INTERNATIONAL SCIENTIFIC JOURNAL OF
Contemporary Pedagogical and Philological Research
2026 / No. 1(1)
SCIENCE CONNECTS PEOPLE

УДК 519.248.6

СТАТИСТИЧЕСКИЙ АНАЛИЗ ТЕКУЩИХ ПЕДАГОГИЧЕСКИХ ЭКСПРИМЕНТОВ ПРИ ПОМОЩИ КОНТРОЛЬНЫХ КАРТ И ИХ СРАВНЕНИЕ С ТРАДИЦИОННЫМ МЕТОДОМ

Юлдашев Х.Д.

Андижанский государственный университет

Аннотация. В статье проведено новая контрольная карта на основе критерия хи-квадрат Пирсона которая проверяет однородность двух независимых выборок в педагогических экспериментах. Метод контрольных карт позволяет анализировать, визуально отслеживать и статистически оценивать педагогические эксперименты не только на основе конечных результатов, но и на протяжении всего процесса их проведения. Для оценки эффективности метода контрольных карт за основу были взяты результаты экспериментов из диссертаций по педагогике. Каждое исследование оценивалось в сравнительном ключе с использованием как традиционных статистических критериев, так и метода контрольных карт. Оценка проводилась по шести параметрам: точность, чувствительность, уровень ошибок, учет динамики, простота интерпретации и возможность практического применения. В работе показана, что метод контрольных карт показал результаты, превышающие показатели традиционного метода критериев на 11,8 %, что в относительном выражении означает повышение эффективности примерно на 20,17%.

Ключевые слова: Критерий Пирсона, контрольные карты, границы контрольных карт, критерий Вилкоксона.

Введение.

В настоящее время в большинстве исследованиях при изучении различных процессов используют статистический инструмент, так называемый контрольные карты Шухарта (ККШ) [1,2,3,4]. Изучению проблем образования

можно найти, например, в следующих работах [3,5,6]. Поэтому в ряде ситуаций при анализе педагогических данных назрела необходимость перехода от классических методов к методу КК. В этой статье мы качестве примера рассмотрим задачу проверки однородности двух выборок: $x_1, x_2, \dots, x_m$ и $y_1, y_2, \dots, y_n$ (т. е. наборы из m и n действительных чисел) - требуется проверить, есть ли различие между выборками.

Качественный и количественный анализ педагогических явлений и процессов методами теории вероятностей и математической статистики уточняют выдвигаемых методов и гипотез. Обычно статистические выводы делаются после окончания экспериментов. Тогда мы имеем результаты типа “Off line control” [8,9,10,11,12,13,14].

На практике важным являются результаты типа “On line control” т.е. полученные статистические выводы используется в ходе эксперимента. Эти результаты одновременно дают проверить правильность проведенного эксперимента (в том числе метода) и своевременно исправить недостатки в планах эксперимента.

Обычно КК называют “голосом процесса”. Она является универсальным языком для разных специалистов использующие статистических выводов в своих исследованиях. Для каждой КК есть соответствующая статистическая гипотеза и критерий для проверки этой гипотезы.

Если исследователь начал действовать по планированной процедуре улучшающий учебный процесс (новые методы преподавания, измененное содержание обучения, правильное повторение учебного материала, некоторый новый способ запоминания, новый учебный план…), то требуется статистическое управление процесса. Мы считаем, что здесь использование метода КК является продуктивным.

SCIENCE CONNECTS PEOPLE

Цель и методы исследования.

Для статистического анализа текущих педагогических экспериментов методом КК сначала нужно найти характеристики используемых КК:

– контрольную величину, связанную с качеством учебного процесса;

– контрольные границы контрольной величины: LCL - нижняя контрольная граница; UCL - верхняя контрольная граница.

Для построения КК имеется два подхода:

Вероятностный подход: здесь КК выполняют задачу последовательной проверки определенных статистических гипотез. Для нахождения контрольных границ требуются знание распределения контрольной величины (точной или предельной) и 95% -й доверительный интервал [2].

Операциональный подход: как в первом подходе КК выполняют задачу последовательной проверки определенных статистических гипотез. Для нахождения контрольных границ не требуются знание распределения контрольной величины, находят по правилу «шесть сигма» [2,4].

В данной статье мы при построении КК используем первый подход. А в качестве используемых критериев берем критерий однородности двух выборок критерий хи-квадрат К. Пирсона [7] для проверки выдвигаемых гипотез.

Здесь коротко изложим традиционный метод проверки однородности и при помощи этих критериев для исследования педагогических процессов. Затем на основе их мы находим подходящие характеристики используемых новых КК.

Критерий однородности хи-квадрат применяют к анализу данных любой природы и при этом можно анализировать одновременно любое конечное число выборок.

Предположим осуществлено две серии независимых наблюдений из контрольной и экспериментальной группы состоящих n_1 и n_2 наблюдений соответственно $\underline{X} = (x_1, x_2, \dots, x_{n_1})$ и $\underline{Y} = (y_1, y_2, \dots, y_{n_2})$. Причем измеряемое

SCIENCE CONNECTS PEOPLE

свойство имеет $N > 2$ категорий. Далее в соответствии со специально разработанными критериями определим частоты попадания элементов выборок в одну из N категорий. Пусть $\underline{v}_1 = (v_{11}, v_{12}, \dots, v_{1N})$, $\underline{v}_2 = (v_{21}, v_{22}, \dots, v_{2N})$ частоты исходов, а $\underline{p}_1 = (p_{11}, p_{12}, \dots, p_{1N})$, $\underline{p}_2 = (p_{21}, p_{22}, \dots, p_{2N})$ - вероятности контрольной и экспериментальной группы. Тогда гипотеза однородности означающее утверждение, что вероятности исходов не менялись в контрольной и экспериментальной группе гласить:

$$H_0: \underline{p}_1 = \underline{p}_2 = \underline{p} \quad (1)$$

где $p = (p_1, p_2, \dots, p_N)$, $(p_1 + p_2 + \cdots + p_N = 1)$ - неизвестный вектор вероятностей.

При этом альтернативная гипотеза имеет вид

$$H_1: \underline{p}_1 \neq \underline{p}_2 \neq \underline{p} \quad (2)$$

хотя бы для одной категории.

Критерий проверки этих гипотез на уровне α даётся следующим соотношением

$$\hat{X}^2_{n_1,n_2} = \frac{1}{\omega(1-\omega)}\left(\sum_{j=1}^{N} \omega_j v_{1j} - n_1\omega\right),$$

где $\omega_j = \frac{v_{1j}}{(v_{1j}+v_{2j})}$, $\omega = \frac{n_1}{(n_1+n_2)}$.

При больших значениях n_1 и n_2 правило проверки однородности хи – квадрат гласить: H_0 отвергается и принимается H_1 если $\hat{X}^2_{n_1,n_2} > \chi^2_{1-\alpha;\,N-1}$, здесь $\chi^2_{1-\alpha;\,N-1}$ – критическое значение критерия хи-квадрат со степенью свободы $N - 1$.

При статистическом анализе результатов педагогического эксперимента широко применяются традиционные статистические критерии. Однако такой подход зачастую ограничивается лишь сравнением итоговых показателей. При

этом могут остаться без должного внимания тонкие изменения, происходящие в ходе эксперимента, динамические сдвиги во времени и постепенное развитие педагогического воздействия. В связи с этим метод КК позволяет анализировать, визуально отслеживать и статистически оценивать педагогические эксперименты не только на основе конечных результатов, но и на протяжении всего процесса их проведения.

Для оценки эффективности метода КК за основу были взяты результаты экспериментов из диссертаций по педагогике. Каждое исследование оценивалось в сравнительном ключе с использованием как традиционных статистических критериев, так и метода контрольных карт. Оценка проводилась по шести параметрам: точность, чувствительность, уровень ошибок, учет динамики, простота интерпретации и возможность практического применения.

Первый критерий - точность, отражающая степень соответствия результата реальной педагогической ситуации; высокий балл присваивается, если результаты подтверждаются при пересчете.

Второй критерий - чувствительность, то есть способность метода фиксировать незначительные изменения; в то время как традиционные критерии обычно выявляют лишь существенные различия, КК позволяют обнаружить даже тонкие сдвиги в педагогическом процессе.

Третий показатель - уровень ошибок; оценка метода повышается по мере снижения доли неверных выводов.

Четвертый критерий - степень учета динамики, предполагающая анализ изменений во времени.

Пятый критерий - простота интерпретации, то есть возможность разъяснить результаты педагогам с помощью понятных графиков и комментариев.

Шестой критерий - практическая применимость, характеризующая возможность внедрения метода в реальный образовательный процесс.

SCIENCE CONNECTS PEOPLE

Общая эффективность определяется по следующей формуле:

$$Q = \frac{1}{6}[A + S + (100 - E) + D + I + P]$$

Здесь A - точность; S - чувствительность; $100 - E$ - величина, обратная уровню ошибок; D - учет динамики; I - простота интерпретации; P - практическая применимость.

Согласно приведенным критериям оценки, общие показатели, рассчитанные на основе традиционных статистических критериев, как правило, находятся в диапазоне 50-80%, тогда как показатели, полученные методом контрольных карт, - в диапазоне 75–95%. Преимущество метода контрольных карт особенно заметно в отношении чувствительности, учета динамики, простоты интерпретации и практической применимости.

Результаты.

КК хи - квадрат (КК-χ^2) построим в случае $N = 4$. При этом X и Y имеют четыре категории: 5 ($j = 1$), 4 ($j = 2$), 3 ($j = 3$) и 2 ($j = 4$) как оценки изучаемых объектов. Тогда гипотезы (1) и (2) принимают следующий вид

$H_0: p_{11} = p_{21} = p_1;\ p_{12} = p_{22} = p_2;\ p_{13} = p_{23} = p_4;\ p_{14} = p_{24} = p_4$

$H_1: p_{11} \neq p_{21} \neq p_1;\ p_{12} \neq p_{22} \neq p_2;\ p_{13} \neq p_{23} \neq p_4;\ p_{14} \neq p_{24} \neq p_4$

хотя бы для одной категории.

При этом

$$\hat{X}^2_{n_1,\,n_2} = \frac{1}{\omega(1-\omega)}\left(\sum_{j=1}^{4} \omega_j \upsilon_{1j} - n_1\omega\right).$$

Верхнюю контрольную границу искомой КК находим из того, что при больших значениях n_1 и n_2 имеет место соотношение

$\hat{X}^2_{n_1,n_2} \to \chi^2_{1-\alpha;\,3}.$ (3)

Это мы сформулируем в виде следующего утверждения.

SCIENCE CONNECTS PEOPLE

Теорема. Пусть $\alpha = 0{,}01$ и верно гипотеза H_0, то для нахождения UCL_{χ^2} имеет место следующая формула:

$$UCL_{\chi^2} = n_1\omega + 0{,}352 \cdot \omega(1-\omega).$$

Следствие. Если $g(\underline{X},\underline{Y}) > UCL_{\chi^2}$, тогда, верно, альтернативная гипотеза H_1.

Доказательства теоремы приведем при помощи построения 95% доверительного интервала для контрольной величины.

Из предельного соотношения (3) мы имеем

$$P\left(\hat{X}^2_{n_1,n_2} < \chi^2_{0{,}95;\,3}\right) = 0{,}95.$$

Это эквивалентно соотношению

$$P\left(\sum_{j=1}^{4} \omega_j v_{1j} < n\omega + \omega(1-\omega)\chi^2_{0{,}95;3}\right) = 0{,}95.$$

Далее из правила выбора контрольной величины и UCL_{χ^2} имеем

$$g(\underline{X},\underline{Y}) = \sum_{j=1}^{4} \omega_j v_{1j}$$

$$UCL_{\chi^2} = n_1\omega + 0{,}352 \cdot \omega(1-\omega).$$

Теорема доказана.

Для иллюстрации теоремы мы рассмотрим эксперимент, направленный на выявление четырёх свойств изучения английского языка RWSL (Reading; Writing; Speaking; Listening) по-старому и новому учебнику для технических специальностей. Образуя контрольную и экспериментальную группу из студентов, которые учились по традиционной и по новой методике, собраны выборки отражающие знания студентов и по ней студентов распределено на четыре категории. Проверка знаний студентов в эксперименте проводились четыре раза в отдельности по свойствам RWSL. Построенная КК-χ^2на основе теоремы и следствия выглядит следующим образом.

SCIENCE CONNECTS PEOPLE

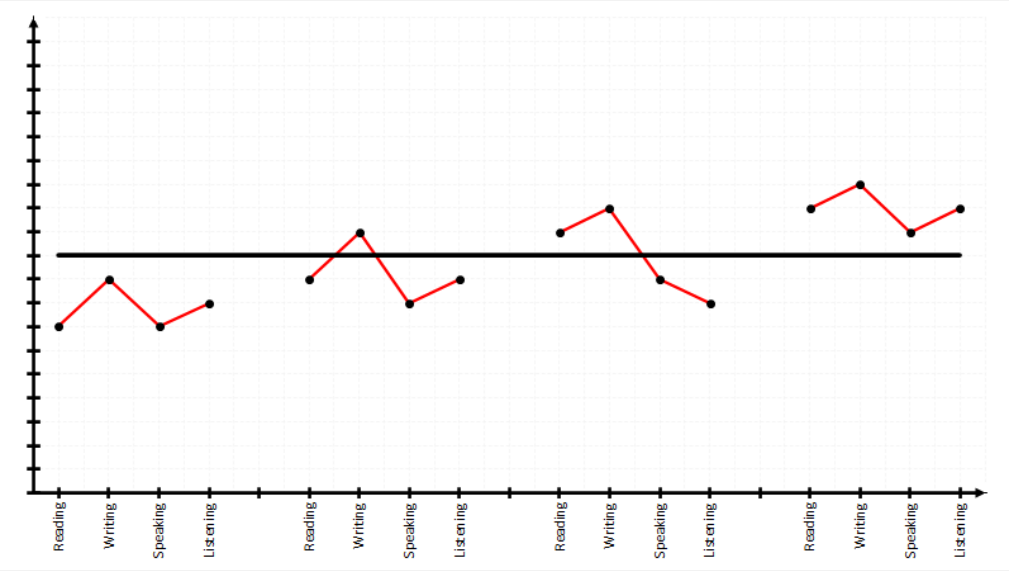

Рис.1.Проверка однородности при помощи КК-χ^2.

Из диаграммы КК мы видим, что до полного восстановления всех свойств RWSL экспериментатору пришлось много работать по двум свойствам S и L

В исследовании сравнивались результаты, полученные с использованием Критерий хи-квадрат и метода КК, на материале 10 PhD диссертаций. Полученные данные представлены в таблице 2. Первый столбик означает номер PhD, второй-балл ($Q_{k,i}$) поставленный за традиционное решение задачи, третий столбик балл ($Q_{nk,i}$) поставленный за решение методом КК, четвертый-разницу оценок ($d_i = Q_{nk,i} - Q_{k,i}$) и пятый столбик знак разности d_i .

Таблица 2. Результаты оценок PhD по шести критериям.

				знак величины
1	56	78	22	+
2	60	75	15	+
3	58	70	12	+
4	62	80	18	+
5	55	65	10	+
6	59	68	9	+
7	61	72	11	+
8	57	66	9	+
9	63	62	-1	-
10	54	67	13	+

SCIENCE CONNECTS PEOPLE

Из таблицы видно, что в 9 из 10 случаев метод КК показал лучшие результаты по сравнению с традиционными критерием. Лишь в одном случае разность оказалась отрицательной, что указывает на необходимость более углубленного статистического анализа превосходства данных методов.

Для определения статистической значимости этих различий используем критерий знаковых рангов Вилкоксона [15] для парных выборок. Этот критерий позволяет установить, обусловлена ли разница в результатах (полученных при оценке одних и тех же испытуемых двумя различными методами) случайностью или же она статистически значима. Здесь формулируются следующие гипотезы:

H_0: $Q_k = Q_{nk}$ (эффективность обоих методов одинакова);

H_1: $Q_{nk} > Q_k$ (метод контрольных карт более эффективен).

С использованием критерия знаковых рангов Вилкоксона сначала вычисляется разность для каждой пары: $d_i = Q_{nk,i} - Q_{k,i}$.

Затем берутся абсолютные значения разностей, располагаются в порядке возрастания, и им присваиваются ранги. Определяются суммы рангов, соответствующие положительным и отрицательным разностям:

$T^+ = 54, T^- = 1.$

Статистика критерия определяется следующим образом:

$T = min(T^+, T^-) = 1.$

При $n = 10$ и уровне значимости $\alpha = 0{,}05$ значение $T = 1$ является очень низким по сравнению с критическим значением критерия Вилкоксона. Следовательно, нулевая гипотеза (H_0) отвергается, а альтернативная гипотеза (H_1) принимается. Таким образом, статистически обосновано, что метод контрольных карт эффективнее традиционных статистических критериев.

Средние показатели эффективности также подтверждают этот вывод. Среднее значение, рассчитанное по традиционным критериям, составляет:

$$Q_k = \frac{56 + 60 + 58 + 62 + 55 + 59 + 61 + 57 + 63 + 54}{10} = 58{,}5\%$$

SCIENCE CONNECTS PEOPLE

Среднее значение, полученное методом контрольных карт, составляет:

$$Q_k = \frac{78 + 75 + 70 + 80 + 65 + 68 + 72 + 66 + 62 + 67}{10} = 70{,}3\%$$

Это представлено на следующей диаграмме.

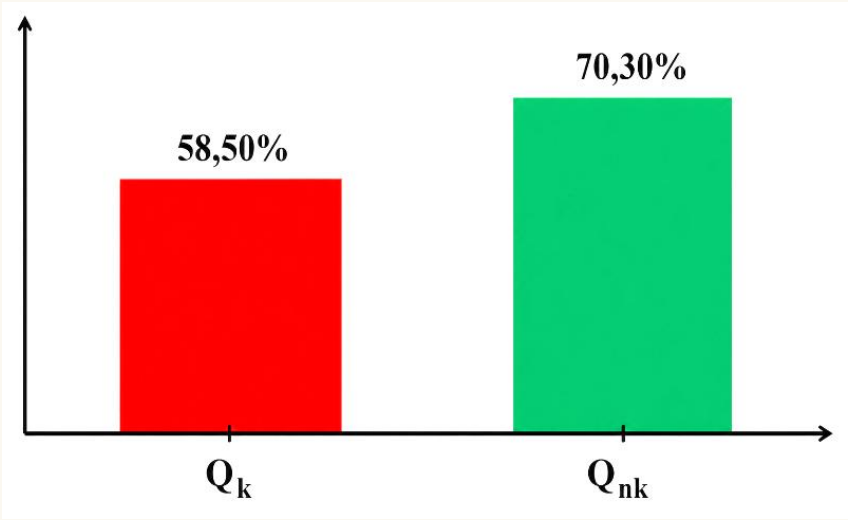

Рис 2. Сравнительный анализ эффективности метода КК и традиционного метода.

Таким образом, метод КК показал результаты, превышающие показатели традиционного метода критериев на 11,8%, что в относительном выражении означает повышение эффективности примерно на 20,17%. Относительное преимущество рассчитывается следующим образом:

$$\eta = \frac{70{,}3 - 58{,}5}{58{,}5} \cdot 100\% \approx 20{,}17\%.$$

Это свидетельствует о том, что метод КК обладает существенным преимуществом при анализе результатов педагогических экспериментов.

ОБСУЖДЕНИЕ.

Преимущество КК заключается в том, что они позволяют осуществлять поэтапный мониторинг экспериментального процесса. Традиционные статистические критерии, как правило, ограничиваются сравнением результатов, полученных в начале и в конце эксперимента. КК, напротив, позволяют выявить, на каком этапе происходят изменения и когда педагогическое воздействие

SCIENCE CONNECTS PEOPLE

усиливается или ослабевает. Это дает исследователю возможность своевременно корректировать методику эксперимента, изменять механизм воздействия или применять дополнительные педагогические меры.

Важность КК также обусловлена тем, что они представляют результаты в графическом виде. Благодаря этому педагогам — даже тем, кто не обладает глубокими знаниями в области математической статистики, — легче понять результаты анализа. Графическое представление наглядно демонстрирует такие тенденции в результатах эксперимента, как стабильность, отклонения, рост или спад. Таким образом, КК важны не только как инструмент статистического анализа, но и как средство педагогической диагностики и мониторинга.

КК кроме последовательной проверки статистических гипотез выполняют и некоторые дополнительные функции:

КК - это удобный метод для контролирования педагогических экспериментов. Их используют прямо на месте, где требуется проверка нестабильности и определяет простые (случайные) и особые причины, приводящие процесс в неуправляемое состояние;

КК помогают для повышения качество учебного процесса;

КК являются общим языком для разных специалистов;

Если на основе диаграмм сделано возможные причины нестабильности, то используя их педагогический процесс можно привести в стабильное состояние.

Метод КК является универсальным в том смысле, что статистические выводы даётся в диаграммах. Кроме того, эти диаграммы сохраняются в виде документа о проведенных экспериментах и анализ результатов нескольких таких экспериментов могут давать прогнозирующих закономерностей.

SCIENCE CONNECTS PEOPLE

Литература

1. Адлер Ю.П., Максимова О.В., Шпер В.А. Контрольные карты Шухарта в России и за рубежом: краткий обзор современного состояния (статистические аспекты). Стандарты и качества, 8. с 1-40. 2011.

2. Х.-Й.Миттаг, Х.Ринне. Статистические методы обеспечения качества. М.: Машиностроение, 1995. - 616 с.

3. Солонин С. И. Метод контрольных карт. Екатеринбург. УрФУ. 2014.

4. Д.Уилер, Д.Чамберс. Статистическое управление процессами. Оптимизация бизнеса с использованием контрольных карт Шухарта. Москва – 2009.

5. Turner R. (1998). The Red Bead Experiment for Educators. - Quality Progress, June, pp. 69-74.

6. Gitlow H.S. (2001) Viewing statistics from a quality control perspective. – International Journal of Quality & Reliability Management, vol.18, N0.2, pp.169-179.

7. Ивченко Г.И., Медведев Ю.И. Введение в математическую статистику. М.: Издательство ЛКИ, 2010. - 600 с.

8. Орлов А.И. О проверке однородности двух независимых выборок - Журнал «Заводская лаборатория». 2003. Т.69. No.1. С.55-60.

9. М.И.Грабарь, К.А.Краснянская. Применение математической статистики в педагогических исследованиях. Непараметрические методы. М., “Педагогика”, 1977.

10. Д.А.Новиков. Статистические методы в педагогических исследованиях (типовые случаи). М.: МЗ-Пресс, 2004. - 67 с.

11. Кондрашова Е.В. Применение методов математической статистики при проведении педагогического эксперимента. Современные наукоемкие технологии. – 2021. – № 2 – С. 162-169.

12. Dalibor Gonda, Gabriela Pavlovičová, Viliam Duriš and Anna Tirpáková. Implementation of Pedagogical Research into Statistical Courses to Develop Students' Statistical Literacy. Mathematics 2022, 10, 1793. https://doi.org/10.3390/math10111793.

13. I. Nikolaeva, Yu. Sleptsov, I. Gogoleva, A. Mirzagitova, N. Bystrova, R. Tsarev. Statistical hypothesis testing as an instrument of pedagogical experiment. AIP Conference Proceedings. Volume 2647, Issue 1. https://doi.org/10.1063/5.0104059.

14. V. Petruk, Yu. Rudenko, A. Yurchenko, I. Kharchenko, S. Kharchenko, O. Semenikhina. Analysis of the Results of the Pedagogical Experiment on the Integrated Analysis of the Average and Dispersions. International Journal of Modern Education and Computer Science (IJMECS). Vol. 14, No. 6, Dec. 2022.

15. Wilcoxon F. Individual comparisons by ranking methods // Biometrics Bulletin. - 1945. - Vol. 1, № 6. - P. 80-83.