УДК 519.25

ВИЗНАЧЕННЯ ВПЛИВУ ОБСЯГУ ВИБІРКИ НА ЧУТЛИВІСТЬ СТАТИСТИЧНОГО КРИТЕРІЮ ПРИ СТАТИСТИЧНОМУ АНАЛІЗУ ПЕДАГОГІЧНОГО ЕКСПЕРИМЕНТУ

Кулибаба Є.С.
Керівник: д. т. н, професор Єремєєв В.С.
Мелітопольський державний педагогічний університет імені Богдана Хмельницкого
e-mail: eremeev@mdpu.org.ua

Постановка проблеми. Перед сучасною педагогічною наукою постає необхідність розв'язати конфлікт між невпинним зростанням обсягу знань та обмеженістю часу. Для перевірки будь якої педагогічної ідеї потрібним є проведення педагогічних експериментів. При цьому для кількісної оцінки ефективності будь-якого педагогічного експерименту застосовують статистичні методи. Зрозуміло, що такі оцінки можуть бути зроблені лише з певною мірою ймовірності, так як на експеримент впливає багато факторів, які не можна врахувати. Але чим більшими є обсяги контрольної та експериментальної групи, тим більше проявляється вплив основного фактору та нівелюються фактори випадкові, не пов'язані з педагогічним експериментом. З іншого боку, проведення експериментів для надто великих груп потребує більше часу, що затримує прийняття рішення щодо необхідності корегування педагогічної методики. Тому представляє інтерес розробка методів, які дозволять встановити оптимальний обсяг груп, при якому з достатньою достовірністю можна виявити заданий щонайменший вплив педагогічного експерименту на фактор успішності учнів.

Аналіз останніх досліджень. Сьогодні відомо багато статистичних критеріїв [1], [2] що дозволяють перевіряти гіпотези про рівність середніх або дисперсій двох вибірок. Але задача порівняння середніх двох нормально розподілених вибірок при невідомих та нерівних дисперсіях (проблема Беренса – Фішера) поки що немає точного розвязку. Однак на практиці використовують різні наближення. Одними з найвідоміших є критерії Кохрена – Кокса, Сатервайта та Крамера – Уелча [3]. Іншим питанням, яке не має однозначної відповіді у загальному випадку, є питання про оптимальний обсяг вибірки.

Мета статті - сформулювати математичний критерій, що характеризує чутливість статистичного критерію до різниці між математичними очікуваннями у контрольній та експериментальній виборках.

Основна частина. Нехай в ході педагогічного експерименту досліджується ефективність засвоєння учнями деяких професійних навиків та знань, яка характеризується деякою сукупною оцінкою (однофакторний експеримент). Ця оцінка визначається для кожного учня з контрольної групи та з експериментальної групи. Експеримент полягає у порівнянні середньої оцінки в обох групах.

В залежності від випадку у математичній статистиці використовують різні статистичні критерії того, що дані групи відрізняються істотно.

Нехай оцінки у контрольній та експериментальній групах із обсягами n_1 та n_2 розподілені нормально із математичними очікуваннями M_1 та M_2 й з відомими середньоквадратичними відхиленнями σ_1 та σ_2. У якості нульової гіпотези потрібно взяти гіпотезу про те, що $M_1 = M_2$ (тобто те, що $M_1 \neq M_2$ зумовлено дією випадкових факторів). Так як в результаті педагогічного експерименту бажаним є відхилення оцінок у експериментальній групі у більшу сторону у порівнянні із контрольною групою, альтернативною гіпотезою слід вважати $M_1 < M_2$. Для перевірки гіпотези із заданим рівнем значущості α використовують z-критерій [2]:

$$Z = \frac{M_2 - M_1}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < z_k \qquad (1)$$

де у якості квантілю цього критерію z_k використовують $\Phi^{-1}\left(\frac{1-2\alpha}{2}\right)$, де

Ф – функція, зворотна до інтегралу похибок

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz \qquad (2)$$

Якщо середньоквадратичні відхилення σ_1 та σ_2 невідомі, але рівні між собою, при тих само нульовій та альтернативній гіпотезах застосовують критерій Стьюдента

$$T = \frac{M_2 - M_1}{\sqrt{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} < t_k \qquad (3)$$

де t_k - правосторонній квантіль розподілу Стьюдента із $n_1 + n_2 - 2$ ступенями свободи для рівня значущості α, в якості σ_1^2 та σ_2^2 взяті виправлені дисперсії

$$\sigma_1^2 = \frac{1}{n_1 - 1} \sum_i (x_i - M_1)^2, \quad \sigma_2^2 = \frac{1}{n_2 - 1} \sum_i (y_i - M_2)^2 \qquad (4)$$

Критерії Кохрена – Кокса, Сатервайта та Крамера – Уелча критерії мають таку ж статистику, як і в критерію Стьюдента, однак відрізняються від цього критерію квантілем.

Квантіль критерію Кохрена-Кокса обчислюється так:

$$t_k = \frac{v_1 t(n_1 - 1, \alpha) + v_2 t(n_2 - 1, \alpha)}{v_1 + v_2} \tag{5}$$

де

$$v_1 = \frac{\sigma_1^{\;2}}{n_1}, \quad v_2 = \frac{\sigma_2^{\;2}}{n_2} \tag{6}$$

t(n, α) – квантіль розподілу Стьюдента із n – ступенями вільності, $\sigma_1^{\;2}$ та $\sigma_2^{\;2}$ обчислені за формулами (4).

Квантіль критерію Сатервайта обчислюється як t(f, α), де

$$f = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^2 \left(\frac{1}{n_1 - 1}\left(\frac{\sigma_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{\sigma_2^2}{n_2}\right)^2\right)^{-1} \tag{7}$$

Квантіль критерію Крамера-Уелча обчислюється як t(f, α), де

$$f = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^2 \left(\frac{1}{n_1 + 1}\left(\frac{\sigma_1^2}{n_1}\right)^2 + \frac{1}{n_2 + 1}\left(\frac{\sigma_2^2}{n_2}\right)^2\right)^{-1} - 2 \tag{8}$$

Нехай оцінки у контрольній та експериментальній групах рівного обсягу n розподілені нормально із середньоквадратичним відхиленням σ, математичне очікування оцінок у контрольній групі M_1 у експериментальній M_2. Тоді статистика z-критерію прийме вигляд

$$Z = \frac{M_2 - M_1}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{n}}} = \frac{M_2 - M_1}{\sigma}\sqrt{\frac{n}{2}} \tag{9}$$

Припустивши, що вибіркові дисперсії є досить близькими до генеральної, статистику критерію Стьюдента (а також критеріїв Кохрена-Кокса, Сатервайта та Крамера-Уелча) можна записати у вигляді

$$T = \frac{M_2 - M_1}{\sigma}\sqrt{\frac{n}{2}} \tag{10}$$

У z-критерії значущим є відношення

$$\tilde{N} = \frac{M_2 - M_1}{\sigma} \tag{11}$$

яке матиме простий зміст: на скільки відрізняється середнє стандартизованого нормального розподілу різниці між оцінками двох груп від нуля.

Статистика (10) залежить від (11). Це відношення можна покласти в основу критерію чутливості статистичного критерію до відносного відхилення математичних очікувань. Дійсно, нехай ми достовірно знаємо значення величин M_1, M_2, σ. Тоді за критерій чутливості можна вважати значення величини (11), при якому ймовірність прийняття альтернативної гіпотези статистичним критерієм буде 1 - α. Критерій можна записати

$$Crit = \left. \frac{M_2 - M_1}{\sigma} \right|_{\lim\limits_{K \to \infty} \frac{k}{K} = 1 - \alpha} \qquad (12)$$

де K – кількість чисельних експериментів (досить велике число), k – кількість експериментів, при якій статистичний критерій вказав на справедливість альтернативної гіпотези.

Чисельні розрахунки показують, що для розглянутих статистичних критеріїв функція (12) є монотонно спадною відносно n, тому у якості оптимального обсягу груп можна вважати мінімальне n таке, що функція (12) менша за відношення очікуваного педагогічного ефекту до апріорного значення середньоквадратичного відхилення, якщо така апріорна інформація є.

Висновки. У статті побудовано функцію, що характеризує чутливість групи статистичних критеріїв до відмінності контрольної та експериментальної вибірок. За наявності апріорної інформації про дисперсію вибірок функція (12) може бути використано для визначення оптимального обсягу груп.

Література

1. Єремєєв В.С. Теорія ймовірностей та математична статистика. Навчальний посібник./ В.С. Єремєєв, Д.О. Сосновських, О.В. Тітова - Мелитополь: ТОВ «Видавничий будинок», 2009. – 187 с.
2. Еремеев В.С. Статистическая обработка педагогического эксперимента в случае неизвестной функции распределения / В.С. Еремеев, В.В. Кузьминов// Сборник трудов Северокавказского филиала Белгородского государственного технологического университета имени В. Г. Шухова. 2013г., вып 13, с.44-51.
3. Критерий_Стьюдента. [Електронний ресурс]. Режим доступу: http://www.machinelearning.ru/wiki/index.php?title =

Анотація: У статті розглянуто проблему визначення чутливості статистичного критерію та її залежності від обсягу вибірки, запропонована функція чутливості.

Ключові слова: квантіль; критерії Стьюдента; педагогічний експеримент; статистика; статистичний критерій.

Аннотация: В статье рассмотрена проблема определения чувствительности статистического критерия и ее зависимости от объема выборки, предложенна функция чувствительности.

Ключевые слова: квантиль; критерий Стьюдента; педагогический эксперимент; статистика; статистический критерий.

***Abstract*:** The article considers the problem of determining the sensitivity of the statistical criterion and its dependence on the sample size, the proposed function of sensitivity.

Keywords: quantile; student test; pedagogical experiment; Student’s criterion; statistics; statistical criterion.