

## ТЕХНОЛОГИЯ ОЦЕНКИ ВЗАИМОСВЯЗИ МЕЖДУ ПРИЗНАКАМИ НА НОМИНАЛЬНЫХ ШКАЛАХ

О.С. Кошевой, М.К. Карпова

ФГБОУ ВПО «Пензенский государственный университет»,  
г. Пенза

Рецензент д-р социол. наук, доцент В.П. Воробьев

**Ключевые слова и фразы:** критерий хи-квадрат; наблюдаемые и ожидаемые частоты; номинальная шкала; уровень значимости.

**Аннотация:** Представлена технология оценки взаимосвязи между независимыми признаками в таблице сопряженности на номинальных шкалах с использованием критерия хи-квадрат. На практическом примере представлена поэтапная реализация алгоритма принятия решения о наличии статистической взаимосвязи между двумя признаками.

В последнее время, при проведении социологического исследования, существенно возрастают требования к достоверности выводов и рекомендаций, полученных на основании выполненных исследований.

Одним из наиболее распространенных способов проведения эмпирического социологического исследования является анкетный опрос. При этом результаты анкетирования представляются в виде линейного, как правило, одномерного распределения частот ответов. Несмотря на то, что результаты одномерного анализа данных часто имеют самостоятельное значение, большинство исследователей уделяют основное внимание анализу связей между переменными [2, 7, 6]. В работе [5, с. 58] прямо указывается: линейное распределение может показывать, что данные группы респондентов действительно различаются (например, мужчин в выборке в два раза больше чем женщин). Однако визуального различия между категориями недостаточно для того, чтобы с уверенностью констатировать наличие статистического значимого различия.

Анализ инструментария, используемого в ходе проведения эмпирического социологического исследования с использованием анкетного опроса показывает, что большинство вопросов анкет представлены на номинальных шкалах. При этом для данного вида шкал, для установления статисти-

---

Кошевой Олег Сергеевич – доктор технических наук, профессор кафедры «Государственное управление и социология региона»; Карпова Маргарита Константиновна – кандидат социологических наук, доцент кафедры «Государственное управление и социология региона», e-mail: karpovamk@mail.ru, ФГБОУ ВПО «Пензенский государственный университет», г. Пенза.

ческой связи между переменными, наибольшее распространение получил критерий хи-квадрат  $\chi^2$ , определяемый по зависимости [3]

$$\chi^2 = \sum_{i=1}^n \frac{(Q_i - E_i)^2}{E_i}, \quad (1)$$

где  $Q_i$  – наблюдаемые частоты;  $E_i$  – ожидаемые частоты;  $n$  – число клеток в таблице сопряженности.

Ожидаемая частота наблюдений  $E_i$  в данной ячейке таблицы сопряженности равняется произведению общего количества наблюдений в строке и общего количества наблюдений в столбце деленному на общее количество наблюдений в таблице сопряженности [8].

Рассмотрим пример использования критерия  $\chi^2$  при анализе статистической связи студентов юношей и девушек по показателю мотивации к успеху. Исследование было проведено в 2010 году на базе Пензенского государственного университета среди третьих курсов специальностей «Государственное и муниципальное управление», «Менеджмент» и «Социология управления». При анкетировании использовалась методика диагностики личности на мотивацию к успеху Т. Элерса [4]. Результаты проведенного исследования представлены в виде одномерного линейного распределения (рис. 1).

Предварительный анализ данных, представленных на рис. 1, свидетельствует о том, что уровень мотивации к успеху у девушек гораздо выше, чем у юношей, однако этот вывод является весьма поверхностным, и связан с тем, что в опросе участвовало гораздо большее количество девушек, чем юношей.

Для формирования более достоверных выводов на первом этапе принятия решения о наличии (отсутствии) статистической связи между изучаемыми признаками формулируем нулевую и альтернативную гипотезы:  $H_0$ : Нет связи между признаками;  $H_1$ : Связь статистически значима.

Далее строим таблицы сопряженности для 57 респондентов с наблюдаемыми и ожидаемыми частотами (табл. 1).

Тогда, в соответствии с зависимостью (1) расчетное значение критерия  $\chi^2_p$  будет равно



**Рис. 1. Результаты обследования студентов:**  
 – пол женский;  – пол мужской

Таблица 1

## Уровень мотивации к успеху студентов

Мотивация к успеху	Пол респондентов		Итого
	женский	мужской	
Средняя	19,0/21,3	8,0/5,7	27,0/27,0
Умеренно высокая	16,0/15,0	3,0/4,0	19,0/19,0
Слишком высокая	10,0/8,7	1,0/2,3	11,0/11,0
Итого	45,0/45,0	12,0/12,0	57,0/57,0

Примечание: в числителе колонок наблюдаемые частоты, в знаменателе – ожидаемые частоты.

$$\chi_p = \frac{(19-21,3)^2}{21,3} + \frac{(8-5,7)^2}{5,7} + \frac{(16-15)^2}{15} + \frac{(3-4)^2}{4} + \frac{(10-8,7)^2}{8,7} + \frac{(1-2,3)^2}{2,3} = 2,42.$$

Для определения критического значения критерия  $\chi_{кр}^2$  необходимо рассчитать число степеней свободы  $df$  по формуле

$$df = (r-1)(c-1), \quad (2)$$

где  $r$  – число строк в таблице сопряженности;  $c$  – число колонок в таблице сопряженности.

$$df = (3-1)(2-1) = 2.$$

Критическое значение критерия  $\chi_{кр}^2$  можно рассчитать с использованием статистикой функции ХИ2ОБР табличного процессора MS Excel. Для условий рассматриваемого примера при уровне значимости 0,05  $\chi_{кр}^2 = 5,991$ .

Поскольку расчетное значение критерия  $\chi_p^2$  меньше критического значения критерия  $\chi_{кр}^2$ , то нулевая гипотеза  $H_0$  о независимости мотивации к успеху у юношей и девушек не отвергается и, следовательно, уровень мотивации к успеху не зависит от полового признака изучаемых переменных.

При использовании критерия  $\chi^2$  существует два ограничения [1]: во-первых, ожидаемые частоты  $< 5$  должны встречаться не более чем в 20 % полей таблицы сопряженности; во-вторых, суммы по строкам и столбцам всегда должны быть больше нуля.

Анализ данных табл. 1 показывает, что первое ограничение на применение критерия  $\chi^2$  не выполняется, так как существует 33,3 % ячеек, у которых ожидаемые частоты менее 5. Второе ограничение выполняется. Следовательно, к полученному выводу о независимости уровня мотивации от полового признака студентов следует относиться с определенной долей осторожности.

Достаточно удобно критерий хи-квадрат реализуется в среде компьютерной системы SPSS.

Для тестирования необходимо выполнить следующие действия [1]:

– в главном меню выбрать команды: *Analyze*→*Descriptive Statistics*→*Crosstabs*;

– переменную *Мотивация* перенести в список строк, а переменную *Пол* перенести в список столбцов;

– щелкнуть по кнопке *Cells* и в открывшемся диалоговом окне установить, кроме предлагаемого по умолчанию флажка *Observed*, флажок *Expected* и далее *Continue*;

– щелкнуть по кнопке *Statistics* и в открывшемся диалоговом окне установить флажок *Chi-square* и далее *Continue*;

– в главном диалоговом окне щелкнуть по кнопке *OK*.

В результате решения будет сформирована таблица сопряженности (табл. 2).

Сравнивая данные табл. 1 с данными табл. 2, видно полное совпадение наблюдаемых и ожидаемых частот.

В результате установлено, что расчетное значение критерия хи-квадрат составляет 2,459 и практически совпадает с засчитанным ранее 2,42. Небольшая разница в расчетах вызвана ошибками округления. В расчетах также показано, что асимптотическая значимость теста составляет 0,292. Учитывая, что эта величина намного превышает пороговый уровень значимости 0,05, можно сделать полученный ранее вывод о том, что уровень мотивации к успеху не зависит от полового признака изучаемых переменных.

Таким образом, представленная технология оценки статистической взаимосвязи между двумя независимыми признаками, измеренными на номинальной шкале, достаточно легко реализуется как аналитически

Таблица 2

**Кросс-таблица признаков «мотивация–пол»**

Мотивация к успеху		Пол		Сумма частот
		мужской	женский	
Средняя	наблюдаемые частоты	8,0	19,0	27,0
	ожидаемые частоты	5,7	21,3	27,0
Умеренно высокая	наблюдаемые частоты	3,0	16,0	19,0
	ожидаемые частоты	4,0	15,0	19,0
Слишком высокая	наблюдаемые частоты	1,0	10,0	11,0
	ожидаемые частоты	2,3	8,7	11,0
Сумма частот	наблюдаемые частоты	12,0	45,0	57,0
	ожидаемые частоты	12,0	45,0	57,0

с использованием простейших зависимостей, так и в среде компьютерной системы SPSS.

Полученные результаты позволяют более обоснованно сформировать выводы по исследуемому социальному явлению.

#### *Список литературы*

1. Бююль, А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей : пер. с нем. / А. Бююль, П. Цефель. – СПб. : ООО «ДиаСофтЮП», 2005. – 608 с.

2. Девятко, И.Ф. Методы социологического исследования / И.Ф. Девятко. – Екатеринбург : Изд-во Урал. ун-та, 1998. – 208 с.

3. Крыштановский, А.О. Анализ социологических данных с помощью пакета SPSS : учеб. пособие / А.О. Крыштановский ; отв. ред. Ю.Н. Толстова, А.В. Рыжова. – М. : Издат. дом гос. ун-та «Высшая школа экономики», 2006. – 283 с.

4. Методика диагностики личности на мотивацию к успеху Т. Элерса [Электронный ресурс]. – Режим доступа : <http://psylist.net/praktikum/23.htm> (дата обращения 21.09.12). – Загл. с экрана.

5. Таганов, Д.Н. SPSS: Статистический анализ в маркетинговых исследованиях / Д.Н. Таганов. – СПб. : Питер, 2005. – 192 с.

6. Татарова, Г.Г. Методология анализа данных в социологии (введение) : учеб. для вузов / Г.Г. Татарова. – М. : NOTA BENE, 1999. – 224 с.

7. Толстова, Ю.Н. Анализ социологических данных. Методология, дескриптивная статистика, изучение связей между номинальными признаками : учеб. пособие для вузов / Ю.Н. Толстова. – М. : Научный мир, 2000. – 352 с.

8. Хили, Дж. Статистика. Социологические и маркетинговые исследования / Дж. Хили ; пер. с англ. под общ. ред. А.А. Руденко. – 6-е изд. – Киев : ООО «ДиаСофтЮП» ; СПб. : Питер, 2005. – 638 с.

---

### **Technology of Assessing Attributes Relation on Nominal Scales**

**O.S. Koshevoy, M.K. Karpova**

*Penza State University, Penza*

**Key words and phrases:** chi-squared test; nominal scale; observed and expected frequencies; significance point.

**Abstract:** The paper describes the technology of assessing independent attributes relation in contingency table on nominal scales using chi-squared test. The authors illustrate the sequential implementation of decision-making algorithm on statistical relation between two attributes.

---

© О.С. Кошевой, М.К. Карпова, 2012