МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М.Є. Жуковського
“Харківський авіаційний інститут”

Б.А. Смірнов, О.Д. Научитель, Л.І. Костенко

МАТЕМАТИЧНІ МЕТОДИ В ПСИХОЛОГІЇ

Навчальний посібник

Харків “ХАІ” 2007

УДК
ББК Ч 31 +Ю9 Я2
К

Математичні методи в психології / Б.А. Смірнов, О.Д. Научитель, Л.І. Костенко. – Навч. посібник.- Харків: Нац. аерокосм. ун-т “Харк. авіац. ін-т”, 2007. – 57 с.

Описано математичні методи, що використовуються під час проведення психологічного дослідження і обробки його результатів. Наведено загальні відомості про статистичну обробку даних практичного дослідження в галузі психології, показано основні завдання математичної обробки даних. Розглянуто правила використання основних параметричних і непараметричних критеріїв, а також загальні положення про використання методів багатовимірної статистики. Дано загальні рекомендації щодо методів обробки результатів наукового дослідження.

Для студентів денної та заочної форм навчання, що вивчають навчальні дисципліни “Математичні методи в психології”, “Психодіагностика”, “Методика організацій наукових досліджень”, а також для тих, хто використовує методи математичної обробки під час виконання науково-дослідних робіт освітньо-кваліфікаційного рівня.

Іл. 2. Табл. 5. Бібліогр.: 19 назв

Рецензенти: канд.техн. наук, проф. В.Г. Іванов,
канд. психол. наук О.Є. Фальова

1. Загальні поняття про математичну обробку експериментальних даних

Курсові, бакалаврські, дипломні роботи дослідницького характеру, як правило, пов'язані з вимірюванням тих або інших психологічних показників (у психології їх називають також змінними, або ознаками). При цьому під вимірюванням у загальному випадку розуміють приписування чисел показникам (психологічним об'єктам), що вивчаються, відповідно до певних правил. Правила вимірювання встановлюють відповідності між властивостями чисел і властивостями вимірюваних об'єктів. У психології виділяють чотири типи такої відповідності і, отже, чотири типи психологічних вимірювальних шкал.

Найпростішою з них є шкала найменувань (номінальна, номінативна, іменна шкали). У ній об'єкти, що вивчаються, розділяються на класи, що мають різні найменування, назви. Звідси – і назва шкали: nomen латинською, означає ім'я, назва. Номінальна шкала дозволяє нам підраховувати частоту зустрічності різних найменувань і навіть працювати з ними за допомогою математичних методів. Одиницею вимірювання тут є кількість (частота) спостережень (випробуваних, реакцій, виборів тощо). Можуть використовуватися також похідні від частоти показники: частість (відносна частота), вірогідність, відсоток. Простим різновидом номінальної шкали є дихотомічна; вона складається всього з двох альтернативних класів, що відрізняються один від одного за принципом «так»-«ні». Проте за допомогою такої найпростішої шкали вже можна проводити статистичну обробку одержаних результатів (кутове перетворення Фішера, критерій Мак-Німара та ін.)

Другою є шкала порядків (порядкова, рангова, ординальна шкали). За допомогою цієї шкали встановлюють порядок у степеня вираженості психологічної якості, що вивчається, від об'єкта з найбільш вираженою якістю до об'єкта з якнайменше вираженою якістю (або навпаки). Цей порядок нумерується звичайно натуральними числами, або рангами (1, 2, 3, …), за допомогою яких і здійснюється подальша математична обробка. Ці дві шкали відносяться до неметричних, оскільки вони безпосередньо самим об'єктам або явищам чисел не приписують (не мають загальноприйнятої міри й одиниць вимірювання).

Наступні два типи шкал відносяться до метричних (від лат. metrio – вимірюю). Такими є шкала однакових інтервалів (інтервальна шкала) і шкала однакових відносин (пропорційна шкала). (Шкала інтервалів грунтується на рівності різниці степенів вираженості психологічної властивості у двох об'єктів різниці двох чисел, що приписуються цим об'єктам за даною властивістю. Шкала відносин припускає рівність відношення (ділення) степенів вираженості психологічної властивості у двох об'єктів відношенню (діленню) двох чисел, приписаним об'єктам за даною властивістю. Основна відмінність між цими шкалами полягає у тому, що в першій з них (інтервальній) відсутній об'єктивний нуль (початок відліку), а в другій (пропорційній) він є. Прикладом шкали інтервалів є шкала вимірювання температури за Цельсієм, шкала вимірювання інтелекту за допомогою стандартного IQ-показника та ін.; прикладом шкали відносин є шкала вимірювання температури за Кельвіном, шкала вимірювання більшості фізичних величин (часу, маси, довжини тощо). З погляду можливостей застосування цих двох шкал для математичної обробки результатів вимірювань вони практично рівноцінні, і в цьому плані відмінності між ними звичайно не роблять.

Одержані в результаті досліду психологічні показники є, як правило, випадковими величинами, тому їх математична обробка повинна вестися за допомогою статистичних методів. Застосування тих або інших з них визначається тим, до якої з розглянутих шкал відноситься одержаний статистичний матеріал. Кожна зі шкал має строго обмежені можливості для проведення статистичних операцій. Тому перш ніж проводити будь-яку з цих операцій, необхідно визначити, чи має вона сенс для використовуваних числових величин. Можливості різних шкал показано на рис. 1. При користуванні ним слід мати на увазі, що операції, які можна проводити зі шкалами нижчого рівня, застосовуються також до шкал вищих рівнів. Наприклад, критерій згоди χ^2 може застосовуватися до даних, одержаних у номінальній шкалі. Тим більше, він може застосовуватися і при використовуванні порядкової й інтервальної шкал.

Тому в необхідних випадках виконуються перетворення даних з однієї шкали в іншу. Наприклад, результати, одержані в інтервальній шкалі, за необхідності можуть бути перетворені в ранги та переведені в номінальну шкалу. Приклад такого перетворення при дослідженні екстраверсії-інтроверсії шести випробуваних наведено в табл. 1 [10].

Розглянуте перетворення називається зниженням потужності шкали. Проте слід мати на увазі, що при цьому втрачається частина психологічно важливої інформації про випробуваних. У той же час таке перетворення дозволяє за необхідності використовувати менш потужні шкали. Ця обставина дозволяє розширити діапазон

Номінальна шкала	Порядкова шкала	Інтервальна шкала	
Мода			**Міри центральної тенденції і розсіювання**
	Медіана		
	Розмах		
	Дисперсія Стандартне відхилення		
	Квартильне відхилення		
Коефіцієнти чотириклітинної та багатоклітинної спряженості (Пірсона, Чупрова, асоціації, контингентності)			**Міри зв'язку**
	Коефіцієнти рангової кореляції Спірмена і Кендала		
	Коефіцієнт конкордації		
Критерій φ^* (кутове перетворення Фішера)			
Біномний m-критерій			
Критерій χ^2 порівняння розподілів			
Критерій Мак-Німара			
	U-критерій Манна – Уїтні, Т-критерій Вілконсона		**Міри перевірки статистичних гіпотез**
	Непараметричні аналоги дисперсійного аналізу: критерій Крускала – Уолліса, Фрідмана		
	Критерії тенденції Пейджа, Джонкіра		
	Л-критерій порівняння розподілів Колмогорова – Смірнова		
	«Швидкі» методи - критерій дисперсійного аналізу Немені, Лінка, Уолліса		
		t-критерій Стьюдента	
		F-критерій Фішера	
		Дисперсійний аналіз	
		Критерій порівняння n дисперсій Бартлета, Кохрана	
		Критерій грубих помилок спостережень	

Рис. 1. Статистичні операції та відповідні їм числові шкали

(номенклатуру) можливих для застосування статистичних методів. Таке найчастіше зустрічається, коли за умовами досліду немає можливості використовувати параметричні методи, і тому замість них ми вимушені застосовувати непараметричні методи.

Таблиця 1

Показники рівня екстраверсії-інтроверсії для шести випробуваних, що виражені в різних числових шкалах

Випробувані	**Шкала інтервалів**	**Шкала порядків**	**Номінальна шкала**
А	20	5	Е
Б	15	4	Е
В	22	6	Е
Г	9	3	І
Д	3	1	І
Е	4	2	І

Розглянемо це питання докладніше. Методи статистичної обробки діляться на параметричні та непараметричні. Застосування параметричних методів можливе лише при виконанні двох умов: 1) показники, що вимірюються, підпорядковуються нормальному закону розподілу; 2) ці показники виміряні як мінімум в інтервальній шкалі. При використовуванні параметричних методів необхідно знати параметри розподілу, тобто середнє значення та дисперсію. Прикладами таких методів є порівняння середніх за допомогою t-критерію Стьюдента, порівняння дисперсій за допомогою F-критерію Фішера, обчислення коефіцієнта лінійної кореляції Пірсона, дисперсійний аналіз тощо.

На жаль, більшість психологічних змінних не задовольняють умови можливості застосування параметричних методів. Це виявляється в такому. По-перше, багато з цих змінних не підпорядковуються нормальному закону розподілу, а сам закон є, як правило, асиметричним. По-друге, ці змінні звичайно не можуть бути подані в інтервальній шкалі. Рівноінтервальними можна вважати лише шкали в одиницях стандартного відхилення (z-оцінки, стени, стенайни тощо) і процентильні шкали, і лише за умови, що початковий розподіл був нормальним. Решта оцінок, у тому числі й «сирі бали», одержані за результатами тестових випробувань, ці умови не задовольняють [16]. На жаль, ця обставина при обробці дослідних даних часто не враховується, тому застосування параметричних методів призводить в цьому випадку до серйозних помилок.

Тому в психологічних дослідженнях широко повинні застосовуватися непараметричні методи. Вони вільні (не залежать) від виду розподілу, а при їх застосуванні використовуються не самі значення спостережуваних величин, а тільки їх частоти або ранги, тобто величини, не залежні від параметрів розподілу. Це робить можливим використання непараметричних методів при поданні початкових даних у порядковій, а часто навіть у номінальній шкалі. Такі методи вельми зручні для практичного використовування, оскільки потребують мінімального обсягу обчислень і апріорних відомостей і можуть застосовуватися навіть при неможливості прямих вимірювань ознак, що вивчаються. Це є перевагою непараметричних методів, а їх недоліком - менша точність порівняно з параметричними методами.

Таким чином, обидві групи методів мають як свої переваги, так і недоліки. Очевидно, тому навіть відомі спеціалісти з використання математичних методів у психології не мають одностайної думки щодо їх можливостей. Так, Є. В. Сидоренко вважає, що можливість застосування параметричних методів у психології надзвичайно обмежена, оскільки початкові дані, як правило, не задовольняють вимоги параметричності. Тому вона рекомендує використовувати їх непараметричні аналоги [16]. На відміну від цього Г.В. Суходольський вважає, що «… не слід користуватися чисельними непараметричними критеріями, потужність яких мала або не визначена зовсім» [20, с. 221].

Такі протиріччя додають труднощів студентам при виборі потрібного методу. Спільні рекомендацій зводяться тут до такого. В першу чергу слід обирати параметричні методи як більш потужні. Якщо початкові дані не задовольняють умови параметричності, необхідно спробувати провести їх штучну нормалізацію та стандартизацію [10, 20]. Якщо це неможливо виконати, слід застосувати відповідний непараметричний аналог. Однак для збереження тієї ж точності при цьому треба збільшувати об'єм вибірки.

2. Основні задачі математичної обробки даних

При математичній обробці дослідних даних виникає необхідність у вирішенні різних задач. Усі вони умовно можуть бути розбиті на три групи.

І. Статистичний опис одержаних результатів. При цьому використовуються методи так званої *описової статистики*. За їх допомогою здійснюється математичний опис будь-якої вибірки. Частинними задачами тут є такі.

1. **Визначення заходів центральної тенденції**, тобто статистичних показників вибірки, що характеризують найбільш виражене репрезентативне значення змінної. Основні з них: середнє значення, медіана, мода, квантиль і його різновиди (квартилі, квантилі, процентилі).
2. **Визначення заходів мінливості**, тобто статистичних показників, що характеризують розкид значень змінної, що вивчається, щодо міри центральної тенденції. Основні з них: розмах (середнє лінійне відхилення), дисперсія, середнє квадратичне відхилення, коефіцієнт варіації, квартильне відхилення.
3. **Визначення заходів асиметрії і ексцесу**, що визначаються відповідно за допомогою третього і четвертого центрального моментів розподілу випадкової величини. Ці показники самостійного значення при обробці даних не мають, вони необхідні лише для наближеної перевірки дослідних даних відповідно до нормального закону розподілу. Близькість значень асиметрії і ексцесу до нуля може служити підтвердженням відповідності дослідних даних нормальному закону розподілу і, отже, можливості (разом з іншими умовами) застосування параметричних методів.
4. **Побудова законів розподілу випадкової величини**. Закон розподілу є найбільш повною характеристикою випадкової величини. Він показує, які значення і як часто вони зустрічаються в генеральній сукупності, що описується за допомогою даної вибірки.

Розглянуті величини, одержані в результаті математичного опису вибірки, являють собою точкові вибіркові оцінки відповідного генерального параметра, інакше – вони є результатом його вимірювання на основі аналізу даної вибірки. Через обмежений об'єм вибірки таке вимірювання завжди здійснюється з деякою похибкою. Міра незбігу вибіркової оцінки і відповідного їй генерального параметра називається *помилкою репрезентативності,* або статистичною помилкою.

Величина цієї помилки зменшується при збільшенні об'єму вибірки n (кількість вимірювань) та наближається до нуля при $n \to \infty$.

У математичній статистиці отримано формули для визначення помилок репрезентативності, які оцінюються за величиною середньоквадратичного (стандартного) відхилення відповідного вибіркового показника. Наприклад, помилки репрезентативності при визначенні середнього значення $\bar{x}$, середньоквадратичного відхилення σ, дисперсії Д, асиметрії А, ексцесу Е, коефіцієнта кореляції r, вірогідності Р відповідно дорівнюють:

$$\sigma_{\bar{x}} = \sqrt{\frac{Д}{n}},\ \sigma_{\sigma} = \sqrt{\frac{Д}{2n}},\ \sigma_{Д} = \frac{Д}{\sqrt{2n}},\ \sigma_{A} = \sqrt{\frac{6}{n}},\ \sigma_{E} = \sqrt{\frac{24}{n}}.\ \sigma_{r} = \sqrt{\frac{1-r^2}{n-1}},\ \sigma_{P} = \frac{\sqrt{P(1-P)}}{n}.$$

Формули для визначення помилок репрезентативності інших статистичних показників наведено в роботі [18]. З точки зору теорії вимірювань грамотно подати їх результат – це означає отримане значення виміряного показника та його помилку репрезентативності подати у вигляді $\bar{x} \pm \sigma_{\bar{x}}$, $p \pm \sigma_{p}$, $Д \pm \sigma_{Д}$. З наведених формул випливає, що при необмеженому збільшенні n ($n \to \infty$) помилка репрезентативності наближується до нуля, а отримана вибіркова оцінка – до свого генерального параметра.

ІІ. **Визначення взаємозв'язку двох і більше психологічних змінних.** Коли змінні носять вірогідний (випадковий) характер, цей взаємозв'язок називається *кореляцією*. У загальному випадку кореляція характеризується п'ятьма властивостями: спрямованістю (односторонньою або взаємною); тіснотою (силою), яка за величиною може лежати в межах від -1 до +1; формою (лінійною або нелінійною); напрямом (додатним або від’ємним); мірністю (парна або множинна). Кількісна оцінка величини кореляції виконується за допомогою різних коефіцієнтів. Кожний з них враховує ті або інші з перерахованих властивостей.

Найвідомішим і найпоширенішим є коефіцієнт лінійної кореляції Пірсона. Проте він відноситься до параметричних, тому його застосування в психологічних дослідженнях не завжди можливе. Це зумовлено, з одного боку, розглянутими вище обмеженнями параметричних методів, з іншого – необхідністю лінійної залежності між змінними, що вивчаються. Якщо ця залежність нелінійна, слід застосовувати більш загальну міру - коефіцієнт детермінації. Він дозволяє виявити двосторонній взаємозв'язок (х від у і у від х), причому в загальному випадку за величиною він може бути різним у двох змінних, зв'язаних між собою нелінійною залежністю. На жаль, дана обставина в психологічних дослідженнях часто ігнорується, і кореляцію без достатніх на те підстав визначають за допомогою коефіцієнта кореляції Пірсона, який не завжди дозволяє її виявити при нелінійній залежності між змінними.

Через зазначені обставини в психологічних дослідженнях більше уваги слід приділяти непараметричним коефіцієнтам кореляції. Якщо змінні, що вивчаються, подані в порядковій шкалі (у рангах), то вельми ефективним є застосування коефіцієнта рангової кореляції Спірмена. Дещо менш ефективним у цьому випадку є коефіцієнт рангової кореляції Кендала.

Якщо порівнювані змінні подані в номінальній шкалі, то стохастичний зв'язок між ними називають *спряженістю*, іноді – кореляцією якісних ознак [1]. Як заходи, що кількісно характеризують степень спряженості, використовують чотири коефіцієнти. Два з них - коефіцієнт асоціації та коефіцієнт контингентності - використовуються тільки для альтернативного (чотириклітинного) розподілу; їх називають коефіцієнтами чотириклітинної (2Х2: дві змінні, кожна з яких подана двома значеннями) спряженості, або дихотомічними коефіцієнтами кореляції. Іноді як різновид дихотомічної класифікації окремо розглядають випадки, коли обидві дихотомічні змінні добуваються з вдало розподіленої сукупності (наприклад, зріст вище 160 см позначається як 1, нижче 160 см – як 0; відповідні йому значення маси тіла більше 60 кг позначаються 1, менше – 0). Такі дані містять більше інформації порівняно з простою дихотомічною класифікацією і тому дозволяють отримати більш точні значення коефіцієнта спряженості. Як міра зв‘язку таких дихотомічних змінних у даному випадку пропонується використовувати так званий тетрохорічний коефіцієнт кореляції [3]. Два інші коефіцієнти (коефіцієнт спряженості Пірсона і коефіцієнт спряженості Чупрова) використовуються до будь-якої m×n системи подій. Їх називають коефіцієнтами m×n спряженості й використовують у випадку, якщо число порівнюваних ознак більше двох, а самі ознаки мають число градацій більше двох.

Усі розглянуті коефіцієнти лежать у межах від 0 до 1, оцінюють степінь взаємної спряженості, не вказуючи її напряму. Відзначимо також, що ці коефіцієнти є мірою взаємозв'язку двох і більш класифікованих подій, тобто подій, розбитих на певні класи, для кожного з яких визначено частоту її появи. Іншими словами, вимірювані властивості таких психологічних змінних подані в номінальній шкалі за допомогою частоти їх зустрічності в досліджуваній вибірці.

При аналізі взаємозв'язку змінних, виражених різними шкалами, застосовують точково-бісеріальний (коли одна змінна виражена в дихотомічній шкалі найменувань, а інша - в шкалі інтервалів) і рангово-бісеріальний (коли одна змінна виражена в дихотомічній шкалі, а інша - в ранговій шкалі) коефіцієнти кореляції. І, нарешті, можливий випадок, коли одна змінна виражена в ранговій (порядкової) шкалі, а інша - в інтервальній. Показника, що безпосередньо оцінює взаємозв'язок між такими змінними, немає; тому у такому разі інтервальну шкалу слід перетворити в порядкову (приклад такого перетворення наведено у табл.1), а після цього обчислити коефіцієнт рангової кореляції за Спірменом або Кендалом.

Як випливає з розглянутого, при обчисленні взаємозв‘язку випадкових змінних можуть використовуватися різні коефіцієнти кореляції. Їх вибір залежить від ряду обставин: шкали, в якій виміряні ці змінні (інтервальна або неметрична); виду зв‘язку (лінійний або нелінійний); числа випадкових величин, що характеризують розглянуту систему (парна або множинна кореляція). Для полегшення вибору необхідного коефіцієнта кореляції слід використовувати дані табл. 2.

Розширенням і продовженням вивчення взаємозв'язку між двома випадковими змінними, що вимірються в інтервальній шкалі, є регресійний аналіз. Це пов'язано з тим, що розглянутий раніше коефіцієнт лінійної кореляції Пірсона вказує лише на наявність або відсутність зв'язку між змінними, але сам цей зв'язок не описує. Для цього саме і служить регресійний аналіз. Регресія виражається або графічно, або аналітично і показує, як в середньому змінюється показник при зміні якогось чинника, що вивчається (незалежної змінної). Так само, як і кореляція, регресія, може бути або парною, або множинною. У прикладних дослідженнях, як правило, використовується лише лінійна регресія, значно рідше - параболічна. Решта видів регресії в психологічних дослідженнях практично не використовується.

Найважливішим елементом кореляційного та регресійного аналізів є перевірка значущості (відмінності від нуля) одержаних коефіцієнтів. Без цієї перевірки (а її часто, на жаль, забувають зробити) одержані значення коефіцієнтів кореляції не мають особливого сенсу, якщо вони значущо не відрізняються від нуля. Ця перевірка здійснюється за допомогою статистичних критеріїв: для коефіцієнта лінійної кореляції Пірсона, рангової кореляції Спірмена, точково-бісеріальної та рангово-бісеріальної кореляцій, а також коефіцієнтів лінійної регресії - за допомогою t-критерію Стьюдента; для коефіцієнта детермінації - за допомогою F-критерію Фішера, для коефіцієнтів спряженості - за допомогою критерію φ^2 Пірсона. Для оцінки відмінності від нуля коефіцієнта рангової кореляції Кендала розраховується величина Zкритерію, основаного на використанні стандартного нормального відхилення. Процедура перевірки полягає у тому, що емпіричне значення знайденого критерію порівнюється з відповідним табличним значенням (а воно залежить від обраного рівня значущості й об'єму вибірки). Набуте значення коефіцієнта кореляції, що вивчається, може вважатися значущим (відмінним від нуля) тільки у тому випадку, якщо воно за абсолютною величиною перевищує табличне значення. Інакше слід визнати, що набуте емпіричне значення істотно не відрізняється від нуля. Табличні значення зазначених критеріїв наведено в довідковій літературі [4, 10, 16, 18].

Таблиця 2

Зведена таблиця коефіцієнтів парної кореляції

Шкала виміру		**Рекомендований коефіцієнт кореляції**	**Число степенів вільності**
Змінна x	**Змінна y**		
Інтервальна	Інтервальна	Коефіцієнт лінійної кореляції Пірсона (при лінійному зв‘язку між x і y)	df=n-2
Інтервальна	Інтервальна	Коефіцієнт детермінації (при нелінійному зв‘язку)	df=n-2
Порядкова	Порядкова	Коефіцієнти рангової кореляції: • Спірмена • Кендала	df=n df=n
Номінативна	Номінативна	Коефіцієнти спряженості для матриць 2x2: • φ-коефіцієнт Пірсона • Q-коефіцієнт Юла • тетрахоричний коефіцієнт кореляції	df=1 df=1 df=1
Номінативна	Номінативна	Коефіцієнти спряженості для матриць n_1 x n_2, m x n: • К-коефіцієнт Чупрова • С-коефіцієнт Пірсона	df=(n_1-1)(n_2-1) df=(n_1-1)(n_2-1)
Номінативна	Інтервальна	Точково-бісеріальний коефіцієнт	df=2
Номінативна	Порядкова	Рангово-бісеріальний коефіцієнт	df=n-2
Інтервальна	Порядкова	Не розроблений	-

На закінчення слід підкреслити, що при інтерпретації результатів кореляційного аналізу необхідна особлива обережність при обліку статистично значущих високих кореляцій: іноді можуть виникнути хибні кореляції за рахунок того, що обидві змінні, що вивчаються, зазнають сильного впливу інших чинників, що не враховуються в дослідженні.

Тому величина коефіцієнта кореляції несе лише інформацію про наявність зв'язку між змінними, що вивчаються, але не несе інформацію про її причини. Іншими словами, корельовані величини не обов'язково є залежними. Їх причинний зв'язок може бути зумовлений якимсь третім чинником. Тому причинно-наслідкові відносини часто можуть бути виявлені тільки за допомогою спеціально організованого експерименту.

Також слід зазначити, що при проведенні кореляційного аналізу в ряді випадків може виникнути задача перевірки гіпотез про відмінність кореляцій, тобто їх порівняння та визначення схожості/відмінності. Найчастіше висновок при цьому робиться інтуїтивно, без достатнього статистичного обґрунтування, що може призвести до помилкових результатів. Задача порівняння кореляцій має два варіанти розв'язання: для залежних і незалежних вибірок. В обох випадках для її вирішення використовується Z-критерій, оснований на аналізі стандартного нормального розподілу. У випадку незалежних вибірок при цьому треба попередньо провести Z-перетворення Фішера коефіцієнтів кореляції. Докладніше про це говориться в [3, 10].

III. Перевірка статистичних гіпотез (статистичний висновок). Вживані при вирішенні задач цієї групи методи складають зміст так званої *перевірної статистики* і базуються в основному на теорії статистичних рішень. Ці задачі зводяться переважно до порівняння двох і більше вибірок з метою визначення їх незалежності або належності до однієї генеральної сукупності, або, інакше, до перевірки гіпотез про відмінність двох і більш вибірок або експериментальних умов. При вирішенні задач цієї групи слід враховувати характер вибірок, які можуть бути незалежними або зв'язаними (окремим випадком останніх є одна і та ж вибірка, яка досліджується в різних експериментальних умовах). Ці два види вибірок відповідають умовам проведення міжгрупового і внутрішньогрупового дослідження. Крім цього, слід враховувати кількість порівнюваних вибірок (дві або більше двох) та шкалу, у якій виміряні дані, що підлягають порівнянню (інтервальна або порядкова). Усього, таким чином, маємо 2х2х2=8 варіантів порівняння (рис. 2). Вибір правильного варіанта, тобто адекватного статистичного методу, є найважливішою умовою грамотного вирішення задачі перевірки статистичних гіпотез.

Статистичні гіпотези, що перевіряються при вирішенні задач третьої групи, розділяються на нульові й альтернативні (конкуруючі).

Нульова гіпотеза – це гіпотеза про відсутність відмінностей; це те, що ми хочемо спростувати, якщо перед нами стоїть задача встановити значущість відмінностей. *Альтернативна гіпотеза* – це гіпотеза про значущість відмінностей; це те, що ми хочемо довести, якщо нам потрібно встановити значущість відмінностей. Обидва ці види гіпотез можуть бути направленими і ненаправленими.

При статистичній перевірці гіпотез у першому випадку використовуються односторонні, а в другому – двосторонні статистичні критерії (критерії згоди).

<table>
<tr><th>Кількість вибірок</th><th colspan="2">Дві вибірки</th><th colspan="2">Більше двох вибірок</th></tr>
<tr><th>Залежність вибірок</th><th>Незалежні</th><th>Залежні</th><th>Незалежні</th><th>Залежні</th></tr>
<tr><td rowspan="2">Вимірювальна величина, виражена в шкалі інтервалів</td><td colspan="4">Параметричні методи порівняння</td></tr>
<tr><td>t-критерій Стьюдента</td><td>t-критерій Стьюдента для парних порівнянь</td><td>Дисперсійний аналіз</td><td>Дисперсійний аналіз з повторними вимірюваннями</td></tr>
<tr><td rowspan="2">Вимірювальна величина, виражена в порядковій шкалі</td><td colspan="4">Непараметричні методи порівняння</td></tr>
<tr><td>U-критерій Манна – Уїтні, критерій серій</td><td>T-критерій Вілконсона, критерій знаків</td><td>H-критерій Крускала – Уолліса, s-критерій тенденцій Джонкіра</td><td>χ_r^2 -критерій Фрідмана, L-критерій тенденцій Пейджа</td></tr>
</table>

Рис. 2. Класифікація методів статистичного висновку про схожість-відмінність вибірок за рівнем вираженості ознаки, що вивчається

Статистичний критерій - це вирішальне правило, що з певною достовірністю забезпечує прийняття істинної і відхилення помилкової гіпотез. Статистичні критерії відповідно до сказаного раніше діляться на параметричні та непараметричні. Параметричні критерії включають в розрахункову формулу параметри розподілу; їх застосування можливе, якщо змінні, що вивчаються, виміряні в інтервальній шкалі і мають нормальний розподіл. На відміну від цього непараметричні критерії не залежать від розподілу, вони оперують даними, поданими в порядковій та номінативній шкалах. Тому вони мають більшу область застосування, ніж параметричні критерії, проте порівняно з ними непараметричні критерії мають меншу потужність. Виходячи з усього сказаного, слід зробити висновок, що при обробці результатів психологічних досліджень основними слід вважати непараметричні критерії. За їх допомогою можна вирішувати практично всі задачі. Виняток становить лише задача оцінки впливу двох і більше факторів на зміну ознаки, що вивчається. Цю задачу можна вирішити тільки за допомогою дисперсійного двофакторного аналізу. Він є єдиним з параметричних критеріїв, аналога якому в непараметричній статистиці немає.

При перевірці статистичних гіпотез можливі два види помилок. Помилка, що полягає у відхиленні нульової гіпотези, в той час як вона правильна, називається помилкою першого роду (у теорії виявлення сигналів вона називається пропуском сигналу). Вірогідність такої події називається рівнем значущості і звичайно позначається буквою α. Протилежна подія є правильне рішення, його вірогідність дорівнює 1 - α; вона називається довірчою вірогідністю.

Помилка, що полягає у тому, що ми прийняли нульову гіпотезу в той час, як вона неправильна, називається помилкою другого роду (пропуском сигналу). Вірогідність такої помилки звичайно позначається буквою β. Вірогідність протилежної події дорівнює 1 - β; ця величина називається потужністю критерію. Потужність критерію - це його здатність виявляти відмінності, якщо вони є. Іншими словами, це його здатність відхиляти нульову гіпотезу про відсутність відмінностей, якщо вона неправильна. Потужність критерію звичайно визначається шляхом досліджень. При зменшенні рівня значущості знижується вірогідність помилки першого роду, але при цьому збільшується вірогідність помилки другого роду. Одночасно зменшити α і β можна тільки за рахунок збільшення об'єму вибірки.

Основні задачі статистичної перевірки гіпотез можуть бути розбиті на декілька груп [16]:

1. Виявлення відмінностей у рівні досліджуваної ознаки. Рівень ознаки може бути виражений: для номінальних шкал - частотами їх зустрічності, для порядкових шкал - їх рангами, для інтервальних шкал – їх середніми значеннями. При вирішенні цих задач використовуються дві або більше незалежних вибірок випробуваних, в кожному з цих випадків застосовуються різні критерії. Параметричними критеріями при вирішенні задач цієї групи є: t- критерій Стьюдента, v-критерій Уелша при вирішенні тієї ж задачі у вибірках з дисперсіями, що розрізняються, F-критерій Фішера при порівнянні двох дисперсій, G-критерій Кохрана та B-критерій Бартлета при порівнянні більше двох дисперсій. Включення трьох останніх критеріїв у цю групу є дещо умовним, проте величину дисперсії також можна вважати рівнем досліджуваної ознаки (що враховує не його величину, а мінливість). Решта критеріїв цієї групи відноситься до непараметричних.

2. Оцінка зрушення досліджуваної ознаки. Зрушенням називається зміна рівня ознаки в одній і тій же вибірці під впливом якихось

факторів: часу, ситуації, умов проведення випробувань тощо. Принциповою відмінністю задач цієї групи від задач першої групи є те, що тут йдеться про одну і ту ж вибірку (внутрішньогрупове дослідження), але поставлену в різні умови. При цьому можуть проводитися два або більше вимірів однієї і тієї ж ознаки. Залежно від цього використовуються різні критерії згоди. При вирішенні задач цієї групи можуть застосовуватися і параметричні критерії. У цьому випадку необхідно, щоб об'єм вибірки не мінявся і кожному елементу первинної вибірки був поставлений у відповідність той же елемент після діяння досліджуваного фактора. У результаті цього проводиться попарне порівняння їх середніх значень або дисперсій за допомогою параметричного t-критерію Стьюдента.

3. Виявлення відмінностей у розподілі ознаки. Ці задачі полягають у зіставленні (перевірці степеня відмінності) емпіричного розподілу з одним із теоретичних, або в зіставленні двох емпіричних розподілів. Тут використовуються тільки непараметричні критерії і серед них найпотужніший χ^2, а також λ-критерій Колмогорова – Смірнова.

4. Аналіз змін ознаки під впливом контрольованих умов. Число цих умов (факторів) може дорівнювати одній або двом. У першому випадку як параметричний метод можуть використовуватися однофакторний дисперсійний аналіз для незалежних або зв'язаних вибірок, а також декілька непараметричних методів: для незалежних вибірок – H-критерій Крускала – Уолліса і S-критерій тенденцій Джонкіра, для зв'язаних вибірок – χ^2_r – Фрідмана або L-критерій тенденцій Пейджа. Усі вони є непараметричними аналогами однофакторного дисперсійного аналізу. При цьому слід мати на увазі, що критерії H і χ^2_r дозволяють тільки констатувати наявність відмінностей (змін) досліджуваної ознаки під впливом будь-яких умов, а критерії S і L, крім цього, вказують і напрям (тенденцію) цих змін [16]. Як уже наголошувалося, параметричним методом у даній групі є однофакторний дисперсійний аналіз, який реалізується в двох модифікаціях - для незалежних і зв'язаних вибірок [4, 10]. Проте необхідно констатувати, що цей метод є досить складним і трудомістким для ручної обробки. Тому для спрощення розрахунків можна використовувати статистичні машинні програми типу Excel чи Statistika або застосовувати так звані «швидкі» методи - критерії дисперсійного аналізу. При розрахунках за допомогою цих методів (хоча вони і відносяться до параметричних) дисперсії не обчислюються, а варіативність оцінюється іншими способами: за допомогою величини розмаху між максимальним і мінімальним

значеннями ознаки (критерії Лінка і Уолліса) або за допомогою оцінки різниці рангів (критерій Немені). Використовування цих критеріїв значно прискорює і спрощує процедуру розрахунків. Слід, проте, мати на увазі, що ці критерії дають лише приблизну оцінку впливу фактора, і якщо вони не підтверджують цього впливу, то дисперсійний аналіз, можливо, його виявить [4]. На закінчення відзначимо, що у разі зміни ознаки під впливом двох факторів (умов) одночасно слід використовувати двофакторний дисперсійний аналіз [16, 18].

Він є, мабуть, єдиним методом, аналога якому в непараметричній статистиці немає. Решта параметричних методів має в ній відповідні аналоги (хоча і менш потужні).

Усі розглянуті критерії є однофункціональними. Вони призначені, як правило, для вирішення лише однієї (або групи схожих) задачі і можуть бути застосовані при певних обмеженнях. Серед них – об'єм вибірки (велика або мала), її характер (незалежна або зв'язана), кількість вибірок (дві або більше), тип використовуваної вимірювальної шкали. Абсолютно особливе значення серед цих критеріїв займають так звані багатофункціональні критерії. Вони можуть використовуватися для найрізноманітніших даних, вибірок і задач. Це означає, що проаналізовані дані можуть бути подані в будь-якій шкалі, починаючи від номінальної (шкали найменувань), а порівнювані вибірки можуть бути як незалежними, так і зв'язаними. Іншими словами, за допомогою цих критеріїв можна порівнювати і різні вибірки випробуваних, і показники однієї і тієї ж вибірки, виміряні в різних умовах із застосуванням будь-яких вимірювальних шкал. До багатофункціональних критеріїв повною мірою відноситься критерій φ^* (кутове перетворення Фішера) і з деякими обмовками – біномний m-критерій [16].

Багатофункціональні критерії дозволяють вирішувати вкрай широке коло задач: зіставлення рівня досліджуваної ознаки (порівняння середніх значень двох незалежних вибірок), зрушення у її значеннях (порівняння середніх значень в двох залежних вибірках, наприклад, в одній і тій же групі випробуваних), порівняння розподілів та ін. Іншими словами, вони можуть замінити переважну більшість з розглянутих вище критеріїв. Для застосування цих критеріїв достатньо підрахувати, як часто зустрічається ознака, що вивчається, при проведенні психологічного дослідження. А це можна зробити практично завжди. Багатофункціональні критерії побудовано на зіставленні значень ознаки, виражених у частках одиниці або відсотках. Суть критеріїв зводиться до визначення того, яка частина спостережень (реакцій, виборів, випробуваних) характеризується (або навпаки, не характеризується) ефектом, що цікавить дослідника. При цьому застосовувати ці критерії можливо до дуже малих вибірок (п'ять чоловік, а за певних умов і дві людини). Верхньої межі у φ^*-критерію не існує, а у m-критерію вона дорівнює 50.

Як бачимо, ці критерії (особливо кутове перетворення Фішера) мають достатньо велику універсальність і широкі можливості їх застосування в психологічній практиці, проте вони, на жаль, дуже рідко в ній використовуються. Одна з причин цього полягає у тому, що ці критерії, незважаючи на їх простоту і універсальність, не знайшли достатнього віддзеркалення в існуючій літературі, і багато фахівців про них просто не знають. Достатньо сказати, що навіть у таких популярних в психології працях, як [6, 10, 18, 19], про них навіть не згадується. Виняток становлять лише окремі праці [4, 16]. З цієї причини трохи пізніше ми на цих критеріях зупинимося докладніше.

Класифікацію розглянутих задач і можливості їх вирішення за допомогою методів перевірної і кореляційної статистики дано в табл. 3. Для зручності та полегшення вибору потрібного методу можна звернутися також до табл. 4, яка доповнює попередню, дає коротку характеристику того або іншого методу і показує умови його застосування. Таким чином, залежно від характеру вирішуваної задачі, умов проведення дослідження (особливості й об'єм вибірок, їх кількість, тип використовуваної шкали та ін.), характеру гіпотези, яку треба перевірити, у кожному конкретному випадку за допомогою табл. 3 і 4 можна обрати метод обробки одержаних даних, найадекватніший і можливий для застосування в даних умовах. Роботу рекомендується виконувати ще до проведення дослідження – на стадії його планування. Це дозволить зібрати всі необхідні для вирішення поставленої задачі дані й одержати потрібний результат (наприклад, перевірити гіпотезу, що висувається). Інакше може виявитися, що зібрані не ті або не всі дані, потрібні для вирішення поставленої задачі, або зібрані такі дані, які взагалі неможливо проаналізувати. Іншими словами, при розробці плану дослідження (і перш за все експерименту) потрібно чітко розуміти концепції, що лежать в основі статистичних операцій, які повинні бути використані для аналізу одержаних даних [8].

Таблиця 3

Можливості застосування різних статистичних методів у психології

Задачі	Умови застосування	Методи вирішення	Розрахункова формула	Література
Виявлення відмінностей у рівні досліджуваної ознаки	Дві вибірки випробуваних	t-критерій Стьюдента v-критерій Уелша q-критерій Розенбаума U-критерій Манна – Уїтні φ*-критерій (кутове перетворення Фішера) M-критерій Мак-Німара F-критерій Фішера	1, 3, 6 - - 13 21 36 7	4,7,10,13,14 12 10, 16 4, 13, 16 4, 16 4 4, 7, 10, 17, 20
	Три і більше вибірок випробуваних	S-критерій тенденцій Джонкіра g-критерій Кохрана H-критерій Крускала – Уолліса B-критерій Бартієта Критерій Лівена	- 8 15 - 10	16 18, 20 7,10,16 7, 18 3
Оцінка зрушення значень досліджуваної ознаки	Два виміри на одній і тій же вибірці	T-критерій Вілконсона G-критерій знаків t-критерій Стьюдента для попарного порівняння: - середніх значень - дисперсій	14 - 4 -	4, 8, 16 4, 8, 16 7,8,9,10,13 7, 10 7, 10
	Три і більше вимірів на одній і тій же вибірці	χ_r^2-критерій Фрідмана L-критерій тенденцій Пейджа	16 -	4, 10, 16 10, 16

Продовження табл. 3

Задачі	Умови застосування	Методи вирішення	Розрахункова формула	Література
Виявлення відмінностей в розподілі ознаки	При зіставленні емпіричного розподілення з теоретичним	φ^2-критерій Пірсона λ-критерій Колмогорова – Смірнова біномний m-критерій	17 18, 19 22	1,4,6,16,18 4, 7, 16, 19 16
	При зіставленні двох емпіричних розподілів	φ^2-критерій Пірсона λ-критерій Колмогорова – Смірнова φ*-критерій (кутове перетворення Фішера)	17 19 21	1,4,6,16,18 4, 7, 16, 19 4, 16
Аналіз зміни ознаки під впливом короткочасних умов	При впливі одного фактора	S-критерій тенденцій Джонкіра H-критерій Крускала – Уолліса ς_r^2-критерій Фрідмана L-критерій тенденції Пейджа Дисперсійний однофакторний аналіз Фішера: - для незалежних вибірок - для зв'язаних вибірок Критерій Лінка і Уолліса Критерій Німені	- 15 16 - 7 - 12 -	16 7, 10, 16 4, 10, 16 10, 16 4, 7, 8, 10, 14, 16, 17 4, 10, 16 4 4
	При впливі двох факторів	Дисперсійний двофакторний аналіз Фішера	-	16, 18

Продовження табл. 3

Задачі	Умови застосування	Методи вирішення	Розрахункова формула	Література
Виявлення степеня зв'язку (узгодженості двох ознак)	Ознаки наведені в шкалі інтервалів	Коефіцієнт лінійної кореляції Пірсона Коефіцієнт детермінації	9 11	1, 4, 7, 9, 10, 14, 18, 19 4, 7, 10, 18
	Ознаки наведені в шкалі порядків	Коефіцієнти рангової кореляції: - Спірмена - Кендала	 23 24	 1, 4, 8, 10, 16 1, 4, 10
	Ознаки подані в шкалі найменувань	Коефіцієнти спряженості: - чотириклітинної - багатоклітинної -тетрахоричний коефіцієнт кореляції	 26, 27 28, 29 -	 1, 10, 12, 18 1, 10, 12, 18 3
	Ознаки подані в різних шкалах (одна – у номінальній, інша – у інтервальній або порядковій)	Коефіцієнти кореляції: - точково-бісеріальної - рангово-бісеріальної	 30 31	 1, 4, 10 1, 4, 10
Те ж, для більше ніж дві ознаки	Ознаки подані в шкалі порядків	Коефіцієнти конкордації: - Спірмена - Кендала і Сміта	 - 32, 33	 18 18
	Ознаки подані в шкалі інтервалів	Множинна регресія Множинна кореляція Парціальна (частинна) кореляція Факторний аналіз Багатовимірне шкалювання Кластерний і дискримінантний аналіз	37 38, 39 40 - - -	4, 7, 10, 19 4, 7, 18 4, 7, 10, 19 4, 8, 10, 18, 19, 20 10, 11, 12 10, 11, 12

Закінчення табл. 3

Задачі	Умови застосування	Методи вирішення	Розрахункова формула	Література
Виявлення результатів вимірювань, що відхиляються	Ознака подана в шкалі інтервалів	Критерій грубих помилок	35	7, 15

Примітка. Таблиця є узагальненням даних, наведених у працях [16, табл. 1.2]; [4, табл. дод. 2]: [9, с. 176].

Таблиця 4

Умови застосування статистичних критеріїв у психології

Критерії	Шкала оцінок			Характер вибірки					Можливість використання при малому n	Література
	Номінальна	Порядкова	Інтервальна	Одна	Дві незалежні	Більше ніж дві незалежні	Дві зв'язані	Більше ніж дві зв'язані		
t-критерій Стьюдента			+		+				+	4, 7, 10, 13, 14
t-критерій Стьюдента для попарного порівняння			+				+		+	4,7, 10,13,14
F-критерій Фішера			+		+				+	4,7, 10,13,14
Критерій Бартлета			+			+			+	7,18
Критерій Кохрана			+			+			+	7,18
Коефіцієнт кореляції Пірсона			+	+	+$^{+)}$				+	1,4,7,10, 14,18, 19
Коефіцієнт детер-мінації			+	+	+$^{+)}$				-	4,7,10, 18
U-критерій Манна – Уїтні		+	+		+				+	4,13,16
q-критерій Розенбаума			+		+				+	10,16
Дисперсійний аналіз			+			+			+	3,4,7,10, 14,16,17, 18
Дисперсійний аналіз для зв'язаних вибірок			+					+	+	4,10,16,1 8
Критерій Лінка і Уолліса			+			+			+	4
Критерій Немені		+	+			+			+	4
Т-критерій Вілконсона Критерій знаків		+					+		+	4, 8, 16
Н-критерій Крускала - Уолліса S-критерій тенденцій		+				+			+	4, 16

Продовження табл. 4

Критерії	Шкала оцінок			Характер вибірки					Можливість використання при малому n	Література
	Номінальна	Порядкова	Інтервальна	Одна	Дві незалежні	Більше ніж дві незалежні	Дві зв'язані	Більше ніж дві зв'язані		
χ^2_r-критерій Фрідмана L-критерій тенденцій		+						+	+	4,16
Критерій χ^2 Пірсона	+			+	$+^{+)}$	$+^{+)}$			-	1, 4, 6, 16, 18
λ-критерій Колмогорова – Смірнова		+		+	$+^{+)}$				+	4, 7, 16, 18
Кутове перетворення Фішера	+			+	+		+		+	4,16
Біномний критерій	+			+	+		+		+	16
Критерії Мак-Німара	+						+			4
Коефіцієнти кореляції Спірмена, Кендала		+		+	$+^{+)}$				+	1, 4, 10, 16
Коефіцієнти зв'язаності	+			+	$+^{+)}$		$+^{+)}$		+	1, 10, 12, 18
Коефіцієнт точково-бісеріальної кореляції	+		+	+					+	1, 4, 10
Коефіцієнт ран-гово-бісеріальної кореляції		+	+	+					+	1, 4, 10
Коефіцієнт кон-кордації		+				$+^{+)}$			+	18
Критерій грубих помилок			+	+					+	7,15
Факторний аналіз			+	+		$+^{+)}$			-	1, 4, 10, 18, 19
Кластерний аналіз			+	+		$+^{+)}$			-	10
Багатовимірне шкалювання			+	+		$+^{+)}$			-	10, 11, 12

Закінчення табл. 4

Критерії	Шкала оцінок			Характер вибірки					Можливість використання при малому n	Література
	Номінальна	Порядкова	Інтервальна	Одна	Дві незалежні	Більше ніж дві незалежні	Дві зв'язані	Більше ніж дві зв'язані		
Багатовимірний регресійний аналіз			+	+		++)			-	4, 7, 10, 18, 19
Багатовимірний кореляційний аналіз			+	+		++)			+	4, 7, 18, 19
Частинна кореляція			+	+					+	4, 7, 10, 19

+) Тут і в подібних цьому випадках йдеться не про кількість вибірок (вона, як правило, одна), а про кількість одночасно аналізованих і вимірюваних незалежно одна від одної ознак.

3. Опис і правила використовування основних параметричних статистичних критеріїв

Подані в табл. 3 і 4 критерії та методи статистичної обробки, що їх реалізують, відносяться до одновимірних. Це означає, що при їх використовуванні вивчення і опис того або іншого психологічного явища ведуться тільки за допомогою одного показника. Якщо ж таких показників декілька, то кожний з них аналізується окремо і незалежно один від одного. У тих випадках, коли аналіз явища потрібно проводити за декількома показниками одночасно, доводиться застосовувати методи багатовимірної статистики. Вони є складнішими порівняно з одновимірними, але іноді без їх допомоги при обробці дослідних даних не обійтися. Тому в наступному розділі дамо їх коротку характеристику, а поки обмежимося розглядом основних одновимірних методів.

Критерій згоди Стьюдента (t-критерій) є параметричним і призначений для встановлення схожості-відмінності середніх значень в двох вибірках. Крім того, він застосовується для перевірки значущості (відмінності від нуля) деяких мір зв'язку, а також для перевірки значущості відмінності дисперсій двох залежних вибірок.

Для перевірки значущості відмінності середніх значень двох незалежних вибірок можуть використовуватися різні формули. Однією з найбільш вживаних є така:

$$t = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}} \,, \qquad (1)$$

де n_1 і n_2 - об'єми вибірок; $\overline{x_1}$ і $\overline{x_2}$ – їх середні значення; S_1^2 і S_2^2 - їх вибіркові дисперсії.

Якщо між вибірковими дисперсіями S_1^2 і S_2^2 немає значущих відмінностей, їх можна використовувати для визначення середньозваженої дисперсії. Розрахунок ведеться за формулою

$$S^2 = \frac{(n_1 - 1)S_1^{\,2} + (n_2 - 1)S_2^{\,2}}{n_1 + n_2 - 2} \,. \qquad (2)$$

У цьому випадку величину t можна визначити за формулою

$$t = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{S\sqrt{\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}}} \,. \qquad (3)$$

Знайдене емпіричне значення t порівнюється з табличним при вибраному рівні значущості та числі степенів вільності $df = n_1 + n_2 - 2$. Якщо виявляється, що $t_е > t_т$, можна зробити висновок про статистично значущу відмінність середніх значень $\overline{x_1}$ і $\overline{x_2}$ і неприйняття нульової гіпотези.

У разі зв'язаних (залежних) вибірок необхідно, щоб об'єми вибірок були однакові і кожному елементу однієї вибірки x_{i1} був поставлений у відповідність аналогічний елемент x_{i2} іншої вибірки. Для кожної пари аналогічних значень обох вибірок визначаються різниці, з якими і ведеться подальша статистична обробка. Величину t знаходять за формулою

$$t = \frac{\bar{d}_i \sqrt{n}}{S_d}, \tag{4}$$

де $\bar{d}$ і S_d - середнє значення та середньоквадратичне відхилення величини $\bar{d}_i$, які знаходять за загальними правилами визначення цих показників, тобто

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}, \quad S_d = \sqrt{\sum_{i=1}^{n} \frac{(d_i - \bar{d})^2}{n-1}}\ . \tag{5}$$

Кількість степенів вільності $df = n - 1$. Набуте емпіричне значення t порівнюється з табличним і робиться висновок про значущість відмінностей середніх значень в обох вибірках.

Критерій Стьюдента застосовується також для порівняння дисперсій двох залежних вибірок [7]. У цьому випадку виникає кореляція результатів, для обліку якої при порівнянні дисперсій залежних вибірок використовують формулу

$$t = \frac{S_1^2 - S_2^2}{\sqrt{\frac{4 S_1^2 S_2^2}{n-2}(1 - r_{12}^2)}}\ , \tag{6}$$

де $S_1^2 - S_2^2$ - оцінки дисперсій вибірок; n - їх об'єм; r_{12} - коефіцієнт кореляції, що визначається за парними значеннями обох вибірок.

Одержана величина t має $df = n - 2$ степенів вільності і порівнюється з табличним значенням на заданому рівні значущості.

Критерій згоди Фішера[*] (F-критерій) застосовується для порівняння дисперсій двох незалежних вибірок. Для цього обчислюється величина

$$F = \frac{S_1^2}{S_2^2} \tag{7}$$

яка порівнюється з табличним значенням. Якщо знайдене емпіричне значення перевищує табличне, то відмінність між порівнюваними дисперсіями вважається значущою. При користуванні формулою (7) потрібно мати на увазі, що індекс 1 відноситься до вибірки з більшою дисперсією, а числа степенів вільності дорівнюють: $df_1 = n_1 - 1$, $df_2 = n_2 - 1$.

Критерій Кохрана (g-критерій) використовується для перевірки однорідності дисперсій m незалежних вибірок (m>2) і визначається за формулою [7, 18, 19, 20]

$$g = \frac{\max S_i^2}{S_1^2 + S_2^2 + \dots S_m^2}\ . \tag{8}$$

Знайдене за формулою (8) емпіричне значення критерію порівнюється з табличним при $df = n - 1$ степенях вільності. Якщо воно менше табличного, то відмінності між порівнюваними дисперсіями можна вважати незначущими (приймається нульова гіпотеза). Така задача часто виникає при проведенні дисперсійного аналізу. Застосовувати критерій можна лише при однаковому об'ємі порівнюваних вибірок ($n_1 = n_1 = \dots n_m = n$). Інакше слід використовувати критерій Бартлета або Лівена. Вони є потужнішими, але й складнішими з погляду їх обчислення [7, 18]. Критерій Лівена також застосовується для перевірки однорідності (гомогенності) n>2 дисперсій і ґрунтується на ідеї однофакторного дисперсійного аналізу для незалежних вибірок. У більшості випадків він стає нечуттєвим до відхилення вихідних даних від нормального закону розподілу та має достатню потужність [3]. Тому він переважно застосовується в сучасних машинних програмах статистичної обробки даних [10].

Коефіцієнт лінійної кореляції Пірсона використовується для встановлення міри взаємозв'язку двох ознак (х і у), виражених в інтервальній шкалі, що підпорядковуються нормальному закону розподілу та зв’язані між собою лінійною залежністю. Він називається також твором моментів Пірсона, і його обчислюють за формулою

[*] F-критерій Фішера не слід змішувати з кутовим перетворенням Фішера (φ^*критерій).

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{(n-1)\sigma_x \sigma_y} \quad , \tag{9}$$

де $\overline{x}$ і $\overline{y}$ – середні значення величин x_i і y_i; σ_x і σ_y – їх стандартні (середньоквадратичне відхилення) відхилення; n - об'єм вибірки.

Величина r_{xy} може лежати в межах від -1 до +1, перевірка значущості (відмінності від нуля) набутого значення rxy проводиться за допомогою t-критерію Стьюдента або спеціальних таблиць. Обчислене значення r_{xy} вважається статистично значущим (достовірним), якщо воно перевищує табличне значення при даному рівні значущості. Інакше навіть при великому абсолютному значенні r_{xy} говорити про наявність взаємозв'язку величин x і y неправомірно. У той же час за інших умов (більший об'єм вибірки або рівень значущості) навіть менші значення r_{xy} можуть свідчити про достовірний (хоча й менш виражений) взаємозв'язок проаналізованих ознак. Число степенів вільності $df = n-1$.

Коефіцієнт детермінації (кореляційне відношення) η є розширенням поняття коефіцієнта лінійної кореляції для випадку нелінійної залежності між величинами x і y. Іншими словами, він є загальнішою і універсальною мірою для визначення ступеня взаємозв'язку двох випадкових величин. За абсолютною величиною він також лежить в межах від -1 до +1; при лінійній залежності величин x і y* $\eta_{xy} > |r_{xy}|$ і $\eta_{xy} \neq \eta_{xy}$. В основі обчислення величини η лежить відоме в теорії вірогідності співвідношення [7, 19], згідно з яким для системи залежних випадкових величин x і y справедливе співвідношення

$$\sigma_y^2 = \sigma_p^2 + \sigma_{y/x}^2 \quad , \tag{10}$$

де σ_y^2 - безумовна (повна) дисперсія величини y; $\sigma_{y/x}^2$ - умовна дисперсія y, обчислена тільки за фіксованими значеннями x (випадкова дисперсія або дисперсія відтворюваності); $\sigma_p^2 = \sigma_y^2 - \sigma_{y/x}^2$ - частина повної дисперсії, що характеризує стохастичний компонент (факторна дисперсія, або дисперсія, що пояснюється регресією). З урахуванням сказаного коефіцієнт детермінації у найзагальнішому вигляді можна визначити як квадратний корінь з відношення

* $\eta_{xy}=r_{xy}$, при нелінійній залежності.

дисперсій умовних середніх (факторних) дисперсій до безумовної, тобто

$$\eta = \sqrt{\frac{\sigma_p^{\ 2}}{\sigma_y^{\ 2}}} = \sqrt{1 - \frac{\sigma^2 {y/x}}{\sigma_y^{\ 2}}}\ . \tag{11}$$

Коефіцієнт детермінації (на відміну від коефіцієнта лінійної кореляції) встановлює два типи зв'язків: залежність у від х, що визначається виразом (11), і залежність х від у, які в загальному випадку можуть не збігатися ($\eta_{xy} \neq \eta_{yx}$). Іншими словами, спочатку одна змінна розглядається як залежна, інша - як незалежна, потім – навпаки [10].

Дисперсійний аналіз (у англійській транскрипції ANOVA – analysis of variance) є параметричним методом, що дозволяє виявити вплив однієї або декількох незалежних змінних (чинників) на залежну змінну (показник, що вивчається). Відповідно до цього розрізняють одно- і багатофакторний аналіз. Він може проводитися відносно незалежних або зв'язаних вибірок однакової або різної чисельності. При проведенні однофакторного аналізу вивчається вплив одного фактора на залежну змінну, багатофакторний аналіз дозволяє виявляти крім дії кожного окремого фактора дію поєднань різних факторів.

Суть дисперсійного аналізу полягає у розчленовуванні загальної дисперсії ознаки, що вивчається, на окремі компоненти. В однофакторному варіанті при цьому використовується вираз, аналогічний (10). Вираз у даному випадку має вигляд

$$S^2_{\text{î áù}} = S^2_{\text{âí óò}} + S^2_{\text{ì ã}}$$

де $S^2_{\text{î áù}}$ – загальна (повна) дисперсія всіх спостережень; $S^2_{\text{âí óò}.}$ - внутрішньогрупова (залишкова) дисперсія; $S^2_{\text{ì ã}}$ - міжгрупова (факторна) дисперсія.

Дисперсія $S^2_{\text{âí óò}.}$ характеризує розсіювання всередині груп; вона викликана дією випадкових (неконтрольованих) явищ. Дисперсія $S^2_{\text{ì ã}}$ зумовлена впливом чинника, що вивчається, і характеризує варіацію групових середніх. У сукупності обидві ці дисперсії дають повну дисперсію. Вплив фактора признається значущим, якщо дисперсія

$S^2_{\grave{i}\,\tilde{a}}$ значуще відрізняється від $S^2_{\hat{a}\acute{i}\,\acute{o}\grave{o}.}$, іншими словами – факторний компонент загальної дисперсії, тобто $S^2_{\grave{i}\,\tilde{a}}$, істотно перевищує її випадковий компонент $S^2_{\hat{a}\acute{i}\,\acute{o}\grave{o}.}$. Порівняння ведеться за критерієм Фішера (7), в якому велика дисперсія (тобто $S^2_{\grave{i}\,\tilde{a}}$) завжди знаходиться в чисельнику. При цьому кількість степенів вільності $df_1 = m, df_2 = m(n-1)$, де m – число експериментальних груп (рівнів зміни фактора), n – число вимірювань (випробуваних) у кожній групі. Ці дані відповідають умовам проведення однофакторного дисперсійного аналізу в незалежних незв'язаних вибірках однакової чисельності. Такий випадок є найпростішим і найпоширенішим. У складніших випадках сама ідея аналізу принципово не відрізняється від розглянутої, змінюються тільки деякі обчислювальні процедури. Детальніше з ними можна ознайомитися у відповідній літературі [3, 14, 16, 18].

Однак у будь-якому випадку слід знати: якщо встановлено вплив фактора на залежну зміну (психологічний показник, що вивчається), іншими словами – встановлено статистично значущу різницю хоча б між двома середніми значеннями (умова $S^2_{мг} > S^2_{внут}$ виконується), то необхідно встановити, які саме середні значення відрізняються. Для цього проводяться їх попарні порівняння (кожного з кожним). У найпростішому випадку вони проводяться за допомогою t-критерію Стьюдента [17]. Однак такий підхід вважається не зовсім коректним, оскільки не враховує одночасно всю отриману інформацію. Більш коректним у цьому випадку є порівняння середніх за допомогою t-критерію Шеффе. Методику та приклади його використання при проведенні дисперсійного аналізу наведено у роботах [3, 10].

Критерій Лінка і Уолліса застосовується з тією ж метою, що й однофакторний дисперсійний аналіз, але є його наближеним і спрощеним аналогом. Тому він відноситься до так званих «швидких» методів – критеріїв дисперсійного аналізу [4]. На відміну від останнього для аналізу варіативності (мінливості) цей критерій використовує не дисперсії, а величини розмаху між максимальним і мінімальним значеннями ознаки, що аналізується. Це істотно спрощує і прискорює процедуру проведення аналізу, але при цьому зменшує його точність. Емпіричне значення критерію розраховується за формулою

$$K_e = \frac{n(\max \bar{x}_i - \min \bar{x}_j)}{сума \quad розмахів}, \tag{12}$$

де $\max \bar{x}_i$, $\min \bar{x}_j$ - відповідно максимальне і мінімальне середні значення аналізованої ознаки для різних рівнів чинника в усій вибірці; сума розмахів обчислюється підсумовуванням різниць між максимальним і мінімальним значеннями для кожного рівня зміни фактора (кожної вибірки).

Одержане емпіричне значення $К_е$ порівнюється з табличним значенням $К_т$, що визначається для даного рівня значущості α і об'єму вибірки n. При виконанні умови $К_е > К_т$ вплив фактора визнається значущим (невипадковим).

Критерій Немені вирішує ті ж задачі, що й критерій Лінка і Уолліса, але оснований на ранжуванні всієї вибірки. Ранги кожної групи підсумовуються і обчислюються абсолютні значення їх різниці. За таблицею критичних значень робиться висновок про рівень схожості або відмінності в групах, тобто про значущість впливу фактора.

Для застосування «швидких» методів - критеріїв дисперсійного аналізу – необхідно дотримуватися таких умов. Вимірювання повинне бути проведене в шкалі інтервалів, а ознака, що вивчається, розподілена за нормальним законом. На кожному рівні зміни фактора досліджуються різні групи, тобто вибірки мають бути незалежними і однаковими за чисельністю. Кількість груп (рівнів зміни чинника) повинна бути не менше трьох, а в останньому випадку – не більше десяти. «Швидкі» методи дають лише приблизну оцінку впливу чинника, і якщо вони не визнають впливу, то однофакторний дисперсійний аналіз може впевненіше виявити його [4].

4. Опис і правила використовування основних непараметричних статистичних критеріїв

Розглянуті методи відносяться до параметричних, що накладає серйозні обмеження на можливості їх використання в психологічних дослідженнях. Це стосується як нормальності розподілів ознак, що вивчаються, так і можливості їх подання в інтервальній шкалі. Про це вже йшлося в розд. 1. До сказаного там слід додати, що навіть у тих випадках, коли психологічна ознака, що вивчається, виражена в метричних одиницях (наприклад, воля за допомогою часу або величини утримуваного мускульного зусилля), формальні ознаки метрики тут мають місце, але з психологічної точки зору це не завжди так. Треба пам’ятати, що повинно йтися перш за все про однаковість інтервалів, коли кожне з можливих значень ознаки має знаходитися від іншого на однаковій відстані. Для психологічних змінних це, як правило, не виконується. Тому тут ми в кращому разі маємо шкалу порядку і мусимо, хочемо ми того чи не хочемо, використовувати непараметричні методи [9, 16]. Вони оперують рангами або частотами. Розглянемо найбільш важливі з непараметричних критеріїв.

Критерій Манна – Уїтні (U-критерій) є непараметричним аналогом t-критерію Стьюдента для незалежних вибірок. Він призначений для оцінки відмінності між двома незалежними вибірками за рівнем будь-якої ознаки, вираженої кількісно. Для проведення розрахунків значення ознаки мають бути виражені у рангах. Розрахункова формула має вигляд

$$U = n_1 \cdot n_2 + \frac{n_x \cdot (n_x + 1)}{2} - T_x, \qquad (13)$$

де, n_1, n_2 - об'єми вибірок; T_x - більша з двох рангових сум; n_x - об'єм вибірки з більшою ранговою сумою (тобто n_1 або n_2).

Знайдене за формулою (13) емпіричне значення U порівнюється з табличним для числа вимірювань n_1 (для вибірки меншого об'єму) і n_2. На відміну від попередніх критеріїв відмінність між вибірками признається істотною на даному рівні значущості, якщо $U < U_T$. Аналогічне U-критерію призначення має непараметричний критерій Розенбаума (q-критерій), який є, проте, дещо менш потужним, ніж U-критерій [10, 16]. Критерій Вілконсона (Т-критерій) вирішує ті ж задачі, що й U-критерій, але у випадку залежних (зв'язаних) вибірок. Він є непараметричним аналогом t-критерію Стьюдента для попарного порівняння. Для проведення розрахунків за допомогою Т-критерію значення ознаки повинні бути виражені в рангах. Розрахункова формула має вигляд

$$T_1 = \sum R_+ , T_2 = \left| \sum R_- \right| , \tag{14}$$

де, R_+, R_- – ранги відповідно додатної та від'ємної різниць.

Як емпіричне значення Т$_е$ береться якнайменше із значень T_1, T_2, яке порівнюється з табличним для даного рівня значущості і об'єму вибірки. Якщо $T_р \le T_э$, відмінності між вибірками визнаються достовірними, тобто нульова гіпотеза відкидається [4, 10, 16]. Аналогічне Т-критерію призначення має критерій знаків, проте має меншу потужність, ніж Т-критерій.

Критерій Крускала ─ Уолліса (Н-критерій) призначений для оцінки відмінностей за рівнем вираженості деякої ознаки в m>2 незалежних вибірках одночасно. Він являє собою розширення критерію Манна – Уїтні на випадок більше двох вибірок і є непараметричним аналогом однофакторного дисперсійного аналізу для незалежних вибірок. Для застосування критерію, що аналізується, ознаки повинні бути виражені в рангах. Розрахункова формула має вигляд

$$H = \left[\frac{12}{N(N+1)} \cdot \sum_{i=1}^{m} \frac{T_i^2}{n_i} \right] - 3(N+1) , \tag{15}$$

де N - загальний об'єм об'єднаної вибірки; m - кількість незалежних вибірок; n_i - об'єм і-ї вибірки; T_i - суми рангів і-ї вибірки.

Знайдене за цією формулою емпіричне значення Н$_е$ порівнюється з табличним Н$_т$ (ним є величина χ^2 для даного рівня значущості та кількості степенів вільності). Якщо Н$_е$>χ^2, відмінності між вибірками мають місце і нульова гіпотеза відкидається [16]. Аналогічно може використовуватися також критерій тенденцій Джонкіра (S-критерій), який порівняно з Н-критерієм дозволяє додатково упорядкувати обстежувані вибірки за ознакою, що вивчається, тобто оцінити тенденцію його зміни [16].

Критерій Фрідмана (критерій χ^2_r) застосовується для зіставлення показників, виміряних у трьох і більше умовах на одній і тій же вибірці випробуваних. Він є розширенням критерію Вілконсона на випадок більше двох залежних вибірок і є непараметричним аналогом однофакторного дисперсійного аналізу для залежних вибірок. Для застосування критерію ознаки, що аналізується, повинні бути виражені в рангах. Розрахунок критерію ведуть за формулою

$$\chi^2_r = \left[\frac{12}{nc(c+1)} \sum_{i=1}^{c} T_i^2 \right] - 3n(c+1), \qquad (16)$$

де n - об'єм вибірки; c - кількість умов; T_i - сума рангів по кожній з умов.

Знайдене за формулою (16) значення при c=3 або c=4 порівнюється з табличним значенням χ^2_r, при c>4 – з табличним значенням χ^2_r при кількості степенів вільності $df = c - 1$. Якщо $\chi^2_{r\ емп}$ дорівнює критичному значенню χ^2 або перевищує його, відмінності достовірні. Аналогічно може використовуватися також критерій тенденцій Пейджа (L-критерій). Він дозволяє виявити тенденції в зміні величин ознаки при переході від умови до умови і є розширенням критерію Фрідмана, тому що не тільки констатує відмінності, але й вказує на напрям їх змін [16].

Критерій згоди χ^2 Пірсона призначений для виявлення відмінностей у розподілі ознаки, зокрема для порівняння емпіричного розподілу з одним з теоретичних або для порівняння між собою двох і більше емпіричних розподілів. Критерій застосовується також для оцінки значущості (відмінності від нуля) деяких коефіцієнтів спряженості. Він є одним з найпотужніших непараметричних критеріїв. Ще одна його перевага полягає у тому, що його можна застосовувати при використанні навіть найпростішої шкали - дихотомічної. Величина χ^2 обчислюється за формулою

$$\chi^2 = \sum_{i=1}^{K} \frac{(f_{ei} - f_{Ti})^2}{f_T} = n \sum_{i=1}^{K} \frac{(P_{ei} - P_{Ti})^2}{P_T}, \qquad (17)$$

де f_{ei} і f_{Ti} - емпірична і теоретична частоти i-го інтервалу групування; К – кількість таких інтервалів; P_{ei} і P_{Ti} - емпіричні та теоретичні частоти або апроксимуюча їх вірогідність; n - загальна кількість спостережень.

Знайдене за формулою (17) емпіричне значення χ^2 порівнюється з табличним для $df = K - 1$ степенів вільності. При виконанні умови $\chi^2_e > \chi^2_т$ відмінність між зрівнюваними розподілами вважається значущою [4, 16]. Формула (17) є найзагальнішим і найтиповішим виразом для обчислення емпіричного значення критерію χ^2. Вона застосовується звичайно для порівняння емпіричного розподілу з одним з теоретичних. Крім цього, критерій χ^2 використовується для порівняння між собою двох і більше емпіричних розподілів, а також для порівняння показників усередині однієї, але досить великої за чисельністю вибірки. У кожному з цих випадків можуть використовуватися свої формули для обчислення величини χ^2. Проте всі вони є окремими випадками формули (17).

Як один з прикладів наведемо вираз для обчислення величини χ^2 при встановленні схожості-відмінності між частотними розподілами декількох вибірок, інакше — для визначення належності — неналежності їх до однієї генеральної сукупності [9]. Обчислення величини χ^2 у цьому випадку ведеться за формулою

$$\chi^2 = \sum_{i=1}^{K}\sum_{j=1}^{S} \frac{(f_{eij} - f_{Tij})}{f_{Oij}},$$

де f_{eij}, $f_{Òij}$ - відповідно емпірична (спостережувана) і теоретична (очікувана) кількість випадків (спостережень), що приходяться на і-й розряд j-ї вибірки; К – кількість розрядів (інтервалів квантування); S - кількість порівнюваних вибірок.

Крім цього, при порівнянні між собою декількох емпіричних розподілів, коли необхідно визначити, чи належать вони до однієї і тієї ж генеральної сукупності, визначення критерію χ^2 багато в чому збігається з розрахунком коефіцієнтів спряженості [9, 13]. Кількість степенів вільності при цьому визначається за формулою $df = (K-1)(\tilde{n}-1)$, де К і с – відповідно кількість стовпців і рядків у таблиці спряженості [4]. Теоретичні частоти $f_т$ при цьому знаходяться за умовою їх незалежності для порівняння вибірок за формулами, наведеними у відповідній літературі [4, 16, 18].

Критерій згоди Смірнова – Колмогорова (λ-критерій) застосовується для зіставлення між собою емпіричного розподілу з одним з теоретичних або двох емпіричних розподілів. На відміну від критерію χ^2, в якому зіставляються частоти двох розподілів окремо по кожному розряду (інтервалу квантування), у критерії λ зіставляються накопичені частоти (інтегральні функції розподілу). В основі розрахунку λ-критерію – обчислення абсолютної величини різниці

$$d = \max\left|F(x) - F^*(x)\right|, \tag{18}$$

де $F(x)$ і $F^*(x)$ - відповідно теоретична й емпірична функції розподілу.

Якщо d_{max} або більше критичного (табличного), або дорівнює йому, або перевищує його при даному об'ємі вибірки, відмінності між розподілами визнаються достовірними [16]. При порівнянні двох емпіричних розподілів розраховують величину

$$\lambda = d\max\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}, \tag{19}$$

де n_1 і n_2 - об'єми порівнюваних вибірок.

Якщо знайдене за формулою (19) $\lambda_{емп} \geq 1{,}36$, відмінності між розподілами вважаються значущими на рівні $ð \leq 0{,}05$ [4, 16]. При застосуванні λ-критерію слід мати на увазі, що не можна накопичувати частоти, які відрізняються лише якісно і не є шкалою порядку (не впорядковані за збільшенням або зменшенням ознаки, що вивчається). Крім того, λ–критерій не рекомендується застосовувати при n<50. У цих випадках слід застосовувати критерій χ^2 [9, 10].

Критерій φ^* (кутове перетворення Фішера) є багатофункціональним (див. розд. 2), він дозволяє оцінити достовірність відмінностей між процентними частками двох вибірок, в яких зареєстрований і, навпаки, не зареєстрований ефект, що цікавить нас. Суть кутового перетворення Фішера полягає у перекладі процентних часток у величини центрального кута, що вимірюється в радіанах. Розрахунок ведеться за формулою

$$\varphi = 2\arcsin\sqrt{p}\,, \tag{20}$$

де р - частка спостережень, виражена у відсотках.

Значення φ^* розраховується за формулою

$$\varphi^* = (\varphi_1 - \varphi_2)\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}\,, \tag{21}$$

де n_1, n_2 - кількість спостережень у порівнюваних вибірках, причому індекс 1 відповідає більшому центральному куту.

При збільшенні розбіжності між кутами φ_1 і φ_2 і збільшенні чисельності порівнюваних вибірок значення критерію зростає. Чим більше величина φ^*, тим більш вірогідно, що відмінності достовірні. Критичні значення φ^* дорівнюють 1,64 при $\alpha \le 0{,}05$ і 2,31 при $\alpha \le 0{,}01$. Якщо знайдене за формулою (21) значення φ^* більше критичного для даного рівня значущості, то відмінності між порівнюваними вибірками слід вважати достовірними [4, 16].

Біномний критерій m відноситься до багатофункціональних і дозволяє вирішувати широке коло задач (див. розд. 2). Його застосування основане на зіставленні частоти зустрічності $m_{емп}$ будь-якого ефекту з теоретичною або заданою (критичною) частотою. Відмінності вважаються достовірними при виконанні умови

$$m_{емп} \ge m_{кр}. \tag{22}$$

Значення $m_{кр}$ залежать від об'єму вибірки і обраного рівня значущості [16].

Коефіцієнт кореляції Спірмена служить для оцінки ступеня взаємозв'язку двох ознак, виміряних у порядковій шкалі (у рангах), і є непараметричним аналогом коефіцієнта кореляції Пірсона.

Для розрахунку коефіцієнта кореляції Спірмена використовують формулу

$$\rho_{xy} = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \qquad (23)$$

де n - об'єм вибірки; d - різниця рангів для і-ї пари порівнюваних ознак. Знайдене за формулою (23) значення ρ_{xy} порівнюється з табличним для $df = n-2$ степенів вільності. Якщо це значення виявляється більше табличного, воно є значущим (відмінним від нуля). Якщо при обчисленні коефіцієнта ρ_{xy} хоча б для одного ряду порівнюваних признаків є збіжні ранги, необхідно вводити спеціальну поправку. Методику та приклади її обчислення і застосування показано в працях [9, 16]. *Коефіцієнт рангової кореляції Кендала* використовується для тих же цілей, що і коефіцієнт кореляції Спірмена, але оснований на іншому принципі. Тут не розраховуються різниці рангів, як це робиться у формулі (23), а визначається кількість збігів та інверсій, які визначаються відносно натурального ряду чисел (1, 2, 3,…). Розрахункова формула має вигляд

$$\tau = \frac{\sum P - \sum Q}{0,5n(n-1)}, \qquad (24)$$

де $\sum P$ - сума кількості збігів; $\sum Q$ - сума кількості інверсій.

При підрахунку величин $\sum P$ і $\sum Q$ для кожного рядка визначається кількість збігів і інверсій відносно всіх рядків, що знаходяться нижче початкової таблиці рангів, або, що те ж саме, всі рядки порівнюються між собою попарно, потім підраховуються суми збігів і інверсій. Коефіцієнт кореляції τ, а також коефіцієнти спряженості, що розглядаються далі, не мають спеціальних таблиць для знаходження їх критичних значень. Тому їх знаходження визначається за допомогою t-критерію Стьюдента за формулою

$$t_{êð} = |\tau| \sqrt{\frac{n-2}{1-\tau^2}}. \qquad (25)$$

Знайдене за формулою (25) емпіричне значення $t_{êð}$ порівнюється з табличним значенням t_T при даному рівні значущості та кількості степенів вільності $df = n-2$. Якщо $t_{êð} > t_T$, знайдене за формулою (25) значення τ достовірно відрізняється від нуля [4, 10].

Однак при цьому слід звернути увагу читача на той факт, що застосування формули (25) для оцінки значущості коефіцієнта кореляції t, пропоноване у праці [4], є не зовсім коректним. Уся справа в тому, що ця формула справедлива тільки для оцінки значущості коефіцієнтів кореляції Пірсона, які визначаються за формулою (9), та всіх інших, отриманих на її основі. До них, зокрема, відносяться коефіцієнт кореляції Спірмена (23) та φ-коефіцієнт чотириклітинної спряженості Пірсона (27). Ці коефіцієнти є частинними випадками коефіцієнта лінійної кореляції Пірсона (9) і отримані на його основі за умови, що порівнювальні признаки подані відповідно у ранговій (коефіцієнт Спірмена) та номінативній (коефіцієнт спряженості φ) шкалах. Коефіцієнт кореляції Кендала ґрунтується на зовсім іншому принципі та має іншу природу. Тому для оцінки його значущості застосування формули (25) не зовсім коректне. У цьому випадку слід використовувати Z-критерій та одиничне вдале розподілення [10].

Коефіцієнти спряженості (кореляції якісних ознак) використовуються для кількісної оцінки ступеня зв'язку (залежності) двох ознак, поданих у номінативних шкалах. У психології використовуються чотири види таких коефіцієнтів: два з них відносяться до дихотомічних змінних, коли в одному явищі виділяються два класи за альтернативним принципом (та – ні), два інших – до випадку, коли в досліджуваних явищах (якостях) виділяються не дві, а більша кількість градацій [10, 18].

Класифікація об'єктів за двома властивостями приводить до побудови таблиці спряженості розміром 2х2, приклад якої має вигляд [10] табл. 5.

Таблиця 5

Класифікація об'єктів за двома властивостями

Відвідування занять	**Варіант здачі заліку**	
	Здав залік з першого разу (х)	Не здав залік з першого разу (х)
Студент відвідував більше 50% занять (У)	a	b
Студент відвідував менше 50% занять ($\overline{Y}$)	c	d

У цій таблиці величини a, b, c, d відповідають кількості випадків, що мають одночасно властивості стовпця і рядка. Маючи такі дані, можна оцінити взаємозв'язок між подіями Х і У (у даному випадку між відвідуванням занять і успішністю). Для цього використовуються два види коефіцієнтів чотириклітинної (2х2) спряженості, які обчислюються за формулами

$$Q = \frac{ad - cb}{ad + cb} \ , \tag{26}$$

$$\varphi = \frac{ad - cb}{\sqrt{(a+b)(c+d)(b+d)(a+c)}} \ . \tag{27}$$

Величина Q назвається коефіцієнтом контингенції, величина φ - коефіцієнтом асоціації (коефіцієнтом чотириклітинної спряженості Пірсона). Обидва коефіцієнти змінюються в межах від -1 до +1, при цьому не завжди можна дати змістовну інтерпретацію додатному і від'ємному зв'язкам. Перевірка значущості відмінності від нуля цих коефіцієнтів виконується за формулою (25). Значення коефіцієнтів Q і φ можуть розрізнятися між собою, досить строгих обгрунтувань того, який з них кращий, немає.

При застосуванні таблиць чотириклітинної (2х2) спряженості й обчисленні коефіцієнтів кореляції дихотомічних змінних за формулою (27), тобто коли число степенів вільності df=1, необхідно звернути увагу на ряд обставин.

По-перше, φ-коефіцієнт Пірсона однозначно зв'язаний з критерієм χ^2. Цей зв'язок визначається формулою $\chi^2 = n\varphi^2$, яка і лежить в основі перевірки значущості коефіцієнтів кореляції (спряженості) дихотомічних змінних. Для цього знайдене за цією формулою емпіричне значення χ^2 порівнюється з табличним для df=1 степенів вільності. Якщо $\chi^2_е > \chi^2_т$, кореляція (спряженість) між; порівнювальними дихотомічними змінними визнається невипадковою на даному рівні значущості. По-друге, при аналізі таблиць спряженості 2х2 і визначенні на їх основі емпіричного значення критерію χ^2 у ряді випадків необхідно вводити поправку Йєтса на неперервність. Вона полягає у зменшенні різниці між емпіричною χ^2_e і теоретичною χ^2_T частотами, взятими за абсолютною величиною, на 0,5 для кожної клітинки чотириклітинної таблиці відповідно за формулою

$$\chi^2_e = \Sigma \frac{\left(\left|\chi_e - \chi_T\right| - 0.5\right)^2}{\chi_T} \ .$$

Величина χ^2_e з поправкою на неперервність визначається у будь-якому випадку при df=1, коли n>40, а якщо жодна з теоретичних частот не менше 5, то – при n≥20 [10]. Якщо таблиця спряженості 2х2 не задовольняє ці умови, то слід скористатися точним методом Фішера (ТМФ). Процедуру та зразок його застосування див. Гублер Е.В., Генкин А.А. Применение непараметрических критериев статистики в медико-биологических исследованиях. – Л.: Медицина, 1973. – С. 33 – 36.

По-третє, слід мати на увазі, що розглянуті в пп. 1 і 2 положення справедливі тільки для **незалежних вибірок**. Коли ж вибірки **залежні**, тобто одна й та сама вибірка класифікована за одною і тією ж дихотомічною основою двічі (наприклад, до і після здійснення деякого впливу), слід застосовувати критерій Мак-Німара. Порядок і приклади його застосування наведено в працях [4, 10]. Також слід зазначити, що у літературі зустрічаються різні назви цього критерію: критерій Мак-Ньюмара [1], критерій Макнамари [4], критерій Макнамара (Бодалев А.А., Столин В.В. Общая психодиагностика. – СПб.: Речь, 2002. – С. 132). В усіх цих випадках йдеться про один і той же критерій (метод), а різниця у назвах викликана, вочевидь, помилками перекладу. По-четверте, як відзначалось у розд. 2, частинним випадком дихотомічної класифікації є такий, коли для встановлення зв‘язку між дихотомічними змінними розглядаються дві нормально розподілені дихотомічні змінні. У цьому випадку більш точне значення для визначення зв‘язку між ними порівняно з φ-коефіцієнтом спряженості Пірсона дає можливість отримати тетрахоричний коефіцієнт кореляції, який обчислюється за формулою

$$r_{tet} = \cos \frac{180^0}{1+\sqrt{bc/ad}}.$$

Ця формула також запропонована Пірсоном, її не рекомендується застосовувати, якщо (a+b)/n або (b+d)/n (тобто P_y або P_x) значно відхиляються від 0,5. Іншими словами, якщо (a+b)n або (b+d)n більше 0,7 або менше 0,3, обчислення r_{tet} може відбуватися з великими похибками [3].

Якщо в досліджуваних ознаках виділяється більше двох градацій, використовуються коефіцієнти багатоклітинної (m·n) спряженості: К – Чупрова і С – Пірсона. Вони обчислюються за такими формулами:

$$K = \sqrt{\frac{\chi^2}{N(n-1)(m-1)}}, \qquad (28)$$

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}, \qquad (29)$$

де N - загальне число випробуваних; m і n - кількість рядків і стовпців у таблиці спряженості; χ^2- критерій згоди, що визначається за формулою (17).

Висновок про відмінність значень цих коефіцієнтів від нуля і, отже, про наявність або відсутність зв'язку робиться за величиною χ^2 [1, 3]. Обидва коефіцієнти змінюються тільки від нуля до одиниці та вимірюють тісноту спряженості, не вказуючи її напряму. Про нього можна судити за формою спільного розподілу вірогідності (частот) подій. Коефіцієнт Чупрова менш чутливий до кількості подій mxn, ніж коефіцієнт Пірсона, тому при малих значеннях m і n краще застосовувати коефіцієнт Чупрова [18, 19].

Бісеріальні коефіцієнти кореляції використовуються для оцінки зв'язку двох ознак, виміряних у різних шкалах: один – у дихотомічній шкалі найменувань, інший – в інтервальній (точково-бісеріальний коефіцієнт $r_{рв}$); або один – у дихотомічній шкалі, інший – у ранговій (рангово-бісеріальний коефіцієнт $r_{гв}$). Ознаку, подану у дихотомічній шкалі, позначимо У, вона може набувати всього двох значень (0 або 1); ознаку, подану в інтервальній або ранговій шкалі, позначимо Х. З урахуванням цих зауважень розрахункові формули матимуть такий вигляд:

$$r_{рв} = \frac{M_{х/1} - M_{х/0}}{\sigma_х}\sqrt{\frac{n_1 \cdot n_0}{n(n-1)}}\,, \tag{30}$$

$$r_{rв} = \frac{2}{n}(R_{x/1} - R_{x/0})\,, \tag{31}$$

де $M_{х/1}$ і $M_{х/0}$ - середнє значення об'єктів, що мають відповідно одиницю і нуль по у; $\sigma_х$ - середньоквадратичне відхилення всіх об'єктів, виражене в інтервальній шкалі; n_1 і n_0 - кількість об'єктів, що мають відповідно одиницю і нуль по У; n - загальна кількість об'єктів (об'єм вибірки); $R_{x/1}$ і $R_{x/0}$ - середній ранг об'єктів, що мають відповідно одиницю і нуль по Х.

Зазначені коефіцієнти можуть набувати значень від -1 до +1. Як і для коефіцієнтів φ і Q, від'ємний знак цих коефіцієнтів, як правило, змістовної інтерпретації не має. Відмінність від нуля перевіряється за допомогою t-критерію Стьюдента за формулою (25) [1, 4, 10].

Коефіцієнт конкордації (узгодженості) є розширенням коефіцієнтів рангової кореляції на випадок n>2 ознак. Необхідність обчислення коефіцієнта конкордації виникає, наприклад, при груповій оцінці особи (ГОО) за декількома ознаками, виявленні професійно важливих якостей людини, вивченні відмінностей між груповою оцінкою і самооцінкою, експертній оцінці якості продукції тощо. В таких випадках створюється група незалежних експертів і кожного з них просять прорангувати за ступенем важливості або вираженості у даної людини кожної з обстежуваних якостей. Якість такої оцінки багато в чому залежить від ступеня узгодженості думок експертів, які й оцінюються за допомогою коефіцієнта конкордації за формулою, запропонованою М. Кендалом і Б. Смітом:

$$W = \frac{12S}{m^2(n^3 - n)} \quad , \qquad (32)$$

де m - кількість експертів; n - кількість параметрів, що рангуються; S – сума квадратів відхилень рангових оцінок, що фактично зустрічаються.

Сума розраховується за формулою

$$S = \sum_{i=1}^{n}\left[\sum_{j=1}^{m} R_{ij} - \frac{1}{2}m(n+1)\right]^2 , \qquad (33)$$

де R_{ij} - ранг, присвоєний j-м експертом за i-ю ознакою.

Коефіцієнт конкордації змінюється в межах від 0 (немає збігу думок експертів) до 1 (повний збіг думок). Перевірка його значущості (відмінності від нуля) проводиться за критерієм χ^2, емпіричне значення якого обчислюється за формулою

$$\chi^2 = \frac{12S}{mn(n+1)} . \qquad (34)$$

Це значення порівнюється з табличним для $df = n - 1$. Якщо $\chi_e^2 \rangle \chi_T^2$, то узгодженість думок експертів на даному рівні значущості є невипадковою [18]. Якщо у деяких експертів за будь-якими ознаками є ранги, що збіглися, то у формули (32) – (34) вносяться відповідні уточнення. Для збільшення ступеня узгодженості експертів по черзі виключаються з розгляду один або декілька експертів і за формулою (32) знов обчислюється величина W. Експерт, думка якого вносить найбільшу розузгодженість, виключається з процесу ранжування або він пояснює, чому так зробив ранжування ознак. На підставі цього вносяться корективи в експертний аналіз [18]. *Критерій грубих помилок спостережень* (т-критерій) дозволяє виявити наявність у вибірці грубих помилок (промахів). Такі помилки (значення, що «вискакують») характеризуються різкою відмінністю від результатів решти дослідів (вимірювань). Вони виникають унаслідок дії випадкових чинників, що істотно виходять за межі норми (неуважність або різкі коливання стану випробуваного, неправильне фіксування даних, дія несподіваного подразника, не передбаченого програмою дослідження, та ін.). Обґрунтований висновок про виключення таких значень з процесу обробки одержаних даних базується на застосуванні методів статистичної перевірки гіпотез. Питання про доцільність бракування значення, що «вискакує», вирішується шляхом порівняння його з результатами решти спостережень. Для цього використовується формула

$$\tau = \frac{|x^* - \bar{x}|}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}}, \tag{35}$$

де x* - значення, що «вискакує»; x_i і $\bar{x}$ - відповідно результат i-го вимірювання і середній результат n - 1 вимірювань (без урахування того, що «вискакує»); n - загальна кількість «правильних» вимірювань.

Знайдене за формулою (35) значення τ порівнюється з табличним. Якщо $\tau \geq \tau_т$ для обраного рівня значущості й об'єму вибірки, то нульова гіпотеза відкидається і значення x*, що розглядається, вважається грубою помилкою і з подальшої обробки виключається, оскільки не є характерним для даної ситуації [7, 15]. Проте у багатьох випадках залишати без уваги цей факт (особливо при неодноразовій його появі) не варто. Наявність значень, що «вискакують», є нехарактерною для даного досліду, проте може свідчити про наявність деяких прихованих (неявних) причин, що не враховуються поки в процесі дослідження. У цих випадках аналіз похибок вимірювань доцільно доповнити методом аналізу одиничного випадку [5].

Критерій Макнамари (М-критерій) застосовується для виявлення схожості-відмінності двох спряжених (залежних) вибірок за рівнем вираженості будь-якої ознаки. Проаналізовані дані повинні бути подані в дихотомічній шкалі. Для їх обробки будується таблиця, аналогічна табл. 4. Спосіб обчислення М-критерію залежить від суми величин b+c. Якщо $b + c \leq 20$, то знаходяться значення $m = \min(b, c)$ і $n = b + c$. Для знайдених значень m і n за спеціальними таблицями обчислюється емпіричне значення M_e [4]. Це значення порівнюється з критичним, що дорівнює $M_{кр}$ =0,025 для α =0,05 або $M_{кр}$ =0,005 для α =0,01.

Якщо b+c>20, емпіричне значення М знаходиться за формулою

$$M_e = \frac{(b-c)^2}{b+c}. \tag{36}$$

Це значення порівнюється з критичним, що визначається розподілом ς^2 для $df = 1$, і таким, що дорівнює 3,841 для α =0,05 або 6,635 для α =0,01. Якщо $M_e < M_{кр}$, відмінності між вибірками визнаються незначущими для даного рівня значущості. На закінчення звернемо увагу на принципову відмінність критерію Макнамари від інших критеріїв: за статистичними таблицями знаходиться не критичне, а емпіричне значення критерію. Величини $M_{кр}$ є сталими і залежать від вибраного рівня значущості і способу розрахунку критерію [4].

5. Методи багатовимірної статистики при обробці дослідних даних у психології

Розглянуті три групи задач і методи їх розв’язання (статистичний опис результатів дослідження, визначення взаємозв'язку двох і більше змінних, перевірка статистичних гіпотез) складають основний зміст відповідно до описової, кореляційної та перевірної статистики. Вони охоплюють переважну більшість методів статистичної обробки одновимірних даних. Проте можна ще окремо виділити і методи обробки багатовимірних даних. Однак вони часто не характерні для психології, оскільки тут багато об'єктів і явищ є багатовимірними, тобто оцінюються не поодинці, а за сукупністю ознак. Для обробки таких даних використовуються методи багатовимірної статистики.

Щодо застосування таких методів потрібно зробити ряд попередніх зауважень. Багатовимірні методи достатньо складні як з погляду їх освоєння і розуміння, так і з погляду обчислювальної роботи. Останнє, правда, може бути зменшене шляхом застосування персональних комп'ютерів і відповідного програмного забезпечення. Багатовимірні методи набагато менше формалізуються, ніж одновимірні, тому їх грамотне і правильне застосування багато в чому залежить від досвіду, інтуїції та знань дослідника, розуміння ним суті вирішуваної задачі, можливостей її коректування в ході рішення, уміння інтерпретувати одержувані проміжні результати, вносити необхідні корективи в процес аналізу даних. Це пов'язано з тим, що будь-який багатовимірний метод потребує циклічної обробки даних, де на кожному етапі сам дослідник повинен ухвалювати рішення про характер обробки. Все це вносить істотний суб'єктивний елемент в процес багатовимірного аналізу. Для його проведення дослідник мусить обов'язково знати загальне значення методу, що використовується, вимоги до початкових даних і основні показники для інтерпретації одержуваних результатів. Крім цього, слід мати на увазі, що реалізація будь-якого методу потребує, як правило, застосування комп'ютера і відповідного програмного забезпечення. Необхідно також чітко розуміти, що задачу багатовимірного аналізу (як і будь-яку іншу задачу обробки дослідних даних із застосуванням комп'ютера) вирішує людина, а комп'ютер з програмним забезпеченням є лише інструментом в її руках.

За допомогою багатовимірних методів розв'язуються такі задачі, як структуризація емпіричної інформації (факторний аналіз, багатовимірне шкалювання), класифікація (кластерний аналіз), екстраполяція (множинний регресійний аналіз), розпізнавання образів (дискримінантний аналіз) та ін. [10]. Коротко розглянемо суть основних з цих методів.

Факторний аналіз застосовується для скорочення числа початкових змінних за рахунок об'єднання їх в деякі сукупності, що виступають як цілісні одиниці, які характеризують об'єкт, що вивчається. Ці одиниці в даному випадку називаються факторами, від яких потрібно відрізняти фактори дисперсійного аналізу, що є окремими ознаками (змінні). Вважається, що саме сукупність ознак у певних комбінаціях може характеризувати явище або закономірності його розвитку, тоді як окремо або в інших комбінаціях ці ознаки не дають інформації [11]. Фактори, як правило, приховані від безпосереднього спостереження, в явному вигляді не виявляються. Тому їх у факторному аналізі називають також латентними змінними, латентною (прихованою) суттю.

Початковим матеріалом для факторного аналізу служать коефіцієнти взаємної (парної) кореляції початкових змінних одна з одною. Ці коефіцієнти зводяться в кореляційну матрицю, або інакше – в матрицю інтеркореляцій. Ця матриця завжди квадратна (розміром nxn, де n - кількість початкових змінних) і симетрична. За допомогою спеціальних прийомів вона піддається факторизації, тобто з неї витягується деяке число m факторів ($m<n$). Фактор є штучним статистичним показником, що виникає в результаті спеціальних перетворень матриці інтеркореляцій. Проблема факторизації не має однозначного вирішення і залежить від досвіду і умінь дослідника. В результаті факторизації скорочується число аналізованих змінних, замість n їх стає m. Одержані фактори можуть мати однозначну психологічну інтерпретацію, але можуть її і не мати (наприклад, чинник G в теорії інтелекту). Одержана в результаті факторизації матриця має розмірність nxm, де n - число початкових змінних, m - число виділених факторів. Ці фактори і є предметом подальшого аналізу.

Ще раз відзначимо, що фактор треба розглядати як штучне поняття, що визначає угруповання змінних (ознак) на підставі зв'язків, що є між ними. Це поняття є умовним, тому що, змінивши певні умови (а вони звичайно суб'єктивні) процедури факторизації матриці інтеркореляцій, можна одержати іншу матрицю фактора (структуру). У новій матриці можуть стати іншими розподіл змінних по факторах і їх факторні навантаження. І багато що тут залежить від досвіду і уміння дослідника, розуміння ним суті вирішуваної задачі. Якщо всього цього у нього буде недостатньо, то одержане рішення може виявитися бездоганним з формально-логічної, але безглуздим зі змістовної психологічної точок зору. Таке рішення може бути не просто безглуздим, але й шкідливим, оскільки примушуватиме дослідника інтерпретувати те, що взагалі не піддається інтерпретації. Умови, що приводять до появи таких рішень, розглянуто в [19].

У зв'язку з цим у факторному аналізі існує поняття «проста структура»; у ній кожна змінна має значущі навантаження тільки по одному з факторів, а самі фактори ортогональні, тобто не залежать один від одного. Матриця фактора з простою структурою дозволяє провести змістовну інтерпретацію одержаного результату і дати найменування кожному фактору [4]. Крім цього, з неї можна витягнути додаткові характеристики. Вони називаються *спільність* і *власне значення* фактора. Спільність дозволяє визначити, яка частина варіативності тієї або іншої змінної пояснюється тим чи іншим фактором; за допомогою власного значення фактора можна визначити, яку частину варіативності (дисперсії) пояснює кожен фактор. Основним прийомом отримання простої структури є ротація (обертання) факторів. Мета ротацій - одержати таку факторну структуру, в якій кожен фактор має деяку кількість великих навантажень і деяку – малих. У результаті цього кожна змінна має істотні навантаження по декількох факторах. Це істотно полегшує інтерпретацію факторів і визначення їх значущості. Відповідна цим умовам структура фактора, як вже згадувалося, називається простою.

Проведення факторного аналізу можливе лише при виконанні цілого ряду обмежуючих умов. По-перше, оскільки початковою для факторного аналізу є матриця інтеркореляцій, то перш за все мають бути виконані умови, що забезпечують розрахунок коефіцієнтів лінійної кореляції (див. вище). По-друге, в початковій кореляційній матриці мають бути декілька кореляцій за модулем вище 0,3, інакше достатньо важко витягнути з матриці будь-які фактори. По-третє, вибірка випробуваних має бути достатньо великою. В ідеальному випадку, вона повинна бути не менше ста, інакше виявляться дуже великі стандартні помилки кореляції. Проте якщо фактори добре визначені (наприклад, з навантаженнями 0,7, а не 0,3), досліднику потрібна менша вибірка, щоб виділити їх [10]. Але у будь-якому випадку об'єм вибірки повинен не менше ніж в три рази перевищувати початкову кількість ознак, що аналізуються [4]. Докладніші відомості щодо застосування факторного аналізу в психології наведено в працях [4, 10, 18, 19].

Дискримінантний аналіз застосовується, коли дослідник має множину об'єктів (наприклад випробуваних), розділених на групи. Для кожного з об'єктів виміряно ряд кількісних характеристик. Тоді за допомогою дискримінантного аналізу можна вирішити дві задачі: а) визначити, які змінні краще за все підходять для віднесення об'єкта до однієї з цих груп; б) зарахувати до однієї з цих груп «новий» об'єкт, для якого відомі лише значення виміряних для нього змінних. Дискримінантний аналіз називають також класифікацією з навчанням або розпізнаванням образів [10].

Кластерний аналіз вирішує задачу побудови класифікації, тобто розділення початкової множини об'єктів на групи (класи, кластери). У цьому плані кластерний аналіз є процедурою впорядковування об'єктів у порівняно однорідні класи на основі попарного порівняння цих об'єктів за заздалегідь визначеними і виміряними ознаками. Синонімами кластерного аналізу є також автоматична класифікація, таксомонічний аналіз, аналіз образів (без навчання). Основу кластерного аналізу складають ієрархічні алгоритмічні методи, в яких класифікація здійснюється шляхом послідовного об'єднання об'єктів в групи, що виявляються в результаті ієрархічно організованими [10, 11].

Багатовимірне шкалювання (БШ) полягає у виявленні структури досліджуваної множини об'єктів. У цьому значенні БШ близьке до мети факторного і кластерного аналізів. Проте на відміну від факторного, але подібно кластерному аналізу початковою інформацією для БШ є дані про близькість (схожість) об'єктів між собою. Початковими даними для БШ є суб'єктивні думки випробуваних (експертів) про схожість (близькість) стимулів (об'єктів). Центральне положення БШ полягає у тому, що в основі таких думок лежить обмежене число суб'єктивних ознак, що визначають розрізнення стимулів, і людина, явно чи неявно виносячи свої думки, враховує ці ознаки. Ґрунтуючись на цьому положенні, розв'язується головна задача БШ - реконструкція психологічного простору, заданого невеликим числом вимірювань (шкал), і розташування в ньому точок (стимулів) так, щоб відстані між ними найкращим чином відповідали початковим суб'єктивним відмінностям [10,11].

Множинний регресійний аналіз (МРА) полягає у вивченні взаємозв'язку однієї змінної (залежної) від декількох інших змінних (незалежних). При цьому хоча б одна з них повинна носити випадковий характер. Найчастіше метод застосовується для прогнозу результату (навчання, діяльності) за рядом заздалегідь виміряних характеристик.

Для цього будується рівняння множинної регресії

$$Y = a + b_1 x_1 + b_2 x_2 + ... + b_n x_n + e, \qquad (37)$$

де y - залежна змінна (наприклад, успішність навчання або діяльності); x_1, $x_2, ... x_n$ - психологічні показники; b_1, $b_2, ... b_n$ - параметри моделі (коефіцієнти регресії); e - помилка прогнозу.

Підставляючи в це рівняння параметри моделі (а вони звичайно знаходяться за методом якнайменших квадратів), дослідник одержує очікуваний прогноз з відомою похибкою. Проведення МРА пов'язане з виконанням ряду умов. Неприпустимо використовувати змінні, що мають лінійний зв'язок одна з одною.

Відповідно неприпустимі змінні, коефіцієнт кореляції яких один з одним дорівнює одиниці. Самі ж незалежні змінні мають бути виміряні в шкалі інтервалів і мати нормальний розподіл.

Бажано відбирати для МРА незалежні змінні, що сильно корелюють із залежною змінною і слабко - один з одним [4, 10].

Множинний кореляційний аналіз на відміну від визначення коефіцієнта парної лінійної кореляції за формулою (9) проводиться з метою визначення ступеня взаємозв'язку в багатовимірній системі, що характеризується ознаками x_1, $x_2,\ldots x_m$. У цьому випадку може виявитися, що одна випадкова величина, наприклад x_1, залежить від декількох інших (x_2, $x_3,\ldots x_m$). Для вивчення такого роду залежностей використовується коефіцієнт множинної кореляції, що визначається за формулою [7]

$$r_{1(2,3,\ldots m)} = \sqrt{1 - \frac{Q_1}{Q_{11}}}\ , \qquad (38)$$

де Q - детермінант кореляційної матриці (поняття про неї було дано при описі факторного аналізу); Q_{11} - мінор цього детермінанта, одержаний викреслюванням першого стовпця і першого рядка.

В окремому випадку при n=3

$$r_{1(2,3)} = \sqrt{\frac{r_{12} + r_{13} - 2r_{12}r_{13}r_{23}}{1 - r_{23}^{2}}}\ . \qquad (39)$$

Коефіцієнт множинної кореляції є мірою лінійного взаємозв'язку n випадкових величин, виміряних за шкалою інтервалів. Він завжди має додатний знак і може набувати значеннь від 0 до 1. Крім того, за своєю величиною він завжди більше будь-якого з парних коефіцієнтів кореляції, що визначаються за формулою (9), а його мала величина може свідчити про відсутність лінійного зв'язку. В цьому випадку потрібно використовувати поняття коефіцієнта множинної детермінації [19]. Для перевірки рівня достовірності слід скористатися формулою (25) і таблицею критичних значень для t-критерію Стьюдента при df=n-3 [4]. Крім коефіцієнта множинної кореляції в багатовимірному аналізі використовують також (парціальні) коефіцієнти кореляції. Вони характеризують тісноту зв'язку між випадковими величинами систем (x_1, $x_2,\ldots x_m$) при виключенні (елімінації) впливу інших величин [4, 7, 10, 19]. В окремому випадку при m=3 парціальний коефіцієнт кореляції, що характеризує тісноту зв'язку між x_1 и x_2 при усуненні впливу величини x_3, має вигляд

$$r_{12(\bar{3})} = \frac{r_{12} - r_{23} r_{13}}{\sqrt{(1 - r_{23}{}^2)(1 - r_{13}{}^2)}} \qquad (40)$$

Аналогічно шляхом перестановки відповідних індексів будуються коефіцієнти частинної кореляції між x_1 і x_3 (при сталій x_2) і x_2 і x_3 (при сталій x_1). Перевірка значущості цих коефіцієнтів здійснюється за формулою (25) для t-критерію Стьюдента з числом степенів вільності df=n-2 (де n - об'єм вибірки) [4, 10].

6. Загальні рекомендації щодо вибору методів обробки дослідних даних

Таким чином, загальний алгоритм обробки дослідних даних в психологічному дослідженні зводиться до такого. Перш за все визначаються конкретні задачі, що підлягають обробці (статистичний опис вибірки, визначення взаємозв'язку змінних, перевірка статистичних гіпотез). Потім за допомогою рис. 1 з'ясовуються можливості застосування тих або інших вимірювальних процедур (шкал) для отримання дослідних даних. За необхідності здійснюється зниження потужності шкали (приклад його показано в табл. 1). Після цього (якщо задача не обмежується тільки описом дослідних даних) за першим стовпцем табл. 3 визначається, яка конкретно задача стоїть у даному дослідженні, а за другим стовпцем уточнюється, які умови її розв'язання.

Після цього за допомогою третього стовпця вибираються конкретні методи вирішення. При цьому слід враховувати розглянуті вище методичні рекомендації. Потім за допомогою табл. 4 визначаються умови (можливості) застосування обраного методу і перевіряється відповідність реальним умовам проведення дослідження. Якщо вони не відповідають можливостям застосування даного методу, слід вибрати інший.

Дану роботу рекомендується проводити до початку проведення дослідження; це дозволить краще спланувати і одержати саме ті дані, які потрібні для вирішення поставленої дослідницької задачі. Розглянутий алгоритм вибору методів обробки не завжди може бути реалізований відразу, з першого разу. У ряді випадків на якомусь проміжному етапі може виникнути необхідність додаткових уточнень, тоді слід повернутися назад і далі продовжити роботу щодо вибору методів обробки з урахуванням одержаних уточнень. Після остаточного вибору алгоритму і методів обробки за допомогою необхідної літератури (вона дана в останній колонці табл. 3) можна приступити до роботи з обробки одержаних даних.

При використанні особливо складних з обчислювальної точки зору параметричних методів (наприклад, кореляційний, дисперсійний аналізи) і також багатовимірних методів доцільно використовувати стандартні машинні програми типу Excel і Statistica.

Проте можливість застосування таких методів повинна бути у кожному випадку цілком обґрунтованою у зв'язку з їх обмеженими можливостями використання в психологічних дослідженнях. Це зумовлено тим, що ці методи є параметричними, а психологічні дані параметричності не задовольняють.

Тому застосування таких методів може привести до серйозних помилок. На закінчення відзначимо, що опис і аналіз результатів психологічного дослідження не повинні проводитися тільки на якісному рівні, а обов'язково доповнюватися якісним аналізом із застосуванням методів математичної статистики. Саме таким чином можуть бути подолані елементи суб'єктивізму, неминуче характерні для якісного опису.

При кількісному аналізі сформульовані судження і висновки, гіпотези, що перевіряються, стають більш незалежними від особи дослідника, при цьому забезпечується можливість їх перевірки. Тому в даний час знання існуючих прийомів обробки і аналіз результатів проведених досліджень за допомогою статистичних показників є обов'язковими для психолога. Сказане, природно, не заперечує необхідності якісного опису і аналізу, а потребує додатково, крім цього, і якісного аналізу. Проте слід мати на увазі, що тільки знайомство з даним навчальним посібником не може замінити студенту необхідності систематичного вивчення математичної статистики [2].

Тому, якщо ви хочете вийти далеко за межі основного курсу щодо застосування математичних методів у психології, необхідно додатково виконати таке [8].

По-перше, слід навчитися використовувати складніші статистичні методи для планів, що мають декілька незалежних змінних і комбінацій внутрішньогрупових і міжгрупових змінних. У цьому випадку неоціненну допомогу можуть надати методи багатовимірної статистики.

По-друге, необхідно навчитися користуватися комп'ютером і пакетами прикладних статистичних програм, щоб економити час і зусилля, розширити можливості застосування статистичних методів у психології.

По-третє, розуміння концепцій, що лежать в основі статистичних методів, дозволить не тільки вибрати найприйнятніший метод аналізу одержаних даних, але й планувати майбутнє дослідження так, щоб ці дані можна було ефективно аналізувати та використовувати. Іншими словами, важливо не тільки одержати ті або інші дані, але й зуміти скористатися ними для досягнення мети дослідження. Для досягнення цього слід чітко розуміти, що статистичний аналіз (обробка одержаних даних) і планування дослідження тісно зв'язані між собою.

По-четверте, застосовуючи персональний комп'ютер і статистичні програми (це зараз є обов'язковим для психолога), слід чітко розуміти, що ту або іншу дослідницьку задачу вирішує дослідник, а не машина. Остання є лише інструментом в його руках, яким можна грамотно (або навпаки – не письмово) скористатися. Для цього потрібно розуміти, що робить комп'ютер, пам'ятати про допущення, обмеження та можливості вживаних статистичних методів. Тому рекомендується, щоб студент хоча б один раз виконав кожну статистичну програму вручну, перш ніж скористатися комп'ютером, для розуміння того, як розв'язується та або інша статистична задача [8].

Бібліографічний список

1. Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психодагностике. – СПб.: Питер, 2001. – 528 с.
2. Гайда В.К. Приёмы измерений и статистические способы обработки их результатов в психологическом исследовании: Практикум по общей, экспериментальной и прикладной психологии. – СПб.: Питер, 2001. – С. 5 – 32.
3. Глас Дж., Стенли Дж. Статистические методы в психологии и педагогике. – М.: Прогресс, 1976. – 426 с.
4. Ермолаев О.Ю. Математическая статистика для психологов. – М.: Флинта, 2002. – 336 с.
5. Климов Е.А.Общая психология. – М.:ЮНИТА –ДАНА,2001.– 511 с.
6. Крачивец А.Н., Шишкин Е.В., Дьячков А.Г. Математика для психологов. – М.: Флинта, 2003. – 376 с.
7. Лямец В.И. Методы статистического анализа. – Х.: ХВВКИУ, 1988. – 227 с.
8. Мартин Д. Психологические эксперименты. Секреты механизмов психики. – СПб.: Прайм - Еврознак, 2002. – 480 с.
9. Манеров В.Х. Математическое обеспечение психологических исследований. – СПб.: РГПУ им. А.И. Герцена, 2005. – 144 с.
10. Наследов А.Д., Тарасов С.Г. Применение математических методов в психологии. – СПб.: Изд-во СПбУ, 2001. – 208 с.
11. Никандров В.В. Экспериментальная психология. – СПб.: Речь, 2003. – 408 с.
12. Ноос И.Н. Руководство по психодиагностике. – М.: Ин-т психиатрии, 2005. – 688 с.
13. Психологическая диагностика / Под ред. М.К. Ахимовой и К.М. Гуревача. – СПб.: Питер, 2006. – 652 с.
14. Пустыльник Е.И. Статистические методы анализа и обработки наблюдений. – М.: Наука, 1968. – 312 с.
15. Румшиский Л.З. Математическая обработка результатов эксперимента. – М.: Наука, 1971. – 168 с.
16. Сидоренко Е.В. Методы математической обработки в психологии. – СПб.: Речь, 2002. – 350 с.
17. Смирнов Б.А., Королёв А.В. Практикум по инженерной психологии. – Х.: Нац. арокосм.ун-т «ХАІ», 2004. – 164 с.
18. Суходольский Г.В. Основы математической статистики для психологов. – 2-е изд. – СПб.: Изд-во СПбУ, 1998. – 418 с.
19. Суходольский Г.В. Лекции по высшей математике для гуманитариев. – Х.: Гуманитарный центр, 201. – 630 с.

ЗМІСТ

Смірнов Борис Анатолійович

Научитель Олена Давидівна

Костенко Лілія Іванівна

МАТЕМАТИЧНІ МЕТОДИ В ПСИХОЛОГІЇ

Редактор С.П. Гевло

Зв. план, 2007
Підписано до друку 12.11.2007
Формат 60x84 1/16. Папір офс. № 2. Офс. друк.
Ум. друк. арк. 3,3. Обл.-вид. арк. 3,5. Наклад 100 прим.
Замовлення 526. Ціна вільна

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»
61070, Харків-70, вул. Чкалова, 17
http://www.khai.edu
Видавничий центр «ХАІ»
61070, Харків-70, вул. Чкалова, 17
izdat@khai.edu