МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
імені В. Н. КАРАЗІНА

ДИСПЕРСІЙНИЙ АНАЛІЗ ТА МНОЖИННИЙ РЕГРЕСІЙНИЙ АНАЛІЗ

Методичні рекомендації
до розв’язання задач
з курсу «Прикладні задачі математичної статистики»

Електронне видання

Харків – 2023

УДК 519.237+519.246.8(075.8)
Д 48

Рецензенти:

О. Л. Вишневецький – канд. фіз.-мат. наук, доцент кафедри вищої математики Харківського національного автомобільно-дорожнього університету;

Т. Б. Фастовська – канд. фіз.-мат. наук, доцент кафедри фундаментальної математики факультету математики і інформатики Харківського національного університету імені В. Н. Каразіна.

Затверджено до розміщення в мережі Інтернет рішенням Науково-методичної ради Харківського національного університету імені В. Н. Каразіна (протокол № 7 від 19 квітня 2023 року)

Д 48 **Дисперсійний** аналіз та множинний регресійний аналіз : методичні рекомендації до розв'язання задач з курсу «Прикладні задачі математичної статистики» [Електронне видання] / уклад. О. Л. Півень, Т. І. Сморцова. – Харків : ХНУ імені В. Н. Каразіна, 2023. – (PDF 40 с.)

Методичні рекомендації з курсу «Прикладні задачі математичної статистики» розроблені для студентів 1 курсу другого магістерського рівня освіти денної форми навчання спеціальності 113 «Прикладна математика». Тут висвітлені питання з розділів математичної статистики «Дисперсійний аналіз» та «Множинний регресійний аналіз». Наведено велику кількість прикладів розв'язання типових задач з цих розділів курсу, а також завдання для самостійної роботи.

Матеріали можуть бути корисними для студентів різних спеціальностей, які використовують методи прикладної математичної статистики для своїх досліджень.

УДК 519.237+519.246.8(075.8)

ЗМІСТ

Вступ 4

1. Однофакторний дисперсійний аналіз 5

2. Двофакторний дисперсійний аналіз...10

3. Множинний регресійний аналіз. Частинний коефіцієнт кореляції 20

Список літератури 33

Додатки. Таблиці математичної статистики 34

ВСТУП

Методичні рекомендації та завдання з курсу «Прикладні задачі математичної статистики» розроблені для студентів 1 курсу другого магістерського рівня освіти денної форми навчання спеціальності 113 «Прикладна математика» факультету математики і інформатики. Тут висвітлені питання з розділів математичної статистики «Дисперсійний аналіз» (розділи 1–2) та «Множинний регресійний аналіз» (розділ 3). Ці розділи математичної статистики застосовуються в різноманітних сферах: економіці, соціології, фінансовому аналізі, техніці та інших. У методичних вказівках наведено велику кількість прикладів розв'язання типових задач із зазначених розділів курсу, а також завдання для самостійної роботи. Усі задачі мають прикладний характер. У додатках наведені таблиці математичної статистики, які необхідні для розв'язання задач.

Студенти, яким викладається курс «Прикладні задачі математичної статистики», повинні мати знання з курсів «Теорія ймовірностей» та «Математична статистика». Зокрема, студенти мають володіти основними законами розподілу, методами оцінювання параметрів, характеристик випадкових величин та випадкових векторів, методами перевірки статистичних гіпотез про параметри нормального розподілу, критеріями узгодженості, основами кореляційного аналізу. Ці питання викладені в [1–4].

Ці матеріали можуть бути корисними для викладачів та студентів різних спеціальностей, які використовують методи прикладної математичної статистики для своїх досліджень, під час проведення практичних занять в аудиторії та для самостійної роботи.

1. ОДНОФАКТОРНИЙ ДИСПЕРСІЙНИЙ АНАЛІЗ

Модель однофакторного дисперсійного аналізу розглядається у [1, 2, 5–7].

Нехай є вибірка x_{ij}, $j = 1, \dots, n_i$ з відповідної нормальної генеральної сукупності X_i, $i = 1, \dots, m$ з математичним сподіванням a_i та сталою дисперсією σ^2. Потрібно перевірити гіпотезу про те, що величини a_i збігаються. Для цього обчислюємо послідовно

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}\,, i = 1, \dots, m; \quad n = \sum_{i=1}^{m} n_i\,; \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i} x_{ij}\,;$$

$$\bar{\sigma}_1^2 = \frac{1}{m-1}\sum_{i=1}^{m} n_i(\bar{x}_i - \bar{x})^2\,; \quad \bar{\sigma}_2^2 = \frac{1}{n-m}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)^2;$$

$$F = \frac{\bar{\sigma}_1^2}{\bar{\sigma}_2^2}. \tag{1.1}$$

Для заданого рівня значущості q та ступенів свободи $m-1$, $n-m$, за табл. 4 Додатку (закон розподілу Фішера) знаходимо величину $F_{m-1;n-m;q}$. Якщо $F < F_{m-1;n-m;q}$, то з довірчою ймовірністю $\alpha = 1 - q$ приймаємо гіпотезу про те, що усі середні a_i збігаються. Якщо $F \geq F_{m-1;n-m;q}$, то на рівні значущості q відхиляємо цю гіпотезу і приймаємо альтернативну гіпотезу про те, що не всі величини a_i однакові.

Довірчий інтервал для $a_i - a_j$, $i \neq j$, який відповідає заданій довірчій ймовірності α, має вигляд

$$\left(\left(\bar{x}_i - \bar{x}_j\right) - t_{n-m;q}\bar{\sigma}_2\sqrt{\frac{n_i + n_j}{n_i n_i}};\ \left(\bar{x}_i - \bar{x}_j\right) + t_{n-m;q}\bar{\sigma}_2\sqrt{\frac{n_i + n_j}{n_i n_i}}\right). \tag{1.2}$$

Для ймовірності $q = 1 - \alpha$ та числа ступенів свободи $n - m$ величина $t_{n-m;q}$ знаходиться за табл. 3 Додатку (закон розподілу Стьюдента). Нагадаємо, що якщо число ступенів свободи $k > 30$, то величина $t_{k;q}$ наближено знаходиться за допомогою функції Лапласа (табл. 1 Додатку) як $t_{k;q} = \Phi_0^{-1}\left(\frac{\alpha}{2}\right)$.

Приклад 1.1. Оцінити вплив рівня реклами всередині магазину на обсяг продажів.

Магазин	Рівень реклами		
	Високий	Середній	Низький
	Продажі, тис. грн		
1	10	8	5
2	9	8	7
3	10	7	6
4	8	9	4
5	9	6	5
6	8	4	2
7	9	5	3
8	7	5	2
9	7	6	1
10	6	4	2

Розв'язання. Є фактор – реклама, який має три різні рівні. Досліджується вплив цього фактору на результуючу ознаку – обсяг продажів. Тут $m = 3$, $n_i = 10$, $i = 1, 2, 3$. Нехай a_1, a_2, a_3 – середні обсяги продажів у магазинах з високим, середнім та низьким рівнями реклами товару відповідно. Провівши послідовно обчислення за формулами (1.1), отримаємо, що

$$n = 30; \bar{x}_1 = 8{,}3; \bar{x}_2 = 6{,}2; \bar{x}_3 = 3{,}7;$$
$$\bar{\sigma}_1^2 = 53{,}03; \bar{\sigma}_2^2 = 2{,}95; F = 17{,}94.$$

За табл. 4 Додатку (закон розподілу Фішера) для рівня значущості $q = 0{,}05$ знаходимо величину $F_{2;27;0{,}05} = 3{,}354$. Оскільки $F > F_{2;27;0{,}05}$, то на рівні значущості 0,05 ми відхиляємо гіпотезу про те, що всі середні a_i збігаються, та робимо висновок про істотний вплив рівня реклами на обсяг продажів. Для довірчої ймовірності $\alpha = 0{,}95$ за формулою (1.2) побудуємо довірчі інтервали для різниць середніх обсягів продажів за різних рівнів реклами. За табл. 3 Додатку (закон розподілу Стьюдента) знаходимо $t_{27;0,05} = 2{,}05$. Тоді шукані довірчі інтервали за формулою (1.2) мають вигляд

$$a_1 - a_2 \in (0{,}5239; 3{,}6761),$$
$$a_2 - a_3 \in (0{,}9239; 4{,}0761),$$
$$a_1 - a_3 \in (3{,}0239; 6{,}1761).$$

Завдання для самостійної роботи

1.1. За даними, які наведені нижче, з'ясувати, чи впливає оснащеність підприємств основними фондами на випуск продукції. Дані про продукцію в порівняних цінах за підприємствами з основними фондами 2–4 млн грн: 3,4; 2,4; 3,5; 4,1; 2,6; за підприємствами з основними фондами 4–6 млн грн: 6,1; 4,5; 10,6; 8,5; 8,8; 10,2; за підприємствами з основними фондами 6–8 млн грн: 19,6; 10,7; 12,4; 13,8; та за підприємствами з основними фондами 8–10 млн грн: 20,7; 18,2; 14,2.

1.2. Протягом чотирьох років використовувались п'ять різних технологій з вирощування сільськогосподарської культури. Перша технологія дала врожайність (в умовних одиницях): 1,2; 1,1; 1; 1,3; друга технологія – 0,6; 1,1; 0,8; 0,7; третя технологія – 1,7; 1,4; 1,3; 1,5; четверта технологія – 0,9; 0,6; 0,8; 1,0; п'ята технологія – 1,0; 1,4; 1,1; 0,9. Встановити, чи впливає застосування різних технологій на врожайність культур.

1.3. На заводі встановлено чотири лінії випуску облицювальної плитки. З кожної лінії випадковим чином відібрано по 10 плиток та виміряно їх товщини (в мм). Відхилення від нормального розміру плиток з першої лінії випуску склали 0,6; 0,2; 0,4; 0,5; 0,8; 0,2; 0,1; 0,6; 0,8; 0,8; з другої – 0,2; 0,2; 0,4; 0,3; 0,3; 0,6; 0,8; 0,2; 0,5; 0,5; з третьої – 0,8; 0,6; 0,2; 0,4; 0,9; 1,1; 0,8; 0,2; 0,4; 0,8; з четвертої – 0,7; 0,7; 0,3; 0,3; 0,2; 0,8; 0,6; 0,4; 0,2; 0,6. За цими даними встановити, чи впливає лінія випуску на якість плиток.

1.4. Нижче наведені дані про терміни служби ламп (в годинах) в кожній з чотирьох партій, виготовлених на одному заводі. Кожна партія ламп була виготовлена за окремою технологією. Тривалість горіння ламп першої партії склала 1660; 1610; 1650; 1680; 1700; 1700; 1700; другої партії – 1580; 1640; 1640; 1700; 1150; третьої партії – 1460; 1540; 1620; 1620; 1640; 1600; 1740; 1820; четвертої партії – 1510; 1520; 1530; 1570; 1000; 1680; 1530. З'ясувати, чи впливає технологія виробництва ламп на їх довговічність.

1.5. Для трьох вимірювальних приладів досліджені помилки вимірювання деякої величини. Нижче наведені одержані дані (в сотих частках мікронів). Дані для першого приладу: –4; –2; –21; –4; –7; для другого приладу: 7; 11; 30; 27, для третього приладу: 19; 2; –13. Перевірити, чи змінюється величина помилки від приладу до приладу.

1.6. Нижче наведені результати тестування чотирьох груп випускників виробничо-технічного училища за кількістю виготовлених однотипних деталей за зміну. Результати тестування першої групи: 60; 80; 75; 80; 85; 70, другої групи: 75; 66; 85; 80; 70; 80; 90, третьої: 60; 80; 65; 60; 60; 86; 75, четвертої: 95; 85; 100; 80. Кожна група навчалась за окремою методикою. З'ясувати, чи впливає методика навчання на якість підготовки учнів.

1.7. На трьох ділянках посівної однакової площі, які знаходились у 2020 році під чорним паром, зафіксовані такі врожайності пшениці (в ц/га): 26,6; 26,6; 30,6. На ще трьох ділянках тієї ж площі, які знаходились в 2020 році під картоплею, зафіксовані врожайності пшениці: 24,3; 25,2; 25,2. Нарешті, на трьох ділянках аналогічної площі, які знаходились під кормовими травами, врожайності пшениці такі: 26,6; 28; 31. З'ясувати, чи впливає спосіб використання землі у 2020 році на врожайність пшениці.

1.8. Дані про рейтинг політичної партії (у %), який вимірювався у навмання обраних шести регіонах Західної України: 14,5; 5,6; 23,8; 6,4; 26,2; 14,5; у навмання обраних шести регіонах Центральної України: 22,5; 12,2; 24,8; 16,8; 11,9; 26,6 та у навмання обраних шести регіонах Східної України: 13,4; 20,8; 30,8; 20,8; 6,4; 12,3. Перевірити, чи впливає регіон України на рейтинг цієї політичної партії.

1.9. Проводилися дослідження розподілу кількості кров'яних тілець в деякій одиниці об'єму крові людей, які знаходилися у певний час в трьох зонах на різних відстанях від Чорнобильської АЕС та в зоні, вільної від радіації. Кількість кров'яних тілець у 5 людей, які заходилися в зоні АЕС: 6; 8; 3; 2; 6; у 6 людей, які знаходилися на відстані 50 км від АЕС: 5; 4; 10; 11; 6; 8; у 6 людей, які знаходилися на відстані 100 км від АЕС: 5; 4; 13; 12; 10; 15 та у 4 людей, які знаходилися в зоні, вільної від радіації: 18; 16; 21; 20. Перевірити, чи впливає віддаленість від Чорнобильської АЕС на кількість кров'яних тілець у крові людей.

1.10. Оцінити істотність відмінностей успішності студентів у трьох групах. Дані про середній бал за сесію студентів першої групи: 3,7; 4,5; 3,6; 4,8; 4,7; 4,5; 5; 4,8; 4,9; середні бали за сесію студентів другої групи: 4; 3; 4,6; 4,3; 4,8; 5; 3,7; 3,7; 3; 3; 3; середні бали за сесію студентів третьої групи: 5; 5; 4,4; 4,4; 3; 3,8; 3,9; 3,9; 3,4; 3,5; 4,6; 4,7; 4,8; 4,6.

Зауваження. Умови задач 1.4, 1.6, 1.8 та 1.9 (або їх варіації) можна знайти у навчальному посібнику [6].

2. ДВОФАКТОРНИЙ ДИСПЕРСІЙНИЙ АНАЛІЗ

Модель двофакторного дисперсійного аналізу розглядається у [2, 5–7].

Нехай є два фактори A та B. Фактор A набуває n різних значень, фактор B набуває m різних значень. Досліджується вплив факторів A та B на деяку випадкову величину, яка має нормальний закон розподілу з математичним сподіванням a_{ij} та сталою дисперсією σ^2. За i-го значення фактору A та j-го значення фактору B є вибірка x_{ijk}, $k = 1, \dots, r$ з нормальної генеральної сукупності X_{ij}, $i = 1, \dots, n$; $j = 1, \dots, m$ з математичним сподіванням a_{ij} та сталою дисперсією σ^2. Потрібно встановити вплив факторів A та B, а за $r > 1$ – і їх взаємодії, на величину, що досліджується.

Випадок $r = 1$. Позначимо $x_{ij} = x_{ij1}$. Обчислюємо послідовно

$$\bar{x} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij};$$

$$\bar{x}_{i\cdot} = \frac{1}{m}\sum_{j=1}^{m} x_{ij}, i = 1, \dots, n, \quad \bar{x}_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n} x_{ij}, j = 1, \dots, m;$$

$$\bar{\sigma}_1^2 = \frac{m}{n-1}\sum_{i=1}^{n}(\bar{x}_{i\cdot} - \bar{x})^2; \quad \bar{\sigma}_2^2 = \frac{n}{m-1}\sum_{j=1}^{m}\left(\bar{x}_{\cdot j} - \bar{x}\right)^2;$$

$$\bar{\sigma}_3^2 = \frac{1}{(n-1)(m-1)}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(x_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x}\right)^2;$$

$$F_A = \frac{\bar{\sigma}_1^2}{\bar{\sigma}_3^2}; F_B = \frac{\bar{\sigma}_2^2}{\bar{\sigma}_3^2}. \qquad (2.1)$$

Задамо рівень значущості q. Для ступенів свободи $n - 1$, $(n-1)(m-1)$ за табл. 4 Додатку (закон розподілу Фішера) знаходимо величину $F_1 = F_{(n-1);(n-1)(m-1);q}$, для ступенів свободи $m - 1$, $(n-1)(m-1)$ – величину $F_2 = F_{(m-1);(n-1)(m-1);q}$. Якщо

$F_A < F_1$, то з довірчою ймовірністю $\alpha = 1 - q$ приймаємо гіпотезу про відсутність впливу фактору A на величину, яка досліджується. якщо, $F_A \geq F_1$, то на рівні значущості q відхиляємо цю гіпотезу та приймаємо альтернативну гіпотезу про вплив фактору A на величину, що досліджується. Якщо $F_B < F_2$, то з довірчою ймовірністю α приймаємо гіпотезу про відсутність впливу фактору B на величину, що досліджується. Якщо $F_B \geq F_2$, то на рівні значущості q відхиляємо цю гіпотезу та приймаємо альтернативну гіпотезу про вплив фактору B на величину, яка досліджується.

Випадок $r > 1$. Обчислюємо послідовно

$$\bar{x} = \frac{1}{nmr}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} x_{ijk};$$

$$\bar{x}_{i\cdot\cdot} = \frac{1}{mr}\sum_{j=1}^{m}\sum_{k=1}^{r} x_{ijk}, \quad \bar{x}_{\cdot j\cdot} = \frac{1}{nr}\sum_{i=1}^{n}\sum_{k=1}^{r} x_{ijk}, \quad \bar{x}_{ij\cdot} = \frac{1}{r}\sum_{k=1}^{r} x_{ijk},$$

$$i = 1, \dots, n, \quad j = 1, \dots, m;$$

$$\bar{\sigma}_1^2 = \frac{mr}{n-1}\sum_{i=1}^{n}(\bar{x}_{i\cdot\cdot} - \bar{x})^2; \quad \bar{\sigma}_2^2 = \frac{nr}{m-1}\sum_{j=1}^{m}(\bar{x}_{\cdot j\cdot} - \bar{x})^2;$$

$$\bar{\sigma}_3^2 = \frac{r}{(n-1)(m-1)}\sum_{i=1}^{n}\sum_{j=1}^{m}(x_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})^2;$$

$$\bar{\sigma}_4^2 = \frac{1}{nm(r-1)}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r}(x_{ijk} - \bar{x}_{ij\cdot})^2$$

$$F_A = \frac{\bar{\sigma}_1^2}{\bar{\sigma}_4^2}; F_B = \frac{\bar{\sigma}_2^2}{\bar{\sigma}_4^2}; F_{AB} = \frac{\bar{\sigma}_3^2}{\bar{\sigma}_4^2}. \quad (2.2)$$

Задамо рівень значущості q. Для ступенів свободи $n - 1$, $nm(r-1)$, за табл. 4 Додатку (закон розподілу Фішера) знаходимо величину $F_1 = F_{(n-1);nm(r-1);q}$, для ступенів свободи $m - 1$, $nm(r - 1)$ знаходимо величину $F_2 = F_{(m-1);nm(r-1);q}$, для ступенів свободи $(n-1)(m-1), nm(r-1)$ знаходимо величину $F_3 = F_{(m-1)(n-1);nm(r-1);q}$. Якщо $F_A < F_1$, то з довірчою ймовірністю

$\alpha = 1 - q$ приймаємо гіпотезу про відсутність впливу фактору A на величину, яка досліджується. Якщо $F_A \geq F_1$, то на рівні значущості q відхиляємо цю гіпотезу та приймаємо альтернативну гіпотезу про вплив фактору A на величину, яка досліджується. Якщо $F_B < F_2$, то з довірчою ймовірністю α приймаємо гіпотезу про відсутність впливу фактору B на величину, яка досліджується. Якщо $F_B \geq F_2$, то на рівні значущості q відхиляємо цю гіпотезу та приймаємо альтернативну гіпотезу про вплив фактору B на величину, яка досліджується. Якщо $F_{AB} < F_3$, то з довірчою ймовірністю α приймаємо гіпотезу про відсутність впливу взаємодії факторів A та B на величину, яка досліджується. Якщо $F_{AB} \geq F_3$, то на рівні значущості q відхиляємо цю гіпотезу та приймаємо альтернативну гіпотезу про вплив взаємодії факторів A та B на величину, яка досліджується.

Приклад 2.1. Встановити вплив факторів A – технології виготовлення (дві технології) та фактору B – матеріалу для виготовлення (три марки сталі) на якість деталей.

Фактор *B* / Фактор *A*	B_1	B_2	B_3
A_1	10	8	15
A_2	12	11	13

Розв'язання. Є два фактори: фактор A – технологія виготовлення деталі та фактор B – марка сталі. Фактор A набуває $n = 2$ значення, фактор B набуває $m = 3$ значення. Величина, яка досліджується, характеризує якість виготовлених деталей. Для будь-яких фіксованих значень факторів A та B отримано $r = 1$ вибіркове значення величини, яка досліджується. Послідовно застосовуючи формули (2.1), маємо

$$\bar{x} = 11{,}5,\ \bar{x}_{1\cdot} = 11,\ \bar{x}_{2\cdot} = 12,\ \bar{x}_{\cdot 1} = 11,\ \bar{x}_{\cdot 2} = 9{,}5,\ \bar{x}_{\cdot 3} = 14,$$
$$\bar{\sigma}_1^2 = 1{,}5,\ \bar{\sigma}_2^2 = 10{,}5,\ \bar{\sigma}_3^2 = 3{,}5,\ F_A = 0{,}4286,\ F_B = 3.$$

За табл. 4 Додатку (закон розподілу Фішера) для рівня значущості $q = 0{,}05$ знаходимо значення $F_1 = F_{1;2;0{,}05} = 18{,}51$, $F_2 = F_{2;2;0{,}05} = 19$. Оскільки виконуються нерівності $F_A < F_1$, $F_B < F_2$, то з довірчою ймовірністю 0,95 ми приймаємо гіпотезу про відсутність впливу як фактору *A*, так і фактору *B*.

Приклад 2.2. Компанія хоче перевірити ефективність своєї реклами. Для цього був обраний продукт та створені два типи рекламних роликів – серйозний та смішний. Рекламу розміщено в робочі та вихідні дні. Також вибрали 16 клієнтів та навмання розподілили їх на 4 групи з чотирьох клієнтів кожна. Кожний клієнт оцінив ефективність ролика за двадцятибальною шкалою. Результати оцінювання наведені в таблиці. Встановити, чи впливає тип ролика та тип дня, а також їх взаємодія, на ефективність реклами.

День / Тип ролика	Робочий	Вихідний
Смішний	6 10 11 9	15 18 14 16
Серйозний	8 13 12 10	19 20 13 17

Розв'язання. У цій задачі є два фактори: фактор *A* – тип ролика (з двома значеннями) та фактор *B* – тип дня (також з двома значеннями). Дослідимо вплив факторів *A* та *B* на величину – ефективність реклами, яка вимірюється в деякій двадцятибальній шкалі. Для будь-яких фіксованих значень факторів *A* та *B* маємо 4 значення ефективності реклами. Тому $r = 4$, $m = n = 2$. Послідовно застосовуючи формули (2.2), маємо

$$\bar{x} = 13{,}1875,\ \bar{x}_{1..} = 12{,}375,\ \bar{x}_{2..} = 14,\ \bar{x}_{.1.} = 9{,}875,\ \bar{x}_{.2.} = 16{,}5,$$
$$\bar{x}_{11.} = 9,\ \bar{x}_{12.} = 15{,}75,\ \bar{x}_{21.} = 10{,}75,\ \bar{x}_{22.} = 17{,}25,$$
$$\bar{\sigma}_1^2 = 10{,}5625,\ \bar{\sigma}_2^2 = 175{,}5625,\ \bar{\sigma}_3^2 = 0{,}0625,\ \bar{\sigma}_4^2 = 5{,}5208,$$
$$F_A = 1{,}91,\ F_B = 31{,}8,\ F_{AB} = 0{,}01.$$

За табл. 4 Додатку (закон розподілу Фішера) для рівня значущості $q = 0{,}05$ знаходимо $F_1 = F_2 = F_3 = F_{1;12;0,05} = 4{,}75$. Оскільки $F_A < F_{1;12;0,05}$ та $F_{AB} < F_{1;12;0,05}$, то з довірчою ймовірністю 0,95 ми приймаємо гіпотезу про відсутність впливу фактору A та гіпотезу про відсутність впливу взаємодії факторів A та B на величину, яка досліджується. А оскільки $F_B > F_{1;12;0,05}$, то на рівні значущості 0,05 ми відхиляємо гіпотезу про відсутність впливу фактору B на результуючу величину. Звідси робимо висновок, що на ефективність реклами впливає лише тип дня, а тип рекламного ролика на її ефективність не впливає.

Завдання для самостійної роботи

2.1. Досліджується вплив на врожайність ячменю (в ц/га) наведених нижче факторів. Фактор A: посів здійснюється після чорного пару – A_1, після коренеплодів – A_2, після колосових культур – A_3; фактор B – сорт ячменю (три сорти). Дані про врожайності ячменю (в ц/га) для будь-яких фіксованих значень факторів A та B наведені в таблиці. Встановити, чи впливають ці фактори та їх взаємодія на врожайність ячменю.

Фактор A / Фактор B	A_1	A_2	A_3
B_1	34,2; 30,6; 36,8; 33,4; 35; 32,5; 34,2; 36	42,5; 40,4; 44,6; 46,8; 39,4; 38,6; 45,8; 49,3	44,2; 46; 45,6; 49,3; 45,8; 42,3; 40,8; 41,4
B_2	32,5; 30,4; 39,4; 40,3; 36,4; 38,9; 39,8; 42	30,3; 35,3; 36,8; 40,5; 28,4; 33,2; 39,1; 26,9	40,3; 45; 46,8; 30,2; 48,8; 50,2; 39; 38,5
B_3	33,3; 34,8; 39,2; 35; 32,4; 34; 39,8; 40,8	30,4; 36; 40,5; 44,4; 30,8; 42,5; 46; 33,5	32,3; 29,8; 34,3; 42; 34,8; 31,6; 40; 29,6

2.2. Досліджується вплив на міцність чавуну двох факторів: A – кількість кремнію в чавуні (A_1 – 0,24 %, A_2 – 0,42 %, A_3 – 0,52 %); B – температурний режим плавлення (два режиму). Дані про міцність чавуну (в кг/м^3) для будь-яких фіксованих значень факторів A та B наведені в таблиці. Встановити, чи впливають ці фактори та їх взаємодія на міцність чавуну.

Фактор A / Фактор B	A_1	A_2	A_3
B_1	40,2; 40,8; 38,2; 39,6; 42,4; 44,5; 40,1; 38,8	42,5; 43,4; 44,5; 46,4; 40,1; 36,5; 40,3; 41,8	49,2; 50,2; 48,4; 50; 52,5; 38,4; 49,8; 50,4
B_2	33,4; 36,5; 34,4; 40,2; 42; 30,2; 31,8; 35,5	31,6; 33,4; 38,4; 35; 38,9; 29,5; 28,4; 30,6	29,3; 35,6; 36; 26,8; 38; 28,5; 30,6; 40,2

2.3. Досліджується вплив на врожайність кукурудзи двох факторів: A – внесення добрив у ґрунт (три типи добрив); B – глибина поливу землі (три значення глибини). Дані про врожайність кукурудзи (в ц/га) для будь-яких фіксованих значень факторів A та B наведені в таблиці. Встановити, чи впливають ці фактори на врожайність кукурудзи.

Фактор A / Фактор B	A_1	A_2	A_3
B_1	30,2	36,2	40,2
B_2	44,2	42,4	42,3
B_3	40,2	38,5	43,2

2.4. Досліджується вплив двох факторів на опріснення морської води: A – тип опріснювача (три типи опріснювачів), B – різні лабораторії (три лабораторії). Дані про концентрацію солі в опрісненій воді (в проміле, ‰) наведені в таблиці. Встановити, чи впливають ці фактори на концентрацію солі в опрісненій воді.

Фактор B \ Фактор A	A_1	A_2	A_3
B_1	3,6	2,92	3,25
B_2	4,2	3,35	3,75
B_3	3,46	3,42	3,63

2.5. Досліджується вплив факторів A та B на міцність сталі: A – процентний вміст нікелю (три значення), B – процентний вміст марганцю (два значення). Дані про границі міцності сталі (в кг/м^3) при розтягуванні для будь-яких фіксованих значень факторів A та B наведені в таблиці. Встановити, чи впливають ці фактори та їх взаємодія на пружність сталі.

Фактор B \ Фактор A	A_1	A_2	A_3
B_1	36,4; 38,7; 36,5	41,2; 42; 42,8	36,5; 39,8; 42,4
B_2	39,2; 42,3; 44,5	39,7; 38,4; 42,5	40,8; 43,2; 41,8

2.6. Досліджується вплив факторів A та B на твердість сплаву: A – процентний вміст нікелю (три значення), B – процентний вміст хрому (три значення). Дані про твердість сплаву (кг·с/мм^2) для будь-яких фіксованих значень факторів A та B наведені в таблиці. Встановити, чи впливають ці фактори та їх взаємодія на твердість сплаву.

Фактор B \ Фактор A	A_1	A_2	A_3
B_1	53,2; 53,9; 54,8; 55,9; 62,2	55,4; 66,7; 77,2; 53,2; 65,4	68,3; 69,8; 74,7; 79,2; 61,5
B_2	67,2; 66,2; 55,3; 53; 72,3	77,2; 65,4; 53,9; 65,1; 63,4	77,9; 62,3; 68,9; 64,5; 73,2
B_3	70,2; 72,1; 54,4; 53,1; 73,4	69,2; 65,4; 70,4; 55,4; 62,3	75,5; 76,4; 54,2; 56,1; 62,3

2.7. Досліджується вплив факторів *A* та *B* на продуктивність праці: *A* – плинність кадрів (три значення), *B* – оплата праці (два значення). Дані про продуктивність праці (кількість випущеної продукції за одиницю часу) для будь-яких фіксованих значень факторів *A* та *B* наведені в таблиці. Встановити, чи впливають ці фактори та їх взаємодія на продуктивність праці.

Фактор *A* / Фактор *B*	A_1	A_2	A_3
B_1	15,62; 14,3; 14,25; 15,81; 16,35; 15,61; 14,3; 12,5; 11,2	10,83; 10,2; 13,4; 16,25; 12,2; 5,45; 6,41; 8,93; 13,44	12,44; 14,5; 7,6; 6,75; 8,96; 16,37; 9,82; 7,83; 10,53
B_2	16,52; 14,21; 6,85; 8,7; 10,43; 3,5; 12,8; 11,6; 6,72	13,24; 8,16; 9,44; 10,8; 14,56; 12,46; 11,83; 10,99; 16,42	6,81; 5,74; 10,36; 14,57; 12,44; 13,47; 15,25; 13,4; 5,07

2.8. Досліджується вплив факторів *A* та *B* на рейтинг політичної партії: *A* – регіони країни, *B* – вікові границі опитаних. Результати дослідження наведені в таблиці. Встановити, чи впливають ці фактори та їх взаємодія на рейтинг політичної партії.

Фактор *A* / Фактор *B*	Західна Україна	Південна Україна	Східна Україна
До 30 років	9,5; 5,5; 4,2; 6,7	4,2; 10,5; 8,9; 9,6	8,6; 7,5; 4,3; 19,8
30–55 років	2,5; 3,4; 7,8; 12,4	6,5; 7,2; 13,6; 22,4	12,5; 10,6; 22,4; 8,5
55–70 років	2,1; 3,3; 7,8; 2,2	4,5; 12,6; 22,5; 40,1	15,8; 35,6; 21,4; 3,2

2.9. Партія м'яса від п'яти різних постачальників завантажується в машину для пакування в банки. Машина має шість циліндрів, які наповнюють банки. Навмання з кожної партії та з кожного циліндру було взято по три наповнені банки. Встановлені після зважування відхилення маси банок від номінального значення (в грамах) наведені в таблиці. Чи істотний влив факторів *A* (постачальника), *B* (тип циліндру) та їх взаємодія на величину зазначених відхилень?

Постачальник / Циліндр	1	2	3	4	5
1	1; 1; 2	4; 3; 5	6; 3; 7	3; 1; 3	1; 3; 3
2	–1; 3;–1	–2; 1; 0	3; 1; 5	2; 0; 1	1; 0; 1
3	1; 1; 1	2; 0; 1	2; 4; 3	1; 3; 3	3; 3; 3
4	–2; 3; 0	–2; 0; 1	3; 3; 4	0; 0; 2	0; 1; 1
5	1; 1; –1	2; 1; 5	0; 1; 2	1; 0; –1	–2; 3; 1
6	0; 1; 1	0; 0; 3	3; 3; 4	3; 0; 2	3; 1; 2

2.10. Під час обробки матеріалів обстеження бюджету часу підраховано час, який в середньому один робочий витрачає на читання газет (в годинах). Встановити вплив доходу та складу сім'ї, а також взаємодії цих двох факторів на середній час читання газет.

Дохід в міс. / Склад сім'ї	До 10000 грн	10000–20000 грн	Більше 20000 грн
З дошкільнятами	0,7; 1; 0,9	1,6; 2,2; 0,8	1,3; 1,7; 0,7
З дітьми шкільного віку	1,5; 1,8; 0,5	1,7; 1,9; 0,5	1,8; 1,9; 0,2
Без дітей	1,5; 1,6; 1,6	2,1; 0,2; 1,4	1,5; 1,5; 0,7

2.11. З 10 магазинів зібрали дані про сумарні обсяги продажів (в тис. грн) товарів кожного з трьох типів: з високим, середнім та низьким рівнями реклами. Встановити, чи впливає на обсяг продажів товару рівень реклами цього товару та магазин, в якому він продається.

Рівень реклами / Магазин	Високий	Середній	Низький
1	10	8	5
2	9	8	7
3	10	7	6
4	8	9	4
5	9	6	5
6	8	4	2
7	9	5	3
8	7	5	2
9	7	6	1
10	6	4	2

2.12. У таблиці містяться дані про врожайність (в ц/га) чотирьох сортів пшениці на виділених п’яти ділянках землі (блоках). Встановити, чи впливають на врожайність сорт пшениці та ділянки землі.

Номер блоку / Сорт	1	2	3	4	5
1	2,87	2,67	2,16	2,50	2,82
2	2,45	2,85	2,77	2,87	3,25
3	2,32	2,47	2,00	2,40	2,40
4	2,90	2,87	2,25	2,80	2,70

Зауваження. Умови деяких із запропонованих задач або їх варіації можна знайти у навчальному посібнику [6].

3. МНОЖИННИЙ РЕГРЕСІЙНИЙ АНАЛІЗ. ЧАСТИННИЙ КОЕФІЦІЄНТ КОРЕЛЯЦІЇ

Модель множинного лінійного регресійного аналізу розглядається у [2, 7, 8].

Припустимо, що випадкові величини $Y, X_1, \dots, X_p$ підпорядковуються закону $(p+1)$-вимірного нормального розподілу та задовольняють рівняння ***лінійної стохастичної залежності:***

$$Y = \theta_0 + \sum_{j=1}^{p} \theta_j X_j + \varepsilon, \tag{3.1}$$

де ε – деяка випадкова величина, яка має нормальний розподіл з нульовим математичним сподіванням та дисперсією σ^2, $\theta_0, \theta_1, \dots, \theta_p$ – числові параметри.

Відомі вибіркові значення $y_1, \dots, y_n$ з генеральної сукупності Y та вибіркові значення $x_{j1}, \dots, x_{jn}$ з відповідної генеральної сукупності $X_j, j = 1, \dots, p$, які відповідають $y_1, \dots, y_n$. За цими вибірковими даними будується матриця розмірності $n \times (p+1)$

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{pmatrix},$$

на яку накладається обмеження

$$rank\, X = p + 1 < n. \tag{3.2}$$

Для оцінки коефіцієнтів $\theta_0, \theta_1, \dots, \theta_p$ рівняння (3.1) за заданими вибірковими даними застосовується ***метод найменших квадратів***. Згідно з цим методом, за виконання обмеження (3.2) оцінки $\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_p$ відповідних параметрів $\theta_0, \theta_1, \dots, \theta_p$ рівняння (3.1) знаходяться за формулою

$$\hat{\theta} = \left(\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_p\right)^T = (X^T X)^{-1} X^T y, \qquad (3.3)$$

де $y = (y_1, \dots, y_n)^T$, символ T позначає операцію транспонування матриці.

Введемо вектор $\hat{y} = (\hat{y}_1, \dots, \hat{y}_n)^T = X\hat{\theta}$ та статистику

$$\hat{s}^2 = \frac{1}{n-p-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2,$$

яка є незсунутою оцінкою дисперсії σ^2.

Довірчий інтервал для параметру θ_j, $j = 0, \dots, p$, ***регресійної моделі*** **(3.1)**, який відповідає заданій довірчій ймовірності α, має вигляд

$$\theta_j = \left(\hat{\theta}_j - \hat{s} t_{n-p-1;q} \sqrt{(X^T X)^{-1}_{j+1,j+1}}, \hat{\theta}_j + \hat{s} t_{n-p-1;q} \sqrt{(X^T X)^{-1}_{j+1,j+1}}\right), \qquad (3.4)$$

де $(X^T X)^{-1}_{ii}$, $i = 1, \dots, p+1$, – i-й елемент головної діагоналі матриці $(X^T X)^{-1}$, а величина $t_{n-p-1;q}$ знаходиться за табл. 3 Додатку (закон розподілу Стьюдента) для ймовірності $q = 1 - \alpha$ та числа ступенів свободи $(n - p - 1)$.

Перевірка статистичної значущості параметру θ_j, $j = 0, \dots, p$, ***регресійної моделі*** **(3.1)**. Задамо довірчу ймовірність α. За табл. 3 Додатку (закон розподілу Стьюдента) для рівня значущості $q = 1 - \alpha$ та числа ступенів свободи $(n - p - 1)$ знаходимо величину $t_{n-p-1;q}$. Обчислюємо $t_j = \frac{\hat{\theta}_j}{\hat{s}\sqrt{(X^T X)^{-1}_{j+1,j+1}}}$. Якщо $|t_j| < t_{n-p-1;q}$, то з довірчою ймовірністю α приймаємо гіпотезу про те, що $\theta_j = 0$, тобто параметр θ_j є статистично незначущим. Якщо $|t_j| \geq t_{n-p-1;q}$, то на рівні значущості q ми відхиляємо гіпотезу про те, що $\theta_j = 0$ та робимо висновок про статистичну значущість параметру θ_j.

Довірчий інтервал для дисперсії залишків σ^2, яка відповідає заданій довірчій ймовірності α, має вигляд

$$\sigma^2 \in \left(\frac{(n-p-1)\hat{s}^2}{\chi^2_{n-p-1;\,\frac{q}{2}}};\ \ \frac{(n-p-1)\hat{s}^2}{\chi^2_{n-p-1;\,1-\frac{q}{2}}} \right). \qquad (3.5)$$

Для ймовірності $q = 1 - \alpha$ та числа ступенів свободи $(n-p-1)$ величини $\chi^2_{n-p-1;\,\frac{q}{2}}$ та $\chi^2_{n-p-1;\,1-\frac{q}{2}}$ визначаються за табл. 2 Додатку (закон розподілу χ^2).

Довірчий інтервал для функції регресії. Нехай виконується рівняння лінійної стохастичної залежності (3.1). Розглянемо функцію регресії

$$f(x_1, \dots, x_p) = \theta_0 + \sum_{j=1}^{p} \theta_j X_j$$

за фіксованих значень $x_1, \dots, x_p$ відповідних ознак $X_1, \dots, X_p$. Довірчий інтервал для функції $f(x_1, \dots, x_p)$, який відповідає заданій довірчій ймовірності α, має вигляд

$$f(x_1, \dots, x_p) \in \Big(\hat{f}(x_1, \dots, x_p) - \hat{s}t_{n-p-1;q}\sqrt{((X^TX)^{-1}z, z)},$$
$$\hat{f}(x_1, \dots, x_p) + \hat{s}t_{n-p-1;q}\sqrt{((X^TX)^{-1}z, z)} \Big), \qquad (3.6)$$

де

$$z = (1, x_1, \dots, x_p)^T, \quad \hat{f}(x_1, \dots, x_p) = \hat{\theta}_0 + \sum_{j=1}^{p} \hat{\theta}_j x_j. \qquad (3.7)$$

Перевірка лінійної регресійної моделі на адекватність. Обчислимо

$$F = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 / p}{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2 / (n-p-1)}, \qquad \text{де } \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \quad (3.8)$$

Для заданого рівня значущості q та ступенів свободи $p, n-p-1$ за табл. 4 Додатку (закон розподілу Фішера) знаходимо величину $F_{p;n-p-1;q}$. Якщо $F < F_{p;n-p-1;q}$, то з довірчою ймовірністю $\alpha = 1 -$

q приймаємо гіпотезу про неадекватність лінійної регресійної моделі. Якщо $F \geq F_{p;n-p-1;q}$, то на рівні значущості q відхиляємо гіпотезу про неадекватність лінійної регресійної моделі та приймаємо альтернативну гіпотезу про адекватність лінійної регресійної моделі.

Вибіркова оцінка частинного коефіцієнту кореляції. Вибіркова оцінка коефіцієнту кореляції між випадковими величинами $X_k, X_j, k = 0, \ldots, p, j = 0, \ldots, p, X_0 = Y$ має вигляд

$$\bar{r}_{kj} = \frac{n\sum_{i=1}^{n} x_{ki}x_{ji} - \sum_{i=1}^{n} x_{ki}\sum_{i=1}^{n} x_{ji}}{\sqrt{\left(n\sum_{i=1}^{n} x_{ki}^2 - \left(\sum_{i=1}^{n} x_{ki}\right)^2\right)\left(n\sum_{i=1}^{n} x_{ji}^2 - \left(\sum_{i=1}^{n} x_{ji}\right)^2\right)}}, \quad (3.9)$$

де $x_{0i} = y_i,\ i = 1, \ldots, n$. Співвідношення (3.9) визначають вибіркову кореляційну матрицю $\bar{R} = \{\bar{r}_{kj}\}$ випадкового вектору $\left(X_0, \ldots, X_p\right)^T$. Вибіркова оцінка частинного коефіцієнту кореляції між випадковими величинами $X_k,\ X_j,\ k = 0, \ldots, p,\ j = 0, \ldots, p,\ k \neq j$, обчислюється за формулою

$$\bar{\rho}_{kj} = -\frac{\bar{R}_{kj}}{\sqrt{\bar{R}_{kk}\bar{R}_{jj}}}, \quad (3.10)$$

де $\bar{R}_{lm}$, $l = 0, \ldots, p$, $m = 0, \ldots, p$ – алгебраїчне доповнення елементу $\bar{r}_{lm}$ матриці $\bar{R}$.

Перевірка статистичної значущості частинного коефіцієнту кореляції. Для перевірки статистичної значущості частинного коефіцієнту кореляції між випадковими величинами X_k, X_j, $k = 0, \ldots, p$, $j = 0, \ldots, p$, $k \neq j$, обчислюється вибіркова оцінка цього коефіцієнту кореляції за формулою (3.10) та статистика:

$$t = \frac{\bar{\rho}_{kj}\sqrt{n-p-1}}{\sqrt{1-\bar{\rho}_{kj}^2}}. \quad (3.11)$$

Для заданого рівня значущості q та числа ступенів свободи $(n-p-1)$ за табл. 3 Додатку (закон розподілу Стьюдента) заходимо величину $t_{n-p-1;q}$. Якщо $|t| < t_{n-p-1;q}$, то з довірчою ймовірністю $\alpha = 1 - q$ приймаємо гіпотезу про те, що відхилення від нуля частинного коефіцієнту кореляції не є значущим. Якщо $|t| \geq t_{n-p-1;q}$, то на рівні значущості q відхиляємо цю гіпотезу та приймаємо альтернативну гіпотезу про те, що відхилення від нуля частинного коефіцієнту кореляції є значущим.

Приклад 3.1. Бюджетне обстеження п'яти випадково вибраних сімей дало такі результати (в тис. грн):

Сім'я	Накопичення, Y	Дохід, X_1	Майно, X_2
1	3	40	60
2	6	55	36
3	5	45	36
4	3,5	30	15
5	1,5	30	90

1. Оцінити параметри лінійної регресійної моделі залежності Y від X_1, X_2 за методом найменших квадратів, виписати рівняння лінійної регресії.
2. Перевірити отримане рівняння регресії на адекватність.
3. Побудувати довірчі інтервали для дисперсії залишків, параметру θ_1 регресійної моделі, а також для функції регресії за рівнем доходу 40 тис. грн та вартістю майна 60 тис. грн. Перевірити статистичну значущість параметру θ_1.
4. Обчислити вибіркову оцінку частинного коефіцієнту кореляції між випадковими величинами Y та X_1. Перевірити статистичну значущість частинного коефіцієнту кореляції між цими величинами.

Розв'язання.

1. Для цього прикладу маємо $n = 5$, $p = 2$. Будуємо матрицю X та вектор y:

$$X = \begin{pmatrix} 1 & 40 & 60 \\ 1 & 55 & 36 \\ 1 & 45 & 36 \\ 1 & 30 & 15 \\ 1 & 30 & 90 \end{pmatrix}, \quad y = \begin{pmatrix} 3 \\ 6 \\ 5 \\ 3{,}5 \\ 1{,}5 \end{pmatrix}.$$

Тоді

$$X^T X = \begin{pmatrix} 5 & 200 & 237 \\ 200 & 8450 & 9150 \\ 237 & 9150 & 14517 \end{pmatrix},$$

$$(X^T X)^{-1} = \begin{pmatrix} 5{,}6916 & -0{,}1074 & -0{,}0252 \\ -0{,}1074 & 0{,}0024 & 0{,}00024 \\ -0{,}0252 & 0{,}00024 & 0{,}00033 \end{pmatrix}$$

Тепер методом найменших квадратів обчислюємо оцінку за формулою (3.3)

$$\hat{\theta} = (X^T X)^{-1} X^T y = (0{,}2787 \quad 0{,}1229 \quad -0{,}0294)^T$$

та отримуємо рівняння регресії

$$\hat{y}(X_1, X_2) = 0{,}2787 + 0{,}1229 X_1 - 0{,}0294 X_2.$$

2. Для перевірки побудованого рівняння регресії на адекватність обчислимо вектор

$$\hat{y} = (\hat{y}_1, \dots, \hat{y}_n)^T = X\hat{\theta} = (3{,}4295; 5{,}9784; 4{,}7496; 3{,}5240; 1{,}3185)^T$$

та середнє значення накопичення $\bar{y} = 3{,}8$. За формулою (3.8) обчислюємо статистику $F = 42{,}7532$. Для заданого рівня значущості $q = 0{,}05$ знаходимо за табл. 4 Додатку (закон розподілу Фішера) $F_{2;2;0{,}05} = 19$. Оскільки $F > F_{2;2;0{,}05}$, то на рівні значущості 0,05 гіпотезу про неадекватність моделі регресії слід відкинути, а отже, побудована модель регресії є адекватною.

3. Використовуючи співвідношення (3.4), побудуємо тепер довірчий інтервал для параметру θ_1, який відповідає довірчій ймовірності $\alpha = 0{,}95$. Покладемо в (3.4) $j = 1$. Обчислимо значення $\hat{s}^2 = 0{,}1406$, $\hat{s} = 0{,}3749$, $(X^T X)^{-1}_{22} = 0{,}0024$. За табл. 3 Додатку (закон розподілу Стьюдента) знаходимо $t_{2;0{,}05} = 4{,}3$. Підставляючи знайдені значення у співвідношення (3.4), отримуємо 95% довірчий інтервал $(0{,}0439;\ 0{,}2018)$ для параметру θ_1.

За табл. 2 Додатку (закон розподілу χ^2) знаходимо, що

$$\chi^2_{2;0,025} = 7{,}38, \quad \chi^2_{2;0,975} = 0{,}05$$

та за формулою (3.5) будуємо 95% довірчий інтервал $(0{,}0381;\ 5{,}6225)$ для дисперсії залишків σ^2.

Побудуємо тепер довірчий інтервал для функції регресії за рівнем доходу 40 тис. грн та вартістю майна 60 тис. грн. Підставляючи в (3.7) значення $x_1 = 40$, $x_2 = 60$, отримуємо $z = (1,\ 40,\ 60)^T$, $\hat{f}(40,\ 60) = 3{,}4295$. Тепер за допомогою співвідношення (3.6) будуємо 95% довірчий інтервал для функції регресії в зазначеній точці $f(x_1, x_2) \in (2{,}6199;\ 4{,}2391)$.

Перевіримо статистичну значущість параметру θ_1. Для цього обчислюємо $t_1 = \frac{\hat{\theta}_1}{\hat{s}\sqrt{(X^TX)^{-1}_{2,2}}} = 6{,}6915$. Оскільки $|t_1| > t_{2;0,05} = 4{,}3$, то на рівні значущості $q = 0{,}05$ ми відкидаємо гіпотезу про статистичну незначущість параметру θ_1 і приймаємо гіпотезу про його статистичну значущість.

4. Для обчислення оцінки частинного коефіцієнту кореляції $\bar{\rho}_{01}$ між випадковими величинами Y та X_1 обчислимо елементи вибіркової кореляційної матриці $\bar{R}$ за формулою (3.9). Отримаємо, що

$$\bar{R} = \begin{pmatrix} 1 & 0{,}8737 & -0{,}6822 \\ 0{,}8737 & 1 & -0{,}2715 \\ -0{,}6822 & -0{,}2715 & 1 \end{pmatrix}.$$

Тепер, з урахуванням нумерації її рядків та стовпців від 0 до p, знаходимо алгебраїчні доповнення

$$\bar{R}_{01} = -0{,}6885; \quad \bar{R}_{00} = 0{,}9263; \quad \bar{R}_{11} = 0{,}5346$$

відповідних елементів матриці $\bar{R}$. Далі за формулою (3.10) знаходимо шукану вибіркову оцінку частинного коефіцієнту кореляції $\bar{\rho}_{01} = 0{,}9784$.

Для перевірки статистичної значущості відповідного частинного коефіцієнту кореляції ρ_{01} за формулою (3.11) обчислюємо статистику $t = 6{,}6915$. Оскільки $|t| > t_{2;0,05} = 4{,}3$, то гіпотеза про незначуще відхиленні від нуля частинного коефіцієнту кореляції ρ_{01} відкидається на рівні значущості 0,05.

Завдання для самостійної роботи

В умовах задач 3.1 – 3.10 виконати такі завдання.

1. Оцінити параметри $\theta_0, \theta_1, \dots, \theta_p$ лінійної регресійної моделі вигляду (3.1) методом найменших квадратів за даними таблиці, записати рівняння лінійної регресії.
2. Перевірити отримане рівняння регресії на адекватність.
3. Побудувати довірчі інтервали для дисперсії залишків, параметрів регресійної моделі, а також для функції регресії за заданих значень регресорів (значення беруться з будь-якого рядка таблиці даних на вибір студента). Перевірити статистичну значущість параметрів регресії.
4. Обчислити оцінки частинних коефіцієнтів кореляції між випадковою величиною Y та випадковими величинами X_j. Перевірити статистичну значущість відповідних частинних коефіцієнтів кореляції.

3.1. В таблиці наведені річні прибутковості за час спостережень акцій компаній A, B, C, що працюють в одній галузі. Досліджується лінійна стохастична залежність прибутковості акцій компанії A від прибутковостей акцій компаній B та C.

Номер спостереження	Прибутковість компанії B	Прибутковість компанії C	Прибутковість компанії A
1	–5,31	–2,07	–2,54
2	16,84	19,34	26,50
3	0,07	3,63	4,44
4	10,03	13,28	17,12
5	4,98	7,68	10,19
6	7,52	10,51	13,88
7	0,23	3,68	4,55
8	5,53	8,78	10,28
9	5,94	8,67	11,76
10	6,09	8,60	11,89
11	0,93	3,74	5,14
12	3,22	6,25	7,70
13	2,08	5,67	7,17
14	2,81	5,84	7,57
15	10,73	13,54	17,46

3.2. На заводі безалкогольних напоїв досліджується залежність місячного обсягу реалізованої продукції Y від X_1, X_2, X_3, X_4 – витрат у попередньому місяці на теле-, радіо-, газетну та зовнішню рекламу відповідно. Є дані за дванадцять місяців.

Місяць	*Обсяг реалізації (тис. грн)*	*Витрати на рекламу (тис. грн.)*			
		Теле-рекламу	*Радіо-рекламу*	*Газетну рекламу*	*Зовнішню рекламу*
1	15304	133	35	38	27
2	17554	152	40	32	29
3	16876	130	48	35	28
4	16435	165	40	44	25
5	15229	125	42	48	18
6	16986	158	37	37	32
7	17914	165	50	43	38
8	16817	149	37	38	29
9	16579	169	33	28	27
10	15330	137	31	39	22
11	16781	178	42	42	18
12	17008	147	49	37	19

3.3. На хлібобулочному підприємстві досліджується залежність місячного обсягу реалізованої продукції Y від X_1, X_2, X_3, X_4 – витрат у попередньому місяці на теле-, радіо-, газетну і зовнішню рекламу відповідно. Є дані за дванадцять місяців.

Місяць	*Обсяг реалізації (тис. грн)*	*Витрати на рекламу (тис. грн)*			
		Теле-рекламу	*Радіо-рекламу*	*Газетну рекламу*	*Зовнішню рекламу*
1	14050	240	42	42	34
2	16310	263	47	44	36
3	15632	241	55	45	35
4	15126	276	47	42	32
5	13972	236	49	47	25
6	15753	272	44	45	39
7	16661	276	57	55	45
8	15584	260	46	47	36
9	15326	280	40	35	34
10	14077	248	38	38	29
11	15528	289	49	45	25
12	15755	258	56	52	26

3.4. Будується модель ціни автомобіля Y на вторинному ринку залежно від пробігу X_1, терміну експлуатації X_2 та об'єму двигуна X_3. Є дані про п'ятнадцять автомобілів тієї самої моделі:

Номер автомобіля	*Ціна автомобіля (у.о.)*	*Пробіг (тис. км)*	*Термін експлуатації (років)*	*Об'єм двигуна (л)*
1	12500	130	12	2,3
2	13700	120	10	1,9
3	9200	300	15	1,8
4	11400	180	13	2,1
5	15800	150	14	2,6
6	12300	80	8	1,7
7	16300	170	10	2,4
8	10200	210	11	1,9
9	11000	250	7	1,9
10	12700	150	9	1,7
11	15000	90	4	2,2
12	10500	230	13	2,4
13	17200	120	8	2,3
14	16000	110	9	2,5
15	17100	120	6	2,6

3.5. За тринадцятьма комерційними банками є дані, які характеризують залежність річного прибутку Y від X_1, X_2, X_3, X_4 – розміру власного капіталу, загальної суми залучених коштів, середньорічних ставок за гривневими депозитами та короткостроковими кредитами відповідно.

№ банку	*Прибуток (млн грн)*	*Власний капітал (млн грн)*	*Залучені кошти (млн грн)*	*Депозитна ставка (% річних)*	*Кредитна ставка (% річних)*
1	115	4428	3278	12,5	17,7
2	80	3756	5696	11,7	18,2
3	97	2970	2210	11,2	19,1
4	92	6231	5823	9,7	15,2
5	129	3960	4569	13,5	18,5
6	223	7354	2896	10,8	18,6
7	251	4662	3526	12,1	15,7
8	267	4760	2259	11,7	16,6
9	137	4569	4596	13,7	17,3

№ банку	*Прибуток (млн грн)*	*Власний капітал (млн грн)*	*Залучені кошти (млн грн)*	*Депозитна ставка (% річних)*	*Кредитна ставка (% річних)*
10	163	5274	3271	12,5	19,3
11	225	5418	4596	12,8	17,8
12	278	5359	3256	11,2	14,5
13	367	8254	5189	10,4	13,7

3.6. Для чотирнадцяти страхових компаній є дані, які характеризують залежність чистого річного прибутку $\boldsymbol{Y}$ від $\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}, \boldsymbol{X_4}$ – річних розмірів власних коштів, страхових резервів, страхових премій та страхових виплат відповідно (все в тис. грн):

Номер компанії	*Річний прибуток*	*Власні кошти*	*Страхові резерви*	*Страхові премії*	*Страхові виплати*
1	92	3444	9563	11456	1659
2	42	2658	6354	5249	2625
3	186	9723	10245	12968	4489
4	48	4526	6398	7589	6896
5	38	5369	5692	7256	5698
6	74	2248	6359	4963	4321
7	48	5671	6892	7259	6692
8	82	4312	7256	6935	756
9	45	2226	8256	2693	5532
10	46	3654	5982	6324	3235
11	65	2635	6359	7853	5325
12	29	2463	7532	8253	6862
13	34	3265	5632	7564	6325
14	66	7546	7625	9638	4569

3.7. На ливарному підприємстві досліджується вплив обсягів виробництва алюмінію $\boldsymbol{X_1}$, міді $\boldsymbol{X_2}$ та олова $\boldsymbol{X_3}$ на технологічні витрати електроенергії $\boldsymbol{Y}$. Є дані за одинадцять місяців.

Місяць	*Витрати електроенергії (тис. кВт·г)*	*Обсяг виробництва (тонн)*		
		Алюмінію	*Міді*	*Олова*
1	286	68	42	7
2	374	45	32	27
3	308	59	36	14
4	319	62	43	18
5	616	72	59	26
6	495	128	63	48
7	825	78	42	23
8	253	42	34	7
9	495	52	26	27
10	1056	96	75	36
11	902	122	56	43

3.8. Досліджується залежність ціни акцій компанії «Ласка» від цін акцій суміжних компаній «Хладик», «Рудь» та «Лімо». Є дані про результати біржових торгів за п'ятнадцять днів.

День	*Ціна акції (грн)*			
	«Ласка»	*«Хладик»*	*«Рудь»*	*«Лімо»*
1	67	30	59	75
2	63	27	55	68
3	58	22	52	78
4	66	27	55	60
5	63	18	78	72
6	67	33	68	73
7	70	32	72	71
8	63	24	75	78
9	60	29	56	72
10	64	27	79	65
11	56	32	62	82
12	68	32	79	71
13	71	56	75	72
14	62	40	72	78
15	64	37	68	75

3.9. На холодокомбінаті досліджується залежність місячного обсягу реалізації морозива $\boldsymbol{Y}$ від середньої ціни продукції $\boldsymbol{X_1}$, що випускається, витрат на рекламу $\boldsymbol{X_2}$, середньомісячної температури повітря $\boldsymbol{X_3}$ та місячного темпу інфляції $\boldsymbol{X_4}$. Є дані за дванадцять місяців.

Місяць	*Обсяг реалізації (тис. грн)*	*Ціна (грн)*	*Витрати на рекламу (тис. грн)*	*Темпера-тура повітря (°С)*	*Темп інфляції (%)*
1	185	8,3	6	2	0,3
2	162	8,3	7	4	0,4
3	182	8,9	5	7	0,3
4	195	10,6	5	10	0,2
5	226	10,7	7	13	0,7
6	279	10,8	22	18	0,9
7	312	12,2	12	22	0,9
8	286	14,2	17	24	0,4
9	212	14,5	22	17	0,1
10	178	13,7	26	13	0,1
11	182	13,3	8	8	0,5
12	173	12,1	4	5	0,9

3.10. Для фірми, що функціонує в умовах конкуренції, досліджується залежність обсягу пропозиції Y (в шт.) деякого товару від ціни X_1 (в у.о.) цього товару та заробітної плати X_2 (в у.о.) співробітників цієї фірми. Є дані за 16 місяців.

Місяць	*Пропозиція (шт.)*	*Ціна товару (у.о.)*	*Зарплата (у.о.)*
1	20	10	12
2	25	15	10
3	30	20	9
4	45	25	9
5	60	40	8
6	69	37	8
7	75	43	6
8	90	35	4
9	105	38	4
10	110	55	5
11	120	50	3
12	130	35	1
13	130	40	2
14	130	55	3
15	135	45	1
16	140	65	2

СПИСОК ЛІТЕРАТУРИ

1. Карташов М. В. Імовірність, процеси, статистика : посібник. Київ: ВПЦ «Київський університет», 2008. 504 с.
2. Руденко В. М. Математична статистика : навчальний посібник. Київ: Центр учбової літератури, 2012. 304 с.
3. Лебедєв Є. О. Математична статистика : навчальний посібник / Є. О. Лебедєв, Г. В. Лівінська, І. В. Розора, М. М. Шарапов. Київ: ВПЦ «Київський університет», 2016. 160 с.
4. Турчин В. М. Теорія ймовірностей та математична статистика. Дніпропетровськ: ІМА-прес, 2014. 556 с.
5. Henry Scheffe. The Analysis of Variance. NY: John Wiley & Sons, 1959. 477 p.
6. Жлуктенко В. І. Теорія ймовірностей і математична статистика : навчальний посібник у 2-х ч. Ч. 2 / В. І. Жлуктенко, С. І. Наконечний, С. С. Савіна. Київ: КНЕУ, 2007. 224 с.
7. Савченко О. Г. Теорія ймовірностей та математична статистика / О. Г. Савченко, Н. В. Валько, Г. М. Кавун, Л. В. Кузьмич. Херсон: РВЦ «Колос», 2017. 406 с.
8. Maurice G. Kendall. Inference and Relationship / Maurice G. Kendall, Alan Stuart. London: Charles Griffin & Co Limited, 1961. 690 p.

ДОДАТКИ
ТАБЛИЦІ МАТЕМАТИЧНОЇ СТАТИСТИКИ

Таблиця 1

Значення функції Лапласа

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

x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916

2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Таблиця 2

Значення $\chi^2_{k;q}$, відповідні ймовірності $q = P\left(\chi^2_k \geq \chi^2_{k;q}\right)$, де випадкова величина χ^2_k має розподіл χ^2 з k ступенями свободи

k	q									
	0,99	0,975	0,95	0,9	0,5	0,25	0,1	0,05	0,025	0,01
1	0,0002	0,0001	0,004	0,02	0,45	1,32	2,71	3,84	5,02	6,63
2	0,02	0,05	0,10	0,21	1,39	2,77	4,61	5,99	7,38	9,21
3	0,11	0,22	0,35	0,58	2,37	4,11	6,25	7,81	9,35	11,34
4	0,30	0,48	0,71	1,06	3,36	5,39	7,78	9,49	11,14	13,28
5	0,55	0,83	1,15	1,61	4,35	6,63	9,24	11,07	12,83	15,09
6	0,87	1,24	1,64	2,20	5,35	7,84	10,64	12,59	14,45	16,81
7	1,24	1,69	2,17	2,83	6,35	9,04	12,02	14,07	16,01	18,48
8	1,65	2,18	2,73	3,49	7,34	10,22	13,36	15,51	17,53	20,09
9	2,09	2,70	3,33	4,17	8,34	11,39	14,68	16,92	19,02	21,67
10	2,56	3,25	3,94	4,87	9,34	12,55	15,99	18,31	20,48	23,21
11	3,05	3,82	4,57	5,58	10,34	13,70	17,28	19,68	21,92	24,73
12	3,57	4,40	5,23	6,30	11,34	14,85	18,55	21,03	23,34	26,22
13	4,11	5,01	5,89	7,04	12,34	15,98	19,81	22,36	24,74	27,69
14	4,66	5,63	6,57	7,79	13,34	17,12	21,06	23,68	26,12	29,14
15	5,23	6,26	7,26	8,55	14,34	18,25	22,31	25,00	27,49	30,58
16	5,81	6,91	7,96	9,31	15,34	19,37	23,54	26,30	28,85	32,00
17	6,41	7,56	8,67	10,09	16,34	20,49	24,77	27,59	30,19	33,41
18	7,01	8,23	9,39	10,86	17,34	21,60	25,99	28,87	31,53	34,81
19	7,63	8,91	10,12	11,65	18,34	22,72	27,20	30,14	32,85	36,19
20	8,26	9,59	10,85	12,44	19,34	23,83	28,41	31,41	34,17	37,57
21	8,90	10,28	11,59	13,24	20,34	24,93	29,62	32,67	35,48	38,93
22	9,54	10,98	12,34	14,04	21,34	26,04	30,81	33,92	36,78	40,29
23	10,20	11,69	13,09	14,85	22,34	27,14	32,01	35,17	38,08	41,64
24	10,86	12,40	13,85	15,66	23,34	28,24	33,20	36,42	39,36	42,98
25	11,52	13,12	14,61	16,47	24,34	29,34	34,38	37,65	40,65	44,31
26	12,20	13,84	15,38	17,29	25,34	30,43	35,56	38,89	41,92	45,64
27	12,88	14,57	16,15	18,11	26,34	31,53	36,74	40,11	43,19	46,96
28	13,56	15,31	16,93	18,94	27,34	32,62	37,92	41,34	44,46	48,28
29	14,26	16,05	17,71	19,77	28,34	33,71	39,09	42,56	45,72	49,59
30	14,95	16,79	18,49	20,60	29,34	34,80	40,26	43,77	46,98	50,89

Таблиця 3

Значення $t_{k;q}$, відповідні ймовірності $q = P(|t_k| \geq |t_{k;q}|)$, де випадкова величина t_k має розподіл Стьюдента з k ступенями свободи

k	q							
	0,2	0,1	0,05	0,02	0,01	0,005	0,002	0,001
1	3,08	6,31	12,71	31,82	63,66	127,32	318,31	636,62
2	1,89	2,92	4,30	6,96	9,92	14,09	22,33	31,60
3	1,64	2,35	3,18	4,54	5,84	7,45	10,21	12,92
4	1,53	2,13	2,78	3,75	4,60	5,60	7,17	8,61
5	1,48	2,02	2,57	3,36	4,03	4,77	5,89	6,87
6	1,44	1,94	2,45	3,14	3,71	4,32	5,21	5,96
7	1,41	1,89	2,36	3,00	3,50	4,03	4,79	5,41
8	1,40	1,86	2,31	2,90	3,36	3,83	4,50	5,04
9	1,38	1,83	2,26	2,82	3,25	3,69	4,30	4,78
10	1,37	1,81	2,23	2,76	3,17	3,58	4,14	4,59
11	1,36	1,80	2,20	2,72	3,11	3,50	4,02	4,44
12	1,36	1,78	2,18	2,68	3,05	3,43	3,93	4,32
13	1,35	1,77	2,16	2,65	3,01	3,37	3,85	4,22
14	1,35	1,76	2,14	2,62	2,98	3,33	3,79	4,14
15	1,34	1,75	2,13	2,60	2,95	3,29	3,73	4,07
16	1,34	1,75	2,12	2,58	2,92	3,25	3,69	4,01
17	1,33	1,74	2,11	2,57	2,90	3,22	3,65	3,97
18	1,33	1,73	2,10	2,55	2,88	3,20	3,61	3,92
19	1,33	1,73	2,09	2,54	2,86	3,17	3,58	3,88
20	1,33	1,72	2,09	2,53	2,85	3,15	3,55	3,85
21	1,32	1,72	2,08	2,52	2,83	3,14	3,53	3,82
22	1,32	1,72	2,07	2,51	2,82	3,12	3,50	3,79
23	1,32	1,71	2,07	2,50	2,81	3,10	3,48	3,77
24	1,32	1,71	2,06	2,49	2,80	3,09	3,47	3,75
25	1,32	1,71	2,06	2,49	2,79	3,08	3,45	3,73
26	1,31	1,71	2,06	2,48	2,78	3,07	3,43	3,71
27	1,31	1,70	2,05	2,47	2,77	3,06	3,42	3,69
28	1,31	1,70	2,05	2,47	2,76	3,05	3,41	3,67
29	1,31	1,70	2,05	2,46	2,76	3,04	3,40	3,66
30	1,31	1,70	2,04	2,46	2,75	3,03	3,39	3,65

Таблиця 4

Значення $F_{n_1;n_2;q}$, відповідні ймовірності $q = P\big(F_{n_1;n_2} > F_{n_1;n_2;q}\big)$, де випадкова величина має $F_{n_1;n_2}$ розподіл Фішера з n_1, n_2 ступенями свободи

$q = 0{,}05$										
$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16

$q = 0{,}05$										
n_2 \ n_1	11	12	13	14	15	16	17	18	19	20
1	242,98	243,91	244,69	245,36	245,95	246,46	246,92	247,32	247,69	248,01
2	19,40	19,41	19,42	19,42	19,43	19,43	19,44	19,44	19,44	19,45
3	8,76	8,74	8,73	8,71	8,70	8,69	8,68	8,67	8,67	8,66
4	5,94	5,91	5,89	5,87	5,86	5,84	5,83	5,82	5,81	5,80
5	4,70	4,68	4,66	4,64	4,62	4,60	4,59	4,58	4,57	4,56
6	4,03	4,00	3,98	3,96	3,94	3,92	3,91	3,90	3,88	3,87
7	3,60	3,57	3,55	3,53	3,51	3,49	3,48	3,47	3,46	3,44
8	3,31	3,28	3,26	3,24	3,22	3,20	3,19	3,17	3,16	3,15
9	3,10	3,07	3,05	3,03	3,01	2,99	2,97	2,96	2,95	2,94
10	2,94	2,91	2,89	2,86	2,85	2,83	2,81	2,80	2,79	2,77
11	2,82	2,79	2,76	2,74	2,72	2,70	2,69	2,67	2,66	2,65
12	2,72	2,69	2,66	2,64	2,62	2,60	2,58	2,57	2,56	2,54
13	2,63	2,60	2,58	2,55	2,53	2,51	2,50	2,48	2,47	2,46
14	2,57	2,53	2,51	2,48	2,46	2,44	2,43	2,41	2,40	2,39
15	2,51	2,48	2,45	2,42	2,40	2,38	2,37	2,35	2,34	2,33
16	2,46	2,42	2,40	2,37	2,35	2,33	2,32	2,30	2,29	2,28
17	2,41	2,38	2,35	2,33	2,31	2,29	2,27	2,26	2,24	2,23
18	2,37	2,34	2,31	2,29	2,27	2,25	2,23	2,22	2,20	2,19
19	2,34	2,31	2,28	2,26	2,23	2,21	2,20	2,18	2,17	2,16
20	2,31	2,28	2,25	2,22	2,20	2,18	2,17	2,15	2,14	2,12
21	2,28	2,25	2,22	2,20	2,18	2,16	2,14	2,12	2,11	2,10
22	2,26	2,23	2,20	2,17	2,15	2,13	2,11	2,10	2,08	2,07
23	2,24	2,20	2,18	2,15	2,13	2,11	2,09	2,08	2,06	2,05
24	2,22	2,18	2,15	2,13	2,11	2,09	2,07	2,05	2,04	2,03
25	2,20	2,16	2,14	2,11	2,09	2,07	2,05	2,04	2,02	2,01
26	2,18	2,15	2,12	2,09	2,07	2,05	2,03	2,02	2,00	1,99
27	2,17	2,13	2,10	2,08	2,06	2,04	2,02	2,00	1,99	1,97
28	2,15	2,12	2,09	2,06	2,04	2,02	2,00	1,99	1,97	1,96
29	2,14	2,10	2,08	2,05	2,03	2,01	1,99	1,97	1,96	1,94
30	2,13	2,09	2,06	2,04	2,01	1,99	1,98	1,96	1,95	1,93

Електронне навчальне видання комбінованого використання
Можна використовувати в локальному та мережному режимі

Півень Олексій Леонідович
Сморцова Тетяна Іванівна

ДИСПЕРСІЙНИЙ АНАЛІЗ ТА МНОЖИННИЙ РЕГРЕСІЙНИЙ АНАЛІЗ

Методичні рекомендації
до розв'язання задач
з курсу «Прикладні задачі математичної статистики»

Електронне видання

Коректор *О. В. Пікалова*
Комп'ютерне верстання *Т. І. Сморцова*

Підписано до видання 20.04.2023 р.
Гарнітура Times New Roman. Обсяг 1,0 Мб. Зам. № 64/23.

Харківський національний університет імені В. Н. Каразіна,
61022, м. Харків, майдан Свободи, 4.
Свідоцтво суб'єкта видавничої справи ДК № 3367 від 13.01.2009

Видавництво ХНУ імені В. Н. Каразіна