ЛЕКЦІЯ 8
ПАРАМЕТРИЧНІ МЕТОДИ ПОРІВНЯННЯ ДВОХ ВИБІРОК

1. Теоретичні засади та сфера застосування t-критерію Стьюдента.
2. Статистичний критерій t-Стьюдента для однієї вибірки.
3. Статистичний критерій t-Стьюдента для незалежних вибірок.
4. Статистичний критерій t-Стьюдента для залежних вибірок.

1. Теоретичні засади та сфера застосування t-критерію Стьюдента

Для порівняння вибіркових середніх величин, які належать до двох сукупностей даних, і для вирішення питання про те, чи відрізняються середні значення статистично достовірно один від одного, використовують t-критерій Стьюдента або його непараметричні аналоги.

Критерій розроблений англійським математиком Вільямом С. Госсетом (1876-1937). «Стьюдент» – це псевдонім, під яким В. Госсет друкував свої роботи, працюючи в Дубліні на пивоварні Артура Гіннеса. Інший дослідник, який свого часу працював на Гіннеса, опублікував статтю, що містила конфіденційну комерційну інформацію, – з того часу працівникам підприємства було заборонено публікуватися (незалежно від змісту публікації). Тож В. Госсет, щоб уникнути розголосу і обійти заборону, друкував свої роботи під псевдонімом.

В. Госсет тісно співпрацював із К. Пірсоном (багато його робіт надруковано у Пірсонівському журналі «Біометрика»), і саме К. Пірсон відправив його роботу «The probable error of a mean», у якій власне і дається опис t-статистики Р. Фішеру. Завдяки пропозиціям Р. Фішера t-критерій набув сучасного вигляду.

У вітчизняній традиції прийнято говорити про t-критерій Стьюдента. В англомовній літературі і в статистичних програмах вживають термін «Т-тест».

Критерій t-Стьюдента використовується у трьох випадках:

1) порівняння середніх показників двох залежних вибірок (t-критерій для залежних вибірок);

2) порівняння середніх показників двох незалежних вибірок (t-критерій для незалежних вибірок);

3) порівняння середнього показника однієї вибірки із певною заданою величною (t-критерій для однієї вибірки).

Застосувавши критерій Стьюдента, ми дізнаємося, наскільки статистично значимі відмінності між двома вибірками і, відповідно, наскільки впевнено можна робити висновки про ці відмінності.

Цей параметричний метод, що використовується для перевірки гіпотез про достовірність різниці середніх при аналізі кількісних даних в генеральних сукупностях з нормальним розподілом і з однаковою дисперсією. На жаль, метод Стьюдента дуже часто використовують для малих вибірок, не впевнившись попередньо в тому, що дані у відповідних генеральних сукупностях нормально розподілені (наприклад, результати виконання дуже легкого завдання, з яким впоралися всі досліджувані, або ж, навпаки, дуже важкого завдання не дають нормального розподілу).

Метод Стьюдента різний для незалежних і залежних вибірок. Незалежні вибірки отримують при дослідженні двох різних груп досліджуваних (наприклад, це контрольна і досліджувана групи). До залежних вибірок відносяться, наприклад, результати однієї і тієї ж групи досліджуваних до і після впливу незалежної змінної.

2. Статистичний критерій t-Стьюдента для однієї вибірки

Алгоритм обчислення t-критерію Стьюдента для однієї вибірки

1. Сформулюйте статистичні гіпотези:

H_0 – відмінності між $\overline{X}$ та A випадкові і незначимі.

H_1 – відмінності між $\overline{X}$ та A достовірні, значимі.

2. Обчисліть емпіричне значення t-критерію ($t_{емп}$) за формулою:

$$t_{емп} = \frac{|\overline{X} - A|}{\sigma / \sqrt{n}}$$

df –число ступенів свободи df = n - 1

$\overline{X}$ – середнє арифметичне;

A – величина, з якою середнє арифметичне порівнюється;

σ – стандартне відхилення;

n – обсяг вибірки.

3. Врахувавши число ступенів свободи (df) за таблицею критичних значень критерію t-Стьюдента, знайдіть $t_{крит}$ і порівняйте отримані значення:

- якщо $t_{емп} \geq t_{табл}$ – приймають гіпотезу H_1;
- якщо $t_{емп} < t_{табл}$ – приймають гіпотезу H_0.

3. Статистичний критерій t-Стьюдента для незалежних вибірок

Призначення критерію: перевірка гіпотези про достовірність різниці середніх на вибірках з розподілом, близьким до нормального. Дозволяє оцінити відмінності між середніми значеннями $\overline{x_1}$ та $\overline{x_2}$ двох вибірок.

Область дії: дві незалежні вибірки (контрольна, експериментальна групи).

Обмеження:

1) дані виміряні у кількісних шкалах
2) розподіл ознаки в обох вибірках відповідає закону нормального розподілу
3) об'єм двох вибірок не повинні суттєво відрізнятися один від одного (не більше ніж у 1,5-2 рази)

Формулювання гіпотез:

H_0 : значення ознаки у контрольній (К) та експериментальній (Е) групах не відрізняються.

H_1 : значення ознаки у контрольній (К) та експериментальній (Е) групах відрізняються.

H_0 : середнє значення ознаки у вибірці 1 достовірно не відрізняється від середнього значення ознаки у вибірці 2

H_1 : середнє значення ознаки у вибірці 1 достовірно відрізняється від середнього значення ознаки у вибірці 2

H_0 : відмінності $\overline{x_1}$ та $\overline{x_2}$ випадкові і незначимі

H_1 : відмінності $\overline{x_1}$ та $\overline{x_2}$ достовірні, значимі

У загальному вигляді формула для підрахунку t-критерія Стьюдента:

$$t_{\text{екс}} = \frac{|\overline{x} - \overline{y}|}{\sigma}$$

Для вибірок $n_1=n_2=n$

$$\sigma = \sqrt{\sigma_x^2 + \sigma_y^2} = \sqrt{\frac{\sum(x-\overline{x})^2 + \sum(y-\overline{y})^2}{(n-1)\cdot n}}$$

У випадку коли $n_1 \neq n_2$

$$\sigma = \sqrt{\sigma_x^2 + \sigma_y^2} = \sqrt{\frac{\sum(x-\overline{x})^2 + \sum(y-\overline{y})^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1\cdot n_2}}$$

Алгоритм:

1. Підрахувати статистики першої вибірки $\overline{x_1}$ та σ_1.
2. Підрахувати статистики другої вибірки $\overline{x_2}$ та σ_2.
3. Підрахувати кількість ступенів свободи для незалежних вибірок об'ємом n_1 та n_2: $df= n_1+n_2-2$.
4. Підрахувати емпіричне значення t-критерію за однією із формул:

   $$t_{\text{екс}} = \frac{|\overline{x} - \overline{y}|}{\sigma}$$

   або

   $$t_{\text{емп}} = \frac{|\overline{x} - \overline{y}|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

5. За таблицею критичних значень $t_{\text{кр}}$ для відповідного числа ступенів свободи df. Якщо $|t_{\text{емп}}| \geq t_{\text{кр}}$ на рівні значимості p=0,05, то гіпотеза H_0 приймається.

Оцінювання результатів:

А) Якщо результат $t_{екс}$, одержаний при обчисленні значення критерію Стьюдента, менше, ніж значення $t_{ст}$ для рівня значущості $p=0,05$ та ступенів свободи n_1+n_2-2 , то *приймаємо нульову гіпотезу*: значення ознаки у контрольній та експериментальній групах не відрізняються; вибірки із однієї популяції;

Б) Якщо результат $t_{екс}$, одержаний при обчисленні значення критерію Стьюдента, більше, ніж значення $t_{ст}$ для рівня значущості $p=0,01$ та ступенів свободи n_1+n_2-2, то приймаємо альтернативну гіпотезу: значення ознаки у контрольній та експериментальній групах відрізняються; вибірки належать до різних популяцій.

Задача 1. Психолог вимірював час складної сенсомоторної реакції вибору (мс) в контрольній та експериментальній групах. В експериментальну групу (X) входило 9 спортсменів, в контрольну групу (Y) – 8 осіб, які активно не займалися спортом. Психолог перевіряв гіпотезу про те, що середня швидкість складної сенсомоторної реакції у спортсменів вища, ніж ця ж величина у людей, які не займаються спортом.

Результати експерименту були представлені в таблиці, зроблено деякі розрахунки:

№	Групи		Відхил від середнього		Квадрати відхилень	
	X	Y	$\sum(x_i - \overline{x})$	$\sum(y_i - \overline{y})$	$\sum(x_i - \overline{x})^2$	$\sum(y_i - \overline{y})^2$
1	504	580	-22	-58	484	3364
2	560	692	34	54	1156	2916
3	420	700	-106	62	11236	3844
4	600	621	74	-17	5476	289
5	580	640	54	2	2916	4
6	530	561	4	-77	16	5929
7	490	680	-36	42	1296	1764
8	580	630	54	-8	2916	64
9	470	-	-56	0	3136	0
Сума	4734	5104	0	0	28632	18174
Середнє	526	638				

Середні значення у кожній вибірці будуть:

$\overline{x} = \frac{4734}{9} = 526$

$\overline{y} = \frac{5104}{8} = 638$

Обчислимо різницю за абсолютною величиною між середніми значеннями:

$$|\overline{x} - \overline{y}| = |526 - 638| = 112$$

Обчислимо дисперсію за формулою:

$$\sigma = \sqrt{\sigma x^2 + \sigma y^2} = \sqrt{\frac{\sum(x - \overline{x})^2 + \sum(y - \overline{y})^2}{n_1 + n_2 - 2} \cdot \frac{n_1 + n_2}{n_1 \cdot n_2}}$$

Отримаємо:

$$\sigma = \sqrt{\frac{28632 + 18171}{9 + 8 - 2} \cdot \frac{9 + 8}{9 \cdot 8}} = \sqrt{\frac{46806}{15} \cdot \frac{17}{72}} = \sqrt{736{,}8} \approx 27{,}14$$

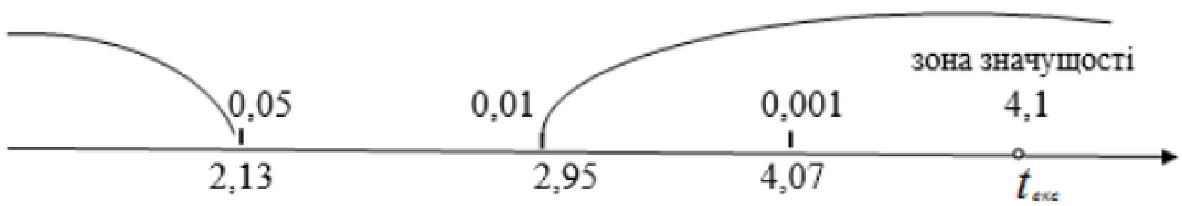
Тоді $t_{\text{екс}} = \frac{|\overline{x} - \overline{y}|}{\sigma} = \frac{|526 - 638|}{27{,}4} = \frac{112}{27{,}4} = 4.1$

Число ступенів свободи df=9+8-2=15

За таблицею критичних значень коефіцієнта Стьюдента (t-критерія) для даного числа ступенів свободи знаходимо:

$$t_{\text{кр}} = \begin{cases} 2{,}13 \text{ для } p \leq 0{,}05 \\ 2{,}95 \text{ для } p \leq 0{,}01 \\ 4{,}07 \text{ для } p \leq 0{,}001 \end{cases}$$

Будуємо вісь значущості:



Висновок: виявлені відмінності між експериментальними і контрольними групами значущі більш, ніж на 0,1% рівні, тобто середня швидкість складної сенсомоторної реакції вибору в групі спортсменів значно вища, ніж в групі людей, які не займаються спортом. Тож, гіпотеза H_0 про схожість відхиляється і на 0,1% рівні значущості приймається альтернативна гіпотеза H_1 – про відмінності між експериментальною і контрольною групами.

Задача 2. При дослідженні психомоторики використовується параметр: час простої реакції (ЧПР). Чи пов'язані ЧПР з успішністю роботи на автотранспорті?

Для дослідження обрали дві групи водіїв:

I група – водії, які протягом року не здійснювали жодної аварії;

II група – водії, які протягом року здійснили дві і більше аварій.

Параметри: I група: $\overline{x_1} = 200$ м/с; σ_1=17 м/с; n_1=20

II група: $\overline{x_2} = 260$ м/с; σ_2=25 м/с; n_2=15

$\overline{x}$ – середнє значення;

σ – середнє квадратичне відхилення;

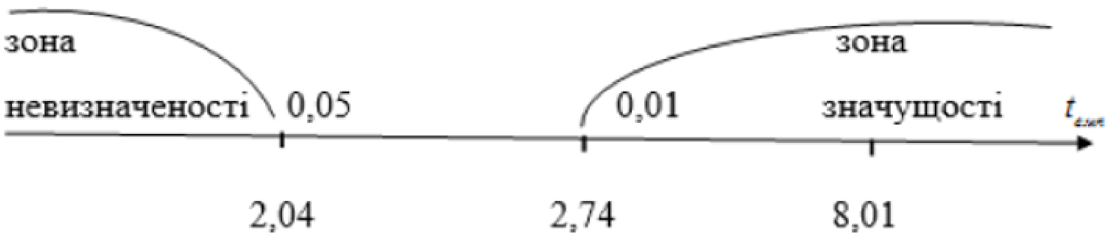
n – обсяг вибірки.

$$t_{\text{емп}} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{60}{\sqrt{\frac{289}{20} + \frac{625}{15}}} \approx 8{,}01$$

Число ступенів свободи df=20+15-2=33.

Критичні значення коефіцієнта Стьюдента знаходимо за таблицею:

$$t_{\text{кр}} = \begin{cases} 2{,}04 \text{ для р} \leq 0{,}05 \\ 2{,}74 \text{ для р} \leq 0{,}01 \\ 3{,}62 \text{ для р} \leq 0{,}001 \end{cases}$$



Висновок: Оскільки $t_{\text{емп}} > t_{\text{кр}}$ (p=0,001), тому І група водіїв суттєво відрізняється за ЧПР від ІІ групи. Тому ЧПР є показником успішності роботи на автотранспорті і може являтися критерієм відбору для водіїв таксі.

4. Статистичний критерій t-Стьюдента для залежних вибірок

Призначення критерію: перевірка гіпотези про достовірність різниці середніх на вибірках із розподілом, близьким до нормального.

Область дії: залежні вибірки („до – після”).

Формулювання гіпотез:

H_0 : значення ознаки між фоновим рівнем і рівнем сформованості ознаки після впливу незалежної змінної на вибірці не відрізняється.

H_1 : значення ознаки у вибірці до та після впливу незалежної змінної відрізняються.

Оцінювання результатів:

А) Якщо результат $t_{\text{емп}}$, одержаний при обчисленні значення критерію Стьюдента, ***менше***, ніж значення $t_{\text{ст}}$ для рівня значущості p=0,05 та ступенів свободи $(n_1+n_2)-2$, то приймаємо нульову гіпотезу: значення ознаки до та після впливу не відрізняється; вибірки із однієї популяції; вплив на вибірку Е випадковий;

Б) Якщо результат $t_{емп}$, одержаний при обчисленні значення критерію Стьюдента, ***більше,*** ніж значення $t_{ст}$ для рівня значущості p=0,01 та ступенів свободи $(n_1+n_2)-2$, то приймаємо альтернативну гіпотезу: значення ознаки між фоном та даними після експерименту відрізняються; вибірки належать до різних популяцій; вплив на вибірку не випадковий.

Коефіцієнт **Стьюдента** обчислюють за формулою:

$$t_{екс} = \frac{\overline{d}}{Sd}$$

$$\overline{d} = \frac{\sum d}{n} = \frac{\sum(x-y)}{n}$$

де $d = x - y$ – різниця між відповідними значеннями змінних x і y;

$\overline{d}$ – середнє значення цих різниць;

$$Sd = \sqrt{\frac{\sum d^2 - \frac{(\sum \overline{d})^2}{n}}{n(n-1)}}$$

Число ступенів свободи df=n-1

Задача 3. Психолого висловив припущення, що в результаті навчання час розв'язання еквівалентних задач в «15» буде значно зменшуватися. Для перевірки гіпотези у 8 осіб порівняли час розв'язування (в хв.) І і ІІІ задач.

Учасник експерименту	І задача	ІІІ задача	$\overline{d}$	d^2
1	4	3	1	1
2	3,5	3	0,5	0,25
3	4,1	3,8	0,3	0,09
4	5,5	2,1	3,4	11,56
5	4,6	4,9	-0,3	0,09
6	6	5,3	0,7	0,49
7	5,1	3,1	2	4
8	4,3	2,7	1,6	2,56
Сума	**37,1**	**27,9**	**9,2**	**20,04**

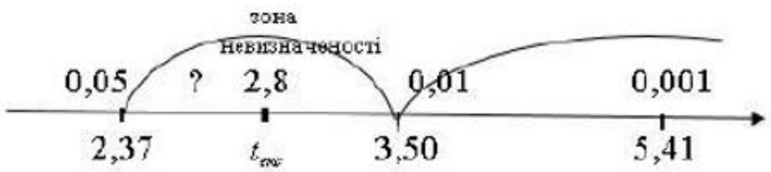
$$\overline{d} = \frac{\sum(x-y)}{n} = \frac{9,2}{8} = 1,15$$

$$Sd = \sqrt{\frac{\sum d^2 - \frac{(\sum \overline{d})^2}{n}}{n(n-1)}} = \sqrt{\frac{20{,}04 - \frac{(9{,}2)^2}{8}}{8(8-1)}} = \sqrt{\frac{20{,}04 - 10{,}58}{56}} = 0{,}41$$

Число ступенів свободи df=8-1=7

$$t_{\text{екс}} = \frac{\overline{d}}{Sd} = \frac{1{,}15}{0{,}41} = 2.8$$

$$t_{\text{кр}} = \begin{cases} 2{,}37 \text{ для p} \leq 0{,}05 \\ 3{,}50 \text{ для p} \leq 0{,}01 \\ 5{,}41 \text{ для p} \leq 0{,}001 \end{cases}$$



Висновок: на 5% рівні гіпотеза *H₀* відхиляються, *H₁* – приймаються (про відмінності).

Задача 4. На курсах післядипломної пед. освіти вчителі опановували прийоми мнемотехніки. Щоб перевірити ефективність навчання, проводиться тестування перед початком курсу і по його закінченню. Планувалося запам'ятати на слух 20 слів, фіксували кількість правильно повторених слів.

№	x_1 (до)	x_2 (після)	$d_1=x_2-x_1$	d^2
1	10	18	8	64
2	6	10	4	16
3	7	7	0	0
4	5	8	3	9
5	8	12	4	16
6	4	7	3	9
7	7	15	8	64
8	6	14	8	64
9	6	11	5	25
10	8	13	5	25
Сума	67	115	48	292

1) $\overline{d} = \frac{115-67}{10} = 4{,}8$ – середнє значення різниці
2) Знаходимо стаендартне відхилення для різниць

$$d = \sqrt{\frac{\sum d^2 - \frac{(\sum \overline{d})^2}{n}}{n(n-1)}} = \sqrt{\frac{292 - \frac{48^2}{10}}{10(10-1)}} = \sqrt{\frac{292 - 230.4}{90}} = 0{,}83$$

3) Розрахуємо $t_{\text{екс}} = \frac{\overline{d}}{Sd} = \frac{4{,}8}{0{,}83} = 5{,}8$

4) Знаходимо число ступенів свободи df=n-1=10-1=9

5) Визначимо $t_{\text{кр}}$ за таблицею

$$t_{\text{кр}} = \begin{cases} 2{,}26 \text{для р} \leq 0{,}05 \\ 3{,}25 \text{ для р} \leq 0{,}01 \\ 4{,}78 \text{ для р} \leq 0{,}001 \end{cases}$$

Висновок: Гіпотеза H_0 про схожість відхиляється і на 0,1% рівні значущості приймаємо альтернативну гіпотезу H_1 – про відмінності в показниках до початку експерименту і після, таким чином отримане емпіричне значення $t_{\text{емп}}$ перевищує табличне, тому робимо висновок про ефективність навчання.

КРИТИЧНІ ЗНАЧЕННЯ Т-КРИТЕРІЯ СТЬЮДЕНТА

Відмінності між двома вибірками можна вважати достовірними, якщо $t_{емп} \geq t_{0,05}$, і тим більш достовірними, якщо $t_{емп} \geq t_{0,01}$

Число степеней свободы K	p			Число степеней свободы K	p		
	0,05	0,01	0,001		0,05	0,01	0,001
1	12 71	63 66	64 60	18	2 10	2 88	3 92
2	4 30	9 92	31 60	19	2 09	2 86	3 88
3	3 18	5 84	12 92	20	2 09	2 85	3 85
4	2 78	4 60	8 61	21	2 08	2 83	3 82
5	2 57	4 03	6 87	22	2 07	2 82	3 79
6	2 45	3 71	5 96	23	2 07	2 81	3 77
7	2 37	3 50	5 41	24	2 08	2 80	3 75
8	2 31	3 36	5 04	25	2 06	2 79	3 73
9	2 26	3 25	4 78	26	2 06	2 78	3 71
10	2 23	3 17	4 59	27	2 05	2 77	3 69
11	2 20	3 11	4 44	28	2 05	2 76	3 67
12	2 18	3 05	4 32	29	2 05	2 76	3 66
13	2 16	3 01	4 22	30	2 04	2 75	3 65
14	2 14	2 98	4 14	40	2 02	2 70	3 55
15	2 13	2 95	4 07	60	2 00	2 66	3 46
16	2 12	2 92	4 02	120	1 98	2 62	3 37
17	2 11	2 90	3 97	∞	1 96	2 58	3 29
P	0 05	0 01	0 001	–	0 05	0 01	0 001

Додаток 1. Критичні значення коефіцієнта Стьюдента (t-критерія) для ступенів свободи та рівнів значущості

η	*P (рівні значущості)*							
	0.2	**0.1**	**0.05**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	3.0770	6.3130	12.7060	31.820	63.656	127.656	318.306	636.619
2	1.8850	2.9200	4.3020	6.964	9.924	14.089	22.327	31.599
3	1.6377	2.35340	3.182	4.540	5.840	7.458	10.214	12.924
4	1.5332	2.13180	2.776	3.746	4.604	5.597	7.173	8.610
5	1.4759	2.01500	2.570	3.649	4.0321	4.773	5.893	6.863
6	1.4390	1.943	2.4460	3.1420	3.7070	4.316	5.2070	5.958
7	1.4149	1.8946	2.3646	2.998	3.4995	4.2293	4.785	5.4079
8	1.3968	1.8596	2.3060	2.8965	3.3554	3.832	4.5008	5.0413
9	1.3830	1.8331	2.2622	2.8214	3.2498	3.6897	4.2968	4.780
10	1.3720	1.8125	2.2281	2.7638	3.1693	3.5814	4.1437	4.5869
11	1.363	1.795	2.201	2.718	3.105	3.496	4.024	4.437
12	1.3562	1.7823	2.1788	2.6810	3.0845	3.4284	3.929	4.178
13	1.3502	1.7709	2.1604	2.6503	3.1123	3.3725	3.852	4.220
14	1.3450	1.7613	2.1448	2.6245	2.976	3.3257	3.787	4.140
15	1.3406	1.7530	2.1314	2.6025	2.9467	3.2860	3.732	4.072
16	1.3360	1.7450	2.1190	2.5830	2.9200	3.2520	3.6860	4.0150