**Національний університет цивільного захисту України**

**Кафедра психології діяльності в особливих умовах**

# МАТЕМАТИЧНІ МЕТОДИ В ПСИХОЛОГІЇ

## *Курс лекцій*

## *Мультимедійне навчальне видання*

**Харків 2020**

**Національний університет цивільного захисту України**

**Кафедра психології діяльності в особливих умовах**

# МАТЕМАТИЧНІ МЕТОДИ В ПСИХОЛОГІЇ

## *Курс лекцій*

## *Мультимедійне навчальне видання*

**Харків 2020**

Рекомендовано до друку вченою радою соціально-психологічного факультету НУЦЗ України
(протокол від 21.05.2020 року № 9)

**Рецензенти:** доктор психологічних наук, доцент **В. О. Олефір**, завідувач кафедри загальної психології Харківського національного університету ім. В.Н. Каразіна;
доктор біологічних наук, професор **Л. А. Перелигіна**, начальник кафедри психології діяльності в особливих умовах НУЦЗ України.

**Боснюк В. Ф.**
**Математичні методи в психології**: курс лекцій. Мультимедійне навчальне видання – Х.: НУЦЗУ, 2020. – 141 с.

У курсі лекцій розглянуто базові параметри описової статистики, розкриті основи застосування методів аналізу даних, алгоритми їх вибору залежно від дослідницької ситуації. Під час аналізу статистичних критеріїв основна увага приділяється межам їх застосування, можливим альтернативам, особливостям інтерпретації результатів.

Лекції забезпечені різноманітними прикладами та питаннями для підготовки, які дозволяють закріпити та покращити теоретичні знання і отримати навики практичного використання статистичних методів. В кінці тексту наведено таблиці критичних значень для різноманітних статистичних критеріїв.

Відмінна особливість навчального видання – використання технології QR-кодування, яка дозволяє суттєво збільшити обсяг наданої інформації, використовувати інтернет-ресурси.

Курс лекцій призначено для здобувачів вищої освіти. Може стати корисним для викладачів психологічних факультетів та для всіх, хто цікавиться методами аналізу даних.

**ЗМІСТ**

# ВСТУП

Знання отримані під час вивчення навчальної дисципліни «Математичні методи в психології» сприяє розвитку професійного мислення в здобувачів вищої освіти. Застосовують для опрацювання результатів досліджень, щоб встановити кількісні залежності між психічними явищами. Вони допомагають критично оцінювати результати наукових публікацій, підвищують надійність висновків, надають підстави для теоретичних узагальнень.

Даний курс формує статистично-математичну грамотність у психологів, розвиває вміння висловлювати обґрунтовані судження. Він передбачає теоретичне і практичне оволодіння студентами та курсантами статистичними процедурами аналізу емпіричних даних для вирішення завдань, які виникають спочатку в межах виконання ними курсових і дипломних робіт, а потім і в процесі проведення власних наукових та науково-практичних психологічних досліджень.

Основні завдання вивчення навчальної дисципліни:

- сформувати у здобувачів вищої освіти позитивну мотивацію до використання сучасних математичних методів в фундаментальних та прикладних психологічних дослідженнях;
- освоїти теоретичні положення та алгоритми розрахунків основних статистичних методів аналізу даних;
- закріпити на практиці знання про організацію та планування психологічного дослідження;
- сформувати навички використання математичних методів для вирішення професійних завдань.

В курсі лекцій використана технологія QR-кодування. В тексті після навчальних прикладів опису алгоритмів застосування статистичних критеріїв аналізу даних розміщені QR-коди (двовимірні штрих-коди швидкого реагування) шифрування інформації, в яких закодовані авторські відео-огляди процедури аналізу даних та інтерпретації результатів у статистичній програмі SPSS Statistics 23.0. Також за допомогою них забезпечено доступ до рекомендованої літератури.

Для того щоб декодувати інформацію, необхідно на декілька секунд піднести до його зображення камеру мобільного телефону (смартфона, планшета), у якого установлена програма-сканер QR-кодів та є доступ до мережі Інтернет. Коли зчитування коду закінчиться на екрані з’явиться інформаційний контент.

## ЛЕКЦІЯ 1
## ОСОБЛИВОСТІ ВИМІРЮВАННЯ ОЗНАК ТА КІЛЬКІСНИЙ ОПИС ДАНИХ В ПСИХОЛОГІЇ

### План лекції

1.1. Генеральна сукупність та вибірка дослідження
1.2. Проблема вимірювання у психології
1.3. Типи та характеристики вимірювальних шкал
1.4. Первинна описова статистика
1.5. Представлення результатів дослідження за допомогою параметрів розподілу

**Час проведення:** 2 учбові години.

### Література

1. Гланц С. Медико-биологическая статистика. Пер. с англ. – М., Практика, 1998. – 459 с.

2. Климчук В.О. Математичні методи у психології. Навчальний посібник. – К.: Освіта України, 2009. – 288 с.

3. Кричевец А. Н., Корнеев А. А., Рассказова Е. И. Основы статистики для психологов. – М.: Акрополь, 2019. – 286 с.

4. Савченко Т.Н., Головина Г.М. Роль математической психологии в гуманитарном знании // Психология. Журнал Высшей школы экономики, 2014. – Т. 11. – № 3. – С. 8-22

5. Телейко А.Б. Чорней Р.К. Математико-статистичні методи в соціології та психології: навч. посібник. – Київ: МАУП, 2007. – 418 с.

### 1.1 Генеральна сукупність та вибірка дослідження

Дослідник-теоретик, розмірковуючи про ті чи інші закономірності психіки та поведінки, як правило, міркує не про конкретний об'єкт свого дослідження, а має на увазі швидше якусь множину об'єктів. Наприклад, досліджуючи пам'ять людини, ми маємо на увазі не когось конкретно, а людину взагалі, усіх людей, які живуть на Землі, жили або ще будуть жити. Очевидно, у даному прикладі мова йде про якусь досить велику множину об'єктів, яка до того ж немає чітко визначених меж. Таку множину об'єктів у математичній статистиці прийнято називати *генеральною сукупністю* або *популяцією (N).* Зазвичай, ми можемо конкретизувати наші уявлення про досліджувану генеральну сукупність, зробити її більш компактною. Наприклад, ми можемо говорити не про пам'ять людини взагалі, а про пам'ять якоїсь більш вузької групи людей – про пам'ять дитини певного віку або про пам'ять людей, які страждають амнезією. Але навіть у цьому випадку, генеральна сукупність виявиться досить розмитою сукупністю, що включає велику кількість об'єктів, над поведінкою яких власне й розмірковує теоретик, намагаючись зрозуміти закони поведінки цих теоретичних об'єктів, передбачити цю поведінку.

Проводячи експериментальне дослідження, дослідник-експериментатор прагне перевірити передбачення дослідника-теоретика. Водночас він не може провести дослідження відразу з усіма об'єктами, які складають всю генеральну сукупність. Адже генеральна сукупність дуже велика, до того ж її межі, як правило, не визначені. Так, ідея об'їхати весь світ, досліджуючи пам'ять усіх людей, які страждають тими чи іншими формами амнезії, експериментатору навряд чи буде цікавою й захоплюючою. Навіть якщо йому і вдасться це зробити, через якийсь час, на жаль, з'являться нові люди, які страждають тією ж недугою, і за такого підходу доведеться знову і знову перевіряти припущення теоретика.

Що ж говорити про ситуацію, коли теоретик міркує про людину взагалі! На щастя, у цьому немає ніякої необхідності. Як прийнято говорити в такому випадку, щоб дізнатися смак супу не обов'язково з'їдати цілий котел. Досить однієї ложки. Але попередньо суп необхідно добре розмішати. Саме тому дослідник-експериментатор має справу не з усією генеральною сукупністю, а лише з невеликою її частиною, так званою вибірковою сукупністю або вибіркою.

*Вибірка (n)*– це обмежена за чисельністю група об'єктів (в психології – досліджуваних, респондентів), спеціально відібрана з генеральної сукупності для вивчення її властивостей.

Ця частина генеральної сукупності повинна в максимальному ступені бути подібною до самої генеральної сукупності. Дана характеристика відображається в понятті репрезентативності.

*Репрезентативність вибірки* – це здатність вибірки представляти досліджувані явища досить повно з погляду їхньої мінливості в генеральній сукупності. Це основний критерієм для визначення меж генералізації висновків дослідження.

Наведемо два методи, що забезпечують репрезентативність вибірки.

Перший метод *формування простої випадкової вибірки (рандомізація)*. У цьому випадку вибірка складається з елементів, відібраних із генеральної сукупності у такий спосіб, щоби кожен елемент цієї сукупності мав би рівні можливості (рівну ймовірність) потрапити до вибірки. Отримана таким чином вибірка називається простою випадковою вибіркою .

Другий *метод формування стратометричної випадкової вибірки*. Для цього необхідно розбити елементи генеральної сукупності на страти (групи) за деякими характеристиками. Стратометричний відбір використовується в тому випадку, якщо у вибірці обов'язково повинні бути представлені досліджувані з певним набором характеристик (стать, вік, рівень освіти тощо).

Крім того *вибірки можуть бути залежними або незалежними*.

Якщо можна встановити гомоморфну пару (тобто, коли одному випадку з вибірки X відповідає один і тільки один випадок із вибірки Y і навпаки) для кожного випадку у двох вибірках, такі вибірки називаються залежними (зв'язаними, парними). Приклади залежних вибірок: пари близнюків, чоловіки і їх дружини, група досліджуваних до проведення психологічного експерименту й після тощо.

Вибірки називаються незалежними (незв'язними), якщо процедура дослідження й отримані результати вимірювання деякої властивості в респондентів однієї вибірки не впливають на особливості протікання цього ж дослідження й результати вимірювання цієї ж властивості в респондентів іншої вибірки.

Відповідно, залежні вибірки завжди мають однаковий обсяг досліджуваних, а обсяг незалежних може відрізнятися.

## 1.2 Проблема вимірювання в психології

У своїй роботі психолог постійно стикається з проблемою вимірювання індивідуально-психологічних особливостей. Для цього в психодіагностиці розробляються спеціальні вимірювальні процедури. Сенс яких полягає в тому, що у процесі дослідження психологічні параметри можуть отримувати кількісне вираження. Кількісні дані, отримані внаслідок ретельно спланованого дослідження за допомогою спеціальних вимірювальних процедур, у подальшому використовуються для математико-статистичної обробки.

Вимірювання можна визначити як приписування чисел за певними правилами об'єктам або подіям. Ці правила повинні встановлювати відповідність між властивостями об'єктів, що розглядаються, з одного боку, і ряду чисел – з іншого. Загалом можна сказати, що вимірювання – це процедура, під час якої об'єкт дослідження, порівнюється з деяким еталоном і отримує чисельне вираження в певному масштабі або шкалі.

У кожному конкретному випадку вимірювання є операцією, за допомогою якої емпіричним даним надається форма зв'язного числового повідомлення. Важливо розуміти, що приписування чисел об'єктам за певними принципами і правилами визначає тип шкали. Саме закодована в числовій формі інформація дозволяє використати математичні методи і виявляти приховані закономірності; крім того, числове представлення об'єктів або подій дозволяє оперувати складними поняттями в більш скороченій формі. Саме це і є причиною використання вимірювань у будь-якій науці, у тому числі і психології.

Будь-який різновид вимірювання передбачає наявність одиниць вимірювання. Одиниця вимірювання – це та «вимірювальна паличка», як говорив С. Стівенс, яка є умовним еталоном для здійснення тих або інших вимірювальних процедур. У природничих науках і техніці є стандартні одиниці вимірювання, наприклад, градус, метр, ампер тощо.

Психологічні змінні за одиничними винятками не мають власних вимірювальних одиниць. Тому здебільшого значення психологічної ознаки визначається за допомогою спеціальних вимірювальних шкал.

Розрізняють декілька типів шкал. Операції, а саме способи вимірювання об'єктів задають тип шкали. Шкала, у свою чергу, характеризується видом перетворень, які можуть бути віднесені до результатів вимірювання. Якщо не дотримуватися цього правила, то структура шкали порушиться, а дані вимірювання не можна буде осмислено інтерпретувати.

Тип шкали визначає сукупність статистичних методів, які можуть бути використані для обробки даних вимірювання.

Шкала (латин. scala) – інструмент для вимірювання властивостей об'єкта, що являє собою числову систему, у якій відносини між різними властивостями об'єктів виражені властивостями числового ряду.

У психології різні шкали використовуються для вивчення різних характеристик соціально-психологічних явищ. Виділяється чотири типи числових систем, що визначають відповідно чотири рівні, або шкали вимірювання:

1) шкала номінальна (номінативна, найменувань, класифікаційна);
2) шкала порядку (рангова, ординарна);
3) шкала інтервалів (рівних інтервалів);
4) шкала відношень (пропорційна).

Низка фахівців виділяють також абсолютну шкалу і шкалу різниць.

Виділяють три основних атрибути вимірювальних шкал, наявність або відсутність яких визначає належність шкали до тієї чи іншої категорії:

1. впорядкованість даних означає, що один пункт шкали більший, менший або дорівнює іншому пункту;

2. інтервальність пунктів шкали означає, що інтервал між будь-якою парою чисел більший, менший або дорівнює інтервалу між іншою парою чисел;

3. нульова точка означає, що набір чисел має позначку відліку, що позначається нулем та відповідає повній відсутності властивості.

Крім того, виділяють такі групи шкал:

- неметричні, у яких відсутні одиниці вимірювань (номінальна й порядкова шкали);

- метричні (шкала інтервалів, шкала відносин).

### 1.3 Типи та характеристики вимірювальних шкал

***Шкала номінальна.*** Шкала номінальна (синоніми: шкала найменувань, номінативна шкала) створюється шляхом привласнення «імен» об'єктам.

У результаті аналізу об'єкти класифікуються за групами (класами), які можуть бути позначені номерами, назвами, іменами тощо. Позначення класу не вимірюється кількісно, воно лише дозволяє відрізнити один об'єкт від іншого щодо вимірюваної властивості. Це і складає сутність номінальної шкали.

Це найпростіша вимірювальна шкала, хоча фактично вона не асоціюється з вимірюванням і не пов'язана з поняттям «величина». Вона використовується тільки щоб відрізнити один об'єкт від іншого. Тут відсутні всі головні атрибути вимірювальних шкал, а саме упорядкованість, інтервальні, нульова точка.

Вимірювання полягає в тому, що після вивчення об'єкта визначають належність результату до того чи іншого стану (якості) і записують це за допомогою символу (набором символів, назвою, словом, числом), що позначає даний стан.

Незважаючи на тенденцію «завищувати» потужність шкали, психологи дуже часто застосовують номінальну шкалу у дослідженнях. Часто вимірювальні процедури в діагностиці особистості зводяться до типологізації: віднесення конкретної особистості до того або іншого типу. Прикладом такої типології є класичні темпераменти: холерик, сангвінік, меланхолік і флегматик.

Найпростіша номінальна шкала називається дихотомічною. У разі вимірювання за дихотомічною шкалою ознаки можна кодувати двома символами або цифрами, наприклад 0 і 1, 2 і 6, або буквами А і Б, а також будь-якими двома відмінними один від одного символами. Ознака, що виміряна за дихотомічною шкалою називається альтернативною.

Необхідно розуміти, що позначення класів – це тільки символи, навіть якщо для цього використані номера. З цими номерами не можна здійснювати операції як із числами. Для прикладу, якщо в одного спортсмена на спині номер 1, а в іншого – 2, то ніяких інших висновків, крім того, що це різні учасники змагань робити не можна: наприклад, не можна сказати, що «другий у два рази кращий».

Байдуже, у якому порядку будуть розташовані класифікаційні назви. Те, що номер одного класу більший або менший, нічого не говорить про властивості об'єктів, за винятком того, що вони різні.

З даними зафіксованих у номінальній шкалі можна здійснювати тільки операцію перевірки їх тотожності або відмінності.

Приклади номінальної шкали у психології: стать, національність, освіта, тип темпераменту, тип особистості тощо.

***Шкала порядку.*** Якщо можна встановити порядок розташування психологічних об'єктів відповідно до вираженості в них якоїсь властивості, то використовується порядкова шкала.

У порядковій шкалі присутня упорядкованість, але відсутні атрибути інтервальності та нульової точки.

Дана шкала дозволяє здійснити лінійну впорядкованість об'єктів на деякій осі ознаки (відношення «більше» або «менше»). Тим самим є можливість вивчити «лінійну» вираженість властивості, тоді як шкала найменувань використовує «категоріальну» властивість (властивість є – властивості немає).

Конструювання шкали порядку – процедура більш складна, ніж створення шкали найменувань. Вона дозволяє зафіксувати ранг (місце) кожного значення змінної по відношенню до інших значень.

Одиницею вимірювання є відстань в один клас, ранг або бал. Водночас відстань між рангами (балами) може бути різною (вона нам невідома).

У порядковій шкалі повинно бути не менше трьох градацій (класів). Чим більше число градацій, на які розбита вся експериментальна сукупність, тим ширші можливості статистичної обробки отриманих даних і перевірки статистичних гіпотез.

Порядкові змінні можна кодувати будь-якими цифрами, головне щоб зберігатися порядок, тобто кожна наступна цифра повинна бути більшою (або меншою) попередньої.

Характерною особливістю порядкових шкал є те, що відношення порядку нічого не говорить про дистанцію між класами. Тому порядкові експериментальні дані, навіть якщо вони зображені цифрами, не можна розглядати як числа. З числовими значеннями порядкової шкали не рекомендується виконувати арифметичні операції: додавати, віднімати, множити, ділити, обчислювати середнє значення. Проте якщо виникла така необхідність, то результати цих обчислень слід інтерпретувати з обережністю, тому що в цій шкалі немає «еталонної» одиниці виміру, яка єдина на всіх ділянках шкали.

Шкали порядку широко використовуються в психології. Ще сам засновник теорії вимірювання С. Стівенс висловив точку зору, що результати психологічних вимірювань здебільшого відповідають шкалам порядку.

***Шкала інтервалів.*** Шкала інтервалів є першою метричною шкалою. Власне, починаючи з неї, доцільно говорити про вимірювання у вузькому значенні цього слова.

У даній шкалі присутні атрибути впорядкованості та інтервальності, проте нульова точка довільна, вона не вказує на відсутність вираженості властивості, це просто точка відліку.

За допомогою шкали інтервалів можна порівнювати два об'єкти, визначити величину відмінностей між ними. Ми дізнаємось на скільки більше або менш виражена певна властивість у одного об'єкта, ніж у іншого. Водночас немає сенсу говорити, у скільки разів більше або менше.

У цій шкалі тільки інтервали відображають значення справжніх чисел і тільки над інтервалами слід виконувати арифметичні операції. Якщо забути про це, то є ризик отримати безглузді результати. Наприклад, не можна говорити, що температура води зросла у два рази при нагріванні від 9 до 18 градусів за шкалою Цельсія, оскільки для того, хто звик користуватися шкалою Фаренгейта, це буде дуже дивно,  так як у цій шкалі температура води в тому ж експерименті зміниться від 37 до 42 градусів.

Інтервальна шкала дозволяє застосовувати практично всю параметричну статистику для аналізу даних. Для характеристики центральної тенденції крім медіани й моди можна розраховувати середнє арифметичне, а для оцінки варіації – дисперсію, стандартне відхилення. Можна визначати коефіцієнти асиметрії й ексцесу та інші параметри розподілу.

У психологічних дослідженнях до інтервальних шкал належать так звані квазіінтервальні (штучно створені) шкали. Це шкала IQ, стени, Т-бали, стенайни й будь-які інші стандартизовані оцінки в методиках.

***Шкала відношень.*** У даній шкалі об'єкти класифікуються пропорційно ступеню вираженості діагностованої властивості. Значення в шкалі відношень є «повноправними» числами, з ними можна здійснювати будь-які арифметичні операції і використовувати різноманітні статистичні методи. Тут присутні всі атрибути вимірювальних шкал: упорядкованість, інтервальність, нульова точка.

Особливістю цієї шкали є наявність абсолютного (істинного) нуля, який означає повну відсутність вираженості ознаки. У шкалі відносин можна стверджувати, у скільки разів властивість одного об'єкта перевершує таку ж властивість у іншого.

Приклади шкал відношень: дистанція, вага, зріст, час, температура за Кельвіном; у психології – шкали порогів абсолютної чутливості, час реакції, кількість об'єктів або суб'єктів.

### 1.4 Первинна описова статистика

Первинна статистична обробка спрямована на впорядкування інформації про об'єкти дослідження, дозволяє в стислому вигляді зрозуміти загальний характер усієї сукупності даних: про їхню однорідність/неоднорідність, компактність/мінливість, відповідність певному закону розподілення - не відповідність тощо.

Первинний статистичний аналіз усієї сукупності отриманих у дослідженні даних дає можливість відповісти на два основних питання: 1) яке значення найбільш характерне для вибірки?; 2) чи велика варіативність даних щодо цього характерного значення?

Для відповіді на перше питання обчислюються міри центральної тенденції, для другого – міри мінливості.

***Мірами центральної тенденції*** називають чисельні показники типових властивостей емпіричних даних. Ці показники дають відповіді на питання про те, наприклад, «який середній рівень інтелекту студентів університету?», «яке типове значення показника відповідальності певної групи осіб?». Є порівняно невелика кількість таких показників-мір і насамперед: мода, медіана, середнє арифметичне. Кожна конкретна міра центральної тенденції має свої особливості, що робить її цінною для характеристики об'єкта дослідження в певних умовах.

*Мода* (*Мо*) – це значення, яке найчастіше трапляється у вибірці, тобто число з найбільшою частотою. Для того щоби знайти моду, необхідно проаналізувати варіаційний ряд і знайти в ньому значення, яке повторюється частіше за інших (в згрупованому розподілі вказується *модальний інтервал* значень ознаки). Зверніть увагу, що модою виступає саме значення ознаки, а не її найбільша частота.

Якщо це значення одне, то такий розподіл ознаки називається унімодальним. Розподіл може мати й декілька мод. У разі наявності двох мод розподіл називається бімодальним. Якщо всі значення в групі трапляються з однаковою частотою, то вважається, що мода відсутня.

Вибираючи моду в якості міри центральної тенденції ознаки у вибірці, слід пам'ятати, що вона дійсно буде відображати основну тенденцію тільки в випадку, якщо модальна частота значно перевищує найближчі до неї частоти (у кілька разів).

**Приклад 1.1.**

1. У дослідженні двох різних вибірок отримані такі результати:

Таблиця 1.1 –Вибірка № 1

| Тип темпераменту | Холерик | Сангвінік | Флегматик | Меланхолік | Σ |
|---|---|---|---|---|---|
| $f_i$ | 16 | 17 | 13 | 14 | 60 |
| $p_i$ | 0,27 | 0,28 | 0,22 | 0,23 | 1,00 |

Вибірка № 2

| Тип темпераменту | Холерик | Сангвінік | Флегматик | Меланхолік | Σ |
|---|---|---|---|---|---|
| $f_i$ | 10 | 22 | 5 | 3 | 40 |
| $p_i$ | 0,25 | 0,55 | 0,12 | 0,08 | 1,00 |

У 1-й вибірці не можна говорити про існування певної тенденції представленості будь-якого типу темпераменту у групі досліджуваних: вони трапляються приблизно однаково. Хоча якщо формально виділяти моду, то сангвінічний тип темпераменту є модальним (частота цього типу в даній вибірці найбільша).

У 2-й вибірці модою теж є сангвінічний тип темпераменту (він також має найбільшу частоту у варіаційному ряду), проте дана частота перевищує найближчу (за величиною) майже удвічі – у даному випадку вже можна говорити про виявлення основної тенденції представленості типів темпераменту в даній вибірці.

*Медіана* (*Me* або *Md*) – це значення ознаки, що ділить впорядковану вибірку досліджуваних на дві рівні частини: 50 % мають значення ознаки менше медіани, 50 % мають значення ознаки більше медіани. Іншими словами, медіана припадає на середину упорядкованої послідовності емпіричних даних.

Якщо всі значення у вибірці упорядкувати, то не слід вважати, що медіана завжди розташована на середині розмаху вибірки. Вона може розташовуватися в будь-якому місці емпіричного ряду. Значення медіани залежить від того, як всередині цього упорядкованого ряду розподілилися досліджувані.

Крім того, медіана не обов'язково повинна співпадати з конкретним значенням. Збіг відбувається у випадку непарного числа значень (відповідей), розбіжність – у разі парної кількості чисел. В останньому випадку медіана обчислюється як середнє арифметичне двох центральних значень у впорядкованому ряді даних.

**Приклад 1.2.** Розрахуємо медіану для розподілу емпіричних даних: 90, 66, 106, 84, 105, 83, 104, 82, 107, 97.

Спочатку необхідно впорядкувати дані: 66, 82, 83, 84, 90, 97, 104, 105, 106, 107.

У нашому прикладі парна кількість чисел, центральними значеннями є 90 та 97.

$$Md = \frac{90+97}{2} = 93,5.$$

Для симетричних розподілів медіана співпадає з модою, для асиметричних – не співпадає.

Медіану використовують як міру мінливості для даних виміряних у порядковій шкалі.

*Середнє арифметичне значення, середнє значення* – це значення ознаки, яке відображає середній рівень вираженості параметра у досліджуваних.

Середнє арифметичне значення визначається за такими формулами:

А) для вибіркової оцінки:

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

де $x_i$ – значення ознаки *x* у конкретного досліджуваного;

$n$ – обсяг вибірки досліджуваних.

Б) для генеральної сукупності:

$$\mu = \frac{\sum x}{N}$$

де *x* – значення ознаки у всіх членів генеральної сукупності;

*N* – число членів генеральної сукупності.

Для аналізу результатів дослідження бажано використовувати всі три міри центральної тенденції – моду, медіану й середнє арифметичне значення.

Особливості мір центральної тенденції:

1. У малих за обсягом вибірках мода може бути не стабільна.

**Приклад 1.3.** Для вибірки 1, 1, 1, 3, 5, 7, 7, 8. *Mo* = 1. Припустимо в цій вибірці один досліджуваний був замінений іншим. Значення стали такими: 1, 1, 3, 5, 7, 7, 7, 8. *Mo* = 7.

2. На медіану не впливають найбільші та найменші значення (аутлаєри).

**Приклад 1.4**. Для вибірки 1, 1, 3, 5, 7. *Md* = 3. Для вибірки 1, 1, 3, 5, 20. *Md* = 3.

3. На величину середнього значення впливає кожен елемент вибірки, якщо який-небудь елемент вибірки зміниться на величину c, то середнє значення зміниться на величину c/n.

**Приклад 1.5.** Для вибірки 1, 1, 3, 5, 7. $\bar{x}$ = 3,4. Для вибірки 1, 1, 3, 5, 16. $\bar{x}$ = 5,2. Одне значення у вибірці збільшилось на 9 одиниць. Середнє арифметичне збільшилось на 9/5 = 1,8.

4. Деякі вибірки взагалі не можна охарактеризувати за допомогою мір центральної тенденції. Особливо це актуально для вибірок, що мають декілька мод.

**Приклад 1.6.** Викладач за допомогою тесту досягнень визначив рівень знань студентів за спеціальною бальною шкалою.

За результатами тестування отримано такий варіаційний ряд:

Таблиця 1.2

| Кількість балів | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| $f_i$ | 15 | 10 | 3 | 0 | 0 | 0 | 4 | 12 | 15 |

Результати можна представити у вигляді гістограми:

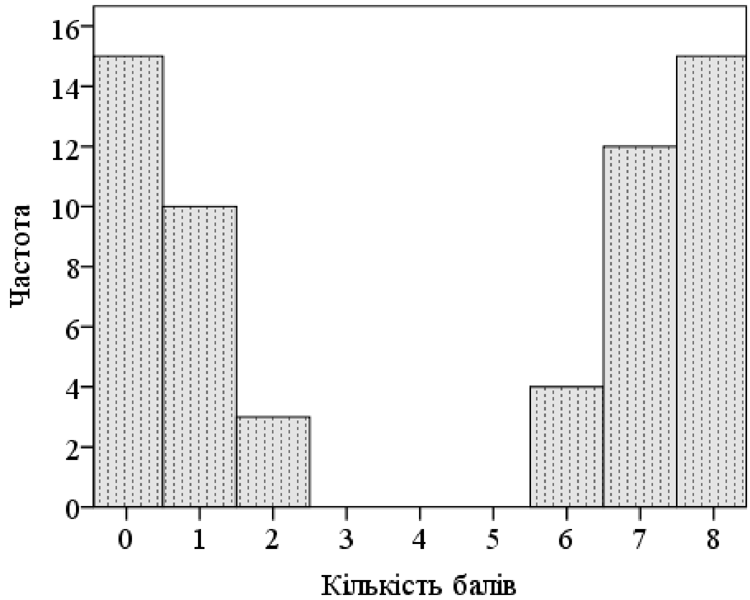


**Рис. 1.1 – Розподіл оцінок**

У даному прикладі середнє арифметичне значення $\overline{x}$ = 3,85, хоча жоден студент не отримав трійки чи четвірки. Медіана цієї вибірки $Md$ = 2,17, хоча є досить велика кількість значень 7 та 8. У даному прикладі ні медіана, ні середнє арифметичне значення не дають адекватного уявлення про досліджувану вибірку. Найбільш коректною характеристикою для даної вибірки є таке твердження: «Гістограма є бімодальною й має V-подібну форму, $Mo_1$ = 0, $Mo_2$ = 8».

5. Якщо вибірка унімодальна та симетрична (відповідає закону нормального розподілу даних), то мода, медіана і середнє значення співпадають.

6. Найпростіше з розглянутих трьох мір центральної тенденції визначається мода, її можна легко обчислити за гістограмою або полігоном частот (рис. 1.2).

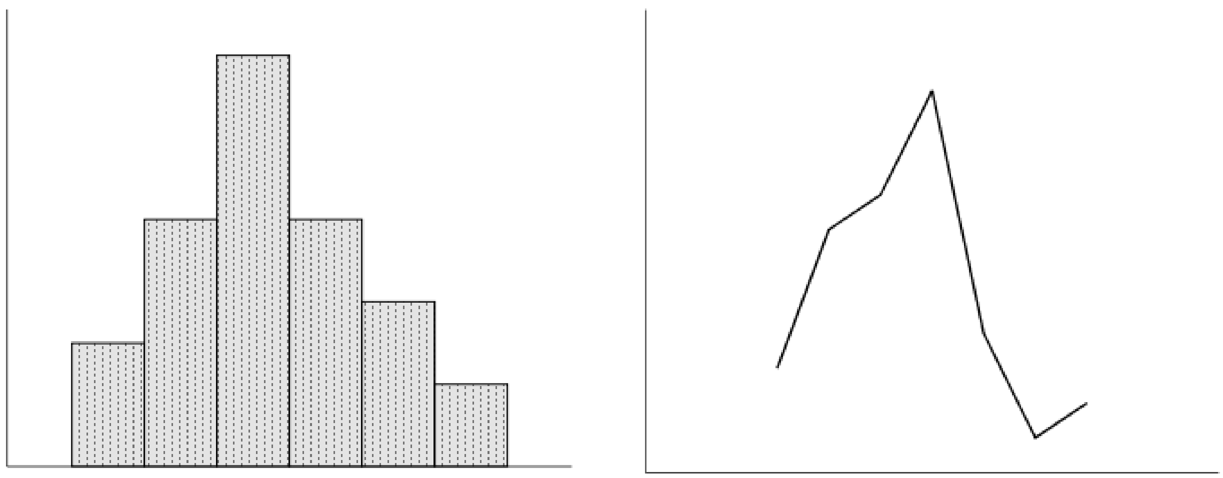

**Рис. 1.2 – Гістограма та полігон частот**

За складністю розрахунку медіана займає проміжне положення між модою й середнім значенням.

Розглянемо приклад, як змінюються міри центральної тенденції, якщо замінити одне значення вибірки.

**Приклад 1.7.** Перша вибірка 1, 3, 3, 5, 6, 7, 8.

$$Mo = 3;\ Md = 5;\ \bar{x} = 4{,}71.$$

Друга вибірка 1, 3, 3, 5, 6, 7, 16

$$Mo = 3;\ Md = 5;\ \bar{x} = 5{,}86.$$

7. У вибірках великого обсягу мода й медіана є більш стійкими характеристиками, ніж середнє значення.

Загалом не можна однозначно сказати, яка із мір центральної тенденції краще характеризує вираженість ознаки у вибірці.

*Середнє геометричне значення (G)* відображає середній приріст значень ознаки, а *середнє гармонійне значення (H)* показує середню швидкість зміни ознаки. Вказані параметри має сенс розраховувати для динамічних ознак, які порівняно швидко змінюються в часі. Наприклад, характеристики психічних станів або ефективність роботи протягом робочого дня. Такого типу ознаки в практиці психологічних досліджень трапляються не часто. Тому формули для розрахунку цих параметрів ми розглядати не будемо (вони достатньо складні). При необхідності слід звернутися до підручника Дж. Гласс та Дж. Стенлі.

***Міри мінливості даних*** – це статистичні показники, що дозволяють судити про ступінь однорідності отриманої множини даних, її компактності. Найбільш поширені в психологічних дослідженнях показники: варіаційний розмах, дисперсія, стандартне відхилення, коефіцієнт варіації, процентилі, квартилі, довірчий інтервал.

*Варіаційний розмах (R)* – це інтервал між максимальним і мінімальним значеннями ознаки. Це найпростіша міра мінливості.

$$R = x_{\max} - x_{\min},$$

де $x_{\max}$ – максимальне значення у вибірці;

$x_{\min}$ – мінімальне значення у вибірці.

Варіаційний розмах показує експериментально виявлену область існування ознаки у вибірці, тобто охоплює 100 % значень ряду. Він визначається легко і швидко, але чутливий до випадкових даних, особливо за малої кількості досліджуваних.

*Дисперсія* – це середній квадрат відхилень всіх значень ознаки від середнього арифметичного. Обчислюються тільки для інтервальних та відносних вимірювальних шкал. Має розмірність значення ознаки у квадраті, через що на практиці використовується не часто. Визначається за такими формулами:

А) для вибіркової оцінки:

$$s^2 = \frac{\sum\left(x_i - \bar{x}\right)^2}{n-1},$$

де $x_i$ – значення ознаки $x$ у конкретного досліджуваного;
$\bar{x}$ – середнє арифметичне значення;
$n$ – обсяг вибірки досліджуваних.

Б) для генеральної сукупності:

$$\sigma^2 = \frac{\sum\left(x_i - \mu\right)^2}{N},$$

де $x_i$ – значення ознаки у всіх членів генеральної сукупності;
$\mu$ – середнє арифметичне генеральної сукупності;
$N$ – число членів генеральної сукупності.

*Стандартне відхилення, середньоквадратичне відхилення* – це середнє відхилення кожного значення ознаки від середнього арифметичного. Має таку ж розмірність, що й сама ознака. Щоби її визначити необхідно з дисперсії розрахувати квадратний корінь.

А) для вибіркової оцінки:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum\left(x_i - \bar{x}\right)^2}{n-1}},$$

де $x_i$ – значення ознаки $x$ у конкретного досліджуваного;
$\bar{x}$ – середнє арифметичне значення;
$n$ – обсяг вибірки досліджуваних.

Б) для генеральної сукупності:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum\left(x_i - \mu\right)^2}{N}},$$

де $x_i$ – значення ознаки у всіх членів генеральної сукупності;
$\mu$ – середнє арифметичне генеральної сукупності;
$N$ – число членів генеральної сукупності.

**Приклад 1.8.** Необхідно розрахувати вибіркову дисперсію та стандартне відхилення тривожності заданого набору даних.

Таблиця 1.3 – Таблиця первинних даних та проміжних розрахунків

| № | *X* | $\left(x_i - \bar{x}\right)$ | $\left(x_i - \bar{x}\right)^2$ |
|---|---|---|---|
| 1 | 7 | 1,7 | 2,89 |
| 2 | 9 | 3,7 | 13,69 |
| 3 | 2 | -3,3 | 10,89 |
| 4 | 4 | -1,3 | 1,69 |
| 5 | 5 | -0,3 | 0,09 |
| 6 | 6 | 0,7 | 0,49 |
| 7 | 7 | 1,7 | 2,89 |
| 8 | 6 | 0,7 | 0,49 |
| 9 | 2 | -3,3 | 10,89 |
| 10 | 5 | -0,3 | 0,09 |
| | $\bar{x}$ = 5,3 | $\Sigma$ = 0 | $\Sigma$ = 44,1 |

$$s^2 = \frac{\sum\left(x_i - \bar{x}\right)^2}{n-1} = \frac{44{,}1}{10-1} = 4{,}9.$$

$$s = \sqrt{\frac{\Sigma\left(x_i - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{44{,}1}{10-1}} = 2{,}21.$$

Таким чином, дисперсія тривожності складає 4,9, а стандартне відхилення – 2,21.

Слід пам’ятати, що стандартне відхилення залежить не тільки від ступеня розсіювання даних, але і від одиниць вимірювання. Тому за допомогою неї можна порівнювати мінливість лише одних і тих же показників, а зіставляти стандартне відхилення різних параметрів за абсолютною величиною не можна. Для того щоби порівняти за рівнем мінливості ознаки будь-якої розмірності (представлені в різних одиницях виміру) та уникнути впливу на величину стандартного відхилення масштабу вимірювань середнього арифметичного значення застосовують коефіцієнт варіації.

*Коефіцієнт варіації* – процентне відношення стандартного відхилення до середньої арифметичної величини ознаки. Застосовується для порівняння мінливості розподілів ознак, що мають різну розмірність. За своєю суттю коефіцієнт варіації призводить до однакового масштабу величину стандартного відхилення.

$$Cv = \frac{s}{\bar{x}} \cdot 100\%,$$

де $s$ – стандартне відхилення;

$\bar{x}$ – середнє арифметичне значення.

Коефіцієнт варіації дозволяє порівнювати мінливість ознак, що представлені в різних шкалах, а також оцінити однорідність вибірки. Орієнтовними критеріями оцінки варіативності за його коефіцієнтом можна вважати: низький рівень – до 10 %; середній рівень – 10-20 %, високий рівень – вище 20 %.

**Приклад 1.9.** Необхідно оцінити варіативність отриманих в результаті дослідження даних: $\bar{x} = 12{,}1$; $s = 3{,}3$.

$$Cv = \frac{s}{\bar{x}} \cdot 100\% = \frac{3{,}3}{12{,}1} \cdot 100\% = 27{,}3\%.$$

Таким чином, коефіцієнт варіації більше 20 %, що свідчить про велику мінливість даних.

У психології також широко використовуються квантилі розподілу. *Квантиль* – це точка на числовій осі виміряної ознаки, яка ділить всю сукупність упорядкованих даних на дві групи з відомим співвідношенням їхньої чисельності. З одним із квантилів ми вже знайомі – це медіана. Це значення ознаки, яке ділить всю сукупність даних на дві групи з рівною кількістю. Крім медіани часто використовуються процентилі та квартилі, хоча є і інші види.

*Процентилі* ($P_1$.......$P_{99}$) – це 99 точок значень ознаки, які ділять упорядковану (за зростанням) множину даних на 100 частин, які рівні за чисельністю. Визначення конкретного значення процентилю аналогічно визначенню медіани. Наприклад, у разі визначення 10-го процентилю ($P_{10}$), спочатку всі значення ознаки упорядковуються за зростанням. Після цього визначається 10% досліджуваних, що мають найменшу вираженість ознаки. $P_{10}$ буде відповідати тому значенню ознаки, яке відокремлює 10% досліджуваних від решти 90%.

*Квартилі* ($Q_1$, $Q_2$, $Q_3$) – це 3 точки значень ознаки, які ділять упорядковану (за зростанням) множину даних на 4 рівні за чисельністю частини. Перший квартиль відповідає 25-му процентилю ($P_{25}$) , другий – 50-му процентилю ($P_{50}$) або медіані ($Md$), третій квартиль відповідає 75-му процентилю ($P_{75}$).

*Децилі* ($D_1$, $D_2$, $D_3$, $D_4$, $D_5$, $D_6$, $D_7$, $D_8$, $D_9$) – це 9 точок значень ознаки, які ділять упорядковану (за зростанням) множину даних на 10 рівних за чисельністю частин.

Процентилі, квартилі й децилі використовуються для визначення частоти тих чи інших значень (або інтервалів) виміряної ознаки або для виділення підгруп і окремих досліджуваних, які найбільш типові чи нетипові для даної множини спостережень.

Інший спосіб оцінки мінливості даних – інтервальне оцінювання. У цьому випадку замість одного числа вказуються нижня і верхня межі інтервалу, тобто група значень на числовій осі.

На практиці для характеристики варіативності порядкових змінних використовують *міжквартильний інтервал* – міра мінливості даних, що охоплює 50 % всіх значень ряду й показує статистичну норму для даної вибірки. Це діапазон між $Q_1$ та $Q_3$.

*Довірчий інтервал (ДІ)* – вид оцінювання, у якому інтервал формується з урахуванням ймовірності попадання значення параметра в межі інтервалу із заданою надійністю. Для середнього значення зазвичай розраховують 95 % довірчий інтервал, який буде містити 95 % середніх значень вибірок, які будуть формуватися із нескінченної генеральної сукупності. Іншими словами можна з 95 % надійністю стверджувати, що істинне значення середнього значення в популяції буде знаходитися в межах 95 % ДІ.

Довірчий інтервал – інтервальна оцінка середнього, яка доповнює точкову оцінку. Чим більший обсяг вибірки, тим вужчий діапазон ознаки.

Якщо розподіл ознак в популяції відповідає закону нормального розподілу, то при розрахунку нижньої та верхньої межі довірчого інтервалу можна застосувати формулу:

$$\bar{x} - t \cdot \frac{s}{\sqrt{n}} \text{ та } \bar{x} + t \cdot \frac{s}{\sqrt{n}},$$

де $\bar{x}$ – середнє арифметичне значення;

$t$ – значення розподілу критерію $t$-Стьюдента;

$s$ – стандартне відхилення;

$n$ – обсяг вибірки досліджуваних.

Значення розподілу критерію $t$-Стьюдента береться із таблиці критичних значення $t$-Стьюдента (додаток 3) з врахуванням ширини діапазону та числа ступенів свободи $df = n - 1$, де $n$ – обсяг вибірки досліджуваних.

**Приклад 1.10.** Необхідно розрахувати 95 % довірчий інтервал екстраверсії нормально розподілених даних з параметрами: $\bar{x} = 17$; $s = 5{,}1$; $n = 30$.

$df = n - 1 = 30 - 1 = 29$.

$t = 2{,}05$ (так як ми розраховуємо 95 % ДІ, то беремо значення $t$ при $p = 0{,}05$).

$$\bar{x} - t \cdot \frac{s}{\sqrt{n}} = 17 - 2{,}05 \cdot \frac{5{,}1}{\sqrt{30}} = 15{,}1.$$

$$\bar{x} + t \cdot \frac{s}{\sqrt{n}} = 17 + 2{,}05 \cdot \frac{5{,}1}{\sqrt{30}} = 18{,}9.$$

Таким чином, з 95 % вірогідністю ми можемо стверджувати, що в діапазон від 15,1 до 18,9 балів попаде  істинне значення середнього значення екстраверсії в генеральній сукупності.

Для будь-якого типу розподілу, включаючи нормальний довірчий інтервал можна розрахувати з використанням технік «чисельного ресамплінгу» або, як їх іноді називають, «методи генерації повторних вибірок», які об'єднують чотири різних підходи, що відрізняються за алгоритмом, але близьких за суттю: рандомізація або переставний тест (permutation), метод «складного ножа» (jackknife), крос-перевірка (cross-validation) та бутстреп (bootstrap). Ці алгоритми моделюють емпіричний розподіл вибіркових характеристик, є сучасною альтернативою параметричних методів і бурхливо розвиваються останнє десятиліття. Найбільш популярний бустреп.

Основна ідея бутстрепа полягає в тому, що методом статистичних випробувань Монте-Карло багаторазово «витягуються» повторні вибірки з емпіричного розподілу. А саме: беруть сукупність з $n$ членів вихідної вибірки $x_1$, $x_2$, ..., $x_{n-1}$, $x_n$, звідки на кожному кроці з $n$ послідовних ітерацій за допомогою датчика випадкових чисел «витягується» довільний елемент $x_k$, який знову «повертається» у вихідну вибірку. Тобто використовується алгоритм «випадкового вибору з поверненням» (random sampling with replacement). У результаті деякі члени в кожній окремій псевдовибірці можуть повторюватися два або більше разів, тоді як інші – відсутні. Таким способом можна сформувати будь-яке завгодно велике число бутстреп-вибірок (зазвичай 5000-10000), що складаються з випадкових комбінацій вихідного набору елементів.

Маючи ці множини даних конструюється розподіл середніх значень і з її кінців відрізається 2,5 % площі розподілу, тобто по 2,5 % процентилю. В результаті залишається 95 % довірчий інтервал для середнього значення, що розрахований за допомогою процедури бустрепа процентильним методом. Крім нього є і інші методи бустрепу, один з найкращих – метод BCa  (Bias Corrected accelerated – прискорений бутстреп з поправкою на зміщення).

## 1.5 Представлення результатів дослідження за допомогою параметрів розподілу

Результати, отримані тим або іншим методом вимірювання, допускають різні математичні процедури. Обробка цих результатів залежить від:

1) шкали, у якій були виміряні психологічні ознаки;
2) від завдань дослідження (що саме хоче відобразити дослідник).

Тому перш ніж приступати до обробки і вибору процедур математико-статистичного аналізу даних, необхідно відповісти собі на питання – в якій шкалі були виміряні ознаки.

Для полегшення прийняття рішення про вибір параметрів розподілу слід скористатися таблицею 1.4. Параметри розподілу – це числові характеристики, які відображають основні тенденції виразності й мінливості ознаки в даній вибірці.

Таблиця 1.4 –Вибір описових параметрів вибірки в залежності від вимірювальної шкали

<table>
<tr><th>Вимірювальна шкала</th><th>Міри центральної тенденції або положення</th><th>Міри мінливості</th></tr>
<tr><td>Шкала номінальна</td><td>$Mo$ – мода</td><td>$f_i$ – абсолютна частота<br>$p_i$ – відносна частота<br>$p_{\%i}$ – процентна частота</td></tr>
<tr><td rowspan="3">Шкала порядку</td><td>$Mo$ – мода<br>$Md$ – медіана<br>$Q_1$, $Q_2$, $Q_3$ – квартилі<br>$D_1$……….$D_9$ – децилі<br>$P_1$………..$P_{99}$ – процентилі</td><td>$f_i$ – абсолютна частота<br>$F_i$ – накопичена абсолютна частота<br>$p_i$ – відносна частота<br>$P_i$ – накопичена відносна частота<br>$p_{\%i}$ – процентна частота<br>$P_{\%i}$ – накопичена процентна частота<br>$IQR$ – міжквартильний розмах<br>$A_{SQ}$ – квартильний коефіцієнт асиметрії</td></tr>
<tr><td colspan="2">У разі відповідності даних закону нормального розподілу, «довгих» шкал і/або великого обсягу вибірки</td></tr>
<tr><td>$\overline{x}$ – середнє арифметичне значення</td><td>$s$ – стандартне відхилення</td></tr>
<tr><td>Шкала інтервалів</td><td>$Mo$ – мода<br>$Md$ – медіана<br>$Q_1$, $Q_2$, $Q_3$ – квартилі<br>$D_1$……….$D_9$ – децилі<br>$P_1$………..$P_{99}$ – процентилі $\overline{x}$ – середнє арифметичне значення</td><td>Усі види частот: $f_i$, $F_i$, $p_i$, $P_i$, $p_{\%i}$, $P_{\%i}$<br>$IQR$ – міжквартильний розмах<br>$A_{SQ}$ – квартильний коефіцієнт асиметрії<br>$s^2$ – дисперсія<br>$s$ – стандартне відхилення<br>$A$ – коефіцієнт асиметрії<br>$E$ – коефіцієнт ексцесу<br>$Cv$ – коефіцієнт варіації</td></tr>
<tr><td>Шкала відношень</td><td>$Mo$ – мода<br>$Md$ – медіана<br>$Q_1$, $Q_2$, $Q_3$ – квартилі<br>$D_1$……….$D_9$ – децилі<br>$P_1$………..$P_{99}$ – процентилі $\overline{x}$ – середнє арифметичне значення<br>$G$ – середнє геометричне значення<br>$H$ – середнє гармонійне значення</td><td>Усі види частот: $f_i$, $F_i$, $p_i$, $P_i$, $p_{\%i}$, $P_{\%i}$<br>$IQR$ – міжквартильний розмах<br>$A_{SQ}$ – квартильний коефіцієнт асиметрії<br>$s^2$ – дисперсія<br>$s$ – стандартне відхилення<br>$A$ – коефіцієнт асиметрії<br>$E$ – коефіцієнт ексцесу<br>$Cv$ – коефіцієнт варіації</td></tr>
</table>

**Завдання на самостійну підготовку**

1. Історія розвитку статистичного аналізу даних.

2. Дайте порівняльну характеристику основним стратегіям формування репрезентативних вибірок дослідження.

3. Особливості вимірювання психічних явищ. «Опосередкованість» вимірювання психічних явищ. Поняття одиниці вимірювання в психології.

4. Основні властивості метричних і неметричних шкал вимірювання.

5. Правила ранжування даних.

6. Особливості попереднього аналізу даних у вибірці. Відновлення пропущених спостережень. Перевірка первинних даних на наявність аутлаєрів.

7. Візуальне представлення результатів дослідження за допомогою графіків.

8. Стандартизація даних і стандартизовані шкали в психології.

9. Характеристика основних методів первинної статистичної обробки даних.

## ЛЕКЦІЯ 2
## СТАТИСТИЧНІ ГІПОТЕЗИ ТА ОСОБЛИВОСТІ ЇХ ПЕРЕВІРКИ

### План лекції

2.1. Статистичні гіпотези та критерії
2.2. Розмір ефекту
2.3. Параметри нормального розподілу даних
2.4. Перевірка даних на відповідність закону нормального розподілу

**Час проведення:** 2 учбові години.

### Література


1. Гржибовский А.М., Унгуряну Т.Н. Анализ биомедицинских данных с использованием пакета статистических программ SPSS: учебное пособие. – Архангельск: Изд-во Северного государственного медицинского университета, 2017. – 293 с.

2. Климчук В.О. Математичні методи у психології. Навчальний посібник. – К.: Освіта України, 2009. – 288 с.

3. Корнеев А.А., Рассказова Е.И., Кричевец А.Н., Койфман А.Я. Критика методологии проверки нулевой гипотезы: ограничения и возможные пути выхода. Часть I // Психологические исследования, 2016. – Т. 9. – № 45. – С. 1.

4. Сивуха С.В., Козяк А.А. О реформе статистического вывода в психологии // Психология. Журнал Высшей школы экономики, 2009. – Т. 6. – № 4. – С. 66-86.

5. Телейко А.Б. Чорней Р.К. Математико-статистичні методи в соціології та психології: навч. посібник. – Київ: МАУП, 2007. – 418 с.

6. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edn. New York: Lawrence Earlbaum Associates, 1988. – 567 p.

7. Tomczak M., Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size // Trends in Sport Sciences, 2014. – 1(21). – p. 19-25.

## 2.1 Статистичні гіпотези та критерії

Отримані в дослідженнях вибіркові дані завжди обмежені і значною мірою мають випадковий характер. Саме тому для аналізу таких даних і використовується математична статистика, що дозволяє узагальнювати закономірності, отримані на вибірці та поширювати їх на всю генеральну сукупність.

Підкреслимо ще раз, що отримані в результаті дослідження на будь-якій вибірці дані служать підставою для судження про генеральну сукупність. Однак через дії випадкових причин оцінка параметрів генеральної сукупності, зроблену на підставі експериментальних (вибіркових) даних завжди буде супроводжуватися похибкою, і тому подібного роду оцінки повинні розглядатися як приблизні, а не як остаточні твердження.

Подібні судження про властивості й параметри генеральної сукупності за оцінками вибірки отримали назву *статистичних гіпотез*.

Суть перевірки статистичної гіпотези полягає в тому, щоб встановити чи узгоджуються експериментальні дані та сформульована гіпотеза, чи можна вважати, що розходження між гіпотезою та результатом статистичного аналізу експериментальних даних викликані випадковими факторами.

Формулювання статистичних гіпотез має певний формальний вигляд, гіпотези розглядаються парами й в кожній парі є основна (нульова) й альтернативна гіпотези.

*Нульова гіпотеза ($H_0$)* – припущенням про рівність значень, що порівнюються.

*Альтернативна гіпотеза ($H_1$)* заперечення твердження нульової гіпотези.

Нульова й альтернативна гіпотези утворюють повну групу несумісних подій: якщо одна з них вірна, то інша є помилковою, і навпаки, тому відхилення однієї з них означає прийняття іншої.

Таблиця 2.1 –Характеристика статистичних гіпотез

| | **Статистичні гіпотези** | |
|---|---|---|
| | **нульова гіпотеза** | **альтернативна гіпотеза** |
| Позначається | $H_0$ | $H_1$ |
| Призначення | гіпотеза про відсутність відмінностей | гіпотеза про існування значущих відмінностей |
| Передбачуваний результат | це те, що ми хочемо спростувати | це те, що ми хочемо довести |
| Математична модель | $\bar{x}_1 - \bar{x}_2 = 0$ | $\bar{x}_1 - \bar{x}_2 \neq 0$ |

У більшості випадків альтернативну гіпотезу можна назвати експериментальною гіпотезою – це ситуація, коли в дослідженні ставиться завдання довести існування відмінностей між вибірками (взаємозв'язку між змінними). Якщо ж дослідник хоче довести саме відсутність відмінностей, то експериментальною гіпотезою є нульова гіпотеза.

Нульова й альтернативна гіпотези можуть бути направленими (однобічними) і ненаправленими (двобічними).

Таблиця 2.2 – Види статистичних гіпотез

| | Статистичні гіпотези | |
|---|---|---|
| | нульова гіпотеза | альтернативна гіпотеза |
| Направлені гіпотези | $\bar{x}_1 \leq \bar{x}_2$<br>( $\bar{x}_1$ не перевищує $\bar{x}_2$ ) | $\bar{x}_1 > \bar{x}_2$<br>( $\bar{x}_1$ перевищує $\bar{x}_2$ ) |
| Ненаправлені гіпотези | $\bar{x}_1 = \bar{x}_2$<br>( $\bar{x}_1$ не відрізняється $\bar{x}_2$ ) | $\bar{x}_1 \neq \bar{x}_2$<br>( $\bar{x}_1$ відрізняється $\bar{x}_2$ ) |

Направлені гіпотези формулюються, якщо необхідно довести, що:

1) в одній із груп індивідуальні значення досліджуваних за ознакою вищі, а в іншій нижчі;

2) в одній із груп під дією експериментальних впливів відбулися більш виражені зміни, ніж в іншій групі.

Ненаправлені гіпотези використовуються, якщо необхідно довести, що форми розподілу ознаки у двох різних групах досліджуваних відрізняються.

Статистична перевірка гіпотез, ґрунтується на вибіркових даних, тому неминуче пов'язана з ризиком (ймовірністю) прийняття хибного рішення для генеральної сукупності. Можливі помилки двох родів – першого та другого.

Таблиця 2.3 –Статистичні рішення при оцінці гіпотез

| **Рішення дослідника** | **В дійсності** | |
|---|---|---|
| | **Істинна гіпотеза $H_0$** | **Істинна гіпотеза $H_1$** |
| Гіпотеза $H_0$ відхиляється | Хибне рішення.<br>Помилка 1-го роду.<br>Ймовірність = α. | Вірне рішення.<br>Статистична потужність.<br>Ймовірність = 1-β. |
| Гіпотеза $H_0$ приймається | Вірне рішення.<br>Довірча ймовірність.<br>Ймовірність = 1-α. | Хибне рішення.<br>Помилка 2-го роду.<br>Ймовірність = β. |

Максимально прийнятна ймовірність помилки першого роду позначається α, чим нижчий її рівень, тим більший обсяг вибірки необхідно залучити в дослідженні. Цей рівень визначає сам експериментатор до початку збору даних, як статистичну межу готовності помилитися при прийнятті рішення. Це ймовірність хибнопозитивного результату, тобто відхилення нульової гіпотези коли в генеральній сукупності вона істинна. У психології найчастіше ймовірність помилки першого роду встановлюється на рівні 0,05, хоча це лише данина традиції, кожен фахівець може вибрати свій рівень значущості в залежності від предмету дослідження. Заповітне значення 0,05 свідчить, що отриманий результат може виявитися випадковим в 1 випадку з 20. А це не так і мало, тобто вплив ймовірності на результат дослідження часто буває недооцінений, особливо в дослідженнях із залученням багатьох діагностичних параметрів.

Необхідно зазначити, що низька величина помилки першого роду не завжди корисна, так як збільшується ймовірність помилки другого роду (β) – ймовірність хибнонегативного результату, тобто прийняття рішення не відхиляти нульову гіпотезу, хоча в дійсності вона невірна. У дослідженнях розмір цієї помилки не зазначається відкрито, але про неї слід обов'язково пам’ятати формулюючи висновок, особливо працюючи з вибірками малого розміру. На практиці рекомендується, щоби цей параметр не перевищував значення в 0,2.

Помилка другого роду зворотно пов’язана зі статистичною потужністю дослідження, розраховується як 1-β. Статистична потужність – здатність виявити статистично значущі відмінності, якщо вони дійсно існують. Іншими словами – ймовірність відхилення нульової гіпотези, коли вона помилкова. Зазвичай вона визначається на рівні 80 - 90 % (0,8 - 0,9). Здебільшого у дослідженнях використовується саме цей діапазон. Проте, якщо досліджуються рідкісні явища та дуже складно сформувати необхідний обсяг вибірки, то потужність можна знизити до 70 %, щоби мати можливість все-таки провести дослідження. Чим вища статистична потужність, тим більшу кількість осіб необхідно набрати в кожну групу, причому це збільшення не пропорційне, а експоненційне, тобто кожне наступне зростання потужності в 1% вимагає все більшу кількість досліджуваних.

Перед початком дослідження критично важливо знати статистичну потужність. Є сенс розпочинати дослідження тільки в ситуації з високою ймовірністю виявлення значущого ефекту. Безвідповідально витрачати ресурси за інших умов, ми ризикуємо проводити зарання приречені на невдачу дослідження. Тобто, якщо не виявлено статистично значущих відмінностей при потужності, наприклад у 40 %, то ми отримали невизначений результат. Це може бути викликане як недостатньою кількістю досліджуваних так і дійсно відсутньою різницею. У наведеному прикладі, ймовірність не помітити ефект складає 60 %. Ми не виявили ефект, але не можемо стверджувати, що його немає. Отже, перш ніж формулювати висновок про відсутність відмінностей, дослідник повинен з’ясувати чи достатньою була потужність для її виявлення.

Необхідно дуже уважно оцінити баланс між помилками першого та другого роду при прийнятті рішення про рівень довіри до отриманих результатів та їхнє подальше використання на практиці. Зазвичай для більшості досліджень оптимальним поєднанням вважається 5 % $\alpha$-рівня та 80 % статистичної потужності. Проте, як уже зазначалося вище, баланс між ризиком прийняття неправильного рішення й дійсністю проведення дослідження визначається фахівцем індивідуально в кожному випадку, тільки йому відомо, наскільки важливі ті чи інші висновки та які можливості формування вибірки необхідного обсягу.

Ці характеристики повинен враховувати дослідник під час планування дослідження, при розрахунку необхідного обсягу вибірки.

Для більшості психологічних досліджень визначення необхідного обсягу вибірки досить проблематичне питання, адекватне обґрунтування якого відображена далеко не в кожному з них. Здебільшого дослідники орієнтуються на умовні рекомендації або на свої можливості зібрати первинні дані. Таке невірне та халатне ставлення до вирішення даного питання може нівелювати результати будь-якого дослідження, піддати під сумнів коректність інтерпретації числових результатів статистичних методів. Насправді, кожен науковець на етапі планування дослідження повинен поставити собі питання наступного типу: «Я можу вивчити Х осіб. Чи є сенс проводити таке дослідження?» або «Наскільки велику вибірку мені необхідно сформувати, щоби виявити очікувану величину ефекту?» На ці сформульовані питання можна відповісти за допомогою аналізу статистичної потужності – набору методів для планування дослідження, які дозволяють оптимізувати використання наявних ресурсів.

Не всі розуміють, що питання розміру вибірки – справа не однієї формули, а досить складна тема, яка потребує аналізу популяції на яку будуть розповсюджуватися результати, детального усвідомлення дослідницьких гіпотез та завдань, розуміння дизайну дослідження, типів вимірювальних шкал, у яких представлені дані, направленість гіпотез, відношення між вибірками, чутливість статистичних критеріїв, визначення найменшої величини ефекту, що має практичне значення, передбачення вибування осіб у тривалих проспективних дослідженнях тощо.

Виявляється, не можна взяти, наприклад 20 найдоступніших осіб, провести з ними дослідження та проаналізувати одержані дані. І причина тут не тільки в тому, що об’єктів може бути недостатньо для якісного статистичного аналізу. Проблема в тому, що за такого підходу вибірка не репрезентативна, отримані результати характерні лице для цих 20 досліджуваних. Отримані висновки є випадковими й не можуть бути поширені на всіх досліджуваних генеральної сукупності.

Формування малої вибірки в дослідженні, збільшує ймовірність прийняття помилкової гіпотези, результат часто буває недооціненим. А це означає, що люди були марно піддані опитуванню та/або експериментальному впливові. Крім того, фінансові та часові ресурси були витрачені даремно, оскільки в кінцевому рахунку дослідження не покращило практику надання психологічної допомоги.

З іншого боку, не завжди коректна широко поширена думка, що великі вибірки ідеально підходять для досліджень та статистичного аналізу. Формування значної вибірки досліджуваних пов'язана з різними труднощами. Не етично залучати до експерименту більше осіб, ніж це необхідно, тому що вони піддаються впливові, результат якого невідомий. Використання великої кількості досліджуваних також вимагає додаткових фінансових, організаційних та людських затрат, ніж це вимагають поставлені завдання дослідження. Крім того, підвищується ймовірність виявлення статистично значущих результатів, які мають низьку практичну цінністю.

Для вирішення цієї дилеми є компромісний і єдиновірний варіант – ретельне планування психологічного дослідження з розрахунком необхідного обсягу досліджуваних. Це є обов'язковою умовою якісного дослідження в межах методології доказової психології та гарантує, якщо практично важливий ефект існує, то він із високим ступенем ймовірності буде виявлений статистичним аналізом. Такий підхід дозволяє приймати рішення у своїх дослідженнях більш раціонально та обґрунтовано. Він також корисний для критичної оцінки нової інформації отриманої за читанням наукової літератури.

Загалом висновки, які можна зробити на основі емпіричних даних, є асиметричними: гіпотеза може відхилятися, але ніколи не може бути остаточно прийнята, будь-яка гіпотеза завжди відкрита для перевірки. Водночас відхилення гіпотези не означає відхилення теорії, з якої вона випливає. На жаль, процедура експерименту ніколи не може доказати абсолютно достовірне психологічне знання.

Перевірку статистичних гіпотез здійснюють на основі статистичних критеріїв. *Статистичний критерій* – інструмент визначення рівня статистичної значущості. За допомогою нього здійснюється прийняття істинної гіпотези і відхилення хибної. Кожний із критеріїв має свої переваги та недоліки, можливості та обмеження.

Статистичні критерії бувають: *параметричні та непараметричні*.

Загалом параметричні критерії вважаються більш потужними, ніж непараметричні, якщо ознака виміряна за кількісною (метричною) шкалою й нормально розподілена. Якщо розподіл ознаки відрізняється від нормального виду, то використовують непараметричні критерії.

Статистичний критерій складається з:

1) формули розрахунку емпіричного значення критерію за вибірковими статистиками;

2) правила (формули) визначення числа ступенів свободи *df*;

3) теоретичного розподілу для даного числа ступенів свободи;

4) правила співвідношення емпіричного значення критерію з критичним значенням теоретичного розподілу для визначеного α-рівня, яке дозволяє визначити статистичну значущість.

Критичні значення критерію визначаються за відповідними таблицями критичних значень залежно від числа ступенів свободи. Позначається буквою ν або *df*. Ми будемо використовувати – *df* (degrees of freedom).

Число ступенів свободи – кількість можливих напрямків мінливості ознаки. Дорівнює числу класів варіаційного ряду мінус число умов, за яких він був сформований. Поняття «число ступенів свободи» можна пояснити на прикладі. Нехай у нас є рівняння X1 + X2 + X3 = 10. Дана сума може бути одержана при різних значеннях змінних, наприклад: 2 + 5 + 3 = = 10; 3 + 3 + 4 = 10; 1 + 0 + 9 = 10; 6 + 7 + (-3) = 10 тощо. Два доданки можуть бути будь-якими числами, а останній має доповнювати суму перших двох до десяти, тобто він не є вільним, а «зв'язаним». Отже, у даному прикладі можливе число змін дорівнює двом, а в загальному випадку для *n* доданків це число дорівнює $n - 1$.

Для кожного статистичного критерію визначення числа ступенів свободи має свою специфіку. Тому в кожному алгоритмі розрахунку статистичного критерію вказується формула для визначення числа ступенів свободи. Як правило, це число лінійно залежить від обсягу вибірки або від числа ознак *k* чи їх градацій – чим більші ці показники, тим більше число ступенів свободи.

*Статистична значущість (p-value)* – це розрахована ймовірність того, що даний або більш екстремальний результат для вибірки визначеного обсягу може бути отриманий за умови, що нуль-гіпотеза вірна в генеральної сукупності. Якщо p-значення менше α-рівня, то гіпотеза вважається відхиленою як малоймовірною для отриманих результатів.

Найбільш розповсюдженими помилками в інтерпретації значення статистичної значущості є розуміння його як (Сивуха, Козяк, 2009):

1) показника істинності нуль-гіпотези;

2) міри ймовірності випадкового отримання результатів;

3) об'єктивного засобу перевірки нуль-гіпотези;

4) міри доказу істинності альтернативної гіпотези;

5) міри відтворення результатів;

6) показника практичної важливості отриманих результатів.

Історично склалося так, що при дослідженні психологічних явищ виділяють три рівні статистичної значущості критерію, який визначається співвідношенням його емпіричного та критичного значення:

- нижчим рівнем статистичної значущості вважають 5 %-й рівень ($p \leq 0,05$);

- достатнім рівнем статистичної значущості вважають 1 %-й рівень ($p \leq 0,01$);

- вищим – 0,1%-й рівень ($p \leq 0,001$).

Тому в таблицях критичних значень вказують значення критеріїв, що відповідають рівням статистичної значущості $p \leq 0,05$, $p \leq 0,01$ та $p \leq 0,001$. Чим менше *p*-значення, тим більше підстав для того, щоби відхилити $H_0$, на користь $H_1$, і підтвердити експериментальну гіпотезу.

Припустимо, що при порівнянні двох середніх значень було отримано $p = 0,05$. Це означає, що ймовірність виявлених таких або більш виражених відмінностей становить не більше 5 % в ситуації вірності нульової гіпотези.

Необхідно пам’ятати, чим більше гіпотез перевіряється, тим вищий шанс отримати випадковий результат. Це відображається в проблемі множинного порівняння, *p*-значення збільшується пропорційно кількості перевірених гіпотез.

## 2.2 Розмір ефекту

Проблему вимірювання розміру ефекту актуалізував Джейкоб Коен, як частини розробленої ним парадигми метааналізу (Cohen, 1988). Він зміг донести сенс та важливість поняття «розмір ефекту» до спеціалістів соціальних наук та розробив найбільш популярні індекси оцінки її величини. Позначається абревіатурою *ES* (effect size).

Розмір ефекту – кількісне відображення величини (ступеня прояву) деякого явища, що представляє науковий інтерес. Він відображає силу взаємозв’язку між явищами, величину різниці ознаки між групами, відповідність теоретичної моделі емпіричним даним та дозволяє оцінити корисність, важливість, цінність, осмисленість такого відношення. Розмір ефекту за своєю суттю близький до поняття практичної значущості дослідження – головного орієнтира наукової роботи.

Розміри ефекту дозволяють дослідникам вийти за межі дихотомічного мислення: «Є відмінність (взаємозв’язок) чи немає?» до більш важливого аспекту: «Яка величина цих відмінностей (сила взаємозв’язку)?», відійти від акцентуванні на ідентифікації статистичної значущості та перейти до більш цінного і простішого для розуміння кількісного опису величини розміру ефекту, більш наукового підходу до накопичення знань. Окрім того класичні показники розміру ефекту не залежить від обсягу вибірки на відміну від статистичної значущості, яку за це постійно критикують. Статистична значущість в експериментах із різною кількістю досліджуваних, але з однаковими базовими описовими характеристиками (наприклад, середніми арифметичними значеннями, стандартними відхиленнями, довірчими інтервалами) буде відрізнятися, а оцінка їх розміру ефекту – ні.

У якості показника розміру ефекту можуть виступати нестандартизовані величини (різниця між двома середніми арифметичними значеннями, відсотками, нестандартизовані регресійні коефіцієнти тощо) та стандартизовані індекси, які розраховуються для перетворення ефекту в зрозумілий масштаб.

Нестандартизовані величини розміру ефекту корисні, коли досліджувані змінні мають безпосередньо зрозуміле значення (наприклад, швидкість прийняття рішення). Стандартизовані індекси величини ефекту необхідно застосовувати, коли вимірювання не має внутрішнього значення, наприклад, числа в шкалі Лайкерта; коли в дослідженнях використовувалися різні шкали, тому пряме порівняння результатів неможливе; або коли розмір ефекту розглядається в контексті мінливості в популяції.

Вибір індексу розміру ефекту в конкретному дослідженні залежить від шкали вимірювання, використаних статистичних критеріїв та дизайну дослідження (табл. 2.4).

Таблиця 2.4 –Індекси кількісної оцінки розміру ефекту

| **Індекс ES** | **Статистичний критерій** | **Формула розрахунку індексу ES** |
|---|---|---|
| *d*-Коена | *t*-Стьюдента для незалежних вибірок (за умови рівності обсягів вибірок та гомогенності дисперсій). | $d = \dfrac{\lvert \overline{X}_1 - \overline{X}_2 \rvert}{\sqrt{\dfrac{s_1^2 + s_2^2}{2}}}$ |
| *g*-Хеджесса | *t*-Стьюдента для незалежних вибірок (за умови різної кількості досліджуваних у групах або/та малі обсяги вибірок). | $g = \dfrac{\lvert \overline{X}_1 - \overline{X}_2 \rvert}{\sqrt{\dfrac{(n_1-1)\cdot s_1^2 + (n_2-1)\cdot s_2^2}{n_1+n_2-2}}}$ |
| *Δs*-Гласса | *t*-Стьюдента для незалежних вибірок (за умови гетерогенності дисперсій у вибірках). | $\Delta s = \dfrac{\lvert \overline{X}_1 - \overline{X}_2 \rvert}{\sqrt{s^2_{контрол.}}}$ |
| $d_z$-Коена | *t*-Стьюдента для залежних вибірок | $d_z = \dfrac{\lvert \overline{X}_d \rvert}{s_d}$ |
| $d_{os}$-Коена | *t*-Стьюдента для однієї вибірки | $d_{os} = \dfrac{\lvert \overline{X} - A \rvert}{s}$ |
| *r* | *U*-Манна-Уітні, *T*-Вілкоксона | $r = \dfrac{z}{\sqrt{n}}$ |
| $\varepsilon^2$ | *H*-Крускала-Уолліса | $\varepsilon^2 = H \cdot \dfrac{n+1}{n^2-1}$ |

Продовження таблиці 2.4

| Індекс ES | Статистичний критерій | Формула розрахунку індексу ES |
|---|---|---|
| $W$-Кендалла | $\chi_r^2$ -Фрідмана | $W = \frac{\chi_r^2}{n(k-1)}$ |
| $w$-Коена | $\chi^2$-Пірсона для однієї вибірки | $w = \sqrt{\sum_{i=1}^{k} \frac{(p_{емп.} - p_{теор.})^2}{p_{теор.}}}$ |
| $\varphi$ | $\chi^2$-Пірсона (таблиця зв'язаності ознак має розмір 2×2) | $\varphi = \sqrt{\frac{\chi^2}{n}}$ |
| $V$-Крамера | $\chi^2$-Пірсона (таблиця зв'язаності ознак має відмінний від 2×2 розмір) | $V = \sqrt{\frac{\chi^2}{n \cdot (c-1)}}$ |
| $g$-Коена | Біноміальний критерій m | $g = \lvert P - 0,5 \rvert$ |
| $r^2$ | Коефіцієнт лінійної кореляції $r_{xy}$ Пірсона | $r^2 = {r_{xy}}^2$ |
| $r$ | Коефіцієнт контингенції К. Пірсона $\varphi$;<br>Коефіцієнт асоціації Д. Юла $Q$;<br>Коефіцієнти взаємної зв'язаності ознак Пірсона $C$, Чупрова $K$ та Крамера $V$;<br>Точково-бісеріальний коефіцієнт кореляції $r_{pb}$;<br>Рангово-бісеріальний коефіцієнт кореляції $r_{rb}$;<br>Коефіцієнти рангової кореляції $r_s$ Спірмена, $\tau$ Кендалла;<br>Коефіцієнт лінійної кореляції $r_{xy}$ Пірсона. | Їхні формули розрахунку коефіцієнтів кореляції |

Необхідно зазначити, що наведені стандартизовані індекси ES не єдино можливі для зазначених статистичних критеріїв. Звернувшись до спеціалізованої літератури можна виявити альтернативні міри оцінки розміру ефекту. Крім того, вони майже всі пов'язані між собою, за необхідності можна трансформувати одні індекси в інші. Така універсальність дозволяє порівнювати результати різних досліджень у метааналітичних дослідженнях для отримання більш надійних оцінок параметрів генеральної сукупності.

Оскільки стандартизовані величини ефекту не мають одиниць вимірювань, їхня інтерпретація для кожного індексу зводиться до питання який ефект вважається малим, середнім або великим. Традиційно для вирішення даної проблеми застосовується два підходи. Перший із них є порівняння виявленого ефекту з типовими ефектами отриманими іншими дослідниками у відповідній галузі знань. Другим підходом при інтерпретації є орієнтація на «еталонні» значення. Найбільш визнані в поведінкових науках є методичні рекомендації Дж. Коена (табл. 2.5).

Таблиця 2.5. –Узагальнена інтерпретація розмірів ефекту

| **Індекси ES** | **Інтерпретація величини розміру ефекту** |
|---|---|
| *d*-Коена, *g*-Хеджесса, *Δs*-Гласса, $d_z$-Коена, $d_{os}$-Коена | 0,00 ≤ ES < 0,20 – несуттєвий;<br>0,20 ≤ ES < 0,50 – малий;<br>0,50 ≤ ES < 0,80 – середній;<br>0,80 ≤ ES – великий. |
| *r, φ, W*-Кендалла, *w*-Коена, | 0,00 ≤ \| ES \| < 0,10 – несуттєвий;<br>0,10 ≤ \| ES \| < 0,30 – малий;<br>0,30 ≤ \| ES \| < 0,50– середній;<br>0,50 ≤ \| ES \| < 1,00 – великий. |
| $\varepsilon^2$ | 0,00 ≤ $\varepsilon^2$ < 0,01 – несуттєвий;<br>0,01 ≤ $\varepsilon^2$ < 0,08 – малий;<br>0,08 ≤ $\varepsilon^2$ < 0,26 – середній;<br>0,26 ≤ $\varepsilon^2$ – великий. |
| *g*-Коена | 0,00 ≤ *g* < 0,05 – несуттєвий;<br>0,05 ≤ *g* < 0,15 – малий;<br>0,15 ≤ *g* < 0,25 – середній;<br>0,25 ≤ *g* – великий. |
| $r^2$ | 0,00 ≤ $r^2$ < 0,01 – несуттєвий;<br>0,01 ≤ $r^2$ < 0,09 – малий;<br>0,09 ≤ $r^2$ < 0,25– середній;<br>0,25 ≤ $r^2$ < 1,00 – великий. |

Тим не більше, жоден із традиційних підходів до інтерпретації розміру ефекту в психологічній галузі знань не позбавлений серйозних проблем. Порівняння виявленого ефекту з іншими дослідниками ускладняється низьким відтворенням результатів експериментів. Наприклад, у проекті над яким працювали 270 осіб був представлений звіт про повторне відтворення 100 досліджень, що були опубліковані в провідних психологічних журналах. Як результат, понад 60 % реплікацій не підтвердили статистичну значущість висновків на рівні 0,05. Середній розмір ефекту реплікацій був удвічі менший у порівнянні з опублікованими результатами, що відображає суттєве зменшення (Open Science Collaboration, 2015).

З іншої сторони, використання еталонних стандартів у психології під час інтерпретації величини розміру ефекту втрачає сенс через значну різнорідність психологічних наукових напрямів. Т. Шефер та М. Шварц виявили значну відмінність у медіанах розподілів розмірів ефекту в різних галузях психології. Найбільші показники розмірів ефекту публікуються в дослідження з експериментальної та біологічної психології, а в соціальній психології та психології розвитку – значно менші.

Одне з прагматичних рішень для розв’язання проблеми інтерпретації запропонував сам Дж. Коен: відображати розмір ефекту в нестандартизованих величинах та інтерпретувати їхнє практичне значення в контексті проведеного дослідження.

Тобто запропоновані узагальнені градації розмірів ефекту необхідно сприймати тільки в якості орієнтирів, бо для кожного окремого дослідження повинен застосовуватися індивідуальний підхід. Оцінюючи необхідно враховувати теоретичні, економічні, етичні та практичні аспекти.

Існує декілька типових ситуацій інтерпретації результатів дослідження з урахуванням оцінки розміру ефекту після тестування гіпотези:

1) якщо немає достатніх доказів для відхилення $H_0$ (не виявили відмінностей), а *ES* середній чи великий, не можна стверджувати про відсутність відмінностей. Це привід здійснити додаткове дослідження збільшивши обсяг вибірки;

2) якщо $H_0$ не відхилено, а *ES* несуттєвий чи малий, можна констатувати, що відмінностей швидше всього дійсно немає;

3) якщо є підстави для відхилення $H_0$ (виявили відмінності), а *ES* несуттєвий чи малий, необхідно добре подумати чи мають такі відмінності психологічне значення (можливо вони взагалі в межах помилки вимірювання);

4) якщо $H_0$ відхилено, а *ES* великий є підстави стверджувати, що відмінності дійсно існують.

Оцінка розміру ефекту до тестування гіпотези особливо актуальна в експериментальних дослідженнях. Вона надає можливість розрахувати необхідний обсяг вибірки під час планування роботи, що дуже важливо при обмеженому доступі до досліджуваних. У такому випадку, необхідно додатково визначити прийнятний рівень помилку першого роду та статистичну потужність дослідження. Для цього можна скористатися безоплатним програмним забезпеченням, таким як пакетом GPower (https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html) або програмно-статистичним середовищем із відкритим кодом R (https://www.r-project.org/).

Окрім того, увага до розміру ефекту не вичерпується цікавістю до конкретного дослідження, важливим аспектом цього інтересу є бажання уніфікувати звіти про результати одиничних досліджень так, щоби їх можна було використовувати в метааналітичних проектах. З цією метою науковому товариству необхідно звернути увагу на стандарти розроблені APA з публікації емпіричних даних, тому що раніше висловлене твердження С.В. Сивухою не втратило своєї актуальності й сьогодні: «У нашому розпорядженні немає вихідного матеріалу для метааналізу – достатньої кількості публікацій, які присвячені одній темі та написані відповідно до норм співставлення: детальним і чітким описом процедурних особливостей дослідження та його результатів в термінах статистичних ефектів. У вітчизняній психології немає репрезентативною бази даних публікацій, що містять релевантну статистичну інформацію».

Отже, є достатньо вагомі причини дослідникам доповнити свої звіти про перевірку значущості нульової гіпотези інформацією про розмір ефекту та його довірчі інтервали. У якості аргументів виділяють такі причини:

1. На відміну від значення р-значення розмір ефекту відображає цінність результатів дослідження.
2. Дозволяє оцінити ступінь прояву явища незалежно від обсягу вибірки.
3. Представлені результати дослідження в стандартизованому вигляді дозволяють порівнювати між собою ознаки, що діагностовано в різних вимірювальних шкалах.
4. Використовується для визначення необхідного обсягу вибірки в запланованих дослідженнях.
5. Дозволяє оцінити статистичну потужності у вже проведених дослідженнях.
6. Розмір ефекту незамінний у проведенні метааналітичних досліджень, дозволяючи порівнювати між собою результати різних авторів.
7. Представлений разом із довірчим інтервалом розмір ефекту може використовуватися для оцінювання параметрів генеральної сукупності.

## 2.3 Параметри нормального розподілу даних

При прийнятті рішення, який із статистичних критеріїв застосувати при аналізі результатів дослідження, важливо оцінити характер розподілу даних. Якщо ознаки  мають нормальний або близький до нормального розподіл, то можна використовувати методи параметричної статистики. У випадках відхилення від нормальності перевагу слід надавати методам менш вибагливим до виду розподілу даних – непараметричним критеріям. Часто в психології перевіряють на відповідність даних закону нормального розподілу щоби визначити у якій вимірювальній шкалі представлена ознака – у інтервальній або порядковій.

Параметри розподілу – це числові характеристики, які вказують, де «в середньому» розташовуються значення ознаки та наскільки ці значення мінливі. Найбільш практично важливими параметрами для оцінки нормальності розподілу даних є середнє арифметичне значення ($\bar{x}$), середньоквадратичне відхилення ($s$), асиметрія ($A$) та ексцес ($E$).

У реальних психологічних дослідженнях ми оперуємо не параметрами, а їхніми наближеними значеннями, так званими оцінками параметрів. Це пояснюється обмеженістю вибірок дослідження. Чим більше вибірка, тим ближча оцінка параметра до його істинного значення. Надалі, кажучи про параметри, ми будемо мати на увазі їхні оцінки.

У статистиці є різні види розподілів випадкових величин: нормальний, рівномірний, біноміальний, хі-квадрат, Пуасона, Стьюдента, Фішера тощо.

Нормальний розподіл найчастіше використовують у психологічних дослідженнях.

Нормальним такий розподіл називається тому, що він найчастіше трапляється в природничо-наукових дослідженнях і є «нормою» будь-якого масового випадкового прояву ознак. Цей вид розподілу відкритого трьома вченими в різний час: Муавром у 1733 році в Англії, Гауссом у 1809 році в Німеччині та Лапласом у 1812 році у Франції. Нормальний розподіл є найбільш теоретично вивченим та зручним для математичного аналізу даних. На основі нього розроблені найбільш потужні прийоми й методи статистичного аналізу.

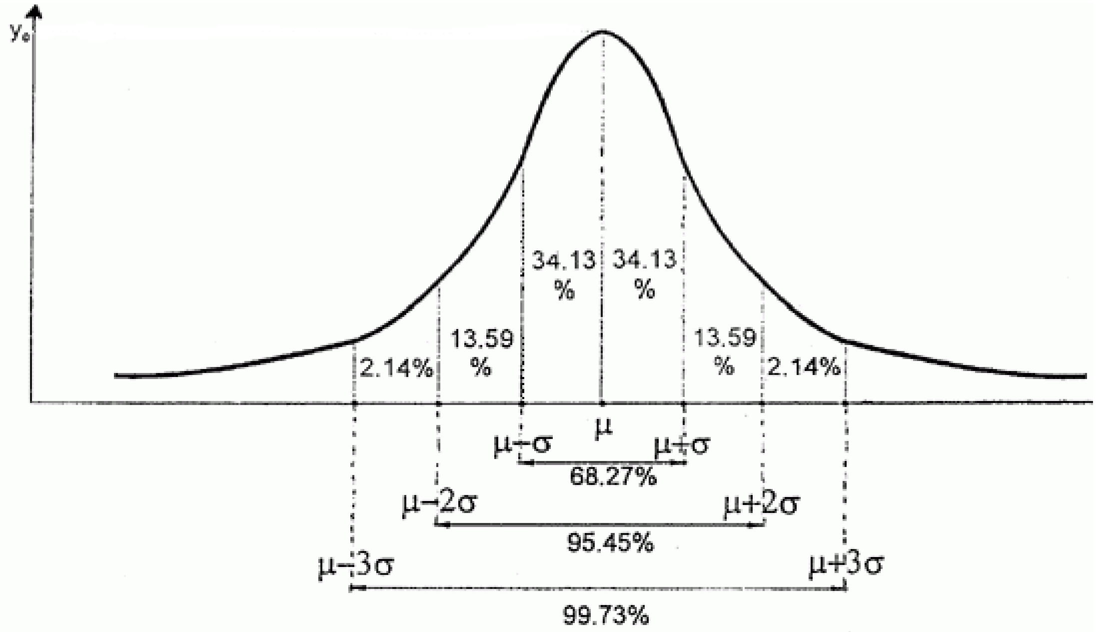


**Рис. 2.1 – Крива нормального розподілу**

Нормальний розподіл має такі характеристики:

1. Крайні значення ознаки в ньому досить рідкісні, а значення, близькі до середньої величини – досить часті.

2. Графічно нормальний розподіл має вигляд кривої дзвоноподібного типу.

3. Крива розподілу симетрична щодо центру.

4. Середнє арифметичне значення, мода і медіана у нормальному розподілі рівні.

5. Форму розподілу можна описати двома параметрами: середнім арифметичним значенням і стандартним відхиленням.

6. Слід запам'ятати, що $\pm 1\sigma = 68{,}27$ % площі розподілу; $\pm 2\sigma = 95{,}45$ %; $\pm 3\sigma = 99{,}73$ %.

Якщо стандартне відхилення стале, а величина середнього значення змінюється, то форма кривої нормального розподілу залишається незмінною, а лише її графік зміщується вправо або вліво на осі абсцис. За умови сталої величини середнього значення зміна стандартного відхилення провокує зміну тільки ширини кривої, зі зменшенням сигми крива стає більш вузькою та піднімається водночас вверх, а зі збільшенням сигми крива розширюється, але опускається вниз. Проте у всіх випадках крива нормального розподілу залишається симетричною щодо середнього значення, зберігаючи правильну дзвоноподібну форму.

Незважаючи на те, що теоретично нормальний закон розподілу передбачає існування будь-яких нескінченно малих і нескінченно великих значень, у психології випадкові змінні мають кінцеві діапазони існування. На практиці, як правило, використовуються функції нормального розподілу, що обмежені зліва і справа відхиленням у діапазоні $3\sigma$.

У ситуаціях, коли які-небудь причини сприяють більш частій появі значень, які вищі або, навпаки, нижчі середнього значення, утворюються асиметричні розподіли.

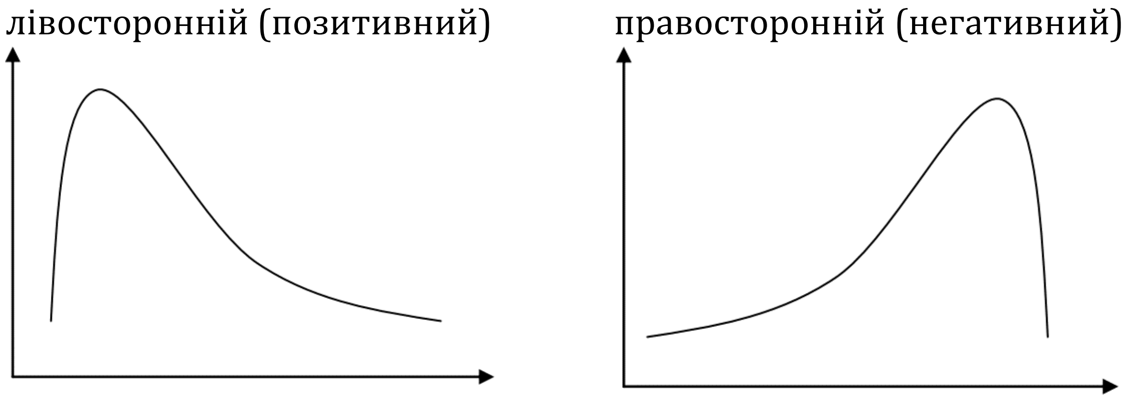


**Рис. 2.2 – Асиметричні розподіли**

У разі лівосторонньої (позитивної) асиметрії в розподілі частіше трапляються більш низькі значення ознаки, а в ситуації правосторонньої (негативної) – більш високі.

Про симетрію можна судити за показником асиметрії. *Асиметрія* ($A$) характеризує ступінь несиметричності розподілу щодо його середнього арифметичного значення. Асиметрія обчислюється за формулою:

$$A = \frac{\Sigma\left(x_i - \bar{x}\right)^3}{n \cdot s^3}.$$

Якщо обсяг вибірки невеликий ($n < 30$) для визначення вибіркового значення асиметрії необхідно застосовувати точну розрахункову формулу:

$$A = \frac{n \cdot \Sigma\left(x_i - \bar{x}\right)^3}{(n-1)\cdot(n-2)\cdot s^3},$$

де $x_i$ – значення ознаки $x$ у конкретного досліджуваного;

$\bar{x}$ – середнє арифметичне значення;

$s$ – середньоквадратичне відхилення;

$n$ – обсяг вибірки досліджуваних.

Для лівосторонніх розподілів $A > 0$; для правосторонніх $A < 0$; для симетричних розподілів $A = 0$.

У ситуаціях, коли які-небудь причини сприяють більш частій появі середніх або близьких до середніх значень, утворюється розподіл із позитивним ексцесом. Якщо ж у розподілі переважають крайні значення, причому одночасно більш низькі й більш високі, то такий розподіл характеризується негативним ексцесом, а в центрі розподілу може утворитися прогин.

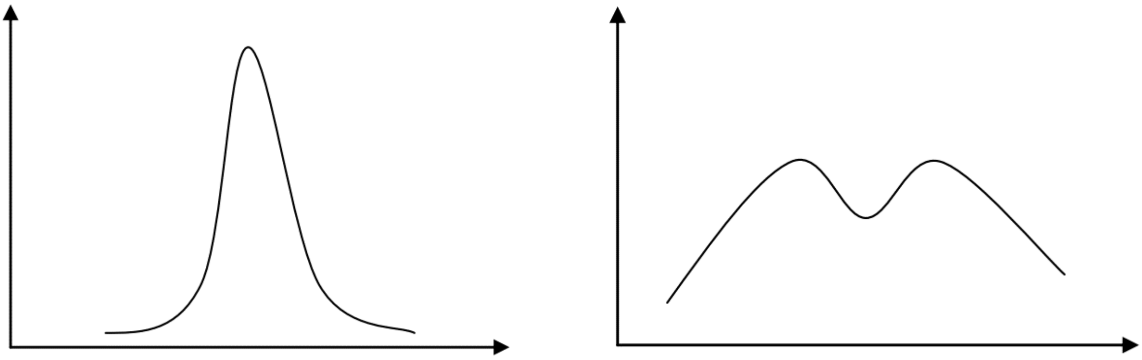


**Рис. 2.3 – Розподіли з вираженим ексцесом**

*Ексцес (E)* характеризує відносну опуклість або згладженість розподілу вибірки в порівнянні з нормальним розподілом. Позитивний ексцес характеризує порівняно загострений розподіл, негативний – порівняно згладжений.

Показник ексцесу визначається за формулою:

$$E = \frac{\Sigma\left(x_i - \bar{x}\right)^4}{n \cdot s^4} - 3.$$

Якщо обсяг вибірки невеликий ($n < 30$) для визначення вибіркового значення ексцесу необхідно застосовувати точну розрахункову формулу:

$$E = \frac{n \cdot (n+1) \cdot \Sigma\left(x_i - \bar{x}\right)^4}{(n-1) \cdot (n-2) \cdot (n-3) \cdot s^4} - \frac{3 \cdot (n-1)^2}{(n-2) \cdot (n-3)},$$

де $x_i$ – значення ознаки $x$ у конкретного досліджуваного;

$\bar{x}$ – середнє арифметичне значення;

$s$ – середньоквадратичне відхилення;

$n$ – обсяг вибірки досліджуваних.

У розподілах із позитивним ексцесом $E > 0$, з негативним – $E < 0$. У нормальному розподілі $E = 0$.

Нормальний розподіл характеризується нульовою асиметрією й ексцесом.

### 2.4 Перевірка даних на відповідність закону нормального розподілу

Нормальний розподіл є одним з емпірично перевірених фактів щодо загальної природи дійсності й може розглядатися як фундаментальний закон природи. Такий розподіл характерний для багатьох ознак генеральної сукупності. Чим більший обсяг вибірки, тим більша ймовірність того, що дані будуть розподілені нормально.

Проблема виникає, якщо застосовувати статистичні критерії, що ґрунтуються на припущенні нормальності, до даних, які не є нормальними. А така ситуація в психології виникає дуже часто, оскільки психологічні дослідження здебільшого проводяться на невеликих вибірках.

У зв’язку з цим виникає необхідність перевірки даних на відповідність закону нормального розподілу.

У ситуації «ручної» перевірки (без застосування спеціалізованих статистичних пакетів обробки даних), оцінка нормальності здійснюється візуально та за допомогою спеціальних розрахунків.

1. *Візуальне визначення типу розподілу.* Необхідно сконструювати гістограму та порівняти її з характеристиками нормального розподілу.

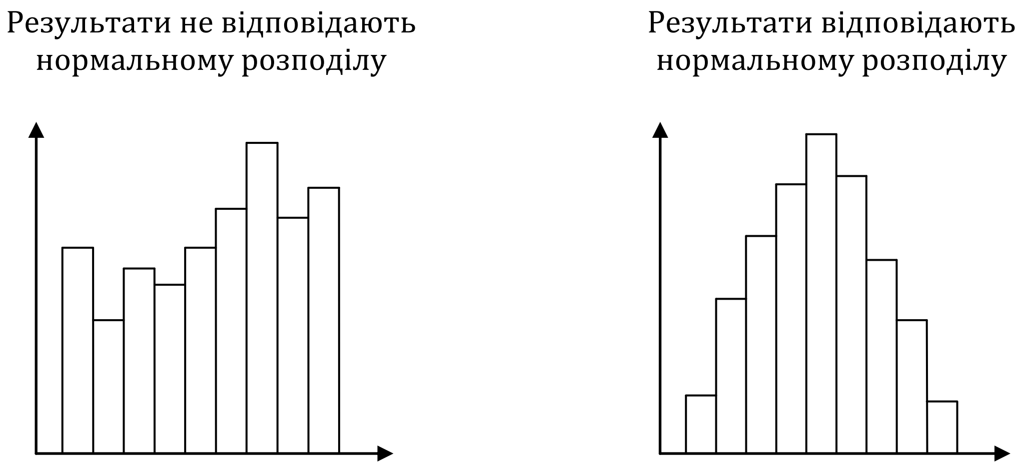


**Рис. 2.4 – Гістограми розподілів даних**

Кількість інтервалів, на які необхідно розділити вибірку досліджуваних для конструювання гістограми розподілу, можна визначити за допомогою формули Стерджеса:

$$m = 1 + 3{,}32 \cdot lg\, n,$$

де – $lg$ – десятинний логарифм;

$n$ – обсяг вибірки досліджуваних.

Для визначення кількості інтервалів можна скористатися таблицею 2.6.:

Таблиця 2.6.

| Обсяг вибірки (від - до) | 6-11 | 12-22 | 23-45 | 46-90 | 91-181 | >182 |
|---|---|---|---|---|---|---|
| Кількість інтервалів | 4 | 5 | 6 | 7 | 8 | 9 |

Вважається, що ця формула дозволяє конструювати задовільні гістограми при обсязі вибірки не більше 200 досліджуваних. В інших випадках необхідно скористатися формулами Скотта або Фрідмана Діаконіса.

Величина інтервалів визначається за формулою:

$$h = \frac{x_{\max} - x_{\min}}{n},$$

де $x_{\max}$ – максимальне значення;

$x_{\min}$ – мінімальне значення;

$n$ – обсяг вибірки досліджуваних.

2. Необхідно пам'ятати, що в нормальному розподілі даних однакові середнє арифметичне значення ($\bar{x}$), мода (Mo) і медіана (Me). Порядок і особливості їх розрахунку зазначено в матеріалі лекції № 1.

3. *Наближена кількісна оцінка нормальності розподілу* здійснюється на основі розрахунку коефіцієнтів асиметрії та ексцесу, а також помилок цих параметрів. Вказані критерії рекомендується застосовувати лише для малих вибірок.

Розглянемо критерії М.А. Плохінського та Є.І. Пустильника.

Згідно з критерієм М.А. Плохінського розподіл вважається нормальним, якщо показники асиметрії ($A$) й ексцесу ($E$) не перевищують у 3 рази свою помилку репрезентативності ($m_A$ і $m_E$).

Тобто, дані відповідають закону нормального розподілу, якщо:

$$t_A = \frac{|A|}{m_A} \leq 3 \text{ та } t_E = \frac{|E|}{m_E} \leq 3,$$

де

$$m_A = \sqrt{\frac{6}{n+3}}, \; m_E = 2 \cdot \sqrt{\frac{6}{n+3}}.$$

У статистиці під «помилкою» слід розуміти не помилку дослідження, а міру представництва даної величини, тобто наскільки величина отримана з вибіркової сукупності відрізняється від істинної в генеральній сукупності.

Згідно з критерієм Є.І. Пустильника розподіл вважається нормальним, якщо показники асиметрії ($A$) й ексцесу ($E$) менші своїх критичних значень ($A_{крит.}$ і $E_{крит.}$).

Тобто, дані відповідають закону нормального розподілу, якщо:

$$|A| \le A_{крит.} \text{ та } |E| \le E_{крит.},$$

де

$$A_{крит.} = 3 \cdot \sqrt{\frac{6 \cdot (n-1)}{(n+1)(n+3)}},$$

$$E_{крит.} = 5 \cdot \sqrt{\frac{24 \cdot n \cdot (n-2) \cdot (n-3)}{(n+1)^2 \cdot (n+3) \cdot (n+5)}}.$$

**Приклад 1.** Чи можна стверджувати, що отримані в дослідженні емпіричні дані відповідають закону нормального розподілу?

Таблиця 2.7 – Таблиця первинних даних та проміжних розрахунків

| № | *X* | $(x_i - \bar{x})$ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x})^3$ | $(x_i - \bar{x})^4$ |
|---|---|---|---|---|---|
| 1 | 4 | -4,43 | 19,62 | -86,94 | 385,14 |
| 2 | 5 | -3,43 | 11,76 | -40,35 | 138,41 |
| 3 | 8 | -0,43 | 0,18 | -0,08 | 0,03 |
| 4 | 7 | -1,43 | 2,04 | -2,92 | 4,18 |
| 5 | 9 | 0,57 | 0,32 | 0,19 | 0,11 |
| 6 | 6 | -2,43 | 5,90 | -14,35 | 34,87 |
| 7 | 8 | -0,43 | 0,18 | -0,08 | 0,03 |
| 8 | 12 | 3,57 | 12,74 | 45,50 | 162,43 |
| 9 | 10 | 1,57 | 2,46 | 3,87 | 6,08 |
| 10 | 9 | 0,57 | 0,32 | 0,19 | 0,11 |
| 11 | 6 | -2,43 | 5,90 | -14,35 | 34,87 |
| 12 | 11 | 2,57 | 6,60 | 16,97 | 43,62 |
| 13 | 10 | 1,57 | 2,46 | 3,87 | 6,08 |
| 14 | 13 | 4,57 | 20,88 | 95,44 | 436,18 |
| | $\bar{x}$ = 8,43 | Σ = -0,02 | Σ =91,43 | Σ = 6,96 | Σ = 1252,13 |

Формулюємо нульову і альтернативну гіпотези:

$H_0$: розподіл емпіричних даних відрізняється від закону нормального розподілу;

$H_1$: розподіл емпіричних даних відповідає закону нормального розподілу.

Застосуємо *критерій М.А. Плохинського*.

Для розрахунку показників асиметрії та ексцесу необхідно розрахувати середнє арифметичне значення ($\bar{x}$) та середньоквадратичне відхилення ($s$).

1. Визначаємо середнє арифметичне значення змінної $X$.

$\bar{x} = 8{,}43$.

2. Для кожного значення змінної $x_i$ знаходимо різницю $\left(x_i - \bar{x}\right)$. Сума відхилень від середнього значення повинна бути рівна нулю (з точністю до похибки обчислень).

3. Кожне відхилення від середнього значення підносимо до квадрату і фіксуємо результат у відповідний стовпчик $\left(x_i - \bar{x}\right)^2$.

Сума квадратів відхилень необхідна для обчислення стандартного відхилення.

$$s = \sqrt{\frac{\Sigma\left(x_i - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{91,43}{14-1}} = 2,65.$$

4. Обчислюємо суми результатів стовпчиків $\left(x_i - \bar{x}\right)^3$ та $\left(x_i - \bar{x}\right)^4$.

5. Розрахуємо показник асиметрії. Застосуємо точну розрахункову формулу.

$$A = \frac{n \cdot \Sigma\left(x_i - \bar{x}\right)^3}{(n-1)\cdot(n-2)\cdot s^3} = \frac{14 \cdot 6{,}96}{(14-1)\cdot(14-2)\cdot 2,65^3} = 0,035.$$

6. Розрахуємо показник ексцесу за точною розрахунковою формулою.

$$E = \frac{n\cdot(n+1)\cdot\Sigma\left(x_i - \bar{x}\right)^4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot s^4} - \frac{3\cdot(n-1)^2}{(n-2)\cdot(n-3)} = \frac{14\cdot(14+1)\cdot 1252{,}13}{(14-1)\cdot(14-2)\cdot(14-3)\cdot 2,65^4} -$$

$$-\frac{3\cdot(14-1)^2}{(14-2)\cdot(14-3)} = 3{,}107 - 3{,}841 = -0{,}734.$$

7. Визначаємо помилки репрезентативності асиметрії та ексцесу.

$$m_A = \sqrt{\frac{6}{n+3}} = \sqrt{\frac{6}{14+3}} = 0{,}59;$$

$$m_E = 2 \cdot \sqrt{\frac{6}{n+3}} = 2 \cdot \sqrt{\frac{6}{14+3}} = 1{,}18.$$

8. Згідно з критерієм М.А. Плохинського показники асиметрії й ексцесу свідчать про статистичну відмінність емпіричного розподілу від нормального в тому випадку, якщо вони перевищують за абсолютною величиною свої помилки репрезентативності в 3 рази.

$$t_A = \frac{|A|}{m_A} = \frac{|0{,}035|}{0{,}59} = 0{,}06;$$

$$t_E = \frac{|E|}{m_E} = \frac{|-0{,}734|}{1{,}18} = 0{,}62.$$

За результатами розрахунків обидва показники не перевищують утричі свої помилки репрезентативності, тому можемо зробити висновок, що розподіл відповідає закону нормального розподілу.

Тепер здійснимо перевірку за *критерієм Є.І. Пустильника*.

1. Розрахуємо критичні значення для показників асиметрії та ексцесу:

$$A_{крит.} = 3 \cdot \sqrt{\frac{6 \cdot (n-1)}{(n+1)(n+3)}} = 3 \cdot \sqrt{\frac{6 \cdot (14-1)}{(14+1) \cdot (14+3)}} = 3 \cdot 0{,}55 = 1{,}65;$$

$$E_{крит.} = 5 \cdot \sqrt{\frac{24 \cdot n \cdot (n-2) \cdot (n-3)}{(n+1)^2 \cdot (n+3) \cdot (n+5)}} = 5 \cdot \sqrt{\frac{24 \cdot 14 \cdot (14-2) \cdot (14-3)}{(14+1)^2 \cdot (14+3) \cdot (14+5)}} = 5 \cdot 0{,}78 = 3{,}9.$$

2. Згідно з критерієм Є.І. Пустильника дані відповідають закону нормального розподілу, якщо:

$$|A| \le A_{крит.} \text{ та } |E| \le E_{крит.}$$

У нашому випадку:

|0,035| < 1,65,

|-0,734| < 3,9.

Можемо зробити висновок, що дані нормально розподілені.

Отже, обидва варіанти перевірки, за критерієм М.А. Плохинського та Є.І. Пустильника, дають один і той же результат: розподіл ознаки не відрізняється від нормального розподілу. До змінної Х можливо застосувати без обмежень будь-які параметричні критерії.

Можна вибрати будь-який із запропонованих варіантів перевірки й дотримуватися його.

У разі великого обсягу вибірки раціонально здійснювати розрахунок оцінок параметрів на ЕОМ, застосовуючи спеціалізовані статистичні пакети. Рекомендуємо SPSS Statistics або Statistica.

В SPSS Statistics зіставлення емпіричного розподілу даних із нормальним розподілом реалізується такими взаємодоповнюючими методами:

1. *Визначення середнього арифметичного значення, моди і медіани*: рівність є доказом нормального розподілу даних.

2. *Конструювання гістограми розподілу*.

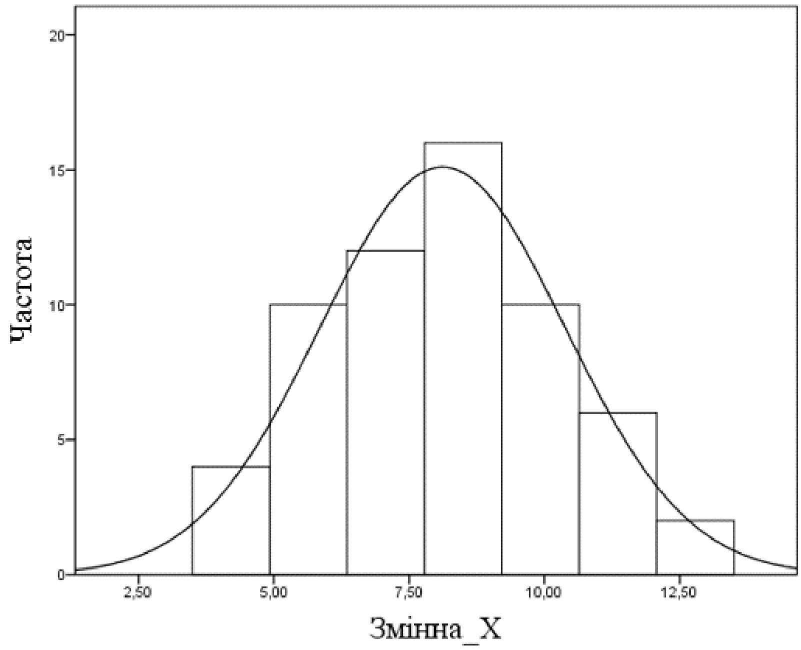


**Рис. 2.2 – Гістограма розподілу даних**

У наведеному прикладі $n$ = 60, тому згідно з формулою Стерджеса сконструйовано гістограму із семи інтервалів розподілу. Дзвоноподібна форма гістограми свідчить на користь нормального розподілу.

3. *Конструювання квантильного графіку (нормально ймовірний графік, графік Q-Q)*.

Цей графіки використовуються для оцінки нормальності розподілу змінної, тобто близькості цього розподілу до нормального. Залежність відображається на квантильному графіку.

Квантильний графік конструюється таким чином. Спочатку всі значення упорядковуються за рангом. За цими рангами розраховуються значення z (стандартизовані значення нормального розподілу) з припущенням, що дані відповідають закону нормального розподілу. Ці значення z відкладаються на осі Y графіка. Якщо емпіричні значення (відкладаються на осі X) розподілені нормально, то всі значення на графіку будуть розміщені на прямій лінії. Якщо значення не є нормально розподіленими, вони будуть відхилятися від лінії. На цьому графіку можна легко виявити аутлаєри.

За графіком видно, що емпіричні дані не суттєво відхиляються від прямої. Розташування точок на прямій лінії свідчить на користь нормального розподілу.

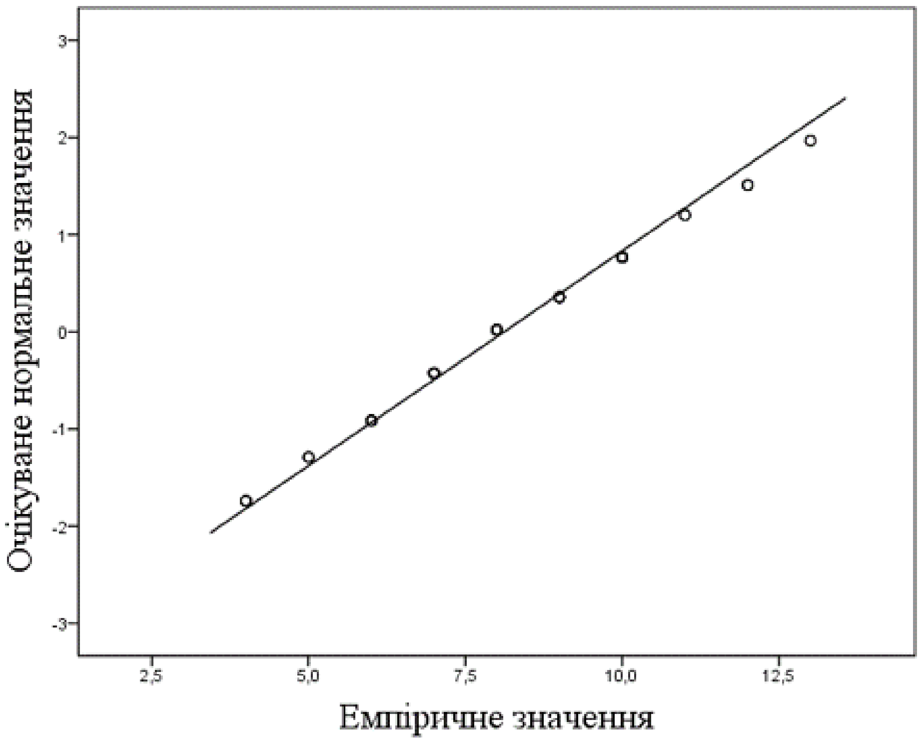


**Рис. 2.3– Квантильний графік**

4. *Конструювання діаграми типу Box Plot* (ящик з вусами, коробчаста діаграма).

Діаграми типу Box Plot – це графік, який дозволяє стисло зобразити основні параметри розподілу ймовірностей. Такий вид діаграми в зручній формі показує медіану, нижній та верхній квартилі, мінімальне та максимальне значення вибірки у діапазоні 1,5 міжквартильного розмаху, а також аутлаєри. Відстань між різними частинами ящика дозволяє визначити ступінь мінливості й асиметрію даних.

На діаграмі типу Box Plot висота прямокутника відповідає ширині міжквартильного розмаху ($IQR$), тобто інтервалу від 25 до 75 процентиля ($P_{25}$ – $P_{75}$) або першого та третього квартилю ($Q_1$ – $Q_3$). Вказана міра мінливості щодо наших даних має такий сенс: у діапазоні від 14 до 20 містяться 50 % результатів, у 25 % досліджуваних показник змінної X нижчий 14, у 25 % – вищий 20.

Горизонтальна лінія в середині прямокутника відповідає медіані ($Md$) або другому квартилю ($Q_2$).

Межі «вусів» визначаються за такою формулою:

$$X_{нижній} = Q_1 - 1{,}5 \cdot (Q_3 - Q_1);$$

$$X_{верхній} = Q_3 + 1{,}5 \cdot (Q_3 - Q_1).$$

Тобто, від першого квартилю віднімається 1,5 міжквартильного розмаху. Якщо є значення, яке менше цього результату, то це вважається аутлаєром. Верхня межа визначається як третій квартиль плюс 1,5 розмаху. Відповідно, якщо є число більше цього значення, це також аутлаєр.

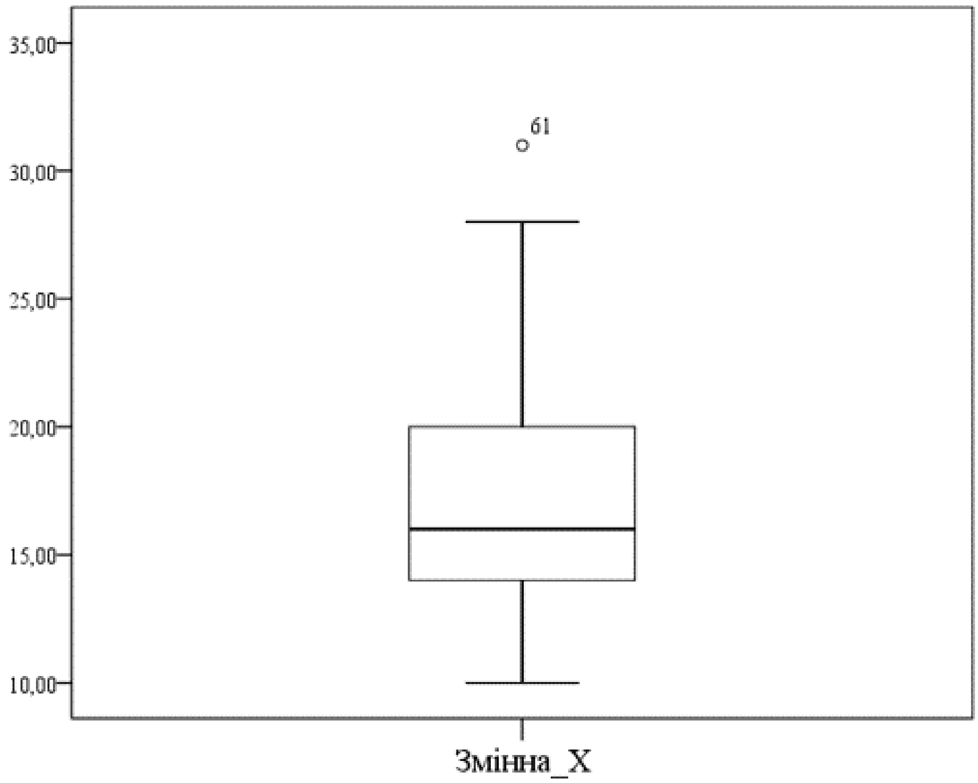


**Рис. 2.4 – Діаграма типу Box Plot**

Останнє значення з представлених даних, яке попадає в цей інтервал є межею «вуса». Значення, що визначено як аутлаєр у розрахунок не береться.

У нашому прикладі, «вусами» відмічено мінімальне (10) та максимальне значення (28) у діапазоні 1,5 міжквартильного розмаху, а кружечком – аутлаєр (результат досліджуваного № 61).

За діаграмою видно, що медіана розміщена не в центрі міжквартильного розмаху, «вуса» асиметричні (нижній коротший верхнього), є аутлаєр у результатах дослідження. Це все є свідченням невідповідності емпіричних даних нормальному розподілу.

5. *Застосування статистичних критеріїв Колмогорова-Смирнова з поправкою Ліллієфорса та Шапіро-Уілка*.

Якщо досягнуте в результаті розрахунку критеріїв значення статистичної значущості перевищує 0,05 ($p > 0,05$), то емпіричний розподіл можна вважати нормально розподіленим. Критерій Колмогорова-Смирнова з поправкою Ліллієфорса рекомендується використовувати для великих за обсягом вибірок, а критерій Шапіро-Уілка для порівняно невеликих.

Результати перевірки розподілу на нормальність слід завжди враховувати в сукупності. Так, результати перевірки за допомогою одних тільки статистичних критеріїв Колмогорова-Смирнова й Шапіро-Уілка слід інтерпретувати з обережністю, оскільки вони чутливі до обсягу вибірок. Ймовірність отримання статистично значущих відмінностей у ситуації перевірки розподілу за умови однакового відхиленні емпіричного розподілу від нормального при $n = 1000$ значно вища, ніж, наприклад, при $n = 30$. Деякі дослідники рекомендують завжди вважати розподіл відмінним від нормального при $n < 30$.

А.М. Гржибовский дає такі практичні рекомендації: при кількості спостережень від 30 до 100, якщо статистичні критерії покажуть відхилення розподілу від нормального ($p < 0,05$), слід вважати, що розподіл відрізняється від нормального, якщо графіки і значення асиметрії та ексцесу не свідчать про зворотне. Якщо кількість спостережень перевищує 100, і статистична значущість критеріїв перевірки розподілу на нормальність перевищує 0,05 ($p > 0,05$), то розподіл вважають нормальним, якщо графіки і значення асиметрії та ексцесу не говорять про зворотне.

Для умовної відповідності емпіричного розподілу нормальному, вважається, що показники асиметрії та ексцесу повинні перебувати в межах від -1 до +1.

**Завдання на самостійну підготовку**

1. Характеристика статистичних гіпотез.
2. Ідея процедури перевірки статистичної значущості нульової гіпотези.
3. Навіщо здійснювати статистичний аналіз даних?
4. Розповсюджені помилки в інтерпретації р-значення статистичної значущості при перевірці гіпотез.
5. Характеристика помилок I та II роду в дослідженні.
6. Розмір ефекту як показник практичної значущості дослідження.
7. Важливість розрахунку обсягу вибірки в дослідженні.
8. Нормальний закон розподілення даних і його застосування. Характеристики параметрів нормального закону розподілення.
9. Методи перевірки даних на відповідність закону нормального розподілу.
10. Які основні параметри z-розподілу?

## ЛЕКЦІЯ 3
## МЕТОДИ АНАЛІЗУ СТАТИСТИЧНОГО ЗВ'ЯЗКУ МІЖ ЗМІННИМИ

### План лекції

3.1. Сутність методів встановлення статистичних зв'язків
3.2. Лінійна кореляція
3.2.1. Коефіцієнт кореляції Пірсона $r_{xy}$
3.3. Рангові коефіцієнти кореляції
3.3.1. Коефіцієнт рангової кореляції $r_s$ Спірмена
3.3.2. Коефіцієнт рангової кореляції τ Кендалла
3.4. Коефіцієнти асоціативної залежності номінальних даних
3.4.1. Коефіцієнт контингенції Пірсона $\varphi$
3.4.2. Коефіцієнт асоціації Юла $Q$
3.4.3. Коефіцієнти взаємної зв'язаності ознак Пірсона $C$, Чупрова $K$ та Крамера $V$
3.5. Бісеріальні коефіцієнти кореляції
3.5.1. Точково-бісеріальний коефіцієнт кореляції $r_{pb}$
3.5.2. Рангово-бісеріальний коефіцієнт кореляції $r_{rb}$

**Час проведення**: 6 учбових годин.

### Література


1. Гланц С. Медико-биологическая статистика. Пер. с англ. – М., Практика, 1998. – 459 с.

2. Гржибовский А.М., Унгуряну Т.Н. Анализ биомедицинских данных с использованием пакета статистических программ SPSS: учебное пособие. – Архангельск: Изд-во Северного государственного медицинского университета, 2017. – 293 с.

3. Климчук В.О. Математичні методи у психології. Навчальний посібник. – К.: Освіта України, 2009. – 288 с.

4. Кричевец А. Н., Корнеев А. А., Рассказова Е. И. Основы статистики для психологов. – М.: Акрополь, 2019. – 286 с.

5. Телейко А.Б. Чорней Р.К. Математико-статистичні методи в соціології та психології: навч. посібник. – Київ: МАУП, 2007. – 418 с.

6. Харькова О.А., Соловьев А.Г. Статистические методы и математическое моделирование: учебное пособие. – Архангельск: Изд-во Северного государственного медицинского университета, 2017. – 164 с.

## 3.1 Сутність методів встановлення статистичних зв'язків

Важливим елементом математико-статистичних досліджень у психології є виявлення специфічного зв'язку між різними змінними, ознаками, факторами тощо. Ця специфічність полягає в тому, що залежність між змінними неможливо описати функціонально, тобто за допомогою однозначно визначеної функції. Зв'язок цих ознак виявляється принципово по-іншому: від того, якого значення набуде одна величина, залежать ймовірна характеристика іншої. Отже, незважаючи на те, що неможливо точно вказати значення залежної ознаки, усе ж таки відома деяка інформація про можливе розташування її значень, певного ймовірного інтервалу, середнього значення, міри мінливості тощо. Такі зв'язки у статистиці між кількісними і порядковими показниками називають кореляцію, а між номінальними показниками – асоціацією. Для загальної характеристики методів встановлення статистичних зв'язків ми будемо використовувати термін – кореляція.

*Кореляція* – статистична залежність, за якої кожному значенню однієї змінної $x$ відповідає певне очікуване значення іншої змінної $y$.

Термін «кореляція» вперше застосував французький палеонтолог Ж. Кюв'є, який вивів «закон кореляції частин і органів тварин» (цей закон дозволяє відновлювати за знайденими частинам тіла вигляд усієї тварини). У статистику зазначений термін ввів англійський біолог і статистик Ф. Гальтон.

Міри зв'язку виявляють співвідношення, як правило, між двома змінними, які виміряні на одній вибірці.

Кореляційний аналіз для двох випадкових величин складається з:

- графічного представлення залежності;
- обчислення коефіцієнта кореляції;
- перевірки статистичної значущості зв'язку.

У якості показника тісноти зв'язку між змінними використовується коефіцієнт кореляції ($r$).

*Коефіцієнт кореляції* – числова характеристика сили й напрямку ймовірного зв'язку двох змінних, що приймає значення в діапазоні від -1 до + 1.

Кореляційні зв'язки розрізняються за силою, формою та напрямком.

Показником сили зв'язку є абсолютна величина коефіцієнта кореляції. Сила зв'язку безпосередньо вказує, наскільки синхронно проявляється спільна мінливість досліджуваних змінних. Силу зв'язку можна оцінити за допомогою шкали Чедока.

Таблиця 3.1 –Класифікація зв'язку за силою (шкала Чедока)

| **Якісна характеристика сили зв'язку** | **Кількісна міра сили зв'язку** |
|---|---|
| дуже слабка | $0,00 < \lvert r \rvert \leq 0,10$ |
| слабка | $0,10 < \lvert r \rvert \leq 0,30$ |
| помірна | $0,30 < \lvert r \rvert \leq 0,50$ |
| середня | $0,50 < \lvert r \rvert \leq 0,70$ |
| сильна | $0,70 < \lvert r \rvert \leq 0,90$ |
| дуже сильна | $0,90 < \lvert r \rvert \leq 1,00$ |

Наочне уявлення про характер зв'язку дає діаграма розсіювання – графік, осі якого відповідають значенням двох змінних, а кожен досліджуваний є точкою. Більшість статистичних кореляційних критеріїв спрямовані на виявлення та інтерпретації зв'язку лінійної форми.

прямолінійна криволінійна

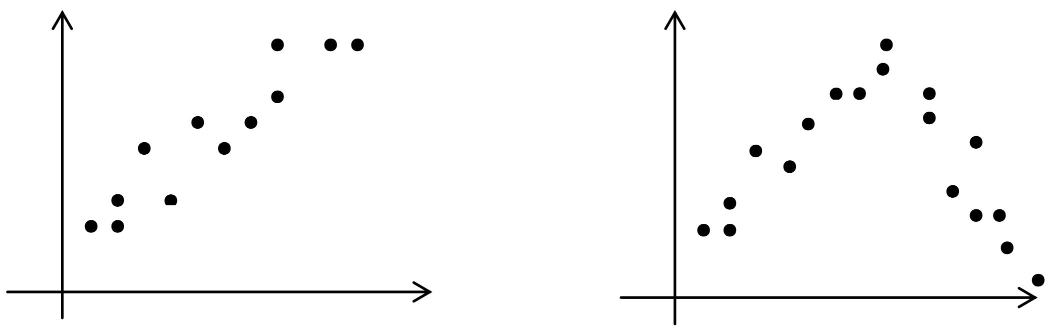

**Рис. 3.1 – Класифікація зв'язку за формою**

Показником напрямку зв'язку є знак коефіцієнта кореляції.

Позитивна (пряма) Негативна (зворотна)

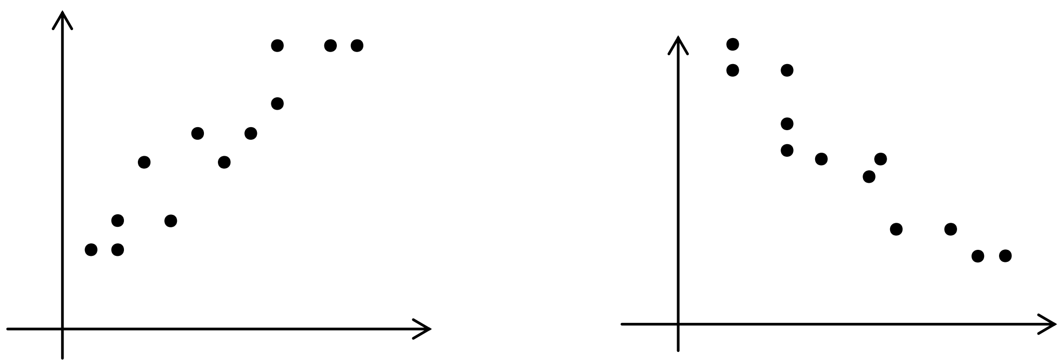

**Рис. 3.2 – Класифікація зв'язку за напрямком**

Коефіцієнти кореляції характеризуються ще і значущістю. Сильний зв'язок може виявитися випадковим у разі малого обсягу вибірки, а слабкий може бути високо значущим через великий обсяг вибірки.

Необхідно пам'ятати, що кореляційний зв'язок не можна вважати свідченням причинно-наслідкового зв'язку. Іншими словами, якщо змінні взаємопов'язані між собою, то за даним фактом ми не маємо підстав для формулювання висновку, що одна змінна впливає на іншу.

Якщо зв'язок не виявлено, але є підстави вважати, що зв'язок насправді є, слід перевірити можливі причини недостовірності зв'язку:

• *нелінійність зв'язку* – для встановлення цього необхідно проаналізувати діаграму розсіювання. Якщо зв'язок нелінійний, але монотонний необхідно перейти до рангових коефіцієнтів кореляції. Якщо зв'язок не монотонний, то доречно поділити вибірку досліджуваних на частини, у яких зв'язок монотонний й обчислити кореляції окремо для кожної частини вибірки або поділити вибірку на контрастні групи й далі порівнювати їх за рівнем вираженості ознаки;

• *наявність аутлаєрів і виражена асиметрія розподілу однієї або обох змінних.* Для цього необхідно проаналізувати гістограми розподілу частот обох ознак. У разі наявності аутлаєрів або асиметрії необхідно видалити аутлаєри або перейти до рангової кореляції;

• *неоднорідність вибірки* (проаналізувати діаграму розсіювання). Спробувати розділити вибірку на частини, у яких зв'язок може мати різні напрямки.

Якщо ж зв'язок статистично значущий, то перш ніж формулювати змістовний висновок, необхідно виключити можливість помилкової кореляції:

• *зв'язок зумовлений аутлаєрами*. У разі наявності аутлаєрів перейти до рангових кореляцій або видалити їх;

• *зв'язок зумовлений впливом третьої змінної*. Якщо є таке явище, необхідно обчислити кореляцію не тільки для всієї вибірки, а і для кожної групи.

При оцінці кореляційного зв'язку необхідно визначити типи шкал, у яких виміряні змінні, вид розподілу первинних даних та форму залежності.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення коефіцієнта $r$, тому додаткових розрахунків здійснювати не потрібно.

Для полегшення завдання вибору коефіцієнта кореляції залежно від типу вимірювальної шкали, у яких представлені змінні, доцільно скористатися таблицею 3.3.

Таблиця 3.2 – Методичні рекомендації інтерпретації величину розміру ефекту кореляційної залежності запропоновані Дж. Коеном

| **Рівень розміру ефекту** | **Величина коефіцієнта кореляції** |
|---|---|
| несуттєвий | $0{,}00 \leq \lvert r \rvert < 0{,}10$ |
| малий | $0{,}10 \leq \lvert r \rvert < 0{,}30$ |
| середній | $0{,}30 \leq \lvert r \rvert < 0{,}50$ |
| великий | $0{,}50 \leq \lvert r \rvert < 1{,}00$ |

Таблиця 3.3 – Статистичні методи знаходження взаємозв’язку між змінними

<table>
<tr><th colspan="2" rowspan="2">Вимірювальні шкали</th><th colspan="2">Шкала номінальна</th><th rowspan="2">Шкала порядку</th><th rowspan="2">Шкали інтервалів та відношень</th></tr>
<tr><th>k = 2</th><th>k > 2</th></tr>
<tr><td rowspan="2">Шкала номіна-льна</td><td>k = 2</td><td>Коефіцієнт контингенції Пірсона $\varphi$<br><br>Коефіцієнт асоціації Юла $Q$</td><td></td><td></td><td></td></tr>
<tr><td>k > 2</td><td>Коефіцієнт взаємної зв’язаності Чупрова $K$<br><br>Коефіцієнт взаємної зв’язаності Крамера $V$</td><td rowspan="3">Коефіцієнт взаємної зв’язаності Пірсона $C$<br><br>Коефіцієнт взаємної зв’язаності Чупрова $K$<br><br>Коефіцієнт взаємної зв’язаності Крамера $V$</td><td></td><td></td></tr>
<tr><td colspan="2">Шкала порядку</td><td>Рангово-бісеріальний коефіцієнт кореляції $r_{rb}$</td><td rowspan="2">Коефіцієнт кореляції $r_s$ Спірмена<br><br>Коефіцієнт кореляції $\tau$ Кендалла</td><td></td></tr>
<tr><td colspan="2">Шкали інтервалів та відношень</td><td>Точково-бісеріальний коефіцієнт кореляції $r_{pb}$</td><td>Коефіцієнт кореляції Пірсона $r_{xy}$<br><br>Кореляційне відношення $\eta$</td></tr>
</table>

Примітка: k – кількість градацій номінальної змінної.

## 3.2 Лінійна кореляція

### 3.2.1 Коефіцієнт лінійної кореляції Пірсона $r_{xy}$

Для вивчення лінійного взаємозв’язку двох змінних, виміряних у кількісних шкалах (інтервальній або відношень) застосовується коефіцієнт кореляції $r_{xy}$ Пірсона. Коли говорять про кореляцію і не уточнюють деталі – мається на увазі саме цей метод. Коефіцієнт кореляції $r_{xy}$ Пірсона фіксує наявність лише лінійного зв'язку між змінними і приймає значення від –1 до +1. Знак вказує на напрямок взаємозв'язку, а абсолютне значення коефіцієнта – на силу зв'язку.

Коефіцієнт кореляції $r_{xy}$ Пірсона є параметричним методом і його коректне застосування можливе тільки у випадку, якщо розподілення даних двох змінних не відрізняється від нормального виду.

Значення коефіцієнта кореляції $r_{xy}$ Пірсона визначається за формулою:

$$r_{xy} = \frac{\Sigma\left(x_i - \overline{x}\right)\cdot\left(y_i - \overline{y}\right)}{\left(n-1\right)\cdot s_x \cdot s_y},$$

де $x_i$, $y_i$ – первинні значення змінних $X$ та $Y$;

$\overline{x}$, $\overline{y}$ – середнє арифметичне змінних $X$ та $Y$;

$s_x$, $s_y$ – стандартне відхилення змінних $X$ та $Y$;

$n$ – обсяг вибірки досліджуваних.

Якщо емпіричне значення $r_{xy}$ дорівнює або перевищує критичне значення для заданого α-рівня, то $H_0$ відхиляється і формулюється висновок, що коефіцієнт кореляції статистично значуще відрізняється від нуля.

Показником розміру ефекту зв'язку між змінними окрім безпосередньо самого емпіричного значення коефіцієнта $r_{xy}$ може виступати й коефіцієнт детермінації $r^2$, тобто коефіцієнт кореляції в квадраті. Він змінюється від 0 до 1 (або від 0 до 100 %) і показує в якій мірі мінливість однієї змінної обумовлена (детермінована) впливом іншої змінної. На підставі показника $r^2$ Дж. Коен виділив наступні градації величини розміру ефекту:

- $0{,}00 \leq r^2 < 0{,}01$ – несуттєвий;
- $0{,}01 \leq r^2 < 0{,}09$ – малий;
- $0{,}09 \leq r^2 < 0{,}25$– середній;
- $0{,}25 \leq r^2 < 1{,}00$ – великий.

**Приклад 3.1.** У 20 досліджуваних виміряли рівень інтернет-залежності й тривожності особистості. Необхідно вияснити чи є статистично значущий зв'язок між цими параметрами.

Таблиця 3.4 –Таблиця первинних даних та проміжних розрахунків

| № | Інтернет-залежність (*X*) | Рівень тривожності (*Y*) | $\left(x_i - \overline{x}\right)$ | $\left(y_i - \overline{y}\right)$ | $\left(x_i - \overline{x}\right)^2$ | $\left(y_i - \overline{y}\right)^2$ | $\left(x_i - \overline{x}\right)\cdot\left(y_i - \overline{y}\right)$ |
|---|---|---|---|---|---|---|---|
| 1 | 84 | 20 | 32,1 | 3,5 | 1030,41 | 12,25 | 112,35 |
| 2 | 57 | 16 | 5,1 | -0,5 | 26,01 | 0,25 | -2,55 |
| 3 | 21 | 7 | -30,9 | -9,5 | 954,81 | 90,25 | 293,55 |
| 4 | 32 | 15 | -19,9 | -1,5 | 396,01 | 2,25 | 29,85 |
| 5 | 69 | 17 | 17,1 | 0,5 | 292,41 | 0,25 | 8,55 |

Продовження таблиці 3.4.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 6 | 37 | 14 | -14,9 | -2,5 | 222,01 | 6,25 | 37,25 |
| 7 | 49 | 17 | -2,9 | 0,5 | 8,41 | 0,25 | -1,45 |
| 8 | 47 | 16 | -4,9 | -0,5 | 24,01 | 0,25 | 2,45 |
| 9 | 82 | 29 | 30,1 | 12,5 | 906,01 | 156,25 | 376,25 |
| 10 | 65 | 22 | 13,1 | 5,5 | 171,61 | 30,25 | 72,05 |
| 11 | 23 | 11 | -28,9 | -5,5 | 835,21 | 30,25 | 158,95 |
| 12 | 48 | 14 | -3,9 | -2,5 | 15,21 | 6,25 | 9,75 |
| 13 | 77 | 25 | 25,1 | 8,5 | 630,01 | 72,25 | 213,35 |
| 14 | 39 | 7 | -12,9 | -9,5 | 166,41 | 90,25 | 122,55 |
| 15 | 45 | 16 | -6,9 | -0,5 | 47,61 | 0,25 | 3,45 |
| 16 | 68 | 19 | 16,1 | 2,5 | 259,21 | 6,25 | 40,25 |
| 17 | 68 | 21 | 16,1 | 4,5 | 259,21 | 20,25 | 72,45 |
| 18 | 27 | 10 | -24,9 | -6,5 | 620,01 | 42,25 | 161,85 |
| 19 | 41 | 16 | -10,9 | -0,5 | 118,81 | 0,25 | 5,45 |
| 20 | 59 | 18 | 7,1 | 1,5 | 50,41 | 2,25 | 10,65 |
| | $\bar{x}$ = 51,9 | $\bar{y}$ = 16,5 | Σ = 0 | Σ = 0 | Σ=7033,8 | Σ = 569 | Σ = 1727 |

Будемо перевіряти первинні дані на можливість використання параметричного коефіцієнта лінійної кореляції $r_{xy}$ Пірсона.

1. З метою перевірки даних у обох змінних на відповідність закону нормального розподілу використаємо критерій М.А. Плохінського.

Розподіл даних відповідає нормальному, якщо:

$$t_A = \frac{|A|}{m_A} \leq 3 \quad \text{та} \quad t_E = \frac{|E|}{m_E} \leq 3.$$

Для розрахунку показників асиметрії та ексцесу необхідно розрахувати середнє арифметичне значення та стандартне відхилення.

$\bar{x}$ = 51,9; $\bar{y}$ = 16,5.

Для кожного значення змінних $x_i$ та $y_i$ знаходимо різницю $\left(x_i - \bar{x}\right)$ та $\left(y_i - \bar{y}\right)$, результат записуємо у відповідні стовпчики. Суми відхилень від середнього значення для кожної змінної повинні бути рівні нулю (з точністю до похибки обчислень).

Кожне відхилення від середнього значення підносимо до квадрату і фіксуємо результат у відповідні стовпчики $\left(x_i - \bar{x}\right)^2$ та $\left(y_i - \bar{y}\right)^2$.

Суми квадратів відхилень необхідні для обчислення стандартних відхилень.

$$s_x = \sqrt{\frac{\Sigma\left(x_i - \bar{s}\right)^2}{n-1}} = \sqrt{\frac{7033,8}{20-1}} = 19,24;$$

$$s_y = \sqrt{\frac{\Sigma\left(y_i - \overline{y}\right)^2}{n-1}} = \sqrt{\frac{569}{20-1}} = 5,47 \cdot$$

Додатково розрахували:

$\Sigma \left(x_i - \overline{x}\right)^3$ = 8006,16;

$\Sigma \left(x_i - \overline{x}\right)^4$ = 4776778,43;

$\Sigma \left(y_i - \overline{y}\right)^3$ = 696;

$\Sigma \left(y_i - \overline{y}\right)^4$ = 50227,25.

Для визначення вибіркового значення асиметрії та ексцесу в обох вибірках дослідження застосуємо точні розрахункові формули.

$$A_x = \frac{n \cdot \Sigma\left(x_i - \overline{x}\right)^3}{(n-1)\cdot(n-2)\cdot s_x^3} = \frac{20 \cdot 8006,16}{(20-1)\cdot(20-2)\cdot 19,24^3} = 0,066.$$

$$E_x = \frac{n\cdot(n+1)\cdot\Sigma\left(x_i - \overline{x}\right)^4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot s_x^4} - \frac{3\cdot(n-1)^2}{(n-2)\cdot(n-3)} = \frac{20\cdot(20+1)\cdot 4776778,43}{(20-1)\cdot(20-2)\cdot(20-3)\cdot 19,24^4} -$$

$$- \frac{3\cdot(20-1)^2}{(20-2)\cdot(20-3)} = 2,518 - 3,539 = -1,021.$$

$$A_y = \frac{n \cdot \Sigma\left(x_i - \overline{y}\right)^3}{(n-1)\cdot(n-2)\cdot s_y^3} = \frac{20 \cdot 696}{(20-1)\cdot(20-2)\cdot 5,47^3} = 0,248.$$

$$E_y = \frac{n\cdot(n+1)\cdot\Sigma\left(x_i - \overline{y}\right)^4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot s_y^4} - \frac{3\cdot(n-1)^2}{(n-2)\cdot(n-3)} = \frac{20\cdot(20+1)\cdot 50227,25}{(20-1)\cdot(20-2)\cdot(20-3)\cdot 5,47^4} -$$

$$- \frac{3\cdot(20-1)^2}{(20-2)\cdot(20-3)} = 4,053 - 3,539 = 0,514.$$

Визначаємо помилки репрезентативності асиметрії та ексцесу для *n* = 20.

$$m_A = \sqrt{\frac{6}{n+3}} = \sqrt{\frac{6}{20+3}} = 0,511;$$

$$m_E = 2 \cdot \sqrt{\frac{6}{n+3}} = 2 \cdot \sqrt{\frac{6}{20+3}} = 1{,}022.$$

Згідно з критерієм М.А. Плохинського показники асиметрії й ексцесу свідчать про статистичну відмінність емпіричного розподілу від нормального в тому випадку, якщо вони перевищують за абсолютною величиною свої помилки репрезентативності в 3 рази.

$$t_{A_x} = \frac{|A_x|}{m_A} = \frac{|0{,}066|}{0{,}511} = 0{,}13; \; t_{E_x} = \frac{|E_x|}{m_E} = \frac{|-1{,}021|}{1{,}022} = 1.$$

$$t_{A_y} = \frac{|A_y|}{m_A} = \frac{|0{,}248|}{0{,}511} = 0{,}49; \; t_{E_y} = \frac{|E_y|}{m_E} = \frac{|0{,}514|}{1{,}022} = 0{,}5.$$

За результатами розрахунків показники асиметрії та ексцесу в обох змінних не перевищують утричі свої помилки репрезентативності, тому можемо зробити висновок, що дані мають нормальний розподіл. Відповідно, ми можемо використати параметричний коефіцієнт кореляції $r_{xy}$ Пірсона.

2. Встановлюємо прийнятний рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: інтернет-залежність й тривожність особистості не корелюють ($r_{xy} = 0$);

$H_1$: інтернет-залежність й тривожність особистості корелюють ($r_{xy} \neq 0$).

3. Розраховуємо емпіричне значення коефіцієнта кореляції $r_{xy}$ Пірсона:

$$r_{xy} = \frac{\Sigma\left(x_i - \bar{x}\right)\cdot\left(y_i - \bar{y}\right)}{(n-1)\cdot s_x \cdot s_y} = \frac{1727}{(20-1)\cdot 19{,}24 \cdot 5{,}47} = 0{,}863.$$

Отримане значення свідчить про пряму сильну кореляцію між змінними. Перевіримо її статистичну значущість.

4. За таблицею критичних значень коефіцієнта кореляції $r_{xy}$ Пірсона (додаток 1) для $df = n = 20$ та заданого $\alpha = 0{,}05$ знаходимо $r_{крит.} = 0{,}444$..

Оскільки $|r_{емп.}| > r_{крит.}$ (0,863 > 0,679), гіпотеза $H_0$ відхиляється.

5. Оцінку розміру ефекту здійснюємо за показником коефіцієнта детермінації $r^2$.

$$r^2 = {r_{xy}}^2 = 0{,}863^2 = 0{,}745.$$

Розмір ефекту відповідає великому рівню згідно інтерпретації Дж. Коена. Іншими словами спільна мінливість змінних інтернет-залежності та тривожності складає 74,5 %.

Висновок. Коефіцієнт кореляції Пірсона дозволив відхилити нульову гіпотезу для заданого набору даних ($r_{xy}$ = 0,863; $p$ < 0,05; $n$ = 20). Існує статистично значуща сильна позитивна кореляція між інтернет-залежністю й тривожністю особистості. Розміру ефекту відповідає великому рівню згідно інтерпретації Дж. Коена ($r^2$ = 0,74).

– відео-огляд процедури аналізу даних Прикладу 3.1. у статистичній програмі SPSS Statistics 23.0.

### 3.3 Рангові коефіцієнти кореляції

На практиці дослідники психологічних явищ та процесів мають справу переважно з неметричними вимірювальними шкалах. Якщо дані в дослідженні не відповідають закону нормального розподіл

у або представлені в порядковій шкалі, то для дослідження зв'язку між ознаками використовують коефіцієнти рангової кореляції $r_s$ Спірмена або $\tau$ Кендалла. Їх переваги – простота, широкі можливості та універсальність.

Принципових відмінностей між цими критеріями немає, але прийнято вважати, що коефіцієнт Кендалла є більш «змістовним», оскільки він більш повно й детально аналізує зв'язку між змінними, перебираючи всі можливі відповідності між парами значень. Коефіцієнт Спірмена більш точно враховує саме кількісну ступінь зв'язку між змінними. При розрахунках одних і тих самих даних $r_s > \tau$.

Коефіцієнти рангової кореляції можуть змінюватися від –1 до +1. Позитивний знак свідчить про прямий взаємозв'язок змінних, негативний знак – про зворотній.

#### 3.3.1 Коефіцієнт рангової кореляції $r_s$ Спірмена

Уперше завдання обчислення коефіцієнта кореляції між двома змінними, що представлені не своїми значеннями, а рангами цих значень було вирішено Френсісом Гальтоном наприкінці XIX століття. Однак такий підхід до обчислення значення коефіцієнта кореляції став популярним тільки після його використання Чарльзом Спірменом на початку XX століття. Остання обставина визначила, що коефіцієнт рангової кореляції носить ім'я Спірмена.

Методом рангової кореляції Спірмена можна визначити силу і напрямок зв'язку між двома ознаками чи профілями (ієрархіями) ознак, що представлені у вимірювальних шкалах порядку, або одна ознака – у шкалі порядку, а друга – у будь-якій кількісній шкалі (інтервальній або відносин).

Для обчислення коефіцієнта рангової кореляції $r_s$ Спірмена (інше позначення ρ) необхідно мати два ряди значень, які можна прорангувати. Такими рядами значень можуть бути:

1) дві ознаки однієї і тієї самої групи досліджуваних;

2) дві індивідуальні ієрархії ознак двох досліджуваних за одним і тим самим набором ознак;

3) дві групові ієрархії ознак;

4) індивідуальна та групова ієрархія ознак.

Емпіричне значення коефіцієнта рангової кореляції $r_s$ Спірмена визначається за формулою:

$$r_s = 1 - \frac{6 \cdot \Sigma d_i^2}{n \cdot (n^2 - 1)}, \text{ де}$$

$d$ – різниця рангів;
n – обсяг вибірки досліджуваних або кількість ознак.

У разі наявності зв'язаних рангів (однакових) однієї чи обох змінних коефіцієнт загальна формула рангової кореляції $r_s$ Спірмена обчислює дещо неточне значення. У такому разі необхідно здійснити поправку на однакові ранги. Скоригована формула має такий вигляд:

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d_i^2}{2 \cdot \sqrt{\Sigma x^2 \cdot \Sigma y^2}},$$

де $\Sigma x^2 = \frac{n^3 - n}{12} - T_x$; $\Sigma y^2 = \frac{n^3 - n}{12} - T_y$; $T_x = \frac{\Sigma(a^3 - a)}{12}$; $T_y = \frac{\Sigma(b^3 - b)}{12}$;

$d$ – різниця рангів;

$n$ – обсяг вибірки досліджуваних або кількість ознак.

$a$ – обсяг кожної групи однакових рангів у першому ранговому ряду X;

$b$ – обсяг кожної групи однакових рангів у другому ранговому ряду Y.

Якщо емпіричне значення $r_s$ дорівнює або більше критичного значення для заданого α-рівня, то $H_0$ відхиляється і формулюється висновок, що коефіцієнт кореляції статистично значуще відрізняється від нуля.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення коефіцієнта $r_s$. При інтерпретації необхідно орієнтуватися на табл. 3.2.

**Приклад 3.2.** Чи є статистичний взаємозв'язок показників значущості типів цінностей на рівні нормативні ідеалів у студентів хімічного та психологічного факультетів?

Таблиця 3.5 –Таблиця первинних даних та проміжних розрахунків

| № | Цінності | Хіміки (X) | Психологи (Y) | *d* | $d^2$ |
|---|---|---|---|---|---|
| 1 | Конформність | 8 | 8 | 0 | 0 |
| 2 | Традиції | 9 | 10 | -1 | 1 |
| 3 | Доброта | 7 | 4 | 3 | 9 |
| 4 | Універсалізм | 10 | 9 | 1 | 1 |
| 5 | Самостійність | 3 | 2 | 1 | 1 |
| 6 | Стимуляція | 4 | 6 | -2 | 4 |
| 7 | Гедонізм | 2 | 3 | -1 | 1 |
| 8 | Досягнення | 5 | 1 | 4 | 16 |
| 9 | Влада | 6 | 5 | 1 | 1 |
| 10 | Безпека | 1 | 7 | -6 | 36 |
| | | | | Σ = 0 | Σ = 70 |

Первинні дані за двома параметрами представлені в шкалі порядку. Ранги не повторюються, відповідно емпіричне значення коефіцієнта кореляції $r_s$ Спірмена будемо визначати за загальною формулою.

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: типи цінностей у студентів хімічного та психологічного факультетів не взаємозв'язані ($r_s = 0$);

$H_1$: типи цінностей у студентів хімічного та психологічного факультетів взаємозв'язані ($r_s \neq 0$).

2. Обчислюємо *d* – різницю рангів за формулою: $d_i = R_{Xi} - R_{Yi}$;

3. Визначаємо $d^2$ – квадрат різниці рангів.

4. Знаходимо суму $d^2$ за вибіркою досліджуваних.

5. Розраховуємо емпіричне значення коефіцієнта рангової кореляції Спірмена за формулою $r_s$:

$$r_s = 1 - \frac{6 \cdot \Sigma d_i^2}{n \cdot (n^2 - 1)} = 1 - \frac{6 \cdot 70}{10 \cdot \left(10^2 - 1\right)} = 0{,}576.$$

Отримане значення свідчить про пряму кореляцію середньої сили між змінними. Перевіримо її статистичну значущість.

6. За таблицею критичних значень коефіцієнта кореляції $r_s$ Спірмена (додаток 1) для $df = n = 10$ та заданого α = 0,05 знаходимо $r_{крит.} = 0{,}632$..

Оскільки $|r_{емп.}| < r_{крит.}$ (0,576 < 0,632), гіпотеза $H_0$ приймається.

7. Оцінюємо розмір ефекту.

Розмір стандартизованого ефекту згідно класифікації Дж. Коена – великий ($r_s = 0{,}576$).

<u>Висновок.</u> Відсутній статистично значущий зв'язок між типами цінностей на рівні нормативні ідеалів у студентів хімічного та психологі-

чного факультетів. Згідно статистики критерію $r_s$ Спірмена ($r_s = 0{,}576$; $p > 0{,}05$; $n = 10$) для відхилення нульової гіпотези немає достатніх підстав, проте виявлений великий розмір ефекту свідчить про доцільність реорганізації дослідження, збільшивши статистичну потужність.

– відео-огляд процедури аналізу даних Прикладу 3.2. у статистичній програмі SPSS Statistics 23.0.

**Приклад 3.3.** Чи можна стверджувати, що різні види пам'яті статистично значуще взаємопов'язані між собою?

Таблиця 3.6 – Таблиця первинних даних та проміжних розрахунків

| № | **Логічна пам'ять (X)** | **Механічна пам'ять (Y)** | $R_{Xi}$ | $R_{Yi}$ | ***d*** | $d^2$ |
|---|---|---|---|---|---|---|
| 1 | 8 | 46 | 3,5 | 14 | -10,5 | 110,25 |
| 2 | 5 | 34 | 1 | 2 | -1 | 1 |
| 3 | 13 | 39 | 11,5 | 7 | 4,5 | 20,25 |
| 4 | 10 | 32 | 6 | 1 | 5 | 25 |
| 5 | 11 | 45 | 7,5 | 12,5 | -5 | 25 |
| 6 | 8 | 36 | 3,5 | 4 | -0,5 | 0,25 |
| 7 | 9 | 40 | 5 | 8 | -3 | 9 |
| 8 | 13 | 38 | 11,5 | 6 | 5,5 | 30,25 |
| 9 | 13 | 41 | 11,5 | 9 | 2,5 | 6,25 |
| 10 | 7 | 42 | 2 | 10 | -8 | 64 |
| 11 | 13 | 44 | 11,5 | 11 | 0,5 | 0,25 |
| 12 | 11 | 37 | 7,5 | 5 | 2,5 | 6,25 |
| 13 | 15 | 45 | 14 | 12,5 | 1,5 | 2,25 |
| 14 | 12 | 35 | 9 | 3 | 6 | 36 |
| 15 | 17 | 47 | 15 | 15 | 0 | 0 |
| | | | $\Sigma_R = 120$ | $\Sigma_R = 120$ | $\Sigma = 0$ | $\Sigma = 336$ |

Підстав вважати, що первинні дані за обома змінними відповідають закону нормального розподілу немає. Вибірка досліджуваних недостатньо велика, щоби зробити якісний висновок при перевірці даних на нормальність розподілу. Отже, для виявлення взаємозв'язку між змінними будемо використовувати непараметричний коефіцієнт кореляції $r_s$ Спірмена.

1. Встановлюємо прийнятний рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: логічна та механічна пам'ять особистості не взаємозв'язані ($r_s = 0$);

$H_1$: логічна та механічна пам'ять особистості взаємозв'язані ($r_s \neq 0$).

1. Прорангуємо обидва показники від найменшого до найбільшого. Для перевірки правильності ранжування можна використовувати той факт, що $\Sigma_R = n \cdot (n+1)/2$. У нашому випадку сума рангів рівна 120 ($\Sigma_R = 120$), $n \cdot (n+1)/2 = 15 \cdot (15+1)/2 = 120$. Тотожність підтвердилась (120 =120).

2. Обчислюємо $d$ – різницю рангів за формулою: $d_i = R_{Xi} - R_{Yi}$;

3. Визначаємо $d^2$ – квадрат різниці рангів.

4. Знаходимо суму $d^2$ за вибіркою досліджуваних.

5. Будемо розраховувати емпіричне значення коефіцієнта рангової кореляції $r_s$ Спірмена за спеціальною формулою через існування зв'язаних рангів. Визначаємо поправки на однакові ранги.

$$T_x = \frac{\Sigma\left(a^3 - a\right)}{12} = \frac{(2^3-2)+(2^3-2)+(4^3-4)}{12} = 6;$$

$$T_y = \frac{\Sigma\left(b^3 - b\right)}{12} = \frac{\left(2^2-2\right)}{12} = 0{,}5;$$

$$\Sigma x^2 = \frac{n^3 - n}{12} - T_x = \frac{15^3 - 15}{12} - 6 = 274;$$

$$\Sigma y^2 = \frac{n^3 - n}{12} - T_y = \frac{15^3 - 15}{12} - 0{,}5 = 279{,}5;$$

6. Розраховуємо емпіричне значення коефіцієнта рангової кореляції $r_s$ Спірмена:

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d_i^2}{2 \cdot \sqrt{\Sigma x^2 \cdot \Sigma y^2}} = \frac{274 + 279{,}5 - 336}{2 \cdot \sqrt{274 \cdot 279{,}5}} = 0{,}393.$$

Отримане значення свідчить про пряму кореляцію помірної сили між змінними. Перевіримо її статистичну значущість.

8. За таблицею критичних значень коефіцієнта кореляції $r_s$ Спірмена (додаток 1) для $df = n = 15$ та заданого $\alpha = 0{,}05$ знаходимо $r_{крит.} = 0{,}514$..

Оскільки $|r_{емп.}| < r_{крит.}$ (0,393 < 0,514), гіпотеза $H_0$ приймається.

9. Оцінюємо розмір ефекту.

Розмір стандартизованого ефекту згідно класифікації Дж. Коена – середній ($r_s = 0{,}393$).

<u>Висновок.</u> За допомогою критерію $r_s$ Спірмена не змогли доказати існування статистично значущого зв'язку між логічною та механічною пам'яттю ($r_s = 0{,}393$; $p > 0{,}05$; $n = 15$), нульова гіпотеза приймається. Ви-

явлений середній розмір ефекту згідно інтерпретації Дж. Коена свідчить про доцільність збільшення статистичної потужності дослідження.

### 3.3.2 Коефіцієнт рангової кореляції *τ* Кендалла

У середині 1940-х рр. Моріс Кендалл запропонував альтернативний підхід до визначення кореляційного зв'язку між ранговими змінними.

Ідея кореляції М. Кендалла полягала в наступному. Уявімо, ми вишикували групу людей за зростом – від найвищого до найнижчого й дізналися вагу кожного з них. Тепер ми маємо два впорядковані ряди – для зросту і ваги. Якщо припустити, що чим більший зріст, тим більша вага, то зміна рангів зросту буде збігатися зі зміною рангів ваги. Тобто, першому рангу за зростом відповідатиме перший ранг за вагою, другому рангу за зростом – другий ранг за вагою й так далі. Однак ця послідовність може бути порушена (людина може бути високою, але худою, або низькорослою, але гладкою), і замість збігу рангів зросту і ваги ми отримаємо інверсію рангів (перестановку): наприклад, зріст зменшився, а вага збільшилася.

Якщо обидва рангові ряди повністю збігаються, то коефіцієнт кореляції дорівнює +1. Якщо один ряд змінюється протилежно в порівнянні до іншого, ми маємо суцільні інверсії, то коефіцієнт кореляції дорівнює –1.

Здебільшого в разі порівняння двох рядів рангів, у них будуть траплятися як збіги, так і інверсії. М. Кендалл запропонував формулу для обчислення коефіцієнта кореляції на основі кількості збігів та інверсій.

Емпіричне значення коефіцієнта рангової кореляції τ Кендалла визначається за формулою:

$$\tau = \frac{P - Q}{n \cdot (n-1)/2},$$

де $P$ – кількість збігів (погоджені пари);
$Q$ – кількість інверсій (непогоджені пари);
$n$ – обсяг вибірки досліджуваних або кількість ознак.

Є декілька модифікацій формули для обчислення коефіцієнта рангової кореляції τ Кендалла, у яких використовується тільки значення кількості збігів або інверсій або (як у нашому випадку), як збіги так і інверсії. Усі вони приводять до однакового результату.

У разі наявності зв'язаних (однакових) рангів однієї чи обох змінних коефіцієнт кореляції τ Кендалла обчислює дещо неточне значення. У такому разі необхідно здійснити поправку на однакові ранги. Формула набуває наступного вигляду:

$$\tau = \frac{P - Q}{\left(\sqrt{\frac{n \cdot (n-1)}{2} - T_x}\right) \cdot \left(\sqrt{\frac{n \cdot (n-1)}{2} - T_y}\right)},$$

де $T_x = \frac{\Sigma(a_x - 1)}{2}$; $T_y = \frac{\Sigma(b_y - 1)}{2}$

*P* – кількість збігів (погоджені пари);

*Q* – кількість інверсій (непогоджені пари);

*n* – обсяг вибірки досліджуваних або кількість ознак.

$a_x$ – обсяг кожної групи однакових рангів у першому ранговому ряду X;

$b_y$ – обсяг кожної групи однакових рангів у другому ранговому ряду Y.

Коефіцієнт кореляції $\tau$ Кендалла немає своєї таблиці для знаходження критичних значень. Оскільки статистика коефіцієнта кореляції відповідає розподілу Стьюдента, то пошук критичних значень здійснюється за таблицею критичних значень критерію *t*-Стьюдента. Для обчислення емпіричного значення використовується формула:

$$t = |\tau| \cdot \sqrt{\frac{n-2}{1-\tau^2}}\text{, де}$$

$\tau$ – емпіричне значення коефіцієнта кореляції $\tau$ Кендалла;

*n* – обсяг вибірки досліджуваних.

Якщо емпіричне значення *t* дорівнює або більше критичного значення для заданого α-рівня, то $H_0$ відхиляється та приймається $H_1$, формулюється висновок, що коефіцієнт кореляції статистично значуще відрізняється від нуля.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення коефіцієнта τ Кендалла. При інтерпретації необхідно орієнтуватися на методичні рекомендації запропоновані Джейкобом Коеном (див. таблицю 3.2.).

При інтерпретації коефіцієнта кореляції τ Кендалла можна врахувати, що для двох рангових рядів ймовірність збігів дорівнює (1 + τ)/2, а ймовірність інверсій дорівнює (1 – τ)/2. Наприклад, якщо взаємозв'язок оцінок жіночої привабливості двох студентів дорівнює τ = 0,6, то це означає 80 % збігів і 20 % інверсій. Іншими словами, думки студентів щодо жіночої привабливості співпадають вчетверо частіше. Інтерпретація коефіцієнтів кореляції Пірсона чи Спірмена має не такий явний характер. Отриманий результат необхідно спочатку звести до квадрату, а потім зробити висновок щодо частки дисперсії, яка може бути пояснена отриманим результатом.

**Приклад 3.4.** Необхідно встановити чи існує статистично значущий взаємозв'язок між відповідальністю особистості та професійною успішністю?

Таблиця 3.7 –Таблиця первинних даних та проміжних розрахунків

| № | **Відповідальність особистості ($X$)** | **Професійна успішність ($Y$)** | **Відповідальність особистості ($X_{ранг}$)** | **Професійна успішність ($Y_{ранг}$)** |
|---|---|---|---|---|
| 1 | 40 | 37 | 7 | 8 |
| 2 | 49 | 42 | 12 | 11 |
| 3 | 44 | 25 | 9 | 5 |
| 4 | 42 | 40 | 8 | 10 |
| 5 | 24 | 19 | 1 | 3 |
| 6 | 48 | 39 | 11 | 9 |
| 7 | 36 | 27 | 5 | 6 |
| 8 | 25 | 14 | 2 | 1 |
| 9 | 45 | 43 | 10 | 12 |
| 10 | 28 | 16 | 3 | 2 |
| 11 | 31 | 20 | 4 | 4 |
| 12 | 39 | 35 | 6 | 7 |
| | | | $\Sigma_R = 78$ | $\Sigma_R = 78$ |

Підстав вважати, що первинні дані за обома змінними представлені в інтервальній вимірювальній шкалі та відповідають закону нормального розподілу немає. Вибірка досліджуваних недостатньо велика, щоби зробити якісний висновок при перевірці даних у двох змінних на нормальність розподілу. Отже будемо вважати, що ми працюємо з порядковими шкалами, тому для виявлення зв'язку між змінними будемо використовувати непараметричний коефіцієнт кореляції τ Кендалла.

1. Встановлюємо прийнятний рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: відповідальність особистості та професійна успішність не взаємозв'язані ($\tau = 0$);

$H_1$: відповідальність особистості та професійна успішність взаємозв'язані ($\tau \neq 0$).

2. Прорангуємо обидва показники від найменшого до найбільшого. Для перевірки правильності ранжування можна використовувати той факт, що $\Sigma_R = n \cdot (n+1)/2$. У нашому випадку сума рангів рівна 78 ($\Sigma_R = 78$), $n \cdot (n+1)/2 = 12 \cdot (12+1)/2 = 78.$ Тотожність підтвердилась (78 =78).

3. Для обчислення коефіцієнта кореляції τ Кендалла необхідно впорядкувати другий стовпчик (відповідальність особистості) за зростанням рангів. Відповідно поміняються місцями ранги показника професійної успішності (див. табл. 3.8).

4. Підраховуємо кількість збігів та інверсій.

Підрахунок кількості збігів ($P$) здійснюється в такий спосіб. Беремо саме верхнє число третього стовпчика – 3. Рахуємо скільки всього чисел більших за 3 трапляються нижче в цьому ж стовпці. Їх – 9 (4, 5, 6, 7, 8, 9, 10, 11, 12). Число дев'ять записуємо навпроти 5 досліджуваного (див. № досліджуваного) у колонці *P.* Потім аналізуємо наступне число з третього стовпчика. Це – одиниця, більше неї нижче за стовпцем розташовані 10 чисел. Навпроти 8 досліджуваного записуємо 10. І так далі. $\Sigma_P = 57$.

Таблиця 3.8 –Таблиця проміжних розрахунків

| № | Відповідальність особистості ($X_{ранг}$) | Професійна успішність ($Y_{ранг}$) | *P* | *Q* |
|---|---|---|---|---|
| 5 | 1 | 3 | 9 | 2 |
| 8 | 2 | 1 | 10 | 0 |
| 10 | 3 | 2 | 9 | 0 |
| 11 | 4 | 4 | 8 | 0 |
| 7 | 5 | 6 | 6 | 1 |
| 12 | 6 | 7 | 5 | 1 |
| 1 | 7 | 8 | 4 | 1 |
| 4 | 8 | 10 | 2 | 2 |
| 3 | 9 | 5 | 3 | 0 |
| 9 | 10 | 12 | 0 | 2 |
| 6 | 11 | 9 | 1 | 0 |
| 2 | 12 | 11 | 0 | 0 |
| | | | $\Sigma_P = 57$ | $\Sigma_Q = 9$ |

Для визначення кількості інверсій знову беремо саме верхнє число третього стовпчика – 3. Підраховуємо скільки усього чисел менших за 3 трапляються нижче в цьому ж стовпці. Це числа 2 і 1, тобто їх два. Двійку записуємо навпроти 5 досліджуваного в колонці *Q.* Потім аналізуємо наступне число з третього стовпчика. Це – одиниця, менше неї нижче за стовпцем немає жодного числа. Навпроти досліджуваного під № 8 записуємо 0. І так далі. $\Sigma_Q = 9$.

Для перевірки правильності підрахунку кількості збігів та інверсій можна використати формулу:

$$\Sigma_P + \Sigma_Q = \frac{(n-1)\cdot n}{2};$$

У нашому випадку $\Sigma_P + \Sigma_Q = 66$; $\frac{(n-1)\cdot n}{2} = \frac{(12-1)\cdot 12}{2} = 66$; Тотожність підтвердилась (66 = 66).

5. Розраховуємо емпіричне значення за формулою критерія τ Кендалла.

$$\tau = \frac{P-Q}{n\cdot(n-1)/2} = \frac{57-9}{12\cdot(12-1)/2} = 0{,}727\,.$$

Отримане значення свідчить про сильну пряму кореляцію між змінними. Перевіримо її статистичну значущість.

6. Пошук критичних значень здійснюється за таблицею критичних значень критерію $t$-Стьюдента. Для обчислення емпіричного значення використовується формула:

$$t = |\tau| \cdot \sqrt{\frac{n-2}{1-\tau^2}} = |0,727| \cdot \sqrt{\frac{12-2}{1-0,727^2}} = 3,348.$$

7. За таблицею критичних значень критерію $t$-Стьюдента (додаток 3) для $df = n - 2 = 12 - 2 = 10$ та заданого $\alpha = 0,05$ знаходимо $t_{крит.} = 2,23$.

Оскільки $t_{емп.} > t_{крит.}$ (3,348 > 3,17), гіпотеза $H_0$ відхиляється.

8. Оцінюємо розмір ефекту.

Розмір стандартизованого ефекту згідно Дж. Коена – великий ($\tau = 0,727$).

Висновок. Коефіцієнт кореляції Кендалла дозволив відхилити нульову гіпотезу ($\tau = 0,727$; $p < 0,05$; $n = 12$). Існує статистично значуща сильна позитивна кореляція між відповідальністю особистості та професійною успішністю. Розміру ефекту відповідає великому рівню згідно інтерпретації Дж. Коена.

– відео-огляд процедури аналізу даних Прикладу 3.4. у статистичній програмі SPSS Statistics 23.0.

## 3.4 Коефіцієнти асоціативної залежності номінальних даних

Дуже часто в психології при виконанні різних видів досліджень фіксуються й аналізуються ознаки, що не мають безпосередньої кількісної міри. Такі дані представлені в номінальній (номінативній, найменувань) вимірювальній шкалі. У межах цієї шкали об’єкти класифікуються, а класи здебільшого позначаються номерами, які нічого не говорить про властивості об’єкта, крім того, що він належить до певної групи.

В такому випадку для аналізу кількісної залежності між змінними використовуються коефіцієнти розроблені на основі тесту $\chi^2$-Пірсона. Значення 0 відповідає повній незалежності змінних, а 1 – їх максимальній залежності. В цих коефіцієнтах знак не інтерпретується. Він не вказує на напрямок залежності, тому що залежить від кодувань (позначень) градацій значень ознак.

### 3.4.1 Коефіцієнт контингенції Пірсона $\varphi$

*Коефіцієнт контингенції Карла Пірсона* $\varphi$ дозволяє визначити асоціацію між якісними характеристиками об'єктів. Використовується для пошуку залежності між змінними, що виміряні в дихотомічній (може мати тільки два значення) номінальній шкалі. Значення цього коефіцієнта може змінюватися в межах від –1 до +1.

Є декілька способів розрахунку коефіцієнта контингенції Пірсона $\varphi$. Найбільш простий за таблицею зв'язаності ознак, яка має такий вигляд (табл. 3.9.):

Таблиця 3.9 – Таблиця зв'язаності ознак розміром 2×2

| **Ознаки** | | **Змінна X** | | **Усього** |
|---|---|---|---|---|
| | | **1** | **0** | |
| **Змінна Y** | **1** | a | b | a+b |
| | **0** | c | d | c+d |
| **Усього** | | a+c | b+d | n |

У такому випадку розрахункове значення коефіцієнта контингенції обчислюється за формулою:

$$\varphi = \frac{a \cdot d - b \cdot c}{\sqrt{(a+c)\cdot(b+d)\cdot(a+b)\cdot(c+d)}}$$

Коефіцієнт контингенції немає своєї таблиці з критичними значеннями. Статистична значущість взаємозв'язку оцінюється за допомогою критерію $\chi^2$-Пірсона.

$$\chi^2 = \varphi^2 \cdot n,$$

де $\varphi$ – емпіричне значення коефіцієнта контингенції;
$n$ – обсяг вибірки досліджуваних.

Для прийняття рішення про рівень статистичної значущості необхідно емпіричне (розрахункове) значення критерію $\chi^2$-Пірсона порівняти з критичним (табличним) значенням. Критичне значення визначається залежно від числа ступенів свободи та заданого α-рівня. Зважаючи на те, що для таблиці розміром 2×2 число ступенів свободи завжди однакове та дорівнює 1 ($df$ = 1), то ці значення завжди сталі (додаток 2): $\chi^2_{0,05.}$ = 3,841; $\chi^2_{0,01.}$= 6,635; $\chi^2_{0,001.}$= 10,829.

Якщо $\chi^2_{емп.} \geq \chi^2_{крит.}$, то залежність між ознаками статистично значуща, тобто ознаки змінюються узгоджено.

Якщо $\chi^2_{емп.} < \chi^2_{крит.}$, то статистично значущої залежності між ознаками немає.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення коефіцієнта контингенції $\varphi$. При інтерпретації необхідно орієнтуватися на рекомендації запропоновані Дж. Коеном (див. таблицю 3.2.).

**Приклад 3.5.** Необхідно перевірити, чи є залежність між статтю та депресивністю особистості? Результати дослідження представлені в таблиці 3.10.

Ознака Х (стать): 1 – жінка; 2 – чоловік.

Ознака У (депресія): 1 – є ознаки; 2 – відсутні ознаки.

Таблиця 3.10 – Таблиця первинних даних

| №п/п | Стать | Депресія | №п/п | Стать | Депресія |
|---|---|---|---|---|---|
| 1 | 2 | 2 | 21 | 1 | 1 |
| 2 | 1 | 1 | 22 | 2 | 2 |
| 3 | 2 | 1 | 23 | 1 | 1 |
| 4 | 1 | 1 | 24 | 2 | 2 |
| 5 | 1 | 2 | 25 | 2 | 2 |
| 6 | 1 | 2 | 26 | 2 | 2 |
| 7 | 2 | 1 | 27 | 2 | 2 |
| 8 | 1 | 1 | 28 | 1 | 2 |
| 9 | 1 | 2 | 29 | 2 | 2 |
| 10 | 1 | 1 | 30 | 1 | 2 |
| 11 | 1 | 2 | 31 | 2 | 2 |
| 12 | 1 | 2 | 32 | 1 | 1 |
| 13 | 1 | 2 | 33 | 1 | 1 |
| 14 | 2 | 2 | 34 | 2 | 2 |
| 15 | 2 | 1 | 35 | 2 | 2 |
| 16 | 1 | 2 | 36 | 1 | 2 |
| 17 | 1 | 2 | 37 | 2 | 2 |
| 18 | 2 | 1 | 38 | 2 | 2 |
| 19 | 1 | 2 | 39 | 2 | 2 |
| 20 | 1 | 1 | 40 | 1 | 2 |

Оскільки обидві ознаки представлені в дихотомічній номінальній шкалі, то для перевірки наявності залежності між ними можна використати коефіцієнт контингенції Пірсона $\varphi$.

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: між статтю та депресивністю особистості відсутня залежність ($\varphi = 0$);

$H_1$: між статтю та депресивністю особистості є залежність ($\varphi \neq 0$).

1. За даними таблиці 3.10. сформуємо таблицю зв'язаності ознак і підрахуємо абсолютні частоти всіх можливих поєднань значень (табл. 3.11.).

Таблиця 3.11 –Таблиця зв’язаності ознак

| Ознаки | | Депресія | | Усього |
|---|---|---|---|---|
| | | Є ознаки | Відсутні ознаки | |
| Стать | Жінки | 9 | 13 | 22 |
| | Чоловіки | 4 | 14 | 18 |
| Усього | | 13 | 27 | 40 |

2. У даній таблиці: $a = 9$, $b = 13$, $c = 4$, $d = 14$.

Підставами ці значення у формулу коефіцієнта контингенції $\varphi$:

$$\varphi = \frac{9 \cdot 14 - 13 \cdot 4}{\sqrt{(9+4) \cdot (13+14) \cdot (9+13) \cdot (4+14)}} = 0{,}198.$$

Отримане значення свідчить про залежність дуже слабкої сили між ознаками. Перевіримо її статистичну значущість.

4. Для оцінки статистичної значущості залежності необхідно розрахувати емпіричне значення критерію $\chi^2$-Пірсона.

$$\chi^2 = \varphi^2 \cdot n = (0{,}198)^2 \cdot 40 = 1{,}57.$$

За таблицею критичних значень критерію $\chi^2$-Пірсона (додаток 2) для $df = 1$ та заданого $\alpha = 0{,}05$ знаходимо $\chi^2_{крит.} = 3{,}841$.

Оскільки $\chi^2_{емп.} < \chi^2_{крит.}$ (1,57 < 3,841), гіпотеза $H_0$ приймається.

5. Оцінюємо розмір ефекту.

Показником розміру ефекту є безпосередньо саме емпіричне значення коефіцієнта *φ.* Згідно Джейкоба Коена 0,198 відповідає малому рівню.

Для таблиці розміром 2×2 можна визначити ще один показник величини розміру ефекту – відношення шансів (odds ratio – OR).

$$OR = \frac{a \cdot d}{b \cdot c} = \frac{9 \cdot 14}{13 \cdot 4} = 2{,}42.$$

Іншими словами, ймовірність мати ознаки депресії у жінок в 2,4 рази вища ніж у чоловіків.

<u>Висновок</u>. Емпіричні дані не дозволяють відхилити нульову гіпотезу коефіцієнтом контингенції ($\varphi = 0{,}198$; $p > 0{,}05$; $n = 40$). Тобто, відсутня статистично значуща асоціативна залежність між статтю та депресивністю особистості. Розмір ефекту відповідає малому рівню згідно Дж. Коена. Ймовірність мати ознаки депресії у жінок в 2,4 рази вища ніж у чоловіків.

– відео-огляд процедури аналізу даних Прикладу 3.5. у статистичній програмі SPSS Statistics 23.0.

### 3.4.2 Коефіцієнт асоціації Юла *Q*

*Коефіцієнт асоціації Юла Q* також використовується для пошуку залежності між змінними, що виміряні в номінальній дихотомічній шкалі. Може набувати значень в інтервалі від −1 до +1. Визначається за формулою:

$$Q = \frac{ad - bc}{ad + bc},$$

де $a$, $b$, $c$, $d$ – дані, що визначаються згідно з таблицею зв'язаності ознак розміром 2×2 (див. табл. 3.9.).

Коефіцієнт асоціації як і коефіцієнт контингенції немає своєї таблиці з критичними значеннями. Статистична значущість залежності оцінюється за допомогою критерію $\chi^2$-Пірсона.

$$\chi^2 = Q^2 \cdot n,$$

де $Q$ – емпіричне значення коефіцієнта асоціації;
$n$ – обсяг вибірки досліджуваних.

Для прийняття рішення про рівень статистичної значущості необхідно емпіричне (розрахункове) значення критерію $\chi^2$ порівняти з критичним (табличним) значенням. Критичне значення визначається залежно від числа ступенів свободи та заданого α-рівня. Зважаючи на те, що для таблиці розміром 2×2 число ступенів свободи завжди дорівнює 1 ($df$ = 1), то ці значення завжди сталі (додаток 2): $\chi^2_{0,05.}$ = 3,841; $\chi^2_{0,01.}$= 6,635; $\chi^2_{0,001.}$= 10,829.

Якщо $\chi^2_{емп.} \geq \chi^2_{крит.}$, то залежність між ознаками статистично значуща, тобто ознаки змінюються узгоджено.

Якщо $\chi^2_{емп.} < \chi^2_{крит.}$, то статистично значущої залежності між ознаками немає.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення коефіцієнта контингенції *Q*. При інтерпретації необхідно орієнтуватися на рекомендації запропоновані Дж. Коеном (див. таблицю 3.2.).

Незважаючи, що коефіцієнт асоціації Юла та коефіцієнт контингенції Пірсона використовуються для вирішення однакових завдань, кожен із них має свої закономірності. Коефіцієнт контингенції вимірює двобічну залежність, а коефіцієнт асоціації Юла – однобічну. Як наслідок, коефіцієнт асоціації завжди більший коефіцієнта контингенції за умови, що вони розраховані за одними й тими даними. Коефіцієнт контингенції дає більш обережну оцінку ступеня тісноти залежності.

Якщо хоча б одне з чотирьох значень у таблиці зв'язаності 2×2 відсутнє, то значення коефіцієнта асоціації Юла буде дорівнювати одиниці, а це перебільшена оцінка ступеня тісноти залежності між ознаками. У даному випадку необхідно використовувати коефіцієнт контингенції.

**Приклад 3.6.** Необхідно оцінити, чи впливають заняття студентів у науково-дослідній лабораторії на впевненість під час захисту кваліфікаційних робіт. Дослідник у своєму розпорядженні має дані результатів дослідження 320 студентів, з яких 240 займаються в науково-дослідній лабораторії.

Таблиця 3.12 – Таблиця зв'язаності ознак

<table>
<tr><th colspan="2" rowspan="2">Ознаки</th><th colspan="2">Впевненість під час захисту кваліфікаційних робіт</th><th rowspan="2">Усього</th></tr>
<tr><th>Впевнені</th><th>Невпевнені</th></tr>
<tr><td rowspan="2">Групи студентів</td><td>Займаються в науково-дослідній лабораторії</td><td>163</td><td>77</td><td>240</td></tr>
<tr><td>Не займаються в науково-дослідній лабораторії</td><td>46</td><td>34</td><td>80</td></tr>
<tr><td colspan="2">Усього</td><td>209</td><td>111</td><td>320</td></tr>
</table>

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: впевненість під час захисту кваліфікаційних робіт не залежить від занять у науково-дослідній лабораторії ($Q = 0$);

$H_1$: впевненість під час захисту кваліфікаційних робіт залежить від занять у науково-дослідній лабораторії ($Q \neq 0$).

2. У даній таблиці: $a = 163$, $b = 77$, $c = 46$, $d = 34$.

Підставами ці значення у формулу коефіцієнта асоціації $Q$.

$$Q = \frac{ad - bc}{ad + bc} = \frac{163 \cdot 34 - 77 \cdot 46}{163 \cdot 34 + 77 \cdot 34} = 0{,}215.$$

Отримане значення свідчить про залежність слабкої сили між ознаками. Перевіримо її статистичну значущість.

2. Для оцінки статистичної значущості залежності необхідно знайти емпіричне значення критерію $\chi^2$-Пірсона.

$$\chi^2 = Q^2 \cdot n = 0{,}215^2 \cdot 320 = 14{,}792.$$

За таблицею критичних значень критерію $\chi^2$-Пірсона (додаток 2) для $df$ = 1 та заданого α = 0,05 знаходимо $\chi^2_{крит.}$ = 3,841.

Оскільки $\chi^2_{емп.} > \chi^2_{крит.}$ (14,792 > 3,841), гіпотеза $H_0$ відхиляється.

5. Оцінюємо розмір ефекту.

Показником розміру ефекту є безпосередньо саме емпіричне значення коефіцієнта $Q$. Згідно Дж. Коена 0,215 відповідає малому рівню.

Висновок. Коефіцієнт асоціації $Q$ дозволив відхилити нульову гіпотезу для заданого набору даних ($Q$ = 0,215; $p$ < 0,05; $n$ = 320). Існує статистично значуща слабкої сили асоціація між заняттями студентами в науково-дослідній лабораторії та впевненістю під час захисту кваліфікаційних робіт. Проте малий рівень розміру ефекту вимагає задуматися про практичну цінність виявленої залежності.

### 3.4.3 Коефіцієнти взаємної зв'язаності ознак Пірсона *C*, Чупрова *K* та Крамера *V*

Коефіцієнти взаємної зв'язаності ознак Пірсона $C$, Чупрова $K$ та Крамера $V$ застосовуються для оцінки залежності в ситуаціях, коли кожна змінна представлена в номінальній шкалі та хоча б одна з них має більш ніж дві градацій. Тобто таблиця зв'язаності ознак повинна мати розмір не 2×2, як у коефіцієнтів асоціації та контингенції, а будь-який інший.

Ці критерії ґрунтуються на безпосередньому розрахунку критерію $\chi^2$-Пірсона. Необхідно зауважити, що із збільшенням значення $\chi^2$ сила залежності коефіцієнтів взаємної зв'язаності зростає.

Розрахункова формула $\chi^2$-Пірсона має такий вигляд:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(f_{емп.} - f_{теор.}\right)^2}{f_{теор.}},$$

де $k$ – кількість класових інтервалів (градацій) досліджуваної ознаки;

$f_{емп.}$ та $f_{теор.}$ – відповідно емпіричні та теоретичні (очікувані) частоти, що відповідають визначеним градаціям змінної.

Для розрахунку теоретичної (очікуваної) частоти для кожної комірки таблиці необхідно знати, що:

$$f_{ij} = \frac{f_i \cdot f_j}{n},$$

де $f_i$ – сума частот у всіх комірках $i$-рядка;

$f_j$ – сума частот у всіх комірках $j$- стовпчика;

$n$ – сума частот у всій таблиці зв'язаності ознак.

Значення коефіцієнтів Пірсона $C$, Чупрова $K$ та Крамера $V$ може бути в межах від 0 до +1, але кожен критерій має свої особливості.

Значення коефіцієнта Пірсона $C$ у разі повної відсутності асоціації між змінними дорівнює нулю, але при абсолютній залежності він не дорівнює одиниці, а лише прагне до неї в ситуації збільшення числа градацій ознаки. Максимум значення $C$ досягається коли $k = m$ ($k$, $m$ – кількість рядків і стовпчиків у таблиці зв'язаності ознак), але величина максимального значення змінюється зі зміною числа категорій: за умови $k$ ($m$) = 3 значення $C$ не може бути більшим 0,8, за умови $k$ ($m$) = 5 максимальне значення $C$ дорівнює 0,89 і так далі. Відповідно, це призводить до виникнення труднощів порівняння таблиць різного розміру. З цієї причини коефіцієнт Пірсона $C$ рекомендується використовувати тільки в разі однакової кількості рядків і стовпчиків таблиці зв'язаності ознак ($k = m$).

Для виправлення цього недоліка коефіцієнта Пірсона Олександр Чупров запропонував інший коефіцієнт (Чупрова $K$), але і його максимальне значення досягає одиниці лише в разі рівності кількості градацій досліджуваних ознак ($k = m$), а в ситуації $k \neq m$ не може досягати одиниці.

З перелічених критеріїв лише коефіцієнт кореляції Крамера $V$ може досягати одиниці незалежно від розміру таблиці зв'язаності. Для квадратних таблиць зв'язаності ($k = m$) коефіцієнти Чупрова $K$ та Крамера $V$ співпадають, в усіх інших випадках $V < K$. При цьому коефіцієнт Чупрова $K$ завжди менший коефіцієнта Пірсона $C$.

Коефіцієнти взаємної зв'язаності Пірсона $C$, Чупрова $K$ та Крамера $V$ визначається за формулами:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}\,;$$

$$K = \sqrt{\frac{\chi^2}{n \cdot \sqrt{(k-1)\cdot(m-1)}}}\,;$$

$$V = \sqrt{\frac{\chi^2}{n \cdot (c-1)}}\,,$$

де $\chi^2$ – емпіричне значення критерію $\chi^2$-Пірсона (у всіх формулах);

$n$ – обсяг вибірки досліджуваних (у всіх формулах);

$k$, $m$ – кількість рядків і стовпчиків у таблиці зв'язаності ознак (Чупрова $K$);

$c$ – найменша кількість градацій із двох змінних (Крамера $V$).

Обмеження у використанні цих коефіцієнтів відповідають обмеженням критерію $\chi^2$-Пірсона, а саме: обсяг вибірки повинен бути $n \geq 30$, не більше 20 % теоретичних (очікуваних) частот можуть бути менші 5. Водночас не повинно бути теоретичних частот менших 1.

Таблиць із критичними значеннями для коефіцієнтів зв'язаності немає. Тому для прийняття статистичного рішення необхідно:

1. Вибрати рівень значущості та сформулювати нульову і альтернативну гіпотези.

2. Обчислити емпіричне значення критерію $\chi^2$-Пірсона.

3. Розрахувати значення коефіцієнта взаємної зв'язаності.

4. Порівняти емпіричне значення критерію $\chi^2$ з критичним значенням (додаток 2) для відповідного числа ступенів свободи. $df = (k-1)\cdot(m-1)$, де $k$, $m$ – кількість рядків і стовпчиків таблиці зв'язаності ознак.

5. Якщо $\chi^2_{емп.} \geq \chi^2_{крит.}$, то залежність між ознаками статистично значуща, тобто ознаки змінюються узгоджено.

Якщо $\chi^2_{емп.} < \chi^2_{крит.}$, то формулюється висновок про відсутність статистично значущої залежності.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення коефіцієнтів взаємної зв'язаності ознак. При інтерпретації необхідно орієнтуватися на рекомендації запропоновані Дж. Коеном (див. таблицю 3.2.).

**Приклад 3.7.** Необхідно встановити чи існує статистично значуща залежність між рівнями розвитку загальних пізнавальних здібностей (ЗПЗ) та професійної успішності?

Таблиця 3.13 – Таблиця зв'язаності ознак

| **Параметри** | | **Рівень професійної успішності** | | |
|---|---|---|---|---|
| | | **Високий** | **Середній** | **Низький** |
| **Рівень ЗПЗ** | **Високий** | 20 | 15 | 6 |
| | **Середній** | 12 | 13 | 8 |
| | **Низький** | 5 | 10 | 13 |

Первинні дані представлені в номінальній шкалі. Таблиця зв'язаності ознак має розмір 3×3. Необхідно визначити асоціативну залежність. У такому випадку можна застосувати коефіцієнти взаємної зв'язаності Пірсона $C$, Чупрова $K$ та Крамера $V$.

1. Встановлюємо рівень ймовірності помилки першого роду $\alpha = 0{,}05$ і формулюємо нульову та альтернативну гіпотези:

$H_0$: між рівнями розвитку загальних пізнавальних здібностей і професійної успішністю відсутня асоціативна залежність ($C, K, V = 0$);

$H_1$: між рівнями розвитку загальних пізнавальних здібностей і професійної успішністю існує асоціативна залежність ($C, K, V \neq 0$).

2. Для розрахунку емпіричного значення $\chi^2$ необхідно спочатку визначити теоретичні (очікувані) частоти для кожної комірки таблиці.

$$f_{ij} = \frac{f_i \cdot f_j}{n},$$

де $f_i$ – сума частот у всіх комірках $i$-рядка;

$f_j$ – сума частот у всіх комірках $j$- стовпчика;
$n$ – сума частот у всій таблиці зв'язаності ознак.

Таблиця 3.14 – Значення очікуваних частот

| ЗПЗ | Рівень професійної успішності | | | Σ |
|---|---|---|---|---|
| | Високий | Середній | Низький | |
| **Високі** | $f_{11}=\frac{41\cdot 37}{102}=14,87$ | $f_{12}=\frac{41\cdot 38}{102}=15,28$ | $f_{13}=\frac{41\cdot 27}{102}=10,85$ | 41 |
| **Середні** | $f_{21}=\frac{33\cdot 37}{102}=11,97$ | $f_{22}=\frac{33\cdot 38}{102}=12,29$ | $f_{23}=\frac{33\cdot 27}{102}=8,74$ | 33 |
| **Низькі** | $f_{31}=\frac{28\cdot 37}{102}=10,16$ | $f_{32}=\frac{28\cdot 38}{102}=10,43$ | $f_{33}=\frac{28\cdot 27}{102}=7,41$ | 28 |
| **Σ** | 37 | 38 | 27 | 102 |

3. Підставимо значення емпіричних і очікуваних частот у формулу для розрахунку $\chi^2$.

$$\chi^2=\sum_{i=1}^{k}\frac{\left(f_{емп.}-f_{теор.}\right)^2}{f_{теор.}}=\frac{(20-14,87)^2}{14,87}+\frac{(15-15,28)^2}{15,28}+\frac{(6-10,85)^2}{10,85}+$$
$$+\frac{(12-11,97)^2}{11,97}+\frac{(13-12,29)^2}{12,29}+\frac{(8-8,74)^2}{8,74}+\frac{(5-10,16)^2}{10,16}+\frac{(10-10,43)^2}{10,43}+$$
$$+\frac{(13-7,41)^2}{7,41}=10,86.$$

4. Обчислюємо значення коефіцієнтів взаємної зв'язаності за відповідними формулами:

$$C=\sqrt{\frac{\chi^2}{\chi^2+n}}=\sqrt{\frac{10,86}{10,86+102}}=0,31;$$

$$K=\sqrt{\frac{\chi^2}{n\cdot\sqrt{(k-1)\cdot(m-1)}}}=\sqrt{\frac{10,86}{102\cdot\sqrt{(3-1)\cdot(3-1)}}}=0,23;$$

$$V=\sqrt{\frac{\chi^2}{n\cdot(c-1)}}=\sqrt{\frac{10,86}{102\cdot(3\text{-}1)}}=0,23.$$

Отримані значення свідчать про помірний та низький рівень сили залежності між ознаками. Перевіримо їхню статистичну значущість.

5. За таблицею критичних значень критерію $\chi^2$-Пірсона (додаток 2) для $df=(k-1)\cdot(m-1)=(3-1)\cdot(3-1)=4$ та заданого $\alpha = 0,05$ визначаємо $\chi^2_{крит.}$ = 9,488..

Оскільки $\chi^2_{емп.} > \chi^2_{крит.}$ (10,86 > 9,488), гіпотеза $H_0$ відхиляється.

5. Оцінюємо розмір ефекту.

Показником розміру ефекту є безпосередньо саме емпіричне значення коефіцієнтів взаємної зв'язаності ознак. Згідно інтерпретації Дж. Коена 0,231 відповідає малому рівню, а 0,311 – середньому.

Висновок. Коефіцієнти взаємної зв'язаності ознак дозволили відхилити нульову гіпотезу ($C$ = 0,31; $K$ = 0,23; $V$ = 0,23; $p < 0{,}05$ $df$ = 4; $n$ = 102). Тобто, існує статистично значуща асоціативна залежність між розвитком загальних пізнавальних здібностей і професійною успішністю. Розміри ефекту відповідають середньому та малому рівню в залежності від статистичного критерію.

– відео-огляд процедури аналізу даних Прикладу 3.7. у статистичній програмі SPSS Statistics 23.0.

### 3.5 Бісеріальні коефіцієнти кореляції

Бісеріальні коефіцієнти кореляції оцінюють залежність двох ознак, одна з яких представлена в номінальній дихотомічній шкалі, а друга в шкалі порядку або в будь-якій кількісній шкалі (інтервальній або відношень).

Значення бісеріальних коефіцієнтів кореляції змінюються в діапазоні від –1 до +1. Водночас знак для інтерпретації результатів немає значення, тому що він залежить від послідовності в якій будуть представлені градації номінальної ознаки.

#### 3.5.1 Точково-бісеріальний коефіцієнт кореляції $r_{pb}$

Точково-бісеріальний коефіцієнт кореляції $r_{pb}$ використовується у випадку, коли одна змінна виміряна у дихотомічній номінальній шкалі, а друга – в шкалі інтервалів чи відношень. Емпіричне значення визначається за формулою:

$$r_{pb} = \frac{\overline{y_1} - \overline{y_0}}{s_y} \cdot \sqrt{\frac{n_1 \cdot n_0}{n \cdot (n-1)}}, \text{ де}$$

$\overline{y_1}$ – середнє арифметичне змінної У для об'єктів, що мають за Х одиницю;

$\overline{y_0}$ – середнє арифметичне змінної У для об'єктів, що мають за Х нуль;

$s_y$ – стандартне відхилення Y;

$n_1$ – кількість об’єктів за X, які мають одиницю;

$n_0$ – кількість об’єктів за X, які мають нуль;

$n$ – обсяг вибірки досліджуваних.

Точково-бісеріальний коефіцієнт кореляції немає своєї таблиці для знаходження критичних значень. Оскільки статистика коефіцієнта кореляції має розподіл Стьюдента, то пошук критичних значень здійснюється за таблицею критичних значень критерію $t$-Стьюдента. Для обчислення емпіричного значення використовується формула:

$$t = \left| r_{pb} \right| \cdot \sqrt{\frac{n-2}{1-r_{pb}^2}}$$ , де

$r_{pb}$ – емпіричне значення точково-бісеріального коефіцієнта кореляції;

$n$ – обсяг вибірки досліджуваних.

Якщо емпіричне значення $t$ дорівнює або більше критичного значення для заданого α-рівня, то $H_0$ відхиляється і формулюється висновок, що коефіцієнт кореляції статистично значуще відрізняється від нуля.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення точково-бісеріального коефіцієнта.

**Приклад 3.8.** Необхідно перевірити, чи є взаємозв’язок між статтю людини й самооцінкою? Параметр «самооцінка» в популяції відповідає закону нормального розподілу, жіноча стать досліджуваного закодована – 0, чоловіча – 1. Результати дослідження представлені в таблиці 3.15.

Таблиця 3.15 –Таблиця первинних даних та проміжних розрахунків

| № | **Стать (X)** | **Самооцінка (Y)** | $y_i - \overline{y}$ | $\left(y_i - \overline{y}\right)^2$ |
|---|---|---|---|---|
| 1 | 1 | 95,7 | 15,9 | 252,81 |
| 2 | 1 | 97 | 17,2 | 295,84 |
| 3 | 1 | 83 | 3,2 | 10,24 |
| 4 | 1 | 98,6 | 18,8 | 353,44 |
| 5 | 1 | 88,6 | 8,8 | 77,44 |
| 6 | 1 | 83 | 3,2 | 10,24 |
| 7 | 1 | 71 | -8,8 | 77,44 |
| 8 | 0 | 90 | 10,2 | 104,04 |
| 9 | 0 | 70 | -9,8 | 96,04 |
| 10 | 0 | 70 | -9,8 | 96,04 |
| 11 | 0 | 70 | -9,8 | 96,04 |
| 12 | 0 | 68,6 | -11,2 | 125,44 |
| 13 | 0 | 70 | -9,8 | 96,04 |
| 14 | 0 | 73 | -6,8 | 46,24 |
| 15 | 0 | 68,5 | -11,3 | 127,69 |
| | | $\overline{y}$ = 79,8 | Σ = 0 | Σ = 1865,02 |

Змінна X представлена в дихотомічній номінальній шкалі, а змінна Y виміряна в шкалі інтервалів, а дані відповідають закону нормального розподілу згідно умови завдання. В такому випадку для знаходження залежності між змінними необхідно використати точково-бісеріальний коефіцієнт кореляції.

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: самооцінка не залежить від статі людини ($r_{pb} = 0$);

$H_1$: самооцінка залежить від статі людини ($r_{pb} \neq 0$).

1. Визначаємо середнє арифметичне значення самооцінки для дівчат та хлопців:

$$\bar{y}_1 = \frac{95{,}7+97+83+98{,}6+88{,}6+83+71}{7} = 88{,}1;$$

$$\bar{y}_0 = \frac{90+70+70+70+68{,}6+70+73+68{,}5}{8} = 72{,}5.$$

2. Обчислюємо стандартне відхилення для Y:

$$s = \sqrt{\frac{\Sigma\left(y_i - \bar{y}\right)^2}{n-1}} = \sqrt{\frac{1865{,}02}{15-1}} = 11{,}54.$$

3. Розраховуємо емпіричне значення точково-бісеріального коефіцієнту кореляції за формулою $r_{pb}$:

$$r_{pb} = \frac{\bar{y}_1 - \bar{y}_0}{s_y} \cdot \sqrt{\frac{n_1 \cdot n_0}{n \cdot (n-1)}} = \frac{88{,}1-72{,}5}{11{,}54} \cdot \sqrt{\frac{7 \cdot 8}{15 \cdot (15-1)}} = 0{,}699.$$

Отримане значення свідчить про кореляцію середньої сили між змінними. Перевіримо її статистичну значущість.

4. Пошук критичних значень здійснюється за таблицею критичних значень критерію $t$-Стьюдента. Для обчислення емпіричного значення використовується формула:

$$t = \left|r_{pb}\right| \cdot \sqrt{\frac{n-2}{1-r_{pb}^2}} = \left|0{,}699\right| \cdot \sqrt{\frac{15-2}{1-0{,}699^2}} = 3{,}53.$$

6. За таблицею критичних значень критерію $t$-Стьюдента (додаток 3) для $df = n - 2 = 15 - 2 = 13$ та заданого $\alpha = 0{,}05$ знаходимо $t_{крит.} = 2{,}16$.

Оскільки $t_{емп.} > t_{крит.}$ (3,53 > 2,16), гіпотеза $H_0$ відхиляється.

7. Оцінюємо розмір ефекту.

Згідно Дж. Коена значення $r_{pb} = 0{,}699$ відповідає великому розміру ефекту.

Висновок. Застосування точково-бісеріального коефіцієнта кореляції дає підстави стверджувати, що існує статистично значуща залежність середньої сили між самооцінкою та статтю людини ($r_{pb} = 0{,}699$; $p < 0{,}05$; $n = 15$). Розмір стандартизованого ефекту – великий.

– відео-огляд процедури аналізу даних Прикладу 3.8. у статистичній програмі SPSS Statistics 23.0.

### 3.5.2 Рангово-бісеріальний коефіцієнт кореляції $r_{rb}$

Рангово-бісеріальний коефіцієнт кореляції застосовується у випадку, коли одна змінна виміряна в дихотомічній номінальній шкалі, а друга – шкалі порядку. Емпіричне значення визначається за формулою:

$$r_{rb} = \frac{(\overline{y_1} - \overline{y_0}) \cdot 2}{n},$$

де $\overline{y_1}$ – середній ранг змінної У для об'єктів, що мають за Х одиницю;

$\overline{y_0}$ – середній ранг змінної У для об'єктів, що мають за Х нуль;

$n$ – обсяг вибірки досліджуваних.

Рангово-бісеріальний коефіцієнт кореляції немає своєї таблиці для знаходження критичних значень. Оскільки статистика коефіцієнта кореляції має розподіл Стьюдента, то пошук критичних значень здійснюється за таблицею критичних значень критерію $t$-Стьюдента. Для обчислення емпіричного значення використовується формула:

$$t = |r_{rb}| \cdot \sqrt{\frac{n-2}{1-r_{rb}^2}},$$

де $r_{rb}$ – емпіричне значення рангово-бісеріального коефіцієнта кореляції;

$n$ – обсяг вибірки досліджуваних.

Якщо емпіричне значення $t$ дорівнює або більше критичного значення для заданого α-рівня, то $H_0$ відхиляється і формулюється висновок, що коефіцієнт кореляції статистично значуще відрізняється від нуля.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення рангово-бісеріального коефіцієнта. При інтерпретації необхідно орієнтуватися на рекомендації запропоновані Дж. Коеном (див. таблицю 3.2.).

**Приклад 3.9.** Необхідно перевірити, чи є зв’язок між статтю досліджуваного і комунікативними здібностями? Результати дослідження представлені в таблиці 3.16. Дівчата – 0; Хлопці – 1.

Таблиця 3.16 –Таблиця первинних даних

| № | Стать (X) | Ранг вербальних здібностей (Y) | № | Стать (X) | Ранг вербальних здібностей (Y) |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 9 | 1 | 4 |
| 2 | 0 | 10 | 10 | 1 | 3 |
| 3 | 1 | 6 | 11 | 1 | 5 |
| 4 | 1 | 9 | 12 | 0 | 11 |
| 5 | 0 | 15 | 13 | 1 | 12 |
| 6 | 1 | 7 | 14 | 1 | 2 |
| 7 | 0 | 8 | 15 | 0 | 14 |
| 8 | 0 | 13 | | | |

Змінна X представлена в дихотомічній номінальній шкалі, а змінна Y в шкалі порядку. Для того, щоб визначити зв’язок між змінними необхідно використати рангово-бісеріальний коефіцієнт кореляції.

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: комунікативні здібності особистості не залежать від статі ($r_{rb} = 0$);

$H_1$: комунікативні здібності особистості залежать від статі ($r_{rb} \neq 0$).

2. Визначаємо середній ранг комунікативних здібностей для дівчат та хлопців:

$$\bar{y}_1 = \frac{1+6+9+7+4+3+5+12+2}{9} = 5{,}44\,;$$

$$\bar{y}_0 = \frac{10+15+8+13+11+14}{6} = 11{,}83.$$

3. Розраховуємо емпіричне значення рангово-бісеріального коефіцієнту кореляції за формулою $r_{rb}$:

$$r_{rb} = \frac{(\overline{y_1} - \overline{y_0}) \cdot 2}{n} = \frac{5{,}44 - 11{,}83}{15} = -0{,}852.$$

Отримане значення свідчить про сильну кореляцію між змінними (знак у бісеріальних коефіцієнтах кореляції не інтерпретується). Перевіримо її статистичну значущість.

4. Пошук критичних значень здійснюється за таблицею критичних значень критерію t-Стьюдента. Для обчислення емпіричного значення використовується формула:

$$t = |r_{rb}| \cdot \sqrt{\frac{n-2}{1-r_{rb}^2}} = |-0,852| \cdot \sqrt{\frac{15-2}{1-(-0,852)^2}} = 5,869.$$

5. За таблицею критичних значень критерію $t$-Стьюдента (додаток 3) для $df = n - 2 = 15 - 2 = 13$ та заданого $\alpha = 0,05$ знаходимо $t_{крит.} = 2,16$.

Оскільки $t_{емп.} > t_{крит.}$ (5,869 > 2,16), гіпотеза $H_0$ відхиляється.

7. Оцінюємо розмір ефекту.

Згідно Дж. Коена значення 0,852 відповідає великому розміру ефекту.

Висновок. Рангово-бісеріальний коефіцієнт кореляції дозволили відхилити нульову гіпотезу ($r_{rb}$ = -0,852; p < 0,05 $n$ = 15). Тобто, існує статистично значуща залежність між комунікативними здібностями особистості та статтю. Розмір стандартизованого ефекту – великий.

**Завдання на самостійну підготовку**

1. Охарактеризуйте поняття функціонального та кореляційного зв’язку між змінними.
2. Вкажіть причини та помилки, які можуть виникати під час інтерпретації коефіцієнтів кореляції.
3. За яких умов для встановлення зв’язку між змінними використовується тетрахоричний коефіцієнт кореляції?
4. Характеристика та особливості використання коефіцієнтів множинної та часткової кореляції.
5. В яких ситуаціях використовується кореляційне відношення η?
6. Сутність методу регресійного аналізу даних
7. Як графічно можна зобразити зв’язок між двома змінними? Кореляційні плеяди.

## ЛЕКЦІЯ 4
## ПАРАМЕТРИЧНІ МЕТОДИ СТАТИСТИЧНОГО ПОРІВНЯННЯ ВИБІРОК ДОСЛІДЖУВАНИХ

### План лекції

4.1. Теоретичні засади та сфера застосування *t*-критерію Стьюдента
4.2. Критерій *t*-Стьюдента для незалежних вибірок
4.3. Критерій *t*-Стьюдента для залежних вибірок
4.4. Критерій *t*-Стьюдента для однієї вибірки

**Час проведення**: 2 учбові години.

### Література


1. Гланц С. Медико-биологическая статистика. Пер. с англ. – М., Практика, 1998. – 459 с.

2. Гржибовский А.М., Унгуряну Т.Н. Анализ биомедицинских данных с использованием пакета статистических программ SPSS: учебное пособие. – Архангельск: Изд-во Северного государственного медицинского университета, 2017. – 293 с.

3. Климчук В.О. Математичні методи у психології. Навчальний посібник. – К.: Освіта України, 2009. – 288 с.

4. Кричевец А. Н., Корнеев А. А., Рассказова Е. И. Основы статистики для психологов. – М.: Акрополь, 2019. – 286 с.

5. Телейко А.Б. Чорней Р.К. Математико-статистичні методи в соціології та психології: навч. посібник. – Київ: МАУП, 2007. – 418 с.

6. Харькова О.А., Соловьев А.Г. Статистические методы и математическое моделирование: учебное пособие. – Архангельск: Изд-во Северного государственного медицинского университета, 2017. – 164 с.


### 4.1 Теоретичні засади та сфера застосування *t*-критерію Стьюдента

*t*-критерій Стьюдента – загальна назва статистичних критеріїв, що ґрунтуються на порівнянні даних з розподілом Стьюдента. Найчастіше застосовується з метою перевірки рівності середніх значень у двох вибірках.

Критерій розроблений англійським вченим Вільямом Госсетом (1876-1937). Він працюючи в Дубліні на пивоварні Guinness, згідно умов контракту, не мав права публікувати результати своїх досліджень. І щоб уникнути розголосу й обійти заборону В. Госсет друкував свої роботи під псевдонімом Student, звідси і назва критерію.

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

Критерій *t*-Стьюдента використовується в трьох випадках:

1) порівняння середніх значень двох незалежних вибірок (*t*-критерій для незалежних вибірок);

2) порівняння середніх значень двох залежних вибірок (*t*-критерій для залежних вибірок);

3) порівняння середнього значення однієї вибірки із заданою величиною (*t*-критерій для однієї вибірки).

*Обмеження до застосування.*

Загальна вимога для всіх видів критеріїв Стьюдента – нормальність розподілу даних у обох вибірках. Якщо дані не відповідають закону нормального розподілу застосовувати критерій не коректно. Відповідно первинні дані повинні бути виміряні в метричній шкалі (інтервалів чи відношень).

Обмеження, яке стосується виключно *t*-критерію Стьюдента для незалежних вибірок – гомогенність дисперсій вибірок порівняння. Якщо дисперсії гетерогенні необхідно обрати *t*-критерій Уелша.

Обмеження щодо *t*-критерію Стьюдента для залежних вибірок – наявність статистично значущого прямого кореляційного зв’язку між залежними вибірками.

Щодо обсягу вибірок, формальних обмежень немає. Однак дуже складно зробити якісний висновок про відповідність емпіричного розподілу нормальному виду на малих вибірках (менше 30 спостережень), тому дослідники в таких ситуаціях досить часто не здійснюють перевірку на нормальність даних, а відразу застосовують непараметричні критерії (Манна-Уітні, Вілкоксона тощо).

*Загальний алгоритм застосування t-критерію Стьюдента.*

1. *Перевірка даних у обох вибірках досліджуваних на відповідність закону нормального розподілу.*

Без застосування спеціалізованих статистичних пакетів обробки даних кількісна оцінка нормальності розподілу здійснюється на основі розрахунку коефіцієнтів асиметрії, ексцесу та помилок цих параметрів (див. лекцію №2, навчальне питання 4: «Перевірка даних на відповідність закону нормального розподілу»).

*2А. Перевірка рівності (гомогенності) дисперсій для незалежних вибірок.*

Використовується *критерій F-Фішера для перевірки гомогенності дисперсій*. Емпіричне значення визначається за формулою:

$$F = \frac{s_1^2}{s_2^2}; \ df_{чис.} = n_1 - 1; \ df_{знам.} = n_2 - 1,$$

де $s_1^2$ – більша дисперсія із двох вибірок;

$s_2^2$ – менша дисперсія із двох вибірок;

$n_1$, $n_2$ – обсяг досліджуваних у двох вибірках.

Критичні значення ($F_{крит.}$) визначаємо за таблицею критичних значень *F*-критерію Фішера для перевірки ненаправлених альтернатив (додаток 4).

Якщо $F_{емп.} \geq F_{крит.}$ для відповідно числа ступенів свободи, то приймаємо рішення про підтвердження гіпотези про відмінність дисперсій на заданому $\alpha$-рівні, та відповідно, слід застосовувати не критерій Стьюдента в стандартній формі, а *t-критерій Уелша* (модифікований критерій Стьюдента для незалежних вибірок із гетерогенними дисперсіями).

*2Б. Перевірка наявності кореляційного зв'язку для залежних вибірок дослідження.*

Дуже часто залежність вибірок визначається дизайном дослідження, хоча формальним критерієм є наявність кореляційного зв'язку.

Використовується лінійний коефіцієнт кореляції $r_{xy}$ Пірсона, який визначається за формулою:

$$r_{xy} = \frac{\Sigma\left(x_i - \overline{x}_1\right)\cdot\left(x_i - \overline{x}_2\right)}{(n-1)\cdot s_1 \cdot s_2},$$

де $x_i$ – первинні значення;

$\overline{x}_1$, $\overline{x}_2$ – середнє арифметичне змінних $X_1$ та $X_2$;

$s_1$, $s_2$ – стандартне відхилення змінних $X_1$ та $X_2$;

$n$ – обсяг вибірки досліджуваних.

Отримане емпіричне значення коефіцієнта кореляції $r_{xy}$ Пірсона необхідно порівняти із критичним значенням. Кількість ступенів свободи дорівнює обсягу вибірки досліджуваних ($df = n$).

2. *Власне обчислення t-критерію Стьюдента.*

Обчислення критерію залежить від:

- вибірки залежні чи незалежні;
- дисперсії гомогенні чи гетерогенні;
- однаковий чи різний обсяг вибірок.

*4. Порівняння отриманого емпіричного значення t-критерію із критичним табличним значенням для формулювання статистичного висновку.*

За таблицею критичних значень критерію $t$-Стьюдента (додаток 3) для заданого α-рівня та визначеного число ступенів свободи ($df$), знаходимо $t_{крит.}$:

- якщо $|t_{емп.}| < t_{крит.}$ – приймаємо гіпотезу $H_0$. Відмінності між вибірками статистично незначущі.

- якщо $|t_{емп.}| \geq t_{крит.}$ – приймаємо гіпотезу $H_1$. Відмінності між вибірками статистично значущі.

*5. Оцінка розміру ефекту.*

Для оцінки розміру ефекту критерію $t$-Стьюдента використовується індекс $d$-Коена та його модифікації:

1) у випадку незалежності вибірок, рівності обсягів вибірок та гомогенності дисперсій:

$$d = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{s_{pooled.}}} = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{\frac{s_1^2 + s_2^2}{2}}},$$

де $\overline{x}_1$, $\overline{x}_2$ – середнє арифметичне значення в першій та другій вибірках;

$s_1, s_2$ – стандартне відхилення в першій та другій вибірках.

2) в ситуації незалежності вибірок, коли у них різна кількість досліджуваних у групах або/та малі обсяги вибірок для уникнення переоцінки розміру ефекту у формулу розрахунку $d$-Коена вноситься поправка $g$-Хеджесса:

$$g = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{\frac{(n_1 - 1)\cdot s_1^2 + (n_2 - 1)\cdot s_2^2}{n_1 + n_2 - 2}}},$$

де $\overline{x}_1$, $\overline{x}_2$ – середнє арифметичне значення в першій та другій вибірках;

$s_1, s_2$ – стандартне відхилення в першій та другій вибірках;

$n_1$, $n_2$ – обсяг досліджуваних у першій та другій вибірках.

3) за умови гетерогенності дисперсій у незалежних вибірках до індексу $d$-Коена вноситься поправка $\Delta s$-Гласса:

$$\Delta s = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{s_{контрол.}^2}},$$

де $\overline{x}_1$, $\overline{x}_2$ – середнє арифметичне значення в першій та другій вибірках;

$s^2_{контрол.}$ – стандартне відхилення в контрольній групі.

4) індекс розміру ефекту для залежних вибірок має позначення $d_z$-Коена:

$$d_z = \frac{|\bar{x}_d|}{s_d},$$

де $|\bar{x}_d|$ – середнє арифметичне різниці пар значень;

$s_d$ – стандартне відхилення різниці пар значень.

5) для критерія $t$-Стьюдента для однієї вибірки використовується індекс $d_{os}$-Коена:

$$d_{os} = \frac{|\bar{x} - A|}{s},$$

де $\bar{x}$ – середнє арифметичне значення;

$A$ – величина, з якою середнє арифметичне порівнюється;

$s$ – стандартне відхилення.

На даний момент не існує узгоджених стандартів при інтерпретації величини розміру ефекту. Дж. Коеном в якості методичних рекомендацій були запропоновані наступні градації значення $d$ та його модифікацій:

$0{,}00 \le d < 0{,}20$ – несуттєвий;

$0{,}20 \le d < 0{,}50$ – малий;

$0{,}50 \le d < 0{,}80$ – середній;

$0{,}80 \le d$ – великий.

Ці діапазони необхідно сприймати тільки в якості орієнтирів, так як для кожного окремого дослідження повинен застосовуватися свій підхід до інтерпретації. При оцінці розміру ефекту необхідно враховувати теоретичні, економічні, етичні та практичні аспекти.

## 4.2 Критерій *t*-Стьюдента для незалежних вибірок

*Обчислення t-критерію Стьюдента для незалежних вибірок (гомогенні дисперсії)*

Критерій призначений для порівняння середніх значень ознак в двох незалежних групах.

Формальним критерієм незалежності вибірок є відсутність кореляції між ними. З погляду змісту – незалежними є ті вибірки, між якими відсутні жодні зв'язки. Прикладом незалежних вибірок можуть бути рандомізовані (сформовані випадковим чином) контрольна та експериментальна група, дві професійні групи тощо.

Якщо обсяги вибірок однакові, то емпіричне значення $t$-критерію Стьюдента обчислюється за формулою:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2 + s_2^2}{n}}}; \; df = 2n - 2.$$

Якщо обсяги вибірок різні:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{(n_1 - 1)\cdot s_1^2 + (n_2 - 1)\cdot s_2^2}{n_1 + n_2 - 2}\cdot\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}; \; df = n_1 + n_2 - 2,$$

де $\bar{x}_1$, $\bar{x}_2$ – середнє арифметичне значення в першій та другій вибірках;

$s_1, s_2$ – стандартне відхилення в першій та другій вибірках;

$n_1$, $n_2$ – обсяг досліджуваних у першій та другій вибірках;

$n$ – загальний обсяг досліджуваних.

Якщо дисперсії гетерогенні, то застосовується $t$-критерій Уелша для незалежних вибірок:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}; \; df = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1 - 1) + \left(s_2^2/n_2\right)^2/(n_2 - 1)},$$

де $\bar{x}_1$, $\bar{x}_2$ – середнє арифметичне значення в першій та другій вибірках;

$s_1, s_2$ – стандартне відхилення в першій та другій вибірках;

$n_1$, $n_2$ – обсяг досліджуваних у першій та другій вибірках;

$n$ – загальний обсяг досліджуваних.

**Приклад 4.1**. Психолог фіксував час прийняття рішення (в секундах) у стресових умовах. Одна група ($X_1$) була сформована із 15 осіб, які регулярно займаються спортом, а інша ($X_2$) – із 13 осіб, які не займаються спортом. Чи відрізняється швидкість прийняття рішення в групах з різним відношенням до спорту?

У таблиці представлені результати тестування двох незалежних груп досліджуваних. Первинні дані представлені в шкалі відношень. З метою знаходження відмінностей між групами є передумови для використання параметричного $t$-критерій Стьюдента для незалежних вибірок. Перевіримо первинні дані на обмеження, щодо застосування статистичного критерію.

Таблиця 4.1 –Таблиця первинних даних та проміжних розрахунків

| № | **Група №1 ($X_1$)** | **Група №2 ($X_2$)** | $\left(x_i-\bar{x}_1\right)$ | $\left(x_i-\bar{x}_2\right)$ | $\left(x_i-\bar{x}_1\right)^2$ | $\left(x_i-\bar{x}_2\right)^2$ | $\left(x_i-\bar{x}_1\right)^3$ | $\left(x_i-\bar{x}_2\right)^3$ | $\left(x_i-\bar{x}_1\right)^4$ | $\left(x_i-\bar{x}_2\right)^4$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 4 | 1,73 | -2,15 | 2,99 | 4,62 | 5,18 | -9,94 | 8,96 | 21,37 |
| 2 | 5 | 6 | -0,27 | -0,15 | 0,07 | 0,02 | -0,02 | 0,00 | 0,01 | 0,00 |
| 3 | 8 | 9 | 2,73 | 2,85 | 7,45 | 8,12 | 20,35 | 23,15 | 55,55 | 65,98 |
| 4 | 6 | 8 | 0,73 | 1,85 | 0,53 | 3,42 | 0,39 | 6,33 | 0,28 | 11,71 |
| 5 | 4 | 4 | -1,27 | -2,15 | 1,61 | 4,62 | -2,05 | -9,94 | 2,60 | 21,37 |
| 6 | 5 | 6 | -0,27 | -0,15 | 0,07 | 0,02 | -0,02 | 0,00 | 0,01 | 0,00 |
| 7 | 7 | 6 | 1,73 | -0,15 | 2,99 | 0,02 | 5,18 | 0,00 | 8,96 | 0,00 |
| 8 | 5 | 5 | -0,27 | -1,15 | 0,07 | 1,32 | -0,02 | -1,52 | 0,01 | 1,75 |
| 9 | 4 | 7 | -1,27 | 0,85 | 1,61 | 0,72 | -2,05 | 0,61 | 2,60 | 0,52 |
| 10 | 5 | 4 | -0,27 | -2,15 | 0,07 | 4,62 | -0,02 | -9,94 | 0,01 | 21,37 |
| 11 | 4 | 7 | -1,27 | 0,85 | 1,61 | 0,72 | -2,05 | 0,61 | 2,60 | 0,52 |
| 12 | 5 | 9 | -0,27 | 2,85 | 0,07 | 8,12 | -0,02 | 23,15 | 0,01 | 65,98 |
| 13 | 5 | 5 | -0,27 | -1,15 | 0,07 | 1,32 | -0,02 | -1,52 | 0,01 | 1,75 |
| 14 | 6 | | 0,73 | | 0,53 | | 0,39 | | 0,28 | |
| 15 | 3 | | -2,27 | | 5,15 | | -11,70 | | 26,55 | |
| $\bar{x}_1$ = 5,27; $\bar{x}_2$ = 6,15. | | Σ = | | | 24,93 | 37,69 | 13,52 | 20,99 | 108,42 | 212,31 |

1. У даному прикладі, для перевірки даних у обох вибірках дослідження на відповідність закону нормального розподілу використаємо критерій Є.І. Пустильника.

Розподіл вважається нормальним, якщо показники асиметрії ($A$) й ексцесу ($E$) менші своїх критичних значень ($A_{крит.}$ і $E_{крит.}$).

Спочатку розрахуємо показники асиметрії та ексцесу в обох вибірках дослідження. Для цього необхідно визначити середнє арифметичне значення та стандартне відхилення.

$$\bar{x}_1 = 5{,}27;\ \bar{x}_2 = 6{,}15.$$

$$s_1 = \sqrt{\frac{\Sigma\left(x_i-\bar{x}_1\right)^2}{n-1}} = \sqrt{\frac{24{,}93}{15-1}} = 1{,}33;$$

$$s_2 = \sqrt{\frac{\Sigma\left(x_i-\bar{x}_2\right)^2}{n-1}} = \sqrt{\frac{37{,}69}{13-1}} = 1{,}77.$$

Для визначення асиметрії та ексцесу в обох вибірках дослідження використаємо точні розрахункові формули:

$$A_1 = \frac{n \cdot \Sigma\left(x_i - \overline{x}_1\right)^3}{(n-1)\cdot(n-2)\cdot s_1^3} = \frac{15 \cdot 13,52}{(15-1)\cdot(15-2)\cdot 1,33^3} = 0,474.$$

$$E_1 = \frac{n\cdot(n+1)\cdot\Sigma\left(x_i - x_1\right)^4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot s_1^4} - \frac{3\cdot(n-1)^2}{(n-2)\cdot(n-3)} = \frac{15\cdot(15+1)\cdot 108,42}{(15-1)\cdot(15-2)\cdot(15-3)\cdot 1,33^4} -$$

$$- \frac{3\cdot(15-1)^2}{(15-2)\cdot(15-3)} = 3,808 - 3,769 = 0,039.$$

$$A_2 = \frac{n \cdot \Sigma\left(x_i - \overline{x}_2\right)^3}{(n-1)\cdot(n-2)\cdot s_2^3} = \frac{13 \cdot 20,99}{(13-1)\cdot(13-2)\cdot 1,77^3} = 0,37.$$

$$E_2 = \frac{n\cdot(n+1)\cdot\Sigma\left(x_i - \overline{x}_2\right)^4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot s_2^4} - \frac{3\cdot(n-1)^2}{(n-2)\cdot(n-3)} = \frac{13\cdot(13+1)\cdot 212,31}{(13-1)\cdot(13-2)\cdot(13-3)\cdot 1,77^4} -$$

$$- \frac{3\cdot(13-1)^2}{(13-2)\cdot(13-3)} = 2,982 - 3,927 = \text{-}0,945.$$

Розрахуємо критичні значення для показників асиметрії та ексцесу для обох вибірок дослідження:

$$A_{крит.1} = 3\cdot\sqrt{\frac{6\cdot(n-1)}{(n+1)(n+3)}} = 3\cdot\sqrt{\frac{6\cdot(15-1)}{(15+1)\cdot(15+3)}} = 3\cdot 0,54 = 1,62\,;$$

$$E_{крит.1} = 5\cdot\sqrt{\frac{24\cdot n\cdot(n-2)\cdot(n-3)}{(n+1)^2\cdot(n+3)\cdot(n+5)}} = 5\cdot\sqrt{\frac{24\cdot 15\cdot(15-2)\cdot(15-3)}{(15+1)^2\cdot(15+3)\cdot(15+5)}} = 5\cdot 0,78 = 3,9.$$

$$A_{крит.2} = 3\cdot\sqrt{\frac{6\cdot(n-1)}{(n+1)(n+3)}} = 3\cdot\sqrt{\frac{6\cdot(13-1)}{(13+1)\cdot(13+3)}} = 3\cdot 0,57 = 1,71\,;$$

$$E_{крит.2} = 5\cdot\sqrt{\frac{24\cdot n\cdot(n-2)\cdot(n-3)}{(n+1)^2\cdot(n+3)\cdot(n+5)}} = 5\cdot\sqrt{\frac{24\cdot 13\cdot(13-2)\cdot(13-3)}{(13+1)^2\cdot(13+3)\cdot(13+5)}} = 5\cdot 0,78 = 3,9.$$

Згідно з критерієм Є.І. Пустильника дані відповідають закону нормального розподілу, якщо:

$|A| \le A_{крит.}$ та $|E| \le E_{крит.}$

У нашому випадку:

$|0,474| < 1,62$ та $|0,039| < 3,9$ – для першої вибірки;

$|0,37| < 1,71$ та $|-0,945| < 3,9$ – для другої вибірки;

Можемо зробити висновок, що розподіл у обох вибірках дослідження відповідає закону нормального розподілу. Відповідно, ми можемо використати параметричний критерій $t$-Стьюдента для незалежних вибірок.

2. Здійснюємо перевірку рівності (гомогенності) дисперсій для незалежних вибірок.

Використовуємо критерій $F$-Фішера для перевірки гомогенності дисперсій. Визначаємо емпіричне значення:

$$F = \frac{s_1^2}{s_2^2} = \frac{1,77^2}{1,33^2} = 1,77.$$

$s_1^2$ – більша дисперсія із двох вибірок;

$s_2^2$ – менша дисперсія із двох вибірок.

Обчислюємо кількість ступенів свободи:

$df_{чис.} = n_1 - 1 = 13 - 1 = 12$;

$df_{знам.} = n_2 - 1 = 15 - 1 = 14$.

Визначаємо $F_{0,05}$, за таблицею критичних значень $F$-критерію Фішера для перевірки ненаправлених альтернатив (додаток 4).

$F_{0,05} = 3,05$.

Оскільки $F_{емп.} < F_{крит.}$ (1,77 < 3,05), гіпотеза $H_0$ підтверджуються. Дисперсії вибірок не відрізняються, тобто гомогенні ($F_{(12;\,14)} = 1,77$; $p > 0,05$). Ми можемо застосувати критерій $t$-Стьюдента для незалежних вибірок

3. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0,05$;

$H_0$: швидкість прийняття рішення у стресових умовах в групах з різним відношенням до спорту не відрізняється ($\mu_1 = \mu_2$);

$H_1$: швидкість прийняття рішення у стресових умовах в групах з різним відношенням до спорту відрізняється ($\mu_1 \neq \mu_2$).

4. Розраховуємо емпіричне значення $t$-критерію Стьюдента для незалежних вибірок.

У нашому випадку обсяги двох вибірок різні, тому емпіричне значення $t$-критерію розраховується за уточненою формулою:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)\cdot s_1^2 + (n_2-1)\cdot s_2^2}{n_1+n_2-2}\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{5,27-6,15}{\sqrt{\frac{(15-1)\cdot 1,33^2 + (13-1)\cdot 1,77^2}{15+13-2}\cdot\left(\frac{1}{15}+\frac{1}{13}\right)}} =$$

$$= \frac{-0,88}{\sqrt{2,4\cdot 0,14}} = -1,51.$$

5. За таблицею критичних значень критерію $t$-Стьюдента (додаток 3) для $df = n_1 + n_2 - 2 = 15 + 13 - 2 = 26$ та заданого $\alpha = 0{,}05$ знаходимо $t_{крит.} = 2{,}06$ (двобічна критична область).

Оскільки $|t_{емп.}| > t_{крит.}$ (1,51 < 2,06), гіпотеза $H_0$ приймається.

5. Оцінка розміру ефекту здійснюється за індексом $g$-Хеджесса, тому що у групах різна кількість досліджуваних:

$$g = \frac{\left|\bar{x}_1 - \bar{x}_2\right|}{\sqrt{\dfrac{(n_1-1)\cdot s_1^2 + (n_2-1)\cdot s_2^2}{n_1+n_2-2}}} = \frac{\left|5{,}27-6{,}15\right|}{\sqrt{\dfrac{(15-1)\cdot 1{,}33^2 + (13-1)\cdot 1{,}77^2}{15+13-2}}} = \frac{0{,}88}{\sqrt{2{,}4}} = 0{,}57.$$

Згідно інтерпретації Дж. Коена розмір ефекту відповідає середньому рівню.

Висновок. Емпіричні дані не дозволяють відхилити нульову гіпотезу критерієм $t$-Стьюдента для незалежних вибірок ($t_{(26)} = -1{,}51$; $p > 0{,}05$; двобічна критична область). Відсутня статистично значуща відмінність в середніх значеннях швидкості прийняття рішення у стресових умовах в групах спортсменів ($\bar{x} = 5{,}27$; $s = 1{,}33$) та осіб, які не займаються спортом ($\bar{x} = 6{,}15$; $s = 1{,}77$). При цьому виявлений середній розмір стандартизованого ефекту ($g = 0{,}57$) свідчить про доцільність збільшення статистичної потужності дослідження.

– відео-огляд процедури аналізу даних Прикладу 4.1. у статистичній програмі SPSS Statistics 23.0.

### 4.3 Критерій *t*-Стьюдента для залежних вибірок

$t$-критерій Стьюдента для залежних вибірок (Т-критерій для парних вибірок) дозволяє перевірити гіпотезу, що середні значення двох залежних вибірок відрізняються один від одного.

Як правило, залежність вибірок визначається дизайном дослідження, хоча формальним критерієм є наявність кореляційного зв’язку. Прикладом залежних вибірок може бути група досліджуваних до та після експериментального впливу, ситуації, коли ознака вимірюється у одних і тих учасників дослідження в два різних моменти часу. До певної міри залежними можна вважати вибірки, одна з яких сформована із чоловіків, а друга – з їх дружин, або одна вибірка – брати, друга – їх сестри.

Оскільки ми маємо справу із залежними вибірками, фактично – із парами значень, то одиницею аналізу є різниця між цими парами значень.

Емпіричне значення $t$-критерію Стьюдента для залежних вибірок визначається за формулою:

$$t = \frac{\overline{x}_d}{s_d / \sqrt{n}}, \ df = n - 1, \text{ де}$$

$\overline{x}_d$ – середнє арифметичне різниці пар значень;
$s_d$ – стандартне відхилення різниці пар значень;
$n$ – обсяг вибірки досліджуваних.

**Приклад 4.2**. Чи статистично значуще змінилися показники агресивності підлітків після спеціальних психологічних корекційних вправ?

Таблиця 4.2.

Таблиця первинних даних та проміжних розрахунків

| № | До вправ ($X_1$) | Після вправ ($X_2$) | $(x_i - \overline{X}_1)$ | $(x_i - \overline{X}_2)$ | $(x_i - \overline{X}_1)^2$ | $(x_i - \overline{X}_2)^2$ | $(x_i - \overline{X}_1)^3$ | $(x_i - \overline{X}_2)^3$ | $(x_i - \overline{X}_1)^4$ | $(x_i - \overline{X}_2)^4$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 24 | 22 | -9,5 | -7,4 | 90,25 | 54,76 | -857,38 | -405,22 | 8145,06 | 2998,66 |
| 2 | 12 | 12 | -21,5 | -17,4 | 462,25 | 302,76 | -9938,38 | -5268,02 | 213675,06 | 91663,62 |
| 3 | 42 | 41 | 8,5 | 11,6 | 72,25 | 134,56 | 614,13 | 1560,90 | 5220,06 | 18106,39 |
| 4 | 34 | 31 | 0,5 | 1,6 | 0,25 | 2,56 | 0,13 | 4,10 | 0,06 | 6,55 |
| 5 | 28 | 32 | -5,5 | 2,6 | 30,25 | 6,76 | -166,38 | 17,58 | 915,06 | 45,70 |
| 6 | 55 | 44 | 21,5 | 14,6 | 462,25 | 213,16 | 9938,38 | 3112,14 | 213675,06 | 45437,19 |
| 7 | 50 | 50 | 16,5 | 20,6 | 272,25 | 424,36 | 4492,13 | 8741,82 | 74120,06 | 180081,41 |
| 8 | 52 | 32 | 18,5 | 2,6 | 342,25 | 6,76 | 6331,63 | 17,58 | 117135,06 | 45,70 |
| 9 | 50 | 42 | 16,5 | 12,6 | 272,25 | 158,76 | 4492,13 | 2000,38 | 74120,06 | 25204,74 |
| 10 | 32 | 21 | -1,5 | -8,4 | 2,25 | 70,56 | -3,38 | -592,70 | 5,06 | 4978,71 |
| 11 | 33 | 34 | -0,5 | 4,6 | 0,25 | 21,16 | -0,13 | 97,34 | 0,06 | 447,75 |
| 12 | 31 | 36 | -2,5 | 6,6 | 6,25 | 43,56 | -15,63 | 287,50 | 39,06 | 1897,47 |
| 13 | 60 | 38 | 26,5 | 8,6 | 702,25 | 73,96 | 18609,63 | 636,06 | 493155,06 | 5470,08 |
| 14 | 25 | 23 | -8,5 | -6,4 | 72,25 | 40,96 | -614,13 | -262,14 | 5220,06 | 1677,72 |
| 15 | 33 | 33 | -0,5 | 3,6 | 0,25 | 12,96 | -0,13 | 46,66 | 0,06 | 167,96 |
| 16 | 29 | 26 | -4,5 | -3,4 | 20,25 | 11,56 | -91,13 | -39,30 | 410,06 | 133,63 |
| 17 | 17 | 16 | -16,5 | -13,4 | 272,25 | 179,56 | -4492,13 | -2406,10 | 74120,06 | 32241,79 |
| 18 | 12 | 18 | -21,5 | -11,4 | 462,25 | 129,96 | -9938,38 | -1481,54 | 213675,06 | 16889,60 |
| 19 | 25 | 25 | -8,5 | -4,4 | 72,25 | 19,36 | -614,13 | -85,18 | 5220,06 | 374,81 |
| 20 | 26 | 12 | -7,5 | -17,4 | 56,25 | 302,76 | -421,88 | -5268,02 | 3164,06 | 91663,62 |
| $\overline{X}_1$=33,5; $\overline{X}_2$=29,4. Σ = | | | 0 | 0 | 3671 | 2210,8 | 17325 | 713,76 | 1502014,25 | 519533,1 |

У таблиці представлені результати однієї й тієї ж групи досліджуваних, з якими двічі проводилося психологічне тестування. Змістовно такі вибірки називають залежними (зв’язаними, парними). Будемо перевіряти первинні дані на можливість застосування параметричного $t$-критерію Стьюдента для залежних вибірок.

1. Для перевірки даних на відповідність закону нормального розподілу використаємо критерій М.А. Плохінського. Дані відповідають закону нормального розподілу, якщо:

$$t_A = \frac{|A|}{m_A} \le 3 \text{ та } t_E = \frac{|E|}{m_E} \le 3.$$

Для розрахунку показників асиметрії та ексцесу необхідно розрахувати середнє арифметичне значення та стандартне відхилення.

$$\bar{x}_1 = 33{,}5; \ \bar{x}_2 = 29{,}4.$$

$$s_1 = \sqrt{\frac{\Sigma\left(x_i - \bar{x}_1\right)^2}{n-1}} = \sqrt{\frac{3671}{20-1}} = 13{,}9;$$

$$s_2 = \sqrt{\frac{\Sigma\left(x_i - \bar{x}_2\right)^2}{n-1}} = \sqrt{\frac{2210{,}8}{20-1}} = 10{,}79.$$

Для визначення асиметрії та ексцесу в обох вибірках дослідження застосуємо точні розрахункові формули:

$$A_1 = \frac{n \cdot \Sigma\left(x_i - \bar{x}_1\right)^3}{(n-1)\cdot(n-2)\cdot s_1^3} = \frac{20 \cdot 17325}{(20-1)\cdot(20-2)\cdot 13{,}9^3} = 0{,}377.$$

$$E_1 = \frac{n\cdot(n+1)\cdot\Sigma\left(x_i - \bar{x}_1\right)^4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot s_1^4} - \frac{3\cdot(n-1)^2}{(n-2)\cdot(n-3)} = \frac{20\cdot(20+1)\cdot 1502014{,}25}{(20-1)\cdot(20-2)\cdot(20-3)\cdot 13{,}9^4} -$$
$$- \frac{3\cdot(20-1)^2}{(20-2)\cdot(20-3)} = 2{,}907 - 3{,}539 = -0{,}632.$$

$$A_2 = \frac{n \cdot \Sigma\left(x_i - \bar{x}_2\right)^3}{(n-1)\cdot(n-2)\cdot s_2^3} = \frac{20 \cdot 713{,}76}{(20-1)\cdot(20-2)\cdot 10{,}79^3} = 0{,}033.$$

$$E_2 = \frac{n\cdot(n+1)\cdot\Sigma\left(x_i - \bar{x}_2\right)^4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot s_2^4} - \frac{3\cdot(n-1)^2}{(n-2)\cdot(n-3)} = \frac{20\cdot(20+1)\cdot 519533{,}10}{(20-1)\cdot(20-2)\cdot(20-3)\cdot 10{,}79^4} -$$
$$- \frac{3\cdot(20-1)^2}{(20-2)\cdot(20-3)} = 2{,}769 - 3{,}539 = -0{,}77.$$

Визначаємо помилки репрезентативності асиметрії та ексцесу для $n = 20$.

$$m_A = \sqrt{\frac{6}{n+3}} = \sqrt{\frac{6}{20+3}} = 0{,}511;$$

$$m_E = 2 \cdot \sqrt{\frac{6}{n+3}} = 2 \cdot \sqrt{\frac{6}{20+3}} = 1{,}022.$$

Згідно з критерієм М.А. Плохинського показники асиметрії й ексцесу свідчать про статистичну відмінність емпіричного розподілу від нормального в тому випадку, якщо вони перевищують за абсолютною величиною свої помилки репрезентативності в 3 рази.

$$t_{A_1} = \frac{|A_1|}{m_A} = \frac{|0{,}377|}{0{,}511} = 0{,}74; \quad t_{E1} = \frac{|E_1|}{m_E} = \frac{|-0{,}632|}{1{,}022} = 0{,}62.$$

$$t_{A_2} = \frac{|A_2|}{m_A} = \frac{|0{,}033|}{0{,}511} = 0{,}06; \quad t_{E2} = \frac{|E_2|}{m_E} = \frac{|-0{,}77|}{1{,}022} = 0{,}75.$$

За результатами розрахунків показники асиметрії та ексцесу в обох вибірках не перевищують утричі свої помилки репрезентативності, тому можемо зробити висновок, що розподіли відповідають закону нормального розподілу. Тобто, можна використовувати параметричний $t$-критерій Стьюдента для залежних вибірок.

2. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: агресивність підлітків не змінилася після спеціальних корекційних вправ ($\mu_1 = \mu_2$);

$H_1$: агресивність підлітків змінилася після спеціальних корекційних вправ ($\mu_1 \neq \mu_2$).

3. Розраховуємо емпіричне значення $t$-критерію Стьюдента для залежних вибірок.Для розрахунку $t$-критерію необхідно розрахувати середнє арифметичне різниці пар значень ($\bar{x}_d$) та стандартне відхилення різниці пар значень ($s_d$).

$$\bar{x}_d = 4{,}1.$$

$$s_d = \sqrt{\frac{\Sigma\left(d_i - \bar{x}_d\right)^2}{n-1}} = \sqrt{\frac{1155{,}8}{20-1}} = 7{,}8.$$

$$t = \frac{\bar{x}_d}{s_d / \sqrt{n}} = \frac{4{,}1}{7{,}8/\sqrt{20}} = 2{,}35.$$

Таблиця 4.3 – Таблиця первинних даних та проміжних розрахунків

| № | **До вправ ($X_1$)** | **Після вправ ($X_2$)** | $d_i = X_1 - X_2$ | $(d_i - \bar{x}_d)$ | $(d_i - \bar{x}_d)^2$ |
|---|---|---|---|---|---|
| 1 | 24 | 22 | 2 | -2,1 | 4,41 |
| 2 | 12 | 12 | 0 | -4,1 | 16,81 |
| 3 | 42 | 41 | 1 | -3,1 | 9,61 |
| 4 | 34 | 31 | 3 | -1,1 | 1,21 |
| 5 | 28 | 32 | -4 | -8,1 | 65,61 |
| 6 | 55 | 44 | 11 | 6,9 | 47,61 |
| 7 | 50 | 50 | 0 | -4,1 | 16,81 |
| 8 | 52 | 32 | 20 | 15,9 | 252,81 |
| 9 | 50 | 42 | 8 | 3,9 | 15,21 |
| 10 | 32 | 21 | 11 | 6,9 | 47,61 |
| 11 | 33 | 34 | -1 | -5,1 | 26,01 |
| 12 | 31 | 36 | -5 | -9,1 | 82,81 |
| 13 | 60 | 38 | 22 | 17,9 | 320,41 |
| 14 | 25 | 23 | 2 | -2,1 | 4,41 |
| 15 | 33 | 33 | 0 | -4,1 | 16,81 |
| 16 | 29 | 26 | 3 | -1,1 | 1,21 |
| 17 | 17 | 16 | 1 | -3,1 | 9,61 |
| 18 | 12 | 18 | -6 | -10,1 | 102,01 |
| 19 | 25 | 25 | 0 | -4,1 | 16,81 |
| 20 | 26 | 12 | 14 | 9,9 | 98,01 |
| | | | $\bar{x}_d = 4,1$ | $\Sigma = 0$ | $\Sigma = 1155,8$ |

4. За таблицею критичних значень критерію $t$-Стьюдента (додаток 3) для $df = n - 1 = 20 - 1 = 19$ та заданого $\alpha = 0,05$ знаходимо $t_{крит.} = 2,09$ (двобічна критична область).

Оскільки $|t_{емп.}| > t_{крит.}$ (2,35 > 2,09), гіпотеза $H_0$ відхиляється.

5. Оцінка розміру ефекту здійснюється за індексом $d_z$-Коена:

$$d_z = \frac{|\bar{x}_d|}{s_d} = \frac{4,1}{7,8} = 0,53.$$

Згідно Дж. Коена значення 0,53 для індексу $d_z$ відповідає середньому рівню.

<u>Висновок:</u> критерій $t$-Стьюдента для залежних вибірок дозволив відхилити нульову гіпотезу на користь альтернативної ($t_{(19)} = 2,35$; $p < 0,05$; двобічна критична область). Статистично значуще змінилося середнє значення агресивності підлітків після спеціальних психологічних корекційних вправ ($\bar{x}_d = 4,1$; $s_d = 7,8$). Розмір стандартизованого ефекту – середній ($d_z = 0,53$).

– відео-огляд процедури аналізу даних Прикладу 4.2. у статистичній програмі SPSS Statistics 23.0.

### 4.4 Критерій *t*-Стьюдента для однієї вибірки

*t*-критерій Стьюдента для однієї вибірки дозволяє перевірити гіпотезу, що середнє значення ознаки відрізняються від деякого іншого визначеного значення.

Емпіричне значення *t*-критерію Стьюдента для однієї вибірки визначається за формулою:

$$t = \frac{\bar{x} - A}{s/\sqrt{n}}, \ \mathrm{df} = \mathrm{n} - 1,$$

де $\bar{x}$ – середнє арифметичне значення;
*A* – величина, з якою середнє арифметичне порівнюється;
*s* – стандартне відхилення;
*n* – обсяг вибірки досліджуваних.

**Приклад 4.3**. Чи статистично значуще перевищує рівень розвитку посттравматичного стресового розладу (ПТСР) нормативний показник у 35 балів у ліквідаторів Чорнобильської катастрофи?

Для того щоби мати підстави застосувати критерій *t*-Стьюдента для однієї вибірки необхідно перевірити результати тестування на відповідність закону нормального розподілу.

1. З метою перевірки даних на нормальність розподілу використаємо критерій М.А. Плохінського, згідно з яким розподіл вважається нормальним, якщо:

$$t_A = \frac{|A|}{m_A} \le 3 \quad \text{та} \quad t_E = \frac{|E|}{m_E} \le 3.$$

Для розрахунку показників асиметрії та ексцесу необхідно розрахувати середнє арифметичне значення та стандартне відхилення.

$$\bar{x} = 36 \qquad s = \sqrt{\frac{\Sigma\left(x_i - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{368}{23-1}} = 4{,}09.$$

Таблиця 4.4 –Таблиця первинних даних та проміжних розрахунків

| **№** | **ПТСР (*X*)** | $\left(x_i - \overline{x}\right)$ | $\left(x_i - \overline{x}\right)^2$ | $\left(x_i - \overline{x}\right)^3$ | $\left(x_i - \overline{x}\right)^4$ |
|---|---|---|---|---|---|
| 1 | 33 | -3 | 9 | -27 | 81 |
| 2 | 34 | -2 | 4 | -8 | 16 |
| 3 | 39 | 3 | 9 | 27 | 81 |
| 4 | 36 | 0 | 0 | 0 | 0 |
| 5 | 41 | 5 | 25 | 125 | 625 |
| 6 | 42 | 6 | 36 | 216 | 1296 |
| 7 | 29 | -7 | 49 | -343 | 2401 |
| 8 | 34 | -2 | 4 | -8 | 16 |
| 9 | 38 | 2 | 4 | 8 | 16 |
| 10 | 32 | -4 | 16 | -64 | 256 |
| 11 | 29 | -7 | 49 | -343 | 2401 |
| 12 | 31 | -5 | 25 | -125 | 625 |
| 13 | 31 | -5 | 25 | -125 | 625 |
| 14 | 38 | 2 | 4 | 8 | 16 |
| 15 | 40 | 4 | 16 | 64 | 256 |
| 16 | 42 | 6 | 36 | 216 | 1296 |
| 17 | 33 | -3 | 9 | -27 | 81 |
| 18 | 35 | -1 | 1 | -1 | 1 |
| 19 | 35 | -1 | 1 | -1 | 1 |
| 20 | 41 | 5 | 25 | 125 | 625 |
| 21 | 40 | 4 | 16 | 64 | 256 |
| 22 | 38 | 2 | 4 | 8 | 16 |
| 23 | 37 | 1 | 1 | 1 | 1 |
| | $\overline{x}$ = 36 | Σ = 0 | Σ = 368 | Σ = -210 | Σ = 10988 |

Для визначення значення асиметрії та ексцесу застосуємо точні розрахункові формули:

$$A = \frac{n \cdot \Sigma\left(x_i - \overline{x}\right)^3}{(n-1)\cdot(n-2)\cdot s^3} = \frac{23\cdot(-210)}{(23-1)\cdot(23-2)\cdot 4,09^3} = -0,153.$$

$$E = \frac{n\cdot(n+1)\cdot\Sigma\left(x_i - \overline{x}\right)^4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot s^4} - \frac{3\cdot(n-1)^2}{(n-2)\cdot(n-3)} = \frac{23\cdot(23+1)\cdot 10988}{(23-1)\cdot(23-2)\cdot(23-3)\cdot 4,09^4} -$$

$$- \frac{3\cdot(23-1)^2}{(23-2)\cdot(23-3)} = 2,346 - 3,457 = -1,111.$$

Визначаємо помилки репрезентативності асиметрії та ексцесу для $n$ = 23.

$$m_A = \sqrt{\frac{6}{n+3}} = \sqrt{\frac{6}{23+3}} = 0{,}48;$$

$$m_E = 2 \cdot \sqrt{\frac{6}{n+3}} = 2 \cdot \sqrt{\frac{6}{23+3}} = 0{,}96.$$

Згідно з критерієм М.А. Плохинського показники асиметрії й ексцесу свідчать про статистичну відмінність емпіричного розподілу від нормального в тому випадку, якщо вони перевищують за абсолютною величиною свої помилки репрезентативності в 3 рази.

$$t_A = \frac{|A|}{m_A} = \frac{|-0{,}153|}{0{,}48} = 0{,}32;$$

$$t_E = \frac{|E|}{m_E} = \frac{|-1{,}111|}{0{,}96} = 1{,}16.$$

За результатами розрахунків показники асиметрії та ексцесу не перевищують утричі свої помилки репрезентативності, тому можемо зробити висновок, що розподіл відповідає закону нормального розподілу. Тобто, можна використовувати параметричний $t$-критерій Стьюдента для однієї вибірки.

2. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha$ = 0,05;

$H_0$: рівень розвитку ПТСР у ліквідаторів Чорнобильської катастрофи не перевищує нормативний показник в 35 балів ($\mu \leq A$);

$H_1$: рівень розвитку ПТСР у ліквідаторів Чорнобильської катастрофи перевищує нормативний показник в 35 балів ($\mu > A$).

3. Обчислюємо емпіричне значення $t$-критерію Стьюдента для однієї вибірки.

$$t = \frac{\bar{x} - A}{s/\sqrt{n}} = \frac{36-35}{4{,}09/\sqrt{23}} = 1{,}173.$$

4. За таблицею критичних значень критерію $t$-Стьюдента (додаток 3) для $df = n - 1 = 23 - 1 = 22$ та заданого $\alpha$ = 0,05 знаходимо $t_{крит.}$= 1,72 (однобічна критична область).

Оскільки $|t_{емп.}| < t_{крит.}$ (1,173 < 1,72), гіпотеза $H_0$ приймається.

5. Визначаємо розмір ефекту за індексом $d_{os}$-Коена:

$$d_{os} = \frac{|\bar{x} - A|}{s} = \frac{|36 - 35|}{4,09} = 0,24.$$

Розмір стандартизованого ефекту – малий.

Висновок. Критерій $t$-Стьюдента для однієї вибірки показав, що для заданого набору даних немає достатніх доказів, щоб стверджувати про статистично значуще перевищення середнього значення розвитку ПТСР у ліквідаторів Чорнобильської катастрофи ($\bar{x}$ = 36; $s$ = 4,09) в порівнянні з нормативним показник в 35 балів ($t_{(22)}$= 1,173; $p$ > 0,05; однобічна критична область). Згідно інтерпретації Дж. Коена розмір ефекту відповідає малому рівню ($d_{os}$ =0,24).

– відео-огляд процедури аналізу даних Прикладу 4.3. у статистичній програмі SPSS Statistics 23.0.

**Завдання на самостійну підготовку**

1. Вимоги до даних при застосуванні параметричних критеріїв статистичного порівняння вибірок досліджуваних.
2. Особливості використання та алгоритм розрахунку критерію $t$-Стьюдента для незалежних вибірок.
3. Особливості використання та алгоритм розрахунку критерію $t$-Стьюдента для залежних вибірок.
4. Особливості використання та алгоритм розрахунку критерію $t$-Стьюдента для однієї вибірки.

## ЛЕКЦІЯ 5
## НЕПАРАМЕТРИЧНІ МЕТОДИ СТАТИСТИЧНОГО ПОРІВНЯННЯ ВИБІРОК ДОСЛІДЖУВАНИХ

### План лекції

5.1. Критерії порівняння ознак
5.1.1. *U*-критерій Манна-Уітні
5.1.2. *H*-критерій Краскела-Уолліса
5.2. Критерії розпізнавання зсувів
5.2.1. Критерій *T*-Вілкоксона
5.2.2. Критерій $\chi_r^2$-Фрідмана
5.3. Критерії порівняння розподілів
5.3.1. Критерії $\chi^2$-Пірсона

**Час проведення**: 4 учбові години.

### Література

1. Гланц С. Медико-биологическая статистика. Пер. с англ. – М., Практика, 1998. – 459 с.

2. Гржибовский А.М., Унгуряну Т.Н. Анализ биомедицинских данных с использованием пакета статистических программ SPSS: учебное пособие. – Архангельск: Изд-во Северного государственного медицинского университета, 2017. – 293 с.

3. Климчук В.О. Математичні методи у психології. Навчальний посібник. – К.: Освіта України, 2009. – 288 с.

4. Кричевец А. Н., Корнеев А. А., Рассказова Е. И. Основы статистики для психологов. – М.: Акрополь, 2019. – 286 с.

5. Телейко А.Б. Чорней Р.К. Математико-статистичні методи в соціології та психології: навч. посібник. – Київ: МАУП, 2007. – 418 с.

6. Харькова О.А., Соловьев А.Г. Статистические методы и математическое моделирование: учебное пособие. – Архангельск: Изд-во Северного государственного медицинского университета, 2017. – 164 с.

7. Tomczak M., Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size // Trends in Sport Sciences, 2014. – 1(21). – p. 19-25.

## 5.1 Критерії порівняння ознак

### 5.1.1 *U*-критерій Манна-Уітні

*U*-критерій вперше був запропонований Френком Вілкоксоном (1945) для аналізу значущості відмінностей між двома однаковими за обсягом незалежними вибірками. Дослідники Г.Б. Манн та Д.Р. Уітні (1947) модифікували його для вибірок різного обсягу. Є декілька способів обчислення *U*-критерію і, відповідно, кілька варіантів таблиць статистичної значущості. Здійснюючи розрахунок *U*-Манна-Уітні варто на це звертати увагу. Також можна зустріти різні назви цього критерію: критерій Манна-Уітні-Вілкоксона (скорочено – MWW), критерій Вілкоксона-Манна-Уітні, критерій рангових сум Вілкоксона.

Критерій Манна-Уітні є непараметричним, тобто його застосування не вимагає нормальності розподілу і рівності дисперсій. У цьому плані критерій менш вимогливий, ніж його параметричний аналог – критерій *t*-Стьюдента для незалежних вибірок. Водночас *U*-критерій менш чутливий, ніж *t*-критерій.

*Обмеження до застосування критерію:*

1. Вибірки дослідження повинні бути незалежними.
2. Дані повинні бути виміряні не нижче порядкової шкали та варіювати в достатньо широкому діапазоні.
3. У кожній вибірці повинно бути щонайменше три спостереження. Якщо в одній вибірці 2 особи, то в іншій має бути мінімум 5.

Емпіричне значення *U*-критерію Манна-Уітні визначається за формулою:

$$U = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x,$$

де $n_1$ – кількість досліджуваних у першій групі;
$n_2$ – кількість досліджуваних у другій групі;
$n_x$ – кількість досліджуваних у групі з більшою сумою рангів;
$T_x$ – більша із двох рангових сум.

Якщо емпіричне значення *U* дорівнює або менше критичного значення для заданого α-рівня, то $H_0$ відхиляється і формулюється висновок, що відмінності між вибірками статистично значущі.

Для оцінки розміру ефекту для критерію *U*-Манна-Уітні існує декілька формул.

Найбільш часто в спеціальній літературі вказується, що значення розміру ефекту оцінюється через *z*-показник. Стандартизована статистика тесту ділиться на квадратний корінь розміру вибірки.

$$r_{effect} = \frac{z}{\sqrt{n}},$$

де $z$ – стандартизований показник для $U$-значення;
$n$ – загальний обсяг досліджуваних.

Б. Кінг та Е. Мініум запропонували використовувати вираз запропонована Г. Глассом, що базується на рангово-бісеріальній кореляції:

$$r_{effect} = \frac{2 \cdot (\overline{R}_1 - \overline{R}_2)}{n_1 + n_2},$$

де $\overline{R}_1$, $\overline{R}_2$ – середній ранг в першій та другій вибірках;
$n_1$, $n_2$ – обсяг досліджуваних у першій та другій вибірках.

Є можливість застосовувати альтернативну формулу через емпіричне значення критерію $U$-Манна-Уітні, що приводить до того самого результату:

$$r_{effect} = \frac{2 \cdot U}{n_1 \cdot n_2},$$

де $U$ – емпіричне значення $U$-критерію Манна-Уітні;
$n_1$, $n_2$ – обсяг досліджуваних у першій та другій вибірках.
$n$ – загальний обсяг досліджуваних в обох вибірках.

Інтерпретація рівня розміру ефекту запропонованих $r$-індексів (згідно Дж. Коена):

$0,00 \leq |r| < 0,10$ – несуттєвий;
$0,10 \leq |r| < 0,30$ – малий;
$0,30 \leq |r| < 0,50$ – середній;
$0,50 \leq |r| < 1,00$ – великий.

**Приклад 5.1**. У трудовому колективі вивчили рівень мотивації досягнення в кожного працівника. Чи є статистичні відмінності між чоловіками та жінками в прояві мотивації досягнення?

Дві вибірки в дослідженні незалежні. Підстав вважати, що первинні дані за обома групами відповідають закону нормального розподілу немає. Вибірки досліджуваних недостатньо великі, щоби зробити якісний висновок при перевірці даних на нормальність розподілу. Отже, для виявлення відмінностей між групами досліджуваних будемо використовувати непараметричний $U$-критерій Манна-Уітні.

Таблиця 5.1 –Таблиця первинних даних та проміжних розрахунків

| № | **Чоловіки ($X_1$)** | **Жінки ($X_2$)** | **$Rx_1$** | **$Rx_2$** |
|---|---|---|---|---|
| 1 | 8 | 10 | 11,5 | 14,5 |
| 2 | 4 | 5 | 3 | 6 |
| 3 | 11 | 12 | 17,5 | 20,5 |
| 4 | 5 | 7 | 6 | 9,5 |
| 5 | 6 | 8 | 8 | 11,5 |
| 6 | 7 | 10 | 9,5 | 14,5 |
| 7 | 3 | 4 | 1 | 3 |
| 8 | 9 | 11 | 13 | 17,5 |
| 9 | 11 | 11 | 17,5 | 17,5 |
| 10 | 4 | 5 | 3 | 6 |
| 11 | 12 | | 20,5 | |
| | | | Σ = 110,5 | Σ = 120,5 |

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: рівень мотивації чоловіків не відрізняється від рівня мотивації в жінок;

$H_1$: рівень мотивації чоловіків відрізняється від рівня мотивації в жінок.

2. Дані обох груп об'єднуємо в один ряд та присвоюємо кожному значенню ранг за порядком зростання. Якщо є декілька однакових числових значень, то використовується правило зв'язаних рангів.

3. Підраховуємо суми рангів для обох груп. $\Sigma Rx_1 = 110{,}5$; $\Sigma Rx_2 = 120{,}5$.

Для перевірки правильності ранжування можна використовувати той факт, що $\Sigma_R = n \cdot (n+1)/2$. У нашому випадку сума рангів усієї вибірки досліджуваних дорівнює 231 ($\Sigma_R = 231$), $n \cdot (n+1)/2 = 21 \cdot (21+1)/2 = 231$. Тотожність підтвердилась (231 =231).

4. Визначаємо більшу із двох рангових сум. $T_x = 120{,}5$.

5. Обчислюємо емпіричне значення критерію $U$-Манна-Уітні за формулою:

$$U = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x = (11 \cdot 10) + \frac{10 \cdot (10+1)}{2} - 120{,}5 = 44{,}5.$$

6. За таблицею критичних значень $U$-критерію Манна-Уітні для ненаправлених альтернатив (додаток 5) для $df_1 = n_1 = 11$, $df_2 = n_2 = 10$ та заданого $\alpha = 0{,}05$ знаходимо $U_{крит.} = 26$.

Оскільки $U_{емп.} > U_{крит.}$ (44,5 > 26), гіпотеза $H_0$ приймається.

1. Визначаємо розмір ефекту.

$$r_{effect} = \frac{z}{\sqrt{n}}$$

$$z=\frac{U-\frac{n_1\cdot n_2}{2}}{\sqrt{\frac{n_1\cdot n_2\cdot(n_1+n_2+1)}{12}}}=\frac{44,5-\frac{11\cdot 10}{2}}{\sqrt{\frac{10\cdot 11(11+10+1)}{12}}}=-0,74.$$

$$r_{effect}=\frac{z}{\sqrt{n}}=\frac{-0,74}{\sqrt{21}}=-0,16.$$ (знак не інтерпретується)

Згідно методичних рекомендацій Дж. Коена розмір ефекту відповідає малому рівню.

Висновок. Представлені дані не дають підстав відхилити нульову гіпотезу критерієм *U*-Манна-Уітні ($U_{(11,\,10)}$ = 44,5; $p$ > 0,05; двобічна критична область). Отже, відсутні статистично значущі відмінності в середніх рангах прояву мотивації досягнення між чоловіками ($\overline{R}=10,05$) та жінками ($\overline{R}=12,05$). Розмір стандартизованого ефекту відповідає малому рівню ($r_{effect}=0,16.$).

– відео-огляд процедури аналізу даних Прикладу 5.1. у статистичній програмі SPSS Statistics 23.0.

### 5.1.2 *H*-критерій Краскела-Уолліса

Непараметричний *H*-критерій Краскела-Уолліса використовують для оцінювання відмінностей між трьома та більше вибірками за рівнем прояву ознаки. Іноді цей критерій ще називають непараметричним дисперсійним аналізом. Він дає змогу встановити, чи змінюється рівень ознаки при переході від однієї групи до іншої, але не вказує напрямок цих змін, тобто не можна дізнатися яка саме група відрізняється від інших. В даній ситуації здійснювати попарне порівняння вибірок за допомогою критерію *U*-Манна-Уітні некоректно, тому що відбудеться завищення статистичної значущості результатів, тобто збільшення ймовірності помилки першого роду. Для цього розроблені непараметричні методи множинного порівняння. Коли кількість досліджуваних у вибірках однакові застосовують непараметричний варіант критерію Ньюмена-Кейлса, а при різних обсягах вибірок – критерій Данна (більш детально Гланц С., с. 351).

Якщо є необхідність виявити тенденції зміни величин ознаки при переході від однієї вибірки до іншої порівнюючи три та більше вибірок необхідно використати критерій тенденцій *S*-Джонкхієра-Терпстри. Цей критерій можна розглядати як продовження тесту *H*-Краскела-Уолліса, оскільки він не лише констатує відмінності, а й зазначає напрямок змін.

*Обмеження до застосування критерію:*

1. Вибірки мають бути незалежними.
2. Дані повинні бути виміряні не нижче порядкової шкали.

Емпіричне значення *H*-критерію Краскела-Уолліса визначається за допомогою формули:

$$H = \frac{12}{n(n+1)} \cdot \Sigma \frac{R_i^2}{n_i} - 3(n+1),$$

де $n$ – загальна кількість досліджуваних у всіх вибірках;

$n_i$ – обсяг вибірки $i$;

$R_i$ – сума рангів для вибірки $i$.

У разі наявності зв'язаних (однакових) рангів загальна формула *H*-критерію Краскела-Уолліса обчислює дещо неточне значення. У такому разі необхідно здійснити поправку на однакові ранги. Скоригована формула набуває такого вигляду:

$$H = \frac{\frac{12}{n(n+1)} \cdot \Sigma \frac{R_i^2}{n_i} - 3(n+1)}{1 - \frac{\Sigma T_i}{n^3 - n}},$$

де $T_i = \left(t_i^3 - t_i\right);$

$t_i$ – розмір $i$-групи зв'язаних рангів.

Для прийняття рішення про рівень статистичної значущості необхідно емпіричне (розрахункове) значення критерію *H* порівняти з критичним (табличним) значенням.

Якщо кількість вибірок $k \geq 3$, $n \geq 6$, то користуємося таблицею критичних значень $\chi^2$, $df = k - 1$, (додаток 2).

Якщо $k = 3$, $n \leq 5$, то користуємося таблицею критичних значень *H*-Краскела-Уолліса (додаток 7).

Якщо емпіричне значення дорівнює або більше критичного значення для заданого α-рівня, то $H_0$ відхиляється і формулюється висновок, що відмінності між вибірками статистично значущі.

Щоб дізнатися, наскільки сильний вплив рівня прояву ознаки на залежну змінну корисно визначити розмір статистичного ефекту. На жаль, для *H*-Краскела-Уолліса немає єдиного узгодженого методу розрахунку. Відомо два індекси ES: $\varepsilon^2$ (епсилон квадрат) та $\eta^2$ (ета квадрат). Більш консервативним є $\eta^2$.

$$\varepsilon^2 = \frac{H}{(n^2-1)/(n+1)},$$

де $H$ – емпіричне значення $H$-критерію Краскеала-Уолліса;
$n$ – загальна кількість досліджуваних у всіх вибірках.

$$\eta^2 = \frac{H-k+1}{n-k},$$

де $H$ – емпіричне значення $H$-критерію Краскела-Уолліса;
$k$ – кількість груп;
$n$ – загальна кількість досліджуваних у всіх вибірках.

В літературних джерелах пропонуються наступні орієнтовні межі для оцінки величини статистичного ефекту для індексу $\varepsilon^2$:

- $0{,}00 \le \varepsilon^2 < 0{,}01$ – несуттєвий;
- $0{,}01 \le \varepsilon^2 < 0{,}08$ – малий;
- $0{,}08 \le \varepsilon^2 < 0{,}26$ – середній;
- $0{,}26 \le \varepsilon^2$ – великий.

**Приклад 5.2.** У здобувачів вищої освіти різних навчальних спеціальностей на 4 курсі навчання визначили інтегральну оцінку професійної ідентичності. Чи відрізняється професійна ідентичність у студентів різних спеціальностей?

Таблиця 5.2 –Таблиця первинних даних та проміжних розрахунків

| № | Психологи | | Політологи | | Соціологи | | Журналісти | |
|---|---|---|---|---|---|---|---|---|
| | ПІ | Ранг | ПІ | Ранг | ПІ | Ранг | ПІ | Ранг |
| 1 | 69 | 16 | 62 | 14 | 29 | 1 | 71 | 18 |
| 2 | 89 | 29 | 45 | 6 | 76 | 21 | 94 | 32 |
| 3 | 91 | 31 | 77 | 22 | 70 | 17 | 100 | 35 |
| 4 | 83 | 27 | 44 | 5 | 42 | 4 | 98 | 33 |
| 5 | 72 | 19 | 54 | 12 | 51 | 9 | 84 | 28 |
| 6 | 46 | 7 | 31 | 3 | 25 | 2 | 82 | 26 |
| 7 | 52 | 10 | 73 | 20 | 53 | 11 | 79 | 23 |
| 8 | 99 | 34 | 81 | 25 | 67 | 15 | | |
| 9 | 90 | 30 | | | 48 | 8 | | |
| 10 | 58 | 13 | | | | | | |
| 11 | 80 | 24 | | | | | | |
| | | $\Sigma = 240$ | | $\Sigma = 107$ | | $\Sigma = 88$ | | $\Sigma = 195$ |

ПІ – рівень професійної ідентичності.

Первинні дані представлені в шкалі порядку. Вибірки в дослідження незалежні. Необхідно оцінити відмінність одночасно між чотирма вибірками за рівнем досліджуваної ознаки. Для вирішення даного завдання розроблений критерій $H$-Краскела-Уолліса.

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: здобувачі вищої освіти різних спеціальностей не відрізняються за параметром професійної ідентичності;

$H_1$: здобувачі вищої освіти різних спеціальностей відрізняються за параметром професійної ідентичності.

2. Дані всіх груп досліджуваних об'єднуємо у один ряд та присвоюємо кожному числу ранг за порядком зростання. Якщо є декілька однакових числових значень, то використовується правило зв'язаних рангів.

3. Підраховуємо суми рангів для всіх груп. $R_1 = 240$; $R_2 = 107$; $R_3 = 88$; $R_4 = 195$.

Для перевірки правильності ранжування можна використовувати той факт, що $\Sigma_R = n \cdot (n+1)/2$. У нашому випадку сума рангів всієї вибірки досліджуваних дорівнює 630 ($\Sigma_R = 630$), $n \cdot (n+1)/2 = 35 \cdot (35+1)/2 = 630$. Тотожність підтвердилась (630 = 630).

4. У нашому прикладі відсутні зв'язані ранги. Емпіричне значення критерію $H$-Краскела-Уолліса будемо визначати за загальною формулою:

$$H = \frac{12}{n(n+1)} \cdot \Sigma \frac{R_i^2}{n_i} - 3(n+1).$$

$$H = \frac{12}{35 \cdot (35+1)} \cdot \left[ \frac{240^2}{11} + \frac{107^2}{8} + \frac{88^2}{9} + \frac{195^2}{7} \right] - 3 \cdot (35+1) = 15{,}43.$$

5. Так як, $k \geq 3$, $n \geq 6$ при прийнятті рішення про рівень статистичної значущості використовуємо таблицю критичних значень $\chi^2$ (додаток 2). За даною таблицею для $df = k - 1 = 4 - 1 = 3$ та заданого $\alpha = 0{,}05$ знаходимо $\chi^2_{крит.} = 7{,}815$.

Оскільки $\chi^2_{емп.} > \chi^2_{крит.}$ (15,43 > 7,815), гіпотеза $H_0$ відхиляється.

6. Оцінюємо величину розміру статистичного ефекту за індексом $\varepsilon^2$:

$$\varepsilon^2 = \frac{H}{(n^2-1)/(n+1)} = \frac{15{,}43}{(35^2-1)/(35+1)} = 0{,}454.$$

Розмір стандартизованого ефекту – великий.

<u>Висновок.</u> Критерій $H$-Краскела-Уолліса дозволив відхилити нульову гіпотезу ($H_{(3)} = 15{,}43$; $p < 0{,}05$). Існує статистично значуща відмінність в середніх рангах параметра професійної ідентичності у здобувачі вищої освіти різних спеціальностей. Розмір ефекту – великий ($\varepsilon^2 = 0{,}454$).

– відео-огляд процедури аналізу даних Прикладу 5.2. у статистичній програмі SPSS Statistics 23.0.

**Приклад 5.3.** У колективах з різним стилем керівництва було виміряно рівень особистісної відповідальності у працівників. Чи впливає стиль керівництва на рівень особистісної відповідальності досліджуваних?

Таблиця 5.3 –Таблиця первинних даних та проміжних розрахунків

| № | Авторитарний стиль керівництва | | Демократичний стиль керівництва | | Ліберальний стиль керівництва | |
|---|---|---|---|---|---|---|
| | **Рівень відповідал.** | **Ранг** | **Рівень відповідал.** | **Ранг** | **Рівень відповідал.** | **Ранг** |
| 1 | 12 | 18,5 | 11 | 16 | 10 | 13 |
| 2 | 10 | 13 | 8 | 6 | 9 | 9,5 |
| 3 | 12 | 18,5 | 8 | 6 | 9 | 9,5 |
| 4 | 9 | 9,5 | 14 | 20 | 11 | 16 |
| 5 | 8 | 6 | 7 | 4 | 6 | 2,5 |
| 6 | 11 | 16 | 10 | 13 | 6 | 2,5 |
| 7 | 15 | 21 | 5 | 1 | | |
| 8 | | | 9 | 9,5 | | |
| | | $\Sigma = 102,5$ | | $\Sigma = 75,5$ | | $\Sigma = 53$ |

Первинні дані представлені в шкалі порядку. Суть завдання полягає в оцінці відмінностей між трьома вибірками за рівнем розвитку ознаки. Вибірки в дослідження незалежні. За таких умов слід користуватися $H$-критерієм Краскела-Уолліса.

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0,05$;

$H_0$: рівень особистісної відповідальності не залежить від стилю керівництва;

$H_1$: рівень особистісної відповідальності залежить від стилю керівництва.

2. Дані всіх груп досліджуваних об'єднуємо в один ряд та присвоюємо кожному числу ранг за порядком зростання. Якщо є декілька однакових числових значень, то використовується правило зв'язаних рангів.

3. Підраховуємо суми рангів для всіх груп. $R_1 = 102,5$; $R_2 = 75,5$; $R_3 = 53$.

Для перевірки правильності ранжування можна використовувати той факт, що $\Sigma_R = n \cdot (n+1)/2$. У нашому випадку сума рангів усієї вибірки досліджуваних дорівнює 231 ($\Sigma_R$ = 229), $n \cdot (n+1)/2 = 21 \cdot (21+1)/2 = 231$. Тотожність підтвердилась (231 = 231).

3. У нашому прикладі є зв'язані ранги. Емпіричне значення критерію *H*-Краскела-Уолліса будемо визначати за скоригованою формулою:

$$H = \frac{\frac{12}{n(n+1)} \cdot \Sigma \frac{R_i^2}{n_i} - 3(n+1)}{1 - \frac{\Sigma T_i}{n^3 - n}},$$

де $T_i = (t_i^3 - t_i)$;

$t_i$ – розмір *i*-групи зв'язаних рангів.

Таблиця 5.4 –Розмір груп зв'язаних рангів

| **Рівень емпатії** | 6 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|
| $t_i$ | 2 | 3 | 4 | 3 | 3 | 2 |
| $T_i = (t_i^3 - t_i)$ | 6 | 24 | 60 | 24 | 24 | 6 |

$$H = \frac{\frac{12}{21 \cdot (21+1)} \cdot \left[\frac{102{,}5^2}{7} + \frac{75{,}5^2}{8} + \frac{53^2}{6}\right] - 3 \cdot (21+1)}{1 - \frac{(6+24+60+24+24+6)}{21^3 - 21}} = 3{,}71.$$

5. Так як, $k \geq 3$, $n \geq 6$ при прийнятті рішення про рівень статистичної значущості використовуємо таблицю критичних значень $\chi^2$ (додаток 2). За даною таблицею для $df = k - 1 = 3 - 1 = 2$ та заданого $\alpha = 0{,}05$ знаходимо $\chi^2_{крит.} = 5{,}991$.

Оскільки $\chi^2_{емп.} < \chi^2_{крит.}$ (3,71 < 5,991), гіпотеза $H_0$ приймається.

6. Оцінюємо величину розміру статистичного ефекту за індексом $\varepsilon^2$:

$$\varepsilon^2 = H \cdot \frac{n+1}{n^2 - 1} = 3{,}71 \cdot \frac{21+1}{21^2 - 1} = 0{,}186.$$

Розмір стандартизованого ефекту – середній.

<u>Висновок.</u> Рівень особистісної відповідальності у працівників колективу не залежить від стилю керівництва. Згідно статистики критерію *H*-Краскела-Уолліса ($H_{(2)} = 3{,}71$; $p > 0{,}05$) для відхилення нульової гіпотези немає достатніх підстав, проте середній розмір ефекту ($\varepsilon^2 = 0{,}186$) свідчить про доцільність збільшення статистичної потужності дослідження.

## 5.2 Критерії розпізнавання зсувів

### 5.2.1 Критерій *T*-Вілкоксона

Критерій *T*-Вілкоксона використовують для порівняння показників, виміряних у двох різних умовах на одній і тій же вибірці досліджуваних. Він дозволяє встановити не лише спрямованість змін, але й їхню виразність, тобто довести, що зсув показників у одному напрямку є більш інтенсивним, ніж у іншому.

Критерій *T*-Вілкоксона є непараметричним, тобто для його застосування знімається вимога нормальності розподілу і вимога рівності дисперсій. В цьому плані критерій менш вимогливий, ніж його параметричний аналог – критерій *t*-Стьюдента для залежних вибірок. Водночас *T*-Вілкоксона менш чутливий, ніж *t*-критерій.

Суть методу полягає в тому, що зміни за абсолютними значеннями між вимірами в тому чи іншому напрямку можуть бути впорядкованими. Якщо зсуви в позитивний чи негативний бік відбуваються випадково, то суми рангів абсолютних значень будуть приблизно рівні. Якщо інтенсивність зсуву в одному з напрямків переважає, то сума рангів абсолютних значень у протилежний бік буде значно меншою, ніж це могло бути у разі випадкових змін.

Доцільно застосовувати цей критерій, коли величина зсувів варіює в достатньо широкому діапазоні. У разі якщо зсуви несуттєво відрізняються між собою (наприклад: +1, -1 та 0), зважаючи на велике число однакових рангів, ранжування втрачає сенс, раціональніше в такій ситуації використати критерій *G*-знаків.

*Обмеження до застосування та особливості критерію:*

1. Вибірка має бути зв'язаною.
2. Дані повинні бути виміряні не нижче порядкової шкали та варіювати в достатньо широкому діапазоні.
3. Мінімальний обсяг вибірки дорівнює 5 досліджуваним.
4. Нульові зсуви (різниця значень дорівнює 0) з розгляду вилучаються. Водночас обсяг вибірки *n* зменшується на кількість нульових зсувів.

Для розрахунку критерію *T*-Вілкоксона не потрібні спеціальні формули, достатньо підрахувати суми рангів для позитивних і негативних величин різниці. За емпіричне значення приймається менша сума рангів.

$T_{емп.} = \min (R_1, R_2)$, де

$R_1$ – сума рангів для позитивних величин різниці;

$R_2$ – сума рангів для негативних величин різниці.

Якщо емпіричне значення *T* дорівнює або менше критичного значення для заданого α-рівня, то $H_0$ відхиляється і формулюється висновок, що відмінності між вибірками статистично значущі.

Для оцінки розміру стандартизованого ефекту після використання критерію $T$-Вілкоксона рекомендується використати одну із запропонованих формул:

$$r_{effect} = \frac{4 \cdot \left| T - \left( \frac{R_1 + R_2}{2} \right) \right|}{n \cdot (n+1)},$$

де $R_1$ – сума рангів для позитивних величин різниці;

$R_2$ – сума рангів для негативних величин різниці;

$T$ – менша сума рангів ($R_1$ або $R_2$);

$n$ – обсяг досліджуваних (обсяг вибірки зменшився на кількість нульових зсувів).

Розмір ефекту для даного критерію найбільш часто оцінюється через $z$-показник. Стандартизована статистика тесту ділиться на квадратний корінь розміру вибірки.

$$r_{effect} = \frac{z}{\sqrt{n}},$$

де $z$ – стандартизована статистика критерію $T$- Вілкоксона;

$n$ – загальний обсяг досліджуваних в обох вибірках.

Дейв Керби запропонував ще один варіант розрахунку розміру стандартизованого ефекту:

$$r_{effect} = \frac{W}{\Sigma R},$$

де $W$ – сума рангів з врахуванням знаку величин різниці;

$\Sigma R$ – сума рангів абсолютних величин різниці.

Інтерпретація рівня розміру ефекту запропонованих $r$-індексів (згідно Дж. Коена):

$0{,}00 \le |r| < 0{,}10$ – несуттєвий;

$0{,}10 \le |r| < 0{,}30$ – малий;

$0{,}30 \le |r| < 0{,}50$– середній;

$0{,}50 \le |r| < 1{,}00$ – великий.

**Приклад 5.4.** Чи змінюється стійкість уваги у студентів впродовж навчального дня? Діагностувалися показники в тій же групі перед першою та після останньої пари занять.

В таблиці представлені результати однієї і тієї ж групи досліджуваних, з якими двічі проводилося психологічне тестування. Змістовно така вибірка називають зв'язаною чи парною. Підстав вважати, що первинні дані відповідають закону нормального розподілу немає. Вибірка досліджуваних недостатньо велика, щоби зробити якісний висновок при перевірці даних на нормальність розподілу. Отже, для виявлення відмінностей між досліджуваними в різний час тестування будемо використовувати непараметричний критерій *T*-Вілкоксона.

Таблиця 5.5 –Таблиця первинних даних та проміжних розрахунків

| **№** | **До занять ($X_1$)** | **Після занять ($X_2$)** | **Різниця $d_i$** | **Ранги абсолютних величин різниці $\|d_i\|$** | **Ранги позитивних величин різниці $d_i$+** | **Ранги негативних величин різниці $d_i$-** |
|---|---|---|---|---|---|---|
| 1 | 23 | 21 | 2 | 3 | 3 | |
| 2 | 17 | 17 | 0 | - | | |
| 3 | 23 | 25 | -2 | 3 | | 3 |
| 4 | 15 | 11 | 4 | 7 | 7 | |
| 5 | 28 | 21 | 7 | 11 | 11 | |
| 6 | 24 | 24 | 0 | - | | |
| 7 | 21 | 16 | 5 | 9,5 | 9,5 | |
| 8 | 22 | 26 | -4 | 7 | | 7 |
| 9 | 17 | 18 | -1 | 1 | | 1 |
| 10 | 26 | 22 | 4 | 7 | 7 | |
| 11 | 25 | 20 | 5 | 9,5 | 9,5 | |
| 12 | 19 | 21 | -2 | 3 | | 3 |
| 13 | 13 | 16 | -3 | 5 | | 5 |
| 14 | 19 | 8 | 11 | 12 | 12 | |
| | | | | $\Sigma_R$ = 78 | $T_1$ = 59 | $T_2$ = 19 |

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha$ = 0,05;

$H_0$: у студентів рівень стійкості уваги не змінюється впродовж навчального дня;

$H_1$: у студентів рівень стійкості уваги змінюється впродовж навчального дня.

2. Обчислюємо різницю між індивідуальними значеннями першого та другого вимірювання («до – після»).

3. Рангуємо абсолютні величини різниці ($|d_i|$) за порядком зростання. Якщо є декілька однакових числових значень, то використовується правило зв'язаних рангів.

Для перевірки правильності ранжування можна використовувати той факт, що $\Sigma_R = n\cdot(n+1)/2$. У нашому випадку сума рангів рівна 78 (ΣR = 78), $n$ = 12 (обсяг вибірки зменшився на кількість нульових зсувів). $n\cdot(n+1)/2 = 12\cdot(12+1)/2 = 78$. Отже, рівність реальної та розрахункової сум рангів підтверджено (78 =78).

4. Підраховуємо суми рангів окремо для позитивних та негативних величин різниці ($d_i$+; $d_i$-):

$R_1 = 59$;

$R_2 = 19$.

5. За емпіричне значення критерію $T_{емп.}$ приймається менша сума:

$T_{емп.} = 19$.

6. За таблицею критичних значень критерію $T$-Вілкоксона (додаток 8) для $df = n = 12$ (обсяг вибірки зменшився на кількість нульових зсувів) та заданого $\alpha = 0{,}05$ знаходимо $T_{крит.} = 13$ (двобічна критична область).

Оскільки $T_{емп.} > T_{крит.}$ (19 > 13), гіпотеза $H_0$ приймається.

7. Визначаємо розмір ефекту:

$$r_{effect} = \frac{z}{\sqrt{n}}$$

$$z = \frac{\min(T_1, T_2) - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} = \frac{19 - \frac{12(12+1)}{4}}{\sqrt{\frac{12(12+1)(2 \cdot 12+1)}{24}}} = -1{,}57.$$

$$r_{effect} = \frac{z}{\sqrt{n}} = \frac{-1{,}57}{\sqrt{12}} = -0{,}45.$$ (знак не інтерпретується)

Розмір стандартизованого ефекту – середній.

<u>Висновок</u>. Емпіричні дані не дозволяють відхилити нульову гіпотезу критерієм $T$-Вілкоксона ($T_{(12)} = 19$; $p > 0{,}05$; двобічна критична область). У студентів рівень стійкості уваги не змінюється впродовж навчального дня. При цьому вище середнього розмір ефекту ($r_{effect} = 0{,}45$) свідчить про доцільність збільшення статистичної потужності тесту.

– відео-огляд процедури аналізу даних Прикладу 5.4. у статистичній програмі SPSS Statistics 23.0.

### 5.2.2 Критерій $\chi_r^2$-Фрідмана

Непараметричний критерій $\chi_r^2$-Фрідмана (читається «хі-ар-квадрат») використовують для порівняння показників, виміряних у трьох і більше умовах на одній і тій же вибірці досліджуваних. Він дає змогу встановити, що показники від однієї умови до іншої змінюються, але не вказує напрямок цих змін, тобто не можна дізнатися яка саме умова відрізняється від інших. В даній ситуації здійснювати попарне порівняння за допомогою критерію *T*-Вілкоксона некоректно, тому що відбудеться завищення статистичної значущості результатів, тобто збільшення ймовірності помилки першого роду. Для цього розроблені непараметричні методи множинного порівняння, як правило, використовують критерій Данна (більш детально Гланц С., с. 351).

Якщо є необхідність виявити тенденції зміни величин ознаки при переході від однієї умови до іншої необхідно використати критерій тенденцій *L*-Пейджа. Цей критерій можна розглядати як продовження критерію Фрідмана, оскільки він не лише констатує відмінності, а й зазначає напрямок змін.

Критерій Фрідмана, як і Вілкоксона використовує процедуру ранжування результатів вимірювань, але ранжування відбувається не за вертикаллю, як у Вілкоксона, а за горизонталлю, від вимірювання до вимірювання. Використання даного критерію не вимагає нормального розподілу варіаційних рядів.

*Обмеження до застосування критерію:*

1. Вибірка має бути зв'язаною.
2. Дані повинні бути виміряні не нижче порядкової шкали.
3. Мінімальна кількість досліджуваних двоє ($n \geq 2$), в яких ознаки вимірювалися в щонайменше трьох умовах ($k \geq 3$).

Емпіричне значення критерію $\chi_r^2$-Фрідмана визначається за допомогою формули:

$$\chi_r^2 = \frac{12}{nk(k+1)} \cdot \Sigma R_i^2 - 3n(k+1),$$

де $n$ – обсяг вибірки досліджуваних;

$k$ – кількість вимірювань;

$R_i$ – сума рангів для умови *i*.

У разі наявності зв'язаних (однакових) рангів загальна формула критерію $\chi_r^2$-Фрідмана обчислює дещо неточне значення. У такому разі необхідно здійснити поправку на однакові ранги. Формула набуває такого вигляду:

$$\chi_r^2 = \frac{12 \cdot \Sigma R_i^2 - 3n^2 k(k+1)^2}{nk(k+1) + \left(nk - \sum_{j=1}^{g} t_j^3\right) \Big/ (k-1)},$$

де $g$ – число груп зв'язаних рангів у кожному рядку таблиці;

$t$ – розмір *j*-групи зв'язаних рангів (скільки однакових рангів у неї входить).

Зверніть увагу! Навіть якщо в рядку відсутні зв'язані ранги, кожен окремий (не зв'язаний) ранг розглядається як група зв'язаних рангів, що має одне значення ($t$ = 1).

Для прийняття рішення про рівень статистичної значущості необхідно емпіричне (розрахункове) значення критерію $\chi_r^2$ порівняти з критичним (табличним) значенням.

Якщо $k$ = 3, $n$ > 9 або $k$ > 3, $n$ > 4, то користуємося таблицею критичних значень $\chi^2$, $df = k - 1$ (додаток 2).

Якщо $k$ = 3, $n$ < 10 або $k$ = 4, $n$ < 5, то користуємося спеціальними таблицями критичних значень $\chi_r^2$-Фрідмана (додатки 10 та 11).

Якщо емпіричне значення дорівнює або більше критичного значення для заданого α-рівня, то $H_0$ відхиляється і формулюється висновок, що відмінності статистично значущі.

Рекомендованим індексом оцінки розміру ефекту для критерію $\chi_r^2$-Фрідмана є $W$-Кендалла.

$$W = \frac{\chi_r^2}{n(k-1)},$$

де $\chi_r^2$ – емпіричне значення $\chi_r^2$-Фрідмана;

$n$ – обсяг вибірки досліджуваних;

$k$ – кількість вимірювань.

Це міра погодженості між досліджуваними, де 0 означає повна відсутність згоди, а 1 – цілковита згода. Інтерпретація розміру ефекту буде залежати від галузі застосування, в психології ми орієнтуємося на показники:

0,00 ≤ $W$ < 0,10 – несуттєвий;

0,10 ≤ $W$ < 0,30 – малий;

0,30 ≤ $W$ < 0,50– середній;

0,50 ≤ $W$< 1,00 – великий.

**Приклад 5.5.** Десяти студентам було запропоновано виконати завдання чотирьох варіантів самостійної роботи з навчальної дисципліни. Чи можна стверджувати, що розроблені викладачем варіанти завдань однакові за складністю? Виконання робіт оцінено за 100 бальною системою.

Таблиця 5.6 –Таблиця первинних даних та проміжних розрахунків

| № | Варіант № 1 | | Варіант № 2 | | Варіант № 3 | | Варіант № 4 | |
|---|---|---|---|---|---|---|---|---|
| | **Оцінка** | **Ранг** | **Оцінка** | **Ранг** | **Оцінка** | **Ранг** | **Оцінка** | **Ранг** |
| 1 | 92 | 3 | 85 | 1 | 96 | 4 | 90 | 2 |
| 2 | 65 | 1 | 73 | 3 | 77 | 4 | 69 | 2 |
| 3 | 59 | 2 | 53 | 1 | 61 | 3 | 63 | 4 |
| 4 | 84 | 2 | 92 | 4 | 86 | 3 | 72 | 1 |
| 5 | 65 | 2 | 68 | 3 | 62 | 1 | 72 | 4 |
| 6 | 71 | 4 | 57 | 1 | 70 | 3 | 63 | 2 |

Продовження таблиці 5.6

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 7 | 86 | 3 | 75 | 1 | 84 | 2 | 91 | 4 |
| 8 | 79 | 2 | 82 | 3 | 72 | 1 | 84 | 4 |
| 9 | 85 | 2 | 86 | 3 | 80 | 1 | 87 | 4 |
| 10 | 100 | 4 | 94 | 2 | 92 | 1 | 97 | 3 |
| | | $\sum = 25$ | | $\sum = 22$ | | $\sum = 23$ | | $\sum = 30$ |

Первинні дані представлені в шкалі порядку. Необхідно порівняти показники, виміряних у чотирьох умовах на одній і тій же вибірці досліджуваних. У такій ситуації слід використовувати критерій $\chi_r^2$-Фрідмана.

1. Встановлюємо рівень рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha = 0{,}05$;

$H_0$: за складністю розроблені викладачем варіанти завдань самостійної роботи не відрізняються;

$H_1$: за складністю розроблені викладачем варіанти завдань самостійної роботи відрізняються.

2. Рангуємо індивідуальні результати першого досліджуваного, отримані ним у першому, другому, третьому і четвертому варіантах завдань. У такий же спосіб необхідно проранжувати індивідуальні значення інших досліджуваних.

3. Розраховуємо суми рангів для кожного варіанту завдань.

4. У нашому прикладі відсутні зв'язані ранги. Емпіричне значення критерію $\chi_r^2$-Фрідмана будемо визначати за загальною формулою:

$$\chi_r^2 = \frac{12}{n \cdot k \cdot (k+1)} \cdot \sum R_i^2 - 3n \cdot (k+1).$$

$$\chi_r^2 = \frac{12}{10 \cdot 4 \cdot (4+1)} \cdot \left[25^2 + 22^2 + 23^2 + 30^2\right] - 3 \cdot 10 \cdot (4+1) = 2{,}28.$$

5. Так як, в нашому прикладі $k > 3$, $n > 4$, то для визначення рівня статистичної значущості користуємося таблицею критичних значень $\chi^2$ (додаток 2). За даною таблицею для $df = k - 1 = 4 - 1 = 3$ та заданого $\alpha = 0{,}05$ знаходимо $\chi^2_{крит.} = 7{,}815$.

Оскільки $\chi^2_{емп.} > \chi^2_{крит.}$ (2,28 < 7,815), гіпотеза $H_0$ приймається.

6. Оцінюємо величину розміру стандартизованого ефекту:

$$W = \frac{\chi_r^2}{n(k-1)} = \frac{2{,}28}{10 \cdot (4-1)} = 0{,}076.$$

Розмір стандартизованого ефекту дуже низький (несуттєвий).

Висновок. Критерій $\chi^2_r$-Фрідмана показав, що не має достатніх доказів для відхилення нульової гіпотези ($\chi^2_{r\,(3)}$ = 2,28; $p$ > 0,05). За складністю розроблені викладачем варіанти завдань самостійної роботи не відрізняються. Розмір стандартизованого ефекту відповідає низькому рівню ($W$ = 0,076).

– відео-огляд процедури аналізу даних Прикладу 5.5. у статистичній програмі SPSS Statistics 23.0.

**Приклад 5.6.** Було виміряно рівень конформізму в 9 досліджуваних у різних експериментальних ситуаціях. Чи впливають умови експерименту на рівень конформізму досліджуваних?

Таблиця 5.7 –Таблиця первинних даних та проміжних розрахунків

| № | **Експеримент № 1** | | **Експеримент № 2** | | **Експеримент № 3** | |
|---|---|---|---|---|---|---|
| | **Рівень конформізму** | **Ранг** | **Рівень конформізму** | **Ранг** | **Рівень конформізму** | **Ранг** |
| 1 | 6 | 1 | 12 | 3 | 8 | 2 |
| 2 | 4 | 1 | 15 | 3 | 9 | 2 |
| 3 | 6 | 1,5 | 11 | 3 | 6 | 1,5 |
| 4 | 6 | 2 | 14 | 3 | 4 | 1 |
| 5 | 5 | 1 | 15 | 3 | 7 | 2 |
| 6 | 8 | 2 | 7 | 1 | 11 | 3 |
| 7 | 10 | 1,5 | 11 | 3 | 10 | 1,5 |
| 8 | 7 | 1 | 13 | 3 | 12 | 2 |
| 9 | 9 | 2,5 | 9 | 2,5 | 8 | 1 |
| | $\Sigma =$ | 13,5 | | 24,5 | | 16 |

Первинні дані представлені в шкалі порядку. Суть завдання полягає в  порівнянні показників, що виміряні в трьох умовах на одній і тій же вибірці досліджуваних. За таких умов слід користуватися критерієм $\chi^2_r$-Фрідмана.

1. Встановлюємо рівень ймовірності помилки першого роду та формулюємо нульову і альтернативну гіпотези:

$\alpha$ = 0,05;

$H_0$: умови експерименту не впливають на рівень конформізму досліджуваних;

$H_1$: умови експерименту впливають на рівень конформізму досліджуваних.

2. Рангуємо індивідуальні результати першого досліджуваного, отримані ним у різних умовах дослідження. У такий же спосіб необхідно проранжувати індивідуальні значення інших досліджуваних.

3. Розраховуємо суми рангів конформізму для кожної експериментальної ситуації.

4. Оскільки в нашому прикладі є зв'язані ранги емпіричне значення критерію $\chi_r^2$-Фрідмана будемо визначати за скоригованою формулою:

$$\chi_r^2 = \frac{12 \cdot \Sigma R_i^2 - 3n^2 k(k+1)^2}{nk(k+1) + \left(nk - \sum_{j=1}^{g} t_j^3\right) \Big/ (k-1)}.$$

Спочатку розрахуємо значення $\sum_{j=1}^{g} t_j^3$:

$$\sum_{j=1}^{g} t_j^3 = (1^3 + 1^3 + 1^3) + (1^3 + 1^3 + 1^3) + (1^3 + 2^3) + (1^3 + 1^3 + 1^3) + (1^3 + 1^3 + 1^3) +$$

$$+ (1^3 + 1^3 + 1^3) + (1^3 + 2^3) + (1^3 + 1^3 + 1^3) + (1^3 + 2^3) = 45.$$

$$\chi_r^2 = \frac{12 \cdot \left[13,5^2 + 24,5^2 + 16^2\right] - 3 \cdot 9^2 \cdot 3 \cdot (3+1)^2}{9 \cdot 3 \cdot (3+1) + (9 \cdot 3 - 45)/(3-1)} = \frac{798}{99} = 8,06.$$

5. Так як, в нашому прикладі $k = 3$, $n < 10$, то для визначення рівня статистичної значущості користуємося спеціальною таблицею критичних значень $\chi_r^2$-Фрідмана (додаток 10). За даною таблицею знаходимо критичне значення близьке до заданого $\alpha = 0,05$, $\chi^2_{0,048.} = 6,222$.

Оскільки $\chi^2_{емп.} > \chi^2_{крит.}$ (8,06 > 6,222), гіпотеза $H_0$ відхиляється.

6. Оцінюємо величину розміру статистичного ефекту:

$$W = \frac{\chi_r^2}{n(k-1)} = \frac{8,06}{9 \cdot (3-1)} = 0,448.$$

Розмір стандартизованого ефекту – середній.

<u>Висновок</u>. Застосування критерію $\chi_r^2$-Фрідмана дає підстави стверджувати, що існують відмінності між середніми рангами конформізму у досліджуваних в різних експериментальних ситуаціях ($\chi_r^2 = 8,06$; $p < 0,05$; $n = 9$). Розмір стандартизованого ефекту відповідає середньому рівню ($W = 0,448$).

## 5.3 Критерії порівняння розподілів

### 5.3.1 Критерій $\chi^2$-Пірсона

Критерій $\chi^2$-Пірсона (критерій узгодженості, критерій $\chi^2$) є найпотужнішим непараметричним критерієм. Його можна застосовувати як до числових, рангових, так і до номінальних даних. До того ж кількість порівнюваних розподілів не обмежується.

Критерій $\chi^2$ відповідає на запитання, чи з однаковою частотою трапляються різні значення ознаки у двох та більше розподілах?

Після встановлення відмінностей, якщо необхідно дізнатися за якими саме ознаками відрізняються розподіли рекомендується застосувати $z$-критерій для порівняння пропорцій або біноміальний критерій $m$. Цінну інформацію можна отримати проаналізувавши скореговані стандартизовані залишки.

*Обмеження до застосування критерію:*

1. Обсяг вибірки повинен бути достатньо великим $n \geq 30$. У разі $n < 30$ критерій $\chi^2$ дає приблизні значення, точність критерію зростає зі збільшенням вибірки.

2. Якщо $df = 1$ (або $k = 2$), усі теоретичні (очікувані) частоти повинні бути більші 5. Тобто, якщо число розрядів задано наперед і не може змінюватися, то ми не можемо застосувати критерій $\chi^2$ не накопичивши мінімальної кількості спостережень.

3. Якщо $df > 1$ (або $k > 2$), не більше 20 % теоретичних (очікуваних) частот можуть бути менші 5. Водночас не повинно бути теоретичних частот менших 1. Іноді, щоби задовольнити цю вимогу доводиться об'єднувати деякі класові інтервали.

3. Розряди не повинні перехрещуватися: якщо спостереження віднести до одного розряду, то його вже не можна віднести ні до якого іншого розряду.

4. У разі співставлення в розподілах ознак, які набувають усього два значення, необхідно вносити поправку на неперервність.

5. Критерій $\chi^2$ дозволяє виявляти відмінність емпіричного розподілу від теоретичного або відмінності між декількома емпіричними розподілами, але не виявляє спрямування цих відмінностей.

*Критерій $\chi^2$-Пірсона застосовують у двох основних випадках:*

1. Для співставлення двох і більше емпіричних розподілів.

2. Для співставлення емпіричного розподілу з теоретичним (нормальним, рівномірним тощо).

Ще одним застосуванням критерія $\chi^2$-Пірсона є перевірка гіпотез про наявність зв'язку між ознаками (змінними). На основі $\chi^2$ створені коефіцієнти взаємної зв'язаності Пірсона $C$, Чупрова $K$ та Крамера $V$.

Основна розрахункова формула має такий вигляд:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(f_{емп.} - f_{теор.}\right)^2}{f_{теор.}},$$

де $k$ – кількість класових інтервалів (градацій) досліджуваної ознаки;

$f_{емп.}$ та $f_{теор.}$ – відповідно емпіричні та теоретичні (очікувані) частоти, що відповідають визначеним градаціям змінної.

У ситуації порівняння між собою двох і більше емпіричних розподілів якісних ознак, отриманих на незалежних вибірках, для розрахунку теоретичної (очікуваної) частоти для кожної комірки таблиці необхідно знати, що:

$$f_{ij} = \frac{f_i \cdot f_j}{n},$$

де $f_i$ – сума частот у всіх комірках $i$-рядка;

$f_j$ – сума частот у всіх комірках $j$- стовпчика;

$n$ – сума частот у всій таблиці зв'язаності ознак.

Для таблиці зв'язаності ознак розміром 2×2 застосовується критерій $\chi^2$-Пірсона з поправкою на неперервність.

$$\chi^2 = \frac{n \cdot \left(\left|ad - bc\right| - \frac{n}{2}\right)^2}{(a+b)\cdot(c+d)\cdot(a+c)\cdot(b+d)},$$

де $a$, $b$, $c$, $d$ – дані, що визначаються згідно з таблицею зв'язаності ознак розміром 2×2;

$n$ – обсяг вибірки досліджуваних.

Обчислене емпіричне значення критерію порівнюють із критичним для заданого α-рівня, що визначається за таблицею критичних значень критерію $\chi^2$ (додаток 2), для кількості ступенів свободи:

$$df = (k - 1)\cdot(m - 1),$$

де $k$ – кількість розрядів (класових інтервалів або градацій ознаки);

$m$ – кількість розподілів, що порівнюються.

Якщо $\chi^2_{емп.}$ менше $\chi^2_{крит.}$, то формулюється висновок про відсутні статистично значущі відмінності між розподілами (приймаємо гіпотезу $H_0$).

Якщо $\chi^2_{емп.}$ більше або дорівнює $\chi^2_{крит.}$, то нульова гіпотеза відхиляється та приймається альтернативна.

Для оцінки розміру стандартизованого ефекту для критерію $\chi^2$-Пірсона в залежності від умов застосовуються різні індекси:

1) $\varphi$ – таблиця зв'язаності ознак має розмір 2×2:

$$\varphi = \sqrt{\frac{\chi^2}{n}},$$

де $\chi^2$ – емпіричне значення критерію $\chi^2$-Пірсона;
$n$ – обсяг вибірки досліджуваних.

2) $V$-Крамера в ситуації, коли таблиця зв'язаності ознак має відмінний від 2×2 розмір:

$$V = \sqrt{\frac{\chi^2}{n \cdot (c-1)}},$$

де $\chi^2$ – емпіричне значення критерію $\chi 2$-Пірсона;
$n$ – обсяг вибірки досліджуваних;
$c$ – найменша кількість градацій із двох змінних.

3) $w$-Коена використовується при співставленні емпіричного розподілу з теоретичним:

$$w = \sqrt{\sum_{i=1}^{k} \frac{(p_{емп.} - p_{теор.})^2}{p_{теор.}}},$$

де $k$ – кількість класових інтервалів (градацій) досліджуваної ознаки;
$p_{емп.}$ та $p_{теор.}$ – відповідно емпіричні та теоретичні (очікувані) відносні частоти.

При інтерпретації величини розміру ефекту для $\varphi$ та $w$-Коена психологи використовують наступні рекомендації:

$0{,}00 \le \varphi, w < 0{,}10$ – несуттєвий;
$0{,}10 \le \varphi, w < 0{,}30$ – малий;
$0{,}30 \le \varphi, w < 0{,}50$– середній;
$0{,}50 \le \varphi, w < 1{,}00$ – великий.

Що стосується інтерпретації для $V$-Крамера, то існують різноманітні емпіричні правила, але одне із них дозволяє інтерпретувати в залежності від числа ступенів свободи *df:*

Таблиця 5.8 –Інтерпретація розміру ефекту індексу $V$-Крамера

| ***df*** | **Несуттєвий** | **Малий** | **Середній** | **Великий** |
|---|---|---|---|---|
| 1 | 0,00 – 0,09 | 0,10 – 0,29 | 0,30 – 0,49 | > 0,50 |
| 2 | 0,00 – 0,06 | 0,07 – 0,20 | 0,21 – 0,34 | > 0,35 |
| 3 | 0,00 – 0,05 | 0,06 – 0,16 | 0,17 – 0,28 | > 0,29 |
| 4 | 0,00 – 0,04 | 0,05 – 0,14 | 0,15 – 0,24 | > 0,25 |
| 5 | 0,00 – 0,04 | 0,05 – 0,12 | 0,13 – 0,21 | > 0,22 |

**Приклад 5.7.** Чи можна стверджувати, що професійні здібності у юнаків та дівчат статистично значуще різні?

Таблиця 5.9 – Таблиця зв’язаності ознак

| **Тип професійних здібностей** | **Юнаки ($f_{емп.}$)** | **Дівчата ($f_{емп.}$)** | **Усього** |
|---|---|---|---|
| Людина-природа | 27 | 41 | 68 |
| Людина-техніка | 29 | 13 | 42 |
| Людина - людина | 32 | 47 | 79 |
| Людина - художній образ | 18 | 30 | 48 |
| Людина - знак | 25 | 24 | 49 |
| Усього | 131 | 155 | 286 |

Первинні дані представлені в номінальній шкалі. Таблиця зв’язаності ознак має розмір відмінний від 2×2, а саме 5×2. У такому випадку застосовується класичний варіант критерію $\chi^2$-Пірсона.

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha$ = 0,05;

$H_0$: між розподілами типів професійних здібностей у юнаків та дівчат відмінностей немає;

$H_1$: існують відмінності між розподілами типів професійних здібностей у юнаків та дівчат.

2. Для розрахунку емпіричного значення $\chi^2$-Пірсона необхідно спочатку визначити теоретичні (очікувані) частоти для кожної комірки таблиці:

$$f_{ij} = \frac{f_i \cdot f_j}{n},$$

де $f_i$ – сума частот у всіх комірках $i$-рядка;

$f_j$ – сума частот у всіх комірках $j$-стовпчика;

$n$ – сума частот у всій таблиці зв’язаності ознак.

Таблиця 5.10 –Значення очікуваних частот

| **Тип професійних здібностей** | **Юнаки ($f_{теор.}$)** | **Дівчата ($f_{теор.}$)** | **Усього** |
|---|---|---|---|
| Людина - природа | $f_{11} = \frac{68 \cdot 131}{286} = 31,1$ | $f_{12} = \frac{68 \cdot 155}{286} = 36,9$ | 68 |
| Людина - техніка | $f_{21} = \frac{42 \cdot 131}{286} = 19,2$ | $f_{22} = \frac{42 \cdot 155}{286} = 22,8$ | 42 |
| Людина - людина | $f_{31} = \frac{79 \cdot 131}{286} = 36,2$ | $f_{32} = \frac{79 \cdot 155}{286} = 42,8$ | 79 |
| Людина - художній образ | $f_{41} = \frac{48 \cdot 131}{286} = 22,0$ | $f_{42} = \frac{48 \cdot 155}{286} = 26,0$ | 48 |
| Людина - знак | $f_{51} = \frac{49 \cdot 131}{286} = 22,4$ | $f_{52} = \frac{49 \cdot 155}{286} = 26,6$ | 49 |
| Усього | 131 | 155 | 286 |

3. Підставимо значення емпіричних і очікуваних частот у формулу для розрахунку $\chi^2$:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(f_{емп.} - f_{теор.}\right)^2}{f_{теор.}} = \frac{(27-31{,}1)^2}{31{,}1} + \frac{(41-36{,}9)^2}{36{,}9} + \frac{(29-19{,}2)^2}{19{,}2} +$$

$$+\frac{(13-22{,}8)^2}{22{,}8} + \frac{(32-36{,}2)^2}{36{,}2} + \frac{(47-42{,}8)^2}{42{,}8} + \frac{(18-22{,}0)^2}{22{,}0} + \frac{(30-26{,}0)^2}{26{,}0} +$$

$$+\frac{(25-22{,}4)^2}{22{,}4} + \frac{(24-26{,}6)^2}{26{,}6} = 12{,}92.$$

4. За таблицею критичних значень критерію $\chi^2$-Пірсона (додаток 2) для заданого α = 0,05 та $df = (k - 1)\cdot(m - 1) = (5 - 1)\cdot(2 - 1) = 4$ знаходимо $\chi^2_{крит.} = 9{,}488$.

Оскільки $\chi^2_{емп.} > \chi^2_{крит.}$ (12,92 > 9,488), гіпотеза $H_0$ відхиляється.

5. Оцінка розміру ефекту здійснюється за індексом *V*-Крамера, тому що таблиця зв'язаності ознак має відмінний від 2×2 розмір:

$$V = \sqrt{\frac{\chi^2}{n\cdot(c-1)}} = \sqrt{\frac{12{,}92}{286\cdot(2-1)}} = 0{,}21.$$

Згідно інтерпретації Дж. Коена розмір ефекту відповідає середньому рівню ($df = 4$).

Висновок. Критерій $\chi^2$-Пірсона дозволив відхилити нульову гіпотезу ($\chi^2_{(4)} = 12{,}92$; $p < 0{,}05$). Типи професійних здібностей у юнаків та дівчат статистично значуще відрізняються. Розмір стандартизованого ефекту – середній ($V = 0{,}21$).

– відео-огляд процедури аналізу даних Прикладу 5.7. у статистичній програмі SPSS Statistics 23.0.

**Приклад 5.8.** За результатами дослідження встановлено, що 97 студентів із 925 з екстраверсією після навчання у закладі вищої освіти отримали диплом з відзнакою, а із інтроверсією – 44 студенти із 661. Чи можна стверджувати, що отримання диплому з відзнакою залежить від характеристики темпераменту.

Таблиця 5.11 –Таблиця зв'язаності ознак

| Параметри | Тип диплому після навчання у ЗВО | | Усього |
|---|---|---|---|
| | з відзнакою | без відзнаки | |
| **Екстраверсія** | 97 | 828 | 925 |
| **Інтроверсія** | 44 | 617 | 661 |
| **Усього** | 141 | 1445 | 1586 |

Суть завдання полягає в порівнянні між собою двох емпіричних розподілів номінальних ознак, отриманих на незалежних вибірках. Таблиця первинних даних має розмір 2×2. У такому випадку застосовується критерій $\chi^2$-Пірсона з поправкою на неперервність.

$$\chi^2 = \frac{n \cdot \left( |ad - bc| - \frac{n}{2} \right)^2}{(a+b) \cdot (c+d) \cdot (a+c)(b+d)}$$

У нашому прикладі: $a$ = 97, $b$ = 828, $c$ = 44, $d$ = 617, $n$ = 1586.

1. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha$ = 0,05;

$H_0$: отримання диплому з відзнакою не залежить від екстраверсії/інтраверсії;

$H_1$: отримання диплому з відзнакою залежить від екстраверсії/інтраверсії.

2. Обчислимо емпіричне значення $\chi^2$-Пірсона з поправкою на неперервність.

$$\chi^2 = \frac{1586 \cdot \left( |97 \cdot 617 - 828 \cdot 44| - \frac{1586}{2} \right)^2}{(97+828) \cdot (44+617) \cdot (97+44) \cdot (828+617)} = 6{,}516.$$

3. За таблицею критичних значень критерію $\chi^2$ (додаток 2) для заданого $\alpha$ = 0,05 та $df = (k-1)\cdot(m-1) = (2-1)\cdot(2-1) = 1$ знаходимо $\chi^2_{крит.}$= 3,841.

Оскільки $\chi^2_{емп.} > \chi^2_{крит.}$ (6,516 > 3,841), гіпотеза $H_0$ відхиляється.

5. Оцінка розміру ефекту здійснюється за індексом $\varphi$, тому що таблиця зв'язаності ознак має розмір 2×2:

$$\varphi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{6{,}516}{1586}} = 0{,}06.$$

Розмір стандартизованого ефекту дуже низький (несуттєвий).

Висновок. Критерій $\chi^2$-Пірсона дає підстави відхилити нульову гіпотезу ($\chi^2_{(1)}$ = 6,516; $p$ < 0,05). Тобто, отримання диплому з відзнакою залежить від екстраверсії/інтраверсії. Проте дуже низький розмір стандартизованого ефекту ($\varphi$ = 0,06) свідчить про низьку практичну цінність отриманих рузультатів.

**Приклад 5.9.** Чи можна на підставі отриманих даних стверджувати, що типи темпераментів у студентів-психологів розподілені з однаковою частотою?

Таблиця 5.12.

Розподіл типів темпераменту в студентів-психологів

| **Тип темпераменту** | **Холерик** | **Флегматик** | **Сангвінік** | **Меланхолік** | **Усього** |
|---|---|---|---|---|---|
| Кількість | 17 | 23 | 32 | 8 | 80 |

Первинні дані представлені в номінальній шкалі. Необхідно перевірити, чи відповідає емпіричний розподіл результатів у вибірці теоретичному в генеральній сукупності. Для вирішення подібних завдань застосовується критерій $\chi^2$-Пірсона для однієї вибірки.

1. Що стосується виду теоретичного розподілу, то в нашому випадку використовується рівномірний розподіл. Суть його в тому, що всі результати вважаються рівноймовірними. За наявності чотирьох типів темпераменту ймовірність діагностувати в студентів-психологів будь-який із них дорівнює 1/4 = 0,25. Іншими словами, якби емпіричний розподіл результатів повністю збігався з теоретичним, то в кожну комірку таблиці потрапило б однакове число типів темпераменту, що дорівнює 80/4 = 20.

З урахуванням даної обставини остаточний варіант розрахункової таблиці даних для нашого прикладу набуде наступного вигляду.

Таблиця 5.13 –Теоретичний і емпіричний розподіл типів темпераменту в студентів-психологів

| **Тип темпераменту** | **Холерик** | **Флегматик** | **Сангвінік** | **Меланхолік** | **Усього** |
|---|---|---|---|---|---|
| Емпіричні частоти ($f_{емп.}$) | 17 | 23 | 32 | 8 | 80 |
| Теоретичні частоти ($f_{теор.}$) | 20 | 20 | 20 | 20 | 80 |

2. Встановлюємо рівень ймовірності помилки першого роду і формулюємо нульову та альтернативну гіпотези:

$\alpha$ = 0,05;

$H_0$: типи темпераменту у студентів-психологів розподілені з однаковою частотою;

$H_1$: типи темпераменту у студентів-психологів розподілені з різною частотою.

3. Обчислюється емпіричне значення критерію :

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(f_{емп.} - f_{теор.}\right)^2}{f_{теор.}} = \frac{(17-20)^2}{20} + \frac{(23-20)^2}{20} + \frac{(32-20)^2}{20} + \frac{(8-20)^2}{20} =$$

$$= 0{,}45 + 0{,}45 + 7{,}2 + 7{,}2 = 15{,}3.$$

4. За таблицею критичних значень критерію $\chi^2$ (додаток 2) для заданого $\alpha$ = 0,05 та $df = (k - 1)\cdot(m - 1) = (4 - 1)\cdot(2 - 1) = 3$ знаходимо $\chi^2_{крит.}$= 7,815.

Оскільки $\chi^2_{емп.} > \chi^2_{крит.}$ (15,3 > 7,815), гіпотеза $H_0$ відхиляється.

5. Оцінка розміру ефекту здійснюється за індексом *w*-Коена.

$$w = \sqrt{\sum_{i=1}^{k} \frac{(p_{емп.} - p_{теор.})^2}{p_{теор.}}} ,$$

де $k$ – кількість класових інтервалів (градацій) досліджуваної ознаки; $p_{емп.}$ та $p_{теор.}$ – емпіричні та теоретичні (очікувані) відносні частоти.

Підраховуємо відносні частоти для всіх категорій ознаки за формулою:

$$p_i = \frac{f_i}{n_i},$$

де $f_i$ – частота для категорії ознаки;

$n_i$ – кількість спостережень у вибірці.

Таблиця 5.14 –Теоретичний і емпіричний розподіл відносних частот типів темпераменту в студентів-психологів

| **Тип темпераменту** | **Холерик** | **Флегматик** | **Сангвінік** | **Меланхолік** | **Усього** |
|---|---|---|---|---|---|
| Емпіричні відносні частоти ( $p_{емп.}$ ) | 17/80 = 0,2125 | 23/80 = 0,2875 | 32/80 = 0,40 | 8/80 = 0,10 | Σ = 1 |
| Теоретичні відносні частоти ( $p_{теор.}$ ) | 20/80 = 0,25 | 20/80 = 0,25 | 20/80 = 0,25 | 20/80 = 0,25 | Σ = 1 |

$$w = \sqrt{\frac{(0,2125-0,25)^2}{0,25} + \frac{(0,2875-0,25)^2}{0,25} + \frac{(0,40-0,25)^2}{0,25} + \frac{(0,10-0,25)^2}{0,25}} = 0,44.$$

Розмір стандартизованого ефекту відповідає середньому рівню.

Висновок. Застосування критерію $\chi^2$-Пірсона дає підстави відхилити нульову гіпотезу та прийняти альтернативну ($\chi^2_{(3)}$= 15,03; $p$ < 0,05). Типи темпераменту розподілені не рівномірно у студентів-психологів. Розмір ефекту відповідає середньому рівню ($w$ = 0,44).

– відео-огляд процедури аналізу даних Прикладу 5.9. у статистичній програмі SPSS Statistics 23.0.

**Завдання на самостійну підготовку**

1. Вимоги до даних при застосуванні непараметричних критеріїв статистичного порівняння вибірок досліджуваних.

2. Особливості використання та алгоритм розрахунку $G$-критерію знаків.

3. Особливості використання та алгоритм розрахунку Q-критерію Розенбаума.

4. Особливості використання та алгоритм розрахунку критерію тенденцій $L$-Пейджа

5. Особливості використання та алгоритм розрахунку критерію тенденцій $S$-Джонкхієра-Терпстри.

6. Особливості використання та алгоритм розрахунку критерію $\varphi$ – кутове перетворення Фішера.

7. Складіть таблицю умов застосування статистичних критеріїв статистичного порівняння вибірок досліджуваних.

# ДОДАТКИ

## Додаток 1

**Критичні значення коефіцієнта кореляції Пірсона $r_{xy}$ (Спірмена $r_s$)**

| *df* | Рівень значущості | | | |
|---|---|---|---|---|
| | **0,10** | **0,05** | **0,01** | **0,001** |
| **5** | 0,805 | 0,878 | 0,959 | 0,991 |
| **6** | 0,729 | 0,811 | 0,917 | 0,974 |
| **7** | 0,669 | 0,754 | 0,875 | 0,951 |
| **8** | 0,621 | 0,707 | 0,834 | 0,925 |
| **9** | 0,582 | 0,666 | 0,798 | 0,898 |
| **10** | 0,549 | 0,632 | 0,765 | 0,872 |
| **11** | 0,521 | 0,602 | 0,735 | 0,847 |
| **12** | 0,497 | 0,576 | 0,708 | 0,823 |
| **13** | 0,476 | 0,553 | 0,684 | 0,801 |
| **14** | 0,458 | 0,532 | 0,661 | 0,780 |
| **15** | 0,441 | 0,514 | 0,641 | 0,760 |
| **16** | 0,426 | 0,497 | 0,623 | 0,742 |
| **17** | 0,412 | 0,482 | 0,606 | 0,725 |
| **18** | 0,400 | 0,468 | 0,590 | 0,708 |
| **19** | 0,389 | 0,456 | 0,575 | 0,693 |
| **20** | 0,378 | 0,444 | 0,561 | 0,679 |
| **21** | 0,369 | 0,433 | 0,549 | 0,665 |
| **22** | 0,360 | 0,423 | 0,537 | 0,652 |
| **23** | 0,352 | 0,413 | 0,526 | 0,640 |
| **24** | 0,344 | 0,404 | 0,515 | 0,629 |
| **25** | 0,337 | 0,396 | 0,505 | 0,618 |
| **26** | 0,330 | 0,388 | 0,496 | 0,607 |
| **27** | 0,323 | 0,381 | 0,487 | 0,597 |
| **28** | 0,317 | 0,374 | 0,479 | 0,588 |
| **29** | 0,311 | 0,367 | 0,471 | 0,579 |
| **30** | 0,306 | 0,361 | 0,463 | 0,570 |
| **31** | 0,301 | 0355 | 0,456 | 0,562 |
| **32** | 0,296 | 0,349 | 0,449 | 0,554 |
| **33** | 0,291 | 0,344 | 0,442 | 0,547 |
| **34** | 0,287 | 0,339 | 0,436 | 0,539 |
| **35** | 0,283 | 0,334 | 0,430 | 0,532 |
| **36** | 0,279 | 0,329 | 0,424 | 0,525 |
| **37** | 0,275 | 0,325 | 0,418 | 0,519 |
| **38** | 0,271 | 0,320 | 0,413 | 0,513 |
| **39** | 0,267 | 0,316 | 0,408 | 0,507 |
| **40** | 0,264 | 0,312 | 0,403 | 0,501 |
| **41** | 0,260 | 0,308 | 0,398 | 0,495 |
| **42** | 0,257 | 0,304 | 0,393 | 0,490 |
| **43** | 0,254 | 0,301 | 0,389 | 0,484 |
| **44** | 0,251 | 0,297 | 0,384 | 0,479 |
| **45** | 0,248 | 0,294 | 0,380 | 0474 |

| | | | | |
|---|---|---|---|---|
| **46** | 0,246 | 0,291 | 0,376 | 0,469 |
| **47** | 0,243 | 0,288 | 0,372 | 0,465 |
| **48** | 0,240 | 0,285 | 0,368 | 0,460 |
| **49** | 0,238 | 0,282 | 0,365 | 0,456 |
| **50** | 0,235 | 0,279 | 0,361 | 0,451 |
| **51** | 0,233 | 0,276 | 0,358 | 0,447 |
| **52** | 0,231 | 0,273 | 0,354 | 0,443 |
| **53** | 0,228 | 0,271 | 0,351 | 0,439 |
| **54** | 0,226 | 0,268 | 0,348 | 0,435 |
| **55** | 0,224 | 0,266 | 0,345 | 0,432 |
| **56** | 0,222 | 0,263 | 0,341 | 0,428 |
| **57** | 0,220 | 0,261 | 0,339 | 0,424 |
| **58** | 0,218 | 0,259 | 0,336 | 0,421 |
| **59** | 0,216 | 0,256 | 0,333 | 0,418 |
| **60** | 0,214 | 0,254 | 0,330 | 0,414 |
| **61** | 0,213 | 0,252 | 0,327 | 0,411 |
| **62** | 0,211 | 0,250 | 0,325 | 0,408 |
| **63** | 0,209 | 0,248 | 0,322 | 0,405 |
| **64** | 0,207 | 0,246 | 0,320 | 0,402 |
| **65** | 0,206 | 0,244 | 0,317 | 0,399 |
| **66** | 0,204 | 0,242 | 0,315 | 0,396 |
| **67** | 0,203 | 0,240 | 0,313 | 0,393 |
| **68** | 0,201 | 0,239 | 0,310 | 0,390 |
| **69** | 0,200 | 0,237 | 0,308 | 0,388 |
| **70** | 0,198 | 0,235 | 0,306 | 0,385 |
| **80** | 0,185 | 0,220 | 0,286 | 0,361 |
| **90** | 0,174 | 0,207 | 0,270 | 0,341 |
| **100** | 0,165 | 0,197 | 0,256 | 0,324 |
| **110** | 0,158 | 0,187 | 0,245 | 0,310 |
| **120** | 0,151 | 0,179 | 0,234 | 0,297 |
| **130** | 0,145 | 0,172 | 0,225 | 0,285 |
| **140** | 0,140 | 0,166 | 0,217 | 0,275 |
| **150** | 0,135 | 0,160 | 0,210 | 0,266 |
| **200** | 0,117 | 0,139 | 0,182 | 0,231 |
| **250** | 0,104 | 0,124 | 0,163 | 0,207 |
| **300** | 0,095 | 0,113 | 0,149 | 0,189 |

## Додаток 2

## Критичні значення критерію $\chi^2$

| *df* | Рівень значущості | | | *df* | Рівень значущості | | |
|---|---|---|---|---|---|---|---|
| | **0,05** | **0,01** | **0,001** | | **0,05** | **0,01** | **0,001** |
| **1** | 3,841 | 6,635 | 10,829 | **39** | 54,572 | 62,428 | 72,080 |
| **2** | 5,991 | 9,210 | 13,817 | **40** | 55,758 | 63,691 | 73,428 |
| **3** | 7,815 | 11,345 | 16,269 | **41** | 56,942 | 64,950 | 74,772 |
| **4** | 9,488 | 13,277 | 18,470 | **42** | 58,124 | 66,206 | 76,111 |
| **5** | 11,070 | 15,086 | 20,519 | **43** | 59,304 | 67,459 | 77,447 |
| **6** | 12,592 | 16,812 | 22,462 | **44** | 60,481 | 68,709 | 78,779 |
| **7** | 14,067 | 18,475 | 24,327 | **45** | 61,656 | 69,957 | 80,107 |
| **8** | 15,507 | 20,090 | 26,130 | **46** | 62,830 | 71,201 | 81,431 |
| **9** | 16,919 | 21,666 | 27,883 | **47** | 64,001 | 72,443 | 82,752 |
| **10** | 18,307 | 23,209 | 29,594 | **48** | 65,171 | 73,683 | 84,069 |
| **11** | 19,675 | 24,725 | 31,271 | **49** | 66,339 | 74,919 | 85,384 |
| **12** | 21,026 | 26,217 | 32,917 | **50** | 67,505 | 76,154 | 86,694 |
| **13** | 22,362 | 27,688 | 34,536 | **51** | 68,669 | 77,386 | 88,003 |
| **14** | 23,685 | 29,141 | 36,132 | **52** | 69,832 | 78,616 | 89,308 |
| **15** | 24,996 | 30,578 | 37,706 | **53** | 70,993 | 79,843 | 90,609 |
| **16** | 26,296 | 32,000 | 39,262 | **54** | 72,153 | 81,069 | 91,909 |
| **17** | 27,587 | 33,409 | 40,801 | **55** | 73,311 | 82,292 | 93,205 |
| **18** | 28,869 | 34,805 | 42,323 | **56** | 74,468 | 83,513 | 94,499 |
| **19** | 30,144 | 36,191 | 43,832 | **57** | 75,624 | 84,733 | 95,790 |
| **20** | 31,410 | 37,566 | 45,327 | **58** | 76,778 | 85,950 | 97,078 |
| **21** | 32,671 | 38,932 | 46,810 | **59** | 77,931 | 87,166 | 98,365 |
| **22** | 33,924 | 40,289 | 48,281 | **60** | 79,082 | 88,379 | 99,649 |
| **23** | 35,172 | 41,638 | 49,742 | **61** | 80,232 | 89,591 | 100,887 |
| **24** | 36,415 | 42,980 | 51,194 | **62** | 81,381 | 90,802 | 102,165 |
| **25** | 37,652 | 44,314 | 52,635 | **63** | 82,529 | 92,010 | 103,442 |
| **26** | 38,885 | 45,642 | 54,068 | **64** | 83,675 | 93,217 | 104,717 |
| **27** | 40,113 | 46,963 | 55,493 | **65** | 84,821 | 94,422 | 105,988 |
| **28** | 41,337 | 48,278 | 56,910 | **66** | 85,965 | 95,626 | 107,257 |
| **29** | 42,557 | 49,588 | 58,320 | **67** | 87,108 | 96,828 | 108,525 |
| **30** | 43,773 | 50,892 | 59,722 | **68** | 88,250 | 98,028 | 109,793 |
| **31** | 44,985 | 52,191 | 61,118 | **69** | 89,391 | 99,227 | 111,055 |
| **32** | 46,194 | 53,486 | 62,508 | **70** | 90,631 | 100,425 | 112,317 |
| **33** | 47,400 | 54,776 | 63,891 | **71** | 91,670 | 101,621 | 113,577 |
| **34** | 48,602 | 56,061 | 65,269 | **72** | 92,808 | 102,816 | 114,834 |
| **35** | 49,802 | 57,342 | 66,641 | **73** | 93,945 | 104,010 | 116,092 |
| **36** | 50,998 | 58,619 | 68,008 | **74** | 95,081 | 105,202 | 117,347 |
| **37** | 52,192 | 59,892 | 69,370 | **75** | 96,217 | 106,393 | 118,599 |
| **38** | 53,384 | 61,162 | 70,728 | **76** | 97,351 | 107,582 | 119,850 |

**Додаток 3**

## Критичні значення критерію *t*-Стьюдента

| *df* | Рівень значущості (двобічна критична область) | | | | | |
|---|---|---|---|---|---|---|
| | **0,10** | **0,05** | **0,02** | **0,01** | **0,002** | **0,001** |
| **1** | 6,31 | 12,70 | 31,82 | 63,70 | 318,30 | 636,61 |
| **2** | 2,92 | 4,30 | 6,97 | 9,92 | 22,33 | 31,60 |
| **3** | 2,35 | 3,18 | 4,54 | 5,84 | 10,22 | 12,92 |
| **4** | 2,13 | 2,78 | 3,75 | 4,60 | 7,17 | 8,61 |
| **5** | 2,01 | 2,57 | 3,37 | 4,03 | 5,90 | 6,87 |
| **6** | 1,94 | 2,45 | 3,14 | 3,71 | 5,21 | 5,96 |
| **7** | 1,89 | 2,36 | 3,00 | 3,50 | 4,79 | 5,41 |
| **8** | 1,86 | 2,31 | 2,90 | 3,36 | 4,50 | 5,04 |
| **9** | 1,83 | 2,26 | 2,82 | 3,25 | 4,30 | 4,78 |
| **10** | 1,81 | 2,23 | 2,76 | 3,17 | 4,14 | 4,59 |
| **11** | 1,80 | 2,20 | 2,72 | 3,11 | 4,03 | 4,44 |
| **12** | 1,78 | 2,18 | 2,68 | 3,05 | 3,93 | 4,32 |
| **13** | 1,77 | 2,16 | 2,65 | 3,01 | 3,85 | 4,22 |
| **14** | 1,76 | 2,14 | 2,62 | 2,98 | 3,79 | 4,14 |
| **15** | 1,75 | 2,13 | 2,60 | 2,95 | 3,73 | 4,07 |
| **16** | 1,75 | 2,12 | 2,58 | 2,92 | 3,69 | 4,02 |
| **17** | 1,74 | 2,11 | 2,57 | 2,90 | 3,65 | 3,97 |
| **18** | 1,73 | 2,10 | 2,55 | 2,88 | 3,61 | 3,92 |
| **19** | 1,73 | 2,09 | 2,54 | 2,86 | 3,58 | 3,88 |
| **20** | 1,73 | 2,09 | 2,53 | 2,85 | 3,55 | 3,85 |
| **21** | 1,72 | 2,08 | 2,52 | 2,83 | 3,53 | 3,82 |
| **22** | 1,72 | 2,07 | 2,51 | 2,82 | 3,51 | 3,79 |
| **23** | 1,71 | 2,07 | 2,50 | 2,81 | 3,49 | 3,77 |
| **24** | 1,71 | 2,06 | 2,49 | 2,80 | 3,47 | 3,74 |
| **25** | 1,71 | 2,06 | 2,49 | 2,79 | 3,45 | 3,73 |
| **26** | 1,71 | 2,06 | 2,48 | 2,78 | 3,44 | 3,71 |
| **27** | 1,71 | 2,05 | 2,47 | 2,77 | 3,42 | 3,69 |
| **28** | 1,70 | 2,05 | 2,46 | 2,76 | 3,41 | 3,67 |
| **29** | 1,70 | 2,05 | 2,46 | 2,76 | 3,40 | 3,66 |
| **30** | 1,70 | 2,04 | 2,46 | 2,75 | 3,39 | 3,65 |
| **40** | 1,68 | 2,02 | 2,42 | 2,70 | 3,31 | 3,55 |
| **60** | 1,67 | 2,00 | 2,39 | 2,66 | 3,23 | 3,46 |
| **120** | 1,66 | 1,98 | 2,36 | 2,62 | 3,17 | 3,37 |
| **∞** | 1,64 | 1,96 | 2,33 | 2,58 | 3,09 | 3,29 |
| | **0,05** | **0,025** | **0,01** | **0,005** | **0,001** | **0,0005** |
| | **Рівень значущості (однобічна критична область)** | | | | | |

**Додаток 4**

## Критичні значення критерію *F*-Фішера для перевірки ненаправлених альтернатив

**$\alpha$ =0,05**

| | | $df$ чисельника | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **10** | **12** | **24** | **…** |
| $df$ знаменника | **3** | 17,4 | 16,0 | 15,4 | 15,1 | 14,9 | 14,7 | 14,6 | 14,5 | 14,4 | 14,3 | 14,1 | 13,9 |
| | **5** | 10,0 | 8,43 | 7,76 | 7,39 | 7,15 | 6,98 | 6,85 | 6,76 | 6,62 | 6,53 | 6,28 | 6,02 |
| | **7** | 8,07 | 6,54 | 5,89 | 5,52 | 5,29 | 5,12 | 5,00 | 4,90 | 4,76 | 4,67 | 4,42 | 4,14 |
| | **10** | 6,94 | 5,46 | 4,83 | 4,47 | 4,24 | 4,07 | 3,95 | 3,86 | 3,72 | 3,62 | 3,37 | 3,08 |
| | **11** | 6,72 | 5,26 | 4,63 | 4,28 | 4,04 | 3,88 | 3,76 | 3,66 | 3,53 | 3,43 | 3,17 | 2,88 |
| | **12** | 6,55 | 5,10 | 4,47 | 4,12 | 3,89 | 3,73 | 3,61 | 3,51 | 3,37 | 3,28 | 3,02 | 2,73 |
| | **13** | 6,41 | 4,97 | 4,35 | 4,00 | 3,77 | 3,60 | 3,48 | 3,39 | 3,25 | 3,15 | 2,89 | 2,60 |
| | **14** | 6,30 | 4,86 | 4,24 | 3,89 | 3,66 | 3,50 | 3,38 | 3,29 | 3,15 | 3,05 | 2,79 | 2,49 |
| | **15** | 6,20 | 4,77 | 4,15 | 3,80 | 3,58 | 3,42 | 3,29 | 3,20 | 3,06 | 2,96 | 2,70 | 2,40 |
| | **16** | 6,12 | 4,69 | 4,08 | 3,73 | 3,50 | 3,34 | 3,22 | 3,13 | 2,99 | 2,89 | 2,63 | 2,32 |
| | **18** | 5,98 | 4,56 | 3,95 | 3,61 | 3,38 | 3,22 | 3,10 | 3,01 | 2,87 | 2,77 | 2,50 | 2,19 |
| | **20** | 5,87 | 4,46 | 3,86 | 3,52 | 3,29 | 3,13 | 3,01 | 2,91 | 2,77 | 2,68 | 2,41 | 2,09 |
| | **30** | 5,57 | 4,18 | 3,59 | 3,25 | 3,03 | 2,87 | 2,75 | 2,65 | 2,51 | 2,41 | 2,14 | 1,79 |
| | **40** | 5,42 | 4,05 | 3,46 | 3,13 | 2,90 | 2,74 | 2,62 | 2,53 | 2,39 | 2,29 | 2,01 | 1,64 |
| | **50** | 5,34 | 3,98 | 3,39 | 3,05 | 2,83 | 2,67 | 2,55 | 2,46 | 2,32 | 2,22 | 1,93 | 1,55 |
| | **70** | 5,25 | 3,89 | 3,31 | 2,98 | 2,75 | 2,60 | 2,47 | 2,38 | 2,24 | 2,14 | 1,85 | 1,44 |
| | **100** | 5,18 | 3,83 | 3,25 | 2,92 | 2,70 | 2,54 | 2,42 | 2,32 | 2,18 | 2,08 | 1,78 | 1,35 |
| | **200** | 5,10 | 3,76 | 3,18 | 2,85 | 2,63 | 2,47 | 2,35 | 2,27 | 2,11 | 2,01 | 1,71 | 1,23 |
| | **…** | 5,03 | 3,69 | 3,12 | 2,79 | 2,57 | 2,41 | 2,29 | 2,19 | 2,05 | 1,95 | 1,64 | |

## Додаток 5

## Критичні значення критерію *U*-Манна-Уїтні для ненаправлених альтернатив (двобічна область)

| $\alpha = 0{,}05$ | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **$df_{1,2}$** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
| **4** | - | 0 | | | | | | | | | | | | | | | | |
| **5** | 0 | 1 | 2 | | | | | | | | | | | | | | | |
| **6** | 1 | 2 | 3 | 5 | | | | | | | | | | | | | | |
| **7** | 1 | 3 | 5 | 6 | 8 | | | | | | | | | | | | | |
| **8** | 2 | 4 | 6 | 8 | 10 | 13 | | | | | | | | | | | | |
| **9** | 2 | 4 | 7 | 10 | 12 | 15 | 17 | | | | | | | | | | | |
| **10** | 3 | 5 | 8 | 11 | 14 | 17 | 20 | 23 | | | | | | | | | | |
| **11** | 3 | 6 | 9 | 13 | 16 | 19 | 23 | 26 | 30 | | | | | | | | | |
| **12** | 4 | 7 | 11 | 14 | 18 | 22 | 26 | 29 | 33 | 37 | | | | | | | | |
| **13** | 4 | 8 | 12 | 16 | 20 | 24 | 28 | 33 | 37 | 41 | 45 | | | | | | | |
| **14** | 5 | 9 | 13 | 17 | 22 | 26 | 31 | 36 | 40 | 45 | 50 | 55 | | | | | | |
| **15** | 5 | 10 | 14 | 19 | 24 | 29 | 34 | 39 | 44 | 49 | 54 | 59 | 64 | | | | | |
| **16** | 6 | 11 | 15 | 21 | 26 | 31 | 37 | 42 | 47 | 53 | 59 | 64 | 70 | 75 | | | | |
| **17** | 6 | 11 | 17 | 22 | 28 | 34 | 39 | 45 | 51 | 57 | 63 | 67 | 75 | 81 | 87 | | | |
| **18** | 7 | 12 | 18 | 24 | 30 | 36 | 42 | 48 | 55 | 61 | 67 | 74 | 80 | 86 | 93 | 99 | | |
| **19** | 7 | 13 | 19 | 25 | 32 | 38 | 45 | 52 | 58 | 65 | 72 | 78 | 85 | 92 | 99 | 106 | 113 | |
| **20** | 8 | 14 | 20 | 27 | 34 | 41 | 48 | 55 | 62 | 69 | 76 | 83 | 90 | 98 | 105 | 112 | 119 | 127 |
| $\alpha = 0{,}01$ | | | | | | | | | | | | | | | | | | |
| **5** | - | - | 0 | | | | | | | | | | | | | | | |
| **6** | - | 0 | 1 | 2 | | | | | | | | | | | | | | |
| **7** | - | 0 | 1 | 3 | 4 | | | | | | | | | | | | | |
| **8** | - | 1 | 2 | 4 | 6 | 7 | | | | | | | | | | | | |
| **9** | 0 | 1 | 3 | 5 | 7 | 9 | 11 | | | | | | | | | | | |
| **10** | 0 | 2 | 4 | 6 | 9 | 11 | 13 | 16 | | | | | | | | | | |
| **11** | 0 | 2 | 5 | 7 | 10 | 13 | 16 | 18 | 21 | | | | | | | | | |
| **12** | 1 | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 21 | 27 | | | | | | | | |
| **13** | 1 | 3 | 7 | 10 | 13 | 17 | 20 | 24 | 24 | 31 | 34 | | | | | | | |
| **14** | 1 | 4 | 7 | 11 | 15 | 18 | 22 | 26 | 26 | 34 | 38 | 42 | | | | | | |
| **15** | 2 | 5 | 8 | 12 | 16 | 20 | 24 | 29 | 29 | 37 | 42 | 46 | 51 | | | | | |
| **16** | 2 | 5 | 9 | 13 | 18 | 22 | 27 | 31 | 31 | 41 | 45 | 50 | 55 | 60 | | | | |
| **17** | 2 | 6 | 10 | 15 | 19 | 24 | 29 | 34 | 34 | 44 | 49 | 54 | 60 | 65 | 70 | | | |
| **18** | 2 | 6 | 11 | 16 | 21 | 26 | 31 | 37 | 37 | 47 | 53 | 58 | 64 | 70 | 75 | 81 | | |
| **19** | 3 | 7 | 12 | 17 | 22 | 28 | 33 | 39 | 39 | 51 | 56 | 63 | 69 | 74 | 81 | 87 | 93 | |
| **20** | 3 | 8 | 13 | 18 | 24 | 30 | 36 | 42 | 42 | 54 | 60 | 67 | 73 | 79 | 86 | 92 | 99 | 105 |

**Додаток 6**

## Критичні значення критерію *U*-Манна-Уітні для направлених альтернатив (однобічна область)

| $\alpha = 0,05$ | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_{1,2}$ | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
| **3** | - | 0 | | | | | | | | | | | | | | | | | |
| **4** | - | 0 | 1 | | | | | | | | | | | | | | | | |
| **5** | 0 | 1 | 2 | 4 | | | | | | | | | | | | | | | |
| **6** | 0 | 2 | 3 | 5 | 7 | | | | | | | | | | | | | | |
| **7** | 0 | 2 | 4 | 6 | 8 | 11 | | | | | | | | | | | | | |
| **8** | 1 | 3 | 5 | 8 | 10 | 13 | 15 | | | | | | | | | | | | |
| **9** | 1 | 4 | 6 | 9 | 12 | 15 | 18 | 21 | | | | | | | | | | | |
| **10** | 1 | 4 | 7 | 11 | 14 | 17 | 20 | 24 | 27 | | | | | | | | | | |
| **11** | 1 | 5 | 8 | 12 | 16 | 19 | 23 | 27 | 31 | 34 | | | | | | | | | |
| **12** | 2 | 5 | 9 | 13 | 17 | 21 | 26 | 30 | 34 | 38 | 42 | | | | | | | | |
| **13** | 2 | 6 | 10 | 15 | 19 | 24 | 28 | 33 | 37 | 42 | 47 | 51 | | | | | | | |
| **14** | 3 | 7 | 11 | 16 | 21 | 26 | 31 | 36 | 41 | 46 | 51 | 56 | 61 | | | | | | |
| **15** | 3 | 7 | 12 | 18 | 23 | 28 | 33 | 39 | 44 | 50 | 55 | 61 | 66 | 72 | | | | | |
| **16** | 3 | 8 | 14 | 19 | 25 | 30 | 36 | 42 | 48 | 54 | 60 | 65 | 71 | 77 | 83 | | | | |
| **17** | 3 | 9 | 15 | 20 | 26 | 33 | 39 | 45 | 51 | 57 | 64 | 70 | 77 | 83 | 89 | 96 | | | |
| **18** | 4 | 9 | 16 | 22 | 28 | 35 | 41 | 48 | 55 | 61 | 68 | 75 | 82 | 88 | 95 | 102 | 109 | | |
| **19** | 4 | 10 | 17 | 23 | 30 | 37 | 44 | 51 | 58 | 65 | 72 | 80 | 87 | 94 | 101 | 109 | 116 | 123 | |
| **20** | 4 | 11 | 18 | 25 | 32 | 39 | 47 | 54 | 62 | 69 | 77 | 84 | 92 | 100 | 107 | 115 | 123 | 130 | 138 |
| $\alpha = 0,01$ | | | | | | | | | | | | | | | | | | | |
| **5** | - | - | 0 | 1 | | | | | | | | | | | | | | | |
| **6** | - | - | 1 | 2 | 3 | | | | | | | | | | | | | | |
| **7** | - | 0 | 1 | 3 | 4 | 6 | | | | | | | | | | | | | |
| **8** | - | 0 | 2 | 4 | 6 | 7 | 9 | | | | | | | | | | | | |
| **9** | - | 1 | 3 | 5 | 7 | 9 | 11 | 14 | | | | | | | | | | | |
| **10** | - | 1 | 3 | 6 | 8 | 11 | 13 | 16 | 19 | | | | | | | | | | |
| **11** | - | 1 | 4 | 7 | 9 | 12 | 15 | 18 | 22 | 25 | | | | | | | | | |
| **12** | - | 2 | 5 | 8 | 11 | 14 | 17 | 21 | 24 | 28 | 31 | | | | | | | | |
| **13** | 0 | 2 | 5 | 9 | 12 | 16 | 20 | 23 | 27 | 31 | 35 | 39 | | | | | | | |
| **14** | 0 | 2 | 6 | 10 | 13 | 17 | 22 | 26 | 30 | 34 | 38 | 43 | 47 | | | | | | |
| **15** | 0 | 3 | 7 | 11 | 15 | 19 | 24 | 28 | 33 | 37 | 42 | 47 | 51 | 56 | | | | | |
| **16** | 0 | 3 | 7 | 12 | 16 | 21 | 26 | 31 | 36 | 41 | 46 | 51 | 56 | 61 | 66 | | | | |
| **17** | 0 | 4 | 8 | 13 | 18 | 23 | 28 | 33 | 38 | 44 | 49 | 55 | 60 | 66 | 71 | 77 | | | |
| **18** | 0 | 4 | 9 | 14 | 19 | 24 | 30 | 36 | 41 | 47 | 53 | 59 | 65 | 70 | 76 | 82 | 88 | | |
| **19** | 1 | 4 | 9 | 15 | 20 | 26 | 32 | 38 | 44 | 50 | 56 | 63 | 69 | 75 | 82 | 88 | 94 | 101 | |
| **20** | 1 | 5 | 10 | 16 | 22 | 28 | 34 | 40 | 47 | 53 | 60 | 67 | 73 | 80 | 87 | 93 | 100 | 107 | 114 |

**Додаток 7**

## Критичні значення критерію *H*-Краскела-Уолліса для трьох вибірок чисельністю $n \le 5$

| Обсяги вибірок | | | | | Обсяги вибірок | | | | | Обсяги вибірок | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | $n_2$ | $n_3$ | $H$ | $p$ | $n_1$ | $n_2$ | $n_3$ | $H$ | $p$ | $n_1$ | $n_2$ | $n_3$ | $H$ | $p$ |
| 2 | 1 | 1 | 2,7000 | 0,500 | 4 | 4 | 1 | 6,6667 | 0,010 | 5 | 4 | 1 | 6,9545 | 0,008 |
| 2 | 2 | 1 | 3,6000 | 0,200 | | | | 6,1667 | 0,022 | | | | 6,8400 | 0,011 |
| 2 | 2 | 2 | 4,5714 | 0,067 | | | | 4,9667 | 0,048 | | | | 4,9855 | 0,044 |
| 3 | 1 | 1 | 3,2000 | 0,300 | | | | 4,8667 | 0,054 | | | | 4,8600 | 0,056 |
| 3 | 2 | 1 | 4,2857 | 0,100 | | | | 4,1667 | 0,082 | | | | 3,9873 | 0,098 |
| | | | 3,8571 | 0,133 | | | | 4,0667 | 0,102 | | | | 3,9600 | 0,102 |
| 3 | 2 | 2 | 5,3572 | 0,029 | 4 | 4 | 2 | 7,0364 | 0,006 | 5 | 4 | 2 | 7,2045 | 0,009 |
| | | | 4,7143 | 0,048 | | | | 6,8727 | 0,011 | | | | 7,1182 | 0,010 |
| | | | 4,5000 | 0,067 | | | | 5,4545 | 0,046 | | | | 5,2727 | 0,049 |
| | | | 4,4643 | 0,105 | | | | 5,2364 | 0,052 | | | | 5,2682 | 0,050 |
| 3 | 3 | 1 | 5,1429 | 0,043 | | | | 4,5545 | 0,098 | | | | 4,5409 | 0,098 |
| | | | 4,5714 | 0,100 | | | | 4,4455 | 0,103 | | | | 4,5182 | 0,101 |
| | | | 4,0000 | 0,129 | 4 | 4 | 3 | 7,1439 | 0,010 | 5 | 4 | 3 | 7,4449 | 0,010 |
| 3 | 3 | 2 | 6,2500 | 0,011 | | | | 7,1364 | 0,011 | | | | 7,3949 | 0,011 |
| | | | 5,3611 | 0,032 | | | | 5,5985 | 0,049 | | | | 5,6564 | 0,049 |
| | | | 5,1389 | 0,061 | | | | 5,5758 | 0,051 | | | | 5,6308 | 0,050 |
| | | | 4,5556 | 0,100 | | | | 4,5455 | 0,099 | | | | 4,5487 | 0,099 |
| | | | 4,2500 | 0,121 | | | | 4,4773 | 0,102 | | | | 4,5231 | 0,103 |
| 3 | 3 | 3 | 7,2000 | 0,004 | 4 | 4 | 4 | 7,6538 | 0,008 | 5 | 4 | 4 | 7,7604 | 0,009 |
| | | | 6,4889 | 0,011 | | | | 7,5385 | 0,011 | | | | 7,7440 | 0,011 |
| | | | 5,6889 | 0,029 | | | | 5,6923 | 0,049 | | | | 5,6571 | 0,049 |
| | | | 5,6000 | 0,050 | | | | 5,6538 | 0,054 | | | | 5,6176 | 0,050 |
| | | | 5,0667 | 0,086 | | | | 4,6539 | 0,097 | | | | 4,6187 | 0,100 |
| | | | 4,6222 | 0,100 | | | | 4,5001 | 0,104 | | | | 4,5527 | 0,102 |
| 4 | 1 | 1 | 3,5714 | 0,200 | 5 | 1 | 1 | 3,8571 | 0,143 | 5 | 5 | 1 | 7,3091 | 0,009 |
| 4 | 2 | 1 | 4,8214 | 0,057 | 5 | 2 | 1 | 5,2500 | 0,036 | | | | 6,8364 | 0,011 |
| | | | 4,5000 | 0,076 | | | | 5,0000 | 0,048 | | | | 5,1273 | 0,046 |
| | | | 4,0179 | 0,114 | | | | 4,4500 | 0,071 | | | | 4,9091 | 0,053 |
| 4 | 2 | 2 | 6,0000 | 0,014 | | | | 4,2000 | 0,095 | | | | 4,1091 | 0,086 |
| | | | 5,3333 | 0,033 | | | | 4,0500 | 0,119 | | | | 4,0364 | 0,105 |
| | | | 5,1250 | 0,052 | 5 | 2 | 2 | 6,5333 | 0,008 | 5 | 5 | 2 | 7,3385 | 0,010 |
| | | | 4,4583 | 0,100 | | | | 6,1333 | 0,013 | | | | 7,2692 | 0,010 |
| | | | 4,1667 | 0,105 | | | | 5,1600 | 0,034 | | | | 5,3385 | 0,047 |
| 4 | 3 | 1 | 5,8333 | 0,021 | | | | 5,0400 | 0,056 | | | | 5,2462 | 0,051 |
| | | | 5,2083 | 0,050 | | | | 4,3733 | 0,090 | | | | 4,6231 | 0,097 |
| | | | 5,0000 | 0,057 | | | | 4,2933 | 0,122 | | | | 4,5077 | 0,100 |
| | | | 4,0556 | 0,093 | 5 | 3 | 1 | 6,4000 | 0,012 | 5 | 5 | 3 | 7,5780 | 0,010 |
| | | | 3,8889 | 0,129 | | | | 4,9600 | 0,048 | | | | 7,5429 | 0,010 |
| 4 | 3 | 2 | 6,4444 | 0,008 | | | | 4,8711 | 0,052 | | | | 5,7055 | 0,046 |
| | | | 6,3000 | 0,011 | | | | 4,0178 | 0,095 | | | | 5,6264 | 0,051 |
| | | | 5,4444 | 0,046 | | | | 3,8400 | 0,123 | | | | 4,5451 | 0,100 |
| | | | 5,4000 | 0,051 | 5 | 3 | 2 | 6,9091 | 0,009 | | | | 4,5363 | 0,102 |

| Обсяги вибірок | | | | | Обсяги вибірок | | | | | Обсяги вибірок | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | $n_2$ | $n_3$ | $H$ | $p$ | $n_1$ | $n_2$ | $n_3$ | $H$ | $p$ | $n_1$ | $n_2$ | $n_3$ | $H$ | $p$ |
| | | | 4,5111 | 0,098 | | | | 6,8218 | 0,010 | 5 | 5 | 4 | 7,8229 | 0,010 |
| | | | 4,4444 | 0,102 | | | | 5,2509 | 0,049 | | | | 7,7914 | 0,010 |
| 4 | 3 | 3 | 6,7455 | 0,010 | | | | 5,1055 | 0,052 | | | | 5,6657 | 0,049 |
| | | | 6,7091 | 0,013 | | | | 4,6509 | 0,091 | | | | 5,6429 | 0,050 |
| | | | 5,7909 | 0,046 | | | | 4,4945 | 0,101 | | | | 4,5229 | 0,099 |
| | | | 5,7273 | 0,050 | 5 | 3 | 3 | 7,0788 | 0,009 | | | | 4,5200 | 0,101 |
| | | | 4,7091 | 0,092 | | | | 6,9818 | 0,011 | 5 | 5 | 5 | 8,0000 | 0,009 |
| | | | 4,7000 | 0,101 | | | | 5,6485 | 0,049 | | | | 7,9800 | 0,010 |
| | | | | | | | | 5,5152 | 0,051 | | | | 5,7800 | 0,049 |
| | | | | | | | | 4,5333 | 0,097 | | | | 5,6600 | 0,051 |
| | | | | | | | | 4,4121 | 0,109 | | | | 4,5600 | 0,100 |
| | | | | | | | | | | | | | 4,5000 | 0,102 |

## Додаток 8

### Критичні значення критерію *T*-Вілкоксона для ненаправлених альтернатив (двобічна область)

| $n$ | Рівень значущості | | $n$ | Рівень значущості | |
|---|---|---|---|---|---|
| | **0,05** | **0,01** | | **0,01** | **0,05** |
| **5** | - | - | **28** | 116 | 91 |
| **6** | 0 | - | **29** | 126 | 100 |
| **7** | 2 | - | **30** | 137 | 109 |
| **8** | 3 | 0 | **31** | 147 | 118 |
| **9** | 5 | 1 | **32** | 159 | 128 |
| **10** | 8 | 3 | **33** | 170 | 138 |
| **11** | 10 | 5 | **34** | 182 | 148 |
| **12** | 13 | 7 | **35** | 195 | 159 |
| **13** | 17 | 9 | **36** | 208 | 171 |
| **14** | 21 | 12 | **37** | 221 | 182 |
| **15** | 25 | 15 | **38** | 235 | 194 |
| **16** | 29 | 19 | **39** | 249 | 207 |
| **17** | 34 | 23 | **40** | 264 | 220 |
| **18** | 40 | 27 | **41** | 279 | 233 |
| **19** | 46 | 32 | **42** | 294 | 247 |
| **20** | 52 | 37 | **43** | 310 | 261 |
| **21** | 58 | 42 | **44** | 327 | 276 |
| **22** | 65 | 48 | **45** | 343 | 291 |
| **23** | 73 | 54 | **46** | 361 | 307 |
| **24** | 81 | 61 | **47** | 378 | 322 |
| **25** | 89 | 68 | **48** | 396 | 339 |
| **26** | 98 | 75 | **49** | 415 | 355 |
| **27** | 107 | 83 | **50** | 434 | 373 |

**Додаток 9**

## Критичні значення критерію *T*-Вілкоксона для направлених альтернатив (однобічна область)

| *n* | Рівень значущості | | *n* | Рівень значущості | |
|---|---|---|---|---|---|
| | **0,05** | **0,01** | | **0,01** | **0,05** |
| **5** | 0 | - | **28** | 130 | 101 |
| **6** | 2 | - | **29** | 140 | 110 |
| **7** | 3 | 0 | **30** | 151 | 120 |
| **8** | 5 | 1 | **31** | 163 | 130 |
| **9** | 8 | 3 | **32** | 175 | 140 |
| **10** | 10 | 5 | **33** | 187 | 151 |
| **11** | 13 | 7 | **34** | 200 | 162 |
| **12** | 17 | 9 | **35** | 213 | 173 |
| **13** | 21 | 12 | **36** | 227 | 185 |
| **14** | 25 | 15 | **37** | 241 | 198 |
| **15** | 30 | 19 | **38** | 256 | 211 |
| **16** | 35 | 23 | **39** | 271 | 224 |
| **17** | 41 | 27 | **40** | 286 | 238 |
| **18** | 47 | 32 | **41** | 302 | 252 |
| **19** | 53 | 37 | **42** | 319 | 266 |
| **20** | 60 | 43 | **43** | 336 | 281 |
| **21** | 67 | 49 | **44** | 353 | 296 |
| **22** | 75 | 55 | **45** | 371 | 312 |
| **23** | 83 | 62 | **46** | 389 | 328 |
| **24** | 91 | 69 | **47** | 407 | 345 |
| **25** | 100 | 76 | **48** | 426 | 362 |
| **26** | 110 | 84 | **49** | 446 | 379 |
| **27** | 119 | 92 | **50** | 466 | 397 |

**Додаток 10**

## Критичні значення критерію $\chi_r^2$-Фрідмана для кількості умов *k* = 3 та кількості досліджуваних *n* < 10

| ***n* = 2** | | ***n* = 3** | | ***n* = 4** | | ***n* = 5** | |
|---|---|---|---|---|---|---|---|
| $\chi_r^2$ | ***p*** | $\chi_r^2$ | ***p*** | $\chi_r^2$ | ***p*** | $\chi_r^2$ | ***p*** |
| 0 | 1,000 | 0,000 | 1,000 | 0,0 | 1,000 | 0,0 | 1,000 |
| 1 | 0,833 | 0,667 | 0,944 | 0,5 | 0,931 | 0,4 | 0,954 |
| 3 | 0,500 | 2,000 | 0,528 | 1,5 | 0,653 | 1,2 | 0,691 |
| 4 | 0,167 | 2,667 | 0,361 | 2,0 | 0,431 | 1,6 | 0,522 |
| | | 4,667 | 0,194 | 3,5 | 0,273 | 2,8 | 0,367 |
| | | 6,000 | 0,028 | 4,5 | 0,125 | 3,6 | 0,182 |
| | | | | 6,0 | 0,069 | 4,8 | 0,124 |
| | | | | 6,5 | 0,042 | 5,2 | 0,093 |
| | | | | 8,0 | 0,0046 | 6,4 | 0,039 |
| | | | | | | 7,6 | 0,024 |
| | | | | | | 8,4 | 0,0085 |
| | | | | | | 10,0 | 0,00077 |

| ***n* = 6** | | ***n* = 7** | | ***n* = 8** | | ***n* = 9** | |
|---|---|---|---|---|---|---|---|
| $\chi_r^2$ | ***p*** | $\chi_r^2$ | ***p*** | $\chi_r^2$ | ***p*** | $\chi_r^2$ | ***p*** |
| 0,00 | 1,000 | 0,000 | 1,000 | 0,00 | 1,000 | 0,000 | 1,000 |
| 0,33 | 0,956 | 0,286 | 0,964 | 0,25 | 0,967 | 0,222 | 0,971 |
| 1,00 | 0,740 | 0,857 | 0,768 | 0,75 | 0,794 | 0,667 | 0,814 |
| 1,33 | 0,570 | 1,143 | 0,620 | 1,00 | 0,654 | 0,889 | 0,865 |
| 2,33 | 0,430 | 2,000 | 0,486 | 1,75 | 0,531 | 1,556 | 0,569 |
| 3,00 | 0,252 | 2,571 | 0,305 | 2,25 | 0,355 | 2,000 | 0,398 |
| 4,00 | 0,184 | 3,429 | 0,237 | 3,00 | 0,285 | 2,667 | 0,328 |
| 4,33 | 0,142 | 3,714 | 0,192 | 3,25 | 0,236 | 2,889 | 0,278 |
| 5,33 | 0,072 | 4,571 | 0,112 | 4,00 | 0,149 | 3,556 | 0,187 |
| 6,33 | 0,052 | 5,429 | 0,085 | 4,75 | 0,120 | 4,222 | 0,154 |
| 7,00 | 0,029 | 6,000 | 0,052 | 5,25 | 0,079 | 4,667 | 0,107 |
| 8,33 | 0,012 | 7,143 | 0,027 | 6,25 | 0,047 | 5,556 | 0,069 |
| 9,00 | 0,0081 | 7,714 | 0,021 | 6,75 | 0,038 | 6,000 | 0,057 |
| 9,33 | 0,0055 | 8,000 | 0,016 | 7,00 | 0,030 | 6,222 | 0,048 |
| 10,33 | 0,0017 | 8,857 | 0,0084 | 7,75 | 0,018 | 6,889 | 0,031 |
| 12,00 | 0,00013 | 10,286 | 0,0036 | 9,00 | 0,0099 | 8,000 | 0,019 |
| | | 10,571 | 0,0027 | 9,25 | 0,0080 | 8,222 | 0,016 |
| | | 11,143 | 0,0012 | 9,75 | 0,0048 | 8,667 | 0,010 |
| | | 12,286 | 0,00032 | 10,75 | 0,0024 | 9,556 | 0,0060 |
| | | 14,000 | 0,000021 | 12,00 | 0,0011 | 10,667 | 0,0035 |
| | | | | 12,25 | 0,00086 | 10,889 | 0,0029 |
| | | | | 13,00 | 0,00026 | 11,556 | 0,0013 |
| | | | | 14,25 | 0,000061 | 12,667 | 0,00066 |
| | | | | 16,00 | 0,0000036 | 13,556 | 0,00035 |
| | | | | | | 14,000 | 0,00020 |
| | | | | | | 14,222 | 0,000097 |
| | | | | | | 14,889 | 0,000054 |
| | | | | | | 16,222 | 0,000011 |
| | | | | | | 18,000 | 0,0000006 |

## Додаток 11

## Критичні значення критерію $\chi_r^2$-Фрідмана для кількості умов $k$ = 4 та кількості досліджуваних $n$ < 5

| $n$ = 2 | | $n$ = 3 | | $n$ = 4 | | | |
|---|---|---|---|---|---|---|---|
| $\chi_r^2$ | $p$ | $\chi_r^2$ | $p$ | $\chi_r^2$ | $p$ | $\chi_r^2$ | $p$ |
| 0 | 1,000 | 0,0 | 1,000 | 0,0 | 1,000 | 5,7 | 0,141 |
| 0,6 | 0,958 | 0,6 | 0,958 | 0,3 | 0,992 | 6,0 | 0,105 |
| 1,2 | 0,834 | 1,0 | 0,910 | 0,6 | 0,928 | 6,3 | 0,094 |
| 1,8 | 0,792 | 1,8 | 0,727 | 0,9 | 0,900 | 6,6 | 0,077 |
| 2,4 | 0,625 | 2,2 | 0,608 | 1,2 | 0,800 | 6,9 | 0,068 |
| 3,0 | 0,542 | 2,6 | 0,524 | 1,5 | 0,754 | 7,2 | 0,054 |
| 3,6 | 0,458 | 3,4 | 0,446 | 1,8 | 0,677 | 7,5 | 0,052 |
| 4,2 | 0,375 | 3,8 | 0,342 | 2,1 | 0,649 | 7,8 | 0,036 |
| 4,8 | 0,208 | 4,2 | 0,300 | 2,4 | 0,524 | 81 | 0,033 |
| 5,4 | 0,167 | 5,0 | 0,207 | 2,7 | 0,508 | 8,4 | 0,019 |
| 6,0 | 0,042 | 5,4 | 0,175 | 3,0 | 0,432 | 8,7 | 0,014 |
| | | 5,8 | 0,148 | 3,3 | 0,389 | 9,3 | 0,012 |
| | | 6,6 | 0,075 | 3,6 | 0,355 | 9,6 | 0,0069 |
| | | 7,0 | 0,054 | 3,9 | 0,324 | 9,9 | 0,0062 |
| | | 7,4 | 0,033 | 4,5 | 0,242 | 10,2 | 0,0027 |
| | | 8,2 | 0,017 | 4,8 | 0,200 | 10,8 | 0,0016 |
| | | 9,0 | 0,0017 | 5,1 | 0,190 | 11,1 | 0,00094 |
| | | | | 5,4 | 0,158 | 12,0 | 0,000072 |

*Навчальне видання*

**Боснюк** Валерій Федорович

# МАТЕМАТИЧНІ МЕТОДИ В ПСИХОЛОГІЇ

***Курс лекцій***
***Мультимедійне навчальне видання***

Підписано до друку 15.06.2020. Формат 60х84 1/16.
Умовн.-друк. арк. 9,0.
Вид. № 26/20.

Сектор редакційно-видавничої діяльності
Національного університету цивільного захисту України
61023 м. Харків, вул. Чернишевська, 94.
www.nuczu.edu.ua