

## МАТЕМАТИЧНА ОБРОБКА РЕЗУЛЬТАТІВ ЕКСПЕРИМЕНТУ

У статті представлено порядок виконання статистичної обробки результатів експерименту та можливі недоліки і помилки у висвітленні остаточних результатів лабораторних досліджень.

**Ключові слова:** достовірні, стандартне відхилення, стандартна помилка середнього.

**Постановка проблеми.** Результати лабораторних досліджень – улюблений об'єкт для статистичної обробки в наукових працях з найрізноманітніших сільськогосподарських спеціальностей. У роботі молодого науковця велика увага надається саме для отримання певних лабораторних даних та методикам їх отримання. Зазвичай на обробку та пояснення результатів залишається менше часу та уваги, ніж на саме виконання досліду. Крім того, спеціаліст, який закінчив сільськогосподарський вищий навчальний заклад, не зовсім компетентний у статистичній обробці результатів експерименту. Як правило, це зводиться до знаходження середнього арифметичного, стандартної помилки середнього та встановлення корелятивних зв'язків між групами [1].

**Мета роботи** – розробити методику виконання математичної обробки даних та обґрунтувати її переваги над існуючими, а також висвітлити можливі недоліки і помилки у представленні остаточних результатів досліджень.

**Матеріал і методи досліджень.** Для представлення порядку статистичної обробки лабораторних даних користувалися вмістом вітаміну Е в жовтку перепелиних яєць породи фараон, що визначали [5] за реакцією окиснення токоферолів хлорним залізом. Статистичну обробку результатів проводили за методикою Плохинського Н.А. [3] з урахуванням деяких поправок [1, 6].

**Результати досліджень та їх обговорення.** Основними задачами статистичного аналізу в більшості випадків є: опис групи чи груп даних з розрахунком параметрів розподілу та порівняння деяких груп даних [2]. Компактним описом даних займається описова статистика, в основі якої лежить поняття нормального розподілу [4]. Нормальний розподіл ознаки зустрічається досить часто: якщо деяка величина, наприклад, концентрація гемоглобіну крові,

відхиляється від середнього значення під дією великої кількості незначних та незалежних один від одного факторів: втрата та надходження заліза, інтенсивність електропоезу, тривалість життя еритроцитів та інше. Тобто, якщо значення ознаки, що нас цікавить у більшості об'єктів, наближаються до їх середнього і з рівною імовірністю відхиляються від нього у більшу чи меншу сторону, розподіл називається *нормальним* [2].

Для опису такого розподілу використовуються наступні параметри: середнє значення  $M$  та стандартне відхилення  $s$  [6]. Скористаємося отриманими лабораторними даними вмісту вітаміну Е в жовтку перепелиних яєць.

Створимо алгоритм розрахунку параметрів розподілу:

1. Розташуємо отримані результати в таблиці у другому рядочку та пронумеруємо їх (табл. 1).

**Таблиця 1. Розрахунок середнього значення та стандартного відхилення**

№	1	2	3	4	5	6	7	8	9	10
$X$	70,0	70,2	70,5	70,4	70,6	70,4	70,8	70,9	70,5	70,7
$X - M$	-0,5	-0,3	0	-0,1	0,1	-0,1	0,3	0,4	0	0,2
$(X - M)^2$	0,25	0,9	0	0,01	0,01	0,01	0,09	0,16	0	0,04

2. Розрахуємо середнє арифметичне вказаних даних:  $M = \frac{\sum_{i=1}^n X_i}{n} = 70,5$  мкг/г

(як суму вмісту вітаміну Е в жовтку перепелиних яєць поділено на число проб).

3. Третій рядок заповнимо відхиленнями даних від середнього значення:  $X - M$  (від кожного значення  $X$  віднімаємо середнє значення  $M$ ). Зауважимо, що сума всіх відхилень повинна дорівнювати нулю.

4. Далі підносимо до квадратів отримані відхилення  $(X - M)^2$ .

5. Розраховуємо стандартне відхилення, яке називають *середнім квадратичним відхиленням*:

$$s = \sqrt{\frac{\sum_{i=1}^n (X - M)^2}{n - 1}} = \frac{1,47}{10 - 1} = 0,4.$$

Таким чином, в результаті нескладних розрахунків отримуємо два основних параметри – середнє значення  $X$  та стандартне відхилення  $(X - M)^2$ , що характеризує розподілення ознаки в сукупності даних, в нашому випадку вмісту вітаміну Е в жовтку перепелиних яєць. Отримані результати ми можемо

записати у форматі  $M \pm s$ , в нашому випадку це:  $70,5 \pm 0,4$  мкг/г. Звернемо увагу на те, що середнє значення та стандартне відхилення вимірюються в одних одиницях, у нашому випадку мкг/г.

На жаль, в більшості наукових робіт отримані результати мали б такий вигляд:  $M \pm m$ , а в цифрах  $70,5 \pm 0,1$  мкг/г, причому науковцю цей результат більш до вподоби, оскільки на його думку отримано менше відхилення і більш достовірний результат.

Що ж таке і звідки береться  $m$  – стандартна помилка середнього? У більшості випадків ми маємо справу не з генеральною сукупністю, а з випадково вибраною вибіркою. Якщо ми візьмемо для статистичної оцінки результати вмісту вітаміну Е в жовтку перепелиних яєць в інший час або в жовтку яєць, знесених іншими перепілками, то середнє значення буде відрізнятися від тієї величини, яку ми щойно отримали. Можна сказати, що стандартна помилка середнього відображає діапазон значень, в якому повинно знаходитися середнє значення будь-якої випадкової вибірки. В нашому прикладі це значить, що в більшій половині випадків середнє значення будь-якої вибірки буде знаходитися в діапазоні від 70,4 до 70,6 мкг/г.

Для розрахунку стандартної помилки середнього використовують формулу:  $m = \frac{s}{\sqrt{n}}$ . Оскільки дана величина менша від стандартного відхилення, автори наукових робіт залюбки використовують її, підтверджуючи цим самим більш точні та достовірні результати. Можна зауважити і те, що стандартна помилка відображає тільки точність оцінки середнього і не дає наочного уявлення про розсіювання даних. Таким чином, описуючи сукупність даних, рекомендується приводити значення стандартного відхилення.

Для більш точних висновків необхідно збільшити об'єм вибірки і перевірити схожість отриманої гістограми з нормальним розподілом.

Якщо аналізу піддаються дві чи більше груп тварин, виникає питання оцінки різниці між групами за аналізованим показником. Наприклад, вміст вітаміну Е в жовтку перепелиних яєць склав  $70,5 \pm 0,4$  мкг/г, а в жовтку яєць після тривалого зберігання –  $65,3 \pm 0,5$  мкг/г. Порівняємо середні значення в

групах, визначивши цю різницю у відсотках. Можна зробити висновок, що вміст вітаміну Е в жовтку свіжих перепелиних яєць вищий на 7% (якщо показник вмісту вітаміну свіжих яєць прийняти 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 статистично значущої різниці між групами за методом Стюдента. Для цього визначається критерій достовірності різниці

за формулою:  $t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{m_1^2 + m_2^2}}$ , або з використанням числа спостережень і

стандартних відхилень:  $t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ . Але для правильності обрахунків варто з

двох квадратів стандартних відхилень  $s_1^2$  та  $s_2^2$  розрахувати об'єднану оцінку

дисперсії для двох даних груп:  $s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ . Тепер, коли відома

об'єднана оцінка дисперсії  $s^2$  для двох довільних вибірок, розраховуємо

«повноцінний» критерій Стюдента:  $t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}}$ .

Повернемося до результатів вмісту вітаміну Е в жовтку перепелиних яєць. Для наших досліджень  $s^2 = 0,45$ , а  $t = 26$ . Отже, критерій  $t = 26$  ми розрахували, але тепер необхідно правильно його оцінити.

Принцип оцінки критерію Стюдента полягає в тому, що чим більший отриманий результат  $t$ , тим більше імовірності вважати, що отримана різниця між групами статистично значуща. Обчислений критерій  $t$  порівнюється зі стандартним (табличним) значенням критерію Стюдента  $t_{st}$  для  $\nu = n_1 + n_2 - 2$ , де  $n_1$  та  $n_2$  – кількість вимірів у групах. Для нашого випадку  $\nu = 10 + 10 - 2 = 18$  і табличне значення критерію  $t_{st}$  дорівнює:  $\{2,101; 2,878; 3,922\}$ . Отримане значення  $t = 26$  перевищує табличні значення, тому можемо стверджувати, про найвищий рівень статистично значущої різниці між групами  $p < 0,001$ .

Якщо вирахований критерій  $t$  більший за стандартне значення критерію Стьюдента  $t_{st}$  для  $p < 0,05$ , це означає, що різниця між групами є статистично значущою з надійністю 95% (тобто різницю можна очікувати у 95 випадках зі 100). Якщо  $t < t_{st}$  для  $p < 0,01$ , різниця статистично значуща з надійністю 99 %, якщо  $t < t_{st}$  для  $p < 0,001$ , різниця статистично значуща 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 відсутність статистично значущої різниці між групами.

Крім того, варто зауважити, що не варто замість терміна «статистично значущі» використовувати популярний термін «достовірні», що має в статистиці інше значення. Відмітимо і те, що критерій Стьюдента застосовується лише для порівняння двох груп, а не декількох груп попарно. У випадку використання критерію Стьюдента для множинних порівнянь необхідно введення поправки Бонфероні, або застосування інших критеріїв.

Виникає ще одне питання: наскільки правомірно використання параметрів нормального розподілу для опису конкретної сукупності даних? Для цього побудуємо відповідну гістограму частот, розділивши весь інтервал на проміжки із кроком 0,2 мкг/г (рис.1).

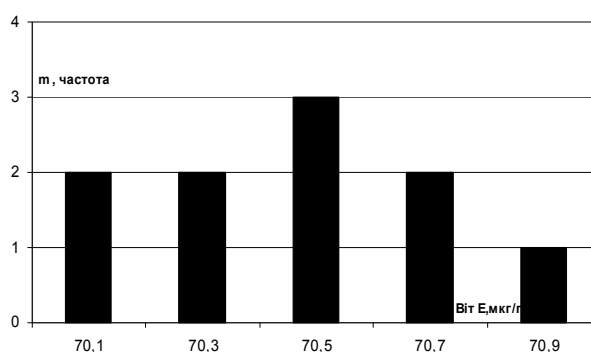


Рис.1. Гістограма розподілення вмісту вітаміну Е в жовтку перепелиних яєць.

Дані для побудови цієї гістограми отримано із табл. 1. Але для більш точного результату необхідно взяти значно більше проб і порівняти утворений графік із графіком нормального розподілу. Лише після відміченої схожості можна використовувати критерій Стьюдента.

## Висновки та перспективи подальших досліджень

1. Необхідно враховувати, що для застосування критерію Стьюдента значення ознак кожної із груп, що порівнюються, повинні мати розподіл, близький до нормального, а їх дисперсії є приблизно рівними [1, 2, 6].

2. Не варто застосовувати замість середнього квадратичного відхилення  $s$  стандартну помилку середнього  $m$ , оскільки одна з цих величин характеризує вибірку, а інша генеральну сукупність [4].

3. Для обрахунку критерію Стьюдента необхідно розрахувати об'єднану оцінку дисперсії для двох даних груп  $s^2$ , а не дисперсії кожної групи окремо.

4. Критерій Стьюдента застосовується лише для порівняння двох груп, а не декількох груп попарно. У випадку використання критерію Стьюдента для множинних порівнянь необхідно введення поправки Бонфероні, або застосування інших критеріїв [1].

## СПИСОК ЛІТЕРАТУРИ

1. Гаркавий В.Г., Ярова В.В. Математична статистика. – К.: Професіонал, 2004. – 484 с.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1999.
3. Плохинский Н.А. Руководство по биометрии для зоотехников. – М.: Колос, 1969. – 256с.
4. Солодовников А.С. Теория вероятностей. – М.: Просвещение, 1983.
5. Сурай П.Ф., Ионов И.А. Методы анализа кормов и продуктов птицеводства: Метод. реком. – Харьков, 1989. – 95 с.
6. Турчин В.М. Математична статистика. – К.: Академія, 1999. – 240 с.

### Математическая обработка результатов эксперимента

**Е.П. Мельниченко**

В работе представлено порядок статистической обработки результатов эксперимента и возможные недостатки и ошибки в представлении окончательных результатов исследований.

**Ключевые слова:** достоверные, стандартное отклонение, стандартная ошибка среднего.

### Mathematical processing of experimental results

**E. Melnishenko**

The paper presents the procedure of statistical processing of experimental results and possible shortcomings and mistakes highlighted in the final results of laboratory studies.

**Key words:** significant standard deviation, average standard error.