

## МЕТОДИ СТАТИСТИЧНОЇ ОБРОБКИ МЕДИЧНИХ ДАНИХ

**Терентьєва В.Ю.**

*Національний авіаційний університет, Київ*

*Науковий керівник: Шевченко І.В., к.е.н., доц.*

Ключові слова: медицина, математико-статистичні методи, полігон частот.

Розглянемо застосування математико-статистичних методів у медицині, зокрема методів обробки медичних даних для медико-соціологічних досліджень. Статистичні методи обробки даних вже давно застосовуються в різних сферах людської діяльності й, насамперед, там, де досліджуються закономірності, властиві великим обсягам об'єктів. Це відбувається й в медичній галузі, де завдяки використанню більш потужної сучасної апаратури та застосуванню нових методів обстеження невпинно зростають інформаційні масиви та кількісні дані про стан здоров'я пацієнта [1].

Проблема аналізу та обробки медичних даних з кожним днем набуває актуальності, оскільки питання, що виникають у лікарів під час прийняття правильного рішення, стають все складнішими. Знайти відповіді можна після обробки й аналізу достатньо великих інформаційних масивів, в основі яких лежать математичні методи [2].

За допомогою методів математичної статистики, можна прорахувати актуальні для медицини проблеми в таких сферах, як: загальна хірургія, травматологія, ортопедія, інфектологія тощо. Ключовим моментом тут є формування вибірки. Так, якщо метою наукового дослідження є з'ясування причини виникнення та розповсюдження певної хвороби в даний час в цілому, то розглядаємо генеральну сукупність (популяцію), тобто все те населення, а не якусь територію чи групу населення яке піддається ризику виникнення цієї хвороби. Якщо метою наукового дослідження є вивчення причини хвороби тільки в даний країні або місті, то за генеральну сукупність приймаємо відповідне населення країни чи міста. Ще менше за обсягом є генеральна сукупність в рутинних аналітичних дослідженнях, наприклад при розслідуванні спалаху хвороби в «організованому» колективі дітей. У цьому випадку, популяція – це всі діти та персонал даної установи, або однієї (кількох) груп, залежно від первісної гіпотези про причини даного спалаху [3].

Зазвичай під час масштабного дослідження можуть виникати певні труднощі, наприклад, великі витрати часу, сил і засобів і, нерідко, неможливість їх проведення. У таких випадках проводять вибіркові епідеміологічні дослідження (sample study), які призначені для того, щоб ґрунтуючись на даних, які отримані при вивченні захворюваності відносно невеликої частини населення – вибірки (sample), зробити висновки про особливості досліджуваного явища у всій популяції (генеральної сукупності), з якої дана вибірка була сформована. Використовуючи полігон частот (відносних частот) (рис. 1), можна отримати наочне уявлення про закономірності можливих змін спостережуваних значень. Полігон часто застосовується у випадках, коли необхідно показати динаміку явища (зміна показників в часі), наприклад зниження захворюваності населення на інфекційні хвороби.

Гістограму частот (відносних частот) можна застосувати для зображення динаміки або іншого стану явища, наприклад, зростання кількості апаратів УЗД за кілька років або зображення рівня захворюваності на діабет у різних країнах за будь-якої рік (статика) (рис. 2).

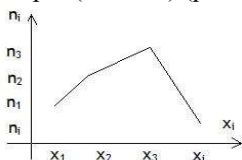


Рис. 1

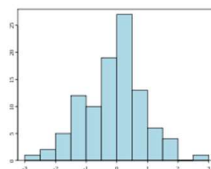


Рис. 2

Досліджуваний масив даних (статистичний розподіл) містить повну інформацію про варіацію ознаки. За великої кількості даних провести аналіз стану певного явища можна за допомогою основних числових характеристик. Так, для обчислення середньої кількості пацієнтів, що викувалися, або захищені від інфекційних захворювань через щеплення, та розсіювання їх значень відносно середньої кількості можна за допомогою таких числових характеристик, як вибіркова середня величин  $\bar{x}_B$ , дисперсія вибірки (вибіркова дисперсія)  $D_B$ , середнє квадратичне відхилення вибірки  $\sigma_B$  :

$$\bar{x}_B = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{n}; D_B = \frac{\sum_{i=1}^k (x_i - \bar{x}_B)^2 n_i}{n}; \sigma_B = \sqrt{D_B}.$$

Застосування сучасних статистичних програм передбачає знання основних методів та етапів статистичного аналізу: їх послідовність, необхідність і достатність. Отже, в медичних дослідженнях, пов'язаних з базами даних, доцільно перетворити медичну проблему на статистичну. Завдяки використанню комп'ютерів і широкому впровадженню сучасних комп'ютерних технологій докорінно змінився процес обробки й аналізу медичних даних. Застосовуючи той чи інший метод статистичного аналізу, необхідно бути впевненим, що допущення, на яких він заснований, виконуються хоча б наближено. Інакше існує велика ймовірність отримати помилкові висновки, незважаючи на правильну послідовність дій.

#### Список використаних джерел:

1. Гойко О. Методичний підхід щодо вибору методу статистичної обробки для медико-соціологічних досліджень / О.В. Гойко // Медична інформатика та інженерія. – 2015. – №2. – С. 52-58.
2. Гойко О. Сучасні технології обробки й аналізу медичних даних / О.В. Гойко // Медична інформатика та інженерія. – 2009. – №4. – С. 39-44.
3. Голованова І. Основи медичної статистики / І. Голованова, І. Белікова, Н. Ляхова – Полтава, 2017. – С. 113.