ПРИМЕНЕНИЕ НЕКОТОРЫХ СТАТИСТИЧЕСКИХ МЕТОДОВ В ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ[2]

Андрей Захарович ХАРИТОН, профессор, Тираспольский Государственный университет, г. Кишинёв, stodolna@yandex.ru

Анна Валерьевна МЕЛЬНИЧУК, доцент кафедры Педагогики и современных образовательных технологий ПГУ, anytik_tiras@mail.ru

Инна Валерьевна КОЛОКОЛОВА, доцент кафедры теории и методики физического воспитания и спортивной тренировки, Педагогического института ФКиС, Московского городского педагогического университета

Аннотация. В статье изложены некоторые рекомендации относительно применения в педагогических исследованиях наиболее подходящих для этих целей статистических методов. Рассматриваютсяконкретныепримерыприменениястатистическихметодов.

Abstract. In this article some methodological instructions on utilization of statistical methods for scientific research in pedagogy are exposed. The publication will be useful to students, graduate students, PhD students, teachers of high school and gymnasium.

[2] Данная статья является продолжением серии публикаций на тему «Применение некоторых статистических методов в педагогических исследованиях». Первая часть опубликована в сб.: «Ştiinţa, educaţie, cultura» (Наука, образование, культура) // Матер. Междунар. научно-практ. конф. – 4 февраля, 2016 /Сост.: Л.В. Федотова и др. – Комрат: КГУ, 2016. – 618 с. – ISBN 978-9975-83-011-9.

В первой статье были рассмотрены применение в педагогических исследованиях статистических методов: среднее арифметическое, математическое ожидание, мода, медиана, дисперсия, критерий Крамера-Уэлча.

Продолжим рассмотрение наиболее значимых статистических методов, которые применяются в педагогических исследованиях.

7. Критерий Вилкоксона-Манна-Уитни.

Этот критерий требует сложные вычисления, однако в педагогических исследованиях его считают более эффективным. Критерий ВМУ оперирует с результатами парных сравнений двух выборок – экспериментальной и контрольной групп до и после эксперимента или экспериментальной – экспериментальной, контрольной – контрольной групп соответственно до и после окончания эксперимента.

Математическая формула применения критерия ВМУ:

$$W_{эмп.} = \frac{\left|\frac{N \cdot M}{2} - U\right|}{\sqrt{\frac{N \cdot M\left(N + M + 1\right)}{12}}}.$$

Алгоритм применения критерия:

1) Для каждого элемента первой выборки x_iопределим число a_i элементов второй выборкиy, которые превосходят его по значению ($y_j > x_i$), число b_iэлементов второй выборки, которое по своему значению равны ему ($y_j = x_i$).

2) Вычисляем сумму: $a_1 + a_2 + \cdots + a_м + \frac{1}{2}\left(b_1 + b_2 + \cdots + b_м\right) = \sum_{i=1}^{м} a_i + \frac{1}{2}\sum_{i=1}^{м} b_i.$

Полученная сумма называется эмпирическим значением критерия Манна-Уитни и обозначается буквой U.

3) Подставляем значения M, N, U в формуле $W_{эмп.}$

4) Сравниваем значения ВМУ с критическим значением $W_{0,05} = 1,96$.

Возможны два случая:

a) $W_{эмп.} \le 1,96$. В таком случае следует вывод: гипотеза о том, что экспериментальная выборка и контрольная совпадают принимается на уровне значимости 0,05 (в пределах 5%).

b) $W_{эмп.} > 1,96$. В таком случае следует вывод: достоверность различий характеристик сравниваемых выборок составляет 95%.

Пример: Применим критерий ВМУ к экспериментальной и контрольной группам до начала эксперимента (данные приведены в таблице 2 предыдущей статьи).

Таблица 3

Участник экспериментальной группы i	Число правильно решенных задач i-участником ЭГ до начала эксперимента x_i	Текущие значения формулы $a_i + \frac{1}{2} b_i$	Участник контрольной группы j	Число правильно решенных задач j-участником КГ до начала эксперимента y_i
1	6	15,5	1	8
2	8	8,5	2	8
3	4	23	3	9
4	6	15,5	4	7
5	5	20	5	6
6	8	8,5	6	6
7	7	12	7	6
8	9	4	8	5
9	3	24,5	9	10
10	8	8,5	10	8
11	6	15,5	11	9
12	5	20	12	5
13	9	4	13	6
14	4	23	14	7
15	7	12	15	8
16	9	4	16	9
17	5	20	17	10
18	6	15,5	18	5
19	5	20	19	5
20	4	23	20	4
21	6	15,5	21	4
22	10	1	22	3
–	–	–	23	6
–	–	–	24	9
–	–	–	25	8

U= 313,5

$$W_{\text{эмп.}} = \frac{\left|\frac{25 \cdot 22}{2} - 313,5\right|}{\sqrt{\frac{25 \cdot 22 \cdot (25 + 22 + 1)}{12}}} = \frac{|-38,5|}{46,90} = 0,018 \ .$$

$W_{\text{эмп.}}$ = 0,018 < 1,96, гипотеза о том, что экспериментальная выборка и контрольная выборка совпадают принимается на уровне значимости 0,05 (в пределах 5%).

8. Критерий χ^2 (хи квадрат).

Этот критерий используется для выяснения достоверности совпадений и различий измерений для контрольной и экспериментальной групп расположенных в виде порядковых шкал. Принята в педагогических исследованиях формула: $\chi^2_{эмп.} = N \cdot M \cdot \sum_{i=1}^{L} \frac{\left(\frac{n_i}{N} - \frac{m_i}{M}\right)^2}{n_i + m_i}$. В этой формуле L – количество групп на которое делятся числовые значения выборки.

Примечание 1: Критические значения $\chi^2_{0,05}$ приведены в таблице 4. Для чисел больше 9 следует обращаться к более полным таблицам.

Таблица 4.

$L-1$	1	2	3	4	5	6	7	8	9
$\chi^2_{0,05}$	3,84	5,99	7,82	9,49	11,07	12,59	14,07	15,52	16,92

Алгоритм применения критерия χ^2:

1) Вычисляем $\chi^2_{эмп.}$ по формуле указанной выше.

2) Сравниваем полученное значение $\chi^2_{эмп.}$ с критическим значением $\chi^2_{0,05}$ (таблица 4).

Возможны два случая:

a) $\chi^2_{эмп.} \leq \chi^2_{0,05}$, то характеристики сравниваемых выборок совпадают с уровнем значимости 0,05;

b) $\chi^2_{эмп.} > \chi^2_{0,05}$, то достоверность различий характеристик сравниваемых выборок составляет 95%.

Пример: Рассмотрим вычисление критерия $\chi^2_{эмп.}$ применительно к окончанию эксперимента, таблица 5, полученной на основании таблицы 3.

Таблица 5.

Уровень знаний	КГ количество участников до начало эксперимента	ЭГ количество участников до начало эксперимента	КГ количество участников после окончания эксперимента	ЭГ количество участников после окончания эксперимента
Низкий	7	8	6	6
Средний	12	10	13	10
Высокий	6	4	6	6

В нашем примере L = 3, M = 25 (количество участников контрольной группы), N = 22 (количество участников экспериментальной группы). Параметры экспериментальной группы после окончания эксперимента: n_1 = 6, n_2 = 10, n_3 = 6. Параметры контрольной группы после окончания эксперимента: m_1 = 6, m_2 = 13, m_3 = 6. Таким образом:

$$\chi^2_{эмп.} = 22 \cdot 25 \cdot \left[\frac{\left(\frac{6}{22} - \frac{6}{25}\right)^2}{6+6} + \frac{\left(\frac{10}{22} - \frac{13}{25}\right)^2}{10+13} + \frac{\left(\frac{6}{22} - \frac{6}{25}\right)^2}{6+6} \right] = 550\left(\frac{0,0011}{12} + \frac{0,0043}{23} + \frac{0,0011}{12}\right) =$$

$$= 550 \cdot (0,00009 + 0,000187 + 0,00009) = 550 \cdot 0,00036 = 0,20 \ .$$

Так как $\chi^2_{эмп.}$ = 0,20, а 0,20 < 5,99 (значение для L – 1 = 3 – 1 = 2 из таблицы 4), следует вывод: характеристики выборок экспериментальной и контрольной групп по окончании эксперимента совпадают с уровнем значимости 0,05 (5%).

Примечание 2: Очень часто возникает необходимость учесть всевозможные отношения между экспериментальной и контрольной групп.

При помощи критерия χ^2 в таком случае заполняется таблица 6.

Таблица 6.

	КГ до эксперимента	ЭГ до эксперимента	КГ после эксперимента	ЭГ после эксперимента
КГ до эксперимента	0	0,46	0,10	0,7
ЭГ до эксперимента	0,46	0	2,64	0,78
КГ после эксперимента	0,10	2,64	0	0,20
ЭГ после эксперимента	0,7	0,78	0,20	0

$\chi^2_{эмп.}$ для клеток (1; 2) и (2; 1) – контрольная группа до начало эксперимента и экспериментальной группы до начало эксперимента:

$$22 \cdot 25 \cdot \left[\frac{\left(\frac{8}{22} - \frac{7}{25}\right)^2}{8+7} + \frac{\left(\frac{10}{22} - \frac{12}{25}\right)^2}{10+12} + \frac{\left(\frac{4}{22} - \frac{6}{25}\right)^2}{4+6} \right] = 550(0,00043 + 0,00004 + 0,00036) = 0,46 \ .$$

$\chi^2_{эмп.}$ для клеток (1; 3) и (3; 1) – контрольная группа до начало эксперимента и контрольная группа после окончания эксперимента:

$$22\cdot 25\cdot\left[\frac{\left(\frac{6}{25}-\frac{7}{25}\right)^2}{6+7}+\frac{\left(\frac{13}{25}-\frac{12}{25}\right)^2}{13+12}+\frac{\left(\frac{6}{25}-\frac{6}{25}\right)^2}{6+6}\right]=550\left(0,00012+0,000064+0\right)=0,10\ .$$

$\chi^2_{эмп.}$ для клеток (1; 4) и (4; 1) – контрольная группа до начало эксперимента и экспериментальная группа после окончания эксперимента:

$$22\cdot 25\cdot\left[\frac{\left(\frac{6}{22}-\frac{7}{25}\right)^2}{6+7}+\frac{\left(\frac{10}{22}-\frac{12}{25}\right)^2}{10+12}+\frac{\left(\frac{6}{22}-\frac{6}{25}\right)^2}{6+6}\right]=550\left(0,000008+0,00004+0,00008\right)=0,7\ .$$

$\chi^2_{эмп.}$ для клеток (2; 3) и (3; 2) – контрольная группа после окончания эксперимента и экспериментальная группа до эксперимента:

$$22\cdot 25\cdot\left[\frac{\left(\frac{6}{25}-\frac{8}{22}\right)^2}{6+8}+\frac{\left(\frac{13}{25}-\frac{10}{22}\right)^2}{13+10}+\frac{\left(\frac{6}{25}-\frac{4}{22}\right)^2}{6+4}\right]=550\left(0,001+0,0002+0,0036\right)=2,64\ .$$

$\chi^2_{эмп.}$ для клеток (2; 4) и (4; 2) – экспериментальная группа до начало эксперимента и экспериментальная группапосле окончания эксперимента:

$$22\cdot 25\cdot\left[\frac{\left(\frac{6}{22}-\frac{8}{22}\right)^2}{8+7}+\frac{\left(\frac{10}{22}-\frac{10}{22}\right)^2}{10+10}+\frac{\left(\frac{6}{22}-\frac{4}{22}\right)^2}{6+4}\right]=550\left(0,00059+0,00082\right)=0,78\ .$$

В таблице 6 все значения $\chi^2_{эмп.}$ меньше значения $\chi^2_{0,05}$ = 5,99 для $L - 1 = 2$. Значит: все значения ЭГ и КГ попарно до и после эксперимента совпадают с уровнем значимости 0,05.

9. Критерий Фишера.

В математической статистике существуют несколько критериев Фишера. Между прочим, такое явление наблюдается и относительно многих других статистических критериев. Однако в педагогических исследованиях чаще всего используется критерий Фишера, который получил практическое название угловое преобразование. Указанный коэффициент Фишера базируется на понятие дихотомическая шкала.

Дихотомическая шкала – порядковая шкала с всего двумя различными упорядоченными баллами – «высокий», «низкий» («справился с заданием», «не справился с заданием»). Характеристикой группы будет число ее членов (или доля, процент от общего числа), получивших, например, максимальный балл.

Пусть для экспериментальной группы числовая характеристика является (n_1, n_2), где n_1 – число членов группы, набравших низкий балл, а n_2 – число членов группы набравших высокий балл и пусть $n_1 + n_2 = N$, $P = n_2 / N$ – доля членов группы набравших максимальный балл.

Рассуждая аналогично, для контрольной группы соответственно имеем: (m_1, m_2), $m_1 + m_2 = M$, $q = m_2 / M$.

Критическое значение $\varphi_{0,05}$ критерия Фишера для уровня значимости 0,05 равно 1,64.

Формула критерия Фишера: $\varphi_{\text{эмп.}} = 2\left|arc\sin\left(\sqrt{p}\right) - arc\sin\left(\sqrt{q}\right)\right| \cdot \sqrt{\dfrac{M \cdot N}{M + N}}$.

Значения функции y = arcsinx можно найти в книге: В.М. Брадис: Четырехзначные математические таблицы, Москва, «Просвещение», любые годы издания.

Алгоритм применения критерия:

1) На основании таблицы, содержащей результаты эксперимента контрольной и экспериментальной групп до начало и после завершения эксперимента (смотри таблицу 3), составить дихотомическую таблицу с указанием $P = n_2 / N$и $q = m_2 / M$;

2) Вычислить для сравниваемых выборок $\varphi_{\text{эмп.}}$ по формуле Фишера;

3) Сравнивать полученные значения с критическим значением $\varphi_{0,05} = 1{,}64$:

а) При $\varphi_{\text{эмп.}} \leq 1{,}64$, характеристики сравниваемых выборок совпадают с уровнем значимости 0,05;

b) При $\varphi_{\text{эмп.}} > 1{,}64$, достоверность различий характеристик сравниваемых выборок составляет 95%.

Применим критерий Фишера к таблице 1 (первая статья). Будем считать возможными два случая, два уровня знаний: «не усвоили материал» (число правильно решенных задач меньше либо равно 6) и «успешно усвоили материал» (число правильно решенных задач строго больше 6).

В таблице 1, N = 22 – экспериментальная группа, M = 25 – контрольная группа.

Таблица 7. (дихотомическая таблица)

	КГ до начала эксперимента (M = 25)	ЭГ до начала эксперимента (N = 22)	КГ после окончания эксперимента	ЭГ после окончания эксперимента
Доля, которую составляют учащиеся, не усвоившие материал	$m_1 = 12$ $\frac{12}{25}=0,48$	$n_1 = 13$ $\frac{13}{22}=0,59$	$m_1 = 12$ $\frac{12}{25}=0,48$	$n_1 = 10$ $\frac{10}{22}=0,45$
Доля, которую составляют учащиеся, усвоившие материал	$m_2 = 13$ $q_1=\frac{13}{25}=0,52$	$n_2 = 9$ $P_1=\frac{9}{22}=0,41$	$m_2 = 13$ $q_1=\frac{13}{25}=0,52$	$n_2 = 12$ $P_2=\frac{12}{22}=0,55$

Применим критерий Фишера к данным таблицы 7. Полученные результаты запишем в таблицу 8.

Таблица 8.

	КГ до начала эксперимента	ЭГ до начала эксперимента	КГ после окончания эксперимента	ЭГ после окончания эксперимента
КГ до начала эксперимента	0	0,75	0	0,21
ЭГ до начала эксперимента	0,75	0	0,75	0,96
КГ после окончания эксперимента	0	0,75	0	0,21
ЭГ после окончания эксперимента	0,21	0,96	0,21	0

Значение функции для клеток (1; 2) и (2; 1):

$$\varphi_{\text{эмп.}} = 2\left|\arcsin\left(\sqrt{0,41}\right)-\arcsin\left(\sqrt{0,52}\right)\right|\cdot\sqrt{\frac{25\cdot 22}{25+22}}=$$

$$=2\left|\arcsin\left(0,6403\right)-\arcsin\left(0,7211\right)\right|\cdot 3,4208=$$

$$=\left|0,70-0,81\right|\cdot 6,8417=0,75\ .$$

Значение функции для клеток (1; 3) и (3; 1):

$$\varphi_{\text{эмп.}} = 2\left|\arcsin\left(\sqrt{0,52}\right) - \arcsin\left(\sqrt{52}\right)\right| \cdot \sqrt{\frac{25 \cdot 22}{25+22}} = 0 .$$

Значение функции для клеток (1; 4) и (4; 1):

$$\varphi_{\text{эмп.}} = 2\left|\arcsin\left(\sqrt{0,55}\right) - \arcsin\left(\sqrt{0,52}\right)\right| \cdot \sqrt{\frac{25 \cdot 22}{25+22}} =$$
$$= \left|\arcsin(0,7416) - \arcsin(0,7211)\right| \cdot 6,8417 =$$
$$= \left|0,84 - 0,81\right| \cdot 6,8417 = 0,03 \cdot 6,8417 = 0,21 .$$

Значение функции для клеток (2; 3) и (3; 2):

$$\varphi_{\text{эмп.}} = 2\left|\sin\left(\sqrt{0,52}\right) - \arcsin\left(\sqrt{0,41}\right)\right| \cdot \sqrt{\frac{25 \cdot 22}{25+22}} =$$
$$= \left|\arcsin(0,7211) - \arcsin(0,6403)\right| \cdot 6,8417 =$$
$$= \left|0,81 - 0,70\right| \cdot 6,8417 = 0,75 .$$

Значение функции для клеток (2; 4) и (4; 2):

$$\varphi_{\text{эмп.}} = 2\left|\arcsin\left(\sqrt{0,55}\right) - \arcsin\left(\sqrt{0,41}\right)\right| \cdot \sqrt{\frac{25 \cdot 22}{25+22}} =$$
$$= \left|\arcsin(0,7416) - \arcsin(0,6403)\right| \cdot 6,8417 =$$
$$= \left|0,84 - 0,70\right| \cdot 6,8417 = 0,96 .$$

Значение функции для клеток (3; 4) и (4; 3):

$$\varphi_{\text{эмп.}} = 2\left|\arcsin\left(\sqrt{0,55}\right) - \arcsin\left(\sqrt{0,52}\right)\right| \cdot \sqrt{\frac{25 \cdot 22}{25+22}} =$$
$$= \left|\arcsin(0,7416) - \arcsin(0,7211)\right| \cdot 6,8417 =$$
$$= \left|0,84 - 0,81\right| \cdot 6,8417 = 0,21 .$$

На базе полученных данных (таблица 8) можно сделать вывод о достоверности различных состояний между контрольной и экспериментальной групп. Данные таблицы говорят о том, что коэффициент Фишера в этом эксперименте меньше критического значения 1,64. Таким образом, следует подвергать сомнению избранную методику обучения.

10. Критерий Стьюдента (t – критерий, – значимости критерий для средних значений нормальных распределений).

В педагогических исследованиях КС применяется для того чтобы установить различие между начальным уровнем знаний испытуемых и конечным их уровнем знаний. Иначе говоря: проверяется гипотеза H_0 о равенстве этих двух результатов (начальный и конечный) совместно с альтернативной гипотезой H_1 о неравенстве этих двух результатов (начальный и конечный).

Проводятся определенные статистические вычисления в результате которых гипотеза H_0 принимается (в таком случае гипотеза H_1 отвергается) или гипотеза H_0 отвергается (в таком случае гипотеза H_1 принимается).

Значение критерия *t* определяется двумя способами: вычислением и табличным способом.

Поясним как это делается на конкретном примере. Рассмотрим статистические данные экспериментальной группы из таблицы 1:

Таблица 9.

количество решенных задач в начале эксперимента	6	8	4	6	5	8	7	9	3	8	6	5	9	4	7	9	5	6	5	4	6	10
количество решенных задач в конце эксперимента	7	8	4	6	5	9	7	9	4	9	7	5	9	4	8	9	6	7	6	5	6	10
испытуемые	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

Чтобы понятнее был процесс вычисления критерия *t*, приведем часть статистической таблицы *t* – критерия Стьюдента при уровне значимости 0,10; 0,05; 0,01.

Таблица 10.

Число степеней свободы *df* →		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
α	0,10	6,3138	2,9200	2,3534	2,1318	2,0150	1,9432	1,8946	1,8595	1,8331	1,8125	1,7959	1,7823	1,7709	1,7613	1,7530	1,7459	1,7396	1,7341	1,7291	1,7247
	0,05	12,706	4,3027	3,1825	2,7764	2,5706	2,4469	2,3646	2,3060	2,2622	2,2281	2,2010	2,1788	2,1604	2,1448	2,1315	2,1199	2,1098	2,1009	2,0930	2,0860
	0,01	63,657	9,9248	5,8409	4,6041	4,0321	3,7074	3,4995	3,3554	3,2498	3,1693	3,1058	3,0545	3,0123	2,9768	2,9467	2,9208	2,8982	2,8784	2,8609	2,8453

Находим табличное значение *t* (пишут $t_{таб.}$), используя таблицы 9 и 10.

В этом случае df (число степеней свободы) равно: $n - 2 = 22 - 2 = 20$ (n – число пар оценок, решенных задач).

$\alpha = 0{,}05$ – уровень значимости для педагогических исследований, следовательно: $t_{таб.} = 2{,}0860$.

Находим вычислением значение t. В этом случае используется формула: $t = \frac{r}{m}$,

где $r = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}$, $\sigma_x^2 = \frac{\sum x^2}{n} - (\bar{x})^2$, $\sigma_y^2 = \frac{\sum y^2}{n} - (\bar{y})^2$, $\sigma_x = \sqrt{\sigma_x^2}$, $\sigma_y = \sqrt{\sigma_y^2}$, $m = \sqrt{\frac{1 - r^2}{n - 2}}$.

Для рационализации вычислений, составим вспомогательную таблицу – таблица 11.

Таблица 11.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	Σ	среднее	σ^2	σ
x	6	8	4	6	5	8	7	9	3	8	6	5	9	4	7	9	5	6	5	4	6	10	140	6,36		
y	7	8	4	6	5	9	7	9	4	9	7	5	9	4	8	9	6	7	6	5	6	10	150	6,82		
xy	42	64	16	36	25	72	49	81	12	72	42	25	81	16	56	81	30	42	30	20	36	100	1028	46,73		
x^2	36	64	16	36	25	64	49	81	9	64	36	25	81	16	49	81	25	36	25	16	36	100	970	44,09		
y^2	49	64	16	36	25	81	49	81	16	81	49	25	81	16	64	81	36	49	36	25	36	100	1096	49,82		

$$\sigma_x^2 = 44{,}09 - (6{,}36)^2 = 44{,}09 - 40{,}45 = 3{,}64 \text{ ; } \sigma_x = 1{,}91 \text{ ;}$$

$$\sigma_y^2 = 49{,}82 - (6{,}82)^2 = 49{,}82 - 46{,}51 = 3{,}31 \text{ ; } \sigma_y = 1{,}82 \text{ ; } r = \frac{46{,}73 - 6{,}36 \cdot 6{,}82}{1{,}91 \cdot 1{,}82} = \frac{3{,}35}{3{,}48} = 0{,}96 \text{ .}$$

$$m = \sqrt{\frac{1 - (0{,}96)^2}{22 - 2}} = \sqrt{\frac{1 - 0{,}92}{20}} = \sqrt{\frac{0{,}08}{20}} = \sqrt{0{,}004} = 0{,}06 \text{ .}$$

$$t = \frac{0{,}96}{0{,}06} = 16 \text{ .}$$

Если $t_{вычисл.} > t_{таб.}$ – исключается гипотеза H_0 и остается гипотеза H_1.

Если $t_{вычисл.} \le t_{таб.}$ – не исключается гипотеза H_0и считается недействительной гипотеза H_1.

В нашем примере ($t_{вычисл.} = 16$) > ($t_{таб.} = 2{,}0860$) – исключается гипотеза о равенстве начальных и конечных результатов и остается гипотеза о неравенстве этих двух результатов, то есть об истинности проведенного эксперимента.