

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
МИКОЛАЇВСЬКИЙ НАЦІОНАЛЬНИЙ АГРАРНИЙ УНІВЕРСИТЕТ

Факультет агротехнологій
Кафедра землеробства, геодезії та землеустрою

ОСНОВИ ДОСЛІДНИЦЬКОЇ СПРАВИ

**Методичні рекомендації до виконання практичних робіт
для здобувачів вищої освіти освітнього ступеня
«Молодший бакалавр» початкового рівня (короткий цикл)
спеціальності 201 «Агрономія» денної форми навчання**



МИКОЛАЇВ
2021

УДК 001.891:631/635

О-75

Друкується за рішенням науково-методичної комісії факультету агротехнологій Миколаївського національного аграрного університету від 13 травня 2021 р., протокол № 9.

Укладач:

І. В. Смірнова – асистент кафедри землеробства, геодезії та землеустрою Миколаївський національний аграрний університет.

Рецензенти:

О. М. Дробітько – канд. с.-г. наук, голова фермерського господарства «Олена» Братського району Миколаївської області;

М. І. Федорчук – д-р с.-г. наук, професор кафедри рослинництва та садово-паркового господарства, Миколаївський національний аграрний університет.

© Миколаївський національний аграрний університет, 2021

ЗМІСТ

Вступ	4
Практична робота 1. Обчислення статистичних характеристик малої вибірки за кількісної мінливості	6
Практична робота 2. Обчислення статистичних характеристик великої вибірки за кількісної мінливості	10
Практична робота 3. Обчислення статистичних характеристик вибірки при якісній мінливості	12
Практична робота 4. Підготовка даних про врожайність до статистичного аналізу	15
Практична робота 5. Оцінити істотність різниці вибіркових середніх за t-критерієм та за найменшою істотною різницею (НІР)	19
Практична робота 6. Дисперсійний аналіз однофакторного вегетаційного дослідження	23
Практична робота 7. Дисперсійний аналіз двофакторного вегетаційного дослідження	29
Практична робота 8. Дисперсійний аналіз однофакторного польового дослідження	32
Практична робота 9. Кореляційний і регресійний аналіз прямолінійної залежності	37
Практична робота 10. Аналіз криволінійної залежності та складання рівнянь регресії для криволінійної залежності	40
Додатки	43
Список рекомендованої літератури	46

ВСТУП

Ефективність і якість наукової роботи, результативність досліджень в агрономії визначається методичним рівнем планування і постановки польових і лабораторних експериментів та методами проведення статистичної обробки експериментальних даних.

Метою навчальної дисципліни «Основи дослідницької справи» є сформувати у здобувачів вищої освіти освітнього ступеня «Молодший бакалавр» початкового рівня (короткий цикл) спеціальності 201 «Агрономія» систему знань і навичок з методів і організації проведення досліджень у сфері землеробства, рослинництва, агрохімії, фізіології рослин.

Завдання курсу – освоїти і закріпити на практичних заняттях найважливіші розділи дисципліни, в тому числі :

- основні поняття і елементи методики польового досліджу;
- розміщення варіантів у польовому досліді;
- планування польового досліджу;
- техніка закладання та проведення польового досліджу;
- документація та звітність в науково-дослідній роботі;
- математична статистика, емпіричні та теоретичні розподіли;
- розрахунки статистичних характеристик;
- статистичні методи перевірки гіпотез;
- дисперсійний аналіз одно- та багатofакторних дослідів;
- кореляція, регресія, складання рівнянь регресії для лінійної та криволінійної залежностей.

У результаті вивчення дисципліни студент повинен **знати**:

- сутність загальнонаукових і спеціальних методів досліджень в агрономії;
- польовий дослід як основний метод в агрономії, принципи його планування та проведення;
- методику і техніку закладання польового досліджу;
- зміст спостережень у польовому досліді;
- особливості закладання та проведення інших спеціальних методів дослідження в агрономії;
- методику виконання статистичного аналізу експериментальних даних і використання його результатів для їх інтерпретації.

На підставі набутих знань студент повинен **уміти**:

- закласти польовий, вегетаційний чи лізиметричний досліді;
- відповідно до програми досліджень провести в них обліки і

- спостереження;
- здійснити статистичний аналіз експериментальних даних відповідно до обраного методу і дати оцінку якості проведеному дослідженню;
 - вести необхідну документацію дослідів та складати на її основі науковий звіт.

Навчальна дисципліна «Основи дослідницької справи» складається з трьох блоків змістових модулів (основних її розділів), які містять у собі близькі за змістом теми лекцій та практичних занять, індивідуальні завдання та інші організаційні форми навчального процесу.

ПРАКТИЧНА РОБОТА № 1

Тема: Обчислення статистичних характеристик малої вибірки за кількісної мінливості

Завдання: Обчислити статистичні характеристики малих вибірок за вихідними даними, наведеними в таблиці 1.1.

Таблиця 1.1

Урожайність соняшнику залежно від рівномірності внесення добрив, т/га

Внесення добрив	Повторність			
	I	II	III	IV
Рівномірне	1,6	1,8	1,5	1,7
Нерівномірне	1,8	1,0	2,2	1,7

Приклад: Обчислити статистичні характеристики малих вибірок за вихідними даними, наведеними в таблиці 1.2.

Для малих вибірок обчислюють такі статистичні характеристики: середні арифметичні, дисперсії, стандартні відхилення, коефіцієнти варіювання, помилки вибірових середніх, граничні оцінки середніх арифметичних, відносні помилки вибірових середніх, точність досліду.

Таблиця 1.2

Урожайність соняшнику залежно від рівномірності внесення добрив, т/га

Внесення добрив	Повторність			
	I	II	III	IV
Рівномірне	1,6	1,8	1,5	1,5
Нерівномірне	1,8	0,7	2,2	1,7

Середня арифметична (\bar{x}). Для обчислення цієї характеристики варіюючі ознаки (результати спостережень) позначають знаком X , а кількість повторностей – n . Тоді середню арифметичну розраховують за формулою

$$\bar{X} = \frac{\sum X}{n}$$

Для варіанта з рівномірним внесенням добрив середня арифметична

$$\bar{X}_1 = \frac{\sum X}{n} = \frac{1,6 + 1,8 + 1,5 + 1,5}{4} = \frac{6,4}{4} = 1,6 \text{ т/га.}$$

Для варіанта з нерівномірним внесенням добрив середня арифметична

$$\bar{X}_2 = \frac{\sum X}{n} = \frac{1,8+0,7+2,2+1,7}{4} = \frac{6,4}{4} = 1,6 \text{ т/га.}$$

Отже, середні арифметичні однакові, але **розмах варіювання (R)** в них різний. За рівномірного внесення добрив $R_1 = X_{max} - X_{min} = 1,8 - 1,5 = 0,3$, а за нерівномірного $R_2 = X_{max} - X_{min} = 2,2 - 0,7 = 1,5$, тобто урожайність соняшнику сильніше варіює за нерівномірного внесення добрив, коли створюється строкатість родючості ґрунту.

Середня арифметична є основною статистичною характеристикою кожного варіаційного ряду, а всі інші характеристики лише пояснюють основну.

Дисперсія (S^2) - це середній квадрат відхилень кожного члена варіаційного ряду (X_1, X_2, \dots, X_n) від середньої арифметичної; це показник, який повніше за розмах варіації характеризує варіаційні ряди. Дисперсія обчислюється за формулою

$$S^2 = \frac{\sum (X - \bar{x})^2}{n - 1}.$$

Для її розрахунків складається допоміжна таблиця 1.3, в яку вносять результати спостережень (дані з таблиці 1.2).

Таблиця 1.3

Обчислення квадратів відхилень від середньої арифметичної

Внесення добрив							
Рівномірне				Нерівномірне			
n	X	$X - \bar{x}$	$(X - \bar{x})^2$	n	X	$X - \bar{x}$	$(X - \bar{x})^2$
1	1,6	0	0	1	1,8	0,2	0,04
2	1,8	0,2	0,04	2	0,7	-0,9	0,81
3	1,5	-0,1	0,01	3	2,2	0,6	0,36
4	1,5	-0,1	0,01	4	1,7	0,1	0,01
	$\bar{x} = 1,6$	$\sum (X - \bar{x}) = 0$	$\sum (X - \bar{x})^2 = 0,06$		$\bar{x} = 1,6$	$\sum (X - \bar{x}) = 0$	$\sum (X - \bar{x})^2 = 1,22$

Підставивши одержані дані (суми квадратів відхилень) у наведену формулу отримаємо такі дисперсії:

$$S_1^2 = \frac{0,06}{4-1} = 0,02; \quad S_2^2 = \frac{1,22}{4-1} = 0,41.$$

Їх порівняння показує, що в другому варіанті, де добрива внесені нерівномірно, дисперсія була в двадцять разів більшою, ніж

за рівномірного внесення добрив.

Дисперсія використовується не лише для характеристики варіювання досліджуваних показників, а й для обчислення стандартного відхилення (S).

Стандартне відхилення (S) обчислюється за формулою

$$S = \sqrt{S^2}.$$

Для першого варіанта $S_1 = \sqrt{S_1^2} = \sqrt{0,02} = 0,14$ т/га, а для другого $S_2 = \sqrt{S_2^2} = \sqrt{0,41} = 0,64$ т/га.

Стандартне відхилення виражається у тих же одиницях, що і середня арифметична.

Якщо у дослідженнях порівнюють мінливість ознак, що мають різні одиниці виміру (центнери, штуки, сантиметри тощо), то дисперсія і також стандартне відхилення для таких порівнянь непридатні. У таких випадках доцільно користуватися коефіцієнтом варіювання.

Коефіцієнт варіювання (V) – це відношення стандартного відхилення до середньої арифметичної, що виражається у відсотках і обчислюється за формулою

$$V\% = \frac{S \cdot 100}{x}.$$

Для першого варіанта $V_1 = \frac{0,14 \cdot 100}{1,6} = 8,75\%$, для другого

$$V_2 = \frac{0,64 \cdot 100}{1,6} = 40,0\%.$$

Отже, строкастість родючості ґрунту, призводить до невіривності урожайності соняшнику, тому що варіювання врожаю значно більше за нерівномірного внесення добрив.

Правило: варіювання умовно вважають незначним, якщо коефіцієнт його не перевищує 10%; середнім, коли коефіцієнт варіювання перебуває в межах 10-20%; значним, коли він перевищує 20%.

Варіювання врожаю більшості польових культур становить 8-12%. Менше варіювання врожаю у культур звичайного рядкового способу сівби, більше – у просапних.

Середні арифметичні мають свої помилки, які спричиняються внаслідок неповного представництва вибіркової сукупності. Ці помилки властиві лише вибіркового методу досліджень, а їх

чисельне значення залежить від ступеня мінливості досліджуваних ознак і обсягів вибірки.

Помилку вибіркової середньої ($S_{\bar{x}}$) обчислюють за формулою

$$S_{\bar{x}} = \sqrt{\frac{\sum (X - \bar{x})^2}{n(n-1)}} = \sqrt{\frac{S^2}{n}}.$$

У наведеному прикладі для першого варіанту

$$S_{\bar{x}_1} = \sqrt{\frac{0,02}{4}} = 0,07, \quad \text{а для другого} \quad S_{\bar{x}_2} = \sqrt{\frac{0,41}{4}} = 0,32.$$

Значення помилок використовують для **інтервальної оцінки середніх арифметичних** за формулою

$$\bar{x} \pm t \cdot S_{\bar{x}}.$$

За оцінки на рівні імовірності $P_{0,95}$ значення t дорівнює 3,18, а за $P_{0,99}$ – 5,84.

З урахуванням цих показників межеве значення середньої арифметичної для першого варіанта становить

$$1,6 \pm 3,18 \cdot 0,07 \quad \text{та} \quad 1,6 \pm 5,84 \cdot 0,07.$$

Це означає, що на рівні $P_{0,95}$ межові величини середньої арифметичної будуть $1,38 \div 1,82$, а на рівні $P_{0,99}$ – $1,21 \div 1,99$.

Відносною помилкою вибіркової середньої називають відношення помилки вибіркової середньої арифметичної, вираженої у відсотках, і визначають за формулою

$$S_{\bar{x}} \% = \frac{S_{\bar{x}} \cdot 100}{\bar{x}}.$$

Для варіанта з рівномірним внесенням добрив відносна помилка

$$\text{вибіркової середньої} \quad S_{\bar{x}_1} \% = \frac{0,07 \cdot 100}{1,6} = 4,38\%, \quad \text{а для варіанта з}$$

$$\text{нерівномірним внесенням добрив} \quad S_{\bar{x}_2} \% = \frac{0,32 \cdot 100}{1,6} = 20,0\%.$$

Залежно від значення відносної помилки роблять висновки про точність дослідження – це різниця між 100 % і відносною помилкою ($T = 100\% - S_{\bar{x}}\%$).

Правило: умовно точність вважають високою, якщо її значення не перевищує 3 %, середньою – коли воно становить 3-6 % і низькою коли перевищує 7 %.

Для збільшення високої точності дослідження, слід дбати про те, щоб дослідження проводилися на вирівняних за родючістю площах.

ПРАКТИЧНА РОБОТА № 2

Тема: Обчислення статистичних характеристик великої вибірки за кількісної мінливості

Завдання: Обчислити статистичні характеристики великої вибірки за даними по кількості личинок клопа черепашки на 1 м². Варіаційний ряд включає 50 повторень: 33, 29, 11, 38, 63, 14, 34, 10, 53, 2, 37, 17, 22, 13, 50, 28, 16, 33, 33, 28, 2, 23, 39, 48, 32, 31, 26, 27, 46, 18, 53, 32, 43, 59, 26, 70, 39, 29, 36, 37, 25, 47, 69, 31, 39, 49, 47, 35, 57, 26.

Приклад: Обчислити статистичні характеристики великої вибірки, яка включає довжину стебла 40 рослин пшениці озимої (таблиця 2.1).

Розміщують довжину стебла 40 рослин у зростаючому порядку: 27, 31, 33, 33, 35, 36, 36, 36, 36, 36, 36, 36, 36, 36, 36, 37, 37, 37, 37, 37, 37, 37, 37, 37, 38, 38, 38, 38, 38, 38, 38, 38, 39, 39, 39, 40, 40, 41, 41, 42, 43, 44.

Визначаємо число груп

$$C_2 = \sqrt{n} = \sqrt{40} = 6.$$

Визначаємо інтервал групи

$$i = \frac{X_{\max} - X_{\min}}{C_2} = \frac{44 - 27}{6} = 3.$$

Таблиця 2.1

Допоміжна таблиця для обробки варіаційного ряду великої вибірки довжини стебла пшениці озимої

Інтервал групи	Середнє значення групи, X	Частота f	f · X	X ²	f · X ²
27-29	28	1	28	784	784
30-32	31	1	31	961	961
33-35	34	3	102	1156	3468
36-38	37	25	925	1369	34225
39-41	40	7	280	1600	11200
42-44	43	3	129	1849	5547
	$\bar{x} = 37,4$	$\Sigma f = 40$	$\Sigma fX = 1495$	$\Sigma X^2 = 7719$	$\Sigma fX^2 = 56185$

Розрахунки проводять у такій послідовності:

Середня арифметична

$$\bar{x} = \frac{\Sigma f \cdot X}{n} = \frac{1495}{40} = 37,4 \text{ см.}$$

Дисперсія

$$S^2 = \frac{\sum f(X - \bar{x})^2}{n-1} = \frac{\sum fX^2 - (\sum fX)^2 \div n}{n-1} = \frac{56185 - 1495^2 \div 40}{39} = 7,95$$

Стандартне відхилення

$$S = \sqrt{S^2} = \sqrt{7,95} = 2,82 \text{ см.}$$

Коефіцієнт варіації

$$V = \frac{S \cdot 100}{\bar{x}} = \frac{2,82 \cdot 100}{37,4} = 7,54\%.$$

Помилка вибіркової середньої

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{2,82}{\sqrt{40}} = \frac{2,82}{6,32} = 0,45 \text{ см.}$$

Відносна помилка вибіркової середньої

$$S_{\bar{x}}\% = \frac{S_{\bar{x}} \cdot 100}{\bar{x}} = \frac{0,45 \cdot 100}{37,4} = 1,20\%.$$

Інтервальну оцінку середньої на рівнях імовірності складають:
за $P_{0,95}$:

$$\begin{aligned} & \bar{x} \pm t_{0,95} \cdot S_{\bar{x}}; \\ & 37,4 \pm 2,04 \cdot 0,45; \\ & 37,4 \pm 0,92(36,48 \div 38,32). \end{aligned}$$

за $P_{0,99}$:

$$\begin{aligned} & \bar{x} \pm t_{0,99} \cdot S_{\bar{x}}; \\ & 37,4 \pm 2,75 \cdot 0,45; \\ & 37,4 \pm 1,24 \cdot (36,16 \div 38,64). \end{aligned}$$

Висновки:

1. Середня арифметична висоти рослин озимої пшениці дорівнює 37,4 см.
2. Коефіцієнт варіації 7,54% свідчить про незначне варіювання висоти рослин.
3. Значення відносної помилки 1,2% свідчить про досить високу точність обчислення середньої арифметичної.
4. До даного варіаційного ряду на рівні $P_{0,95}$ належать рослини висотою 36,5-38,3 см, а на рівні $P_{0,99}$ – 36,2-38,6 см. Усі інші дані, що не ввійшли до інтервалу оцінки на обох рівнях надійної імовірності, не належать до даного варіаційного ряду і вважаються нехарактерними для нього.

ПРАКТИЧНА РОБОТА №3

Тема: Обчислення статистичних характеристик вибірки за якісної мінливості

Завдання: Після збирання картоплі виявилось, що у сорту Луговська із 100 бульб (N_1) не ураженими фітофторозом було 85 (n_1), а у сорту Слов'янка із 100 бульб (N_2) – лише 97 (n_2).

Приклад: Після збирання картоплі виявилось, що у сорту Гатчинський із 100 бульб (N_1) якісними були 80 (n_1), а у сорту Іскра із 100 бульб (N_2) – лише 70 (n_2).

Обчислити статистичні характеристики.

Для аналізу варіаційних рядів якісної мінливості обчислюють такі статистичні характеристики: частку наявності ознаки (p); частку відсутності ознаки (q); показник мінливості якісної ознаки – стандартне відхилення (S); коефіцієнт варіації (V) і помилку частки (S_p).

Частка наявності ознаки (p) – це відношення кількості об'єктів з даною ознакою (n) до загального обсягу вибірки (N). Обчислюють її за формулою $p = n : N$.

Визначають частку наявності ознаки p_1 і p_2 у кожного з наведених сортів:

$$P_1 = n_1 \div N_1 = 80 \div 100 = 0,80 \text{ або } 80\%;$$

$$P_2 = n_2 \div N_2 = 70 \div 100 = 0,70 \text{ або } 70\%.$$

Частка відсутності ознаки – це різниця між одиницею і часткою наявності ознаки. Цю частку обчислюють за формулою $q = 1 - p_1$. Для досліджуваних сортів Гатчинська і Іскра частка відсутності ознаки q_1 і q_2 відповідно становитиме:

$$q_1 = 1 - p_1 = 1 - 0,8 = 0,2 \text{ частка або } 20\%;$$

$$q_2 = 1 - p_2 = 1 - 0,7 = 0,3 \text{ частка або } 30\%.$$

Якщо досліджуваний об'єкт має лише дві градації, **показник мінливості S** обчислюють за формулою $S = \sqrt{P \cdot q}$.

Для сорту Гатчинська $S = \sqrt{p_1 \cdot q_1} = \sqrt{0,2 \cdot 0,8} = \sqrt{0,16} = 0,40$, а для сорту Іскра $S = \sqrt{p_2 \cdot q_2} = \sqrt{0,3 \cdot 0,7} = \sqrt{0,21} = 0,46$.

Максимальне значення (0,5) мінливості мають за умови, якщо $p = q = 0,5$ ($S_{max} = \sqrt{p \cdot q} = \sqrt{0,5 \cdot 0,5} = 0,5$).

Якщо об'єкт досліджень має не дві, а більше градацій, то для такої вибірки показник якісної мінливості (S) обчислюють за

формулою $S = \sqrt[n]{p_1 p_2 \dots p_n}$, де p_1, p_2, \dots, p_n – частка ознак із загального обсягу вибірки; n – кількість градацій ознак.

Показник мінливості використовується для визначення **коефіцієнта варіювання (V_p)** як відношення показника мінливості (S) до його максимального значення (S_{max}) вираженого в процентах.

Обчислюється коефіцієнт варіювання в процентах за формулою

$$V_p = \frac{S \cdot 100}{S_{max}}$$

Для сорту Гатчинська коефіцієнт варіювання становитиме:

$$V_{p1} = \frac{S_1 \cdot 100}{S_{max}} = \frac{0,4 \cdot 100}{0,5} = 80\%,$$

для сорту Іскра

$$V_{p2} = \frac{S_2 \cdot 100}{S_{max}} = \frac{0,46 \cdot 100}{0,5} = 92\%.$$

Максимальне значення коефіцієнта варіювання – 100% буває за $S = S_{max} = 0,5$.

Для оцінки точності у визначенні вибірових середніх арифметичних за якісної мінливості вираховують **помилку вибіркової середньої арифметичної** за формулою

$$S_p = \frac{S}{\sqrt{N}}$$

Стосовно до альтернативної мінливості вона матиме вигляд

$$S_p = \frac{S}{\sqrt{N}} = \frac{\sqrt{pq}}{N} = \sqrt{\frac{pq}{N}}$$

Для сорту Гатчинська $S_{p1} = \sqrt{\frac{p_1 q_1}{N}} = \sqrt{\frac{0,8 \cdot 0,2}{100}} = 0,040$, а для

сорту Іскра $S_{p2} = \sqrt{\frac{p_2 q_2}{N}} = \sqrt{\frac{0,7 \cdot 0,3}{100}} = 0,046$.

Інтервальну оцінку помилки вибіркової середньої роблять за формулою $P \pm t \cdot S_p$.

На рівні імовірності $P_{0,95}$ значення $t=1,98$, а на рівні $P_{0,99}$ $t=2,63$.

Для сорту Гатчинська ці інтервали становитимуть для рівня $P_{0,95}$ $0,80 \pm 1,98 \cdot 0,04$ тобто $0,80 \pm 0,08 (0,72 \div 0,88)$, а для рівня $P_{0,99}$ $0,80 \pm 2,63 \cdot 0,04$ тобто $0,8 \pm 0,105 (0,72 \div 0,90)$.

Для сорту Іскра на цих рівнях, відповідно, інтервали

становитимуть $0,70 \pm 1,98 \cdot 0,046$ або $0,70 \pm 0,09(0,61 \div 0,79)$ і $0,70 \pm 2,63 \cdot 0,046$ або $0,70 \pm 0,12(0,58 \div 0,82)$.

Висновок: результати спостережень дозволять вважати, що частка здорових бульб картоплі сорту Гатчинська може становити 72-88%, а сорту Іскра – 61-79%. Це означає, що сорт Гатчинський стійкіший до хвороб, ніж сорт Іскра.

Для визначення достовірності різниці між частками наявності ознак обчислюють **фактичний критерій Стюдента (t_{ϕ})**

$$t_{\phi} = \frac{|P_1 - P_2|}{\sqrt{\frac{p_1 q_1}{N_1} + \frac{p_2 q_2}{N_2}}} = \frac{0,80 - 0,70}{\sqrt{\frac{0,80 \cdot 0,20}{100} + \frac{0,70 \cdot 0,3}{100}}} = \frac{0,10}{\sqrt{0,0016 + 0,0021}} = \frac{0,10}{0,06} = 1,66.$$

Фактичний критерій Стюдента порівнюють з теоретичним, який приймають за додатком 1 за числом ступенів вільності $\gamma = 100$.

$$\gamma_p = (N_1 - 1) + (N_2 - 1) = (100 - 1) + (100 - 1) = 99 + 99 = 198.$$

При γ_p , що дорівнює 100 і більше, критерій $t_{0,95} = 1,96$, а $t_{0,99} = 2,58$.

Правило: якщо фактичний критерій Стюдента t_{ϕ} дорівнює теоретичному або більший за нього, то різниця достовірна і навпаки.

Висновок: оскільки критерій Стюдента фактичний між варіантами становить 1,66, що значно менше теоретичних критеріїв на обох рівнях надійної імовірності, то у картоплі сорту Іскра зниження здорових бульб недостовірне порівняно з сортом Гатчинський.

ПРАКТИЧНА РОБОТА №4

Тема: Підготовка даних про врожайність до статистичного аналізу

Завдання:

1. Обчислити середню арифметичну зважену, якщо з площі 5 га зібрали по 4,23 тон пшениці озимої, а з площі 8 га – по 5,37 тон.
2. Провести бракування сумнівних дат, якщо у досліді, де вивчався вплив доз добрив на урожайність пшениці озимої за повторностями, вона була 6,58; 3,35; 5,62; 5,58 т/га.
3. Відновити втрачену дату, якщо у досліді, де вивчався вплив різних попередників на урожайність пшениці озимої, цю урожайність у різних повтореннях показано у таблиці 4.1.

Таблиця 4.1

Урожайність пшениці озимої після попередників, т/га

Номер варіанта	Попередники	Повторність			
		I	II	III	IV
1	Багаторічні трави на один укіс	$X_{\text{відн}}$	5,27	5,31	5,29
2	Вико-овсяна сумішка	4,58	4,63	4,89	4,87
3	Горох	4,53	4,67	4,75	4,91
4	Кукурудза на силос	2,53	2,64	2,78	2,45

Заокруглення чисел експерименту слід представляти тризначними числами. Наприклад: урожайність буряків цукрових 52,8 т/га; пшениці озимої – 5,34 т/га; насіння люцерни 0,317 т/га. Показники, менші за одиницю, виражаються тисячними – 0,529.

Число округлюється до більшого, якщо після нього стоять цифри 5 і більше, та навпаки.

Наприклад, число 0,8523 округлюється до 0,852, а число 0,8545 – до 0,855.

Обчислення середніх арифметичних

Прості середні арифметичні обчислюються як результат ділення суми спостережень на їх кількість

$$\bar{x} = \frac{\sum X}{n}$$

Однак, у досліді трапляються ситуації, коли різні рівні врожайності культури стосуються різних площ. Наприклад, з площі 3 га зібрали по 3,73 центнерів пшениці озимої, а з площі 5 га – по 4,87

центнерів. Проста середня арифметична склала б $\frac{3,73+4,87}{2} = 4,3$ т/га.

Але оскільки площі різні, то слід обчислювати середню арифметичну зважену ($x_{зв}$) за формулою

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{\sum f},$$

де x_1, x_2, \dots, x_n – варіююча ознака (у нашому прикладі врожайність пшениці озимої);

f – частота (площа посіву пшениці озимої певного варіанта).

Підставивши у формулу чисельні значення цих показників отримуємо

$$\bar{x}_{зв} = \frac{3,73 \cdot 3 + 4,87 \cdot 5}{3 + 5} = \frac{11,19 + 24,35}{8} = \frac{35,54}{8} = 4,44 \text{ т/га.}$$

Отриманий результат суттєво відрізняється від визначеного за формулою середньої арифметичної простої, що вказує на необхідність користування в цих випадках формулою середньої зваженої.

Бракування сумнівних дат. За аналізу даних у межах кожного варіанта (за повторностями) деякі з них можуть значно відрізнятися від інших і викликати сумнів щодо їх належності до певних варіаційних рядів. Сумнівні дати можна об'єктивно бракувати лише методами математичної статистики.

Наприклад, у досліді, де вивчався вплив доз добрив на урожайність пшениці озимої за повторностями, вона була 6,24; 4,57; 5,32; 5,58 т/га.

Щоб установити, що всі ці дані належать до одного варіаційного ряду, їх числові значення розміщують у зростаючому порядку: 4,57; 5,32; 5,58; 6,24.

Найбільш сумнівними є найменша дата – 4,57 та найбільша – 6,24.

Для перевірки їх сумнівності кожній з дат дають відповідний номер – 4,57 (X_1); 5,32 (X_2); 5,58 (X_{n-1}); 6,24 (X_n) і обчислюють критерій τ_n за формулами:

$$\tau_1 = \frac{X_2 - X_1}{X_{n-1} - X_1} = \frac{5,32 - 4,57}{5,58 - 4,57} = \frac{0,75}{1,01} = 0,743,$$

$$\tau_n = \frac{X_n - X_{n-1}}{X_n - X_2} = \frac{6,24 - 5,58}{6,24 - 5,32} = \frac{0,66}{0,92} = 0,717.$$

Розрахункові критерії τ порівнюють з їх теоретичними значеннями і роблять висновки за таким правилом: якщо розрахункові критерії ($\tau_1 \tau_2$) більші за теоретичні або дорівнюють їм, то дата (спостереження), що перевіряється, є сумнівною і її треба вибракувати. Теоретичні значення критеріїв в τ приймають за додатком 2 згідно з числом повторностей (n) і рівнем надійної імовірності $P_{0,95}$ чи $P_{0,99}$. При $n = 4$ критерії τ теоретичні відповідно становлять $\tau_{0,95}=0,955$ і $\tau_{0,99}=0,991$.

Висновки:

1. Оскільки $\tau_1=0,743$ менше $\tau_{0,95}(0,955)$ та $\tau_{0,99}(0,991)$, то дата 4,57 не викликає сумніву і її не слід вибракувувати.
2. Оскільки $\tau_n = 0,717$ теж менше $\tau_{0,95}(0,955)$ і $\tau_{0,99}(0,991)$, то дата 6,24 теж не викликає сумніву і її не слід вибракувувати.

Слід зазначити, що бракування дат за наведеними формулами можливе, якщо кількість повторностей у досліді становить не менше 4 та коли $X_1 \neq X_2$, а $X_n \neq X_{n-1}$, тому що при цьому дати не можуть бути сумнівними, отже і не потребують перевірки.

Відновлення втрачених дат. Унаслідок випадання дат на деяких ділянках певною мірою ускладнюється статистичний аналіз дослідів.

Причинами випадання можуть бути сильні зливи (дуже замулюють окремі ділянки), град (випадає смугами), випадкове пошкодження зернових культур і соняшнику птахами, шкідниками, хворобами, наїзди транспорту на придорожні ділянки тощо. Випадання дат можливе і в результаті їх бракування. Це може сильно вплинути на зміну середніх збільшуючи їх, що, в свою чергу, призводить до виникнення помилок. Проте, їм можна запобігти, відновлюючи втрачені дати за формулою

$$X_{\text{відн}} = \frac{lV + nP - \sum X}{(l-1)(n-1)},$$

де $X_{\text{відн}}$ – дата, що відновлюється; l – кількість варіантів; V – сума дат у тому варіанті, де є втрачена дата; n – кількість повторностей; P – сума дат у повторенні, де є втрачена дата; $\sum X$ – сума дат у досліді, за винятком втраченої дати ($X_{\text{відн}}$).

Так, наприклад, у досліді, де вивчався вплив різних попередників на урожайність пшениці озимої, цю урожайність у різних повтореннях показано у табл.(4.2).

Підставивши у наведену формулу замість букв їх числові

значення, отримуємо

$$X_{\text{відн}} = \frac{4 \cdot 80 + 4 \cdot 154 - 670}{(4-1)(4-1)} = \frac{2660}{9} = 29,6 \text{ ц/га.}$$

Таблиця 4.2

Урожайність озимої пшениці після попередників, т/га

Номер варіанта	Попередники	Повторність			
		I	II	III	IV
1	Багаторічні трави на один укіс	4,92	5,13	5,32	5,34
2	Вико-овсяна сумішка	4,62	4,81	4,91	5,14
3	Горох	4,48	4,70	4,71	4,92
4	Кукурудза на силос	2,50	2,74	2,76	$X_{\text{відн}}$

Відновлену дату 2,96 т/га ставлять на місце втраченої і проводять далі відповідну статистичну обробку.

При втраті одночасно кількох дат в одному досліді можна використовувати метод статистичної обробки для дослідів з неповним числом дат.

ПРАКТИЧНА РОБОТА №5

Тема: Оцінити істотність різниці вибірових середніх за t-критерієм та за найменшою істотною різницею (НІР)

Завдання:

Вихідні дані для виконання завдання наведено в таблиці 5.1.

Таблиця 5.1

Значення вихідних дат для завдання

Передостання цифра шифру (X_1)									
1	2	3	4	5	6	7	8	9	0
26	22	23	26	24	25	37	31	34	33
20	26	25	27	26	27	38	35	37	39
26	29	27	28	28	29	41	39	40	42
21	23	29	29	30	31	43	43	43	45
32	27	31	31	32	33	46	47	46	48
Остання цифра шифру (X_2)									
1	2	3	4	5	6	7	8	9	0
25	20	18	23	28	45	54	64	57	70
32	36	41	43	38	22	29	36	42	53
35	44	47	51	45	14	18	25	31	36
33	35	41	43	39	21	25	34	38	47
27	26	28	30	33	40	51	67	54	73

Для розрахунків завдання необхідно виписати із таблиці 5.1 дві вибірки вихідних даних відповідно до шифру залікової книжки.

Приклад.

Вибірка 1: 36, 39, 42, 45, 48.

Вибірка 2: 27, 36, 42, 38, 31.

Позначають спостереження вибірки 1 через X_1 , а вибірки 2 – через X_2 , складають допоміжну таблицю 5.2, в яку заносять значення вибірок 1 та 2.

В наведеній таблиці підраховуємо окремо суми по обох вибірках, тобто $\sum X_1$ та $\sum X_2$. Визначаємо середні арифметичні значення по кожній вибірці \bar{x}_1 та \bar{x}_2 , відхилення кожного значення вибірки від її середньоарифметичного, тобто $X_1 - \bar{x}_1$; $X_2 - \bar{x}_2$ та квадрати цих відхилень - $(X_1 - \bar{x}_1)^2$; $(X_2 - \bar{x}_2)^2$.

Таблиця 5.2

Обчислення квадратів відхилень від середньої арифметичної

№ п/п	Вибірка 1 (X_1)	Вибірка 2 (X_2)	Відхилення від середніх		Квадрати відхилень	
			$X_1 - \bar{x}_1$	$X_2 - \bar{x}_2$	$(X_1 - \bar{x}_1)^2$	$(X_2 - \bar{x}_2)^2$
1	36	27	-6	-7,8	36	60,84
2	39	36	-3	1,2	9	1,44
3	42	42	0	7,2	0	51,84
4	45	38	3	3,2	9	10,24
5	48	31	6	-3,8	36	14,44
Суми	210	174	0	0	90	138,8
\bar{x}	42	34,8				

Середня арифметична (\bar{x}). Для обчислення цієї характеристики варіюючі ознаки (результати спостережень) позначають знаком X , а кількість повторностей – n .

Для вибірки 1 середня арифметична становитиме:

$$\bar{X}_1 = \frac{\sum X}{n} = \frac{36 + 39 + 42 + 45 + 48}{5} = \frac{210}{5} = 42,$$

для вибірки 2 $\bar{X}_2 = \frac{\sum X}{n} = \frac{27 + 36 + 42 + 38 + 31}{5} = \frac{174}{5} = 34,8.$

Різниця середніх (d_x)

$$d_x = \bar{x}_1 - \bar{x}_2 = 42 - 34,8 = 7,2.$$

Дисперсія (S^2) - це середній квадрат відхилень кожного члена варіаційного ряду (X_1, X_2, \dots, X_n) від середньої арифметичної; це показник, який повніше за розмах варіації характеризує варіаційні ряди. Дисперсія обчислюється за формулою

$$S^2 = \frac{\sum (X - \bar{x})^2}{n - 1}.$$

Підставивши суми квадратів відхилень з таблиці 4 у наведену формулу отримаємо такі дисперсії:

$$S_1^2 = \frac{90}{5 - 1} = 22,5; \quad S_2^2 = \frac{138,8}{5 - 1} = 34,7.$$

Дисперсія використовується не лише для характеристики варіювання досліджуваних показників, а й для обчислення стандартного відхилення (S).

Стандартне відхилення (S) обчислюється за формулою

$$S = \sqrt{S^2}.$$

Для першої вибірки $S_1 = \sqrt{S_1^2} = \sqrt{22,5} = 4,74$, а для другої $S_2 = \sqrt{S_2^2} = \sqrt{34,7} = 5,89$.

Середні арифметичні мають свої помилки, які спричиняються внаслідок неповного представництва вибіркової сукупності. Ці помилки властиві лише вибірковому методу досліджень, а їх чисельне значення залежить від ступеня мінливості досліджуваних ознак і обсягів вибірки.

Похибку вибіркової середньої (S_x) обчислюють за формулою

$$S_x = \sqrt{\frac{\sum (X - \bar{x})^2}{n(n-1)}} = \sqrt{\frac{S^2}{n}}.$$

Для першої вибірки $S_{x1} = \sqrt{\frac{22,5}{5}} = 2,12$,

для другої $S_{x2} = \sqrt{\frac{34,7}{5}} = 2,63$.

Помилка різниці середніх (S_d) обчислюється за формулою

$$S_d = \sqrt{S_{x1}^2 + S_{x2}^2};$$

$$S_d = \sqrt{2,12^2 + 2,63^2} = 3,38.$$

Довірчий інтервал розраховують за формулою

$$\bar{X} \pm t_{05} S_x.$$

Для вибірки 1 довірчий інтервал

$$x_1 \pm t_{05} S_{x1};$$

$$42 \pm 2,78 \cdot 2,12;$$

$$42 \pm 5,9 (36,1 \div 47,9).$$

Для вибірки 2 довірчий інтервал

$$x_2 \pm t_{05} S_{x2};$$

$$34,8 \pm 2,78 \cdot 2,63;$$

$$34,8 \pm 7,3 (27,5 \div 42,1).$$

Фактичний критерій Стюдента розраховують за формулою

$$t_{\phi} = \frac{d_x}{S_d} = \frac{7,2}{3,38} = 2,13.$$

Число ступенів свободи $\gamma = n_1 + n_2 - 2 = 5 + 5 - 2 = 8$.

Правило: якщо значення фактичного критерія Стьюдента t_{ϕ} , розрахованого відповідно до таблиці 3, більше критерія t_{05} або t_{01} , то різниця між вибірковими середнім істотна, а якщо t_{ϕ} менше t_{05} або t_{01} , то різниця між вибірковими середнім неістотна.

Порівнюючи фактичне значення критерію Стьюдента $t_{\phi} = 2,13$ з теоретичним при різних рівнях значущості і при ступені свободи $\gamma = 8$ $t_{05} = 2,31$, $t_{01} = 3,36$, приходимо до **висновку**, що t_{ϕ} менше від теоретичного на 5% рівні значущості і менше на 1% рівні і таким чином різниця вибірових середніх за t-критерієм на обох рівнях неістотна.

Найменшу істотну різницю (НІР) розраховуємо за формулами:

$$НІР_{05} = t_{05} S_d = 2,31 \cdot 3,38 = 7,81;$$

$$НІР_{01} = t_{01} S_d = 3,36 \cdot 3,38 = 11,37.$$

Правило: якщо різниця між середніми (d_x) більша або дорівнює НІР, то вона істотна, якщо менше НІР – неістотна.

Порівнюємо різницю середніх із значенням НІР ($7,2 < 7,81$) на 5% рівні значущості та ($7,2 < 11,37$) на 1% рівні значущості й приходимо до **висновку**, що різниця вибірових середніх по НІР на обох рівнях значущості неістотна.

ПРАКТИЧНА РОБОТА № 6

Тема: Дисперсійний аналіз однофакторного вегетаційного дослідю

Суть дисперсійного аналізу: У польовому досліді, розміщеному методом рендомізованих повторень, урожай змінюється залежно від варіантів, повторень, а також від випадкових причин – неврахованої зміни умов навколишнього середовища або індивідуальної мінливості самих рослин. Останні дві причини також впливають на помилки дослідю. Англійський математик Р. Фішер виразив ці зміни сумами квадратів таких розсіювань: варіантів – C_V ; повторень – C_P ; помилки – C_Z . Їх сума і є сумою квадратів загального розсіювання (C_Y). Тоді $C_Y = C_V + C_P + C_Z$.

Для кожного розсіювання обчислюють число ступенів свободи (γ):

варіантів - $\gamma_v = l - 1$;

повторень - $\gamma = n - 1$;

помилки - $\gamma_z = (l - 1)(n - 1)$;

загального розсіювання – $N - 1$;

де $N = l \cdot n$, (l – кількість варіантів, n – кількість повторень).

Діленням певної суми квадратів відхилень на число ступенів свободи отримують дисперсію – S^2 . Дисперсія – це розсіювання даних дослідю і розчленування загального варіювання врожаю чи інших показників на складові частини. Звідси і назва методу – дисперсійний аналіз. Найбільш застосовувані дисперсія варіантів (S_V^2) та дисперсія помилки (S_Z^2), яку ще називають дисперсією залишку.

Співвідношення цих двох дисперсій є тим основним критерієм, що має змогу дати загальну оцінку достовірності різниці між середніми арифметичними або загальну оцінку достовірності дослідю. Цей критерій позначають першою літерою прізвища автора дисперсійного аналізу Фішера. Критерій Фішера визначають за формулою

$$F_{\text{факт}} = S_V^2 \div S_Z^2.$$

Розрахувавши критерій Фішера фактичний ($F_{\text{факт}}$), його порівнюють із теоретичним критерієм ($F_{\text{теор}}$) на певних рівнях надійної імовірності (додатки 3 і 4) і роблять висновки. Якщо критерій Фішера фактичний (розрахований) дорівнює критерію

теоретичному або більший від нього ($F_{факт} \geq F_{теор}$), достовірність різниць між середніми арифметичними доведена. Це означає, що в досліді є одна пара або кілька пар варіантів, між середніми арифметичними яких є достовірна різниця. Якщо $F_{факт} < F_{теор}$, то достовірних різниць між середніми арифметичними немає.

Буває, що $F_{факт}$ лише дещо менший від $F_{теор}$. Дотримуючись вищенаведеного правила, слід робити висновок про те, що достовірних різниць у досліді немає. Однак продовження аналізу часто дає змогу знайти цю різницю хоч між однією парою варіантів. Тому в таких випадках не варто зупинятися лише на розрахунках критерію F , а треба знаходити найменшу істотну різницю (НІР). З цим статистичним показником порівнюють різницю (d) між середніми арифметичними. Якщо $d \geq$ НІР, то між варіантами доведена істотність різниці. Докази ведуть, як правило, на рівнях надійної імовірності $P_{0,95}$ та $P_{0,99}$.

Дисперсійний аналіз є найдосконалішим методом статистичної обробки даних. Його переваги полягають у виділенні із загального варіювання його компонентів, розрахуванні узагальненої помилки всього досліді (S_x) на основі більшої кількості спостережень, ніж для індивідуальних помилок окремих пар варіантів у недисперсійних методах. Так, при 5-ти варіантах і 4-х повторностях число ступенів свободи помилки ν_z становить $(5-1)(4-1)=12$, у той час як для кожного варіанта досліді окремо воно становить $4-1=3$, тобто в 4 рази менше, а для пари варіантів $(4-1)+(4-1)=6$.

Дисперсійний аналіз досить ефективний для багатofакторних дослідів, оскільки дає змогу визначити достовірність не лише дії факторів окремо, а й їх взаємодії.

Висновок про точність усього досліді роблять наприкінці дисперсійного аналізу на основі числового значення відносної помилки $S_x \%$, яку визначають за формулою

$$S_x \% = \frac{S_x \cdot 100}{\overline{X_N}},$$

де $S_x \%$ - узагальнена помилка досліді, $\overline{X_N}$ - середня арифметична всього досліді.

Без обчислення помилки досліді дисперсійний аналіз вважається незакінченим, а висновки неповними.

Основна відміна дисперсійних аналізів полягає у формулах і в

переліках тих сум квадратів, що розраховуються:

- 1) неповна рендомізація - $C_y = C_V + C_P + C_Z$;
- 2) повна рендомізація - $C_y = C_V + C_Z$;
- 3) латинський квадрат і латинський прямокутник

$$C_y = C_V + C_P + C_Z + C_C;$$

- 4) двофакторний дослід

$$C_y = C_P + C_a + C_b + C_{ab} + C_{ac} + C_{bc} + C_{abc} + C_Z.$$

Завдання: Виконати дисперсійний аналіз даних однофакторного вегетаційного дослідження по вивченню дії добрив на урожай зеленої маси гороху (таблиця 6.1).

Таблиця 6.1

Урожай зеленої маси гороху, г/посудину

Номер варіанта	Повторення		
	I	II	III
1	24,3	23,4	25,5
2	23,7	24,3	25,5
3	24,0	24,3	25,5
4	24,0	24,9	25,5
5	24,0	25,2	25,5

Приклад: Виконати дисперсійний аналіз даних однофакторного вегетаційного дослідження по вивченню дії добрив на урожай зеленої маси гороху (таблиця 6.2).

Дози добрив по варіантах: 1 варіант – 0, 2 – 10, 3 – 20, 4 – 30, 5 – 40 г на посудину.

Кожний варіант вивчається в трьох посудинах.

Таблиця 6.2

Урожай зеленої маси гороху, г/посудину

Номер варіанта	Повторення			Сума за варіантом, $\sum V$	Середня за варіантом, \bar{X}
	I	II	III		
1	16,5	17,3	17,5	51,3	17,1
2	18,7	20,4	18,5	57,6	19,2
3	20,7	21,2	19,0	60,9	20,3
4	21,3	23,8	22,1	67,2	22,4
5	23,1	24,3	25,5	72,9	24,3
				$\sum X = 309,9$	$\bar{X}_{N=20} = 20,7$

В однофакторному вегетаційному досліді загальне варіювання результативної ознаки розкладається на два компоненти:

$C_y = C_v + C_z$, тобто на варіювання варіантів і випадкове варіювання.

Статистичний аналіз даних проводять в три етапи:

1 етап. Складають розрахункову таблицю 6.2, в якій визначають суми і середні по варіантах, загальну суму ($\sum X$) і середнє арифметичне по всьому досліді ($\bar{X}_N = \sum X:N$).

2 етап. Розраховують суми квадратів відхилень по формулах та визначають фактичне значення критерія Фішера $F_{факт}$:

- загальне число спостережень у досліді:

$$N = l \cdot n = 5 \cdot 3 = 15;$$

- корегуючий фактор (поправка):

$$C = (\sum X)^2 : N = (309,9)^2 : 15 = 6402,5;$$

- загальна сума квадратів відхилень:

$$C_y = \sum X^2 - C = (16,5^2 + 17,3^2 + 17,5^2 + \dots + 25,5^2) - 6402,5 = 6507,71 - 6402,5 = 105,21;$$

- сума квадратів відхилень для варіантів:

$$C_v = \sum V^2 : n - C = (51,3^2 + 57,6^2 + 60,9^2 + 67,2^2 + 72,7^2) : 3 - 6402,5 = 19488,51 : 3 - 6402,5 = 93,67;$$

- сума квадратів відхилень залишку (помилки):

$$C_z = C_y - C_v = 105,21 - 93,67 = 11,54.$$

Одержані дані розрахунків заносять до таблиці 6.3, на основі яких обчислюють дисперсію варіантів (S_V^2), дисперсію помилки (S_Z^2) та критерій Фішера фактичний ($F_{факт}$).

Таблиця 6.3

Результати дисперсійного аналізу

Розсіювання	Суми квадратів	Ступені вільності γ	Дисперсія, S^2	$F_{факт}$	$F_{теор}$	
					$P_{0,95}$	$P_{0,99}$
Загальне	105,21	14	-	-	-	-
Варіантів	93,67	4	23,42	20,3	3,48	5,99
Помилки	11,54	10	1,15			

Дисперсії розраховують за такими формулами:

$$S_V^2 = C_v : \gamma_v = 93,67 : 4 = 23,42; \quad S_Z^2 = C_z : \gamma_z = 11,54 : 10 = 1,15.$$

Критерій Фішера фактичний розраховують за формулою:

$$F_{факт} = S_V^2 : S_Z^2 = 23,42 : 1,15 = 20,3.$$

Теоретичне значення критерію Фішера знаходять у додатках 3 і 4 за числом ступеня вільності варіантів $\gamma=3$ (більша дисперсія) та помилки $\gamma_Z=9$ (менша дисперсія). На перетині цих чисел теоретичне значення критерію Фішера становить при $P_{0,95}$ 3,48 і при $P_{0,99}$ – 5,99.

Правило: якщо критерій Фішера фактичний дорівнює теоретичному або більший від нього, то різниця між усіма чи окремими варіантами дослідження вважається достовірною. У нашому прикладі $F_{\text{факт}}=1694,93$, що значно більше за $F_{0,95}$ і $F_{0,99}$, що становлять відповідно 3,48 і 5,99, свідчить про достовірність цих різниць на обох рівнях надійної імовірності.

3 етап.

Для оцінки конкретних різниць розраховують:

а) помилку дослідження ($S_{\bar{X}}$)

$$S_{\bar{X}} = \sqrt{S_Z^2 \div n} = \sqrt{0,29 \div 3} = 0,27 .$$

б) помилку різниці S_d

$$S_d = S_{\bar{X}} \cdot 1,41 = 0,27 \cdot 1,41 = 0,381 \text{ ц/га } (1,41=\sqrt{2}).$$

в) найменшу істотну різницю (НІР) в абсолютних і відносних показниках:

$$НІР = S_d \cdot t_{05} = 2,23 \cdot 0,88 = 1,96;$$

$$НІР\% = \frac{t_{05} \cdot S_d}{\bar{X}} \cdot 100 = \frac{2,23 \cdot 0,88}{20,66} \cdot 100 = 9,5\%.$$

Теоретичне значення критерію Стюдента знаходять у додатку 1 за числом ступенів вільності залишкового розсіювання (помилки), яке у нашому випадку становить 2,23.

Одержані дані заносять до підсумкової таблиці 6.4.

Таблиця 6.4

Підсумкова таблиця дисперсійного аналізу

Номер варіанта	Дози добрив, г/посудину	Середній урожай, \bar{X}	Різниця d	
			г/посудину	%
1	-	17,1	-	-
2	10	19,2	2,1	12,28
3	20	20,3	3,2	18,71
4	30	22,4	5,3	30,99
5	40	24,3	7,2	42,1
НІР ₀₅		-	1,96	9,5

Порівнюючи різниці між дослідними варіантами і контролем та різниці між окремими дослідними варіантами із зазначенням НР, роблять висновок про істотність цих різниць, дотримуючись **ПРАВИЛА: якщо різниці більші за значення НР або дорівнюють йому, то ці різниці істотні.**

Висновок: внесення добрив в дозах 10, 20, 30, 40 г/посудину істотно підвищує врожайність зеленої маси гороху (у всіх варіантах різниця з контролем – d більше НР).

ПРАКТИЧНА РОБОТА № 7

Тема: *Дисперсійний аналіз двофакторного вегетаційного дослідю*

Завдання: Виконати дисперсійний аналіз даних однофакторного вегетаційного дослідю по вивченню дії добрив на урожай зеленої маси гороху (таблиця 7.1).

Таблиця 7.1

Урожай зеленої маси гороху, г/посудину

Номер варіанта	Повторення		
	I	II	III
1	24,3	23,4	25,5
2	23,7	24,3	25,5
3	24,0	24,3	25,5
4	24,0	24,9	25,5
5	24,0	25,2	25,5

Приклад: Виконати дисперсійний аналіз даних однофакторного вегетаційного дослідю по вивченню дії добрив на урожай зеленої маси гороху (таблиця 7.2).

Таблиця 7.2

Урожай зеленої маси гороху, г/посудину

Номер варіанта	Повторення			Сума за варіантом, $\sum V$	Середня за варіантом, \bar{X}
	I	II	III		
1	16,5	17,3	17,5	51,3	17,1
2	18,7	20,4	18,5	57,6	19,2
3	20,7	21,2	19,0	60,9	20,3
4	21,3	23,8	22,1	67,2	22,4
5	23,1	24,3	25,5	72,9	24,3
				$\sum X = 309,9$	$\bar{X}_{N=5} = 20,7$

Дози добрив по варіантах: 1 варіант – 0, 2 – 10, 3 – 20, 4 – 30, 5 – 40 г на посудину. Кожний варіант вивчається в трьох посудинах.

В однофакторному вегетаційному досліді загальне варіювання результативної ознаки розкладається на два компоненти:

$S_y = S_v + S_z$, тобто на варіювання варіантів і випадкове варіювання.

Статистичний аналіз даних проводять в три етапи:

1 етап. Складають розрахункову таблицю 7.2, в якій визначають суми і середні по варіантах, загальну суму ($\sum X$) і середнє арифметичне по всьому досліді ($\bar{X}_N = \sum X:N$).

2 етап. Розраховують суми квадратів відхилень по формулах та визначають фактичне значення критерія Фішера $F_{факт}$:

- загальне число спостережень у досліді: $N = l \cdot n = 5 \cdot 3 = 15$;

- корегуючий фактор (поправка):

$$C = (\sum X)^2:N = (309,9)^2:15 = 6402,5;$$

- загальна сума квадратів відхилень:

$$C_y = \sum X^2 - C = (16,5^2 + 17,3^2 + 17,5^2 + \dots + 25,5^2) - 6402,5 = 6507,71 - 6402,5 = 105,21;$$

- сума квадратів відхилень для варіантів:

$$C_v = \sum V^2:N = (51,3^2 + 57,6^2 + 60,9^2 + 67,2^2 + 72,7^2):3 - 6402,5 = 19488,51:3 - 6402,5 = 93,67;$$

- сума квадратів відхилень залишку (помилки):

$$C_z = C_y - C_v = 105,21 - 93,67 = 11,54.$$

Одержані дані розрахунків заносять до таблиці 7.3, на основі яких обчислюють дисперсію варіантів (S_V^2), дисперсію помилки (S_Z^2) та критерій Фішера фактичний ($F_{факт}$).

Таблиця 7.3

Результати дисперсійного аналізу

Розсіювання	Суми квадратів	Ступені вільності γ	Дисперсія, S^2	$F_{факт}$	$F_{теор}$	
					$P_{0,95}$	$P_{0,99}$
Загальне	105,21	14	-	-	-	-
Варіантів	93,67	4	23,42	20,3	3,48	5,99
Помилки	11,54	10	1,15			

Дисперсії розраховують за такими формулами:

$$S_V^2 = C_v : \gamma_v = 93,67:4 = 23,42; \quad S_Z^2 = C_z : \gamma_z = 11,54:10 = 1,15.$$

Критерій Фішера фактичний розраховують за формулою:

$$F_{факт} = S_V^2 : S_Z^2 = 23,42:1,15 = 20,3.$$

Теоретичне значення критерію Фішера знаходять у додатках 3 і 4 за числом ступеня вільності варіантів $\gamma_v=3$ (більша дисперсія) та помилки $\gamma_z=9$ (менша дисперсія). На перетині цих чисел теоретичне значення критерію Фішера становить при $P_{0,95}$ 3,48 і при $P_{0,99}$ – 5,99.

Правило: якщо критерій Фішера фактичний дорівнює теоретичному або більший від нього, то різниця між усіма чи

окремими варіантами досліду вважається достовірною. У нашому прикладі $F_{\text{факт}}=1694,93$, що значно більше за $F_{0,95}$ і $F_{0,99}$, що становлять відповідно 3,86 і 6,99, свідчить про достовірність цих різниць на обох рівнях надійної імовірності.

3 етап. Для оцінки конкретних різниць розраховують:

а) помилку досліду ($S_{\bar{x}}$)

$$S_{\bar{x}} = \sqrt{S_Z^2 \div n} = \sqrt{0,29 \div 4} = 0,27.$$

б) помилку різниці S_d

$$S_d = S_{\bar{x}} \cdot 1,41 = 0,27 \cdot 1,41 = 0,381 \text{ ц/га } (1,41=\sqrt{2}).$$

в) найменшу істотну різницю (НІР) в абсолютних і відносних показника:

$$НІР = S_d \cdot t_{05} = 2,23 \cdot 0,88 = 1,96;$$

$$НІР\% = \frac{t_{05} \cdot S_d}{\bar{X}} \cdot 100 = \frac{2,23 \cdot 0,88}{20,66} \cdot 100 = 9,5\%.$$

Теоретичне значення критерію Стюдента знаходять у додатку 1 за числом ступенів вільності залишкового розсіювання (помилки), яке у нашому випадку становить 2,23. Одержані дані заносять до підсумкової таблиці 7.4.

Таблиця 7.4

Підсумкова таблиця дисперсійного аналізу

Номер варіанта	Дози добрив, г/посудину	Середній урожай, \bar{X}	Різниця d	
			г/посудину	%
1	-	17,1	-	-
2	10	19,2	2,1	12,28
3	20	20,3	3,2	18,71
4	30	22,4	5,3	30,99
5	40	24,3	7,2	42,1
НІР ₀₅		-	1,96	9,5

Порівнюючи різниці між дослідними варіантами і контролем та різниці між окремими дослідними варіантами із зазначенням НІР, роблять висновок про істотність цих різниць, дотримуючись **ПРАВИЛА: якщо різниці більші за значення НІР або дорівнюють йому, то ці різниці істотні.**

Висновок: внесення добрив в дозах 10, 20, 30, 40 г/посудину істотно підвищує врожайність зеленої маси гороху (у всіх варіантах різниця з контролем – d більше НІР).

ПРАКТИЧНА РОБОТА №8

Тема: Дисперсійний аналіз однофакторного польового досліду

Завдання: Виконати дисперсійний аналіз одно факторного польового досліду. Виконується на базі даних про врожайність пшениці озимої після різних попередників, які наведені в таблиці 8.1.

Таблиця 8.1

Урожайність озимої пшениці після різних попередників, ц/га

Номер варіанта	Попередник	Повторення			
		I	II	III	IV
1	Багаторічні трави на один укіс	49,2	51,3	53,2	53,4
2	Вико-вівсяна сумішка	46,2	48,1	49,1	51,4
3	Горох	44,8	47,0	47,1	49,2
4	Кукурудза на силос	25,0	27,4	27,6	29,6

Приклад: Дисперсійний аналіз виконується на базі даних про врожайність озимої пшениці після різних попередників в 3 етапи.

1 етап.

Таблиця 8.2

Урожайність озимої пшениці після різних попередників, ц/га

Номер варіанта	Попередник	Повторення				Сума за варіантом, $\sum V$	Середня за варіантом, \bar{X}
		I	II	III	IV		
1	Багаторічні трави на один укіс	49,2	51,3	53,2	53,4	207,1	51,8
2	Вико-вівсяна сумішка	46,2	48,1	49,1	51,4	194,8	48,7
3	Горох	44,8	47,0	47,1	49,2	188,1	47,0
4	Кукурудза на силос	25,0	27,4	27,6	29,6	109,6	27,4
Сума за повторенням, $\sum P$	-	165,2	173,8	177,0	183,6	$\sum X = \sum \sum P =$ $= \sum \sum V = 699,6$	$\bar{X}_{N=}$ $= 43,7$

2 етап.

Вибирають довільний початок, значення якого має бути цілим числом, близьким до загальної середньої ($A=44$) і складають таблицю 8.3 відхилень подільночних урожаїв від довільного початку ($X-A$) та визначають суми відхилень за варіантами ($\sum V_A$), повтореннями ($(\sum P_A)$) і суму сум за варіантами і повтореннями (L).

Таблиця 8.3**Відхилення дат від довільного початку ($X-A$)**

Номер варіанта	Повторення				$\sum V_A$
	I	II	III	IV	
1	5,2	7,3	9,2	9,4	31,1
2	2,2	4,1	5,1	7,4	18,8
3	0,8	3,0	3,1	6,2	13,1
4	-19,0	-16,6	-16,4	-14,4	-66,4
$\sum P_A$	-10,8	-2,2	1,0	8,6	$L=-3,4$

Всі одержані відхилення та їх суми підносять до квадрата і записують до таблиці 8.4.

Таблиця 8.4**Квадрат відхилень дат від довільного початку**

Номер варіанта	Повторення				$(\sum V_A)^2$
	I	II	III	IV	
1	27,04	53,29	84,64	88,36	967,21
2	4,84	16,81	26,01	54,76	353,44
3	0,64	9,0	9,61	38,44	171,61
4	361	275,56	268,96	207,36	4408,96
$(\sum P_A)^2$	116,64	4,84	1,0	73,96	$L^2=11,56$

Корегуючий фактор $C = \frac{L^2}{N} = \frac{11,56}{16} = 0,72$.

Загальна сума квадратів відхилень

$$C_Y = \sum (X - A)^2 - C = (27,04 + 53,29 + 84,64 + 88,36 + 4,84 + 16,81 + 26,01 + 54,76 + 0,64 + 9,0 + 9,61 + 38,44 + 361 + 275,56 + 268,96 + 207,36) - 0,72 = 1526,32 - 0,72 = 1525,6$$

Сума квадратів повторень

$$C_P = (\sum P_A)^2 \div l - C = (116,64 + 4,84 + 1 + 73,96) \div 4 - 0,72 = 196,44 \div 4 - 0,72 = 48,39$$

Сума квадратів відхилень варіантів

$$C_V = (\sum V_A)^2 \div n - C = (967,21 + 353,44 + 171,61 + 4408,96) \div 4 - 0,72 = 5901,22 \div 4 - 0,72 = 1475,31 - 0,72 = 1474,59$$

Сума квадратів відхилень залишку

$$C_Z = C_y - C_P - C_V = 1525,59 - 48,39 - 1474,59 = 2,61.$$

Обчислюють число ступенів вільності загального розсіювання (γ_Y), повторень (γ_P), варіантів (γ_V) та помилки (γ_Z):

$$\gamma_Y = N - 1 = 16 - 1 = 15;$$

$$\gamma_P = n - 1 = 4 - 1 = 3;$$

$$\gamma_V = l - 1 = 4 - 1 = 3;$$

$$\gamma_Z = (l - 1)(n - 1) = (4 - 1)(4 - 1) = 3 \cdot 3 = 9;$$

Одержані дані розрахунків заносять до лівої частини таблиці 8.5, на основі яких обчислюють дисперсію варіантів (S_V^2), дисперсію помилки (S_Z^2) та критерій Фішера фактичний ($F_{\text{факт}}$)

Таблиця 8.5

Результати дисперсійного аналізу

Розсіювання	Суми квадратів	Ступені вільності γ	Дисперсія, S^2	$F_{\text{факт}}$	$F_{\text{теор}}$	
					$P_{0,95}$	$P_{0,99}$
Загальне	1525,59	15	-	-	-	-
Повторень	48,39	3	-	-	-	-
Варіантів	1474,59	3	491,53	1694,93	3,86	6,99
Помилки	2,61	9	0,29	-	-	-

Дисперсії розраховують за такими формулами:

$$S_V^2 = C_V : \gamma_V = 1474,59 : 3 = 491,5 ;$$

$$S_Z^2 = C_Z : \gamma_Z = 2,61 : 9 = 0,29.$$

Критерій Фішера фактичний розраховують за формулою

$$F_{\text{факт}} = S_V^2 : S_Z^2 = 491,53 : 0,29 = 1694,93.$$

Теоретичне значення критерію Фішера знаходять у додатках 3 і 4 за числом ступеня вільності варіантів $\gamma=3$ (більша дисперсія) та помилки $\gamma_Z=9$ (менша дисперсія). На перетині цих чисел теоретичне значення критерію Фішера становить при $P_{0,95}$ 3,86 і при $P_{0,99}$ – 6,99.

Правило: якщо критерій Фішера фактичний дорівнює теоретичному або більший від нього, то різниця між усіма чи окремими варіантами дослідів вважається достовірною.

Висновок: у нашому прикладі $F_{\text{факт}}=1694,93$, що значно більше за $F_{0,95}$ і $F_{0,99}$, що становлять відповідно 3,86 і 6,99, свідчить про достовірність цих різниць на обох рівнях надійної імовірності.

3 етап. Далі розраховують узагальнену помилку дослідів ($S_{\bar{X}}$) та помилку різниці S_d за формулами:

$$S_{\bar{x}} = \sqrt{S_z^2 \div n} = \sqrt{0,29 \div 4} = 0,27;$$

$$S_d = S_{\bar{x}} \cdot 1,41 = 0,27 \cdot 1,41 = 0,381 \text{ ц/га } (1,41 = \sqrt{2}).$$

Найменшу істотну різницю (НІР) розраховують, як правило, на двох рівнях надійної імовірності ($P_{0,95}$ і $P_{0,99}$) за такими формулами:

$$\text{НІР}_{0,95} = S_d \cdot t_{0,95};$$

$$\text{НІР}_{0,99} = S_d \cdot t_{0,99}.$$

Теоретичне значення критерію Стюдента знаходять у додатку 1 за числом ступенів вільності залишкового розсіювання (помилки), яке у нашому випадку становить на рівнях імовірності $P_{0,95}$ і $P_{0,99}$, відповідно 2,26 і 3,25.

Для характеристики істотності приватних різниць і точності дослідів розраховують:

- **найменшу істотну різницю:**

$$\text{НІР}_{0,95} = S_d \cdot t_{0,95} = 0,381 \cdot 2,26 = 0,86 \text{ ц/га};$$

$$\text{НІР}_{0,99} = S_d \cdot t_{0,99} = 0,381 \cdot 3,25 = 1,24 \text{ ц/га};$$

- **відносну помилку всього дослідів**

$$S_{\bar{x}} \% = \frac{S_{\bar{x}} \cdot 100}{X_N} = \frac{0,27 \cdot 100}{43,7} = 0,62\%;$$

- **точність дослідів**

$$T \% = 100 - 0,62 = 99,32\%.$$

Одержані дані заносять до підсумкової таблиці 8.6.

Таблиця 8.6

Підсумкова таблиця дисперсійного аналізу

Номер варіанта	Попередник	\bar{X}	Різниця d	НІР		$S_{\bar{x}} \%$	T%
				0,95	0,99		
1	Багаторічні трави на один укіс	51,8	-	-	-	-	-
2	Вико-вівсяна сумішка	48,7	-3,1	0,86	1,24	0,62	99,36
3	Горох	47,0	-4,8				
4	Кукурудза на силос	27,4	-24,4				

Порівнюючи різниці між дослідними варіантами і контролем та різниці між окремими дослідними варіантами із зазначенням НІР на обох рівнях надійної імовірності, роблять висновок про істотність цих різниць, дотримуючись **ПРАВИЛА**: *якщо різниці більші за значення НІР або дорівнюють йому, то ці різниці істотні.*

Висновки:

- оскільки критерій Фішера фактичний становить 1694,93, що значно більше за $F_{0,95}(3,86)$ і $F_{0,99}(6,99)$, то різниця між усіма окремими варіантами дослідів достовірна на обох рівнях надійної імовірності;
- точність обчислення середніх арифметичних висока, тому що значення $S_{\bar{x}}\%$ не перевищує 3 %;
- зниження врожайності в другому варіанті проти контролю істотне, тому що різниця між середніми цих варіантів більша від НІР на обох рівнях імовірності, у третьому і четвертому варіантах зниження врожайності тим більш істотне на обох рівнях імовірності.

ЛАБОРАТОРНА РОБОТА №9

Тема: *Кореляційний та регресійний аналіз прямолінійної залежності*

Завдання: Провести аналіз залежності між довжиною 10 окремих листків пшениці озимої та їх площами (табл. 9.1) визначених на основі індивідуальних вимірів.

Таблиця 9.1

Данні для проведення розрахунків

Номери листків (пар)	Довжина листа, см X	Площа листа, см ² Y
1	16,1	7,4
2	17,3	8,7
3	18,6	10,3
4	20,0	11,2
5	21,3	12,9
6	21,6	13,2
7	21,8	13,7
8	22,0	14,1
9	22,4	14,3
10	22,8	14,8

Приклад: Виконати кореляційний та регресійний аналізи даних таблиці 9.2, в якій наведено результати визначення відносної вологості (X) та липкості (Y) чорнозему.

Таблиця 9.2

Обчислення кореляційної залежності між відотною вологістю та липкістю ґрунту

Номер пари	Вологість, % X	Липкість, г/см ² Y	X ²	y ²	X·Y
1	19,9	0,0	396,01	0,00	0,00
2	20,9	0,6	436,81	0,36	12,54
3	26,1	1,1	681,21	1,21	28,71
4	29,4	1,2	864,36	1,44	35,28
5	30,5	1,7	930,25	2,89	51,85
6	40,3	1,7	1624,09	2,89	68,51
7	44,8	2,6	2007,04	6,76	116,48
8	47,8	3,4	2284,84	11,56	162,52
9	55,6	4,2	3091,36	17,64	233,52
10	58,3	5,8	3398,89	33,64	3387,14
11	64,5	6,3	4160,25	39,69	406,35
12	76,6	7,3	5867,56	53,29	559,18
N=12	∑X=514,7	∑Y=35,9	∑X ² =25742,67	∑Y ² =171,37	∑X·Y=2013,08

Кореляційний аналіз:

Визначають шість допоміжних величин:

$$n = 12;$$

$$\bar{X} = \sum X \div n = 514,7 \div 12 = 42,89\%;$$

$$\bar{Y} = \sum Y \div n = 35,9 \div 12 = 2,992 / \text{см}^2.$$

$$\sum (X - \bar{x})^2 = \sum X^2 - (\sum X)^2 \div n = 257742,67 - 514,7^2 \div 12 = 3666,33;$$

$$\sum (Y - \bar{y})^2 = \sum Y^2 - (\sum Y)^2 \div n = 171,37 - 35,9^2 \div 12 = 63,97;$$

$$\sum (X - \bar{x})(Y - \bar{y}) = \sum XY - (\sum X \sum Y) \div n = 2013,08 - (514,7 \cdot 35,9) \div 12 = 473,27.$$

Далі визначають:

- **коефіцієнт кореляції r**

$$r = \frac{\sum (X - \bar{x})(Y - \bar{y})}{\sqrt{\sum (X - \bar{x})^2 \sum (Y - \bar{y})^2}} = \frac{473,27}{\sqrt{3666,33 \cdot 63,97}} = 0,977 \approx 0,98.$$

- **помилку коефіцієнта кореляції**

$$S_r = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-0,98^2}{12-2}} = \sqrt{\frac{0,02}{10}} = 0,045.$$

- **критерій достовірності коефіцієнта кореляції**

$$t_r = \frac{r}{S_r} = \frac{0,98}{0,045} = 21,78.$$

Теоретичне значення критерію Стюдента знаходять за числом ступенів вільності

$$\gamma_r = n - 2 = 12 - 2 = 10,$$

$$t_{0,95}=2,23; t_{0,99}=3,17.$$

Про силу зв'язку роблять висновок за таким правилом: якщо коефіцієнт кореляції дорівнює одиниці, то зв'язок - повний; якщо він становить 0,66-0,99, то зв'язок - сильний; якщо він перебуває в межах 0,33-0,66 – середній і якщо коефіцієнт кореляції менший за 0,33, то зв'язок - слабкий.

Оскільки в нашому прикладі $r=0,98$, то зв'язок між відносною вологістю та липкістю чорнозему сильний.

Про напрям зв'язку висновок роблять за правилом залежно від знака при коефіцієнті кореляції: якщо він плюсовий, то кореляція пряма, а якщо мінусовий, то зворотна.

У нашому прикладі кореляція пряма.

Про достовірність зв'язку висновок роблять за правилом: якщо критерій достовірності коефіцієнта кореляції фактичний

більший за теоретичні його значення або дорівнює їм, то зв'язок достовірний.

Висновок: оскільки критерій фактичний (t_r) становить 21,78, що значно більше теоретичних значень $t_{0,95}$ (2,23) і $t_{0,99}$ (3,17), то зв'язок між відносною вологістю чорнозему і його липкістю достовірний на обох рівнях надійної імовірності.

Регресійний аналіз. Його здійснюють за сильного та достовірного зв'язку і будь-якого напрямку (прямого чи зворотного). Під регресією розуміють зміну результативної ознаки Y (функції) при певній зміні однієї або декількох факторіальних (аргумента). Зв'язок між функцією та аргументом виражають рівнянням регресії, що має такий вигляд

$$Y = \bar{y} + b_{yx} (X - \bar{x}),$$

де b_{yx} - коефіцієнт регресії, який визначається за формулою

$$b_{yx} = \frac{\sum (X - \bar{x})(Y - \bar{y})}{\sum (X - \bar{x})^2} = \frac{473,27}{3666,33} = 0,13 \text{ г/см}^3.$$

Отже, в нашому прикладі за зміни відносної вологості чорнозему на 1 % його липкість змінюється на 0,13 г/см³.

Підготувавши значення коефіцієнта регресії у рівняння регресії, одержимо робоче рівняння за яким, знаючи відносну вологість чорнозему, можна визначити його липкість.

$$y = \bar{y} + b_{yx} (X - \bar{x}) = 2,99 + 0,13(X - 42,89) = 0,13X - 2,58;$$

таким чином $y = 0,13X - 2,58$.

Розрахуємо липкість чорнозему за відносної вологості його 76,6% (12 пара) $y = 0,13 \cdot 76,6 - 2,58 = 9,96 - 2,58 = 7,38 \text{ г/см}^3$, а фактична липкість склала 7,30 г/см³. Точність прогнозування липкості чорнозему за його відносною вологістю розраховують за формулою.

Різниця між розрахунковою липкістю і фактичною становить $7,38 - 7,30 = 0,08 \text{ г/см}^3$ або $0,08 \cdot 100 \div 7,30 = 1,1 \%$, тому точність досліду дорівнюватиме

$$T\% = 100 - 1,1 = 98,9\%.$$

Таким чином, точність прогнозування липкості чорнозему за його відносною вологістю є високою.

ПРАКТИЧНА РОБОТА № 10

Тема: Криволінійна залежність

Завдання: На основі дослідних матеріалів залежність урожайності пшениці озимої від норми висіву (X) за таблицею 10.1 визначити кореляційне відношення та скласти рівняння регресії.

Таблиця 10.1

Данні для проведення розрахунків

№ пар	1	2	3	4	5	6	7	8	9	10
Норми висіву, млн/га X	3,5	3,6	3,7	3,8	3,9	4	4,1	4,2	4,3	4,4
Урожайність, ц/га Y	33	35	37	43	52	56	55	51	46	42

Наявність криволінійного зв'язку можна визначити за графіком, коли за зростання X спостерігається спочатку зростання Y, а потім його зменшення (або навпаки).

Приклад: На основі дослідних матеріалів залежність урожайності озимої пшениці від норми висіву (X) за таблицею 10.2 визначити кореляційне відношення та скласти рівняння регресії.

Таблиця 10.2

Урожайність зерна пшениці озимої (Y) залежно від норми висіву (X)

№ пар	Норми висіву, млн/га X	Урожайність, ц/га Y	\bar{Y}_x	$y - \bar{y}_x$	$(y - \bar{y}_x)^2$	$y - \bar{y}$	$(y - \bar{y})^2$
1	3,5	32	33	-1	1	-14	196
2	3,6	34		1	1	-12	144
3	3,7	38	42	-4	16	-8	64
4	3,8	46		4	16	0	0
5	3,9	51	53	-2	4	5	25
6	4,0	55		2	4	9	81
7	4,1	57	55	2	4	11	121
8	4,2	53		-2	4	7	49
9	4,3	50	47	3	9	4	16
10	4,4	44		-3	9	-2	4
		$\bar{y} = 46$		$\sum (y - \bar{y}_x) = 0$	$\sum (y - \bar{y}_x)^2 = 68$	$\sum (y - \bar{y}) = 0$	$\sum (y - \bar{y})^2 = 700$

Аналіз криволінійної залежності. Для визначення аналізу криволінійної залежності користуються не коефіцієнтом кореляції, а

кореляційним відношенням η_{xy} або η_{yx} , яке розраховується за формулою

$$\eta_{yx} = \sqrt{\frac{\sum (y - \bar{y})^2 - (y - y_x)^2}{\sum (y - \bar{y})^2}} = \sqrt{\frac{700 - 68}{700}} = 0,95.$$

Суть розрахунків таблиці 18. При 10 варіантах норм висіву їх доцільно розділити на 5 груп і розраховувати середнє значення y_x для кожної групи. Подальші розрахунки зрозумілі, їх суми використовують для обчислення кореляційного відношення.

Висновок: оскільки кореляційне відношення становить 0,95 і не виходить за межі 0,66-0,99, то між нормами висіву і врожайністю зерна озимої пшениці зв'язок сильний.

Далі розраховують:

- помилку кореляційного відношення

$$S\eta_{yx} = \sqrt{\frac{1 - \eta_{yx}^2}{n - 2}} = \sqrt{\frac{1 - 0,95^2}{10 - 2}} = 0,11.$$

- критерій достовірності фактичний

$$t_\eta = \frac{\eta_{yx}}{S\eta_{yx}} = \frac{0,95}{0,11} = 8,64.$$

Критерій Стюдента теоретичний знаходять за числом ступенів вільності

$$\gamma_\eta = n - 2 = 10 - 2 = 8; \quad t_{0,95} = 2,31; \quad t_{0,99} = 3,36.$$

Висновок: оскільки критерій Стюдента фактичний $t_\eta = 8,64$ більший за $t_{0,95}$ і $t_{0,99}$, то зв'язок достовірний на обох рівнях надійної імовірності.

Складання рівняння регресії для криволінійної залежності. Оскільки у наведеному прикладі зв'язок сильний і достовірний це дає нам підставу скласти рівняння регресії. Графічне зображення залежності урожайності зерна озимої пшениці від норми висіву має форму параболи.

Рівняння параболи загалом має вигляд

$$y = a + b_1 x + b_2 X^2.$$

Вона розраховується за формулою

$$y = \bar{y} + \frac{\sum (X - \bar{x})y}{\sum (X - \bar{x})^2} (X - \bar{x}) + \left[\frac{\sum (X - \bar{x})^2 y - n\bar{c}y}{\sum (X - \bar{x})^4 - n\bar{c}^2} \right] \cdot [(X - \bar{x})^2 - \bar{c}],$$

де \bar{c} – корегуючий фактор, що визначається як $\sum (X - \bar{x})^2 : n$.

Для розрахунків рівняння параболи складають допоміжну таблицю 10.3.

Таблиця 10.3

Розрахунки вихідних даних для складання квадратичної параболі

Норма висіву, млн./га, X	Урожай ц/га, Y	$(X - \bar{x})$	$(X - \bar{x})^2$	$(X - \bar{x})^4$	$(X - \bar{x})y$	$(X - \bar{x})^2 y$
3,5	32,	-0,45	0,2025	0,0410	-14,4	6,480
3,6	34	-0,35	0,1225	0,0150	-11,9	4,165
3,7	38	-0,25	0,0625	0,0039	-9,5	2,375
3,8	46	-0,15	0,0225	0,0005	-6,9	1,035
3,9	51	-0,05	0,0025	0	-2,55	0,128
4,0	55	0,05	0,0025	0	2,75	0,138
4,1	57	0,15	0,0225	0,0005	8,55	1,283
4,2	53	0,25	0,0625	0,0039	13,25	3,313
4,3	50	0,35	0,1225	0,0150	17,50	6,125
4,4	44	0,45	0,2025	0,0410	19,80	8,910
$\bar{x}=3,95$	$\bar{y}=46,0$	$\Sigma(X - \bar{x})=0$	$(X - \bar{x})^2 = 0,825$	$(X - \bar{x})^4 = 0,1208$	$(X - \bar{x})y = 16,60$	$(X - \bar{x})^2 y = 33,95$

Корегуючий фактор

$$C = \frac{\sum (X - \bar{x})^2}{n} = \frac{0,825}{10} = 0,0825.$$

Далі підставляють значення сум з таблиці 10.3 та корегуючий фактор у вищенаведену формулу для одержання рівняння регресії

$$\begin{aligned} \bar{y} &= 46 + \frac{16,6}{0,825}(X - 3,95) + \left[\frac{33,95 - 10 \cdot 0,0825 \cdot 46}{0,1208 - 10 \cdot 0,0825^2} \right] \cdot [(X - 3,95)^2 - 0,0825] = 46 + 20,1212X - \\ &- 79,4787 + \frac{-4}{0,0527} \cdot (X^2 - 7,9X + 15,6025 - 0,0825) = 46 + 20,1212X - 79,4787 - 75,9X^2 + \\ &+ 599,61X - 1184,2298 + 6,2618 = 619,7312X - 75,9X^2 - 1211,4467. \end{aligned}$$

Для перевірки точності прогнозування врожайності зерна пшениці озимої за розрахованим рівнянням регресії, підставляємо норму висіву $X=4$ млн шт/га.

$$\begin{aligned} y &= 619,7312 \cdot 4 - 75,9 \cdot 4^2 - 1211,4467 = 2478,9248 - 1214,4 - 1211,4467 = \\ &= 53,0781 \approx 53,1 \text{ ц / га.} \end{aligned}$$

Точність прогнозування урожайності зерна пшениці озимої за нормою висіву розраховують за формулою

$$T\% = \frac{53,1 \cdot 100}{55} = 96,54\%.$$

Отже прогнозування за отриманою формулою точне, що вказує на можливість практичного використання виведеного рівняння.

Додаток 1

Значення критерію t для 5 і 1 % рівня значущості

Число ступенів вільності	Рівень значущості		Число ступенів вільності	Рівень значущості	
	0,05	0,01		0,05	0,01
1	12,71	63,66	18	2,10	2,88
2	4,30	9,93	19	2,09	2,86
3	3,18	5,84	20	2,09	2,85
4	2,78	4,60	21	2,08	2,83
5	2,57	4,03	22	2,07	2,82
6	2,45	3,71	23	2,07	2,81
7	2,37	3,50	24	2,06	2,80
8	2,31	3,36	25	2,06	2,79
9	2,26	3,25	26	2,06	2,78
10	2,23	3,17	27	2,05	2,77
11	2,20	3,11	28	2,05	2,76
12	2,18	3,06	29	2,05	2,76
13	2,16	3,01	30	2,04	2,75
14	2,15	2,98	50	2,01	2,68
15	2,13	2,95	100	1,98	2,63
16	2,12	2,92	∞	1,96	2,58
17	2,11	2,90			

Додаток 2

Значення критерію τ для 5 і 1 % рівня значущості

n	τ		n	τ	
	0,05	0,01		0,05	0,01
4	0,955	0,991	14	0,395	0,502
5	0,807	0,916	16	0,369	0,472
6	0,689	0,805	18	0,349	0,449
7	0,610	0,740	20	0,334	0,430
8	0,554	0,683	22	0,320	0,414
9	0,512	0,635	24	0,309	0,400
10	0,477	0,597	26	0,299	0,389
11	0,450	0,566	28	0,291	0,378
12	0,428	0,541	30	0,283	0,369

Додаток 3

Значення критерію F на 5 % рівні значущості (імовірності 95 %)

Ступені вільності для меншої дисперсії (знаменника)	Ступені вільності для більшої дисперсії (чисельника)													
	1	2	3	4	5	6	7	8	9	10	12	24	50	100
1	161	200	216	225	230	234	237	239	241	242	244	249	252	253
2	18,5	19	19,2	19,3	19	19,3	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5
3	10,1	9,6	9,28	9,12	9	8,94	8,88	8,84	8,81	8,78	8,74	8,64	8,58	8,56
4	7,71	6,9	6,59	6,39	6,3	6,16	6,09	6,04	6	5,96	5,91	5,77	5,7	5,66
5	6,61	5,8	5,41	5,19	5,1	4,95	4,88	4,82	4,78	4,74	4,68	4,53	4,44	4,4
6	5,99	5,1	4,76	4,53	4,4	4,27	4,21	4,15	4,1	4,06	4	3,84	3,75	3,71
7	5,59	4,7	4,35	4,12	4	3,87	3,79	3,73	3,68	3,63	3,57	3,41	3,32	3,28
8	5,32	4,5	4,07	3,84	3,7	3,58	3,5	3,44	3,39	3,34	3,28	3,12	3,03	2,98
9	5,12	4,3	3,86	3,63	3,5	3,37	3,29	3,23	3,18	3,13	3,07	2,9	2,8	2,76
10	4,96	4,1	3,71	3,48	3,3	3,22	3,14	3,07	3,02	2,97	2,91	2,74	2,64	2,59
11	4,84	4	3,59	3,36	3,2	3,09	3,01	2,95	2,9	2,86	2,79	2,61	2,5	2,45
12	4,75	3,9	3,49	3,26	3,1	3	2,92	2,85	2,8	2,76	2,69	2,5	2,4	2,35
13	4,46	3,8	3,41	3,18	3	2,92	2,84	2,77	2,72	2,67	2,6	2,42	2,32	2,26
14	4,6	3,7	3,34	3,11	3	2,85	2,77	2,7	2,65	2,6	2,53	2,35	2,24	2,19
15	4,54	3,6	3,29	3,06	2,9	2,79	2,7	2,64	2,59	2,55	2,48	2,29	2,18	2,12
16	4,49	3,6	3,24	3,01	2,9	2,74	2,66	2,59	2,54	2,49	2,42	2,24	2,13	2,07
17	4,45	3,6	3,2	2,96	2,8	2,7	2,62	2,55	2,5	2,45	2,38	2,19	2,08	2,02
18	4,41	3,6	3,16	2,94	2,8	2,66	2,58	2,51	2,46	2,41	2,34	2,15	2,04	1,98
19	4,38	3,5	3,13	2,9	2,7	2,63	2,55	2,48	2,43	2,38	2,31	2,11	2	1,94
20	4,35	3,5	3,1	2,87	2,7	2,6	2,52	2,45	2,4	2,35	2,28	2,08	1,96	1,9
21	4,32	3,5	3,07	2,84	2,7	2,57	2,49	2,42	2,37	2,32	2,25	2,05	1,93	1,87
22	4,3	3,4	3,05	2,82	2,7	2,55	2,47	2,4	2,35	2,3	2,23	2,03	1,91	1,84
23	4,28	3,4	3,03	2,8	2,6	2,53	2,45	2,38	2,32	2,28	2,2	2	1,88	1,82
24	4,26	3,4	3,01	2,78	2,6	2,51	2,43	2,36	2,3	2,26	2,18	1,98	1,86	1,8
25	4,24	3,4	2,99	2,76	2,6	2,49	2,41	2,34	2,25	2,24	2,16	1,96	1,84	1,77
26	4,22	3,4	2,98	2,74	2,6	2,47	2,39	2,32	2,27	2,22	2,15	1,95	1,82	1,76
28	4,2	3,3	2,95	2,71	2,6	2,44	2,36	2,29	2,24	2,19	2,12	1,91	1,78	1,72
30	4,17	3,3	2,92	2,69	2,5	2,42	2,34	2,27	2,21	2,12	2,09	1,89	1,76	1,69
40	4,08	3,2	2,84	2,61	2,5	2,34	2,25	2,18	2,12	2,07	2	1,79	1,66	1,59
50	4,03	3,2	2,79	2,56	2,4	2,29	2,2	2,13	20,7	2,02	1,95	1,74	1,6	1,52
100	3,94	3,1	2,7	2,46	2,3	2,19	2,1	2,03	1,97	1,92	1,85	1,63	1,48	1,39

Додаток 4

Значення критерія F на 5 % рівні значущості (імовірності 99 %)

Ступені свободи для меншої дисперсії (знаменника)	Ступені свободи для більшої дисперсії (чисельника)													
	1	2	3	4	5	6	7	8	9	10	12	24	50	100
1	4052	4999	5403	5625	5764	5889	5928	5981	6022	6,56	6106	6324	6302	6334
2	98,5	99	99,2	99,3	99,3	99,3	99,3	99,4	99,4	99,4	99,4	99,5	99,5	99,5
3	34,1	30,8	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	27,1	26,6	26,4	26,2
4	21,2	18	16,7	16	15,5	15,2	15	14,8	14,7	14,5	14,4	13,9	13,7	13,6
5	16,3	13,3	12,1	11,4	11	10,7	10,5	10,3	10,2	10,1	9,89	9,47	9,24	9,13
6	13,7	10,9	9,78	9,15	8,75	8,47	8,26	8,1	7,98	7,87	7,72	7,31	7,09	6,99
7	12,3	9,55	8,45	7,85	7,46	7,19	7	6,84	6,71	6,62	6,47	6,07	5,85	5,75
8	11,3	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,67	5,28	5,06	4,96
9	10,6	8,02	6,99	6,42	6,06	5,8	5,62	5,47	5,35	5,26	5,11	4,73	4,51	4,41
10	10	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,71	4,33	4,12	4,01
11	9,85	7,2	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,4	4,02	3,8	3,7
12	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,5	4,39	4,3	4,16	3,78	3,56	3,46
13	9,07	6,7	5,74	5,2	4,86	4,62	4,44	4,3	4,19	4,1	3,96	3,59	3,37	3,27
14	8,86	6,51	5,56	5,03	4,69	4,46	4,29	4,14	4,03	3,94	3,8	3,43	3,21	3,11
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4	3,89	3,8	3,67	3,29	3,07	2,97
16	8,53	6,23	5,29	4,77	4,44	4,2	3,89	3,78	3,69	3,61	3,45	3,18	2,96	2,86
17	8,4	6,11	5,18	4,67	4,34	4,1	3,93	3,79	3,68	3,59	3,45	3,08	2,86	2,76
18	8,28	6,01	5,09	5,58	4,25	4,01	3,85	3,71	3,6	3,51	3,37	3	2,78	2,68
19	8,18	5,93	5,01	4,5	4,17	3,94	3,77	3,68	3,52	2,43	3,3	2,92	2,7	2,63
20	8,1	5,85	4,94	4,43	4,1	3,87	3,71	3,56	3,45	3,37	3,23	2,86	2,63	2,53
21	8,02	5,78	4,87	4,37	4,04	3,81	3,65	3,51	3,4	3,31	3,17	2,8	2,58	2,47
22	7,94	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,75	2,53	2,42
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,3	3,21	3,07	2,7	2,48	2,37
24	7,82	5,61	4,72	4,22	3,9	3,67	3,5	3,36	3,25	3,17	3,03	2,66	2,44	2,33
25	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,21	3,13	2,99	2,62	2,4	2,29
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,17	3,09	2,96	2,58	2,36	2,25
28	7,64	5,45	4,57	4,07	3,76	3,53	3,36	3,23	3,11	3,03	2,9	2,52	2,3	2,18
30	7,56	5,39	4,51	4,02	3,7	3,47	3,3	3,17	3,06	2,98	2,84	2,47	2,24	2,13
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,88	2,8	2,66	2,29	2,05	1,94
50	7,17	5,06	4,2	3,72	3,41	3,18	3,02	2,88	2,78	2,7	2,56	2,18	1,94	1,81
100	6,9	4,82	3,98	3,51	3,2	2,99	2,82	2,69	2,59	2,51	2,36	1,98	1,73	1,59

СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Дідора В. Г., Смаглий О. Ф., Ермантраут Е. Р. Методика наукових досліджень в агрономії : навч. посіб. Київ : Центр учбової літератури, 2013. 264 с.
2. Основи наукових досліджень в агрономії : підруч. / В. О. Єщенко та ін. Вінниця : ПП «ГД «Едельвейс і К»», 2014. 332 с.
3. Дослідна справа в агрономії: навч. посібник у 2 кн. – Кн. 1. Теоретичні аспекти дослідної справи / А. О. Рожков та ін.; за ред. А. О. Рожкова. Харків : Майдан, 2016. 316 с.
4. Дисперсійний і кореляційний аналіз результатів польових дослідів : навч. посіб. / В. О. Ушкаренко та ін. Херсон : Айлант , 2009. 372 с.
5. Манько Ю. П., Цюк О. А., Павлов О. С. Методологія, методи і методика досліджень в агрономії : навч. посіб. Вінниця : Тов «Нілан-ЛТД», 2016. 96 с.
6. Партика З. В. Основи наукових досліджень: підготовка дисертацій : навч. посіб. Київ : Видавництво Ліра-К, 2017. 232 с.
7. Колесников О. В. Основи наукових досліджень : навч. посіб. Київ : Центр навчальної літератури, 2019. 144 с.
8. Євтушенко М. Ю., Хижняк М. І. Методологія та організація наукових досліджень : навч. посіб. Київ : Центр навчальної літератури, 2019. 350 с.
9. Надикто В. Т. Основи наукових досліджень : підруч. Херсон, 2017. 268 с.

Навчальне видання

ОСНОВИ ДОСЛІДНИЦЬКОЇ СПРАВИ

Методичні рекомендації

Укладач: Смірнова Ірина Вікторівна

Формат 60x84 1/16. Ум. друк. арк. 3,0.

Тираж 20 прим. Зам. № _____

Надруковано у видавничому відділі
Миколаївського національного аграрного університету
54020, м. Миколаїв, вул. Георгія Гонгадзе, 9

Свідоцтво суб'єкта видавничої справи ДК № 4490 від

20.02.2013р.

