

УДК 37.012.4
ББК 74 в6
Ш 42
Л.В. Шелехова

**Статистическая проверка простых гипотез
в педагогических исследованиях**
(Рецензирована)

Аннотация:

В статье рассматривается статистическая проверка простых гипотез как процедура обоснованного сопоставления высказанной гипотезы с имеющейся выборкой, осуществляемая при помощи того или иного статистического критерия; приводится классификация задач педагогического исследования и непараметрических методов их решения; отмечается, что статистическая гипотеза: а) появляется одновременно с ведущей идеей исследования и ею определяется; б) возникает на основании известных знаний, но выходит за их пределы; в) позволяет систематизировать предположения исследователя и представить их в чётком и лаконичном виде; г) позволяет соединить известные познания с новым, требующим обоснования; д) экспериментально и практически проверяема.

Ключевые слова:

Статистическая гипотеза, статистическая проверка простых гипотез, статистический критерий, задача статистического исследования,

Использование математических методов в педагогических исследованиях предполагает создание формального математического аппарата, пригодного для изучения педагогических явлений и процессов на специальном объекте – модели, являющейся промежуточным звеном между исследователем и предметом исследования. Подобное замещение допускается при условии, что реальный объект изучения и его математическая модель имеют сходство в основополагающих элементах и отношениях между ними. Поэтому, изучая модель, можно получить новые данные о предмете исследования, которые в обычных условиях определить достаточно сложно, а в некоторых случаях и невозможно.

Построение математической модели предполагает качественное описание предмета исследования, формулирование статистической гипотезы и ее проверки.

Статистическая гипотеза – это предположение об определенных эмпирических характеристиках распределения в данной совокупности, которое: а) появляется одновременно с ведущей идеей исследования и ею определяется; б) возникает на основании известных знаний, но выходит за их пределы; в) позволяет систематизировать предположения исследователя и представить их в чётком и лаконичном виде; г) позволяет соединить известные познания с новым, требующим обоснования; д) экспериментально и практически проверяема.

Статистическая проверка гипотезы представляет собой процедуру обоснованного сопоставления высказанной гипотезы с имеющейся выборкой, осуществляемая при помощи того или иного статистического критерия (метода). Использование статистических методов, к сожалению, не всегда сопровождается четкими представлениями о возможностях и границах их применения, так как в научно-методической литературе нет единого подхода, позволяющего исследователю определиться с выбором статистического критерия, соответствующего задачам педагогического исследования. Рассмотрим в качестве примера некоторые подходы

решения проблемы, которые отражают основные аспекты данного вопроса.

Б.А. Сосновский при решении данной проблемы предлагает основываться на мощности критерия – чем выше мощность критерия, тем он предпочтительней. Однако при этом не указывается, как данный критерий будет соответствовать математической модели эмпирического исследования [1].

Е.В. Сидоренко полагает, что выбор статистического критерия следует, прежде всего, определять, исходя из задач исследования и количества элементов (иерархий) выборки. Однако при этом: 1) соответствие гипотез исследования и статистических гипотез критерия должно устанавливаться самим исследователем, что затрудняет выбор критерия; 2) классификация критериев относительно количества элементов (иерархий) представлена согласно используемым автором статистических таблиц, что неоправданно сужает границы применимости критериев [2].

Ю.Н. Толстова, основываясь на анализе условий (ограничений) применимости статистических методов, предлагает типологию общепринятых статистических критериев, сводящуюся к простому их перечислению, что мало похоже на классификацию и поэтому затрудняет ее использование. При этом не учитываются задачи и гипотезы исследования, что может привести к некорректным выводам в психолого-педагогических исследованиях [3].

А.Д. Наследов отмечает, что выбор критерия определяется статистической гипотезой, но при этом не приводит видов возможных гипотез, ссылаясь на то, что их количество сводится к небольшому числу типичных исследовательских ситуаций, в основу классификаций которых, прежде всего, должны быть положены типы шкал, количество и соотношение сравниваемых групп. Андрей Дмитриевич приводит обоснование этих утверждений, но при этом не дает четких рекомендаций по использованию тех или иных методов, применяемых в педагогическом исследовании [4].

Анализ научной и учебно-методической литературы [1, 2, 3, 4, 5, 6, 7, 8, 9] позволил выявить необходимые критерии при выборе статистического метода:

1. Тип переменных (признаков).
2. Шкала, которая использовалась при измерении психолого-педагогических показателей (номинативная, порядковая, интервальная и отношений).
3. Тип распределения данных, который получился в исследовании:
 - а) нормальное распределение (в этом случае используют параметрические критерии);
 - б) тип распределения данных не известен (не имеет значения) (применяют непараметрические критерии).
4. Количество элементов в выборке (при небольших объемах выборки испытуемых целесообразно использовать непараметрические критерии, которые дают большую достоверность выводам, независимо от того, получено ли в исследовании нормальное распределение данных или нет).
5. Тип исследовательской задачи (для непараметрических методов: выявление различий в уровне исследуемого признака; оценка сдвига значений исследуемого признака; выявление различий в распределении признака; выявление степени согласованности изменений).
6. Формулировка основной (нулевой) и альтернативной статистических гипотез.
7. Ограничения, которые имеет каждый критерий.

Учитывая вышесказанное, можно предложить следующую классификацию исследовательских задач и непараметрических методов их решения, учитывающую условия применимости критерия (тип шкалы, количество выборок и замеров) и соответствующие формулировки нулевой и альтернативной статистических гипотез.

Классификация исследовательских задач и непараметрических методов их решения

задача	условия	гипотеза	шкала	критерий
1. Выявление различий в уровне исследуемого признака				
оценка различий между несколькими выборками по уровню какого-либо признака, количественно измеренного.	две независимые выборки	H_0 : уровень признака в первой выборке не превышает уровня признака во второй выборке. H_1 : уровень признака в первой выборке превышает уровень признака во второй выборке.	порядковая, интервальная	Q - критерий Розенбаума U - критерий Манна-Уитни W - критерий Вилкоксона
	более двух выборок	H_0 : между выборками 1, 2, 3 и т.д. существуют лишь случайные различия по уровню исследуемого признака. H_1 : между выборками 1, 2, 3 и т.д. существуют неслучайные различия по уровню исследуемого признака.	порядковая, интервальная	H - критерий Крускала-Уолиса
			номинативная, порядковая, интервальная	φ^* - критерий Фишера
		H_0 : тенденция возрастания значений признака при переходе от выборки к выборке является случайной. H_1 : тенденция возрастания значений признака при переходе от выборки к выборке не является случайной.	порядковая, интервальная	S - критерий тенденций Джонкира
2. Оценка сдвига значений исследуемого признака				
(сдвиг - это разность между вторым и первым замерах одного признака на одной и той же выборке испытуемых)				
а) временные, ситуационные, умозрительные, измерительные (одни и те же показатели, измеренные у одних и тех же испытуемых в разное время, в ситуациях в разных представляемых условиях или разными способами).	два замера одного признака на одной и той же выборке	H_0 : отсутствие значимых различий в состоянии изучаемого свойства при первичном и вторичном измерениях его состояния у респондентов рассматриваемой совокупности. H_1 : состояния изучаемого свойства значимо различны в одной и той же совокупности респондентов при первичном измерении этого свойства и при вторичном его измерении.	номинативная	Критерий Макнамары
			порядковая, интервальная	G - двухсторонний критерий знаков
			интервальная	T - двухсторонний критерий Вилкоксона

б) сдвиги под влиянием экспериментальных воздействий (одни и те же показатели, измеренные у одних и тех же испытуемых до и после воздействия: при отсутствии или при наличии контрольной группы).	два замера одного признака на одной и той же выборке	<p>Вариант 1</p> <p>H_0: результаты второго измерения изучаемого свойства у одних и тех же объектов - y_i имеют тенденцию быть меньше результатов первичного измерения - x_i.</p> <p>H_1: результаты второго измерения изучаемого свойства у одних и тех же объектов - y_i имеют тенденцию превышать результаты первичного измерения - x_i.</p> <p>Вариант 2</p> <p>H_0: результаты второго измерения изучаемого свойства у одних и тех же объектов - y_i имеют тенденцию быть больше результатов первичного измерения - x_i.</p> <p>H_1: результаты второго измерения изучаемого свойства у одних и тех же объектов - y_i имеют тенденцию быть меньше результатов первичного измерения - x_i.</p>	<p>порядковая, интервальная</p>	G - односторонний критерий знаков
			интервальная	T – односторонний критерий Вилкоксона
в) структурные сдвиги (разные показатели одних и тех же испытуемых, если они измерены в одних и тех же единицах, по одной той же шкале).	более двух замеров одного признака на одной и той же выборке	<p>H_0: увеличение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему и далее, случайно.</p> <p>H_1: увеличение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему и далее, неслучайно.</p>	порядковая, интервальная	L - критерий тенденций Пейджа
		<p>H_0: изменение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему и далее, случайно.</p> <p>H_1: изменение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему и далее, неслучайно.</p>	интервальная	χ^2_r - критерий Фридмана
			номинативная, порядковая, интервальная	ϕ^* - критерий Фишера

3. Выявление различий в распределении признака

Распределение могут различаться по средним, дисперсиям, асимметрии, эксцессу и по сочетанию данных параметров

сопоставление эмпирического распределения с теоретическим.	один замер одного признака на одной выборке	<p>H_0: полученное эмпирическое распределение не отличается от теоретического распределения.</p> <p>H_1: полученное эмпирическое распределение отличается от теоретического распределения.</p>	номинативная порядковая, интервальная	χ^2 - критерий Пирсона с поправкой на непрерывность m –биномиальный критерий
сопоставление эмпирических распределений одного и того же признака.	две независимые выборки одинаковой или различной численности	<p>H_0: эмпирическое распределение 1 не отличается от эмпирического распределения 2.</p> <p>H_1: эмпирическое распределение 1 отличается от эмпирического распределения 2.</p>	номинативная порядковая, интервальная	χ^2 - критерий Пирсона
			порядковая, интервальная	Медианный критерий λ - критерий Колмогорова-Смирнова
сопоставление трех или более эмпирических распределений одного и того же признака.	более двух независимых выборок одинаковой или различной численности	<p>H_0: эмпирические распределения не различаются между собой.</p> <p>H_1: эмпирические распределения различаются между собой</p>	номинативная порядковая, интервальная	φ^* - критерий Фишера

4. Выявление степени согласованности изменений

определение степени тесноты связи между качественными признаками, показателем которой является абсолютная величина коэффициента корреляции.	замеры двух признаков на одной и той же выборке	вариант 1	номинативная	коэффициенты ассоциации Д.Юла и контингенции К.Пирсона Коэффициенты взаимной сопряженности К.Пирсона и А.Чупрова
		<p>H_0: корреляция между переменными А и Б не отличается от нуля.</p> <p>H_1: корреляция между переменными А и Б достоверно отличается от нуля.</p>		
		вариант 2		
		<p>H_0: корреляция между иерархиями А и Б не отличается от нуля.</p> <p>H_1: корреляция между иерархиями А и Б достоверно отличается от нуля.</p>	порядковая	коэффициенты ранговой корреляции Спирмена, Кенделла и Гудмана
			интервальная	Коэффициент линейной корреляции К.Пирсона

Только при соблюдении всех выше перечисленных положений статистическая проверка гипотезы как элемент метода научного исследования позволяет объединить эмпирическое и теоретическое в педагогическом исследовании, т.е. сочетать в ходе изучения педагогического объекта прямое наблюдение, факты, эмпирический уровень исследования с построением логических конструкций и научных абстракций. Таким образом, практическая и познавательная ценность статистической проверки гипотезы определяется ее адекватностью изучаемым сторонам объекта, а также тем, насколько правильно выбран метод для его обоснования, т.е. насколько правильно построено педагогическое исследование.

Примечания:

1. Сосновский Б.А. Лабораторный практикум по общей психологии. М., 1979. С.18; С.137-138.
2. Сидоренко Е.В. Методы математической обработки в психологии. СПб.: Речь. С. 33-34 .
3. Толстова Ю.Н. Анализ социологических данных. М., 2000. С. 87, 95.
4. Наследов, А.Д. Математические методы психологического исследования. Анализ и интерпретация данных / А.Д. Наследов. – СПб.: Речь, 2006. – С. 111-122.
5. Созонова М.С. Математические методы в психологии. Тобольск: Изд-во ТГПИ, 2006. 172 с.
6. Тимошенко А.И. Математические методы исследования в психологии. Иркутск: Иркут. гос. ун-т, 2006. 207 с.
7. Гуревич К.М. Проблемы современной психологической диагностики // Психологическая диагностика: Проблемы и исследования. М.,1981. С. 14, 19.
8. Глас Дж., Стэтли Дж. Статистические методы в педагогике и психологии. М., 1976. С. 222, 270.
9. Суходольский, Г.В. Основы математической статистики для психологов. Л., 1972. С. 294.