

ОХОРОНА ЗДОРОВ'Я

© І.М. Рогач, А.О. Керецман, А.Д. Сіткар, 2017

УДК 614.1+311.21:61:001.891+303.832/.833.5

І.М. РОГАЧ, А.О. КЕРЕЦМАН, А.Д. СІТКАР

Ужгородський національний університет, медичний факультет, кафедра соціальної медицини та гігієни, Ужгород

ПРАВИЛЬНО ВИБРАНИЙ МЕТОД СТАТИСТИЧНОГО АНАЛІЗУ – ШЛЯХ ДО ЯКІСНОЇ ІНТЕРПРЕТАЦІЇ ДАНИХ МЕДИЧНИХ ДОСЛІДЖЕНЬ

Вказано типові помилки при виборі статистичних методів для аналізу медичних даних, їх проведення, шляхи вирішення проблемних питань та їх усунення. Проаналізовано більше 100 методів аналізу даних, проведено класифікацію критеріїв щодо статистичних задач, які вони вирішують (за типами даних, нормальністю розподілу, кількістю груп), створено алгоритм вибору методу у вигляді таблиць. Визначено нульову та альтернативну гіпотези щодо найпоширеніших методів статистичного аналізу, які використовуються в медичних, біологічних та соціологічних наукових дослідженнях.

Ключові слова: статистика, методи аналізу даних, статистичне дослідження, достовірність результатів досліджень, статистичні гіпотези

Вступ. Недосконалість або неправильність статистичного аналізу веде до неправильної інтерпретації даних наукового дослідження (особливо при замовленні статистичної обробки та нерозуміння суті методу аналізу), помилкових результатів та висновків, подальше використання яких веде до виникнення хибних тверджень в науці та медицині. Такий аспект особливо важливий для доказової медицини, на засадах якої впроваджуються нові методи діагностики, лікування та профілактики, що ґрунтуються на принципах охорони здоров'я в цілому.

Особливу актуальність визначає перехідний період до ринкових відносин, де медична галузь має величезні проблеми тіньової економіки, яка визначає специфіку надання медичних послуг. Лікарі не відчувають достатньої мотивації в якості надання медичної допомоги, причиною чого є незнання та нерозуміння суті статистичного аналізу, на основі якого визначаються показники діяльності системи охорони здоров'я та параметри стану здоров'я населення.

Мета дослідження. Визначити типові помилки при виборі та застосуванні статистичних методів аналізу медичних даних, шляхи вирішення проблемних питань та їх усунення, з метою розробки алгоритму вибору доцільного методу проведення статистичного дослідження.

Матеріали та методи. Використано бібліосемантичний метод та контент-аналіз літератури з питань дослідження, в яких представлено най-

більш поширені методики статистичної обробки медичних даних.

Результати досліджень та їх обговорення. Однією з перших проблем, з якою стикається науковець, є визначення необхідного для дослідження об'єму вибірки. Для цього доцільно використовувати спеціальні швидкі формули, формулу Лера, номограму Альтмана [5]. При цьому, слід пам'ятати, що хоча швидкі формули є простими у виконанні, вони повинні відповідати певному статистичному методу (наприклад, t-критерію Стьюдента або критерію χ^2 -Пірсона). Тобто дослідник повинен наперед знати, з яким типом даних він буде працювати, яким є розподіл цих даних та який метод аналізу він хоче використати. Проте на практиці через певні обставини не завжди вдається набрати необхідний об'єм досліджуваних одиниць. У такій ситуації потрібно використовувати непараметричні критерії, які, за певних умов, дають вищу достовірність та можуть бути більш потужними, ніж параметричні. Найдоступнішим та прийнятним із вказаних методів є номограма Альтмана.

Друга проблема – перевірка нормальності розподілу, який є необхідною умовою для використання параметричних критеріїв. Для його визначення доступним і простим є використання коефіцієнтів асиметрії та ексцесу [4]. Найбільш достовірним методом для перевірки нормальності вважається критерій Шапіро-Уїлка. Так, методом аналізу, який використовується у більш ніж 50% наукових досліджень, є t-критерій Стьюдента (табл. 1).

Таблиця 1

Методи аналізу кількісних даних при нормальному їх розподілі (параметричні критерії)

Порівняння середніх значень	Одна група		Одновибірковий t-кр Стьюдента
	Дві групи	Пов'язані	Парний t-кр Стьюдента
		Незалежні	Непарний t-кр Стьюдента: а) при рівних дисперсіях; б) при нерівних дисперсіях; в) без припущення про дисперсії
	Більше двох груп	Пов'язані	Однофакторний дисперсійний аналіз повторних вимірювань; двофакторний дисперсійний аналіз із повторними вимірюваннями
		Незалежні	Однофакторний дисперсійний аналіз без повторів (однакові або різні розміри груп); двофакторний дисперсійний аналіз без повторів (однакові або різні розміри груп); багатофакторний дисперсійний аналіз; кр Діксона; LSD-кр; метод латинських квадратів
	Множинні порівняння груп	Між собою	кр Стьюдента з поправкою Бонферроні; кр Ньюмена-Кейсла (параметричний); кр Тьюкі; Кр Шеффе
З контролем		кр Стьюдента з поправкою Бонферроні; кр Даннета (параметричний)	
Порівняння дисперсій	Дві групи	F-кр Фішера	
	Більше двох груп	Кр Бартлетта; Q(G)-кр Кохрена / Кочрена (при однакових розмірах груп)	
Аналіз взаємозв'язку між явищами	Два ряди змінних	Коеф кореляції r-Пірсона	
	Більше двох рядів змінних	Параметричний дисперсійний аналіз (кр Фішера); багатомірний параметричний дисперсійний аналіз	

Примітки: кр – критерій, коеф – коефіцієнт.

Однак, t-критерій Стьюдента дуже чутливий до нормальності розподілу даних і має багато модифікацій: одновибірковий, парний, непарний (при рівних або нерівних дисперсіях, без припущення про дисперсії) [3]. При цьому, слід пам'ятати, що просте усереднення результатів з використанням довірчих інтервалів не дає змогу стверджувати, що існує статистично значуща відмінність між двома групами (це не критерій Стьюдента, це його розподіл)!

При порівнянні показників двох незалежних груп, які не підпорядковуються нормальному розподілу, використовують критерій Манна-Уїтні [6]. Хоча він і є непараметричним методом, але має певні вимоги – бажана рівність об'єму груп та рівність дисперсій (за критерієм Левене, а у випадку нормального розподілу даних – за критерієм Фішера). На практиці цими вимогами часто нехтують, що призводить до подальших помилок.

Іноколи виникає ситуація, коли потрібно не просто оцінити різницю між декількома досліджуваними групами (за допомогою дисперсійного аналізу), а виявити, де саме є різниця. Для цього використовують методи множинних порівнянь, тобто попарне порівняння між групами. Типовою помилкою в таких випадках стає використання більшості дослідників t-критерію Стьюдента. Але якщо проаналізувати доцільність його використання, то виявляється, що при кожному попарному порівнянні достовірність результатів (при рівні значущості $p < 0,05$) знижується на 5%. Тобто попарно порівнюючи три групи, рівень достовірності дося-

гає лише 85%. Для вирішення даної проблеми необхідно використовувати спеціальний t-критерій Стьюдента з поправкою Бонферроні [1] або критерій Ньюмена-Кейсла. У випадках попарного порівняння груп з контролем необхідно використовувати критерій Даннета. Останні два методи мають однойменні непараметричні аналоги.

Однофакторний дисперсійний аналіз також має чіткі вимоги: рівність дисперсій між всіма групами та нормальність розподілу в кожній з них. Непараметричним аналогом дисперсійного аналізу повторних вимірювань є критерій рангових сум Фрідмана (табл. 2).

Критерій рангових сум Фрідмана дає змогу оцінити достовірність різниці більше двох груп показників однієї вибірки [3]. Типовим прикладом може стати глюкозотолерантний тест: чи відрізняються рівні глюкози крові через одну та дві години після цукрового навантаження в досліджуваній групі пацієнтів? Проте, якщо завдання дослідження передбачає оцінити, чи є статистично значущою тенденція до зниження, підвищення або коливання рівня глікемії в ході експерименту, то більш доцільно використати критерій Пейджа. Він не просто вказує на достовірність різниці, але й дає змогу виконати поставлене завдання дослідження.

При використанні критерію χ^2 -Пірсона (таблиця 2x2) або критерію χ^2 (великі таблиці спряження) існує вимога, що при побудові таблиць кожна клітинка повинна мати значення не менше 5 (табл. 3).

Методи аналізу кількісних даних при довільному їх розподілі (непараметричні критерії)

Порівняння середніх значень	Одна група		G-кр знаків; Кр Гупта; медіанний одновибірковий кр
	Дві групи	Пов'язані	G-кр знаків; W(T)-кр знакових рангів Вілкоксона
		Незалежні	Двовибірковий кр Вілкоксона рангових сум; U(T)-кр Манна-Уїтні (при рівних дисперсіях); Q-кр Розенбаума; кр серій Вальда-Вольфовіца; двовибірковий кр Колмогорова-Смірнова
		Неоднакові розміри груп	Кр Уайта
	Більше двох груп	Пов'язані	Кр рангових сум Фрідмана; L-кр Пейджа; кр Холлендера
		Незалежні	H-кр Крускала-Уолліса; кр Неменьї; Медіанний кр (тест)
		Неоднакові розміри груп	Медіанний кр (тест)
	Множинні порівняння груп	Між собою	Кр Ньюмена-Кейсла (непараметричний); кр Данна
		З контролем	Кр Даннета (непараметричний)
		Неоднакові розміри груп	Кр Данна
Порівняння дисперсій	Дві групи	Кр Зігеля-Тьюкі; кр Мозеса; Левене	
	Більше двох груп	Кр Фрідмана; кр Левене	
Аналіз взаємозв'язку між явищами	Два ряди змінних	Коеф кореляції $r(\rho)$ -Спірмена; коеф кореляції τ -Кендалла; Q-коеф асоціації Юла; Со-коеф коваріації; ϕ -коеф спряженості, контингенції; V-коеф Крамера	
	Більше двох рядів змінних	Непараметричний дисперсійний аналіз (кр Зігеля-Тьюкі); C-коеф взаємної спряженості Пірсона; k-коеф Чупрова	

Примітки: кр – критерій, коеф – коефіцієнт.

Методи аналізу якісних даних

Порівняння середніх значень	Одна група		G-кр знаків для пропорцій; z-кр однієї пропорції (з поправкою Йейтса)
	Дві групи	Пов'язані	Кр Мак-Немара
		Незалежні	Кр χ^2 -Пірсона: таблиця 2x2 (з поправкою Йейтса); точний кр Фішера (кр Фішера-Ірвіна); z-кр двох пропорцій
	Більше двох груп	Пов'язані	Q(G)-кр Кохрена / Кочрена (при однакових розмірах груп)
		Незалежні	Кр χ^2 -Пірсона: великі таблиці спряження; Кр χ^2 -Пірсона для впорядкованих градацій (для тренда)
Множинні порівняння груп		Процедура Мараскуїло-Ляха-Гурьянова (М-Л-Г)	
Порівняння дисперсій			Дисперсійний аналіз (альтернативний комплекс)
Аналіз взаємозв'язку між явищами	Два ряди змінних	Кр χ^2 -Пірсона: таблиця 2x2 (з поправкою Йейтса); точний кр Фішера (кр Фішера-Ірвіна); кр χ^2 -Пірсона: великі таблиці спряження; коеф кореляції $r(\rho)$ -Спірмена; коеф кореляції τ -Кендалла; тетрагоричний показник зв'язку; полігоричний показник зв'язку	
	Більше двох рядів змінних	Коеф конкордації Кендалла; багатомірний непараметричний дисперсійний аналіз Фрідмана	

Примітки: кр – критерій, коеф – коефіцієнт.

Якщо дане допущення (вимоги при побудові таблиць) не виконується, то потрібно використовувати точний критерій Фішера [1], який вираховується через факторіали.

Щодо аналізу наявності лінійного зв'язку між явищами існує безліч методів (коефіцієнти кореляції r -Пірсона, $r(\rho)$ -Спірмена, τ -Кендалла, Q-коефіцієнт асоціації Юла, C-коефіцієнт взаємної спряженості Пірсона, K-коефіцієнт Чупрова), кожен

з яких потрібно використовувати у певному конкретному випадку, в залежності від очікуваних результатів [6].

Проведений нами контент-аналіз літературних джерел дав змогу класифікувати методи статистичного аналізу відповідно до статистичних задач, які вони вирішують, та розробити спеціальний алгоритм вибору доцільного методу дослідження (рис. 1).

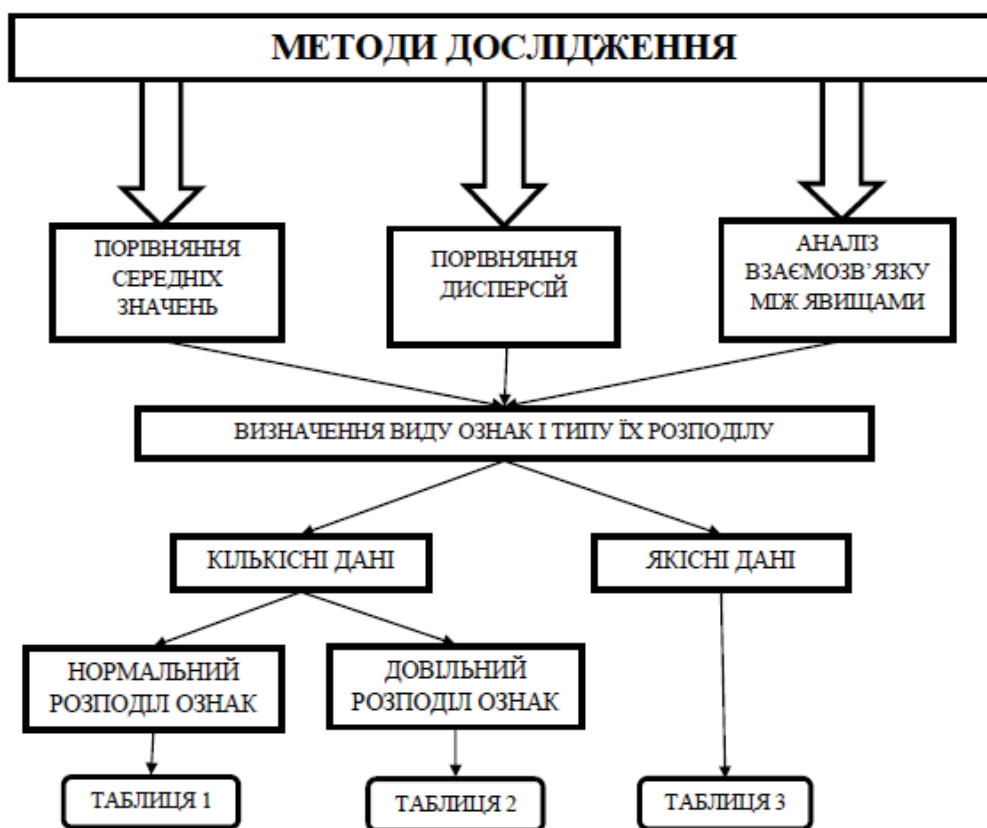


Рис 1. Алгоритм вибору методу статистичного аналізу даних.

Перед тим, як зробити вибір методу, зокрема за допомогою створеного алгоритму, необхідно визначити тип досліджуваних даних, обчислити відносні та середні показники, довірчі інтервали (описова статистика) та оцінити нормальність розподілу даних (табл. 1, 2, 3). Та попри те, що обчислення відносних та середніх величин є початковим етапом статистичного аналізу, важливим моментом залишається визначення доцільності їх використання. Адже, наприклад, у лікарняній палаті, де один пацієнт мертвий і вже холодний, а інший живий і в лихоманці, середня температура тіла пацієнтів буде нормальною [2].

Крім представлених методів статистичного аналізу існує низка спеціальних, вибір яких визначається метою дослідження. До спеціальних методів відносять: когортні дослідження (показник відносного ризику); дослідження випадок-контроль (показник відношення шансів); оцінка згоди, порівняння двох методів виміру (метод Бленда-Алтмана, Каппа (κ) Кохена, критерій згоди χ^2 -Пірсона); методи класифікації даних (LSD-критерій, кластерний аналіз, дискримінантний аналіз); регресійний аналіз (проста, множинна, логістична та Пуассоновська регресія); аналіз виживаності (критерій Гехана, метод Каплана-Мейера, модель пропорційних ризиків Кокса, Логранговий критерій, критерій Пето, стратифікований аналіз, експоненціальна регресія); аналіз нормальності розподілу (коефіцієнт асиметрії та ексцесу, критерій Шапіро-Уїлка, Колмогорова-

Смірнова; ω^2 -Смірнова-Крамера-фон Мізеса, спрощені формули, наближені графічні методи); визначення необхідного об'єму вибірки (швидкі формули, формула Лера, номограма Альтмана, спеціальні таблиці); діагностичні інструменти (чутливість, специфічність, валідність та ефективність тесту); гармонійний аналіз, ряди Фур'є; пробіт-аналіз (визначення летальної дози та коефіцієнта кумуляції); методи послідовного аналізу; метод Байєса; аналіз динамічних (часових) рядів; метод нейронних сіток; стандартизація даних; моделювання.

Далі, в дужках, наведено пояснення / формулювання альтернативної гіпотези H_1 (достовірність якої намагаються довести у більшості випадків) щодо найпоширеніших методів статистичного аналізу, які використовуються в медичних, біологічних та соціологічних наукових дослідженнях [1, 5, 6]: одновибірковий t-критерій Стюдента (середнє значення в популяції не рівне константі (певному значенню)); парний t-критерій Стюдента (різниця середніх у популяції не дорівнює нулю); непарний t-критерій Стюдента (різниця середніх у двох популяціях не дорівнює нулю); однофакторний дисперсійний аналіз без повторів (хоча б одне групове середнє в популяції відрізняється від інших); G-критерій знаків (медіана в популяції не дорівнює константі (певному значенню)); двовибірковий критерій Вілкоксона рангових сум (дві групи мають різний розподіл у популяції); U(T)-критерій Манна-Уїтні (різниця медіан у

популяції не рівна нулю); критерій χ^2 -Пірсона: таблиця 2×2 (пропорції досліджуваних із певною ознакою не рівні у двох популяціях); критерій χ^2 -Пірсона: великі таблиці спряження (два фактори в популяції взаємопов'язані); F-критерій Фішера (дві популяційні дисперсії не рівні); коефіцієнт кореляції r-Пірсона (величина коефіцієнта кореляції (наявність лінійного зв'язку) в популяції не дорівнює нулю); коефіцієнт кореляції r(p)-Спірмена (величина коефіцієнта кореляції (наявність лінійного зв'язку) в популяції не рівна нулю); коефіцієнти асиметрії та ексцесу (використовуються для наближеної перевірки нормальності розподілу даних у вибірці (при нормальному законі розподілу значення коефіцієн-

тів дорівнюють нулю)); показник відношення шансів (спеціальний показник, який використовується в дослідженнях типу «випадок-контроль» і виражає ефект дії чинника ризику).

Таким чином, вибір методу статистичного аналізу залежить від розуміння дослідником поставлених завдань, які сприятимуть досягненню мети дослідження.

Висновки. Не обов'язково знати методику проведення всіх видів статистичного аналізу (а їх є безліч), потрібно зробити правильний вибір методу – такий принцип, на нашу думку, має стати одним із основних щодо наукової діяльності в медичній галузі. При цьому, кожен випадок статистичного дослідження потребує свій, специфічний для досліджуваних явищ та параметрів, метод.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Гланц С. Медико-биологическая статистика / С. Гланц. — М.: Практика, 1998. — 459 с.
2. Ланг Т.А. Как описывать статистику в медицине / Т. Ланг, М. Сесик. — М.: Практическая медицина, 2011. — 480 с.
3. Лапач С.Н. Статистика в науке и бизнесе / С.Н. Лапач, А.В. Чубенко, П.Н. Бабич. — К.: МОРИОН, 2002. — 640 с.
4. Мінцер О.П. Оброблення клінічних і експериментальних даних у медицині / О. П. Мінцер, Ю.В. Вороненко, В.В. Власов. — К.: Вища школа, 2003. — 350 с.
5. Петри А. Наглядная медицинская статистика / А. Петри, К. Сэбин. — М.: ГЭОТАР-Медиа, 2009. — 168 с.
6. Статистика в медицине и биологии: руководство в 2 т. / [сост. Медик В.А., Токмачев М.С., Фишман Б.Б.; ред. Комарова Ю.М.]. — Т. 1. — М.: Медицина, 2000. — 412 с.

I.M. RONACH, A.O. KERETSMAN, A.D. SITKAR

Uzhhorod National University, Medical Faculty, Department of Social Medicine and Hygiene, Uzhhorod

CORRECT CHOICE OF STATISTICAL ANALYSIS METHOD IS THE KEY WAY TO HIGH-QUALITY INTERPRITATION OF DATA OF MEDICAL RESEARCH

The typical errors in the choice of statistical methods for the analysis of medical data, their carrying out, ways of solving problematic issues and their elimination are indicated. More than 100 methods of data analysis have been analyzed, the classification of criteria with regard to statistical tasks that they solve (by data types, normality of distribution, number of groups) has been carried out, an algorithm of method selection in the form of tables has been created. The zero and alternative hypotheses concerning the most widely used methods of statistical analysis, which are used in medical, biological and sociological research, have been determined.

Key words: statistics, methods of data analysis, statistical research, reliability of research results, statistical hypotheses

Стаття надійшла до редакції: 28.08.2017 р.