УДК 519.25

Т.А. Юрьева, А.П. Филимонова, Н.А. Чалкина

СТАТИСТИЧЕСКАЯ ОЦЕНКА СВЯЗИ МЕЖДУ КАЧЕСТВЕННЫМИ ПРИЗНАКАМИ В ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

Проникновение математики в педагогическую науку неизбежно сопровождается рядом трудностей, обусловленных в первую очередь сложностью и трудоемкостью вычислительных процедур. Развитие статистических методов анализа качественных переменных позволяет в значительной степени облегчить вычисления. В статье рассматриваются способы статистического анализа таблиц сопряженности и иллюстрируются на примерах педагогического содержания.

Ключевые слова: номинальная шкала, качественные переменные, таблица сопряженности, статистические критерии.

STATISTICAL EVALUATION OF THE LINK BETWEEN QUALITY CHARACTERISTICS IN EDUCATIONAL RESEARCH

Use of mathematics in pedagogical science is inevitably accompanied by a number of difficulties in the first place due to the complexity and labor-intensive computational procedures. Development of statistical methods for analyzing qualitative variables would greatly facilitate the calculations. The article discusses a reception of statistical analysis of contingency tables and illustrated by examples of pedagogical content.

Key words: nominal scale, qualitative variables, contingency table, statistical criteria.

По мере интеграции педагогики с другими науками особое значение приобретает постепенный переход от качественного описания явлений и процессов педагогики к их количественному моделированию. Вместе с тем следует отметить, что в силу особенностей природы и сложности характера педагогических явлений применение методов математической статистики в педагогике ограниченно. Классический аппарат статистических методов подчас оказывается не приспособленным для количественного измерения результатов педагогических исследований.

В то же время развитие статистической методологии позволяет говорить о правомерности ее грамотного применения и в педагогических исследованиях. Методы анализа не только количественных, но и качественных данных используются во многих современных педагогических исследованиях ([1], [2], [3] и др.). В одном исследовании могут встречаться данные, измеренные в качественных и количественных шкалах. Например, в работе Н.Н. Двоерядкиной [4] для изучения качества математической подготовки студентов применялись следующие показатели: мотивация (ранговая шкала), знания и умения (шкала отношений), способность к исследовательской деятельности (дихотомическая шкала). Совместная статистическая обработка признаков, измеренных в разных шкалах, в большинстве случаев вынуждает преобразовывать переменные к номинальной шкале (шкале наименований).

Построение номинальной шкалы, ее использование при измерении некоторого свойства рассматриваемых объектов возможно, если установлен критерий, позволяющий по состоянию измеряемого свойства распределить эти объекты на несколько классов, причем каждый объект должен попасть только в один класс. Число «делений» в шкале наименований соответствует числу классов.

Несмотря на кажущуюся примитивность шкалы наименований, измерения по этой шкале могут быть использованы для проверки некоторых статистических гипотез и для вычисления показателей корреляции качественных признаков. Одним из инструментов, позволяющих проводить анализ связей между двумя и более качественными переменными, является таблица сопряженности.

Простейшая таблица сопряженности 2x2 (табл. 1) содержит распределение частот двух качественных признаков, измеренных в номинальной шкале с двумя «делениями» (дихотомическая шкала).

Таблица 1

Общий вид таблицы сопряженности 2x2

	Признак *B*	
Признак *A*	*A*	*b*
	c	*d*

Здесь a – число элементов выборки, обладающих признаками A и B одновременно; b – число элементов выборки, обладающих признаком A, но не обладающих признаком B; c – число элементов выборки, обладающих признаком B, но не обладающих признаком A; d – число элементов выборки, не обладающих ни одним из признаков A и B.

Для установления силы связи между признаками A и B можно использовать коэффициент контингенции

$$V = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}.$$

Для проверки гипотезы о статистической значимости взаимосвязи между изучаемыми признаками используется величина $\chi^2 = nV^2$, имеющая при отсутствии связи распределение χ^2 с f=1 степенью свободы. С учетом поправки на непрерывность статистика критерия контингенции для проверки связи признаков имеет вид $\chi^2 = \dfrac{n\left(|ad - bc| - \frac{n}{2}\right)^2}{(a+b)(a+c)(b+d)(c+d)}$, где $n=a+b+c+d$.

Если $\chi^2 > \chi^2_\alpha(1)$, то с достоверностью α зависимость признаков A и B признается значимой. Этот критерий применим при $n \geq 40$ и a, b, c, $d \geq 5$. Если эти условия не выполняются, то следует воспользоваться точным критерием Фишера, основанном на статистике

$$p = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{(a+b+c+d)!} \sum_{i=0}^{a} \frac{1}{(a+b-i)!(a+c-i)!(a+d-i)!}$$

Если $p > 1-\alpha$, то с вероятностью α связь признается значимой (α=0,95).

Пример 1. Допустим, нас интересует, зависит ли успешность прохождения студентами второго курса Интернет-тестирования по математике от прохождения ими адаптивного курса в начале обучения в вузе. В группе из 20 человек, посещавших адаптивный курс, успешно справились с тестированием 12 человек, а в группе из 25 человек, не посещавших адаптивный курс, – 10 человек. В нашем случае признак A – посещение адаптивного курса, B – успешность прохождения Интернет-тестирования.

n=45, a=12, b=8, c=10, d=15.

Коэффициент контингенции $V = \dfrac{12 \cdot 15 - 8 \cdot 10}{\sqrt{(12+8)(12+10)(8+15)(10+15)}}$ =0,199.

Статистика критерия контингенции $\chi^2 = \dfrac{45\left(12 \cdot 15 - \dfrac{45}{2}\right)^2}{(12+8)(12+10)(8+15)(10+15)}$ =4,412.

Так как $\chi^2 > \chi^2_{0,95}(1)$ =3,84, то связь, несмотря на ее слабость, статистически значима.

Наиболее распространенной в педагогических исследованиях является проблема установления статистической значимости различий в распределении частот в одной и той же группе объектов до и после применения контролируемых воздействий. В этом случае также можно составить таблицу сопряженности 2х2, но составляющие ее данные уже не будут независимыми. В этом случае *a* – число объектов выборки с положительным эффектом до и после воздействия, *d* – число объектов выборки с отрицательным эффектом до и после воздействия, *b* – число объектов выборки с положительным эффектом до воздействия и отрицательным эффектом после, *c* – наоборот.

Для проверки статистической значимости равенства *b*=*c* используется критерий Макнамары $\chi^2 = \dfrac{(b-c)^2}{b+c+1}$, а при *b*+*c*<30 $\chi^2 = \dfrac{(|b-c|-1)^2}{b+c+1}$. Если $\chi^2 > \chi^2_\alpha(1)$, разница между *b* и *c* признается значимой ($\chi^2_\alpha(1)$ – α-квантиль распределения хи-квадрат с *f*=1 степенью свободы).

Пример 2. Изучалось влияние олимпиады в Сочи на развитие патриотизма среди российских подростков. Для этого до и после ее проведения 120 подросткам был задан вопрос: «Гордитесь ли вы российским флагом?». В выборке из 10 подростков положительный ответ до и отрицательный ответ после олимпиады дал 1 человек, положительный ответ после и отрицательный ответ до – 9 человек. В этом случае *b*=1, *c*=9, $\chi^2 = \dfrac{(|1-9|-1)^2}{1+9+1}$ =4,45. Сравнивая с $\chi^2_{0,95}(1)$ =3,84, делаем вывод о статистически значимом влиянии олимпиады на развитие патриотизма.

Помимо рассмотренных методов, анализ таблиц сопряженности 2х2 можно проводить с помощью критерия Вульфа, коэффициента коллигации Юла, коэффициента ассоциации [5]. Последние два позволяют грубо оценить наличие связи.

Для проверки гипотезы о статистической неразличимости двух таблиц сопряженности 2х2 используется критерий Ле Роя:

$$R = (a_1+b_1+c_1+d_1+a_2+b_2+c_2+d_2) \cdot \left(\frac{1}{a_2+b_2+c_2+d_2}\left(\frac{a_2^2}{a_1+a_2}+\frac{b_2^2}{b_1+b_2}+\frac{c_2^2}{c_1+c_2}+\frac{d_2^2}{d_1+d_2}\right) + \right.$$
$$\left. + \frac{1}{a_1+b_1+c_1+d_1}\left(\frac{a_1^2}{a_1+a_2}+\frac{b_1^2}{b_1+b_2}+\frac{c_1^2}{c_1+c_2}+\frac{d_1^2}{d_1+d_2}\right) - 1 \right).$$

Индекс 1 показывает принадлежность первой таблице сопряженности, 2 – второй. Гипотеза о статистической неразличимости таблиц отклоняется с достоверностью *α*, если R> $\chi^2_\alpha(3)$, ($\chi^2_\alpha(3)$ – α-квантиль распределения хи-квадрат с *f*=3 степенями свободы). *R*-критерий применим, если все числа в таблицах превышают 3.

Пример 3. Необходимость использования различных методик преподавания математики на экономическом и энергетическом факультетах оценивалась с помощью тестирования. При этом были сформированы две выборки из студентов экономического и энергетического факультетов. Результаты тестирования после применения разработанных методик обучения приведены в табл. 2 и 3.

Таблица 2

Экономический факультет

Результаты тестирования	Методика 1	Методика 2
Справились	8	15
Не справились	11	11

Таблица 3

Энергетический факультет

Результаты тестирования	Методика 1	Методика 2
Справились	10	20
Не справились	9	5

Рассчитаем эмпирическое значение критерия Ле Роя:

$$R=(8+15+11+11+10+20+9+5)\cdot\left(\frac{1}{10+20+9+5}\left(\frac{10^2}{8+10}+\frac{20^2}{15+20}+\frac{9^2}{11+9}+\frac{5^2}{11+5}\right)+\right.$$

$$\left.+\frac{1}{8+15+11+11}\left(\frac{8^2}{8+10}+\frac{15^2}{15+20}+\frac{11^2}{11+9}+\frac{11^2}{11+5}\right)-1\right)=3{,}376.$$

Так как $\chi^2_{0,95}(3)$ =7,815, то студентов экономического и энергетического факультета можно рассматривать как выборки одной совокупности. Таким образом, разработанные методики не дают разных результатов в тестировании на разных факультетах.

Если результаты наблюдений могут быть классифицированы по трем или более качественным признакам, рассматриваются так называемые таблицы сопряженности *r*х*c* (табл. 4).

Таблица 4

Общий вид таблицы сопряженности *rxc*

	Категории признака *B*					
Категории признака *A*	1	2	…	*j*	…	*c*
1	n_{11}	n_{12}	…	n_{1j}	…	n_{1c}
2	n_{21}	n_{22}	…	n_{2j}	…	n_{2c}
…	…	…	…	…	…	…
i	n_{i1}	n_{i2}	…	n_{ij}	…	n_{ic}
…	…	…	…	…	…	…
R	n_{r1}	n_{r2}	…	n_{rj}	…	n_{rc}

Здесь n_{ij} – эмпирическая частота.

Для проверки гипотезы о взаимосвязи изучаемых признаков используется критерий Пирсона: $\chi^2=\sum_{i=1}^{k}\frac{(f_э-f_m)^2}{f_m}$, где $k=rc$; f_m – теоретическая частота; $f_э$ – эмпирическая частота.

Для каждой ячейки таблицы соответствующая теоретическая частота рассчитывается по формуле: $f_{mij}=\frac{\left(\sum_{j=1}^{c}n_{ij}\right)\cdot\left(\sum_{i=1}^{r}n_{ij}\right)}{\sum_{i=1}^{r}\sum_{j=1}^{c}n_{ij}}$.

Если $\chi^2>\chi^2_\alpha(f)$, то взаимосвязь признается статистически значимой (число степеней свободы *f=(r-1)(c-1)*.

Статистика хи-квадрат неудобна при оценке связи признаков, так как ее значения не нормированы. Поэтому используется коэффициент сопряженности Чупрова: $K_R = \sqrt{\dfrac{\chi^2}{n\sqrt{(r-1)(c-1)}}}$.

<u>Пример 4</u>. По данным диагностики темперамента у 93 подростков с помощью опросника Г. Айзенка и диагностики профессиональных предпочтений с помощью опросника ДДО Климова определяли, влияет ли тип темперамента на профессиональные предпочтения. Результаты исследования представлены в табл. 5.

Таблица 5

Профессиональные предпочтения		Человек – техника	Человек – знаковая система	Человек – человек	Всего
Тип темперамента	№ столбца / № строки	1	2	3	-
Холерик	1	1	5	15	21
Сангвиник	2	13	5	9	27
Меланхолик	3	2	17	3	22
Флегматик	4	16	3	4	23
Итого	-	32	30	31	93

Вычислим сначала теоретические частоты:

f_{m11}=(21*32)/93=7,2; f_{m12}=(21*30)/93=6,8; f_{m13}=(21*31)/93=7; f_{m21}=(27*32)/93=9,3; f_{m22}=(27*30)/93=8,7; f_{m23}=(27*31)/93=9; f_{m31}=(22*32)/93=7,6; f_{m32}=(22*30)/93=7,1; f_{m33}=(22*31)/93=7,3; f_{m41}=(23*32)/93=7,9; f_{m42}=(23*30)/93=7,4; f_{m43}=(23*31)/93=7,7.

Дальнейшие расчеты оформим в табл. 6.

Таблица 6

Альтернативы	$f_э$	f_m	$f_э - f_m$	$(f_э - f_m)^2$	$\dfrac{(f_э - f_m)^2}{f_m}$
n_{11}	1	7,2	-6,2	38,44	5,34
n_{12}	5	6,8	-1,8	3,24	0,48
n_{13}	15	7	8	64	9,14
n_{21}	13	9,3	3,7	13,69	1,47
n_{22}	5	8,7	-3,7	13,69	1,57
n_{23}	9	9	0	0	0
n_{31}	2	7,6	-5,6	31,36	4,13
n_{32}	17	7,1	9,9	98,01	13,80
n_{33}	3	7,3	-4,3	18,49	2,53
n_{41}	16	7,9	8,1	65,61	8,31
n_{42}	3	7,4	-4,4	19,36	2,62
n_{43}	4	7,7	-3,7	13,69	1,78
Суммы	93	93	0	-	$\chi^2_{эмп}$ =51,17

Число степеней свободы будет равно f *=(4-1)*(3-1)=6*. $\chi^2_{0,99}(6)$=16,812. Следовательно, тип темперамента влияет на профессиональные предпочтения.

Коэффициент сопряженности Чупрова $K_R = \sqrt{\dfrac{51{,}17}{93\sqrt{(4-1)(3-1)}}}$ =0,47 показывает умеренную связь между исследуемыми признаками.

Представленный обзор методов статистической оценки связи между качественными признаками не претендует на полноту. Мы лишь попытались проиллюстрировать возможность и простоту их применения в педагогических исследованиях.

---

1. Лебедь, О.А. Методологические проблемы обучения прикладной статистике бакалавров гуманитарных направлений / Т.А. Макарчук, Т.А. Юрьева, О.А. Лебедь // Гуманизация образования. – 2009. – № 5. – С. 28-33.
2. Макарчук, Т.А. Методология и методика обучения прикладной статистике бакалавров гуманитарных направлений подготовки: Монография / Т.А. Макарчук, Н.Н. Двоерядкина, Н.А. Чалкина. – Владивосток: Дальнаука, 2010. – 152 с.
3. Юрьева, Т.А. Профессиональная направленность обучения математическим дисциплинам психологов в вузе /Т.А. Юрьева, А.П. Филимонова, С.В. Костенко // Международный научно-исследовательский журнал. Екатеринбург. – 2013.– Ч. 2. – №1(8) – С.46-47.
4. Двоерядкина, Н.Н., Макарчук, Т.А. Содержательное обеспечение становления математико-статистического мышления социологов // Информатика и системы управления. – 2005. – № 1 (09). – С. 162-168.
5. Кобзарь, А.И. Прикладная математическая статистика. Для инженеров и научных работников. – М.: Физматлит, 2006. – 816 с.