

МІНІСТЕРСТВО АГРАРНОЇ ПОЛІТИКИ УКРАЇНИ
БІЛОЦЕРКІВСЬКИЙ ДЕРЖАВНИЙ АГРАРНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ВЕТЕРИНАРНОЇ МЕДИЦИНИ
Кафедра фізики та вищої математики

**СТАТИСТИЧНА
ОБРОБКА
ЕКСПЕРИМЕНТАЛЬНИХ
ДАНИХ**

Навчальний посібник

Біла Церква
2006

УДК 517 (075.8)

Рекомендовано до видання вченою радою
факультету ветеринарної медицини
(Протокол №9 від 19.04.2006р.)

Укладачі: **О.П. Мельниченко**, асистент; **І.Л. Якименко**, д-р біол. наук, професор;
Р.Л. Шевченко, канд. фіз.-мат. наук, доцент.

Статистична обробка експериментальних даних: Навчальний посібник / О.П. Мельниченко, І.Л. Якименко, Р.Л. Шевченко – Біла Церква, 2006.– с.

Розглянуто методи обробки і аналізу статистичної інформації, зокрема методи Стьюдента та Фішера; введені поняття середньої арифметичної величини, середнього квадратичного відхилення, визначення помилки репрезентативності та коефіцієнта кореляції.

Кожен параграф містить стисло викладений теоретичний матеріал, проілюстрований прикладами та зразками розв'язування типових вправ. Обов'язковий елемент – задачі для самостійного розв'язування та теоретичні питання. До посібника входить алфавітний покажчик, що сприятиме можливості швидко відшукати відповідь на те чи інше запитання.

Рекомендований для студентів факультету ветеринарної медицини.

Рецензент: **М.І. Трегуб**, зав. кафедри механізації, канд. техн. наук

© БДАУ, 2006

ВСТУП

Статистика (від лат. *status* – стан) вивчає кількісний бік суспільних явищ і процесів у нерозривному зв'язку з їх якісним змістом. Інакше кажучи, статистика – це наука, що збирає, обробляє різні дані, пов'язані з масовими явищами, процесами, подіями та вивчає їх закономірності.

Статистику поділяють на *описову* та *пояснювальну*. Добір необхідної інформації для проведення досліджень – справа описової статистики. Прикладом може бути Книга рекордів Гіннеса, в якій збирається інформація про якісні та кількісні характеристики виняткових подій та явищ.

Статистика дає стисло й чітку характеристику досліджуваного масового явища і водночас можливість виявити та пояснити тенденції його розвитку. За даними пояснювальної статистики формулюються певні висновки, робиться прогноз.

I. Основні поняття теорії ймовірностей

Випадковою називається змінна величина, яка в результаті дослідження може з'явитися, а може не з'явитися, незалежно від бажання дослідника. Будь-який результат проведеного випробування називають *подією*. Наприклад, з яйця вилупилося курча – подія. Події позначають літерами A, B, C тощо. Ймовірність появи події A позначається $P(A)$. Поняття ймовірності є базовим поняттям у теорії ймовірностей і широко використовується в математичній статистиці.

Ймовірність P появи події A називається відношення числа m випробувань, що сприяють появі події A , до загального числа n всіх можливих випробувань:

$$P(A) = \frac{m}{n}.$$

Наприклад, свиноматка привела 10 поросят, з них 6 свинок. З цього опоросу можна по одному вилучити всі 10 поросят, тобто загальне число випробувань по вилученню поросят дорівнює десяти: $n=10$. Серед цих можливих випробувань сприяють появі свинки тільки шість випробувань, тобто $m=6$. Якщо подія A полягає у вилученні свинки, то $P(A) = \frac{m}{n} = \frac{6}{10} = 0,6$.

Ймовірність змінюється від нуля (якщо $P(A)=0$, то подія неможлива) до одиниці (якщо $P(A)=1$, то подія достовірна).

Отже визначити ймовірність появи будь-якої події не викликає ніяких труднощів, якщо відомі значення всіх можливих подій та значення їх сприяючої частини. Існують умови, за яких значення m та n невідомі і безпосередньо не можуть бути знайдені.

Нехай маємо скриньку з 10 кульками, з яких 4 білих (отже імовірність дістати одну білу кульку буде 0,4). Скринька закрита і має невеликий люк, через який можна дістати одну кульку. Перед дослідженням ставимо задачу: знайти можливість дістати білу кульку, якщо дозволяється діставати необмежену

кількість разів по одній кульці і після запису її кольору повертати у скриньку. Нехай дослідник проведе 10000 вилучень, серед яких біла кулька зустрінеться 4125 разів. У цьому випадку відношення $\frac{m}{n} = \frac{4125}{10000} = 0,4125$ буде відносною

частотою появи білої кульки у сьогоднішньому досліді. Зрозуміло, що в наступному аналогічному досліді число білих кульок буде іншим (наприклад, 3997). Значення частоти появи кульки білого кольору буде 0,3997. Таких дослідів можна провести досить багато. Наприклад, якщо цей дослід проводити кожен день протягом, то будемо мати 365 значень відносних частот, деякі з них навіть можуть повторюватися. Аналіз отриманих значень показує, що є таке число, навколо якого групуються всі отримані значення відносних частот. У нашому випадку це число буде близьким до числа 0,4.

Число, навколо якого групуються значення відносних частот, називається *статистичною ймовірністю*.

Всі змінні величини, в тому числі і випадкові, можуть змінюватись лише двома способами: дискретно або неперервно.

Якщо між двома найближчими значеннями випадкової величини інших значень не може бути в принципі, то така величина називається *дискретною* (від лат. *discretus* – розривний). Наприклад, число поросят – величина дискретна, свиноматка може привести тільки одне ціле число поросят.

Якщо ж між двома як завгодно близькими значеннями випадкової величини існує безліч інших значень, то така величина називається *неперервною*. Наприклад, врожайність пшениці (в ц/га) є величиною неперервною, оскільки між двома її навмання взятими значеннями 30 ц/га і 31 ц/га вміщується безліч інших значень врожайності, яку можна отримати на цій ланці.

Нехай вздовж лінії довжиною 1м рухається точковий промінь. Відстань від лівого краю лінії до променя змінюється неперервним чином. Якщо промінь переміщується хаотично, міняючи швидкість та напрям руху незалежно від дослідника, то вимірювання відстані буде величиною випадковою. За умови, що промінь не зникає і не виходить за межі лінії, можемо стверджувати, що промінь знаходиться на лінії гарантовано, тобто ймовірність цієї події – одиниця. Ймовірність того, що промінь у вибраний момент часу знаходиться на будь-якій половині лінії буде дорівнювати 0,5, на будь-якій четвертій частині лінії – 0,25, на десятій частині – 0,1 і т.д. Чим вужчий інтервал, тим менша ймовірність попадання випадкової величини в цей інтервал.

Для дискретної величини завжди можна вказати ймовірність появи кожного з її значень, тоді як для неперервної величини можна вказати лише ймовірність її попадання в деякий інтервал.

Закон зміни ймовірностей появи випадкових величин називається *законом розподілу*.

Будь-яка величина X характеризується трьома числовими характеристиками: математичним сподіванням $M(X)$, дисперсією $D(X)$ та середнім квадратичним відхиленням $\sigma(X)$.

Математичним сподіванням $M(X)$ називається число, до якого прагнуть всі середні арифметичні значення випадкової величини (істинне середнє).

Дисперсією $D(X)$ (від лат. *dispersion* – розсіювання) називається математичне сподівання квадрата відхилення значень випадкової величини від значення математичного сподівання.

Середнім квадратичним відхиленням $\sigma(X)$ називається квадратний корінь з дисперсії.

Будь-яка випадкова величина, дискретна чи неперервна, має закон розподілу. Серед законів розподілу зустрічаються такі, які характеризують велику кількість різноманітних випадкових величин. Такі закони називають *розповсюдженими*.

Серед законів розподілу неперервних величин найбільш розповсюдженим є так званий *нормальний розподіл*. Цей розподіл задається формулою:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

де a – математичне сподівання, σ – середнє квадратичне відхилення, x – довільне значення випадкової величини. Функція $f(X)$ характеризує закон зміни ймовірностей випадкової величини при зміні X на невелику величину ΔX .

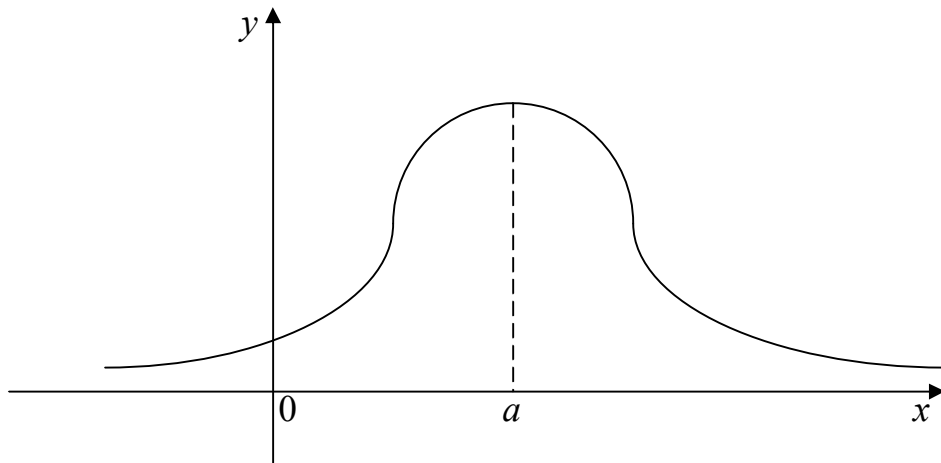


Рисунок. **Нормальний розподіл**

Закон нормального розподілу має наступну особливість: якщо існує n випадкових величин $x_1, x_2, x_3, \dots, x_n$, кожна з яких має деякий свій закон розподілу з обмеженою дисперсією, то в сумі всі ці величини утворюють нову величину, закон розподілу якої буде наближатися до нормального.

Розглянемо, наприклад, закони розподілу успішності навчальних груп (класів), якщо використовується стабільна система оцінок. Можуть бути групи, які вчаться виключно відмінно, а можуть бути групи, що вчаться вкрай незадовільно. До того, ж в одній і тій же групі закон розподілу оцінок буде різним для різних предметів. Але якщо ці всі групи зібрати разом і підрахувати розподіл кожної оцінки, то виявиться, що графік розподілу майже не буде відрізнятися від нормального розподілу. Власне, за цієї причини закон було названо нормальним.

Вказана вище особливість нормального розподілу дає можливість широкого його застосування у математичній статистиці до випадкових величин з невідомим законом розподілу. З нормального розподілу витікають розподіли "хі-квадрат", t -розподіл Стьюдента, F -розподіл Фішера-Снедекора та інші:

а) розподіл "хі-квадрат"

Якщо існує n незалежних величин $x_1, x_2, x_3, \dots, x_n$, що підпорядковуються нормальному закону розподілу і мають характеристики $a=0$ і $\sigma=1$, то випадкова величина $x^2 = x_1^2 + x_2^2 + \dots + x_n^2$ має новий закон розподілу, який називається розподілом "хі-квадрат" і має один параметр $k=n$. Параметр розподілу називається *числом степенів свободи*. Але якщо відоме ще якесь співвідношення (наприклад, $\sum x_i = n \cdot \bar{x}$), то $k = n - 1$. З ростом n числа випадкових величин розподіл повільно наближається до нормального. Така величина, як дисперсія, підкоряється розподілу "хі-квадрат".

б) t -розподіл Стьюдента

Якщо задана нормально розподілена величина x з параметрами $a=0$ і $\sigma=1$, а також незалежна від неї y , яка розподілена за законом "хі-квадрат" з параметром k , то випадкова величина $T = \frac{x}{\sqrt{\frac{y}{k}}}$ має новий розподіл, який

називається t -розподілом Стьюдента з параметром $k=n-1$. З ростом n числа випадкових величин t -розподіл досить швидко наближається до нормального, тому він широко використовується у статистиці при малих об'ємах дослідницьких значень.

в) F -розподіл Фішера-Снедекора

Якщо незалежні випадкові величини x та y розподілені за законом "хі-квадрат" з параметрами розподілу k_1 і k_2 , то випадкова величина $F = \frac{x}{y} \cdot \frac{k_1}{k_2}$ має розподіл, що називається F -розподілом Фішера-Снедекора з двома параметрами k_1 і k_2 .

Суттєвою особливістю наведених розподілів є те, що вони залежать тільки від числа степенів свободи і зовсім не залежать від параметрів випадкових величин, що досліджуються.

Питання для самоперевірки

1. Яка величина називається випадковою?
2. Що таке подія?
3. Наведіть визначення ймовірностей.
4. Які величини називають дискретними?
5. Які величини називають неперервними?
6. Що називають законом розподілу?
7. Яка особливість нормального закону розподілу?
8. Як утворюється "хі-квадрат"?
9. Що називають розподілом Стьюдента?

10. Який розподіл називається розподілом Фішера-Снедекора?

Задачі для самостійного розв'язування

1. Група з 24-х студентів складається з 15 хлопців та 9 дівчат. Яка імовірність обрати старостою дівчину?
2. Яка імовірність, що студент витягне одну з чотирьох карток з цифрами "1", "2", "3", "4"?
3. З 10 дослідних цуценят у чотирьох виявлено стоматит. Яка імовірність, що перше відібране цуценя не хворіє на стоматит?
4. Яка імовірність, що з яйця вилупиться півник або курочка, вважаючи, що ймовірність появи як півника, так і курочки 0,5?
5. Яка імовірність, що завтра наступить новий день?

II. Генеральна сукупність. Частота

Метою статистичної обробки експериментальних даних є представлення їх у формі, зручній для аналізу та проведення такого аналізу з метою виявлення певних закономірностей у процесах, стан яких відображають експериментальні дані.

У прикладному аспекті найчастіше статистична обробка даних спрямована на виявлення відмінностей між відповідними показниками та оцінку достовірності таких відмінностей. Наприклад, на оцінку різниці між дослідною та контрольною групами тварин за певною ознакою та достовірність (статистична значимість) такої різниці. Зрозуміло, що порівнюватись можуть не тільки дослідна і контрольна групи тварин, а й, наприклад, групи тварин з різних господарств тощо.

Загальна кількість об'єктів, які мають деяку спільну ознаку, називається *генеральною сукупністю*. Деяка випадковими чином відібрана частина генеральної сукупності називається *вибірковою сукупністю*, або просто *вибіркою*. Об'єкти для дослідження, що довільним чином відбираються з генеральної сукупності, утворюють *вибіркову сукупність (вибірку)*. Кількість об'єктів генеральної сукупності називають *обсягом генеральної сукупності*. Число об'єктів вибіркової сукупності утворює обсяг вибірки.

Приклад 1. При аналізі маси 200 свиней перед забоем у господарстві першим етапом, очевидно, буде вимірювання маси тварин. При цьому, як правило, проводиться вибіркоче зважування тварин, оскільки за наявності великої кількості тварин в господарстві всіх їх зважити технічно неможливо. Так, може бути проведено зважування 10 тварин. У цьому випадку 200 – це обсяг генеральної сукупності, а 10 – обсяг вибірки.

Якщо властивості об'єктів вибірки правильно відображають властивості об'єктів генеральної сукупності, то кажуть, що вибірка репрезентативна (представницька). У цьому випадку статистичні дані вважаються *репрезентативними*.

Статистичні дослідження проводять за певним планом, в якому можна виділити такі етапи:

1) формулюють завдання дослідження та визначають обсяг потрібної вибірки (мета, об'єкти вивчення, їх кількість, які ознаки беруться до уваги, які характеристики об'єктів оцінюються);

2) збирають потрібні дані, вибирають доцільну форму їх подання для подальшого дослідження, здійснюють їх наочне зображення після часткової обробки;

3) проводять остаточну обробку статистичного матеріалу та його вивчення. За результатами формулюють певні висновки.

Компактною та наочною формою подання даних, отриманих в результаті статистичного дослідження, є статистичні таблиці, графіки, стовбчасті та кругові діаграми.

Звичайно об'єкти обстежують за їх певними характеристиками. Так, у біохімічних дослідженнях встановлюються такі ознаки, як вміст гемоглобіну, лейкоцитів, еритроцитів та ін. Значення ознак можуть бути найрізноманітніші. В статистиці їх називають *варіантами* (лат. *variantis* – змінна), тобто значення ознаки – *варіанта*.

Сукупність варіант, записаних у порядку їх зростання, називають *варіаційним рядом*. Операцію запису варіант у порядку зростання називають *рангуванням*.

Приклад 2. У прикладі 1, аналізуючи живу масу свиней, отримали: $x_1; x_2; x_3; x_4; x_5; x_6; x_7; x_8; x_9; x_{10}$ – маси десяти відібраних свиней. Ці впорядковані результати є варіаційним рядом.

Коли ставиться питання про те, як часто зустрічаються ті чи інші дані, користуються терміном *частота*, розуміючи при цьому кількість даних, що трапляються. Частоту позначають буквою m .

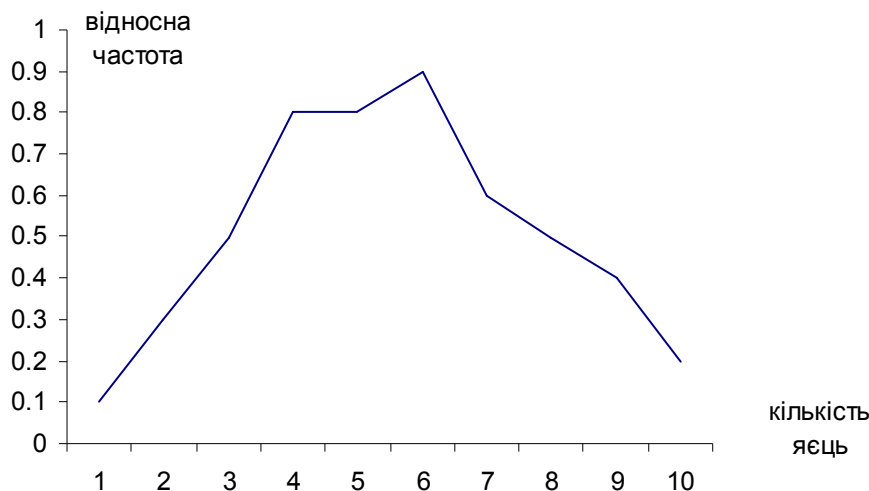
Відношення частоти (m) даних певного класу вибірки до кількості даних усієї вибірки (n) називають *відносною частотою* $\left(\frac{m}{n}\right)$ появи варіанти.

Властивість відносних частот: сума відносних частот даних будь-якої вибірки дорівнює 1. (Для уникнення помилок ця умова завжди перевіряється.)

Приклад 3. Для дослідження яйценоскості курей певного господарства проводився підрахунок знесених за добу яєць від курей, що знаходилися в клітках по 10 голів. Результати записували в таблицю:

Кількість яєць	1	2	3	4	5	6	7	8	9	10
Підрахунок випадків, m	1	3	5	8	8	9	6	5	4	2
Частота $\frac{m}{n}$	0,1	0,3	0,5	0,8	0,8	0,9	0,6	0,5	0,4	0,2

Побудуємо за результатами досліджень графік залежності відносної частоти від кількості яєць. Відкладемо по горизонтальній осі кількість яєць, а по вертикальній – відносні частоти:



Побудовану ламану лінію називають *полігоном частот*, або *многокутником розподілу*.

Статистичним розподілом випадкової величини називається залежність значень варіант від їх відносних частот.

Приклад 4. Результати 20 спостережень над кількісною варіантою зведено в таблицю:

Варіанта, X	2	6	10	12	14
Частота, m	1	5	7	3	4

Визначимо відносні частоти спостережуваної варіанти:

$$\frac{m_1}{n} = \frac{1}{20}; \quad \frac{m_2}{n} = \frac{5}{20}; \quad \frac{m_3}{n} = \frac{7}{20}; \quad \frac{m_4}{n} = \frac{3}{20}; \quad \frac{m_5}{n} = \frac{4}{20}.$$

Запишемо статистичний розподіл у вигляді таблиці:

X	2	6	10	12	14
$\frac{m}{n}$	$\frac{1}{20}$	$\frac{5}{20}$	$\frac{7}{20}$	$\frac{3}{20}$	$\frac{4}{20}$

Контроль: $\frac{1}{20} + \frac{5}{20} + \frac{7}{20} + \frac{3}{20} + \frac{4}{20} = \frac{20}{20} = 1$ (сума відносних частот даних будь-якої вибірки дорівнює одиниці).

Питання для самоперевірки

1. Що вивчає математична статистика і які завдання ставить перед собою як наука?
2. Що називають генеральною сукупністю?

3. Вибірка та її обсяг.
4. За яким планом проводять статистичні дослідження?
5. Що називають варіантою та варіаційним рядом?
6. Що називають частотою?
7. Як створити частотну таблицю?
8. Які відомі зручні представлення отриманої інформації?
9. Що називають відносною частотою? Властивість відносних частот.
10. Що називають статистичним розподілом випадкової величини?

Задачі для самостійного розв'язування

1. Для визначення задовільного стану худоби було обстежено 20 з 500 корів. Знайдіть обсяг генеральної сукупності та вибірки.

2. Отримані наступні дані про виробництво молока в 17 господарствах області:

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Надої мол. фур. корову (ц)	23	30	30	35	27	37	27	27	40	42	33	23	37	32	30	35	28

Отримати за цими даними варіаційний ряд і побудувати полігон розподілення частот. Записати статистичний розподіл у вигляді таблиці і побудувати полігон розподілення відносних частот.

3. За спостереженнями деякого собачого розплідника було встановлено, одноразовий приплід на одну самку: 3, 5, 3, 7, 5, 6, 8, 5, 4, 8, 9, 7, 6, 7, 9, 4, 2, 4 цуценят. Отримати за цими даними варіаційний ряд і побудувати полігон розподілення частот та відносних частот.

4. Результати 25 спостережень над кількісною варіантою генеральної сукупності зведено в таблицю:

Варіаційний ряд	3	6	11	13	15	16
Частота, m	2	3	6	8	4	2

Визначити статистичний розподіл вибірки.

5. Результати обстежень взятих навмання 15 свиноматок за кількістю приплоду виявилися наступними: 6; 7; 12; 11; 12; 9; 10; 11; 9; 8; 10; 8; 7; 8; 7. Побудувати за даними варіаційний ряд і частотну ламану.

III. Середня арифметична величина

Середня арифметична є здавна відомою величиною. Вона широко застосовується у науці і техніці. Немає буквально жодної біологічної роботи, де б не зустрічалася в тій чи іншій формі середня арифметична.

Середня арифметична – це узагальнююча величина, яка носить і абстрактний, і конкретний характер. Абстрактність середньої арифметичної полягає в тому, що вона може не дорівнювати за своїм числовим значенням жодному елементу вибірки. А конкретність виявляється в тому, що вона

виражається в тих же одиницях виміру, що і варіанти вибірки. Її значення є середнім між найменшим та найбільшим значеннями вибірки.

Розглянемо вибірку з n вимірів, кожен з яких позначимо $x_1, x_2, x_3, \dots, x_n$. Тоді середнє арифметичне \bar{X} значення для n вимірів:

$$\bar{X} = \frac{x_1 + x_2 + \dots + x_n}{n},$$

де n – це кількість вимірів певної ознаки.

Середня величина характеризує групу в цілому і вона більш об'єктивно відбиває аналізований показник, ніж будь-який один вимір (маса однієї тварини). Середнє значення величини X позначають буквою \bar{X} (для будь якої ознаки).

Приклад.

Було зроблено 5 вимірів вмісту каротину у крові телят (в мкг/100мл): 21,27; 21,36; 21,09; 21,16; 21,47. Обчислити середнє арифметичне:

$$\bar{X} = \frac{21,27 + 21,36 + 21,09 + 21,16 + 21,47}{5} = 21,27.$$

Отже, середнє арифметичне значення вмісту каротину у крові телят дорівнює 21,27.

Питання для самоперевірки

1. Що називають середньою арифметичною величиною?
2. Як обчислити середню арифметичну величину?
3. У чому полягає абстрактність та узагальненість середньої арифметичної величини?

Задачі для самостійного розв'язування

1. Було зроблено 5 вимірів вмісту селену в крові (в умовних одиницях): 17,24; 21,66; 19,39; 21,06; 18,47. Записати варіаційний ряд та обчислити середнє арифметичне.
2. Було зроблено 7 вимірів маси цуценят одного віку певної породи (в кг): 5,62; 3,44; 5,17; 4,71; 4,26; 5,79; 3,41. Записати варіаційний ряд та обчислити середнє арифметичне.
3. Було зроблено 7 вимірів вмісту вітаміну А у корів в літній період (в мкг/100мл): 66; 72; 132; 101; 94, 71, 83. Записати варіаційний ряд та обчислити середнє арифметичне.
4. Було зроблено 10 вимірів загального білка в сиворотці крові кроликів (в г/100мл): 7,24; 6,36; 7,99; 7,06; 6,47; 7,04; 7,78; 8,00; 6,98; 7,33. Записати варіаційний ряд та обчислити середнє арифметичне.
5. Було зроблено 5 вимірів відносної густини крові собак (в кг/см³): 1,024; 1,056; 1,049; 1,046; 1,047. Записати варіаційний ряд та обчислити середнє арифметичне.

IV. Середнє квадратичне відхилення. Дисперсія

Знаючи середнє арифметичне значення даних експерименту, виникає питання: як обчислити середню величину, на яку відрізняються дані від середнього арифметичного?

Різницю між будь-яким виміром з вибірки і середнім арифметичним цієї ж вибірки називають *відхиленням* варіанти x_i від \bar{X} : $x_i - \bar{X}$.

Якщо обчислити відхилення для усіх варіант, то серед отриманих значень будуть від'ємні і додатні, які у сумі даватимуть 0, тобто взаємно компенсуються. Це означає, що беззмістовно обчислювати середнє відхилення як середнє арифметичне відхилень. Для того, щоб уникнути компенсації додатних і від'ємних значень, існує декілька способів. Найпоширеніший – піднесення кожної різниці $(x_i - \bar{X})$ до квадрату. (Квадрати як від'ємних, так і додатних величин є величинами додатними.) Додаючи квадрати усіх різниць і ділячи на кількість цих різниць, отримаємо величину, яка називається *дисперсією*. Фактично вона показує *середнє арифметичне квадратів відхилень*. Для того, щоб позбутися квадрату величини, обчислюємо корінь квадратний з дисперсії. Отримане значення називається *середнім квадратичним відхиленням*. Розрізняють формули середнього квадратичного відхилення для генеральної і вибіркової сукупностей.

При існуючих даних генеральної сукупності використовують таку формулу:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}},$$

де X_i – значення i -тої варіанти, $i=1, \dots, n$, \bar{X} – середнє арифметичне, n – об'єм генеральної сукупності.

Оскільки $D(X) = \sigma^2$, то $D(X) = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}$ – зміщена дисперсія.

Якщо ж є тільки дані вибірки, то для визначення середнього квадратичного відхилення застосовується така формула:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1}}.$$

А дисперсія, що називають незміщеною, відповідно:

$$D(X) = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1},$$

де X_i – значення i -тої варіанти, $i=1, \dots, n$, \bar{X} – середнє арифметичне, n – об'єм вибіркової сукупності.

Приклад.

Було зроблено 5 вимірів вмісту кальцію в крові (в умовних одиницях): 11,27; 11,36; 11,09; 11,16; 11,47. Обчислити середнє квадратичне відхилення та

дисперсію вмісту кальцію у крові (5 даних вимірів розглядаємо як вибірку сукупність):

$$\bar{X} = \frac{11,27 + 11,36 + 11,09 + 11,16 + 11,47}{5} = \frac{56,35}{5} = 11,27.$$

Отже, середнє арифметичне значення вмісту кальцію у крові дорівнює 11,27.

Тоді середнє квадратичне відхилення:

$$\sigma = \sqrt{\frac{(11,27 - 11,27)^2 + (11,36 - 11,27)^2 + (11,09 - 11,27)^2 + (11,16 - 11,27)^2 + (11,47 - 11,27)^2}{5 - 1}} \approx 0,152.$$

А дисперсія $D(X) = \sigma^2 = 0,152^2 \approx 0,023$.

Питання для самоперевірки

1. Що називають відхиленням варіанти від середнього арифметичного?
2. Чому неможливо обчислити середнє відхилення як середнє арифметичне відхилень?
3. Що називають дисперсією?
4. Що показує дисперсія?
5. Формула для обчислення дисперсії при існуючих даних генеральної сукупності.
6. Формула для обчислення дисперсії, якщо відомі лише дані вибірки.

Задачі для самостійного розв'язування

1. Було зроблено 7 вимірів глюкози в крові курей (ммоль/л): 7,77; 5,98; 4,67; 7,42; 6,66; 7,43; 4,99. Обчислити середнє квадратичне відхилення вмісту глюкози в крові птиці.
2. Було зроблено 7 вимірів вмісту вітаміну А у сировотці крові свиноматок (в мкг/100мл): 66; 72; 92; 100; 94, 79, 93. Записати варіаційний ряд, обчислити середнє арифметичне та середнє квадратичне відхилення вмісту вітаміну А в сировотці крові свиноматок.
3. Було зроблено 7 вимірів маси тіла дворічних бугаїв-плідників симентальської породи (кг): 497,7; 481,4; 480,9; 488,1; 501,6; 500,9; 499,1. Записати варіаційний ряд, обчислити середнє арифметичне та середнє квадратичне відхилення вмісту вітаміну А в сировотці крові свиноматок.
4. Було зроблено 7 вимірів вмісту жиру в молоці першої лактації корів шведської породи (%): 3,78; 3,63; 3,68; 3,70; 3,79; 3,85; 4,05. Обчислити середнє арифметичне та середнє квадратичне відхилення вмісту жиру.
5. Було зроблено 10 вимірів маси яєць однієї курки (в г): 56, 54, 51, 54, 59, 54, 55, 55, 56, 57. Обчислити середнє арифметичне та середнє квадратичне відхилення маси яєць.

V. Визначення помилки репрезентативності

Вивчаючи певну ознаку, неможливо дослідити усі об'єкти генеральної сукупності тому, що вона, як правило, є дуже численною, можливо навіть складається з нескінченно великого числа членів.

Тому робиться вибірка об'єктів, які і досліджуються. При цьому постає таке питання: чи можливо за результатами, отриманими при вивченні вибірки, робити висновки про всю генеральну сукупність?

Характеризуючи цілу сукупність лише за її частиною, неможливо уникнути помилок, які називаються *помилками репрезентативності*. Навіть за ідеальної організації дослідницької роботи з'являються помилки такого типу.

Помилка репрезентативності середньої арифметичної залежить від двох величин: від різноманітності ознаки у генеральній сукупності і від чисельності вибірки. Чим менша степінь різноманітності (на її величину вказує середнє квадратичне відхилення) і чим більша кількість вибраних для дослідження об'єктів, тим меншою є величина помилки репрезентативності вибіркового середнього арифметичного.

Для розрахунку величини помилки використовується формула:

$$\Delta = \frac{\sigma}{\sqrt{n}},$$

де $\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1}}$, X_i – значення i -тої варіанти, $i=1, \dots, n$, \bar{X} – середнє арифметичне вибірки, n – об'єм вибіркової сукупності.

Тепер будь-яку ознаку можна представити у формі:

$$x = \bar{X} \pm \Delta.$$

Приклад.

Було зроблено 5 вимірів вмісту кетонівих тіл в крові (в умовних одиницях): 13,27; 14,36; 14,19; 13,46; 13,47. Обчислити помилку середнього арифметичного.

Обчислимо середнє арифметичне вмісту кетонівих тіл у крові:

$$\bar{X} = \frac{13,27 + 14,36 + 14,19 + 13,46 + 13,47}{5} = 13,75.$$

Обчислимо середнє квадратичне відхилення:

$$\sigma = \sqrt{\frac{(13,27 - 13,75)^2 + (14,36 - 13,75)^2 + (14,19 - 13,75)^2 + (13,46 - 13,75)^2 + (13,47 - 13,75)^2}{5-1}} \approx 0,494.$$

$$\text{Тоді: } \Delta = \frac{0,494}{\sqrt{5}} \approx 0,221.$$

Отже, вміст кетонівих тіл у крові можна представити у вигляді: $13,75 \pm 0,221$.

Питання для самоперевірки

1. Від чого залежить помилка репрезентативності?
2. Якою формулою користуються для розрахунку величини помилки репрезентативності?
3. Як можна представити будь-яку ознаку?

Задачі для самостійного розв'язування

1. Було зроблено 7 вимірів глюкози в крові телят (ммоль/л): 2,77; 3,98; 3,67; 2,42; 3,66; 2,43; 2,99. Обчислити помилку середньої арифметичної.

2. Було зроблено 7 вимірів вмісту вітаміну Е у сиворотці крові свиноматок (в мкг/100мл): 96; 82; 102; 100; 94, 99, 98. Обчислити помилку середньої арифметичної.

3. Було зроблено 7 вимірів маси тіла п'ятирічних бугаїв-плідників симентальської породи (кг): 97,7; 981,4; 980,9; 988,1; 901,6; 900,9; 999,1. Обчислити помилку середньої арифметичної.

4. Було зроблено 10 вимірів довжини (в мм) яєць однієї курки: 58, 60, 61, 54, 59, 57, 55, 56, 56, 57. Обчислити помилку середньої арифметичної.

5. Було зроблено 7 вимірів вмісту білка в молоці першої лактації корів шведської породи (%): 3,28; 3,10; 3,18; 3,21; 3,29; 3,32; 3,45. Обчислити помилку середньої арифметичної.

VI. Оцінка достовірності різниць між групами за критерієм Стьюдента

Якщо аналізу піддаються дві (чи більше) групи тварин, виникає питання оцінки різниці між групами за аналізованим показником. Наприклад, у першому господарстві жива маса свиней перед забоєм склала $115,7 \pm 1,2$ кг, а в другому господарстві $125,5 \pm 2,5$ кг. Порівняємо середні значення в групах, визначивши цю різницю у відсотках. Ми виявляємо, що у другому господарстві маса свиней на 8,5% більша, ніж у першому (якщо показник першого господарства взяти за 100%). Але залишається питання, чи є ця різниця статистично достовірною, тобто є не випадковою, а закономірною за виявлених характеристик груп.

Найбільш поширеною є оцінка достовірності різниць між групами за методом Стьюдента. Для цього визначається критерій достовірності різниці за формулою:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\Delta_1^2 + \Delta_2^2}},$$

де \bar{X}_1 і \bar{X}_2 – середні значення в групах, Δ_1 і Δ_2 – помилки репрезентативності в групах.

Обчислений критерій t порівнюється зі стандартним (табличним) значенням критерія Стьюдента t_{st} для $n = n_1 + n_2 - 2$, де n_1 та n_2 – кількість вимірів у групах (див. у табл. 1 додатку).

Якщо вирахований критерій t більший стандартного значення критерію Стьюдента t_{st} для $p < 0,05$, це означає, що різниця між групами є достовірною з надійністю 95% (тобто різницю можна очікувати у 95 випадках із 100). Якщо $t < t_{st}$ для $p < 0,01$, різниця достовірна з надійністю 99 %, якщо $t < t_{st}$ для $p < 0,001$, різниця достовірна з максимальною надійністю 99,9 %.

У випадку, якщо $t > t_{st}$ – це означає, що за цією різницею між групами не можна зробити висновок про наявність чи відсутності достовірної різниці між групами (потрібні додаткові дослідження).

Приклад. У досліді перевіряли дію добрив на урожайність картоплі (ц/га):

Контрольна група	121	143	145	149	150
Дослідна група	170	180	188	196	187

Визначити оцінку достовірності за критерієм Стьюдента.

Знайдемо середні арифметичні та середні квадратичні відхилення урожайності картоплі обох груп. Для першої групи:

$$\bar{X}_1 = \frac{121+143+145+149+150}{5} = 141,6.$$

Обчислимо середнє квадратичне відхилення:

$$\sigma_1 = \sqrt{\frac{(121-141,6)^2 + (143-141,6)^2 + (145-141,6)^2 + (149-141,6)^2 + (150-141,6)^2}{5-1}} \approx 11,87.$$

$$\text{Тоді: } \Delta_1 = \frac{11,87}{\sqrt{5}} \approx 5,31.$$

Для другої групи:

$$\bar{X}_2 = \frac{170+180+188+196+187}{5} = 184,2.$$

Обчислимо середнє квадратичне відхилення:

$$\sigma_2 = \sqrt{\frac{(170-184,2)^2 + (180-184,2)^2 + (188-184,2)^2 + (196-184,2)^2 + (187-184,2)^2}{5-1}} \approx 9,76.$$

$$\text{Тоді: } \Delta_2 = \frac{9,76}{\sqrt{5}} \approx 4,36.$$

Отже, критерій достовірності різниці дорівнює:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\Delta_1^2 + \Delta_2^2}} = \frac{141,6 - 184,2}{\sqrt{5,31^2 + 4,36^2}} = \frac{42,6}{\sqrt{47,2057}} \approx 6,2.$$

Обчислимо n : $n = n_1 + n_2 - 2 = 5 + 5 - 2 = 8$.

Знайдемо t_{st} за таблицею 1 додатку для $n=8$: $t_{st} = \{2,3 - 3,4 - 5,0\}$. Оскільки $t < t_{st}$ для $p < 0,001$, різниця достовірна з максимальною надійністю 99,9 %.

Питання для самоперевірки

1. Що показує метод Стьюдента?
2. За якою формулою визначається критерій достовірності різниці?
3. Коли можна говорити про достовірну різницю між групами?
4. Коли різниця між групами не достовірна?

Задачі для самостійного розв'язування

1. Активність каталази (мккат) у тканинах печінки добових перепелів м'ясної та яєчної порід представлена в наступній таблиці:

М'ясна порода	0,94	0,93	0,92	0,93	0,95	0,95
Яєчна порода	0,89	0,88	0,88	0,86	0,87	0,89

Визначити оцінку достовірності за критерієм Стьюдента.

2. У таблиці представлені дані урожайності цукрового буряка (ц/га) двох різних господарств:

I господарство	221	243	245	248	249	250
II господарство	309	280	288	306	287	309

Визначити оцінку достовірності за критерієм Стьюдента.

3. У таблиці представлені дані перекисного окиснення ліпідів (мкМ/л) у тканинах мозку 15-денних ембріонів курей різних порід:

I порода	0,86	0,88	0,85	0,87	0,89
II порода	0,78	0,76	0,79	0,78	0,81

Визначити оцінку достовірності за критерієм Стьюдента.

4. У таблиці представлено висоту в холці (см) двох різних порід собак одного віку:

I порода	50	53	55	56	56
II порода	57	60	62	62	63

Визначити оцінку достовірності за критерієм Стьюдента.

5. У таблиці представлено швидкість зсідання еритроцитів свиней та собак:

Свині	2,2	4,8	5,2	5,6	5,7
Собаки	2,7	6,0	5,6	4,2	4,3

Визначити оцінку достовірності за критерієм Стьюдента.

VII. Оцінка достовірності різниць між групами за критерієм Фішера

Визначити достовірність різниці між групами можна за критерієм Фішера за формулою:

$$F = \frac{(\bar{X}_1 - \bar{X}_2)^2}{\sigma_z^2} \cdot \frac{n_1 n_2}{n_1 + n_2} \geq F_{st} \left\{ \begin{array}{l} v_1 = 1 \\ v_2 = n_1 + n_2 - 2 \end{array} \right\},$$

де F – критерій Фішера; \bar{X}_1 і \bar{X}_2 – середні арифметичні;
 σ_z – випадкова варіанса, яку обраховують за формулою:

$$\sigma_z^2 = \frac{n_1(n_1 - 1)\Delta_1^2 + n_2(n_2 - 1)\Delta_2^2}{n_1 + n_2 - 2};$$

n_1, n_2 – кількість вимірів (тварин) у групах.

Вирахований критерій F порівнюється зі стандартним критерієм Фішера F_{st} (табл. 2 додатку) для трьох значень рівня надійності 0,95, 0,99, 0,999, як і у попередньому методі.

Приклад. Порівнюючи масу (кг) дорослих індиків двох порід після однакової відгодівлі, отримані наступні показники:

I порода	4,1	3,9	4,3	3,8	3,8	4,0
II порода	4,3	4,4	4,3	4,5	4,6	4,5

Визначити оцінку достовірності за критерієм Фішера.

Знайдемо середні арифметичні та середні квадратичні відхилення мас індиків обох порід. Для першої породи:

$$\bar{X}_1 = \frac{4,1 + 3,9 + 4,3 + 3,8 + 3,8 + 4,0}{6} = 3,98.$$

Обчислимо середнє квадратичне відхилення:

$$\sigma_1 = \sqrt{\frac{(3,98 - 4,1)^2 + (3,98 - 3,9)^2 + (3,98 - 4,3)^2 + (3,98 - 3,8)^2 + (3,98 - 3,8)^2 + (3,98 - 4,0)^2}{6 - 1}} \approx 0,194.$$

$$\text{Тоді: } \Delta_1 = \frac{\sigma_1}{\sqrt{6}} = \frac{0,194}{\sqrt{6}} \approx 0,079.$$

Для другої групи:

$$\bar{X}_2 = \frac{4,3 + 4,4 + 4,3 + 4,5 + 4,6 + 4,5}{6} = 4,43.$$

Обчислимо середнє квадратичне відхилення:

$$\sigma_2 = \sqrt{\frac{(4,43 - 4,3)^2 + (4,43 - 4,4)^2 + (4,43 - 4,3)^2 + (4,43 - 4,5)^2 + (4,43 - 4,6)^2 + (4,43 - 4,5)^2}{6 - 1}} \approx 0,121.$$

$$\text{Тоді: } \Delta_2 = \frac{\sigma_2}{\sqrt{6}} = \frac{0,121}{\sqrt{6}} \approx 0,049.$$

Обрахуємо випадкову варіанту σ_z за формулою:

$$\sigma_z^2 = \frac{n_1(n_1 - 1)m_1^2 + n_2(n_2 - 1)m_2^2}{n_1 + n_2 - 2} = \frac{6 \cdot 5 \cdot 0,079^2 + 6 \cdot 5 \cdot 0,049^2}{6 + 6 - 2} \approx 0,012.$$

$$\text{Тоді, } F = \frac{(\bar{X}_1 - \bar{X}_2)^2}{\sigma_z^2} \cdot \frac{n_1 n_2}{n_1 + n_2} = \frac{(3,98 - 4,43)^2}{0,012} \cdot \frac{6 \cdot 6}{6 + 6} = 50,6.$$

$$\text{Обчислимо } \nu_2: \nu_2 = n_1 + n_2 - 2 = 5 + 5 - 2 = 8.$$

Знайдемо F_{st} (табл. 3 додатку) для $\begin{cases} \nu_1 = 1 \\ \nu_2 = 10 \end{cases}$: $F_{st} = \{25,0 - 10 - 21,0\}$. Отже

різниця достовірна з максимальною надійністю 99,9 %.

Питання для самоперевірки

1. За якою формулою визначається достовірність різниці за критерієм Фішера?
2. Як обрахувати випадкову варіансу?
3. Коли можна говорити про достовірну різницю між групами?
4. Коли різниця між групами не є достовірною?

Задачі для самостійного розв'язування

1. Активність каталази (мккат) у тканинах печінки добових перепелів м'ясної та ячної порід представлено в наступній таблиці:

М'ясна порода	0,87	0,80	0,82	0,81	0,83	0,84
Ячна порода	0,79	0,78	0,78	0,76	0,77	0,7

Визначити оцінку достовірності за критерієм Фішера.

2. У таблиці представлені дані урожайності цукрового буряка (ц/га) двох різних господарств:

I господарство	291	283	295	288	289	290
II господарство	309	300	289	306	287	309

Визначити оцінку достовірності за критерієм Фішера.

3. У таблиці представлені дані перекисного окиснення ліпідів (мкМ/л) у тканинах мозку 15-денних ембріонів курей різних порід:

I порода	0,76	0,78	0,75	0,77	0,79
II порода	0,79	0,82	0,79	0,81	0,81

Визначити оцінку достовірності за критерієм Фішера.

4. У таблиці представлено висоту в холці (см) двох різних порід собак одного віку:

--	--	--	--	--	--

I порода	60	63	65	66	66
II порода	67	69	64	65	63

Визначити оцінку достовірності за критерієм Фішера.

5. У таблиці представлено швидкість зсідання еритроцитів свиней та собак:

Свині	2,2	4,8	5,2	5,6	5,7
Собаки	3,7	6,0	5,6	4,2	4,3

Визначити оцінку достовірності за критерієм Фішера.

VIII. Оцінка достовірності різниць між групами за методом кутів (φ -метод)

Якщо показник групи визначається у відсотках (наприклад, частка хворих тварин у групі), то для оцінки достовірності різниці між групами неможливо застосувати попередні критерії. (Наприклад, при порівнянні двох груп тварин по 100 голів у кожній, у першій виявлено 15 інфікованих тварин, у другій – 25 голів – 15% та 25%, відповідно.) У цьому випадку використовують метод φ , що дозволяє оцінити достовірність різниць між частками груп.

Метод визначення вірогідності різниці часток (відсотка) появи ознаки у вибіркових сукупностях, запропонований англійським статистиком Р.Е.Фішером (1890–1962). Метод може бути застосований за будь-яких значень часток, але найчастіше ним користуються, якщо $p < 0,2$ або $p > 0,8$.

Фішер показав, що визначити вірогідність різниці часток точніше і простіше, якщо замість кожної частки взяти кут, синус якого дорівнює кореню квадратному з цієї частки. Тоді частки перетворюються в кути φ за формулою:

$$\varphi = 0,0349 \arcsin \sqrt{p}, \text{ де } p - \text{ частка, } \varphi - \text{ кут у радіанах.}$$

Для полегшення обрахунків створена таблиця 2 додатку, де представлено кути φ (φ_1) в радіанах від частки.

Вірогідність різниці часток методом φ визначається за допомогою такої формули: $F = (\varphi_1 - \varphi_2)^2 \cdot \frac{n_1 n_2}{n_1 + n_2} \geq F_{st} \left\{ \begin{array}{l} \nu_1 = 1 \\ \nu_2 = n_1 + n_2 - 2 \end{array} \right\}$,

де F – критерій Фішера; n_1, n_2 – кількість вимірів (тварин) у групах.

Обрахований критерій F порівнюється зі стандартним критерієм Фішера (табл. 3) за трьома рівнями надійності, як і у попередньому методі.

Приклад. 3 1000 курчат, що отримували опромінені дріжджі, хворіло рахітом 10, а з 2000 курчат, що не отримували опромінені дріжджі, хворіло 80. Визначити ефективність добавки у корм курчатам опромінених дріжджів.

Тоді, $n_1=1000$; $a_1=10$; $p_1 = \frac{a_1}{n_1} = 0,01$; $\varphi_1=0,2003$ (кут визначаємо за табл. 2 додатку). Відповідно $n_2=2000$; $a_2=80$; $p_2 = \frac{a_2}{n_2} = 0,04$; $\varphi_2=0,4027$.

Вірогідність різниці часток:

$$F = (\varphi_1 - \varphi_2)^2 \cdot \frac{n_1 n_2}{n_1 + n_2} = (0,2003 - 0,4027)^2 \cdot \frac{1000 \cdot 2000}{1000 + 2000} = 41,0.$$

$$F_{st} = \{3,8 - 6,6 - 10,8\} \text{ для } v_1=1, v_2=1000+2000-2=\infty.$$

Перевірка показала високу імовірність того, що вся генеральна сукупність курчат, що отримувала опромінені дріжджі, значно менше вразлива рахітом порівняно з генеральною сукупністю контрольних курчат.

Питання для самоперевірки

1. Коли використовують метод φ ?
2. Як визначити кут φ ?
3. За допомогою якої формули визначають достовірність різниці частин?

Задачі для самостійного розв'язування

1. При дослідженні розповсюдження крипторхізму в приплоді різних плідників було виявлено, що серед 20 нащадків барана А крипторхів виявилось 5, а у барана В з 30 синів крипторхів – 6. Встановити достовірність різниці.
2. При дослідженні впровадження нової вакцини для кролів було встановлено, що з 40 тварин дослідної групи не стійкими до міксоматозу виявилось дві особини, а з 100 тварин контрольної 21 кроль захворів на міксоматоз. Визначити ефективність нової вакцини.
3. Досліджуючи ефективність нової добавки на яйценоскість курей, було встановлено, що з 30 дослідних курочок на певний період вже неслися 25, а з 50 контрольних – 34. Визначити ефективність нової добавки.
4. При утриманні гусей в нових умовах було встановлено, що з 50 дослідних гусей стійкими до інфекційних захворювань виявилось 94%, а з 50 контрольних – 82%. Визначити ефективність нових умов утримання?
5. З 500 курчат, що отримували вітамінізовані добавки, хворіло рахітом 5, а з 1000 курчат, що не отримували вітамінізовані добавки, хворіло 50. Визначити ефективність добавки в корм курчатам опромінених дрожжей.

ІХ. Кореляція. Коефіцієнт кореляції

Кореляція (від лат. *correlation* – відповідність) – статистична залежність між величинами, яка не має, взагалі кажучи, строго функціонального характеру. Кореляційна залежність виникає тоді, коли одна з величин залежить не тільки від заданої другої, а й від деяких випадкових факторів; або, коли серед умов, від яких залежать обидві величини, є загальні для них обох.

Кореляційний зв'язок – це не точна залежність однієї величини від іншої. Числовим значенням однієї змінної ставиться у відповідність середнє декількох значень інших. Наприклад, між кількістю внесених на поле добрив і

врожайністю пшениці існує незаперечна залежність. Але це не означає, що конкретній кількості добрив відповідає визначена величина урожаю. На урожай впливає багато інших факторів: склад і структура ґрунту, різні методи посіву і таке інше.

Кореляційний зв'язок виявляється у середньому для усієї сукупності спостережень. По відношенню ж до окремих спостережень цей зв'язок є дуже неповним і неточним. Відомо, наприклад, що існує кореляція між вагою тварини і її висотою. Це означає, що більш високі тварини звичайно важчі за низьких. Та в деяких випадках низька тварина може виявитися важчою за високу.

Кореляційний зв'язок може мати різну степінь – від повної незалежності до функціональної залежності. Крім того, характер зв'язку між різними величинами може бути різний. Тому виникає необхідність визначити форму, напрям і степінь кореляційних зв'язків.

За формою кореляція може бути *прямолінійною* і *криволінійною*, за напрямком – *прямою* і *оберненою*.

При додатній кореляції залежність між величинами буде прямою: при збільшенні однієї величини, збільшується й інша. При від'ємній кореляції залежність обернена: збільшення однієї величини пов'язано зі зменшенням другої. Степінь кореляції вимірюється різними показниками зв'язку. Такими показниками є коефіцієнт кореляції, кореляційне відношення та ін.

Кореляційне відношення визначається як середнє значення добутків відхилень величин X та Y від їх середніх:

$$\mu = \frac{\sum_{i=1}^n (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{n}.$$

Числове значення μ не може бути мірою тісноти зв'язку, бо залежить від одиниць вимірювання величини X та Y . Тому вводиться поняття коефіцієнта кореляції:

$$r = \frac{\mu}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^n (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}},$$

де X_i, Y_i – числові значення величин, між якими встановлюється кореляційний зв'язок, \bar{X}, \bar{Y} – їх середні арифметичні значення величин.

Для незалежних величин $r=0$, для функціональних залежностей $r=\pm 1$. Якщо зростання X призводить до зростання Y , то r – додатне, якщо до зменшення – то r – від'ємне. Щоб нехтувати знаком r , вводять поняття коефіцієнта детермінації $R = r^2$, який завжди додатний.

Кореляція вважається сильною, якщо $r < 0,375$.

Приклад. Дві випадково розподілені величини X і Y . Знайти коефіцієнт кореляції:

X	40	40	39	40	41	38	42	40	42	38
Y	35	37	35	36	35	36	37	36	38	35

Обчислимо середні значення \bar{X} і \bar{Y} :

$$\bar{X} = \frac{40 + 40 + 39 + 40 + 41 + 38 + 42 + 40 + 42 + 38}{10} = 40;$$

$$\bar{Y} = \frac{35 + 37 + 35 + 36 + 35 + 36 + 37 + 36 + 38 + 35}{10} = 36.$$

Для обчислення r необхідно, крім \bar{X} і \bar{Y} , визначити наступні величини: $(X_i - \bar{X})$, $(Y_i - \bar{Y})$, $(X_i - \bar{X}) \cdot (Y_i - \bar{Y})$, $(X_i - \bar{X})^2$, $(Y_i - \bar{Y})^2$. Тому результати обчислень зведемо в таблицю:

X	Y	$(X_i - \bar{X})$	$(Y_i - \bar{Y})$	$(X_i - \bar{X}) \cdot (Y_i - \bar{Y})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$
40	35	0	-1	0	0	1
40	37	0	1	0	0	1
39	35	-1	-1	1	1	1
40	36	0	0	0	0	0
41	35	1	-1	1	1	1
38	36	-2	0	2	4	1
42	37	2	1	0	0	0
40	36	0	0	0	4	0
42	38	2	2	4	4	4
38	35	-2	-1	2	4	1
$\bar{X}=40$	$\bar{Y}=36$	$\sum=0$	$\sum=0$	$\sum=8$	$\sum=18$	$\sum=10$

З таблиці легко отримати:

$$\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) = 8 \sum_{i=1}^n (X_i - \bar{X})^2 = 18; \sum_{i=1}^n (Y_i - \bar{Y}) = 10.$$

$$\text{Тоді } r = \frac{8}{\sqrt{18 \cdot 10}} \approx 0,596.$$

Питання для самоперевірки

1. Коли виникає кореляційна залежність?
2. Що виражає кореляційний зв'язок?
3. Що показує кореляція?
4. Яка кореляція може бути за формою і за напрямком?
5. За якою формулою обчислюється кореляція?

Задачі для самостійного розв'язування

1. У таблиці представлено довжини X в міліметрах і маса Y в грамах 10 штук яєць однієї курки:

X	60	58	57	55	56	58	55	57	55	59
Y	56	53	4	51	54	59	55	55	56	57

Обчислити коефіцієнт кореляції.

2. У таблиці представлені дані щодо вмісту жиру (X) і білка (Y) в молоці першої лактації корів швейцарської породи:

X	3,78	3,63	3,68	3,65	3,70	3,79	3,85	3,92	4,05	4,06
Y	3,24	3,10	3,18	3,21	3,25	3,32	3,39	3,40	3,47	3,56

Обчислити коефіцієнт кореляції.

3. У таблиці представлені результати вимірювання діаметра сосни X в сантиметрах і її висоти Y в метрах:

X	15	20	25	30	35	40	45
Y	18	19	20	21	22	23	24

Обчислити коефіцієнт кореляції.

4. Обчислити коефіцієнт кореляції між показником гематокрита і середнім об'ємом еритроцитів у собак:

X	42,0	42,3	43,1	43,5	45,7	45,9	46,8
Y	68	69	68	69	70	72	73

5. Обчислити коефіцієнт кореляції між масою (X) і довжиною тіла (Y) собаки:

X	49	50	52	54	55	56	57
Y	70	81	94	102	120	145	160

ДОДАТКИ

Таблиця 1

Стандартні значення критерію Стьюдента (t)

n	$p_1 = 0,05$	$p_2 = 0,01$	$p_3 = 0,001$	n	$p_1 = 0,05$	$p_2 = 0,01$	$p_3 = 0,001$
1	12,7	63,7	637,0	13	2,2	3,0	4,1
2	4,3	9,9	31,6	14 – 15	2,1	3,0	4,1
3	3,2	5,8	12,9	16 – 17	2,1	2,9	4,0
4	2,8	4,6	8,6	18 – 20	2,1	2,9	3,9
5	2,6	4,0	6,9	21 – 24	2,1	2,8	3,8
6	2,4	3,7	6,0	25 – 28	2,1	2,8	3,7
7	2,4	3,5	5,3	29 – 30	2,0	2,8	3,7
8	2,3	3,4	5,0	31 – 34	2,0	2,7	3,7
9	2,3	3,3	4,8	35 – 42	2,0	2,7	3,6
10	2,2	3,2	4,6	43 – 62	2,0	2,7	3,5
11	2,2	3,1	4,4	63 – 175	2,0	2,6	3,4
12	2,2	3,1	4,2	176 – ∞	2,0	2,6	3,3

Таблиця 2

Кути φ (фі) в радіанах: $\varphi = 2 \times \frac{\pi}{180} \arcsin \sqrt{p}$

p	Останні цифри значення частки										p%
	0	1	2	3	4	5	6	7	8	9	
1	2	3	4	5	6	7	8	9	10	11	12
0,01	0,2003	0,2102	0,2195	0,2285	0,2372	0,2456	0,2587	0,2615	0,2691	0,2765	1
0,02	0,2838	0,2909	0,2978	0,3045	0,3111	0,3176	0,3239	0,3301	0,3363	0,3423	2
0,03	0,3482	0,3540	0,3597	0,3654	0,3700	0,3764	0,3818	0,3871	0,3924	0,3976	3
0,04	0,4027	0,4078	0,4128	0,4178	0,4227	0,4275	0,4323	0,4371	0,4418	0,4464	4
0,05	0,4510	0,4556	0,4601	0,4646	0,4690	0,4735	0,4778	0,4822	0,4865	0,4907	5
0,06	0,4949	0,4991	0,5033	0,5074	0,5115	0,5156	0,5196	0,5237	0,5276	0,5316	6
0,07	0,5355	0,5394	0,5433	0,5472	0,5510	0,5548	0,5586	0,5624	0,5661	0,5698	7
0,08	0,5735	0,5772	0,5808	0,5845	0,5881	0,5917	0,5953	0,5988	0,6024	0,6059	8
0,09	0,6094	0,6129	0,6163	0,6198	0,6232	0,6267	0,6301	0,6334	0,6368	0,6402	9
0,10	0,6435	0,6468	0,6501	0,6534	0,6567	0,6600	0,6632	0,6665	0,6697	0,6729	10
0,11	0,6761	0,6793	0,6825	0,6857	0,6888	0,6920	0,6951	0,6982	0,7013	0,7044	11
0,12	0,7075	0,7107	0,7136	0,7167	0,7197	0,7228	0,7258	0,7288	0,7318	0,7348	12
0,13	0,7377	0,7407	0,7437	0,7466	0,749	0,7525	0,7554	0,7583	0,7612	0,7641	13
0,14	0,7670	0,7699	0,7727	0,7756	0,7785	0,7813	0,7841	0,7800	0,7898	0,7926	14
0,15	0,7954	0,7982	0,8010	0,8038	0,8065	0,8093	0,8121	0,8148	0,8176	0,8203	15
0,16	0,8230	0,8258	0,8285	0,8312	0,8339	0,8366	0,8393	0,8420	0,8446	0,8473	16
0,17	0,8500	0,8526	0,8553	0,8579	0,8606	0,8632	0,8658	0,8685	0,8711	0,8737	17
0,18	0,8763	0,8789	0,8815	0,8841	0,8867	0,8892	0,8918	0,8944	0,8970	0,8995	18
0,19	0,9021	0,9046	0,9071	0,9097	0,9122	0,9147	0,9173	0,9198	0,9223	0,9248	19
0,20	0,9273	0,9298	0,9323	0,9348	0,9373	0,9397	0,9422	0,9447	0,9472	0,9496	20

1	2	3	4	5	6	7	8	9	10	11	12
0,21	0,9521	0,9545	0,9570	0,9594	0,9619	0,9643	0,9667	0,9692	0,9716	0,9740	21
0,22	0,9764	0,9788	0,9813	0,9836	0,9860	0,9884	0,9908	0,9932	0,9956	0,9980	22
0,23	1,0004	1,0027	1,0051	1,0075	1,0098	1,0122	1,0146	1,0169	1,0193	1,0216	23
0,24	1,0240	1,0262	1,0286	1,0310	1,0333	1,0356	1,0379	1,0403	1,0426	1,0449	24
0,25	1,0472	0,0495	1,0518	1,0541	1,0564	1,0587	1,0610	1,0633	1,0656	1,0679	25
0,26	1,0701	0,0724	1,0747	1,0770	1,0793	1,0815	1,0838	1,0860	1,0883	1,0906	26
0,27	1,0928	0,0951	1,0973	1,0996	1,1018	1,1040	1,1063	1,1085	1,1107	1,1130	27
0,28	1,1152	0,1174	1,1197	1,1219	1,1241	1,1263	1,1285	1,1307	1,1329	1,1352	28
0,29	1,1374	0,1396	1,1418	1,1440	1,1462	1,1483	1,1505	1,1527	1,1549	1,1571	29
0,30	1,1593	1,1615	1,1636	1,1658	1,1680	1,1702	1,1723	1,1745	1,1767	1,1788	30
0,31	1,1810	1,1832	1,1853	1,1875	1,1896	1,1918	1,1969	1,1961	1,1982	1,2004	31
0,32	1,2025	1,2045	1,2068	1,2090	1,2111	1,2132	1,2154	1,2175	1,2196	1,2218	32
0,33	1,2239	1,2260	1,2281	1,2303	1,2324	1,2345	1,2366	1,2387	1,2408	1,2430	33
0,34	1,2451	1,2472	1,2493	1,2514	1,2535	1,2556	1,2577	1,2598	1,2619	1,2640	34
0,35	1,2661	1,2682	1,2703	1,2724	1,2745	1,2766	1,2787	1,2808	1,2828	1,2849	35
0,36	1,2870	1,2891	1,2912	1,2933	1,2953	1,2974	1,2995	1,3016	1,3036	1,3057	36
0,37	1,3078	1,3098	1,3119	1,3140	1,3161	1,3181	1,3202	1,3222	1,3243	1,3264	37
0,38	1,3284	1,3305	1,3326	1,3346	1,3367	1,3387	1,3408	1,3428	1,3449	1,3469	38
0,39	1,3490	1,3510	1,3531	1,3551	1,3572	1,3592	1,3613	1,3633	1,3654	1,3674	39
0,40	1,3694	1,3715	1,3735	1,3756	1,3776	1,3796	1,3817	1,3837	1,3857	1,3877	40
0,41	1,3898	1,3918	1,3939	1,3959	1,3979	1,3980	1,4020	1,4040	1,4061	1,4080	41
0,42	1,4101	1,4121	1,4142	1,4162	1,4182	1,4202	1,4223	1,4243	1,4263	1,4283	42
0,43	1,4303	1,4324	1,4344	1,4364	1,4384	1,4404	1,4424	1,4445	1,4465	1,4485	43
0,44	1,4505	1,4525	1,4545	1,4566	1,4586	1,4606	1,4627	1,4646	1,4666	1,4686	44
0,45	1,4706	1,4726	1,4748	1,4767	1,4787	1,4807	1,4827	1,4847	1,4867	1,4887	45

1	2	3	4	5	6	7	8	9	10	11	12
0,46	1,4907	1,4927	1,4947	1,4967	1,4987	1,5007	1,5027	1,5048	1,5068	1,5088	46
0,47	1,5108	1,5128	1,5148	1,5168	1,5188	1,5208	1,5228	1,5248	1,5268	1,5288	47
0,48	1,5308	1,5328	1,5348	1,5368	1,5388	1,5408	1,5428	1,5448	1,5468	1,5488	48
0,49	1,5508	1,5528	1,5548	1,5568	1,5588	1,5608	1,5628	1,5648	1,5668	1,5688	49
0,50	1,5708	1,5728	1,5748	1,5768	1,5788	1,5808	1,5828	1,5848	1,5868	1,5888	50
0,51	1,5908	1,5928	1,5948	1,5968	1,5988	1,6008	1,6028	1,6048	1,6068	1,6088	51
0,52	1,6108	1,6128	1,6148	1,6168	1,6188	1,6208	1,6228	1,6248	1,6268	1,6288	52
0,53	1,6308	1,6328	1,6348	1,6368	1,6388	1,6409	1,6428	1,6449	1,6469	1,6489	53
0,54	1,6509	1,6530	1,6549	1,6569	1,6589	1,6609	1,6629	1,6649	1,6669	1,6689	54
0,55	1,6710	1,6730	1,6750	1,6770	1,6790	1,6810	1,6830	1,6850	1,6871	1,6891	55
0,56	1,6911	1,6931	1,6951	1,6971	1,6992	1,7012	1,7032	1,7052	1,7076	1,7092	56
0,57	1,7113	1,7133	1,7153	1,7173	1,7193	1,7214	1,7234	1,7254	1,7274	1,7295	57
0,58	1,7315	1,7335	1,7355	1,7376	1,7396	1,7416	1,7437	1,7457	1,7477	1,7498	58
0,59	1,7518	1,7538	1,7559	1,7579	1,7599	1,7620	1,7640	1,7660	1,7681	1,7701	59
0,60	1,7722	1,7742	1,7762	1,7783	1,7803	1,7824	1,7844	1,7865	1,7885	1,7906	60
0,61	1,7926	1,7947	1,7967	1,7988	1,8008	1,8029	1,8049	1,8070	1,8090	1,8111	61
0,62	1,8132	1,8152	1,8172	1,8194	1,8214	1,8234	1,8255	1,8276	1,8297	1,8318	62
0,63	1,8338	1,8359	1,8380	1,8400	1,8421	1,8442	1,8463	1,8484	1,8504	1,8525	63
0,64	1,8546	1,8567	1,8588	1,8609	1,8629	1,8650	1,8671	1,8692	1,8713	118734	64
0,65	1,8755	1,8776	1,8797	1,8818	1,8839	1,8860	1,8881	1,8902	1,8923	1,8944	65
0,66	1,8965	1,8986	1,9008	1,9029	1,9050	1,9071	1,9092	1,9113	1,9135	1,9156	66
0,67	1,9177	1,9198	1,9220	1,9241	1,9262	1,9284	1,9305	1,9326	1,9348	1,9369	67
0,68	1,9391	1,9412	1,9434	1,9455	1,9477	1,9498	1,9520	1,9541	1,9563	1,9584	6
0,69	1,9606	1,9628	1,9650	1,9671	1,9693	1,9714	1,9736	1,9756	1,9780	1,9801	69
0,70	1,9823	1,9845	1,9867	1,9889	1,9911	1,9933	1,9954	1,9976	1,9998	2,0020	70

1	2	3	4	5	6	7	8	9	10	11	12
0,71	2,0042	2,0065	2,0087	2,0109	2,0131	2,0153	2,0175	2,0197	2,0220	2,0242	71
0,72	2,0264	2,0286	2,0309	2,0331	2,0353	2,0376	2,0398	2,0421	2,0443	2,0465	72
0,73	2,0488	2,0511	2,0533	2,0556	2,0578	2,0601	2,0624	2,0646	2,0669	2,0692	73
0,74	2,0715	2,0737	2,0760	2,0783	2,0806	2,0829	2,0852	2,0875	2,0898	2,0921	74
0,75	2,0944	2,0967	2,0990	2,1013	2,1037	2,1060	2,1083	2,1106	2,1130	2,1153	75
0,76	2,1177	2,1200	2,1223	2,1247	2,1270	2,1294	2,1388	2,1341	2,1365	2,1389	76
0,77	2,1412	2,1436	2,1460	2,1484	2,1508	2,1532	2,1556	2,1580	2,1604	2,1628	77
0,78	2,1652	2,1676	2,1700	2,1724	2,1749	2,1773	2,1797	2,1822	2,1846	2,1871	78
0,79	2,1895	2,1920	2,1544	2,1969	2,1994	2,2019	2,2043	2,2068	2,2093	2,2118	79
0,80	2,2143	2,2168	2,2193	2,2218	2,2243	2,2269	2,2294	2,2319	2,2345	2,2370	8
0,81	2,2395	2,2421	2,2447	2,2472	2,2498	2,2524	2,2549	2,2575	2,2601	2,2627	81
0,82	2,2653	2,2680	2,2705	2,2731	2,2758	2,2784	2,2810	2,2837	2,2863	2,2870	82
0,83	2,2916	2,2943	2,2970	2,2996	2,3023	2,3050	2,3077	2,3104	2,3131	2,3158	83
0,84	2,3186	2,3213	2,3240	2,3268	2,3295	2,3323	2,3351	2,3378	2,3406	2,3434	84
0,85	2,3462	2,3490	2,3518	2,3546	2,3575	2,3603	2,3631	2,3660	2,3689	2,3717	85
0,86	2,3746	2,3775	2,3804	2,3833	2,3862	2,3891	2,3921	2,3950	2,3979	2,4009	86
0,87	2,4039	2,4068	2,4098	2,4128	2,4158	2,4189	2,4219	2,4249	2,4280	2,4310	87
0,88	2,4341	2,4372	2,4403	2,4436	2,4465	2,4496	2,4528	2,4559	2,4591	2,4623	88
0,89	2,4655	2,4687	2,4719	2,4751	2,4784	2,4816	2,4849	2,4882	2,4915	2,4948	89
0,90	2,4981	2,5014	2,5048	2,5082	2,5115	2,5150	2,5184	2,5218	2,5253	2,5287	90
0,91	2,5322	2,5357	2,5392	2,5428	2,5463	2,5500	2,5535	2,5571	2,5608	2,5644	91
0,92	2,5681	2,5718	2,5755	2,5792	2,5830	2,5868	2,5906	2,5944	2,5983	2,6022	92
0,93	2,6061	2,6100	2,6140	2,6179	2,6220	2,6256	2,6301	2,6342	2,6383	2,6425	93
0,94	2,6467	2,6509	2,6552	2,6594	2,6638	2,6681	2,6726	2,6770	2,6815	2,6856	94
0,95	2,6906	2,6952	2,6998	2,7045	2,7093	2,7441	2,7189	2,7238	2,7288	2,7338	95

1	2	3	4	5	6	7	8	9	10	11	12
0,96	2,7389	2,7440	2,7492	2,7545	2,7598	2,7652	2,7707	2,7763	2,7819	2,7876	96
0,97	2,7934	2,7993	2,8054	2,8115	2,8177	2,8240	2,8305	2,8371	2,8439	2,8507	97
0,98	2,8578	2,8650	2,8725	2,8801	2,8879	2,8960	2,9044	2,9131	2,9221	2,9314	98
0,990	2,94126	2,9422	2,94332	2,94444	2,94538	2,94642	2,94745	2,94848	2,94951	2,95054	99,0
0,991	2,95157	2,95266	2,95375	2,95484	2,95593	2,95702	2,95811	2,95920	2,96029	2,96136	99,1
0,992	2,96247	2,96363	2,96479	2,96595	2,96711	2,96827	2,96942	2,97058	2,97174	2,97290	99,2
0,993	2,97406	2,97531	2,97655	2,97780	2,97904	2,98029	2,98154	2,98278	2,98403	2,98527	99,3
0,994	3,98652	2,98787	2,98923	2,99058	2,99193	2,99329	2,99464	2,99599	2,99734	2,99870	99,4
0,995	3,00005	3,00155	3,00304	3,00454	3,00604	3,00754	3,00903	3,01053	3,01203	3,01352	99,5
0,996	3,01502	3,01672	3,01847	3,02011	3,02181	3,02351	3,02521	3,02691	3,02860	3,03080	99,6
0,997	3,03200	3,03401	3,03602	3,03804	3,04005	3,04206	3,04407	3,04608	3,04810	3,05011	99,7
0,998	3,05212	3,05474	3,05736	3,05999	3,06261	3,06523	3,06785	3,07047	3,07310	3,07522	99,8
0,9990	3,07834	3,07897	3,07961	3,08024	3,08087	3,08151	3,08214	3,08277	3,08340	3,08404	99,90
0,9991	3,08467	3,08530	3,08593	3,08657	3,08720	3,08783	3,08846	3,08909	3,08973	3,09036	99,91
0,9992	3,09099	3,09162	3,09226	3,09289	3,09352	3,09416	3,09479	3,09542	3,09605	3,09669	99,92
0,9993	3,09732	3,09795	3,09858	3,09922	3,09985	3,10048	3,10111	3,10174	3,10238	3,10301	99,93
0,9994	3,10364	3,10427	3,10491	3,10554	3,10617	3,10681	3,10744	3,10807	3,10870	3,10934	99,94
0,9995	3,10997	3,11060	3,11123	3,11187	3,11250	3,11313	3,11376	3,11439	3,11503	3,11566	99,95
0,9996	3,11629	3,11692	3,11756	3,11819	3,11882	3,11946	3,12009	3,12072	3,12135	3,12199	99,96
0,9997	3,12262	3,12325	3,12388	3,12452	3,12515	3,12578	3,12641	3,12704	3,12768	3,12831	99,97
0,9998	3,12894	3,12957	3,13020	3,13084	3,13147	3,13210	3,13273	3,13336	3,13399	3,13463	99,98
0,9999	3,13526	3,13589	3,13653	3,13716	3,13779	3,13843	3,13906	3,13969	3,14032	3,14096	99,99
1.0000	3,14159										100.00

Таблиця 3

Стандартні значення критерія Фішера

v ₂ \ v ₁	1			v ₂ \ v ₁	1			v ₂ \ v ₁	1		
	p=0,001	p=0,01	p=0,05		p=0,001	p=0,01	p=0,05		p=0,001	p=0,01	p=0,05
10	21,0	10,0	5,0	25	13,9	7,8	4,2	55	12,1	7,1	4,0
11	19,7	9,7	4,8	26	13,7	7,7	4,2	60	12,0	7,1	4,0
12	18,6	9,3	4,8	27	13,6	7,7	4,2	65	11,9	7,0	4,0
13	17,8	9,1	4,7	28	13,5	7,6	4,2	70	11,6	7,0	4,0
14	17,1	8,9	4,6	29	13,4	7,6	4,2	80	11,6	7,0	4,0
15	16,6	8,7	4,5	30	13,3	7,6	4,2	100	11,5	6,9	3,9
16	16,1	8,5	4,5	32	13,2	7,5	4,1	125	11,4	6,8	3,9
17	15,7	8,4	4,5	34	13,1	7,4	4,1	150	11,3	6,8	3,9
18	15,4	8,3	4,4	36	13,0	7,4	4,1	200	11,2	6,8	3,9
19	15,1	8,2	4,4	38	12,9	7,3	4,1	200	11,2	6,8	3,9
20	14,8	8,1	4,3	40	12,8	7,3	4,1	400	11,0	6,7	3,9
21	14,6	8,0	4,3	42	12,7	7,3	4,1	1000	10,9	6,7	3,8
22	14,4	7,9	4,3	46	12,4	7,2	4,0	∞	10,8	6,6	3,8
23	14,2	7,9	4,3	48	12,3	7,2	4,0				
24	14,0	7,8	4,3	50	12,2	7,2	4,0				

Абетковий покажчик

Варіанта
Варіаційний ряд
Відносна частота
Відхилення
Вибірка
Випадкова величина
Генеральна сукупність
Дисперсія
Ймовірність
Закон розподілу
Критерій Стьюдента
Критерій Фішера
Критерій φ (фі)
Математичне сподівання
Нормальний розподіл
Обсяг вибірки
Обсяг генеральної сукупності
Статистика
Статистичний розподіл
Середнє арифметичне
Середнє квадратичне відхилення
Помилка репрезентативності
Розподіл
 "хі-квадрат"
 t-розподіл Стьюдента
 F-розподіл Фішера-Снедекора
Частота
Частотні таблиці

Рекомендована література

1. Гаркавий В.Г., Ярова В.В. Математична статистика. – К: Професіонал, 2004. – 484 с.
2. Плохинский Н.А. Руководство по биометрии для зоотехников. – М: Колос, 1969. – 256с.
3. Турчин В.М. Математична статистика. – К: Академія, 1999. – 240 с.

ЗМІСТ

ВСТУП

- I. Основні поняття теорії ймовірностей
- II. Генеральна сукупність. Частота
- III. Середня арифметична величина
- IV. Середнє квадратичне відхилення. Дисперсія
- V. Визначення помилки репрезентативності
- VI. Оцінка достовірності різниць між групами за критерієм Стьюдента
- VII. Оцінка достовірності різниць між групами за критерієм Фішера
- VIII. Оцінка достовірності різниць між групами за методом кутів (φ -метод).
- IX. Кореляція. Коефіцієнт кореляції

ДОДАТКИ

Абетковий покажчик

Рекомендована література

СТАТИСТИЧНА ОБРОБКА ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ
Навчальний посібник

Мельниченко Олена Петрівна
Якименко Ігор Леонідович
Шевченко Ростислав Леонідович

Редактор О.М. Трегубова
Комп'ютерна верстка

Здано до складання 25.05.2006. Підписано до друку. Формат 60x84 ¹/₁₆.
Ум. Друк. арк. Тираж 300 примірників. Зам. Ціна. Сектор оперативної
поліграфії РВІКВ БДАУ 09117 Біла Церква, Соборна пл. 8, тел. 3-11-01