

**МІНІСТЕРСТВО ВНУТРІШНІХ СПРАВ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ВНУТРІШНІХ СПРАВ**

*Факультет № 6
Кафедра соціології та психології*

ТЕКСТ ЛЕКЦІЇ

з навчальної дисципліни

«Комп'ютерні методи практичної психології»
обов'язкових компонент освітньої програми
першого (бакалаврського) рівня вищої освіти

053 Психологія (практична психологія)

**Тема №6. Аналіз психологічних даних в SPSS:
параметричні критерії**

Харків 2023

ЗАТВЕРДЖЕНО

Науково-методичною радою
Харківського національного
університету внутрішніх справ
Протокол від 30.08.2023 р. №7

СХВАЛЕНО

Вченою радою факультету №6
Протокол від 25.08.2023 р. №7

ПОГОДЖЕНО

Секцією Науково-методичної
ради ХНУВС з гуманітарних та
соціально-економічних дисциплін
Протокол від 29.08.2023 р. №7

Розглянуто на засіданні кафедри соціології та психології
Протокол від 15.08.2023 р. №8

Розробник:

Професор кафедри соціології та психології факультету №6
д-р соціол. н., професор Нечитайло Ірина Сергіївна

Рецензенти:

1. Керівник психологічної служби Харківського гуманітарного університету «Народна українська академія», доцент кафедри соціології та гуманітарних дисциплін, к. психол. н., Гога Н. П.;
2. Доцент кафедри соціології та психології факультету №6, к. психол. н., доцент Філоненко В. М.

ТЕМА №6. АНАЛІЗ ПСИХОЛОГІЧНИХ ДАНИХ В SPSS: ПАРАМЕТРИЧНІ КРИТЕРІЇ

План

- 6.1.** Теоретичні аспекти і значення параметричних критеріїв.
- 6.2.** Практичні аспекти застосування параметричних тестів.

Рекомендована література

Основна

1. Максименко С. Д., Носенко Е. Л. Експериментальна психологія : підручник Київ : Центр учб. літ., 2017. 359 с.
2. Нечитайло І. С., Бірюкова М. В. Математичні методи в соціології : підручник для студентів ВНЗ / Нар. укр. акад., [каф. соціології]. Харків : Вид-во НУА, 2012. 243 с.

Допоміжна

3. Застосування математично-статистичних методів аналізу у психологічних вимірюваннях [Електронний ресурс]. Режим доступу: <http://surl.li/aghly> . Дата звернення: 31.07.2023.
4. Катаєв Є. С. Використання статистичних методів обробки даних у дослідженнях “я-концепції” особистості. Вісник Національного університету оборони України. 2012. №2 (27) /2012. С. 171-176.
5. Кислова О. М. Методи аналізу та комп’ютерної обробки соціологічної інформації. Харків : Вид-во ХНУ імені В.Н. Каразіна, 2020. 165 с.
6. Салюк М. А. Статистична обробка даних експериментального дослідження. Методичний посібник з курсу «Експериментальна психологія» / за ред. Е.Л. Носенко. Дніпропетровськ: Інновація, 2010. 26 с.
7. Старушенко Г. А. Статистична обробка даних в системі публічного управління : навч. посіб. Дніпро : ГРАНІ, 2018. 144 с.

ТЕКСТ ЛЕКЦІЇ

6.1. Теоретичні аспекти і значення параметричних критеріїв

Критерії й тести, що застосовують для порівняння вибірок, поділяють на дві групи: параметричні й непараметричні.

Особливістю параметричних критеріїв є припущення, що розподіл ознаки в генеральній сукупності підпорядковується певному відомому закону. Ця відповідність має бути доведена до застосування будь-якого з параметричних тестів.

Переважає більшість параметричних тестів розроблена для нормально розподілених даних. Але для деяких типів гіпотез існують параметричні тести, призначені для вибірок, що підпорядковуються іншим законам розподілу.

Як правило, параметричні критерії є потужнішими за непараметричні.

Застосування непараметричних критеріїв у випадках, коли можна використовувати параметричні, призводить до збільшення ймовірності прийняття помилкової нульової гіпотези (тобто помилки другого роду).

Z-критерій є випадковою величиною, що підпорядковується стандартному нормальному розподілу. *Z-критерій* можна застосовувати також для порівняння середніх значень довільно розподілених незалежних вибірок великого обсягу ($n \geq 30$), враховуючи, що в цьому разі вибіркві середні мають приблизно нормальний розподіл, а вибіркві дисперсії є достатньо точними оцінками генеральних дисперсій.

Для порівняння середніх значень вибірок застосовують *t-критерій Стюдента*. Його запропоновано американським статистиком Уїльямом Госсетом в 1908 р. за результатами дослідження проблеми скорочення кількості проб, які потрібно взяти при контролі за якістю продукції пивоварного заводу за умови забезпечення виконання вимог стандартів. Розглядають дві незалежні нормальні вибірки з генеральних сукупностей, що мають рівні або нерівні, але відомі чи рівні невідомі дисперсії.

При аналізі спряжених вибірок їх порівняння здійснюють з метою визначення наявності ефекту від певного фактору, наприклад, впливу змін у технології на якість виробленої продукції. Вимога щодо рівності дисперсій при цьому не висувається. Нульова гіпотеза полягає у відсутності різниці між середніми.

Іноді цей критерій називають одновибірковим критерієм Стюдента. Відповідна статистика має розподіл Стюдента з кількістю степенів вільності $n-1$. Якщо дисперсії або їх відношення невідомі й припущення про рівність дисперсій є необґрунтованим, то виникає так звана проблема Беренса – Фішера, що полягає у перевірці нульової гіпотези про рівність вибірквих середніх за таких умов.

Одним з підходів до її вирішення є застосування *критерію Уелча (Крамера–Уелча)*, запропонованого Б. Уелчем в 1947 р. Статистика цього критерію є приблизно такою самою, як для розподілу Стюдента з кількістю степенів вільності. Порівняння вказує, що основною відмінністю критерію Уелча з погляду прикладного аналізу є зміна кількості степенів вільності.

F-критерій Фішера запропоновано британським біологом і статистиком Рональдом Фішером в 1920 р. Його використовують для порівняння дисперсій двох вибірок. Гіпотезу про рівність дисперсій порівнюваних сукупностей відхиляють, якщо обчислене значення перевищує табличне при заданому довірчому рівні. При цьому, якщо конкуруючою є одnobічна гіпотеза $\sigma_1^2 / \sigma_2^2 > s$, то як критичну точку беруть значення оберненого розподілу Фішера, що відповідає рівню значущості α при заданій кількості степенів вільності. Якщо ж конкуруючою є двобічна гіпотеза $\sigma_1^2 / \sigma_2^2 \neq s$, то критичною точкою буде значення оберненого розподілу Фішера, що відповідає рівню значущості $\alpha / 2$.

Для оцінювання значущості отриманого значення F використовують *критерій В. І. Романовського* (запропонований радянським математиком Всеволодом Івановичем Романовським в 1928 р.). При $F \geq F_{\alpha}$ значення F вважають значущим, а порівнювані вибірки – істотно різними.

Параметричні тести можна застосовувати також при множинних порівняннях, тобто при порівнянні двох груп вибірок одна з одною. Кожну групу задають подібно до того, як задають параметри масивів даних у методах двофакторного дисперсійного аналізу. При множинних порівняннях використовують багатовимірні узагальнення тестів, що були розглянуті вище.

6.2. Практичні аспекти застосування параметричних тестів

Порівняння двох середніх за допомогою t-критерію

Різні варіанти обробки даних із застосуванням t-критерію дозволяють зробити висновок про відмінність двох середніх значень. Наприклад, у разі t-критерію для незалежних вибірок перевіряється достовірність відмінності двох вибірок по змінній, вимірюваній у представників цих двох вибірок. Для цих вибірок обчислюються середні значення, але потім t-критерію визначається статистична значимість їх відмінності. Застосування t-критерію, мабуть, найпоширеніший метод статистичного висновку, оскільки він дозволяє відповісти на просте питання про відмінність двох вибірок за рівнем вираженості вимірюваної ознаки. Основна вимога до даних для застосування цього критерію – представлення змінних, за якими порівнюються вибірки, у метричній шкалі.

SPSS дозволяє застосовувати 3 варіанти t-критерію: Independent-Samples T Test (t-критерій для незалежних вибірок), Paired-Samples T Test (t-критерій для залежних вибірок), One-Samples T Test (t-критерій для однієї вибірки).

Перший з варіантів t-критерію, t-критерій для незалежних вибірок (independent samples), призначений для порівняння середніх значень двох вибірок. Для порівнюваних вибірок повинні бути визначені значення однієї і тієї ж змінної. За допомогою t-критерію для незалежних вибірок можна порівняти успішність студентів і студенток, ступінь задоволеності життям неодружених і одружених, середній зріст футболістів двох команд тощо.

Обов'язковою умовою для проведення t-критерію є незалежність вибірок.

Другий з t-критеріїв, t-критерій для залежних вибірок (paired samples, або dependent samples), дозволяє порівняти середні значення двох вимірювань однієї ознаки для однієї й тієї ж вибірки, наприклад, результати першого і останнього іспитів групи студентів або значення показника до і після впливу на групу. Обов'язковою умовою застосування t-критерію для залежних вибірок є наявність повторного вимірювання для однієї вибірки.

Останній з t-критеріїв, t-критерій для однієї вибірки (one sample), дає змогу порівняти середнє значення цієї вибірки з певною еталонною величиною. Наприклад, чи відрізняється середнє значення деякого тесту у даній вибірці від нормативної величини (наприклад, чи відрізняється час, показаний бігунами під час змагання, від 17 хвилин тощо).

Рівень значущості

Результат порівняння середніх значень із застосуванням t-критерію оцінюється за рівнем значущості. Нагадаємо, що рівень значущості є мірою

статистичної достовірності результату обчислень, в даному випадку – відмінностей середніх, і служить підставою для інтерпретації. Якщо дослідження показало, що p -рівень значущості відмінностей не перевищує 0,05, це означає, що з імовірністю не більше 5% різниці є випадковими. Зазвичай це є підставою для висновку щодо статистичної достовірності відмінностей. В іншому випадку ($p > 0,05$) відмінність визнається статистично недостовірною і не підлягає змістовній інтерпретації.

SPSS дозволяє визначати два рівня значущості: односторонній (one-tailed) і двосторонній (two-tailed). Зазвичай використовується двосторонній рівень значущості. Але якщо ви заздалегідь знаєте напрямків відмінностей і вас цікавить тільки цей напрям, то можна використовувати односторонній критерій значущості. Однак така ситуація зустрічається рідко, а якщо і зустрічається, то правомірність односторонньої перевірки насилу піддається обґрунтуванню.

Істотною відмінністю між двома типами рівнів значущості є те, що двосторонній p -рівень виявляється вдвічі більшим, ніж односторонній.

За замовчуванням SPSS обчислює двосторонній рівень значущості, але легко можна отримати односторонній рівень значущості, розділивши обчислене програмою значення на 2.

Покрокові алгоритми обчислень

Для застосування t -критерію у SPSS спочатку необхідно виконати три підготовчих кроки, наведені нижче.

Крок 1. Створіть новий файл даних або підготуйте існуючий.

Крок 2. Запустіть програму SPSS за допомогою значка на робочому столі або команди Пуск * Програми t SPSS for Windows > SPSS 11.5 for Windows (Start > is\$| Programs > SPSS for Windows > SPSS 11.5 for Windows) головного меню Windows.

У відкритому після запуску програми діалоговому вікні SPSS for Windows клацніть на кнопці Cancel (Скасувати).

Після виконання цього кроку на екрані з'явиться вікно редактора даних SPSS.

Крок 3. Відкрийте файл даних, з якими ви маєте намір працювати.

Якщо він розташований в поточній папці, то виконаєте наступні дії:

- 1). Виберіть у меню File (Файл) команду Open > Data (Відкриття > Дані) або клацніть на кнопці OpenFile (відкрити файл) панелі інструментів;
- 2). У діалоговому вікні двічі клацніть потрібний файл або введіть з клавіатури і клацніть на кнопці OK.

Незалежно від того, відкрита програма SPSS тільки що або якісь процедури вже виконувалися, у верхній частині головного вікна має бути присутній рядок меню. Поки рядок меню присутній на екрані, доступні всі команди аналізу даних. При цьому вікно з даними бачити не обов'язково.

При роботі з таблицями результатів чи під час редагування діаграм деякі пункти меню можуть зникати або змінюватися. Щоб повернутися до головного вікна і стандартної панелі меню, клацніть на кнопці згортання або відновлення поточного вікна.

Застосування t-критерію для незалежних вибірок

Після завершення кроку 3 на екрані має з'явитися вікно редактора даних з рядком меню.

Крок 4. В меню Analyze (Аналіз) виберіть команду Compare Means > Independent-Samples T Test (Порівняння середніх > t-критерій для незалежних вибірок). На екрані з'явиться діалогове вікно Independent-Samples T Test (t-критерій для незалежних вибірок).

У представленому діалоговому вікні ви можете вказати параметри для застосування t-критерію. Ліва частина вікна містить список всіх доступних змінних файлу, а права частина – список Test Variable(s) (Досліджувані змінні), в який вам необхідно помістити імена змінних для застосування t-критерію.

Список може містити будь-яку кількість змінних, а самі змінні повинні бути метричного типу (це змінні «позначка-1», «позначка-2» тощо).

У полі Grouping Variable (Групуюча змінна) указується ім'я змінної, градаціям якої відповідають дві незалежні вибірки для t-критерію. Як правило, групуюча змінна є дискретною і має щонайменше 2 градації. Прикладом дворівневої змінної є змінна статі, яка має дві градації: 1) чоловіча; 2) жіноча. Групуюча змінна, насправді, може мати і більше двох градацій, але в обох випадках необхідно розбити множину значень змінної на дві категорії або задати дві градації з існуючих. Так, ми можемо взяти змінну «клас» як групуючу змінну і вказати, що в одну вибірку слід включати всі об'єкти, що мають значення 1, а в іншу вибірку – всі об'єкти, які мають значення 3. Це дозволить порівняти середні значення для двох з трьох класів.

Отже, після того як ім'я групуючої змінної задано, необхідно задати її категорії. Це робиться клацанням на кнопці Define Groups (Визначення груп).

Зверніть увагу на те, що даний крок є необхідним навіть для дворівневих змінних. На екрані з'явиться діалогове вікно Define Groups (Визначення груп).

Якщо встановлено перемикач Use specified values (Використовувати задані значення), то в полях Group1 (Група 1) і Group2 (Група 2) потрібно перерахувати значення групуючої змінної, які формують першу і другу вибірки. Якщо встановлений перемикач Cut Point (Точка поділу), то в полі праворуч вказується значення, розділяє весь діапазон змінної на два непересічних інтервали.

У наступному прикладі у якості групуючої ми використовуємо змінну «Заклад вищої освіти». Вона має 4 градації, тому нам потрібно розбити ці градації на 2 групи.

Для цього ми встановимо в якості точки розділу 3. Це буде означати, що в першу вибірку потраплять учні, які вибирають гуманітарні і економічні спеціальності, а в другу – учні, які вибирають технічні та природничі спеціальності.

Крок 5. На цьому кроці ми застосуємо t-критерій для порівняння юнаків та дівчат (градацій змінної «стать») зі змінною 2.

1. Клацніть спочатку на змінну 2, щоб виділити її, а потім – на кнопці вгорі зі стрілками, щоб перемістити змінну в список Test Variable(s) (Досліджувані змінні).

2. Клацніть спочатку на змінній статі, щоб виділити її, а потім – на нижній кнопці зі стрілкою, щоб перемістити змінну в поле Grouping Variable (Групуюча змінна).

3. Клацніть на кнопці Define Groups (Визначення груп), щоб відкрити однойменне діалогове вікно.

4. Введіть число 1 в полі Group1 (Група 1), натисніть клавішу Табуляції, введіть число 2 у полі Group2 (Група 2) і клацніть па кнопці Continue (Продовжити), щоб повернутися в діалогове вікно Independent-Samples T Test (t-критерій для незалежних вибірок).

5. Клацніть на кнопці ОК, щоб відкрити вікно виведення.

Крок 5а. Після виконання кроку 4 має відкритися діалогове вікно Independently Samples T Test (t-критерій для незалежних вибірок).

Якщо ви вже встигли попрацювати з цим вікном, очистіть його клацанням на кнопці Reset (Скидання) і виконайте такі дії:

1. Клацніть спочатку па змінній «поначка2», щоб виділити її, а потім – на кнопці вгорі зі стрілками, щоб перемістити змінну в список Test Variable(s) (Досліджувані змінні).

2. Клацніть спочатку на змінній «Заклад вищої освіти», щоб виділити її, а потім – на нижній кнопці зі стрілкою, щоб перемістити змінну в поле I Grouping Variable (Групуюча змінна).

3. Клацніть на кнопці Define Groups (Визначення груп), щоб відкрити однойменне діалогове вікно.

4. Встановіть перемикач Cut point (Точка поділу), введіть в розташоване у полі поруч значення 3 і клацніть на кнопці Continue (Продовжити), щоб повернутися в діалогове вікно Independent-Samples T Test (t-критерій для незалежних вибірок).

5. Клацніть на кнопці ОК, щоб відкрити вікно виведення результатів.

Застосування t-критерію для залежних вибірок

Тепер розглянемо два приклади застосування t-критерію для залежних вибірок. Після завершення кроку 3 на екрані має з'явитися вікно редактора даних з рядком меню.

Крок 5б. У меню Analyze (Аналіз) виберіть команду Compare Means > Paired-Samples T Test (Порівняння середніх > t-критерій для залежних вибірок). На екрані з'явиться діалогове вікно Paired-Samples T Test (t-критерій для залежних вибірок).

Застосування t-критерію для залежних вибірок дещо простіше, ніж для незалежних. Процедура керується єдиним вікном і не вимагає вказування групуючої змінної. У лівій частині вікна знаходиться вже знайомий список доступних змінних файлу, в якому ви можете вибрати змінні для застосування t-критерію. Незвичайним тут є те, що потрібно вказувати пари змінних – звідси назва цільового списку: Paired Variables (Пари змінних).

Спочатку потрібно клацнути на імені першої змінної, потім – на імені другої змінної, і після цього за допомогою кнопки зі стрілкою пара заноситься в список. Кількість обраних пар може бути будь-якою.

Крок 5в. Наприклад, щоб порівняти оцінки учнів 10 і 11 класів («Позначка 1» і «Позначка 2»), необхідно виконати наступні дії:

1. Клацніть на змінній «Позначка 1». Її ім'я з'явиться в області Current Selections (Поточне виділення) поряд з міткою Variable 1 (Змінна 1).
2. Клацніть на змінній «Позначка 2». Її ім'я з'явиться в області Current Selections (Поточне виділення) поряд з міткою Variable 2 (Змінна 2).
3. Клацніть на кнопці зі стрілкою, щоб перемістити вибрану пару змінних в список Paired Variables (Пари змінних).
4. Клацніть на кнопці ОК, щоб відкрити вікно виводу.

Щоб застосувати t-критерій одночасно для декількох пар змінних, їх все потрібно перемістити в список Paired Variables (Пари змінних). У наступному прикладі проводиться порівняння результатів першого тесту (тест1) з результатами інших чотирьох тестів.

Крок 5г. Після виконання кроку 4 має відкритися діалогове вікно Paired-Samples t-Test (t-критерій для залежних вибірок).

Якщо ви вже встигли попрацювати з цим вікном, очистіть його клацанням на кнопці Reset (Скидання) і виконайте наступні дії:

1. Клацніть на змінній «Тест 1». Її ім'я з'явиться в області Current Selections (Поточне виділення) поряд з міткою Variable 1 (Варіант 1).
2. Клацніть на змінній «Тест 2». Її ім'я з'явиться в області Current Selections (Поточне виділення) поряд з міткою Variable 2 (Варіант 2).
3. Клацніть на кнопці зі стрілкою, щоб перемістити вибрану пару змінних в список Paired Variables (Пари змінних).
4. Повторіть три попередні дії, тільки замість змінної «Тест 2» у другій дії послідовно потрібно використовувати змінні «Тест 3», «Тест 4» і «Тест 5».
5. Клацніть на кнопці ОК, щоб відкрити вікно виведення.

Застосування t-критерію для однієї вибірки

Іноді буває необхідно порівняти середнє значення розподілу з якоюсь фіксованою величиною. Уявімо собі таку ситуацію. Дослідник вирішив перевірити, чи відрізняються результати тестування його вибірки від нормативних показників. Припустимо, нормативний показник тесту дорівнює 10. Для того щоб перевірити результат вибірки на відповідність нормі, потрібно обчислити середнє значення для вибірки і порівняти його з числом 10. Подібні порівняння в програмі SPSS реалізуються за допомогою t-критерію для однієї вибірки.

Крок 5д. В меню Analyze (Аналіз) виберіть команду Compare Means > One-Sample t-Test (Порівняння середніх > t-критерій для однієї вибірки). На екрані з'явиться діалогове вікно One-Sample T Test (t-критерій для однієї вибірки).

Застосування t-критерію для однієї вибірки є найбільш простим з точки

зору як обчислень, так і управління. Все, що вам потрібно зробити, – це вибрати одну або кілька змінних для аналізу і, помістивши її (їх) в список Test Variable(s) (Досліджувані змінні), визначити константу для порівняння у полі Test Value (Тестове значення), після чого клацнути на кнопці ОК. Зверніть увагу, що у разі завдання у списку Test Variable(s) (Досліджувані змінні) декількох змінних, їх середні значення будуть порівнюватися з однією і тією ж величиною, зазначеною в полі Test Value (Тестове значення). Якщо потрібно провести порівняння з різними константами, необхідно застосувати t-критерій для кожної з констант.

Крок 5е. На цьому етапі порівнюються середні значення змінних «Тест 2» і «Тест 3» з числом 10.

1. Клацніть спочатку на змінну «Тест 2», щоб виділити її, а потім – на кнопці зі стрілкою, щоб перемістити змінну в список Test Variable(s) (Досліджувані змінні).

2. Повторіть попередній крок для змінної «Тест 3».

3. Натисніть клавішу Tab, щоб перевести курсор в полі Test Value 1 (Тестове значення) і введіть число 10.

4. Клацніть на кнопці ОК, щоб відкрити вікно виведення. Після виконання кроку 5, 5а, 5б, 5в або 5г програма автоматично активізує вікно виведення. Для перегляду результатів, при необхідності, можна скористатися вертикальною і горизонтальною смугами прокрутки. Необхідно звернути увагу на стандартний рядок меню у верхній частині вікна виведення: його присутність дозволяє виконувати будь-які статистичні операції, не перемикаючись назад у вікно редактора даних.