

**П.Е. ПУСТОВОЙТОВ**, канд. техн. наук, доц., НТУ "ХПИ", Харьков,  
**Ж.Д. КРАВЧЕНКО**, зам. главного врача, ХГСБ, Харьков

## **АНАЛИЗ СТОХАСТИЧЕСКОЙ ЗАВИСИМОСТИ НОЗОЛОГИЙ ПРИ ПЛАНИРОВАНИИ ЭФФЕКТИВНОГО СТРУКТУРНОГО ОБЕСПЕЧЕНИЯ МЕДИЦИНСКОГО ЦЕНТРА ВУЗА**

В работе предлагается методика планирования нагрузки врача-специалиста, основанная на анализе парной стохастической зависимости количества обращений по двум выбранным нозологиям за сутки. На примере двух заболеваний был проведен анализ парной корреляции, проверена значимость коэффициента корреляции, построена модель регрессии. В работе используются реальные данные со статистических талонов Харьковской городской студенческой больницы. Ил.: 1. Табл.: 1. Библиогр.: 9 назв.

**Ключевые слова:** коэффициент корреляции, стохастическая зависимость нозологий, парная корреляция, нозология.

**Постановка задачи и анализ литературы.** При эффективном планировании структурно-функционального обеспечения медицинского центра ВУЗа необходимо учитывать стохастическую зависимость между некоторыми нозологиями [1]. Принимая во внимание такую зависимость, можно делать выводы как относительно качества выставления диагнозов, так и эффективного планирования штатных единиц медицинского центра [2]. Использование математического аппарата в медицине позволяет предсказать вспышки эпидемии и снизить вред от их последствий [3 – 4].

Для примера рассмотрим зависимость между классами нозологий J35 (хронический тонзиллит) и J32 (хронический синусит). Общую картину их взаимной изменчивости можно получить, изобразив на координатной плоскости все пары чисел, соответствующие количеству обращений по каждому виду нозологий в день.

Полученное корреляционное поле построено по данным со статистических талонов Харьковской городской студенческой больницы за период с 01.01.2011 по 31.12.2011 (рис.).

По графику видно, что в целом для всей совокупности значений увеличение количества случаев появления больных заболеванием J35 приводит к увеличению количества больных заболеванием J32.

**Целью статьи** является определение известными математическими методами функциональной зависимости между количествами обращений за день по двум классам болезней.

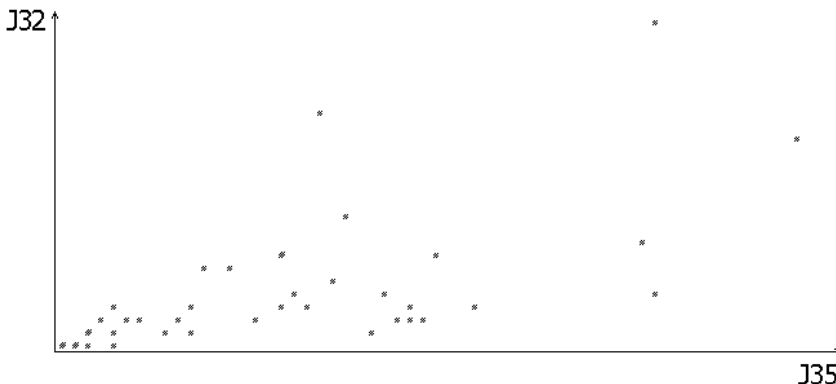


Рис. Взаимная зависимость количества нозологий J35 и J32 в день

**Основные результаты.** Для построения корреляционной таблицы разобьем данные о количестве обращений пациентов по каждой из рассматриваемых болезней на интервалы. Длина интервала вычисляется по формуле Стерджесса [5 – 7]:

$$l = \frac{x_{\max} - x_{\min}}{1 + 3.322 * \lg n},$$

где  $x_{\max}$  и  $x_{\min}$  – максимальное и минимальное значения выборки соответственно,  $n$  – объем выборки.

В рассматриваемой задаче получены следующие параметры:

*Всего пар значений: 76.*

*Длина интервала по J35: 8.*

*Длина интервала по J32: 4.*

*Количество интервалов по J35: 7.*

*Количество интервалов по J32: 6.*

По полученным данным со статистических талонов, учитывая указанные выше параметры, строится корреляционная таблица [8 – 9] (табл.). Здесь  $\bar{x}_i$  и  $\bar{y}_j$  – средние значения случайной величины для интервалов, а  $n_{i\bullet}$  и  $n_{\bullet j}$  абсолютная частота попадания случайных величин в соответствующие интервалы.

Коэффициент корреляции Пирсона вычисляется по формуле

$$\rho_{X,Y}^* = \frac{S_{X,Y}}{S_X S_Y},$$

где  $S_{X,Y}$  – ковариация случайных величин [8], а  $S_X^2$  и  $S_Y^2$  – соответственно дисперсии случайных величин [8].

Таблица

Корреляционная таблица

$\bar{x}_i$ $\bar{y}_j$	3	11	19	27	33	46	58	$n_{\bullet j}$
2	21	6	2	8	0	0	0	37
6	2	6	8	4	2	2	0	24
10	0	0	5	2	0	2	0	9
14	0	0	0	0	0	0	0	0
18	0	2	0	0	0	0	2	4
22	0	0	0	0	0	0	0	0
26	0	0	0	0	0	2	0	2
$n_{i\bullet}$	23	14	15	14	2	6	2	

В примере получен коэффициент корреляции Пирсона [10]  $\rho^* \approx 0.627$ , что говорит о наличии прямой линейной зависимости.

Для проверки значимости коэффициента корреляции выдвинем основную гипотезу ( $H_0: \rho = 0$ ). В роли альтернативной гипотезы выступит ( $H_1: \rho \neq 0$ ). Используем статистику Фишера [11]. Известно, что случайная величина

$$V = \frac{1}{2} \ln \frac{1 + \rho^*}{1 - \rho^*}$$

распределена по нормальному закону, а ее математическое ожидание и дисперсия равны соответственно

$$m_V \approx \frac{1}{2} \cdot \ln \frac{1 - \rho}{1 + \rho}, \quad \sigma_V^2 = \frac{1}{n - 3}.$$

Тогда, стандартизуя  $V$ , получаем нужную статистику Фишера:

$$U = \frac{V - m_V}{\sigma_V} = \frac{\sqrt{n - 3}}{2} \ln \frac{1 + \rho^*}{1 - \rho^*} \sim N(0, 1),$$

$$U = \frac{\sqrt{76 - 3}}{2} \cdot \ln \frac{1 + 0.627}{1 - 0.627} = \frac{8.544}{2} \cdot \ln \frac{1.627}{0.373} \approx 6.293.$$

Квантиль нормального стандартного распределения уровня 6.293 равен 0,9999999998443, откуда следует, что полученный коэффициент корреляции значим на уровне значимости 0,000000000078.

Предложим для данного случая модель линейной регрессии

$$y(x) = a_0 x + a_1. \quad (1)$$

Найдем МНК-оценки [12] параметров регрессии из соотношения

$$\psi(a_0, a_1) = \sum_{k=1}^n [y_k - a_1 - a_0 x_k]^2 \rightarrow \min. \quad (2)$$

Продифференцируем соотношение (2) по параметрам

$$\begin{cases} \frac{d\psi}{da_0} = 0, \\ \frac{d\psi}{da_1} = 0. \end{cases} \quad \begin{cases} a_0 \sum_{k=1}^n x_k^2 + a_1 \sum_{k=1}^n x_k = \sum_{k=1}^n x_k y_k, \\ a_0 \sum_{k=1}^n x_k^2 + a_1 n = \sum_{k=1}^n y_k. \end{cases} \quad (3)$$

Получим систему 2-х уравнений с двумя неизвестными

$$\begin{cases} a_0 \alpha_{2,0}^{\bullet} + a_1 \bar{x} = \alpha_{1,1}^{\bullet}, \\ a_0 \bar{x} + a_1 = \bar{y}, \end{cases} \quad (4)$$

где  $\alpha_{r,s}^{\bullet}$  – начальный выборочный момент порядка  $r + s$ , а  $\bar{x}$  и  $\bar{y}$  – средние значения.

Решение системы имеет вид

$$\tilde{a}_0 = \frac{S_{X,Y}}{S_X} = \rho^{\bullet} \cdot \frac{S_Y}{S_X}, \quad (5)$$

$$\tilde{a}_1 = \bar{y} - \bar{x} \rho^{\bullet} \cdot \frac{S_Y}{S_X}. \quad (6)$$

Тогда подставляя (5) и (6) в (1) получаем соотношение для уравнения регрессии

$$y(x) = \bar{y} + \rho^{\bullet} \cdot \frac{S_Y}{S_X} \cdot (x - \bar{x}). \quad (7)$$

Для предлагаемого примера параметры регрессии имеют следующие значения

$$\tilde{a}_0 = 0.627 \cdot \frac{27}{201} = 0.084, \quad (8)$$

$$\tilde{a}_1 = 5.289 - 18.447 \cdot 0.627 \cdot \frac{27}{201} = 3.735. \quad (9)$$

Тогда, с учетом (8) и (9), уравнение регрессии для примера задается соотношением

$$y(x) = 0.084x + 3.735. \quad (10)$$

**Выводы.** Таким образом, был проведен анализ зависимости нозологий для двух выбранных видов болезней по данным Харьковской городской студенческой больницы. Вычисленный коэффициент корреляции Пирсона позволил предположить наличие линейной зависимости между количествами обращений по каждому из рассмотренных видов болезней за сутки. Была проверена значимость коэффициента корреляции, построена модель регрессии, найдены параметры функции регрессии.

Оценивание парных зависимостей между нозологиями позволяет планировать эффективное структурно-функциональное обеспечение медицинского центра ВУЗа, вычислять нагрузку врачей-специалистов, медицинских сестер.

Рассмотренная методика может быть внедрена в виде программного обеспечения для полного охвата нозологий международного классификатора болезней МКБ-10.

**Список литературы:** 1. *Баевский Р.М.* Прогнозирование состояний на грани нормы и патологии / *Р.М. Баевский*. – М.: "Знание", 1979. – 74 с. 2. *Бароян О.В.* Математика и эпидемиология / *О.В. Бароян, Л.А. Рвачев*. – М.: "Знание", 1977. – 63 с. 3. *Бароян О.В.* Моделирование и прогнозирование эпидемий гриппа / *О.В. Бароян, Л.А. Рвачев, Ю.Г. Иванников*. – М.: ИЭМ им. Н.Ф. Гамалеи, 1977. – 546 с. 4. *Беляков В.Д.* Эпидемиология / *В.Д. Беляков, Р.Х. Яфаев*. – М.: Медицина, 1989. – 104 с. 5. *Айвазян С.А.* Прикладная статистика и основы эконометрики / *С.А. Айвазян, В.С. Мхитарян*. – М.: ЮНИТИ, 1998. – 1005 с. 6. *Боровков А.А.* Математическая статистика. Оценка параметров. Проверка гипотез / *А.А. Боровков*. – М.: Наука, 1984. – 472 с. 7. *Гнеденко Б.В.* Элементарное введение в теорию вероятностей / *Б.В. Гнеденко, А.Я. Хинчин*. – М.: Наука, 1970. – 169 с. 8. *Гмурман В.Е.* Теория вероятностей и математическая статистика: Учеб. пособие для вузов / *В.Е. Гмурман*. – М.: Высш. шк., 2003. – 479 с. 9. *Рыков В.В.* Математическая статистика и планирование эксперимента / *В.В. Рыков, В.Ю. Иткин*. – М.: РГУ нефти и газа им. И.М. Губкина, 2009. – 303 с. 10. *Вентцель Е.С.* Теория вероятностей и ее инженерные приложения / *Е.С. Вентцель, Л.А. Овчаров*. – М.: Академия, 2003. – 464 с. 11. *Айвазян С.А.* Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное изд. / *С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин*. – М.: Финансы и статистика, 1983. – 471 с. 12. *Горяинов В.Б.* Математическая статистика: Учеб. для вузов / *В.Б. Горяинов, И.В. Павлов, Г.М. Цветкова* и др. / Под ред. В.С. Зарубина, А.П. Крищенко. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2001. – 424 с.

*Статью представил д.т.н., проф. НТУ "ХПИ" Серков А.А.*

УДК 519.2

**Аналіз стохастичної залежності нозологій при плануванні ефективного структурного забезпечення медичного центра ВНЗ / Пустовойтов П.Е., Кравченко Ж.Д. // Вісник НТУ "ХПИ". Серія: Інформатика та моделювання. – Харків: НТУ "ХПИ". – 2012. – № 38. – С. 142 – 147.**

У роботі пропонується методика планування навантаження лікаря-спеціаліста, заснована на аналізі парних стохастичних залежностей кількості звернень по двом обраним нозологіям за добу. На прикладі двох захворювань було проведено аналіз парної кореляції, перевірено значимість коефіцієнту кореляції, побудовано модель регресії. У роботі використовуються реальні дані зі статистичних талонів Харківської міської студентської лікарні. Лл.: 1. Табл.: 1. Бібліогр.: 12 назв.

**Ключові слова:** коефіцієнт кореляції, стохастична залежність нозологій, парна кореляція, нозологія.

UDC 519.2

**Stochastic analysis of nosologies dependences during the effective structure supply planning in campus medical center / Pustovoitov P., Kravchenko Z. // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkov: NTU "KhPI". – 2012. – №. 38. – P. 142 – 147.**

It was suggested the technique of doctor loading planning, based on the analysis of pair stochastic dependence of statistic talons number during the day. It was showed the example of two diseases. The analysis of pair correlation was carried out, the importance of factor of correlation is checked, the regression model is constructed. It was used real data from statistical talons of the Kharkiv municipal student's hospital. Figs.: 1. Tab.: 1. Refs.: 12 titles.

**Keywords:** correlation coefficient, stochastic dependence, pair correlation, nosology.

*Поступила в редакцію 20.04.2012*