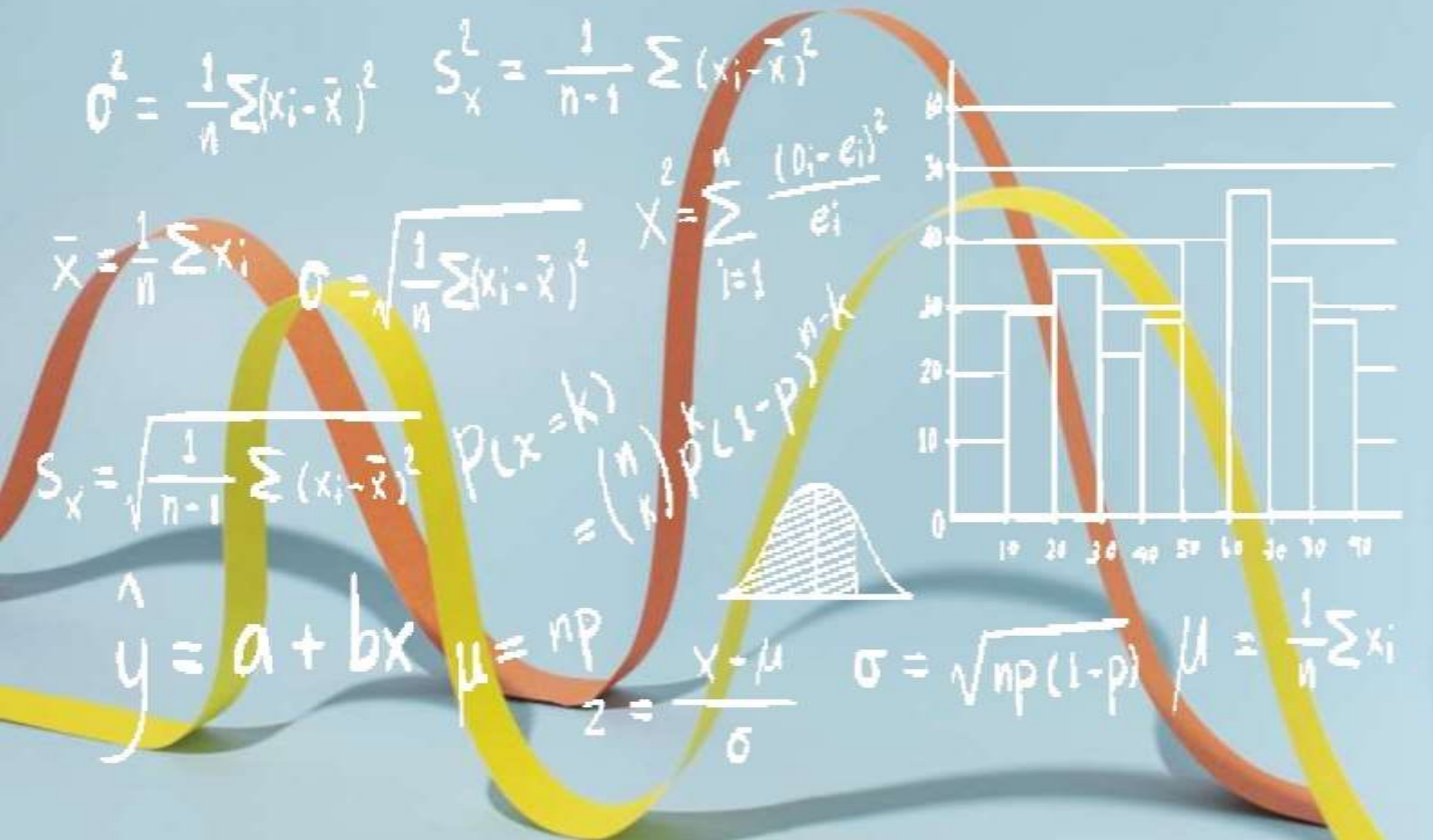


ЗАСТОСУВАННЯ МЕТОДІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ У ПСИХОЛОГО-ПЕДАГОГІЧНИХ ДОСЛІДЖЕННЯХ



НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ОБОРОНИ УКРАЇНИ
Кафедра суспільних наук

**ЗАСТОСУВАННЯ МЕТОДІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ
У ПСИХОЛОГО-ПЕДАГОГІЧНИХ ДОСЛІДЖЕННЯХ**

Видання університету – 2023

УДК 159.96:378.6:519.23

Затверджено Вченою радою Національного університету оборони України для використання у навчальному процесі (протокол № 11 від 28 листопада 2022 року)

Рецензенти: *Миронець С.М.*, доктор психологічних наук, доцент, завідувач кафедри психології Державного торговельно-економічного університету;
Капосльоз Г.В., кандидат психологічних наук, старший науковий співробітник, заступник начальника науково-дослідного відділу прикладних військово-соціологічних досліджень Науково-дослідного центру гуманітарних проблем Збройних Сил України, підполковник.

Москальов І.О., Лисенко Д.П. Застосування методів математичної статистики у психолого-педагогічних дослідженнях: навч. посіб. Київ : НУОУ, 2023. 187 с.

Автори: *Москальов І.О.*, доктор філософії у галузі соціальних та поведінкових наук зі спеціальності 053 Психологія, старший викладач кафедри суспільних наук гуманітарного інституту Національного університету оборони України, полковник;
Лисенко Д.П., доктор філософії у галузі соціальних та поведінкових наук зі спеціальності 053 Психологія, старший науковий співробітник науково-дослідної лабораторії кафедри морально-психологічного забезпечення діяльності військ (сил) гуманітарного інституту Національного університету оборони України, підполковник.

Навчальний посібник розроблений для використання в освітньому процесі під час підготовки на другому і третьому рівнях. Містить практичні основи застосування методів математичної статистики у психології та педагогіці, рекомендації з обробки даних експериментальних досліджень. У посібнику наведено правила графічного відображення статистичної інформації, обчислення середнього арифметичного, моди, медіани, дисперсії. Теоретичний матеріал супроводжується великою кількістю професійно спрямованих прикладів. Навчальний посібник спрямований на набуття вмінь користування математико-статистичними методами обробки інформації у психологічних дослідженнях. Описано основи кореляційного аналізу, роботу з методами статистичного висновку. Книга містить практикум зі статистичної обробки даних на базі програми SPSS.

Для здобувачів другого і третього рівнів освіти за спеціальностями 011 «Освітні, педагогічні науки» та 053 «Психологія», психологів-дослідників, педагогів і всіх, хто проводить психологічні дослідження.

© Москальов І.О., Лисенко Д.П., 2023

© Національний університет оборони України, 2023

ЗМІСТ

ПЕРЕДМОВА. ПСИХОЛОГІЧНИЙ ГОРИЗОНТ ЗАСТОСУВАННЯ МЕТОДІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ У ПСИХОЛОГІЇ	5
ВСТУП	8
РОЗДІЛ 1. ПРОБЛЕМА ВИМІРЮВАННЯ У НАУКОВИХ ДОСЛІДЖЕННЯХ	11
1.1. Основні завдання та методи математичної статистики.	11
1.2. Сутність вимірювань у психологічних та педагогічних дослідженнях.	18
1.3. Типи і характеристики вимірювальних шкал.	23
РОЗДІЛ 2. ПОНЯТТЯ ВИБІРКИ В НАУКОВИХ ДОСЛІДЖЕННЯХ	33
2.1. Генеральна сукупність та вибірка дослідження.	33
2.2. Вимоги до вибірки.	35
2.3. Методи формування вибірки.	39
РОЗДІЛ 3. ПЕРВИННА ОПИСОВА СТАТИСТИКА	41
3.1. Міри центральної тенденції.	41
3.2. Міри мінливості.	46
3.3. Міри положення.	48
3.4. Міри зсуву.	50
РОЗДІЛ 4. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ	55
4.1. Поняття статистичної гіпотези.	55
4.2. Рівні статистичної значущості.	57
4.3. Логіка та процедура перевірки гіпотез.	58

4.4. Статистичні критерії розрізень.	61
РОЗДІЛ 5. МЕТОДИ ПОРІВНЯННЯ ВИБІРОК ДОСЛІДЖУВАНИХ	69
5.1. Оцінювання відповідності емпіричного закону нормальному розподілу.	69
5.2. Оцінювання достовірності зсуву.	77
5.3. Оцінювання достовірності відмінностей.	91
РОЗДІЛ 6. ОСНОВИ КОРЕЛЯЦІЙНОГО АНАЛІЗУ	114
6.1. Поняття про кореляцію.	114
6.2. Моделі зв'язку у шкалах інтервалів і відношень.	119
6.3. Моделі зв'язку у порядкових вимірюваннях.	124
РОЗДІЛ 7. СПОСОБИ ВІЗУАЛІЗАЦІЇ І ПОДАННЯ ДАНИХ	138
7.1. Основні правила візуалізації даних.	138
7.2. Таблиця, як техніка візуалізації.	141
7.3. Графічні способи подання даних.	144
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	152
ДОДАТКИ	155
1. Статистичні таблиці критичних значень.	155
2. Таблиця вибору техніки візуалізації.	171
ГЛОСАРІЙ	176

ПСИХОЛОГІЧНИЙ ГОРИЗОНТ ЗАСТОСУВАННЯ МЕТОДІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ У ПСИХОЛОГІЇ

Ситуація в якій перебуває Україна з моменту отримання статусу незалежної держави потребує підготовки високоосвічених професійних кадрів не тільки для різноманітних сфер економіки, бізнесу, але й для Збройних Сил України, особливо це питання постало гостро з часів військового вторгнення «рф» на територію України, незаконної анексії АР Крим, територій Луганського та Донецького регіону, а з лютого 2022 року повномасштабної військової агресії проти України.

Сучасна психологія потребує доказовості й має бути орієнтованою на розв'язання актуальних проблем, що постають перед особистістю в умовах кризового суспільства, особливою ланкою у цій ситуації стають фахівців ризиконебезпечних професій, зокрема військовослужбовці Збройних Сил України та інших військових формувань.

Загальним для всіх психологів–практиків, незалежно від їх спеціалізації, слугують актуальні знання психодіагностики, психологічного прогнозування й управління. Психодіагностика, психологічне прогнозування, психологічне управління є не тільки функціями психологічної служби, але й дозволяють розв'язувати взаємозалежні задачі, й послідовні етапи практичної й науково-дослідної діяльності у всіх галузях психології. Саме тому вони дозволяють природнім шляхом структурувати різноманіття математико-психологічних моделей і методів, як у психологічному, так і в математичному аспектах.

З огляду на зазначене – навчальний посібник «Застосування методів математичної статистики у психолого-педагогічних дослідженнях», підготовлений авторами Ігорем Москальовим та Дмитром Лисенком є актуальним, значно доповнить навчально-методичний арсенал в системі підготовки фахівців Збройних Сил України та буде корисним для застосування у навчальному процесі інших закладів вищої освіти. Автори вдало сформували зміст і структуру навчального посібника, яка складається зі вступу, семи розділів, списку використаних джерел та додатків та повною мірою відповідає методичним рекомендаціям МОН

України та Положенням, що визначають порядок розробки і видання такого виду навчально-методичних видань в Національного університету оборони України.

Позитивним є й те, що автори надають уточнений глосарій понять, термінів та визначень в межах тематики навчального видання, що значно підвищує комфортність у використанні даного видання у навчальному процесі.

У першому розділі наведено основний понятійний апарат та визначеннями математичної статистики, а також особливості отримання даних у психолого-педагогічних дослідженнях: а саме багатофакторністю та варіативністю.

Другий розділ висвітлює основні категорії та поняття в межах загально-наукового експериментального методу дослідження та розкриває особливості формування саме дослідження сфери психології, дає розуміння таким категоріям, як типи шкал, спостереження, вибірка, генеральна сукупність, вибіркові статистичні характеристики, порядкова статистика, а також методи їх оцінювання.

Третій розділ дає уявлення про формування основних параметрів та величин описової статистики: міри центральної тенденції, міри мінливості, міри положення. Та тим, як відповідно до оцінювання цих параметрів, робити висновки стосовно закону розподілу даних. Наявність інформації про закон розподілу даних вимірювань дозволяє коректно підсумовувати похибки або складові невизначеності вимірювань та визначати відповідні довірчі інтервали обґрунтовувати достовірність результатів експерименту.

Четвертий та п'ятий розділи присвячені аналітичній статистиці: правилам висунення та перевірки статистичних гіпотез, визначенню рівнів статистичної значущості та опису основних статичних методів.

У шостому розділі наведено основні процедури кореляційного аналізу.

Сьомий розділ присвячено документуванню та представленню результатів наукової звітності, авторами запропоновані до використання деякі програмні продукти для створення якісної інфографіки.

Отже можемо констатувати, що навчальний посібник повною мірою відповідає вимогам стандартів вищої освіти з підготовки фахівців зі спеціальності 053 – Психологія та знайде широкий схвальний відгук серед психологів практиків та фахівців, що відповідають за формування управлінських рішень на оперативно-тактичному та стратегічному рівнях військового управління.

Завідувач кафедри психології

*Державного торговельно-економічного
університету, доктор психологічних наук*

Сергій МИРОНЕЦЬ

ВСТУП

Розвиток психології та педагогіки на сучасному етапі супроводжується стрімким розширенням галузей застосування результатів вимірювань. Дослідження й експерименти, психодіагностика, підготовка наукових статей та виступів на конференціях – із цим далеко не вичерпним переліком робіт зіштовхуються всі наукові та науково-педагогічні працівники, здобувачі наукового ступеня, психологи і педагоги. У всіх цих випадках потрібно об'єктивно довести свою точку зору, переконати в її істинності, адже здебільшого науковець може обґрунтувати власну точку зору, лише посилаючись на результати конкретних досліджень. Зробити це він може лише в разі уміння адекватно використовувати математичний апарат для опису, обробки та аналізу емпіричних даних. Одним із основних шляхів забезпечення належної якості результатів вимірювань є коректне застосування методів математичної статистики та опрацювання даних вимірювань.

Під час розроблення цього посібника автори прагнули систематично викласти основні положення та процедури математичної статистики, застосовані для вирішення широкого кола завдань, пов'язаних із вимірювальним експериментом. Навчальний посібник має чітку практичну спрямованість, майже всі положення статистичного опрацювання даних проілюстровано числовими прикладами і детальними поясненнями. Для кращого розуміння і засвоєння матеріалу та набуття практичних навичок у кожному розділі передбачено контрольні запитання і завдання. Посібник дає змогу опанувати сучасні методи статистичного аналізу даних вимірювань на рівні, достатньому для самостійного застосування цих методів у практичній та науковій діяльності.

Матеріали посібника структуровані у сім розділів, які охоплюють основні поняття, використовувані у психологічній та педагогічній практиці під час опрацювання експериментальних даних.

Перший розділ посібника є суто теоретичним та ознайомлює з основними поняттями та визначеннями математичної статистики, а також особливостями отримання даних у психолого-педагогічних дослідженнях:

багатофакторністю і варіативністю. У другому розділі розглянуто основні категорії та поняття такі як типи шкал, спостереження, вибірка, генеральна сукупність, вибіркові статистичні характеристики, порядкова статистика, а також методи їх оцінювання.

Третій розділ ознайомлює з основними параметрами та величинами описової статистики: мірами центральної тенденції, мірами мінливості, мірами положення, а також тим, як за оцінюванням цих параметрів робити висновки стосовно закону розподілу даних. Наявність інформації про закон розподілу даних вимірювань дає змогу коректно підсумовувати похибки або складові невизначеності вимірювань, визначати відповідні довірчі інтервали, обґрунтовувати достовірність результатів експерименту.

Четвертий і п'ятий розділи присвячені аналітичній статистиці: правилам висунення та перевірки статистичних гіпотез, визначенню рівнів статистичної значущості та опису основних статичних методів.

Розглянуті у п'ятому розділі методи поділено на дві групи: методи з групуванням вибірових даних та без групування. Методи з групуванням даних ґрунтуються на аналізі варіаційних рядів. Як засоби аналізу використовують емпіричний закон розподілу, гістограму або полігон частот. Методи без групування даних ґрунтуються на спеціалізованих (до конкретного закону розподілу) параметричних критеріях згоди.

У шостому розділі наведено основні процедури кореляційного аналізу. Кореляційний аналіз застосовують для визначення наявності і ступеня стохастичного взаємозв'язку між складовими похибки або невизначеності вимірювань, виявлення впливу різних чинників на процес вимірювання; обґрунтування рівнянь непрямих вимірювань тощо. У розділі розглянуто параметричні та непараметричні методи кореляційного аналізу.

Сьомий розділ присвячено документуванню і поданню експериментальних даних, детальному розкриттю основних правил табличної та графічної візуалізації результатів експерименту.

Проаналізовано основні помилки, а також запропоновано кілька програмних продуктів для створення якісної інфографіки.

Автори сподіваються, що посібник буде корисним насамперед для здобувачів освітньо-наукового ступеня доктора філософії, а також курсантів, студентів, слухачів, наукових та науково-педагогічних працівників, які працюють у галузі психології та педагогіки.

РОЗДІЛ 1

ПРОБЛЕМА ВИМІРЮВАННЯ У НАУКОВИХ ДОСЛІДЖЕННЯХ

- 1.1. Основні завдання та методи математичної статистики.*
- 1.2. Сутність вимірювань у психологічних та педагогічних дослідженнях.*
- 1.3. Типи і характеристики вимірювальних шкал.*

1.1. Основні завдання та методи математичної статистики

В умовах широкого застосування методів сучасної математики в усіх галузях наукових досліджень, фундаментальних і прикладних, а також у вирішенні низки практичних проблем суспільного життя особливу увагу надають математичній статистиці. Як галузь математичних знань статистика базується на теорії ймовірності і є наукою про методи формулювання умовиводу про властивості дослідженої статистичної сукупності. Математична статистика гармонійно поєднана із загальною науковою методологією, інтерпретацією явищ з позицій їх діалектичного розвитку та особливими методами спеціальних галузей статистичної науки. Пропонуючи математичну техніку для вивчення імовірнісних властивостей досліджуваних явищ і процесів, статистика стає методом спеціальних наук, а її математичний апарат плідно застосовують під час вивчення явищ і процесів, що відбуваються в житті суспільства.

На цей час математичну статистику широко застосовують у психології, економіці, фізиці, хімії, медицині та інших науках. На основі її методів можна розв'язувати багато аналітичних задач у галузі психології та педагогіки. Зокрема, кількісні характеристики, одержані за результатами математично-статистичного аналізу, дають змогу мати глибше уявлення про причинно-наслідкові зв'язки у психологічних явищах та процесах, а також одержувати стійкі надійні параметри для проведення розрахунків і, насамперед, прогнозування та формулювання висновків. Впровадження комп'ютерних технологій створює реальні можливості для широкого

впровадження методів математичної статистики у вирішенні різних завдань.

Як окрема наукова дисципліна математична статистика виникла у другій половині XIX ст. До цього часу значні успіхи вже були досягнуті в теорії ймовірностей, яка стала теоретичною основою математичної статистики. Поштовхом до її розвитку були експериментальні дослідження в галузі наук про природу.

Математична статистика є складовою статистичної науки взагалі, її специфічним методом дослідження. Відповідаючи на питання різних наук, математична статистика містить найбагатший арсенал математико-статистичних прийомів обробки емпіричних даних. При цьому у ній гармонійно поєднані наукова методологія, інтерпретація явищ на основі філософії і специфічні методи статистичної науки.

Для сучасних наукових досліджень характерне взаємопроникнення різних наук. Математична статистика є науковою дисципліною, метод якої застосовують для потреб інших наук. Однак не слід ототожнювати поняття «статистичні методи» і «математико-статистичні методи». Як математико-статистичні, розуміють методи безпосередньо пов'язані із ймовірнісним оцінюванням результатів статистичного спостереження, що передбачають визначення величини математичної ймовірності. Іншими словами, якщо мова йде не про описові функції статистики, то в усякому статистичному висновку визначальною є теорія ймовірностей. Математична статистика з її ймовірнісною методологією породжує навіть деякі галузі теоретичних знань, відокремлені від певних наук.

Предмет математичної статистики – це формальна математична сторона статистичних методів дослідження, байдужа до специфічної природи об'єктів, які вивчають. Згідно з цим визначенням предмет математичної статистики є чисто математичною теорією математико-статистичних методів незалежно від специфіки і сфери їх застосування.

Характеризуючи математичну статистику як науку, яка розробляє методи одержання, опису й обробки даних статистичних спостережень для вивчення закономірностей випадкових масових явищ, зазначимо, що

категорія «статистична закономірність» нерозривно пов'язана з основною категорією теорії статистики – «статистичною сукупністю». Оскільки надалі у викладенні різних питань цю категорію будемо застосовувати часто, надамо її визначення.

Статистична сукупність – це сукупність однорідних об'єктів чи явищ, об'єднаних за певними ознаками у єдине ціле. Окремі одиниці статистичних сукупностей відрізняються між собою. Тому виникає необхідність застосування деяких описуваних параметрів статистичної сукупності – **статистичних характеристик**. Найважливішими з них є середні величини (арифметична, квадратична, мода, медіана та ін.) міри мінливості (дисперсія, середнє квадратичне відхилення та ін.), міри зсуву, моменти сукупності і низка інших.

Завдання математичної статистики в цілому пов'язані з вирішенням питань обробки даних статистичних спостережень. Залежно від особливостей статистичного виміру явищ, які вивчають, і мети аналізу їм надають певної форми.

Основні **завдання** математичної статистики можна об'єднати у три великі групи:

Перша група завдань стосується встановлення законів розподілу різних випадкових змінних, одержаних у результаті статистичного спостереження. Оскільки аналізована сукупність є не повним обсягом даних вибіркою і результатами, одержаними на їх основі, зумовленими елементами випадковості, потрібно знати, які саме риси досліджуваних явищ стійкі, а які – випадкові, адже дані беруть не в повному обсязі. Вирішення цього завдання можливе за умови правильного вибору методів обробки даних.

Використовувані методи, мають сприяти встановленню і збереженню характерних (типових) рис явища. Отже, ця група завдань математичної статистики містить певну систему методів систематизації і перетворення даних статистичного спостереження.

Друга група завдань стосується перевірки статистичних гіпотез. Маючи визначену сукупність даних (як правило, невелику за обсягом),

дослідник зобов'язаний висунути гіпотезу про закономірності, притаманні явищу, яке вивчає. Висунуту гіпотезу необхідно перевірити, тобто з'ясувати чи підтверджують дані спостереження гіпотезу про те, що середня їх величина дорівнюватиме відповідній середній для всієї сукупності, із якої проведена вибірка.

Одне з основних завдань цієї групи – перевірка гіпотез стосовно законів розподілу, тобто з'ясування чи підтверджують дані вибірки гіпотезу про те, що досліджуване явище підпорядковане закону нормального розподілу (чи будь-якому іншому закону). У математичній постановці ця задача може бути сформульована так: на підставі деяких передбачень можна вважати, що функція розподілу досліджуваної випадкової змінної величини $x \in f(x)$. Чи збігаються спостережувані значення з гіпотезою, якщо випадкова величина x дійсно має розподіл $f(x)$.

До *третьої групи* завдань, належить оцінювання невідомих параметрів різних розподілів. Необхідність вирішення цих завдань можна обґрунтувати так. Оскільки дослідник працює не з усією сукупністю одиниць явища, яке вивчає, а тільки з їх частиною (вибіркою), рівень деяких статистичних характеристик для всієї сукупності залишається невідомим (наприклад, середнє, дисперсія та ін.). У цьому разі варто застосувати специфічний метод для оцінювання характеристики, одержаної за даними вибіркового спостереження. Математична статистика має у своєму розпорядженні цілий арсенал методів для вирішення завдань оцінювання розподілу, а також їх надійності (точності).

У математичній постановці задачу оцінювання невідомих параметрів розподілу можна сформулювати так: випадкова змінна x має функцію розподілу певного виду, зумовлену деякими параметрами з невідомими значеннями. За даними спостереження необхідно оцінити ці параметри.

Статистична методологія – це сукупність прийомів, правил і методів дослідження. Як метод розуміють спосіб теоретичного дослідження або практичного здійснення чого-небудь (наприклад, філософський метод, евристичний метод). Коли мова йде про метод науки взагалі, то мають на увазі найбільш загальні способи підходу до вивчення будь-яких явищ. Як

метод конкретної науки розуміють окремі специфічні прийоми і методи, пристосовані до дослідження її предмета. Загальнонауковим вважається діалектичний метод. Керуючись його принципами, статистика розробляє свої специфічні прийоми і методи, а також відпрацьовує систему показників, яка дає змогу вивчати кількісні сторони суспільних явищ і процесів.

Особливість статистичних методів полягає в їх комплексності, що зумовлено як різноманітністю форм статистичних закономірностей, так і складністю самого процесу статистичного дослідження. Специфіка методів пояснюється змістом виконуваної роботи у процесі дослідження певних соціально-економічних явищ, природа яких досить складна і непередбачувана, тому вивчати їх треба взаємопов'язано і взаємообумовлено. Для цього у статистичній теорії розроблено широке коло методологічних і методичних засобів, які дають змогу кількісно вимірювати досліджувані зв'язки (насамперед причинно-наслідкові).

Філософський підхід як методологічна основа статистичної науки передбачає розгляд явищ і процесів в їх русі, постійних змінах і розвитку. З цією метою розроблена відповідна система показників, які дають змогу охарактеризувати варіації змін у рівнях явищ, визначати тенденції і закономірності їх розвитку. Важливим моментом філософського підходу до вивчення суспільних явищ є визначення меж переходу кількісних явищ у якісні форми їх прояву. Філософська основа статистики втілюється в її специфічних методах.

Методи статистичного спостереження застосовуються на першому етапі статистичного дослідження для збирання та оцінювання їх якості. Через масовість зібраних даних застосовують метод масового статистичного спостереження - спостереження над множиною елементів, які становлять статистичну сукупність. Вивчаючи сукупності за допомогою масових спостережень, статистика виявляє притаманні їм загальні риси, процеси, закономірності. Робота ця досить кропітка і складна, тому потребує наукової організації. До масових спостережень належать обстеження, збирання звітності, переписи та ін.

Методи зведення і групування первинного статистичного матеріалу, які застосовують на другому етапі дослідження, поєднують методи перевірки, систематизації, обробки, підсумовування даних і подання їх у формі зручної інфографіки.

Зведення забезпечує систематизацію первинної інформації, підрахунок кількості одиниць сукупності й обсягу ознак, що їх характеризують. Важливим етапом цього етапу дослідження є розподіл інформації (розчленування) за ознаками її відмінності, тобто групування статистичних даних. Специфіка методів групування зумовлена завданнями дослідження і якістю первинної інформації. Методи групування дають змогу розділити досліджувану сукупність на однорідні за певними ознаками в певному відношенні частини.

Методи (прийоми) визначення узагальнюючих зведених синтетичних показників застосовують на третьому етапі статистичного дослідження для вирішення завдання визначення певних параметрів.

Етап узагальнення й аналізу зведеного матеріалу передбачає виявлення характерних властивостей і закономірностей досліджуваних явищ, взаємозалежностей факторних та результативних ознак і т. ін. На цьому етапі використовують весь арсенал статистичних прийомів і методів дослідження, розраховуються узагальнюючі статистичні показники: абсолютні, відносні й середні величини. Окремі загальні риси формування названих вище показників визначають вимірюванням їх варіацій, що дає змогу характеризувати умови перебігу масових процесів, для чого досліджують закони розподілу.

Тенденції і закономірності у русі показників у часовому розрізі вивчають дослідженням рядів динаміки, спеціальних (математичних) прийомів їх обробки і моделювання.

Причинно-наслідкові зв'язки психологічних явищ досліджують методами аналітичних групувань, кореляційного, регресійного і дисперсійного аналізу. Аналітичні можливості статистичних методів значно розширюються в разі умілого використання статистичних графіків

(графо-аналітичного методу), номограм, табличного, балансового і багатовимірних методів аналізу.

Зазначені методи і прийоми досліджень та принципи їх використання у статистичному аналізі становлять предмет курсу загальної математичної статистики.

Щодо теоретичних аспектів питання про статистичний метод зазначимо, що він має ту ж основу, що й форми наукового методу, в яких застосовують індукцію і дедукцію. Дослідник спочатку веде спостереження, потім у ході експерименту (чи аналізу) переходить за допомогою індукції і починає будувати теорію. Далі за допомогою дедукції дослідник передбачає результати подальших експериментів, які в інших умовах передбачити було б неможливо. Потім науковець збирає фактичні дані, які підтверджують або спростовують правильність передбачень, і розвиває основну теорію або будує нову. У статистичних дослідженнях наведена спрощена схема реального процесу має свої особливості, адже статистика має справу переважно з невизначеними ситуаціями. Якщо ступінь невизначеності досить значний, результати матимуть обмежену сферу застосування. Наявність невизначеності ситуацій зумовлює у статистичному аналізі нескінченні труднощі, тому вчені-статистики постійно працюють над винайденням нових або удосконаленням наявних методів. Нові вимоги орієнтують на застосування досить складних і витончених методів статистики, зокрема методів математичної статистики.

Значення математики для розвитку статистики зросло останнім часом через широке застосування електронної техніки. Однак це не означає, що використання складного математичного апарату може перетворити статистику в математику, яка досліджує просторові форми та кількісні співвідношення реального світу у «чистому вигляді», а статистика вивчає матеріальний зміст явищ, використовуючи математику як інструмент дослідження. Це дає змогу упорядкувати інформаційні масиви, прискорити їх обробку, передачу та зберігання, а також перевести

фактичні дані на формалізовану мову для проведення масових розрахунків.

Статистичні дослідження завжди цікаві, навіть якщо не доведені до кінця намальованого дослідником шляху (в науці це інколи трапляється), але на самому шляху досліджень пізнано дуже багато.

Майбутнє статистики як науки очевидне, хоча треба завжди пам'ятати, що будь-якій науці притаманні помилки, а тому вона завжди потребує удосконалення. Помилки, які трапляються в статистиці, найчастіше виникають не через недосконалість статистичної науки, а через невміле (неправильне) застосування статистичних методів.

1.2. Сутність вимірювань у психологічних та педагогічних дослідженнях

Початковим етапом математичної обробки результатів будь-якого дослідження, зокрема і психолого-педагогічного, є вимірювання. Інакше кажучи, досліджувана ознака (властивість, риса, характеристика) може бути виміряна, тобто виражена у кількісній (чисельній) формі. Чисельне вираження ознаки може бути різним: від представлення її в бінарній системі (1 – наявність ознаки, 0 – відсутність ознаки) до точних кількісних значень (наприклад, максимальна амплітуда альфа-ритму електроенцефалограми для випробуваного становить 95 мікрвольт).

Одним із складних завдань у психології та педагогіці є завдання математичної формалізації виразності досліджуваних ознак, тобто переведення їх у кількісні вирази. Про це йтиметься далі у відповідних розділах. А в цьому розділі наведемо загальні відомості про вимірювання взагалі та особливості вимірювання психологічних властивостей (ознак, рис, характеристик) зокрема.

Існує низка визначень поняття **вимірювання**. Так, вимірювання іноді трактують як пізнавальний процес, що включає дослідження кількісних характеристик матеріальних об'єктів за допомогою відповідних вимірювальних приладів. Таке формулювання цілком може бути використане для фізичних вимірювань, але не годиться для вимірювання

психологічних величин. Найчастіше, кажучи про процедури психологічного вимірювання, мають на увазі наявність не вимірювальних приладів, а сукупності завдань, питань або тверджень. Проте в деяких галузях психологічної науки (психофізиці, психофізіології та ін.) передбачене використання і приладових (апаратурних) методів вимірювання.

Також визначають вимірювання як присвоєння чисел певним об'єктам, властивостям, ознакам, подіям чи змінам відповідно до певних правил. Це визначення більш доречне для вимірювання у психології, хоча задля справедливості зазначимо, що не всі психологічні величини можна виразити числом – деякі з них виражають якісними визначеннями, назвами, символами та ін.

Нарешті, вимірювання можна визначити як побудову шкал у вигляді ізоморфного відображення емпіричної системи з відносинами до чисельної системи з відносинами, тобто фактично ототожнити поняття вимірювання та шкалювання. На перший погляд, це так, хоча в деяких випадках поняття шкалювання ширше за поняття вимірювання, адже передбачає впорядкування не тільки чисельних (кількісних), а й якісних характеристик.

Будь-який вид вимірювання передбачає наявність певних вимірювальних одиниць. Одиниця вимірювання – це «вимірювальна паличка» (за висловом С. Стівенса), яка є своєрідним еталоном для проведення вимірювальних операцій. У фізиці та інших природничо-наукових дисциплінах використовують основні і похідні одиниці вимірювання. Основних одиниць вимірювання у Міжнародній системі одиниць (SI) небагато: це кілограм (*кг*) – одиниця маси, метр (*м*) – одиниця відстані, секунда (*с*) – одиниця часу, градус Кельвіна (*К*) – одиниця температури, ампер (*А*) – одиниця сили струму, кандела (*кд*) – одиниця сили світла, моль – одиниця кількості речовини. Всі інші одиниці (швидкість, щільність, освітленість, тиск та ін.) є похідними та виводяться з основних одиниць вимірювання.

Крім загальноприйнятих (міжнародних) одиниць вимірювання у наукових дослідженнях дуже обмежено застосовують традиційні (національні) одиниці (фунти, унції, дюйми, ярди, фути та ін.).

Поряд із найбільш поширеною десятковою системою числення, деякі одиниці обчислюють в шістдесятковій системі, започаткованій в Ассирії та Вавилоні. У цій системі обчислюють інтервали часу (години, хвилини), а також кутові міри (кутові градуси, хвилини і секунди). Попри деякі незручності, що виникають у ході переведення одиниць вимірювання з шістдесяткової системи в десяткову, вавилонсько-ассирійська система настільки міцно увійшла в нашу свідомість, що міняти її на сьогодні, мабуть, немає сенсу.

Фізичні одиниці вимірювання використовують у психолого-педагогічних дослідженнях не завжди. У деяких випадках, наприклад у психофізиці, доцільно, щоб суб'єкт оцінював будь-які величини або знаходив ступінь відмінності між ними в загальноприйнятих одиницях (наприклад, проміжки часу – у хвилинах і секундах, довжину ліній – у сантиметрах, відстань до об'єкта – в метрах та ін.). Загальноприйняті одиниці вимірювання застосовуються й у психофізіології. Так, час сенсомоторних реакцій і час розпізнання образів прийнято вимірювати в секундах або мілісекундах, амплітуду викликаних потенціалів – в мікрвольтах, частоту ритмів електроенцефалограми – в кількості коливань за секунду та ін. Проте найчастіше психологи у вимірюваннях користуються умовними одиницями («сирими» балами, стенами тощо). Так, у більшості тестів-опитувальників як одиницю вимірювання застосовують відповіді «так» чи «ні». Досліджувану властивість обчислюють з урахуванням співвідношення цих відповідей (їх суми, різниці тощо). Під час виконання «інтелектуальних» тестів одиницею вимірювання є розв'язок кожної задачі (виконання окремого завдання), а досліджувану ознаку (коефіцієнт інтелекту та ін.) визначають за кількістю виконаних завдань.

Вперше у сучасному трактуванні думку про принципову можливість вимірювання психічних явищ, процесів і станів висловив відомий

німецький філософ Густав Теодор Фехнер (1801 – 1887). У фундаментальній праці «Елементи психофізики» він писав: «...важко заперечити, що духовне взагалі підпорядковане кількісним відносинам. Адже можна говорити не тільки про більшу чи меншу силу відчуття, а й про різну силу потягу, про те, що існує більший чи менший ступінь уваги, жвавості спогадів чи образів фантазії, ясності свідомості загалом, а також інтенсивності окремих думок... Таким чином, вище духовне не меншою мірою, аніж чуттєве... може бути охарактеризовано кількісно».

Попри тривалу полеміку щодо можливості кількісного опису психічних явищ, процесів і станів, що розгорнулася після виходу у світ книги Фехнера, на сьогодні важко уявити психологічну та педагогічну науку без вимірювання. Психофізика, психофізіологія, психометрика, психодіагностика, педагогіка, андрагогіка – ось далеко не повний перелік психологічних та педагогічних дисциплін, у яких вимірювання є необхідним інструментом.

Іноді кажуть, що вимірювання психічних величин, що ґрунтується на суб'єктивних звітах досліджуваних, не варте довіри лише через їх суб'єктивне. Не вдаючись у філософський бік проблеми, можна сказати, що психологічні (педагогічні) вимірювання так само надійні й валідні, як і вимірювання фізичні, але мають свої особливості. Основними властивостями психолого-педагогічного вимірювання є їх багатофакторність і варіативність.

Багатофакторність вимірювання у психології полягає в тому, що на психологічні величини впливає безліч різних чинників, одні з яких (релевантні) безпосередньо пов'язані з ознакою, вимірюваного, а інші (іррелевантні) не пов'язані з нею або пов'язані побічно. Вплив усіх іррелевантних чинників врахувати неможливо. Проте чим більшу їх кількість буде враховано, тим дієвішою можна вважати методіку, більш валідною математичну модель, точнішим психологічний прогноз.

Існує оптимальний, на наш погляд, спосіб подолання труднощів, пов'язаних із багатофакторністю психологічних вимірювань. Так, якщо на вимірювану психологічну ознаку діє велика кількість різноманітних

чинників, то апріорно приймають точку зору, що всі ці різноспрямовані чинники зрештою врівноважують один одного і досліджувана ознака варіює випадковим чином. Відомо, що на принципі випадковості засновано цілу галузь математичної науки – теорію ймовірностей. Не дивно, що багато математичних методів, які використовують у психології, засновані саме на ймовірнісній теорії. Крім того, існують спеціальні методи та прийоми, що дають змогу визначити випадковість чи не випадковість змін у досліджуваній ознаці. Якщо психологічна властивість (ознака) – випадкова величина, то до неї застосовні основні статистичні критерії, якщо ознака змінюється не випадково, слід виявити і по можливості усунути (чи мінімізувати) чинник, який спричиняє виникненню систематичної помилки. Якщо це неможливо, слід використовувати так звані непараметричні методи статистичної обробки отриманих результатів.

Варіабельність (варіативність) психологічних вимірювань полягає в тому, що психологічні величини (ознаки, змінні) найчастіше набувають значень, дуже відмінних одне від одного, тому, поряд із мірами центральної тенденції (модю, медіаною, середнім значенням), у психології завжди доводиться враховувати і варіабельність (мінливість) вимірюваної ознаки. Доведено, що варіабельність змінних сама по собі є інформативним показником. Розроблено велику кількість статистичних методів, заснованих саме на аналізі варіабельності: дисперсійний, кореляційний, факторний аналіз та ін. Іншими словами, варіабельність психологічних ознак із противника перетворилася на союзника, і на її підґрунті стало можливим вирішення досить складних статистичних завдань. У різних галузях і розділах психології вимірювання має власну специфіку. Так, психофізичні вимірювання передбачають, як правило, використання двох шкал: першої – шкали фізичних одиниць (сили світла, звуку, просторових, часових параметрів сигналу та ін.), другої – суб'єктивної (шкали суджень, оцінок тощо), яка може бути виражена у термінах номінальної, порядкової, інтервальної шкали чи шкали відносин

(див. підрозд. 1.3). У разі застосування неметричного шкалювання дослідник, як правило, оперує лише суб'єктивними шкалами.

Подвійна метрика передбачена й у психофізіологічних дослідженнях. Фізіологічні процеси в організмі людини вимірюють спеціальними приладами і виражають в загальноприйнятих фізичних одиницях – секундах, герцах, мікро- і мілівольтах та ін. Водночас психічні процеси, що супроводжують фізіологічні зміни в організмі, вимірюють у термінах суб'єктивного самозвіту.

Особливе місце вимірювання посідають у психодіагностиці, оскільки включають, з одного боку, систему суб'єктивних звітів або невербальних операцій суб'єкта, з іншого – систему умовних прийомів та методів оцінювання психологічних показників. Наголосимо, що, незважаючи на варіативність і багатофакторність психологічних величин, вимірювання у психології є невід'ємним етапом дослідження, що дає змогу із певним ступенем точності і надійності описати різноманітні психічні явища, процеси і стани.

1.3. Типи і характеристики вимірювальних шкал

Оскільки психологія вивчає психологічні процеси, то, за потребою, вона оперує різними числовими показниками, що виражають частоту, тривалість і ступінь прояву пов'язаності різних характеристик. Передумовою усіх операцій із кількісним вираженням властивостей психологічних процесів і показників є початкове вимірювання якісних ознак або його квантифікація. Проблема початкового виміру лише частково математична. Щоб за певними правилами приписати числа властивостям об'єкта психології, треба усвідомити їх змістовну структуру, знайти відповідність між нею та інструментом вимірювання. Це завдання якісно-кількісного аналізу. Вимірюванню підлягають будь-які властивості психологічних об'єктів: якісні і кількісні. З кількісними усе зрозуміло, ними вже є загальноприйняті зразки вимірювання (рік, маса, вік та інше). Якісні характеристики не мають встановлених еталонів вимірювання, тому їх

доводиться конструювати відповідно до природи об'єкта, який вивчають. Отже, введемо деякі визначення.

Ознаки та змінні – це вимірювані психологічні явища, наприклад, час вирішення завдання, кількість помилок.

Значення ознак визначають за допомогою спеціальних шкал спостереження. Психологічні змінні є *випадковими величинами*, оскільки невідомо заздалегідь, якого саме значення вони набудуть.

Змінна величина (або просто змінна) – кількісно вимірювана властивість або ознака, що набуває різних значень. Змінними можуть бути різні психічні ознаки: час вирішення завдання, кількість допущених помилок, рівень тривожності або нейротизму, коефіцієнт інтелекту та ін. Значення змінних можуть змінюватися безперервно, або дискретно. Так, у більшості психофізіологічних досліджень вимірювані величини в принципі безперервні, і точність їх вимірювання залежить від точності вимірювального пристрою (приладу). Дискретні значення змінних застосовують у більшості психодіагностичних процедур, де вимірюваний параметр найчастіше набуває значень цілих чисел: кількість позитивних і негативних відповідей, кількість правильно виконаних завдань тощо.

Вважають, що психологічні змінні є випадковими величинами, оскільки на них впливають різноманітні чинники (див. підрозд. 1.2) і неможливо передбачити заздалегідь, якого значення вони набудуть.

Дані у статистиці – це основні елементи, що підлягають аналізу. Даними можуть бути кількісні результати, властивості, місце у послідовності, тобто будь-яка інформація, яка може бути класифікована чи розподілена за категоріями для обробки.

Побудова **розподілу низки даних** – це поділ первинних даних, отриманих на вибірці, на класи чи категорії для отримання узагальненої, впорядкованої картини, і подальшого їх аналізу. Існують два типи даних (рисунок 1.1):

Кількісні дані – одержані в ході вимірювань (наприклад, дані про вагу, розміри, температуру, час, результати тестування та ін.), їх можна розподілити за шкалою з рівними інтервалами;

Якісні дані – відображають якість властивості елементів вибірки чи популяції, їх не можна виміряти, єдиною їх кількісною оцінкою є частота прояву.

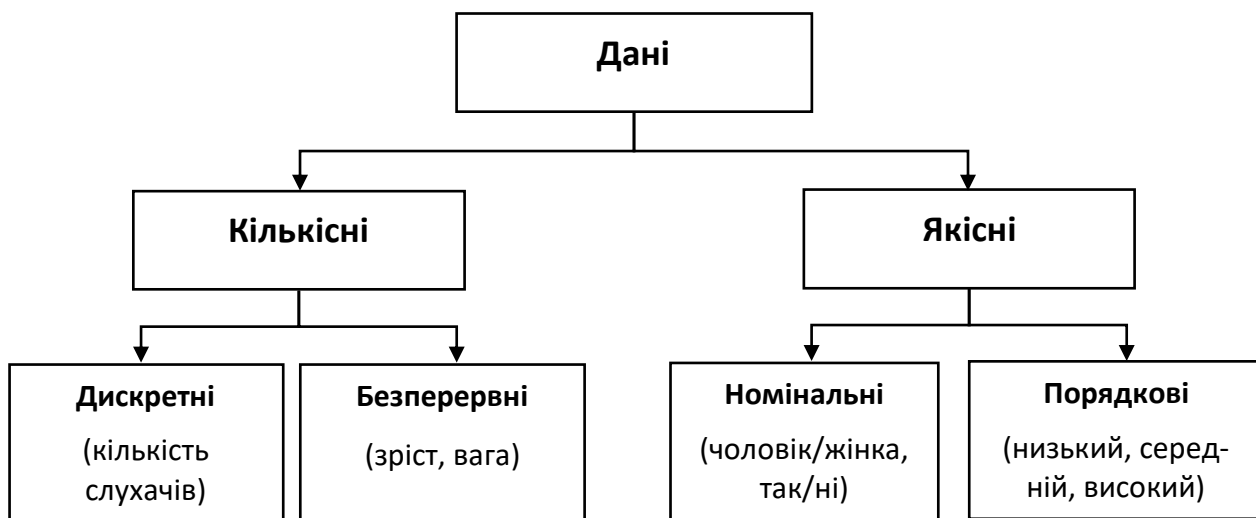


Рис. 1.1. Розподіл типів даних.

З усіх цих типів лише кількісні дані можна аналізувати за допомогою методів, основу яких становлять *параметри* (середнє арифметичне, мода, дисперсія тощо). Однак навіть до кількісних даних такі методи можна застосувати лише якщо кількість достатня, щоб з'явився нормальний розподіл.

Математична обробка – це оперування значеннями ознаки (змінної), отриманими у досліджуваних осіб у процесі психологічного дослідження. Методи математичної обробки дуже різноманітні: побудова розподілу частот вимірюваної ознаки, обчислення мір центральної тенденції, мір мінливості (варіабельності) ознаки, визначення специфіки зв'язку між різними змінними, встановлення форми залежності однієї ознаки від іншої, вплив чинників на величину ознаки та ін.

Оскільки більшість змінних, які вивчають у психології, не є жорстко детермінованими величинами, то більшість математичних методів засновані на постулатах теорії ймовірностей. Це стосується і висновків, які робить дослідник за результатами математичної обробки отриманих даних. Будь-який висновок чи прогноз може бути зроблено лише з певною ймовірністю. Для характеристики цієї ймовірності використовують поняття рівнів значущості (p).

Вимірювання – це приписування числових форм об'єктам чи подіям відповідно до певних правил.

Шкала в широкому розумінні цього слова є впорядкованою сукупністю даних. Інакше кажучи, якщо у психологічному експерименті (спостереженні, опитуванні тощо) ми отримуємо якісь результати (дані) і належним чином впорядковуємо їх, ми конструюємо шкалу. Так, С. Стівенс запропонував класифікацію із 4 типів шкал вимірювання (таблиця 1.1):

- 1) номінативна, номінальна, чи шкала найменувань;
- 2) порядкова, чи рангова, шкала;
- 3) інтервальна, або шкала рівних інтервалів;
- 4) шкала рівних відносин.

Таблиця 1.1

Класифікація вимірювальних шкал

Шкала	Характеристика	Приклади
Номінативна	Полягає у наданні будь-якій властивості або ознаці певного позначення або символу. Вимірювання за цією шкалою передбачає класифікацію або розподіл на класи, що не перетинаються. Самі по собі символи не містять ніякої інформації, операції з ними не мають сенсу	Типи темпераменту; типи акцентуації характеру, варіанти відповідей досліджуваних: дихотомічні (двійкові) так – ні, за – проти, інтроверт – екстраверт, повна сім'я – неповна сім'я.
Порядкова	Класифікує сукупність виміряних ознак за принципом "більше – менше", "вище – нижче", "сильніше – слабше".	Шкільні оцінки від 1 до 12, закодовані рівні від низького до високого; ранжовані ієрархії (наприклад, військові звання)
Інтервальна	Кожне із можливих значень виміряних величин віддалене від найближчого на рівну відстань. Нуль умовний. Для вимірювання за допомогою шкали інтервалів встановлюють спеціальні одиниці вимірювання – стени. Вимірюваній властивості або предмету присвоюють число, що дорівнює кількості одиниць вимірювання, еквівалентне кількості вимірюваної властивості	Семантичний диференціал Ч. Осгуда, IQ Векслера, Т-шкала, 16-факторний опитувальник Кеттела та інші тестові шкали, які спеціально вводять для обґрунтування їх рівноінтервальності
Рівних відносин	Має всі властивості інтервальної шкали і чітко фіксований нуль, який означає повну відсутність властивості.	Зріст, вага, кількість реакцій, показник сили, витривалості

Номінативна шкала – це шкала, що класифікує за назвою, тобто не вимірює кількісно, а дає змогу відрізнити один об'єкт від іншого, або одного суб'єкта від іншого. Номінативна шкала – це спосіб класифікації об'єктів чи суб'єктів, їх розподілу в осередках класифікації.

Найпростішим, окремим випадком номінативної шкали є дихотомічна шкала, що складається з двох осередків, наприклад: "має братів і сестер – єдина дитина у сім'ї"; "чоловік – жінка"; "проголосував "За" – проголосував "Проти"" та ін.

Після класифікації всіх об'єктів, реакцій чи досліджуваних можна від найменувань перейти до чисел, щоб підрахувати кількість спостережень у кожному класі. Номінативна шкала дозволяє підраховувати частоти різних найменувань або значень ознаки, а потім працювати з ними. Одиниця вимірювання, якою ми оперуємо – одне спостереження.

На номінативній шкалі проводять такі операції із числами.

1) Знаходять частоти розподілу за пунктами шкали за допомогою процентування або натуральних одиниць, підраховують чисельність кожної групи та відношення цієї чисельності до загального ряду розподілу;

2) Шукають середню тенденцію за модальною частотою (M_o) – групою з найбільшою чисельністю, що дає змогу розподілити психологічні характеристики у кількісних показниках, для наочності – відобразити у діаграмах;

3) Встановлюють взаємозв'язки між рядами властивостей, розташованими невпорядковано. З цією метою створюють перехресні таблиці. Крім простого відсоткового співвідношення, за таблицях перехресної класифікації підраховують критерій кореляції ознак за Пірсоном.

Порядкова шкала – це шкала, що класифікує за принципом "більше – менше". Якщо у шкалі найменувань байдуже, у якому порядку розташовані класифікаційні осередки, то у порядковій шкалі вони утворюють послідовність від осередку «найменше значення» до осередку «найбільше значення» (чи навпаки).

Це повністю впорядкована шкала найменувань, що встановлює відносини рівності між явищами у кожному класі та відносини послідовності у поняттях «більше – менше» між усіма класами.

Упорядковані порядкові шкали загальноповживані під час опитувань громадської думки. З їх допомогою вимірюють інтенсивність оцінок якихось психологічних властивостей, суджень, подій, ступеня згоди чи незгоди із запропонованими твердженнями. Дуже часто застосовують різновид шкал цього виду – рангові, що передбачають повне впорядкування якихось об'єктів.

На порядковій шкалі проводять такі операції з числами:

Інтервали в цій шкалі не рівні, тому числа позначають лише порядок розташування ознак, а операції з числами – кількісно виражають властивості у кожному пункті:

1) числа піддають монотонним перетворенням, тобто замінюють іншими зі збереженням колишнього порядку. Так, замість ранжування від 1 до 5 можна впорядкувати той самий ряд у числах від 2 до 10, при цьому відношення між рангами залишаться незмінними;

2) застосовують сумарні оцінки за рядом впорядкованих номінальних шкал для вимірювання тієї самої властивості набору різних індикаторів;

3) для роботи з матеріалом, зібраним за впорядкованою шкалою, використовують, крім модальних показників (M_o), пошук середньої тенденції за допомогою медіани (M_e або M), обчислення середнього арифметичного (\bar{X}) та оцінювання розкиду даних за допомогою дисперсії (D) і стандартного відхилення (δ);

4) застосовують кореляцію рангів за Спірменом або Кендалом, щоб вказати на наявність або відсутність функціональних зв'язків у двох рядах ознак, виміряних упорядкованими шкалами.

Інтервальна шкала – це шкала, що класифікує за принципом «більше на певну кількість одиниць – менше на певну кількість одиниць». Кожне з можливих значень ознаки відстає від іншого на рівній відстані.

Шкала інтервалів є повністю впорядкованим рядом з виміряними інтервалами між пунктами, причому відлік починається довільно від обраної величини (немає абсолютного нуля).

В інтервальній метричній шкалі виконують такі операції з числами:

- 1) точку відліку вибирають довільно;
- 2) застосовують методи описової статистики;
- 3) застосовують кореляційний та регресійний аналіз, обчислюють коефіцієнт парної кореляції Пірсона та коефіцієнти множинної кореляції, що може передбачати зміни в одній змінній залежно від змін в іншій або в ряді змінних.

Шкала рівних відносин – це шкала, що класифікує об'єкти або суб'єктів пропорційно до ступеня вираженості вимірюваної властивості. У шкалах відносин класи позначають числами, пропорційними одне одному: 2 так належить до 4, як 4 до 8. Це передбачає наявність абсолютної нульової точки відліку. Вважають, що у психології прикладами шкал рівних відносин є шкали порогів абсолютної чутливості. Можливості людської психіки настільки великі, що важко уявити абсолютний нуль будь-якої вимірюваної психологічної змінної. Абсолютна дурість та абсолютна чесність – поняття швидше життєвої психології.

Можливі перетворення з однієї шкали в іншу. Результати, отримані за шкалою інтервалів, можуть бути перетворені на ранги або переведені в номінативну шкалу. Розглянемо, наприклад, первинні результати шести досліджуваних за шкалою екстраверсії-інтроверсії тесту Айзенка (таблиця 1.2).

Таблиця 1.2

Досліджувані особи	Інтервальна шкала	Порядкова (рангова) шкала	Номінативна шкала
Антоненко	20	5	Е
Бевз	15	4	Е
Василюк	22	6	Е
Іваненко	9	3	І
Костюк	3	1	І
Яцко	4	2	І

Перший стовпець – імена досліджуваних осіб, другий стовпець – бал за виразність якості (реалізована шкала інтервалів), третій стовпець – відповідно до вихідного балу досліджуваним особам приписані ранги (перший ранг отримує досліджуваний, що має найменший бал, другий ранг – досліджувана особа, що має наступний за величиною бал, і т. д.), четвертий стовпець – відповідно до вихідних балів досліджувані особи розподілені на два класи: інтроверти (І) – бали від 0 до 12, екстраверти (Е) – від 13 до 24. Зазначимо, що кожен перехід від однієї шкали до іншої означає втрату частини інформації про досліджуваних осіб. Так, у ході ранжування (застосування порядкової шкали) зникне, наприклад, та інформація, що випробувані особи Костюк і Яцко, розташовані один за одним, мають відмінність у первинних оцінках в один бал, а Іваненко та Бевз – у шість балів. У разі розподілу досліджуваних осіб за класами (застосування номінативної шкали) в один клас потраплять ті, які сильно розрізняються за первинними оцінками.

Ми розглянули різні прийоми переведення якісних психологічних ознак у кількісне вираження. Зазначимо, що під час опису психологічних явищ необхідно завжди усвідомлювати, яку саме шкалу використано, оскільки кожен спосіб обробки експериментальних даних розрахований на певний тип шкали. Застосування математичних методів до неадекватних даних призводить до отримання дивних, а часто й хибних результатів. Квантифікація складних і неоднозначних психологічних характеристик накладає чимало обмежень на математичні операції з їх вимірами. Математик працює з абстрактними числами, психолог має пам'ятати, що насправді стоїть за величинами, якими він оперує.

Існують такі **обмеження** у застосування різних шкал:

1) пропорційними кількісні показники мають бути фіксовані різними шкалами в межах одного дослідження. Сильніша шкала відрізняється від слабшої тим, що допускає ширший діапазон математичних операцій з числами. Все, що допустимо для слабкої шкали, допустимо і для сильної, але не навпаки. Змішування в аналізі вимірювальних еталонів різного типу

призводить до того, що дослідники не використовують можливості сильних шкал,

2) форму розподілу величин, фіксованих описаними вище шкалами, передбачають нормальною. У разі нормального розподілу оцінки заходи розсіювання збігаються: $\bar{X} = Me = Mo$, у скошеного хвості розподілу впливають на середню (\bar{X}).

Таким чином, необхідно уважно вивчати форму розподілу з точки зору його відхилення від нормального.

Питання для самоконтролю

1. Дайте визначення поняттю вимірювання?
2. Розкрийте сутність поняття «статистична сукупність»?
3. Які завдання вирішує математична статистика?
4. Назвіть основні типи шкал?
5. Розкрийте властивості психолого-педагогічних вимірювань?

Завдання для самостійної роботи

1. Які із зазначених змінних є дискретними?
 - а) стать людини;
 - б) кількість домашніх тварин у родині;
 - в) дальність прямого пострілу гармати;
 - г) кількість військовослужбовців у підрозділі;
 - д) імена Ваших друзів.
2. Які Ви знаєте основні два типи змінних у статистиці?
 - а) якісні і номінативні;
 - б) безперервні та кількісні;
 - в) рангові і номінативні;
 - г) кількісні і якісні.

3. Якщо зріст 10 учасників дослідження наведений у ранговій шкалі (за спаданням: 1 – найвищий, 2 – нижче і т. д.), тоді правильним твердженням буде:

- а) досліджуваний із рангом 4 на 2 см нижчий, ніж досліджуваний із рангом 6;
- б) досліджуваний із рангом 4 вищий, ніж досліджуваний із рангом 6, але нижчий, ніж досліджуваний із рангом 2;
- в) у досліджуваних із рангами 1 і 3 така сама різниця у рості, як і у досліджуваних із рангами 3 і 5;
- г) жодне твердження не є правильним.

4. Знайдіть співвідношення між змінними та їх типами.

Кількісна (дискретна) змінна	Група крові
Кількісна (безперервна) змінна	Зріст (у мм)
Рангова змінна	Успішність тих, хто навчається
Номінативна змінна	Кількість публікацій

РОЗДІЛ 2

ПОНЯТТЯ ВИБІРКИ В НАУКОВИХ ДОСЛІДЖЕННЯХ

2.1. Генеральна сукупність та вибірка дослідження.

2.2. Вимоги до вибірки.

2.3. Методи формування вибірки.

2.1. Генеральна сукупність та вибірка дослідження

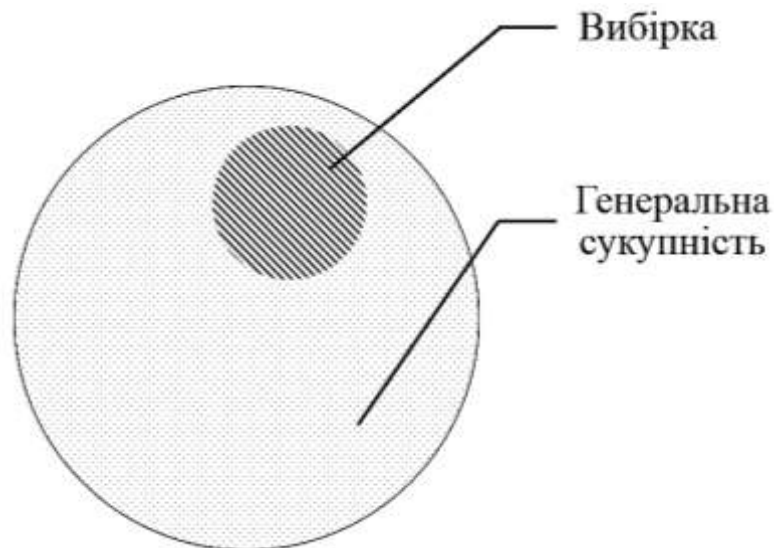
У математичній статистиці виокремлюють два фундаментальні поняття: генеральну сукупність і вибірку.

Генеральна сукупність (N) – множина усіх об'єктів, властивості яких цікавлять дослідника. Теоретично вважають, що обсяг генеральної сукупності необмежений, практично ж обсяг генеральної сукупності завжди обмежений і може залежати від предмета спостереження і завдання, яке необхідно вирішити.

Приклад: *військовий психолог має за мету дослідити емоційний стан конкретного військовослужбовця протягом бойового чергування. У цьому разі генеральною сукупністю є сам військовослужбовець. Якщо психолог досліджує емоційний стан особового складу чергових бойових обслуг певного роду військ, то сукупність буде набагато більшою, і ще більшою, якщо досліджує емоційний стан чергових бойових обслуг у виді збройних сил.*

На практиці дослідити усю генеральну сукупність або занадто трудомістко, або майже неможливо, тому психолог, як правило, досліджує певну вибірку людей, яку добирає з більшої за чисельністю групи.

Вибірка (n) – це обрана для дослідження частина генеральної сукупності. Вибірку можна тлумачити як деякий емпіричний аналог генеральної сукупності. Задачі математичної статистики фактично зводяться до обґрунтування висновку про об'єктивні властивості генеральної сукупності за результатами дослідження вибірки.



Кількість елементів вибірки n називають **обсягом вибірки**, який може становити не менше двох досліджуваних. У статистиці розрізняють малу ($n < 30$), середню $30 < n < 100$ і велику вибірку ($n > 100$). Цей поділ достатньо умовний.

Зауважимо, що обсяг вибірки залежить насамперед від завдань дослідження. Військовий психолог може досліджувати поодинокі випадки, якщо ті становлять особливий інтерес для науки, наприклад, психофізіологічні особливості професійної адаптації бойових плавців. У таких випадках, зрозуміло, генеральна сукупність і, відповідно, вибірка будуть малими за обсягом. Коли психолог ставить за мету вивчити характеристики, властиві для багатьох представників генеральної сукупності, більший обсяг вибірки дасть змогу отримати надійніші результати. Крім того, обсяг вибірки залежить від статистичних методів, які передбачають використовувати. Одні методи потребують великої кількості досліджуваних осіб у вибірці, інші можуть бути застосовані при відносно невеликій їх кількості.

Малі вибірки зазвичай використовують для статистичного контролю відомих властивостей вже досліджених сукупностей, великі вибірки - для визначення невідомих властивостей і параметрів сукупності.

Загальні рекомендації досліднику під час визначення обсягу вибірки:

– якщо потрібно порівняти дві вибірки, їх загальна чисельність не має бути меншою за 50 досліджуваних осіб;

- для вивчення взаємозв'язку між будь-якими властивостями обсяг вибірки має становити більше 30–35 досліджуваних;
- у ході розроблення психодіагностичних методик обсяг вибірки має становити від 200 до 2000 досліджуваних осіб;
- чисельність порівнюваних вибірок має бути приблизно однаковою;
- чим більша мінливість досліджуваної властивості, тим більшим має бути обсяг вибірки. Мінливість можна зменшити, збільшуючи однорідність вибірки, наприклад, за статтю, віком тощо, але при цьому можливість генералізації висновків зменшиться.

Залежно від схеми експерименту розрізняють незалежні та залежні вибірки. Вибірki є **незалежними**, якщо процедура експерименту й отримані результати вимірювання деякої властивості у досліджуваних осіб однієї вибірки не впливають на особливості перебігу цього експерименту і результати вимірювання цієї ж властивості у досліджуваних осіб іншої вибірки.

Приклад: *курсанти першого курсу різних спеціальностей.*

Вибірki є **залежними**, якщо процедура експерименту й отримані результати вимірювання деякої властивості, проведені на одній вибірці, впливають на іншу. Зауважимо, що група досліджуваних осіб, у якій двічі проводилося психологічне обстеження (нехай навіть різних психологічних властивостей, особливостей, якостей), за визначенням є залежною вибіркою.

Приклад: *повторне дослідження властивостей на тій самій вибірці після дії (ситуація “до – після”).*

Отже, залежні вибірки завжди мають однаковий обсяг, а незалежні – різний.

2.2. Вимоги до вибірки

Вибірka, що адекватно відображатиме властивості генеральної сукупності, має відповідати таким вимогам:

однорідність – включає досліджуваних осіб, характеристики яких відповідають меті експерименту (якщо вивчають курсантів першого курсу,

то у вибірці не можуть бути курсанти інших курсів). Водночас має бути дотримано однорідності вибірки і за іншими критеріями, наприклад статтю. Підставами для формування однорідної вибірки можуть бути різні характеристики залежно від цілей дослідження: вік, стать, освіта, рівень інтелекту, національність тощо;

репрезентативність – представленість всіх основних ознак генеральної сукупності приблизно у тій самій пропорції і з тією ж частотою, з якою ця ознака наявна в генеральній сукупності. Іншими словами, репрезентативна вибірка є меншою за розміром, але точною моделлю генеральної сукупності, яку вона має відображати. Залежно від репрезентативності вибірки висновки, зроблені на основі її дослідження, можна з більшою мірою впевненості застосовувати до всієї генеральної сукупності.

Поняття “репрезентативний” означає типовий, характерний, адекватний, подібний до чого-небудь. Репрезентативності досягають завдяки побудові вибірки, яка найкраще відображає генеральну сукупність і дає змогу обґрунтовано переносити висновки, отримані під час її дослідження, на генеральну сукупність.

Приклади порушення репрезентативності вибірки

“Передвиборче опитування”

Найвідомішим прикладом порушення репрезентативності вибірки є історія провалу американського журналу “Літературний дайджест”. У 1936 році журнал вкотре провів поштове опитування громадської думки про ймовірні результати майбутніх президентських виборів у США. До 1936 року опитування завжди правильно передбачало переможця. Опитування у 1936 році показало, що переможцем із великим відривом стане кандидат від республіканців, але переможцем виявився представник демократів. Таким чином, гігантська вибірка (близько 2,4 млн осіб) не забезпечила достовірних результатів.

Визначено дві основні причини провалу: зміщення під час формування вибірки та зміщення внаслідок відмови респондентів від участі в опитуванні.

Перша проблема полягала в тому, що журнал включив своїх передплатників до списку для розсилки анкет і, маючи бажання розширити вибірку, використав два інші доступні тоді списки громадян: зареєстрованих автовласників та користувачів телефонів. Але за часів Великої депресії представники цих груп відрізнялися від решти населення вищим прибутком, як і передплатники журналу. Таким чином, отримана база для розсилки була некоректним відображенням структури населення США.

Друга проблема з опитуванням полягала в тому, що з 10 млн осіб, чийі імена були в початковому списку розсилки, лише 2,4 млн взяли участь в опитуванні. Ймовірно, високий відсоток відмов був пов'язаний з тим, що опитування проводилося поштою. Вже на той час американці ставилися до поштових розсилок як до спаму. Таким чином, розмір вибірки становив приблизно одну чверть запланованої. Коли частка тих, хто відповів, низька, вважають, що дослідження містить необ'єктивні відповіді.

У цій історії дві моралі: велика, але неправильно сформована вибірка набагато гірша за маленьку, але правильно сформовану. Під час проведення опитування не слід нехтувати зсувом у відборі та зміщенням, що виникло через відмови.

Систематична помилка того, хто вижив

Під час Другої світової війни американські військові зіштовхнулися з тим, що не всі американські бомбардувальники після виконання завдання поверталися на базу. На літаках, що повернулися, залишалось безліч пробоїн від пострілів противника, але їх розподіл був нерівномірним: більше – на фюзеляжі та інших частинах, менше – в паливній системі, набагато менше – у двигуні (рисунок 2.1).

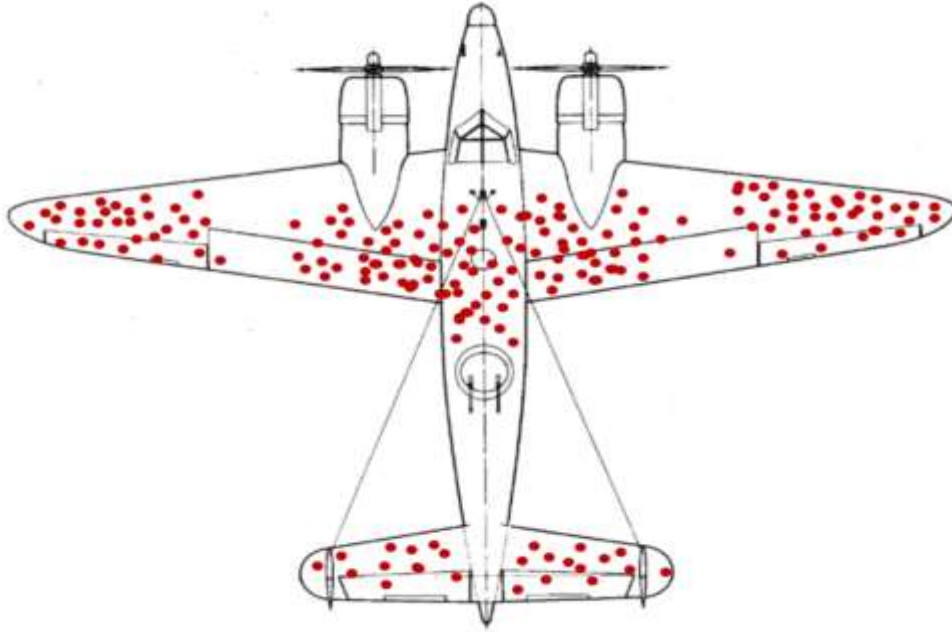


Рис. 2.1. Розподіл пробоїн від пострілів противника на американських бомбардувальниках після бойових вильотів.

Командування вирішило, що на найбільш ушкоджених місцях доцільно встановити більше броні. Залучений до вирішення завдання математик заперечив: дані вказують на те, що літак, який має пробоїни в цих місцях, може повернутися на базу, а літак, якому влучили в бензобак чи двигун, виходить із ладу і не повертається, тому зміцнювати слід ті місця, які у літаків, що повернулися, пошкоджені найменше.

Це завдання є прикладом порушення репрезентативності вибірки, коли до неї включені не ті досліджувані (літаки, що повернулися), а ті, що не повернулися – проігноровані.

Існують різні ступені репрезентативності: жоден дослідник не в змозі сформулювати вибірку, яка б абсолютно точно відтворювала структуру генеральної сукупності. Це призвело б до того, що вибірка дорівнювала б за обсягом генеральній сукупності. З огляду на це дослідник мусить обирати: великі обсяги вибірки і дуже високу репрезентативність або малі за обсягом вибірки, але нижчу репрезентативність.

Таким чином, найважливішою властивістю статистичної обробки даних є проблема репрезентативності: вибірка має адекватно відображати властивості генеральної сукупності.

2.3. Методи формування вибірки

Виникає цілком логічне питання: як сформувати репрезентативну вибірку? З точки зору статистики репрезентативність вибірки означає, що наведений у вибірці розподіл досліджуваних ознак відповідає (з певною часткою похибки) їх розподілу в генеральній сукупності. Розглянемо основні методи, що забезпечують репрезентативність вибірки.

Метод формування простої випадкової вибірки означає, що у вибірку включають досліджуваних осіб, відібраних із генеральної сукупності так, щоб кожна досліджувана особа сукупності мала рівні можливості (рівну ймовірність) потрапити у вибірку. Отримати просту випадкову вибірку можна за допомогою спеціальних таблиць випадкових чисел.

Метод формування стратифікованої випадкової вибірки передбачає поділ елементів генеральної сукупності на страти (підгрупи) за певними характеристиками. Вибірку формують так, щоб у ній були досліджувані особи з кожної страти (наприклад, за статтю, віком, освітою тощо).

Метод попарного відбору передбачає, що в експериментальну і контрольну групи включають індивідів, еквівалентних за значущими для експерименту побічними параметрами. Ідеальний варіантом було б використання близнюкових пар.

Метод наближеного моделювання полягає у доборі однорідних підгруп, у яких досліджувані мають однакові характеристики, крім тих, що становлять інтерес для психолога.

Метод репрезентативного моделювання передбачає виокремлення додаткової значущої змінної. Усі досліджувані особи беруть участь у тестуванні і надалі підлягають ранжуванню за ступенем виразності певної змінної. Групи формують так, щоб досліджувані з однаковими або близькими значеннями цієї змінної увійшли до складу різних груп.

Метод формування реальних груп реалізують шляхом добору бажаючих для участі в експерименті, однак це, безумовно, порушує репрезентативність вибірки.

З точки зору зовнішньої і внутрішньої валідності оптимальною є стратифікована випадкова вибірка. Досліджувані, що становлять певну

генеральну сукупність, можуть різнитися за статтю, віком, рівнем освіти, професійним досвідом тощо. Кожному з них слід надати рівну можливість стати учасником експерименту.

Сформувати вибірку можна за допомогою MS Excel: пакет “Аналіз даних” містить відповідні розділи “Генерація випадкових чисел” та “Вибірка”.

Питання для самоконтролю

1. Дайте визначення генеральної сукупності.
2. Дайте визначення вибірки.
3. Що таке обсяг вибірки і від чого він залежить?
4. Яка вибірка є малою, великою? Який рекомендований обсяг вибірки?
5. Дайте визначення залежної та незалежної вибірок.
6. Яку вибірку називають однорідною?
7. Яку вибірку називають репрезентативною?
8. Назвіть основні методи, що забезпечують репрезентативність вибірки.

Завдання для самостійної роботи

1. Позначте генеральну сукупність і вибірку:
 - курсанти технічних спеціальностей, курсанти-інженери;
 - механіки-водії танкових екіпажів, військовослужбовці сухопутних військ;
 - льотний склад авіаційних частин, льотчики авіаційної бригади.
2. Зазначте, які вибірки є залежними, а які незалежними:
 - група військовослужбовців до та після участі у тренінгу;
 - військовослужбовці 1-го та 2-го взводів однієї роти;
 - офіцери-психологи та офіцери-зв'язківці.
3. Самостійно сформулюйте проблему дослідження, визначте генеральну сукупність та оптимальний обсяг вибірки. Вкажіть, як зробити вибірку репрезентативною та однорідною.

РОЗДІЛ 3

ПЕРВИННА ОПИСОВА СТАТИСТИКА

3.1. Міри центральної тенденції.

3.2. Міри мінливості.

3.3. Міри положення.

3.4. Міри зсуву.

3.1. Міри центральної тенденції

Всі методи кількісної обробки поділяють на первинні і вторинні. Первинна статистична обробка має за мету впорядкування інформації про об'єкт і предмет вивчення. На цьому етапі «сирі» відомості групують за певними критеріями і вписують у зведені таблиці. Первинно оброблені дані дають змогу дослідникові у загальних рисах побачити всю сукупність даних: її однорідність – неоднорідність, компактність – розкиданість, чіткість – розмитість та ін. Цю інформацію легше простежити у візуальній формі наведення даних і зрозуміти їх розподіл.

Основними методами первинної статистичної обробки є: обчислення мір центральної тенденції, мір зсуву, мір розкиду (мінливості) даних та квантиля розподілу.

Первинний статистичний аналіз всієї сукупності отриманих у дослідженні даних дає змогу охарактеризувати її у стислому вигляді і з'ясувати: яке значення найбільш характерне для вибірки і чи великий розкид даних щодо характерного значення, тобто яка їх варіативність. Для цього обчислюються міри центральної тенденції і міри мінливості (або розкиду) та міри зсуву.

Міри центральної тенденції (МЦТ) – найважливіші характеристики сукупності статистичних даних. Використовуючи їх, можна за допомогою одного числа навести найбільш загальний опис досліджуваного набору даних. Проте часом МЦТ (точніше їх некоректне чи недоцільне використання) можуть істотно спотворити уявлення про сукупність даних.

Як правило, використовують три види мір центральної тенденції: середнє арифметичне, моду та медіану. Ці показники дають відповіді на питання, наприклад, про «середній рівень інтелекту слухачів університету», чи «типове значення показника успішності певної навчальної групи осіб» тощо. Кожна конкретна міра центральної тенденції має свої особливості, що роблять її цінною для характеристики об'єкта дослідження в певних умовах.

Міри центральної тенденції потрібні з наступних міркувань:

1. Щоб отримати загальну картину розподілу. Ми не можемо запам'ятати кожен факт, що стосується сфери дослідження.
2. Щоб отримати чітку картину щодо досліджуваної сфери для розуміння та отримання потрібних висновків.
3. Щоб отримати чіткий опис групи в цілому та мати змогу порівнювати дві або більше груп у термінах типової «поведінки».

Мода (*Mo*) – це числове значення, яке найбільш часто зустрічається у вибірці, тобто число з найбільшою частотою. Якщо всі значення в групі зустрічаються з однаковою частотою, то вважається, що моди немає (унімодальний розподіл). Якщо два сусідніх значення мають однакову частоту і більше частоти будь-якого іншого значення, мода є середнє цих двох значень. Якщо те ж саме відноситься до двох несуміжних значень, то існує дві моди, а група оцінок є бімодальною. На моду ніколи не впливають екстремальні значення в розподілі, а впливають – екстремальні частоти значень, наскільки часто те чи інше значення змінної зустрічається в розподілі.

Мода використовується:

1. Коли нам потрібна швидка і приблизна міра центральної тенденції.
2. Коли потрібна міра центральної тенденції, що має бути типовим значенням.

Переваги моди:

1. Мода показує найбільш поширене значення в розподілі.
2. На моду не впливають екстремальні значення – так як на середнє.
3. Моду можна визначити для відкритих інтервалів/категорій.

4. Допомогає аналізувати якісні дані.

5. Моду можна виявити просто побудувавши графік розподілу чи стовпчасту діаграму.

Обмеження:

1. Не включає до визначення/розрахунку всі спостереження розподілу, а лише концентрацію частот.

2. Подальші алгебраїчні перетворення неможливі – на відміну від середнього.

3. Буває важко визначити моду у випадку багатомодального чи бімодального розподілу.

Розподіл може мати більше двох популярних значень, але якщо має більше ніж трьох мод, опис такого розподілу в термінах найбільш частих значень може втрачати будь-який сенс.

Середнє арифметичне значення (вибіркове середнє або середнє) \bar{X} – це результат ділення суми всіх значень (X) на їх кількість (n).

$$\bar{X} \equiv M = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Середнє, або ж просте середнє, або ж арифметичне середнє (arithmetic mean) – просто середнє значення ряду даних. Найвідомішою мірою центральної тенденції – і найбільш вживаною в повсякденному побуті.

Для його обчислення досить скласти разом всі значення в розподілі, і поділити на кількість спостережень. В Excel чи Google Spreadsheets для цього є функція MEAN. Є різні математичні способи підрахунку середнього, але в усіх сучасних електронних таблицях та спеціальних програмних пакетах для роботи з даними і статистикою є ця функція.

Є певні загальні правила для використання середнього, зокрема:

1. Середнє – це «центр тяжіння» розподілу, і кожне значення дає внесок у визначення середнього значення, коли поширення значень є симетричними довкола центральної точки.

2. Середнє значення більш стабільне, ніж медіана чи мода. Тому, коли потрібно знайти найбільш стабільну міру центральної тенденції,

використовують середнє.

Переваги середнього:

1. Середнє визначене дуже жорстко, тому не виникає питань чи нерозуміння щодо його значення та суті.

2. Це найбільш поширена міра центральної тенденції, оскільки її легко зрозуміти.

3. Середнє легко підрахувати.

4. Враховує всі значення розподілу.

Обмеження чи недоліки середнього:

1. На значення середнього впливають екстремальні значення (відомий іронічний жарт про «середню температуру по лікарні»).

2. Часом середнім є значення, що не присутнє в розподілі.

3. Часом результатом можуть бути абсурдні значення. Наприклад, маємо 41, 44, та 42 військовослужбовців у кожному підрозділі. Виходить, що середня кількість військовослужбовців у трьох підрозділах – 42,3. А так не буває.

Медіана (*Me*) визначається як середня точка у ряді розподілу (впорядкований набір значень змінної для різних спостережень – наприклад від найменшого до найбільшого значення) – до цієї точки розташовано половина всіх значень, і після цієї точки теж половина значень. Тобто, медіана, це значення, що ділить впорядкований ряд навпіл (може приймати значення з точністю до 0,5 наприклад 29,5). Якщо кількість значень непарна, то береться одне зі значень – те, що стоїть у розподілі рівно по центру.

Коли значень парна кількість, то беруть два центральні значення, і знаходять їхнє середнє.

Для чого використовують медіану?

1. Коли потрібно знайти точну середню точку, точку на «півдорозі» від найменшого значення до найбільшого.

2. Коли екстремальні значення впливають на середнє – медіана є найкращою мірою центральної тенденції.

3. Медіану використовують коли потрібно, щоб певні значення

впливали на центральну тенденцію, але все, що про них відомо – що вони «нижче» або «вище» медіани.

Переваги медіани:

1. Легко вирахувати та зрозуміти.
2. Для підрахунку медіани не потрібні всі значення в розподілі.
3. Екстремальні значення розподілу не впливають на медіану.
4. Її можна визначити і для «відкритих» категорій/класів інтервалів.

Обмеження медіани:

1. Вона не так жорстко визначена як середнє, оскільки її значення не так вираховується, як знаходиться (серед значень в розподілі).

2. Не враховує всі спостереження (значення для всіх спостережень).

3. З медіаною потім не можна робити алгебраїчні перетворення так, як із середнім.

4. Потребує впорядкування значень або класів інтервалів у висхідному чи спадному порядку.

5. Часом медіаною може бути значення, яке не присутнє у самому розподілі.

Особливості мір центральної тенденції:

– мода вибірки обчислюється просто, її можна визначити «на око».

Для дуже великих груп даних мода є досить стабільною мірою центру розподілу;

– медіана займає проміжне положення між модою і середнім з погляду її підрахунку. Ця міра особливо легко визначається у разі ранжованих даних;

– середнє арифметичне передбачає використання усіх значень вибірки, причому всі вони впливають на значення цієї міри.

Зазвичай вибіркоче середнє застосовують, щоб якнайточніше визначити найбільшої центральну тенденцію. Медіану обчислюють, якщо у серії є «нетипові» дані (так звані «викиди»), які істотно впливають на середнє. Моду використовують, коли не потрібна висока точність, але важлива швидкість визначення міри центральної тенденції.

Обчислюють усі три показники також, щоб оцінити розподіл даних. На

нормальний розподіл даних вказує те, що середнє арифметичне значення, медіана і мода однакові або дуже близькі.

3.2. Міри мінливості

Міри мінливості даних – це статистичні показники, що дають змогу оцінити ступінь однорідності отриманої множини даних, її компактність. Найбільш поширеними у психологічних дослідженнях є такі показники: розмах, дисперсія, стандартне відхилення.

Міри мінливості – такі як дисперсія і стандартне відхилення дають нам розуміння, наскільки добре, наприклад, середнє представляє весь набір даних. Якщо розкид значень у розподілі великий, то середнє не є таким репрезентативним, ніж якщо розкид даних малий. Тобто, великий розкид даних означає, що, ймовірно, є великі відмінності між значеннями.

Розмах (R) – це інтервал між максимальним і мінімальним значеннями ознаки. Його визначають легко і швидко, але розмах чутливий до випадкових даних (викидів), особливо за малої кількості досліджуваних осіб. Це найпростіша міра відхилення.

Розмах визначають лише крайніми значеннями ознаки:

$$R = X_{max} - X_{min}$$

Дисперсія (D) є мірою відхилення значень випадкової величини від центру розподілу (середнього). Більші значення дисперсії свідчать про більші відхилення значень випадкової величини від центру розподілу. Дисперсія підраховується як середнє квадратів різниці значень розподілу і середнього розподілу. Тобто, спочатку виявляємо середнє розподілу, потім від кожного значення віднімаємо значення середнього, підносимо його у квадратний ступінь, для всіх цих квадратів вираховуємо середнє.

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

Це й буде дисперсія. Для того, щоб отримати стандартне відхилення,

що вимірюється в тих же одиницях, що й значення змінної, яку ми спостерігаємо, береться квадратний корінь з дисперсії. Дисперсія обчислюється тільки для інтервальних та абсолютних шкал.

Нехай необхідно дослідити статистичну сукупність даних X_1, X_2, \dots, X_n . Припустимо, що середнє арифметичне сукупності становить \bar{X} . Для обчислення дисперсії потрібно виконати такі дії:

- визначити всі індивідуальні відхилення $X_i - \bar{X}$ від середнього значення;
- обчислити квадрати індивідуальних відхилень $(X_i - \bar{X})^2$;
- розрахувати середнє значення отриманих квадратів.

Іншими словами, дисперсія – це середнє арифметичне квадратів індивідуальних відхилень від середнього значення сукупності.

Зауважимо, що дисперсія ігнорує знаки відхилень, адже їх було піднесено до квадрата. За цією властивістю дисперсія подібна до середнього абсолютного відхилення.

Крім того, завдяки піднесенню до квадрата дисперсія підсилює вплив великих відхилень (>1) та певною мірою нехтує малими відхиленнями (<1), тому є надійнішою характеристикою, ніж середнє абсолютне відхилення.

Проте використання дисперсії має й недоліки. На відміну від середнього абсолютного відхилення, дисперсія характеризує розсіювання даних у термінах квадратних, а не лінійних одиниць. Наприклад, якщо інтерес становлять відхилення в індивідуальних прибутках від середньогрупових, то за допомогою дисперсії можна отримати їх характеристику у квадратних гривнях. Зрозуміло, природніше описувати відхилення прибутків за допомогою показника, вимірюваного у гривнях.

Щоб врахувати цей недолік, у статистиці замість дисперсії використовують поняття стандартного відхилення (δ).

Стандартне (середньоквадратичне) відхилення – це квадратний корінь із дисперсії:

$$\delta = \sqrt{D} = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}}$$

Стандартне відхилення – зручніший показник, аніж дисперсія. Для багатьох розподілів можна приблизно визначити відсоток даних в одному, двох, трьох і більше стандартних відхиленнях від середнього.

Додатково зауважимо також, що для «не дуже» асиметричних розподілів буде правильною наближена тотожність, яка пов'язує середньоарифметичне та стандартне відхилення: $\bar{X} \approx 0,8 \delta$.

3.3. Міри положення

У психології також широко застосовують **міри положення**, які називають квантилями розподілу. **Квантиль** – це точка на числовій осі вимірної ознаки, яка ділить всю сукупність упорядкованих даних на дві групи з відомим співвідношенням їх чисельності. З одним із квантилів ми вже ознайомилися – це медіана – значення, яке ділить усю сукупність вимірювань на дві групи з рівною кількістю одиниць. Крім медіани, часто використовують процентилі і квантилі.

Процентилі – це 99 точок – значень ознаки ($P_1 \dots P_{99}$), які ділять упорядковану (за зростанням) множину даних на 100 частин, рівних за чисельністю. Визначення конкретного значення процентиля аналогічне визначенню медіани. Наприклад, у ході визначення 10-го процентиля, P_{10} , спочатку всі значення ознаки упорядковують за зростанням, потім відокремлюють 10% досліджуваних, із найменшою вираженістю ознаки. P_{10} відповідатиме тому значенню ознаки, яке відокремлює 10% досліджуваних осіб від решти 90%.

Квантилі – це 3 точки – значення ознаки (P_{25}, P_{50}, P_{75}), які ділять упорядковану (за зростанням) множину даних на 4 рівні за чисельністю частини. Перший квантиль відповідає 25-му процентилю, другий – 50-му процентилю, або медіані, третій квантиль відповідає 75-му процентилю.

Процентилі і квантилі використовуються для визначення частоти виникнення певних значень (або інтервалів) вимірної ознаки чи виокремлення підгруп та окремих досліджуваних осіб, найбільш типових або нетипових для наявної множини спостережень.

Ступінь свободи. Кількість ступенів свободи – це кількість одиниць, що вільно варіюють у складі вибірки. Якщо вибірка складається з n елементів та характеризується середнім \bar{X} , то будь-який елемент цієї безлічі може бути наведено як різницю між цим елементом і сумою решти елементів, крім нього самого.

Приклад: розглянемо ряд чисел 4, 7, 15, 19, 25. Середнє арифметичне $\bar{X} = 14$. У ряду 5 чисел, тобто $n = 5$. Виразимо перший елемент ряду через інші елементи та середнє арифметичне:

$$4 = (5 \cdot 14) - 25 - 19 - 7$$

Далі виразимо другий елемент ряду через інші елементи та середнє арифметичне:

$$7 = (5 \cdot 14) - 25 - 19 - 25 - 15 \text{ і т. д.}$$

Таким чином, один елемент вибірки не має свободи варіації та завжди може бути виражений через інші елементи і середнє арифметичне. У розглянутому прикладі кількість ступенів свободи k визначатиметься як;

$$k = n - 1$$

де n – загальне число елементів вибірки.

За наявності кількох обмежень свободи варіації, кількість ступенів свободи позначають як ν («ню»), що дорівнюватиме $\nu = n - k$, де k – кількість обмежень свободи варіації.

У загальному випадку для таблиці експериментальних даних кількість ступенів свободи визначають за такою формулою:

$$\nu = (m - 1) (n - 1),$$

де m – кількість стовпців, n – кількість рядків.

Знаходження кількості ступенів свободи для кожного методу має свої специфічні особливості.

3.4. Міри зсуву

Стандартом розподілів вважають нормальний розподіл із нульовою асиметрією та ексцесом, проте на практиці цього не завжди досягають. Емпіричний розподіл завжди буде дещо відрізнятися від ідеального (теоретичного). Для врахування цих відхилень застосовують міри зсуву: асиметрію, ексцес та коефіцієнт варіації.

Асиметрія (A) характеризує ступінь несиметричності розподілу відносно його середнього (рисунок 3.1). Позитивна асиметрія (рисунок 3.1б) вказує на відхилення вершини розподілу в бік від'ємних значень, негативна (рис. 3.1в) – у бік додатних.

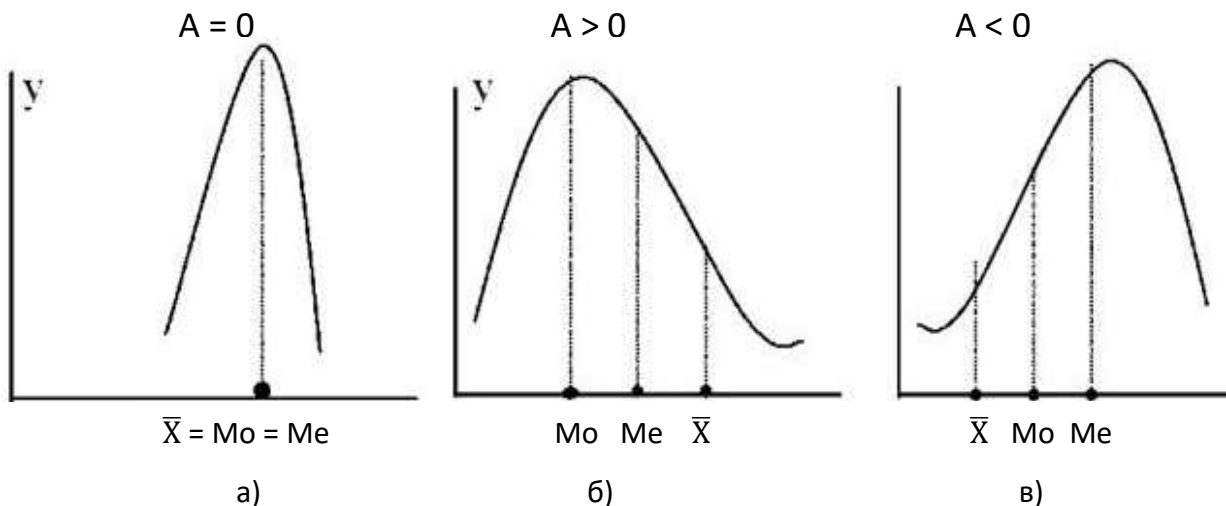


Рис. 3.1. Форми розподілу для різних значень асиметрії.

Коефіцієнт асиметрії (A) розраховують за такою формулою:

$$A = \sqrt{\frac{\sum(x_i - \bar{x})^3}{n * \delta^3}}$$

Вважають, що криві з абсолютною величиною показника асиметрії $A > \pm 0,5$ характеризуються значним зміщенням. Якщо $A \pm 0,25$ - асиметрія незначна.

Графічно (див. рис. 3.1.) асиметрію можна описати напрямом довшої гілки кривої ("дзвін"). Криві, зображені на рисунку 3.1, відображають симетрію і два найпоширеніші види асиметрії розподілу. У разі симетричного розподілу (рисунок 3.1а) середнє арифметичне, мода і медіана рівні між собою. Для асиметричних кривих ці статистичні величини неоднакові, причому середнє арифметичне і медіана зміщені від центра в бік довшої гілки кривої. Оскільки середнє арифметичне (\bar{X}) "чутливе" до "точного" положення більш віддалених від моди (Mo) точок кривої, а медіана (Me) "нечутлива", то середнє (\bar{X}) зрушене більше, ніж медіана (Me). У цьому випадку медіана розміщена між модою і середнім арифметичним.

Як бачимо, напрямок асиметрії геометрично встановити дуже просто, кількісну ж форму ступеня асиметрії обчислюють алгебраїчно.

Недоліком коефіцієнта асиметрії як міри асиметрії є те, що цей показник не має ані верхньої, ані нижньої меж. Особливо великого розміру коефіцієнта асиметрії практично не існує, тобто $+\infty > A > -\infty$.

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

Для встановлення міри відхилення від нормального розподілу обчислюють показник ексцесу (E), що характеризує відхилення від нормального розподілу варіанта із виступанням (рисунок 3.2б) або падінням (рисунок 3.2в) вершини кривої розподілу. У разі виступання вершини ексцес називають додатним, у разі падіння – від'ємним.

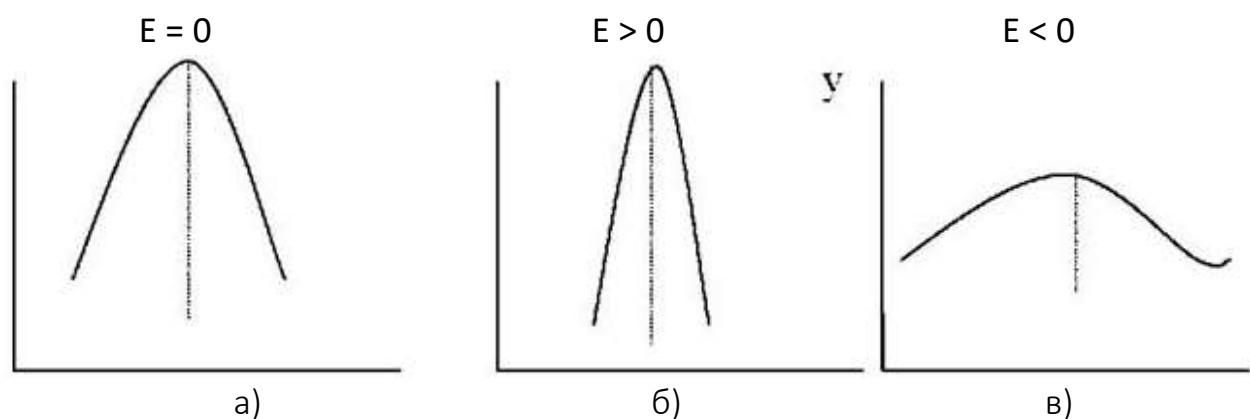


Рис. 3.2. Форми розподілу для різних значень ексцесу.

Для кількісного вимірювання гостровершинності використовують центральний момент четвертого порядку (μ_4), відношення якого до середньоквадратичного відхилення в четвертому ступені називають коефіцієнтом гостровершинності (ексцесом). Отже, коефіцієнт гостровершинності, або просто "ексцес" (E) можна виразити такою формулою:

$$E = \sqrt{\frac{\sum(x_i - \bar{x})^4}{n * \delta^4}} - 3$$

Якщо величина показника ексцесу приблизно дорівнює 0,4, криву вважають слабоексцесивною. Найбільша абсолютна величина від'ємного ексцесу становить мінус 2. При такому значенні вершина кривої опускається до осі абсцис і крива розподілу ділиться на дві самостійні одновершинні криві.

Зазначимо, що термін "ексцес" походить від грецького слова (*Kurtosis*), тому назви форми ексцесу в різних кривих розподілу мають корінь цього слова: стрічкокуртичні (рисунок 3.2б) і платокуртичні (рисунок 3.2в), мезокуртичні (рисунок 3.2а).

Відповідно до графіка для стрічкокуртичної кривої характерне розміщення більшості одиниць спостережень поблизу центра, для платокуртичної кривої (форма силуету плато) варіанти значно віддалені від центра розподілу. Помірне розміщення навколо центра розподілу визначає форма ексцесу у вигляді мезокуртичної кривої.

Асиметрія та ексцес нормального розподілу мають дорівнювати нулю. Якщо вони суттєво відрізняються від нуля (хоча б один із двох параметрів), це означає що отриманий розподіл відрізняється від нормального.

Якщо перевірка узгодженості емпіричного розподілу з нормальним теоретичним вказує на те що емпіричний розподіл відповідає нормальному, ми можемо вважати що отриманий розподіл даних можна розглядати як стійкий та репрезентативний стосовно генеральної сукупності, тобто на його основі можна визначити репрезентативні тестові

норми. Якщо перевірка не виявляє нормальності на заданому рівні, це може свідчити про те, що вибірка або мала та нерепрезентативна до заданої генеральної сукупності, або вимірювана властивість та будова тесту (спосіб підрахунку) взагалі не дають нормального розподілу.

Перевірити розподіл на відповідність нормальному можна за таблицями критичних значень асиметрії та ексцесу (додаток 1 таблиці 1 та 2), але цей спосіб є досить малопотужним та дозволяє лише передбачити що наш емпіричний розподіл є нормальним. Під час аналізу емпіричного розподілу може виникнути така ситуація, коли один показник відповідає нормальності а інший ні. У цьому випадку діє правило: якщо хоча б один показник виходить за межі критичних значень – розподіл відрізняється від нормального. Статистичні критерії, які дозволяють порівняти емпіричний розподіл з теоретичним нормальним, будуть нами детально розглянуті у розділі 5.1.

Питання для самоконтролю

1. Поняття міри центральної тенденції.
2. Поняття мода, середнє арифметичне значення, медіана.
3. Квантілі розподілу. Процентілі та квантілі.
4. Міри мінливості. Поняття розмах, дисперсія, стандартне відхилення.
5. Властивості дисперсії. Стандартизація.

Завдання для самостійної роботи

1. При визначенні ступеня вираженості деякої психічної властивості у двох групах, експериментальній та контрольній, бали розподілилися наступним чином: 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.

Дайте порівняльну характеристику ступеня виразності цієї властивості даних групах.

2. Під проведення дослідження особливостей прийняття рішень офіцерами була використана методика “Скриті фігури” Готтшальда, отримані наступні результати: 19, 16, 17, 12, 15, 16, 17, 17, 21, 23, 18, 13, 12, 13, 19, 20, 21, 27, 9, 12, 13, 26, 23, 14, 15, 22, 21, 16, 16, 18, 17, 10, 12, 17.

Розрахуйте аналітично та за допомогою програмних продуктів міри центральної тенденції, міри мінливості та міри зсуву. Відповідно до отриманих результатів зробіть висновок стосовно характеристик розподілу даних.

РОЗДІЛ 4

ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

- 4.1. Поняття статистичної гіпотези.*
- 4.2. Рівні статистичної значущості.*
- 4.3. Логіка та процедура перевірки гіпотез.*
- 4.4. Статистичні критерії розрізень.*

4.1. Поняття статистичної гіпотези

Отримані за результатами експерименту дані є підставою для судження про генеральну сукупність. Внаслідок випадкових ймовірнісних причин оцінки параметрів генеральної сукупності, зроблені на основі експериментальних даних, завжди матимуть похибку, тому їх розглядають як ймовірні. Ці припущення називають статистичними гіпотезами.

Для кращого розуміння розмежуємо поняття “наукова гіпотеза” і “статистична гіпотеза”. Наукові гіпотези формулюють як передбачуване вирішення проблеми. Статистична гіпотеза – це твердження про невідомий параметр, викладене мовою математичної статистики. Будь-яка наукова гіпотеза потребує перекладу на мову статистики. Після проведення конкретного експерименту перевіряють численні статистичні гіпотези, оскільки у кожному психологічному дослідженні реєструють не один, а безліч параметрів, і кожен з них характеризується кількома статистичними мірами: центральної тенденції, мінливості, розподілу. Можна також обчислити міри зв'язку між параметрами та оцінити їх значущість.

Отже, експериментальна гіпотеза слугує для організації експерименту, а статистична – для організації процедури порівняння реєстрованих параметрів. Статистична гіпотеза необхідна на етапі математичної інтерпретації даних емпіричних досліджень. Велика кількість статистичних гіпотез необхідна для підтвердження чи

спростування основної – експериментальної гіпотези. Таким чином, експериментальна гіпотеза є первинною, статистична – вторинною.

Статистична гіпотеза – це припущення на певному рівні статистичної значущості про властивості генеральної сукупності за оцінками вибірки. Іншими словами, це наукова гіпотеза, яка припускає статистичну перевірку. Сформульовану статистичну гіпотезу позначають літерою H (*Hypothesis*).

Розрізняють нульову й альтернативну гіпотези.

Нульова гіпотеза (H_0) є припущенням про відсутність відмінностей у значеннях ознак, які зіставляють (гіпотеза про схожість).

Альтернативна гіпотеза (H_1) – це припущення про існування відмінностей у значеннях ознак, які зіставляють (гіпотеза про відмінність).

Статистичні гіпотези можуть бути спрямованими і неспрямованими. Види статистичних гіпотез наведено на рисунку 4.1.

Статистичні гіпотези

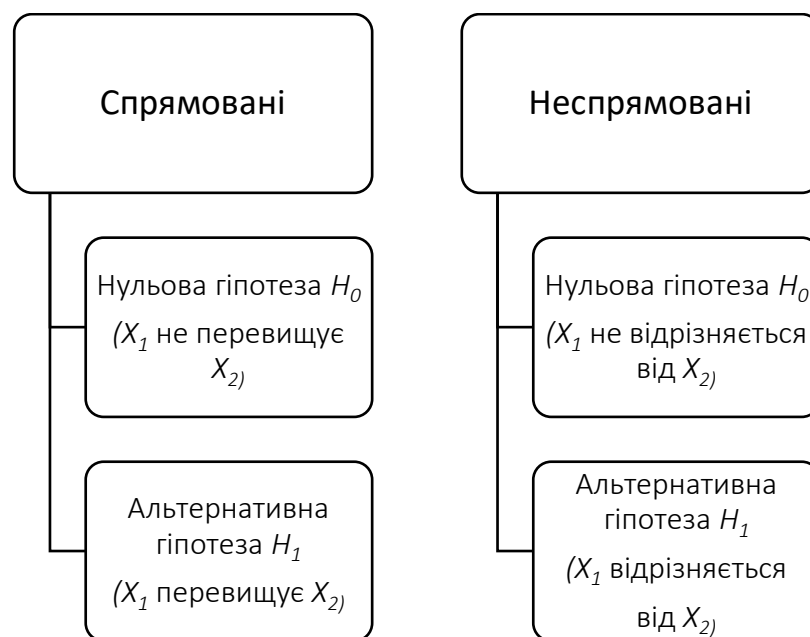


Рис. 4.1. Види статистичних гіпотез.

Спрямовані гіпотези формулюють, якщо необхідно перевірити значущість відмінностей між групою, в якій індивідуальні значення досліджуваних осіб за якою-небудь ознакою вищі, і групою, в якій

значення нижчі; якщо необхідно довести, що у групі А під впливом експериментальних дій відбулися більш виражені зміни, ніж у групі Б.

Неспрямовані гіпотези формулюють, якщо необхідно довести, що у групах А і Б форми розподілу ознаки відрізняються.

4.2. Рівні статистичної значущості

Провести лінію між прийняттям або неприйняттям нульової гіпотези неможливо абсолютно точно через наявність в експерименті випадкових впливів. Цю межу визначають поняттям рівня значущості.

Рівень значущості – це можливість помилкового відхилення нульової гіпотези, тобто поріг достовірності, показник ймовірності безпомилкових висновків і прогнозів.

У психології розрізняють такі рівні статистичної значущості: низький – 5% ($p \leq 0,05$); достатній – 1% ($p \leq 0,01$); високий – 0,1% ($p \leq 0,001$). У таблицях критичних значень зазвичай наводять значення критеріїв, що відповідають вказаним рівням статистичної значущості.

Коли зазначають, що відмінності достовірні на 5% рівня значущості (при $p = 0,05$), то мають на увазі, що ймовірність того, що відмінності все-таки недостовірні, становить 0,05. Іншими словами, висновки та прогнози дослідника справедливі у 95% випадків.

Коли зазначають, що відмінності достовірні на 1% рівня значущості (при $p = 0,01$), то мають на увазі, що ймовірність того, що відмінності все-таки недостовірні, становить 0,01.

Чим менший рівень значущості, тим менша ймовірність того, що ми припустилися помилки і тим надійніший зв'язок. Рівень $p = 0,05$ означає, що ймовірність припуститися помилки дорівнює п'яти помилкам у вибірці зі ста елементів (випадків, досліджуваних осіб).

Правила відхилення H_0 та прийняття H_1 :

якщо емпіричне значення критерію дорівнює критичному значенню, що відповідає $p = 0,05$ або перевищує його, то H_0 відхиляють, але ще не можна прийняти H_1 ;

якщо емпіричне значення критерію дорівнює критичному значенню, що відповідає $p = 0,01$ або перевищує його, то H_0 відхиляють і приймають H_1 .

Винятки: G -критерій знаків, T -критерій Вілкоксона та U -критерій Манна-Уїтні. Для них встановлені зворотні співвідношення.

Статистична перевірка гіпотез, заснована на експериментальних, вибірових даних, неминуче пов'язана з ризиком (ймовірністю) прийняти хибне рішення. При цьому можливі помилки двох родів.

Помилку 1-го роду допускають, якщо приймають рішення відхилити гіпотезу H_0 , хоча насправді вона правильна. Ця помилка може виникнути, якщо дослідник вважає, що виявив реальні результати, яких насправді немає.

Помилку 2-го роду допускають, якщо приймають рішення не відхилити гіпотезу H_0 , хоча насправді вона неправильна. У цьому разі дослідник може вважати, що не виявив результату, але насправді результат існує. Можливі помилки у процесі прийнятті статистичних рішень наведено в таблиці 4.1.

Таблиця 4.1

Помилки у процесі прийняття статистичних рішень

Результат перевірки гіпотези H_0	Можливі стани гіпотези, яку перевіряють	
	Гіпотеза H_0 правильна	Гіпотеза H_1 правильна
Гіпотеза H_0 відхилена	Помилка 1-го роду	Правильне рішення
Гіпотеза H_0 не відхилена	Правильне рішення	Помилка 2-го роду

На основі досвіду практичних досліджень зазначимо, що для уникнення помилок 1-го та 2-го роду під час обґрунтування висновків рекомендовано приймати рішення про наявність відмінностей (зв'язку), спираючись на рівень $p \leq 0,01$.

4.3. Логіка та процедура перевірки гіпотез

Прийняття статистичного рішення. На підставі отриманих експериментальних даних за попередньо обраним статистичним методом

обчислюють емпіричне значення ($X_{емп}$). Потім його порівнюють із критичними, які відповідають рівню значущості для обраного статистичного методу. Критичні величини для цього статистичного методу знаходять за допомогою відповідних таблиць.

Порівнювати отримане емпіричне значення із критичними (табличними) зручно за допомогою “осі значущості” – прямої, яка має три відокремлені зони: зону незначущості, зону невизначеності і зону значущості. Межами трьох зон є критичні значення $X_{кр1}$ та $X_{кр2}$ для рівнів $p = 0,05$ та $p = 0,01$ відповідно.



Праворуч від критичного значення $X_{кр2}$ розміщена “зона значущості”, до якої потрапляють емпіричні значення, що перевищують $X_{кр2}$, а отже, безумовно значущі.

Ліворуч від критичного значення $X_{кр1}$ розміщена “зона незначущості”, до якої потрапляють емпіричні значення, нижчі за $X_{кр1}$, а отже, безумовно незначущі.

Між критичними значеннями $X_{кр1}$ та $X_{кр2}$ розміщена “зона невизначеності”. У разі потрапляння до неї емпіричного значення можна відхилити гіпотезу про недостовірність відмінностей (H_0), але ще можна прийняти гіпотезу про їх достовірність (H_1).

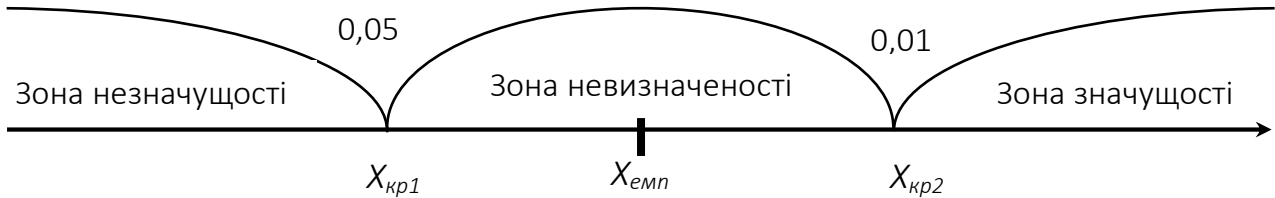
Отже, обчислене за допомогою певного статистичного методу $X_{емп}$ має потрапити до однієї з трьох зон. Розглянемо можливі варіанти.

Варіант 1: емпіричне значення потрапило до зони незначущості.



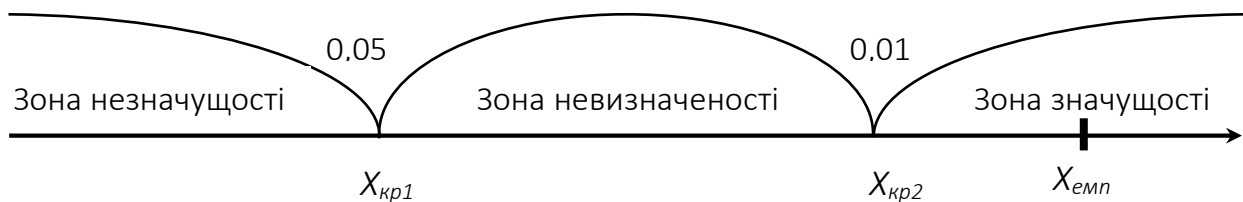
У цьому випадку приймають нульову гіпотезу H_0 про відсутність відмінностей (зв'язку).

Варіант 2: емпіричне значення потрапило до зони невизначеності.



У цьому випадку залежно від важливості вирішуваного завдання дослідник може вважати отриману статистичну оцінку достовірною на рівні 5% і прийняти гіпотезу H_1 , відхиливши гіпотезу H_0 , або, навпаки, визначити отриману статистичну оцінку як недостовірну на рівні 1%, прийнявши гіпотезу H_0 . Зазначимо, що це саме той випадок, коли психолог може припуститися помилок першого чи другого роду. За цих обставин найкращий вихід – це збільшення обсягу вибірки.

Варіант 3: емпіричне значення потрапило до зони значущості.



У цьому випадку приймають альтернативну гіпотезу H_1 про наявність відмінностей (зв'язку) на рівні значущості $p < 0,01$, а гіпотезу H_0 відхиляють.

Зауважимо, що величина $X_{емп}$ може точно збігтися з $X_{кр1}$ або $X_{кр2}$. У першому випадку можна вважати, що оцінка достовірна точно на рівні 5% і прийняти гіпотезу H_1 , або, навпаки, гіпотезу H_0 . У другому випадку, як - правило, приймають альтернативну гіпотезу H_1 про наявність відмінностей (зв'язку), а гіпотезу H_0 відхиляють.

Прийняття статистичного рішення можна навести в такому вигляді (рисунок 4.2).



Рис. 4.2. Алгоритм прийняття статистичного рішення.

4.4. Статистичні критерії розрізень

У психологічних та педагогічних дослідженнях для доказу ефективності програм, тренінгів, вправ, крім відстеження на певних етапах якісних змін, використовують математичну статистику. За допомогою математичних методів зіставляють результати “до” і “після” певного впливу, виявляють динаміку зміни показників під впливом експериментальних впливів, порівнюють контрольну й експериментальну

групи, оцінюють зміни за будь-яким психологічним показником у кількох групах та ін. Для вирішення такого великого класу завдань використовують статистичні критерії розрізень.

Статистичний критерій – інструмент для визначення рівня статистичної значущості. Статистичні критерії дають змогу оцінити рівень статистичної достовірності відмінностей між показниками, виміряними відповідно до плану проведення психологічного дослідження.

Статистичні критерії розрізняють за такими характеристиками:

тип вимірювальної шкали (для певних критеріїв можна застосовувати лише номінальні шкали, для інших – порядкову, інтервальну та шкалу відношень);

максимальний обсяг вибірки та кількість одночасно порівнюваних вибірок (існують критерії, що дають змогу оцінити відмінності відразу у трьох і більшій кількості вибірок, деякі – зіставляти нерівні за чисельністю вибірки);

якість вибірки (застосування критеріїв для незалежних та залежних вибірок);

потужність (здатність виявляти відмінності або відхиляти нульову гіпотезу, якщо вона неправильна).

Психолог може вирішувати експериментальні завдання з використанням різних статистичних критеріїв. При цьому можливо, що один критерій дає змогу виявити відмінності, а інший їх не виявляє, це означає, що перший критерій потужніший, ніж другий. Водночас, чим потужніший критерій, тим він більш трудомісткий, тому у виборі статистичного критерію слід керуватися простотою його обчислення. Зазначимо, що відсутність відмінностей, виявлена за допомогою критерію, не гарантує, що за допомогою потужнішого критерію їх теж не можна встановити.

Статистичні критерії розрізень поділяють на параметричні та непараметричні.

Параметричні критерії засновані на конкретному типі розподілу генеральної сукупності (як правило, нормальному) або використовують параметри цієї сукупності (середні, дисперсії тощо). Це *t*-критерій Стьюдента, *F*-критерій Фішера тощо.

Непараметричні критерії не базуються на припущенні про тип розподілу генеральної сукупності, не використовують параметри цієї сукупності та засновані на оперуванні частотами або рангами. Це *U*-критерій Манна-Уїтні, *T*-критерій Вілкоксона, *G*-критерій знаків, *Q*-критерій Розенбаума та ін.

При нормальному розподілі генеральної сукупності параметричні критерії мають більшу потужність порівняно з непараметричними, тобто здатні з більшою достовірністю відкидати нульову гіпотезу, якщо вона неправильна. Відповідно, коли вибірки взяті з нормально розподілених генеральних сукупностей, слід віддавати перевагу параметричним критеріям.

Однак, як показує практика, переважна більшість емпіричних даних, отриманих у психологічних експериментах, не розподілена нормально, тому застосування параметричних критеріїв під час аналізу результатів психологічних досліджень може призвести до помилок у статистичних висновках. У таких випадках непараметричні критерії виявляються потужнішими, тобто здатними з більшою достовірністю відкидати нульову гіпотезу.

Отже, під час оцінювання відмінностей у розподілах, далеких від нормального, непараметричні критерії можуть виявити значущі відмінності, а параметричні критерії таких відмінностей не виявляють. Зауважимо, що непараметричні критерії виявляють значущі відмінності, навіть якщо розподіл близький до нормального, а під час обчислень вручну непараметричні критерії значно менш трудомісткі, ніж параметричні.

Можливості й обмеження параметричних і непараметричних критеріїв наведено у таблиці 4.2.

Порівняльна характеристика статистичних критеріїв
(за О.В. Сидоренко)

№ п/п	Параметричні критерії	Непараметричні критерії
1	Дають змогу прямо оцінити різниці середніх, отриманих у двох вибірках (<i>t</i> -критерій Стьюдента)	Дають змогу оцінити лише середні тенденції, відповісти на питання, чи частіше у вибірці А трапляються вищі, а у вибірці Б – нижчі значення ознаки (критерії <i>Q</i> , <i>U</i> , φ^* тощо)
2	Дають змогу прямо оцінити відмінності в дисперсіях (<i>F</i> -критерій Фішера)	Уможливають оцінювання відмінностей лише в діапазонах варіативності ознаки (критерій φ^*)
3	Уможливають виявлення тенденції зміни ознаки у разі переходу від умови до умови при нормальному розподілі ознаки (однофакторний дисперсійний аналіз)	Уможливають виявлення тенденції зміни ознаки у разі переходу від умови до умови за будь-якого розподілу ознаки (критерії тенденцій <i>L</i> і <i>S</i>)
4	Уможливають оцінювання взаємодії двох і більше чинників у їх впливі на зміни ознаки (двофакторний дисперсійний аналіз)	Ця можливість відсутня
5	Експериментальні дані мають відповідати умовам: а) значення ознаки виміряні за інтервальною шкалою; б) розподіл ознаки є нормальним; в) у дисперсійному аналізі слід дотримуватися вимоги рівності дисперсій у комірках комплексу	Експериментальні дані можуть не відповідати жодній із цих умов: а) значення ознаки можуть бути відображені на будь-якій шкалі; б) розподіл ознаки може бути будь-яким, його збіг з теоретичним законом розподілу не потребує перевірки; в) вимогу до рівності дисперсій не висувають
6	Математичні розрахунки дуже складні	Математичні розрахунки здебільшого прості і можуть бути виконані швидко
7	Якщо умови, перераховані у п. 5, задоволено, параметричні критерії є дещо потужнішими, ніж непараметричні	Якщо умови, перераховані у п. 5, не задоволено, непараметричні критерії є потужнішими, ніж параметричні, оскільки менш чутливі до “засмічень”

Таким чином, у процесі підготовки експериментального дослідження необхідно завчасно визначити залежність та незалежність вибірок, їх обсяг, тип шкали та вид статистичного критерію.

Вибір статичного критерію визначається двома факторами: кількістю оброблюваних вибірок даних і завданням, що стоїть перед дослідником. Кількість вибірок, своєю чергою, залежить від кількості груп, у яких проводиться випробування.

Як було зазначено у розділі 2.1 дві (і більше) вибірки експериментальних даних називаються *залежними* (пов'язаними), якщо вони отримані на одній і тій же групі випробуваних; вибірки, отримані різних груп досліджуваних, називаються *незалежними*.

З певною часткою умовності можна виділити **дві базові схеми** організації дослідження:

схема 1 – є одна група досліджуваних; дослідник надає вплив сам або спостерігає за впливом інших факторів (середовища, віку та ін.); проводиться два або більше вимірів одного (декількох) критеріальних параметрів в різні моменти часу (тобто формуються залежні вибірки); можливе вирішення наступних завдань:

- виявлення достовірності зміни (зсуву) одного параметра (2 виміри) або тенденції його зміни (3 і більше вимірів);

- при одночасному (паралельному) вимірі двох параметрів – виявлення ступеня узгодженості зміни цих параметрів (кореляційний аналіз);

- при одночасному (паралельному) вимірі одного параметра та інтенсивності одного або двох контрольованих (вимірюваних факторів – виявлення впливу факторів на параметр (дисперсійний аналіз).

схема 2 – є дві групи (найчастіше – це експериментальна, на яку здійснювався вплив, і контрольна, яка навчалася за "традиційною" схемою); одноразово вимірюється один параметр; можливе вирішення наступних завдань:

- виявлення відмінностей значення параметра між вибірками;

– виявлення відмінностей у характері розподілу вимірюваного параметра за випробуваними у вибірках.

Безумовно, можливі якісь комбіновані схеми проведення дослідження, наприклад проведення рознесених у часі вимірювань як в експериментальній, так і в контрольній групах – у цьому випадку відкривається можливість зіставляти динамічні показники між групами. Інший приклад – одночасне дослідження в декількох експериментальних і кількох контрольних групах.

Вибір статистичного критерію, що відповідає задачі, може бути здійснений згідно з таблицею 4.3:

Таблиця 4.3

Розподіл статистичних критеріїв відповідно до завдань

№	Завдання		Характер вибірок	
			Залежні	Незалежні
1	Виявлення відмінностей у рівні ознаки	2 вибірки	–	Q, U, φ^*, t
		3 і більше вибірок	–	S, H
2	Оцінка достовірності зсуву рівнів ознаки	2 вимірювання	M, G, T, φ^*, t	–
		3 і більше вимірювань	L	–
3	Оцінювання відповідності емпіричного розподілу теоретичному (критерії згоди)		$A, E, t_A, t_E, \chi^2, \lambda$	$A, E, t_A, t_E, \chi^2, \lambda$
4	Виявлення відмінностей у двох емпіричних розподілах ознаки		$\chi^2, \lambda, \varphi^*$	$\chi^2, \lambda, \varphi^*$
5	Виявлення ступеня узгодженості змін двох ознак (кореляційний аналіз)		R, r_s, r	R, r_s, r
6	Аналіз змін ознаки під впливом контролюємих факторів (дисперсійний аналіз)	1 фактор	L, F_1	S, F_1
		2 фактори одночасно	F_2	F_2

В таблиці прийнято такі позначення критеріїв:

Q – критерій Розенбаума;

U – критерій Манна-Уїтні;

φ^* – критерій (кутове перетворення) Фішера;

S – критерій тенденцій Джонкіра;

H – критерій Крускала-Уолліса;
 M – критерій Макнамари;
 T – критерій Вілкоксона;
 G – критерій знаків;
 t – критерій Стюдента;
 χ^2 – критерій Пірсона;
 L – критерій тенденцій Пейджа;
 λ – критерій Колмогорова-Смирнова;
 A та E – значення асиметрії та ексцесу;
 t_A і t_E – критерій асиметрії та ексцесу;
 R – узагальнене позначення критеріїв кореляційного аналізу
 r_s – Коефіцієнт рангової кореляції Спірмена;
 r – Коефіцієнт лінійної кореляції Пірсона;
 F_1 – критерій Фішера для однофакторного дисперсійного аналізу;
 F_2 – критерій Фішера для двофакторного дисперсійного аналізу.

Очевидно, кращим слід вважати підхід, у якому дослідник вибирає критеріальні параметри, організаційну схему досліджень і методи обробки на етапі планування експерименту. Однак у практиці психологічних та педагогічних досліджень дуже поширена ситуація, коли спочатку отримують експериментальні дані, а потім здійснюють вибір статистичного методу – у цьому випадку наведена вище таблиця і опис можливих схем допоможуть такий вибір здійснити. Якщо одне й те завдання вирішується декількома методами, потрібно починати з найпростішого, і тільки якщо він не дає підстав прийняти експериментальну гіпотезу (H_1), слід звернутися до більш складних і, як правило, більш потужних методів. Якщо навіть найбільш потужний з однотипних методів не дає підстав прийняти H_1 вона повинна бути відхилена в даному формулюванні і для даної організації експерименту. Однак навіть у цьому випадку нічого не заважає досліднику змінити формулювання гіпотези, схему експерименту, обсяги вибірок або використовуваний статистичний критерій і постаратися отримати докази. Іншими словами, процедура вибору та застосування критеріального

статистичного методу – процес неоднозначний і не самоочевидний – він вимагає, з одного боку, знання цих методів, їх можливостей та умов застосовності, а, з іншого боку, критичного та сумлінного відношення до власних експериментальних даних, щоб бачити в них не те, що хотілося б, а те, що вони реально представляють.

Питання для самоконтролю

1. Дайте визначення статистичної гіпотези.
2. Дайте визначення нульової та альтернативної гіпотез.
3. У якому випадку досліднику необхідно сформулювати спрямовану гіпотезу?
4. У якому випадку досліднику необхідно сформулювати неспрямовану гіпотезу?
5. Дайте визначення рівня статистичної значущості. Які рівні значущості використовують у психологічних дослідженнях?
6. Охарактеризуйте помилки 1-го та 2-го роду.
7. Розкрийте алгоритм прийняття статистичного рішення.
8. Дайте визначення статистичного критерію.
9. За якими характеристиками розрізняють статистичні критерії?
10. Охарактеризуйте параметричні і непараметричні статистичні критерії розрізень.

Завдання для самостійної роботи

1. Сформулюйте нульову та альтернативну гіпотези для вказаної проблеми: офіцер-психолог дослідив психофізіологічну адаптацію молодих офіцерів до умов військової служби. Чи можна стверджувати, що існують стійкі індивідуально-психологічні якості, сформованість яких необхідно враховувати у разі призначення на нову посаду?

2. Офіцер-психолог провів корекційну роботу зі зниження реактивної тривожності у підрозділі. Чи сприяла корекційна робота зниженню реактивної тривожності у досліджуваних військово-службовців? Сформулюйте статистичні гіпотези.

РОЗДІЛ 5

МЕТОДИ ПОРІВНЯННЯ ВИБІРОК ДОСЛІДЖУВАНИХ

5.1. Оцінювання відповідності емпіричного закону нормальному розподілу.

5.2. Оцінювання достовірності зсуву.

5.3. Оцінювання достовірності відмінностей.

5.1. Оцінювання відповідності емпіричного закону нормальному розподілу

Вибірковий емпіричний розподіл та його графічне зображення не дають змоги зробити висновки щодо закону розподілу ознаки у сукупності, з якої взята вибірка. Знання цього закону є необхідною умовою застосування багатьох математичних методів. Наприклад, параметричні критерії, дисперсійний аналіз та інші методи використовують лише за умови нормального закону досліджуваної ознаки.

Тому будь-яка статистична обробка починається, як правило, зі спроби оцінити закон розподілу. Прагнення застосувати методи, які розроблено для певного закону розподілу, в умовах, коли реальний розподіл відрізняється від гіпотетичного, є найбільш розповсюдженою помилкою, що призводить у підсумку і до помилкових висновків.

Критерії перевірки гіпотез щодо закону розподілу прийнято називати **критеріями згоди**, які можна розділити на дві групи: загальні та спеціальні. Загальні критерії застосовують до формулювань гіпотез про згоду спостережень з будь-яким можливим розподілом. Спеціальні критерії згоди використовують у разі перевірки гіпотези щодо конкретної форми розподілу – нормальної, рівномірної, експоненціальної тощо. Такі критерії носять відповідну назву – критерії нормальності, критерії рівномірності та ін.

Розрахунки емпіричного розподілу та його графічна візуалізація не дають надійних підстав для висновку щодо закону розподілу ознаки у сукупності, з якої взята вибірка. Тим часом знання цього закону є

необхідною умовою використання багатьох математичних методів. Наприклад, застосування параметричних критеріїв, дисперсійного аналізу вимагає попередньої перевірки нормальності розподілу досліджуваної ознаки.

Серед методів оцінювання законів розподілу ймовірностей випадкових величин біля двох десятків було спеціально розроблено для перевірки нормальності. Найбільш розповсюдженими вважаються критерії асиметрії й ексцесу, λ -критерій Колмогорова-Смірнова, χ^2 -квадрат Пірсона та ін. Спробуємо розглянути деякі із них.

Критерії асиметрії та ексцесу t_A і t_E

Призначення. Критерії асиметрії та ексцесу застосовують для приблизної перевірки гіпотези про нормальність емпіричного розподілу. Асиметрія характеризує ступінь несиметричності, ексцес – ступінь загостреності (згладженості) кривої диференціальної функції емпіричного розподілу у порівнянні з функцією щільності нормального розподілу. У підрозділі 3.4. ми вже розглядали як безпосередньо за вирахованими значеннями асиметрії та ексцесу зробити висновки стосовно відповідності розподілу нормальному закону, але існують і спеціальні критерії, які дозволяють це зробити.

Для ідеального (теоретичного) нормального розподілу критерії асиметрії та ексцесу дорівнюють нулю. Таблиць критичних значень для вказаних критеріїв немає, про достовірну відмінність емпіричного розподілу від нормального (теоретичного) свідчитиме якщо критерії t_A і t_E прийматимуть значення **3 та більше**.

Приклад: отримано наступні емпіричні дані: 1, 2, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6. Перевірити відповідність розподілу цих даних нормальному закону розподілу ознаки.

Розв'язання.

1. Сформулюємо статистичні гіпотези:

H_0 – емпіричний розподіл не відрізняється від нормального.

H_1 – емпіричний розподіл відрізняється від нормального.

2. Перевірити припущення: вимірювання зроблено за шкалою інтервалів.

3. Статистичний критерій розраховується за формулами:

$$t_A = \frac{|A|}{m_A} ; \quad t_E = \frac{|E|}{m_E}$$

де A і E – значення асиметрії та ексцесу; m_A і m_E – відповідні похибки:

$$m_A = \sqrt{\frac{6*(n-1)}{(n+1)*(n+3)}} ; \quad m_E = \sqrt{\frac{24*n*(n-2)*(n-3)*(n-5)}{(n+1)^2*(n+3)*(n+5)}}$$

Формули для розрахунку A і E наведені у підрозділі 3.4 цього посібника.

4. Для розрахунку критеріїв можна скористатися формулами, або табличним оператором Excel. Готових розрахунків цих критеріїв у статистичних пакетах SPSS або Statistika не передбачено.

5. За результатами розрахунків отримано результати $t_A \approx 1,31$ та $t_E \approx 0,09$, тобто не існує статистично значущих підстав стверджувати про достовірну відмінність емпіричного розподілу від теоретичного нормального.

Відповідь: Оскільки t_A і $t_E \leq 3$, приймається нульова гіпотеза, тобто емпіричний розподіл не відрізняється від нормального.

λ-критерій Колмогорова-Смірнова

Призначення. Найбільш ефективним критерієм при перевірці нормальності розподілу вважається λ-критерій Колмогорова-Смірнова. Критерій дозволяє оцінити ймовірність того, що вибірка належить до генеральної сукупності з нормальним розподілом. Якщо ймовірність $p \leq 0,05$, то даний емпіричний розподіл суттєво відрізняється від нормального, а якщо $p > 0,05$, даний розподіл приблизно відповідає нормальному.

Ідея застосування даного критерію полягає у знаходженні точки, в якій накопичене відхилення одного розподілу від іншого виявляється

максимальним (порівнюватися можуть емпіричний розподіл з теоретичним, або ж два емпіричних), і оцінки статистичної достовірності цього відхилення. Накопиченими відхилення (відмінності) двох розподілів називаються з тієї причини, що у даному методі проводиться послідовне порівняння частот попадання в першу градацію, потім в першу і другу, далі в першу, другу і третю і т. д. Якщо відмінності між двома розподілами статистично достовірні, то на якомусь кроці накопичене відхилення перевищить критичне значення (яке, як завжди, визначається за таблицею).

Приклад: на етапі проходження психологічного обстеження до вступу у ВНЗ психологом була застосована методика «Кількісні відношення», результати наведені у таблиці 5.1. З метою подальшого вибору статистичних критеріїв необхідно визначити чи відповідає сукупність отриманих даних закону нормального розподілу?

Таблиця 5.1

№ спост.	Результат	№ спост	Результат	№ спост	Результат	№ спост	Результат
1	9	9	12	17	12	25	12
2	13	10	16	18	10	26	13
3	10	11	13	19	14	27	13
4	11	12	11	20	11	28	11
5	10	13	13	21	15	29	13
6	12	14	15	22	15	30	15
7	13	15	12	23	14	31	13
8	14	16	13	24	14	32	14

Значення λ -критерій Колмогорова-Смірнова нескладно розрахувати за допомогою програми Excel Microsoft:

1. Створити таблицю частот випадіння діапазону значень.
2. Для емпіричних значень порахувати відносну частоту випадіння шляхом ділення частоти випадіння окремого значення на суму усіх вимірів.
3. Порахувати накопичену відносну частоту для кожного значення.
4. Ті самі дії здійснити стосовно теоретичного розподілу.

5. Розрахувати абсолютні (по модулю) різниці накопичених частот шляхом віднімання накопиченої відносної частоти емпіричного розподілу від теоретичного, для кожного значення яке приймає змінна, позначити їх як d .

6. Серед усіх отриманих різниць значень обрати найбільше, позначити його як d_{max} , воно ж і є нашим $\lambda_{емп}$. У нашому випадку це значення дорівнює 0,125 та отримано аж у трьох випадках, тому приймаємо будь яке з них.

7. За таблицею критичних значень (додаток 1 таблиця 3) знаходимо $\lambda_{0,01} = 0,297$; $\lambda_{0,05} = 0,248$.

8. Наше значення не перевищує критичних, тому різниці не можуть вважатися достовірними, тобто підтверджується H_0 : наш емпіричний розподіл відповідає нормальному.

Результат (кількість балів)	Емпіричні дані			Теоретичні дані		Різниця накоплених частот
	Частота	Відносна частота	Накопичена відносна частота	Відносна частота	Накопичена відносна частота	
9	1	0,031	0,031	0,125	0,125	0,094
10	3	0,094	0,125	0,125	0,250	0,125
11	4	0,125	0,250	0,125	0,375	0,125
12	5	0,156	0,406	0,125	0,500	0,094
13	10	0,313	0,719	0,125	0,625	0,094
14	5	0,156	0,875	0,125	0,750	0,125
15	3	0,094	0,969	0,125	0,875	0,094
16	1	0,031	1,000	0,125	1,000	0,000
Суми	32	1,000		1,000	$d_{max} =$	0,125

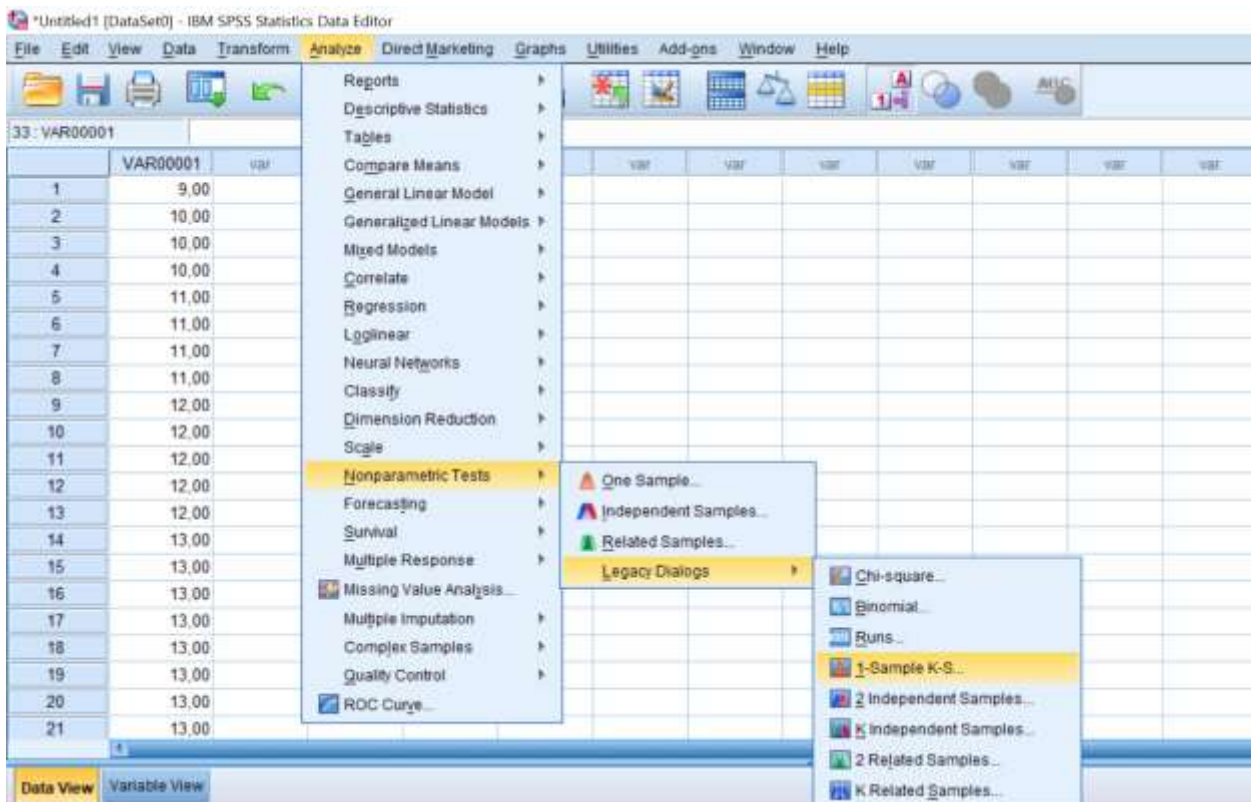
Обчислення λ -критерію Колмогорова-Смірнова за допомогою SPSS.

Розв'яжемо попереднє завдання за допомогою комп'ютерної програми SPSS.

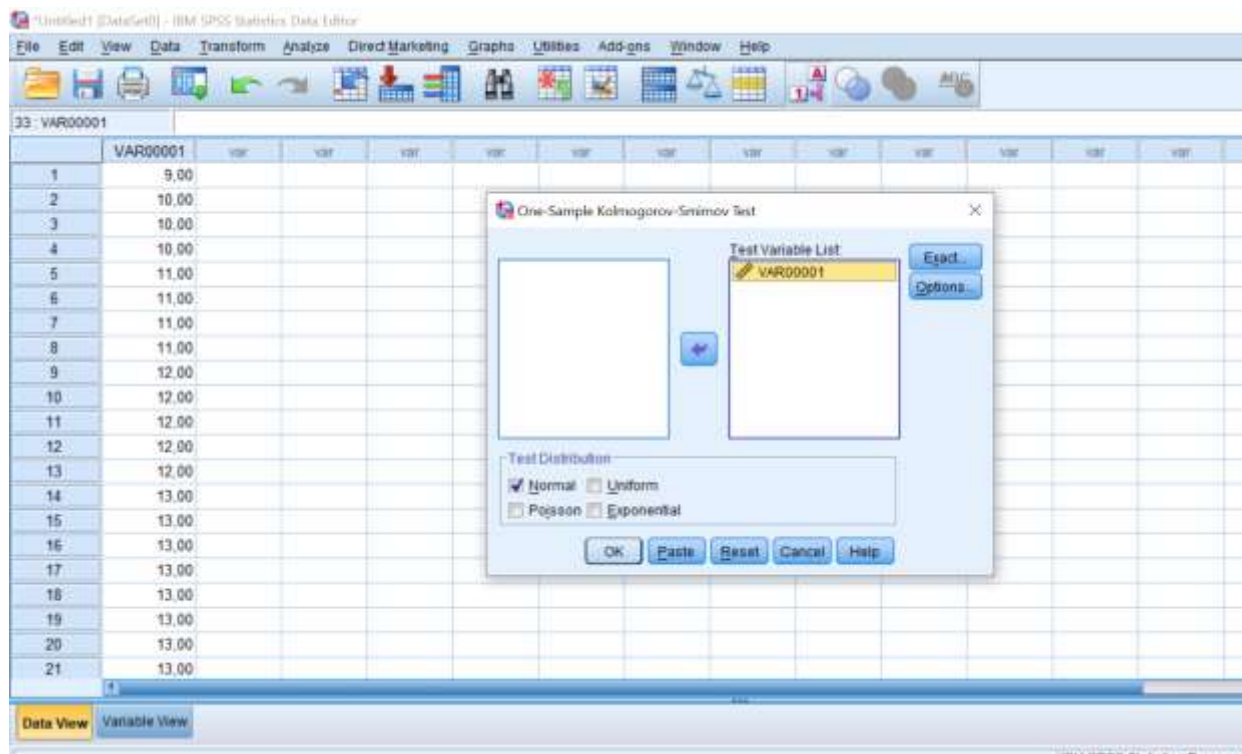
1. Вводимо дані в таблицю у стовпець: var1.

	VAR00001	var1	var2	var3	var4	var5	var6	var7	var8	var9	var10	var11	var12	var13
1	9,00													
2	10,00													
3	10,00													
4	10,00													
5	11,00													
6	11,00													
7	11,00													
8	11,00													
9	12,00													
10	12,00													
11	12,00													
12	12,00													
13	12,00													
14	13,00													
15	13,00													
16	13,00													
17	13,00													
18	13,00													
19	13,00													
20	13,00													
21	13,00													

2. У верхньому меню обираємо Analyze → Nonparametric Tests → Legacy Dialogs → 1-Sample K-S....



3. У меню, яке з'явилося, виділяємо змінну var1 і за допомогою кнопки ► переносимо її у праве вікно. У цьому вікні у розділі Test Distribution ставимо прапорець на Normal (див. рисунок нижче).



4. Натискаємо ОК та отримуємо такий результат:

One-Sample Kolmogorov-Smirnov Test

		VAR00001
N		32
Normal Parameters ^{a,b}	Mean	12,6250
	Std. Deviation	1,66074
Most Extreme Differences	Absolute	,183
	Positive	,129
	Negative	-,183
Kolmogorov-Smirnov Z		1,036
Asymp. Sig. (2-tailed)		,234

a. Test distribution is Normal.

b. Calculated from data.

Таблиця Test Statistics^a містить результат перевірки гіпотези: р-рівень значущості (Asymp. Sig. (2-tailed)), який дорівнює 0,234. Так як $0,234 > 0,01$, приймаємо гіпотезу H_0 про відсутність статистично значущих відмінностей нашого емпіричного розподілу від нормального.

Алгоритм розрахунку λ -критерію Колмогорова-Смірнова.

1. Перевірити, чи виконуються умови застосування критерію λ .
2. Сформулювати основну і альтернативну гіпотези.

H_0 : відмінності між двома розподілами недостовірні;

H_1 : відмінності між двома розподілами достовірні.

3. Занести в таблицю розряди і відповідні частоти порівнюваних розподілів.

4. Обчислити відносні частоти порівнюваних розподілів.

5. Розрахувати накопичені відносні частоти порівнюваних розподілів.

6. Обчислити абсолютні значення різниць відповідних накопичених відносних частот порівнюваних розподілів. Найбільшу різницю позначити d_{max} .

7. При порівнянні двох емпіричних розподілів розрахувати емпіричне значення критерію λ за формулою:

$$\lambda_{емп} = d_{max} \sqrt{\frac{n_1 * n_2}{n_1 + n_2}}$$

де n_1, n_2 – обсяги вибірок.

8. При порівнянні двох емпіричних розподілів за (додаток 1 таблиця 4) визначити рівень статистичної значущості, який відповідає отриманому значенню $\lambda_{емп}$ і на цій основі дійти висновку про прийняття чи відхилення основної гіпотези. Якщо рівень статистичної значущості не перевищує 0,01, гіпотеза H_0 відхиляється, а якщо перевищує 0,05, ця гіпотеза приймається. В інших випадках гіпотеза H_0 відхиляється на рівні значущості, що відповідає отриманому значенню $\lambda_{емп}$.

У разі порівняння емпіричного розподілу з теоретичним для заданого обсягу вибірки n визначити критичне значення d_{max} . Якщо $d_{max} > d_{0,01}$, гіпотеза H_0 відхиляється, якщо $d_{max} < d_{0,05}$ – приймається. Якщо $d_{0,05} < d_{max} < d_{0,01}$, гіпотеза H_0 відхиляється на рівні значущості 0,05 (або 5 %).

Обмеження λ -критерію Колмогорова-Смірнова.

1. Критерій вимагає, щоб вибірка була досить великою. При зіставленні двох емпіричних розподілів необхідно, щоб $n > 50$. Зіставлення емпіричного розподілу з теоретичним іноді допускається за $n > 5$.

2. Розряди мають бути впорядковані за наростанням чи спаданням будь-якої ознаки. Вони обов'язково повинні відображати якусь

односпрямовану його зміну. Наприклад, ми можемо за розряди прийняти дні тижня, 1-й, 2-й, 3-й місяці після проходження курсу терапії, підвищення температури тіла, посилення почуття недостатності тощо. У той же час, якщо ми візьмемо розряди, які випадково виявилися вибудованими в дану послідовність, то і накопичення частот буде відображати лише цей елемент випадкового сусідства розрядів. Наприклад, якщо шість стимульних картин у методиці Хекхаузена різним випробуванням пред'являються в різному порядку, ми не маємо права говорити про накопичення реакцій при переході від картини №1 стандартного набору до картини №2 та ін. Ми не можемо говорити про односпрямовану зміну ознаки при зіставленні категорій "черговість народження", "національність", "специфіка здобутої освіти" та ін. Ці дані являють собою номінативні шкали: у них немає жодної однозначної односпрямованої зміни ознаки.

Отже, ми не можемо накопичувати частоти за розрядами, які відрізняються лише якісно і не є шкалами порядку. У всіх тих випадках, коли розряди є невпорядкованими за зростанням або спаданням якоїсь ознаки категорії, нам слід застосовувати метод χ^2 .

5.2. Оцінювання достовірності зсуву

Критерій знаків G

Призначення. За допомогою критерію знаків G визначають загальний напрям зсуву досліджуваної ознаки. Цей критерій дає змогу встановити, в який бік у вибірці загалом змінюються значення ознаки під час переходу від першого вимірювання до другого: чи змінюються показники у бік поліпшення, підвищення чи посилення або, навпаки, у бік погіршення, зниження чи ослаблення. Критерій знаків G є непараметричним, його застосовують до даних, отриманих у ранговій, інтервальної шкалах та шкалі відношень тільки у залежних (зв'язаних) вибірках.

Критерій знаків G достатньо простий для обчислення і передбачає лише встановлення знаку відмінностей між парами порівнюваних величин досліджуваної ознаки до та після впливу. Після підсумовування кількості позитивних та негативних зсувів визначають їх співвідношення, яке порівнюють із критичним значенням. Вплив чинника вважають статистично достовірним, якщо розбіжності перевищують критичне значення. Якщо в експериментальних даних є нульові зсуви (досліджувана ознака не змінюється під впливом чинника), такі результати виключають із розгляду, а кількість спостережень n зменшують на кількість нульових зсувів.

Приклад: визначити, чи сприяв проведений тренінг підвищенню показників за досліджуваною ознакою. Результати, отримані до та після експериментального впливу, наведено в таблиці 5.2.

Таблиця 5.2

Результати діагностики до та після експериментального впливу

n	Показники досліджуваної ознаки	
	До участі у тренінгу	Після участі у 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

Розв’язання. Сформулюємо статистичні гіпотези:

H_0 – зсув показників після тренінгу випадковий;

H_1 – зсув показників після тренінгу не випадковий.

Визначимо “зсув”, тобто різницю між показниками кожного учасника “після” та “до” участі у тренінгу (таблиця 5.3).

Таблиця 5.3

Результати діагностики до та після експериментального впливу

n	Показники досліджуваної ознаки		Зсув
	До участі у тренінгу	Після участі у тренінгу	
1	3	8	+5
2	5	8	+3
3	5	9	+4
4	8	9	+1
5	6	7	+1
6	4	8	+4
7	8	9	+1
8	3	3	0
9	5	6	+1
10	5	8	+3

Обчислимо загальну кількість нульових, позитивних і негативних зсувів: нульових зсувів - 1, позитивних зсувів - 9, негативних зсувів немає. Як вказано вище, враховують суми лише позитивних і негативних зсувів, а нульові зсуви не враховують.

Суму зсувів, яка виявилася найбільшою, незалежно від того, позитивні зсуви чи негативні, називають *типовим зсувом* і позначають літерою n . Типовий зсув використовують під час роботи з таблицею критичних значень для критерію знаків G , в якій наводять критичні величини 5% і 1% рівнів значущості. Суму зсувів, яка виявилася найменшою, називають *нетиповим зсувом* і позначають G_{emn} . Нетиповий зсув G_{emn} вказують на осі значущості.

NB! Якщо загальна кількість позитивних і негативних зсувів однакова, то критерій G -знаків застосовувати не можна. У цих випадках застосовують інші критерії, наприклад T -критерій Вілкоксона.

Для нашого прикладу типовий зсув $n = 9$, а нетиповий $G_{emn} = 0$. Для оцінювання статистичної достовірності зсуву знаходимо критичні значення критерію G -знаків для $n = 9$:

n	p	
	0,05	0,01
9	1	0

Побудуємо “вісь значущості”, де розташуємо критичні значення $G_{0,05} = 1$, $G_{0,01} = 0$ та емпіричне значення $G_{емп} = 0$.



NB! Для критерію G -знаків “вісь значущості” перевернута. Нуль на числовій осі розташований справа, а збільшення числового ряду відбувається у протилежний бік, тобто справа наліво. Вказане пов’язано з тим, що чим більша є кількість нетипових зсувів, тим менша ймовірність того, що сумарний зсув виявиться статистично достовірним. Такий тип розташування “осі значущості” справедливий також для T -критерію Вілкоксона та U -критерію Манна-Уїтні.

У нашому прикладі отримане $G_{емп}$ збіглося з критичним значенням зони значущості $G_{0,01} = 0$. Отже, відхиляємо гіпотезу H_0 та приймаємо гіпотезу H_1 : зсув показників після тренінгу не випадковий і є статистично значущим на рівні $p = 0,01$. Таким чином, проведений тренінг статистично достовірно сприяв підвищенню показників за досліджуваною ознакою.

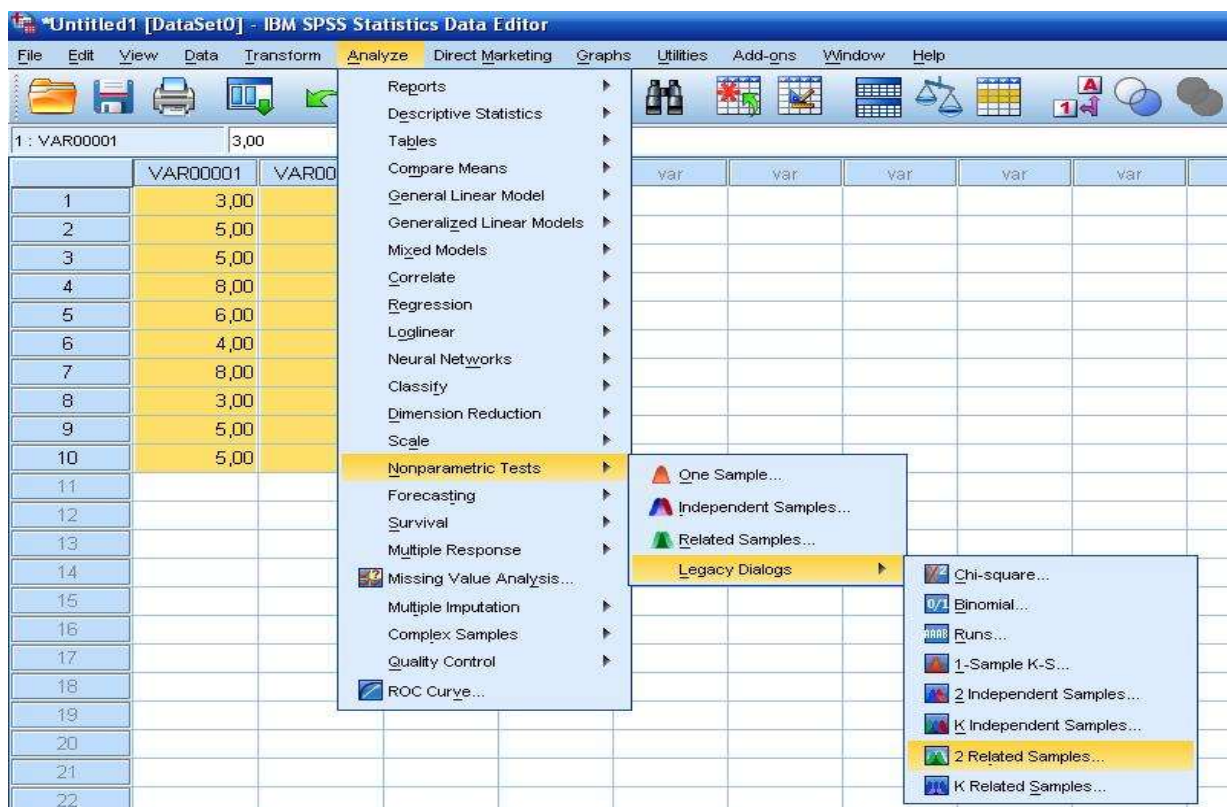
Обчислення критерію знаків G за допомогою SPSS

Розв’яжемо попереднє завдання за допомогою комп’ютерної програми SPSS.

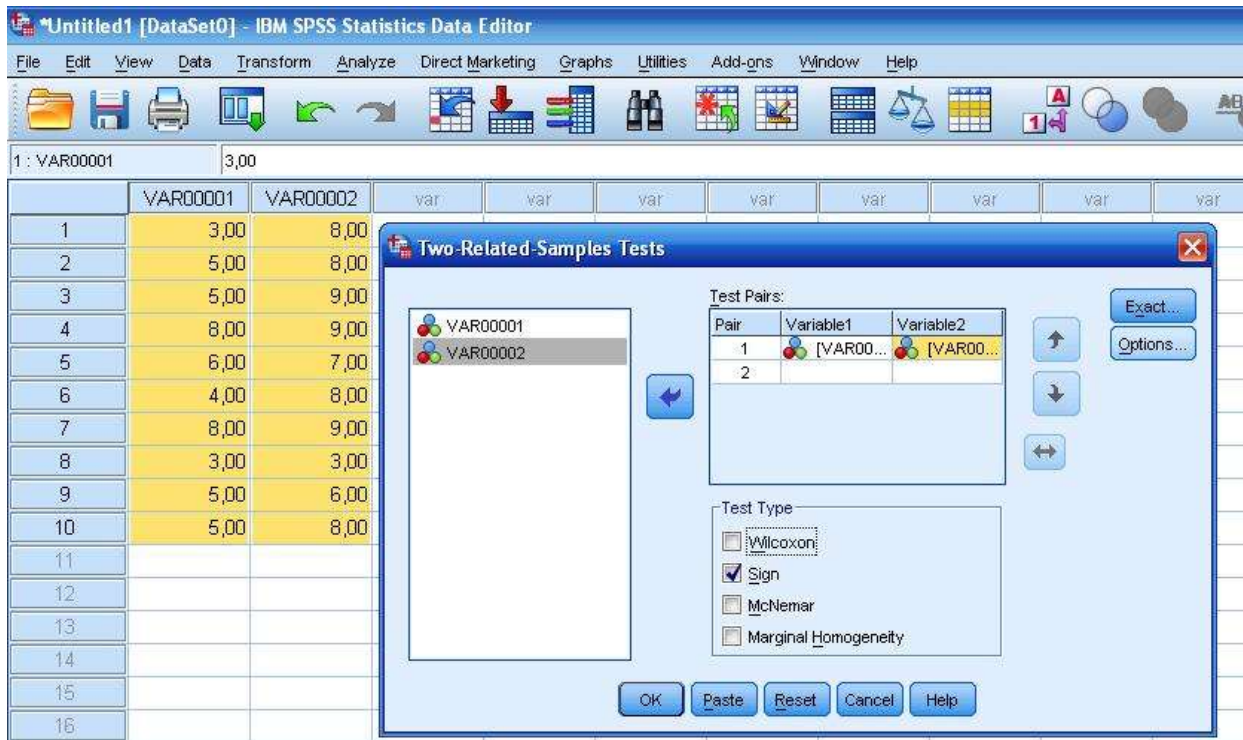
1. Вводимо дані в таблицю у два стовпці: var1 та var2.

	VAR00001	VAR00002	var
1	3,00	8,00	
2	5,00	8,00	
3	5,00	9,00	
4	8,00	9,00	
5	6,00	7,00	
6	4,00	8,00	
7	8,00	9,00	
8	3,00	3,00	
9	5,00	6,00	
10	5,00	8,00	
11			

2. У верхньому меню обираємо Analyze → Nonparametric Tests → Legacy Dialogs → 2 Related Samples.



3. У меню, яке з'явилося, виділяємо дві змінні var1 і var2, за допомогою кнопки ► переносимо їх у праве вікно, де у розділі Test Type ставимо прапорець на Sign.



4. Натискаємо OK та отримуємо такий результат.

Frequencies

	N
VAR00002 - VAR00001 Negative Differences ^a	0
Positive Differences ^b	9
Ties ^c	1
Total	10

a. VAR00002 < VAR00001

b. VAR00002 > VAR00001

c. VAR00002 = VAR00001

Test Statistics^a

	VAR00002 - VAR00001
Exact Sig. (2-tailed)	,004 ^b

a. Sign Test

b. Binomial distribution used.

Таблиця Frequencies містить дані про загальну кількість негативних (Negative Differences^a), позитивних (Positive Differences^b) та нульових (Ties^c) зсувів.

Таблиця Test Statistics^a містить результат перевірки гіпотези: p -рівень значущості (Exact Sig. (2-tailed)) дорівнює 0,004. Через те, що $0,004 < 0,01$, приймаємо гіпотезу H_1 про наявність статистично достовірного зсуву.

Алгоритм розрахунку критерію знаків G :

1. Обчислити загальну кількість нульових, позитивних та негативних зсувів (різниць між показниками “після” та “до” експериментального впливу). Виключити з розрахунку нульові зсуви. У разі однакової кількості позитивних і негативних зсувів обрати інші критерії.

2. Визначити переважний напрям змін (суму зсувів, яка виявилася найбільшою). Вважати зсуви у переважному напрямі “типовими”. Типовий зсув позначити літерою n .

3. Визначити кількість “нетипових” зсувів (суму зсувів, яка виявилася найменшою). Вважати це число емпіричним значенням G .

4. Визначити критичні значення критерію G -знаків для отриманого n за таблицею.

5. Порівняти $G_{емп}$ та $G_{кр}$ для $p = 0,05$ та $p = 0,01$ (додаток 1 таблиця 5). Якщо $G_{емп}$ менше $G_{кр}$ або принаймні дорівнює йому, типовий зсув можна вважати достовірним.

Критерій знаків G є одним із найпростіших за способом обчислення. Традиційно його вважають одним із найменш потужних. Однак можна стверджувати, що якщо критерій знаків показав значущі відмінності на рівні $p = 0,01$, то інші, потужніші критерії, також підтвердять ці відмінності. Водночас, якщо за допомогою критерію знаків G значущих відмінностей не виявлено, то цілком можливо, що потужніші критерії зможуть встановити відмінності.

У застосуванні критерію знаків G необхідно дотримуватися таких умов:

1. Вимірювання проводити за ранговою, інтервальною шкалами, чи шкалою відношень.

2. Вибірка створювати однорідну та залежну (зв'язану).

3. Кількість елементів у порівнюваних вибірках забезпечують однаковою.

4. Критерій знаків G застосовують для вибірки від 5 до 300 досліджуваних осіб (на більшу величину значень типових зсувів не розрахована таблиця критичних значень).

5. У разі великої кількості парних значень, які порівнюють, критерій знаків G є достатньо ефективним.

6. У разі виявлення однакової кількості типових і нетипових зсувів критерій знаків G не застосовують, а беруть інші критерії.

T -критерій Вілкоксона

Призначення. T -критерій Вілкоксона застосовують для оцінювання відмінностей в експериментальних даних, отриманих за двох різних умов на тій самій вибірці досліджуваних. Критерій дає змогу виявити не тільки спрямованість змін (критерій G -знаків), а й встановити, наскільки зсув показників в одному напрямі інтенсивніший, ніж в іншому. T -критерій Вілкоксона потужніший за критерій G -знаків та може бути використаний, якщо досліджувану ознаку виміряно за шкалою порядку або шкалою інтервалів. Обсяг вибірки – від 5 до 50 осіб.

T -критерій Вілкоксона заснований на ранжуванні абсолютних величин різниці між двома рядами вибірових значень у першому та другому експерименті (наприклад “до” та “після” будь-якого впливу). Ранжування абсолютних величин означає, що знаки різниць не враховують, проте надалі разом із загальною сумою рангів розподіляють суму рангів позитивних і негативних зсувів. Якщо інтенсивність зсуву в одному з напрямів виявляється більшою, відповідна сума рангів також більша. Цей зсув, як і у випадку з критерієм знаків G , називають типовим, а протилежний, менший за сумою рангів, – нетиповим. Як і для критерію знаків, ці два зсуви виявляються додатковими один до одного. T -критерій Вілкоксона базується на величині нетипового зсуву – $T_{emп}$.

Приклад: *Чи сприяла корекційна робота офіцера-психолога зниженню реактивної тривожності у досліджуваних військовослужбовців? Результати, отримані до та після експериментального впливу, наведено в таблиці 5.4.*

Таблиця 5.4

Результати діагностики до та після експериментального впливу

n	Показники реактивної тривожності за методикою Спілбергера	
	До	Після
1	69	51
2	73	76
3	56	45
4	63	51
5	71	63
6	69	42
7	69	57
8	71	63
9	70	61
10	71	60
11	67	68
12	54	49

Розв'язання. Сформулюємо статистичні гіпотези:

H_0 – зсув показників після корекційної роботи випадковий;

H_1 – зсув показників після корекційної роботи не випадковий.

Побудуємо додаткові стовпці, необхідні для подальшої роботи за T -критерієм Вілкоксона (таблиця 5.5).

Таблиця 5.5

Показники реактивної тривожності за методикою Спілбергера до та після експериментального впливу

n	Показники реактивної тривожності		Зсув	Абсолютний зсув	Ранг абсолютного зсуву
	До	Після			
1	69	51	-18	18	11
2	73	76	+3	3	2
3	56	45	-11	11	7,5
4	63	51	-12	12	9,5
5	71	63	-8	8	4
6	69	42	-27	27	12
7	69	57	-12	12	9,5
8	71	63	-9	9	5,5
9	70	61	-9	9	5,5
10	71	60	-11	11	7,5
11	67	68	+1	1	1
12	54	49	-5	5	3

Значення у стовпці “Зсув” отримуємо відніманням чисел стовпця “до” від чисел стовпця “після”. У стовпець “Абсолютний зсув” переносимо числа зі стовпця “Зсув” без знаків. У стовпці “Ранг абсолютного зсуву” найменшому значенню надаємо ранг 1 (у нашому випадку це 1). Наступному за величиною абсолютному зсуву зі значенням 3 надаємо ранг 2, зсуву зі значенням 5 – ранг 3 і т. д. Якщо виявлено однакові абсолютні зсуви, то приписувані ранги усереднюють між собою. Наприклад, два абсолютних зсуви зі значеннями 11 і 11 мають ранги 7 і 8. Усереднюємо ранги і приписуємо вказаним зсувам ранг 7,5.

T_{emn} дорівнює сумі рангів нетипових зсувів. У нашому випадку нетипових зсувів два: +3 та +1. Їх ранги дорівнюють 2 і 1 відповідно. Отже, $T_{emn} = 2+1=3$.

Статистичну достовірність зсуву за критерієм оцінюють за таблицею його критичних значень (додаток 1 таблиця 6). Критичні величини у таблиці шукають за кількістю досліджуваних. У нашому прикладі $n = 12$, тому частина таблиці має такий вигляд:

n	p	
	0,05	0,01
12	17	9

Побудуємо “вісь значущості”, де розташуємо критичні значення $T_{0,05} = 17$, $T_{0,01} = 9$ і емпіричне значення $T_{emn} = 3$.



NB! Зазначимо, що напрям “осі значущості” для T -критерію Вілкоксона, як і для критерію знаків G , має положення нуля праворуч, на відміну від традиційного – зліва, і числовий ряд збільшується у протилежний бік.

Значення критерію T_{emn} потрапило у зону значущості. Таким чином, гіпотезу H_0 відхиляємо і приймаємо гіпотезу H_1 , тому що зсув показників після корекційної роботи не випадковий. Отриманий за результатами експерименту зсув показників є статистично значущим на рівні $p < 0,01$. Отже,

корекційна робота офіцера-психолога статистично достовірно сприяла зниженню реактивної тривожності у досліджуваних військовослужбовців.

Обчислення T-критерію Вілкоксона за допомогою SPSS

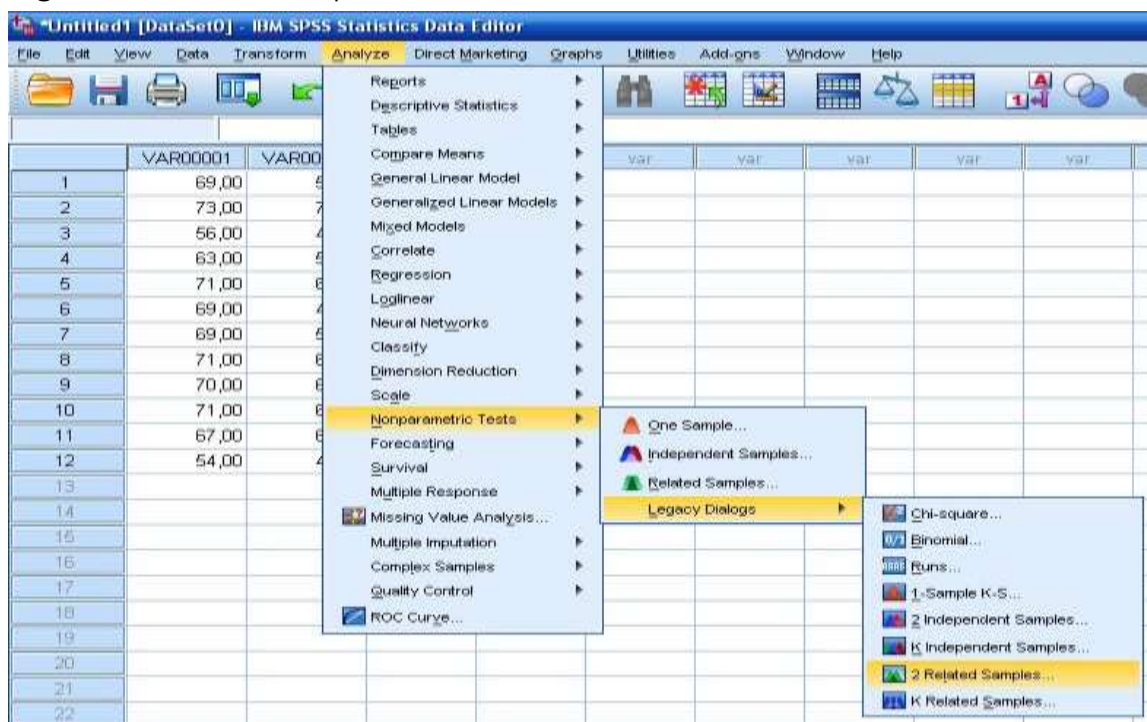
Розв'яжемо попереднє завдання за допомогою комп'ютерної програми SPSS.

1. Вводимо дані в таблицю у два стовпці: var1 та var2.

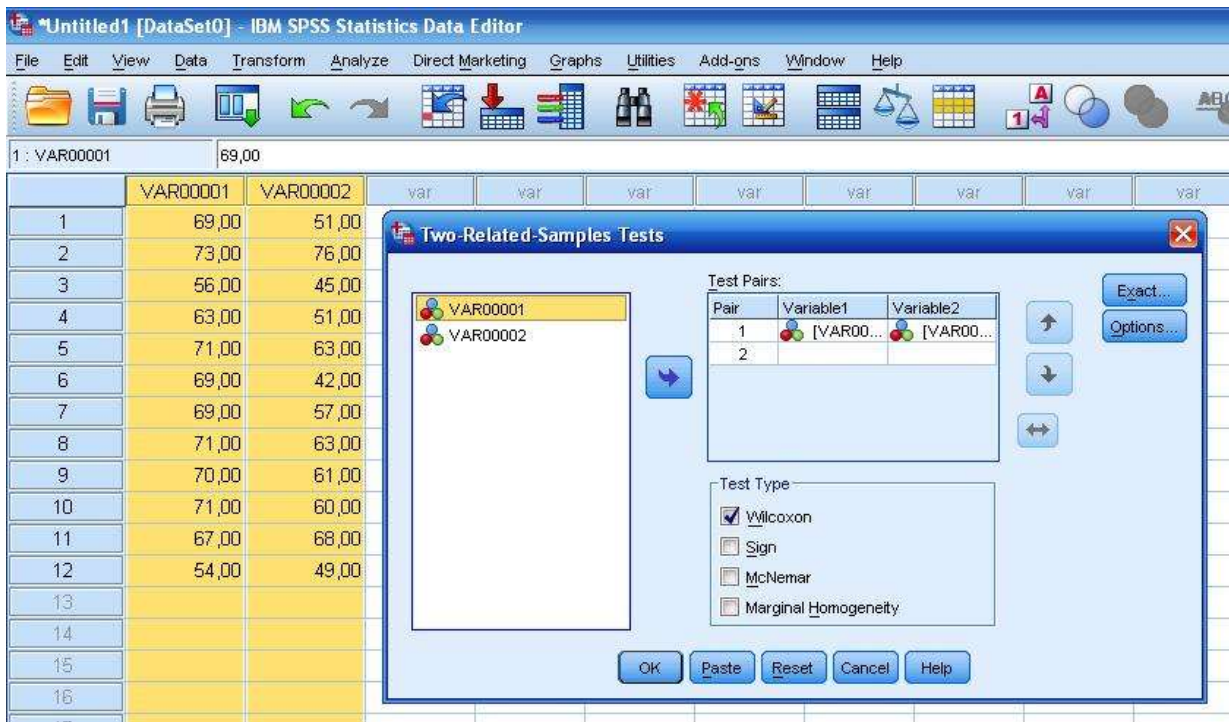


	VAR00001	VAR00002
1	69,00	51,00
2	73,00	76,00
3	56,00	45,00
4	63,00	51,00
5	71,00	63,00
6	69,00	42,00
7	69,00	57,00
8	71,00	63,00
9	70,00	61,00
10	71,00	60,00
11	67,00	68,00
12	54,00	49,00

2. У верхньому меню обираємо Analyze → Nonparametric Tests → Legacy Dialogs → 2 Related Samples



3. У меню, яке з'явилося, виділяємо дві змінні var1 і var2 і за допомогою кнопки ► переносимо їх у праве вікно. У цьому вікні у розділі Test Type ставимо прапорець на Wilcoxon.



4. Натискаємо ОК , отримуємо такий результат:

Wilcoxon Signed Ranks Test

		N	Mean Rank	Sum of Ranks
VAR00002 - VAR00001	Negative Ranks	10 ^a	7,50	75,00
	Positive Ranks	2 ^b	1,50	3,00
	Ties	0 ^c		
	Total	12		

a. VAR00002 < VAR00001

b. VAR00002 > VAR00001

c. VAR00002 = VAR00001

		VAR00002 - VAR00001
Z		-2,827 ^b
Asymp. Sig. (2-tailed)		,005

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Таблиця Ranks містить дані про кількість негативних (Negative Ranks), позитивних (Positive Ranks) та нульових (Ties) рангів.

Таблиця Test Statistics^a містить результат перевірки гіпотези: емпіричне значення критерію (Z) та p-рівень значущості (Asymp. Sig. (2-tailed)). У нашому випадку рівень значущості дорівнює 0,005. Через те, що 0,005 < 0,01, приймаємо гіпотезу H_1 , зсув показників є статистично значущим.

Алгоритм розрахунку T -критерію Вілкоксона:

1. Обчислити різницю між показниками “після” і “до” експериментального впливу. Визначити, що вважатиметься “типовим” зсувом і сформулювати відповідні гіпотези.

2. Перевести різниці в абсолютні величини та записати їх в окремий стовпець.

3. Ранжувати абсолютні величини різниць, надаючи меншому значенню менший ранг. Перевірити збіг отриманої суми рангів із розрахунковою.

4. Позначити ранги, які відповідають зсувам у “нетиповому” напрямі. Обчислити $T_{емп}$ – суму рангів нетипових зсувів.

5. Визначити критичні значення T -критерію Вілкоксона для кількості досліджуваних n за таблицею.

6. Зіставити $T_{емп}$ з $T_{кр}$ для $p = 0,05$ та $p = 0,01$. Якщо $T_{емп}$ менше за $T_{кр}$ або принаймні дорівнює йому, типовий зсув за інтенсивністю можна вважати достовірним.

У застосуванні T -критерію Вілкоксона необхідно дотримуватися таких умов:

1. Вимірювання проводять за всіма шкалами, крім номінальної.
2. Вибірki мають бути залежними.
3. Кількість елементів у вибірках, які порівнюють, має бути однаковою.
4. T -критерій Вілкоксона можна застосовувати, якщо чисельність вибірки становить від 5 до 50 (на більші величини таблиця критичних значень не розрахована).

Питання для самоконтролю

1. Охарактеризуйте критерій знаків G . Для чого він призначений і що дає змогу встановити?
2. Які зсуви вважають типовими у разі застосування критерію знаків G ? Які зсуви є нетиповими?
3. Який алгоритм розрахунку критерію знаків G ?

4. Перерахуйте основні умови, яких необхідно дотримуватись у застосуванні критерію знаків G .

5. Охарактеризуйте T -критерій Вілкоксона. Для чого він призначений і що дає змогу встановити?

6. Який алгоритм розрахунку T -критерію Вілкоксона?

7. Перерахуйте основні умови, яких необхідно дотримуватись у застосуванні T -критерію Вілкоксона.

Завдання для самостійної роботи

1. Офіцер-психолог провів груповий тренінг з військовослужбовцями підрозділу для зниження рівня тривожності. З'ясуйте, наскільки проведений тренінг є ефективним. Результати вимірювання рівня тривожності в учасників тренінгу наведено в таблиці 5.6.

Таблиця 5.6

Показники діагностики рівня тривожності за опитувальником Тейлора

n	Показники досліджуваної ознаки	
	До участі у тренінгу	Після участі у тренінгу
1	29	33
2	38	38
3	34	25
4	34	33
5	40	34
6	25	30
7	22	25
8	25	26
9	30	31
10	20	21
11	15	14
12	30	23
13	31	33
14	30	33

2. Офіцер-психолог провів груповий гештальт-тренінг з військовослужбовцями підрозділу, які брали участь у бойових діях. Визначити, чи підвищився рівень довіри та доброзичливості в учасників тренінгу? Результати вимірювання наведені у таблиці 5.7.

Таблиця 5.7

Показники діагностики доброзичливості за шкалою Д. Кемпбелла

n	Показники досліджуваної ознаки	
	До участі у тренінгу	Після участі у тренінгу
1	4	4
2	6	7
3	5	5
4	4	5
5	3	6
6	5	6
7	3	4
8	4	2
9	3	5
10	6	7

5.3. Оцінювання достовірності відмінностей

У психології та педагогіці часто доводиться проводити процедуру виявлення відмінностей між двома, трьома і більше вибірками досліджуваних осіб. Це може бути, наприклад, завдання з визначення психологічних особливостей військовослужбовців які брали участь у бойових діях, порівняно з тими, які не брали, або відмінностей між військовослужбовцями різних родів військ за певною ознакою тощо. Залежно від завдання дослідження однаково важливо буває довести нульову гіпотезу (наприклад, у ході формування контрольної та експериментальної груп на початку експерименту) або альтернативну (наприклад, коли потрібно порівняти незалежні вибірки між собою, щоб виявити певні відмінності).

Іноді за виявленими у дослідженні статистично достовірними відмінностями формують "груповий профіль" або "психологічний портрет" людини певної професії, статусу, соматичного захворювання та ін.

Останнім часом дедалі частіше постає завдання з виявлення психологічного портрета спеціалістів різних військових професій: "успішного оператора РЛС", "успішного розвідника" та ін. Такі дослідження не завжди передбачають застосування двох чи більше

вибірок. Іноді обстежують одну, але досить представницьку вибірку чисельністю не менш ніж 60 осіб, потім всередині цієї вибірки виокремлюють групи більш і менш успішних фахівців, їх дані за дослідженими змінними зіставляються між собою. У найпростішому випадку критерієм для поділу вибірки на "успішних" та "неуспішних" буде середня величина за показником успішності. Однак такий поділ є досить грубим: особи, які отримали близькі оцінки успішності, можуть опинитися в протилежних групах, а особи, які помітно відрізняються за оцінками успішності, – в одній і тій же групі. Це може спотворити результати зіставлення груп чи, у крайньому випадку, зробити різницю між групами менш помітними.

Щоб уникнути цього, можна спробувати виділити групи "успішних" і "неуспішних" фахівців суворіше, включаючи в першу з них лише тих, чії значення перевищують середню величину не менше ніж на $1/4$ стандартного відхилення, а в другу групу – лише тих, чії значення не менше ніж на $1/4$ стандартного відхилення нижче за середню величину. При цьому всі, хто опиняється в зоні середніх величин, $M \pm 1/4\sigma$ виключаються з подальших зіставлень. Якщо розподіл близький до нормального, випадіє приблизно 19,8% піддослідних. Якщо розподіл відрізняється від нормального, таких випробуваних може бути і більше. Щоб уникнути втрат, можна зіставляти не дві, а три групи досліджуваних: з високою, середньою та низькою професійною успішністю.

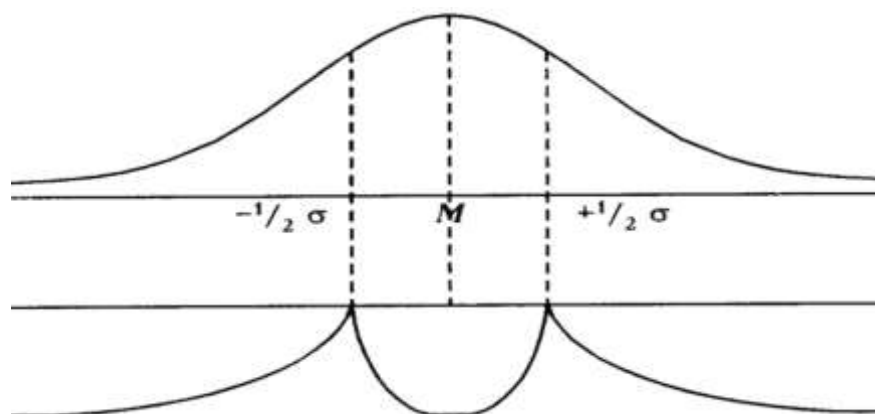


Рис. 5.1. Схематичне зображення процесу поділу вибірки на групи з низькою, середньою та високою професійною успішністю.

На рисунку 5.1 наведено схему поділу вибірки на групи з низькою, середньою та високою професійною успішністю за критерієм відхилення значень від середньої величини на $\frac{1}{2}$ від стандартного відхилення. У разі такого суворого розподілу в "середню" групу потрапляють (при нормальному розподілі) близько 38,2% всіх досліджуваних осіб, а в крайніх групах виявляється по 30,9% досліджуваних.

Чим менше випробовуваних виявляється у групах, тим менше в нас можливостей виявлення достовірних відмінностей, оскільки критичні значення більшості критеріїв за малих n суворіше, ніж за великих n . Таким чином, у разі несуворого поділу досліджуваних на групи ми втрачаємо у точності, а в разі суворого – у кількості досліджуваних осіб.

Зіставлення рівневих показників у різних вибірках може бути необхідною частиною комплексних діагностичних, навчальних, психокорекційних та інших програм. Воно допомагає звернути увагу на ті особливості обстежених вибірок, які мають бути враховані та використані при адаптації програм до цієї групи у процесі їхнього конкретного втілення.

З цією метою застосовується низка статистичних критеріїв. Рішення про вибір того чи іншого критерію приймається на основі того, скільки вибірок зіставляється та який їх обсяг.

Q-критерій Розенбаума

Призначення. Критерій використовується для оцінки відмінностей між двома вибірками за рівнем якоїсь ознаки, кількісно виміряної. У кожній із вибірок має бути не менше 11 досліджуваних.

Це дуже простий непараметричний критерій, який дозволяє швидко оцінити різницю між двома вибірками за якоюсь ознакою. Однак якщо критерій Q не виявляє достовірних відмінностей, це ще не означає, що їх насправді немає.

У цьому випадку варто застосувати критерій φ^* -Фішера. Якщо ж Q -критерій виявляє достовірну різницю між вибірками з рівнем значимості

$p \leq 0,01$, можна обмежитися лише ним і уникнути труднощів застосування інших критеріїв.

Критерій застосовується у тих випадках, коли дані представлені принаймні у порядковій шкалі. Ознака повинна варіювати у якомусь діапазоні значень, інакше зіставлення з допомогою Q -критерію просто неможливі. Наприклад, якщо ми маємо лише три значення ознаки, 1, 2 і 3, – нам дуже важко буде встановити відмінності. Метод Розенбаума потребує досить тонко виміряних ознак.

Застосування критерію починається з упорядкування значень ознаки в обох вибірках щодо наростання (або спадання) ознаки. Для того щоб не заплутатися, в цьому і в багатьох інших умовах рекомендується першим рядом (вибіркою, групою) вважати той ряд, де значення вище, а другим рядом – той, де значення нижче.

Гіпотези.

H_0 : Рівень ознаки у вибірці 1 не перевищує рівня ознаки у вибірці 2.

H_1 : Рівень ознаки у вибірці 1 перевищує рівень ознаки у вибірці 2.

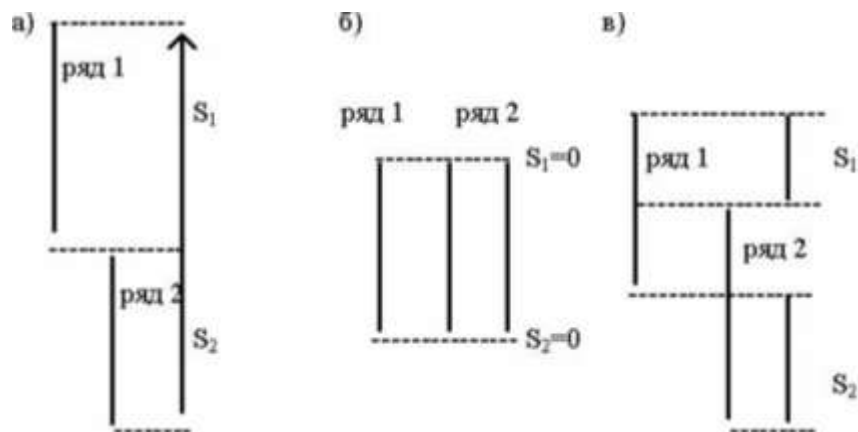


Рис. 5.2. Графічне уявлення критерію Q . Можливі співвідношення рядів значень у двох вибірках; S_1 – зона значень першого ряду, які вищі за максимальне значення 2-го ряду; S_2 – зона значень другого ряду, які менші за мінімальне значення 1-го ряду.

На рисунку 5.2. представлені три варіанти співвідношення рядів значень у двох вибірках. У варіанті (а) всі значення першого ряду вищі за всі значення другого ряду. Відмінності, безумовно, достовірні, за умови

дотримання умови, що $n_1, n_2 \geq 11$. У варіанті (б), навпаки, обидва ряди знаходяться на тому самому рівні: відмінності недостовірні. У варіанті (в) ряди частково перехрещуються, але все ж таки перший ряд виявляється набагато вище другого. Чи достатньо великі зони S_1 і S_2 у сумі складові Q , можна визначити за таблицею 7 додатка 1, де наведені критичні значення Q для різних n . Чим величина Q більше, тим більш достовірні відмінності ми зможемо констатувати.

Алгоритм розрахунку Q -критерію Розенбаума

1. Перевірити, чи виконуються обмеження: $n_1, n_2 \geq 11, n_1 \approx n_2$.
2. Упорядкувати значення окремо у кожній вибірці за ступенем зростання ознаки. Вважати вибіркою 1 ту вибірку, значення у якій імовірно вище, а вибіркою 2 таку, де значення імовірно нижче.
3. Визначити найвище (максимальне) значення вибірці 2.
4. Підрахувати кількість значень у вибірці 1, які вищі за максимальне значення у вибірці 2. Позначити отриману величину як S_1 .
5. Визначити найнижче (мінімальне) значення у вибірці 1.
6. Підрахувати кількість значень у вибірці 2, які нижчі від мінімального значення вибірки. Позначити отриману величину як S_2 .
7. Підрахувати емпіричне значення Q за такою формулою:
$$Q_{емп} = S_1 + S_2.$$
8. За таблицею критичних значень визначити критичні значення $Q_{кр}$ даних n_1 і n_2 . Якщо $Q_{емп}$ дорівнює $Q_{кр}$ на рівні 0,05 або перевищує його, H_0 відкидається.
9. При $n_1, n_2 > 26$ зіставити отримане емпіричне значення $Q_{кр} = 8$ ($p \leq 0,05$) і $Q_{кр} = 10$ ($p \leq 0,01$). Якщо $Q_{емп}$ перевищує чи принаймні дорівнює $Q_{кр} = 8$, H_0 відкидається.

Для застосування Q -критерію Розенбаума необхідно дотримуватись наступних умов:

1. У кожній із незалежних вибірок має бути не менше 11 спостережень. При цьому обсяги вибірок мають бути приблизно однакові.
2. Якщо в обох вибірках менше 50 спостережень, абсолютна величина різниці між n_1 і n_2 не має бути більше 10 спостережень.

3. Якщо в кожній з вибірок більше 51 спостереження, але менше 100, то абсолютна величина різниці між n_1 і n_2 не повинна бути більшою за 20 спостережень.

4. Якщо в кожній із вибірок більше 100 спостережень, то допускається, щоб одна з вибірок була більшою за іншу не більше ніж у 1,5-2 рази.

5. Діапазони розкиду значень у двох вибірках повинні не збігатися між собою, інакше застосування критерію є безглуздим. Тим часом, можливі випадки, коли діапазони розкиду значень збігаються, але, внаслідок різнобічної асиметрії двох розподілів, відмінності середніх величин ознак істотні (рисунки 5.3. та 5.4).

6. Якщо одна вибірка має два «хвости», використання критерію заборонено.

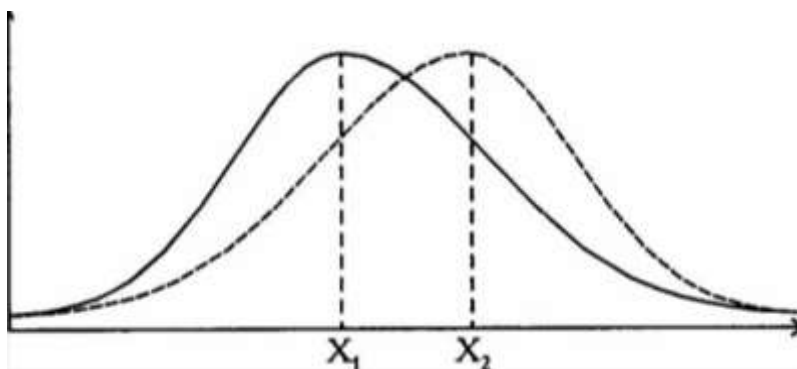


Рис. 5.3. Варіант співвідношення розподілів ознаки у двох вибірках, у якому Q-критерій безпорадний.

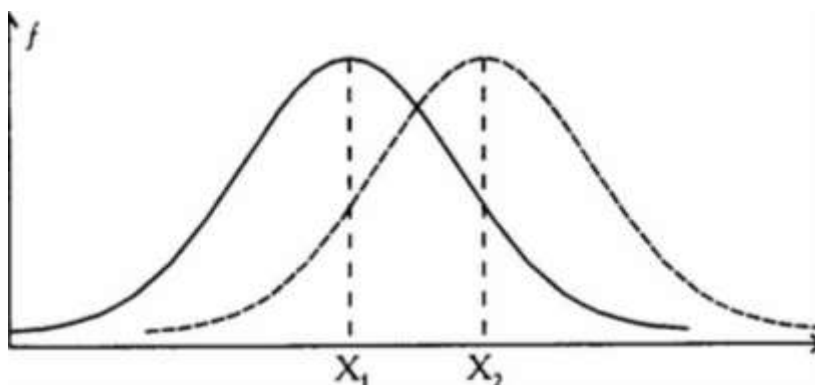


Рис. 5.4. Варіант співвідношення розподілів ознаки у двох вибірках, у якому Q-критерій може бути могутнім.

Приклад: На етапі вступу до ВВНЗ двох групах абітурієнтів, було виміряно рівень вербального інтелекту за допомогою методики Д. Векслера. Було обстежено 26 юнаків віком від 18 до 24 років (середній вік 20,5 років). 14 із них були військовослужбовцями служби за контрактом, а 12 – такими, які закінчили середню школу. Показники вербального інтелекту представлені у таблиці 5.5. Чи можна стверджувати, що одна з груп перевершує іншу за рівнем вербального інтелекту?

Таблиця 5.5.

Індивідуальні значення вербального інтелекту у вибірках військовослужбовців служби за контрактом ($n_1 = 14$) та випускників шкіл ($n_2 = 12$) факультетів

№	Військовослужбовці служби за контрактом ($n_1 = 14$)	Випускники шкіл ($n_2 = 12$)
1	132	126
2	134	127
3	124	132
4	132	120
5	135	119
6	132	126
7	131	120
8	132	123
9	121	120
10	127	116
11	136	123
12	129	115
13	136	-
14	136	-

Упорядкуємо значення обох вибірок, та був сформулюємо гіпотези:
 H_0 : Військовослужбовці служби за контрактом не перевершують випускників шкіл за рівнем вербального інтелекту.

H_1 : Військовослужбовці служби за контрактом перевершують випускників шкіл за рівнем вербального інтелекту.

Таблиця 5.6.

**Впорядковані за спаданням вербального інтелекту ряди
індивідуальних значень у двох вибірках**

Військовослужбовці служби за контрактом ($n_1 = 14$)		Випускники шкіл ($n_2 = 12$)	
1	136		
2	136		
3	136		
4	135		
5	134		
6	132	1	132
7	132		
8	132		
9	132		
10	131		
11	129		
12	127	2	127
		3	126
		4	126
13	124		
		5	123
		6	123
14	121		
		7	120
		8	120
		9	120
		10	119
		11	116
		12	115

Як видно з таблиці 5.6. ми позначили ряди: перший, той, що "вище" – ряд військовослужбовців, а другий, той, що "нижче" – ряд випускників шкіл.

За таблицею 5.6. визначаємо кількість значень першого ряду, які більші за максимальне значення другого ряду: $S_1 = 5$.

Тепер визначаємо кількість значень другого ряду, які менші за мінімальне значення першого ряду: $S_2 = 6$.

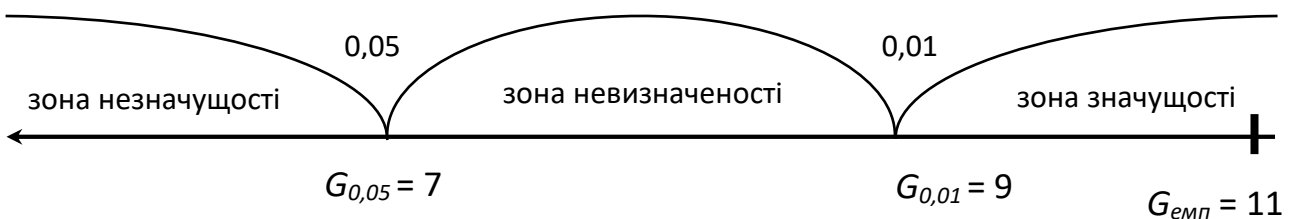
Обчислюємо $Q_{емп}$ за формулою:

$$Q_{емп} = S_1 + S_2 = 5 + 6 = 11$$

За таблицею 7 додатка 1 визначаємо критичні значення $Q_{кр}$ для $n_1 = 14, n_2 = 12$:

$$Q_{кр} = 7 (p \leq 0,05) \text{ та } Q_{кр} = 9 (p \leq 0,01)$$

Чим більша розбіжність між вибірками, то більше вписувалося величина Q . Побудуємо "вісь значимості".



$$Q_{емп} > Q_{кр} (p \leq 0,01)$$

Відповідь: H_0 відхиляється. Приймається H_1 . Військовослужбовці служби за контрактом перевершують випускників шкіл за рівнем вербального інтелекту ($p < 0,01$).

NB! У тих випадках, коли емпірична величина критерію виявляється на межі зони незначущості, ми маємо право стверджувати лише, що відмінності достовірні при $p \leq 0,05$, якщо воно виявляється між двома критичними значеннями, то ми можемо стверджувати, що $p < 0,05$.

U-критерій Манна-Уїтні

Призначення. Критерій використовують для оцінки розбіжностей між двома незалежними вибірками за рівнем деякої ознаки, яка кількісно вимірюється. Він дозволяє виявити відмінності між двома малими вибірками, коли $n_1, n_2 \geq 3$, або $n_1 = 2, n_2 \geq 5$. Даний критерій є більш потужним, ніж критерій Розенбаума. Ідея цього методу полягає у визначенні зони перехресних значень між двома рядами.

Перший ряд – це той ряд значень, у якому значення, за попередньою оцінкою, вищі, а другий ряд – той, де значення за припущенням нижчі.

Чим меншою буде зона перехресних значень, тим вища ймовірність, що відмінності достовірні. Інколи ці розбіжності називають розбіжностями у розташуванні двох вибірок. Емпіричне значення критерію U_{emn} відображає те, наскільки велика зона співпадіння між рядами. Тому чим менше U_{emn} , тим більш ймовірно, що відмінності достовірні.

Формулювання гіпотези

H_0 : Рівень ознаки у групі 2 не нижчий рівня ознаки у групі 1.

H_1 : Рівень ознаки у групі 2 нижчий рівня ознаки у групі 1.

Графічне представлення критерію. На рисунку представлені три можливих варіанти співвідношення двох рядів значень.

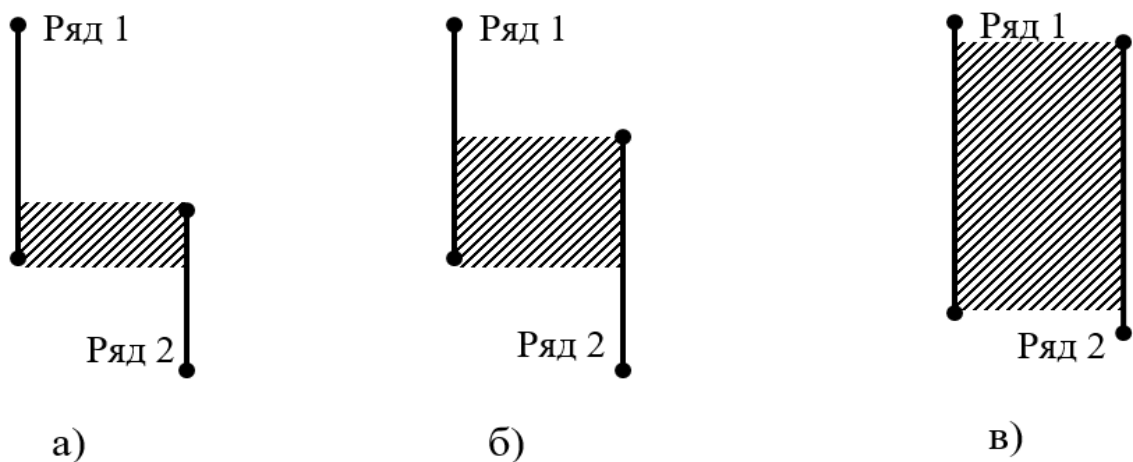


Рис. 5.5. Графічне представлення U -критерію Манна-Уїтні.

Варіант А: другий ряд нижче першого, і ряди майже не накладаються. Область накладання занадто мала, щоб приховати відмінності між рядами, тому є шанс, що розбіжності між рядами достовірні. Точно визначити це можна за допомогою критерію U .

Варіант Б: другий ряд теж нижче першого, але і область накладання значень у двох рядів достатньо велика. Вона може ще не досягати критичної величини, коли розбіжності необхідно визнати неістотними. Для уточнення можна використати критерій U .

Варіант В: другий ряд нижче першого, але область накладання настільки велика, що відмінності між рядами зникають.

Правила ранжування:

1. Впорядкувати дані експериментального дослідження за спаданням (зростанням) ознаки спочатку для кожної із вибірок, а потім спільно для обох вибірок.

2. Провести ранжування даних ознаки обох вибірок, тобто присвоїти кожному значенню ознаки ранг. Меншому значенню нараховується менший ранг. Найменшому значенню нараховується ранг 1. Найбільшому значенню нараховується ранг, що дорівнює кількості ранжованих значень.

3. У випадку, якщо кілька значень рівні, їм нараховується ранг, що є середнім значенням із тих рангів які вони би дістали, якби не були рівними.

4. Загальна сума рангів має співпадати з розрахунковою, яка обчислюється за формулою. Неспівпадіння реальної і розрахункової сум рангів свідчить про допущену помилку, яку треба знайти і виправити.

Для застосування U -критерію Манна-Уїтні необхідно дотримуватись наступних умов:

1. У кожній вибірці повинно бути не менше 3 спостережень: $n_1, n_2 \geq 3$; припускається, щоб в одній виборці було 2 спостереження, але тоді у другій їй повинно бути не менше 5 ($n_1 = 2, n_2 \geq 5$).

2. У кожній вибірці повинно бути не більше 60 спостережень. Якщо $n_1, n_2 > 20$, краще обрати інший критерій (наприклад, кутове перетворення Фішера), оскільки ранжування буде досить трудомістким.

Оцінювання результатів:

а) Якщо результат $U_{емп} > U_{кр}$, для рівня значущості $p = 0,05$ та обсягів вибірок n_1 і n_2 , то приймаємо нульову гіпотезу: рівень ознаки у першій вибірці не перевищує рівень ознаки у другій;

б) Якщо результат $U_{емп} < U_{кр}$, для рівня значущості $p = 0,01$ та обсягів вибірок n_1 і n_2 , то приймаємо альтернативну гіпотезу: рівень ознаки у першій вибірці перевищує рівень ознаки у другій.

Приклад: У двох групах слухачів з $n_1 = 8$, $n_2 = 10$ проводився диференційований залік з максимально можливим числом балів 100. Фактично набрані слухачами бали представлені 1 таблиці 5.7. Чи можна

стверджувати, що рівні засвоєння навчального предмету у цих групах різняться?

Таблиці 5.7

Навчальна група 1

№	В/звання, прізвище, ініціали	Бали
1	майор Авраменко М.В.	75
2	майор Бігун В.О.	87
3	майор Когут Т.Ю.	95
4	капітан Полюляк А.Р.	85
5	підполковник Мисько О.М.	88
6	майор Повз В.О.	85
7	майор Шворак В.Я.	89
8	капітан Яворницький А.Д.	82

Навчальна група 2

№	Прізвище	Бали
1	майор Бандера С.А.	95
2	капітан Веремей А.П.	84
3	майор Ласий М.В.	72
4	майор Найдюк В.О.	70
5	майор Опря Т.Ю.	65
6	капітан Полюляк А.Р.	68
7	підполковник Ровз Р.Д.	68
8	майор Слука В.О.	72
9	майор Тичний В.Я.	45
10	капітан Ярош А.Д.	56

H_0 : Відмінність у рівні засвоєння предмету між обома групами відсутня.

H_1 : Слухачі групи 1 мають вищий рівень засвоєння навчального предмету.

Алгоритм використання U -критерію Манна-Уїтні.

1. Перенести дані вибірок на індивідуальні картки.
2. Помітити картки вибірки 1 одним кольором (червоним), а всі картки з вибірки 2 – іншим (синім).
3. Розкласти всі картки в єдиний ряд за зростанням ознаки, не звертаючи уваги на те, до якої вибірки вони відносяться, так, якщо б ми працювали з однією великою вибіркою.
4. Проранжувати значення на картках, приписуючи меншому значенню менший ранг. Всього рангів має вийти $n_1 + n_2 = 8 + 10 = 18$. В нашій задачі деякі значення повторюються, тому цим значенням присвоюють ранг, що є середнім значенням із тих рангів які вони би дістали, якби не були рівними. Наприклад, значення 85 зустрічається двічі:

на 12 та 13 місцях: $(12+13):2=12,5$, тому обом значенням присвоїмо ранг 12,5.

5. Знову розкласти картки на дві групи за кольором.

6. Полічити суму рангів окремо на червоних картках (група 1) (у нашому прикладі дорівнює 106,5) і на синіх картках (група 2) (у нашому прикладі дорівнює 64,5). Перевірити, чи співпадає загальна сума рангів із розрахунковою.

7. Визначити більшу з двох рангових сум (у нашому прикладі дорівнює 105,5 для першої групи).

8. Обчислити U -критерій за формулою:

$$U_{\text{емп}} = (n_1 * n_2) + \frac{n_x * (n_x + 1)}{2} - T_x$$

де n_1 – кількість досліджуваних у вибірці 1;

n_2 – кількість досліджуваних у вибірці 2;

n_x – кількість досліджуваних у вибірці з більшою сумою рангів;

T_x – більша із двох рангових сум;

У нашому прикладі $U_{\text{емп}} = 9,5$

9. Визначити $U_{\text{кр}}$ за таблицею 8 додатку 1 для рівнів статистичної значущості $p \leq 0,05$ та $0,01$ (для рівня $0,05$ $U_{\text{кр}} = 20$ та рівня $0,01$ $U_{\text{кр}} = 13$ відповідно)

10. Будуємо вісь значимості:

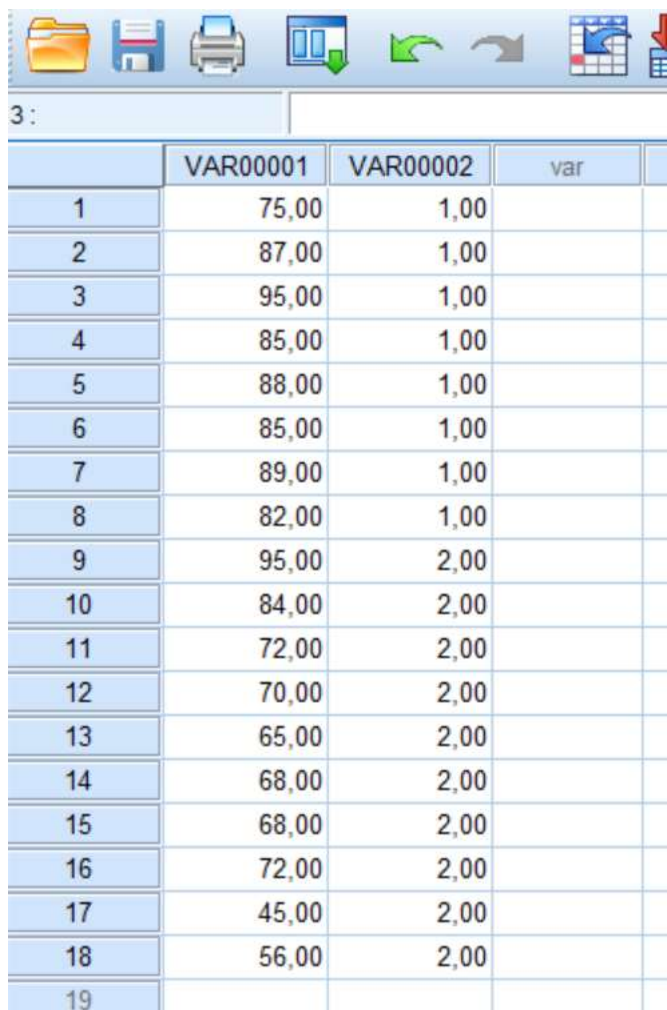


11. Висновок: H_0 відкидається, навчальна група 1 має вищий рівень засвоєння предмету.

Обчислення *U*-критерію Манна-Уїтні за допомогою SPSS.

Розв'яжемо попереднє завдання за допомогою комп'ютерної програми SPSS.

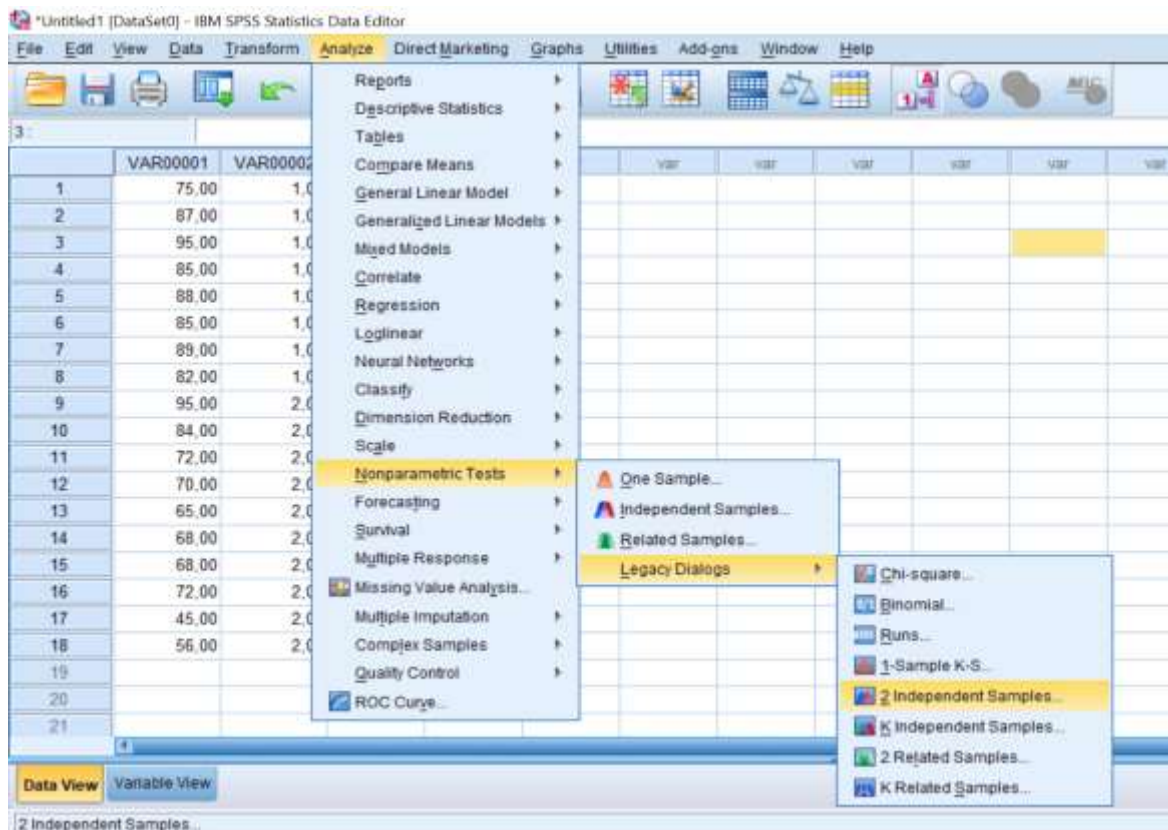
1. Вводимо дані в таблицю у два стовпці var1 та var2 наступним чином: у стовпчик var1 вводимо дані вибірки 1 а одразу під нею вибірки 2. У стовпчик var2 вводимо нумерацію вибірок, тобто напроти кожного значення першої вибірки вводимо значення 1, а напроти значень вибірки 2 цифру 2. Таким чином ми отримуємо у першому стовпчику 18 значень напроти кожного з яких буде зазначено до якої з двох вибірок воно належить.



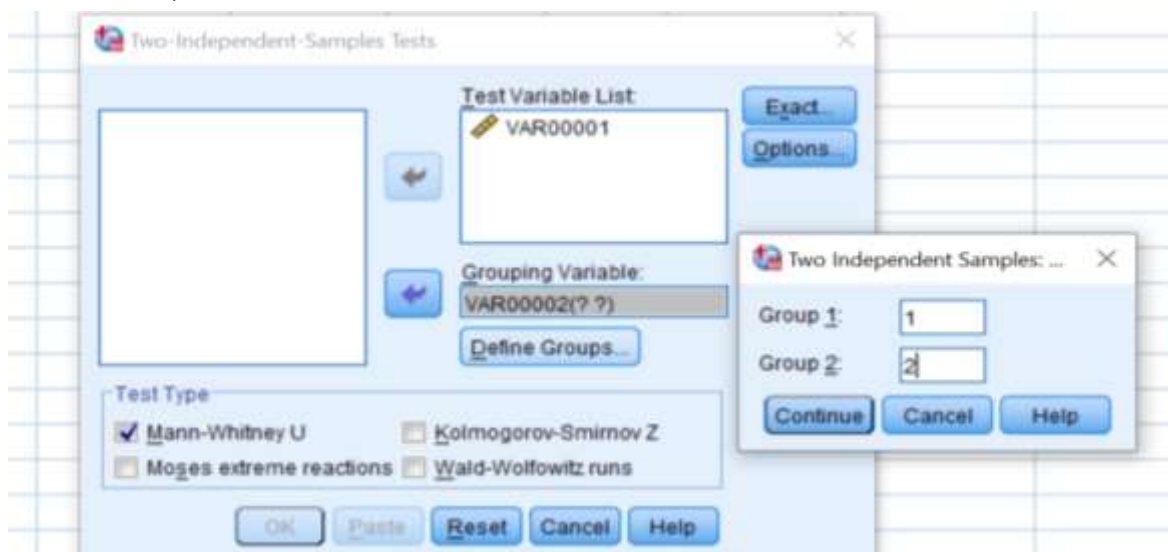
The screenshot shows a spreadsheet application window with a toolbar at the top containing icons for file operations (folder, save, print, copy, paste, delete) and a search icon. Below the toolbar, the spreadsheet is displayed with the following data:

	VAR00001	VAR00002	var
1	75,00	1,00	
2	87,00	1,00	
3	95,00	1,00	
4	85,00	1,00	
5	88,00	1,00	
6	85,00	1,00	
7	89,00	1,00	
8	82,00	1,00	
9	95,00	2,00	
10	84,00	2,00	
11	72,00	2,00	
12	70,00	2,00	
13	65,00	2,00	
14	68,00	2,00	
15	68,00	2,00	
16	72,00	2,00	
17	45,00	2,00	
18	56,00	2,00	
19			

2. У верхньому меню обираємо Analyze → Nonparametric Tests → Legacy Dialogs → 2 Independent Samples.



3. У меню, яке з'явилося, виділяємо дві змінні var1 і за допомогою кнопки ► переносимо її у праве верхнє вікно Test variable list. Змінну var2 переносимо у вікно, яке нижче Grouping Variable, натискаємо Define Groups та вносимо означення тих вибірок, які будуть оброблятися, тобто та 2. У цьому вікні у розділі Test Type ставимо прапорець на Mann-Whitney U (див. малюнок).



4. Натискаємо ОК та отримуємо наступний результат:

Mann-Whitney Test

Ranks				
	VAR00002	N	Mean Rank	Sum of Ranks
VAR00001	1,00	8	13,31	106,50
	2,00	10	6,45	64,50
	Total	18		

Test Statistics ^a	
	VAR00001
Mann-Whitney U	9,500
Wilcoxon W	64,500
Z	-2,716
Asymp. Sig. (2-tailed)	,007
Exact Sig. [2*(1-tailed Sig.)]	,004 ^b

a. Grouping Variable: VAR00002

b. Not corrected for ties.

Таблиця Ranks містить дані про середнє значення рангів (Mean Ranks), та суми рангів (Sum of Ranks).

Таблиця Test Statistics^a містить результат перевірки гіпотези: емпіричне значення критерію (Z) та р-рівень значущості (Asymp. Sig. (2-tailed)). У нашому випадку рівень значущості дорівнює 0,007. Так як $0,007 < 0,01$,

приймаємо гіпотезу H_1 , відмінності показників є статистично значущими.

H-критерій Крускала-Уолліса

Призначення. Критерій призначений для оцінки відмінностей одночасно між трьома, чотирма та більше вибірками за рівнем будь-якої ознаки. Він дозволяє встановити, що рівень ознаки змінюється під час переходу від групи до групи, але не вказує на напрямок цих відмінностей.

Цей критерій іноді розглядають як непараметричний аналог методу дисперсійного однофакторного аналізу для незалежних вибірок. Іноді його називають критерієм "суми рангів". Цей критерій є продовженням U-критерію Манна-Уїтні на більше чим дві вибірки. Усі індивідуальні значення ранжуються так, як ніби то це є одна велика вибірка. Потім усі індивідуальні значення повертаються у свої початкові вибірки і ми підраховуємо суми отриманих ними рангів окремо у кожній вибірці. Якщо відмінності між вибірками випадкові, суми рангів не будуть відрізнятися суттєво, тому що високі та низькі ранги розподіляться рівномірно між вибірками. Якщо ж в одній із вибірок будуть переважати низькі значення

рангів, в іншій високі, в третій середні, то критерій H дозволить встановити ці відмінності.

Приклад формулювання гіпотез:

H_0 : між вибірками 1, 2, 3 і т. д. існують лише випадкові відмінності за рівнем досліджуваної ознаки.

H_1 : між вибірками 1, 2, 3 і т. д. існують не випадкові відмінності за рівнем досліджуваної ознаки.

Графічно цей критерій представляє собою суму зон, які перехрещуються у вибірках, як відображено на рисунку 5.6:

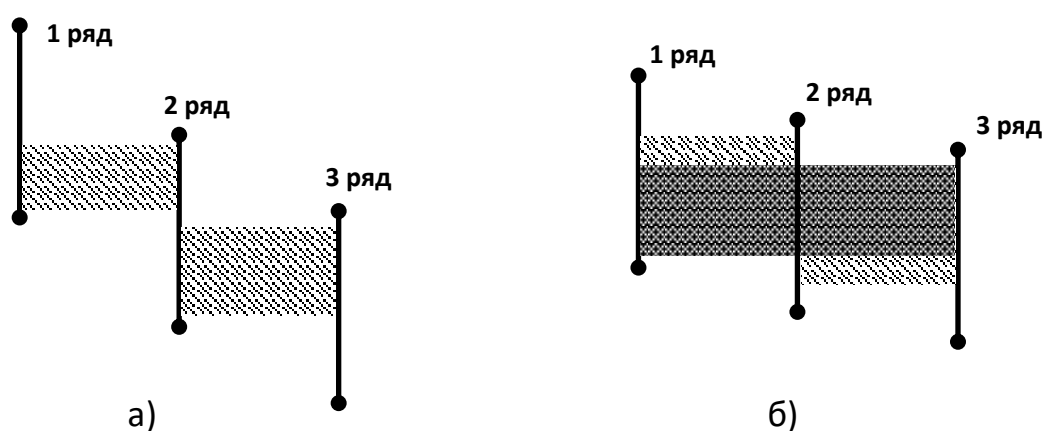


Рис. 5.6. Два можливих варіанти накладення співвідношень рядів значень у трьох вибірках.

Критерій H оцінює загальну суму зон, які перехрещуються, під час співставлення усіх вибірок. Якщо сумарна область накладення мала (Рис. 5.6 а), різниця достовірна; якщо вона досягає визначеної критичної величини та перевищує її (Рис. 5.6 б), то різниця між вибірками не є достовірною.

Алгоритм підрахунку H -критерію Крускала-Уолліса

1. Перенести всі показники досліджуваних на індивідуальні картки.
2. Позначити картки випробуваних групи 1 певним кольором, наприклад червоним, картки випробуваних групи 2 – синім, картки випробуваних груп 3 і 4 – відповідно, зеленим і жовтим кольором і т. ін. (можна використовувати будь-які інші позначення).

3. Розкласти всі картки в єдиний ряд за ступенем наростання ознаки, не зважаючи на те, до якої групи відносяться картки, тобто так як би ми працювали з однією об'єднаною вибіркою.

4. Проранжувати значення на картках, приписуючи меншому значенню менший ранг. Надписати на кожній картці її ранг. Загальна кількість рангів дорівнюватиме кількості досліджуваних в об'єднаній вибірці.

5. Знову розкласти картки по групах, орієнтуючись на кольорові, або інші прийняті позначення.

6. Підрахувати суми рангів окремо з кожної групи. Перевірити збіг загальної суми рангів із розрахунковою.

7. Підрахувати значення критерію H за такою формулою:

$$H = \left[\frac{12}{N(N+1)} * \sum \frac{T_i^2}{n} \right] - 3(N+1)$$

де N – загальна кількість піддослідних в об'єднаній вибірці;

n – кількість випробуваних у кожній вибірці;

T – суми рангів за кожною групою.

8. При кількості груп $c = 3$, $n_1, n_2, n_3 \leq 5$, визначити критичні значення і відповідний їм рівень значущості за таблицею 9 додатку 1. Якщо $H_{емп}$ дорівнює, або перевищує критичне значення $H_{0,05}$, H_0 відкидається.

9. При кількості груп $c > 3$ або кількості піддослідних $n_1, n_2, n_3 > 5$, визначити критичні значення χ^2 по таблиці 10 додатку 1. Для цього попередньо необхідно визначити кількість ступенів свободи u за формулою:

$$u = c - 1$$

Якщо $H_{емп}$ дорівнює або перевищує критичне значення χ^2 , H_0 відкидається.

Обмеження критерію:

1. При зіставленні 3-х вибірок допускається, щоб у одній з них $n = 3$, а двох інших $n = 2$. Однак відмінності можна буде встановити лише на нижчому рівні значущості ($p \leq 0,05$).

2. Для того, щоб виявилось можливим діагностувати відмінності на вищому рівні значущості ($p \leq 0,01$), необхідно, щоб в одній вибірці було 4 спостереження, а в двох інших не менше двох.

При великих вибірках $n_1, n_2, n_3 > 5$ необхідно користуватися таблицею критичних значень критерію χ^2 . Оскільки критерій Крускала-Уолліса асимптотично наближається до розподілу χ^2 . Великою величиною входу у таблицю є кількість ступенів свободи u , яка розраховується за формулою:

$$u = c - 1$$

Приклад: Чотири групи досліджуваних військовослужбовців виконували психофізіологічний тест "Кільця Ландольта" у різних експериментальних умовах. Завданням дослідження було встановити – чи залежить ефективність виконання тесту від зовнішніх умов. Тобто необхідно встановити чи існують статистично значущі відмінності у виконанні тесту між групами. До кожної групи входило чотири досліджуваних.

Рішення: Число помилок показника перемикання уваги (у відсотках) зазначено у таблиці 5.8.

Таблиця 5.8

№ дослідженого	1 група	2 група	3 група	4 група
1	23	45	34	21
2	20	12	24	22
3	34	34	25	26
4	35	11	40	27
Сума	112	102	123	96

Для подальшої роботи з критерієм нам необхідно вибудувати усі значення в один стовпець у порядку збільшення та присвоїти їм ранги.

NB! Нагадаємо правила ранжування. Кожному значенню присвоюється ранг відповідно до його порядкового номеру у шкалі за зростанням або убутанням. У випадку якщо два або більше значень мають однаковий показник, їм присвоюється однаковий ранг який розраховується як середньоарифметичне суми їх позицій. Наприклад: у розподілі ознаки 51 бал набрали два респонденти з позиціями 12 та 13

відповідно, тоді $(12+13)/2 = 12,5$. Тобто їм присвоюється однаковий ранг 12,5.

1. Побудуємо таблицю ранжування:

Дані	Ранги	Дані	Ранги
11	1	26	9
12	2	27	10
20	3	34	12
21	4	34	12
22	5	34	12
23	6	35	14
24	7	40	15
25	8	45	16
Загальна сума рангів: 136			

NB! Перевірити правильність підрахунку суми рангів можна за формулою:

$$\frac{N(N + 1)}{2}$$

де N – загальна кількість досліджуваних в об'єднаній вибірці;

Тобто, для нашого прикладу це $(16*17)/2 = 136$.

2. Розподілимо наші дані знов за чотирма групами, але вже із рангами:

№ дослідженого	1 група	Ранг	2 група	Ранг	3 група	Ранг	4 група	Ранг
1	23	6	45	16	34	12	21	4
2	20	3	12	2	24	7	22	5
3	34	12	34	12	25	8	26	9
4	35	14	11	11	40	15	27	10
Сума	112	35	102	31	123	42	96	31

3. Розрахуємо емпіричне значення критерію за формулою:

$$H_{emn} = \left[\frac{12}{N(N+1)} * \sum \frac{T_i^2}{n} \right] - 3(N + 1);$$

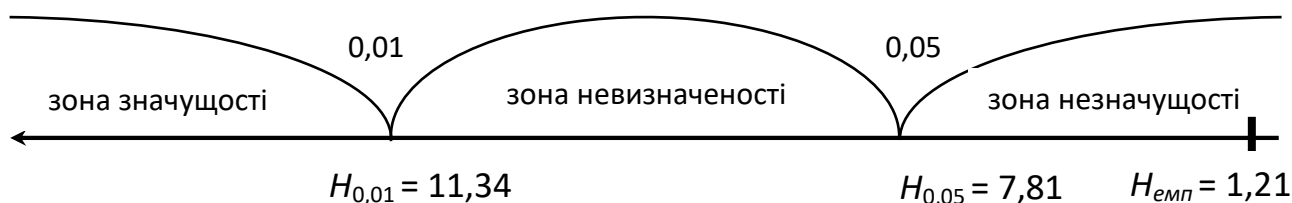
$$H_{emn} = \frac{12}{16*17} * \left[\frac{35*35}{4} + \frac{31*31}{4} + \frac{42*42}{4} + \frac{28*28}{4} \right] - 3*17 = 1,21$$

4. Під час визначення критичних значень H -критерія, якщо кількість вибірок більше трьох, користуйтеся таблицею критичних значень для χ^2 . Для цього попередньо розраховують кількість ступенів свободи u для $c = 4$. Тоді $u = c - 1 = 3$.

5. За таблицею критичних значень χ^2 знаходимо (дод. 1 таблиця 10):

$$H_{кр} = \begin{cases} 7,81 \text{ для } p \leq 0,05 \\ 11,34 \text{ для } p \leq 0,01 \end{cases}$$

6. Будуємо «вісь значущості»



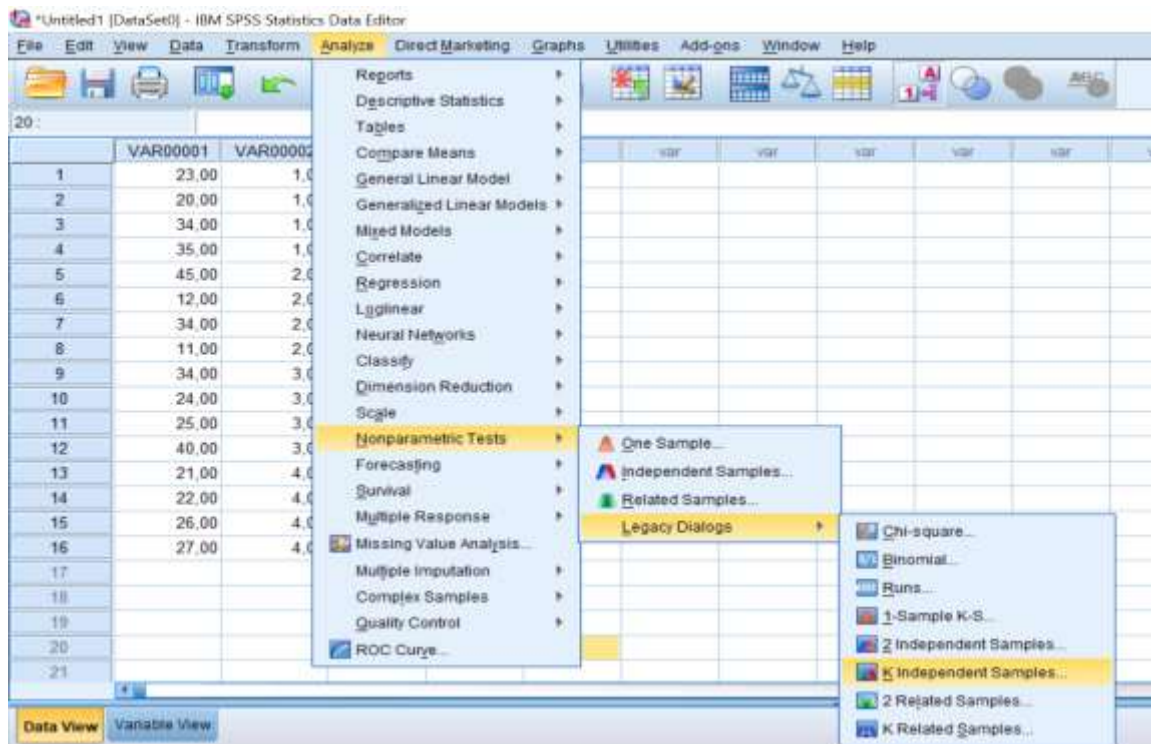
Обчислення H -критерію Крускала-Уоллеса за допомогою SPSS.

Розв'яжемо попереднє завдання за допомогою комп'ютерної програми SPSS.

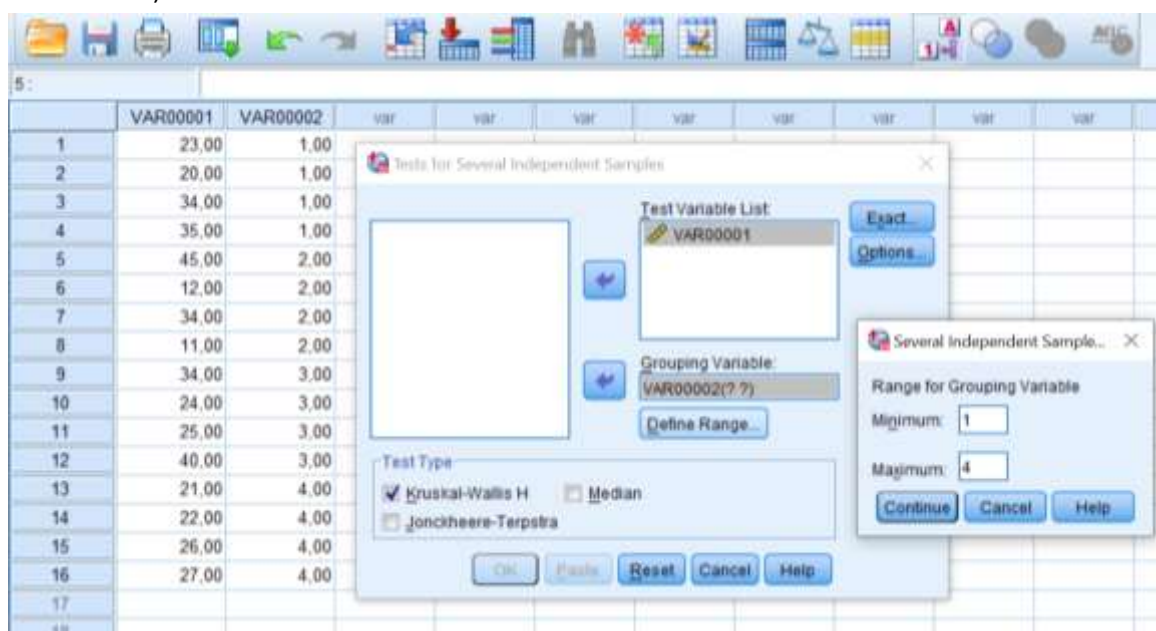
1. Вводимо дані в таблицю у два стовпці var1 та var2 наступним чином: у стовпчик var1 вводимо дані чотирьох вибірок одна під одною. У стовпчик var2 вводимо нумерацію вибірок, тобто напроти кожного значення першої вибірки вводимо значення 1, а напроти значень вибірки 2 цифру 2 і так далі. Таким чином ми отримуємо у першому стовпчику 16 значень напроти кожного з яких буде зазначено до якої з двох вибірок воно належить.

	VAR00001	VAR00002	var
1	23,00	1,00	
2	20,00	1,00	
3	34,00	1,00	
4	35,00	1,00	
5	45,00	2,00	
6	12,00	2,00	
7	34,00	2,00	
8	11,00	2,00	
9	34,00	3,00	
10	24,00	3,00	
11	25,00	3,00	
12	40,00	3,00	
13	21,00	4,00	
14	22,00	4,00	
15	26,00	4,00	
16	27,00	4,00	

5. У верхньому меню обираємо Analyze → Nonparametric Tests → Legacy Dialogs → K Independent Samples.



6. У меню, яке з'явилося, виділяємо дві змінні var1 і за допомогою кнопки ► переносимо її у праве верхнє вікно Test variable list. Змінну var2 переносимо у вікно, яке нижче Grouping Variable, натискаємо Define Groups та вносимо туди значення тих вибірок, які будуть оброблятися, тобто від 1 до 4. У цьому вікні у розділі Test Type ставимо прапорець на Kruskal-Wallis H (див. малюнок).



7. Натискаємо ОК та отримуємо наступний результат:

Kruskal-Wallis Test

	VAR00002	N	Mean Rank
VAR00001	1,00	4	8,75
	2,00	4	7,75
	3,00	4	10,50
	4,00	4	7,00
	Total	16	

	VAR00001
Chi-Square	1,220
df	3
Asymp. Sig.	,748

a. Kruskal Wallis Test

b. Grouping Variable: VAR00002

Таблиця Ranks містить дані про середнє значення рангів (Mean Ranks).

Таблиця Test Statistics^{a,b} містить результат перевірки гіпотези: емпіричне значення критерію (H) Chi-Square 1,220 та p -рівень значущості (Asymp. Sig.). У нашому випадку рівень значущості дорівнює 0,748. Так як $0,748 < 0,01$ та $0,05$, приймаємо гіпотезу H_0 , відмінності показників є статистично незначущими.

РОЗДІЛ 6

ОСНОВИ КОРЕЛЯЦІЙНОГО АНАЛІЗУ

- 6.1. *Поняття про кореляцію.*
- 6.2. *Моделі зв'язку у шкалах інтервалів і відношень.*
- 6.3. *Моделі зв'язку у порядкових вимірюваннях.*

6.1. Поняття про кореляцію

Значна кількість психологічних рис, властивостей, ознак військовослужбовців не є незалежними, а певним чином взаємопов'язані. Військовий психолог-дослідник часто має виявити наявність та специфіку зв'язку між цими ознаками, властивостями, рисами. Під час вивчення кореляцій намагаються встановити, чи існує зв'язок між двома показниками в одній вибірці (наприклад, показниками професійної готовності операторів радіолокаційних станцій та показниками успішності бойового чергування) або між двома різними вибірками (наприклад, порівняння структури ціннісних орієнтацій військовослужбовців із повних і неповних сімей), якщо цей зв'язок існує, то чи супроводжується підвищення одного показника зростанням (позитивною кореляцією) або зменшенням (негативною кореляцією) іншого.

Іншими словами, кореляційний аналіз допомагає встановити, чи можна передбачати можливі значення одного показника, знаючи величину іншого. У математичному сенсі завдання полягає у знаходженні зв'язку між двома рядами змінних (x_i та y_i), виміряних на тій самій вибірці досліджуваних осіб. Про наявність зв'язку (кореляції) між цими змінними можна говорити, коли зміна величини x веде до закономірної зміни величини y , а специфіка змін передбачувана.

Кореляція (від лат. *correlatio* – співвідношення, зв'язок, залежність) – це узгоджена зміна ознак. Якщо зі зміною однієї (чи кількох) величин змінюється інша, між показниками цих явищ спостерігатиметься кореляція. Поняття “кореляції” увійшло в науковий обіг у ХІХ ст. завдяки

роботам англійського антрополога та психолога Френсіса Гальтона (Galton) (1822–1911) та англійського математика Карла Пірсона (Pearson) (1857–1936).

NB! *Наявність кореляції між двома змінними не означає наявності причинно-наслідкової залежності між ними, проте дає змогу висунути таку гіпотезу.*

Кореляційний зв'язок – це узгоджена зміна двох або більше ознак (множинний кореляційний зв'язок). Кореляційний зв'язок вказує на те, що мінливість однієї ознаки певним чином відповідає мінливості іншої.

Кореляційна залежність – це зміни, які вносять значення однієї ознаки у ймовірність появи різних значень іншої ознаки.

У кореляційних зв'язках кожному значенню однієї ознаки може відповідати певний розподіл значень іншої ознаки, але не певне його значення.

Поняття “кореляційний зв'язок” та “кореляційна залежність” часто використовують як синоніми. Водночас узгоджені зміни ознак та кореляційний зв'язок між ними можуть вказувати не на залежність цих ознак між собою, а на їх залежність від якоїсь третьої ознаки або поєднання ознак, які не розглядають у дослідженні. Наприклад, успішність навчання курсантів і рівень їх абстрактного мислення корелюють між собою, але не можна стверджувати, що на навченість впливає лише чинник абстрактного мислення, оскільки неврахованими залишаються мотивація, пам'ять тощо.

Залежність має на увазі вплив, зв'язок – будь-які узгоджені зміни, які можуть бути пояснені сотнями причин. Кореляційні зв'язки не можна розглядати як підтвердження причинно-наслідкового зв'язку, вони вказують лише на те, що зміни однієї ознаки, як правило, супроводжують певні зміни іншої, але чи є причина змін в одній з ознак, чи вона опиняється за межами досліджуваної пари ознак, нам невідомо.

Основними показниками кореляційного зв'язку є напрям, сила (тіснота) і надійність (значущість).

Напря́м зв'язку визначають за знаком кореляції: позитивна – зв'язок прямий; негативна – зв'язок зворотний.

Позитивною кореляцією називають такий зв'язок між змінними, коли значення обох змінних зростають або зменшуються пропорційно: зі зменшенням (збільшенням) однієї зменшується (збільшується) інша (наприклад, чим вища особистісна тривожність, тим вищий ризик захворіти на виразку шлунку).

Негативною кореляцією називають такий зв'язок між змінними, коли зростання однієї змінної супроводжується спаданням іншої: зв'язок є обернено пропорційним (наприклад, чим більш боязливий військовослужбовець, тим менше в нього шансів стати лідером у підрозділі).

Силу (тісноту) зв'язку визначають абсолютною величиною кореляції r (змінюється від 0 до 1). Ступінь зв'язку між ознаками виражають **коефіцієнтом кореляції** і позначають r . Значення коефіцієнта можуть перебувати в діапазоні від -1 до $+1$. Використовують дві системи класифікації кореляційного зв'язку за його силою: загальна та часткова (таблиці 6.1. та 6.2.).

Таблиця 6.1

Загальна класифікація кореляційного зв'язку

Кореляційний зв'язок	Значення коефіцієнта кореляції
Сильний	$r > 0,70$
Середній	$0,50 < r < 0,69$
Помірний	$0,30 < r < 0,49$
Слабкий	$0,20 < r < 0,29$
Дуже слабкий	$r < 0,19$

Таблиця 6.2

Часткова класифікація кореляційного зв'язку

Кореляція	Рівень статистичної значущості
Високозначуща	$p \leq 0,001$
Значуща	$p \leq 0,01$
Тенденція до достовірного зв'язку	$p \leq 0,05$

Вказані класифікації не збігаються. Перша орієнтована на величину коефіцієнта кореляції, друга – визначає рівень значущості, якого досягає ця величина коефіцієнта кореляції при певному обсязі вибірки. Чим більший обсяг вибірки, тим меншої величини коефіцієнта кореляції виявляється достатньо, щоб кореляція була визнана достовірною, тобто при малому обсязі вибірки може виявитися так, що сильна кореляція виявиться недостовірною, а при великих обсягах вибірки навіть слабка кореляція може виявитися достовірною.

Зазвичай прийнято орієнтуватися на другу класифікацію, що враховує обсяг вибірки. Водночас необхідно пам'ятати, що сильна або висока кореляція – це кореляція з коефіцієнтом $r > 0,70$, а не просто кореляція високого рівня значущості.

Надійність (значущість) зв'язку визначають за рівнем статистичної значущості: чим менший p -рівень, тим вища статистична значущість, достовірність зв'язку.

Отже, завдання кореляційного аналізу полягає у встановленні напряму зв'язку між ознаками, вимірюванні його тісноти та перевірці рівня значущості одержаних коефіцієнтів кореляції.

Кореляції можуть бути лінійними та нелінійними. Якщо зі збільшенням чи зменшенням однієї змінної інша змінна зростає або зменшується, то *зв'язок лінійний* (наприклад, між особистісною пластичністю та схильністю до зміни соціальних установок). Якщо зі збільшенням однієї величини зміна іншої величини нелінійна, її можна описати іншими законами, то *зв'язок нелінійний* (прикладом нелінійної кореляції є перший закон Єркаса–Додсона: у міру збільшення інтенсивності мотивації якість діяльності змінюється за дзвіноподібною кривою: спочатку підвищується, а потім поступово знижується). На практиці у дослідженнях сильний лінійний кореляційний зв'язок трапляється досить рідко.

У психологічних дослідженнях найчастіше використовують коефіцієнт лінійної кореляції r Пірсона і методи рангової кореляції Спірмена та Кендала. Метод Пірсона є параметричним, тому не позбавлений

недоліків, властивих параметричним методам (необхідно, щоб дані були виміряні в інтервальних шкалах або розподіл не відрізнявся від нормального).

Метод рангової кореляції Спірмена є непараметричним, універсальним і зручним для роботи з даними, виміряними в будь-яких шкалах. Унікальність методу рангової кореляції полягає в можливості зіставлення не індивідуальних показників, а індивідуальних ієрархій або профілів, що недоступно жодному з інших статистичних методів, у тому числі методу лінійної кореляції. Коефіцієнт рангової кореляції рекомендовано використовувати, коли потрібно перевірити, чи узгоджено змінюються різні ознаки в того самого досліджуваного і наскільки збігаються індивідуальні рангові показники у двох досліджуваних або у досліджуваного і групи.

Вибір міри зв'язку. Для правильного вибору коефіцієнта кореляції необхідно визначити тип шкали, за допомогою якого наводять змінні. Можливі поєднання різних типів шкал і відповідні їм коефіцієнти кореляції наведено у таблиці 6.3.

Таблиця 6.3

Типи шкал та відповідні їм коефіцієнти кореляції

Тип шкал даних		Коефіцієнт кореляції	Приклад
Перша	Друга		
Інтервальна (відношень)	Інтервальна	Коефіцієнт Пірсона r_{xy}	Визначити зв'язок між рівнем інтернальності-екстернальності та рівнем професійного вигорання у командирів військових підрозділів
Рангова (порядкова)	Рангова	Коефіцієнт Спірмена r_s	Визначити зв'язок між ранговими оцінками якостей особистості, що за уявленнями досліджуваної особи мають міститися в "Я реальному" і "Я ідеальному"
		Коефіцієнт Кендала τ	Визначити ступінь збігу оцінок, визначених рядовим і сержантським складом у ході ранжування критеріїв боєздатності військового підрозділу

Тип шкал даних		Коефіцієнт кореляції	Приклад
Перша	Друга		
		Коефіцієнт конкордації W	Визначити ступінь узгодженості думок експертів щодо моделі вантажного військового автомобіля за певною кількістю критеріїв оцінювання
Номінальна (дихотомічна)	Номінальна	Коефіцієнт асоціації Q Коефіцієнт контингенції Φ	Визначити ставлення досліджуваних військовослужбовців до соціальної мережі (позитивне / негативне) залежно від статі
Номінальна	Інтервальна	Точково-бісеріальний r_{pb}	Визначити ступінь зв'язку та його статистичну значущість між типом темпераменту і рівнем особистої тривожності у групі досліджуваних військовослужбовців
Номінальна	Рангова	Рангово-бісеріальний r_{rb}	Визначити зв'язок між статтю і конфліктністю у підрозділі у групі досліджуваних військовослужбовців (чоловіків та жінок), проранжованій за проявами конфліктності стосовно товаришів по службі

6.2. Моделі зв'язку у шкалах інтервалів і відношень

Коефіцієнт кореляції Пірсона r

Призначення. Коефіцієнт кореляції r Пірсона характеризує наявність лише лінійного зв'язку між ознаками. Формула для обчислення коефіцієнта кореляції r Пірсона має такий вигляд:

$$r_{xy} = \frac{\sum(x_i - \bar{X}) \cdot (y_i - \bar{Y})}{(n-1) \cdot \sigma_x \cdot \sigma_y}$$

де X , Y – середні арифметичні елементів рядів x та y відповідно; n – кількість пар вимірювань; σ_x , σ_y – стандартні відхилення елементів рядів x та y відповідно.

Обчислення коефіцієнта кореляції r Пірсона – процес досить складний, тому розглянемо алгоритм його розрахунку у програмах MS Excel та SPSS.

Приклад: 22 досліджуваних військовослужбовця взяли участь у тестуванні для визначення рівня суб'єктивного контролю та мотивації до успіху. Отримані результати наведено в таблиці 6.4.

Таблиця 6.4

Результати дослідження рівня суб'єктивного контролю та мотивації до успіху

n	Рівень суб'єктивного контролю	Рівень мотивації до успіху	n	Рівень суб'єктивного контролю	Рівень мотивації до успіху
1	27	18	12	24	12
2	24	19	13	27	15
3	27	16	14	25	15
4	30	13	15	37	23
5	25	17	16	35	24
6	18	13	17	25	20
7	28	19	18	22	14
8	31	19	19	26	21
9	31	10	20	34	24
10	30	24	21	25	17
11	18	13	22	31	17

Необхідно перевірити гіпотезу про те, що досліджуваним із високим рівнем інтернальності притаманний вищий рівень мотивації до успіху.

Обчислення коефіцієнта кореляції r Пірсона за допомогою MS Excel:

1. Вводимо дані в таблицю у два стовпці.
2. Довільно обираємо вільну комірку у таблиці. У верхньому меню натискаємо кнопку “fx” (вставити функцію), у категорії “Статистичні” обираємо функцію “ПІРСОН” і натискаємо ОК.
3. У вікні “Масив 1” виділяємо дані першого стовпчика. У вікні “Масив 2” виділяємо дані другого стовпчика. Натискаємо ОК. У довільно обраній перед цим комірці отримуємо значення коефіцієнта кореляції r Пірсона.

B23		✕ ✓ f_x		=PEARSON(A1:A22;B1:B22)						
	A	B	C	D	E	F	G	H	I	
1	27	18								
2	24	19								
3	27	16								
4	30	13								
5	25	17								
6	18	13								
7	28	19								
8	31	19								
9	31	10								
10	30	24								
11	18	13								
12	24	12								
13	27	15								
14	25	15								
15	37	23								
16	35	24								
17	25	20								
18	22	14								
19	26	21								
20	34	24								
21	25	17								
22	31	17								
23		0,576								
24										

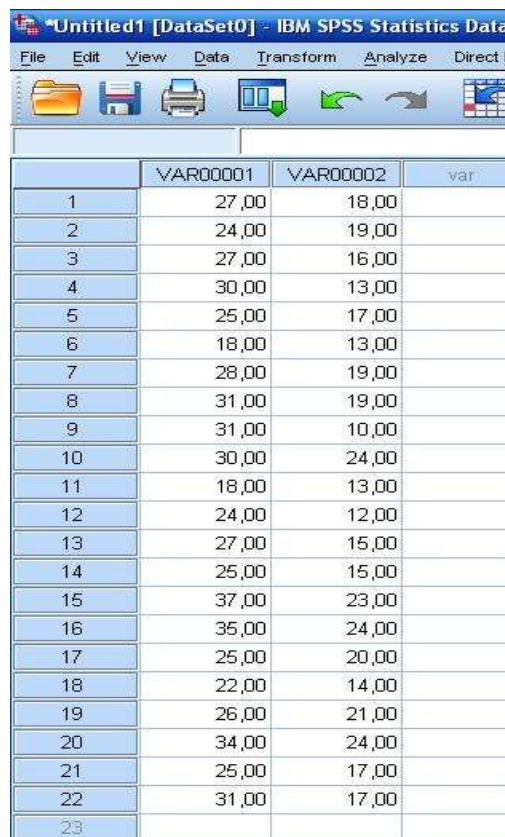
4. Отриманий коефіцієнт кореляції Пірсона $r = 0,58$. Знаходимо критичні значення коефіцієнта кореляції r Пірсона для нашої кількості досліджуваних:

$k = n - 2$	p	
	0,05	0,01
20	0,42	0,54

Порівнюємо отримане значення коефіцієнта кореляції Пірсона з критичними. Робимо висновок про те, що $r_{xy} > r_{кр}$, отже кореляція є статистично значущою для $p = 0,05$ та $p = 0,01$.

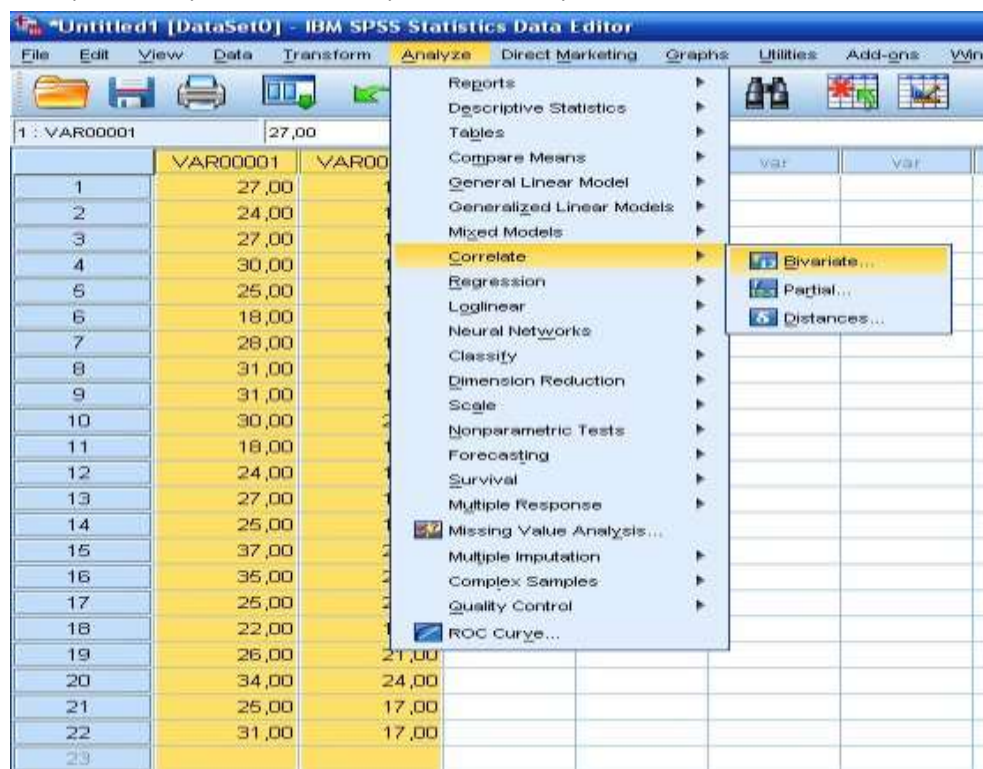
Обчислення коефіцієнта кореляції r Пірсона за допомогою SPSS:

1. Вводимо дані в таблицю у два стовпці: var1 та var2.

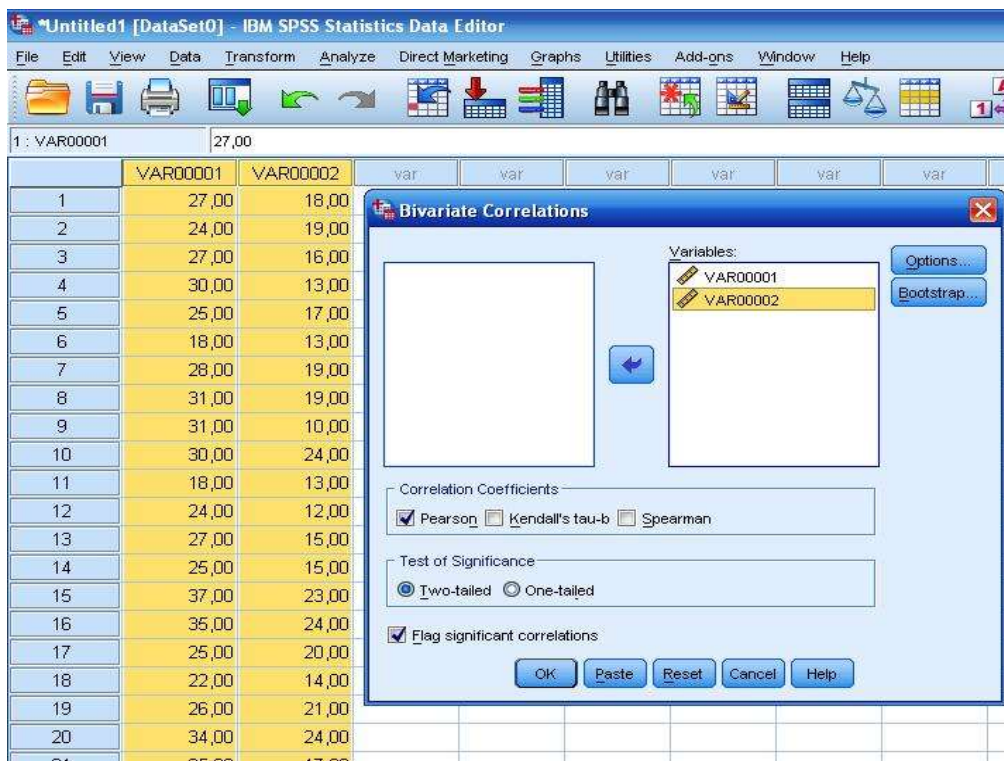


	VAR00001	VAR00002	var
1	27,00	18,00	
2	24,00	19,00	
3	27,00	16,00	
4	30,00	13,00	
5	25,00	17,00	
6	18,00	13,00	
7	28,00	19,00	
8	31,00	19,00	
9	31,00	10,00	
10	30,00	24,00	
11	18,00	13,00	
12	24,00	12,00	
13	27,00	15,00	
14	25,00	15,00	
15	37,00	23,00	
16	35,00	24,00	
17	25,00	20,00	
18	22,00	14,00	
19	26,00	21,00	
20	34,00	24,00	
21	25,00	17,00	
22	31,00	17,00	
23			

2. У верхньому меню вибираємо Analyze → Correlate → Bivariate.



3. У вікні, що відкривається, виділяємо дві змінні var1 і var2 і за допомогою кнопки ► переносимо їх у праве вікно. У цьому вікні в розділі Correlation Coefficients перевіряємо, що за замовчуванням обраний прапорець, розміщений на позначці “Pearson”.



Correlations

		VAR00001	VAR00002
VAR00001	Pearson Correlation	1	,576**
	Sig. (2-tailed)		,005
	N	22	22
VAR00002	Pearson Correlation	,576**	1
	Sig. (2-tailed)	,005	
	N	22	22

** Correlation is significant at the 0.01 level (2-tailed).

4. Натискаємо ОК та отримуємо результат:

Отримуємо симетричну матрицю, що містить коефіцієнт кореляції (Pearson Correlation) із значенням 0,576 і p -рівень значущості (Sig. (2-tailed)), що дорівнює 0,005. Значення

отриманого p -рівня менше за 0,05 та 0,01, а отже, кореляція є статистично значущою для обох рівнів.

Висновок: кореляція між показниками рівня суб’єктивного контролю та мотивації до успіху досліджуваних військовослужбовців є статистично значущою для $p = 0,05$ та $p = 0,01$.

У ході застосування коефіцієнта кореляції r Пірсона необхідно дотримуватися таких умов:

1. Змінні X та Y розподілені нормально.
2. Кількість значень змінної X має дорівнювати кількості значень змінної Y .
3. Ознака має бути виміряна за шкалою інтервалів або відношень.
4. Кількість значень n має бути від 5 до 1000.

6.3. Моделі зв'язку у порядкових вимірюваннях

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 визначення тісноти зв'язку між ознаками, виміряними у порядкових шкалах, застосовують методи рангової кореляції. До них належать: коефіцієнти рангової кореляції Спірмена та Кендала (використовують для визначення тісноти зв'язку між двома величинами) та коефіцієнт конкордації (встановлює статистичний зв'язок між кількома ознаками). Використання коефіцієнта лінійної кореляції Пірсона у разі, коли про закон розподілу і 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 їх значення можна упорядкувати та проранжувати. Ці методи також можуть бути використані стосовно ознак, вимірянних в інтервальних шкалах, проте їх ефективність у цьому разі буде нижчою.

Коефіцієнт кореляції Спірмена r_s

Призначення. Метод рангової кореляції Спірмена дає змогу визначити тісноту (силу) та напрям кореляційного зв'язку між двома

ознаками, виміряними у ранговій шкалі, або двома профілями (ієрархіями) ознак.

Для підрахунку рангової кореляції необхідно мати два ряди проранжованих значень. Спочатку показники ранжують окремо за кожною ознакою. Як правило, меншим значенням ознаки надають менший ранг, однаковим значенням ряду – середнє рангове число. Порівнювані ознаки можна ранжувати в будь-якому напрямі: у бік погіршення якості (ранг 1 отримує найшвидка, найрозумніша тощо досліджувана особа), і в бік поліпшення якості (ранг 1 отримує найповільніша досліджувана особа). Головне, щоб обидві змінні були проранжовані однаковим способом. Такими рядами значень можуть бути:

дві ознаки, виміряні в тій самій групі досліджуваних осіб;

дві індивідуальні ієрархії ознак, виявлені у двох досліджуваних осіб за тим самим набором ознак;

дві групові ієрархії ознак;

індивідуальна та групова ієрархії ознак.

Дві ознаки. Ранжують індивідуальні значення за однією ознакою, отриманою різними досліджуваними особами, а потім індивідуальні значення за іншою ознакою. Якщо дві ознаки пов'язані позитивно, то досліджувані особи, що мають низькі ранги за однією з них, матимуть низькі ранги й за іншою, а досліджувані, що мають високі ранги за однією з ознак, матимуть за іншою ознакою також високі ранги. Для підрахунку r_s необхідно визначити різницю (d) між рангами, отриманими певною досліджуваною особою за обома ознаками. Потім показники d певним чином перетворюють і віднімають від 1. Чим менша різниця між рангами, тим більшим буде r_s і тим ближчим до +1.

Якщо кореляції немає, всі ранги будуть перемішані і між ними не буде відповідності. Формула складена так, що в цьому разі r_s виявиться близьким до 0.

У разі негативної кореляції низьким рангам досліджуваних за однією ознакою відповідатимуть високі ранги за іншою ознакою, і навпаки. Чим більша розбіжність між рангами досліджуваних осіб за двома змінними, тим ближче r_s до -1 .

Два індивідуальні профілі. Ранжують індивідуальні значення, отримані кожною з двох досліджуваних осіб за певним (однаковим для них обох) набором ознак. Перший ранг отримує ознаку з найнижчим значенням, другий ранг – ознаку з вищим значенням. Очевидно, що всі ознаки мають бути виміряні в тих самих одиницях, інакше ранжування неможливе. Наприклад, неможливо проранжувати показники за опитувальником Кеттелла (16PF), якщо їх виражено в “сірих” балах, оскільки за різними факторами діапазони значень різні: від 0 до 13, від 0 до 20 і від 0 до 26. Ми не можемо сказати, який із факторів посяде перше місце за виразністю, поки не приведемо всі значення до єдиної шкали (найчастіше – шкали стенів).

Якщо індивідуальні ієрархії двох досліджуваних осіб пов'язані позитивно, то ознаки, що мають низькі ранги в однієї з них, матимуть низькі ранги і в іншій, і навпаки. Наприклад, якщо в однієї досліджуваної особи фактор E (домінантність) має найнижчий ранг, то й в іншій досліджуваної особи він матиме низький ранг, якщо в однієї досліджуваної особи фактор C (емоційна стійкість) має вищий ранг, то й в іншій досліджуваної особи матиме високий ранг тощо.

Два групових профілі. Ранжують середньогрупові значення, отримані у двох групах досліджуваних осіб за певним, однаковим для двох груп, набором ознак. Надалі застосовують логіку обчислень таку ж, як і в попередніх двох випадках.

Індивідуальний та груповий профілі. Ранжують окремо індивідуальні значення досліджуваної особи та середньогрупові значення за тим же набором ознак, отримані, як правило, за винятком цієї досліджуваної особи – вона не бере участі в середньогруповому профілі, з яким зіставлятиметься її індивідуальний профіль. Рангова кореляція дасть змогу перевірити, наскільки узгоджені індивідуальні та груповий профілі.

У всіх чотирьох випадках значущість отриманого коефіцієнта кореляції визначають за кількістю ранжованих значень N . У першому випадку ця кількість збігатиметься з обсягом вибірки n , у другому – кількістю спостережень буде кількість ознак ієрархії, у третьому і четвертому N є кількістю ознак, які зіставляють, а не кількість досліджуваних у групах.

Коефіцієнт рангової кореляції r Спірмена обчислюють за такою формулою:

$$r_s = 1 - \frac{6 \sum(D^2)}{n(n^2-1)},$$

де n – кількість ознак, які підлягають ранжуванню (показників, досліджуваних осіб); D – різниця між рангами за двома змінними для кожної досліджуваної особи; $\sum(D^2)$ – сума квадратів різниць рангів.

Приклад: 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 на цей час (ранг 20 отримує найвиразніша якість). Результати наведено в таблиці 6.5.

Таблиця 6.5

Результати ранжування якостей за ступенем бажаності

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

Необхідно розрахувати коефіцієнт кореляції Спірмена між виразністю якостей у досліджуваної особи на цей час та її ідеальними уявленнями, за підсумками розрахунку коефіцієнта рангової кореляції Спірмена – зробити висновок про рівень самооцінки досліджуваної особи.

Розв’язання. Сформулюємо статистичні гіпотези:

H_0 – кореляція між виразністю якостей у досліджуваної особи на цей час та її ідеальними уявленнями не відрізняється від нуля;

H_1 – кореляція між виразністю якостей у досліджуваної особи на цей час та її ідеальним уявленням відрізняється від нуля і є статистично значущою.

За допомогою MS Excel складаємо таблицю, до якої вносимо дві рангові послідовності: бажану N і реальну N' . Побудуємо додаткові стовпці, необхідні для подальшої роботи з методом: різниці рангів d і квадрату різниці рангів d^2 .

	Якості	N	N'	$d = N - N'$	d^2
1	Поступливість	14	15	-1	1
2	Сміливість	15	18	-3	9
3	Запальність	2	16	-14	196
4	Наполегливість	13	13	0	0
5	Нервозність	1	7	-6	36
6	Терплячість	17	10	7	49
7	Захопленість	12	20	-8	64
8	Пасивність	8	2	6	36
9	Холодність	10	19	-9	81
10	Ентузіазм	9	17	-8	64
11	Обережність	16	4	12	144
12	Примхливість	3	1	2	4
13	Повільність	18	6	12	144
14	Нерішучість	7	11	-4	16
15	Енергійність	20	12	8	64
16	Життєрадісність	19	8	11	121
17	Недовірливість	4	3	1	1
18	Впертість	5	9	-4	16
19	Безтурботність	11	14	-3	9
20	Сором'язливість	6	5	1	1
					1056

Розраховуємо за формулою значення коефіцієнта кореляції Спірмена:

$$r_s = 1 - \frac{6 \cdot 1056}{20 \cdot (20^2 - 1)} = 0,206.$$

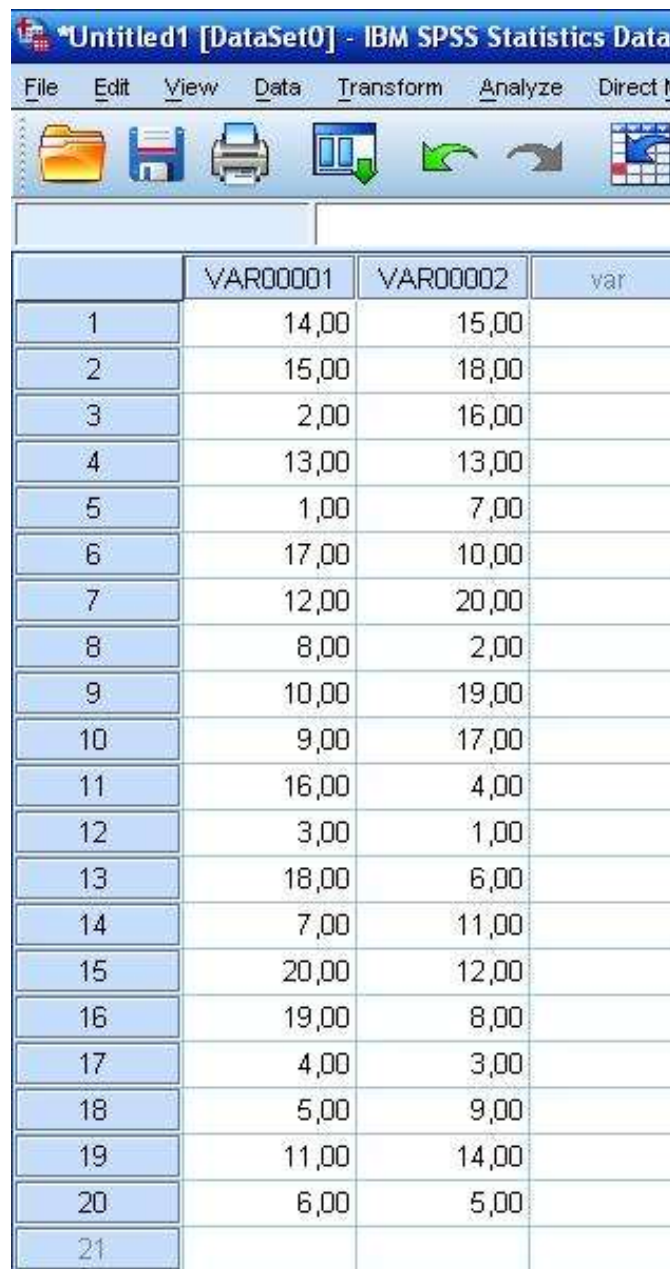
Отриманий коефіцієнт кореляції Спірмена $r = 0,206$. Критичні значення для коефіцієнта кореляції r Спірмена знаходимо за таблицею Критичних значень коефіцієнта кореляції r Спірмена. Пошук критичних величин ведемо за кількістю якостей у досліджуваній особі. У нашому прикладі $n = 20$, тому частина таблиці має такий вигляд:

n	p	
	0,05	0,01
20	0,45	0,57

Порівнюємо отримане значення коефіцієнта кореляції Спірмена із критичними. Робимо висновок про те, що $r_s < r_{крит}$, отже кореляція не є статистично значущою для $p = 0,05$ та $p = 0,01$. Приймаємо гіпотезу H_0 . Кореляція між виразністю якостей у досліджуваній особі на цей час та її ідеальними уявленнями не досягає рівня статистичної значущості, що вказує на відсутність зв'язку між виразністю якостей у досліджуваній особі на цей час та її ідеальними уявленнями.

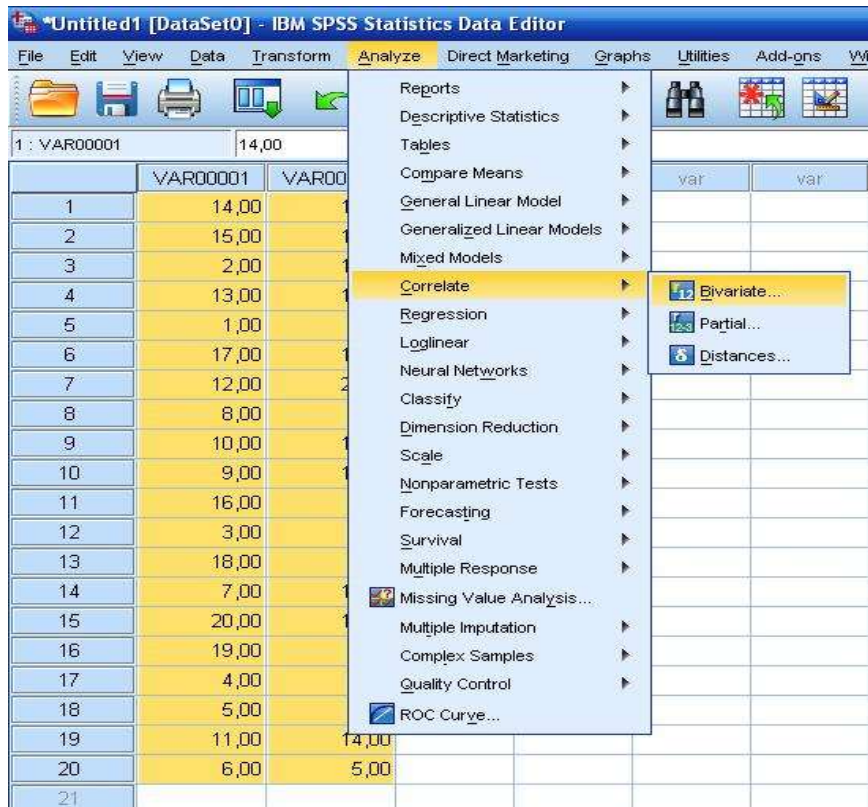
Обчислення коефіцієнта кореляції r Спірмена за допомогою SPSS:

1. Вводимо дані в таблицю у два стовпці: var1 та var2.

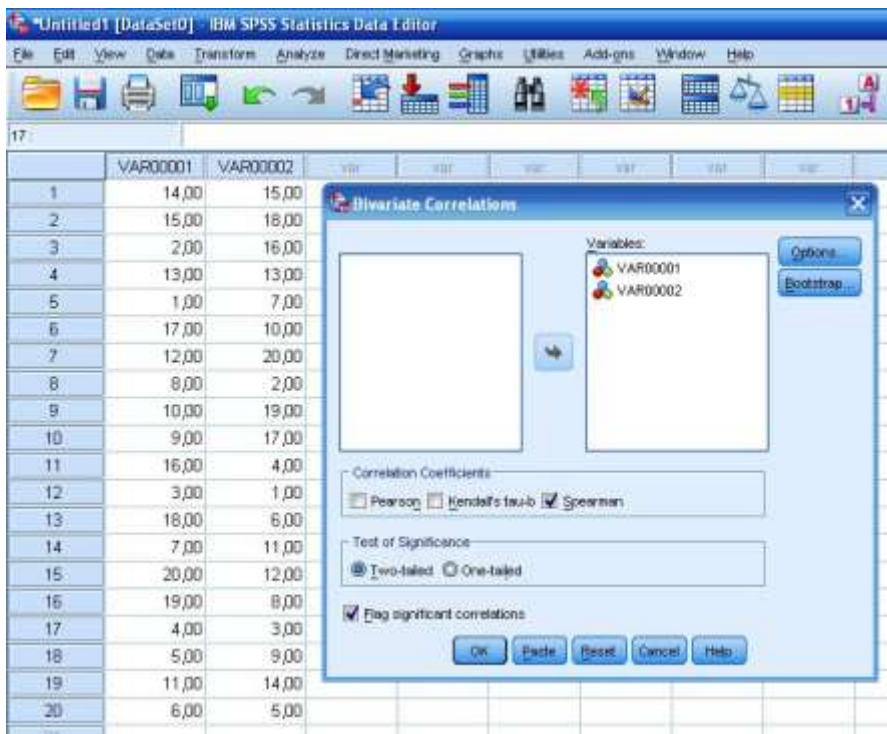


	VAR00001	VAR00002	var
1	14,00	15,00	
2	15,00	18,00	
3	2,00	16,00	
4	13,00	13,00	
5	1,00	7,00	
6	17,00	10,00	
7	12,00	20,00	
8	8,00	2,00	
9	10,00	19,00	
10	9,00	17,00	
11	16,00	4,00	
12	3,00	1,00	
13	18,00	6,00	
14	7,00	11,00	
15	20,00	12,00	
16	19,00	8,00	
17	4,00	3,00	
18	5,00	9,00	
19	11,00	14,00	
20	6,00	5,00	
21			

2. У верхньому меню обираємо Analyze → Correlate → Bivariate.



3. У меню, що з'явилося, виділяємо дві змінні var1 і var2, за допомогою кнопки ► переносимо їх у праве вікно. У цьому вікні в розділі Correlation Coefficients ставимо прапорець на Spearman.



4. Натискаємо ОК та отримуємо результат.

			VAR00001	VAR00002
Spearman's rho	VAR00001	Correlation Coefficient	1,000	,206
		Sig. (2-tailed)	.	,384
		N	20	20
	VAR00002	Correlation Coefficient	,206	1,000
		Sig. (2-tailed)	,384	.
		N	20	20

Отримуємо симетричну матрицю, яка містить коефіцієнт кореляції (Correlation Coefficient), що дорівнює 0,206 і p -рівень значущості (Sig. (2-tailed)), що у нашому випадку дорівнює 0,384. Рівень значущості $0,384 > 0,05$, отже, приймаємо гіпотезу H_0 про відсутність статистично достовірного кореляційного зв'язку.

У застосуванні коефіцієнта кореляції Спірмена r_s необхідно дотримуватися таких умов:

1. Розподіл не має значення.
2. Кількість значень змінної X має дорівнювати кількості значень змінної Y .
3. Ознака може бути виміряна за будь-якою кількісною шкалою або за ранговою шкалою.
4. За кожною змінною має бути проведено не менше 5 спостережень. Верхню межу вибірки визначають за наявними таблицями критичних значень.

Коефіцієнт конкордації (узгодженості) Кендала W

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

Коефіцієнт конкордації Кендала визначають за такою формулою:

$$W = \frac{12 \sum_{i=1}^n D_i^2}{m^2 (n^3 - n)}$$

де: n – кількість об'єктів (досліджуваних осіб), які оцінюють; m – кількість рангових послідовностей (кількість експертів); $D_i = d_i - \bar{d}$ – відхилення суми рангів i -го об'єкта від середньої суми рангів усіх об'єктів.

Значення коефіцієнта конкордації, на відміну коефіцієнта кореляції, вкладаються в інтервалі $0 \leq W \leq 1$. Коефіцієнт конкордації дорівнює одиниці за повного збігу всіх рангових послідовностей. Якщо думки експертів (рангові послідовності) повністю протилежні, коефіцієнт конкордації дорівнює нулю (коефіцієнт кореляції в цьому разі дорівнює -1).

Приклад: експертна комісія військової частини з 5 осіб ранжувала 7 типів тактичних штурмових рюкзаків (ранг 1 присвоювали найкращому рюкзаку). Необхідно визначити ступінь узгодженості думок експертів за результатами оцінювання.

Розв'язання. Сформулюємо статистичні гіпотези:

H_0 – думки експертів не узгоджуються між собою;

H_1 – думки експертів узгоджуються між собою.

Створюємо таблицю, до якої вносимо оцінки, які виставили експерти, рангові суми d_i , відхилення D_i суми рангів від середньої суми рангів усіх об'єктів і D_i^2 .

Тип рюкзака (n)	Експерти (m)					d_i	$D_i = d_i - \bar{d}$	D_i^2
	1	2	3	4	5			
1	1	1	2	1	3	8	-12	144
2	3	2	1	2	1	9	-11	121
3	4	5	7	4	5	25	5	25
4	2	3	5	6	4	20	0	0
5	6	6	6	3	2	23	3	9
6	7	4	4	5	6	26	6	36
7	5	7	3	7	7	29	9	81
						140		416

Середня сума рангів усіх об'єктів $\bar{d} = \frac{140}{7} = 20$.

Коефіцієнт конкордації Кендала визначаємо за такою формулою:

$$W = \frac{12 * 416}{5^2 (7^3 - 7)} = 0,594.$$

Висновок: значення коефіцієнта конкордації Кендала становить 0,59, що вказує на достатній рівень узгодженості думок експертів щодо оцінювання тактичних штурмових рюкзаків.

Обчислення коефіцієнта конкордації Кендала за допомогою SPSS:

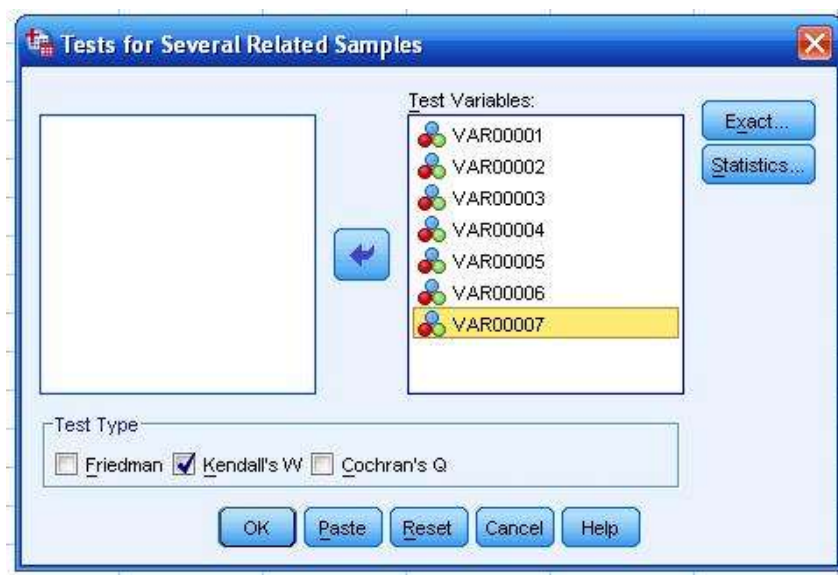
1. Вводимо дані в таблицю у стовпці: var1 – var7.

	VAR00001	VAR00002	VAR00003	VAR00004	VAR00005	VAR00006	VAR00007	var
1	1,00	3,00	4,00	2,00	6,00	7,00	5,00	
2	1,00	2,00	5,00	3,00	6,00	4,00	7,00	
3	2,00	1,00	7,00	5,00	6,00	4,00	3,00	
4	1,00	2,00	4,00	6,00	3,00	5,00	7,00	
5	3,00	1,00	5,00	4,00	2,00	6,00	7,00	
6								

2. У верхньому меню обираємо Analyze → Nonparametric Tests → Legacy Dialogs → K Related Samples.

The screenshot shows the SPSS interface with the 'Analyze' menu open. The path 'Nonparametric Tests' > 'Legacy Dialogs' > 'K Related Samples...' is selected. The background data table shows columns VAR00005, VAR00006, VAR00007, and 'var' highlighted in yellow.

3. У меню, що з'явилося, виділяємо змінні var 1 – var 7 і за допомогою кнопки ► переносимо їх у праве вікно. У цьому вікні в розділі Test Type ставимо прапорець на Kendall's W.



4. Натискаємо ОК та отримуємо результат.

Kendall's W Test

Ranks	
	Mean Rank
VAR00001	1,60
VAR00002	1,80
VAR00003	5,00
VAR00004	4,00
VAR00005	4,60
VAR00006	5,20
VAR00007	5,80

Test Statistics	
N	5
Kendall's W ^a	,594
Chi-Square	17,829
df	6
Asymp. Sig.	,007

a. Kendall's Coefficient of Concordance

Значення коефіцієнта конкордації Кендала становить 0,594. Також у віконці Test Statistics є значення критерію Пірсона χ^2 , яке дорівнює 17,829. За допомогою критерію Пірсона χ^2 проведемо додаткову перевірку гіпотези H_0 про те, що думки експертів не узгоджені між собою. Знаходимо критичні значення критерію Пірсона χ^2 для числа ступенів свободи $df = n - 1 = 6$:

df	p	
	0,05	0,01
20	12,592	16,812

Порівнюємо отримане значення критерію Пірсона χ^2 з критичними. Робимо висновок про те, що $\chi^2_{emn} > \chi^2_{кр}$, кореляція є статистично значущою для $p = 0,05$ та $p = 0,01$. Отже, маємо

достатньо тісну узгодженість думок експертів щодо оцінювання тактичних штурмових рюкзаків.

Питання для самоконтролю

1. Що таке кореляція?
2. Розмежуйте поняття “кореляційний зв’язок” та “кореляційна залежність”. Чому ці поняття не можна вважати синонімами?
3. Перерахуйте основні показники кореляційного зв’язку, розкрийте значення кожного з них.
4. Що таке коефіцієнт кореляції і яких значень він може набувати?
5. Що необхідно визначити для правильного вибору необхідного коефіцієнта кореляції?
6. Охарактеризуйте коефіцієнт кореляції Пірсона.
7. Перерахуйте основні умови, необхідні для використання коефіцієнта кореляції Пірсона.
8. Охарактеризуйте коефіцієнт кореляції Спірмена.
9. Перерахуйте основні умови, необхідні для використання коефіцієнта кореляції Спірмена.
10. Охарактеризуйте коефіцієнт конкордації (узгодженості) Кендала.

Завдання для самостійної роботи

1. У тестуванні з визначення стилю поведінки у конфліктній ситуації за методикою Томаса – Кілмана взяли участь 10 операторів радіолокаційних станцій. Визначити, чи існує статистично достовірний кореляційний зв’язок між показниками за шкалами “суперництво” та “компроміс”, виміряними серед операторів. Отримані дані наведено в таблиці.

Вихідні дані для вирішення завдання

n	“Суперництво”	“Компроміс”	n	“Суперництво”	“Компроміс”
1	8	5	6	2	7
2	4	5	7	4	5
3	0	8	8	4	6
4	4	6	9	2	8
5	7	3	10	6	3

2. На основі психологічного дослідження отримано емпіричні дані про 10 військовослужбовців підрозділу: оцінювання власних лідерських якостей та психологічного клімату у підрозділі. Кожна оцінка надана у вигляді натурального числа, де 1 – максимальний бал, 10 – мінімальній. Необхідно визначити, чи існує достовірний зв'язок між самооцінюванням військовослужбовців та їх міжособистісними взаєминами у військовому підрозділі. Отримані дані наведено в таблиці.

Вихідні дані для вирішення завдання

Показники оцінювання	Досліджувані <i>n</i>									
	1	2	3	4	5	6	7	8	9	10
Рівень лідерських якостей	2	1	3	2	3	5	3	6	5	2
Психологічний клімат у підрозділі	4	7	2	3	1	2	4	2	3	6

3. П'ятьом експертам-психологам було надано чотири різних моделі організації психологічної підготовки військовослужбовців для визначення кращої. Найбільший бал відповідає найкращому варіанту. Визначте ступінь узгодженості думок експертів за результатами оцінювання. Бали, виставлені експертами, наведено в таблиці.

Вихідні дані для вирішення завдання

Експерти	Варіанти організації психологічної підготовки			
	1	2	3	4
1	3	4	2	1
2	4	3	1	2
3	3	4	2	1
4	3	4	1	2
5	4	3	2	1

РОЗДІЛ 7

СПОСОБИ ВІЗУАЛІЗАЦІЇ І ПОДАННЯ ДАНИХ

7.1. Основні правила візуалізації даних.

7.2. Таблиця, як техніка візуалізації.

7.3. Графічні способи подання даних.

7.1. Основні правила візуалізації даних

Як було зазначено у розділі 1 цього посібника, процес аналізу даних містить три етапи. Спочатку дані потрібно підготувати, тобто зібрати, очистити і відібрати потрібні для дослідження. Тут ми тільки починаємо занурюватися у цифри. Далі ми опрацьовуємо результати, тобто висуваємо статистичні гіпотези та опрацьовувати масиви даних за допомогою різноманітних статистичних критеріїв. Це викликає у багатьох значні труднощі, часто стає навіть спричиняє виникнення жартів про гуманітаріїв і математику. На останньому етапі нам потрібно синтезувати отримані результати, тобто надати певних значень і категорій нашим науковим здобуткам і зробити висновки. На цьому етапі може здаватися що усі проблеми позаду, але насправді тут дуже легко можна помилитися і потрапити в зону ілюзій, адже кожному досліднику після титанічної роботи, так кортить нарешті отримати бажаний результат.

Найпростішими способами статистичного аналізу є описові статистики, які, як видно із назви, на елементарному рівні кількісно описують або підсумовують масив даних. У попередніх розділах ми розглядали, як за допомогою порівняння й аналізу мір центральної тенденції та мір мінливості можна з'ясувати тип і форму розподілу із яким ми працюємо. Хоча такий спосіб може надати лише загальне уявлення про розподіл масиву даних, ним необхідно користуватися як на початковому етапі опрацювання даних, щоб коректно вибрати статистичні критерії, так і наприкінці для порівняння побаченого й отриманого фактично. В науці іноді стається так, що безглузда помилка сприяла геніальним відкриттям,

але швидше за все, це не наш випадок, тому завжди порівнюйте і сумнівайтеся.

Крім візуального аналізу співвідношень числових параметрів розподілу даних, корисно також переводити їх у табличні та графічні форми візуалізації, що дає змогу побачити отримані дані з іншої точки зору і з'ясувати більше інформації. Порядок розташування, колір, розмір, співвідношення, форми – усе має значення для отримання інформації у символічній формі та її закодування. Візуалізація дає змогу охоплювати великі обсяги інформації, наче стискаючи її, і роблячи компактною, робити доступнішим сприйняття складної інформації, пришвидшуючи порівняння величин і полегшуючи виявлення ознак у даних. Аналіз даних, формулювання певних висновків на основі масивів даних – це, насамперед, стиснення інформації, її узагальнення. Так, побудова простої стовпчастої діаграми засобами будь-якої електронної таблиці, дає змогу швидко і просто (візуально) визначити моду або медіану. Детальніше способи та форми візуалізації даних буде розглянуто далі, але загалом, треба пам'ятати, що візуалізація – не лише важливий елемент подання даних, наприклад, під час публікації результатів дослідження, а й важливий елемент аналізу для швидкого знаходження особливостей у масивах даних, формулювання певних висновків, пошуку шляхів подальшого аналізу.

Важливою рисою візуалізованих даних є їх переконливість, тому дуже важливо уникати викривлення інформації у процесі візуалізації, зокрема, в хорошій візуалізації наочність загальної картини не заважає сприйняттю деталей.

Відповідно до заданої мети розрізняють **два типи візуалізації**: як техніку аналізу даних (побачити у графічній або табличній формі те, що не побачили у числовій) і як техніку подання даних, тобто інфопродукту. Головне, що вирізняє візуалізацію як інфопродукт – чітке повідомлення. Якщо в даних немає нічого такого, що хотілося би в них показати сприймачеві, жодна візуалізація не зробить їх цікавими.

Побудові візуалізації має передувати момент осяяння, усвідомлення в даних чогось такого, чим варто поділитися: масштабом явища, якоюсь залежністю, закономірністю, співвідношенням, аномалією, тощо.

Далі треба обрати метод візуалізації, який допоможе якнайдоступніше донести повідомлення. Слід пам'ятати, що увага людини, опрацьовуючи патерни (шукані ознаки) в полі уваги, завжди розділяє патерн на дві сутності – фігуру і тло (важливе і те, серед чого воно важливе). У разі тривалого споглядання багатого на деталі патерна можливе розкладання його на субпатерни, в кожному з яких так само знаходять фігуру і тло.

У хорошій візуалізації головна ознака (патерн) відразу помітна і сприймається фігурою, а її пов'язаність із тлом передає чи принаймні відтіняє головне повідомлення. Відповідно різні повідомлення потребують різних технік візуалізації. Однак незалежно від техніки візуалізації говорити мають самі дані.

Існує чимало технік візуалізації, і серед них трапляються дуже складні, але для передання багатьох повідомлень досить дуже простих технік, інколи трохи доопрацьованих та конкретизованих відповідно до мети дослідження.

У цьому розділі розглянемо, як правильно та коректно візуалізувати дані за допомогою табличних і графічних форм. Таких форм існує досить багато, але часто дослідник-початківець не вміє правильно обрати потрібну і далі звичайних тривимірних стовпчиків, заретушованих у різні кольори, справа не йде. Такий спосіб візуалізації може справити перше враження на людей, не компетентних у цій сфері, але досвідчені науковці навряд чи сприймуть його всерйоз, адже до форм подання результатів дослідження висувають низку **вимог**: інформативність, лаконічність, коректність, простота та ін. Отже, розглянемо способи подання, або візуалізації даних.

7.2. Таблиця, як техніка візуалізації

Систематизацію даних у табличній формі (табулювання) проводять відповідно до макетів таблиць, які дослідник склав ще на етапі розроблення форми для збирання інформації. Однак у процесі дослідження може виникнути ідея зіставлення даних, які раніше не планували зіставляти, що потребує додаткових обчислень. Зазвичай ідея додаткового табулювання виникає через необхідність аналізу отриманих даних і характерна для пошукових досліджень. Розрізняють одновимірне (просте) і перехресне (крос) табулювання.

Просте табулювання – це підрахунок кількості подій, які потрапляють у кожну категорію, коли категорії базуються на одній змінній. Одновимірний розподіл дає змогу підсумувати частоту, з якою різні значення певної змінної спостерігаються в наборі даних. Наприклад, ми маємо набір даних про 26 військовослужбовців за термінами проходження ними військової служби від 14 до 26 років. Частотний розподіл змінної "термін проходження військової служби" у таблиці має такий вигляд:

Термін	14	15	16	17	18	19	20	21	22	23	24	25	26
Частота	2	–	1	1	2	3	5	4	–	5	–	2	1

Якщо змінні інтервальні і значень дуже багато, доцільно значення змінних розбити на певні інтервали, які об'єднати за певними категоріями:

Термін	<15	15 – 20	21-25	>25
Частота	2	12	11	1

Можна також визначити частоту певного значення у вигляді частки чи відсотка від усіх спостережень. У вигляді таблиці нижче наведено частотний, відносний і накопичений частотний розподіли.

Термін	Частота	Відносна частота	Накопичена частота
<15	2	0,08	0,08
15 – 20	12	0,46	0,54
21 – 25	11	0,42	0,96
>25	1	0,4	1,00
Разом	16	1,00	–

Як видно з таблиці, підсумок відносної частоти в кожній з категорій дорівнює 1 (або 100%) – оскільки включає всі спостереження.

Одновимірний розподіл може бути реалізований щодо якісних та кількісних змінних, дискретних і неперервних. Можна класифікувати дані за кількісними характеристиками, наприклад віковими групами, рівнем доходу тощо, а також класифікувати за якісними – статтю, професією, національністю тощо.

Ще одним способом узагальнення даних у таблиці є перехресне, або "крос-табулювання". По суті, це таблиця, у якій зіставлено дві окремі характеристики (змінні) – по колонках і рядках. Так, у колонках розміщені різні значення (чи групи значень, категорії) однієї змінної, а в рядках – значення іншої змінної. У SPSS ця функція реалізована в меню *Analyze* → *Descriptive Statistics* → *Crosstabs*.

Кількість військовослужбовців, призваних на військову службу за контрактом (за областями)	2020	2021	2022
Вінницька	3544	3732	3918
Волинська	1954	2060	2197
Дніпропетровська	6534	6999	6095
Донецька	1287	1353	1141

Ще до початку аналізу важливо розуміти: грамотне й охайне структурування даних передбачає, що в таблиці всі змінні мають бути в колонках, всі спостереження – в рядках, а на їх перетині – значення змінних для спостережень.

Для аналізу даних, можна порівнювати одні змінні з іншими, робити крос-табуляцію, будувати двовимірні розподіли, наприклад, щоб на найпростішому рівні виявити залежності між змінними.

У процесі створення таблиці розподілу для набору даних, що містить велику кількість спостережень, варто дотримуватися таких основних правил:

- знайти найменше та найбільше значення змінної;
- визначити "ширину" категорій, за якими будуть згруповані значення змінної;
- включити всі можливі значення змінної.

Визначення ширини інтервалу, потребує знаходження компромісного рішення: інтервали мають бути достатньо вузькими, щоб всі або більшість спостережень не потрапили в один інтервал, і при цьому достатньо широкими, щоб не було по одному спостереженню в інтервалі. Також важливо, щоб інтервали були взаємовиключними.

Належно побудована таблиця може ефективно унаочнювати певні властивості даних. Важко давати загальні поради щодо організації таких таблиць, адже кожен випадок може потребувати спеціального рішення, але деякі принципи досить-таки прості, хоча й не абсолютні:

- заголовок має містити згадку про основну змінну, тобто відповідати сутності таблиці;
- одиниці вимірювання вказують у заголовку чи заголовкових комірках у головці чи боковику, іншими словами, в комірках мають бути тільки значення;
- об'єднані комірки в головці чи боковику таблиці мають відповідати структурі даних;
- таблицю не слід переускладнювати, краще навести кілька простих таблиць, а якщо уникнути значного розміру таблиці неможливо, дублювати головку на кожній сторінці і ввести елементи візуального структурування, наприклад, підзаголовки в об'єднаних рядках;

– комірки не обов'язково мають містити цифри чи слова – таблиця може бути носієм простих діаграм, графіків і гістограм, крім того, використання піктограм в таблиці може пришвидшити її сприйняття.

Оформлення таблиці як засобу візуалізації може бути різноманітним, але слід уникати зайвих ліній: достатніми є заливка білим фоном і лінії, що відокремлюють головку і боковик від даних. Вирівнювати текст у таблиці потрібно по лівому краю, але якщо наведено багато числових значень, що відрізняються на кілька розрядів, їх варто вирівнювати по правому краю, а якщо це десяткові дробки – за десятковим символом.

Грамотно і лаконічно побудовані таблиці легко сприймаються і прикрашають сторінку. Якщо око "намагається оминати" вашу таблицю, перевірте, чи не переобтяжена вона зайвими графічними елементами, чи достатні відбивки тощо.

7.3. Графічні способи подання даних

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

Кінцеві візуалізації від робочих з процесу аналізу відрізняють не лише більша охайність і естетичність. Якщо в ході аналізу триває пошук закономірностей і цікавих тенденцій, то у кінцевій візуалізації ми, як правило, наперед знаємо, що хочемо показати, знаємо, які емоції викликати у споживача і чим його вразити.

Якщо робочі візуалізації зрозумілі для людини, яка підготовлена до сприйняття даних і перебуває в контексті роботи з ними, то кінцеві візуалізації мають занурювати у контекст непідготовлену людину.

Методи вирішення цих задач різноманітні, і частина з них виходить за межі візуалізації даних. Для занурення непідготовленої людини в контекст, виокремлення з даних головного, найчастіше використовують засоби інфографіки, яка може містити елементи візуалізації даних, або проілюструючи певний набір даних, але не відобразити його.

Техніки візуалізації. Серед великої кількості технік візуалізації трапляються дуже складні, але для передавання багатьох повідомлень досить простих технік, інколи трохи доопрацьованих. Наприклад, вікостатева (популяційна, демографічна) піраміда, стандартний інструмент візуалізації даних про демографічні тенденції популяції та порівняння популяцій – це дві гістограми, поставлені вертикально нулем до нуля.

Таким чином пошук правильної техніки варто починати від простіших, а якщо вони не передають потрібну ідею – переходити до складніших. Для вибору належної техніки візуалізації, яка б унаочнювала ту властивість даних, яку хотілося б показати сприймачеві, доречно скористатися класифікацією Джина Желязни (додаток 2). Техніки візуалізації в ній організовано у вигляді дерева, коренем якого є питання: "Що ми хочемо показати?".

Гістограма показує розподіл значень у множині об'єктів. На одній осі, як правило, горизонтальній, відкладають значення, а на вертикальній – показують, скільком об'єктам у множині відповідає це значення. Оскільки кількість об'єктів – натуральне число, гістограми завжди будують від нуля.

За виглядом гістограми можуть бути стовпчастими (тут та нижче за текстом див. додаток 2), або лінійними, що зближує їх зі стовпчастими та лінійними діаграмами.

Стовпчасті діаграми. У звичайній стовпчастій діаграмі призначеній для порівняння кількох значень, значуща лише одна вісь, а вздовж іншої розташовано стовпці зручної для сприйняття ширини, висота яких кодує значення. Стовпчасті діаграми часто будують не від нуля.

Складнішим різновидом стовпчастої діаграми є стовпчаста діаграма зі змінною шириною стовпців, у якій ширина стовпців не фіксована, а кодує другу змінну. Проблемою сприйняття такої діаграми є те, що

співвідношення за однією змінною оцінити легко, але другу змінну споживач сприймає опосередковано за площею стовпців. Якщо обидві змінні дають спільний ефект, цей тип діаграми придатний для використання, якщо змінні і їхній вплив незалежні – ні.

Ще складніший варіант стовпчастої діаграми – **стосова**. У ній кожен стовпець складається з шарів, товщина кожного з яких кодує змінну, а їх сума становить певне значущє ціле.

Існує вид стосової діаграми, коли всі стовпці мають рівну висоту. У цьому разі стосова діаграма показує лише різний розподіл змінних між об'єктами-стовпцями.

Лінійні діаграми (графіки), діаграми розсіювання і бульбашкові діаграми. Відома зі шкільного курсу алгебри декартова система координат є одним із найпотужніших інструментів візуалізації даних. Просте розміщення точок даних у координатній сітці дає діаграму розсіювання, яка, будучи зручним інструментом пошуку кореляції під час аналізу, може бути адекватним засобом візуалізації даних, якими можна відобразити як тенденцію, так і типові відхилення від неї.

Якщо треба показати відповідність чи відхилення експериментальних даних від теоретичної (обчисленої) кривої, до точок додають "вуса", що показують межі ймовірної похибки, або будують експериментальну криву – смугу, ширина якої дорівнює похибці вимірювання.

Якщо точок дуже багато і тенденція чітка, лишається власне лінія, яку називають лінійною діаграмою. У лінійних діаграмах можна поєднувати декілька ліній, що дуже зручно для порівняння параметрів кількох подібних об'єктів. Кодують лінії кольором або різними пунктирами. При цьому важливо не перевантажувати діаграму: для більш ніж чотирьох ліній краще сприйматимуться кілька менших графіків, наведених поруч в одному масштабі.

Якщо треба показати залежність трьох змінних, точки на діаграмі розсіювання замінюють на кола, діаметр яких кодує додаткову змінну. Важливо правильно обрати масштаб для цієї змінної: різниці в діаметрі

мають бути помітними, але кола не слід робити завеликими, щоб не заважати сприйняттю позицій їх центрів у координатній площині.

Колові діаграми. Аналогічно поділу на частини стовпців у стосовій діаграмі, кола можна ділити на сектори, щоб показати частки у цілому. Така "тортова" діаграма належно відображає співвідношення між кількома значними частками, але стає неефективною, якщо часток дуже багато, тоді дрібні частки об'єднують у групу "інші".

Інші техніки візуалізації. Існує ще багато технік візуалізації – від досить строгих (Ейлерові кола та споріднені з ними діаграми Венна, які використовуються для унаочнення відносин між множинами і їх перетинами до екзотики типу личок Чернова, інформація про практичну користь яких досить неоднозначна). Щоб продемонструвати структури і залежності, використовують графи і організаційні діаграми. Для деяких візуалізацій використовують не декартові координати, а, наприклад, полярні.

Інструменти візуалізації. Прості засоби візуалізації включено 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, таких як Excel Microsoft, SPSS або Statistika, вже закладені пакети візуалізації даних. Деякі засоби побудови діаграм вбудовано також у графічні пакети, як от Adobe Illustrator, їх використовують для графічно інтенсивних візуалізацій. Ілюстраторні діаграми підтримують колові, стовпчасті та деякі інші види діаграм, у них включено графічні примітиви (наприклад, заміну прямокутного стовпця на фігурний), а також передбачено можливість оперативно перемалювати діаграму-шаблон, змінивши файл даних із роздільниками-табуляторами.

Однак для побудови якісних візуалізацій краще використовувати спеціалізовані інструменти (табл. 7.1), тим більше, що серед них є безкоштовні і досить прості у використанні.

Програмні продукти для роботи із візуалізацією даних

№ п/п	Назва програмного продукту	Короткий опис можливостей використання	Інтернет-адреса
1.	RAW	Онлайн-сервіс швидкої візуалізації даних. Дає змогу швидко будувати досить екзотичні діаграми, зокрема bump chart, alluvial chart, streamgraph тощо. Дані можна завантажити з кліпборду CSV, обрати тип візуалізації і прив'язку змінних до наявних способів кодування інформації. Далі можна детальніше налагодити можливість передання кольорів, розмірів і специфічних для поточного типу діаграми параметрів. а результат експортувати у SVG для подальшої обробки в Inkscape чи Illustrator, у зображення PNG або об'єкт JSON	http://app.raw.densitydesign.org
2.	Chartbuilder	Дуже простий у використанні інструмент, що приймає дані у JSON або копіюванням (роздільник — табулятор). Будує прості XY графіки. Експортує у PNG, SVG і JSON	https://quartz.github.io/Chartbuilder/
3.	Silk	Платформа публікації даних із можливостями їх обробки і візуалізації	www.silk.co
4.	Tableau Public	Десктопна програма для Windows та Mac OS, призначена для створення інтерактивних комбінованих візуалізацій без необхідності програмувати чи малювати. Має безплатну версію, у разі використання якої «воркбуки» зберігаються тільки на серверах Tableau. У програми велике ком'юніті	http://www.tableau.com/
5.	Quadrigram	Сервіс побудови інтерактивних сторінок із тексту, зображень і діаграм	http://www.quadrigram.com/

Як обманювати за допомогою візуалізації даних? Існує досить багато прикладів, коли за допомогою математичної статистики намагаються видати бажане за дійсне. Щоб уникнути цього, корисно прочитати книгу Даррела Хаффа "Як брехати за допомогою статистики". Однак трапляється й так, що правильне дослідження можна подвійно трактувати.

Як приклад наведемо "Парадокс Сімпсона", названий на честь дослідника Едварда Сімпсона, який у 1951 описав цей феномен. Університет Берклі у 1973 році звинуватили у ґендерній нерівності. Для ілюстрації дещо спростимо вихідні умови. Нехай в університеті є два факультети: А та В.

На факультеті А:

Стать	Подали заяви	Прийняті	Відсоток прийнятих
Чоловіки	900	450	50
Жінки	100	80	80

На факультеті В:

Стать	Подали заяви	Прийняті	Відсоток прийнятих
Чоловіки	100	10	10
Жінки	900	180	20

Якщо подивитися на відсоток прийнятих окремо по факультетах А та В, то можна зробити висновок, що дискримінують чоловіків. Однак, якщо підсумувати кількість прийнятих по факультетах, то ситуація виявиться зовсім іншою.

Разом факультетах А та В:

Стать	Подали заяви	Прийняті	Відсоток прийнятих
Чоловіки	1000	460	46
Жінки	1000	260	26

У цьому разі можна припустити факт дискримінації жінок.

Будь-які масиви даних мають бути візуалізовані коректно й об'єктивно. Розглянемо детальніше **найпоширеніші помилки** (або ж хитрощі) у візуалізації даних.

Відсотки замість абсолютних величин. Дуже часто, щоб надати своїм даним значущості, свідомо або несвідомо, дослідники переводять абсолютну кількість значень даних у відсотки, тобто відносну кількість, що особливо відчутно на малих вибірках. Наприклад, коли зазначають, що у 2020 році у військовій частині зафіксовано одне правопорушення, а у 2021 році – два, переводять кількість з абсолютної величини у відносну, тобто у відсотки, і роблять хибний висновок, що кількість правопорушень зростає аж (!) на 200%. Такий висновок є хибним і таким, який ні про що не говорить, адже тут справа у звичайній випадковості. Інший приклад: 20 листопада 2022 року Міністр оборони України Олексій Резніков публікує на своїй сторінці наступну інформацію (рисунок 7.1):

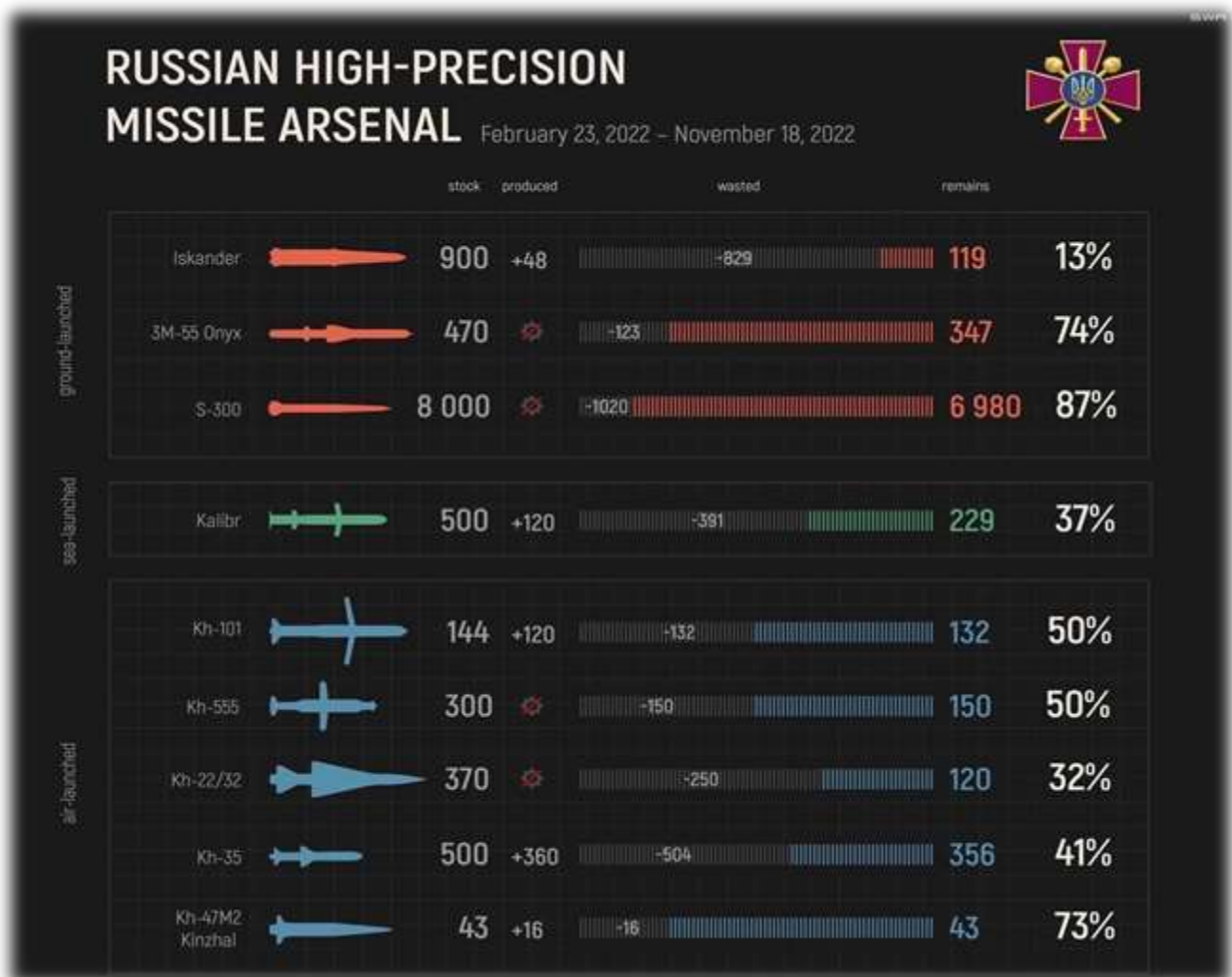


Рис. 7.1. Кількість ракет за типами, яка залишилася у рф з моменту повномасштабного вторгнення.

Цікава інфографіка, якісна і змістовна. Проте після перепосту на одному із інформаційних джерел під цією інфографікою з'явився цікавий висновок: "Таким чином, в російській федерації залишилося найбільше ракет ЗМ-55 "Онікс" – 74%, S-300 – 87% та Х47-М2 "Кінжал" – 73%". У цьому випадку було грубо застосовано відносні одиниці вимірювання, тобто відсотки. Але ж ерефія обстрілює територію України не відсотками, а реальними ракетами!

Зсув шкали. Щоб продемонструвати значні відмінності там, де їх немає, дослідники ніби «зсовують» шкали, починаючи відлік не з нуля, а зі зручнішого для них числа. Хоча лінійні розміри діаграми залишаться незмінними, але інтервали між значеннями шкал зростуть, що зробить криву більш крутою та візуально ніби збільшить різницю між окремими її значеннями.

Приховування даних. Якщо мета полягає у приховуванні значних відмінностей в даних, їх розміщують на одній шкалі з іншими даними, які на порядок відрізняються. На їх фоні будь-які відмінності або зміни будуть сприйняті як незначні.

Зміна масштабів. М'якшим варіантом створення ілюзії значущості є зміна масштабів шкал. Залежно від масштабу ті самі дані матимуть різний вигляд.

Таким чином, треба бути дуже обережним, інтерпретуючи дані у вигляді графіків та діаграм. Набагато складніше маніпулювати даними, наведеними в табличних формах, однак і тут можна використати деякі хитрощі, які можуть ввести в оману непідготовлену людину.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Авраменко В.І. Теорія ймовірностей і математична статистика : навч. посіб. 2-ге вид., перероб. і доп. Дніпродзержинськ : ДДТУ, 2013. 245 с.
2. Бабенко В.В. Основи теорії ймовірностей і статистичні методи обробки даних у психологічних і педагогічних експериментах. Львів: Видавничий центр ЛНУ імені Івана Франка, 2006. 168 с.
3. Барковський В.В., Барковська Н.В., Лопатін О.К. Теорія ймовірностей та математична статистика. 5-те вид. Київ: Центр учбової літератури, 2010. 424 с.
4. Бахрушин В.Є. Методи аналізу даних : навч. посіб. для студентів. Запоріжжя : КПУ, 2011. 268 с.
5. Боснюк В.Ф. Математичні методи в психології : курс лекцій. Харків : НУЦЗУ, 2016. 56 с.
6. Браун П. Редігер Г., Макденіел М. Засіло в голові. Наука успішного навчання. Київ : Наш формат, 2019. 240 с.
7. Важинський С.Е., Щербак Т.І. Методика та організація наукових досліджень : навч. посіб. Суми : СумДПУ, 2016. 260 с.
8. Валєєв К.Г., Джалладова І.А. Збірник задач з теорії ймовірностей та математичної статистики: навч. посіб. Київ : КНЕУ, 2006. 352 с.
9. Вдовенко В.В. Математичні методи в психології: Навчально-методичний посібник. Кіровоград: ПП «Авангард», 2017. 112 с.
10. Вітченко А.О. Теорія і методика наукових досліджень у вищій військовій школі: підруч. Київ : НУОУ, 2020. 268 с.
11. Галян О.І. Експериментальна психологія: навчальний посібник. Київ : Академвидав, 2012. 400 с.
12. Галян І.М. Психодіагностика. Навч. посіб. 2-ге видання, Київ : Академвидав, 2011. 464 с.
13. Грабовецький Б.Є. Методи експертних оцінок: теорія, методологія, напрямки використання : монографія. Вінниця : ВНТУ, 2010. 171 с.

14. Горонескуль М.М. Таблиці функцій і критичних точок розподілів. Теорія ймовірностей. Математична статистика. Математичні методи у психології. Харків : УЦЗУ, 2009. 90 с.

15. Грицюк П.М., Остапчук О.П. Аналіз даних: навч. посіб. Рівне : НУВГП, 2008. 218 с.

16. Диховичний О.О., Круглова Н.В., Москальов І.О. Використання математичних моделей для аналізу результатів психологічного тесту Гілфорда. Mathematics in Modern Technical University. Київ : КПІ, 2018. С. 75–89.

17. Еленберг Дж. Як ніколи не помилятися. Сила математичного мислення. Київ : Наш Формат, 2017. 576 с.

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

19. Костюк В.О. Статистика: навч. посібник. Харків : ХНУМГ ім. О.М. Бекетова, 2015. 166 с.

20. Лиходєєва Г.В. Комп'ютерний практикум з математичної статистики : навч. посіб. Київ : Центр учбової літератури, 2018. 98 с.

21. Мальська М.П. Організація наукових досліджень : навч. посіб. Київ : «Центр учбової літератури», 2017. 136 с.

22. Математичні методи в психології. Курс відео лекцій. URL: <https://www.youtube.com/channel/UCu7Tj9z4NnyenGXmkvMGEEnA/videos>

23. Опря А.Т. Статистика. Навч. посіб. Київ : Центр учбової літератури, 2012. 448 с.

24. Партико З.В. Основи наукових досліджень: підготовка дисертації: навч. посіб. 2-ге вид., перероб. і доп. Київ : Ліра-К, 2018. 232 с.

25. Прикладна математична статистика для економістів : навч. посіб. /Копич І. М., Копитко Б.І., Сороківський В.М., Бабенко В.В., Стефаняк В. І. Львів : Новий Світ 2000, 2012. 408 с.

26. Приставка П.О. Аналіз даних. Навч. посіб. МОН України. Донецьк : РВВ ДНУ, 2008. 92 с.

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

28. Руденко В.М., Руденко Н.М. Математичні методи в психології : підручник. Київ : Академвидав, 2009. 384 с.
29. Руська Р.В.: Теорія імовірності та математична статистика в психології: навч. посіб. Тернопіль, 2020. 112 с.
30. Сердюк А.С., Сердюк Л.З. Методи математичної статистики у психології. Київ : КВГІ, 1998. 76 с.
31. Статистичний аналіз даних вимірювань: навч. посіб. / Єременко В.С., Куц Ю.В., Мокійчук В.М., Самойліченко О.В. Київ : НАУ, 2013. 320 с.
32. Сидоренко Е.В. Методы математической статистики в психологии. СПб: ООО «Речь», 2000. 350 с.
33. Теорія ймовірностей та математична статистика в задачах військового спрямування : навч. посіб. / Сокіл Б. І. та ін. Львів : НАСВ, 2020. 304 с.
34. Ягупов В.В. Теорія і методологія військово-педагогічних досліджень: підручник. Київ : НУОУ ім. І. Черняхівського, 2019. 444 с.

ДОДАТКИ

Додаток 1. Таблиці критичних значень

Таблиця 1.

Критичні значення коефіцієнта асиметрії (A_x)

Об'єм 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

Таблиця 2.

Критичні значення коефіцієнта ексцесу (E_x)

Об'єм 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

Таблиця 3.

Критичні значення d_{max} , що відповідають рівням статистичної значущості $p \leq 0,05$ і $p \leq 0,01$ у порівнянні емпіричного розподілу з теоретичним

Відмінності між розподілами можуть вважатися достовірними, якщо абсолютна величина максимальної різниці $d_{емп}$ досягає або перевищує $d_{0,05}$ і тим більше достовірними, якщо $d_{емп}$ досягає або перевищує $d_{0,01}$

n	Максимальний модуль різниці накопичених частот d_{max}	
	0,05	0,01
	5	0,6074
10	0,4295	0,5147
15	0,3507	0,4202
20	0,3037	0,3639
25	0,2716	0,3255
30	0,2480	0,2972
40	0,2147	0,2574

n	Максимальний модуль різниці накопичених частот d_{max}	
	0,05	0,01
	50	0,1921
60	0,1753	0,2101
70	0,1623	0,1945
80	0,1518	0,1820
90	0,1432	-
100	0,1358	-
>100	$1,36/\sqrt{n}$	$1,63/\sqrt{n}$

Таблиця 4.

Критерій λ -Колмогорова-Смірнова для зіставлення емпіричного розподілу з теоретичним (при $n > 50$) або двох емпіричних розподілів між собою (при $n > 50$): рівні статистичної значимості різних значень $\lambda_{емп}$

За отриманим значенням $\lambda_{емп}$ визначається рівень значущості відмінностей між двома розподілами

λ	λ , останній десятковий знак									
	0	1	2	3	4	5	6	7	8	9
	p - десяткові знаки ("0," не зазначається)									
0,3	99999	99998	99995	99991	99983	99970	99949	99917	99872	99807
0,4	99719	99603	99452	99262	99027	98741	98400	97998	97532	96998
0,5	96394	95719	94969	94147	93250	92282	91242	90134	88960	87724
0,6	86428	85077	83678	82225	80732	79201	77636	76042	74422	72781
0,7	71124	69453	67774	66089	64402	62717	61036	59363	57700	56050
0,8	54414	52796	51197	49619	48063	46532	45026	43545	42093	40668
0,9	39273	37907	36571	35266	33992	32748	31536	30356	29206	28087
1,0	27000	25943	24917	23922	22957	22021	21114	20236	19387	18566
1,1	17772	17005	16264	15550	14861	14196	13556	12939	12345	11774
1,2	11225	10697	10190	09703	09235	08787	08357	07944	07550	07171
1,3	06809	06463	06132	05815	05513	05224	04949	04686	04435	04196
1,4	03968	03751	03545	03348	03162	02984	02815	02655	02503	02359
1,5	02222	02092	01969	01852	01742	01638	01539	01446	01357	01274
1,6	01195	01121	01051	00985	00922	00864	00808	00756	00707	00661
1,7	00618	00577	00539	00503	00469	00438	00408	00380	00354	00330
1,6	00307	00285	00265	00247	00229	00213	00198	00186	00170	00158
1,9	00146	00136	00126	00116	00108	00100	00092	00085	00079	00073
2,0	00067	00062	00057	00053	00048	00045	00041	00038	00035	00032
2,1	00030	00027	00025	00023	00021	00019	00018	00016	00015	00014
2,2	00013	00011	00010	00010	00009	00008	00007	00007	00006	00006
2,3	00005	00005	00004	00004	00004	00003	00003	00003	00002	00002
2,4	00002	00002	00002	00001	00001	00001	00001	00001	00001	00001

Таблиця 5.

**Критичні значення G-критерій знаків
для рівнів статистичної значущості $p \leq 0,05$ та $p \leq 0,01$**

Переважання «типового» зсуву є достовірним, якщо G_{emp} нижче або дорівнює $G_{0,05}$ і тим більше достовірними, якщо G_{emp} нижче або дорівнює $G_{0,01}$.

<i>n</i>	<i>p</i>	
	0,05	0,01
5	0	-
6	0	-
7	0	0
8	1	0
9	1	0
10	1	0
11	2	1
12	2	1
13	3	1
14	3	2
15	3	2
16	4	2
17	4	3
18	5	3
19	5	4
20	5	4
21	6	4
22	6	5
23	7	5
24	7	5
25	7	6
26	8	6

<i>n</i>	<i>p</i>	
	0,05	0,01
27	8	7
28	8	7
29	9	7
30	10	8
31	10	8
32	10	8
33	11	9
34	11	9
35	12	10
36	12	10
37	13	10
38	13	11
39	13	11
40	14	12
41	14	12
42	15	13
43	15	13
44	16	13
45	16	14
46	16	14
47	17	15
48	17	15

<i>n</i>	<i>p</i>	
	0,05	0,01
49	18	15
50	18	16
52	19	17
54	20	18
56	21	18
58	22	19
60	23	20
62	24	21
64	24	22
66	25	23
68	26	23
70	27	24
72	28	25
74	29	26
76	30	27
78	31	28
80	32	29
82	33	30
84	33	30
86	34	31
88	35	32
90	36	33

<i>n</i>	<i>p</i>	
	0,05	0,01
92	37	34
94	38	35
96	39	36
98	40	37
100	41	37
110	45	42
120	50	46
130	55	51
140	59	55
150	64	60
160	69	64
170	73	69
180	78	73
190	83	78
200	87	83
220	97	92
240	106	101
260	116	110
280	125	120
300	135	129

Таблиця 6.

**Критичні значення T -критерію Вілкоксона для рівнів
статистичної значущості $p \leq 0,05$ та $p \leq 0,01$**

«Типовий» зсув є достовірно переважним за інтенсивністю, якщо $T_{емп}$ нижче або дорівнює $T_{0,05}$, і тим більше достовірно переважним, якщо $T_{емп}$ нижче або дорівнює $T_{0,01}$.

<i>n</i>	<i>P</i>	
	<i>0,05</i>	<i>0,01</i>
5	0	-
6	2	-
7	3	0
8	5	1
9	8	3
10	10	5
11	13	7
12	17	9
13	21	12
14	25	15
15	30	19
16	35	23
17	41	27
18	47	32
19	53	37
20	60	43
21	67	49
22	75	55
23	83	62
24	91	69
25	100	76
26	110	84
27	119	92

<i>n</i>	<i>P</i>	
	<i>0,05</i>	<i>0,01</i>
28	130	101
29	140	110
30	151	120
31	163	130
32	175	140
33	187	151
34	200	162
35	213	173
36	227	185
37	241	198
38	256	211
39	271	224
40	286	238
41	302	252
42	319	266
43	336	281
44	353	296
45	371	312
46	389	328
47	407	345
48	426	362
49	446	379
50	466	397

Таблиця 7.

**Критичні значення Q-критерій Розенбауму
для рівнів статистичної значущості $p \leq 0,05$ та $p \leq 0,01$**

Відмінності між двома вибірками можна вважати достовірними ($p \leq 0,05$), якщо $Q_{емп}$ дорівнює або вище критичного значення $Q_{0,05}$, і тим більше достовірними ($p \leq 0,01$), якщо $Q_{емп}$ дорівнює або вище критичного значення $Q_{0,01}$

<i>n</i>	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
<i>p = 0,05</i>																
11	6															
12	6	6														
13	6	6	6													
14	7	7	6	6												
15	7	7	6	6	6											
16	8	7	7	7	6	6										
17	7	7	7	7	7	7	7									
18	7	7	7	7	7	7	7	7								
19	7	7	7	7	7	7	7	7	7							
20	7	7	7	7	7	7	7	7	7	7						
21	8	7	7	7	7	7	7	7	7	7	7					
22	8	7	7	7	7	7	7	7	7	7	7	7				
23	8	8	7	7	7	7	7	7	7	7	7	7	7			
24	8	8	8	8	8	8	8	8	8	8	7	7	7	7		
25	8	8	8	8	8	8	8	8	8	7	7	7	7	7	7	
26	8	8	8	8	8	8	8	8	8	8	7	7	7	7	7	7
<i>p = 0,01</i>																
11	9															
12	9	9														
13	9	9	9													
14	9	9	9	9												
15	9	9	9	9	9											
16	9	9	9	9	9	9										
17	10	9	9	9	9	9	9									
18	10	10	9	9	9	9	9	9								
19	10	10	10	9	9	9	9	9	9							
20	10	10	10	10	9	9	9	9	9	9						
21	11	10	10	10	9	9	9	9	9	9	9					
22	11	11	10	10	10	9	9	9	9	9	9	9				
23	11	11	10	10	10	10	9	9	9	9	9	9	9			
24	12	11	11	10	10	10	10	9	9	9	9	9	9	9		
25	12	11	11	10	10	10	10	10	9	9	9	9	9	9	9	
26	12	12	11	11	10	10	10	10	10	9	9	9	9	9	9	9

Таблиця 8.

**Критичні значення U -критерій Манна-Уїтні для рівнів
статистичної значимості $p \leq 0,05$ та $p \leq 0,01$**

Відмінності між двома вибірками можна вважати значущими ($p < 0,05$), якщо $U_{емп}$ нижче або дорівнює $U_{0,05}$ і тим більше достовірними ($p < 0,01$), якщо $U_{емп}$ нижче або дорівнює $U_{0,01}$

n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_2	$p = 0,05$																		
3	–	0																	
4	–	0	1																
5	0	1	2	4															
6	0	2	3	5	7														
7	0	2	4	6	8	11													
8	1	3	5	8	10	13	15												
9	1	4	6	9	12	15	18	21											
10	1	4	7	11	14	17	20	24	27										
11	1	5	8	12	16	19	23	27	31	34									
12	2	5	9	13	17	21	26	30	34	38	42								
13	2	6	10	15	19	24	28	33	37	42	47	51							
14	3	7	11	16	21	26	31	36	41	46	51	56	61						
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72					
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96			
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	$p = 0,01$																		
5	–	–	0	1															
6	–	–	1	2	3														
7	–	0	1	3	4	6													
8	–	0	2	4	6	7	9												
9	–	1	3	5	7	9	11	14											
10	–	1	3	6	8	11	13	16	19										
11	–	1	4	7	9	12	15	18	22	25									
12	–	2	5	8	11	14	17	21	24	28	31								
13	0	2	5	9	12	16	20	23	27	31	35	39							
14	0	2	6	10	13	17	22	26	30	34	38	43	47						
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56					
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66				
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77			
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88		
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Продовження таблиці 8

n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
n_2	$p = 0,05$																	
21	19	26	34	41	49	57	65	73	81	89	97	105	113	121	130	138	146	154
22	20	28	36	44	52	60	69	77	85	94	102	111	119	128	136	145	154	162
23	21	29	37	46	55	63	72	81	90	99	107	116	125	134	143	152	161	170
24	22	31	39	48	57	66	75	85	94	103	113	122	131	141	150	160	169	179
25	23	32	41	50	60	69	79	89	98	108	118	128	137	147	157	167	177	187
26	24	33	43	53	62	72	82	93	103	113	123	133	143	154	164	174	185	195
27	25	35	45	55	65	75	86	96	107	118	128	139	150	160	171	182	193	203
28	26	36	47	57	68	79	89	100	111	122	133	144	156	167	178	189	200	212
29	27	38	48	59	70	82	93	104	116	127	139	150	162	173	185	196	208	220
30	28	39	50	62	73	85	96	108	120	132	144	156	168	180	192	204	216	228
31	29	41	52	64	76	88	100	112	124	137	149	161	174	186	199	211	224	236
32	30	42	54	66	78	91	103	116	129	141	154	167	180	193	206	219	232	245
33	31	43	56	68	81	94	107	120	133	146	159	173	186	199	213	226	239	253
34	32	45	58	71	84	97	110	124	137	151	164	178	192	206	219	233	247	261
35	33	46	59	73	86	100	114	128	142	156	170	184	198	212	226	241	255	269
36	35	48	61	75	89	103	117	132	146	160	175	189	204	219	233	248	263	278
37	36	49	63	77	92	106	121	135	150	165	180	195	210	225	240	255	271	286
38	37	51	65	79	94	109	124	139	155	170	185	201	216	232	247	263	278	294
39	38	52	67	82	97	112	128	143	159	175	190	206	222	238	254	270	286	302
40	39	53	69	84	100	115	131	147	163	179	196	212	228	245	261	278	294	311
	$p = 0,01$																	
21	10	16	22	29	35	42	49	56	63	70	77	84	91	98	105	113	120	127
22	10	17	23	30	37	45	52	59	66	74	81	89	96	104	111	119	127	134
23	11	18	25	32	39	47	55	62	70	78	86	94	102	109	117	125	133	141
24	12	19	26	34	42	49	57	66	74	82	90	98	107	115	123	132	140	149
25	12	20	27	35	44	52	60	69	77	86	95	103	112	121	130	138	147	156
26	13	21	29	37	46	54	63	72	81	90	99	108	117	126	136	145	154	163
27	14	22	30	39	48	57	66	75	85	94	103	113	122	132	142	151	161	171
28	14	23	32	41	50	59	69	78	88	98	108	118	128	138	148	158	168	178
29	15	24	33	42	52	62	72	82	92	102	112	123	133	143	154	164	175	185
30	15	25	34	44	54	64	75	85	95	106	117	127	138	149	160	171	182	192
31	16	26	36	46	56	67	77	88	99	110	121	132	143	155	166	177	188	200
32	17	27	37	47	58	69	80	91	103	114	126	137	149	160	172	184	195	207
33	17	28	38	49	60	72	83	95	106	118	130	142	154	166	178	190	202	214
34	18	29	40	51	62	74	86	98	110	122	134	147	159	172	184	197	209	222
35	19	30	41	53	64	77	89	101	114	126	139	152	164	177	190	203	216	229
36	19	31	42	54	67	79	92	104	117	130	143	156	170	183	196	210	223	236
37	20	32	44	56	69	81	95	108	121	134	148	161	175	189	202	216	230	244
38	21	33	45	58	71	84	97	111	125	138	152	166	180	194	208	223	237	251
39	21	34	46	59	73	86	100	114	128	142	157	171	185	200	214	229	244	258
40	22	35	48	61	75	89	103	117	132	146	161	176	191	206	221	236	251	266

Продовження таблиці 8

n_1	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
n_2	$p = 0,05$																		
21																			
22	171																		
23	180	189																	
24	188	198	207																
25	197	207	217	227															
26	206	216	226	237	247														
27	214	225	236	247	258	268													
28	223	234	245	257	268	279	291												
29	232	243	255	267	278	290	302	314											
30	240	252	265	277	289	301	313	326	338										
31	249	261	274	287	299	312	325	337	350	363									
32	258	271	284	297	310	323	336	349	362	375	389								
33	266	280	293	307	320	334	347	361	374	388	402	415							
34	275	289	303	317	331	345	359	373	387	401	415	429	443						
35	284	298	312	327	341	356	370	385	399	413	428	442	457	471					
36	292	307	322	337	352	367	381	396	411	426	441	456	471	486	501				
37	301	316	332	347	362	378	393	408	424	439	454	470	485	501	516	531			
38	310	325	341	357	373	388	404	420	436	452	467	483	499	515	531	547	563		
39	318	335	351	367	383	399	416	432	448	464	481	497	513	530	546	562	579	595	
40	327	344	360	377	394	410	427	444	460	477	494	511	527	544	561	578	594	611	628
	$p = 0,01$																		
21																			
22	142																		
23	150	158																	
24	154	166	174																
25	165	174	183	192															
26	173	182	191	201	210														
27	180	190	200	209	219	229													
28	188	198	208	218	229	239	249												
29	196	206	217	227	238	249	259	270											
30	203	214	225	236	247	258	270	281	292										
31	211	223	234	245	257	268	280	291	303	314									
32	219	231	242	254	266	278	290	302	314	326	338								
33	227	239	251	263	276	288	300	313	325	337	350	362							
34	234	247	260	272	285	298	311	323	336	349	362	375	387						
35	242	255	268	281	294	308	321	334	347	360	374	387	400	413					
36	250	263	277	290	304	318	331	345	358	372	386	399	413	427	440				
37	258	271	285	299	313	327	341	355	370	384	398	412	426	440	454	468			
38	265	280	294	308	323	337	352	366	381	395	410	424	439	453	468	482	497		
39	273	288	303	317	332	347	362	377	392	407	422	437	452	467	482	497	512	527	
40	281	296	311	326	342	357	372	388	403	418	434	449	465	480	495	511	526	542	557

Продовження таблиці 8

n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
n_2	$p = 0,05$																	
41	40	55	70	86	102	118	135	151	168	184	201	218	234	251	268	285	302	319
42	41	56	72	88	105	121	138	155	172	189	206	223	240	258	275	292	310	327
43	42	58	74	91	107	124	142	159	176	194	211	229	247	264	282	300	318	335
44	43	59	76	93	110	128	145	163	181	199	216	235	253	271	289	307	325	344
45	44	61	78	95	113	131	149	167	185	203	222	240	259	277	296	315	333	352
46	45	62	80	97	115	134	152	171	189	208	227	246	265	284	303	322	341	360
47	46	64	81	100	118	137	156	175	194	213	232	251	271	290	310	329	349	369
48	47	65	83	102	121	140	159	178	198	218	237	257	277	297	317	337	357	377
49	48	66	85	104	123	143	163	182	202	222	243	263	283	303	324	344	365	385
50	49	68	87	106	126	146	166	186	207	227	248	268	289	310	331	352	372	393
51	50	69	89	109	129	149	170	190	211	232	253	274	295	316	338	359	380	402
52	51	71	91	111	131	152	173	194	215	237	258	280	301	323	345	366	388	410
53	52	72	92	113	134	155	177	198	220	241	263	285	307	329	352	374	396	418
54	53	74	94	115	137	158	180	202	224	246	269	291	313	336	359	381	404	427
55	54	75	96	118	139	161	184	206	228	251	274	297	319	342	365	389	412	435
56	55	76	98	120	142	164	187	210	233	256	279	302	326	349	372	396	420	443
57	57	78	100	122	145	167	191	214	237	261	284	308	332	355	379	403	427	451
58	58	79	102	124	147	171	194	218	241	265	289	314	338	362	386	411	435	460
59	59	81	103	127	150	174	198	222	246	270	295	319	344	369	393	418	443	468
60	60	82	105	129	153	177	201	225	250	275	300	325	350	375	400	426	451	476
	$p = 0,05$																	
41	23	36	49	63	77	91	106	121	136	151	166	181	196	211	227	242	258	273
42	23	37	50	65	79	94	109	124	139	155	170	186	201	217	233	249	265	280
43	24	38	52	66	81	96	112	127	143	159	175	190	207	223	239	255	271	288
44	25	39	53	68	83	99	115	130	146	163	179	195	212	228	245	262	278	295
45	25	40	54	70	85	101	117	134	150	167	183	200	217	234	251	268	285	303
46	26	41	56	71	87	104	120	137	154	171	188	205	222	240	257	275	292	310
47	27	42	57	73	90	106	123	140	157	175	192	210	228	245	263	281	299	317
48	27	43	58	75	92	109	126	143	161	179	197	215	233	251	269	288	306	325
49	28	44	60	77	94	111	129	147	165	183	201	220	238	257	276	294	313	332
50	29	45	61	78	96	114	132	150	168	187	206	225	244	263	282	301	320	339
51	29	46	63	80	98	116	135	153	172	191	210	229	249	268	288	307	327	347
52	30	47	64	82	100	119	137	157	176	195	215	234	254	274	294	314	334	354
53	31	48	65	83	102	121	140	160	179	199	219	239	259	280	300	320	341	361
54	31	49	67	85	104	114	143	163	183	203	224	244	265	285	306	327	348	369
55	32	50	68	87	106	126	146	166	187	207	228	249	270	291	312	333	355	376
56	33	51	69	89	108	129	149	177	190	211	233	254	275	297	318	340	362	384
57	33	52	71	90	111	131	152	173	194	215	237	259	281	302	324	347	369	391
58	34	53	72	92	113	133	155	176	198	220	242	264	286	308	331	353	376	398
59	34	54	73	94	115	136	158	179	201	224	246	268	291	314	337	360	383	406
60	35	55	75	96	117	138	160	183	205	228	250	273	296	320	343	366	390	413

Продовження таблиці 8

n_1	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
n_2	$p = 0,05$																		
41	336	353	370	387	404	421	438	456	473	490	507	524	541	559	576	593	610	628	645
42	345	362	380	397	415	432	450	467	485	503	520	538	556	573	591	609	626	644	662
43	353	371	389	407	425	443	461	479	497	515	533	552	570	588	606	624	642	660	679
44	362	380	399	417	436	454	473	491	510	528	547	565	584	602	621	640	658	677	695
45	371	390	408	427	446	465	484	503	522	541	560	579	598	617	636	655	674	693	712
46	380	399	418	437	457	476	495	515	534	554	573	593	612	631	651	670	690	709	729
47	388	408	428	447	467	487	507	527	547	566	586	606	626	646	666	686	706	726	746
48	397	417	437	458	478	498	518	539	559	579	600	620	640	661	681	701	722	742	763
49	406	426	447	468	488	509	530	550	571	592	613	634	654	675	696	717	738	759	780
50	414	435	457	478	499	520	541	562	583	605	626	647	669	690	711	732	754	775	796
51	423	445	466	488	509	531	553	574	596	618	639	661	683	704	726	748	770	791	813
52	432	454	476	498	520	542	564	586	608	630	652	675	697	719	741	763	786	808	830
53	441	463	485	508	530	553	575	598	620	643	666	688	711	734	756	779	802	824	847
54	449	472	495	518	541	564	587	610	633	656	679	702	725	748	771	794	818	841	864
55	458	481	505	528	551	575	598	622	645	669	692	716	739	763	786	810	834	857	881
56	467	491	514	538	562	586	610	634	657	681	705	729	753	777	801	825	850	874	898
57	476	500	524	548	572	597	621	645	670	694	719	743	768	792	816	841	865	890	915
58	484	509	534	558	583	608	633	657	682	707	732	757	782	807	832	856	881	906	931
59	493	518	543	568	594	619	644	669	694	720	745	770	796	821	847	872	897	923	948
60	502	527	553	578	604	630	655	681	707	733	758	784	810	836	862	888	913	939	965
	$p = 0,01$																		
41	289	304	320	336	351	367	383	398	414	430	446	462	477	493	509	525	541	557	573
42	296	312	328	345	361	377	393	409	425	442	458	474	490	507	523	539	556	572	588
43	304	321	337	354	370	387	403	420	437	453	470	487	503	520	537	553	570	587	604
44	312	329	346	363	380	397	414	431	448	465	482	499	516	533	550	568	585	602	619
45	320	337	354	372	389	407	424	441	459	476	494	511	529	547	564	582	599	617	635
46	328	345	363	381	399	416	434	452	470	488	506	524	542	560	578	596	614	632	650
47	335	353	372	390	408	426	445	463	481	500	518	536	555	573	592	610	629	647	666
48	343	362	380	399	418	436	455	474	492	511	530	549	568	587	606	625	643	662	681
49	351	370	389	408	427	446	465	484	504	523	542	561	581	600	619	639	658	678	697
50	359	378	398	417	437	456	476	495	515	535	554	574	594	613	633	653	673	693	713
51	366	386	406	526	446	466	486	506	526	546	566	587	607	627	647	667	688	708	728
52	374	395	415	435	456	476	496	517	537	558	578	599	620	640	661	682	702	723	744
53	382	403	423	444	465	486	507	528	549	570	591	612	633	654	675	696	717	738	759
54	390	411	432	453	475	496	517	538	560	581	603	624	646	667	689	710	732	753	775
55	398	419	441	462	484	506	527	549	571	593	615	637	659	680	702	724	746	768	790
56	405	427	449	471	494	516	538	560	582	605	627	649	679	694	716	738	761	784	806
57	413	436	458	581	503	526	548	571	593	616	639	662	684	707	730	753	776	799	822
58	421	444	467	490	513	536	559	582	605	628	651	674	697	721	744	767	790	814	837
59	429	452	475	499	522	545	569	592	616	640	663	687	710	734	758	781	805	829	853
60	437	460	484	508	532	555	579	603	627	651	675	699	723	747	772	796	820	844	868

Закінчення таблиці 8

n_1	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
n_2	$p = 0,05$																			
41	662																			
42	679	697																		
43	697	715	733																	
44	714	733	751	770																
45	731	750	769	789	808															
46	749	768	788	807	827	846														
47	766	786	806	826	846	866	886													
48	783	804	824	845	865	886	906	927												
49	800	821	842	863	884	905	926	947	968											
50	818	839	861	882	903	925	946	968	989	1010										
51	835	857	879	901	922	944	966	988	1010	1032	1054									
52	852	875	897	919	942	964	986	1009	1031	1053	1076	1098								
53	870	893	915	938	961	934	1006	1029	1052	1075	1098	1120	1143							
54	887	910	934	957	980	1003	1026	1050	1073	1096	1119	1143	1166	1189						
55	901	928	952	975	999	1023	1046	1070	1094	1113	1141	1165	1189	1213	1236					
56	922	946	970	994	1018	1042	1067	1091	1115	1139	1163	1187	1212	1236	1260	1284				
57	939	964	988	1013	1037	1062	1087	1111	1136	1161	1185	1210	1235	1259	1284	1309	1333			
58	956	981	1007	1032	1057	1082	1107	1132	1157	1182	1207	1232	1257	1283	1308	1333	1358	1383		
59	974	999	1025	1050	1076	1101	1127	1152	1178	1204	1229	1255	1280	1306	1331	1357	1383	1408	1434	
60	991	1017	1043	1069	1095	1121	1147	1173	1199	1225	1251	1277	1303	1329	1355	1381	1407	1433	1460	1486
	$p = 0,01$																			
41	589																			
42	605	621																		
43	621	637	654																	
44	636	654	671	688																
45	652	670	688	706	723															
46	668	687	705	723	741	759														
47	684	703	722	740	759	777	796													
48	700	719	738	757	776	795	814	834												
49	716	736	755	775	794	814	833	853	872											
50	732	752	772	792	812	832	852	872	892	912										
51	748	769	789	809	830	850	870	891	911	932	952									
52	764	785	806	827	847	868	889	910	931	951	972	993								
53	780	802	823	844	865	886	908	929	950	971	993	1014	1035							
54	796	818	840	861	883	905	926	948	970	991	1013	1035	1057	1078						
55	812	834	857	879	901	923	945	967	989	1011	1034	1056	1078	1100	1122					
56	828	851	873	896	919	941	964	986	1009	1031	1054	1077	1099	1122	1145	1167				
57	844	867	890	913	936	959	982	1005	1028	1051	1074	1098	1121	1141	1167	1190	1213			
58	861	884	907	931	954	978	1001	1024	1048	1071	1095	1118	1142	1165	1189	1213	1236	1260		
59	877	900	924	948	972	996	1020	1044	1068	1091	1115	1139	1163	1187	1211	1235	1259	1283	1307	
60	893	917	941	965	990	1014	1038	1063	1087	1111	1136	1160	1185	1209	1234	1258	1282	1307	1331	1356

**Критичні значення H -критерію Крускала-Уолліса
для різних поєднань n_1, n_2 та n_3**

Відмінності між трьома вибірками можна вважати достовірними на вказаному в таблиці рівні значущості, якщо $H_{емп}$ досягає відповідного критичного значення або перевищує його.

Об'єми вибірок			Значення H	p	Об'єми вибірок			Значення H	p
n_1	n_2	n_3			n_1	n_2	n_3		
2	1	1	2,7000	0,500	4	3	1	5,8333	0,021
2	2	1	3,6000	0,200				5,2083	0,050
2	2	2	4,5714	0,067				5,0000	0,057
3	1	1	3,2000	0,300				4,0556	0,093
3	2	1	4,2857	0,100				3,8889	0,129
			3,8571	0,133	6,4444	0,008			
3	2	2	5,3572	0,029	4	3	2	6,3000	0,011
			4,7143	0,048				5,4444	0,046
			4,5000	0,067				5,4000	0,051
			4,4643	0,105				4,5111	0,098
3	3	1	5,1429	0,043	4	3	3	4,4444	0,102
			4,5714	0,100				6,7455	0,010
			4,0000	0,129				6,7091	0,013
3	3	2	6,2500	0,011	4	3	3	5,7909	0,046
			5,3611	0,032				5,7273	0,050
			5,1389	0,061				4,7091	0,092
			4,5556	0,100				4,7000	0,101
			4,2500	0,121				6,6667	0,010
3	3	3	7,2000	0,004	4	4	1	6,1667	0,022
			6,4889	0,011				4,9667	0,048
			5,6889	0,029				4,8667	0,054
			5,6000	0,050				4,1667	0,082
			5,0667	0,086				4,0667	0,102
			4,6222	0,100				7,0364	0,006
4	1	1	3,5714	0,200	4	4	2	6,8727	0,011
4	2	1	4,8214	0,057				5,4545	0,046
			4,5000	0,076				5,2364	0,052
			4,0179	0,114				4,5545	0,098
4	2	2	6,0000	0,014				4,4455	0,103
			5,3333	0,033					
			5,1250	0,052					
			4,4583	0,100					
			4,1667	0,105					

Об'єми вибірок			Значення H	p	Об'єми вибірок			Значення H	p
n_1	n_2	n_3			n_1	n_2	n_3		
4	4	3	7,1439	0,010	5	4	2	7,2045	0,009
			7,1364	0,011				7,1182	0,010
			5,5985	0,049				5,2727	0,049
			5,5758	0,051				5,2682	0,050
			4,5455	0,099				4,5409	0,098
			4,4773	0,102				4,5182	0,101
4	4	4	7,6538	0,008	5	4	3	7,4449	0,010
			7,5385	0,011				7,3949	0,011
			5,6923	0,049				5,6564	0,049
			5,6538	0,054				5,6308	0,050
			4,6539	0,097				4,5487	0,099
			4,5001	0,104				4,5231	0,103
5	1	1	3,8571	0,143	5	4	4	7,7604	0,009
5	2	1	5,2500	0,036				7,7440	0,011
			5,0000	0,048				5,6571	0,049
			4,4500	0,071				5,6176	0,050
			4,2000	0,095				4,6187	0,100
			4,0500	0,119				4,5527	0,102
5	2	2	6,5333	0,008	5	5	1	7,3091	0,009
5	2	2	6,1333	0,013				6,8364	0,011
			5,1600	0,034				5,1273	0,046
			5,0400	0,056				4,9091	0,053
			4,3733	0,090				4,1091	0,086
			4,2933	0,122				4,0364	0,105
5	3	1	6,4000	0,012	5	5	2	7,3385	0,010
5	3	1	4,9600	0,048				7,2692	0,010
			4,8711	0,052				5,3385	0,047
			4,0178	0,095				5,2462	0,051
			3,8400	0,123				4,6231	0,097
								4,5077	0,100
5	3	2	6,9091	0,009	5	5	3	7,5780	0,010
5	3	2	6,8218	0,010				7,5429	0,010
			5,2509	0,049				5,7055	0,046
			5,1055	0,052				5,6264	0,051
			4,6509	0,091				4,5451	0,100
			4,4945	0,101				4,5363	0,102
5	3	3	7,0788	0,009	5	5	4	7,8229	0,010
5	3	3	6,9818	0,011				7,7914	0,010
			5,6485	0,049				5,6657	0,049
			5,5152	0,051				5,6429	0,050
			4,5333	0,097				4,5229	0,099
			4,4121	0,109				4,5200	0,101
5	4	1	6,9545	0,008	5	5	5	8,0000	0,009
5	4	1	6,8400	0,011				7,9800	0,010
			4,985	0,044				5,7800	0,049
			4,860	0,056				5,6600	0,051
			3,987	0,098				4,5600	0,100
			3,960	0,102				4,5000	0,102

Таблиця 10

Критичні значення критерію χ^2 для рівнів статистичної значущості
 $p \leq 0,05$ та $p \leq 0,01$ при різному числі ступенів свободи ν

Різниця між двома розподілами може вважатися достовірною, якщо $\chi^2_{емп}$ досягає або перевищує $\chi^2_{0,05}$, і тим більше достовірними, якщо $\chi^2_{емп}$ досягає або перевищує $\chi^2_{0,01}$

p			p			p		
ν	0,05	0,01	ν	0,05	0,01	ν	0,05	0,01
1	3,841	6,635	35	49,802	57,342	69	89,391	99,227
2	5,991	9,210	36	50,998	58,619	70	90,631	100,425
3	7,815	11,345	37	52,192	59,892	71	91,670	101,621
4	9,488	13,277	38	53,384	61,162	72	92,808	102,816
5	11,070	15,086	39	54,572	62,428	73	93,945	104,010
6	12,592	16,812	40	55,758	63,691	74	95,081	105,202
7	14,067	18,475	41	56,942	64,950	75	96,217	106,393
8	15,507	20,090	42	58,124	66,206	76	97,351	107,582
9	16,919	21,666	43	59,304	67,459	77	98,484	108,771
10	18,307	23,209	44	60,481	68,709	78	99,617	109,958
11	19,675	24,725	45	61,656	69,957	79	100,749	111,144
12	21,026	26,217	46	62,830	71,201	80	101,879	112,329
13	22,362	27,688	47	64,001	72,443	81	103,010	113,512
14	23,685	29,141	48	65,171	73,683	82	104,139	114,695
15	24,996	30,578	49	66,339	74,919	83	105,267	115,876
16	26,296	32,000	50	67,505	76,154	84	106,395	117,057
17	27,587	33,409	51	68,669	77,386	85	107,522	118,236
18	28,869	34,805	52	69,832	78,616	86	108,648	119,414
19	30,144	36,191	53	70,993	79,843	87	109,773	120,591
20	31,410	37,566	54	72,153	81,069	88	110,898	121,767
21	32,671	38,932	55	73,311	82,292	89	112,022	122,942
22	33,924	40,289	56	74,468	83,513	90	113,145	124,116
23	35,172	41,638	57	75,624	84,733	91	114,268	125,289
24	36,415	42,980	58	76,778	85,950	92	115,390	126,462
25	37,652	44,314	59	77,931	87,166	93	116,511	127,633
26	38,885	45,642	60	79,082	88,379	94	117,632	128,803
27	40,113	46,963	61	80,232	89,591	95	118,752	129,973
28	41,337	48,278	62	81,381	90,802	96	119,871	131,141
29	42,557	49,588	63	82,529	92,010	97	120,990	132,309
30	43,773	50,892	64	83,675	93,217	98	122,108	133,476
31	44,985	52,191	65	84,821	94,422	99	123,225	134,642
32	46,194	53,486	66	85,965	95,626	100	124,342	135,807
33	47,400	54,776	67	87,108	96,828			
34	48,602	56,061	68	88,250	98,028			

Критичні значення коефіцієнта кореляції r_{xy} Пірсона

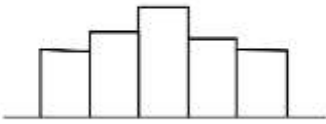


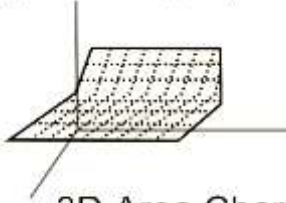
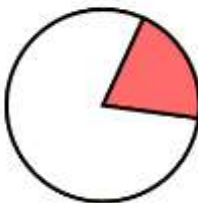
df ($k=n-2$)	p		df ($k=n-2$)	p	
	0,05	0,01		0,05	0,01
5	0,75	0,87	27	0,37	0,47
6	0,71	0,83	28	0,36	0,46
7	0,67	0,80	29	0,36	0,46
8	0,63	0,77	30	0,35	0,45
9	0,60	0,74	35	0,33	0,42
10	0,58	0,71	40	0,30	0,39
11	0,55	0,68	45	0,29	0,37
12	0,53	0,66	50	0,27	0,35
13	0,51	0,64	60	0,25	0,33
14	0,50	0,62	70	0,23	0,30
15	0,48	0,61	80	0,22	0,28
16	0,47	0,59	90	0,21	0,27
17	0,46	0,58	100	0,20	0,25
18	0,44	0,56	125	0,17	0,23
19	0,43	0,55	150	0,16	0,21
20	0,42	0,54	200	0,14	0,18
21	0,41	0,53	300	0,11	0,15
22	0,40	0,52	400	0,10	0,13
23	0,40	0,51	500	0,09	0,12
24	0,39	0,50	700	0,07	0,10
25	0,38	0,49	900	0,06	0,09
26	0,37	0,48	1000	0,06	0,09

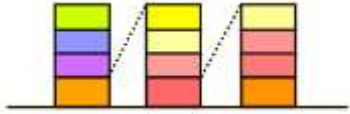
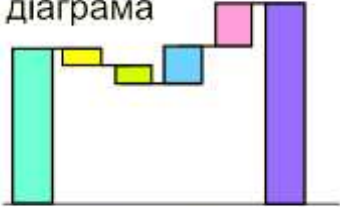
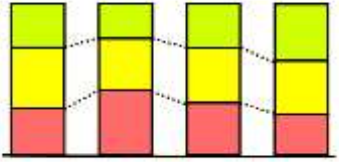
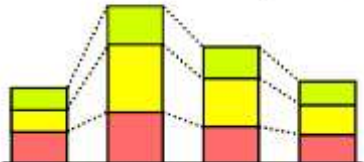
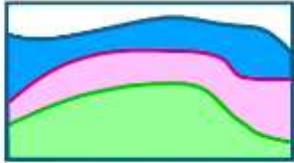
Критичні значення вибіркового коефіцієнта кореляції рангів Спірмена

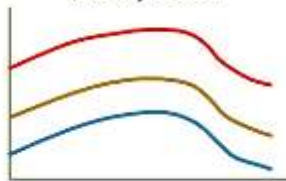

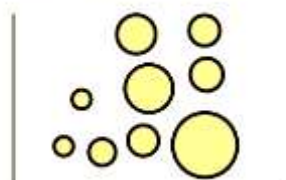
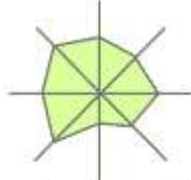

Зв'язок достовірний, якщо $r_{s\text{емп}} \geq r_{s\text{0,05}}$ і тим більше достовірна, якщо $r_{s\text{емп}} \geq r_{s\text{0,01}}$.

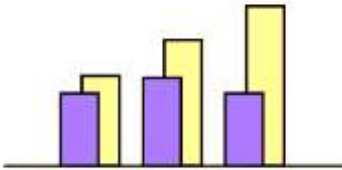

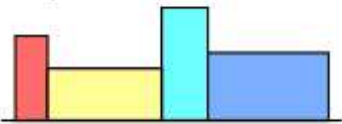

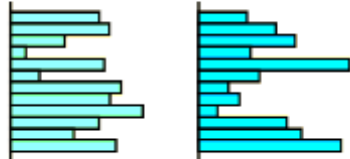
<i>n</i>	<i>p</i>		<i>n</i>	<i>p</i>		<i>n</i>	<i>p</i>	
	0,05	0,01		0,05	0,01		0,05	0,01
5	0,94	-	17	0,48	0,62	29	0,37	0,48
6	0,85	-	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	21	0,36	0,46
8	0,72	0,88	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,54	34	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

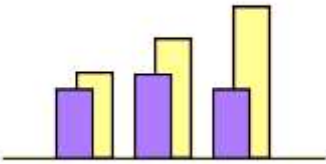
Таблиця вибору техніки візуалізації
(за Джином Желязни)

Що ми хочемо показати?				Техніка візуалізації
1	Розподіл	Одна змінна	Небагато точок даних / категорій / інтервалів	Стовпчикова гістограма  Column Histogram
2	Розподіл	Одна змінна	Багато точок даних	Лінійна гістограма  Line Histogram
3	Розподіл	Дві змінні	Взаємозв'язок	Діаграма розкиду  Scatter Chart
4	Розподіл	Три змінні	Взаємозв'язок	Тривимірна ареальна діаграма  3D Area Chart
5	Композицію (будову)	Статичну	Проста частина в цілому	Кругова діаграма  Pie Chart

6	Композицію (будову)	Статичну	Компоненти компонентів	<p>Стосова 100% стовпчаста діаграма з підкомпонентами</p>  <p>Stacked 100% Column Chart with Subcomponents</p>
7	Композицію (будову)	Статичну	Накопичення додавань і віднімань в цілому	<p>Водоспадна діаграма</p>  <p>Waterfall Chart</p>
8	Зміну протягом часу	Небагато періодів	Важливі тільки відносні відмінності.	<p>Стосова 100% стовпчаста діаграма</p>  <p>Stacked 100% Column Chart</p>
9	Зміну протягом часу	Небагато періодів	Важливі і абсолютні, і відносні відмінності	<p>Стосова стовпчаста діаграма</p>  <p>Stacked Column Chart</p>
10	Зміну протягом часу	Багато періодів	Важливі тільки відносні відмінності	<p>Стосова 100% ареальна діаграма</p>  <p>Stacked 100% Area Chart</p>

11	Зміну протягом часу	Багато періодів	Важливі і абсолютні, і відносні відмінності	<p>Стосова ареальна діаграма</p>  <p>Stacked Area Chart</p>
12	Зв'язок чи залежність	Двох змінних	Взаємозв'язок	<p>Діаграма розкиду</p>  <p>Scatter Chart</p>
13	Зв'язок чи залежність	Трьох змінних	Взаємозв'язок	<p>Бульбашкова діаграма</p>  <p>Bubble Chart</p>
14	Порівняння	В часі	Багато періодів — циклічні дані	<p>Кругова діаграма «Павутиння»</p>  <p>Circular Area Chart</p>
15	Порівняння	В часі	Багато періодів — ациклічні дані	<p>Лінійна діаграма</p>  <p>Line Chart</p>

16	Порівняння	В часі	Мало періодів, одна чи декілька категорій	<p>Стовпчаста діаграма</p>  <p>Column Chart</p>
17	Порівняння	В часі	Мало періодів, багато категорій	<p>Лінійна діаграма</p>  <p>Line Chart</p>
18	Порівняння	Між об'єктами	Дві змінні на об'єкт	<p>Стовпчаста діаграма зі змінною шириною стовпчиків</p>  <p>Variable Width Column Chart</p>
19	Порівняння	Між об'єктами	Одна змінна на об'єкт, багато категорій	<p>Таблиця або таблиця з вбудованими діаграмам</p>  <p>Table or Table with Embedded Charts</p>
20	Порівняння	Між об'єктами	Одна змінна на об'єкт, мало категорій, багато об'єктів	<p>Горизонтальна стовпчаста діаграма</p>  <p>Bar Chart</p>

21	Порівняння	Між об'єктами	Одна змінна на об'єкт, мало категорій, мало об'єктів	Стовпчаста діаграма  Column Chart
----	------------	---------------	--	--

ГЛОСАРІЙ

Абсолютне відхилення – абсолютне значення індивідуального відхилення.

Альтернативна гіпотеза – одна зі статистичних гіпотез, правильність яких перевіряється статистичним критерієм; будується за основною і, як правило, є її логічним доповненням.

Багаторівнева вибірка – вибірка, яку формують, попередньо зобразивши генеральну сукупність у вигляді ієрархічно-групової структури. Далі з усіх груп найвищого рівня вибирають декілька, у кожній з них вибирають декілька груп нижчого рівня і так аж до найнижчого рівня, на якому вибирають елементи генеральної сукупності.

Відносна частота – частота варіанти, виражена у відсотках до загального обсягу сукупності.

Відносне відхилення – відношення абсолютного відхилення до значення відповідної усередненої характеристики.

Варіанта – окреме значення варіаційного ряду.

Варіаційний ряд – впорядкований список усіх можливих значень ознаки, які потенційно можуть зустрічатися в сукупності.

Вибірка зі зміненими пропорціями – вибірка, в якій збільшено пропорцію елементів тих груп генеральної сукупності, властивості яких потрібно вивчити якомога точніше.

Вибірка – сукупність (як правило, невелика) даних, вибраних за допомогою певної процедури з генеральної сукупності. Аналізуючи вибірку, дослідник намагається встановити властивості генеральної сукупності. Основна вимога до вибірки – репрезентативність.

Вибірковий параметр – параметр частотного розподілу вибіркової сукупності.

Вибірковий розподіл – розподіл, отриманий за такою процедурою: з генеральної сукупності беруть всі можливі прості випадкові вибірки заданого розміру і для кожної з цих вибірок обчислюють значення деякого параметра. Розподіл отриманих значень називають вибірковим розподілом цього параметра. Теоретичні результати про властивості

вибіркових розподілів використовують для вмотивування точкових оцінок, а також для визначення надійності інтервальних оцінок. Найчастіше використовують вибіркові розподіли середнього значення, пропорції, різниці середніх значень та різниці пропорцій.

Випадкова вибірка – вибірка, для якої ймовірність бути нерепрезентативною мала, обчислювана і швидко зменшується зі збільшенням обсягу вибірки.

Виправлене стандартне відхилення – незміщена точкова оцінка генерального стандартного відхилення.

Гістограма частот – графічне зображення емпіричного частотного розподілу сукупності даних. Формально гістограму утворюють прямокутники, висоти яких дорівнюють частотам варіант, а самі варіанти лежать всередині їх основ.

Генеральна сукупність – сукупність (як правило, великого обсягу) усіх елементів з певною властивістю. У практичних дослідженнях намагаються отримати якомога точнішу інформацію про властивості генеральної сукупності.

Генеральний параметр – параметр частотного розподілу генеральної сукупності.

Двобічний тест – параметричний тест, основна гіпотеза якого стверджує рівність генерального параметра деякому фіксованому числу.

Дискретна шкала – числова шкала, для кожного значення якої можна вказати його найближчі сусідні значення.

Дисперсія – середнє арифметичне значення квадратів індивідуальних відхилень від середнього значення сукупності.

Дихотомічна шкала – номінальна шкала, яка складається лише з двох значень.

Довірчий інтервал – інтервал для реального значення оцінюваного параметра, який вказується інтервальною оцінкою.

Експериментальна вибірка – вибірка, з якою експериментують з метою порівняння статистичних даних, отриманих до і після експерименту у вибірці.

Ексцес – параметр частотного розподілу, що вказує, які значення переважають у сукупності, близькі чи далекі від середнього.

Емпіричний частотний розподіл – таблиця, сформована варіантами варіаційного ряду сукупності та їх частотами.

Ефективна оцінка – оцінка, стандартна помилка якої є най-меншою у класі незміщених оцінок.

Індивідуальне відхилення – різниця між значенням конкретного елемента сукупності та деякою її усередненою характеристикою.

Інтервальна оцінка – оцінка генерального параметра вигляду “реальне значення генерального параметра належить такому-то інтервалу”. При цьому вказується рівень довіри до інтервальної оцінки.

Інтервальна шкала – числова шкала з відносним нулем. Спосіб компактнішого запису емпіричного частотного розподілу. Таблиця інтервального частотного розподілу складається з неперетинних інтервалів, які вичерпують всі варіанти, та інтервальних частот.

Інтервальна частота – кількість значень сукупності, які потрапляють до інтервалу. При переході до інтервального частотного розподілу частина інформації про вихідну сукупність втрачається.

Кількісна ознака – ознака, для якої природне використання операції підсумовування її значень. Скажімо, природно казати про сукупний прибуток індивідуумів у сукупності.

Кластерна вибірка – багаторівнева вибірка з генеральної сукупності, ієрархічну структуру якої сформовано за географічним принципом.

Коефіцієнт асиметрії – параметр частотного розподілу, який характеризує те, якою мірою елементи сукупності зміщені відносно середнього рівня.

Коефіцієнт варіації – показник у відсотках розсіювання даних, який обчислюється як відношення стандартного відхилення сукупності до її середнього значення.

Коефіцієнт кореляції – показник сили кореляційного зв'язку чи залежності. Найбільше його абсолютне значення одиниця відповідає простому функціональному зв'язку між значеннями досліджуваних ознак,

найменше нуль відсутності або функціонального зв'язку, або зв'язку взагалі.

Контрольна вибірка – вибірка, на яку впливають лише зовнішні щодо експерименту умови. Порівняння результатів, отриманих на експериментальній та контрольній вибірках, дає змогу сформулювати висновок про дієвість експерименту.

Кореляційна залежність – це залежність, яка вказує на вплив однієї з досліджуваних ознак на ймовірність появи різних значень іншої (залежної) ознаки. Кореляційна залежність є свідченням причинно-наслідкового зв'язку.

Кореляційний зв'язок – це взаємозалежність або узгоджена зміна ймовірнісних характеристик кількох ознак. Кореляційний зв'язок не можна вважати свідченням причинно-наслідкового зв'язку.

Кореляція – ймовірнісна або статистична залежність, або зв'язок. На відміну від функціональної залежності кореляція виникає тоді, коли залежність однієї з ознак від іншої ускладнюється наявністю низки випадкових факторів.

Критичні значення – значення, з якими порівнюється емпірична статистика для того, щоб прийняти або відхилити основну гіпотезу.

Кумулята частот – графічне зображення частотного розподілу, яке будують у вигляді ламаної, що послідовно з'єднує точки вигляду (варіанта; накопичена частота).

Лінійна оцінка – точкова оцінка, формула для розрахунку якої лінійна відносно даних вибірки. Умова лінійності свідчить про простоту розрахунку оцінки.

Медіана – усереднений показник, значення якого перевищує одну половину даних сукупності й одночасно менше від іншої.

Міри розсіювання – числові параметри частотного розподілу, які характеризують ступінь розсіювання даних сукупності. У статистиці найважливішою мірою розсіювання вважають стандартне відхилення.

Мода – значення, яке найчастіше зустрічається в сукупності статистичних даних.

Модель прогнозу – модель, в якій зв'язність досліджуваних ознак означає, що реалізоване значення однієї з ознак дає змогу доволі добре передбачити значення іншої. При цьому може розрізнятися напрямленість зв'язку: добре передбачення значення однієї з досліджуваних змінних за значеннями іншої не завжди означає таку саму зворотну передбачуваність.

Накопичена частота – сумарна частота всіх варіант, що не перевищують задану.

Незміщена оцінка – така точкова оцінка, для якої середнє значення параметра, отримане з вибірок однакового обсягу однієї і тієї самої генеральної сукупності, збігається з реальним значенням оцінюваного генерального параметра. Незміщеність одна з найважливіших характеристик точкових оцінок.

Неперервна шкала – числова шкала з властивістю, що між будь-якими двома її значеннями знайдеться інше значення шкали. Сукупності з неперервною шкалою можуть моделюватися неперервними розподілами.

Номінальна шкала – шкала, усі можливі значення якої суть деякі назви, імена. Усі значення номінальної шкали рівноправні.

Обсяг сукупності – кількість елементів сукупності.

Однобічний тест – параметричний тест, основна гіпотеза якого стверджує, що генеральний параметр не перевищує (або не менший) деяке фіксоване число.

Ознака – деяка властивість елементів сукупності. Кожний елемент сукупності має деяке значення (не обов'язково числове) за ознакою. Ці значення породжують нову сукупність, а саме сукупність даних.

Основна гіпотеза – статистична гіпотеза, правильність якої перевіряють статистичним критерієм. Часом її називають нульовою.

Оцінка найбільшої правдоподібності – метод точкового оцінювання параметрів генеральної сукупності.

Полігон частот – графічне зображення частотного розподілу сукупності даних. Формально це ламана, яка послідовно з'єднує точки вигляду (варіанта; частота варіанти).

Порівняння ознак – порівняння кількох генеральних сукупностей за величиною ознаки.

Порівняння розподілів – розпізнавання відмінності між розподілами кількох генеральних сукупностей або між генеральним розподілом і деяким теоретичним.

Порядкова шкала – шкала, значення якої можна впорядкувати за відношенням “менше більше”.

Потужність критерію – здатність статистичного критерію розпізнати правильність альтернативної гіпотези. Формально потужність дорівнює одиниці мінус ймовірність помилки II роду. Помилка II роду полягає у прийнятті неправильної основної гіпотези.

Пропорційна шкала – числова шкала, нульове значення якої абсолютне.

Пропорція – відсоток елементів сукупності з певною властивістю.

Проста випадкова вибірка – вибірка, ймовірність вибору якої з генеральної сукупності дорівнює ймовірності отримати довільну вибірку такого самого обсягу.

Ранжування – процедура, яка допомагає оцифрувати (тобто перетворити на числову) порядкову шкалу. При цьому кожне значення порядкової шкали отримує певний ранг.

Регресія – залежність середнього значення однієї з досліджуваних ознак від значення іншої (інших) ознаки. Розрізняють регресії за їх функціональним виглядом: лінійна, параболічна тощо.

Рівень довіри – показник надійності інтервальної оцінки, зміст якого полягає в такому. Якби було виконано одне й те саме вибіркоче дослідження багаторазово, у результаті було б отримано різні довірчі інтервали. Рівень довіри фактично вказує відсоток правильних з них, тобто відсоток інтервалів, які містять реальне значення оцінюваного генерального параметра.

Рівень значущості – ймовірність помилки I роду при тестуванні статистичної гіпотези. Ця помилка полягає в тому, що статистичний критерій відхиляє правильну основну гіпотезу. Рівень значущості є основною характеристикою надійності тестування. Вважають, що нижчий рівень значущості, то надійніші висновки статистичного критерію.

Рівномірна шкала – спеціальна числова шкала з властивістю, що відстань між будь-якими двома її сусідніми значеннями є однією і тією ж. На практиці рівномірні шкали майже не зустрічаються. Проте є спеціальні процедури рівноміризації нерівномірних шкал. Найчастіше рівномірні шкали застосовують у психології.

Розмах варіації – показник розсіювання даних у сукупності, який обчислюється як різниця між найбільшим та найменшим значеннями в ній.

Розпізнавання зсувів – розпізнавання зміни генеральної сукупності з часом або під зовнішнім впливом.

Розподіл вибіркового середнього – інший термін для вибіркового розподілу середнього значення.

Середнє абсолютне відхилення – середнє значення всіх абсолютних відхилень від середнього арифметичного сукупності.

Середнє значення – найуживаніший усереднений показник, який обчислюється як відношення суми всіх даних до їх кількості.

Систематична вибірка – вибірка за деяким систематичним правилом. Наприклад, можна опитати кожного десятого перехожого на вулиці. Систематичні вибірки часто дають змогу отримати достатню репрезентативність генеральної сукупності. Як правило, їх застосовують тоді, коли випадковий вибір утруднений.

Спроможна оцінка – оцінка, математичне сподівання вибіркового розподілу якої наближається до реального значення оцінюваного генерального параметра зі збільшенням вибірки. Іншими словами, для великих вибірок “змістовність” оцінки означає її “незміщеність”.

Стандартне відхилення – найуживаніший показник розсіювання даних у сукупності, який обчислюється як квадратний корінь з дисперсії.

Статистична гіпотеза – гіпотеза про параметри чи форму розподілу генеральної сукупності. Правильність статистичних гіпотез перевіряють за допомогою статистичних критеріїв. Як правило, перед застосуванням критерію потрібно навести дві гіпотези: нульову (основну) та альтернативну.

Статистичний критерій – правило, яке за даними вибірки дає змогу прийняти чи відхилити статистичну гіпотезу з певною надійністю (рівнем значущості).

Стратифікаційна вибірка – вибірка, яка враховує кількісне співвідношення між обсягами певних частин генеральної сукупності.

Сукупність – початкове поняття статистики, під яким розуміють набір однорідних елементів з подібними властивостями будь-якої природи.

Таблиця спряженості – частотна таблиця, що містить сукупність частот усіх можливих комбінацій значень досліджуваних ознак.

Точкова оцінка – формула, яка за даними будь-якої вибірки породжує деяке значення. Це значення вважають певною оцінкою генерального параметра. Існує низка властивостей точкових оцінок, які характеризують їх надійність та простоту застосування.

Точність інтервальної оцінки – половина ширини довірчого інтервалу. Як правило, вона фіксується наперед. Потрібно вміти знаходити мінімальний розмір вибірки, здатної забезпечити бажану точність. Не варто плутати точність оцінки з її надійністю. Що вища точність дослідження (тобто довірчий інтервал менший, вужчий), то нижча його надійність.

Умовні варіанти – варіанти, застосування яких при обчисленні числових характеристик розподілу дає змогу перейти до простіших (щодо складності обчислень) розподілів.

Усереднений показник – значення, яке характеризує середній рівень даних у сукупності. Усереднений показник не обов'язково є числовим значенням.

Частота варіанти – кількість потраплянь варіанти до заданої сукупності.

Числова шкала – шкала, значення якої є числами.

Шкала ознаки – список усіх можливих значень ознаки. Найуживаніші типи шкал: номінальна, порядкова та числова.

Якісна ознака – ознака, для якої неприродньо застосовувати операцію додавання значень елементів сукупності. Наприклад, неприродньо казати про сумарний зріст індивідумів сукупності.

ДЛЯ НОТАТОК

Навчальний посібник

Москальов І.О., Лисенко Д.П.

ЗАСТОСУВАННЯ МЕТОДІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ
У ПСИХОЛОГО-ПЕДАГОГІЧНИХ ДОСЛІДЖЕННЯХ

Підписано до друку 11.01.23 Формат 60x84 1/16
Обл-вид. арк. 10,488. Друк. арк. 11,5. Тираж 10 прим.
Друкарня НУОУ

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру
видавців, виготівників і розповсюджувачів видавничої продукції,
серія ДК № 2205 від 02.06.2005



Москальов Ігор Олександрович

доктор філософії у галузі соціальних та поведінкових наук зі спеціальності 053 Психологія, старший викладач кафедри суспільних наук гуманітарного інституту Національного університету оборони України, полковник



Лисенко Дмитро Павлович

доктор філософії у галузі соціальних та поведінкових наук зі спеціальності 053 Психологія, старший науковий співробітник науково-дослідної лабораторії кафедри морально-психологічного забезпечення діяльності військ (сил) гуманітарного інституту Національного університету оборони України, підполковник