Б. В. ПЛЕСКАЧ, Я. О. КОРКОС

МАТЕМАТИЧНІ МЕТОДИ В ПСИХОЛОГІЇ

НАВЧАЛЬНО-МЕТОДИЧНИЙ ПОСІБНИК

2025

УДК 159.9:519.2(072)

П 38

Рекомендовано до друку рішенням Вченої ради
Київського столичного університету імені Бориса Грінченка
(протокол №9 від 25 вересня 2025 року)

Рецензенти:

Кокун О. М., доктор психологічних наук, професор, дійсний член (академік) НАПН України, заступник директора Інституту психології імені Г. С. Костюка НАПН України з науково-інноваційної роботи;
Виноградов О. Г., кандидат психологічних наук, доцент кафедри соціальної психології Київського інституту сучасної психології і психотерапії

Плескач Б. В., Коркос Я. О.

П 38 Математичні методи в психології: навчально-методичний посібник. – Х.: ФОП Панов А. М., 2025. – 180 с.

ISBN 978-617-8534-52-3

У посібнику, укладеному на основі навчального курсу «Математичні методи в психології», розглянуто місце математичних методів у психології. Наведено інформацію про основну ідею перевірки статистичних гіпотез на основі вибіркового підходу. Детально розглянуто методи описової статистики, критерії вивчення взаємозв'язку між змінними, методи прогнозування та класифікації, критерії перевірки статистичних гіпотез. Посібник адресовано студентам психологічних факультетів для допомоги в засвоєнні курсу «Математичні методи в психології». Посібник вводить читача в термінологію дисципліни та детально розглядає основи математичного аналізу даних.

УДК 159.9:519.2(072)

ISBN 978-617-8534-52-3

ЗМІСТ

ВСТУП

Актуальність вивчення дисципліни «Математичні методи в психології» зумовлено необхідністю підготовки сучасного висококваліфікованого фахівця-психолога, який здатний як надавати психологічну допомогу, так і здійснювати власні дослідження, збирати та аналізувати отримані результати, презентувати результати власних досліджень усно та письмово.

У Київському столичному університеті імені Бориса Грінченка дисципліна «Математичні методи в психології» викладається в студентів першого (бакалаврського) рівня освіти спеціальності С4 «Психологія» за освітніми програмами «Практична психологія» та «Консультаційна психологія». У межах дисципліни студенти знайомляться з описовими методами узагальнення первинних даних дослідження, можливостями графічного представлення отриманих результатів, методами вивчення зв'язку між змінними, складними методами класифікації - досліджених (за спільними ознаками) та змінних (факторний аналіз), критеріями здійснення статистичних висновків.

Застосування математичних методів лежить в основі наукового підходу, дозволяє виявляти та описувати закономірності, які мають високу статистичну ймовірність. Успішне засвоєння курсу сприятиме набуттю студентом здатності самому здійснювати дослідження та аналізувати отримані емпіричні дані, представляти отримані результати у фахових періодичних виданнях, дозволить студентам самостійно орієнтуватись у представлених в науковій літературі результатах дослідження, розуміти лінію математичного аналізу, усвідомити переваги та обмеження представлених результатів.

Пропонований навчально-методичний посібник призначено для допомоги студентам у засвоєнні знань та практичних навичок з дисципліни «Математичні методи в психології». У посібнику доповнюється основна теоретична інформація, навчальний курс представлено в структурованому логічному вигляді – для полегшення оволодіння дисципліною.

1. Основний зміст курсу

Тема 1 Вступ до математичної статистики

Лекція 1. Математичні методи в психології: загальна характеристика

Мета: сформувати уявлення студентів про математичну статистику як сучасну галузь математичної науки, ознайомити з основними її розділами. Сформувати розуміння логіки проведення вибіркового дослідження. Розглянути поняття змінної та шкали, класифікацію шкал у психології. Створити загальне уявлення про нормальний розподіл, вибірковий метод, принципи обчислення необхідного об'єму вибірки та способи її формування. Розглянути логіку наукового дослідження від формулювання наукової гіпотези до перевірки статистичних гіпотез. Сформувати загальне уявлення про перевірку статистичних гіпотез, місце статистичних методів у психологічній науці.

Основні поняття: математична статистика, описова статистика, статистичні висновки, генеральна сукупність, вибірка, статистики, параметри, ймовірнісна вибірка, змінна, вимірювання, шкала, шкала найменувань, шкала порядку, шкала інтервалів, шкала відношення, нормальний розподіл, об'єм імовірнісної вибірки, репрезентативність вибірки, наукова гіпотеза, теоретична гіпотеза, емпірична гіпотеза, статистична гіпотеза, статистичний критерій, H_0 та H_1 гіпотези, потужність критерію.

План.

1.1. Основні завдання та методи математичної статистики.

1.2. Типи змінних: дискретні та неперервні

1.3. Вимірювання та типи шкал

1.4. Закон нормального розподілу

1.5. Вибірковий метод, обчислення мінімального об'єму вибірки для підтвердження/спростування наукових гіпотез на заданому рівні достовірності

1.6. Наукові гіпотези

1.7. Статистичні критерії: загальне уявлення про перевірку статистичних гіпотез

1.8. Математичні методи в психології

1.1. Основні завдання та методи математичної статистики

В. М. Руденко відзначає, що *математична статистика* – це сучасна галузь математичної науки, яка займається статистичним описом результатів експериментів і спостережень, а також побудовою математичних моделей на основі опису ймовірнісних залежностей між певними явищами [12, с. 9]. Математична статистика ґрунтується на теорії імовірності, яка є її теоретичною базою. У найбільш узагальненому вигляді в структурі математичної статистики виділяють дві складові: описову статистику і статистичні висновки (див. рис 1.1).

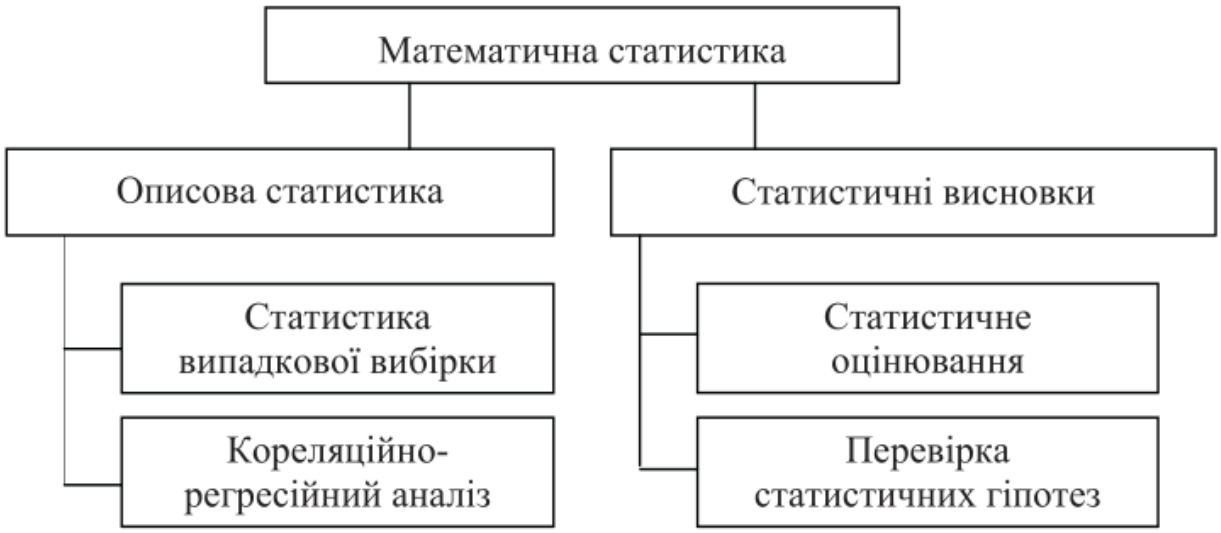

Рис. 1.1. Основні розділи математичної статистики

На думку Дж. Гласса, ці складові склались історично. Описова статистика походить із завдань, які виконували органи державного управління, складаючи звіти та здійснюючи планування: перепис, вимірювання, опис, упорядкування даних, табулювання [20, с. 7].

Теорія імовірності та статистичного висновку створювалась під впливом завдань із оцінки імовірності здобути виграш в азартних іграх [20, с. 8].

Описова статистика займається:

- отриманням описових статистичних характеристик одновимірних вибіркових даних;
- дослідженням взаємозв'язків між двома чи більше змінними за допомогою таблиць зв'язаності та кореляційного аналізу.

Описова статистика включає в себе математичні інструменти для обробки великого та достатньо різноманітного масиву даних, що дозволяє їх описати, узагальнити та виділити найважливіші властивості, оцінити зв'язок між певними змінними. Таким чином, описова статистика дозволяє узагальнити дані, що сприяє їх кращому розумінню.

Описовими методами обчислюються статистики випадкової вибірки, які служать підставою для здійснення статистичних оцінок та висновків про генеральну сукупність.

Статистичні висновки дозволяють:

- на основі статистик випадкових вибірок робити припущення про параметри генеральної сукупності, здійснити оцінку здатності вибіркових статистик відображати параметри генеральної сукупності; оцінити похибку, яка допускається під час екстраполяції вибіркових статистик на генеральну сукупність.
- здійснювати ймовірнісні висновки про зв'язок між певними параметрами генеральної сукупності та їх особливості (перевірка статистичних гіпотез).

Дж. Гудвин відзначає, що математична статистика є однією з важливих складових наукового пізнання, яка дозволяє отримувати знання про світ, у якому людина живе, описувати його закономірності, встановлювати причинно-наслідкові зв'язки між певними явищами [21].

Генеральна сукупність – це певна кількість (кінцева або нескінченна) об'єктів, які є носіями особливості, про яку дослідник збирається зробити

висновки [20, с. 217]. Так, окрему властивість (наприклад, інтелект), яка досліджується в певної множини об’єктів (наприклад, у дітей 10 років), називають генеральною сукупністю.

Термін *вибірка* означає певну частку генеральної сукупності. Для наукового дослідження більш бажано, щоб вибірка була сформована особливим чином, була випадковою (імовірнісною), оскільки дотримання випадковості при формуванні вибірки створює можливість статистичної оцінки, тобто ймовірнісного приписування отриманих на вибірці статистик генеральній сукупності.

Використовуючи статистичний підхід, дослідники намагаються зробити висновок про особливості генеральної сукупності на основі обстеженої вибірки (або вибірок). Такий висновок завжди має певну ймовірність достовірності (наприклад, 95%) та помилки (наприклад 5%).

Вибірковий підхід застосовується через те, що вивчення всіх об’єктів генеральної сукупності зазвичай є неможливим через їх велику кількість, витрати часу та ресурсів дослідників [12, с. 10].

Іншими важливими термінами, які необхідно розглянути, є терміни «статистики» та «параметр», вони вже згадувались у попередньому тексті.

Під *статистиками* розуміють описові значення, отримані на конкретній вибірці, як-от: середнє, медіана, мода, дисперсія, процентилі тощо [20, с. 218].

Параметрами називаються описові значення, отримані на обстеженій вибірці, які обчислені для генеральної сукупності [20, с. 218].

Отже, термін *параметр* застосовуються для генеральної сукупності, *статистики* – для конкретної вибірки. Для опису вибірки застосовуються латинські символи, для генеральної сукупності – грецькі.

На основі вибіркових статистик можна здійснити оцінку параметра генеральної сукупності.

У роботі В. Руденко представлено узагальнену схему проведення дослідження на основі вибіркового підходу (див. рис. 1.2).

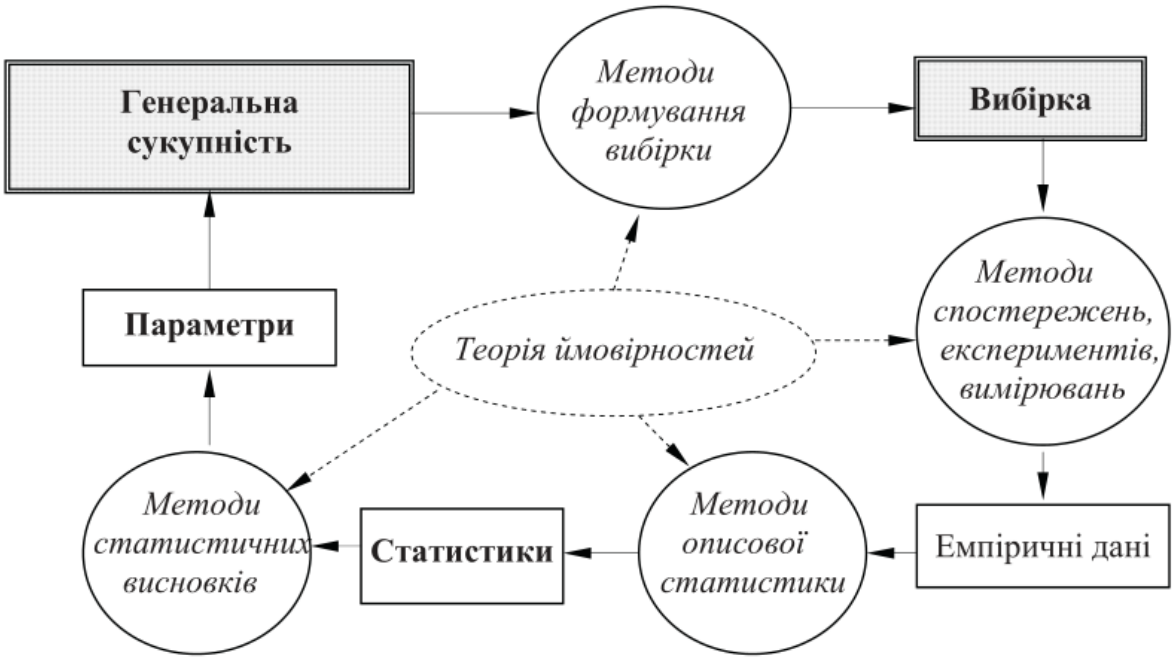

Рис. 1.2. Схема застосування методів математичної статистики

При проведенні дослідження спочатку проводиться теоретичний аналіз вже відомої інформації про якість, яку планується дослідити, генеральну сукупність, яку планується вивчати, про параметри, які вивчаються (середнє, дисперсію).

Використовуючи певні методи, з генеральної сукупності отримують вибірку, яка буде вивчатись.

Вибірка вивчається на основі дослідницьких методів. Це може бути спеціальним чином організоване спостереження, психодіагностичне вимірювання, експеримент. На основі дослідницьких методів отримують емпіричні дані.

Отримані емпіричні дані оброблюються методами описової статистики та отримуються вибіркові показники – статистики.

На основі використання методів статистичних висновків здійснюється оцінка здатності вибіркових статистик описувати параметри генеральної сукупності з певною долею імовірності. Перевіряються статистичні гіпотези (з точністю, яка визначається заданим рівнем імовірності) про певні особливості, які властиві генеральній сукупності.

1.2. Типи змінних: дискретні та неперервні

Під *змінною* розуміють певну властивість, значення якої вимірюються. У психології досліджуються люди, які є носіями певних особливостей. Саме люди певної групи як носії цікавих досліднику психологічних властивостей представляють генеральну сукупність. Таким чином, змінна, яка вивчається, є функцією, яку дає конкретна людина з цієї генеральної сукупності [20]. Наприклад, для дослідження емпатії в підлітків 12-14 років досліднику слід дослідити репрезентативну вибірку з цієї генеральної сукупності. Для цього йому потрібно сформувати випадкову (імовірнісну) групу дітей з усієї сукупності дітей визначеного віку.

Імовірнісна вибірка формується з конкретних осіб, тоді як емпатія (яка досліджується певною шкалою або декількома шкалами) є функцією кожного представника цієї вибірки. Таким чином, рівень емпатії, який потім вимірюється, є похідною змінною, певною функцією в математичному сенсі цього терміну.

Змінну можна визначити як певне явище, яке підлягає дослідженню і яке має певні конкретні прояви, на основі яких його можна виміряти (тобто це явище, яке операціоналізовано).

Самі по собі змінні, які вимірюються, можуть бути *неперервні* та *дискретні*. Під *неперервними* змінними розуміють такі, що можуть набувати будь-яких значень на певній осі чисел. Наприклад, час може вимірюватись годинами, хвилинами, секундами, мілісекундами й так далі. У такому разі явище, яке вимірюється, неперервне і точність його вимірювання залежить від точності обладнання, яке для цього призначено.

Дискретні змінні визначаються як такі, що можуть мати лише певне значення на осі чисел. Наприклад, дитина правильно або неправильно вирішила завдання. Особа відповіла на певне питання або «так», або «ні». У такому разі можливе лише певне значення змінної, що вивчається.

1.3. Вимірювання та типи шкал

Вимірювання можна визначити як процедуру опису певних властивостей об’єкту чи явища за допомогою символьних або числових значень. У психології вимірювання зазвичай здійснюється на основі шкал.

Термін *шкала* походить від латинського scalae, яке можна перекласти як «драбина», та являє собою числову систему, призначену для впорядкування та кількісного або якісного представлення властивостей досліджуваних об'єктів. У шкалі властивості об’єктів описуються за допомогою певних числових або якісних значень.

За допомогою шкали з’являється можливість порівняти вираженість психологічного явища в різних людей. Стенлі Стівенс запропонував розрізняти 4 типи вимірювальних шкал [8, с. 16]:

- шкала найменувань;
- шкала порядку;
- шкала інтервалів;
- шкала відношення.

Шкали *найменувань* та *порядку* називаються також *неметричними (нечисловими) типами даних* (що означає, що у них є лише два значення: так або ні; або більше чи менше).

Шкала *інтервалів* та *шкала відношення* називаються *метричними* або *числовими* типами шкал, вони дозволяють виміряти ступінь вираженості певної психологічної якості.

Шкала найменувань – це нечислова (*номінальна***)** шкала, вона дозволяє відносити прояви об’єкту, який вимірюється, до певної категорії (наприклад, чоловік або жінка). При цьому категорії між собою є рівнозначними.

Шкала порядку. Під порядковою, або *ординальною*, розуміють шкалу, значення якої можна порівнювати. Але з порядкової шкали можна зробити лише висновок, чи більше А від Б, чи навпаки менше. Висновки, на скільки одиниць більше або менше, зробити неможливо. Відстань між об’єктами, виміряна в шкалі порядку, не обов’язково є рівною.

Наприклад, коли учня просять розташувати цінності зі списку за мірою їх важливості, отримані в такому вимірюванні дані будуть ранговою, порядковою шкалою. Найбільш поширеною порядковою шкалою в системі освіти є п'ятибальна. Вона складається з якісних категорій, таких як "незадовільно", "задовільно", "добре" та "відмінно". Ці категорії можна легко ранжувати за ступенем зростання, що відповідає підвищенню рівня знань та компетентності. Утім, ця шкала не є інтервальною, тобто не дозволяє визначити точну різницю між сусідніми значеннями.

Метричні шкали, також відомі як *числові*, є найбільш придатними для статистичного аналізу. Особливість шкал цього типу полягає в тому, що властивості об’єктів вони вимірюють на основі спеціально розроблених *еталонних значень*, які приймаються за одиницю в шкалі. Використання еталонних значень дозволяє кількісно описати певний об’єкт чи явище (наприклад, у скількох еталонах довжини, сантиметрах, описується зріст тої або іншої людини). Метричні шкали дозволяють кількісно порівнювати виміряні об’єкти один з одним. У шкалах такого типу певне значення приймається як нульове. С. Стівенс запропонував розрізняти два види метричних шкал: *інтервальні* та *шкали відношення (пропорційні)*. Підставами для такої класифікації є використання нульового значення, яке може бути відносним (умовним) або абсолютним (повна відсутність властивості, яка вимірюється).

Шкала інтервалів використовує еталонне значення. Наприклад, умовний рівень зміни температури, який отримав назву градус. Але в шкалі інтервалів за нуль приймається відносне значення. Наприклад, у шкалі температури Цельсія за нульове значення прийнята температура переходу води з твердого у рідкий стан. Шкала інтервалів дозволяє зробити висновок, на скільки еталонних значень один об’єкт відрізняється від іншого.

У *шкалі відношення* використовується абсолютне нульове значення, і це дозволяє робити висновок, не лише на яку кількість еталонів відрізняється один об’єкт від іншого, а й у скільки разів.

Основна відмінність між рівнями вимірювання визначається спектром допустимих математичних операцій, які можна застосовувати до отриманих даних. Номінальний рівень обмежується класифікацією об'єктів за категоріями без можливості впорядкування, що дозволяє лише підраховувати частоти та визначати моду. Порядковий рівень, зберігаючи категоріальний характер даних, вводить можливість ранжування, проте без фіксованих інтервалів між значеннями, що розширює аналітичні можливості до визначення медіани. Інтервальний рівень, крім упорядкованості, забезпечує рівномірність інтервалів, але не має абсолютного нуля, що дозволяє використовувати арифметичні операції додавання та віднімання, а також обчислювати середнє значення. Найвищий рівень – рівень відношення – відрізняється наявністю абсолютного нуля, що дає змогу не лише вимірювати різницю між значеннями, а й аналізувати їхні співвідношення через множення та ділення, що робить його найбільш універсальним для кількісного аналізу.

1.4. Закон нормального розподілу

Змінна, яку потрібно дослідити в генеральній сукупності, вимірюється в певній шкалі. Значення за цією шкалою трапляються з певною частотою. Так, є значення, які трапляються найчастіше, та значення, які трапляються дуже рідко. У генеральній сукупності частоти значень змінної зазвичай розподілені у вигляді дзвоноподібної кривої, яка має назву *нормального розподілу,* або розподілу Гауса (див. рис. 1.3).

Рис. 1.3. Розподіл частот (полігон) для зросту 8585 дорослих людей, народжених в Англії у 19 столітті.

На рис. 1.3 зображено частоту, з якою траплявся певний ріст людей, народжених в Англії у XIX столітті. Якщо намалювати лінію між частотами кожної дискретної *події* (якою є певний інтервал росту), вийде дзвоноподібна крива, яка називається нормальним розподілом.

Як можна побачити з рис. 1.3, певні події (тобто зареєстроване за шкалою значення) трапляються з різною частотою. Так, найчастіше спостерігається ріст, який становить приблизно 170 сантиметрів. Так, у наближеному до нього інтервалі росту було зафіксовано 1329 подій. Це означає, що в 1329 людей з 8585 обстеженої вибірки був зафіксований означений інтервал, що становить приблизно 15,5% усієї вибірки. Водночас ріст у межах 152 сантиметрів трапляється рідко, у 41 людини, що становить менше від 0,5% всіх обстежених.

Перший опис закономірностей, які згодом здобули назву нормального розподілу, здійснив А. Кетле в роботі "Досвід соціальної фізики" (1835) [1]. Учений описав важливу закономірність, властиву соціальним процесам (тривалість життя, вік вступу у шлюб, народжуваність): усі вони

демонстрували стійку тенденцію до концентрації навколо певного середнього значення.

А. Кетле показав, що екстремальні значення соціальних показників трапляються значно рідше, ніж значення, близькі до середнього. При цьому відхилення в бік збільшення і зменшення проявлялися з приблизно однаковою імовірністю. Цю властивість соціальних явищ він схарактеризував як "закон відхилення від середньої величини". Подальші дослідження, особливо роботи Ф. Гальтона, продемонстрували універсальність цього явища. Було встановлено, що будь-які масові вимірювання антропометричних чи демографічних характеристик у популяціях людей дають характерний графічний розподіл, який має дзвоноподібну форму.

Форму таких розподілів можна описати математичною формулою, яку запропонував у 19 столітті математик А. де Муавр.

$$f(x_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i - M)^2 / 2\sigma^2}$$

Рис. 1.4. Формула нормального розподілу

У формулі символи означають таке: *$f(x_i)$ – висота підйому кривої; e – основа натурального логарифму (приблизно 2,718); π – число «пі» (приблизно 3,14); M і σ – середнє і стандартне відхилення для змінної x_i, які визначають положення кривої на числовій осі й задають її розмах.*

Вважається, що коли варіація певної ознаки обумовлена впливом численних факторів, її розподіл у генеральній сукупності набуває форми класичної гаусової кривої (нормального розподілу).

Отже, вважається, що розподіли в генеральній сукупності найчастіше наближаються до кривої нормального розподілу. Завдяки цьому вона використовується в теорії математичної статистики для того, щоб передбачувати частоту певної події.

1.5. Вибірковий метод, обчислення мінімального об’єму вибірки для підтвердження/спростування наукових гіпотез на заданому рівні достовірності

Виходячи з закону нормального розподілу, для того щоб вибірка адекватно відображала генеральну сукупність, її слід сформувати випадковим чином. Випадковий спосіб дозволяє зібрати вибірку, у якій точно відображаються частоти, з якими трапляються події в генеральній сукупності, що дозволяє сформувати *репрезентативну вибірку*.

Репрезентативною вибіркою називають таку вибірку, яка відтворює структуру генеральної сукупності за ключовими параметрами, тобто основні ознаки генеральної сукупності представлено в репрезентативній вибірці з тією самою частотою та в тих самих пропорціях.

Під *випадковим (імовірнісним, рандомним) способом утворення вибірки* мається на увазі, що кожна подія має однакову (рівнозначну) ймовірність потрапити у вибірку.

На практиці це означає, що кожна особа із визначеної генеральної сукупності (яка є носієм певної змінної, певного значення за шкалою, на основі якої вимірюються її характеристики) має однакову ймовірність потрапити у вибірку, що формується. У такому разі можна очікувати, що отримана вибірка відображатиме частотний розподіл змінної в генеральній сукупності.

У роботі Г. Шварц досить ретельно розглядаються принципи побудови ймовірнісних вибірок [29, с. 13]. Так, для їхнього формування потрібно мати список всіх учасників, які складають генеральну сукупність, та визначитись з кількістю осіб (склад вибірки), яких необхідно відібрати. Так, наприклад, нехай генеральна сукупність (обмежена за віком, статтю та професією) становить 1000 осіб. З цієї сукупності нам потрібно обстежити 140 осіб, що становить 14%.

Для здійснення випадкового відбору особам, які є в списку генеральної сукупності, привласнюється номер від першого до останнього учасника. Після цього використовується *таблиця випадкових чисел* (див. рис. 1.5)

73735	45963	78134	63873
02965	58303	90708	20025
98859	23851	27965	62394
33666	62570	64775	78428
81666	26440	20422	05720
15838	47174	76866	14330
89793	34378	08730	56522
78155	22466	81978	57323
16381	66207	11698	99314
75002	80827	53867	37797
99982	27601	62686	44711
84543	87442	50033	14021
77757	54043	46176	42391
80871	32792	87989	72248
30500	28220	12444	71840

Рис. 1.5. Фрагмент таблиці випадкових п’ятизначних чисел

На рис. 1.5. наведено випадкові п’ятизначні числа, згенеровані в 1950-х роках на спеціально сконструйованому комп’ютері [20]. Для формування вибірки дослідник має у будь-якому стовпці обрати перші цифри, які відповідають його генеральній сукупності.

Наприклад, якщо ми обрали перший стовпчик для генеральної сукупності 1000 осіб, то у випадкову вибірку ввійдуть опитані під номерами: 737; 29; 988; 336; 816; 158 і т. д.

У наш час для утворення випадкової вибірки рідко звертаються до таблиць випадкових чисел, оскільки майже кожен пакет статистичної обробки даних має генератор випадкових чисел. Так, у SPSS утворити випадкову вибірку в 140 осіб зі списку 1000 осіб достатньо легко та швидко.

Інший спосіб формування випадкової вибірки – це *систематичний відбір*. Для його здійснення спочатку визначається крок або проміжок відступу в списку. Для визначення проміжку загальна кількість об'єктів генеральної сукупності ділиться на кількість вибірки. У нашому випадку проміжок між об'єктами становить 7,14:

$$\frac{\text{Загальна кількість осіб у сукупності}}{\text{Кількість осіб у вибірці}} = \frac{N}{n} = \frac{1000}{140} = 7{,}14$$

Перший номер обирається випадково, за допомогою таблиці випадкових чисел від 1 до 9. Наприклад, нехай це число становить 4. Тоді, додаючи до першого випадкового числа визначений проміжок, ми формуємо випадкову вибірку.

Так, у цьому випадку нашими номерами ймовірнісної вибірки будуть: 4, 11, 18, 25, 32, 39…

Систематичний відбір дозволяє сформувати випадкову вибірку. Винятком є випадки, коли список генеральної сукупності циклічний. Наприклад, це список родин з дітьми. У цьому списку члени сім'ї записуються в певній послідовності: тато, матір, дитина. Коли інтервал оцінювання збігається з циклом, отримана вибірка може бути не ймовірною. Наприклад, у вибірку потрапляють переважно чоловіки. Через це для циклічних списків метод систематичного відбору не застосовується.

Часто під час формування вибірки використовується принцип стратифікації для того, щоб забезпечити її репрезентативність. Під стратифікацією мається на увазі, що в утвореній вибірці буде досягнуто співвідношення груп таке саме, як і в генеральній сукупності. Для цього на основі попередніх досліджень вивчається співвідношення груп у генеральній сукупності (наприклад, за даними переписів населення, співвідношення жінок та чоловіків становить 55% до 45% відповідно). Після того як визначено необхідний загальний обсяг вибірки, на основі пропорцій (у генеральній сукупності) визначається потрібна кількість кожної групи у вибірці.

Наприклад, у генеральній сукупності жінок – 55%, а чоловіків – 45%. Якщо запланований обсяг вибірки – 1000 осіб, то в стратифікованій за статтю вибірці має бути 550 жінок та 450 чоловіків. Після визначення необхідного обсягу кожної групи у вибірці представники кожної групи обираються у вибірку рандомним чином з генеральної сукупності. Вибірка, сформована таким чином, називається *стратифікованою імовірнісною вибіркою.*

Серійна вибірка (інші назви: кластерна, гніздова). Цей метод утворення вибірки передбачає відбір не окремих елементів досліджуваної сукупності, а певних груп (кластерів), що утворюють структурні одиниці генеральної сукупності. Відібрані кластери формуються випадковим чином з метою забезпечення репрезентативності, після чого всі елементи в межах кожної відібраної групи підлягають повному обстеженню. Такими кластерами можуть бути школи, окремі класи. Існують й інші способи формування вибірки [16; 29; 6, с. 31-32].

Потрібно відзначити, що *події* (цим терміном позначаються певні конкретні значення змінної, що вивчається), присутні в генеральній сукупності, мають нерівнозначну ймовірність потрапити у випадкову вибірку. Чим частіше певна подія трапляється в генеральній сукупності, тим її ймовірність потрапити у вибірку є вищою.

У межах *теорії імовірності* математики розробили формули оцінки ймовірності потрапляння подій у випадкову вибірку, а також правила складання, множення, перестановки та поєднання ймовірностей [7; 20, с. 182].

Розглянемо найпростіший випадок, коли можливими є лише дві події, при цьому поява першої події (А) виключає імовірність появи другої події (В).

Наведемо приклад обчислення ймовірності вибору однієї з 9 куль з урни, який наводиться в роботі Дж. Гласс [20, с. 180]. Нехай генеральна сукупність утворена 9 кулями. З них 6 білих (подія А) та 3 чорні (подія В). Якою є імовірність вибору білої кулі, коли спосіб вибору куль є випадковим?

Для розв’язання цього прикладу слід користуватись визначенням: і*мовірність події А є відношення всіх подій А (у генеральній сукупності) до*

загальної кількості всіх можливих подій [20, с. 180]. У статистиці ймовірність позначається символом Р. Отже, маємо таку формулу:

$$P(A) = \frac{\text{Можлива кількість подій A}}{\text{Загальна кількість всіх можливих подій}} = \frac{6}{9} = \frac{2}{3}$$

Так, загальна кількість всіх можливий подій становить 9 (усього в урні є 9 куль). Загальна кількість всіх можливих подій А (імовірність вибору білої кулі) становить 6. З формули видно, що ймовірність вибору білої кулі становить 2/3. Це означає, що найімовірніше, за випадкового вибору (коли куля після вибору буде повертатись в урну) у трьох виборах буде обрано дві білі кулі та одну чорну. Утім, конкретний вибір куль може відрізнятись від теоретичної імовірності. Вона проявить себе за великої кількості повторів. Наприклад, при 100 повторах співвідношення вибору білих куль до чорних буде наближатись до 2/3.

Якщо після першого вибору біла куля не буде повертатись в урну, то ймовірність повторного випадкового вибору білої кулі змінюється. Така ймовірність буде становити 5/8 (приблизно 0,6), оскільки в урні залишаться 5 білих куль, за 8 можливих подій. Отже, імовірність, що буде витягнено білу кулю, буде трохи вищою за ймовірність події В (витягнення чорної кулі).

Отже, для формування репрезентативної вибірки важливою вимогою є її ймовірнісний характер, що забезпечує однакові шанси всіх подій у генеральній сукупності потрапити у вибірку. Тоді як імовірність потрапляння подій, які є в генеральній сукупності, у випадкову вибірку не є однаковою та пов’язана з їх частотою присутності в генеральній сукупності. Ці закономірності забезпечують результат: імовірнісна вибірка достатнього обсягу відображає частоту подій (певних значень) в генеральній сукупності.

Необхідний *розмір випадкової вибірки* обчислюється, виходячи з того, наскільки точними мають бути результати перевірки статистичних гіпотез, якою є допустима ймовірність помилки.

Спосіб обчислення обсягу вибірки відрізняється залежно від того, що є предметом дослідження: конкретна подія (у шкалі найменувань) чи значення за метричною шкалою. Наприклад, конкретною подією (у шкалі найменувань) може бути однозначна відповідь, яка отримана від опитаного. Для обчислення точної кількості вибірки для подій у номінальній шкалі застосовується формула, наведена на рис. 1.6:

$$n = \frac{\frac{Z^2 pq}{\Delta^2}}{1 + \frac{\frac{Z^2 pq}{\Delta^2} - 1}{N}}$$

Рис. 1.6. Обчислення обсягу вибірки для дискретної події

У формулі символи означають таке: *n – обсяг вибірки; N – обсяг генеральної сукупності; Z – коефіцієнт обраного дослідником довірчого інтервалу, за довірчого інтервалу 0,95 (95%) приймається значення 1,96; p – доля опитаних з наявною ознакою: за невідомої імовірності застосовується значення 0,5 (тобто ознака наявна в 50% опитаних), яка дає найвище значення вибірки; q = 1-p – частка опитаних, у яких відсутня ознака. Якщо не відомо, у якого відсотка ознака наявна та відсутня, також приймається значення 0,5; Δ - максимальна помилка вибірки. За 95% інтервалу приймається за 0,04.*

Наприклад, досліднику з високою долею імовірності необхідно виявити частоту студентів-психологів третіх курсів, які у вільний час ходять до кінотеатру на перегляд фільму. Анкета передбачає деякі варіанти проведення вільного часу, але ми будемо оцінювати лише події «піду в кінотеатр» проти всіх інших подій.

За запропонованою формулою (рис. 1.6) для обчислення обсягу вибірки нам потрібно знати лише приблизну кількість генеральної сукупності та ймовірність події, яка досліджується. Припустимо, що генеральна сукупність усіх студентів-психологів 3-го курсу в Україні становить 50 000 осіб. Наступним кроком нам потрібно дізнатись, наскільки часто студенти третього

курсу проводять вільний час в кінотеатрі – для того, щоб обчислити ймовірність цієї події. Найвища вибірка потрібна за ймовірності 50% (p = 0,5). Якщо ж імовірність події буде меншою, наприклад p = 0,3, обсяг потрібної вибірки буде меншим.

Проводячи обчислення за формулою (для ймовірності події p = 0,5), наведеною на рис. 1.6, отримаємо обсяг вибірки n = 593. Це означає, що для того, щоб з точністю 95% (0,05% помилки) дізнатись, скільки студентів-психологів 3-го курсу ходять у вільний час у кінотеатр, потрібно дослідити випадкову вибірку обсягом 593 особи.

Г. Шварц наводить також інші формули (див. рис. 7) для обчислення обсягу вибірки, якщо відомий обсяг генеральної сукупності, які дають ті ж самі результати для нашого прикладу [29, с. 59]:

$$n = \frac{u_\alpha^2 PQN}{e_p^2(N-1) + u_\alpha^2 PQ}$$

Рис. 1.7. Обчислення об’єму вибірки для дискретної події

У формулі символи означають: *n – обсяг вибірки; N – обсяг генеральної сукупності;* U_a *– коефіцієнт обраного дослідником довірчого інтервалу, за довірчого інтервалу 0,95 (95%) приймається значення 1,96; P – доля опитаних з наявною ознакою. За невідомої імовірності застосовується значення 0,5 (тобто ознака наявна в 50% опитаних), яка дає найвище значення вибірки; Q = 1-p – доля опитаних, у яких відсутня ознака: якщо невідомо, у якого відсотка ознака наявна та відсутня, також приймається значення 0,5;* e_p *- максимальна помилка вибірки: за 95% інтервалу приймається за 0,04.*

Для випадку, який наведено вище, обчислення за наведеною на рис. 7 формулою показує ідентичний обсяг вибірки n = 593. Утім, формула, запропонована Г. Шварц, є більш зручною для обчислень.

Також можна проводити обчислення вибірки лише на основі ймовірної помилки, вважаючи, що обсяг генеральної сукупності є нескінченним, за формулою, наведеною на рис. 1.8 [29, с. 60]:

$$n = \frac{Z^2 pq}{\Delta^2}$$

Рис.1.8. Обчислення обсягу вибірки для дискретної події за нескінченної генеральної сукупності

Г. Шварц попереджає, що підрахунок за вказаною формулою дає великий обсяг вибірки, що не має сенсу, якщо генеральна сукупність мала. Зробімо обчислення для нашого випадку:

$n = 1{,}96^2 * 0{,}5 * 0{,}5 / 0{,}04^2 = 0{,}9604 / 0{,}0016 = 600$

Отже, обчислення за точною формулою та формулою для нескінченної генеральної сукупності для встановленої нами ймовірності похибки та ймовірності виникнення події 0,5 дає зіставні результати.

За ймовірності виникнення події 0,3 (або 0,7) потрібний розмір вибірки зменшиться та буде становити 504 особи. Так само розмір вибірки буде зменшуватись у разі збільшенні рівня можливої похибки, наприклад, з 0,04 до 0,049. У такому разі обсяг необхідної вибірки (за ймовірності події p = 0,3) становитиме 336 осіб.

Обчислення обсягу випадкової вибірки для виявлення середнього значення генеральної сукупності за інтервальною шкалою (яка, наприклад, має значення від 0 до 9) здійснюється за іншою формулою. Для цього обчислення слід знати інформацію про дисперсію змінної, що вивчається в генеральній сукупності.

Якщо інформації про дисперсію змінної в генеральній сукупності немає в доступній літературі, Г. Шварц рекомендує проведення пілотажного дослідження на випадковій вибірці обсягом 30 осіб [29, с. 78]. Отримані на цій

невеликій вибірці значення дисперсії пропонується вважати такими, що гіпотетично належать генеральній сукупності. Далі на основі цих припущень треба здійснювати обчислення мінімального обсягу більшої вибірки для отримання достовірних результатів.

Для обчислення обсягу вибірки із заданими параметрами достовірності пропонується формула, представлена на рис. 1.9 [29, с. 79]:

$$n = \frac{N}{1 + \left(\frac{e_{\bar{x}}}{u_\alpha s}\right)^2 N}$$

Рис. 1. 9 Обчислення обсягу вибірки для з допустимим рівнем помилки середнього генеральної сукупності для інтервальної шкали

У формулі символи означають таке: *N – обсяг генеральної сукупності;* e_x *– гранична помилка вибірки; у цій формулі гранична помилка вибірки задається дослідником (наприклад, гранична помилка може бути 0,5 бала для дискретної інтервальної шкали з кроком в 1 бал);* u_a *– коефіцієнт обраного дослідником довірчого інтервалу, за довірчого інтервалу 0,95 (95%) приймається значення 1,96; s – квадратний корінь з дисперсії, тобто стандартне відхилення вибірки (або дані про s в генеральній сукупності).*

Під час визначення обсягу вибірки використовується термін *гранична помилка вибірки*, яка становить половину довірчого інтервалу для середнього генеральної сукупності. В формулі, на рис. 1.9, чим меншою є гранична помилка тим більш точно можна здійснювати висновки про середнє генеральної сукупності. Гранична помилка вибірки впливає на об’єм вибірки: чим менше помилка тим має бути більший об’єм вибірки.

Наприклад, є генеральна сукупність дівчат – студенток психологічних факультетів бакалаврського рівня освіти (1-4 курси). Припустимо, що ця сукупність по всій Україні становить 150 000 осіб. Нам потрібно обчислити обсяг вибірки для створення нормативних значень для фактору А опитувальника 16PF (форма С) Р. Кеттелла. На випадковій вибірці 34 студенток-психологів було встановлено, що середнє вибіркове становить (х)

6,69, стандартне відхилення (s) = 1,88. Граничну помилку вибірки встановлюємо в 0,4 бала за ймовірної помилки 0,05 (z = 1,96 або приблизно 2).

У такому разі мінімальний обсяг випадкової вибірки згідно з формулою на рис. 1.9 буде становити: n = 89 осіб. Якщо гранична помилка буде задана меншою за 0,4 бала, то кількість осіб у вибірці буде збільшуватись.

Ще один спосіб обчислення обсягу вибірки для дослідження середнього значення інтервальної шкали ґрунтується на формулі *варіації вибірки* (v) та задає допустиму *варіацію середнього значення* (v_x). Спосіб називається «обчислення обсягу вибірки за заданої відносної точності». Цей спосіб обчислення вибірки вважається прийнятним для проведення різноманітних порівнянь з середнім значенням. Прийнятним рівнем коефіцієнта варіації середнього є $v_x < 0,05$ (за бажаного $v_x < 0,01$). У представленій нижче формулі (рис. 1.10) варіація середнього задається через відносну граничну помилку.

У формулу підставляється значення варіації, яке отримується з невеликої випадкової вибірки під час пілотажного дослідження (n = 30), та заданий (бажаний) рівень варіації середнього.

Варіація вибірки або *коефіцієнт варіації* обчислюється за формулою [29, с. 72]:

$$v = s/x$$

У формулі символи означають: *s – стандартне відхилення; x – середнє генеральної сукупності (вибіркове середнє, яке екстраполюється).*

Формулу для обчислення обсягу вибірки за заданої відносної точності представлено на рис. 1.10 [29, с. 81]:

$$n = \frac{u_\alpha^2 \, v^2}{e_{xr}^2 + \frac{u_\alpha^2 \, v^2}{N}}$$

Рис. 1.10. Обчислення обсягу вибірки для метричної шкали за заданої відносної точності

У формулі символи означають таке: u_a *– коефіцієнт обраного дослідником довірчого інтервалу, за довірчого інтервалу 0,95 (95%) приймається значення 1,96; v – варіація вибірки;* e_{xr} *– відносна гранична помилка (допустимими є значення від 0,01 до 0,04): у цій формулі гранична помилка вибірки задається дослідником, як допустима ймовірність помилки; N – обсяг генеральної сукупності.*

Для наведеного вище прикладу зі шкалою А опитувальника Р. Кеттелла v = 1,88/6,69 = 0,281. Як і в попередньому обчисленні, довірчий інтервал (u_a) встановлюємо на рівні 0,05. Тоді u_a = 1,96 (приблизно 2), відносну граничну помилку (e_{xr}) встановлюємо на рівні 0,04. Генеральна сукупність становить 150 000. Провівши обчислення за формулою на з рис. 1.10, отримаємо обсяг вибірки: n = 197.

Класифікація вибірок за обсягом передбачає їх поділ на *малі* та *великі*, при цьому критерії розмежування мають умовний характер. *Малими* вважаються вибірки з кількістю елементів $n \leq 30$, які зазвичай застосовуються для статистичного контролю вже вивчених характеристик сукупностей. До категорії *великих* відносять вибірки з $n > 200$, тоді як *середні вибірки* – в інтервалі $30 \leq n \leq 200$. Великі вибірки потрібні для виявлення невідомих параметрів та властивостей досліджуваної сукупності, оскільки забезпечують вищу точність і надійність статистичних висновків.

Вважається, що вибірка обсягом 30 осіб може за певних умов характеризувати генеральну сукупність [29]. Дж. Гласс вважає, що для адекватного відображення частот у генеральній сукупності потрібна вибірка обсягом 100 осіб [20]. З іншого боку, обсяг вибірки, як уже було зазначено, обчислюється, виходячи не лише з її здатності відображати частоти

генеральної сукупності, а й із допустимої помилки при оцінці параметрів генеральної сукупності.

Згідно з *вибірковим підходом,* застосування математико-статистичних процедур реалізується через чітку послідовність дій.

По-перше, з генеральної сукупності, параметри якої підлягають аналізу, здійснюється відбір репрезентативної імовірнісної вибірки оптимального обсягу, що визначається цілями дослідження.

По-друге, шляхом експериментальних процедур, спостережень та вимірювальних операцій щодо об'єктів вибірки формується масив емпіричних даних.

На третьому етапі за допомогою методів описової статистики здійснюється первинна обробка отриманих даних, результатом чого є розрахунок вибіркових статистик.

На завершальному етапі застосування методів статистичного висновку дозволяє на підставі вибіркових статистик оцінити параметри та виявити закономірності, що характеризують властивості генеральної сукупності в цілому.

1.6. Наукові гіпотези

Першим етапом дослідження є формулювання проблеми. Складнощі, які виникають у людини в буденному житті, які вона намагається вирішити, можна розглядати як спрощений приклад наукової проблеми. На відміну від життєвих труднощів, наукова проблема визначається в межах термінології конкретної дисципліни. Під час наступного кроку наукові терміни повинні бути операціоналізованими. Це означає, що має бути описано конкретні, доступні для вимірювання прояви явищ, позначених поняттями.

Під *науковою проблемою* розуміють певне невирішене завдання, яке має теоретичне або практичне значення. Це може бути невирішене прикладне завдання, дефіцит знань чи їх суперечність тощо. Розв’язання проблеми

передбачає здійснення відкриття, яке має велике теоретичне чи практичне значення. Наукова проблема вміщує в себе наукові теми та питання, які потрібно вирішити для її розкриття.

Постановка проблеми має низку етапів: 1) аналіз наявних суперечностей та виділення невирішеної області; 2) розбиття проблеми на *теми та наукові питання*, які необхідно вирішити, розробка структури проблеми; 3) аналіз актуальності розв'язання сформульованої проблеми для практичної діяльності.

Постановка проблеми тягне за собою формулювання гіпотези про існування зв'язку між явищами, які описуються поняттями (термінами) науки. Гіпотеза про зв'язок між явищами є шляхом до розв'язання проблеми, до наукового відкриття.

Наукова гіпотеза являє собою теоретично обґрунтоване припущення, яке ще не пройшло емпіричну перевірку і, відповідно, не було остаточно підтверджене або спростоване.

Для того, щоб гіпотеза була науковою, вона має відповідати:

– *принципу фальсифікації* – згідно з яким, мають існувати факти (результати спостережень, експерименту тощо), на основі яких гіпотезу можна спростувати;

– *принципу верифікації* — згідно з яким, можна отримати певний результат, який підтверджує гіпотезу. При цьому підтвердження гіпотези не виключає її можливого спростування в майбутніх емпіричних дослідженнях.

До наукових гіпотез відносять теоретичні, експериментальні та статистичні. *Теоретичні гіпотези* стосуються зв'язку між поняттями дослідження. Вони можуть формулюватись як запитання або як твердження про існування зв'язку. *Емпірична гіпотеза* необхідна для організації самого дослідження. Вона містить у собі припущення про існування зв'язку між конкретними змінними. Для формулювання емпіричної гіпотези теоретичні поняття має бути операціоналізовано, тобто переведено на рівень змінних, між конкретними проявами яких допускається зв'язок.

Статистична гіпотеза необхідна на етапі математичної інтерпретації даних емпіричних досліджень. Статистична гіпотеза містить два поняття: H_0 – нуль-гіпотеза та H_1 (альтернативна) – гіпотеза.

H_0 – гіпотеза перевіряє відповідність емпіричних вибіркових даних певному теоретичному розподілу, на перевірку якого спрямований критерій (наприклад, середні двох груп на рівні генеральної сукупності належать до можливих варіацій генерального середнього значення, кореляція в генеральній сукупності відсутня, вибірка відповідає закону нормального розподілу тощо).

H_1 – гіпотеза приймається в разі спростування нуль-гіпотези. Альтернативна гіпотеза означає, що емпіричні вибіркові дані перевищують встановлені довірчі інтервали для теоретичного розподілу, який перевірявся критерієм, і отримані вибіркові дані не належать до цього теоретичного розподілу (з відомою імовірністю помилки: 5%, 1% або 0,1%). Наприклад, H_1 може бути такою: середні двох вибірок не рівні та належать до різних генеральних середніх; кореляція в генеральній сукупності наявна; вибіркові дані не підпадають під теоретичні частоти нормального розподілу тощо.

Статистичний висновок робиться стосовно генеральної сукупності, для якої обстежена вибірка репрезентативна. Наприклад, якщо емпірична гіпотеза припускає, що втома буде впливати на результати коректурної проби (увагу), то статистична гіпотеза перевіряє, чи будуть достовірні відмінності у двох групах досліджених (без втоми та втомлені) за кількістю помилок, швидкістю виконання завдання.

Статистична перевірка гіпотези дозволяє зробити судження про властивості генеральної сукупності, описати її. Завдання самого наукового дослідження полягає в тому, щоб описати потрібні властивості генеральної сукупності. Оскільки генеральна сукупність є умовно нескінченною, її опис можливий лише на основі вибіркового методу. Для примноження наукового знання, важливо після проведення емпіричного дослідження на ймовірнісній вибірці:

- описати залучену в дослідження вибірку, спосіб її формування та обсяг;

- порівняти результати дослідження з результатами дослідження інших авторів.

Для підтвердження або спростування експериментальної гіпотези зазвичай перевіряється велика кількість статистичних гіпотез (про відповідність даних теоретичному нормальному розподілу, про існування кореляції між двома змінними, про вплив на зв'язок між двома змінними третьої змінної тощо). Таким чином, експериментальна гіпотеза про зв'язок між змінними виступає як первинна, тоді як статистична гіпотеза є похідною від експериментальної, більш конкретизована та орієнтована на знайдення певних достовірних закономірностей між шкалами, використаними в дослідженні.

Збіг результатів багатьох вибіркових досліджень, здійснених для перевірки певної гіпотези та представлених у науковій літературі, підтверджує існування особливостей генеральної сукупності, яка вивчалась.

Діяльність дослідника з формулювання теоретичної гіпотези про зв'язок між двома явищами та операціоналізація теоретичної гіпотези на рівні змінних, уточнення чи спростування гіпотези вважається найкреативнішим етапом наукової діяльності, пов'язаним із загальною творчою здатністю дослідника.

Наукова гіпотеза, навіть у разі її тимчасового підтвердження, не набуває статусу остаточно прийнятої істини, залишаючись відкритою для подальших емпіричних перевірок і уточнень. Її формулювання виконує функцію систематизації дослідницьких припущень, забезпечуючи їх чітке, лаконічне і концептуально узгоджене представлення.

Статистична гіпотеза визначається як формалізоване припущення щодо параметрів або форми розподілу ймовірностей у генеральній сукупності, яке підлягає перевірці на основі наявних вибіркових даних. Залежно від рівня специфікації, статистичні гіпотези поділяються на прості та складні: проста гіпотеза повністю задає конкретний розподіл імовірностей, тоді як складна охоплює множину можливих розподілів, що відповідають заданим умовам.

Найчастіше під час перевірки статистичних гіпотез потрібно довести значущість відмінностей, оскільки ця інформація більш інформативна для пошуку нового, або констатувати наявність чи відсутність певного зв'язку. Перевірка гіпотез здійснюється за допомогою критеріїв статистичної оцінки відмінностей.

1.7. Статистичні критерії: загальне уявлення про перевірку статистичних гіпотез

Статистичний критерій являє собою формалізоване вирішальне правило, що дозволяє з високим ступенем достовірності приймати істинну гіпотезу та відхиляти хибну. Під статистичним критерієм також розуміють метод обчислення певного числового показника, який слугує основою для ухвалення рішення щодо гіпотези. Здебільшого *значущість емпіричних відмінностей* встановлюється за умови перевищення критичного порогу емпіричного значення критерію, хоча для окремих критеріїв застосовуються альтернативні правила, які детально описуються в їхній методологічній специфікації.

На основі спеціалізованих таблиць критичних значень визначається рівень статистичної значущості, якому відповідає отримане емпіричне значення, що дозволяє інтерпретувати результати дослідження в контексті заданого рівня α-помилки.

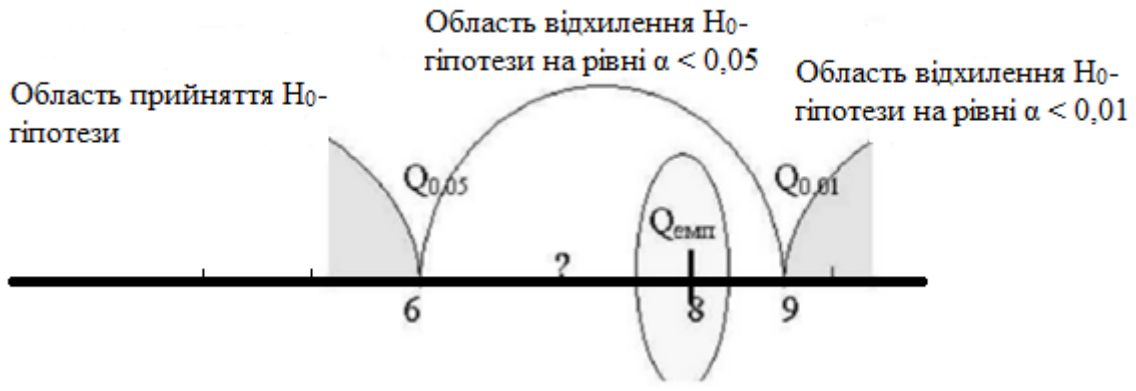

Рис. 1.11. Правило прийняття статистичних гіпотез

На рис. 1.11 відображено емпіричне значення статистичного критерію ($Q_{емп}$ = 8), яке є більшим за критичне значення критерію для рівня достовірності 0,05 ($Q_{0,05}$ = 6), але меншим для *критичного значення критерію* для рівня достовірності 0,01 ($Q_{0,01}$ = 9).

Правило відхилення H_0 і прийняття H_1.

Доти, поки рівень значущості не досягне p = 0,05, дослідник не має права відхилити нульову гіпотезу (наприклад, про відсутність відмінностей у середніх на рівні генеральної сукупності). Якщо емпіричне значення критерію дорівнює критичному значенню, відповідному $p \leq 0,05$ (або перевищує його), то – H_0 відхиляється з допустимою імовірністю помилки 5% та приймається H_1 – гіпотеза. Якщо емпіричне значення критерію дорівнює критичному значенню, відповідному $p < 0,01$ (або перевищує його), – H_0 відхиляється з допустимою імовірністю помилки 1% та приймається альтернативна гіпотеза H_1. Деякі науковці наголошують на необхідності використання більш жорсткого критерію для спростування H0 на рівні $\alpha<0,01$, однак поки процедура спростування гіпотез є незмінною.

Виняток становлять деякі непараметричні критерії, зокрема критерій знаків, критерій T Вілкоксона та критерій U Манна–Уїтні, для яких застосовуються зворотні правила інтерпретації критичних значень. Для спрощення процесу ухвалення рішення іноді використовують графічне представлення у вигляді «осі значущості».

Однією з ключових характеристик статистичного критерію є його *потужність*, тобто здатність виявляти реальні відмінності між порівнюваними групами. *Потужність критерію* визначає його ефективність у відкиданні нульової гіпотези, якщо вона є хибною. Потужність критеріїв перевіряється емпірично. До найбільш потужних відносять: t-критерій Стюдента, F-критерій Фішера, критерій χ^2 Пірсона (для таблиць 2x2 з поправкою Фішера). Менш потужними є непараметричні критерії, наприклад, U-Манна-Уїтні.

При прийнятті статистичних гіпотез можливі статистичні помилки, які здобули назву: *помилка першого роду* (А – помилка) та *помилка другого роду* (β – помилка) [12, с. 170].

Помилка першого роду полягає в помилковому відкиданні нульової гіпотези через помилковий висновок про те, що відмінності в генеральній сукупності існують. Альфа-помилка контролюється обраним рівнем достовірності. Так, за $p < 0,05$ імовірність помилкового висновку про те, що в генеральній сукупності є відмінності, наприклад між певними групами, становить 5%. Цей помилковий висновок пов'язаний з тим, що в 5% випадків може бути отримано вибіркові статистики, які некоректно описують генеральну сукупність.

Помилка другого роду (β – помилка) полягає в неправильному прийнятті нульової гіпотези, тобто судження про те, що в генеральній сукупності відмінності відсутні. Імовірність помилки другого роду висока, коли: 1) відмінності в генеральній сукупності існують, але невеликі, 2) за використання статистичного критерію з низькою потужністю; 3) за недостатнього обсягу вибірки.

1.8. Шляхи використання математичних методів у психологічних дослідженнях

У психологічній науці математичні методи відіграють важливу роль, що зумовлено низкою факторів.

По-перше, вони забезпечують структурованість, логічну впорядкованість і раціональність дослідницького процесу, сприяючи формалізації аналізу психічних явищ.

По-друге, математичні методи потрібні для обробки значних обсягів емпіричних даних, представлених у кількісній формі, а також для їх узагальнення та інтеграції в цілісну емпіричну модель досліджуваного феномену.

Основною метою використання статистичних методів у психології є підвищення достовірності інтерпретацій шляхом застосування ймовірнісних моделей для оцінки емпіричних закономірностей.

У психології виділяють кілька основних напрямів застосування статистичних методів (див. рис. 1.12).

Рис. 1.12. Напрями застосування математичних методів у психології

По-перше, це – *описова статистика*, яка охоплює процедури групування, табулювання, графічного представлення та кількісної оцінки даних, що дозволяє структурувати емпіричну інформацію.

По-друге, це *теорія статистичного висновку*, яка використовується для екстраполяції закономірностей, виявлених у вибірці, на певну генеральну сукупність, що забезпечує узагальнення результатів.

По-третє, це *теорія планування експериментів,* спрямована на розробку експериментальних планів (дослідницьких дизайнів), які дозволяють виявляти та перевіряти причинно-наслідкові зв’язки між психологічними змінними.

Серед найбільш поширених статистичних методів у психології слід виокремити кореляційний, регресійний та факторний аналіз.

Історія використання математичних методів у психології демонструє періодичні коливання: від їхньої абсолютизації та вимог обов’язкового застосування статистичних методів до їх повного ігнорування в емпіричній практиці. Загалом у наукових дослідженнях вважається доцільним збереження рівнозначності теоретичного аналізу та дослідницьких методів, які дозволяють отримати емпіричні дані для подальшої обробки методами математичної статистики. Така рівнозначність має ґрунтуватися на принципі змістової та процедурної відповідності між досліджуваним феноменом і обраними методами емпіричного дослідження. Статистичний аналіз дає змогу кількісно описати залежності між явищами, але не забезпечує їх змістової інтерпретації.

Таким чином, дотримання принципів наукової організації психологічного дослідження, зокрема узгодження методів із предметом і метою дослідження, є необхідною умовою для забезпечення наукової продуктивності, уникнення методологічних помилок та підвищення ефективності емпіричних процедур.

Лекція 2. Основи поняття математичної обробки психологічних даних

Мета: розглянути описову статистику як галузь статистики. Ознайомити студентів зі способами обробки первинних даних емпіричного дослідження, зокрема впорядкування та табулювання даних, групування частот, створення діаграми розподілу, гістограми та полігону. Розглянути способи опису вибірки за допомогою квантилів, візуалізації вибіркового розподілу за допомогою коробчастого графіку. Сформувати уявлення про основні статистики для опису вибірки, статистики мір центральної тенденції та міри розсіювання. Сформувати розуміння, що таке нормальний розподіл та які його закономірності, способи перевірки форми вибіркового розподілу.

Основні поняття: описова статистика, дані, впорядкування, ранжування, табулювання даних, частота події, групування частот, гістограма, полігон, коробчастий графік («ящик з вусами»), міри центральної тенденції, мода, медіана, середнє арифметичне, міри мінливості, варіаційний розмах, дисперсія та стандартне відхилення, закон нормального розподілу, z-перетворення, асиметрія, ексцес.

План

1. Описова статистика.
2. Міри центральної тенденції
3. Міри мінливості
4. Нормальний розподіл, характеристики форм розподілу.

2.1. Описова статистика. Табулювання даних, квантилі, графічне представлення розподілу частот

Описова статистика становить окрему галузь статистичного аналізу, яка спрямована на систематизацію емпіричних даних, вона займається збором, упорядкуванням, узагальненням та графічним представленням даних. Її функціональне призначення полягає у виявленні базових характеристик

сукупності отриманих даних, що дозволяє сформувати загальне уявлення про структуру вибірки та тенденції в ній. Основна мета описової статистики полягає в тому, щоб забезпечити наочне й доступне представлення даних, придатне для подальшої інтерпретації та формулювання висновків.

Для зручності опрацювання отриманих у дослідженні *даних* їх прийнято впорядковувати, представляти у вигляді частот, групувати, графічно репрезентувати та описувати вибірку на основі квантилів. У статистиці під терміном *дані* розуміють первинні (необроблені) результати вимірювань.

Найпершим кроком для *впорядкування даних* є сортування отриманих результатів за мірою їхнього зростання. Таке сортування є важливим кроком для привласнення рангових значень, обчислення частот подій, визначення медіани тощо.

Після того як дані впорядковано від найменшого значення до найвищого, можна надати їм рангові значення та обчислити частоти подій. В табл. 2.1 представлено впорядкований за зростанням список опитаних за шкалою «Посттравматичне зростання»

Таблиця 2.1

Результати опитування вибірки за шкалою посттравматичного зростання, впорядковані за зростанням

Порядковий номер	Досліджений	Результати	Ранг	Значення, подія	Частота	Сумарний відсоток
1	**2**	**3**	**4**	**5**	**6**	**7**
1	Учасник 22	16,00	1,0	16,00	1	2,4
2	Учасник 37	19,00	2,0	19,00	1	4,9
3	Учасник 38	20,00	3,5	20,00	2	9,8
4	Учасник 41	20,00	3,5	21,00	2	14,6
5	Учасник 17	21,00	5,5	23,00	1	17,1
6	Учасник 18	21,00	5,5	25,00	2	22,0
7	Учасник 2	23,00	7,0	28,00	1	24,4
8	Учасник 8	25,00	8,5	30,00	1	26,8
9	Учасник 39	25,00	8,5	31,00	1	29,3
10	Учасник 3	28,00	10,0	32,00	2	34,1
11	Учасник 4	30,00	11,0	33,00	3	41,5
12	Учасник 10	31,00	12,0	34,00	1	43,9
13	Учасник 12	32,00	13,5	35,00	1	46,3
14	Учасник 33	32,00	13,5	36,00	5	58,5
15	Учасник 13	33,00	16,0	37,00	1	61,0

16	Учасник 21	33,00	16,0	38,00	2	65,9
17	Учасник 29	33,00	16,0	39,00	2	70,7
18	Учасник 27	34,00	18,0	40,00	2	75,6
19	Учасник 36	35,00	19,0	42,00	1	78,0
20	Учасник 6	36,00	22,0	43,00	1	80,5
21	Учасник 9	36,00	22,0	44,00	1	82,9
22	Учасник 16	36,00	22,0	45,00	1	85,4
23	Учасник 19	36,00	22,0	47,00	3	92,7
24	Учасник 35	36,00	22,0	48,00	1	95,1
25	Учасник 32	37,00	25,0	49,00	1	97,6
26	Учасник 1	38,00	26,5	52,00	1	100,0
27	Учасник 34	38,00	26,5			
28	Учасник 5	39,00	28,5			
29	Учасник 40	39,00	28,5			
30	Учасник 30	40,00	30,5			
31	Учасник 31	40,00	30,5			
32	Учасник 23	42,00	32,0			
33	Учасник 15	43,00	33,0			
34	Учасник 24	44,00	34,0			
35	Учасник 20	45,00	35,0			
36	Учасник 7	47,00	37,0			
37	Учасник 14	47,00	37,0			
38	Учасник 25	47,00	37,0			
39	Учасник 28	48,00	39,0			
40	Учасник 11	49,00	40,0			
41	Учасник 26	52,00	41,0			

Таблицю побудовано аналогічно до даних, представлених Дж. Гласс [20, с. 32]. Представлення даних у таблиці називається *табулюванням.*

У табл. 2.1 видно, що після впорядкування даних за шкалою посттравматичного зростання (наведені в колонці 3) від найнижчого до найвищого значення змінився порядок учасників дослідження (колонка 2).

Після впорядкування легко побачити медіанне значення (36 балів), яке належить дослідженому з порядковим номером 21 (Учасник 9). *Медіанне значення* належить особі, яка в упорядкованій вибірці знаходиться посередині загальної кількості учасників.

У колонці 4 (табл. 2.1) наведено *рангові значення* опитаних. Ранжування – це надання значення в порядку зростання (або зменшення) певної якості. Дані у колонці 3 є метричними даними (шкала інтервалів) про міру вираженості посттравматичного зростання. Надаючи учасникам дослідження

ранг, ми переводимо дані зі шкали інтервалів у порядкову шкалу, яка відображає місце дослідженого в числовому ряду за його результатами тесту.

Рангове значення надається за певними правилами. Це порядковий номер особи у впорядкованому ряді певних значень (у нашому випадку – результатів за шкалою). Але якщо учасники мають однаковий результат, ранг визначається як середнє значення їх суми порядкових номерів. Наприклад, як можна побачити з табл. 2.1, учасники 38 та 41 отримали за шкалою «Посттравматичне зростання» по 20 балів. Їх порядкові номера у впорядкованому ряді даних становлять відповідно 3 та 4. У такому разі їм слід надати ранг на основі середнього значення їх порядкових номерів: 3+4 = 7/2 = 3,5. Як можна побачити з табл. 2.1 (в стовпчик 4), ці опитані й отримали ранг 3,5.

У подальшому аналізі табульований список даних скорочується до *розподілу частот* (див. табл. 2.1, колонки 5, 6). *Розподіл частот* – це впорядкований перелік значень за певною шкалою із зазначенням частоти, з якою це значення трапляється. Так, у табл. 2.1 видно значення за шкалою посттравматичного зростання (колонка 5) та частота цього значення (колонка 6). *Частотою* називають кількість осіб, у яких це значення трапляється, або кількість певних подій (кількість появи певного значення).

При подальшому узагальненні будується таблиця *згрупованих частот*. *Групування частот* корисне, коли є велика кількість значень за шкалою. Для групування частот утворюється *розряд оцінок*, який є одиницею для групування [20, с. 32]. Розряд оцінок є групою, яка вміщує в себе задану кількість оцінок. Побудова розподілу згрупованих частот є корисною для візуального відображення великого масиву даних і лежить в основі побудови гістограми.

Побудова таблиці згрупованих частот передбачає 4 етапи [20, с. 34].

1. Визначення загального *включеного розмаху* в середині вибірки. *Включений варіаційний розмах* є найпростішою мірою мінливості даних – це інтервал між максимальним і мінімальним значеннями ознаки, до якого

додається одиниця. Для даних, представлених у табл. 2.1, максимальне значення становить 52 бали; мінімальне – 16 балів. Відповідно включений варіаційний розмах становить 52 – 16 = 36+1 = 37 балів. Якщо значення за шкалою є дискретними (тобто передбачають ціле число без десятих), то границі значень зазвичай встановлюють у нуль цілих п’ять десятих до та після значення. Тобто, щоб при групуванні вмістити містити мінімальне значення в 16 балів, вважається, що воно варіюється в діапазоні від 15,5 до 16,5 балів.

2. Вибір інтервалу групування розрядів. Для групування частот мінімальна кількість розрядів становить 12, а максимальна 15. Якщо варіаційний розмах у вибірці невеликий (менше 16), то підстави для групування частот відсутні. У роботі Дж. Гласс представлено спосіб визначення ширини розрядів як оптимальна кількість між 12 та 15 розрядами [20, с. 34]. Тут обмежимось визначенням ширини мінімальної кількості розрядів з максимально можливим інтервалом. Для обчислення ширини розряду слід розмах розділити на кількість розрядів, що дозволяє визначити *інтервал розряду*: 37/12 = 3,08 балів. Отже, для групування даних табл. 2.1 мінімальний обсяг розряду має становити 3 оцінки.

3. Визначення границь розряду. Групування частот завжди починається зі значення, яке кратне інтервальному розряду і є меншим за мінімальне значення. 15 є числом, кратним виділеному нами інтервалу в 3 бали; також 15 є числом, меншим за мінімальне значення – 16 балів. Так, щодо нашої шкали, то для того, щоб вмістити мінімальне значення, будується інтервальний діапазон між 14,5 по 17,5, який вміщує у себе дискретну оцінку 16 балів. Для дробових значень інтервальний діапазон будується інакше [20, с. 34].

4. Табулювання *групованих частот*. Таке табулювання починається із занесення в таблицю границь розрядів з подальшим обчисленням частот, з якими трапляються значення в кожному розряді (див. табл. 2.2).

Таблиця 2.2

Табуляція згрупованих частот

Номер розряду	Границі розрядів	Значення, які потрапили в розряд	Частота подій в розряді	Відсоток
1	**2**	**3**	**4**	**5**
1	14,5 – 17,5	15, 16, 17	1	2,4
2	17,5 – 20,5	18, 19, 20	3	7,3
3	20,5 – 23,5	21, 22, 23	3	7,3
4	23,5 – 26,5	24, 25,26	2	4,9
5	26,5 – 29,5	27, 28, 29	1	2,4
6	29,5 – 32,5	30, 31, 32	4	9,8
7	32,5 – 35,5	33, 34, 35	5	12,2
8	35,5 – 38,5	36, 37, 38	8	19,5
9	38,5 – 41,5	39, 40, 41	4	9,8
10	41,5 – 44,5	42, 43, 44	3	7,3
11	44,5 – 47,5	45, 46, 47	4	9,8
12	47,5 – 50,5	48, 49, 50	2	4,9
13	50,5 – 53,5	51, 52, 53	1	2,4

З табл. 2.1 можна побачити, що границю найменшого розряду розпочато зі значення 14,5, яке дозволяє ввійти встановленій мінімальній оцінці 15. Додаючи до значення 14,5 установлений нами інтервал групування в 3 бали, обчислюємо верхню границю першого розряду: 14,5 + 3 = 17,5. Таким чином можна утворити 13 розрядів (див. табл. 2.2). У колонці 2 (табл. 2.2) відображено границі розрядів та значення за шкалою, які входять у розряд (колонка 3). Як можна побачити, у кожен розряд потрапили по три дискретні оцінки, які складають інтервал розряду.

На основі згрупованих частот за рівнозначними інтервалами легко побачити, які значення трапляються з високою частотою, а які є рідкими подіями.

На рис. 2.1 представлено *діаграму частот*, табульовану простим способом, а на рис. 2.2. представлено *гістограму*, яка є графічним відображенням групованих частот (частоти оцінок, які згруповані в розряди).

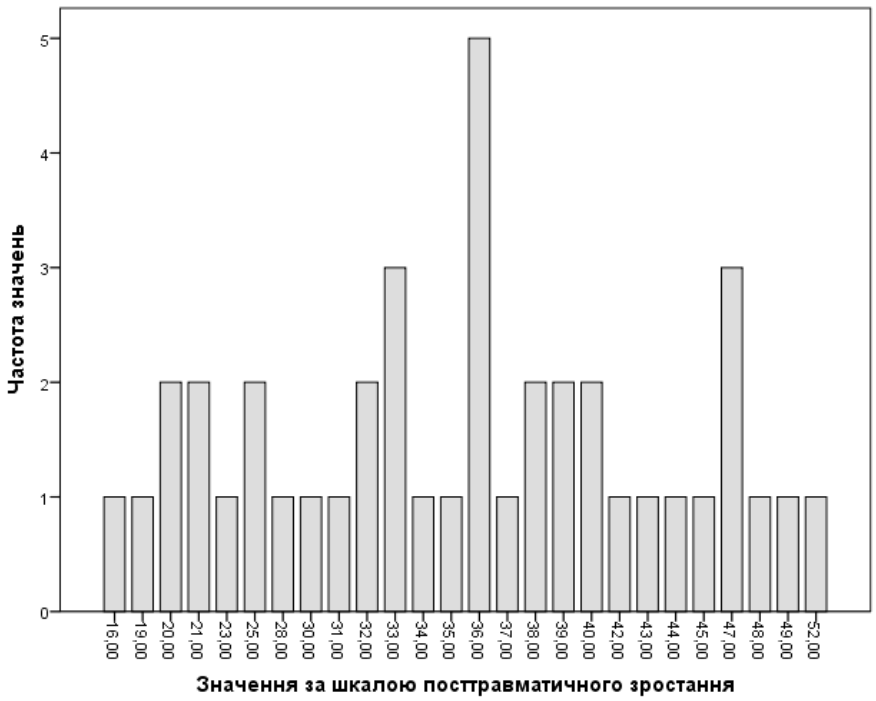

Рис. 2.1 Діаграма розподілу частот

Діаграма розподілу рис 2.1 графічно відображає розподіл частот у простому табулюванні (див. табл. 2.1). З її допомогою достатньо легко визначити *моду* – значення, яке найчастіше трапляється, має найвищу частоту (у нашому випадку це 36 балів, яке має частоту 5). Діаграма розподілу корисна, коли шкала має невисоку кількість балів (менше, ніж 16), і в такому випадку вона може відображати розподіл значень. Один із недоліків діаграми розподілу полягає в тому, що значення за горизонтальною шкалою X представлено за мірою їх появи без дотримання рівного інтервалу, що не дозволяє візуально відобразити міру розсіювання значень у вибірці.

За високої кількості балів за шкалою для презентації даних доцільніше використовувати гістограму.

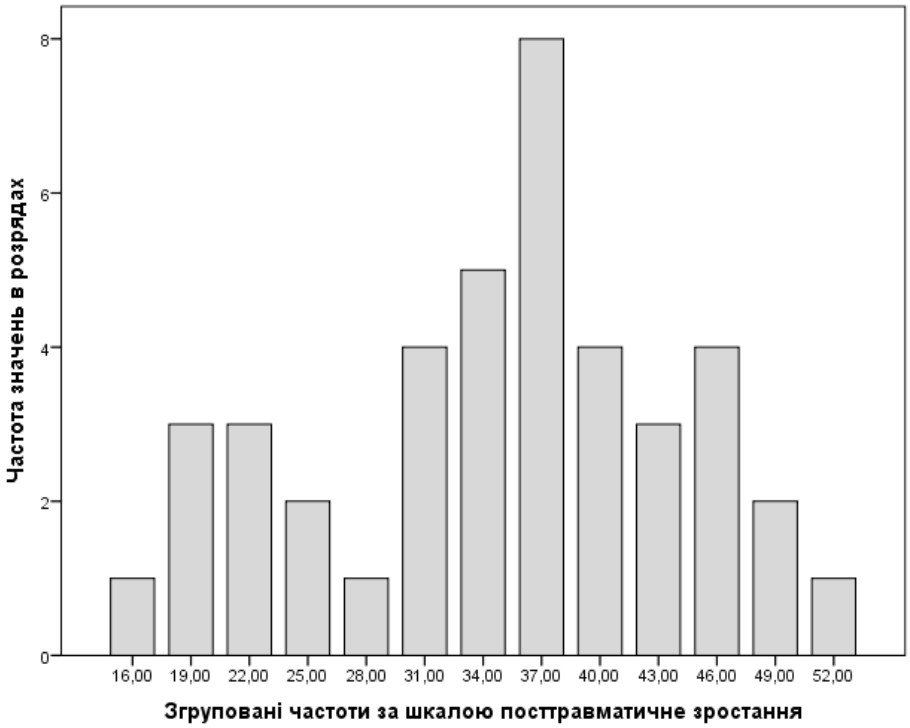

Рис. 2.2 Гістограма результатів за шкалою посттравматичного зростання

На рис. 2.2 наведено гістограму, побудовану на основі згрупованих частот (див. табл. 2.2) за шкалою посттравматичного зростання. Розряди в побудованій гістограмі позначено їх середніми значеннями. Презентація даних у гістограмі дозволяє одразу зробити висновок про те, які інтервали значень мають найбільшу та найменшу частоту, зробити попередній висновок про характер розподілу. Слід зазначити, що існують і інші вихідні принципи визначення інтервалів під час побудови гістограми. Так, В. Боснюк пропонує використовувати формулу Стерджеса та наводить таблицю з рекомендованою кількістю інтервалів [2, с. 41-42].

До іншого поширеного типу графічного відображення частот можна віднести *полігон* (див. рис. 2.3). Побудова полігону є схожою до побудови гістограми: напочатку частоти також необхідно згрупувати. У гістограмі кожний стовпчик закінчується горизонтальною лінією на висоті, яка відповідає частоті розряду. У полігоні кожний умовний стовпчик закінчується точкою, яка виставляється над середнім значенням розряду на висоті, що відповідає частоті розряду. Далі отримані точки з’єднуються прямими.

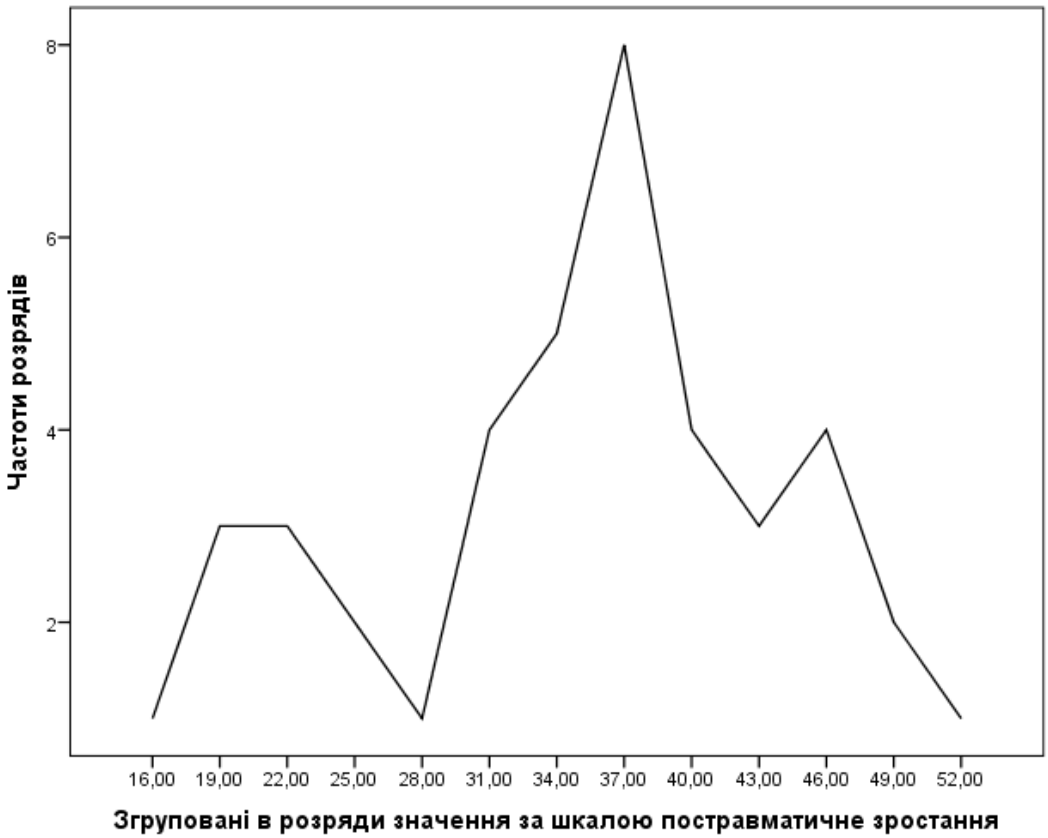

Рис. 2.3 Полігон згрупованих частот за шкалою «Посттравматичне зростання»

Перевага полігону полягає в тому, що він може візуально характеризувати розподіл у вибірці та дозволяє наочно порівнювати розподіли двох вибірок між собою [20, с. 46]. Так, на графіку можна намалювати полігони для кількох вибірок одночасно, і їх легко можна зіставити. На рис. 1.3 також було наведено полігон, а в описі до рисунка обговорювались частоти, які лежали в межах центрального розряду з середнім у 170 см. Є також багато інших способів графічного відображення частот, деякі з них наведено в роботі Дж. Гласс.

Для опису вибірки часто використовуються квантилі [12, с. 51; 2, с. 20]. *Квантиль* – загальне поняття, що означає точку на шкалі значень, яка ділить вибірку упорядкованих даних на дві групи з заданим співвідношенням за чисельністю (вимірюється у відсотках) [20, с. 20]. Одним із квантилів є медіана, яка ділить вибірку на дві групи з рівною кількістю опитаних.

До найуживаніших квантилів відносяться *квартилі, процентилі* та *децилі. Квартилі* – це три точки на шкалі значень, які ділять упорядковану за зростанням вибірку на чотири рівні за чисельністю частини по 25% опитаних в кожній. Так, нижче за точку Q_1 – 25% вибірки, нижче за Q_2 – 50% вибірки; нижче за Q_3 – 75% вибірки. Середній квартиль Q_2 збігається з медіаною, ділить вибірку навпіл та для шкали посттравматичного зростання становить 36 балів (досліджений з 21 порядковим номером). Для обчислення першого квартилю Q_1 слід визначити медіану у половини вибірки, яка менша за медіану (якщо кількість учасників парна, обчислюють середнє значення за шкалою у двох середніх учасників). Для значень (з табл. 2.1) медіана меншої половини вибірки є середнім значенням за шкалою у двох учасників, які перебувають посередині впорядкованого розподілу (це учасники дослідження з порядковими номерами 10 та 11): 28+30 = 58/2 = 29. Отже, Q_1 = 29. Значення Q_3 є середнім значенням за шкалою від досліджених з порядковими номерами 31 та 32: 40+42 = 82/2 = 41. Отже, Q_3 = 41.

Графічне відображення квартилів має назву «*коробчастий графік*», або «ящик з вусами» (див. рис. 2.4).

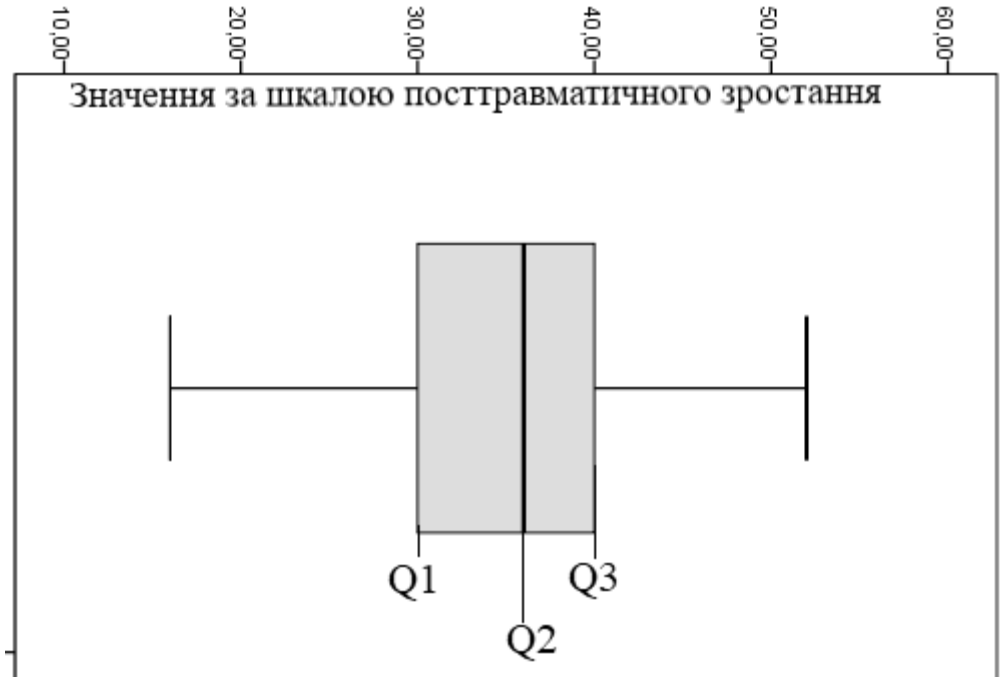

Рис. 2.4 Коробчастий графік розподілу значень

У середині рис. 2.4 можна бачити коробку, стінки якої відображають: ліворуч – значення нижнього, лівого квартиля Q_1; праворуч – значення верхнього квартиля Q_3. Лінія посередині відображає медіану (Q_2). Довжина «вусів» може варіюватись. Часто вони закінчуються найменшими (зліва) та найбільшими (справа) значеннями, які трапляються у вибірці. Інколи довжина «вусів» обчислюється за певними правилами й це дозволяє відобразити появу у вибірці значень за шкалою, які істотно відхиляються від теоретичних очікувань, є аномальними. Найчастіше довжина «вусів» обчислюється як 1,5*IQR (тобто 1,5 помножити на міжквартильний розмах). IQR обчислюється за формулою: IQR = Q3−Q1, де Q3 – точка третього квартилю, яка відділяє 75% спостережень; Q1 – точка першого квартилю, яка відділяє 25% спостережень. Коробчаста діаграма дозволяє візуально оцінити розподіл даних, його симетричність, міру однорідності результатів, а також зробити наочне порівняння характеру розподілу даних у декількох вибірках.

Процентилі є статистичними показниками, що поділяють впорядковану за зростанням вибірку на 100 рівних за чисельністю частин. Кожен процентиль позначає порогове значення ознаки, яке відповідає певному проценту розподілу. Наприклад, 10-й процентиль (P_{10}),) визначає значення, нижчі за яке – 10% спостережень, тобто він відокремлює нижню десяту частину вибірки від решти 90%. Процедура обчислення процентиля є аналогічною до визначення медіани, але з урахуванням відповідного процентного порогу. *Децилі* (D1–D9) поділяють упорядковану множину даних на 10 рівних частин, і кожен дециль репрезентує межу між відповідними інтервалами розподілу. Зокрема, D1 відповідає першій десятій частині вибірки, D2 — другій, і так далі до D9. У деяких психометричних тестах інтерпретація здійснюється на основі нормативних значень, виділених за допомогою процентилів та квартилів, що дозволяє здійснювати порівняльну оцінку індивідуальних показників у контексті популяційного розподілу.

2.2. Міри центральної тенденції

Міри центральної тенденції (МЦТ) – це числові показники, що відображають типові або узагальнені характеристики розподілу емпіричних даних. Вони можуть характеризувати генеральну сукупність та давати відповідь, наприклад, на питання: «Який середній рівень стресу у студентів під час війни?», визначити найпоширеніше ім'я серед українців у 2024 році тощо. До найбільш поширених МЦТ відносять: *моду*, *медіану*, *середнє арифметичне*.

Мода (Mo) – це статистичний показник, що репрезентує значення ознаки, яке найчастіше трапляється в емпіричному розподілі. Вона характеризує найтиповіше або найпоширеніше значення в сукупності даних. Важливо розрізняти саму моду як значення, а не її частоту. Для ряду значень: 1, 2, 2, 2, 3, 4, 4, 5, 6, 6, 6, 6, 7 — мода дорівнює 6, оскільки це значення повторюється чотири рази, що є найбільшою частотою серед усіх елементів вибірки. На рис. 2.1 мода добре видна та становить 36 балів. В. Боснюк зазначає, що мода може відображати центральну тенденцію лише тоді, коли частота цього значення в рази перевищує частоти суміжних значень [2, с. 13]. В. Руденко наводить низку правил для виділення моди [12, с. 36]. Ось ці правила:

– частоти подій можна розподілити таким чином, що моду неможливо виділити (мода відсутня). Наприклад, у розподілі 0, 0, 2, 2, 4, 4, 6, 6 усі значення мають однакову частоту;

– коли у впорядкованому розподілі значення, які знаходяться поряд, мають однакову частоту, мода обчислюється як середнє арифметичне цих значень. Наприклад, для розподілу 0, 0, 1, 2, 2, 3, 3, 3, 4, 4, 4 мода $Mo = (3+4):2 = 3,5$;

– коли значення не стоять поряд, але мають однакову частоту, розподіл описується як такий, що має декілька мод. Як приклад можна навести такі частоти значень: 5, 5, 6, 6, 6, 7, 8, 8, 8. У цьому розподілі виділяються дві моди (бімодальність): $Mo_1=6$ й $Mo_2=8$;

– у вибірці дані можуть бути розподілені таким чином, що доречним є виділення великих та малих мод. Наприклад, у розподілі: 9, 10, 10, 10, 10, 11, 11, 12, 12, 13, 13, 13, 13, 14, 14, 15, 15, 15, 15, 15, 15 – виділяється одна велика мода Mo_1=15 та дві малі моди: Mo_2=10 та Mo_3=13.

Мода використовується для дискретних шкал (у яких є лише цілі значення) та для номінальних шкал (наприклад, для вивчення, яке ім'я є найбільш поширеним).

*Медіана (*позначається як *Me* або *Md)* – це показник центральної тенденції, який визначає серединне значення у впорядкованій за зростанням вибірці. Вона поділяє сукупність даних на дві рівні частини: 50% спостережень мають значення ознаки, менше за медіану, а решта 50% — більше. Таким чином, медіана є стійким індикатором центру розподілу, особливо у випадках, коли дані містять екстремальні значення або мають асиметричний характер.

На графіку упорядкованих даних (без групування) медіану легко визначити візуально. У разі, якщо кількість досліджених непарна, медіана – це значення особи, яка перебуває посередині впорядкованої вибірки. Якщо ж кількість досліджених парна, медіана – це середнє значення за шкалою від двох осіб по центру розподілу впорядкованих даних.

Формула визначення медіани для непарної вибірки наведена на рис. 2.5:

$$Md = x_{(n+1)/2}$$

Рис. 2.5 Формула визначення медіани для непарної вибірки

У формулі *x – це значення за шкалою, за якою було опитано вибірку; нижній індекс* $_{(n+1)/2}$ *– повідомляє, які обчислення потрібно здійснити, щоб виявити значення x;* n *– обсяг вибірки.*

Отже, для того, щоб дізнатись медіальне значення в непарній вибірці, потрібно до кількості опитаних у вибірці додати одиницю та отриману суму поділити на два. У результаті буде отримано порядковий номер учасника дослідження, який є носієм медіального значення у впорядкованій вибірці.

Формулу обчислення медіани для вибірки з парною кількістю обстежених показано на рис. 2.6:

$$Md = \frac{x_{n/2} + x_{n/2+1}}{2}$$

Рис. 2.6 Формула визначення медіани для парної вибірки

З неї видно, що спочатку визначаються два порядкових номери учасників, які перебувають посередині впорядкованого ряду. Їх значення за шкалою потрібно просумувати та поділити на два (обчислити середнє), яке і буде медіаною.

Середнє значення (X) вважається найбільш вдалою, точною та незміщеною оцінкою, яка може характеризувати генеральну сукупність [20, с. 228]. На думку С. Бігун, середнє значення являє собою типового представника, загальну та закономірну особливість, у межах якої варіюються індивідуальні (випадкові) значення [1, с. 51]. Середнє значення застосовується, коли обсяг ознаки, яка усереднюється, дорівнює сумі індивідуальних значень окремих елементів.

В. Боснюк дає таке визначення: «С*ереднє арифметичне значення, середнє значення* – це значення ознаки, яке відображає середній рівень вираженості параметра у досліджуваних» [2, с. 14].

Є різні види обчислення середніх значень: гармонійне, кубічне, квадратичне, геометричне, арифметичне [1, с. 53]. Формули цих обчислень призначені для різних видів даних. Наприклад, середнє можна обчислити для згрупованих даних, коли відсутні індивідуальні значення, для результатів, переведених у стандартні бали. Найчастіше використовується середнє арифметичне. Формули для обчислення середнього арифметичного наведено на рис. 2.7:

$$\overline{X} = \frac{x_1 + x_2 + \ldots + x_n}{n} \text{ або } \overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ скорочено } \overline{X} = \frac{1}{n}\sum x_i$$

Рис. 2.7 Формула середнього значення

У формулі використано такі позначення: *x_1 – значення першого опитаного вибірки і т.д.; x_i – значення ознаки в конкретного опитаного з-поміж ознак усіх представників вибірки;* n *– обсяг (загальна кількість) вибірки.*

Середнє арифметичне є сумою всіх значень вибірки, яка ділиться на загальну кількість опитаних. Вважається, що статистичне середнє буде об'єктивною і типовою ознакою лише в тому разі, якщо воно обчислюється для якісно однорідної сукупності на великій вибірці [1, с. 51]. Якщо ж середні обчислюються для представників двох якісно різних сукупностей, отримані результати не матимуть наукового значення і їх потрібно вважати фіктивними [1, с. 52]. Для генеральної сукупності середнє обчислюється схожим чином та позначається μ.

2.3. Міри мінливості

Міри мінливості даних – це статистичні показники, які дозволяють оцінити ступінь однорідності, компактності та варіативності розподілу даних у вибірці [2, с. 18]. У психологічних дослідженнях найчастіше застосовуються такі показники, як варіаційний розмах, дисперсія, стандартне відхилення, коефіцієнт варіації, процентилі, квартилі та довірчий інтервал [2, с. 18].

Варіаційний розмах (R) є найпростішою мірою мінливості, визначається як інтервал між максимальною та мінімальною ознакою. Для його обчислення від найвищого значення шкали в обстеженій вибірці віднімається найменше значення. Цей вид розмаху також називається виключеним, і він найчастіше використовується [20, с. 75]. Для розбиття вибірки на розряди застосовується обчислення *включеного розмаху:* до різниці між максимальним та мінімальним значенням додається один бал.

Обчислення процентилів та квартилів вже розглянуто вище.

Дисперсія (s^2) характеризує міру однорідності даних. Чим вищим є значення дисперсії, тим більш різнорідними є дані. Для обчислення дисперсії використовується поправка: від загальної кількості вибірки (n) віднімається одиниця, що робить цю статистику незміщеною характеристикою генеральної сукупності. Обчислення дисперсії використовується виключно для даних, виміряних у шкалі інтервалів та відношення.

За визначенням В. Боснюк, *дисперсія* (s^2) – це середній квадрат відхилень усіх значень ознаки від середнього арифметичного [2, с. 18]. В формулі, кожне індивідуальне відхилення від середнього підноситься у квадрат для того, щоб у значенні статистики відображались абсолютні відмінності (тобто не було негативних значень). Потім, згідно формули дисперсії, сума всіх квадратів відхилень ділиться на загальну кількість вибірки плюс один. Тобто, обчислюється середній квадрат відхилення.

Для вибірки дисперсія обчислюється за такою формулою (див. рис. 2.8):

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

Рис. 2.8 Формула вибіркової дисперсії

У формулі x_i *– це значення ознаки за шкалою в конкретного досліджуваного;* $\bar{x}$ *– середнє арифметичне значення за шкалою;* n *– обсяг вибірки досліджуваних.*

Дисперсія для генеральної сукупності (δ^2) обчислюється за схожою формулою, але поправка (n-1, поправка Бесселя) не використовується. Сума квадратів відхилень ділиться просто на загальну кількість.

Дисперсія – дуже важливий показник, зокрема, її обчислення передбачає коефіцієнт кореляції Пірсона, дисперсійний аналіз тощо.

Стандартне відхилення, або середньоквадратичне відхилення (s), – це середнє відхилення кожного значення ознаки від середнього арифметичного.

Щоб його визначити, слід із дисперсії розрахувати квадратний корінь. Формулу наведено на рис. 2.9:

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

Рис. 2.9. Формула стандартного відхилення для вибірки

Для генеральної сукупності стандартне відхилення обчислюється схожим чином та позначається δ.

Стандартне відхилення є однією з найбільш корисних та часто вживаних статистик, оскільки за умови нормального розподілу даних ця статистика дає можливість прогнозувати ймовірність певного випадку [2, с. 80]. Так, у нормальному розподілі між x + s та x – s містяться 68% усієї вибірки.

Стандартне відхилення вважають статистикою, яка вказує на точність середнього значення [12, с. 59]. Що меншим є значення стандартного відхилення, то більш надійним є середнє.

На основі дисперсії та стандартного відхилення неможливо порівнювати між собою вибірки, виміряні в різних шкалах, – для таких порівнянь розроблено коефіцієнт варіації.

Коефіцієнт варіації, або квадратичний коефіцієнт варіації (V або Cv), – це відношення стандартного відхилення до середньої арифметичної величини ознаки [2, с. 19]. Коефіцієнт варіації вимірюється завжди у відсотках. Його формулу ми вже наводили, але представимо її ще раз (див. рис. 2.10):

$$V = \frac{\sigma}{\bar{x}} \quad \text{та} \quad V = s \,/\, \bar{X} \cdot 100\%$$

Рис. 2.10. Коефіцієнт варіації

Коефіцієнт варіації дозволяє оцінити однорідність вибірки та типовість середнього. Так, В. Боснюк наводить орієнтовні значення V, які вказують на однорідність: до 10% – низький рівень варіативності (однорідність); 10%– 20% – середній рівень варіативності; 20% і вище – висока варіативність вибірки (неоднорідність) [2, с. 20]. С. Бігун стверджує, що за V понад 33% вибірка вважається неоднорідною, а середнє значення нетиповим [1, с. 59].

Коефіцієнт варіації використовується для порівняння між собою розсіювання у вибірках, що виміряні в різних шкалах. Коефіцієнт варіації, так само як і дисперсія та стандартне відхилення, використовують лише для метричних шкал. Для шкал порядку використовується квартильний коефіцієнт варіації [1, с. 59].

2.4. Нормальний розподіл, характеристики форм розподілу

Крива нормального розподілу – це модель, винайдена математиками (А. Муавр, К. Гаусс, П. Лаплас) для опису розподілу даних великої кількості незалежних випадкових спостережень. Ця модель розподілу даних є найбільш теоретично вивченою та зручною для статистичного аналізу. Припущення, що значення нормально розподілені, дозволяє здійснювати теоретичний прогноз імовірності появи певної події (певного значення за шкалою) та частоти, з якою ця подія може траплятись. У параметричних статистичних критеріях припущення про нормальний розподіл даних є вихідним положенням і закономірності нормального розподілу використовуються для статистичних висновків. Під час створення тестових норм також часто виходять із закономірностей нормального розподілу, що дозволяє розподілити результати на типові (середні) та рідкісні, які свідчать про виражені індивідуальні особливості, відхилення від статистичної норми.

Нормально розподілені дані мають такі особливості (за В. Боснюк):

1. Найбільшу частоту мають події, близькі до середнього; при цьому частота крайніх значень ознаки наближається до нуля.

2. Відображені графічно частоти вибіркових даних мають вигляд кривої дзвоноподібного типу (що перевіряється візуально за оцінки гістограми).

3. Крива розподілу симетрична щодо центру.

4. Середнє арифметичне значення, мода і медіана в нормальному розподілі є рівними.

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

6. Частота розподілу даних близька до частоти нормального розподілу й підпорядковується їй. Вона є такою: ±1σ = 68,27 % площі розподілу; ±2σ = 95,45 %; ±3σ = 99,73 % [2, с. 38] (див. рис. 2.11).

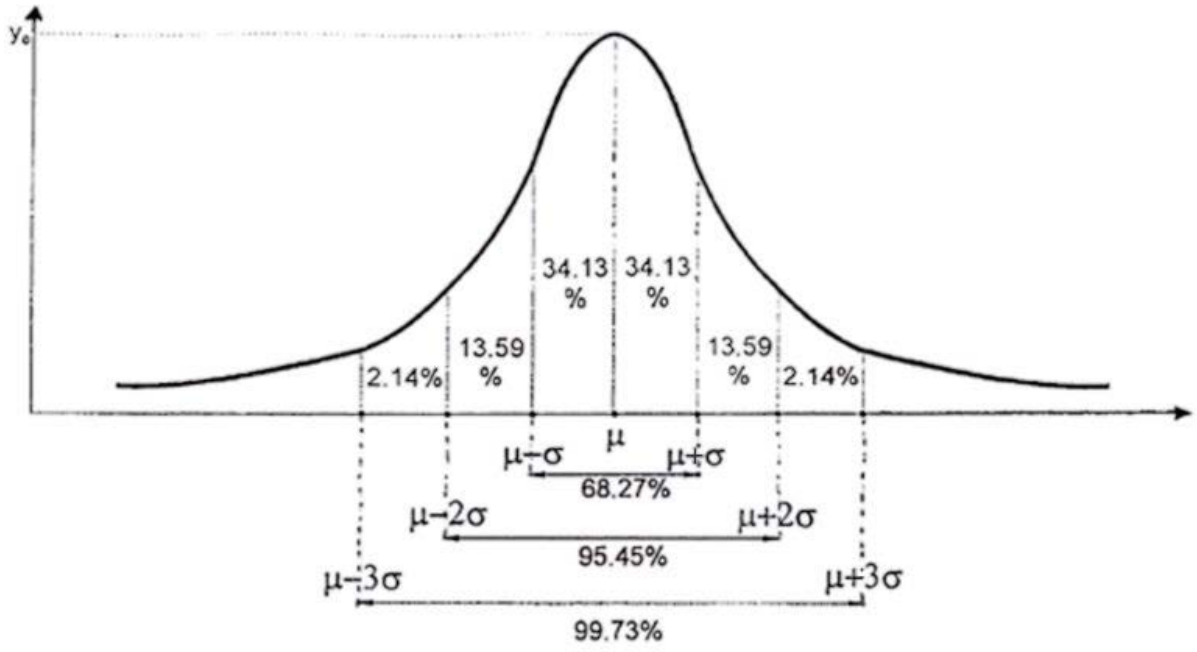

Рис. 2.11. Розподіл частот у теоретичному нормальному розподілі

Зазвичай, дані отримані в конкретному дослідженні на всі 100% не відповідають ідеальному нормальному розподілу, лише наближаються до повної відповідності. Дж. Гласс пише, що ніколи не була знайдена сукупність даних, яка має ідеальний нормальний розподіл. Тим не менш, для здійснення математичного аналізу прийнято, припускаючи незначну помилку,

використовувати криву нормального розподілу для опису даних конкретної вибірки та генеральної сукупності [20, с. 99].

Виходячи з припущення, що дані за певною шкалою є нормально розподіленими на основі середнього та стандартного відхилення, можна здійснювати прогноз (оцінку ймовірності) появи певного значення.

Якщо дані нормально розподілені, для здійснення оцінки ймовірності появи певного значення за шкалою здійснюють *z-перетворення* (рис. 2.12).

$$z_j = \frac{x_j - \overline{X}}{s}$$

Рис. 2.12. Перетворення значення в стандартизоване

У формулі використано такі позначення: x_j *– значення конкретного опитаного J за шкалою X; X – середнє арифметичне значення;* z_j *– нормоване значення опитаного J за шкалою.*

Z-перетворення переводить оцінки за шкалою в так звану одиничну нормальну криву розподілу, яка є еталонною. *Одиничний нормальний розподіл* має нульове середнє значення (X = 0) та стандартне відхилення, яке дорівнює одиниці (S = 1). Зіставляючи значення z зі значеннями спеціально розроблених таблиць, можна з точністю визначити, з якою частотою це значення трапляється (згідно з теоретичними припущеннями) та у скількох відсотків осіб значення за шкалою більше або менше [20, с. 99].

Z-перетворення не перетворює дані на нормально розподілені, як деякі вважають помилково. Якщо розподіл (до z-перетворення) сильно відрізнявся від нормального, то застосування z-перетворення для прогностичних цілей є некоректним. Z-перетворення переводить дані за нормально розподіленою шкалою в одиничний розподіл, що є корисним для розв’язання деяких статистичних завдань.

Розгляньмо статистики, які характеризують форму розподілу даних та можуть бути використані для наближеної оцінки *відповідності результатів за*

шкалою нормальному розподілу [2, с. 60]. До цих статистик належить значення *асиметрії* та *ексцесу*.

Асиметрія (А) характеризує ступінь несиметричності розподілу щодо його середнього арифметичного значення [2, с. 40]. Для обчислення асиметрії використовують формули, представлені на рис. 2.13:

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

Рис. 2.13. Формули обчислення асиметрії

У формулі використано такі позначення: x_i – *значення ознаки x у конкретного досліджуваного;* $\bar{x}$ *– середнє арифметичне значення;* s *– середньоквадратичне відхилення;* n *– обсяг вибірки досліджуваних.*

Представлена на рис. 2.13 права формула призначена для обчислення асиметрії в малих за обсягом вибірках за n < 30, для інших випадків використовується формула, наведена зліва. Коли А = 0, дані є ідеально симетричними. За позитивних значень статистики асиметрії (А > 0) асиметрія вважається зсунутою вправо, оскільки в такому випадку кількість значень за шкалою (подій), які більші за середнє, буде більшою (див. рис. 2.14). Відповідно, кількість значень за шкалою менших за середнє менша.

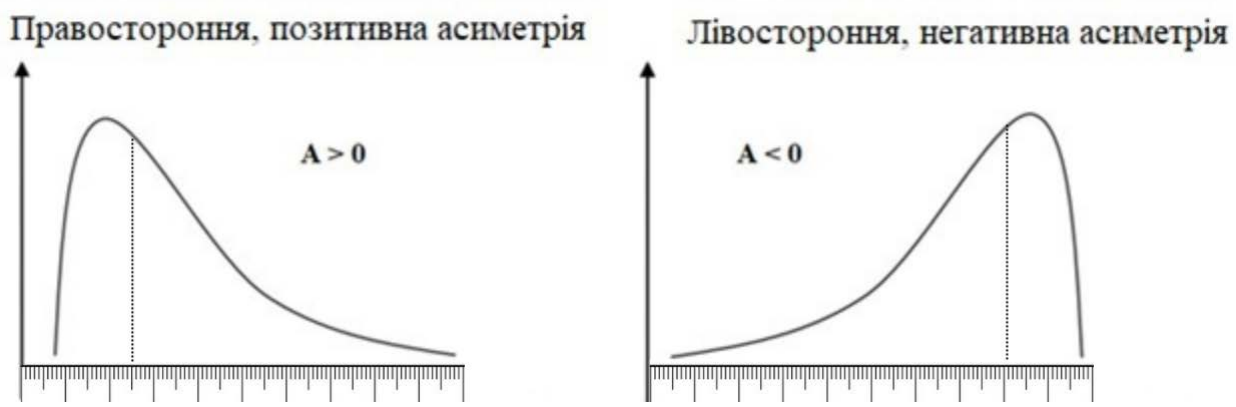

Рис. 2.14. Асиметричність розподілу

За негативних значень асиметрії (A < 0) спостерігається лівостороння асиметрія, бо буде більше подій, які мають значення, менше за середнє (див. рис. 2.14).

Ексцес (E) характеризує відносну опуклість або згладженість розподілу вибірки порівняно з нормальним розподілом [2, с. 40]. В. Боснюк зазначає, що «позитивний ексцес (E > 0) характеризує порівняно загострений розподіл, негативний (E < 0) – порівняно згладжений» [2, с. 41].

Формули для обчислення ексцесу представлено на рис. 2.15, для середнього та великого обсягу вибірки використовується формула справа, для вибірок з n < 30 – з лівої сторони.

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

Рис. 2.15 Формули обчислення ексцесу

Позначення, використані у формулі, аналогічні до формули розрахунку асиметрії, тому нами не наводяться. Графічно розподіли з позитивним та негативним ексцесом представлено на рис. 2.16.

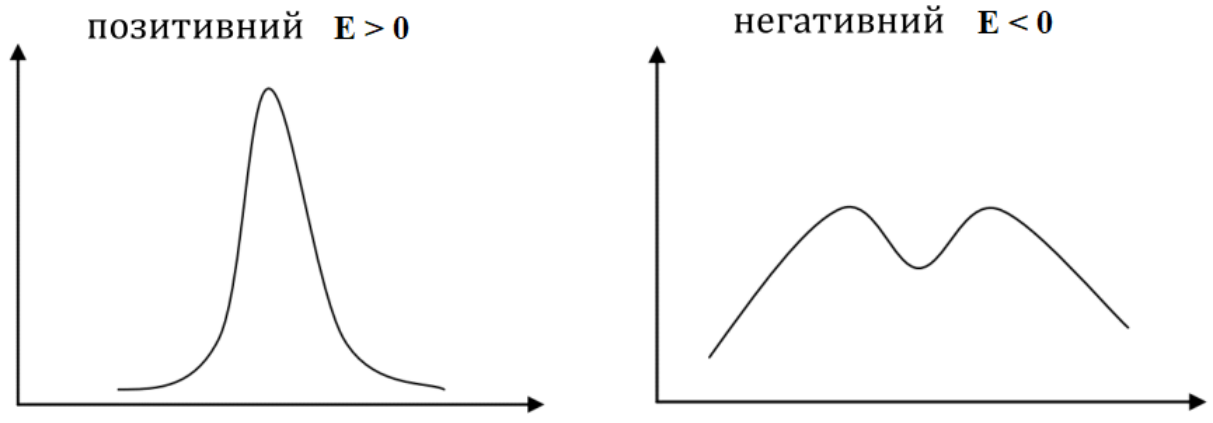

Рис. 2.16 Розподіли з вираженим ексцесом

Для приблизної оцінки нормальності вибіркового розподілу вважається, що обидва показники асиметрії та ексцесу мають бути в межах від -1 до +1 [2, с. 50]. Крім того, для оцінки нормального розподілу за показниками асиметрії

та ексцесу розраховуються спеціальні критичні значення, про що можна детальніше дізнатись з роботи В. Боснюк [2, с. 43-45].

Деякі автори стверджують, що вибірка обсягом до 30 учасників не може мати нормального розподілу, тому до таких вибірок потрібно використовувати непараметричні критерії [2, с. 41].

Найбільш поширеним способом перевірки вибірки на відповідність нормальному розподілу є обчислення критеріїв Шапіро-Уїлка (для вибірок 30-100 осіб) та Колмогорова-Смірнова (для вибірок обсягом у понад 100).

Якщо рівень достовірності за зазначеними критеріями є меншим за 5% ($p<0,05$), то відхилення від теоретичного нормального розподілу є достовірним, отриманий у дослідженні вибірковий розподіл неправильно описувати за допомогою кривої нормального розподілу та її закономірностей; при статистичному аналізі слід використовувати непараметричні критерії.

Тема 2. Кореляційний та регресійний аналіз

Лекція 3. Кореляційний та регресійний аналіз

Мета: сформувати ґрунтовне розуміння концепції кореляційного аналізу та методології його статистичного проведення. Ознайомити студентів з принципами та застосуванням одномірної лінійної регресії. Поняття кореляції. Коефіцієнти кореляції Пірсона і Спірмена. Інтерпретація коефіцієнтів кореляції: лінійна і нелінійна кореляція. Коефіцієнти взаємної зв'язаності. Регресія як засіб прогнозування. Лінійні моделі регресії. Одномірна лінійна регресія.

Основні поняття теми: коефіцієнт кореляції, напрям зв'язку, сила зв'язку, значущість зв'язку, графік розсіювання, лінійна кореляція, регресійне рівняння, залежна змінна, незалежна змінна, загальна лінійна модель.

План

1. Кореляційний аналіз: поняття та основні завдання
2. Типи коефіцієнтів кореляції
3. Коефіцієнти взаємної зв'язаності
4. Регресійний аналіз

3.1. Кореляційний аналіз: поняття та основні завдання

Кореляційний аналіз є одним із найпоширеніших методів математичної статистики, який використовується для вивчення взаємозв'язків між змінними. У психології цей метод має широке застосування, оскільки більшість досліджень передбачає оцінку зв'язків між характеристиками особистості, поведінковими проявами чи психофізіологічними параметрами.

Кореляція – це статистичний показник, який характеризує силу та напрям взаємозв'язку між двома змінними. Вона вказує на те, чи існує взаємозв'язок між змінними, але не визначає, який з факторів є причиною, а який наслідком [2]. Якщо дві змінні корелюють, це означає, що вони змінюються разом, але це не обов'язково передбачає причинно-наслідкові

зв’язки. У цьому контексті кореляція є цінною, оскільки виявлення певного взаємозв’язку може бути корисним для подальшого дослідження причинно-наслідкових взаємозв’язків. Для цього вчені організовують експериментальне або квазіекспериментальне дослідження. Таким чином, хоча кореляція може допомогти виявити потенційні зв’язки між змінними, для встановлення причинності необхідно додатково застосовувати інші методи й обережно трактувати отримані результати.

Народження методології вимірювання взаємозв'язків припадає на кінець XIX століття – період інтенсивного становлення статистичних методів дослідження. Центральними постатями цього наукового напрямку стали британські вчені Ф. Гальтон та К. Пірсон, які заклали фундаментальні основи кореляційного аналізу.

Ф. Гальтон, двоюрідний брат Ч. Дарвіна, вважається засновником антропометрії та психометрики. У 1870-1880-х роках він розпочав систематичні дослідження індивідуальних відмінностей, запровадивши кількісні методи вимірювання людських здібностей. Саме Ф. Гальтон першим увів термін «кореляція» та почав розробляти статистичні методи для виявлення взаємозв'язків між різними характеристиками.

Карл Пірсон, учень та послідовник Гальтона, значно просунув кореляційний аналіз далі, розробивши найвідоміші статистичні коефіцієнти кореляції. Зокрема, коефіцієнт кореляції Пірсона (r) став стандартним інструментом для вимірювання лінійного зв'язку між двома змінними.

З моменту свого виникнення методи кореляційного аналізу швидко стали надзвичайно популярними в психологічних дослідженнях. Вони дозволили науковцям:

- кількісно оцінювати взаємозв'язки між психологічними характеристиками;
- виявляти приховані залежності між різними особистісними властивостями;
- здійснювати статистичне прогнозування на основі емпіричних даних.

На сьогодні коефіцієнт кореляції залишається одним з найважливіших інструментів наукового пізнання в психології, соціології, педагогіці та багатьох інших гуманітарних і природничих науках.

Коефіцієнт кореляції – це числова міра, що визначає силу та напрямок ймовірнісного взаємозв’язку між двома змінними. Його значення лежить у межах від -1 до +1.

Сила зв’язку досягає максимуму, коли між змінними існує однозначна відповідність: кожному значенню однієї змінної відповідає єдине значення іншої. У такому випадку емпіричний зв’язок відповідає функціональній лінійній залежності. Показником сили є абсолютне значення коефіцієнта кореляції: що ближче його модуль до 1, тим сильніший зв’язок.

Напрямок зв’язку визначається залежністю між змінними:

1. Прямий (позитивний) зв’язок: зі збільшенням значень однієї змінної зростають значення іншої.

2. Зворотний (негативний) зв’язок: збільшення значень однієї змінної супроводжується зменшенням значень іншої.

Знак коефіцієнта кореляції вказує на напрямок зв’язку. Якщо коефіцієнт кореляції близький до 0, це свідчить про відсутність залежності між змінними.

Монотонний та немонотонний взаємозв’язок. У статистичних дослідженнях взаємозв'язки між змінними можуть мати принципово різну природу, яка визначається характером монотонності або немонотонності.

Монотонний взаємозв'язок характеризується послідовною, передбачуваною зміною однієї змінної відносно іншої – коли збільшення або зменшення першої змінної однозначно викликає пропорційне збільшення чи зменшення другої. На противагу цьому, *немонотонний* взаємозв'язок являє собою складнішу картину, де зміна однієї змінної може викликати непередбачувану, нелінійну реакцію іншої: графік такої залежності може мати хвилеподібний, стрибкоподібний або циклічний характер (див. рис 3.1).

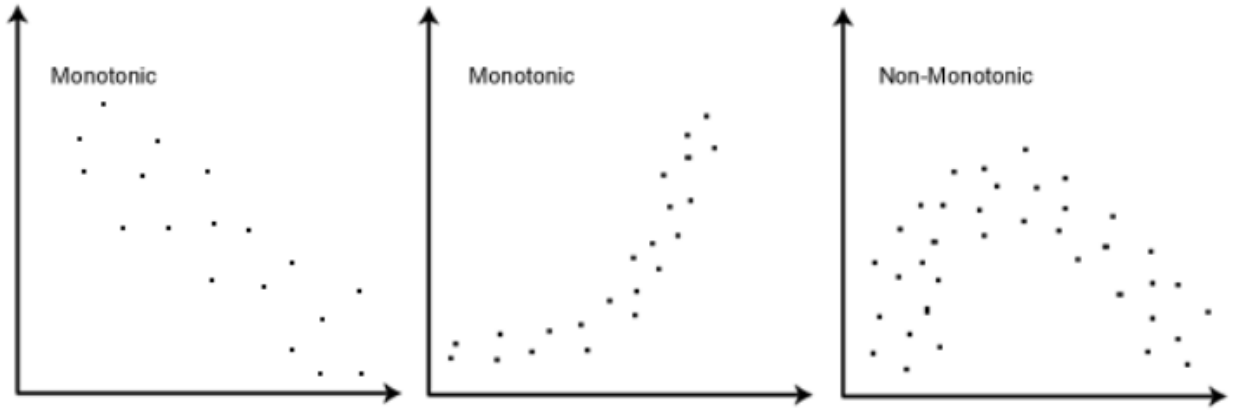

Рис. 3.1. Приклади монотонного та немонотонного взаємозв’язку

Діаграми розсіювання як метод візуалізації зв’язку. Діаграми розсіювання є інструментом для візуалізації взаємозв’язку між двома змінними. Це діаграма, на якій кожна точка відображає значення двох змінних. Одну змінну відкладають по горизонтальній осі (X), а іншу – по вертикальній осі (Y). Завдяки цьому можна швидко оцінити тип зв’язку між змінними: лінійний, нелінійний або відсутність зв’язку. Такі графіки допомагають не лише візуально оцінити кореляцію між змінними, а й виявити можливі екстремальні значення або нестандартні патерни в даних.

Використання графіків розсіювання є важливим на етапі попереднього аналізу даних, коли необхідно оцінити, які статистичні методи варто застосовувати для вивчення взаємозв’язків, а також визначити, чи є потреба в очищенні даних.

Класифікація кореляційних зв’язків. Є дві основні класифікації кореляційних зв’язків: за величиною коефіцієнта та за рівнем статистичної значущості. Типи кореляційних зв’язків за величиною коефіцієнта кореляції зображено на рис. 3.2.

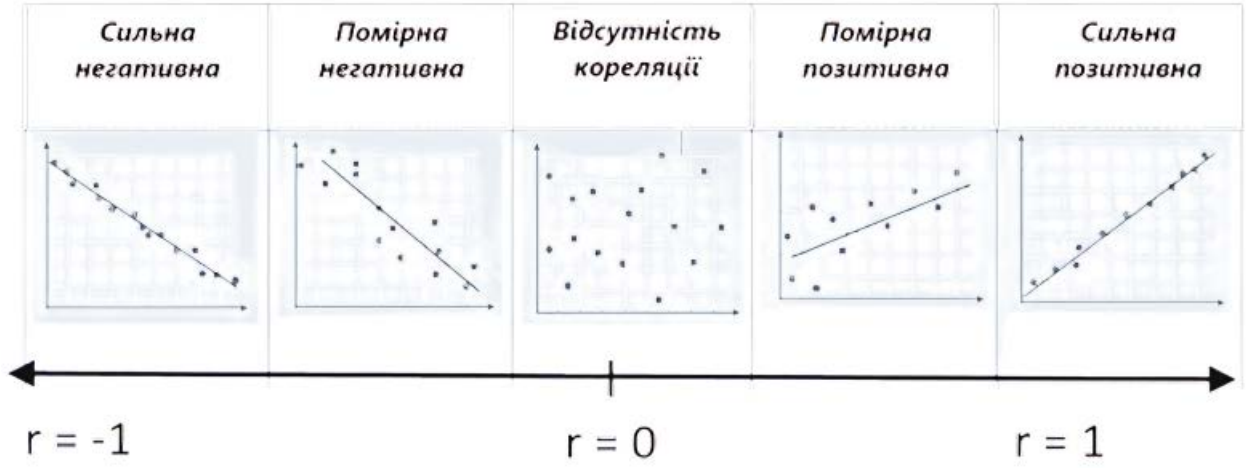

Рис 3.2. Діаграма розсіювання при різних значеннях сили кореляції

Залежно від коефіцієнта кореляції сила взаємозв'язку може бути:

а) сильна : $r > 0,70$ або $r < -0,70$;

б) середня : $0,50 < r < 0,69$ або $-0,69 < r < -0,50$;

в) помірна $0,30 < r < 0,49$ або $-0,49 < r < -0,30$;

г) слабка $0,20 < r < 0,29$ або $-0,29 < r < -0,20$;

д) дуже слабка $r < 0,19$ або $r > -0,19$.

Оцінюючи силу кореляційного зв'язку, ми повинні також враховувати *рівень статистичної значущості*, який вказує на ймовірність того, що отриманий коефіцієнт кореляції є не випадковим, а справді відображає залежність між змінними. Рівні значущості допомагають визначити, наскільки надійними є результати аналізу та чи можна їх узагальнювати на всю генеральну сукупність.

Рівень значущості (α) є критичним порогом, який встановлюється заздалегідь і визначає допустиму ймовірність помилки першого роду, тобто відхилення нульової гіпотези, коли вона правильна. Найчастіше використовують такі його значення: 0,05, 0,01, 0,001. Ці рівні свідчать про готовність дослідника прийняти ризик помилки в 5%, 1% або 0,1 % випадків відповідно. *P-значення,* яке обчислюється в ході статистичного тестування, вказує на ймовірність отримання спостережуваного або більш екстремального результату за умови істинності нульової гіпотези. Якщо p-значення (p) є

меншим за встановлений рівень альфа, це свідчить про статистичну значущість результатів, що дає підстави для відхилення нульової гіпотези на користь альтернативної. Такими чином:

❖ якщо $p \leq \alpha$, то ми *відхиляємо* нульову гіпотезу, кореляція є статистично значущою: між змінними, імовірно, існує кореляційний зв'язок;

❖ якщо $p > \alpha$, ми *не відхиляємо* нульову гіпотезу, немає достатніх підстав стверджувати про наявність статистично значущої кореляції між змінними.

3.2. Типи коефіцієнтів кореляції

У психологічних дослідженнях широко застосовують *коефіцієнт лінійної кореляції Пірсона*, який вимірює силу та напрямок лінійного зв'язку між двома змінними. Окрім цього, використовують рангові методи кореляції, зокрема Спірмена та Кендалла, які допомагають виявити монотонні залежності, особливо у випадках, коли дані не відповідають припущенню про нормальний розподіл або містять викиди (екстремальні значення). Вибір критерію кореляції залежить від *типу даних*, *розподілу змінних і* вимог до статистичних припущень (обраний *рівень значущості*).

Коефіцієнт лінійної кореляції Пірсона (r). Критерій лінійної кореляції, розроблений британським математиком Карлом Пірсоном у 1896 році, є статистичним методом, який використовується для вимірювання сили та напрямку лінійної залежності між двома неперервними змінними. Представлений коефіцієнтом кореляції Пірсона, що позначається як r, статистичний тест кількісно визначає, наскільки тісно точки даних групуються навколо прямої лінії, що дозволяє дослідникам оцінювати зв'язки між змінними в різних наукових галузях, таких як психологія, медицина, економіка та соціальні науки. Він став основним інструментом статистичного аналізу, інтегральним елементом для перевірки гіпотез і перевірки дослідницьких теорій, заснованих на емпіричних даних.

Коефіцієнт лінійної кореляції розраховується за формулою, представленою на рис. 3.3:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Рис. 3.3. Формула коефіцієнта лінійної кореляції Пірсона

У формулі використано такі позначення: *n – розмір вибірки;* x_i , y_i *– індивідуальні значення за індексом і змінних x та y;* $\bar{x}$, $\bar{y}$ *– середні значення для змінних x та y.*

Перед проведенням кореляційного аналізу дослідник повинен ретельно перевірити відповідність емпіричних даних певним статистичним критеріям, які забезпечують валідність та надійність результатів дослідження. Саме тому принципово важливо дотримуватися таких *умов застосування* коефіцієнта лінійної кореляції Пірсона:

1. *Лінійність зв'язку.* Першочерговою умовою є припущення про лінійний зв'язок між двома змінними. Лінійність зв'язку передбачає, що зміна змінної X на одну одиницю завжди призводить до зміни змінної Y на одну і ту ж величину. Якщо залежність має нелінійний характер, коефіцієнт Пірсона може недооцінити силу зв'язку.

2. *Шкали вимірювання*. Дані повинні бути отримані за допомогою *метричних шкал* (інтервальної, відношень). Використання порядкових даних не завжди відповідає вимогам для коректного застосування цього методу.

3. *Нормальність розподілу*. Для коректного проведення статистичних висновків (наприклад, за оцінки значущості) кожна змінна має приблизно відповідати нормальному розподілу. У разі значних відхилень можуть виникнути проблеми з точністю висновків. Для перевірки припущень про нормальність розподілу використовують критерії Шапіро-Уілка та Колмогорова-Смірнова.

4. *Обсяг вибірки:* $n_1 \geq 30$ і $n_2 \geq 30$

Для забезпечення наукової коректності та достовірності дослідження важливо розуміти обмеження в застосуванні коефіцієнта лінійної кореляції Пірсона:

1. *Чутливість до викидів*. Аномальні значення можуть суттєво вплинути на величину коефіцієнта, спотворюючи реальну картину зв’язку.

2. *Обмеження на тип зв’язку*. Коефіцієнт Пірсона відображає лише лінійну залежність. У випадку *нелінійних* відносин його значення може бути близьким до нуля навіть за наявності сильної залежності між змінними.

3. *Порушення нормальності*. Якщо розподіл даних значно відхиляється від нормального, результати аналізу можуть бути ненадійними, і в таких випадках доцільно застосовувати непараметричні методи (наприклад, рангову кореляцію Спірмена або Кендала).

4. *Вплив розміру вибірки*. Невеликий розмір вибірки може призвести до нестабільних, нерепрезентативних оцінок коефіцієнта.

Коефіцієнт рангової кореляції Спірмена (r_s). Коефіцієнт рангової кореляції Спірмена, що позначається як r_s є непараметричним статистичним показником, який оцінює силу та напрямок зв'язку між двома ранжованими змінними. Введений Чарльзом Спірменом у 1904 році, цей коефіцієнт є особливо корисним для аналізу порядкових даних і ненормально розподілених (див. *нормальний розподіл*) наборів даних, забезпечуючи надійну альтернативу коефіцієнта кореляції Пірсона, який передбачає лінійні зв'язки й нормальність розподілу даних.

Коефіцієнт Спірмена розраховується за формулою на рис. 3.4:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

Рис. 3.4. Коефіцієнт рангової кореляції Спірмена

У формулі використано такі позначення: d_i – *різниця парних рангів,* n *– кількість респондентів.*

Коефіцієнт кореляції Спірмена використовується для вимірювання монотонного (але не обов’язково лінійного) зв’язку між двома змінними. Він базується на рангах даних, а не на їхніх фактичних значеннях.

Коефіцієнт рангової кореляції Спірмена має кілька важливих переваг, які роблять його корисним у психологічних дослідженнях. По-перше, цей коефіцієнт підходить для роботи з непараметричними даними, оскільки не потребує припущення про нормальний розподіл досліджуваних змінних. По-друге, він є стійким до впливу аномальних значень (викидів), що робить його надійним у випадках, коли дані містять екстремальні спостереження. По-третє, коефіцієнт Спірмена здатен виявляти монотонні зв’язки між змінними навіть тоді, коли ці зв’язки не є лінійними, що розширює можливості його застосування для аналізу складних взаємозв’язків.

Водночас коефіцієнт рангової кореляції Спірмена має кілька важливих обмежень, які слід враховувати під час його застосування. По-перше, цей критерій не враховує силу нелінійного, але немонотонного зв’язку між змінними, тому може не виявити взаємозв’язки, які не мають чіткої монотонності. По-друге, він є менш потужним порівняно з коефіцієнтом лінійної кореляції Пірсона у випадках, коли між змінними справді існує лінійна залежність. Крім того, для його застосування рекомендовано мінімальний обсяг вибірки, не менший за п’ять учасників дослідження. До того ж, якщо у вибірці спостерігається велика кількість однакових рангів, показник може втрачати точність, тому в таких випадках обов’язково потрібно використовувати поправку на однакові ранги.

Залежно від поставленого завдання дослідження та особливостей емпіричних даних обирається відповідний статистичний метод оцінки взаємозв'язку. Детальне порівняння коефіцієнтів кореляції Пірсона та Спірмена представлено в табл. 3.1

Таблиця 3.1

Порівняння критеріїв кореляції Пірсона та Спірмена

Характеристика	**Коефіцієнт кореляції Пірсона**	**Коефіцієнт кореляції Спірмена**
Тип зв'язку	лінійна залежність	лінійна або нелінійна
Тип даних	інтервальний або відношень	порядковий, інтервальний, відношень
Стійкість до екстремальних значень	чутливий	менш чутливий
Обчислення	використовує початкові значення змінних	використовує ранги змінних
Умови застосування	-нормальний розподіл даних	- не потребує нормального розподілу даних
Переваги	- забезпечує точніший результат за нормального розподілу даних і лінійності зв'язку	- гнучкість у застосуванні до порядкових даних або нелінійних взаємозв'язків; - більш стійкий до викидів та розподілу даних, що суттєво відрізняється від нормального
Недоліки	-непридатний для виявлення нелінійних зв'язків; -чутливий до екстремальних значень (викидів)	-втрачається частина інформації при ранжування.

3.3. Коефіцієнти взаємної зв'язаності

У статистичному аналізі категоріальних даних застосовується *коефіцієнт кореляції Пірсона (φ), коефіцієнт контингенції Пірсона (C)* та *коефіцієнт Чупрова-Крамера (Cramer's V).* Зазначені коефіцієнти виступають інструментами дослідження взаємозв'язків між змінними в *шкалах найменувань*. На відміну від кореляційних коефіцієнтів, призначених для кількісних шкал, ці статистичні методи спеціально розроблені для роботи з якісними характеристиками, представленими в перехресних таблицях. Коефіцієнт кореляції *φ* спрямований на перевірку міри взаємозв'язку в шкалах

найменувань, коефіцієнт контингенції С дозволяє перевірити гіпотезу про наявність статистично значущого зв'язку між категоріями, а коефіцієнт Чупрова-Крамера додатково оцінює силу цього зв'язку.

Формули обчислення критеріїв зв’язаності оперують частотами в таблицях узгодженості (узагальнені дані щодо декількох змінних).

Коефіцієнт кореляції Пірсона (позначається як φ) – це міра зв’язку між двома категорійними змінними, яка базується на порівнянні їх частот. Коефіцієнт φ визначається за формулою на рис. 3.5:

$$\varphi = \frac{p_{xy} - p_x p_y}{\sqrt{p_x q_x p_y q_y}}$$

Рис. 3.5. Формула коефіцієнта кореляції Пірсона φ

У формулі використано такі позначення: p_{xy} – *загальна частка спільних подій (поява значення 1) одночасно в стовпчиках X та Y;* p_x – *загальна частка подій (поява значення 1) у стовпці X;* p_y – *загальна частка подій (поява значення 1) у стовпці Y;* q_x – *загальна частка відсутності події (поява значення 0) у стовпці X;* q_y – *загальна частка відсутності події (поява значення 0) у стовпці Y.*

Наведена формула використовується для невпорядкованих в таблицю даних під час порівняння двох стовпців з результатами опитаних за двома змінними. У формулі на рис. 3.5 під часткою події мають на увазі відношення кількості подій до загальної кількості. Наприклад, якщо в стовпчику X є 12 спостережень, серед яких 6 є подіями зі значенням 1, то частка події 1 буде становити 6/12 = 0,5.

Дж. Гласс зазначає, що коефіцієнт кореляції φ – це адаптація коефіцієнта кореляції Пірсона для дихотомічних шкал [20, с. 146]. Простішим та зрозумілішим може бути обчислення φ на основі частот, що представлені в таблиці узгодженості 2x2.

Як приклад використаємо дві змінні: фактор MD, який вимірює адекватність самооцінки, та фактор С, який вимірює особистісну зрілість. На основі середніх значень (отриманих на вибірці 92 студентів) дані з інтервальної шкали було перекодовано в дихотомічну таким чином:

- MD < 5,5 = 0 – самооцінка адекватна, значення менше за середньо-вибіркове.

- MD > 5,5 = 1 – самооцінка завищена, значення більше за середньо-вибіркове.

- C < 5,5 = 0 – емоційна нестійкість, значення менше за середньо-вибіркове.

- C > 5,5 = 1 – емоційна стійкість, значення більше за середньо-вибіркове.

Результати діагностики за двома зазначеними факторами представлено в табл. 3.2

Таблиця 3.2

Результати діагностики студентів за факторами MD та С методики 16PF

Значення за фактором MD (**X**)	Значення за фактором С (**Y**)		Сума частот в строчках
	Емоц. нестабільність C<5,5 (0 балів)	Емоц. стабільність C>5,5 (1 бал)	
Адекватна самооцінка MD<5,5 (0 балів)	30 (67%) **A**	15 (33%) **B**	45
Завищена самооцінка MD>5,5 (1 бал)	9 (19%) **C**	38 (81%) **D**	47
Сума частот в стовпчиках	39	53	92

На перший погляд, табл. 3.2 (після впорядкування даних в таблицю та обчислення відсотків кожної події для рядків з адекватною самооцінкою та завищеною) дозволяє зробити припущення, що студенти з умовно адекватною самооцінкою оцінюють себе як більш емоційно нестабільних. Так, емоційна нестабільність характерна для 67% студентів з-поміж тих, що мають адекватну самооцінку. Водночас студенти з умовно завищеною самооцінкою схильні

отримувати результати, які свідчать про їх високу емоційну стабільність. Так, 87% студентів з завищеною самооцінкою вважають себе емоційно стабільними і лише 18% оцінюють себе як схильних до емоційної нестійкості. Можна допустити, що в студентів з умовно завищеною самооцінкою є схильність переоцінювати свої емоційно-вольові особливості.

Утім, для того, щоб зробити висновок про існування цієї закономірності в генеральній сукупності, слід скористатись статистичними методами, у нашому випадку коефіцієнтом φ Пірсона. Формулу для обчислення впорядкованих в таблицю даних показано на рис. 3.6

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

Рис. 3.6. Формула φ для таблиць 2x2

У формулі використано такі позначення: *a, d, b, c – частоти у відповідних клітинках (див. табл. 3.2)*

Отже, для того, щоб обчислити φ, потрібно спочатку перемножити збіги частот подій за двома шкалами: A (X=0, Y=0) та D (X=1, Y=1), а потім від отриманої суми відняти добуток частот з двох клітинок, які не збігаються: B (X=0, Y=1) та C (X=1, Y=0). Обчислення проводиться згідно з формулою 3.6.

Підрахунок за даними таблиці 3.2 показує, що φ = 0,48. Рівень достовірності оцінюється на основі переведення значення φ в статистику χ^2 за формулою на рис. 3.7, за якою здійснюється перевірка рівня достовірності.

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

Рис. 3.7. Формула переведення значення коефіцієнта φ в значення критерію χ^2

Для даних, наведених у табл. 3., рівень достовірності становить 0,1% ($p < 0,001$). Таким чином, можна зробити висновок, що студенти із

значеннями, вищими за середнє за шкалою «Адекватність самооцінки», схильні до перебільшення своїх емоційно-вольових рис.

φ коефіцієнт має певні обмеження, які слід враховувати під час його застосування. По-перше, він чутливий до розподілу змінних: якщо одна з категорій має дуже малу або надто велику частку спостережень (наприклад, 90% в одній категорії), значення φ може недооцінювати реальну силу зв’язку.

По-друге, фі-коефіцієнт виявляє лише лінійні асоціації між двома змінними і не в змозі виявити складніші нелінійні залежності, які можуть мати психологічне чи поведінкове значення. Також варто пам’ятати, що метод застосовний виключно для бінарних змінних і не підходить для даних із більшим числом категорій. У таких випадках доцільно звертатися до альтернативних статистичних методів, наприклад до коефіцієнта Крамера V для номінальних змінних або до коефіцієнта τ Кендалла для порядкових змінних.

Коефіцієнт контингенції Пірсона (позначається як C) призначений для порівняння міри зв’язаності в таблицях з рівною кількістю рядків і стовпців [11, с. 65]. Коефіцієнт ґрунтується на критерії χ^2, тож вважаємо за доцільне спочатку навести інформацію про нього.

*Критерій Хі квад*рат (χ^2) є непараметричним статистичним методом, призначеним для аналізу відмінностей між емпіричними та теоретичними частотами розподілу досліджуваних ознак. Він дозволяє перевірити статистичні гіпотези про взаємозв'язок між категоріальними змінними, оцінити значущість розбіжностей у перехресних таблицях та визначити, чи є спостережувані відмінності випадковими або наслідком суттєвих закономірностей. Розрахунок χ^2 базується на порівнянні очікуваних та фактичних частот, де статистично значущі результати (низький рівень p-значення) свідчать про наявність статистично достовірного зв'язку між досліджуваними категоріями. Формулу для розрахунку критерію χ^2 для таблиць 3x3 і більше представлено ліворуч, для таблиць 2x2 – праворуч на рисунку 3.8:

$$\chi^2 = n\left[\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{f_{ij}^2}{f_{i.}f_{.j}} - 1\right], \quad \chi^2 = \frac{n(f_{11}f_{22} - f_{12}f_{21})^2}{f_{1.}f_{2.}f_{.1}f_{.2}}$$

Рис. 3.8. Формули для обчислення значення за критерієм χ^2

У формулі використано такі позначення: *n – кількість опитаних;* f_{ij} *– спостережувана частота в комірці (i, j);* $f_{i.}$ *– загальна частота в стовпчику i;* $f_{.j}$ *– загальна частота в рядку j;* f_{11} *– комірка зі збігом частот i=0 j=0;* f_{22} *- комірка зі збігом частот i=1 j=1;* f_{12} *та* f_{21} *– комірки з відсутніми збігами: i=0 j=1 та i=1 j=0;* $f_{.1}$ *– загальна частота в стовпчику 1;* $f_{.2}$ *– загальна частота в стовпчику 2;* $f_{1.}$ *– загальна частота в рядку 1; ;* f_2 *– загальна частота в рядку 2.*

Як приклад обчислення критерію χ^2 наведемо ще раз дані з таблиці 3.2 з нанесеним у них позначенням комірок відповідно до формули на рис. 3.8 (див. табл. 3.3). Для обчислення значення χ^2 потрібно скористатись формулою з правого боку.

Таблиця 3.3

Результати студентів за факторами MD та C методики 16PF

Значення за фактором MD (**X**)	Значення за фактором C (**Y**)		Сума частот у рядках
	Емоц. нестабільність C<5,5 (0 балів)	Емоц. нестабільність C>5,5 (1 бал)	
Адекватна самооцінка MD<5,5 (0 балів)	30 (67%) f_{11}	15 (33%) f_{12}	45 (100%) $f_{1.}$
Завищена самооцінка MD>5,5 (1 бал)	9 (19%) f_{21}	38 (81%) f_{22}	47 (100%) $f_{2.}$
Сума частот у стовпчиках	39 $f_{.1}$	53 $f_{.2}$	92

Обчислення значення χ^2 за формулою на рис. 3.8 для таблиць 2x2 праворуч показує: $\chi^2 = 21{,}25$ ($p<0{,}001$). Це означає, що розподіл частот достовірно відхиляється від теоретично-прогнозованого, є закономірним, а не випадковим.

Критерій χ^2 використовується для перевірки наявності статистично значущої залежності між двома категоріальними змінними. Якщо значення p-рівня менше за обраний рівень значущості, це означає, що зв'язок між змінними є значущим, тобто спостережувані частоти суттєво відрізняються від очікуваних за умов незалежності змінних. Проте саме значення χ^2 не показує силу цього зв'язку, оскільки воно залежить від розміру вибірки.

Коефіцієнт контингенції Пірсона (позначається як C) – це міра зв'язку між двома категорійними змінними, яка базується на статистиці χ^2-квадрат. Він визначається за формулою (рис. 3.9):

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

Рис. 3.9. Формула обчислення критерію контингенції С Пірсона

У формулі використано такі позначення: *χ^2 – значення критерію хі-квадрат, n – загальна кількість спостережень.*

Значення критерію С, обчислене для даних, наведених у табл. 3.2 та 3.3, становить 0,433 ($p < 0{,}001$). Значення коефіцієнта С варіюється від 0 до деякого максимального значення, яке залежить від розмірності таблиці спряженості, тому його часто нормалізують. Що більше значення С, то сильніший зв'язок між змінними, проте він не досягає 1 навіть за повної залежності. Через цю особливість коефіцієнт контингенції інколи доповнюють іншими показниками, такими як коефіцієнт Чупрової-Крамера, для більш коректного порівняння між таблицями різних розмірів.

Коефіцієнт Чупрової-Крамера (V) нормує значення χ^2 у діапазоні від 0 до 1 (див. рис. 3.10).

$$V = \sqrt{\frac{\chi^2/n}{\min(k-1, r-1)}}$$

Рис. 3.10. Формула обчислення коефіцієнта Крамера V

У формулі використано такі позначення: χ^2 *– значення критерію хі-квадрат, n – загальна кількість досліджуваних, k – кількість категорій у стовпцях, r – кількість категорій у рядках, (k−1, r−1) – мінімальне з двох значень (кількість ступенів свободи для меншого виміру).*

Якщо V близьке до 0, це означає слабкий зв’язок, а якщо значення наближається до 1, це свідчить про сильний зв'язок між змінними. Орієнтовна інтерпретація: $V < 0,1$ – дуже слабкий зв’язок; $0,1 \leq V < 0,30$ – слабкий; $0,3 \leq V < 0,50$ – помірний, $V \geq 0,5$ – сильний зв’язок. Для даних у табл. 3.3 коефіцієнт Крамера становить: $V = 0,48$, як і для φ.

Таким чином, описані коефіцієнти: – φ Пірсона, контингенції С Пірсона та Чупрової-Крамера V – становлять базовий інструментарій для кількісного аналізу взаємозв’язків між категоріальними змінними у психологічних дослідженнях. Вони дозволяють не лише виявити наявність статистично значущого зв'язку, а й оцінити його силу, що є важливим для інтерпретації отриманих результатів.

Вибір конкретного коефіцієнта залежить від типу даних та розмірності таблиці спряженості:

- φ найбільш адекватний для аналізу двох дихотомічних змінних,
- коефіцієнт контингенції С застосовується для квадратних таблиць із рівною кількістю категорій,
- коефіцієнт Чупрової-Крамера V є універсальним показником сили зв’язку для таблиць будь-якого розміру, оскільки нормалізує значення χ^2 в діапазоні від 0 до 1.

Загалом, застосування цих методів підвищує об’єктивність і точність досліджень у психології, дозволяючи систематично вивчати закономірності взаємодії між якісними ознаками та формувати на їх основі науково обґрунтовані висновки.

3.4. Регресійний аналіз

Регресія – це статистичний метод дослідження залежності між змінними, який дозволяє не лише встановити наявність зв'язку, але й побудувати математичну модель прогнозування значень однієї змінної на основі відомих значень інших змінних. На відміну від кореляційного аналізу, який лише оцінює силу та напрямок зв'язку, *регресійний аналіз* надає можливість кількісно передбачати зміни залежної змінної під впливом незалежних змінних. Залежно від кількості змінних та характеру їхнього взаємозв'язку розрізняють *лінійну та нелінійну, просту та множинну регресію*. Ключова перевага регресійного аналізу полягає в його здатності не лише описувати статистичні закономірності, а й слугувати інструментом прогнозування та моделювання.

У регресійному аналізі змінні поділяються на залежні та незалежні. *Залежна змінна* (Y) — це величина, яку намагаються пояснити або передбачити, і вона залежить від інших факторів. Наприклад, рівень тривожності може залежати від кількості годин сну, а успішність у навчанні – від рівня мотивації та когнітивних здібностей. *Незалежні змінні* (X) – це фактори, які можуть впливати на залежну змінну. Наприклад, рівень соціальної підтримки та фінансовий стан можуть пояснювати рівень психологічного добробуту, а регулярність фізичних вправ і здорове харчування – когнітивну ефективність [12, с. 68].

Є чимало різних видів регресійного аналізу, кожен із яких використовується залежно від характеру даних і поставлених завдань. У цьому посібнику буде розглянуто дві основні моделі*: просту* та *множинну лінійну регресію*, які є основою для багатьох інших регресійних методів.

Проста лінійна регресія застосовується для моделювання залежності між двома змінними: незалежною (предиктором) і залежною [20, с. 110]. Її можна описати рівнянням (див. рис. 3.11):

$$Y = \beta_0 + \beta_1 X_1$$

Рис. 3.11. Формула простої лінійної регресії

У формулі використано такі позначення: *Y – залежна змінна, X_1 – незалежна змінна, β_0 – константа (інтерсепт), β_1 – коефіцієнт регресії (нахил)*.

Проста лінійна регресія підходить для випадків, у яких між змінними існує лінійний зв'язок. Наприклад, можна використовувати цю модель для оцінки впливу витрачених годин навчання на підсумкову оцінку студента або залежності ціни товару від його ваги.

Приклад візуального зображення регресійної моделі представлено на рис. 3.12.

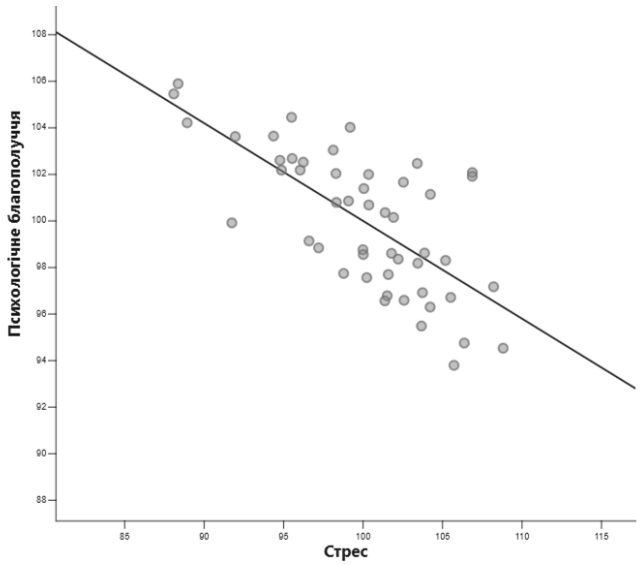

Рис. 3.12. Регресійна модель залежності психологічного благополуччя від стресу

Представлена регресійна модель описує залежність психологічного благополуччя від рівня стресу. Кожна точка на графіку відповідає окремому респонденту. По горизонтальній осі відкладено значення стресу, а по вертикальній – рівень психологічного благополуччя.

З діаграми видно, що спостерігається чітка негативна лінійна залежність: зі збільшенням рівня стресу рівень психологічного благополуччя знижується. Про це свідчить нахил лінії регресії вниз, що ілюструє загальну тенденцію.

Така залежність свідчить про те, що стрес є статистично значущим предиктором зниження психологічного добробуту. Інакше кажучи, підвищення рівня стресу в учасників дослідження пов'язується з погіршенням їхнього суб’єктивного самопочуття. Отриману регресійну модель можна використати для прогнозування рівня психологічного благополуччя за відомим значенням стресу.

Для оцінки якості лінійної регресії використовують низку основних метрик:

1. *Середньоквадратична помилка* (*англ*. Mean Squared Error, MSE), рис. 3.13

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Рис. 3.13. Формула обчислення середньоквадратичної помилки

У формулі використано такі позначення*: n – кількість спостережень (респондентів); y_i– фактичне значення для і-го спостереження; $\hat{y}$ – прогнозоване значення для і-го спостереження.*

Вимірює середню величину квадрата різниць між фактичними та передбаченими значеннями. Вона надає більшу вагу великим помилкам через піднесення до квадрата, що робить її чутливою до викидів. Менше значення MSE свідчить про кращу модель, але її інтерпретація може бути складною, оскільки ця метрика вимірюється в квадратичних одиницях залежної змінної.

2. *Корінь середньоквадратичної помилки* (*англ.* Root Mean Squared Error, RMSE), див. рис. 3.14:

$$RMSE = \sqrt{MSE}$$

Рис. 3.14. Формула обчислення RMSE

RMSE є квадратним коренем із MSE і повертає похибку в тих самих одиницях, що й прогнозована змінна. Це спрощує інтерпретацію, оскільки

RMSE показує середній розмір відхилення прогнозу від фактичного значення. Чим нижчий RMSE, тим точніша модель, а його чутливість до викидів залишається такою ж, як у MSE.

3. *Середня абсолютна помилка* (*англ. Mean Absolute Error*, MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Рис. 3.15. Формула обчислення MAE

У формулі використано такі позначення*: n – кількість спостережень (респондентів); y_i– фактичне значення для і-го спостереження; $\hat{y}_i$ – прогнозоване значення для і-го спостереження.*

MAE обчислює середню абсолютну різницю між прогнозами та фактичними значеннями. Вона менш чутлива до великих помилок, ніж MSE, оскільки не використовує піднесення до квадрата. MAE корисна для випадків, коли потрібно оцінити середню похибку в зрозумілих одиницях без значного впливу екстремальних значень.

4. *Коефіцієнт детермінації (R^2)*

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

Рис. 3.16. Формула обчислення коефіцієнта детермінації

У формулі використано такі позначення: *y_i– фактичне значення для і-го спостереження; $\hat{y}_i$ – прогнозоване значення для і-го спостереження; $\bar{y}$ – середнє значення фактичних спостережень.*

R^2 показує, яку частку варіації залежної змінної пояснює модель, порівняно з середнім значенням цієї змінної. Він змінюється від 0 (модель нічого не пояснює) до 1 (модель ідеально прогнозує залежну змінну). Високий R^2 означає, що модель добре підходить для даних, але не гарантує її узагальненості й застосування щодо інших вибірок.

5. *Скоригований коефіцієнт детермінації* (R^2_a)

$$R^2_a = 1 - \left(\frac{(1 - R^2)(n - 1)}{n - p - 1} \right)$$

Рис. 3.17. Формула скоригованого коефіцієнта детермінації

У формулі використано такі позначення: n *– кількість спостережень (респондентів);* p *– кількість предикторів (незалежних змінних) у моделі;* R^2 *– коефіцієнт детермінації.*

R^2_a враховує кількість незалежних змінних у моделі та вказує зайві предиктори, які не покращують пояснювальну здатність. Це важливо в множинній регресії, оскільки простий R^2 завжди зростає за додавання нових змінних, навіть якщо вони незначно впливають на залежну змінну.

Скоригований R^2_a допомагає об'єктивніше оцінювати якість моделі.

6. *F-критерій Фішера (F)*

$$F = \frac{MSR}{MSE} = \frac{RSS/k}{ESS/(n - k - 1)}$$

Рис. 3.18. Формула обчислення критерію Фішера

У формулі використано такі позначення: *RSS – сума квадратів поясненої варіації, ESS – сума квадратів залишків (непоясненої варіації), k – кількість предикторів (пояснювальних змінних), n – обсяг вибірки, MSR – середній квадрат регресії, MSE –* середньоквадратична помилка.

Критерій Фішера в регресії використовується для перевірки значущості всієї моделі або для порівняння моделей лінійної регресії. Він допомагає визначити, чи пояснюють незалежні змінні значну частину варіації залежної змінної. Результат тесту порівнюється з критичним значенням з таблиці розподілу Фішера, або перевіряється p-значення. Якщо p-значення < α, то відхиляємо нульову гіпотезу та вважаємо, що модель є статистично значущою і пояснює значну частину варіації залежної змінної.

Тема 3. Багатомірні методи в психології: прогнозування та класифікації

Лекція 4. Багатомірні методи в психології: прогнозування та класифікації

Мета: ознайомити студентів з призначенням та процедурою застосування в психології таких методів багатомірної статистики, як множинний регресійний аналіз, кластерний аналіз та факторний аналіз.

Призначення і класифікація методів багатомірної статистики. Множинна регресія як засіб прогнозування. Призначення кластерного аналізу. Методи кластерного аналізу: метод одиничного зв'язку, метод повного зв'язку, метод середнього зв'язку. Графічне представлення результатів кластерного аналізу: дендрограма. Алгоритм прийняття рішення про вибір кількості класів в кластерному аналізі. Кластерний аналіз об'єктів обробка на комп'ютері. Факторний аналіз: призначення. Аналіз головних компонентів і факторний аналіз. Проблема числа факторів. Методи факторного аналізу. Обертання факторів: ортогональні та косокутні. Факторні навантаження після варімакс-обертання. Послідовність факторного аналізу.

Основні поняття теми: множинна регресія, ієрархічний кластерний аналіз, дендрограма, матриця відмінностей, фактор, варімакс-обертання, навантаження змінної.

План

1. Багатомірні методи в психології
2. Множинний регресійний аналіз
3. Кластерний аналіз
4. Факторний аналіз
5. Структурне моделювання рівнянь (SEM)

4.1. Багатомірні методи в психології.

Багатомірні методи в психології використовуються для аналізу складних даних, що містять багато змінних, і дозволяють виявляти складні взаємозв'язки між ними. Ці методи дозволяють аналізувати, як кілька факторів одночасно впливають на поведінку, когнітивні процеси, емоції та інші психологічні явища. Основними багатомірними методами, що використовуються в психології, є:

1. *Множинна регресія.* Множинна регресія дозволяє вивчати, як декілька незалежних змінних одночасно впливають на одну залежну змінну. Метод дозволяє досліджувати, як демографічні (вік, стать) та психофізіологічні (рівень стресу) характеристики співвідносяться з показниками психологічного стану, такими як рівень депресивних симптомів або суб'єктивне відчуття добробуту.

2. *Кластерний аналіз.* Цей метод використовується для групування об'єктів або людей за подібними ознаками в кластери. Наприклад, кластерний аналіз може допомогти науковцям класифікувати осіб зі схожими симптомами або поведінковими патернами задля визначення типів розладів або психічних станів.

3. *Факторний аналіз.* Це статистичний метод, який використовується для виявлення латентних змінних (факторів), які пояснюють кореляції між багатьма спостережуваними змінними. Факторний аналіз може використовуватися для дослідження особистісних рис або когнітивних здібностей, щоб з'ясувати, які фактори (наприклад, екстраверсія або емоційна стабільність) лежать в основі різних поведінкових або психічних характеристик.

4. *Моделювання структурних рівнянь (SEM).* SEM дозволяє створювати складні моделі, що включають латентні змінні, та вивчати, як ці змінні взаємодіють між собою. Метод є ефективним для досліджень, які вивчають взаємозв'язки між різними психічними процесами, як-от вплив емоцій на поведінку чи мотивацію.

4.2. Множинний регресійний аналіз

Множинна регресія розширює просту лінійну регресію шляхом включення кількох незалежних змінних для прогнозування значення залежної змінної [19]. Цей метод дозволяє аналізувати складні залежності між багатьма факторами. Її рівняння виглядає так:

$$\mathbf{Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n}$$

Рис. 4.1. Формула множинної регресії

У формулі використано такі позначення: *Y – залежна змінна, X_1, X_2,…, X_n – незалежні змінні; β_0 – константа, β_1, β_2,…,β_n – коефіцієнти регресії*

Коефіцієнт β_n показує, наскільки в середньому зміниться залежна змінна Y за збільшення незалежної змінної X_i на одиницю (за умови, що інші незалежні змінні залишаються незмінними). Цей метод дозволяє будувати складніші моделі та враховувати одночасно кілька факторів. У психології множинна лінійна регресія може використовуватися, наприклад, для прогнозування рівня стресу на основі таких факторів, як робоче навантаження, рівень соціальної підтримки, фінансовий стан і фізична активність [14, с. 53].

Однак множинна регресія має свої обмеження. Одним із ключових викликів є *мультиколінеарність*, що передбачає високий рівень кореляції між незалежними змінними та призводить до нестабільності оцінок параметрів моделі [32]. Для подолання цієї проблеми часто використовують спеціальні методи регуляризації, такі як рідж-регресія або ласо-регресія.

Для доєднання незалежних змінних у регресійну модель використовують такі *методи відбору*: вихідний (Enter), прямий покроковий (Forward), зворотний покроковий (Backward).

Вихідний метод (Enter). Цей метод передбачає одночасне включення всіх незалежних змінних у регресійну модель. Усі предиктори вводяться в рівняння регресії одним кроком без жодного відбору змінних. Дослідник самостійно визначає, які змінні включити до моделі. Оцінюються коефіцієнти

для всіх предикторів, незалежно від їх статистичної значущості. Перевагами методу є простота застосування, можливість перевірки впливу всіх змінних одночасно, а також його доцільність за наявності теоретичного обґрунтування для включення всіх змінних. Недоліками є можливе включення незначущих предикторів, ризик мультиколінеарності за великої кількості змінних та потенційно менша прогностична здатність моделі.

*Прямий покроковий метод (Forward)***.** Цей метод починає з порожньої моделі і послідовно додає змінні по одній, базуючись на статистичних критеріях. Початкова модель не містить предикторів. На кожному кроці до моделі додається змінна з найбільшим статистичним внеском. Процес додавання змінних продовжується, доки не буде досягнуто певного критерію зупинки. Змінна додається, якщо її p-значення менше від установленого порогу (зазвичай 0,05). Перевагами методу є відбір найбільш значущих предикторів, зменшення ризику мультиколінеарності, а також отримання більш економічної й зрозумілої моделі. Недоліками є можливість пропуску важливих змінних, неврахування взаємодії між ними та залежність від порядку включення змінних.

Зворотний покроковий метод (Backward). Цей метод починає з повної моделі, що включає всі предиктори, і послідовно видаляє найменш значущі змінні. Початкова модель містить усі потенційні предиктори. На кожному кроці з моделі видаляється змінна з найменшим статистичним внеском. Змінна видаляється за умови, якщо її p-значення перевищує встановлений поріг. Процес завершується, коли в моделі залишаються лише статистично значущі змінні. Перевагами методу є менший ризик пропустити важливі взаємодії між змінними, можливість виявити «замаскований» вплив деяких предикторів та ефективність для великих наборів даних. Недоліками можна назвати потребу в більших обчислювальних ресурсах для початкової повної моделі, можливі проблеми зі збіжністю за сильної мультиколінеарності та чутливість до викидів у початковій моделі.

Оцінка якості моделі здійснюється за допомогою тих самих метрик, що й за одномірної лінійної регресії (див розділ 3.3).

4.3. Кластерний аналіз

Кластерний аналіз – це метод багатовимірного статистичного аналізу, що дозволяє класифікувати об'єкти у відносно однорідні групи (кластери) на основі набору вимірюваних ознак. На відміну від багатьох інших статистичних методів, кластерний аналіз зазвичай використовується, коли відсутні апріорні гіпотези щодо групування і дослідження має пошуковий характер.

Кластерний аналіз базується на кількох *ключових принципах*. По-перше, об'єкти всередині одного кластера повинні бути максимально подібними між собою. По-друге, об'єкти з різних кластерів повинні максимально відрізнятися один від одного. По-третє, групування об'єктів відбувається на основі декількох характеристик одночасно, що забезпечує багатовимірність аналізу.

Для визначення подібності або відмінності між об'єктами використовуються різні міри *відстані*. Найпоширенішою є *евклідова відстань* – геометрична метрика, що обчислює довжину прямої між точками в багатовимірному просторі. Також застосовуються *манхеттенська відстань* (відстань міських кварталів), яка є сумою абсолютних різниць координат; *відстань Махаланобіса*, що враховує кореляції між змінними; *косинусна подібність*, яка вимірює косинус кута між векторами ознак, та к*ореляційні міри*, що використовують коефіцієнт кореляції Пірсона або Спірмена як міру подібності.

Є два основних типи *алгоритмів кластеризації*: ієрархічні та неієрархічні методи. *Ієрархічні методи* створюють ієрархічне дерево (*дендрограму*, див. рис. 4.2), що показує процес об'єднання кластерів. Вони поділяються на агломеративні методи, які починають з окремих об'єктів як окремих кластерів і послідовно об'єднують найближчі кластери, та дивізимні

методи, які починають з одного кластера, що містить усі об'єкти, і послідовно розділяють їх.

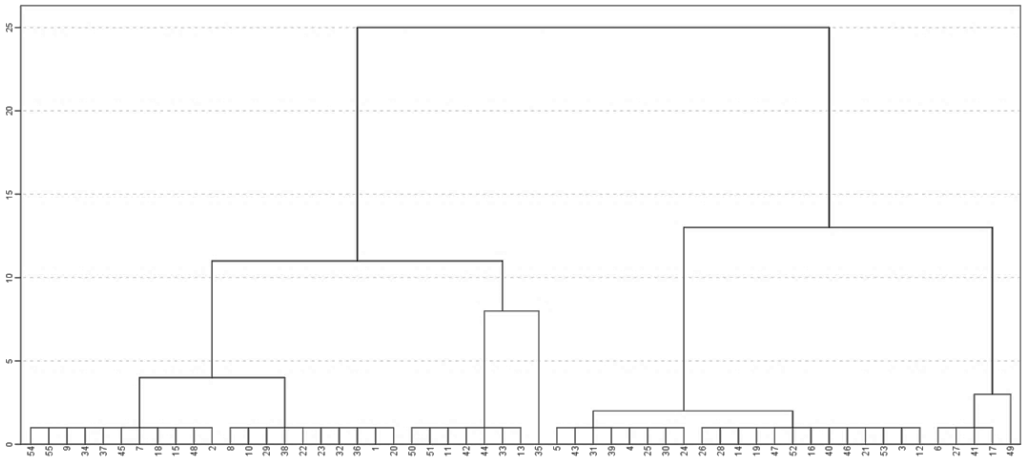

Рис. 4.2. Приклад дендрограми

Неієрархічні методи включають такі популярні алгоритми, як k-середніх (k-means), який вимагає попереднє визначення кількості кластерів k і ітеративно призначає об'єкти до найближчого центроїда (див. рис 4.3).

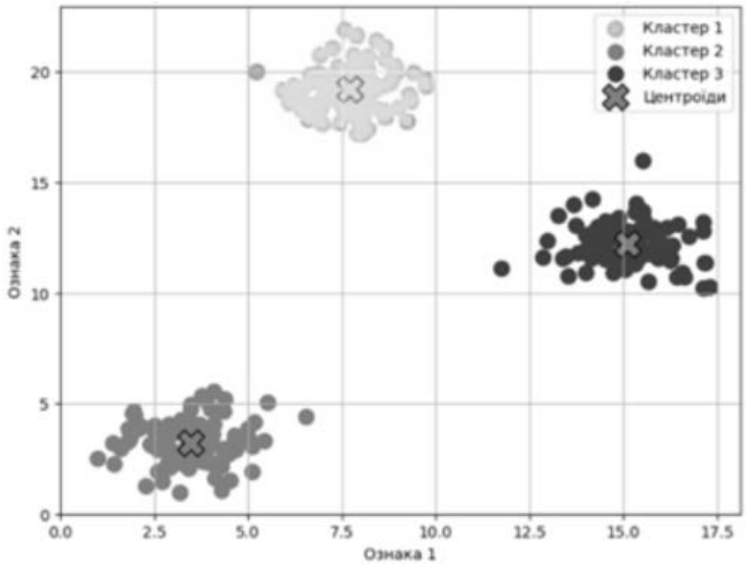

Рис. 4.3. Кластерний аналіз методом k-means

Для визначення оптимальної кількості кластерів використовуються різні методи. *Метод ліктя* (*англ. Elbow Method*) аналізує графік залежності внутрішньокластерної дисперсії від кількості кластерів (*див. рис. 4.4*). За збільшення кількості кластерів, внутрішньокластерна дисперсія (WCSS)

зменшується, однак після певного значення k (у цьому випадку після k = 2) зменшення дисперсії стає менш значним, адже додавання нових кластерів не дає істотного покращення в межах розподілу даних.

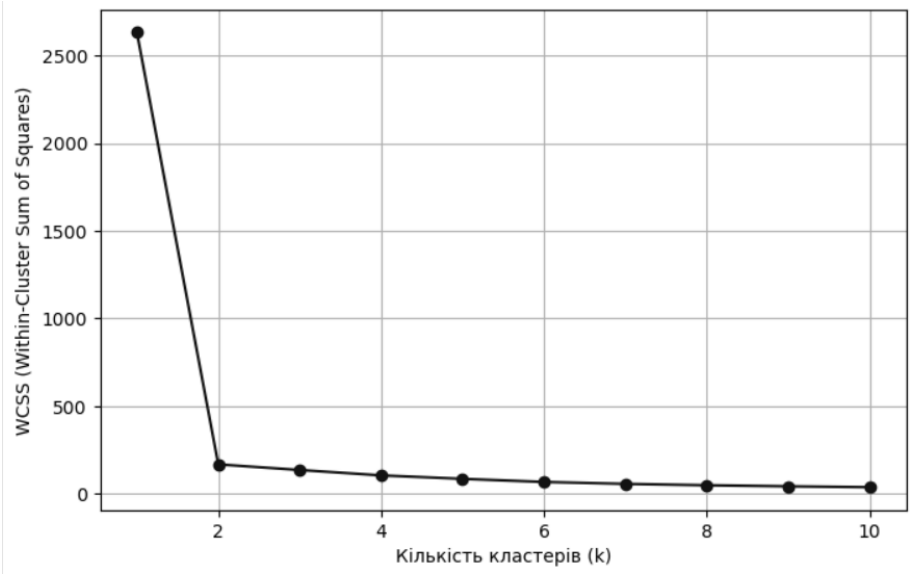

Рис. 4.4. Метод ліктя для вибору оптимальної кількості кластерів

Метод силуету – це техніка для оцінки якості кластеризації, що дозволяє виміряти, наскільки добре кожен об'єкт був кластеризований. Силуетний коефіцієнт для кожної точки розраховується за такою формулою:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Рис. 4.5 Формула обчислення силуетного коефіцієнта

У формулі використано такі позначення*: a(i) – середня відстань між і-м об'єктом та всіма іншими об'єктами в його кластері (внутрішньокластерна відстань); b(i) – найменша середня відстань між і-м об'єктом та об'єктами в іншому кластері (найближчий сусідній кластер); max(a(i),b(i)) – нормалізаційний знаменник, що забезпечує значення в межах [-1, 1]*;

Середнє значення силуетного коефіцієнта по всіх точках використовується для вибору оптимального K. Значення силуету варіюється від -1 до 1, де 1 означає, що об'єкт міститься в правильному кластері, а значення, ближче до -1, — що об'єкт було неправильно класифіковано.

Індекс Девіса-Болдіна оцінює якість кластеризації, враховуючи внутрішню дисперсію кластерів і відстань між їхніми центрами. Чим менший цей індекс, тим краща кластеризація.

$$DB = \frac{1}{k} \sum_{i=1}^{k} \max_{j \neq i} \left(\frac{s_i + s_j}{d_{ij}} \right)$$

Рис. 4.6 Формула обчислення індексу Девіса-Болдіна

У формулі використано такі позначення: *k – кількість кластерів,* s_i *– середня відстань між точками в кластері i,* d_{ij} *– відстань між центрами кластерів i та j.*

Критерій Калінського-Харабаза оцінює якість кластеризації за допомогою співвідношення між міжкластерною дисперсією та внутрішньокластерною дисперсією. Чим вищий цей індекс, тим краща кластеризація.

$$CH = \frac{\mathrm{Tr}(B_k)/(k-1)}{\mathrm{Tr}(W_k)/(n-k)}$$

Рис. 4.7 Формула обчислення *критерію Калінського-Харабаза*

У формулі використано такі позначення: B_k *– матриця міжкластерної дисперсії,* W_k *– матриця внутрішньокластерної дисперсії, k – кількість кластерів, n – загальна кількість зразків.*

Кластерний аналіз широко застосовується в різних галузях. У маркетингу він використовується для сегментації споживачів за поведінкою та уподобаннями, а в біології допомагає класифікувати організми та аналізувати генетичні дані. У медицині використовується в діагностиці захворювань та класифікації симптомів. У психології допомагає виявляти типи особистості та класифікувати поведінкові патерни. В економічній географії дає змогу групувати країни за економічними показниками, тоді як в інформатиці застосовується для розпізнавання образів, аналізу тексту та пошуку інформації.

Кластерний аналіз не вимагає попередніх припущень про структуру даних, дозволяє виявляти приховані структури в даних та спрощує інтерпретацію багатовимірних даних. Однак метод має і певні обмеження: результати можуть бути нестабільними в разі зміни алгоритму або параметрів, існує суб'єктивність у виборі кількості кластерів. Загалом, метод чутливий до викидів та шуму в даних, а різні алгоритми можуть давати різні результати на одних і тих самих даних.

Попри обмеження, кластерний аналіз залишається важливим інструментом дослідження складних даних, оскільки дозволяє виявляти приховані закономірності та структури в багатовимірному просторі ознак. Він допомагає дослідникам знаходити певні доречні групування даних, що можуть бути неочевидними в разі безпосереднього спостереження або за умови використання інших аналітичних методів.

Прикладом використання кластерного аналізу є дослідження А. Блачніо, А. Пзепіорката та І. Пантік присвячене зв’язку між залежністю від Facebook, самооцінкою та задоволеністю життям [15]. У дослідженні взяли участь 381 користувач соціальної мережі з Польщі. Дослідники застосували k-середніх кластеризацію, щоб виявити групи користувачів зі схожими профілями за двома змінними: рівнем залежності від Facebook та інтенсивністю його використання. У результаті було виділено три кластери: звичайні, проблемні та залежні користувачі (див. табл. 4.1)

Вивчення особливостей виділених груп та інших змінних використаних в дослідженні, таких як самооцінка та задоволення життям. Аналіз отриманих даних показав, що звичайні користувачі (низькі показники як залежності, так і інтенсивності використання) характеризувалися вищим рівнем самооцінки та життєвої задоволеності. На думку авторів, це вказує на здорове використання соціальних мереж як інструменту для спілкування.

До проблемних користувачів були віднесені особи з високою інтенсивністю використання Facebook, але середнім рівнем залежності. Попри знижену самооцінку, вони демонстрували порівняно високий рівень задоволеності

життям. На думку авторів, отримані особливості можна пояснити як своєрідний “ефект ейфорії”, характерний для початкової фази формування залежності: інтенсивне користування ще приносить позитивні емоції, хоча вже з’являються труднощі з контролем поведінки.

Таблиця 4.1

Кластери користувачів Facebook за рівнем залежності та інтенсивності

Показник	Звичайні користувачі (n = 212)	Проблемні користувачі (n = 57)	Залежні користувачі (n = 123)
	M	M	M
Facebook залежність	-0,73	0,24	1,89
Facebook інтенсивність	-0,30	0,48	0,73

До залежних користувачів були віднесені особи з високими показниками шкали залежності та середніми показники інтенсивності використання. Ця група вирізнялася нижчим рівнем як самооцінки, так і життєвої задоволеності. Користування Facebook стало для них проблемним і негативно впливало на якість життя [15].

Як видно з наведеного дослідження, використання кластерного аналізу дозволило дослідникам розбити вибірку на три групи на основі значень за шкалами залежності від соціальної мережі та інтенсивності її використання. В подальшому особливості виділених груп були якісно описані, для кожної групи вивчені особливості самооцінки та задоволення життям.

4.4. Факторний аналіз

Факторний аналіз є статистичним методом, який використовується для виявлення прихованих взаємозв'язків між змінними та їх групування у фактори. Цей метод спрямований на редукцію даних – зменшення кількості змінних шляхом об'єднання корельованих змінних у меншу кількість

латентних факторів, які пояснюють значну частину дисперсії первинних змінних.

Основна ідея факторного аналізу полягає в тому, що спостережувані змінні можуть бути частково пояснені меншою кількістю неспостережуваних (латентних) змінних, які називаються факторами. Ці фактори відображають фундаментальні конструкти, що лежать в основі досліджуваного явища. Наприклад, відповіді на різні запитання тесту інтелекту можуть відображати базові когнітивні здібності, такі як вербальний інтелект, просторове мислення, математичні здібності тощо [31, с. 4].

Математично факторний аналіз представляє кожну змінну як лінійну комбінацію факторів плюс унікальну дисперсію (помилка), що не пояснюється факторами . Для кожного фактора модель виглядає так (рис. 4.8):

$$X_i = \lambda_{i1}F_1 + \lambda_{i2}F_2 + \cdots + \lambda_{ik}F_k + E_i$$

Рис. 4.8 Формула обчислення значення змінної як залежної величини, яка визначається впливом латентних факторів

У формулі використано такі позначення: $\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{ik}$ *— це факторні навантаження, які можна розглядати як коефіцієнти лінійної регресії між кожною змінною* X_i *й латентними факторами* $F_1, F_2, \ldots, F_k$*;* E_i *— залишковий компонент, який містить ту частину варіації змінної* X_i*, яку не можна пояснити факторами.*

У процесі факторного аналізу дослідники визначають оптимальну кількість факторів, використовуючи різні критерії, такі як *критерій Кайзера* (власні значення більші, ніж 1) та *критерій «кам'янистого осипу»* (див. рис. 4.9).

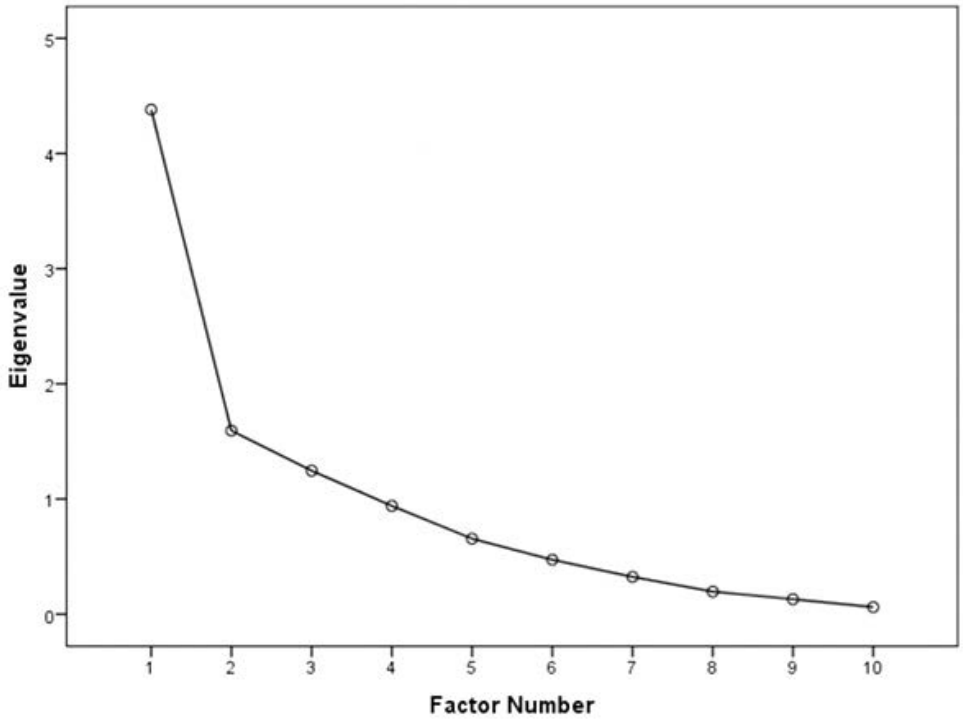

Рис. 4.9. Приклад графіку «кам’яного осипу»

Відповідного до наведеного графіку, оптимальна кількість факторів визначається в точці його «зламу» – це місце, де крива, що спадає, різко змінює кут нахилу і стає майже горизонтальною. На графіку «злам» спостерігається приблизно на факторі 2. Це означає, що, імовірно, оптимальна кількість факторів 2±1, оскільки після другого фактору власні значення знижуються повільніше.

Після визначення кількості факторів здійснюється їх інтерпретація на основі факторних навантажень – коефіцієнтів кореляції між змінними та факторами.

Є два основних підходи до факторного аналізу: *дослідницький* (EFA) та *підтверджувальний, або конфірматорний* (CFA). Дослідницький факторний аналіз використовується, коли дослідник не має чітких гіпотез щодо структури факторів і прагне виявити приховані конструкти (приклад – табл. 4.2).

Таблиця 4.2

Матриця повернутих компонентів (метод: PCA, обертання: Varimax with Kaiser Normalization)

	Фактори					
	1	2	3	4	5	6
q15	,838					
q14	,783					
q16	,781					
q13	,757					
q23		,829				
q21		,804				
q24		,774				
q22		,752				
q10			,817			
q9			,735			
q11			,705			
q12			,698			
q2				,805		
q3				,766		
q4				,622		
q1				,582		
q18					,762	
q19					,695	
q20					,641	
q17					,575	
q6						,829
q5						,760
q7						,714
q8						,486

Відповідно до прикладу, виявлено 6 факторів, які є латентними змінними, що описують різні аспекти конструкта чи групи змінних.

Кожен фактор має сильні навантаження на певні змінні, що дозволяє зробити висновок, що ці групи змінних мають спільну варіацію і можна описати їх за допомогою відповідних факторів.

Підтверджувальний факторний аналіз застосовується для перевірки конкретних гіпотез щодо факторної структури, коли дослідник має теоретичні припущення про кількість факторів та їх зв'язки зі змінними.

Різноманіття *методів факторного* аналізу, від класичного методу головних компонент до сучасних байєсівських підходів, надає широкі можливості для роботи з даними різної природи та якості [19; 31]. Кожен із

цих методів має свої особливості, переваги та обмеження, які важливо враховувати при аналізі.

Метод головних компонент (Principal Components Analysis, PCA) є одним із найпопулярніших підходів до факторного аналізу. PCA перетворює набір потенційно корельованих змінних у набір лінійно некорельованих змінних, які називаються головними компонентами. Перша головна компонента пояснює найбільшу частку дисперсії даних, кожна наступна компонента пояснює максимальну дисперсію, що залишилася. У SPSS цей метод використовується за замовчуванням у процедурі факторного аналізу. PCA особливо корисний для зменшення розмірності даних та ідентифікації прихованих структур.

Метод головних осей (Principal Axis Factoring, PAF) фокусується на спільній дисперсії й використовує для її оцінки ітеративний процес. Метод краще підходить, коли мета полягає у виявленні латентних конструкцій, що лежать в основі вимірюваних змінних, особливо коли дані не задовольняють умови багатовимірного нормального розподілу.

Метод максимальної правдоподібності (Maximum Likelihood, ML) – оцінює параметри факторної моделі шляхом максимізації імовірності спостереження наявних кореляцій, якщо вибірка походить з багатовимірного нормального розподілу. Цей метод дозволяє перевіряти статистичні гіпотези про кількість факторів та будувати довірчі інтервали для факторних навантажень. ML найкраще працює з великими вибірками і вимагає багатовимірного нормального розподілу даних.

Узагальнений метод найменших квадратів (Generalized Least Squares, GLS) – оптимізує факторну структуру, надаючи більшої ваги змінним з вищою унікальністю. GLS мінімізує зважені квадрати різниць між спостереженими та відтвореними кореляціями. Цей метод є менш обмежувальним щодо припущення про нормальність розподілу, ніж метод максимальної правдоподібності, тому може бути кращим вибором для даних, які дещо відхиляються від нормального розподілу.

Метод незваженого найменшого квадрата (Unweighted Least Squares, ULS) – мінімізує суму квадратів різниць між спостереженими та відтвореними кореляціями, не використовуючи вагові коефіцієнти. ULS не вимагає припущення про нормальний розподіл даних і добре працює з порядковими змінними. Цей метод особливо корисний, коли розмір вибірки невеликий або коли дані значно відхиляються від нормального розподілу.

Альфа-факторний аналіз (Alpha Factoring) – розглядає змінні у наборі даних як вибірку з більшого набору потенційних змінних. Метод максимізує надійність факторів, використовуючи коефіцієнт альфа Кронбаха. Альфа-факторинг особливо корисний, коли дослідник зацікавлений у виявленні факторів, які мають високу внутрішню узгодженість і могли б бути надійними під час застосування до іншої вибірки змінних з тієї ж основної популяції.

Факторизація образів (Image Factoring) – ґрунтується на теорії образів, яка розрізняє образи (частини змінних, що можуть бути передбачені іншими змінними) та антиобрази (частини, що не можуть бути передбачені). Цей метод фокусується на кореляціях образів, які представляють регресійні прогнози кожної змінної на основі всіх інших змінних. Факторизація образів може бути корисною, коли дослідник зацікавлений у спільній дисперсії змінних та хоче мінімізувати вплив унікальної дисперсії.

Важливим кроком у факторному аналізі є вибір *методу обертання факторів*, який допомагає отримати більш інтерпретовану факторну структуру. *Ортогональні обертання* (наприклад, Varimax) передбачають незалежність факторів, тоді як *косокутні обертання* (наприклад, Promax, Direct Oblimin) дозволяють факторам корелювати між собою, що часто більш реалістично відображає дійсність, особливо в психологічних та соціальних дослідженнях.

Факторний аналіз знаходить широке застосування в різних галузях. У психометрії він використовується для розробки та валідації психологічних тестів, виявлення основних вимірів особистості, здібностей та установок. У маркетингових дослідженнях він допомагає виявити ключові фактори, що

впливають на споживчу поведінку. В економіці він застосовується для аналізу економічних індикаторів та виявлення прихованих економічних факторів. У соціології факторний аналіз допомагає виявити базові соціальні установки та цінності.

Перед проведенням факторного аналізу *необхідно перевірити кілька припущень:* наявність достатньої кількості спостережень (зазвичай рекомендується мінімум 5-10 спостережень на змінну), наявність кореляцій між змінними (перевіряється за допомогою кореляційної матриці, *критерію сферичності Бартлетта* та *міри адекватності вибірки Кайзера-Мейєра-Олкіна*), багатовимірну нормальність даних.

Хоча факторний аналіз є потужним інструментом, він має певні *обмеження*. Інтерпретація факторів часто суб'єктивна і залежить від досвіду дослідника. Різні методи екстракції факторів та обертання можуть давати різні результати. Крім того, факторний аналіз чутливий до викидів, пропущених даних та порушень його основних припущень.

Попри ці обмеження, факторний аналіз залишається одним із найважливіших методів для розуміння складних взаємозв'язків між змінними та виявлення прихованих структур у даних. Він допомагає дослідникам спростити складні набори даних, виявити фундаментальні конструкти та розробити обґрунтовані теорії в різних галузях науки.

4.5. Структурне моделювання рівнянь (SEM)

Структурне моделювання рівнянь (Structural Equation Modeling, SEM) – це статистичний метод, що поєднує факторний аналіз і множинну регресію. З його допомогою можна досліджувати складні причинно-наслідкові зв'язки між спостережуваними та латентними змінними. SEM вважається одним із найбільш комплексних і гнучких підходів до аналізу складних взаємозв'язків у даних, особливо в соціальних, поведінкових і економічних науках [19, с. 731].

На відміну від традиційних статистичних методів, SEM дозволяє одночасно аналізувати цілу мережу взаємозв'язків між змінними, перевіряти складні теоретичні моделі, враховувати похибки вимірювання та вивчати як прямі, так і непрямі ефекти між змінними [23]. Це робить SEM особливо цінним для дослідження складних психологічних, соціальних та економічних феноменів, які не можна адекватно описати простішими статистичними методами.

SEM складається з двох основних компонентів: вимірювальної (*англ. measurement model*) та структурної моделі (*англ. structural model*). Вимірювальна модель являє собою *конфірматорний факторний аналіз*, який визначає зв'язки між латентними змінними (неспостережуваними конструктами) та їхніми індикаторами (спостережуваними змінними). Структурна модель визначає причинно-наслідкові зв'язки між латентними змінними, подібно до шляхового аналізу або системи регресійних рівнянь.

Процес застосування SEM зазвичай включає кілька етапів. Спочатку дослідник формулює теоретичну модель, що описує гіпотетичні взаємозв'язки між змінними. Потім цю модель слід виразити у вигляді системи рівнянь та/або за допомогою шляхової діаграми. Наступним кроком є збір даних та оцінка параметрів моделі. Для оцінки параметрів найчастіше використовуються методи максимальної правдоподібності, узагальнених найменших квадратів або зважених найменших квадратів.

Після оцінки параметрів модель оцінюється на відповідність даним за допомогою різних індексів, таких як χ^2 тест, CFI, TLI, RMSEA та SRMR. χ^2 тест використовується для перевірки гіпотези про відповідність моделі даним. Низькі значення χ^2 і відсутність статистичної значущості ($p > 0,05$) свідчать про значну відповідність між спостережуваними та очікуваними значеннями (різниця між ними є статистично несуттєвою). На практиці показник χ^2 зазвичай доповнюють іншими індексами якості моделі, оскільки χ^2-тест є чутливим до розміру вибірки і може виявляти статистично значущі відмінності навіть за незначних відхилень у великих вибірках [19; 23].

CFI (*англ. Comparative Fit Index*) – порівняльний індекс відповідності, TLI (*англ. Tucker-Lewis Index*) – модифікована версія CFI, значення, вищі за 0,90, вказують на прийнятну модель, а вищі за 0,95 – на достатню відповідність моделі. RMSEA (*англ. Root Mean Square Error of Approximation*) – середньоквадратична помилка апроксимації, SRMR (*англ. Standardized Root Mean Square Residual*) – стандартизований середньоквадратичний залишок. Значення < 0,08 за цими показниками вказують на прийнятну модель, < 0,05 – на достатню відповідність, а > 0,08 – на слабку відповідність.

Якщо модель не має достатньої відповідності даним, дослідник може модифікувати її на основі теоретичних міркувань та індексів модифікації, які вказують на можливі шляхи покращення моделі.

SEM має багато переваг порівняно з традиційними статистичними методами. Зокрема, він дозволяє враховувати похибки вимірювання, що робить результати надійнішими, а також дає можливість тестувати складні теоретичні моделі з багатьма змінними та шляхами впливу. Крім того, SEM дозволяє порівнювати альтернативні моделі, що допомагає вибрати найбільш відповідну теоретичну структуру.

SEM широко застосовується в різних галузях науки. У психології цей метод використовується для дослідження особистості, когнітивних процесів, розвитку, психологічного благополуччя. В освіті SEM допомагає вивчати фактори, що впливають на академічну успішність, мотивацію до навчання, ефективність освітніх інтервенцій. У маркетингу метод застосовується для аналізу споживчої поведінки, лояльності до бренду, впливу реклами. В економіці та фінансах SEM використовується для моделювання складних економічних взаємозв'язків, ринкових механізмів, фінансових рішень.

Розглянемо приклад застосування SEM для вибору та перевірки оптимальної факторної структури опитувальника, така процедура також називається конфірматорним факторним аналізом. У своїй статті О. Вельдбрехт, Н. Зінченко та Н. Тавровецька здійснили адаптацію шкали самооцінки Розенберга та

дослідили її психометричні характеристики [4]. Зокрема, автори порівнювали дві альтернативні факторні моделі: одновимірну та двовимірну (див. табл. 4.3).

Таблиця 4.3

Порівняння факторних моделей шкали Розенберга

Опис моделі	χ^2	df	P	CFI	TLI	SRMR	RMSEA
Модель 1. Одновимірне рішення	152	35	< .001	0.909	0.883	0.052	0.101
Модель 2. Двовимірне рішення: позитивно та негативно сформульовані пункти	122	34	< .001	0.931	0.909	0.047	0.089

Порівняння наведених індексів показує, що «Модель 2» (двовимірна) демонструє кращу відповідність емпіричним даним за більшістю критеріїв. Так, значення χ^2 у «Моделі 2» є нижчим (122 проти 152), що свідчить про менший розрив між теоретичною та фактичною матрицею коваріацій. Індекс CFI підвищується з 0.909 до 0.931 — обидва значення перевищують пороговий рівень 0.90, однак у «Моделі 2» він є вищим. Показник TLI зростає з 0.883 до 0.909, переходячи у зону прийнятної відповідності (>0.90). Значення SRMR зменшується з 0.052 до 0.047, що вказує на зниження середніх залишкових розбіжностей і досягнення рівня гарної відповідності (<0.05). Нарешті, показник RMSEA знижується з 0.101 (низька відповідність) до 0.089, тобто опускається до верхньої межі прийнятного діапазону.

Узагальнюючи, «Модель 2» демонструє кращу узгодженість з емпіричними даними, насамперед за показниками CFI, TLI та SRMR. Водночас значення RMSEA у двофакторному рішенні все ще перебуває на межі прийнятності, що не дозволяє вважати модель оптимальною. Наведені вище критерії вказують на більшу прийнятність двофакторної структури опитуваника (поділу на позитивно та негативно сформульовані пункти), проте запропонована модель потребує додаткової перевірки та, можливо, подальшої модифікації самої методики [4].

Структурне моделювання (SEM) також використовується для побудови моделей психологічних структур їх взаємодії з іншими факторами. Як приклад такого застосування можна розглянути дослідження проведене О. Савченко,

Л. Колесніченко, В. Чужиковою та О. Береженною [13]. Гіпотеза дослідження полягала в тому, що рівень переживання цінності власного життя має статистично значущі зв'язки з проявами резильєнтності та використанням копінг-поведінки, причому ці зв'язки опосередковуються рівнем ясності їхнього самоусвідомлення (медіатор).

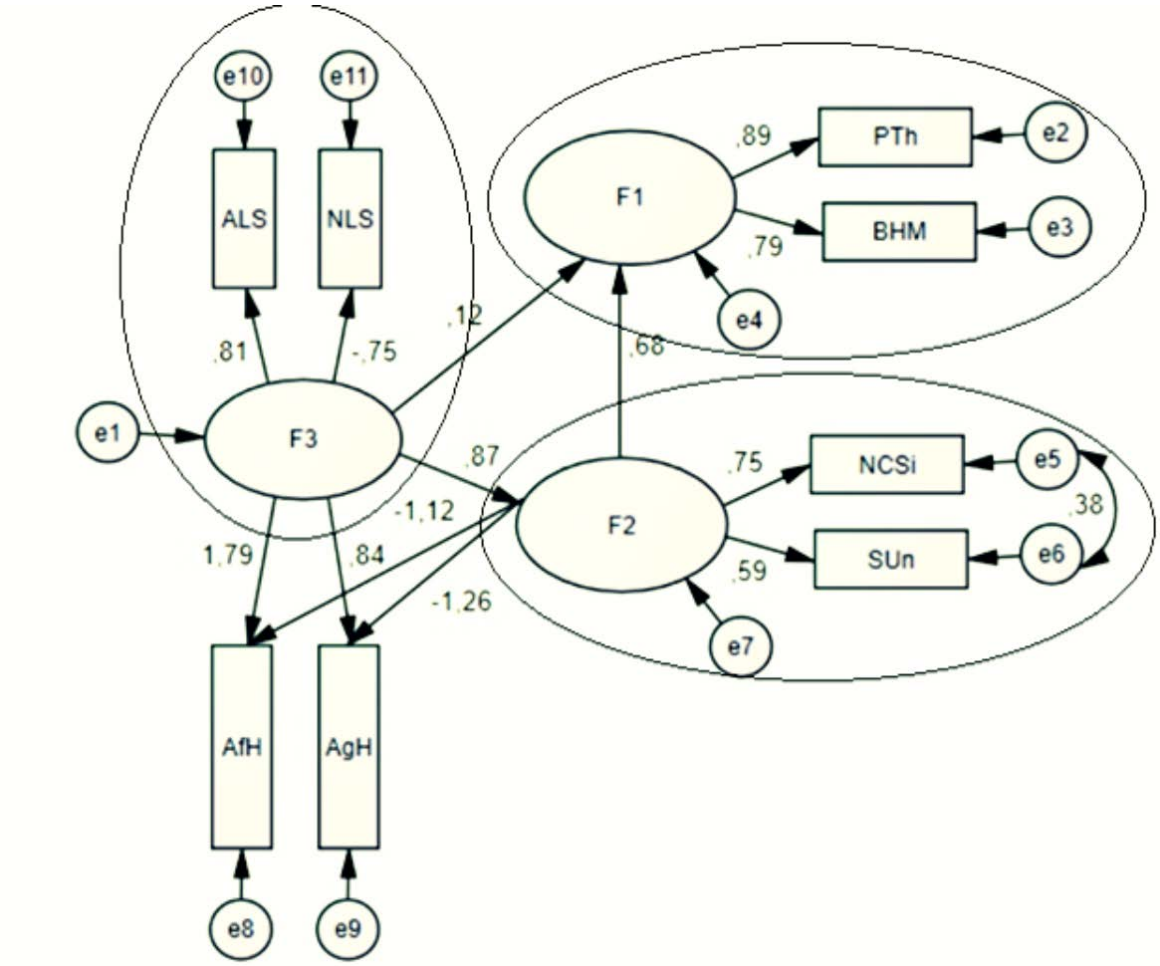

Рис 4.10. Структурна модель зв'язку міри цінності життя з резильєнтністю та стилями гумору

Як результат дослідження автори запропонували модель (рис. 4.10), яка включає три фактори другого порядку:

- F3 – цінність життя (предиктор), який утворюють змінні ALS – активний пошук цінності життя та NLS – знецінення життя;
- F2 – ясність самоусвідомлення (медіатор), що складається з показників NCSi – несуперечливість образу Я та SUn – розуміння себе;

- F1 – резильєнтність (залежна змінна), яка об’єднує PTh – позитивне мислення та BHM – мислення, орієнтоване на віру та надію.
- Окремо відображено зв’язки з AfH – афіліативним стилем гумору та AgH – агресивним стилем гумору.

Стрілками позначено напрямки впливів, біля яких наведено стандартизовані бета-коефіцієнти. Аналіз бета-коефіцієнтів свідчить, що цінність життя має слабкий прямий вплив (β = 0,12) на резильєнтність, а основний ефект реалізується опосередковано – через опосередкування (медіацію) зі сторони ясності усвідомлення (β = 0,87 та β = 0,68).

Розглянемо детально фактор «резильєнтність», який об’єднує позитивне мислення (PTh, β = 0,89) та мислення, орієнтоване на віру та надію (BHM, β = 0,79). Високі факторні навантаження у цьому випадку означають сильний взаємозв’язок. Пояснена дисперсія (частина загальної мінливості даних, яка описується певним фактором) вимірюється через β^2, у нашому випадку: позитивне мислення – 0,79 ($0,89^2$) та мислення, орієнтоване на віру та надію – 0,62 ($0,79^2$). Іншу частину дисперсію позначають похибкою вимірювання, що вимірюється через формулу зображену нижче:

$$e = 1 - \beta^2$$

Рис 4.11. Формула похибки вимірювання

У формулі використано такі позначення: e – похибка вимірювання; β – факторне навантаження.

У цьому випадку похибка вимірювання позитивного мислення дорівнює 0,21,а для мислення, орієнтованого на віру та надію – 0,38. Похибка вимірювання позначає той відсоток варіації, який модель не може пояснити й до якого може відноситись статистичний шум та вплив інших змінних, не врахованих в моделі. Отже, латентний конструкт «резильєнтність» пояснює 79% дисперсії фактору «позитивне мислення» та 62% дисперсії «мислення, орієнтоване на

віру та надію». Залишкова дисперсія (21% та 38% відповідно) не пояснюються поточною структурною моделлю.

Водночас SEM має певні обмеження і вимоги. Він потребує відносно великих вибірок для здобуття стабільних і надійних результатів. Хоча SEM передбачає нормальний розподіл даних, є методи, здатні працювати за умов його відсутності. Крім того, SEM є конфірматорним, а не дослідницьким методом, тобто потребує чіткої теоретичної основи перед аналізом [19].

З розвитком обчислювальних потужностей та статистичного програмного забезпечення, SEM став більш доступним для дослідників. Сучасні програми, такі як AMOS, Mplus, LISREL, lavaan (в R) та інші, спростили побудову та оцінку структурних моделей, що сприяло поширенню методу в різних галузях науки.

Останні розробки в SEM включають байєсівський підхід до оцінки параметрів, методи аналізу категоріальних даних, багаторівневі структурні моделі для ієрархічних даних, моделі зростання для лонгітюдних досліджень. Ці інновації ще більше розширили можливості SEM для вирішення складних дослідницьких завдань у сучасній науці.

Тема 4. Перевірка статистичних гіпотез

Лекція 5. Перевірка статистичних гіпотез: параметричні та непараметричні критерії

Мета: сформувати у студентів уявлення про логіку перевірки статистичних гіпотез в метричних та неметричних шкалах. Ознайомити студентів зі способом перевірки нульової гіпотези для коефіцієнта кореляції Пірсона. Створити уявлення про теоретичні розподіли: хі-квадрат, t та F розподіли як основу для перевірки статистичних гіпотез. Ознайомити зі способом перевірки H_0 гіпотези для критерію хі-квадрат. Розглянути t-критерії Стюдента; F-критерій Фішера; непараметричні критерії: U Манна-Уїтні, H Краскела-Уолліса, T Уілкоксона.

Основні поняття теми: центральна гранична теорема, довірчий інтервал, стандартна помилка середнього генеральної сукупності, помилка першого роду, помилка другого роду, нормальний розподіл, хі-квадрат розподіл; t – розподіл, F-розподіл, критерій χ^2; t-критерій Стюдента для незалежних вибірок; t-критерій Стюдента для пов'язаних вибірок; F-критерій Фішера; U- критерій Манна-Уітні; H-критерій Краскела-Уолліса; T-критерій Уілкоксона.

План

1. Центральна гранична теорема, довірчий інтервал середнього генеральної сукупності.
2. Порівняння двох вибірок на основі обчислення довірчих інтервалів
3. Критерій χ^2 – Пірсона
4. t – критерій Стюдента
5. F-критерій Фішера для порівняння дисперсії двох незалежних вибірок
6. Непараметричні критерії

5.1. Центральна гранична теорема, довірчий інтервал середнього генеральної сукупності

Центральна гранична теорема є однією з основних, на яких ґрунтується теорія статистичного висновку. На її основі легко зрозуміти основну ідею перевірки статистичних гіпотез. Теорему можна сформулювати так: *середнє всіх вибіркових середніх (взятих із однієї генеральної сукупності) буде дорівнювати середньому генеральної сукупності (μ), а дисперсія вибіркових середніх буде становити δ^2/n, де δ^2 – дисперсія сукупності, отриманої після складання всіх отриманих вибірок* [20, с. 221].

При цьому ймовірнісні вибірки, виділені з генеральної сукупності, мають бути достатньо великі для того, щоб їх вибіркове середнє описувалось законом нормального розподілу. Дж. Гласс вважає, що вибірки мають бути обсягом у принаймні 100 спостережень для того, щоб їх розподіл частот підпорядковувався закону нормального розподілу.

Пояснюючи положення центральної граничної теореми, Дж. Гласс зазначає, що розподіл середніх значень випадкових вибірок, узятих з генеральної сукупності, також буде підпорядковуватись закону нормального розподілу (див. рис. 5.1).

Рис. 5.1. Розподіл середніх значень 100 вибірок, узятих з однієї генеральної сукупності

При цьому середнє значення генеральної сукупності буде дорівнювати середньому значенню, обчисленому на основі середніх, отриманих із окремих вибірок.

Відповідно, оскільки розподіл середніх значень певної кількості випадкових вибірок підпорядковується закону нормального розподілу, значення середньо-квадратичного відхилення сукупності вибірок дозволяє оцінити ймовірність значення, яке приймає середнє. Таким чином, центральна гранична теорема відкриває можливості побудови *довірчого інтервалу для середнього значення генеральної сукупності* на основі вибіркових значень дисперсії та стандартного відхилення.

Середнє квадратичне відхилення для середнього (*стандартна помилка середнього*) обчислюється за формулою на рис. 5.2:

$$\sigma_{\bar{x}.} = \sigma / \sqrt{\bar{n}}.$$

Рис. 5.2. Формула обчислення стандартної помилки середнього

У формулі використано такі позначення: δ_x – середнє-квадратичне відхилення середнього (стандартна помилка середнього); δ - стандартне відхилення вибірки; n – кількість учасників дослідження.

Як вже зазначалось, стандартне відхилення середнього дозволяє встановити діапазон, довірчий інтервал, в який з певною долею ймовірності, може потрапити середнє значення з множини вибірок (тобто, середнє генеральної сукупності). Таким чином, використовуючи стандартне відхилення середнього, знайденого на основі вибіркових статистик, можна стверджувати (див. також рис. 5.3), що:

- 68% всіх середніх значень вибірок, які виділені з генеральної сукупності, будуть у діапазоні «середнє генеральної вибірки» ($\mu \pm \delta_x$), тобто в інтервалі від $\mu - (\delta/\sqrt{n})$ до $\mu + (\delta/\sqrt{n})$;

- 95% середніх значень будуть в інтервалі $\mu - (2\delta/\sqrt{n})$ до $\mu + (2\delta/\sqrt{n})$

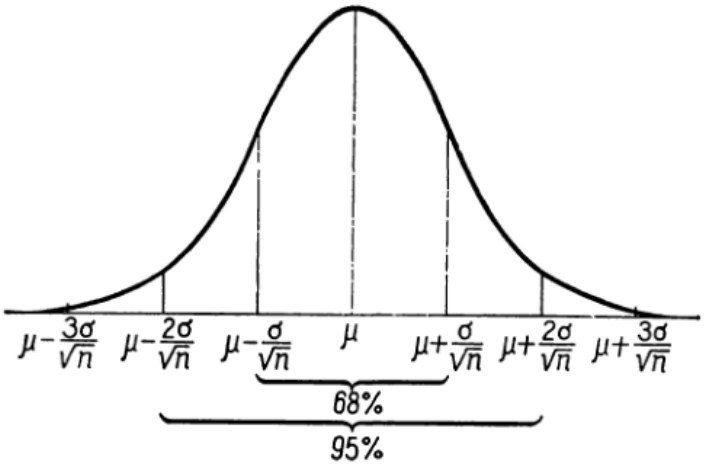

Рис. 5.3. Довірчий інтервал для середнього генеральної сукупності

Отже, розподіл середніх значень вибірок на рівні генеральної сукупності також підпорядковується закону нормального розподілу, що дає змогу оцінювати ймовірність отримати те чи інше значення середнього.

Використовуючи центральну граничну теорему стосовно конкретної вибірки, можемо стверджувати, що з імовірністю 95% середнє значення генеральної сукупності буде перебувати в інтервалі «вибіркове середнє плюс/мінус дві стандартні помилки середнього ($x – (2s/\sqrt{n})$)».

5.2. Порівняння двох вибірок на основі обчислення довірчих інтервалів

Якщо вибірка сформована за ймовірнісним принципом і має достатній обсяг, її статистики можуть бути використані для оцінки параметрів генеральної сукупності.

Тоді, розраховуючи середні значення за певною шкалою у двох групах опитаних та довірчі інтервали середнього, можна зробити висновок про належність середніх значень до однієї генеральної сукупності, або про низьку ймовірність, що середні значення в генеральній сукупності є рівними в цих двох вибірках.

Як приклад перевіримо гіпотезу, що середні значення двох груп достовірно відрізняються. Скористаємось даними з нашої третьої лекції. Вибірка складається з 92 студентів, які пройшли тестування за короткою формою опитувальника Р. Кеттелла. Візьмемо дві шкали: MD – Адекватність самооцінки та С – Емоційна стабільність.

За шкалою «Адекватність самооцінки» у нас є дві групи студентів: 1 – з умовно адекватною самооцінкою (MD < 5,5) та 2 – умовно завищеною самооцінкою (MD > 5,5). Умовою перевірки статистичного висновку є те, що опитувані у двох вибірках обирались на основі ймовірнісного вибору та шкала є метричною. Друга умова - розподіл за шкалою, для якої здійснюється порівняння, є нормальним.

У групі 1 з адекватною самооцінкою маємо 31 студента, середнє значення за шкалою С становить 5 балів (n=31; x_1 = 5; s=2,08).

У групі 2 з завищеною самооцінкою маємо 47 студентів, середнє значення за шкалою С становить 6,894 бала (n=47; x_2 = 6,894; s=2,01).

Статистична гіпотеза, яка перевіряється: H_0: середні значення двох груп належать до однієї генеральної сукупності; H_1: середні значення двох груп не належать до однієї генеральної сукупності. Рівень достовірності встановлюємо ймовірності помилки 5%.

Для перевірки цієї гіпотези потрібно побудувати довірчий інтервал для середнього генеральної сукупності. Згідно з центральною граничною теоремою, середнє значення генеральної сукупності є середнім значенням усіх вибірок. Отже, якщо гіпотеза H_0 правильна і вибірки справді належать до однієї генеральної сукупності, то її середнє значення $\mu = (x_1 + x_2)/2 = (5+6,894)/2 = 5,945$.

Дисперсія генеральної сукупності буде становити дисперсію всіх вибірок при їх складанні. Оскільки загальна вибірка у нас достатньо велика (n=92), можна прийняти її дисперсію як параметр імовірної генеральної сукупності. Для побудови довірчого інтервалу слід скористатись дисперсією,

стандартним відхиленням та стандартною помилкою середнього для шкали С «Емоційна стабільність». Статистики є такими: s^2=5,631; s=2,373; δ_x=0,2474.

На основі обчислених параметрів будуємо довірчі інтервали для середнього генеральної сукупності для випадку, якщо коректна H0 гіпотеза (див. рис. 5.4).

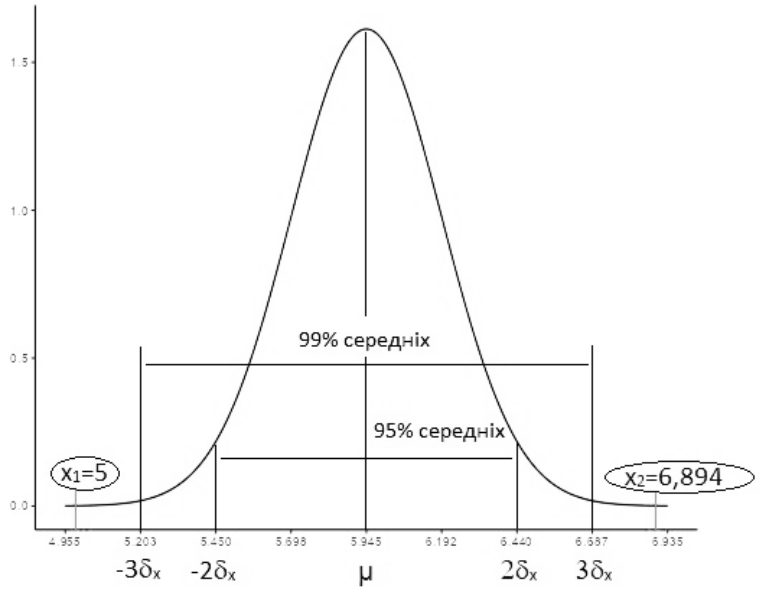

Рис. 5.4. Приклад застосування довірчого інтервалу для μ з метою перевірки гіпотези про приналежність середніх двох груп до однієї генеральної сукупності

Обчислення показує, що якщо вибірки отримано з однієї генеральної сукупності, то в 95% відсотків випадків їх середні будуть знаходитись в інтервалі від 5,45 бала до 6,44 бала; а в 99% – в інтервалі між 5,203 бала до 6,687 бала.

Оскільки середні значення двох досліджених груп (x_1 = 5 та x_2 = 6,894) – за межами 99% довірчого інтервалу для середнього генеральної сукупності, можна зробити висновок, що ймовірність їх приналежності до однієї генеральної сукупності є меншою за 1%.

Отже, H_0-гіпотеза про відсутність відмінностей у середніх значеннях між двома групами в генеральній сукупності відкидається.

Для статистичних висновків використовується не лише довірчий інтервал для середнього значення, а й прогностична здатність самого нормального розподілу. Наприклад, Дж. Гласс наводить вирішення завдання про доведення або спростування гіпотези про відсутність кореляції в генеральній сукупності для вибірки в 200 учасників [20, с. 249].

Гіпотеза про відсутність кореляції в такому випадку називається H_0 – гіпотезою. Гіпотеза про наявність ненульової кореляції (дійсного існування зв'язку) називається H_1 – гіпотезою.

Стандартне відхилення (δ) кореляції для такої вибірки становить 0,071. 95% довірчий інтервал для кореляції вибірки зазначеного обсягу лежить на відстані 0±1,96δ та становить 0,071*1,96 = 0,14. Довірчий інтервал представлено на рис. 5.5.

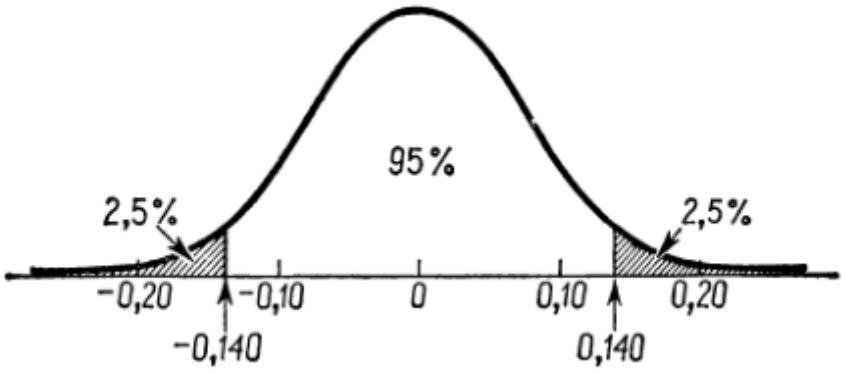

Рис. 5.5. 95% довірчий інтервал для вибіркового розподілу нульового коефіцієнта кореляції для вибірки 200 учасників.

З побудованого довірчого інтервалу видно, що 95% ймовірність відсутності кореляції в генеральній сукупності між двома змінними (зв'язок яких перевіряється) для вибірки 200 досліджених знаходиться в інтервалі від r = -0,14 до r = 0,14. Якщо у випадковій вибірці (обсягом 200 осіб) було отримано кореляцію 0,20, ймовірність, що ця вибірка належала до генеральної сукупності, у якій була відсутня кореляція, є дуже низькою (становить менше 0,05).

Відповідно спростовується H_0 гіпотеза та приймається H_1 – гіпотеза: з великою долею імовірності між змінними, що розглядаються, в генеральній

сукупності існує прямо-пропорційна кореляція. Довірчий інтервал показує ймовірність ***помилки першого роду*** **(α)** – у нашому випадку неправильне спростування гіпотези про відсутність кореляції. Отже, імовірність неправильного спростування нульової гіпотези, якщо на вибірці в 200 опитаних знайдено кореляцію 0,20, дуже мала (менша за 5%).

Помилка другого роду (β) полягає у цьому випадку в імовірності неправильного прийняття нульової гіпотези. У нашому випадку помилка другого роду є високоймовірною, якщо реальна кореляція в генеральній сукупності невисока та наближається до 95% довірчого інтервалу, побудованого на основі відсутньої кореляції. На рис. 5.6 представлено накладання довірчих інтервалів за нульової кореляції та кореляції 0,20, для вибірок обсягом у 200.

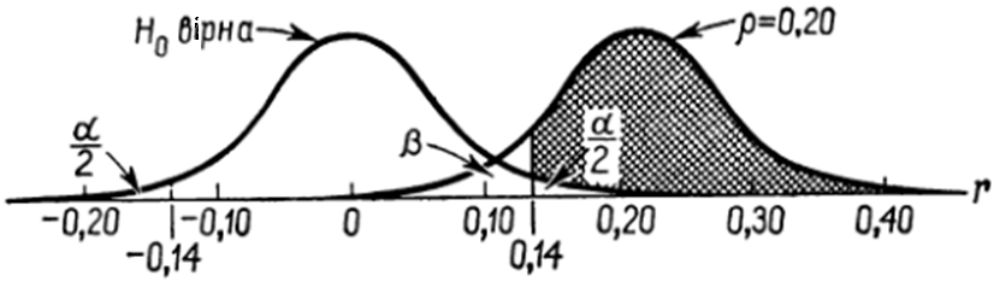

Рис. 5.6. Довірчі інтервали для нульової кореляції та R=0,20 у вибірці об’ємом 200

З рис. 5.6 видно, що довірчі інтервали частково перекриваються. Дж. Гласс [20, с. 259] пише, що 82% значень кореляції (за кореляції, рівної 0,20 в генеральній сукупності) будуть більшими за значення 0,14, яке є граничним інтервалом для нульової кореляції. Це означає, що ймовірність помилкового прийняття нульової гіпотези (β – помилки) за істинного значенні кореляції в генеральній сукупності, яке дорівнює r = 0,20, становить приблизно 100 - 82 = 18%. Отже, за умови істинного r = 0,20 в генеральній сукупності, з високою долею ймовірності (18%) ми отримаємо значення коефіцієнта кореляції нижче від критичного рівня r = 0,14 і зробимо хибний висновок про відсутність

кореляції в генеральній сукупності. Водночас за істинної кореляції r = 0,28 в генеральній сукупності ймовірність хибного прийняття H0-гіпотези є меншим за 5% (тобто прийнятним).

Отже, випадки перевірки статистичних гіпотез, у яких гіпотеза H_0 не була спростована, можуть містити помилку другого роду (β – помилку).

Критерії перевірки статистичних гіпотез використовують закони ймовірності, на яких побудована логіка довірчих інтервалів, водночас спрощують обчислення. Отже, розглянемо основні, найбільш популярні критерії.

5.3 Критерій χ^2 – Пірсона

Критерій використовується для порівняння частот у двох групах, створених за певними ознаками (класифікацією, шкалою найменувань). Критерій спрямований на підтвердження або спростування гіпотези про те, що розподіл у двох групах за ознакою, що вивчається, є випадковим (відповідає частоті, яка теоретично очікується).

Умови застосування критерію такі:

1. Обсяг вибірки має бути не меншим за 30 осіб.
2. У разі використання для таблиць 2x2 в кожній комірці має бути мінімум 5 спостережень (результатів 5 учасників).

Формула критерію вже обговорювалась у лекції 3, зараз ми зупинимось на основній ідеї перевірки статистичної гіпотези.

Теорія ймовірності дозволяє прогнозувати частоту появи певної події у випадковому виборі за умови, що відомий відсоток подій у генеральній сукупності (див. лекцію 1). Приймається, що за багаторазового рандомного вибору з генеральної сукупності утворена вибірка буде в наближеному вигляді відображати частоти подій, як вони існують у генеральній сукупності.

У теорії ймовірності є правила, які дозволяють прогнозувати ймовірність появи поєднаних подій. Так, імовірність появи поєднаної події А

і Б дорівнює добутку ймовірності події А та ймовірності події Б (за умови, що ці події незалежні та ймовірнісні) [20, с. 186]. І саме це правило дозволяє визначати, чи є розподіл частот випадковим, чи він порушує правило ймовірнісного поєднання частот. У таблиці 5.1 наведено дані, які вже обговорювались.

Таблиця 5.1

Поєднання частот за шкалами «Адекватна самооцінка» та «Емоційна стабільність»

Значення за фактором MD	Значення за фактором С		Сума частот в строчках
	Емоц. нестабільність С<5,5 (0 балів)	Емоц. стабільність С>5,5 (1 бал)	
Адекватна самооцінка MD<5,5 (0 балів)	30 (67%) f_{11}	15 (33%) f_{12}	45 (100%) $f_{1.}$
Завищена самооцінка MD>5,5 (1 бал)	9 (19%) f_{21}	38 (81%) f_{22}	47 (100%) $f_{2.}$
Сума частот в стовпчиках	39 $f_{.1}$	53 $f_{.2}$	92

Теоретичне правило виражено у формулі (див. рис. 5.7).

$$H_0 : P_{ij} = P_{i.} \; P_{.j}$$

Рис. 5.7. Нульова гіпотеза χ^2 – частота в комірці має відповідати теоретичному очікуванню

У формулі використано такі позначення: H_0 - нульова гіпотеза; P_{ij} – імовірність події в комірці подій I та J; $P_{i.}$ – загальна ймовірність подій I; $P_{.j}$ – загальна ймовірність подій J.

Провівши обчислення, отримаємо, що при ймовірнісному розподілі в комірці f_{11} мало бути 19 осіб (42% осіб з адекватною самооцінкою). Для пояснення наведемо обчислення теоретичної ймовірності $f_{11.}$ Ймовірність події «Емоційна нестабільність» становить 39 / 92 = 0,42. Імовірність появи Адекватної самооцінки – 45 / 92 = 0,49. Таким чином, теоретична ймовірність f_{11} (адекватна самооцінка поєднується з емоційною нестабільністю) становить

0,42 * 0,49 = 0,21. Тобто, якщо частоти є незалежними, теоретично потрібно очікувати частоту події f_{11} у 92 * 0,21 = 19 осіб. Насправді ж в емпіричних даних частота цієї події на 11 одиниць вища. Виникає питання, наскільки це значення лежить у межах ймовірного відхилення частот?

У критерії χ^2 розбіжність частот теоретичного і емпіричного розподілу перевіряється за допомогою зіставлення з критичними значеннями χ^2-*розподілу*.

Значення конкретного опитаного в χ^2-розподілі дорівнює його z^2 [20, с. 207]. Таким чином, для того, щоб побудувати хі-квадрат розподіл за певною змінною, потрібно значення опитаних перевести в стандартні z-бали та піднести їх у квадрат. На рис 5.8 вказано формули обчислення z-бала для конкретного опитаного та переведення z-балів у значення χ^2 – розподілу.

$$z_i = \frac{X_i - \overline{X}}{s} \qquad \chi_1^2 = z^2$$

Рис. 5.8. Формули обчислення z-значення та хі-квадрат.

У формулі використано такі позначення: z_i – z-значення в конкретного опитаного; X_i – значення за шкалою x у конкретного опитаного; $\overline{X}$ – середнє значення за шкалою x; s – стандартне відхилення шкали x; χ^2 – значення хі-квадрат.

Перетворення z-розподілу на χ^2-розподіл змінює вид кривої розподілу частот. Для шкали з одним ступенем свободи гістограма χ^2-розподілу має вигляд, представлений на рис. 5.9:

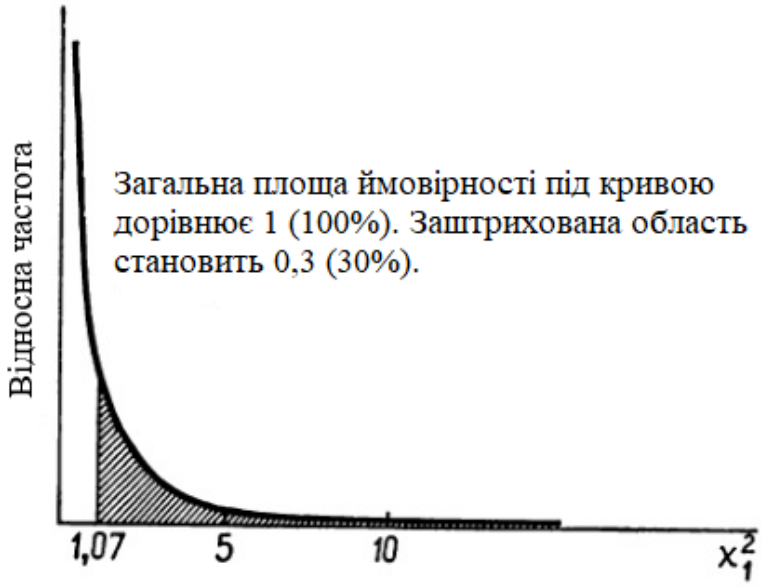

Рис. 5.9. Гістограма розподілу значень χ^2 – розподілу з одним ступенем свободи.

Дж. Гласс зауважує, що зі зростанням ступенів свободи крива χ^2-розподілу починає нагадувати дзвоноподібну форму. Крива χ^2-розподілу, подібно до одиничного нормального розподілу, має здатність передбачати частоту виникнення події та має ряд описаних ученими властивостей, які дозволяють здійснювати ці прогнози.

Потрібно зазначити, що для здійснення статистичних висновків χ^2-розподіл будується для поєднаних шкал, наприклад, $\chi_2^2 = z_A^2 + z_B^2$. Такі теоретично-прогнозовані криві χ^2 будуються для випадково розподілених даних та незалежних одна від одної змінних і відображають, з якою частотою події будуть траплятися за теоретичного розподілу на основі закону ймовірностей. Оскільки в основі χ^2-розподілу – одиничний розподіл (z-значення), у якому нульове значення є середнім, то найбільш поширені частоти в χ^2-розподілі будуть розташовані біля нульових значень. При цьому чим вищими будуть значення χ^2–розподілу, тим менш імовірною є подія.

На рис. 5.10 наведено формулу обчислення *критерію* χ^2 для таблиць 2х2. На відміну від формули на рис. 3.8 (зліва, для метричної шкали), значення χ^2

за формулою, наведеною на рис. 5.10, обчислюється для двох шкал, які мають біноміальний розподіл (тобто кожна шкала фіксує лише дві можливі події).

$$\chi^2 = \frac{n(f_{11}f_{22} - f_{12}f_{21})^2}{f_1 . f_2 . f_{.1} f_{.2}}.$$

Рис. 5.10 Формула критерію χ^2 для таблиць 2х2

У формулі використано такі позначення: *n – кількість опитаних;* f_{11} *– комірка зі збігом частот i=0 j=0 (див. табл. 5.1);* f_{22} *- комірка зі збігом частот i=1 j=1;* f_{12} *та* f_{21} *– комірки з відсутніми збігами: i=0 j=1 та i=1 j=0;* $f_{.1}$ *– загальна частота в стовпчику 1;* $f_{.2}$ *– загальна частота в стовпчику 2;* $f_{1.}$ *– загальна частота в рядку 1; ;* $f_{2.}$ *– загальна частота в рядку 2*

У наведеній формулі за повного збігу емпіричних частот з теоретичними очікуваннями (що є дуже рідкісним випадком) значення χ^2 буде дорівнювати нулю. У разі наближення частот до теоретично очікуваних значення χ^2 буде наближатись до нуля. Чим сильніше проявлятиметься розбіжність емпіричних частот з теоретичними очікуваннями, тим більшим буде значення критерію χ^2.

Отримане для конкретної таблиці емпіричних даних значення χ^2-критерію являє собою певну точку в χ^2-розподілі (відповідного ступеню свободи), і на основі χ^2-розподілу можна оцінити ймовірність виникнення такої події. У випадках, коли значення χ^2 має імовірність виникнення, меншу від 5%, робиться висновок про відхилення емпіричного розподілу від теоретично очікуваного. Для такого зіставлення створено спеціальні таблиці з *критичними значеннями* χ^2 для кожного ступеню свободи. Під *критичним значенням критерію* розуміють його певне число, яке дозволяє на заданому рівні ймовірності (α) відхилити нуль-гіпотезу. Зазвичай критичні значення з відповідними ступенями свободи (чи обсягом вибірки) вже обчислені статистиками для кожного критерію і опубліковані в таблицях. Часто до опису критеріїв додаються формули для можливості самостійного обчислення критичних значень за потреби.

Для даних з табл. 5.1 χ^2 = 21,255 імовірність появи такого поєднання частот менша за 0,001. Отже, імовірність отримати такі результати за

незалежного випадкового розподілу вкрай низька. Висновок: частоти в групах відрізняються від теоретично очікуваних.

5.4. t – критерій Стьюдента

Назва t – критерій Стьюдента – використовується як загальна для цілої низки критеріїв, які ґрунтуються на порівнянні даних з розподілом Стьюдента [2, с. 84]. До родини t – критеріїв належить: одно вибірковий t – критерій, призначений для перевірки гіпотези про відповідність середнього генеральної сукупності певному значенню; t – критерій для незалежних вибірок для перевірки гіпотези про приналежність середніх значень за певною шкалою у двох групах одній генеральній сукупності; t – критерій для пов'язаних вибірок – для вивчення наскільки суттєвими є відмінності між двома групами (наприклад, до та після експериментального впливу). Перелічені критерії використовують імовірнісні закономірності нормального розподілу. Водночас імовірність оцінюється не на основі обчислення довірчих інтервалів, а на основі зіставлення значень з критичними показниками t – розподілу Стьюдента.

Умови застосування критерію такі [2, с. 85]:

1. Вибірки, які порівнюються, мають відповідати моделі нормального розподілу. Оскільки ж висновок про відповідність вибірки нормальному розподілу важко зробити щодо вибірки, у якій менше, ніж 30 учасників, мінімальна кількість учасників дослідження в групах, які порівнюються, повинна становити не менше, ніж 30 осіб.

2. Дані слід вимірювати в метричній шкалі (інтервальній або відношення).

3. Обмеження для t – критерію для незалежних вибірок: дисперсія шкали у двох групах має бути гомогенною, що перевіряється на основі тесту Ліveня (перевіряє гіпотезу про приналежність дисперсії двох груп одній генеральній сукупності).

4. Обмеження для t – критерію для залежних вибірок: групи, які порівнюються, повинні мати достовірну лінійну кореляцію.

t – критерії дають виражене в цифрах значення співвідношення вимірів за шкалою між двома групами (за результатами конкретного експерименту) на кривій t – розподілу Стьюдента. Крива *t – розподілу* відображає імовірність появи певного значення залежно від ступенів свободи [15, с. 213].

Формула t – розподілу представлена на рис. 5.11:

$$t_n = \frac{z}{\sqrt{\chi_n^2 / n}}$$

Рис. 5.11. Формула t-розподілу

У формулі використано такі позначення: z – стандартизовані в одиничному нормальному розподілі значення за першою шкалою, χ^2 – значення за другою шкалою, стандартизовані в z^2 – значеннях; n – кількість ступенів свободи шкали в χ^2 – значеннях.

З формули 5.11 видно, що в основі значень *t-розподілу* лежить співвідношення стандартизованих значень, отриманих у двох групах. Виміри у двох групах, які порівнюються, здійснюються в єдиній метричній шкалі.

Криву t-розподілу представлено на рис. 5.12

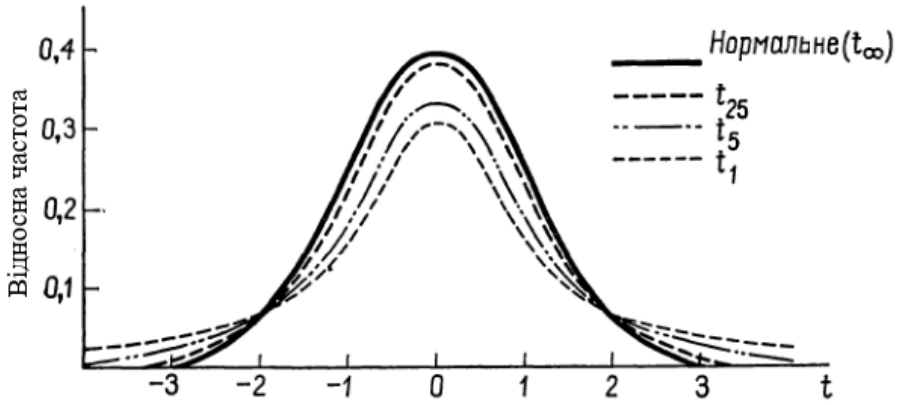

Рис. 5.12. t-розподіл з 1, 5, 25 та нескінченною кількістю ступенів свободи

t-розподіл використовується для оцінки достовірності гіпотези про рівність середнього значення у двох вибірках (незалежних і залежних), а також

для оцінки достовірності спростування нульової гіпотези в кореляції Пірсона за кількості опитаних понад 20.

t – критерій Стьюдента для незалежних вибірок.

t – критерій Стьюдента для незалежних вибірок перевіряє гіпотезу про рівність середніх значень у двох сукупностях. Гіпотези мають такий вигляд:

$H_0 : \mu_1 - \mu_2 = 0$

$H_1 : \mu_1 - \mu_2 \neq 0$

Як бачимо, різниця між середніми значеннями двох генеральних сукупностей дорівнює нулю (нуль-гіпотеза) або ж відрізняється від нуля (альтернативна гіпотеза). Значення t – критерію обчислюється за формулою (рис. 5.13):

$$t = \frac{\bar{X}_{.1} - \bar{X}_{.2}}{\sqrt{\frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Рис. 5.13 Формула t-критерію Стьюдента для незалежних вибірок з різною кількістю учасників

У формулі використано такі позначення: $\bar{X}_{.1}$ – середнє значення першої вибірки; $\bar{X}_{.2}$ – середнє значення другої вибірки; S_1^2 - дисперсія першої вибірки; S_2^2 - дисперсія другої вибірки; n_1 – кількість осіб у першій вибірці; n_2 – кількість осіб у другій вибірці.

Вимоги для застосування критерію наведені вище на сторінці 116. Перед застосуванням критерію Стьюдента для незалежних вибірок перевіряють дисперсію двох груп за критерієм Лівеня [2, с. 85-86]. У SPSS така перевірка здійснюється автоматично.

У випадку, якщо дисперсії у двох вибірках неоднорідні, для порівняння цих вибірок застосовується t – критерій Уелча [2, с. 90].

Отриманий коефіцієнт порівнюють з критичними значеннями для відповідного рівня достовірності t – розподілу зі ступенями свободи: загальна кількість опитаних у двох вибірках мінус два опитаних. Якщо отримане

значення t – критерію менше від критичного, це означає, що різниця між середніми значеннями в двох групах не істотна (варіюється в межах довірчого інтервалу до генерального середнього), приймається нульова гіпотеза.

Якщо ж значення t – критерію більше від критичного, середні двох вибірок відрізняються на заданому дослідником рівні достовірності.

В якості прикладу застосування t – критерію для порівняння двох незалежних вибірок можна навести результати отримані іранськими дослідникам В. Юсеф та співавторами [33]. В проведеному дослідженні вчені дійшли висновку, що для чоловіків (у порівнянні з жінками) більш характерні копінг-стратегії вирішення проблем та уникнення. Відмінності в середніх значеннях за зазначеними шкалами між чоловіками (n = 171) та жінками (n = 174) перевірялась на основі критерію Стюдента. Так, середнє значення за копінгом вирішення проблем в групі чоловіків становило 25,72 бала, а в групі жінок – 23,45 бала. Рівень достовірності за t – критерієм становив $p < 0,001$. Це означає – отримані в двох групах середні значення не належать до однієї генеральної сукупності. Схожі відмінності були знайдені і для копінгу уникнення. Середнє значення за шкалою уникнення в групі чоловіків становило 18,47 бала; в групі жінок – 17,36 бала ($p < 0,05$). Тобто, копінг уникнення серед чоловіків переважає, чоловіки та жінки не належать до однієї генеральної сукупності за цією шкалою (імовірність помилки менше 5%).

t – критерій Стьюдента для залежних вибірок. Залежними вибірками називають такі, які певним чином пов'язані [20, с. 270]. Наприклад, це можуть бути рандомна вибірка клієнтів до корекційного тренінгу та ті ж самі особи після здійсненого втручання. Або вибірка хлопчиків та їхніх сестер. Вимоги для застосування критерію наведені на початку підрозділу 5.3.

Як і в попередньому випадку, перевіряється припущення про рівність середніх значень двох вибірок в генеральній сукупності або їх відмінність:

$H_0 : \mu_1 - \mu_2 = 0$

$H_1 : \mu_1 - \mu_2 \neq 0$

Т - критерій для пов’язаних вибірок ґрунтується на ідеї, що за умови рівності середніх двох вибірок під час обчислення суми різниці значень у пов’язаних парах утвориться t – розподіл значень з середнім, яке дорівнює нулю.

Т - критерій для пов’язаних вибірок обчислюється за формулою (рис. 5.14):

$$t = \frac{\bar{d}.}{s_d/\sqrt{n}}, \text{ де}$$

$$\bar{d}. = \sum_{i=1}^{n} (X_{i1} - X_{i2})/n = \sum_{i=1}^{n} d_i/n, \qquad s_d = \sqrt{\sum_{i=1}^{n} \frac{(d_i - \bar{d}.)^2}{n-1}}$$

Рис. 5.14 Формула t - критерію для пов’язаних вибірок

У формулі використано такі позначення: $\bar{d}$ – середнє арифметичне різниці пар значень; S_d – стандартне відхилення різниці пар значень; X_{i1} – значення за шкалою в конкретного опитаного з першої вибірки; X_{i2} – значення за шкалою у того ж самого опитаного з другої вибірки n – кількість учасників; d_i – сума різниці пар першого та другого дослідження.

Якщо значення t – критерію перевищують встановлений (зазвичай 95%) інтервал (емпіричне значення t – критерію перевищує критичне), відмінності в середніх значеннях вважаються на обраному рівні ймовірності невипадковими та властивими для генеральної сукупності. Слід пам’ятати, що t – розподіл є симетричним, тому є верхні та нижні критичні значення.

В якості прикладу застосування t – критерію для залежних вибірок наведемо дослідження А. Гупта та С. Кумарі, в якому вивчається ефективність впливу конітивно-поведінкової терапії (КПТ) на метакогнітивні переконання при депресивних розладах [22].

Дизайн дослідження передбачав порівняння ефективності 10 сеансів КПТ у 20 пацієнтів з депресивними розладами до та після втручання. В якості контрольної групи виступили 20 пацієнтів з депресивними розладами, які отримували КПТ та медикаментозне лікування. Тобто, перевірялась ефективність впливу КПТ при депресивних розладах як монотерапія.

Критеріями виключення були особи, які вже отримували психотерапію або використовують психотерапевтичні технології самодопомоги, отримували лікування електро-судомною терапією, мали супутні психічні або важкі соматичні захворювання, важку депресію, суїцидальні схильності, інтелектуальну недостатність, неврологічні порушення та травми. В дослідженні були використані такі методики як: Коротка форма міжнародного нейропсихіатричного інтерв'ю (MINI 7.0.2), Шкала депресії Бека (BDI 2), Метакогнітивний опитувальник (MCQ-30), Короткий опитувальник якості життя ВОЗ (WHOQOL-BREF), Глобальна оцінка функціонування (GAF).

За всіма вказаними методиками монотерапія КПТ та комбінована КПТ з медикаментозним лікуванням показала високу ефективність. Для прикладу наведемо лише зміни значень за шкалою депресії Бека. Перед початком монотерапії КПТ середнє значення за шкало BDI становило 20,95 бала, після втручання BDI = 7,25 бала ($t = 9{,}38$; $p < 0{,}001$). Отже, різниця за шкалою депресії Бека перед початком та після завершення КПТ становила більше 13 балів, рівень депресії істотно зменшився. Ймовірність отримати таку сильну різницю між замірами в обстеженій вибірці, якщо в генеральній сукупності різниця між замірами дорівнює нулю, є дуже малою (менше однієї з тисячі обстежених вибірок). Тому, можна зробити висновок, що в генеральній сукупності пацієнтів з депресивними розладами КПТ як монотерапія з високою долею ймовірності зменшує інтенсивність депресії.

5.5. F-критерій Фішера для порівняння дисперсії двох незалежних вибірок

Критерій Фішера використовується в одно- та багатофакторному дисперсійному аналізі (ANOVA). У цьому підрозділі розглянемо лише застосування критерію Фішера для порівняння двох незалежних вибірок.

Завдання критерію полягає в оцінці того, наскільки дисперсія у двох незалежних групах за певною шкалою є тотожною або відрізняється. Якщо

дисперсія у двох групах однакова, вони належать до однієї генеральної сукупності, тобто між ними немає суттєвих відмінностей.

Вимоги до застосування F-критерію є такими [20, с. 278]:

1. Шкала є метричною.

2. Бажаний мінімальний обсяг у двох групах – 10-12 учасників.

3. Вибірки мають бути випадковими.

4. Генеральні сукупності двох груп, які порівнюються, повинні мати нормальний розподіл.

В основі використання критерію Фішера лежить перетворення даних на *F-розподіл*, який є похідним χ^2-розподілу та обчислюється за формулою (рис. 5.15) [20, с. 211]:

$$F_{n_1,\,n_2} = \frac{\chi^2_{n_1} / n_1}{\chi^2_{n_2} / n_2},$$

Рис. 5.15 Формула F-розподілу

У формулі використано такі позначення: $F_{n1,\,n2}$ *– значення точки на F-розподілі, нижніми індексами позначаються ступені свободи для чисельника (перший індекс) та знаменника (другий індекс);* χ^2_{n1} *- значення хі-квадрата чисельника в першій вибірці (значення хі-квадрату за змінною всіх опитаних у вибірці додаються), нижнім індексом вказується ступінь свободи чисельника;* χ^2_{n2} *значення хі-квадрата знаменника в другій вибірці (значення хі-квадрата за змінною всіх опитаних у другій вибірці додаються), нижнім індексом вказується ступінь свободи знаменника;* n_1 *– ступінь свободи чисельника;* n_2 *– ступінь свободи знаменника.*

Ступені свободи обчислюються за принципом: обсяг вибірки мінус один (df = n-1) [2, с. 93]. F-розподіл дозволяє зробити прогноз, чи підпорядковується співвідношення розсіювання даних у двох вибірках закону нормального розподілу, чи є невипадковим. Зображення F-розподілу для вибірок з 4 та 4 ступенями свободи ($F_{4,4}$) та вибірок з 4 та 25 ступенями свободи ($F_{4,25}$) представлено на рис. 5.16.

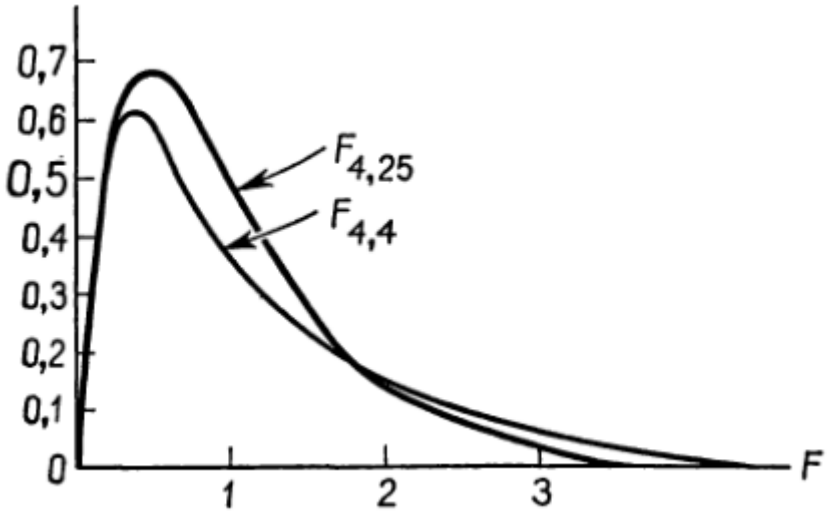

Рис. 5.16. Крива F-розподілу із зазначеними ступенями свободи

Для порівняння дисперсій двох незалежних вибірок використовується критерій пропорцій двох вибіркових дисперсій (рис. 5.17) :

$$F = \frac{s_1^2}{s_2^2}.$$

Рис. 5.17. Формула F-критерію для двох незалежних вибірок

У формулі використано такі позначення: S_1^2 – дисперсія чисельника; S_2^2 – дисперсія знаменника.

Дж. Гласс зазначає, що дисперсія двох вибірок, яка буде чисельником чи знаменником, обирається випадково [20, с. 278]. У будь-якому разі, якщо дисперсії не рівні, значення будуть перевищувати критичні. За рівності дисперсій значення *F-критерію* буде наближатись до одиниці.

Отримане емпіричне значення F-критерію зіставляється з критичними значеннями F-розподілу з відповідними ступенями свободи. Ступінь свободи чисельника дорівнює обсягу вибірки (чия дисперсія виступила чисельником) мінус один ($df_{чисельника}$ = n-1). Так само обчислюється і рівень свободи для знаменника: $df_{знаменника}$ = n-1. Якщо значення F-критерію перевищують граничні, тобто дисперсія у двох вибірках відрізняється на заданому рівні ймовірності, то гіпотеза про рівність дисперсій відкидається. Як і t-критерій

Стюдента, F-розподіл є симетричним. Тому є нижні критичні значення (F-емпіричне має бути нижчим) та верхні критичні значення (F-емпіричне має бути вищим).

Як приклад застосування F – критерію наведемо вже цитовані раніше результати дослідження В. Юсеф та співавторів [33]. Мета дослідження полягала у вивченні зв'язку між стресовими подіями та копінг-стратегіями, які застосовують особи для їх подолання. Вчені підкреслюють, що природа цього зв'язку не однозначна. Так, копінг вирішення проблем пов'язаний з більш кращим запам'ятовуванням стресових ситуацій, які вирішувались. Тому, саме по-собі вимірювання стресових подій може бути неточним через особливості сприйняття та пам'яті тісно-пов'язаними з типом копінгу.

Вчені виділили три групи осіб за різною кількістю стресових подій, використовуючи для їх позначення імовірність захворіти протягом 2 років за даними Т. Холмса та Р. Раге. В табл. 5.2 представлені середні значення в трьох виділених групах за шкалами копінг-стратегій та наведене значення F – критерію та його рівень достовірності.

Таблиця 5.2

Результати дослідження зв'язку стресових подій та стратегій подолання В. Юсеф та співавторів

Група	N	Ризик захворіти	Середні значення за шкалами		
			Копінг-стратегії орієнтовані на:		
			Проблему	Емоційне відреагування	Уникнення
1	252	30%	25,3	22,8	20,1
2	75	50%	24,8	20,8	18,2
3	18	80%	23,5	18,3	16,4
F – критерій			F = 1,72; p = 0,18	F = 4,25; p < 0,05	F = 7,52; p < 0,001

З табл. 5.2 видно, що між групами були виявлені відмінності в середніх значеннях майже за всіма шкалами (окрім орієнтованого на проблему копінгу). За результатами дисперсійного аналізу В. Юсеф та співавтори зробили висновок, що висока кількість стресових подій пов'язана з певною

нездатністю долати стресові ситуації. На це вказують низькі значення за шкалами використання копінг-стратегій орієнтованих на емоційне відреагування та уникнення. Відповідно, практичним завданням може бути навчання вразливих верств населення долати стресові події з використанням найбільш адекватного до ситуації копінгу.

Застосування F – критерію для порівняння дисперсії в трьох і більше групах дозволяє зробити висновок, про присутність достовірних відмінностей між певними групами серед усіх груп. Проте, не вказує між якими саме групами ці відмінності присутні. Для статистичного висновку про достовірні відмінності між конкретними групами разом з F – критерієм, як правило, застосовується критерій Шеффе.

5.6. Непараметричні критерії

Непараметричні критерії застосовуються для порівняння вибірок, які виміряні в порядковій шкалі (ранги) або вище (інтервальна шкала та шкала відношень). При цьому розподіл даних у генеральній сукупності може відрізнятись від нормального. Непараметричні критерії є менш потужними порівняно з параметричними.

Критерій U-Манна-Уїтні

U-критерій застосовується для порівняння незалежних вибірок, є непараметричним аналогом t-критерію Стьюдента. Критерій може застосовуватись для вибірок невеликого обсягу.

Вимоги до його застосування такі [2, с. 103]:

1. Мінімальний обсяг вибірки становить 3 особи в кожній групі. Допускається, щоб в одній групі було дві особи, тоді в іншій має бути 5 осіб.

2. Максимальна кількість осіб у кожній з груп має бути до 60. Загальна вибірка має бути не більшою за 100 осіб. Якщо це значення перевищене, критерій помилково відхиляє нульову гіпотезу.

3. Вибірки слід сформувати на основі ймовірнісного принципу.

4. Дві групи порівнюються на основі однієї шкали, яка має бути порядковою або метричною та мати високу варіацію значень. Не бажаним є збіг значень у двох групах.

Формула обчислення U-критерію така: (рис. 5.18):

$$U = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x,$$

Рис. 5.18. Формула обчислення U-критерію Манна-Уітні

У формулі використано такі позначення: 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.*

Отримане значення за критерієм Манна-Уїтні порівнюється з критичним значенням. Емпіричне значення критерію має бути меншим за критичне. Відсутність відмінностей між двома групами спостерігається, коли емпіричне значення критерію наближається до нуля. У разі значення загальної вибірки 20 осіб і вище рівень імовірності відмінностей обчислюється на основі стандартизованих значень Z_u.

Нульова гіпотеза для U-критерію має таке формулювання: різниця між двома групами в рангових розподілах відсутня. У разі її спростування приймається H_1 -гіпотеза: ознака розподілена нерівномірно у двох групах, і це перевищує можливі ймовірнісні відхилення. Спростування нульової гіпотези за критерієм Манна-Уїтні вказує на відмінність медіани у двох групах.

Під час обчислення критерію перш за все потрібно скласти єдиний ранговий ряд, утворений значеннями двох вибірок. Ранги мають надаватися за зростанням ознаки (перший ранг надається найменшому значенню. Утворена кількість рангів має дорівнювати загальній кількості учасників дослідження (сумі осіб у двох групах).

Окремо обчислюється сума рангів у першій та другій групах. Визначається найбільша сума рангів, яку потрібно підставити у формулу U-критерію (рис. 5.18). За формулою обчислюється значення U-критерію.

В якості прикладу використання U-критерію наведемо результати дослідження Х. Озоріо-Шпулер та співавторів [27]. Мета дослідження полягала у вивченні емоційного виснаження медичних сестер в залежності від рівня стресу, статі, курсу та місця навчання. До проявів емоційного виснаження вчені відносили: втрату енергії, відчуття фізичної та фізіологічної втоми, втрату мотивації до навчання, цинічне та безвідповідальне ставленням до студентських обов'язків. Актуальність дослідження пов'язана з негативним впливом емоційного виснаження на успішність в навчанні та психічний стан студентів. В якості гіпотези дослідження допускалось, що більш високий рівень емоційного виснаження характерний для працюючих студенток, які мають дітей.

Було обстежено 1368 студентів з чотирьох вишів (3 – іспанських, 1 – чилійського). Автори повідомляють, що обстежені вибірки не є випадковими, тому дослідження має певні обмеження та репрезентативне лише по відношенню до обстеженої вибірки.

В дослідженні використана метрична шкала Емоційного виснаження (EES). Метричними змінними також були: вік та рік навчання. Рівень стресу вимірювався за шкалою порядку, інші змінні вимірювались за шкалами найменувань та бінарними шкалами: університет навчання, сімейний стан, стать, наявність потребуючих піклування осіб, робота, час на дорогу, фінансування навчання, стипендія, повторне перескладання дисциплін. Переважна більшість студентів мала середній рівень емоційного виснаження (від 26,1 бала до 36,9 бала). В табл. 5.3 представлені деякі результати порівняння за U-критерієм Манна-Уїтні двох груп студентів виділеними на основі соціодемографічних та академічних ознак.

Таблиця 5.3

Середні значення емоційного виснаження в групах студентів за соціодемографічними та академічними ознаками за Х. Озоріо-Шпулер та співавторами

Ознака, яка утворює групи	Категорії в групі	Кількість студентів	Середнє значення	Рівень достовірності за U-критерієм
Стать	Чоловік	214	27,9	$p < 0,05$
	Жінка	1077	31,9	
Діти	Так	36	35,7	$p < 0,05$
	Ні	1071	30,9	
Робота, крім навчання	Так	428	32,1	$p < 0,05$
	Ні	892	30,8	
Час на дорогу	<30 хвилин	1034	30,9	$p < 0,05$
	>30 хвилин	259	32,7	
Повторні перескладання	Так	259	32,5	$p < 0,05$
	Ні	979	30,8	
Стипендія	Так	719	31,9	$p < 0,05$
	Ні	604	30,5	

З табл. 5.3 видно, що окремі групи мають більш високі значення емоційного вигорання: жінки; студенти з дітьми; працюючі; студенти, які витрачають на дорогу більше 30 хвилин; студенти, які мали досвід повторного перескладання предметів; студенти, які отримують стипендію. Для порівняння груп був застосований непараметричний критерій Манна-Уїтні, який показав, що міри центральної тенденції виділених груп з високою долею імовірності не належать до спільної генеральної сукупності. В подальшому дослідники використали виявлені предиктори емоційного виснаження для побудови моделі множинної регресії (дозволила передбачити 25% значень виснаження).

Н-критерій Краскела-Уолліса.

Н-критерій вважається непараметричним аналогом однофакторного дисперсійного аналізу. Призначений для порівняння медіани в трьох групах за шкалою порядку (або вище).

Умови застосування критерію:

1. Наявність трьох або більше груп, незалежних одна від одної.

2. Результати за всіма трьома групами виміряно за однією шкалою (ранговою або метричною).

3. Мінімальна кількість учасників у вибірках: у першій – 4 особи, у другій – 2 особи, у третій – 2 особи (4/2/2).

4. Вибірки сформовано на основі ймовірнісного принципу.

H-критерій не показує, у якій саме групі медіана є вищою, лише вказує наявність відмінності між групами.

Формулу обчислення H-критерію наведено на рис. 5.19:

$$H = \frac{12}{n(n+1)} \cdot \Sigma \frac{R_i^2}{n_i} - 3(n+1)$$

Рис. 5.19 Формула H-критерію Краскела-Уолліса

У формулі використано такі позначення: n *– загальна кількість учасників у всіх вибірках;* n_i *– кількість осіб у групі i;* R_i *– сума рангів для вибірки i.*

У разі, якщо значення в групах мають однакові ранги, використовується модифікована формула [2, с. 108].

Для обчислення значень за формулою дані всіх груп розміщуються в один стовпчик від меншого до більшого та ранжуються. Для кожної групи обчислюється сума рангів. Після цього проводять обчислення за формулою. У формулу (рис. 5.19) підставляються суми рангів за кожною групою, та діленням на обсяг вибірки обчислюється квадрат середнього значення рангу в кожній групі. Після цього обчислюється сума квадратів середніх рангових значень. Отримане значення множать на результат ділення 12/(n(n+1)). Від результату множення віднімається 3(n+1).

Емпіричне значення зіставляється з критичним. Якщо емпіричне дорівнює критичному або більше за нього на заданому рівні значущості, нуль-гіпотеза відхиляється та робиться висновок про відмінність між групами в генеральній сукупності [2, с. 108].

В дослідженні іранських вчених Л. Резаї та співавторів вивчались особистісні риси осіб, які страждають двома типами безсоння та здоровими особами [28]. Статистичний аналіз був спрямований на порівняння середніх значень за шкалами короткої форми MMPI в трьох групах: пацієнтів, які страждають парадоксальним безсонням, психофізіологічним безсонням та осіб без розладів сну. Допускалося, що для парадоксального безсоння характерно когнітивно-емоційне гіперзбудження та неправильне сприйняття, для його корекції можуть бути використані психотерапевтичні методи. Психофізіологічне безсоння пов'язувалось вченими з надмірним фізіологічним збудженням центральної нервової системи, допускалась, що цей розлад потребує медикаментозного лікування. Метою дослідження було виявлення профілів особистісних рис за різних видів безсоння, для покращення якості діагностики розладів сну та розробки стратегій корекції.

В статистичному аналізі дослідники використовували F – критерій у випадках коли дані в трьох групах відповідали закономірностям нормального розподілу. У випадках коли дані достовірно відхилялись від нормального розподілу застосовувався H-критерій (див. табл. 5.4)

Таблиця 5.4

Середні значення особистісних рис за MMPI в групах пацієнтів з безсонням та осіб без розладів сну за Л. Резаї та співавторами

Шкала MMPI-2	Середні значення в групах			Стати-стичний критерій	Рівень достові-рності
	Парадоксальне безсоння, n = 20	Психофізіологічне безсоння, n = 20	Здоровий сон, n = 60		
Істерія	6,9	2,52	4,5	F = 15,41	p<0,001
Депресія	10,85	10,5	7,1	H = 22,62	p<0,001
Іпохондрія	12,4	12,75	9,3	F = 14,79	p<0,001
Психопатія	7,58	8,1	6,05	F = 8,4	p<0,001
Параноя	6,85	6,7	4,45	H = 18,66	p<0,001
Гіпоманія	6,0	5,05	3,9	H = 13,79	p<0,001
Психоастенія	8,85	8,65	5,88	H = 13,76	p<0,001
Шизофренія	8,2	7,85	5,9	F = 3,65	P=0,06

З табл. 5.4 видно існування достовірних відмінностей в середніх значеннях між трьома групами дослідження майже за всіма шкалами MMPI (за

виключенням шкали Шизофренія). Критерій Краскела-Уолліса використовувався для порівняння медіан в шкалах: Депресія, Параноя, Гіпоманія та Психоастенія. Для перелічених шкал Н-критерій показав достовірні відмінності, що означає – медіани в деяких з груп не належать до спільної генеральної сукупності (з допустимою помилкою α). Але, Н-критерій лише констатує наявність відмінностей, проте не вказує між якими групами вони існують.

Для висновку між якими саме групами спостерігається відмінність вчені використали критерій Тьюкі. Для порівняння об'єднаної групи всіх пацієнтів з інсомнією (парадоксальне та фізіологічне безсоння) та осіб без розладів сну – критерій Манна-Уітні. В результаті проведених порівнянь, вчені дійшли висновку, що пацієнти з психофізіологічним порушенням сну (при порівняні з групою осіб зі здоровим сном) мали більш високі середні значення за шкалами: Істерія, Іпохондрія та Психопатія. Пацієнти з парадоксальним безсонням (при порівняні з групою осіб зі здоровим сном) мали більш високі середні значення за шкалами: Істерія та Гіпоманія. Обидві групи пацієнтів з розладами сну – значимо не відрізнялись між собою за використаними шкалами MMPI. Об'єднана група пацієнтів з інсомнією відзначалась більш високими балами за шкалами: Депресія, Параноя та Психоастенія у порівнянні з особами зі здоровим сном.

Критерій Вілкоксона. Критерій Т Вілкоксона (іноді W-критерій Вілкоксона) є непараметричним та використовується для порівняння двох пов'язаних вибірок (зазвичай склад вибірок є ідентичним, порівнюються результати до та після певного впливу). Критерій показує не лише існування достовірної відмінності між групами, але й напрямок «зсуву» [2, с. 112]. Т Вілкоксона є непараметричним аналогом t-критерію Стьюдента для пов'язаних вибірок.

Критерій перевіряє гіпотезу (H_0): середні значення двох залежних вибірок належать до одного генерального середнього.

Вимоги до застосування критерію Вілкоксона [2, с. 115]:

1. Мінімальний обсяг вибірки – 5 учасників, значення яких порівнюється до та після певного впливу.

2. Вибірки мають бути пов’язані, дані мають бути представлені в шкалі порядку або метричних шкалах.

3. Під час підрахунків вилучаються випадки з «зсувом», який дорівнює нулю, через це на відповідну кількість вилучених осіб зменшується обсяг вибірки.

4. Змінна має варіюватись у широкому діапазоні значень, різниця рангів має бути вищою, ніж зсуви в межах +1 та -1.

Формула для обчислення критерію Вілкоксона відсутня – є загальний алгоритм, якого потрібно дотримуватись:

1. Результати тестування учасників дослідження впорядковуються в алфавітному порядку.

2. Для кожного учасника обчислюється різниця між сирими значеннями за шкалою. Це означає, що від значення першого тестування віднімаються значення повторного та записуються у відповідний стовпчик.

3. Вилучаються опитувані з нульовою різницею, зменшується обсяг вибірки.

4. В окремому стовпчику різниця величин переводиться в абсолютні значення (без мінуса), які ранжуються від найменшого до найбільшого.

5. Створюються два нових стовпчики: до першого вносяться всі абсолютні ранги для зсуву в позитивний бік (коли різниця сирих балів між першим та другим тестуванням дала позитивне число).

6. У другий стовпчик вносяться абсолютні ранги для зсуву в негативний бік (коли різниця сирих балів між першим та другим тестуванням дала негативне число).

7. Обчислюється сума рангів для позитивного та негативного зсуву. Отримана сума рангів, яка є меншою, є критичним значенням критерію Вілкоксона.

Емпіричне значення критерію Вілкоксона зіставляється з критичним на заданому рівні значущості. Якщо емпіричне менше за критичне або дорівнює йому, H_0-гіпотеза відкидається. Якщо обсяг вибірки є вищим за 25 осіб, для визначення рівня достовірності використовується стандартизована статистика Z [2, с. 113].

В якості прикладу застосування W-критерію для пов'язаних вибірок можна навести дослідження У. Сенгупта та А. Р. Сінгха спрямоване на вивчення ефективності функціональної аналітичної терапії для зменшення симптомів шизофренії [30]. Функціональна аналітична терапія описується вченими як сучасне поєднання когнітивно-поведінкової та психоаналітичної терапії. Вона спрямована на усвідомлення особистісних причин сприйняття соціальних ситуацій, корекцію сприйняття та поведінки в соціальних ситуаціях.

В дослідженні прийняло участь 10 пацієнтів психіатричного стаціонару чоловічої статі, віком від 20 до 40 років. Об'єм експериментальної та контрольної групи склав по 5 пацієнтів. Дизайн дослідження передбачав психодіагностику до та після психотерапевтичної допомоги. В контрольній групі проводилась повторна психодіагностика без впливу. Ефективність психотерапевтичного впливу вивчалась за допомогою аналізу змін за використаними тестами (був застосований W-критерій). Для доведення, що ефект зумовлений психотерапією, а не іншими факторами дані зіставлялися з контрольною групою (використаний критерій Манна-Уітні).

В досліджені застосовані наступні методики: Шкала позитивних і негативних синдромів (PANSS), Шкала оцінки апатії, Шкала сімейної адаптації та згуртованості (FACES), Шкала якості життя пацієнтів з шизофренією (SQLF), опитувальник Копінг-поведінка у стресових ситуаціях (CRI-A). Перед початком психотерапії між контрольною та експериментальною групами були відсутні достовірні відмінності за використаними шкалами (групи були співставними).

В табл. 5.5 представлені результати порівняння шкал використаних методик за W-критерієм.

Таблиця 5.5

Середні значення за батареєю тестів в групі пацієнтів хворих шизофренією до та після психотерапії за У. Сенгупта, А. Р. Сінгх

Назва методики/шкали	Середні значення за шкалою		Результати W-критерію	
	До психотерапії	Після втручання	Z	p
PANSS				
Позивні симптоми	34,6	19,8	2,03	< 0,05
Негативні симптоми	69,8	62,6	2,02	< 0,05
SQLS				
Психосоціальний домен	41	13	2,02	< 0,05
Домен симптомів	19	3,6	3,03	< 0,05
CRI-A				
Логічний аналіз	28,6	60,8	2,06	< 0,05
Позитивна переоцінка	35	58	2,03	< 0,05
Пошук підтримки	33,4	62,6	2,03	< 0,05
Вирішення проблем	34,4	61,2	2,03	< 0,05
Заперечення	67,6	41,4	2,03	< 0,05
Емоційне відреагування	76	49,6	2,02	< 0,05

Як видно з табл. 5.5 пацієнти, які отримували психотерапію значно зменшили позитивні та негативні симптоми захворювання, покращили якість життя та посилили стратегії подолання стресових ситуацій. За іншими шкалами достовірних відмінностей до та після втручання не виявлено.

Також, достовірні відмінності спостерігались з контрольною групою. Як зазначають автори – основним недоліком дослідження була мала вибірка, яку вчені планують збільшити в повторному дослідженні. Іншим недоліком роботи є недостатній опис медичного лікування, яке отримували пацієнти в двох групах.

ПЛАНИ СЕМІНАРСЬКИХ ЗАНЯТЬ

Семінар 1. Основні показники, одержувані в результаті первинної обробки експериментальних даних.

Мета: закріпити основні поняття математичної статистики, освоїти зміст та призначення показників, які отримуються в результаті первинної обробки експериментальних даних, розширити знання про вибірковий метод дослідження.

Питання до семінару:

1. Розкрийте основну ідею математичної статистики. З яких основних блоків складається застосування методів математичної статистики? Охарактеризуйте їх.
2. Розкрийте основні завдання математичної статистики.
3. Розрийте поняття змінна, шкала, дані. Розкрийте класифікацію шкал в психології, особливості кожної із шкал.
4. Охарактеризуйте співвідношення генеральної та вибіркової сукупностей.
5. Розкрийте способи формування ймовірнісної вибірки. Розкрийте основний алгоритм формування ймовірнісної стратифікованої вибірки.
6. Розкрийте способи формування вибірки, які не належать до ймовірнісних, в яких випадках вони застосовуються і які у них обмеження.
7. Розкрийте способи обробки первинних даних: просте табулювання, ранжування, групування частот, види графічного відображення розподілу частот у вибірці.
8. Розкрийте спосіб підрахунку середнього, моди та медіани. Яке значення вони мають для характеристики вибірки?
9. Розкрийте поняття теоретичної, емпіричної та статистичної гіпотез.
10. Розкрийте спосіб обчислення дисперсії та стандартного відхилення, інших мір розсіювання. Для чого вони застосовуються?
11. Розкрийте основні закономірності нормального розподілу. Для чого вони застосовуються?

Основні поняття: математична статистика, описова статистика, статистичні висновки, генеральна сукупність, вибірка, статистики, параметри, ймовірнісна вибірка, змінна, вимірювання, шкала, шкала найменувань, шкала порядку, шкала інтервалів, шкала відношення, нормальний розподіл, об'єм ймовірнісної вибірки, репрезентативність вибірки, наукова гіпотеза, теоретична гіпотеза, емпірична гіпотеза, статистична гіпотеза, статистичний критерій, H0 та H1 гіпотези, дані, впорядкування, ранжування, табулювання даних, частота події, групування частот, гістограма, полігон, коробчастий графік (ящик з вусами), міри центральної тенденції, мода, медіана, середнє арифметичне, міри мінливості, варіаційний розмах, дисперсію та стандартне відхилення, закон нормального розподілу, z-перетворення, асиметрія, ексцес.

Рекомендовані джерела:

1. Боснюк В.Ф. Математичні методи в психології: курс лекцій. Мультимедійне видання. Харьків : НУЦЗУ, 2020. 141 с.

2. Руденко В.М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.

Семінарське 2. Кореляційний аналіз та його основні параметри. Одномірна лінійна регресія.

Мета: сформувати у студентів наукове уявлення про кореляційний аналіз та його основні параметри, модель лінійної регресії та її призначення.

Питання до семінару:

1. Розкрийте поняття кореляція. Охарактеризуйте сутність кореляційного зв'язку.
2. Проаналізуйте випадки застосування коефіцієнтів кореляції Пірсона та Спірмена.
3. Який кореляційний зв'язок називають прямим, а який зворотнім? Наведіть приклади.
4. Охарактеризуйте основні параметри оцінки кореляційного відношення. Проведіть аналіз різних діаграм розсіювання.

5. Опишіть особливості використання коефіцієнтів оцінки зв'язку для шкал найменувань: ф кореляцію К. Пірсона; критерій χ^2, коефіцієнт зв'язаності V Крамера.

6. Для чого застосовується одномірна лінійна регресія, на що вказує коефіцієнт детермінації R^2.

7. Які показники якості моделі одномірної лінійної регресії? Наведіть приклад складання формули регресії на основі значення незалежної змінної, регресійного коефіцієнта та константи.

Основні поняття: коефіцієнт кореляції, напрям зв'язку, сила зв'язку, значущість зв'язку, графік розсіювання, лінійна кореляція, регресійне рівняння, залежна змінна, незалежна змінна, загальна лінійна модель.

Рекомендована література:

1. Руденко В.М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.

2. Wasserman L. All of statistics: a concise course in statistical inference. Springer, 2010. 468 p.

Семінар 3. Методи багатомірної статистики та особливості їх застосування у психології

Мета: сформувати у студентів наукове уявлення про методи багатомірної статистики та особливості їх застосування у психології.

Питання до семінару:

1. Дайте характеристику багатомірних методів математичної статистики. Розкрийте найбільш поширені методи багатомірної статистики, та наведіть приклади їх застосування в психології.

2. Наведіть приклади застосування методів багатомірної статистики в психології на основі публікацій за останні 5 років.

3. Охарактеризуйте основні методи відбору змінних та параметри оцінки показників множинного регресійного аналізу. Які сфери використання моделей множинної регресії?

4. Розкрийте призначення кластерного аналізу, тип даних з яким працює кластерний аналіз та способи оцінки відстані між дослідженими.

5. Розкрийте алгоритми кластерізації опитаних та способи оцінки числа кластерів.

6. Розкрийте основне призначення пошукового (екплуаторного) та конфірматорного факторного аналізу.

7. Розкрийте сутність різних моделей факторного аналізу, призначення та види обертання кореляційної матриці, проблему визначення числа факторів. Що є основним результатом факторного аналізу і в якому вигляді цей результат подається?

8. Розкрийте основні показники, які вказують на придатність факторного аналізу для пояснення кореляційних зв'язків у вибірці. Розкрийте основні показники, які вказують на високу пояснювальну здатність фактору в конфірматорному факторному аналізі (структурна модель).

Основні поняття: множинна регресія, ієрархічний кластерний аналіз, дендрограма, матриця відмінностей, фактор, варімакс-обертання, навантаження змінної, структурні моделі.

Рекомендована література:

1. Руденко В.М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.

2. Климчук В.О. Математичні методи у психології. Навч. посіб. для студентів психологічних спеціальностей. Київ : Освіта України, 2009. 288 с.

3. Fidell L. S., Tabachnick B. G. Using multivariate statistics. 7th ed. Pearson Education, 2019. 832 p.

ПЛАНИ ПРАКТИЧНИХ ЗАНЯТЬ

Практичне заняття 1. Способи формування ймовірнісної (випадкової) вибірки.

Мета: сформувати практичні навички утворювати випадкову (ймовірнісну) вибірку.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення Microsoft Excel, SPSS або Jamovi (на вибір).

Завдання:

1. Візьміть список Вашої групи. На основі таблиць випадкових чисел (або генератору випадкового вибору) сформуйте випадкову вибірку з навчальної групи.
2. За посиланням знаходяться значення 500 опитаних за методикою дослідження рівня стресу: https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical1.docx Розрахуйте для вказаної вибірки такі міри центральної тенденції: мода, медіана та середнє арифметичне. Напишіть формули, за якими здійснюються ці обчислення.
3. Сформуйте з наведеної вибірки ймовірнісну вибірку з 30 опитаних та обчисліть значення: моди, медіани та середньо арифметичного значення.
4. Порівняйте між собою повну та вибірку на основі середнього, моди та медіани.

Проблемні питання. Якими можуть бути наслідки для результатів дослідження, якщо замість ймовірнісної вибірки використати довільну (невипадкову) підгрупу учасників? У яких ситуаціях середнє арифметичне може вводити в оману дослідника у порівнянні з медіаною або модою? Наведіть приклади. Чому при малих вибірках (наприклад, n = 30) мода, медіана та середнє арифметичне можуть суттєво відрізнятися від аналогічних показників для повної вибірки? Чому у психології так важливо дотримуватися принципів випадковості вибірки?

Рекомендована література:

1. Руденко В.М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.

Практичне заняття 2. Табулювання даних, графічне відображення розподілу частот.

Мета: сформувати практичні навички впорядковувати дані: сортувати за зростанням, привласнювати ранги, обчислювати частоти значень, укрупнювати дані, графічно репрезентувати отримані дані.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення Microsoft Excel, SPSS або Jamovi (на вибір).

Завдання:

1. Завантажте за посиленням результати тестування вибірки (n=43) за допомогою тесту прогресивні матриці Равена: https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical2.docx
Зробіть сортування опитаних по зростанню (від меншого до більшого).
2. Привласніть опитаним ранги. Пам’ятайте, що спочатку привласнюються ранги від першого до останнього опитаного.
3. Обчисліть частоти, для кожного значення, яке спостерігається у вибірці.
4. Здійсніть групування частот у розряди оцінок.
5. Побудуйте графік розподілу частот: гістограму та полігон частот.

Проблемні питання. У яких випадках використання рангів може бути більш інформативним, ніж робота з «сирими» балами? Як вибір ширини інтервалів (розрядів) при групуванні частот впливає на форму гістограми? Які можуть бути спотворення? Чим полігон частот відрізняється від гістограми? У яких випадках полігон частот дає більше інформації?

Рекомендована література:

1. Руденко В.М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.
2. Glass G. V., Hopkins K. D. Statistical methods in education and psychology. Boston : Allyn & Bacon, 1996. 674 p.

Практичне заняття 3. Розрахунки та інтерпретація ММ та МЦТ.

Мета: сформувати практичні навички розрахунку ММ та МЦТ на основі результатів психологічного дослідження.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення Microsoft Excel, SPSS або Jamovi (на вибір).

Завдання:

1. Завантажте дані за посиланням: https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical3.docx у відповідне програмне забезпечення (Microsoft Excel, SPSS або Jamovi).
2. Розрахуйте в програмному забезпеченні: міри центральної тенденції та міри мінливості.
3. Проаналізуйте отримані результати.
4. Розрахуйте письмово показники моди, медіани та середнього для вказаної в табл. 1 сукупності. Напишіть висновок з якісним аналізом змін, які відбулися при зміні сьомого елемента сукупності. Поясніть причини змін чи їх відсутності.

Таблиця 1

Описові статистики емпіричних даних

Спостережувані дані													**МЦТ**		
I	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*Mo*	*Md*	*X*
B.1	*1*	*2*	*2*	*3*	*4*	*5*	*5*	*5*	*5*	*6*	*7*	*8*			
B.2	*1*	*2*	*2*	*3*	*4*	*5*	*15*	*5*	*5*	*6*	*7*	*8*			

Проблемні питання. Які переваги і недоліки кожної міри центральної тенденції – моди, медіани та середнього? Які висновки можна зробити про стійкість центральних тенденцій до аномальних або екстремальних значень? Чи можна стверджувати, що медіана завжди краще за середнє арифметичне? Які ситуації можуть бути винятками?

Рекомендована література:

1. Руденко В.М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.

Практичне заняття 4. Закон нормального розподілу та його застосування.

Мета: сформувати практичні навички обчислення показників, які дають інформацію про відповідність вибіркового розподілу теоретичному нормальному розподілу. Формування уявлення про принципи, на основі яких створюються норми.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. Імпортуйте дані розміщені за посиланням: https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical4.docx
Розрахуйте для шкал показники середнього, стандартного відхилення, асиметрії та ексцесу, статистики Шапіро-Уілка та Колмогорова-Смірнова. Побудуйте гістограму з кривою нормального розподілу. Зробіть висновок про відповідність вибіркового розподілу теоретичному нормальному.
2. Напишіть формулу перекодування шкал методики в інші змінні для розрахунку нормативних значень на основі наступної логіки: менше $(x - \sigma)$ = низькі значення, від $(x - \sigma)$ до $(x + \sigma)$ = середні значення, більше $(x + \sigma)$ = високі значення. Зробіть z-перетворення шкал методики та порівняйте Ваші нормативні значення з z-балами.

3. Виділіть групи опитаних, які мають високі та низькі значення за шкалами методики на основі виділених Вами нормативних значень. Інсталюйте в Jamovi модуль DistrACTION. Відкрийте в Jamovi базу даних «Амтхауер_Равен.xlsx», яка прикріплена до практичної. Зробіть передбачення з якою ймовірністю у вибірці зустрічалось би у вибірці значення за методикою Равена менше 39 балів (за умови, якщо вибірка є нормально-розподілена). А яке значення Ви спостерігаєте в розподілі частот таких балів? Зробіть передбачення з якою ймовірністю має у Вас зустрічатись частота значень більше 55,7? А яке значення Ви спостерігаєте в розподілі частот таких балів?

Проблемні питання. Чому у психологічних дослідженнях часто припускають нормальний розподіл даних. Що може вказувати на те, що ваші дані не відповідають нормальному розподілу, навіть якщо гістограма «схожа» на дзвіноподібну криву? Що показує асиметрія і чому важливо розуміти її знак і величину? Яке практичне значення показника ексцесу? Чим розподіл з високим ексцесом відрізняється від нормального? Чому одночасно використовують кілька тестів нормальності? Чи може їхній результат відрізнятися і чому?

Рекомендована література:

1. Glass G. V., Hopkins K. D. Statistical methods in education and psychology. Boston : Allyn & Bacon, 1996. 674 p.

2. Боснюк В. Ф. Математичні методи в психології: курс лекцій. Мультимедійне видання. Харьків : НУЦЗУ, 2020. 141 с.

3. Лабораторний практикум «Статистичний аналіз SPSS» / Упор. Матковський С. О. та інш. Львів : Львівський національний університет імені Івана Франка, 2020. 126 с.

Практичне заняття 5. Коефіцієнт лінійної кореляції Пірсона.

Мета: закріпити теоретичні знання та сформувати навички проведення кореляційного аналізу з використанням коефіцієнта лінійної кореляції Пірсона.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. За посиланням https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical5.docx – результати тестування студентів за: методикою соціального інтелекту Гілфорда, емоційного інтелекту Холла, 6 субтест тесту Структура інтелекту Амтхауера, результатами тесту Прогресивні матриці Равена. Збережіть дані на жорсткий диск комп'ютера та відкрийте в програмі SPSS або Jamovi.
2. Перевірте чи є дані за шкалами нормально розподілені. Для цього обчисліть статистику Колмогорова-Смірнова та Шапіро-Уілка для кожної окремої шкали.
3. Побудуйте коробковий графік для всіх змінних та проаналізуйте чи є опитані, чиї значення є екстремальними (виходять за межі 1,5*IQR).
4. Якщо є спостереження, які суперечать ймовірнісним очікуванням – відфільтруйте, або виключіть таких опитаних з підрахунку. Обчисліть заново статистику Колмогорова-Смірнова.
6. Обчисліть кореляції за Пірсоном між змінними, які відповідають вимогам до застосування критерію. Проаналізуйте отримані дані (оцініть силу зв'язку та ймовірність помилки першого порядку) та зробіть їх психологічну інтерпретацію.

Проблемні питання. Чому важливо перевіряти нормальність розподілу перед застосуванням коефіцієнта Пірсона? Які наслідки можливі, якщо проігнорувати цю умову? Які альтернативні методи можна застосувати, якщо дані не відповідають нормальному розподілу? Чому поодинокі аномальні значення можуть суттєво спотворювати коефіцієнт Пірсона? Як це впливає на

достовірність висновків? Чому важливо аналізувати не тільки числове значення кореляції, а й коробковий графік?

Рекомендована література:

1. Жалдак М. І. Теорія ймовірностей і математична статистика : підручник для студентів. Київ : Національний педагогічний університет імені Н. М. Драгоманова, 2017. 707 с.

Практичне заняття 6. Коефіцієнт рангової кореляції Спірмена та міри зв'язку для номінальних шкал.

Мета: закріпити теоретичні знання та сформувати навички проведення кореляційного аналізу з використанням коефіцієнта рангової кореляції Спірмена. Отримати навики обчислення міри зв'язку між номінальними шкалами на основі критерію Пірсона та Крамера-V.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. За посиланням https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical6.docx результати тестування студентів за: методикою соціального інтелекту Гілфорда (+ 4 його субшкали), емоційного інтелекту Холла (+ 5 його субшкал), 6 субтестом тесту Структура інтелекту Амтхауера, результатами тесту Прогресивні матриці Равена. Збережіть дані на жорсткий диск комп'ютера та відкрийте в програмі SPSS або Jamovi.
2. Ваше перше завдання перевірити нормальність розподілу даних змінних за критерієм Шапіро-Уілка й зробити висновок про розподіл даних за шкалами.
3. Обрахуйте коефіцієнт рангової кореляції Спірмена в SPSS або Jamovi між усіма змінними у наборі даних.
4. Проаналізуйте отримані дані (оцініть силу зв'язку та ймовірність помилки першого порядку) та зробіть їх психологічну інтерпретацію.

5. Обчисліть описову статистику (середнє, та стандартне відхилення) для шкал: Загальна шкала соціального інтелекту Гілфорда та Загальна шкала емоційного інтелекту Холла.

6. Перекодуйте значення за цими двома шкалами у нормативні значення на основі середнього та стандартного відхилення (на 3 групи: «високий» - [x + σ, x + 3σ], «середній» - [x − σ, x + σ], «низький» - [x − σ, x − 3σ]).

7. Обрахуйте коефіцієнт взаємної зв'язаності Пірсона та Крамера-V в SPSS за вказаними вище змінними перекодованими змінними.

8. Проаналізуйте та зробіть психологічну інтерпретацію отриманих даних.

Проблемні питання. Чим коефіцієнт рангової кореляції Спірмена принципово відрізняється від коефіцієнта лінійної кореляції Пірсона? У яких ситуаціях Спірмена слід використовувати обов'язково? Як впливають екстремальні значення на результати Пірсона та Спірмена? Чому Спірмен стійкіший? Чи можна порівнювати між собою результати Пірсона і Спірмена для однієї й тієї ж пари змінних? Що робити, якщо вони сильно відрізняються? Чому для номінальних шкал коефіцієнт Спірмена непридатний? Які помилки виникають, якщо використовувати рангові методи для номінальних даних?

Рекомендована література:

1. Жалдак М. І. Теорія ймовірностей і математична статистика : підручник для студентів. Київ : Національний педагогічний університет імені Н. М. Драгоманова, 2017. 707 с.

Практичне заняття 7. Побудова моделі лінійної та множинної регресії.

Мета: закріпити знання про сутність одномірної та множинної лінійної регресії, сформувати практичні навички обробки психологічних даних з допомогою регресійного аналізу.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. У розміщеному за посиланням файлі https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical7.docx дані дослідження мотивації, когнітивних здібностей, продуктивності та рівня стресу співробітників. Завантажте файли на комп'ютер та відкрийте їх у SPSS або Jamovi.
2. Дослідить нормальність розподілу змінних. Вивчіть кореляцію між ними, врахувавши нормальність розподілу та обравши відповідний кореляційний тест (Спірмена чи Пірсона).
3. Побудуйте модель лінійної регресії (Залежна зміна – продуктивність). Для цього – оберіть фактор з найвищою кореляцією. Побудуйте регресійну модель. Вивчіть наскільки вона є прогностичною та точною.
4. Побудуйте регресійне рівняння за отриманими даними.
5. Проаналізуйте отримані дані та зробіть їх психологічну інтерпретацію.
6. Додайте до регресійної моделі інші фактори, які мають достовірну кореляцію з залежною змінною, яка вивчається. За допомогою використання різних моделей побудови множинної регресії оберіть кількість змінних з максимальним значенням передбачення та валідності.
7. Побудуйте рівняння множинної регресії за отриманими даними.
8. Проаналізуйте отримані дані та зробіть їх психологічну інтерпретацію.

Проблемні питання. Які припущення лежать в основі методу лінійної регресії? Що буде, якщо вони порушуються? Чому перед побудовою регресійної моделі важливо перевіряти нормальність змінних та силу кореляцій? Чим регресія відрізняється від кореляції? Які додаткові питання вона дозволяє вирішувати у порівнянні з кореляційним аналізом? Чому для лінійної регресії важливо вибирати незалежну змінну, яка має найсильніший зв’язок із залежною? Що станеться, якщо зв’язок дуже слабкий? Що таке коефіцієнт детермінації (R^2) і як його значення впливає на довіру до моделі?

Рекомендована література:

1. Руденко В.М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.

Практичне заняття 8. Методи застосування кластерного аналізу у психології

Мета: закріпити у студентів теоретичні знання та сформувати практичні навички обробки психологічних даних з допомогою кластерного аналізу.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. Завантажте файл за посиланням: https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical8.docx Відкрийте файли додані до практичної використовуючи SPSS Jamovi. Для роботи з кластерним аналізом в програмі Jamovi потрібно інсталювати модуль SnowCluster. Завантажте файл.
2. За допомогою ієрархічного кластерного аналізу виокремте типи (кластери) студентів за змінними критичного мислення та психологічного благополуччя. Визначитесь з оптимальною кількістю кластерів на основі вивчення діаграм розсіювання на кожному кроці групування опитаних в кластери.
3. Порівняйте та опишіть середні значення досліджуваних змінних для кожного типу (кластеру) студентів.
4. За допомогою кластерного аналізу к-середніх виокремте типи (кластери) студентів за змінними стресу та психологічного благополуччя. Визначитесь з оптимальною кількістю кластерів на основі вивчення діаграм розсіювання на кожному кроці групування опитаних в кластери.
5. Проаналізуйте отримані дані та зробіть їх психологічну інтерпретацію.

Проблемні питання. Чим кластерний аналіз принципово відрізняється від регресійного та кореляційного аналізу? Які запитання він дозволяє дослідникові вирішити? Чому при ієрархічному кластерному аналізі важливо правильно вибрати метод об’єднання і метрику відстані? Які помилки можуть виникнути? Які критерії дослідник може використати для визначення «оптимальної» кількості кластерів? Чому це рішення не завжди однозначне?

Які переваги і недоліки має метод к-середніх у порівнянні з ієрархічним кластерним аналізом?

Рекомендована література:

1. Боснюк В. Ф. Математичні методи в психології : курс лекцій. Мультимедійне навчальне видання. Харків : НУЦЗУ, 2020. 141 с.

2. Лабораторний практикум «Статистичний аналіз SPSS» / Упор. Матковський С. О. та інш. Львів : Львівський національний університет імені Івана Франка, 2020. 126 с.

Практичне заняття 9. Факторний аналіз та його призначення.

Мета: закріпити теоретичні знання та сформувати навички обробки психологічних даних з допомогою факторного аналізу.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. За посиланням: https://osf.io/e8tj9 наведені дані зібрані для адаптації О. Литвиненко та Я. Коркосом методики «Шкала автономної кар'єрної мотивації. Відкрийте прикріплені файли в SPSS чи в Jamovi.
2. Оцініть чи данні підходять для факторного аналізу: наявність достатньої кількості спостережень (зазвичай рекомендується мінімум 5-10 спостережень на змінну), критерій сферичності Бартлетта, міра адекватності вибірки Кайзера-Мейєра-Олкіна (КМО)
3. Перевірте факторну структуру опитувальника. Визначити, які фактори визначають відповіді на запитання. Проведіть факторизацію відповідей на запитання використовуючи метод головних компонент та обертання «Varimax».
4. Порівняйте виділені фактори з оригінальними ключами до методики (прикріплені до практичного) та зробіть висновок про факторну структуру опитувальника та конструктну валідність методики.

5. Спробуйте використати косоугольний метод обертання Direct Oblimin. Порівняйте виділені фактори на основі обертання Verimax та Direct Oblimin. Спробуйте застосувати інші методи факторного аналізу.

6. Проекспериментуйте з різними типами виділення факторів (альфа-факторнизація, метод головних осей) використовуючи метод обертання Varimax.

7. Проаналізуйте отримані дані та зробіть їх психологічну інтерпретацію.

Проблемні питання. Чим факторний аналіз відрізняється від кластерного? Які питання дозволяє дослідити факторний аналіз у психології? Чому важливо перевіряти критерій сферичності Бартлетта та KMO перед проведенням факторного аналізу? Що може статися, якщо ці умови ігнорувати? Що означає «виділення головних компонент»? Чому цей метод часто використовують для зведення даних? Яка роль обертання факторів (Varimax, Direct Oblimin) у факторному аналізі? Яка принципова різниця між ортогональним (Varimax) та косокутним (Direct Oblimin) обертанням? Коли варто використовувати одне, а коли інше?

Рекомендована література:

1. Лабораторний практикум «Статистичний аналіз SPSS» / Упор. Матковський С. О. та інш. Львів : Львівський національний університет імені Івана Франка, 2020. 126 с.

2. Климчук В.О. Математичні методи у психології. Навч. посіб. для студентів психологічних спеціальностей. Київ : Освіта України, 2009. 288 с.

Практичне заняття 10. Утворення факторів другого порядку. Конфірматорний факторний аналіз.

Мета: сформувати у студентів практичні навички використання конфірматорного факторного аналізу для перевірки факторних структур у психологічних дослідженнях.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у

психології»; 2) програмне забезпечення SPSS (модуль Amos Graphicts) або Jamovi (модуль SEM).

Завдання:

1. Відкрийте за посиланням https://osf.io/e8tj9 базу даних з відповідями на запитання опитувальника «Шкала автономної кар'єрної мотивації».
2. Проведіть факторний аналіз питань опитувальника методом Головні компоненти з обертанням Direct Oblimin. Автоматично збережіть виділені фактори.
3. Проведіть вторинний факторний аналіз факторів першого порядку. Опишіть виділені вторинні фактори та дайте їм психологічну інтерпретацію.
4. Перевірте факторну модель виділених вторинних факторів методом конфірматорного факторного аналізу (CFA) використовуючи опцію SEM в Jamovi або Amos Graphicts в SPSS. Перевірте фактори відповідно до критеріїв CFI, TLI, RMSEA, SRMR (CFI/TLI мають бути більше 0,90, RMSEA/SRMR мають бути менше 0,08).
5. Уточніть психологічні інтерпретації виділених вторинних факторів.

Проблемні питання. Чим принципово відрізняється конфірматорний факторний аналіз (CFA) від експлораторного (EFA)? Чому обидва підходи важливі? Які підстави можуть виправдовувати утворення факторів другого порядку? Наведіть приклади, де це має сенс. Чому просто виділити фактори першого порядку недостатньо для перевірки валідності складної структури? Які проблеми можуть свідчити про низькі значення CFI/TLI або високі RMSEA/SRMR? Як це інтерпретувати у контексті опитувальника? Чому у CFA модель задається заздалегідь, а не виводиться автоматично, як в EFA?

Рекомендована література:

1. Лабораторний практикум «Статистичний аналіз SPSS» / Упор. Матковський С. О. та інш. Львів : Львівський національний університет імені Івана Франка, 2020. 126 с.
2. Руденко В.М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.

Практичне заняття 11. Порівняння незалежних вибірок за параметричними критеріями.

Мета: закріпити у студентів теоретичні знання та сформувати практичні навички обробки психологічних даних за допомогою параметричних критеріїв: одно та двох вибіркового критерію Стюдента, дисперсійного аналізу та однофакторного дисперсійного аналізу.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. Відкрийте прикріплені файли за посиланням https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical11.xlsx у SPSS або в Jamovi. У файлі представлені дані п'ятифакторного опитувальника, шкали прокрастинації, шкали стресу. У змінній «sex» чоловіча стать закодована цифрою «0», жіноча – «1».
2. Здійсніть перевірку на нормальність розподілу (тести Колмогорова-Смірнова, Шапіро-Вілка) за всім кількісними шкалами. В основі групування даних повинен бути фактор статі (*sex*).
3. Виберіть 2 змінні, що мають нормальний розподіл. Наприклад, Самоконтроль (*Self control*) та Емоційна стабільність (*Emotional stability*). Порівняйте групи за статтю на основі критерію Стюдента для незалежних вибірок та F-критерієм Фішера.
4. Дайте психологічну інтерпретації отриманих даних.
5. Проведіть однофакторний дисперсійний аналіз ANOVA для груп з різним рівнем прокрастинації за всіма змінними з нормальним розподілом. Оберіть критерій Шеффе (*Scheffé*) та Геймса-Хауелла (*Games-Howell*). для порівняння середніх. Опишіть, між якими групами виявлена достовірна відмінність. Дайте психологічну інтерпретацію отриманих результатів.

Проблемні питання. Яке припущення про дисперсії лежить в основі t-критерію для незалежних вибірок? Як перевірити це припущення? Чим відрізняється F-критерій Фішера від t-критерію Стюдента? Чому обидва можуть використовуватися у порівнянні двох груп? Які наслідки має використання ANOVA при порушенні припущення про рівність дисперсій у групах? Чому після проведення ANOVA використовують постхок тести (Шеффе, Геймс-Гауелл)? Чим вони відрізняються і для чого потрібні? Які помилки можна припустити при трактуванні ANOVA без постхок тестів?

Рекомендована література:

1. Боснюк В. Ф. Математичні методи в психології : курс лекцій. Мультимедійне навчальне видання. Харків : НУЦЗУ, 2020. 141 с.
2. Лабораторний практикум «Статистичний аналіз SPSS» / Упор. Матковський С. О. та інш. Львів : Львівський національний університет імені Івана Франка, 2020. 126 с.
3. Фетісов В. С. Пакет статистичного аналізу даних STATISTICA : навчальний посібник. Ніжин : НДУ ім. М. Гоголя, 2018. 114 с.

Практичне заняття 12. Порівняння незалежних вибірок за непараметричними критеріями.

Мета: закріпити теоретичні знання та сформувати навички застосування непараметричних критеріїв для обробки результатів психологічного дослідження.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. Завантажте за посиланням https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical12.docx результати тестування опитувальником Р. Кеттелла та відкрийте дані в SPSS або в Jamovi.

2. Ваше перше завдання полягає у дослідженні нормальності розподілу (тест Шапіро-Уілка, Колмогорова-Смірнова) факторів 16-факторного опитувальника Р. Кеттелла, групуючи дані за ознакою гендеру. Серед змінних є стовпчик підписаний gender, в ньому зазначена стать опитаного. Цифрою «1» позначені чоловіки, «2» – жінки.

3. Порівняйте психологічні особливості чоловіків та жінок для ознак, що маю не нормальний розподіл даних. Використайте критерій Манна-Уітні для того, щоб знайти, які особливості (фактори Кеттелла) відрізняють групи чоловіків та жінок з великою достовірністю. Опишіть середні значення чоловіків та жінок за факторами, за якими знайдена відмінність та дайте психологічну інтерпретацію.

4. Оберіть до вподоби дві будь-які змінні, які можуть класифікувати опитаних (*наприклад: Ket_E та Ket_F*). За допомогою ієрархічного кластерного аналізу сформуйте 3-4 кластери опитаних з притаманними їм унікальними характеристиками. Опишіть виділені групи.

5. За допомогою критерію Краскелла-Уолліса дослідить чи є специфічні особливості виділених груп за іншими факторами Кеттелла. Опишіть середні значення у групах за факторами для яких Ви знайшли відмінності. Дайте психологічну інтерпретацію знайдених особливостей.

6. Перекодуйте значення за фактором С Кеттелла у дихотомічні значення згідно з прикріпленим до практичного файлу «*Практична робота 12.Ключ.sps*». Використовуючи критерій хі-квадрат (та поправку Фішера для таблиць 2х2) порівняйте групу чоловіків та жінок за перекодованим фактором С Кеттелла. Що характерно для групи чоловіків та для групи жінок з отриманих результатів? Яка достовірність отриманих результатів виходячи з вимог до застосування критерію хі-квадрат?

7. Проаналізуйте аналогічним чином і інші достовірні відмінності для груп студентів, які на момент тестування використовували для відповідей комп'ютер або смартфон. Для цього Ви можете скористатись змінною:

"comp_planchet" (використання комп'ютера або смартфона) в якому відповіді закодовані наступним чином: 1 – я виконую завдання на персональному комп'ютері або іншому пристрої для введення даних за допомогою клавіатури; 2 – я виконую завдання на планшеті; 3 – я виконую завдання на телефоні.

Проблемні питання. Чим непараметричні критерії (Манна-Уітні, Краскела-Уолліса, хі-квадрат) принципово відрізняються від параметричних? У яких ситуаціях вони обов’язково потрібні? Чому нормальність розподілу є критичною умовою для параметричних тестів, але не для непараметричних? Які сильні та слабкі сторони непараметричних критеріїв у порівнянні з параметричними? Чому критерій Манна-Уітні підходить для порівняння двох незалежних груп? Що він насправді порівнює – середні чи ранги? Чим відрізняється Краскелла-Уолліса від Манна-Уітні? Чому іноді їх обидва використовують у межах одного дослідження?

Рекомендована література.

1. Математичні методи в психології: методичні вказівки з організації та планування самостійної роботи студентів для здобувачів освітньокваліфікаційного рівня бакалавр за спеціальністю 053 – психологія / Упор. : В. О. Олефір. Харків, 2016. 59 с.
2. Фетісов В. С. Пакет статистичного аналізу даних STATISTICA : навчальний посібник. Ніжин : НДУ ім. М. Гоголя, 2018. 114 с.

Практичне заняття 13. Порівняння залежних вибірок.

Мета: закріпити теоретичні знання та сформувати навички застосування параметричних та непараметричних критеріїв для порівняння пов'язаних вибірок.

Перелік необхідного обладнання: 1) стандартизований робочий зошит**,** що доданий до електронного курсу «Математичні методи у психології»; 2) програмне забезпечення SPSS або Jamovi (на вибір).

Завдання:

1. Завантажте базу даних 10 студентів до та після релаксації за посиланням: https://pleskach.wordpress.com/wp-content/uploads/2025/10/practical13.docx Відкрийте її в SPSS або в Jamovi. Доповніть базу даних своїми результатами тестування. Обчисліть результати тестування по шкалам методики САН, використовуючи прикріплений до файлу ключ.

2. Здійсніть перевірку розподілу (тест Шапіро-Уілка, Колмогорова-Смірнова) даних за шкалами методики до та після релаксаційного втручання.

3. Використовуючи t-критерій для пов'язаних вибірок порівняйте результати до та після релаксації.

4. Використовуючи непараметричний критерій Уілкоксона порівняйте отримані результати до та після релаксації.

5. Дайте психологічну інтерпретацію отриманим результатам. Для інтерпретації результатів можна скористатись описом методики у збірнику, який прикріплений до практичного.

Проблемні питання. Чим принципово відрізняється порівняння незалежних і залежних вибірок? Як це впливає на вибір статистичного тесту? Чому важливо перевіряти нормальність розподілу різниць значень у залежних вибірках, а не просто окремих груп? Які припущення лежать в основі t-критерію для залежних вибірок? Що буде, якщо їх порушити? У чому сенс непараметричного критерію Вілкоксона для пов’язаних вибірок? Коли його краще використовувати, ніж t-критерій? Чому для одного й того ж набору даних результати t-критерію та критерію Вілкоксона можуть відрізнятись?

Рекомендована література.

1. Фетісов В. С. Пакет статистичного аналізу даних STATISTICA : навчальний посібник. Ніжин : НДУ ім. М. Гоголя, 2018. 114 с.

ПЛАНИ САМОСТІЙНИХ РОБІТ

Самостійна робота 1. Математичні методи в психології: загальна характеристика.

Мета: доповнити знання про місце математичних методів в психології та надійність вимірювань.

Завдання:

1. Розкрийте відмінності між буденним та науковим пізнанням та місце математичних методів в науковому пізнанні.
2. Проблема надійності та валідності вимірювань у психології. Приклади з історії психології відомих теорій та оцінка надійності/валідності пропонованих ними вимірювань (френологія В. Галля, конституційна теорія темпераменту Е. Кречмера).
3. Визначить, що саме може порушувати надійність вимірювань у психології. Складіть перелік. Чим відрізняються загрози ненадійності в експерименті і при тестуванні?
4. Розкрийте поняття змінна, шкала, дані. Розкрийте класифікацію емпіричних даних в психології на: "L","Q", "T" – типи даних. Наведіть приклади таких вимірювань.

Рекомендована література:

1. Goodwin C.J., Goodwin K.A. Research in psychology. Methods and design. New Jersey: Wiley, 2012. 480 p.

Самостійна робота 2. Основні поняття математичної обробки психологічних даних.

Мета: розширити знання про візуалізацію розподілу даних у вибірці.

Завдання:

1. Для наочного зображення статистичних розподілів використовують графіки та діаграми (діаграма розподілу, гістограма, полігон, кумулята, коробчастий графік тощо) – визначте їх особливості та області застосування.

2. Ознайомтеся із можливостями модулю побудови графіків SPSS: опишіть алгоритм побудови графіків в SPSS.

Рекомендована література:

1. Вдовенко В. В. Математичні методи в психології: навчально-методичний посібник. Кіровоград : ПП «Центр оперативної поліграфії» Авангард», 2017. 112 с.

2. Волков В. О. Математичні методи в психології: практикум. Горлівка : Ліхтар, 2013. 188 с.

Самостійна робота 3. Кореляційний аналіз.

Мета: проаналізувати методи кореляційного аналізу та особливості їх застосування в психологічних дослідженнях .

Завдання:

1. Проаналізуйте, які відмінності виникають при аналізі зв'язку двох змінних, що мають нормальний розподіл, виникають у випадку застосовування кореляції за Пірсоном за кореляції за Спірменом. Зробіть те саме для даних, де відсутній нормальний розподіл.
2. Доведіть, що вибірковий коефіцієнт кореляції є випадковою величиною.
3. Сутність та алгоритм процедури оцінки значимості параметрів взаємозв'язку.
4. Непараметричні методи оцінки взаємозв'язку та причини їхньої популярності у психології.

Рекомендована література:

1. Боснюк В.Ф. Математичні методи в психології: курс лекцій. Мультимедійне видання. Харьків : НУЦЗУ, 2020. 141 с.

2. Glass G.V., Hopkins K.D. Statistical methods in education and psychology. Boston : Allyn & Bacon, 1996. 674 p.

Самостійна робота 4. Математичні методи прогнозування

Мета: ознайомитись з особливостями застосування у психології таких багатомірних методів прогнозування, як одномірна лінійна регресія та множинний регресійний аналіз.

Завдання:

1. Розкрийте лінійний регресійний аналіз як метод прогнозування.
2. Проаналізуйте лінійні моделі регресії: одномірна лінійна регресія та множинна лінійна регресія. Обґрунтуйте відмінності між ними.
3. Використовуючи дані отримані під час практичних занять спробуйте побудувати модель прогнозу. Наприклад, між особистісними рисами за 16PF та (залежна змінна) соціометричним статусом студента.

Рекомендована література:

1. Климчук В.О. Математичні методи у психології. Навч. посібник для студентів психологічних спеціальностей. Київ : Освіта України, 2009. 288 с.
2. Fidell L. S., Tabachnick B. G. Using multivariate statistics. 7th ed. Pearson Education, 2019. 832 p.

Самостійна робота 5. Порівняння незалежних вибірок за непараметричними критеріями

Мета: закріпити практичні навики порівняння двох незалежних вибірок в статистичних пакетах обробки даних Jamovi чи SPSS.

Завдання:

1. Використовуючи дані з Додатку 2, проаналізувати як змінна "самопочуття" впливає на результати тестування за деякими шкалами опитувальника Р. Кеттелла.
2. Інформація про самопочуття опитаних знаходиться у змінній "самопочуття". Відповіді закодовані наступним чином: 3 = в доброму самопочутті; 2 = самопочуття погіршене, або почуваю себе погано.
3. Для порівняння використайте непараметричний критерій Манна-Уітні та порівняйте як відрізняються дві групи опитаних (в доброму і

погіршеному/поганому самопочутті) за факторами опитувальника Кеттелла, за якими факторами між групами є достовірна відмінність?

4. Опишіть середні значення за факторами опитувальника Кеттелла (для яких знайдені достовірні відмінності) та зробіть інтерпретацію отриманих результатів.

Рекомендована література:

1. Боснюк В.Ф. Математичні методи в психології: курс лекцій. Мультимедійне видання. Харьків : НУЦЗУ, 2020. 141 с.

2. Вдовенко В. В. Математичні методи в психології: навчально-методичний посібник. Кіровоград : ПП «Центр оперативної поліграфії» Авангард», 2017. 112 с.

ПИТАННЯ ДО ЕКЗАМЕНУ

1. Представлена невпорядкована вибірка, завдання – розрахувати моду.
2. Процедура упорядкування об’єктів в порівняно однорідні класи на основі попарного порівняння цих об’єктів за попередньо визначеним і виміряним критерієм - це: … (обрати правильну альтернативу).
3. Графічне зображення послідовності об’єднання об’єктів в кластери називається: … (обрати правильну альтернативу).
4. Представлена невпорядкована вибірка, завдання – визначити медіану.
5. Визначення p-рівня значущості при кореляційному аналізі Пірсона відбувається за допомогою: … (обрати правильну альтернативу).
6. Зменшення розмірності вихідних даних з ціллю їх економного опису за умови мінімальних втрат вихідної інформації - це головне завдання: … (обрати правильну альтернативу).
7. Представлена невпорядкована вибірка, завдання – розрахувати середнє арифметичне.
8. Якщо використовується зв'язок між трьома змінними, то можлива перевірка припущення про те, що зв'язок між двома змінними не залежить від впливу третьої змінної за допомогою: … (обрати правильну альтернативу).
9. До багатомірних методів прогнозування відносяться: … (обрати правильну альтернативу, яка містить лише методи прогнозування).
10. Чи правильне таке твердження "До видів регресійного аналізу відносяться парна і множинна регресія"?
11. Встановіть відповідність між поняттями математичної статистики та їх визначенням: гомогенність, гетерогенність, контрольна група, експериментальна група.
12. Встановіть відповідність між типами залежностей змінних X та У і їх визначенням: функціональна залежність, лінійна залежність в регресивній моделі, стохастична та кореляційна залежності.

13. Встановіть відповідність між різновидами функцій та їх тлумаченням: лінійний зв’язок, позитивний зв’язок, негативний зв’язок, монотонний зв’язок, немонотонний зв’язок.

14. Встановіть відповідність між різновидами графічних функцій та їх назвами: нелінійна немонотонна з негативним зв'язком, нелінійна монотонна з негативним зв'язком, лінійна з негативним зв'язком, лінійна з позитивним зв'язком, нелінійна монотонна з позитивним зв'язком.

15. Встановіть відповідність між силою кореляційного зв'язку та його показниками: слабкий, сильний, помірний, середній.

16. Встановіть відповідність між рівнями статистичної значущості та їх показниками: висока значущість кореляції, незначуща кореляція, тенденція достовірного зв'язку.

17. Встановіть відповідність між методами математичної статистики та їх призначенням: кореляційний аналіз, регресійний аналіз, факторний аналіз.

18. Знайдіть відповідність між показниками кореляції та особливостями їх вимірювання: абсолютна величина кореляції, знак кореляції, рівень статистичної значущості.

19. Встановіть відповідність між МЦТ та їх значенням: мода, медіана, середнє арифметичне.

20. Встановіть відповідність: наукова (теоретична) гіпотеза, статистична гіпотеза, експериментальна гіпотеза.

21. Основними показниками кореляційного аналізу є: … (обрати правильну альтернативу).

22. На дистантній моделі в математиці засновані: … (обрати правильну альтернативу).

23. Який з критеріїв дозволяє перевірити нормальність розподілу? … (обрати правильну альтернативу).

24. Для кореляції рангових змінних використовується: … (обрати правильну альтернативу).

25. Яка зі статистик не відноситься до мір центральної тенденції? … (обрати правильну альтернативу).

26. Конфірматорний факторний аналіз відокремлює дисперсію змінних, які пояснює фактор від: … (обрати правильну альтернативу).

27. Причина спільної мінливості кількох вихідних змінних - це: … (обрати правильну альтернативу).

28. Косокутні методи обертання дозволяють: … (обрати правильну альтернативу).

29. Показниками доцільності проведення факторного аналізу є: … (обрати правильну альтернативу).

30. Який тип графіка зображено? … (обрати правильну назву).

31. В основі множинного регресійного аналізу лежить: … (обрати правильну альтернативу).

32. Яке з тверджень про моду є хибним? … (обрати правильну альтернативу).

33. Для графічного аналізу кореляційного зв'язку між змінними застосовуються: … (обрати правильну альтернативу).

34. Оптимальне число кластерів в ієрархічному кластерному аналізі визначається: … (обрати правильну альтернативу).

35. Математична статистика спирається на: … (обрати правильну альтернативу).

36. Варімакс-обертання в факторному аналізі використовується для: … (обрати правильну альтернативу).

37. Як якісно оцінити лінійність (нелінійність) кореляції? … (обрати правильну альтернативу).

38. Факторні навантаження розрахуються на основі: … (обрати правильну альтернативу).

39. Яка міра центральної тенденції у розподілі має максимальну частоту? … (обрати правильну альтернативу).

40. Середину варіаційного ряду визначає … (обрати правильну альтернативу).

Рекомендована література

1. Бегун С. І. Статистика: навчальний посібник. Луцьк : Волинський національний університет імені Лесі Українки, 2022. 230 с.
2. Боснюк В. Ф. Математичні методи в психології: курс лекцій. Мультимедійне видання. Харьків : НУЦЗУ, 2020. 141 с.
3. Вдовенко В. В. Математичні методи в психології: навчально-методичний посібник. Кіровоград : ПП «Центр оперативної поліграфії» Авангард», 2017. 112 с.
4. Вельдбрехт О., Зінченко, Н., Тавровецька Н. Шкала самооцінки Розенберга: українська адаптація та психометричні властивості. Інсайт: психологічні виміри суспільства. 2025. № 13. С. 142–172. https://doi.org/10.32999/2663-970X/2025-13-7
5. Волков В. О., Яковець В. П. Математичні методи в психології: практикум. Горлівка : Ліхтар, 2013. 188 с.
6. Експериментальна психологія: методичні матеріали навчальної дисципліни/ розроб. Горобець Т.В., Нечипоренко О.В. Черкаси : Черкаський національний університет імені Богдана Хмельницького, 2012. 120 с.
7. Жалдак М. І. Теорія ймовірностей і математична статистика : підручник для студентів. Київ : Національний педагогічний університет імені Н. М. Драгоманова, 2017. 707 с.
8. Климчук В.О. Математичні методи у психології. Навчальний посібник для студентів психологічних спеціальностей. Київ : Освіта України, 2009. 288 с.
9. Лабораторний практикум «Статистичний аналіз SPSS» / Упор. Матковський С. О. та інш. Львів : Львівський національний університет імені Івана Франка, 2020. 126 с.
10. Малхазов О. Р. Експериментальна психологія: практичний посібник. Київ : Талком, 2020. 321 с.

11. Математичні методи в психології: методичні вказівки з організації та планування самостійної роботи студентів для здобувачів освітнього кваліфікаційного рівня бакалавр за спеціальністю 053 – психологія / Упор. : В. О. Олефір. Харків, 2016. 59 с.
12. Руденко В. М. Математична статистика. Навч. посіб. Київ: Центр учбової літератури, 2012. 304 с.
13. Савченко О., Колесніченко Л., Чужикова В., Береженна О. Цінність власного життя в підтримці резильєнтності військових: побудова медіаторної моделі. Інсайт: психологічні виміри суспільства. *Інсайт: психологічні виміри суспільства*, 2024. № 12. С. 66–95.
14. Фетісов В. С. Пакет статистичного аналізу даних STATISTICA : навчальний посібник. Ніжин : НДУ ім. М. Гоголя, 2018. 114 с.
15. Błachnio A., Przepiorka A., Pantic I. Association between Facebook addiction, self-esteem and life satisfaction: a cross-sectional study. *Computers in human behavior*. 2016. Vol. 55. P. 701–705. URL: https://doi.org/10.1016/j.chb.2015.10.026
16. Cochran W. G. Sampling techniques. New York : John Wile & Sons, 1977. 428 p.
17. Denniss R., Naneva S. Statistics with JASP: first steps for psychology students. Sheffield : University of Sheffield, 2025. 87 p.
18. Exploring diversity with statistics: step-by-step JASP guide / ed. by R. V. Walker. 2022 : University of Tennessee at Chattanooga. URL: https://scholar.utc.edu/open-textbooks/1.
19. Fidell L. S., Tabachnick B. G. Using multivariate statistics. 7th ed. Pearson Education, 2019. 832 p.
20. Glass G. V., Stanley J. C.. Statistical methods in education and psychology. New Jersey : Prentice-Hall Inc, 1970. 495 p.
21. Goodwin C. J., Goodwin K. A. Research in psychology. Methods and design. New Jersey: Wiley, 2012. 480 p.

22. Gupta A., Kumari S. Effect of CBT on metacognitive beliefs in depressive disorders. *Turk psikiyatri derg.* 2023. Vol 34, No. 2. Pp 8-88. DOI: 10.5080/u26398

23. Kline R. B. Principles and practice of structural equation modeling. 5th ed. London : Guilford Press, 2023. 494 p.

24. Mallery P. IBM SPSS statistics 29 step by step: a simple guide and reference. Taylor & Francis Group, 2024. 426 p.

25. Navarro D., Foxcroft D. Learning statistics with Jamovi. A tutorial for beginners in statistical analysis. Cambridge : Open Book Publishers, 2025. 490 p. URL: https://doi.org/10.11647/OBP.0333.

26. Navarro D., Foxcroft D., Faulkenberry T. J. Learning statistics with JASP: a tutorial for psychology students and other beginner. 429 p. URL: https://learnstatswithjasp.com.

27. Osorio-Spuler X., Illesca-Pretty M., Gonzalez-Osorio L. et al. Emotional exhaustion in nursing students. A multicenter study. *Revista da escola de enfermagem da USP*. 2023. No. 57. https://doi.org/10.1590/1980-220X-REEUSP-2022-0319en

28. Rezaie L., Vakili-Amani Y., Paschall E., Khazaie H. Personality profiles in paradoxical insomnia: a case-control study. *Sleep Science*. 2020. Vol. 14, No. 4. Pp. 242-248. DOI: 10.5935/1984-0063.20190148

29. Schwarz H. Stichprobenverfahren. Ein leitfaden zur anwendung statistischer schatzverfahren. Berlin: Oldenbourg Wissenschaftsverlag, 1975. 194 p.

30. Sengupta U., Singh A. R. Application of functional analytic psychotherapy to manage schizophrenia. *Indian psychiatry journal*. 2021. Vol. 30, No. 1. Pp. 55-61. DOI: 10.4103/ipj.ipj_10_20

31. Shrestha N. Factor analysis as a tool for survey analysis. *American journal of applied mathematics and statistics*. 2021. Vol. 9, no. 1. P. 4–11.

32. Wasserman L. All of statistics: a concise course in statistical inference. Springer, 2010. 468 p.

33. Yousef V., Zahra J., Yasaman-Zahra S., Fathola M. Association between common stressful life events and coping strategies in adults. *Journal of education and health promotion*. 2021. Vol. 10, Is. 1(301). DOI: 10.4103/jehp.jehp_519_20

Додаток 1

Схема вибору статистичного критерію

Задача / питання	**Шкала**	**Нормальність розподілу**	**Гомогенність дисперсій**	**Критерій**
Порівняння середнього з відомим значенням	Інтервальна/відношення	Нормальний розподіл	—	Одновибірковий t-критерій
	Інтервальна/відношення	Не нормальний	—	Критерій знаків або одновибірковий критерій Вілкоксона
Порівняння середніх двох незалежних груп	Інтервальна/відношення	Нормальний	Гомоскедастичність	Незалежний t-критерій
	Інтервальна/відношення	Нормальний	Гетероскедастичність	t-критерій з поправкою Уелча
	Порядкова	—	—	Манна-Уітні (U-критерій)
	Номінальна	—	—	χ^2-критерій
Порівняння двох залежних вимірювань	Інтервальна/відношення	Нормальний	—	Парний t-test
	Порядкова	—	—	Критерій Вілкоксона
Порівняння трьох і більше груп (незалежних)	Інтервальна/відношення	Нормальний	Гомогенність дисперсій	Однофакторна ANOVA
	Інтервальна/відношення	Нормальний	Гетерогенність дисперсій	Однофакторний дисперсійний аналіз з поправкою Уелча
	Порядкова	—	—	Критерій Краскела – Уолліса
	Номінальна	—	—	χ^2-критерій
Порівняння трьох і більше пов’язаних груп	Інтервальна/відношення	Нормальний	—	ANOVA для повторних вимірювань,
	Порядкова	—	—	Критерій Фрідмана
Перевірка нормальності	—	—	—	Критерій Колмогорова-Смірнова, Шапіро-Уілка
Перевірка гомогенності дисперсій	—	—	—	Критерій Левене,
Зв’язок між двома змінними	Інтервальна/відношення	Нормальний	—	Критерій лінійної кореляції Пірсона
	Порядкова	—	—	Критерій рангової кореляції Спірмена
	Номінальна	—	—	χ^2-критерій

Додаток 2

Дані до самостійної 5

Опитаний	Самопочуття	Деякі фактори опитувальника Р. Кеттелла											
		MD	A	B	C	E	F	G	H	I	L	M	N
1	3	5	7	6	7	8	7	4	8	10	7	6	3
2	3	5	6	6	5	1	5	8	4	7	9	6	8
3	3	3	4	4	5	8	4	8	7	10	4	8	4
4	2	6	5	4	7	6	5	9	8	9	4	4	6
5	2	5	8	5	4	5	7	6	7	10	6	9	5
6	3	5	6	5	2	6	2	6	9	8	7	8	4
7	3	4	10	4	3	2	8	3	10	8	8	7	5
8	2	5	7	7	9	6	6	9	9	9	8	5	5
9	3	5	9	4	3	5	4	7	5	8	4	12	2
10	3	5	7	4	5	3	5	9	1	9	4	9	3
11	2	6	6	5	5	5	6	7	4	7	5	6	4
12	2	2	6	4	5	6	6	10	6	9	8	6	5
13	2	4	6	5	3	4	3	9	9	9	4	7	4
14	3	7	6	6	7	9	6	11	11	8	6	4	1
15	3	4	4	3	2	6	6	8	10	8	4	8	4
16	3	4	6	1	6	8	6	8	8	8	4	6	2
17	3	6	4	1	6	2	2	8	4	10	8	6	6
18	3	6	8	3	6	8	8	10	6	8	2	4	4
19	3	2	8	1	4	2	4	10	6	8	4	8	2
20	3	8	10	3	10	10	6	12	12	8	6	6	2
21	3	10	8	3	6	8	2	4	10	10	4	8	2
22	3	6	6	3	6	4	0	4	6	8	6	4	8
23	3	6	6	2	4	8	0	6	4	6	4	6	4
24	2	4	8	4	8	4	8	8	6	6	2	4	4
25	3	6	10	2	2	8	2	10	8	6	4	4	6
26	3	2	6	1	4	6	4	8	2	12	4	6	6
27	3	6	8	4	6	6	6	6	10	6	8	6	4
28	2	6	5	7	9	7	2	8	6	9	3	6	2
29	3	3	6	4	7	7	5	10	8	5	6	5	8
30	3	6	8	4	9	1	4	10	2	9	8	10	7
31	3	5	6	7	4	4	3	7	2	9	4	6	5
32	3	5	6	5	11	5	4	8	7	8	4	4	6
33	2	5	3	4	2	4	1	10	3	7	6	11	3
34	3	8	3	5	7	7	1	8	9	4	5	8	0
35	3	6	8	5	3	7	4	5	11	7	8	7	6
36	2	5	4	4	5	2	2	7	3	10	7	9	6
37	2	6	8	3	6	10	8	7	9	7	7	6	5
38	3	8	10	5	9	5	6	6	9	9	4	7	2
39	3	3	4	6	5	9	1	1	12	10	5	4	2
40	2	3	5	7	8	7	9	8	8	9	6	5	4
41	3	4	8	4	4	6	6	6	5	10	6	8	6
42	2	10	2	5	9	7	6	10	10	5	7	6	2
43	3	5	7	4	8	10	8	10	11	8	5	5	4
44	2	6	6	4	8	5	5	10	10	7	7	9	4
45	2	3	8	5	2	8	2	12	7	7	3	6	2

46	3	7	10	5	7	2	5	5	2	8	6	7	6
47	2	3	6	3	2	6	7	10	10	8	4	5	6
48	3	4	8	5	3	4	1	10	4	9	4	8	6
49	3	2	4	5	6	2	0	6	3	11	6	6	10
50	3	3	5	4	2	6	9	9	8	6	7	8	6
51	2	4	0	6	3	7	1	10	4	10	8	8	10
52	3	9	6	2	8	4	7	6	9	10	7	7	5
53	2	4	7	3	6	8	5	5	8	9	5	8	8
54	2	5	4	5	2	4	0	7	6	8	4	10	3
55	3	7	5	6	7	8	3	11	7	7	3	7	6
56	2	8	6	5	7	5	3	10	6	10	7	6	5
57	3	5	8	3	7	8	5	6	6	7	5	8	2
58	3	6	5	4	4	6	5	11	8	5	7	6	4
59	2	6	6	6	6	7	2	10	8	7	3	7	5
60	2	6	6	5	5	5	0	10	1	4	7	8	4
61	3	3	7	4	4	9	8	3	10	3	4	2	4

Предметний покажчик

Н

О

П

Р

С

Т

Для нотаток

Для нотаток

Навчальне видання

ПЛЕСКАЧ Богдан Вадимович
КОРКОС Ярослав Олександрович

МАТЕМАТИЧНІ МЕТОДИ В ПСИХОЛОГІЇ

Навчально-методичний посібник

Підписано до друку 07.11.2025. Формат 60×84/16.
Папір офсетний. Друк цифровий.
Гарнітура Times New Roman. Ум. друк. арк. 10,46.
Наклад 50 прим. Зам. №2/11/25.

Видавець і виготовлювач: ФОП Панов А. М.
Свідоцтво серії ДК №4847 від 06.02.2015 р.
м. Харків, вул. Жон Мироносиць, 10, оф. 6,
тел.: +38(057)714-06-74, +38(050)976-32-87,
copy@vlavke.com