

## ПРОВЕРКА НОРМАЛЬНОСТИ ЭКСПЕРИМЕНТАЛЬНЫХ МЕДИЦИНСКИХ ДАННЫХ С ПОМОЩЬЮ КРИТЕРИЯ ШАПИРО-УИЛКА

*А.Л. Бурцева*

*(г. Томск, Томский политехнический университет)*

*E-mail: anechkabv@mail.ru*

## CHECK THE NORMALITY OF EXPERIMENTAL MEDICAL DATA BY MEANS OF SHAPIRO-WILKS TEST

*A.L. Burceva*

*(Tomsk, Tomsk Polytechnic University)*

**Abstract.** The statistical methods for data analysis are used for the solution of a set of various tasks in human activity. The important condition which defining the possibility of application of parametrical methods is submission of the analyzed data to the law of normal distribution which has a characteristic bell-shaped appearance. We will conduct research on normality by means of the Shapiro-Wilk W-test because this method is more powerful.

**Keywords:** Normal distribution, Shapiro-Wilk W-test, experimental medical data, data analysis, bronchial asthma, Statistica 8.0.

Актуальность данной работы обусловлена необходимостью решения проблем связанных с широким распространением бронхиальной астмы. В последние годы эти проблемы приобретают все больший размах.

Исходные данные: данные традиционных показателей вентиляции легких и параметры механики дыхания [1, 2].

Объект исследования: больные с различными формами бронхиальной астмы и группа больных с психогенной одышкой, у которых измерены различные физиологические показатели.

Статистические методы анализа данных широко используются для решения множества различных задач в медико-биологической деятельности человека. Как известно, существующие методы статистического анализа можно подразделить на две группы – параметрические и непараметрические. Важным условием, определяющим возможность применения параметрических методов, является подчинение анализируемых данных закону нормального распределения, которое имеет характерный колоколообразный вид. В то же время непараметрические методы выполнения этого условия не требуют. Установлено, что в подавляющем большинстве случаев (около 75 %) распределения признаков существенно отличаются от нормального. Во избежание ошибок, любой анализ признаков должен сопровождаться проверкой нормальности их распределения [3].

Существует несколько подходов, реализованных в пакете Statistica 8.0 [3-6]:

- модуль Distribution fitting. Этот модуль позволяет проверить соответствие анализируемых данных целому ряду математических распределений с помощью критерия  $\chi^2$ . Мощность теста относительно невысока.
- модуль Descriptive Statistics с применением теста Колмогорова-Смирнова и Лиллифорса. Этот тест проверяет нулевую гипотезу об отсутствии различий между наблюдаемым распределением признака и теоретическим ожидаемым нормальным распределением;
- график нормальных вероятностей. Такой график изображает зависимость ожидаемых нормальных частот значений признака от их реальных частот.
- модуль Descriptive Statistics с применением W-критерия Шапиро-Уилка. Данный критерий обладает наибольшей мощностью и является наиболее предпочтительным, особенно при небольших выборках (около 50-80 наблюдений).

Проведем исследование на нормальность с помощью W-теста Шапиро-Уилка, так как этот метод обладает большей мощностью.

Создадим гистограмму распределения значений признака и ожидаемую нормальную кривую. Результаты выбранных тестов на нормальность автоматически располагаются в за-

головке этого графика. При  $p > 0,05$  можно заключить, что анализируемое распределение не отличается от нормального.

Как видно на рис. 1 с данными МОД1 для теста Шапиро-Уилка получаем  $p = 0,43146$ , что свидетельствует о нормальности распределения данных МОД1. На рис. 2 видно, что распределение C dyn. ненормально.

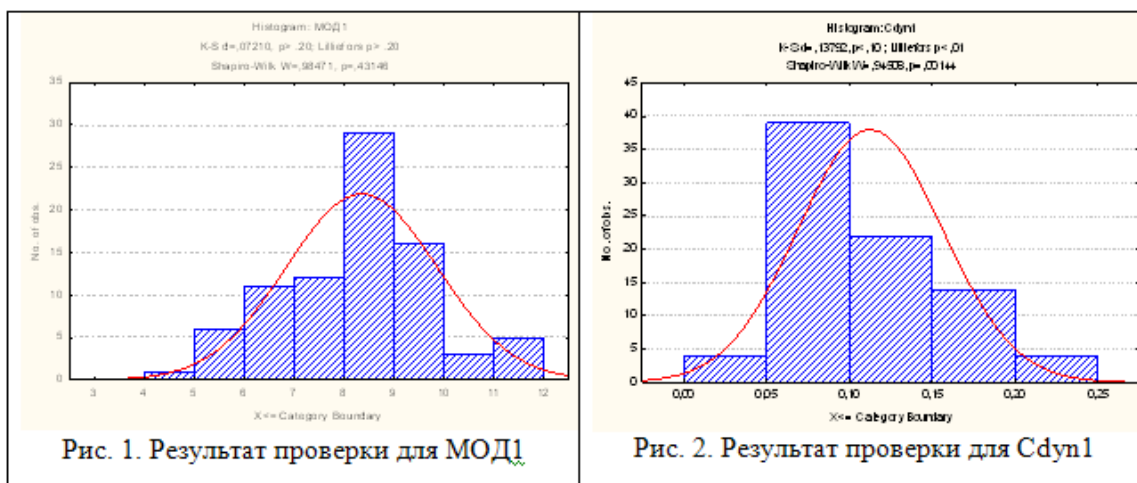


Рис. 1. Результат проверки для МОД1

Рис. 2. Результат проверки для Cdyn1

Анализ всех традиционных показателей вентиляции легких и параметров механики дыхания до (1) и после (2) лечения (МОД, ЖЕЛ, ФЖЕЛ, ОФВ-1, ОФВ-1/ЖЕЛ, МВЛ, ПОС выд., МОС-25, МОС-50, МОС-75 и W общ., W н.эл., W эл., W уд., W МОД 10, W МОД 15, W МВЛ общ., W МВЛуд.1, C dyn., C stat., Рвд., Рвыд.) позволяет сделать вывод о том, что далеко не все данные распределены нормально, так как во многих случаях не выполняется условие  $p < 0.05$ .

Большая часть данных имеет распределение отличное от нормального, среди них: ОФВ-1/ЖЕЛ1, Cdyn.1, Cstat.1, Рвд.1, Рвыд.1, Wобщ.1, Wуд.1, Wн.эл.1, Wэл.1, W МОД10\_1, W МОД15\_1, W МВЛобщ.1, W МВЛуд.1.

**Публикация подготовлена в рамках проектов РФФИ №15-07-08922 и № 14-07-00675.**

### Список литературы

1. Немеров Е.В., Языков К.Г. К вопросу изучения личностных свойств в психофизиологической реактивности больных бронхиальной астмой на аудиовизуальную стимуляцию // Вестник ТГПУ. – 2011. – Вып. 6 (108). – С. 134 – 137.
2. Берестнева О.Г., Осадчая И.А., Прокопьев Р.О. Разработка интеллектуальной системы исследования психогенных и непсихогенных форм бронхиальной астмы // Известия Южного федерального университета. Технические науки. – 2014. – № 10 (159). – С. 61-68.
3. Боровков А.А. Математическая статистика. Оценка параметров. Проверка гипотез / А. А. Боровков. – М.: ФИЗМАТЛИТ, 2007. – 704 с.
4. Берестнева О. Г. , Муратова Е. А. , Шелехов И. Л. , Жаркова О. С. , Уразаев А. М. Математические методы в психологии и педагогике: Учебное пособие. – Томск : ТГПУ, 2012. – 276 с.
5. Сидоренко, Е.В. Методы математической обработки в психологии. /Е.В.Сидоренко. – СПб.: Социально-психологический центр, 2006. – 352 с.