УДК 519.235

ВЫЯВЛЕНИЕ НЕЗАВИСИМОСТИ МЕЖДУ КАЧЕСТВЕННЫМИ ПЕРЕМЕННЫМИ

О. А. Бакаева, В. Н. Щенников

Мордовский государственный университет им. Н. П.Огарева

Поступила в редакцию 29.04.2011 г.

Аннотация. В данной статье формулируются теоремы и приводятся способы определения независимости качественных переменных, в основе которых лежат различные соотношения частот, а также введен коэффициент, характеризующий степень независимости. Рассмотрены примеры, иллюстрирующие применение этих методов при обработке результатов лечения вируса гриппа A(H1N1).
Ключевые слова: Таблицы сопряженности, зависимость, независимость, обработка данных, имеющих качественную природу, частоты, коэффициент независимости, вирус гриппа A(H1N1).

Abstract. In this article theorems are formulated and ways of definition of independence of qualitative variables in which basis various parities of frequencies lie are resulted, and also the factor characterizing degree of independence is entered. The examples illustrating application of these methods at processing of results of treatment of a virus of flu A(H1N1) are considered.
Key words: Cross-tabulation tables, dependence, independence, the data processing, having the qualitative nature, frequencies, independence factor, a virus of flu A(H1N1).

Основной целью выявления связи между переменными независимо от их природы, количественной или качественной, является прогнозное значение одной переменной по известному значению другой. Поэтому первый вопрос, который возникает при работе с таблицей сопряженности и относящийся к переменным A и B: «Зависимы ли признаки A и B друг от друга?» Если ответ на него положительный, то возникают следующие вопросы: «Каким образом связаны эти переменные?» и «Как оценить силу их связи?». Ответить на все эти вопросы, если данные имеют качественную природу, помогает аппарат таблиц сопряженности.

Таблицы сопряженности — это инструмент, позволяющий проводить анализ связей между категоризованными переменными (двумя и более). Данный способ сводится к построению таблиц, которые отражают совместное распределение двух и более переменных, обладающих ограниченным количеством категорий или принимающих определенные значения. Категории одной переменной помещают в таблицу так, чтобы они размещались в ней в соответствии с категориями другой (или несколькими другими) переменной. Таким образом, распределение частот одной переменной подразделяется на группы в зависимости от категорий других переменных. Обычно таблица совместного распределения частот для двух категориальных или дискретных переменных строится так: строки соответствуют значениям одной переменной, а столбцы — другой.

Простейший вид таблицы сопряженности — это таблица сопряженности (матрица) размерности 2×2, в которой значения двух дихотомических переменных "пересечены" (сопряжены) на разных уровнях и каждая переменная принимает только два значения, соответственно A_1, A_2 и B_1, B_2. Тогда таблица сопряженности будет иметь вид

Таблица 1

Частоты 2×2

	B_1	B_2	Всего
A_1	f_{11}	f_{12}	f_{10}
A_2	f_{21}	f_{22}	f_{20}
Всего	f_{01}	f_{02}	f_{00}

где f_{ij} — это наблюдаемая частота респондентов или каких-то определенных признаков, попавших в ячейку $(A_i,\ B_j)$. Символ f_{i0} означает сумму всех значений признака, попавших в категорию A_i, и аналогично f_{0j} — сумму значений сопряженного ему признака, относящихся к категории B_j, а f_{00} — это общий итог всех рассмотренных случаев. Математически это означает:

$$f_{i0} = \sum_j f_{ij}, \quad f_{0j} = \sum_i f_{ij},$$
$$f_{00} = \sum_i f_{i0} = \sum_j f_{0j} = \sum_i \sum_j f_{ij}.$$

Таким образом, исходя из некоторых соотношений значений частот признаков, можно сделать вывод о зависимости или независимости между ними. Обычно при исследовании переменных задача состоит в выявлении зависимости между ними. Но иногда выявить зависимость и тем более оценить силу связи – это более сложная и трудоемкая задача, чем доказать обратное, т.е. что признаки независимы. Если между переменными связи нет, то это означает, что знание об отношении объекта исследования к категории A не дает никакой информации относительно категории B. Хотя довольно трудно найти переменные, которые были бы совершенно независимы, поэтому сама идея независимости (отсутствия связи) не менее важна, чем идея зависимости между переменными. Математически это означает, что если A и B независимы, то должно ожидать, что отношение $\frac{f_{11}}{f_{01}}$ должно быть приближенно равно $\frac{f_{12}}{f_{02}}$, а $\frac{f_{11}}{f_{10}}$ будет приближенно равно $\frac{f_{21}}{f_{20}}$ [1, с. 18]. Как видно, в данном случае для определения независимости были использованы только частоты, находящиеся в первой строке (категория A_1 признака A), и частоты, стоящие в первом столбце (категория B_1 признака B). Получается, что «работают» только частоты f_{11}, f_{12} и f_{21}, тогда как значение частоты f_{22} в правой нижней ячейке вообще не задействовано. Если же переменные независимы, то, как уже отмечалось выше, знание об отношении объекта исследования к категории A не дает никакой информации относительно категории B, поэтому для выявления зависимости или независимости можно использовать как первые категории, так и вторые. Следовательно, если A и B независимы, то должно ожидать, что отношение $\frac{f_{21}}{f_{01}}$ должно быть приближенно равно $\frac{f_{22}}{f_{02}}$, а $\frac{f_{12}}{f_{10}}$ будет приближенно равно $\frac{f_{22}}{f_{20}}$. Тогда значение частоты f_{22} будет учтено, но, по аналогии с первым случаем, потеряется значение f_{11}.

Определение 1. Переменные A (A_1; A_2) и B (B_1; B_2) называются независимыми, если одновременно выполняется хотя бы одна из пар следующих равенств:

1) $\frac{f_{11}}{f_{01}} = \frac{f_{12}}{f_{02}}, \quad \frac{f_{11}}{f_{10}} = \frac{f_{21}}{f_{20}}$;

2) $\frac{f_{21}}{f_{01}} = \frac{f_{22}}{f_{02}}, \quad \frac{f_{12}}{f_{10}} = \frac{f_{22}}{f_{20}}$.

Определение 2. Переменные A и B называются нестрого независимыми, если одновременно выполняется хотя бы одна из пар следующих приближенных равенств:

1) $\frac{f_{11}}{f_{01}} \approx \frac{f_{12}}{f_{02}}, \quad \frac{f_{11}}{f_{10}} \approx \frac{f_{21}}{f_{20}}$;

2) $\frac{f_{21}}{f_{01}} \approx \frac{f_{22}}{f_{02}}, \quad \frac{f_{12}}{f_{10}} \approx \frac{f_{22}}{f_{20}}$.

Теорема 1 (необходимое условие независимости).

Если переменные A (A_1; A_2) и B (B_1; B_2) независимы, то должно выполняться хотя бы одно из следующих равенств:

$$\frac{f_{01}}{f_{02}} f_{12} = \frac{f_{10}}{f_{20}} f_{21}, \tag{1}$$

$$\frac{f_{02}}{f_{01}} \cdot f_{21} = \frac{f_{20}}{f_{10}} \cdot f_{12}. \tag{2}$$

Доказательство. Пусть переменные A и B независимы. Докажем выполнение равенства (1). По определению 1, если переменные независимы, то одновременно должны выполняться два условия $\frac{f_{11}}{f_{01}} = \frac{f_{12}}{f_{02}}$ и $\frac{f_{11}}{f_{10}} = \frac{f_{21}}{f_{20}}$. Выразим f_{11} из первого и второго равенств и приравняем эти два выражения друг другу, получим:

$$f_{11} = \frac{f_{01} \cdot f_{12}}{f_{02}} = \frac{f_{01}}{f_{02}} \cdot f_{12},$$
$$f_{11} = \frac{f_{10} \cdot f_{21}}{f_{20}} = \frac{f_{10}}{f_{20}} \cdot f_{21}.$$

Следовательно, если A и B независимы, то

$$\frac{f_{01}}{f_{02}} f_{12} = \frac{f_{10}}{f_{20}} f_{21}.$$

Далее докажем независимость переменных, если выполняется условие (2). Опять же по Определению 1, если переменные независимы, то одновременно должны выполняться условия: $\frac{f_{21}}{f_{01}} = \frac{f_{22}}{f_{02}}$, $\frac{f_{12}}{f_{10}} = \frac{f_{22}}{f_{20}}$ (рассматриваем вторую пару равенств). Если выразить f_{22} из первого и второго равенств и затем приравнять эти два выражения друг другу, то будем иметь:

$$f_{22} = \frac{f_{02} \cdot f_{21}}{f_{01}} = \frac{f_{02}}{f_{01}} \cdot f_{21},$$

$$f_{22} = \frac{f_{20} \cdot f_{12}}{f_{10}} = \frac{f_{20}}{f_{10}} \cdot f_{12}.$$

Приравнивая правые части, получим

$$\frac{f_{02}}{f_{01}} \cdot f_{21} = \frac{f_{20}}{f_{10}} \cdot f_{12}.$$

Теорема доказана. ▲

Теорема 1 о необходимом условии независимости получена на основе определения 1, т.е. равенства между отношениями частот. Выясним, какие необходимые условия должны выполняться, чтобы нестрогая независимость перешла в независимость. Для этого, во-первых, необходимо ввести коэффициент независимости, т.е. константу, которая будет приближенные равенства отношений частот сводить к точным.

Определение 3. Константа $c \geq 0$, характеризующая степень независимости переменных A и B, называется *коэффициентом независимости* ($c \in R^{+}$).

Теорема 2 (необходимое условие нестрогой независимости).

Если переменные A (A_1; A_2) и B (B_1; B_2) нестрого независимы, то должно выполняться хотя бы одно из следующих равенств:

$$\frac{f_{01}}{f_{02}} f_{12} = c_1 \frac{f_{10}}{f_{20}} f_{21}, \tag{3}$$

$$\frac{f_{02}}{f_{01}} \cdot f_{21} = c_2 \frac{f_{20}}{f_{10}} \cdot f_{12}, \tag{4}$$

где c_1 и c_2 — коэффициенты независимости.

Доказательство. Пусть переменные A и B независимы. Докажем выполнение равенства (3). Если из первой пары приближенных равенств Определения 2 выразить f_{11} и затем приравнять эти два выражения друг другу, то получится:

$$f_{11} \approx \frac{f_{01}}{f_{02}} \cdot f_{12}, \quad f_{11} \approx \frac{f_{10}}{f_{20}} \cdot f_{21}.$$

Следовательно, если A и B нестрого независимы, то $\frac{f_{01}}{f_{02}} f_{12} \approx \frac{f_{10}}{f_{20}} f_{21}$. Чтобы это нестрогое равенство преобразовать в строгое, введем константу c_1. Получим $\frac{f_{01}}{f_{02}} f_{12} = c_1 \frac{f_{10}}{f_{20}} f_{21}$.

Доказательство нестрогой независимости в случае выполнения условия (4) будет аналогичным. ▲

Нетрудно заметить, что выражения (1) и (2) можно получить из выражений (3) и (4), если приравнять c_1 и c_2 к 1.

Утверждение 1. Если переменные A и B независимы, то $c_1 = 1$.

Чтобы оценить на практике, насколько независимы переменные, выведем формулу для расчета коэффициента независимости. Выразив c_1 из равенства (3), получим:

$$c_1 = \frac{f_{01}\, f_{12}\, f_{20}}{f_{10}\, f_{21}\, f_{02}}. \tag{5}$$

Учитывая, что частоты в ячейках либо положительные числа, либо 0, значение $c_1 \in R^{+}$. Формула (5) получена при рассмотрении двух первых равенств в Определении 1.

В Утверждении 1 говорится только о равенстве единице коэффициента c_1. Но нетрудно заметить, что если переменные независимы, то и $c_2 = 1$. Аналогично предыдущему случаю, выразим c_2 из равенства (4):

$$c_2 = \frac{f_{10}\, f_{21}\, f_{02}}{f_{01}\, f_{12}\, f_{20}}. \tag{6}$$

Данные рассуждения доказывают, что первая и вторая строки значений равносильны, а также равносильны первый и второй столбцы, поэтому в формулировках первой и второй теоремы достаточно рассмотреть только первую пару равенств и указать значение только коэффициента c_1.

Замечание 1. Нетрудно заметить, что между коэффициентами независимости c_1 и c_2 существует связь, причем обратно пропорциональная, т.е. $c_1 \cdot c_2 = 1$.

Данное замечание подтверждает тот факт, что если переменные A и B независимы, то и $c_1 = 1$ и $c_2 = 1$. Поэтому далее достаточно рассматривать только свойства коэффициента c_1, для c_2 они аналогичны, но с учетом обратной пропорциональности.

Замечание 2. Как видно, коэффициент, характеризующий степень независимости между переменными, представляет собой отношение, из чего следует, что может возникнуть ряд трудностей для его нахождения и затем последующей оценки силы независимости. Поэтому целесообразно описать его вычисление и использование для некоторых особых случаев, к которым можно отнести нулевое значение частоты в одной из ячеек в таблице сопряженности. Рассмотрим 4 таких возможных случая:

1) Если $f_{11} = 0$, то $f_{10} = f_{12}$, $f_{01} = f_{21}$ и тогда $c_1 = \dfrac{f_{20}}{f_{02}}$.

2) Если $f_{12} = 0$, то $f_{10} = f_{11}$, $f_{02} = f_{22}$ и тогда числитель в выражении (5) обращается в 0. Однако это не значит, что переменные абсолютно зависимы. В этом случае для того чтобы получить значение коэффициента независимости, необходимо поменять местами столбцы в таблице (матрице) сопряженности, т.е. категории B_1 и B_2 между собой.

3) Если $f_{21} = 0$, то $f_{20} = f_{22}$, $f_{01} = f_{11}$ и тогда знаменатель выражения (5) обращается в 0, и дробь теряет смысл. Поэтому следует, как и в предыдущем случае, поменять местами столбцы в матрице сопряженности.

4) Если $f_{22} = 0$, то $f_{20} = f_{21}$, $f_{02} = f_{12}$ и тогда $c_1 = \dfrac{f_{01}}{f_{10}}$.

Как доказано выше, если переменные A (A_1; A_2) и B (B_1; B_2) независимы, то коэффициент независимости равен 1. Далее будет целесообразным рассмотреть случаи, когда такое возможно.

Теорема 3. Если $f_{12} = f_{21}$, то коэффициент независимости $c_1 = 1$.

Доказательство: Пусть выполняется условие $f_{12} = f_{21}$. Тогда $c_1 = \dfrac{f_{01}\, f_{12}\, f_{20}}{f_{10}\, f_{21}\, f_{02}} = \dfrac{f_{01} f_{20}}{f_{10} f_{02}} =$

$$= \frac{(f_{11}+f_{21})(f_{21}+f_{22})}{(f_{11}+f_{12})(f_{12}+f_{22})} = \frac{(f_{11}+f_{12})(f_{12}+f_{22})}{(f_{11}+f_{12})(f_{12}+f_{22})} = 1.$$

Доказательство завершено ▲. Рассмотрим еще один случай, когда коэффициент $c_1 = 1$.

Теорема 4. Коэффициент независимости $c_1 = 1$, тогда и только тогда, когда справедливо условие $f_{11}\, f_{22} = f_{12}\, f_{21}$.

Доказательство: 1) Если $c_1 = 1$, то справедливо $f_{01} f_{12} f_{20} = f_{10} f_{21} f_{02}$. Представим в виде суммы маргинальные частоты по строкам и столбцам.

$$(f_{11}+f_{21}) f_{12} (f_{21}+f_{22}) = (f_{11}+f_{12}) f_{21} (f_{12}+f_{22}).$$

Разделив полученное равенство на $f_{12} \cdot f_{21}$, имеем:

$$\left(\frac{f_{11}+f_{21}}{f_{21}}\right)(f_{21}+f_{22}) = (f_{11}+f_{12})\left(\frac{f_{12}+f_{22}}{f_{12}}\right),$$

$$\left(\frac{f_{11}}{f_{21}}+1\right)(f_{21}+f_{22}) = (f_{11}+f_{12})\left(1+\frac{f_{22}}{f_{12}}\right),$$

$$f_{11} + \frac{f_{11}f_{22}}{f_{21}} + f_{21} + f_{22} = f_{11} + f_{12} + \frac{f_{11}f_{22}}{f_{12}} + f_{22},$$

$$\frac{f_{11}f_{22}}{f_{21}} + f_{21} = f_{12} + \frac{f_{11}f_{22}}{f_{12}},$$

$$f_{21} - f_{12} = f_{22}\left(\frac{f_{11}}{f_{12}} - \frac{f_{11}}{f_{21}}\right),$$

$$f_{21} - f_{12} = f_{22}\left(\frac{f_{11}f_{21} - f_{11}f_{12}}{f_{12}f_{21}}\right),$$

$$f_{21} - f_{12} = f_{11}f_{22}\left(\frac{f_{21} - f_{12}}{f_{12}f_{21}}\right).$$

Разделим правую и левую части равенства на $(f_{21} - f_{12})$, В результате получим: $\dfrac{f_{11}f_{22}}{f_{12}f_{21}} = 1$, т.е. $f_{11}\, f_{22} = f_{12}\, f_{21}$.

2) Пусть $f_{11}\, f_{22} = f_{12}\, f_{21}$, тогда $f_{11} = \dfrac{f_{21}f_{12}}{f_{22}} \Rightarrow$ $\Rightarrow \dfrac{f_{11}}{f_{21}} = \dfrac{f_{12}}{f_{22}}$ (*). А коэффициент независимости будет равен: $c_1 = \dfrac{f_{01}\, f_{12}\, f_{20}}{f_{10}\, f_{21}\, f_{02}} =$ $= \dfrac{(f_{11}+f_{21}) f_{12} (f_{21}+f_{22})}{(f_{11}+f_{12}) f_{21} (f_{12}+f_{22})}$. Умножим числитель и знаменатель на f_{11}, получим $c_1 = \dfrac{(f_{11}+f_{21}) f_{12} f_{11} (f_{21}+f_{22})}{(f_{11}+f_{12}) f_{21}\, f_{11} (f_{12}+f_{22})}$. Учитывая соотношение (*) коэффициент независимости примет вид:

$$c_1 = \frac{(f_{11}+f_{21}) f_{12} f_{12} (f_{21}+f_{22})}{(f_{11}+f_{12}) f_{22} f_{11} (f_{12}+f_{22})} =$$

$$= \frac{(f_{11}f_{12}+f_{21}f_{12})(f_{12}f_{21}+f_{12}f_{22})}{(f_{11}f_{22}+f_{12}f_{22})(f_{11}f_{12}+f_{11}f_{22})}.$$

Но так как $f_{11}\, f_{22} = f_{12}\, f_{21}$, то

$$c_1 = \frac{(f_{11}f_{12}+f_{21}f_{12})(f_{12}f_{21}+f_{12}f_{22})}{(f_{21}f_{12}+f_{12}f_{22})(f_{11}f_{12}+f_{21}f_{12})} = 1.$$

Теорема полностью доказана ▲.

Из Теоремы 4 следует, что произведение частот на главной диагонали должно быть в точности равно произведению частот на побочной диагонали в четырехклеточной таблице (матрице) сопряженности. Следовательно, определитель данной таблицы будет равен 0, т.е. $f_{11}\, f_{22} - f_{12}\, f_{21} = 0$.

Значение коэффициента c_1 характеризует силу независимости между переменными. В случае независимости переменных $c_1 = 1$. Однако в действительности независимые переменные практически не встречаются. Поэтому имеет смысл вычислять величину абсолютной ошибки независимости. Абсолютной ошибкой независимости (отклонением независимости) будет выступать величина $|1 - c_1|$. Чем она меньше, тем с большей «силой» переменные независимы.

Для оценки точности выводов будем использовать не только абсолютную ошибку независимости, но и абсолютную ошибку отклонения частот. Для этого нужно определить, насколько отличаются величины $\frac{f_{01}}{f_{02}} f_{12}$ и $\frac{f_{02}}{f_{01}} \cdot f_{21}$ соответственно от $\frac{f_{10}}{f_{20}} f_{21}$ и $\frac{f_{20}}{f_{10}} \cdot f_{12}$.

Определение 4. Если выполняется условие

$$\left| \frac{f_{01}}{f_{02}} f_{12} - \frac{f_{10}}{f_{20}} f_{21} \right| < \varepsilon_1 , \tag{7}$$

то переменные A и B независимы с абсолютной ошибкой отклонения частот ε_1.

Определение 5. Если выполняется условие

$$\left| \frac{f_{02}}{f_{01}} f_{21} - \frac{f_{20}}{f_{10}} f_{12} \right| < \varepsilon_2 , \tag{8}$$

то переменные A и B независимы с абсолютной ошибкой отклонения частот ε_2.

Для выявления независимости двух признаков из выражений (7) и (8) следует брать то, для которого абсолютная ошибка отклонения частот минимальна. Причем при решении этого вопроса рекомендуется рассматривать не сами абсолютные величины разности частот ε_1 и ε_2, а величины относительных ошибок $\frac{\varepsilon_1}{f_{00}} \cdot 100\ \%$ и $\frac{\varepsilon_2}{f_{00}} \cdot 100\ \%$. Данная оценка намного точнее указывает величину ошибки в процентном отношении и при этом учитывается объем выборки.

В классической теории измерений приемлемой стандартной ошибкой считается относительная ошибка, составляющая 5—10 %, но не более [3].

Но даже если между A и B нет никакой связи, из этого еще не следует, что они независимы на данном множестве $\{f_{ij}\}$, поскольку может вмешаться случайная вариация [1]. Поэтому считается целесообразным использовать несколько различных способов, чтобы подтвердить или опровергнуть независимость переменных. Для формирования новых и использования уже существующих способов проверки независимости необходимо вернуться к теоретическому распределению частот, часто являющемуся основой того или иного критерия.

Пример 1. По данным Управления Роспотребнадзора в Республике Мордовия в 2009—2010 гг. было зарегистрировано 75 случаев заболевания вирусом A(H1N1) («свиной грипп») среди взрослых людей (18—58 лет) и 45 случаев среди детей (до 18 лет). Среди взрослых — было 9 летальных исходов, среди детей летальных исходов не было. Данные по заболеваемости и результатам лечения представлены в таблице 2.

Таблица 2

Частоты результатов лечения вируса A(N1H1) в Республике Мордовия в 2009—2010 гг.

	Результат лечения		
Возраст	Летальный	Выздоровление	Всего
Дети 0—17 лет	0	45	45
Взрослые 18—58 лет	9	66	75
Всего	9	111	120

Очевидно, что пациентов с диагнозом выздоровел большинство — 111 из 120 человек. Однако, нужно доказать статистически, что выздоровление носит неслучайный характер и понять, значимо ли отличается количество выздоровевших детей и взрослых. Для этого следует найти значения выражений $\frac{f_{01}}{f_{02}} f_{12}$ и $\frac{f_{10}}{f_{20}} f_{21}$, а затем сравнить их, а также сравнить значения $\frac{f_{02}}{f_{01}} f_{21}$ и $\frac{f_{20}}{f_{10}} f_{12}$. Таким образом, будет получена предварительная оценка коэффициента, характеризующего степень независимости признаков "Возраст" и "Результат лечения". Соответствующие значения будут следующими: $\frac{f_{01}}{f_{02}} f_{12} = 3.649$, $\frac{f_{10}}{f_{20}} f_{21} = 5.400$, $\frac{f_{02}}{f_{01}} \cdot f_{21} = 111.000$,

$\frac{f_{20}}{f_{10}} \cdot f_{12} = 75.000$. Тогда абсолютная ошибка отклонения частот для элементов первой строки и первого столбца составляет $\varepsilon_1 = |5{,}400 - 3{,}649| = 1{,}751$, а ошибка для элементов второй строки и второго столбца равна $\varepsilon_2 = |111{,}000 - 75{,}000| = 36{,}000$. Величина первой ошибки достаточно мала. Но это без учета объема выборки, и поэтому, чтобы точнее оценить данную ошибку, необходимо полученную величину поделить на объем выборки. В первом случае относительная ошибка $\frac{1{,}751}{120} \cdot 100 = 1{,}46\ \%$ незначительна, поэтому можно сделать предварительный вывод, что значения $\frac{f_{01}}{f_{02}} f_{12}$ и $\frac{f_{10}}{f_{20}} f_{21}$ приближенно равны, и, следовательно, переменные независимы «на уровне» $1 - 0{,}015 = 0{,}985$ или в 98,5 % случаев. Для второго случая данная величина будет равна $\frac{36}{120} \cdot 100 = 30\ \%$, т.е. достаточно велика, поэтому можно сделать предварительный вывод, что значения $\frac{f_{02}}{f_{01}} \cdot f_{21}$ и $\frac{f_{20}}{f_{10}} \cdot f_{12}$ далеко не равны, и, следовательно, переменные независимы «на уровне» $1 - 0{,}3 = 0{,}7$.

Далее, чтобы убедиться в правоте этого вывода, необходимо рассчитать c – коэффициент, характеризующий степень независимости признаков. Так как значение частоты в левой верхней ячейке равно 0, то воспользуемся формулой, приведенной в Замечании 2. Получается, что коэффициент независимости $c_1 = \frac{75}{111} = 0{,}676$, что свидетельствует о том, что переменные независимы с абсолютной ошибкой независимости $\alpha_1 = |1 - 0{,}676| = 0{,}324$. Если рассматривать выражение (6), то коэффициент независимости $c_2 = \frac{111}{75} = 1{,}480$ или по формуле, приведенной в Замечании 1.1. $c_2 = \frac{1}{0{,}676} = 0{,}480$. В этом случае абсолютная ошибка независимости $\alpha_2 = |1 - 1{,}480| = 0{,}480$.

Учитывая, что одна из величин абсолютной ошибки отклонения частот приемлемо мала, а значения коэффициентов независимости выше среднего 0.5, предварительно можно сделать вывод, о том, что результат лечения не зависит от возраста пациента. Данный способ проверки не претендует на абсолютную точность выводов, но его преимущества заключаются в том, что он достаточно прост и не требует больших математических расчетов.

ЛИТЕРАТУРА

1. *Аптон Г.* Анализ таблиц сопряженности / перевод с англ. и предисл. Ю. П. Адлера. — М. : Финансы и статистика, 1982. — 143 с.
2. *Кобзарь А. И.* Прикладная математическая статистика. — М. : Физматлит, 2006. — 816 с.
3. *Тейлор Дж.* Введение в теорию ошибок / пер. с англ. — М.: Мир, 1985. — 272 с. ил.
4. *Фишер Р. А.* Статистические методы для исследователей. — М. : Госстатиздат, 1958. – 268 с.

Бакаева Ольга Александровна — аспирантка, Мордовский государственный университет им. Н. П.Огарева

Тел. 8 (4342) 29-06-49

E-mail: helga_rm@rambler.ru

Щенников Владимир Николаевич – профессор, доктор физико-математических наук, Мордовский государственный университет им. Н. П.Огарева

Тел. 8 (4342) 23-32-05

Bakaeva Olga Aleksandrovna — Post-Graduate Student, Mordovian State University named after N. P. Ogarev

Tel. 8 (4342) 29-06-49

E-mail: helga_rm@rambler.ru

Shchennicov Vladimir Nikolaevich — Professor, Doctor of Physics-math. Sciences, Mordovian State University named after N. P. Ogarev

Tel. 8 (4342) 23-32-05