



DOI: <https://doi.org/10.15688/mpcm.jvolsu.2021.3.3>

УДК 510:796.071.3

ББК 22.1

Дата поступления статьи: 14.06.2021

Дата принятия статьи: 20.08.2021

ОБ ИСПОЛЬЗОВАНИИ КРИТЕРИЯ ВИЛКОКСОНА В ПРАКТИКЕ ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЙ

Елена Владимировна Бондарева

Кандидат педагогических наук,
доцент кафедры математического анализа и теории функций,
Волгоградский государственный университет
ele-bondareva@yandex.ru, matf@volsu.ru
<https://orcid.org/0000-0001-7580-6760>
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

Елена Алексеевна Мазепа

Кандидат физико-математических наук,
доцент кафедры математического анализа и теории функций,
Волгоградский государственный университет
elena.mazepa@volsu.ru, matf@volsu.ru
<https://orcid.org/0000-0001-7603-4133>
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

Аннотация. В статье рассматривается вопрос об использовании непараметрических мер влияния в педагогическом исследовании. Выделен критерий Вилкоксона для сравнения показателей, полученных в результате измерения в связанной выборке. Указаны необходимые условия для применения критерия, а также преимущества его использования. Возможность использования критерия Вилкоксона проиллюстрирована на примере определения эффективности предложенной методики обучения с использованием онлайн-платформы Учи.ру в образовательной деятельности.

Ключевые слова: непараметрические меры влияния, связанная выборка, ранговая сумма, критерий Вилкоксона, критерий Z .

Введение

Любой педагогический эксперимент, проводимый в ходе педагогического исследования, подразумевает измерение различных параметров, факторов, признаков и т. п. Далее проводится качественный анализ результатов диагностики. Однако ни одно современное исследование не ограничивается качественными описаниями результатов изучаемого явления, поскольку они не могут отражать полной картины проводимого исследования. Большинство современных исследований свидетельствуют о том, что более достоверные выводы можно получить с помощью количественного анализа, который подразумевает математическую обработку с использованием статистических методов [1; 2]. В психолого-педагогических исследованиях часто возникает необходимость изучить влияние определенного фактора на исследуемый признак. Так, например, анализ научных работ в области физической культуры и спорта показал, что тренеру в процессе подготовки спортсменов необходимы знания и умения, связанные с обработкой большого объема числовых данных измерений. Работа с данными подразумевает выявление факторов и измерение степени их влияния на эффективность тренировочного процесса. Чаще всего для принятия обоснованного решения в спортивных исследованиях используют метод проверки статистических гипотез, корреляционный, регрессионный и дисперсионный анализы, а также метод принятия коллективных решений [3]. Если же речь идет о процессе обучения, фактором, влияющим на какой-либо признак, может выступать разработанная методика обучения, которая может привести к изменениям исследуемого признака. На основании этих изменений определяется статистическая достоверность (значимость) влияния данного фактора. Для анализа достоверности влияния того или иного фактора, а также для сравнения между собой силы влияния разных факторов используются как параметрические (дисперсионный анализ), так и непараметрические меры влияния (критерий знаков, критерий Вилкоксона) [4]. В данной статье рассматриваются непараметрические меры влияния в условиях проведения педагогического эксперимента, а именно, критерий Вилкоксона.

1. Особенности применения критерия Вилкоксона в педагогико-психологических исследованиях

Критерий Вилкоксона используется для сравнения показателей, которые измеряются в двух различных условиях на одной и той же выборке испытуемых, то есть речь идет о связанной выборке. С помощью этого критерия можно диагностировать как направленность изменений, так и их выраженность. Помимо этого, преимуществом использования данного критерия является тот факт, что для его применения не требуется выполнение какого-то определенного закона распределения. Являясь ранговым критерием, критерий Вилкоксона позволяет присваивать значения признака как положительные, так и отрицательные ранги. К условиям, необходимым для использования критерия Вилкоксона, относятся следующие: измерение можно проводить в любой шкале, кроме номинальной (чаще используется шкала порядка или шкала интервалов); объем выборки должен составлять от 5 до 50 человек. Чем больше объем выборки, тем более выражено проявляют себя отклонения, связанные с воздействием эксперимента [5]. Измерения проводятся до и после проведения эксперимента. Экспериментом выступает некоторое воздействие на процесс обучения (определенная методика). Критерий позволяет установить эффективность воздействия по статистическим данным. По исходным данным (измерения до и

после эксперимента) составляется выборка разности наблюдаемых показателей для каждого объекта группы. Причем значения в этой выборке могут быть как положительными (в случае уменьшения показателя), так и отрицательными (в случае увеличения). Если разность получается равной нулю, то ее далее не учитывают. Следующий шаг — это ранжирование по возрастанию полученных значений по абсолютной величине. Ранги расставляются в соответствии с порядковыми номерами наблюдаемых показателей, упорядоченных по возрастанию. При этом, если некоторые показатели имеют одинаковые значения, то всем им предписывается одинаковый ранг, равный среднему арифметическому их порядковых номеров. Это промежуточные ранги. Далее каждому рангу следует приписать знак «+» или «-» согласно знака соответствующей ему разности. Таким образом, получают так называемые *знаковые ранги*. Просуммировав знаковые ранги, получают случайную величину $W_{\text{набл}}$, которая называется критерием Вилкоксона. Остается найти $W_{\text{кр}}$ для рассматриваемой выборки по соответствующим таблицам и сравнить эти два полученных значения для получения вывода о подтверждении или опровержении гипотезы исследования.

Следует заметить, что в таблицах критических значений критерия Вилкоксона, как правило, указаны значения критерия для $n \leq 27$, большинство таблиц содержат значения до $n = 20$ включительно. Это связано с достаточно громоздкими вычислениями для нахождения критических значений при увеличении объема выборки. Однако отслеживается такая закономерность, что при увеличении n распределение случайной величины W стремится к нормальному с нулевым средним и вычисляемым стандартным отклонением по формуле (см., например, [5]):

$$\sigma_w = \sqrt{\frac{n(n+1)(2n+1)}{6}}.$$

Практика вычислений показывает, что уже при $n > 20$ выполняется приближенное соответствие $W \sim N(0; \sigma_w)$. Это означает, что уже при $n > 20$ табличные вычисления можно значительно упростить с помощью критерия Z вместо критерия W . Расчетная формула для критерия Z имеет вид:

$$Z = \frac{W}{\sigma_w}.$$

Однако статистика Вилкоксона W является дискретной величиной, в отличие от нормального, которое относят к непрерывным распределениям. Погрешность при переходе от дискретного распределения к непрерывному распределению Z сглаживается с помощью поправки Йейтса. Тогда используется следующая расчетная формула для Z -критерия (см. также [5]):

$$Z = \frac{|W| - 1/2}{\sigma_w}, \quad \text{где} \quad \sigma_w = \sqrt{\frac{n(n+1)(2n+1)}{6}}.$$

Затем вместо $W_{\text{набл}}$ вычисляется $Z_{\text{набл}}$, а критические точки при уровне значимости α находят в соответствующей таблице. Заметим, что для непрерывного критерия Z уровень значимости α можно выбирать любой. А при дискретном критерии W уровень значимости выбирают из набора фиксированных дискретных значений. В следующем разделе рассмотрим алгоритм приведенных выше вычислений на конкретном примере.

2. Педагогический эксперимент и его результаты

Педагогический эксперимент проводился на базе одной из школ города Волгограда в течение 2018–2019 учебного года. В диссертационном исследовании на присуждение степени магистра была выдвинута гипотеза — применение онлайн-платформы Учи.ру в образовательной деятельности положительно влияет на усвоение материала по математике. Эта гипотеза была принята за нулевую. В качестве экспериментальной группы был выбран 4-й класс в количестве 27 человек. Ученикам экспериментальной группы было рекомендовано в качестве домашней работы по математике заниматься на образовательной платформе Учи.ру, выполняя по 20 заданий по предмету [6]. Для данного эксперимента была создана система педагогических тестов для определения процента верно решенных задач по математике до (x_i) и после (x'_i) проведения эксперимента. Результаты входного и итогового тестирования представлены в таблице 1.

Таблица 1

i	1	2	3	4	5	6	7	8	9
x_i	58,33	91,67	75,00	100,00	100,00	83,33	83,33	100,00	75,00
x'_i	50,00	70,00	40,00	80,00	80,00	30,00	60,00	100,00	90,00

i	10	11	12	13	14	15	16	17	18
x_i	100,00	91,67	75,00	100,00	100,00	91,67	100,00	75,00	100,00
x'_i	100,00	90,00	50,00	100,00	100,00	70,00	90,00	60,00	80,00

i	19	20	21	22	23	24	25	26	27
x_i	100,00	100,00	100,00	100,00	91,67	75,00	100,00	75,00	100,00
x'_i	100,00	80,00	100,00	70,00	100,00	60,00	80,00	50,00	100,00

Используя эти данные, необходимо определить эффективность предложенной методики. Для удобства проведения расчетов по критерию Вилкоксона оформляем данные в виде таблицы 2.

Введем следующие обозначения: i — номер испытуемого по списку, x_i — процент правильно решенных задач теста до эксперимента, x'_i — процент правильно решенных задач теста после эксперимента, $x'_i - x_i$ — изменение показателя.

Прокомментируем полученные расчеты. Первый столбец содержит порядковый номер участника эксперимента — ученика. Во втором и третьем столбцах записаны исходные данные наблюдений — результаты входного и итогового тестирования (до и после эксперимента). Данные остальных столбцов заполняются на их основе с помощью описанных выше действий. Значения разностей уменьшения (изменения) показателя, подсчитанные в четвертом столбце, — есть выборка генеральной совокупности случайной величины X . Поскольку закон распределения X не известен в условиях данного эксперимента, проводится расчет случайной величины W — суммы знаковых рангов, полученных путем сравнения различных значений случайной величины X с учетом знаков этих значений. Наблюдаемое значение случайной величины W представляет собой сумму значений последнего столбца и равно, согласно расчетам: $W_{\text{набл}} = -194$.

Таблица 2

i	x_i	x'_i	$x'_i - x_i$	$ x'_i - x_i $	Ранг	Знаковый ранг
1	58,33	50,00	-8,33	8,33	3,0	-3,0
2	91,67	70,00	-21,67	21,67	13,5	-13,5
3	75,00	40,00	-35,00	35,00	19,0	-19,0
4	100,00	80,00	-20,00	20,00	10,0	-10,0
5	100,00	80,00	-20,00	20,00	10,0	-10,0
6	83,33	30,00	-53,33	53,33	20,0	-20,0
7	83,33	60,00	-23,33	23,33	15,0	-15,0
8	100,00	100,00	0	0	0	0
9	75,00	90,00	15,00	15,00	6,0	6,0
10	100,00	100,00	0	0	0	0
11	91,67	90,00	-1,67	1,67	1,0	-1,0
12	75,00	50,00	-25,00	25,00	16,5	-16,5
13	100,00	100,00	0	0	0	0
14	100,00	100,00	0	0	0	0
15	91,67	70,00	-21,67	21,67	13,5	-13,5
16	100,00	90,00	-10,00	10,00	4,0	-4,0
17	75,00	60,00	-15,00	15,00	6,0	-6,0
18	100,00	80,00	-20,00	20,00	10,0	-10,0
19	100,00	100,00	0	0	0	0
20	100,00	80,00	-20,00	20,00	10,0	-10,0
21	100,00	100,00	0	0	0	0
22	100,00	70,00	-30,00	30,00	18,0	-18,0
23	91,67	100,00	8,33	8,33	2,0	2,0
24	75,00	60,00	-15,00	15,00	6,0	-6,0
25	100,00	80,00	-20,00	20,00	10,0	-10,0
26	75,00	50,00	-25,00	25,00	16,5	-16,5
27	100,00	100,00	0	0	0	0

В таблице критических значений для $n = 20$ (нулевые результаты не учитываются) при уровне значимости 0,02 находим $W_{кр} = 124$. Так как $W_{набл} > W_{кр}$, то результаты неоднородны, но поскольку $W_{набл} < 0$, то увеличения процента решенных задач не произошло. Таким образом, нулевая гипотеза отвергается. Можно сделать вывод о неэффективности использования онлайн-платформы Учи.ру для усвоения материала по математике для младших школьников. Полученные выводы были использованы далее в рамках проводимого педагогического исследования.

СПИСОК ЛИТЕРАТУРЫ

1. Бондарева, Е. В. Дисперсионный анализ в психолого-педагогических исследованиях / Е. В. Бондарева, Н. В. Стеценко // The First European Conference on Physics and Mathematics. Proceedings of the Conference. — Vienna : East West Association for Advanced Studies and Higher Education GmbH, 2015. — С. 12–17.
2. Бондарева, Е. В. Об особенностях подготовки бакалавров и магистров в институте математики и информационных технологий / Е. В. Бондарева, И. А. Романова, Н. В. Сте-

ценко // Вестник Волгоградского государственного университета. Серия 6, Университетское образование. — 2015. — Т. 16, № 1. — С. 37–42.

3. Бондарева, Е. В. О возможности использования метода коллективного принятия решений в спортивных исследованиях / Е. В. Бондарева, Н. В. Стеценко // Вестник Волгоградского государственного университета. Серия 1, Математика. Физика. — 2016. — Т. 33, № 2. — С. 6–12.

4. Лупандин, В. И. Математические методы в психологии / В. И. Лупандин. — Екатеринбург : Изд-во Урал. ун-та, 2009. — 196 с.

5. Медик, В. А. Математическая статистика в медицине / В. А. Медик, М. С. Токмачев. — М. : Финансы и статистика, 2007. — 800 с.

6. Наволокина, И. В. Информационные технологии в обучении математике (на примере образовательной платформы Учи.ру): магистерская диссертация / И. В. Наволокина. — Волгоград : Изд-во ВолГУ, 2019. — 54 с.

REFERENCES

1. Bondareva E.V., Stetsenko N.V. Dispersionnyy analiz v psikhologo-pedagogicheskikh issledovaniyakh [Analysis of Variance in Psychological Pedagogical Research]. *The First European Conference on Physics and Mathematics. Proceedings of the Conference*. Vienna, East West Association for Advanced Studies and Higher Education GmbH, 2015, pp. 12-17.

2. Bondareva E.V., Romanova I.A., Stetsenko N.V. Ob osobennostyakh podgotovki bakalavrov i magistrrov v institute matematiki i informatsionnykh tekhnologiy [On the Peculiarities of Training Bachelors and Masters at the Institute of Mathematics and Information Technologies]. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 6, Universitetskoe obrazovanie*, 2015, vol. 16, no. 1, pp. 37-42.

3. Bondareva E.V., Stetsenko N.V. O vozmozhnosti ispolzovaniya metoda kollektivnogo prinyatiya resheniy v sportivnykh issledovaniyakh [On the Possibility of Using the Method of Collective Decision-Making in Sports Research]. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 1, Matematika. Fizika* [The Science Journal of Volgograd State University. Mathematics. Physics], 2016, vol. 33, no. 2, pp. 6-12.

4. Lupandin V.I. *Matematicheskie metody v psikhologii* [Mathematical Methods in Psychology]. Ekaterinburg, Izd-vo Ural. Un-ta, 2009. 196 p.

5. Medik V.A., Tokmachev M.S. *Matematicheskaya statistika v meditsine* [Mathematical Statistics in Medicine]. Moscow, Finansy i statistika Publ., 2007. 800 p.

6. Navolokina I.V. *Informatsionnye tekhnologii v obuchenii matematike (na primere obrazovatelnoy platformy Uchi.ru): masterskaya dissertatsiya* [Information Technologies in Teaching Mathematics (On the Example of the Educational Platform Uchi.ru). Master's Thesis]. Volgograd, Izd-vo VolSU, 2019. 54 p.

ON THE USE OF THE WILCOXON CRITERION IN THE PRACTICE OF PEDAGOGICAL RESEARCH

Elena V. Bondareva

Candidate of Pedagogical Sciences, Associate Professor,
Department of Mathematical Analysis and Function Theory,
Volgograd State University
ele-bondareva@yandex.ru, matf@volsu.ru
<https://orcid.org/0000-0001-7580-6760>
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

Elena A. Mazepa

Candidate of Physical and Mathematical Sciences, Associate Professor,
Department of Mathematical Analysis and Function Theory,
Volgograd State University
elena.mazepa@volsu.ru, matf@volsu.ru
<https://orcid.org/0000-0001-7603-4133>
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

Abstract. The article discusses the use of nonparametric measures of influence in pedagogical research. The Wilcoxon criterion for comparing indicators obtained as a result of measurements in a connected sample is highlighted. The necessary conditions for the application of the criterion, as well as the advantages of its use are indicated. The possibility of using the Wilcoxon criterion is illustrated by the example of determining the effectiveness of the proposed teaching methodology using the Uchi.ru online platform in educational activities. Most modern studies in the field of pedagogy and psychology, along with qualitative descriptions of the results of the research, actively use quantitative analysis. Last analysis includes processing of the data using methods of mathematical statistics. In order to show the effectiveness of the developed methodology for the learning process, it should be considered as a factor influencing changes in the studied trait. Further study of these changes allows the researcher to determine the statistical reliability (significance) of the influence of the factor under consideration. The subsequent analysis of the reliability of the influence of a particular factor, the comparison of the measure of influence of various factors is carried out using parametric (variance analysis) and nonparametric measures of influence (the sign criterion, the Wilcoxon criterion). This article discusses the use of nonparametric measures of influence in pedagogical research. The Wilcoxon criterion is selected for comparing the indicators obtained as a result of measurement in a connected sample and allowing to diagnose not only the direction of changes, but also their severity. The necessary conditions for the application of the criterion, as well as the advantages of its use, are indicated. The possibility of using the Wilcoxon criterion is illustrated by the example of determining the effectiveness of the proposed teaching methodology using the Uchi.ru online platform in educational activities. Conclusions are made about its inefficiency for mastering the material in mathematics for younger schoolchildren.

Key words: nonparametric measures of influence, connected sample, rank sum, Wilcoxon test, Z -test.