УДК 303.4 : 519.2

5.2.2. Математические,
статистические и инструментальные методы экономики (физико-математические науки, экономические науки)

# ПРОВЕРКА СОГЛАСИЯ С БЕТА-РАСПРЕДЕЛЕНИЕМ МЕТОДОМ МОМЕНТОВ

Орлов Александр Иванович
д.э.н., д.т.н., к.ф.-м.н., профессор
РИНЦ SPIN-код: 4342-4994
prof-orlov@mail.ru
*Московский государственный технический университет им. Н.Э. Баумана, Россия, 105005, Москва, 2-я Бауманская ул., 5*

Проверка согласия опытного распределения с теоретическим - одна из основных задач прикладной математической статистики. Проверяют, можно ли в качестве функции распределения элементов выборки рассматривать конкретное теоретическое распределение или распределение из того или иного параметрического семейства. Пример - проверка нормальности (в качестве параметрического семейства рассматривается семейство нормальных распределений). Настоящая статья посвящена проверке согласия с семейством бета-распределений. Анализируем место рассматриваемой постановки задачи в многообразии современных статистических методов. Кратко обсуждаем основные этапы развития прикладной статистики: описательная (до 1900 г.), параметрическая (1900 - 1933), непараметрическая статистика (1933 - 1979), статистика нечисловых данных (с 1979 г.). Речь идет о переднем крае фронта научных исследований. "В тылу" остается ряд теоретических задач предыдущих этапов, а при анализе конкретных статистических данных используют методы всех четырех этапов. Проверка согласия с бета-распределением относится к параметрической статистике, но оставалась не изученной до настоящего времени. Как один из подходов к решению этой задачи предлагаем использовать метод моментов. Для проверки согласия с бета-распределением используем третий центральный момент. Его оцениваем как непосредственно по выборке, так и с помощью выражения через параметры бета-распределения, в котором вместо неизвестных теоретических значений подставлены их выборочные оценки, в их качестве используем оценки метода моментов. При

UDC 303.4 : 519.2

5.2.2. Mathematical,
statistical and instrumental methods of economics (physical and mathematical sciences, economic sciences)

# GOOD-OF-FIT TESTING WITH BETA DISTRIBUTION BY THE METHOD OF MOMENTS

Orlov Alexander Ivanovich
Dr.Sci.Econ., Dr.Sci.Tech., Cand.Phys-Math.Sci., professor
RSCI SPIN-code: 4342-4994
*Bauman Moscow State Technical University, Moscow, Russia*

Good-of-fit testing of the experimental distribution with the theoretical one is one of the main tasks of applied mathematical statistics. It is checked whether it is possible to consider a specific theoretical distribution or a distribution from one or another parametric family as a distribution function of the elements of the sample. An example is a test of normality (the family of normal distributions is considered as a parametric family). This article is devoted to good-of-fit testing with the family of beta distributions. Let us analyze the place of the problem statement under consideration in the variety of modern statistical methods. We briefly discuss the main stages in the development of applied statistics: descriptive (before 1900), parametric (1900 - 1933), non-parametric statistics (1933 - 1979), statistics of non-numerical data (since 1979). This is at the forefront of scientific research. "In the rear" remains a number of theoretical problems of the previous stages, and in the analysis of specific statistical data, methods of all four stages are used. Testing for goodness-of-fit with the beta distribution is related to parametric statistics, but has remained unexplored to date. As one of the approaches to solving this problem, we propose to use the method of moments. To goodness-of-fit testing with the beta distribution, we use the third central moment. We estimate it both directly from the sample and using an expression through the parameters of the beta distribution, in which instead of unknown theoretical values their sample estimates are substituted, and we use the estimates of the method of moments as them. If the fit hypothesis is true, the difference between the indicated values is asymptotically normal with expectation 0. To goodness-of-fit testing with the beta distribution, we also suggest using the skewness coefficient. The results of calculations based on real data are carried справедливости гипотезы согласия разность указанных величин является асимптотически нормальной с математическим ожиданием 0. Для проверки согласия с бета-распределением предлагаем также использовать коэффициент асимметрии. Проведены результаты расчетов по реальным данным. Сформулирован ряд нерешенных задач. Так, на основе асимптотического распределения коэффициента асимметрии можно проверять непараметрическую гипотезу симметрии функции распределения с целью обнаружения эффекта в связанных выборках. Асимптотические дисперсии рассматриваемых статистик могут быть найдены методом линеаризации

out. A number of unsolved problems are formulated. Thus, on the basis of the asymptotic distribution of the skewness, one can test the nonparametric hypothesis of the symmetry of the distribution function in order to detect the effect in paired samples. The asymptotic variances of the considered statistics can be found by the linearization method

Ключевые слова: СТАТИСТИЧЕСКИЕ МЕТОДЫ, БЕТА-РАСПРЕДЕЛЕНИЕ, МЕТОД МОМЕНТОВ, ПРОВЕРКА СОГЛАСИЯ, ТРЕТИЙ ЦЕНТРАЛЬНЫЙ МОМЕНТ, КОЭФФИЦИЕНТ АСИММЕТРИИ

Keywords: STATISTICAL METHODS, BETA DISTRIBUTION, METHODS OF MOMENTS, GOODNESS-OF-FIT TESTING, THIRD CENTRAL MOMENT, ASYMMETRY COEFFICIENT

http://dx.doi.org/10.21515/1990-4665-189-011

## Introduction

One of the main tasks of applied mathematical statistics is to check the agreement between the experimental distribution and the theoretical one. A statistical hypothesis is tested that a specific theoretical distribution or a distribution from one or another parametric family can be considered as a distribution of sample elements. The normality test hypothesis is widely known, in which the family of normal distributions is considered as a parametric family. This article discusses testing for goodness-of-fit with a family of beta distributions. First, it is necessary to discuss the place of the considered formulation of the problem in the variety of modern statistical methods, and for this it is necessary to give an idea of the main stages in the development of applied statistics.

## Stages of development of applied statistics

What is Applied Statistics? We put the following meaning into this term: "Applied statistics is the science of how to process data" [1, 2]. The basics of the scientific apparatus of this science are presented in the handbook [3].

The analysis of methods of statistical data analysis carried out by us in [4], [5, part 1] made it possible to single out the main stages in the development of this scientific field. We believe that there are four such stages. Let's briefly discuss their features.

Until the beginning of the twentieth century. - stage of descriptive statistics. Typical methods of statistical analysis are the construction of tables, the calculation of the arithmetic mean of measurement results, the restoration of dependencies by the least squares method created by K. Gauss in 1794. Descriptive statistics is not a single science, it is a set of separate algorithms. Nevertheless, descriptive statistics allows solving a number of practical problems, and its algorithms are still used today. Thus, the construction of tables is the main intellectual apparatus used by Rosstat (the Federal State Statistics Service of Russia).

Mathematical statistics as a single science was created at the beginning of the 20th century. To describe statistical data, K. Pearson suggested using distribution functions from the four-parameter family he introduced. The most well-known subfamilies are normal, logarithmically normal, exponential, gamma distributions, which are of interest to us in this article beta distributions. An advanced mathematical theory of parameter estimation and hypothesis testing has been developed for samples from parametric families. We call it parametric statistics. She was at the forefront of the development of statistical theory in 1900-1933. Note that the methods of parametric statistics were actively used to solve practical problems. So, the first scientific journal on mathematical statistics was called "Biometrics". By the middle of the twentieth century. Numerous textbooks on probability theory and mathematical statistics were prepared, the content of which is primarily parametric statistics. And at present, the teaching and practical application of statistical methods is carried out primarily on their basis.

However, parametric statistics have a fundamental drawback - the distributions of real data, as a rule, are not normal and do not belong to any other parametric distribution [6]. The works of the next stage in the development of methods of statistical data analysis - nonparametric statistics (1933 - 1979) are aimed at overcoming this shortcoming. As the beginning of this stage, we take the moment of the appearance of the A.N. Kolmogorov to check the agreement between the experimental distribution and the given theoretical one. In nonparametric statistics, problems of statistical analysis of data with arbitrary distributions are solved (for proving theorems, it may be necessary to require only the continuity of distribution functions). To date, nonparametric methods have been able to solve almost all problems of data analysis, considered in parametric statistics. But the conclusions obtained are more justified, since unrealistic assumptions are not made that the distributions of the results of observations belong to one or another parametric family.

The elements of the samples at the considered stages of the development of statistics are numbers or vectors, i.e. elements of linear spaces. With the development of science and technology, samples from distributions on nonlinear spaces began to play an increasingly important role. There was a statistics of objects of non-numerical nature. This term was first introduced in 1979 (in monograph [7] and article [8]). Its synonyms are used - statistics of non-numerical data, non-numerical statistics [9, 10]. This is the fourth stage in the development of statistical data analysis methods. It continues at the present time. Some results of its development are summarized in [11, 12]. The fourth stage also includes the statistical component of a new theoretical and applied area of mathematics - systemic fuzzy interval mathematics [13 - 15].

As already noted, the methods of statistical data analysis obtained at the previous stages continue to be actively and successfully used. In particular, judging by the common textbooks on probability theory and mathematical statistics, the system of teaching applied statistics still corresponds to parametric statistics, i.e. remains at the level of the middle of the twentieth century. Naturally, the scientific and practical workers who have undergone such training are simply not familiar with non-parametric and non-numerical statistics.

At the same time, we have to state that by no means all the scientific problems of parametric statistics have been solved at the appropriate stage in the development of statistical science. Thus, only relatively recently it has been established that in many cases one-step estimates should be used instead of maximum likelihood estimates [16]. Since parametric statistics is widely used, it seems worthwhile to make efforts to solve the "missing" problems, although the cutting edge of scientific research has moved far ahead.

The sections of parametric statistics that have not been worked out so far include the statistical analysis of data with a beta distribution. In textbooks on probability theory and mathematical statistics and related reference books, the family of beta distributions is mentioned, some of its properties are indicated [17 - 19], however, a number of issues related to parameter estimation and hypothesis testing on samples from distributions of this family remain unresolved.

We have begun to develop a system of methods for statistical data analysis, the distribution functions of which are included in the family of beta distributions. We will use the concepts and notation introduced in [20], the direct continuation of which is the present paper. In this article, we found asymptotic distributions of beta distribution parameter estimates obtained by the method of moments; in it, algorithms for point and interval estimation were developed. Obviously, estimates by the method of moments can be reasonably used only when the hypothesis is true: the sample elements are distributed according to the beta distribution with some parameter values. These values are unknown to us, we estimate them. But is this hypothesis really true? To answer this question, we consider in this article a method for testing agreement between an empirical distribution and a theoretical beta distribution based on the use of sample moments. First, we need to discuss the statement of the problem of checking the agreement.

## Methods for checking agreement with a parametric family

A fit test is a test of the statistical hypothesis that the distribution function of the sample elements is included in the considered parametric family (although the values of the distribution parameters themselves remain unknown).

By analogy with the formulation of the problems of checking the homogeneity of two independent samples [21, 22], we can distinguish between checking the agreement of characteristics and checking absolute agreement. Let us discuss these settings using the example of checking the agreement with a normal distribution.

Let us introduce the main object of study in mathematical statistics - the sample$X_1$, X2, ..., Xn, i.e., (finite) sequence of independent identically distributed random variables with distribution function F(x) = P(X < x), where X is a random variable having the same distribution as the sample elements.

Consider the theoretical central moments of the random variable $X$ order 2, 3, 4, i.e.

$$M_2 = M[(X - M(X))^2] = \sigma^2, M_3 = M[(X - M(X))^3], \\ M_4 = M[(X - M(X))^4] \quad (1)$$

and the corresponding selective central moments

$$m_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X - \overline{X}\right)^2 = s^2, m_3 = \frac{1}{n}\sum_{i=1}^{n}\left(X - \overline{X}\right)^3, m_4 = \frac{1}{n}\sum_{i=1}^{n}\left(X - \overline{X}\right)^4. \quad (2)$$

As is known [1 - 3], as the sample size tends to infinity, the sample central moments converge (in probability) at the corresponding theoretical central moments:

$$\lim_{n\to\infty} m_i = M_i(P), i = 2,3,4. \quad (3)$$

In formula (3), the symbol (P) means that convergence in probability is considered.

According to [19], the asymmetry coefficient of the random variable $X$ is called the quantity

$$\gamma_1 = \frac{M_3}{\sigma^3}. \qquad (4)$$

For a symmetric distribution (for example, if the distribution density is an even function when the median of the distribution is chosen as the reference point) with a finite third moment, the skewness is 0.

Kurtosis coefficient $X$ is the value

$$\gamma_2 = \frac{M_4}{\sigma^4} - 3. \qquad (5)$$

The sample coefficients of skewness and kurtosis are statistics obtained by substituting sample central moments instead of theoretical ones into formulas (4) and (5), i.e.

$$g_1 = \frac{m_3}{s^3} \qquad (6)$$

And

$$g_2 = \frac{m_4}{s^4} - 3 \qquad (7)$$

respectively.

With an increase in the sample size, the coefficients$g_1$and g2 converge (if the fourth moment of the random variable X is finite) to the theoretical skewness and kurtosis coefficients:

$$\lim_{n\to\infty} g_j = \gamma_j, j = 1,2 \qquad (8)$$

(convergence in probability).

It is known that for any normal distribution, the skewness and kurtosis coefficients are equal to 0. Naturally, the idea arises to use this information to test the agreement with the family of normal distributions.

If the hypothesis that the theoretical skewness coefficient is equal to 0 is rejected, then the distribution of the random variable $X$ is not normal. The converse is not true, since the theoretical skewness is 0 not only for normal distributions, but for all symmetrical ones.

Similarly, if the hypothesis that the theoretical kurtosis is equal to 0 is rejected, then the distribution of the random variable *X* is not normal. The converse is not true, since the theoretical coefficient of kurtosis is 0 not only for normal distributions, but also for many others.

The rules for testing these hypotheses are based on checking whether the sample coefficients of skewness and kurtosis fall into the corresponding critical regions. Detailed information on checking the agreement with a normal distribution based on these coefficients is contained in the fundamental monograph by L.N. Bolsheva and N.V. Smirnov "Tables of Mathematical Statistics" [23]. We consider this monograph to be the pinnacle of Russian statistical science in the first two thirds of the 20th century.

We have discussed checking the agreement of characteristics. As already noted, if the test results show that the characteristics are significantly different from those that correspond to the normal distribution, then the hypothesis that the sample is taken from the normal distribution should be rejected. However, the absence of a significant difference does not allow us to conclude that the distribution of sample elements is normal, since the same values of the characteristics (in this case, the skewness and kurtosis coefficients) also have many distributions that are not normal.

Therefore, let's move on to methods for checking absolute agreement. Consider criteria that are consistent, i.e. any difference from normality will be detected with a sufficiently large sample size. Consistent are analogues of the Kolmogorov criteria and omega-square. In the original formulation, these criteria are intended to test the agreement of the empirical distribution (sometimes written - experimental distribution) with a fixed theoretical one. When checking normality, the theoretical distribution is known only up to parameters, therefore, in the analogs of the Kolmogorov tests and the omega-square, the theoretical distribution is replaced by a normal distribution, in which the arithmetic mean of the sample elements is substituted for the mathematical expectation, and the sample variance is substituted for the variance. The asymptotic distributions of these analogs are found. They differ from the distributions of the Kolmogorov and omega-square statistics. Critical values differ many times [1 - 3]. A common mistake is that when checking the agreement with a normal distribution, the critical values of the distributions of the Kolmogorov and omega-square statistics are used, rather than the critical values of their counterparts [24].

Other criteria are also used to test normality, in particular, the Shapiro-Wilk test [25]. However, this criterion is not consistent.

## Application to beta distribution

Third central moment$M_3$can be estimated both directly from the sample (see formula (2)), and using its expression in terms of the parameters of the beta distribution, in which their sample estimates are substituted for the unknown theoretical values. One can use the estimates of the method of moments [20].

Consider a random variable *X*, which has a beta distribution on the interval (0; 1). For it, the probability distribution density is given by formula (1) in [20]. It is known [17, p.146] that the third central moment of the random variable X is expressed in terms of the parameters p and q of the beta distribution as follows:

$$M_3 = \frac{2\mathrm{pq}(q-p)}{(p+q)^3(p+q+1)(p+q+2)}. \quad (9)$$

We will use the estimates of the method of moments of the parameters p and q of the beta distribution

$$p * \overline{X}\left\{\frac{\overline{X}(1-\overline{X})}{s^2} - 1\right\} \quad (10)$$

And

$$q * (1-\overline{X})\left\{\frac{\overline{X}(1-\overline{X})}{s^2} - 1\right\} \quad \text{(eleven)}$$

respectively [20]. Substituting these estimates into (9), we obtain an estimate of the theoretical third central moment as a function of the sample arithmetic mean and sample variance. This estimate has the form

$$M_3 * \frac{2p*q*(q*-p*)}{(p*+q*)^3(p*+q*+1)(p*+q*+2)} = f(\overline{X}, s^2).(12)$$

There is no need to give here a rather complicated expression for the function f in (12).

To test the hypothesis of goodness of fit with the beta distribution, you can use a test based on statistics

$$Z = m_3 - M_3 * \quad (13)$$

If the sample elements are distributed according to the beta distribution, i.e. If there is absolute agreement, then

$$\lim_{n\to\infty} Z = 0. \quad (14)$$

(convergence in probability). It can be shown that the statistics indicated in (13) is asymptotically normal with mathematical expectation 0 and variance A/n, the constant A can be found based on the linearization method, demonstrated in detail in [20] when finding the asymptotic distribution of the estimates of the method of moments p* and q * parameters p and q of the beta distribution. On the basis of asymptotic normality, a criterion for testing the statistical hypothesis of agreement with a family of beta distributions is constructed in a standard way.

To check the agreement with the beta distribution, one can use not only the third central moment, but also other characteristics of the distribution. Thus, it is known [18] that for the beta distribution, the skewness coefficient is as follows:

$$\gamma_1 = \frac{2(q-p)\sqrt{p+q+1}}{(p+q+2)\sqrt{\mathrm{pq}}}. \quad (15)$$

Statistics can be used to check for agreement with the beta distribution.

$$W = g_1 - \frac{2(q*-p*)\sqrt{p*+q*+1}}{(p*+q*+2)\sqrt{p*q*}} \quad (16)$$

If the hypothesis of absolute agreement is true, the statistic W tends to 0 (in probability) with an unlimited increase in the sample size. The statistic W is asymptotically normal with zero mean and variance B/n. As in the previous case, the constant B can be found based on the linearization method described in [20].

Let us pass to beta distributions on the interval [*a*, *b*] (see formulas (60) and (61) in [20]. Let the sample$Y, Y_2,..., Y_n$.To check the agreement with the family of beta distributions, let us pass to the sample, in the distribution of which the parameters a and b are excluded, namely, to the sample$X_1, X_2,..., X_n$, Where

$$X_i = \frac{Y_i - a}{h} = \frac{Y_i - a}{b - a}, i = 1,2,..., n. \quad (17)$$

The elements of this sample have a beta distribution on the interval [0, 1] (see f-lu (1) in [20]), and you can check the agreement with the family of beta distributions using the above methods using the third central moment and the skewness coefficient .

To do this, it is necessary to pass from the sample moments calculated from the sample$Y, Y_2,..., Y_n$, to the sampling moments$X_1, X_2,..., X_n$. As indicated in formula (63) of article [20],

$$\overline{X} = \frac{\overline{Y} - a}{h}, s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{(b-a)^2}\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2\right) \quad (18).$$

Consider such a transition for a selective third central moment. It is easy to see that

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}\left(X - \overline{X}\right)^3 = \frac{1}{(b-a)^3}\left(\frac{1}{n}\sum_{i=1}^{n}(Y - Y)^3\right). \quad (18)$$

Let us analyze the statistical data given in Table 1 of the article [20]. For them, the boundaries of the interval are set*a*= 5, b = 135, then b - a = h = 130, the sample arithmetic mean is 57.88, and the sample variance s2 is 663.00, the sample standard deviation s = 25.75. Let's move on to a sample from the beta distribution with values in (0; 1) according to (17). After such a transition, the sample arithmetic mean becomes equal to 0.407, and the sample variance becomes equal to 0.0392.

By direct calculation according to the table. 1 in [20], we find that the sample third central moment for these data is 14927.91. When passing to a sample from the beta distribution on the interval (0, 1), in accordance with formula (18), we obtain that

$$m_3 = \frac{14927{,}91}{130^3} = \frac{14927{,}91}{2197000} = 0{,}006795 \qquad (19)$$

The estimates of the method of moments of parameters of the beta distribution are as follows: $p^*$ = 2.10, q* = 3.06 (see formulas (64) and (65) in [20]). Let us find an estimate of the theoretical third central moment by the formula (12)

$$M_3 * \frac{2\times 2{,}10\times 3{,}06(3{,}06-2{,}10)}{(2{,}10+3{,}06)^3(2{,}10+3{,}06+1)(2{,}10+3{,}06+2)} = 0{,}001879 \qquad (20)$$

As shown in [20], common recommendations are based on the fact that the parameters of the beta distribution take the values $p$= 2 and q = 3. Let us substitute these values into formula (9):

$$M_3 = \frac{2\times 2\times 3(3-2)}{(2+3)^3(2+3+1)(2+3+2)} = 0{,}002286. \qquad (21)$$

In formulas (19), (20). (21) we observe the difference only in the third decimal place.

Consider an estimate of the asymmetry coefficient. Direct calculation by formula (6) gives

$$g_1 = \frac{m_3}{s^3} = \frac{14927{,}91}{25{,}75^3} = \frac{14927{,}91}{17073{,}86} = 0{,}8743. \qquad (22)$$

When using the values p = 2 and q = 3 of the parameters of the beta distribution according to formula (15)

$$\gamma_1 = \frac{2(3-2)\sqrt{2+3+1}}{(2+3+2)\sqrt{2\times 3}} = \frac{2}{7} = 0{,}2857. \qquad (23)$$

If we use the values of the parameter estimates p* = 2.10, q* = 3.06, then according to formula (16) we obtain the asymmetry estimate

$$g_1 * \frac{2(3{,}06-2{,}10)\sqrt{2{,}10+3{,}06+1}}{(2{,}10+3{,}06+2)\sqrt{2{,}10\times 3{,}06}} = \frac{4{,}7653}{18{,}150} = 0{,}2626. \qquad (24)$$

The right parts of formulas (23) and (24) give close values, but differ from the value in the right part of (22). Table data. 1 in [20] correspond to a distribution with positive skewness, which is noticeably larger than for the beta distribution.

As already noted, a significant difference between the data in Table. 1 in [20] and a family of beta distributions, one could speak on the basis of the properties of the asymptotic distributions of the considered statistics (sample third central moment and sample skewness). However, at present, the parameters of the asymptotically normal distributions of these statistics remain unknown.

The large scatter of possible values of the parameters of beta distributions demonstrated in [20] (confidence intervals are very wide) gives reason to believe that the statistics considered for checking the agreement - the sample third central moment and the sample skewness coefficient - also have a wide confidence interval, and therefore reject the hypothesis on agreement with the data in Table. 1 with a family of beta distributions, there is no reason. From what has been said, it is clear that further research is needed in this direction.

**Conclusion**

In this article, the problem of testing the points of agreement between an experimental (sample) distribution and a family of beta distributions is posed and partially solved. It is shown that there are still many unsolved problems in the considered section of parametric statistics.

Some unsolved problems of nonparametric statistics are also revealed (the modern idea of nonparametric statistics is disclosed in [29]). So, it is of interest to check the symmetry of the distribution by the coefficient of asymmetry. Such a setting does not imply a normal distribution of the elements of the sample, from which the well-known methods start (see, for example, [23]). A successful solution of the problem under consideration can find application in the statistical analysis of coupled samples, give a new criterion for the homogeneity of coupled samples, supplementing the previously developed ones, in particular, those obtained in [30].

## Literature

1. Orlov A.I. Applied statistics. - M .: Exam, 2006. - 671 p.

2. Orlov A.I. Applied Statistical Analysis: textbook. - M .: AI Pi Ar Media, 2022. - 812 p.

3. Orlov A.I. Probability and applied statistics: basic facts: a handbook. — M.: KnoRus, 2010. — 192 p.

4. Orlov A.I. The main stages of the formation of statistical methods // Scientific journal of KubSAU. 2014. No. 97. S. 73-85.

5. Loiko V.I., Lutsenko E.V., Orlov A.I. High statistical technologies and system-cognitive modeling in ecology. - Krasnodar: KubGAU, 2019. - 258 p.

6. Orlov A.I. Distributions of real statistical data are not normal // Scientific journal of KubSAU. 2016. No. 117. P. 71–90.

7. Orlov A.I. Sustainability in socio-economic models. - M.: Nauka, 1979. - 296 p.

8. Orlov A.I. Statistics of objects of non-numerical nature and expert assessments // Expert assessments / Questions of cybernetics. Issue 58. - M.: Scientific Council of the Academy of Sciences of the USSR on the complex problem "Cybernetics", 1979. P.17-33.

9. Orlov A.I. Organizational and economic modeling: : textbook: in 3 hours Part 1: Non-numerical statistics. - M .: Publishing house of MSTU im. N. E. Bauman, 2009. - 542 p.

10. Orlov A.I. Artificial intelligence: non-numeric statistics: textbook. - M .: AI Pi Ar Media, 2022. - 446 p.

11. eleven.Orlov A.I. Statistics of non-numerical data - the central part of modern applied statistics // Scientific journal of KubSAU. 2020. No. 156. pp. 111–142.

12. 12.Orlov A.I.Statistics of non-numerical data for forty years (review) // Factory laboratory. material diagnostics. 2019. V.85. No. 11. - S. 69-84.

13. Orlov A.I., Lutsenko E.V. System fuzzy interval mathematics. Monograph (scientific edition). - Krasnodar, KubGAU. 2014. - 600 p.

14. Orlov A.I. System fuzzy interval mathematics - the basis of mathematics of the XXI century // Scientific journal of KubSAU. 2021. No. 165. pp. 111–130.

15. Orlov A.I., Lutsenko E.V. Analysis of data, information and knowledge in systemic fuzzy interval mathematics: scientific monograph. - Krasnodar: KubGAU, 2022. - 405 p.

16. Orlov A.I. Parameter estimation: one-step estimates are preferable to maximum likelihood estimates // Scientific journal of KubSAU.2015. No. 109. pp. 208 - 237.

17. Balakrishnan N., Nevzorov VB A primer of statistical distributions. - New Jersey: Wiley-Interscience, 2003. - 328 p.R.

18. Hastings N., Peacock J. Handbook of statistical distributions - M.: Statistics, 1980. - 95 p.

19. Probability and mathematical statistics: Encyclopedia / Ch. ed. Yu.V. Prokhorov. - M. : Bolshaya Ros. Encycl., 1999. - 910 p.

20. 20.Orlov A.I. Statistical analysis of samples from the beta distribution // Scientific journal of KubSAU. 2023. No. 187. S.

21. Orlov A.I. System of models and methods for checking the homogeneity of two independent samples // Scientific journal of KubGAU. 2020. No. 157. pp. 145–169.

22. Orlov A.I. On methods for checking the homogeneity of two independent samples //Factory laboratory. material diagnostics. 2020. V.86. No. 3. S. 67-76.

23. Bolshev L.N., Smirnov N.V. Tables of mathematical statistics. Ed. 3. - M.: science, 1983. - 417 p.

24. Orlov A.I. Non-parametric criteria of agreement of Kolmogorov, Smirnov, omega-square and errors in their application // Scientific journal of KubSAU.2014. No. 97. S. 32-45.

25. 25 Shapiro SS, Wilk MBAn analysis of variance test for normality // Biometrika. 1965. V.52. No. 3. P. 591-611.

26. Skvortsov Yu.V. Organizational and economic issues in graduation design. - M.: Higher school, 2006. - 399 p.
27. Organization and planning of machine-building production (production management): a textbook for universities / Nekrasov L. A., Postnikova E. S., Skvortsov Yu. V., Ukhanova T. V.; ed. Skvortsov Yu. V. - 3rd ed., revised. and additional - M. : Student, 2019. - 412 p.
28. 28.Golenko D.I. Statistical methods of network planning and management. - M.: Science. Ch. ed. Phys.-Math. lit., 1968. - 400 p.
29. Orlov A.I. Current state of nonparametric statistics //Scientific journal KubSAU. 2015. No. 106. S. 239 - 269.
30. Orlov A.I. On checking the homogeneity of coupled samples // Scientific journal of KubSAU. 2016. No. 123. P. 708–726.

## References

1. Orlov AI Prikladnaya statistika. — M.: Ekzamen, 2006. — 671 s.
2. Orlov AI Prikladnoj statisticheskij analiz: uchebnik. — M.: Aj Pi Ar Media, 2022. — 812 p.
3. Orlov AI Veroyatnost' i prikladnaya statistika: basic facts: spravochnik. — M.: KnoRus, 2010. — 192 s.
4. Orlov AI Osnovnye etapy stanovleniya statisticheskih metodov // Nauchnyj zhurnal KubGAU. 2014. No. 97. S. 73-85.
5. Lojko VI, Lucenko EV, Orlov AI – Krasnodar : KubGAU, 2019. – 258 s.
6. Orlov AI Raspredeleniya real'nyh statisticheskih dannyh ne yavlyayutsya normal'nymi // Nauchnyj zhurnal KubGAU. 2016. No. 117. S. 71–90.
7. Orlov AI Ustojchivost' v social'no-ekonomicheskih modelyah. - M.: Nauka, 1979. - 296 s.
8. Orlov AI Statistika ob "ektov nechislovoj prirody i ekspertnye ocenki // Ekspertnye ocenki / Voprosy kibernetiki. Vyp.58. - M .: Nauchnyj Sovet AN SSSR po kompleksnoj probleme "Kibernetika", 1979. S.17-33.
9. Orlov AI Organizacionno-ekonomicheskoe modelirovanie:: uchebnik: v 3 ch. CH.1: Nechislovaya statistika. — M.: Izd-vo MGTU im. NE Baumana, 2009. - 542 s.
10. Orlov AI Iskusstvennyj intellekt: nechislovaya statistika: uchebnik. — M.: Aj Pi Ar Media, 2022. — 446 p.
11. Orlov AI Statistika nechislovyh dannyh - central'naya chast' sovremennoj prikladnoj statistiki // Nauchnyj zhurnal KubGAU. 2020. No. 156. S. 111–142.
12. Orlov AI Statistika nechislovyh dannyh za sorok let (obzor) // Zavodskaya laboratoriya. Diagnostic materialov. 2019. T.85. No. 11. - S. 69-84.
13. Orlov AI, Lucenko EV Sistemnaya nechetkaya interval'naya matematika. Monografiya (scientific edition). – Krasnodar, KubGAU. 2014. - 600 s.
14. Orlov AI Sistemnaya nechetkaya interval'naya matematika - osnova matematiki XXI century // Nauchnyj zhurnal KubGAU. 2021. No. 165. S. 111–130.
15. Orlov AI, Lucenko EV Analiz dannyh, informacii i znanij v sistemnoj nechetkoj interval'noj matematike: nauchnaya monografiya. - Krasnodar: KubGAU, 2022. - 405 s.
16. Orlov AI Ocenivanie parametrov: odnoshagovye ocenki predpochtitel'nee ocenok maksimal'nogo pravdopodobiya // Nauchnyj zhurnal KubGAU. 2015. No. 109. S. 208-237.
17. Balakrishnan N., Nevzorov VB A primer of statistical distributions. - New Jersey: Wiley-Interscience, 2003. - 328 pr.
18. Hastings N., Pikok J. Spravochnik po statisticheskim raspredeleniyam - M.: Statistika, 1980. - 95 s.

19. Veroyatnost' i matematicheskaya statistika : Enciklopediya / Gl. red. YU.V. Prohorov. - M. : Bol'shaya Ros. Encikl., 1999. - 910 s.

20. Orlov AI Statisticheskij analiz vyborok iz beta-raspredeleniya // Nauchnyj zhurnal KubGAU. 2023. No. 187. S.

21. Orlov AI Sistema modelej i metodov proverki odnorodnosti dvuh nezavisimyh vyborok // Nauchnyj zhurnal KubGAU. 2020. No. 157. S. 145–169.

22. Orlov AI O metodah proverki odnorodnosti dvuh nezavisimyh vyborok // Zavodskaya laboratoriya. Diagnostic materialov. 2020. T.86. No. 3. S. 67-76.

23. Bol'shev LN, Smirnov NV Tablicy matematicheskoj statistiki. Izd. 3. - M.: nauka, 1983. - 417 s.

24. Orlov AI Neparametricheskie kriterii soglasiya Kolmogorova, Smirnova, omega-kvadrat i oshibki pri ih primenenii // Nauchnyj zhurnal KubGAU. 2014. No. 97. S. 32-45.

25. Shapiro SS, Wilk MB An analysis of variance test for normality // Biometrika. 1965. V.52. No. 3. P. 591-611.

26. Skvorkov YU.V. Organizacionno-ekonomicheskie voprosy v diplomnom proektirovanii. - M.: Vysshaya shkola, 2006. - 399 s.

27. Organizaciya i planirovanie mashinostroitel'nogo proizvodstva (proizvodstvennyj management): uchebnik dlya vuzov / Nekrasov LA, Postnikova ES, Skvorcov YU. V., Uhanova TV; red. Skvorkov YU. V. - 3rd izd., pererab. i dop. - M. : Student, 2019. - 412 s.

28. Golenko DI Statisticheskie metody setevogo planirovaniya i upravleniya. - M.: Nauka. gl. red. fiz.-mat. lit., 1968. - 400 s.

29. Orlov AI Sovremennoe sostoyanie neparametricheskoj statistiki // Nauchnyj zhurnal KubGAU. 2015. No. 106. S. 239 - 269.

30. Orlov AI O proverke odnorodnosti svyazannyh vyborok // Nauchnyj zhurnal KubGAU.2016. No. 123. S. 708–726.