УДК 517(075.8)

Кавецька Анастасія Вячеславівна[1]
Сачанюк-Кавецька Наталія Василівна[2]

МЕТОД ШЕФФЕ В СТАТИСТИЧНІЙ ОБРОБЦІ МЕДИЧНИХ ДАНИХ

[1] Вінницький національний медичний університет ім. М. І. Пирогова
[2] Вінницький національний технічний університет

Анотація

Статистичний аналіз є одним з елементів доказової медицини, оскільки питання, що виникають у лікарів при прийнятті рішення, щодо діагностики та лікування, стають все складнішими. Знайти ці рішення можна лише після обробки й аналізу достатньо великих інформаційних масивів. Формулювання мети, вибір методу і способу організації дослідження, особливості відбору пацієнтів і характер отриманих даних – все це визначає вибір адекватного методу статистичного аналізу і впливає на достовірність отриманих результатів. В роботі розглянуто можливість застосування методу Шеффе в статистичній обробці даних для з'ясування контрастів щодо залежності швидкості осідання еритроцитів (ШОЕ) від кліматичної зони.

Ключові слова: статистика Фішера-Снедекора, метод Шеффе, дисперсійний аналіз.

Abstract

Statistical analysis is one of the elements of evidence-based medicine, as questions that arise in physicians when making decisions about diagnosis and treatment are becoming increasingly complex. These solutions can be found only after processing and analysis of large enough information arrays. Formulation of the purpose, the choice of method and method of organization of research, features of selection of patients and character of the received data - all this defines a choice of an adequate method of the statistical analysis and influences reliability of the received results. The paper considers the possibility of using the Scheffe method in statistical data processing to determine the contrasts of the dependence of erythrocyte sedimentation rate (ESR) on the climatic zone.

Key words: Fischer-Snedekor statistics, Scheffe method, analysis of variance.

Вступ

Статистичні методи в клінічних дослідженнях стали використовуватися відносно недавно, починаючи з 1930-х років. Саме біостатистика [1] відіграла важливу роль у переході медичних досліджень від опису окремих спостережень і серій випадків до проведення експериментальної роботи з використанням контрольних груп і випробувань. Постановка мети, вибір методу і способу організації дослідження, особливості відбору пацієнтів і характер отриманих даних – все це визначає вибір адекватного методу статистичного аналізу і впливає на достовірність отриманих результатів. Суть статистичного підходу полягає в тому, щоб замінити дослідження великої кількості об'єктів вивченням значно меншої її частини та здійснити подальше «поширення» результатів дослідження на всю множину. Одержані висновки мають бути обґрунтованими та надійними, що в основному залежить від якості вихідних даних [2, 3].

У наукових роботах часто доводиться демонструвати статистичні відмінності чи схожості між групами об'єктів дослідження. Порівнюються експериментальні та контрольні зразки, результати лікування різними препаратами, показники у хворих та здорових тощо. Дослідник спочатку постулює нульову гіпотезу про те, що групи не відрізняються. Якщо імовірність того, що ці групи ідентичні, менше 5%, дослідник може впевнено відкинути гіпотезу групової подібності і стверджувати, що групи ймовірно відрізняються. Цей п'ятивідсотковий поріг називається рівнем значущості. Досить часто, саме застосування статистичних методів дозволяє з'ясувати наявність зв'язків між різними показниками чи присутність контрастів щодо залежності досліджуваних показників.

Результати дослідження

Розглянемо ситуацію з дослідженням ШОЕ у людей, що перенесли майже рік тому гостре інгаляційне отруєння парами отруйної речовини, не вдаючись у детальний опис типу речовини

та інших змін показників крові. Постраждалі знаходились в чотирьох різних кліматичних умовах (Спекотний, Помірний, Холодний, Полярний). Результати ШОЕ (y_i , мм/г) наведено в таблиці 1.

Таблиця 1. Початкові дані ШОЕ у людей після гострого отруєння отруйними парами

Спекотний	Помірний	Холодний	Полярний
8	7	5	4
7	6	5	3
6	5	4	3
5	5	4	2
4	4	3	
	3	3	

Суто візуально з таблиці 1 можна припустити, що у потерпілих в умовах спекотного клімату ШОЕ вище, ніж у потерпілих, які мали інтоксикацію в районах з холодних і помірних кліматом. Одержимо статистичний висновок на основі дисперсійного аналізу, який широко застосовується при обробці клінічних лабораторних даних.

Застосуємо статистику Фішера-Снедекора [4] із рівнем значущості $\alpha = 0,05$ та ступенями вільності $k_1 = 4-1=3, k_2 = 21-4=17$, де число 4 – це кількість кліматичних зон, 21 – загальне число статистичних даних. Дана статистика передбачає порівняння критичного значення функції Фішера, знайдене за статистичними таблицями. При цьому, критичне значення функції Фішера становить $f_{кр} = f(p=0,95; k_1=3; k_2=17) = 3,592$. Для зручності використання даної статистики, заповнюємо таблицю проміжних розрахунків (таблиця 2).

Таблиця 2. Проміжні розрахунки для статистики Фішера-Снедекора

Клімат	n_i	$\sum_i y_i$	$\sum_i y_i^2$	$\overline{y_i} = \frac{\sum_i y_i}{n_i}$	$n_i \cdot \left(\overline{y_i}\right)^2$
Спекотний	5	30	190	6	180
Помірний	6	30	160	5	150
Холодний	6	24	100	4	96
Полярний	4	12	38	3	13
Сума	**21**	**96**	**488**	-	**439**

Загальний середній показник : $\overline{Y} = \frac{1}{21}96 = 4,571$. Однофакторний дисперсійний аналіз зручно подавати у вигляді таблиці (табл. 3):

Таблиця 3 Однофакторний дисперсійний аналіз

Компоненти мінливості	Сума квадратів відхилень	Число ступенів вільності	Середній квадрат	Відношення Фішера
Факторна	$SS_A = 343,009$	3	$MS_A = \frac{SS_A}{3} = 114,34$	$f = \frac{MS_A}{MS_W} = 39,7$
Залишкова	$SS_W = 49$	17	$MS_W = \frac{SS_W}{17} = 2,88$	-

Оскільки розраховане значення функції Фішера більше за критичне значення ($f > f_{кр}$), то можна зробити висновок, що ШОЕ істотно відрізняється для різних кліматичних зон на рівні значущості $\alpha = 0,05$.

З’ясуємо контрасти щодо залежності ШОЕ від кліматичної зони. Оскільки обсяги вибірок у групах є різними, то доцільно використати метод Шеффе, який реалізується так.

1. Формулюємо нульову та альтернативну гіпотези, причому H_1 : є ненульові контрасти.
2. Критерієм перевірки є статистика $s_{kl} = \overline{y_k} - \overline{y_l}$, з критичним значенням $s_{кр} = \sqrt{f(p = 1-\alpha; k_1; k_2)(I-1)}$, де I - кількість рівнів досліджуваної ознаки (в нашому випадку їх чотири).
3. За даними експерименту обчислюють емпіричні значення статистичного критерію. Множину емпіричних значень зручно представляти у вигляді таблиці.
4. Якщо емпіричні значення статистики більші критичного значення, то відхиляється нульова гіпотеза й відмінність k - ї та l - ї умов на рівні значущості α вважають істотною.

В нашому випадку $S_{кр} = \sqrt{3 \cdot 3,592} = 3,283$. Складемо матрицю значень даного критерію (табл. 4), додатні різниці більші критичного значення виділяємо.

Таблиця 4. Матриця значень критерію Шеффе

Клімат			Полярний	Холодний	Помірний	Спекотний
	Обсяг вибірки		4	6	6	5
		Середнє значення	3	4	5	6
Спекотний	5	6	3	2	1	-
Помірний	6	5	2	1	-	-1
Холодний	6	4	1	-	-1	-2
Полярний	4	3	-	-1	-2	-3

Висновки

Оскільки жодне із значень критерію Шеффе не більше за критичне значення, то не існує ненульових контрастів. Це означає, що ШОЕ не залежить від кліматичної зони.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Гланц С. Медико-биологическая статистика / С. Гланц. - М. : Практика, 1999. - 461 с.
2. Сергиенко В. И. Математическая статистика в клинических исследованиях / В. И. Сергиенко, И. Б. Бондарева - М. : ГЭОТАР МЕДИЦИНА, 2000. - 256 с.
3. Мінцер О.П. Оброблення клінічних і експериментальних даних у медицині: навч. посібник / О.П. Мінцер, Ю.В. Вороненко, В.В. Власов - К.: Вища шк., 2003. - 350 с.
4. Личковський Е. І. Вища математика. Теорія наукових досліджень у фармації та медицині: підручник / Е. І. Личковський, П. Л Свердан. – К.: Знання, 2021. – 476 с.

Кавецька Анастасія Вячеславівна, Вінницький національний медичний університет, 1-й медичний факультет, 2-й курс, група 13 А, ankavecka@gmail.com
Сачанюк-Кавецька Наталія Василівна, к. .т. н., доцент, Вінницький національний технічний університет, кафедра вищої математики, skn1901@gmail.com

Kavetska Anastasia V. – student, medical faculty, group 13-A, Vinnytsia National Medical University, Vinnytsia, e-mail: ankavecka@gmail.com
Sachaniuk-Kavets`ka Natalia V. Candidate of Technical Sciences, Associate Professor, Department of Higher Mathematics, Vinnytsia National Technical University, skn1901@gmail.com