

КРИТЕРІЙ ЗГОДИ МІЗЕСА-СМІРНОВА В СТАТИСТИЧНІЙ ОБРОБЦІ МЕДИЧНИХ ДАНИХ

¹ Вінницький національний медичний університет ім. М. І. Пирогова

² Вінницький національний технічний університет

Анотація

Математичні методи аналізу даних широко використовують при дослідженні медичних систем і процесів. У клінічних дослідженнях часто виникає необхідність вибрати метод описової статистики та визначити, чи підпорядковуються кількісні дані закону нормального розподілу, оскільки розподіл багатьох статистичних даних є нормальним або може бути отриманий із нормального розподілу за допомогою деяких перетворень. В роботі розглянуто можливість аналітичного з'ясування, чи має ознака, задана варіаційним рядом, нормальний розподіл.

Ключові слова: критерій згоди, варіаційний ряд, мембранний потенціал, рівень значущості, нормальний розподіл

Abstract

Mathematical methods of data analysis are widely used in the study of medical systems and processes. In clinical trials, it is often necessary to choose the method of descriptive statistics and determine whether quantitative data are subject to the law of normal distribution, as the distribution of many statistics is normal or can be obtained from the normal distribution by some transformations. The paper considers the possibility of analytical determination of whether a feature given by a variation series has a normal distribution.

Keywords: agreement criterion, variation series, membrane potential, level of significance, normal distribution

Вступ

Математичні методи аналізу даних широко використовують при дослідженні різноманітних систем і процесів – природних, технічних, екологічних, економічних, соціальних, медичних тощо [1, 2]. Усі медичні тести базуються на порівнянні зразків. Можна порівняти експериментальні та контрольні зразки, результати лікування різними препаратами, показники у хворих і здорових людей, чоловіків і жінок тощо.

У клінічних дослідженнях часто виникає необхідність вибору методу описової статистики та визначення, чи підпорядковуються кількісні дані закону нормального розподілу [3]. Розподіл багатьох статистичних даних є нормальним або може бути отриманий із нормального розподілу за допомогою деяких перетворень. У результаті перевірки вибіркового розподілу було виявлено, що 69 (39,88 %) з них розподілені за нормальним законом, а розподіл 104 (60,12 %) вибірок не є нормальним [4]. Форма нормального розподілу залежить від двох параметрів: математичного сподівання (μ) та середнього квадратичного відхилення (σ^2). Характеристикою нормального розподілу є те, що 68% усіх його спостережень лежать у діапазоні ± 1 стандартне відхилення середнього, а діапазон стандартного відхилення ± 2 містить 95% значень. Іншими словами, у нормальному розподілі нормовані спостереження, менше ніж -2 або більше $+2$ мають відносну частоту менше 5%. Крім того, у нормальному розподілі ознак параметричні критерії є більш потужними, ніж непараметричні, тобто вони мають менше шансів на помилку другого типу. При розподілах, відмінних від нормального, параметричні критерії можуть не застосовуватись. У цьому випадку використовують непараметричні критерії або вибірка змінюється до нормального закону. Знання закону розподілу вибірки дозволяє досліднику правильно вибрати критерії перевірки гіпотез та отримати статистично достовірні дані.

Існує міжнародний стандарт ISO 5479:1997 [5] для перевірки чи мають статистичні дані нормальний закон розподілу. Стандарт враховує графічний метод перевірки, критерії тесту на симетрію та на значення ексцесу, критерії Шапіро-Уїлка та Епса-Паллі. Тому актуальним є застосування критеріїв, які дозволяють аналітично з'ясувати, чи має ознака, задана варіаційним рядом, нормальний розподіл.

Результати дослідження

Розглянемо ситуацію з дослідженням значення мембранного потенціалу м'язової клітини (x , мВ). Якщо позначити кількість досліджуваних з відповідним значенням потенціалу через m_x , то результати утворюють варіаційний ряд, який наведено в таблиці 1.

Таблиця 1. Варіаційний ряд значення мембранного потенціалу м'язової клітини

x , мВ	29	30	31	32	33	34	35	36	37
m_x	1	7	6	9	0	10	8	9	0

Потрібно з'ясувати, чи має місце нормальний закон розподілу досліджуваної ознаки на рівні значущості $\alpha = 0,05$.

Визначаємо спочатку оцінки математичного сподівання й стандартного відхилення:

$$n = \sum_{i=1}^9 m_i = 50; \quad \bar{x} = \frac{1}{n} \sum_{i=1}^9 m_i x_i = 33,14; \quad \overline{x^2} = \frac{1}{n} \sum_{i=1}^9 m_i x_i^2 = 1103; \quad \sigma = \sqrt{\frac{n(\overline{x^2} - \bar{x}^2)}{n-1}} = 2,42. \quad (1)$$

Формулюємо гіпотези і задаємо рівень значущості:

H_0 : мембранний потенціал (мВ) м'язової клітини має функцію нормального розподілу:

$$F_N(x; 33,14; 2,42) = \Phi\left(\frac{x - 33,14}{2,42}\right); \quad (2)$$

H_1 : мембранний потенціал м'язової клітини має інший закон розподілу; рівень значущості $\alpha = 0,05$.

Критерієм перевірки гіпотез обираємо критерій Мізеса-Смірнова [6]. Оскільки параметри нормального розподілу визначені за даними вибірки, то критичне значення визначається за формулою:

$$(n\omega^2)^* = n\omega_1^2(\alpha = 0,05) \left(1 + \frac{1}{2n}\right) = 0,126 \cdot \left(1 + \frac{1}{100}\right) = 0,127. \quad (3)$$

Доповнимо варіаційний ряд, заданий таблицею 1, номерами значень ознаки та значеннями нормальної функції розподілу (див. табл. 2)

Таблиця 2. Доповнений номерами значень та значеннями нормальної функції розподілу варіаційний ряд

x , мВ	m_x	i	$\Phi\left(\frac{x - 33,14}{2,42}\right)$
29	1	1	$\Phi(-1,7) = 0,045$
30	7	$(\overline{2,8})$	$\Phi(-1,3) = 0,097$
31	6	$(\overline{9,14})$	$\Phi(-0,88) = 0,189$
32	9	$(\overline{15,23})$	$\Phi(-0,5) = 0,309$
34	10	$(\overline{24,33})$	$\Phi(0,35) = 0,637$
35	8	$(\overline{34,41})$	$\Phi(0,76) = 0,776$
36	9	$(\overline{42,50})$	$\Phi(1,18) = 0,881$

Розрахуємо емпіричне значення критерію $n\omega_e^2$:

$$n\omega_e^2 = \frac{1}{12 \cdot 50} + \left(0,045 - \frac{1}{100}\right)^2 + \sum_{i=2}^8 \left(0,097 - \frac{2i-1}{100}\right)^2 + \sum_{i=9}^{14} \left(0,189 - \frac{2i-1}{100}\right)^2 + \sum_{i=15}^{23} \left(0,309 - \frac{2i-1}{100}\right)^2 + \sum_{i=24}^{33} \left(0,637 - \frac{2i-1}{100}\right)^2 + \sum_{i=34}^{41} \left(0,776 - \frac{2i-1}{100}\right)^2 + \sum_{i=42}^{50} \left(0,881 - \frac{2i-1}{100}\right)^2 = 0,0016 + 0,001225 + 0,011543 + 0,012766 + 0,057489 + 0,09229 + 0,027168 + 0,031569 = 0,23565.$$

Висновки

Оскільки емпіричне значення більше критичного $n\omega_e^2 > (n\omega^2)^*$, то нульгіпотеза відхиляється і розподіл досліджуваної ознаки не описується функцією розподілу $F_N(x; 33,14; 2,42)$, а отже даний розподіл не має всіх властивостей нормального закону.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. . Громадське здоров'я: підручник для студ. вищих мед. навч. закладів. - Вид. 3 – Вінниця: «Нова книга», 2013. – 560 с.
2. . 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.
3. Medical Statistics at a Glance Text and Workbook. Aviva Petria, Caroline Sabin. – Wiley-Blackwell, 2013. – 288 p.
4. М. А. Іванчук, П. Р. Іванчук Нормальний закон розподілу в медичних дослідженнях/ *Медична інформатика та інженерія*. -№1, 2013. – С. 48-52.
5. ISO 5479:1997(en) Statistical interpretation of data – Tests for departure from the normal distribution: веб-сайт. URL: :<https://www.iso.org/obp/ui/#iso:std:iso:5479:ed-1:v1:en> (дата звернення: 10.05.2022).
6. 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-Kavetska Natalia V. Candidate of Technical Sciences, Associate Professor, Department of Higher Mathematics, Vinnytsia National Technical University, skn1901@gmail.com