

МАТЕМАТИКА

МАТЕМАТИЧНІ МЕТОДИ АНАЛІЗУ РЕЗУЛЬТАТІВ ПЕДАГОГІЧНОГО ЕКСПЕРИМЕНТУ

Гончаренко Я.В.,

*кандидат фіз.-мат. наук, доцент,
Національний педагогічний університет імені М.П.Драгоманова*

Горбачук В.О.,

*аспірант,
Національний педагогічний університет імені М.П.Драгоманова*

Стаття присвячена проблемі використання різних математичних методів обробки статистичних результатів експерименту та перевірки статистичних гіпотез, що підтверджують ефективність проведеного педагогічного дослідження.

Статья посвящена проблеме использования различных математических методов обработки статистических результатов эксперимента и проверки статистических гипотез, подтверждающих эффективность проведенного педагогического исследования.

This paper is devoted to problem of using different math methods to analyzing of statistic results of experiment and checking of statistic hypothesis, that confirming the effectiveness of done pedagogical experiment.

Постановка проблеми. Існує велика кількість визначень поняття «педагогічний експеримент». У всіх цих визначеннях поняття «педагогічний експеримент» присутня основна думка про те, що педагогічний експеримент — це науково обгрунтована і добре продумана система організації педагогічного процесу, спрямована на відкриття нового педагогічного знання, перевірку і обгрунтування наукових припущень або гіпотез.

Педагогічний експеримент передбачає активний вплив на педагогічне явище чи процес шляхом створення нових умов, що відповідають меті дослідження. Експеримент, на відміну від інших методів, створює умови для:

- 1) перевірки ефективності запроваджень нових методичних систем, методик, методів, форм, засобів тощо у навчально-виховний процес;
- 2) порівняння ролі та впливу різних факторів на педагогічний процес;
- 3) вибору оптимальних факторів для організації певних ситуацій навчання та виховання;
- 4) виявлення умов реалізації певних педагогічних задач;
- 5) виявлення специфіки та закономірностей протікання педагогічного процесу в конкретних, в тому числі й наперед заданих, умовах.

Отже, експеримент у загальній системі методів дослідження допомагає встановити наукові факти, пояснити та узагальнити нові дані з позицій більш загальних теорій; будувати на базі одержаних результатів нові гіпотези та теорії.

Як правило, розрізняють такі види педагогічного експерименту:

1. Констатуючий експеримент полягає в тому, що дослідник експериментальним шляхом встановлює лише стан педагогічної системи, що вивчається: констатує наявність зв'язків, залежностей між явищами, визначає вихідні дані для подальшого дослідження.

2. Формуючий експеримент супроводжується застосуванням спеціально розробленої системи заходів, спрямованих на формування в учнів (студентів) певних якостей, на покращення результатів їх навчання, виховання, трудової діяльності тощо.

3. Контролюючий експеримент визначає наявність змін у рівні знань, умінь та навичок за матеріалами формуючого експерименту.

Для одержання достовірної інформації важливим є виділення з об'єкту дослідження тої сукупності, яка буде підлягати вивченню, тобто формування вибіркової сукупності.

Найчастіше для формуючого експерименту утворюється дві незалежні вибірки (учнів або студентів) – так звані експериментальна та контрольна групи. Навчання (розвиток, виховання) в контрольній групі здійснюється за традиційною методикою, тоді як в експериментальній групі впроваджуються результати відповідного наукового дослідження. В результаті оцінювання рівнів сформованості певних знань (компетенцій, умінь, навичок), на які здійснювався педагогічний вплив в ході дослідження, на початку та в кінці експерименту, формулюються та перевіряються статистичні гіпотези, які відображають ефективність впровадження результатів наукового дослідження.

Аналіз останніх досліджень та публікацій. Існує велика кількість монографій, підручників та навчальних посібників з теорії ймовірностей, математичної статистики та їх застосувань [1-7], в яких ґрунтовно викладено основні статистичні методики досліджень та розглянуто приклади їх застосувань, але публікацій, присвячених застосуванням математичних методів саме для обробки експериментальних даних в педагогіці, не так багато. Однією з кращих, на нашу думку, є [2], в якій розкриваються особливості застосування методів математичної статистики в психології.

З досвіду ознайомлення з статистичною обробкою результатів педагогічних експериментів в багатьох дисертаціях з педагогіки, можемо сказати, що автори дисертацій досить часто невиправдано застосовують математичні методи, зокрема, використовують параметричні критерії без встановлення типу розподілу генеральної сукупності.

Мета статті. Дана стаття присвячена проблемі використання різних математичних методів обробки статистичних результатів експерименту та перевірки статистичних гіпотез, що підтверджують ефективність проведеного педагогічного дослідження.

Виклад основного матеріалу. Нагадаємо, що статистичні критерії поділяються на параметричні та непараметричні [1,2,4]. Параметричними називають критерії, які включають у формулу розрахунку параметри розподілу. Тобто середні і дисперсії (нормальний розподіл, t - критерій Стьюдента, F - критерій та ін.) Критерії, які не включають в формулу розрахунку параметри розподілу і базуються на оперуванні частотами або рангами (критерій Q Розенбаума, критерій T Вілкоксона та ін.) називають непараметричними. І ті і інші критерії мають свої переваги та недоліки. На основі декількох посібників [1,2,7] можна оцінити можливості і обмеження тих і інших.

Таблиця 1. Можливості та обмеження параметричних і непараметричних критеріїв.

Параметричні критерії	Непараметричні критерії
-----------------------	-------------------------

1. Дозволяють прямо оцінити різницю в середніх, отриманих у двох вибірках (t - Критерій Стьюдента).	Дають можливість оцінити лише середні тенденції, наприклад, відповіді на питання, чи частіше в вибірці А зустрічаються більш високі, а в вибірці Б – більш низькі значення ознаки (критерії Q Розенбаума, U, ϕ^* - кутове перетворення Фішера та ін.)
2. Дають можливість прямо оцінити відмінності в дисперсіях (критерій Фішера).	Дають можливість оцінити лише відмінності в діапазонах варіативності ознаки (критерій ϕ^*)
3. Дають можливість виявити тенденції зміни ознаки при переході від умови до умови (дисперсійний однофакторний аналіз), але лише при умові нормального розподілу ознаки.	Дають можливість виявити тенденції зміни ознаки при переході від умови до умови при будь-якому розподілі ознаки (критерії тенденцій L Пейджа та S Джонкіра).
4. Дають можливість оцінити взаємодію двох і більше факторів в їх впливі на зміну ознаки (двофакторний дисперсійний аналіз).	Дана можливість відсутня.
5. Експериментальні дані мають відповідати двом, а іноді трьом, умовам: <ul style="list-style-type: none"> – значення ознаки виміряні по інтервальній шкалі; – розподіл ознаки є нормальним; – в дисперсійному аналізі потрібно дотримуватися вимоги рівності дисперсій в осередках комплексу. 	Експериментальні дані можуть не відповідати жодній з цих умов: <ul style="list-style-type: none"> – значення ознаки можуть бути представлені в будь-якій шкалі, починаючи зі шкали найменувань; – розподіл ознаки може бути будь-яким і спів падання його з яким-небудь теоретичним законом розподілу є необов'язковим і не потребує перевірки; – вимога рівності дисперсій відсутня.
6. Математичні розрахунки доволі складні.	Математичні розрахунки в більшості випадків прості і займають мало часу (за виключенням критеріїв згоди χ^2 і λ Колмогорова-Смірнова).
7. Якщо умови перераховані у пункті 5 виконуються, то параметричні критерії виявляються більш потужними ніж непараметричні.	Якщо умови перераховані у пункті 5 не виконуються, то непараметричні критерії виявляються більш потужними ніж параметричні, тому що вони менш чутливі до «засмічень».

Найчастіше в педагогічних дослідженнях виникають наступні види гіпотез: 1) гіпотеза про наявність; 2) порівняння розподілів.

Нижче ми пропонуємо алгоритми вибору статистичного критерію, який найдоцільніше використати в дослідженні в залежності від гіпотези та особливостей отриманих вибіркового даних.

Таблиця 2. Алгоритм вибору критерію для оцінки зсуву показника.

Скільки замірів (<i>c</i>) співставляється?							
2 заміри				3 і більше замірів			
Скільки виборок досліджено?				Скільки виборок досліджено?			
1 – експериментальна		2 – експериментальна і контрольна		1 – експериментальна		2 – експериментальна і контрольна	
При якісних зсувах або кількісних в вузькому діапазоні	При кількісних зсувах, які можна ранжувати по інтенсивності	Варіант 1: Незалежна обробка двох виборок	Варіант 2: співставлення двох виборок	При $c \leq 6$, $n \leq 12$	При $c > 6$ та/або $n > 12$	При $c \leq 6$, $n \leq 12$ окремо по кожній вибірці	При $c > 6$ та/або $n > 12$ окремо по кожній вибірці
Критерій знаків G	Критерій Вілкоксона T	Критерій знаків G або критерій Вілкоксона T	Q – критерій, U – критерій, ϕ^* – критерій,	L – критерій тенденцій Пейджа	критерій Фрідмана	L – критерій тенденцій Пейджа	критерій Фрідмана

Таблиця 3. Алгоритм вибору критерію для порівняння розподілів.

Скільки розрядів (<i>c</i>) має ознака?							
$c = 2$ (альтернативні розподіли)				$c \geq 3$			
Які розподіли співставляються?				Яку шкалу являють собою розряди?			
Емпіричне з теоретичним		2 емпіричні між собою		Номінальну розряди відрізняються лише по якості	Порядкову розряди впорядковані за зростанням ознаки		
$n \leq 50$	$n > 50$	Критерій Фішера	Критерій χ^2 Пірсона	Критерій χ^2 Пірсона	Виявлення схожості та відмінностей	Виявлення точки максимального накопичення відмінностей	
Біноміальний критерій	Критерій χ^2 Пірсона				Критерій χ^2 Пірсона	Критерій λ Колмогорова-Смірнова	

Розглянемо приклад застосування непараметричного критерію. Методами статистичного аналізу вивчимо наскільки суттєвими є відмінності між двома статистичними вибірками (КГ та ЕГ). Як правило для цього проводять перевірку статистичних гіпотез про належність обох виборок одній генеральній сукупності або про рівність середніх. При цьому використання параметричних критеріїв (таких як критерії Фішера або Стьюдента) передбачає встановлення виду розподілу генеральної сукупності, що не завжди є можливим. Тому в практиці педагогічних досліджень доцільніше використовувати непараметричні критерії статистики, такі, як критерій знаків, критерії Вілкоксона, Ван-дер-Вардена, Спірмена, кутове перетворення Фішера, які вільні від припущення про закон розподілу виборок і ґрунтуються на припущенні про незалежність спостережень.

Оскільки в нашому експерименті розглядаються дві незалежні вибірки (контрольна та експериментальна групи), які необхідно співставити за частотою досліджуваного

показника (рівні сформованості рефлексії), то найдоцільніше використати кутове перетворення Фішера φ^* . Цей критерій оцінює достовірність відмінностей між процентними долями двох вибірок, в яких реєструється досліджуваний показник.

Суть кутового перетворення Фішера полягає в переводі процентних долей в величини центрального кута, що вимірюється в радіанах. Більшій процентній долі відповідатиме більший кут φ_1 , а меншій — менший кут φ_2 , але співвідношення між ними нелінійне: $\varphi = 2 \arcsin \sqrt{P}$, де P — процентна доля, виражена в долях одиниці. При збільшенні розходжень між кутами φ_1 і φ_2 та збільшенні чисельності виборок значення критерію зростатимуть. Чим більше значення величини φ^* , тим більш ймовірно, що відмінності між вибірками є не випадковими (статистично значущими).

Гіпотези критерію Фішера:

H_0 : доля осіб, у яких проявляється досліджуваний ефект, у виборці 1 не більша, ніж у виборці 2.

H_1 : доля осіб, у яких проявляється досліджуваний ефект, у виборці 1 більша, ніж у виборці 2.

Алгоритм перевірки гіпотез за допомогою φ^* -критерію Фішера:

1) процентні співвідношення переводяться в долі одиниці, які, в свою чергу переводяться в радіани за формулою кутового перетворення Фішера: $\varphi_1 = 2 \arcsin \sqrt{P_1}$, $\varphi_2 = 2 \arcsin \sqrt{P_2}$, де P_1 і P_2 — відповідні долі, що порівнюються.

2) обчислюється спостережуване значення за формулою: $\varphi_{ait}^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$,

де n_1 і n_2 — обсяги досліджуваних вибірок.

3) здійснюється перевірка значущості отриманого критерію шляхом знаходження ймовірності отриманого емпіричного значення в t -розподілі Стюдента.

Приклад. Нехай в експерименті досліджувались рівні сформованості певних знань, вмій та компетенцій студентів контрольної та експериментальної груп до та після формувального етапу педагогічного експерименту. Виділялось три рівні сформованості відповідних знань, вмій та навичок: низький, середній та високий. В результаті дослідження були отримані наступні результати:

Таблиця 4.

Рівні	КГ (% студентів)		ЕГ (% студентів)	
	До експерименту	Після експерименту	До експерименту	Після експерименту
Низький	38,2	36,4	38,8	4,2
Середній	50,3	50,9	51,5	32,7
Високий	11,5	12,7	9,7	63,1

На основі результатів порівняльного аналізу рівнів сформованості досліджуваних показників можна сформулювати наступні гіпотези:

- 1) рівні сформованості досліджуваних знань, умінь та компетенцій до проведення експерименту в контрольній та експериментальній групах суттєво не відрізнялись (див. рис. 1);

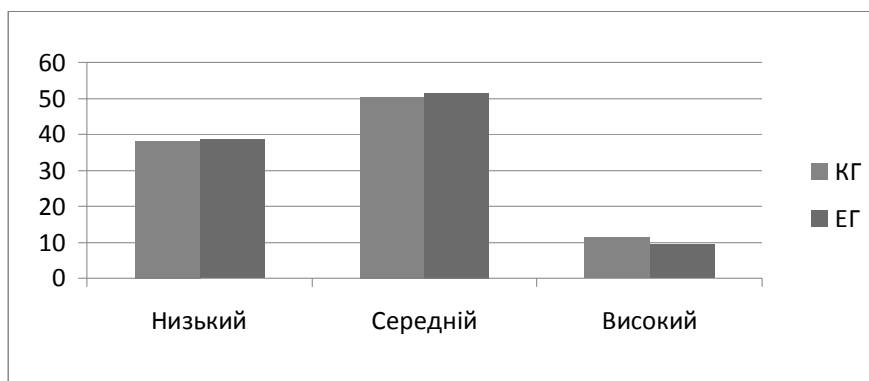


Рис. 1.

- 2) рівні сформованості досліджуваних знань, умінь та компетенцій в контрольній групі після проведення експерименту не зазнали суттєвих структурних змін (див. рис. 2);

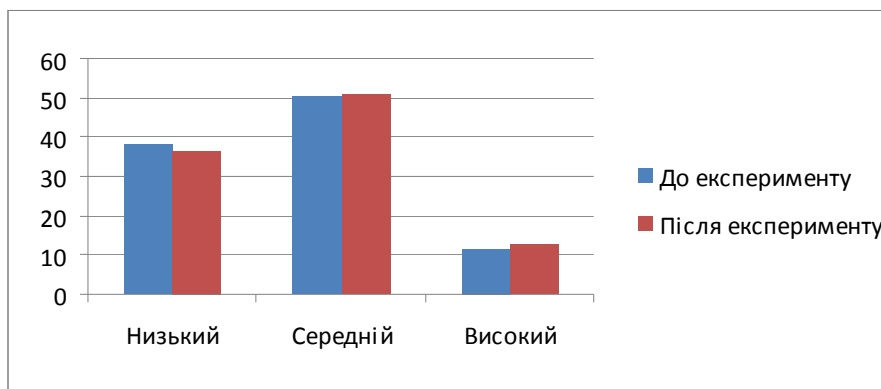


Рис.2.

- 3) рівні сформованості досліджуваних знань, умінь та компетенцій в експериментальній групі після проведення експерименту змінились суттєво – кількість студентів з низьким рівнем зменшилась, а з високим та середнім зростає (див. рис. 3).

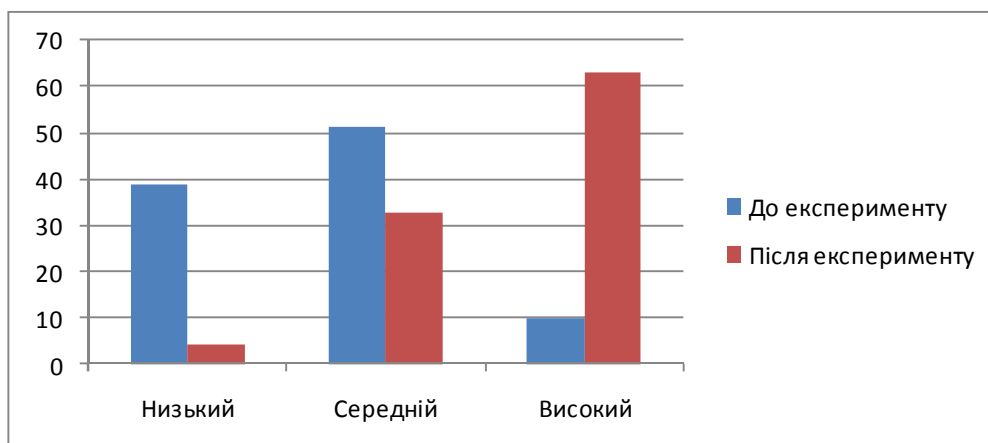


Рис.3.

Статистичну перевірку сформульованих гіпотез виконаємо за критерієм Фішера (кутове перетворення φ).

В ході обробки експериментальних даних були отримані наступні результати.

Таблиця 5. Порівняння рівнів сформованості досліджуваних знань, умінь та компетенцій до проведення експерименту в контрольній та експериментальній групах

	Рівні					
	Низький		Середній		Високий	
	КГ	ЕГ	КГ	ЕГ	КГ	ЕГ
Долі одиниці	0,382	0,388	0,503	0,515	0,115	0,097
Кутове перетворення	1,2521	1,2625	1,4520	1,4711	0,6790	0,6234
Спостережуване значення критерію φ^*	0,0695		0,1282		0,3736	
Значущість з якою приймається гіпотеза H_0	0,9447		0,8983		0,7095	
Значущість з якою приймається гіпотеза H_1	0,0553		0,1017		0,2905	

Таблиця 6. Порівняння рівнів сформованості досліджуваних знань, умінь та компетенцій в контрольній групі до та після проведення експерименту

	Рівні					
	Низький		Середній		Високий	
	До	Після	До	Після	До	Після
Долі одиниці	0,382	0,364	0,503	0,509	0,115	0,127
Кутове перетворення	1,333	1,295	1,577	1,589	0,692	0,729
Спостережуване значення критерію φ^*	0,2632		0,0849		0,2603	
Значущість з якою приймається гіпотеза H_0	0,7950		0,9332		0,7973	
Значущість з якою приймається гіпотеза H_1	0,2050		0,0668		0,2027	

Таблиця 7. Порівняння рівнів сформованості досліджуваних знань, умінь та компетенцій в експериментальній групі до та після проведення експерименту

	Рівні					
	Низький		Середній		Високий	
	До	Після	До	Після	До	Після
Долі одиниці	0,388	0,042	0,515	0,327	0,097	0,631
Кутове перетворення	1,3449	0,4128	1,6008	1,2175	0,6334	1,8359
Спостережуване значення критерію φ^*	6,5908		2,7104		8,5027	

Значущість з якою приймається гіпотеза H_0	$2,02 \cdot 10^{-6}$	0,0135	$4,49 \cdot 10^{-8}$
Значущість з якою приймається гіпотеза H_1	≈ 1	0,9865	≈ 1

Отже, в результаті статистичної перевірки результатів, можна зробити висновок, що в контрольній групі не відбулось суттєвих змін в прояві досліджуваних рівнів, відмінності між рівнями показників до та після проведення експерименту є статично незначущими (з рівнями ймовірності 0,2050, 0,0668 та 0,2027 для низького, середнього та високого рівнів відповідно). В експериментальній же групі зміни у прояві досліджуваних рівнів є суттєвими, ймовірність з якою можна прийняти гіпотезу про наявність суттєвих змін в проявах рівнів прямує до 1. Особливо суттєвими є зміни в показниках, які стосуються низького та високого рівнів. Тобто з ймовірністю близькою до 1 (100%) можна стверджувати, що доля студентів, що мали низький рівень сформованості рефлексії до експерименту, суттєво зменшилась, а доля студентів з високим рівнем суттєво зростає.

Висновки. Застосування методів математичної статистики до обробки результатів педагогічного експеримента передбачає ґрунтовне знання не тільки самих методів, а й умов (обмежень) їх застосовності. Необхідно, щоб дослідник-педагог мав уявлення про такі поняття, як вибірка та генеральна сукупність, репрезентативність вибірки, похибки вибіркового статистичного дослідження, вибірккові статистичні оцінки, їх види та властивості, статистичні гіпотези, похибки першого та другого роду, статистичні критерії, потужність критерію, параметричні та непараметричні критерії, їх властивості та області застосовності. Ці знання та вміння застосувати їх на практиці є необхідною складовою загальної дослідницької культури.

Список використаної літератури

1. Ядренко М.Й., Гихман И.И., Скороход А.В. Теория вероятностей и математическая статистика – К.: Вища школа, 1988. - 440с.
2. Сидоренко Е.В. Математические методы обработки в психологии. СПб.: ООО «Речь», 2000. — 350с.
3. Гончаренко Я.В. Теорія ймовірностей і математична статистика: //Практикум — К.: НПУ імені Драгоманова, 2011. — 146с.
4. Крамер Г. Математические методы статистики. — М.: Мир, 1975. — 648с.
5. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. - М.: ЮНИТИ-ДАНА, 2004. - 573 с.
6. Руденко В.М. Математична статистика. Навч. посіб. – К.: Центр учбової літератури, 2012. – 304 с.
7. Турчин В.Н. Теория вероятностей и математическая статистика. — Д.: Изд-во Днепропетровского нац. ун-та, 2008 — 656с.