот (рис. 2).



Рис. 2

У випадку неперервної випадкової величини X елементи вибірки, зазвичай, об'єднують у групи і подають результати спостережень у вигляді **групованим статистичним ряду**.

Для цього інтервал, який містить всі елементи вибірки, розбивають на k часткових інтервалів, що не перетинаються. У більшості випадків ці інтервали вибирають однакової довжини $b \approx \frac{R}{k}$, де R – розмах вибірки. Після того, як часткові інтервали вибрано, визначають частоти n_i^* – кількість елементів вибірки, що попали в i -й інтервал, а потім $\frac{n_i^*}{b}$ – щільність частот. Одночасно підраховують $\frac{w_i}{b}$ – щільність відносних частот, де $w_i = \frac{n_i^*}{n}$ – відносна частота. Якщо елемент вибірки співпадає з верхньою межею часткового інтервалу, то його відносять до наступного інтервалу. За отриманими даними будують так звану **гістограму частот**.

Гістограмою частот називають східчасту фігуру, що складається з прямокутників, основами яких є часткові інтервали довжиною b , а висоти дорівнюють $\frac{n_i^*}{b}$. Площа гістограми дорівнює сумі всіх частот, тобто обсягу вибірки: $\sum_i b \frac{n_i^*}{b} = \sum_i n_i^* = n$.

Аналогічно будують **гістограму відносних частот** – східчасту фігуру, що складається з прямокутників, в основі яких лежать часткові інтервали довжиною b , а висоти дорівнюють $\frac{w_i}{b}$. Площа гістограми відносних частот дорівнює сумі всіх відносних частот, тобто одиниці: $\sum_i b \frac{w_i}{b} = \sum_i w_i = \sum_i \frac{n_i^*}{n} = 1$. Якщо точки гістограми відносних частот

з'єднати плавною лінією, то ця лінія у першому наближенні буде зображати графік щільності $f(x)$ розподілу ймовірностей випадкової величини

Приклад 3. Побудувати гістограму відносних частот за даним розподілом вибірки обсягу $n = 50$.

Номер інтервалу	1	2	3	4	5
Частковий інтервал (x_i, x_{i+1})	(1;3)	(3;5)	(5;7)	(7;9)	(9;11)
Частоти n_i^*	7	10	20	8	5

Розв'язання. Знайдемо відносні частоти:

$$w_1 = \frac{7}{50} = 0,14; \quad w_2 = \frac{10}{50} = 0,2; \quad w_3 = \frac{20}{50} = 0,4; \quad w_4 = \frac{8}{50} = 0,16;$$

$w_5 = \frac{5}{50} = 0,1$. Знайдемо щільності відносних частот, враховуючи, що

довжина часткового інтервалу $b = 2$:

$\frac{w_1}{b} = 0,07$; $\frac{w_2}{b} = 0,1$; $\frac{w_3}{b} = 0,2$; $\frac{w_4}{b} = 0,08$; $\frac{w_5}{b} = 0,05$. Будуємо гістограму частот (рис. 3).

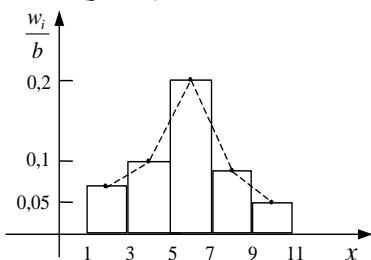


Рис. 3

Зауважимо, що за даними групованого статистичного ряду можна побудувати і полігон частот як ламану з вершинами в точках x_i^*, n_i^* , де x_i^* — середина i -го часткового інтервалу. Аналогічно будеться полігон відносних частот (ламана з вершинами в точках x_i^*, w_i , $w_i = \frac{n_i^*}{n}$).

Особливе значення має ламана з вершинами в точках $\left(x_i^*, \frac{w_i}{b} \right)$, яка є наближенням графіка щільності розподілу ймовірностей випадкової величини (рис. 3).

Нехай x_1, x_2, \dots, x_n – вибірка обсягу n з генеральної сукупності X . Якщо на цю вибірку умовно дивитись як на дискретну випадкову величину, що приймає значення x_1, x_2, \dots, x_n з ймовірностями $\frac{1}{n}$, то

можна ввести числові характеристики, які називають ***вибірковими числовими характеристиками*** (характеристиками вибірки):

$$M^* X = m_x^* = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^k n_i x_i - \text{вибіркове середнє},$$

$$D^* X = D_x^* = \frac{1}{n} \sum_{i=1}^n |x_i - m_x^*|^2 - \text{вибіркова дисперсія}.$$

Аналогічно, можна записати співвідношення для вибіркових початкових та центральних моментів:

$$M^* [X^k] = \frac{1}{n} \sum_{i=1}^n x_i^k, \quad M^* [X - M^*]^k = \frac{1}{n} \sum_{i=1}^n |x_i - m_x^*|^k.$$

Слід мати на увазі, що вибіркові числові характеристики є характеристиками даної вибірки, а не характеристиками розподілу генеральної сукупності. Тому ми й позначаємо їх “зірочкою”. Вибіркові характеристики зі збільшенням обсягу вибірки (числа дослідів) збігаються за ймовірністю до відповідних числових характеристик генеральної сукупності і при достатньо великому числі дослідів можуть прийматись за наближені значення останніх.

Нехай досліжується двовимірна випадкова величина (X, Y) (двохвимірна генеральна сукупність) і в результаті n спостережень над цією величиною одержано значення $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ – двовимірну вибірку обсягу n .

Розподілом двовимірної вибірки називають розподіл двовимірного дискретного випадкового вектора, який приймає значення (x_i, y_i) ,

$i = 1, 2, \dots, n$ з одинаковими ймовірностями $\frac{1}{n}$. Вибіркові числові

характеристики двовимірної вибірки обчислюються як відповідні числові характеристики двовимірного випадкового вектора дискретного типу (див. лекцію 13).

Основні вибіркові характеристики двовимірної вибірки такі:

$$\bar{x} = m_x^* = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = m_y^* = \frac{1}{n} \sum_{i=1}^n y_i - \text{вибіркові середні};$$

$$s_x^2 = D_x^* = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|^2, \quad s_y^2 = D_y^* = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}|^2 - \text{вибіркові дисперсії};$$

$$K_{xy}^* = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) - \text{вибіркова коваріація або}$$

$$k_{xy}^* = \frac{K_{xy}^*}{s_x s_y} - \text{вибірковий коефіцієнт кореляції}.$$

Лекція № 19

Оцінка числових характеристик та параметрів розподілу випадкової величини за результатами вибірки.

Точкові оцінки та їх властивості

Нехай закон розподілу випадкової величини відомий і необхідно знайти лише параметри, від яких він залежить (наприклад, математичне сподівання m та середнє квадратичне відхилення σ у випадку нормальну розподіленої випадкової величини). Слід зазначити, що в деяких випадках закон розподілу випадкової величини взагалі є несуттєвим, а необхідно знати лише числові характеристики цієї величини.

Отже, нехай у результаті n незалежних дослідів отримано вибірку (x_1, x_2, \dots, x_n) . Виникає запитання, як на основі цього експериментального матеріалу наближено оцінити деякий параметр a розподілу випадкової величини X ? Оцінку параметра будемо позначати \tilde{a} . Міркуємо так: припустимо, що досліди ще не проведені і їх результати нам не відомі, тобто є випадковими. Позначимо через X_i значення, яке прийме випадкова величина X в i -му досліді. Тоді конкретну вибірку (x_1, x_2, \dots, x_n) можна розглядати як деяку реалізацію випадкового вектора з n незалежними компонентами X_1, X_2, \dots, X_n , кожна з яких має розподіл генеральної сукупності, тобто самої випадкової величини X , що спостерігається. Оскільки всяка оцінка \tilde{a} параметра a є деякою функцією (комбінацією) елементів вибірки, які випадковим чином змінюються від вибірки до вибірки, то цю оцінку можна розглядати як функцію n випадкових величин X_i (випадкового вектора), тобто вважати випадковою величиною

$$\tilde{a} = \tilde{a}(X_1, X_2, \dots, X_n). \quad (1)$$

Наприклад, середнє арифметичне n значень випадкової величини X є функцією випадкових величин X_i , що має вигляд $\tilde{m} = \frac{1}{n} \sum_{i=1}^n X_i$.

Невипадкова функція (1) випадкових аргументів X_i , які є можливими значеннями елементів вибірки, називається **статистикою**. Наприклад, $\tilde{m} = \frac{1}{n} \sum_{i=1}^n X_i$, $\tilde{D} = \frac{1}{n} \sum_{i=1}^n (X_i - \tilde{m})^2$ – статистики.

Отже, кожна оцінка \tilde{a} параметра a розглядається як випадкова величина. Закон розподілу цієї величини залежить від закону

розподілу випадкової величини X , що спостерігається, вигляду функції (1), а тому і від числа дослідів n . Для того щоб оцінка \tilde{a} була “доброякісною”, до неї висуваються певні вимоги:

1. По-перше, ця оцінка повинна бути **змістовною**. Це означає, що зі збільшенням числа дослідів n оцінка \tilde{a} повинна наблизатися (збігатися за ймовірністю) до параметра a , що оцінюється, тобто

$$\lim_{n \rightarrow \infty} P |\tilde{a} - a| < \varepsilon = 1 \text{ для будь-якого } \varepsilon > 0. \quad (2)$$

2. По-друге, бажано, щоб при використанні величини \tilde{a} замість a не реалізовувалась систематична похибка у бік завищення чи зниження. Для цього необхідно вимагати, щоб виконувалась умова

$$M \tilde{a} = a. \quad (3)$$

Якщо умова (3) виконується, то оцінка \tilde{a} називається **незміщеною**. У протилежному разі вона називається **зміщеною**.

3. Оцінки, які є незміщеними і змістовними, можуть відрізнятись дисперсіями. Чим менша дисперсія оцінки, тим менша ймовірність грубої похибки при визначенні параметра a . Тому необхідно, щоб дисперсія оцінки була мінімальною, тобто

$$D \tilde{a}(X_1, X_2, \dots, X_n) = D_{\min}. \quad (4)$$

Незміщена оцінка, яка має властивість (4), називається **ефективною**. Якщо \tilde{a}_1, \tilde{a}_2 – дві різні оцінки параметра a , то при $D \tilde{a}_1 < D \tilde{a}_2$ кажуть, що оцінка \tilde{a}_1 є більш ефективною, ніж оцінка \tilde{a}_2 .

Якщо параметр a водночас є і числововою характеристикою випадкової величини X , наприклад, математичним сподіванням, дисперсією тощо, то за оцінку цього параметра природно взяти відповідну вибіркову числову характеристику. Проте такий спосіб оцінювання не завжди призводить до найкращих оцінок у розумінні сформульованих до них вище вимог. Зупинимось на знаходженні оцінок для математичного сподівання $m = M X$ та дисперсії $D = D X$ випадкової величини X (для нормального розподілу, наприклад, характеристики m і \sqrt{D} є параметрами розподілу). Отже, нехай є випадкова величина X з математичним сподіванням $M X$ та дисперсією $D X$. Обидві характеристики невідомі. За оцінку математичного сподівання візьмемо середнє арифметичне

$$\tilde{m} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (5)$$

Оцінка (5) є змістовою. Дійсно, згідно з законом великих чисел

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - M[X] \right| < \varepsilon \right\} = 1. \text{ Знайдемо } M[\tilde{m}] :$$

$$M[\tilde{m}] = M \left[\frac{1}{n} \sum_{i=1}^n X_i \right] = \frac{1}{n} \sum_{i=1}^n M[X_i] = \frac{1}{n} n M[X] = M[X] = m. \quad \text{Тут ми}$$

скористалися тим, що X_i – випадкові величини, кожна з яких розподілена за тим же законом, що й випадкова величина X , тобто $M[X_i] = M[X]$. Отже, оцінка (5) є також незміщеною.

Знайдемо дисперсію оцінки (5):

$$D[\tilde{m}] = \frac{1}{n^2} D \left[\sum_{i=1}^n X_i \right] = \frac{1}{n^2} \sum_{i=1}^n D[X_i] = \frac{1}{n^2} n D[X] = \frac{D[X]}{n}, \text{ оскільки } X_i –$$

незалежні і $D[X_i] = D[X]$.

Ефективність чи неефективність оцінки (5) залежать від закону розподілу випадкової величини X . Доведено, що для нормально розподіленої випадкової величини X оцінка (5) є ефективною.

Перейдемо до оцінки дисперсії. За таку оцінку візьмемо вибіркову дисперсію

$$D^* = \frac{1}{n} \sum_{i=1}^n (X_i - \tilde{m})^2. \quad (6)$$

Перевіримо її змістовність.

$$D^* = \frac{1}{n} \sum_{i=1}^n X_i^2 - 2\tilde{m} \frac{1}{n} \sum_{i=1}^n X_i + \tilde{m}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \tilde{m}^2. \quad (7)$$

Перший доданок в (7) є середнім арифметичним n значень випадкової величини X^2 і тому збігається за ймовірністю до її математичного сподівання $M[X^2]$. Другий доданок збігається за ймовірністю до $\tilde{m}^2 = M[X]^2$. Оскільки $D[X] = M[X^2] - M[X]^2$, то D^* збігається за ймовірністю до $D[X]$. Отже, оцінка (6) є змістовою. Перевіримо її на незміщеність. Для цього спершу перетворимо вираз для D^* , використовуючи (5):

$$\begin{aligned} D^* &= \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \frac{1}{n^2} \sum_{i=1}^n X_i^2 - \frac{2}{n^2} \sum_{i < j} X_i X_j = \\ &= \frac{n-1}{n^2} \sum_{i=1}^n X_i^2 - \frac{2}{n^2} \sum_{i < j} X_i X_j. \end{aligned}$$

Оскільки дисперсія не залежить від вибору початку координат, остання рівність не зміниться, якщо ми замістимо величини X_i візьмемо

величини $\overset{\circ}{X}_i = X_i - m$, тобто відцентруємо всі випадкові величини X_i :

$$D^* = \frac{n-1}{n^2} \sum_{i=1}^n \overset{\circ}{X}_i^2 - \frac{2}{n^2} \sum_{i < j} \overset{\circ}{X}_i \overset{\circ}{X}_j.$$

Зайдемо математичне сподівання $M[D^*]$:

$$\begin{aligned} M[D^*] &= \frac{n-1}{n^2} \sum_{i=1}^n M[\overset{\circ}{X}_i^2] - \frac{2}{n^2} \sum_{i < j} M[\overset{\circ}{X}_i \overset{\circ}{X}_j] = \frac{n-1}{n^2} \sum_{i=1}^n D[\overset{\circ}{X}_i] = \\ &= \frac{n-1}{n^2} D[X] n = \frac{n-1}{n} D[X]. \end{aligned}$$

Тут ми скористалися тим, що X_i та X_j (або $\overset{\circ}{X}_i$ та $\overset{\circ}{X}_j$) є незалежними випадковими величинами і тим, що $M[\overset{\circ}{X}_i] = 0$. Отже,

$M[D^*] = \frac{n-1}{n} D[X]$. Тому оцінка (6) є зміщеню. Користуючись цією оцінкою замість $D[X]$, ми будемо робити систематичну похибку.

Незміщену оцінку дисперсії отримаємо, якщо домножимо D^* на $\frac{n}{n-1}$:

$$\tilde{D} = D^* \frac{n}{n-1} = \frac{1}{n-1} \sum_{i=1}^n X_i - \tilde{m}^2. \quad (8)$$

Оскільки при $n \rightarrow \infty$, $\frac{n}{n-1} \rightarrow 1$, то оцінка (8) буде також і змістовною.

Нехай в результаті n експериментів отримано конкретну вибірку (x_1, x_2, \dots, x_n) . Тепер нас вже не цікавить, що результати експериментів випадкові. На основі викладеного вище за наближені значення математичного сподівання та дисперсії випадкової величини ми беремо значення:

$$M[X] \approx \tilde{m} = \frac{1}{n} \sum_{i=1}^n x_i; \quad D[X] \approx \tilde{D} = \frac{1}{n-1} \sum_{i=1}^n x_i - \tilde{m}^2. \quad (9)$$

У розрахунках другу з формул (9) зручніше використовувати у вигляді (див. (7), (8))

$$D X \approx \tilde{D} = \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^n x_i^2 - \tilde{m}^2 \right). \quad (9')$$

Рекомендується показати **самостійно**, що при відомому математичному сподіванні m випадкової величини X оцінку дисперсії слід брати у вигляді вибіркової дисперсії, тобто

$$D X \approx D^* = \frac{1}{n} \sum_{i=1}^n x_i - m^2.$$

Іноді знаходження оцінок можна значно спростити, якщо скористатись формулами

$$\begin{aligned} M X &\approx \tilde{m} = \frac{1}{n} \sum_{i=1}^n (x_i - a) + a, \\ D X &\approx \tilde{D} = \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^n (x_i - a)^2 - (\tilde{m} - a)^2 \right). \end{aligned} \quad (10)$$

Правильний вибір величини a спрощує обробку статистичного матеріалу. Рекомендується показати **самостійно**, що формули (10) рівносильні вищенаведеним формулам.

Якщо вибірку оброблено і записано у вигляді статистичного ряду

x_i	x_1	x_2	\dots	x_k
n_i	n_1	n_2	\dots	n_k

де $\sum_{i=1}^k n_i = n$, то оцінки (9) та (9') можна записати у вигляді

$$\begin{aligned} M X &\approx \tilde{m} = \frac{1}{n} \sum_{i=1}^k n_i x_i, \quad D X \approx \tilde{D} = \frac{1}{n-1} \sum_{i=1}^k n_i (x_i - \tilde{m})^2 = \\ &= \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^k n_i x_i^2 - \tilde{m}^2 \right). \end{aligned} \quad (11)$$

Якщо $m = M X$ – відоме, то $D X \approx \frac{1}{n} \sum_{i=1}^k n_i (x_i - m)^2$.

У випадку групованої вибірки у формулах (11) під x_i слід розуміти середини інтервалів, а під n_i – відповідні частоти (у позначеннях лекції 18 це x_i^* та n_i^*).

Нехай X – неперервна зі щільністю розподілу $f(x; a)$ або дискретна з розподілом $P(X=x) = p(x; a)$ випадкова величина, причому параметр a – довільний, тобто не обов'язково є математичним сподіванням, дисперсією чи іншою числововою

характеристикою випадкової величини. Розглянемо один з найпоширеніших загальних методів знаходження оцінки \tilde{a} параметра a , а саме *метод максимальної правдоподібності*.

Означення. Функцію

$$L = L(x_1, x_2, \dots, x_n; a) = f(x_1; a) f(x_2; a) \dots f(x_n; a) \quad (12)$$

у випадку неперервної випадкової величини або функцію

$$L = L(x_1, x_2, \dots, x_n; a) = p(x_1; a) p(x_2; a) \dots p(x_n; a) \quad (13)$$

у випадку дискретної випадкової величини називають *функцією правдоподібності*.

Очевидно, що (12) є спільною щільністю розподілу випадкового вектора X_1, X_2, \dots, X_n , а (13) – ймовірністю того, що реалізацією вибіркового вектора є вибірка x_1, x_2, \dots, x_n .

Оцінкою максимальної правдоподібності \tilde{a} називають таке значення параметра a , при якому функція правдоподібності для фіксованих x_1, x_2, \dots, x_n досягає максимуму: $L(\tilde{a}) = \max L(a)$.

Зауваження. Згідно з (13) оцінкою параметра a є таке значення \tilde{a} , при якому ймовірність одержання даної конкретної вибірки максимальна. У випадку (12) слід говорити про максимум ймовірності попадання вибірки в n -вимірний паралелепіпед з центром в точці x_1, x_2, \dots, x_n і ребрами dx_1, dx_2, \dots, dx_n .

Оскільки точки максимальних значень для функцій $L(a)$ і $\ln L(a)$ співпадають, зручніше знаходити оцінку максимальної правдоподібності з умови

$$\frac{d \ln L(a)}{da} = 0. \quad (14)$$

Саму функцію $\ln L(a)$ називають *логарифмічною функцією правдоподібності*.

Якщо невідомих параметрів розподілу декілька, тобто $a = a_1, a_2, \dots, a_s$, то оцінки $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_s$ цих параметрів знаходять як розв'язок системи рівнянь

$$\frac{\partial \ln L(a_1, a_2, \dots, a_s)}{\partial a_k} = 0, \quad k = 1, 2, \dots, s. \quad (15)$$

Оцінки, одержані методом максимальної правдоподібності для широкого класу розподілів, є змістовними і мають найменшу дисперсію в порівнянні з іншими можливими оцінками тих же параметрів. Іноді оцінка максимальної правдоподібності може бути зміщеною.

Приклад 1. Знайти оцінку максимальної правдоподібності для ймовірності p випадкової події A .

Розв'язання. Нехай проведено n випробувань Бернуллі і в m з них спостерігалася подія A . Розглянемо випадкову величину X – число наставань події A в окремому випробуванні. Маємо розподіл $P(X=0) = 1-p$; $P(X=1) = p$ з невідомим параметром p . Довільна вибірка X_1, \dots, X_n з даного розподілу – це послідовність чисел 0 і 1, причому 1 зустрічається m разів, 0 – $(n-m)$ разів. Запишемо функцію правдоподібності (13):

$$L(x_1, \dots, x_n; p) = p^m (1-p)^{n-m}.$$

Тоді рівняння (14) приймає вигляд

$$\frac{d \ln L}{dp} = \frac{m}{p} - \frac{n-m}{1-p} = 0.$$

Звідси знаходимо єдиний розв'язок

$$\tilde{p} = \frac{m}{n}. \quad (16)$$

Отже, оцінкою максимальної правдоподібності ймовірності події є частота цієї події в послідовності випробувань Бернуллі.

Рекомендується обґрунтувати **самостійно**, що оцінка (16) є змістовою і незміщеною оцінкою ймовірності події.

Приклад 2. Знайти оцінки максимальної правдоподібності для математичного сподівання m і дисперсії σ^2 нормальному розподіленої випадкової величини.

Розв'язання. Щільність розподілу має вигляд

$$f(x; m, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}, \text{ де } m \text{ і } \sigma^2 \text{ – невідомі параметри.}$$

Знаходимо функцію правдоподібності (12):

$$L(x_1, \dots, x_n; m, \sigma^2) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{\sum_{i=1}^n (x_i - m)^2}{2\sigma^2}}.$$

Для логарифмічної функції правдоподібності можна записати

$$\ln L(x_1, \dots, x_n; m, \sigma^2) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{\sum_{i=1}^n (x_i - m)^2}{2\sigma^2}.$$

Система (15) набуває вигляду

$$\begin{cases} \frac{\partial \ln L}{\partial m} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - m) = 0 \\ \frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - m)^2 = 0. \end{cases}$$

З першого рівняння системи знаходимо $\tilde{m} = \frac{1}{n} \sum_{i=1}^n x_i$. Підставляючи це значення в друге рівняння, одержуємо $\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \tilde{m})^2$.

Отже, шуканими оцінками є вибіркове середнє та вибіркова дисперсія. Остання, як ми вже знаємо, є зміщеною оцінкою дисперсії.

Лекція № 20

Інтервалальні оцінки параметрів розподілу

Оцінки, з якими ми мали справу у попередній лекції, називаються точковими оцінками. Такі оцінки визначаються одним числом, тобто вказують на точку числової осі, в якій повинно знаходитись значення невідомого параметра. У більшості задач необхідно не тільки знайти відповідне числове значення параметра a , а й оцінити точність та надійність цього значення. Такого роду задачі дуже важливі при малому числі спостережень (малому обсязі вибірки), оскільки точкова оцінка \tilde{a} у значній мірі є випадковою і заміна a на \tilde{a} може привести до грубих помилок. Для визначення точності оцінки \tilde{a} в математичній статистиці користуються **довірчими інтервалами**, а для визначення надійності – **довірчими ймовірностями**.

Нехай для невипадкового параметра a одержана з досліду незміщена оцінка \tilde{a} . Необхідно оцінити можливу при цьому похибку. Очевидно, що чим менша за абсолютною величиною різниця $|\tilde{a} - a|$, тим точніше оцінка \tilde{a} визначає параметр a . Водночас категорично стверджувати, що оцінка \tilde{a} задовольняє нерівності $|\tilde{a} - a| < \varepsilon$ для деякого наперед заданого числа ε не можна, бо ця оцінка є випадковою величиною. Таке твердження можна зробити лише з деякою ймовірністю β , тобто

$$P |\tilde{a} - a| < \varepsilon = \beta. \quad (1)$$

Якщо закон розподілу точкової оцінки \tilde{a} відомий, то ймовірність β можна знайти.

Зазвичай, задача ставиться дещо по іншому, а саме, наперед задається не число ε , а близька до 1 ймовірність β . Потім знаходиться таке значення $\varepsilon > 0$, для якого виконується рівність (1). Перепишемо цю рівність у таких рівносильних формах:

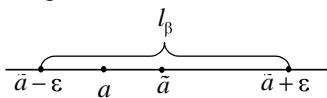
$$P \ a - \varepsilon < \tilde{a} < a + \varepsilon = \beta, \quad P \ \tilde{a} - \varepsilon < a < \tilde{a} + \varepsilon = \beta. \quad (2)$$

Згідно з першою рівністю (2) ймовірність β – це ймовірність попадання випадкової величини \tilde{a} у повністю визначений інтервал $a - \varepsilon, a + \varepsilon$. І таке тлумачення ймовірності β є осмисленим і змістовним.

На відміну від \tilde{a} , невідоме значення параметра a не є випадковим. Випадковими в другій рівності (2) є величини $\tilde{a} - \varepsilon$ та $\tilde{a} + \varepsilon$, а отже, і інтервал

$$l_\beta = \tilde{a} - \varepsilon, \tilde{a} + \varepsilon. \quad (3)$$

Тому ймовірність β в другій рівності (2) слід трактувати не як ймовірність попадання невідомого значення параметра a в інтервал l_β , а як ймовірність того, що випадковий інтервал l_β “покриє” цей параметр a .



Очевидно, що наведені вище тлумачення ймовірності β є рівносильними. Інтервал l_β називають **довірчим або надійним інтервалом**, а ймовірність β – **надійністю оцінки або довірчою ймовірністю**.

Довірчу ймовірність слід розуміти так: якщо, провівши досліди і одержавши n реалізацій випадкової величини \tilde{a} , побудувати дляожної з цих реалізацій інтервал l_β (3), то деякі з цих інтервалів “покриють” точку a , а інші – не “покриють”. Частка інтервалів, що “покривають” a , буде приблизно дорівнювати β .

Як приклад, знайдемо з надійністю β довірчий інтервал для математичного сподівання нормально розподіленої випадкової величини X у припущені, що дисперсія $D X = \sigma^2$ відома. У попередній лекції було показано, що незміщеною оцінкою математичного сподівання є вибіркове середнє

$$\tilde{m} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (4)$$

Випадкові величини X_i незалежні і мають такий же розподіл, що й випадкова величина X , тобто нормальній. З теорії ймовірностей

відомо, що сума незалежних і нормальню розподілених випадкових величин також розподілена нормально (див. лекцію 15). Тому \tilde{m} є нормальню розподіленою випадковою величиною з характеристиками $M \tilde{m} = m$ та $D \tilde{m} = \frac{\sigma^2}{n}$ (див. лекцію 19). Скористаємося формулou для ймовірності попадання нормально розподіленої випадкової величини на симетричний відносно математичного сподівання інтервал

$$P |\tilde{m} - m| < \varepsilon = 2\Phi\left(\frac{\varepsilon}{\sigma/\sqrt{n}}\right) - 1, \quad (5)$$

де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ – функція розподілу стандартизованої

нормальної величини. Для значень цієї функції, як ми вже знаємо, складені таблиці (див. Додаток 2, табл. 1). Оскільки ймовірність у лівій частині (5) задано (надійність β), а σ вважається відомим, то невідоме число ε можна знайти з рівняння

$$\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = \frac{1+\beta}{2}. \quad (6)$$

Для цього за таблицею значень функції $\Phi(x)$ знаходиться аргумент x_0 , якому відповідає задане значення функції $\frac{1+\beta}{2}$, а потім знаходиться і невідоме ε

$$\varepsilon = \frac{\sigma}{\sqrt{n}} x_0. \quad (7)$$

Нагадаємо, що n – кількість спостережень, а σ – відоме.

Рівність (6) можна переписати в іншому вигляді, якщо скористатись поняттям **квантілі**.

Означення 1. *Квантіллю порядку p розподілу випадкової величини X називається число x_p , для якого $P X < x_p = F(x_p) = p$,* де $F(x)$ – функція розподілу.

Нехай x_p та x_{1-p} – квантілі деякого розподілу, щільність ймовірностей $f(x)$ якого симетрична відносно осі ординат. Рекомендується показати **самостійно**, що тоді $x_p = -x_{1-p}$.

Позначимо квантіль порядку p стандартизованого нормальногорозподілу $N(0, 1)$ через u_p . Тоді з рівності (6) випливає, що $\frac{\varepsilon\sqrt{n}}{\sigma} = u_{\frac{1+\beta}{2}}$. Отже,

$$\varepsilon = \frac{\sigma}{\sqrt{n}} u_{\frac{1+\beta}{2}}. \quad (8)$$

Для квантілів розподілу $N(0, 1)$ складена таблиця (див. Додаток 2, табл. 3). Враховуючи (8), довірчий інтервал для математичного сподівання можна записати у вигляді

$$\tilde{m} - \frac{\sigma}{\sqrt{n}} u_{\frac{1+\beta}{2}} < m < \tilde{m} + \frac{\sigma}{\sqrt{n}} u_{\frac{1+\beta}{2}}. \quad (9)$$

На практиці більш природною є ситуація, коли обидва параметри m та σ нормальногорозподілу невідомі. З метою побудови довірчих інтервалів у цьому випадку познайомимось спочатку з розподілами χ^2 ("xi"- квадрат) та Стьюдента, які, як і нормальногорозподіл випадкової величини, відіграють важливу роль у математичній статистиці.

Нехай X_1, X_2, \dots, X_n – незалежні нормальногорозподілені за законом $N(0, 1)$ випадкові величини.

Означення 2. Розподіл випадкової величини

$$\chi^2 = X_1^2 + X_2^2 + \dots + X_n^2 \quad (10)$$

називається розподілом χ^2 ("xi" - квадрат) з n ступенями вільності і позначається $\chi^2(n)$.

З означення випливає, що щільність $f(x)$ розподілу $\chi^2(n)$ лежить у 1-му квадранті (з виразом для функції $f(x)$ та процедурою його отримання можна познайомитися у спеціальній літературі з математичної статистики).

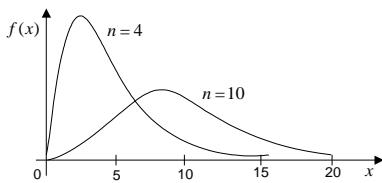


Рис. 1

На рис.1 наведені криві розподілу $\chi^2(n)$ для $n = 4$ та $n = 10$. Для всіх n крива $f(x)$ – це одновершинна асиметрична крива, що асимптотично наближається

до осі Ox . Зі збільшенням n вершина кривої $f(x)$ зсувається вправо від початку координат і розподіл $\chi^2(n)$ прямує до нормальногорозподілу. При $n > 30$ розподіл $\chi^2(n)$ з достатньою для практичних цілей точністю апроксимується нормальним розподілом. Математичне сподівання та дисперсія розподілу $\chi^2(n)$ відповідно дорівнюють:

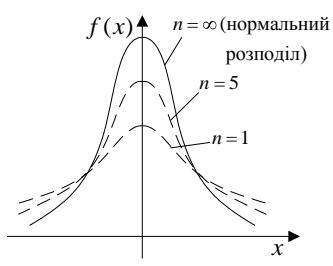
$$M[\chi^2(n)] = n, \quad D[\chi^2(n)] = 2n.$$


Рис. 2

Означення 3. Розподілом Стьюдента з n ступенями вільності називається розподіл випадкової величини $T(n)$, що дорівнює відношенню двох незалежних випадкових величин X та $\sqrt{\frac{\chi^2(n)}{n}}$, тобто

$$T(n) = \frac{X}{\sqrt{\frac{\chi^2(n)}{n}}} = \frac{X\sqrt{n}}{\sqrt{\chi^2(n)}}, \quad (11)$$

де випадкова величина X має нормальній розподіл $N(0;1)$.

Щільність розподілу Стьюдента симетрична відносно осі Oy (рис. 2), а тому $t_p(n) = -t_{1-p}(n)$, де $t_p(n)$ – квантіль розподілу Стьюдента порядку p . Зі збільшенням числа ступенів вільності n розподіл $T(n)$ наближається до нормальногорозподілу швидше, ніж розподіл $\chi^2(n)$. При $n > 30$ можна користуватись наближеною рівністю $t_p(n) \approx u_p$ для квантілів. Математичне сподівання та дисперсія

розподілу $T(n)$ дорівнюють: $M[T(n)] = 0$, $D[T(n)] = \frac{n}{n-2}$, $n > 2$.

Для квантілів розподілів $\chi^2(n)$ та $T(n)$ складені таблиці (див. Додаток 2, табл. 4, 5).

Сформулюємо та приймемо без доведення таку теорему.

Теорема. Нехай X_1, X_2, \dots, X_n – вибірковий вектор з нормальнорозподіленої генеральної сукупності $N(m, \sigma^2)$. Тоді вибіркове середнє

$$\tilde{m} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \text{ та виправлена вибіркова дисперсія}$$

$\tilde{D} = S^2 = \frac{1}{n-1} \sum_{i=1}^n X_i - \bar{X}^2$ є незалежними випадковими величинами,

причому статистика $\frac{n-1}{\sigma^2} S^2$ має розподіл $\chi^2(n-1)$.

Розглянемо декілька прикладів використання цієї теореми.

Приклад 1. Нехай за вибіркою з нормальним розподіленої генеральної сукупності $N(m, \sigma)$ розглядається статистика

$$Y = \frac{\bar{X} - m}{\sqrt{\frac{S^2}{n}}}, \quad (12)$$

де \bar{X} та $S = \sqrt{D}$ – величини, що фігурують у наведеній вище теоремі. Покажемо, що статистика Y має розподіл $T(n-1)$.

Дійсно, випадкова величина $U = \frac{\bar{X} - m}{\sqrt{\frac{S^2}{n}}}$ розподілена за законом

$N(0, 1)$, оскільки $M[\bar{X}] = m$, а $D[\bar{X}] = \frac{\sigma^2}{n}$. Випадкова величина

$V = \frac{n-1}{\sigma^2} S^2$, згідно з теоремою, має розподіл $\chi^2(n-1)$. Отже,

враховуючи означення (11), одержуємо для статистики

$\frac{U}{\sqrt{V}} \sqrt{n-1} = \frac{\bar{X} - m}{\sqrt{\frac{S^2}{n}}} = Y$ розподіл Стьюдента з $n-1$ ступенями

вільності.

Приклад 2. Нехай в умовах попереднього прикладу розглядається статистика $D^* = \frac{1}{n} \sum_{i=1}^n X_i - \bar{X}^2$ (вибіркова дисперсія). Рекомендується

показати **самостійно**, що статистики $\frac{n}{\sigma^2} D^*$ та $\frac{\bar{X} - m}{\sqrt{D^*} \sqrt{n-1}}$ мають

розподіли відповідно $\chi^2(n-1)$ та $T(n-1)$.

Розглянемо ще один приклад, який безпосередньо не має відношення до сформульованої вище теореми, але буде використовуватись у наступній лекції.

Приклад 3. За вибіркою з генеральної сукупності з розподілом $N(m, \sigma)$ розглядається статистика $S_0^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$. Покажемо, що статистика

$$Y = \frac{nS_0^2}{\sigma^2} \quad (13)$$

розподілена за законом $\chi^2(n)$.

Для цього запишемо очевидну рівність

$$\frac{nS_0^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - m}{\sigma} \right)^2 = \sum_{i=1}^n U_i^2. \quad (14)$$

Оскільки випадкові величини $U_i = \frac{X_i - m}{\sigma}$, $i = 1, 2, \dots, n$ розподілені за законом $N(0, 1)$ і незалежні, то згідно з означенням (10) випадкова величина (14) має розподіл $\chi^2(n)$.

Лекція № 21

Приклади на побудову довірчих інтервалів

На практиці надійність оцінки β беруть такою, щоб подію з ймовірністю β можна було вважати практично вірогідною, наприклад, $\beta = 0,95; 0,99; 0,999$. Іноді величину β записують у вигляді $\beta = 1 - \alpha$. Мале число α називається рівнем значущості оцінки.

Нехай $y_{\frac{\alpha}{2}}$ та $y_{1-\frac{\alpha}{2}}$ – квантилі розподілу деякої неперервної випадкової величини Y порядку $\frac{\alpha}{2}$ та $1 - \frac{\alpha}{2}$. Очевидно, що тоді

$$P\left\{ y_{\frac{\alpha}{2}} < Y < y_{1-\frac{\alpha}{2}} \right\} = F\left(y_{1-\frac{\alpha}{2}}\right) - F\left(y_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha. \quad (1)$$

Познайомимось з одним із можливих методів побудови довірчих інтервалів. Нехай \tilde{a} – незміщена оцінка параметра a розподілу деякої випадкової величини X . Допустимо, що існує статистика $Y = Y(\tilde{a}, a)$, для якої:

- 1) закон розподілу відомий і не залежить від a ;
- 2) функція $Y = Y(\tilde{a}, a)$ неперервна і строго монотонна відносно a .

Нехай $\beta = 1 - \alpha$ – задана довірча ймовірність, а $y_{\frac{\alpha}{2}}$ та $y_{1-\frac{\alpha}{2}}$ – квантілі розподілу статистики Y . Тоді на підставі рівності (1) з ймовірністю $\beta = 1 - \alpha$ виконуються нерівності

$$y_{\frac{\alpha}{2}} < Y < y_{1-\frac{\alpha}{2}}. \quad (2)$$

Довірчий інтервал для параметра a можна побудувати, розв’язавши нерівності (2) відносно a .

Приклад 1. Довірчий інтервал для математичного сподівання m нормально розподіленої генеральної сукупності при відомій дисперсії σ^2 .

Даний довірчий інтервал побудовано у попередній лекції. Отримаємо його з використанням нерівностей (2). Нехай надійність $\beta = 1 - \alpha$ задано. Розглянемо статистику Y

$$Y = \frac{\bar{X} - m}{\sigma / \sqrt{n}}, \quad \bar{X} = \tilde{m} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (3)$$

Легко показати, що вона розподілена за законом $N(0, 1)$ і задовольняє сформульованим вище вимогам. Отже,

$$u_{\frac{\alpha}{2}} < \frac{\bar{X} - m}{\sigma / \sqrt{n}} < u_{1-\frac{\alpha}{2}}, \quad (4)$$

де u_p – квантіль порядку p розподілу $N(0, 1)$. Розв’язуючи нерівності (4), одержуємо

$$\bar{X} - \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} < m < \bar{X} + \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}}. \quad (5)$$

Враховуючи рівність $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$, довірчий інтервал запишемо у вигляді

$$\bar{X} - \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} < m < \bar{X} + \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}}. \quad (6)$$

Оскільки $1 - \frac{\alpha}{2} = \frac{1+\beta}{2}$, $\tilde{m} = \bar{X}$, нерівності (6) з точністю до позначень співпадають з нерівностями (9) лекції 20.

Приклад 2. Довірчий інтервал для математичного сподівання нормально розподіленої генеральної сукупності при невідомій дисперсії σ^2 .

Скористаємось статистикою (12) лекції 20. Отже,

$$t_{\frac{\alpha}{2}} n-1 < \frac{\bar{X} - m}{S / \sqrt{n}} < t_{1-\frac{\alpha}{2}} n-1 . \quad (7)$$

Розв'язавши (7) відносно m і врахувавши рівність для квантілів $t_{\frac{\alpha}{2}} n-1 = -t_{1-\frac{\alpha}{2}} n-1$, одержуємо довірчий інтервал

$$\bar{X} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}} n-1 < m < \bar{X} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}} n-1 . \quad (8)$$

Нагадаємо, що тут $S^2 = \frac{1}{n-1} \sum_{i=1}^n X_i - \bar{X}^2$.

Якщо порівняти довірчі інтервали (6) та (8), то можна відмітити, що при заміні дисперсії σ^2 її оцінкою S^2 нормальній розподіл слід замінити на розподіл Стьюдента.

Приклад 3. Довірчий інтервал для дисперсії σ^2 нормальню розподіленої генеральної сукупності при відомому математичному сподіванні m .

Скористаємось статистикою (13) лекції 20. Згідно з (2) маємо

$$\chi_{\frac{\alpha}{2}}^2 n < \frac{nS_0^2}{\sigma^2} < \chi_{1-\frac{\alpha}{2}}^2 n , \quad (9)$$

де $\chi_p^2 n$ – квантіль розподілу χ^2 з n ступенями вільності.

Розв'язуючи (9) відносно σ^2 , одержуємо довірчий інтервал

$$\frac{nS_0^2}{\chi_{1-\frac{\alpha}{2}}^2 n} < \sigma^2 < \frac{nS_0^2}{\chi_{\frac{\alpha}{2}}^2 n} . \quad (10)$$

Тут $S_0^2 = \frac{1}{n} \sum_{i=1}^n X_i - m^2$.

Приклад 4. Довірчий інтервал для дисперсії σ^2 нормальню розподіленої генеральної сукупності при невідомому математичному сподіванні m .

Згідно з теоремою лекції 20 скористаємось статистикою $\frac{n-1}{\sigma^2} S^2$.

Тоді

$$\chi_{\frac{\alpha}{2}}^2 n-1 < \frac{n-1 S^2}{\sigma^2} < \chi_{1-\frac{\alpha}{2}}^2 n-1 \quad (11)$$

і для довірчого інтервалу одержуємо

$$\frac{n-1 S^2}{\chi_{1-\frac{\alpha}{2}}^2 n-1} < \sigma^2 < \frac{n-1 S^2}{\chi_{\frac{\alpha}{2}}^2 n-1}. \quad (12)$$

Для середнього квадратичного відхилення маємо

$$\frac{\sqrt{n-1}S}{\sqrt{\chi_{1-\frac{\alpha}{2}}^2 n-1}} < \sigma < \frac{\sqrt{n-1}S}{\sqrt{\chi_{\frac{\alpha}{2}}^2 n-1}}. \quad (13)$$

Як і раніше,

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n X_i - \bar{X}^2}. \quad (14)$$

Приклад 5. Довірчий інтервал для математичного сподівання генеральної сукупності, розподіл якої не є нормальним.

Статистика

$$\bar{m} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (15)$$

З точністю до невипадкового множника $\frac{1}{n}$ є сумою n незалежних та

однаково розподілених випадкових величин X_i . Тому, яким би не був закон розподілу генеральної сукупності X , для достатньо великих n закон розподілу цієї статистики згідно з центральною граничною теоремою теорії ймовірностей близький до нормального з математичним сподіванням $M[\bar{X}] = m$ та дисперсією $D[\bar{X}] = \frac{\sigma^2}{n}$, а

отже, середнім квадратичним відхиленням $\sigma[\bar{X}] = \frac{\sigma}{\sqrt{n}}$. Тут m та σ – характеристики генеральної сукупності X . Отже, статистика

$$Y = \frac{\bar{X} - m}{\sigma/\sqrt{n}} \quad (16)$$

для достатньо великих n має розподіл, близький до розподілу $N(0, 1)$. Тому з надійністю, близькою до $\beta = 1 - \alpha$, має місце довірчий інтервал (6). На практиці, зазвичай, величина σ є невідомою. Тому в (6) її замінюють оцінкою (14). Якщо n достатньо велике, така заміна мало впливає на точність довірчого інтервалу та його надійність.

Отже, для великих n можна користуватись таким наближенім

довірчим інтервалом для математичного сподівання генеральної сукупності, розподіл якої не є нормальним:

$$\bar{X} - \frac{S}{\sqrt{n}} u_{\frac{1-\alpha}{2}} < m < \bar{X} + \frac{S}{\sqrt{n}} u_{\frac{1-\alpha}{2}}. \quad (17)$$

Довірча ймовірність при цьому близька до $\beta = 1 - \alpha$, тобто

$$P\left\{\bar{X} - \frac{S}{\sqrt{n}} u_{\frac{1-\alpha}{2}} < m < \bar{X} + \frac{S}{\sqrt{n}} u_{\frac{1-\alpha}{2}}\right\} \approx \beta = 1 - \alpha. \quad (18)$$

Приклад 6. Довірчий інтервал для ймовірності успіху в схемі Бернуллі.

Задача ставиться так: нехай в n незалежних випробуваннях успіх з'явився X разів. Необхідно побудувати довірчий інтервал для ймовірності p успіху в одному випробуванні. Як ми вже знаємо, випадкова величина X має біномний розподіл. Разом з тим, для достатньо великих n згідно з інтегральною теоремою Муавра-Лапласа (див. лекцію 6) має місце така наближена рівність

$$P\left\{x_1 \leq \frac{X - np}{\sqrt{npq}} < x_2\right\} \approx \Phi(x_2) - \Phi(x_1), \quad q = 1 - p, \quad (19)$$

де $\Phi(x)$ – функція стандартизованого нормального розподілу $N(0, 1)$.

Рекомендується скористатись наближеною рівністю з лекції 6 і одержати рівність (19) **самостійно**.

З (19) випливає, що для великих n випадкова величина

$$Y = \frac{X - np}{\sqrt{npq}} = \frac{\frac{X}{n} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{p^* - p}{\sqrt{n}}}{\sqrt{\frac{pq}{n}}} \quad (20)$$

розподілена приблизно за нормальним законом $N(0, 1)$. Тут ми ввели відносну частоту успіху $p^* = \frac{X}{n}$, яка є змістовою та незміщеною оцінкою ймовірності p успіху в одному випробуванні. Те, що ця оцінка є змістовою, випливає з теореми Бернуллі. Незміщеність $M[p^*] = p$ є наслідком розподілу випадкової величини (20) за законом $N(0, 1)$.

Рекомендується показати **самостійно**, що статистика (20) є строго монотонною функцією p .

3 (2) одержуємо нерівності

$$u_{\frac{\alpha}{2}} < \frac{p^* - p}{\sqrt{pq} / \sqrt{n}} < u_{\frac{1-\alpha}{2}}, \quad (21)$$

які після елементарних перетворень подамо у вигляді

$$p^* - \sqrt{\frac{pq}{n}} u_{\frac{1-\alpha}{2}} < p < p^* + \sqrt{\frac{pq}{n}} u_{\frac{1-\alpha}{2}}. \quad (22)$$

Звичайно, ці нерівності виконуються з ймовірністю, що наближено дорівнює надійності $\beta = 1 - \alpha$.

Далі, замінимо в (22) значення p та q під коренем їх оцінками відповідно p^* та $1 - p^*$. В результаті одержуємо такий наближений довірчий інтервал для ймовірності успіху в схемі Бернуллі:

$$p^* - \sqrt{\frac{p^* (1-p^*)}{n}} u_{\frac{1-\alpha}{2}} < p < p^* + \sqrt{\frac{p^* (1-p^*)}{n}} u_{\frac{1-\alpha}{2}}. \quad (23)$$

Якщо n має порядок сотень, нерівність (23) дає задовільні результати.

У всіх наведених вище прикладах оцінка \tilde{a} параметра a розглядалась як випадкова величина. У практичному використанні одержаних довірчих інтервалів ці оцінки, звичайно, конкретизуються за результатами дослідів.

Розглянемо приклад на використання довірчого інтервалу (23).

Приклад 7. Нехай при 600 підкиданнях дефектної (несиметричної) монети герб випав 240 разів. За даною інформацією необхідно: а) побудувати 95% наближений довірчий інтервал для ймовірності p випадання герба при одному підкиданні; б) наблизено оцінити, яку мінімальну кількість підкидань монети необхідно зробити, щоб з надійністю 95% можна було стверджувати, що відносна частота появи герба відхиляється від ймовірності p не більше ніж на 1%.

Розв'язання: а) Відносна частота за даними досліду дорівнює $p^* = \frac{240}{600} = 0,4$; надійність $\beta = 1 - \alpha = 0,95$. За таблицею квантілів

нормального розподілу $N(0,1)$ знаходимо $u_{\frac{1-\alpha}{2}} = u_{0,975} = 1,96$. Тоді за формулою (23) наближений довірчий інтервал для p матиме вигляд $0,36 < p < 0,44$.

б) Довірчий інтервал (23) запишемо у вигляді

$$\left| p^* - p \right| < \sqrt{\frac{p^* (1-p^*)}{n}} u_{1-\frac{\alpha}{2}}.$$

Остання нерівність виконується з ймовірністю, близькою до $\beta = 1 - \alpha = 0,95$. За умовою $|p^* - p| \leq 0,01$, а тому для знаходження n одержуємо нерівність

$$\sqrt{\frac{p^* (1-p^*)}{n}} u_{0,975} = \sqrt{\frac{0,4 \cdot 0,6}{n}} 1,96 \leq 0,01.$$

Отже, $n \geq 196^2 \cdot 0,24 = 9219,84$. Тому мінімальна кількість підкідань $n = 9220$.

Розглянутий приклад показує, яку величезну кількість випробувань необхідно здійснити, щоб із задовільною точністю оцінити ймовірність події за її відносною частотою. На практиці, зазвичай, використовують значно меншу кількість випробувань. Тому необхідно пам'ятати, що точність наближеної рівності $p^* \approx p$, як правило, невелика.

Лекція № 22

Основні поняття статистичної перевірки статистичних гіпотез

У природознавстві, техніці, економіці досить часто при вивченні того чи іншого випадкового явища використовують різного роду припущення (гіпотези), які можна перевірити статистично, тобто спираючись на результати спостережень у вигляді випадкової вибірки. Таким, наприклад, є припущення, що розподіл продуктивності праці робітників, які працюють в однакових організаційно-технічних умовах, описується нормальним законом або припущення, що середні розміри деталей, які виготовляються на однотипних паралельно працюючих верстатах, не розрізняються між собою.

Під **статистичною гіпотезою** далі будемо розуміти довільне припущення щодо вигляду невідомого розподілу випадкової величини або щодо параметрів відомих розподілів. В останньому випадку гіпотеза називається **параметричною**. У даний і наступній лекціях будуть розглядатися лише параметричні гіпотези.

Статистична гіпотеза називається **простою**, якщо вона однозначно визначає розподіл випадкової величини. У протилежному разі гіпотеза називається **складною**. Наприклад, припущення, що випадкова

величина X має нормальній розподіл $N(0, 1)$ є простою гіпотезою, тоді як гіпотеза “випадкова величина X розподілена за законом $N(m, 1)$, де $m_1 \leq m \leq m_2$ ” – складна.

До простих гіпотез відносяться, наприклад, такі припущення:

1. Параметр відомого однопараметричного розподілу, наприклад, параметр λ показникового розподілу, дорівнює деякому фіксованому числу λ_0 .

2. Математичне сподівання нормального розподілу дорівнює 2 (середнє квадратичне відхилення σ при цьому відоме).

Якщо у другому прикладі σ – невідоме, то гіпотеза є складною.

Гіпотезу, яку перевіряють, називають **нульовою** або **основною** гіпотезою і позначають H_0 . Для параметричної нульової гіпотези H_0 розглядають одну з **альтернативних** (конкурентних) гіпотез H_1 . Наприклад, якщо перевіряють гіпотезу про рівність параметра a деякому фіксованому значенню a_0 (будемо писати $H_0 : a = a_0$), то альтернативною може бути одна з таких гіпотез: $H_1 : a > a_0$; $H_1 : a < a_0$; $H_1 : a \neq a_0$, зокрема проста альтернативна гіпотеза $H_1 : a = a_1$, де a_1 – фіксоване число, $a_1 \neq a_0$.

Зauważення 1. Математична статистика не дає ніяких рекомендацій щодо вибору основної та альтернативної гіпотез. Цей вибір повністю визначається дослідником і залежить від конкретної задачі. Математична статистика пропонує методи перевірки статистичних гіпотез, при застосуванні яких можна приймати лише одну з гіпотез H_0 або H_1 , одночасно відхиляючи іншу.

Оскільки рішення про прийняття гіпотези H_0 чи її відхилення на користь альтернативної гіпотези H_1 повинно прийматися на основі вибірки, що містить n незалежних спостережень X_1, X_2, \dots, X_n над випадковою величиною X , то зрозуміло, що множина всіх можливих вибірок обсягу n повинна бути певним чином розділена на дві підмножини, які не перетинаються (позначимо ці підмножини вибірок через S та \bar{S}) і такі, що гіпотеза H_0 відхиляється, якщо отримана в експерименті конкретна вибірка належить до однієї з цих підмножин (nehай це підмножина S), і приймається, якщо – до іншої підмножини (підмножини \bar{S}).

На практиці реалізація даного підходу до перевірки статистичних гіпотез проводиться за допомогою правильно підібраної одновимірної

випадкової величини (статистики) $Z = Z(X_1, X_2, \dots, X_n)$ з відомим за умови справедливості гіпотези H_0 розподілом. Ця статистика відображає множину всіх можливих вибірок обсягу n на деяку множину V числової осі R так, що образами вказаних вище множин вибірок S та \bar{S} є відповідно числові множини $V_k \subset V$ та $\bar{V}_k = V \setminus V_k$. Тоді висновок про відхилення чи прийняття гіпотези H_0 може прийматись за допомогою вибікового значення z_e статистики Z , обчисленого за результатами конкретної вибірки x_1, x_2, \dots, x_n . А саме: відхилити гіпотезу H_0 , якщо $z_e \in V_k$; прийняти гіпотезу H_0 , якщо $z_e \in \bar{V}_k$.

Множину V_k називають **критичною областю**, а множину \bar{V}_k – **областю прийняття гіпотези**. Саму статистику Z , яка використовується для перевірки нульової гіпотези, називають **статистичним критерієм** (або просто критерієм).

Іноді поняття “статистичний критерій” наділяють більш широким смыслом, розуміючи під ним вцілому все правило, згідно з яким приймається рішення про прийняття чи відхилення гіпотези H_0 . Статистику Z у цьому випадку називають статистикою критерію, а область V_k – критичною областю критерію.

При перевірці простої параметричної гіпотези $H_0: a = a_0$ використовують, як правило, ту ж статистику, що й при оцінці параметра a , тобто статистику $\tilde{a} = \tilde{a}(X_1, X_2, \dots, X_n)$.

Отже, задача перевірки гіпотези H_0 з використанням статистики Z зводиться до побудови критичної області V_k . Ясна річ, що це можна зробити не єдиним способом. Виникає питання, якими принципами слід користуватись при побудові області V_k , щоб вона була найбільш “случиною” для даної гіпотези H_0 ? Ці принципи були сформульовані в роботах відомих математиків Неймана та Пірсона. При побудові критичної області V_k слід мати на увазі, що приймаючи чи відхиляючи гіпотезу H_0 , можна припуститись помилок двох типів.

Нехай деяка чисрова множина V_k розглядається як критична область при перевірці гіпотези H_0 з використанням статистики Z : якщо $z_e \in V_k$, то гіпотеза H_0 відхиляється; якщо $z_e \notin V_k$ – приймається. При цьому можлива одна з чотирьох ситуацій у

залежності від того, справедлива гіпотеза H_0 чи ні, попадає вибіркове значення z_e статистики Z в область V_k чи ні. Докладніше:

1. Гіпотеза H_0 справедлива; $z_e \notin V_k$. Гіпотеза H_0 приймається.
2. Гіпотеза H_0 несправедлива; $z_e \notin V_k$. Гіпотеза H_0 приймається.
3. Гіпотеза H_0 справедлива; $z_e \in V_k$. Гіпотеза H_0 відхиляється.
4. Гіпотеза H_0 несправедлива; $z_e \in V_k$. Гіпотеза H_0 відхиляється.

З чотирьох ситуацій дві (1 та 4) – задовільні і дві (2 та 3) – нездовільні. У випадках 2 та 3 говорять, що внаслідок використання критичної області V_k допускається помилка. У випадку 2: приймається несправедлива гіпотеза; у випадку 3: відхиляється справедлива гіпотеза. Ці помилки мають свої назви.

Означення 1. Помилка, яка полягає в тому, що гіпотеза H_0 відхиляється, коли вона справедлива, називається ***помилкою першого роду***.

Означення 2. Помилка, яка полягає в тому, що гіпотеза H_0 приймається, коли вона несправедлива, називається ***помилкою другого роду***.

Чи можна побудувати критичну область V_k так, щоб уникнути вказаних помилок? Ні, не можна. Якою б не була критична область $V_k \neq \emptyset$, вибіркове значення z_e через випадковість вибірки може потрапити при справедливій гіпотезі H_0 до V_k , що приведе до помилки першого роду. При несправедливій гіпотезі H_0 значення z_e може не потрапити до V_k , що приведе до помилки другого роду. Отже, позбутися помилок в принципі неможливо. Оскільки помилки першого та другого родів є випадковими, то можна говорити про ймовірності цих помилок. Тому природно намагатися побудувати таку критичну область V_k , використовуючи яку матимемо мінімально можливі ймовірності помилок. Подивимося, чи можна це зробити.

Розглянемо ***просту*** параметричну гіпотезу $H_0 : a = a_0$ та альтернативну їй ***просту*** гіпотезу $H_1 : a = a_1$, $a_1 \neq a_0$. Для наочності будемо вважати, що статистика Z , яка використовується для перевірки гіпотези H_0 , є неперервною випадковою величиною. Вище, говорячи про статистику Z , ми завжди мали на увазі її розподіл за умови справедливості гіпотези H_0 , що перевіряється. Позначимо у нашому випадку щільність розподілу статистики Z за справедливої гіпотези H_0

через $f(z|H_0)$, а через $f(z|H_1)$ – щільність розподілу цієї статистики за умови справедливості гіпотези H_1 . Вважаємо, цю щільність також відомою. Нехай тепер деяка неперервна числова множина V_k використовується як критична область.

Ймовірність помилки першого роду (позначимо цю ймовірність через α) дорівнює ймовірності того, що статистика Z за умови справедливості гіпотези H_0 попаде в критичну область V_k . Отже,

$$P(Z \in V_k | H_0) = \alpha. \quad (1)$$

Зауважимо, що $P(Z \in V_k | H_0)$ є ймовірністю події $Z \in V_k$, яка обчислюється при справедливій гіпотезі H_0 , тобто

$$P(Z \in V_k | H_0) = \int_{V_k} f(z|H_0) dz$$

і не має нічого спільного з умовною ймовірністю.

Ймовірність помилки другого роду (позначимо її через β) можна знайти (при простій гіпотезі H_1) за формулою

$$\beta = P(Z \notin V_k | H_1) = P(Z \in \bar{V}_k | H_1) = \int_{\bar{V}_k} f(z|H_1) dz. \quad (2)$$

Очевидно, що \bar{V}_k завжди можна побудувати, якщо відома область V_k . Отже, бачимо, що ймовірність помилок першого та другого роду однозначно визначається критичною областю V_k . Тому, вибираючи V_k , скажімо, за умови, що ймовірність α помилки першого роду мала, ми водночас одержуємо ймовірність β помилки другого роду такою, якою вона виходить. Можна вибирати V_k і за умови, що ймовірність помилки другого роду мала, але при цьому ймовірність помилки першого роду визначається вибраним V_k і буде такою, якою вийде. Отже, вибрати V_k так, щоб одночасно були контролювані ймовірності помилок першого та другого роду, не вдається. Тому будемо діяти наступним чином.

Як зазначалось раніше, нульову гіпотезу з сукупності всіх гіпотез ми вибираємо самі. Так ось, за нульову (основну) гіпотезу H_0 ми домовляємося вибирати ту, для якої важливіше уникнути помилки, що полягає в її відхиленні, коли вона справедлива. Помилка першого роду – це помилка, якої важливіше уникнути – це помилка, “ціна” якої вища.

Тому нашою **першою вимогою** щодо вибору критичної області V_k є вимога, щоб ймовірність помилки першого роду (1) була малою.

Означення 3. Ймовірність α помилки першого роду називається **рівнем значущості критерію**.

Отже, перед аналізом вибірки фіксується деяке мале число α і вибирається критична область V_k з умовою, що задовольняється рівність (1). Наскільки малим потрібно брати α , залежить від серйозності "наслідків" помилки першого роду. Зазвичай, вибирають одне з стандартних значень α , таких як 0,005, 0,01, 0,05, 0,1. Ця стандартизація має ту перевагу, що дозволяє скоротити об'єм таблиць, які використовуються статистиком у роботі. Ніякої іншої спеціальної причини для вибору саме цих значень немає. Якщо, наприклад, $\alpha = 0,01$, то це означає, що в одному випадку зі 100 є ризик допустити помилку першого роду (відхилити гіпотезу H_0 , коли вона справедлива).

Рівністю (1) при малих α реалізується так званий принцип практичної неможливості здійснитися малоймовірній випадковій події при одній спробі. Якщо вибіркове значення z_e статистики Z потрапляє до критичної області, а це подія малоймовірна і вона все-таки здійснилася, то нульова гіпотеза H_0 відхиляється. У противному разі – приймається.

Зауважимо, що прийняття гіпотези зовсім не означає її доведення. Дійсно, одиничний приклад, який підтверджує справедливість деякого твердження, ще не доводить це твердження. Тому при прийнятті гіпотези правильніше було б говорити, що експериментальні дані (вибірка) не суперечать основній гіпотезі і тому не дають підстав до її відхилення.

Означення 4. Ймовірність $P(Z \in V_k / H_1)$ відхилити основну гіпотезу H_0 , коли справедлива **проста** альтернативна гіпотеза H_1 , називають **потужністю критерію** відносно гіпотези H_1 .

Очевидно, що

$$P(Z \notin V_k / H_1) + P(Z \in V_k / H_1) = 1, \quad (3)$$

тобто сума ймовірності помилки другого роду і потужності критерію відносно простої гіпотези H_1 дорівнює одиниці.

Якщо для простої гіпотези $H_0: a = a_0$ розглядається складна альтернативна гіпотеза $H_1: a \in A$, де A – деяка підмножина множини значень параметра a , то очевидно, що за справедливості цієї складної гіпотези буде справедливою одна з простих гіпотез $H_1^{(a_1)}: a = a_1$, $a_1 \in A$. Для кожного $a \in A$ можна обчислити ймовірність помилки

другого роду $P Z \in V_k / H_1^{(a)} = \beta(a)$ або ймовірність $P Z \in V_k / H_1^{(a)} = 1 - \beta(a)$ відхилення основної гіпотези H_0 за умови справедливості простої гіпотези $H_1^{(a)}$.

Означення 5. Функція $M V_k; a = P Z \in V_k / H_1^{(a)}$, $a \in A$

називається функцією потужності критерію.

Якщо H_1 – проста альтернативна гіпотеза, причому $H_1: a = a_1$, то $M V_k; a_1 = 1 - \beta$ є потужністю критерію відносно H_1 . Тут $\beta = \beta(a_1)$.

Нехай, як раніше, статистика Z – неперервна випадкова величина зі щільністю $f(z/H_0)$. Зрозуміло, що на множині значень цієї статистики можна вибрати скільки завгодно критичних областей V_k для заданого рівня значущості α .

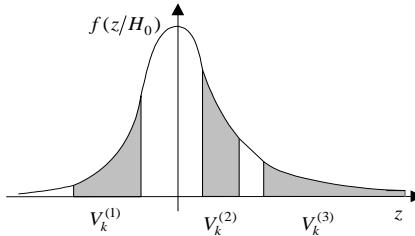


Рис. 1

Так, на рис. 1, наприклад, зображені три області $V_k^{(1)}, V_k^{(2)}, V_k^{(3)}$ на числовій осі, в які статистика Z попадає з однаковою ймовірністю α (площи заштрихованих областей дорівнюють α). Проте ймовірності помилок другого роду для всіх цих областей, взагалі кажучи, різні.

Означення 6. Найкращою критичною областю (НКО) називають таку критичну область, яка для заданого рівня значущості α забезпечує мінімальну ймовірність помилки другого роду.

Враховуючи (3), можна стверджувати, що критерій перевірки простої гіпотези H_0 проти простої альтернативної гіпотези H_1 , в якому використовується НКО, має максимальну потужність.

Щоб критична область була найкращою, інтуїтивно представляється найбільш природним вибрати її якомога “широкою”, наскільки це можливо (див. (2)). Точну відповідь на питання про НКО для заданого рівня значущості α дає теорема Неймана-Пірсона, з якою можна познайомитися в повних курсах математичної статистики. Ми лише відмітимо, що для тих критеріїв перевірки гіпотез, які будуть

розділяється нижче, НКО згідно з теоремою Неймана-Пірсона розміщується на “хвостах” розподілів статистик даних критерій, як то, наприклад, область $V_k^{(3)}$ на рис. 1. При цьому положення критичної області V_k на множині значень статистики Z залежить від вигляду альтернативної гіпотези H_1 . Якщо, наприклад, перевіряється гіпотеза $H_0 : a = a_0$, а альтернативною гіпотезою є гіпотеза $H_1 : a < a_0$ (зокрема, проста альтернативна гіпотеза $H_1 : a = a_1, a_1 < a_0$), то критична область розміщується на лівому “хвості” розподілу статистики Z (рис. 2). Враховуючи (1) і пригадавши означення квантілі, цю область можна подати у вигляді нерівності $Z < z_\alpha$, де z_α – квантіль порядку α розподілу статистики Z за умови справедливості гіпотези H_0 . Якщо альтернативною гіпотезою є гіпотеза $H_1 : a > a_0$ (зокрема, $H_1 : a = a_1, a_1 > a_0$), то критична область розміщується на правому “хвості” розподілу статистики Z (рис. 3) і має вигляд $Z > z_{1-\alpha}$.

У розглянутих випадках критерій називається *одностороннім*, відповідно лівостороннім та правостороннім.

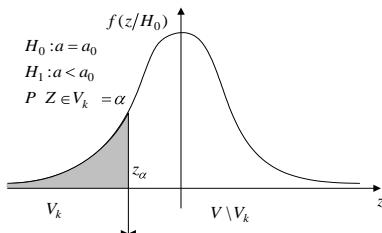


Рис. 2

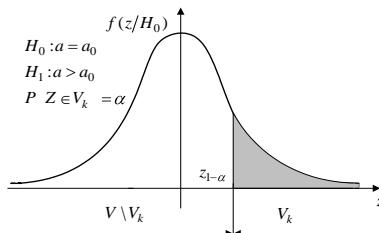


Рис. 3

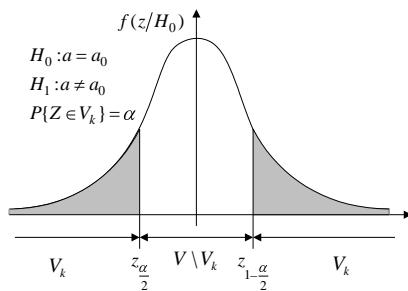


Рис. 4

Якщо альтернативна гіпотеза формулюється як $H_1 : a \neq a_0$, то критична область розміщується на двох “хвостах” розподілу статистики Z (рис. 4). Максимальна потужність критерію

забезпечується тоді, коли V_k визначається системою нерівностей
 $Z < z_{\frac{\alpha}{2}}$ і $Z > z_{1-\frac{\alpha}{2}}$. Критерій у цьому випадку називається
двостороннім.

Зauważення 2. Розглянемо для випадку лівостороннього критерію (рис. 2) ймовірність $p = P\{Z < z_\alpha\} = F(z_\alpha)$, де $F(x)$ – функція розподілу статистики Z , а z_α – вибіркове значення цієї статистики. Оскільки від вибірки до вибірки z_α змінюється, значення p є випадковою величиною.

Знайшовши p для конкретної вибірки, можна прийняти чи відхилити гіпотезу H_0 для будь-якого рівня значущості α шляхом простого порівняння p і α , а саме:

$$\begin{aligned} \text{якщо } \alpha < p, \text{ то гіпотеза } H_0 \text{ приймається}; \\ \text{якщо } \alpha > p, \text{ то гіпотеза } H_0 \text{ відхиляється}. \end{aligned} \quad (4)$$

Процедура порівняння (4) залишається справедливою і у випадку правостороннього критерію (рис. 3), якщо покласти $p = P\{Z \geq z_\alpha\} = 1 - P\{Z < z_\alpha\} = 1 - F(z_\alpha)$.

Для двостороннього критерію (рис. 4) візьмемо

$$p = \min(F(z_\alpha); 1 - F(z_\alpha)). \quad (5)$$

Тоді перевірка гіпотези H_0 зведеться до такої процедури порівняння:

$$\begin{aligned} \text{якщо } \frac{\alpha}{2} < p, \text{ то гіпотеза } H_0 \text{ приймається}; \\ \text{якщо } \frac{\alpha}{2} > p, \text{ то гіпотеза } H_0 \text{ відхиляється}. \end{aligned} \quad (6)$$

Отже, чим менше значення p , тим сильніше свідчить вибірка проти гіпотези H_0 . Якщо користуватись значеннями p , то немає потреби перед аналізом вибірки фіксувати певний рівень значущості α , як це передбачалось раніше. Число p має назву фактично досягнутого рівня значущості на вибірці. Про поняття досягнутого рівня значущості можна прочитати, наприклад, в книзі А.А. Боровкова “Математична статистика”, 1984, гл. 3, §4.

Нижче, як і до цього зауваження, ми будемо використовувати фіксований рівень значущості α , оперуючи не розподілами статистик, а їхніми квантілями.

Лекція № 23

Перевірка гіпотез про параметри нормальну розподілених випадкових величин

З врахуванням викладеного у попередній лекції, перевірка параметричної статистичної гіпотези зводиться до таких етапів:

- 1) формулюються основна H_0 та альтернативна H_1 гіпотези;
- 2) вибирається рівень значущості α ;
- 3) вибирається статистика критерію Z і знаходиться її розподіл за умови, що справедлива гіпотеза H_0 ;
- 4) залежно від вигляду альтернативної гіпотези визначається критична область V_k однією з нерівностей $Z < z_\alpha$, $Z > z_{1-\alpha}$ або системою нерівностей $Z < z_{\frac{\alpha}{2}}$ і $Z > z_{\frac{1-\alpha}{2}}$;
- 5) отримується вибірка спостережень і обчислюється вибіркове значення z_e статистики критерію;
- 6) приймається статистичне рішення: якщо $z_e \in V_k$, то гіпотеза H_0 відхиляється; якщо $z_e \notin V_k$, то – приймається.

Як правило, на етапах 3) – 6) використовується статистика, для якої значення квантилів табульовані: статистика з нормальним розподілом $N(0,1)$, статистика Стьюдента, статистика χ^2 – “ χ^2 ”-квадрат або статистика Фішера, з якою ми познайомимося нижче. Проте інтерпретацію результатів і обчислення ймовірностей помилок, що допускаються при перевірці гіпотези, зручно проводити для статистики, що є безпосередньою оцінкою параметра a , тобто статистики \tilde{a} .

Приклад 1. З нормально розподіленої генеральної сукупності з відомим середнім квадратичним відхиленням $\sigma = 2$ і невідомим математичним сподіванням m отримана вибірка обсягу $n = 25$. За результатами цієї вибірки знайдено вибіркове середнє $\bar{x}_e = 9,3$. Необхідно для рівня значущості $\alpha = 0,05$ перевірити основну гіпотезу $H_0 : m = 10$ при альтернативній гіпотезі $H_1 : m < 10$.

Розв'язання. 1) Основна та альтернативна гіпотези сформульовані в умові задачі: $H_0 : m = 10$, $H_1 : m < 10$;
2) рівень значущості задано: $\alpha = 0,05$;
3) скористаємося оцінкою математичного сподівання – вибірковим

середнім $\tilde{m} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Оскільки вибірка отримується з нормальним розподіленої генеральної сукупності, статистика \bar{X} також має нормальний розподіл з дисперсією $\frac{\sigma^2}{n} = \frac{4}{25}$. За умови, що гіпотеза H_0 справедлива, математичне сподівання статистики \bar{X} дорівнює 10. За статистику критерію візьмемо нормовану статистику $U = \frac{\bar{X} - 10}{\sqrt{\frac{4}{25}}}$.

Очевидно, що вона має розподіл $N(0, 1)$;

- 4) враховуючи вигляд альтернативної гіпотези, критичну область задаємо нерівністю $U < u_\alpha$. За таблицею квантилів розподілу $N(0, 1)$ знаходимо $u_\alpha = u_{0,05} = -u_{0,95} = -1,645$;
- 5) вибіркове значення нормованої статистики критерію дорівнює $u_e = \frac{9,3 - 10}{\sqrt{\frac{4}{25}}} = -1,75$;
- 6) приймаємо статистичне рішення: оскільки $u_e < -1,645$, тобто $u_e \in V_k$, гіпотеза H_0 відхиляється.

Зauważення (див. зауваження 2 попередньої лекції). Для фактично досягнутого рівня значущості p на даній вибірці маемо $p = P\{U < -1,75\} = \Phi(-1,75) = 1 - \Phi(1,75) = 0,0401$. Оскільки $\alpha = 0,05 > 0,0401 = p$, гіпотезу H_0 слід відхилити.

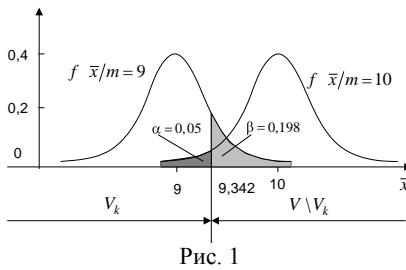
Границю \bar{x}_k критичної області для вихідної статистики \bar{X} можна отримати із співвідношення $\frac{\bar{x}_k - 10}{\sqrt{\frac{4}{25}}} = -1,645$. Отже, $\bar{x}_k = 9,342$, тобто

критична область для статистики \bar{X} визначається нерівністю $\bar{X} < 9,342$.

В умовах розглянутого прикладу знайдемо ймовірність помилки другого роду β для випадку простої альтернативної гіпотези $H_1: m = 9$. За умови, що ця гіпотеза справедлива, статистика \bar{X} має розподіл $N(9, \sqrt{\frac{4}{25}})$. Отже,

$$\beta = P \bar{X} \geq 9,342 / H_1 : m = 9 = 1 - \Phi \left(\frac{9,342 - 9}{\sqrt{\frac{4}{25}}} \right) = 1 - \Phi 0,855 \approx 0,198.$$

Тут $\Phi(x)$ – функція розподілу нормального закону $N(0,1)$. Ймовірності помилок першого та другого роду показані у вигляді заштрихованих площ під кривими щільностей розподілу статистики критерію на рис. 1.



Для заданої ймовірності α помилки першого роду ймовірність помилки другого роду може бути зменшена шляхом збільшення обсягу вибірки. Якщо при цьому ймовірність помилки другого роду не повинна перевищувати заданого значення β ,

то мінімальний обсяг вибірки n можна знайти, розв'язавши систему

$$P(Z \in V_k / H_0) = \alpha; P(Z \notin V_k / H_1) \leq \beta.$$

Аналітичний розв'язок цієї системи існує лише в найпростіших випадках.

Приклад 2. Який мінімальний обсяг вибірки n потрібно взяти в умовах прикладу 1, щоб при перевірці гіпотези $H_0 : m = 10$ при альтернативній гіпотезі $H_1 : m = 9$ ймовірність помилки першого роду дорівнювала $\alpha = 0,01$, а ймовірність помилки другого роду не перевищувала $0,1$? Яка критична область в цьому випадку?

Розв'язання. Критична область визначається нерівністю $\bar{X} < \bar{x}_k$.

За умовою

$$P(\bar{X} < \bar{x}_k / H_0 : m = 10) = \Phi \left(\frac{\bar{x}_k - 10}{\sqrt{\frac{4}{n}}} \right) = 0,01,$$

$$P(\bar{X} \geq \bar{x}_k / H_1 : m = 9) = 1 - \Phi \left(\frac{\bar{x}_k - 9}{\sqrt{\frac{4}{n}}} \right) \leq 0,1.$$

Перепишемо цю систему так:

$$\frac{\bar{x}_k - 10}{2} \sqrt{n} = u_{0,01} = -2,326; \quad \frac{\bar{x}_k - 9}{2} \sqrt{n} \geq u_{0,9} = 1,282.$$

Виключаючи \bar{x}_k , отримуємо $n > 53$. Підставляючи найменше значення $n = 53$ у перше рівняння системи, знаходимо границю критичної області: $\bar{x}_k = 10 - \frac{2 \cdot 2,326}{\sqrt{53}} = 9,361$. Отже, критична область визначається нерівністю $\bar{X} < 9,361$.

Приклад 3. Показати **самостійно**, що в умовах прикладу 1 гіпотезу $H_0: m = 10$ при альтернативній гіпотезі $H_1: m \neq 10$ слід прийняти.

Перевірка параметричної гіпотези $H_0: a = a_0$ при альтернативній гіпотезі $H_1: a \neq a_0$ може бути проведена з використанням довірчого інтервалу для параметра a . Дійсно, для статистики Z критерію з ймовірністю $1 - \alpha$ виконується нерівність

$$z_{\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}, \quad (1)$$

яка описує область прийняття гіпотези H_0 (саме цю нерівність ми використовували для побудови довірчих інтервалів в лекції 21). Розв'язуючи (1) відносно значення a_0 параметра a , отримаємо довірчий інтервал, який з надійністю $1 - \alpha$ покриває це значення.

Отже, щоб для рівня значущості α перевірити гіпотезу $H_0: a = a_0$ при альтернативній гіпотезі $H_1: a \neq a_0$, потрібно з надійністю $1 - \alpha$ побудувати довірчий інтервал для параметра a . Якщо значення a_0 покривається цим довірчим інтервалом, гіпотеза H_0 приймається, якщо ні – відхиляється.

Приклад 4. В умовах прикладу 1 перевірити гіпотезу $H_0: m = 10$ при альтернативній гіпотезі $H_1: m \neq 10$, скориставшись довірчим інтервалом для m .

Розв'язання. У нашому випадку довірчий інтервал для m має вигляд $\bar{X} - \frac{\sigma}{\sqrt{n}} u_{\frac{1-\alpha}{2}} < m < \bar{X} + \frac{\sigma}{\sqrt{n}} u_{\frac{1-\alpha}{2}}$ (див. (5) з лекції 21).

Підставляючи $\bar{X} = 9,3$; $\sigma = 2$; $n = 25$; $u_{\frac{1-\alpha}{2}} = u_{\frac{1-0,05}{2}} = u_{0,975} = 1,96$,

отримуємо $8,516 < m < 10,084$. Оскільки знайдений довірчий інтервал покриває значення $m = 10$, гіпотезу H_0 слід прийняти.

Приклад 5. За вибіркою обсягу $n = 20$, отриманою з нормально розподіленої генеральної сукупності з невідомими математичним

сподіванням m та дисперсією σ^2 , знайдені вибіркове середнє $\bar{x} = 16$ і виправлене вибіркове середнє квадратичне відхилення $s = 4,5$. Необхідно для рівня значущості $\alpha = 0,05$ перевірити гіпотезу $H_0: m = m_0 = 15$ при: 1) альтернативній гіпотезі $H_1: m > 15$; 2) альтернативній гіпотезі $H_1: m \neq 15$.

Розв'язання пропонується провести **самостійно**, скориставшись, як статистикою критерію, статистикою Стьюдента (12) з лекції 20

$$Z = T_{n-1} = \frac{\bar{X} - m_0}{\sqrt{\frac{s^2}{n}}}.$$

Приклад 6. Перевірити нульову гіпотезу H_0 з попереднього прикладу при альтернативній гіпотезі 2), скориставшись довірчим інтервалом для m .

Розв'язання провести **самостійно** (див. (8) з лекції 21).

Перевірка гіпотези про рівність центрів розподілу двох нормальним розподіленим генеральних сукупностей, дисперсії яких відомі

Розглянемо дві випадкові величини X та Y з розподілами відповідно $N(m_x, \sigma_x)$ та $N(m_y, \sigma_y)$. За умовою σ_x та σ_y – відомі. Потрібно за двома незалежними вибірками з обсягами n_1 та n_2 , отриманими відповідно з генеральних сукупностей X та Y , перевірити гіпотезу $H_0: m_x = m_y$. Оскільки про значення математичних сподівань m_x та m_y нам нічого невідомо, скористаємося їх найкращими оцінками $\tilde{m}_x = \bar{X}$ та $\tilde{m}_y = \bar{Y}$. Вибіркові середні \bar{X} та \bar{Y} мають закони розподілу відповідно $N\left(m_x, \frac{\sigma_x^2}{n_1}\right)$ та $N\left(m_y, \frac{\sigma_y^2}{n_2}\right)$.

Оскільки вибірки незалежні, то \bar{X} та \bar{Y} також незалежні і різниця $\bar{X} - \bar{Y}$ має нормальний розподіл, причому

$$D[\bar{X} - \bar{Y}] = D[\bar{X}] + D[\bar{Y}] = \frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}. \quad \text{Якщо гіпотеза } H_0: m_x = m_y$$

справедлива, то $M[\bar{X} - \bar{Y}] = m_x - m_y = 0$. Отже, статистика

$$Z = U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}} \tag{2}$$

є випадковою величиною з розподілом $N(0, 1)$.

Статистика (2) використовується як статистика критерію перевірки нульової гіпотези. Критична область для заданого рівня значущості α будеться в залежності від вигляду альтернативної гіпотези (див. лекцію 22). Для двостороннього критерію, наприклад, критичну область можна подати у вигляді

$$\frac{|\bar{x} - \bar{y}|}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}} > u_{\frac{1-\alpha}{2}}, \quad (3)$$

де \bar{x} та \bar{y} – вибіркові значення середніх \bar{X} та \bar{Y} , а $u_{\frac{1-\alpha}{2}}$ – квантіль

розділу $N(0,1)$.

Вибіркові значення \bar{x} та \bar{y} , як правило, не дорівнюють одне одному. Якщо в результаті перевірки гіпотеза $H_0: m_x = m_y$ приймається, то відмінність значень \bar{x} та \bar{y} називають **незначимою** (випадковою). Ця відмінність пояснюється випадковими причинами, зокрема випадковістю вибірок. Якщо ж гіпотеза $H_0: m_x = m_y$ відхиляється, то відмінність вибіркових значень \bar{x} та \bar{y} називають **значимою** (невипадковою). Це означає, що її не можна пояснити випадковими причинами, оскільки самі генеральні середні (математичні сподівання m_x і m_y) різні.

Приклад 7. У результаті двох серій вимірювань з кількістю вимірювань $n_1 = 25$ та $n_2 = 50$ отримані такі середні значення досліджуваної величини: $\bar{x} = 9,79$ і $\bar{y} = 9,60$. Чи можна з надійністю 0,99 (на рівні значущості $\alpha = 0,01$) пояснити таку розбіжність значень випадковими причинами, якщо відомо, що середні квадратичні відхилення в обох серіях вимірювань $\sigma_x = \sigma_y = 0,3$?

Розв'язання. Знайдемо вибіркове значення критерію (2):
 $|z_e| = \frac{|9,79 - 9,60|}{0,3 \sqrt{\frac{1}{25} + \frac{1}{50}}} = 2,59$. Квантіль $u_{\frac{1-\alpha}{2}} = u_{0,995} = 2,576$. Оскільки $|z_e| > u_{\frac{1-\alpha}{2}}$, то з надійністю 0,99 слід вважати розбіжність середніх невипадковим (значимим).

На практиці досить часто виникає задача порівняння точності приладів, інструментів, самих методів вимірювань тощо. Очевидно, що кращим є той прилад, інструмент і метод, який забезпечує найменше розсіювання результатів вимірювань. Якщо говорити мовою

математичної статистики, то ця задача означає порівняння дисперсій випадкових величин. Початок наступної лекції буде присвячений перевірці гіпотези про рівність дисперсій двох нормальних генеральних сукупностей. Статистикою критерію перевірки даної гіпотези є статистика Фішера. Тому познайомимося спочатку з **розділом Фішера**.

Означення. Розподілом Фішера з n_1 і n_2 ступенями вільності називається розподіл випадкової величини F_{n_1, n_2} , яка дорівнює відношенню двох незалежних випадкових величин $\frac{\chi^2_{n_1}}{\chi^2_{n_2}}$ та $\frac{\chi^2_{n_2}}{\chi^2_{n_1}}$, тобто

$$F_{n_1, n_2} = \frac{\chi^2_{n_1} / n_1}{\chi^2_{n_2} / n_2}. \quad (4)$$

Нагадаємо, що χ^2_n – випадкова величина, що має “ χ^2 ”-квадрат розподіл з n ступенями вільності. Розподіл Фішера залежить лише від чисел ступенів вільності і не залежить від інших параметрів. Вираз для щільності $f(x)$ (рис. 2) розподілу Фішера можна знайти у спеціальній літературі. Далі сам розподіл Фішера, як і випадкову величину, що має цей розподіл, позначатимемо також F_{n_1, n_2} . Квантилі розподілу F_{n_1, n_2} табулюовані і наведені в таблиці 6 Додатку 2.

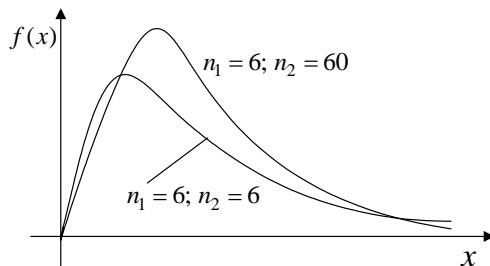


Рис. 2

Для квантилів розподілу Фішера порядків p і $1-p$ справедлива формула

$$F_{1-p, n_1, n_2} = \frac{1}{F_p, n_2, n_1}. \quad (5)$$

Пропонується перевірити **самостійно** справедливість співвідношення

$$F(1, k) = T^2(k),$$

де $T(k)$ – статистика Стьюдента (див. лекцію 20).

Лекція № 24

Перевірка гіпотези про рівність дисперсій двох нормальних генеральних сукупностей. Гіпотези про закони розподілу.

Критерій згоди χ^2

Нехай випадкові величини X та Y розподілені за законами відповідно $N(m_x, \sigma_x)$ та $N(m_y, \sigma_y)$. Параметри цих розподілів нам невідомі. З генеральної сукупності X отримується вибірка обсягу n_1 , а з генеральної сукупності Y – вибірка обсягу n_2 . Вибірки незалежні. За цими вибірками необхідно перевірити гіпотезу про рівність дисперсій генеральних сукупностей X та Y , тобто гіпотезу $H_0 : \sigma_x^2 = \sigma_y^2$.

Незалежно від того, справедлива гіпотеза H_0 чи ні, виправлені вибіркові дисперсії $S_x^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$ та $S_y^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$ є змістовними і незміщеними оцінками генеральних дисперсій σ_x^2 та σ_y^2 (див. лекцію 19). Тому критерій для перевірки гіпотези $H_0 : \sigma_x^2 = \sigma_y^2$ можна будувати, порівнюючи відношення S_x^2 / S_y^2 з 1. Якщо це відношення “значно” відхиляється від 1, то гіпотезу H_0 природно відхилити. У противному разі гіпотезу H_0 природно не відхилити.

Знайдемо розподіл випадкової величини $Z = S_x^2 / S_y^2$, скориставшись теоремою з лекції 20. Згідно з цією теоремою випадкові величини $\frac{n_1 - 1}{\sigma_x^2} S_x^2$ та $\frac{n_2 - 1}{\sigma_y^2} S_y^2$ мають “хi”-квадрат розподілі

відповідно $\chi^2(n_1 - 1)$ та $\chi^2(n_2 - 1)$, тобто $\frac{n_1 - 1}{\sigma_x^2} S_x^2 = \chi^2(n_1 - 1)$, а

$$\frac{n_2 - 1}{\sigma_y} S_y^2 = \chi^2 \quad n_2 - 1 . \text{ Звідси } Z = S_x^2 / S_y^2 = \frac{\chi^2}{\chi^2} \frac{n_1 - 1}{n_2 - 1} / \frac{n_1 - 1}{n_2 - 1} \cdot \frac{\sigma_x}{\sigma_y} . \text{ Якщо}$$

гіпотеза $H_0 : \sigma_x^2 = \sigma_y^2$ – справедлива, то згідно з означенням розподілу Фішера

$$Z = S_x^2 / S_y^2 = F \quad n_1 - 1, n_2 - 1 . \quad (1)$$

Отже, за умови справедливості гіпотези $H_0 : \sigma_x^2 = \sigma_y^2$ відношення виправлених вибіркових дисперсій має розподіл Фішера з $n_1 - 1$ та $n_2 - 1$ ступенями вільності.

Вибираємо статистику (1) за статистику критерію перевірки нульової гіпотези.

Критична область для заданого рівня значущості α будується в залежності від вигляду альтернативної гіпотези H_1 . Розглянемо два випадки.

Випадок 1. Альтернативна гіпотеза має вигляд $H_1 : \sigma_x^2 > \sigma_y^2$. Як ми вже знаємо, найкраща критична область у цьому випадку визначається нерівністю $Z > z_{1-\alpha}$ (див. лекцію 22) або

$$\frac{S_x^2}{S_y^2} = F \quad n_1 - 1, n_2 - 1 > F_{1-\alpha} \quad n_1 - 1, n_2 - 1 . \quad (2)$$

Позначимо через s_x^2 та s_y^2 виправлені вибіркові дисперсії, обчислені з результатами конкретних вибірок.

Тоді, якщо $s_x^2 / s_y^2 > F_{1-\alpha} \quad n_1 - 1, n_2 - 1$, гіпотеза $H_0 : \sigma_x^2 = \sigma_y^2$ відхиляється; у протилежному разі відхиляти гіпотезу немає підстав.

Випадок 2. Альтернативна гіпотеза має вигляд $H_1 : \sigma_x^2 \neq \sigma_y^2$. Для зручності скористаємося нерівністю $z_{\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}$ для області прийняття гіпотези H_0 , яку з врахуванням властивості (5) лекції 23 подамо у вигляді

$$\frac{1}{F_{1-\frac{\alpha}{2}} \quad n_2 - 1, n_1 - 1} \leq \frac{S_x^2}{S_y^2} \leq F_{\frac{\alpha}{2}} \quad n_1 - 1, n_2 - 1 . \quad (3)$$

Отже, якщо вибіркове відношення s_x^2 / s_y^2 задовольняє нерівностям (3), гіпотеза $H_0 : \sigma_x^2 = \sigma_y^2$ приймається; у протилежному разі – відхиляється.

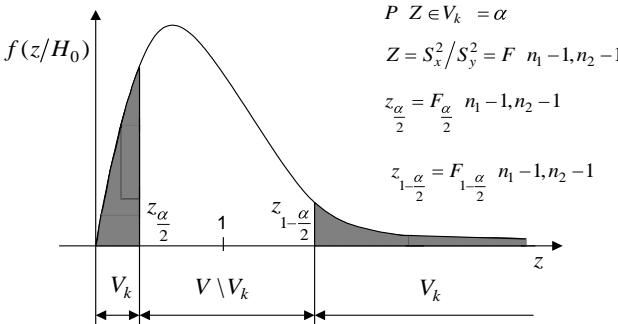


Рис. 1

Положення критичної області у випадку альтернативної гіпотези $H_1: \sigma_x^2 \neq \sigma_y^2$ показано на рис. 1.

Приклад 1. Деяка фізична величина вимірюється двома приладами: А та В. Нижче наведені вибірки значень цієї величини, які одержані даними приладами.

Прилад А: 1,32; 1,35; 1,32; 1,35; 1,30; 1,30; 1,37; 1,31; 1,39; 1,39.

Прилад В: 1,35; 1,31; 1,31; 1,41; 1,39; 1,37; 1,32; 1,34.

З'ясувати, чи можна на рівні значущості $\alpha = 0,02$ вважати, що точність вимірювання приладами А та В однаакова?

Розв'язання. У термінах перевірки статистичних гіпотез цю задачу можна сформулювати так. Маємо дві реалізації незалежних вибірок з нормальних розподілів $N m_x, \sigma_x$, $N m_y, \sigma_y$. Нехай для визначеності x_1, x_2, \dots, x_{10} – вибірка, одержана приладом А; y_1, y_2, \dots, y_8 – приладом В. Висувається гіпотеза про однакову точність вимірювань. Це означає, що необхідно перевірити гіпотезу $H_0: \sigma_x^2 = \sigma_y^2$. За альтернативну природно розглянути гіпотезу $H_1: \sigma_x^2 \neq \sigma_y^2$. Проведемо необхідні обчислення. У нас $n_1 = 10$, $n_2 = 8$,

$$\bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i = 1,34, \quad \bar{y} = \frac{1}{n_2} \sum_{i=1}^{n_2} y_i = 1,35, \quad s_x^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2 = 12,1 \cdot 10^{-4},$$

$$s_y^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2 = 14,0 \cdot 10^{-4}, \quad \frac{s_x^2}{s_y^2} = \frac{12,1 \cdot 10^{-4}}{14,0 \cdot 10^{-4}} = 0,86,$$

$$F_{\frac{1-\alpha}{2}}(n_1-1, n_2-1) = F_{0,99}(9, 7) = 6,72,$$

$$\frac{1}{F_{\frac{n_2-\alpha}{2}, n_1-1}} = \frac{1}{F_{0,99} \quad 7, 9} = \frac{1}{5,61} = 0,18 . \quad \text{Оскільки значення}$$

$\frac{s_x^2}{s_y^2} = 0,86$ належить проміжку $0,18; 6,72$, то на 2%-рівні значущості

гіпотеза $H_0: \sigma_x^2 = \sigma_y^2$ приймається. Для приладів А та В це означає, що гіпотеза про однакову точність вимірювань цими приладами не суперечить експериментальним даним.

У попередній та даній лекціях ми розглянули ряд прикладів перевірки гіпотез про параметри нормальних розподілів. Зрозуміло, що цими прикладами не вичерпуються всі можливі гіпотези такого роду. У таблицях 1, 2 Додатку 3 наведені розглянуті нами та інші можливі гіпотези про дисперсії та математичні сподівання нормальних генеральних сукупностей, описані статистики критеріїв, їх розподіли, а також області прийняття нульової гіпотези для різних альтернативних гіпотез.

Перевірка гіпотез про закон розподілу випадкової величини. Критерій χ^2

Нехай X – випадкова величина з невідомим законом розподілу і виникає необхідність визначення цього закону на основі дослідних даних. Зразу ж відмітимо, що розв'язати цю задачу в такій загальній постановці досить складно, а в більшості випадків робити цього, навіть, і не потрібно. Досить часто на основі аналізу вибірки і деяких додаткових міркувань можна висунути щодо вигляду закону розподілу певну гіпотезу. Найбільш простою і в той же час найбільш сильною гіпотезою подібного роду є гіпотеза, що функція розподілу випадкової величини має повністю визначений вигляд $F(x)$. Виникає питання, чи не спростовують дослідні дані цю гіпотезу?

Отже, якщо закон розподілу випадкової величини X невідомий, але є підстави припустити, що він має певний вигляд $F(x)$, то перевіряється нульова гіпотеза H_0 : генеральна сукупність X має розподіл $F(x)$. Будемо писати $H_0: F_x(x) = F(x)$.

Незалежно від того, справедлива гіпотеза H_0 чи ні, за вибіркою можна завжди знайти емпіричну функцію розподілу $F^*(x)$. Гіпотезу H_0 приймають, якщо “відхилення” емпіричного розподілу $F^*(x)$ від

гіпотетичного F_x мале. У протилежному разі гіпотезу H_0 відхиляють. Міру такого “відхилення” можна визначити багатьма способами, у відповідності до яких отримуються різні критерії для перевірки нульової гіпотези. Ці критерії мають назву **критерій згоди**. Найширше застосування має критерій згоди χ^2 Пірсона, який ми розглянемо детально.

Розіб'ємо область значень випадкової величини X на скінченне число l множин $\Delta_1, \Delta_2, \dots, \Delta_l$, які не перетинаються. Якщо X – неперервна випадкова величина, то це l інтервалів (рис. 2),

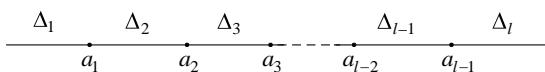


Рис. 2

причому для визначеності вважаємо, що правий кінець виключається з відповідної множини, а лівий – включається. Якщо X – дискретна випадкова величина, то $\Delta_1, \Delta_2, \dots, \Delta_l$ – це групи, що містять окремі значення цієї випадкової величини.

Далі припускаємо, що гіпотеза $H_0: F_x(x) = F(x)$ справедлива. За функцією розподілу $F(x)$ знаходимо ймовірності попадання випадкової величини у множини розбиття, тобто ймовірності $p_k = P(X \in \Delta_k), k = 1, 2, \dots, l$. Очевидно, що $\sum_{k=1}^l p_k = 1$. У позначеннях рис. 2, наприклад,

$$p_k = F(a_k) - F(a_{k-1}), \quad k = 1, 2, \dots, l, \quad a_0 = -\infty, \quad a_l = \infty. \quad (4)$$

Множини Δ_k вибираються таким чином, що всі $p_k > 0$.

Позначимо через v_k – число елементів вибірки X_1, X_2, \dots, X_n , які потрапляють до множини $\Delta_k, k = 1, 2, \dots, l$. Тоді $\frac{v_k}{n}$ – відносна частота попадання випадкової величини X до множини Δ_k при n спостереженнях над нею. Очевидно, що $\sum_{k=1}^l v_k = n, \quad \sum_{k=1}^l \frac{v_k}{n} = 1$.

Зauważення 1. Поки що вибірка розглядається нами як випадковий вектор. Тому v_k – випадкова величина, що має біномний розподіл з математичним сподіванням np_k та дисперсією $np_k q_k$, $q_k = 1 - p_k$. Дійсно, кожна компонента цього випадкового вектора є однією і тією ж

випадковою величиною X і вибірку можна інтерпретувати як n -кратне повторення одного і того ж експерименту з ймовірністю “успіху” p_k в кожному експерименті.

Повернемось до величин p_k та $\frac{v_k}{n}$, що є відповідно

ймовірностями та відносними частотами події $X \in \Delta_k$. Згідно з теоремою Бернуллі при справедливій гіпотезі H_0 і достатньо великому n ці величини повинні відрізнятися мало. Якщо це не так, то гіпотезу природно відхилити.

Гіпотетичний розподіл характеризується величинами p_k , а емпіричний (розподіл вибірки) – величинами $\frac{v_k}{n}$. Тому міру відхилення цих розподілів будемо шукати як міру відхилення величин p_k та $\frac{v_k}{n}$. Згідно з методом найменших квадратів за таку міру

візьмемо вираз $\sum_{k=1}^l c_k \left(\frac{v_k}{n} - p_k \right)^2$, де коефіцієнти c_k можуть бути більш-менш довільними. Проте, як показав Пірсон, міра відхилення з найпростішими властивостями отримується тоді, коли $c_k = \frac{n}{p_k}$. Отже, будемо користуватися величиною

$$\chi_n^2 = \sum_{k=1}^l \frac{(v_k - np_k)^2}{np_k}, \quad (5)$$

що досить просто виражається через величини v_k та np_k . Останні іноді називають відповідно частотами, що спостерігаються, та частотами, на які сподіваються.

Якщо розподіл величини χ_n^2 відомий, то для заданого рівня значущості α можна побудувати критичну область і область прийняття гіпотези H_0 . Проте на практиці розподіл величини χ_n^2 для кожного значення n , як правило, не знаходять, а користуються граничним (при $n \rightarrow \infty$) розподілом цієї величини.

Зауваження 2. Розподіл випадкової величини $U_k = \frac{v_k - np_k}{\sqrt{np_k q_k}}$ при

$n \rightarrow \infty$ наближається до нормальногорозподілу $N(0, 1)$ (див зауваження 1

і лекцію 17). Тому для достатньо великих n близьким до нормального буде розподіл випадкової величини $U_k \sqrt{q_k} = \frac{v_k - np_k}{\sqrt{np_k}}$. Випадкові величини U_k пов'язані лінійною залежністю, дійсно

$$\sum_{k=1}^l U_k \sqrt{np_k q_k} = \sum_{k=1}^l v_k - np_k = \sum_{k=1}^l v_k - n \sum_{k=1}^l p_k = n - n = 0.$$

Пірсон довів таку теорему.

Теорема (Пірсона). При $n \rightarrow \infty$ розподіл випадкової величини χ_n^2 з (5) збігається до розподілу χ^2 з $l-1$ ступенями вільності.

Використовуючи цю теорему, введемо критерій згоди χ^2 для перевірки гіпотези $H_0: F_x(x) = F(x)$.

Нехай (x_1, \dots, x_n) – конкретна вибірка, отримана в результаті спостереження над випадковою величиною X . Тоді, провівши описану вище процедуру розбиття множини значень X і знайшовши величини p_k і v_k , обчислюємо вибіркове значення статистики критерію χ^2 за формулою (5):

$$\chi_{\alpha}^2 = \sum_{k=1}^l \frac{v_k - np_k}{np_k}^2. \quad (6)$$

Нехай $\chi_{1-\alpha}^2(l-1)$ – квантиль розподілу χ^2 порядку $1-\alpha$. Тоді гіпотеза H_0 на рівні значущості α відхиляється, якщо $\chi_{\alpha}^2 \geq \chi_{1-\alpha}^2(l-1)$. Якщо $\chi_{\alpha}^2 < \chi_{1-\alpha}^2(l-1)$, то гіпотеза вважається такою, що не суперечить дослідним даним (приймається).

На практиці гіпотетичний розподіл має чітко визначений вигляд $F(x)$ не так часто. Більш вживаним є випадок, коли гіпотетичний розподіл містить один чи декілька невідомих параметрів. Тоді гіпотеза H_0 формулюється так: функція розподілу випадкової величини X дорівнює $F(x; \alpha_1, \alpha_2, \dots, \alpha_s)$ при деяких значеннях параметрів $\alpha_1, \alpha_2, \dots, \alpha_s$.

За значення невідомих параметрів природно взяти їх точкові оцінки $\tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_s$, побудовані за вибіркою. Тоді гіпотетичний розподіл конкретизується: $F(x; \tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_s) = F(x)$ і можна згідно з описаним вище побудувати величину χ_n^2 з (5).

Фішер показав при досить загальних умовах, що й у цьому випадку статистика χ_n^2 при $n \rightarrow \infty$ має χ^2 -розподіл, але з $l-s-1$ ступенями вільності. Тому вибіркове значення статистики χ_s^2 з (6) повинно порівнюватись з квантиллю $\chi_{1-\alpha}^2(l-s-1)$.

Отже, щоб скористатись критерієм згоди χ^2 у випадку, коли гіпотетичний розподіл містить s невідомих параметрів, потрібно замінити ці параметри їх точковими оцінками і зменшити число ступенів вільності χ^2 -розподілу на s .

Зauważення 3. Необхідно умовою застосування критерію χ^2 є умова $np_i \geq 5$. Це пов'язано з тим, що використовується граничний розподіл випадкової величини (5) при $n \rightarrow \infty$. Якщо умова $np_i \geq 5$ для деяких інтервалів Δ_k не виконується, то їх потрібно об'єднати з сусідніми інтервалами.

Лекція № 25

Приклади на застосування критерію χ^2

Розглянемо випадкову величину $X = \{0; 1\}$ – індикатор випадкової події A (X приймає значення 1, якщо подія A при проведенні експерименту відбувається, та 0, якщо – не відбувається). Сформулюємо гіпотезу H_0 : випадкова величина X має розподіл $P\{X=1\}=p$, $P\{X=0\}=1-p=q$, де p – фіксоване число, $0 < p < 1$. Іншими словами, сформульовано гіпотезу H_0 : ймовірність події A дорівнює p . Нехай для перевірки гіпотези H_0 проведено n експериментів, в результаті яких подія A настала m разів. По відношенню до випадкової величини X результати експериментів подамо у вигляді:

0	1
$v_1 = n - m$	$v_2 = m$

.

Область можливих значень X розбита на $l=2$ множини: $\Delta_1 = \{0\}$, $\Delta_2 = \{1\}$. За умови, що гіпотеза H_0 справедлива, $p_1 = P\{X \in \Delta_1\} = 1 - p = q$, $p_2 = P\{X \in \Delta_2\} = p$. Отже, вибіркове значення статистики критерію χ^2 приймає вигляд

$$\chi_e^2 = \frac{n - m - nq}{nq} + \frac{m - np}{np} = \frac{m - np}{npq}^2. \quad (1)$$

Це значення порівнюється з квантіллю $\chi_{1-\alpha}^2 l-1 = \chi_{1-\alpha}^2 1$ і приймається відповідне рішення.

Зauważення 1. Перевіряючи з використанням критерію χ^2 гіпотезу H_0 про рівність ймовірності події A деякому числу p , тобто гіпотезу $H_0: P A = p$, альтернативною гіпотезою слід вважати гіпотезу $H_1: P A \neq p$.

Розглянемо декілька цікавих прикладів на використання критерію χ^2 , які не потребують громіздких обчислень.

Приклад 1. При 50 підкиданнях монети “герб” випав 20 разів. Чи можна з надійністю 0,9 стверджувати, що монета симетрична?

Розв’язання. Перевіримо для рівня значущості $\alpha = 0,1$ гіпотезу H_0 : ймовірність випадання “герба” дорівнює $p = \frac{1}{2}$. Підставляючи в (1) $n = 50$, $m = 20$, $p = q = \frac{1}{2}$, одержуємо

$$\chi_e^2 = \left(20 - 50 \cdot \frac{1}{2} \right)^2 / 50 \cdot \frac{1}{2} \cdot \frac{1}{2} = 2. \text{ Оскільки } \chi_{1-\alpha}^2 1 = \chi_{0,9}^2 1 = 2,71 > \chi_e^2,$$

гіпотеза H_0 приймається, тобто дослідні дані не суперечать припущенняю, що монета симетрична.

Приклад 2. При 120 підкиданнях грального кубика “шістка” з’явилась 40 разів. Чи узгоджується цей результат з припущенням на рівні $\alpha = 0,05$, що кубик “правильний”?

Розв’язання. Перевіримо з використанням критерію χ^2 гіпотезу H_0 : ймовірність випадання “шістки” дорівнює $p = \frac{1}{6}$. Підставляючи в

$$(1) \quad n = 120, \quad m = 40, \quad p = \frac{1}{6}, \quad q = \frac{5}{6}, \quad \text{одержуємо } \chi_e^2 = \frac{(40 - 120 \cdot 1/6)^2}{120 \cdot 1/6 \cdot 5/6} = 24.$$

За таблицею квантілів $\chi_{1-\alpha}^2(1) = \chi_{0,95}^2(1) = 3,84$. Отже, $\chi_e^2 > \chi_{0,95}^2 1$. Тому гіпотеза H_0 відхиляється. Оскільки дослідні дані спростовують

гіпотезу, що ймовірність випадання “шістки” дорівнює $\frac{1}{6}$, то тим самим вони спростовують і припущення, що кубик “правильний”.

Перейдемо до застосування критерію χ^2 для перевірки гіпотез про закони розподілу.

Приклад 3. Спостерігались покази годинникової стрілки 500 навмання вибраних годинників після їх зупинки. Результати спостережень зведені до таблиці

$[i, i+1)$	$[0,1)$	$[1,2)$	$[2,3)$	$[3,4)$	$[4,5)$	$[5,6)$	$[6,7)$
n_i	41	34	54	39	49	45	41
$[i, i+1)$	$[7,8)$	$[8,9)$	$[9,10)$	$[10,11)$	$[11,12)$	Всього	
n_i	33	37	41	47	39	500	

Тут $[i, i+1)$ – проміжок часу від i -ї години до $i+1$ -ї години, $i = 0, 1, 2, \dots, 11$; n_i – кількість годинників, покази годинникової стрілки яких на момент зупинки належать проміжку $[i, i+1)$.

Позначимо через X – положення годинникової стрілки на момент зупинки годинника. Перевірити на рівні значущості $\alpha = 0,05$ гіпотезу H_0 : випадкова величина X має рівномірний розподіл на $[0, 12]$.

Розв’язання. Область значень випадкової величини X розбита на $l=12$ множин $[i, i+1)$, $i = 0, 1, \dots, 11$. За умови справедливості гіпотези H_0 величина X має щільність розподілу

$$f(x) = \begin{cases} \frac{1}{12} & \text{при } x \in [0, 12] \\ 0 & \text{при } x \notin [0, 12]. \end{cases}$$

Тоді $p_k = P(X \in [k, k+1]) = \int_k^{k+1} \frac{1}{12} dx = \frac{1}{12}$ і $np_k = 500 \cdot \frac{1}{12} = 41,7 \geq 5$, $k = 0, 1, \dots, 11$. Отже, можна скористатись критерієм χ^2 . Обчислимо вибіркове значення χ^2_ϵ за формулою (6) лекції 24:

$$\chi^2_\epsilon = \frac{1}{41,7} \left[(41 - 41,7)^2 + (34 - 41,7)^2 + \dots + (39 - 41,7)^2 \right] = 9,99.$$

За таблицею квантілів χ^2 -розподілу (Додаток 2, табл. 4) знаходимо $\chi^2_{1-\alpha}(l-1) = \chi^2_{0,95}(11) = 19,7$. Отже, $\chi^2_\epsilon < \chi^2_{0,95}(11)$. Тому гіпотеза H_0 не відхиляється.

Приклад 4 (див. “Сборник задач по математике для втузов”. Ч.3. Под ред. А.В. Ефимова. – М.: Наука, 1990). У перших двох стовпцях табл. 1 наведені дані про відмову апаратури за 10000 годин роботи.

Табл. 1.

Число відмов, k	Кількість випадків, в яких спостерігалось k відмов, v_k	$p_k = \frac{0,6^k}{k!} e^{-0,6}$	Очікуване число випадків з k відмовами, np_k
0	427	0,54881	416
1	235	0,32929	249
2	72	0,09879	75
3	21	0,01976	15
4	1	0,00296	2
5	1	0,00036	0
≥ 6	0	0,00004	0
Сума	757	–	–

Загальне число випробуваних екземплярів апаратури $n = 757$, при цьому спостерігалась $0 \cdot 427 + 1 \cdot 235 + 2 \cdot 72 + 3 \cdot 21 + 4 \cdot 1 + 5 \cdot 1 = 451$ відмова. Перевірити на рівні значущості $\alpha = 0,01$ гіпотезу H_0 : число відмов має розподіл Пуассона, тобто

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

Розв’язання. Оцінка параметра λ дорівнює середньому числу відмов: $\tilde{\lambda} = 451/757 \approx 0,6$. За формулою (*) знаходимо ймовірності p_k і величини np_k (третій і четвертий стовпці табл. 1). Для $k = 4, 5$ і 6 значення $np_k < 5$. Тому об’єднуємо ці рядки з рядком для $k = 3$. В результаті одержуємо значення, наведені в табл. 2.

Табл. 2

k	v_k	np_k	$\frac{v_k - np_k}{np_k}^2$
0	427	416	0,291
1	235	249	0,787
2	72	75	0,120
≥ 3	23	17	2,118
–	–	–	$\chi^2_e = 3,316$

Оскільки за вибіркою оцінюється $s = 1$ параметр, то число ступенів вільності $l - s - 1 = 4 - 1 - 1 = 2$. За таблицею квантілів

знаходимо $\chi^2_{0,99} \cdot 2 = 9,21$. Отже, $\chi^2 < \chi^2_{0,99} \cdot 2$ і гіпотеза про розподіл числа відмов за законом Пуассона приймається.

Приклад 5 (взято звідти ж, що й приклад 4). В результаті спостереження над випадковою величиною X одержано вибірку обсягом $n=55$ і знайдено оцінки математичного сподівання і дисперсії

$$\tilde{m} = \bar{x} = \frac{1}{n} \sum_{i=1}^{55} x_i \approx 17,84; \quad \tilde{\sigma}^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^{55} (x_i - \bar{x})^2 \approx 8,53.$$

Вибірку групували; результати групування наведені в другому і третьому стовпцях табл. 3.

Перевірити на рівні значущості $\alpha=0,1$ гіпотезу H_0 : випадкова величина X має нормальній розподіл.

Розв'язання. Знаходимо ймовірності p_k (четвертий стовпчик табл. 3) попадання випадкової величини в інтервалі Δ_k за формулою

$$p_k = P\{X \in \Delta_k\} = \Phi\left(\frac{b_k - \bar{x}}{s}\right) - \Phi\left(\frac{a_k - \bar{x}}{s}\right), \quad k = 1, 2, \dots, 7,$$

де a_k і b_k – відповідно нижня та верхня межа інтервалу. $\Phi(x)$ – функція розподілу $N(0, 1)$, значення якої наведені в табл. 1 Додатку 2. У п'ятому стовпці наведені частоти np_k , на які сподіваються, а в шостому – значення v_k після об'єднання перших двох і останніх двох інтервалів.

Табл.3

Номер інтервалу k	Межі інтервалу Δx	Кількість елементів вибірки в інтервалі v_k	Ймовірність попадання в інтервалі p_k	Очікувана частота np_k	$v_k - np_k$	$\frac{v_k - np_k}{np_k}$ 2	
1	$-\infty - 12$	2	0,0228	1,254	5,274	0,725	0,010
2	12–14	4	0,0731	4,020			
3	14–16	8	0,1686	9,273	9,273	-1,273	0,175
4	16–18	12	0,2576	14,168	14,168	-2,168	0,332
5	18–20	16	0,2484	13,662	13,662	2,338	0,400
6	20–22	10	0,1519	8,354	12,633	0,366	0,011
7	$22 - +\infty$	3	0,0778	4,279			
	Сума	55	1,0001	55	55	–	0,928

Оскільки після об'єднання залишилось $l = 5$ інтервалів, а за вибіркою оцінено $s = 2$ параметри, то число ступенів вільності $l - s - 1 = 5 - 2 - 1 = 2$. За табл. 4 Додатку 2 знаходимо $\chi^2_{0,9} | 2 = 4,61$. Вибіркове значення статистики критерію $\chi^2_{\text{e}} = 0,928$. Отже, $\chi^2_{\text{e}} < \chi^2_{0,9} | 2$ і тому гіпотеза про нормальний розподіл приймається.

Критерій χ^2 як критерій незалежності випадкових величин

Нехай в результаті експерименту спостерігаються дві дискретні випадкові величини X та Y з можливими значеннями x_1, x_2, \dots, x_k та y_1, y_2, \dots, y_l . Виникає питання: чи залежні ці випадкові величини? Як ми вже знаємо (див. лекцію 12), на пару X, Y можна дивитись як на двовимірну випадкову величину (випадковий вектор) з розподілом у вигляді табл. 4, де $p_{ij} = P\{X = x_i, Y = y_j\}$.

Табл. 4

$x_i \backslash y_i$	x_1	x_2	\dots	x_k
y_1	p_{11}	p_{21}	\dots	p_{k1}
y_2	p_{12}	p_{22}	\dots	p_{k2}
\vdots	\vdots	\vdots	\dots	\vdots
y_l	p_{1l}	p_{2l}	\dots	p_{kl}

Розглянемо випадок, коли ані розподіл вектора (величини p_{ij}), ані розподіли

x_i	x_1	x_2	\dots	x_k	i	y_j	y_1	y_2	\dots	y_l
p_{x_i}	p_{x_1}	p_{x_2}	\dots	p_{x_k}		p_{y_j}	p_{y_1}	p_{y_2}	\dots	p_{y_l}

його компонент (величини p_{x_i} та p_{y_j}) нам невідомі. Нагадаємо, що

$$p_{x_i} = P\{X = x_i\} = \sum_{j=1}^l p_{ij}, \quad i = 1, 2, \dots, k; \quad (2)$$

$$p_{y_j} = P\{Y = y_j\} = \sum_{i=1}^k p_{ij}, \quad j = 1, 2, \dots, l.$$

Стосовно сумісного розподілу вектора X, Y висувається гіпотеза про незалежність його компонент, тобто гіпотеза $H_0: p_{ij} = p_{x_i} \cdot p_{y_j}$ (див. лекцію 12).

Для перевірки гіпотези H_0 проводяться n експериментів, результати яких можна подати у вигляді табл. 5.

Табл. 5

$y_i \backslash x_i$	x_1	x_2	...	x_k
y_1	v_{11}	v_{21}	...	v_{k1}
y_2	v_{12}	v_{22}	...	v_{k2}
\vdots	\vdots	\vdots	...	\vdots
y_l	v_{1l}	v_{2l}	...	v_{kl}

Тут v_{ij} – число експериментів, в яких спостерігалась пара чисел x_i, y_j . Зрозуміло, що $\sum_{i=1}^k \sum_{j=1}^l v_{ij} = n$. По аналогії з (2) введемо

означення

$$v_{x_i} = \sum_{j=1}^l v_{ij}, \quad i = 1, 2, \dots, k; \quad v_{y_j} = \sum_{i=1}^k v_{ij}, \quad j = 1, 2, \dots, l.$$

Очевидно, що v_{x_i} – число експериментів, в яких спостерігалась подія $\{X = x_i\}$, а v_{y_j} – подія $\{Y = y_j\}$.

Критерій χ^2 для перевірки гіпотези $H_0: p_{ij} = p_{x_i} \cdot p_{y_j}$ буде використовуватися за тією ж схемою, що й раніше. Множинами розбиття області значень випадкового вектора X, Y є множини $\Delta_{ij} = x_i, y_j, \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, l$. Очевидно, що таких множин $k \cdot l$.

Оскільки гіпотетичний розподіл $p_{ij} = p_{x_i} \cdot p_{y_j}$ містить невідомі параметри p_{x_i} та p_{y_j} , потрібно побудувати точкові оцінки цих параметрів. Такими оцінками є відносні частоти $\tilde{p}_{x_i} = \frac{v_{x_i}}{n}$ та $\tilde{p}_{y_j} = \frac{v_{y_j}}{n}$.

Зauważення 2. Оскільки $\sum_{i=1}^k p_{x_i} = 1$ і $\sum_{j=1}^l p_{y_j} = 1$, то очевидно, що число незалежних параметрів p_{x_i}, p_{y_j} , які оцінюються за двовимірною вибіркою, дорівнює $k - 1 + l - 1 = k + l - 2$.

Після заміни невідомих параметрів їх точковими оцінками, гіпотетичний розподіл конкретизується і за умови, що він має місце

$$P(X, Y \in \Delta_{ij}) = p_{ij} = \tilde{p}_{x_i} \cdot \tilde{p}_{y_j} = \frac{v_{x_i}}{n} \cdot \frac{v_{y_j}}{n}. \quad (3)$$

За міру відхилення емпіричного розподілу $\frac{v_{ij}}{n}$, $i = 1, 2, \dots, k$, $j = 1, 2, \dots, l$ від гіпотетичного (3) вибирається статистика

$$\chi_n^2 = \sum_{i=1}^k \sum_{j=1}^l \frac{\left| v_{ij} - n \tilde{p}_{x_i} \tilde{p}_{y_j} \right|^2}{n \tilde{p}_{x_i} \tilde{p}_{y_j}} = \sum_{i=1}^k \sum_{j=1}^l \frac{\left| v_{ij} - v_{x_i} v_{y_j} / n \right|^2}{v_{x_i} v_{y_j} / n}. \quad (4)$$

Фішер показав, що при $n \rightarrow \infty$ статистика (4) має χ^2 -розподіл з $kl - k + l - 2 - 1 = k - 1 \cdot l - 1$ ступенями вільності (див. зауваження 2).

Отже, вибіркове значення χ_n^2 статистики (4) порівнюється з квантиллю $\chi_{1-\alpha}^2$ $k-1$ $l-1$ і приймається відповідне рішення про прийняття чи відхилення гіпотези H_0 .

Зауваження 3. Оскільки, знову ж таки, використовується граничний розподіл статистики (4) при $n \rightarrow \infty$, вимагається виконання умови

$$n \tilde{p}_{x_i} \tilde{p}_{y_j} = \frac{v_{x_i} v_{y_j}}{n} \geq 5. \text{ Якщо ця умова для деяких клітинок табл. 5 не}$$

виконується, то відповідні рядки і стовпці потрібно об'єднати з сусіднimi рядками i стовпцями.

Приклад 6. (Взято з підручника “Вища математика”. Книга 2. За ред. Г.Л. Кулініча. – К. Либідь, 1996). Нижче наведені дані про колір волосся на голові (світле, темне) і колір очей (голубі, карі) 147 випадково вибраних юнаків.

Очі Волосся \	Голубі	Карі	Разом
Темне	31	41	72
Світле	40	35	75
Разом	71	76	147

Чи можна на підставі цих даних зробити висновок, що колір очей пов’язаний з кольором волосся на голові?

У термінах перевірки статистичних гіпотез ця задача формулюється як задача перевірки гіпотези про незалежність випадкових величин (кольору очей і кольору волосся). В умові задачі $k = l = 2$, $n = 147$. Значення v_{ij} , v_{x_i} , v_{y_j} подані в таблиці. Легко перевірити, що $\frac{v_{x_i} v_{y_j}}{n} > 5$, $i, j = 1, 2$. Знаходимо значення статистики (4) :

$$\chi_n^2 = \frac{\left(31 - \frac{72 \cdot 71}{147} \right)^2}{\frac{72 \cdot 71}{147}} + \frac{\left(41 - \frac{72 \cdot 76}{147} \right)^2}{\frac{72 \cdot 76}{147}} + \frac{\left(40 - \frac{75 \cdot 71}{147} \right)^2}{\frac{75 \cdot 71}{147}} + \frac{\left(35 - \frac{75 \cdot 76}{147} \right)^2}{\frac{75 \cdot 76}{147}} \approx 1,51.$$

Перевіримо гіпотезу на рівні значущості $\alpha = 0,05$. Знаходимо квантіль $\chi^2_{1-\alpha} \quad k-1 \quad l-1 = \chi^2_{0,95} \quad 1 = 3,84$. Оскільки $\chi^2_e < \chi^2_{0,95} \quad 1$, гіпотеза приймається, тобто колір очей і колір волосся можна вважати незалежними випадковими величинами.

Зauważення 4. Перевірку гіпотези про незалежність можна проводити у випадку неперервних випадкових величин X та Y . Для цього область значень кожної з них розбивається на скіченне число інтервалів (наприклад, k інтервалів для X і l інтервалів для Y). Тоді v_{ij} – число експериментів, в яких X потрапляє до i -го інтервалу, а Y – до j -го інтервалу. Очевидний смисл мають і величини v_{x_i} та v_{y_j} . За значеннями v_{ij}, v_{x_i}, v_{y_j} обчислюється статистика (4), знаходиться квантіль $\chi^2_{1-\alpha} \quad k-1 \quad l-1$ і приймається відповідне рішення.

Розділ 3. ЕЛЕМЕНТИ КОРЕЛЯЦІЙНОЇ ТЕОРІЇ ВИПАДКОВИХ ФУНКІЙ

Лекція № 26

Означення випадкової функції. Закони розподілу та числові характеристики випадкових функцій

Нехай X – деяка випадкова величина. У результаті окремого експерименту над цією величиною ми спостерігаємо її конкретне, наперед невідоме, але єдине значення. Якщо одночасно розглядаються декілька випадкових величин, результатом окремого експерименту є набір конкретних значень цих величин.

У багатьох випадках результатом спостереження в окремому експерименті є чисрова функція деякого параметра t . Нехай, наприклад, експеримент полягає у спостереженні за коливанням напруги в електричній мережі навколо її номінального значення впродовж деякого інтервалу часу. Тоді результатом кожного такого окремого експерименту є конкретна функція часу t . Очевидно, що під дією випадкових впливів вигляд цієї функції від експерименту до експерименту буде змінюватись. У подібних випадках користуються поняттям **випадкової функції**, яке узагальнює поняття випадкової величини та системи випадкових величин.

Дамо формальне означення випадкової функції. Нехай T – деяка множина дійсних чисел; Ω, F, P – ймовірнісний простір.

Означення 1. Випадковою функцією називається функція невипадкового аргументу t , яка при кожному фіксованому значенні цього аргументу є випадковою величиною, заданою на ймовірнісному просторі Ω, F, P (див. лекцію 7).

Наприклад, якщо X_0 – випадкова величина, то функції $X(t) = t^3 X_0$; $Y(t) = \sin X_0 t$; $Z(t) = X_0 \sin t$ є випадковими.

Множину T дійсних чисел називають областю визначення випадкової функції. У випадку, коли параметр t відіграє роль часу, випадкову функцію ще називають **випадковим процесом**. Випадкові функції і далі будемо позначати великими буквами $X(t)$, $Y(t)$ тощо.

Оскільки випадкова величина сама є функцією на множині елементарних подій Ω випадкового експерименту, то для випадкової функції $X(t)$ можна записати

$$X(t) = X(\omega, t) \quad t \in T, \omega \in \Omega. \quad (1)$$

При фіксованому значенні $\omega = \omega_0$ маємо невипадкову функцію $x(t) = X(\omega_0, t)$ однієї дійсної змінної t , яку називають **реалізацією** або **траекторією** випадкової функції. Спостерігаючи випадкову функцію при деякому експерименті, ми фактично спостерігаємо одну з її реалізацій. Якщо над випадковою функцією $X(t)$ провести цілу серію експериментів, то в результаті будемо мати сімейство її реалізацій $x_1(t), x_2(t), \dots, x_n(t)$, де індекс вказує на номер експерименту (рис. 1). Реалізації випадкової функції і надалі позначатимемо малими буквами.

Якщо в (1) зафіксувати $t = t_0$, то утвориться випадкова величина $X(t_0) = X(\omega, t_0)$, яку називають **перерізом** випадкової функції (рис. 1).

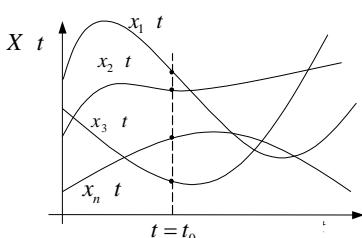


Рис. 1

Отже, випадкову функцію можна розглядати як сукупність усіх її траекторій або як сукупність усіх її перерізів. В останньому випадку випадкову функцію можна розглядати як узагальнення поняття системи випадкових величин, коли ця система є нескінченною (в загальному випадку – незчисленною).

Зрозуміло, що задати випадкову функцію аналітично (у вигляді формул), взагалі кажучи, неможливо. Це можна зробити в частинних випадках, коли вигляд випадкової функції, як у наведених після означення 1 прикладах, відомий, а визначальні параметри є випадковими величинами.

Область визначення випадкової функції T може бути як неперервною (деяким числовим проміжком), так і дискретною множиною дійсних чисел. В останньому випадку випадкову функцію ще називають випадковою послідовністю. Перерізи випадкової функції, як випадкові величини, можуть бути як дискретними, так і неперервними. У зв'язку з цим наведемо ряд прикладів випадкових функцій:

1. $X(t)$ – величина атмосферного тиску в залежності від висоти t над рівнем моря;

$Y(t)$ – діаметр ткацької нитки в залежності від довжини нитки t в процесі її виготовлення;

$Z(t)$ – відхилення швидкості літака від номінальної швидкості в залежності від часу t .

2. $N(t)$ – кількість викликів на телефонній станції за час t ;

$M(t)$ – зміна в часі кількості автомашин, що прибувають до бензоколонки.

3. $X(m)$ – результати вимірювання розміру однотипних деталей;

$Y(m)$ – час безперебійної роботи приладу (тут перехід від деталі до деталі або від попередньої відмови приладу до наступної приймається за зміну дискретного аргументу $t = m$).

4. $N(t_1, t_2, \dots, t_n, \dots)$ – кількість очок, що випадає при підкиданні грального кубика в дискретні моменти часу $t_1, t_2, \dots, t_n, \dots$;

$M(m)$ – результати перевірки виробів на можливість браку.

Наведені приклади відображають чотири можливі типи випадкових функцій в залежності від того, належать можливі значення аргументу t і функції $X(t)$ дискретні множині чисел чи числовому проміжку. Нижче розглядаються випадкові функції неперервного аргументу t .

Перейдемо до описування випадкової функції з ймовірнісної точки зору.

Нехай задано випадкову функцію $X(t)$. Зафіксуємо деяке значення аргументу $t = t_1$ і розглянемо значення $X(t_1)$ випадкової функції. Згідно з означенням $X(t_1)$ є випадковою величиною, а тому характеризується своїм законом розподілу, наприклад функцією розподілу $F_1(x; t_1) = P(X(t_1) < x)$ або щільністю $f_1(x; t_1) = \frac{\partial F_1(x; t_1)}{\partial x}$,

якщо переріз $X(t_1)$ – випадкова величина неперервного типу і функція $F_1(x; t_1)$ – диференційовна. Добуток $f_1(x; t_1)dx$ дає наближене значення ймовірності $P(x < X(t_1) < x + dx)$ того, що графік реалізації $x(t)$ перетинає пряму $t = t_1$ між точками x та $x + dx$ (рис. 2).

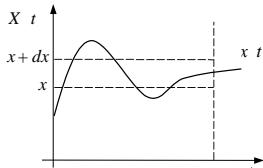


Рис. 2

Функції $F_1(x; t_1)$ та $f_1(x; t_1)$ називаються відповідно одновимірною функцією розподілу та одновимірною щільністю розподілу випадкової функції $X(t)$ при $t = t_1$.

Отже, випадкова функція $X(t)$ при фіксованому значенні аргументу t повністю описується одновимірним законом розподілу її можливих значень у формі функції розподілу $F_1(x; t_1)$ або щільності $f_1(x; t_1)$.

Однак, одновимірним законом розподілу не можна охарактеризувати зв'язок між значеннями випадкової функції для різних значень аргументу. Тому введемо поняття багатовимірного закону розподілу випадкової функції (будемо оперувати тільки щільностями ймовірності).

Зафіксуємо два значення аргументу: t_1 і t_2 . Відповідні перерізи випадкової функції утворюють систему двох випадкових величин $X(t_1), X(t_2)$. Для характеристики ймовірності зв'язку між цими величинами слід скористатися спільною щільністю розподілу $f_2(x_1, x_2; t_1, t_2)$, яку називають двовимірною щільністю ймовірності. Добуток $f_2(x_1, x_2; t_1, t_2)dx_1dx_2$ дає наближене значення ймовірності $P(x_1 < X(t_1) < x_1 + dx_1; x_2 < X(t_2) < x_2 + dx_2)$ того, що графік реалізації буде перетинати пряму $t = t_1$ на висоті між x_1 і $x_1 + dx_1$, а пряму $t = t_2$ – на висоті між x_2 і $x_2 + dx_2$ (рис. 3).

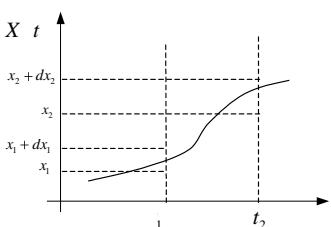


Рис. 3

Знаючи двовимірну щільність, можна за загальним правилом (див. лекцію 12) обчислити одновимірну щільність:

$$f_1(x_1; t_1) = \int_{-\infty}^{\infty} f_2(x_1, x_2; t_1, t_2)dx_2.$$

Зрозуміло, що й двовимірна щільність ймовірності не є повною характеристикою випадкової функції. Можна було б ввести тривимірну щільність розподілу, а також щільність розподілу будь-якого виміру. Разом з тим, відмітимо, що знання двовимірної щільності $f_2(x_1, x_2; t_1, t_2)$

достатньо для так званої **кореляційної теорії випадкових функцій**, елементи якої ми розглядаємо.

Використання багатовимірних щільностей розподілу для описування випадкових функцій спряжено з громіздкими математичними перетвореннями. Тому на практиці користуються ймовірнісними характеристиками випадкових функцій, аналогічними числовим характеристикам випадкових величин (математичними сподіваннями, дисперсіями, кореляційними функціями тощо).

Означення 2. Математичним сподіванням випадкової функції $X(t)$ називається така невипадкова функція $m_x(t)$ аргументу t , яка при кожному фіксованому значенні цього аргументу дорівнює математичному сподіванню відповідного перерізу випадкової функції. З означення математичного сподівання випливає, що

$$m_x(t) = M X(t) = \int_{-\infty}^{\infty} x f_1(x; t) dx. \quad (2)$$

Математичне сподівання $m_x(t)$ є деякою середньою функцією, навколо якої групуються і відносно якої коливаються всі можливі реалізації випадкової функції (рис. 4). Функцію $m_x(t)$ ще називають невипадковою складовою випадкової функції $X(t)$, а різницю

$$\overset{\circ}{X}(t) = X(t) - m_x(t) - \text{суть флюктуаційною частиною цієї функції.}$$

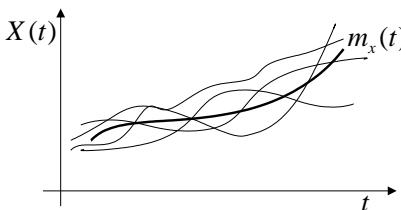


Рис. 4

Означення 3. Дисперсією випадкової функції $X(t)$ називається невипадкова функція $D_x(t)$ аргументу t , яка при кожному фіксованому значенні цього аргументу дорівнює дисперсії відповідного перерізу випадкової функції:

$$D_x(t) = D X(t) = \int_{-\infty}^{+\infty} X(t) - m_x(t)^2 f_1(x; t) dx. \quad (3)$$

Дисперсія випадкової функції характеризує розсіювання її реалізацій навколо математичного сподівання.

Середнє квадратичне відхилення випадкової функції визначається як квадратний корінь з дисперсії, тобто

$$\sigma_x(t) = \sqrt{D_x(t)}. \quad (4)$$

Математичне сподівання та дисперсія характеризують випадкову функцію далеко не повністю. Можна навести приклади, коли

випадкові функції мають однакові математичні сподівання та дисперсії, але значно відрізняються своєю поведінкою. Зокрема, знаючи математичне сподівання та дисперсію випадкової функції, нічого не можна сказати про рівень залежності двох її перерізів. Для оцінки такої залежності користуються характеристикою, яку називають **кореляційною функцією**.

Означення 4. Кореляційною функцією випадкової функції $X(t)$ називається невипадкова функція двох аргументів $K_x(t_1, t_2)$, яка при кожній парі значень t_1 і t_2 дорівнює коваріації відповідних перерізів випадкової функції (див. лекцію 13).

Якщо відома двовимірна щільність розподілу $f_2(x_1, x_2; t_1, t_2)$, то кореляційну функцію можна знайти за формулою

$$\begin{aligned} K_x(t_1, t_2) &= M \left[X(t_1) - m_x(t_1) \quad X(t_2) - m_x(t_2) \right] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 - m_x(t_1) \quad x_2 - m_x(t_2) f_2(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned} \quad (5)$$

З (5) видно, що у випадку $t_1 = t_2 = t$

$$K_x(t, t) = D_x(t). \quad (6)$$

Отже, за основні характеристики випадкової функції можна взяти її математичне сподівання та кореляційну функцію.

Наведемо такі властивості останньої:

1. Симетричність:

$$K_x(t_1, t_2) = K_x(t_2, t_1). \quad (7)$$

Це випливає з того, що коваріація випадкових величин не залежить від послідовності, в якій ці величини розглядаються.

2. Абсолютна величина кореляційної функції не перевищує добутку середніх квадратичних відхилень відповідних перерізів:

$$K_x(t_1, t_2) \leq \sigma_x(t_1) \sigma_x(t_2) = \sqrt{K_x(t_1, t_1) K_x(t_2, t_2)}.$$

Це випливає з відповідної нерівності для коваріації випадкових величин, а саме: коваріація випадкових величин за абсолютною величиною не перевищує добутку середніх квадратичних відхилень цих величин (див. лекцію 14, приклад 3).

3. Якщо до випадкової функції $X(t)$ додати невипадкову функцію $\phi(t)$, то її кореляційна функція не зміниться. Дійсно, нехай $Y(t) = X(t) + \phi(t)$. Тоді з властивості математичного сподівання $m_y(t) = m_x(t) + \phi(t)$. Отже, $Y(t) - m_y(t) = X(t) - m_x(t)$. Враховуючи (5), отримуємо $K_y(t_1, t_2) = K_x(t_1, t_2)$.

4. При множенні випадкової функції $X(t)$ на невипадкову функцію $\varphi(t)$ її кореляційну функцію слід помножити на $\varphi(t_1) \varphi(t_2)$.

Дійсно, нехай $Y(t) = \varphi(t)X(t)$. Тоді $m_y(t) = \varphi(t)m_x(t)$.

Далі, $Y(t) - m_y(t) = \varphi(t)X(t) - \varphi(t)m_x(t)$. Отже,

$$K_y(t_1, t_2) = M \begin{bmatrix} Y(t_1) - m_y(t_1) & Y(t_2) - m_y(t_2) \end{bmatrix} = \\ = M \begin{bmatrix} \varphi(t_1)X(t_1) - \varphi(t_1)m_x(t_1) & \varphi(t_2)X(t_2) - \varphi(t_2)m_x(t_2) \end{bmatrix} = \varphi(t_1)\varphi(t_2)K_x(t_1, t_2).$$

Іноді розглядають нормовану кореляційну функцію

$$k_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sigma_x(t_1)\sigma_x(t_2)}, \quad (8)$$

що є, очевидно, коефіцієнтом кореляції перерізів $X|t_1$ і $X|t_2$ випадкової функції.

Як і коефіцієнт кореляції, функція $k_x(t_1, t_2)$ змінюється у проміжку $-1; 1$ (див. лекцію 14). При $t_1 = t_2 = t$ $k_x(t, t) = 1$.

Нормована кореляційна функція має той же ймовірнісний смисл, що і коефіцієнт кореляції, а саме: чим більше модуль цієї функції до 1, тим **лінійний** зв'язок між відповідними перерізами сильніший; чим більше модуль до нуля, тим цей зв'язок слабший.

Нехай одночасно розглядаються дві випадкові функції $X|t$ та $Y(u)$. Оцінку рівня залежності перерізів цих функцій проводять за допомогою **взаємної кореляційної функції**.

Означення 5. Взаємною кореляційною функцією двох випадкових функцій $X|t$ і $Y(u)$ називається невипадкова функція двох аргументів $K_{xy}(t, u)$, яка при кожній парі значень t, u дорівнює коваріації відповідних перерізів випадкових функцій $X|t$ і $Y(u)$:

$$K_{xy}(t, u) = M \begin{bmatrix} X(t) - m_x(t) & Y(u) - m_y(u) \end{bmatrix}.$$

Якщо випадкові функції є функціями одного й того ж аргументу, то $K_{xy}(t_1, t_2) = M \begin{bmatrix} X(t_1) - m_x(t_1) & Y(t_2) - m_y(t_2) \end{bmatrix} = M[\overset{\circ}{X}(t_1)\overset{\circ}{Y}(t_2)]$, де $\overset{\circ}{X}(t) = X(t) - m_x(t)$, $\overset{\circ}{Y}(t) = Y(t) - m_y(t)$ – центровані випадкові функції.

Функції $X|t$ і $Y(u)$ називають **корельованими**, якщо їх взаємна кореляційна функція не дорівнює тотожно нулю. У протилежному разі, коли $K_{xy}(t, u) \equiv 0$, функції $X|t$ і $Y(u)$ називають **некорельованими**.

Нормована взаємна кореляційна функція визначається так:

$$k_{xy}(t, u) = \frac{K_{xy}(t, u)}{\sigma_x(t) \sigma_y(u)}.$$

Розглянемо ряд прикладів.

Приклад 1. Випадкова функція має вигляд $X(t) = Vt + b$, де V – випадкова величина, розподілена за нормальним законом $N(m, \sigma)$, а b – невипадкова постійна. Знайти одновимірну щільність $f_1(x; t)$ і основні характеристики випадкової функції: $m_x(t)$, $K_x(t_1, t_2)$.

Розв'язання. Очевидно, що перерізи випадкової функції $X(t)$ мають нормальній розподіл. Знайдемо $m_x(t)$, $\sigma_x(t)$, скориставшись властивостями математичного сподівання та дисперсії випадкових величин: $m_x(t) = M[Vt + b] = tM[V] + b = tm + b$; $D_x = t^2 D[V] = t^2 \sigma^2$;

$$\sigma_x(t) = |t| \sigma. \text{ Отже, } f_1(x; t) = \frac{1}{|t| \sigma \sqrt{2\pi}} e^{-\frac{[x - mt + b]^2}{2t^2 \sigma^2}}.$$

Кореляційну функцію можна знайти безпосередньо за означенням (5). Проте зручніше скористатися властивостями кореляційної функції 3,4, врахувавши, що $\varphi(t) = t$; $K_V(t_1, t_2) = D[V] = \sigma^2$. Отже, $K_x(t_1, t_2) = t_1 t_2 \sigma^2$. Звідси, зокрема, $D_x(t) = K_x(t, t) = t^2 \sigma^2$.

Приклад 2. Випадковий процес має вигляд випадкової гармоніки $X(t) = A \cos(\omega t + \Phi)$, де $A > 0$ – випадкова амплітуда, для якої існує $M[A^2]$; Φ – незалежна від A випадкова фаза з рівномірним розподілом на відрізку $0; 2\pi$; ω – невипадкова постійна (частота гармоніки). Знайти нормовану кореляційну функцію процесу.

Розв'язання. Перепишемо випадковий процес у вигляді $X(t) = V_1 \cos(\omega t) - V_2 \sin(\omega t)$, де $V_1 = A \cos \Phi$, $V_2 = A \sin \Phi$.

Оскільки A і Φ – незалежні випадкові величини, то

$$M[V_1] = M[A \cos \Phi] = M[A] \frac{1}{2\pi} \int_0^{2\pi} \cos \varphi d\varphi = 0. \text{ Аналогічно,}$$

$M[V_2] = 0$. Тому $M[X(t)] = m_x(t) \equiv 0$. Отже, випадкова гармоніка є чисто флюктуаційним процесом. За означенням (5) знаходимо

$$K_x(t_1, t_2) = M[V_1 \cos(\omega t_1) - V_2 \sin(\omega t_1) V_1 \cos(\omega t_2) - V_2 \sin(\omega t_2)] =$$

$$= M \left[V_1^2 \cos \omega t_1 \cos \omega t_2 + V_2^2 \sin \omega t_1 \sin \omega t_2 - V_1 V_2 \sin \omega t_1 + t_2 \right].$$

Знайдемо $M[V_1^2]$, $M[V_2^2]$, $M[V_1 V_2]$.

$$M[V_1^2] = M[A^2] M[\cos^2 \Phi] = M[A^2] \frac{1}{2\pi} \int_0^{2\pi} \cos^2 \varphi d\varphi = \frac{1}{2} M[A^2].$$

Аналогічно, $M[V_2^2] = \frac{1}{2} M[A^2]$.

$$M[V_1 V_2] = M[A^2] M[\cos \Phi \sin \Phi] = M[A^2] \frac{1}{2\pi} \int_0^{2\pi} \cos \varphi \sin \varphi d\varphi = 0.$$

Отже, $K_x[t_1, t_2] = \frac{1}{2} M[A^2] \cos \omega t_1 - t_2$.

При $t_1 = t_2 = t$ отримуємо дисперсію випадкової гармоніки

$$D[X(t)] = K_x(t, t) = \frac{1}{2} M[A^2].$$

Для нормованої кореляційної функції $k_x[t_1, t_2]$ маємо

$$k_x[t_1, t_2] = \cos \omega t_1 - t_2.$$

Приклад 3. Нехай випадкові процеси $X(t)$ та $Y(t)$ мають вигляд

$X(t) = V_1 \cos t - V_2 \sin t$, $Y(t) = V_2 \cos t + V_1 \sin t$, причому V_1 та V_2 – некорельовані стандартизовані випадкові величини. Знайти взаємну кореляційну функцію даних процесів.

Розв'язання. За умовою $M[V_1] = M[V_2] = 0$; $D[V_1] = D[V_2] = 1$;

$$M[V_1^2] = D[V_1] + M^2[V_1] = 1; \text{ аналогічно, } M[V_2^2] = 1;$$

$$M[V_1 V_2] = M[V_1] M[V_2] = 0.$$

Тому $m_x(t) \equiv 0$; $m_y(t) \equiv 0$. За означенням взаємної кореляційної функції $K_{xy}[t_1, t_2] = M[V_1 \cos t_1 - V_2 \sin t_1 | V_2 \cos t_2 + V_1 \sin t_2] =$

$$= M[V_1 V_2 \cos t_1 \cos t_2 + V_1^2 \cos t_1 \sin t_2 - V_2^2 \sin t_1 \cos t_2 - V_1 V_2 \sin t_1 \sin t_2] =$$

$$= \cos t_1 \sin t_2 - \sin t_1 \cos t_2 = \sin(t_2 - t_1).$$

Легко перевірити, що $K_{xy}[t_1, t_2]$ є одночасно і нормованою взаємною кореляційною функцією $k_{xy}[t_1, t_2]$.

Лекція № 27

Невід'ємна визначеність кореляційної функції. Канонічний розклад випадкової функції. Похідна та інтеграл від випадкової функції, що має канонічний розклад

Характерною рисою кореляційної функції $K_x(t_1, t_2)$ будь-якої випадкової функції $X(t)$ є її **невід'ємна визначеність**. Це означає, що для довільного натурального n , довільних дійсних чисел x_1, x_2, \dots, x_n і довільного набору значень аргументів t_1, t_2, \dots, t_n з області визначення кореляційної функції виконується нерівність

$$\sum_{i=1}^n \sum_{j=1}^n K_x(t_i, t_j) x_i x_j \geq 0. \quad (1)$$

$$\begin{aligned} \text{Дійсно, } & \sum_{i=1}^n \sum_{j=1}^n K_x(t_i, t_j) x_i x_j = \sum_{i=1}^n \sum_{j=1}^n M \left[\begin{array}{cc} \overset{\circ}{X}(t_i) & \overset{\circ}{X}(t_j) \end{array} \right] x_i x_j = \\ & = M \left[\sum_{i=1}^n \overset{\circ}{X}(t_i) x_i \cdot \sum_{j=1}^n \overset{\circ}{X}(t_j) x_j \right] = M \left[\left(\sum_{i=1}^n \overset{\circ}{X}(t_i) x_i \right)^2 \right] \geq 0. \end{aligned}$$

Тут ми скористались тим, що знаки суми і математичного сподівання можна міняти місцями.

У лекції 26 безпосередньо за означенням ми довели ряд властивостей кореляційної функції. Легко показати, що деякі з цих властивостей є простими наслідками невід'ємності (1). Покладемо, наприклад, в (1) $n=1$. Тоді $K_x(t_1, t_1) x_1^2 \geq 0$ і $K_x(t_1, t_1) \geq 0$.

При $n=2$ з (1) і властивості симетрії $K_x(t_1, t_2) = K_x(t_2, t_1)$ одержуємо

$$K_x(t_1, t_1) x_1^2 + 2K_x(t_1, t_2) x_1 x_2 + K_x(t_2, t_2) x_2^2 \geq 0. \quad (2)$$

Умовою невід'ємної визначеності квадратичної форми (2) є

$$K_x^2(t_1, t_2) \leq K_x(t_1, t_1) K_x(t_2, t_2). \quad (3)$$

Взагалі, можна показати, що невід'ємна визначеність (1) є необхідною і достатньою умовою для того, щоб функція $K_x(t_1, t_2)$ могла бути кореляційною функцією. Необхідність нами вже показана. Достатність випливає з того, що для довільної невід'ємної визначененої функції $K_x(t_1, t_2)$ можна побудувати випадкову функцію з нормальними багатовимірними щільностями, для якої $K_x(t_1, t_2)$ буде кореляційною функцією (див. лекції 10, 13).

Розглянемо суму двох випадкових функцій $V_1(t)$ та $V_2(t)$:

$$X(t) = V_1(t) + V_2(t). \quad (4)$$

Очевидно, що $m_x(t) = m_{V_1}(t) + m_{V_2}(t)$. Знайдемо кореляційну функцію суми (4), скориставшись рівністю $\dot{X}(t) = \dot{V}_1(t) + \dot{V}_2(t)$, де

$\dot{X}(t) = X(t) - m_x(t)$, $\dot{V}_i(t) = V_i(t) - m_{V_i}(t)$ — центровані випадкові функції. Отже,

$$\begin{aligned} K_x(t_1, t_2) &= M\left[\begin{array}{cc} \dot{X}(t_1) & \dot{X}(t_2) \end{array}\right] = M\left[\left(\begin{array}{c} \dot{V}_1(t_1) \\ \dot{V}_2(t_1) \end{array}\right) \left(\begin{array}{c} \dot{V}_1(t_2) \\ \dot{V}_2(t_2) \end{array}\right)\right] = \\ &= M\left[\begin{array}{cc} \dot{V}_1(t_1) & \dot{V}_1(t_2) \end{array}\right] + M\left[\begin{array}{cc} \dot{V}_2(t_1) & \dot{V}_2(t_2) \end{array}\right] + M\left[\begin{array}{cc} \dot{V}_1(t_1) & \dot{V}_2(t_2) \end{array}\right] + M\left[\begin{array}{cc} \dot{V}_1(t_2) & \dot{V}_2(t_1) \end{array}\right] = \\ &= K_{V_1}(t_1, t_2) + K_{V_2}(t_1, t_2) + K_{V_1 V_2}(t_1, t_2) + K_{V_2 V_1}(t_2, t_1). \end{aligned}$$

У випадку, коли $V_1(t)$ та $V_2(t)$ — **некорельовані** випадкові функції (див. лекцію 26),

$$K_x(t_1, t_2) = K_{V_1}(t_1, t_2) + K_{V_2}(t_1, t_2). \quad (5)$$

Властивість (5) легко узагальнюється на довільну кількість доданків, а саме, кореляційна функція суми довільної кількості **попарно некорельованих** випадкових функцій дорівнює сумі кореляційних

функцій доданків. Отже, якщо $X = \sum_{k=1}^n V_k(t)$ і

$$K_{V_i V_j}(t_1, t_2) = M\left[\begin{array}{cc} \dot{V}_i(t_1) & \dot{V}_j(t_2) \end{array}\right] \equiv 0 \text{ при } i \neq j, \text{ то}$$

$$K_x(t_1, t_2) = \sum_{k=1}^n K_{V_k}(t_1, t_2). \quad (6)$$

Нехай випадкова функція $X(t)$ має вигляд

$$X(t) = m_x(t) + \sum_{k=1}^n \dot{V}_k(t) \phi_k(t), \quad (7)$$

де $\phi_k(t)$, $k = 1, 2, \dots, n$ — невипадкові функції, а \dot{V}_k — центровані $\left(M\left[\dot{V}_k\right] = 0\right)$ і попарно некорельовані $\left(M\left[\dot{V}_i \dot{V}_j\right] = 0, i \neq j\right)$

випадкові величини з дисперсіями $M\left[\dot{V}_k \dot{V}_k\right] = D_k$, $k = 1, 2, \dots, n$.

Означення 1. Зображення випадкової функції у вигляді (7) називають **канонічним розкладом** цієї функції, а невипадкові функції $\varphi_k t$, $k = 1, 2, \dots, n$ – **координатними функціями** канонічного розкладу.

Оскільки випадкові функції $\overset{\circ}{V}_k \varphi_k t$, $k = 1, 2, \dots, n$ попарно некорельовані і мають кореляційні функції $\varphi_k t_1 \varphi_k t_2 D_k$, $k = 1, 2, \dots, n$ (див. лекцію 26, власт. 4 кореляційної функції), то згідно з (6) кореляційна функція для $X t$ з (7) приймає вигляд

$$K_x(t_1, t_2) = \sum_{k=1}^n \varphi_k(t_1) \varphi_k(t_2) D_k. \quad (8)$$

Якщо покласти $t_1 = t_2 = t$, то отримаємо вираз для дисперсії

$$D_x(t) = K_x(t, t) = \sum_{k=1}^n \varphi_k^2(t) D_k. \quad (9)$$

Означення 2. Вираз (8) називається **канонічним розкладом кореляційної функції**.

Отже, якщо канонічний розклад випадкової функції відомий, то відразу ж знаходиться і канонічний розклад її кореляційної функції. Справедливе обернене твердження: якщо кореляційна функція допускає канонічний розклад (8), то існує канонічний розклад випадкової функції (7), в якому випадкові величини $\overset{\circ}{V}_k$ мають дисперсії D_k , $k = 1, 2, \dots, n$. Доведення цього факту ми не наводимо.

Приклад 1. Нехай випадкова функція $X t$ має вигляд

$$X(t) = \sum_{k=1}^n U_k \psi_k(t), \quad (10)$$

де $\psi_k(t)$, $k = 1, 2, \dots, n$ – невипадкові функції, а U_k – довільні випадкові величини. Чи існує для випадкової функції $X t$ канонічний розклад (7)?

Розв'язання. Якщо U_k , $k = 1, 2, \dots, n$ – попарно некорельовані, то шляхом їх центрування за допомогою тотожного перетворення приходимо до канонічного розкладу (7):

$$X(t) = \sum_{k=1}^n m_k \psi_k(t) + \sum_{k=1}^n U_k - m_k \psi_k(t) = m_x(t) + \sum_{k=1}^n \overset{\circ}{U}_k \psi_k(t), \quad (11)$$

де $m_k = M(U_k)$, $k = 1, 2, \dots, n$. Тому цей випадок не є для нас цікавим.

Нехай $\mathbf{K}_u = \begin{pmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ \dots & K_{22} & \dots & K_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & K_{nn} \end{pmatrix}$ – коваріаційна матриця системи

випадкових величин U_1, U_2, \dots, U_n , причому існують $K_{ij} = M \left[\overset{\circ}{U_i} \overset{\circ}{U_j} \right]$, відмінні від нуля при $i \neq j$.

Очевидно, що в цьому випадку вираз (11) не є канонічним розкладом.

Знайдемо кореляційну функцію $K_x(t_1, t_2)$. З врахуванням (11) маємо

$$\begin{aligned} K_x(t_1, t_2) &= M \left[\sum_{i=1}^n \overset{\circ}{U_i} \psi_i(t_1) \cdot \sum_{j=1}^n \overset{\circ}{U_j} \psi_j(t_2) \right] = \sum_{i=1}^n \sum_{j=1}^n M \left[\overset{\circ}{U_i} \overset{\circ}{U_j} \right] \psi_i(t_1) \psi_j(t_2) = \\ &= \sum_{i=1}^n \sum_{j=1}^n K_{ij} \psi_i(t_1) \psi_j(t_2). \end{aligned} \quad (12)$$

Останній вираз можна розглядати як симетричну $K_{ij} = K_{ji}$ білінійну форму. З курсу лінійної алгебри відомо, що цю форму завжди можна звести шляхом лінійних перетворень до канонічного вигляду типу (8). Звідси випливає, що для випадкової функції (10) з довільними випадковими коефіцієнтами U_k існує канонічний розклад (7).

Приклад 2. За допомогою лінійного перетворення звести вираз (10) з прикладу 1 до канонічного вигляду (7).

Розв'язання. Перейдемо до матричних позначень, ввівши вектор-стовпці

$$\overset{\circ}{\mathbf{U}} = \begin{pmatrix} \overset{\circ}{U_1} \\ \overset{\circ}{U_2} \\ \vdots \\ \overset{\circ}{U_n} \end{pmatrix}, \quad \boldsymbol{\psi}(t) = \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \\ \vdots \\ \psi_n(t) \end{pmatrix}. \quad \text{Тоді вираз (11) можна переписати у}$$

вигляді скалярного добутку

$$\overset{\circ}{X}(t) = \overset{\circ}{\mathbf{U}}^T \boldsymbol{\psi}(t), \quad (13)$$

де символ T означає транспонування, тобто $\overset{\circ}{\mathbf{U}}^T = \begin{pmatrix} \overset{\circ}{U}_1 & \overset{\circ}{U}_2 & \dots & \overset{\circ}{U}_n \end{pmatrix}$ – вектор-рядок. Коваріаційна матриця приймає вигляд

$$\mathbf{K}_u = M \begin{bmatrix} \overset{\circ}{\mathbf{U}} \overset{\circ}{\mathbf{U}}^T \end{bmatrix}. \quad (14)$$

Нехай $\overset{\circ}{V} = \overset{\circ}{A} \overset{\circ}{U}$ – новий вектор випадкових величин. Матрицю $\overset{\circ}{A}$ вибираємо так, щоб випадкові компоненти $\overset{\circ}{V}_k$ вектора $\overset{\circ}{V}$ були попарно некорельованими. Маємо

$$\mathbf{K}_v = M \begin{bmatrix} \overset{\circ}{V} \overset{\circ}{V}^T \end{bmatrix} = M \begin{bmatrix} \overset{\circ}{A} \overset{\circ}{U} \overset{\circ}{U}^T \overset{\circ}{A}^T \end{bmatrix} = A M \begin{bmatrix} \overset{\circ}{\mathbf{U}} \overset{\circ}{\mathbf{U}}^T \end{bmatrix} A^T = A \mathbf{K}_u A^T = \mathbf{D}_v, \quad (15)$$

де \mathbf{D}_v – діагональна матриця дисперсій нових компонент. Отже, матриця A повинна бути такою, що зводить коваріаційну матрицю \mathbf{K}_u до діагонального вигляду. В курсі лінійної алгебри доводиться, що матриця A будеться так:

1. Знаходяться власні значення матриці \mathbf{K}_u шляхом розв'язування рівняння

$$\det \mathbf{K}_u - \lambda \mathbf{I} = 0, \quad (*)$$

де \mathbf{I} – одинична матриця. Оскільки \mathbf{K}_u – симетрична матриця, всі власні значення $\lambda_k \geq 0$, $k = 1, 2, \dots, n$.

2. Для кожного власного значення λ_k , $k = 1, 2, \dots, n$ знаходиться відповідний йому нормований власний вектор $\mathbf{a}_k^T = [a_{k1}, a_{k2}, \dots, a_{kn}]$, $k = 1, 2, \dots, n$ шляхом розв'язування системи лінійних рівнянь

$$\mathbf{K}_u - \lambda_k \mathbf{I} \mathbf{a}_k = 0 \quad (**)$$

за умови $a_{k1}^2 + a_{k2}^2 + \dots + a_{kn}^2 = 1$.

3. Матриця A – це матриця, k -ий рядок якої містить координати k -го нормованого власного вектора $\mathbf{a}_k^T = [a_{k1}, a_{k2}, \dots, a_{kn}]$, $k = 1, 2, \dots, n$.

Якщо всі власні значення λ_k , $k = 1, 2, \dots, n$ різні, то матрицю A можна сформувати $n!$ способами. Якщо серед власних значень є кратні, існує безліч способів формування матриці A .

Нехай матрицю A побудовано і знайдено новий вектор випадкових величин $\vec{V} = A \vec{U}$ з коваріаційною матрицею $K_v = D_v$ (на

діагоналі матриці D_v стоять власні значення λ_k , $k = 1, 2, \dots, n$). При цьому повинен залишитись незмінним скалярний добуток (13). Це досягається вибором нового вектора координатних функцій $\varphi(t) = A\psi(t)$. Дійсно,

$$\dot{\vec{X}}(t) = \vec{U}^T \vec{\psi}(t) = \left(A^{-1} \vec{V} \right)^T \vec{\psi}(t) = \vec{V}^T A^{-1} \vec{\psi}(t) = \vec{V}^T A \vec{\psi}(t) = \vec{V}^T \varphi(t). \quad (16)$$

Звідси одержуємо вираз для $\varphi(t)$. В (16) враховано, що $A^{-1} = A^T$ внаслідок ортогональності матриці A . За знайденими векторами випадкових величин \vec{V} і координатних функцій $\varphi(t)$ будуємо канонічний розклад (7). Відмітимо, що задача має не єдиний розв'язок. Результат залежить від способу формування ортогональної матриці A .

Приклад 3. Випадковий процес $X(t)$ задано у вигляді:

$$X(t) = t + U_1 \cos t + U_2 \sin t, \quad M(U_1) = 1, \quad M(U_2) = 2, \quad K_u = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Звести $X(t)$ до канонічного вигляду.

Розв'язання.

$$m_x(t) = M[X(t)] = t + M(U_1) \cos t + M(U_2) \sin t = t + \cos t + 2 \sin t.$$

Перетворимо випадкову функцію $X(t)$ до вигляду (11):

$$\begin{aligned} X(t) &= m_x(t) + t + U_1 \cos t + U_2 \sin t - t - \cos t - 2 \sin t = \\ &= m_x(t) + U_1 - 1 \cos t + U_2 - 2 \sin t = m_x(t) + \dot{U}_1 \cos t + \dot{U}_2 \sin t, \end{aligned}$$

$$\text{де } \dot{U}_1 = U_1 - 1, \quad \dot{U}_2 = U_2 - 2.$$

1. Знайдемо власні значення матриці K_u . Рівняння (*) приймає вигляд $\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0 \Leftrightarrow 2-\lambda^2-1=0 \Leftrightarrow \lambda_1=1; \lambda_2=3$.
2. Знайдемо нормовані власні вектори. Система (**) приймає вигляд $\begin{cases} 2-\lambda_1 a_1 + a_2 = 0 \\ a_1 + 2-\lambda_2 a_2 = 0 \end{cases}$. При цьому повинна виконуватись умова $a_1^2 + a_2^2 = 1$.

$$a) \quad \lambda = 1: \begin{cases} a_1 + a_2 = 0 \\ a_1^2 + a_2^2 = 1 \end{cases} \Leftrightarrow \begin{cases} a_1 = -a_2 \\ 2a_1^2 = 1 \end{cases} \Leftrightarrow \begin{cases} a_1 = \frac{1}{\sqrt{2}} \\ a_2 = -\frac{1}{\sqrt{2}} \end{cases}. \quad \text{Отже,}$$

першим власним нормованим вектором є $\mathbf{a}_1^T = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$.

$$b) \lambda = 3: \begin{cases} -a_1 + a_2 = 0 \\ a_1^2 + a_2^2 = 1 \end{cases} \Leftrightarrow \begin{cases} a_1 = a_2 \\ 2a_1^2 = 1 \end{cases} \Leftrightarrow \begin{cases} a_1 = \frac{1}{\sqrt{2}} \\ a_2 = \frac{1}{\sqrt{2}} \end{cases}.$$

Маємо $\mathbf{a}_2^T = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$ – другий власний ортонормований вектор.

$$3. \quad \text{Будуємо матрицю } \mathbf{A}: \mathbf{A} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Знаходимо новий вектор випадкових величин:

$$\overset{\circ}{\mathbf{V}} = \mathbf{A} \overset{\circ}{\mathbf{U}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \overset{\circ}{U}_1 \\ \overset{\circ}{U}_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \overset{\circ}{U}_1 - \overset{\circ}{U}_2 \\ \overset{\circ}{U}_1 + \overset{\circ}{U}_2 \end{pmatrix} = \begin{pmatrix} \overset{\circ}{V}_1 \\ \overset{\circ}{V}_2 \end{pmatrix}.$$

Знаходимо новий вектор координатних функцій:

$$\overset{\circ}{\boldsymbol{\varphi}}(t) = \mathbf{A} \overset{\circ}{\boldsymbol{\psi}}(t) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos t - \sin t \\ \cos t + \sin t \end{pmatrix} = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \end{pmatrix}.$$

Отже, канонічним розкладом випадкової функції є

$$X(t) = m_x(t) + \overset{\circ}{\mathbf{V}}^T \overset{\circ}{\boldsymbol{\varphi}}(t) = m_x(t) + \overset{\circ}{V}_1 \varphi_1(t) + \overset{\circ}{V}_2 \varphi_2(t),$$

$$\text{де } \overset{\circ}{V}_1 = \frac{1}{\sqrt{2}} \left(\overset{\circ}{U}_1 - \overset{\circ}{U}_2 \right), \quad \overset{\circ}{V}_2 = \frac{1}{\sqrt{2}} \left(\overset{\circ}{U}_1 + \overset{\circ}{U}_2 \right), \quad \varphi_1(t) = \frac{1}{\sqrt{2}} \cos t - \sin t,$$

$\varphi_2(t) = \frac{1}{\sqrt{2}} \cos t + \sin t$. Для кореляційної функції канонічний розклад

має вигляд

$$K_x(t_1, t_2) = \lambda_1 \varphi_1(t_1) \varphi_1(t_2) + \lambda_2 \varphi_2(t_1) \varphi_2(t_2) = \varphi_1(t_1) \varphi_1(t_2) + 3 \varphi_2(t_1) \varphi_2(t_2).$$

Похідна та інтеграл від випадкової функції, що має канонічний розклад

У даній лекції обмежимося класом випадкових функцій, що допускають канонічний розклад (7). Похідну та інтеграл для таких випадкових функцій природно ввести як випадкові функції, що визначаються формулами

$$Y(t) = \frac{dX(t)}{dt} = \frac{dm_x(t)}{dt} + \sum_{k=1}^n \overset{\circ}{V}_k \frac{d\phi_k(t)}{dt}, \quad (17)$$

$$Z(t) = \int_a^t X(s) ds = \int_a^t m_x(s) ds + \sum_{k=1}^n \overset{\circ}{V}_k \int_a^t \phi_k(s) ds.$$

Очевидно, що

$$M\left[\frac{dX(t)}{dt}\right] = \frac{dm_x(t)}{dt}; \quad M\left[\int_a^t X(s) ds\right] = \int_a^t m_x(s) ds. \quad (18)$$

Оскільки під знаками сум в (17) стоять попарно некорельовані випадкові функції, то для кореляційних функцій похідної і інтеграла згідно з (6) можна записати їх канонічні розклади

$$K_y(t_1, t_2) = \sum_{k=1}^n \frac{d\phi_k(t_1)}{dt} \frac{d\phi_k(t_2)}{dt} D_k, \quad (19)$$

$$K_z(t_1, t_2) = \sum_{k=1}^n \int_a^{t_1} \phi_k(s_1) ds_1 \int_a^{t_2} \phi_k(s_2) ds_2 \cdot D_k.$$

Якщо врахувати (8), то останні рівності можна подати у вигляді

$$K_y(t_1, t_2) = \frac{\partial}{\partial t_1} \frac{\partial}{\partial t_2} K_x(t_1, t_2), \quad K_z(t_1, t_2) = \int_a^{t_1} \int_a^{t_2} K_x(s_1, s_2) ds_1 ds_2. \quad (20)$$

Отже, введення понять похідної та інтеграла від випадкової функції, яка допускає канонічний розклад (7), не викликає ніяких ускладнень. Для цього нам знадобились звичайні операції математичного аналізу. Питання про диференційовність та інтегровність випадкової функції зводиться до питання про диференційовність та інтегровність невипадкових функцій $m_x(t)$ і $\phi_k(t)$, $k=1, 2, \dots, n$ у звичайному розумінні. Випадок, коли в (7) $n=\infty$, ми поки що не розглядаємо.

Щоб знайти похідну чи інтеграл від випадкової функції, яка допускає канонічний розклад, не обов'язково шукати цей розклад. Так, випадкову функцію вигляду (10) можна диференціювати чи інтегрувати безпосередньо, не зводячи її до канонічного вигляду (7). Легко перевірити, що властивості (18) і (20) залишаються в силі.

Приклад 4. На вхід диференціюючого пристрою подається випадковий процес, що має канонічний розклад $X(t) = t + \overset{\circ}{V}_1 \cos t + \overset{\circ}{V}_2 \sin t$, $D\left[\overset{\circ}{V}_1\right] = 1$, $D\left[\overset{\circ}{V}_2\right] = 2$. Знайти математичне сподівання, дисперсію та кореляційну функцію процесу на виході цього пристрою.

Розв'язання. Оскільки $M[X] = m_x(t) = t$, а

$K_x(t_1, t_2) = \cos t_1 \cos t_2 + 2 \sin t_1 \sin t_2$, то згідно з (18), (20)

$$M[Y(t)] = M\left[\frac{dX(t)}{dt}\right] = \frac{dm_x(t)}{dt} = 1;$$

$$K_y(t_1, t_2) = \frac{\partial}{\partial t_1} \frac{\partial}{\partial t_2} K_x(t_1, t_2) = \frac{\partial}{\partial t_2} -\sin t_1 \cos t_2 + 2 \cos t_1 \sin t_2 =$$

$$= \sin t_1 \sin t_2 + 2 \cos t_1 \cos t_2; D[Y(t)] = K_y(t; t) = \sin^2 t + 2 \cos^2 t = 1 + \cos^2 t.$$

На виході диференціюючого пристрою маємо випадковий процес $Y(t) = \frac{dX(t)}{dt} = 1 - \overset{\circ}{V}_1 \sin t + \overset{\circ}{V}_2 \cos t$. Тому характеристики цього процесу можна знайти безпосередньо, не користуючись співвідношеннями (18), (20).

Приклад 5. На вхід інтегратора, що працює за принципом $z(t) = \int_0^t x(s) ds$, де $x(s)$ – довільна реалізація випадкового процесу на вході, подається випадковий процес $X(t)$ з математичним сподіванням $m_x(t) = 1 + \sin^2 \omega t$ і кореляційною функцією $K_x(t_1, t_2) = 4 \cos \omega t_1 \cos \omega t_2$. Знайти математичне сподівання і кореляційну функцію випадкового процесу $Z(t)$ на виході інтегратора.

Розв'язання. Згідно з (18) маємо: $M[Z(t)] = m_z(t) = \int_0^t m_x(s) ds =$

$$= \int_0^t 1 + \sin^2 \omega s ds = \int_0^t ds + \int_0^t \frac{1 - \cos 2\omega s}{2} ds = \frac{3}{2}t - \frac{1}{4\omega} \sin 2\omega t.$$

Кореляційну функцію на виході знаходимо, скориставшись другим співвідношенням (20):

$$K_z(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} K_x(s_1, s_2) ds_1 ds_2 = \int_0^{t_1} \int_0^{t_2} 4 \cos \omega s_1 \cos \omega s_2 ds_1 ds_2 =$$

$$= 4 \int_0^{t_1} \cos \omega s_1 ds_1 \int_0^{t_2} \cos \omega s_2 ds_2 = \frac{4}{\omega^2} \sin \omega t_1 \sin \omega t_2.$$

Лекція № 28

Диференцювання та інтегрування випадкових функцій у середньоквадратичному

Клас випадкових функцій, що допускають канонічний розклад, є надто вузьким, тоді як поняття похідної та інтеграла бажано поширити на довільні випадкові функції. Оскільки математичне сподівання та кореляційна функція розглядаються нами як основні характеристики випадкової функції $X(t)$, то було б, звичайно, природним ввести

узагальнені поняття похідної $Y(t) = \frac{dX(t)}{dt}$ та інтеграла $\int_a^t X(s) ds$

таким чином, щоб співвідношення

$$M\left[\frac{dX(t)}{dt} \right] = \frac{dm_x(t)}{dt}; \quad K_y(t_1, t_2) = \frac{\partial}{\partial t_1} \frac{\partial}{\partial t_2} K_x(t_1, t_2), \quad (1)$$

$$M\left[\int_a^t X(s) ds \right] = \int_a^t m_x(s) ds; \quad K_z(t_1, t_2) = \int_a^{t_1} \int_a^{t_2} K_x(t_1, t_2) ds_1 ds_2, \quad (2)$$

одержані нами в лекції 27 для частинного випадку, не змінювались.

В основі загальних означенень похідної та інтеграла від випадкової функції лежить поняття *середньоквадратичної збіжності*.

Означення 1. Кажуть, що випадкова функція $X(t)$ збігається у середньоквадратичному при $t \rightarrow t_0$ до випадкової величини X_0 (коротко це позначають так: $\lim_{t \rightarrow t_0} l.i.m. X(t) = X_0$), якщо

$$\lim_{t \rightarrow t_0} M[X(t) - X_0]^2 = 0. \quad (3)$$

Означення 2. Випадкова функція $X(t)$ називається *неперервною у середньоквадратичному* в точці t_0 , якщо

$$\lim_{t \rightarrow t_0} l.i.m. X(t) = X(t_0). \quad (4)$$

Зауваження 1. Запишемо рівність

$$M[X(t) - X(t_0)]^2 = D[X(t) - X(t_0)] + m_x(t) - m_x(t_0)^2, \quad (5)$$

де, зокрема,

$$D[X(t) - X(t_0)] = K_x(t, t) + K_x(t_0, t_0) - 2K_x(t, t_0). \quad (6)$$

З (5), (6) видно, що як тільки математичне сподівання $m_x t$ є неперервною функцією в точках $t_0 \in a, b$, а кореляційна функція $K_x(t_1, t_2)$ – неперервною функцією в точках t_0, t_1, t_2 , $t_0 \in a, b$, то випадкова функція $X(t)$ буде неперервною у середньоквадратичному в точках $t_0 \in a, b$. Можна показати, що справедливе й обернене твердження, а саме: з неперервності випадкової функції $X(t)$ у середньоквадратичному в точках $t_0 \in a, b$ випливає відмічена вище неперервність математичного сподівання та кореляційної функції. Більше того, неперервність кореляційної функції має місце у всіх точках (t_1, t_2) , $t_1, t_2 \in (a, b)$. Отже, неперервність у середньоквадратичному випадкової функції зводиться фактично до неперервності невипадкових функцій $m_x t$ і $K_x(t_1, t_2)$ у звичайному розумінні.

Перейдемо до означення похідної випадкової функції.

Означення 3. Похідною (похідною у середньоквадратичному або середньоквадратичною похідною) випадкової функції $X(t)$

називається випадкова функція $Y(t) = \frac{dX(t)}{dt}$, яка визначається як границя у середньоквадратичному відношенні приросту випадкової функції до приросту невипадкового аргумента

$$Y(t) = \frac{dX(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{X(t + \Delta t) - X(t)}{\Delta t}, \quad (7)$$

тобто така випадкова функція $Y(t) = \frac{dX(t)}{dt}$, для якої

$$\lim_{\Delta t \rightarrow 0} M \left[\left(\frac{X(t + \Delta t) - X(t)}{\Delta t} - Y(t) \right)^2 \right] = 0. \quad (8)$$

Якщо $X(t)$ – невипадкова функція, то очевидно, що середньоквадратична похідна є похідною $\frac{dX(t)}{dt}$ у звичайному розумінні.

Якщо $X(t) \equiv X_0$, де X_0 – випадкова величина, то за означенням

$$(8) \text{ слід прийняти, що } Y(t) = \frac{dX(t)}{dt} = \frac{dX_0}{dt} \equiv 0.$$

Безпосередньо за означенням можна одержати ряд інших властивостей похідної у середньоквадратичному.

Покажемо, наприклад, що похідна суми двох випадкових функцій дорівнює сумі похідних доданків. Скористаємося нерівністю

$$M \left[\left(\left(\frac{X_1(t + \Delta t) - X_1(t)}{\Delta t} - X'_1(t) \right)^2 + \left(\frac{X_2(t + \Delta t) - X_2(t)}{\Delta t} - X'_2(t) \right)^2 \right) \right] \leq \\ \leq 2M \left[\left(\frac{X_1(t + \Delta t) - X_1(t)}{\Delta t} - X'_1(t) \right)^2 + \left(\frac{X_2(t + \Delta t) - X_2(t)}{\Delta t} - X'_2(t) \right)^2 \right],$$

з якої видно, що як тільки похідні $X'_1(t) = \frac{dX_1(t)}{dt}$ і $X'_2(t) = \frac{dX_2(t)}{dt}$

випадкових функцій $X_1(t)$ і $X_2(t)$ існують, то існує похідна їх суми $X(t) = X_1(t) + X_2(t)$, причому

$$\frac{d(X_1(t) + X_2(t))}{dt} = \frac{dX_1(t)}{dt} + \frac{dX_2(t)}{dt}. \quad (9)$$

Очевидно, що випадкову постійну (випадкову величину) можна виносити за знак похідної.

Зauważення 2. Нехай випадкова функція $X(t)$ має канонічний розклад

$$X(t) = m_x(t) + \sum_{k=1}^n \overset{\circ}{V}_k \phi_k(t), \quad (10)$$

$$M\left[\overset{\circ}{V}_k\right] = 0, k = 1, 2, \dots, n; \quad M\left[\overset{\circ}{V}_i \overset{\circ}{V}_j\right] = 0 \text{ при } i \neq j; \quad M\left[\overset{\circ}{V}_k \overset{\circ}{V}_k\right] = D_k.$$

У лекції 27 похідну цієї функції ми визначили як випадкову функцію

$$Y(t) = \frac{dX(t)}{dt} = \frac{dm_x(t)}{dt} + \sum_{k=1}^n \overset{\circ}{V}_k \frac{d\phi_k(t)}{dt}. \quad (11)$$

Якщо підставити (10), (11) у (8), то легко переконатись, що збіжність у середньоквадратичному має місце. Отже, вираз (11) повністю узгоджується з означенням похідної у середньоквадратичному. Разом з тим, вираз для похідної (11) можна одержати безпосередньо з (10), скориставшись відмінними вище властивостями похідної у середньоквадратичному.

Перейдемо до характеристик похідної як випадкової функції.
Нехай для випадкової функції $X(t)$ існує її середньоквадратична

похідна $Y(t) = \frac{dX(t)}{dt}$ з характеристиками $M[Y(t)] = m_y(t)$ і

$K_y \ t_1, t_2$. Тоді з очевидної нерівності

$$M \left[\left(\frac{X \ t + \Delta t - X \ t}{\Delta t} - Y \ t \right)^2 \right] \geq \left(M \left[\frac{X \ t + \Delta t - X \ t}{\Delta t} - Y \ t \right] \right)^2 = \\ = \left(\frac{m_x \ t + \Delta t - m_x \ t}{\Delta t} - M[Y \ t] \right)^2 \quad \text{випливає, що існує звичайна}$$

похідна $\frac{dm_x \ t}{dt}$ і

$$m_y \ t = M \left[\frac{dX \ t}{dt} \right] = \frac{d}{dt} M[X \ t] = \frac{dm_x \ t}{dt}. \quad (12)$$

Згідно з (9), (12), якщо $X \ t = \overset{\circ}{X} \ t + m_x \ t$, то

$$Y \ t = \frac{dX \ t}{dt} = \frac{d\overset{\circ}{X} \ t}{dt} + m_y \ t.$$

$$\text{Тоді } K_y \ t_1, t_2 = M \left[\frac{d\overset{\circ}{X} \ t_1}{dt} \cdot \frac{d\overset{\circ}{X} \ t_2}{dt} \right] = M \left[\frac{\partial^2 \left(\overset{\circ}{X} \ t_1 \overset{\circ}{X} \ t_2 \right)}{\partial t_1 \partial t_2} \right] = \\ = \frac{\partial}{\partial t_1} \frac{\partial}{\partial t_2} M \left[\overset{\circ}{X} \ t_1 \overset{\circ}{X} \ t_2 \right] = \frac{\partial}{\partial t_1} \frac{\partial}{\partial t_2} K_x \ t_1, t_2. \quad (13)$$

Тут ми скористалися тим, що згідно з (12) знаки математичного сподівання та похідної можна міняти місцями.

Отже, приходимо до висновку, що співвідношення (1) для похідної випадкової функції залишаються в силі.

Іноді потрібно знати взаємну кореляційну функцію випадкової функції $X \ t$ та її похідної $X' \ t = \frac{dX \ t}{dt}$. Знайдемо цю функцію:

$$K_{xx'} \ t_1, t_2 = M \left[\overset{\circ}{X} \ t_1 \frac{d\overset{\circ}{X} \ t_2}{dt} \right] = M \left[\frac{\partial \left(\overset{\circ}{X} \ t_1 \overset{\circ}{X} \ t_2 \right)}{\partial t_2} \right] = \\ = \frac{\partial}{\partial t_2} M \left[\overset{\circ}{X} \ t_1 \overset{\circ}{X} \ t_2 \right] = \frac{\partial}{\partial t_2} K_x \ t_1, t_2. \quad (14)$$

Одержануши співвідношення (12) і (13), ми тим самим показали, що як тільки для випадкової функції $X(t)$ на інтервалі a, b існує середньоквадратична похідна $\frac{dX}{dt}$, то існує похідна від математичного $\frac{dm_x}{dt}$ на a, b і друга мішана похідна кореляційної функції $\frac{\partial^2 K_x}{\partial t_1 \partial t_2} t_1, t_2$ для $t_1, t_2 \in a, b$. Справедливе обернене твердження. При цьому достатньо вимагати існування другої мішаної похідної кореляційної функції лише на діагоналі $t_1 = t_2 = t \in a, b$, тобто існування $\left. \frac{\partial^2 K_x}{\partial t_1 \partial t_2} t_1, t_2 \right|_{t_1=t, t_2=t}$. Доведення цього факту ми не наводимо.

Зauważення 3. При вивчені випадкових величин ми зустрічались з поняттям збіжності за ймовірністю. А саме, послідовність випадкових величин $X_1, X_2, \dots, X_n, \dots$ називається збіжною за ймовірністю до випадкової величини X_0 , якщо для довільного додатного числа $\varepsilon > 0$ виконується рівність

$$\lim_{n \rightarrow \infty} P |X_n - X_0| > \varepsilon = 0. \quad (15)$$

Збіжність у середньоквадратичному послідовності випадкових величин $X_1, X_2, \dots, X_n, \dots$ до випадкової величини X_0 означає:

$$\lim_{n \rightarrow \infty} M [X_n - X_0]^2 = 0. \quad (16)$$

Виникає природне питання: як співвідносяться між собою ці збіжності? Щоб відповісти на це питання, скористаємося нерівністю Чебишова у формі

$$P |X| > \varepsilon < \frac{M [X^2]}{\varepsilon^2}, \quad (17)$$

де X – довільна випадкова величина, а $\varepsilon > 0$ – довільне число.

Якщо покласти $X = X_n - X_0$, то нерівність (17) набуде вигляду

$$P |X_n - X_0| > \varepsilon < \frac{M [X_n - X_0]^2}{\varepsilon^2}. \quad (18)$$

Отже, зі збіжності в середньоквадратичному випливає збіжність за ймовірністю. Обернене твердження, взагалі кажучи, невірне.

Якщо в (17) замість X покласти $X - M[X] = X - m$, то отримаємо нерівність Чебишова у формі, якою ми користувалися раніше (див. лекцію 16).

Перейдемо до інтеграла від випадкової функції $X(t)$. Розіб'ємо відрізок a, b на n частин точками $s_0 = a, s_1, s_2, \dots, s_n = b$ і введемо інтегральні суми

$$S_n = \sum_{k=1}^n X(s_k) \Delta s_k, \quad (19)$$

де $\Delta s_k = s_k - s_{k-1}$, причому $\max_k |\Delta s_k| \rightarrow 0$ при $n \rightarrow \infty$. Очевидно, що S_n , $n = 1, 2, \dots$ є послідовністю випадкових величин.

Означення 4. Інтегралом (інтегралом у середньоквадратичному) від випадкової функції $X(t)$ в межах від a до b називається **випадкова величина**, що дорівнює границі у середньоквадратичному послідовності інтегральних сум (19):

$$\int_a^b X(t) dt = \lim_{n \rightarrow \infty} S_n. \quad (20)$$

якщо одна з меж a або b – змінна, то результатом інтегрування буде **випадкова функція**. Наприклад,

$$Z(t) = \int_a^t X(s) ds = \lim_{n \rightarrow \infty} S_n, \quad s_0 = a, \quad s_n = t. \quad (21)$$

Згідно з означенням збіжності у середньоквадратичному

$$\lim_{n \rightarrow \infty} M[S_n - Z(t)^2] = 0. \quad (22)$$

Безпосередньо за означенням можна показати, що для інтеграла у середньоквадратичному зберігається ряд властивостей звичайного інтеграла, а саме: інтеграл суми випадкових функцій дорівнює сумі інтегралів доданків; випадкову постійну можна виносити за знак інтеграла. Тому, якщо випадкова функція має канонічний розклад, то її інтегрування зводиться до інтегрування математичного сподівання та координатних функцій у звичайному розумінні.

Покажемо, що співвідношення (2) залишаються в силі. Оскільки

$$\begin{aligned} M[S_n - Z(t)^2] &\geq M[S_n - Z(t)]^2 = M[S_n] - M[Z(t)]^2 = \\ &= \left(M\left[\sum_{k=1}^n X(s_k) \Delta s_k \right] - m_z(t) \right)^2 = \left(\sum_{k=1}^n M[X(s_k)] \Delta s_k - m_z(t) \right)^2 = \\ &= \left(\sum_{k=1}^n m_x(s_k) \Delta s_k - m_z(t) \right)^2, \quad \text{то з інтегровності } X(t) \text{ у} \end{aligned}$$

середньоквадратичному випливає інтегровність математичного

сподівання $m_x(t)$, причому

$$m_z(t) = M \left[\int_a^t X(s) ds \right] = \int_a^t M[X(s)] ds = \int_a^t m_x(s) ds. \quad (23)$$

Тут ми скористалися тим, що знак суми і математичного сподівання можна міняти місцями і що $\sum_{k=1}^n m_x(s_k) \Delta s_k$ є інтегральною сумаю математичного сподівання.

Знайдемо кореляційну функцію $K_z(t_1, t_2) = M[\overset{\circ}{Z}(t_1) \overset{\circ}{Z}(t_2)]$.

Якщо $X(t) = \overset{\circ}{X}(t) + m_x(t)$, то

$$\overset{\circ}{Z}(t) = \int_a^t X(s) ds = \int_a^t \overset{\circ}{X}(s) ds + m_y(t).$$

Тому

$$\overset{\circ}{Z}(t_1) \overset{\circ}{Z}(t_2) = \int_a^{t_1} \overset{\circ}{X}(s_1) ds_1 \cdot \int_a^{t_2} \overset{\circ}{X}(s_2) ds_2 = \int_a^{t_1} \int_a^{t_2} \overset{\circ}{X}(s_1) \overset{\circ}{X}(s_2) ds_1 ds_2$$

і

$$\begin{aligned} K_z(t_1, t_2) &= M \left[\int_a^{t_1} \int_a^{t_2} \overset{\circ}{X}(s_1) \overset{\circ}{X}(s_2) ds_1 ds_2 \right] = \\ &= \int_a^{t_1} \int_a^{t_2} M[\overset{\circ}{X}(s_1) \overset{\circ}{X}(s_2)] ds_1 ds_2 = \int_a^{t_1} \int_a^{t_2} K_x(s_1, s_2) ds_1 ds_2. \end{aligned} \quad (24)$$

Тут ми скористалися тим, що згідно з (23) знаки інтеграла та математичного сподівання можна міняти місцями.

Рекомендується показати **самостійно**, що взаємна кореляційна функція випадкової функції $X(t)$ та її інтеграла $Z(t) = \int_a^t X(s) ds$ визначається формулою

$$K_{xz}(t_1, t_2) = \int_a^{t_2} K_x(t_1, s) ds. \quad (25)$$

Зауваження 4. Необхідно і достатньою умовами інтегровності у середньоквадратичному випадкової функції $X(t)$ на інтервалі a, b є інтегровність математичного сподівання $m_x(t)$ на даному інтервалі і інтегровність кореляційної функції $K_x(t_1, t_2)$ при $t_1, t_2 \in [a, b]$. Необхідну

умову нами фактично доведено при виведенні спiввiдношень (23), (24). Доведення достатньої умови ми проводити не будемо.

Зaуваження 5. Умову невiд'ємностi визначеностi кореляцiйної функцiї $K_x(t_1, t_2)$ (див. (1) з лекцiї 27) можна подати у виглядi

$$I = \int_a^b \int_a^b K_x(t_1, t_2) \varphi(t_1) \varphi(t_2) dt_1 dt_2 \geq 0. \quad (26)$$

Тут $\varphi(t)$ – довiльна невипадкова функцiя. Хоча функцiя $K_x(t_1, t_2)$ i може приймати вiд'ємнi значення для деяких значень аргументiв t_1 i t_2 , iнтеграл (26) niкoli не буде вiд'ємним.

Нерiвнiсть (26) можна отримати з нерiвностi (1) лекцiї 27, переходячи при $n \rightarrow \infty$ вiд подвiйної суми до подвiйного iнтеграла. Проте це легко зробити безпосередньо за означенням кореляцiйної функцiї. Diйсно,

$$\begin{aligned} I &= \int_a^b \int_a^b M \left[\overset{\circ}{X}(t_1) \overset{\circ}{X}(t_2) \right] \varphi(t_1) \varphi(t_2) dt_1 dt_2 = M \left[\int_a^b \overset{\circ}{X}(t_1) \varphi(t_1) dt_1 \cdot \int_a^b \overset{\circ}{X}(t_2) \varphi(t_2) dt_2 \right] = \\ &= M \left[\left(\int_a^b \overset{\circ}{X}(t) \varphi(t) dt \right)^2 \right] \geq 0. \end{aligned}$$

Tут mi скористались тим, що знаки математичного сподiвання та iнтеграла можна мiняти мiсцями. Вiдмiтимо, що межi iнтегрування a, b можуть бути нескiнченними.

Лекцiя № 29

Визначення характеристик випадкової функцiї з дослiду. Стaцiонарнi випадковi функцiї

Viще mi оперували ймовiрнiстiми характеристиками випадкових функцiй (математичними сподiваннями, дисперсiями, кореляцiйними функцiями) u припущеннi, що вони обчислюються за вiдомими одновимiрними та двовимiрними законами розподiлу випадкових функцiй згiдно з формулами (2), (3), (5) лекцiї 26. Aле на практицi цi закони розподiлу не завжди бувають вiдомими. U таких випадках замiсть невiдомих точних ймовiрнiстiх характеристик випадкової функцiї користуються iх oцiнками, знайденими за даними спостережень.

Нехай в результатi проведення n незалежних дослiдiв одержано n реалiзацiй випадкової функцiї $X(t)$. Розiб'ємо промiжок часу $[0, T]$, протягом якого проводяться спостереження, на rivni iнтервали точками $0 = t_1, t_2, \dots, t_m = T$. Tут i нижче вважаємо для простоти

аргумент t випадкової функції часом. Для кожного значення аргументу t_j , $1 \leq j \leq m$ знайдемо значення всіх n реалізацій випадкової функції. Дані дослідів подамо у вигляді таблиці

№ досліду	аргумент перерізу реалізація	t_1	t_2	...	t_j	...	t_m
1	$x_1 t$	$x_1 t_1$	$x_1 t_2$...	$x_1 t_j$...	$x_1 t_m$
2	$x_2 t$	$x_2 t_1$	$x_2 t_2$...	$x_2 t_j$...	$x_2 t_m$
\vdots	\vdots	\vdots	\vdots	...	\vdots	...	\vdots
n	$x_n t$	$x_n t_1$	$x_n t_2$...	$x_n t_j$...	$x_n t_m$

Тут $x_i t_j$ – значення i -ої реалізації випадкової функції $X t$ в момент часу t_j , $1 \leq i \leq n$, $1 \leq j \leq m$.

Оцінкою математичного сподівання перерізу $X t_j$ є середнє арифметичне

$$\tilde{m} t_j = \bar{x} t_j = \frac{1}{n} \sum_{i=1}^n x_i t_j . \quad (1)$$

За оцінку математичного сподівання $m_x t$ випадкової функції $X t$ приймається сукупність оцінок (1), тобто $\bar{x} t_1$, $\bar{x} t_2$, ..., $\bar{x} t_m$. Позначимо цю сукупність через $\bar{x} t$:

$$\bar{x} t = \bar{x} t_1, \bar{x} t_2, \dots, \bar{x} t_m . \quad (2)$$

Оцінкою дисперсії перерізу $X t_j$ є виправлена дисперсія

$$s^2 t_j = \frac{1}{n-1} \sum_{i=1}^n [x_i t_j - \bar{x} t_j]^2 . \quad (3)$$

За оцінку дисперсії $D_x t$ випадкової функції $X t$ приймається сукупність оцінок (3), яку позначимо через $s_x^2 t$. Отже,

$$s_x^2 t = s^2 t_1, s^2 t_2, \dots, s^2 t_m . \quad (4)$$

Оцінкою коваріації двох перерізів $X t_k$ і $X t_l$ є

$$\tilde{K} t_k, t_l = \frac{1}{n-1} \sum_{i=1}^n [x_i t_k - \bar{x} t_k] [x_i t_l - \bar{x} t_l] . \quad (5)$$

На відміну від вибіркової коваріації (див. лекцію 18), оцінка (5) є незміщеною.

За оцінку кореляційної функції $K_x(t_1, t_2)$ приймається сукупність оцінок (5), обчислених для всіх можливих пар аргументів t_k, t_l з m аргументів t_1, t_2, \dots, t_m . Враховуючи симетричність кореляційної функції $K_x(t_1, t_2) = K_x(t_2, t_1)$, будемо мати C_m^2 таких пар.

Знаючи оцінки (1), (3), (5) для всіх t_1, t_2, \dots, t_m , характеристики випадкової функції можна апроксимувати тими чи іншими аналітичними виразами.

Проте слід відмітити, що достатньо точні оцінки математичного сподівання, дисперсії та кореляційної функції можна отримати лише при великій кількості реалізацій випадкової функції і великій кількості точок t_1, t_2, \dots, t_m (розуміло, що число реалізацій n і число точок m залежать від характеру поведінки випадкової функції). Тому обчислення характеристик випадкової функції за результатами спостережень є досить трудомістким процесом. Для деяких класів випадкових функцій цей процес можна спростити.

Стаціонарні випадкові функції

Однією з основних ознак, за якою проводять класифікацію випадкових функцій, є залежність чи незалежність властивостей цих функцій від початку відліку часу. За цією ознакою розрізняють *стаціонарні* та *нестаціонарні* випадкові функції. Для стаціонарних випадкових функцій всі n -вимірні закони розподілу не змінюються при зсуві в часі на довільну величину τ . В межах *кореляційної теорії*, елементи якої ми розглядаємо, користуються лише одновимірними та двовимірними законами розподілу. Тому *стаціонарність* випадкової функції $X(t)$ в межах цієї теорії означає:

$$f_1(x_1; t_1) = f_1(x_1; t_1 + \tau), \quad f_2(x_1, x_2; t_1, t_2) = f_2(x_1, x_2; t_1 + \tau, t_2 + \tau). \quad (6)$$

Тут для прикладу розглядається випадкова функція з перерізами неперервного типу, так що f_1 та f_2 – одновимірна та двовимірна щільності розподілу.

Якщо в (6) покласти $\tau = -t_1$, то будемо мати

$$f_1(x_1; t_1) = f_1(x_1; 0), \quad f_2(x_1, x_2; t_1, t_2) = f_2(x_1, x_2; 0, t_2 - t_1). \quad (7)$$

Отже, для стаціонарної випадкової функції одновимірний закон розподілу не залежить від моменту часу t_1 , для якого він розглядається, а двовимірний закон розподілу залежить від різниці моментів часу і не залежить від початку відліку часу. Це означає (див. лекцію 26), що *математичне сподівання і дисперсія стаціонарної*

випадкової функції $X(t)$ стала, а кореляційна функція залежить від різниці моментів часу:

$$M[X(t)] = m_x(t) = m_x = \text{const}; \quad D[X(t)] = D_x(t) = D_x = \text{const}; \quad (8)$$

$$K_x(t_1, t_2) = K_x(t_2 - t_1) = K_x(\tau), \quad \tau = t_2 - t_1.$$

Випадкову функцію $X(t)$, для якої виконуються рівності (6), ще називають *строго стаціонарною* (або *стаціонарно у вузькому розумінні*). Якщо для випадкової функції виконуються співвідношення (8), то таку функцію називають *стаціонарно в широкому розумінні*. Оскільки співвідношення (8) є необхідними, але не є в загальному випадку достатніми для співвідношень (6), то наведені поняття стаціонарності, взагалі кажучи, не співпадають. Проте ці поняття ототожнюються у випадку, коли закони розподілу випадкової функції $X(t)$ є *нормальними* (у цьому випадку випадкову функцію називають *гаусівською*). Дійсно, для гаусівської випадкової функції щільності (6), як ми знаємо (див. лекції 10, 13), повністю визначаються характеристиками (8). Оскільки гаусівські випадкові функції досить часто зустрічаються при розв'язуванні практичних задач, будемо розглядати далі випадкові функції, стаціонарні в широкому розумінні, причому будемо називати їх просто стаціонарними.

Означення 1. *Стаціонарною* випадковою функцією назовемо таку випадкову функцію $X(t)$, математичне сподівання якої стало, а кореляційна функція залежить лише від різниці аргументів, тобто

$$K_x(t_1, t_2) = K_x(t_2 - t_1) = K_x(\tau), \quad \tau = t_2 - t_1.$$

В означенні опущена вимога сталості дисперсії, оскільки вона випливає з вигляду кореляційної функції:

$$D_x(t) = K_x(t, t) = K_x(t - t) = K_x(0) = D_x = \text{const}.$$

З наведеного означення і загальних властивостей кореляційної функції (див. лекцію 26) випливають такі *властивості кореляційної функції для стаціонарних випадкових функцій*:

- 1) $K_x(\tau) = K_x(-\tau)$. Дійсно, $K_x(\tau) = K_x(t_2 - t_1) = K_x(t_1 - t_2) = K_x(-\tau)$.
 - 2) $K_x(0) \geq 0$. Дійсно, $K_x(0) = D_x \geq 0$.
 - 3) $|K_x(\tau)| \leq K_x(0) = D_x$. Дійсно,
- $$|K_x(\tau)| = |K_x(t_2 - t_1)| \leq \sqrt{D_x(t_1) D_x(t_2)} = D_x = K_x(0).$$

Отже, графік кореляційної функції $K_x(\tau)$ симетричний відносно осі ординат і знаходиться в горизонтальній смузі $-D_x, D_x$.

Приклад 1. Випадкова гармоніка

$$\begin{aligned} X(t) &= U \cos \omega t + V \sin \omega t, \\ M[U] &= 0, \quad M[V] = 0, \quad M[U^2] = M[V^2] = D, \quad M[UV] = 0 \end{aligned} \quad (9)$$

є стаціонарним випадковим процесом.

Дійсно, оскільки $m_x(t) = 0$, а

$K_x(t_1, t_2) = D \cos \omega t_1 \cos \omega t_2 + D \sin \omega t_1 \sin \omega t_2 = D \cos \omega |t_2 - t_1| = D \cos \omega \tau$,
процес (9) є стаціонарним. Тут кореляційна функція знайдена нами як кореляційна функція канонічного розкладу (9).

Випадкову гармоніку (9) можна подати у вигляді

$$X(t) = A \sin \omega t + \varphi, \quad (9')$$

де $A = \sqrt{U^2 + V^2}$ – випадкова амплітуда, а $\varphi = \arctg \frac{V}{U}$ – випадкова фаза (рекомендується провести перетворення **самостійно**). З іншого боку, $U = A \sin \varphi$, а $V = A \cos \varphi$.

Приклад 2. Чи є випадкова гармоніка (9') з випадковою амплітудою A і фіксованою фазою $\varphi = \varphi_0 = \text{const}$ стаціонарним випадковим процесом?

Дати обґрунтовану відповідь **самостійно**.

У практичних застосуваннях досить часто досліджуються стаціонарні випадкові функції, які є результатом впливу великого числа незалежних випадкових факторів, серед яких немає домінуючих. Розподілі таких випадкових функцій близькі до нормального, а типові кореляційні функції мають вигляд

$$K_x(\tau) = D_x e^{-\alpha|\tau|}, \quad (10)$$

$$K_x(\tau) = D_x e^{-\alpha|\tau|} \cos \omega \tau. \quad (11)$$

Графіки цих функцій показано на рис. 1, 2.

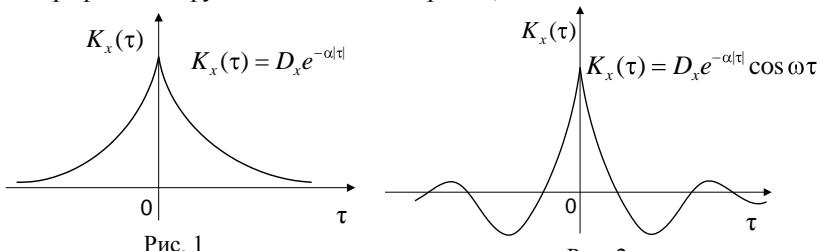


Рис. 1

Рис. 2

Більш загальною є кореляційна функція вигляду

$$K_x(\tau) = D_x e^{-\alpha|\tau|} \cos \omega \tau + \gamma \sin \omega |\tau|. \quad (12)$$

Функції (10) та (11) мають розривну похідну в початку координат. Для апроксимації кореляційних функцій з неперервною похідною можна скористатись виразом (12), підібравши в ньому коефіцієнт γ так, щоб похідна цієї кореляційної функції дорівнювала нулю при $\tau = 0$. Рекомендується показати **самостійно**, що потрібно взяти $\gamma = \frac{\alpha}{\omega}$.

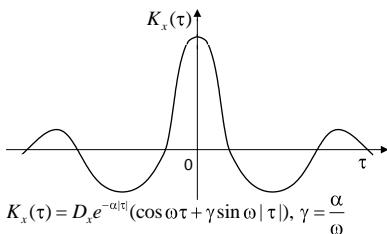


Рис. 3

На рис. 3 наведено графік такої кореляційної функції. Відмітимо, що вираз (12) може бути кореляційною функцією лише тоді, коли $|\gamma| \leq \frac{\alpha}{\omega}$, оскільки при $|\gamma| > \frac{\alpha}{\omega}$ не виконується властивість 3) кореляційної функції.

Як і в загальному випадку, для стаціонарної випадкової функції можна ввести нормовану кореляційну функцію, яка є функцією однієї змінної τ :

$$k_x(\tau) = \frac{K_x(\tau)}{D_x}. \quad (13)$$

Очевидно, що $k_x(\tau)$ – це коефіцієнт кореляції перерізів, розділених інтервалом τ . Як і в загальному випадку, $k_x(\tau) \leq 1$. Крім того, $k_x(0) = 1$.

Нехай $X(t)$ – диференційовна стаціонарна випадкова функція з похідною $Y(t) = \frac{dX(t)}{dt}$. Оскільки $m_y(t) = \frac{dm_x(t)}{dt}$, а $m_x(t) = \text{const}$, то $m_y(t) \equiv 0$. Знайдемо кореляційну функцію $K_y(t_1, t_2)$, врахувавши, що $K_x(t_1, t_2) = K_x(\tau)$; $\tau = t_2 - t_1$; $\frac{d\tau}{dt_1} = -1$; $\frac{d\tau}{dt_2} = 1$. Маємо:

$$K_y(t_1, t_2) = \frac{\partial}{\partial t_1} \frac{\partial}{\partial t_2} K_x(\tau) = \frac{\partial}{\partial t_1} \left[\frac{dK_x(\tau)}{d\tau} \cdot \frac{d\tau}{dt_2} \right] = \frac{d^2 K_x(\tau)}{d\tau^2} \frac{d\tau}{dt_1} = -\frac{d^2 K_x(\tau)}{d\tau^2}.$$

Отже, похідна $Y(t) = \frac{dX(t)}{dt}$ стаціонарної випадкової функції $X(t)$ також є стаціонарною випадковою функцією з характеристиками

$$m_y \equiv 0, \quad K_{yy}(\tau) = -\frac{d^2 K_x(\tau)}{d\tau^2}. \quad (14)$$

Зauważення 1. Необхідною і достатньою умовою диференційовності довільної випадкової функції $X(t)$ є існування похідної від математичного

сподівання і існування мішаної похідної $\frac{\partial^2 K_x(t_1, t_2)}{\partial t_1 \partial t_2}$ при $t_1 = t_2 = t$ (див.

лекцію 28). Якщо $X(t)$ – стаціонарна випадкова функція, то питання про

існування похідної $\frac{dm_x(t)}{dt}$ знімається, оскільки $m_x(t) \equiv const$. Тому

необхідною і достатньою умовою диференційовності стаціонарної випадкової функції є існування другої похідної $\frac{d^2 K_x(\tau)}{d\tau^2}$ при $\tau = t_2 - t_1 = 0$.

Але оскільки $\left. \frac{d^2 K_x(\tau)}{d\tau^2} \right|_{\tau=0} = -K_{yy}(0) = -D_y$, то диференційовність стаціонарної випадкової функції означає фактично скінченність дисперсії похідної.

Означення 2. Дві стаціонарні випадкові функції $X_1(t)$ і $X_2(t)$ називаються стаціонарно зв'язаними, якщо їх взаємна кореляційна функція залежить тільки від різниці аргументів:

$$K_{x_1 x_2}(t_1, t_2) = K_{x_1 x_2}(\tau), \quad \tau = t_2 - t_1 \quad (15)$$

Приклад 3. Випадкові функції з прикладу 3 лекції 26 є стаціонарними і стаціонарно зв'язаними.

Рекомендується показати **самостійно**, що взаємна кореляційна функція стаціонарної випадкової функції $X(t)$ та її похідної

$Y(t) = \frac{dX(t)}{dt}$ має вигляд

$$K_{xy}(t_1, t_2) = \frac{dK_x(\tau)}{d\tau}, \quad (16)$$

тобто $X(t)$ та $Y(t)$ зв'язані стаціонарно.

Зauważenie 2. Нехай $X(t)$ – стаціонарна випадкова функція з кореляційною функцією $K_x(\tau) = K_x(t_2 - t_1)$. Розглянемо інтеграл від цієї функції

$$Z(t) = \int_0^t X(s) ds.$$

Згідно з формулою (24) лекції 28 для дисперсії випадкової функції $Z(t)$ можна записати

$$D_z(t) = \int_0^t \int_0^t K_x(s_2 - s_1) ds_1 ds_2.$$

Перейдемо у подвійному інтегралі до нових змінних $\tau = s_2 - s_1$ і $\theta = s_2$. В результаті нескладних перетворень (рекомендується провести ці перетворення **самостійно**) вираз для дисперсії випадкової функції $Z(t)$ можна подати у вигляді

$$D_z(t) = 2 \int_0^t t - \tau K_x(\tau) d\tau. \quad (*)$$

Нехай, наприклад, $K_x(\tau) = \frac{1}{1 + \tau^2}$. Тоді згідно з $(*)$

$$D_z(t) = 2 \int_0^t t - \tau \frac{1}{1 + \tau^2} d\tau = 2t \arctg t - \ln(1 + t^2).$$

Отже, на відміну від похідної, інтеграл від стаціонарної випадкової функції, взагалі кажучи, не є стаціонарною випадковою функцією, оскільки для останньої дисперсія стала.

Приклад 4. Випадковий процес $X(t)$ має вигляд

$$X(t) = m_x + \sum_{k=1}^n U_k \cos \omega_k t + V_k \sin \omega_k t, \quad (17)$$

де $m_x = \text{const}$; $M[U_k] = M[V_k] = 0$; $M[U_i V_j] = 0$; $M[U_k^2] = M[V_k^2] = D_k$ для всіх $i, j, k = 1, 2, \dots, n$; $M[U_i U_j] = M[V_i V_j] = 0$ при $i \neq j$.

Показати, що цей процес є стаціонарним і знайти його дисперсію.

Розв'язання. За наведених умов вираз (17) є канонічним розкладом випадкової функції $X(t)$. Тому для кореляційної функції можна записати (див. лекцію 27):

$$K_x(t_1, t_2) = \sum_{k=1}^n D_k \cos \omega_k t_1 \cos \omega_k t_2 + D_k \sin \omega_k t_1 \sin \omega_k t_2 = \sum_{k=1}^n D_k \cos \omega_k(t_2 - t_1).$$

Отже, кореляційна функція випадкової функції (17) має вигляд

$$K_x(\tau) = \sum_{k=1}^n D_k \cos \omega_k \tau, \quad (18)$$

а тому ця випадкова функція є стаціонарною.

Поклавши в (18) $\tau = 0$, отримуємо вираз для дисперсії

$$D_x = \sum_{k=1}^n D_k. \quad (19)$$

Ввівши випадкові амплітуди $A_k = \sqrt{U_k^2 + V_k^2}$ і випадкові фази $\varphi_k = \arctg U_k/V_k$ (див. приклад 1), випадкову функцію (17) можна переписати у вигляді

$$X(t) = m_x + \sum_{k=1}^n A_k \sin \omega_k t + \varphi_k. \quad (20)$$

Отже, якщо випадкову функцію $X(t)$ можна подати у вигляді суми гармонік різних частот з випадковими амплітудами і випадковими фазами, то $X(t)$ – стаціонарна випадкова функція. Звичайно, при цьому повинні виконуватись умови, що накладаються в (17) на випадкові величини U_k, V_k . Рекомендується показати **самостійно**, що як тільки ці умови не виконуються, випадкова функція $X(t)$ з (17) стаціонарною не буде.

Лекція № 30

Ергодичність стаціонарних випадкових функцій. Комплексні випадкові функції та їх характеристики

Стаціонарні випадкові процеси досить часто зустрічаються на практиці, особливо у фізиці і техніці. Такі процеси відбуваються в часі приблизно однорідно і мають вигляд неперервних випадкових коливань навколо деякого середнього значення. Наприклад, коливання напруги в електричній мережі, випадкові шуми в приладах радіоелектроніки тощо. Як правило, процеси, які відбуваються в системах, що пропрацювали достатньо тривалий час, описуються стаціонарними випадковими функціями.

Нехай в результаті проведення дослідів одержано декілька реалізацій випадкової функції $X(t)$ і знайдені оцінки математичного сподівання, дисперсії та кореляційної функції описані на початку

минулої лекції способом. Використовуючи теорію статистичної перевірки гіпотез, можна вияснити, випадкові чи ні розбіжності в оцінках математичних сподівань і оцінках дисперсій перерізів X t_j випадкової функції, а також розбіжності в оцінках коваріації при фіксованих різницях моментів часу $t_k - t_l = const$ (див. оцінки (1), (3), (5) з лекції 29). Якщо ці розбіжності носять випадковий характер, то математичне сподівання і дисперсію випадкової функції можна вважати постійними величинами, а кореляційну функцію такою, що залежить лише від різниці аргументів $t_k - t_l$, тобто вважати, що випадкова функція є стаціонарною.

Проте при розв'язуванні практичних задач досить часто доводиться обмежуватись невеликим числом реалізацій випадкової функції. Це може привести до того, що теорія статистичної перевірки гіпотез, яка передбачає достатньо велике число спостережень, не підтверджує гіпотезу про стаціонарність випадкової функції, хоча природа явища дозволяє стверджувати, що воно повинно описуватись стаціонарною випадковою функцією. У цьому випадку, вважаючи випадкову функцію стаціонарною, оцінку математичного сподівання можна знайти, усереднивши оцінки для математичних сподівань перерізів (див. оцінку (1) з лекції 29):

$$m_x = \bar{x} = \frac{\bar{x}(t_1) + \bar{x}(t_2) + \dots + \bar{x}(t_m)}{m}. \quad (1)$$

Аналогічно, одержуємо оцінку для дисперсії (див. (3) з лекції 29):

$$D_x = s_x^2 = \frac{s^2 t_1 + s^2 t_2 + \dots + s^2 t_m}{m}. \quad (2)$$

Усереднюючи оцінки (див. (5) з лекції 29) коваріації для одинакових різниць $t_k - t_l = const$ моментів часу, отримаємо оцінку значень кореляційної функції $K_x(\tau)$ на дискретній множині значень аргумента τ . Очевидно, що для великих τ ці оцінки стають ненадійними, оскільки отримуються усередненням невеликої кількості даних. Далі емпіричну кореляційну функцію можна апроксимувати тим чи іншим аналітичним виразом. Зрозуміло, що цей вираз повинен відображати найбільш характерні риси емпіричної кореляційної функції і згладжувати випадкові коливання при великих τ .

Розглянутий метод експериментального визначення характеристик стаціонарної випадкової функції є досить громіздким і включає два етапи: оцінку характеристик перерізів випадкової функції за реалізаціями і усереднення цих оцінок. Проте серед стаціонарних

випадкових функцій можна виділити клас функцій, для яких процедуру експериментальної оцінки характеристик можна значно спростити.

Нехай $X(t)$ – стаціонарна випадкова функція з математичним сподіванням m_x і кореляційною функцією $K_x(\tau)$, $\tau = t_2 - t_1$. Розглянемо випадкову функцію

$$Z(t) = \frac{1}{t} \int_0^t X(s) ds, \quad t > 0. \quad (3)$$

Очевидно, що математичне сподівання цієї випадкової функції

$$M[Z(t)] = m_z = m_x. \quad (4)$$

Для дисперсії $D_z(t)$ випадкової функції $Z(t)$ згідно з формулою (*) лекції 29 можна записати

$$D_z(t) = M[(Z(t) - m_x)^2] = \frac{2}{t^2} \int_0^t t - \tau |K_x(\tau)| d\tau. \quad (5)$$

Будемо розглядати стаціонарні випадкові функції, неперервні у середньоквадратичному, тобто такі, кореляційна функція $K_x(\tau)$ яких є неперервною функцією. Поставимо таку **задачу**: знайти умови, за яких $\lim_{t \rightarrow \infty} D_z(t) = 0$. Для цього введемо функцію

$$F(t) = 2 \int_0^t |t - \tau| |K_x(\tau)| d\tau \quad (6)$$

і запишемо очевидну нерівність

$$D_z(t) \leq \frac{F(t)}{t^2}. \quad (7)$$

Оскільки похідна

$$F'(t) = 2 \int_0^t |K_x(\tau)| d\tau > 0, \quad (8)$$

то функція $F(t)$ монотонно зростає при $t \rightarrow \infty$. Крім того, похідна $F'(t)$ сама є монотонно зростаючою функцією при $t \rightarrow \infty$.

Розглянемо випадок, коли

$$\lim_{t \rightarrow \infty} F(t) = \infty, \quad \lim_{t \rightarrow \infty} F'(t) = \infty \quad (9)$$

і знайдемо для цього випадку границю $\lim_{t \rightarrow \infty} \frac{F(t)}{t^2}$, скориставшись правилом Лопітала:

$$\lim_{t \rightarrow \infty} \frac{F(t)}{t^2} = \lim_{t \rightarrow \infty} \frac{F'(t)}{2t} = \lim_{t \rightarrow \infty} \frac{F''(t)}{2} = \lim_{t \rightarrow \infty} |K_x(t)|. \quad (10)$$

З нерівності (7) і співвідношень (10) випливає, що як тільки

$$\lim_{t \rightarrow \infty} K_x(t) = 0, \text{ то } \lim_{t \rightarrow \infty} D_z(t) = 0.$$

У випадку, коли кореляційна функція $K_x(\tau)$ прямує до нуля при $\tau \rightarrow \infty$ досить швидко, так що інтегали (6) і (8) збігаються, прямування до нуля дисперсії $D_z(t)$ при $t \rightarrow \infty$ є очевидним. Якщо збігається лише інтеграл (8), рівність $\lim_{t \rightarrow \infty} D_z(t) = 0$ випливає з нерівності (7) і першої рівності (10).

Отже, ми показали, що достатньою умовою необмеженого прямування до нуля дисперсії $D_z(t)$ при $t \rightarrow \infty$ є необмежене прямування до нуля кореляційної функції $K_x(\tau)$ при $|\tau| \rightarrow \infty$ (тут враховано, що функція $K_x(\tau)$ є парною). Це означає, що за умови $\lim_{|\tau| \rightarrow \infty} K_x(\tau) = 0$ випадкова функція $Z(t)$ з (3) зі збільшенням t втрачає свою випадковість і для достатньо великого T можна записати наблизену рівність

$$m_x \approx m_x^* = \frac{1}{T} \int_0^T x(t) dt, \quad (11)$$

де $x(t)$ – реалізація стаціонарної випадкової функції $X(t)$. Величина m_x^* є середнім значенням реалізації $x(t)$ на інтервалі часу $[0, T]$. Розсіювання результатів обчислень за формулою (11) для різних реалізацій прямує до нуля при необмеженому збільшенні довжини реалізації T .

Стаціонарні випадкові функції $X(t)$, для яких оцінку математичного сподівання можна знайти шляхом усереднення в часі лише однієї, але достатньо тривалої реалізації $x(t)$, будемо називати **ергодичними відносно математичного сподівання**.

Отже, для класу ергодичних відносно математичного сподівання стаціонарних випадкових функцій усереднення по множині реалізацій

при знаходженні математичного сподівання можна замінити простим усередненням в часі лише однієї реалізації.

Кореляційну функцію $K_x \tau$ стаціонарної випадкової функції $X(t)$ для довільного $\tau = \tau_0$ можна записати у вигляді

$$K_x \tau_0 = M \left[\begin{array}{cc} \dot{\bar{X}}(t) & \dot{\bar{X}}(t + \tau_0) \end{array} \right], \quad (12)$$

де $\dot{\bar{X}}(t) = \bar{X}(t) - m_x$, $\dot{\bar{X}}(t + \tau_0) = \bar{X}(t + \tau_0) - m_x$. Тому знайти кореляційну функцію $K_x \tau$ стаціонарної випадкової функції $X(t)$ означає знайти математичне сподівання випадкової функції $Y(t) = \dot{\bar{X}}(t) \dot{\bar{X}}(t + \tau_0)$ для всіх τ_0 . При певних умовах це можна зробити, усереднюючи в часі одну з реалізацій

$$\bar{y}(t) = \dot{\bar{x}}(t) \dot{\bar{x}}(t + \tau_0) = \bar{x}(t) - m_x \quad \bar{x}(t + \tau_0) - m_x \quad (13)$$

випадкової функції $Y(t)$.

Скористаємось тим, що у випадку стаціонарної гаусівської функції $X(t)$ (див. лекцію 29) випадкова функція $Y(t)$ також є стаціонарною і має кореляційну функцію

$$K_y \tau = K_x^2 \tau + K_x \tau + \tau_0 K_x \tau - \tau_0. \quad (14)$$

Доведення цього факту можна знайти, наприклад, в книзі В.С. Пугачєва “Введение в теорию вероятностей”, 1968. З (14) бачимо, що умова $\lim_{|\tau| \rightarrow \infty} K_x \tau = 0$ є достатньою для того, щоб $\lim_{|\tau| \rightarrow \infty} K_y \tau = 0$.

Тому за цієї умови можна скористатись наведеними вище результатами і записати з врахуванням (12) наближену рівність (11) для реалізації (13) випадкової функції $Y(t)$

$$K_x \tau_0 \approx K_x^* \tau_0 = \frac{1}{T - \tau_0} \int_0^{T - \tau_0} \bar{x}(t) - m_x \quad \bar{x}(t + \tau_0) - m_x dt. \quad (15)$$

Тут $K_x^* \tau_0$ – середнє значення добутку $\bar{x}(t) - m_x \quad \bar{x}(t + \tau_0) - m_x$ на інтервалі $[0, T - \tau_0]$, оскільки функції $\bar{x}(t) - m_x$ і $\bar{x}(t + \tau_0) - m_x$ відомі разом тільки в цьому інтервалі. Формула (15) дає оцінку кореляційної функції для $\tau_0 > 0$. Для $\tau_0 < 0$ оцінка кореляційної функції визначається з умови її парності.

Стаціонарні випадкові функції $X(t)$, для яких оцінку кореляційної функції можна знайти за однією достатньо тривалою реалізацією $x(t)$ з використанням наближеної рівності (15), назовемо **ергодичними відносно кореляційної функції**.

Отже, умова $\lim_{|\tau| \rightarrow \infty} K_x(\tau) = 0$ є достатньою для ергодичності довільної стаціонарної випадкової функції $X(t)$ відносно математичного сподівання. Якщо, крім того, $X(t)$ – гаусівська випадкова функція, то ця умова є достатньою і для ергодичності відносно кореляційної функції.

Приклад 1. Нормально розподілені стаціонарні випадкові функції з розглянутими у попередній лекції трьома типовими кореляційними функціями

$$\begin{aligned} K_x(\tau) &= D_x e^{-\alpha|\tau|}; \quad K_x(\tau) = D_x e^{-\alpha|\tau|} \cos \omega \tau; \\ K_x(\tau) &= D_x e^{-\alpha|\tau|} \cos \omega \tau + \gamma \sin \omega |\tau| \end{aligned} \quad (16)$$

ергодичні як відносно математичного сподівання, так і відносно кореляційної функції, оскільки всі ці кореляційні функції прямають до нуля при $|\tau| \rightarrow \infty$. Тому їх математичні сподівання і кореляційні функції можна визначити за однією достатньо тривалою реалізацією, користуючись формулами (11) і (15). Відмітимо, що формулою (15) рекомендується користуватись для $\tau_0 < \frac{T}{5}$. У протилежному разі похибка оцінки кореляційної функції за формулою (15) різко зростає.

Приклад 2. Випадкова гармоніка

$$\begin{aligned} X(t) &= U \cos \omega t + V \sin \omega t, \\ M(U) &= 0, \quad M(V) = 0, \quad M[U^2] = M[V^2] = D, \quad M(UV) = 0 \end{aligned} \quad (17)$$

ергодична відносно математичного сподівання, але не є ергодичною відносно кореляційної функції.

Дійсно, для середнього значення реалізації $x(t)$ на скінченному проміжку маємо:

$$\frac{1}{T} \int_0^T x(t) dt = \frac{1}{T} \int_0^T u \cos \omega t + v \sin \omega t dt = \frac{1}{T\omega} [u \sin \omega T - v \cos \omega T + v].$$

Оскільки $\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt = 0 = m_x$, то процес (17) є ергодичним

відносно математичного сподівання (відмітимо, що достатня умова цієї ергодичності не виконується. Дійсно, $\lim_{\tau \rightarrow \infty} K_x(\tau) = \lim_{\tau \rightarrow \infty} D \cos \omega \tau \neq 0$ (границя не існує)).

Подамо процес (17) у вигляді $X(t) = A \sin \omega t + \varphi$, де $A = \sqrt{U^2 + V^2}$, а $\varphi = \operatorname{arctg} U/V$ (див. приклад 1 лекції 29). Нехай $x(t) = a_1 \sin \omega t + \varphi_1$ – довільна реалізація цього процесу. Тоді

$$\begin{aligned} \lim_{T \rightarrow \infty} \left(\frac{1}{T - \tau_0} \int_0^{T - \tau_0} x(t) x(t + \tau_0) dt \right) &= \lim_{T \rightarrow \infty} \left(\frac{a_1^2}{T - \tau_0} \int_0^{T - \tau_0} \sin \omega t + \varphi_1 \sin \omega(t + \tau_0) + \varphi_1 dt \right) = \\ &= \frac{a_1^2}{2} \cos \omega \tau_0 \quad (\text{обчислення інтеграла і перехід до границі рекомендується провести} \text{ самостійно}). \end{aligned}$$

Отже, при використанні формули (15) різним реалізаціям $x(t)$ (a_1^2 – різні) відповідатимуть різні значення функції $K_x^*(\tau_0)$. Тому процес (17) не є ергодичним відносно кореляційної функції.

Зauważення 1. З теорії стаціонарних випадкових функцій відомо, що будь-якій кореляційній функції $K_x(\tau)$ завжди можуть відповідати як ергодичні, так і неергодичні по відношенню до кореляційної функції стаціонарні випадкові функції. Тому, знаючи лише кореляційну функцію стаціонарної випадкової функції, не можна в загальному випадку визначити, є вона ергодичною відносно кореляційної функції чи ні. Для відповіді на це питання необхідно мати додаткову інформацію про статистичні характеристики цієї стаціонарної випадкової функції, наприклад, знати, що випадкова функція розподілена нормальню. Для нормально розподілених стаціонарних випадкових функцій, як ми вже відмічали, умова $\lim_{|\tau| \rightarrow \infty} K_x(\tau) = 0$ є достатньою для їх ергодичності відносно кореляційної функції.

У результаті досліду значення реалізації $x(t)$ становить відомими для скінченної множини значень аргумента t . Тому при використанні формул (11) і (15) інтеграли обчислюються наблизено, для чого замінюються скінченими сумами. Розіб'ємо інтервал $[0, T]$ на n рівних частин довжиною $\Delta t = \frac{T}{n}$. Позначимо середини отриманих

частинних інтервалів через t_1, t_2, \dots, t_n . Тоді для оцінки (11) математичного сподівання можна записати

$$m_x \approx \frac{1}{T} \sum_{i=1}^n x(t_i) \Delta t = \frac{1}{n} \sum_{i=1}^n x(t_i) = \bar{x}. \quad (18)$$

Аналогічно можна обчислити кореляційну функцію для $\tau_0 = \Delta t, 2\Delta t, \dots$. Нехай, наприклад, $\tau_0 = m\Delta t = \frac{mT}{n}$. Якщо в (15),

розділити інтервал інтегрування $0, T - \tau_0$, де $T - \tau_0 = T - \frac{mT}{n} = \frac{n-m}{n}T$, на $n-m$ рівних частин довжиною Δt і замінити m_x на \bar{x} , то будемо мати

$$\begin{aligned} K_x(m\Delta t) &\approx \frac{1}{\frac{n-m}{n}T} \sum_{i=1}^{n-m} [x(t_i) - \bar{x}] [x(t_{i+m}) - \bar{x}] \frac{T}{n} = \\ &= \frac{1}{n-m} \sum_{i=1}^{n-m} [x(t_i) - \bar{x}] [x(t_{i+m}) - \bar{x}]. \end{aligned} \quad (19)$$

Для оцінки дисперсії $D_x = K_x(0)$ покладемо $m=0$:

$$D_x \approx \frac{1}{n} \sum_{i=1}^n [x(t_i) - \bar{x}]^2. \quad (20)$$

Комплексні випадкові функції та їх характеристики

Для дослідження стаціонарних випадкових функцій широко використовуються поняття комплексної випадкової функції та її характеристик. Це пов'язано не стільки з намаганням узагальнити поняття дійсної випадкової функції, а, в першу чергу, з самою структурою стаціонарної випадкової функції, яку математично зручніше описувати в комплексній формі. Ситуація тут та ж сама, що й у звичайному аналізі періодичних функцій. Розклад цих функцій в ряд Фур'є, наприклад, краще виглядає в комплексній формі. Крім того, математичні операції значно простіше виконувати з показниковими функціями, ніж з тригонометричними.

Введемо основні поняття.

Означення 1. Комплексною випадковою функцією називається функція

$$Z(t) = X(t) + iY(t), \quad (21)$$

де $X(t)$, $Y(t)$ – дійсні випадкові функції дійсного аргументу t , а $i = \sqrt{-1}$ – уявна одиниця.

Означення 2. Комплексна випадкова функція

$$\bar{Z}(t) = X(t) - iY(t) \quad (22)$$

називається **спряженою** з комплексною випадковою функцією (21).

Поняття математичного сподівання та дисперсії комплексної випадкової функції вводяться таким чином, щоб при $Y(t) \equiv 0$ ці характеристики співпадали з характеристиками для випадкової функції $X(t)$, тобто виконувались рівності $m_z(t) = m_x(t)$, $D_z(t) = D_x(t)$.

Означення 3. Математичним сподіванням комплексної випадкової функції (21) називається комплексна (невипадкова) функція

$$m_z(t) = m_x(t) + im_y(t). \quad (23)$$

Якщо $Y(t) \equiv 0$, одержуємо $m_z(t) = m_x(t)$.

Введемо центровану комплексну випадкову функцію $\hat{Z}(t)$:

$$\begin{aligned} \hat{Z}(t) &= Z(t) - m_z(t) = X(t) + iY(t) - m_x(t) - im_y(t) = \\ &= X(t) - m_x(t) + i(Y(t) - m_y(t)) = \hat{X}(t) + i\hat{Y}(t). \end{aligned}$$

Означення 4. Дисперсією комплексної випадкової функції (21) називається математичне сподівання квадрата модуля центрованої випадкової функції $\hat{Z}(t)$:

$$D_z(t) = M\left[\left|\hat{Z}(t)\right|^2\right] = M\left[\left(\hat{X}(t)\right)^2 + \left(\hat{Y}(t)\right)^2\right] = D_x(t) + D_y(t). \quad (24)$$

Отже, при такому означенні **дисперсія комплексної випадкової функції дорівнює сумі дисперсій її дійсної та уявної частин**. Зокрема, при $Y(t) \equiv 0$ одержуємо $D_z(t) = D_x(t)$.

Кореляційна функція $K_z(t_1, t_2)$ комплексної випадкової функції (21) вводиться таким чином, щоб при $t_1 = t_2 = t$ вона дорівнювала дисперсії $D_z(t)$, тобто, щоб зберіглась вже відома нам властивість кореляційної функції для дійсної випадкової функції.

Означення 5. Кореляційною функцією комплексної випадкової функції (21) називається математичне сподівання добутку $\hat{Z}(t_1)\hat{Z}(t_2)$,

тобто

$$K_z(t_1, t_2) = M \begin{bmatrix} \overset{\circ}{Z} & t_1 & \overset{\circ}{\bar{Z}} & t_2 \end{bmatrix}. \quad (25)$$

$$\text{Якщо } t_1 = t_2 = t, \text{ то } K_z(t, t) = M \begin{bmatrix} \overset{\circ}{Z} & t & \overset{\circ}{\bar{Z}} & t \end{bmatrix} = M \left[\left| \overset{\circ}{Z} t \right|^2 \right] = D_z(t).$$

З означення 5 випливає, що

$$K_z(t_1, t_2) = K_x(t_1, t_2) + K_y(t_1, t_2) + i K_{xy}(t_2, t_1) - K_{xy}(t_1, t_2), \quad (26)$$

зокрема для некорельзованих $X(t)$ і $Y(t)$

$$K_z(t_1, t_2) = K_x(t_1, t_2) + K_y(t_1, t_2). \quad (27)$$

Рекомендується довести рівність (26) **самостійно**. Зокрема, одержати рівність

$$K_z(t_1, t_2) = \overline{K_z(t_2, t_1)}. \quad (28)$$

Нехай $Z_1(t) = X_1(t) + iY_1(t)$ і $Z_2(t) = X_2(t) + iY_2(t)$ – дві комплексні випадкові функції.

Означення 6. Взаємною кореляційною функцією для $Z_1(t)$ і $Z_2(t)$ назовемо невипадкову функцію

$$K_{z_1 z_2}(t_1, t_2) = M \begin{bmatrix} \overset{\circ}{Z_1} & t_1 & \overset{\circ}{\bar{Z}_2} & t_2 \end{bmatrix}. \quad (29)$$

Комплексні випадкові функції $Z_1(t)$, $Z_2(t)$ назовемо **некорельзованими**, якщо $K_{z_1 z_2}(t_1, t_2) = 0$.

Якщо $Y_1(t) \equiv 0$ і $Y_2(t) \equiv 0$, то з (29) одержуємо

$$K_{z_1 z_2}(t_1, t_2) = M \begin{bmatrix} \overset{\circ}{X_1} & t_1 & \overset{\circ}{X_2} & t_2 \end{bmatrix} = K_{x_1 x_2}(t_1, t_2). \quad (30)$$

Рекомендується виразити взаємну кореляційну функцію двох комплексних випадкових функцій через взаємні кореляційні функції їх дійсних та уявних частин **самостійно**.

Зауваження 2. Якщо в наведених вище означеннях 1-4 випадкові функції $X(t)$ та $Y(t)$ вважати випадковими постійними (випадковими величинами), то прийдемо до понять комплексної випадкової величини та її характеристик: математичного сподівання та дисперсії. Аналогічно, з означення 6 приходимо до поняття коваріації двох комплексних випадкових величин.

Зауваження 3. Властивість **невід'ємної визначеності** кореляційної функції формулюється так: для довільних комплексних чисел c_j і довільних t_j з області визначення кореляційної функції $K_z(t_1, t_2)$ сума $\sum_i \sum_j K_z(t_i, t_j) c_i \bar{c}_j$ дійсна і невід'ємна. Рекомендується довести цю властивість **самостійно**.

Поширимо поняття **стационарності** на комплексні випадкові функції.

Означення 7. Комплексна випадкова функція $Z(t)$ називається **стационарною**, якщо її математичне сподівання стало, а кореляційна функція залежить лише від різниці аргументів:

$$m_z(t) = m_x + i m_y = \text{const}, \\ K_z(t_1, t_2) = M \begin{bmatrix} \overset{\circ}{Z}(t_1) & \overset{\circ}{Z}(t_2) \end{bmatrix} = K_z(t_1 - t_2) = K_z(\tau), \quad \tau = t_1 - t_2. \quad (31)$$

З (28) випливає, що

$$K_z(\tau) = \overline{K_z(-\tau)}. \quad (32)$$

Властивість невід'ємної визначеності кореляційної функції приймає вигляд

$$\sum_i \sum_j K_z(t_i - t_j) c_i \bar{c}_j \geq 0. \quad (33)$$

Лекція № 31

Стационарні випадкові функції з дискретним та неперервним спектрами. Спектральна щільність

Розглянемо клас випадкових функцій, які можна утворити з випадкових гармонік різних частот з випадковими амплітудами і випадковими фазами за допомогою формул (див. приклад 4 з лекції 29)

$$X(t) = m_x + \sum_{k=1}^n U_k \cos \omega_k t + V_k \sin \omega_k t, \quad (1)$$

де $\omega_1, \omega_2, \dots, \omega_n$ – довільні частоти; m_x – стала (математичне сподівання); $U_1, U_2, \dots, U_n, V_1, V_2, \dots, V_n$ – випадкові величини з нульовими математичними сподіваннями. Якщо, зокрема, ці випадкові величини не корельовані і мають попарно рівні дисперсії

$D U_k = D V_k = D_k$, то випадкова функція (1) є стаціонарною і має кореляційну функцію (див. приклад 4 з лекції 29)

$$K_x \tau = \sum_{k=1}^n D_k \cos \omega_k \tau. \quad (2)$$

Формули (1) та (2) називаються *спектральними розкладами* відповідно стаціонарної випадкової функції $X(t)$ та її кореляційної функції $K_x \tau$. Дисперсія випадкової функції (1) є сума дисперсій окремих гармонік на частотах $\omega_1, \omega_2, \dots, \omega_n$:

$$D_x = K_x 0 = \sum_{k=1}^n D_k \quad D_k > 0. \quad (3)$$

Означення 1. Сукупність всіх частот $\omega_1, \omega_2, \dots, \omega_n$ і відповідних їм дисперсій D_1, D_2, \dots, D_n назовемо *дискретним спектром* стаціонарної випадкової функції $X(t)$, а саму функцію $X(t)$ – стаціонарною випадковою функцією з дискретним спектром.

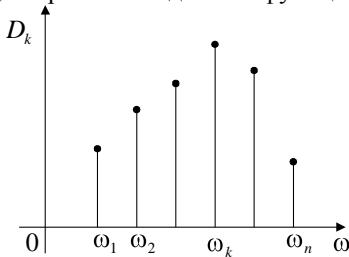


Рис. 1

Дискретний спектр можна зобразити графічно (рис. 1). Ординати D_k називають *спектральними лініями*.

Формули (1), (2), (3) можуть розглядатись для довільного n , зокрема, для $n = \infty$. Проте, якщо ми

хочемо отримати стаціонарну випадкову функцію зі скінченою дисперсією, то випадкові величини U_k та V_k повинні бути такими, щоб ряд (3) для $n = \infty$ збігався. Зрозуміло, що тоді буде збігатися і ряд (2).

До цього часу ми не накладали ніяких обмежень на частоти ω_k в формулі (1). Ці частоти могли бути довільними. Розглянемо тепер частинний випадок нескінченної кількості рівновіддалених одна від одної частот

$$\omega_k = \frac{k\pi}{T}, \quad k = 1, 2, \dots, \quad \Delta\omega = \omega_{k+1} - \omega_k = \frac{\pi}{T}, \quad (4)$$

де T – довільне додатне число. У цьому випадку функція (2) є періодичною з періодом $2T$, а ряд (2) – рядом Фур'є цієї функції по косинусах. Нагадаємо, що функція $K_x \tau$ – парна. Коєфіцієнти D_k

згідно з теорією рядів Фур'є визначаються формулою

$$D_k = \frac{2}{T} \int_0^T K_x(\tau) \cos \omega_k \tau d\tau \quad k = 1, 2, \dots . \quad (5)$$

Якщо кожній частоті $\omega_k = \frac{k\pi}{T}$, $k = 1, 2, \dots$ поставити у відповідність дисперсію D_k , то отримаємо дискретний спектр з нескінченною (численною) кількістю спектральних ліній (ординат D_k). Сусідні спектральні лінії (рис. 1) віддалені одна від одної на одну і ту ж відстань $\Delta\omega = \frac{\pi}{T}$.

Якщо кореляційна функція $K_x(\tau)$ не є періодичною, то її не можна на всій осі $-\infty < \tau < \infty$ подати у вигляді (2), а саму випадкову функцію $X(t)$ на всій осі $-\infty < t < \infty$ – у вигляді (1). У цьому випадку $X(t)$ не є випадковою функцією з дискретним спектром. Проте у ряд Фур'є з коефіцієнтами (5) можна розвинути парну неперіодичну функцію $K_x(\tau)$ за умови, що вона розглядається лише на скінченному інтервалі $-T, T$ (у цьому випадку вважають функцію $K_x(\tau)$ періодично продовженою з інтервалу $-T, T$ на всю вісь). Це означає, що й стаціонарну випадкову функцію $X(t)$, для якої така функція $K_x(\tau)$ є кореляційною, можна розвинути на “скінченному інтервалі” $t \in -\theta, \theta$ (оскільки $\tau = t_2 - t_1$, то очевидно, що $\theta = T/2$) у ряд зчисленного числа випадкових гармонік.

Хоча з практичної точки зору можливість таких розвинень у ряди на скінчених інтервалах мало що дає, нею можна скористатись для граничного переходу у наведених вище співвідношеннях при $T \rightarrow \infty$, оскільки ці співвідношення справедливі для довільних T . Нижче будемо цікавитись лише кореляційною функцією $K_x(\tau)$ стаціонарної випадкової функції $X(t)$, а не самою функцією $X(t)$.

Перепишемо спочатку вираз (2) для кореляційної функції, підставивши замість D_k вираз (5):

$$\begin{aligned}
K_x \tau &= \sum_{k=1}^{\infty} \left[\frac{2}{T} \int_0^T K_x(t) \cos \omega_k t dt \right] \cos \omega_k \tau = \\
&= \sum_{k=1}^{\infty} \left[\frac{2}{\pi} \int_0^T K_x(t) \cos \omega_k t dt \right] \cos \omega_k \tau \cdot \Delta \omega = \\
&= \sum_{k=1}^{\infty} S^T \omega_k \cos \omega_k \tau \cdot \Delta \omega.
\end{aligned} \tag{6}$$

Тут введено позначення

$$S^T \omega_k = \frac{2}{\pi} \int_0^T K_x(\tau) \cos \omega_k \tau d\tau = \frac{D_k}{\Delta \omega} \geq 0. \tag{7}$$

Нижче розглядаються лише абсолютно інтегровні кореляційні функції, тобто такі, для яких інтеграл

$$\int_0^\infty |K_x(\tau)| d\tau = c < \infty - збігається. \tag{8}$$

Якщо $T \rightarrow \infty$, то $\Delta \omega \rightarrow 0$ і частота стає неперевною. Тому її можна позначити просто ω (без індексу).

Функція $S^T \omega_k$ з (7) при $T \rightarrow \infty$ переходить у функцію

$$S_x \omega = \frac{2}{\pi} \int_0^\infty K_x(\tau) \cos \omega \tau d\tau. \tag{9}$$

Відмітимо, що згідно з (8) інтеграл (9) збігається і функція $S_x \omega$ існує.

Зauważення 1. При $\Delta \omega \rightarrow 0$ кількість частот у довільному скінченному інтервалі ω', ω'' нескінченно зростає. Щоб сума дисперсій окремих гармонік на цих частотах залишалась скінченою, необхідно нескінченно зменшувати дисперсії окремих гармонік. Тому в (7) $D_k \rightarrow 0$ при $\Delta \omega \rightarrow 0$.

Вираз (6) для кореляційної функції при $T \rightarrow \infty$ приймає вигляд

$$K_x \tau = \int_0^\infty S_x \omega \cos \omega \tau d\omega. \tag{10}$$

Означення 2. Стационарна випадкова функція $X(t)$, для якої існує така функція $S_x \omega$, що виконуються співвідношення (9) і (10), називається стационарною випадковою функцією з **неперевним спектром**. Функція $S_x \omega$ називається **спектральною щільністю**.

Інтегральні формули (9) і (10) носять назву формул Вінера-Хінчина і є фактично формулами взаємно оберненого косинус-перетворення Фур'є.

З (9) випливає, що функція $S_x(\omega)$, якщо її розглядати на всій осі $-\infty < \omega < \infty$, є парною

$$S_x(-\omega) = S_x(\omega). \quad (11)$$

Далі, виходячи з (7), можна записати:

$$S_x(\omega) \geq 0. \quad (12)$$

Покладемо в (10) $\tau = 0$. В результаті отримаємо формулу для дисперсії стаціонарної випадкової функції з неперервним спектром:

$$D_x = K_x(0) = \int_0^{\infty} S_x(\omega) d\omega. \quad (13)$$

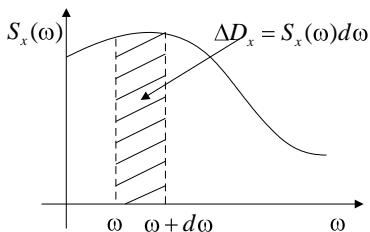


Рис. 2

З (13) видно, що функція $S_x(\omega)$ характеризує розподіл дисперсії стаціонарної випадкової функції по частотах, а саме, є щільністю дисперсії на неперервному спектрі частот. Тому елементарному частотному інтервалу $\Delta\omega = d\omega$ відповідає елементарна дисперсія $\Delta D_x = S_x(\omega)d\omega = S_x(\omega)\Delta\omega$ (рис. 2).

Отже, ми ввели нову і дуже важливу характеристику стаціонарної випадкової функції – спектральну щільність $S_x(\omega)$. Знаючи цю характеристику, можна за формулою (10) знайти кореляційну функцію і навпаки, знаючи кореляційну функцію, можна за формулою (9) визначити спектральну щільність. Тому для стаціонарної випадкової функції зовсім не важливо, яку характеристику задавати, кореляційну функцію чи спектральну щільність. У цьому відношенні ці характеристики рівносильні. На практиці у більшості випадків зручніше задавати спектральну щільність $S_x(\omega)$.

Зауваження 2. Ми прийшли до співвідношень (9) і (10), а отже, і до поняття спектральної щільності $S_x(\omega)$ за допомогою граничного переходу. Проте функцію $S_x(\omega)$ можна ввести формально, беручи за основу означення 2. При цьому невід'ємність функції $S_x(\omega)$ можна

довести, скориставшись невід'ємною визначеністю кореляційної функції $K_x \tau$. Дійсно, для $K_x \tau$ справедлива нерівність

$$\int_0^\infty \int_0^\infty K_x(t_2 - t_1) \varphi(t_1) \varphi(t_2) dt_1 dt_2 \geq 0, \quad (14)$$

де $\varphi(t)$ – довільна невипадкова функція (див. лекцію 28, зауваження 5). Запишемо спочатку нерівність (14) для функції $\varphi(t) = e^{-\varepsilon t} \cos \omega t$, а потім – для функції $\varphi(t) = e^{-\varepsilon t} \sin \omega t$, де ε – довільне додатне число. В результаті почлененного додавання одержаних таким чином нерівностей приходимо до нерівності

$$\int_0^\infty \int_0^\infty K_x(t_2 - t_1) \cos \omega(t_2 - t_1) e^{-\varepsilon(t_2 + t_1)} dt_1 dt_2 \geq 0. \quad (15)$$

Ввівши нові змінні $\tau = t_2 - t_1$, $\theta = t_1 + t_2$ і провівши необхідні викладки, можна звести нерівність (15) до вигляду

$$\int_0^\infty K_x(\tau) \cos \omega \tau \cdot e^{-\varepsilon \tau} d\tau \geq 0. \quad (16)$$

З останньої нерівності при $\varepsilon \rightarrow 0$ отримуємо, враховуючи (9), нерівність (12). Рекомендується провести необхідні викладки **самостійно**.

Поряд зі спектральною щільністю часто використовується нормована спектральна щільність

$$s_x(\omega) = \frac{S_x(\omega)}{D_x}. \quad (17)$$

Нормована спектральна щільність і нормована кореляційна функція $k_x(\tau) = K_x(\tau)/D_x$ зв'язані між собою тими ж перетвореннями Фур'є (див. формули (9), (10)):

$$s_x(\omega) = \frac{2}{\pi} \int_0^\infty k_x(\tau) \cos \omega \tau d\tau, \quad k_x(\tau) = \int_0^\infty s_x(\omega) \cos \omega \tau d\omega. \quad (18)$$

Оскільки $k_x(0) = 1$, з другої формули (18) одержуємо

$$\int_0^\infty s_x(\omega) d\omega = 1, \quad (19)$$

тобто площа, обмежена графіком нормованої спектральної щільності, дорівнює одиниці.

Спектральна щільність похідної стаціонарної випадкової функції

Як ми вже знаємо, похідна $Y(t) = \frac{dX(t)}{dt}$ диференційованої

стаціонарної випадкової функції $X(t)$ є стаціонарною випадковою функцією. Кореляційна функція похідної $Y(t)$ знаходиться за формулою (див. формулу (14) лекції 29)

$$K_y(\tau) = -\frac{d^2 K_x(\tau)}{d\tau^2} = -K''_x(\tau). \quad (20)$$

Підставимо в (20) вираз для кореляційної функції (10):

$$K_y(\tau) = \int_0^\infty \omega^2 S_x(\omega) \cos \omega \tau d\omega. \quad (21)$$

З іншого боку, згідно з (10)

$$K_y(\tau) = \int_0^\infty S_y(\omega) \cos \omega \tau d\omega. \quad (22)$$

Порівнюючи формули (21) і (22), отримуємо

$$S_y(\omega) = \omega^2 S_x(\omega). \quad (23)$$

Отже, диференціювання стаціонарної випадкової функції зводиться до множення її спектральної щільності на ω^2 .

Закінчимо дану лекцію наступним прикладом.

Приклад (випадковий імпульсний сигнал). Розглянемо випадкову функцію $X(t)$, реалізації якої будується наступним чином. У деякий випадковий момент часу T_1 з'являється прямоугільний імпульс довжиною t_0 з випадковою амплітудою A_1 , що розподілена незалежно від T_1 і має характеристики $m_{A_1} = 0$, $\sigma_{A_1} = \sigma$. У момент часу $T_1 + t_0$ цей імпульс закінчується і з'являється новий імпульс довжиною t_0 з випадковою амплітудою A_2 , розподіленою однаково з A_1 , але незалежно від A_1 і T_1 і т.д. Одну з реалізацій даного процесу показано на рис. 3 (при цьому вважається, що початок процесу (момент T_1) в ідеальній моделі імпульсного сигналу знаходиться на мінус нескінченості).

Знайти характеристики процесу: $m_x(t)$, $K_x(t_1, t_2)$ і, якщо процес є стаціонарним, обчислити спектральну щільність $S_x(\omega)$.

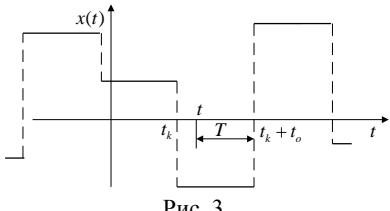


Рис. 3

Розв'язання. Зафіксуємо довільний момент часу t . Тоді цей момент часу буде належати деякому інтервалу $t_k, t_k + t_0$, де t_k – момент початку k -го імпульсу (рис. 3). Для всіх точок цього

інтервалу спостерігається випадкова величина A_k . Отже, $X|t = A_k$ і $M[X|t] = M[A_k] = m_{A_k} = 0$. Позначимо через T проміжок часу від моменту спостереження t до моменту закінчення k -го імпульсу $t_k + t_0$. Нехай тепер $t_2 = t + \tau$, $0 < \tau < t_0$. Маємо два випадки: $\tau \leq T$ і $\tau > T$, для яких за умовою

$$M[X|t, X|t+\tau | \tau \leq T] = M[A_k^2] = \sigma^2;$$

$$M[X|t, X|t+\tau | \tau > T] = M[A_k A_{k+1}] = 0.$$

Тому (див. лекцію 14, формулу (21)) при $0 < \tau < t_0$

$$M[X|t, X|t+\tau] = P\{\tau \leq T\}\sigma^2 + P\{\tau > T\} \cdot 0 = \sigma^2 P\{\tau \leq T\} = \sigma^2 (1 - P\{T < \tau\}).$$

Очевидно, що при $\tau \geq t_0$ $M[X|t, X|t+\tau] = 0$.

Оскільки моменти часу t і t_k випадкові, то можна вважати випадкову величину T розподіленою рівномірно на $0, t_0$. Тому

$P\{T < \tau\} = F(\tau) = \frac{\tau}{t_0}$ (див. лекцію 8, формулу (*)). Отже, маємо такий вираз для кореляційної функції:

$$K_x(t_1, t_2) = K_x(\tau) = \begin{cases} \sigma^2 \left(1 - \frac{\tau}{t_0}\right), & 0 < \tau < t_0 \\ 0, & \tau \geq t_0. \end{cases} \quad (24)$$

Оскільки випадкова функція є стаціонарною, знайдемо її спектральну щільність, скориставшись формулою (9):

$$\begin{aligned} S_x(\omega) &= \frac{2}{\pi} \int_0^\infty K_x(\tau) \cos \omega \tau d\tau = \\ &= \frac{2\sigma^2}{\pi} \int_0^{t_0} \left(1 - \frac{\tau}{t_0}\right) \cos \omega \tau d\tau = \frac{4\sigma^2}{\pi t_0 \omega^2} \sin^2 \frac{\omega t_0}{2}. \end{aligned} \quad (25)$$

Відмітимо, що розглянутий нами процес не є диференційовним. Чому?

Лекція № 32

Комплексна форма запису спектральних співвідношень стационарної випадкової функції. Деякі приклади стационарних випадкових функцій

З точки зору простоти математичних перетворень іноді зручніше використовувати комплексну форму запису спектральних співвідношень минулої лекції. Як і раніше, розглянемо два випадки.

I. Стационарна випадкова функція з дискретним спектром

Запишемо канонічний розклад стационарної випадкової функції

$$X(t) = m_x + \sum_{k=1}^n U_k \cos \omega_k t + V_k \sin \omega_k t \quad (1)$$

(див. (1) з лекції 31) у комплексній формі, скориставшись формулами Ейлера

$$\cos \omega_k t = \frac{e^{i\omega_k t} + e^{-i\omega_k t}}{2}, \quad \sin \omega_k t = \frac{e^{i\omega_k t} - e^{-i\omega_k t}}{2i}. \quad (2)$$

Підставимо формули (2) в (1):

$$X(t) = m_x + \sum_{k=1}^n \frac{1}{2} (U_k - iV_k) e^{i\omega_k t} + \frac{1}{2} (U_k + iV_k) e^{-i\omega_k t}. \quad (3)$$

Введемо комплексні випадкові величини

$$\Phi_k = \frac{1}{2} (U_k - iV_k), \quad \Phi_{-k} = \frac{1}{2} (U_k + iV_k), \quad k = 1, 2, \dots, n \quad (4)$$

і умовно поширимо область частот на від'ємні значення, поклавши $-\omega_k = \omega_{-k}$. Формула (3) набуває вигляду

$$X(t) = m_x + \sum_{k=-n}^n \Phi_k e^{i\omega_k t}, \quad \Phi_{-k} = \bar{\Phi}_k. \quad (5)$$

Якщо поширювати суму і на $k = 0$, то слід покласти $\omega_0 \equiv 0$ і $\Phi_0 \equiv 0$. Очевидно, що $M[\Phi_k] = 0$, $k = \pm 1, \pm 2, \dots, \pm n$ і $M[\Phi_k \bar{\Phi}_j] = 0$ при $k \neq j$, тобто Φ_k , $k = \pm 1, \pm 2, \dots, \pm n$ – некорельовані комплексні випадкові величини з нульовими математичними сподіваннями. Крім того, виконуються рівності

$$D[\Phi_k] = D[\Phi_{-k}] = M[|\Phi_k|^2] = \frac{1}{4} M[U_k^2 + V_k^2] = \frac{1}{2} D_k, \text{ де}$$

$$D_k = M[U_k^2] = M[V_k^2] = D[U_k] = D[V_k], \quad k = 1, 2, \dots, n.$$

Легко знайти кореляційну функцію випадкової функції (5):

$$K_x(t_1, t_2) = M \left[\sum_{k=-n}^n \Phi_k e^{i\omega_k t_1} \cdot \sum_{j=-n}^n \bar{\Phi}_j e^{-i\omega_j t_2} \right] = \sum_{k=-n}^n M[\Phi_k \bar{\Phi}_k] e^{i\omega_k t_1 - i\omega_k t_2}.$$

Отже,

$$K_x(\tau) = \sum_{k=-n}^n D[\Phi_k] e^{i\omega_k \tau}, \quad \tau = t_1 - t_2. \quad (6)$$

Якщо в (6) покласти $\tau = 0$, отримаємо вираз для дисперсії:

$$D_x = \sum_{k=-n}^n D[\Phi_k] = \sum_{k=1}^n D_k. \quad (7)$$

Отже, на відміну від дійсного спектрального розкладу (1) стаціонарної випадкової функції $X(t)$, де кожна гармоніка частотою $\omega_k > 0$ зображається одним доданком $U_k \cos \omega_k t + V_k \sin \omega_k t$, у комплексному спектральному розкладі (5) кожна гармоніка зображається двома некорельованими комплексними доданками $\Phi_k e^{i\omega_k t}$ та $\Phi_{-k} e^{i\omega_{-k} t}$ і дисперсія цієї гармоніки ділиться порівну між цими доданками.

Відмітимо, що поширення області частот на від'ємні значення позбавлене будь-якого фізичного смыслу, а є лише математичною абстракцією, що застосовується для подання співвідношень у зручній для використання формі.

Зauważення 1. Використовуючи (5), введемо випадкову функцію

$$\overset{\circ}{X}(t) = \sum_{k=-n}^n \overset{\circ}{X}_k(t), \quad \overset{\circ}{X}_k(t) = \Phi_k e^{i\omega_k t}. \quad (8)$$

Якщо домовитись середньоквадратичне значення $M\left[\left(\overset{\circ}{X}(t)\right)^2\right] = D_x$

називати **середньою енергією** стаціонарної випадкової функції $\overset{\circ}{X}(t)$, а рівність (7) переписати у вигляді

$$M\left[\left(\overset{\circ}{X}(t)\right)^2\right] = \sum_{k=-n}^n M\left[\left|\Phi_k e^{i\omega_k t}\right|^2\right] = \sum_{k=-n}^n M\left[\left|\overset{\circ}{X}_k(t)\right|^2\right], \quad (9)$$

то можна стверджувати, що середня енергія стаціонарної випадкової функції $\overset{\circ}{X}(t)$ складається з енергій окремих гармонік $\overset{\circ}{X}_k(t) = \Phi_k e^{i\omega_k t}$, $k = \pm 1, \pm 2, \dots, \pm n$.

II. Стационарна випадкова функція з неперервним спектром

У минулій лекції ми отримали формули, що пов'язують спектральну щільність $S_x(\omega)$ і кореляційну функцію $K_x(\tau)$ стационарної випадкової функції $X(t)$:

$$S_x(\omega) = \frac{2}{\pi} \int_0^\infty K_x(\tau) \cos \omega \tau d\tau, \quad K_x(\tau) = \int_0^\infty S_x(\omega) \cos \omega \tau d\omega. \quad (10)$$

Формулу для спектральної щільності можна записати у вигляді

$$S_x(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} K_x(\tau) \cos \omega \tau d\tau, \quad (11)$$

оскільки підінтегральний вираз є парною функцією. З іншого боку, можна записати

$$0 = \frac{1}{\pi} \int_{-\infty}^{\infty} K_x(\tau) \sin \omega \tau d\tau, \quad (12)$$

оскільки підінтегральна функція в (12) непарна, а інтеграл береться у симетричних межах. Якщо помножити формулу (12) на i і результат відняти від формули (11), отримаємо

$$S_x(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} K_x(\tau) e^{-i\omega\tau} d\tau. \quad (13)$$

Тут враховано, що $\cos \omega \tau - i \sin \omega \tau = e^{-i\omega\tau}$.

Далі, перепишемо формулу (10) для кореляційної функції у вигляді

$$K_x(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} S_x(\omega) \cos \omega \tau d\omega, \quad (14)$$

оскільки $S_x(-\omega) = S_x(\omega)$. З іншого боку,

$$0 = \frac{1}{2} \int_{-\infty}^{\infty} S_x(\omega) \sin \omega \tau d\omega. \quad (15)$$

Помножимо формулу (15) на i і результат додамо до формули (14). З врахуванням рівності $\cos \omega \tau + i \sin \omega \tau = e^{i\omega\tau}$ будемо мати

$$K_x(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} S_x(\omega) e^{i\omega\tau} d\omega. \quad (16)$$

Отже, ми одержали спектральний розклад кореляційної функції стационарної випадкової функції $X(t)$ у комплексній формі. Якщо в

(16) покласти $\tau = 0$, матимемо вираз для дисперсії:

$$D_x = K_x \cdot 0 = \frac{1}{2} \int_{-\infty}^{\infty} S_x(\omega) d\omega. \quad (17)$$

Формули (13) і (16), які пов'язують кореляційну функцію і спектральну щільність, є комплексною формою формул Вінера-Хінчина (10).

Приклад 1 (телеграфний сигнал). Розглянемо випадковий процес $X(t)$, який з однаковою ймовірністю може приймати лише два значення $+\sigma$ і $-\sigma$, причому число змін знака за час $\tau = t_2 - t_1$ не залежить від поведінки процесу до моменту t_1 і саме є випадковим процесом з одновимірним законом розподілу

$$P(N(\tau) = m) = P_m(\tau) = \frac{\mu^\tau}{m!} e^{-\mu\tau}, \quad \tau > 0, \quad m = 0, 1, 2, \dots, \quad (18)$$

де μ – заданий параметр, що трактується як середнє число змін знака за одиницю часу. Процес $N(\tau)$ з розподілами перерізів (18) носить назву Пуассона (див. формулу (13) лекції 11), а процес $X(t)$ називають **телеграфним сигналом**.

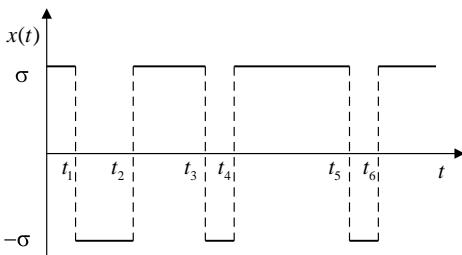


Рис. 1

Одну з можливих реалізацій телеграфного сигналу показано на рис. 1.

Необхідно знайти: а) кореляційну функцію телеграфного сигналу; б) спектральну щільність телеграфного сигналу.

Розв'язання. а) Очевидно, що $M[X(t)] = \sigma \cdot \frac{1}{2} + -\sigma \cdot \frac{1}{2} = 0$. Для двох довільних моментів часу t і $t + \tau$, $\tau > 0$ добуток $X(t)X(t + \tau)$ може прийняти одне з двох можливих значень: σ^2 або $-\sigma^2$. Ймовірність першого з цих значень дорівнює ймовірності того, що число змін знака за час τ є парним, другого – непарним. Тому $M[X(t)X(t + \tau)] =$

$$= \sigma^2 \sum_{m=0}^{\infty} P_{2m} \tau + -\sigma^2 \sum_{m=0}^{\infty} P_{2m+1} \tau = \sigma^2 \sum_{m=0}^{\infty} \frac{-\mu \tau^m}{m!} e^{-\mu \tau} = \sigma^2 e^{-2\mu \tau}, \text{ оскільки}$$

$$\sum_{m=0}^{\infty} \frac{-\mu \tau^m}{m!} = e^{-\mu \tau}.$$

Поширюючи область значень τ на всю вісь і враховуючи парність кореляційної функції, отримуємо

$$K_x(\tau) = M[X(t)X(t+\tau)] = \sigma^2 e^{-2\mu|\tau|}. \quad (19)$$

б) Згідно з формулою (13)

$$\begin{aligned} S_x(\omega) &= \frac{\sigma^2}{\pi} \int_{-\infty}^{\infty} e^{-2\mu|\tau|-i\omega\tau} d\tau = \frac{\sigma^2}{\pi} \left[\int_{-\infty}^0 e^{2\mu\tau-i\omega\tau} d\tau + \int_0^{\infty} e^{-2\mu\tau-i\omega\tau} d\tau \right] = \\ &= \frac{\sigma^2}{\pi} \left[\int_{-\infty}^0 e^{2\mu-i\omega\tau} d\tau + \int_0^{\infty} e^{-2\mu+i\omega\tau} d\tau \right] = \frac{\sigma^2}{\pi} \left[\frac{1}{2\mu-i\omega} + \frac{1}{2\mu+i\omega} \right]. \end{aligned}$$

Отже, спектральна щільність телеграфного сигналу має вигляд:

$$S_x(\omega) = \frac{\sigma^2}{\pi} \cdot \frac{4\mu}{\omega^2 + 4\mu^2}. \quad (20)$$

Приклад 2 (білий шум). Розглянемо параметричну сім'ю $X_\theta(t)$ стаціонарних випадкових функцій, спектральні щільності яких залежать від параметра $\theta > 0$:

$$S_x(\omega; \theta) = S_0 e^{-\theta|\omega|}, \quad S_0 = \text{const} > 0. \quad (21)$$

Використовуючи формулу (16), знаходимо відповідні кореляційні

$$\begin{aligned} \text{функції: } K_x(\tau; \theta) &= \frac{1}{2} \int_{-\infty}^{\infty} S_0 e^{-\theta|\omega|+i\omega\tau} d\omega = \frac{S_0}{2} \left[\int_{-\infty}^0 e^{\theta\omega+i\omega\tau} d\omega + \int_0^{\infty} e^{-\theta\omega+i\omega\tau} d\omega \right] = \\ &= \frac{S_0}{2} \left[\int_0^{\infty} e^{-\omega(\theta+i\tau)} d\omega + \int_0^{\infty} e^{-\omega(\theta-i\tau)} d\omega \right] = \frac{S_0}{2} \left[\frac{1}{\theta+i\tau} + \frac{1}{\theta-i\tau} \right]. \end{aligned}$$

Отже,

$$K_x(\tau; \theta) = S_0 \frac{\theta}{\theta^2 + \tau^2}. \quad (22)$$

Означення. *Ідеальним білим шумом* називають стаціонарну випадкову функцію, спектральна щільність якої стала на всій осі частот:

$$S_x(\omega) = S_0 = \text{const} > 0, \quad -\infty < \omega < \infty. \quad (23)$$

Очевидно, що стала спектральна щільність отримується з (21) при $\theta \rightarrow 0$:

$$S_x(\omega) = S_0 = \lim_{\theta \rightarrow 0} S_x(\omega; \theta). \quad (24)$$

Кореляційною функцією ідеального білого шуму називемо границю сім'ї кореляційних функцій (22) при $\theta \rightarrow 0$:

$$K_x(\tau) = \lim_{\theta \rightarrow 0} K_x(\tau; \theta) = \begin{cases} \infty, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}. \quad (25)$$

Зauważення 2. Кореляційну функцію ідеального білого шуму можна отримати з використанням інших параметричних сімей спектральних щільностей, наприклад, сім'ї

$$S_x(\omega; \theta) = S_0 e^{-\theta^2 \omega^2}, \quad S_0 > 0, \quad \theta > 0. \quad (26)$$

Зайдемо відповідну сім'ю кореляційних функцій:

$$\begin{aligned} K_x(\tau; \theta) &= \frac{S_0}{2} \int_{-\infty}^{\infty} e^{-\theta^2 \omega^2 + i\omega\tau} d\omega = \left| -\theta^2 \omega^2 + i\omega\tau = -\left(\theta\omega - \frac{i\tau}{2\theta} \right)^2 - \frac{\tau^2}{4\theta^2} \right| = \\ &= \frac{S_0}{2} e^{-\frac{\tau^2}{4\theta^2}} \int_{-\infty}^{\infty} e^{-\theta^2 \omega^2 - \frac{\tau^2}{4\theta^2}} d\omega = \left| \begin{array}{l} \theta\omega - \frac{i\tau}{2\theta} = t \\ d\omega = \frac{dt}{\theta} \end{array} \right| = \frac{S_0}{2\theta} e^{-\frac{\tau^2}{4\theta^2}} \int_{-\infty}^{\infty} e^{-t^2} dt. \end{aligned}$$

Враховуючи, що $\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$ – інтеграл Пуассона, отримуємо

$$K_x(\tau; \theta) = \frac{S_0 \sqrt{\pi}}{2\theta} e^{-\frac{\tau^2}{4\theta^2}}. \quad (27)$$

Переходячи до границі при $\theta \rightarrow 0$, будемо мати кореляційну функцію (25).

Відмітимо, що у розглянутому випадку кореляційна функція і спектральна щільність є показниковими функціями.

Зauważення 3. Якщо кореляційну функцію $K_x(\tau)$ для випадку (23) знаходить безпосередньо, то отримаємо (див. формули (16), (14)):

$$K_x(\tau) = \frac{S_0}{2} \int_{-\infty}^{\infty} e^{i\omega\tau} d\omega = S_0 \int_0^{\infty} \cos \omega \tau d\omega. \quad (28)$$

Але інтеграли в (28) розбігаються. У той час як ліва частина в (28) з точки зору фізичних міркувань має бути конкретною функцією від τ , справа в межах класичної математики маємо невизначеність. Це протиріччя означає, що в даному випадку ми виходимо за межі можливого використання відомого нам математичного апарату і що необхідно вводити нові математичні поняття. Отже, виходячи з “фізичних” міркувань, природно вимагати такого розширення поняття інтеграла, щоб рівності (28) мали певний смисл.

Подібного роду протиріччя у математиці та її застосуваннях привели до розробки нового математичного апарату – апарату узагальнених функцій. З використанням цього апарату ідеальний білий шум описується досить просто. Ми не будемо знайомитись з даним апаратом. Відмітимо лише, що його застосування для ідеального білого шуму приводить до кореляційної функції (25), а основна ідея, що полягає в тлумаченні розбіжного інтегралу (28) як границі збіжних інтегралів, продемонстрована нами вище.

Рівність нулю кореляційної функції ідеального білого шуму (25) означає некорельованість будь-яких його перерізів $X(t_1)$, $X(t_2)$, $t_1 \neq t_2$. Цей факт, як і нескінченість дисперсії $D_x = K_x \cdot 0$, свідчить про те, що реалізувати ідеальний білий шум неможливо. У всякому фізичному явищі значення випадкової функції в даній точці у деякій мірі визначає і її значення в близьких до неї точках, тобто значення реальних випадкових функцій в достатньо близьких точках корельовані.

Хоча у природі ідеальних білих шумів і не існує, це поняття є зручною математичною абстракцією, що використовується як в самій теорії випадкових функцій, так і в її застосуваннях. З практичної точки зору випадкову функцію можна вважати білим шумом, якщо кореляція її перерізів поширюється лише на такі інтервали зміни аргументу, які не можна розрізнати при даній точності вимірювань.

Очевидно, що розглянутий вище телеграфний сигнал та інші випадкові функції, кореляційна функція і спектральна щільність яких мають вигляд

$$K_x(\tau) = \sigma^2 e^{-\alpha|\tau|}, \quad S_x(\omega) = \frac{\sigma^2}{\pi} \cdot \frac{2\alpha}{\omega^2 + \alpha^2}, \quad (29)$$

можна практично вважати білим шумом (реальним білим шумом), якщо α достатньо велике. Дійсно, для достатньо великого α значення показникової кореляційної функції буде як завгодно малим при $\tau \neq 0$. При цьому спектральна щільність буде практично постійною у значному діапазоні частот. Отже, для реального білого шуму спектральна щільність не є сталою, але може вважатись сталою у деякому діапазоні частот, причому за межами цього діапазону поведінкою спектральної щільності, як правило, не цікавляється.

Щоб відобразити обмеженість діапазону частот, в якому спектральна щільність стала, вводять поняття *смугового білого шуму*.

Приклад 3 (смуговий низькочастотний білий шум). Знайти кореляційну функцію стаціонарної випадкової функції $X(t)$,

спектральна щільність якої має вигляд

$$S_x(\omega) = \begin{cases} S_0, & |\omega| < \omega_0 \\ 0, & |\omega| \geq \omega_0 \end{cases}, \quad S_0 = \text{const} > 0. \quad (30)$$

Розв'язання. Скористаємося формулою (10):

$$K_x(\tau) = \int_0^\infty S_0 \cos \omega \tau d\tau = S_0 \int_0^{\omega_0} \cos \omega \tau d\tau = S_0 \frac{\sin \omega_0 \tau}{\tau}. \quad (31)$$

Для дисперсії отримуємо

$$D_x = K_x(0) = \lim_{\tau \rightarrow 0} K_x(\tau) = S_0 \omega_0 \lim_{\tau \rightarrow 0} \frac{\sin \omega_0 \tau}{\omega_0 \tau} = S_0 \omega_0. \quad (32)$$

Зauważення 4. Якщо в (30) $\omega_0 \rightarrow \infty$, то дисперсія (32) стає нескінченно великою. У всіх точках $\tau_k = \frac{k\pi}{\omega_0}$, $k = \pm 1, \pm 2, \dots$ функція

$K_x(\tau)$ з (31) дорівнює нулю. Якщо $\omega_0 \rightarrow \infty$, то сусідні точки τ_k (нулі функції) зближуються і в граничному випадку “неперервно” заповнюють всю вісь $-\infty < \tau < \infty$.Хоча з точки зору класичної математики наведені міркування позбавлені смислу (границя $\lim_{\omega_0 \rightarrow \infty} \frac{\sin \omega_0 \tau}{\tau}$ не існує у звичайному розумінні), вони також приводять до кореляційної функції вигляду (25).

Лекція № 33

Лінійні перетворення випадкових функцій.

Перетворення стаціонарної випадкової функції лінійною динамічною системою з постійними коефіцієнтами

При проходженні через той чи інший пристрій, систему автоматичного управління, регулювання тощо випадкова функція $X(t)$ підлягає тим чи іншим перетворенням, в результаті чого на виході отримується нова випадкова функція $Y(t)$. Символічно це можна записати так: $Y(t) = AX(t)$. Перетворення A (назвемо його **оператором**) може мати будь-який вигляд, наприклад, диференціювання (див. приклад 4 лекції 27), інтегрування (див. приклад 5 лекції 27), множення на невипадкову функцію тощо. Подібні перетворення або оператори бувають лінійними і нелінійними. У свою чергу лінійні оператори поділяються на однорідні та

неоднорідні. Лінійним однорідним оператором L_t^0 називається таке перетворення, яке задовольняє принципу суперпозиції:

$$L_t^0 c_1 x_1(t) + c_2 x_2(t) = c_1 L_t^0 x_1(t) + c_2 L_t^0 x_2(t), \quad (1)$$

де $x_1(t)$ і $x_2(t)$ – довільні невипадкові функції (реалізації випадкової функції $X(t)$), а c_1, c_2 – довільні невипадкові сталі. Прикладами лінійних однорідних операторів є:

- 1) оператор диференціювання

$$y(t) = \frac{dx(t)}{dt}, \quad L_t^0 \equiv \frac{d}{dt};$$

- 2) оператор інтегрування

$$y(t) = \int_0^t x(t) dt, \quad L_t^0 \equiv \int_0^t \cdot dt;$$

- 3) оператор множення на невипадкову функцію $g(t)$

$$y(t) = g(t) x(t), \quad L_t^0 \equiv g(t) \cdot;$$

- 4) диференціальний оператор

$$\begin{aligned} y(t) &= b_m \frac{d^m x(t)}{dt^m} + b_{m-1} \frac{d^{m-1} x(t)}{dt^{m-1}} + \dots + b_0 x(t), \\ L_t^0 &\equiv b_m \frac{d^m \cdot}{dt^m} + b_{m-1} \frac{d^{m-1} \cdot}{dt^{m-1}} + \dots + b_0 \cdot, \end{aligned}$$

де b_0, b_1, \dots, b_m – невипадкові сталі.

Якщо лінійний оператор є сумою лінійного однорідного оператора і деякої невипадкової функції $f(t)$, то він буде неоднорідним:

$$L_t = L_t^0 + f(t) \text{ – неоднорідний лінійний оператор.} \quad (2)$$

Наведемо приклади лінійних неоднорідних перетворень:

$$\text{а) } y(t) = \frac{dx(t)}{dt} + f(t); \quad \text{б) } y(t) = \int_0^t x(t) dt + f(t); \quad \text{в) } y(t) = g(t) \cdot x(t) + f(t).$$

Зручним апаратом при виконанні лінійних перетворень над випадковими функціями є зображення останніх у вигляді канонічного розкладу

$$X(t) = m_x(t) + \sum_{k=1}^n \overset{\circ}{V}_k \varphi_k(t). \quad (3)$$

Дійсно, у цьому випадку дія оператора L_t^0 на випадкову функцію $X(t)$ фактично зводиться до дій цього оператора на невипадкові

функції – математичне сподівання $m_x t$ та координатні функції $\varphi_k t$:

$$Y t = L_t^0 X t = L_t^0 m_x t + \sum_{k=1}^n \overset{\circ}{V}_k L_t^0 \varphi_k t. \quad (4)$$

Оскільки справа в (4) знову маємо канонічний розклад, одразу ж знаходимо і характеристики випадкової функції $Y t$:

$$\begin{aligned} m_y t &= L_t^0 m_x t; \quad K_y t_1, t_2 = \sum_{k=1}^n L_{t_1}^0 \varphi_k t_1 L_{t_2}^0 \varphi_k t_2 \overset{\circ}{V}_k = \\ &= L_{t_1}^0 L_{t_2}^0 \sum_{k=1}^n \varphi_k t_1 \varphi_k t_2 \overset{\circ}{V}_k = L_{t_1}^0 L_{t_2}^0 K_x t_1, t_2. \end{aligned} \quad (5)$$

Формули (5) узагальнюють одержані нами раніше формулі (18) і (20) з лекції 27 для операцій диференціювання та інтегрування випадкових функцій.

Очевидно, що дія лінійного неоднорідного оператора (2) не змінить кореляційної функції (5). Математичне сподівання при цьому зміниться на величину $f x$.

Нехай на вхід деякого пристрою, системи тощо (далі будемо вживати термін **динамічна система**) надходить випадкова функція $X t$. Система перетворює цю функцію і на виході маємо випадкову функцію $Y t$.

Під **лінійною динамічною системою з постійними коефіцієнтами** розуміють систему, яка описується лінійним диференціальним рівнянням з постійними коефіцієнтами

$$a_n \frac{d^n y t}{dt^n} + a_{n-1} \frac{d^{n-1} y t}{dt^{n-1}} + \dots + a_0 y t = b_m \frac{d^m x t}{dt^m} + b_{m-1} \frac{d^{m-1} x t}{dt^{m-1}} + \dots + b_0 x t, \quad (6)$$

де $x t$ – реалізація вхідної випадкової функції $X t$, а $y t$ – реалізація випадкової функції $Y t$ на виході системи. Очевидно, що зліва і справа в (6) стоять лінійні однорідні оператори, що діють відповідно на функції $y t$ та $x t$. Далі вважаємо, що на вхід системи надходить **стационарна** випадкова функція $X t$. Введемо формальні позначення $p = \frac{d}{dt}$, $p^2 = \frac{d^2}{dt^2}$ і т.д. Тоді рівняння (6) можна

записати у вигляді

$$a_n p^n + a_{n-1} p^{n-1} + \dots + a_0 \quad y = \quad b_m p^m + b_{m-1} p^{m-1} + \dots + b_0 \quad x, \quad (7)$$

звідки отримуємо “умовний розв’язок”

$$y = H \quad p \quad x, \quad (8)$$

де

$$H \quad p = \frac{b_m p^m + b_{m-1} p^{m-1} + \dots + b_0}{a_n p^n + a_{n-1} p^{n-1} + \dots + a_0}. \quad (9)$$

Зauważення 1. Проведений нами формальний перехід від рівняння (6) до співвідношення (7) з подальшим введенням функції $H p$ обґрунтovується наступним: якщо подіти на рівняння (6) перетворенням Лапласа при нульових початкових умовах на функції $y t$ і $x t$, то прийдемо до співвідношення (7), де під x і y слід розуміти перетворення Лапласа функцій $x t$ і $y t$. Отже, функція $H p$ є відношенням перетворених по Лапласу вихідного та вхідного сигналів. Тому змінну p в (9) слід вважати комплексною.

Функція $H p$ називається **передатною функцією** лінійної динамічної системи.

Якщо динамічна система стійка, то через деякий час з моменту початку її роботи перехідні процеси в ній закінчаться і вона вийде на стаціонарний режим, тобто вихідну випадкову функцію $Y t$ можна вважати стаціонарною.

Зauważення 2. Кажуть, що лінійна динамічна система стійка, якщо функція $H p$, як функція комплексної змінної p , не має полюсів у правій півплощині комплексної площини p (див. Додаток 4).

Тому нижче як випадкову функцію $X t$, так і випадкову функцію $Y t$ вважаємо стаціонарними.

Нехай на вхід системи надходить гармоніка $x t = e^{i\omega t}$. Реакцію системи у стаціонарному режимі також шукаємо у вигляді гармоніки $y t = A e^{i\omega t}$, де коефіцієнт A потрібно визначити. Підставляючи

вирази $x t$ і $y t$ у рівняння (7) і враховуючи, що $\frac{d^n}{dt^n} e^{i\omega t} = i\omega^n e^{i\omega t}$,

після скорочення на $e^{i\omega t}$ одержуємо $A \ a_n \ i\omega^n + \dots + a_0 = b_m \ i\omega^m + \dots + b_0$, звідки $A = H \ i\omega$.

Отже,

$$y(t) = H \ i\omega \ e^{i\omega t}. \quad (10)$$

Передатна функція $H(p)$ при $p = i\omega$ називається **частотною характеристикою** лінійної динамічної системи. Модуль $|H(i\omega)|$ є коефіцієнтом підсилення амплітуди гармоніки з частотою ω .

Нехай тепер на вхід системи подається стаціонарна випадкова функція $X(t)$ з канонічним розкладом (див. лекцію 32)

$$X(t) = m_x + \sum_{k=-n}^n \Phi_k e^{i\omega_k t}. \quad (11)$$

При проходженні через систему в силу її лінійності k -ий член розкладу (11) множиться на частотну характеристику $H(i\omega_k)$.

Математичне сподівання $m_x = \text{const}$ можна розглядати як гармоніку нульової частоти (в (11) покладаємо $\Phi_0 \equiv 0$). Отже,

$$Y(t) = H(0)m_x + \sum_{k=-n}^n \Phi_k H(i\omega_k) e^{i\omega_k t} = m_y + \sum_{k=-n}^n \Phi_k H(i\omega_k) e^{i\omega_k t}, \quad (12)$$

де

$$m_y = H(0)m_x = \frac{b_0}{a_0}m_x. \quad (13)$$

Рекомендується переконатись у справедливості формули (12) **самостійно**.

Вираз (12) є канонічним розкладом стаціонарної випадкової функції $Y(t)$. Знайдемо дисперсію комплексної випадкової величини $\Phi_k H(i\omega_k)$:

$$\begin{aligned} D[\Phi_k H(i\omega_k)] &= M[|\Phi_k H(i\omega_k)|^2] = M[|\Phi_k|^2 |H(i\omega_k)|^2] = \\ &= |H(i\omega_k)|^2 M[|\Phi_k|^2] = |H(i\omega_k)|^2 D_k. \end{aligned} \quad (14)$$

Отже, при перетворенні стаціонарної випадкової функції лінійною системою кожна ордината її спектра множиться на квадрат модуля частотної характеристики для відповідної частоти.

Переходячи до неперервного спектра шляхом міркувань, аналогічних міркуванням з лекції 31, можна отримати з (14) відповідне співвідношення між спектральними щільностями вхідної та вихідної стаціонарних випадкових функцій, а саме

$$S_y(\omega) = |H(i\omega)|^2 S_x(\omega). \quad (15)$$

Знаючи спектральну щільність (15), можна знайти кореляційну функцію $K_y(0)$ та стаціонарної випадкової функції на вихіді системи, скориставшись формулою (16) з лекції 32. Зокрема, для дисперсії на вихіді системи можна записати

$$D_y = K_y(0) = \frac{1}{2} \int_{-\infty}^{\infty} |H(i\omega)|^2 S_x(\omega) d\omega. \quad (16)$$

Щоб дисперсія D_y була скінченою, інтеграл (16) повинен збігатися. З виразу (9) для функції $H(p)$ випливає, наприклад, що достатньою умовою цієї збіжності є умова $m \leq n$.

Знаходження кореляційної функції за відомою спектральною щільністю у багатьох випадках ускладнюється математичними труднощами при обчисленні інтегралів. У таких випадках, як правило, звертаються до методів теорії функцій комплексної змінної, а саме, до теорії лишків. Оскільки спектральні щільності у більшості є дробово-раціональними функціями частоти (або апроксимуються такими функціями), то в Додатку 4 наведені елементи теорії лишків для знаходження інтегралів саме від таких функцій.

Приклад 1. На вхід лінійної динамічної системи, що описується диференціальним рівнянням

$$y''(t) + 3y'(t) + 2y(t) = x'(t) + 2x(t), \quad (17)$$

надходить випадковий телеграфний сигнал (див. приклад 1 з лекції 32).

Знайти кореляційну функцію та дисперсію стаціонарного процесу на вихіді системи.

Розв'язання. Оскільки $m=1 < n=2$, то на вихіді системи дисперсія буде скінченою. Передатна функція системи має вигляд

$$H(p) = \frac{p+2}{p^2+3p+2} = \frac{1}{p+1}. \quad (18)$$

Тому характеристики випадкового сигналу на вихіді даної системи будуть такими ж, як і на вихіді лінійної динамічної системи, що описується диференціальним рівнянням

$$y' t + y t = x t . \quad (19)$$

Оскільки полюс передатної функції $p = -1$ знаходиться у лівій півплощині комплексної площини, то система стійка.

Знаходимо квадрат модуля частотної характеристики:

$$|H i\omega|^2 = H i\omega H -i\omega = \frac{1}{i\omega+1} \cdot \frac{1}{-i\omega+1} = \frac{1}{\omega^2+1}. \quad (20)$$

Для телеграфного сигналу спектральна щільність

$$S_x \omega = \frac{\sigma^2}{\pi} \cdot \frac{4\mu}{\omega^2 + 4\mu^2}.$$

Тому згідно з формулою (15)

$$S_y \omega = \frac{S_x \omega}{\omega^2 + 1} = \frac{4\mu\sigma^2}{\pi} \cdot \frac{1}{\omega^2 + 4\mu^2 \omega^2 + 1}. \quad (21)$$

Знаходимо кореляційну функцію вихідного сигналу:

$$K_y \tau = \frac{1}{2} \int_{-\infty}^{\infty} S_y \omega e^{i\omega\tau} d\omega = \frac{2\mu\sigma^2}{\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega\tau} d\omega}{\omega^2 + 4\mu^2 \omega^2 + 1}. \quad (22)$$

Останній інтеграл обчислюється за допомогою формул (9) Додатку 4 при $\tau > 0$ (див. приклад 2 цього Додатку):

$$\int_{-\infty}^{\infty} \frac{e^{i\omega\tau} d\omega}{\omega^2 + 4\mu^2 \omega^2 + 1} = \frac{\pi e^{-2\mu\tau} [2\mu e^{\tau/2\mu-1} - 1]}{2\mu 4\mu^2 - 1}. \quad (23)$$

Враховуючи парність кореляційної функції, остаточно отримуємо

$$K_y \tau = \sigma^2 \frac{e^{-2\mu|\tau|} [2\mu e^{|\tau|/2\mu-1} - 1]}{4\mu^2 - 1}. \quad (24)$$

Дисперсія сигналу на вихіді дорівнює

$$D_y = K_y 0 = \frac{\sigma^2}{1 + 2\mu}. \quad (25)$$

Отже, дисперсія сигналу на вихіді системи в $1 + 2\mu$ разів менша за дисперсію вхідного телеграфного сигналу.

Приклад 2. Нехай лінійна динамічна система описується рівнянням

$$y'' t + 2hy' t + \omega_0^2 y t = x t . \quad (26)$$

Це може бути резонансний контур, механічний маятник з внутрішнім

тертім тощо. Власні коливання цієї системи (при $x(t) = 0$) описуються функцією вигляду $y(t) = Ae^{-ht} \sin(\omega_B t + \theta)$, де $\omega_B = \sqrt{\omega_0^2 - h^2}$ – власна частота системи. Запишемо вирази для передатної функції та квадрата модуля частотної характеристики:

$$H(p) = \frac{1}{p^2 + 2h + \omega_0^2};$$

$$|H(i\omega)|^2 = H(i\omega)H(-i\omega) = \frac{1}{\omega^2 - \omega_0^2 + 4h^2\omega^2}. \quad (27)$$

Нехай на вхід системи надходить ідеальний білий шум ($S_x(\omega) = c^2$, $-\infty < \omega < \infty$). Тоді спектральна щільність на виході має вигляд

$$S_y(\omega) = \frac{c^2}{\omega^2 - \omega_0^2 + 4h^2\omega^2}. \quad (28)$$

Максимум спектральної щільності $S_y(\omega)$ (різко виражений для малого “коєфіцієнта тертя” h) досягається при $\omega^2 = \omega_0^2 - 2h^2 = \omega_B^2 - h^2$. При віддаленні від точки максимуму функція $S_y(\omega)$ швидко спадає (рис. 1). Це свідчить про те, що у вихідному стаціонарному процесі $Y(t)$ різко переважають “елементарні коливання” з частотами, близькими до власної частоти системи. Рекомендується показати **самостійно**, що дисперсія D_y на виході системи дорівнює $D_y = \frac{\pi c^2}{4\omega_0^2 h}$.

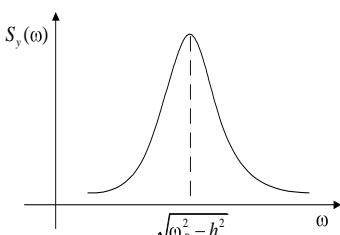


Рис. 1

З певним наближенням збудження системи можна вважати білим шумом, якщо це, наприклад, випадкові імпульси з прикладу лекції 31 при достатньо малому t_0 .

На наведеному прикладі можна прослідкувати деякі особливості поведінки лінійних систем під дією випадкових коливань. Так, якщо на

вхід системи (26) надходить випадкова гармоніка з довільною частотою Ω , то в усталеному режимі система буде здійснювати вимушені коливання з тією ж частотою Ω . Проте, якщо при великій частоті Ω (набагато більшій, ніж власна частота системи: $\Omega \gg \omega_B$)

замінити гармонічне збудження збудженням випадковими імпульсами, що виникають приблизно через малі проміжки часу $t_0 \ll \frac{\pi}{\Omega}$, і вважати це останнє збудження білим шумом, то отримаємо результат, який якісно відрізняється від попереднього. А саме, матимемо “коливання”, в яких різко переважають “елементарні коливання” з околу власної частоти ω_B .

Зміст

Від автора	3
Вступ	4
Розділ 1. ЕЛЕМЕНТИ ТЕОРІЇ ЙМОВІРНОСТЕЙ	6
Лекція №1 Поняття випадкового експерименту і випадкової події. Схематизація випадкових подій. Операції над подіями. Відносна частота випадкової події	6
Лекція №2 Аксіоматичне означення ймовірності події. Ймовірнісний простір. Класична ймовірнісна схема	13
Лекція №3 Комбінаторний метод знаходження ймовірності в класичній схемі	21
Лекція №4 Дискретний ймовірнісний простір. Геометричні ймовірності. Неперервний ймовірнісний простір Умовні ймовірності та незалежність подій	30
Лекція №5 Формула повної ймовірності. Формула Байеса. Послідовності незалежних випробувань. Формула Бернуллі	37
Лекція №6 Формули наближених обчислень в схемі Бернуллі	43
Лекція №7 Поняття випадкової величини	49
Лекція №8 Дискретні та неперервні випадкові величини	57
Лекція №9 Числові характеристики випадкових величин	64
Лекція №10 Нормальний закон розподілу випадкової величини	71
Лекція №11 Деякі інші важливі для практики розподіли випадкових величин	78
Лекція №12 Спільний розподіл декількох випадкових величин (випадкові вектори)	85
Лекція №13 Числові характеристики системи випадкових величин (X, Y). Умовні числові характеристики. Регресія. Нормальний розподіл на площині	94
Лекція №14 Числові характеристики функцій випадкових величин.....	102
Лекція №15 Закони розподілу функцій випадкових величин. Задача композиції.....	110
Лекція №16 Закон великих чисел. Нерівність Чебишова. Теореми Чебишова та їх наслідки	119
Лекція №17 Характеристичні функції. Центральна гранична теорема.....	127
Розділ 2. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ	136
Лекція №18 Поняття вибірки. Методи описування вибірки. Характеристики вибірки	136

<i>Лекція №19</i> Оцінка числових характеристик та параметрів розподілу випадкової величини за результатами вибірки. Точкові оцінки та їх властивості	144
<i>Лекція №20</i> Інтервальні оцінки параметрів розподілу	151
<i>Лекція №21</i> Приклади на побудову довірчих інтервалів	157
<i>Лекція №22</i> Основні поняття статистичної перевірки статистичних гіпотез	163
<i>Лекція №23</i> Перевірка гіпотез про параметри нормально розподілених випадкових величин	172
<i>Лекція №24</i> Перевірка гіпотези про рівність дисперсій двох нормальних генеральних сукупностей. Гіпотези про закони розподілу. Критерій згоди χ^2	179
<i>Лекція №25</i> Приклади на застосування критерію χ^2	186
Розділ 3. ЕЛЕМЕНТИ КОРЕЛЯЦІЙНОЇ ТЕОРІЇ ВИПАДКОВИХ ФУНКІЙ	195
<i>Лекція №26</i> Означення випадкової функції. Закони розподілу та числові характеристики випадкових функцій	195
<i>Лекція №27</i> Невід'ємна визначеність кореляційної функції. Канонічний розклад випадкової функції. Похідна та інтеграл від випадкової функції, що має канонічний розклад	204
<i>Лекція №28</i> Диференціювання та інтегрування випадкових функцій у середньоквадратичному	213
<i>Лекція №29</i> Визначення характеристик випадкової функції з досліду. Стационарні випадкові функції	220
<i>Лекція №30</i> Ергодичність стационарних випадкових функцій. Комплексні випадкові функції та їх характеристики	228
<i>Лекція №31</i> Стационарні випадкові функції з дискретним та неперервним спектрами. Спектральна щільність	238
<i>Лекція №32</i> Комплексна форма запису спектральних співвідношень стационарної випадкової функції. Деякі приклади стационарних випадкових функцій	246
<i>Лекція №33</i> Лінійні перетворення випадкових функцій. Перетворення стационарної випадкової функції лінійною динамічною системою з постійними коефіцієнтами	253
Додаток 1 Елементарні поняття теорії множин. Метод математичної індукції.....	262
Додаток 2 Таблиця 1. Функція розподілу нормального закону $N(0,1)$	271
Таблиця 2. Значення функції щільності нормальногорозподілу $N(0,1)$	272
Таблиця 3. Квантіліального розподілу $N(0,1)$	272

Таблиця 4. Квантілі хі-квадрат розподілу	273
Таблиця 5. Квантілі розподілу Стьюдента.....	275
Таблиця 6. Квантілі розподілу Фішера	276
Додаток 3 Таблиця 1. Критерій значущості для перевірки гіпотез про дисперсії нормально розподілених генеральних сукупностей	282
Таблиця 2. Критерій значущості для перевірки гіпотез про середні нормально розподілених генеральних сукупностей	283
Додаток 4 Знаходження визначених інтегралів за допомогою лишків.....	285
Список використаної та рекомендованої літератури.....	287
Зміст.....	288