

and psychological assistance), resource assistance (help in natural resources – clothes, food products and money) to instrumental support (providing services) and possessing connections, suggesting exchange of protections, recommendations, information, etc. The last two types of support refer to the active usage of social capital of ties in which respondents are included (these are help in employment, studying, mediation in approach to officials) and are available to the minority of population.

I will continue research this topic. Global gender gap, chances for happy motherhood and worthy life in old age as a three difference profiles in opportunities will be main focuses in the next article.

Reference

1. Saakova, I. (2013), "Where in the Ukraine to live well", http://www.consumerinfo.org.ua/must_know/quality/935/7585/
2. Muradyan, Ye. S., Salnikova, S. A., Titarenko, L. G., Shirokanova, A. O. (2014), "The dynamics of value and norm system and life chances: the experience of post-Soviet transformation in the Borderlands", Vilnius : YeGU, 366 p.

Мурадян Олена. Життєві шанси й можливості в Україні: суб'єктивна оцінка об'єктивної реальності. У цій статті проаналізовано стан невизначеності та наявність великої кількості соціальних груп з обмеженим життєвим ресурсом і невисоким статусом як характеристики українського суспільства, що актуалізує увагу на адаптаційному підході до аналізу життєвих стратегій та пошуку відповіді на питання: яка система ціннісних орієнтацій допомагає успішній внутрішній і зовнішній адаптації до існуючих умов та соціальних змін, коли можливість довгострокового планування обмежена.

Ключові слова: життєві шанси, життєві можливості, життєві ресурси.

Мурадян Елена. Жизненные шансы и возможности в Украине: субъективная оценка объективной реальности. В статье анализируются состояние неопределенности и наличие большого количества социальных групп с ограниченным жизненным ресурсом и невысоким статусом как характеристики украинского общества, что актуализирует внимание к адаптационному подходу к анализу жизненных стратегий и поиску ответа на вопрос: какая система ценностных ориентаций помогает успешной внутренней и внешней адаптации к существующим условиям и социальным изменениям, когда возможность долгосрочного планирования ограничена. Анализ данной работы направлен на рассмотрение жизненных стратегий в зависимости от факторов региональной социокультурной среды. Это позволяет предполагать варианты перспективного развития социума в целом, и определенного специфического региона в частности. При этом рассматривается соотношение жизненных шансов со структурой социальных возможностей, что происходит в процессе оценки индивидом своей исходной позиции – процессе, необходимом ему для построения жизненных стратегий. Учитываются данные об уровне жизни и социальной защищенности жителей Украины до 2014 г.

Ключевые слова: жизненные шансы, жизненные возможности, жизненные ресурсы.

Article was received 2015/06/12

УДК 316.2:303.448

Світлана Сальнікова

Застосування критерію Пірсона до визначення структури тестових завдань

У статті продемонстровано застосування критерію Пірсона, а саме процедури побудови двовимірної таблиці для двох ознак із відсутністю статистичного зв'язку між ними, до розв'язання актуального прикладного завдання, зокрема визначення кількісних параметрів структури тестових завдань як першого кроку в забезпеченні їх валідності. У встановленні структури тестових завдань послуговуємося державними стандартами вищої освіти України та орієнтуємося на створення бази завдань для комп'ютерного тестування. Процедури обчислення кількісних параметрів структури тестових завдань представлені в табличному форматі й можуть використовуватися розробником тестів як зразок.

Ключові слова: критерій Пірсона, Хі-квадрат, валідність, статистична незалежність, тестові завдання.

Постановка наукової проблеми та її значення. Створення потужних статистичних пакетів для обробки й аналізу соціологічних даних, і не лише соціологічних на кшталт OCA, SPSS, R, Statistica, значно спростило роботу із великими масивами даних, проте лишається значний спектр завдань, розв'язання котрих потребує здійснення звичайних «паперових» розрахунків. Отже, актуальними лишаються знання процедури таких обчислень, з одного боку, і сфери задач, на котрі вони поширюються, – з іншого.

Однією із таких процедур є критерій Пірсона, або критерій Хі-квадрат, застосування якого, наприклад, у статистичному пакеті SPSS є вкрай простим [1, с. 64–66], а обчислення самотужки потребує численних арифметичних розрахунків, через які фахівцям (чи майбутнім фахівцям) з обробки даних знання цього критерію видається зайвим.

З іншого боку, масовість використання тестових завдань, їх популяризація в усіх сферах життя створює іншу наукову проблему – проблему валідності та надійності розроблених тестів. У межах цієї статті не можна вирішити це питання повністю, але перший крок до цього буде зроблено.

Аналіз досліджень цієї проблеми. На сьогодні існує велика кількість літератури з математичного аналізу великих масивів даних, у тому числі спеціалізованої під певний статистичний пакет, у якій критерій Пірсона розглянуто як статистичний метод із поданням процедури його реалізації та інтерпретації (див. [1, с. 161], а також [2, с. 65–80]). Значно менше публікацій стосовно створенню валідних тестових завдань; переважно це стосується психодіагностичних методик¹, адже в їх основі лежать, зазвичай, тести. При створенні цільової бази тестових завдань послуговуються державним стандартом вищої освіти України [3], на нього ми й посилатимемося.

Мета цього дослідження – застосування критерію Пірсона до визначення кількісних параметрів структури тестових завдань як першого кроку в забезпеченні їх валідності. Досягнення мети передбачає виконання таких **завдань**: (1) представити ідею Пірсона щодо пошуку наявності / відсутності зв'язку між двома ознаками, а також деякі процедури відповідного критерію, (2) визначити структуру тестових завдань та її кількісні параметри, застосувавши критерій Пірсона.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Попередньо зауважимо, що стаття орієнтована не стільки на фахівців у сфері математичного аналізу емпіричних даних (тут вони не зовсім будуть вдоволені термінологічними спрощеннями й неповнотою подання критерію Пірсона), скільки на майбутніх фахівців для покращення їхнього сприйняття у вивченні критерію Пірсона, а також усіх тих, хто причетний до формування бази тестових завдань із будь-яких навчальних дисциплін для різних освітніх рівнів.

Основна мета *критерію Пірсона* – установлення наявності / відсутності зв'язку між двома ознаками, емпіричний двовимірний розподіл для яких можна подати у вигляді таблиці.

Таблиця 1

Таблиця емпіричного двовимірного розподілу двох ознак, $\{N_{ij}\}$

X	Y				N(X)
	y_1	y_2	...	y_l	
x_1	N_{11}	N_{12}	...	N_{1l}	$N(x_1)$
x_2	N_{21}	N_{22}	...	N_{2l}	$N(x_2)$
...	...	...	...	...	...
x_k	N_{k1}	N_{k2}	...	N_{kl}	$N(x_k)$
N(Y)	N(y₁)	N(y₂)	...	N(y_l)	N

У табл. 1 представлено розподіл між ознаками X (має k категорій) та Y (має l категорій). Внутрішньоклітинні частоти N_{ij} , де i – будь-яка категорія ознаки X (від 1 до k), j – будь-яка категорія ознаки Y (від 1 до l), таблиці $\{N_{ij}\}$ будуть числами цілими й неоднаковими. І хоча кореляційний зв'язок спостерігається тоді, коли зі зміною (зростанням чи спаданням) однієї ознаки інша також змінюється (зростає і/або спадає чи спадає і/або зростає), установити його наявність / відсутність, на

¹ Стандартні вимоги до валідності та надійності можна подивитись у: Рукавишников А. А., Соколова М. В. Нормы профессиональной этики для разработчиков и пользователей психодиагностических методик. Стандартные требования к психологическим тестам. – Ярославль : Дебют, 1991.

перший погляд, неможливо. По-перше, тому що навіть різні, з математичної точки зору, числа статистично можуть бути однаковими (наприклад, $12 \neq 15$, але, порівняно з обсягом N чи маргінальними сумами $N(x_i)$, $N(y_l)$, ці числа будуть статистично однаковими). По-друге, одні частоти, порівняно з іншими, можуть бути значно меншими не тому, що зміна однієї ознаки приводить до зміни іншої, а тому, що відповідна категорія є малонаповненою (наприклад когорта з найменшим освітнім рівнем).

Основна ідея Пірсона полягає в тому, щоб порівняти емпіричну таблицю двовимірного розподілу $\{N_{ij}\}$ з теоретичною таблицею $\{N_{ij}^0\}$ такої ж розмірності (відповідної кількості рядків і стовпчиків), для якої статистичної залежності між ознаками не існуватиме (або, що те ж саме, порівняти емпіричний і теоретичний двовимірні розподіли). Логіка така:

- якщо таблиці однакові, то з відсутності статистичного зв'язку між ознаками в теоретичній таблиці впливає відсутність статистичного зв'язку в емпіричній;
- якщо таблиці різні, то з відсутності статистичного зв'язку між ознаками в теоретичній таблиці впливає наявність статистичного зв'язку в емпіричній.

Тепер процедури критерію полягають у (1) побудові теоретичної таблиці двовимірного розподілу та (2) порівнянні двох таблиць.

Теоретична таблиця двовимірного розподілу двох ознак $\{N_{ij}^0\}$ матиме таку ж кількість рядків і стовпців (k і l , відповідно), ті ж маргінали (суми по рядках $N(x_i)$ і стовпцях $N(y_l)$), а отже, загальну кількість спостережень N ; іншими будуть тільки внутрішньоклітинні частоти N_{ij} – тепер вони мають розподілитися так, аби між X та Y не було залежності.

Арифметика обчислення теоретичних внутрішньоклітинних частот N_{ij}^0 проста: потрібно перемножити відповідні маргінали й поділити на загальну кількість (див. табл. 2).

Таблиця 2

Таблиця теоретичного двовимірного розподілу двох ознак, $\{N_{ij}^0\}$

X	Y				N(X)
	y_1	y_2	...	y_l	
x_1	$N_{11}^0 = \frac{N(y_1) \times N(x_1)}{N}$	$N_{12}^0 = \frac{N(y_2) \times N(x_1)}{N}$	...	$N_{1l}^0 = \frac{N(y_l) \times N(x_1)}{N}$	$N(x_1)$
x_2	$N_{21}^0 = \frac{N(y_1) \times N(x_2)}{N}$	$N_{22}^0 = \frac{N(y_2) \times N(x_2)}{N}$	...	$N_{2l}^0 = \frac{N(y_l) \times N(x_2)}{N}$	$N(x_2)$
...	...	...	...	...	...
x_k	$N_{k1}^0 = \frac{N(y_1) \times N(x_k)}{N}$	$N_{k2}^0 = \frac{N(y_2) \times N(x_k)}{N}$	...	$N_{kl}^0 = \frac{N(y_l) \times N(x_k)}{N}$	$N(x_k)$
N(Y)	N(y₁)	N(y₂)	...	N(y_l)	N

Маємо дві таблиці. Установлення статистичної однаковості / неоднаковості двох різних таблиць, «на перший погляд», знову ж таки є неможливим; тут використовуємо спеціальну процедуру порівняння: обчислення деякого числа χ^2 (читається Хі-квадрат; квадрат означає, що число додатне), за значенням¹ якого встановлюють, чи є відмінність між таблицями випадковою (тобто вони неоднакові виключно з математичної точки зору), чи вони дійсно різні.

Отже, обчислюємо емпіричне значення χ^2 , потім у таблиці критичних значень критерію t Стюдента [2, с. 259] шукаємо теоретичне значення χ_0^2 на перетині стовпців df (ступінь свободи² від англ. *degree of freedom*) та рівня значущості (прийнятна для нас величина помилки).

¹ Для порівняння двох таблиць використовується ідея знання різниці двох чисел: якщо різниця дорівнює нулю, то числа однакові. Але тут статистично порівнюються не два числа, а дві таблиці з «числами-внутрішньоклітинними частотами», тому для початку знаходимо попарно всі різниці між відповідними внутрішньоклітинними частотами обох таблиць ($N_{ij} - N_{ij}^0$), потім – їх суму; попередньо різниці треба піднести до квадрата, аби уникнути сумарного нуля, а потім розділити на N_{ij}^0 , щоб позбутися величини у квадраті. Отже, формула для розрахунку числа, яке відіграватиме роль нульової точки, є такою $\chi^2 = \sum_{i=1}^k \sum_{j=1}^l \frac{(N_{ij} - N_{ij}^0)^2}{N_{ij}^0}$. Це число називається

мірою відхилення емпіричної таблиці від теоретичної.

² Це число, яке обчислюється як $(k-1) \cdot (l-1)$, – важливий показник, тому що критерій чутливий до розміру (кількість рядків, кількість стовпців) і форми (квадратна при $k = l$, видовжена при $k \neq l$) таблиці.

Маємо два числа, як і дві таблиці, – емпіричне χ^2 і теоретичне χ_0^2 – їх і порівнюємо. Фактично, число χ_0^2 є своєрідним нулем на осі, праворуч якого містяться більші за нього значення (значимі), а ліворуч – менші (незначимі). Тобто, якщо розраховане на основі емпіричних даних число виявиться меншим за табличне чи таким самим ($\chi^2 \leq \chi_0^2$), то різниця між таблицями є незначимою, тобто випадковою. Якщо ж розраховане число перевищує табличне ($\chi^2 > \chi_0^2$), то різниця між таблицями є значимою, невідповідною. Тільки тоді, коли емпіричне значення χ^2 значно перевищує теоретичне χ_0^2 , можемо з певною ймовірністю стверджувати про наявність статистичного зв'язку між двома досліджуваними ознаками X та Y.

Отже, підсумуємо: основна ідея – це порівняння двох двовимірних таблиць однакового розміру, на основі знання про одну з них отримуємо знання про іншу. Цим знанням виступає статистична незалежність або статистична залежність (іншими словами, відсутність або наявність кореляційного зв'язку). Задля реалізації цієї ідеї потрібно побудувати двовимірну таблицю для двох статистично незалежних ознак. Власне цю процедуру ми й зrealізуємо під час виконання другого завдання статті.

Для початку визначимо структуру тестових завдань, тут для зразка послуговуватимемося державним стандартом вищої освіти України [3], згідно з яким визначено два формати (форми) тестових завдань (тестів) – (1) завдання відкритої форми з вільно конструйованими відповідями та (2) закритої форми із запропонованим «віялом» відповідей, однієї, декількох (не більше п'яти), установленим відповідностей. Ураховуючи впровадження комп'ютерного тестування як більш надійної та ефективної форми контролю знань, обмежимося закритою формою тестових завдань; ураховуючи також диференційну особливість знань тих, у кого їх (знання) передбачається виміряти, зупинимося на багаторівневих тестах. Існують різні практики виокремлення кількості рівнів, ми ж візьмемо класичні три рівні складності.

І типи тестових завдань, і рівні їх складності повинні бути в прийнятному співвідношенні у загальній базі даних. При цьому кожен тип тесту має бути представлений усіма трьома рівнями складності. Співвідношення типів та рівнів багато в чому залежить від цільового призначення бази тестових знань у цілому або ж значимості (вагомості) окремого розділу чи навчальної дисципліни в межах програми тестування. Для прикладу встановимо такі співвідношення для трьох типів тестових завдань і трьох рівнів складності як 25 % / 35 % / 40 % та 50 % / 30 % / 20 % відповідно.

Отже, маємо дві ознаки: X – «типи тестів» із трьома категоріями x_1 – «одна правильна відповідь», x_2 – «дві правильні відповіді», x_3 – «три-чотири правильні відповіді»; Y – «рівні складності тестів» також із трьома категоріями y_1 – «перший рівень складності», y_2 – «другий рівень складності», y_3 – «третій рівень складності».

Обсяг бази тестових завдань певної навчальної дисципліни має становити не менше 100 тестів¹. Нехай у цьому випадку нам потрібно 120. Розрахуємо кількість тестів по кожній категорії кожної з ознак.

Таблиця 3

Приклад розрахунку кількостей типів і рівнів складності тестових завдань

Ознака	Категорія	%	$N(x_i)$	Ознака	Категорія	%	$N(y_j)$
X – «типи тестів»	x_1 – «1 правильна відповідь»	25	$\frac{25 \cdot 120}{100} = 30$	Y – «рівні складності тестів»	y_1 – «I рівень складності»	50	$\frac{50 \cdot 120}{100} = 60$
	x_2 – «2 правильні відповіді»	35	$\frac{35 \cdot 120}{100} = 42$		y_2 – «II рівень складності»	30	$\frac{30 \cdot 120}{100} = 36$
	x_3 – «3–4 правильні відповіді»	40	$\frac{40 \cdot 120}{100} = 48$		y_3 – «III рівень складності»	20	$\frac{20 \cdot 120}{100} = 24$
Загалом		100	120	Загалом		100	120

¹ Зауваження щодо мінімального обсягу бази тестових завдань ґрунтується на одній з умов застосування критерію Пірсона: загальний обсяг повинен задовольняти умову $N \geq 100$. З іншого боку, є також вимога до загальної кількості тестів, яка ґрунтується на таких показниках, як кількість тестів у варіанті, кількість варіантів, кількість осіб, які одночасно проходять тестування.

Знаючи маргінальні суми й послуговуючись процедурою побудови двовимірної таблиці з розподіленими в ній таким чином частотами, аби між ознаками не існувало статистичної залежності, розрахуємо кількісні параметри заданої структури тестових завдань у вже знайомому табличному форматі.

Таблиця 4

Приклад розрахунку кількісних параметрів структури тестових завдань

X	Y			N(X)
	Y ₁	Y ₂	Y ₃	
x ₁	$\frac{60 \times 30}{120} = 15$	$\frac{36 \times 30}{120} = 9$	$\frac{24 \times 30}{120} = 6$	30
x ₂	$\frac{60 \times 42}{120} = 21$	$\frac{36 \times 42}{120} = 13$	$\frac{24 \times 42}{120} = 8$	42
x ₃	$\frac{60 \times 48}{120} = 24$	$\frac{36 \times 48}{120} = 14$	$\frac{24 \times 48}{120} = 10$	48
N(Y)	60	36	24	120

Отже, згідно з даними табл. 4, тести розподілено відповідно до вказаного співвідношення за кожною ознакою, і за таких їх кількісних параметрів не буде надмірного навантаження певними типами завдань на певному рівні складності. Тепер розробник чітко знатиме, що зі 120 тестів у нього має бути 15 завдань з однією правильною відповіддю першого рівня складності, дев'ять завдань з однією правильною відповіддю другого рівня складності та шість завдань з однією правильною відповіддю третього рівня складності й т. ін.

Висновки та перспективи подальших досліджень. Сфера застосування математичних методів не обмежується великими дослідницькими проектами, обробка та аналіз котрих здійснюються за допомогою спеціалізованих статистичних пакетів; ідеї їх створення можуть бути корисними в багатьох інших прикладних задачах.

У межах цієї статті продемонстровано застосування критерію Пірсона, а саме процедури побудови двовимірної таблиці для двох ознак із відсутністю статистичного зв'язку між ними, до визначення кількісних параметрів структури тестових завдань як першого кроку в забезпеченні їх валідності. І хоча структура тестових завдань обрана автором статті як приклад і може варіювати залежно від цільового призначення бази таких завдань, усе ж тут послуговувалися державними стандартами вищої освіти України й орієнтувалися на комп'ютерне тестування. Наочні розрахунки кількісних параметрів структури тестових завдань представлено в табл. 3 і 4. Вони можуть слугувати зразком для випадку створення реальної бази тестів, єдине обмеження для котрої – одночасний розрахунок кількісних параметрів принаймні для 100 тестів.

Подальші дослідження можна акцентувати як на розробці тих прикладних завдань, у виконанні яких сприяли б певні ідеї при створенні математичних методів чи самі методи, так і на розширенні дослідницьких можливостей математичних методів.

Джерела та література

1. Горбачик А. П. Аналіз даних соціологічних досліджень засобами SPSS : навч. посіб. / А. П. Горбачик, С. А. Сальнікова. – Луцьк : РВВ «Вежа» Волин. нац. ун-ту ім. Лесі Українки, 2008. – 164 с.
2. Паніотто В. І. Статистичний аналіз соціологічних даних / В. І. Паніотто, В. С. Максименко, Н. М. Харченко. – К. : Вид. дім «КМ Академія», 2004. – 270 с.
3. Комплекс нормативних документів для розробки складових систем стандартів вищої освіти. Засоби діагностики якості вищої освіти // Інформаційний вісник «Вища освіта». – 2003. – № 10. – С. 67–82.

References

1. Gorbachyk, A. P., Salnikova, S. A. (2008), "Analysis of the data of sociological research by means of SPSS", Lutsk : «Vezha» Lesya Ukrainka Volyn National University, 164 p.
2. Paniotto, V. I., Maksymenko, V. S., Kharchenko, N. M. (2004), "Statistical analysis of sociological data", K. : Ed. House «KM Academy», 270 p.

3. "Complex of normative documents for the development of components of higher education. Diagnostic Quality of Higher Education" (2003), *Information Bulletin «Higher Education»*, No. 10, Pp. 67–82.

Сальникова Светлана. Применение критерия Пирсона к определению структуры тестовых заданий.

Демонстрируется применение критерия Пирсона, а именно процедуры построения двумерной таблицы для двух статистически независимых переменных, к определению количественных параметров структуры тестовых заданий. Структура тестов определяется согласно государственным стандартам высшего образования Украины и является ориентированной на компьютерное тестирование. Процедуры расчета количественных параметров структуры тестовых заданий представлены в табличном формате.

Ключевые слова: критерий Пирсона, Хи-квадрат, валидность, статистическая независимость, тестовые задания.

Salnikova Svitlana. The Application of Pearson's Criterion for Determination of Structure of Tests. The emergence of international projects and the availability of powerful statistical packages for data analysis contributed to displacement of attention of scientists from development and expanding the scope of application of mathematical methods to their applicability in comparative projects and ways of realization in computer programs.

The author shows the application of Pearson criterion, exactly - the procedure of construction of the two dimensional table for two indicators with the lack of statistical relationship between them, for determination of quantitative parameters of structure of tests as the first step in ensuring their validity. The structure of the tests is determined according to state standards of higher education in Ukraine and is focused on the computer test. Procedures for calculating the quantitative parameters of structure of tests are presented in table format and can be used by developer of tests as a sample. There is some limitation of the proposed procedure; this is the calculation of quantitative parameters for over 100 tests for one an academic subject or its parts.

Key words: Pearson criterion, chi-square, the validity, statistical independency, tests.

Статтю отримано 18.01.2015 р.