

О ВЗАИМОСВЯЗИ МЕЖДУ ЭЛЕМЕНТАМИ ВЫБОРКИ

Коваленко В.И., Шичинов А.И. (ДонНТУ, г. Донецк, Украина)
Тел./Факс: +38 (062) 3010805; E-mail: tm@mech.dgtu.donetsk.ua

Abstract: The work is devoted correlation detection between the elements characterising sample. For this purpose the correlation analysis which is one of widely applied methods of an estimation of statistical connections is used. As the elementary numerical performance of connection the covariance (expectation) serves. For definition of the statistical importance of the received outcome the hypothesis about equality to zero of an estimation of a coefficient of correlation is checked. In work the flow chart of carrying out of a linear correlation analysis of the statistical data is reduced.

Keywords: sample, the flow chart, the analysis, correlation.

Корреляционный анализ – один из широко применяемых методов оценки статистических связей [1]. Он отвечает на вопросы: влияет ли данная величина на выходную и какова степень (теснота) связи между величинами?

Простейшей численной характеристикой связи [2, 3] служит ковариация (математическое ожидание) произведения отклонений x и y от их центров

$$\text{cov}(x; y) = \frac{\sum_{i=1}^n (x_i - \bar{X}) \cdot (y_i - \bar{Y})}{n} \quad (1)$$

где: n – количество пар значений величин x и y .

Однако ковариация зависит от размерности переменных. Для перехода к безразмерной характеристике отношения переменных нормируют

$$x^* = \frac{x_i - \bar{X}}{\sigma_x}; \quad y^* = \frac{y_i - \bar{Y}}{\sigma_y}$$

Ковариацию нормированных отклонений называют коэффициентом корреляции

$$\rho_{x,y} = \text{cov}(x^*; y^*) = \frac{\sum_{i=1}^n (x_i - \bar{X}) \cdot (y_i - \bar{Y})}{n \cdot \sigma_x \cdot \sigma_y}, \quad (2)$$

где: σ_x, σ_y – стандартные отклонения случайных величин x и y соответственно.

Так как σ_x и σ_y неизвестны, то определяют не коэффициент корреляции, а его оценку:

$$r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{X}) \cdot (y_i - \bar{Y})}{n \cdot S_x \cdot S_y} \quad (3)$$

где: S_x, S_y – оценки стандартного отклонения, мм.

Оценка коэффициента корреляции может принимать значения от -1 до +1.

Для определения статистической значимости полученного результата проверяется гипотеза о равенстве нулю оценки коэффициента корреляции. В качестве критерия при проверке гипотезы используется случайная величина которая подчиняется распределению Стьюдента с $k=n-2$ степенями свободы.

$$T = \frac{r_{x,y} \cdot \sqrt{n-1}}{\sqrt{1-r_{x,y}^2}}, \quad (4)$$

Для проверки гипотезы о равенстве нулю оценки коэффициента корреляции необходимо по формуле (4) вычислить наблюдаемое значение критерия $T_{набл}$, а затем по заданному уровню значимости α , числу степеней свободы k и таблице критических точек распределения Стьюдента найти критическую точку $t_{кр}(\alpha; k)$. Если $|T_{набл}| < t_{кр}$, то нет оснований отвергать гипотезу.

Если $|T_{набл}| > t_{кр}$, то гипотезу отвергаем, то есть оценка коэффициента корреляции статистически значима и между исследуемыми величинами (торцового и радиального биений) наблюдается линейная зависимость.

Оценка коэффициента корреляции может быть определена из выражения

$$r_{x,y} = \sqrt{1 - (2\gamma)^2}, \quad (5)$$

где: γ - относительная погрешность оценки коэффициента корреляции.

В программе «Выборка» реализован линейный корреляционный анализ данных, блок – схема которого приведена на рис. 1.

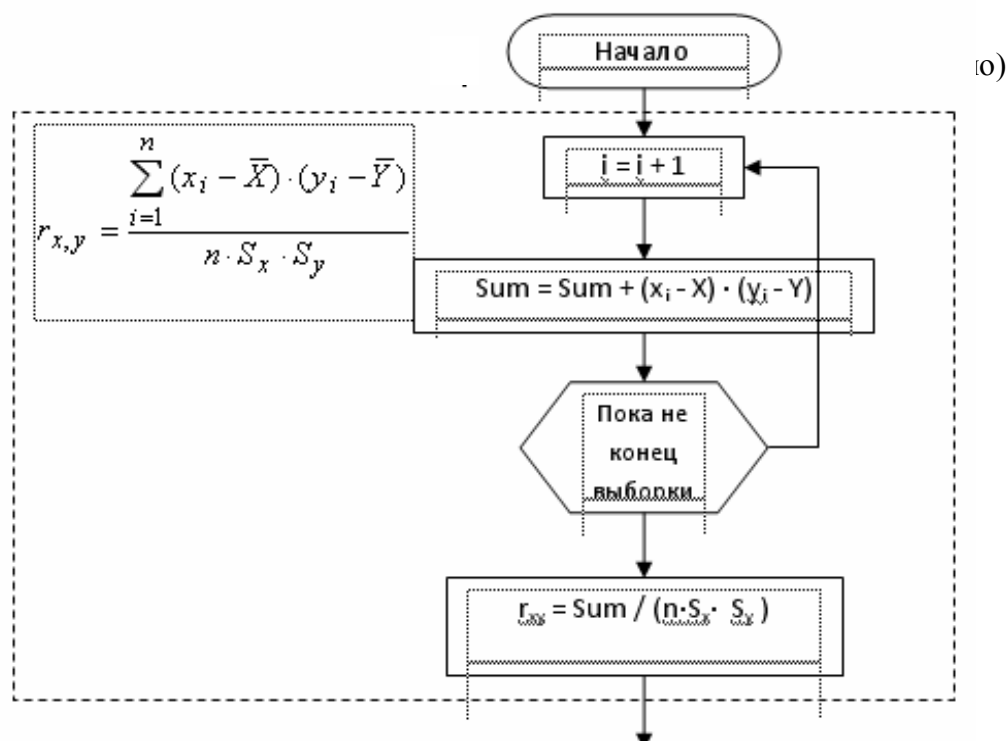


Рис. 1,а. Блок – схема проведения корреляционного анализа (начало)

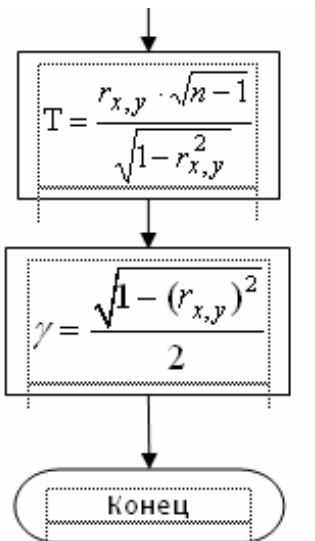


Рис. 1,б. Блок – схема проведения корреляционного анализа (окончание)

Конечным результатом для работы данного модуля программного обеспечения является коэффициент корреляции и проверка гипотезы о равенстве нулю оценки коэффициента корреляции.

Вычисления, проведенные с помощью программы обработки статистических данных, с использованием зависимостей (4) – (5), показали, что нулевая гипотеза о равенстве нулю оценки коэффициента корреляции не подтверждается (рис. 2).

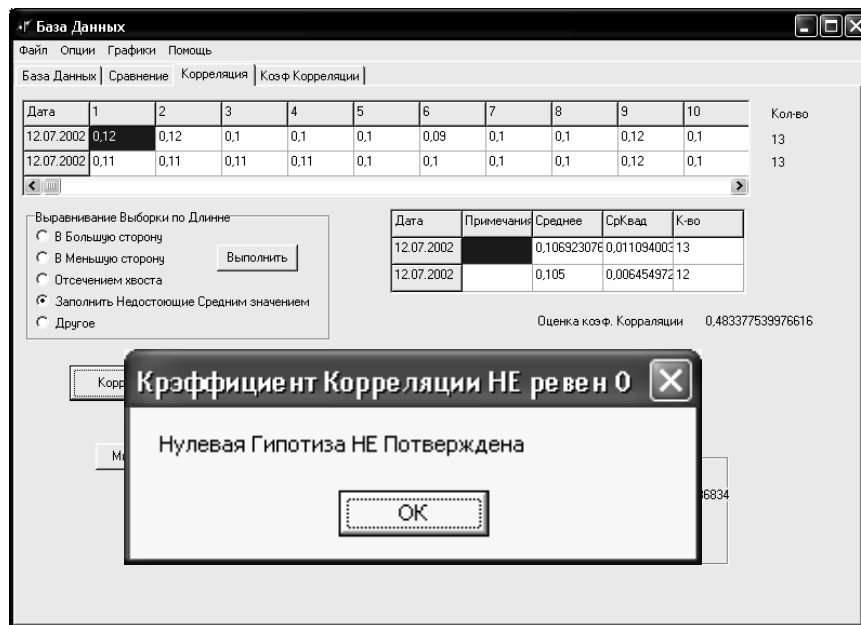


Рис. 2. Результаты проведения корреляционного анализа

Список литературы: 1. Орлов А. И. Прикладная статистика. Учебник. — М.: Экзамен, 2006. — 671 с. 2. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход : монография / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. – Новосибирск : Изд-во НГТУ, 2011. – 888 с. 3. Лекции по эконометрике: учебное пособие/Шанченко Н.И., Ульяновск: УлГТУ, 2008.