

## ЛАБОРАТОРНЕ ЗАНЯТТЯ № 14

### ТЕМА: ОЦІНКА ДОСТОВІРНОСТІ РЕЗУЛЬТАТІВ ДОСЛІДЖЕННЯ: ПЕРВИННІ МЕТОДИ СТАТИСТИЧНОГО АНАЛІЗУ

**Мета:** ознайомитися з первинними методами обробки результатів експерименту і застосувати їх при обчисленні середньої арифметичної, дисперсії, моди та медіани печінкових показників у здорових осіб та хворих на гепатит С.

#### **Значення теми:**

Статистика в біології та медицині є одним з інструментів аналізу експериментальних даних і клінічних спостережень, а також мовою, за допомогою якої повідомляються отримані математичні результати.

**Знати:** первинні методи статистичного аналізу.

**Вміти:** визначати середню арифметичну, дисперсію, моду та медіану.

#### ***Короткий зміст теми***

Методами статистичної обробки результатів експерименту називаються математичні прийоми, формули, способи кількісних розрахунків, за допомогою яких показники, що одержані в ході експерименту, можна узагальнювати, приводити в систему, виявляючи приховані в них закономірності.

Головна мета будь-якого статистичного методу - представити кількісні дані в систематизованій і стислій формі з тим, щоб полегшити їх розуміння.

**До первинних методів**, за допомогою яких можна отримати показники, що безпосередньо відображають результати отриманих в експерименті вимірювань належать: визначення середньої арифметичної, дисперсії, моди та медіани.

#### **Хід роботи**

При проведенні будь-якого дослідження можливі помилки, які не можна врахувати математичними методами, але при гарній організації дослідження їх можна уникнути або звести до мінімуму:

- а) помилки методичні (неправильна методика збору і обробки матеріалу);
- б) помилки точності (неточність приладів, недостатня точність розрахунків, неточність первинної реєстрації фактів);
- в) помилки уваги (описки, прорахунки, помилки);
- г) помилки типовості (відбір групи об'єктів, нетипових для всієї генеральної сукупності, тенденційний підбір первинних даних).

Для зменшення розмірів помилок необхідно дотримуватися об'єктивності відбору одиниць спостереження, використовувати контроль за якістю матеріалу на кожному етапі роботи.

#### **Методика обробки результатів вимірювань**

*1. Проводять n незалежних дослідів і визначають n значень величини, яку визначають*

Наприклад: Вміст С-реактивного білка у хворих на ревматоїдний артрит при загостренні та після лікування наданий у таблиці:

Хворі із загостренням РА		Хворі після лікування	
№	СРБ, мг/л	№	СРБ, мг/л
1	60	1	0
2	120	2	60
3	90	3	0
4	120	4	60
5	120	5	30
6	160	6	60
7	192	7	0
8	60	8	0
9	60	9	0
10	60	10	0

2. Розраховують середнє арифметичне значення величини, яку досліджують

Середня арифметична (М) (далі просто «середня») – це найбільш часто використовуваний показник центральної тенденції. Її застосовують, зокрема, в розрахунках, необхідних для опису розподілу і для його подальшого аналізу. Її обчислюють, розділивши суму всіх значень даних на число цих даних.

$$\bar{x} = \frac{\sum x_i}{n},$$

де  $\bar{x}$  – середньоарифметичне значення, n – кількість досліджуваних, або кількість приватних показників;  $\sum x_i$  – сума всіх результатів.

Нехай n = 10:  $x_1, x_2, x_3, \dots, x_n$ .

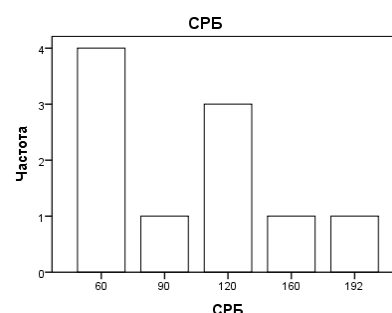
$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10}}{10}$$

У нашому прикладі у хворих з РА при загостренні середньоарифметичне значення дорівнюватиме 104,2 мг / л

Варіаційний ряд, крім середньої арифметичної, може бути охарактеризований ще двома величинами – *медианою і модою*.

Модою (позначається  $M_o$ ) називається та варіанта, якій відповідає найбільша кількість частот варіаційного ряду. Так, в нашому прикладі для експериментальної 1 групи мода для фону буде дорівнювати 60, тому що вона зустрічається 4 рази.

Первинний результат	Частота
60	4
90	1
120	3
160	1
192	1



Мода використовується рідко і головним чином для того, щоб дати загальне уявлення про розподіл. Модою можна користуватися для оцінки середньої тривалості захворювань, особливо при малій кількості хворих на цю хворобу.

Медианою (позначається  $Me$ ) називається варіанта, що ділить варіаційний ряд на дві рівні половини, тобто та варіанта, менше якої мають розміри половини частот.

Медіана застосовується в біологічній статистиці рідко. За допомогою медіани визначається, наприклад, так звана ймовірна тривалість майбутнього життя в таблицях смертності.

Основною відмінністю медіани і моди від середньої арифметичної є те, що на їх розміри не впливають величина крайніх значень варіант, наявних у варіаційному ряду, тоді як при визначенні середньої арифметичної беруться до уваги значення всіх варіант.

У нашому прикладі у хворих з РА при загостренні медіана дорівнює:  $(120 + 90) : 2 = 105$  мг / л

3. Розраховують відхилення кожного результату від середнього значення (середньої арифметичної) ( $\Delta x_i = x_i - \bar{x}$ ):

$$\Delta x_1 = |x_1 - \bar{x}|$$

$$\Delta x_2 = |x_2 - \bar{x}|$$

$\vdots$

$$\Delta x_{10} = |x_{10} - \bar{x}|$$

4. Для виявлення розкиду приватних даних щодо середньої використовують величину, яку називають стандартним (або середньоквадратичним) відхиленням

Визначають стандартне відхилення від середнього за формулою:

$$S_{\bar{x}} = \sqrt{\frac{\Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_n^2}{n}}$$

Для того, щоб більш точно оцінити стандартне відхилення для малих вибірок (з числом елементів менше 30), в знаменнику виразу під коренем треба використовувати не  $n$ , а  $n-1$ :

$$S_{\bar{x}} = \sqrt{\frac{\Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_n^2}{n-1}}$$

У нашому прикладі у хворих при загостренні РА стандартне відхилення дорівнює 46,65. Воно дозволяє сказати, що більша частина результатів розташовується в межах 46,652 від середньої, тобто між 57,55 ( $104,2 - 46,65$ ) і 150,85 ( $104,2 + 46,65$ )

#### 5. Визначення помилки показника

Величина помилки репрезентативності визначається як обсягом вибірки, так і різноманітністю ознаки. Чим більше число спостережень, тим менше помилка; чим більш мінлива ознака, тим більше величина статистичної

похибки. Розглянемо обчислення середніх помилок відносного показника і середньої величини.

1. Середня помилка показника обчислюється за формулою:  $m = \pm \sqrt{\frac{p * q}{n}}$ , де  $m$  – середня помилка;  $p$  – статистичний коефіцієнт (відносна величина);  $q$  – величина, зворотна  $p$  (альтернативний показник), і виражена як  $(1-p)$ ,  $(100-p)$ ,  $(1000-p)$  і т.д. в залежності від основи, на яку розрахований коефіцієнт;  $n$  – число спостережень у вибірковій сукупності.

Якщо число спостережень недостатньо велике (менше 30), у формулу вводиться правка:

$$m = \pm \sqrt{\frac{p * q}{n - 1}}$$

**Приклад:** Розрахувати середню помилку показника летальності в лікувальному закладі, якщо відомо: всього вибуло зі стаціонару 317 хворих, з них померло 13.

Летальність складе:

$$\begin{aligned} \frac{13 * 100}{317} &= 4.1\% \\ p &= 4,1 \quad q = 100 - 4,1 = 95,9 \quad n = 317 \\ m &= \pm \sqrt{\frac{pq}{n}} = \pm \sqrt{\frac{4,1 * 95,9}{317}} = \pm 1,11 \end{aligned}$$

Таким чином, показник летальності дорівнює:  $4,1 \pm 1,11\%$

2. Розрахунок помилки середньої величини здійснюється за формулою:

$m = \pm \frac{\sigma}{\sqrt{n}}$  і  $m = \pm \frac{\sigma}{\sqrt{n-1}}$ , якщо  $n \leq 30$ , де  $m$  – середня помилка;  $\sigma$  – середнє квадратичне відхилення;  $n$  - число спостережень.

**Приклад:** У результаті визначення СРБ в крові у 10 хворих на ревматоїдний артрит були отримані наступні дані: середній вміст СРБ ( $\bar{x}$ ) склав 104,2 мг/л; середнє відхилення ( $\sigma$ ) - 46,65 мг / л. Визначити помилку змісту СРБ в крові.

$$m = \pm \frac{\sigma}{\sqrt{n}} = \frac{46,65}{\sqrt{10}} = \pm 14,76 \text{ г} \quad \bar{x} = 104,2 \pm 14,76 \text{ мг/л.}$$

**Завдання:** провести обчислення середньої арифметичної, дисперсії, моди та медіани за даними таблиці для АСТ, глюкози, білірубіну та тимолової проби у здорових осіб (група 1) та хворих на гепатит С (група 2).

групи	№	АЛТ	АСТ	глюкоза	білірубін	тимолова проба
1	1	26	30	5,3	10,0	1,6
1	2	27	31	4,4	13,0	2,4
1	3	31	28	5,2	15,0	3,2
1	4	24	36	4,8	10,0	1,8
1	5	21	27	5,0	13,0	0,8
1	6	23	25	4,6	15,5	2,4

1	7	24	23	4,8	10,0	1,6
1	8	21	28	5,2	13,0	1,2
1	9	29	32	5,3	16,8	3,6
1	10	20	24	4,5	12,0	1,8
$\bar{X}_1$		<b>24,6</b>				
$m_1$		<b>1,15</b>				
2	1	85	76	6,0	17,0	2,4
2	2	60	48	6,2	14,0	3,2
2	3	48	34	4,3	18,0	1,8
2	4	46	30	5,0	14,0	2,4
2	5	54	33	5,2	12,0	1,8
2	6	44	28	5,8	14	1,8
2	7	44	30	5,2	18,0	1,8
2	8	70	50	4,8	15,0	2,4
2	9	40	32	4,8	16,0	1,8
2	10	68	42	5,0	14,0	2,4
$\bar{X}_2$		<b>55,9</b>				
$m_2$		<b>4,6</b>				

Зробити висновки.