Билятдинов К.З., Фортинский А.Г.

Применение многофункциональных статистических критериев при контроле качества обучения

При решении задач всестороннего контроля качества обучения возникает необходимость сопоставления двух (или более) рядов выборочных значений по частоте встречаемости какого-либо признака. В этом случае считается рациональным выбор и применение многофункциональных статистических критериев, которые помогут решить эту задачу по отношению к самым разнообразным данным, выборкам и задачам.

В частности, к числу многофункциональных критериев относится критерий φ* Фишера (угловое преобразование Фишера).

При использовании многофункциональных статистических критериев данные могут быть представлены в любой шкале [1, 2, 3,4].

Например, угловое преобразование Фишера может обоснованно выбираться и эффективно применяться в исследованиях, когда используется дихотомическая шкала измерений.

Дихотомическая шкала – это порядковая шкала с всего двумя различными упорядоченными баллами (или событиями) – "высокий"-"низкий", "справился с заданием"-"не справился", "прошел тест"-"не прошел" и т.д. Характеристикой данных выборок (групп), помимо общего числа ее членов, будет число членов (или доля, процент от общего числа), набравших заданный, например – максимальный, балл (в общем случае – число членов, обладающих заданным признаком) [3].

Выборки данных могут быть как независимыми, так и связанными по каким либо показателям и признакам между собой.

Критерий φ* Фишера целесообразно применять в следующих случаях:

1) для сравнения разных выборок данных;
2) для сравнения признаков (показателей) одной и той же выборки данных, измеренных в разных условиях;
3) для решения задачи сопоставления уровней исследуемого признака, сдвигов в значениях исследуемого признака и сравнения распределений.

При этом в критерии φ* Фишера верхней границы не существует, то есть выборки данных или массивы данных, которые необходимо сравнить, могут быть очень большими [4].

Кроме того, при использовании критерия вначале можно проверить, различаются ли выборки данных по уровню какого-либо признака, а затем сравнить распределения признака в двух выборках.

Такая задача может быть актуальной при анализе различий в диапазонах или форме распределения оценок, получаемых испытуемыми по какой-либо новой методике.

Однако, при выборе и применении критерия φ* Фишера необходимо учитывать следующие ограничения:

1. ни одна из сопоставляемых долей не должна быть равной нулю, в

этом случае результат сравнения будет достоверным;

2. нижний предел применения критерии φ* Фишера равен двум наблюдениям в одной из выборок, но должны соблюдаться следующие соотношения в численности двух выборок;

2.1. если в одной выборке данных всего 2 наблюдения (или события), то во второй должно быть не менее 30 наблюдений;

2.2. в случае если в одной из выборок три наблюдения, то во второй должно быть не менее семи;

2.3. если в одной из выборок всего 4 наблюдения, то во второй должно быть не менее 5;

2.4. в случае если в первой и второй выборках данных число наблюдений больше или равно пяти, то тогда возможны любые их сравнения.

Если не учитывать вышеприведенные ограничения, то в этих случаях не удастся выявить достоверных различий. Других ограничений у критерия Фишера нет [1, 4].

Многофункциональные критерии построены на сопоставлении долей, выраженных в долях единицы или в процентах.

Суть критериев состоит в определении того, какая доля наблюдений (событий, реакций, выборов, испытуемых) в данной выборке характеризуется интересующим исследователя эффектом и какая доля этим эффектом не характеризуется.

Таким эффектом может быть:

а) определенное значение качественно определяемого признака;

б) определенный уровень количественно измеряемого признака;

в) определенное соотношение значений или уровней исследуемого признака.

Итак, путем сведения любых данных к альтернативной шкале "Есть эффект - нет аффекта" многофункциональные критерии позволяют решать все три задачи сопоставлений - сравнения "уровней", оценки "сдвигов" и сравнения распределений.

Критерий φ* применяется в случаях, когда обследованы две выборки данных. В общем, назначение Критерий Фишера – это сопоставления двух выборок по частоте встречаемости интересующего исследователя эффекта.

Критерий оценивает достоверность различий между процентными долями двух выборок, в которых зарегистрирован интересующий нас эффект.

При использовании углового преобразования Фишера происходит перевод процентных долей в величины центрального угла, измеряемого в радианах.

Большей процентной доле будет соответствовать больший угол, а меньшей доле - меньший угол, но соотношения между ними не линейные.

В общем случае в литературе рекомендуется следующая последовательность расчета критерия Фишера [1, 2, 3,4]:

1. Определить значения признака, которые будут составляющих выборок на тех у кого есть эффект и нет эффекта;

2. Если признак измерен количественно, то следует использовать критерий Колмогорова-Смирнова для поиска оптимальной точки разделения, а если нет, то продолжить применение выбранного критерия Фишера.

3. Начертить таблицу 1;

Таблица 1.

Есть эффект. 1 выборка	Нет эффекта. 1 выборка
Есть эффект. 2 выборка	Нет эффекта. 2 выборка

4. Подсчитать количество значений, соответствующих ячейкам этой таблицы, и занести числа в таблицу. Сумма чисел по строкам должна совпадать с общим числом составляющих первой и второй выборки соответственно;

5. Подсчитать процентные доли значений имеющих эффект для первой и второй выборки путем деления содержимого ячеек левого столбца на объем соответствующей выборки;

Занести полученные значения в таблицу в соответствующие места (таблицу 2).

Таблица 2.

Есть эффект. 1 выборка (…%)	Нет эффекта. 1 выборка
Есть эффект. 2 выборка (…%)	Нет эффекта. 2 выборка

6. Проверить, не равняется ли одна из сопоставляемых долей нулю;

7. Определить из справочных таблиц или по формуле (1) значение угла для каждой из сопоставляемых процентных долей;

8. По формуле (2) рассчитать эмпирическое значение φ*;

9. Полученный результат по таблицам сравнить с критическим значением.

В качестве примера обоснованного выбора и эффективного применения углового преобразования Фишера можно привести вариант решения задачи сравнения двух рядов выборочных значений, полученных при оценке двух учебных групп (кафедр, подразделений), по частоте встречаемости заданного показателя.

Показателем могут быть различны: ответы на вопросы анкеты, результаты обучения, результаты тестирования и т.д.

Далее в данном примере результаты опроса – это ответы обучаемых на вопросы анкеты и (или) тестов.

В случае применения критерия Фишера сравниваемые подразделения будут выступать как выборки данных.

При условии, что численность различных подразделений, участвовавших в опросе, может быть не одинакова, то поэтому для достоверного сравнивания результатов опросов целесообразно выбрать многофункциональные статистические критерии.

В частности критерий φ^* Фишера, где, как уже было выше изложено, верхней границы не существует – выборки могут быть сколь угодно большими [1, 3, 4].

Это обстоятельство позволит использовать его при сравнении результатов опросов в подразделениях, где численность личного состава может достигать сотен сотрудников.

Итак, при сравнительном анализе результатов опросов использовался критерий Фишера (φ^*) для сопоставления двух выборок по частоте встречаемости интересующего исследователя ответа на заданный вопрос в анкете [4]. Критерий Фишера будет оценивать достоверность различий между процентными долями двух выборок, в которых зарегистрирован интересующий ответ на вопрос анкеты.

В данном примере происходит преобразование процентных долей в величины центрального угла, который измеряется в радианах (1).

Очевидно, что большей процентной доле ответов на интересующий исследователя вопрос анкеты, будет соответствовать больший угол φ, а меньшей доле – меньший угол, но соотношения здесь, как уже известно, не линейные:

$$\varphi = 2\arcsin(\sqrt{P}), \qquad (1)$$

где P – процентная доля, выраженная в долях единицы.

При увеличении расхождения между углами φ_1 и φ_2, и увеличения численности выборок значения критерия возрастает. Чем больше величина φ^*, тем более вероятно, что различия достоверны.

Отсюда целесообразно сформулировать следующие гипотезы:

H_0: Доля обучаемых, выбравших i-ответ на j-вопрос анкеты, в выборке n не больше, чем в выборке n+1.

H_1: Доля обучаемых, выбравших i-ответ на j-вопрос анкеты, в выборке n больше, чем в выборке n+1.

Эмпирическое значение φ^* вычислено по формуле (2):

$$\varphi^*_{эмп} = (\varphi_1 - \varphi_2) \bullet \sqrt{\frac{n_1 \bullet n_2}{n_1 + n_2}}, \qquad (2)$$

где φ_1 - угол, соответствующей большей процентной доле, φ_2 - угол, соответствующей меньшей процентной доле, n_1 - количество наблюдений в выборке 1, n_2 - количество наблюдений в выборке 2.

Далее сравнивалось полученное значение $\varphi^*_{эмп}$ с установленными критическими значениями φ^*, соответствующие принятым уровням статистической значимости $\varphi^*_{кр.нижн.} \leq 1{,}64 (p \leq 0{,}05)$ и $\varphi^*_{кр.верх.} \leq 2{,}31 (p \leq 0{,}01)$. Если значение $\varphi^*_{эмп} > \varphi^*_{кр.верх.}$, то H_0 отвергается и принимается H_1, если $\varphi^*_{эмп} < \varphi^*_{кр.нижн.}$, то H_0 принимается и отвергается H_1.

Если необходимо сравнить результаты опросов в более чем двух подразделениях, то тогда в общем случае процедура опроса и сравнительного анализа результатов опроса может заключатся в следующем: при помощи критерия φ^* Фишера поочередно сравнивались две группы опрошенных (или две выборки) для проверки достоверности ранее сформулированных гипотез: H_0 и H_1.

В результате последовательных парных сравнений имеющихся групп могут быть получены достоверные сведения в какой группе больше доля обучаемых, выбравших 1-й ответ на *j*-вопрос анкеты, затем в какой группе больше доля обучаемых, выбравших 2-й ответ на *j*-вопрос анкеты и так далее.

Таким образом, критерий Фишера позволяет эффективно решать подобного рода задачи.

Используемая литература:

1. Ермолаев О.Ю. Математическая статистика для психологов. М.: Флинта, 2002г.;
2. Закс Л. Статистическое оценивание.- М.: Статистика, 1976. – 595с.;
3. Новиков Д.А. Статистические методы в педагогических исследованиях (типовые случаи). М.: МЗ-Пресс, 2004. – 67 с.;
4. Сидоренко Е.В. Методы математической обработки в психологии, СПб., Речь, 2000 г.