В. М. Краєвський, Я. О. Остапенко,
Т. М. Паянок, Н. В. Параниця

МЕТОДИ АНАЛІЗУ ДАНИХ У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ



ДЕРЖАВНИЙ ПОДАТКОВИЙ УНІВЕРСИТЕТ

В. М. Краєвський, Я. О. Остапенко,
Т. М. Паянок, Н. В.Параниця

МЕТОДИ АНАЛІЗУ ДАНИХ У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ

Навчальний посібник

Ірпінь
2024

УДК [33.021+005.52]:159.9(075.8)
ББК 65.053+88я73
М 54

Рекомендовано до друку Вченою радою
Державного податкового університету
(протокол № 5 від 30 жовтня 2024 року)

Рецензенти:

В. К. Савчук, д-р екон. наук, професор, професор кафедри статистики та економічного аналізу Національного університету біоресурсів і природокористування України;

А. В. Лісовий, д-р екон. наук, професор, завідувач кафедри аудиту, державного фінансового контролю та аналізу Державного податкового університету.

Автори:

Краєвський В. М., д-р екон. наук, професор, професор кафедри облікових технологій та бізнес-аналітики, декан факультету податкової, справи, обліку та аудиту; **Остапенко Я. О.**, канд. екон. наук, доцент, доцент кафедри облікових технологій та бізнес-аналітики; **Паянок Т. М.**, канд. екон. наук, доцент, завідувач кафедри облікових технологій та бізнес-аналітики; **Параниця Н. В.**, канд. екон. наук, доцент, доцент кафедри облікових технологій та бізнес-аналітики.

.

М 54 **Методи аналізу даних у психологічних дослідженнях :** навчальний посібник / В. М. Краєвський, Я. О. Остапенко, Т. М. Паянок, Н. В. Параниця ; Державний податковий університет. – Ірпінь, 2024. – 144 с.
ISBN 978-966-337-743-8

Навчальний посібник спрямований на формування професійних практичних компетенцій щодо використання методів аналізу даних у психологічних дослідженнях з метою прийняття ефективних професійних рішень. Охоплює основні теми навчальної дисципліни «Методи аналізу даних у психологічних дослідженнях». Кожна з тем має теоретичні питання, практичні завдання з типовими прикладами розв'язку, питання та тести для самоконтролю. Наприкінці подано глосарій.

Структура посібника дозволяє ефективно його використовувати для очного, заочного та дистанційного навчання.

Розрахований на здобувачів вищої освіти психологічних спеціальностей, слухачів післядипломної освіти, курсів перепідготовки, працівників підприємницьких структур, широкого кола фахівців з психологічних досліджень і для самостійного вивчення дисципліни.

УДК [33.021+005.52]:159.9(075.8)
ББК 65.053+88я73

© В. М. Краєвський, Я. О. Остапенко,
Т. М. Паянок, Н. В. Параниця, 2024

ISBN 978-966-337-743-8

© Державний податковий університет, 2024

ЗМІСТ

ПЕРЕДМОВА 6

**ТЕМА 1. ОСНОВНІ ПОНЯТТЯ,
ЯКІ ВИКОРИСТОВУЮТЬ ПІД ЧАС АНАЛІЗУ
ДАНИХ У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ** 8
1.1. Поняття аналізу даних у психологічних дослідженнях 8
1.2. Методи аналізу даних 9
1.3. Вимірювання та шкали в психологічних дослідженнях 12
Перелік питань для самоконтролю 17
Тести для самоконтролю 17

**ТЕМА 2. ВИБІРКОВИЙ МЕТОД
У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ** 20
2.1. Поняття вибіркового методу 20
2.2. Види вибірки та способи відбору 21
2.3. Обчислення похибок вибірки та визначення
потрібного обсягу вибірки 23
Приклади розв'язку типових задач 25
Перелік питань і завдань для самоконтролю 27
Тести для самоконтролю 28

**ТЕМА 3. МЕТОДИ ОПИСОВОЇ СТАТИСТИКИ
В ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ** 31
3.1. Міри центральної тенденції 31
3.2. Міри мінливості даних 32
3.3. Міри положення 34
Приклади розв'язку типових задач 35
Перелік питань і завдань для самоконтролю 38
Тести для самоконтролю 39

**ТЕМА 4. МЕТОДИ АНАЛІЗУ ВЗАЄМОЗВ'ЯЗКУ
В ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ** 42
4.1. Поняття взаємозв'язків між явищами
та процесами психології. Види взаємозв'язків 42
4.2. Основи кореляційного аналізу 43

4.3. Основи регресійного аналізу 46
4.4. Моделі зв'язку в порядкових вимірюваннях.
Рангові кореляції 48
4.5. Оцінка тісноти взаємозв'язку
між номінальними даними 50
Приклади розв'язку типових задач 53
Перелік питань і завдань для самоконтролю 58
Тести для самоконтролю 59

**ТЕМА 5. ДИСПЕРСІЙНИЙ АНАЛІЗ
У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ** 63
5.1. Основні поняття дисперсійного аналізу 63
5.2. Однофакторний дисперсійний аналіз 64
5.3. Двофакторний дисперсійний аналіз 67
Приклади розв'язку типових задач 68
Перелік питань і завдань для самоконтролю 70
Тести для самоконтролю 71

**ТЕМА 6. ФАКТОРНИЙ І КЛАСТЕРНИЙ АНАЛІЗИ
В ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ** 73
6.1. Факторний аналіз 73
6.2. Кластерний аналіз 75
Приклади розв'язку типових задач 77
Перелік питань і завдань для самоконтролю 87
Тести для самоконтролю 88

**ТЕМА 7. МЕТОДИ ДОСЛІДЖЕННЯ
Й АНАЛІЗУВАННЯ ПСИХОЛОГІЧНИХ ПРОЦЕСІВ
У ДИНАМІЦІ ТА ТЕНДЕНЦІЇ РОЗВИТКУ** 91
7.1. Методи дослідження й аналізування
психологічних процесів у динаміці 91
7.2. Методи аналізування тенденції розвитку
в психологічних дослідженнях 95
Приклади розв'язку типових задач 97
Перелік питань і завдань для самоконтролю 101
Тести для самоконтролю 103

ТЕМА 8. МЕТОДИ ВІЗУАЛІЗАЦІЇ ТА ПОДАННЯ ДАНИХ ПСИХОЛОГІЧНИХ ДОСЛІДЖЕНЬ 106
8.1. Статистична таблиця як метод подання даних 106
8.2. Графічні методи візуалізації даних 108
Приклади розв'язку типових задач.....114
Перелік питань і завдань для самоконтролю 117
Тести для самоконтролю.....118

ГЛОСАРІЙ.....121

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 127

ДОДАТКИ 131

ПЕРЕДМОВА

Входження в бізнес і його розвиток здійснюється в певному психологічному та соціокультурному середовищі. На ефективність бізнесу впливає ряд чинників. Особливої актуальності набувають уміння грамотно опрацьовувати й аналізувати дані психологічних досліджень.

Психолог, який володіє всебічними знаннями про психологічну реальність, має широкі можливості ефективно впливати на неї. Здобути такі знання за допомогою методології та методів лише гуманітарних наук неможливо. Щоб отримати об'єктивну її картину, потрібне вміння застосовувати й методи аналізу даних. У психології методи аналізу даних допомагають зробити процес дослідження явищ більш чітким, структурованим і раціональним. Засвоївши основні положення й методи, фахівець з психології буде спроможним обробляти дані, які отримані в результаті проведення різних психологічних експериментів, анкетування, встановлювати взаємозв'язки між різними чинниками, аналізувати отримані дані, прогнозувати.

Навчальний посібник «Методи аналізу даних у психологічних дослідженнях» сприятиме оволодінню навичками аналізу даних результатів психологічних досліджень.

Посібник складається з 8 тем, які містять детальний теоретичний опис основних питань використання методів аналізу даних у психології з потрібними визначеннями та формулами. Прикладний аспект використання методів аналізу в психологічних дослідженнях відображують приклади розв'язку типових задач. Для закріплення отриманих знань і навичок у навчальному посібнику пропонується перелік питань і завдань, а також тести для самоконтролю. Наведений у кінці посібника глосарій допоможе детальніше розібратись у математичних і статистичних термінах.

Навчальний посібник «Методи аналізу даних у психологічних дослідженнях» призначений для здобувачів вищої освіти психологічних спеціальностей усіх форм навчання, слухачів післядипломної освіти, курсів перепідготовки, працівників підприємницьких

структур, широкого кола фахівців з психологічних досліджень і для самостійного вивчення дисципліни.

Успішне оволодіння матеріалами навчального посібника сприятиме розвитку аналітичного мислення та дозволить використовувати методи аналізу даних під час прийняття ефективних професійних рішень.

ТЕМА 1
ОСНОВНІ ПОНЯТТЯ, ЯКІ ВИКОРИСТОВУЮТЬ ПІД ЧАС АНАЛІЗУ ДАНИХ У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ



1.1. Поняття аналізу даних у психологічних дослідженнях.

1.2. Методи аналізу даних.

1.3. Вимірювання та шкали в психологічних дослідженнях.

1.1. Поняття аналізу даних у психологічних дослідженнях

У психологічних дослідженнях широко використовуються якісний і кількісний аналізи. Кількісний аналіз насамперед застосовується для опису даних і визначення значущості результатів.

Кількісний аналіз – це сукупність методів опису та процедур перетворення даних дослідження за допомогою математичних і статистичних інструментів.

Кількісні методи інтерпретації та аналізу даних містять у собі:

– *дисперсійний аналіз*, що визначає зміну дисперсії (розсіяння) залежно від змінно співвідношення зі змінами незалежної змінної, визначає змінну, яка викликає зміни залежної змінної (f-критерій Фішера, критерій Стьюдента);

– *кореляційний аналіз*, який встановлює взаємозв’язок і направленість змін залежної змінної та незалежної змінної;

– *факторний аналіз*, що призначений для визначення впливу фактора (набору корельованих незалежних змінних) на залежну змінну. Фактор можна сформувати за допомогою лінійного рівняння регресії;

– *регресійний аналіз* дозволяє пов'язати кілька залежних змінних з одним фактором, який відображає вплив незалежних змінних на об'єкт;

– *кластерний аналіз*, який ґрунтується на «схожості» змінних (характеристик), якими володіють різні суб'єкти, для визначення взаємозв'язку та ступеня подібності їхніх поведінкових реакцій, і об'єднує різних «схожих» суб'єктів (досліджуваних) в один клас (групу, кластер тощо) відповідно до певних статистичних критеріїв.

Якісний аналіз у психологічних дослідженнях використовується для опису та характеристики зв'язків між досліджуваними явищами.

Якісний аналіз результатів – це множина процедур і методів опису даних дослідження, що ґрунтуються на теоретичних висновках і узагальненнях, особистому досвіді, інтуїції та методах логічного мислення.

Типами якісного дослідження результатів у психологічних дослідженнях є:

– *систематизація та диференціація дослідницького матеріалу* за типами, видами та варіантами, розробка схем і структур;

– *психологічна казуїстика* – систематичний опис типових та особливих випадків.

Інтерпретація (тлумачення, пояснення) результатів здійснюється на основі теоретичних моделей і систематизації результатів якісного та кількісного аналізу дослідницького матеріалу.

1.2. Методи аналізу даних

Методи аналізу даних – це математичні формули, прийоми, кількісні розрахунки, які дозволяють систематизувати отримані в ході дослідження дані, виявити наявні закономірності.

Первинні методи дозволяють встановити показники, які відображають безпосередні результати досліджень.

З їхньою допомогою психолог може сформувати своє перше уявлення про об'єкт: про його характеристики, про наявні закономірності тощо.

Первинні методи:

Середнє арифметичне. Це одне з найпростіших арифметичних дій. Для отримання цього кількісного показника дослідникові достатньо скласти всі виявлені числові значення й поділити отриману суму на кількість цих значень.

Медіана. Під медіаною розуміється числова величина, яка займає центральне положення в послідовному ряду даних. Інакше кажучи, з усього масиву чисел половина виявляється менше медіани, а половина – більше. Відповідно, вище й нижче медіани кількість значень однакова.

Мода. Цей метод передбачає виділення числа, яке частіше за інших присутнє у вибірці – саме «модне» число.

Наприклад, якщо більшість випробовуваних демонструють під час експерименту одну й ту саму реакцію на будь-який подразник, то кількісний показник за цією реакцією буде ставитися до моди.

Дисперсія – це міра мінливості, яка дозволяє судити про ступінь варіації ознаки.

Вторинні методи спрямовані на більш глибоке вивчення питання.

Вони допомагають виявляти приховані закономірності, встановлювати взаємозв'язки.

Вторинні методи: кореляційний аналіз, регресійний аналіз, факторний аналіз та ін.

Кореляційний аналіз – це комплекс процедур статистичного дослідження взаємозалежності змінних, що перебувають у кореляційних відношеннях: водночас переважає нелінійна їхня залежність, тобто значенню будь-якої окремо взятої змінної може відповідати деяка кількість значень змінної іншого ряду, що відхиляються від середнього в певний бік.

Кореляційний аналіз є одним із допоміжних методів розв'язання теоретичних проблем психодіагностики і містить у собі набір статистичних процедур, які широко використовуються для розробки тестів та інших психодіагностичних методик, визначення їхньої надійності та валідності.

У прикладних психологічних дослідженнях кореляційний аналіз є основним методом статистичної обробки кількісного емпіричного матеріалу.

Регресійний аналіз у психології – це метод аналізу, який дозволяє дослідити, чи залежить середнє значення певної змінної від варіації іншої змінної або декількох змінних (у цьому випадку використовується множинний регресійний аналіз).

Поняття регресійного аналізу ввів Ф. Гальтоп, який встановив певну кореляцію між зростом батьків та їхніх дорослих дітей. Науковець помітив, що низькорослі батьки мають дещо вищих дітей, а високі батьки – нижчих дітей. Він назвав цю закономірність регресією.

Регресійний аналіз використовується переважно в емпіричних психологічних дослідженнях для розробки психологічних тестів для вирішення завдань, пов'язаних з оцінкою певного впливу (наприклад, вплив інтелектуальних здібностей на академічну успішність, мотивацію поведінки тощо).

Факторний аналіз – це метод багатовимірної математичної статистики, який використовується в процесі дослідження статистично значущих характеристик і дозволяє виявити кілька чинників, які не видно під час безпосереднього спостереження. Факторний аналіз не тільки встановлює взаємозв'язки між досліджуваними змінними, що перебувають у стані трансформації, а й визначає ступінь цього зв'язку та виявляє основні чинники, які лежать в основі цих трансформацій.

Факторний аналіз особливо ефективний на ранніх стадіях дослідження, коли потрібно виявити попередні закономірності в досліджуваній галузі. Це дає змогу уточнити подальші експерименти порівняно з експериментами, заснованими на випадково або довільно обраних змінних.

Загалом математичні методи можуть бути досить ефективними та корисними в організації й проведенні психологічних досліджень, проте варто пам'ятати, що математичний метод, як і будь-який інший, має свою сферу прикладання та певні дослідницькі можливості. Застосування методу зумовлене природою предмета дослідження та завданнями пізнавальних дій дослідника.

1.3. Вимірювання та шкали в психологічних дослідженнях

У своїй роботі психологи часто стикаються з проблемою вимірювання психологічних характеристик людей, як-от креативність, нервовість, імпульсивність та особливості нервової системи. Для цього психодіагностика розробляє спеціальні вимірювальні процедури, зокрема тести. Крім того, психологія широко використовує експериментальні методи та моделі для вивчення психічних явищ у когнітивній та особистісній сфері. Це можуть бути моделі когнітивних процесів (сприйняття, пам'яті, мислення) або таких специфічних явищ, як мотивація, ціннісні орієнтації та особистість. Важливо, що під час експерименту характеристики досліджуваного можуть бути кількісно виміряні. Кількісні дані, отримані в результаті ретельно спланованих експериментів із застосуванням спеціальних процедур вимірювання, використовуються для математичної та статистичної обробки.

Вимірювання визначається як присвоєння числового значення об'єкту або події. Воно здійснюється за певними правилами. Ці правила повинні встановлювати відповідність між деякими властивостями досліджуваного об'єкта, з одного боку, й набором числових значень – з іншого.

Загалом ***вимірювання*** – це процедура, за допомогою якої суб'єкт, що досліджується, порівнюється з певним еталоном і виражається чисельно в певній шкалі або масштабі.

Розрізняють декілька типів шкал. Тип шкали визначає групу методів, які використовують для обробки вимірювання даних. Для вивчення різних особливостей соціально-психологічних явищ у психології використовують різні шкали.

Виділяється чотири шкали вимірювання:

1) шкала номінальна (номінативна, найменувань, класифікаційна);

2) шкала порядку (рангова, ординарна);

3) шкала інтервалів (рівних інтервалів);

4) шкала відношень (пропорційна).

Ряд фахівців вирізняють шкалу абсолютних значень і шкалу різниць.

Виділяють *три основних атрибути вимірювальних шкал*, наявність або відсутність яких визначає належність шкали до певної категорії:

- впорядкованість даних означає, що один пункт шкали більший, менший або дорівнює іншому пункту;
- інтервальність пунктів шкали означає, що інтервал між будь-якою парою чисел більший, менший або дорівнює інтервалу між іншою парою чисел;
- нульова точка означає, що набір чисел має позначку відліку, яка позначається нулем і відповідає повній відсутності властивості.

Крім того, виділяють такі *групи шкал*:

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

– метричні (шкала інтервалів, шкала відносин).

Шкала номінальна. Шкала номінальна (синоніми: *шкала найменувань, номінативна шкала*) створюється, присвоюючи об’єкту «ім’я».

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

Це й становить сутність номінальної шкали. Є найпростішою вимірювальною шкалою, хоча фактично вона не асоціюється з вимірюванням і не пов’язана з поняттям «величина». Вона використовується, тільки щоб відрізнити один об’єкт від іншого. Тут відсутні всі основні атрибути вимірювальних шкал, а саме упорядкованість, інтервальні, нульова точка.

Вимірювання заключається в тому, що після вивчення об’єкта визначають належність результату до певного стану (якості) й записують це за допомогою символу (набором символів, назвою, словом, числом), що позначає цей стан.

Незважаючи на тенденцію переоцінювати силу шкал, психологи часто використовують номінальні шкали в своїх дослідженнях. Вимірювальні процедури в оцінці особистості часто призводять до створення типологій, які відносять конкретні особистості до певних типів. Приклади цієї типології охоплюють такі темпераменти: холерик, сангвінік, меланхолік і флегматик.

Найпростіша номінальна шкала відома як дихотомічна шкала. Під час вимірювання за дихотомічною шкалою ознаки можуть кодуватися двома літерами або числами, наприклад 0 і 1, 2 і 6, літерами А і В або будь-якими двома різними літерами. Ознаки, виміряні за дихотомічною шкалою, називаються альтернативними ознаками.

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

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

За даними, зафіксованими у номінальній шкалі, можна здійснювати лише операцію перевірки їхньої тотожності або відмінності.

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

Шкала порядку. Шкала порядку використовується тоді, коли є можливість встановити порядок психологічних об’єктів відповідно до тяжкості їхнього характеру.

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

Ця шкала може бути використана для лінійного впорядкування об’єктів на певних осях ознак (відношення «більше» або «менше»). Тим самим є можливість вивчити «лінійну» вираженість властивості тоді, як шкалою найменувань використовується «категоріальна» властивість: властивість є – властивості немає.

Конструювання шкали порядку – більш складна процедура, ніж створення іменної шкали. Вона дозволяє зафіксовувати ранг (розташування) кожного значення змінної відносно інших значень. Одиницею вимірювання є відстань в один клас, ранг або бал. Водночас відстань між рангами (балами) може бути різною (вона нам невідома).

У порядковій шкалі має бути не менше трьох градацій (класів). Чим більше число градацій, на які поділяється сукупність, що досліджується, тим більші можливості для статистичної обробки даних і перевірки статистичних гіпотез. Порядкові змінні можна кодувати будь-якими цифрами, загалом щоб зберігався порядок, тобто кожна наступна цифра повинна бути більшою (або меншою) попередньої.

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

Порядкова шкала досить часто використовується в психологічних дослідженнях. Ще сам засновник теорії вимірювання С. Стівенс вважав, що більшість психологічних вимірювань відповідає шкалам порядку.

Шкала інтервалів. Інтервальна шкала – це перша метрична шкала. Краще починати з цієї шкали та свідчити про вимірювання у вужчому сенсі.

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

Інтервальні шкали можна використовувати для порівняння двох об’єктів і визначення величини їхніх відмінностей. Це спосіб з’ясувати, наскільки більше або менше властивість виражена в одному об’єкті, ніж в іншому.

У цій шкалі тільки інтервали відображають значення справжніх чисел і лише над інтервалами варто виконувати арифметичні операції. Якщо забути про це, то є ризик отримати безглузді результати. Наприклад, не можна наголошувати, що температура води зросла у два рази під час нагрівання від 9 до 18 градусів за шкалою Цельсія, оскільки для того, хто звик користуватися шкалою Фаренгейта, це буде досить дивно, оскільки в цій шкалі температура води в тому самому експерименті зміниться від 37 до 42 градусів.

З інтервальними шкалами для аналізу даних можна використовувати майже будь-яку параметричну статистику. Для характеристики центральної тенденції, крім медіани й моди, можна розраховувати середнє арифметичне, а для оцінки варіації – дисперсію, стандартне відхилення. Можна визначати коефіцієнти асиметрії та ексцесу й інші параметри розподілу.

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

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

Особливість цієї шкали полягає в наявності абсолютного (істинного) нуля. У шкалі відносин можна стверджувати, в скільки разів властивість одного об’єкта перевершує таку саму властивість у іншого.

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

Перелік питань для самоконтролю



1. Що вивчає дисципліна «Методи аналізу даних у психологічних дослідженнях»?
2. Опишіть предмет дисципліни.
3. Назвіть, які методи використовують для аналізу даних.
4. Які існують типи даних у психологічних дослідженнях?
5. Які особливості використання методів аналізу даних у психології?
6. Назвіть шкали вимірювання в психології.

Тести для самоконтролю

1. Чим … тип шкали, тим більше інформації можна одержати про властивості об'єкта вимірювання:

а) нижчий;
б) досконаліший;
в) вищий;
г) якісніший.
Відповідь: 1) а, б; 2) а; 3) б; 4) –.

2. Шкали найменувань і порядку є:

а) розрахунковими шкалами;
б) числовими шкалами;
в) кількісними шкалами;
г) якісними шкалами.
Відповідь: 1) а, б; 2) а; 3) б; 4) –.

3. Шкали інтервалів і відношень є:

а) метричними шкалами;
б) категоріальними шкалами;
в) якісними шкалами;
г) атрибутивними шкалами.
Відповідь: 1) а, б; 2) а; 3) б; 4) –.

4. Чим складніший рівень психіки, тим … його можна виміряти:

а) менш точно;

б) більш точно;

в) точно;

г) неточно.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

5. Відповідно до типів шкал використовують … міри центральної тенденції:

а) найпростіші;

б) будь-які;

в) однакові;

г) різні.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

6. Для шкал певного рівня можна використовувати статистичні міри шкал … рівнів, але не навпаки:

а) сусідніх;

б) усіх попередніх;

в) усіх наступних;

г) будь-яких.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

7. Шкала найменувань – це шкала, яка класифікує:

а) за кількісною ознакою;

б) за назвою;

в) за величиною;

г) за порядком.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

8. Якщо використовується шкала найменувань, то класам об’єктів для їхнього позначення можна приписувати:

а) будь-які символи: словесні позначки, числа, букви чи інші знаки;

б) тільки числа;

в) тільки букви;

г) тільки словесні позначки.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

9. У порядковій шкалі класам можна приписувати:

а) точку абсолютного нуля;

б) дисперсію;

в) середні арифметичні значення;

г) словесні або числові позначення.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

10. Якщо кількість рангів рівна кількості впорядкованих об’єктів, то таке ранжування називається:

а) ранговим;

б) примусовим рангуванням;

в) якісним;

г) кількісним.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

ТЕМА 2
ВИБІРКОВИЙ МЕТОД У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ



2.1. Поняття вибіркового методу.
2.2. Види вибірки та способи відбору.
2.3. Обчислення похибок вибірки та визначення потрібного обсягу вибірки.

2.1. Поняття вибіркового методу

Дослідник-теоретик, розмірковуючи про певні закономірності психіки та поведінки зазвичай міркує не про конкретний об'єкт свого дослідження, а має на увазі насамперед якусь множину об'єктів. Наприклад, досліджуючи пам'ять людини, маємо на увазі не когось конкретно, а людину взагалі, усіх людей, які живуть на Землі.

Очевидно, в цьому прикладі йдеться про множину об'єктів, яка до того ж немає чітко визначених меж. Таку множину об'єктів у аналізі даних прийнято називати ***генеральною сукупністю*** або популяцією (N). Зазвичай можемо конкретизувати наші уявлення про досліджувану генеральну сукупність, зробити її більш компактною. Наприклад, свідчимо не про пам'ять людини взагалі, а про пам'ять якоїсь більш вузької групи людей – про пам'ять дитини певного віку або людей, які страждають амнезією. Але навіть у цьому разі генеральна сукупність виявиться досить розмитою сукупністю, що охоплює значну кількість об'єктів, над поведінкою яких, власне, й розмірковує теоретик, намагаючись зрозуміти закони поведінки цих теоретичних об'єктів, передбачити цю поведінку.

Проводячи експериментальне дослідження, дослідник-експериментатор прагне перевірити передбачення дослідника-теоретика. Водночас він не може провести дослідження відразу з усіма об'єктами, які становлять усю генеральну сукупність. Адже генеральна сукупність надто велика, до того ж її межі зазвичай невизначені. Так, ідея об'їхати весь світ, досліджуючи пам'ять усіх

людей, які страждають певними формами амнезії, експериментатору навряд чи буде цікавою й захоплюючою. Навіть якщо йому і вдасться це зробити, через якийсь час, на жаль, з'являться нові люди, які страждають тією самою недугою, і за такого підходу доведеться знову перевіряти припущення теоретика.

Що ж свідчити про ситуацію, коли теоретик міркує про людину взагалі. На щастя, у цьому немає жодної потреби. Як варто наголошувати в такому разі, щоб дізнатися суп на смак, не обов'язково з'їдати всю каструлю. Досить однієї ложки. Але попередньо суп потрібно добре розмішати. Саме тому дослідник-експериментатор має справу не з усією генеральною сукупністю, а лише з невеликою її частиною, так званою ***вибірковою сукупністю*** або вибіркою.

Вибірка (n) – це обмежена за кількістю група одиниць (досліджуваних, респондентів), спеціально відібрана із загальної сукупності для дослідження її властивостей.

2.2. Види вибірки та способи відбору

Вибіркова сукупність повинна бути максимально подібною до самої генеральної сукупності. Ця характеристика відображається в понятті репрезентативності.

Репрезентативність вибірки – це властивість вибірки репрезентувати досліджувані явища з погляду їхньої мінливості в генеральній сукупності.

Наведемо два методи забезпечення репрезентативності вибірки.

Перший метод полягає у формуванні простої випадкової вибірки (рандомізація). Він передбачає, що вибірка складатиметься з елементів, які відібрані з генеральної сукупності так, що кожен елемент генеральної сукупності має рівні шанси (ймовірність) потрапити до вибірки. Отримана вибірка називається простою випадковою вибіркою.

Другий спосіб – формування стратифікованої випадкової вибірки. Для цього елементи генеральної сукупності потрібно розділити на страти (групи) за певними ознаками.

Стратифікована вибірка використовується, коли вибірка повинна охоплювати респондентів з певними характеристиками (наприклад, стать, вік, рівень освіти). Вибірки є *залежні* або *незалежні*. Якщо можна встановити гомоморфну пару (тобто коли одному випадку з вибірки X відповідає один і тільки один випадок із вибірки Y і навпаки) для кожного випадку в двох вибірках, такі вибірки називаються ***залежними*** (пов'язаними, парними). Приклади залежних вибірок: пари близнюків, чоловіки та їх дружини, група досліджуваних до проведення психологічного експерименту й після тощо.

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

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

За обсягом вибірки ділять на *малі й великі*. До ***малих*** відносять вибірки, в яких число елементів $n \leq 30$. Поняття великої вибірки не визначено, але ***великою*** вважається вибірка, в якій число елементів > 200 і ***середня*** вибірка задовольняє умові $30 \leq n \leq 200$. Цей поділ є умовним. Малі вибірки використовують під час статистичного контролю відомих властивостей уже вивчених сукупностей. Великі вибірки використовують для установки невідомих властивостей і параметрів сукупності.

Вибірка елементів для вибіркового спостереження може здійснюватися способом *повторного й безповторного* відбору.

Повторним відбором називають такий відбір, у разі якого кожна обстежувана одиниця знову повертається до генеральної сукупності, продовжує брати участь у подальшому відборі й може потрапити повторно у вибірку для обстеження.

Безповторним називається такий відбір, під час якого один раз описані одиниці спостереження в подальшому відборі участі не беруть. Безповторний відбір зазвичай дає більш точні результати, ніж повторний.

2.3. Обчислення похибок вибірки та визначення потрібного обсягу вибірки

Особливістю вибіркової сукупності є оцінки відповідних параметрів генеральної сукупності. Однак вибірка не є точно репрезентативною для генеральної сукупності, тому оцінки не відповідають самим параметрам.

Різниця між ними називається ***похибкою репрезентативності***. Ця похибка може бути як систематичною, так і випадковою.

Систематичні помилки виникають, коли під час формування вибіркової сукупності не дотримуються принципу випадковості відбору, який забезпечує рівні можливості для всіх елементів генеральної сукупності потрапити до вибірки. Систематичні помилки для всіх елементів генеральної сукупності називаються *помилками зміщення*, оскільки вони є односпрямованими.

На відміну від них помилки, які неминуче виникають навіть за умови дотримання принципу випадкового відбору та є незміщеними, називаються ***випадковими*** помилками й не зміщують оцінки.

Розмір випадкової помилки вибірки визначається згідно із граничними теоремами ймовірностей. Розрізняють *середню* та *граничну* похибки вибірки.

Під ***середньою (стандартною) помилкою*** вибірки розуміють таке розходження між вибірковою та генеральною сукупністю, яке не перевищує розмір середнього квадратичного відхилення. Максимально можливе розходження між вибірковою та генеральною сукупністю називають ***граничною помилкою*** вибірки, тобто це максимум помилки в разі заданої ймовірності її появи.

Існують дві формули середньої помилки вибірки. Одна з них використовується під час вимірювання середнього значення ознаки, друга – коли вибірково вимірюється частка ознаки.

Помилка вибірки обчислюється за формулами

Повторна вибірка	Безповторна вибірка
Для середньої: $\mu_x = \sqrt{\frac{\sigma^2}{n}}$ (2.1)	$\mu_x = \sqrt{\frac{\sigma^2}{n}(1-\frac{n}{N})}$. (2.2)
Для частки: $\mu_p = \sqrt{\frac{pq}{n}}$ (2.3)	$\mu_p = \sqrt{\frac{pq}{n}(1-\frac{n}{N})}$, (2.4)

де μ – середня помилка вибірки;

σ^2 – дисперсія;

n – число одиниць вибіркової сукупності;

N – число одиниць генеральної сукупності;

W – частка одиниць з однаковою ознакою в загальній чисельності одиниць вибірки.

Гранична помилка вибірки обчислюється за формулою:

$$\Delta = t\mu, \qquad (2.5)$$

де t – коефіцієнт кратності помилки;

μ – середня помилка вибірки.

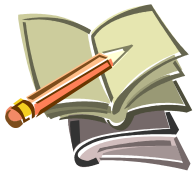
Значення t за ймовірністю: 0,683 дорівнює 1;
0,954 дорівнює 2;
0,997 дорівнює 3;
0,999 дорівнює 4.

Вищезазначені формули граничних помилок вибірки після відповідних перетворень можна використати для обчислення потрібної чисельності вибірки (n).

Формули для обчислення вибіркової сукупності:

	Повторна вибірка	Безповторна вибірка
Для середньої	$n = \frac{t^2\sigma^2}{\Delta^2}$ (2.6)	$n = \frac{t^2\sigma^2 N}{\Delta^2 N + t^2\sigma^2}$. (2.7)
Для частки	$n = \frac{t^2 W(100-W)}{\Delta^2}$ (2.8)	$n = \frac{t^2 W(100-W)N}{\Delta^2 N + t^2 W(100-W)}$. (2.9)

Приклади розв’язку типових задач



1. Серед студентства закладу вищої освіти 30 % студентів і 70 % студенток. Оцінити ймовірність отримання нерепрезентативної вибірки, що складається лише зі студентів.

Розв’язання:

Ймовірність вибрати першого студента становить 0,3, другого – 0,3 · 0,3 = 0,09, третього – 0,09 · 0,3 = 0,027 і так до кінця. Ймовірність отримати вибірку обсягом 10 з одних студентів становить $0{,}3^{10} = 0{,}095 < 0{,}15 = 0{,}00001$ і є вже доволі малою. Водночас ця ймовірність швидко зменшується зі збільшенням обсягу вибірки.

2. Генеральна сукупність складається з чотирьох осіб різного зросту.

Таблиця 2.1 – Показники зросту досліджуваних

Номер особи в списку	1	2	3	4
Зріст особи, см	172	176	180	180

Визначити середній зріст і стандартне відхилення зросту досліджуваних.

Розв’язання:

Припустимо, що для оцінки середнього значення зросту цієї сукупності формують випадкову вибірку обсягом 2 і визначають її середнє значення. Спробуємо оцінити рівень потенційної помилки такого дослідження.

Для цього побудуємо двовимірний розподіл вибіркового середнього. Спочатку сформуємо всі прості випадкові вибірки обсягом 21 і визначимо середнє значення зросту в кожній з них.

Таблиця 2.2 – Двовимірний розподіл
вибіркового середнього значення зросту

Номери в списку	Значення зросту у вибірці, см	Вибіркове середнє, см
1; 1	172; 172	172
1; 2	172; 174	173
1; 3	172; 180	176
1; 4	172; 180	176
2; 1	174; 172	173
2; 2	174; 174	174
2; 3	174; 180	177
2; 4	174; 180	177
3; 1	180; 172	176
3; 2	180; 174	177
3; 3	180; 180	180
3; 4	180; 180	180
4; 1	180; 172	176
4; 2	180; 174	177
4; 3	180; 180	180
4; 4	180; 180	180

Знайшовши частоти отриманих вибіркових середніх, утворимо двовимірний розподіл вибіркового середнього.

Таблиця 2.3 – Двовимірний розподіл вибіркового середнього

Вибіркове середнє	Частота
172	1
173	2
174	1
176	4
177	4
180	4

Середнє значення цього розподілу:

$$\boldsymbol{\mu_x} = \frac{172 \cdot 1 + 173 \cdot 2 + 174 \cdot 1 + 176 \cdot 4 + 177 \cdot 4 + 180 \cdot 4}{16} = 176{,}5.$$

Середнє значення вихідної генеральної сукупності:

$$\boldsymbol{\mu_x} = \frac{172 + 176 + 180 + 180}{4} = 176{,}5.$$

Отже, середнє значення розподілу вибіркового середнього збігається із середнім значенням генеральної сукупності. Водночас середнє значення жодної з вибірок не збігається з генеральним середнім. Отже, будь-який фахівець, який прагнув би оцінити середній зріст генеральної сукупності за вибіркою з двох осіб, не мав би жодних шансів отримати точне значення генерального середнього.

Кожна вибірка продукує помилку для того, щоб врахувати рівень цих помилок, природно визначити стандартне відхилення розподілу вибіркового середнього, інакше кажучи, стандартну помилку. Опускаючи обчислення, одразу наведемо її значення: $\sigma X \approx 2{,}52$.

Водночас стандартне відхилення генеральної сукупності: $\sigma \approx 3{,}57$.

Чи можна знайти якийсь змістовний зв'язок між значеннями σX та σ? Виявляється, можна: $3{,}57 \approx 2{,}52\ \sqrt{2}$, де 2 – обсяг вибірки.

Отже, наведені закономірності невипадкові.

Перелік питань і завдань для самоконтролю



1. Які існують переваги та недоліки вибіркового методу?
2. Яка відмінність між генеральною сукупністю та вибірковою?
3. Назвіть види формування вибірки.
4. Опишіть алгоритм розрахунку мінімального обсягу вибірки.
5. Проводимо соціологічне опитування киян. Обираємо випадкових перехожих на Хрещатику під час обідньої перерви. Чи можна вважати цю вибірку простою випадковою? Чи це репрезентативна вибірка? Поясніть свою відповідь.
6. Для тестування нового методу лікування випадково відбирають 25 пацієнтів із загальної популяції добровольців, які являють собою усіх потенційних пацієнтів. Це проста випадкова вибірка? Чи можна вважати її репрезентативною? Обґрунтуйте свою відповідь.

7. Частка жінок серед учителів середніх шкіл міста становить 80 %. Маючи цю інформацію, як би ви відібрали репрезентативну вибірку із 50 осіб з цієї сукупності?

8. Розглянемо дві генеральні сукупності: одну з чотирьох осіб, іншу – з трьох. Припустимо, IQ-індекси осіб у цих сукупностях такі: 113, 120, 121, 124 у першій та 100, 104, 129 у другій. Розгляньте (2, 1) – вимірний вибірковий розподіл різниці середніх.

Тести для самоконтролю

1. Поняття, яке відображає рівень відповідності структури вибірки структурі генеральної сукупності, – це:

а) адекватність;

б) релевантність;

в) значимість;

г) репрезентативність.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

2. Показник, який ділить упорядкований ряд даних на дві пропорційні за кількістю значень частини, називається:

а) стандартним відхиленням;

б) модою;

в) дисперсією;

г) медіаною.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

3. Статистика висновку спрямована на:

а) визначення взаємозв’язків між досліджуваними явищами в генеральній сукупності;

б) пошук зв’язків у досліджуваній сукупності;

в) пошук форми і напрямку зв’язку у генеральній сукупності;

г) розкриття властивостей генеральної сукупності на основі вибіркових даних.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

4. Показник, який у ряді даних зустрічається найчастіше, – це:

а) стандартне відхилення;

б) мода;

в) дисперсія;

г) медіана.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

5. Показник, який належить до характеристики центральної тенденції, – це:

а) стандартне відхилення;

б) розмах варіації;

в) дисперсія;

г) середнє арифметичне.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

6. Розділом описової статистики є:

а) перевірка статистичних гіпотез;

б) статистика випадкової вибірки;

в) статистика висновку;

г) статистичне оцінювання.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

7. Репрезентативність вибірки – це:

а) поняття, яке відображає рівень відповідності обсягу вибірки обсягу генеральної сукупності;

б) поняття, яке відображає рівень відповідності методу дослідження вибірки методу дослідження генеральної сукупності;

в) поняття, яке відображає рівень відповідності структури вибірки структурі генеральної сукупності;

г) поняття, яке відображає рівень відповідності результатів дослідження вибірки результатам дослідження генеральної сукупності.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

8. Чим більша вибіркова сукупність, тим вищою буде:
а) прогностична значимість наукових результатів;
б) статистична значимість наукових результатів;
в) наукова значимість результатів;
г) точність наукових результатів.
Відповідь: 1) а, б; 2) а; 3) б; 4) –.

ТЕМА 3
МЕТОДИ ОПИСОВОЇ СТАТИСТИКИ В ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ



3.1. Міри центральної тенденції.
3.2. Міри мінливості даних.
3.3. Міри положення.

3.1. Міри центральної тенденції

Методи обробки кількісних даних поділяють на *первинні* та *вторинні*.

Метою первинної статистичної обробки є впорядкування інформації про предмет та об'єкт дослідження. Водночас дані групуються за певними критеріями, які заносяться в зведені таблиці. Оброблена інформація дає змогу отримати уяву про характер усього набору даних: їх однорідність чи неоднорідність, компактність чи розкиданість, чіткість чи розмитість тощо. Така інформація відображується у візуальному представленні даних і надає інформацію щодо їх розподілу.

О*сновними методами первинної статистичної обробки* є визначення показників центральної тенденції, мінливості (розкиду) даних і квантилі розподілу.

Первинна статистична обробка даних дозволяє охарактеризувати сукупність у стислому вигляді та визначити:

1) найбільш характерне значення для вибірки;

2) варіативність даних.

Для визначення найбільш характерного значення для вибірки, визначають міри центральної тенденції, для визначення варіативності даних – розраховують показники мінливості (або розкиду).

Числові показники типових характеристик емпіричних даних називаються ***мірами центральної тенденції***. Такими показниками можуть бути, наприклад, середній рівень інтелекту студента університету. До таких показників-мір належать: середнє арифметичне, медіана та мода.

Мода (Mo) – це значення, що зустрічається найчастіше у вибірці, тобто число, яке має найбільшу частоту. Якщо всі значення в групі мають однакову частоту, мода не передбачається. Якщо два сусідні значення з однаковою частотою й зустрічаються частіше за інші значення, то найчастішим значенням вважається середнє з цих двох значень. Якщо однакову частоту мають два несуміжні значення, то сукупність є бімодальною.

Середнє арифметичне $\bar{X}$ – це частка суми значень сукупності (X) на їх кількість (N):

$$\bar{X} = \frac{\sum x_i}{N} \tag{3.1}$$

Медіана (Me) – це показник, який міститься посередині впорядкованих даних. Медіана може не збігатися з конкретним показником. Збіг відбувається для непарної кількості значень (відповідей), а розбіжність – для парної кількості. У цьому випадку медіана обчислюватиметься як середнє арифметичне двох центральних значень. Водночас сукупність має бути впорядкованою.

Міри центральної тенденції мають свої особливості:

- моду вибірки можна визначити «на око». Для великих сукупностей мода є характеристикою центру розподілу;
- медіана перебуває в проміжному положенні між середнім арифметичним та модою. Її особливо легко визначити в ранжованих даних;
- для визначення середнього арифметичного використовують всі значення вибірки, водночас усі вони мають вплив на значення цієї міри.

Визначення таких характеристик центральної тенденції здійснюють і для дослідження розподілу даних. За нормального розподілу вони мають бути однаковими або досить близькими.

3.2. Міри мінливості даних

Міри мінливості даних – це статистичні показники, які дозволяють визначити ступінь однорідності отриманого набору даних і його компактність.

Найпоширенішими показниками мінливості даних у психологічних дослідженнях є ексцес, дисперсія та середньоквадратичне відхилення.

Розмах (R) – це інтервал між максимальним і мінімальним значеннями ознаки.

Розмах (R) – це інтервал між найбільшим і найменшим значеннями ознаки. Розмах визначається за крайніми значеннями ознаки:

$$R = X_{max} - X_{min} \qquad (3.2)$$

Дисперсія (σ^2) – це середнє арифметичне квадратів відхилень від середнього арифметичного індивідуальних значень ознаки.

Дисперсія має розмірність, що дорівнює квадрату ознаки, і обчислюється залежно від того, чи дані згруповані (зважена) чи ні (проста):

Проста

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n} \qquad (3.3)$$

Зважена

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} \qquad (3.4)$$

Стандартне (середньоквадратичне) відхилення (σ) – це середнє відхилення кожного значення ознаки від середнього арифметичного. Має таку саму розмірність, що й сама ознака.

Щоб знайти його, потрібно обчислити корінь квадратний з дисперсії. Його додатне значення приймається за міру мінливості, яка називається середньоквадратичним або стандартним відхиленням:

$$\sigma = \sqrt{\sigma^2}. \qquad (3.5)$$

Коефіцієнт варіації (v) – відсоткове відношення стандартного відхилення до середньої арифметичної величини ознаки:

$$v = \frac{\sigma * 100}{\bar{x}}. \qquad (3.6)$$

Застосовується для порівняння мінливості розподілів ознак, що мають різну розмірність. За своєю сутністю коефіцієнт варіації призводить до однакового масштабу величину стандартного відхилення.

Коефіцієнт варіації дозволяє порівнювати мінливість ознак, які представлені в різних шкалах, а також оцінити однорідність вибірки. *Орієнтовними критеріями оцінки варіативності за коефіцієнтом* можна вважати:

- низький рівень – до 10 %;
- середній рівень – 10–20 %,
- високий рівень – вище 20 %.

Асиметрія (А) характеризує ступінь асиметрії розподілу щодо його середнього значення:

$$A = \frac{\bar{X} - Мо}{\sigma} \quad (3.7) \qquad \text{або} \qquad A = \frac{\bar{X} - Ме}{\sigma} \quad (3.8)$$

Позитивна асиметрія вказує на те, що верхня частина розподілу зміщена в бік негативного значення, а негативна асиметрія вказує на те, що вона зміщена в бік позитивного значення.

Якщо A| < 0,25 асиметрія слабка, в разі 0,25 < |A| < 0,5 – середня, в разі |A| > 0.5 – сильна.

Ексцес (Е) характеризує випуклість або згладженість розподілу вибірки порівняно з нормальним розподілом:

$$E = \frac{m_4}{\sigma^4} - 3, \quad (3.9)$$

де $m_4 = i^4 \frac{\sum\left(\frac{x-\bar{x}}{i}\right)^4 f}{\sum f}$ – центральний момент 4-го порядку, i – ширина інтервалу.

E > 0 для гостровершинного розподілу, E < 0 для плосковершинного розподілу і E = 0 для нормального розподілу.

3.3. Міри положення

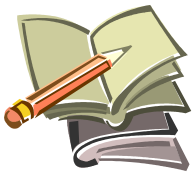
У психології досить широко використовують таку ***міру положення***, як квантиль розподілу.

Квантиль – це значення виміряної ознаки на числовій осі, яке ділить усю сукупність ранжованих показників з відомим співвідношенням їхньої чисельності на дві групи. До квантилів належить медіана, тобто ознака, яка поділяє досліджувану сукупність на дві рівні групи. Також часто використовують процентилі та квартилі.

Процентилі – значення ознаки (P_1 ... P_{99}), які поділяють упорядковану (за зростанням) множину даних на 100, рівних за чисельністю частин. Визначення значення конкретного процентиля подібно до визначення медіани. Наприклад, щоб визначити 10-й процентиль P_10, усі значення ознаки спочатку впорядковують у порядку зростання. Потім підраховують 10 % випробовуваних з найменшою виразністю ознаки. P_10 відповідає значенню характеристики, яка відокремлює 10 % суб'єкта від інших 90 %.

Квартилі – це три значення ознаки (P_{25}, P_{50}, P_{75}), які ділять впорядкований (за зростанням) набір даних на 4 рівні частини. Перший квартиль дорівнює двадцять п'ятому процентилю, 2-й – відповідає 50-му процентилю або центру, а 3-й квартиль дорівнює сімдесят п'ятому процентилю. Процентилі та квартилі використовують для виявлення підгруп та окремих суб'єктів, які є найбільш типовими або нетиповими для цієї сукупності, або для визначення частоти виникнення певного значення (або інтервалу) вимірюваної ознаки.

Приклади розв'язку типових задач



1. У дослідженні двох різних вибірок отримано результати:

Таблиця 3.1 – Вибірка № 1

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

Таблиця 3.2 – Вибірка № 2

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

Розв'язання:

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

У 2-й вибірці модою теж є сангвінічний тип темпераменту (він також має найбільшу частоту у варіаційному ряду), проте ця частота перевищує найближчу (за величиною) майже вдвічі, в цьому випадку вже можна зазначати про виявлення основної тенденції представленості типів темпераменту в цій вибірці.

2. Для вибірки визначити моду: 1, 1, 1, 3, 5, 7, 7, 8.

Розв'язання:

Мо = 1, бо 1 найчастіше повторюється, тобто має найбільшу частоту.

Припустимо, що в цій вибірці один досліджуваний був замінений іншим. Значення стали такими: 1, 1, 3, 5, 7, 7, 7, 8.

Мо = 7, бо 7 найчастіше повторюється або має найбільшу частоту.

3. Для вибірки 1, 1, 3, 5, 7 визначити медіану.

Розв'язання:

Ме = 3, оскільки 3 міститься посередині впорядкованого (ранжованого) ряду розподілу.

4. Викладач за допомогою тесту досягнень визначив рівень знань здобувачів вищої освіти за спеціальною бальною шкалою. За результатами тестування, отримано такий варіаційний ряд.

Таблиця 3.3 – Розподіл здобувачів вищої освіти за кількістю балів

Кількість балів	0	1	2	3	4	5	6	7	8
Кількість здобувачів вищої освіти	15	10	3	0	0	0	4	12	5

Визначити $\bar{X}$, Мо, Ме.

Розв'язання:

$$\bar{X} = \frac{0 \cdot 15 + 1 \cdot 10 + 2 \cdot 3 + 3 \cdot 0 + 4 \cdot 0 + 5 \cdot 0 + 6 \cdot 4 + 7 \cdot 12 + 8 \cdot 5}{15 + 10 + 3 + 0 + 0 + 0 + 4 + 12 + 5} = \frac{164}{49} = 3{,}35;$$

Mo = 0; Me = 3,5.

Отже, середній бал, отриманий здобувачами вищої освіти, становить 3,35. Найчастіше здобувачі отримували 0 балів, водночас половина студентів отримала менше 3,5 бала, інша – більше.

5. Розрахувати показники описової статистики тривожності досліджуваних.

Таблиця 3.4 – Показники тривожності досліджуваних

№ з/п	Показник тривожності
1	7
2	9
3	2
4	4
5	5
6	6
7	7
8	6
9	2
10	5

Розв'язання:

Розрахуємо показники описової статистики за допомогою Пакету аналізу Microsoft Excel.

Таблиця 3.5 – Результати описової статистики

Показник тривожності	
1	2
Середнє значення	5,3
Стандартна похибка	0,7
Медіана	5,5
Мода	7
Стандартне відхилення	2,213594
Дисперсія вибірки	4,9
Ексцес	–0,2973

Продовження таблиці 3.5

1	2
Асиметрія	– 0,16595
Діапазон	7
Мінімальний	2
Максимальний	9
Сума	53
Кількість	10
Найбільше (1)	9
Найменша (1)	2
Довірчий рівень (95,0 %)	1,58351

Отже, середній показник тривожності становив 5,3 бала. Найчастіше спостерігалась тривожність у 7 балів. Половина досліджуваних мала показник тривожності менше 5,5, половина – більше. Стандартна похибка становить 0,7, стандартне відхилення – 2,21. Показник тривожності кожного досліджуваного відхиляється від середнього значення на 4,9 бала. Спостерігається слабка асиметрія даних та плосковершинний розподіл.

Перелік питань і завдань для самоконтролю



1. Складові міри центральної тенденції. Поняття моди, середнього арифметичного значення, медіани.

2. Поняття квантиля розподілу. Процентилі та квартилі.

3. Складові міри мінливості. Поняття розмаху, дисперсії, стандартного відхилення. Дисперсія та її властивості. Поняття стандартизації.

4. Під час визначення ступеня виразності певної психічної властивості в експериментальній і контрольній групах бали розподілилися так: експериментальна група – 16, 18, 17, 13, 15, 16, 17, 17, 22, 21, 18, 13, 12, 13, 17, 20, 20; контрольна – 27, 9, 12, 13, 26, 23, 14, 15, 22, 21, 16, 16, 18, 17, 10, 11, 16.

Дати порівняльну характеристику ступеня виразності такої властивості в цих групах.

5. Розрахувати показники описової статистики тривожності досліджуваних.

Таблиця 3.6 – Показники тривожності досліджуваних

№ з/п	Показник тривожності
1	9
2	5
3	3
4	4
5	7
6	4
7	8
8	9
9	2
10	5

Тести для самоконтролю

1. На підставі співвідношення між значеннями характеристик центру розподілу (середньої, моди та медіани) можна виміряти:

а) асиметрію розподілу;

б) ексцес розподілу.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

2. Коефіцієнт асиметрії розподілу областей України за часткою україномовного населення становить 0,7. Це означає, що:

а) регіональний розподіл України за часткою україномовного населення зміщений праворуч;

б) регіональний розподіл України за часткою україномовного населення зміщений ліворуч;

в) частка україномовного населення є високою в більшості регіонів і низькою в деяких регіонах;

г) низька частка україномовного населення в більшості регіонів і висока – лише в деяких регіонах;

Відповідь: 1) а, в; 2) б, в; 3) а, г; 4) б, г.

3. Інтервальні варіаційні ряди зображують за допомогою графіків:

а) полігон;

б) гістограма;

в) кумулята;

г) огіва.

Відповідь: 1) а; 2) б; 3) в; 4) г.

4. Під час плосковершинного розподілу коефіцієнт ексцесу:

а) $E_к = 0$;

б) $E_к < 0$;

в) $E_к > 0$.

Відповідь: 1) а, б; 2) б; 3) в; 4) б, в.

5. Значення медіани збігається із значенням середньої:

а) у симетричному розподілі;

б) в асиметричному розподілі;

в) зміну медіани передбачити неможливо.

Відповідь: 1) а; 2) б; 3) в; 4) –.

6. Коефіцієнт асиметрії розраховується за формулою:

а) $A = \frac{\bar{x} - M_0}{\sigma}$;

б) $V = \frac{\sigma}{\bar{x}} \cdot 100\%$;

в) $A = \frac{\sum \left(x_i - \bar{x}\right)^2}{n}$;

г) $f_t = \frac{y_t \cdot i \cdot \sum f}{\sigma}$.

Відповідь: 1) а, б; 2) а; 3) б; 4) –.

7. Співвідношення x < Me < Mo притаманне розподілам:

а) симетричним;

б) з правосторонньою асиметрією;

в) з лівосторонньою асиметрією;

г) з двома і більше вершинами.

Відповідь: 1) а; 2) б; 3) в; 4) г.

8. Середнє квадратичне та середнє лінійне відхилення ідентичні за змістом:

а) так;

б) ні.

Середнє квадратичне та середнє лінійне відхилення однакові за абсолютною величиною:

в) так; г) ні.

Відповідь*:* 1) а, в; 2) а, г; 3) б, в; 4) б, г.

9. На підставі співвідношення між значеннями характеристик центру розподілу (середньої, моди та медіани) можна виміряти:

а) асиметрію розподілу;

б) ексцес розподілу.

Відповідь: 1)а, б; 2) а; 3) б; 4) –.

ТЕМА 4
МЕТОДИ АНАЛІЗУ ВЗАЄМОЗВ'ЯЗКУ В ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ



4.1. Поняття взаємозв'язків між явищами та процесами психології. Види взаємозв'язків.

4.2. Основи кореляційного аналізу.

4.3. Основи регресійного аналізу.

4.4. Моделі зв'язку в порядкових вимірюваннях. Рангові кореляції.

4.5. Оцінка тісноти взаємозв'язку між номінальними даними.

4.1. Поняття взаємозв'язків між явищами та процесами психології. Види взаємозв'язків

Важливим елементом математико-статистичних досліджень у психології є виявлення специфічного зв'язку між різними змінними, ознаками, чинниками тощо. Ця специфічність полягає в тому, що залежність між змінними неможливо описати функціонально, тобто за допомогою однозначно визначеної функції. Зв'язок цих ознак виявляється принципово по-іншому: від того, якого значення набуде одна величина, залежить імовірна характеристика іншої. Отже, незважаючи на те, що неможливо точно вказати значення залежної ознаки, все ж таки відома деяка інформація про можливе розташування її значень, певного ймовірного інтервалу, середнього значення, міри мінливості тощо. Такі зв'язки між кількісними й порядковими показниками називають кореляцією, а між номінальними показниками – асоціацією.

Усі взаємозв'язки явищ і процесів суспільного життя поділяють на два види: *функціональні (повні)* та *стохастичні (нефункціональні)*.

Функціональний – це такий зв'язок, у разі якого кожному конкретному показнику ознаки чинника x відповідають певні ознаки результату y.

Функціональні зв’язки можуть бути виражені алгебраїчно за допомогою виразу, який встановлює атрибути чинника й результату, тобто повну відповідність між причиною та наслідком.

Стохастичний – це зв’язок, де певному показнику ознаки *x* відповідають кілька значень ознаки результату *y* (функцією не виражається).

Підвидом стохастичного зв’язку є *кореляційний.*

4.2. Основи кореляційного аналізу

Аналіз взаємозв’язків між досліджуваними ознаками – основний вид завдань, що зустрічається практично в будь-якому емпіричному дослідженні. Визначення взаємозв’язків між досліджуваними змінними цікавить дослідника не саме по собі, а як відображення відповідних причинно-наслідкових відносин.

Під час вивчення кореляцій намагаються встановити, чи існує якийсь зв’язок між двома показниками в одній вибірці (наприклад, між зростом і вагою дітей або між рівнем IQ і шкільною успішністю) або між двома різними вибірками (наприклад, у разі порівняння пар близнюків), і якщо цей зв’язок існує, то чи супроводжується збільшення одного показника зростанням (позитивна кореляція) або зменшенням (негативна кореляція) іншого.

Інакше кажучи, кореляційний аналіз допомагає встановити, чи можна прогнозувати можливі значення одного показника, знаючи величину іншого.

Кореляція (від англ. *співвідношення, відповідність*) – взаємозв’язок між показниками, що полягає в зміні середнього значення однієї з них залежно від зміни іншої.

Кореляційний аналіз охоплює:

- ❖ графічне представлення залежності;
- ❖ обчислення коефіцієнта кореляції;
- ❖ перевірку статистичної значущості зв’язку.

Як показник тісноти взаємозв’язку між досліджуваними показниками, використовується *коефіцієнт кореляції* (r).

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

Кореляційні взаємозв'язки відрізняються за силою, формою та напрямом.

Критерієм тісноти взаємозв'язку є абсолютне значення коефіцієнта кореляції. Сила цього зв'язку є прямим показником того, до якої міри спільна варіація досліджуваних змінних виявляється синхронізованою. Силу зв'язку можна оцінити за допомогою шкали Чедока.

Таблиця 4.1 – Сила зв'язку за шкалою Чедока

Характеристика сили зв'язку	Міра сили зв'язку
досить слабка	$0{,}00 < \lvert r \rvert \le 0{,}10$
слабка	$0{,}10 < \lvert r \rvert \le 0{,}30$
помірна	$0{,}30 < \lvert r \rvert \le 0{,}50$
середня	$0{,}50 < \lvert r \rvert \le 0{,}70$
сильна	$0{,}70 < \lvert r \rvert \le 0{,}90$
досить сильна	$0{,}90 < \lvert r \rvert \le 1{,}00$

Наочне уявлення про характер зв'язку дає діаграма розсіювання – графік, осі якого відповідають значенням двох змінних, а кожен досліджуваний є точкою.

Кореляції можуть мати дві форми: *прямолінійну* та *криволінійну (обернену).*

У разі ***прямолінійного*** *зв'язку* однакові зміни факторних ознак відповідають однаковим змінам результативних ознак.

Його можна точно або наближено зобразити рівнянням будь-якої прямої лінії.

У разі ***криволінійного*** *(оберненого)* зв'язку рівні зміни факторних ознак відповідають нерівній зміні результативних, тобто ознаки результату змінюється в протилежному напрямі відносно зміни значення факторної.

Криволінійні зв'язки можуть бути виражені рівнянням будь-якої кривої лінії: параболи, гіперболи, степеневої функції.

Критерієм напряму зв'язку є знак коефіцієнта кореляції.

Отже, завданням кореляційного аналізу є встановлення напряму (позитивний чи негативний) і форми (лінійна чи нелінійна) взаємозв'язку між показниками, вимірювання їхньої тісноти та перевірка рівня статистичної значущості отриманого коефіцієнта кореляції.

Наразі розроблено ряд різних коефіцієнтів кореляції. Найчастіше використовують коефіцієнт кореляції Пірсона, Спірмена та r-Кендалла. Зазвичай комп'ютерні статистичні програми пропонують ці три коефіцієнти як основні в меню «Кореляція» та надають методи для порівняння груп з метою вирішення інших дослідницьких завдань.

Вибір методу розрахунку коефіцієнтів кореляції залежить насамперед від:

- ✓ типу шкали вимірювання змінної;
- ✓ виду нормальності розподілу даних.

Визначення взаємозв'язків між досліджуваними ознаками, які приймають випадкові значення, починається з оцінки їхньої лінійності.

Для вивчення лінійного взаємозв'язку змінних, виміряних у кількісних шкалах (інтервальній або відношень), застосовується коефіцієнт кореляції r_{xy} Пірсона. Коли йдеться про кореляцію та не уточнюють деталі, мається на увазі саме цей метод.

Коефіцієнт кореляції r_{xy} Пірсона фіксує наявність лише лінійного взаємозв'язку між досліджуваними даними й варіює в межах від –1 до +1. Знак вказує на напрям взаємозв'язку, а абсолютне значення коефіцієнта – на силу зв'язку. Коефіцієнт кореляції Пірсона є параметричним методом і його коректне застосування можливе лише, якщо розподілення даних двох змінних не відрізняється від нормального виду.

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

$$r_{xy} = \frac{\overline{XY} - \overline{X} \cdot \overline{Y}}{\sigma_X \sigma_Y} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{(n\sum X^2 - \sum X \sum X)(n\sum Y^2 - \sum Y \sum Y)}}, \quad (4.1)$$

де X, Y – значення змінних;

$\overline{X}, \overline{Y}$ – середні значення змінних;

σ_X, σ_Y – стандартне відхилення змінних;

n – обсяг вибірки досліджуваних.

Показником розміру ефекту взаємозв'язку між досліджуваними показниками, крім безпосередньо самого емпіричного значення коефіцієнта r_{xy}, може виступати й коефіцієнт детермінації $r^2 = R$, тобто коефіцієнт кореляції в квадраті. Він змінюється від 0 до 1 (або від 0 до 100 %) і показує якою мірою мінливість однієї змінної обумовлена (детермінована) впливом іншої змінної. На підставі показника r_{xy} Дж. Коен виділив такі градації величини розміру ефекту:

- 0,00 ≤ r 2 < 0,01 – несуттєвий;
- 0,01 ≤ r 2 < 0,09 – малий;
- 0,09 ≤ r 2 < 0,25 – середній;
- 0,25 ≤ r 2 < 1,00 – великий.

4.3. Основи регресійного аналізу

Регресійний аналіз ґрунтується на припущенні, що взаємозв'язок між показниками дослідження чинника (ознака, яка впливає на іншу ознаку) та ознаками результату (ознака, яка змінюється в результаті впливу чинника) може бути виражений функцією *Y = f(x)*, яка називається рівнянням регресії.

Рівняння регресії – рівняння, яке виражає взаємозв'язок між даними, що досліджуються. Інакше кажучи, економіко-схематична модель залежності результативних ознак від факторних ознак.

Графік рівняння регресії – це *лінія регресії*, що відображує кореляційну залежність. На графіку по горизонтальній осі (OX) відкладаються значення факторних ознак, а на вертикальній осі (OY) – показник ознаки результату.

Згідно з аналітичними рівняннями ця залежність може бути лінійною або нелінійною.

Найпоширенішими рівняннями регресії є такі:

$y = a + bx$ – лінійне;

$y = ab^x$ – показникове;

$y = ax^b$ – степеневе;

$y = a + bx + cx^2$ – параболічне;

$y = a + \frac{b}{x}$ – гіперболічне,

де y – теоретичні дані ознаки результату;

a, b і c – коефіцієнти регресії (параметри рівняння регресії).

Першим етапом регресійного аналізу під час обґрунтування моделі є вирішення двох завдань: вибір факторних і результативних змінних та вибір типу рівняння регресії.

Правильний підбір ознаки й типу рівняння регресії вимагає теоретичного аналізу зв'язку. Для перевірки правильності вибору типу рівняння регресії часто використовують графічне представлення зв'язку. Під час побудови цього графіка на горизонтальній осі відкладаються показники факторного чинника *x*, а на вертикальній осі – значення результативного чинника *y*. На графіку кожній одиниці сукупності відповідає окрема точка. За формою точок на кореляційному полі роблять висновки про тип рівняння регресії. У разі великого обсягу сукупності на графік варто наносити середню групову величину аналітичного групування. Лінія групових середніх називається емпіричною лінією регресії. Для того щоб визначити тип рівняння регресії, використовують методи пошуку функцій. У цьому разі розраховують різні типи рівнянь регресії, й на основі статистичних та математичних критеріїв обирається найкраще з них. На етапі оцінювання лінії регресії параметри обраних рівнянь визначають з використанням методу найменших квадратів за допомогою побудови та розв'язання системи нормальних рівнянь. Лінійна функція відповідає такій системі рівнянь:

$$\begin{cases} na + b\sum x = \sum y \\ a\sum x + b\sum x^2 = \sum xy \end{cases}.$$

Особливої уваги заслуговує інтерпретація параметрів *a* та *b* у рівнянні лінійної регресії. Параметр *b* показує зміну середнього значення ознаки у власному вимірі під час збільшення факторної на одиницю власного виміру. Параметр *a* – це теоретичне значення Y у разі x = 0.

4.4. Моделі зв'язку в порядкових вимірюваннях. Рангові кореляції

У психології часто виникає потреба проаналізувати зв'язок між змінними, які неможливо виміряти в інтервальній шкалі, проте можна впорядкувати та проранжувати за ступенем послаблення або посилення прояву ознаки. Для того щоб визначити щільність зв'язку між даними, виміряними в порядкових шкалах, застосовують методи рангової кореляції. До них належать: коефіцієнти Спірмена та Кендала (використовують для визначення тісноти взаємозв'язку між величинами) й коефіцієнт конкордації (встановлює взаємозв'язок між декількома ознаками). Використання коефіцієнта лінійної кореляції Пірсона у разі, коли про закон розподілу й тип вимірювальної шкали надійної інформації немає, може призвести до виникнення значних помилок.

Методи рангової кореляції можуть бути використані для визначення тісноти зв'язку не тільки між кількісними змінними, а й між якісними ознаками за умови, що їхнє значення можна упорядкувати та проранжувати. Ці методи також можуть бути використані стосовно ознак, виміряних в інтервальних шкалах, проте їхня ефективність у цьому разі буде нижчою.

Непараметричним аналогом класичного коефіцієнта кореляції Пірсона є *коефіцієнт кореляції рангів Спірмена.* Його особливістю є те, що під час розрахунку враховуються ранги, а не середнє арифметичне та дисперсія змінних. Емпіричне значення коефіцієнта рангової кореляції Спірмена визначається за формулою:

$$r_s = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)}, \qquad (4.2)$$

де d – різниця рангів;

n – обсяг вибірки досліджуваних або кількість ознак.

Для рангів альтернативою кореляції за Спірменом є кореляція τ-Кендала.

Коефіцієнт конкордації Кендала використовують, коли сукупність об'єктів характеризується кількома послідовностями рангів і потрібно встановити взаємозв'язок між послідовностями. Такі завдання виникають, наприклад, під час аналізу експертних оцінок: кілька експертів ранжують тих самих досліджуваних осіб за певною якістю, а психологу варто визначити ступінь узгодженості думок групи експертів.

Згідно з кореляцією Кендала напрям зв'язку можна визначити, порівнюючи піддослідних попарно. Збігання зміни за *X* за направленістю зі зміною за *У* відображує позитивний зв'язок, незбігання – негативний зв'язок.

Емпіричне значення коефіцієнта рангової кореляції τ Кендала визначається за формулою:

$$\tau = \frac{P-Q}{n\cdot(n-1)/2}, \qquad (4.3)$$

де P – кількість збігів (погоджені пари);

Q – кількість інверсій (непогоджені пари);

n – обсяг вибірки досліджуваних або кількість ознак.

Показником розміру ефекту виступає безпосередньо саме емпіричне значення коефіцієнта τ-Кендала. У разі інтерпретації потрібно орієнтуватися на методичні рекомендації, запропоновані Джейкобом Коеном.

Таблиця 4.2 – Значення коефіцієнта Кендалла за Джейкобом Коеном

Рівень розміру ефекту	Величина коефіцієнта
Несуттєвий	$0{,}00 \le \lvert \tau \rvert < 0{,}10$
Малий	$0{,}10 \le \lvert \tau \rvert < 0{,}30$
Середній	$0{,}30 \le \lvert \tau \rvert < 0{,}50$
Великий	$0{,}50 \le \lvert \tau \rvert < 1{,}00$

4.5. Оцінка тісноти взаємозв'язку між номінальними даними

Досить часто в психології під час виконання різних видів досліджень фіксують й аналізують ознаки, що не мають безпосередньої кількісної міри. Такі дані представлені в номінальній (номінативній, найменувань) вимірювальній шкалі. У межах цієї шкали об'єкти класифікують, а класи здебільшого позначають номерами, які нічого не свідчать про властивості об'єкта, крім того, що він належить до певної групи.

У такому випадку для аналізу кількісної залежності між змінними використовують коефіцієнти, розроблені на основі тесту χ^2-Пірсона. Значення 0 відповідає повній незалежності змінних, а 1 – їхній максимальній залежності. У цих коефіцієнтах знак не інтерпретується. Він не вказує на напрям залежності, тому що залежить від кодувань (позначень) градацій значень ознак.

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

Є декілька способів розрахунку коефіцієнта контингенції Пірсона φ. Найбільш простий існує за таблицею зв'язаності ознак.

Таблиця 4.3 – Таблиця зв'язаності

a	b	a + b
c	d	c + d
a + c	b + d	a + b + c + d

У такому разі розрахункове значення коефіцієнта контингенції обчислюється за формулою:

$$\varphi = \frac{ad - bc}{\sqrt{(a+b)\cdot(c+d)\cdot(a+c)\cdot(b+d)}}. \qquad (4.4)$$

Коефіцієнт контингенції не має своєї таблиці з критичними значеннями. Статистична значущість взаємозв'язку оцінюється за допомогою критерію χ^2-Пірсона.

$$\chi^2 = \varphi^2 \cdot n, \tag{4.5}$$

де φ – емпіричне значення коефіцієнта контингенції;

n – обсяг вибірки досліджуваних.

Для прийняття рішення про рівень статистичної значущості потрібно емпіричне (розрахункове) значення критерію χ^2-Пірсона порівняти з табличним (критичним) значенням. Критичне значення визначається залежно від числа ступенів свободи та заданого α-рівня.

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

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

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

$$Q = \frac{ad-bc}{ad+bc}, \tag{4.6}$$

де *a, b, c, d* – дані, що визначаються згідно з таблицею зв'язаності ознак розміром 2 × 2.

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

Для оцінки залежності в ситуаціях, коли кожна змінна представлена в номінальній шкалі та хоча б одна з них має більш ніж дві градацій, застосовуються коефіцієнти взаємної зв'язаності ознак Пірсона *C*, Чупрова *K* та Крамера *V*.

Ці критерії ґрунтуються на безпосередньому розрахунку критерію χ^2-Пірсона. Потрібно зауважити, що із збільшенням значення χ^2 сила залежності коефіцієнтів взаємної пов'язаності зростає. Розрахункова формула χ^2-Пірсона має такий вигляд:

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_{\text{емп.}} - f_{\text{теор.}})^2}{f_{\text{теор.}}}, \quad (4.7)$$

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

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

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

$$\mathrm{f_{ij}} = \frac{\mathrm{f_i \cdot f_j}}{\mathrm{n}}, \quad (4.8)$$

де $\mathrm{f_i}$ – сума частот у всіх комірках *i*-рядка;

$\mathrm{f_j}$ – сума частот у всіх комірках *j*-стовпчика;

n – сума частот у всій таблиці пов'язаності ознак.

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

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

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}; \quad (4.9.)$$

$$K = \sqrt{\frac{\chi^2}{n \cdot \sqrt{(k-1)\cdot(m-1)}}}; \quad (4.10)$$

$$V = \sqrt{\frac{\chi^2}{n \cdot (c-1)}}, \quad (4.11)$$

де χ^2 – емпіричне значення критерію χ^2-Пірсона (у всіх формулах);

n – обсяг вибірки досліджуваних (у всіх формулах);

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

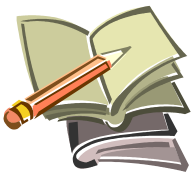
c – найменша кількість градацій із двох змінних (формула Крамера *V*).

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

Таблиць із критичними значеннями для коефіцієнтів пов'язаності немає. Тому для прийняття статистичного рішення потрібно:

- вибрати рівень значущості та сформулювати нульову й альтернативну гіпотези;
- обчислити емпіричне значення критерію χ^2-Пірсона;
- розрахувати значення коефіцієнта взаємної пов'язаності;
- порівняти емпіричне значення критерію χ^2 з критичним значенням для відповідного числа ступенів свободи. df = (k –1)·(m –1), де k, m – кількість рядків і стовпчиків таблиці пов'язаності ознак.
- якщо χ^2 емп. $\geq \chi^2$ крит., то залежність між ознаками статистично значуща, тобто ознаки змінюються узгоджено. Якщо χ^2 емп. $< \chi^2$ крит., то формулюється висновок про відсутність статистично значущої залежності.

Приклади розв'язку типових задач



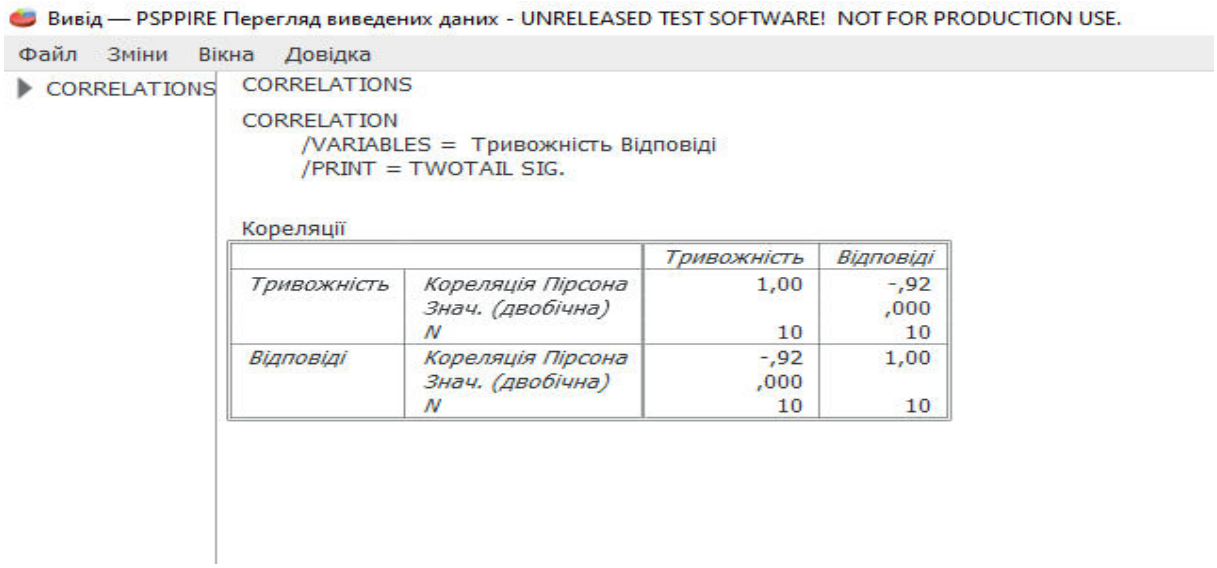
1. За результатами тестування досліджуваних, визначити взаємозв'язок між показником тривожності та кількістю правильних відповідей.

Таблиця 4.4 – Розподіл правильних відповідей за показником тривожності

№ з/п	Показник тривожності	Кількість правильних відповідей
1.	2,8	8,0
2.	1,9	13,0
3.	2,9	5,0
4.	2,0	16,0
5.	3,0	11,0
6.	3,1	6,0
7.	2,8	9,0
8.	1,6	18,0
9.	3,2	5,0
10.	3,3	2,0

Розв'язання:

За допомогою програми PSPP розрахуємо коефіцієнт кореляції Пірсона.


Вивід — PSPPIRE Перегляд виведених даних - UNRELEASED TEST SOFTWARE! NOT FOR PRODUCTION USE.

Файл Зміни Вікна Довідка

CORRELATIONS

```
CORRELATIONS
CORRELATION
      /VARIABLES = Тривожність Відповіді
      /PRINT = TWOTAIL SIG.
```

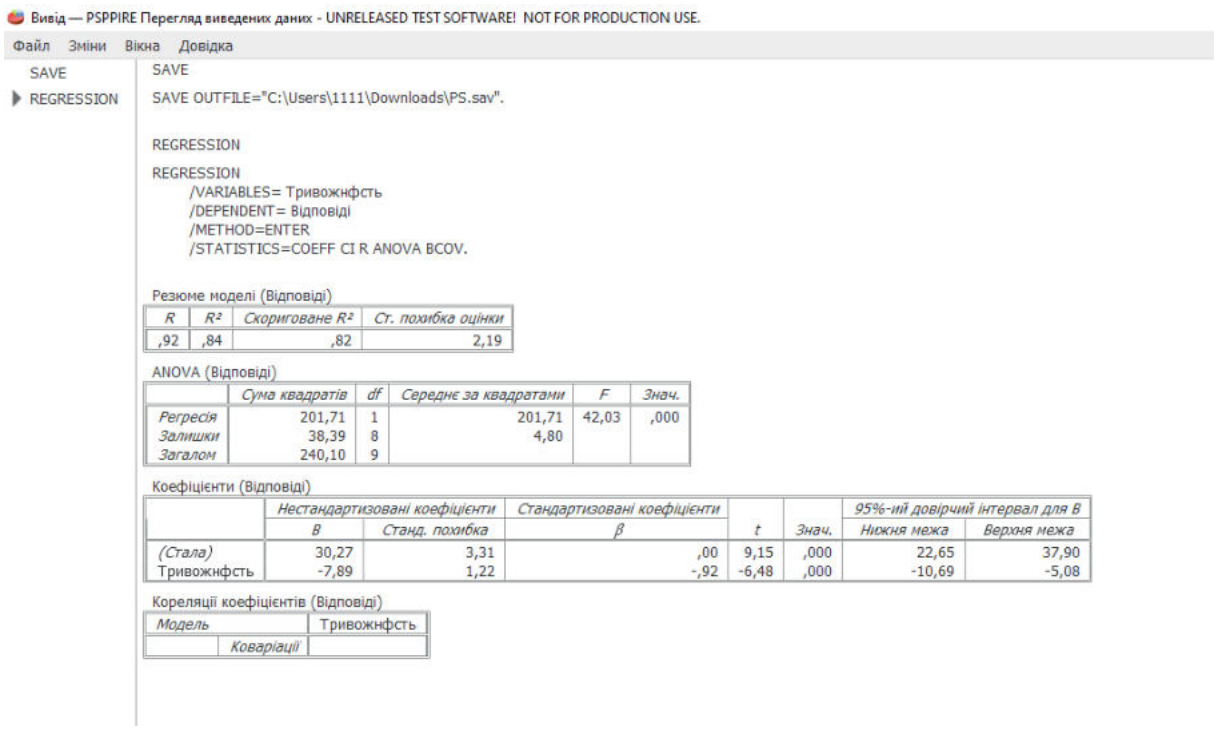
Кореляції

		Тривожність	Відповіді
Тривожність	Кореляція Пірсона	1,00	-,92
	Знач. (двобічна)		,000
	N	10	10
Відповіді	Кореляція Пірсона	-,92	1,00
	Знач. (двобічна)	,000	
	N	10	10

Рисунок 4.1 – Результати розрахунку коефіцієнта кореляції в PSPP

Отже, між показником тривожності й кількістю правильних відповідей існує тісний обернений зв'язок. На 92 % кількість правильних відповідей залежить від рівня тривожності.

Результати регресійного аналізу визначать ступінь впливу показника тривожності на кількість правильних відповідей.


Вивід — PSPPIRE Перегляд виведених даних - UNRELEASED TEST SOFTWARE! NOT FOR PRODUCTION USE.

Файл Зміни Вікна Довідка

SAVE

REGRESSION

```
SAVE
SAVE OUTFILE="C:\Users\1111\Downloads\PS.sav".

REGRESSION
REGRESSION
      /VARIABLES= Тривожнфсть
      /DEPENDENT= Відповіді
      /METHOD=ENTER
      /STATISTICS=COEFF CI R ANOVA BCOV.
```

Резюме моделі (Відповіді)

R	R^2	Скориговане R^2	Ст. похибка оцінки
,92	,84	,82	2,19

ANOVA (Відповіді)

	Сума квадратів	df	Середнє за квадратами	F	Знач.
Регресія	201,71	1	201,71	42,03	,000
Залишки	38,39	8	4,80		
Загалом	240,10	9			

Коефіцієнти (Відповіді)

	Нестандартизовані коефіцієнти		Стандартизовані коефіцієнти			95%-ий довірчий інтервал для B	
	B	Станд. похибка	β	t	Знач.	Нижня межа	Верхня межа
(Стала)	30,27	3,31	,00	9,15	,000	22,65	37,90
Тривожнфсть	-7,89	1,22	-,92	-6,48	,000	-10,69	-5,08

Кореляції коефіцієнтів (Відповіді)

Модель		Тривожнфсть
	Коваріації	

Рисунок 4.2 – Результати регресійного аналізу в PSPP

Отримали лінійне рівняння регресії: Y = 30,27 –7,89x.

Тобто в разі зростання показника тривожності на одиницю кількість правильних відповідей зменшується приблизно на 8.

2. Розглянемо дві дихотомічні ознаки: стать (1 – чоловік, 0 – жінка) і паління (1 – палить, 0 – не палить). Припустимо, для перевірки зв'язку між статтю і палінням досліджено 100 осіб.

Таблиця 4.5 – Чотириклітинна таблиця спряженості між ознаками «стать» і «паління»

Схильність до паління	Стать		Разом
	Чоловік	Жінка	
Палить	42	18	60
Не палить	13	27	40
Разом	55	45	100

Чи взаємопов'язані досліджувані ознаки «стать» і «паління»?

Розв'язання:

Обчислимо коефіцієнти асоціації Юла та контингенції:

$$Q = \frac{ad-bc}{ad+bc} = \frac{42 \cdot 27 - 18 \cdot 13}{42 \cdot 27 + 18 \cdot 13} = \frac{900}{1368} = 0{,}6579;$$

$$\Phi = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}} = \frac{42 \cdot 27 - 18 \cdot 13}{\sqrt{(42 + 18)(13 + 27)(42 + 13)(18 + 27)}} = \frac{3}{\sqrt{66}} = 0{,}3693.$$

Як бачимо, ці коефіцієнти відмінні від нуля й водночас суттєво різняться. Зважаючи на однобічність коефіцієнта Юла і двобічність коефіцієнта контингенції, доходимо висновку, що залежність між ознаками «стать» і «паління» існує, причому вона швидше однобічна, ніж багатобічна, тобто серед чоловіків більшість палить, а серед жінок такої явної переваги до непаління, як у чоловіків до паління, не спостерігається.

3. У дослідженнях, які моделюють діяльність авіадиспетчерів, група досліджуваних здобувачів вищої освіти фізичного факультету державного університету проходила практику перед початком роботи на тренажері. Досліджувані мали розв'язати задачі вибору оптимального типу злітно-посадочної смуги для заданого типу літака.

Таблиця 4.6 – Показники кількості помилок,
зроблених під час тренувальної сесії,
і рівня вербального інтелекту студентів-фізиків

№ з/п	Код імені спостережуваного	Кількість помилок	Показник вербального інтелекту
1.	А. Т.	29	131
2.	А. П.	54	132
3.	І. Ч.	13	121
4.	А. Ц.	8	127
5.	А. С.	14	136
6.	Є. К.	26	124
7.	А. К.	9	134
8.	Л. Б.	20	136
9.	А. І.	2	132
10.	В. Ф.	17	136
Сума		192	1309
Середнє		19,2	130,9

Чи пов’язана кількість помилок, зроблених досліджуваними під час тренувальної сесії, з показником вербального інтелекту, обчисленим за методикою Д. Векслера?

Розв’язання:

Сформулюємо основну й альтернативну гіпотези.

H0: кореляція між кількістю помилок і рівнем вербального інтелекту не відрізняється від нуля;

H1: кореляція між кількістю помилок і рівнем вербального інтелекту відрізняється статистично значуще від нуля.

Проранжуємо обидва показника від найменшого до найбільшого, підрахуємо різниці між рангами та піднесемо їх до квадрату. Дані занесемо в таблицю.

Таблиця 4.7 – Розрахунок для рангового коефіцієнта кореляції Спірмена під час порівняння показників кількості помилок і вербального інтелекту здобувачів-фізиків

№ з/п	Код імені спостережуваного	Кількість помилок		Показник вербального інтелекту		Різниця рангів d	d^2
		Значення	Ранг	Значення	Ранг		
1.	А. Т.	29	9	131	4	5	25
2.	А. П.	54	10	132	5,5	4,5	20,25
3.	І. Ч.	13	4	121	1	3	9
4.	А. Ц.	8	2	127	3	–1	1
5.	А. С.	14	5	136	9	–4	16
6.	Є. К.	26	8	124	2	6	36
7.	А. К.	9	3	134	7	–4	16
8.	Л. Б.	20	7	136	9	–2	4
9.	А. І.	2	1	132	5,5	–4,5	20,25
10.	В. Ф.	17	6	136	9	–3	9
Разом			55		55	0	156

Оскільки деякі показники вербального інтелекту мають однакові ранги, розрахуємо поправку:

$$\mathrm{T_b} = \sum \frac{\mathrm{b}^3 - \mathrm{b}}{12} = \frac{2^3 - 2}{12} + \frac{3^3 - 3}{12} = 2{,}5.$$

Емпіричне значення коефіцієнта рангової кореляції:

$$r_s^{\text{емп}} = 1 - 6 \cdot \frac{\sum d_i^2 + T_a + T_b}{n \cdot (n^2 - 1)} = 1 - 6 \cdot \frac{156{,}5 + 0 + 2{,}5}{10 \cdot (10^2 - 1)} = 0{,}036.$$

Визначимо критичні значення коефіцієнта для заданого n = 10:

$$r_s^{0,05} = 0{,}64; \qquad r_s^{0,01} = 0{,}79.$$

Оскільки $|r_s^{\text{емп}}| \geq r_s^{0,05}$ гіпотеза Н0 приймається. Кореляція між показниками кількості помилок і рівнем вербального інтелекту не відрізняється від нуля.

Перелік питань і завдань для самоконтролю



1. Види взаємозв’язків між досліджуваними явищами.
2. Сутність стохастичної залежності.
3. Сутність кореляційної залежності.
4. Модель аналітичного групування.
5. Теоретичне обґрунтування моделі.
6. Поняття регресії.
7. Оцінка лінії регресії.
8. Вимірювання щільності зв’язку.
9. Дослідження істотності зв’язку в психологічних дослідженнях.
10. Поняття критерію Фішера.
11. Досліджується читацька аудиторія деякого журналу за гендерною ознакою. У результаті опитування 100 осіб отримали такі дані:

Таблиця 4.8 – Чотириклітинна таблиця спряженості між ознаками «стать» і «читач журналу»

Ставлення до журналу	Стать		Разом
	Чоловік	Жінка	
Читач	12	25	37
Нечитач	31	32	63
Разом	43	57	100

Чи взаємопов’язані статистично досліджувані ознаки «стать» і «читач журналу»? Зробіть висновок, розрахувавши коефіцієнти асоціації Юла та контингенції.

12. У дослідженні, присвяченому проблемі ціннісної переорієнтації, було визначено ієрархію термінальних цінностей у батьків та їхніх дорослих дітей за методикою М. Рокіча. У таблиці наведено ранжування кінцевих значень, отриманих у результаті опитування пари мати – дочка (мати – 66 років, дочка – 42 роки).

Таблиця 4.9 – Ранги термінальних цінностей
за списком М. Рокича в індивідуальних ієрархіях матері та дочки

№ з/п	Термінальні цінності	Ряд цінностей в ієрархії	
		матері	дочки
1.	Активне діяльне життя	15	15
2.	Життєва мудрість	1	3
3.	Здоров’я	7	14
4.	Цікава робота	8	12
5.	Краса природи й мистецтво	16	17
6.	Любов	11	10
7.	Матеріально забезпечене життя	12	13
8.	Суспільне визнання	17	5
9.	Пізнання	5	1
10.	Продуктивне життя	2	2
11.	Розвиток	6	8
12.	Розваги	18	18
13.	Свобода	4	6
14.	Щасливе подружнє життя	13	4
15.	Щастя інших	14	16
16.	Творчість	10	9
17.	Упевненість у собі	3	7
Разом		171	171

Чи пов’язані ці ієрархії цінностей? Щоб відповісти на це запитання, скористайтеся ранговою кореляцією Спірмена.

Тести для самоконтролю

1. Кореляційним називається:

а) повний зв’язок між двома або більше ознаками;

б) повний взаємозв’язок між даними, що досліджуються;

в) неповний зв’язок між ознаками, встановлений на основі одного спостереження;

г) неповний зв’язок між ознаками, який виникає під час спостереження значної кількості даних.

Відповідь: 1) а; 2) б; 3) в; 4) г.

2. За межі правильного визначення поняття «кореляція» виходить:

а) залежність між випадковими величинами, що не має функціонального характеру;

б) середнє значення однієї ознаки, що змінюється залежно від значення іншої ознаки;

в) визначення форми зв'язку;

г) неповна залежність між характеристиками.

Відповідь: 1) а; 2) б; 3) в; 4) г.

3. Регресія – це:

а) математичне очікування значення змінної, спричинене зміною випадкової змінної;

б) тіснота зв'язку;

в) різновид пропорційної залежності двох змінних;

г) лінія, вид залежності середнього значення ознаки результату від факторної.

Відповідь: 1) а; 2) б; 3) в; 4) г.

4. Стохастичний зв'язок – це:

а) форма кореляційного зв'язку;

б) вид кореляції;

в) взаємозв'язок між досліджуваними випадковими величинами, в разі якого зміна однієї з них викликає зміну закону розподілу інших;

г) тип зв'язку між випадковими величинами.

Відповідь: 1) а; 2) б; 3) в; 4) г.

5. Поняття «форма кореляційного зв'язку» – це:

а) аналітичне рівняння зв'язку;

б) тип аналітичної формули, яка відображує залежність між досліджуваними ознаками;

в) вид дослідження взаємозалежностей між ознаками;

г) кутовий коефіцієнт у прямолінійному рівнянні зв'язку.

Відповідь: 1) а; 2) б; 3) в; 4) г.

6. За межі визначення завдань кореляційного аналізу виходить:

а) оцінка параметрів (середнє, дисперсія, коефіцієнт кореляції) для нормально розподілених сукупностей;

б) визначення індивідуального спільного впливу чинників на ознаку результату;

в) виявлення взаємозалежної структури ознак;

г) перевірка значущості оцінених параметрів.

Відповідь: 1) а; 2) б; 3) в; 4) г.

7. «Мультиколінеарність» – це:

а) кореляції між чинниками, які не були теоретично доведені;

б) кореляція між досліджуваними ознаками;

в) кореляція між чинниками;

г) поєднання декількох рядів розподілу для побудови кореляційної моделі.

Відповідь: 1) а; 2) б; 3) в; 4) г.

8. За межі визначення переваг кореляційно-регресійного методу перед методом статистичних угрупувань виходить:

а) можливість дослідження взаємозв'язків між декількома ознаками одночасно;

б) раціональне та наочне представлення цифрових характеристик досліджуваного явища;

в) отримання показника тісноти зв'язку та оцінка параметрів генеральної сукупності на основі вибіркових даних;

г) елімінування випадкової варіації досліджуваних ознак.

Відповідь: 1) а; 2) б; 3) в; 4) г.

9. Кореляція, в разі якої ознака розглядається як результат дії двох і більше чинників, називається:

а) криволінійною;

б) прямолінійною;

в) множинною;

г) простою.

Відповідь: 1) а; 2) б; 3) в; 4) г.

10. Кореляційний зв'язок, у разі якого значення результативної оцінки змінюється в протилежному напрямі щодо факторної, називається:

а) обернений;

б) криволінійний;

в) прямий;

г) прямолінійний.

Відповідь: 1) а; 2) б; 3) в; 4) г.

ТЕМА 5
ДИСПЕРСІЙНИЙ АНАЛІЗ У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ



5.1. Основні поняття дисперсійного аналізу.
5.2. Однофакторний дисперсійний аналіз.
5.3. Двофакторний дисперсійний аналіз.

5.1. Основні поняття дисперсійного аналізу

Дисперсійний аналіз, або ANOVA (абревіатура англійського виразу *Analysis of Variance*), є одним із найбільш широко використовуваних методів статистичного аналізу в психології. Його популярність обумовлена принаймні двома причинами:

1. ANOVA, подібно t-критерієм Стьюдента, дозволяє оцінити відмінності між вибірковими середніми, проте, на відміну від t-критерію, в ньому немає обмежень на кількість порівнюваних середніх.

2. ANOVA дозволяє психологу мати справу з двома або більш незалежними змінними (ознаками, чинниками) одночасно, оцінюючи ефект як кожної з них окремо.

Сутність дисперсійного аналізу полягає в розкладанні (аналізі) дисперсії на складові компоненти, порівнюючи, які одна з одною за допомогою F-критерію, можна оцінити їхній внесок у загальну варіацію даних. Завдяки тому, що дисперсійний аналіз є методом оцінки різних експериментальних взаємодій, він вимагає планування емпіричного дослідження.

Зазвичай в експерименті дослідника цікавлять деякі причини або чинники, які впливають на спостережувану ознаку. Вони називаються *контрольованими чинниками* на відміну від інших дій, які теж впливають на спостережувану ознаку, але не контролюються. Контрольованість чинника заключається у виборі певної градації його зміни дослідником. Коли контрольованих чинників два або більше, використовують багатофакторний дисперсійний аналіз. У

дослідженні додатковими досліджуваними чинниками можуть бути: стать, професійний досвід, рівень освіти та ін. У разі наявності більш ніж одного чинника з'являється можливість оцінити вплив на залежну змінну міжфакторної взаємодії. Отже, дисперсійний аналіз дозволяє встановити факт впливу чинника на ознаку або, якщо контрольованих чинників декілька, виявити й оцінити їх взаємодію.

5.2. Однофакторний дисперсійний аналіз

Основною метою однофакторного аналізу зазвичай є оцінка величини впливу певного фактора на досліджувану реакцію. Іншою метою є порівняння кількох факторів між собою, щоб визначити відмінності в їхньому впливі на відповідь, які часто називають факторними контрастами. Попереднім етапом є перевірка нульової гіпотези про відсутність впливу досліджуваного фактора. Тобто гіпотеза полягає в тому, що зміни значення ознаки в порівнюваних вибірках є випадковими й що всі дані належать до однієї генеральної сукупності.

Якщо нульова гіпотеза відхиляється, наступним кроком є кількісна оцінка впливу досліджуваного фактора та побудова інтервалів довіри для одержаних характеристик. Якщо нульова гіпотеза не відхиляється, то зазвичай приймається нульова гіпотеза й робиться висновок про невпливовість. Якщо є підстави вважати, що такий ефект має бути (наприклад, він може випливати з теоретичних уявлень про предмет дослідження), варто перевірити інші фактори, які можуть його маскувати. В однофакторному дисперсійному аналізі вихідна інформація подається як таблиця, де кількість стовпчиків відповідає кількості факторних рівнів, а кількість значень у кожному стовпчику – кількості спостережень відповідному факторному рівню. Різні факторні рівні можуть мати різний обсяг спостережень. У такому разі результати спостережень на різних рівнях вважають вибіркою з нормально розподіленої генеральної сукупності, середнє й дисперсія якої однакові та не залежать від рівня. Метою аналізу є перевірка нульової гіпотези про те, що середні значення розглянутих сукупностей є рівними.

Основною метою дисперсійного аналізу є виявлення впливу окремих чинників, які зумовлюють варіацію ознаки. В основі дисперсійного аналізу лежить закон розкладання загальної дисперсії на складові, згідно з яким загальна дисперсія результативної ознаки складається з міжгрупової дисперсії та середньої дисперсії з групових (залишкової):

$$\sigma^2 = \overline{\sigma^2} + \delta^2, \tag{5.1}$$

де

загальна дисперсія – характеризує варіацію значень ознаки y_s навколо загальної середньої $\bar{y}$:

$$\sigma^2 = \frac{\sum_1^n (y_j - \bar{y})2}{n}; \tag{5.2}$$

середня з групових дисперсій – узагальнює внутрішньо групову варіацію:

$$\overline{\sigma^2} = \frac{\sum_1^m \sigma_j^2 \cdot f_j}{\sum_1^m f_j}; \tag{5.3}$$

міжгрупова дисперсія – характеризує варіацію групових середніх $\bar{y}_j$ навколо загальної середньої:

$$\delta^2 = \frac{\sum_1^m (\bar{y}_j - \bar{y})^2 \cdot f_j}{\sum_1^m f_j}. \tag{5.4}$$

Відношення міжгрупової дисперсії до дисперсії загальної визначає щільність кореляційного зв’язку та називається *коефіцієнтом детермінації.*

Його позначають буквою *R* й обчислюють за формулою:

$$R^2 = \frac{\delta^2}{\sigma^2} \cdot \tag{5.5}$$

Кореляційне відношення визначає щільність взаємозв'язку між факторною (групувальною) ознакою та результативною й відсоток міжгрупової дисперсії у загальній. Показник позначають грецькою літерою η – (тета).

Визначається:

$$\eta = \sqrt{\frac{\delta^2}{\sigma^2}}, \tag{5.6}$$

де δ^2 – міжгрупова дисперсія;

σ^2 – загальна дисперсія.

Кореляційне відношення варіює в межах від 0 до 1. Якщо $\eta = 0$, міжгрупова дисперсія дорівнюватиме 0. Це буває в разі *однаковості* всіх групових середніх і *відсутності* кореляційної залежності між ознаками.

Якщо $\eta = 1$, загальна дисперсія дорівнює міжгруповій, а групова середня – нулю. У цьому разі конкретному показнику значення факторної ознаки відповідатиме одне значення результативної ознаки, тобто має місце *функціональний зв'язок.*

Індекс кореляції обчислюється співвідношенням внутрішньогрупової дисперсії до загальної. Визначається за формулою:

$$R = \sqrt{1 - \frac{\overline{\sigma^2}}{\sigma^2}}, \tag{5.7}$$

де σ^2 – внутрішньогрупова дисперсія,

σ^2 – загальна дисперсія.

Чим ближче індекс до 1, тим взаємозв'язок між показниками більш щільний.

5.3. Двофакторний дисперсійний аналіз

Для відповідної нормально розподіленої вибірки використовується двофакторний дисперсійний аналіз. Дані представлені в таблиці, де стовпчики представляють дані, які відповідають певному рівню першого чинника, а рядки – дані, які відповідають рівню другого.

Під час двофакторного дисперсійного аналізу використовують модель, за якою значення відгуку є сумою внесків окремо кожного із чинників b_i і t_j, а також незалежної від чинників випадкової компоненти ε_{ij}:

$$x_{ij} = b_{i+} + t_j + \varepsilon_{ij}. \quad (5.8)$$

У разі коли випадкова компонента ε_{ij} підпорядковується нормальному розподілу з нульовим середнім і рівними для всіх i, j дисперсіями, σ^2 застосовують двофакторний дисперсійний аналіз (дисперсійний аналіз за двома ознаками).

Нульова гіпотеза може полягати в рівності ефектів стовпчиків між собою H_1: $\tau_1 = \tau_2 = \cdots = \tau_к = 0$ або рівності ефектів рядків між собою H_2: $\beta_1 = \beta_2 = \cdots = \beta_n = 0$, тобто в першому випадку припускають відсутність впливу чинника А, а у другому – чинника В. Як і у випадку однофакторного дисперсійного аналізу, в цьому разі розраховують дві оцінки дисперсії.

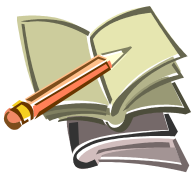
Для перевірки справедливості гіпотези H_1 потрібно розраховувати відношення дисперсій:

$$F = \frac{\sigma_2^2}{\sigma_1^2}. \quad (5.9)$$

Нульову гіпотезу приймають у разі рівня значущості α, якщо $F < F_{1-\alpha}$, де α-квантиль F-розподілу з відповідними кількостями ступенів вільності.

Якщо припущення, потрібні для двофакторного дисперсійного аналізу, не можуть бути виконані, використовується непараметричний ранговий тест Фрідмана (Friedman, Kendall and Smith). Цей критерій не залежить від типу розподілу. Передбачається лише, що розподіл величин ε_{ij} є однаковим і неперервним, а самі вони незалежні одна від одної.

Приклади розв’язку типових задач



1. Заклад вищої освіти проводив дослідження щодо кількості художньої літератури, яку читають дві академічні групи (гуманітарна й технічна). Кожна група складалася з 15 здобувачів, які навчаються в ЗВО. Читання літератури оцінювалося за загальною кількістю прочитаних сторінок за останній рік. Дані опитування представлені в таблиці нижче в порядку зростання. За допомогою індивідуальних абсолютних відхилень від середнього значення $\bar{x} = 115$ визначимо дисперсію та стандартне відхилення для обох груп.

Таблиця 5.1 – Кількість друкованих сторінок художньої літератури, прочитаної здобувачами закладів вищої освіти

Позиція в упорядкованому списку	Кількість друкованих сторінок літератури, прочитаної за останній рік серед здобувачів групи	
	Гуманітарної	Технічної
1	99	64
2	102	65
3	103	67
4	106	68
5	108	70
6	112	90
7	115	105
8	115	115
9	117	115
10	120	135
11	121	150
12	122	165
13	126	167
14	128	172
15	131	177
Разом	1 725	1 725

Розв’язання:

Обчислимо квадрати абсолютних відхилень та їхні суми для кожної групи і заведемо їх у таблицю.

Таблиця 5.2 – Кількість аркушів художньої літератури, прочитаної студентом за рік

Абсолютні відхилення та квадрати абсолютних відхилень від усередненого значення для студентів груп спрямованості					
Гуманітарної			Технічної		
Значення (x_i)	Абсолютне відхилення $\lvert x_i - \bar{x}\rvert$	Квадрат абсолютного відхилення $\lvert x_i - \bar{x}\rvert^2$	Значення (x_i)	Абсолютне відхилення	Квадрат абсолютного відхилення
99	16	256	64	51	2 601
102	13	169	65	50	2 500
103	12	144	67	48	2 304
106	9	81	68	47	2 209
108	7	49	70	45	2 025
112	3	9	90	25	625
115	0	0	105	10	100
115	0	0	115	0	0
117	2	4	115	0	0
120	5	25	135	20	400
121	6	36	150	35	1 225
122	7	49	165	50	2 500
126	11	121	167	52	2 704
128	13	169	172	57	3 249
131	16	256	177	62	3 844
Разом		1 368	Разом		26 286

За отриманими даними обчислюємо дисперсії (як середнє значення квадратів абсолютних відхилень) і стандартні відхилення (як корінь, квадратний з дисперсії):

$$S^2_{\text{гум}} = \frac{1368}{15} = 91{,}2; \qquad S^2_{\text{тех}} = \frac{26286}{15} = 1752{,}4;$$

$$\sigma_{\text{гум}} = \sqrt{91{,}2} \approx 9{,}55; \qquad \sigma_{\text{гум}} = \sqrt{1752{,}4} \approx 41{,}86.$$

З цього прикладу видно, що дисперсію та стандартні відхилення доцільно визначати, записуючи результати проміжних обчислень у таблиці.

Перелік питань і завдань для самоконтролю



1. Поняття дисперсійного аналізу.
2. Сутність дисперсійного аналізу.
3. Види дисперсій.
4. Математичні властивості дисперсії.
5. Поняття загальної дисперсії.
6. Поняття групової дисперсії.
7. Поняття середньої дисперсії з групових.
8. Правило додавання дисперсій.
9. Однофакторний дисперсійний аналіз у психології.
10. Багатофакторний дисперсійний аналіз у психології.
11. За допомогою дисперсійного однофакторного аналізу на 1 % рівня значущості зробити висновки щодо впливу рівня відповідальності на показники успішності виконання завдань учнями.

Таблиця 5.3 – Розподіл показників успішності виконання завдання за рівнем відповідальності

Відповідальність	Показники успішності виконання завдань					
Низька	3	1	2	3	3	2
Середня	4	5	3	5	4	3
Висока	6	4	5	6	5	6

12.Зробити висновки на 5 % рівні значущості щодо впливу факторів А і В на значення 16 об’єктів дослідження.

Таблиця 5.4 – Розподіл об’єктів дослідження за факторами впливу

Фактор В (рівень)	Фактор А (рівень)	
	А1 (високий)	А2 (низький)
В1 (низький)	8, 5, 4, 6	4, 4, 5, 3
В2 (високий)	5, 4, 5, 3	5, 7, 7, 6

Тести для самоконтролю

1. Зміст причинно-наслідкового характеру зв'язку заключається в тому, що за певних умов:

а) кореляційне відношення наближається до 0;

б) одні явища (факторні) спричиняють інші (результативні);

в) одному значенню фактора відповідає одне значення результату;

г) одні явища (результативні) спричиняють варіацію інших явищ (факторних).

Відповідь: 1) а; 2) б; 3) в; 4) г.

2. За стохастичного зв'язку кожному значенню факторної ознаки відповідає:

а) одне значення ознаки результату;

б) множина значень ознаки результату.

в) середнє значення ознаки результату;

г) модальне значення результату.

Відповідь: 1) а; 2) б; 3) в; 4) г.

3. Функціональними є:

а) площа круга від радіуса;

б) сума амортизації від норми амортизаційних відрахувань;

в) вартість квартири від типу будинку.

Відповідь: 1) а, б; 2) а, в; 3) б, в; 4) а, б, в.

4. Кореляція – це підвид зв'язку:

а) функціонального;

б) стохастичного.

Соціально-економічним явищам він:

в) притаманний;

г) непритаманний.

Відповідь: 1) а, в; 2) б, в; 3) а, г; 4) б, г.

5. За лінійного зв'язку однаковим змінам факторної ознаки відповідає зміна результативної ознаки:

а) рівномірна;

б) різна.

За оберненого зв'язку в разі збільшення факторної ознаки результативна ознака:

в) зменшується;

г) збільшується.

Відповідь: 1) а, в; 2) б, в; 3) а, г; 4) б; г.

6. Щільність зв'язку в аналітичних групуваннях оцінюють:

а) коефіцієнтом спряженості;

б) коефіцієнтом детермінації;

в) кореляційним відношенням;

г) коефіцієнтом контингенції.

Відповідь: 1) а; 2) б; 3) в; 4) г.

7. η^2 змінюється:

а) від О до 1;

б) від –1 до 1.

Якщо $\eta^2 = 0$, то зв'язок:

в) відсутній;

г) функціональний.

Відповідь: 1) а, в; 2) а, г; 3) б, в; 4) б, г.

8. Кореляційне відношення є часткою:

а) міжгрупової дисперсії в залишковій;

б) міжгрупової дисперсії в загальній;

в) залишкової дисперсії в загальній.

Відповідь: 1) а, в; 2) а; 3) б, в; 4) б.

ТЕМА 6
ФАКТОРНИЙ І КЛАСТЕРНИЙ АНАЛІЗИ В ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ



6.1. Факторний аналіз.
6.2. Кластерний аналіз.

6.1. Факторний аналіз

Факторний аналіз являє собою сукупність методів, покликаних визначити, наскільки пов'язані (корелюють) змінні можуть бути згруповані так, щоб кожну групу можна було розглядати як одну складову змінну або фактор, а не як ряд окремих змінних. Найбільш поширене застосування факторного аналізу в психології пояснюється можливістю об'єднати сукупність пунктів опитувальника, що оцінюють конкретну характеристику, так, щоб отримати загальний індикатор цієї характеристики.

Основними завданнями факторного аналізу є:

1. Дослідити структуру взаємозв'язку досліджуваних змінних. У цьому разі кожна група змінних визначається фактором, відповідно до якого ці дані є максимально навантаженими.

2. Ідентифікація латентних (прихованих) факторів змінних як чинників взаємозв'язку вихідних даних.

3. Обчислення значень факторів. Водночас кількість факторів значно менша від кількості вихідних змінних.

Рекомендований мінімальний обсяг вибірки, яку можна використовувати для проведення факторного аналізу, по-різному визначається різними авторами, але загальноприйнято, що він повинен бути більше кількості змінних.

У психології використовують різні методи факторного аналізу.

Метод факторного аналізу – це способи визначення факторного складу за відомої кількості чинників. До них належать: метод основних чинників, метод основних компонент, метод максимальної правдоподібності, факторний аналіз образів, метод найменших квадратів.

Аналіз основних компонент визначає спільність кожної змінної через сумування квадратів тридцяти шести навантажень за основними компонентами.

Факторний аналіз образів застосовується до матриці коефіцієнтів кореляції, у якій на основній діагоналі замість одиниць розташовані спільні оцінки. Водночас спільні змінних визначають попередньо як квадрат коефіцієнта множинної кореляції. У цьому разі оцінка надає більш точні результати, ніж при методі основних компонент. Але водночас значення спільностей може недооцінюватися, що зі свого боку приводить до спотворень факторної структури.

Метод основних осей дає змогу обчислити спільності за методом основних компонент спочатку, що надає більш точні результати. Далі факторні навантаження та власні значення обчислюють відповідно до попередніх значень спільностей. Кінцеве рішення отримують за результатами мінімальних відмінностей між спільними ознаками.

Метод незважених найменших квадратів сприяє мінімізації квадратів залишків вихідної та репродукованої матриці коефіцієнтів кореляції. Спочатку оцінюють спільності через квадрат коефіцієнта множинної кореляції, потім визначається факторна структура й оновлюються коефіцієнти кореляції.

Узагальнений метод найменших квадратів передбачає введення спеціальних вагових коефіцієнтів для кожної змінної окремо. За цим методом вплив на факторну структуру залежить від величини спільної змінної.

Метод максимальної правдоподібності направлений на зменшення різниці кореляцій між даними. Цей метод дозволяє статистично оцінити показник повноти факторизації, тобто «якість підгонки».

Факторний аналіз є покроковою процедурою, під час якої на кожному етапі дослідником приймаються рішення про наступне перетвореннях показників. Основною метою водночас залишається можливість отримати змістовну інтерпретацію кінцевих результатів.

Факторний аналіз здійснюється за такими етапами:

1. Обрання вихідної інформації.
2. Попереднє дослідження проблеми кількості чинників.
3. Розкладання матриці коефіцієнтів кореляції.
4. Обертання чинників та їхнє попереднє тлумачення.
5. Визначення якості факторного складу.
6. Розрахунок оцінок факторів і коефіцієнтів.

Залежно від своїх цілей дослідник визначає, скільки разів повторити таку послідовність етапів, які етапи пропустити та наскільки глибоко потрібно опрацювати кожен з них.

6.2. Кластерний аналіз

Багато досліджень мають на меті організувати отримані дані в певну структуру. Наприклад, у біології часто ставлять за мету розділити популяції тварин на види й підвиди; в психології – класифікувати типи поведінки; в освіті – освітні цілі. Кластерний аналіз може допомогти в цьому.

Кластерний аналіз є методом багатомірного дослідження даних. Він передбачає збір даних, які містять інформацію про досліджувані одиниці та зведення їх до однорідних, схожих між собою класів, груп, кластерів. Зазвичай водночас дослідник не володіє вихідними даними ні про відмінності класів, ні про їхній склад. Дослідник має лише інформацію про характеристики чи ознаки об’єктів, на основі яких робить висновок про схожість чи відмінність об’єктів дослідження. У літературі можна зустріти поняття, тотожні кластерному аналізу: *таксономічний аналіз, автоматична класифікація, аналіз образів.*

Варіантами кластерного аналізу є безліч обчислювальних процедур, які використовують під час класифікації досліджуваних об’єктів. Така класифікація передбачає групування їх у класи, водночас об’єкти в кожному класі мають бути більш схожими між собою, ніж інших кластерів. Тобто ***кластерний аналіз*** – це процедура об’єднання досліджуваних об’єктів в однорідні класи за допомогою попарного порівняння їх за попередньо визначеними критеріями.

Етапи кластерного аналізу:

- Проведення дослідження.
- Підготовка показників для проведення кластерного аналізу.
- Вибір методів кластерного аналізу.
- Вибір і розрахунок відстаней між об'єктами.
- Вибір стратегій кластеризації.
- Формування кластерів за обраною стратегією.
- Перевірка й інтерпретація результатів кластерного аналізу.

Існують три основні методи кластерного аналізу: деревоподібна кластеризація, K-середні та кластеризація з двома входами.

Методи деревоподібної кластеризації дозволяють будувати ієрархічні дерева кластерів.

Метод кластеризації K-середніх використовується, коли дослідник уже має певну апріорну гіпотезу про кількість кластерів. У цьому методі дослідник апріорі визначає кількість кластерів, а алгоритм кластеризації дозволяє зробити так, щоб ці кластери максимально відрізнялися між собою. Перевагою методу є можливість перевірити статистичну значущість відмінностей між обраними кластерами.

Метод двовходового об'єднання (two-way joining) використовується, коли потрібно кластеризувати об'єкти (змінні) та спостереження одночасно.

В. Ю. Крилов та Т. В. Острякова рекомендують кілька стратегій кластеризації, які можна досить легко реалізувати вручну. Цікаво також зазначити, що ці стратегії базуються на стадіях психічного розвитку, виділених Л. С. Виготським. Це такі етапи:

- асоціативні комплекси;
- ланцюговий комплекс;
- асоціативно-ланцюгові комплекси;
- комплекси-збірки;
- тощо.

Для створення ***асоціативного кластера*** один елемент обирається як ядро. Потім вибираються всі елементи, найближчі за відстанню до ядра. Ядер буде стільки, скільки елементів у матриці. Для вибору певного елемента як ядра потрібно враховувати характер проблеми, яку намагається вирішити дослідник, і властивості матриці відстаней.

Складні асоціативні кластери відрізняються від простих тим, що мають багатоелементне ядро. Для побудови такого кластера вибирається складне ядро (зазвичай це кілька елементів, розташованих близько один до одного), а потім для кожного елемента вибирається його найближчий сусід. Потім вибирається певна мінімальна відстань, і в кластері залишаються тільки елементи з цією мінімальною відстанню від ядра.

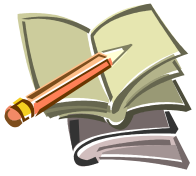
Для побудови ***ланцюгового кластера*** з матриці вибирається один елемент (№ 1), шукається наступний найближчий елемент (№ 2) і так повторюється. Побудова ланцюгового кластера завершується, коли всі найближчі елементи вилучені.

Побудова асоціативних і ланцюгових кластерів має різні цілі. Метою асоціативної кластеризації є знадення елементу, найближчого до ядра, тоді як метою ланцюгової кластеризації є знайдення послідовного зв'язку між початковим елементом і всіма іншими матричними елементами.

Наприклад, використання ланцюгових кластерів є дослідженням циркуляції чуток у приватних компаніях. Для цього потрібно створити матрицю, де кількість щоденних контактів між працівниками в організації є мірою близькості. Потім для кожного працівника можна створити ланцюгові кластери, щоб визначити найбільш імовірний спосіб поширення чуток.

Однак цікаво розробляти стратегії, які поєднують сильні сторони як асоціативних, так і ланцюгових кластерів. Це ***асоціативно-ланцюговий*** кластер. Щоб його побудувати, спочатку використовують стратегію створення асоціативного кластера, а потім – ланцюгового кластера.

Приклади розв'язку типових задач



1. Встановити, чи можна виділити тривогу й депресію, одержувані за результатами самозвіту, в окремі чинники. Протестовано досліджуваних трьома короткими питаннями щодо тривоги (Anxiety, A1 – A3) і трьома – щодо депресії (Depression, Dl – D3):

A1 Я відчуваю тривогу;

A2 Я стаю напруженим;

A3 Я спокійний;

D1 Я пригнічений;

D2 Я відчуваю себе марним;

D3 Я щасливий.

Відповіді на кожне з цих питань даються за 5-бальною шкалою, де 1 означає «ніколи», 2 – «іноді», 3 – «часто», 4 – «значну частину часу» і 5 – «завжди».

Таблиця 6.1 – Результати тестування досліджуваних

Спостереження	A1 Тривожний	A2 Напружений	A3 С покійний	D1 Подавлений (пригнічений)	D2 Марний	D3 Щасливий
1	2	1	3	1	2	5
2	1	2	3	4	3	3
3	3	3	4	2	1	4
4	4	4	3	3	2	3
5	5	5	2	3	4	4
6	4	5	2	4	3	1
7	4	3	2	5	4	1
8	3	3	4	4	4	3
9	3	5	3	3	4	1

Розв'язання:

Першим кроком під час здійснення факторного аналізу є створення кореляційної матриці, в якій містяться коефіцієнти кореляції всіх змінних один з одним.

Таблиця 6.2 – Кореляційна матриця

	A1 Тривожний	A2 Напружений	A3 Спокійний	D1 Подавлений (Пригнічений)	D2 Марний	D3 Щасливий
A1 Тривожний	1					
A2 Напружений	0,73849625	1				
A3 Спокійний	–0,5026155	–0,399252103	1			
D1 Подавлений (пригнічений)	0,22115385	0,300267706	–0,36957025	1		
D2 Марний	0,27907815	0,392567942	–0,42905817	0,651182357	1	
D3 Щасливий	–0,2496346	–0,539894472	0,40777823	–0,741102694	–0,52831	1

Найбільші за абсолютним значенням коефіцієнти кореляції дорівнюють 0,74 (кореляція між «Тривожний» і «Напружений») і –0,74 (кореляція між «Пригнічений» і «Щасливий»). Наступна за абсолютним значенням кореляція між «Пригнічений» і «Даремний» дорівнює 0,65. Найменший за абсолютним значенням коефіцієнт кореляції між «Тривожний» і «Пригнічений» дорівнює 0,22.

Величина спільної дисперсії для двох змінних є квадрат коефіцієнта кореляції цих змінних. Так, величина спільної дисперсії, що пояснюється тим загальним, що присутній і у відчутті тривожності, й у відчутті напруженості, дорівнює 0,742, або наближено 0,55, тоді як величина загальної дисперсії між тривожністю і пригніченістю дорівнює 0,222, або близько 0,05. Інакше кажучи, величина спільної дисперсії між тривожністю та напруженістю в 11 разів більша величини спільної дисперсії між тривожністю й пригніченістю. Максимальне значення величини спільної дисперсії одно 1,00 (± 1,002 = 1,00), а мінімальне – 0,00 (0,002 = 0,00).

Якщо подивитися на значення коефіцієнтів кореляції, можна побачити, що є деяка тенденція, яка виявляється в тому, що пункти опитувальника, пов'язані з тривогою, більш тісно пов'язані між собою, ніж з даними, пов'язаними з депресією, а останні зі свого боку сильніше корелюють між собою, ніж з пунктами, пов'язаними з тривогою. Отже, можна припустити існування двох окремих груп для пунктів, пов'язаних з тривогою й депресією. Наприклад, коефіцієнт кореляції між відчуттям неспокою та напруженістю дорівнює 0,74, тоді як коефіцієнт кореляції між відчуттям неспокою та пригніченістю дорівнює всього лише 0,22. Проте це є не зовсім зрозумілим, оскільки абсолютне значення коефіцієнта кореляції між «Напружений» і «Щасливий», що входять до групи депресивних пунктів, вище (–0,54), ніж абсолютне значення коефіцієнта кореляції між «Напружений» і «Спокійний» (–0,40).

Для дослідження використаємо метод основних компонент.

У методі основних компонент величина дисперсії дорівнює числу змінних, оскільки дисперсія чи спільність кожної змінної приймається рівною 1,00. Отже, в разі шести змінних сумарна (загальна) дисперсія дорівнює 6,00. Число освічених, або видобутих, компонент математично завжди дорівнює числу змінних, що беруть участь в аналізі. Так, у разі шести змінних буде витягнуто шість чинників.

Щоб обчислити частку загальної дисперсії, що пояснюється кожним чинником, треба підсумувати квадрати навантажень. У результаті отримуємо значення чинника, яке ділимо на число змінних.

За результатами факторного аналізу з використанням програми PSPP маємо таблиці шести основних компонент для набору даних із шести змінних та обчислених квадратів навантажень.

Таблиця 6.3 – Таблиця основних компонент

	1	2	3	4	5	6
A1 Тривожний	0,66	0,67	0,03	0,11	0,29	–0,14
A2 Напружений	0,76	0,46	0,37	0,02	–0,20	0,18
A3 Спокійний	–0,69	–4,20	0,64	0,25	0,10	–0,05
D1 Подавлений	0,76	–0,53	0,05	–0,04	0,35	0,14
D2 Марний	0,75	–0,34	–0,15	0,52	–0,17	–0,06
D3 Щасливий	–0,80	0,34	–0,27	0,35	0,14	0,17

За результатами розрахунків, усі шість змінних найсильніше корелюють з першим чинником, за винятком відчуття тривоги, яке трохи сильніше корелює з другим чинником.

Таблиця 6.4 – Частка загальної дисперсії

	1	2	3	4	5	6	Спільності
A1 Тривожний	0,43	0,45	0,00	0,01	0,08	0,02	1,00
A2 Напружений	0,58	0,21	0,14	0,00	0,04	0,03	1,00
A3 Спокійний	0,48	0,04	0,41	0,06	0,01	0,00	1,00
D1 Подавлений	0,57	0,28	0,00	0,00	0,12	0,02	1,00
D2 Марний	0,56	0,11	0,02	0,27	0,03	0,00	1,00
D3 Щасливий	0,64	0,12	0,07	0,02	0,02	0,03	1,00
Власні значення	**3,26**	**1,21**	**0,64**	**0,47**	**0,31**	**0,11**	**6,00**
Доля поясненої дисперсії	*0,54*	*0,20*	*0,11*	*0,08*	*0,05*	*0,02*	

Таблиця відображує результат, за яким перша основна компонента (чинник) пояснює 54 % загальної дисперсії шести змінних, другий чинник – ще 20 % загальної дисперсії тощо.

Власні значення головних компонент відобразимо графічно:

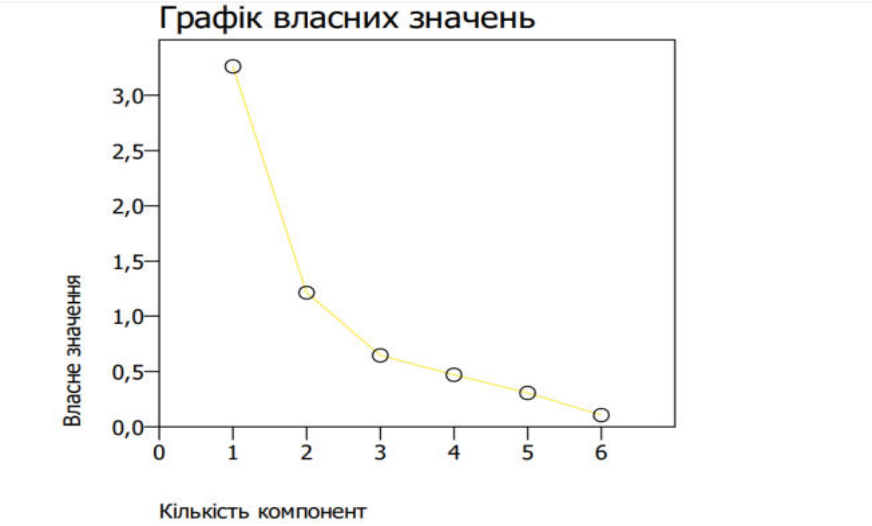


Рисунок 6.1 – Графік власних значень
для шести основних компонент

Використовуючи метод варімакс, розрахуємо обертання двох виділених основних компонент:

Таблиця 6.5 – Перші дві основні компоненти
після обертання за методом варімакс

	1	2
A1 Тривожний	0,05	0,94
A2 Напружений	0,27	0,85’
A3 Спокійний	–0,39	–0,60
D1 Подавлений	0,92	0,10
D2 Марний	0,79	0,24
D3 Щасливий	–0,83	–0,27

Перший з факторів після обертання варімакс можна інтерпретувати як чинник депресії, оскільки всі три відповідних пункти мають з цього чинника навантаження ± 0,79 і вище, тоді як усі три питання, які стосуються тривожності, мають навантаження ± 0,39 і нижче. Другий фактор після обертання варімакс можна інтерпретувати як чинник тривожності, оскільки пункти, відповідні тривозі, мають з цього чинника навантаження.

Частка загальної дисперсії, яка пояснюється кожним з двох виділених факторів, після обертання за методом варімакс дорівнює їхнім власним значенням або сумі квадратів факторних навантажень для кожного фактора, поділеній на число змінних.

Таблиця 6.6 – Частки загальної дисперсії,
що пояснюється першими двома компонентами
після обертання за методом варімакс

	1	2
A1 Тривожний	0,00	0,88
A2 Напружений	0,07	0,72
A3 Спокійний	0,15	0,36
D1 Подавлений	0,84	0,01
D2 Марний	0,62	0,06
D3 Щасливий	0,69	0,07
Власні значення	2,37	2,10
Частка поясненої дисперсії	0,40	0,35

Частка загальної дисперсії, що пояснюється першим фактором, дорівнює 0,40, другим фактором – 0,35. Ці частки дисперсії відрізняються від відповідних величин для початкових, що не піддавалися обертанню, основних компонент, оскільки в результаті обертання змінилися навантаження змінних на ці фактори.

2. Провести дослідження позитивного ставлення здобувачів вищої освіти психологічного факультету до здобувачів з інших факультетів і до обраних професій, об’єднати здобувачів у групи зі схожими поглядами. З цією метою розподілено факультети та можливі спеціалізації для випускників-психологів. Здобувачам запропоновано визначити своє відношення до всіх за 10-бальною шкалою.

Таблиця 6.7 – Результати опитування здобувачів вищої освіти

		Об’єкти (досліджувані здобувачі)										
		О. А.	В. О.	П. М.	П. К.	А. Р.	В. К.	Д. В.	Р. П.	К. Д.	О. Е.	А. З.
Спостереження (результати оцінювання)	Здобувач психологічного факультету	5	9	9	5	5	10	3	5	1,5	1,5	1,5
	Здобувач педагогічного факультету	2	0,5	3	1	2	3	4	2	9	9	9
	Здобувач фізико-математичного факультету	3	1	2	4	3	4	6	3,5	10	10	10
	Здобувач природничого факультету	9	2	1,5	3	4	2	7	7	9,5	9,5	9,5
	Здобувач історичного факультету	4	3	2,5	3,5	6	5	9	9	7	7	7
	Здобувач філологічного факультету	10	4	1	2	1	6	1,5	1	8	8	8
	Психолог	4,5	9,5	10	10	10	9	2,5	5,5	2	2	2
	Соціальний педагог	5,5	10	8	9,5	9,5	9,5	3,5	4	1	1	1
	Вчитель	8	5	5	7	5	7	5	3	6	6	6
	Викладач	6	6	7	9	8	5,5	8	8	5	5	5
	Керівник	3,5	7	6	8	7	4,5	10	10	3	3	3
	Консультант	7	8	4	6	6	8	2	6	4	4	4

Розв’язання:

Під час виконання кластерного аналізу потрібно врахувати, що кластеризація традиційно здійснюється за об’єктами (стовпчиками масивів даних). Зупинимося на методі деревовидної кластеризації. Виберемо Евклідову відстань і обчислимо її. Розрахуємо відстань між першим і другим об’єктами (суб’єктами О. А. та В. О. вихідного набору даних). Для цього використаємо формулу:

$$L = \sqrt{\sum(x_i - y_i)^2}.$$

Результати обчислення: $\sum(x_i - y_i)^2$ = 175,75; L = 13,26.

Після визначення дистанції між усіма об'єктами створюється таблиця відстаней (distances matrix).

Таблиця 6.8 – Таблиця відстаней між усіма об'єктами

	О. А.	В.О.	П. М.	П. К.	А. Р.	В. К.	Д. В.	Р. П.	К. Д.	О. Е.	А. З.
О. А	0,00	13,26	14,81	1 13,41	13,51	11,51	14,37	13,63	12,82	12,82	12,82
В. О.	13,26	0,00	**6,24**	7,00	7,26	5,94	16,76	12,94	**21,66**	**21,66**	**21,66**
П. М.	14,81	6,24	0,00	6,82	6,69	8,02	14,94	12,37	**20,41**	**20,41**	**20,41**
П. К.	13,41	7,00	6,82	0,00	4,03	8,82	13,44	10,95	19,62	19,62	19,62
А. Р.	13,51	7,26	6,69	4,03	0,00	8,86	12,35	9,04	18,89	18,89	18,89
В. К.	11,51	5,94	8,02	8,82	8,86	0,00	16,43	13,83	18,69	18,69	18,69
Д. В.	14,37	16,76	14,94	13,44 1	12,35	16,43	0,00	6,61	12,85	12,85	12,85
Р. П.	13,63	12,94	12,37	10,95	9,04	13,83	6,61	0,00	15,97	15,97	15,97
К. Д.	12,82	21,66	20,41	19,62	18,89	18,69	12,85	15,97	0,00	0,00	0,00
О. Е.	12,82	21,66	20,41	19,62	18,89	18,69	12,85	15,97	0,00	0,00	0,00
А. З.	12,82	21,66	20,41	19,62	18,89	18,69	12,85	15,97	0,00	0,00	0,00

Найбільша дистанція існує між здобувачами В. О. та К. Д., О. Е., А. З. (L = 21,66), а також між П. М. та К. Д., О. Е., А. З. (L = 20,41). У разі такого аналізу можна дійти висновку, що існує найменше дві групи здобувачів з різним відношенням до навчання й обраної професії – група В. О., П. М. та група К. Д., О. Е., А. З. Розглянемо відстань усередині кожної групи. Насправді виявилися досить близько один до одного здобувачі В. О. та П. М. (L = 6,24), ідентичні один одному (на нульовій відстані) здобувачі К. Д., О. Е. та А. З. (L = 0).

До даних таблиці відстаней можна застосувати стратегію найближчого сусіда.

Найменші відстані в таблиці вказані між досліджуваними К. Д., О. Е., А. З. (L = 0). Зрозуміло, що вони увійшли до першого кластера. Другий кластер з найбільшою відстанню – це П. К. і А. Р. (L = 4,03). Їх об'єднаємо в другий кластер.

Третій кластер формують здобувачі В. О. й В. К. (L = 5,94). Відстань від В. О. дорівнює L = 6,25, тому до третього кластера входить П. М. Отже, четвертий кластер матиме власну структуру. Це первинний кластер, що містить об'єкти В. О. і В. К. і вторинний

кластер, що містить об’єкт П. М. П’ятий кластер утворять об’єкти Д. В. і Р. П. (L = 6,61).

Усі об’єкти увійшли до кластерів, тому тепер потрібно пов’язати між собою всі отримані кластери. Наступна за величиною відстань – L = 6,69.Така відстань спостерігається між П. М. (четвертим кластером) і А. Р. (другим кластером). Відповідно, другий і третій кластери варто об’єднати в шостий кластер. Далі потрібно знайти відстань від п’ятого до шостого кластера. До неї належить L = 9,04 між здобувачами Р. П. і А. Р., що становитиме сьомий кластер. Подальший аналіз відстаней показує, що об’єкти О. А. та В. К. додаються до сьомого кластера з найменшою дистанцією L = 11,51. Так утворюється восьмий кластер. Дев’ятий кластер утворено за допомогою з’єднання першого та восьмого кластерів на відстані L = 12,81 (здобувачі К. Д., О. Е., А. З. та О. А.).

Описані вище процеси кластеризації варто підсумувати як зведену таблицю (amalgamation schedule).

Таблиця 6.9 – Таблиця об’єднань

Крок	Відстань до найближчого сусіда	Об’єкти										
		1	2	3	4	5	6	7	8	9	10	11
1	0,00	К. Д.	О. Е.									
2	0,00	К. Д.	О. Е.	А. З.								
3	4,03	П. К.	А. Р.									
4	5,94	В. О.	В. К.									
5	6,25	В. О.	В. К.	П. М.								
6	6,61	Д. В.	П. Р.									
7	6,69	В. О.	В. К.	П. М.	П. К.	А. Р.						
8	9,04	В. О.	В. К.	П. М.	П. К.	А. Р.	Д. В.	П. Р.				
9	11,51	О. А.	В. О.	В. К.	П. М.	П. К.	А. Р.	Д. В.	П. Р.			
10	12,82	О. А.	В. О.	В. К.	П. М.	П. К.	А. Р.	Д. В.	П. Р.	К. Д.	О. Е.	А. З.

Можна створити дерево кластерів з графічним представленням отриманих кластерів (hierarchical tree plot).

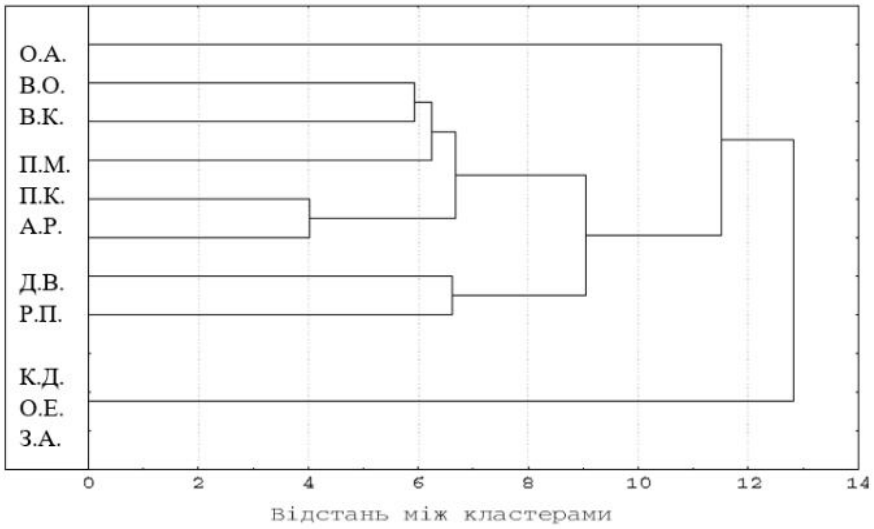


Рисунок 6.2 – Кластерне дерево

Отже, до першого кластера увійшли здобувачі: К. Д., О. Е., А. З. Таблиця вихідних даних показує, що це здобувачі найбільш позитивно ставляться до здобувачів з інших факультетів, але водночас менш позитивно – до своїх майбутніх рольових позицій. І навпаки, здобувачі з другого кластера (П. К. та А. Р.) більш позитивно ставляться до своєї майбутньої професійної діяльності та менш позитивно – до студентської спільноти. Здобувачі третього кластера (В. О., В. К., П. М.) позитивно ставляться до професії психолога та до здобувачів своєї спеціальності. Четвертий кластер здобувачів (Д. В. та Р. П.) має найбільш позитивне ставлення до професії вчителя та здобувачів такої самої орієнтації. Досить складно оцінювати ставлення О. А., спостерігається професійна невизначеність.

Отже, результати аналізу свідчать про те, що є підстави для класифікації здобувачів як таких, що професійно ідентифікуються. Здобувачів, які не сформували професійну ідентичність, поділяють на два класи: тих, хто найбільш позитивно ставиться до свого життя (К. Д., О. Е., А. З.), і тих, хто позитивно ставиться до майбутньої професійної діяльності без розмежування на види (П. К., А. Р.). Здобувачів, у яких сформована професійна ідентифікація, можна поділити на дві групи: тих, хто орієнтується на психологічну діяльність (В. О., В. К., П. М.) та тих, хто на освітню (Д. В., Р. П.)

Таблиця 6.10 – Розподіл на кластери здобувачів вищої освіти за сформованістю професійної ідентифікації

Несформована професійна ідентифікація	Позитивно ставляться до студентського життя	К. Д., О. Е., А. З.
	Позитивно ставляться до майбутньої професії, не поділяючи її на види	П. К., А. Р.
Сформована професійна ідентифікація	Орієнтовані на психолого-педагогічну діяльність	В. О., В. К., П. М.
	Орієнтовані на керівну діяльність	Д. В., Р. П.

Перелік питань і завдань для самоконтролю



1. Особливості факторного аналізу.
2. Місце факторного аналізу в структурі експерименту.
3. Мета та методи кластерного аналізу.
4. Підготовка показників до проведення кластерного аналізу.
5. Деревоподібна кластеризація: розташування відстаней між об’єктами.
6. Кластеризація деревоподібна: стратегії та інтерпретація.
7. Результати кластерного аналізу.
8. Провести дослідження на основі даних про те, як молоді люди сприймають своїх близьких.

Таблиця 6.11 – Матриця якості «Значимі особи»

Якість	Я	Мати	Батько	Брат	Сестра	Кохана людина
Досягнення	2	5	6	1	3	4
Мудрість	5	3	6	2	4	1
Чіткість планів	4	5	3	1	6	2
Пошук	2	4	5	1	3	6
Рух	3	4	5	2	1	6
Досягнення вищого	1	4	6	2	3	5
Визначеність	5	4	3	1	6	2
Хитрість	5	4	6	2	3	1

9. Дослідити сприймання здобувачами вищої освіти першого курсу психологічних теорій після вивчення курсу «Теорії особистості», об’єднати здобувачів у певні групи за схожим ставлення.

Таблиця 6.12 – Результати опитування здобувачів вищої освіти

	А. О.	О. В.	М. П.	К. П.	Р. А.	К. В.	В. Д.	П. Р.	Д. К.	Е. О.	З. А.
З. Фройд	2	5	4	6	8	2	5	5	8	5	9
К. Юнг	3	2	6	4	5	1	8	9	6	7	3
А. Адлер	1	4	2	5	7	3	3	8	6	7	9
Дж. Уотсон	4	2	6	1	7	5	8	9	10	3	7
Б. Скіннер	10	7	3	4	9	5	6	8	2	9	9
К. Халл	3	2	4	5	1	6	7	2	3	9	9
А. Маслоу	5	3	6	1	3	8	5	6	7	4	7
К. Роджерс	4	2	7	3	10	8	9	5	10	8	9
Р. Мей	5	2	6	8	2	8	9	10	6	7	8

Тести для самоконтролю

1. Факторний аналіз використовується для:

а) виявлення факторів впливу;

б) зменшення розмірності масиву даних;

в) факторного диференціювання інтегральних ознак;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

2. Під час підготовки даних до факторного аналізу бажаним є, щоб кількість стовпчиків:

а) була меншою за кількість рядків;

б) була більшою за кількість рядків;

в) дорівнювала кількості рядків;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

3. Факторні ваги – це:

а) коефіцієнти кореляції кожної із аналізованих змінних з кожним з виділених чинників;

б) кількісні показники зв'язку виділених чинників з об'єктами;

в) дисперсії, які обумовлюються чинниками;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

4. Для оцінки значимості фактора використовують:

а) факторні навантаження;

б) факторні ваги;

в) власні значення факторів;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

5. Осмислена інтерпретація результатів кластерного аналізу можлива за умови, якщо вона ґрунтується:

а) на теоретичній моделі;

б) багатьох повторних дослідженнях;

в) використанні методів перевірки статистичної значимості;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

6. Не існує методу кластерного аналізу:

а) деревоподібної кластеризації;

б) двовходового об'єднання;

в) імовірнісних кластерів;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

7. Для пошуку відстаней між об'єктами в кластерному аналізі використовують:

а) відстань Спірмена;

б) відстань Йокогами;

в) відстань Чебишева;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

8. Під час опису результатів кластерного аналізу в статті в її текст потрібно внести:

а) кластерні дерева;

б) кластерні дерева та таблиці об’єднань;

в) кластерні дерева, таблиці об’єднань і таблиці відстаней;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

9. Факторний аналіз використовується для:

а) пошуку прихованих змінних;

б) виявлення чинників впливу;

в) факторного диференціювання інтегральних ознак;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

10. Вимогою до підготовки даних до факторного аналізу є представлення масиву даних у вигляді:

а) двохвимірної матриці;

б) N-вимірної матриці;

в) будь-як;

г) правильна відповідь відсутня.

Відповідь: 1) а; 2) б; 3) в; 4) г.

ТЕМА 7
МЕТОДИ ДОСЛІДЖЕННЯ Й АНАЛІЗУВАННЯ ПСИХОЛОГІЧНИХ ПРОЦЕСІВ У ДИНАМІЦІ ТА ТЕНДЕНЦІЇ РОЗВИТКУ



7.1. Методи дослідження й аналізування психологічних процесів у динаміці.

7.2. Методи аналізування тенденції розвитку в психологічних дослідженнях.

7.1. Методи дослідження й аналізування психологічних процесів у динаміці

Психологічні явища та процеси з часом змінюються.

Процес розвитку психологічних явищ у часі називають динамікою. Для її відображення будують ряди динаміки.

Рядом динаміки називається ряд чисел, які характеризують зміну величини психологічного явища в часі.

Розрізняють ряди динаміки *абсолютних, середніх* та *відносних величин*.

Ряди динаміки абсолютних величин бувають *моментними* та *періодичними*.

Кожний ряд динаміки можна охарактеризувати за допомогою таких показників: *абсолютний приріст, коефіцієнт зростання, темп зростання, темп приросту, абсолютне значення одного відсотка приросту, середній рівень ряду динаміки, середній абсолютний приріст, середній темп приросту* та *середній темп зростання.*

Абсолютний приріст визначається різницею між поточним (звітним) і базисним рівнями.

Базисний

$$A = У_n - У_o \quad (7.1)$$

Ланцюговий

$$A = У_n - У_{n-1}, \quad (7.2)$$

де A – абсолютний приріст;

$У_n$ – поточний (звітний) рівень;

$У_0$ – базисний рівень;

$У_{n-1}$ – попередній рівень.

Коефіцієнт зростання обчислюється як відношення рівня досліджуваного періоду ($У_п$) до рівня, з яким порівнюють.

Базисний

$$К_р = \frac{У_п}{У_о} \quad (7.3)$$

Ланцюговий

$$К_р = \frac{У_п}{У_{п-1}} \quad (7.4)$$

Темп зростання обчислюється як процентне відношення рівня досліджуваного періоду ($У_п$) до рівня, з яким порівнюють.

Базисний

$$Т_р = \frac{У_п}{У_0}100 \quad (7.5)$$

Ланцюговий

$$Т_р = \frac{У_п}{У_{п-1}}100 \quad (7.6)$$

Не важко помітити, що $Т_р = К_р \times 100$ (7.7)

Темп приросту обчислюється за допомогою ділення абсолютного приросту, помноженого на 100, на величину рівня, з яким порівнюють.

Обчислюється також базисним і ланцюговим способами.

Базисний

$$Т_{пр} = \frac{У_п - У_0}{У_0}100 \quad (7.8)$$

Ланцюговий

$$Т_{пр} = \frac{У_п - У_{п-1}}{У_{п-1}}100 \quad (7.9)$$

Темп приросту можна обчислити за допомогою віднімання 100 від значення темпів зростання:

$$Т_{пр} = Т_р - 100. \quad (7.10)$$

Абсолютне значення одного відсотка приросту обчислюється за допомогою ділення абсолютного приросту на темп приросту.

$$А\,\% = \frac{У_п - У_{п-1}}{Т_{пр}}. \quad (7.11)$$

Для узагальнення динаміки досліджуваного явища за декілька періодів визначають різні середні показники. До них належать: *середні рівні ряду* та *середні показники змін рівнів ряду*.

Середні рівні загалом використовують для узагальнення рядів, що коливаються, для забезпечення порівняності чисельника та знаменника, а також для побудови динамічних рядів похідних показників.

Метод розрахунку середнього рівня динамічного ряду залежить від виду динамічного ряду.

Для *інтервальних рядів* абсолютних значень з однаковою тривалістю використовується:

$$\bar{y} = \frac{\sum x}{n} \ , \tag{7.12}$$

де n – кількість рівнів ряду;

x – значення рівнів ряду.

У *моментному ряді,* де показники змінюються рівномірно між датами (рівні рівновіддалені), середнє значення між двома датами розраховується як сума половин значень на початок і кінець періоду (середня арифметична проста):

$$\bar{y} = \frac{y_0 + y_n}{2}; \tag{7.13}$$

у разі різних проміжків часу – середня арифметична зважена.

Якщо в моментному ряді динаміки інтервал між датами однаковий, то середній рівень розраховується за формулою середньої хронологічної:

$$\bar{y} = \frac{(\frac{y_1}{2} + y_2 + y_3 + ... + y_{n-1} + \frac{y_n}{2})}{n-1} . \tag{7.14}$$

У моментних та інтервальних рядах динаміки з нерівномірно розподіленими періодами (інтервалами) для визначення середнього рівня ряду використовують середню арифметичну зважену:

$$\bar{y} = \frac{\sum yt}{\sum t} , \tag{7.15}$$

де yt – рівні ряду;

t – проміжки часу між суміжними датами або періоди часу.

Середній абсолютний приріст обчислюється за допомогою середньої арифметичної простої з ланцюгових абсолютних приростів.

$$\overline{A} = \frac{\sum A}{n}, \tag{7.16}$$

де A – величини ланцюгових приростів;

n – число ланцюгових приростів.

За відсутності ланцюгових приростів середній абсолютний приріст можна обчислити за формулою:

$$\overline{A} = \frac{Y_n - Y_0}{n-1}, \tag{7.17}$$

де n – число календарних дат.

Середні темпи (коефіцієнти) зростання обчислюються за формулами середньої геометричної:

$$\overline{K}_p = \sqrt[n]{K_1 K_2 ..... K_n} = \sqrt[n]{ПК}, \tag{7.18}$$

де $\overline{K}$ – середній коефіцієнт зростання;

K – ланцюгові коефіцієнти зростання;

n – число ланцюгових коефіцієнтів;

$П$ – знак добутку, який свідчить про те, що дані, які стоять після нього, потрібно перемножити та знайти їхній добуток.

Якщо відсутні ланцюгові коефіцієнти, то можна скористатися іншою формулою:

$$\overline{K} = \sqrt[n-1]{\frac{Y_n}{Y_0}}, \tag{7.19}$$

де n – число календарних дат;

Y_n і Y_0 – рівні щодо звітного й базисного періодів.

Одержані коефіцієнти перетворюють у темпи зростання за формулою $T_{пр} = K_p \times 100$ і в темпи приросту за формулою $T_{пр} = K_p \times 100 - 100$.

Якщо швидкість розвитку явища в межах розглядуваного періоду неоднакова, то зіставленням однойменних характеристик за різні інтервали часу визначають **прискорення** чи **сповільнення динаміки**. Різниця абсолютних ланцюгових приростів $\Delta_t - \Delta_{t-1}$ характеризує **абсолютне прискорення** (+) чи **уповільнення** (–). Зіставленням темпів зростання (приросту) дістають **відносне прискорення (уповільнення).**

7.2. Методи аналізування тенденції розвитку в психологічних дослідженнях

У процесі аналізу рядів динаміки важливо виявити загальну тенденцію розвитку психологічного явища або процесу, тобто встановити, чи змінюється воно у висхідному або низхідному напрямі.

Тенденція – це довгострокова еволюція з порівняно плавною траєкторією.

Серед методів опису трендів особливо широко використовуються *трендові криві*.

Вибір функції залежить від характеру динаміки:

– у разі порівняно стабільних абсолютних приростів обирають лінійний тренд:

$$Y_t = a + bt; \qquad (7.20)$$

– якщо темпи зростання є стабільними, обирається експоненціальна функція:

$$Y_t = ab^t \qquad (7.21)$$

У лінійному рівнянні параметр *b* характеризує середній абсолютний приріст, в експоненті – середній темп зростання. Параметр *a* в обох функціях є теоретичним значенням рівня під час t = 0. У лінійному рівнянні параметр *b* характеризує швидкість динаміки. У лінійній функції – абсолютна середня (середній абсолютний приріст), а в експоненті – відносна середня (середній темп зростання). Коли характеристики швидкості розвитку зростають (чи зменшуються), використовують інші функції (парабола 2-го ступеня, модифікована експонента тощо).

Визначають параметри трендових кривих, розв'язуючи системи нормальних рівнянь. Для лінійної функції вона має такий вигляд:

$$\begin{cases} an + b\sum t = \sum y \\ a\sum t + b\sum t^2 = \sum yt \end{cases}. \qquad (7.22)$$

Параметри *a* і *b* лінійного рівняння легше обчислити за допомогою методу відліку від умовного нуля. Умовний нуль – це значення *t* в середині ряду. Тоді значення *t*, розміщені вище середини, будуть від'ємними, а нижче – додатними. У разі непарного числа членів ряду (наприклад, $n = 5$) змінній *t* надаються значення з інтервалом одиниця: -2, –1, 0, 1, 2; у разі парного (наприклад, $n = 6$) – з інтервалом два: –5, –3, –1, 1, 3, 5. В обох випадках $\sum t = 0$, а система рівнянь набуває:

$$\begin{cases} na = \sum y \\ b\sum t^2 = \sum yt \end{cases} \tag{7.23}$$

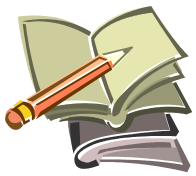
$$a = \frac{\sum_1^n y_i}{n}. \tag{7.24}$$

Тоді

$$b = \frac{\sum_1^n y_i t_i}{\sum_1^n t^2}. \tag{7.25}$$

Під час прогнозування економічних показників використовують метод ***екстраполяції,*** в разі якого значення розраховують на рівнях, що виходять за межі наявних фактичних даних. Екстраполяція ґрунтується на припущенні, що виявлена тенденція збережеться в майбутньому. Щоб зробити прогноз, варто підставити в рівняння тренду значення, потрібні для продовження вихідного ряду, й обчислити *y*.

Приклади розв’язку типових задач



1. За наведеними даними про психотичні розлади визначити показники динаміки.

Таблиця 7.1 – Психотичні розлади за 2019–2023 рр.

Рік	Психотичні розлади
2019	23
2020	21
2021	28
2022	34
2023	34
Разом	140

Розв’язання:

Абсолютний приріст (зменшення)

Базисний

$\Delta_t = y_t - y_{t-1}$

2020 р.: 21 – 23 = –2
2021 р.: 28 – 23 = 5
2022 р.: 34 – 23 = 11
2023 р.: 34 – 23 = 11

Ланцюговий

$\Delta_t = y_t - y_0$

2020 р.: 21 – 23 = –2
2021 р.: 28 – 21 = 7
2022 р.: 34 – 28 = 6
2023 р.: 34 – 34 = 0

Коефіцієнт зростання

Базисний

$K_t = \frac{y_t}{y_{t-1}}$

2020 р.: $\frac{21}{23} = 0{,}91$
2021 р.: $\frac{28}{23} = 1{,}22$
2022 р.: $\frac{34}{23} = 1{,}48$
2023 р.: $\frac{34}{23}1{,}48$

Ланцюговий

$K_t = \frac{y_t}{y_0}$

2020р.: $\frac{21}{23} = 0{,}91$
2021 р.: $\frac{28}{21} = 1{,}33$
2022 р.: $\frac{34}{28} = 1{,}21$
2023 р.: $\frac{34}{34} = 1$

Темп зростання

$$\text{Т}_{\text{зр.}} = \text{К}_t \cdot 100\ \%$$

Базисний:

2020 р.: 0,91 × 100 % = 91 %
2021 р.: 1,22 × 100 % = 122 %
2022 р.: 1,48 × 100 % = 148 %
2023 р.: 1,48 × 100 % = 148 %

Ланцюговий

2020 р.: 0,91 × 100 % = 91 %
2021 р.: 1,33 × 100 % = 133 %
2022 р.: 1,21 × 100 % = 121 %
2023 р.: 1 × 100 % = 100 %

Темп приросту

$$T_t = \frac{100 \bullet \Delta_t}{y_{t-1}} = \frac{100(y_t - y_{t-1})}{y_{t-1}} = 100(K_t - 1)$$

Базисний

2020 р.: 91 – 100 % = –9 %
2021 р.: 122 – 100 % = 22 %
2022 р.: 148 – 100 % = 48 %
2023 р.: 148 – 100 % = 48 %

Ланцюговий

2020 р.: 91 – 100 % = 9 %
2021 р.: 133 – 100 % = 33 %
2022 р.: 121 – 100 % = 21 %
2023 р.: 100 – 100 % = 0 %

Абсолютне значення 1 % приросту

$$A\% = \frac{\Delta_t}{T_t} = \frac{y_{i-1}}{100} = 0{,}01 y_{i-1}$$

Ланцюговий

2020 р.: $\frac{21}{100\%} = 0{,}21$
2021 р.: $\frac{28}{100\%} = 0{,}28$
2022 р.: $\frac{34}{100\%} = 0{,}34$
2023 р.: $\frac{34}{100\%} = 0{,}34$

Середній абсолютний приріст

$$\bar{\Delta} = \frac{\sum_1^n \Delta_t}{n} = \frac{1}{n}\sum(y_t - y_{t-1})$$

Базисний

$\bar{\Delta} = \frac{-2+5+11+11}{4} = 6{,}25$

Ланцюговий

$\bar{\Delta} = \frac{-2+7+6+0}{4} = 2{,}75$

Середній коефіцієнт (темп) зростання:

$$\bar{K} = \sqrt[n]{K_1 K_{2'''} K_n} = \sqrt[n]{\frac{y_n}{y_0}}$$

Базисний

$\sqrt[4]{0{,}91 \times 1{,}22 \times 1{,}48 \times 1{,}48} = 1{,}25$

Ланцюговий

$\sqrt[4]{0{,}91 \times 1{,}33 \times 1{,}21 \times 1} = 1{,}10$

Середній темп приросту

$$\bar{T} = \bar{K} \times 100\ \% - 100\ \%$$

Базисний

1,25 × 100 % – 100 % = 25 %

Ланцюговий

1,10 × 100 % – 100 % = 10 %

Висновок. Найбільший приріст психотичних розладів, порівняно з 2019, відбувся 2022 та 2023 рр.: на 11 розладів, тобто на 48 %. Порівняно з 2021 р.: на 7 розладів, тобто 33 %.

У середньому, порівняно з 2019 роком, психотичних розладів збільшилося майже на 6 випадків, або на 25 %.

Найбільший спад психотичних розладів, порівняно з 2019 р. і попереднім періодом, був 2020 р.: на 2 розлади, тобто 9 %. 2023 р., порівняно з попереднім роком, зміни в психотичних розладах не спостерігались.

У середньому, порівняно з попереднім періодом, психотичних розладів збільшилося майже на 3 випадки, або 10 %.

2. За даними про психічні розлади в судово-психіатричній експертизі в цивільних справах визначити тенденцію розвитку та спрогнозувати показники на найближчі періоди.

Таблиця 7.2 – Динаміка алкогольних психічних розладів у судово-психіатричній експертизі в цивільних справах

Рік	Психотичні розлади, осіб
1	2
2013	5
2014	6
2015	11
2016	12
2017	17

Продовження таблиці 7.2

1	2
2018	17
2019	23
2020	21
2021	28
2022	34
2023	34
Разом	208

Розв'язання:

Використовуючи Майстер діаграм Microsoft Excel, побудуємо графік динаміки психотичних розладів за 2013–2023 рр.:

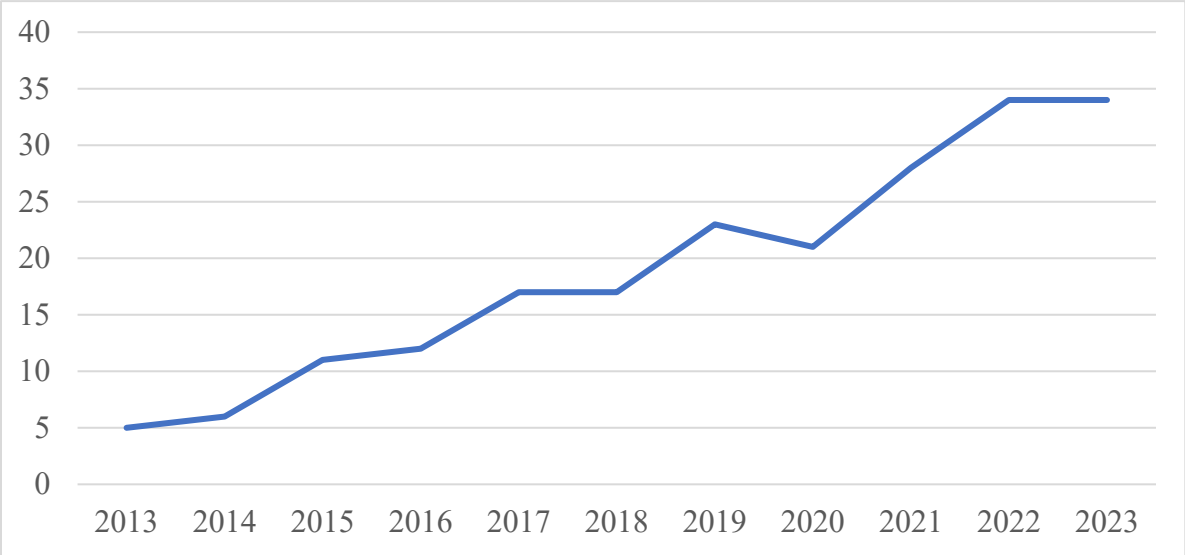


Рисунок 7.1 – Динаміка алкогольних психічних розладів у судово-психіатричній експертизі в цивільних справах, осіб

За рівнянням тренду визначимо тенденцію алкогольних психотичних розладів.

Таблиця 7.3 – Визначення тенденції за рівнянням тренду

Вид тренду	Рівняння тренду	Коефіцієнт детермінації, R^2
Лінійна	y = 3,0182x + 0,8	0,9664
Логарифмічна	y = 12,52ln(x) – 1,0117	0,8393
Степенева	$y = 4{,}1272x^{0{,}8526}$	0,9607
Експоненційна	$y = 51678{,}6e^{0{,}18866x}$	0,9316

Для дослідження тенденції розладів найбільш якісним є лінійний тренд ($R^2 = 0{,}9664$).

За допомогою Microsoft Excel спрогнозуємо подальшу тенденцію явища.

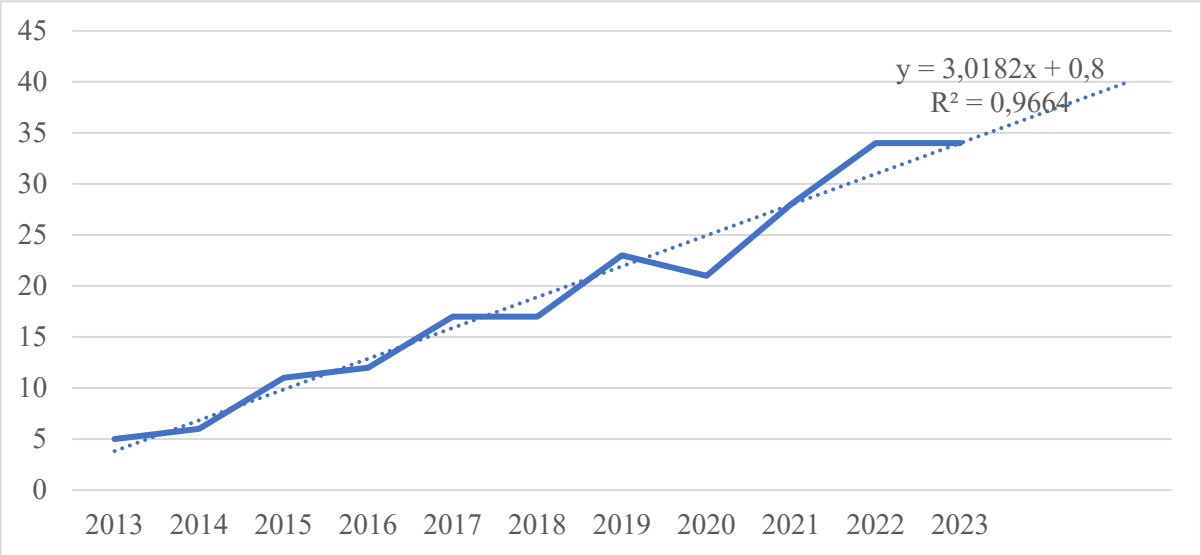


Рисунок 7.2 – Прогноз алкогольних психотичних розладів на найближчі періоди

Висновок. Отже, якщо на ситуацію не впливати, то в найближчі роки показник алкогольних психотичних розладів збільшуватиметься.

Перелік питань і завдань для самоконтролю



1. Абсолютний приріст і його значення в психологічних дослідженнях.
2. Темп зростання та його значення в психологічних дослідженнях.
3. Темп приросту та його значення в психологічних дослідженнях.
4. Абсолютне значення 1 % приросту та його значення в психологічних дослідженнях.
5. Методи виявлення тенденції в психології.
6. Інтерполяція в психології.
7. Прогнозування за тенденцією розвитку.
8. За наведеними даними проаналізувати динаміку справ про визнання особи недієздатною.

Таблиця 7.4 – Розподіл алкогольних психічних розладів залежно від категорії цивільних справ

Рік	Справи про визнання особи недієздатною
2016	14
2017	30
2018	21
2019	26
2020	23
2021	30
2022	28
2023	21
Разом	?

9. За наведеними даними визначити тенденцію розвитку явища та спрогнозувати його розвиток на найближчі періоди.

Таблиця 7.5 – Розподіл алкогольних психічних розладів залежно від категорії цивільних справ

Рік	Справи про визнання особи недієздатною
2009	6
2010	5
2011	8
2012	10
2013	13
2014	14
2015	20
2016	14
2017	30
2018	21
2019	26
2020	23
2021	30
2022	28
2023	21
Разом за 15 років	269

Тести для самоконтролю

1. Абсолютну величину розміру змін явища характеризує:

а) темп зростання;

б) коефіцієнт приросту;

в) абсолютний приріст;

г) абсолютне значення 1 % приросту.

Відповідь: 1) а; 2) б; 3) в; 4) г.

2. За яким видом розраховують середній коефіцієнт зростання:

а) арифметичним; геометричним;

б) геометричним;

в) хронологічним;

г) квадратичним.

Відповідь: 1) а; 2) б; 3) в; 4) г.

3. Яка кількісна статистична характеристика ряду динаміки визначає тенденцію розвитку явища:

а) автокореляція;

б) автоковаріація;

в) регресія;

г) тренд.

Відповідь: 1) а; 2) б; 3) в; 4) г.

4. Який вид тенденції розвитку явища характеризує тенденцію змін зв'язку між окремими рівнями ряду:

а) тенденція дисперсії;

б) тенденція середнього рівня;

в) тенденція в русі показників приросту;

г) тенденція автокореляції.

Відповідь: 1) а; 2) б; 3) в; 4) г.

5. У чому полягає сутність завдання щодо прийомів оброблення рядів динаміки з метою виявлення основної тенденції розвитку явища:

а) елімінування дії випадкових, другорядних причин, що визначають динаміку явища;

б) встановлення характеру дії основних причин, що визначають динаміку явища;

в) побудова математичних функцій динаміки;

г) елімінування дії випадкових причин і встановлення характеру дії основних причин, що визначають динаміку явища.

Відповідь: 1) а; 2) б; 3) в; 4) г.

6. Прийом змикання рядів динаміки використовують під час:

а) виявлення закономірності розвитку явища;

б) непорівняності рівнів рядів динаміки;

в) виявлення типу загальної тенденції динаміки;

г) виявлення характеру основної тенденції динаміки.

Відповідь: 1) а; 2) б; 3) в; 4) г.

7. Під загальною тенденцією динаміки розуміють:

а) тенденцію до зростання рівня явища;

б) тенденцію в русі показників динаміки;

в) тенденцію до зростання, стабільності або зниження рівня цього явища;

г) тенденцію до зростання або зниження рівнів ряду.

Відповідь: 1) а; 2) б; 3) в; 4) г.

8. Яка з відповідей виходить за межі типів динаміки:

а) абсолютні прирости стабільні;

б) абсолютні прирости зростають;

в) темпи зростання збільшуються;

г) темпи зростання зменшуються.

Відповідь: 1) а; 2) б; 3) в; 4) г.

9. Які з прийомів виявлення загальної тенденції розвитку й характеру динаміки варто використовувати, коли рівні ряду динаміки значно варіюють:

а) побудова графіків рядів динаміки;

б) згладжування через укрупнення інтервалів, згладжування за допомогою ковзної середньої;

в) визначення автокореляції у рядах динаміки;

г) змикання рядів динаміки.

Відповідь: 1) а; 2) б; 3) в; 4) г.

10. Який тип аналітичної функції використовують для вирівнювання ряду динаміки під час рівномірного збільшення абсолютних приростів:

а) рівняння параболи;

б) рівняння прямої;

в) ряд Фур’є;

г) рівняння показової функції.

Відповідь: 1) а; 2) б; 3) в; 4) г.

ТЕМА 8
МЕТОДИ ВІЗУАЛІЗАЦІЇ ТА ПОДАННЯ ДАНИХ ПСИХОЛОГІЧНИХ ДОСЛІДЖЕНЬ



8.1. Статистична таблиця як метод подання даних.

8.2. Графічні методи візуалізації даних.

8.1. Статистична таблиця як метод подання даних

Для кращого сприйняття й розуміння закономірностей психологічних явищ і процесів широко використовують статистичні таблиці, які забезпечують повне унаочнення інформації.

Статистичні таблиці призначені для зручного, систематизованого й наочного відображення даних зведення та групування даних, що сприяє кращому їх розумінню й аналізу.

Таблиця складається з горизонтальних рядків і вертикальних граф, які перетинаються. Назва таблиці, сукупність горизонтальних рядків, вертикальних граф та їхніх заголовків без числових даних *становлять макет таблиці.*

Деякі таблиці супроводжуються примітками, в яких пояснюється зміст окремих показників або заголовків, наводиться інша додаткова інформація.

Статистична таблиця має *підмет* і *присудок*.

Підмет – це перелік одиниць сукупності або їхніх груп, які характеризуються показниками. Розміщений найчастіше ліворуч підмет (назви горизонтальних рядків) вказує на об’єкт, що вивчається й відповідає на запитання: хто? що?

Присудок – це система показників, якими характеризується підмет. Він розміщується праворуч (назви вертикальних граф) і відповідає на запитання: якими показниками характеризується об’єкт?



Рисунок 8.1 – Модель статистичної таблиці

Належно побудована таблиця може ефективно унаочнювати певні властивості даних.

Під час складання статистичних таблиць варто дотримуватися правил:

- таблиця повинна бути компактною та містити лише ті цифрові дані, які потрібні для вивчення певного явища чи сукупності;
- заголовки таблиці, рядків, граф варто формулювати чітко й лаконічно, використовуючи лише загальноприйняті скорочення одиниць виміру та слів у тексті заголовка;
- якщо показників у присудку багато, їх треба пронумерувати. Водночас графи, в яких наведено перелік об’єктів або груп, позначають великим літерами алфавіту, а графи з показниками присудка – арабськими цифрами;
- якщо явище відсутнє, то у відповідній клітинці ставлять тире; якщо про нього нема інформації – ставлять три крапки; якщо виразити його у числовій формі неможливо або вона не має логічного змісту, ставлять «х»;
- кількісні показники у межах однієї графи варто наводити з однаковою точністю, тобто до 0,1; 0,01; 0,001 тощо;

- якщо наведені показники мають різні одиниці виміру, то їх вказують у відповідних графах; якщо одиниця виміру для всіх показників одна, її вказують під загальним заголовком у правому кутку таблиці;
- таблиці мають бути замкненими, тобто з підсумковими результатами.

Залежно від побудови підмета (присудка) статистичні таблиці поділяють на *прості, групові* та *комбінаційні*.

Прості таблиці містять зведені показники щодо переліку одиниць спостереження, переліку одиниць часу (хронологічних дат) або сукупностей, які вивчаються. У підметі простої таблиці перелічують одиниці сукупності або одиниці часу.

Групові таблиці містять зведення про сукупність, розподілену на окремі групи за однією ознакою. Водночас об'єкт дослідження характеризує систему показників.

Комбінаційні таблиці – таблиці, в підметі яких виділені групи за двома й більше ознаками.

8.2. Графічні методи візуалізації даних

Таблична форма подання даних доречна для їхніх зберігання й обробки, але на етапі аналізу більш доцільно використовувати графічне подання (візуалізацію) даних, наприклад діаграми розсіювання, графіки, гістограми та стовпчасті діаграми. Графічна форма полегшує сприйняття інформації як цілого, вочевиднює її особливості, тенденції та аномалії, що буде цікаво й зрозуміло не лише аналітику, а й кінцевому споживачеві.

Статистичний графік – умовне відображення числових величин за допомогою ліній, геометричних фігур та інших графічних засобів.

За допомогою графіків:

- ✓ узагальнюють і аналізують дані;
- ✓ вивчають структуру, динаміку явищ;
- ✓ визначають взаємозв'язки між показниками.

Кожен статистичний графік має такі елементи:

- ❖ заголовок (назву);
- ❖ координатну сітку;
- ❖ шкалу;
- ❖ масштаб;
- ❖ графічний образ;
- ❖ пояснення умовних позначень;
- ❖ числові дані.

Заголовок графіка повинен точно й чітко розкривати його основний зміст, давати характеристику місця та часу. Його розміщують у нижній частині графіка.

Координатна сітка – це прямокутна система координат. На горизонтальній і вертикальній осях координатної сітки будують шкали.

Шкалою називають лінію з масштабними позначеннями. Позначення відповідають цілим числам. За допомогою шкали вимірюють графічні образи.

Масштаб – міра перетворення статистичної величини в графічну.

Графічний образ – це сукупність точок, ліній, геометричних фігур, призначених для зображення відповідних статистичних величин. Він є основою графіка, його мовою.

Пояснення умовних позначень розкривають зміст окремих елементів графічного образу. Вони розташовуються на самому графіку або за його межами.

Числові дані – розміри показників, які зображують на графіку. Під час побудови графіка важливо правильно обрати масштаб, спосіб зображення, раціонально розташувати елементи графіка, не перевантажити графік матеріалом.

Характерною особливістю статистичних графіків є те, що масштаби на осях можуть бути різними, тобто шкала на вертикальній осі розграфлена інакше, ніж шкала на горизонтальній. Графік може супроводжуватися примітками, в яких вказують джерела статистичних даних, розкривають зміст і методику їхнього одержання.

За загальним призначенням графіки поділяють на:

- аналітичні;
- ілюстративні;
- інформаційні.

До *аналітичних* графіків належать графіки групувань, рядів розподілу й графіки взаємозв'язку.

Графіки групувань і рядів розподілу зі свого боку можуть бути поділені на графіки дискретних рядів розподілу та графіки інтервальних рядів розподілу.

Ілюстративні графіки показують зростання ознаки, динаміку розвитку, структуру явища.

До *інформаційних* графіків можна віднести графіки порівняння.

За способом побудови статистичні графіки поділяють на такі види:

- діаграми;
- картограми;
- картодіаграми.

Найпоширенішими графіками є діаграми.

Діаграми бувають: лінійні, стовпчикові, рядкові, секторні, фігурні.

Діаграми можна будувати в двомірному (на площині) та тримірному просторах. Побудова просторових (об'ємних) діаграм потребує додаткових знань з нарисної геометрії, й тому нами не розглядається.

Картограми та *картодіаграми* будують на основі географічної карти або топографічного плану. У картодіаграмах поєднують карти з діаграмами або фігурами. Їхня побудова потребує чималих затрат часу й особливого хисту до малювання.

За формою графічного образу виділяють графіки:

– крапкові;
– лінійні;
– площинні;
– просторові (об'ємні).

Крапковим графіком зображають дані дискретного ряду розподілу.

Лінійні графіки (лінійні діаграми) показують тенденцію розвитку явища.

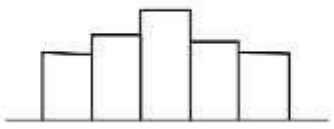
До *площинних* графіків належать діаграми стовпчикові, рядкові, секторні, фігурні.

Усі площинні діаграми можуть зі свого боку бути зображені в тримірному просторі. Тоді вони матимуть вид *просторових* графіків.

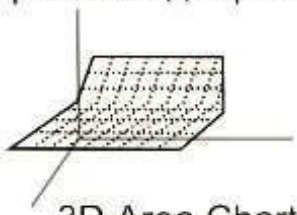
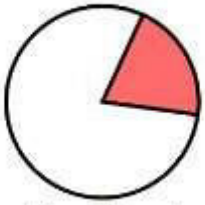
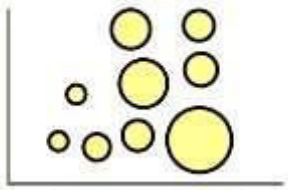
Лінійна діаграма відображує розміри показників у формі лінії, яка з'єднує розміщені в координатному полі точки. Лінія як графічний образ найбільше придатна для зображення поступального процесу розвитку, тому лінійні діаграми застосовують для вивчення динаміки показників.

Для полегшення вибору належної техніки візуалізації, яка б унаочнювала ту властивість даних, яку хотілося б показати сприймачеві, доречно скористатися класифікацією Джина Желязни. Техніки візуалізації в ній організовано як дерево, коренем якого є запитання: «Що ми хочемо показати?».

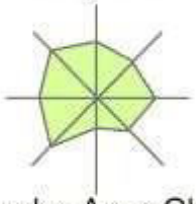
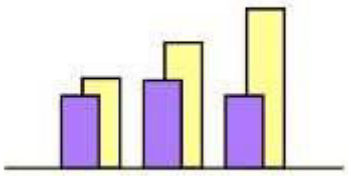
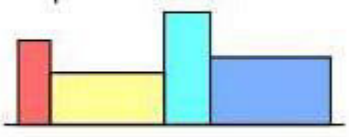
Таблиця 8.1 – Таблиця вибору техніки візуалізації (за Джином Желязни)

Що ми хочемо показати?				Техніка візуалізації
1	2	3	4	5
1.	Розподіл	Однією змінною	Небагато точок даних / категорій / інтервалів	Стовпчикова гістограма  Column Histogram
2.	Розподіл	Однією змінною	Багато точок даних	Лінійна гістограма  Line Histogram

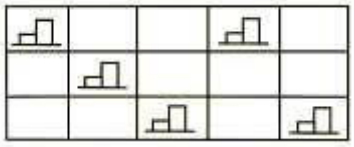
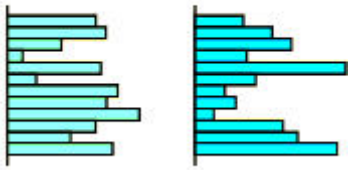
Продовження таблиці 8.1

1	2	3	4	5
3.	Розподіл	Двох змінних	Взаємозв’язок	Діаграма розкиду  Scatter Chart
4.	Розподіл	Трьох змінних	Взаємозв’язок	Тривимірна ареальна діаграма  3D Area Chart
5.	Композицію (будову)	Статичну	Просту частину загалом	Кругова діаграма  Pie Chart
6.	Зв’язок чи залежність	Двох змінних	Взаємозв’язок	Діаграма розкиду  Scatter Chart
7.	Зв’язок чи залежність	Трьох змінних	Взаємозв’язок	Бульбашкова діаграма  Bubble Chart

Продовження таблиці 8.1

1	2	3	4	5
8.	Порівняння	У часі	Багато періодів – циклічні дані	Кругова діаграма «Павутиння»  Circular Area Chart
9.	Порівняння	У часі	Багато періодів – ациклічні дані	Лінійна діаграма  Line Chart
10.	Порівняння	У часі	Мало періодів, одна чи декілька категорій	Стовпчаста діаграма  Column Chart
11.	Порівняння	У часі	Мало періодів, багато категорій	Лінійна діаграма  Line Chart
12.	Порівняння	Між об’єктами	Дві змінні на об’єкт	Стовпчаста діаграма зі змінною шириною стовпчиків  Variable Width Column Chart

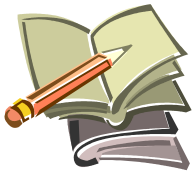
Продовження таблиці 8.1

1	2	3	4	5
13.	Порівняння	Між об’єктами	Одна змінна на об’єкт, багато категорій	Таблиця або таблиця з вбудованими діаграмам  Table or Table with Embedded Charts
14.	Порівняння	Між об’єктами	Одна змінна на об’єкт, мало категорій, багато об’єктів	Горизонтальна стовпчаста діаграма  Bar Chart

Прості засоби візуалізації внесено в сучасні електронні таблиці, які хоч і не охоплюють всього різноманіття технік, але для простих задач та оперативного прототипування цілком придатні.

Так, у програмних продуктах, як-от Microsoft Excel, SPSS або Statistika, вже закладені пакети візуалізації даних. Деякі засоби побудови діаграм вбудовано також у графічні пакети, наприклад Adobe Illustrator, їх використовують для графічно інтенсивних візуалізацій.

Приклади розв’язку типових задач



1. У групі з 27 осіб здійснено опитування, одним із запитань якого було таке: «Як часто ви відвідували кінотеатр минулого року?»

Результати опитування: 1, 1, 1, 1, 1, 2, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 5, 5, 6, 6, 7, 7, 7, 7, 7, 7.

За результатами опитування скласти ряд розподілу за частотою відвідування кінотеатру й відобразити його статистичною таблицею та графічно.

Розв'язання:

Таблиця 8.2 – Розподіл опитуваних
за частотою відвідування кінотеатру

Відвідування кінотеатру за рік	1	2	3	4	5	6	7
Кількість опитаних	5	1	8	3	2	2	6

За результатами таблиці найбільша кількість опитаних за рік відвідала кінотеатр 3 рази.

Відобразимо ряд розподілу графічно. Для такого розподілу доцільно побудувати стовпчикову діаграму (рис. 8.2).

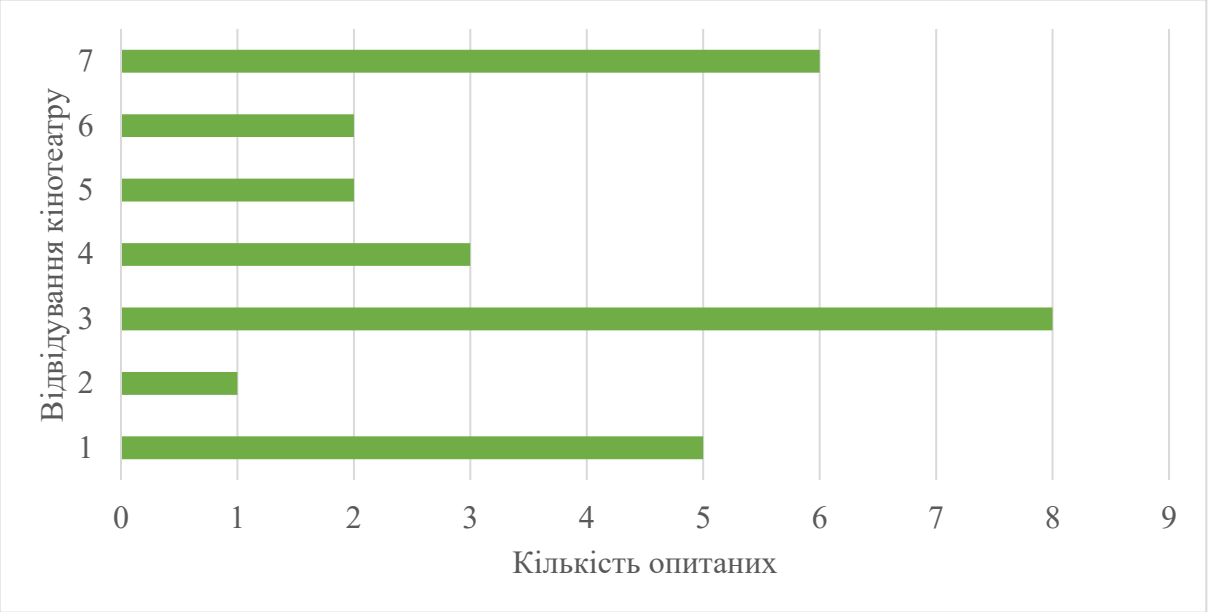


Рисунок 8.2 – Відвідування кінотеатру за рік опитаними

Графік наочно відображує, що переважною більшістю опитаних відвідано кінотеатр 3 рази за рік.

2. Під час визначення ступеня певних психічних характеристик у двох групах: експериментальній та контрольній, бали розподілилися так: експериментальна група – 16, 18, 17, 13, 15, 16, 17, 17, 22, 21, 18, 13, 12, 13, 17, 20, 20; контрольна група – 27, 9, 12, 13, 26, 23, 14, 15, 22, 21, 16, 16, 18, 17, 10, 11, 16. Проаналізуйте ступінь вираженості психічної властивості в групах.

Розв'язання:

Побудуємо стовпчикові діаграми ступеня вираженості психічної властивості за групами:



Рисунок 8.3 – Ступінь вираженості психічної властивості в експериментальній групі



Рисунок 8.4 – Ступінь вираженості психічної властивості в контрольній групі

Отже, найбільша кількість досліджуваних в експериментальній групі має ступінь вираженості психічної властивості в інтервалі 15–19. У контрольній групі є найбільше досліджуваних зі ступенем вираженості від 9 до 15.

Перелік питань і завдань для самоконтролю



1. Статистичні таблиці. Правила та види їхнього оформлення.
2. Роль і значення графічного методу в зображенні статистичних даних. Елементи статистичного графіка.
3. Класифікація видів графіків.
4. Використання графічного методу в аналізі психологічних явищ і процесів.
5. За наведеними даними, використовуючи методи візуалізації даних, визначити, чи сприяв проведений тренінг підвищенню показників за досліджуваною ознакою. Результати, отримані до та після експериментального впливу, наведені в таблиці 8.3.

Таблиця 8.3 – Результати діагностики до та після експериментального впливу

№ з/п	Показники досліджуваної ознаки	
	До участі в тренінгу	Після участі в тренінгу
1.	3	8
2.	5	8
3.	5	9
4.	8	9
5.	6	7
6.	4	8
7.	8	9
8.	3	3
9.	5	6
10.	5	8

6. Досліджуваній особі пропонують ранжувати 20 якостей за ступенем бажаності (ранг 20 присвоюють найбажанішій якості). В іншому стовпчику пропонують ранжувати ті самі якості за ступенем виразності (ранг 20 отримує найвиразніша якість).

Таблиця 8.4 – Результати ранжування якостей досліджуваної особи

№ з/п	Якості	Ранжування за ступенем бажаності	Ранжування за ступенем виразності
1.	Поступливість	14	15
2.	Сміливість	15	18
3.	Запальність	2	16
4.	Наполегливість	13	13
5.	Нервозність	1	7
6.	Терплячість	17	10
7.	Захопленість	12	20
8.	Пасивність	8	2
9.	Холодність	10	19
10.	Ентузіазм	9	17
11.	Обережність	16	4
12.	Примхливість	3	1
13.	Повільність	18	6
14.	Нерішучість	7	11
15.	Енергійність	20	12
16.	Життєрадісність	19	8
17.	Недовірливість	4	3
18.	Впертість	5	9
19.	Безтурботність	11	14
20.	Сором’язливість	6	5

Використовуючи методи візуалізації даних, зробити висновок про рівень самооцінки досліджуваної особи.

Тести для самоконтролю

1. Яка з відповідей не є правилом оформлення статистичних таблиць:

а) статистичні таблиці повинні бути оформлені різними кольорами;

б) назви таблиці, її граф і рядків повинні бути стислими;

в) для усіх показників таблиці потрібно вказувати одиниці виміру;

г) відсутність явища в таблиці позначається символом «–».

Відповідь: 1) а; 2) б; 3) в; 4) г.

2. Яка статистична таблиця називається комбінаційною:

а) підмет містить одну або більше ознак;

б) підмет містить групи за двома й більше ознаками;

в) підмет містить групи одиниць спостереження;

г) підмет містить групи одиниць спостереження за однією ознакою.

Відповідь: 1) а; 2) б; 3) в; 4) г.

3. Вид, до якого відносять статистичну таблицю, побудовану за трьома групувальними ознаками є:

а) складним;

б) груповим;

в) аналітичним;

г) комбінаційним.

Відповідь: 1) а; 2) б; 3) в; 4) г.

4. Статистичні графіки від графіків взагалі відрізняє:

а) безперервність виразу;

б) двомірність графічних знаків;

в) предмет дослідження – масові статистичні дані;

г) наявність обов’язкових елементів.

Відповідь: 1) а; 2) б; 3) в; 4) г.

5. До елементів статистичних графіків не відносять:

а) поле графіка, графічні знаки;

б) суб’єкт дослідження;

в) масштабні або просторові орієнтири;

г) експлікацію графіка.

Відповідь: 1) а; 2) б; 3) в; 4) г.

6. В основу класифікації статистичних графіків покладено:

а) форми й типи графіків;

б) умовні зображення та загальне призначення;

в) загальне призначення, види, форми й типи основних елементів;

г) предмет дослідження.

Відповідь: 1) а; 2) б; 3) в; 4) г.

7. За видами поля статистичні графіки класифікують на:

а) діаграми, картограми, картодіаграми;

б) лінійні, стовпчикові, стрічкові;

в) прямокутні, колові;

г) фігурні, лінійні.

Відповідь: 1) а; 2) б; 3) в; 4) г.

8. Які існують види діаграм:

а) лінійні, стовпчикові;

б) лінійні, стовпчикові, стрічкові, прямокутні, колові, секторні, радіальні, фігурні;

в) лінійні, фігурні;

г) прямокутні, колові.

Відповідь: 1) а; 2) б; 3) в; 4) г.

9. Варіаційні ряди графічно можна зобразити як:

а) статистичні таблиці;

б) діаграми, картограми, картодіаграми;

в) полігон, гістограма, кумулята, огіва;

г) всього вище переліченого.

Відповідь: 1) а; 2) б; 3) в; 4) г.

10. Який тип графіка графічно являє собою ряди розподілу за допомогою ламаної лінії, що зображує варіанти та частоти ряду:

а) полігон;

б) гістограма;

в) кумулята;

г) огіва.

Відповідь: 1) а; 2) б; 3) в; 4) г.

ГЛОСАРІЙ

Абсолютне значення 1 % приросту – це співвідношення показника абсолютного приросту та сотої частини базового рівня індикатора або відповідного темпу приросту.

Абсолютний показник – це показник у формі абсолютної величини, яка відображає фізичні властивості, часові та вартісні характеристики соціально-економічних процесів і явищ.

Абсолютний приріст – це індикатор часового ряду, який показує, на скільки одиниць змінився поточний рівень індикатора порівняно з рівнем попереднього або базового періоду.

Абсолютне прискорення приросту – різниця між наступним абсолютним приростом і попереднім абсолютним приростом, що показує, на скільки один показник вищий (нижчий) за попередній.

Аналітичне вирівнювання – метод виявлення основних тенденцій динаміки, який здійснюється за допомогою формули, що відображає загальну тенденцію ряду.

Варіанти – індивідуальні значення ознаки, отримані у варіаційній серії розподілів, тобто конкретне значення змінної.

Варіаційний ряд розподілу – ряд розподілу, який будується за кількісною ознакою.

Варіація – коливання значення ознаки окремих одиниць сукупності.

Вибіркова сукупність – сукупність одиниць, відібраних для дослідження.

Вибіркова частка – це відсоток одиниць за певною ознакою у вибірковій сукупності.

Відносна похибка показує, який відсоток вибіркових оцінок відхиляється від параметрів генеральної сукупності.

Відносна стандартна похибка середньої – це коефіцієнт варіації вибіркових середніх.

Відносне прискорення – це відношення абсолютного прискорення до базового абсолютного приросту. Обчислюється лише в разі, якщо абсолютний приріст бази порівняння є додатнім числом.

Відносні показники варіації – це лінійний коефіцієнт варіації, квадратичний коефіцієнт варіації та коефіцієнт осциляції.

Генеральна сукупність – множина одиниць, з якої обираються елементи для дослідження.

Гранична похибка вибірки – це максимально можлива помилка для заданої ймовірності, яка визначає розмір помилки залежно від імовірності її виявлення.

Групова статистична таблиця – це таблиця, підметом якої є групування одиниць сукупності за однією (кількісною або атрибутивною) ознакою.

Груповий добір передбачає формування вибіркової сукупності на основі добору груп одиниць із загальної сукупності.

Децилі – варіанти, які ділять ранжований ряд на 10 рівних частин.

Дискретний варіаційний ряд – ряд одиниць сукупності за дискретними характеристиками.

Дисперсія є середнім квадратом відхилення від середнього індивідуальних значень ознаки.

Ексцес розподілу – це ступінь концентрації елементів поблизу центру розподілу.

Емпіричне кореляційне відношення – квадратний корінь з коефіцієнта детермінації.

Інтервальний ряд динаміки – числовий ряд, що характеризує часові зміни величини суспільного явища, рівень якого представлений за певний період часу.

Інтерполяція – розрахунок приблизних рівнів у середині динамічного ряду, але невідомих з будь-яких причин.

Екстраполяція – визначення рівнів за межами ряду, що досліджується.

Квадратичний коефіцієнт варіації – це відношення у відсотках стандартного відхилення до середнього значення ознаки.

Квартилі – це варіанти, які поділяють упорядковані дані сукупності на рівновеликі чотири частини.

Коефіцієнт детермінації – це співвідношення міжгрупової дисперсії та загальної.

Коефіцієнт осциляції – це відсоткове відношення величини варіації до середнього значення ознаки.

Комбінаційна статистична таблиця – це таблиця, підметом якої є групування одиниць сукупності одночасно за двома й більше ознаками.

Кореляційна залежність передбачає зміну групових середніх результативної ознаки в результаті зміни факторної.

Лінійний коефіцієнт варіації – відсоткове відношення середнього значення лінійного відхилення до середнього значення ознаки.

Лінія регресії є основною характеристикою кореляції. Лінія регресії y від x – це функція, яка пов'язує середнє значення y із значенням x. За формою лінії регресії розрізняють лінійні та нелінійні зв'язки.

Медіана – значення ознаки в середині впорядкованого ряду в цій одиниці сукупності, варіанта, яка міститься в середині впорядкованого ряду та ділить його на дві рівні частини.

Метод аналітичного групування групує всі елементи вихідної інформації за факторною ознакою x та обчислює середнє ознаки результату y для кожної групи.

Метод плинних (ковзних) середніх – один із прийомів виявлення тенденції.

Міжгрупова варіація є результатом впливу чинників, які лежать в основі групування, внутрішньогрупова – це результат впливу інших чинників, ніж фактори групування.

Основна тенденція (тренд) – це зміна в часі рівня явища, яка порівняно незалежна від випадкових коливань. Основна тенденція виражається як трендове рівняння (моделі) або графіка.

Показники варіації – це показники ступеня, до якого індивідуальні значення ознаки коливаються (осцилюють) від середнього значення.

Правило розкладання (декомпозиції) варіації: для статистичної сукупності, згрупованої за певними ознаками, можна визначити три типи дисперсії: загальну, внутрішньогрупову та міжгрупову. Загальна дисперсія характеризує відхилення всіх одиниць сукупності від загального середнього значення, тобто відхилення ознак навколо загального середнього значення; внутрішньогрупова дисперсія – відхилення ознак у групі від середнього значення групи; міжгрупова дисперсія-відхилення середнього значення групи від загального середнього значення.

Простий випадковий добір – це вибірка, за якої відбір одиниць (або груп одиниць) для дослідження з генеральної сукупності здійснюється випадково.

Рівень істотності – це ймовірність того, що значення η^2 є достатньо малим, щоб бути більшим за критичне значення (коли зв'язок між атрибутами відсутній).

Рівень ряду динаміки – абсолютні (відносні та середні) значення для кожного члена часового ряду.

Рівняння регресії показує типові кореляції між величиною фактора та результату за певних умов, але функціональні рівняння дійсні лише для окремих випадків.

Розмах варіації – це різниця між максимальним і мінімальним значеннями змінної.

Розшарований (районований) добір – це метод формування, заснований на структурі загальної сукупності. За ним відбувається попереднє структурування сукупності та незалежний відбір елементів кожної складової.

Ряд розподілу – це впорядкований поділ одиниць групи на групи за певними ознаками, що варіюються.

Ряди динаміки – це значення явища на часовій осі, послідовність чисел, що характеризує зміну величини явища в часі.

Середнє квадратичне відхилення – корінь квадратний з дисперсії.

Середнє лінійне відхилення – середнє арифметичне абсолютного значення відхилення від середнього значення варіанта ознаки.

Середній або середньорічний абсолютний приріст – показник ряду динаміки, який показує, на скільки одиниць за одиницю часу (за рік) змінився рівень аналізованого показника за певний період.

Середній або середньорічний темп зростання – показник динамічного ряду, який відображує зміну рівня аналізованого показника в середньому за одиницю часу (рік) за певний період.

Середній або середньорічний темп приросту – показник динамічного ряду, який відображує середню відсоткову зміну за одиницю часу (за рік) за певний період.

Середній рівень динамічного ряду – середнє значення, розраховане на основі рівня часового ряду.

Середня арифметична зважена використовується, коли індивідуальні значення ознаки, для якої обчислюється середня, повторюються кілька разів, тобто коли розрахунок базується на згрупованих даних.

Середня арифметична проста використовується, коли розрахунок базується на незгрупованих даних.

Середня величина – це узагальнена характеристика сукупності подібних явищ за змінними, тобто узагальнений показник, який характеризує типовий рівень змінних на одиницю однорідної сукупності.

Середня гармонічна розраховується з відносних значень усереднених ознак. Може бути простою та зваженою.

Середня інтервального ряду розраховується за формулою простої середньої арифметичної. Зважені середні арифметичні з рівнями усереднення, тобто модифіковані середні використовуються для часових рядів, коли даних недостатньо, а інтервали між наданими часовими точками нерівномірні.

Середня хронологічна моментного ряду визначається сумуванням усіх даних ряду, поділена на кількість членів ряду без одиниці, водночас перший та останній дані ряду вдвічі менші за недостатніх даних часового ряду, але інтервали між заданими часовими точками рівні.

Серійний добір – добір, за якого вивчають не окремі одиниці, а серії або гнізда.

Симетричний розподіл – елементи, рівновіддалені від центра значення, мають однакову частоту.

Стандартна (середня) похибка вибірки – стандартне відхилення вибіркової оцінки від параметра в загальній сукупності, що відображує середнє значення вибіркової оцінки та можливе відхилення від середнього значення в генеральній сукупності.

Стохастичний зв'язок – у цьому зв'язку значенню ознаки x відповідає певний набір значень ознаки y.

Структурне групування – це поділ однорідної сукупності на групи, структура яких характеризується певними варіативними (змінними) ознаками.

Темп зростання або коефіцієнт зростання – характеристика ряду динаміки, яка показує, у скільки разів поточний рівень показника змінився порівняно з рівнем попереднього або базового періоду.

Темп приросту – показник ряду динаміки, який показує відсоткову зміну поточного (порівнянного) рівня аналізованого показника порівняно з рівнем попереднього або базового періоду.

Функціональний зв'язок має повну відповідність факторної та результативної ознаки, тобто кожному значенню ознаки *x* чітко має відповідне значення ознаки *y*.

Частка – це частота, виражена як відносна величина як коефіцієнт або відсоток від загальної кількості.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Боснюк В. Ф. Математичні методи в психології : курс лекцій. Мультимедійне навчальне видання. Х. : НУЦЗУ, 2020. 141 с.
2. Климчук В. О. Математичні методи у психології : навчальний посібник для студентів психологічних спеціальностей. К. : Освіта України, 2009. 288 с.
3. Краєвський В. М., Остапенко Я. О., Параниця Н. В. Статистика. Ірпінь : Університет ДФС України, 2019. 218 с.
4. Москальов І. О., Лисенко Д. П. Застосування методів математичної статистики у психолого-педагогічних дослідженнях : навч. посіб. Київ : НУОУ, 2023. 187 с.
5. Корнієнко І. О., Воронова О. Ю. Статистичні методи у психології : курс лекцій з дисципліни для студентів денної та заочної форм навчання спеціальності 053 Психологія ОС «Магістр». Мукачево : МДУ, 2019. 44 с.
6. Руська Р. В. Теорія імовірності та математична статистика в психології : навч. посіб. Тернопіль, 2020. 112 с.
7. Дорогань-Писаренко Л. О., Єгорова О. В., Рудич А. І. Статистика : навч. посіб. Полтава : РВВ ПДАУ, 2021. 300 с.
8. Краєвський В. М., Параниця Н. В., Остапенко Я. О. Фінансова статистика : навчальний посібник. Ірпінь : Університет ДФС України, 2018. 308 с. URL : https://ir.dpu.edu.ua/items/829abde6-d06f-4ad1-924b-232a8ced0a01
9. Майборода Р. Є., Сугакова О. В. Статистичний аналіз даних за допомогою пакета STATISTICA. URL : http://matphys.rpd.univ.kiev.ua/ downloads/courses/mmatstat/StatAn.doc
10. Митна статистика України : навч. посібн. / Пашко П. В., Ушаков О. Ю., Краєвський В. М., Паянок Т. М., Остапенко Я. О., Параниця Н. В., Богдан С. В. / Університет державної фіскальної служби України. Ірпінь : УДФСУ, 2020. 380 с. (Серія «Податкова та митна справа в Україні ; т. 135). URL : https://shron1.chtyvo.org.ua/Pashko_Pavlo/Mytna_statystyka_Ukrainy.pdf?
11. Найко Д. А., Шевчук О. Ф. Теорія ймовірностей та математична статистика : навч. посіб. Вінниця : ВНАУ, 2020. 382 с.

12. Карпенко Л. М. Статистика : навч. посіб. Одеса : ОРІДУ НАДУ, 2019. 184 с.

13. Раєвнєва О. В., Аксьонова І. В., Бровко О. І. Статистика : навчальний посібник / за заг. ред. д-ра екон. наук, професора О. В. Раєвнєвої. Харків : ХНЕУ ім. С. Кузнеця, 2019. 389 с.

14. Дорогань-Писаренко Л. О., Єгорова О. В., Рудич А. І. Статистика : навч. посібник. Полтава : РВВ ПДАУ, 2021. 300 с.

15. Остапенко Я. О. Використання ІКТ для статистичного аналізу психологічних процесів майбутніх психологів. *Інженерні та освітні технології.* 2021. Т. 9, № 1. С. 102–113. DOI : https://doi.org/10.30929/2307-9770.2021.09.01.09

16. Остапенко Я. О. Використання PSPP під час статистичного аналізу. *Східна Європа: економіка, бізнес та управління.* 2018. Вип. 2 (13). С. 254–265. URL : http://ir.nusta.edu.ua/handle/doc/3309

17. Остапенко Я. О. Економіко-математичне моделювання як інструментарій прогнозування: теоретичний аспект. *International Scientific – Practical Conference Modern Transformation of Economics and Management in the Era of Globalization* : Conference Proceedings. January 29, 2016. Klaipeda : Baltija Publishing, 2016. P. 321–323.

18. Остапенко Я. О. Програмні продукти статистичних та економіко-математичних досліджень в економіці. *International Scientific Conference Anti-Crisis Management: State, Region, Enterprise* : Conference Proceedings. Part III. November 17th, 2017. Le Mans, France : Baltija Publishing, 2017. P. 48–50.

19. Паянок Т. М., Задорожня Т. М. Статистичний аналіз даних : навчальний посібник. Ірпінь : Університет державної фіскальної служби України, 2020. 312 с.

20. Аналіз даних за допомогою R : навчально-методичний посібник / Новицька Н., Паянок Т., Параниця Н., Богдан С. Ірпінь : Університет державної фіскальної служби України, 2021. 324 с. (Серія «На допомогу студенту УДФСУ» ; т. 87).

21. Програми для математичної і статистичної обробки даних. URL : http://chembio.com.ua/aspirant/grant/item/

22. Функції Excel (за категоріями). URL : https://support.office.com/uk-ua/article

23. Математичні методи в психології. Курс відеолекцій. URL : https://www.youtube.com/channel/UCu7Tj9z4NnyenGXmkvMGEnA/videos

24. Параниця Н. В., Скасків Л. В., Чернобай О. Б. Статистичний аналіз даних при використанні хмарних технологій. *Наука і техніка сьогодні.* 2023. № 5(19). С. 42–49. DOI : https://doi.org/10.52058/2786-6025-2023-5(19)-42-49

25. Параниця Н. В., Параниця С. П., Таращенко В. А. Когнітивне моделювання сталого розвитку регіонів України. *Ефективна економіка.* 2023. № 4. DOI : http://doi.org/10.32702/2307-2105.2023.4.21

26. Параниця Н. В., Параниця С. П. Системний аналіз та моделювання війни в Україні. *Наукові перспективи.* 2022. № 5(23). С. 401–408. DOI : https://doi.org/10.52058/2708-7530-2022-5(23)

27. Борейко Н. М., Параниця Н. В., Параниця С. П. Оцінювання ризиків у сфері трансфертного ціноутворення на прикладі податкових відомств країн-членів Європейського Союзу. *Інвестиції: практика та досвід.* 2022. № 5–6. С. 17–23. DOI 10.32702/2306-6814.2022.5-6.17

28. Параниця Н. В., Буличов О. С., Охмак О. М. Переваги і ризики застосування хмарних технологій в сфері бухгалтерського обліку. *Економіка та держава* : міжнародний науково-практичний журнал. Київ, 2021. № 4. С. 128–131. DOI 10.32702/23066806.2021.4.128

29. Задорожня Т. М., Параниця Н. В., Таращенко В. А. Статистичний аналіз динаміки паритету. *Східна Європа: Економіка, бізнес та управління* : електронне фахове видання. 2021. Випуск 1 (28). С. 123–128.

30. Andry Field. Discovering Statistics using SPSS. URL : https://in.sagepub.com/.../-discoveringstatistics-using...spss-statastics/

31. Bentler P. M. EQS 6 Structural Equations Program Manual. Encino, CA: Multivariate Software, Inc., 2006. 418 p.

32. Byrne B. M. Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming. 2nd ed. Multivariate applications series. New York : Taylor & Francis Group, 2010. 396 p.

33. Harrington D. Confirmatory factof analysis. New York : Oxford University Press, Inc., 2009. 122 p.

34. Joseph F. Healey. Statistics: A tool for Social Research. URL : https://www.amazon.-com/Statistics-Research-Joseph-F-Healey

35. Kline R. B. Principles and practice of structural equation modeling. 3nd ed. New York : The Guilford Press, 2011. 432 p.

36. European Statistical System Committee. (2018). Bucharest Memorandum on Official Statistics in a Datafied Society (Trusted Smart Statistics). URL : https://ec.europa.eu/eurostat/web/ess/-/dgins2018-bucharest-memorandum-adopted

37. Promoting reproducibility-by-design in statistical offices / Luhmann, S., Grazzini, J., Ricciato, F., Meszaros, M., Giannakouris, K., Museux, J-M., Hahn, M. *In Proceedings for New Techniques and Technologies for Statistics (NTTS), Eurostat*. 2019. DOI : https://doi.org/10.2901/EUROSTAT.C2017Google Scholar

38. Ruppert, E., Gromme, F., Cakici, B. (2018). Citizen data and trust in official statistics. *Economie et Statistique / Economics and Statistics.* P. 505–506, 179–193. DOI : https://doi.org/10.24187/ecostat.2018.505d.1971Google Scholar

39. Ricciato, F. (2018). Towards a reference methodological framework for processing mno data for official statistics. *In 15th Global Forum on Tourism Statistics.* Cusco, Peru. URL : https://tinyurl.com/ycgvx4m6Google Scholar

40. Show me your code, and then I will trust your figures: Towards software-agnostic open algorithms in statistical production / Grazzini, J., Lamarche, P., Gaffuri, J., Museux, J-M. *In Proceedings of the Quality Conference.* 2018. DOI : https://doi.org/10.5281/zenodo.3240282

ДОДАТКИ

Додаток А

Значення критерію Пірсона χ^2 для рівня істотності 0,10; 0,05; 0,01

k	*0,10*	*0,05*	*0,01*	*k*	*0,10*	*0,05*	*0,01*
1	2,71	3,84	6,63	21,00	29,62	32,67	38,93
2	4,61	5,99	9,21	22,00	30,81	3392,	40,29
3	6,25	7,81	11,34	23,00	32,01	35,17	41,64
4	7,78	9,49	13,28	24,00	33,20	36,42	42,98
5	9,24	11,07	15,09	25,00	34,38	37,65	44,31
6	10,64	12,59	16,81	26,00	35,56	38,89	45,64
7	12,02	14,07	18,48	27,00	36,74	40,11	46,96
8	13,36	15,51	20,09	28,00	37,92	41,34	48,28
9	14,68	16,92	21,67	29,00	39,09	42,56	49,59
10	50,99	18,31	23,21	30,00	40,26	43,77	50,89
11	17,28	19,68	24,72	40,00	51,80	55,76	63,69
12	18,55	21,03	26,22	50,00	63,17	67,50	76,15
13	19,81	22,36	27,69	60,00	74,40	79,08	88,38
14	21,06	23,68	29,14	70,00	85,53	90,53	100,42
15	22,31	25,00	30,58	80,00	96,58	101,88	112,33
16	23,54	26,30	32,00	90,00	107,56	113,14	124,12
17	24,77	27,59	33,41	100,00	118,50	124,34	135,81
18	25,99	28,87	34,81				
19	27,20	30,14	36,19				
20	28,41	31,41	37,57				

Додаток Б

Критичні значення кореляційного відношення η^2 і коефіцієнта детермінації R^2 у разі рівня значимості 0,05

k_2	k_1						
	1	2	3	4	5	6	7
5	0,569	0,699	0,764	0,806	0,835	0,854	0,872
6	0,500	0,632	0,704	0,751	0,785	0,811	0,831
7	0,444	0,575	0,651	0,702	0,739	0,768	0,791
8	0,399	0,527	0,604	0,657	0,697	0,729	0,754
9	0,362	0,488	0,563	0,618	0,659	0,692	0,719
10	0,332	0,451	0,527	0,582	0,624	0,659	0,687
12	0,283	0,394	0,466	0,521	0,564	0,600	0,630
14	0,247	0,348	0,417	0,471	0,514	0,550	0,580
16	0,219	0,312	0,378	0,429	0,477	0,507	0,538
18	0,197	0,283	0,345	0,394	0,435	0,470	0,501
20	0,179	0,259	0,318	0,364	0,404	0,432	0,468
22	0,164	0,238	0,294	0,339	0,377	0,410	0,439
24	0,151	0,221	0,273	0,316	0,353	0,385	0,414
26	0,140	0,206	0,256	0,297	0,332	0,363	0,391
28	0,130	0,193	0,240	0,279	0,314	0,344	0,371
30	0,122	0,182	0,227	0,264	0,297	0,326	0,353
40	0,093	0,139	0,176	0,207	0,234	0,259	0,282
50	0,075	0,113	0,143	0,170	0,194	0,216	0,235
60	0,063	0,095	0,121	0,144	0,165	0,184	0,202
80	0,047	0,072	0,093	0,110	0,127	0,142	0,156
100	0,038	0,058	0,075	0,090	0,103	0,116	0,128
120	0,032	0,049	0,063	0,080	0,087	0,098	0,109
200	0,019	0,030	0,038	0,046	0,053	0,060	0,067
400	0,010	0,015	0,019	0,023	0,027	0,031	0,034

Додаток В

Таблиця значень F-критерію Фішера у разі рівня значимості α = 0,05

*k*1 / *k*2	1	2	3	4	5	6	8	12	24	∞
1	161,45	199,50	215,72	224,57	230,17	233,97	238,89	243,91	249,04	254,32
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62

	1	2	3	4	5	6	8	12	24	∞
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Критичні значення t-критерію Стьюдента в разі рівня значимості 0,10; 0,05; 0,01

Число ступенів вільності	α			Число ступенів вільності	α		
	0,10	0,05	0,01		0,10	0,05	0,01
1	6,3138	12,706	63,657	18	1,7341	2,1009	2,8784
2	2.9200	4,3027	9,9248	19	1,7291	2,0930	2,8609
3	2,3534	3,1825	5,8409	20	1,7247	2,0860	2,8453
4	2,1318	2,7764	4,6041	21	1,7207	2,0796	2,8314
5	2,0150	2,5706	4,0321	22	1,7171	2,0739	2,8188
6	1,9432	2,4469	3,7074	23	1,7139	2,0687	2,8073
7	1,8946	2,3646	3,4995	24	1,7109	2,0639	2,7969
8	1,8595	2,3060	3,3554	25	1,7081	2,0595	2,7874
9	1,8331	2,2622	3,2498	26	1,7056	2,0555	2,7787
10	1,8125	2,2281	3,1693	27	1,7033	2,0518	2,7707
11	1,7959	2,2010	3,1058	28	1,7011	2,0484	2,7633
12	1,7823	2,1788	3,0545	29	1,6991	2,0452	2,7564
13	1,7709	2,1604	3,0123	30	1,6973	2,0423	2,7500
14	1,7613	2,1448	2,9768	40	1,6839	2,0211	2,7045
15	1,7530	2,1315	2,9467	60	1,6707	2,0003	2,6603
16	1,7459	2,1199	2,9208	120	1,6577	1,9799	2,6174
17	1,7396	2,1098	2,8982	00	1,6449	1,9600	2,5758

Додаток Д

Критичні значення вибіркового коефіцієнта кореляції рангів Спірмена

Зв'язок достовірний, якщо $r_{s\ емп} \geq r_{s\ 0,05}$, і тим більше достовірний, якщо $r_{s\ емп} \geq r_{s\ 0,01}$.

n	*p*		*n*	*p*		*n*	*p*	
	0,05	*0,01*		*0,05*	*0,01*		*0,05*	*0,01*
5	0,94	-	17	0,48	0,62	29	0,37	0,48
6	0,85	-	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	21	0,36	0,46
8	0,72	0,88	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,54	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,58	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,49	0,51	37	0,33	0,43
14	0,54	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,66	27	0,38	0,49	39	0,32	0,41
16	0,50	0,64	28	0,38	0,48	40	0,31	0,40

Додаток Е

Критичні значення коефіцієнта асиметрії (Ах)

Обсяг вибірки *(n)*	Рівні значимості (*p*)		Об'єм вибірки *(n)*	Рівні значимості (*p*)	
	0,05	*0,01*		*0,05*	*0,01*
25	0,711	1,061	250	0,251	0,360
30	0,661	0,982	300	0,230	0,329
35	0,621	0,921	350	0,213	0,305
40	0,587	0,869	400	0,200	0,285
45	0,558	0,825	450	0,188	0,269
50	0,533	0,787	500	0,179	0,255
60	0,492	0,723	550	0,171	0,243
70	0,459	0,673	600	0,163	0,233
80	0,432	0,631	650	0,157	0,224
90	0,409	0,596	700	0,151	0,215
100	0,389	0,567	750	0,146	0,208
125	0,350	0,508	800	0,142	0,202
150	0,321	0,464	850	0,138	0,196
175	0,298	0,430	900	0,134	0,190
200	0,280	0,403	950	0,130	0,185

Критичні значення коефіцієнта ексцесу (Ех)

Обсяг вибірки *(n)*	Рівні значимості (*p*)		
	0,1	*0,05*	*0,01*
11	0,890	0,907	0,936
16	0,873	0,888	0,914
21	0,863	0,877	0,900
26	0,857	0,869	0,890
31	0,851	0,863	0,883
36	0,847	0,858	0,877
41	0,844	0,854	0,872
46	0,841	0,851	0,868
51	0,839	0,848	0,865
61	0,835	0,843	0,859
71	0,832	0,840	0,855
81	0,830	0,838	0,852
91	0,828	0,835	0,848
101	0,826	0,834	0,846
201	0,818	0,823	0,832
301	0,814	0,818	0,826
401	0,812	0,816	0,822
501	0,810	0,814	0,820

Додаток Ж

Середні коефіцієнти зростання

Коефіцієнт зростання (К)	Середній коефіцієнт динаміки								
	$\sqrt{}$	$\sqrt[3]{}$	$\sqrt[4]{}$	$\sqrt[5]{}$	$\sqrt[6]{}$	$\sqrt[7]{}$	$\sqrt[8]{}$	$\sqrt[9]{}$	$\sqrt[10]{}$
0,860	0,9274	0,9510	0,9630	0,9703	0,9752	0,9787	0,9813	0,9834	0,9850
0,865	0,9301	0,9528	0,9644	0,9714	0,9761	0,9795	0,9820	0,9840	0,9856
0,870	0,9327	0,9546	0,9658	0,9725	0,9771	0,9803	0,9827	0,9846	0,9862
0,875	0,9354	0,9565	0,9672	0,9736	0,9780	0,9811	0,9834	0,9853	0,9867
0,880	0,9381	0,9583	0,9685	0,9748	0,9789	0,9819	0,9841	0,9859	0,9873
0,885	0,9407	0,9601	0,9699	0,9759	0,9798	0,9827	0,9848	0,9865	0,9879
0,890	0,9434	0,9619	0,9713	0,9770	0,9808	0,9835	0,9855	0,9871	0,9884
0,895	0,9460	0,9637	0,9726	0,9781	0,9817	0,9843	0,9862	0,9877	0,9890
0,900	0,9487	0,9655	0,9740	0,9791	0,9826	0,9851	0,9869	0,9884	0,9895
0,905	0,9513	0,9673	0,9754	0,9802	0,9835	0,9858	0,9876	0,9890	0,9901
0,910	0,9539	0,9691	0,9767	0,9813	0,9844	0,9866	0,9883	0,9896	0,9906
0,915	0,9566	0,9708	0,9780	0,9824	0,9853	0,9874	0,9890	0,9902	0,9912
0,920	0,9592	0,9726	0,9794	0,9835	0,9862	0,9882	0,9896	0,9908	0,9917
0,925	0,9618	0,9743	0,9807	0,9845	0,9871	0,9889	0,9903	0,9914	0,9922
0,930	0,9644	0,9761	0,9820	0,9856	0,9880	0,9897	0,9910	0,9920	0,9928
0,935	0,9670	0,9778	0,9833	0,9866	0,9889	0,9904	0,9916	0,9926	0,9933
0,940	0,9695	0,9796	0,9847	0,9877	0,9897	0,9912	0,9923	0,9931	0,9938
0,945	0,9721	0,9813	0,9860	0,9887	0,9906	0,9920	0,9930	0,9937	0,9944
0,950	0,9747	0,9830	0,9873	0,9898	0,9915	0,9927	0,9936	0,9943	0,9949
0,955	0,9772	0,9848	0,9886	0,9908	0,9924	0,9934	0,9943	0,9949	0,9954
0,960	0,9798	0,9865	0,9898	0,9919	0,9932	0,9942	0,9949	0,9955	0,9959
0,965	0,9823	0,9882	0,9911	0,9929	0,9941	0,9949	0,9956	0,9960	0,9964
0,970	0,9849	0,9899	0,9924	0,9939	0,9949	0,9957	0,9962	0,9966	0,9970
0,975	0,9874	0,9916	0,9937	0,9949	0,9958	0,9964	0,9968	0,9972	0,9975
0,980	0,9899	0,9933	0,9950	0,9960	0,9966	0,9971	0,9975	0,9978	0,9980
0,985	0,9925	0,9950	0,9962	0,9970	0,9975	0,9978	0,9981	0,9983	0,9985
0,990	0,9950	0,9967	0,9975	0,9980	0,9983	0,9986	0,9987	0,9989	0,9990
0,995	0,9975	0,9983	0,9987	0,9990	0,9992	0,9993	0,9994	0,9994	0,9995
1,000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
1,005	1,0025	1,0017	1,0012	1,0010	1,0008	1,0007	1,0006	1,0006	1,0005
1,010	1,0050	1,0033	1,0025	1,0020	1,0017	1,0014	1,0012	1,0011	1,0010
1,015	1,0075	1,0050	1,0037	1,0030	1,0025	1,0021	1,0019	1,0017	1,0015
1,020	1,0100	1,0066	1,0050	1,0040	1,0033	1,0028	1,0025	1,0022	1,0020
1,025	1,0124	1,0083	1,0062	1,0050	1,0041	1,0035	1,0031	1,0027	1,0025
1,030	1,0149	1,0099	1,0074	1,0059	1,0049	1,0042	1,0037	1,0033	1,0030
1,035	1,0173	1,0115	1,0086	1,0069	1,0058	1,0049	1,0043	1,0038	1,0034
1,040	1,0198	1,0132	1,0099	1,0079	1,0066	1,0056	1,0049	1,0044	1,0039
1,045	1,0223	1,0148	1,0111	1,0088	1,0074	1,0063	1,0055	1,0049	1,0044
1,050	1,0247	1,0164	1,0123	1,0098	1,0082	1,0070	1,0061	1,0054	1,0049
1,055	1,0271	1,0180	1,0135	1,0108	1,0090	1,0077	1,0067	1,0060	1,0054

(K)	$\sqrt{}$	$\sqrt[3]{}$	$\sqrt[4]{}$	$\sqrt[5]{}$	$\sqrt[6]{}$	$\sqrt[7]{}$	$\sqrt[8]{}$	$\sqrt[9]{}$	$\sqrt[10]{}$
1,060	1,0296	1,0196	1,0147	1,0117	1,0098	1,0084	1,0073	1,0065	1,0058
1,065	1,0320	1,0212	1,0159	1,0127	1,0106	1,0090	1,0079	1,0070	1,0063
1,070	1,0344	1,0228	1,0171	1,0136	1,0113	1,0097	1,0085	1,0075	1,0068
1,075	1,0368	1,0244	1,0182	1,0146	1,0121	1,0104	1,0091	1,0081	1,0073
1,080	1,0392	1,0260	1,0194	1,0155	1,0129	1,0111	1,0097	1,0086	1,0077
1,085	1,0416	1,0276	1,0206	1,0164	1,0137	1,0117	1,0102	1,0091	1,0082
1,090	1,0440	1,0291	1,0218	1,0174	1,0145	1,0124	1,0108	1,0096	1,0087
1,095	1,0464	1,0307	1,0229	1,0183	1,0152	1,0130	1,0114	1,0101	1,0091
1,100	1,0488	1,0323	1,0241	1,0192	1,0160	1,0137	1,0120	1,0106	1,0096
1,105	1,0512	1,0338	1,0253	1,0202	1,0168	1,0144	1,0126	1,0112	1,0100
1,110	1,0536	1,0354	1,0264	1,0211	1,0175	1,0150	1,0131	1,0117	1,0105
1,115	1,0559	1,0370	1,0276	1,0220	1,0183	1,0157	1,0137	1,0122	1,0109
1,120	1,0583	1,0385	1,0287	1,0229	1,0191	1,0163	1,0143	1,0127	1,0114
1,125	1,0607	1,0400	1,0299	1,0238	1,0198	1,0170	1,0148	1,0132	1,0118
1,130	1,0630	1,0416	1,0310	1,0247	1,0206	1,0176	1,0154	1,0137	1,0123
1,135	1,0654	1,0431	1,0322	1,0256	1,0213	1,0183	1,0160	1,0142	1,0127
1,140	1,0677	1,0446	1,0333	1,0266	1,0221	1,0189	1,0165	1,0147	1,0132
1,145	1,0700	1,0462	1,0344	1,0275	1,0228	1,0195	1,0171	1,0152	1,0136
1,150	1,0724	1,0477	1,0356	1,0283	1,0236	1,0202	1,0176	1,0157	1,0141
1,155	1,0747	1,0492	1,0367	1,0292	1,0243	1,0208	1,0182	1,0161	1,0145
1,160	1,0770	1,0507	1,0378	1,0301	1,0250	1,0214	1,0187	1,0166	1,0150
1,165	1,0794	1,0522	1,0389	1,0310	1,0258	1,0221	1,0193	1,0171	1,0154
1,170	1,0817	1,0537	1,0400	1,0319	1,0265	1,0227	1,0198	1,0176	1,0158
1,175	1,0840	1,0552	1,0411	1,0328	1,0272	1,0233	1,0204	1,0181	1,0163
1,180	1,0863	1,0567	1,0422	1,0337	1,0280	1,0239	1,0209	1,0186	1,0167
1,185	1,0886	1,0582	1,0433	1,0345	1,0287	1,0245	1,0214	1,0190	1,0171
1,190	1,0909	1,0597	1,0444	1,0354	1,0294	1,0252	1,0220	1,0195	1,0175
1,195	1,0932	1,0612	1,0455	1,0363	1,0301	1,0258	1,0225	1,0200	1,0180
1,200	1,0954	1,0627	1,0466	1,0371	1,0309	1,0264	1,0231	1,0205	1,0184
1,205	1,0977	1,0641	1,0477	1,0380	1,0316	1,0270	1,0236	1,0209	1,0188
1,210	1,1000	1,0656	1,0488	1,0389	1,0323	1,0276	1,0241	1,0214	1,0192
1,215	1,1023	1,0671	1,0499	1,0397	1,0330	1,0282	1,0246	1,0219	1,0197
1,220	1,1045	1,0685	1,0510	1,0406	1,0337	1,0288	1,0252	1,0223	1,0201
1,225	1,1068	1,0700	1,0520	1,0414	1,0344	1,0294	1,0257	1,0228	1,0205
1,230	1,1091	1,0714	1,0531	1,0423	1,0351	1,0300	1,0262	1,0233	1,0209
1,235	1,1113	1,0729	1,0542	1,0431	1,0358	1,0306	1,0267	1,0237	1,0213
1,240	1,1136	1,0743	1,0553	1,0440	1,0365	1,0312	1,0273	1,0242	1,0217
1,245	1,1158	1,0758	1,0563	1,0448	1,0372	1,0318	1,0278	1,0246	1,0222
1,250	1,1180	1,0772	1,0574	1,0456	1,0379	1,0324	1,0283	1,0251	1,0226
1,255	1,1203	1,0787	1,0584	1,0465	1,0386	1,0330	1,0288	1,0256	1,0230
1,260	1,1225	1,0801	1,0595	1,0473	1,0393	1,0336	1,0293	1,0260	1,0234
1,265	1,1247	1,0815	1,0605	1,0481	1,0400	1,0342	1,0298	1,0265	1,0238
1,270	1,1269	1,0829	1,0616	1,0490	1,0406	1,0347	1,0303	1,0269	1,0242
1,275	1,1292	1,0844	1,0626	1,0498	1,0413	1,0353	1,0308	1,0274	1,0246

(K)	$\sqrt{}$	$\sqrt[3]{}$	$\sqrt[4]{}$	$\sqrt[5]{}$	$\sqrt[6]{}$	$\sqrt[7]{}$	$\sqrt[8]{}$	$\sqrt[9]{}$	$\sqrt[10]{}$
1,280	1,1314	1,0858	1,0637	1,0506	1,0420	1,0359	1,0313	1,0278	1,0250
1,285	1,1336	1,0872	1,0647	1,0514	1,0427	1,0365	1,0318	1,0283	1,0254
1,290	1,1358	1,0886	1,0657	1,0522	1,0434	1,0370	1,0323	1,0287	1,0258
1,295	1,1380	1,0900	1,0668	1,0531	1,0440	1,0376	1,0328	1,0291	1,0262
1,300	1,1402	1,0914	1,0678	1,0539	1,0447	1,0382	1,0333	1,0296	1,0266
1,305	1,1424	1,0928	1,0688	1,0547	1,0454	1,0388	1,0338	1,0300	1,0270
1,310	1,1446	1,0942	1,0698	1,0555	1,0460	1,0393	1,0343	1,0305	1,0274
1,315	1,1467	1,0956	1,0709	1,0563	1,0467	1,0399	1,0348	1,0309	1,0278
1,320	1,1489	1,0970	1,0719	1,0571	1,0474	1,0405	1,0353	1,0313	1,0282
1,325	1,1511	1,0983	1,0729	1,0579	1,0480	1,0410	1,0358	1,0318	1,0285
1,330	1,1533	1,0997	1,0739	1,0587	1,0487	1,0416	1,0363	1,0322	1,0289
1,335	1,1554	1,1011	1,0749	1,0595	1,0493	1,0421	1,0368	1,0326	1,0293
1,340	1,1576	1,1025	1,0759	1,0603	1,0500	1,0427	1,0373	1,0331	1,0297
1,345	1,1597	1,1038	1,0769	1,0611	1,0506	1,0433	1,0377	1,0335	1,0301
1,350	1,1619	1,1052	1,0779	1,0619	1,0513	1,0438	1,0382	1,0339	1,0305
1,355	1,1640	1,1066	1,0789	1,0626	1,0519	1,0444	1,0387	1,0343	1,0308
1,360	1,1662	1,1079	1,0799	1,0634	1,0526	1,0449	1,0392	1,0348	1,0312
1,365	1,1683	1,1093	1,0809	1,0642	1,0532	1,0455	1,0397	1,0352	1,0316
1,370	1,1705	1,1106	1,0819	1,0650	1,0539	1,0460	1,0401	1,0356	1,0320
1,375	1,1726	1,1120	1,0829	1,0658	1,0545	1,0465	1,0406	1,0360	1,0324
1,380	1,1747	1,1133	1,0839	1,0665	1,0551	1,0471	1,0411	1,0364	1,0327
1,385	1,1769	1,1147	1,0848	1,0673	1,0558	1,0476	1,0416	1,0369	1,0331
1,390	1,1790	1,1160	1,0858	1,0681	1,0564	1,0482	1,0420	1,0373	1,0335
1,395	1,1811	1,1174	1,0868	1,0688	1,0571	1,0487	1,0425	1,0377	1,0338
1,400	1,1832	1,1187	1,0878	1,0696	1,0577	1,0492	1,0430	1,0381	1,0342
1,405	1,1853	1,1200	1,0887	1,0704	1,0583	1,0498	1,0434	1,0385	1,0346
1,410	1,1874	1,1213	1,0897	1,0711	1,0589	1,0503	1,0439	1,0389	1,0350
1,415	1,1895	1,1227	1,0907	1,0719	1,0596	1,0508	1,0443	1,0393	1,0353
1,420	1,1916	1,1240	1,0916	1,0726	1,0602	1,0514	1,0448	1,0397	1,0357
1,425	1,1937	1,1253	1,0926	1,0734	1,0608	1,0519	1,0453	1,0401	1,0361
1,430	1,1958	1,1266	1,0935	1,0742	1,0614	1,0524	1,0457	1,0405	1,0364
1,435	1,1979	1,1279	1,0945	1,0749	1,0620	1,0529	1,0462	1,0409	1,0368
1,440	1,2000	1,1292	1,0954	1,0757	1,0627	1,0535	1,0466	1,0413	1,0371
1,445	1,2021	1,1305	1,0964	1,0764	1,0633	1,0540	1,0471	1,0417	1,0375
1,450	1,2042	1,1319	1,0973	1,0771	1,0639	1,0545	1,0475	1,0421	1,0379
1,455	1,2062	1,1332	1,0983	1,0779	1,0645	1,0550	1,0480	1,0425	1,0382
1,460	1,2083	1,1344	1,0992	1,0786	1,0651	1,0556	1,0484	1,0429	1,0386
1,465	1,2104	1,1357	1,1002	1,0794	1,0657	1,0561	1,0489	1,0433	1,0389
1,470	1,2124	1,1370	1,1011	1,0801	1,0663	1,0566	1,0493	1,0437	1,0393
1,475	1,2145	1,1383	1,1020	1,0808	1,0669	1,0571	1,0498	1,0441	1,0396
1,480	1,2166	1,1396	1,1030	1,0816	1,0675	1,0576	1,0502	1,0445	1,0400
1,485	1,2186	1,1409	1,1039	1,0823	1,0681	1,0581	1,0507	1,0449	1,0403
1,490	1,2207	1,1422	1,1048	1,0830	1,0687	1,0586	1,0511	1,0453	1,0407
1,495	1,2227	1,1434	1,1058	1,0837	1,0693	1,0591	1,0516	1,0457	1,0410

(K)	$\sqrt{}$	$\sqrt[3]{}$	$\sqrt[4]{}$	$\sqrt[5]{}$	$\sqrt[6]{}$	$\sqrt[7]{}$	$\sqrt[8]{}$	$\sqrt[9]{}$	$\sqrt[10]{}$
1,500	1,2247	1,1447	1,1067	1,0845	1,0699	1,0596	1,0520	1,0461	1,0414
1,505	1,2268	1,1460	1,1076	1,0852	1,0705	1,0601	1,0524	1,0465	1,0417
1,510	1,2288	1,1473	1,1085	1,0859	1,0711	1,0606	1,0529	1,0469	1,0421
1,515	1,2309	1,1485	1,1094	1,0866	1,0717	1,0611	1,0533	1,0472	1,0424
1,520	1,2329	1,1498	1,1104	1,0873	1,0723	1,0616	1,0537	1,0476	1,0428
1,525	1,2349	1,1510	1,1113	1,0881	1,0729	1,0621	1,0542	1,0480	1,0431
1,530	1,2369	1,1523	1,1122	1,0888	1,0735	1,0626	1,0546	1,0484	1,0434
1,535	1,2390	1,1535	1,1131	1,0895	1,0740	1,0631	1,0550	1,0488	1,0438
1,540	1,2410	1,1548	1,1140	1,0902	1,0746	1,0636	1,0555	1,0491	1,0441
1,545	1,2430	1,1560	1,1149	1,0909	1,0752	1,0641	1,0559	1,0495	1,0445
1,550	1,2450	1,1573	1,1158	1,0916	1,0758	1,0646	1,0563	1,0499	1,0448
1,555	1,2470	1,1585	1,1167	1,0923	1,0764	1,0651	1,0567	1,0503	1,0451
1,560	1,2490	1,1598	1,1176	1,0930	1,0769	1,0656	1,0572	1,0507	1,0455
1,565	1,2510	1,1610	1,1185	1,0937	1,0775	1,0661	1,0576	1,0510	1,0458
1,570	1,2530	1,1623	1,1194	1,0944	1,0781	1,0666	1,0580	1,0514	1,0461
1,575	1,2550	1,1635	1,1203	1,0951	1,0786	1,0670	1,0584	1,0518	1,0465
1,580	1,2570	1,1647	1,1212	1,0958	1,0792	1,0675	1,0588	1,0521	1,0468
1,585	1,2590	1,1659	1,1220	1,0965	1,0798	1,0680	1,0593	1,0525	1,0471
1,590	1,2610	1,1672	1,1229	1,0972	1,0804	1,0685	1,0597	1,0529	1,0475
1,595	1,2629	1,1684	1,1238	1,0979	1,0809	1,0690	1,0601	1,0532	1,0478
1,600	1,2649	1,1696	1,1247	1,0986	1,0815	1,0694	1,0605	1,0536	1,0481
1,605	1,2669	1,1708	1,1256	1,0992	1,0820	1,0699	1,0609	1,0540	1,0484
1,610	1,2689	1,1720	1,1264	1,0999	1,0826	1,0704	1,0613	1,0543	1,0488
1,615	1,2708	1,1733	1,1273	1,1006	1,0832	1,0709	1,0617	1,0547	1,0491
1,620	1,2728	1,1745	1,1282	1,1013	1,0837	1,0713	1,0622	1,0551	1,0494
1,625	1,2748	1,1757	1,1291	1,1020	1,0843	1,0718	1,0626	1,0554	1,0497
1,630	1,2767	1,1769	1,1299	1,1026	1,0848	1,0723	1,0630	1,0558	1,0501
1,635	1,2787	1,1781	1,1308	1,1033	1,0854	1,0728	1,0634	1,0561	1,0504
1,640	1,2806	1,1793	1,1316	1,1040	1,0859	1,0732	1,0638	1,0565	1,0507
1,645	1,2826	1,1805	1,1325	1,1047	1,0865	1,0737	1,0642	1,0569	1,0510
1,650	1,2845	1,1817	1,1334	1,1053	1,0870	1,0742	1,0646	1,0572	1,0514
1,655	1,2865	1,1829	1,1342	1,1060	1,0876	1,0746	1,0650	1,0576	1,0517
1,660	1,2884	1,1840	1,1351	1,1067	1,0881	1,0751	1,0654	1,0579	1,0520
1,665	1,2903	1,1852	1,1359	1,1073	1,0887	1,0756	1,0658	1,0583	1,0523
1,670	1,2923	1,1864	1,1368	1,1080	1,0892	1,0760	1,0662	1,0586	1,0526
1,675	1,2942	1,1876	1,1376	1,1087	1,0898	1,0765	1,0666	1,0590	1,0529
1,680	1,2961	1,1888	1,1385	1,1093	1,0903	1,0769	1,0670	1,0593	1,0532
1,685	1,2981	1,1900	1,1393	1,1100	1,0909	1,0774	1,0674	1,0597	1,0536
1,690	1,3000	1,1911	1,1402	1,1107	1,0914	1,0778	1,0678	1,0600	1,0539
1,695	1,3019	1,1923	1,1410	1,1113	1,0919	1,0783	1,0682	1,0604	1,0542
1,700	1,3038	1,1935	1,1419	1,1120	1,0925	1,0788	1,0686	1,0607	1,0545
1,705	1,3058	1,1947	1,1427	1,1126	1,0930	1,0792	1,0690	1,0611	1,0548
1,710	1,3077	1,1958	1,1435	1,1133	1,0935	1,0797	1,0694	1,0614	1,0551
1,715	1,3096	1,1970	1,1444	1,1139	1,0941	1,0801	1,0698	1,0618	1,0554

(K)	$\sqrt{}$	$\sqrt[3]{}$	$\sqrt[4]{}$	$\sqrt[5]{}$	$\sqrt[6]{}$	$\sqrt[7]{}$	$\sqrt[8]{}$	$\sqrt[9]{}$	$\sqrt[10]{}$
1,720	1,3115	1,1981	1,1452	1,1146	1,0946	1,0806	1,0701	1,0621	1,0557
1,725	1,3134	1,1993	1,1460	1,1152	1,0951	1,0810	1,0705	1,0625	1,0560
1,730	1,3153	1,2005	1,1469	1,1159	1,0957	1,0815	1,0709	1,0628	1,0563
1,735	1,3172	1,2016	1,1477	1,1165	1,0962	1,0819	1,0713	1,0631	1,0566
1,740	1,3191	1,2028	1,1485	1,1171	1,0967	1,0823	1,0717	1,0635	1,0570
1,745	1,3210	1,2039	1,1493	1,1178	1,0972	1,0828	1,0721	1,0638	1,0573
1,750	1,3229	1,2051	1,1502	1,1184	1,0978	1,0832	1,0725	1,0642	1,0576
1,755	1,3248	1,2062	1,1510	1,1191	1,0983	1,0837	1,0728	1,0645	1,0579
1,760	1,3266	1,2074	1,1518	1,1197	1,0988	1,0841	1,0732	1,0648	1,0582
1,765	1,3285	1,2085	1,1526	1,1203	1,0993	1,0845	1,0736	1,0652	1,0585
1,770	1,3304	1,2096	1,1534	1,1210	1,0998	1,0850	1,0740	1,0655	1,0588
1,775	1,3323	1,2108	1,1542	1,1216	1,1004	1,0854	1,0744	1,0658	1,0591
1,780	1,3342	1,2119	1,1551	1,1222	1,1009	1,0859	1,0747	1,0662	1,0594
1,785	1,3360	1,2131	1,1559	1,1229	1,1014	1,0863	1,0751	1,0665	1,0597
1,790	1,3379	1,2142	1,1567	1,1235	1,1019	1,0867	1,0755	1,0668	1,0599
1,795	1,3398	1,2153	1,1575	1,1241	1,1024	1,0872	1,0759	1,0672	1,0602
1,800	1,3416	1,2164	1,1583	1,1247	1,1029	1,0876	1,0762	1,0675	1,0605
1,805	1,3435	1,2176	1,1591	1,1254	1,1034	1,0880	1,0766	1,0678	1,0608
1,810	1,3454	1,2187	1,1599	1,1260	1,1039	1,0885	1,0770	1,0681	1,0611
1,815	1,3472	1,2198	1,1607	1,1266	1,1045	1,0889	1,0774	1,0685	1,0614
1,820	1,3491	1,2209	1,1615	1,1272	1,1050	1,0893	1,0777	1,0688	1,0617
1,825	1,3509	1,2220	1,1623	1,1279	1,1055	1,0897	1,0781	1,0691	1,0620
1,830	1,3528	1,2232	1,1631	1,1285	1,1060	1,0902	1,0785	1,0695	1,0623
1,835	1,3546	1,2243	1,1639	1,1291	1,1065	1,0906	1,0788	1,0698	1,0626
1,840	1,3565	1,2254	1,1647	1,1297	1,1070	1,0910	1,0792	1,0701	1,0629
1,845	1,3583	1,2265	1,1655	1,1303	1,1075	1,0914	1,0796	1,0704	1,0632
1,850	1,3601	1,2276	1,1663	1,1309	1,1080	1,0919	1,0799	1,0707	1,0635
1,855	1,3620	1,2287	1,1670	1,1315	1,1085	1,0923	1,0803	1,0711	1,0637
1,860	1,3638	1,2298	1,1678	1,1321	1,1090	1,0927	1,0807	1,0714	1,0640
1,865	1,3657	1,2309	1,1686	1,1328	1,1095	1,0931	1,0810	1,0717	1,0643
1,870	1,3675	1,2320	1,1694	1,1334	1,1100	1,0935	1,0814	1,0720	1,0646
1,875	1,3693	1,2331	1,1702	1,1340	1,1105	1,0940	1,0817	1,0723	1,0649
1,880	1,3711	1,2342	1,1710	1,1346	1,1109	1,0944	1,0821	1,0727	1,0652
1,885	1,3730	1,2353	1,1717	1,1352	1,1114	1,0948	1,0825	1,0730	1,0654
1,890	1,3748	1,2364	1,1725	1,1358	1,1119	1,0952	1,0828	1,0733	1,0657
1,895	1,3766	1,2375	1,1733	1,1364	1,1124	1,0956	1,0832	1,0736	1,0660
1,900	1,3784	1,2386	1,1741	1,1370	1,1129	1,0960	1,0835	1,0739	1,0663
1,905	1,3802	1,2396	1,1748	1,1376	1,1134	1,0964	1,0839	1,0742	1,0666
1,910	1,3820	1,2407	1,1756	1,1382	1,1139	1,0969	1,0842	1,0745	1,0668
1,915	1,3838	1,2418	1,1764	1,1388	1,1144	1,0973	1,0846	1,0749	1,0671
1,920	1,3856	1,2429	1,1771	1,1394	1,1149	1,0977	1,0850	1,0752	1,0674
1,925	1,3874	1,2440	1,1779	1,1400	1,1153	1,0981	1,0853	1,0755	1,0677
1,930	1,3892	1,2450	1,1787	1,1405	1,1158	1,0985	1,0857	1,0758	1,0680
1,935	1,3910	1,2461	1,1794	1,1411	1,1163	1,0989	1,0860	1,0761	1,0682

(K)	$\sqrt{}$	$\sqrt[3]{}$	$\sqrt[4]{}$	$\sqrt[5]{}$	$\sqrt[6]{}$	$\sqrt[7]{}$	$\sqrt[8]{}$	$\sqrt[9]{}$	$\sqrt[10]{}$
1,940	1,3928	1,2472	1,1802	1,1417	1,1168	1,0993	1,0864	1,0764	1,0685
1,945	1,3946	1,2483	1,1809	1,1423	1,1173	1,0997	1,0867	1,0767	1,0688
1,950	1,3964	1,2493	1,1817	1,1429	1,1177	1,1001	1,0871	1,0770	1,0691
1,955	1,3982	1,2504	1,1825	1,1435	1,1182	1,1005	1,0874	1,0773	1,0693
1,960	1,4000	1,2515	1,1832	1,1441	1,1187	1,1009	1,0878	1,0776	1,0696
1,965	1,4018	1,2525	1,1840	1,1446	1,1192	1,1013	1,0881	1,0779	1,0699
1,970	1,4036	1,2536	1,1847	1,1452	1,1196	1,1017	1,0884	1,0782	1,0702
1,975	1,4053	1,2546	1,1855	1,1458	1,1201	1,1021	1,0888	1,0786	1,0704
1,980	1,4071	1,2557	1,1862	1,1464	1,1206	1,1025	1,0891	1,0789	1,0707
1,985	1,4089	1,2568	1,1870	1,1470	1,1211	1,1029	1,0895	1,0792	1,0710
1,990	1,4107	1,2578	1,1877	1,1475	1,1215	1,1033	1,0898	1,0795	1,0712
1,995	1,4124	1,2589	1,1885	1,1481	1,1220	1,1037	1,0902	1,0798	1,0715
2,000	1,4142	1,2599	1,1892	1,1487	1,1225	1,1041	1,0905	1,0801	1,0718

Навчальне видання

Краєвський Володимир Миколайович,
Остапенко Яна Олександрівна,
Паянок Тетяна Миколаївна,
Параниця Надія Володимирівна

МЕТОДИ АНАЛІЗУ ДАНИХ У ПСИХОЛОГІЧНИХ ДОСЛІДЖЕННЯХ

Навчальний посібник

Відповідальний за випуск *А. В. Лавренюк*

Відповідальний редактор *М. М. Грабарчук*
Редактор *М. П. Клименко*

Форматування та
комп'ютерна верстка *Д. П. Завальніцька*

Здано до друку 29.11.2024. Формат 60 × 84 / 14
Папір офсетний № 1. Гарнітура «Times New Roman»
Ум. друк. арк. 8.32
Наклад 300 прим. Замовлення № 1118

Підготовлено до друку редакційно-видавничим відділом
Державного податкового університету
08205, вул. Університетська, 31, м. Ірпінь, Київська область,
Україна

*Свідоцтво про внесення суб'єкта видавничої справи
до державного реєстру видавців, виготовлювачів і
розповсюджувачів видавничої продукції
Серія ДК № 7669 від 20.09.2022*