МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Івано-Франківський національний технічний
університет нафти і газу
Кафедра вищої математики

М.М.Осипчук

ТЕОРІЯ ЙМОВІРНОСТЕЙ, ВИПАДКОВІ ПРОЦЕСИ ТА МАТЕМАТИЧНА СТАТИСТИКА

КОНСПЕКТ ЛЕКЦІЙ

ЧАСТИНА 2

МАТЕМАТИЧНА СТАТИСТИКА

м. Івано-Франківськ
2003

Осипчук М.М. Теорія ймовірностей, випадкові процеси та математична статистика Ч. 2 Математична статистика: Конспект лекцій. — Івано-Франківськ: Факел, 2003. - 85 с.

Конспект лекцій призначений для студентів всіх спеціальностей денної та заочної форм навчання, які вивчають дисципліни "Теорія ймовірностей та математична статистика"і "Вища математика"(розділ "Теорія ймовірностей та математична статистика").

Рецензент: канд. фіз.-мат. наук, доцент Я.І.Савчук

Зміст

Розділ 1

Елементарні статистичні методи

1.1 Вибірка, її характеристики та методи одержання

1.1.1 Статистика, статистичні дані

Поняття *статистика* для більшості людей пов'язане з наборами різного роду інформації, з методами її аналізу та прийняття рішень. В залежності від джерел інформації часто виділяють економічну, правову, промислову, транспортну та інші статистики. Основою для всіх різновидностей статистики є *математична статистика.*

Відомий англійський статистик Р. Фішер відзначав, що "статистика як наука, є одним із розділів прикладної математики, і її можна розглядати як математику застосовувану при обробці результатів масового спостереження. У статистиці, як і в інших математичних науках та сама формула однаковою мірою стосується найрізноманітніших матеріальних об'єктів".

Загальна "теорія статистики є галуззю прикладної математики; вона має свої корені в тій галузі чистої математики, що відома як теорія ймовірностей, і фактично весь будинок математичної теорії [статистики] в широкому розумінні включає теорію ймовірностей поряд з іншими аксіомами, що посилюють аксіоми теорії ймовірностей", — пише американський статистик Аманфарлан Муд.

Математична статистика має справу з, так званою, числовою (статистичною) інформацією. Проте не завжди зібрана інформація виражається числами. Тому доволі часто виникає потреба в перетворенні якісної статистичної інформації в числову. Для цього використовують процедуру *виміру* — приписування чисел об'єктам і їх властивостям відповідно до деяких правил. Ці правила встановлюють відповідність між деякими властивостями чисел і деякими властивостями об'єктів. Для виміру об'єктів та їх властивостей часто вживають такі шкали:

- шкала найменувань;
- шкала порядку;
- шкала інтервалів;
- шкала відношень.

Найбільш простий і, можливо, найбільш поширений вимір у шкалі найменувань, що є групуванням об'єктів на класи на підставі наявності в них спільної ознаки або властивості. Класам даються найменування і присвоюються числові значення (номери). Номер класу нічого не говорить про властивості об'єктів за винятком того, що вони відрізняються. Вимір у шкалі порядку — це групування об'єктів на класи на підставі наявності в них спільної ознаки та приписування класам чисел в порядку зростання чи спадання цієї ознаки. Такий вид виміру ще називають ранжуванням. Слід зауважити, що в шкалі порядку рівні різниці чисел не обов'язково відповідають рівним різницям кількості ознаки. Таку відповідність забезпечує вимір у шкалі інтервалів. У цій шкалі існують одиниці виміру. Особливістю шкал інтервалів є довільність у виборі нульвої точки на шкалі, що зовсім не вказує на відсутність ознаки (літочислення, температура, час). Вимір у шкалі відношень — це таке присвоювання чисел об'єктам, при якому одиниця виміру дозволяє впорядкувати об'єкти так, що можна визначити, на скільки і у скільки в одного об'єкта ознаки більше, ніж в іншого (вага, відстань, площа, об'єм).

Вибір методу вимірювання залежить від конкретної задачі та від вподобань дослідника. Надалі будемо вважати, що вся статистична інформація, з якою ми будемо мати справу, є числовою.

1.1.2 Вибірка та способи її описання

Враховуючи те, що статистична інформація одержується в результаті спостережень, вимірювань та інших дій, на результат яких впливає багато мало доступних для врахування факторів, можна вважати кожну одиницю статистичної інформації випадковою величиною. Оскільки досліджувана статистична інформація є набором даних одержаних в результаті проведення однотипних дій, то матимемо справу з набором однаково розподілених незалежних в сукупності випадкових величин.

Отже, нехай випадкова величина ξ спостерігається в деякому стохастичному експерименті. Повторимо цей експеримент в незмінних умовах n раз. З, таким чином організованим, дослідженням пов'язаний n-вимірний випадковий вектор

$$(\xi_1, \xi_2, \ldots, \xi_n). \tag{1.1}$$

Очевидно, що випадкові величини ξ_i, $i = 1, 2, \ldots, n$,— незалежні в сукупності та розподілені за тим же законом, що і випадкова величина ξ.

Розподіл випадкової величини ξ називається *розподілом генеральної сукупності*, а випадковий вектор (1.1) — *вибірковим вектором.* Набір чисел, що одержуються на практиці при повторенні експерименту є реалізацією вибіркового вектора і називається *вибіркою* x_1, x_2,..., x_n об'єму n. Вибірку x_1, x_2,..., x_n при необхідності можна розглядати як точку *вибіркового простору*, тобто множини, на якій задано розподіл вибіркового вектора. Аналогічно визначається вибірка і у випадку, коли випадковий експеримент пов'язаний з кількома випадковими величинами.

Результати спостережень x_1, x_2,..., x_n генеральної сукупності випадкової величини ξ, записані в порядку їх реєстрації, звичайно незручні для подальшого аналізу. Завданням статистичного описання вибірки є одержання такого її задання, яке дозволяло би виявити характерні особливості сукупності даних.

Варіаційним рядом вибірки x_1, x_2,..., x_n називається її запис, при якому елементи впорядковані по зростанню, тобто, вибірка записується у вигляді $x^{(1)}$, $x^{(2)}$,..., $x^{(n)}$, де $x^{(1)} \leq x^{(2)} \leq \ldots \leq x^{(n)}$. Різниця між максимальним та мінімальним елементами вибірки $x^{(n)} - x^{(1)} = w$ називається *розмахом варіації* вибірки.

Нехай вибірка x_1, x_2,..., x_n містить k різних чисел z_1, z_2,..., z_k, причому z_i у вибірці зустрічається n_i раз ($i = 1, 2, \ldots, k$). Число n_i називається *частотою* елемента z_i, а число $\nu_i = \frac{n_i}{n}$ — *відносною частотою* цього елемента. Очевидно, що $\sum\limits_{i=1}^{k} n_i = n$, а $\sum\limits_{i=1}^{k} \nu_i = 1$. *Статистичним рядом* вибірки називається набір пар (z_i, n_i), $i = 1, 2, \ldots, k$. Статистичний ряд зручно записувати у вигляді таблиці, в першому рядку якої записують z_i, а в другому n_i.

При великому об'ємі вибірки її елементи об'єднують в групи, задаючи результати експерименту у вигляді *згрупованого статистичного ряду*. Для цього інтервал, що містить всі елементи вибірки, розбивають на k інтервалів, що не перетинаються. Для спрощення надалі будемо вважати, що такі інтервали мають однакову довжину. Після того як інтервали вибрано, визначають *частоти* — кількості n_i елементів вибірки, що попали в i-тий інтервал. Значення, що характеризує інтервал є його серединою. Слід зауважити, що при такому підході вноситься деяка додаткова похибка в подальші обчислення.

Розподілом вибірки називається розподіл дискретної випадкової величини, що приймає значення x_1, x_2,..., x_n з ймовірностями $\frac{1}{n}$. Відповідна функція розподілу називається *емпіричною функцією розподілу* та позначається $F_n(x)$. Очевидно, що для статистичного ряду (z_i, n_i)

$$F_n(x) = \frac{1}{n} \sum_{z_i < x} n_i.$$

Аналогічно задається емпірична функція розподілу для згрупованої вибірки.

Теорема (Глівенко) Нехай $F_n(x)$— емпірична функція розподілу, побудована за вибіркою об'єму n з генеральної сукупності розподілу $F(x)$. Тоді для будь-якого $x \in (-\infty, +\infty)$ і кожного $\varepsilon > 0$

$$\lim_{n \to \infty} \mathbf{P}(|F_n(x) - F(x)| < \varepsilon) = 1.$$

Запис цієї теореми стає зрозумілим, якщо зауважити, що значення емпіричної функції розподілу залежать від реалізації вибіркового вектора і тому є випадковими. Отже, при кожному x $F_n(x)$ збігається за ймовірністю до $F(x)$ і тому при великому об'ємі вибірки може бути наближеним значенням функції розподілу генеральної сукупності в кожній точці x.

Для наглядного описання вибірки часто використовують гістограму та полігон частот.

Гістограмою частот згрупованої вибірки називається кусочно стала функція, яка на інтервалах групування приймає значення $\frac{n_i}{h}$, де h— довжина інтервалів групування. Площа ступінчатої фігури під гістограмою дорівнює об'єму вибірки n. Аналогічно визначається *гістограма відносних частот.* При великих об'ємах вибірки та малих інтервалах групування гістограма відносних частот є статистичним аналогом щільності розподілу генеральної сукупності.

Полігоном частот називається ламана з вершинами в точках $\left(z_i, \frac{n_i}{h}\right)$, $i = 1, 2, \ldots, k$, а *полігоном відносних частот* називається ламана з вершинами в точках $\left(z_i, \frac{\nu_i}{h}\right)$, $i = 1, 2, \ldots, k$. Якщо щільність розподілу генеральної сукупності є досить гладкою функцією, то полігон відносних частот є кращим наближенням щільності, ніж гістограма.

Приклад. Для даної вибірки побудувати варіаційний ряд, знайти розмах варіації, побудувати згруповану вибірку та відповідні емпіричну функцію розподілу, полігон та гістограму частот, використавши 7 інтервалів групування. Вибірка:

20,3, 15,4, 17,2, 19,2, 23,3, 18,1, 21,9, 15,3, 16,8, 13,2, 20,4, 16,5, 19,7, 20,5, 14,3, 20,1, 16,8, 14,7, 20,8, 19,5, 15,3, 19,3, 17,8, 16,2, 15,7, 22,8, 21,9, 12,5, 10,1, 21,1, 18,3, 14,7, 14,5, Таблиця 1.1

Номер інтервала i	Межі інтервала	Середини інтервалів z_i	Частота n_i	Відносна частота ν_i
1	10–12	11	2	0,0364
2	12–14	13	4	0,0727
3	14–16	15	8	0,1455
4	16–18	17	12	0,2182
5	18–20	19	16	0,2909
6	20–22	21	10	0,1818
7	22–24	23	3	0,0545

18,1, 18,4, 13,9, 19,1, 18,5, 20,2, 23,8, 16,7, 20,4, 19,5, 17,2, 19,6, 17,8, 21,3, 17,5, 19,4, 17,8, 13,5, 17,8, 11,8, 18,6, 19,1.

Розв'язок. Варіаційний ряд:

10,1, 11,8, 12,5, 13,2, 13,5, 13,9, 14,3, 14,5, 14,7, 14,7, 15,3, 15,3, 15,4, 15,7, 16,2, 16,5, 16,7, 16,8, 16,8, 17,2, 17,2, 17,5, 17,8, 17,8, 17,8, 17,8, 18,1, 18,1, 18,3, 18,4, 18,5, 18,6, 19,1, 19,1, 19,2, 19,3, 19,4, 19,5, 19,5, 19,6, 19,7, 20,1, 20,2, 20,3, 20,4, 20,4, 20,5, 20,8, 21,1, 21,3, 21,9, 21,9, 22,8, 23,5, 23,8.

Розмах варіації $w = 23,8 - 10,1 = 13,7$.

Довжину інтервалу групування візьмемо рівною $h = 2$, бо $\frac{13,7}{7} \approx 2$. Виберемо 7 інтервалів довжиною 2, перший з яких $[10; 12)$, а останній $[22; 24]$ та знайдемо їх середини, відповідні частоти, відносні частоти. Результати занесемо в таблицю 1.1.

За результатами групування будуємо гістограму частот (рис. 1.1). З'єднуючи відрізками середини верхніх основ прямокутників, з яких складається гістограма, одержимо

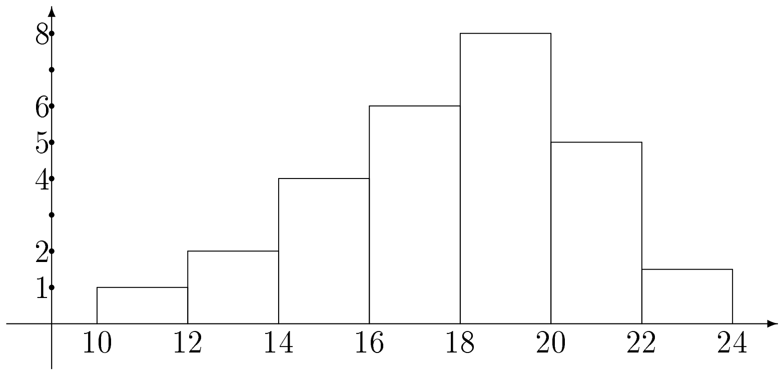


Рисунок 1.1 - Гістограма частот

відповідний полігон частот (рис. 1.2).

Емпірична функція розподілу є східчастою функцією, яка має скачки величини ν_i в точках z_i (при $x \leq 11$ $F_n(x) = 0$, а при $x > 23$ $F_n(x) = 1$). Графік функції $F_n(x)$ зображено на рисунку 1.3.

Вправи

1. Дані значення залишків (в см) від кожного з 20 сувоїв, одержаних при розкрою матеріалу:

№рулона	1	2	3	4	5	6	7	8	9	10
Залишок	14.8	12	2.9	2.4	11.7	8.1	10.7	2.7	9.7	14.3
№рулона	11	12	13	14	15	16	17	18	19	20
Залишок	2.3	2.2	11.3	15.7	6.3	0.8	2.6	4.8	8.7	4.5

Побудувати емпіричну функцію розподілу та гістограму частот. Знайти вибіркову середню залишків та їх вибіркову дисперсію.

2. Визначити незсунені оцінки математичного сподівання та дисперсії максимальної швидкості літака за даними 20 вимірів

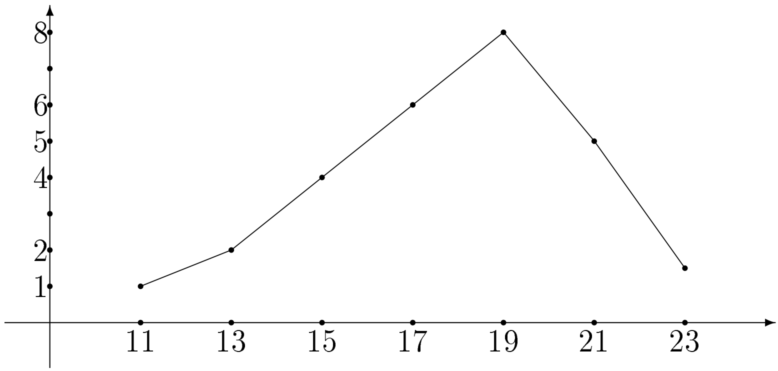


Рисунок 1.2 - Полігон частот

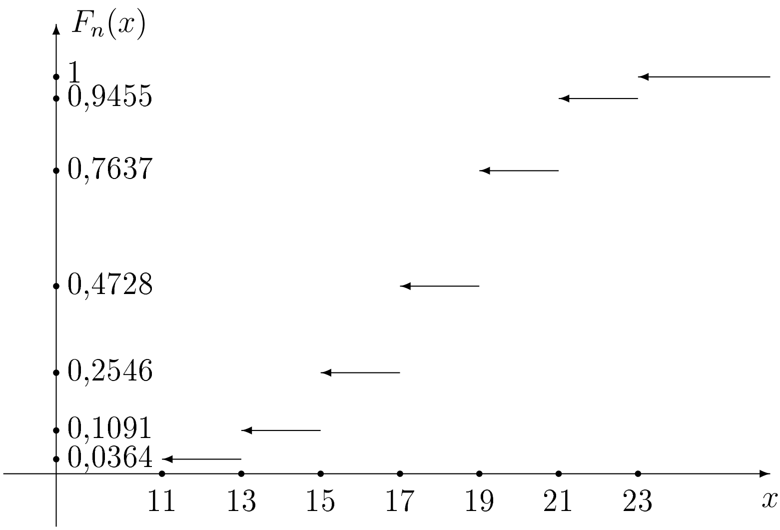


Рисунок 1.3 - Емпірична функція розподілу

його максимальної швидкості (м/с): 485.2, 512.0, 497.1, 502.4, 488.3, 491.9, 489.3, 502.7, 509.2, 514.3, 502.3, 497.8, 511.3, 515.7, 506.3, 504.5, 499.2, 497.4, 485.2, 507.8 .

3. Залишок ξ палива в баках літака після кожного з п'ятнадцяти рейсів за заданим маршрутом виражається числами (кг): 340.2, 316.8, 325.3, 329.2, 351.8, 348.7, 330.1, 345.6, 352.3, 331.7, 318.4, 332.2, 341.4, 318.2, 341.4 . Визначити незсунені оцінки $\mathbf{M}\,\xi$ та $\mathbf{D}\,\xi$.

4. За даною вибіркою з нормально розподіленої генеральної сукупності знайти:

- об'єм вибірки n;
- розмах варіації;
- медіану;
- точкову оцінку математичного сподівання (вибіркове середнє);
- точкову оцінку дисперсії (вибіркову дисперсію);
- побудувати гістограму, полігон частот та емпіричну функцію розподілу, розбивши інтервал, який містить варіаційний ряд, на $\sqrt{n}$ рівних частин.

$3,7,\ 2,9,\ 3,9,\ 6,1,\ 4,5,\ -0,7,\ 3,0,\ 1,2,\ 0,9,\ 2,4,\ 4,0,\ 3,8,\ 0,2,\ 0,2,$ $1,5,\ 2,7,\ 4,0,\ 2,6,\ 6,0,\ -0,8,\ -0,2,\ -2,2,\ 2,0,\ 6,1,\ 3,1,\ 3,7,\ 2,0,$ $4,4,\ 5,5,\ 4,3,\ 2,6,\ -0,7,\ 6,5,\ 3,7,\ 0,3,\ -0,3.$

Розділ 2

Теорія оцінювання параметрів розподілів

2.1 Точкові оцінки, їх властивості та методи побудови

Нехай x_1, x_2,..., x_n — вибірка з генеральної сукупності деякої випадкової величини ξ.

2.1.1 Загальна теорія оцінювання параметрів.

Нехай функція розподілу (розподіл) залежить від деякого параметра θ. Оцінка (наближене значення) $\hat{\theta}$ цього параметра очевидно залежить від вибіркових значень. Тому задачу оцінювання параметрів розподілу потрібно розглядати як задачу відшукання такої функції $h : R^n \to R$, що $\hat{\theta} = h(x_1, x_2, \ldots, x_n)$. Такі функції будемо називати статистиками, а їх значення — оцінками.

Оскільки вибірка є реалізацією деякого випадкового вектора, то оцінка параметра розподілу побудована за вибіркою є випадковою величиною.

Означення. Оцінку $\hat{\theta}$ будемо називати *незміщеною* оцінкою параметра θ, якщо $\mathbf{M}\,\hat{\theta} = \theta$.

Якщо розуміти математичне сподівання як деяке середнє значення випадкової величини, то незміщеність оцінки означає, що її середнє значення співпадає з дійсним значенням параметра, або, що те ж саме різниця $\hat{\theta} - \theta$ в середньому дорівнює нулю. Незміщена оцінка $\hat{\theta}$ найменше відхиляється від значення θ серед всіх оцінок з однією і тією ж дисперсією. Дійсно для будь-якої оцінки $\hat{\theta}$

$$\mathbf{M}\,|\hat{\theta} - \theta|^2 = \mathbf{M}\,|(\hat{\theta} - \mathbf{M}\,\hat{\theta}) + (\mathbf{M}\,\hat{\theta} - \theta)|^2 =$$
$$= \mathbf{M}(\hat{\theta} - \mathbf{M}\,\hat{\theta})^2 + 2\,\mathbf{M}(\hat{\theta} - \mathbf{M}\,\hat{\theta})(\mathbf{M}\,\hat{\theta} - \theta) + \mathbf{M}(\mathbf{M}\,\hat{\theta} - \theta)^2 =$$
$$= \mathbf{D}\,\hat{\theta} + (\mathbf{M}\,\hat{\theta} - \theta)^2.$$

Розглядаючи оцінку параметра розподілу як функцію від вибірки, зауважимо, що вона залежить і від об'єму вибірки n. Тому кожного разу, коли можна одержувати нескінченно великі за об'ємом вибірки, ми можемо розглядати цілу послідовність оцінок $\hat{\theta}_n = h(x_1, x_2, \ldots, x_n)$. Хоча часто про цю послідовність будемо говорити як про одну оцінку.

Означення. Оцінку $\hat{\theta} = \hat{\theta}_n$ будемо називати *асимптотично незміщеною* оцінкою параметра θ, якщо $\mathbf{M}\,\hat{\theta} \to \theta$ при $n \to \infty$.

Означення. Оцінка $\hat{\theta} = \hat{\theta}_n$ називається *консистентною (слушною)* оцінкою параметра θ, якщо для кожного $\varepsilon > 0$

$$\mathbf{P}(|\hat{\theta} - \theta| > \varepsilon) \to 0$$

при $n \to \infty$.

Звідси можна зробити висновок, що з ростом об'єму вибірки конзистентна оцінка параметра розподілу прямує за ймовірністю до дійсного значення цього параметра.

Вище було зауважено, що незміщена оцінка серед оцінок з тією ж дисперсією найменше відхиляється від значення параметра. Серед незміщених оцінок найкращою буде та, в якої найменша дисперсія.

Означення. Незміщену оцінку θ^* будемо називати *ефективною*, якщо

$$\mathbf{D}\,\theta^* = \inf_{\hat{\theta}:\mathbf{M}\,\hat{\theta}=\theta} \mathbf{D}\,\hat{\theta}.$$

Припустимо, що досліджувана випадкова величина має щільність розподілу $f(x;\theta)$. Розглянемо незміщену оцінку параметра θ: $\hat{\theta} = h(x_1, x_2, \ldots, x_n)$. При деяких умовах регулярності функцій f і h (див.[1]) має місце *нерівність Крамера-Рао*

$$\mathbf{D}\,\hat{\theta} \geq \frac{1}{I(\theta)}, \tag{2.1}$$

де $I(\theta) = n\,\mathbf{M}\left(\frac{\partial \ln f(\xi;\theta)}{\partial\theta}\right)^2$ — *кількість інформації за Фішером*, що міститься у вибірці.

Рівність в (2.1) досягається тоді і тільки тоді, коли

$$\sum_{k=1}^{n} \frac{\partial \ln f(x_k;\theta)}{\partial\theta} = \lambda(h(x_1, x_2, \ldots, x_n) - \theta). \tag{2.2}$$

Тут λ не залежить від $x_1, x_2, \ldots, x_n$.

У випадку дискретної випадкової величини ξ

$$I(\theta) = n\,\mathbf{M}\left(\frac{\partial \ln p(\xi;\theta)}{\partial\theta}\right)^2,$$

[1] Турчин В.М. Математична статистика. Навч. посібник. – К.: Академія, 1999. – 240 с.

де $p(x,\theta) = \mathbf{P}(\xi = x)$. І умова досягнення рівності в (2.1) виглядає так:

$$\sum_{k=1}^{n} \frac{\partial \ln p(x_k;\theta)}{\partial\theta} = \lambda(h(x_1, x_2, \ldots, x_n) - \theta).$$

Приклад. Нехай вибірка x_1, x_2,..., x_n одержана з генеральної сукупності нормально розподіленої з параметрами a та σ^2 випадкової величини. Перевірити незміщеність, консистентність та ефективність оцінок

$$\bar{x} = \hat{a} = \frac{1}{n}\sum_{k=1}^{n} x_k, \quad \hat{\sigma^2} = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2.$$

Розв'язок. Нехай $y_i = x_i - a$, $i = 1, 2, \ldots, n$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. Оскільки вибіркові значення є реалізаціями незалежних, однаково $N(a,\sigma^2)$-розподілених випадкових величин, то

$$\mathbf{M}\, y_i = 0, \quad \mathbf{D}\, y_i = \mathbf{M}\, y_i^2 = \mathbf{D}\, x_i = \sigma^2 \quad \mathbf{M}\, y_i y_j = 0 \quad (i \neq j)$$

$$\text{і} \quad \mathbf{M}\,\hat{a} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{M}\, x_k = \frac{1}{n} na = a;$$

$$\begin{aligned}
\mathbf{M}\,\hat{\sigma^2} &= \frac{1}{n}\sum_{k=1}^{n} \mathbf{M}(x_k - \bar{x})^2 = \frac{1}{n}\sum_{k=1}^{n} \mathbf{M}(y_k - \bar{y})^2 = \\
&= \frac{1}{n}\mathbf{M}\sum_{k=1}^{n} y_k^2 - \frac{2}{n}\mathbf{M}\sum_{k=1}^{n} y_k\bar{y} + \frac{1}{n}\mathbf{M}\, n(\bar{y})^2 = \\
&= \frac{1}{n}\sum_{k=1}^{n}\mathbf{M}\, y_k^2 - 2\,\mathbf{M}(\bar{y})^2 + \mathbf{M}(\bar{y})^2 = \\
&= \frac{1}{n} n\sigma^2 - \frac{1}{n^2}\mathbf{M}\left(\sum_{k=1}^{n} y_k^2 + 2\sum_{i<j} y_i y_j\right) =
\end{aligned}$$

$$= \sigma^2 - \frac{1}{n^2}\sum_{k=1}^{n} \mathbf{M}\, y_k^2 = \sigma^2 - \frac{1}{n^2} n\sigma^2 =$$

$$= \frac{n-1}{n}\sigma^2.$$

Звідси $\bar{x}$ є незміщеною оцінкою параметра a, а $\hat{\sigma^2}$ — зміщена оцінка[2] параметра σ^2 (зміщення становить $\frac{\sigma^2}{n}$).

Консистентність оцінки $\bar{x}$ є суттю закону великих чисел (теорема Чебишова). Дійсно, за законом великих чисел, для незалежних однаково розподілених випадкових величин x_1, x_2,...,x_n

$$\frac{1}{n}\sum_{i=1}^{n} x_i \overset{\mathbf{P}}{\to} a \quad , n \to \infty.$$

Оцінка $\hat{\sigma^2}$ є також консистентною, як і оцінка $s^2 = \frac{n}{n-1}\hat{\sigma^2}$. Доведення цього факту ми пропустимо.

Перевіримо ефективність оцінки $\bar{x}$. Для цього перевіримо виконання рівності (2.2). Щільність $N(a;\sigma^2)$-розподілу має вигляд

$$f(x;a,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{\frac{(x-a)^2}{2\sigma^2}\right\}.$$

Тоді $\ln f(x;a,\sigma) = -\ln\sqrt{2\pi} - \ln\sigma - \frac{(x-a)^2}{2\sigma^2}$ і

$$\sum_{k=1}^{n} \frac{\partial \ln f(x_k;a,\sigma)}{\partial a} = \sum_{k=1}^{n} \frac{x-a}{\sigma^2} = \frac{n}{\sigma^2}(\bar{x}-a) \qquad (2.3)$$

$$\sum_{k=1}^{n} \frac{\partial \ln f(x_k;a,\sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{k=1}^{n}(x_k-a)^2 =$$

$$= \frac{n}{\sigma^3}\left(\frac{1}{n}\sum_{k=1}^{n}(x_k-a)^2 - \sigma^2\right). \qquad (2.4)$$

---

[2] Незміщеною оцінкою σ^2 є статистика $s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k-\bar{x})^2$

Рівність (2.3) означає, що оцінка $\bar{x}$ є ефективною, а з рівності (2.4) випливає, що оцінка $\frac{1}{n}\sum_{k=1}^{n}(x_k-a)^2$ є ефективною оцінкою параметра σ^2 при відомому значенні параметра a. Можна довести, що такою є і оцінка s^2.

2.1.2 Метод максимальної правдоподібності

Розглянемо найбільш застосовуваний метод побудови точкових оцінок параметрів розподілу.

Це *метод максимальної правдоподібності*, запропонований Р. Фішером. Він полягає в тому, що для отримання оцінки невідомого параметру θ розподілу генеральної сукупності шукають таке значення $\hat{\theta}$, при якому ймовірність реалізації випадкового вектора $(\xi_1;\xi_2;\ldots;\xi_n)$ була б максимальною.

Нехай отримано вибірку x_1, x_2,..., x_n з генеральної сукупності неперервної випадкової величини з щільністю розподілу $f(x;\theta)$, що залежить від параметра θ.

Означення. *Функцією правдоподібності* вибірки x_1, x_2, ..., x_n будемо називати функцію

$$L(\theta)=f(x_1;\theta)\cdot f(x_2;\theta)\cdot\ldots\cdot f(x_n;\theta).$$

Значення $\hat{\theta}(x_1,\ldots,x_n)$ параметра θ, при якому досягається максимум функції правдоподібності, називається *оцінкою максимальної правдоподібності* (МП-оцінкою) параметра θ.

Якщо досліджувана випадкова величина дискретна з розподілом $(x_i;p_i(\theta))$, а вибірка задана статистичним рядом $(x_i;n_i)$, то функцією правдоподібності є функція $(i=1,2,\ldots,k)$

$$L(\theta)=p_1^{n_1}(\theta)\cdot p_2^{n_2}(\theta)\cdot\ldots\cdot p_k^{n_k}(\theta).$$

Приклад. Оцінка $\overline{x}$ є МП-оцінкою параметра a нормального розподілу. Дійсно,

$$L(a) = \frac{1}{\sqrt{(2\pi)^n}\sigma^n} \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k - a)^2\right\}$$

Ця функція досягає свого максимального значення в точці, де її похідна дорівнює нулю. Оскільки

$$L'(a) = \frac{1}{\sqrt{(2\pi)^n}\sigma^{n+2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k - a)^2\right\} \times$$

$$\times \left(\sum_{k=1}^{n} x_k - na\right),$$

то $\widehat{a} = \frac{1}{n}\sum_{k=1}^{n} x_k = \overline{x}$.

2.2 Інтервальні оцінки

2.2.1 Надійний інтервал

Точкові оцінки параметра розподілу генеральної сукупності не дають можливості встановити точність, з якою вони наближають відповідний параметр. Вирішити цю проблему можна визначаючи так звані інтервальні оцінки параметра.

Означення. *Інтервальною оцінкою (надійним інтервалом)* параметра θ розподілу при рівні надійності γ називається випадковий інтервал $(\hat{\theta}_1; \hat{\theta}_2)$, де $\hat{\theta}_1$ і $\hat{\theta}_2$ — оцінки параметра θ, для якого виконується [3]

$$\mathbf{P}(\hat{\theta}_1 \le \theta \le \hat{\theta}_2) \ge \gamma.$$

[3] Іноді пишуть $\mathbf{P}(\hat{\theta}_1 \le \theta \le \hat{\theta}_2) = \gamma$ або $\mathbf{P}(\hat{\theta}_1 < \theta < \hat{\theta}_2) = \gamma$

Використовуючи центральну граничну теорему, у випадку великого об'єму вибірки можна наближено знайти межі надійного інтервалу для математичного сподівання.

Нехай x_1, x_2,...,x_n — вибірка з генеральної сукупності розподіленої за довільним законом з математичним сподіванням a та дисперсією σ^2. Тоді за центральною граничною теоремою при $n \to \infty$

$$\mathbf{P}\left(\frac{\bar{x}-a}{\sigma/\sqrt{n}} < t\right) \to \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{t} e^{-\frac{z^2}{2}}\,dz$$

Звідси легко одержати, що

$$\mathbf{P}\left(|\bar{x}-a| < \frac{\sigma}{\sqrt{n}}\,t\right) \to 2\Phi(t),$$

де $\Phi(t)$ — функція Лапласа.

Тоді при досить великих n має місце наближена рівність

$$\mathbf{P}\left(|\bar{x}-a| < \frac{\sigma}{\sqrt{n}}\,t\right) \approx 2\Phi(t),$$

Отже, при рівні надійності $\gamma = 2\Phi(t)$ надійним інтервалом для математичного сподівання a є проміжок

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}}\,t; \bar{x} + \frac{\sigma}{\sqrt{n}}\,t\right). \tag{2.5}$$

Зауваження. У вираз для одержаного надійного інтервалу входить точне значення середньоквадратичного відхилення генеральної сукупності. Якщо воно невідоме, то при великих n можна замінити його на незміщену оцінку s.

Зауваження. Інтервал (2.5) є точним у випадку, коли розподіл генеральної сукупності є нормальним з відомою дисперсією σ^2. Дійсно, якщо x_1, x_2,...,x_n — незалежні $N(a;\sigma^2)$-розподілені випадкові величини, то випадкова величина $\frac{\bar{x}-a}{\sigma/\sqrt{n}}$ має стандартний нормальний розподіл.

2.2.2 Інтервальні оцінки параметрів нормального розподілу

Нехай $x_1, ..., x_n$ — вибірка з генеральної сукупності нормального розподілу $N\left(a, \sigma^2\right)$. Розглядаючи x_i як однаково $N\left(a, \sigma^2\right)$ розподілені незалежні випадкові величини, можна записати $Mx_i = a$, $Dx_i = \sigma^2$.

Розглянемо окремі випадки.

Надійний інтервал для математичного сподівання при відомій дисперсії

Точкова оцінка математичного сподівання $\bar{x} = \frac{1}{n}\sum\limits_{k=1}^{n} x_k$. Оскільки

$$M\sum_{k=1}^{n} x_k = na, \quad D\sum_{k=1}^{n} x_k = n\sigma^2,$$

то

$$M\bar{x} = a, \quad D\bar{x} = \frac{\sigma^2}{n}.$$

Отже, $\bar{x}$ має нормальний розподіл $N\left(a, \frac{\sigma^2}{n}\right)$, як лінійна комбінація нормально розподілених випадкових величин.

Знайдемо таке ε, що $P\left(|\bar{x}-a| < \varepsilon\right) = \gamma$, де γ вибраний рівень надійності. Оскільки $P\left(|\bar{x}-a| < \varepsilon\right) = 2F_0\left(\frac{\sqrt{n}}{\sigma}\varepsilon\right) - 1$, де F_0 — функція стандартного нормального розподілу, то $F_0\left(\frac{\sqrt{n}}{\sigma}\varepsilon\right) = \frac{\gamma+1}{2}$. Звідси $\frac{\sqrt{n}}{\sigma}\varepsilon = u_{\frac{\gamma+1}{2}}$, де $u_{\frac{\gamma+1}{2}}$- квантиль порядку $\frac{\gamma+1}{2}$ стандартного нормального розподілу і $\varepsilon = u_{\frac{\gamma+1}{2}}\frac{\sigma}{\sqrt{n}}$

Отже, надійний інтервал для математичного сподівання при відомій дисперсії має вигляд:

$$\left(\bar{x} - u_{\frac{\gamma+1}{2}}\frac{\sigma}{\sqrt{n}}; \bar{x} + u_{\frac{\gamma+1}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

Надійний інтервал для математичного сподівання при невідомій дисперсії

Поряд з точковою оцінкою математичного сподівання $\bar{x} = \frac{1}{n}\sum\limits_{k=1}^{n} x_k$ розглянемо оцінку дисперсії $s^2 = \frac{1}{n-1}\sum\limits_{k=1}^{n}(x_k - \bar{x})^2$.

Можна довести, що випадкова величина $\frac{\bar{x}-a}{s}\sqrt{n}$ має розподіл Стьюдента з $n-1$ ступенем вільності. Тоді

$$P\left(|\bar{x} - a| < \varepsilon\right) = 2F_{St}\left(\frac{\sqrt{n}}{s}\varepsilon\right) - 1,$$

де F_{St} — функція розподілу Стьюдента з $n-1$ ступенем вільності.

Звідси $\varepsilon = \frac{s}{\sqrt{n}}t_{\frac{\gamma+1}{2}}(n-1)$, де $t_{\frac{\gamma+1}{2}}(n-1)$- квантиль порядку $\frac{\gamma+1}{2}$ розподілу Стьюдента з $n-1$ ступенями вільності.

Отже, надійний інтервал для математичного сподівання при невідомій дисперсії має вигляд:

$$\left(\bar{x} - \frac{s}{\sqrt{n}}t_{\frac{\gamma+1}{2}}(n-1); \bar{x} + \frac{s}{\sqrt{n}}t_{\frac{\gamma+1}{2}}(n-1)\right)$$

Надійний інтервал для дисперсії при відомому математичному сподіванні

Найкращою оцінкою дисперсії при відомому математичному сподіванні a є $\hat{\sigma^2} = \frac{1}{n}\sum\limits_{k=1}^{n}(x_k - a)^2$. Розглянемо

$$\frac{\hat{\sigma^2}}{\sigma^2}n = \sum_{k=1}^{n}\left(\frac{x_k - a}{\sigma}\right)^2$$

Сума квадратів незалежних нормально розподілених з параметрами 0 і 1 випадкових величин розподілена за законом Пірсона з n степенями вільності. Таким чином

$$P\left(\sigma_1^2 < \sigma^2 < \sigma_2^2\right) = P\left(n\frac{\hat{\sigma}^2}{\sigma_2^2} < n\frac{\hat{\sigma}^2}{\sigma^2} < n\frac{\hat{\sigma}^2}{\sigma_1^2}\right) = \gamma$$

$$= F_{\chi^2}\left(n\frac{\hat{\sigma}^2}{\sigma_1^2}\right) - F_{\chi^2}\left(n\frac{\hat{\sigma}^2}{\sigma_2^2}\right),$$

де $F_{\chi^2}(x)$ — функція розподілу Пірсона (χ^2) з n ступенями вільності. Вибравши σ_1 і σ_2 так, щоб

$$F_{\chi^2}\left(n\frac{\hat{\sigma}^2}{\sigma_2^2}\right) = \frac{1-\gamma}{2} \quad ; \quad F_{\chi^2}\left(n\frac{\hat{\sigma}^2}{\sigma_1^2}\right) = \frac{1+\gamma}{2},$$

одержимо $n\frac{\hat{\sigma^2}}{\sigma_2^2} = \chi^2_{\frac{1-\gamma}{2}}(n)$; $n\frac{\hat{\sigma^2}}{\sigma_1^2} = \chi^2_{\frac{1+\gamma}{2}}(n)$, де $\chi^2_{\frac{1-\gamma}{2}}(n)$ та $\chi^2_{\frac{1+\gamma}{2}}(n)$ — квантилі розподілу Пірсона з n степенями вільності порядку $\frac{1-\gamma}{2}$ та $\frac{1+\gamma}{2}$, відповідно.

Отже,

$$\sigma_1^2 = \frac{n\hat{\sigma^2}}{\chi^2_{\frac{1+\gamma}{2}}(n)} \quad \sigma_2^2 = \frac{n\hat{\sigma^2}}{\chi^2_{\frac{1-\gamma}{2}}(n)}$$

Таким чином надійний інтервал з надійністю γ для дисперсії нормального розподілу при відомому математичному сподіванні має вигляд:

$$\left(\frac{n\hat{\sigma^2}}{\chi^2_{\frac{1+\gamma}{2}}(n)};\frac{n\hat{\sigma^2}}{\chi^2_{\frac{1-\gamma}{2}}(n)}\right)$$

Надійний інтервал для дисперсії при невідомому математичному сподіванні

Оцінка дисперсії при невідомому математичному сподіванні має вигляд $s^2=\frac{1}{n-1}\sum\limits_{k=1}^{n}(x_k-\bar{x})^2$. Розглянемо

$$\frac{s^2}{\sigma^2}(n-1)=\sum_{k=1}^{n}\left(\frac{x_k-\bar{x}}{\sigma}\right)^2.$$

Аналогічно до пункту 2.2.2 отримаємо:

$$P\left(\sigma_1^2<\sigma^2<\sigma_2^2\right)=$$

$$P\left((n-1)\frac{s^2}{\sigma_2^2}<(n-1)\frac{s^2}{\sigma^2}<(n-1)\frac{s^2}{\sigma_1^2}\right)=\gamma$$

$$=F_{\chi^2}\left((n-1)\frac{s^2}{\sigma_1^2}\right)-F_{\chi^2}\left((n-1)\frac{s^2}{\sigma_2^2}\right),$$

де σ_1 і σ_1 вибрані так, щоб

$$F_{\chi^2}\left((n-1)\frac{s^2}{\sigma_2^2}\right)=\frac{1-\gamma}{2}\quad;\quad F_{\chi^2}\left((n-1)\frac{s^2}{\sigma_1^2}\right)=\frac{1+\gamma}{2}.$$

Отже,

$$\sigma_1^2=\frac{(n-1)s^2}{\chi^2_{\frac{1+\gamma}{2}}(n-1)},\quad \sigma_2^2=\frac{(n-1)s^2}{\chi^2_{\frac{1-\gamma}{2}}(n-1)},$$

де $\chi^2_{\frac{1-\gamma}{2}}(n-1)$ та $\chi^2_{\frac{1+\gamma}{2}}(n-1)$- квантилі розподілу Пірсона з $n-1$ степенями вільності порядків $\frac{1-\gamma}{2}$ та $\frac{1+\gamma}{2}$ відповідно.

Тому надійний інтервал з надійністю γ для дисперсії нормального розподілу при невідомому математичному сподіванні має вигляд:

$$\left(\frac{(n-1)\,s^2}{\chi^2_{\frac{1+\gamma}{2}}(n-1)};\frac{(n-1)\,s^2}{\chi^2_{\frac{1-\gamma}{2}}(n-1)}\right)$$

Вправи

1. Для вибірки об'єму $n = 15$ для нормально розподіленої випадкової величини:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.3	-1.6	-0.8	0.3	-1	0.1	0.4	-1.1	0	-1.1	-1.6	0.2	0	-1.6	-1.5

Знайти вибіркові середнє і дисперсію та вказати точність наближення математичного сподівання вибірковим середнім з надійністю 0.96 . З тією ж надійністю знайти довірчий інтервал для дисперсії.

2. Скільки вимірів фізичної величини потрібно провести, щоб точність інтервальної оцінки математичного сподівання стала рівною 0.5, а надійність цієї оцінки була не меншою 0.9 . Середня квадратична похибка вимірювального приладу $\sigma = 0.9$.

3. Нехай в n незалежних випробуваннях успіх відбувся x раз. Знайти довірчий інтервал для ймовірності p успіху в одному випробуванні.

4. При перевірці 100 деталей з великої партії виявлено 10 бракованих деталей.

1. Знайти 95%-ний наближений довірчий інтервал частини бракованих деталей у всій партії.

2. Який мінімальний об'єм вибірки потрібно взяти для того, щоб з ймовірністю 0.95 можна було стверджувати, що частина бракованих деталей у всій партії відрізняється від частоти їх появи у вибірці не більше ніж на 1%.

5. Нехай з одної генеральної сукупності нормального розподілу одержано дві вибірки об'ємів n_1 і n_2 відповідно. Вибіркові оцінки середніх та дисперсій за цими вибірками дорівнюють $\bar{x}_1$, $\bar{x}_2$, s_1^2, s_2^2. Об'єднані оцінки середнього та дисперсії за вибіркою об'єму $n_1 + n_2$ обчислюються за формулами

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} \quad , \quad s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Довести, що якщо дисперсія генеральної сукупності відома і дорівнює σ^2, то довірчий інтервал з надійністю α для математичного сподівання визначається таким чином:

$$\bar{x} - \frac{\sigma}{\sqrt{n_1 + n_2}} u_{(1+\alpha)/2} < m < \bar{x} + \frac{\sigma}{\sqrt{n_1 + n_2}} u_{(1+\alpha)/2} \quad ;$$

якщо дисперсія генеральної сукупності невідома, то довірчий інтервал з надійністю α для математичного сподівання визначається таким чином:

$$\bar{x} - \frac{s}{\sqrt{n_1 + n_2}} t_{(1+\alpha)/2}(n_1 + n_2 - 2) < m <$$
$$< \bar{x} + \frac{s}{\sqrt{n_1 + n_2}} t_{(1+\alpha)/2}(n_1 + n_2 - 2),$$

де u_γ — квантиль порядку γ стандартного нормального розподілу, $t_\gamma(n)$ — квантиль порядку γ розподілу Стьюдента з n ступенями вільності.

6. Знайти довірчий інтервал та точності відповідних інтервальних оцінок для невідомого математичного сподівання нормально розподіленої випадкової величини ξ, якщо $\sigma^2 = \mathbf{D}\,\xi = 25$, об'єм вибірки $n = 36$, точкова оцінка математичного сподівання $\bar{x} = 7$. Довірчі інтервали та відповідні їм точності побудувати для надійностей $\gamma = 0.95$, $\gamma = 0.99$,$\gamma = 0.999$.

Розділ 3

Теорія перевірки гіпотез

3.1 Статистичні гіпотези та критерії їх перевірки

3.1.1 Загальні поняття

Статистичною гіпотезою (гіпотезою) будемо називати твердження стосовно розподілу та їх параметрів випадкових величин. Істинність статистичної гіпотези перевіряється з використанням одержаної вибірки з генеральної сукупності значень випадкових величин, щодо яких сформульована гіпотеза. Слід відмітити, що результати перевірки статистичної гіпотези не є логічними доказами її істинності чи хибності. Ми будемо говорити, що одержана вибірка суперечить чи не суперечить сформульованій гіпотезі, а гіпотеза, відповідно, узгоджується чи не узгоджується з наявними статистичними даними. В першому випадку гіпотеза приймається (не відхиляється), в другому – відхиляється.

Гіпотезу, яка підлягає перевірці, будемо називати *основ-*

ною чи *нульовою* та позначатимемо H_0. Гіпотезу, що суперечить H_0, називатимемо *конкуруючою* чи *альтернативною*. Вона позначається H_1.

Правило, за яким приймається рішення відхилити чи не відхилити основну гіпотезу, називається *критерієм*. Оскільки рішення приймається на основі спостережень над значеннями випадкової величини ξ, то потрібно вибрати відповідну статистику Z, яка в цьому випадку називається, *статистикою критерію*. Перевірка статистичної гіпотези базується на принципі, згідно якого малоймовірні події є неможливими, а події з великими ймовірностями є, відповідно, достовірними. Тобто перед початком аналізу вибірки вибирається деяке число, що є верхним рівнем малих ймовірностей. Це число називається рівнем значущості і частіше всього позначається буквою α. Нехай V — множина значень статистики Z, а $V_k \subseteq V$ — така множина, що при умові істинності гіпотези H_0 ймовірність попадання статистики критерію в V_k дорівнює α. Множина V_k називається критичною областю, а множина $V \setminus V_k$ — областю прийняття гіпотези. Тобто, якщо вибіркове значення статистики критерію попадає в критичну область, то основна гіпотеза повинна бути відхилена як така, що не узгоджується з одержаною вибіркою. Якщо ж вибіркове значення цієї статистики не попадає в V_k, то гіпотеза приймається, тобто вважається, що вона не суперечить даній вибірці.

Очевидно, оскільки перевірка правильності статистичної гіпотези на основі вибірки містить долю випадковості в своїх результатах, то при цьому можливі помилки. Розрізняють *помилку першого типу*, що полягає у відхиленні основної гіпотези тоді, коли вона насправді правильна, і *помилку другого типу*, коли основна гіпотеза приймається тоді, як правильною є альтернативна гіпотеза. Ймовірність помилки першого типу дорівнює ймовірності попадання статистики критерію в критичну область при умові, що правильна основна гіпотеза, тобто дорівнює рівню значущості: $\mathbf{P}(Z \in V_k/H_0) = \alpha$. Оскільки ця ймовірність мала, то ми вважаємо, що значення статистики критерію не може попасти в критичну область, якщо основна гіпотеза правильна.

Ймовірність помилки другого типу

$$\beta = \mathbf{P}(Z \in V \setminus V_k/H_1),$$

нажаль, не може бути зроблена як завгодно малою для даної вибірки. Зменшити ймовірність помилки другого роду можна шляхом збільшення об'єму вибірки. Якщо ж вибірка задана, бажано основною гіпотезою вибирати те припущення, для якого помилкове відхилення є більш небажаним, ніж його прийняття коли воно неправильне.

Приклад. Перевіряється шкідливість нового лікарського препарату на основі спостережень за тими, хто його приймав. Основною гіпотезою H_0 можна вибрати припущення, що препарат шкідливий (тоді альтернативна гіпотеза $H_1 = \{$препарат нешкідливий$\}$) або, що препарат нешкідливий (тоді альтернативна гіпотеза $H_1 = \{$препарат шкідливий$\}$). В першому випадку помилка першого типу означає, що шкідливий препарат буде визнано нешкідливим, а в другому випадку — що нешкідливий препарат буде визнано шкідливим. Наслідком такої помилки в першому випадку є загроза здоров'ю і життю людей, а в другому випадку — тільки збитки фірми, що розробила цей препарат. Очевидно, що перший спосіб вибору основної та альтернативної гіпотез є більш прийнятним.

3.1.2 Критерії, що базуються на інтервальних оцінках

Нехай $x_1, x_2, ..., x_n$ – вибірка об'єму n з генеральної сукупності значень випадкової величини, розподіл якої задається функцією розподілу $F(x, \theta)$, що залежить від параметра θ. Розглянемо гіпотезуH_0, яка полягає в тому, що $\theta = \theta_0$. Альтернативною гіпотезою до H_0 будемо вважати одну з гіпотез $H_1' : \theta \neq \theta_0$, $H_1'' : \theta < \theta_0$, $H_1''' : \theta > \theta_0$. Критерії з альтернативними гіпотезами H_1'', H_1''' називаються односторонніми, відповідно, лівостороннім та правостороннім. Перевірку гіпотези H_0 можна проводити, базуючись на інтервальних оцінках (довірчих інтервалах) параметра θ.

Як відомо, для побудови інтервальної оцінки параметра θ потрібно побудувати деяку (бажано ефективну) оцінку $\widehat{\theta}$ параметра θ та знайти таку статистику $g\left(\widehat{\theta}, \theta\right)$, розподіл якої відомий і не залежить від θ. Крім того функція $g\left(\widehat{\theta}, \theta\right)$ повинна бути непервна та строго монотонна по θ.

Зафіксуємо деякий рівень α — рівень значущості та побудуємо довірчий інтервал (θ_1, θ_2) чи лівосторонній $(-\infty; \theta_2)$ або правосторонній $(\theta_1; +\infty)$ довірчі інтервали з надійністю $1 - \alpha$. Після цього критерій перевірки гіпотези H_0 виглядатиме так: якщо альтернативною гіпотезою є H_1', то коли $\theta_0 \notin (\theta_1, \theta_2)$ гіпотеза відхиляється, а коли $\theta_0 \in (\theta_1, \theta_2)$, гіпотеза H_0 не відхиляється; якщо альтернативною гіпотезою є H_1'', то коли $\theta_0 \geq \theta_2$ гіпотеза H_0 відхиляється, а коли $\theta_0 < \theta_2$ не відхиляється; якщо альтернативною гіпотезою є H_1''', то коли $\theta_0 \leq \theta_1$, гіпотеза H_0 відхиляється, а коли $\theta_0 > \theta_1$ не відхиляється. При цьому в усіх випадках ймовірність помилки першого роду (відхилити правильну гіпотезу) дорівнює α (не більша, ніж α).

Зауважимо, що рівень значущості α можна вибрати неоднозначно. Тому для кожного фіксованого θ_0 верхня границя ймовірності помилки першого роду буде визначатися тим значенням α, при якому θ_0 співпаде з однією із границь довірчого інтервалу.

3.2 Гіпотези про параметри нормального розподілу

Через центральну граничну теорему нормальний розподіл відіграє особливу роль в математичній статистиці. Тому розглянемо гіпотези про параметри a та σ^2 нормального розподілу.

Нехай дана вибірка $x_1, x_2, ..., x_n$ з генеральної сукупності значень випадкової величини з розподілом $N(a, \sigma^2)$. Параметри розподілу можуть бути як невідомими, так і відомими (один з них). Випадок, коли відомі обидва параметри інтересу не представляє.

Розглянемо гіпотезу $H_0 : a = a_0$ при умові, що дисперсія відома. Значення σ характеризує точність приладу, з допомогою якого одержані вибіркові значення, і тому може бути відомим. Тоді статистика $\frac{\bar{x}-a}{\sigma/\sqrt{n}}$ має розподіл $N(0;1)$ і двосторонній довірчий інтервал при рівні значущості α має вигляд:

$$\left(\bar{x} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}; \bar{x} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

Односторонні довірчі інтервали такі:

$$\left(-\infty; \bar{x} + u_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right) \text{та} \quad \left(\bar{x} - u_{1-\alpha}\frac{\sigma}{\sqrt{n}}; +\infty\right)$$

Тут u_γ- квантиль порядку γ стандартного нормального розподілу $N(0;1)$.

Отже, якщо $\bar{x} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < a_0 < \bar{x} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ чи $a_0 < \bar{x} - u_\alpha\frac{\sigma}{\sqrt{n}}$ або $a_0 > \bar{x} - u_{1-\alpha}\frac{\sigma}{\sqrt{n}}$, то гіпотеза H_0 не відхиляється (не суперечить емпіричному розподілу вибірки), відповідно до альтернативних гіпотез $H_1' : a \neq a_0$, $H_1'' : a < a_0$, $H_1''' : a > a_0$.

Розглянемо тепер ту ж саму гіпотезу H_0 при умові, що дисперсія σ^2 невідома (що частіше всього буває). Тоді статистика $\frac{\bar{x}-a}{s/\sqrt{n}}$, де $s^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i - \bar{x})^2$ — точкова оцінка дисперсій σ^2, має розподіл Стьюдента з $n-1$ ступенем вільності $(T(n-1))$.

Довірчі інтервали в розглядуваних нами трьох випадках мають такий вигляд (рівень значущості α):
двосторонній $\left(\bar{x} - t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}; \bar{x} + t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}\right)$,
лівосторонній $\left(-\infty; \bar{x} - t_\alpha(n-1)\frac{s}{\sqrt{n}}\right)$,
правосторонній $\left(\bar{x} - t_{1-\alpha}(n-1)\frac{s}{\sqrt{n}}; +\infty\right)$,
де $t_\gamma(n-1)$ — квантиль порядку γ розподілу Стьюдента з $n-1$ ступенем вільності.

Тепер очевидно, якими є критерії перевірки гіпотези H_0 в цих випадках.

Перейдемо до гіпотези про дисперсію нормального розподілу. Нехай $H_0 : \sigma^2 = \sigma_0^2$ при можливих альтернативах $H_1' : \sigma^2 \neq \sigma_0^2$, $H_1'' : \sigma^2 < \sigma_0^2$, $H_1''' : \sigma^2 > \sigma_0^2$:.

Розглянемо перший варіант: параметр a відомий. Це може бути у випадку, коли вимірюється деяка еталонна величина з метою визначення точності приладу (в припущенні, що прилад систематичної помилки не дає). Статистика $\frac{n\cdot s_0^2}{\sigma^2}$, де $s_0^2 = \frac{1}{n}\sum\limits_{i=1}^{n}(x_i - a)^2$ — точкова оцінка дисперсій σ^2, має розподіл Пірсона (χ^2) з n ступенями вільності.

При рівні значущості α довірчі інтервали для σ^2, мають такий вигляд:
двосторонній $\left(\frac{n\cdot s_0^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}; \frac{n\cdot s_0^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right)$,
лівосторонній $\left(0; \frac{n\cdot s_0^2}{\chi^2_{\alpha}(n)}\right)$,
правосторонній $\left(\frac{n\cdot s_0^2}{\chi^2_{1-\alpha}(n)}; +\infty\right)$.

Тут $\chi^2_\gamma(n)$ – квантиль порядку γ розподілу Пірсона з n ступенями вільності.

Отже, якщо σ_0 попадає в один з цих інтервалів, то гіпотеза H_0 не суперечить емпіричному розподілу вибірки при відповідній альтернативі, в іншому випадку гіпотезу H_0 потрібно відхилити як таку, що не узгоджується з вибіркою. При цьому, як і раніше, ймовірність помилки першого роду дорівнює α (не перевищує α).

Нехай тепер математичне сподівання не відоме (основний випадок). Тоді статистика $\frac{(n-1)s^2}{\sigma^2}$ має розподіл Пірсона з $n-1$ ступенем вільності. Відповідні довірчі інтервали такі:
двосторонні $\left(\frac{(n-1)s^2}{\chi_{1-\frac{\alpha}{2}}(n-1)}; \frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}}(n-1)}\right)$,
лівосторонні $\left(0; \frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)}\right)$,
правосторонні $\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)}; +\infty\right)$.

Тут $s^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2$ — точкова оцінка дисперсій σ^2, $\chi^2_\gamma(n-1)$ – квантиль порядку γ розподілу Пірсона з $n-1$ ступенем вільності. Таким чином критерієм перевірки гіпотези H_0, при розглянутих альтернативних гіпотезах, є попадання значення σ_0^2 у відповідний довірчий інтервал.

Приклад. Для дослідження ефекту від використання нової технології у виробництві деяких виробів в порівнянні з існуючою технологією виробничий підрозділ працював

10 днів за новою технологією та 10 днів — за старою. На початку кожної пари днів з допомогою підкидання монети вирішувалось за якою технологією працювати першого дня. Знайдені різниці обсягів виробництва продукцій для кожної пари днів (від обсягу продукції, виготовленої за новою технологією віднімався обсяг продукції, виготовленої за старою технологією) наведено в таблиці:

№ пари	Різниця	№ пари	Різниця
1	2,4	6	1,6
2	1,0	7	-0,4
3	0,7	8	1,1
4	0,0	9	0,1
5	1,1	10	0,7

Чи підтверджують одержані дані наявність ефекту від впровадження нової технології?

Розв'язок. У термінах перевірки статистичних гіпотез задачу можна сформулювати так. Маємо 10 незалежних спостережень нормально розподіленої ($N(a; \sigma^2)$) випадкової величини. Причому, параметри цього розподілу невідомі. Стосовно параметра a висувається гіпотеза $H_0 : a = 0$. Невідхилення цієї гіпотези будемо трактувати як відсутність ефекту від впровадження нової технології. Як альтернативу розглядаємо однобічну гіпотезу $H_1 : a > 0$. Вибір альтернативи продиктований намаганням встановити наявність ефекту (приросту виробництва продукції) від використання нової технології. Відхилення гіпотези H_0 будемо інтерпретувати як наявність ефекту.

Знайдемо

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

$$= \frac{1}{10}(2,4 + 1,0 + ... + 0,7) = 0,83,$$

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

$$= \frac{1}{9}((2,4 - 0,83)^2 + ... + (0,7 - 0,83)^2) = 0,667$$

та за таблицею розподілу Стьюдента (візьмемо рівень значущості $\alpha = 0,05$)

$$t_{1-0,05}(10 - 1) = 1,833.$$

Оскільки

$$\bar{x} - t_{1-\alpha}\,(n-1)\,\frac{s}{\sqrt{n}} = 0,83 - 1,833 \cdot \sqrt{\frac{0,667}{10}} = 0,356 > 0,$$

то на рівні значущості 0,05 гіпотеза $H_0 : a = 0$ відхиляється. Тобто припущення про відсутність ефекту нової технології суперечить наявним даним. При цьому ми могли помилитися (помилка першого типу) з ймовірністю 0,05.

3.3 Дисперсійний аналіз

Припустимо, що певна величина вимірюється під впливом деякого фактора. Це може бути фактор "прилад", якщо вона вимірюється різними приладами, чи фактор "умова", якщо вона вимірюється в різних умовах і т. д. Нас цікавитиме питання про рівність середніх значень, одержаних при різних значеннях фактора. Для вивчення цього питання застосуємо *однофакторний дисперсійний аналіз*.

Нехай результати спостережень утворюють l незалежних вибірок, одержаних з генеральних сукупностей значень l нормально розподілених випадкових величин з математичними сподіваннями a_1, a_2,..., a_l та рівними дисперсіями σ^2.

Розглянемо гіпотезу $H_0 : a_1 = a_2 = ... = a_l$. Якщо гіпотеза H_0 не суперечить вибіркам, то можна стверджувати, що фактор не впливає на середнє значення результатів вимірів. В іншому випадку середні значення деяких вибірок мають суттєві відмінності, що свідчить про вплив даного фактора. Позначимо через x_{ij} $(i = 1, 2, ..., n_j;\ j = 1, 2, ..., l)$ i-те вибіркове значення j-тої вибірки. Нехай $\overline{x_j} = \frac{1}{n_j}\sum\limits_{i=1}^{n_l} x_{ij}$ — вибіркове середнє j-тої вибірки; $\overline{x} = \frac{1}{n}\sum\limits_{j=1}^{l}\sum\limits_{i=1}^{n_j} x_{ij}$ — вибіркове середнє об'єднаної вибірки (тут $n = \sum\limits_{j=1}^{l} n_j$).

Розглянемо

$$\sum_{j=1}^{l}\sum_{i=1}^{n_l}(x_{ij} - \bar{x})^2 = \sum_{j=1}^{l}\sum_{i=1}^{n_l}(x_{ij} - \overline{x_j} + \overline{x_j} - \bar{x})^2 =$$

$$= \sum_{j=1}^{l}\sum_{i=1}^{n_l}(x_{ij} - \overline{x_j})^2 + 2\sum_{j=1}^{l}\sum_{i=1}^{n_l}(x_{ij} - \overline{x_j})(\overline{x_j} - \bar{x}) +$$

$$+ \sum_{j=1}^{l}\sum_{i=1}^{n_l}(x_j - \bar{x})^2 =$$

$$= \sum_{j=1}^{l}\sum_{i=1}^{n_l}(x_{ij} - \overline{x_j})^2 + \sum_{i=1}^{l} n_i(\overline{x_i} - \bar{x})^2.$$

Тут враховано, що

$$\sum_{j=1}^{l}\sum_{i=1}^{n_l}\left(x_{ij}-\overline{x_j}\right)\left(\overline{x_j}-\bar{x}\right)=\sum_{j=1}^{l}\left(\overline{x_j}-\bar{x}\right)\left(\sum_{i=1}^{n_l}x_{ij}-n_j\overline{x_j}\right)=0$$

за означенням $\overline{x_j}$.

Позначивши

$$\sum_{j=1}^{l}\sum_{i=1}^{n_l}\left(x_{ij}-\overline{x_j}\right)^2=Q_1, \tag{3.1}$$

$$\sum_{j=1}^{l}n_j\left(\overline{x_j}-\bar{x}\right)^2=Q_2, \tag{3.2}$$

$$\sum_{j=1}^{l}\sum_{i=1}^{n_l}\left(x_{ij}-\bar{x}\right)^2=Q, \tag{3.3}$$

одержимо рівність

$$Q=Q_1+Q_2 \tag{3.4}$$

Цю рівність називають *основним рівнянням дисперсійного аналізу.* Сума Q описує загальну мінливість вибіркових значень відносно загального середнього значення. Сума Q_1 задає мінливість вибіркових значень всередині кожної вибірки відносно свого вибіркового середнього. Сума Q_2 характеризує відхилення вибіркових середніх кожної вибірки від загального середнього. Значне переважання Q_2 над Q_1 в рівності (3.4) свідчить не на користь гіпотези H_0 (тобто є вплив фактора на результати вимірювань).

Вияснимо, що означають слова “значне переважання”. Якщо гіпотеза H_0 правильна, то статистики $\frac{Q_1}{\sigma^2}$ та $\frac{Q_2}{\sigma^2}$ незалежні та мають розподіли Пірсона з $n-l$ та $l-1$ ступенями вільності, відповідно. Звідси випливає, що статистики

$S_1^2 = \frac{Q_1}{n-l}$, $S_2^2 = \frac{Q_2}{l-1}$ є незміщеними оцінками параметра σ^2. Відношення цих оцінок

$$\frac{S_2^2}{S_1^2} = \frac{\frac{Q_2}{l-1}}{\frac{Q_1}{n-l}} = F\left(l-1, n-l\right) \tag{3.5}$$

має розподіл Фішера з $l-1$ та $n-l$ ступенями вільності. Тому статистику (3.5) можна використати для перевірки гіпотези H_0. А саме, якщо вибіркове значення $F_{\text{В}}$ статистики (3.5) є малими (тобто $F_{\text{В}} < F_{1-\alpha}(l-1, n-l)$), то гіпотеза H_0 не суперечить результатам спостережень, а якщо $F_{\text{В}}$ велике (тобто $F_{\text{В}} \geq F_{1-\alpha}\left(l-1, n-l\right)$), то гіпотеза H_0 відхиляється. Тут $F_{1-\alpha}\left(l-1, n-l\right)$ — квантиль порядку $1-\alpha$ розподілу Фішера з $l-1$ та $n-l$ ступенями вільності. При цьому ймовірність помилки першого типу дорівнює α.

У випадку, коли гіпотеза не відхиляється, незміщеною оцінкою спільного значення математичних сподівань є статистика $\overline{x}$. Якщо гіпотеза відхиляється, то хоча б два значення a_j є різними. Для того, щоб визначити, які саме математичні сподівання є різними, можна застосувати метод *лінійних контрастів*. Лінійним контрастом називається лінійна комбінація математичних сподівань $L = \sum\limits_{j=1}^{l} c_j a_j$, де c_j, $j = 1, 2, ..., l$ деякі сталі, $\sum\limits_{j=1}^{l} c_j = 0$. Оцінкою величини L є, очевидно, статистика

$$\hat{L} = \sum_{j=1}^{l} c_j \overline{x_j}. \tag{3.6}$$

Оскільки

$$D(\hat{L}) = \sum_{j=1}^{l} c_j^2 D\overline{x_j} = \sum_{j=1}^{l} c_j^2 \frac{1}{n_j^2} \sum_{i=1}^{n_l} Dx_{ij} =$$

$$= \sum_{j=1}^{l} c_j^2 \frac{1}{n_j^2} n_j \sigma^2 = \sigma^2 \sum_{j=1}^{l} \frac{c_j^2}{n_j},$$

то оцінкою дисперсії $\hat{L}$ є статистика

$$s_L^2 = \frac{Q_2}{n-l} \sum_{j=1}^{l} \frac{c_j^2}{n_j}. \tag{3.7}$$

Межі довірчого інтервалу (при значущості α) для лінійного контраста L мають вигляд:

$$\hat{L} \pm s_L \sqrt{(l-1)\, F_{1-\alpha}\,(l-1, n-l)}, \tag{3.8}$$

де $F_{1-\alpha}\,(l-1, n-l)$ — квантиль порядку $1-\alpha$ розподілу Фішера з $l-1$ та $n-l$ ступенями вільності.

Перевіряючи гіпотези типу $a_m = a_k$, розглядаємо відповідно лінійні контрасти $L = a_m - a_k$. Якщо 0 попадає у відповідний довірчий інтервал, то гіпотеза не відхиляється, а в іншому випадку — відхиляється і значення a_m та a_k слід вважати різними.

Приклад. Три групи студентів навчались за різними методиками. Після закінчення певного періоду навчання було проведено тестовий контроль над випадково вибраними студентами з кожної групи. Кількість помилок, зроблених студентами наведено в таблиці:

№ групи	Кількість помилок
1	1 3 2 1 0 2 1
2	2 3 2 1 4
3	4 5 3

При рівні значущості $\alpha = 0,05$ перевірити гіпотезу про відсутність впливу різних методик на результати навчання. Якщо гіпотеза буде відхилена, визначити, які саме методики мають різний вплив на результати студентів.

Розв'язок. Очевидно задача полягає в перевірці гіпотези $H_0 : a_1 = a_2 = a_3$, де a_k — математичне сподівання кількості помилок в k-тій групі. В нашому випадку $l = 3$, $n = 15$.

Обчислимо

$$\overline{x_1} = \frac{1}{7}(1 + 3 + ... + 1) = 1,428,$$

$$\overline{x_2} = \frac{1}{5}(2 + 3 + ... + 4) = 2,4,$$

$$\overline{x_3} = \frac{1}{3}(4 + 5 + 3) = 4,$$

$$\overline{x} = \frac{1}{34}(1 + 3 + ... + 3) = 2,267.$$

Далі з (3.1) — (3.3)

$$Q = (1 - 2,267)^2 + (3 - 2,267)^2 + ... + (3 - 2,267)^2 = 26,93,$$

$$Q_1 = (1 - 1,428)^2 + ... + (1 - 1,428)^2 + (2 - 2,4)^2 + ... \\ +(4 - 2,4)^2 + +(4 - 4)^2 + ... + (3 - 4)^2 = 14,02,$$

$$Q_2 = 7 \cdot (1,428 - 2,267)^2 + 5 \cdot (2,4 - 2,267)^2 + 3 \cdot (4 - 2,267)^2 = \\ = 12,91.$$

Обчислимо вибіркове значення статистики (3.5):

$$F_B = \frac{Q_1/(l-1)}{Q_2/(n-l)} = \frac{14,02/2}{12,91/12} = 6,52.$$

З таблиці розподілу Фішера знаходимо $F_{0,95}(2, 12) = 3,89$. Оскільки $F_B = 6,52 > 3,89$, то гіпотеза про рівність середніх суперечить наявним даним і повинна бути відхилена. Тобто є вплив різних методик на результати навчання.

Перевіримо які ж саме з трьох середніх суттєво відрізняються. Для цього розглянемо три гіпотези $H_0^{(1)} : a_1 = a_2$, $H_0^{(2)} : a_1 = a_2$, $H_0^{(3)} : a_1 = a_2$ з двосторонніми альтернативами. Побудуємо відповідні лінійні контрасти:

$$L_1 = a_1 - a_2,\ L_2 = a_1 - a_3,\ L_3 = a_2 - a_3.$$

Обчислимо оцінки лінійних контрастів (див. (3.6)):

$$\hat{L_1} = 1,428 - 2,4 = -0,972,$$

$$\hat{L_2} = 1,428 - 4 = -2,572,$$

$$\hat{L_3} = 2,4 - 4 = -1,6;$$

та їх дисперсій (див. 3.7):

$$s_{L_1}^2 = \frac{12,91}{15-3} \cdot \left(\frac{1}{7} + \frac{1}{5}\right) = 0,37,$$

$$s_{L_2}^2 = \frac{12,91}{15-3} \cdot \left(\frac{1}{7} + \frac{1}{3}\right) = 0,51,$$

$$s_{L_3}^2 = \frac{12,91}{15-3} \cdot \left(\frac{1}{5} + \frac{1}{3}\right) = 0,58.$$

Вибравши рівень значущості $\alpha = 0,05$, з таблиці розподілу Фішера знаходимо $F_{1-\alpha}(l-1, n-l) = F_{0,95}(2,12) = 3,89$. Отже, межі довірчих інтервалів для розглянутих лінійних контрастів, відповідно, дорівнюють (використаємо формулу (3.8)):

$$-0,97 \pm 1,67,\ -2,57 \pm 2,00,\ -1,6 \pm 2,12.$$

Звідси видно, що гіпотези $a_1 = a_2$, $a_2 = a_3$ не відхиляються (нуль міститься у відповідних довірчих інтервалах);

гіпотеза $a_1 = a_3$ — відхиляється (нуль не належить відповідному довірчому інтервалу). Отже, відмінності середніх значень кількості помилок в першій та другій групах, в другій та третій групах є несуттєвими. Відмінність між середніми значеннями в першій та третій групах істотна, що говорить про вплив різних методик навчання, застосованих в цих групах. Тут варто нагадати, що прийняття чи відхилення статистичної гіпотези не є доведенням логічної істинності відповідного твердження.

3.4 Критерій χ^2

Нехай $x_1, x_2, ..., x_n$– вибірка об'єму n з генеральної сукупності значень деякої випадкової величини ξ. Розглянемо гіпотезу H_0 яка полягає в тому, що розподіл випадкової величини ξ задається, наприклад, функцією розподілу $F(x)$. Розподіл в формулюванні гіпотези повністю визначений, тому у випадку, коли він залежить від деяких параметрів попередньо оцінимо ці параметри за вибіркою і значеннями параметрів візьмемо їхні оцінки. Тому гіпотезу H_0 можна формулювати як належність розподілу випадкової величини до певного параметричного класу розподілів, якщо тільки параметри цих розподілів можуть бути оцінені за вибіркою.

Отже, нехай розподіл заданий функцією $F(x)$ залежить від l параметрів оцінених за даною вибіркою. Ідея критерію Пірсона (χ^2) полягає в тому, щоб порівняти частоти попадання вибіркових значень в інтервали числової осі з частотами попадання в ці інтервали випадкової величини ξ з функцією розподілу $F(x)$.

Розіб'ємо множину можливих значень випадкової величини ξ на k інтервалів, якщо ξ неперервна випадкова величина, чи на k груп окремих значень, якщо ξ дискретна. Нехай це будуть множиниΔ_1, Δ_2,..., Δ_k. Визначимо частоти n_i, $i = 1, 2, ..., k$ попадання вибіркових значень в множини Δ_i та частоти $n\,p_i$ попадання випадкової величини ξ в ці множини, де $n = \sum\limits_{i=1}^{k} n_i$, $p_i = P\,(\xi \in \Delta_i)$, $i = 1, 2, ..., k$. Мірою відхилення значень n_i від $n\,p_i$ вибирається статистика

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}.$$

При умові, що всі Δ_i такі, що $n\,p_i \geq 10$, то з достатньою точністю можна вважати, що розподіл статистик $\frac{n_i - np_i}{\sqrt{np_i}}$ є близьким до стандартного нормального $N\,(0, 1)$. Звідси випливає, що статистика χ^2 має розподіл Пірсона з $k - l - 1$ ступенем вільності, де l — кількість параметрів розподілу оцінених за вибіркою.

Зафіксувавши деякий рівень значущості α, знайдемо квантиль $\chi^2_{1-\alpha}\,(k - l - 1)$ порядку $1 - \alpha$ розподілу Пірсона з $k - l - 1$ ступенем вільності. Тоді якщо для $\chi^2_{\text{В}}$ — значення статистики χ^2 на даній вибірці,

$$\chi^2_{\text{В}} < \chi^2_{1-\alpha}\,(k - l - 1)\,,$$

то гіпотеза H_0 не суперечить вибірці, а якщо

$$\chi^2_{\text{В}} \geq \chi^2_{1-\alpha}\,(k - l - 1)\,,$$

то гіпотеза H_0 суперечить вибірці і повинна бути відхилена. При цьому ймовірність відхилити правильну гіпотезу дорівнює α.

Зауважимо, що якщо вимога $n\,p_i \geq 10$, $i = 1, 2, ..., k$ не виконується для всіх Δ_i, то потрібно об'єднати деякі Δ_i з сусідніми так, щоб ця вимога була виконана. При цьому k

Межі	3,0-3,6	3,6-4,2	4,2-4,8	4,8-5,4	5,4-6,0	6,0-6,6	6,6-7,2
x_i	3,3	3,9	4,5	5,1	5,7	6,3	6,9
x_i^2	10,89	15,21	20,25	26,01	32,49	39,69	47,61
n_i	2	11	30	40	28	12	7

Табл. 3.1:

повинно бути більшим за $l+1$. Якщо об'єм вибірки настільки малий, що цього добитися неможливо, то застосовувати критерій Пірсона не рекомендується.

Приклад. Виміряно вхідний опір 130 електронних ламп (Ом):

Межі	3,0-3,6	3,6-4,2	4,2-4,8	4,8-5,4	5,4-6,0	6,0-6,6	6,6-7,2
Частота	2	11	30	40	28	12	7

Вибравши 10%-ний рівень значущості, перевірити гіпотезу H_0 про те, що вхідний опір електронних ламп нормально розподілений.

Розв'язок. Оцінимо параметри гіпотетичного розподілу (математичне сподівання та дисперсію). Для цього знайдемо середини x_i даних інтервалів та їх квадрати x_i^2 (табл. 3.1).

Отже,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{j} x_i n_i = \frac{1}{130}\cdot(3,3\cdot 2+...+6,9\cdot 7) = 5,169,$$

$$s^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{j} x_i^2 n_i - \overline{x}^2\right) =$$

$$= \frac{130}{129}\left(\frac{1}{130}(10,89\cdot 2+...+47,61\cdot 7)-(5,169)^2\right) = 0,64,$$

$$s = \sqrt{0,64} = 0,8.$$

Тут $n=130$ — об'єм вибірки, $j=7$ — кількість інтервалів.

α_i	β_i	p_i	np_i	np_i	n_i	n_i	$\frac{(n_i-np_i)^2}{np_i}$
3	3.6	0.02	2.80	$+\downarrow$	2	$+\downarrow$	
3.6	4.2	0.09	11.43	14.23	11	13	0.11
4.2	4.8	0.21	27.22	27.22	30	30	0.28
4.8	5.4	0.29	37.87	37.87	40	40	0.12
5.4	6	0.24	30.81	30.81	28	28	0.26
6	6.6	0.11	14.65	18.71	12	19	0.004
6.6	7.2	0.03	4.06	$+\uparrow$	7	$+\uparrow$	

Табл. 3.2:

Для кожного з розглянутих інтервалів $(\alpha_i; \beta_i)$ обчислимо $p_i = \Phi\left(\frac{\beta_i - \overline{x}}{s}\right) - \Phi\left(\frac{\alpha_i - \overline{x}}{s}\right)$, np_i та об'єднаємо з сусідніми ті інтервали, для яких $np_i < 10$ (табл. 3.2).

Звідси $\chi^2_{\text{В}} \approx 0,77$. Оскільки $\chi^2_{0,9}(5-2-1) = 4,61$, то гіпотеза H_0 підтверджується статистичними даними.

Критерій Пірсона (χ^2) можна застосовувати і до перевірки гіпотези про незалежність двох випадкових величин.

Припустимо, що проведено n експериментів, результати яких є значеннями дискретних випадкових величин ξ та η, які приймають відповідно значення $x_1, x_2, ..., x_k$ та $y_1, y_2, ..., y_l$. Якщо ξ та η неперервні випадкової величини, то область значень кожної з них розбиваємо на скінченну кількість інтервалів. Нехай n_{ij} – кількість експериментів, в яких випадкова величина ξ прийняла значення x_i, чи попала в і-тий інтервал, а випадкова величина η прийняла значення y_i, чи попала в j-тий інтервал своїх значень. Розглянемо гіпотезу H_0 про те, що випадкові величини ξ та η незалежні. Якщо гіпотеза H_0 правильна, то за означенням незалежності для дискретних випадкових величин

$$P\left(\xi = x_i; \eta = y_i\right) = P\left(\xi = x_i\right) P\left(\eta = y_i\right).$$

Оцінками ймовірностей, що входять в цю рівність є, відповідно, статистики: $\frac{n_{ij}}{n}$, $\frac{n_{i\cdot}}{n}, \frac{n_{\cdot j}}{n}$, де $n_{i\cdot} = \sum\limits_{j=1}^{l} n_{ij}$, $n_{\cdot j} = \sum\limits_{j=1}^{k} n_{ij}$.

Мірою близькості $\frac{n_{ij}}{n}$ та $\frac{n_{i\cdot} n_{\cdot j}}{n^2}$ можна взяти статистику:

$$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{\left(n_{ij} - \hat{n}_{ij}\right)^2}{\hat{n}_{ij}},$$

де $\hat{n}_{ij} = n \frac{n_{i\cdot}}{n} \frac{n_{\cdot j}}{n} = \frac{n_{i\cdot} n_{\cdot j}}{n}$. Аналогічну рівність можна записати і для неперервних випадкових величин.

Якщо гіпотеза H_0 правильна, то статистика χ^2 має розподіл Пірсона з $(k-1)(l-1)$ ступенем вільності за умови, що n достатньо велике.

Умовою, за якої є допустимими застосування критерію Пірсона, як і в попердньому випадку, є виконання нерівностей $\hat{n}_{ij} \geq 10$ для всіх $i = 1, 2, ..., k$, $j = 1, 2, ..., l$. Якщо ця вимога не виконується, то відповідні інтервали значень випадкових величин треба об'єднати з сусідніми.

На практиці проведення описаної процедури перевірки гіпотези про незалежність зручно починати з побудови таблиці спряженості ознак:

$N\eta \setminus N\xi$	1	2	...	l	$n_{i\cdot}$
1	n_{11}	n_{12}	$\ldots$	n_{1l}	$n_{1\cdot}$
2	n_{21}	n_{22}	$\ldots$	n_{2l}	$n_{2\cdot}$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
k	n_{k1}	n_{k2}	$\ldots$	n_{kl}	$n_{k\cdot}$
$n_{\cdot j}$	$n_{\cdot 1}$	$n_{\cdot 2}$	$\ldots$	$n_{\cdot l}$	n

В таблиці $N\xi$ та $N\eta$ означає номери значень чи інтервалів значень, відповідно, випадкових величин ξ та η.

Зауваження. Випадкові величини ξ та η можна розглядати як дві ознаки, за якими класифікуються елементи деякої вибірки. Незалежність ξ та η означає незалежність ознак.

Приклад. Маючи результати опитування 100 студентів перших трьох курсів на питання "чи вважаєте Ви, що куріння заважає навчанню?", на рівні значущості $\alpha = 0,01$ перевірити гіпотезу про те, що відношення до куріння студентів різних курсів різне.

Відповідь	Курс		
	I	II	III
Ні	15	10	—
Не знаю	8	5	7
Так	—	30	25

Розв'язок. Перевіримо незалежність ознак "курс"та "відношення до куріння". Якщо ці ознаки будуть признані незалежними, то не можна говорити про різне ставлення до куріння студентів різних курсів. Якщо ж вони залежні, то ставлення до куріння студентів різних курсів є різним.

Складемо таблицю спряженості ознак.

$N\eta \setminus N\xi$	1	2	3	$n_{i\cdot}$
1	15	10	0	25
2	8	5	7	20
3	0	30	25	55
$n_{\cdot j}$	23	45	32	100

Обчислимо значення $\hat{n_{ij}}$ (табл. 3.3).

Тому вибіркове значення статистики χ^2 дорівнює 44,24.

$i \setminus j$	1	2	3
1	5,75	11,25	8
2	4,6	9	6,4
3	12,65	24,75	17,6

Табл. 3.3:

Квантиль розподілу Пірсона з $(3-1)(3-1) = 4$ ступенями вільності знаходимо з таблиці значень його функції розподілу: $\chi^2_{0,99}(4) = 13,277$.

Отже, гіпотеза про незалежність розглянутих ознак на рівні значущості $0,01$ відхиляється (бо $44,24 > 13,277$). Тобто відношення до куріння у студентів різних курсів різне.

Вправи

1. Для наведеної згрупованої вибірки вхідного опору електронних ламп (Ом), припустивши 10%-ний рівень значущості, перевірити гіпотезу H_0 про те, що вона одержана з нормально розподіленої генеральної сукупності.

Границі інтервалів	3.0-3.6	3.6-4.2	4.2-4.8	4.8-5.4	5.4-6.0	6.0-6.6	6.6-7.2
Частота	2	8	35	43	22	15	5

2. Під час епідемії грипу вивчалась ефективність щеплень проти цього захворювання. Одержані наступні результати:

	Після щеплення	Без щеплення
Захворіли	4 чол.	34 чол.
Не захворіли	192 чол.	111 чол.

Чи вказують ці результати на ефективність прищеплень? Прийняти $\alpha = 0.01$.

3. На хімічному заводі розроблено два нові варіанти технологічного процесу. Щоб оцінити, як зміниться денна продуктиність при переході на роботу за новими варіантами технологічного процесу, завод протягом 10 днів пряцював за кожним варіантом, включаючи існуючий варіант. Денна продуктивність заводу (в умовних одиницях) наводиться в таблиці. Припустивши, що вибірки зроблені з нормально розподілених генеральних сукупностей з рівними дисперсіями, перевірити гіпотезу про рівність середніх. Якщо гіпотеза правильна, то знайти незсувні оцінки середнього та дисперсії, в іншому випадку провести попарне порівнянння середніх з допомогою метода лінійних контрастів.

День роботи	Добова продуктивність		
	Існуюча схема	Варіант 1	Варіант 2
1	46	74	52
2	48	82	63
3	73	64	72
4	52	72	64
5	72	84	48
6	44	68	70
7	66	76	78
8	46	88	68
9	60	70	70
10	48	60	54
Сумма	555	738	639

$\sum\sum x_{ik} = 1932$, $\sum\sum x_{ik}^2 = 128810$. Прийняти $\alpha = 0.10$.

4. В цеху з 10 станками кожний день реєструвалась кількість станків, що вийшли з ладу. Всього було проведено 200 спостережень, результати яких приведені нижче:

Кількість станків, що вийшли з ладу	0	1	2	3	4	5	6	7	8	9	10
Кількість зареєстрованих випадків	41	62	45	22	16	8	4	2	0	0	0

Перевірити гіпотезу H_0 про те, що кількість станків, що вийшли з ладу, має розподіл Пуасона. Прийняти $\alpha = 0.05$.

5. Досліджуються два виробничих процеси виготовлення поршневих кілець. Використовуючи критерій χ^2, перевірити гіпотезу про рівність процента браку в обидвох процесах за наведеними даними при $\alpha = 0.01$:

Кільця	Процес	
	1	2
Якісні	195	149
Браковані	5	2

6. В трьох магазинах, що продають товари одного виду, дані товарообігу за 8 місяців роботи (в тис гр. од.) склали таке зведення:

Магазин	Місяць								$\sum_{i=1}^{8} x_{ik}$
	1	2	3	4	5	6	7	8	
I	19	23	26	18	20	20	18	35	179
II	20	20	32	27	40	24	22	18	203
III	16	15	18	26	19	17	19	18	148

$\sum_{k=1}^{3} \sum_{i=1}^{8} x_{ik}^2 = 12592.$

Припустивши, що вибірки зроблені з нормально розподілених генеральних сукупностей з рівними дисперсіями, перевірити гіпотезу про рівність середніх. Якщо гіпотеза правильна, то знайти незсувні оцінки середнього та дисперсії, в іншому випадку провести попарне порівнянння середніх з допомогою метода лінійних контрастів. Прийняти $\alpha = 0.1$.

Розділ 4

Кореляційний та регресійний аналізи

4.1 Кореляційний аналіз

Нехай в результаті спостереження над випадковими величинами ξ та η одержано двовимірну вибірку (x_i, y_i), $i = 1, 2, ..., n$. Нас цікавитиме питання про корельованість випадкових величин.

Як відомо, коефіцієнт кореляції випадкових величин ξ та η є число

$$r\left(\xi, \eta\right) = \frac{M\xi\eta - M\xi M\eta}{\sigma\left(\xi\right)\sigma\left(\eta\right)}.$$

Замінюючи числові характеристики випадкових величин ξ та η їх точковими оцінками, одержимо точкову оцінку коефіцієнта кореляції:

$$\hat{r} = \frac{\overline{xy} - \overline{x}\;\overline{y}}{s_x s_y}$$

Якщо $\hat{r}$ близьке до 0, то можна говорити про майже некорельованість випадкових величин ξ та η. Для нормально розподілених величин це означає майже незалежність. У випадку, коли $|\hat{r}|$ близький до 1, то можна стверджувати про залежність між ξ та η як майже лінійну.

Зрозуміло, що поняття близькості $\hat{r}$ до 0 чи 1 є нечітким, тому бажано було б мати критерії перевірки відповідних статистичних гіпотез. Робити це можна, наприклад, з допомогою довірчих інтервалів. Якщо вибірка, одержана з генеральної сукупності, що має двовимірний нормальний розподіл, то при досить великих значеннях об'єму вибірки n статистика $Z = \frac{1}{2}\ln\frac{1+\hat{r}}{1-\hat{r}} = \operatorname{arth}\hat{r}$ має приблизно нормальний розподіл $N\left(\operatorname{arth}\hat{r}, \frac{1}{\sqrt{n-3}}\right)$

Тому довірчий інтервал для r має вигляд

$$\left(\operatorname{th}\left(\operatorname{arth}\hat{r} - \frac{u_{\frac{1+\gamma}{2}}}{\sqrt{n-3}}\right); \operatorname{th}\left(\operatorname{arth}\hat{r} + \frac{u_{\frac{1+\gamma}{2}}}{\sqrt{n-3}}\right)\right) \tag{4.1}$$

Довірча ймовірність дорівнює γ. Якщо 0 чи ± 1 попадає в інтервал (4.1), то можна стверджувати, що випадкові величини незалежні чи лінійно залежні. Ймовірність помилки першого роду в цьому випадку дорівнює $\alpha = 1-\gamma$. Тут, як і раніше, $u_{\frac{1+\gamma}{2}}$ — квантиль порядку $\frac{1+\gamma}{2}$ стандартного нормального розподілу $N(0,1)$.

Іноді, особливо коли вибірка одержана з генеральної сукупності двовимірного дискретного випадкового вектора, зручно задавати вибірку у вигляді *кореляційної таблиці.*

Нехай випадкова величина ξ приймає такі різні значення $x_1, x_2, ..., x_k$, а випадкова величина η має значення $y_1, y_2, ..., y_l$. Пара (x_i, y_i) у вибірці зустрічається n_{ij} раз $(i = 1, 2, ..., k;\ j = 1, 2, ..., l)$. Кореляційна таблиця має такий вигляд (табл. 4.1). В цьому випадку оцінки коефіцієн-

	y_1	y_2	$\dots$	y_l
x_1	n_{11}	n_{12}	$\dots$	n_{1l}
x_2	n_{21}	n_{22}	$\dots$	n_{2l}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_k	n_{k1}	n_{k2}	$\dots$	n_{kl}

Табл. 4.1:

та кореляції залишаються тими ж самими, якщо врахувати, що:

$$n = \sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij}, \quad n_{i\cdot} = \sum_{j=1}^{l} n_{ij}, \quad n_{\cdot j} = \sum_{i=1}^{k} n_{ij},$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_{i\cdot}, \quad \overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_{i\cdot}, \quad \overline{xy} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{l} x_i y_j n_{ij}, \tag{4.2}$$

$$\bar{y} = \frac{1}{n}\sum_{j=1}^{l} y_j n_{\cdot j}, \quad \overline{y^2} = \frac{1}{n}\sum_{j=1}^{l} y_j^2 n_{\cdot j} \tag{4.3}$$

Приклад. Нехай за вибіркою об'єму $n = 300$ ($n = 500$) одержано оцінку коефіцієнта кореляції $\hat{r} = 0,1$ між двома нормально розподіленими випадковими величинами. Чи можна стверджувати про наявність залежності між цими величинами?

Розв'язок. Виберемо довірчу ймовірність $\gamma = 0,95$ та знайдемо довірчі інтервали для коефіцієнта кореляції у випадках коли $n = 300$ та $n = 500$.

Оскільки $u_{\frac{1+0,95}{2}} = 1,96$, то

$$\operatorname{th}\left(\operatorname{arth} 0,1 - \frac{1,96}{\sqrt{300-3}}\right) = -0,013;$$

$$\text{th}\left(\text{arcth}\,0,1 + \frac{1,96}{\sqrt{300-3}}\right) = 0,211.$$

Тобто довірчий інтервал для коефіцієнта кореляції (у випадку $n = 300$) $(-0,013; 0,211)$ і говорити про залежність випадкових величин не можна (бо нуль міститься в довірчому інтервалі). Коли ж $n = 500$, тоді

$$\text{th}\left(\text{arth}\,0,1 - \frac{1,96}{\sqrt{500-3}}\right) = 0,012;$$

$$\text{th}\left(\text{arth}\,0,1 + \frac{1,96}{\sqrt{500-3}}\right) = 0,186$$

і довірчий інтервал для коефіцієнта кореляції $(0,012; 0,186)$. Отже, в цьому випадку існує залежність між випадковими величинами.

З розглянутого прикладу видно, що значення вибіркового коефіцієнта кореляції мало говорить про наявність залежності між випадковими величинами. Приклад демонструє наявність суттєвого впливу об'єму вибірки.

4.2 Лінійна регресія. Метод найменших квадратів

4.2.1 Парна лінійна регресія

Нехай вивчається залежність величин y від величин x (x вважається незалежною змінною). Щоб визначити цю залежність будемо надавати величині x деяких різних значень $x_1, x_2, ..., x_n$ та вимірювати відповідні значення величини y: $y_1, y_2, ..., y_n$. Оскільки будь-які результати вимірів містять в собі випадкові помилки, то значення y_i будемо вважати реалізаціями випадкових величин η_i. Нас цікавитиме можливість побудови такої залежності між x та y виду $y = ax+b$, для якої помилки ε_i в рівностях $y_i = ax_i+b+\varepsilon_i$ були б в деякому розумінні найменшими. Розглядаючи випадок, коли вимірювання проводиться без систематичних помилок можна записати, що

$$M_x(\eta) = ax + b, \tag{4.4}$$

де η випадкова величина — результат виміру величини y, M_x — умовне математичне сподівання при умові, що незалежна змінна прийняла значення x. Рівність (4.4) називається лінійною регресією η на x (або y на x).

Основним методом побудови оцінок параметрів a і b лінійної регресії є метод *найменших квадратів*. Він полягає в тому, що значення a і b знаходимо з умови мінімізації суми квадратів помилок ε_i.

Розглянемо функцію

$$f(a,b) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2.$$

Оскільки функція f є квадратичною (сума квадратів), то вона має єдину точку екстремуму (мінімуму), яку можна знайти з необхідних умов локального екстремуму $\frac{\partial f}{\partial a} = 0$, $\frac{\partial f}{\partial b} = 0$.

Таким чином ми одержуємо, систему рівнянь

$$\left\{ \begin{array}{l} -2\sum\limits_{i=1}^{n} (y_i - ax_i - b)\, x_i = 0 \\ -2\sum\limits_{i=1}^{n} (y_i - ax_i - b) = 0 \end{array} \right. .$$

Звідси

$$\begin{cases} a\sum\limits_{i=1}^{n} x_i^2 + b\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i y_i \\ a\sum\limits_{i=1}^{n} x_i + nb = \sum\limits_{i=1}^{n} y_i \end{cases} \tag{4.5}$$

Систему рівнянь (4.5) називають нормальною системою рівнянь для параметрів лінійної регресії. Розв'язавши систему (4.5), знайдемо точкові оцінки параметрів лінійної регресії

$$\hat{a} = \frac{\overline{xy} - \overline{x}\ \overline{y}}{\overline{x^2}}, \quad \hat{b} = \overline{y} - \hat{a}\overline{x}, \tag{4.6}$$

де середні значення, задаються тими ж виразами, що і в попередній лекції.

Іноді залежність між x та y задається у вигляді кореляційної таблиці. Це буває, коли значення величини є дискретними і вимірювання величини y при кожному даному x_i проводиться декілька раз. Тоді оцінки параметрів лінійної регресії задаються тими ж формулами (4.6), в яких середні значення обчислюються за (4.2), (4.3).

Зауважимо, що залежність y від x може бути і нелінійною. Тому лінійна регресія може неадекватно описувати цю залежність. Вибір виду залежності корисно зробити попередньо, наприклад, побудувати на координатній площині точки з координатами (x_i, y_i) (чи $(x_i; y_{ij})$, коли для одного значення x_i є кілька значень y_{ij}). Якщо ж розміщення точок на площині явно не може бути задано лінійною залежністю, то потрібно підібрати таку нелінійну регресією, яку можна було б звести до лінійної і застосувати метод найменших квадратів. Серед таких нелінійних регресій варто виділити:

показникову $y = be^{ax} \Leftrightarrow \ln y = \ln b + az$;

степеневу $y = bx^2 \Leftrightarrow \ln y = \ln b + a \ln x$;

логарифмічну $y = \ln(ax + b) \Leftrightarrow e^y = ax + b$;

гіперболічну $y = \frac{a}{x+b} \Leftrightarrow \frac{1}{y} = \frac{1}{a}x + \frac{b}{a}$;

поліноміальну $y = a_n x^n + a_{n-1}x^{n-1} + ... + a_0$.

Перші чотири залежності легко зводяться до лінійних відносно деяких нових змінних. Поліноміальна регресія є лінійною відносно зележної змінної y та змінних $x_1 = x$, $x_2 = x^2$,..., $x_n = x^n$. Отже, ми одержуємо множинну лінійну регресію $y = a_0 + a_1x_1 + ... + a_nx_n$. Аналогічно випадку парної лінійної регресії оцінки параметрів регресії одержуємо з нормальної системи рівнянь

$$\begin{pmatrix} 1 & \overline{x_1} & \overline{x_2} & ... & \overline{x_n} \\ \overline{x_1} & \overline{x_1^2} & \overline{x_1x_2} & ... & \overline{x_1x_n} \\ \overline{x_2} & \overline{x_1x_2} & \overline{x_2^2} & ... & \overline{x_2x_n} \\ ... & ... & ... & ... & ... \\ \overline{x_n} & \overline{x_1x_n} & \overline{x_2x_n} & ... & \overline{x_n^2} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ ... \\ a_n \end{pmatrix} = \begin{pmatrix} \overline{y} \\ \overline{x_1y} \\ \overline{x_2y} \\ ... \\ \overline{x_ny} \end{pmatrix}.$$

4.2.2 Властивості лінійних регресій

Розглянемо значення $\widehat{y}_i = \widehat{a}x_i + \widehat{b}$ одержані за оціненою формулою регресії. Параметри оціночної регресії знайдені за результатами спостережень над величинами x та y з допомогою метода найменших квадратів. Різниці $e_i = y_i - \widehat{y}_i$ називаються залишками, які показують наскільки добре оціночна (вибіркова) регресія апроксимує результати спостережень (x_i, y_i), $i = 1, 2, ..., n$.

Залишковою дисперсією називається величина

$$s^2 = \frac{\sum\limits_{i=1}^{n} e_i^2}{n-2}. \tag{4.7}$$

Величина $Q_e = \sum\limits_{i=1}^{n} e_i^2$ називається залишковою сумою квадратів.

Припустимо, що помилки спостережень ε_i є незалежними нормально розподіленими ($\varepsilon_i \sim N(0; \sigma^2)$) випадковими величинами. Звідси випливає, що у випадку адекватності лінійної моделі результати спостережень y_i, $i = 1, 2, ..., n$ є реалізаціями незалежних нормально розподілених величин η_i ($\eta_i \sim N(ax_i + b, \sigma^2)$, $i = 1, 2, ..., n$).

Точковою оцінкою дисперсії σ^2 є залишкова дисперсія s^2. Крім того статистика $\frac{Q_e}{\sigma^2}$ має розподіл χ^2 з $n-2$ ступенями вільності незалежно від розподілу оцінок $\hat{a}$, $\hat{b}$.

Використовуючи це твердження, можна побудувати довірчі інтервали для параметрів лінійної регресії та перевіряти гіпотези про параметри.

Довірчий інтервал для параметра a має вигляд

$$\left(\widehat{a} - t_{\frac{1+\gamma}{2}}(n-2)\, s\sqrt{\frac{1}{Q_x}};\ \widehat{a} + t_{\frac{1+\gamma}{2}}(n-2)\, s\sqrt{\frac{1}{Q_x}}\right), \quad (4.8)$$

а для парметра b —

$$\left(\widehat{b} - t_{\frac{1+\gamma}{2}}(n-2)s\sqrt{\frac{\overline{x^2}}{Q_x}}; \widehat{b} + t_{\frac{1+\gamma}{2}}(n-2)s\sqrt{\frac{\overline{x^2}}{Q_x}}\right), \quad (4.9)$$

де $t_{\frac{1+\gamma}{2}}(n-2)$ — квантиль порядку $\frac{1+\gamma}{2}$ розподілу Стьюдента з $n-2$ ступенями вільності, $Q_x = \sum\limits_{i=1}^{n}(x_i - \overline{x})^2$. Довірчі ймовірності цих інтервалів дорівнюють γ.

Лінійна регресійна модель називається незначимою, якщо параметр $a = 0$. Для перевірки гіпотези $H_0 : a = 0$ можна використати довірчий інтервал (4.8). Таким чином, якщо 0 накривається інтервалом (4.8), то при рівні значущості α регресійну модель вважаємо незначимою. Якщо гіпотеза H_0 відхиляється, то модель вважається статистично значимою. Це, звичайно, не означає, що модель добре узгоджується з результатами спостережень.

Розглянемо

$$Q_y = \sum_{i=1}^{n} (y_i - \overline{y})^2 = \sum_{i=1}^{n} (y_i - \widehat{y}_i + \widehat{y}_i - \overline{y})^2 =$$

$$= \sum_{i=1}^{n} (\widehat{y}_i - \overline{y})^2 + 2\sum_{i=1}^{n} (y_i - \widehat{y}_i)(\widehat{y}_i - \overline{y}) + \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2$$

Оскільки $\sum\limits_{i=1}^{n} (y_i - \widehat{y}_i)(\widehat{y}_i - \overline{y}) = \sum\limits_{i=1}^{n} e_i (\widehat{y}_i - \bar{y}) = \widehat{a} \sum\limits_{i=1}^{n} e_i x_i + \widehat{b} \sum\limits_{i=1}^{n} e_i - \bar{y} \sum\limits_{i=1}^{n} e_i = 0$, то

$$Q_y = Q_R + Q_e,$$

де $Q_R = \sum\limits_{i=1}^{n} (\widehat{y}_i - \bar{y})$.

В цій рівності Q_y називається загальною сумою квадратів відхилень значень величини y від їх середнього значення, Q_R – сума квадратів, пояснена регресією, Q_e – як і раніше, залишкова сума квадратів. Домінування Q_R в сумі, що визначає Q_y, може свідчити про адекватність лінійної моделі результатам спостережень. Відношення $R^2 = \frac{Q_R}{Q_y}$ називається *коефіцієнтом детермінації*. Коефіцієнт детермінації дорівнює тій частині розкиду результатів спостереження відносно горизонтальної прямої $y = \bar{y}$, що пояснюється рівнянням регресії $y = \widehat{a}x + \widehat{b}$. Величина $R = \sqrt{R^2}$ є оцінкою коефіцієнта кореляції між результатами спостережень y_i та знайденими $\widehat{y}_i$.

Наведемо довірчий інтервал для прогнозу $y_0 = ax_0 + b$

$$\left(\widehat{y}_0 - t_{\frac{1+\gamma}{2}}(n-2) \cdot s \cdot \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Q_x}};\right.$$

$$\left.\widehat{y}_0 + t_{\frac{1+\gamma}{2}}(n-2) \cdot s \cdot \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{Q_x}}\right), \qquad (4.10)$$

де $\widehat{y}_0 = \widehat{a}x_0 + \widehat{b}$.

Довірчий інтервал для дисперсії помилок спостережень σ^2 має вигляд

$$\left(\frac{(n-2)\, s^2}{\chi^2_{\frac{1+\gamma}{2}}(n-2)};\ \frac{(n-2)\, s^2}{\chi^2_{\frac{1-\gamma}{2}}(n-2)}\right),$$

де $\chi^2_\beta(n-2)$ — квантиль порядку β розподілу Пірсона (χ^2) з $n-2$ ступенями вільності.

Приклад. В таблиці 4.2 наведено результати вимірювання чутливості відеоканалу ξ та звукового каналу η. Знайти оцінки (точкові та інтервальні) параметрів лінійної регресії, що описує залежність між ξ та η; при рівні значущості $0,05$ перевірити значимість лінійної моделі; з надійністю $0,95$ знайти довірчий інтервал для прогнозованого значення чутливості звукового каналу при чутливості відеоканалу 220 одиниць.

Розв'язок. Знайдемо об'єм вибірки $n = 20$ і:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{20} \cdot 2405 = 120,25,$$

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{1}{20} \cdot 350875 = 17543,75,$$

$\xi = x_i$	240	200	190	180	170	160	150	140	130	120
$\eta = y_i$	170	180	200	230	240	250	280	300	310	320
$\xi = x_i$	110	100	90	80	70	65	60	55	50	45
$\eta = y_i$	330	350	380	400	410	420	430	440	450	460

Табл. 4.2:

$$\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i = \frac{1}{20} \cdot 686400 = 34320,$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{20} \cdot 6550 = 327,5.$$

Тому оцінки параметрів лінійної регресії (див. (4.6)):

$$\hat{a} \approx -1,64, \quad \hat{b} \approx 524,89.$$

Знайдемо залишкову дисперсію за (4.7) та суму $Q_x = \sum\limits_{i=1}^{n} (x_i - \overline{x})^2$:

$$s^2 \approx \frac{2793,58}{18} \approx 155,20, \quad Q_x \approx 61673,75.$$

Виберемо довірчу ймовірність $\gamma = 0,95$ та знайдемо з таблиць квантиль порядку $\frac{1+\gamma}{2}$ розподілу Стьюдента з $n-2$ ступенями вільності $t_{\frac{1+\gamma}{2}}(n-2) = 2,101$. Тому довірчі інтервали для параметрів лінійної регресії (див. (4.8), (4.9)):

- для a

$$\Bigg(-1.64 - 2,101 \cdot \sqrt{155,20 \cdot \frac{17543,75}{61673,75}};$$

$$-1.64 + 2,101 \cdot \sqrt{155,20 \cdot \frac{17543,75}{61673,75}}\Bigg) = (-1,75; -1,53);$$

- для b

$$\left(524,89 - 2,101 \cdot \sqrt{155,20 \cdot \frac{1}{61673,75}};\right.$$

$$\left. 524,89 + 2,101 \cdot \sqrt{155,20 \cdot \frac{1}{61673,75}}\right) = (510,93; 538,85).$$

Оскільки нуль не попадає в довірчий інтервал для параметра a лінійної регресії, то при рівні значущості $0,05$ лінійна модель є значимою.

Довірчий інтервал для прогнозу значення величини чутливості звукового каналу при чутливості відеоканалу 220 одиниць знайдемо з (4.10) $\widehat{y}_0 = -1,64 \cdot 220 + 524,89 = 164,09$:

$$\left(164,09 - 2,101 \cdot \sqrt{155,2} \cdot \sqrt{\frac{1}{20} + \frac{(220-120,25)^2}{61673,75}};\right.$$

$$\left. 164,09 + 2,101 \cdot \sqrt{155,2} \cdot \sqrt{\frac{1}{20} + \frac{(220-120,25)^2}{61673,75}}\right) =$$

$$= (152,06; 176,12).$$

Вправи

1. За заданою вибіркою знайти оцінки для параметрів лінійної регресії Y на x, перевірити значущість лінійної регресії і обчислити коефіцієнт кореляції. Знайти границі довірчих інтервалів для параметрів лінійної моделі і для середнього значення Y при $x = x_0$

Припускається, що похибки спостережень незалежні і мають нормальний розподіл $N(0, \sigma)$. Рівень значущості α задається. При обчисленні слід використовувати значення сум змінних, їх квадратів і попарних добутків.

x	2	3	8	10	14	15	4	12	3	7	6
y	14.39	9.45	7.05	5.32	16.94	1.97	8.75	3.41	13.37	8.22	9.39

$\sum x_i = 86$, $\sum y_i = 98.26$, $\sum x_i^2 = 868$, $\sum y_i^2 = 1087.91$, $\sum x_i y_i = 682.25$, $\alpha = 0.05$.

2. Перевірити адекватність лінійної регресії. Побудувати графік залишків. Прийняти $\alpha = 0.05$. Вибірка задана у вигляді таблиці частот.

y \ x	1	2	3	4	5	6
1	2	1	-	-	-	-
2	1	2	-	-	-	-
3	-	3	1	-	-	-
4	-	1	3	-	-	-
5	-	-	2	2	2	1
6	-	-	-	1	1	1

3. Вважаючи, що залежність між змінними x і $\overline{y}$ має вигляд $y = \beta_0 + \beta_1 x + \beta_2 x^2$, знайти оцінки параметрів за наступною вибіркою.

x	0	2	4	6	8	10
y	5	-1	-0.5	1.5	4.5	8.5

4. За вибіркою спостережень потрібно:

а)знайти оцінки параметрів моделі $y = \beta_0 + \beta_1 x + \beta_2 x^2$;

б)перевірити значущість моделі;

в) знайти оцінки дисперсії похибок спостережень та їх коваріаційної матриці;

г) визначити довірчі інтеравали для параметрів в дисперсії похибок спостережень при заданому рівні значущості α.

Припускається, що похибки спостережень не корельовані і мають нормальний розподіл $N(0, \sigma)$.

x	-3	-2	-1	0	1	2	3
y	6	0	-1	-1	1	5	12

, $\alpha = 0,10$

5. Дослідження залежності тривалості t розв'язку систем лінійних рівнянь від порядку системи n дало такі результати:

n	2	3	4	5	6	7	8	9	10
t, хв	12	35	75	130	210	315	445	600	800

Припустивши, $t = An^{\gamma}$, знайти оцінки параметрів A і γ.

6. За заданою вибіркою знайти оцінки для параметрів лінійної регресії Y на x, перевірити значущість лінійної регресії і обчислити коефіцієнт кореляції. Знайти границі довірчих інтервалів для параметрів лінійної моделі і для середнього значення Y при $x = x_0$

Припускається, що похибки спостережень незалежні і мають нормальний розподіл $N(0, \sigma)$. Рівень значущості α задається. При обчисленні слід використовувати значення сум змінних, їх квадратів і попарних добутків.

x	2.7	4.6	6.3	7.8	9.2	10.6	12.0	13.4	14.7
y	17.0	16.2	13.3	13.0	9.7	9.9	6.2	5.8	5.7

$\sum x_i = 81.3$, $\sum y_i = 96.8$, $\sum x_i^2 = 865.63$, $\sum y_i^2 = 1194$, $\sum x_i y_i = 735.7$, $\alpha = 0.1$.

7. Вважаючи, що залежність між змінними x і y має вигляд $y = \beta_0 + \beta_1 x + \beta_2 x^2$, зайти оцінки параметрів за наведеною вибіркою:

x	-2	-1	0	1	2
y	4.8	0.4	-3.4	0.8	3.2

8. Отримана вибірка спостережень змінних x і y:

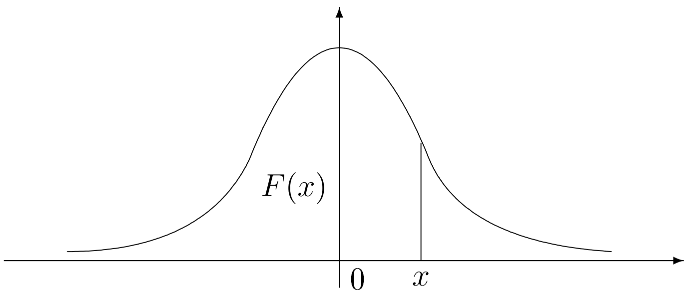
x	1	2	3	4	5	6	7
y	62.1	87.2	109.3	127.3	134.7	136.2	136.9

Для зображення цих даних пропонується використовувати модель $y = \frac{x}{\beta_0 + \beta_1 x}$. Знайти оцінки параметрів β_0 і β_1.

Додаток А

Статистичні таблиці

Значення функції розподілу $N(0,1)$



$$F(-x) = 1 - F(x)$$

Таблиця А.1

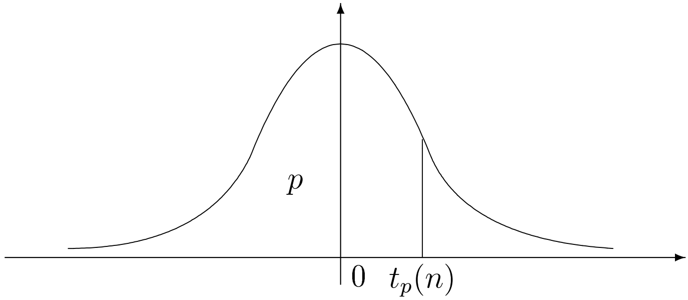
	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	,5000	,5040	,5080	,5120	,5160	,5199	,5239	,5279	,5319	,5359
0,1	,5398	,5438	,5478	,5517	,5557	,5596	,5636	,5675	,5714	,5753
0,2	,5793	,5832	,5871	,5910	,5948	,5987	,6026	,6064	,6103	,6141
0,3	,6179	,6217	,6255	,6293	,6331	,6368	,6406	,6443	,6480	,6517
0,4	,6554	,6591	,6628	,6664	,6700	,6736	,6772	,6808	,6844	,6879
0,5	,6915	,6950	,6985	,7019	,7054	,7088	,7123	,7157	,7190	,7224
0,6	,7257	,7291	,7324	,7357	,7389	,7422	,7454	,7486	,7517	,7549
0,7	,7580	,7611	,7642	,7673	,7704	,7734	,7764	,7794	,7823	,7852
0,8	,7881	,7910	,7939	,7967	,7995	,8023	,8051	,8078	,8106	,8133
0,9	,8159	,8186	,8212	,8238	,8264	,8289	,8315	,8340	,8365	,8389

Продовження таблиці А.1

	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
1	,8413	,8438	,8461	,8485	,8508	,8531	,8554	,8577	,8599	,8621
1,1	,8643	,8665	,8686	,8708	,8729	,8749	,8770	,8790	,8810	,8830
1,2	,8849	,8869	,8888	,8907	,8925	,8944	,8962	,8980	,8997	,9015
1,3	,9032	,9049	,9066	,9082	,9099	,9115	,9131	,9147	,9162	,9177
1,4	,9192	,9207	,9222	,9236	,9251	,9265	,9279	,9292	,9306	,9319
1,5	,9332	,9345	,9357	,9370	,9382	,9394	,9406	,9418	,9429	,9441
1,6	,9452	,9463	,9474	,9484	,9495	,9505	,9515	,9525	,9535	,9545
1,7	,9554	,9564	,9573	,9582	,9591	,9599	,9608	,9616	,9625	,9633
1,8	,9641	,9649	,9656	,9664	,9671	,9678	,9686	,9693	,9699	,9706
1,9	,9713	,9719	,9726	,9732	,9738	,9744	,9750	,9756	,9761	,9767
2	,9772	,9778	,9783	,9788	,9793	,9798	,9803	,9808	,9812	,9817
2,1	,9821	,9826	,9830	,9834	,9838	,9842	,9846	,9850	,9854	,9857
2,2	,9861	,9864	,9868	,9871	,9875	,9878	,9881	,9884	,9887	,9890
2,3	,9893	,9896	,9898	,9901	,9904	,9906	,9909	,9911	,9913	,9916
2,4	,9918	,9920	,9922	,9925	,9927	,9929	,9931	,9932	,9934	,9936
2,5	,9938	,9940	,9941	,9943	,9945	,9946	,9948	,9949	,9951	,9952
2,6	,9953	,9955	,9956	,9957	,9959	,9960	,9961	,9962	,9963	,9964
2,7	,9965	,9966	,9967	,9968	,9969	,9970	,9971	,9972	,9973	,9974
2,8	,9974	,9975	,9976	,9977	,9977	,9978	,9979	,9979	,9980	,9981
2,9	,9981	,9982	,9982	,9983	,9984	,9984	,9985	,9985	,9986	,9986
3	,9987	,9987	,9987	,9988	,9988	,9989	,9989	,9989	,9990	,9990
3,5	$,9_3 77$	$,9_3 78$	$,9_3 78$	$,9_3 79$	$,9_3 80$	$,9_3 81$	$,9_3 81$	$,9_3 82$	$,9_3 83$	$,9_3 83$
4	$,9_4 68$	$,9_4 70$	$,9_4 71$	$,9_4 72$	$,9_4 73$	$,9_4 74$	$,9_4 75$	$,9_4 76$	$,9_4 77$	$,9_4 78$
4,5	$,9_5 66$	$,9_5 68$	$,9_5 69$	$,9_5 70$	$,9_5 72$	$,9_5 73$	$,9_5 74$	$,9_5 76$	$,9_5 77$	$,9_5 78$
5	$,9_6 71$	$,9_6 73$	$,9_6 74$	$,9_6 75$	$,9_6 77$	$,9_6 78$	$,9_6 79$	$,9_6 80$	$,9_6 81$	$,9_6 82$

9_n означає n-кратне повторення цифри 9

Квантилі $t_p(n)$ порядку p розподілу Стьюдента з n ступенями вільності



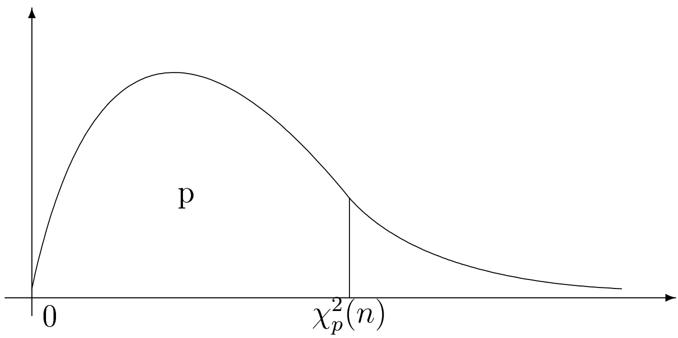
Таблиця А.2

$n \setminus p$	0,75	0,8	0,85	0,9	0,95	0,955	0,96	0,965	0,97	0,975	0,98	0,985	0,99	0,995
1	1,00	1,38	1,96	3,08	6,31	7,03	7,92	9,06	10,6	12,7	15,9	21,2	31,8	63,7
2	0,82	1,06	1,39	1,89	2,92	3,10	3,32	3,58	3,90	4,30	4,85	5,64	6,96	9,92
3	0,76	0,98	1,25	1,64	2,35	2,47	2,61	2,76	2,95	3,18	3,48	3,90	4,54	5,84
4	0,74	0,94	1,19	1,53	2,13	2,23	2,33	2,46	2,60	2,78	3,00	3,30	3,75	4,60
5	0,73	0,92	1,16	1,48	2,02	2,10	2,19	2,30	2,42	2,57	2,76	3,00	3,36	4,03
6	0,72	0,91	1,13	1,44	1,94	2,02	2,10	2,20	2,31	2,45	2,61	2,83	3,14	3,71
7	0,71	0,90	1,12	1,41	1,89	1,97	2,05	2,14	2,24	2,36	2,52	2,71	3,00	3,50
8	0,71	0,89	1,11	1,40	1,86	1,93	2,00	2,09	2,19	2,31	2,45	2,63	2,90	3,36
9	0,70	0,88	1,10	1,38	1,83	1,90	1,97	2,06	2,15	2,26	2,40	2,57	2,82	3,25
10	0,70	0,88	1,09	1,37	1,81	1,88	1,95	2,03	2,12	2,23	2,36	2,53	2,76	3,17
11	0,70	0,88	1,09	1,36	1,80	1,86	1,93	2,01	2,10	2,20	2,33	2,49	2,72	3,11
12	0,70	0,87	1,08	1,36	1,78	1,84	1,91	1,99	2,08	2,18	2,30	2,46	2,68	3,05
13	0,69	0,87	1,08	1,35	1,77	1,83	1,90	1,97	2,06	2,16	2,28	2,44	2,65	3,01
14	0,69	0,87	1,08	1,35	1,76	1,82	1,89	1,96	2,05	2,14	2,26	2,41	2,62	2,98
15	0,69	0,87	1,07	1,34	1,75	1,81	1,88	1,95	2,03	2,13	2,25	2,40	2,60	2,95
16	0,69	0,86	1,07	1,34	1,75	1,80	1,87	1,94	2,02	2,12	2,24	2,38	2,58	2,92
17	0,69	0,86	1,07	1,33	1,74	1,80	1,86	1,93	2,02	2,11	2,22	2,37	2,57	2,90
18	0,69	0,86	1,07	1,33	1,73	1,79	1,86	1,93	2,01	2,10	2,21	2,36	2,55	2,88
19	0,69	0,86	1,07	1,33	1,73	1,79	1,85	1,92	2,00	2,09	2,20	2,35	2,54	2,86
20	0,69	0,86	1,06	1,33	1,72	1,78	1,84	1,91	1,99	2,09	2,20	2,34	2,53	2,85
22	0,69	0,86	1,06	1,32	1,72	1,77	1,84	1,90	1,98	2,07	2,18	2,32	2,51	2,82
24	0,68	0,86	1,06	1,32	1,71	1,77	1,83	1,90	1,97	2,06	2,17	2,31	2,49	2,80
26	0,68	0,86	1,06	1,31	1,71	1,76	1,82	1,89	1,97	2,06	2,16	2,30	2,48	2,78
28	0,68	0,85	1,06	1,31	1,70	1,76	1,82	1,88	1,96	2,05	2,15	2,29	2,47	2,76
30	0,68	0,85	1,05	1,31	1,70	1,75	1,81	1,88	1,95	2,04	2,15	2,28	2,46	2,75
32	0,68	0,85	1,05	1,31	1,69	1,75	1,81	1,87	1,95	2,04	2,14	2,27	2,45	2,74
34	0,68	0,85	1,05	1,31	1,69	1,75	1,80	1,87	1,95	2,03	2,14	2,27	2,44	2,73
36	0,68	0,85	1,05	1,31	1,69	1,74	1,80	1,87	1,94	2,03	2,13	2,26	2,43	2,72
38	0,68	0,85	1,05	1,30	1,69	1,74	1,80	1,86	1,94	2,02	2,13	2,25	2,43	2,71
40	0,68	0,85	1,05	1,30	1,68	1,74	1,80	1,86	1,94	2,02	2,12	2,25	2,42	2,70
45	0,68	0,85	1,05	1,30	1,68	1,73	1,79	1,86	1,93	2,01	2,12	2,24	2,41	2,69
50	0,68	0,85	1,05	1,30	1,68	1,73	1,79	1,85	1,92	2,01	2,11	2,23	2,40	2,68
55	0,68	0,85	1,05	1,30	1,67	1,73	1,78	1,85	1,92	2,00	2,10	2,23	2,40	2,67
60	0,68	0,85	1,05	1,30	1,67	1,72	1,78	1,84	1,92	2,00	2,10	2,22	2,39	2,66
65	0,68	0,85	1,04	1,29	1,67	1,72	1,78	1,84	1,91	2,00	2,10	2,22	2,39	2,65
70	0,68	0,85	1,04	1,29	1,67	1,72	1,78	1,84	1,91	1,99	2,09	2,22	2,38	2,65
75	0,68	0,85	1,04	1,29	1,67	1,72	1,77	1,84	1,91	1,99	2,09	2,21	2,38	2,64
80	0,68	0,85	1,04	1,29	1,66	1,72	1,77	1,84	1,91	1,99	2,09	2,21	2,37	2,64
85	0,68	0,85	1,04	1,29	1,66	1,71	1,77	1,84	1,91	1,99	2,09	2,21	2,37	2,63
90	0,68	0,85	1,04	1,29	1,66	1,71	1,77	1,83	1,90	1,99	2,08	2,21	2,37	2,63
95	0,68	0,85	1,04	1,29	1,66	1,71	1,77	1,83	1,90	1,99	2,08	2,20	2,37	2,63
100	0,68	0,85	1,04	1,29	1,66	1,71	1,77	1,83	1,90	1,98	2,08	2,20	2,36	2,63
110	0,68	0,84	1,04	1,29	1,66	1,71	1,77	1,83	1,90	1,98	2,08	2,20	2,36	2,62
120	0,68	0,84	1,04	1,29	1,66	1,71	1,77	1,83	1,90	1,98	2,08	2,20	2,36	2,62
130	0,68	0,84	1,04	1,29	1,66	1,71	1,76	1,83	1,90	1,98	2,07	2,19	2,36	2,61
140	0,68	0,84	1,04	1,29	1,66	1,71	1,76	1,83	1,90	1,98	2,07	2,19	2,35	2,61
150	0,68	0,84	1,04	1,29	1,66	1,71	1,76	1,82	1,90	1,98	2,07	2,19	2,35	2,61
160	0,68	0,84	1,04	1,29	1,65	1,71	1,76	1,82	1,89	1,97	2,07	2,19	2,35	2,61

Продовження таблиці A.2

$n \setminus p$	0,75	0,8	0,85	0,9	0,95	0,955	0,96	0,965	0,97	0,975	0,98	0,985	0,99	0,995
170	0,68	0,84	1,04	1,29	1,65	1,71	1,76	1,82	1,89	1,97	2,07	2,19	2,35	2,61
180	0,68	0,84	1,04	1,29	1,65	1,70	1,76	1,82	1,89	1,97	2,07	2,19	2,35	2,60
190	0,68	0,84	1,04	1,29	1,65	1,70	1,76	1,82	1,89	1,97	2,07	2,19	2,35	2,60
200	0,68	0,84	1,04	1,29	1,65	1,70	1,76	1,82	1,89	1,97	2,07	2,19	2,35	2,60
250	0,68	0,84	1,04	1,28	1,65	1,70	1,76	1,82	1,89	1,97	2,06	2,18	2,34	2,60
300	0,68	0,84	1,04	1,28	1,65	1,70	1,76	1,82	1,89	1,97	2,06	2,18	2,34	2,59
350	0,68	0,84	1,04	1,28	1,65	1,70	1,76	1,82	1,89	1,97	2,06	2,18	2,34	2,59
400	0,68	0,84	1,04	1,28	1,65	1,70	1,76	1,82	1,89	1,97	2,06	2,18	2,34	2,59
450	0,68	0,84	1,04	1,28	1,65	1,70	1,75	1,82	1,89	1,97	2,06	2,18	2,33	2,59
500	0,67	0,84	1,04	1,28	1,65	1,70	1,75	1,82	1,89	1,96	2,06	2,18	2,33	2,59
550	0,67	0,84	1,04	1,28	1,65	1,70	1,75	1,82	1,88	1,96	2,06	2,18	2,33	2,58

Квантилі $\chi^2_p(n)$ порядку p розподілу Пірсона з n ступенями вільності



Таблиця A.3

$n \setminus p$	,005	0,01	,015	0,02	,025	0,03	,035	0,04	,045	0,05	0,1	0,15	0,2	0,25
1	$,0_44$	$,0_32$	$,0_34$	$,0_36$	$,0_21$	$,0_21$	$,0_22$	$,0_23$	$,0_23$	$,0_24$	,016	,036	,064	,102
2	,010	,020	,030	,040	,051	,061	,071	,082	,092	,103	,211	,325	,446	,575
3	,072	,115	,152	,185	,216	,245	,273	,300	,326	,352	,584	,798	1,01	1,21
4	,207	,297	,368	,429	,484	,535	,582	,627	,670	,711	1,06	1,37	1,65	1,92
5	,412	,554	,662	,752	,831	,903	0,969	1,03	1,09	1,15	1,61	1,99	2,34	2,67
6	,676	,872	1,02	1,13	1,24	1,33	1,41	1,49	1,57	1,64	2,20	2,66	3,07	3,45
7	,989	1,24	1,42	1,56	1,69	1,80	1,90	2,00	2,08	2,17	2,83	3,36	3,82	4,25
8	1,34	1,65	1,86	2,03	2,18	2,31	2,43	2,54	2,64	2,73	3,49	4,08	4,59	5,07
9	1,73	2,09	2,33	2,53	2,70	2,85	2,98	3,10	3,22	3,33	4,17	4,82	5,38	5,90
10	2,16	2,56	2,84	3,06	3,25	3,41	3,56	3,70	3,82	3,94	4,87	5,57	6,18	6,74

Продовження таблиці А.3

$n \setminus p$	0,005	0,01	0,015	0,02	0,025	0,03	0,035	0,04	0,045	0,05	0,1	0,15	0,2	0,25
11	2,60	3,05	3,36	3,61	3,82	4,00	4,16	4,31	4,45	4,57	5,58	6,34	6,99	7,58
12	3,07	3,57	3,91	4,18	4,40	4,60	4,78	4,94	5,09	5,23	6,30	7,11	7,81	8,44
13	3,57	4,11	4,48	4,77	5,01	5,22	5,41	5,58	5,74	5,89	7,04	7,90	8,63	9,30
14	4,07	4,66	5,06	5,37	5,63	5,86	6,06	6,24	6,41	6,57	7,79	8,70	9,47	10,2
15	4,60	5,23	5,65	5,98	6,26	6,50	6,72	6,91	7,09	7,26	8,55	9,50	10,3	11,0
16	5,14	5,81	6,26	6,61	6,91	7,16	7,39	7,60	7,79	7,96	9,31	10,3	11,2	11,9
17	5,70	6,41	6,88	7,25	7,56	7,83	8,07	8,29	8,49	8,67	10,1	11,1	12,0	12,8
18	6,26	7,01	7,52	7,91	8,23	8,51	8,76	8,99	9,20	9,39	10,9	11,9	12,9	13,7
19	6,84	7,63	8,16	8,57	8,91	9,20	9,46	9,70	9,92	10,1	11,7	12,8	13,7	14,6
20	7,43	8,26	8,81	9,24	9,59	9,90	10,2	10,4	10,6	10,9	12,4	13,6	14,6	15,5
21	8,03	8,90	9,47	9,91	10,3	10,6	10,9	11,1	11,4	11,6	13,2	14,4	15,4	16,3
22	8,64	9,54	10,1	10,6	11,0	11,3	11,6	11,9	12,1	12,3	14,0	15,3	16,3	17,2
23	9,26	10,2	10,8	11,3	11,7	12,0	12,3	12,6	12,9	13,1	14,8	16,1	17,2	18,1
24	9,89	10,9	11,5	12,0	12,4	12,8	13,1	13,3	13,6	13,8	15,7	17,0	18,1	19,0
25	10,5	11,5	12,2	12,7	13,1	13,5	13,8	14,1	14,4	14,6	16,5	17,8	18,9	19,9
26	11,2	12,2	12,9	13,4	13,8	14,2	14,6	14,9	15,1	15,4	17,3	18,7	19,8	20,8
27	11,8	12,9	13,6	14,1	14,6	15,0	15,3	15,6	15,9	16,2	18,1	19,5	20,7	21,7
28	12,5	13,6	14,3	14,8	15,3	15,7	16,1	16,4	16,7	16,9	18,9	20,4	21,6	22,7
29	13,1	14,3	15,0	15,6	16,0	16,5	16,8	17,1	17,4	17,7	19,8	21,2	22,5	23,6
30	13,8	15,0	15,7	16,3	16,8	17,2	17,6	17,9	18,2	18,5	20,6	22,1	23,4	24,5
31	14,5	15,7	16,4	17,0	17,5	18,0	18,3	18,7	19,0	19,3	21,4	23,0	24,3	25,4
32	15,1	16,4	17,2	17,8	18,3	18,7	19,1	19,5	19,8	20,1	22,3	23,8	25,1	26,3
33	15,8	17,1	17,9	18,5	19,0	19,5	19,9	20,2	20,6	20,9	23,1	24,7	26,0	27,2
34	16,5	17,8	18,6	19,3	19,8	20,3	20,7	21,0	21,4	21,7	24,0	25,6	26,9	28,1
35	17,2	18,5	19,4	20,0	20,6	21,0	21,4	21,8	22,2	22,5	24,8	26,5	27,8	29,1
36	17,9	19,2	20,1	20,8	21,3	21,8	22,2	22,6	23,0	23,3	25,6	27,3	28,7	30,0
37	18,6	20,0	20,9	21,5	22,1	22,6	23,0	23,4	23,8	24,1	26,5	28,2	29,6	30,9
38	19,3	20,7	21,6	22,3	22,9	23,4	23,8	24,2	24,6	24,9	27,3	29,1	30,5	31,8
39	20,0	21,4	22,4	23,1	23,7	24,2	24,6	25,0	25,4	25,7	28,2	30,0	31,4	32,7
40	20,7	22,2	23,1	23,8	24,4	24,9	25,4	25,8	26,2	26,5	29,1	30,9	32,3	33,7
41	21,4	22,9	23,9	24,6	25,2	25,7	26,2	26,6	27,0	27,3	29,9	31,7	33,3	34,6
42	22,1	23,7	24,6	25,4	26,0	26,5	27,0	27,4	27,8	28,1	30,8	32,6	34,2	35,5
43	22,9	24,4	25,4	26,2	26,8	27,3	27,8	28,2	28,6	29,0	31,6	33,5	35,1	36,4
44	23,6	25,1	26,2	26,9	27,6	28,1	28,6	29,0	29,4	29,8	32,5	34,4	36,0	37,4
45	24,3	25,9	26,9	27,7	28,4	28,9	29,4	29,8	30,2	30,6	33,4	35,3	36,9	38,3
46	25,0	26,7	27,7	28,5	29,2	29,7	30,2	30,7	31,1	31,4	34,2	36,2	37,8	39,2
47	25,8	27,4	28,5	29,3	30,0	30,5	31,0	31,5	31,9	32,3	35,1	37,1	38,7	40,1
48	26,5	28,2	29,3	30,1	30,8	31,3	31,8	32,3	32,7	33,1	35,9	38,0	39,6	41,1
49	27,2	28,9	30,0	30,9	31,6	32,1	32,7	33,1	33,5	33,9	36,8	38,9	40,5	42,0
50	28,0	29,7	30,8	31,7	32,4	33,0	33,5	33,9	34,4	34,8	37,7	39,8	41,4	42,9
55	31,7	33,6	34,8	35,7	36,4	37,0	37,6	38,1	38,5	39,0	42,1	44,2	46,0	47,6
60	35,5	37,5	38,7	39,7	40,5	41,2	41,7	42,3	42,7	43,2	46,5	48,8	50,6	52,3
65	39,4	41,4	42,8	43,8	44,6	45,3	45,9	46,5	47,0	47,4	50,9	53,3	55,3	57,0
70	43,3	45,4	46,8	47,9	48,8	49,5	50,1	50,7	51,3	51,7	55,3	57,8	59,9	61,7
75	47,2	49,5	50,9	52,0	52,9	53,7	54,4	55,0	55,5	56,1	59,8	62,4	64,5	66,4
80	51,2	53,5	55,1	56,2	57,2	58,0	58,7	59,3	59,9	60,4	64,3	67,0	69,2	71,1
85	55,2	57,6	59,2	60,4	61,4	62,2	63,0	63,6	64,2	64,7	68,8	71,6	73,9	75,9
90	59,2	61,8	63,4	64,6	65,6	66,5	67,3	67,9	68,6	69,1	73,3	76,2	78,6	80,6

Продовження таблиці А.3

$n \setminus p$	0,005	0,01	0,015	0,02	0,025	0,03	0,035	0,04	0,045	0,05	0,1	0,15	0,2	0,25
95	63,2	65,9	67,6	68,9	69,9	70,8	71,6	72,3	72,9	73,5	77,8	80,8	83,2	85,4
100	67,3	70,1	71,8	73,1	74,2	75,1	75,9	76,7	77,3	77,9	82,4	85,4	87,9	90,1
110	75,5	78,5	80,3	81,7	82,9	83,8	84,7	85,5	86,2	86,8	91,5	94,7	97,4	99,7
120	83,9	86,9	88,9	90,4	91,6	92,6	93,5	94,3	95,0	95,7	101	104	107	109
130	92,2	95,5	97,5	99,1	100	101	102	103	104	105	110	113	116	119
140	101	104	106	108	109	110	111	112	113	114	119	123	126	128
150	109	113	115	117	118	119	120	121	122	123	128	132	135	138
160	118	121	124	125	127	128	129	130	131	132	138	142	145	148
170	126	130	132	134	136	137	138	139	140	141	147	151	154	157
180	135	139	141	143	145	146	147	148	149	150	156	160	164	167
190	144	148	150	152	154	155	156	157	158	159	165	170	173	177
200	152	156	159	161	163	164	165	166	167	168	175	179	183	186
210	161	165	168	170	172	173	174	176	177	177	184	189	193	196
220	170	174	177	179	181	182	184	185	186	187	194	198	202	206
230	179	183	186	188	190	191	193	194	195	196	203	208	212	215
240	187	192	195	197	199	201	202	203	204	205	212	217	221	225
250	196	201	204	206	208	210	211	212	213	214	222	227	231	235
260	205	210	213	215	217	219	220	221	223	224	231	236	241	244
270	214	219	222	224	226	228	229	231	232	233	241	246	250	254
280	223	228	231	234	236	237	239	240	241	242	250	256	260	264
290	232	237	240	243	245	246	248	249	250	252	260	265	270	273
300	241	246	249	252	254	256	257	259	260	261	269	275	279	283
350	286	291	295	298	300	302	304	305	306	308	317	323	328	332
400	331	337	341	344	346	349	350	352	353	355	364	371	376	381
450	376	383	387	391	393	395	397	399	400	402	412	419	425	429
500	422	429	434	437	440	442	444	446	448	449	460	467	473	478

Продовження таблиці А.3

$n \setminus p$	,995	,99	,985	,98	,975	,97	,965	,96	,955	,95	,9	,85	,8	,75
1	7,88	6,63	5,92	5,41	5,02	4,71	4,45	4,22	4,02	3,84	2,71	2,07	1,64	1,32
2	10,60	9,21	8,40	7,82	7,38	7,01	6,70	6,44	6,20	5,99	4,61	3,79	3,22	2,77
3	12,8	11,3	10,5	9,84	9,35	8,95	8,61	8,31	8,05	7,81	6,25	5,32	4,64	4,11
4	14,9	13,3	12,3	11,7	11,1	10,7	10,4	10,0	9,74	9,49	7,78	6,74	5,99	5,39
5	16,8	15,1	14,1	13,4	12,8	12,4	12,0	11,6	11,3	11,1	9,24	8,12	7,29	6,63
6	18,6	16,8	15,8	15,0	14,5	14,0	13,6	13,2	12,9	12,6	10,6	9,45	8,56	7,84
7	20,3	18,5	17,4	16,6	16,0	15,5	15,1	14,7	14,4	14,1	12,0	10,8	9,80	9,04
8	22,0	20,1	19,0	18,2	17,5	17,0	16,6	16,2	15,8	15,5	13,4	12,0	11,0	10,2
9	23,6	21,7	20,5	19,7	19,0	18,5	18,0	17,6	17,2	16,9	14,7	13,3	12,2	11,4
10	25,2	23,2	22,0	21,2	20,5	19,9	19,4	19,0	18,6	18,3	16,0	14,5	13,4	12,5
11	26,8	24,7	23,5	22,6	21,9	21,3	20,8	20,4	20,0	19,7	17,3	15,8	14,6	13,7
12	28,3	26,2	25,0	24,1	23,3	22,7	22,2	21,8	21,4	21,0	18,5	17,0	15,8	14,8
13	29,8	27,7	26,4	25,5	24,7	24,1	23,6	23,1	22,7	22,4	19,8	18,2	17,0	16,0
14	31,3	29,1	27,8	26,9	26,1	25,5	25,0	24,5	24,1	23,7	21,1	19,4	18,2	17,1

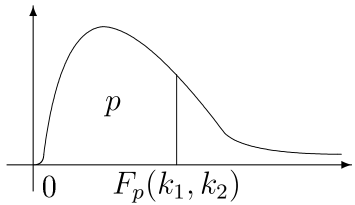
Продовження таблиці А.3

$n \setminus p$	,995	,99	,985	,98	,975	,97	,965	,96	,955	,95	,9	,85	,8	,75
15	32,8	30,6	29,2	28,3	27,5	26,8	26,3	25,8	25,4	25,0	22,3	20,6	19,3	18,2
16	34,3	32,0	30,6	29,6	28,8	28,2	27,6	27,1	26,7	26,3	23,5	21,8	20,5	19,4
17	35,7	33,4	32,0	31,0	30,2	29,5	28,9	28,4	28,0	27,6	24,8	23,0	21,6	20,5
18	37,2	34,8	33,4	32,3	31,5	30,8	30,3	29,7	29,3	28,9	26,0	24,2	22,8	21,6
19	38,6	36,2	34,7	33,7	32,9	32,2	31,6	31,0	30,6	30,1	27,2	25,3	23,9	22,7
20	40,0	37,6	36,1	35,0	34,2	33,5	32,9	32,3	31,8	31,4	28,4	26,5	25,0	23,8
21	41,4	38,9	37,4	36,3	35,5	34,8	34,1	33,6	33,1	32,7	29,6	27,7	26,2	24,9
22	42,8	40,3	38,8	37,7	36,8	36,0	35,4	34,9	34,4	33,9	30,8	28,8	27,3	26,0
23	44,2	41,6	40,1	39,0	38,1	37,3	36,7	36,1	35,6	35,2	32,0	30,0	28,4	27,1
24	45,6	43,0	41,4	40,3	39,4	38,6	38,0	37,4	36,9	36,4	33,2	31,1	29,6	28,2
25	46,9	44,3	42,7	41,6	40,6	39,9	39,2	38,6	38,1	37,7	34,4	32,3	30,7	29,3
26	48,3	45,6	44,0	42,9	41,9	41,1	40,5	39,9	39,4	38,9	35,6	33,4	31,8	30,4
27	49,6	47,0	45,3	44,1	43,2	42,4	41,7	41,1	40,6	40,1	36,7	34,6	32,9	31,5
28	51,0	48,3	46,6	45,4	44,5	43,7	43,0	42,4	41,8	41,3	37,9	35,7	34,0	32,6
29	52,3	49,6	47,9	46,7	45,7	44,9	44,2	43,6	43,1	42,6	39,1	36,9	35,1	33,7
30	53,7	50,9	49,2	48,0	47,0	46,2	45,5	44,8	44,3	43,8	40,3	38,0	36,3	34,8
31	55,0	52,2	50,5	49,2	48,2	47,4	46,7	46,1	45,5	45,0	41,4	39,1	37,4	35,9
32	56,3	53,5	51,8	50,5	49,5	48,6	47,9	47,3	46,7	46,2	42,6	40,3	38,5	37,0
33	57,6	54,8	53,0	51,7	50,7	49,9	49,1	48,5	47,9	47,4	43,7	41,4	39,6	38,1
34	59,0	56,1	54,3	53,0	52,0	51,1	50,4	49,7	49,1	48,6	44,9	42,5	40,7	39,1
35	60,3	57,3	55,6	54,2	53,2	52,3	51,6	50,9	50,3	49,8	46,1	43,6	41,8	40,2
36	61,6	58,6	56,8	55,5	54,4	53,6	52,8	52,1	51,5	51,0	47,2	44,8	42,9	41,3
37	62,9	59,9	58,1	56,7	55,7	54,8	54,0	53,3	52,7	52,2	48,4	45,9	44,0	42,4
38	64,2	61,2	59,3	58,0	56,9	56,0	55,2	54,5	53,9	53,4	49,5	47,0	45,1	43,5
39	65,5	62,4	60,6	59,2	58,1	57,2	56,4	55,7	55,1	54,6	50,7	48,1	46,2	44,5
40	66,8	63,7	61,8	60,4	59,3	58,4	57,6	56,9	56,3	55,8	51,8	49,2	47,3	45,6
41	68,1	64,9	63,1	61,7	60,6	59,6	58,8	58,1	57,5	56,9	52,9	50,4	48,4	46,7
42	69,3	66,2	64,3	62,9	61,8	60,8	60,0	59,3	58,7	58,1	54,1	51,5	49,5	47,8
43	70,6	67,5	65,5	64,1	63,0	62,1	61,2	60,5	59,9	59,3	55,2	52,6	50,5	48,8
44	71,9	68,7	66,8	65,3	64,2	63,3	62,4	61,7	61,1	60,5	56,4	53,7	51,6	49,9
45	73,2	70,0	68,0	66,6	65,4	64,5	63,6	62,9	62,2	61,7	57,5	54,8	52,7	51,0
46	74,4	71,2	69,2	67,8	66,6	65,7	64,8	64,1	63,4	62,8	58,6	55,9	53,8	52,1
47	75,7	72,4	70,4	69,0	67,8	66,8	66,0	65,3	64,6	64,0	59,8	57,0	54,9	53,1
48	77,0	73,7	71,7	70,2	69,0	68,0	67,2	66,4	65,8	65,2	60,9	58,1	56,0	54,2
49	78,2	74,9	72,9	71,4	70,2	69,2	68,4	67,6	67,0	66,3	62,0	59,2	57,1	55,3
50	79,5	76,2	74,1	72,6	71,4	70,4	69,6	68,8	68,1	67,5	63,2	60,3	58,2	56,3
55	85,7	82,3	80,2	78,6	77,4	76,3	75,5	74,7	74,0	73,3	68,8	65,9	63,6	61,7
60	92,0	88,4	86,2	84,6	83,3	82,2	81,3	80,5	79,7	79,1	74,4	71,3	69,0	67,0
65	98,1	94,4	92,2	90,5	89,2	88,1	87,1	86,3	85,5	84,8	80,0	76,8	74,4	72,3
70	104	100	98,1	96,4	95,0	93,9	92,9	92,0	91,2	90,5	85,5	82,3	79,7	77,6
75	110	106	104	102	101	99,7	98,6	97,8	96,9	96,2	91,1	87,7	85,1	82,9
80	116	112	110	108	107	105	104	103	103	102	96,6	93,1	90,4	88,1
85	122	118	116	114	112	111	110	109	108	108	102	98,5	95,7	93,4
90	128	124	122	120	118	117	116	115	114	113	108	104	101	98,6
95	134	130	127	125	124	123	121	120	120	119	113	109	106	104
100	140	136	133	131	130	128	127	126	125	124	118	115	112	109
110	152	147	145	143	141	140	138	137	136	135	129	125	122	120

Продовження таблиці А.3

$n \setminus p$	,995	,99	,985	,98	,975	,97	,965	,96	,955	,95	,9	,85	,8	,75
120	164	159	156	154	152	151	150	148	147	147	140	136	133	130
130	175	170	167	165	163	162	161	160	159	158	151	147	143	140
140	187	182	179	176	175	173	172	171	170	169	162	157	154	151
150	198	193	190	188	186	184	183	182	181	180	173	168	164	161
160	210	205	201	199	197	195	194	193	192	191	183	179	175	172
170	221	216	212	210	208	206	205	204	202	201	194	189	185	182
180	233	227	224	221	219	217	216	215	213	212	205	200	196	192
190	244	238	235	232	230	228	227	225	224	223	215	210	206	203
200	255	249	246	243	241	239	238	236	235	234	226	221	217	213
210	267	261	257	254	252	250	249	247	246	245	237	231	227	223
220	278	272	268	265	263	261	259	258	257	256	247	242	237	234
230	289	283	279	276	274	272	270	269	268	266	258	252	248	244
240	300	294	290	287	285	283	281	280	278	277	268	263	258	254
250	311	305	301	298	296	294	292	290	289	288	279	273	269	265
260	322	316	312	309	307	305	303	301	300	299	290	284	279	275
270	334	327	323	320	317	315	314	312	311	309	300	294	289	285
280	345	338	334	331	328	326	324	323	321	320	311	305	300	296
290	356	349	345	342	339	337	335	334	332	331	321	315	310	306
300	367	360	356	352	350	348	346	344	343	341	332	325	320	316
350	422	414	410	406	404	401	399	398	396	395	384	377	372	367
400	477	469	464	460	457	455	453	451	449	448	437	429	424	419
450	531	523	518	514	511	508	506	504	502	500	489	481	475	470
500	585	576	571	567	564	561	559	557	555	553	541	533	526	521

Квантилі $F_p(k_1, k_2)$ порядку p розподілу Фішера з k_1 і k_2 ступенями вільності



Таблиця А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,9$																	
1	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,19	60,71	61,22	61,74	62,05	62,26	62,53	62,79	63,06
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,41	9,42	9,44	9,45	9,46	9,47	9,47	9,48
3	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,22	5,20	5,18	5,17	5,17	5,16	5,15	5,14
4	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,90	3,87	3,84	3,83	3,82	3,80	3,79	3,78
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,27	3,24	3,21	3,19	3,17	3,16	3,14	3,12
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,90	2,87	2,84	2,81	2,80	2,78	2,76	2,74
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,67	2,63	2,59	2,57	2,56	2,54	2,51	2,49
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,50	2,46	2,42	2,40	2,38	2,36	2,34	2,32
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,38	2,34	2,30	2,27	2,25	2,23	2,21	2,18
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,28	2,24	2,20	2,17	2,16	2,13	2,11	2,08
11	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25	2,21	2,17	2,12	2,10	2,08	2,05	2,03	2,00
12	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21	2,19	2,15	2,10	2,06	2,03	2,01	1,99	1,96	1,93
13	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16	2,14	2,10	2,05	2,01	1,98	1,96	1,93	1,90	1,88

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,9$																	
14	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15	2,12	2,10	2,05	2,01	1,96	1,93	1,91	1,89	1,86	1,83
15	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06	2,02	1,97	1,92	1,89	1,87	1,85	1,82	1,79
16	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06	2,03	1,99	1,94	1,89	1,86	1,84	1,81	1,78	1,75
17	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06	2,03	2,00	1,96	1,91	1,86	1,83	1,81	1,78	1,75	1,72
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00	1,98	1,93	1,89	1,84	1,80	1,78	1,75	1,72	1,69
19	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98	1,96	1,91	1,86	1,81	1,78	1,76	1,73	1,70	1,67
20	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94	1,89	1,84	1,79	1,76	1,74	1,71	1,68	1,64
21	2,96	2,57	2,36	2,23	2,14	2,08	2,02	1,98	1,95	1,92	1,87	1,83	1,78	1,74	1,72	1,69	1,66	1,62
22	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97	1,93	1,90	1,86	1,81	1,76	1,73	1,70	1,67	1,64	1,60
23	2,94	2,55	2,34	2,21	2,11	2,05	1,99	1,95	1,92	1,89	1,84	1,80	1,74	1,71	1,69	1,66	1,62	1,59
24	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94	1,91	1,88	1,83	1,78	1,73	1,70	1,67	1,64	1,61	1,57
25	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,89	1,87	1,82	1,77	1,72	1,68	1,66	1,63	1,59	1,56
26	2,91	2,52	2,31	2,17	2,08	2,01	1,96	1,92	1,88	1,86	1,81	1,76	1,71	1,67	1,65	1,61	1,58	1,54
27	2,90	2,51	2,30	2,17	2,07	2,00	1,95	1,91	1,87	1,85	1,80	1,75	1,70	1,66	1,64	1,60	1,57	1,53
28	2,89	2,50	2,29	2,16	2,06	2,00	1,94	1,90	1,87	1,84	1,79	1,74	1,69	1,65	1,63	1,59	1,56	1,52
29	2,89	2,50	2,28	2,15	2,06	1,99	1,93	1,89	1,86	1,83	1,78	1,73	1,68	1,64	1,62	1,58	1,55	1,51
30	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82	1,77	1,72	1,67	1,63	1,61	1,57	1,54	1,50
40	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76	1,71	1,66	1,61	1,57	1,54	1,51	1,47	1,42
50	2,81	2,41	2,20	2,06	1,97	1,90	1,84	1,80	1,76	1,73	1,68	1,63	1,57	1,53	1,50	1,46	1,42	1,38
75	2,77	2,37	2,16	2,02	1,93	1,85	1,80	1,75	1,72	1,69	1,63	1,58	1,52	1,48	1,45	1,41	1,37	1,32
100	2,76	2,36	2,14	2,00	1,91	1,83	1,78	1,73	1,69	1,66	1,61	1,56	1,49	1,45	1,42	1,38	1,34	1,28
150	2,74	2,34	2,12	1,98	1,89	1,81	1,76	1,71	1,67	1,64	1,59	1,53	1,47	1,43	1,40	1,35	1,30	1,25

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,95$																	
1	161,5	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	246,0	248,0	249,3	250,1	251,1	252,2	253,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,46	19,46	19,47	19,48	19,49
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,63	8,62	8,59	8,57	8,55
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,52	4,50	4,46	4,43	4,40
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,83	3,81	3,77	3,74	3,70
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,40	3,38	3,34	3,30	3,27
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,11	3,08	3,04	3,01	2,97
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,89	2,86	2,83	2,79	2,75
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,73	2,70	2,66	2,62	2,58
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,60	2,57	2,53	2,49	2,45
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,50	2,47	2,43	2,38	2,34
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,41	2,38	2,34	2,30	2,25
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,34	2,31	2,27	2,22	2,18
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,28	2,25	2,20	2,16	2,11
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,23	2,19	2,15	2,11	2,06
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,18	2,15	2,10	2,06	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,14	2,11	2,06	2,02	1,97
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,07	2,04	1,99	1,95	1,90
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,95$																	
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,02	1,98	1,94	1,89	1,84
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,00	1,96	1,91	1,86	1,81
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,97	1,94	1,89	1,84	1,79
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,94	1,90	1,85	1,80	1,75
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,92	1,88	1,84	1,79	1,73
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,89	1,85	1,81	1,75	1,70
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,88	1,84	1,79	1,74	1,68
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,78	1,74	1,69	1,64	1,58
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,95	1,87	1,78	1,73	1,69	1,63	1,58	1,51
75	3,97	3,12	2,73	2,49	2,34	2,22	2,13	2,06	2,01	1,96	1,88	1,80	1,71	1,65	1,61	1,55	1,49	1,42
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,85	1,77	1,68	1,62	1,57	1,52	1,45	1,38
150	3,90	3,06	2,66	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,82	1,73	1,64	1,58	1,54	1,48	1,41	1,33

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,975$																	
1	647,8	799,5	864,2	899,6	921,9	937,1	948,2	956,7	963,3	968,6	976,7	984,9	993,1	998,1	1001,	1006,	1010,	1014,
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,43	39,45	39,46	39,46	39,47	39,48	39,49
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,34	14,25	14,17	14,12	14,08	14,04	13,99	13,95
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,75	8,66	8,56	8,50	8,46	8,41	8,36	8,31
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,52	6,43	6,33	6,27	6,23	6,18	6,12	6,07

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,975$																	
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,37	5,27	5,17	5,11	5,07	5,01	4,96	4,90
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,67	4,57	4,47	4,40	4,36	4,31	4,25	4,20
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,20	4,10	4,00	3,94	3,89	3,84	3,78	3,73
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,87	3,77	3,67	3,60	3,56	3,51	3,45	3,39
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,62	3,52	3,42	3,35	3,31	3,26	3,20	3,14
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,43	3,33	3,23	3,16	3,12	3,06	3,00	2,94
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,28	3,18	3,07	3,01	2,96	2,91	2,85	2,79
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25	3,15	3,05	2,95	2,88	2,84	2,78	2,72	2,66
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15	3,05	2,95	2,84	2,78	2,73	2,67	2,61	2,55
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,96	2,86	2,76	2,69	2,64	2,59	2,52	2,46
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99	2,89	2,79	2,68	2,61	2,57	2,51	2,45	2,38
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,98	2,92	2,82	2,72	2,62	2,55	2,50	2,44	2,38	2,32
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,87	2,77	2,67	2,56	2,49	2,44	2,38	2,32	2,26
19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,88	2,82	2,72	2,62	2,51	2,44	2,39	2,33	2,27	2,20
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,68	2,57	2,46	2,40	2,35	2,29	2,22	2,16
21	5,83	4,42	3,82	3,48	3,25	3,09	2,97	2,87	2,80	2,73	2,64	2,53	2,42	2,36	2,31	2,25	2,18	2,11
22	5,79	4,38	3,78	3,44	3,22	3,05	2,93	2,84	2,76	2,70	2,60	2,50	2,39	2,32	2,27	2,21	2,14	2,08
23	5,75	4,35	3,75	3,41	3,18	3,02	2,90	2,81	2,73	2,67	2,57	2,47	2,36	2,29	2,24	2,18	2,11	2,04
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64	2,54	2,44	2,33	2,26	2,21	2,15	2,08	2,01
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61	2,51	2,41	2,30	2,23	2,18	2,12	2,05	1,98
26	5,66	4,27	3,67	3,33	3,10	2,94	2,82	2,73	2,65	2,59	2,49	2,39	2,28	2,21	2,16	2,09	2,03	1,95
27	5,63	4,24	3,65	3,31	3,08	2,92	2,80	2,71	2,63	2,57	2,47	2,36	2,25	2,18	2,13	2,07	2,00	1,93
28	5,61	4,22	3,63	3,29	3,06	2,90	2,78	2,69	2,61	2,55	2,45	2,34	2,23	2,16	2,11	2,05	1,98	1,91

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,975$																	
29	5,59	4,20	3,61	3,27	3,04	2,88	2,76	2,67	2,59	2,53	2,43	2,32	2,21	2,14	2,09	2,03	1,96	1,89
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,57	2,51	2,41	2,31	2,20	2,12	2,07	2,01	1,94	1,87
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39	2,29	2,18	2,07	1,99	1,94	1,88	1,80	1,72
50	5,34	3,97	3,39	3,05	2,83	2,67	2,55	2,46	2,38	2,32	2,22	2,11	1,99	1,92	1,87	1,80	1,72	1,64
75	5,23	3,88	3,30	2,96	2,74	2,58	2,46	2,37	2,29	2,22	2,12	2,01	1,90	1,82	1,76	1,69	1,61	1,52
100	5,18	3,83	3,25	2,92	2,70	2,54	2,42	2,32	2,24	2,18	2,08	1,97	1,85	1,77	1,71	1,64	1,56	1,46
150	5,13	3,78	3,20	2,87	2,65	2,49	2,37	2,28	2,20	2,13	2,03	1,92	1,80	1,72	1,67	1,59	1,50	1,40

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,99$																	
1	4052,	5000,	5403,	5625,	5764,	5859,	5928,	5981,	6023,	6056,	6106,	6157,	6209,	6240,	6261,	6287,	6313,	6339,
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,58	26,50	26,41	26,32	26,22
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,91	13,84	13,75	13,65	13,56
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,45	9,38	9,29	9,20	9,11
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,30	7,23	7,14	7,06	6,97
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,06	5,99	5,91	5,82	5,74
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,26	5,20	5,12	5,03	4,95
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,71	4,65	4,57	4,48	4,40
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,31	4,25	4,17	4,08	4,00
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,01	3,94	3,86	3,78	3,69
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,76	3,70	3,62	3,54	3,45

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,99$																	
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,57	3,51	3,43	3,34	3,25
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,41	3,35	3,27	3,18	3,09
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,28	3,21	3,13	3,05	2,96
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,16	3,10	3,02	2,93	2,84
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,07	3,00	2,92	2,83	2,75
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	2,98	2,92	2,84	2,75	2,66
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,91	2,84	2,76	2,67	2,58
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,84	2,78	2,69	2,61	2,52
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,79	2,72	2,64	2,55	2,46
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,73	2,67	2,58	2,50	2,40
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,69	2,62	2,54	2,45	2,35
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,64	2,58	2,49	2,40	2,31
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,60	2,54	2,45	2,36	2,27
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,57	2,50	2,42	2,33	2,23
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,54	2,47	2,38	2,29	2,20
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,51	2,44	2,35	2,26	2,17
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,48	2,41	2,33	2,23	2,14
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,45	2,39	2,30	2,21	2,11
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,27	2,20	2,11	2,02	1,92
50	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,78	2,70	2,56	2,42	2,27	2,17	2,10	2,01	1,91	1,80
75	6,99	4,90	4,05	3,58	3,27	3,05	2,89	2,76	2,65	2,57	2,43	2,29	2,13	2,03	1,96	1,87	1,76	1,65
100	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59	2,50	2,37	2,22	2,07	1,97	1,89	1,80	1,69	1,57
150	6,81	4,75	3,91	3,45	3,14	2,92	2,76	2,63	2,53	2,44	2,31	2,16	2,00	1,90	1,83	1,73	1,62	1,49

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,995$																	
1	1621†	2000†	2162†	2250†	2306†	2344†	2372†	2393†	2409†	2422†	2443†	2463†	2484†	2496†	2504†	2515†	2525†	2536†
2	198,5	199,0	199,2	199,3	199,3	199,3	199,4	199,4	199,4	199,4	199,4	199,4	199,5	199,5	199,5	199,5	199,5	199,5
3	55,55	49,80	47,47	46,19	45,39	44,84	44,43	44,13	43,88	43,69	43,39	43,08	42,78	42,59	42,47	42,31	42,15	41,99
4	31,33	26,28	24,26	23,15	22,46	21,97	21,62	21,35	21,14	20,97	20,70	20,44	20,17	20,00	19,89	19,75	19,61	19,47
5	22,78	18,31	16,53	15,56	14,94	14,51	14,20	13,96	13,77	13,62	13,38	13,15	12,90	12,76	12,66	12,53	12,40	12,27
6	18,63	14,54	12,92	12,03	11,46	11,07	10,79	10,57	10,39	10,25	10,03	9,81	9,59	9,45	9,36	9,24	9,12	9,00
7	16,24	12,40	10,88	10,05	9,52	9,16	8,89	8,68	8,51	8,38	8,18	7,97	7,75	7,62	7,53	7,42	7,31	7,19
8	14,69	11,04	9,60	8,81	8,30	7,95	7,69	7,50	7,34	7,21	7,01	6,81	6,61	6,48	6,40	6,29	6,18	6,06
9	13,61	10,11	8,72	7,96	7,47	7,13	6,88	6,69	6,54	6,42	6,23	6,03	5,83	5,71	5,62	5,52	5,41	5,30
10	12,83	9,43	8,08	7,34	6,87	6,54	6,30	6,12	5,97	5,85	5,66	5,47	5,27	5,15	5,07	4,97	4,86	4,75
11	12,23	8,91	7,60	6,88	6,42	6,10	5,86	5,68	5,54	5,42	5,24	5,05	4,86	4,74	4,65	4,55	4,45	4,34
12	11,75	8,51	7,23	6,52	6,07	5,76	5,52	5,35	5,20	5,09	4,91	4,72	4,53	4,41	4,33	4,23	4,12	4,01
13	11,37	8,19	6,93	6,23	5,79	5,48	5,25	5,08	4,94	4,82	4,64	4,46	4,27	4,15	4,07	3,97	3,87	3,76
14	11,06	7,92	6,68	6,00	5,56	5,26	5,03	4,86	4,72	4,60	4,43	4,25	4,06	3,94	3,86	3,76	3,66	3,55
15	10,80	7,70	6,48	5,80	5,37	5,07	4,85	4,67	4,54	4,42	4,25	4,07	3,88	3,77	3,69	3,58	3,48	3,37
16	10,58	7,51	6,30	5,64	5,21	4,91	4,69	4,52	4,38	4,27	4,10	3,92	3,73	3,62	3,54	3,44	3,33	3,22
17	10,38	7,35	6,16	5,50	5,07	4,78	4,56	4,39	4,25	4,14	3,97	3,79	3,61	3,49	3,41	3,31	3,21	3,10
18	10,22	7,21	6,03	5,37	4,96	4,66	4,44	4,28	4,14	4,03	3,86	3,68	3,50	3,38	3,30	3,20	3,10	2,99
19	10,07	7,09	5,92	5,27	4,85	4,56	4,34	4,18	4,04	3,93	3,76	3,59	3,40	3,29	3,21	3,11	3,00	2,89
20	9,94	6,99	5,82	5,17	4,76	4,47	4,26	4,09	3,96	3,85	3,68	3,50	3,32	3,20	3,12	3,02	2,92	2,81
21	9,83	6,89	5,73	5,09	4,68	4,39	4,18	4,01	3,88	3,77	3,60	3,43	3,24	3,13	3,05	2,95	2,84	2,73
22	9,73	6,81	5,65	5,02	4,61	4,32	4,11	3,94	3,81	3,70	3,54	3,36	3,18	3,06	2,98	2,88	2,77	2,66
23	9,63	6,73	5,58	4,95	4,54	4,26	4,05	3,88	3,75	3,64	3,47	3,30	3,12	3,00	2,92	2,82	2,71	2,60
24	9,55	6,66	5,52	4,89	4,49	4,20	3,99	3,83	3,69	3,59	3,42	3,25	3,06	2,95	2,87	2,77	2,66	2,55
25	9,48	6,60	5,46	4,84	4,43	4,15	3,94	3,78	3,64	3,54	3,37	3,20	3,01	2,90	2,82	2,72	2,61	2,50

†означає $\times 10$

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,995$																	
26	9,41	6,54	5,41	4,79	4,38	4,10	3,89	3,73	3,60	3,49	3,33	3,15	2,97	2,85	2,77	2,67	2,56	2,45
27	9,34	6,49	5,36	4,74	4,34	4,06	3,85	3,69	3,56	3,45	3,28	3,11	2,93	2,81	2,73	2,63	2,52	2,41
28	9,28	6,44	5,32	4,70	4,30	4,02	3,81	3,65	3,52	3,41	3,25	3,07	2,89	2,77	2,69	2,59	2,48	2,37
29	9,23	6,40	5,28	4,66	4,26	3,98	3,77	3,61	3,48	3,38	3,21	3,04	2,86	2,74	2,66	2,56	2,45	2,33
30	9,18	6,35	5,24	4,62	4,23	3,95	3,74	3,58	3,45	3,34	3,18	3,01	2,82	2,71	2,63	2,52	2,42	2,30
40	8,83	6,07	4,98	4,37	3,99	3,71	3,51	3,35	3,22	3,12	2,95	2,78	2,60	2,48	2,40	2,30	2,18	2,06
50	8,63	5,90	4,83	4,23	3,85	3,58	3,38	3,22	3,09	2,99	2,82	2,65	2,47	2,35	2,27	2,16	2,05	1,93
75	8,37	5,69	4,63	4,05	3,67	3,41	3,21	3,05	2,93	2,82	2,66	2,49	2,31	2,19	2,10	1,99	1,88	1,74
100	8,24	5,59	4,54	3,96	3,59	3,33	3,13	2,97	2,85	2,74	2,58	2,41	2,23	2,11	2,02	1,91	1,79	1,65
150	8,12	5,49	4,45	3,88	3,51	3,25	3,05	2,89	2,77	2,67	2,51	2,33	2,15	2,03	1,94	1,83	1,70	1,56

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,999$																	
1	4052‡	4999‡	5403‡	5625‡	5764‡	5859‡	5928‡	5981‡	6022‡	6056‡	6106‡	6157‡	6209‡	6240‡	6261‡,	6287‡	6313‡	6339‡
2	998,5	999,0	999,2	999,35	999,3	999,3	999,4	999,4	999,4	999,4	999,4	999,4	999,5	999,5	999,5	999,5	999,5	999,5
3	167,0	148,5	141,1	137,1	134,6	132,9	131,6	130,6	129,9	129,3	128,3	127,4	126,4	125,8	125,5	125,0	124,5	124,0
4	74,14	61,25	56,18	53,44	51,71	50,53	49,66	49,00	48,47	48,05	47,41	46,76	46,10	45,70	45,43	45,09	44,75	44,40

‡означає $\times 10^2$

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,999$																	
5	47,18	37,12	33,20	31,09	29,75	28,83	28,16	27,65	27,24	26,92	26,42	25,91	25,39	25,08	24,87	24,60	24,33	24,06
6	35,51	27,00	23,70	21,92	20,80	20,03	19,46	19,03	18,69	18,41	17,99	17,56	17,12	16,85	16,67	16,44	16,21	15,98
7	29,25	21,69	18,77	17,20	16,21	15,52	15,02	14,63	14,33	14,08	13,71	13,32	12,93	12,69	12,53	12,33	12,12	11,91
8	25,41	18,49	15,83	14,39	13,48	12,86	12,40	12,05	11,77	11,54	11,19	10,84	10,48	10,26	10,11	9,92	9,73	9,53
9	22,86	16,39	13,90	12,56	11,71	11,13	10,70	10,37	10,11	9,89	9,57	9,24	8,90	8,69	8,55	8,37	8,19	8,00
10	21,04	14,91	12,55	11,28	10,48	9,93	9,52	9,20	8,96	8,75	8,45	8,13	7,80	7,60	7,47	7,30	7,12	6,94
11	19,69	13,81	11,56	10,35	9,58	9,05	8,66	8,35	8,12	7,92	7,63	7,32	7,01	6,81	6,68	6,52	6,35	6,18
12	18,64	12,97	10,80	9,63	8,89	8,38	8,00	7,71	7,48	7,29	7,00	6,71	6,40	6,22	6,09	5,93	5,76	5,59
13	17,82	12,31	10,21	9,07	8,35	7,86	7,49	7,21	6,98	6,80	6,52	6,23	5,93	5,75	5,63	5,47	5,30	5,14
14	17,14	11,78	9,73	8,62	7,92	7,44	7,08	6,80	6,58	6,40	6,13	5,85	5,56	5,38	5,25	5,10	4,94	4,77
15	16,59	11,34	9,34	8,25	7,57	7,09	6,74	6,47	6,26	6,08	5,81	5,54	5,25	5,07	4,95	4,80	4,64	4,47
16	16,12	10,97	9,01	7,94	7,27	6,80	6,46	6,19	5,98	5,81	5,55	5,27	4,99	4,82	4,70	4,54	4,39	4,23
17	15,72	10,66	8,73	7,68	7,02	6,56	6,22	5,96	5,75	5,58	5,32	5,05	4,78	4,60	4,48	4,33	4,18	4,02
18	15,38	10,39	8,49	7,46	6,81	6,35	6,02	5,76	5,56	5,39	5,13	4,87	4,59	4,42	4,30	4,15	4,00	3,84
19	15,08	10,16	8,28	7,27	6,62	6,18	5,85	5,59	5,39	5,22	4,97	4,70	4,43	4,26	4,14	3,99	3,84	3,68
20	14,82	9,95	8,10	7,10	6,46	6,02	5,69	5,44	5,24	5,08	4,82	4,56	4,29	4,12	4,00	3,86	3,70	3,54
21	14,59	9,77	7,94	6,95	6,32	5,88	5,56	5,31	5,11	4,95	4,70	4,44	4,17	4,00	3,88	3,74	3,58	3,42
22	14,38	9,61	7,80	6,81	6,19	5,76	5,44	5,19	4,99	4,83	4,58	4,33	4,06	3,89	3,78	3,63	3,48	3,32
23	14,20	9,47	7,67	6,70	6,08	5,65	5,33	5,09	4,89	4,73	4,48	4,23	3,96	3,79	3,68	3,53	3,38	3,22
24	14,03	9,34	7,55	6,59	5,98	5,55	5,23	4,99	4,80	4,64	4,39	4,14	3,87	3,71	3,59	3,45	3,29	3,14
25	13,88	9,22	7,45	6,49	5,89	5,46	5,15	4,91	4,71	4,56	4,31	4,06	3,79	3,63	3,52	3,37	3,22	3,06
26	13,74	9,12	7,36	6,41	5,80	5,38	5,07	4,83	4,64	4,48	4,24	3,99	3,72	3,56	3,44	3,30	3,15	2,99

Продовження таблиці А.4

k_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120
k_2	$p = 0,999$																	
27	13,61	9,02	7,27	6,33	5,73	5,31	5,00	4,76	4,57	4,41	4,17	3,92	3,66	3,49	3,38	3,23	3,08	2,92
28	13,50	8,93	7,19	6,25	5,66	5,24	4,93	4,69	4,50	4,35	4,11	3,86	3,60	3,43	3,32	3,18	3,02	2,86
29	13,39	8,85	7,12	6,19	5,59	5,18	4,87	4,64	4,45	4,29	4,05	3,80	3,54	3,38	3,27	3,12	2,97	2,81
30	13,29	8,77	7,05	6,12	5,53	5,12	4,82	4,58	4,39	4,24	4,00	3,75	3,49	3,33	3,22	3,07	2,92	2,76
40	12,61	8,25	6,59	5,70	5,13	4,73	4,44	4,21	4,02	3,87	3,64	3,40	3,14	2,98	2,87	2,73	2,57	2,41
50	12,22	7,96	6,34	5,46	4,90	4,51	4,22	4,00	3,82	3,67	3,44	3,20	2,95	2,79	2,68	2,53	2,38	2,21
75	11,73	7,58	6,01	5,16	4,62	4,24	3,96	3,74	3,56	3,42	3,19	2,96	2,71	2,55	2,44	2,29	2,13	1,96
100	11,50	7,41	5,86	5,02	4,48	4,11	3,83	3,61	3,44	3,30	3,07	2,84	2,59	2,43	2,32	2,17	2,01	1,83
150	11,27	7,24	5,71	4,88	4,35	3,98	3,71	3,49	3,32	3,18	2,96	2,73	2,48	2,32	2,21	2,06	1,89	1,70