

УГЛОВОЕ ПРЕОБРАЗОВАНИЕ Р. ФИШЕРА В СОВРЕМЕННЫХ ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

Гончарь П.С.¹

¹ФГБОУ ВО «Уральский государственный университет путей сообщения», Екатеринбург, e-mail: gonchar@usurt.ru

Сравнение групп, типичных по объёму для исследований в педагогике, по успеваемости (или аналогичным показателям) с помощью статистических критериев удобно проводить по данным, собираемым для других целей, но оно сталкивается с необходимостью применения непараметрических приемов из-за дискретности распределений соответствующих случайных величин. «Угловое преобразование» Р. Фишера, применяемое для стабилизации показателей (квантилей) распределения обобщающей величины φ^i нуждается в пересмотре процедуры применения и ревизии таблицы критических значений. В статье предлагается способ точного и имитационного моделирования ситуаций сопоставления групп разного объёма и контролируемой вероятности успешности по эмпирическим результатам успеваемости. Рассмотрены некоторые типовые по объёму сопоставляемых групп и предельные по реализуемым вероятностям «успеха» ситуации, позволяющие применять их как основу для принятия решений в похожих ситуациях, результаты представлены в компактной таблице. Даются практические рекомендации по использованию φ^i -критерия с минимальной трудоёмкостью и/или с глубоким анализом ситуации, интерпретации результатов его применения. Изложение иллюстрируется примером прикладного исследования, проведённого в Уральском государственном университете путей сообщения в связи с вынужденным переходом к дистанционным формам педагогического взаимодействия в 2019/2020 учебном году. Оказалось, что доля студентов, вовремя допущенных к первому массовому тестированию, проводимому в конце семестра, заметно и статистически значимо уменьшилась при введении дистанционных форм взаимодействия. Однако резко возросшая результативность тестирования (которое тоже стало проводиться дистанционно) не только компенсировала эти потери, но даже показала некоторое статистически малозначимое увеличение успеваемости по результату теста в общем списочном составе.

Ключевые слова: сравнение долей, угловое преобразование, статистический критерий, успеваемость, дистанционное обучение.

R. FISHER'S ANGULAR TRANSFORMATION IN MODERN PEDAGOGICAL RESEARCH

Gonchar P.S.¹

¹FGBOU VO "Ural State University of Railway Transport", Ekaterinburg, e-mail: gonchar@usurt.ru

Comparison of groups typical in terms of research in pedagogy, in terms of academic performance (or similar indicators) using statistical criteria, faces the need to use non-parametric methods due to the discreteness of the distributions of the corresponding random variables. "Angular transformation" by R. Fisher, used to stabilize the indicators (quantiles) of the distribution of the generalizing value φ^* , needs to be revised in the application procedure and revision of the table of critical values. The article proposes a method for groups comparing based on empirical results of academic performance and modeling. Some typical situations in terms of the volume of compared groups are considered, the results are presented in a compact table. Practical recommendations are given on the use of the φ^* -criterion. The presentation is illustrated by an example of applied research conducted at the Ural State University of Communications in connection with the forced transition to distance forms of pedagogical interaction in the 2019/2020 academic year. It turned out that the proportion of students admitted on time to the first mass testing, held at the end of the semester, significantly and statistically significantly decreased with the introduction of remote forms of interaction. However, the sharply increased effectiveness of testing (which also began to be carried out remotely) not only compensated for these losses, but even showed some statistically insignificant increase in performance according to the test result.

Keywords: comparison of shares, angular transformation, statistical criterion, performance, distance learning.

В практике принятия решений по рациональной типизации учебных ученических или студенческих групп, выявлению существенных различий между группами, оценке динамики

в группах и т.п., часто используются показатели, выраженные бинарным признаком для индивидуального случая и обобщающими долями (процентными долями) в группе, имеющие смысл успеваемости или результативности тестов, контрольно-оценочных мероприятий. Данные такого характера часто бывают доступны в открытых источниках или получаются в результате несложной обработки других данных, собираемых для иных целей. Кроме очевидных ситуаций полного совпадения групповых показателей, либо ситуаций, когда одна из групп продемонстрировала проявление (отсутствие) нужного признака во всех индивидуальных случаях, может быть поставлен вопрос о статистике различий в групповых показателях. Иными словами, если в групповых показателях есть разница, то правомерен вопрос о том, является ли она типичной для сопоставляемых групп данного объёма или свидетельствует о существенных различиях контингента, отсюда и привлечение статистических приёмов для решения этого вопроса.

При общем рассмотрении обозначенной проблемы выявляются следующие особенности, которые должны быть учтены в процедурах сопоставления групп. Во-первых, произвольность выбора того, что при принятии решений считается индивидуальным «успехом», проявление признака в индивидуальном испытании или его отсутствие, налагает требование независимости вычислительных процедур и их результатов от такого выбора. Во-вторых, если вероятности появления и отсутствия признака в индивидуальном случае примерно равны, то ожидается более широкий симметричный естественный разброс значений эмпирических долей, чем при экстремальных вероятностях, близких к крайним значениям 0 и 1 (100%). В-третьих, для сравнительно больших групп ожидается усреднение влияния случайных факторов, которое нужно учитывать. В-четвёртых, в практическом исследовании сопоставляться могут и группы примерно одинакового объёма, и существенно разного объёма.

Применение к данной проблеме подходов, основанных на использовании центральной предельной теоремы (утверждение о том, что «z-критерий» для долей имеет нормальное распределение), представлено в ряде источников [1-3], но представляется сомнительным, когда анализируются группы небольшого (порядка десяти индивидуальных случаев) или среднего (порядка тридцати индивидуальных случаев) объёма, из-за явной дискретности распределений. По опыту, даже при сопоставлении групп объёмом по 90–150 индивидуальных случаев, довольно часто проявляются эффекты, связанные с дискретностью распределений, а это уже объёмы, характерные для диссертационных исследований в педагогике. Кроме того, существенная асимметрия распределений долей при малых (0–20%) и больших (80–100%) случаях фиксации «успехов» приводит к применению специальных статистических критериев непараметрического характера. В качестве такого специального

метода в [4] предлагается использовать критерий Смирнова, но чаще в отечественной медицине [5-7], психологии и педагогике [8; 9] рекомендуется (с рядом ограничений на объёмы обследованных групп) использовать «угловое преобразование» эмпирических долей $\rho_{1,2}$, предложенное Р. Фишером $\varphi_{1,2} = 2 \arcsin \sqrt{\rho_{1,2}}$, и вычисление обобщающей величины, учитывающей объёмы обеих обследованных групп n_1 и n_2 , по формуле

$$\varphi_{\text{эмп}}^i = |\varphi_2 - \varphi_1| \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$$

Для сравнения применяется единая шкала (не совпадающая со стандартным нормальным распределением) критических значений (квантилей накопительной функции распределения), отражающих вероятность того, что эмпирические доли ρ_1 и ρ_2 , реализованные в группах, при одинаковых шансах проявления признака в них, продемонстрируют значение φ^i меньше указанного: $\varphi_{\text{кр}0,9}^i = 1,29$, $\varphi_{\text{кр}0,95}^i = 1,64$, $\varphi_{\text{кр}0,99}^i = 2,31$ и т.д. При больших значениях $\varphi_{\text{эмп}}^i$, превышающих критические, предлагается, с такой доверительной вероятностью, признать различие эмпирических долей существенным, нетипичным при единых шансах проявления признака успешности в двух группах на уровне значимости 10%, 5% или 1% соответственно.

По описанной процедуре, φ^i -критерий выглядит привлекательным: лёгкий способ вычисления главной статистики, компактная таблица критических значений, легко достигаемых в практике применения критерия, применимость к малым и большим группам, очень мягкие ограничения (по описанию критерия, при объёмах групп больше пяти случаев возможны любые сопоставления). Расчётная процедура обладает хорошей симметрией, при замене долей ρ_1 и ρ_2 на их дополнения $\bar{\rho}_{1,2} = 1 - \rho_{1,2}$ значение $\varphi_{\text{эмп}}^i$ гарантированно сохраняется. Скорость возрастания функции $\varphi_{1,2}(\rho_{1,2}) = 2 \arcsin \sqrt{\rho_{1,2}}$ существенно возрастает при приближении к экстремальным значениям $\rho_{1,2} = 0,1$ или $\rho_{1,2} = 0,9$ по сравнению с точкой симметрии $\rho_{1,2} = 0,5$. Значения $\varphi_{\text{эмп}}^i$ возрастают в связи с объёмами сопоставляемых групп, что удобно для планирования дальнейших исследований по предыдущим. Обзор практики использования φ^i -критерия в научных работах психолого-педагогического характера показывает, что он используется для сопоставления групп разного объёма (например, в [10-12]). Для сопоставления групп примерно одинакового малого объёма около 20 индивидуальных случаев критерий использован в [13], но выборки такого объёма редко являются основанием для научного обобщения, хотя могут возникать в педагогической практике, и их не следует игнорировать. Наиболее часто в сопоставлении используются группы, соответствующие примерно одной (в [14-16]) или двум-трём (в [17-19])

академическим учебным группам. Есть примеры ([20] и др.) применения φ^i -критерия для сопоставления групп объёмами порядка одной сотни индивидуальных случаев.

Принципиальные препятствия к широкому применению φ^i -критерия видятся в следующем. Во-первых, неясное происхождение таблицы критических значений, которая переходит из источника в источник около полувека. Во-вторых, неясность в способности критерия различить группы с исходно различными показателями, то есть отсутствие сведений о мощности рассматриваемого критерия, в зависимости от величины этого различия, асимметрии доли и объёмов групп. В-третьих, неуверенность в том, что критерий действительно настолько универсален в отношении асимметричности долей и объёмов сопоставляемых групп, как это декларируется.

Цель исследования

Проверка φ^i -критерия в контролируемых модельных абстрактных ситуациях в контролируемых условиях, с заранее заданными вероятностями проявления «успешности» в сопоставляемых группах, имеющих объёмы, характерные для научных и практических исследований в педагогике; разработка и обоснование практических рекомендаций по его прикладному применению.

Материал и методы исследования

Аналитическое моделирование. При объёмах сопоставляемых групп n_1 и n_2 все возможные доли $p_{1,2}^m = \frac{m}{n_{1,2}}$, где $m=0,1,2,\dots,n_{1,2}$, а их вероятности реализации в испытании определялись по точной формуле Бернулли $p_{n_{1,2}}(m) = C_{n_{1,2}}^m p_{1,2}^m (1-p_{1,2})^{n_{1,2}-m}$. Заполнение двумерной таблицы позволяет вычислить вероятности всех комбинаций долей через произведение вероятностей этих долей, как независимых событий. Параллельно во второй двумерной таблице такой же размерности вычисляются все возможные значения φ^i , они могут дублироваться в таблице, и им приписываются соответствующие вероятности реализации долей из первой таблицы. Далее в третьей таблице сохраняются и суммируются только вероятности из первой таблицы для ограниченных сверху произвольно заданным уровнем φ_0^i из второй таблицы – так получается точная теоретическая вероятность того, что в испытании φ^i не превысит φ_0^i . Для более сложной обработки, значения φ^i сортируются по возрастанию, вероятности дублирующихся значений φ^i складываются, и результаты моделирования могут быть представлены в виде накопительной (интегральной) функции распределения случайной величины φ^i . Для расчётов могут быть использованы электронные таблицы или пакеты прикладных программ для технических расчётов (независимо от платформы), а производительности современных персональных компьютеров достаточно для

точного моделирования любой ситуации с объёмами сопоставляемых групп до 300 индивидуальных случаев. Для получения всех результатов для этой статьи использован табличный процессор MS Excel.

Имитационное моделирование. Случайные величины с заранее заданными вероятностями «успеха» многократно разыгрываются и группируются в пакеты объёмом n_1 или n_2 , имитируя реализацию испытания. Такие пары «выборок» сопоставляются с помощью φ^i -критерия, проявляя закономерности распределения случайной величины φ^i и позволяя сравнить заранее заданные вероятности «успеха» с картиной распределения после реализации в выборке и обработки в виде накопительной (интегральной) функции распределения. Применялось только для ситуации сопоставления групп по 10 индивидуальных случаев (парное сопоставление групп разыгрывалось более трёх тысяч раз) и показало неплохое совпадение с точным теоретическим распределением, но потребовало расчётов в количестве, на десятичный порядок большем, чем аналитическое моделирование.

Применение φ^i -критерия в контролируемых условиях (основные результаты исследования)

Для аналитического получения распределения φ^i в паре сопоставляемых групп по 10 случаев с заранее заданными вероятностями «успеха» $p_{1,2}=0,5$ составим двумерную таблицу (табл. 1) возможных значений долей $\rho_{1,2}$, соответствующих этим парам $\rho_{1,2}$ значений φ^i и (с помощью формулы Бернулли для независимых испытаний и произведения вероятностей) вероятностей появления такого сочетания в обследовании. Аналогично моделируются другие ситуации, различающиеся заранее заданной вероятностью «успеха» в группах постоянного объёма.

Таблица 1

Данные для распределения величины φ^i

(сопоставление двух групп по 10 случаев, заранее заданная вероятность 50%) –

в скобках указана вероятность появления такого частного значения

	ρ_1	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
ρ_2	φ_1 φ_2	0 (0,001)	0,64 (0,010)	0,93 (0,044)	1,16 (0,117)	1,37 (0,205)	1,57 (0,246)	1,77 (0,205)	1,98 (0,117)	2,21 (0,044)	2,5 (0,01)	3,14 (0,001)
0	0 (0,001)	0 (0)	1,44 (0)	2,07 (0)	2,59 (0)	3,06 (0)	3,51 (0)	3,96 (0)	4,43 (0)	4,95 (0)	5,59 (0)	7,02 (0)
0,1	0,64 (0,010)	1,44 (0)	0 (0)	0,63 (0)	1,15 (0,001)	1,62 (0,002)	2,07 (0,002)	2,52 (0,002)	2,99 (0,001)	3,51 (0)	4,15 (0)	5,59 (0)
0,2	0,93 (0,044)	2,07 (0)	0,63 (0)	0 (0,002)	0,52 (0,005)	0,99 (0,009)	1,44 (0,011)	1,89 (0,009)	2,36 (0,005)	2,88 (0,002)	3,51 (0)	4,95 (0)
0,3	1,16 (0,117)	2,59 (0)	1,15 (0,001)	0,52 (0,005)	0 (0,014)	0,47 (0,024)	0,92 (0,029)	1,37 (0,024)	1,84 (0,014)	2,36 (0,005)	2,99 (0,001)	4,43 (0)
0,4	1,37 (0,205)	3,06 (0)	1,62 (0,002)	0,99 (0,009)	0,47 (0,024)	0 (0,042)	0,45 (0,051)	0,9 (0,042)	1,37 (0,024)	1,89 (0,009)	2,52 (0,002)	3,96 (0)
0,5	1,57 (0,246)	3,51 (0)	2,07 (0,002)	1,44 (0,011)	0,92 (0,029)	0,45 (0,051)	0 (0,061)	0,45 (0,051)	0,92 (0,029)	1,44 (0,011)	2,07 (0,002)	3,51 (0)
0,6	1,77	3,96	2,52	1,89	1,37	0,9	0,45	0	0,47	0,99	1,62	3,06

	(0,205)	(0)	(0,002)	(0,009)	(0,024)	(0,042)	(0,051)	(0,042)	(0,024)	(0,009)	(0,002)	(0)
0,7	1,98 (0,117)	4,43 (0)	2,99 (0,001)	2,36 (0,005)	1,84 (0,014)	1,37 (0,024)	0,92 (0,029)	0,47 (0,024)	0 (0,014)	0,52 (0,005)	1,15 (0,001)	2,59 (0)
0,8	2,21 (0,044)	4,95 (0)	3,51 (0)	2,88 (0,002)	2,36 (0,005)	1,89 (0,009)	1,44 (0,011)	0,99 (0,009)	0,52 (0,005)	0 (0,002)	0,63 (0)	2,07 (0)
0,9	2,5 (0,01)	5,59 (0)	4,15 (0)	3,51 (0)	2,99 (0,001)	2,52 (0,002)	2,07 (0,002)	1,62 (0,002)	1,15 (0,001)	0,63 (0)	0 (0)	1,44 (0)
1	3,14 (0,001)	7,02 (0)	5,59 (0)	4,95 (0)	4,43 (0)	3,96 (0)	3,51 (0)	3,06 (0)	2,59 (0)	2,07 (0)	1,44 (0)	0 (0)

Для имитационного моделирования (многократного розыгрыша) сопоставления двух групп с заранее заданными вероятностями «успеха» $p_{1,2}=0,5$ бинарная случайная величина разыгрывалась более шести тысяч раз и из этих результатов были сформированы три тысячи «выборок», характеризуемые долей «успешных» опытов. На рисунке 1 представлены результаты сравнения распределений величины φ^i , полученные аналитически, с помощью имитации (многократного розыгрыша) и табличных данных из [8] – красной линией.

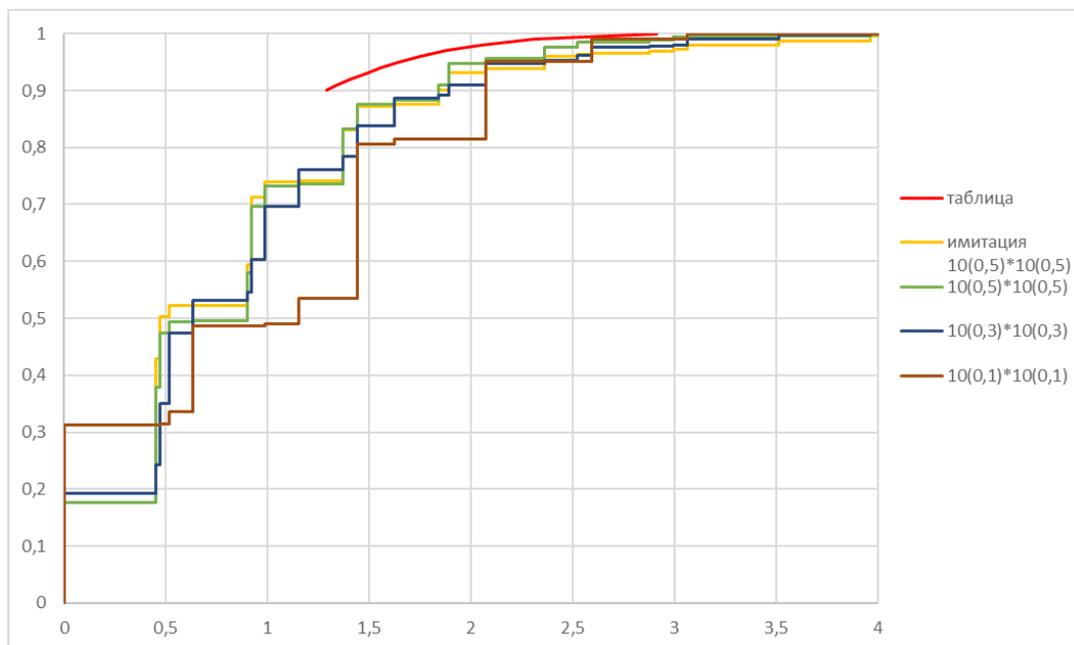


Рис. 1. Сравнение накопительных функций распределения величины φ^i с заранее заданными совпадающими вероятностями «успеха» $p_{1,2}=0,1-0,5$ в группах по 10 индивидуальных случаев

По результатам сравнения, распределение значений φ^i в имитационном моделировании для рассмотренной частной ситуации неплохо совпало с аналитической зависимостью (желтая и зелёная линии на рис. 1 выше), но табличные значения не выдерживают сравнения с расчётными, при их использовании доверительная вероятность принимаемых решений оказывается гораздо ниже декларируемой, и наибольшие расхождения с табличными значениями наблюдаются при экстремальной заранее заданной

вероятности «успеха» $p_{1,2}=0,1(0,9)$. Сравнительные данные для групп с одинаковой заранее заданной вероятностью «успеха» в индивидуальном случае приведены в п.п. 2 – 4 таблицы 2.

Таблица 2

Сопоставление результатов, ожидаемых при применении φ^i -критерия

№	Сопоставляемые группы по 10 случаев	Значение φ^i , не более		
		1,29	1,64	2,31
1	Любые – по таблице критических значений из [8]	90%	95%	99%
2	Одинаковая индивидуальная вероятность «успеха» 0,5	74%	88%	96%
3	Одинаковая индивидуальная вероятность «успеха» 0,3 (0,7)	76%	88%	95%
4	Одинаковая индивидуальная вероятность «успеха» 0,1 (0,9)	53%	81%	95%
5	Индивидуальные вероятности «успеха» 0,3 (0,7) и 0,5	57%	74%	87%
6	Индивидуальные вероятности «успеха» 0,1 (0,9) и 0,3 (0,7)	47%	62%	76%
7	Индивидуальные вероятности «успеха» 0,3 и 0,7	22%	39%	58%
8	Индивидуальные вероятности «успеха» 0,1 (0,9) и 0,5	19%	33%	49%

Для качественной оценки мощности φ^i -критерия используем некоторые ситуации, когда заранее заданные вероятности и, соответственно, проявленные доли «успеха» ожидаются различными, наиболее яркие результаты сравнения соответствуют экстремальным значениям заранее заданной вероятности «успеха» 0,1 (0,9) в одной из групп и приведены далее на рисунке 2 и выше в таблице 2.

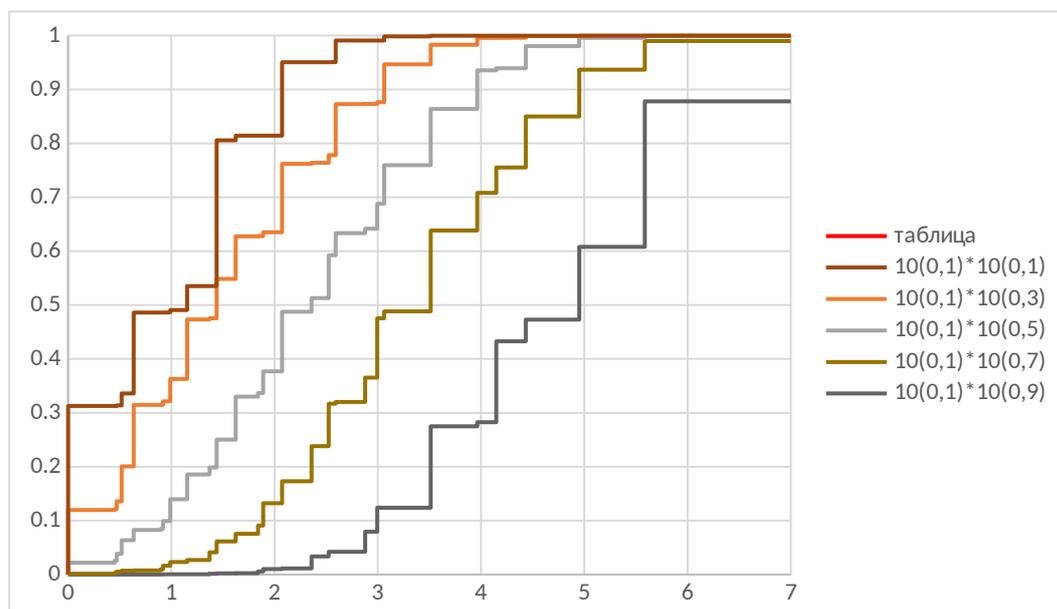


Рис. 2. Накопительная функция распределения величины φ^i для заранее заданной вероятности «успехов» $p_1=0,1(0,9)$ в одной из групп и разной вероятности «успеха» в другой группе

Визуально графики различаются друг от друга гораздо сильнее, чем графики для совпадающих заранее заданных вероятностей «успеха», но при практическом применении φ^i -критерия приходится различать ситуации не по «развёрнутой» картине накопительных

функций распределения, а по единственному значению $\varphi_{эмп}^i$, и если диапазоны наиболее вероятных значений φ^i в двух ситуациях существенно пересекаются, то различить эти ситуации может оказываться затруднительным. Например (п. 4 таблицы 2), если значение $\varphi_{эмп}^i$ окажется равным 1,29, то его можно считать типичным и для ситуации с заранее заданными равными индивидуальными вероятностями «успеха», так как примерно с равной вероятностью φ^i может быть и больше, и меньше этого значения. Но (п. 6 таблицы 2) и при заранее заданном различии 0,2 в индивидуальных вероятностях «успеха» это значение тоже типично, то есть сохраняется неопределённость в выборе одной из этих альтернатив. Однако основной алгоритм принятия решения в процедуре применения φ^i -критерия (с учётом необходимой поправки к табличным значениям) приводит здесь к утверждению «данных за существенное расхождение долей не получено», что не противоречит приведённым в таблице 2 точным результатам. Если рассмотреть возможное значение $\varphi_{эмп}^i=2,31$, или немного больше, то такое значение придётся признать нетипично высоким для совпадающих долей (пп. 2 – 4 таблицы 2), но оно остаётся достаточно типичным для случая с заранее заданными разными вероятностями «успеха» в двух группах этого объёма. Далее, при значениях $\varphi_{эмп}^i=3,5$ и выше до максимально возможного значения $\varphi_{max}^i=\pi\cdot\sqrt{5}=7,02$, наблюдаются значения, нетипично высокие для обеих ситуаций: и с заранее заданной одинаковой вероятностью «успеха» в двух группах, и для различающейся. То есть можно заключить, что различие в долях отражает более яркое различие в вероятностях «успеха» в группах, чем между 0,1 и 0,3 (либо при сборе данных произошла какая-то серьёзная ошибка).

Заметим, что «ошибка второго рода», то есть шансы ошибочно признать ситуации с различными вероятностями «успеха», схожими между собой по этому параметру, оказались малы из-за аккуратной традиционной формулировки вывода в φ^i -критерии: отсутствие данных за различие выборок не даёт основания признавать выборки схожими без специального дополнительного исследования - надо задуматься о том, какие параметры (разные вероятности «успеха» в группах) должны считаться неприемлемыми, и статистически исключить такую ситуацию по причине нетипично малого значения $\varphi_{эмп}^i$, что не входит в традиционный алгоритм использования φ^i -критерия.

В таблице 3, с целью взаимного сравнения, приведены значения φ^i , соответствующие некоторым квантилям в ситуациях с заданным объёмом сопоставляемых групп, типичным для исследований в педагогике, и разными заранее заданными вероятностями «успеха»: 0,1 и 0,5 в обеих группах, 0,1 в одной и 0,5 в другой группе. Интерпретация результатов дана ниже.

Таблица 3

Параметры распределений (квантили) случайной величины φ^i

n_1 n_2	$n_1=10$	$n_1=30$	$n_1=60$	$n_1=90$	$n_1=150$	$n_1=300$
$n_2=10$	1,44-2,08-2,08 1,38-1,85-2,10 1,44-2,36-3,52	1,31-2,05-2,31 1,31-1,68-2,10 1,75-4,31	1,89-2,10-2,33 1,31-1,70-2,08 1,87-4,51	1,94-2,15-2,25 1,35-1,74-2,07 1,97-4,65		
$n_2=30$		1,48-1,89-2,50 1,31-1,60-2,10 1,60-4,58	1,33-1,75-2,32 1,33-1,66-2,10 3,31-5,2	1,32-1,80-2,28 1,29-1,7-2,02 3,60-5,4	1,34-1,74-2,31 1,29-1,67-2,00 3,79-5,7	1,37-1,72-2,28 1,39-1,68-2,00 4,01-5,9
$n_2=60$			1,35-1,75-2,09 1,29-1,66-2,03 4,30-6,0	1,36-1,66-1,98 1,28-1,68-1,97 4,80-6,5	1,31-1,68-2,05 1,28-1,67-1,98 5,3-7,0	1,32-1,71-2,04 1,28-1,66-1,99 5,8-7,5
$n_2=90$				1,29-1,71-2,1 1,35-1,65-1,96 5,5-7,2	1,31-1,69-2,01 1,29-1,65-1,98 6,2-7,9	1,31-1,67-2,01 1,29-1,66-1,97 6,9-8,6
$n_2=150$		$A_1-A_2-A_3$ $B_1-B_2-B_3$ C_1-C_2			1,30-1,64-2,00 1,28-1,63-1,97 7,3-9,0	1,30-1,67-2,00 1,27-1,67-1,95 8,5-11,0
$n_2=300$						1,30-1,65-1,98 1,31-1,64-1,97 10,6-12,3

Для двух сопоставляемых групп с указанными в заголовках объёмами n_1 и n_2 и заранее заданной вероятностью «успеха» $p_1=p_2=0,1(0,9)$ с доверительной вероятностью 0,8, 0,9 или 0,95 не характерны значения φ^i , превышающие A_1 , A_2 или A_3 соответственно, а при $p_1=p_2=0,5$ критические значения B_1 , B_2 или B_3 ; если $\varphi_{эмп}^i$ заметно превышает указанные критические значения – можно утверждать, что сопоставляемые группы различаются по частоте «успеха».

Для заранее заданных вероятностей успеха $p=0,5$ в одной группе и $p=0,1(0,9)$ в другой группе менее 20% реализаций показывают значение φ^i меньше C_1 , и менее 20% реализаций – более C_2 ; если $\varphi_{эмп}^i$ заметно превышает указанные критические значения – можно утверждать, что сопоставляемые группы различаются по частоте «успеха» значительней, чем эти заранее заданные вероятности, либо при фиксации и обработке эмпирических данных допущены ошибки.

Если зафиксировать объём одной из сопоставляемых групп и рассмотреть критические значения, соответствующие квантилям распределений, в зависимости от объёма второй группы, то значения φ^i , характерные и критические для различающихся заранее заданных вероятностей «успеха» в сопоставляемых группах, устойчиво возрастают, а рассчитанные критические значения в ситуациях с одинаковыми заранее заданными вероятностями демонстрируют более сложную картину (рис. 3). При объёме хотя бы одной из сопоставляемых групп до 30 случаев критические значения слабо предсказуемы, но не превышают значения $\varphi^i=2,5$, сказывается дискретность распределений. Если объёмы обеих сопоставляемых групп от 60 случаев, то критические показатели распределений оказываются

достаточно стабильными. На рисунке 3 отдельно цветом показаны эти два типа ситуаций, а данные для 300 случаев в одной из групп выделены отдельно и показаны чёрным пунктиром.

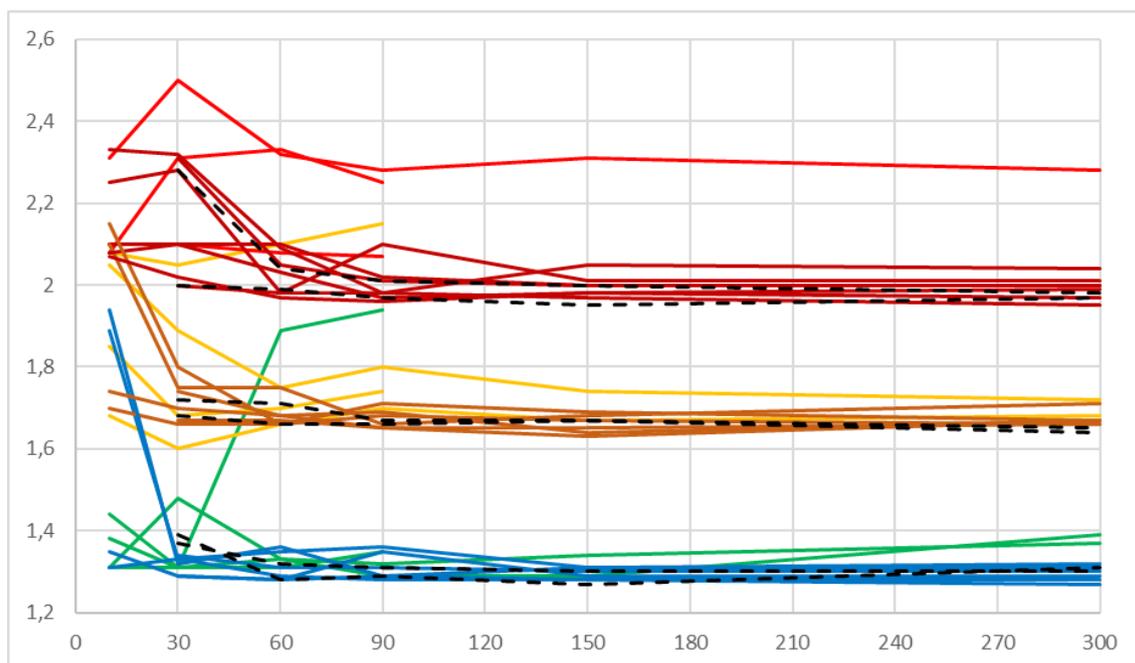


Рис. 3. Динамика критических значений φ^i при фиксированном значении объёма одной группы и одинаковой заранее заданной вероятности «успеха»

Практические рекомендации

1. Несмотря на сложности в применении, не указанные в традиционном алгоритме применения, полностью отказываться от углового преобразования не надо, так как оно сближает между собой показатели при сопоставлении групп объёмом от 60 случаев в каждой. Наоборот, можно утверждать, что значение $\varphi_{эмп}^i > 2,5$ свидетельствует о различной вероятности «успеха» в группах (без уточнения, насколько велика эта разница) с доверительной вероятностью не менее 95%.

2. Лучший способ применения углового преобразования – точное аналитическое моделирование случайной величины φ^i (процедура описана выше) для тех объёмов сопоставляемых групп, возможных (совпадающих или различных, по логике проводимого исследования) вероятностей «успеха», оно позволяет ответить на вопрос, насколько $\varphi_{эмп}^i$ типично для этих гипотетических значений.

3. Возможно использование рассмотренных в таблице 3 ситуаций и критических значений, приведённых в ней, для этого надо выбрать одно или несколько близких по объёму групп позиций и качественно обобщить выводы.

4. Выводы, сделанные с помощью малых по объёму групп и в ситуациях, когда частота успеха менее 0,1, являются ненадёжными, и φ^i -критерий не способен чудесным образом изменить это. Для научного педагогического обобщения лучше отказаться от

сопоставления групп менее 60 индивидуальных случаев. Примером сомнительного использования ϕ^i -критерия может служить [21], где сопоставлялись группы объемом от 2 до 15 случаев.

5. Материалы статьи могут быть использованы как основа для лабораторной работы по основам статистики в образовательной практике вузов.

Пример

Студенты специальности «Эксплуатация железных дорог» ФГБОУ ВО «Уральский государственный университет путей сообщения» каждый семестр изучения курса математики заканчивают тестированием с использованием материала Единого портала интернет-тестирования в сфере образования [22]. Для выявления особенностей влияния вынужденного перехода к дистанционным формам взаимодействия были сопоставлены данные студентов до этого перехода и после него по трём показателям успешности: доля допущенных студентов от общего списочного состава, доля успешных (зачтённых) случаев тестирования среди допущенных и в общем списочном составе. Симметрия расчётной процедуры ϕ^i -критерия позволила использовать его для значений долей «успеха» в сопоставляемых группах (обобщённых потоках студентов), близких к 0,9, так же как для значений, близких к 0,1. Результаты представлены в таблице 4. В рассмотренном примере оказалось, что доля студентов, вовремя допущенных к первому массовому тестированию, заметно уменьшилась при введении дистанционных форм взаимодействия. Однако резко возросшая результативность тестирования (которое тоже стало проводиться дистанционно) не только компенсировала эти потери, но даже показала некоторое увеличение успеваемости по результату теста в общем списочном составе.

Таблица 4

Сопоставление успешности тестирования студентов при введении дистанционных форм педагогического взаимодействия

Показатель	Зимняя сессия 2019/20 уч. года	Летняя сессия 2019/20 уч. года и зимняя сессия 2020/21 уч. года	ϕ^i - статистика
Общее количество студентов	176	352	
Допущено к первому массовому тестированию	159	266	
Успешно прошли тест	105	237	
Доля (от общего количества) допущенных к тесту студентов	0,90	0,76	4,35*
Доля (от допущенных) студентов, успешно прошедших тест	0,66	0,89	5,74**

Доля (от общего количества) студентов, успешно прошедших тест	0,60	0,67	1,73***
---	------	------	---------

* близкие данные таблицы 2 для $p_1=p_2=0,1$ в группах объёма $n_1=150$ и $n_2=300$, в сравнении со значением $\varphi_{эмп}^{\dot{c}} > 2,00$, позволяют утверждать, что сопоставляемые группы **различаются** по доле «успеха» с доверительной вероятностью не менее чем 95% (на уровне значимости менее 5%). С другой стороны, фактические значения $\varphi_{эмп}^{\dot{c}} < 8,5$ ожидаемо малы для яркого (как 0,1 и 0,5) различия долей. Точное моделирование для данных объёмов групп и успешности показывает, что такое значение $\varphi_{эмп}^{\dot{c}}$ характерно для реализации случайной величины $\varphi^{\dot{c}}$ с заранее заданными вероятностями успеха $p_1=0,90$ и $p_2=0,76$ (58% случаев показывают $\varphi^{\dot{c}}$ менее этого значения, 42% – больше). А для усреднённого по выборкам значения $p_1=p_2=0,80$ это значение $\varphi_{эмп}^{\dot{c}}$ оказывается нехарактерно высоким (99% реализаций $\varphi^{\dot{c}}$ демонстрируют значение $\varphi^{\dot{c}} < \varphi_{0,99}^{\dot{c}} = 2,60 \dot{c}$, всё это подтверждает предшествующий вывод о **различии** двух сопоставляемых групп;

** близкие данные таблицы 2 для $p_1=p_2=0,1$ в группах объёма $n_1=150$ и $n_2=300$, в сравнении со значением $\varphi_{эмп}^{\dot{c}} > 2,00$, позволяют утверждать, что сопоставляемые группы **различаются** по доле «успеха» с доверительной вероятностью не менее чем 95% (на уровне значимости менее 5%). С другой стороны, фактические значения $\varphi_{эмп}^{\dot{c}} < 8,5$ ожидаемо малы для яркого (как 0,1 и 0,5) различия долей. Точное моделирование для данных объёмов групп и успешности показывает, что такое значение $\varphi_{эмп}^{\dot{c}}$ характерно для реализации случайной величины $\varphi^{\dot{c}}$ с заранее заданными вероятностями успеха $p_1=0,66$ и $p_2=0,89$ (52% случаев показывают $\varphi^{\dot{c}}$ менее этого значения, 48% – больше). А для усреднённого по выборкам значения $p_1=p_2=0,64$ это значение $\varphi_{эмп}^{\dot{c}}$ оказывается нехарактерно высоким (99% реализаций $\varphi^{\dot{c}}$ демонстрируют значение $\varphi^{\dot{c}} < \varphi_{0,99}^{\dot{c}} = 2,60 \dot{c}$, всё это подтверждает предшествующий вывод о **различии** двух сопоставляемых групп;

*** близкие данные таблицы 2 для $p_1=p_2=0,5$ в группах объёма $n_1=150$ и $n_2=300$, в сравнении со значением $\varphi_{эмп}^{\dot{c}} > \varphi_{0,90}^{\dot{c}} = 1,67$ (но $\varphi_{эмп}^{\dot{c}} < \varphi_{0,95}^{\dot{c}} = 1,95 \dot{c}$, позволяют утверждать, что сопоставляемые группы **совпадают либо слабо различаются** по доле «успеха» с доверительной вероятностью не менее чем 90% (на уровне значимости менее 10%). Точное моделирование для данных объёмов групп и успешности показывает, что такое значение $\varphi_{эмп}^{\dot{c}}$ характерно для реализации случайной величины $\varphi^{\dot{c}}$ с заранее заданными вероятностями успеха $p_1=0,60$ и $p_2=0,67$ (56% случаев показывают $\varphi^{\dot{c}}$ менее этого значения, 44% – больше). А для усреднённого по выборкам значения $p_1=p_2=0,64$ это значение $\varphi_{эмп}^{\dot{c}}$

оказывается довольно высоким (91% реализаций φ^i демонстрируют значение менее $\varphi_{эмп}^i$, 9% – больше), всё это подтверждает предшествующий вывод о *совпадении либо слабом различии* двух сопоставляемых групп.

Заключение

Использование углового преобразования для сопоставления групп может принести значительную пользу в прикладных и научных педагогических задачах. Проведённое исследование уточнило границы применимости и особенности применения этого приёма. Для «сверхбыстрого» решения о различии двух групп по доле «успеха» можно пользоваться значением $\varphi_{кр}^i = 2,5$, более детальное понимание ситуации достигается использованием табличных данных или точным моделированием ситуации.

Список литературы

1. Гржибовский А.М. Выбор статистического критерия для проверки гипотез // Экология человека. 2008. № 11. С. 48-57.
2. Левин Д.М., Стефан Д., Кребель Т.С., Беренсон М.Л. Статистика для менеджеров с использованием Microsoft Excel. М.: Издательский дом «Вильямс», 2004. 1312 с.
3. Теория статистики / под ред. проф. Г.Л. Громыко. М.: ИНФРА-М, 2021. 465 с.
4. Фокин Р.Р., Атоян А.А. Проверка статистических гипотез при обосновании результатов квалификационных работ по некоторым направлениям и специальностям высшей школы и ее компьютерная реализация // Современные наукоемкие технологии. 2016. № 5-3. С. 614-618.
5. Гублер Е.В. Вычислительные методы анализа и распознавания патологических процессов. Л.: Медицина, 1978. 296 с.
6. Барникова И.Э., Самсонова А.В. Информационные технологии в обработке анкетных данных в педагогике и биомеханике спорта. СПб.: Национальный государственный университет физической культуры, спорта и здоровья им. П.Ф. Лесгафта, 2017. 103 с.
7. Гржибовский А.М., Унгурияну Т.Н. Анализ биомедицинских данных с использованием пакета статистических программ SPSS. Архангельск: Изд-во Северного государственного медицинского университета, 2017. 293 с.
8. Стариченко Б.Е. Обработка и представление данных педагогических исследований с помощью компьютера. Екатеринбург, Урал. гос. пед. ун-т, 2004. 218 с.
9. Сидоренко Е.В. Методы математической обработки в психологии. СПб.: Речь, 2010. 350 с.

10. Гончарь П.С., Нематова Е.Д., Петухова Ю.С. Электронное портфолио для студентов университета с разным учебным опытом // Международный журнал экспериментального образования. 2021. № 1. С. 47-51.
11. Евтушенко И.И. Внеучебная деятельность как фактор формирования правовой культуры старшеклассников // Современные проблемы науки и образования. 2013. № 5. [Электронный ресурс]. URL: <https://science-education.ru/ru/article/view?id=10104> (дата обращения: 25.12.2022).
12. Клименских М.В., Корепина Н.А., Шека А.С., Виндекер О.С. Особенности восприятия дистанционного обучения студентами и преподавателями вуза // Современные проблемы науки и образования. 2018. № 1. [Электронный ресурс]. URL: <https://science-education.ru/ru/article/view?id=27421> (дата обращения: 25.12.2022).
13. Карманова А.В., Кондратенко Л.Н. Исследование эффективности структурирования на основе герменевтического подхода профильных задач по математике в аграрном ВУЗе // Современные проблемы науки и образования. 2020. № 5. [Электронный ресурс]. URL: <https://science-education.ru/ru/article/view?id=30108> (дата обращения: 25.12.2022).
14. Ковчина Н.В., Шушерина О.А. Изучение подготовленности будущего бакалавра к социальному взаимодействию в профессиональной сфере методами математической статистики // Современные наукоемкие технологии. 2016. № 6-2. С. 349-353.
15. Аль-Янаи Е.К. Диагностика ценностного отношения старшеклассников к русскому языку // Современные проблемы науки и образования. 2021. № 2. [Электронный ресурс]. URL: <https://science-education.ru/ru/article/view?id=30620> (дата обращения: 25.12.2022).
16. Лукьянова М.И., Галацкова И.А. Изучение особенностей эмоционального состояния обучающихся экспериментальной образовательной организации // Современные проблемы науки и образования. 2019. № 5. [Электронный ресурс]. URL: <https://science-education.ru/ru/article/view?id=29232> (дата обращения: 25.12.2022).
17. Липунова О.В. Формирование у студентов педагогического вуза готовности к работе с детьми с ограниченными возможностями здоровья // Современные проблемы науки и образования. 2018. № 3. [Электронный ресурс]. URL: <https://science-education.ru/ru/article/view?id=27567> (дата обращения: 25.12.2022).
18. Бакланов И.О., Бирюкова И.П. Методы и средства диагностики формирования компетенций в лабораторных практикумах по естественно-научным дисциплинам // Современные проблемы науки и образования. 2020. № 4. [Электронный ресурс]. URL: <https://science-education.ru/ru/article/view?id=30016> (дата обращения: 25.12.2022).
19. Абрамова И.Е., Шишмолина Е.П. Формирование навыков самоорганизации и самооценивания студентов в конкурентной иноязычной обучающей среде: практический

опыт // Образование и наука. 2020. № 22 (10). С. 161-185. DOI: 10.17853/1994-5639-2020-10-161-185 30620.

20. Моторная С.Е. Будущий специалист-профессионал в условиях пандемии: исследование выбора стратегии поведения в конфликте и этнотолерантности // Современные наукоемкие технологии. 2021. № 4. С. 194-198.

21. Бужинская Н.В., Васева Е.С., Шкабара И.Е. Когнитивный стиль будущего специалиста ит-сферы в процессе организации командной работы // Образование и наука. 2022. № 24 (4). С. 79-111. DOI: 10.17853/1994-5639-2022-4-79-111.

22. Оценка и мониторинг образовательных достижений студентов. [Электронный ресурс]. URL: <https://i-exam.ru/> (дата обращения: 25.12.2022).