ЛЕКЦІЯ 5

ДОСТОВІРНІСТЬ РІЗНИЦЬ СЕРЕДНІХ ТА ПЕРЕВІРКА ЗВ'ЯЗКУ МІЖ ЗМІННИМИ

Мета:

Навчальна: Навчитися встановлювати достовірність різниць між середніми показниками вибірок та виявляти наявність можливих зв'язків між змінними.

Розвиваюча: Удосконалити навички первинного аналізу даних.

Виховна: Виховати уміння робити аналітико-логічні висновки.

План

1. Параметричні критерії перевірки гіпотез про середні і дисперсії;
   a) *критерій Фішера,*
   b) *t-критерій Стьюдента.*
2. Непараметричні критерії перевірки достовірності різниць;
   a) поняття рангів,
   b) критерій Уілкоксона,
   c) критерій Манна-Уітні.
3. Параметрична кореляція (коефіцієнт кореляції Пірсона).
4. Поняття про приватну кореляцію.
5. Непараметричні методи кореляції (коефіцієнт кореляції Спірмена).

Питання для самостійного опрацювання

1. Аналіз однорідності дисперсій за Кохреном.
2. Критерій Бартлета.
3. Критерій знаків.
4. Критерій Зігеля-Тьюкі.
5. Критерій Краскела-Уоліса.
6. Коефіцієнт кореляції Кендала.
7. Конкордація Кендала.
8. Перевірка рівності декількох середніх та фактору впливу на змінні методом дисперсійного аналізу.
9. Таблиці спряження.

Параметричні критерії перевірки гіпотез про середні і дисперсії

Параметричними називаються такі методи дослідження, у яких усі дані, що входять до вибірок чи генеральних сукупностей попадають під закон нормального розподілу.

При параметричних методах дослідження оперують та порівнюють між собою, в основному, показники дисперсій (або середньоквадратичних відхилень) та середньоарифметичні значення (також використовують інші середні величини: геометричні, гармонійні, квадратичні та кубічні).

Критерій Фішера

$$F_{розр} = \sigma_1{}^2 / \sigma_2{}^2$$

$F_{розр}$ - обов’язково повинно бути **більше** одиниці

Тому, або на початку **більшу** дисперсію ділять на **меншу**, або якщо результат уже відомий і він менший від одиниці (наприклад 0,843), то одиницю **ділять на результат** (в нашому випадку **1/0,843**).

Критичне значення критерію Фішера перевіряють за спеціальними таблицями. Причому для перевірки беруться ступені вільності чисельника і знаменника окремо ($n_{чис}$-1, $n_{знам}$-1).
*Якщо розрахункове (емпіричне) значення критерію Фішера (при заданому ступені вірогідності) **більше** за табличне (критичне), то дисперсії **різні**.*

Приклад.
Дисперсія такого показника як стресостійкість для учителя фізкультури в школі складала 6,17 (n_1=32), а для спортивного менеджера 4,41 (n_2=33). Дані в обох вибірках попадали під закон нормального розподілу.
Чи можна вважати, що рівень дисперсій у них приблизно однаковий?

Розв'язок.
Нульова гіпотеза. Дисперсії вибірок не відрізняються.
Визначаємо емпіричне (розрахункове) значення критерію за Фішером.

$$F_{\text{емп}} = \frac{\sigma_1{}^2}{\sigma_2{}^2} = \frac{6{,}17}{4{,}41} \approx \mathbf{1{,}4}$$

Перевіривши критичне значення за таблицею побачимо, що при рівні значимості 0,05 критичним значенням для даних вибірок з відповідними ступенями вільності (n_1-1=31 та n_2-1=32) буде $F_{кр}.(0{,}05;\ 31;\ 32)=1{,}81$.
Таким чином, $F_{емп.}<F_{кр.}$, тому дисперсії у вибірках вважатимуться рівними. Тобто, нульова гіпотеза приймається.

t-критерій Стьюдента

Особливості застосування:

1. Усі вибірки мають **нормальний** розподіл
2. Якщо застосовується t-критерій Стьюдента для **незалежних** вибірок, то потрібно, щоб дисперсії були **рівними**.
3. Потрібно враховувати **розмір** та співвідношення розмірів вибірок.

Випадки, які потрібно враховувати:

1. Об’єм вибірки
   a) обидві групи великі (n>30)
   b) обидві групи малі (n<30)
   c) одна група велика, інша – мала
2. За складом
   a) вибірки залежні
   b) вибірки незалежні
3. Критерій стійкий і при малих відхиленнях від нормального розподілу

Для незалежних вибірок та рівних дисперсій (гомоскедастичний тест)

Для 2-х великих, або великої та малої вибірок:

$$t_{розр} = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}} \quad \text{або} \quad t_{розр} = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sqrt{m_1^{\ 2} + m_2^{\ 2}}}$$

де:

$\overline{X}_1$ та $\overline{X}_2$ - середні арифметичні 1-ї та 2-ї вибірок

$\sigma_1^{\ 2}$ та $\sigma_2^{\ 2}$ - дисперсії 1-ї та 2-ї вибірок

$m_1^{\ 2}$ та $m_2^{\ 2}$ - похибки середніх арифметичних 1-ї та 2-ї вибірок

n_1 та n_2 - об’єми 1-ї та 2-ї вибірок

Критичне значення t-критерію Стьюдента перевіряють за спеціальними таблицями. Для перевірки беруться ступені вільності, які розраховуються за формулою:

$k=n_1+n_2-2$, або $k=n_1-1+n_2-1$.

*Якщо розрахункове (емпіричне) значення t-критерію Стьюдента (при заданому ступені вірогідності) **більше** за табличне (критичне) то середні **різні**.*

Для 2-х малих вибірок:

$$t_{\text{розр}} = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\left[\sum\left((x_{1_i} - \bar{X}_1\right)^2 + \sum\left(x_{2_i} - \bar{X}_2\right)^2\right] \cdot (n_1 + n_2)}} \cdot \sqrt{n_1 \cdot n_2(n_1 + n_2 - 2)}$$

де:

$\bar{X}_1$ та $\bar{X}_2$ - середні арифметичні 1-ї та 2-ї вибірок

σ_1^2 та σ_2^2 - дисперсії 1-ї та 2-ї вибірок

n_1 та n_2 - об’єми 1-ї та 2-ї вибірок

x_{1_i} та x_{2_i}- кожне із значень 1-ї та 2-ї вибірок

Критичне значення t-критерію Стьюдента перевіряють за спеціальними таблицями. Для перевірки беруться ступені вільності, які розраховуються за формулою:

$k=n_1+n_2-2$, або $k=n_1-1+n_2-1$.

*Якщо розрахункове (емпіричне) значення t-критерію Стьюдента (при заданому ступені вірогідності) **більше** за табличне (критичне) то середні **різні**.*

Для залежних вибірок без уяви про дисперсії

Даний критерій досить точний і може застосовуватися як для великих так і малих за розміром вибірок

$$t_{\text{розр}} = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\sum\left[(x_{1_i} - x_{2_i}) - (\bar{X}_1 - \bar{X}_2)\right]^2}} \cdot \sqrt{n \cdot (n-1)}$$

де:

$\bar{X}_1$ та $\bar{X}_2$ - середні арифметичні 1-ї та 2-ї вибірок

$\sigma_1^{\ 2}$ та $\sigma_2^{\ 2}$ - дисперсії 1-ї та 2-ї вибірок

n_1 та n_2 - об'єми 1-ї та 2-ї вибірок

x_{1_i} та x_{2_i}- кожне із значень 1-ї та 2-ї вибірок

Критичне значення t-критерію Стьюдента перевіряють за спеціальними таблицями. Для перевірки беруться ступені вільності, які розраховуються за формулою:

$k=n_1+n_2-2$, або $k=n_1-1+n_2-1$.

*Якщо розрахункове (емпіричне) значення t-критерію Стьюдента (при заданому ступені вірогідності) **більше** за табличне (критичне) то середні **різні**.*

Для незалежних вибірок нерівних дисперсій (гетероскедастичний тест)

На даний момент даний вид порівняння середніх має лише наближене вирішення і в математиці називається проблемою Беренса-Фішера.
Можна наближено застосувати формулу (як і для рівних дисперсій):

$$t_{розр} = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}}} \quad \text{або} \quad t_{розр} = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sqrt{m_1{}^2 + m_2{}^2}}$$

де:

$\overline{X}_1$ та $\overline{X}_2$ - середні арифметичні 1-ї та 2-ї вибірок

$\sigma_1{}^2$ та $\sigma_2{}^2$ - дисперсії 1-ї та 2-ї вибірок

$m_1{}^2$ та $m_2{}^2$ - похибки середніх арифметичних 1-ї та 2-ї вибірок

n_1 та n_2 - об’єми 1-ї та 2-ї вибірок

Критичне значення t-критерію Стьюдента перевіряють за спеціальними таблицями. Для перевірки беруться ступені вільності, які розраховуються за формулою:

$$k = \frac{\left(\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}\right)^2}{\frac{\sigma_1{}^2}{n_1 - 1} + \frac{\sigma_2{}^2}{n_2 - 1}}.$$

*Якщо розрахункове (емпіричне) значення t-критерію Стьюдента (при заданому ступені вірогідності) **більше** за табличне (критичне) то середні **різні**.*

Для порівняння із заданим значенням

Якщо дані попадають під закон нормального розподілу і їх потрібно порівняти із якоюсь стандартною величиною, то використовують наступну формулу:

$$t_{розр} = \frac{\left|\overline{X} - A\right| \cdot \sqrt{n}}{\sigma^2}$$

де:

A - величина з якою порівнюється середнє арифметичне вибірки

$\overline{X}$ - середнє арифметичне вибірки

σ^2 - дисперсія вибірки

Критичне значення t-критерію Стьюдента перевіряють за спеціальними таблицями. Для перевірки беруться ступені вільності вибірки, якій розраховується за формулою:

k=n-1.

*Якщо розрахункове (емпіричне) значення t-критерію Стьюдента (при заданому ступені вірогідності) **більше** за табличне (критичне) то середні **різні**.*

Приклад.

Проводилися дослідження серед 16-річних школярів-хлопчиків з бігу на 100 м. В школі №1 середній результат серед 36 юнаків був 14,8±1,7 с, а в школі №2 серед 38 юнаків результат був 15,1±1,6 с.

Чи можна вважати, що в школі №1 результати з бігу кращі ніж у школі №2 (дані попадали під закон нормального розподілу)?

Розв'язок.

Нульова гіпотеза. Різниці у показниках немає.

Оскільки кількість даних більша за 30, дані у вибірках незалежні, то задача на t-критерій Стьюдента для 2-х великих вибірок.

Але виникає питання чи рівні у цих вибірках дисперсії, від цього буде залежати та чи інша формула (а скоріше формула розрахунку ступенів вільності)?

Дисперсію ми можемо визначити, виходячи із значення похибки середнього арифметичного. $m = \frac{\sigma}{\sqrt{n}}$. Звідси $\sigma = m \cdot \sqrt{n}$. Підставивши дані з першої та другої вибірок ми знайдемо, що $\sigma_1 = 1{,}7 \cdot \sqrt{36} = 10{,}2$ с, а $\sigma_2 = 1{,}6 \cdot \sqrt{38} \approx 9{,}86$ с.

Провівши порівняння дисперсій за критерієм Фішера (F-тест) визначимо: $F_{розр} = \sigma_1{}^2 / \sigma_2{}^2$

Або $F_{розр} = \frac{10{,}2^2}{9{,}86^2} \approx 1{,}07$

Перевіривши критичне значення за таблицею побачимо, що при рівні значимості 0,05 критичним значенням для даних вибірок з відповідними ступенями вільності (n_1-1=35 та n_2-1=37) буде $F_{кр.}(0{,}05;\ 35;\ 37) = 1{,}74$.

Таким чином, $F_{емп.} < F_{кр.}$, тому дисперсії у вибірках вважатимуться рівними.

Отже, вибірки незалежні, дисперсії рівні. Значить підходить формула:

$$t_{розр} = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sqrt{m_1^{\ 2} + m_2^{\ 2}}}$$

А ступені вільності обрахуємо за формулою:

$$k=n_1+n_2-2$$

Тобто $t_{розр} = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sqrt{m_1^{\ 2} + m_2^{\ 2}}} = \frac{\left|14{,}8 - 15{,}1\right|}{\sqrt{1{,}7^2 + 1{,}6^2}} \approx 0{,}13$, а k=38+36-2=72

За таблицею для даних ступенів вільності критичне значення буде:
$t_{кр}$.(0,05; 37; 35)=1,99.

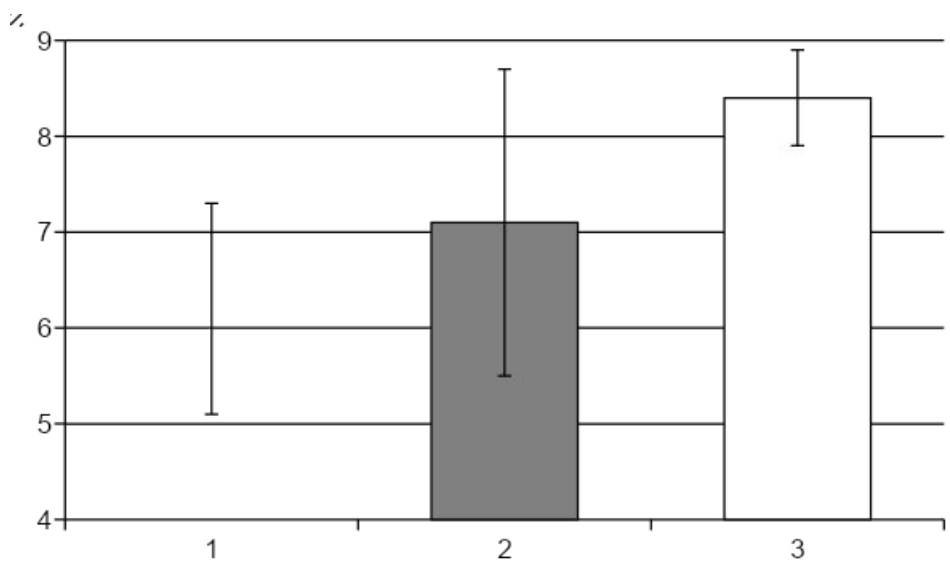
Оскільки розрахункове значення (емпіричне) менше за критичне (табличне), тобто $t_{розр.}<t_{кр.}$ то гіпотеза про рівність середніх **приймається**. Або, іншими словами, достовірних різниць між показниками бігу на 100 м в різних школах **не має**.

Можливі парадокси при перевірці гіпотез про середні

$$\overline{X}_1 = \overline{X}_2 \quad \text{і} \quad \overline{X}_2 = \overline{X}_3 \text{, але } \overline{X}_1 \neq \overline{X}_3$$

$$\overline{X}_1 = A \quad \text{і} \quad \overline{X}_2 = A \text{, але } \overline{X}_1 \neq \overline{X}_2$$

Наприклад:
6,2±1,1 ≈ 7,1±1,6. У свою чергу **7,1±1,6 ≈ 8,4±0,5**. Але **6,2±1,1≠8,4±0,5**



Для перевірки рівності **декількох** середніх потрібно користуватися **спеціальними** критеріями

Непараметричні критерії перевірки достовірності різниць

Поняття про ранги та ранжування

Ранг – це той номер значення у варіаційному ряду, який би воно (значення) отримало, якби ці значення були відсортовані (розставлені) по порядку.

Наприклад
Є ряд даних:
34, 17, 25, 43, 15, 26, 14, 20

Відсортуємо їх у порядку зростання
14, 15, 17, 20, 25, 26, 34, 43

Пронумеруємо ці дані у порядку зростання:

14,	**15,**	**17,**	**20,**	**25,**	**26,**	**34,**	**43**
1	***2***	***3***	***4***	***5***	***6***	***7***	***8***

Отримані номера (з ***1*** по ***8***) і будуть називатися **рангами,** а сам процес їх встановлення – **ранжуванням.**

Часто ранги можна і не визначати, а скористатися готовими експериментальними даними, якщо ці дані являють собою цілі числа і їх розмах однаковий для всіх порівнювальних груп.
Це можуть бути шкільні оцінки, суддівські бали, кількість вистрелів і т.п.

Правила ранжування

1. Меншому значенню нараховується менший ранг. Найменшому значенню нараховується ранг 1. Найбільшому значенню нараховується ранг, що відповідає кількості значень, які ранжуються.

   Наприклад, якщо n=7, то найбільше значення отримає ранг 7, за можливим виключенням для тих випадків, які передбачені правилом 2.

2. У разі, якщо декілька значень рівні, їм нараховується ранг, що є середнім значенням з тих рангів, які вони отримали б, якби не були рівні.

   Наприклад, 3 найменші значення дорівнюють 10 секундам. Якби ми вимірювали час точніше, то ці значення могли б розрізнятися і складали б, скажімо, 10,2 с; 10,5 с; 10,7 с. В цьому випадку вони отримали б ранги, відповідно, 1, 2 і 3. Але оскільки отримані нами значення рівні, кожне з них отримує середній ранг: **(1 + 2 + 3)/3 = 6/3 = 2.** Допустимо, наступні 2 значення дорівнюють 12 сек. Вони повинні були б отримати ранги 4 і 5, але, оскільки вони рівні, то отримують середній ранг: **(4 + 5) / 2 = 4,5**

3. Загальна сума рангів повинна співпадати з розрахунковою, яка визначається по формулі: $\boldsymbol{\Sigma_{\textit{рангів}} = (n^2+n)/2}$,

   де ***n*** - загальна кількість спостережень (значень), що ранжуються.

   Неспівпадання реальної і розрахункової сум рангів свідчить про помилку, допущену при нарахуванні рангів або їх підсумовуванні.

Критерій Уілкоксона (Вілкоксона, W-критерій)

Застосовується для аналізу достовірності різниць переважно у **залежних** вибірках, коли немає відомостей про рівність дисперсій.

Вимоги:

1. Шкала вимірювання повинна бути порядковою, інтервальною або відносною (тобто критерій не можна застосовувати до номінальних змінних).
2. Досліджувані значення повинні бути неперервними і симетричними відносно своєї медіани
3. Кількість значень для аналізу повинна бути не менше 5.

Послідовність виконання розрахунків:

1. Скласти список досліджуваних вимірів у будь-якому порядку, наприклад алфавітному (повинно вийти 2 стовпчика: до дослідження та після).
2. Вирахувати різницю між індивідуальними значеннями у другому і першому вимірюваннях («після» - «до»). Якщо в різниця рівна **0**, то її до уваги не беруть.
3. Проранжувати ці дані не звертаючи уваги на знаки (по модулю).
4. Просумувати ранги окремо для позитивних і негативних різниць.
5. Яка сума буде меншою, ту й вважають емпіричним (експериментальним) критерієм Уілкоксона.
6. Якщо емпіричне значення **більше** критичного (табличного) то різниця **недостовірна**.

Приклад

Вивчалася кількість попадань у мішень з 10 патронів при кульовій стрільбі. Обстежувані стріляли в сонячну погоду і, наступного дня, в похмуру. Дані занесено в таблицю:

Погода	**Кількість попадань**								
Сонячна	5	7	8	2	5	4	1	4	9
Похмура	8	7	5	5	4	2	7	8	9

Чи можна вважати, що дані у сонячну погоду відрізнялися від даних отриманих у похмуру?

Розв'язок.

Нульова гіпотеза: Різниць немає.

1. Розраховуємо різницю між окремими показниками в сонячну погоду та, відповідно, похмуру.

Погода	**Кількість попадань**								
Сонячна	5	7	8	2	5	4	1	4	9
Похмура	8	7	5	5	4	2	7	8	9
Різниця	***-3***	***0***	***3***	***-3***	***1***	***2***	***-6***	***-4***	***0***

2. Відкидаємо дані з нульовими значеннями

Погода	**Кількість попадань**								
Сонячна	5	7	8	2	5	4	1	4	9
Похмура	8	7	5	5	4	2	7	8	9
Різниця	***-3***	***0***	***3***	***-3***	***1***	***2***	***-6***	***-4***	***0***

Погода	**Кількість попадань**						
Сонячна	5	8	2	5	4	1	4
Похмура	8	5	5	4	2	7	8
Різниця	***-3***	***3***	***-3***	***1***	***2***	***-6***	***-4***

3. Сортуємо дані **за різницею** від найменшого до найбільшого **без врахування знаку** (по модулю)

Погода	**Кількість попадань**						
Сонячна	5	4	5	8	2	4	1
Похмура	4	2	8	5	5	8	7
Різниця	***1***	***2***	***-3***	***3***	***-3***	***-4***	***-6***

4. Назначаємо ранги відсоротованим **різницям**

Погода	**Кількість попадань**						
Сонячна	5	4	5	8	2	4	1
Похмура	4	2	8	5	5	8	7
Різниця	***1***	***2***	***-3***	***3***	***-3***	***-4***	***-6***
Ранги	***1***	***2***	***4***	***4***	***4***	***6***	***7***

5. розраховуємо суму окремо рангів для позитивних і негативних значень

$\boldsymbol{\Sigma W_{+}=1+2+4=7}$, $\boldsymbol{\Sigma W_{-}=4+4+6+7=21}$

6. Для порівняння беремо менше із значень сум рангів. В даному випадку це **7**

7. Перевіряємо його із табличним (критичним значенням). Для нашого випадку для 7 значень критичне значення буде рівне **2.**

8. **Відповідь.** Оскільки емпіричне (розрахункове, експериментальне) значення більше за табличне (критичне), то нульова гіпотеза приймається, тобто середні значення у двох вибірках суттєво **не відрізняються**.

Критерій Манна-Уітні (U-критерій, Уілкоксона-Манна-Уітні)

Застосовується для аналізу достовірності різниць переважно у **незалежних** вибірках, коли немає відомостей про рівність дисперсій. Вважається найпотужнішим із непараметричних критеріїв.

Вимоги:

1. Шкала вимірювання повинна бути порядковою, інтервальною або відносною (тобто критерій не можна застосовувати до номінальних змінних).
2. Досліджувані значення повинні бути повинні бути неперервними і симетричними відносно своєї медіани
3. Кількість значень для аналізу у меншій вибірці повинна бути не менше 5.

Послідовність виконання розрахунків:

1. Скласти список досліджуваних вимірів у будь-якому порядку в один ряд або стовпчик наприклад алфавітному (повинно вийти 2 стовпчика).
2. Об'єднати 2 стовпчика в один.
3. Відсортувати дані у стовпчику за зростанням.
4. Проранжувати ці дані.
5. Провести розрахунки за формулами: $U_1=n_1n_2+0{,}5n_1(n_1+1)-R_1$, $U_2=n_1n_2+0{,}5n_2(n_2+1)-R_2$ де R_1 і R_2 - суми рангів, розрахованих для значень, які належать першій і другій вибіркам відповідно, n_1 і n_2 – об'єми першої та другої вибірок.
6. Яке значення U_1 чи U_2 буде меншим, те значення вважають емпіричним (експериментальним) критерієм Манна-Уітні.
7. Якщо емпіричне значення **більше** критичного (табличного) то різниця **недостовірна**.

Приклад

Аналізувалися показники часу затримки дихання на вдиху (в секундах) у хлопців-баскетболістів та хлопців-футболістів. Дані занесено в таблицю:

Спортсмени	Час затримки дихання (с)								
Баскетбол.	120	150	132	147	120	149	156	133	118
Футбол.	145	158	112	145	124	142	161		

Чи можна вважати, що середні значення між баскетболістами і футболістами відрізнялися?

Розв'язок.

Нульова гіпотеза: Різниць немає.

1. Об'єднуємо дані в одну спільну вибірку.

Об'єдн. дані	120	150	132	147	120	149	156	133	118	145	158	112	145	124	142	161

2. Відсортовуємо дані за зростанням

Об'єдн. дані	112	118	120	120	124	132	133	142	145	145	147	149	150	156	158	161

3. Присвоюємо даним відповідні ранги

Об'єдн. дані	112	118	120	120	124	132	133	142	145	145	147	149	150	156	158	161
Ранги	1	2	3,5	3,5	5	6	7	8	9,5	9,5	11	12	13	14	15	16

4. Підсумовуємо ранги з першої та другої вибірок.

$R_1=2+3,5+3,5+6+7+11+12+13+14=72$

$R_2=1+5+8+9,5+9,5+15+16=64$

5. Розраховуємо U_1 та U_2

$U_1=n_1n_2+0,5n_1(n_1+1)-R_1 =9\cdot 7+0,5\cdot 9\cdot (9+1)-72=36;$

$U_2=n_1n_2+0,5n_2(n_2+1)-R_2=9\cdot 7+0,5\cdot 7\cdot (7+1)-64=27.$

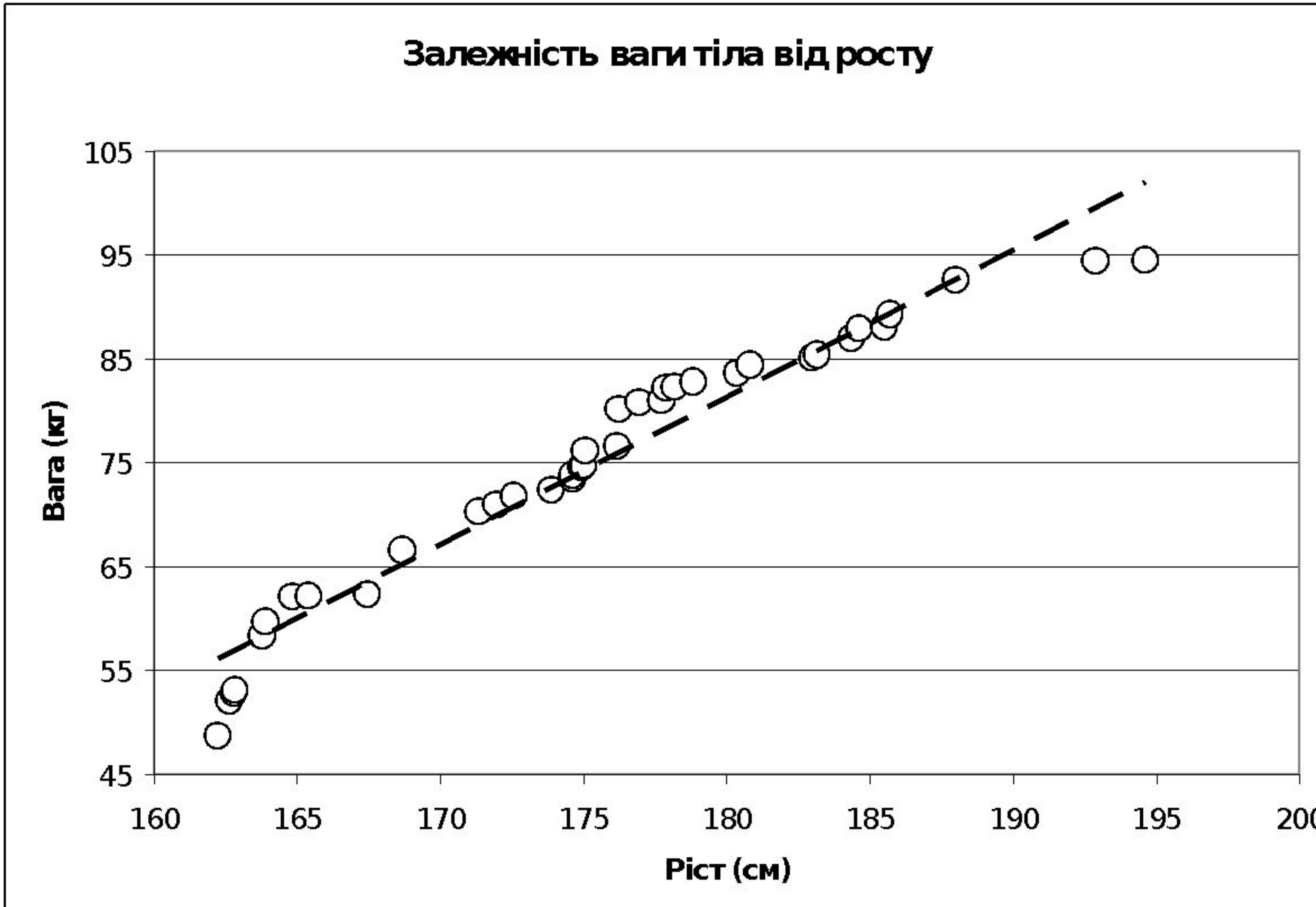
5. Для порівняння беремо менше із значень U_1 чи U_2. В нашому випадку це **27.**

6. Перевіряємо його із табличним (критичним даним). Для нашого випадку для **9** та **7** значень критичне значення буде рівне **12.**

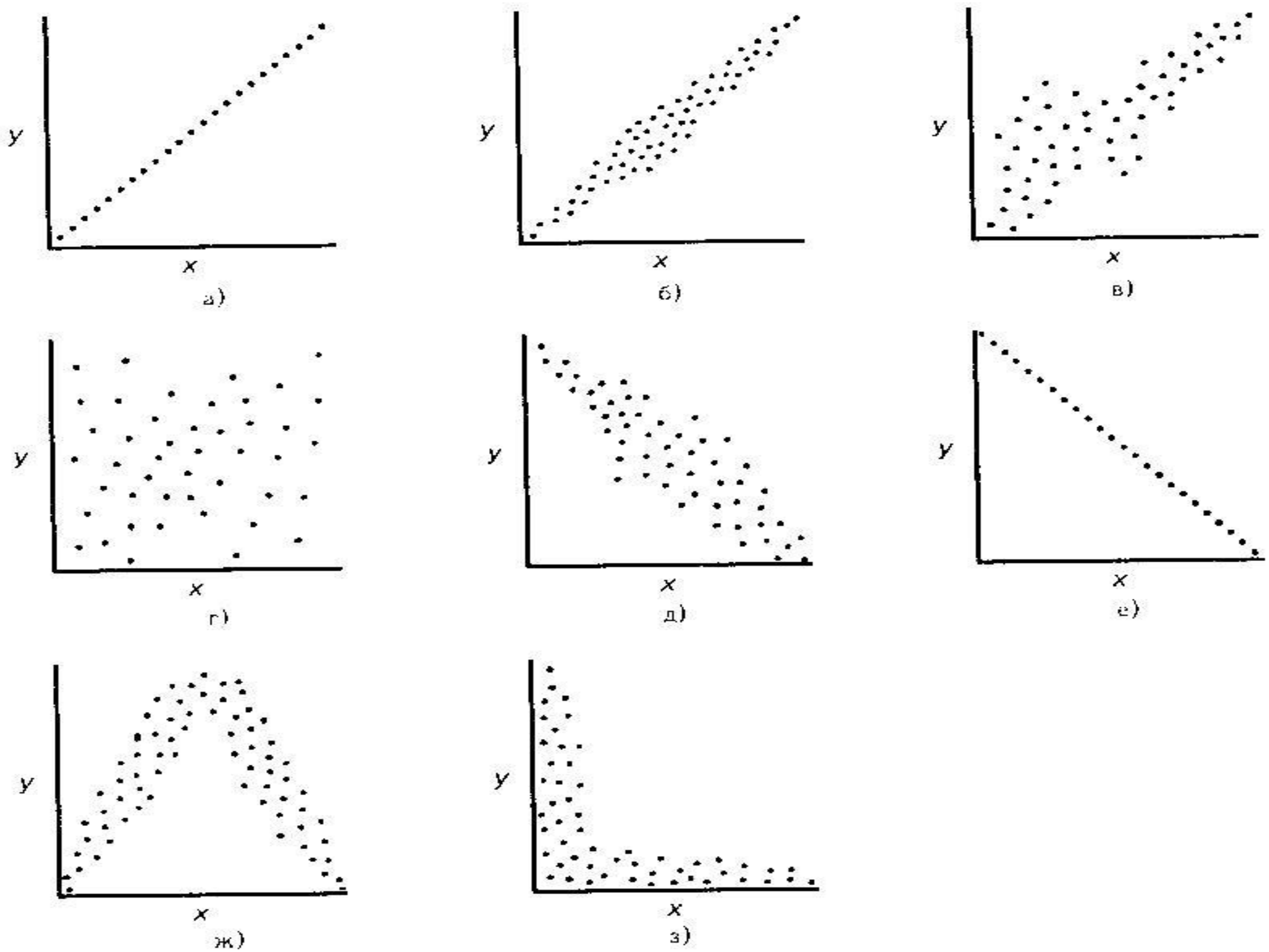
7. **Відповідь.** Оскільки емпіричне (розрахункове, експериментальне) значення більше за табличне (критичне), то нульова гіпотеза приймається, тобто середні значення у двох вибірках суттєво **не відрізняються**.

Кореляція

Кореляція (англ. correlation) в прямому перекладі означає "співвідношення". Якщо зміна однієї змінної супроводжується зміною іншої, то можна говорити про кореляцію цих змінних.



Види кореляційних зв’язків



Види кореляцій

а) строга позитивна, б) сильна позитивна, в) слабка позитивна,
г) нульова, д) сильна негативна, е) строга негативна,
ж) нелінійна, з) нелінійна.

Значення коефіцієнта кореляції **завжди** перебуває в межах від **-1** до **1**.
Якщо коефіцієнт кореляції зі знаком «-», то кореляція негативна (від'ємна, мінусова, обернена), якщо зі знаком «+», то позитивна (додатна, плюсова, пряма).

Коефіцієнт кореляції часто показує лише на залежність величини ***X*** від велечини ***Y***, але це не значить, що величина **Y** теж обов'язково буде залежати від величини **X**.

Кількість обстежуваних у обох вибірках повинна бути строго **однакова**. В противному випадку, або програма відмовиться рахувати, або недостаючі (порожні) дані будуть замінені нульовими значеннями, що призведе до значних викривлень у розрахунках.

Достовірність коефіцієнту кореляції залежить від об'єму вибірки. Чим **більша** за об'ємом вибірка, тим **менше** значення (за абсолютною величиною) коефіцієнту кореляції буде вважатися достовірним.

Параметрична кореляція (коефіцієнт кореляції Пірсона)

Умови застосування.

Лише тоді коли дані в **обох** вибірках попадають під закон нормального розподілу

$$r_{xy} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}}$$

де

X_i та Y_i - кожне значення із вибірки ***X*** та вибірки ***Y*** відповідно,

$\bar{X}$ та $\bar{Y}$ - середні арифметичні із вибірки ***X*** та вибірки ***Y*** відповідно.

Значимість коефіцієнта кореляції

$$t_{\text{кор}} = \frac{r_{xy}\sqrt{(n-2)}}{\sqrt{\left(1 - {r_{xy}}^2\right)}}$$

де

r_{xy} - коефіцієнт кореляції,

n – кількість обстежуваних у вибірці (об'єм вибірки).

Значення $t_{кор}$ перевіряється із табличним (критичним). Якщо розраховане (експериментальне) значення коефіцієнту кореляції **більше** за табличне (критичне), то така кореляція вважається **достовірною**.

Приватна кореляція

Кореляція між двома змінними, вирахувана після усунення впливу усіх інших змінних, називається **приватною кореляцією**.

Наприклад

Довжина волосся може корелювати з ростом людини (чим **вища** людина, тим **коротше** волосся), проте ця залежність стає слабкою або зовсім зникає, якщо усунути вплив **статі** спостережуваних людей, оскільки жінки зазвичай *нижче* ростом і частіше мають *довше* волосся, ніж чоловіки.

$$r_{xyz} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{(1 - r_{xz}{}^2)(1 - r_{yz}{}^2)}}$$

де

r_{xy} - коефіцієнт кореляції між вибірками ***X*** та ***Y***,

r_{xz} - коефіцієнт кореляції між вибірками ***X*** та **Z**,

r_{yz} - коефіцієнт кореляції між вибірками ***Y*** та ***Z***,

Необхідно пам'ятати:

- впливаючих змінних може бути не три, а скільки завгодно;
- ви можете не знати про всі впливаючі змінні;
- деякі автори стверджують, що для коректного використання приватного коефіцієнта кореляції необхідна наявність так званого багатомірного нормального закону розподілу

Рангова кореляція Спірмена (ρ-кореляція)

Умови застосування

1. Змінні мають рангову шкалу виміру;
2. Розподіл даних хоча б у одній вибірці занадто відрізняється від нормального або взагалі невідомий;
3. Вибірки мають невеликий об'єм (n<30).

$$\rho_{xy} = 1 - \frac{6\sum(R_{x_i} - R_{y_i})^2}{n^3 - n}$$

де

R_{x_i} та R_{y_i}- ранги кожного із значень з вибірки ***X*** та ***Y***

n – кількість обстежуваних у вибірці

Перевіряється достовірність коефіцієнту кореляції Спірмена аналогічно як і коефіцієнт кореляції Пірсона за відповідними таблицями.

Якщо вибірки попадали під закон нормального розподілу (або мали незначне відхилення) то із коефіцієнту кореляції Спірмена можна розрахувати коефіцієнт кореляції Пірсона за формулою:

$$r_{xy} = 2sin\frac{\pi}{6}\rho_{xy}$$

де

r_{xy} - коефіцієнт кореляції Пірсона

ρ_{xy} - коефіцієнт кореляції Спірмена

Приклад:

Визначити, чи залежить результат стрибка в довжину з розгону (ознака ***X***) від величини кінцевої швидкості розгону (ознака ***Y***).

№ учасника	1	2	3	4	5	6	7	8
X_i (см)	890	820	825	790	795	802	702	730
Yi (м/с)	10,7	10,5	10,1	9,8	10,1	10,5	9,1	9,6

Нульова гіпотеза: результат стрибка **не залежить** від швидкості розгону

Розв'язок

Даних мало. Отже під закон нормального розподілу вони не попадуть.

Тому для встановлення залежності застосуємо коефіцієнт кореляції Спірмена.

1. Замінимо всі дані на їхні ранги (проранжуємо дані за ознакою ***X*** та ознакою ***Y***).

№ учасника	1	2	3	4	5	6	7	8
X_i (см)	890	820	825	790	795	802	702	730
Yi (м/с)	10,7	10,5	10,1	9,8	10,1	10,5	9,1	9,6
Ранг Xi	8	6	7	3	4	5	1	2
Ранг Yi	8	6,5	4,5	3	4,5	6,5	1	2

2. У відповідності до формули знайдемо квадрат різниці рангів.

$$\rho_{xy} = 1 - \frac{6\sum(\boldsymbol{R_{x_i}} - \boldsymbol{R_{y_i}})^2}{n^3 - n}$$

№ учасника	1	2	3	4	5	6	7	8
X_i (см)	890	820	825	790	795	802	702	730
***Yi* (м/с)**	10,7	10,5	10,1	9,8	10,1	10,5	9,1	9,6
Ранг *Xi*	**8**	**6**	**7**	**3**	**4**	**5**	**1**	**2**
Ранг *Yi*	**8**	**6,5**	**4,5**	**3**	**4,5**	**6,5**	**1**	**2**
	0	0,25	6,25	0	0,25	2,25	0	0

$$\rho_{xy} = 1 - \frac{6\sum(R_{x_i} - R_{y_i})^2}{n^3 - n}$$

3. Знаходимо суму квадратів різниць рангів:
0+0,25+6,25+0+0,25+2,25+0+0=9

4. Розраховуємо власне коефіцієнт кореляції Спірмена:

$$\rho_{xy} = 1 - \frac{6 \cdot \mathbf{9}}{\mathbf{8}^3 - \mathbf{8}} = 1 - \frac{54}{512 - 8} \approx 0{,}89$$

5. Перевіряємо достовірність розрахованого нами коефіцієнту кореляції. Табличне значення для 8 досліджуваних 0,64.

6. **Відповідь.** Оскільки емпіричне (розрахункове, експериментальне) значення більше за табличне (критичне), то нульова гіпотеза відкидається (не приймається), тобто результат стрибка в довжину з розгону **залежить** від величини кінцевої швидкості розгону

ВИСНОВКИ

1. Параметричні методи можна застосовувати **лише** тоді, коли усі досліджувані дані або показники попадають під закон **нормального розподілу**.
2. У випадку коли є сумніви, що до відповідності розподілу даних за законом нормального розподілу, то в такому разі доцільніше використовувати **непараметричні** (рангові) методи аналізу даних.
3. Усі непараметричні методи аналізу базуються на перетворенні досліджуваних величин у рангові показники.

Вибір методу для рішення задач про порівняння параметрів розподілу вибірок

Формулювання задачі в прикладній постановці	Формулювання задачі в статистичній постановці	Додаткові умови		Метод, що застосовується
Порівняння показників контрольної і експериментальної вибірок	Перевірка гіпотези про рівність середніх (центрів розподілу[1]) в двох незалежних вибірках	**Нормальний** закон розподілу	Дисперсії вибірок рівні	t-критерій (Стьюдента) при рівних дисперсіях
			Дисперсії вибірок не рівні	t-критерій (Стьюдента) при нерівних дисперсіях
			Без уяви про дисперсії (але при однакових розмірах вибірок)	t-критерій (Стьюдента) без уяви про дисперсії
		Закон розподілу **відмінний від нормального,** або дані вимірюються **в дискретній шкалі**	Дисперсії вибірок рівні	Манна-Уітні (U-критерій Уілкінсона-Манна-Уітні)
			Без уяви про дисперсії	Двухвибірковий Уілкінсона, медіальний
Порівняння показників вибірок до і після експерименту	Перевірка гіпотези про рівність середніх в двох залежних вибірках	**Нормальний** закон розподілу		t-критерій (Стьюдента) для зв'язаних вибірок
		Закон розподілу **відмінний від нормального,** або дані вимірюються **в дискретній шкалі**		Знаковий одновиборочний критерій Уілкоксона

[1] В ряді випадків фактично перевіряється гіпотеза про рівність медіан або обох мір положення

<table>
<tr><th>Формулювання задачі в прикладній постановці</th><th>Формулювання задачі в статистичній постановці</th><th>Додаткові умови</th><th>Метод, що застосовується</th></tr>
<tr><td rowspan="2">Чи можна вважати, що середні значення показника дорівнюють деякому номінальному значенню</td><td rowspan="2">Перевірка гіпотези про рівність середнього константі</td><td>Нормальний закон розподілу</td><td>t-критерій (Стьюдента)</td></tr>
<tr><td>Закон розподілу відмінний від нормального, або дані вимірюються в дискретній шкалі</td><td>Гупта, знаковий</td></tr>
<tr><td rowspan="2">Порівняння розсіювання показника в двох вибірках</td><td rowspan="2">Перевірка гіпотези про рівність дисперсій (про належність дисперсій до однієї генеральної сукупності)</td><td>Нормальний закон розподілу</td><td>F-критерій (Фішера)</td></tr>
<tr><td>Закон розподілу відмінний від нормального, або дані вимірюються в дискретній шкалі</td><td>Зігеля-Тьюки, Мозеса</td></tr>
<tr><td rowspan="2">Чи можна вважати, що в кількох вибірках має місце одне і те ж значення показника?</td><td rowspan="2">Перевірка гіпотези про рівність дисперсій (про належність дисперсій до однієї генеральної сукупності)</td><td>Нормальний закон розподілу</td><td>G-критерій (Кохрена) при рівному розмірі вибірок, Барлета</td></tr>
<tr><td>Закон розподілу відмінний від нормального, або дані вимірюються в дискретній шкалі</td><td>Фрідмана</td></tr>
<tr><td rowspan="2">Чи можна вважати, що в кількох вибірках має місце одне і те ж значення розсіювання показника?</td><td rowspan="2">Перевірка гіпотези про рівність дисперсій (про належність середніх до однієї генеральної сукупності)</td><td>Нормальний закон розподілу</td><td>Шеффе, Діксона, дисперсійний аналіз, LSD</td></tr>
<tr><td>Закон розподілу відмінний від нормального, або дані вимірюються в дискретній шкалі</td><td>Краскела-Уолліса, медіальний, рангових сум Фрідмана</td></tr>
</table>

Вибір методу аналізу зв'язку між ознаками

Загальна кількість змінних	Шкали вимірювань		Закон розподілу	Метод
	Впливаючі змінні	Залежні змінні		
Дві	Інтервалів або відношень		Нормальний	Параметрична кореляція Персона
			Відмінний від нормального	Непараметрична кореляція Спірмена
	Хоча б одна шкала порядку		-	Непараметрична кореляція Спірмена або Кендалла
Три і більше	Порядку		-	Конкордація
Дві	Обидві шкали найменувань, одна виду „Є/Нема", друга має лише 2 значення		-	Чотирьох клітинні таблиці спряженості
	Обидві шкали найменувань, одна виду „Є/Нема", друга має К значень		-	Таблиці спряженості виду 2×К
	Обидві шкали найменувань, одна має К значень рівнів, друга – L		-	Таблиці спряженості виду К×L

<table>
<tr><th rowspan="2">Загальна кількість змінних</th><th colspan="2">Шкали вимірювань</th><th rowspan="2">Закон розподілу</th><th rowspan="2">Метод</th></tr>
<tr><th>Впливаючі змінні</th><th>Залежні змінні</th></tr>
<tr><td rowspan="3">Дві і більше</td><td rowspan="6">Найменувань</td><td rowspan="3">Інтервалів і відношень</td><td>Нормальний для залежної змінної</td><td>Параметричний дисперсійний аналіз (критерій Фішера)</td></tr>
<tr><td>Порядку для залежної змінної</td><td>Непараметричний дисперсійний аналіз Зігеля і Тьюкі</td></tr>
<tr><td>Відмінний від нормального для залежної змінної</td><td>Непараметричний дисперсійний аналіз Зігеля і Тьюкі</td></tr>
<tr><td rowspan="3">Три і більше</td><td>Порядку</td><td>-</td><td>Багатомірний непараметричний дисперсійний аналіз Фрідмана</td></tr>
<tr><td rowspan="2">Інтервалів або відношень</td><td>Відмінний від нормального для залежної змінної</td><td>Багатомірний непараметричний дисперсійний аналіз Фрідмана</td></tr>
<tr><td>Нормальний для залежної змінної</td><td>Багатомірний параметричний дисперсійний аналіз</td></tr>
</table>

Додаткова література

1. Гланц Стентон. Медико-биологическая статистика. – М.: Практика, 1999, - 459 с.
2. Годик М.А. Спортивная метрология. – М.: Физкультура и спорт, 1988, - 192 с.
3. Рокицкий П.Ф. Биологическая статистика. – Минск: Вышэйшая школа, 1973, 320 с.
4. Сидоренко Е.В. Методы математической обработки в психологии. – Санкт-Петербург: Социально-психологический центр, 1996, 350 с.