

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
МИКОЛАЇВСЬКИЙ НАЦІОНАЛЬНИЙ АГРАРНИЙ УНІВЕРСИТЕТ

Факультет агротехнологій  
Кафедра виноградарства та плодовоовочівництва

# **БІОМЕТРИЧНА ГЕНЕТИКА СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР**

## **Методичні рекомендації**

до виконання практичних робіт для здобувачів вищої освіти ступеня

“магістр” спеціальності 201 “Агрономія”

денної форми навчання

МИКОЛАЇВ

2020

УДК 631.52

М25

Друкується за рішенням науково-методичної комісії факультету агротехнологій Миколаївського національного аграрного університету від 20.02.2020 року, протокол № 6.

Укладач:

І. М. Марценюк – канд. біол. наук, доцент кафедри виноградарства та плодовоовочівництва, Миколаївський національний аграрний університет.

Рецензенти:

М. М. Корхова – канд. с.-г. наук, доцент кафедри рослинництва та садово-паркового господарства, Миколаївський національний аграрний університет.

Л. Б. Ушакова – вчитель біології вищої категорії, Миколаївська ЗОШ І-ІІІ ступенів №50 ім. Г.Л. Дівіної.

© Миколаївський національний аграрний університет, 2020

## ЗМІСТ

<b>Передмова</b> .....	4
<b>1. МОДУЛЬ I. Біометрико-генетичний аналіз ознак</b>	
1.1. Практична робота № 1. Визначення середніх величин методом малих вибірок. Показники мінливості та типи варіацій кількісних і якісних ознак.....	5
1.2. Практична робота № 2. Визначення середніх величин методом великих вибірок.....	12
1.3. Практична робота № 3. Типи розподілу, графічне зображення варіаційних рядів .....	16
1.4. Практична робота № 4. Методи вивчення взаємозв'язку між ознаками. Розрахунок коефіцієнта кореляції ознак для малих вибірок.....	19
1.5. Практична робота № 5. Проведення дисперсійного аналізу однофакторного рівномірного комплексу.....	22
1.6. Практична робота № 6. Розрахунок коефіцієнта успадкованості та прогнозування селекційного ефекту .....	25
<b>2. МОДУЛЬ II. Біометричні методи у селекції сільськогосподарських культур</b>	
2.1. Практична робота № 7. Аналіз генетичних процесів у популяції. Встановлення генетичної рівноваги та подібності популяцій.....	29
2.2. Практична робота №8. Розв'язання задач біометричної генетики засобами MS EXCEL .....	33
<b>Додатки</b> .....	37
<b>Список рекомендованої літератури</b> .....	40

## ПЕРЕДМОВА

Розвиток і вдосконалення таких сучасних наук, як генетика, селекція рослин та рослинництво, неможливі без використання генетичних параметрів селекції: мінливості, успадкованості, повторюваності, кореляції, регресії тощо. Їх вивченням займається наука біометрія (в перекладі з грецького «*біос*» – життя і «*метрон*» – міра). Сучасна біометрія – це предмет, заснований на застосуванні принципів і методів теорії ймовірності та математичної статистики для вивчення біологічних об'єктів.

Біометрична генетика – це один із напрямів сучасної генетики – науки, що вивчає механізми і закономірності спадковості і мінливості організмів.

Основними завданнями у біометричній генетиці є:

- визначення ступеня фенотипової і генотипової мінливості ознаки;
- виявлення особливостей і типів варіювання кількісних і якісних ознак і визначення характеру розподілу особин з різним рівнем ознаки;
- визначення величини фенотипового і генотипового корелятивного зв'язку між різними ознаками;
- визначення частки впливу різних чинників на фенотипову і генотипову мінливість ознаки з використанням дисперсійного і факторного аналізу.

Програма практичних занять спрямована на закріплення здобувачами вищої освіти теоретичних положень лекційного курсу в процесі вирішення генетичних задач, а також генетичного експерименту з використанням рослинних об'єктів.

Методичні рекомендації – це засіб навчання, створений з метою поліпшення засвоєння, повторення, узагальнення, систематизації та перевірки знань здобувачів вищої освіти із дисципліни. У посібнику представлені різнопланові та різнорівневі завдання різних типів. Розроблені завдання допоможуть придбати навички та вміння у використанні біометричної термінології при виконанні типових завдань і вправ курсу генетики сільськогосподарських рослин.

Методичні рекомендації включає 8 практичних занять, згідно з розробленою робочою програмою дисципліни, які коротко відображають основні її розділи, призначений як для аудиторної, так і для самостійної роботи студентів.

## МОДУЛЬ I. Біометрико-генетичний аналіз ознак

### ПРАКТИЧНА РОБОТА № 1

- Тема.** Визначення середніх величин методом малих вибірок. Показники мінливості та типи варіацій кількісних і якісних ознак.
- Мета:** засвоєння методик розрахунку середніх величин; показників мінливості в селекції рослин. Ознайомлення з різними типами розподілу ознак, біометричною символікою.

*Біометрична генетика* – це один із напрямів сучасної генетики – науки, що вивчає механізми і закономірності спадковості і мінливості організмів.

*Біометрія* (біологічна, варіаційна статистика, від *біо...* та грецького *μετρέω* – вимірюю) – наука про статистичний аналіз групових властивостей біологічних об'єктів. Біометрію визначають також як науку про рух групової інформації в популяціях. Підставою для використання біометрії у генетиці було встановлення факту, що у своїх проявах явища спадковості та мінливості характеризуються також статистичними закономірностями (ймовірністю прояву, законом розподілення варіант за наявності їхньої великої кількості у вибірковій сукупності тощо). Предметом біометрії є група біологічних об'єктів, яка називається *сукупністю*. Розрізняють генеральну сукупність та вибірку з неї.

*Генеральна сукупність* – це весь масив об'єктів однієї категорії, подібних за однаковими ознаками і що різняться за іншими. Наприклад, під генеральною сукупністю сорту озимої м'якої пшениці Миронівська 808 слід розуміти всі рослини, які відносяться до цього сорту. Розмір генеральної сукупності позначається великою літерою *N*.

*Вибірка* – це частина типових представників генеральної сукупності, яка її відображає, тобто, це частина від цілого. Відбір даних до неї за особинами, членами вибірки, проводиться за принципом рендемічності, випадковості. Вибірка складається з окремих *варіантів*, під якими розуміють величину ознаки для одиниці сукупності (величина ознаки – маса, одиниця сукупності – група рослин певного сорту). Варіанти позначають символами  $X_1, X_2 \dots X_n$ . Число варіант у вибірці називають її *обсягом* і позначають символом *n*. Розрізняють вибірки малі (при  $n < 30$ ) і великі ( $n > 30$ ).

*Ознака* – елементарна особливість чи властивість організму, яка може бути визначена і відокремлена від багатьох інших. До варіант можуть бути віднесені кількісні й якісні ознаки. До *кількісних ознак* відносять такі, що можна безпосередньо виміряти або підрахувати (наприклад, масу зерна з колоса або висоту рослини). Ці ознаки характеризуються безперервною мінливістю, їх величини залежать від дії багатьох генів.

*Якісні ознаки* характеризуються окремими описовими визначеннями – забарвлення, здорова чи хвора рослинна, з опушеними листками або без опушення. У випадку, коли існують два взаємовиключаючі варіанти, ознаки називаються *альтернативними* (чорна-біла, гладенька-зморшкувата).

При роботі з кількісними ознаками і при великому числі варіант групування виконують за класами варіаційного ряду.

**Варіаційний (ранжирований) ряд** – це упорядковане розміщення варіант у сукупності відповідно до наростання або спадання їх чисельних значень.

*Середні величини* – це абстрактні, цілі або дробові кількості, які кількісно характеризують ознаки тієї чи іншої сукупності. Існує декілька видів середніх показників. До них належать: середня арифметична, середня зважена, середня геометрична, середня квадратична, середня гармонічна, середня кубічна, мода і медіана.

**Середня арифметична**, позначається буквою  $M$  або  $\bar{X}$  – це одна з основних характеристик вибіркової сукупності, яку найчастіше визначають у біологічних дослідженнях. Вона має кілька властивостей. Головна властивість стверджує, що сума відхилення всіх варіант від середньої арифметичної завжди дорівнює нулю, тобто можна скласти формулу:

$$\sum (X_i - \bar{X}) = 0$$

де:  $\Sigma$  – знак суми,  $X_i$  – індивідуальні значення варіант.

Середня арифметична є величина поіменована і виражається у тих самих одиницях, що й варіанти.

Для вирахування середньої арифметичної для малих вибірок використовують наступну формулу:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum X}{n}$$

Визначення середніх характеристик вибірових сукупностей дає можливість порівнювати між собою окремі групи тварин за господарськи корисними ознаками. Хоча середні величини є важливими статистичними показниками, але вони не дозволяють оцінювати гетерогенність (різноманітність) селекціонованих ознак у популяції. Середні величини нічого не говорять про основну властивість живих організмів – мінливість, без обліку якої неможливо скласти повну характеристику вибіркової сукупності. Ознаки, які досліджуються, при однакових середніх величинах можуть істотно відрізнитися за рівнем варіації. Тому, поряд із середньою величиною ознаки визначаються параметри її мінливості. Одним із показників мінливості ознак є *ліміти*, що з латинської означає – межі, які вказують на мінімальні ( $V_{min}$ ) та максимальні ( $V_{max}$ ) варіанти, між якими знаходяться усі члени даної вибіркової сукупності.

Різниця між лімітами ознаки дозволяє визначити наступний показник мінливості – амплітуду або *розмах варіювання* ( $R = V_{max} - V_{min}$ ). Більший розмах свідчить про вищу мінливість однієї із оцінюваних ознак.

Головними показниками, які характеризують мінливість селекційних ознак, є *дисперсія*, *середньоквадратичне відхилення* та *коефіцієнт варіації*. Ці показники вказують на ступінь різноманітності варіант у вибірковій сукупності.

Варіанти ( $X$ ) різняться між собою за рівнем своєї вираженості, а щоб величину кожної з них можна було б порівняти із середньою всієї вибіркової сукупності ( $\bar{X}$ ) необхідно визначити середню величину їхнього відхилення від середньої всієї вибірки. Цього можна досягти, якщо від кожної варіанти відняти величину середньої арифметичної ( $X - \bar{X}$ ). При цьому варіаційний ряд буде мати як від'ємні, так і позитивні числа, які у своїй сумі  $-\Sigma(X - \bar{X})$  завжди будуть дорівнювати нулю.

Саме з метою уникнення мінусових величин усі відхилення підносяться до квадрату  $\Sigma(X - \bar{X})^2$ . Щоб визначити середню величину відхилень потрібно розділити отриману суму на загальну кількість варіант  $- n$ .

В такому випадку *дисперсія* (варіанса) завжди має квадратичне значення і вираховується за наступною формулою:

$$\sigma^2 = \frac{\Sigma(X - \bar{X})^2}{n}$$

*Середнє квадратичне відхилення* ( $\sigma$ ) для малої вибірки одержують при вирахуванні кореня квадратного із дисперсії ( $\sigma^2$ ):

$$\sigma = \pm \sqrt{\frac{\Sigma(X - \bar{X})^2}{n - 1}}$$

Ця величина ( $\sigma$ ) може мати додатне і від'ємне значення, що свідчить про мінливість ознаки як у бік зменшення варіантів від середньої, так і в бік збільшення. За кількісним значенням сигми у відносному порівнянні до середньої величини вибірки ( $\bar{X}$ ) можна зробити висновок про ступінь мінливості ознаки. Чим більша сигма, тим більша мінливість і навпаки.

Середнє квадратичне відхилення показує на стільки в середньому кожна варіанта відхиляється від своєї середньої арифметичної. У так званих нормальних варіаційних рядах весь розмах мінливості обмежений мінімальним та максимальним значенням варіюючої ознаки і містить у собі шестикратну величину середньоквадратичного відхилення. При цьому максимальна варіанта віддалена від середньої арифметичної на  $+3\sigma$ , а мінімальна на  $-3\sigma$ . Тому весь розмах мінливості прийнято виражати як  $\bar{X} \pm 3\sigma$ . На основі цієї властивості мінливості у нормальних варіаційних рядах цілком можливо передбачати окремі показники без додаткових експериментальних спостережень.

*Коефіцієнт мінливості або варіації*. Серед констант варіативності ознак досить наглядно і більш біологічно обґрунтовано характеризує мінливість коефіцієнт варіації ( $C_v$ ), тобто середнє квадратичне відхилення, виражене у відсотках від середнього показника оцінюваної групи тварин за даною ознакою:

$$C_v = \frac{\sigma}{\bar{X}} \times 100$$

Використання коефіцієнта варіації дозволяє порівняти ступінь мінливості різних за одиницями виміру ознак (маси 1000 зерен у г і вмісту клейковини в зерні у % тощо).

Вважають, що ознаки із слабкою мінливістю характеризуються коефіцієнтом варіації на рівні меншому за 10 %, з середньою – 11–20 і при значній мінливості ознаки –  $C_v$  становить більше 21 %.

Численні дослідження показали, що правильно сформована вибіркова сукупність вірно відображає або репрезентує структуру, стан і властивості генеральної сукупності. Існує загальна система добору об'єктів у вибірку сукупність. Головним для цієї системи є правило відповідної ймовірності вибору якого завгодно об'єкта із генеральної сукупності.

**Репрезентативність** – головна властивість вибіркової сукупності характеризувати генеральну сукупність з відповідною точністю та достатньою надійністю. Репрезентативністю позначають і ступінь відповідності вибіркового показника генеральним параметрам. Репрезентативність з'ясовується лише тоді, коли необхідно охарактеризувати усю велику сукупність особин, що цікавить дослідника, на основі вивчення лише якоїсь її частини і коли уже відібрана відповідна вибірка. При цьому можливі певні відхилення вибіркового показника від параметрів генеральної сукупності. Такі відхилення в біометрії називаються *похибками репрезентативності*.

**Похибка середньої арифметичної.** Вибірка, мала чи велика, є лише часткою генеральної сукупності, тому вона не може з абсолютною точністю відобразити її параметри та властивості, тому усі статистичні показники будуть у тій чи іншій мірі відрізнятися від параметрів генеральної сукупності.

Стандартну статистичну похибку позначають буквою  $s$  або  $m$ , а під нею, як нижній індекс, записують знак того біометричного параметру, до якого вона належить. **Похибка середньої арифметичної** для малої вибірки визначається за формулою:

$$s_{\bar{X}} = \frac{\sigma}{\sqrt{n-1}}$$

Похибка вибірки виражається у тих самих одиницях, що й оцінювані показники, які вона характеризує. Похибка має два об'єднані знаки – плюс і мінус ( $\pm$ ), які вказують на те, що відхилення вибіркового показника можуть відбуватися як у бік більших, так і в бік менших значень по відношенню до параметра генеральної сукупності. Показник середньої арифметичної разом зі своєю похибкою записується таким чином:  $\bar{X} \pm s_{\bar{X}}$

**Похибка середнього квадратичного відхилення** обчислюється за наступною формулою:

$$s_{\sigma} = \frac{\sigma}{\sqrt{2n}}$$



**Похибка коефіцієнта варіації** вираховується за формулою:

$$s_{C_v} = \frac{C_v}{\sqrt{2n}}$$

В усіх випадках спостережень, експериментальних досліджень параметри генеральної сукупності невідомі, тому їх характеризують на основі аналізу вибірових сукупностей. Оскільки показники вибірки випадкові, тому й отримані результати вважаються гіпотетичними, або з певною часткою припущення. Для оцінки величини параметрів генеральної сукупності за даними вибірки у біології існує **нульова гіпотеза**. Вона ґрунтується на припущенні, що між показниками двох вибірових сукупностей достовірної різниці не існує, тобто припускається, що оцінювані групи складають однорідний матеріал, тобто, одну сукупність.

Виходячи із цього, математичний аналіз повинен підтвердити існування нульової гіпотези, або, навпаки, відхилити її та довести, що між показниками дійсно існує достовірна різниця. Якщо в процесі порівняльного аналізу результатів досліджень біометричні методи доводять достовірність різниці на користь експериментального варіанту, в такому разі розробку можна впроваджувати у виробництво. У разі, коли достовірність різниці не доведена, вона вважається випадковою і в дії залишається нульова гіпотеза, тому ніяких впроваджень результатів досліджень робити не можна.

Відхилення нульової гіпотези повинно бути обґрунтоване відповідним рівнем значущості. У біологічних дослідженнях приймають щонайменше 5%-й рівень, якому відповідає ймовірність  $P=0,95$ . У більш відповідальних дослідженнях, коли висновки повинні бути більш точними і надійними приймається величина значущості в 1,0% або 0,01%, яким відповідає рівень ймовірності – 0,99 та 0,999.

Найчастіше у науковій практиці використовується три пороги достовірної (довірчої) ймовірності, кожний із яких пов'язаний з відповідною величиною нормованого відхилення:

$P = 0,95$  відповідає –  $x = 1,96 \sigma$ ;

$P = 0,99$  відповідає –  $x = 2,58 \sigma$ ;

$P = 0,999$  відповідає –  $x = 3,29 \sigma$ .

Таким чином, завдяки рівню ймовірності, одержані характеристики вибірки переносяться на аналогічні показники генеральної сукупності з відповідним порогом достовірності. Наприклад ймовірність на рівні 0,95 ( $P \geq 0,95$ , більше або дорівнює) означає, що близькі величини до одержаних у досліді результативних показників повинні повторитися в 95 вибірках із 100.

**Приклади визначення біометричних показників у генетиці рослин.**

**1. Розрахунок  $M$ ,  $\sigma$ ,  $C_v$ ,  $s_{\bar{x}}$ ,  $s_{\sigma}$ ,  $s_{C_v}$ ,  $t_{\bar{x}}$  методом малих вибірок.**

Для обчислення складається простий варіаційний ряд, наприклад, рендомізовано виписуються величини довжини (у см) 20 колосів озимої пшениці (n=20), і дані (X) розміщують у порядку зростання: 7,2; 7,9; 8,4; 8,8; 9,0; 9,0; 9,3; 9,4; 9,7; 9,9; 10,1; 10,2; 10,2; 10,3; 10,4; 10,8; 10,9; 11,7; 12,7; 13,1 (таблиця 1).

Таблиця 1

**Допоміжна таблиця для обробки варіаційного ряду**

№ рослини	Довжина колоса (X), см	Відхилення від середньої арифметичної (X - $\bar{X}$ )	Квадрати відхилень (X - $\bar{X}$ ) <sup>2</sup>
1.	7,2	-2,75	7,56
2.	7,9	-2,05	4,20
3.	8,4	-1,55	2,40
4.	8,8	-1,15	1,32
5.	9,0	-0,95	0,90
6.	9,0	-0,95	0,90
7.	9,3	-0,65	0,42
8.	9,4	-0,55	0,30
9.	9,7	-0,25	0,06
10.	9,9	-0,05	0,00
11.	10,1	0,15	0,02
12.	10,2	0,25	0,06
13.	10,2	0,25	0,06
14.	10,3	0,35	0,12
15.	10,4	0,45	0,20
16.	10,8	0,85	0,72
17.	10,9	0,95	0,90
18.	11,7	1,75	3,06
19.	12,7	2,75	7,56
20.	13,1	3,15	9,92
n = 20	$\bar{X} = 9,95$	0	$\sum (X - \bar{X})^2 = 40,73$

Потім складають варіанти і ділять на кількість (n) варіант, тобто

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

$$\bar{X} = \frac{7,2 + 7,9 + \dots + 13,1}{20} = 9,92$$

У даному випадку середня арифметична  $\bar{X} = 9,92$  см показує, якою би була довжина колоса озимої пшениці, якби не було мінливості.

Як відомо, усі біологічні ознаки змінюються. Тому потрібно обчислити середнє квадратичне відхилення  $\sigma$ , яке вказує на ступінь мінливості, тобто на яку величину в середньому кожна варіанта може відхилятися від середньої арифметичної у бік збільшення або зменшення. Для цього спочатку записують варіанти X, потім відхилення кожної

варіанти від середньої арифметичної  $(X-\bar{X})$ , підносять ці відхилення до квадрату  $(X-\bar{X})^2$ , обчислюють суму квадратів відхилень  $\Sigma(X-\bar{X})^2$  і за формулою обчислюють  $\sigma$ , яка виражається в натуральних величинах:

$$\sigma = \pm \sqrt{\frac{\Sigma(X - \bar{X})^2}{n - 1}}$$

Визначаємо середнє квадратичне відхилення  $\sigma$ :

$$\sigma = \pm \sqrt{\frac{40,73}{19}} = \pm \sqrt{2,14} = \pm 1,46 \text{ см}$$

Отже, кожна варіанта, тобто довжина колоса кожної рослини, в середньому на 1,46 см відхиляється від середньої арифметичної ( $\bar{X}=9,95$  см) у бік збільшення та зменшення. Це вказує на різномірність рослин озимої пшениці даної генеральної сукупності у генетичному аспекті.

### Контрольні запитання:

1. Що таке сукупність? Наведіть приклад різних сукупностей.
2. Що таке варіанти? Випадкова змінна?
3. Якими можуть бути відмінності між окремими значеннями (варіантами) випадкової змінної?
4. Що таке варіаційний ряд? У чому полягають особливості розподілу варіант у варіаційному ряду.
5. Які властивості середньої арифметичної ви знаєте?
6. У чому полягає непрямий спосіб обчислення  $\bar{x}$ ?
7. Що таке варіанса?
8. Яка різниця між  $\sigma$  та  $C_v$ ? В яких випадках важливо використання  $C_v$ ?
9. Чому  $\bar{x}$  та  $\sigma$  є основними характеристиками варіаційного ряду?
10. Чи можна Г. Менделя, засновника науки генетики, вважати раннім вченим-біометриком?

## ПРАКТИЧНА РОБОТА № 2

**Тема.** Визначення середніх величин методом великих вибірок.  
**Мета:** засвоєння методики розрахунку середніх величин показників мінливості для великих вибірок.

Якщо маємо справу з великою кількістю рослин ( $n > 30$ ), тобто з великою вибіркою, обробку цифрових даних проводять не прямим способом, а за допомогою складніших специфічних методів вирахування середньої арифметичної. Одним із таких найчастіше вживаних є *метод умовної середньої*.

В основі цього методу закладені математичні операції, у яких за основу взято модальний клас та відхилення від нього. При цьому модальний клас використовується як умовна середня. Остання здебільшого добирається за найбільшою частотою варіант. У тих випадках, коли модальних класів декілька або він знаходиться не посередині варіаційного ряду, за умовну середню добирають той клас, котрий розділяє варіаційний ряд на більш-менш рівні частини за своїми частотами. На конкретному прикладі розглянемо порядок визначення середньої арифметичної.

**Приклад виконання типового завдання.** Визначити середню величину загальної довжини 100 рослин льону за наступними даними (см):

90	109	99	100	115	68	70	72	73	70
76	82	80	68	69	74	72	69	80	79
79	84	84	108	83	84	99	98	102	101
45	59	60	63	78	87	94	91	88	90
72	68	80	81	84	77	79	81	84	76
70	67	100	103	69	72	74	66	67	72
79	78	83	92	93	81	82	86	89	93
77	76	88	89	94	82	80	81	77	80
92	91	76	79	73	84	79	84	79	84
89	85	93	90	79	83	91	87	89	94

Для складання й опрацювання такого варіаційного ряду будемо класи згідно з таблицею 2.

**Допоміжна таблиця для обробки варіаційного ряду**

Класи ( $w_0 + w_t$ ), см,	Частота варіант, $f$	Умовне відхилення від модального класу ( $a$ )	Добуток, $fa$	Добуток, $fa^2$
45-54	1	-3	-3	9
55-64	3	-2	-6	12
65-74	21	-1	-21	21
<b>75-84</b>	<b>40</b>	<b>0</b>	<b>0</b>	<b>0</b>
85-94	24	+1	24	24
95-104	8	+2	16	32

105-115	3	+3	9	27
k=10	$\Sigma f=100$		$\Sigma fa=19$	$\Sigma (fa)^2=125$

1. Визначаємо ліміти – найбільше і найменше значення даної ознаки:  
найменше значення – 45 см;  
найбільше значення – 115 см.

2. Визначаємо розмах варіювання: від максимального значення у вибірці

віднімаємо мінімальне:  $115-45 = 70$  кг.

3. Визначаємо кількість класів для побудови майбутнього варіаційного ряду. Класи – це градації, в які згруповані близькі за значенням варіанти. Кількість класів у кожному випадку береться довільно, але в межах від 5 до 15.

У даному разі взято 7 класів.

4. Визначаємо класовий проміжок (k) діленням величини розмаху варіювання на кількість взятих класів.

$$k = \frac{70}{7} = 10 \text{ см}$$

5. Додаємо класовий проміжок до найменшого ліміту ( $45+10=55$ ), а потім до нових значень додають заплановану кількість класів і щоб в останній увійшла максимальна варіанта.

6. Знаходимо частоти  $f$ , тобто кількість рослин, яких залежно від довжини розміщують у відповідні класи.

7. Визначаємо умовний середній клас A. Виділяємо його жирними лініями, позначаємо нулем. Величина умовної середньої розраховується за формулою:

$$A = \frac{w_0 + w_t}{2}$$

$$A = \frac{75 + 84}{2} = 79,5 \text{ см}$$

8. Знаходимо умовне відхилення (a) від модального класу, що позначили нулем, усіх інших класів варіаційного ряду у цілих числах. Відхилення, що мають менше значення за умовну середню величину будуть із знаком мінус з відхиленням у бік зменшення на три класових проміжки, а ті що мають більше значення – будуть із знаком плюс з відхилення у бік зростання класів.

9. Знаходимо добуток частот на відхилення  $fa$  та їх суму:  $\Sigma fa = 19$

10. Визначаємо добуток частот на квадрат відхилень  $fa^2$  та їх суму:  
 $\Sigma fa^2=125$

На цьому обробка варіаційного ряду закінчується. Підставляючи в формули дані з варіаційного ряду розрахуємо необхідні показники.

а) Вираховуємо середнє арифметичне значення за формулою:

$$\bar{X} = A + kb$$

де: A – середнє значення умовного середнього класу;

$k$  – класовий проміжок;

$b$  – поправка, яка визначається за формулою:

$$b = \frac{\sum fa}{n}$$
$$b = \frac{19}{100} = 0,19$$

Звідси:  $\bar{X} = 79,5 + 7 \times 0,19 = 79,5 + 1,33 = 80,83$  см.

б) вираховуємо середнє квадратичне відхилення:

$$\sigma = \pm k \sqrt{\frac{\sum (fa)^2}{n}} \pm 10 \sqrt{\frac{125}{100}} = \pm 10 \sqrt{1,25} = \pm 11,2 \text{ см}$$

в) коефіцієнт варіації:

$$C_v = \frac{\sigma}{\bar{X}} \times 100 = \frac{11,2}{80,3} \times 100 = 13,9 \%$$

г) похибки середньої арифметичної:

$$s_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{11,2}{\sqrt{100}} = 1,12 \text{ см}$$

д) похибку середньоквадратичного відхилення:

$$s_{\sigma} = \frac{\sigma}{\sqrt{2n}} = \frac{11,2}{\sqrt{2 \times 100}} = 0,79 \text{ см}$$

ж) похибки коефіцієнта варіації:

$$s_{C_v} = \frac{C_v}{\sqrt{2n}} = \frac{13,9}{\sqrt{2 \times 100}} = 0,98 \text{ см}$$

з) критерій достовірності середнього арифметичного значення:

$$t_{s_{\bar{X}}} = \frac{\bar{X}}{s_{\bar{X}}} = \frac{80,83}{1,12} = 72,2$$

11. Остаточною є така форма запису статистичних даних:

$$\bar{X} \pm s_{\bar{X}} = 80,83 \pm 1,12$$

$$\sigma \pm s_{\sigma} = 11,2 \pm 0,79$$

$$C_v \pm s_{C_v} = 13,9 \pm 0,98$$

12. Знаходимо ймовірність середньоарифметичного значення. При великій вибірці ступені свободи, практично, не мають значення, бо

розподіл Стьюдента наближається до нормального. Тому кількість ступенів свободи буде  $v=n-1=100$ . Тому отримані фактично значення порівнюємо зі стандартними значеннями, які в таблиці (додаток 1) розміщені у рядку навпроти числа 100 ( $\infty$ ).

Число ступенів свободи	Вірогідність (P)		
	0,95	0,99	0,999
$\infty$	1,96	2,58	3,29

Якщо порівняти фактичний критерій вірогідності ( $t_{s\bar{x}} = 72,2$ ) із найвищим стандартним значенням  $t = 3,29$ , що знаходяться у вірогідних межах  $P = 0,999$ , помітно, що він є істотно вищим. Це означає, що *середня арифметична вибірки вірогідно відображає рівень середньої арифметичної загальної довжини рослин льону генеральної сукупності*.

### Контрольні запитання:

1. Що таке "мала" та "велика" вибірки?
2. Які ви знаєте способи обрахунку середньої арифметичної?
3. Що таке помилки репрезентативності даних? Як вони виникають?
4. Як здійснюється оцінка достовірності експериментальних даних?

## ПРАКТИЧНА РОБОТА № 3

- Тема.** Типи розподілу, графічне зображення варіаційних рядів
- Мета:** засвоєння методики побудови різних типів графічного відображення мінливості ознак у культурних рослин.

Більшості кількісних ознак сільськогосподарських рослин властивий нормальний тип розподілу. При ньому найбільше число варіант розміщено в центрі близько середніх значень ознаки. Чим більше відхиляються значення окремих варіант, тим рідше вони зустрічаються, тобто ймовірність повторюваності варіанти зменшується при віддаленні її від середніх величин. Тому в конкретній популяції особин з середніми значеннями продуктивності завжди більше, ніж з мінімальними й максимальними. Нормальний розподіл повністю характеризується середньою величиною і стандартним відхиленням.

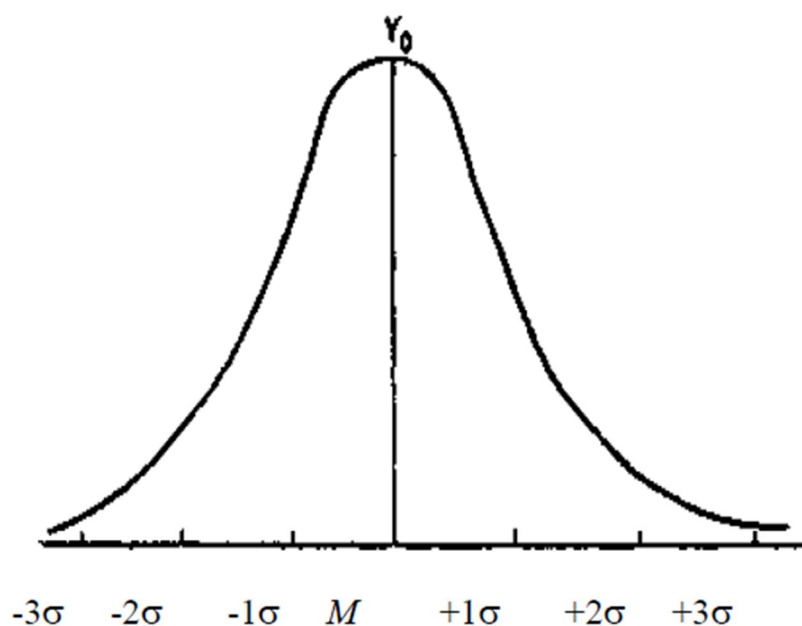
Існує правило трьох сигм або нормального розподілу ознак, яке ґрунтується на кривій Гауса (німецького вченого), який разом з Лапласом (французьким вченим) відкрили це правило.

Для нормального розподілу при  $n \rightarrow \infty$  характерним є наступне:

- теоретична крива має дзвіноподібний симетричний вигляд;
- фундаментом кривої служать класи варіюючої ознаки ( $x$ );
- гілки дзвіноподібної кривої утворені перпендикулярами (ординатами), відновленими з точок значень  $x$ ;
- кінці гілок не зливаються з віссю абсцис, а наближаються до неї в нескінченності (асимптотично).
- вершина нормальної кривої ( $y_0$ ) визначається перпендикуляром, спрямованим з точки  $x$ , яка відповідає величині середній арифметичній даної ознаки.
- $y_0$  відповідає найбільшій частоті зустрічі ( $p$ ) тих особин популяції, у яких величина ознаки дорівнює середній арифметичній;
- у нормальному розподілі точка  $x$  співпадає з величиною моди ( $M_0$ ) і медіани ( $Me$ );
- чим більша мінливість ознаки та відповідно значення сигми, тим більшою буде основа у кривої і тим нижче буде розміщуватися вершина.

Вся різноманітність ознаки як правило входить у розмах  $\bar{X} \pm 3\sigma$ , тобто в інтервалі від самого мінімального до самого максимального значення входить шість сигм. При цьому в межах  $\bar{X} \pm 1\sigma$  знаходиться 68,3% особин, у межах  $\bar{X} \pm 2\sigma$  – 95,5 та  $\bar{X} \pm 3\sigma$  – 99,7% (правило трьох сигм).





**Рис. 1. Крива нормального розподілу частот (Гауса–Лапласа)**

Отже, будь-який варіаційний ряд в загальному вигляді описується графічно варіаційною кривою, яка має не лише ілюстративне значення але і використовується для статистичного аналізу варіюючих ознак. При побудові графіків варіаційних рядів масштаби на осях координат вибираються довільно, але з таким розрахунком, щоб висота варіаційної кривої (тобто максимальна ордината) відносилася до її основи приблизно, як 5:8. Відкладаючи по осі абсцис класи варіаційної лави, треба доводити їх зліва і справа до нульових класів, які не містять жодної варіанти. Недотримання цих правил приводить до небажаних результатів: графік виходить із загостреною вершиною і вузькою основою або у вигляді надмірно розтягнутої по ширині фігури сплющення. У обох випадках варіаційна крива виявляється такою, що недостатньо чітко відображує характерні риси варіювання ознаки, яка вивчається.

Окрім варіаційної кривої емпіричний ряд розподілу можна представити у вигляді *кумуляти* та *огіви*. Кумуляту отримують якщо по осі абсцис відкладати значення класів, а по осі ординат – накопичення частоти з подальшим з'єднанням геометричних точок прямими лініями.

На відміну від куполоподібної варіаційної кривої кумулята має S-образну форму. Накопичені частоти отримують послідовним підсумовуванням або кумуляцією (від лат. *cumulato* – збільшення, накопичення) частот в напрямку від мінімальної класової варіанти або від першого класу до кінця варіаційного ряду.



**Рис. 2. Приклад графіка кумуляти**

Якщо накопичені частоти відкласти по осі абсцис, а значення класів – по осі ординат з подальшим з'єднанням геометричних точок прямими лініями, вийде графік, що називається огівом (рис. 1.4). Огіва дозволяє порівнювати один з одним одночасно декілька емпіричних розподілів неоднакового розміру.



**Рис. 3. Приклад графіка огіви**

### Контрольні запитання:

1. Наведіть формулу і покажіть графічне зображення функції нормального розподілу.
2. Наведіть формулу, покажіть графічне зображення і охарактеризуйте щільність нормального розподілу.
3. Сформулюйте і проілюструйте «правило трьох сигм».
4. Яка частина варіант нормального розподілу знаходяться інтервалах  $\bar{x} \pm \sigma$ ,  $\bar{x} \pm 2\sigma$ ,  $\bar{x} \pm 3\sigma$ ?
5. Чим відрізняється графік кумуляти від графіку огіви?

## ПРАКТИЧНА РОБОТА № 4

- Тема.** Методи вивчення взаємозв'язку між ознаками. Розрахунок коефіцієнта кореляції ознак для малих вибірок.
- Мета:** Оволодіти методами обчислення показників зв'язку між ознаками. Навчитись обчислювати коефіцієнт кореляції ( $r$ ) методом малих вибірок

Із метою підвищення ефективності селекції одночасно за кількома ознаками рекомендується насамперед враховувати їхню взаємну зумовленість, тобто кореляцію.

Для цього виникла необхідність математичного усвідомлення такого явища та вираження його числовим показником, коефіцієнтом, за величиною якого можна було б говорити про тісноту чи силу зв'язку між окремими ознаками.

*Коефіцієнт кореляції* - числовий показник простої лінійної кореляції, який описує напрям і тісноту зв'язку між досліджуваними величинами, вимірює зв'язок лише при лінійній формі залежності, а його абсолютне значення знаходиться в межах від  $-1$  до  $+1$ . При значенні  $r = 0$  – зв'язок відсутній; при  $+1$  – пряма кореляційна залежність; а при  $-1$  – зворотня.

Прийнято вважати, що залежність низька, якщо  $r=0,1-0,3$ ; середня при  $r=0,4-0,6$  і висока – при  $0,7$  і більше. При коефіцієнті кореляції  $r=0,1$  мінливість другої ознаки лише на  $1\%$  залежить від мінливості першої, а в  $99\%$  випадках – від випадкових факторів.

Розрізняють *фенотипову*, *генотипову* та *середовищну* кореляції.

**Фенотипова кореляція** дає змогу встановити зв'язок між різними господарськи корисними ознаками і використати їх з метою селекції.

**Генотипова кореляція** – це форма зв'язку між двома ознаками, зумовлена адитивною дією і взаємодією генів (плейотропія, епістаз тощо). Генотипова кореляція визначає вплив спадкових факторів на ступінь фенотипової кореляції.

**Паратипова (середовищна) кореляція** зумовлена силою і спрямованістю впливу умов середовища на дві ознаки, що вивчаються.

**Розрахунок коефіцієнта кореляції для малих вибірок.**

**Приклад.** Визначено ураженість льону фузаріозом (ряд  $y$ ) у залежності від інтервалу між посівами (у роках) на одному і тому ж полі сприйнятливих до грибних патогенів (фузаріозу) сортів льону (ряд  $x$ ). Провести кореляційний аналіз даних.

Для цього використовуємо допоміжну таблицю 3, шаблон якої поданий нижче.

**Форма запису для розрахунку коефіцієнта кореляції між інтервалом посівів (x) льону на одному і тому ж полі та ураженістю фузаріозом (y)**

Номери пар	Значення ознак		x <sup>2</sup>	y <sup>2</sup>	xy
	x, роки	y, %			
1.	1	88	1	7744	88
2.	2	76	4	5776	152
3.	2	70	4	4900	140
4.	7	5	49	25	35
5.	6	12	36	144	72
6.	5	28	25	784	140
7.	3	45	9	2025	135
8.	4	45	16	2025	180
9.	6	9	36	81	54
10.	3	62	9	3844	186
n=10	∑ x = 39	∑ y = 440	∑ x <sup>2</sup> = 189	∑ y <sup>2</sup> = 27348	∑ xy = 1182

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{C_x \cdot C_y}};$$

$$\text{де: } C_x = \sum x^2 - \frac{(\sum x)^2}{n} = 189 - \frac{(39)^2}{10} = 36,9$$

$$C_y = \sum y^2 - \frac{(\sum y)^2}{n} = 27348 - \frac{(440)^2}{10} = 7988$$

$$r = \frac{1182 - \frac{39 \cdot 440}{10}}{\sqrt{36,9 \cdot 7988}} = \frac{-534}{543} = -0,98$$

Розраховуємо помилку коефіцієнта кореляції ( $s_r$ ), критерій значущості ( $t_r$ ), довірчий інтервал та перевіряємо нульову гіпотезу ( $H_0$ ) про відсутність кореляційного зв'язку між явищами.

$$s_r = \sqrt{\frac{1 - r^2}{n - 2}} = \sqrt{\frac{1 - 0,98^2}{10 - 2}} = 0,07$$

$$t_r = \frac{r}{s_r} = \frac{0,98}{0,07} = 14,00$$

$$v = n - 2 = 10 - 2 = 8 \quad t_{05} = 2,31$$

$$r \pm t_{05} s_r = -0,98 \pm 2,31 \times 0,07 = -0,98 \pm 0,16$$

$$r \pm t_{05} s_r = (-1,14 \div 0,82)$$

За критерієм  $t_r$  ( $t_{\text{факт}} > t_{05}$ ) і довірчим інтервалом, який не включає нульового значення, кореляція суттєва, отже, нульова гіпотеза на 5%-му рівні відкидається.

### **Контрольні запитання:**

1. Що таке кореляція?
2. Яка різниця між позитивною і негативною кореляціями
3. Які можливі значення коефіцієнта кореляції?
4. Які значення коефіцієнта кореляції слід вважати високими, середніми і чому?
5. Чому дорівнює коефіцієнт кореляції при повному кореляційному зв'язку?
6. Як треба розуміти нульову гіпотезу стосовно коефіцієнту кореляції?

## ПРАКТИЧНА РОБОТА № 5

- Тема.** Проведення дисперсійного аналізу однофакторного рівномірного комплексу.
- Мета:** отримати навички вивчення статистичного впливу одного або декількох факторів на результативну ознаку.

У випадку, коли число вибірок в дослідженнях більше двох, метод попарних порівнянь вимагає великої обчислювальної роботи. У цьому випадку використовують метод *дисперсійного аналізу*, заснований на розкладанні загальної дисперсії статистичного комплексу на її складові компоненти.

При утворенні дисперсійних комплексів необхідно дотримуватися двох умов:

- 1) діючі на ознаку регульовані фактори повинні бути незалежні один від одного (багатофакторні комплекси);
- 2) вибірки, що групуються в статистичний комплекс, повинні здійснюватися за принципом рендомізації.

Дисперсійний аналіз використовує властивість адитивності дисперсії випадкової величини, що обумовлено дією незалежних факторів. В залежності від числа джерел дисперсії розрізняють *однофакторний* та *багатофакторний* дисперсійний аналіз.

*Однофакторний* дисперсійний аналіз дозволяє виявити вплив тільки одного фактора. При цьому загальне варіювання результативної ознаки розкладається на два компоненти – варіювання варіантів та випадкове варіювання:  $C_y = C_v + C_z$

Статистичний аналіз даних однофакторного експерименту проводять у три етапи:

**1 етап.** Складають розрахункову таблицю, розміщуючи у ній вихідні дані за рядками та стовпчиками, визначають суми та середні за варіантами, загальну суму та середнє значення результативної ознаки в експерименті:

Таблиця 4

Варіанти	Вихідні дані, X	Число спостережень, n	Суми за варіантами, V	Середні за варіантами, $\bar{x}$
1	$X_1, X_2, X_3, \dots, X_{1n}$	$n_1$	$V_1$	$\bar{x}_1$
2	$X_{11}, X_{12}, X_{13}, \dots, X_{2n}$	$n_2$	$V_2$	$\bar{x}_2$
3	$X_{21}, X_{22}, X_{23}, \dots, X_{3n}$	$n_3$	$V_3$	$\bar{x}_3$
...	...	...	...	...
l	$X_{l1}, X_{l2}, X_{l3}, \dots, X_{ln}$	$n_l$	$V_l$	$\bar{x}_l$
Загальна сума		$N = \sum n$	$\sum X = \sum V$	$\bar{x} = \sum X / N$

**2 етап.** Розраховують суми квадратів відхилень та F-критерій за такими формулами та заносять до таблиці:

Таблиця 5

Дисперсія	Сума квадратів	Ступені свободи	Середній квадрат	$F_{\text{факт.}}$	$F_{\text{теор.}}$
Загальна $C_y$	$\sum X^2 - C$	$N-1$	-	-	-
Варіантів $C_v$	$\sum \frac{v^2}{n} - C$	$l-1$	$s_v^2 = \frac{C_v}{l-1}$	$\frac{s_v^2}{s^2}$	за таблицею
Остаток $C_z$	$C_y - C_v$	$N-l$	$s^2 = \frac{C_z}{N-l}$	-	-

Коригуючий фактор (С) розраховують за формулою:

$$C = (\sum X)^2 / N$$

**3 етап.** Визначають помилку досліду та істотність окремих різниць.

помилка середньої:  $s_x = \sqrt{\frac{s^2}{n}}$ ;

помилка різниці середніх:  $s_d = \sqrt{\frac{2s^2}{n}}$ ;

$НІР_{05} = t_{05} \times s_d$

**4 етап.** Результати експерименту та статистичної обробки даних записують до підсумкової таблиці.

**Приклад.** Встановлення селекційної цінності п'яти сортів за врожайністю зерна.

1. Складаємо розрахункову таблицю:

Таблиця 6

Варіанти	Врожайність зерна, г/ділянку				Число спостережень, $n$	Суми за варіантами, $\sum$	Середні за варіантами, $\bar{x}$
1(st)	454	470	430	500	4	1854	463,5
2	502	550	490	507	4	2049	512,2
3	601	670	550	607	4	2428	607,0
4	407	412	475	402	4	1696	424,0
5	418	470	460	412	4	1760	440,0
<b>Загальна сума</b>					<b><math>N = \sum n = 20</math></b>	<b><math>\sum X = 9787</math></b>	<b><math>\bar{x} = 489,4</math></b>

2. Розраховуємо суми квадратів відхилень та F-критерій.

$$C = \frac{(\sum X)^2}{N} = \frac{(9787)^2}{20} = 4789268$$

$$\sum X^2 - C = (454^2 + 470^2 + \dots + 412^2) - 4789268 = 104941$$

$$\sum \frac{v^2}{n} - C = \frac{(1854^2 + 2049^2 + \dots + 1760^2)}{4} - 4789268 = 86961$$

$$C_y - C_v = 104941 - 86961 = 17980$$

$$F_{\text{факт.}} = \frac{21740}{1199} = 18,13$$

Одержані результати заносимо до таблиці дисперсійного аналізу.

Таблиця 7

Дисперсія	Сума квадратів	Ступені свободи	Середній квадрат	$F_{\text{факт.}}$	$F_{\text{теор.}}$
Загальна $C_y$	104941	19	-	-	-
Варіантів $C_v$	86961	4	21740	18,13	3,06
Залишок $C_z$	17980	15	1199	-	-

Здійснюємо оцінку нульової гіпотези. Оскільки  $F_{\text{факт.}} > F_{\text{теор.}}$ , то нульова гіпотеза відкидається, тобто у досліді є істотні відмінності за варіантами на 5%-му рівні значущості ( $t_{05}$ ).

3. Встановлюємо помилку дослідів та істотність різниць:

$$\text{помилка середньої: } s_x = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{1199}{4}} = 17,3 \text{ г};$$

$$\text{помилка різниці середніх: } s_d = \sqrt{\frac{2s^2}{n}} = \sqrt{\frac{2 \times 1199}{4}} = 24,5 \text{ г};$$

$$\text{НІР}_{05} = t_{05} \times s_d = 2,13 \times 24,5 = 52,2 \text{ г},$$

$$\text{або НІР у відсотках: } \text{НСР}_{05} = \frac{t_{05} \times s_d}{\bar{x}} \times 100 = \frac{52,2}{489,4} \times 100 = 10,7\%$$

4. Результати експерименту та статистичної обробки даних записуємо до підсумкової таблиці.

Таблиця 8

Сорт	Врожайність зерна, г/ділянку	Різниця зі стандартом		Група за врожайністю зерна
		г/ділянку	%	
Саратівська 55 (st)	463,5	-	-	1 (st)
Саратівська 29	424,0	-39,5	-8,5	1
Саратівська 46	440,0	-23,5	-5,1	1
Саратівська 58	512,2	48,7	10,5	1
К 503	607,0	143,5	30,9	2

**Висновок:** у сортовипробуванні встановлені істотні різниці за врожайністю зерна між сортами м'якої пшениці (варіантами). Істотно перевищив за врожайністю зерна сорт-стандарт лише сорт Л 503.

### Контрольні запитання:

1. Що вивчають методом дисперсійного аналізу?
2. Що таке дисперсія, та на які види вона розподіляється?
3. Що таке дисперсійний комплекс, і які види комплексів бувають?
4. Що таке градація, і які комплекси бувають у залежності від розмірів градацій?
5. За якою формулою визначається показник сили впливу фактора, дія якого вивчається?
6. Як визначити достовірність розрахованого показника сили впливу фактора?



## ПРАКТИЧНА РОБОТА № 6

**Тема.** Розрахунок коефіцієнта успадкованості та прогнозування селекційного ефекту

**Мета:** з'ясування значення типів мінливості при доборі рослин, опанування методами визначення коефіцієнта успадкованості та здійснення генетичного прогнозу селекційного поліпшення ознак.

*Успадкованість* – це статистичний термін, який застосовують для позначення частки загальної фенотипічної мінливості, зумовленої генотипічними факторами.

Мірою вираження генотипічної детермінації мінливості ознаки є *коефіцієнт успадкованості ( $h^2$ )*.

$$h^2 = \frac{s_G^2}{s_P^2}, \text{ де}$$

$s_P^2$  – фенотипічна мінливість,

$s_G^2$  – генотипічна.

Величина  $h^2$  не може бути більше 1,0 і менше 0, тобто від'ємною. Чим більша величина  $h^2$ , тим більше мінливість ознаки зумовлена генотипом і, тим менше – факторами середовища.

Коефіцієнт успадкованості ( $h^2$ ) є показником для прогнозування ефективності селекції за фенотипічними параметрами ознаки. При  $h^2$  більше 0,3 є можливість поліпшення ознаки, а найбільш ефективною селекція буде при  $h^2=0,7$  і вище. При значенні  $h^2<0,3$  селекція за даною ознакою неефективна.

Відомо кілька способів розрахунку коефіцієнта успадкованості:

1) за допомогою коефіцієнта кореляції або коефіцієнта регресії між фенотипами споріднених груп (приклад 1);

2) за допомогою дисперсійного аналізу (приклад 2).

**Приклад 1.** Визначити ефективність добору рослин кукурудзи з високим умістом олії у зерні (таблиця 9).

Таблиця 9

Номери пар	Значення ознак		$x^2$	$y^2$	$xy$
	$x$ , материнська форма	$y$ , гібрид			
1	6,99	9,02	48,86	81,36	63,05
2	6,99	6,94	48,86	48,16	48,51
3	7,84	8,13	61,47	66,10	63,74
...	...	...	...	...	...
51	6,57	6,42	43,17	41,22	42,16
n=51	$\sum x = 352,22$	$\sum y = 396,41$	$\sum x^2 = 2502,52$	$\sum y^2 = 3134,14$	$\sum xy = 2762,33$

*Розв'язання.* 1. За вихідними даними розраховують допоміжні величини ( $\sum x, \sum y, \sum x^2, \sum y^2$  та  $\sum xy$ ), коефіцієнт кореляції та критерій  $t_r$ :

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{(\sum x^2 - (\sum x)^2 : n)(\sum y^2 - (\sum y)^2 : n)}} =$$

$$= \frac{2762,33 - (352,22 \times 396,41) : 51}{\sqrt{(2502,52 - (352,22)^2 : 51)(3134,14 - (396,41)^2 : 51)}} =$$

$$= 0,40;$$

$$s_r = \sqrt{\frac{1 - r^2}{n - 2}} = \sqrt{\frac{1 - 0,40^2}{51 - 2}} = 0,13;$$

$$t_r = \frac{r}{s_r} = \frac{0,40}{0,13} = 3,08; \quad v = n - 2 = 51 - 2 = 49; \quad t_{05} = 1,96$$

За  $t$ -критерієм ( $t_r > t_{05}$ ) кореляція є істотною на 5%-му рівні, нульова гіпотеза про незалежність  $y$  від  $x$  відкидається. Отже, можна розрахувати коефіцієнт успадкованості  $h^2$ .

2. Коефіцієнт  $h^2$  розраховують як подвоєний коефіцієнт кореляції між фенотипами батьків та нащадків:

$$h^2 = 2r = 2 \times 0,40 = 0,80$$

Таким чином, близько 80 % фенотипічної мінливості у вмісті олії зумовлено спадковою мінливістю рослин кукурудзи. Отже, добір зерна за цією ознакою може бути ефективним.

**Приклад 2.** За даними маси зерна з одного колоса гібридів пшениці (таблиця 10) розрахувати коефіцієнт успадкованості методом дисперсійного аналізу.

Таблиця 10

**Маса зерна з одного колоса гібридів (г)**

Гібрид (варіанти, $l$ )		Повторення, $x$					Число спостережень, $n$	Суми, $V$	Середні, $\bar{x}$
♀	♂	1	2	3	4	5			
1	1	5,5	4,1	3,0	2,2	3,8	5	18,6	3,72
1	2	5,0	3,1	3,0	2,1	4,0	5	17,2	3,44
1	3	5,2	3,7	3,2	2,0	3,9	5	18,0	3,60
1	4	5,4	5,0	3,4	2,7	4,74	5	20,9	4,18
<b>Загальна сума, <math>P</math></b>		<b>21,1</b>	<b>15,9</b>	<b>12,6</b>	<b>9,0</b>	<b>16,1</b>	<b><math>N =</math> <math>\sum n = 20</math></b>	<b><math>\sum X = 74,7</math></b>	<b>-</b>

*Розв'язання.* 1. Здійснюють дисперсійний аналіз для однофакторного комплексу:

$$C = \frac{(\sum X)^2}{N} = \frac{(74,7)^2}{20} = 279;$$

$$C_y = \sum X^2 - C = (5,5^2 + 4,1^2 + \dots + 4,4^2) - 279 = 22,91;$$

$$C_p = \sum \frac{P^2}{l} - C = (21,1^2 + 15,9^2 + \dots + 16,1^2) : 4 - 279 = 20,25;$$

$$C_v = \sum \frac{V^2}{n} - C = (18,6^2 + 17,2^2 + \dots + 20,9^2) : 5 - 279 = 1,52$$

$$C_z = C_y - C_p - C_v = 22,91 - 20,25 - 1,51 = 1,14$$

2. Одержані результати записують у таблицю дисперсійного аналізу та визначають значущість дії генотипів на фенотипічну мінливість ознаки за  $F$ -критерієм.

Таблиця 11

Дисперсія	Сума квадратів	Ступені свободи	Середній квадрат	$F_{\text{факт.}}$	$F_{05}$
Загальна $C_y$	22,91	19	–	–	–
Повторень, $C_p$	20,25	4	–	–	–
Варіантів $C_v$	1,52	3	$s_v^2 = 0,5066$	5,33	3,49
Залишок $C_z$	1,14	12	$s^2 = 0,0950$	–	–

Вплив батьківської форми на середню масу зерна з одного колоса виявився суттєвим ( $F_{\text{факт.}} > F_{05}$ ). Отже, є доцільність розрахувати коефіцієнт успадкованості  $h^2$ , який характеризує ступінь генетичної мінливості.

Необхідно розкласти фенотипічну мінливість, виражену через дисперсію ( $s_p^2$ ) на варіанти, що її складають: *генотипічну* ( $s_G^2$ ) та *паратипічну* або середовищну ( $s^2$ ). Це дає можливість установити в групі особин частку мінливості, що зумовлена їх генетичною різноманітністю, та частку мінливості, пов'язану із факторами середовища.

Таким чином, середній квадрат для варіантів у нашому прикладі батьківських форм  $s_v^2 = s^2 + ns_G^2$ . Визначаємо фенотипічну та загальну фенотипічну дисперсію:

$$s_G^2 = \frac{s_v^2 - s^2}{n}; \quad s_p^2 = s_G^2 + s^2$$

3. Розрахунок фенотипічної та фенотипічної мінливості, а також коефіцієнта успадкованості за результатами дисперсійного аналізу:

$$s_G^2 = \frac{s_v^2 - s^2}{n} = \frac{0,5066 - 0,0950}{5} = 0,0823$$

$$s_p^2 = s_G^2 + s^2 = 0,0823 + 0,0950 = 0,1773$$

$$h^2 = \frac{s_G^2}{s_p^2} = \frac{0,0823}{0,1773} = 0,46 \text{ або } 46 \%$$

Таким чином, коефіцієнт успадкованості  $h^2$ , що характеризує ступінь передачі ознаки – маса зерна з одного колоса – від батьківської форми до гібрида, становить 0,46 та, відповідно, добір за цією ознакою може бути ефективним.

Зміна генетичного складу популяції під впливом добору, називається генетичним зрушенням або *селекційним ефектом* ( $SE$ ).

Мірою інтенсивності відбору є *селекційний диференціал* ( $SD$ ), тобто різниця між середньою величиною ознаки у особин, що відбираються в якості батьків, і середньої величиною ознаки в популяції.

$$(S_d = X_1 - X_0),$$

Для прогнозування ефективності добору із врахуванням коефіцієнту успадкованості користуються наступною формулою:

$$SE = SD \times H,$$

де  $SE$  – селекційний ефект,

$SD$  – селекційний диференціал,

$H$  – коефіцієнт успадкованості ознаки.

**Приклад 3.** Кількість абдомінальних щетинок в одній з ліній *Drosophila melanogaster* становить 38. Мух з середньою кількістю щетинок 42,8, схрестили для отримання наступного покоління; отже,  $SD = 42,8 - 38 = 4,8$ . Середня кількість щетинок в потомстві відібраних мух була 40,6; звідси  $SE = 40,6 - 38 = 2,6$ . Успадкування оцінюється як

$$H = SE / SD = 2,6 / 4,8 = 0.54$$

### Контрольні запитання:

1. Що таке генотипова мінливість?
2. Що таке фенотипова мінливість?
3. У чому полягає відмінність між термінами "спадковість", "успадкування" та "успадковуваність"?
4. Що таке коефіцієнт успадкованості?
5. Назвіть значення  $h^2$  основних селекційних ознак культурних рослин.
6. Які фактори впливають на значення коефіцієнта успадкованості?

## МОДУЛЬ 2. Біометричні методи у селекції сільськогосподарських культур

### ПРАКТИЧНА РОБОТА № 7

- Тема.** Аналіз генетичних процесів у популяції. Встановлення генетичної рівноваги та подібності популяцій.
- Мета:** опанувати методикою визначення генетичної структури популяцій

*Популяцією* називається сукупність особин одного виду, що займають визначений ареал, вільно схрещують між собою, мають спільне походження і в тій або іншій мірі ізольовані від інших популяцій даного виду.

*Генетична структура популяції* – частота перебування в ній особин, що мають усі можливі поєднання в своєму генотипі домінантних і рецесивних алелів відповідних генів –  $AA$ ,  $Aa$ ,  $aa$ , або частоту знаходження кожного алеля даного гену.

Генетична структура популяції підпорядковується закону Харді-Вайнберга:

*„У необмежено великій популяції при відсутності відбору і мутуванні даних генів та відсутності міграції, числові співвідношення генотипів  $AA$ ,  $aa$  й  $Aa$  залишаються з покоління в покоління постійними.“*

Згідно з законом Харді-Вайнберга в панміктичній популяції концентрація (частота) алелів залишається постійною нескінченне число поколінь.

*Панміксія* – однакова імовірність перебування гамет при вільному перезапиленні всіх рослин у популяції.

Якщо концентрацію алелів  $A$  і  $a$  позначити відповідно  $p$  і  $q$  ( $p+q=1$ ), то математичним вираженням співвідношення генотипів у популяції внаслідок панміктичного схрещування буде формула:

$$p^2 AA + 2pqAa + q^2 aa = 1.$$

Закон Харді-Вайнберга використовується для математичних прогнозів кількісних співвідношень особин з різними генотипами за одним алелем, або для визначення зустрічальності цього гену серед даної популяції. Такі прогнози знаходять практичне застосування в екології, соціології (наука про охорону природи), сільськогосподарській практиці та селекції, в медичній практиці тощо.

Визначити генетичну структуру популяції – означає знайти концентрацію алелів ( $pA$ ,  $qa$ ) у ній або визначити частоти гомозиготних ( $AA$ ,  $aa$ ) або гетерозиготних ( $Aa$ ) генотипів у першому поколінні.

Концентрацію алелів у популяції найбільш точно можна обчислити прямим шляхом. Це можливо в тому разі, якщо фенотипово ідентифікуються всі генотипи (як гомо- так і гетерозиготні). Розв'язання задачі зводиться до обліку особин у кожному фенотиповому класі й

підрахунку кількості алелів у особин, які складають цей клас. Так, гомозиготи несуть два однакових алеля, а гетерозиготи – два різних (приклад 1). Частоти генів можуть бути виражені у долях одиниці або у відсотках, тоді:  $pA + qa = 1$  або  $pA + qa = 100\%$ .

Якщо особини з генотипами  $AA$  і  $Aa$  фенотипово не розрізнені, то концентрацію алелів обчислюють за формулою Харді-Вайнберга. Формулу можна використовувати тільки тоді, коли є впевненість, що популяція наближається до ідеальної, у протилежному разі результат буде перевернутий. Згідно із формулою Харді-Вайнберга частота гомозигот у панміктичній популяції становить  $p^2AA$  або  $q^2aa$ , а гетерозигот  $2pqAa$  (приклад 2).

За множинного алелізму частоти генотипів такі: частка гомозигот дорівнює квадратам частот відповідних алелів, частка гетерозигот – це подвоєний добуток частот відповідних алелів. Зокрема, для триалельної системи ( $pA_1, qA_2, rA_3$ ) формула така:

$$p^2A_1A_1 + q^2A_2A_2 + r^2A_3A_3 + 2pqA_1A_2 + 2prA_1A_3 + 2qrA_2A_3 = 1.$$

**Приклад 1.** У рослини нічна красуня червоний колір квіток неповністю домінує над білим. Гібриди від схрещування червоноквіткових і білокріткових рослин мають рожеві квітки. На ділянці площею 100 м<sup>2</sup> росте 4169 червоноквіткових, 3780 рожевих та 756 білокріткових рослин. Визначте частоту алелів червоного і білого забарвлення квітки нічної красуні у даній популяції.

Дано:

Розв'язання:

Забарвлення квіток:

червоних ( $AA$ ) – 4169

рожевих ( $Aa$ ) – 3780

білих ( $aa$ ) – 756

---

Знайти:  $p_A$  і  $q_a$

Кількість алелів

$$A = 4169 \times 2 + 3780 = 12118$$

Кількість алелів

$$a = 756 \times 2 + 3780 = 5292$$

$$\text{Сума} = 17410 \text{ (100 \% або 1)}$$

$$p_A = \frac{12118}{17410} = 0,696;$$

$$p_A + q_a = 1$$

$$q_a = 1 - 0,696 = 0,304$$

$$\text{Відповідь: } p_A = 0,696; q_a = 0,304$$

**Приклад 2.** В озимого жита опушення соломини під колосом домінує над відсутністю опушення та успадковується моногенно. У популяції жита сорту В'ятка при аналізі апробаційного снопа було виявлено 4 рослини із 500, у яких відсутнє опушення під колосом. Яка частота рецесивного та домінантного алеля у популяції?

Жито – перехреснозапильна культура, розмножується статеву, тому вважаємо, що популяція панміктична. Зважаючи на те, що мутації виникають рідко, ними можна знехтувати. Аналіз рослин з апробаційного снопа дозволить зробити висновок про генетичну структуру всієї популяції жита цього сорту.

Дано:  
 Опушення соломини жита:  
 $A$  – опушена  
 $a$  – неопушена  
 $A > a$   
 Об'єм популяції – 500  
 неопушених рослин – 4

---

Знайти:  $p_A$  і  $q_a$

Розв'язання

$$p + q = 1$$

$$p^2 AA + 2pqAa + q^2 aa = 1$$

опуш.    опуш.    неопуш.

$$q^2 aa = \frac{4}{500}; q_a = \sqrt{0,008} = 0,09$$

$$p_A = 1 - q = 0,91$$

Відповідь:  $p_A = 0,91$ ;  $q_a = 0,9$

**Приклад 3.** Як зміниться кожний розподіл генотипів у популяції ( $p^2 AA = 0,36$ ) + ( $2pqAa = 0,48$ ) + ( $q^2 aa = 0,16$ ) при встановленні нової концентрації алелів:  $p_A = 0,7$ ,  $q_a = 0,3$ .

Дано:  
 Початкове співвідношення частот генотипів:  
 $0,36AA + 0,48Aa + 0,16aa = 1$   
 Остаточна концентрація алелів:  
 $p_A = 0,7$   
 $q_a = 0,3$

---

Остаточне співвідношення частот генотипів – ?

Розв'язання.

- 1) Визначаємо частоту генотипу  $AA$ :  
 $p^2 AA = (p_A)^2$   
 $p^2 AA = 0,7^2 = 0,49$
- 2) Визначаємо частоту генотипу  $aa$ :  
 $q^2 aa = (q_a)^2$   
 $q^2 aa = 0,3^2 = 0,09$
- 3) Визначаємо частоту генотипу  $Aa$ :  
 $2pqAa = 2 \times 0,7 \times 0,3 = 0,42$
- 4) Визначаємо остаточне співвідношення частот генотипів:  
 $p^2 AA + 2pqAa + q^2 aa = 1$   
 $0,49AA + 0,42Aa + 0,09aa = 1$

Відповідь: остаточне співвідношення частот генотипів виражається рівнянням:  $0,49AA + 0,42Aa + 0,09aa = 1$ . Отже, частота гомозиготних рецесивів зменшилася, а гомозиготних домінант збільшилася у порівнянні з вихідними даними.

**Приклад 4.** У популяції, яка розмножується шляхом вільного схрещування, існує така частота генотипів:  $0,2AA$  і  $0,8Aa$ . Визначити, які частоти генотипів  $AA$ ,  $Aa$  і  $aa$  встановляться у першому поколінні після схрещування у даній популяції.

Дано:

$$p^2 AA = 0,2 \text{ (20\%)}$$

$$2pqAa = 0,8 \text{ (80\%)}$$

В F<sub>1</sub>:

$$p^2 AA - ?$$

$$2pqAa - ?$$

$$q^2 aa - ?$$

Розв'язання.

В даній популяції за кожним генотипом утворюється однакова чисельність яйцеклітин і спермій. Оскільки відбувається вільне схрещування, чисельність ймовірних співвідношень знаходимо алгебраїчним шляхом:

Утворення яйцеклітин:

$$p^2 AA = 0,2 \quad 2pqAa = 0,8 \quad \begin{cases} \text{гамет}(A) = 0,4 \\ \text{гамет}(a) = 0,4 \end{cases}$$

Стільки ж утворюється і сперматозоїдів.

Звідси співвідношення особин, які утворилися з різними генотипами:

G	♀ \ ♂	$pA(0,6)$	$qa(0,4)$
F <sub>1</sub>	$pA(0,6)$	$p^2 AA$ 0,36	$pqAa$ 0,24
	$qa(0,4)$	$pqAa$ 0,24	$q^2 aa$ 0,16

Відповідь:  $p^2 AA = 0,36$ ,  $2pqAa = 0,48$ ,  $q^2 aa = 0,16$ .

### Контрольні запитання:

1. Поняття про популяцію як елементарну одиницю еволюції.
2. Яку популяцію прийнято вважати панміктичною?
3. Сформулюйте закон Харді-Вайнберга.



## ПРАКТИЧНА РОБОТА № 8

**Тема.** Розв'язання задач біометричної генетики засобами MS EXCEL

**Мета:** освоїти застосування стандартних функцій MS Excel для вирішення задач описової статистики.

### Завдання 1.

Проведіть аналіз даних в рамках описової статистики з використанням засобів **Вставка функцій** і **Майстер діаграм** MS Excel

1. Запустіть MS Excel: **Пуск / Програми / MS Excel** і перейменуйте ярличок робочого аркуша **Лист 1**: подвійне клацання по ярличку і надрукуйте поверх виділення **Статистика 1**. Введіть вихідні дані та заголовки статистичної таблиці за **зразком 1**: виділіть комірку A1 клацанням миші / введіть текст заголовка і зафіксуйте клацанням по інструменту Enter / розташуйте заголовок по центру стовпців A-E - виділіть комірки A1: E1 і натисніть інструмент **Об'єднати і помістити в центрі** / аналогічно виділяючи послідовно комірки A2 - E11 введіть числа вихідних даних таблиці 12:

Таблиця 12

### Заміри висоти травостою лучного газону

18	21	20	17	24
17	18	23	24	18
22	22	27	23	17
16	17	22	24	25
22	21	19	20	16
21	23	26	17	22
18	24	15	27	22
24	22	21	24	21
18	19	22	27	24
25	22	17	21	21

В комірках G2 - G14 заголовки рядків статистичної таблиці і число вибірок:

Таблиця 13

Середнє	СРЗНАЧ
Середньоквадратичне відхилення	СТАНДОТКЛОН
Дисперсія	ДИСПА
Медіана	МЕДИАНА
Мода	МОДА
Асиметрія	СКОС
Ексцес	ЕКСЦЕС

Найменше значення	МИН
Найбільше значення	МАКС
Кількість вибірок	50

Виконайте розрахунки зазначених в заголовках рядків статистичної таблиці параметрів, вставляючи при допомозі засобів **Вставка функцій** розрахункові формули. Наприклад, для розрахунку середнього виділіть клацанням миші комірку Н2, клацніть інструмент **Вставка функцій** / у вікні **Майстер функцій** в полі **Категорії** клацніть **Статистичні**, в полі **Функція** за допомогою смуги прокрутки перегорніть список назви функцій, знайдіть і клацніть **СРЗНАЧ** і **ОК** / у вікні вставки функції праворуч від поля **Число 1** клацніть кнопку згортання / виділіть мишею діапазон комірок А2: Е11 (утримуючи ліву кнопку миші) / в згорнутому вікні вставки функції клацніть кнопку розгортання / **ОК**; аналогічно вставте інші формули.

2. Сформууйте таблицю частот досліджуваної величини, виконавши угруповання даних і розрахунки безпосереднім введенням формул і за допомогою засобу **Вставка функцій**:

- вставте формулу для обчислення мінімального числа інтервалів групування за допомогою засобу **Вставка функцій**: виділіть комірку А14 і введіть «мін. кількість інтервалів », виділіть комірку В14 / інструмент **Вставка функцій** / в полі **Категорії** клацніть **Математичні** / в полі **Функція** знайдіть і виберіть **ОКРУГЛ** і **ОК** / у вікні вставки функції встановіть курсор в полі **Число розрядів** і введіть 0 (округлення до цілого числа), встановіть курсор в поле **Число** і введіть 5 \* (множник) / в інструменті вибору функції (лівий верхній кут робочої книги) клацніть кнопку списку і виберіть позицію **Інші функції ...** / у вікні **Майстер функцій** виберіть категорію **LOG10** з категорії **Математичні** і **ОК** / у вікні вставки функції в поле **Число 1** введіть посилання з числом вибірок Н14 і **ОК**.

- вставте формулу для розрахунку ширини інтервалу за допомогою введення з клавіатури: виділіть комірку А15 і введіть «ширина інтервалу», виділіть комірку В15 / введіть знак = (дорівнює) і знак ((дужка) / клацніть комірку з максимальним значенням Н10 і натисніть клавішу F4 для переходу до абсолютної посилання / введіть знак - / клацніть клітинку з мінімальним значенням Н9 і натисніть клавішу F4 / введіть знак) (дужка) і знак / (похила риска) і клацніть комірку В14 з числом інтервалів / **Enter**.

- аналогічно в комірці А20-А27 вставте формули для обчислення правих меж інтервалів: клацніть комірку А20, введіть знак = (дорівнює) / клацніть клітинку з мінімальним значенням Н9 і натисніть клавішу F4 для переходу до абсолютної посилання / введіть знак + (плюс) і клацніть комірку з значенням ширини інтервалу В15 / **Enter**; в комірці А21 введіть формулу = А20 + \$ В \$ 15; в комірці А22 і нижче розтягніть формулу з комірки А20 за допомогою автозаповнення: після введення формули в А21 вкажіть на нижній правий кут комірки А21 до появи маркера автозаповнення у формі +, натисніть ліву кнопку миші і, утримуючи її,

протягніть виділення комірки до A27 і відпустіть кнопку миші.

- вставте формулу для розрахунку частот із застосуванням функції масивів: виділіть діапазон комірок B20-B27 / інструмент **Вставка функцій** / знайдіть і виберіть функцію **ЧАСТОТА** з категорії **Статистичні** та **ОК** / у вікні вставки функції праворуч від поля **Масив даних** клацніть кнопку згортання / виділіть мишею діапазон комірок A2: E11 / клацніть кнопку розгортання / справа від поля **Масив інтервалів** клацніть кнопку згортання / виділіть мишею діапазон комірок A20: A27 / клацніть кнопку розгортання / одночасно натисніть клавіші **Ctrl, Shift, Enter** для фіксації функції масиву.

3. Побудуйте гістограму для досліджуваної величини із застосуванням майстра діаграм: виділіть діапазон комірок з таблицею частот A20: B27 / інструмент **Майстер діаграм** / на вкладці **Нестандартні** в полі **Тип** виберіть **Графік / Гістограма 2** і кнопка **Далі** / в вікні **Вихідні дані** на вкладці **Діапазон даних** включіть перемикач в стовпцях / на вкладці **Ряд** клацніть кнопку згортання праворуч від поля **Підписи по осі X** / виділіть діапазон комірок A20:A27 і клацніть кнопку розгортання / в полі **Підписи другої осі X** внесіть діапазон комірок B20:B27 і кнопка **Далі** / в вікні розміщення діаграми включіть перемикач на наявному листі і **ОК**.

**Завдання 2.** Виконайте процедуру генерації випадкових чисел і проаналізуйте їх за допомогою засобів **Аналіз даних** і **Майстер діаграм MS Excel**.

1. Перейдіть на вільний робочий лист книги і перейменуйте його в **Генерація даних**.

2. Підключіть надбудову **Пакет аналізу MS Excel: Сервіс / Надбудови** / в вікні Надбудови встановіть прапорець **Пакет аналізу** та **ОК**.

3. Виконайте генерацію 30 випадкових чисел, розподілених відповідно до нормального закону з нульовим середнім і дисперсією 1: клацніть комірку A1 і **Сервіс / Аналіз даних** / в полі зі списком **Інструмент аналізу** клацніть позицію **Генерація випадкових чисел** і **ОК** / в полі **Число змінних** введіть 1, в поле **Число випадкових чисел** введіть 30, розкрийте список поля **Розподіл** і виберіть позицію **Нормальне**, введіть в полях **Середнє** - 0, **Стандартне відхилення** - 1, в розділі **Параметри виведення** включіть перемикач **вихідний інтервал**, клацніть кнопку згортання / клацніть комірку A1 і кнопку розгортання, **ОК**.

4. Змініть розрядність даних, зменшіть число знаків після коми до двох: виділіть діапазон комірок A1: A30 / клацніть інструмент **Зменшити розрядність** чотири рази.

5. Виконайте процедуру описової статистики по згенерованим даним: **Сервіс / Аналіз даних / Описова статистика** і **ОК** / в вікні Описова статистика в полі **Вхідний інтервал** введіть посилання на діапазон комірок A1: A30 / в розділі **Групування** включіть перемикач по стовпцях і приберіть прапорець **Мітки в першому рядку** / в розділі **Параметри**

**виведення** включите перемикач **Вихідний інтервал** і клацніть комірку C1 / встановіть прапорець **Підсумкова статистика** і **ОК**.

6. Побудуйте гістограму за даними стовпчика A: **Сервіс / Аналіз даних / Гістограма** і **ОК** / в вікні Гистограмма в розділі **Вхідні дані** в поле **Вхідний інтервал** введіть посилання на діапазон комірок A1: A30 і встановіть прапорець **Мітки** / в розділі параметри виводу включіть перемикач **Вихідний інтервал** і вкажіть будь-яку вільну комірку робочого листа / встановіть прапорець **Інтегральний відсоток** і **Висновок графіка** і **ОК**. Додайте на побудовану гістограму 2 лінії тренда: поліноміальної зі ступенем 4 і змінного середнього на 2 точки: клацніть правою клавішею миші по рядах значень, на контекстному меню вибрати позицію **Додати лінію тренда** / на вкладці тип вибрати **Поліноміальна**, в поле **ступінь** ввести 4 і т.д.

### **Контрольні запитання:**

1. Побудова діаграми розсіювання.
2. Розрахунок кореляції за допомогою стандартних функцій MS Excel.
3. Однофакторний дисперсійний аналіз засобами пакету аналізу MS Excel.

Додаток 1

**Стандартні значення критерію достовірності  $t_d$  для малих вибірок (за Стьюдентом)**

Число ступенів свободи	Вірогідність (P)		
	0,95	0,99	0,999
1	12,70	63,66	-
2	4,30	9,93	31,60
3	3,18	5,84	12,94
4	2,78	4,60	8,61
5	2,57	4,03	6,86
6	2,45	3,71	5,96
7	2,37	3,50	5,41
8	2,31	3,36	5,04
9	2,26	3,25	4,78
10	2,23	3,17	4,59
11	2,20	3,11	4,44
12	2,18	3,06	4,32
13	2,16	3,01	4,22
14	2,15	2,98	4,14
15	2,13	2,95	4,07
16	2,12	2,92	4,02
17	2,11	2,90	3,97
18	2,10	2,88	3,92
19	2,09	2,86	3,88
20	2,09	2,85	3,85
21	2,08	2,83	3,82
22	2,07	2,82	3,79
23	2,07	2,81	3,77
24	2,06	2,80	3,75
25	2,06	2,79	3,73
26	2,06	2,78	3,71
27	2,05	2,77	3,69
28	2,05	2,76	3,67
29	2,05	2,75	3,66
30	2,04	2,75	3,65
$\infty$	1,96	2,58	3,29

## Додаток 2

Таблиця значень *F*-критерію Фішера при 5%-му рівні значущості

		Ступені свободи для більшої дисперсії (чисельника)									
		1	2	3	4	5	6	8	12	24	∞
Ступені свободи для меншої дисперсії (знаменника)	1	161,45	199,50	215,72	224,57	230,17	233,97	238,89	243,91	249,04	234,52
	2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
	3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
	4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
	5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
	6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
	7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
	8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
	9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
	10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
	11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
	12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
	13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
	14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
	15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
	16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
	17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
	18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
	19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
	20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
	21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
	22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
	23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
	24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
	25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
	26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
	27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
	28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
	29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
	30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
	35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
	40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
	45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44	
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39	
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35	
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31	
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,65	1,31	
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26	
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21	
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18	
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14	
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10	
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07	
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06	
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03	
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00	

Додаток 3

**Стандартні значення  $\chi^2$  за різних ступенів свободи**

(за Р. Фішером, зі скороченням)

Число ступенів свободи ( <i>df</i> )	Вірогідність (P)		
	0,05	0,01	0,001
1	3,84	6,63	10,83
2	5,99	9,21	13,82
3	7,81	11,34	16,27
4	9,49	13,28	18,47
5	11,07	15,00	20,50
6	12,59	16,81	22,50
7	14,07	18,48	24,30
8	15,51	20,09	26,10
9	16,92	21,67	27,90
10	18,31	23,31	29,60

## СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Биометрия в генетике и селекции растений / А. В. Смиряев, С. П. Мартынов, А. В. Кильчевский. – М. : Изд-во МСХА, 1992. – 269 с.
2. Бугайов В. Д. Спеціальна селекція польових культур : навч. посіб. / В. Д. Бугайов, С. П. Васильківський, В. А. Власенко [та ін.] ; за ред. М. Я. Молоцького. – Біла Церква, 2010. – 368 с.
3. Генетика з основами селекції : підручник / С. І. Стрельчук, С. В. Демидов, Г. Д. Бердишев та ін. – Київ, 2000. – 292 с.
4. Горошко М. П. Біометрія: навчальний посібник / М. П. Горошко, С. І. Миклуш, П. Г. Хомюк. – Львів : Камула, 2004. – 236 с.
5. Доспехов Б. А. Методика полевого опыта / Б. А. Доспехов. – М. : Агропромиздат, 1985. – 351 с.
6. Калінін М. І. Біометрія : підручник для студентів вузів біологічних і екологічних напрямків / М. І. Калінін, В. В. Єлісеєв. Миколаїв : Вид-во МФ НаУКМА, 2000. – 204 с.
7. Марценюк І. М. Генетика. Практикум : навчальний посібник / І. М. Марценюк. – Миколаїв : МНАУ, 2014. – 148 с.
8. Меркурьева Е. К. Генетика с основами биометрии / Е. К. Меркурьева, Г. Н. Шангин-Березовский. – М. : Колос, 1983. – С. 243 – 260.
9. Методика наукових досліджень в агрономії: навч. посіб. / В. Г. Дідора, О. Ф. Смаглий, Е. Р. Ермантраут [та ін.] ; за ред. В. Г. Дідори. – К. : «Центр учбової літератури», 2013. – 264 с.
10. Основи наукових досліджень в агрономії : підручник / В. О. Єщенко, П. Г. Копитко, В. П. Опришко, П. В. Костогриз. – К. : «Дія», 2005. – 288 с.
11. Сиволоб А. В. Генетика : підручник / [А. В. Сиволоб, С. Р. Рушковський, С. С. Кир'яченко та ін.] ; за ред. А. В. Сиволоба. – К. : Видавничо-поліграфічний центр "Київський університет", 2008. – 320 с.
12. Ушкаренко В. О. Дисперсійний і кореляційний аналіз результатів польових дослідів: навч. посіб. / В. О. Ушкаренко. – Херсон : Айлант, 2009. – 372 с.



13. Хмельничий Л. М. Основи біометрії : методичні вказівки до виконання лабораторних та самостійних робіт студентам освітньо-кваліфікаційного рівня «Бакалавр» спеціальність – 6.090102“ Технологія виробництва і переробки продукції тваринництва” та 6.110.100 “ Ветеринарна медицина” / уклад. : Л. М. Хмельничий, І. О. Супрун. – К. : Видавничий центр НУБіП України, 2010. – 69 с.
14. Хмельничий Л. М. Основи генетики з біометрією / Л. М. Хмельничий, І. О. Супрун, А. М.Салогуб. – Суми : ПП Вінниченко М.Д., ФОП Дьоменко В.В. 2011. – 344 с.
15. Чекалин Н. М. Простые и частные коэффициенты генетической корреляции между урожаем и признаками продуктивности колоса у линий и сортов озимой пшеницы / Н. М. Чекалин, В. Н. Тищенко, М. Е. Зюков // Збірник наукових праць СГІ. – Одеса : 2004. – Вип. 6 (46). – С. 103-110.
- 16.Чепур С. С. Біометрія : методичний посібник. / С. С. Чепур. – Ужгород: Видавництво УжНУ «Говерла», 2015. – 40 с.
17. Боровиков В. П. STATISTICA. Искусство анализа данных на компьютере / В. П. Боровиков. 2-е изд. – Спб : Питер, 2003. – 688 с.
- 18.Халафян А. А. STATISTICA 6. Статистический анализ данных : учебник / А. А. Халафян. 3-е изд. – М. : ООО «Бином-Пресс», 2008 г. – 512 с.



Навчальне видання

# **БІОМЕТРИЧНА ГЕНЕТИКА СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР**

Методичні рекомендації

Укладач: **Марценюк** Ігор Михайлович

Формат 60x84 1/16. Ум. друк. арк. \_\_\_\_.

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 20.02.2013 р.

