

Scientific journal  
**PHYSICAL AND MATHEMATICAL EDUCATION**  
Has been issued since 2013.

ISSN 2413-158X (online)  
ISSN 2413-1571 (print)

Науковий журнал  
**ФІЗИКО-МАТЕМАТИЧНА ОСВІТА**  
Видається з 2013.



<http://fmo-journal.fizmatsspu.sumy.ua/>

*Пудова С., Соласа Г., Метваллі Ш. Індивідуальна самостійна робота студентів-медиків з медичної інформатики по вивченню критерію  $\chi^2$ -квадрат. Фізико-математична освіта. 2019. Випуск 3(21). С. 113-120.*

*Pudova S., Solasa G., Metwally S. Medical students' individual independent work on medical informatics for studying the chi-square test. Physical and Mathematical Education. 2019. Issue 3(21). P. 113-120.*

DOI 10.31110/2413-1571-2019-021-3-017  
УДК 61:004:616-053.82:378.147.88

**С.С. Пудова**

Вінницький національний медичний університет ім. М.І. Пирогова, Україна  
svetlana\_pudova@hotmail.com  
ORCID: 0000-0001-5231-400X

**Г.В. Соласа**

Вінницький національний медичний університет ім. М.І. Пирогова, Україна  
giridhar2136@gmail.com  
ORCID: 0000-0003-2278-9092

**Ш. Метваллі**

Вінницький національний медичний університет ім. М.І. Пирогова, Україна  
abdelhai25@icloud.com  
ORCID: 0000-0003-4807-4730

#### ІНДИВІДУАЛЬНА САМОСТІЙНА РОБОТА СТУДЕНТІВ-МЕДИКІВ З МЕДИЧНОЇ ІНФОРМАТИКИ ПО ВИВЧЕННЮ КРИТЕРІЮ $\chi^2$ -КВАДРАТ

##### АНОТАЦІЯ

**Формулювання проблеми.** Відповідно до навчальних програм майбутні лікарі починають вивчати елементи статистики з перших курсів навчання. Зокрема, методи біостатистики вони вивчають на заняттях з медичної інформатики на другому курсі навчання. Водночас більшість вступників у вищі медичні заклади України мають низьку фізико-математичну підготовку. Чи здатні сучасні студенти-медики виконувати складні математичні завдання і чи зацікавлені вони у таких завданнях?

**Матеріали і методи.** У ході педагогічного дослідження було здійснено аналіз літератури з педагогіки та статистики. Оброблену теоретичну інформацію зі статистики стосовно критерію  $\chi^2$ -квадрат використано для створення проблемного завдання з індивідуальної самостійної роботи студентів з медичної інформатики. Додатково використано програму MS Excel для створення демонстраційних матеріалів щодо розрахунків емпіричного значення коефіцієнта  $\chi^2$ -квадрат. На заключному етапі здійснено перевірку виконаних студентами завдань та аналіз педагогічної ситуації щодо проведеної роботи.

**Результати.** Англійські студенти-медики другого курсу лікувального факультету виконали поставлене перед ними завдання значно швидше, ніж ми припускали. Під час виконання продемонстровано досить високий рівень математичних знань, умінь та навичок, а також зацікавленість у виконанні завдання. Студенти представили два різних доведення рівності двох формул, записаних у завданні, стосовно критерію  $\chi^2$ -квадрат. Використовуючи способи доведення студентів, можна вивести формулу для розрахунку емпіричного значення коефіцієнта  $\chi^2$ -квадрат для трьох вибірок, яку також представлено у цій статті.

**Висновки.** Проведена самостійна робота студентів-медиків, з одного боку, продемонструвала готовність майбутніх лікарів до вирішення нестандартних для їхнього профілю завдань, які вимагають поглибленого рівня знань із математики. З іншого боку, зміст завдань та представлені студентами доведення дають можливість глибше розібратися з вибраним статистичним методом та використовувати отримані знання на практиці.

**КЛЮЧОВІ СЛОВА:** індивідуальна самостійна робота студентів, медична інформатика, майбутні лікарі, критерій  $\chi^2$ -квадрат (критерій Пірсона), формула статистики  $\chi^2$ -квадрат.

##### ВСТУП

**Постановка проблеми.** Здатність до абстрактного мислення, аналізу та синтезу, здатність вчитися і оволодівати сучасними знаннями, навички використання інформаційних і комунікаційних технологій, здатність до проведення епідеміологічних та медико-статистичних досліджень здоров'я населення та низку інших складових професійної культури лікаря сучасні студенти-медики мають розвинути під час навчання у вищій медичній школі, щоб бути конкурентоспроможними на ринку праці (Галузевий стандарт, 2016). Аналіз літературних джерел, робота над

проектами, невеликі за часом та обсягом дослідження зі статистичною обробкою даних тощо – це завдання, звичні для більшості студентів. При цьому рівень навчальної компетентності останніх в середньому знижується з кожним роком (Беззуб; Іщейкіна&Сілкова, 2011; МОЗ України). Цікавим для майбутніх досліджень є питання про те, чи є сучасний рівень знань і навичок студентів фактором, який зупиняє викладача створювати завдання високої складності для студентів. Іншими словами, чи готові студенти виконувати складні завдання і чи є такі завдання невиконуваними для них. Розвиток лікаря, котрий мислить творчо і має високий рівень професійної культури, потребує саме нестандартних завдань. Відтак про здатність сучасних студентів виконати певне завдання можна говорити після отримання відповідного педагогічного досвіду.

**Актуальність дослідження.** Питання професійної підготовки майбутніх лікарів досліджували та розкривали у своїх працях І. Є. Булах, Ю. В. Вороненко, І. А. Голованова, С. С. Касинець, К. Г. Магрламова, М. Р. Мруга, Я. В. Цехмістер, К. П. Хоменко, О. Г. Шекера тощо. Зокрема, шляхи організації самостійної роботи студентів при вивченні дисциплін теоретичного та клінічного профілю описали А. М. Гоголь, Т. М. Гонтова, І. А. Колісник, К. П. Локес, Н. В. Пасечко, Л. В. Радецька, І. В. Шлімкевич тощо. Розробкою навчальних завдань із медичної інформатики для самостійної (поза)аудиторної роботи студентів-медиків з метою формування певних професійних компетентностей займалися О. В. Бойко, А. М. Добровольська, В. І. Доценко, В. П. Марценюк, С. К. Мисловська, Г. Ю. Мороховець, С. Ю. Оленець, С. В. Різничок та інші.

Проблеми організації начального процесу з вивчення питань статистики та шляхів формування компетентності майбутніх лікарів у галузі статистичної науки розкривали В. А. Огнев, Т. М. Литвінова, О. І. Панченко, О. Є. Прокопченко, О. В. Сілкова, Н. В. Стучинська тощо. Безпосередньо в публікаціях можна знайти інформацію щодо викладання певних тем зі статистики на перших курсах під час вивчення медичної і біологічної фізики та медичної інформатики відповідно до навчальних програм. При цьому ми не зустрічали наукових робіт, у яких описуються завдання для студентів-медиків, спрямовані на глибше вивчення статистики в позааудиторний час. У той же час певні методичні розробки або згадування про них можна знайти на сайтах кафедр медичної інформатики та фізики. Наприклад, на сайті кафедри медичної інформатики Тернопільського державного медичного університету ім. І. Я. Горбачевського в розділі «Інновації, здобутки» інформується про розроблену методичну вказівку на тему «Дисперсійний аналіз у фармацевтичних дослідженнях» з дисципліни «Вища математика» та методичну вказівку «Побудова кореляційних портретів...» з дисципліни «Медична інформатика» з розділу «Методи біостатистики. Статистичний аналіз даних» (<https://medinf.tdmu.edu.ua/education/innovacie-zdobutki>). Раніше на сайті цієї ж кафедри можна було знайти інформацію про можливість набути уміння та навички з обробки статистичних даних в середовищі R для всіх охочих студентів.

Варто звернути увагу на те, що однією з причин браку наукової інформації в педагогічній галузі з вивчення дисциплін у медичній вищій школі є напрямки наукових досліджень викладачів. Зрозуміло, що більшість з них працює у галузі медицини. Якщо взяти кафедри медичної фізики та інформатики, які часто об'єднані в одну в медичних університетах, то значна кількість викладачів має науковий ступінь в галузі фізико-математичних, біологічних, технічних наук, і, відповідно, результати їхніх досліджень оминають напрямок педагогіки. До того ж далеко не завжди на сайтах кафедр можна побачити інформацію про навчально-методичні посібники, а особливо – їх зміст. Внаслідок цього можемо говорити про недостатню дослідженість питання навчально-виховного процесу у вищих освітніх медичних закладах, в тому числі – при вивченні медичної інформатики.

Повертаючись до індивідуальної самостійної роботи студентів, відзначимо, що в попередніх наших публікаціях (Пудова&Воєвода, 2018) ми згадували, зокрема, про можливі шляхи формування компетентності майбутніх лікарів із питань біостатистики на заняттях з медичної інформатики та типи завдань, які виконують наші студенти. Серед таких завдань зазначався пошук інформації про певний статистичний метод дослідження, наведення прикладів використання вибраного студентами методу в медичній практиці; представлення одного зі способів обробки даних в одній із комп'ютерних програм (Пудова&Воєвода, 2018). Продовжуючи працювати у цьому напрямку, у 2018-2019 навчальному році ми вирішили урізноманітнити завдання для студентів та трохи заглибитися в один зі статистичних методів – критерій Пірсона (критерій  $\chi^2$ -квадрат). Метод порівняння розподілів об'єктів двох і більше вибірок використовують не лише в медичній, а й у будь-якій іншій галузі. Наприклад, в психолого-педагогічних дослідженнях за допомогою цього методу доводиться порівнювати результати, отримані в контрольній та експериментальній групі. Таким чином, завдання з індивідуальної самостійної роботи студентів-медиків, про яке йтиметься в цій статті, а особливо результати виконання цього завдання можуть бути цікавими також для нинішніх і майбутніх педагогів.

**Мета статті** – навести приклад завдання для індивідуальної самостійної роботи студентів-медиків з питань біостатистики при вивченні медичної інформатики, а також представити розв'язання цього завдання, надане студентами.

## ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

У науковій літературі можна зустріти декілька формул, які використовуються для обчислення емпіричного значення коефіцієнта  $\chi^2$ -квадрат (статистики критерію  $\chi^2$ -квадрат). Зокрема, дві з них мають наступний вигляд:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (1),$$

де  $r$  – кількість вибірок або кількість рядочків у таблиці спряженості,  $c$  – кількість категорій або кількість стовпчиків у таблиці спряженості,  $O_{ij}$  – емпірична частота або число, яке спостерігається (отримане в дослідженні), у клітинці таблиці спряженості,  $E_{ij}$  – теоретична частота або очікуване число в тій же клітинці таблиці (Гланц, 1999, С. 142; Грабарь&Краснянская, 1977, С. 97; Суходольський, 1998, С. 284; Стариченко, 2004, С. 59; Толкачев, 2017, С. 14; Franke, Ho&Christie, 2012, P. 449; Harris&Fu, 2018, P. 8; Hugh, 2013, P. 145; Onchiri, 2013, P. 1234; Pandis, 2016, P. 898; Rana&Singhal, 2019, P. 69; Ugoni&Walker, 1995 та ін.);

$$\chi^2 = \frac{1}{n_1 \cdot n_2} \times \sum_{j=1}^c \frac{(n_1 \cdot O_{2j} - n_2 \cdot O_{1j})^2}{O_{1j} + O_{2j}} \quad (2),$$

де  $n_1$  – розмір першої вибірки,  $n_2$  – розмір другої вибірки (Грабарь&Краснянская, 1977, С. 101; Набиулина, 2015, С. 190; Стариченко, 2004, С. 60; Новиков, 2004, С. 52; та ін.).

Формула (1) є загальною для критерію Пірсона, тобто може використовуватися при кількості вибірок більше двох, тоді як формула (2) передбачає роботу з даними по двох вибірках. В англомовних дослідженнях зазвичай використовують формулу (1) (Гланц, 1999, С. 142; Franke, Ho&Christie, 2012, Р 449; Harris&Fu, 2018, Р. 8; Hugh, 2013, Р. 145; Onchiri, 2013, Р. 1234; Pandis, 2016, Р. 898; Rana&Singhal, 2019, Р. 69; Ugoni&Walker, 1995; та ін.), тоді як в україномовних та російськомовних статтях, зокрема, з педагогіки, можна знайти формулу (2), яку записують також у підручники та навчальні посібники (Грабарь&Краснянская, 1977, С. 101; Набиулина, 2015, С. 190; Стариченко, 2004, С. 60; Новиков, 2004, С. 52; та ін.).

Крім того, можна зустріти інформацію, що друга формула впливає з першої, тобто вони є взаємозамінними для випадку з двома вибірками (Грабарь&Краснянская, 1977, С. 101). При цьому в підручниках та статтях ми не знайшли доведення цього факту, тобто він не знаходиться на поверхні. Однак такий факт є цікавим для дослідників, оскільки, володіючи цією інформацією та знаючи, як працювати з формулами, можна без жодних сумнівів використовувати будь-яку з них під час обробки статистичних даних.

**МЕТОДИ ДОСЛІДЖЕННЯ**

Дослідження умовно можна розділити на два етапи. Перший етап пов’язаний з роботою викладача, коли здійснено аналіз педагогічної літератури з визначеної теми та аналіз наукових робіт в напрямку статистики стосовно критерію хі-квадрат. Оброблену теоретичну інформацію зі статистики використано для створення завдань з індивідуальної самостійної роботи для майбутніх лікарів з медичної інформатики. Додатково використано програму MS Excel для створення демонстраційних матеріалів для студентів щодо здійснення розрахунків емпіричного значення статистики хі-квадрат за двома вибраними формулами. Другий етап включає роботу студентів з виконання розробленого завдання, перевірку його викладачем, а також аналіз педагогічної ситуації щодо проведеної роботи.

Завдання з індивідуальної самостійної роботи студентів потребувало доведення рівності двох формул, згаданих вище. Перед виконанням завдання студенти були ознайомлені з: 1) суттю статистичного методу порівняння результатів двох незалежних вибірок за критерієм хі-квадрат; 2) розрахунковими формулами, включаючи їх запис, пояснення, приклад розрахунку за кожною з них; 3) записом статистичної інформації у вигляді таблиці для цих формул. Крім того, на декількох прикладах було продемонстровано ідентичність (до будь-якого знаку після коми) отриманих емпіричних значень коефіцієнта хі-квадрат за двома запропонованими формулами в програмі MS Excel.

**РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ**

У 2018-2019 навчальному році ми працювали в англомовних групах, тому представимо завдання для студентів та його розв’язання англійською мовою.

**Problem (task) for students.** Prove the equality of the two formulas (1) and (2):

$$1) \chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} - \text{general formula of chi-square test,}$$

where  $r$  – number of rows,  $c$  – number of columns,  $O_{i,j}$  – observed value of two nominal variables,  $E_{i,j}$  – expected value of two nominal variables,

$$E_{i,j} = \frac{\text{row total} \times \text{column total}}{\text{sample size}} \tag{3};$$

$$2) \chi^2 = \frac{1}{n_1 \cdot n_2} \times \sum_{j=1}^c \frac{(n_1 \cdot O_{2j} - n_2 \cdot O_{1j})^2}{O_{1j} + O_{2j}} - \text{formula of chi-square test when } r = 2, \text{ where } n_1 - \text{size of the first sample, } n_2 - \text{size of the second sample.}$$

Examples of using formulas:

The first table shows the empirical frequencies (observed values) – the data are based on the conducted experiment (e.g., survey).

Table 1

**Empirical frequencies**

Sample/category	Category 1	Category 2	Category 3	...	Category k	Totally (in sum)
The first sample	$O_{1,1}$	$O_{1,2}$	$O_{1,3}$	...	$O_{1,k}$	$n_1 = O_{1,1} + O_{1,2} + \dots + O_{1,k}$
The second sample	$O_{2,1}$	$O_{2,2}$	$O_{2,3}$	...	$O_{2,k}$	$n_2 = O_{2,1} + O_{2,2} + \dots + O_{2,k}$
Totally (in sum)	$m_1 = O_{1,1} + O_{2,1}$	$m_2 = O_{1,2} + O_{2,2}$	$m_3 = O_{1,3} + O_{2,3}$	...	$m_k = O_{1,k} + O_{2,k}$	$n = n_1 + n_2 = m_1 + m_2 + \dots + m_k$

We can use the data from the Table 1 to calculate  $\chi^2$  by the second formula. For example, we write the observed values in the Table 2 and calculate the value  $\chi^2$ .

Table 2

**Empirical frequencies (observed values) from the experiment**

Sample	Point						Totally
	1	2	3	...	k-1	k	
Sample 1	20	42	12	...	74	95	456
Sample 2	25	40	17	...	88	96	502
Totally	45	82	29	...	162	191	958

$$\chi^2 = \frac{1}{n_1 \cdot n_2} \sum_{j=1}^c \frac{(n_1 \cdot O_{2,j} - n_2 \cdot O_{1,j})^2}{O_{1,j} + O_{2,j}}$$

$$= \frac{1}{456 \times 502} \times \left( \frac{(456 \times 25 - 502 \times 20)^2}{20 + 25} + \frac{(456 \times 40 - 502 \times 42)^2}{42 + 40} + \frac{(456 \times 17 - 502 \times 12)^2}{12 + 17} + \dots + \frac{(456 \times 88 - 502 \times 74)^2}{74 + 88} + \frac{(456 \times 96 - 502 \times 95)^2}{95 + 96} \right) = 10.22$$

Create a Table 3 of theoretical frequencies (expected values) to calculate  $\chi^2$  by the first formula.

Table 3

Theoretical frequencies

Sample/category	Category 1	Category 2	Category 3	...	Category k	Totally (in sum)
The first sample	$E_{1,1} = \frac{n_1 \times m_1}{n}$	$E_{1,2} = \frac{n_1 \times m_2}{n}$	$E_{1,3} = \frac{n_1 \times m_3}{n}$	...	$E_{1,k} = \frac{n_1 \times m_k}{n}$	$n_1$
The second sample	$E_{2,1} = \frac{n_2 \times m_1}{n}$	$E_{2,2} = \frac{n_2 \times m_2}{n}$	$E_{2,3} = \frac{n_2 \times m_3}{n}$	...	$E_{2,k} = \frac{n_2 \times m_k}{n}$	$n_2$
Totally (in sum)	$m_1 = E_{1,1} + E_{2,1}$	$m_2 = E_{1,2} + E_{2,2}$	$m_3 = E_{1,3} + E_{2,3}$	...	$m_k = E_{1,k} + E_{2,k}$	$n$

Create a Table 4 of the theoretical frequencies according to our example.

Table 4

Calculated theoretical frequencies

Sample	Point						Totally
	1	2	3	...	k-1	k	
Sample 1	$456 * 45$	$456 * 82$	$456 * 29$	...	$456 * 162$	$456 * 191$	456
	$\frac{958}{= 21.42}$	$\frac{958}{= 39.03}$	$\frac{958}{= 13.80}$		$\frac{958}{= 77.11}$	$\frac{958}{= 90.91}$	
Sample 2	$502 * 45$	$502 * 82$	$502 * 29$	...	$502 * 162$	$502 * 191$	502
	$\frac{958}{= 23.58}$	$\frac{958}{= 42.97}$	$\frac{958}{= 15.20}$		$\frac{958}{= 84.89}$	$\frac{958}{= 100.09}$	
Totally	45	82	29	...	162	191	958

We use the data in Table 2 and the data in Table 4 to calculate  $\chi^2$  by the first formula (1). For example:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}$$

$$= \frac{(20 - 21.42)^2}{21.42} + \frac{(42 - 39.03)^2}{39.03} + \frac{(12 - 13.80)^2}{13.80} + \dots + \frac{(74 - 77.11)^2}{77.11} + \frac{(95 - 90.91)^2}{90.91} + \frac{(25 - 23.58)^2}{23.58} + \frac{(40 - 42.97)^2}{42.97} + \dots + \frac{(88 - 84.89)^2}{84.89} + \frac{(96 - 100.09)^2}{100.09} = 10.22$$

**Proof from Shaimaa Metwally, Egypt.**

From formula (3) and Table 1, expected value is  $E_{1,j} = \frac{n_1 \cdot (O_{1,j} + O_{2,j})}{n_1 + n_2}$  when  $i = 1$ .

Find the difference  $(O_{1,j} - E_{1,j})^2$ .

$$(O_{1,j} - E_{1,j})^2 = \left( O_{1,j} - \frac{n_1 \cdot (O_{1,j} + O_{2,j})}{n_1 + n_2} \right)^2 = \left( \frac{(n_1 + n_2) \cdot O_{1,j} - n_1 \cdot (O_{1,j} + O_{2,j})}{n_1 + n_2} \right)^2 = \left( \frac{n_1 O_{1,j} + n_2 O_{1,j} - n_1 O_{1,j} - n_1 O_{2,j}}{n_1 + n_2} \right)^2 = \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{(n_1 + n_2)^2}$$

Find the ratio  $\frac{(O_{1,j} - E_{1,j})^2}{E_{1,j}}$ .

$$\frac{(O_{1,j} - E_{1,j})^2}{E_{1,j}} = \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{(n_1 + n_2)^2} \div \frac{n_1 \cdot (O_{1,j} + O_{2,j})}{n_1 + n_2} = \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{(n_1 + n_2)^2} \times \frac{n_1 + n_2}{n_1 \cdot (O_{1,j} + O_{2,j})} = \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{n_1 \cdot (n_1 + n_2) \cdot (O_{1,j} + O_{2,j})} \tag{4}$$

Similar to the previous one we have:

$E_{2,j} = \frac{n_2 \cdot (O_{1,j} + O_{2,j})}{n_1 + n_2}$  is expected value when  $i = 2$ .

Find the difference  $(O_{2,j} - E_{2,j})^2$ .

$$(O_{2,j} - E_{2,j})^2 = \left( O_{2,j} - \frac{n_2 \cdot (O_{1,j} + O_{2,j})}{n_1 + n_2} \right)^2 = \left( \frac{(n_1 + n_2) \cdot O_{2,j} - n_2 \cdot (O_{1,j} + O_{2,j})}{n_1 + n_2} \right)^2 = \left( \frac{n_1 O_{2,j} + n_2 O_{2,j} - n_2 O_{1,j} - n_2 O_{2,j}}{n_1 + n_2} \right)^2 = \frac{(n_1 O_{2,j} - n_2 O_{1,j})^2}{(n_1 + n_2)^2}$$

Find the ratio  $\frac{(O_{2,j} - E_{2,j})^2}{E_{2,j}}$ .

$$\frac{(O_{2,j} - E_{2,j})^2}{E_{2,j}} = \frac{(n_1 O_{2,j} - n_2 O_{1,j})^2}{(n_1 + n_2)^2} \div \frac{n_2 \cdot (O_{1,j} + O_{2,j})}{n_1 + n_2} = \frac{(n_1 O_{2,j} - n_2 O_{1,j})^2}{(n_1 + n_2)^2} \times \frac{n_1 + n_2}{n_2 \cdot (O_{1,j} + O_{2,j})} = \frac{(n_1 O_{2,j} - n_2 O_{1,j})^2}{n_2 \cdot (n_1 + n_2) \cdot (O_{1,j} + O_{2,j})} \tag{5}$$

Taking into account formulas (4) and (5), we write formula (1) for the case when  $r = 2$ .

$$\begin{aligned} \chi^2 &= \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\ &= \sum_{j=1}^c \left( \frac{(O_{1,j} - E_{1,j})^2}{E_{1,j}} + \frac{(O_{2,j} - E_{2,j})^2}{E_{2,j}} \right) \\ &= \sum_{j=1}^c \left( \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{n_1 \cdot (n_1 + n_2) \cdot (O_{1,j} + O_{2,j})} + \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{n_2 \cdot (n_1 + n_2) \cdot (O_{1,j} + O_{2,j})} \right) \\ &= \sum_{j=1}^c \frac{n_2 \cdot (n_2 O_{1,j} - n_1 O_{2,j})^2 + n_1 \cdot (n_2 O_{1,j} - n_1 O_{2,j})^2}{n_1 \cdot n_2 \cdot (n_1 + n_2) \cdot (O_{1,j} + O_{2,j})} \\ &= \sum_{j=1}^c \frac{(n_1 + n_2) \cdot (n_2 O_{1,j} - n_1 O_{2,j})^2}{n_1 \cdot n_2 \cdot (n_1 + n_2) \cdot (O_{1,j} + O_{2,j})} = \sum_{j=1}^c \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{n_1 n_2 (O_{1,j} + O_{2,j})} = \frac{1}{n_1 \cdot n_2} \times \sum_{j=1}^c \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{(O_{1,j} + O_{2,j})} \end{aligned}$$

Thus formula (1) is equal to formula (2).

Hence proved.

**Proof from Giridhar Solasa, India.**

We know that  $n_1 = O_{1,1} + O_{1,2} + \dots + O_{1,k}$ ;  $n_2 = O_{2,1} + O_{2,2} + \dots + O_{2,k}$ ;  $m_1 = O_{1,1} + O_{2,1}$ ;  $m_2 = O_{1,2} + O_{2,2}$ ;  $E_{i,j} = \frac{n_i \times m_j}{n}$ .

Taking  $r = c = \{1,2\}$  we have

$$\begin{aligned} \chi^2 &= \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(O_{1,1} - E_{1,1})^2}{E_{1,1}} + \frac{(O_{2,1} - E_{2,1})^2}{E_{2,1}} + \frac{(O_{1,2} - E_{1,2})^2}{E_{1,2}} + \frac{(O_{2,2} - E_{2,2})^2}{E_{2,2}} = \frac{(O_{1,1} - \frac{n_1 \times m_1}{n})^2}{\frac{n_1 \times m_1}{n}} + \frac{(O_{2,1} - \frac{n_2 \times m_1}{n})^2}{\frac{n_2 \times m_1}{n}} + \\ &\frac{(O_{1,2} - \frac{n_1 \times m_2}{n})^2}{\frac{n_1 \times m_2}{n}} + \frac{(O_{2,2} - \frac{n_2 \times m_2}{n})^2}{\frac{n_2 \times m_2}{n}} = \frac{(n \cdot O_{1,1} - n_1 m_1)^2}{n \cdot n_1 m_1} + \frac{(n \cdot O_{2,1} - n_2 m_1)^2}{n \cdot n_2 m_1} + \frac{(n \cdot O_{1,2} - n_1 m_2)^2}{n \cdot n_1 m_2} + \frac{(n \cdot O_{2,2} - n_2 m_2)^2}{n \cdot n_2 m_2} = \\ &\frac{n_2((n_1 + n_2)O_{1,1} - n_1(O_{1,1} + O_{2,1}))^2 + n_1((n_1 + n_2)O_{2,1} - n_2(O_{1,1} + O_{2,1}))^2}{n \cdot n_1 n_2 m_1} + \frac{n_2((n_1 + n_2)O_{1,2} - n_1(O_{1,2} + O_{2,2}))^2 + n_1((n_1 + n_2)O_{2,2} - n_2(O_{1,2} + O_{2,2}))^2}{n \cdot n_1 n_2 m_2} = \\ &\frac{n_2(n_2 O_{1,1} - n_1 O_{2,1})^2 + n_1(n_1 O_{2,1} - n_2 O_{1,1})^2}{n \cdot n_1 n_2 m_1} + \frac{n_2(n_2 O_{1,2} - n_1 O_{2,2})^2 + n_1(n_1 O_{2,2} - n_2 O_{1,2})^2}{n \cdot n_1 n_2 m_2} = \frac{(n_1 + n_2)(n_2 O_{1,1} - n_1 O_{2,1})^2}{(n_1 + n_2)n_1 n_2 (O_{1,1} + O_{2,1})} + \frac{(n_1 + n_2)(n_2 O_{1,2} - n_1 O_{2,2})^2}{(n_1 + n_2)n_1 n_2 (O_{1,2} + O_{2,2})} = \\ &\frac{1}{n_1 n_2} \times \left( \frac{(n_2 O_{1,1} - n_1 O_{2,1})^2}{(O_{1,1} + O_{2,1})} + \frac{(n_2 O_{1,2} - n_1 O_{2,2})^2}{(O_{1,2} + O_{2,2})} \right) = \frac{1}{n_1 n_2} \sum_{j=1}^2 \frac{(n_2 O_{1,j} - n_1 O_{2,j})^2}{(O_{1,j} + O_{2,j})} \end{aligned}$$

If the above said equation is said to be true for two rows and two columns (i.e.,  $O_{1,1}, O_{2,1}, O_{1,2}, O_{2,2}$ ), it should be equal for infinite series.

$$\chi^2 = \frac{1}{n_1 \cdot n_2} \times \sum_{j=1}^c \frac{(n_1 \cdot O_{2,j} - n_2 \cdot O_{1,j})^2}{O_{1,j} + O_{2,j}}$$

Hence proved.

**ОБГОВОРЕННЯ**

Ми підтверджуємо факт, згаданий у вступі статті, щодо загального низького рівня знань, умінь та навичок українських студентів, зокрема, на перших курсах у напрямку фізико-математичної підготовки. Водночас констатуємо: щороку на лікувальний факультет вступають окремі абітурієнти, досить добре підготовлені з різних дисциплін. Крім того, знаємо, що такі студенти ще до вступу в університет здійснювали певні дослідження з використанням статистичної обробки даних. Деякі іноземні студенти до вступу в університет пройшли навчальні курси зі статистики. Таким чином, маємо певний відсоток студентів-медиків, яким може бути цікаво заглибитися в питання статистики та продемонструвати свої здібності. Інші ж можуть ознайомитися з результатами виконаного завдання та використовувати інформацію щодо критерію хі-квадрат.

Завдання було оголошено для студентів усіх англomовних груп лікувального факультету (майже 250 студентів) в кінці жовтня 2018 року. Пропонували виконати його до початку лютого 2019 року (до початку наступного семестру), але не обмежували цим терміном, оскільки ми припускали, що виконання може зайняти тривалий час. Нагадаємо, що в англomовних джерелах поширеним є використання загальної формули для обчислення емпіричного значення хі-квадрат, і знайти готову відповідь на наше завдання – не так легко. Пошук відповіді у книжках та статтях для українськомовних студентів теж не є простим і потребує опрацювання значної кількості друкованих праць.

Результатом виконання завдання вважалось власне математичне доведення студентів або доведення, знайдене в літературних джерелах, з посиланням на джерело. Представлення результату у вигляді власного доведення орієнтоване на студентів з сильною математичною підготовкою. Пошук доведення в літературних джерелах збільшує коло студентів, які бажають виконати завдання і здатні це зробити. Студентам було дозволено працювати в парах.

На додачу до згаданих умов було використано елемент змагання, який, на нашу думку, мав зменшити тривалість виконання завдання та обмежити кількість повторів у представлених студентами відповідях. Завдання зраховувалось тому виконавцю, який представив певний метод розв'язання першим. Якщо інший виконавець надавав таке саме розв'язання, то завдання йому не зраховувалось. Таким чином, розв'язків завдання могло бути декілька, але вони мали бути різні. Такий підхід також зменшив ймовірність передачі відповідей від одних студентів до інших.

Несподіваним для нас виявився фактичний час виконання завдання. Власні доведення студентів були представлені протягом двох тижнів, тобто набагато раніше, ніж ми очікували. Цей факт підтвердив те, що є студенти, які

мають сильну математичну підготовку й зацікавлені в різнопланових завданнях. Доведення формул, яке наявне в літературних джерелах, не було представлено.

Повертаючись до доведень студентів, звернемо увагу, що, використовуючи один із представлених способів доведення, можна виводити формули  $\chi^2$  через  $n$  та  $O_{i,j}$  для трьох і більше вибірок. Наприклад, коли  $n=3$ , маємо:

$$\chi^2 = \sum_{j=1}^c \frac{n_2 n_3 (n_2 + n_3) O_{1,j}^2 + n_1 n_3 (n_1 + n_3) O_{2,j}^2 + n_1 n_2 (n_1 + n_2) O_{3,j}^2 - 2 n_1 n_2 n_3 (O_{1,j} O_{2,j} + O_{1,j} O_{3,j} + O_{2,j} O_{3,j})}{n_1 n_2 n_3 (O_{1,j} + O_{2,j} + O_{3,j})} \text{ або}$$

$$\chi^2 = \sum_{j=1}^c \frac{n_3 (n_2 O_{1,j} - n_1 O_{2,j})^2 + n_2 (n_3 O_{1,j} - n_1 O_{3,j})^2 + n_1 (n_3 O_{2,j} - n_2 O_{3,j})^2}{n_1 n_2 n_3 (O_{1,j} + O_{2,j} + O_{3,j})}.$$

Однак варто зазначити, що, по-перше, формули стають громіздкими, по-друге, в дослідженнях зазвичай використовують дві вибірки (наприклад, контрольна та експериментальна групи).

## ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Проведена самостійна робота студентів-медиків показує, що серед майбутніх лікарів є представники, готові й здатні виконувати завдання високої складності, зокрема, математичного характеру. Крім того, застосування аналізу літературних джерел для вирішення поставленої проблеми може збільшити кількість студентів, охочих виконувати завдання, надаючи можливість тим, хто має слабку математичну підготовку. Представлені завдання та доведення дають змогу глибше розібратися з вибраним статистичним методом та використовувати отримані знання на практиці.

Ми практикуємо різні види завдань для самостійної індивідуальної роботи студентів. Завдання подібного типу, тобто орієнтоване на добру математичну підготовку, використали вперше. Позитивний результат проведеної роботи відкриває перспективи подальших досліджень, які будуть орієнтовані на створення нових навчальних завдань з метою перевірки й розвитку професійних здібностей, умінь та навичок майбутніх лікарів.

### Список використаних джерел

1. Беззуб І. Сучасний стан та перспективи розвитку медичної освіти в Україні. *Сайт Центру досліджень соціальних комунікацій НБУВ*. URL: [http://nbuviar.gov.ua/index.php?option=com\\_content&view=article&id=2855:suchasnij-stan-ta-perspektivi-rozvitku-medichnoji-osviti-v-ukrajini&catid=8&Itemid=350](http://nbuviar.gov.ua/index.php?option=com_content&view=article&id=2855:suchasnij-stan-ta-perspektivi-rozvitku-medichnoji-osviti-v-ukrajini&catid=8&Itemid=350) (Дата звернення 02.08.2019).
2. Гланц С. А. *Медико-биологическая статистика* / ред.: Н. Е. Бузикашвили, Д. В. Самойлова; пер.: Ю. А. Данилова. Москва: Практика, 1999. 459 с.
3. Грабарь М. И., Краснянская К. А. *Применение математической статистики в педагогических исследованиях. Непараметрические методы*. Москва: Педагогика, 1977. 136 с.
4. Іщейкіна Ю. О., Сілкова О. В. Проблеми відготовки з медичної інформатики студентів вищих навчальних закладів. *Вісник проблем біології і медицини*, 2011. Випуск 3, Т. 3(89). С. 128-129.
5. МОЗ України. *Стратегія розвитку медичної освіти в Україні*. 42 с. URL: [http://moz.gov.ua/uploads/1/8475-medical\\_education\\_analytics.pdf](http://moz.gov.ua/uploads/1/8475-medical_education_analytics.pdf) (Дата звернення 02.08.2019).
6. МОН України. *Галузевий стандарт вищої освіти підготовки на другому (магістерському) рівні Магістра у галузі знань 22 Охорона здоров'я за спеціальністю 222 Медицина*. К., 2016. 48 с.
7. Набиуліна Л. М. Использование критерия  $\chi^2$  (хи-квадрат) для проведения статистической обработки данных педагогического эксперимента. *Проблемы современной науки и образования*, 2015. № 11(41). С. 187-191.
8. Новиков Д. А. *Статистические методы в педагогических исследованиях (типовые случаи)*. Москва: МЗ-Пресс, 2004. 67 с.
9. Пудова С. С., Воевода А. Л. Формування компетентності майбутніх лікарів у питаннях біостатистики при вивченні медичної інформатики. *Матеріали III Міжнародної науково-методичної конференції «Розвиток інтелектуальних умінь і творчих здібностей учнів та студентів у процесі навчання дисциплін природничо-математичного циклу «ІТМ\*плюс – 2018» (Суми, 8-9 листопада 2018 року): у 2 томах. Т. 1*. Суми: ФОП Цьома С. П., 2018. С.129-131.
10. Стариченко Б. Е. Обработка и представление данных педагогических исследований с помощью компьютера. Урал. гос. пед. ун-т. Екатеринбург, 2004. 218 с.
11. Суходольский Г. В. *Основы математической статистики для психологов : учебник*. Санкт-Петербург: Изд-во С.-Петербург. Ун-та, 1998. 464 с.
12. Толкачев А. С. Автоматизированная проверка статистических гипотез в педагогике. *Наука и образование*, 2017. №7(18). С. 13-15.
13. Franke T. M., Ho T., Christie C. A. The Chi-Square Test: Often Used and More Often Misinterpreted. *American Journal of Evaluation*, 2012. Vol. 33, Issue 3. Pp. 448-458. DOI: 10.1177/1098214011426594.
14. Harris C. W., Fu S. The blend justifies the means: the relationships between lecturer use of online content in-class, student age, and student self-stated understanding, with students' use of online content out-of-class in a blended commerce course. *Journal of Applied Learning & Teaching*, 2018. Vol. 1, #1. Pp. 5-12. URL: <http://journals.sfu.ca/jalt/index.php/jalt/article/view/5/1> (Дата звернення 02.08.2019).
15. Hugh M. L. The Chi-square test of independence. *Biochemia Medica*, 2013. Vol. 23, Issue 2. Pp. 143-149. DOI: 10.11613/BM.2013.018.
16. Onchiri S. Conceptual model on application of chi-square test in education and social sciences. *Academic Journals*, 2013. Vol.8(15). Pp. 1231-1241. DOI: 10.5897/ERR11.0305.
17. Pandis N. The chi-square test. *American Journal of Orthodontics and Dentofacial Orthopedics*, 2016. Vol. 150, Issue 5. Pp. 898-899. DOI: 10.1016/j.ajodo.2016.08.009.
18. Rana R., Singhal R. Chi-square Test and its Application in Hypothesis Testing. *Journal of the Practice of Cardiovascular Sciences*, 2015. Vol. 1, Issue 1. Pp. 69-71. DOI: 10.4103/2395-5414.157577.
19. Ugoni A., Walker B. F. The Chi square test. *COMSIG Review*, 1995. Vol. 4, #3. Pp. 61-64. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2050386/pdf/cr043-061b.pdf> (Дата звернення 02.08.2019).

## References

1. Bezzub, I. Suchasnyi stan ta perspektyvy rozvytku medychnoi osvity v Ukraini [Current state and prospects of medical education development in Ukraine]. Sait Tsentru doslidzhen sotsialnykh komunikatsii NBUV [Site of Social Communications Research Center VNLU]. Retrieved from [http://nbuviap.gov.ua/index.php?option=com\\_content&view=article&id=2855:suchasnij-stan-ta-perspektivi-rozvitku-medichnoji-osvity-v-ukrajini&catid=8&Itemid=350](http://nbuviap.gov.ua/index.php?option=com_content&view=article&id=2855:suchasnij-stan-ta-perspektivi-rozvitku-medichnoji-osvity-v-ukrajini&catid=8&Itemid=350) [in Ukrainian].
2. Glantz, S. A. (1999). *Mediko-biologicheskaja statistika [Primer of Biostatistics]*. (Ju. A. Danilova). Moskva: Praktika [in Russian].
3. Grabar', M. I. & Krasnjanskaja, K. A. (1977). *Primenenie matematicheskoy statistiki v pedagogicheskikh issledovanijah. Neparаметричeskie metody [The Use of Mathematical Statistics in Pedagogical Research. Non-parametric Methods]*. Moskva: Pedagogika [in Russian].
4. Ishcheikina, Yu. O. & Silkova, O. V. (2011). Problemy vidhotovky z medychnoi informatyky studentiv vyshchych navchalnykh zakladiv [Training Problems in Medical Informatics of the Higher Educational Institutions Students]. *Visnyk problem biologii i medytsyny – Bulletin of Problems in Biology and Medicine*, (3), V.3(89), 128-129 [in Ukrainian].
5. MOZ Ukrainy [MH of Ukraine]. Stratehiia rozvytku medychnoi osvity v Ukraini [Strategy for the development of medical education in Ukraine]. Retrieved from [http://moz.gov.ua/uploads/1/8475-medical\\_education\\_analytics.pdf](http://moz.gov.ua/uploads/1/8475-medical_education_analytics.pdf) [in Ukrainian].
6. MON Ukrainy [MES of Ukraine]. (2016). *Haluzevyi standart vyshchoi osvity pidhotovky na druhomu (mahisterskomu) rivni Mahistra u haluzi znan 22 Okhorona zdorovia za spetsialnistiu 222 Medytsyna [Branch Standard of Higher Education of Preparation at the Second (Master's) Level of Master in the Field of Knowledge 22 Healthcare by Specialty 222 Medicine]*. Kyiv [in Ukrainian].
7. Nabiulina, L. M. (2015). Ispol'zovanie kriterija  $\chi^2$  (hi-kvadrat) dlja provedenija statisticheskoy obrabotki dannyh pedagogicheskogo jeksperimenta [Using the  $\chi^2$  (Chi-square) Test for the Statistical Processing of the Pedagogical Experiment Data]. *Problemy sovremennoj nauki i obrazovanija – Problems of Modern Science and Education*, 11(41), 187-191 [in Russian].
8. Novikov, D. A. (2004). *Statisticheskie metody v pedagogicheskikh issledovanijah (tipovye sluchai) [Statistical Methods in Pedagogical Research (Typical Cases)]*. Moskva: MZ-Press. [in Russian].
9. Pudova, S. S. & Voievoda, A. L. (2018). Formuvannia kompetentnosti maibutnykh likariv u pytanniakh biostatystyky pry vyvchenni medychnoi informatyky [Forming the Competence of Medical Students in Questions of Biostatistics in the Process of Studying Medical Informatics]. *Materialy III Mizhnarodnoi naukovo-metodychnoi konferentsii «Rozvytok intelektualnykh umin i tvorchykh zdbnostei uchniv ta studentiv u protsesi navchannia dystsyplin pryrodnycho-matematichnoho tsykladu «ITM\*plus – 2018» – The Third International Scientific and Methodical Conference “Development of intellectual skills and creative abilities of students in the process of teaching disciplines of the natural-mathematical cycle “ITM\*plus – 2018”* (pp. 129-131). Sumy: FOP Tsoma S. P [in Ukrainian].
10. Starichenko, B. E. (2004). *Obrabotka i predstavlenie dannyh pedagogicheskikh issleovanij s pomoshh'ju komp'jutera [Processing and Presentation of Pedagogical Research Data Using a Computer]*. Ural. gos. ped. un-t. Ekaterinburg [in Russian].
11. Suhodol'skij, G. V. (1998). *Osnovy matematicheskoy statistiki dlja psihologov : uchebnik [Basics of Mathematical Statistics for Psychologists : a textbook]*. Sankt-Peterburg: Izd-vo S.-Peterburg. Un-ta [in Russian].
12. Tolkachev, A. S. (2017). Avtomatizirovannaja proverka statisticheskikh gipotez v pedagogike [Automated Verification of Statistical Hypotheses in Pedagogy]. *Nauka i obrazovanie – Science and Education*, 7(18), 13-15 [in Russian].
13. Franke, T. M., Ho, T. & Christie, C. A. (2012). The Chi-Square Test: Often Used and More Often Misinterpreted. *American Journal of Evaluation*. 33(3). 448-458. DOI: 10.1177/1098214011426594.
14. Harris, C. W. & Fu, S. (2018). The Blend Justifies the Means: the Relationships Between Lecturer Use of Online Content In-Class, Student Age, and Student Self-Stated Understanding, with Students' Use of Online Content Out-of-Class in a Blended Commerce Course. *Journal of Applied Learning & Teaching*, 1(1), 5-12. Retrieved from <http://journals.sfu.ca/jalt/index.php/jalt/article/view/5/1>.
15. Hugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, 23(2), 143-149. DOI: 10.11613/BM.2013.018.
16. Onchiri, S. (2013). Conceptual Model on Application of Chi-Square Test in Education and Social Sciences. *Academic Journals*, 8(15), 1231-1241. DOI: 10.5897/ERR11.0305.
17. Pandis, N. (2016). The Chi-Square Test. *American Journal of Orthodontics and Dentofacial Orthopedics*, 150(5), 898-899. DOI: 10.1016/j.ajodo.2016.08.009.
18. Rana, R. & Singhal, R. (2015). Chi-square Test and its Application in Hypothesis Testing. *Journal of the Practice of Cardiovascular Sciences*, 1(1), 69-71. DOI: 10.4103/2395-5414.157577.
19. Ugoni, A. & Walker, B. F. (1995). The Chi Square Test. *COMSIG Review*, 4(3), 61-64. Retrieved from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2050386/pdf/cr043-061b.pdf>.

## MEDICAL STUDENTS' INDIVIDUAL INDEPENDENT WORK ON MEDICAL INFORMATICS FOR STUDYING THE CHI-SQUARE TEST

S. Pudova, G. Solasa, S. Metwally

National Pirogov Memorial Medical University, Vinnytsa, Ukraine

**Abstract.**

**Formulation of the problem.** According to the curriculum, future doctors begin to learn the elements of statistics from the first courses of study. In particular, they study biostatistics methods in medical informatics classes in the second year of study. At the same time, most applicants to higher medical institutions in Ukraine have low physical and mathematical training. Are modern medical students capable of performing complex mathematical problems and are they interested in such tasks?

**Materials and methods.** In the process of the pedagogical study, the literature on pedagogy and statistics was analyzed. Processed theoretical information from statistics on the chi-square test was used to create a problematic task for individual independent work of students in medical informatics. MS Excel was additionally used to create demo materials for calculating the empirical value of the

*chi-square coefficient. At the final stage, the completed tasks were checked and the pedagogical situation regarding the work done was analyzed.*

**Results.** *English-speaking second year medical students of the Faculty of Medicine have completed their task much faster than we had assumed. During the task performance, students demonstrated a fairly high level of mathematical knowledge, skills and interest in the task. Future medical doctors presented two different proofs of the equality of the two formulas (written in the task) for calculation empirical value of the chi-square coefficient. Using student proof ways, it is possible to derive the formula for calculating the empirical value of the chi-square coefficient for three samples, which is also presented in this article.*

**Conclusions.** *On the one hand, individual independent work of medical students demonstrated the willingness of future doctors to solve problems that are not standard for their profile and have an immersion in the mathematics. On the other hand, the content of the task and the mathematical proofs of the students make it possible to understand more deeply the chosen statistical method and to use the acquired knowledge in practice.*

**Key words:** *individual independent work of students, medical informatics, future doctors, chi-square test (Pearson's chi-squared test), chi-square statistic formula.*