

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені І. І. МЕЧНИКОВА  
ФАКУЛЬТЕТ МАТЕМАТИКИ, ФІЗИКИ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

**Л. Г. Коваленко**

# **МАТЕМАТИЧНА СТАТИСТИКА**

*Частина 1*

НАВЧАЛЬНИЙ ПОСІБНИК  
для здобувачів 3-го курсу  
першого (бакалаврського) рівня вищої освіти  
спеціальності 111 «Математика»

ОДЕСА  
ОНУ  
2024

**УДК 519.2 (075.8)**  
**K562**

**Автор:**

**Л. Г. Коваленко**, кандидат фізико-математичних наук, доцент.

**Рецензенти:**

**О. І. Третяк**, доктор фізико-математичних наук, професор, завідувач кафедри вищої математики Державного університету інтелектуальних технологій та зв'язку;

**М. О. Малаксіано**, доктор технічних наук, професор, завідувач кафедри технічної кібернетики й інформаційних технологій імені проф. Р. В. Меркта Одеського національного морського університету.

*Рекомендовано до видання вченою радою  
ОНУ імені І. І. Мечникова  
Протокол № 8 від 26 березня 2024 р.*

**Коваленко Л. Г.**

**K562** Математична статистика (частина 1) : навч. посіб. для здобув. 3-го курсу першого (бакалавр.) рівня вищ. освіти спец. 111 «Математика» / Л. Г. Коваленко. Електронні текстові дані (1 файл : 0,6 МБ). Одеса : Одес. нац. ун-т ім. І. І. Мечникова, 2024. – 120 с.  
ISBN 978-617-689-463-6

*Навчальний посібник містить програмний матеріал перших двох змістових модулів з математичної статистики, що є складовою частиною дисципліни «Теорія ймовірності та математична статистика» для студентів третього курсу спеціальності 111 «Математика».*

*Розглядаються базові поняття статистики, методи здобуття та аналізу властивостей точкових, інтервальних оцінок невідомих параметрів розподілу, способи формування та перевірки статистичних гіпотез, класичні критерії щодо гіпотез про розподіл, незалежність та параметри розподілів.*

*Усі означення і теореми ілюструються прикладами.*

*Кожна тема доповнена вправами та контрольними питаннями.*

*Для студентів та аспірантів математичних спеціальностей.*

**УДК 519.2 (075.8)**

# Зміст

|                                                            |           |
|------------------------------------------------------------|-----------|
| Передмова . . . . .                                        | 6         |
| Вступ . . . . .                                            | 8         |
| <i>Контрольні запитання</i> . . . . .                      | 12        |
| <b>1 ОСНОВИ МАТЕМАТИЧНОЇ СТАТИСТИКИ</b>                    | <b>13</b> |
| §1. Первинна обробка вибірових даних . . . . .             | 13        |
| 1.1. Ази статистики . . . . .                              | 13        |
| 1.2. Вибірковий розподіл . . . . .                         | 15        |
| 1.3. Інтервальний ряд . . . . .                            | 17        |
| 1.4. Емпірична щільність . . . . .                         | 20        |
| <i>Контрольні запитання</i> . . . . .                      | 21        |
| <i>Вправи та задачі</i> . . . . .                          | 21        |
| §2. Оцінки та їх властивості . . . . .                     | 22        |
| 2.1. Види статистичних оцінок . . . . .                    | 22        |
| 2.2. Властивості емпіричної функції розподілу . . . . .    | 23        |
| 2.3. Властивості емпіричної щільності . . . . .            | 26        |
| 2.4. Властивості вибірових моментів . . . . .              | 26        |
| 2.5. Формування припущення щодо розподілу . . . . .        | 29        |
| <i>Контрольні запитання</i> . . . . .                      | 30        |
| <i>Вправи та задачі</i> . . . . .                          | 30        |
| §3. Методи здобуття точкових оцінок параметрів розподілу . | 32        |
| 3.1. Метод моментів . . . . .                              | 32        |
| 3.2. Метод максимальної правдоподібності . . . . .         | 36        |

|                                                                                                            |    |
|------------------------------------------------------------------------------------------------------------|----|
| <i>Контрольні запитання</i> . . . . .                                                                      | 46 |
| <i>Вправи та задачі</i> . . . . .                                                                          | 46 |
| §4. Оцінки з мінімальною дисперсією . . . . .                                                              | 47 |
| 4.1. Порівняння оцінок . . . . .                                                                           | 47 |
| 4.2. Нерівність Рао-Крамера . . . . .                                                                      | 49 |
| 4.3. Способи подання інформації Фішера . . . . .                                                           | 53 |
| <i>Контрольні запитання</i> . . . . .                                                                      | 59 |
| <i>Вправи та задачі</i> . . . . .                                                                          | 59 |
| §5. Інтервальні оцінки параметрів розподілу . . . . .                                                      | 60 |
| 5.1. Поняття довірчого інтервалу . . . . .                                                                 | 60 |
| 5.2. Довірчий інтервал для математичного сподівання нормального розподілу з відомою дисперсією . . . . .   | 62 |
| 5.3. Розподіли, пов'язані з нормальним . . . . .                                                           | 64 |
| 5.4. Довірчий інтервал для математичного сподівання нормального розподілу з невідомою дисперсією . . . . . | 67 |
| 5.5. Довірчий інтервал для середньоквадратичного відхилення нормального розподілу . . . . .                | 70 |
| <i>Контрольні запитання</i> . . . . .                                                                      | 71 |
| <i>Вправи та задачі</i> . . . . .                                                                          | 72 |

## **2 ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ 73**

|                                                                      |    |
|----------------------------------------------------------------------|----|
| §6. Гіпотези та критерії . . . . .                                   | 73 |
| 6.1. Поняття статистичної гіпотези. Класифікація помилок . . . . .   | 73 |
| 6.2. Перевірка статистичних гіпотез . . . . .                        | 77 |
| <i>Контрольні запитання</i> . . . . .                                | 79 |
| <i>Вправи та задачі</i> . . . . .                                    | 79 |
| §7. Критерій згоди $\chi^2$ . . . . .                                | 80 |
| 7.1. Критерій $\chi^2$ про розподіл (непараметричний) . . . . .      | 80 |
| 7.2. Критерій $\chi^2$ для складної параметричної гіпотези . . . . . | 83 |
| 7.3. Критерій $\chi^2$ як критерій незалежності . . . . .            | 85 |
| <i>Контрольні запитання</i> . . . . .                                | 88 |
| <i>Вправи та задачі</i> . . . . .                                    | 88 |
| §8. Види критеріїв . . . . .                                         | 89 |

|                                                                                |     |
|--------------------------------------------------------------------------------|-----|
| 8.1. Рівень значущості та потужність критерію . . . . .                        | 89  |
| 8.2. Мінімаксні, баєсовські та найбільш потужні критерії . . . . .             | 90  |
| 8.3. Критерій відношення правдоподібності . . . . .                            | 92  |
| 8.4. Асимптотичні характеристики критеріїв . . . . .                           | 97  |
| 8.5. Критерії згоди . . . . .                                                  | 97  |
| 8.6. Фактично досяжний рівень значущості . . . . .                             | 99  |
| <i>Контрольні запитання . . . . .</i>                                          | 100 |
| <i>Вправи та задачі . . . . .</i>                                              | 100 |
| §9. Перевірка гіпотез про параметри нормальних розподілів . . . . .            | 101 |
| 9.1. Гіпотези про середнє значення при відомій дисперсії . . . . .             | 101 |
| 9.2. Гіпотези про середнє при невідомій дисперсії . . . . .                    | 102 |
| 9.3. Гіпотези про значення дисперсії . . . . .                                 | 102 |
| 9.4. Критерій Фішера . . . . .                                                 | 103 |
| 9.5. Критерій Стюдента . . . . .                                               | 106 |
| <i>Контрольні запитання . . . . .</i>                                          | 108 |
| <i>Вправи та задачі . . . . .</i>                                              | 109 |
| Додаток . . . . .                                                              | 109 |
| Коефіцієнти асиметрії $A$ та ексцесу $E$ основних розподілів . . . . .         | 110 |
| Квантилі стандартного розподілу $N(0, 1)$ . . . . .                            | 110 |
| Квантилі $t_{1-\alpha/2}(n)$ розподілів Стюдента $T(n)$ . . . . .              | 111 |
| Розв'язки $q(\gamma, n)$ інтегральних рівнянь розподілів $\chi(n-1)$ . . . . . | 112 |
| Верхні $\alpha$ -границі розподілів $\chi^2(n)$ . . . . .                      | 113 |
| Верхні $\alpha$ -границі розподілів Фішера $F(n, m)$ . . . . .                 | 115 |

# Передмова

На третьому курсі навчання студенти спеціальності 111 "Математика" вивчають дисципліну "Теорія ймовірностей та математична статистика", яка розрахована на два семестри (8 кредитів, 240 год). Перший семестр присвячений теорії ймовірностей, а другий – математичній статистиці.

Теорія ймовірностей розробляє математичні моделі стохастичних експериментів, вивчає їх властивості та закономірності. Математична статистика ж відіграє особливу роль, оскільки відкриває можливості розпізнавати ті чи інші ймовірнісні моделі у реальному житті.

Даний навчальний посібник охоплює програмний матеріал з математичної статистики. Перша його частина складається з двох розділів, назви яких "Основи математичної статистики" та "Перевірка статистичних гіпотез" збігаються з назвами змістовних модулів курсу<sup>1</sup>.

У першому розділі посібника розглядаються основні поняття вибіркового методу, точкові та інтервальні оцінки невідомих параметрів розподілу, методи їх здобуття та аналізу. Наведені приклади побудови оцінок з детальним дослідженням властивостей.

У другому розділі вивчаються поняття статистичної гіпотези, види гіпотез, критерії їх перевірки, класифікація помилок прийнятих рішень. Увага зосереджена на видах критеріїв, методах їх побудови та аналізу властивостей. Розглядаються критерії щодо гіпотез про розподіл, незалежність випадкових величин та параметри розподілів.

Наведені поняття, теореми та твердження ілюструються прикладами.

---

<sup>1</sup>У робочій програмі дисципліни це змістовні модулі 5 та 6.

Скрізь подвійна нумерація теорем, прикладів, формул і вправ, що відповідає номерам відповідних параграфів.

Курс "Теорія ймовірностей та математична статистика" є базовим для всіх математичних дисциплін про випадкові явища, які у комплексі звуть *стохастикою*<sup>2</sup>. Зокрема, глибоке розуміння ймовірнісних і статистичних моделей необхідне для подальшого вивчення багатовимірного статистичного аналізу, стохастичної фінансової математики, прикладної статистики, аналізу даних та машинного навчання – дисциплін, що передбачені навчальним планом спеціальності 111 "Математика".

Сподіваємось, посібник буде цікавим та корисним усім, хто глибоко цікавиться статистикою.

---

<sup>2</sup>Стохастичність (від грец. *στόχος* [стохос] – припущення, здогад) означає випадковість.

# Вступ

Математична статистика, яка виникла спочатку для цілей демографії і страхування, перетворилась в один з основних методів кількісного дослідження явищ природи, технічних процесів, економіки й лінгвістики.

Гнеденко Б.В.

Висновки практичних досліджень у будь-якій галузі, як правило, ґрунтуються на фактах, що можна подати у вигляді даних (кількісних або якісних). Просто зібрати дані недостатньо. Щоб вони про щось розповіли, необхідно знати: а) як їх зібрати; б) як їх обробити; в) як проаналізувати та зробити змістовні висновки.

Розділ математики, який розробляє методи збору, обробки та аналізу даних спостережень, називають *математичною статистикою*. Статистичні методи тісно пов'язані з імовірнісними, хоча створені для розв'язання в певному сенсі протилежних задач.

Якщо стохастичний експеримент вкладається в одну з відомих імовірнісних моделей, то ймовірнісні методи дозволяють спрогнозувати його наслідки. Статистичні ж методи спрямовані на те, щоб за результатом стохастичного експерименту підібрати найбільш слушну ймовірнісну модель та оцінити ймовірність можливої помилки. Можна сказати, що математична статистика прокладає "місток" з реального життя у світ теорії ймовірностей.

Наприклад, у разі, коли проводиться п'ять випробувань за схемою



Бернуллі<sup>1</sup> з імовірністю  $P(\text{"успіх"}) = 1/3$  і треба знайти ймовірність того, що три рази відбудеться подія *"успіх"* та два рази *"невдача"*, – це задача теорії ймовірностей.

А як деякий експеримент за схемою Бернуллі після п'яти випробувань закінчився подією *"успіх, успіх, успіх, невдача, невдача"*, ймовірність  $P(\text{"успіх"})$  невідома і треба зрозуміти, яке її значення найбільш правдоподібне для цієї моделі, то це вже задача математичної статистики. Один із способів її розв'язання наведений у п. 3.2, прикл. 3.4.

Наскільки корисним може стати статистично-ймовірнісний аналіз, розглянемо на прикладі такої ситуації: деяке підприємство виробляє прогоничі класу міцності  $5,8^2$  і треба перевірити, чи відповідає його продукція стандарту якості міцності на розтяг, номінальне значення якого становить  $500 \text{ н/мм}^2$ .

Вочевидь, перевірка кожного прогонича економічно недоцільна. Шляхом випадкового вибору візьмемо на складі, наприклад, 30 прогоничів і лабораторним долідженням визначимо для кожного з них міцність на розтяг. Отримаємо 30 значень:

$$x_1, \dots, x_{30} \text{ (н/мм}^2\text{)}. \quad (0.1)$$

Статистичні методи дозволяють на основі аналізу числових даних (0.1) зробити висновок про міцність на розтяг усієї продукції підприємства, оцінити ймовірність можливої помилки (все ж таки не кожен прогонич перевіряли) і спосіб її мінімізації. Так, якщо, знову-таки випадково, візьмемо і проаналізуємо 100 прогоничів, ймовірність зробити помилковий висновок буде меншою.

---

<sup>1</sup>Якоб Бернуллі (27.12.1654, Базель, Швейцарія – 16.08.1705, Базель) – видатний математик, фізик, лікар. Відомий значними результатами в багатьох галузях наукознавства. У теорії ймовірностей вивчав просту, але надто важливу схему незалежних випробувань, яку на його честь називають схемою Бернуллі, і шляхом аналізу її властивостей сформував базові поняття, які і сьогодні є основними в сучасній теорії ймовірностей, довів першу теорему із серії законів великих чисел.

<sup>2</sup>Про маркування класів міцності можна дізнатися, наприклад, на сайті <https://krepsila.com/ua/a457534-klassy-prochnosti-boltov.html>.

Більш загально. Припустимо, що випадково обрали  $n$  одиниць продукції і для кожного виробу визначили граничне значення міцності. Маємо  $n$  чисел  $x_1, x_2, \dots, x_n$ , загалом, різних, за якими й розв'яжемо поставлену задачу.

У термінах математичної статистики всю продукцію відповідного класу на складі називають *генеральною сукупністю*, а  $n$  її представників, випадковим чином обраних для досліджень, – *вибіркою обсягу  $n$* .

Нехай  $\xi$  – функція міцності на розтяг, що задана на елементах генеральної сукупності. Тоді на елементах випадкової вибірки вона має статус випадкової величини, яка в  $n$  експериментах набула значень  $x_1, x_2, \dots, x_n$ .

Відповідь на питання, чи має продукція відповідні стандарти якості, можна отримати, якщо визначити розподіл величини  $\xi$ . Дійсно, згідно з центральною граничною теоремою, добре налагоджений процес виробництва характеризується нормальною випадковою величиною  $\xi \in N(a, \sigma^2)$ . Звичайно, для певних прогоничів показники будуть відрізнятися від номінальних, але якщо  $\xi$  має нормальний розподіл, то така відмінність сумарно обумовлена великою кількістю несуттєвих випадкових чинників, кожен з яких "потроху" впливає на результат. Це природно для будь-якого процесу і є нормальним. Якщо ж і  $a \approx 500$  н/мм<sup>2</sup>, та ще й  $\sigma^2$  – невелике значення, то на підприємстві взагалі все гаразд. Крім того, за визначенням розподілом з необхідною точністю не складно розрахувати ймовірність  $P\{\xi \approx 500 \text{ н/мм}^2\}$ , котра також надасть інформативну відповідь про виробництво.

Таким чином, аналіз якості виробничого процесу зводиться до питання про розподіл випадкової величини. Зазначимо, що всі відомості про випадкову величину – це скінченна кількість її значень  $x_1, \dots, x_n$  і апріорна інформація щодо неї (яку саме характеристику генеральної сукупності вивчаємо, який розподіл очікуємо, тощо).

Визначення невідомого розподілу – одна з задач математичної статистики і основна (безпосередньо розглядається в пп. 2.5, 7.1 та 7.2). Якщо ж треба знати, чи покращить міцність прогоничів нова технологія виготовлення – то це вже задача про порівняння двох вибірок (див. пп.

9.4, 9.5). А от питання про те, чи залежить міцність прогонича від часу його загартування вирішується шляхом аналізу вибірових даних за двома ознаками: міцність ( $\xi$ ) та час загартування ( $\eta$ ). Спосіб її розв'язання можна знайти в п. 7.3. Дослідження наявності лінійного зв'язку між  $\xi$  та  $\eta$  – це задача кореляційного аналізу, важливого розділу математичної статистики.

Багато різних за своїм змістом як практичних, так і теоретичних питань, що виникають при перевірці наукових тверджень, формуванні нових гіпотез, можна вирішити за допомогою однієї із статистичних моделей: про розподіл, про порівняння вибірок, про параметри розподілу, про незалежність, вид залежності, тощо. Тому так важливо добре розуміти суть та специфіку статистичних методів, основні поняття, на яких вони ґрунтуються, щоб коректно й доцільно їх застосовувати. Мета даного посібника – допомогти в цьому студентам.

Будь-яку статистичну задачу розв'язують за вибірковими даними.

**Означення.** Сукупність однорідних об'єктів, усі елементи якої підлягають вивченню, називають *генеральною сукупністю*.

Множину  $n$  об'єктів, вибраних *випадковим чином* з генеральної сукупності, називають *вибіркою обсягу  $n$* .

Умова випадкового вибору обов'язкова, оскільки забезпечує репрезентативність вибірки і є основною передумовою статистичного аналізу.

*Репрезентативність* вибірки означає здатність вибіркової сукупності відтворювати характеристики генеральної сукупності.

Основні етапи статистичного аналізу:

- визначення генеральної сукупності відносно задачі, що розглядається;
- формування репрезентативної вибірки обсягу  $n$ ;
- аналіз вибірки і створення на його основі припущень (гіпотез) щодо генеральної сукупності;
- перевірка узгодженості гіпотез з експериментальними даними;
- оцінка ймовірності хибного рішення.

Досить часто ознака генеральної сукупності вимірюється числом (є кількісною) і на елементах вибірки її можна вважати випадковою вели-

чиною, як у прикладі про прогоничі. Числові значення такої величини на елементах вибірки називають *вибірковими даними*. Статистичний аналіз саме кількісних ознак розглядається в першій частині посібника. Його перший розділ охоплює базові поняття статистики, а у другому вивчаються способи розв'язання задач про розподіл, про порівняння вибірок, про параметри розподілу, про незалежність, важливість яких обговорювалась вище. Дисперсійний, кореляційний та регресійний аналіз, а також деякі рангові критерії – матеріал другої частини посібника.

Посібник теоретично спрямований для студентів математичних спеціальностей. Для повного розуміння матеріалу необхідні глибокі знання з теорії ймовірності.

### **Контрольні запитання:**

1. Які задачі математичної статистики?
2. Що називають генеральною сукупністю, вибіркою обсягу  $n$ , вибірковими даними?
3. У чому полягає основна ідея математичної статистики? Схема статистичних міркувань?
4. Охарактеризуйте взаємозв'язок теорії ймовірностей та математичної статистики.
5. Сформулюйте, будь ласка, приклади різних задач (маркетингу, екології, геології, біології, медицини, економіки, комп'ютерних наук тощо), з математичної точки зору близьких трьом задачам про прогоничі, наведеним вище. Опишіть відповідні генеральні сукупності та спосіб збору вибірових даних. Подібна практика буде корисна і надалі. <sup>3</sup>

---

<sup>3</sup>Статистика пізнається на прикладах.

# 1. ОСНОВИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

Уся глибина думки, закладена у формулювання математичних понять, згодом розкривається тим умінням, із яким ці поняття використовуються.

Е. Вігнер

## §1. Первинна обробка вибірових даних

Експеримент – справжній посередник між людиною та природою.

Леонардо да Вінчі

### 1.1. Ази статистики

Припустимо, що вивчається деяка кількісна ознака  $\xi$  генеральної сукупності шляхом аналізу репрезентативної вибірки обсягу  $n$ . На елементах вибірки доцільно вважати  $\xi$  випадковою величиною, що спостерігається шляхом проведення  $n$  незалежних експериментів. Вибіркові дані

$$x_1, x_2, \dots, x_n \tag{1.1}$$

– її значення після першого, другого, ...,  $n$ -го випробування.

Розподіл  $P_\xi$  невідомий і числа (1.1) – це все, що ми маємо, щоб розв'язати основну задачу математичної статистики: знайти невідому функцію

розподілу  $F_\xi(x)$ . Як ми незабаром зрозуміємо, "знайти" – сміливе слово, його треба замінити на "оцінити" і додати "з деякою ймовірністю".

Спершу познайомимось із базовими поняттями та методами математичної статистики, а безпосередньо до розв'язання задачі про визначення невідомого розподілу перейдемо лише в другому розділі посібника, п. 7.1 та 7.2 (див. також п. 2.5).

Почнемо з наступного важливого моменту. Вибіркові дані (1.1) після проведення  $n$  незалежних спостережень – це набір чисел. До початку випробувань невідомо, якими саме вони будуть. Але ми знаємо, що в кожному експерименті очікуваним є значення саме випадкової величини  $\xi$ . Тому в статистиці заведено вважати, що до проведення  $i$ -того експерименту  $x_i$  – випадкова величина, розподілена так само як і  $\xi$ , а після – число, що спостерігається в  $i$ -тому випробуванні,  $i = \overline{1, n}$ . Незважаючи на однакові позначення, з контексту завжди зрозуміло, про що йдеться.

**Означення.** Випадковий вектор  $(x_1, \dots, x_n)$ , компоненти якого – незалежні випадкові величини, розподілені однаково з  $\xi$ , називають *вибіркою обсягу  $n$  з розподілу  $P_\xi$* .

Числовий вектор  $(x_1, \dots, x_n)$ , що стає відомим після  $n$  випробувань, називають *реалізацією вибірки*. Через  $X \subseteq \mathbb{R}^n$  позначають множину всіх можливих реалізацій вибірки і називають *вибірковим простором*.

Якщо елементи реалізації вибірки впорядкувати за зростанням, то отримаємо набір числових значень:

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}, \quad (1.2)$$

котрий називають *варіаційним рядом*, а самі  $x_{(i)}$  – *варіантами*,  $i = \overline{1, n}$ . Вочевидь,  $x_{(1)} = \min\{x_1, \dots, x_n\}$ ,  $x_{(n)} = \max\{x_1, \dots, x_n\}$ .

Число  $v = x_{(n)} - x_{(1)}$  називають *розмахом* вибірки.

Серед варіант в (1.2) можуть бути однакові. Нехай  $x_1^*, \dots, x_l^*$  – ті з них, які різні, причому

$$x_1^* < x_1^* < \dots < x_l^*,$$

а  $m_i$  – кількість варіант в (1.2), що збігаються з  $x_i^*$ ,  $i = \overline{1, l}$ . Число  $m_i$

називають *частотою*  $i$ -тої варіанти. Вочевидь,

$$\sum_{i=1}^l m_i = n.$$

Відношення частоти  $m_i$  до обсягу вибірки  $n$  називають *відносною частотою* варіанти  $x_i^*$  і позначають:

$$w_i = \frac{m_i}{n}, \quad \sum_{i=1}^l w_i = 1.$$

Варіанту з найбільшою частотою називають *модой*  $Mo$  вибірки:

$$Mo = x_j^* \Leftrightarrow m_j = \max\{m_1, \dots, m_l\}.$$

Визначають також *медіану* вибірки  $Me$  за правилом:

$$Me = \begin{cases} x_{(k+1)}, & \text{якщо } n = 2k + 1, \\ (x_k + x_{(k+1)}) / 2, & \text{якщо } n = 2k. \end{cases}$$

## 1.2. Вибірковий розподіл

Про невідомий розподіл  $P_\xi$ , його функцію розподілу  $F_\xi$ , відповідні числові характеристики  $M\xi$ ,  $D\xi$ ,  $\alpha_k = M\xi^k$ ,  $\beta_k = M(\xi - M\xi)^k$  будемо говорити як про *істинні* або *теоретичні* і мати на меті знайти для них відповідні оцінки за вибірковими даними.

Нехай  $(\Omega, \mathfrak{A}, P)$  – імовірнісний простір, за ймовірнісною мірою якого визначаються теоретичні моменти випадкової величини  $\xi : \Omega \rightarrow \mathbb{R}$ . Тоді  $n$  незалежним спостереженням випадкової величини  $\xi$  відповідає прямий добуток імовірнісних просторів  $(\Omega \times \dots \times \Omega, \mathfrak{A} \times \dots \times \mathfrak{A}, P \times \dots \times P)$  і вибірка обсягу  $n$  з розподілу  $P_\xi$  – це випадковий вектор

$$(x_1, \dots, x_n) : \Omega \times \dots \times \Omega \rightarrow X \subseteq \mathbb{R}^n.$$

На одній з елементарних подій  $\omega_0 = \omega_0^1 \times \dots \times \omega_0^n$ , що є наслідком певних  $n$  незалежних випробувань, маємо реалізацію вибірки –  $n$ -вимірний числовий вектор

$$(x_1 = x_1(\omega_0), \dots, x_n = x_n(\omega_0)). \quad (1.3)$$

В іншому ймовірнісному просторі  $(\bar{\Omega}, \bar{\mathfrak{A}}, \bar{P})$  розіграємо нову випадкову величину  $\bar{\xi}$ , яка приймає значення  $x_1, \dots, x_n$  з однаковими ймовірностями:

$$\bar{P}\{\bar{\xi} = x_i\} = \frac{1}{n}, \quad i = \overline{1, n}.$$

Нехай  $x_1^*, x_2^*, \dots, x_l^*$  – різні компоненти вектора (1.3) з абсолютними частотами  $m_1, m_2, \dots, m_l$ , відповідно. Закон розподілу випадкової величини  $\bar{\xi}$  подамо таблицею

|             |                 |                 |         |                 |
|-------------|-----------------|-----------------|---------|-----------------|
| $\bar{\xi}$ | $x_1^*$         | $x_2^*$         | $\dots$ | $x_l^*$         |
| $\bar{P}$   | $\frac{m_1}{n}$ | $\frac{m_2}{n}$ | $\dots$ | $\frac{m_l}{n}$ |

(1.4)

і, як зазвичай для дискретної випадкової величини, запишемо її функцію розподілу:

$$F_{\bar{\xi}}(t) = \sum_{i: x_i^* < t} \frac{m_i}{n} = \frac{n_t}{n} \equiv \bar{F}_n(t), \quad \forall t \in \mathbb{R}. \quad (1.5)$$

Через  $n_t$  в формулі (1.5) позначили кількість варіант, не більших за  $t$ .

Таблицю (1.4) називають *вибірковим* або *емпіричним* розподілом, функцію  $\bar{F}_n(t)$  – *вибірковою* або *емпіричною функцією розподілу*, а числові характеристики випадкової величини  $\bar{\xi}$  (математичне сподівання, дисперсію, моменти) – *емпіричними* або *вибірковим* характеристиками. Емпіричну функцію розподілу називають ще *функцією накопичення відносних частот*.

Емпіричні характеристики позначають так:

$$M\bar{\xi} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^l m_i x_i^* \equiv \bar{x}, \quad (1.6)$$

$$D\bar{\xi} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \frac{1}{n} \sum_{i=1}^l m_i (x_i^*)^2 - \bar{x}^2 \equiv \bar{S}^2, \quad (1.7)$$

$$M\bar{\xi}^k = \frac{1}{n} \sum_{i=1}^n x_i^k = \frac{1}{n} \sum_{i=1}^l m_i (x_i^*)^k \equiv \bar{a}_k, \quad k \in \mathbb{N}, \quad (1.8)$$

$$M(\bar{\xi} - M\bar{\xi})^k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k = \frac{1}{n} \sum_{i=1}^l m_i (x_i^* - \bar{x})^k \equiv \bar{b}_k, \quad k \in \mathbb{N}. \quad (1.9)$$



Читають:  $\bar{x}$  – *емпіричне* або *вибіркове* середнє значення,  $\bar{S}^2$  – *емпірична* або *вибіркова* дисперсія,  $\bar{a}_k$  – *k-тий емпіричний* або *вибірковий* початковий і  $\bar{b}_k$  – *k-тий емпіричний* або *вибірковий* центральний моменти.

Емпіричне середнє значення, мода та медіана характеризують локацію вибіркового розподілу і, як побачимо в подальшому, теоретичного також. Розмах вибірки, емпірична дисперсія – то емпіричні міри розсіювання.

Розглядають також *емпіричне* або *вибіркове* середньоквадратичне відхилення  $\sqrt{\bar{S}^2} = \bar{S}$  та *емпіричні* або *вибіркові* коефіцієнти асиметрії та ексцесу:

$$\bar{A} = \bar{b}_3 / \bar{S}^3, \quad \bar{E} = \bar{b}_4 / \bar{S}^4 - 3. \quad (1.10)$$

– емпіричні характеристики форми.

На відміну від емпіричної дисперсії  $\bar{S}^2$ , яка вимірюється в квадратних одиницях теоретичної випадкової величини, одиниці вимірювання емпіричного середньоквадратичного відхилення  $\bar{S}$  збігаються з одиницями вимірювання  $\xi$ .

Величини  $\bar{A}$  та  $\bar{E}$  безрозмірні, тому і називаються *коефіцієнтами*.

### 1.3. Інтервальний ряд

У сенсі питання, що вивчається за вибіркою, зрозуміло, яку саме випадкову величину  $\xi$  розглядаємо, дискретну чи неперервну. Наприклад, якщо йдеться про "кількість випадків", "кількість одиниць" і таке інше, то очікуємо дискретну випадкову величину. Якщо ж  $\xi$  – "міцність на розтяг", "час реакції", "температура", "прибуток компанії" і т. і., то – неперервну.

Різниця у тому, що множина значень дискретної випадкової величини дискретна, а для неперервної – це деякий проміжок  $E$  і у будь-якому як завгодно малому інтервалі  $(a, b) \subset E$  міститься нескінченна кількість її значень.

Припустимо, що спостережна випадкова величина  $\xi$  абсолютно неперервна і (1.1) – вибірка з невідомого розподілу  $P_\xi$ . Для вирішення основної задачі математичної статистики за скінченною кількістю значень (1.1) неперервної величини необхідно відновити її розподіл. Це досить не просте питання. Побудова емпіричної щільності розглядається в наступному

пункті, а зараз познайомимось з необхідними для того поняттями.

По-перше, зазначимо, що, задля практичних застосувань, неперервну величину завжди квантують. Так називають наближення квантованою величиною, яка визначається скінченним числом зазвичай рівних (але не обов'язково) інтервалів  $(a_i, b_i)$  і в середині цих проміжків є сталою.

Наприклад, щось вимірюємо ми в будь-якому разі лише з точністю до абсолютної похибки  $\Delta x$ . Якщо вимірювання в мм, то  $\Delta x = 0,5$  мм і за умови, що результати належать проміжку  $[5 \text{ мм}, 8 \text{ мм}]$ , маємо інтервали квантування (в мм):

$$[4, 5; 5, 5), \quad [5, 5; 6, 5), \quad [6, 5; 7, 5), \quad [7, 5; 8, 5).$$

Кінець одного проміжку квантування завжди є початком наступного і об'єднання таких проміжків містить усі результати спостережень.

За аналогією квантуємо неперервну випадкову величину  $\xi$ , задану своїми значеннями (1.1). Якщо  $n \leq 10$  і відстані між варіантами усі однакові:

$$x_i - x_{i-1} = \Delta, \quad i = \overline{1, n},$$

то інтервали квантування побудуємо, як при вимірюванні в мм:

$$(x_1 - \Delta/2, x_1 + \Delta/2), \quad (x_1 + \Delta/2, x_2 + \Delta/2), \dots, (x_n - \Delta/2, x_n + \Delta/2).$$

У загальному випадку оберемо інтервал  $(a, b)$ , якому належать усі вибіркові дані (1.1), і розділимо його на  $s = s(n)$  не обов'язково рівних частин. Кількість проміжків квантування  $s(n)$  розрахуємо за правилом Стерджеса<sup>1</sup>:

$$s = 1 + \log_2 n. \tag{1.11}$$

Квантовану випадкову величину  $\xi^*$  задамо проміжками

$$(a_0, a_1), \quad [a_1, a_2), \dots, [a_{s-1}, a_s), \tag{1.12}$$

які назвемо *проміжками групування* ( $a_0 = a, a_s = b$ ).

---

<sup>1</sup>Герберт Стерджес (1882, Огайо, США - 1958, Санта-Рита, США) – математик та статистик.

Нехай  $m_i$  – кількість варіант (1.1), що потрапили в  $i$ -тий проміжок,  $i = \overline{1, s}$ . Розподіл частот випадкової величини  $\xi^*$  характеризує таблиця

|         |              |              |         |                  |
|---------|--------------|--------------|---------|------------------|
| $\xi^*$ | $(a_0, a_1)$ | $[a_1, a_2)$ | $\dots$ | $[a_{s-1}, a_s)$ |
| $m$     | $m_1$        | $m_2$        | $\dots$ | $m_s$            |

(1.13)

яку називають *інтервальним рядом*.

За інтервальним рядом побудуємо розподіл дискретної величини  $\bar{\xi}$ , значення якої в проміжках групування є сталими і збігаються з серединами цих проміжків:

|             |                     |                     |         |                         |
|-------------|---------------------|---------------------|---------|-------------------------|
| $\bar{\xi}$ | $\frac{a_0+a_1}{2}$ | $\frac{a_1+a_2}{2}$ | $\dots$ | $\frac{a_{s-1}+a_s}{2}$ |
| $\bar{P}$   | $\frac{m_1}{n}$     | $\frac{m_2}{n}$     | $\dots$ | $\frac{m_s}{n}$         |

(1.14)

Зазначимо, що розбиття на проміжки можна проводити за одним з двох принципів:  $[a_i, a_{i+1})$  або  $(a_i, a_{i+1}]$ . Відповідні дискретні розподіли можуть відрізнятися зсувом, і що більше "межових" значень  $a_i$ ,  $i = \overline{1, s-1}$  серед вибірових даних, то більший зсув.<sup>2</sup> виправити ситуацію можна шляхом вибору інтервалів. Треба враховувати, що число "межових" значень проміжків групування серед варіант в (1.1) повинно бути мінімальним. Найкращим буде такий вибір, коли проміжки групування можна звести до інтервалів і не втратити при цьому жодного вибіркового значення. Тобто коли кожна варіанта міститься в *середині* якогось із проміжків.

Інтервальний ряд (1.13) якнайкраще відповідає вибіровим даним (1.1) тоді, коли середнє арифметичне, розраховане за таблицею (1.14), майже не відрізняється від середнього

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

розрахованого за реалізацією вибірки. Щоб досягти цього, інтервали будують так, щоб середнє значення одного з них збігалось з середнім  $\bar{x}$ .

---

<sup>2</sup> А ймовірність влучити в точку для неперервної випадкової величини дорівнює нулеві.

#### 1.4. Емпірична щільність

Нехай (1.1) – вибірка з розподілу  $P_\xi$ , де  $\xi$  – абсолютна неперервна випадкова величина з щільністю  $p_\xi(t)$ . Оберемо проміжок  $(a, b)$ , що містить усі значення  $x_i$ ,  $i = \overline{1, n}$ , і побудуємо відповідний інтервальний ряд (1.13) та дискретний емпіричний розподіл (1.14). Далі за емпіричним розподілом (1.14), як і в (1.5), запишемо емпіричну функцію розподілу:

$$\overline{F}_n(t) = \frac{n_t}{n}, \quad t \in \mathbb{R}.$$

У цій формулі  $n_t$  – кількість елементів множини

$$\left\{ \frac{a_0 + a_1}{2}, \dots, \frac{a_{s-1} + a_s}{2} \right\},$$

не більших за  $t$  для кожного  $t \in \mathbb{R}$ .

Оскільки щільність та функцію розподілу пов'язує співвідношення  $p_\xi(t) = F'_\xi(t)$  м.с., то емпіричним аналогом щільності можна вважати функцію

$$\bar{p}_n(t) = \begin{cases} 0, & \text{якщо } t \notin (a_0, a_s), \\ \frac{\overline{F}_n(a_{i-1}) - \overline{F}_n(a_i)}{a_i - a_{i-1}} = \frac{m_i}{n(a_i - a_{i-1})} & \text{при } t \in (a_{i-1}, a_i), \quad i = \overline{1, s}. \end{cases}$$

Графік емпіричної щільності  $\bar{p}_n(t)$  називають *гістограмою відносних частот* або просто *гістограмою*.

Часто *гістограмою* ще називають фігуру, що утворює об'єднання прямокутників з основами на проміжках  $(a_{i-1}, a_i)$  і висотами  $p_i = m_i$ , відповідно,  $i = \overline{1, s}$ .

Гістограма частот – один із способів графічного зображення вибіркового даних, важливий в імовірнісному сенсі, бо дозволяє отримати уявлення про теоретичну щільність. Розглядають також *діаграми*, *полігони* або *многокутники частот*, *секторні*, *променеві діаграми* та ін. (див. про те [5, с. 304]). Діаграми, полігони та многокутники частот корисні для визначення аномальних вибірових значень. Так називають варіанти, які суттєво відрізняються від інших в реалізації.

Аномальне вибірове значення може з'являтися у двох випадках: якщо воно належить генеральній сукупності і є досить малоімовірним або коли

має місце випадкове порушення умов експерименту. Про спосіб видалення аномальних варіант див. [3, с. 220].

**Контрольні запитання:**

1. Що називають реалізацією вибірки, варіаційним рядом, варіантою, розмахом вибірки, частотою та відносною частотою варіанти, модою та медіаною вибірки?

2. У чому полягає відмінність між вибіркою обсягу  $n$  з розподілу  $P_\xi$  і реалізацією вибірки?

3. Як за варіаційним рядом будують вибірковий розподіл, вибіркову функцію розподілу?

4. У чому полягає відмінність між істинною та вибірковою функціями розподілу?

5. Як визначають емпіричні аналоги основних імовірнісних характеристик випадкової величини?

6. Що називають інтервальним рядом і яка характеристика його найкращої відповідності вибірковим даним?

7. Яка відмінність між гістограмою та гістограмою відносних частот? Назвіть, будь ласка, імовірнісний аналог останньої.

**Вправи та задачі: 1.1.** Нехай  $(x_1, \dots, x_n)$  – вибірка обсягу  $n$ .

У разі, коли вибіркові дані скупчені у деякому проміжку (наприклад, досить малі за модулем з великою кількістю значущих цифр після коми) або навпаки, розріджених на числовій осі, суттєво скоротити час при обчисленні емпіричних характеристик може лінійне перетворення вибірових даних за формулою

$$y_i = cx_i + d, \quad c, d \in \mathbb{R}, \quad i = \overline{1, n}. \quad (1.15)$$

Знайдіть, будь ласка, співвідношення між вибірковими моментами, розрахованими за вибіркою  $(x_1, \dots, x_n)$ , і вибірковими моментами, розрахованими за її лінійним перетворенням  $(y_1, \dots, y_n)$ . Переконайтеся, що асиметрія та ексцес інваріантні відносно перетворення (1.15) (відповіді можна звірити в [5, с. 326]).

## §2. Оцінки та їх властивості

### 2.1. Види статистичних оцінок

Ми розглянули різні вибіркові характеристики невідомого розподілу  $P_\xi$ : емпіричну функцію розподілу  $\bar{F}_n(t)$ , у разі абсолютно неперервного розподілу – емпіричну щільність  $\bar{p}_n(t)$ , розмах вибірки  $v$ , моду  $Mo$ , медіану  $Me$ , емпіричні моменти  $\bar{x}$ ,  $\bar{S}^2$ ,  $\bar{a}_k$ ,  $\bar{b}_k$ ,  $k \in N$ ; коефіцієнти асиметрії  $\bar{A}$  та ексцесу  $\bar{E}$ .

На числовому векторі  $x = (x_1(\omega_0), \dots, x_n(\omega_0))$ , тобто на конкретній реалізації,  $\bar{F}_n(t)$ ,  $\bar{p}_n(t)$  – числові функції, а інші емпіричні характеристики – деякі числа. Але всі вони змінюються від вибірки до вибірки і на випадковому векторі  $x = (x_1(\omega), \dots, x_n(\omega))$ , вочевидь,  $\bar{F}_n(t)$ ,  $\bar{p}_n(t)$  – випадкові функції, а  $v$ ,  $Mo$ ,  $Me$ ,  $\bar{x}$ ,  $\bar{S}^2$ ,  $\bar{a}_k$ ,  $\bar{b}_k$ ,  $k \in N$ ;  $\bar{A}$  та  $\bar{E}$  – випадкові величини (як борелівські функції від випадкових величин).

**Означення.** Будь-яку борелівську функцію  $\gamma(x_1, \dots, x_n)$ , що залежить лише від випадкового вектора  $(x_1, \dots, x_n)$ , називають *статистикою*.

Таким чином, усі емпіричні характеристики, розглянуті вище, – деякі статистики.

За означенням,  $\gamma(x_1, \dots, x_n)$  – випадкова величина, а, враховуючи залежність від  $n$ , маємо послідовність випадкових величин. Нас будуть цікавити не будь-які статистики, а тільки ті, які близькі у певному розумінні до відповідних істинних характеристик. А саме, в сенсі збіжності за ймовірністю та майже напевне.

**Означення.** Статистику  $\gamma(x_1, \dots, x_n)$  називають *спроможною* (конзистентною) *оцінкою* параметра  $\theta$  або просто *оцінкою*, якщо

$$\gamma(x_1, \dots, x_n) \xrightarrow{P} \theta, \quad \text{коли } n \rightarrow \infty.$$

Спроможну оцінку параметра  $\theta$  будемо позначати  $\hat{\theta}$ . За означенням,

$$\forall \varepsilon > 0 \quad P\{\omega : |\hat{\theta}(x_1, x_2, \dots, x_n) - \theta| \leq \varepsilon\} \rightarrow 1, \quad \text{коли } n \rightarrow \infty.$$

Вимога спроможності ставиться до всіх оцінок, що використовуються.

У випадку, коли

$$\widehat{\theta}(x_1, \dots, x_n) \xrightarrow{\text{м.н.}} \theta \quad \text{при } n \rightarrow \infty,$$

оцінку  $\widehat{\theta}(x_1, \dots, x_n)$  називають *сильно спроможною*.

Корисною також є така властивість незміщеності.

**Означення.** Оцінку  $\widehat{\theta}$  називають *незміщеною* оцінкою параметра  $\theta$ , якщо для будь-якого  $n \in \mathbb{N}$  маємо:  $M\widehat{\theta}(x_1, \dots, x_n) = \theta$ .

Таким чином, для незміщеної оцінки відхилення  $\widehat{\theta} - \theta$  мають несистематичний характер, бо  $M(\widehat{\theta} - \theta) = M\widehat{\theta} - \theta = \theta - \theta = 0$ .

Наприклад, за умови, що випадкова величина  $\xi$  має скінченне математичне сподівання ( $M|\xi| < +\infty$ ), середнє емпіричне  $\bar{x}$  є сильно спроможною незміщеною оцінкою для параметра  $\theta = M\xi$ , оскільки за посиленням законом великих чисел Колмогорова

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \xrightarrow{\text{м.н.}} M\xi = \theta$$

як середнє значення незалежних однаково розподілених випадкових величин зі скінченним математичним сподіванням. І, крім того,

$$M\bar{x} = \frac{1}{n} \sum_{i=1}^n Mx_i = \frac{1}{n} \cdot n \cdot M\xi = M\xi = \theta.$$

## 2.2. Властивості емпіричної функції розподілу

Невідомий розподіл  $P_\xi$  повністю описує його функція розподілу  $F_\xi(t)$ , означена для будь-якого  $t \in \mathbb{R}$ . Покажемо, що оцінкою для  $F_\xi(t)$  є емпірична функція  $\bar{F}_n(t)$ , визначена за вибіркою обсягу  $n$  формулою (1.5).

На конкретній реалізації вибірки  $\bar{F}_n : \mathbb{R} \rightarrow [0, 1]$  – східчата функція зі стрибками в точках, що відповідають різним значенням варіант  $x_i^*$ . Величина стрибка в точці  $x_i^*$  дорівнює відносній частоті  $w_i = m_i/n$ .

Коли ж  $(x_1, \dots, x_n)$  – вибірка обсягу  $n$  з розподілу  $P_\xi$ , тобто випадковий вектор, то  $\bar{F}_n : \mathbb{R} \times \Omega \rightarrow [0, 1]$  – випадкова функція, а за фіксованого  $t \in \mathbb{R}$  – випадкова величина.

Поведінку емпіричної функції розподілу  $\bar{F}_n(t)$  у відношенні до теоретичної  $F_\xi(t)$  характеризують дві теореми.

**Теорема 2.1.** *Нехай  $x_1, \dots, x_n$  – вибірка з розподілку  $P_\xi$ ,  $F_\xi(t)$  – теоретична і  $\bar{F}_n(t)$  – емпірична функції розподілу. Тоді*

$$\bar{F}_n(t) \xrightarrow{\mathcal{M.H.}} F_\xi(t), \quad \text{коли } n \rightarrow \infty \text{ при будь-якому } t \in \mathbb{R}.$$

**Доведення.** Нехай  $t \in \mathbb{R}$  – деяке фіксоване число. Запишемо емпіричну функцію розподілу у вигляді

$$\bar{F}_n(t) = \frac{n_t}{n} = \frac{1}{n} \sum_{i=1}^n I_{\{x_i < t\}}, \quad (2.1)$$

де  $n_t$ , як і в формулі (1.5), – це кількість варіант, не більших за  $t$ , а індикатори визначаються рівністю

$$I_{\{x_i < t\}} = \begin{cases} 1, & x_i < t; \\ 0, & x_i \geq t. \end{cases}$$

Кожен з індикаторів  $I_{\{x_i < t\}} \in B(1, p)$  – бернуллівська випадкова величина з параметром

$$p = MI_{\{x_i < t\}} = MI_{\{\xi < t\}} = P\{\xi < t\}, \quad i = \overline{1, n}$$

і тоді  $\bar{F}_n(t)$  – середнє значення  $n$  незалежних однаково розподілених бернуллівських випадкових величин зі скінченними математичними сподіваннями.

Згідно з посиленням законом великих чисел Колмогорова, для будь-якого  $t \in \mathbb{R}$

$$\bar{F}_n(t) = \frac{1}{n} \sum_{i=1}^n I_{\{x_i < t\}} \xrightarrow{\mathcal{M.H.}} P\{\xi < t\} = F_\xi(t), \quad n \rightarrow \infty. \quad \square$$

Таким чином, для будь-якого  $t \in \mathbb{R}$  послідовність випадкових значень емпіричної функції розподілу  $\bar{F}_n(t)$  збігається майже напевне до значення істинної функції розподілу  $F_\xi(t)$  в кожній точці  $t \in \mathbb{R}$ . Тобто,  $\bar{F}_n(t)$  є сильно спроможною оцінкою істинної функції  $F_\xi(t)$  для будь-якого  $t \in \mathbb{R}$ .



Наступна теорема свідчить про рівномірний характер збіжності.

**Теорема 2.2 (Глівенко-Кантеллі).** *За умов теореми 2.1,*

$$\sup_{t \in \mathbb{R}} |\bar{F}_n(t) - F_\xi(t)| \xrightarrow{M.H.} 0, \quad \text{коли } n \rightarrow \infty.$$

Теорему Глівенко<sup>1</sup>-Кантеллі<sup>2</sup> називають фундаментальним результатом статистики. Вона означає, що невідомий розподіл кількісної ознаки генеральної сукупності (дискретний чи абсолютно неперервний) можна відновити за вибірковими даними великого обсягу.

Важливі ще такі властивості емпіричної функції розподілу.

**Теорема 2.3.** *За умов теореми 2.1, для будь-якого  $t \in \mathbb{R}$*

- 1)  $\bar{F}_n(t)$  – незміщена оцінка для  $F_\xi(t)$ ;
- 2)  $D\bar{F}_n(t) = F_\xi(t)(1 - F_\xi(t))/n$ ;
- 3) якщо  $F_\xi(t) \neq 0$  і  $F_\xi(t) \neq 1$ , то

$$\sqrt{n} (\bar{F}_n(t) - F_\xi(t)) \xrightarrow{d} \eta(t), \quad \text{де } \eta(t) \in N(0, F_\xi(t)(1 - F_\xi(t))).$$

**Доведення.** 1) Оскільки  $I_{\{x_i < t\}} \in B(1, p)$ ,  $i = \overline{1, n}$  – однаково розподілені бернуллівські величини з параметром  $p = F_\xi(t)$ , то

$$M\bar{F}_n(t) = M \frac{1}{n} \sum_{i=1}^n I_{\{x_i < t\}} = M I_{\{\xi < t\}} = F_\xi(t).$$

Що і означає незміщеність за означенням.

- 2) Випадкові величини  $I_{\{x_i < t\}}$ ,  $i = \overline{1, n}$  незалежні. Тому

$$D\bar{F}_n(t) = \frac{1}{n^2} \sum_{i=1}^n D I_{\{x_i < t\}} = \frac{D I_{\{\xi < t\}}}{n} = \frac{F_\xi(t)(1 - F_\xi(t))}{n}.$$

- 3) Нехай  $F_\xi(t) \neq 0$ . За центральною граничною теоремою,

$$\sqrt{n} (\bar{F}_n(t) - F_\xi(t)) = \left( \sum_{i=1}^n I_{\{x_i < t\}} - n F_\xi(t) \right) / \sqrt{n} =$$

<sup>1</sup>Валерій Іванович Глівенко (21.12.1896, Київ – 15.02.1940, Москва) – математик, логік. Отримав значні результати в терії функцій дійсного змінного, теорії ймовірностей та математичній статистиці, логіці.

<sup>2</sup>Франческо-Паоло Кантеллі (20.12.1875, Палермо, Італія – 21.07.1966, Рим) – математик, статистик, економіст. Відомий роботами з теорії ймовірностей (лема Бореля-Кантеллі, посилений закон великих чисел) та математичної статистики.

$$= \left( \sum_{i=1}^n I_{\{x_i < t\}} - nMI_{\{x_1 < t\}} \right) / \sqrt{n} \xrightarrow{d} \eta(t) \in N(0, F_\xi(t)(1 - F_\xi(t))).$$

Остання властивість також доведена.  $\square$

**Означення.** Оцінку  $\hat{\theta}$  параметра  $\theta$  називають *асимптотично нормальною* з коефіцієнтом  $\sigma^2 > 0$ , якщо

$$(\hat{\theta} - \theta)\sqrt{n} \xrightarrow{d} \eta \quad \text{і} \quad \eta \in N(0, \sigma^2).$$

Твердження 3) теореми 2.3 означає, що  $\bar{F}_n(t)$  – асимптотично нормальна оцінка для  $F_\xi(t)$  з коефіцієнтом  $\sigma^2 = F_\xi(t)(1 - F_\xi(t)) > 0$ , як тільки  $F_\xi(t) \neq 0$  і  $F_\xi(t) \neq 1$ .

### 2.3. Властивості емпіричної щільності

**Теорема 2.4 (Глівенко).** Нехай  $x_1, \dots, x_n$  – вибірка з абсолютно неперервного розподілу  $P_\xi$  і  $\bar{p}_n(t)$  – емпірична щільність, побудована за інтервальним рядом, кількість проміжків групування якого дорівнює  $s(n)$ . Якщо виконується умова

$$\sum_{n=1}^{\infty} \frac{s^4(n)}{n^2} < \infty, \quad (2.2)$$

то

$$\sup_{t \in \mathbb{R}} |\bar{p}_n(t) - p_\xi(t)| \xrightarrow{P} 0, \quad \text{коли } n \rightarrow \infty.$$

Теорема 2.4 свідчить про те, що за умови вибору кількості проміжків групування так, щоб була гарантована збіжність числового ряду (2.2), емпірична щільність  $\bar{p}_n(t)$  є спроможною оцінкою для істинної щільності  $p_\xi(t)$  і збіжність за ймовірністю рівномірна на  $\mathbb{R}$ .

Легко бачити, що величина  $s(n)$ , розрахована за правилом Стердже-са (1.11), задовольняє вимогу (2.2).

### 2.4. Властивості вибірових моментів

Як вже зазначалось вище, за умови скінченного математичного сподівання ( $M|\xi| < +\infty$ ), середнє емпіричне  $\bar{x}$  є сильно спроможною незміщеною оцінкою для невідомого значення  $M\xi$ .

Наступна теорема встановлює більш загальне твердження.

**Теорема 2.5.** *Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi$ , що задовольняє умову  $M|\xi^k| < +\infty$  для деякого  $k \in \mathbb{N}$ ,  $\alpha_k$  та  $\bar{a}_k$  –  $k$ -ті початкові моменти, теоретичний і емпіричний, відповідно. Тоді*

- 1)  $\bar{a}_k \xrightarrow{M.H.} \alpha_k$ , коли  $n \rightarrow \infty$ ;
- 2)  $M\bar{a}_k = \alpha_k$ ;
- 3) якщо  $0 < D\xi^k < +\infty$ , то

$$\sqrt{n}(\bar{a}_k - \alpha_k) \xrightarrow{d} \eta_k \in N(0, D\xi^k), \quad \text{коли } n \rightarrow \infty.$$

Доведення теореми 2.5 пропонується як вправа.

Таким чином, для будь-якого  $k$  натурального  $k$ -тий початковий емпіричний момент  $\bar{a}_k$  є сильно спроможною незміщеною та асимптотично нормальною оцінкою з коефіцієнтом  $D\xi^k$  для  $k$ -того теоретичного моменту  $\alpha_k$ , якщо  $M|\xi^k| < +\infty$ .

Для перевірки спроможності оцінок часто корисною буває така лема.

**Лема 2.1.** *Якщо послідовності випадкових величин  $\mu_n$  та  $\eta_n$  такі, що*

$$\mu_n \xrightarrow{P} a, \quad \eta_n \xrightarrow{P} b, \quad a, b \in \mathbb{R} \quad (2.3)$$

*і функція  $g : X \subseteq \mathbb{R}^2 \rightarrow Y \subseteq \mathbb{R}$  неперервна в точці  $(a, b) \in X$ , то послідовність  $\{g(\mu_n, \eta_n)\}_{n=1}^\infty$  збігається за ймовірністю до числа  $g(a, b)$ :*

$$g(\mu_n, \eta_n) \xrightarrow{P} g(a, b), \quad n \rightarrow \infty.$$

**Доведення.** За означенням неперервної в точці  $(a, b)$  функції, маємо:

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall (x, y) \in X : |x - a| < \delta \text{ і } |y - b| < \delta \Rightarrow |g(x, y) - g(a, b)| < \varepsilon.$$

Звідси випливає включення множин:

$$\{(\mu_n, \eta_n) : |g(\mu_n, \eta_n) - g(a, b)| \geq \varepsilon\} \subseteq \{\mu_n : |\mu_n - a| \geq \delta\} \cup \{\eta_n : |\eta_n - b| \geq \delta\}.$$

Отже,

$$P\{\omega : |g(\mu_n, \eta_n) - g(a, b)| > \varepsilon\} \leq P\{\omega : |\mu_n - a| \geq \delta \text{ або } |\eta_n - b| \geq \delta\} \leq$$

$$\leq P\{\omega : |\mu_n - a| \geq \delta\} + P\{\omega : |\eta_n - b| \geq \delta\} \rightarrow 0, \quad \text{коли } n \rightarrow \infty,$$

за умовою (2.3).  $\square$

**Теорема 2.6.** *Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi$ , для якого  $D\xi = \sigma^2 < +\infty$  і  $\bar{S}^2$  – вибіркова дисперсія. Тоді*

- 1)  $\bar{S}^2 \xrightarrow{P} D\xi$ , коли  $n \rightarrow \infty$ ;
- 2)  $M\bar{S}^2 = \frac{n-1}{n} D\xi$ .

**Доведення.** 1) Застосуємо лему 2.1 для функції  $f(x, y) = x - y^2$  і випадкових величин

$$\mu_n = \frac{1}{n} \sum_{i=1}^n x_i^2 \xrightarrow{\mathcal{M.H.}} M\xi^2, \quad \eta_n = \bar{x} \xrightarrow{\mathcal{M.H.}} M\xi, \quad \text{коли } n \rightarrow \infty,$$

збіжність майже напевне для яких відома з теореми 2.5. Маємо:

$$\bar{S}^2 = f(\mu_n, \eta_n) \xrightarrow{P} f(M\xi^2, M\xi) = M\xi^2 - (M\xi)^2 = D\xi.$$

2) Розглянемо лінійне перетворення вибірки за формулами

$$y_i = x_i - a, \quad \text{де } a = Mx_i = M\xi, \quad i = \overline{1, n}.$$

Вочевидь,  $My_i = 0$ ,  $Dy_i = Dx_i = D\xi = \sigma^2$  і  $\bar{y} = \bar{x} - a$ . Тоді

$$\bar{S}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2.$$

Крім того, випадкові величини  $y_i$ ,  $i = \overline{1, n}$  попарно незалежні і тому

$$M(y_i y_j) = My_i My_j = 0, \quad i, j = \overline{1, n} \quad i \neq j.$$

Отже,

$$\begin{aligned} M \sum_{i=1}^n (y_i - \bar{y})^2 &= M \left( \sum_{i=1}^n y_i^2 - n\bar{y}^2 \right) = nMy_1^2 - nM\bar{y}^2 = \\ &= n\sigma^2 - n\frac{1}{n^2} M \sum_{i=1}^n y_i^2 = n\sigma^2 - \sigma^2 = (n-1)\sigma^2. \end{aligned}$$

Тобто

$$M\bar{S}^2 = \frac{n-1}{n} \sigma^2. \quad \square$$

**Наслідок.** У випадку, коли  $D\xi = \sigma^2 < +\infty$ , статистика  $\overline{S}^2$  є спроможною, але зміщеною оцінкою істинної дисперсії. Спроможною та незміщеною оцінкою дисперсії є випадкова величина

$$\overline{s}^2 = \frac{n}{n-1} \overline{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2,$$

яку називають *виправленою емпіричною дисперсією* або *варіансою*<sup>3</sup>.

Поправку  $n/(n-1)$  ввів Бессель<sup>4</sup>. За малих обсягів вибірки  $n$  вона суттєва, а от при  $n > 50$  значення  $\overline{s}$  та  $\overline{S}$  практично не відрізняються.

Величину  $\sqrt{\overline{s}^2} = \overline{s}$  називають *флуктуацією*<sup>5</sup>.

**Зауваження 2.1.** У багатьох питаннях цілком достатньо, щоб

$$M\gamma(x_1, \dots, x_n) \rightarrow \theta, \quad \text{коли } n \rightarrow \infty.$$

Таку оцінку називають *асимптотично незміщеною*.

Емпірична дисперсія  $\overline{S}^2$  є прикладом асимптотично незміщеної оцінки.

## 2.5. Формування припущення щодо розподілу

Властивості емпіричної функції розподілу, емпіричної щільності (для абсолютно неперервної ознаки генеральної сукупності), вибірових характеристик, встановлені в теоремах 2.1 – 2.6, свідчать про те, що на основі аналізу значень вибіркової моди  $Mo$ , медіани  $Me$ , середнього  $\bar{x}$ , варіанси  $\overline{s}^2$ , коефіцієнтів  $\overline{A}$  та  $\overline{E}$ , за формою гістограми і певною апіорною інформацією<sup>6</sup> про вид невідомого розподілу можна навести певні науково-обґрунтовані припущення.

Так, якщо  $\bar{x} \approx \overline{s}^2$ , то логічно припустити, що теоретичний розподіл – пуассонівський. Підгрунтя можливої апіорної інформації: це розподіл

<sup>3</sup>Варіанса (від італ. *varianza*) – дисперсія, відхилення, коливання.

<sup>4</sup>Фрідріх-Вільгельм Бессель (22.07.1784, Мінден, Німеччина – 17.03.1846, Кенігсберг, Королівство Пруссія, Німецький союз) – видатний астроном та математик, фізик і геодезист.

<sup>5</sup>Флуктуація (від лат. *fluctuatio*) – хвилювання, безперервний рух.

<sup>6</sup>Йдеться про аналіз задачі, що розглядається, з точки зору теорії ймовірностей, зокрема, граничних теорем.

рідкісних подій (як граничний в сенсі слабкої збіжності для послідовності біноміальних випадкових величин з нескінченно малою імовірністю "успіху").

Для абсолютно неперервних розподілів попередній висновок про вид розподілу можна зробити за формою гістограми та шляхом порівняння вибірових коефіцієнтів асиметрії та ексцесу з відповідними коефіцієнтами основних відомих розподілів (див. [3, с. 223]). Деякі коефіцієнти наведені в табл. I додатку.

З класами розподілів випадкових величин (не всі вивчалися в теорії ймовірностей) і способом їх розпізнавання за видом гістограми можна познайомитись в [3, с. 241 – 273]. Корисно також звернути увагу на те, що у кожного з наведених розподілів є своє теоретичне підґрунтя для створення апріорної інформації.

### Контрольні запитання:

1. Що називають статистикою, оцінкою невідомого параметра розподілу; незміщеною, асимптотично незміщеною оцінкою?
2. Наведіть означення збіжності за ймовірністю, майже напевне та слабкої збіжності. Порівняйте їх силу.
3. Які властивості має емпірична функція розподілу? Емпірична щільність?
4. Як обирають кількість проміжків групування при побудові інтервального ряду і чому саме так?
5. Перечисліть, будь ласка, властивості вибірових моментів.
6. Як за емпіричними характеристиками, гістограмою та апріорною інформацією про випадкову величину, що вивчається, зробити попередні висновки (припущення) щодо невідомого розподілу.

**Вправи та задачі: 2.1.** Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу Бернуллі  $B(0, 1)$  з невідомим параметром  $\theta \in (0, 1)$ . Чи будуть статистики

$$x_1, \quad x_1 x_2, \quad x_1(1 - x_2)$$

незміщеними оцінками, відповідно, для  $\theta$ ,  $\theta^2$ ,  $\theta(1 - \theta)$ ?

**2.2.** Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу Пуассона  $\Pi(\theta)$  з невідомим параметром  $\theta > 0$ . Чи будуть статистики  $x_1, I_{\{x_1=k\}}$  незміщеними оцінками для  $\theta, \theta^k e^{-\theta}/k!$ , відповідно.

Для якого параметра статистика  $\gamma = \bar{x} e^{\bar{x}}$  є спроможною оцінкою?

**2.3.** За умови скінченного  $k$ -того центрального моменту розподілу  $P_\xi$  ( $\beta_k = M|\xi - M\xi|^k < +\infty$ ) для деякого  $k \in \mathbb{N}$ , перевірте спроможність оцінки

$$\bar{b}_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k.$$

**2.4.** Чи будуть статистики  $\bar{A}$  та  $\bar{E}$  спроможними та незміщеними оцінками істинних коефіцієнтів асиметрії та ексцесу у разі скінченних моментів  $\beta_3 < +\infty$  та  $\beta_4 < +\infty$ , відповідно.

### §3. Методи здобуття точкових оцінок параметрів розподілу

Продовжимо вивчення основної задачі математичної статистики про визначення невідомого розподілу  $P_\xi$  за вибіркою. Як зазначалось у п. 2.5, на основі аналізу емпіричних характеристик розподілу і апіорної інформації щодо нього, про вид розподілу можна сформулювати певні припущення і тоді невідомими залишаються його параметри. Зазначимо, що параметрами розподілів є, як правило, їх моменти, початкові або центральні.

Розглянемо два класичні методи побудови оцінок невідомих параметрів розподілу, форма якого заздалегідь визначена. Перший з них – метод моментів, запропонований К. Пірсоном<sup>1</sup> у 1894 р.

#### 3.1. Метод моментів

Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi$ , що залежить від  $s$  невідомих параметрів  $\theta_1, \theta_2, \dots, \theta_d$ . Будемо позначати:

$$P_\xi = P_\xi(\theta), \quad \text{де } \theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d.$$

Від  $\theta$  залежить істинна функція розподілу  $F_\xi(t, \theta)$ , теоретичні початкові та центральні моменти  $\alpha_k(\theta)$  і  $\beta_k(\theta)$ ,  $k \in \mathbb{N}$ . Необхідно оцінити невідомі параметри за вибіркою.

Згідно з теоремами 2.1 та 2.2, емпірична функція розподілу  $\bar{F}_n(t)$  зі збільшенням обсягу вибірки  $n$  майже напевне<sup>2</sup> як завгодно мало відрізняється від істинної функції розподілу  $F_\xi(t, \theta)$ . Як наслідок, емпіричні та теоретичні початкові моменти також при великих обсягах вибірки  $n$  відрізняються несуттєво майже напевне, а центральні – з ймовірністю, близькою до 1 (див. теореми 2.5, 2.6 та задачу 2.3).

Для знаходження оцінок параметрів  $\theta_1, \dots, \theta_d$  скористаємось наближеними рівностями для  $d$  перших істинних та емпіричних початкових

---

<sup>1</sup>Карл Пірсон (27.03.1857, Лондон – 27.04.1936, Лондон) – математик, статистик, біолог, історик, психолог та філософ; один із засновників математичної статистики. Докладав багато зусиль до застосування своїх результатів у прикладних задачах.

<sup>2</sup>Тобто завжди, окрім випадків імовірнісної міри нуль.



моментів:

$$\alpha_k = \alpha_k(\theta_1, \dots, \theta_s) \approx \bar{a}_k \quad \text{м.н. при } n \rightarrow \infty; \quad k = \overline{1, d}. \quad (3.1)$$

Чи можна підібрати оцінки  $\hat{\theta}_1, \dots, \hat{\theta}_d$  так, щоб виконувались точні рівності

$$\alpha_k(\hat{\theta}_1, \dots, \hat{\theta}_d) = \bar{a}_k, \quad k = \overline{1, d}? \quad (3.2)$$

Для цього достатньо розв'язати систему рівнянь (3.2) відносно змінних  $\hat{\theta}_1, \dots, \hat{\theta}_d$ . Розв'язком, як правило, будуть спроможні статистики  $\hat{\theta}_1, \dots, \hat{\theta}_d$  (див. лему 2.1), які називають *оцінками методу моментів* (ОММ) параметрів  $\theta_1, \dots, \theta_d$ , відповідно.

Залежно від кількості невідомих параметрів необхідно використати стільки ж моментів (кількість рівнянь в (3.2) збігається з кількістю невідомих  $d$ ).

**Зауваження 3.1.** Метод моментів містить у собі невизначеність, тому що для пошуку невідомих  $\theta_1, \dots, \theta_d$  можна скористатись як першими  $d$  початковими моментами  $\alpha_k(\theta)$  та  $\bar{a}_k$ ,  $k = \overline{1, d}$ , так і, наприклад, починаючи з  $(d+1)$ -го, тобто для  $k = \overline{d, 2d}$ . Замість початкових можна розглядати центральні моменти  $\beta_k(\theta)$  і  $\bar{b}_k$ ,  $k = \overline{1, d}$  чи  $k = \overline{d, 2d}$ , т.і.

У такий спосіб з'явиться багато різних оцінок одного й того ж параметра (і це добре, оскільки є можливість обирати оцінки з найкращими властивостями).

Розглянемо декілька прикладів застосування методу моментів.

**Приклад 3.1.** Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу Пуассона<sup>3</sup>  $\Pi(\theta)$ , параметр  $\theta > 0$  якого невідомий. Знайдемо ОММ для  $\theta$ .

Перший початковий момент випадкової величини  $\xi \in \Pi(\theta)$  знаємо з теорії ймовірностей:  $\alpha_1(\theta) = \theta$ , а перший емпіричний запишемо за формулою (1.8):  $\bar{a}_1 = \bar{x}$ . Тоді в (3.2) для даного випадку маємо єдине рівняння:

$$\theta = \bar{x}$$

---

<sup>3</sup>Сімеон Дені Пуассон (21.06.1781, Пітів'є, Франція - 25.04.1840, Париж) – видатний математик, астроном, фізик, політик, статистик. У теорії ймовірностей добре відома теорема Пуассона – важливий результат про граничний розподіл у схемі Бернуллі.

і  $\hat{\theta} = \bar{x}$  буде ОММ для  $\theta$ , спроможна та незміщена оцінка.

Якщо ж скористатися рівнянням (3.2) для других моментів

$$\alpha_2(\theta) = \theta + \theta^2 \quad \text{та} \quad \bar{a}_2 = \frac{1}{n} \sum_{i=1}^n x_i^2,$$

яке має вигляд

$$\theta^2 + \theta = \bar{a}_2$$

(квадратне рівняння відносно  $\theta$ ), то його коренями будуть

$$\theta_{1,2} = \frac{-1 \pm \sqrt{1 + 4\bar{a}_2}}{2}.$$

Звідси маємо іншу ОММ  $\hat{\theta} = (\sqrt{1 + 4\bar{a}_2} - 1) / 2$  для  $\theta$  (спроможність даної статистики перевірте, будь ласка, як вправу).

**Приклад 3.2.** Нехай  $(x_1, \dots, x_n)$  – вибірка з нормального розподілу  $P_\xi(\theta) = N(\theta_1, \theta_2^2)$ , параметри якого  $\theta_1 \in \mathbb{R}$  і  $\theta_2 > 0$  невідомі.

Знайдемо ОММ параметрів  $\theta_1$  та  $\theta_2$ .

Для нормального розподілу  $\alpha_1 = M\xi = \theta_1$ , а  $\alpha_2 = M\xi^2 = \theta_2^2 + \theta_1^2$ .

Отже, система (3.2) для цього випадку така:

$$\begin{cases} \theta_1 = \bar{a}_1, \\ \theta_2^2 + \theta_1^2 = \bar{a}_2. \end{cases}$$

Звідси

$$\hat{\theta}_1 = \bar{x}, \quad \hat{\theta}_2^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \bar{S}^2$$

– ОММ невідомих параметрів, обидві статистики спроможні та перша з них незміщена.

**Приклад 3.3.** Припустимо, що  $(x_1, \dots, x_n)$  – вибірка з розподілу Релея<sup>4</sup>  $P_\xi(\theta) = R(\theta)$ , щільність якого

$$p(t, \theta) = \begin{cases} \frac{t}{\theta} \exp\left(-\frac{t^2}{2\theta}\right), & t \geq 0, \\ 0, & t < 0 \end{cases} \quad (3.3)$$

---

<sup>4</sup>Джон Вільям Стретт (лорд Релей) (12.11.1842, Ессекс, Англія – 30.06.1919, Ессекс) – фізик, політик, математик, хімік. Один із засновників теорії коливань.

залежить від одного невідомого параметра  $\theta > 0$ .

Знайдемо ОММ параметра  $\theta$ .

Спочатку обчислимо перший початковий момент розподілу Релея:

$$\begin{aligned}\alpha_1(\theta) &= M\xi = \int_{-\infty}^{+\infty} tp(t, \theta) dt = \int_0^{+\infty} \frac{t^2}{\theta} e^{-\frac{t^2}{2\theta}} dt = \\ &= \int_0^{+\infty} t d\left(-e^{-\frac{t^2}{2\theta}}\right) = -te^{-\frac{t^2}{2\theta}} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-\frac{t^2}{2\theta}} dt = \sqrt{\frac{\pi\theta}{2}}.\end{aligned}$$

Отже, оцінку параметра  $\theta$  знайдемо з рівняння:  $\sqrt{\frac{\pi\theta}{2}} = \bar{a}_1 = \bar{x}$ .

Звідси  $\hat{\theta}_1 = \frac{2}{\pi} (\bar{x})^2$  – ОММ для  $\theta$ .

Якщо ж обчислити другий початковий момент:

$$\begin{aligned}\alpha_2(\theta) &= M\xi^2 = \int_{-\infty}^{+\infty} t^2 p(t, \theta) dt = \int_0^{+\infty} \frac{t^3}{\theta} e^{-\frac{t^2}{2\theta}} dt = \\ &= \int_0^{+\infty} t^2 d\left(-e^{-\frac{t^2}{2\theta}}\right) = -t^2 e^{-\frac{t^2}{2\theta}} \Big|_0^{+\infty} + 2 \int_0^{+\infty} t e^{-\frac{t^2}{2\theta}} dt = \\ &= 2\theta \int_{-\infty}^{+\infty} p(t, \theta) dt = 2\theta\end{aligned}$$

і прирівняти його з другим емпіричним моментом  $\bar{a}_2$ , то отримаємо рівняння

$$2\theta = \bar{a}_2.$$

Тоді

$$\hat{\theta}_2 = \frac{1}{2} \bar{a}_2 = \frac{1}{2n} \sum_{i=1}^n x_1^2$$

– ще одна ОММ параметра  $\theta$ .

Маємо  $\hat{\theta}_1$  та  $\hat{\theta}_2$  – дві різні оцінки методу моментів одного і того ж невідомого параметра  $\theta$ . Можна також знайти та прирівняти центральні моменти другого порядку:

$$\beta_2(\theta) = D\xi = M\xi^2 - (M\xi)^2 = 2\theta - \frac{\pi\theta}{2} \quad \text{і} \quad \bar{b}_2 = \bar{S}^2,$$

тобто записати рівняння  $(4 - \pi)\theta = 2\bar{S}^2$  і знайти третю оцінку

$$\hat{\theta}_3 = 2\bar{S}^2 / (4 - \pi).$$

Залишається визначитись, яка з оцінок  $\hat{\theta}_1$ ,  $\hat{\theta}_2$  чи  $\hat{\theta}_3$  найкраща. Будь ласка, ті, кому цікаво, доведіть спроможність отриманих статистик та перевірте їх незміщеність.

**Зауваження 3.2.** Імовірнісне підґрунтя апіорної інформації про реївську випадкову величину: величина  $\xi \in R(\theta)$  характеризує відстань від двовимірної точки  $(\eta_1, \eta_2)$ , координати якої незалежні, однаково нормально розподілені з параметрами  $a = 0$ ,  $\sigma^2 = \theta$ , до не випадкової точки  $(0, 0)$ . Тобто  $\xi = \sqrt{\eta_1^2 + \eta_2^2}$ , де  $\eta_1, \eta_2 \in N(0, \theta)$ .

### 3.2. Метод максимальної правдоподібності

Принцип максимальної правдоподібності:

те, що спостерігається, є найбільш імовірним серед усіх можливих альтернатив.

Р. Є. Ямненко.

Ідею методу максимальної правдоподібності проілюструємо на прикладі статистичної задачі про схему Бернуллі, що розглядалась у вступі.

**Приклад 3.4.** Припустимо, що п'ять незалежних випробувань за схемою Бернуллі закінчилися подією "*успіх, успіх, успіх, невдача, невдача*" і зазделегідь імовірність  $\theta = P(\text{"успіх"})$  невідома.

Необхідно оцінити значення  $\theta$ .

Нехай  $\xi$  – кількість "*успіхів*" у одному випробуванні. Отже  $\xi \in B(1, \theta)$ . За умовою, випадкова величина  $\xi$  спостерігалась п'ять разів, в результаті чого маємо вибірку  $(1, 1, 1, 0, 0)$ . Як для схеми Бернуллі, запишемо ймовірність події  $(1, 1, 1, 0, 0) \equiv \text{"успіх, успіх, успіх, невдача, невдача"}$ :

$$P\{(1, 1, 1, 0, 0)\} = \theta^3(1 - \theta)^2 \equiv P(\theta). \quad (3.4)$$

Для різних значень  $\theta$  маємо різні ймовірності  $P(\theta)$ . Наприклад,

$$P(1/3) = 4/243, \quad \text{а} \quad P(1/2) = 1/32.$$

Оскільки подія  $(1, 1, 1, 0, 0)$ , за умовою, відбулася, значення  $\theta = 1/2$  має правдоподібніший вигляд ніж  $\theta = 1/3$ , бо  $P(1/2) > P(1/3)$ .

Логічно оцінкою невідомого параметра покласти те значення  $\theta^o$ , за якого ймовірність події  $(1, 1, 1, 0, 0)$  була б максимальною.

З точки зору аналізу, задача зводиться до визначення точки максимуму функції  $P(\theta) = \theta^3(1 - \theta)^2$  на проміжку  $(0, 1)$ . Слушно перейти до функції

$$l(\theta) = \ln P(\theta) = 3 \ln \theta + 2 \ln(1 - \theta),$$

оскільки точки екстремуму  $P(\theta)$  і  $l(\theta)$  збігаються. Далі знайдемо похідну  $l'(\theta)$ , стаціонарні точки як корені рівняння  $l'(\theta) = 0$  та перевіримо для них достатні умови максимуму. Маємо:

$$l'(\theta) = \frac{3}{\theta} - \frac{2}{1 - \theta} = 0 \Leftrightarrow \theta = \frac{3}{5}.$$

Оскільки

$$l''(\theta) = -\frac{3}{\theta^2} - \frac{2}{(1 - \theta)^2} < 0 \quad \text{для будь-якого } \theta \in (0, 1),$$

то

$$\max_{\theta \in (0, 1)} P(\theta) = P(3/5).$$

Таким чином, значення  $\theta = 3/5$  найбільш правдоподібне і є розв'язком цієї задачі.

Перейдемо до загальних міркувань. Припустимо, що  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi(\theta)$ , що має  $d$  невідомих параметрів,  $\theta = (\theta_1, \dots, \theta_d)$ . Отримаємо оцінки для  $\theta_1, \dots, \theta_d$  методом, запропонованим Р. Фішером<sup>5</sup> у 1912 р. [7], який і називають методом максимальної правдоподібності. Назва відповідає дійсності, оскільки основна ідея Фішера полягала якраз в тому, щоб оцінкою невідомого параметра вважати те значення  $\theta$ , за якого спостережна вибірка була б максимально правдоподібною.

*Правдоподібність* означає, що ймовірність отримати ту саму реалізацію  $(x_1, \dots, x_n)$  при іншому значенні параметра свідомо менша.

---

<sup>5</sup>Рональд Ейлмер Фішер (17.02.1890, Лондон – 29.07.1962, Аделаїда, Австралія) – видатний статистик, біолог-еволюціоніст, генетик. Вперше запропонував методологію планування експерименту, знаходження оцінок невідомих параметрів розподілу, спосіб порівняння генеральних сукупностей, ортогональність як некорельованість незалежно розподілених випадкових величин.

Спочатку розглянемо випадок, коли  $P_\xi(\theta)$  – дискретний розподіл, і запишемо ймовірність події

$A = \text{"в } n \text{ незалежних випробуваннях спостерігались дані } x_1, \dots, x_n \text{"}$ .

Маємо:

$$P(A) = P\{\xi = x_1, \xi = x_2, \dots, \xi = x_n\} = \prod_{i=1}^n P\{\xi = x_i\}.$$

Значення  $x_1, \dots, x_n$  можуть змінюватись від вибірки до вибірки, ймовірності  $P\{\xi = x_i\}$ , вочевидь, залежні від  $\theta$  і, таким чином, функція

$$P(x, \theta) = \prod_{i=1}^n P\{\xi = x_i\}, \text{ де } x = (x_1, \dots, x_n) \in X \subseteq \mathbb{R}^n. \quad (3.5)$$

характеризує ймовірність реалізації  $(x_1, \dots, x_n)$ .

У разі, коли  $P_\xi(\theta)$  – абсолютно неперервний розподіл зі щільністю  $p_\xi(t, \theta)$ , аналогом (3.5) буде щільність розподілу вектора  $(x_1, \dots, x_n)$ :

$$p(t, \theta) = \prod_{i=1}^n p_\xi(t_i, \theta), \text{ де } t = (t_1, \dots, t_n) \in \mathbb{R}^n. \quad (3.6)$$

Для різних значень  $\theta$  маємо різні закони  $P(x, \theta)$  та  $p(t, \theta)$ , відповідно. Знайдемо серед них такий, за якого ймовірність події  $A$  була б найбільшою.

**Означення.** *Функцією правдоподібності* вибірки  $(x_1, \dots, x_n)$  називають функцію  $L(\theta) = L(\theta_1, \dots, \theta_s)$ , що визначається рівністю

$$L(\theta) = \begin{cases} P(x, \theta), & \text{якщо } \xi - \text{дискретна випадкова величина;} \\ p(x, \theta), & \text{якщо } \xi - \text{абсолютно неперервна,} \end{cases}$$

де  $x = (x_1, \dots, x_n)$ , а залежності  $P(x, \theta)$  та  $p(t, \theta)$  визначені в (3.5) та (3.6).

Якщо  $\Theta$  – область припустимих значень параметра  $\theta$ , то задача зводиться до пошуку такого значення  $\theta^o = (\theta_1^o, \dots, \theta_d^o)$ , за якого

$$L(\theta^o) = \max_{\theta \in \Theta} L(\theta).$$

**Означення.** Оцінкою максимальної правдоподібності (ОМП) параметра  $\theta_i$ ,  $i = \overline{1, d}$  називають  $i$ -ту координату  $\theta_i^o$  точки максимуму функції  $L(\theta)$ .

Якщо  $L(\theta) > 0$  для  $\theta \in \Theta$  і  $x \in X$ , перейдемо до функції  $l(\theta) \equiv \ln L(\theta)$ , яку називають *логарифмічною функцією правдоподібності*.

Найбільшого значення  $L(\theta)$  та  $l(\theta)$  досягають в одній і тій самій точці, але дослідження на екстремум останньої в даних задачах зручніше.

Розглянемо приклади здобуття ОМП.

**Приклад 3.5.** Нехай  $(x_1, \dots, x_n)$  – вибірка з нормального розподілу  $N(\theta_1, \theta_2^2)$ , обидва параметри якого  $\theta_1 \in \mathbb{R}$ ,  $\theta_2 > 0$  невідомі.

Знайдемо ОМП для  $\theta_1$  та  $\theta_2$ .

Функцією правдоподібності вибірки  $(x_1, \dots, x_n)$  є

$$L(\theta_1, \theta_2) = \left(\sqrt{2\pi}\theta_2\right)^{-n} \exp\left(-\frac{1}{2\theta_2^2} \sum_{i=1}^n (x_i - \theta_1)^2\right).$$

Тоді

$$l(\theta_1, \theta_2) = -\frac{n}{2} \ln(2\pi) - n \ln \theta_2 - \frac{1}{2\theta_2^2} \sum_{i=1}^n (x_i - \theta_1)^2$$

– логарифмічна функція правдоподібності. Вона диференційована по  $\theta_1$  та  $\theta_2$  у своїй області визначення.

Для пошуку стаціонарних точок знайдемо частинні похідні:

$$\frac{\partial}{\partial \theta_1} l(\theta_1, \theta_2) = \frac{1}{\theta_2^2} \sum_{i=1}^n (x_i - \theta_1), \quad \frac{\partial}{\partial \theta_2} l(\theta_1, \theta_2) = -\frac{n}{\theta_2} + \frac{1}{\theta_2^3} \sum_{i=1}^n (x_i - \theta_1)^2$$

і розв'яжемо систему рівнянь

$$\begin{cases} \sum_{i=1}^n x_i - n\theta_1 = 0, \\ \sum_{i=1}^n (x_i - \theta_1)^2 - n\theta_2^2 = 0. \end{cases}$$

Маємо єдину стаціонарну точку  $(\theta_1^o, \theta_2^o)$ :

$$\theta_1^o = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}, \quad (\theta_2^o)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \theta_1)^2 = \bar{S}^2.$$

Перевіримо для неї достатню умову максимуму. Обчислимо для цього другі похідні:

$$\frac{\partial^2 l}{\partial \theta_1^2}(\theta_1, \theta_2) = -\frac{n}{\theta_2^2}, \quad \frac{\partial^2 l}{\partial \theta_2 \partial \theta_1}(\theta_1, \theta_2) = -\frac{2}{\theta_2^3} \sum_{i=1}^n (x_i - \theta_1),$$

$$\frac{\partial^2 l}{\partial \theta_2^2}(\theta_1, \theta_2) = \frac{n}{\theta_2^2} - \frac{3}{\theta_2^4} \sum_{i=1}^n (x_i - \theta_1)^2$$

та їх значення в стаціонарній точці:

$$\frac{\partial^2 l}{\partial \theta_1^2}(\theta_1^o, \theta_2^o) = -\frac{n}{\bar{S}^2}, \quad \frac{\partial^2 l}{\partial \theta_2 \partial \theta_1}(\theta_1^o, \theta_2^o) = 0$$

а

$$\frac{\partial^2 l}{\partial \theta_2^2}(\theta_1^o, \theta_2^o) = \frac{n}{\bar{S}^2} - \frac{3n}{\bar{S}^2} = -\frac{2n}{\bar{S}^2}.$$

Оскільки

$$\frac{\partial^2 l}{\partial \theta_1^2}(\theta_1^o, \theta_2^o) < 0$$

і визначник матриці других похідних, вочевидь, додатний, то  $\hat{\theta}_1 = \bar{x}$  та  $\hat{\theta}_2 = \bar{S}^2$  – ОМП невідомих параметрів  $\theta_1$  і  $\theta_2$ .

**Приклад 3.6.** Нехай  $(x_1, \dots, x_n)$  – вибірка з пуассонівського розподілу  $\Pi(\theta)$ ,  $\theta > 0$ . Знайдемо ОМП невідомого параметра  $\theta$ .

Для розподілу Пуассона маємо:

$$P\{\xi = x_i\} = \frac{\theta^{x_i}}{x_i!} e^{-\theta}, \quad i = \overline{1, n},$$

тому функція правдоподібності така:

$$L(\theta) = \prod_{i=1}^n \left( \frac{\theta^{x_i}}{x_i!} e^{-\theta} \right) = \frac{\theta^{n\bar{x}}}{\prod_{i=1}^n (x_i!)} e^{-n\theta}$$

і

$$l(\theta) = \ln L(\theta) = n\bar{x} \ln \theta - n\theta - \ln \prod_{i=1}^n (x_i!).$$

З рівняння

$$l'(\theta) = \frac{n\bar{x}}{\theta} - n = 0$$



знайдемо стаціонарну точку  $\theta^o = \bar{x}$ .

Оскільки  $l''(\theta) = -n\bar{x}/\theta^2 < 0$  для будь-якого  $\theta > 0$ , то ОМП  $\hat{\theta} = \bar{x}$ .

**Приклад 3.7.** Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу Бернуллі  $B(1, \theta)$ , параметр  $\theta \in (0, 1)$  якого невідомий. Побудуємо ОММ та ОМП для  $\theta$ .

ОММ зовсім легко знайти, бо для  $\xi \in B(1, \theta)$

$$\alpha_1 = M\xi = \theta, \quad \bar{\alpha}_1 = \bar{x}; \quad \text{і} \quad \hat{\theta} = \bar{x} - \text{шукана оцінка.}$$

Тепер отримаємо ОМП. Функція  $P(x, \theta)$  (див. (3.5)) має вигляд

$$P(x, \theta) = \prod_{i=1}^n P\{\xi = x_i\} = \theta^{n\bar{x}}(1 - \theta)^{n - n\bar{x}}, \quad (3.7)$$

оскільки розподіл бернуллівської випадкової величини такий:

$$P\{\xi = x_i\} = \begin{cases} \theta, & \text{якщо } x_i = 1; \\ 1 - \theta, & \text{якщо } x_i = 0. \end{cases} = \theta^{x_i}(1 - \theta)^{1 - x_i}, \quad i = \overline{1, n}.$$

Звідси

$$l(\theta) = n\bar{x} \ln \theta + (n - n\bar{x}) \ln(1 - \theta)$$

– відповідна логарифмічна функція правдоподібності.

Далі,

$$l'(\theta) = \frac{n\bar{x}}{\theta} - \frac{n - n\bar{x}}{1 - \theta} = 0 \Leftrightarrow \theta = \bar{x}$$

і

$$l''(\theta) = -\frac{n\bar{x}}{\theta^2} - \frac{n - n\bar{x}}{(1 - \theta)^2} < 0, \quad \forall \theta \in (0, 1).$$

Таким чином,  $\hat{\theta} = \bar{x} \in$  також і ОМП невідомого параметра  $\theta$ .

При розв'язанні наступного прикладу в пригоді стане ймовірнісний аналог **нерівності Йенсена**<sup>6</sup>: для будь-якої випадкової величини  $\eta : \Omega \rightarrow E$  і опуклої на множині  $E$  функції  $\varphi$  виконується нерівність

$$M\varphi(\eta) \leq \varphi(M\eta)$$

---

<sup>6</sup>Йоган Віллем Людвіг Вальдемар Йенсен (08.05.1859, Наксков, Данія – 05.03.1925, Копенгаген) – математик та інженер. Найбільш відомий завдяки доведеній нерівності, яку називають його ім'ям, і важливому співвідношенню комплексного аналізу – формулі Йенсена.

*і рівність має місце тоді і тільки тоді, коли  $\varphi$  лінійна.*

**Приклад 3.8.** Нехай  $P_\xi(\theta) = U[0, \theta]$  – рівномірний розподіл на проміжку  $[0, \theta]$ ,  $\theta > 0$  і  $(x_1, \dots, x_n)$  – вибірка з даного розподілу.

Отримаємо ОММ та ОМП невідомого параметра  $\theta$  та перевіримо їх властивості.

Знайдемо ОММ параметра  $\theta$  за першими початковим моментом:

$$\alpha_1(\theta) = M\xi = \frac{\theta}{2}, \quad \bar{a}_1 = \bar{x}; \quad \frac{\theta}{2} = \bar{x} \Leftrightarrow \theta = 2\bar{x}.$$

Якщо ж за другим початковим моментом, то

$$\alpha_2(\theta) = M\xi^2 = \frac{1}{\theta} \int_0^\theta t^2 dt = \frac{\theta^2}{3} = \bar{a}_2 \Leftrightarrow \theta = \sqrt{3\bar{a}_2}.$$

Легко бачити, що обидві знайдені статистики – спроможні ОММ параметра  $\theta$ , як і те, що оцінка  $\hat{\theta}_1 = 2\bar{x}$  незміщена. Перевіримо незміщеність  $\hat{\theta}_2 = \sqrt{3\bar{a}_2}$ .

При обчисленні математичного сподівання випадкової величини  $\hat{\theta}_2$  виникають труднощі, бо з'являється математичне сподівання від кореня:

$$M\hat{\theta}_2 = M\sqrt{3\bar{a}_2}.$$

Але можна його оцінити за нерівністю Йєнсена. Маємо:

$$M\hat{\theta}_2 = M\sqrt{3\bar{a}_2} < \sqrt{3M\xi^2} = \theta$$

з точним знаком  $<$ . Тобто  $\hat{\theta}_2 = \sqrt{3\bar{a}_2}$  – спроможна, але зміщена оцінка параметра  $\theta$ .

Далі знайдемо ОМП того ж  $\theta$ . Функція правдоподібності вибірки з розподілу  $U[0, \theta]$  така:

$$L(\theta) = \begin{cases} \theta^{-n}, & \text{якщо } x_i \in [0, \theta] \text{ для всіх } i = \overline{1, n}; \\ 0, & \text{в іншому випадку.} \end{cases}$$

Але як для функції від  $\theta$ , треба інакше задати умову на  $x_i$ ,  $i = \overline{1, n}$ . Запишемо її в еквівалентній формі:

$$x_i \in [0, \theta] \text{ для всіх } i = \overline{1, n} \Leftrightarrow \theta \geq \max_{i=\overline{1, n}} x_i = x_{(n)}$$

і тоді

$$L(\theta) = \begin{cases} \theta^{-n}, & \text{якщо } \theta \in [x_{(n)}, +\infty); \\ 0, & \text{якщо } \theta \in (0, x_{(n)}). \end{cases}$$

Функція  $L(\theta)$  розривна. Легко бачити, що

$$\max_{[0, +\infty)} L(\theta) = L(x_{(n)}).$$

Тобто метод максимальної правдоподібності призводить до статистики  $\theta_3 = x_{(n)}$ . Перевіримо її спроможність та незміщеність.

Знайдемо функцію та щільність розподілу випадкової величини  $\theta_3$ :

$$\begin{aligned} F_{\theta_3}(t) &= P\{x_{(n)} < t\} = P\{x_1 < t, x_2 < t, \dots, x_n < t\} = \\ &= P^n\{\xi < t\} = \begin{cases} 0, & t < 0; \\ (t/\theta)^n, & t \in [0, \theta]; \\ 1, & t > \theta \end{cases} \end{aligned}$$

і

$$p_{\theta_3}(t) = \begin{cases} 0, & t \notin (0, \theta); \\ nt^{n-1}\theta^{-n}, & t \in [0, \theta]. \end{cases}$$

Тепер з'ясуємо, чи є  $\theta_3$  спроможною оцінкою параметра  $\theta$ , тобто чи буде  $\theta_3$  збігатися за ймовірністю до  $\theta$ . Для  $\forall \varepsilon > 0$  маємо:

$$\begin{aligned} P\{|\theta_3 - \theta| \geq \varepsilon\} &= \int_{-\infty}^{\theta-\varepsilon} p_{\theta_3}(t) dt + \int_{\theta+\varepsilon}^{+\infty} p_{\theta_3}(t) dt = \\ &= \begin{cases} n\theta^{-n} \int_0^{\theta-\varepsilon} t^{n-1} dt = (1 - \varepsilon/\theta)^n & \text{при } 0 < \varepsilon < \theta; \\ 0, & \text{якщо } \varepsilon > \theta. \end{cases} \end{aligned}$$

Вочевидь,

$$\forall \varepsilon > 0 \quad P\{|\theta_3 - \theta| \geq \varepsilon\} \rightarrow 0, \quad \text{коли } n \rightarrow \infty.$$

Отже,  $\hat{\theta}_3 = x_{(n)}$  – спроможна ОМП параметра  $\theta$ .

Для перевірки незміщеності оцінки  $\hat{\theta}_3$  за відомою щільністю розподілу обчислимо математичне сподівання:

$$M\hat{\theta}_3 = \int_{-\infty}^{+\infty} tp_{\hat{\theta}_3}(t) dt = \frac{n}{\theta^n} \int_0^{\theta} t^n dt = \frac{n}{n+1} \theta \rightarrow \theta, \quad \text{коли } n \rightarrow \infty.$$

Тоді  $\hat{\theta}_3$  є асимптотичною незміщеною ОМП невідомого параметра  $\theta$ .

Розглянемо ще один цікавий приклад на рівномірний розподіл.

**Приклад 3.9.** Нехай  $P_\xi(\theta) = U[\theta, \theta + 1]$  і  $\theta \in \mathbb{R}$  – невідомий параметр.

Знайдемо ОМП  $\theta$  за вибіркою  $(x_1, \dots, x_n)$  з даного розподілу.

Запишемо функцію правдоподібності вибірки:

$$L(\theta) = \begin{cases} 1, & \text{якщо } \theta \leq x_i \leq \theta + 1 \text{ для всіх } i = \overline{1, n}; \\ 0, & \text{якщо інакше,} \end{cases}$$

або, як для функції змінної  $\theta$ ,

$$L(\theta) = \begin{cases} 1, & \text{якщо } x_{(n)} - 1 \leq \theta \leq x_{(1)}; \\ 0, & \text{в іншому разі.} \end{cases}$$

Функція правдоподібності є сталою на проміжку  $[x_{(n)} - 1, x_{(1)}]$ . У такому разі будь-яка точка  $\theta \in [x_{(n)} - 1, x_{(1)}]$  може бути ОМП параметра  $\theta$ .

У наступному прикладі розглянемо розподіл Парето, що визначається щільністю

$$p(t, m, \alpha) = \begin{cases} \frac{\alpha}{m} \cdot \left(\frac{m}{t}\right)^{\alpha+1}, & t \geq m; \\ 0, & t < m \end{cases}$$

з двома параметрами  $\alpha, m > 0$ .

Цей розподіл виник в роботах Вільфредо Парето<sup>4</sup> як розподіл статків суспільства з більшою частиною капіталу, що зосереджена в руках невеликої кількості людей. Параметр  $\alpha$  отримав назву *індекса* Парето. Наприклад, при  $\alpha = \log_4 5$  це "правило 80 – 20". Як виявилося згодом, розподіл "80 – 20" має досить багато природних, економічних, соціальних (і т. і.) явищ.

**Приклад 3.10.** Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу Парето.

Розглянемо два випадки: 1) спочатку будемо вважати, що параметр  $\alpha > 0$  відомий, а параметр  $0 < m \equiv \theta$  – ні. Побудуємо ОМП для  $\theta$ .

---

<sup>4</sup>Вільфредо Парето (15.07.1848, Париж, Франція – 19.08.1923, Женева, Швейцарія) – відомий економіст, філософ і соціолог, один із засновників теорії еліт. Розподіл Парето широко застосовується в економіці та інших наукових дисциплінах.

Запишемо функцію правдоподібності цієї вибірки:

$$L(\theta) = \begin{cases} \frac{\alpha\theta^\alpha}{x_1^{\alpha+1}} \cdot \frac{\alpha\theta^\alpha}{x_2^{\alpha+1}} \cdots \frac{\alpha\theta^\alpha}{x_n^{\alpha+1}}, & \text{якщо } x_i \geq \theta \text{ для всіх } i = \overline{1, n}; \\ 0, & \text{коли хоча б одне значення } x_i < \theta \end{cases}$$

та проаналізуємо її поведінку як функції від  $\theta$ .

За означенням,  $L(\theta) > 0$  лише тоді, коли  $\theta < x_{(1)}$ . Але зі спаданням  $\theta$  при  $\alpha > 0$  величина  $\theta^\alpha$  також спадає, тому  $\max L(\theta)$  досягається в точці  $\theta^o = x_{(1)}$ . Тобто ОМП  $\theta = x_{(1)}$ .

2) Нехай тепер параметр  $m > 0$  відомий, а  $0 < \alpha \equiv \theta$  – ні. Побудуємо ОМП для  $\theta$  в такому випадку.

Функція правдоподібності тепер має вигляд:

$$L(\theta) = \begin{cases} \frac{\theta^n m^{n\theta}}{(x_1 \cdots x_n)^{\theta+1}}, & \text{якщо } x_i \geq m \text{ для всіх } i = \overline{1, n}; \\ 0, & \text{коли хоча б одне значення } x_i < m. \end{cases}$$

Припустимо, що всі вибіркові дані задовольняють нерівності  $x_i \geq m$ ,  $i = \overline{1, n}$  і за цієї умови визначимо логарифмічну функцію правдоподібності:

$$l(\theta) = n \ln \theta + n\theta \ln m - (\theta + 1) \sum_{i=1}^n \ln x_i.$$

Далі, рівняння

$$l'(\theta) = \frac{n}{\theta} + n \ln m - \sum_{i=1}^n \ln x_i = 0$$

виконується, коли

$$\frac{n}{\theta} = \sum_{i=1}^n \ln \frac{x_i}{m}.$$

Звідси

$$\theta = n \left( \sum_{i=1}^n \ln \frac{x_i}{m} \right)^{-1}.$$

Крім того,  $l''(\theta) = -\frac{n}{\theta^2} < 0$  для будь-якого  $\theta > 0$ , тому

$$\text{ОМП } \theta = n \left( \sum_{i=1}^n \ln \frac{x_i}{m} \right)^{-1}, \quad \text{як тільки } x_i > m, \quad \forall i = \overline{1, n}.$$

**Контрольні запитання:**

1. Охарактеризуйте суть методу моментів та методу максимальної правдоподібності.
2. У чому полягає невизначеність методу моментів і чи є ця характеристика відповідним недоліком?
3. Що таке функція правдоподібності, логарифмічна функція правдоподібності? Чому при пошуку ОМП розглядають останню?
4. Нагадайте, будь ласка, достатні умови екстремуму функції однієї та двох змінних.

**Вправи та задачі: 3.1.** Побудуйте ОММ невідомих параметрів наступних розподілів: а) геометричного ; б) рівномірного  $U[\theta_1, \theta_2]$ ; в) гамма-розподілу  $\Gamma(\theta_1, \theta_2)$ ; г) нормального  $N(a, \theta^2)$  ( $a \in \mathbb{R}$  –відоме). Перевірте властивості отриманих оцінок.

**3.2.** Побудуйте ОМП невідомих параметрів розподілів: а) біноміального  $B(m, \theta)$  ( $m \in \mathbb{N}$  – відоме); б) рівномірних  $U[0, 2\theta]$  та  $U[-\theta, \theta]$ ; в) нормального  $N(a, \theta^2)$  ( $a \in \mathbb{R}$  –відоме). Перевірте, будь ласка, властивості отриманих оцінок.

## §4. Оцінки з мінімальною дисперсією

### 4.1. Порівняння оцінок

Розглянемо випадок однопараметричного розподілу  $P_\xi(\theta)$ , параметр якого  $\theta \in R$  невідомий. За допомогою методів моментів та максимальної правдоподібності можна отримати досить багато різних оцінок для  $\theta$ , зокрема спроможних та незміщених. Крім того, якщо вже маємо хоча б одну незміщену оцінку, то на її основі можна будувати інші незміщені оцінки. Наприклад, у такий спосіб:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{\theta}_1 = \frac{1}{[n/2]} \sum_{i=1}^{[n/2]} x_i, \quad \hat{\theta}_j = \frac{1}{n-j} \sum_{i=j+1}^n x_i,$$

де  $1 \leq j < n$  – деяке натуральне число.

Позначимо через  $K_\theta$  клас незміщених оцінок невідомого параметра  $\theta$ . Нагальне питання: яка оцінка з  $K_\theta$  краща? Як їх порівнювати?

Зрозуміло, що кращою є та оцінка  $\hat{\theta}$ , для якої відхилення від істинного значення параметра  $\theta$  менші. Відхилення від  $\theta$ , тобто модуль різниці  $|\hat{\theta} - \theta|$  є випадковою величиною. Відтак порівнюють оцінки в середньому, і як міру порівняння беруть  $M|\hat{\theta} - \theta|$ , а краще – в середньоквадратичному, і тоді розглядають  $M(\hat{\theta} - \theta)^2$ . Крім того, параметр  $\theta$  невідомий, але у разі незміщеної оцінки  $\theta = M\hat{\theta}$  і тоді

$$M(\hat{\theta} - \theta)^2 = M(\hat{\theta} - M\hat{\theta})^2 = D\hat{\theta} \quad \text{для } \hat{\theta} \in K_\theta.$$

Отже, порівняння оцінок у класі  $K_\theta$  у середньоквадратичному є простим порівнянням їх дисперсій. Кращою серед них буде та, дисперсія якої менша.

Логічно цікавитись тими оцінками із  $K_\theta$ , які мають мінімально можливу дисперсію.

**Означення.** Оцінку  $\theta^* \in K_\theta$  називають *ефективною* в класі  $K_\theta$ , якщо для будь-якої іншої оцінки  $\hat{\theta} \in K_\theta$  справедлива нерівність  $D\theta^* < D\hat{\theta}$ .

Ефективну оцінку в класі  $K_\theta$  називають просто *ефективною* або *найкращою*.

Зауважимо, що ефективної оцінки може не бути, тому що множина  $\{D\hat{\theta}, \hat{\theta} \in K_\theta\}$  не завжди має мінімальний елемент. Але якщо мінімум існує, тобто

$$\exists \theta^* \in K_\theta : \quad \forall \hat{\theta} \in K_\theta \quad D\theta^* < D\hat{\theta},$$

то  $\theta^*$  – єдина ефективна (найкраща) оцінка.

Про те наступна теорема.

**Теорема 4.1.** Якщо ефективна оцінка в класі  $K_\theta$  існує, то вона єдина.

**Доведення.** Припустимо супротивне: існують дві різні незміщені ефективні оцінки  $\theta_1^*$  і  $\theta_2^*$  одного й того ж параметра  $\theta$ :

$$M\theta_1^* = M\theta_2^* = \theta, \quad D\theta_1^* = D\theta_2^*.$$

Розглянемо оцінку  $\theta_3 = (\theta_1^* + \theta_2^*)/2$ . Вочевидь,  $M\theta_3 = \theta$ , тобто  $\theta_3 \in K_\theta$ .

Але

$$\begin{aligned} D\theta_3 &= M \left( \frac{\theta_1^* + \theta_2^*}{2} - \theta \right)^2 = M \left( \frac{\theta_1^* - \theta}{2} + \frac{\theta_2^* - \theta}{2} \right)^2 = \\ &= \frac{1}{4} D\theta_1^* + \frac{1}{4} D\theta_2^* + \frac{1}{2} M (\theta_1^* - \theta) (\theta_2^* - \theta) \leq \\ &\leq \frac{1}{2} D\theta_1^* + \frac{1}{2} D\theta_1^* = D\theta_1^*. \end{aligned} \quad (4.1)$$

Оцінка в (4.1) є наслідком нерівності Коші або властивості коефіцієнта кореляції  $\rho$ , згідно з якою

$$|\rho(\theta_1^*, \theta_2^*)| = \frac{|cov(\theta_1^*, \theta_2^*)|}{\sqrt{D\theta_1^* D\theta_2^*}} = \frac{|M(\theta_1^* - \theta)(\theta_2^* - \theta)|}{D\theta_1^*} \leq 1.$$

Знак  $<$  в нерівності (4.1) неможливий, оскільки  $\theta_3, \theta_1^* \in K_\theta$  і  $\theta_1^*$  – ефективна оцінка. Але рівність в (4.1) може бути лише тоді, коли  $\theta_1^*$  і  $\theta_2^*$  лінійно залежні:  $\theta_1^* = c\theta_2^* + d$ .

У такому разі

$$D\theta_1^* = c^2 D\theta_2^* \quad \Leftrightarrow \quad c^2 = 1.$$

Випадок  $c = 1$  не цікавий, бо тоді  $d = M(\theta_1^* - \theta_2^*) = 0$  і звідси  $\theta_1^* = \theta_2^*$ .

Коли ж  $c = -1$ , маємо:

$$d = M(\theta_1^* + \theta_2^*) = 2\theta \quad \text{і} \quad \theta_1^* + \theta_2^* = 2\theta.$$



Отже  $D\theta_3 = 0$  і, як наслідок,  $D\theta_1^* = D\theta_2^* = 0$ . Тобто оцінки  $\theta_1^*$ ,  $\theta_2^*$ ,  $\theta_3$  – деякі сталі в класі  $K_\theta$ . Вочевидь,  $\theta_1^* = \theta_2^* = \theta_3 = \theta$  (тривіальний випадок).

□

Задача про пошук інфімуму множини  $\{D\hat{\theta} | \hat{\theta} \in K_\theta\}$  на перший погляд здається надто складною. Але за певних умов регулярності щодо розподілу  $P_\xi(\theta)$ , нижню межу дисперсій усіх незміщених оцінок параметра  $\theta$  визначає нерівність, отримана незалежно Крамером, Рао та Фреше.<sup>1</sup>

## 4.2. Нерівність Рао-Крамера

**Розподіл дискретний.** Розглянемо спочатку сім'ю дискретних однопараметричних розподілів  $\{P(\theta), \theta \in \Theta\}$ , де  $\Theta \subseteq \mathbb{R}$  – деякий інтервал, скінченний або нескінченний. Нехай  $x = (x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi(\theta) \subset P(\theta)$ , параметр  $\theta \in \Theta$  якого невідомий,  $X \subseteq \mathbb{R}^n$  – вибіровий простір, а

$$L(\theta) \equiv L(x, \theta) = P\{\xi = x_1, \xi = x_2, \dots, \xi = x_n\}$$

– функція правдоподібності вибірки (у цьому питанні важливо слідкувати за залежністю функції  $L$  не тільки від  $\theta$ , а ще й від  $x$ ). Тоді

$$\sum_{x \in X} L(x, \theta) = 1. \quad (4.2)$$

Нехай  $\hat{\theta}(x_1, \dots, x_n) \in K_\theta$  – будь-яка незміщена оцінка параметра  $\theta$ . Її розподіл визначається розподілом вектора  $x = (x_1, \dots, x_n)$ , тобто функцією правдоподібності  $L(x, \theta)$  і тоді

$$M\hat{\theta} = \sum_{x \in X} \hat{\theta}(x) L(x, \theta) = \theta. \quad (4.3)$$

---

<sup>1</sup>Гаральд Крамер (25.09.1893, Стокгольм - 05.10.1985, Стокгольм) – математик, актуарій і статистик, спеціаліст з імовірнісної теорії чисел; один із гігантів галузі статистики.

Кальямпузи Радхакришна Рао (10.09.1920, Ховіна Хадагалі, Індія – 22.08.2023, Нью-Йорк, США) – математик, відомий спеціаліст галузі статистики.

Моріс Рене Фреше (02.09.1878, Маліньї, Франція – 04.09.1973, Париж) – математик, есперантист, статистик, працював в галузі топології і функціонального аналізу, теорії ймовірностей та математичної статистики.

Надалі будемо припускати, що функція правдоподібності  $L(x, \theta)$  задовольняє умовам:

- 1) множина  $X^0 \subseteq X$ , на якій  $L(x, \theta) > 0$  не залежить від  $\theta$ ;
- 2) функція правдоподібності  $L(x, \theta)$  неперервно диференційована за параметром  $\theta$ ;
- 3) логарифмічна похідна за параметром  $\theta$  функції правдоподібності на множині  $X^0 \times \Theta$  ненульова:  $h(x, \theta) \equiv \frac{\partial}{\partial \theta} \ln L(x, \theta) \neq 0$  і

$$0 < Mh^2(x, \theta) < +\infty, \quad \forall \theta \in \Theta;$$

- 4) ряди в (4.3) та (4.2) можна почленно диференціювати за  $\theta$ .

**Означення.** У разі, коли сім'я дискретних розподілів  $\{P(\theta), \theta \in \Theta\}$  задовольняє умовам 1)–4), її називають *дискретним регулярним* або просто *регулярним сімейством*.

Отже, якщо  $\{P(\theta), \theta \in \Theta\}$  – регулярне сімейство, то диференціюванням тотожності (4.3) маємо:

$$\frac{d}{d\theta} \sum_{x \in X} \hat{\theta}(x) L(x, \theta) = \sum_{x \in X} \hat{\theta}(x) \frac{\partial}{\partial \theta} L(x, \theta) = 1, \quad \forall \theta \in \Theta. \quad (4.4)$$

Доцільно тепер подати  $\frac{\partial}{\partial \theta} L(x, \theta)$  через логарифмічну похідну функції  $L(x, \theta)$  за параметром  $\theta$ :

$$h(x, \theta) = \frac{\partial}{\partial \theta} \ln L(x, \theta) = \frac{\partial}{\partial \theta} L(x, \theta) \frac{1}{L(x, \theta)}, \quad x \in X^0, \quad \theta \in \Theta. \quad (4.5)$$

Тоді

$$\frac{\partial}{\partial \theta} L(x, \theta) = h(x, \theta) L(x, \theta) \quad \text{на множині } X^0 \times \Theta \quad (4.6)$$

і праву частину співвідношення (4.4) можна переписати так:

$$1 = \sum_{x \in X^0} \hat{\theta}(x) h(x, \theta) L(x, \theta) = M \left( \hat{\theta}(x) h(x, \theta) \right). \quad (4.7)$$

Крім того, диференціюванням тотожності (4.2) отримаємо, що

$$\frac{d}{d\theta} \sum_{x \in X} L(x, \theta) = \sum_{x \in X^0} \frac{\partial}{\partial \theta} L(x, \theta) = 0. \quad (4.8)$$

Тобто

$$Mh(x, \theta) = \sum_{x \in X^0} h(x, \theta) L(x, \theta) = \sum_{x \in X^0} \frac{\partial}{\partial \theta} L(x, \theta) = 0. \quad (4.9)$$

Звідси і з (4.8) за нерівністю Коші маємо:

$$1 = M \left( \hat{\theta}(x) h(x, \theta) \right) = M \left( (\hat{\theta}(x) - \theta) h(x, \theta) \right) \leq \sqrt{D\hat{\theta}(x) M h^2(x, \theta)}. \quad (4.10)$$

Отже,

$$D\hat{\theta} \geq \frac{1}{M h^2(x, \theta)} = \frac{1}{M \left( \frac{\partial}{\partial \theta} \ln L(x, \theta) \right)^2} \quad \text{для будь-якого } \hat{\theta} \in K_\theta. \quad (4.11)$$

**Розподіл абсолютно неперервний.** У випадку, коли  $\{P(\theta), \theta \in \Theta\}$  – сім'я абсолютно неперервних розподілів ( $\Theta \subseteq \mathbb{R}$  – скінченний або нескінченний інтервал), проведемо аналогічні міркування. Нехай  $x = (x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi(\theta) \subset P(\theta)$ , параметр  $\theta \in \Theta$  якого невідомий,  $X \subseteq \mathbb{R}^n$  – вибірковий простір, а

$$L(x, \theta) = \prod_{i=1}^n p_\xi(x_i, \theta)$$

– функція правдоподібності вибірки. Тоді

$$\int_X L(x, \theta) dx = 1. \quad (4.12)$$

Нехай тепер  $\hat{\theta}(x_1, \dots, x_n) \in K_\theta$  – будь-яка незміщена оцінка параметра  $\theta$ . Як і в дискретному випадку, її розподіл визначається функцією правдоподібності і

$$M\hat{\theta} = \int_X \hat{\theta}(x) L(x, \theta) dx = \theta. \quad (4.13)$$

Тотожності (4.12) і (4.13) можна розглядати як неперервні аналоги (4.2) та (4.3), відповідно.

Для функції правдоподібності вибірки з сім'ї абсолютно неперервних розподілів  $P(\theta)$  знову припустимо виконання умов регулярності 1)–3) та умову 4) у вигляді:

4) інтеграли в (4.12) і (4.13) можна почленно диференціювати по  $\theta$ .

**Означення.** Сім'ю абсолютно неперервних розподілів  $\{P_\xi(\theta), \theta \in \Theta\}$ , що задовольняє умовам 1)–4), називають *абсолютно неперервним регулярним сімейством* або просто *регулярним сімейством*.

За умови регулярності сімейства  $\{P_\xi(\theta), \theta \in \Theta\}$ , продиференціюємо тотожність (4.12). Маємо:

$$\frac{d}{d\theta} \int_X L(x, \theta) dx = \int_{X^0} \frac{\partial}{\partial \theta} L(x, \theta) dx = 0, \quad \forall \theta \in \Theta. \quad (4.14)$$

Якщо ж продиференціювати (4.13), то

$$\frac{d}{d\theta} \int_X \hat{\theta}(x) L(x, \theta) dx = \int_{X^0} \hat{\theta}(x) \frac{\partial}{\partial \theta} L(x, \theta) dx = 1, \quad \forall \theta \in \Theta. \quad (4.15)$$

Далі, як і в (4.5), через  $h(x, \theta)$  позначимо логарифмічну похідну функції  $L(x, \theta)$  і, скориставшись співвідношенням (4.6), перепишемо (4.14) та (4.15) в еквівалентній формі: для будь-якого  $\forall \theta \in \Theta$

$$\int_{X^0} \frac{\partial}{\partial \theta} L(x, \theta) dx = \int_{X^0} h(x, \theta) L(x, \theta) dx = 0 \Leftrightarrow Mh(x, \theta) = 0$$

і

$$\int_{X^0} \hat{\theta}(x) \frac{\partial L(x, \theta)}{\partial \theta} dx = \int_{X^0} \hat{\theta}(x) h(x, \theta) L(x, \theta) dx = 1 \Leftrightarrow M(\hat{\theta}(x) h(x, \theta)) = 1.$$

Звідси легко отримати нерівність (4.10), а з неї і (4.11) в термінах і позначеннях для абсолютно неперервного розподілу.

**Зауваження 4.1.** За означенням математичного сподівання як інтегралу Лебега за ймовірнісною мірою, наведені вище доведення можна записати єдиним чином.

**Загальний випадок.**

**Означення.** Нехай  $\theta \in \Theta$ , функцію

$$I(\theta) \equiv Mh^2(x, \theta) = M \left( \frac{\partial}{\partial \theta} \ln L(x, \theta) \right)^2 \quad (4.16)$$

називають *інформацією Фішера* вибірки  $x = (x_1, \dots, x_n)$  з розподілу  $P(\theta)$ .

А функцію

$$h(x, \theta) = \frac{\partial}{\partial \theta} \ln L(x, \theta)$$

звуть *функцією впливу  $n$  спостережень*.

В термінах інформації Фішера  $I(\theta)$  нерівність (4.11) набуває вигляду

$$D\hat{\theta} \geq \frac{1}{I(\theta)} \quad \text{для будь-якого } \hat{\theta} \in K_{\theta}. \quad (4.17)$$

Співвідношення (4.17) називають *нерівністю Рао-Крамера*.

Таким чином, окремо для сімейств дискретних та абсолютно неперервних розподілів за відповідних умов регулярності розглянуто доведення нерівності Рао-Крамера.

**Теорема 4.2 (Рао, Крамер, Фреше).** *Нехай маємо регулярне сімейство розподілів  $\{P(\theta), \theta \in \Theta\}$ , де  $\Theta \subseteq \mathbb{R}$  – скінченний або нескінченний інтервал,  $x = (x_1, \dots, x_n)$  – вибірка з розподілу  $P_{\xi}(\theta) \subset P(\theta)$ , параметр  $\theta \in \Theta$  якого невідомий і  $\hat{\theta} \in K_{\theta}$  – незміщена оцінка параметра  $\theta$ .*

*Тоді справедлива нерівність (4.17), де інформація Фішера  $I(\theta)$  визначена в (4.16).*

**Зауваження 4.2.** Нерівність Рао-Крамера дає оцінку знизу дисперсії  $D\hat{\theta}$  для будь-якої статистики  $\hat{\theta} \in K_{\theta}$ . Нижня межа не обов'язково досягається. Але якщо для деякої незміщеної оцінки  $\hat{\theta}$

$$D\hat{\theta} = \frac{1}{I(\theta)},$$

то оцінка  $\hat{\theta}$  ефективна, іншої такої не знайти.

### 4.3. Способи подання інформації Фішера

**Теорема 4.3.** *За умови, що тотожність (4.8) у дискретному випадку і (4.14) – у абсолютно неперервному, можна диференціювати ще раз по  $\theta$ , тобто*

$$\sum_{x \in X^0} \frac{\partial^2}{\partial \theta^2} L(x, \theta) dx = 0 \quad \text{і} \quad \int_{X^0} \frac{\partial^2}{\partial \theta^2} L(x, \theta) dx = 0, \quad \text{відповідно,}$$

*інформацію Фішера можна записати у вигляді:*

$$I(\theta) = -M \frac{\partial^2}{\partial \theta^2} \ln L(x, \theta), \quad \forall \theta \in \Theta. \quad (4.18)$$

**Доведення.** Розглянемо доведення для випадку абсолютно неперервного розподілу. Оскільки

$$\frac{\partial^2}{\partial \theta^2} \ln L(x, \theta) = \frac{\partial}{\partial \theta} h(x, \theta) = \frac{\frac{\partial^2}{\partial \theta^2} L(x, \theta)}{L(x, \theta)} - \left( \frac{\frac{\partial}{\partial \theta} L(x, \theta)}{L(x, \theta)} \right)^2$$

(як похідна добутку), то

$$\begin{aligned} M \frac{\partial^2}{\partial \theta^2} \ln L(x, \theta) &= \int_{X^0} \left( \frac{\partial^2}{\partial \theta^2} \ln L(x, \theta) \right) L(x, \theta) dx = \\ &= \int_{X^0} \frac{\partial^2}{\partial \theta^2} L(x, \theta) dx - \int_{X^0} \left( \frac{\frac{\partial}{\partial \theta} L(x, \theta)}{L(x, \theta)} \right)^2 L(x, \theta) dx = \\ &= - \int_{X^0} h^2(x, \theta) L(x, \theta) dx = -Mh^2(x, \theta) = -I(x, \theta). \quad \square \end{aligned}$$

**Теорема 4.4.** За умов регулярності 1)–4), для будь-якого  $\theta \in \Theta$

$$Mh(x, \theta) = \frac{\partial}{\partial \theta} \ln L(x, \theta) = 0 \quad \text{і} \quad I(\theta) = Dh(x, \theta) = D \frac{\partial}{\partial \theta} \ln L(x, \theta).$$

Твердження теореми випливає з тотожностей (4.9) та (4.14), відповідно.

**Теорема 4.5.** Інформація Фішера  $I(\theta)$  залежить від  $n$  лінійно:

$$I(\theta) = nI_1(\theta), \tag{4.19}$$

де

$$I_1(\theta) = Dh(x_1, \theta) = D \frac{\partial}{\partial \theta} \ln L(x_1, \theta) \tag{4.20}$$

– інформація Фішера одного спостереження  $x_1$ .

Дійсно, наприклад, для дискретного розподілу маємо:

$$I(\theta) = \sum_{i=1}^n D \left( \frac{\partial}{\partial \theta} \ln L(x_i, \theta) \right) = nI_1(\theta),$$

оскільки  $x_i, i = \overline{1, n}$  – незалежні однаково розподілені випадкові величини.

Перед тим як перейти до прикладів пошуку ефективних оцінок, зробимо деякі зауваження стосовно умов регулярності. Умови 4), що потребують почленного диференціювання кратних рядів та інтегралів, можна

перевірити за допомогою теорем математичного аналізу, а можна скористатися лемою.

**Лема 4.1.** [1, стор. 151] *Припустимо, що  $x = (x_1, \dots, x_n)$  – вибірка з розподілу сім'ї  $\{P(\theta), \theta \in \Theta\}$ ,  $L(x, \theta)$  – функція правдоподібності вибірки і виконуються умови:*

- (i) *функції  $\sqrt{L(x, \theta)}$  для м. в.  $x \in X$  неперервно диференційовані по  $\theta$ ;*
- (ii) *інформація Фішера  $I(\theta)$  існує, додатна і неперервна по  $\theta$ ;*
- (iii)  *$\hat{\theta}$  – будь-яка статистика, для якої  $M\hat{\theta}^2 < c < +\infty$  при  $\theta \in \Theta$  з деякою сталою  $c \in \mathbb{R}$ .*

*Тоді у разі дискретного розподілу  $P(\theta)$  справедливе співвідношення (4.4) і (4.15) – для неперервного.*

**Приклад 4.1.** Нехай  $(x_1, \dots, x_n)$  – вибірка з нормального розподілу  $P_\xi(\theta) = N(\theta, \sigma^2)$ , де  $\sigma > 0$  – відоме число, а параметр  $\theta \in \mathbb{R}$  невідомий.

Знайдемо ефективну оцінку для  $\theta$ .

Функцією правдоподібності вибірки  $x = (x_1, \dots, x_n)$  з розподілу  $N(\theta, \sigma^2)$  є функція

$$L(x, \theta) = \left(\sqrt{2\pi}\sigma\right)^{-n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta)^2\right),$$

що задовольняє умову (i) леми 4.1.

Інформацію Фішера знайдемо за формулою (4.19):

$$h(x_1, \theta) = \frac{\partial}{\partial \theta} \ln L(x_1, \theta) = \frac{x_1 - \theta}{\sigma^2},$$

$$I_1(\theta) = Dh(x_1, \theta) = \frac{1}{\sigma^4} D(x_1 - \theta) = \frac{1}{\sigma^2}.$$

Отже,  $I(\theta) = n/\sigma^2$  і умова (ii) леми 4.1 також виконується.

Незміщеною оцінкою для  $\theta$  є вибіркове середнє:  $\hat{\theta} = \bar{x} \in K_\theta$  і

$$D\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{\sigma^2}{n}.$$

Тобто в нерівності Рао-Крамера для оцінки  $\hat{\theta} = \bar{x}$  маємо знак "=", вона і є ефективною:  $\theta^* = \bar{x}$ .

**Приклад 4.2.** Нехай тепер  $(x_1, \dots, x_n)$  – вибірка з нормального розподілу  $P_\xi(\theta) = N(a, \theta^2)$ , де  $a \in \mathbb{R}$  – відоме число, а  $\theta^2$  – невідомий параметр. Знайдемо ефективну оцінку для  $\theta^2$ .

Функцією правдоподібності даної вибірки є

$$L(x, \theta) = \left(\sqrt{2\pi}\theta\right)^{-n} \exp\left(-\frac{1}{2\theta^2} \sum_{i=1}^n (x_i - a)^2\right)$$

і вона задовольняє умову (i) леми 4.1.

Функція впливу одного спостереження має вигляд:

$$h(x_1, \theta^2) = \frac{\partial}{\partial(\theta^2)} \ln L(x_1, \theta) = -\frac{1}{2\theta^2} + \frac{1}{2\theta^4}(x_1 - a)^2.$$

Тоді

$$I_1(\theta^2) = Dh(x_1, \theta^2) = \frac{1}{4\theta^8} D(x_1 - a)^2.$$

Знайдемо  $D(x_1 - a)^2$ :

$$\begin{aligned} M(x_1 - a)^4 &= \frac{1}{\sqrt{2\pi}\theta} \int_{-\infty}^{+\infty} (t - a)^4 e^{-\frac{(t-a)^2}{2\theta^2}} dt = \frac{\theta^4}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^4 e^{-\frac{t^2}{2}} dt = \\ &= \frac{\theta^4}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^3 d\left(-e^{-\frac{t^2}{2}}\right) = \frac{3\theta^4}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} dt = 3\theta^4 \end{aligned}$$

і  $D(x_1 - a)^2 = 2\theta^4$ . Остаточно,

$$I_1(\theta^2) = \frac{2\theta^4}{4\theta^8} = \frac{1}{2\theta^4}, \quad \text{а} \quad I(\theta^2) = \frac{n}{2\theta^4}.$$

За лемою 4.1 і теоремою Роа-Крамера-Фреше, дисперсія будь-якої незміщеної оцінки параметра  $\theta^2$  буде не менша ніж  $2\theta^4/n$ .

Неважко перевірити, що для нормального розподілу  $N(a, \theta^2)$  ОММ за другим центральним моментом та ОМП параметра  $\theta^2$  буде оцінка

$$\hat{\theta}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - a)^2 \in K_\theta$$

і

$$D\hat{\theta}^2 = \frac{1}{n^2} \sum_{i=1}^n D(x_i - a)^2 = \frac{2\theta^4}{n}.$$



Тому така оцінка і є ефективною,

$$(\theta^2)^* = \frac{1}{n} \sum_{i=1}^n (x_i - a)^2.$$

**Приклад 4.3.** Розглянемо випадок нерегулярного сімейства, а саме, нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi(\theta)$  з щільністю

$$p_\xi(t, \theta) = \begin{cases} e^{\theta-t}, & \text{якщо } t \geq \theta; \\ 0 & \text{при } t < \theta \end{cases} = \begin{cases} e^{\theta-t}, & \text{якщо } \theta \leq t; \\ 0 & \text{при } \theta > t \end{cases}$$

і параметр  $\theta$  невідомий.

Для цього розподілу носій щільності залежить від  $\theta$ , тобто порушується вже умова 1) регулярності; не має місця також тотожність (4.14).

У таких випадках досить часто можна знайти незміщені оцінки з "аномально малою" дисперсією, тобто такі, для яких дисперсія менша ніж та нижня межа, яку гарантує нерівність Рао-Крамера.

Дійсно, для цього випадку

$$h(x_1, \theta) = \begin{cases} 1, & \text{якщо } \theta \leq x_1; \\ 0 & \text{при } \theta > x_1, \end{cases}$$

$$I_1(\theta) = Dh(x_1, \theta) = \int_{\theta}^{+\infty} e^{\theta-t} dt = 1; \quad I(\theta) = n$$

і формальний запис нерівності Рао-Крамера призводить до такого:

$$D\hat{\theta} \geq 1/n \quad \text{для будь-якого } \hat{\theta} \in K_\theta.$$

Покажемо, що існує оцінка параметра  $\theta$ , дисперсія якої дорівнює  $1/n^2$ . З цією метою розглянемо статистику  $\theta^0 = x_{(1)}$ .

Для перевірки її властивостей визначимо функцію розподілу  $\theta^0$ :

$$\begin{aligned} F_{\theta^0}(t, \theta) &= P\{x_{(1)} < t\} = 1 - P\{x_{(1)} \geq t\} = \\ &= 1 - \prod_{i=1}^n P\{x_i \geq t\} = 1 - (1 - F_\xi(t, \theta))^n. \end{aligned}$$

Але

$$F_{\xi}(t, \theta) = \begin{cases} 1 - e^{\theta-t}, & \text{якщо } t \geq \theta; \\ 0 & \text{при } t < \theta \end{cases}$$

і, остаточно,

$$F_{\theta^0}(t, \theta) = \begin{cases} 1 - e^{n(\theta-t)}, & \text{якщо } t \geq \theta; \\ 0 & \text{при } t < \theta. \end{cases}$$

Звідси

$$p_{\theta^0}(t, \theta) = \begin{cases} ne^{n(\theta-t)}, & \text{якщо } t \geq \theta; \\ 0 & \text{при } t < \theta \end{cases}$$

і

$$M\theta^0 = n \int_{\theta}^{+\infty} te^{n(\theta-t)} dt = \theta + \frac{1}{n}.$$

Отже, незміщеною оцінкою параметра  $\theta$  буде  $\hat{\theta} = x_{(1)} - \frac{1}{n}$ . Спроможність статистики  $x_{(1)} - \frac{1}{n}$  перевірте, будь ласка, самостійно, як і те, що

$$D\left(x_{(1)} - \frac{1}{n}\right) = \frac{1}{n^2}.$$

Наразі розглянемо приклад дискретного розподілу.

**Приклад 4.4.** Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу Пуассона  $\Pi(\theta)$ , параметр якого невідомий. Перевіримо, що  $\theta^* = \bar{x}$ .

Запишемо функцію правдоподібності

$$L(x, \theta) = e^{-n\theta} \prod_{k=1}^n \frac{\theta^{x_k}}{x_k!},$$

функцію впливу одного спостереження

$$h(x_1, \theta) = -\frac{1}{\theta} + \frac{x_1}{\theta}$$

та його інформацію Фішера

$$I_1(\theta) = \frac{Dx_1}{\theta^2} = \frac{1}{\theta}.$$

Звідси  $I(\theta) = n/\theta$ .

Виконуються умови леми 4.1 і за нерівністю Рао-Крамера дисперсія будь-якої незміщеної оцінки не менша ніж  $\theta/n$ . Але

$$D\bar{x} = D\left(\frac{1}{n}\sum_{i=1}^n x_i\right) = \frac{n\theta}{n^2} = \frac{\theta}{n},$$

тому  $\theta^* = \bar{x}$  – найкраща оцінка для  $\theta$ .

**Контрольні запитання:**

1. Скільки існує можливих оцінок невідомого параметра розподілу?
2. Як порівнюють оцінки?
3. Яку оцінку в класі незміщених вважають найкращою?
4. В чому полягає алгоритм пошуку ефективних оцінок параметрів дискретних розподілів? Абсолютно неперервних?
5. Нагадайте, будь ласка, означення інтегралу Лебега за ймовірнісною мірою та способи його обчислення для дискретних та абсолютно неперервних мір.
6. Запишіть нерівність Рао-Крамера та різні формули подання інформації Фішера.

**Вправи та задачі: 4.1.** Нехай  $(x_1, \dots, x_n)$  – вибірка з рівномірного розподілу  $U[\theta, \theta + 5]$ , де  $\theta \in \mathbb{R}$  – невідомий параметр.

Порівняйте оцінки  $\hat{\theta}_1 = x_{(n)}$  та  $\hat{\theta}_1 = x_{(1)}$  в середньоквадратичному. Порівняйте ці оцінки з ОММ  $\hat{\theta}_3 = \bar{x} - 2, 5$ .

Чи буде сімейство  $U[\theta, \theta + 5]$ ,  $\theta \in \mathbb{R}$  регулярним?

**4.2.** Нехай  $(x_1, \dots, x_n)$  – вибірка з показникового розподілу  $\Gamma(1/\theta)$ , де  $\theta > 0$  – невідомий параметр.

Знайдіть, будь ласка, ефективну оцінку для  $\theta$ .

**4.3.** Нехай  $(x_1, \dots, x_n)$  – вибірка з біноміального розподілу  $B(m, \theta)$ , де  $m \in \mathbb{N}$  – відоме число, а  $0 < \theta < 1$  – невідомий параметр. Знайдіть ефективну оцінку для  $\theta$ .

—

## §5. Інтервальні оцінки параметрів розподілу

### 5.1. Поняття довірчого інтервалу

Методи моментів та максимальної правдоподібності дозволяють знаходити точкові оцінки параметрів за вибірковими даними. На конкретній реалізації  $(x_1, \dots, x_n)$  точкова оцінка  $\hat{\theta}(x_1, \dots, x_n)$  – число. Наприклад,  $\hat{\theta}(x_1, \dots, x_n) = 2,652$ .

Навіть якщо  $\hat{\theta}$  – незміщена, спроможна та ефективна оцінка параметра  $\theta \in R$ , обсяг  $n$  досить великий, чи можна стверджувати, що істинне значення  $\theta = 2,652$ ? Певно, що ні. Оскільки  $\hat{\theta}$  – випадкова величина і всі її характеристики не більше ніж імовірнісні, а реалізація  $(x_1, \dots, x_n)$  – одна з багатьох можливих. Якщо  $\hat{\theta}$  має абсолютно неперервний розподіл, то взагалі  $P\{\hat{\theta} = \theta\} = 0$ .

Оцінка похибки ? – питання ймовірнісне. Для заданого  $\varepsilon > 0$  можна знаходити ймовірність, з якою припускаємось похибки величиною  $\varepsilon$ :

$$\forall \varepsilon > 0 \quad P\{|\hat{\theta} - \theta| < \varepsilon\} = \gamma \quad | \quad \gamma = ?$$

А можна розглядати зворотню задачу: для заданого  $\gamma \in (0, 1)$  будемо шукати таке  $\varepsilon$ , що абсолютна похибка величиною в  $\varepsilon$  можлива з імовірністю  $\gamma$ :

$$\forall \gamma \in (0, 1) \quad P\{|\hat{\theta} - \theta| < \varepsilon\} = \gamma \quad | \quad \varepsilon = ? \quad (5.1)$$

Саме цим і зацікавимося.

Перепишемо ймовірність в (5.1) у вигляді

$$P\{\hat{\theta} - \varepsilon < \theta < \hat{\theta} + \varepsilon\} = \gamma$$

і задачу сформулюємо так: для заданого  $\gamma$  треба знайти такий випадковий інтервал  $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$ , котрий містить у собі (кажуть ще – покриває) невідомий параметр  $\theta$  з імовірністю  $\gamma$ .

Випадковий інтервал  $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$  можна розглядати як *інтервальну* оцінку невідомого параметра  $\theta$  надійності  $\gamma$ . Важливі значення  $\gamma$ , близькі до 1, тому покладають  $\gamma = 1 - \alpha$ , де  $\alpha \in (0, 1)$ .

**Означення.** Нехай  $\hat{\theta}_1$  та  $\hat{\theta}_2$  – деякі оцінки, тобто  $\hat{\theta}_i = \hat{\theta}_i(x_1, \dots, x_n)$ ,  $i = \overline{1, 2}$  – спроможні статистики;  $\alpha \in (0, 1)$ . Інтервал  $(\hat{\theta}_1, \hat{\theta}_2)$  називають *довірчим інтервалом* невідомого параметра  $\theta$  надійності  $\gamma = 1 - \alpha$ , якщо

$$P\{\hat{\theta}_1 < \theta < \hat{\theta}_2\} = \gamma.$$

Надійність  $\gamma$ , яку ще називають *довірчою ймовірністю*, задає дослідник. Беруть, наприклад, значення 0,95; 0,99 або 0,999 залежно від того, наскільки високим є очікуваний рівень довіри такій оцінці.

При заданому  $\gamma$  довжина довірчого інтервалу характеризує точність, з якою локалізовано значення параметра. Цінність мають найменші інтервали. Вже незабаром буде зрозуміло, що цього можна досягти збільшенням обсягу вибірки  $n$ . Точність також забезпечується зменшенням розсіювання оцінки  $\hat{\theta}$ . Тому довірчі інтервали будують на основі ефективних оцінок.

Довірчі інтервали не обов'язково симетричні, але розглянемо спочатку саме симетричний випадок (5.1). Припустимо, що  $\hat{\theta}$  – ефективна оцінка,  $\gamma \in (0, 1)$  – фіксоване.

Звернемо увагу на ймовірність в (5.1) та перепишемо її у вигляді:

$$F_{|\hat{\theta} - \theta|}(\varepsilon) = \gamma,$$

де  $F_{|\hat{\theta} - \theta|}(x)$  – функція розподілу випадкової величини  $|\hat{\theta} - \theta|$ .

Звідси, як тільки розподіл випадкової величини  $|\hat{\theta} - \theta|$  відомий, задача пошуку  $\varepsilon$  в (5.1) зводиться до розв'язання рівняння  $F_{|\hat{\theta} - \theta|}(t) = \gamma$ . У зв'язку з останнім розглянемо поняття квантиля.

**Означення.** Нехай  $P_\xi$  – абсолютно неперервний розподіл. Розв'язок рівняння  $F_\xi(t) = p$  називають *p-квантилем* або *квантилем рівня p* розподілу  $P_\xi$  і позначають, наприклад,  $k_p$ , якщо

$$F_\xi(k_p) = p.$$

### 5.2. Довірчий інтервал для математичного сподівання нормального розподілу з відомою дисперсією

Нехай  $\xi \in N(\theta, \sigma^2)$ , де  $\sigma^2 > 0$  – відомий, а  $\theta$  – невідомий параметри;  $(x_1, \dots, x_n)$  – вибірка з даного розподілу. Задамо деяке  $\gamma \in (0, 1)$  і побудуємо довірчий інтервал параметра  $\theta$  надійності  $\gamma$ .

Ефективною оцінкою математичного сподівання нормальної випадкової величини за відомої дисперсії є

$$\hat{\theta} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

За умовою, кожна випадкова величина  $x_i$ ,  $i = \overline{1, n}$  однаково розподілена з  $\xi$ , тому  $\bar{x}$  також має нормальний розподіл з параметрами

$$M\bar{x} = \frac{1}{n} \sum_{i=1}^n Mx_i = \theta \quad \text{і} \quad D\bar{x} = \frac{1}{n^2} \sum_{i=1}^n Dx_i = \frac{\sigma^2}{n}.$$

Відповідно,  $(\bar{x} - \theta)\sqrt{n}/\sigma \in N(0, 1)$  – стандартна випадкова величина.

Тоді

$$\begin{aligned} \forall \gamma \in (0, 1) \quad P\{|\bar{x} - \theta| < \varepsilon\} &= P\left\{\frac{|\bar{x} - \theta|\sqrt{n}}{\sigma} < \frac{\varepsilon\sqrt{n}}{\sigma}\right\} = \\ &= F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - F\left(-\frac{\varepsilon\sqrt{n}}{\sigma}\right) = 2\left(F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - \frac{1}{2}\right) = \\ &= 2F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1 = 1 - \alpha \end{aligned}$$

і шукане  $\varepsilon$  є розв'язком рівняння

$$F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = 1 - \frac{\alpha}{2},$$

де  $F$  – функція розподілу стандартної випадкової величини.

За означенням,

$$\frac{\varepsilon\sqrt{n}}{\sigma} = d_{1-\frac{\alpha}{2}}$$

– квантиль рівня  $(1 - \frac{\alpha}{2})$  розподілу  $N(0, 1)$ . Звідси знайдемо  $\varepsilon$  і запишемо шуканий довірчий інтервал надійності  $\gamma = 1 - \alpha$ :

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}}d_{1-\frac{\alpha}{2}}, \bar{x} + \frac{\sigma}{\sqrt{n}}d_{1-\frac{\alpha}{2}}\right). \quad (5.2)$$

Тобто

$$P \left\{ \bar{x} - \frac{\sigma}{\sqrt{n}} d_{1-\frac{\alpha}{2}} < \theta < \bar{x} + \frac{\sigma}{\sqrt{n}} d_{1-\frac{\alpha}{2}} \right\} = 1 - \alpha.$$

Квантилі стандартного розподілу наведені в табл. II додатку.

**Приклад 5.1.** Нехай  $n = 400$ ,  $\bar{x} = 25,5$ , а  $\sigma^2 = 2,25$ . Побудуємо довірчі інтервали для математичного сподівання розподілу  $N(\theta; 2, 25)$  надійності 0,9; 0,95 та 0,98.

За таблицею II знайдемо потрібні квантилі:

$$d_{1-\frac{0,1}{2}} = 1,64, \quad d_{1-\frac{0,05}{2}} = 1,96, \quad \text{і} \quad d_{1-\frac{0,02}{2}} = 2,33.$$

Відповідно, маємо інтервали:

$$\left( 25,5 - \frac{1,5}{20} \cdot 1,64; 25,5 + \frac{1,5}{20} \cdot 1,64 \right) = (25,38; 25,62) \quad \text{для } \gamma = 0,9,$$

$$\left( 25,5 - \frac{1,5}{20} \cdot 1,96; 25,5 + \frac{1,5}{20} \cdot 1,96 \right) = (25,35; 25,65), \quad \text{для } \gamma = 0,95,$$

та

$$\left( 25,5 - \frac{1,5}{20} \cdot 2,33; 25,5 + \frac{1,5}{20} \cdot 2,33 \right) = (25,33; 25,67) \quad \text{для } \gamma = 0,98.$$

Зрозуміло, що зі збільшенням надійності довірчий інтервал збільшується.

При заданій ймовірності  $\gamma = 1 - \alpha$  зменшити довжину довірчого інтервалу, вочевидь, можна за рахунок збільшення числа спостережень  $n$ . Це підвищить точність інтервальної оцінки.

Припустимо, що є можливість зробити стільки спостережень, скільки необхідно. Тоді виникає така задача: при заданій надійності  $\gamma = 1 - \alpha$  довірчого інтервалу в (5.2) треба визначити найменше число спостережень  $n$ , яке забезпечує точність оцінки  $\varepsilon$ .

За умовою необхідно, щоб виконувалась нерівність

$$\sigma d_{1-\frac{\alpha}{2}} / \sqrt{n} < \varepsilon, \quad \text{звідки } n \geq (\sigma d_{1-\frac{\alpha}{2}} / \varepsilon)^2.$$

Зокрема, для  $\gamma = 0,9$ ,  $\sigma^2 = 2,25$  і  $\varepsilon = 0,1$  маємо:  $n \geq 606$ .

Ключову роль при побудові довірчого інтервалу (5.2) відіграла статистика  $(\bar{x} - \theta)\sqrt{n}/\sigma$ , розподіл якої виявився відомим і, що важливо, не

залежним від параметра  $\theta$ . Наразі, довірчі інтервали будують саме за такою схемою:

1) підбирають статистику  $g(x_1, \dots, x_n; \theta)$ , розподіл якої не залежить від невідомого параметра  $\theta$ . Необхідно, щоб  $g(x_1, \dots, x_n; \theta)$  мала обернену відносно  $\theta$  для будь-яких реалізацій  $(x_1, \dots, x_n) \in X$ ;

2) знаходять квантілі  $g_1$  і  $g_2$  деяких рівнів значущості розподілу  $P_g$ , для яких

$$P\{g_1 < g(x_1, \dots, x_n; \theta) < g_2\} = \gamma = 1 - \alpha;$$

3) розв'язують нерівність  $g_1 < g(x_1, \dots, x_n; \theta) < g_2$  відносно  $\theta$  і отримують довірчий інтервал для параметра  $\theta$  надійності  $\gamma$ .

Найскладніший етап – вибір потрібної статистики та визначення її розподілу. Задача ускладнюється, якщо теоретичний розподіл має більше одного невідомого параметра (розподіл необхідної статистики не повинен залежати від жодного з них).

Поставимо собі за мету побудову довірчих інтервалів для параметрів нормального розподілу, коли вони обидва невідомі. Далі розглядаються розподіли, тісно пов'язані з нормальним, які вирішують це питання.

### 5.3. Розподіли, пов'язані з нормальним

**Означення.** Розподілом Пірсона  $\chi^2(n)$  (або розподілом  $\chi^2$ ) з  $n$  ступенями свободи називають розподіл випадкової величини

$$\chi^2(n) = \xi_1^2 + \dots + \xi_n^2,$$

де  $\xi_1, \dots, \xi_n$  – незалежні випадкові величини, кожна з яких стандартна,  $\xi_i \in N(0, 1)$ ,  $i = \overline{1, n}$ .

Розподіл випадкової величини  $\chi^2(n)$  визначається єдиним параметром  $n$ , а її щільність має вигляд:

$$p(t) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} t^{n/2-1} e^{-t/2}, & t \geq 0; \\ 0, & t < 0, \end{cases}$$

де  $\Gamma(\alpha) = \int_0^{+\infty} u^{\alpha-1} e^{-u} du$  – гамма-функція.



Випадкова величина  $\chi^2(n)$  задовольняє ЦГТ, згідно з якою

$$\frac{\chi^2(n) - M\chi^2(n)}{\sqrt{D\chi^2(n)}} \xrightarrow{d} \eta \in N(0, 1).$$

Неважко підрахувати, що

$$M\chi^2(n) = \sum_{i=1}^n M\xi_i^2 = n,$$

а

$$D\chi^2(n) = \sum_{i=1}^n D\xi_i^2 = \sum_{i=1}^n M\xi_i^4 - n = 3n - n = 2n,$$

оскільки

$$\begin{aligned} M\xi_i^4 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^4 e^{-\frac{x^2}{2}} dx = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^3 d\left(e^{-\frac{x^2}{2}}\right) = \\ &= \frac{3}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx = 3D\xi = 3. \end{aligned}$$

У такий спосіб, випадкова величина

$$\frac{\chi^2(n) - n}{\sqrt{2n}}$$

асимптотично нормальна з коефіцієнтом 1. Тому, якщо ступенів свободи більше ніж 30, то замість розподілу  $\chi^2$  з  $n$  ступенями свободи використовують розподіл  $N(n, 2n)$ .

**Означення.** Розподілом  $\chi(n)$  (або розподілом  $\chi$ ) з  $n$  ступенями свободи називають розподіл випадкової величини  $\chi(n) = \sqrt{\chi^2(n)}$ .

Розподілом Стьюдента<sup>1</sup>  $T(n)$  (або  $t$ -розподілом) з  $n$  ступенями свободи називають розподіл випадкової величини

$$t(n) = \frac{\eta}{\sqrt{\chi^2(n)/n}},$$

де  $\eta \in N(0, 1)$  та  $\chi^2(n)$  – незалежні випадкові величини.

---

<sup>1</sup>Стьюдент - псевдонім В. С. Госета.

Щільність розподілу Стюдента також визначається єдиним параметром  $n$  і має не зовсім простий вигляд:

$$p(t) = \frac{\Gamma((n+1)/2)}{\sqrt{\pi n} \Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}. \quad (5.3)$$

За посиленням законом великих чисел,

$$\frac{1}{n} \chi^2(n) = \frac{1}{n} \sum_{i=1}^n \xi_i^2 \xrightarrow{\text{М.Н.}} \frac{1}{n} \sum_{i=1}^n M \xi_i^2 = 1,$$

тому випадкова величина  $t(n)$  зі зростанням  $n$ ,  $n \rightarrow \infty$ , має розподіл, близький до стандартного  $N(0, 1)$ .

Наступний розподіл також тісно пов'язаний з нормальним. Він знадобиться пізніше, у другому розділі, п. 9.4 при побудові статистичних критеріїв.

**Означення.** Нехай випадкові величини  $\chi^2(n)$  та  $\eta^2(m)$  мають розподіл  $\chi^2$  з  $n$  та  $m$  ступенями свободи і є незалежними. Розподіл випадкової величини

$$f(n, m) = \frac{\frac{1}{n} \chi^2(n)}{\frac{1}{m} \chi^2(m)}$$

називають *розподілом Фішера з  $(n, m)$  ступенями свободи* і позначають  $\Phi(n, m)$ .

Властивості розподілу  $\Phi(n, m)$ :

- 1) якщо  $f(n, m) \in \Phi(n, m)$ , то  $\frac{1}{f(n, m)} \in \Phi(m, n)$ ;
- 2) розподіли  $\Phi(n, m)$  та  $\Phi(m, n)$  пов'язані співвідношеннями:

$$\begin{aligned} F_{n,m}(x) &= P\{f(n, m) < x\} = P\left\{\frac{1}{f(n, m)} > \frac{1}{x}\right\} = \\ &= 1 - F_{m,n}\left(\frac{1}{x}\right); \end{aligned}$$

- 3) якщо випадкова величина  $t(k)$  має розподіл Стюдента з ступенями свободи, то  $t^2(k) \in \Phi(1, k)$ .

#### 5.4. Довірчий інтервал для математичного сподівання нормального розподілу з невідомою дисперсією

Нехай  $(x_1, \dots, x_n)$  – вибірка з нормального розподілу  $N(\theta_1, \theta_2^2)$ , обидва параметри  $\theta_1, \theta_2$  якого невідомі. Визначимо довірчий інтервал для  $\theta_1$ .

При розв'язанні аналогічної задачі з відомою дисперсією, основою наших міркувань була статистика  $(\bar{x} - \theta_1)\sqrt{n}/\sigma$ . У випадку невідомої дисперсії скористаємось її найкращою оцінкою – варіансою

$$\bar{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

і розглянемо статистику

$$\frac{\bar{x} - \theta_1}{\bar{s}} \sqrt{n} = \frac{(\bar{x} - \theta_1)\sqrt{n}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}} = \frac{\frac{\bar{x} - \theta_1}{\theta_2} \sqrt{n}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\theta_2}\right)^2}}. \quad (5.4)$$

Оскільки

$$\frac{\bar{x} - \theta_1}{\theta_2} \sqrt{n} \in N(0, 1),$$

статистика (5.4) нагадує розподіл Стюдента. Перша відмінність – в знаменнику у сумі  $n$  доданків, а множник перед сумою  $1/(n-1)$ . Більш глибока відмінність – випадкові величини  $x_i - \bar{x}$ ,  $i = \overline{1, n}$  не є незалежними. Вони алгебраїчно лінійно залежні, оскільки

$$\sum_{i=1}^n (x_i - \bar{x}) = 0.$$

Але справедлива наступна лема.

**Лема 5.1.** Якщо  $(x_1, \dots, x_n)$  – вибірка з розподілу  $N(\theta_1, \theta_2^2)$ ,  $\bar{x}$  – вибіркове середнє значення, то

$$\frac{1}{\theta_2^2} \sum_{i=1}^n (x_i - \bar{x})^2 \in \chi^2(n-1).$$

**Доведення.** Побудуємо нові випадкові величини:

$$y_1 = \frac{1}{\sqrt{1 \cdot 2}}(x_1 - x_2),$$

$$y_2 = \frac{1}{\sqrt{2 \cdot 3}}(x_1 + x_2 - 2x_3),$$

$$\vdots$$

$$y_{n-1} = \frac{1}{\sqrt{(n-1) \cdot n}}(x_1 + x_2 + \dots + x_{n-1} - (n-1)x_n),$$

$$y_n = \frac{x_1 + x_2 + \dots + x_n}{\sqrt{n}}.$$

Неважко довести, що

$$\sum_{i=1}^n y_i^2 = \sum_{i=1}^n x_i^2.$$

Наприклад, коефіцієнт при  $x_1^2$  у сумі зліва дорівнює:

$$\frac{1}{2} + \frac{1}{2 \cdot 3} + \dots + \frac{1}{n(n-1)} + \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{n-1} - \frac{1}{n} + \frac{1}{n} = 1,$$

а коефіцієнт при  $x_1 \cdot x_2$  нульовий:

$$\begin{aligned} & 2 \left[ -\frac{1}{2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \dots + \frac{1}{n(n-1)} + \frac{1}{n} \right] = \\ & = 2 \left[ -\frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{n-1} - \frac{1}{n} + \frac{1}{n} \right] = 0. \end{aligned}$$

Аналогічно для решти.

Далі,

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2 = \sum_{i=1}^n y_i^2 - n\bar{x}^2 = \sum_{i=1}^n y_i^2 - y_n^2 = \sum_{i=1}^{n-1} y_i^2.$$

Таким чином,

$$\frac{1}{\theta_2^2} \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^{n-1} \left( \frac{y_i}{\theta_2^2} \right)^2.$$

Перевіримо незалежність випадкових величин  $y_i$ ,  $i = \overline{1, n-1}$ . По-перше, легко бачити, що випадкові величини  $y_i$ ,  $i = \overline{1, n-1}$  лінійно залежними алгебраїчно бути не можуть.

Далі, оскільки незалежність гаусівського вектора еквівалентна їх некорильованості,<sup>2</sup> достатньо проаналізувати коваріаційну матрицю. Підрахуємо, наприклад,

$$\begin{aligned}\text{cov}(y_1, y_2) &= M(y_1 y_2) - M y_1 M y_2 = \\ &= \frac{1}{2\sqrt{3}} M(x_1^2 + x_1 x_2 - 2x_1 x_3 - x_1 x_2 - x_2^2 + 2x_2 x_3) = 0.\end{aligned}$$

Аналогічно для кожної іншої пари. Таким чином, випадкові величини  $y_i$ ,  $i = \overline{1, n-1}$ , які є гаусівськими та некорильованими, – попарно незалежні.

Знайдемо параметри їх розподілу. Вочевидь,

$$M y_i = \frac{\theta_1}{\sqrt{i(i-1)}}(i - i) = 0, \quad i = \overline{1, n-1},$$

а

$$D y_i = \frac{\theta_2^2}{i(i+1)}(i + i^2) = \theta_2^2, \quad i = \overline{1, n-1}.$$

У підсумку, випадкові величини  $y_1, y_2, \dots, y_{n-1}$  попарно незалежні і  $y_i \in N(0, \theta_2^2)$ ,  $i = \overline{1, n-1}$ . Тому

$$\frac{1}{\theta_2^2} \sum_{i=1}^{n-1} y_i^2 \in \chi^2(n-1). \quad \square$$

**Наслідок.**

$$\frac{\bar{x} - \theta_1}{\bar{s}} \sqrt{n} \in T(n-1).$$

**Зауваження 5.1.** Квадратична форма  $\sum_{i=1}^n (x_i - \bar{x})^2$  має ранг  $n-1$ , оскільки в ході доведення леми шляхом лінійного перетворення зведена до  $(n-1)$ -ї кількості незалежних змінних. Тобто, мовою теорії ймовірностей, має  $n-1$  ступенів свободи. До речі, це пояснює незміщеність саме варіанси як оцінки дисперсії розподілу.

Для побудови довірчого інтервалу параметра  $\theta_1$  розглянемо ймовірність

$$P \left\{ \frac{|\bar{x} - \theta_1|}{\bar{s}} < \varepsilon \right\} = 1 - \alpha.$$

---

<sup>2</sup>На згадку з теорії ймовірностей: із незалежності випадкових величин завжди ви-тікає їх некорильованість, а для гаусівських величин справедливо і навпаки. Тобто нормальні випадкові величини можуть бути тільки лінійно залежними.

Далі, міркуючи так, як і при побудові інтервалу (5.2) (щільність розподілу Стюдента – парна функція, як і для стандартного розподілу) отримаємо довірчий інтервал надійності  $1 - \alpha$  для математичного сподівання нормального розподілу, коли дисперсія невідома:

$$\left( \bar{x} - \frac{\bar{s}}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1), \bar{x} + \frac{\bar{s}}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \right). \quad (5.5)$$

Тут  $n$  – обсяг вибірки,  $t_{1-\frac{\alpha}{2}}(n-1)$  – квантиль рівня  $(1 - \frac{\alpha}{2})$  розподілу Стюдента з  $(n-1)$  ступенями свободи,  $\bar{x}$  – вибіркове середнє значення, а  $\bar{s}$  – флуктуація.

Квантилі розподілу Стюдента наведені в таблиці III додатку.

При  $n \geq 30$  довірчий інтервал (5.5) трансформується в такий:

$$\left( \bar{x} - \frac{\bar{S}}{\sqrt{n}} d_{1-\frac{\alpha}{2}}, \bar{x} + \frac{\bar{S}}{\sqrt{n}} d_{1-\frac{\alpha}{2}} \right),$$

де  $\bar{S}$  – емпіричне середньоквадратичне відхилення,  $d_{1-\frac{\alpha}{2}}$  – квантиль рівня  $(1 - \frac{\alpha}{2})$  стандартного розподілу.

### 5.5. Довірчий інтервал для середньоквадратичного відхилення нормального розподілу

Нехай  $(x_1, \dots, x_n)$  – вибірка з нормального розподілу  $N(\theta_1, \theta_2^2)$ , обидва параметри  $\theta_1, \theta_2$  якого невідомі. Побудуємо довірчий інтервал надійності  $\gamma = 1 - \alpha$  для середньоквадратичного відхилення  $\theta_2$ .

Розглянемо статистику

$$\frac{\bar{s}^2}{\theta_2^2} = \frac{1}{n-1} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{\theta_2} \right)^2,$$

де  $\bar{x}$  – вибіркове середнє, а  $\bar{s}^2$  – варіанса.

За лемою 5.1,

$$(n-1) \frac{\bar{s}^2}{\theta_2^2} \in \chi^2(n-1),$$

а

$$\sqrt{n-1} \frac{\bar{s}}{\theta_2} \in \chi(n-1).$$

Відмінність цього випадку від попередніх у тому, що випадкова величина з розподілом  $\chi(n-1)$  приймає тільки невід'ємні значення і довірчий інтервал вийде несиметричний. Точніше, симетричний, але не відносно нуля.

Задамо довірчу ймовірність  $\gamma = 1 - \alpha$  і нехай  $z_1, z_2$  такі, що

$$P \left\{ z_1 < \sqrt{n-1} \frac{\bar{s}}{\theta_2} < z_2 \right\} = \gamma. \quad (5.6)$$

Звідси

$$P \left\{ \sqrt{n-1} \frac{\bar{s}}{z_2} < \theta_2 < \sqrt{n-1} \frac{\bar{s}}{z_1} \right\} = \gamma.$$

У цьому співвідношенні дві невідомі маємо на меті звести до однієї. Це можна зробити, якщо шукати довірчий інтервал виду  $(\bar{s} - \delta, \bar{s} + \delta)$ . Тоді

$$\sqrt{n-1} \frac{\bar{s}}{z_2} = \bar{s} - \delta,$$

$$z_2 = \sqrt{n-1} \frac{\bar{s}}{\bar{s} - \delta} = \frac{\sqrt{n-1}}{1 - \delta/\bar{s}} \equiv \frac{\sqrt{n-1}}{1 - q},$$

де  $q = \delta/\bar{s}$ . Аналогічно,

$$z_1 = \frac{\sqrt{n-1}}{1 + q}.$$

Отже, співвідношення (5.6) тепер можна переписати як рівняння відносно  $q$ :

$$\int_{\frac{\sqrt{n-1}}{1+q}}^{\frac{\sqrt{n-1}}{1-q}} p_{\chi}(t) dt = \gamma, \quad (5.7)$$

де  $p_{\chi}(t)$  – щільність розподілу  $\chi(n-1)$ .

Розв'язки  $q = q(\gamma, n)$  таких рівнянь табульовані. З таблиці IV додатку знайдемо відповідне значення  $q = q(\gamma, n)$ . Тоді  $\delta = sq$  і, отже,

$$P \{ \theta_2 \in (\max\{0, s(1 - q)\}, s(1 + q)) \} = 1 - \alpha.$$

### Контрольні запитання:

1. Охарактеризуйте, будь ласка, суть інтервальної оцінки.
2. Що називають довірчим інтервалом заданої надійності для невідомого параметра розподілу?

**3.** Що характеризує точність інтервальної оцінки і від чого ця точність залежить, як її підвищити?

**4.** Чи можна сформулювати основну схему побудови інтервальної оцінки?

**Вправи та задачі: 5.1.** Який квантиль стандартного нормального розподілу більше: рівня 0,05 чи рівня 0,1? Чому?

**5.2.** Нехай  $(x_1, \dots, x_n)$  – вибірка з рівномірного розподілу  $U[0, \theta]$ . Побудувати довірчі інтервали для параметра  $\theta$  на основі статистик  $x_{(1)}$  і  $x_{(n)}$ .

**5.3.** Нехай  $(x_1, \dots, x_n)$  – вибірка з показникового розподілу  $\Gamma(\theta)$ . Побудувати довірчі інтервали для параметра  $\theta$  на основі статистик  $x_1$  і  $x_{(n)}$ .

**5.4.** Знайти математичне сподівання та дисперсію довжини довірчого інтервалу для середнього нормального розподілу при невідомій дисперсії.



## 2. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

Науку треба вважати системою гіпотез,  
з якими ми погоджуємося доти,  
доки вони витримують перевірку.

К. Поппер

### §6. Гіпотези та критерії

Щоб правильно поставити запитання,  
потрібно знати більшу частину відповіді.

Р. Шеклі

#### **6.1. Поняття статистичної гіпотези. Класифікація помилок.**

Статистичні методи дозволяють отримати відповіді на різні питання дійсності на основі вибірових даних. Деякі приклади розглядались у вступі. Повернемося до них більш ґрунтовно і наведемо ще один новий.

Припустимо, що нас цікавить:

- 1) чи зменшує вакцинація певною вакциною  $B$  ризик захворюваності на COVID-19?
- 2) чи стабільна робота певного підприємства із виробництва прогоничів класу міцності 5,8 стосовно відповідності їх міцності на розтяг заявленому

номінальному значенню  $500 \text{ н/мм}^2$ ?

3) чи покращить міцність прогоничів нова технологія виготовлення?

Щоб відповісти на поставлені питання, треба добре зрозуміти як проблему в цілому, так і її математичну суть, тобто побудувати слушну математичну модель.

Стосовно питання 1). Якщо частка людей, що захворіли в певних групах і в різні моменти часу  $t \in T$ , коливається навколо деякого числа  $p$ , то  $p$  називають *рівнем захворюваності* в цих групах за період  $T$ .

Таким чином, треба зіставити рівні захворюваності на COVID-19 серед нещеплених та щеплених вакциною  $B$  людей за один і той самий проміжок часу  $T$  (за тиждень, місяць і т.і.)

З точки зору теорії ймовірностей, рівень захворюваності в певній групі людей за період – це ймовірність виявити хворого при випадковому виборі когось одного з цієї групи в будь-який момент  $t \in T$ .

Якщо ж для обстеження випадково обирається одна людина, то наслідком такого експерименту є одне випробування в схемі Бернуллі з імовірністю успіху, що дорівнює  $p$  ('успіх' полягає у виявленні хворого на COVID-19). Тоді задача, що нас цікавить, зводиться до порівняння ймовірностей  $p_0$  та  $p_1$  двох різних схем Бернуллі: для невакцинованих і вакцинованих людей, відповідно.

Інший спосіб з'ясувати питання 1) – перевірити, чи залежні між собою дві ознаки:  $\xi$  – наявність щеплення вакциною  $B$  у людини та  $\eta$  – показник того, чи хворіла вона на COVID-19 протягом деякого проміжку часу  $T$  та в якій формі.

Стосовно питання 2). Як зазначалось у вступі, ознакою стабільності вказаного виробництва є нормальний розподіл випадкової величини, що характеризує міцність виготовленого прогонича. Нормальний розподіл, згідно з ЦГТ, свідчить про те, що відхилення від стандарту (які, звичайно, бувають у будь-якому виробництві) мають дійсно випадковий характер. Тобто, з математичної точки зору, 2) – це питання про розподіл відповідної випадкової величини.

І, нарешті, щодо питання 3). Тут доцільно в'яснити розподіли двох

випадкових величин  $\xi_0$  та  $\xi_1$ , що характеризують міцність прогоничів за старою і новою технологією, відповідно. Якщо обидва процеси виробництва стабільні, очікувані нормальні розподіли. Тоді задача зводиться до порівняння їх середніх значень і, за необхідності, дисперсій.

Наступний етап – формування репрезентативних вибірок та обробка вибірових даних з метою визначення емпіричних характеристик розподілів. Як вже зазначалось, емпіричні характеристики можуть багато розповісти (див. п. 2.5). На їх основі формують припущення (гіпотези) щодо задач, які розглядаються:

1) щоб порівняти рівні захворюваності нещеплених і щеплених людей, зіставимо спочатку частку захворілих  $\bar{p}_0$  і  $\bar{p}_1$  у вибірках з відповідних генеральних сукупностей. Якщо  $\bar{p}_0 > \bar{p}_1$ , то це ще не означає, що вакцина дієва (оскільки виборки випадкові), але дає підґрунтя для формування основної гіпотези у вигляді:  $H_0 : p_0 > p_1$  за альтернативи;  $H_1 : p_0 < p_1$ . Якщо ж  $\bar{p}_0 \approx \bar{p}_1$ , то основною гіпотезою буде:  $H_0 : p_0 = p_1$  проти альтернативи  $H_1 : p_0 > p_1$ ;

при перевірці залежності наявності щеплення  $\xi$  вакциною  $B$  у людини та показника  $\eta$ , чи перехворіла вона на COVID-19, гіпотезами будуть  $H_0 : \xi$  та  $\eta$  залежні, а  $H_1 = \bar{H}_0$ .

2) у задачі про розподіл основну гіпотезу  $H_0$  формують на основі апріорної інформації шляхом аналізу емпіричних характеристик (див. п. 2.5)

Коли немає додаткової інформації про параметри, гіпотези висувають, наприклад, таким чином:

$$H_0 : P_\xi = N(\bar{x}, \bar{s}^2) \text{ проти альтернативи } H_1 = \bar{H}_0 : P_\xi \neq N(\bar{x}, \bar{s}^2)$$

(за параметри гіпотетичного розподілу  $P_\xi$  обирають значення найкращих для них оцінок на конкретній реалізації);

3) при порівнянні середніх значень нормальних розподілів  $N(\theta_1, \theta_2^2)$  та  $N(\theta_3, \theta_4^2)$ , параметри яких невідомі, або відомі лише дисперсії, підставою для формування основної гіпотези  $H_0$  є порівняння середніх емпіричних  $\bar{x}_0$  та  $\bar{x}_1$  реалізацій вибірок з цих розподілів. Якщо виявилось, що  $\bar{x}_0 < \bar{x}_1$ , то як основну гіпотезу беремо:  $H_0 : \theta_1 < \theta_3$  за альтернативної  $H_1 : \theta_1 > \theta_3$ .

Можливо і так, що  $\bar{x}_0 \approx \bar{x}_1$ . Тоді основною буде  $H_0 : \theta_1 = \theta_3$ , а альтернативною, залежно від наших уявлень про задачу, може бути  $H_1 : \theta_1 \neq \theta_3$  (так звана двостороння альтернатива) або одна з односторонніх:  $H_1 : \theta_1 < \theta_3$  чи  $H_1 : \theta_1 > \theta_3$ .

Перейдемо до загальних означень.

**Означення.** Гіпотези, які можна перевірити на основі вибірових даних, називають *статистичними*.

Саме статистичні гіпотези вивчатимемо далі.

Одну з гіпотез виділяють як основну і позначають  $H_0$ . Інші називають *альтернативними* і позначають  $H_1$ . У подальшому обмежимося випадком однієї альтернативи  $H_1$ .

Якщо гіпотетичне твердження  $H_i$  однозначне, то гіпотезу  $H_i$  називають *простою* ( $i = 0, 1$ ). В іншому разі – складною.

Наприклад,  $H_0 : P_\xi = N(2, 1)$  – проста гіпотеза, оскільки вказує на єдиний розподіл, а  $\bar{H}_0 : P_\xi \neq N(2, 1)$  – складна, бо вказує на множину розподілів, відмінних від  $N(2, 1)$ .

Складними будуть обидві гіпотези  $H_0 : \theta_1 = \theta_3$  та  $H_1 : \theta_1 \neq \theta_3$  якщо порівнювати середні значення двох нормальних розподілів  $N(\theta_1, \theta_2^2)$  і  $N(\theta_3, \theta_4^2)$ , параметри яких невідомі.

Серед альтернативних гіпотез розрізняють *односторонні* та *двосторонні*. Так  $H_1 : \theta \neq 2$  – двостороння альтернатива до  $H_0 : \theta = 2$ , а  $H_1 : \theta > 2$  – правостороння складна альтернативна гіпотеза.

**Означення.** Гіпотезу називають *параметричною*, якщо вона містить деяке твердження про значення параметра певного розподілу.

Основну гіпотезу  $H_0$  та альтернативу до неї  $H_1$  формують на основі апріорної інформації про генеральну сукупність та випадкову величину, що вивчається, шляхом аналізу емпіричних характеристик, гістограми репрезентативної вибірки з даної генеральної сукупності.

Наступний крок – перевірка пари гіпотез  $\{H_0, H_1\}$  (основної і альтернативної до неї). Особливість статистичних методів полягає у тому, що необхідний висновок треба зробити за однією реалізацією вибірових даних (яка є просто однією з можливих реалізацій). Зрозуміло, що у такому

разі важко зробити безпомилкові висновки.

**Означення.** Помилку, яка полягає в тому, що відхиляється справедлива основна гіпотеза (і приймається хибна альтернативна), називають *помилкою першого роду*.

Якщо ж відхиляється правильна альтернатива  $H_1$  (тобто приймається хибна основна гіпотеза  $H_0$ ), то таку помилку називають *помилкою другого роду*.

Наслідки помилок першого та другого роду можуть бути різними і дослідник завжди добре розуміє, яка з помилок для нього найбільш небажана. Основну гіпотезу із сукупності усіх гіпотез він обирає сам і, як правило, це та гіпотеза, для якої важливіше уникнути помилки 1-го роду.

## 6.2. Перевірка статистичних гіпотез

Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi$ , відносно якого висунуто дві гіпотези:  $H_0$  – основна і  $H_1$  – альтернативна до неї.

**Означення.** *Статистичним критерієм* перевірки гіпотези  $H_0$  при альтернативі  $H_1$  називають вимірне відображення

$$\delta : X \rightarrow \{H_0, H_1\} \quad (6.8)$$

з множини всіх можливих вибірових значень у множину гіпотез. Позначають  $\delta = \delta(x_1, \dots, x_n)$ . Вимірність розуміється, як зазвичай: у ймовірнісному просторі  $(\Omega, \mathfrak{A}, P)$

$$\{\omega : \delta(x_1, \dots, x_n) = H_i\} \in \mathfrak{A}, \quad \forall i = 0, 1,$$

тобто визначені ймовірності відповідних подій.

Побудова критерію (6.8) зводиться до того, щоб розділити вибіровий простір  $X \subseteq \mathbb{R}^n$  на дві частини  $S$  та  $\bar{S}$ ,  $S \cap \bar{S} = \emptyset$ ,  $S \cup \bar{S} = X$  так, що

$$\delta(x_1, \dots, x_n) = \begin{cases} H_0, & (x_1, \dots, x_n) \in \bar{S}, \\ H_1, & (x_1, \dots, x_n) \in S. \end{cases}$$

У такому разі  $S$  називають *критичною областю*.

**Означення.** Імовірність помилки 1-го роду, тобто ймовірність відхилити правильну основну гіпотезу  $H_0$ , називають *рівнем значущості* критерію і позначають через  $\alpha$ . Тобто

$$\alpha = P\{\delta(x_1, \dots, x_n) = H_1 | H_0\} = P\{(x_1, \dots, x_n) \in S | H_0\}.$$

Рівень значущості  $\alpha$  задає дослідник. Якщо є впевненість, що гіпотеза  $H_0$  справедлива, то потрібні переконливі докази, щоб змінити свою думку і має сенс обирати зовсім маленький рівень значущості: 0,05; 0,01 чи 0,001. Тоді влучити в  $S$  буде вкрай неправдоподібним, якщо гіпотеза  $H_0$  правдива. Коли ж реалізація вибірки таки потрапила в  $S$ , основну гіпотезу відхиляють на користь альтернативної. Зрозуміло, що для кожного рівня значущості  $\alpha$  критична область, загалом, своя:  $S = S(\alpha)$ .

Певний рівень значущості  $\alpha$  критерію означає, що при багаторазовому застосуванні цього критерію для перевірки гіпотези  $H_0$  за альтернативи  $H_1$  помилки 1-го роду можна припуститися не більше ніж у  $100\alpha$  випадках із 100.

Загальний спосіб перевірки основної гіпотези  $H_0$  за альтернативи  $H_1$ :

- згідно з апріорною інформацією про задачу і на підставі емпіричних характеристик за реалізацією вибірки означають основну гіпотезу  $H_0$  та її альтернативу  $H_1$ ;
- задають рівень значущості критерію  $\alpha$ ;
- при заданому рівні значущості  $\alpha$  будують критичну область  $S$ ;
- за умови, що конкретна реалізація вибірки  $(x_1, \dots, x_n)$  належить критичній області  $S$ , основну гіпотезу  $H_0$  відхиляють на користь альтернативи  $H_1$ . Якщо ж реалізація не потрапила в критичну область, то висновок такий: вибіркові дані не протирічать основній гіпотезі.

Для кожної пари  $\{H_0, H_1\}$  на заданому рівні значущості  $\alpha$ , у цілому, будують свій статистичний критерій. До вибору альтернативи  $H_1$  необхідно ставитись уважно, оскільки вона стане основною після відхилення  $H_0$ .

У разі, коли під альтернативною гіпотезою розуміють те, що просто не виконується основна:  $H_1 = \bar{H}_0$ , критерій перевірки гіпотез  $\{H_0, \bar{H}_0\}$

називають *критерієм згоди*. Для критерію згоди альтернативу явно не записують і питання формулюють так: чи узгоджуються результати спостережень з основною гіпотезою?

**Контрольні запитання:**

1. Що таке статистична гіпотеза? На яких засадах формулюють статистичні гіпотези?
2. Охарактеризуйте види статистичних гіпотез. Наведіть, будь ласка, відповідні приклади.
3. Як перевіряють статистичні гіпотези? Чому важливо уважно ставитись до формування і основної гіпотези і альтернативної також?
4. Як класифікують помилки прийняття рішень про справедливість гіпотез?
5. Надайте означення статистичного критерію, критичної області, рівня значущості критерію.
6. Охарактеризуйте загальний спосіб перевірки основної гіпотези  $H_0$  за однієї альтернативи  $H_1$ .

## §7. Критерій згоди $\chi^2$

### 7.1. Критерій $\chi^2$ про розподіл (непараметричний)

Розглянемо один з найстаріших статистичних критеріїв перевірки гіпотези про розподіл – критерій згоди К. Пірсона  $\chi^2$ .

Нехай  $(x_1, \dots, x_n)$  – вибірка з невідомого розподілу  $P_\xi$ . Основна гіпотеза полягає у тому, що спостережувана випадкова величина  $\xi$  має функцію розподілу  $G(x)$  (гіпотетичну):

$$H_0 : F_\xi(x) = G(x), \quad x \in X.$$

Гіпотеза  $H_0$  проста непараметрична. Треба з'ясувати, чи узгоджуються результати спостережень з  $H_0$ .

Позначимо через  $P_G$  розподіл, що визначається функцією  $G(x)$ . Невідомий розподіл  $P_\xi$  і гіпотетичний  $P_G$  збігаються, коли

$$P_\xi(B) = P_G(B), \quad \forall B \in \mathfrak{B}(\mathbb{R}),$$

де  $\mathfrak{B}(\mathbb{R})$  – борелевська  $\sigma$ -алгебра на  $\mathbb{R}$ . Якщо ж  $P_\xi$  та  $P_G$  різні, то знайдеться борелева множина  $B^o \in \mathfrak{B}(\mathbb{R})$  така, що  $P_\xi(B^o) \neq P_G(B^o)$ .

Незалежно від гіпотез, емпірична функція розподілу  $\bar{F}_n(x)$  мало відрізняється від істинної  $F_\xi(x)$ , тому що є для неї спроможною та незміщеною оцінкою (див. теореми 2.1–2.3). Отже, в розумінні сказаного, достатньо порівняти

$$\bar{P}_n(B) \text{ та } P_G(B), \quad \forall B \in \mathfrak{B}(\mathbb{R}) \quad (7.1)$$

за емпіричним та гіпотетичним розподілом, відповідно.

Нехай  $(a, b)$  – інтервал, що містить усі вибіркові дані:  $X \subset (a, b)$ , де  $X$  – вибірковий простір. За борелеві множини для порівняння ймовірностей (7.1) візьмемо розбиття

$$(a_0, a_1] \cup (a_1, a_2] \cup \dots \cup (a_{s-1}, a_s] \quad (a_0 = a, a_s = b),$$



і побудуємо інтервальний ряд:

|         |              |              |         |                  |
|---------|--------------|--------------|---------|------------------|
| $\xi^*$ | $(a_0, a_1)$ | $[a_1, a_2)$ | $\dots$ | $[a_{s-1}, a_s)$ |
| $m$     | $m_1$        | $m_2$        | $\dots$ | $m_s$            |

(7.2)

(число проміжків групування, наприклад, розрахуємо за правилом Стерджеса (1.11)).

За умов справедливості гіпотези  $H_0$ ,

$$P_G(a_{i-1}, a_i] = G(a_i) - G(a_{i-1}) \equiv p_i, \quad i = \overline{1, s}, \quad (7.3)$$

а

$$\bar{P}_n(a_{i-1}, a_i] = \bar{F}_n(a_i) - \bar{F}_n(a_{i-1}) = \frac{m_i}{n}, \quad i = \overline{1, s}.$$

Для міри порівняння Пірсон запропонував величину

$$q_n(x_1, \dots, x_n) = \sum_{i=1}^s \frac{n}{p_i} \left( \frac{m_i}{n} - p_i \right)^2. \quad (7.4)$$

Оскільки  $\bar{F}_n(x)$  – незміщена і спроможна оцінка для  $F_\xi(x)$ , то

$$\frac{m_i}{n} \xrightarrow{P} F_\xi(a_i) - F_\xi(a_{i-1}), \quad \text{коли } n \rightarrow \infty.$$

Якщо ж гіпотеза  $H_0$  правдива, то

$$F_\xi(a_i) - F_\xi(a_{i-1}) = G(a_i) - G(a_{i-1}) = p_i$$

і, отже,

$$\frac{m_i}{n} - p_i \xrightarrow{P} 0 \quad \text{при } n \rightarrow \infty \text{ для всіх } i = \overline{1, s}.$$

Тобто за умови справедливості гіпотези  $H_0$  значення  $q_n(x_1, \dots, x_n)$  на будь-яких вибіркових даних малі і тоді критична область цього критерію, вочевидь, виглядає так:

$$S = \{(x_1, \dots, x_n) \in X \subseteq \mathbb{R}^n : q_n(x_1, \dots, x_n) \geq l\},$$

де  $l > 0$  – деяке фіксоване значення, що визначає критерій.

Питання пошуку межі  $l$  при заданому рівні значущості  $\alpha$  вирішується мовою квантилів, як тільки розподіл випадкової величини  $q_n$  відомий. У цьому випадку зручно використати наступне поняття верхньої  $\alpha$ -границі.

**Означення.** Значення  $u_\alpha$  називають *верхньою  $\alpha$ -границею* розподілу  $P_\eta$ , якщо  $P\{\eta \geq u_\alpha\} = \alpha$ .

**Зауваження 8.1.** Квантилі  $k_{1-\alpha}$  та верхні  $\alpha$ -границі  $u_\alpha$  одного й того ж розподілу, вочевидь, пов'язані співвідношенням:  $k_{1-\alpha} = u_\alpha$ .

Продовжимо побудову критерію. За означенням,

$$\alpha = P\{(x_1, \dots, x_n) \in S | H_0\} = P\{q_n \text{ gel}\},$$

а отже шукане значення  $l$  – верхня  $\alpha$ -границя розподілу  $P_{q_n}$ .

Стосовно розподілу випадкової величини  $q_n$  відома теорема.

**Теорема 7.1 (Пірсон).** Нехай  $(x_1, \dots, x_n)$  – вибірка з невідомого розподілу  $P_\xi$  для перевірки непараметричної гіпотези  $H_0 : F_\xi(x) = G(x)$  і (7.2) – відповідний інтервальний ряд з  $s$  проміжками групування;  $m_i$  – емпіричні частоти, а ймовірності  $p_i$  розраховані за формулою (7.3),  $i = \overline{1, s}$ .

Якщо справедлива гіпотеза  $H_0$  і  $q_n(x_1, \dots, x_n)$  – випадкова величина, визначена в (7.4), то

$$q_n(x_1, \dots, x_n) \xrightarrow{d} \chi^2(s-1), \quad \text{коли } n \rightarrow \infty.$$

Розглянемо доведення теореми Пірсона при  $s = 2$ . Оскільки для двох проміжків групування  $m_2 = n - m_1$  і  $p_2 = 1 - p_1$ , то

$$\begin{aligned} q_n(x_1, \dots, x_n) &= \frac{n}{p_1} \left( \frac{m_1}{n} - p_1 \right)^2 + \frac{n}{p_2} \left( \frac{m_2}{n} - p_2 \right)^2 = \\ &= \frac{n}{p_1} \left( \frac{m_1}{n} - p_1 \right)^2 + \frac{(m_1 - np_1)^2}{n(1-p_1)} = \frac{(m_1 - np_1)^2}{np_1(1-p_1)} = \left( \frac{m_1 - np_1}{\sqrt{np_1(1-p_1)}} \right)^2. \end{aligned}$$

Крім того, частоту  $m_1$  можна подати у вигляді

$$m_1 = \eta_1 + \dots + \eta_n, \quad \text{де } \eta_i \in B(1, p_1), \quad i = \overline{1, n}$$

і тоді за ЦГТ

$$\frac{m_1 - np_1}{\sqrt{np_1(1-p_1)}} \xrightarrow{d} \psi \in N(0, 1).$$

Звідси  $q_n(x_1, \dots, x_n) \xrightarrow{d} \psi^2 \in \chi^2(1)$ .

**Теорема 7.2 (критерій  $\chi^2$  непараметричний).** *За умов теореми 7.1, якщо основну гіпотезу  $H_0$  відхиляти при*

$$q_n = \sum_{i=1}^s \frac{(m_i - np_i)^2}{np_i} > \chi_\alpha^2(s-1),$$

де  $\chi_\alpha^2(s-1)$  – верхня  $\alpha$ -границя розподілу  $\chi^2(s-1)$  і не відхиляти в супротивному разі, то ймовірність можливої помилки буде не більшою ніж  $\alpha$ .

Вперше значення верхніх  $\alpha$ -границь розподілів  $\chi^2(n)$  при різних значеннях  $n$  були табульовані Ельдертоном<sup>1</sup> в 1902 р. (див. табл.V).

## 7.2. Критерій $\chi^2$ для складної параметричної гіпотези

Розглянемо тепер випадок складної параметричної гіпотези. А саме, припустимо, що вид гіпотетичної функції  $G$  повністю визначений, але її параметри (усі або деякі) залишаються невідомими. Тобто  $G(x) = G(x, \theta)$ , де  $\theta \in \Theta \subset \mathbb{R}^d$  – вектор невідомих параметрів. У такому разі гіпотеза  $H_0 : F_\xi(x) = G(x, \theta)$  для деякого  $\theta \in \Theta \subset \mathbb{R}^d$  – складена параметрична.

Ідея перевірки даної гіпотези  $H_0$  така ж, як і в п. 7.1. Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi$ , оберемо інтервал  $(a, b)$ , що містить усі вибіркові дані і побудуємо інтервальний ряд (7.2). Далі визначимо випадкову величину  $q_n(x_1, \dots, x_n)$  за формулою (7.4).

Складність в тому, що тепер випадкова величина  $q_n$  також залежить від  $\theta$ , оскільки залежні від  $\theta$  імовірності  $p_i(\theta)$ :

$$q_n(x_1, x_2, \dots, x_n; \theta) = \sum_{i=1}^s \frac{n}{p_i(\theta)} \left( \frac{m_i}{n} - p_i(\theta) \right)^2.$$

Як параметри, що невідомі, Фішер розглянув їх оцінки методу максимальної правдоподібності  $\hat{\theta}_1, \hat{\theta}_1, \dots, \hat{\theta}_d$ , знайдені за вибіркою, звів задачу

---

<sup>1</sup>Уільям Пейлін Ельдертон (1877 – 1962) – британський актуарій.

до перевірки гіпотези  $H'_0 : F_\xi(x) = G(x, \hat{\theta})$  і показав, що, за умови правдивої гіпотези  $H'_0$ , випадкова величина

$$q_n(x_1, x_2, \dots, x_n; \hat{\theta}) = \sum_{i=1}^s \frac{n}{p_i(\hat{\theta})} \left( \frac{m_i}{n} - p_i(\hat{\theta}) \right)^2, \quad (7.5)$$

де

$$p_i(\hat{\theta}) = G(a_i, \hat{\theta}) - G(a_{i-1}, \hat{\theta}), \quad i = \overline{1, s}, \quad (7.6)$$

має асимптотичний розподіл  $\chi^2(s - d - 1)$ .

**Теорема 7.3 (Фішер).** *Нехай  $(x_1, \dots, x_n)$  – вибірка з невідомого розподілу  $P_\xi(\theta)$ ,  $\theta \in R^d$  для перевірки параметричної гіпотези  $H_0 : F_\xi(x) = G(x, \hat{\theta})$ , де  $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_d)$  і кожна компонента  $\hat{\theta}_i$  є ОМП відповідного параметра  $\theta_i$ ,  $i = \overline{1, d}$ , (7.2) – інтервальний ряд за вибіркою з  $s$  проміжками групування,  $m_i$  – емпіричні частоти, а ймовірності  $p_i(\hat{\theta})$  розраховані за формулою (7.6),  $i = \overline{1, s}$ . Якщо справедлива гіпотези  $H_0$  і  $q_n(x_1, \dots, x_n; \hat{\theta})$  – випадкова величина, визначена в (7.5), то*

$$q_n(x_1, \dots, x_n; \hat{\theta}) \xrightarrow{d} \chi^2(s - d - 1), \quad \text{коли } n \rightarrow \infty.$$

Тепер межова точка  $l$  критичної області  $S$ , що визначає параметричний критерій  $\chi^2$ , є верхньою  $\alpha$ -границею розподілу  $\chi^2$  з  $(s - d - 1)$  ступенями свободи.

**Теорема 7.4 (критерій  $\chi^2$  параметричний).** *За умов теореми 7.3, якщо основну гіпотезу  $H_0$  відхиляти при*

$$q_n = \sum_{i=1}^s \frac{(m_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})} > \chi_\alpha^2(s - d - 1),$$

де  $\chi_\alpha^2(s - d - 1)$  – верхня  $\alpha$ -границя розподілу  $\chi^2(s - d - 1)$ , і не відхиляти в супротивному разі, то ймовірність можливої помилки першого роду буде не більшою ніж  $\alpha$ .

#### Зауваження до теорем 7.2 та 7.4.

- 1) Інтервали розбиття не обов'язково брати однаковими за довжиною.
- 2) Критерій  $\chi^2$  не відрізняє основну гіпотезу  $H_0$  від альтернативної  $H_1$ , якщо ймовірності попадання в проміжки групування для них однакові. Крім того, асимптотичний характер критерію призводить до того,

що для його застосування важливо мати реалізацію великого обсягу  $n$  і велику кількість проміжків групування  $s = s(n)$ . Кількість проміжків групування  $s = s(n)$  обирають так, щоб  $np_i \geq 10$ ,  $i = \overline{1, s}$ .

Манн<sup>2</sup> і Вальд<sup>3</sup> довели, що оптимальне число  $s = s(n)$  інтервалів задовольняє співвідношенню

$$s(n) \sim 4\sqrt[5]{2} \left( \frac{n}{d_{1-\alpha}} \right)^{2/5},$$

де  $d_{1-\alpha}$  – квантиль рівня  $(1-\alpha)$  стандартного розподілу (див., наприклад, [7], стор. 318).

3) У разі, коли для окремих інтервалів розбиття не виконується нерівність  $np_i \geq 5$ , їх об'єднують з сусідніми. Або розглядають нове розбиття так, щоб нерівність  $np_i \geq 5$  виконувалась для всіх інтервалів  $i = \overline{1, s}$ .

### 7.3. Критерій $\chi^2$ як критерій незалежності

Критерій  $\chi^2$  можна застосовувати також для перевірки гіпотези про незалежність випадкових величин.

Припустимо, що на елементах вибірки обсягу  $n$  з деякої генеральної сукупності спостерігались дві числові ознаки  $\xi$  та  $\eta$ . Наприклад, для кожного прогонича, що підлягав обстеженню на міцність  $\xi$  (н/мм)<sup>2</sup> визначалась ще і його вага  $\eta$  (г). Або для випадково обраної людини фіксувались наявність щеплення  $\xi$  та відповідної хвороби  $\eta$ .

У загальному випадку маємо  $n$  незалежних спостережень випадкового вектора  $(\xi, \eta)$  та відповідні вибіркові дані:

$$(x_i, y_i), \quad i = \overline{1, n}. \quad (7.7)$$

Як і при побудові критерію  $\chi^2$  для перевірки гіпотези про розподіл, обемо інтервали

$$(a, b) \supset \{x_1, \dots, x_n\} \quad \text{і} \quad (c, d) \supset \{y_1, \dots, y_n\}$$

<sup>2</sup>Генрі Бертольд Манн (27.10.1905, Вена – 1.02.2000, Тусон, США) – математик та статистик. Опублікував першу математичну книгу з планування експерименту.

<sup>3</sup>Абрахам Вальд (31.10.1902, Клуж-Напока, Австро-Венгрія – 13.12.1950, Траванкор, Індія) – математик, статистик, економіст. Використовував статистичні методи при аналізі втрат бойових літаків під час Другої світової війни.

та кожен з них розіб'ємо на  $s$  та  $k$  частин, відповідно точками

$$a = a_0 < a_1 < a_2 < \dots < a_s = b, \quad c = c_0 < c_1 < c_2 < \dots < c_k = d$$

(кількість проміжків  $s$  та  $k$  не обов'язково збігаються). Позначимо

$$B_{i,j} = [a_{i-1}, a_i) \times [c_{j-1}, c_j), \quad i = \overline{1, s}, \quad j = \overline{1, k} \quad (7.8)$$

і нехай  $m_{i,j}$  – кількість варіант серед вибірових даних (7.7), що належать проміжку  $B_{i,j}$ ,  $i = \overline{1, s}$ ,  $j = \overline{1, k}$ . Складемо таблицю емпіричних частот випадкового вектора:

| $i, j$   | 1        | 2        | ... | k        | сума     |                                                                                           |
|----------|----------|----------|-----|----------|----------|-------------------------------------------------------------------------------------------|
| 1        | $m_{11}$ | $m_{12}$ | ... | $m_{1k}$ | $m_{1.}$ | $m_{i.} = \sum_{j=1}^k m_{ij}, i = \overline{1, s}$                                       |
| 2        | $m_{21}$ | $m_{22}$ | ... | $m_{2k}$ | $m_{2.}$ |                                                                                           |
| $\vdots$ |          |          |     |          |          | $m_{.j} = \sum_{i=1}^s m_{ij}, j = \overline{1, k}$                                       |
| s        | $m_{s1}$ | $m_{s2}$ | ... | $m_{sk}$ | $m_{s.}$ |                                                                                           |
| сума     | $m_{.1}$ | $m_{.2}$ | ... | $m_{.k}$ | $n$      | $\sum_{i=1}^s m_{i.} = \sum_{j=1}^k m_{.j} =$<br>$= \sum_{i=1}^s \sum_{j=1}^k m_{ij} = n$ |

Табл. 7.1

Розподіл вектора  $(\xi, \eta)$  невідомий і щодо нього висувається гіпотеза про незалежність його компонент:

$$H_0 : \xi \text{ та } \eta \text{ незалежні.} \quad (7.9)$$

Шляхом аналізу вибірових даних таблиці 7.1 дійдемо висновку, чи узгоджуються результати спостережень з основною гіпотезою.

За умови справедливості гіпотези  $H_0$ , для будь-яких  $i = \overline{1, s}$ ,  $j = \overline{1, k}$

$$P\{(\xi, \eta) \in B_{i,j}\} = P\{\xi \in [a_{i-1}, a_i)\}P\{\eta \in [c_{j-1}, c_j)\}. \quad (7.10)$$

Тобто треба перевірити, чи протирічають дані з таблиці 7.1 рівності (7.10).

Значення

$$p_{ij} = P\{(\xi, \eta) \in B_{i,j}\}, \quad p_{i\cdot} = P\{\xi \in [a_{i-1}, a_i]\}, \quad p_{\cdot j} = P\{\eta \in [c_{j-1}, c_j]\}$$

невідомі,  $i = \overline{1, s}$ ,  $j = \overline{1, k}$ . Скористаємось для них оцінками методу максимальної правдоподібності:

$$\hat{p}_{ij} = \frac{m_{ij}}{n}, \quad \hat{p}_{i\cdot} = \frac{m_{i\cdot}}{n}, \quad \hat{p}_{\cdot j} = \frac{m_{\cdot j}}{n}, \quad i = \overline{1, s}, \quad j = \overline{1, k}.$$

За законом великих чисел,

$$\frac{m_{ij}}{n} \xrightarrow{P} p_{ij}, \quad \frac{m_{i\cdot}}{n} \xrightarrow{P} p_{i\cdot}, \quad \text{а} \quad \frac{m_{\cdot j}}{n} \xrightarrow{P} p_{\cdot j}.$$

Таким чином, достатньо порівняти значення

$$m_{ij} \quad \text{та} \quad \frac{m_{i\cdot} m_{\cdot j}}{n}.$$

Велика різниця між ними буде підґрунтям для відхилення гіпотези  $H_0$  і як міру розбіжності розглянемо випадкову величину

$$q_n = \sum_{i=1}^s \sum_{j=1}^k \frac{\left(m_{ij} - \frac{m_{i\cdot} m_{\cdot j}}{n}\right)^2}{m_{i\cdot} m_{\cdot j} / n}.$$

Кількість параметрів, що оцінювались за вибіркою, становить  $s - 1 + k - 1 = s + k - 2$ , а кількість підмножин  $B_{i,j}$ ,  $i = \overline{1, s}$ ,  $j = \overline{1, k}$  дорівнює  $sk$ . Як і в п. 7.1, випадкова величина  $q_n$  має розподіл  $\chi^2(sk - 1 - (s + k - 2))$ . Після очевидних перетворень, ступені свободи рахують за формулою  $(s - 1)(k - 1)$ . Тобто

$$q_n(x_1, \dots, x_n) \xrightarrow{d} \chi^2((s - 1)(k - 1)), \quad n \rightarrow \infty.$$

Тоді якщо основну гіпотезу (7.9) відхиляти при

$$q_n = \sum_{i=1}^s \sum_{j=1}^k \frac{(m_{ij} - m_{i\cdot} m_{\cdot j} / n)^2}{m_{i\cdot} m_{\cdot j} / n} > \chi_\alpha^2((s - 1)(k - 1)),$$

де  $\chi_\alpha^2((s - 1)(k - 1))$  – верхня  $\alpha$ -границя розподілу  $\chi^2((s - 1)(k - 1))$ , і не відхиляти в супротивному разі, то ймовірність можливої помилки буде не більша ніж  $\alpha$ .

**Зауваження 7.1.** Під час побудови критерію вибіркового простір  $\mathbb{R}^2$  необхідно розбити на проміжки  $B_{i,j}$ ,  $i = \overline{1, s}$ ,  $j = \overline{1, k}$  (див. (7.8)) так, щоб виконувались нерівності  $m_i \cdot m_j / n \geq 10$  для всіх  $i = \overline{1, s}$ ,  $j = \overline{1, k}$ .

**Контрольні запитання:**

1. Сформулюйте основну ідею побудови критерію  $\chi^2$  для перевірки непараметричної гіпотези про розподіл.
2. У чому відмінність у разі складної параметричної гіпотези? Як розв'язав цю задачу Фішер?
3. Чому при застосуванні критерію Пірсона необхідно мати вибірку великого обсягу  $n$  і число  $s = s(n)$  – кількість проміжків групування повинно бути не малим? Які відомі рекомендації з цього приводу?
4. Як працюють ідеї Пірсона при перевірці гіпотези про незалежність випадкових величин.

**Вправи та задачі: 7.1.** "Випадковість" не означає в "будь-який спосіб". Випадковість підпорядковується своїм законам. Випишіть, будь ласка, 20 випадкових, на ваш погляд, чисел з відрізка  $[0, 1]$  (для визначеності, розглядаємо 3 цифри після коми). Якщо отримані числа дійсно випадкові, то вони утворюють вибірку з рівномірного розподілу  $U[0, 1]$ . За допомогою критерію Пірсона перевірте, чи це так<sup>4</sup>.

**7.2.** Побудуйте критерій  $\chi^2$  для перевірки гіпотези про однорідність вибірок (див. також [5, с. 382]).

**7.3.** Побудуйте критерій  $\chi^2$  для перевірки гіпотези про незалежність подій.

---

<sup>4</sup>Вказівка: під час запису випадкових чисел не передивляйтесь і не використовуйте записану вже частину.



## §8. Види критеріїв

### 8.1. Рівень значущості та потужність критерію

Нехай  $\delta$  – деякий критерій перевірки основної гіпотези  $H_0$  проти альтернативи  $H_1$ ;  $S$  – його критична область. Якість критерію характеризують імовірності помилок першого та другого роду:

$$\alpha = P\{x \in S | H_0\} \quad \text{та} \quad \beta = P\{x \in \bar{S} | H_1\}, \quad (8.1)$$

відповідно.

Якщо  $\alpha$  – це рівень значущості критерію, то величина  $\beta$  характеризує *потужність* критерію, що розраховується як  $1 - \beta$ .

Рівень значущості критерію задає дослідник і тим самим контролює критичну область, а потужність  $1 - \beta$  – характеристика критерію, який при цьому отримуємо.

Зауважимо, що ймовірності помилок першого та другого роду визначаються умовними ймовірностями (8.1) за різних припущень. Рівень значущості  $\alpha$  розраховують при справедливій гіпотезі  $H_0$ , а  $\beta$  знаходять за умови істинної альтернативи  $H_1$ . Тому фіксованому співвідношенню  $\alpha = 1 - \beta$  ці помилки не задовольняють.

Бажано, звичайно, провести перевірку гіпотез  $\{H_0, H_1\}$  так, щоб мінімізувати ймовірності обох помилок, тобто забезпечити малість  $\alpha$  та  $\beta$  одночасно. Розглянемо два різних критерії  $\delta_1$  та  $\delta_2$  з ймовірностями помилок  $\alpha(\delta_i)$ ,  $\beta(\delta_i)$ ,  $i = \overline{1, 2}$ , відповідно.

**Означення.** Вважають, що *критерій  $\delta_1$  краще критерію  $\delta_2$* , якщо

$$\begin{cases} \alpha(\delta_1) \leq \alpha(\delta_2), \\ \beta(\delta_1) \leq \beta(\delta_2) \end{cases}$$

і хоча б одна з цих нерівностей є точною.

Але не всі критерії можна порівняти за цим означенням.

**Приклад 8.1.** Нехай критерій  $\delta_1$  завжди вибирає  $H_0$ . Для нього, вочевидь,  $\alpha(\delta_1) = 0$ ,  $\beta(\delta_1) = 1$ . А критерій  $\delta_2$  – навпаки. Тоді  $\alpha(\delta_2) = 1$ ,  $\beta(\delta_2) = 0$ .

**Приклад 8.2.** Припустимо, що  $x_1$  – вибірка обсягу  $n = 1$  з розподілу  $P_\xi = N(\theta, 1)$ , для якого треба перевірити основну гіпотезу

$$H_0 : \theta = 0 \text{ проти альтернативи } H_1 : \theta = 1.$$

Вочевидь, критична область даного критерію має вигляд:

$$S = \{(x_1, \dots, x_n) \in X : \bar{x} > c\} = S(c).$$

Вибір сталої  $c$  визначає якість критерію. Як бачимо, зі зростанням числа  $c$  імовірність помилки першого роду зменшується, але ймовірність помилки другого роду збільшується. Якщо  $\delta_1$  та  $\delta_2$  – критерії з критичними областями  $S_1(c_1)$  та  $S_2(c_2)$  відповідно і  $c_1 < c_2$ , то  $\alpha(\delta_1) > \alpha(\delta_2)$ , а  $\beta(\delta_1) < \beta(\delta_2)$ .

Прості приклади 1 та 2 ілюструють загальну тенденцію: при спробі зменшити одну з імовірностей помилок в (8.1) друга, як правило, збільшується.

## 8.2. Мінімаксні, баєсовські та найбільш потужні критерії

Розглянемо два різних критерії  $\delta_1$  та  $\delta_2$  для перевірки простої основної гіпотези  $H_0$  за альтернативи  $H_1$  з імовірностями помилок:  $\alpha(\delta_i)$ ,  $\beta(\delta_i)$ ,  $i = 1, 2$  відповідно.

**Означення.** Кажуть, що критерій  $\delta_1$  *не гірший* за критерій  $\delta_2$  в мінімаксному значенні, якщо

$$\max\{\alpha(\delta_1), \beta(\delta_1)\} \leq \max\{\alpha(\delta_2), \beta(\delta_2)\}.$$

**Означення.** Критерій називають *мінімаксним*, якщо він не гірший від усіх інших критеріїв в мінімаксному сенсі. Позначають  $\delta^{\min \max}$ .

Інакше кажучи, мінімаксний критерій має найменшу з можливих "найбільшу помилку"

$$\max\{\alpha(\delta), \beta(\delta)\}$$

серед усіх інших критеріїв.

У прикладі 8.2 мінімаксним буде критерій

$$\delta^{\min \max} = \{(x_1, \dots, x_n) \in X : \bar{x} \geq \frac{1}{2}\}.$$

Розглянемо ще один підхід до порівняння критеріїв – баєсовський. Припускається, що розподіл  $P_\xi$ , з якого вилучена вибірка, обрали випадково. Тоді і гіпотези  $H_0$  та  $H_1$  – випадкові події.

Імовірності  $P(H_0) = p_0$  та  $P(H_1) = p_1$  розподіляємо апіорі. Не обов'язково  $p_0 + p_1 = 1$ , але їх можна нормувати, і тоді

$$P(H_0) = \frac{p_0}{p_0 + p_1} = q_0, \quad P(H_1) = \frac{p_1}{p_0 + p_1} = q_1. \quad (8.2)$$

У такому разі кажуть про *лінійну функцію втрат*. Втрати при помилковому рішенні дорівнюють  $q_0$ , якщо справедлива основна гіпотеза  $H_0$ , і  $q_1 = 1 - q_0$ , якщо істинна  $H_1$ .

За умови заданих апіорних імовірностей (8.2) імовірність помилки критерію розраховують за формулою повної ймовірності:

$$\begin{aligned} P\{\delta \text{ помиляється}\} &= \sum_{i=0}^1 P\{\delta_1 \text{ помиляється} | H_i\} P\{H_i\} = \\ &= \alpha(\delta_1)q_0 + \beta(\delta_2)q_1. \end{aligned} \quad (8.3)$$

**Означення.** Критерій, який мінімізує ймовірність помилки (8.3), називають *баєсовським*, що відповідає апіорному розподілу ймовірностей гіпотез (8.2). Позначають:  $\delta_{q_0, q_1}^{baies}$ .

Серед усіх можливих критеріїв баєсовський має найменшу "середньовагову" помилку.

При порівнянні критеріїв можливий ще такий підхід: одну з помилок фіксують і мінімізують іншу. Помилки першого та другого роду нерівноправні і за основну гіпотезу ми обираємо ту, для якої бажаною є досить мала ймовірність помилки першого роду. Тому, зазвичай, задають досить мале значення рівня значущості  $\alpha = \varepsilon$  і розглядають клас критеріїв  $K_\varepsilon$ , що задовольняють умові

$$K_\varepsilon = \{\delta \mid \alpha(\delta) \leq \varepsilon\}.$$

**Означення.** Критерій  $\delta_0 \in K_\varepsilon$  називають *найбільш потужним* критерієм (НПК) рівня  $\varepsilon$ , якщо  $\beta(\delta_0) \leq \beta(\delta)$  для будь-якого іншого критерію  $\delta \in K_\varepsilon$ . Позначають:  $\delta^{\text{НПК}}(\varepsilon) = \delta_0$ .

Виявляється, що всі оптимальні критерії (мінімаксні, байєсовські, найбільш потужні) для перевірки гіпотези про розподіл можна побудувати на основі простого вибору сталих в деякому універсальному критерії – критерії відношення правдоподібності.

### 8.3. Критерій відношення правдоподібності

Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi$ , відносно якого висувається основна гіпотеза

$$H_0 : P_\xi = P_0 \text{ за альтернативи } H_1 : P_\xi = P_1,$$

де  $P_0, P_1$  – деякі розподіли.

Домовимось, що  $P_0$  і  $P_1$  або обидва дискретні, або обидва абсолютно неперервні (у випадку, коли один розподіл дискретний, а інший абсолютно неперервний, завжди існує критерій з нульовими ймовірностями помилок, дивись задачу 8.2).

Побудуємо функції правдоподібності розподілів  $P_0$  та  $P_1$ :

$$L_0(x) = \prod_{i=1}^n p_0(x_i), \quad L_1(x) = \prod_{i=1}^n p_1(x_i),$$

де

$$p_j(t) = \begin{cases} p_j(t) - \text{щільність абсолютно неперервного розподілу } P_j; \\ P_j\{\xi = t\} - \text{імовірність, якщо } P_j - \text{дискретний, } j = 0, 1. \end{cases}$$

Ідею побудови критерію відношення правдоподібності проілюструємо на прикладі 8.2, що розглядався вище. А саме, за умови, що  $x_1$  – вибірка обсягу  $n = 1$  з нормального розподілу  $P_\xi = N(\theta, 1)$  треба перевірити гіпотезу  $H_0 : \theta = 0$  за альтернативи  $H_1 : \theta = 1$ .

Логічно відхилити основну гіпотезу  $H_0$ , якщо  $x_1 \geq 1/2$ . Зазначимо, що при  $x_1 = 1/2$  виконується точна рівність для щільностей:  $p_1(1/2) =$

$p_0(1/2)$ , а при  $x > 1/2$  – нерівність  $p_1(x) > p_0(x)$ . Тобто там, де  $p_1(x) \geq p_0(x)$ , ми відхиляємо  $H_0$ ; якщо ж навпаки – відхиляємо  $H_1$ . Критична область такого критерію має вигляд:

$$S = \left\{ x \in X \subseteq \mathbb{R} : \frac{p_1(x)}{p_0(x)} \geq 1 \right\}. \quad (8.4)$$

Порівняння відношення  $p_1(x)/p_0(x)$  з одиницею дає рівно один критерій з фіксованим рівнем значущості та потужністю. Якщо ж маємо на меті будувати критерії з наперед заданим рівнем значущості  $\alpha$ , то в (8.4) замість одиниці розглядають вільний параметр  $c$ :

$$S = \left\{ x \in X \subset \mathbb{R} : \frac{p_1(x)}{p_0(x)} \geq c \right\}.$$

У загальному випадку зіставляють зі сталою  $c$  частку відповідних функцій правдоподібності.

**Означення.** Частку  $L_1(x)/L_0(x)$  називають *відношенням правдоподібності*.

Розглядають тільки такі значення  $x \in X$ , для яких хоча б одна з функцій правдоподібності не дорівнює нулю.

Критерій з критичною областю

$$S = \left\{ x \in X : \frac{L_1(x)}{L_0(x)} \geq c \right\} \quad (8.5)$$

називають *критерієм відношення правдоподібності* і позначають  $\delta_c$ .

Як зазначалось, усі оптимальні критерії перевірки гіпотез про розподіл, означені в п. 8.2, є критеріями відношення правдоподібності. Про те наступна теорема.

**Теорема 8.1.** Нехай  $(x_1, \dots, x_n)$  – вибірка обсягу  $n$  з невідомого розподілу  $P_\xi$ , відносно якого формулюється гіпотеза  $H_0 : P_\xi = P_0$  за альтернативи  $H_1 : P_\xi = P_1$ . Припустимо також, що виконуються умови:

(i) розподіли  $P_0$  та  $P_1$  або обидва дискретні, або обидва абсолютно неперервні;

(ii)  $P \left\{ \frac{L_1(x)}{L_0(x)} = c \mid H_1 \right\} = 0, \quad \forall c > 0.$

Тоді критерій відношення правдоподібності  $\delta_c$  є

- 1) мінімаксним критерієм при такій сталій  $c$ , що  $\alpha(\delta_c) = \beta(\delta_c)$ ;
- 2) байєсовським з апіорними ймовірностями  $q_0$  та  $q_1 = 1 - q_0$ , якщо  $c = q_0/q_1$ :  $\delta_{q_0, q_1}^{baies} = \delta_{q_0/q_1}$ ;
- 3) НПК рівня  $\varepsilon$ , де  $0 < \varepsilon \leq P\{p_2(x) > 0 | H_1\}$ , якщо сталу  $c$  обрано так, що  $\alpha(\delta_c) = \varepsilon$ .

Третє твердження теореми називають *лемою Неймана<sup>1</sup>-Пірсона*.

Розглянемо приклад застосування теореми 8.1.

**Приклад 8.3.** Нехай  $(x_1, \dots, x_n)$  – вибірка з нормального розподілу  $P_\xi = N(\theta, \sigma^2)$ , математичне сподівання  $\theta$  якого невідоме, а дисперсія  $\sigma^2$  відома. Побудуємо мінімаксний, байєсовський з апіорними ймовірностями  $q_0 = 2/3$ ,  $q_1 = 1/3$  та найбільш потужний рівня  $\varepsilon$  критерії для перевірки гіпотези  $H_0 : \theta = a_0$  проти альтернативи  $H_1 : \theta = a_1 > a_0$ .

Розглянемо відношення правдоподібності:

$$\begin{aligned} \frac{L_1(x)}{L_0(x)} &= \exp \left\{ -\frac{1}{2\sigma^2} \left( \sum_{i=1}^n (x_i - a_1)^2 - (x_i - a_0)^2 \right) \right\} = \\ &= \exp \left\{ \frac{a_1 - a_0}{\sigma^2} \sum_{i=1}^n x_i - \frac{a_1^2 - a_0^2}{2\sigma^2} n \right\} \geq c, \end{aligned}$$

де  $c$  – деяка невідома стала, що визначає критичну область шуканого критерію. Звідси маємо:

$$\frac{n(a_1 - a_0)}{\sigma^2} \bar{x} \geq \frac{a_1^2 - a_0^2}{2\sigma^2} n + \ln c.$$

Тобто нерівність  $L_1(x)/L_0(x) \geq c$  рівносильна наступній:

$$\bar{x} \geq \frac{a_0 + a_1}{2} + \frac{\sigma^2 \ln c}{n(a_1 - a_0)}. \quad (8.6)$$

Розподіл випадкової величини  $\bar{x}$  абсолютно неперервний, тому ймовірність точної рівності в (8.6) дорівнює нулеві. Таким чином, виконуються

---

<sup>1</sup>Джон фон Нейман (28.12.1903, Будапешт – 08.02.1957, Вашингтон) – математик, інформатик, фізик, економіст. Один із розробників принципу дії комп'ютера, теорії ігор та її застосування до розв'язання економічних задач.

обидві умови теореми 8.1 і тепер за рахунок вибору сталої  $c$  отримаємо потрібні критерії.

Спочатку побудуємо баєсовський критерій  $\delta_{2/3,1/3}^{baies}$ . За теоремою 8.1, це критерій відношення правдоподібності

$$S = \left\{ x \in X : \bar{x} \geq \frac{a_0 + a_1}{2} + \frac{\sigma^2 \ln c}{n(a_1 - a_0)} \right\}.$$

зі сталою  $c = q_0/q_1 = 2$ . Тобто критерій  $\delta_{2/3,1/3}^{baies}$  характеризується критичною областю

$$\delta_{2/3,1/3}^{baies} : S = \left\{ x \in X : \bar{x} \geq \frac{a_0 + a_1}{2} + \frac{\sigma^2 \ln 2}{n(a_1 - a_0)} \right\}.$$

Якщо ж гіпотези  $H_0$  та  $H_1$  рівноможливі, тобто  $q_0 = q_1 = 1/2$ , то

$$\delta_{1/2,1/2}^{baies} : S = \left\{ x \in X : \bar{x} \geq \frac{a_0 + a_1}{2} \right\},$$

як у наших міркуваннях щодо прикладу 8.2.

Для того щоб побудувати мінімаксий та найбільш потужний критерії, запишемо нерівність (8.6) у вигляді

$$\bar{x} \geq \frac{a_0 + a_1}{2} + \frac{\sigma^2 \ln c}{n(a_1 - a_0)} = C.$$

де  $c$  – невідоме, а, отже, невідома і стала  $C$ . Зазначимо, що  $C$  залежить від  $n$ . Тепер

$$\begin{aligned} \alpha(\delta_C) &= P\{\bar{x} \geq C | H_0\} = P\left\{ \frac{(\bar{x} - a_0)\sqrt{n}}{\sigma} \geq \frac{C - a_0}{\sigma} \sqrt{n} \right\} = \\ &= 1 - F\left( \frac{C - a_0}{\sigma} \sqrt{n} \right), \end{aligned}$$

а

$$\begin{aligned} \beta(\delta_C) &= P\{\bar{x} < C | H_1\} = P\left\{ \frac{(\bar{x} - a_1)\sqrt{n}}{\sigma} < \frac{C - a_1}{\sigma} \sqrt{n} \right\} = \\ &= F\left( \frac{C - a_1}{\sigma} \sqrt{n} \right), \end{aligned}$$

де  $F$  – функція розподілу стандартної випадкової величини.

За умови фіксованого рівня значущості  $\alpha = \varepsilon$  (будуємо НПК), маємо:

$$\alpha(\delta_C) = 1 - F\left(\frac{C - a_0}{\sigma}\sqrt{n}\right) = \varepsilon,$$

звідки

$$\frac{(C - a_0)}{\sigma}\sqrt{n} = d_{1-\varepsilon},$$

де  $d_{1-\varepsilon}$  – квантиль стандартного розподілу рівня  $1 - \varepsilon$  і

$$C = a_0 + \frac{d_{1-\varepsilon}}{\sqrt{n}}\sigma.$$

Тепер критерій  $\delta^{\text{НПК}}(\varepsilon)$  повністю визначений:

$$\delta^{\text{НПК}}(\varepsilon) : S = \left\{ x \in X : \bar{x} \geq a_0 + \frac{d_{1-\varepsilon}}{\sqrt{n}}\sigma \right\}$$

і його потужність наступна:

$$\beta(\delta^{\text{НПК}}(\varepsilon)) = F\left(d_{1-\varepsilon} - \frac{a_1 - a_0}{\sigma}\sqrt{n}\right). \quad (8.7)$$

Легко бачити, що для критерію  $\delta^{\text{НПК}}(\varepsilon)$ :

- (i) за фіксованого  $0 < \varepsilon < 1$  зі збільшенням обсягу вибірки  $n$  імовірність помилки другого роду зменшується;
- (ii) потужність критерію зменшується і тоді, коли при фіксованому  $n$  збільшувати рівень  $\varepsilon$ , оскільки при цьому  $1 - \varepsilon \downarrow$  і  $d_{1-\varepsilon} \downarrow$  також;
- (iii) при фіксованому  $0 < \varepsilon < 1$  імовірність помилки другого роду, меншої за  $\beta_0$ , забезпечує обсяг вибірки

$$n \geq \frac{(d_{1-\varepsilon} - d_{\beta_0})^2 \sigma^2}{(a_1 - a_0)^2}.$$

Це пояснюється ланцюжком нерівностей:

$$\beta(\delta^{\text{НПК}}(\varepsilon)) = F\left(d_{1-\varepsilon} - \frac{a_1 - a_0}{\sigma}\sqrt{n}\right) \leq \beta_0,$$

$$d_{1-\varepsilon} - \frac{a_1 - a_0}{\sigma}\sqrt{n} \leq d_{\beta_0}, \text{ оскільки } F \nearrow;$$

$$\sqrt{n} \geq \frac{(d_{1-\varepsilon} - d_{\beta_0})\sigma}{a_1 - a_0}.$$



Тепер побудуємо  $\delta^{\min \max}$ . Критерій  $\delta_c$  буде мінімаксним, за теоремою, якщо  $\alpha(\delta_c) = \beta(\delta_c)$ . Тобто

$$1 - F\left(\frac{c - a_0}{\sigma}\sqrt{n}\right) = F\left(\frac{c - a_1}{\sigma}\sqrt{n}\right) = 1 - F\left(\frac{a_1 - c}{\sigma}\sqrt{n}\right).$$

Звідси  $c - a_0 = a_1 - c$  і  $c = (a_0 + a_1)/2$ . Остаточно, критична область мінімаксного критерію має вигляд:

$$\delta^{\min \max} : S = \left\{ x \in X : \bar{x} \geq \frac{a_0 + a_1}{2} \right\}.$$

#### 8.4. Асимптотичні характеристики критеріїв

Критична область критерію, як правило, залежить від обсягу вибірки  $n$ . У зв'язку з чим виникає питання про асимптотичну поведінку ймовірностей помилок.

**Означення.** Критерій  $\delta_n = \delta(x_1, \dots, x_n)$  називають *критерієм асимптотичного рівня значущості*  $\varepsilon$ , якщо  $\alpha(\delta_n) \rightarrow \varepsilon$  при  $n \rightarrow \infty$ .

Критерій  $\delta_n$  називають *спроможним*, якщо  $\beta(\delta_n) \rightarrow 0$ , коли  $n \rightarrow \infty$ .

Наприклад, критерій  $\delta^{\text{НПК}}(\varepsilon)$  з прикладу 8.3 є спроможним. Дійсно, знайдений там критерій  $\delta^{\text{НПК}}(\varepsilon)$  визначається критичною областю

$$S = \left\{ x \in X : \bar{x} \geq a_0 + \frac{d_{1-\varepsilon}}{\sqrt{n}}\sigma \right\}$$

і має заданий рівень значущості  $\varepsilon$  (як НПК рівня  $\varepsilon$ ). Імовірність помилки другого роду знайдена в (8.7) і

$$\beta(\delta^{\text{НПК}}(\varepsilon)) = F\left(d_{1-\varepsilon} - \frac{a_1 - a_0}{\sigma}\sqrt{n}\right) \rightarrow 0 \quad \text{при } n \rightarrow \infty$$

(врахувували властивості функції розподілу).

Прикладом спроможного критерію асимптотичного рівня значущості є критерій Пірсона. Більш детально про те в наступному пункті.

#### 8.5. Характеристики критеріїв згоди

Критерії згоди, як і критерії правдоподібності, мають один і той самий принцип побудови, який можна описати за такою схемою.

За означенням, критерій згоди перевіряє, чи узгоджується реалізація вибірки  $(x_1, \dots, x_n)$  з основною гіпотезою  $H_0$ ? Щоб отримати відповідь на це запитання, підбирають тестову статистику  $q_n(x_1, \dots, x_n)$ , розподіл якої або відомий, або відома його асимптотична поведінка. Тоді критичну область будують за визначеним рівнем значущості  $\alpha$  на основі квантилів розподілу  $P_{q_n}$ , який зазвичай абсолютно неперервний, чи його верхніх границь.

Слова "відома асимптотична поведінка" розподілу  $P_{q_n}$  означають:

1) за умови правдоподібності основної гіпотези  $H_0$ ,

$$q_n(x_1, \dots, x_n) \xrightarrow{d} \eta, \quad \text{коли } n \rightarrow \infty \quad (8.8)$$

і  $\eta$  має певний відомий абсолютно неперервний розподіл;

2) у випадку хибної основної гіпотези  $H_0$ ,

$$|q_n(x_1, \dots, x_n)| \xrightarrow{P} +\infty \quad \text{при } n \rightarrow \infty. \quad (8.9)$$

Тоді для заданого рівня значущості  $\alpha$  знаходять сталу  $c$  з рівняння

$$P\{|\eta| \geq c\} = \alpha. \quad (8.10)$$

Вочевидь,  $c = k_{1-\alpha}$  –  $(1 - \alpha)$ -квантиль розподілу  $P_\eta$  і асимптотичний критерій

$$\delta_n(x_1, \dots, x_n) = \begin{cases} H_0, & \text{якщо } |q_n(x_1, \dots, x_n)| < k_{1-\alpha}, \\ \overline{H}_0, & \text{якщо } |q_n(x_1, \dots, x_n)| \geq k_{1-\alpha}. \end{cases} \quad (8.11)$$

має таку критичну область:

$$S = \{(x_1, \dots, x_n) \in X : |q_n(x_1, \dots, x_n)| \geq k_{1-\alpha}\}.$$

Таким чином, якщо  $H_0$  правдива, то статистика  $q_n(x_1, \dots, x_n)$  за великих обсягів  $n$  має майже розподіл  $P_\eta$ . Тоді вона повинна себе поводити як випадкова величина з таким розподілом. Але стала  $c = k_{1-\alpha}$  підбиралась так, щоб імовірність в (8.10) дорівнювала рівню значущості. Тому, якщо на реалізації виконується нерівність  $|q_n(x_1, \dots, x_n)| \geq k_{1-\alpha}$ , то є підстави вважати гіпотезу  $H_0$  хибною з імовірністю помилитися, що дорівнює  $\alpha$ .

**Теорема 8.2.** Критерій згоди, визначений в (8.11), має асимптотичний рівень значущості  $\alpha$  і є спроможним.

**Доведення.** Оскільки за умови справедливої гіпотези  $H_0$  має збіжність за розподілом (8.8), то

$$\alpha(\delta_n) = P\{|q_n(x_1, \dots, x_n)| \geq c | H_0\} \rightarrow P\{|\eta| \geq c\} = \alpha, \quad n \rightarrow \infty.$$

Тобто  $\delta_n$  – критерій асимптотичного рівня значущості  $\alpha$ .

Перевіримо спроможність. З умови (8.9) випливає, що

$$\forall c > 0 \quad P\{|q_n| < c\} \rightarrow 0, \quad \text{коли } n \rightarrow \infty.$$

Але тоді

$$\beta(\delta_n) = P\{|q_n(x_1, \dots, x_n)| < c \mid \overline{H}_0\} \rightarrow 0 \quad \text{при } n \rightarrow \infty. \quad \square$$

**Зауваження 8.2.** Якщо розподіл випадкової величини  $q_n(x_1, \dots, x_n)$  при побудові критерію згоди точно відомий, то відповідний критерій буде мати точний рівень значущості  $\alpha$ .

## 8.6. Фактично досяжний рівень значущості

Існує також точка зору про те, що при перевірці гіпотези  $H_0$  немає необхідності попередньо задавати фіксований рівень значущості. Краще визначити "фактично досяжний".

Нехай  $S_\alpha$  – критична область критерію рівня значущості  $\alpha$ . Припустимо, що  $S_{\alpha_2} \subset S_{\alpha_1}$  при  $\alpha_2 < \alpha_1$  (як у критерії Пірсона).

**Означення.** *Фактично досяжним рівнем значущості* критерію на реалізації  $(x_1, \dots, x_n)$  називають величину

$$\alpha^* = \inf\{\alpha : x \in \overline{S}_\alpha\}.$$

З огляду на те, наскільки малим є значення  $\alpha^*$ , гіпотезу  $H_0$  відхиляють або ні.

Саме фактично досяжний рівень значущості є результатом перевірки гіпотез в будь-якому статистичному пакеті.

**Контрольні запитання:**

1. Які відомі способи порівняння критеріїв? Що таке потужність критерію? Чому, загалом, не виходить одночасно мінімізувати рівень значущості і потужність критерію?
2. Охарактеризуйте мінімаксні, баєсовські та найбільш потужні критерії як способи порівняння критеріїв.
3. Сформулюйте означення критерію відношення правдоподібності та теорему про його властивості.
4. Чому важливо знати асимптотичну поведінку помилок першого та другого роду критерію?
5. Який загальний принцип побудови критеріїв згоди та які у них властивості?
6. Відповіді на які питання можна отримати за допомогою фактично досяжного рівня значущості?

**Вправи та задачі: 8.1.** Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $N(0, \theta^2)$ ,  $\theta > 0$ . Побудувати НПК рівня  $\varepsilon$  для перевірки гіпотези

$$H_0 : \theta = \sigma_1 \text{ проти альтернативи } H_1 : \theta = \sigma_2, \sigma_1 < \sigma_2.$$

**8.2.** Побудувати критерій перевірки гіпотез

$$H_0 : P_\xi = N(a, \sigma^2) \text{ за альтернативи } H_1 : P_\xi = \Pi(\lambda)$$

з нульовими ймовірностями першого та другого роду.

**8.3.** Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi = B(m, \theta)$ , де  $\theta$  може приймати значення  $1/3$  та  $2/3$  з апріорними ймовірностями  $1/5$  та  $4/5$ , а параметр  $m$  відомий. Побудувати баєсовський критерій для перевірки гіпотез  $H_0 : \theta = a$  проти альтернативи  $H_1 : \theta = a_1$ ,  $0 < a < a_1 < 1$ .

## §9. Перевірка гіпотез про параметри нормальних розподілів

### 9.1. Перевірка гіпотез про параметри нормальних розподілів

Нехай  $(x_1, \dots, x_n)$  — вибірка обсягу  $n$  з розподілу  $N(\theta_1, \sigma^2)$ , дисперсія якого відома. У прикладі 8.3 розглядався критерій перевірки гіпотези  $H_0 : \theta_1 = a_0$  проти простої альтернативи  $H_1 : \theta_1 = a_1 > a_0$ .

Але досить часто більш прийнятною є складна альтернатива про те, що  $\theta_1 \neq a_0$ , або  $\theta_1 > a_0$ , чи  $\theta_1 < a_0$ . Проаналізуємо випадок двосторонньої альтернативи  $H_1 = \bar{H}_0 : \theta_1 \neq a_0$ , тобто побудуємо критерій згоди, який би за однією реалізацією вибірки з даною ймовірністю помилки першого роду визначав, чи узгоджуються результати спостережень з основною гіпотезою  $H_0 : \theta_1 = a_0$ .

Як зазвичай для критеріїв згоди, підберемо необхідну тестову статистику. Найкращою оцінкою параметра  $\theta \in \bar{x}$ . У разі, коли гіпотеза  $H_0$  правдива, відхилення  $|\bar{x} - a_0|$  набувають як завгодно малих значень зі зростанням обсягу вибірки  $n$ . Якщо ж  $H_0$  — хибна, то  $x_n$  збігається м.с. до іншого значення, відмінного від  $a_0$ , і відхилення  $|\bar{x} - a_0|$ , вочевидь, будуть суттєві.

Нормовані відхилення мають стандартний нормальний розподіл, тому як тестову статистику оберемо

$$q_n(x_1, \dots, x_n) = \frac{\bar{x} - a_0}{\sigma} \sqrt{n} \in N(0, 1).$$

Далі задамо рівень значущості  $\alpha$  і з умови  $P\{|q_n| > c\} = \alpha$  знайдемо сталу  $c$  як квантиль рівня  $1 - \alpha/2$  стандартного розподілу:  $c = d_{1-\alpha/2}$ .

Таким чином, критична область шуканого критерію має вигляд:

$$S = \{(x_1, \dots, x_n) \in X : \frac{|\bar{x} - a_0|}{\sigma} \sqrt{n} > d_{1-\frac{\alpha}{2}}\}. \quad (9.1)$$

Критична область критерію для перевірки гіпотези  $H_0 : \theta_1 = a_1$  за складної правосторонньої альтернативи  $H_1 : \theta_1 > a_0$  у разі відомої дисперсії має вигляд:

$$S = \{(x_1, \dots, x_n) \in X : \frac{\bar{x} - a_0}{\sigma} \sqrt{n} > d_{1-\alpha}\} \quad (9.2)$$

(у даному випадку природно скористатися  $\alpha$ -верхньою границею стандартного розподілу, але в додатках (табл. II), наведені його квантилі, тому в правій частині нерівності квантиль  $d_{1-\alpha}$  (див. заув. 8.1.)).

### 9.2. Гіпотези про середнє при невідомій дисперсії

Розглядається та ж задача, що і в попередньому пункті, але у випадку невідомої дисперсії:  $P_\xi = N(\theta_1, \theta_2^2)$ .

За тестову статистику візьмемо випадкову величину

$$q_n = \sqrt{n} \frac{\bar{x} - a_0}{\bar{s}}, \text{ де } \bar{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

– варіанса. З'ясуємо розподіл  $q_n$ . Для цього запишемо:

$$q_n = \frac{\sqrt{n}(\bar{x} - a_0)}{\theta_2} \cdot \frac{\theta_2}{\bar{s}},$$

де  $\eta = \sqrt{n}(\bar{x} - a_0)/\theta_2 \in N(0, 1)$ , а

$$\frac{\theta_2}{\bar{s}} = \frac{1}{\sqrt{\frac{(n-1)s^2}{\theta_2^2} \cdot \frac{1}{n-1}}}$$

і  $(n-1)\bar{s}^2/\theta_2^2 \in \chi^2(n-1)$  за лемою 5.1. Тоді випадкова величина  $q_n \in T(n-1)$  має розподіл Стюдента з  $n-1$  ступенями свободи.

Побудова критичних областей відбувається так як і в (9.1) та (9.2), відповідно, тільки замість квантилів стандартного розподілу потрібно взяти квантилі розподілу Стюдента.

### 9.3. Перевірка гіпотези про значення дисперсії

Нехай  $(x_1, \dots, x_n)$  – вибірка з розподілу  $P_\xi = N(\theta_1, \theta_2^2)$ , обидва параметри якого невідомі. За вибіркою перевіряється гіпотеза  $H_0 : \theta_2 = \sigma_0$  за складної двосторонньої альтернативи:  $H_1 = \bar{H}_0 : \theta_2 \neq \sigma_0$ .

Незалежно від гіпотез, варіанса є найкращою оцінкою невідомого параметра  $\theta_2$ . Тому, якщо гіпотеза  $H_0$  справедлива, відношення  $\bar{s}^2/\sigma_0^2$  мало відрізняється від одиниці. А саме, якщо  $\theta_2^2/\sigma_0^2 = 1$ , то

$$M(\bar{s}^2/\sigma_0^2) = M(\bar{s}^2/\theta_2^2) = 1$$

і, за лемою 2.1,

$$\frac{\bar{s}^2}{\sigma_0^2} \xrightarrow{P} \frac{\theta_2^2}{\sigma_0^2} = 1.$$

Якщо ж  $\theta_2^2/\sigma_0^2 \neq 1$ , то

$$M\left(\frac{\bar{s}^2}{\sigma_0^2}\right) = \frac{\theta_2^2}{\sigma_0^2} \neq 1 \quad \text{і} \quad \frac{\bar{s}^2}{\sigma_0^2} \xrightarrow{P} \frac{\theta_2^2}{\sigma_0^2} \neq 1.$$

Тому за тестову статистику беремо  $q_n = \bar{s}^2/\sigma_0^2$  і критична область такого критерію має вигляд

$$S = \left\{ x \in X : 0 < \frac{\bar{s}^2}{\sigma_0^2} < c_1 < 1 \text{ або } \frac{\bar{s}^2}{\sigma_0^2} > c_2 > 1 \right\}.$$

Залишається визначитись зі сталими  $c_1$  та  $c_2$ .

Випадкова величина  $(n-1)\bar{s}^2/\sigma_0^2$  має розподіл  $\chi^2(n-1)$ , тому необхідні сталі легко знаходимо за допомогою верхніх границь даного розподілу. На заданому рівні значущості  $\alpha$  маємо:

$$c_1 = \frac{1}{n-1} \chi_{1-\alpha/2}^2(n-1), \quad \text{а} \quad c_2 = \frac{1}{n-1} \chi_{\alpha/2}^2(n-1).$$

Остаточно, якщо на реалізації вибірки

$$\frac{\bar{s}^2}{\sigma_0^2} \notin \left( \frac{1}{n-1} \chi_{1-\alpha/2}^2(n-1), \frac{1}{n-1} \chi_{\alpha/2}^2(n-1) \right),$$

то основна гіпотеза відхиляється на рівні значущості  $\alpha$ . Для вправи пропонується побудувати критерії для односторонніх альтернатив.

#### 9.4. Критерій Фішера

Досить часто в практичних задачах виникає необхідність порівняти деякі генеральні сукупності, які мають нормальний розподіл. Наприклад, якщо

цікавить результативність нової технології виготовлення певної продукції, або продуктивність праці після навчання персоналу і т.і.

З математичної точки зору, треба забезпечити дві репрезентативні вибірки з відповідних генеральних сукупностей і порівняти їх параметри розподілу. Кажуть, перевірити вибірки на однорідність.

Іноді достатньо порівняти середнє значення, але цю задачу розв'язує критерій Стюдента, який ми розглянемо у наступному пункті. Для застосування критерію Стюдента потрібно знати, чи рівні відповідні дисперсії. Тому почнемо з задачі про порівняння дисперсій.

Нехай  $(x_1, \dots, x_n)$  і  $(y_1, \dots, y_n)$  — дві незалежні вибірки з розподілів  $N(\theta_1, \theta_2^2)$  і  $N(\theta_3, \theta_4^2)$ , параметри яких невідомі. Треба перевірити основну гіпотезу  $H_0 : \theta_2 = \theta_4$  за складної двосторонньої альтернативи  $H_1 = \overline{H_0}$ .

Як зазвичай для критерію згоди, оберемо тестову статистику. Найкращими оцінками параметрів  $\theta_2^2$  та  $\theta_4^2$  є варіанси  $\bar{s}_x^2$  та  $\bar{s}_y^2$ , тому розглянемо статистику  $q_{n,m} = \bar{s}_x^2 / \bar{s}_y^2$ .

**Теорема 9.1 (Фішер).** *Якщо  $(x_1, \dots, x_n)$  і  $(y_1, \dots, y_n)$  — дві незалежні вибірки з розподілів  $N(\theta_1, \theta_2^2)$  та  $N(\theta_3, \theta_4^2)$ , параметри яких невідомі,  $\bar{s}_x^2$  і  $\bar{s}_y^2$  — вибіркові варіанси і справедлива гіпотеза  $H_0 : \theta_2 = \theta_4$ , то випадкова величина  $\bar{s}_x^2 / \bar{s}_y^2$  має розподіл Фішера з  $(n-1, m-1)$  ступенями свободи.*

За лемою 5.1, відношення

$$\frac{(n-1)\bar{s}_x^2}{\theta_2^2} \quad \text{і} \quad \frac{(m-1)\bar{s}_y^2}{\theta_4^2}$$

мають розподіли  $\chi^2(n-1)$  та  $\chi^2(m-1)$ , відповідно.

Але тоді за умови справедливості основної гіпотези  $H_0$  маємо:

$$q_{n,m} = \frac{\bar{s}_x^2}{\bar{s}_y^2} = \frac{\frac{1}{n-1}\chi^2(n-1)}{\frac{1}{m-1}\chi^2(m-1)} \in \Phi(n-1, m-1).$$

Якщо гіпотеза  $H_0$  правдива, тобто  $\theta_2/\theta_4 = 1$ , то для більшості реалізацій  $\bar{s}_x^2/\bar{s}_y^2 \approx 1$ . Отже, критична область має такий вигляд:

$$S = \left\{ (x_1, \dots, x_n) \in X : \frac{\bar{s}_x^2}{\bar{s}_y^2} < c_1 < 1 \text{ або } \frac{\bar{s}_x^2}{\bar{s}_y^2} > c_2 > 1 \right\}.$$



Залишається обрати сталі  $c_1$  та  $c_2$ , щоб забезпечити необхідний рівень значущості  $\alpha$ . З'ясуємо це питання мовою верхніх  $\alpha$ -границь розподілу Фішера з  $(n-1, m-1)$  ступенями свободи  $f_\alpha(n-1, m-1)$  :

$$c_1 = f_{1-\alpha/2}(n-1, m-1), \quad c_2 = f_{\alpha/2}(n-1, m-1).$$

Зазначимо, що для верхніх  $\alpha$ -границь розподілу Фішера справедливе співвідношення

$$f_\alpha(n, m) = \frac{1}{f_{1-\alpha}(m, n)}. \quad (9.3)$$

Дійсно, нехай  $f(n, m) \in \Phi(n, m)$ . За означенням верхньої границі  $f_{1-\alpha}(m, n)$  і властивостями 1,2 розподілу Фішера (п. 5.3), маємо

$$\begin{aligned} 1 - \alpha &= P\{f(n, m) \geq f_{1-\alpha}(n, m)\} = \\ &= P\left\{\frac{1}{f(n, m)} \leq \frac{1}{f_{1-\alpha}(n, m)}\right\} = 1 - P\left\{f(m, n) \geq \frac{1}{f_{1-\alpha}(n, m)}\right\}. \end{aligned}$$

Звідси

$$P\left\{f(m, n) \geq \frac{1}{f_{1-\alpha}(n, m)}\right\} = \alpha,$$

що і означає (9.3).

Залишається надати точне формулювання критерію Фішера.

**Теорема 9.2 (Критерій Фішера).** *Нехай виконуються умови теореми 9.1 і  $f_{\alpha/2}(n-1, m-1)$  — квантиль рівня  $\alpha/2$  розподілу Фішера з  $(n-1, m-1)$  ступенями свободи.*

*Тоді основну гіпотезу  $H_0 : \theta_2 = \theta_4$  відхиляють на користь альтернативи  $H_1 = \bar{H}_0$ , коли*

$$\frac{\bar{s}_x^2}{\bar{s}_y^2} \notin \left( \frac{1}{f_{\alpha/2}(m-1, n-1)}, f_{\alpha/2}(n-1, m-1) \right),$$

*з ймовірністю помилки першого роду, що дорівнює  $\alpha$ .*

**Зауваження 9.1.** У випадку, коли  $\bar{s}_x^2$  суттєво більша за  $\bar{s}_y^2$  і, можливо, з апріорної інформації про задачу, що розглядається, як альтернативу розглядають  $H_1 : \theta_1 > \theta_2$ . Тоді основну гіпотезу  $H_0 : \theta_1 = \theta_2$  відхиляють, коли

$$\frac{\bar{s}_x^2}{\bar{s}_y^2} > f_\alpha(n-1, m-1)$$

*з ймовірністю помилки першого роду, що дорівнює  $\alpha$ .*

### 9.5. Критерій Стюдента

У цьому пункті розглядається критерій перевірки гіпотези про рівність середніх значень двох незалежних нормальних розподілів, який належить Вільяму Госсету. Він у 1908 році опублікував свій результат у статті журналу «Біометрика» під псевдонімом Student. Звідси назва критерію та розподілу Стюдента. Вибір псевдоніма був пов'язаний з його роботою, яка потребувала нерозголошення комерційної таємниці. Госсет працював над оцінкою якості пива в компанії Гіннес.

Нехай  $(x_1, \dots, x_n)$  і  $(y_1, \dots, y_n)$  – дві незалежні вибірки з розподілів  $N(\theta_1, \theta_1^2)$  та  $N(\theta_3, \theta_3^2)$ , параметри яких невідомі, але відомо, що дисперсії рівні (наприклад, за критерієм Фішера). Перевіряється гіпотеза  $H_0 : \theta_1 = \theta_3$  або, що те саме,  $(x_1, \dots, x_n)$  і  $(y_1, \dots, y_n)$  – вибірки з одного й того самого розподілу. Альтернатива може бути складною двосторонньою:  $H_1 = \bar{H}_0 : \theta_1 \neq \theta_3$ , або односторонньою  $H_1 : \theta_1 > \theta_3$  чи  $H_1 : \theta_1 < \theta_3$  залежно від задачі, що вивчається.

Розглянемо складну двосторонню альтернативу  $H_1 = \bar{H}_0$ , тобто побудуємо критерій згоди. Незалежно від гіпотез, найкращими оцінками невідомих параметрів  $\theta_1$  та  $\theta_3 \in \bar{x}$  і  $\bar{y}$ . Тому, коли гіпотеза  $H_0$  правдива і  $\theta_1 = \theta_3$ , значення  $\bar{x}$  і  $\bar{y}$  повинні мало відрізнятися. Як тестову статистику Госсет запропонував нормовану різницю

$$q_{n,m} = \frac{\bar{x} - \bar{y}}{s\sqrt{(n+m)/nm}}, \text{ де } s = \frac{(n-1)\bar{s}_x^2 + (m-1)\bar{s}_y^2}{n+m-2}. \quad (9.4)$$

Розподіл випадкової величини  $q_{n,m}$  визначає така теорема.

**Теорема 9.3 (Госсет).** *Нехай  $(x_1, \dots, x_n)$  і  $(y_1, \dots, y_n)$  – дві незалежні вибірки з розподілів  $N(\theta_1, \theta_1^2)$  та  $N(\theta_3, \theta_3^2)$ , параметри яких невідомі, але відомо, що дисперсії рівні,  $\bar{x}$  і  $\bar{y}$  – емпіричні середні, а  $\bar{s}_x^2$  та  $\bar{s}_y^2$  – варіанси, розраховані за реалізаціями, відповідно.*

*Тоді за умови справедливості основної гіпотези  $H_0 : \theta_1 = \theta_3$  статистика  $q_{n,m}$ , визначена в (9.4), розподілена за законом Стюдента з  $(n+m-2)$  ступенями свободи.*

Оскільки критична область шуканого критерію має вигляд:

$$S = \left\{ (x_1, \dots, x_n) \in X : \frac{|\bar{x} - \bar{y}|}{s\sqrt{n+m}} \sqrt{nm} \geq c \right\},$$

де величина  $s$  визначена в (9.4), задамо рівень значущості  $\alpha$  і знайдемо необхідну сталу  $c$ . Для випадкової величини  $t(n+m-2) \in T(n+m-2)$  маємо:

$$\alpha = P \{ |t(n+m-2)| > c \} = 2 - 2F_t(c)$$

(врахували симетричність розподілу Стьюдента). Звідси

$$2F(c) = 2 - \alpha \Leftrightarrow F(c) = 1 - \alpha/2$$

і тоді  $c = t_{1-\alpha/2}(n+m-2)$  — квантиль розподілу Стьюдента з  $(n+m-2)$  ступенями свободи.

**Теорема 9.4(критерій Стьюдента).** *Нехай виконуються умови теореми 9.10 і  $t_{1-\alpha/2}(n+m-2)$  — квантиль розподілу Стьюдента з  $(n+m-2)$  ступенями свободи.*

*Основну гіпотезу  $H_0 : \theta_1 = \theta_3$  відхиляють на користь альтернативи  $H_1 = \bar{H}_0$ , коли*

$$\frac{|\bar{x} - \bar{y}|}{s\sqrt{n+m}} \sqrt{nm} \geq t_{1-\frac{\alpha}{2}}(n+m-2)$$

*з імовірністю помилки першого роду, що дорівнює  $\alpha$ .*

**Зауваження 9.2.** У разі, коли середнє значення  $\bar{x}$  суттєво більше ніж  $\bar{y}$ , або з апіорної інформації про задачу, що розглядається, висувають односторонню альтернативу. Наприклад,  $H_1 : \theta_1 > \theta_3$ .

Тоді критична область рівня значущості  $\alpha$  визначається як

$$S = \left\{ (x_1, \dots, x_n) \in X : \frac{|\bar{x} - \bar{y}|}{s\sqrt{n+m}} \sqrt{nm} \geq t_{1-\alpha}(n+m-2) \right\}$$

(усі позначення збігаються з позначеннями теореми 9.4).

**Приклад 9.1.** За двома реалізаціями незалежних вибірок обсягами  $n = 5$  та  $m = 6$  з нормальних розподілів  $N(\theta_1, \theta_1^2)$  та  $N(\theta_3, \theta_3^2)$ , параметри яких невідомі, розрахували середні вибірові  $\bar{x} = 3,3$ ,  $\bar{y} = 2,48$  та

виправлені емпіричні дисперсії  $\bar{s}_x^2 = 0,25$  та  $\bar{s}_y^2 = 0,108$ . При рівні значущості  $\alpha = 0,05$  перевірити основну гіпотезу  $H_0 : \theta_1 = \theta_3$  за альтернативи  $H_1 = \bar{H}_0$ .

Виправлена вибіркова дисперсія  $\bar{s}_x^2$  суттєво більша за  $\bar{s}_y^2$ . Перевіримо спочатку гіпотезу про рівність дисперсій:

$$H_0 : \theta_2 = \theta_4 \text{ проти альтернативи } H_1 = \bar{H}_0.$$

Розрахуємо  $q_{n,m} = 0,25/0,108 \approx 2,31$ .

Критична область – правостороння і при  $\alpha = 0,05$ ,

$$f_{0,05}(4,5) = 6,3 > 2,31.$$

Таким чином, вибіркові дані не протирічать гіпотезі про рівність дисперсій. Можна застосувати критерій Стюдента для перевірки розподілів на однорідність. Розглянемо

$$H_0 : \theta_1 = \theta_3 \text{ за альтернативи } H_1 = \bar{H}_0.$$

Тепер, за формулами (9.4),  $q_{n,m} = 0,17$ , критична область двостороння і при  $\alpha = 0,05$ ,  $n = 5$ ,  $m = 6$

$$t_{1-0,05/2}(9) = 2,26 > 0,17.$$

Значення тестової статистики критерію Стюдента на реалізаціях вибірок  $(x_1, \dots, x_n)$  і  $(y_1, \dots, y_n)$  виявилось меншим, ніж критичне значення, тобто основна гіпотеза про рівність середніх не відхиляється.

Усі наведені в цьому параграфі критерії є точними (точного рівня значущості  $\alpha$ ) та спроможними, тобто найкращими. У випадку вибірок із сукупностей, які не мають нормального розподілу, їх також можна застосовувати при великих обсягах вибірок, оскільки середні вибіркові  $\bar{x}$  та  $\bar{y}$  будь-яких генеральних сукупностей розподілені асимптотично нормально.

### Контрольні запитання:

1. Проаналізуйте, будь ласка, чому критерії перевірки гіпотез про математичне сподівання нормального розподілу за відомою і невідомою дисперсіями будують на основі різних статистик?

**2.** Чим відрізняється перевірка гіпотези про значення дисперсії в п. 9.3 від критерія Фішера?

**3.** Чи можна задачу п. 9.5 розв'язати за допомогою критерію Пірсона?

**4.** Наведіть приклади практичних задач на застосування критеріїв Фішера та Стюдента.

**Вправи та задачі: 9.1.** Проаналізуйте, будь ласка, зв'язок між критеріями перевірки гіпотез про параметри нормального розподілу та довірчими інтервалами відповідних параметрів.

**9.2.** Покажіть, що якщо за умов критерію Стюдента гіпотеза  $H_0 : \theta_1 = \theta_3$  не справджується (наприклад,  $\theta_1 > \theta_3$ ) то

$$\begin{aligned}\xi_{n,m} &= \frac{\bar{x} - \bar{y}}{s\sqrt{n+m}}\sqrt{nm} = \frac{(\bar{x} - \theta_1) - (\bar{y} - \theta_2)}{s\sqrt{n+m}}\sqrt{nm} + \\ &+ \frac{\theta_1 - \theta_2}{s\sqrt{n+m}}\sqrt{nm} \xrightarrow{P} \infty.\end{aligned}$$

**9.3.** Побудуйте критерій Стюдента за складної односторонньої альтернативи  $H_1 : \theta_1 > \theta_2$ .

Таблиця I

**Коефіцієнти асиметрії  $A$  та ексцесу  $E$  основних розподілів**

| Розподіл         | $A$     | $E$     |
|------------------|---------|---------|
| Нормальний       | 0       | 0       |
| Експотенціальний | 2       | 6       |
| Максвелла        | 0,06544 | 1,56997 |
| Рівномірний      | 0       | 1,2     |
| Лапласа          | 1,12132 | 3       |
| Релея            | 0,631   | 0,245   |

Таблиця II

**Квантилі стандартного розподілу  $N(0, 1)$**

| $\alpha$ | $d_{1-\frac{\alpha}{2}}$ |
|----------|--------------------------|
| 0,80     | 0,25                     |
| 0,50     | 0,67                     |
| 0,40     | 0,84                     |
| 0,30     | 1,04                     |
| 0,25     | 1,15                     |
| 0,20     | 1,28                     |
| 0,15     | 1,44                     |
| 0,10     | 1,64                     |
| 0,05     | 1,96                     |
| 0,04     | 2,05                     |
| 0,02     | 2,33                     |
| 0,0100   | 2,58                     |
| 0,0050   | 2,81                     |
| 0,0020   | 3,09                     |
| 0,0010   | 3,29                     |
| 0,0001   | 3,89                     |

Таблиця III

**Квантилі  $t_{1-\alpha/2}(n)$  розподілів Стьюдента  $T(n)$** 

| Ступені<br>свободи, n | Рівень значущості $\alpha$ |      |       |       |       |        |        |
|-----------------------|----------------------------|------|-------|-------|-------|--------|--------|
|                       | 0,20                       | 0,10 | 0,05  | 0,02  | 0,01  | 0,005  | 0,001  |
| 1                     | 3,08                       | 6,31 | 12,71 | 31,82 | 63,66 | 127,32 | 636,62 |
| 2                     | 1,89                       | 2,92 | 4,30  | 6,97  | 9,93  | 14,09  | 31,60  |
| 3                     | 1,64                       | 2,35 | 3,18  | 4,54  | 5,84  | 7,45   | 12,94  |
| 4                     | 1,53                       | 2,13 | 2,78  | 3,75  | 4,60  | 5,60   | 8,61   |
| 5                     | 1,48                       | 2,02 | 2,57  | 3,37  | 4,03  | 4,47   | 6,86   |
| 6                     | 1,44                       | 1,94 | 2,45  | 3,14  | 3,71  | 4,32   | 5,96   |
| 7                     | 1,42                       | 1,90 | 2,37  | 3,00  | 3,50  | 4,03   | 5,41   |
| 8                     | 1,40                       | 1,86 | 2,31  | 2,90  | 3,36  | 3,84   | 5,04   |
| 9                     | 1,38                       | 1,83 | 2,26  | 2,82  | 3,25  | 3,69   | 4,78   |
| 10                    | 1,37                       | 1,81 | 2,23  | 2,76  | 3,17  | 3,58   | 4,59   |
| 11                    | 1,36                       | 1,80 | 2,20  | 2,72  | 3,11  | 3,50   | 4,44   |
| 12                    | 1,36                       | 1,78 | 2,18  | 2,68  | 3,06  | 3,43   | 4,32   |
| 13                    | 1,35                       | 1,77 | 2,16  | 2,65  | 3,01  | 3,37   | 4,22   |
| 14                    | 1,34                       | 1,76 | 2,15  | 2,65  | 2,98  | 3,33   | 4,14   |
| 15                    | 1,34                       | 1,75 | 2,13  | 2,60  | 2,95  | 3,29   | 4,07   |
| 16                    | 1,34                       | 1,75 | 2,12  | 2,58  | 2,93  | 3,25   | 4,02   |
| 17                    | 1,33                       | 1,74 | 2,11  | 2,57  | 2,90  | 3,22   | 3,97   |
| 18                    | 1,33                       | 1,73 | 2,10  | 2,55  | 2,88  | 3,20   | 3,92   |
| 19                    | 1,33                       | 1,73 | 2,09  | 2,54  | 2,86  | 3,17   | 3,88   |
| 20                    | 1,33                       | 1,73 | 2,09  | 2,53  | 2,85  | 3,15   | 3,85   |
| 21                    | 1,32                       | 1,72 | 2,08  | 2,52  | 2,83  | 3,14   | 3,82   |
| 22                    | 1,32                       | 1,72 | 2,07  | 2,51  | 2,82  | 3,12   | 3,79   |
| 23                    | 1,32                       | 1,71 | 2,07  | 2,50  | 2,81  | 3,10   | 3,77   |
| 24                    | 1,32                       | 1,71 | 2,06  | 2,49  | 2,80  | 3,09   | 3,75   |
| 25                    | 1,32                       | 1,71 | 2,06  | 2,48  | 2,79  | 3,08   | 3,73   |
| 26                    | 1,32                       | 1,71 | 2,06  | 2,48  | 2,78  | 3,07   | 3,71   |
| 27                    | 1,31                       | 1,70 | 2,05  | 2,47  | 2,77  | 3,06   | 3,69   |
| 28                    | 1,31                       | 1,70 | 2,05  | 2,47  | 2,76  | 3,05   | 3,67   |

## Продовження таблиці III

| Ступені<br>свободи, f | Рівень значущості $\alpha$ |      |      |      |      |       |       |
|-----------------------|----------------------------|------|------|------|------|-------|-------|
|                       | 0,20                       | 0,10 | 0,05 | 0,02 | 0,01 | 0,005 | 0,001 |
| 29                    | 1,31                       | 1,70 | 2,04 | 2,46 | 2,76 | 3,04  | 3,66  |
| 30                    | 1,31                       | 1,70 | 2,04 | 2,46 | 2,75 | 3,03  | 3,65  |
| 40                    | 1,30                       | 1,68 | 2,02 | 2,42 | 2,70 | 2,97  | 3,55  |
| 60                    | 1,30                       | 1,67 | 2,00 | 2,39 | 2,66 | 2,91  | 3,46  |
| 120                   | 1,29                       | 1,66 | 1,98 | 2,36 | 2,62 | 2,86  | 3,37  |
| $\infty$              | 1,28                       | 1,64 | 1,96 | 2,33 | 2,58 | 2,81  | 3,29  |

## Таблиця IV

**Розв'язки  $q = q(\gamma, n)$  інтегральних рівнянь<sup>1</sup> розподілів**

$\chi(n-1)$

| n  | Довірча ймовірність $\gamma$ |      |       | n   | Довірча ймовірність $\gamma$ |       |       |
|----|------------------------------|------|-------|-----|------------------------------|-------|-------|
|    | 0,95                         | 0,99 | 0,999 |     | 0,95                         | 0,99  | 0,999 |
| 5  | 1,37                         | 2,67 | 5,64  | 20  | 0,37                         | 0,58  | 0,88  |
| 6  | 1,09                         | 2,01 | 2,88  | 25  | 0,32                         | 0,49  | 0,73  |
| 7  | 0,92                         | 1,62 | 2,98  | 30  | 0,28                         | 0,43  | 0,63  |
| 8  | 0,80                         | 1,38 | 2,42  | 35  | 0,26                         | 0,38  | 0,56  |
| 9  | 0,71                         | 1,20 | 2,06  | 40  | 0,24                         | 0,35  | 0,50  |
| 10 | 0,65                         | 1,08 | 1,80  | 45  | 0,22                         | 0,32  | 0,46  |
| 11 | 0,59                         | 0,98 | 1,60  | 50  | 0,21                         | 0,30  | 0,43  |
| 12 | 0,55                         | 0,90 | 1,45  | 60  | 0,188                        | 0,209 | 0,38  |
| 13 | 0,52                         | 0,83 | 1,33  | 70  | 0,174                        | 0,245 | 0,34  |
| 14 | 0,48                         | 0,78 | 1,23  | 80  | 0,101                        | 0,226 | 0,31  |
| 15 | 0,46                         | 0,73 | 1,15  | 90  | 0,151                        | 0,211 | 0,29  |
| 16 | 0,44                         | 0,70 | 1,07  | 100 | 0,143                        | 0,198 | 0,27  |
| 17 | 0,42                         | 0,66 | 1,01  | 150 | 0,115                        | 0,160 | 0,211 |
| 18 | 0,40                         | 0,63 | 0,96  | 200 | 0,099                        | 0,136 | 0,185 |
| 19 | 0,39                         | 0,60 | 0,92  | 250 | 0,089                        | 0,120 | 0,162 |

<sup>1</sup>див. (5.7)



Таблиця V

Верхні  $\alpha$ -границі розподілів  $\chi^2(n)$

| n  | Значення $\alpha$ |       |       |       |       |       |       |       |
|----|-------------------|-------|-------|-------|-------|-------|-------|-------|
|    | 0,990             | 0,975 | 0,950 | 0,900 | 0,100 | 0,050 | 0,025 | 0,010 |
| 1  | 0,00              | 0,00  | 0,00  | 0,02  | 2,71  | 3,84  | 5,02  | 6,64  |
| 2  | 0,02              | 0,05  | 0,10  | 0,21  | 4,61  | 5,99  | 7,38  | 9,21  |
| 3  | 0,12              | 0,22  | 0,35  | 0,58  | 6,23  | 7,82  | 9,35  | 11,34 |
| 4  | 0,30              | 0,48  | 0,71  | 1,06  | 7,78  | 9,48  | 11,14 | 13,28 |
| 5  | 0,55              | 0,83  | 1,14  | 1,61  | 9,24  | 11,07 | 12,83 | 15,09 |
| 6  | 0,87              | 1,24  | 1,64  | 2,20  | 10,64 | 12,59 | 14,45 | 16,81 |
| 7  | 1,24              | 1,69  | 2,17  | 2,83  | 12,02 | 14,07 | 16,01 | 18,48 |
| 8  | 1,65              | 2,18  | 2,73  | 3,49  | 13,36 | 15,51 | 17,53 | 20,09 |
| 9  | 2,90              | 2,70  | 3,32  | 4,17  | 14,68 | 16,92 | 19,02 | 21,67 |
| 10 | 2,56              | 3,25  | 3,94  | 4,36  | 15,99 | 18,31 | 20,48 | 23,21 |
| 11 | 3,05              | 3,82  | 4,58  | 5,58  | 17,28 | 19,68 | 21,92 | 24,72 |
| 12 | 3,57              | 4,40  | 5,23  | 6,30  | 18,55 | 21,03 | 23,34 | 26,22 |
| 13 | 4,11              | 5,01  | 5,89  | 7,04  | 19,81 | 22,36 | 24,74 | 27,69 |
| 14 | 4,66              | 5,63  | 6,57  | 7,79  | 21,06 | 23,68 | 26,12 | 29,14 |
| 15 | 5,23              | 6,26  | 7,26  | 8,55  | 22,31 | 25,00 | 27,49 | 30,08 |
| 16 | 5,81              | 6,91  | 7,96  | 9,31  | 23,54 | 26,30 | 28,85 | 32,00 |
| 17 | 6,41              | 7,56  | 8,67  | 10,09 | 24,77 | 27,59 | 30,19 | 33,41 |
| 18 | 7,02              | 8,23  | 9,39  | 10,86 | 25,99 | 28,87 | 31,53 | 34,81 |
| 19 | 7,63              | 8,91  | 10,12 | 11,65 | 27,20 | 30,14 | 32,85 | 36,19 |
| 20 | 8,26              | 9,59  | 10,85 | 12,44 | 28,41 | 31,41 | 34,17 | 37,57 |
| 21 | 8,90              | 10,28 | 11,59 | 13,24 | 29,62 | 32,67 | 35,48 | 38,93 |
| 22 | 9,54              | 10,98 | 12,34 | 14,04 | 30,81 | 33,92 | 36,78 | 40,29 |
| 23 | 10,20             | 11,69 | 13,09 | 14,85 | 32,01 | 35,17 | 38,08 | 41,64 |
| 24 | 10,86             | 12,40 | 13,85 | 15,66 | 33,20 | 36,42 | 39,36 | 42,92 |
| 25 | 11,52             | 13,12 | 14,61 | 16,47 | 34,38 | 37,65 | 40,65 | 44,31 |
| 26 | 12,20             | 12,84 | 15,38 | 17,29 | 35,56 | 38,89 | 41,92 | 45,64 |
| 27 | 12,88             | 14,57 | 16,15 | 18,11 | 36,74 | 40,11 | 43,19 | 46,96 |
| 28 | 13,56             | 15,31 | 16,93 | 18,94 | 37,92 | 41,34 | 44,46 | 48,26 |

Продовження таблиці V

| n   | Значення $\alpha$ |       |       |       |       |       |       |       |
|-----|-------------------|-------|-------|-------|-------|-------|-------|-------|
|     | 0,990             | 0,975 | 0,950 | 0,900 | 0,100 | 0,050 | 0,025 | 0,010 |
| 29  | 14,26             | 16,05 | 17,71 | 19,77 | 39,09 | 42,56 | 45,72 | 49,59 |
| 30  | 14,95             | 16,79 | 18,49 | 20,60 | 40,26 | 43,77 | 46,98 | 50,89 |
| 40  | 22,16             | 24,43 | 26,51 | 29,05 | 51,80 | 55,76 | 59,34 | 63,69 |
| 50  | 29,71             | 32,36 | 34,76 | 37,69 | 63,17 | 67,50 | 71,42 | 76,15 |
| 60  | 37,48             | 40,48 | 43,19 | 46,46 | 74,40 | 79,08 | 83,30 | 88,38 |
| 70  | 45,44             | 48,76 | 51,74 | 55,33 | 85,53 | 90,53 | 95,02 | 100,4 |
| 80  | 53,54             | 57,15 | 60,39 | 64,28 | 96,58 | 101,9 | 106,6 | 112,3 |
| 90  | 61,75             | 65,65 | 69,13 | 73,29 | 107,6 | 113,1 | 118,1 | 124,1 |
| 100 | 70,06             | 74,22 | 77,93 | 82,36 | 118,5 | 124,3 | 129,6 | 135,8 |

Таблиця VI

Рівень значущості  $\alpha = 0,05$

**Верхні  $\alpha$ -границі розподілів Фішера  $F(n, m)$**

| m  | n     |       |       |       |       |       |       |       |          |
|----|-------|-------|-------|-------|-------|-------|-------|-------|----------|
|    | 1     | 2     | 3     | 4     | 5     | 6     | 12    | 24    | $\infty$ |
| 1  | 164,4 | 199,5 | 215,7 | 224,6 | 230,2 | 234,0 | 244,9 | 249,0 | 254,3    |
| 2  | 18,5  | 19,2  | 19,2  | 19,3  | 19,3  | 19,3  | 19,4  | 19,5  | 19,5     |
| 3  | 10,1  | 9,6   | 9,3   | 9,1   | 9,0   | 8,9   | 8,7   | 8,6   | 8,5      |
| 4  | 7,7   | 6,9   | 6,6   | 6,4   | 6,3   | 6,2   | 5,9   | 5,8   | 5,6      |
| 5  | 6,6   | 5,8   | 5,4   | 5,2   | 5,1   | 5,0   | 4,7   | 4,5   | 4,4      |
| 6  | 6,0   | 5,1   | 4,8   | 4,5   | 4,4   | 4,3   | 4,0   | 3,8   | 3,7      |
| 7  | 5,6   | 4,7   | 4,4   | 4,1   | 4,0   | 3,9   | 3,6   | 3,4   | 3,2      |
| 8  | 5,3   | 4,5   | 4,1   | 3,8   | 3,7   | 3,6   | 3,3   | 3,1   | 2,9      |
| 9  | 5,1   | 4,3   | 3,9   | 3,6   | 3,5   | 3,4   | 3,1   | 2,9   | 2,7      |
| 10 | 5,0   | 4,1   | 3,7   | 3,5   | 3,3   | 3,2   | 2,9   | 2,7   | 2,5      |
| 11 | 4,8   | 4,0   | 3,6   | 3,4   | 3,2   | 3,1   | 2,8   | 2,6   | 2,4      |
| 12 | 4,8   | 3,9   | 3,5   | 3,3   | 3,1   | 3,0   | 2,7   | 2,5   | 2,3      |
| 13 | 4,7   | 3,8   | 3,4   | 3,2   | 3,0   | 2,9   | 2,6   | 2,4   | 2,2      |
| 14 | 4,6   | 3,7   | 3,3   | 3,1   | 3,0   | 2,9   | 2,5   | 2,3   | 2,1      |
| 15 | 4,5   | 3,7   | 3,3   | 3,1   | 2,9   | 2,8   | 2,5   | 2,3   | 2,1      |
| 16 | 4,5   | 3,6   | 3,2   | 3,0   | 2,9   | 2,7   | 2,4   | 2,2   | 2,0      |
| 17 | 4,5   | 3,6   | 3,2   | 3,0   | 2,8   | 2,7   | 2,4   | 2,2   | 2,0      |
| 18 | 4,4   | 3,6   | 3,2   | 2,9   | 2,8   | 2,7   | 2,3   | 2,1   | 1,9      |
| 19 | 4,4   | 3,5   | 3,1   | 2,9   | 2,7   | 2,6   | 2,3   | 2,1   | 1,8      |
| 20 | 4,4   | 3,5   | 3,1   | 2,9   | 2,7   | 2,6   | 2,3   | 2,1   | 1,8      |
| 22 | 4,3   | 3,4   | 3,1   | 2,8   | 2,7   | 2,6   | 2,2   | 2,0   | 1,8      |
| 24 | 4,3   | 3,4   | 3,0   | 2,8   | 2,6   | 2,5   | 2,2   | 2,0   | 1,7      |

Продовження таблиці VI  
Рівень значущості  $\alpha = 0,05$

| m        | n   |     |     |     |     |     |     |     |          |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|----------|
|          | 1   | 2   | 3   | 4   | 5   | 6   | 12  | 24  | $\infty$ |
| 26       | 4,3 | 3,4 | 3,0 | 2,7 | 2,6 | 2,4 | 2,1 | 1,9 | 1,7      |
| 28       | 4,2 | 3,3 | 2,9 | 2,7 | 2,6 | 2,4 | 2,1 | 1,9 | 1,6      |
| 30       | 4,2 | 3,3 | 2,9 | 2,7 | 2,5 | 2,4 | 2,1 | 1,9 | 1,6      |
| 40       | 4,1 | 3,2 | 2,9 | 2,6 | 2,5 | 2,3 | 2,0 | 1,8 | 1,5      |
| 60       | 4,0 | 3,2 | 2,8 | 2,5 | 2,4 | 2,3 | 1,9 | 1,7 | 1,4      |
| 120      | 3,9 | 3,1 | 2,7 | 2,5 | 2,3 | 2,2 | 1,8 | 1,6 | 1,3      |
| $\infty$ | 3,8 | 3,0 | 2,6 | 2,4 | 2,2 | 2,1 | 1,8 | 1,5 | 1,0      |

Рівень значущості  $\alpha = 0,01$

| m  | n     |      |      |      |      |      |      |      |          |
|----|-------|------|------|------|------|------|------|------|----------|
|    | 1     | 2    | 3    | 4    | 5    | 6    | 12   | 24   | $\infty$ |
| 2  | 98,5  | 99,0 | 99,2 | 99,3 | 99,3 | 99,4 | 99,4 | 99,5 | 99,5     |
| 3  | 34,1  | 30,8 | 29,5 | 28,7 | 28,2 | 27,9 | 27,1 | 26,6 | 26,1     |
| 4  | 21,2  | 18,0 | 16,7 | 16,0 | 15,5 | 15,2 | 14,4 | 13,9 | 13,5     |
| 5  | 16,3  | 13,3 | 12,1 | 11,4 | 11,0 | 10,7 | 9,9  | 9,5  | 9,0      |
| 6  | 13,70 | 10,9 | 9,8  | 9,2  | 8,8  | 8,5  | 7,7  | 7,3  | 6,9      |
| 7  | 12,3  | 9,6  | 8,5  | 7,9  | 7,5  | 7,2  | 6,5  | 6,1  | 5,7      |
| 8  | 11,3  | 8,7  | 7,6  | 7,0  | 6,6  | 6,4  | 5,7  | 5,3  | 4,9      |
| 9  | 10,6  | 8,0  | 7,0  | 6,4  | 6,1  | 5,8  | 5,1  | 4,7  | 4,3      |
| 10 | 10,0  | 7,6  | 6,6  | 6,0  | 5,6  | 5,4  | 4,7  | 4,3  | 3,9      |
| 11 | 9,7   | 7,2  | 6,2  | 5,7  | 5,3  | 5,1  | 4,4  | 4,0  | 3,6      |
| 12 | 9,3   | 6,9  | 6,9  | 5,4  | 5,1  | 4,8  | 4,2  | 3,8  | 3,4      |
| 13 | 9,1   | 6,7  | 5,7  | 5,2  | 4,9  | 4,6  | 4,0  | 3,6  | 3,2      |
| 14 | 8,9   | 6,5  | 5,6  | 5,0  | 4,7  | 4,5  | 3,8  | 3,4  | 3,0      |
| 15 | 8,7   | 6,4  | 5,4  | 4,9  | 4,6  | 4,3  | 3,7  | 3,3  | 2,9      |

Продовження таблиці VI  
Рівень значущості  $\alpha = 0,01$

| m        | n   |     |     |     |     |     |     |     |          |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|----------|
|          | 1   | 2   | 3   | 4   | 5   | 6   | 12  | 24  | $\infty$ |
| 16       | 8,5 | 6,2 | 5,3 | 4,8 | 4,4 | 4,2 | 3,6 | 3,2 | 2,8      |
| 17       | 8,4 | 6,1 | 5,2 | 4,7 | 4,3 | 4,1 | 3,5 | 3,1 | 2,7      |
| 18       | 8,3 | 6,0 | 5,1 | 4,6 | 4,3 | 4,0 | 3,4 | 3,0 | 2,6      |
| 19       | 8,2 | 5,9 | 5,0 | 4,5 | 4,2 | 3,9 | 3,3 | 3,0 | 2,4      |
| 20       | 8,1 | 5,9 | 4,9 | 4,4 | 4,1 | 3,9 | 3,2 | 2,9 | 2,4      |
| 22       | 7,9 | 5,7 | 4,8 | 4,3 | 4,0 | 3,8 | 3,1 | 2,9 | 2,3      |
| 24       | 7,8 | 5,6 | 4,7 | 4,2 | 3,9 | 3,7 | 3,0 | 2,8 | 2,2      |
| 26       | 7,7 | 5,5 | 4,6 | 4,1 | 3,8 | 3,6 | 3,0 | 2,7 | 2,1      |
| 28       | 7,6 | 5,5 | 4,6 | 4,1 | 3,8 | 3,5 | 2,9 | 2,6 | 2,1      |
| 30       | 7,6 | 5,4 | 4,5 | 4,0 | 3,7 | 3,5 | 2,8 | 2,5 | 2,0      |
| 40       | 7,3 | 5,3 | 4,3 | 3,8 | 3,5 | 3,3 | 2,7 | 2,5 | 1,8      |
| 60       | 7,1 | 5,0 | 4,1 | 3,7 | 3,3 | 3,1 | 2,5 | 2,3 | 1,8      |
| 120      | 6,9 | 4,8 | 4,0 | 3,5 | 3,2 | 3,0 | 2,3 | 2,1 | 1,4      |
| $\infty$ | 6,6 | 4,6 | 3,8 | 3,3 | 3,0 | 2,8 | 2,2 | 2,0 | 1,0      |

# Література

## **Підручники, посібники, статті та монографії**

1. Боровков А. А. Математическая статистика. М: Наука, 1984. 472 с.
2. Карташов М. В. Ймовірність, процеси, статистика: посібник. К: ВПЦ Київський університет, 2008. 494 с.
3. Основи теорії ймовірностей та математичної статистики: навч. посібник./ Бабак В.П. та ін. К., 2003. 432 с.
4. Приймак В. І., Голубник О. Р.. Теорія ймовірностей та математична статистика: підручник. Л: ЛНУ ім. Франка, 2011. 556 с.
5. Сеньо П. С. Теорія ймовірностей та математична статистика: підручник. К.: Знання, 2007. 556 с.
6. Турчин В. М. Математична статистика: навч. посібник. К.: ВЦ "Академія 1999. 240 с.
7. Турчин В. М. Теорія ймовірностей та математична статистика. Основні поняття, приклади і задачі: підручник. Д.: ІМА-прес, 2014. 556 с.
8. Fisher R. A. On the mathematical foundations of theoretical statistics. RTRS, 1921. - P. 155.

**Збірники та посібники з розв'язування задач**

9. В. В. Голомозий, М. В. Карташов, К. В. Ральченко. Збірник задач з теорії ймовірностей та математичної статистики: навч. посіб. К: ВПЦ Київський університет, 2019. 367 с.
10. Кармелюк Г. І. Теорія ймовірностей та математична статистика. Посібник з розв'язування задач: Навч. посібник. – К: Центр учбової літератури, 2007 -576 с.
11. David F. N., Pearson E. S. Elementary Statistical Exercises. Cambridge: University Press, 1961. 241 p.

*Навчальне видання*

**Коваленко Лариса Григорівна**

**МАТЕМАТИЧНА  
СТАТИСТИКА**

*Частина 1*

**НАВЧАЛЬНИЙ ПОСІБНИК**

для здобувачів 3-го курсу  
першого (бакалаврського) рівня вищої освіти  
спеціальності 111 «Математика»

**Електронне видання мережевого використання**

*В авторській редакції*

Затвердж. авт. 30.09.2024. Шрифт Times New Roman.  
Системні вимоги: операційна система сумісна з програмним забезпеченням  
для читання файлів формату PDF.  
Обсяг 0,6 МБ. Зам. № 2842.

Видавець і виготовлювач  
Одеський національний університет імені І. І. Мечникова  
Свідоцтво суб'єкта видавничої справи ДК № 4215 від 22.11.2011 р.  
вул. Університетська, 12, м. Одеса, 65082, Україна  
Тел.: (048) 723 28 39, e-mail: druk@onu.edu.ua