



Національний університет
водного господарства та природокористування

Міністерство освіти і науки України
Національний університет водного господарства та
природокористування

Кафедра агрохімії, ґрунтознавства та землеробства

05-01-40

МЕТОДИЧНІ ВКАЗІВКИ
до виконання практичних робіт з навчальної
дисципліни “Основи наукових досліджень”
студентами напряму підготовки 6.090101 “Агрономія”
денної форми навчання

Рівне – 2016



Методичні вказівки до виконання практичних робіт з навчальної дисципліни “Основи наукових досліджень” студентами напряму підготовки 6.090101 “Агрономія” денної форми навчання / Трушева С.С. – Рівне: НУВГП, 2016. – 27 с.

Упорядник: С.С.Трушева, кандидат сільськогосподарських наук, доцент.

Відповідальний за випуск: С.Т. Вознюк, доктор сільськогосподарських наук, професор, завідувач кафедри агрохімії, ґрунтознавства та землеробства.

З М І С Т

	стр.
Практичні роботи №№ 1-2. Статистичні показники якісної та кількісної мінливості	3
Практична робота № 3. Основи дисперсійного аналізу...	9
Практична робота № 4. Дисперсійний аналіз даних однофакторного вегетаційного дослідження.....	12
Практична робота № 5. Дисперсійний аналіз даних багатофакторного вегетаційного дослідження	16
Практична робота № 6. Дисперсійний аналіз даних однофакторного польового дослідження з однорічними культурами	18
Практична робота № 7. Дисперсійний аналіз даних багатофакторного польового дослідження	22
Практична робота № 8. Пробіт-аналіз	25
Література	27

© Трушева С.С., 2016

© Національний університет
водного господарства та
природокористування, 2016



Практичні роботи №№ 1-2. СТАТИСТИЧНІ ПОКАЗНИКИ КІЛЬКІСНОЇ ТА ЯКІСНОЇ МІНЛИВОСТІ

Мета роботи: ознайомитися з поняттями: мінливість, вибірка, генеральна сукупність, варіаційний ряд, а також з типами мінливості та статистичними показниками, що їх характеризують.

Математична статистика – один з розділів математики. Вона дозволяє роботи висновки щодо всієї (генеральної) сукупності на основі спостережень за вибірковою сукупністю, або вибіркою. Усі статистичні методи базуються **на теорії ймовірності** – науці, яка вивчає загальні закономірності в масових випадкових явищах різної природи, і використовуються всюди, де доводиться мати справу з плануванням експериментів і досліджень, з оцінкою параметрів і перевіркою гіпотез, з прийняттям рішень щодо вивчення складних систем.

Будь-яке масове, множинне явище, наприклад, група рослин на полі або тварин на фермі, уявляє собою сукупність особин, випадків, фактів, предметів, тобто деяких умовних одиниць, кожна з яких, взята окремо, строго індивідуальна і відрізняється від інших рядом ознак – висотою, масою, кількістю продукції тощо. Кожна з ознак може мати у різних особин різний ступінь вираженості, тому говорять, що ознака варіює. **Властивість умовних одиниць** – рослин, урожаїв на паралельних ділянках польового дослідження і т.п. – відрізнятися одна від одної навіть в однорідних сукупностях називається **мінливістю** або **варіюванням**.

Мінливість – це властивість, яка притаманна усім предметам природи: двох абсолютно однакових предметів не існує, хоча відмінності між ними можуть бути непомітними для неозброєного ока. Ознаками, що варіюють, у рослин є, наприклад, їх висота, кількість і маса зерен в колосі, вміст протеїну і т.д. Варіювання виникає внаслідок того, що рослини одного й того ж сорту завжди відрізняються своєю спадковістю, крім того, формування їх часто відбувається за відносно різних умов зовнішнього середовища. Відповідно, при будь-якому дослідженні дані дослідів будуть завжди варіювати в тих чи інших межах.



Мінливість ознак створює відому складність у тих випадках, коли вимагається дати загальну характеристику певній групі рослин, тварин, ґрунтів, що варіює (сукупності), за окремими ознаками або порівняти дві такі групи і знайти відмінності між ними. Абсолютно очевидно, що не завжди можливо (а практично дуже рідко) дослідити за тією чи іншою ознакою всі особини, усю сукупність. В цих випадках вдаються до вивчення лише її частини, за якою роблять загальний висновок. Такий метод називається **вибірковим** і вважається основним при статистичному вивченні сукупності.

Таким чином, *усю групу об'єктів, що підлягає вивченню, називають сукупністю або генеральною сукупністю, а ту частину об'єктів, котра потрапила на перевірку, дослідження, – вибірковою сукупністю (вбіркою)*. Число елементів в генеральній сукупності і вибірці називають їх об'ємом.

Головна мета вибіркового методу – за статистичними показниками малої вибірки (середній пробі) якомога точніше охарактеризувати всю сукупність об'єктів.

В результаті спостережень ми отримуємо відомості щодо чисельної величини ознаки, що вивчається, у кожного члена даної вибіркової сукупності. Можливі значення варіюючої ознаки X називають **варіантами** і позначають X_1, X_2, \dots, X_n . Отриманий таким чином ряд варіюючих величин можна впорядкувати – розташувати значення ознаки (варіанти) в порядку їх зростання (або убутання). Таке упорядкування ряду називають його **ранжуванням**. Після ранжування неважко помітити, що кожне значення ознаки зустрічається неоднакову кількість разів – одні рідко, інші частіше. *Числа, котрі характеризують, скільки разів повторюється кожне значення ознаки у членів даної сукупності, називаються частотами ознаки* і позначаються **f**.

Сума всіх частот ($\sum f$) дорівнює об'єму вибірки, тобто числу членів ряду – n . У результаті такої обробки первинних спостережень отримують так званий варіаційний ряд. **Варіаційний рядом** називається *такий ряд даних, в котрих вказані можливі значення варіюючої ознаки в порядку зростання або убутання і відповідні їм частоти*.

Розрізняють 2 типи мінливості: **кількісну**, котра може бути виміряна, і **якісну**, котра не підлягає вимірюванню.



Під кількісною мінливістю розуміють таку, в котрій відмінності між варіантами виражаються кількістю, наприклад, масою, висотою, урожаєм, числом зерен і т.д. Розрізняють 2 види кількісної мінливості: *переривчасту* (або дискретну) і *безперервну*.

У першому випадку відмінності між варіантами виражаються цілими числами, між якими немає і не може бути переходів, наприклад, число рослин на 1 м², число зерен у колосі і т.д. У другому випадку значення варіант виражаються мірами об'єму, довжини, маси і т.д., між якими можливі будь-які переходи з необмеженим числом можливих значень; усе залежить від ступеня точності, який приймається для характеристики даної кількісної ознаки.

Якісною мінливістю називається таке варіювання, коли відмінності між варіантами виражаються якісними показниками, котрі одні варіанти мають, а інші ні. (колір, смак, форма об'єкта, що вивчається). Якщо ознака приймає тільки два значення (*хворий – здоровий, остистий – безостий*), що взаємно виключають один одного, то мінливість називається **альтернативною**.

Основними статистичними характеристиками кількісної мінливості є: середнє арифметичне (\bar{x}); дисперсія (S^2); стандартне відхилення (S); похибка середнього арифметичного ($S\bar{x}$); коефіцієнт варіації (V); відносна похибка вибіркового середнього ($S\bar{x}\%$).

Середнє арифметичне представляє собою узагальнену, абстрактну характеристику всієї сукупності в цілому. Якщо суму всіх варіантів ($X_1+X_2+\dots+X_n$) позначити через $\sum X$, а число всіх варіантів через n , то формула для визначення простого середнього арифметичного буде мати наступний вигляд:

$$\bar{x} = \sum_{i=1}^n \frac{X}{n} \quad (1.1)$$

Основна властивість середнього арифметичного полягає в рівності суми всіх позитивних і всіх від'ємних відхилень від нього, тобто сума центральних відхилень всіх окремих варіантів від $\bar{x} = 0$. Якщо $\sum (X - \bar{x}) \neq 0$, то значить, допустилися помилки в розрахунках.

Дисперсія і стандартне відхилення є основними мірами варіації, розсіювання ознаки, що вивчається. Дисперсія представляє собою:



$$s^2 = \sum_{i=1}^n \frac{(X - \bar{x})^2}{n - 1} \quad (1.2)$$

де $\sum(X - \bar{x})^2$ - сума квадратів відхилень; n - число всіх вимірювань.

Розмірність дисперсії дорівнює квадрату розмірності ознаки, що вивчається, що незручно і змушує ввести для вимірювання розсіювання іншу характеристику, що має розмірність величини, що варіює, і називається **стандартним** або **середнім квадратичним відхиленням**. Його отримують вилученням квадратного кореня з дисперсії:

$$s = \sqrt{s^2} \quad (1.3)$$

Із математичної статистики відомо, що при визначенні будь-яких середніх величин суми всіх показників необхідно ділити на число незалежних одна від одної величин. У зв'язку з цим в формулах суму квадратів відхилень ділять не на загальне число спостережень, а на число без 1, так як одне будь-яке відхилення залежне і може бути знайдене з рівняння $\sum(X - \bar{x}) = 0$. Інші відхилення можуть вільно варіювати, приймати будь-які значення. Число величин, що вільно варіюють, називають **числом степенів свободи** (ν).

Стандартне відхилення слугує показником, який дає уявлення щодо найбільш вірогідної похибки окремого, одиничного спостереження, взятого з даної сукупності. В межах одного значення ($\pm 1s$) вкладається приблизно 2/3 усіх спостережень, або, точніше, 68,3% всіх варіантів, тобто основне ядро ряду величин, що вивчаються. Тому стандартне відхилення називають також **основним відхиленням варіаційного ряду**.

Відповідно, можливі відхилення від \bar{x} , які перевищують $\pm 1s$, але вірогідність їх по мірі віддалення відхилень \bar{x} від $\pm 1s$ увесь час зменшується. Так, вірогідність зустріти варіанту, що відхиляється від \bar{x} на величину більшу за $\pm 3s$, складає всього близько 0,3%. Тому потроєне значення стандартного відхилення прийнято вважати граничною похибкою окремого спостереження, і, відповідно, майже всі значення варіант у варіаційному ряду вкладаються в межах $\pm 3s$.

Коефіцієнт варіації – стандартне відхилення, виражене в відсотках до середнього арифметичного даної сукупності:

$$V = \frac{s}{\bar{x}} \cdot 100 \quad (1.4)$$



Коефіцієнт варіації є відносним показником мінливості.

Використання коефіцієнту має зміст при вивченні варіації ознаки, яка приймає тільки позитивні значення. Немає сенсу, наприклад, коефіцієнт варіації, вирахований для характеристики коливання середньорічної температури, наближеної до 0°C , коли варіююча ознака приймає як позитивні, так і від'ємні значення.

Мінливість прийнято вважати *незначною*, якщо коефіцієнт варіації не перевищує 10% , *середньою* – якщо V вище 10%, але менше 20% та *значною*, якщо V більше 20%.

Похибка вибіркового середнього є мірою відхилення вибіркового середнього від середньої усієї (генеральної) сукупності.

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \sqrt{\frac{s^2}{n}} \quad (1.5)$$

де s – стандартне відхилення; n – число всіх вимірювань.

Похибки вибірки виникають внаслідок неповної репрезентативності вибіркової сукупності і притаманні тільки вибіркового методу дослідження. Величина похибок залежить від ступеня мінливості ознаки, що вивчається, і від об'єму вибірки.

Похибки вибірки виражають в тих же одиницях вимірювання, що й варіюючу ознаку, і дописують до відповідних середніх зі знаками \pm , тобто $\bar{x} \pm s_{\bar{x}}$.

Похибка вибірки, виражена у відсотках від відповідного середнього, називається **відотною похибкою вибіркового середнього**:

$$s_{\bar{x}\%} = \frac{s_{\bar{x}}}{\bar{x}} \cdot 100 \quad (1.6)$$

У біологічних і агрономічних дослідженнях часто доводиться мати справу з **якісною мінливістю** ознак. Основними статистичними показниками якісної мінливості є: частка ознаки; показник мінливості; коефіцієнт варіації; похибка вибіркової частки.

Частка ознаки позначається через p_1, p_2, p_3 і т.д. і може бути виражена в долях одиниці або у відсотках. В першому випадку сума усіх часток в межах даної сукупності або ряду розподілення дорівнює 1, а в другому – 100%. **Частка ознаки** – це відношення чисельності кожного з членів ряду n_1, n_2, n_3 і т.д. до чисельності сукупності N , тобто вірогідність появи даної ознаки у сукупності, що вивчається:



$$p_1 = \frac{n_1}{N} \quad (1.7)$$

При альтернативній мінливості частка однієї ознаки позначається через p , а другої через q . На основі рівняння $p + q = 1,0$ (або 100%), оскільки вірогідність двох протилежних подій завжди дорівнює 1 (100%), значення $q = 1 - p$.

Показник мінливості якісної ознаки s характеризує варіювання величин ряду відносно одна одної. Значення показника мінливості визначають за формулою:

$$S = \sqrt[k]{p_1 p_2 \dots p_k} \quad (1.8)$$

де p_1, p_2, p_k і т.д. – частки ознаки (або їх процентні значення) в загальній сукупності; k – число градацій ознаки.

Максимальна мінливість (s_{\max}) спостерігається при $p=q=0,5$.

Користуючись значеннями s_{\max} , можна вирахувати **коефіцієнт варіації якісних ознак** – фактичний показник мінливості, виражений в % до максимально можливої мінливості:

$$V_p = \frac{S}{S_{\max}} \cdot 100 \quad (1.9)$$

Коефіцієнт варіації характеризує відносний ступінь мінливості ознак, що вивчаються, і широко використовується для порівняльної оцінки вирівненості різних сукупностей. Максимальне значення $V_p = 100\%$ спостерігається при $s = s_{\max}$.

Похибка вибіркової частки – міра відхилення частки ознаки вибіркової сукупності p від частки його по всій генеральній сукупності P внаслідок неповного представництва (репрезентативності) вибірки. Похибку частки вираховують за формулою:

$$s_p = \frac{S}{\sqrt{N}} \quad (1.10)$$

де S – показник мінливості якісної ознаки; N - об'єм вибірки.

Для альтернативного варіювання, коли значення $s_p = \sqrt{pq}$, формула похибки вибіркової частки прийме вигляд:

$$s_p = \sqrt{\frac{pq}{N}} \quad (1.11)$$



В цій формулі p і q можуть бути виражені в частках одиниці або у %. Вірогідність зустріти p (або q) в інтервалі $p \pm s_p$ складає близько 68%, в інтервалі $p \pm 2s_p$ – 95%, в інтервалі $p \pm 3s_p$ – близько 99%. Відповідно, подібно до кількісної мінливості, усі значення p з вірогідністю 99% вкладаються в межах потрібної похибки вибіркової частки.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Що таке варіація, варіаційний ряд, об'єм вибірки?
2. Яка відмінність між кількісною та якісною мінливістю?
3. Назвіть статистичні показники кількісної мінливості.
4. Яка мінливість називається альтернативною?
5. Що називають часткою ознаки?

Практична робота № 3. ОСНОВИ ДИСПЕРСІЙНОГО АНАЛІЗУ

Мета роботи: ознайомитися з сутністю та основними поняттями дисперсійного аналізу.

Дисперсійний аналіз розроблений і введений в практику сільськогосподарських досліджень англійським вченим Р.Фішером (1925), який відкрив закон розподілення відношення середніх квадратів (дисперсій). **Сутністю дисперсійного аналізу є розчленування загальної суми квадратів відхилень і загального числа степенів свободи на частини – компоненти, що відповідають структурі експерименту, та оцінка значимості дії і взаємодії факторів, що вивчаються, за F -критерієм.**

Якщо обробляють однофакторні статистичні комплекси, що складаються з декількох незалежних вибірок, наприклад l -варіантів в вегетаційному досліді, то загальна мінливість результативної ознаки, що вимірюється загальною сумою квадратів S_y , розчленовується на 2 компоненти:

- 1 – варіювання між вибірками (варіантами) – S_v ;
- 2 – варіювання в середині вибірок – S_z .



Відповідно, в загальному вигляді мінливість ознаки може бути представлена виразом: $C_y = C_v + C_z$.

Тут варіація між варіантами представляє ту частину загальної дисперсії, котра обумовлена дією факторів, що вивчаються, а дисперсія всередині вибірок характеризує похибку експерименту.

Загальне число степенів свободи ($n - 1$) також розчленовується на дві частини:

- ступені свободи для варіантів ($l - 1$)
- ступені свободи для випадкового варіювання ($N - l$):

$$n - 1 = (l - 1) + (N - l) \quad (2.1)$$

Якщо обробляти однофакторні статистичні комплекси, коли варіанти (вибірки) пов'язані будь-якою спільною умовою, що контролюється, наприклад, наявністю організованих повторень (n) в польовому досліді, загальна сума квадратів розкладається на 3 частини :

1 – варіювання повторень – C_p ;

2 – варіювання варіантів – C_v ;

3 – випадкове варіювання - C_z .

У подібних випадках загальна мінливість і загальне число степенів свободи можуть бути представлені виразами:

$$C_y = C_p + C_v + C_z \quad (2.2)$$

$$n - 1 = (n - 1) + (l - 1) + (n - l) \cdot (l - 1) \quad (2.3)$$

Суми квадратів відхилень за даними польового досліді – статистичного комплексу з l – варіантами та n – повтореннями – знаходять в такій послідовності: в вихідній таблиці визначають суми по повтореннях, варіантах і загальну суму всіх спостережень $\sum X$. Потім вираховують:

1) загальне число спостережень $N = ln$

2) поправку $C = (\sum X)^2 : N$

3) загальну суму квадратів $C_y = \sum X^2 - C$

4) суму квадратів для повторень $C_p = \sum P^2 : l - C$

5) суму квадратів для варіантів $C_v = \sum V^2 : n - C$

6) суму квадратів для похибки (залишок) $C_z = C_y - C_p - C_v$.

Дві останні суми C_v і C_z ділять на відповідні їм ступені свободи. В результаті отримують два середніх квадрата (дисперсії): варіантів та похибки. Ці середні квадрати і використовують в дисперсійному аналізі для оцінки значимості дії факторів, що вивчаються. Оцінка



здійснюється шляхом порівняння дисперсії варіантів з дисперсією похибки за критерієм $F = s_v^2 / s^2$.

Таким чином за базу - одиницю порівняння приймають середній квадрат випадкової дисперсії, котра визначає випадкову похибку експерименту.

При цьому нульовою гіпотезою, що перевіряється, слугує припущення: усі вибіркові середні є оцінками однієї генеральної середньої і, відповідно, різниці між ними неістотні. Якщо $F_{\text{факт}} \leq F_{\text{теор}}$, то нульова гіпотеза $H_0 : d = 0$ не відкидається; між усіма вибірковими середніми немає істотних різниць, і на цьому перевірка закінчується. Нульова гіпотеза відкидається, коли $F_{\text{факт}} > F_{\text{теор}}$.

Теоретичні значення критерію F для прийнятого в дослідженні рівня значимості знаходять за довідковими таблицями з урахуванням числа степенів свободи для дисперсії варіантів і випадкової дисперсії. У більшості випадків обирають 5% рівень, а при більш точному підході – 1% рівень значимості. 5% рівню значимості відповідає 95% рівень ймовірності, 1% – 99% рівень.

Критерій F встановлює тільки факт наявності істотної різниці між середніми, але не вказує, між якими середніми є ці відмінності. Тому, коли $F_{\text{факт}} \geq F_{\text{теор}}$ і нульова гіпотеза відкидається, треба визначити, до яких варіантів належать істотні різниці.

У практиці дослідної роботи використовується декілька методів для оцінки істотності різниці між середніми. Обираємо найбільш поширений, а саме: оцінка значимості різниці між середніми за **найменшою істотною різницею (НІР)**. **НІР** – це величина, що вказує межі граничним випадковим відхиленням.

Критерій $\text{НІР} = t \cdot s_d$ показує граничну похибку для різниці двох вибіркових середніх. Якщо фактична різниця $d \geq \text{НІР}$, то вона істотна (значима), а якщо $d < \text{НІР}$, то неістотна, не значима.

Щоб визначити НІР, необхідно за даними дисперсійного аналізу вирахувати:

- узагальнену похибку середнього;
- похибку різниці середніх.

Підставляючи значення s_d в формулу для розрахунку НІР і отримуємо:

$$\text{НІР}_{05} = t_{05} \cdot s_d \quad (2.4)$$

$$\text{НІР}_{01} = t_{01} \cdot s_d \quad (2.5)$$



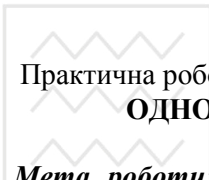
де s_d – похибка різниці середніх незалежних вибірок; t_{05} , t_{01} – критерій Стьюдента.

Індексами при НІР і t записані показники рівня значимості (5% або 1%).

Різниці між середніми, які є більшими за НІР₀₅, вважаються істотними з 5%-ним рівнем значимості, більше НІР₀₁ – істотними з 1%-ним рівнем значимості.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. У чому сутність дисперсійного аналізу?
2. Чим відрізняється статистична обробка даних однофакторного статистичного комплексу від обробки багатфакторного статистичного комплексу?
3. Що таке найменша істотна різниця?



Практична робота № 4. ДИСПЕРСІЙНИЙ АНАЛІЗ ДАНИХ ОДНОФАКТОРНОГО ВЕГЕТАЦІЙНОГО ДОСЛІДУ

Мета роботи: оволодіти практичними навичками проведення дисперсійного аналізу урожайних даних однофакторного вегетаційного дослідження.

Вегетаційні дослідження частіше за все представляють собою статистичні комплекси, які складаються з декількох незалежних варіантів-випірок. Незалежність варіантів, що порівнюються, досягається регулярним переміщенням посудин на вагонетці. Відповідно, в вегетаційних дослідженнях зазвичай немає територіально організованих повторень. У таких випадках дисперсійний аналіз слід проводити як для не поєднаних (незв'язаних) вибірок. Коли в вегетаційному дослідженні варіанти поєднують територіально у повторення, то статистичний аналіз проводять так само, як і для польового дослідження.

Перед дисперсійним аналізом даних вегетаційного дослідження поставлено завдання перевірити статистичну нульову гіпотезу H_0 , котра формулюється таким чином: між середніми по варіантах немає істотної різниці. Коротко нульова гіпотеза записується $H_0 : d = 0$.



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

- варіювання варіантів;
- випадкове варіювання: $C_y = C_v + C_z$.

Статистичний аналіз даних проводять в 3 етапи:

1. Складають розрахункову таблицю, розташовуючи в ній вихідні дані по рядках і стовбцях, визначають суми і середні по варіантам, загальну суму і середнє значення результативної ознаки по досліді.

2. Вираховують суми квадратів відхилень за формулами табл. 4.1 і визначають фактичне значення критерію F_ϕ .

3. Визначають похибку досліді та істотність часткових відмінностей.

Таблиця 4.1

Формули для обрахунку сум квадратів відхилень, дисперсій і критерію F_ϕ

Дисперсія	Сума квадратів	Степені свободи	Середній квадрат	F_ϕ
Загальна C_y	$\sum X^2 - C$	$N - l$	-	
Варіантів C_v	$\sum V^2 : n - C$	$l - 1$	$s_v^2 = \frac{C_v}{l - 1}$	s_v^2 / s^2
Залишок C_z (похибки)	$C_y - C_v$	$N - l$	$s^2 = \frac{C_z}{N - 1}$	
Поправка $C = \sum (X)^2 : N$				

Завдання № 1

Обробити дані вегетаційного досліді по вивченню дії співвідношення № : $P_2O_5 : K_2O$ при живленні розсади томатів на врожай плодів. Нульова гіпотеза $H_0 : d = 0$, тобто всі відмінності між середніми за варіантами статистично не істотні. Варіанти: **1** – контроль (1:1:1); **2** – 1:2:1; **3** – 1:2:2; **4** – 2:1:1; **5** – 2:2:1.

Хід рішення

1. Заповнюємо таблицю 4.2. Підраховуємо суми і середні за варіантами, загальну суму і середній врожай у досліді.
2. Суми квадратів та поправку обраховують безпосередньо за даними табл. 4.1.



3. Після оброблення сум квадратів складають таблицю дисперсійного аналізу (табл. 4.3).

Таблиця 4.2

Ранній врожай плодів томатів (г на посудину)

Варианти	Урожай, X				Число спостережень, n	Сума по варіантах, V	Середні по варіантах
1 (st)	454	470	430	500			
2	502	550	490	507			
3	601	670	550	607			
4	407	412	475	402			
5	418	470	460	412			
Загальна сума					$\sum n = N$	$\sum X =$	$\bar{x} =$

Таблиця 4.3

Результати дисперсійного аналізу

Дисперсія	Сума квадратів	Степені свободи	Середній квадрат	Критерій F	
				факт.	05
C_v				-	-
C_v					
C_z				-	-

- Порівнюємо значення $F_{\text{факт}}$ і F_{05} . Якщо нульова гіпотеза відкидається, то розраховують похибку досліду, s_d – похибку різниці середніх та HP_{05} .
- Складають підсумкову таблицю і роблять висновок.

Таблиця 4.4

Підсумкова таблиця дисперсійного аналізу

Співвідношення N:P:K	Урожай, г/посудину	Різниця зі стандартом		Група
		г	%	
1:1:1 (контроль)				
1:2:1				
1:2:2				
2:1:1				
2:1:1				
HP_{05}				



Завдання № 2

Встановити, чи є істотною різниця в дії форм азотних добрив на врожай вівсяниці лучної. Нульова гіпотеза $H_0 : d = 0$. Варіанти: 1 – без добрив; 2 – сульфат амонію; 3 – аміачна селітра; 4 – сечовина.

Хід рішення

1. Заповнюємо таблицю 4.5. Підраховуємо суми і середні за варіантами, загальну суму і середній врожай у досліді. **Звертаємо увагу на різну кількість повторень по варіантах в досліді. Це вимагає обрахування декількох значень НІР, тому що не всі середні рівноточні.**

Таблиця 4.5

Варі- анти	Врожай вівсяниці, г/посудину						Сума по варі- антах, V	Серед- ні по варіан- тах	
	Урожай, X								
1	16,0	17,2	14,4	15,8	-	-			
2	29,4	30,4	30,3	28,1	-	-			
3	26,0	29,2	26,7	27,1	26,0	28,1			
4	25,3	24,8	26,1	23,2	25,7	24,0			
Загальна сума							$\sum n = N$	$\sum X =$	$\bar{x} =$

- Суми квадратів обраховують безпосередньо за даними табл. 4.1.
- Після обрахування сум квадратів складають таблицю дисперсійного аналізу (табл. 4.6).

Таблиця 4.6

Результати дисперсійного аналізу					
Дисперсія	Сума квадратів	Степені свободи	Середній квадрат	Критерій F	
				факт.	05
C_v				-	-
C_v					
C_z				-	-

- Порівнюємо значення $F_{\text{факт}}$ і F_{05} . Якщо нульова гіпотеза відкидається, то розраховують похибку досліді, s_d – похибку різниці середніх та НІР₀₅.
- Складають підсумкову таблицю і роблять висновок.



Практична робота № 5. ДИСПЕРСІЙНИЙ АНАЛІЗ ДАНИХ БАГАТОФАКТОРНОГО ВЕГЕТАЦІЙНОГО ДОСЛІДУ

Мета роботи: оволодіти практичними навичками проведення дисперсійного аналізу урожайних даних багатofакторного вегетацийного дослiду.

Дисперсійний аналіз даних багатofакторного дослiду проводять в 2 етапи. Перший етап – розчленування загальної варіації результативної ознаки на варіювання варіантів і залишку: $C_y = C_v + C_z$. На другому етапі сума квадратів відхилення для варіантів розчленовується на компоненти, що відповідають джерелам варіювання, - головні ефекти факторів, що вивчаються, та їх взаємодії: $C_v = C_A + C_B + C_{AB}$.

Завдання.

У двофакторному дослiді 2×3 з ґрунтовою культурою ячменю вивчена дія двох доз азоту і трьох доз фосфору. Провести дисперсійний аналіз результатів дослiду. Нульова гіпотеза $H_0: d = 0$.

Хід рішення

1. Визначають суми і середні по варіантам, загальну суму і середній врожай по дослiду (табл. 5.1).

Таблиця 5.1

Урожай зерна ячменю, г/посудину

Азот (А)	Фосфор (В)	Урожай, X				Суми	Середні
a ₀	b ₀	24,1	25,8	23,0	27,0		
	b ₁	28,4	29,7	30,1	27,4		
	b ₂	28,7	30,4	32,0	27,0		
a ₁	b ₀	30,7	34,4	34,0	31,0		
	b ₁	46,7	45,4	47,1	46,3		
	b ₂	59,4	50,7	64,5	60,1		
Загальна сума						$\sum X =$	$\bar{x} =$

2. Обраховують загальну суму квадратів відхилень, суму квадратів для варіантів та залишку (табл. 4.1).

3. Для розрахунку сум квадратів по факторам А,В і взаємодії АВ складають допоміжну табл. 5.2.



Таблиця 5.2

Визначення сум для головних ефектів і взаємодії

Азот (А)	Фосфор (В)			Суми А
	b_0	b_1	b_2	
a_0				
a_1				
Суми В				$\sum X=$

4. Розраховуємо суму квадратів для фактора А:

$$C_A = \sum A^2 \div (l_B n) - C \quad (5.1)$$

5. Розраховуємо суму квадратів для фактора В:

$$C_B = \sum B^2 \div (l_A n) - C \quad (5.2)$$

6. Суму квадратів для взаємодії АВ знаходять за формулою:

$$C_{AB} = C_v - C_A - C_B \quad (5.3)$$

при $(l_A - 1) \cdot (l_B - 1)$ степенях свободи.

7. Заповнюють таблицю дисперсійного аналізу (табл. 5.3).

Таблиця 5.3

Результати дисперсійного аналізу

Дисперсія	Сума квадратів	Степені свободи	Середній квадрат	Критерій F	
				факт.	05
C_v				-	-
C_A					
C_B				-	-
C_{AB}					
C_z					

8. Для оцінки істотності часткових різниць обраховують:

$$s_{\bar{x}} = \sqrt{s^2 \div n} \quad (5.4)$$

де s^2 – дисперсія; n – число повторень.

$$s_d = \sqrt{2s^2 \div n} \quad (5.5)$$

$$HIP_{05} = t_{05} \cdot s_d \quad (5.6)$$

Результати статистичної обробки даних досліджу можна представити у вигляді табл. 5.4. та сформулювати висновки.



Дія азоту і фосфору на врожай ячменю, г/посудину

Дози азоту	Дози фосфору					
	0		1		2	
	урожай	різниця зі st	урожай	різниця зі st	урожай	різниця зі st
0		-				
1						
НІР ₀₅	-		-		-	

Практична робота № 6. **ДИСПЕРСІЙНИЙ АНАЛІЗ ДАНИХ
ОДНОФАКТОРНОГО ПОЛЬОВОГО
ДОСЛІДУ З ОДНОРІЧНИМИ КУЛЬТУРАМИ**

Мета роботи: оволодіти практичними навичками проведення дисперсійного аналізу урожайних даних однофакторного польового досліді з однорічними культурами.

Статистичну обробку даних проводять у такій послідовності:

1 – вихідні дані заносять в таблицю врожайів, визначають суми і середні;

2 – обраховують суми квадратів для всіх джерел варіювання;

3 – складають таблицю дисперсійного аналізу і перевіряють нульову гіпотезу щодо істотності дії і взаємодії факторів за F – критерієм. Якщо нульова гіпотеза не відкидається, то всі різниці між вибірковими середніми знаходяться в межах випадкових відхилень, і в цьому випадку обраховують тільки похибку досліді s_x^2 .

Завдання.

Провести дисперсійний аналіз даних польового досліді з різними сортами капусти. В варіанті 2 випала з обліку ділянка в IV, а в варіанті 5 – ділянка в III та IV повтореннях. Відновити дані, що "випали" і перевірити нульову гіпотезу. Згрупувати варіанти по відношенню до стандарту.

Хід рішення

Перш ніж проводити дисперсійний аналіз даних, необхідно привести результати досліді до порівняльного вигляду. Для цього:



1. В табл. 6.1. записують суми по повтореннях, включаючи ті варіанти, котрі мають повний набір ділянок; розраховують середні по повтореннях шляхом ділення сум на число варіантів, що мають повний набір дат.

Таблиця 6.1

Урожай стандартних головок капусти, ц/га

Варіанти	Повторення, X					Число спостережень
	I	II	III	IV	V	
1	560	542	574	537	510	
2	548	509	560	-	497	
3 (st)	595	569	631	515	501	
4	607	594	612	586	574	
5	629	601	-	-	597	
6	518	502	549	518	499	
Суми по повтореннях з повним набором варіантів						
Середні по 4 варіантах						

2. Для обрахунку теоретично очікуваних урожаїв на ділянках, що випали з обліку, складають допоміжну табл. 6.2., куди вносять врожаї по варіантах, в яких ділянки випали, і середні по повтореннях, обраховані для варіантів з повним набором дат.

Таблиця 6.2

Допоміжна таблиця для відновлення дат, що випали

Варіанти	Повторення, X					Суми	Середні для варіанта	
	I	II	III	IV	V		2	5
2								-
5							-	
Середні по 4 варіантах						-		
Ефекти варіантів	-	-	-	-	-	-		
Відновлений врожай								
2	-	-	-		-	-	-	-
5	-	-			-	-	-	-

3. Складають розрахункову таблицю для дисперсійного аналізу, в котрій відновлені врожаї вказані в дужках. Підраховують суми і середні по варіантах, суми по повтореннях, загальну суму і загальний врожай по досліду (табл. 6.3).



Таблиця 6.3

Урожай стандартних головок капусти, ц/га

Варіанти	Повторення, X					Суми, V	Середні
	I	II	III	IV	V		
1							
2							
3 (st)							
4							
5							
6							
Суми P						$\sum X =$	$\bar{x} =$

4. Для обрахунку сум квадратів вихідні дані доцільно перетворити за співвідношенням $X_l = X - A$, прийнявши за мовне середнє число, максимально наближене до середнього врожаю по досліді. Перетворені дані записують у табл. 6.4, сумують дані по варіантах, повтореннях і знаходять загальну суму $\sum P = \sum V = \sum X_l$.

Таблиця 6.4

Таблиця перетворених дат

Варіанти	$X_l = X - A$					Суми V
	I	II	III	IV	V	
1						
2						
3 (st)						
4						
5						
6						
Суми P						$\sum X_l =$

5. Обраховують суми квадратів відхилень за формулами, наведеними в табл. 6.5.

Таблиця 6.5

Формули для обрахунку сум квадратів відхилень, дисперсій і F_ϕ

Дисперсія	Сума квадратів	Степені свободи	Середній квадрат	F_ϕ
C_y	$\sum X^2 - C$	$N - 1$	-	
C_p	$\sum P^2 : l - C$	$n - 1$	-	
C_v	$\sum V^2 : n - C$	$l - 1$	$s_v^2 = C_v \div l - 1$	s_v^2 / s^2
C_z	$C_y - C_p - C_v$	$(l - 1) \cdot (n - 1)$	$s^2 = C_z \div (l - 1)(n - 1)$	



6. Отримані дані заносять в таблицю дисперсійного аналізу і обраховують значення F-критерію (табл. 6.6). При розрахунку числа степенів свободи для залишку необхідно залишкове число степенів свободи, яке визначається звичайним шляхом, зменшити на число дат, що випали.

Таблиця 6.6

Результати дисперсійного аналізу

Дисперсія	Сума квадратів	Степені свободи	Середній квадрат	Критерій F	
				факт.	05
C_v			-	-	-
C_p			-	-	-
C_v					
C_z				-	-

7. Визначення істотності часткових різниць у досліді з відновленими врожаєми має ту особливість, що необхідно враховувати число фактичних дат, які лежать в основі обрахунку статистичних показників. Обраховують:

а) середню похибку досліду

$$s_{\bar{x}} = \sqrt{\frac{s^2}{(n_1 + n_2 + \dots + n_n) \div l}} \quad (6.1)$$

б) похибки різниці середніх при порівнянні варіантів 1, 3, 4, 6 (n=5).

$$s'_d = \sqrt{2s^2 \div n} \quad (6.2)$$

варіанта 2 (n₁=4) з варіантами 1, 3, 4, 6 (n₂=5).

$$s''_d = \sqrt{s^2 \frac{n_1 + n_2}{n_1 n_2}} \quad (6.3)$$

варіанта 5 (n₁=3) з варіантами 1, 3, 4, 6 (n₂=5).

$$s'''_d = \sqrt{s^2 \frac{n_1 + n_2}{n_1 n_2}} \quad (6.4)$$

в) найменші істотні різниці для 5%-ного рівня значимості:

$$HIP'_{05} = t_{05} \cdot s'_d \quad (6.5)$$

$$HIP''_{05} = t_{05} \cdot s''_d \quad (6.6)$$

$$HIP'''_{05} = t_{05} \cdot s'''_d \quad (6.7)$$

8. Складають підсумкову таблицю і роблять висновок.



Практична робота № 7. ДИСПЕРСІЙНИЙ АНАЛІЗ ДАНИХ БАГАТОФАКТОРНОГО ПОЛЬОВОГО ДОСЛІДУ

Мета роботи: оволодіти практичними навичками проведення дисперсійного аналізу урожайних даних багатofакторного польового дослідy.

Статистичну обробку даних проводять в такій послідовності:

- 1 – вихідні дані заносять в таблицю врожаїв, визначають суми і середні;
- 2- обраховують суми квадратів для загального варіювання S_y , варіювання повторень S_p , варіантів S_v і залишку S_z ;
- 3 – загальне варіювання варіантів S_v розкладають на компоненти – головні ефекти факторів, що вивчаються, та їх взаємодії;
- 4 – складають таблицю дисперсійного аналізу і перевіряють нульову гіпотезу щодо істотності дії і взаємодії факторів за F – критерієм.

Багатofакторний дисперсійний комплекс – це сукупність вихідних спостережень (дані), які дозволяють статистично оцінити дію і взаємодію декількох факторів, що вивчаються, на мінливість результативної ознаки.

Завдання.

У двофакторному досліді 3×4 , який поставлений в 4-х рендомізованих повтореннях, вивчена дія 3-х градацій зрошення (0 – без зрошення; 1 – помірна; 2 – надмірна зрошення) та 4-х доз азоту (0 – без азоту; 1 – 60; 2 – 120; 3- 240 кг/га д.р.) на врожай насіння бавовнику. Провести дисперсійний аналіз даних.

Хід рішення

1. У табл. 7.1 визначають суми і середні. Правильність обрахунків перевіряють за співвідношенням $\sum P = \sum V = \sum X$.

2. Визначають суми квадратів відхилень (табл. 6.5). Загальне число спостережень визначають за формулою:

$$N = l_A \cdot l_B \cdot n \quad (7.1)$$

де l_A – число градацій фактору A ; l_B – число градацій фактору B ; n – число повторень.

3. Визначають суми квадратів для факторів A, B і взаємодії AB . Для цього складають табл. 7.2, в яку вписують суми врожаїв по



варіантах і знаходять необхідні для розрахунку головних ефектів суми А і В.

Таблиця 7.1

Вплив зрошення і доз азоту на врожай насіння бавовнику, ц/га

Зрошення А	Дози азоту В	Повторення Х				Суми V	Середні
		I	II	III	IV		
0	0	14	15	11	13		
	1	15	13	16	16		
	2	13	11	18	16		
	3	16	12	17	10		
1	0	27	24	16	16		
	1	35	34	28	29		
	2	34	33	35	32		
	3	39	40	37	34		
2	0	25	26	19	22		
	1	37	30	23	28		
	2	33	33	31	30		
	3	45	47	49	43		
Суми P						$\sum X =$	$\bar{x} =$

Таблиця 7.2

Визначення головних ефектів і взаємодій

Зрошення А	Дози азоту, В				Суми А
	0	1	2	3	
0					
1					
2					
Суми В					$\sum X =$

4. Розраховують суму квадратів для фактору А:

$$C_A = \sum A^2 \div (l_B n) - C \text{ при } l_A - 1 \text{ степенів свободи} \quad (7.2)$$

де С – поправка.

5. Розраховують суму квадратів для фактору В:

$$C_B = \sum B^2 \div (l_A n) - C \text{ при } l_B - 1 \text{ степенів свободи} \quad (7.3)$$

6. Розраховують суму квадратів для взаємодії АВ:

$$C_{AB} = C_V - C_A - C_B \text{ при } (l_A - 1) \cdot (l_B - 1) \text{ степенів свободи} \quad (7.4)$$

7. Складають таблицю дисперсійного аналізу (табл. 7.3). Степені свободи для залишку:



Практична робота № 8. ПРОБІТ – АНАЛІЗ

Мета роботи: *познайомитися зі статистичною обробкою результатів вивчення сили дії пошкоджуючи факторів на біологічні об'єкти методом пробіт-аналізу.*

При вивченні сили дії пошкоджуючих факторів (опромінення, хімічні засоби боротьби зі шкідниками, хворобами і бур'янами) на біологічні об'єкти широко використовується спеціальний статистичний метод – пробіт –аналіз. **Пробіт – це умовні імовірнісні величини** (від англ. probability unit – імовірнісна одиниця).

Наприклад, чутливість певного виду шкідників до інсектицидів або опромінення може характеризуватися дозою, що спричиняє їх повну загибель. Проте вимірювання летальної дози (ЛД) для окремої особини практично неможливе, тому що загибель шкідника від дози, навіть більшої, ніж летальна, настає не відразу, а через декілька днів або навіть тижнів. Якщо ж доза недостатня, щоб спричинити загибель піддослідної особини, це з'ясується також лише через деякий час. За цей період в організмі відбуваються відновні процеси. Відновлення, проте, не буває повним, і тому повторювати дослід на одних і тих самих об'єктах з деякими інтервалами, кожного разу збільшуючи дозу, неможна.

Точно встановити дозу, яка спричиняє 100%-ну летальність, при порівняльній оцінці різних факторів не тільки не просто, але й не виправдано у зв'язку з затратою експериментальних об'єктів та біологічно активних речовин. Практично достатньо встановити дозу, за якої гине 50% особин, яку й приймають за усереднену характеристику летальної дії пошкоджуючого фактора і позначають ЛД₅₀.

Критерій ЛД₅₀ показує, яка доза препарату (або опромінення) необхідна для даної популяції, щоб імовірність загибелі особин склала 50%. Цей критерій визначають статистичним шляхом. Для цього всю піддослідну сукупність поділяють на групи і на кожен незалежну групу, що складається з великої кількості особин, діють препаратом, що вивчається, в різних, усе зростаючих дозах. В першій групі з мінімальною дозою загинуть (або захворіють), напевно, найменш стійкі особини; в другій групі, де препарат застосований в більшій дозі, окрім найменш стійких особин, загине ще й частина більш



стійких і т.д. Таким чином, отримуємо статистично зростаючий ряд, в якому загинуть збільшується із збільшенням дози препарату.

Для визначення LD_{50} будують графік. На осі абсцис відкладають логарифми доз препарату, що вивчається (або дози опромінення), а по осі ординат – процентне вираження загинуть піддослідних особин, трансформоване в пробіти. Значення пробіт, яке відповідає даному відсотку загинуть особин, знаходять за таблицею 8.1.

Таблиця 8.1

Таблиця для перетворення % частоти загинуть особин у пробіти

Загинуть, %	0	1	2	3	4	5	6	7	8	9
0	-	2,67	2,95	3,12	3,25	3,36	3,45	3,52	3,59	3,66
10	3,72	3,77	3,82	3,87	3,92	3,96	4,01	4,05	4,08	4,12
20	4,16	4,19	4,23	4,26	4,29	4,33	4,36	4,39	4,42	4,45
30	4,48	4,50	4,53	4,56	4,59	4,61	4,64	4,67	4,69	4,72
40	4,75	4,77	4,80	4,82	4,85	4,87	4,90	4,92	4,95	4,97
50	5,00	5,03	5,05	5,08	5,10	5,13	5,15	5,18	5,20	5,23
60	5,25	5,28	5,31	5,33	5,36	5,39	5,41	5,44	5,47	5,50
70	5,52	5,55	5,58	5,61	5,64	5,67	5,71	5,74	5,77	5,81
80	5,84	5,88	5,92	5,95	5,99	6,04	6,08	6,13	6,18	6,23
90	6,28	6,34	6,41	6,48	6,55	6,64	6,75	6,88	7,05	7,33

Завдання.

Виходячи з вихідних даних (табл.8.2), обробити дані спостережень, побудувати графік, що характеризує залежність "ефект-доза" і визначити LD_{50} , LD_{55} , LD_{99} .

Таблиця 8.2

Вплив доз дихлорофосу на загинуть рисового довгоносика

Доза,мг/л	1	4	8	10	16	32	46	64
Загинуть,%	16	20	58	72	82	88	93	96

Хід рішення

1. Відсоток загинуть рисового довгоносика трансформувати за табл. 8.1 в мовні одиниці – пробіти і занести до допоміжної табл. 8.3. Також до таблиці занести дози і відсоток загинуть, перетворені в десяткові логарифми.



Допоміжна таблиця

Доза, мг/л	Середня загибель, %	Перетворені значення доз і відсотку загибелі	
		lgD (вісь X)	значення пробіт (вісь Y)

2. За даними допоміжної таблиці побудувати графік, який виражає залежність "ефект-доза". На осі абсцис відкладають логарифми доз, а на осі ординат – значення пробіт.

3. За знайденими точками провести пряму лінію, по якій шляхом інтерполяції визначити ЛД₅₀, ЛД₅₅, ЛД₉₉. Зробити висновки.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Для чого використовується пробіт-аналіз?
2. Назвіть особливості визначення ЛД₅₀.
3. Як за графіком залежності "ефект-доза" визначити значення доз, що відповідають певному відсотку летальності?

ЛІТЕРАТУРА

1. Доспехов Б.А. Методика полевого опыта : учебник. / Б.А. Доспехов. – М.: Агропромиздат, 1985. – 351 с.
2. Глуховцев В.В., Практикум по основам научных исследований в агрономии : учебн. пособие. / В.В. Глуховцев, В.Г. Кириченко, С.Н. Зудилин. – М.: Колос, 2006. – 240 с.
3. Лакин Г.Ф. Биометрия: учебник. / Г.Ф. Лакин. – М.: Высшая школа, 1980. – 292 с.
4. Литтл Т. Сельскохозяйственное опытное дело. Планирование и анализ. Пер. с англ. / Т. Литтл, Ф. Хиллз. – М.: Колос, 1981. – 296 с.
5. Мойсейченко В.Ф. Основи наукових досліджень в агрономії: навч. посіб. / В.Ф. Мойсейченко, В.О. Єщенко. – К.: Вища школа, 2014. – 333 с.
6. Тимошенко І.І., Основи наукових досліджень в агрономії: навч. посіб. / І.І. Тимошенко, З.М. Майшук, Г.О. Касилович. – Львів: ЛДАУ, 2004. – 111 с.



Національний університет
водного господарства
та природокористування



Національний університет
водного господарства
та природокористування