УДК 303.4 : 519.2

UDC 303.4 : 519.2

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

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

ОЦЕНИВАНИЕ ПАРАМЕТРОВ ГАММА-РАСПРЕДЕЛЕНИЯ

ESTIMATION OF GAMMA DISTRIBUTION PARAMETERS

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

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

Постановки задач статистического анализа данных, имеющих гамма-распределение, относятся к классической математической статистике. Как ни странно, не все одни были решены в рамках параметрической статистики, находившейся на переднем крае развития статистической науки в первой трети XX в. Как и в случае бета-распределения, необходимо заполнить лакуны. Это необходимо, поскольку в настоящее время гамма-распределение широко используется в теоретических и прикладных работах. Примером является ГОСТ 11.011-83 "Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения". Стандартное гамма-распределение определяется параметром формы. При переходе к масштабно-сдвиговому семейству добавляются параметры масштаба и сдвига. Рассмотрены семь постановок задач оценивания параметров, поскольку каждый из трех параметров может быть как неизвестным, так и известным. Для каждой из постановок найдены оценки метода моментов и их асимптотические дисперсии. При известном параметре сдвига получены оценки максимального правдоподобия. Одношаговые оценки, асимптотически эквивалентные оценкам максимального правдоподобия, используем при неизвестном параметре сдвига. Наличие погрешностей измерения отражается на точности оценок параметров при применении тех или иных алгоритмов расчетов. В ГОСТ 11.011-83 на основе модели интервальных данных даны правила выбора метода оценивания при неизвестных параметрах формы и масштаба и известном параметре сдвига. При разработке ГОСТ 11.011-83 были выявлены проблемы, для решения которых предложены новые с научной точки зрения методы. Дальнейшее развитие новых научных результатов, полученных в ходе решения

Statements of problems of statistical analysis of data with a gamma distribution are related to classical mathematical statistics. Oddly enough, not all alone were solved within the framework of parametric statistics, which was at the forefront of the development of statistical science in the first third of the 20th century. As with the beta distribution, gaps need to be filled. This is necessary because the gamma distribution is currently widely used in theoretical and applied work. An example is GOST 11.011-83 "Applied statistics. Rules for determining estimates and confidence limits for gamma distribution parameters". The standard gamma distribution is determined by the shape parameter. When switching to a scale-shift family, scale and translation parameters are added. Seven formulations of parameter estimation problems are considered, since each of the three parameters can be either unknown or known. For each of the formulations, the estimates of the method of moments and their asymptotic variances are found. For a known shift parameter, maximum likelihood estimates are obtained. One-step estimates, asymptotically equivalent to maximum likelihood estimates, are used for an unknown shift parameter. The presence of measurement errors affects the accuracy of parameter estimates when applying certain calculation algorithms. In GOST 11.011-83, based on the interval data model, rules are given for choosing an estimation method for unknown shape and scale parameters and a known shift parameter. During the development of GOST 11.011-83, problems were identified, for the solution of which new methods from a scientific point of view were proposed. Further development of new scientific results obtained in the course of solving a practical problem (development of GOST 11.011-83) led to the creation of new scientific directions. We are talking about the statistics of interval data, as well as one-step estimates. To date, the statistics of interval data as a branch of практической задачи (разработки ГОСТ 11.011-83), привело к созданию новых научных направлений. Речь идет о статистике интервальных данных, а также об одношаговых оценках. К настоящему времени статистика интервальных данных как раздел математической статистики достаточно развита и охватывает все основные области статистических методов. Она является важной составной частью системной нечеткой интервальной математики

mathematical statistics is quite developed and covers all the main areas of statistical methods. It is an important part of systemic fuzzy interval mathematics

Ключевые слова: СТАТИСТИЧЕСКИЕ МЕТОДЫ, ГАММА-РАСПРЕДЕЛЕНИЕ, ОЦЕНИВАНИЕ ПАРАМЕТРОВ, МЕТОД МОМЕНТОВ, ОДНОШАГОВЫЕ ОЦЕНКИ, СТАТИСТИКА ИНТЕРВАЛЬНЫХ ДАННЫХ, АСИМПТОТИЧЕСКИЕ РАСПРЕДЕЛЕНИЯ, ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

Keywords: STATISTICAL METHODS, GAMMA-BETA DISTRIBUTION, ESTIMATION OF PARAMETERS, METHODS OF MOMENTS, ONE-STEP ESTIMATES, STATISTICS OF INTERVAL DATA, ASYMPTOTIC DISTRIBUTIONS, CONFIDENCE INTERVALS

http://dx.doi.org/10.21515/1990-4665-192-012

Introduction

Mathematical, statistical and instrumental methods of economics are based on the scientific discipline "Probability Theory and Mathematical Statistics". Currently it is widely known to the scientific community. In textbooks and reference books, when considering continuous probability distributions, they usually mention the family of gamma distributions (see, for example, [1 - 3], [38][1]). Known methods for estimating the parameters of probability distributions can be applied to this family. Thus, in the series of state standards "Applied Statistics" [4], we developed GOST 11.011-83, dedicated to algorithms for obtaining point estimates and confidence limits for family parameters and subfamilies of gamma distributions [5]. In preparing this normative and technical document, a number of research works were carried out, which made it possible to obtain sufficiently advanced calculation algorithms in the area under consideration. However, this standard was canceled in 1987, along with the whole series "Applied Statistics" (the reasons for this voluntaristic decision are considered in sufficient detail in [6]). After this point,

[1] Here and below, all references to sources are given in accordance with [1].

the pamphlet [5] could only be regarded as a scientific publication. However, this was hampered by its original status as an official regulatory and technical document (in accordance with it, the developers were not even named in the second edition). Brochure [5] was excluded from the libraries of standards (disposed of), but did not get into scientific use. In our opinion, the scientific results on the basis of which it was developed deserve attention. Subsequently, these results were generalized and widely developed. This article is devoted to them, in which for the first time the problem of estimating the parameters of gamma distributions is systematically considered. on which it was developed. Subsequently, these results were generalized and widely developed. This article is devoted to them, in which for the first time the problem of estimating the parameters of gamma distributions is systematically considered. on which it was developed. Subsequently, these results were generalized and widely developed. This article is devoted to them, in which for the first time the problem of estimating the parameters of gamma distributions is systematically considered.

Statements of problems of statistical analysis of data with a gamma distribution are related to classical mathematical statistics. Oddly enough, not all of them were solved within the framework of parametric (in the terminology of [7]) statistics, which was at the forefront of the development of statistical science in the first third of the 20th century. As in the case of the beta distribution [8, 9], it is necessary to fill in the gaps. This is necessary because the gamma distribution is currently widely used in theoretical and applied works (see, for example, [10–16]).

Family of gamma distributions

The density of the standard gamma distribution is:

$$f(x;a)=\begin{cases}\dfrac{1}{\Gamma(a)}x^{a-1}\exp(-x), & x\geq 0,\\ 0, & x<0\end{cases}, \tag{1}$$

where Γ(a) is the gamma function,

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} exp(-x) dx \quad (2)$$

The standard gamma distribution is defined by a single parameter a > 0, which is called the shape parameter.

For a random variable X with a standard gamma distribution, the mathematical expectation and variance are equal to a. It has a third central moment M(X – M(X))3 = 2a, asymmetry $2/\sqrt{a}$ and kurtosis 6/a (see [1 - 3]).

The standard gamma distribution generates a scale-shift family of distributions of random variables

$$Y = bX + c, \quad (3)$$

where b is the scale parameter and c is the shift parameter. As follows from formula (1), the random variable Y given by formula (3) has a density

$$f(x; a, b, c) = \begin{cases} \frac{1}{\Gamma(a)} (x-c)^{a-1} b^{-a} \exp\left[-\frac{x-c}{b}\right], & x \geq c, \\ 0, & x < c. \end{cases} \quad (4)$$

According to (1) and (4) f(x; a) = f(x; a, 1, 0). Then

M(Y) = ab + c, D(Y) = ab2, M(Y – M(Y))3 = 2ab3. (5)

Let Y1, Y2, ..., Yn be a sample of independent identically distributed random variables with gamma distribution (4). In this article, we consider the problem of estimating the parameters of this distribution.

The gamma distribution (4) has three parameters (shape, scale and shift). Each of them can be both known and unknown., total 23 = 8 options. One of them is not related to the problems under consideration, since all parameters are known. To ensure the completeness of the study, it is necessary to consider 23 - 1 = 7 formulations of estimation problems. They are listed in Table 1.

Table 1.

Production options assessment tasks

No. p / p	Form parameter	Scale parameter	Shift parameter
1	Estimated	Famous	Famous
2	Famous	Estimated	Famous
3	Famous	Famous	Estimated
4	Famous	Estimated	Estimated
5	Estimated	Famous	Estimated
6	Estimated	Estimated	Famous
7	Estimated	Estimated	Estimated

Consider all 7 productions.

Estimates of the method of moments

The simplest method for estimating distribution parameters is the method of moments[17, 18]. It consists in the fact that the theoretical moments are expressed in terms of the distribution parameters (see, for example, (5)), then the inverse problem is solved - the functional dependence of the distribution parameters on the theoretical moments is found, then, to obtain statistical estimates of the parameters, sample moments are substituted into this functional dependence instead of theoretical ones. As the sample size grows, the statistical estimates are asymptotically normal and asymptotically unbiased, and the linearization method is used to obtain their asymptotic variances [17, Sec. 4.4]. For samples from beta distributions, this research program was implemented in [8].

The ambiguity of the method of moments is determined by the fact that different sample moments can be used to estimate a certain parameter. For example, in formulation 1 (Table 1), one can proceed from any of the three relations given in (5). We obtain three estimates for the method of moments:

$$a_1 = a_1(\omega) = \frac{\overline{Y} - c}{b},\ \ a_2 = a_2(\omega) = \frac{s^2}{b^2},\ \ a_2 = a_2(\omega) = \frac{m_3}{2b^3}, \qquad (6)$$

Where

$$\overline{Y} = \frac{Y_1 + Y_2 + ... + Y_n}{n},\ s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2,\ m_3 = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^3. \qquad (7)$$

The choice between the three estimates (6) is carried out in favor of the first of them as the most resistant to outliers[17, section 7.2]. Similar results are obtained by comparing the asymptotic variances of estimates (6) found using linearization, as explained above.

Similarly, for statement 2 (see Table 1), we use the estimate

$$b^* = \frac{\overline{Y} - c}{a},\ \ D(b^*) = \frac{1}{n}\frac{b^2}{a}.$$

and for setting 3 - an estimate

$$c^* = \overline{Y} - ab,\ \ D(c^*) = \frac{1}{n}ab^2.$$

In statements 4 - 6, two parameters are unknown, so two sample moments must be used. We will use the sample arithmetic mean and sample variance. We start from the equalities

$$\overline{Y} = ab + c.\ s^2 = ab^2. \qquad (8)$$

In statement 4, based on the sample characteristics and the shape parameter, the scale and shift parameters are estimated. Namely,

$$b^* = \frac{s}{\sqrt{a}},\ \ c^* = \overline{Y} - ab^* = \overline{Y} - s\sqrt{a}.$$

In statement 5, based on the sample characteristics and the scale parameter, the shape and shift parameters are estimated. Namely,

$$a^* = \frac{s^2}{b^2},\ c^* = \overline{Y} - a^* b = \overline{Y} - \frac{s^2}{b}.$$

In Statement 6, based on the sample characteristics and the shift parameter, the shape and scale parameters are estimated. Namely,

$$\overline{Y} - c = ab.\ s^2 = ab^2;\ b^* = \frac{s^2}{\overline{Y} - c},\ a^* = \frac{\overline{Y} - c}{b^*} = \frac{(\overline{Y} - c)^2}{s^2}.$$

In Statement 7, all three parameters are unknown, so three sample moments must be used. We will use the sample arithmetic mean, variance and the third central moment. We have

$$\overline{Y} = ab + c.\ s^2 = ab^2,\ m_3 = 2ab^3,$$

where

$$\frac{m_3}{2s^2} = b^*,\ a^* = \frac{s^2}{(b^*)^2} = \frac{4s^6}{m_3^2},\ c^* = \overline{Y} - a^* b^* = \overline{Y} - \frac{2s^4}{m_3}.$$

The results of estimating the parameters of the gamma distribution by the method of moments for all 7 settings (Table 1) are given in Table. 2. The asymptotic variances of these estimates were obtained by us by the linearization method, which is considered in detail in[8] and [17, Sec. 4.4]. The results are shown in the right column of Table 2.

Table 2.

Estimates of the method of moments and their asymptotic variances

No.	staging tasks	Assessed parameter	View estimates	Asymptotic variance of the estimate
1	1	a	$\frac{\overline{Y} - c}{b}$	$\frac{a}{n}$
2	2	b	$\frac{\overline{Y} - c}{a}$	$\frac{1}{n}\frac{b^2}{a}$
3	3	c	$\overline{Y} - ab$	$\frac{1}{n}ab^2$

4	4	b	$\frac{s}{\sqrt{a}}$	$\frac{b}{2n}(a+3)$
5	4	c	$\overline{Y}-s\sqrt{a}$	$\frac{ab^2}{2n}(a+1)$
6	5	a	$\frac{s^2}{b^2}$	$\frac{2a}{n}(a+3)$
7	5	c	$\overline{Y}-\frac{s^2}{b}$	$\frac{ab^2}{n}(2a+3)$
8	6	a	$\frac{(\overline{Y}-c)^2}{s^2}$	$\frac{2a(a+1)}{n}$
9	6	b	$\frac{s^2}{\overline{Y}-c}$	$\frac{b^2}{n}\left(2+\frac{3}{a}\right)$
10	7	a	$\frac{4s^6}{m_3^2}$	$\frac{6a}{n}(a^2+6a+5)$
eleven	7	b	$\frac{m_3}{2s^2}$	$\frac{b^2}{2an}(6a^2+25a+24)$
12	7	c	$\overline{Y}-\frac{2s^4}{m_3}$	$\frac{ab^2}{n}(3a^2+13a+10)$

Asymptotic confidence bounds for the parameters of the gamma distribution are constructed based on the asymptotic normality of the estimates of these parameters. They look like

$$(\theta_n - C(\gamma)\sqrt{D^*(\theta_n)};\ \theta_n + C(\gamma)\sqrt{D^*(\theta_n)}),$$

Where θ - parameter under consideration, θ_n - its assessment (see Table 2), $D^*(\theta_n)$ - variance estimation θ_n, obtained by replacing in the formula for the asymptotic variance (the right column of Table 2) the unknown values of the parameters with their estimates, $C(\gamma)$ - coefficient corresponding to the confidence level γ (For $\gamma = 0{,}95$ we have $C(\gamma) = C(0{,}95) = 1{,}96$).

Here and below, confidence intervals are constructed on the basis of the asymptotic normality of parameter estimates. For some samples Y_1, Y2, ..., Yn, the confidence limits calculated by asymptotic formulas can go beyond the limits of the domains of definition of the corresponding parameters. Then you need to make adjustments. If the lower confidence limit for the shape parameter or the scale parameter turns out to be negative, then it should be replaced by 0. If the upper confidence limit for the shift parameter is greater than any of the sample elements, then it should be replaced by min(Y1, Y2, ... , Yn). The probabilities of such events tend to 0 as the sample size increases.

As in the case of beta distributions [9], the method of moments can be used to check the agreement with a family of gamma distributions. So, to check the agreement with the subfamily of gamma distributions f(x; a, b, 0) with the shift parameter c = 0, you can use the test with statistics

$$Z = \sqrt{n}\left(\frac{\overline{Y} m_3}{2s^4} - 1\right)$$

Using the method described above for finding the asymptotic distribution of the function of sample moments, we obtain that with an increase in the sample size, the distribution Z converges to the normal one, and if the hypothesis that the shift parameter is equal to 0 is true, the limit distribution Z has a zero mathematical expectation and variance

$$\frac{1}{2a}\left(3a^2 + 13a + 10)\right).$$

Since the shape parameter a is unknown, we should substitute its consistent estimate a* into the expression for the variance. The statistical test for testing the null hypothesis c = 0 with the alternative c > 0 has a critical region

$$\left\{ Z:\ Z > u(1-\alpha)\sqrt{\frac{3(a^*)^2 + 13a^* + 10}{2a^*}} \right\},$$

Where $u(1-\alpha)$ - order quantile $1-\alpha$ standard normal distribution with mean 0 and variance 1. When $n \to \infty$ the significance level of the criterion under consideration tends to α. The use of the method of moments to test the hypothesis of agreement with a parametric family of distributions is considered in more detail in [18, Sec. 7.1] and [39].

Maximum Likelihood Estimates(with a known shift parameter)

Moment method estimates are the simplest type of parameter estimates. However, in many cases they are not the best asymptotically normal estimators (NAS estimators) because their variance is larger. Among the NAS estimators, the best known are the maximum likelihood estimators. They are obtained by maximizing the product of densities with respect to the parameters, in which the elements of the sample are substituted instead of arguments. In case of gamma distributionwith density (4) we are talking about maximizing the likelihood function

$$F=\prod_{1\le i\le n} f(Y_i;a,b,c)=\prod_{1\le i\le n}\frac{1}{\Gamma(a)}(Y_i-c)^{a-1}b^{-a}\exp\left[-\frac{Y_i-c}{b}\right] \quad (9)$$

parameters a > 0, b > 0, c < min (Y1, Y1, ..., Yn).

It's easier to deal with the log-likelihood function

$$L(a,b,c)=\ln F=-n\ln\Gamma(a)+(a-1)\sum_{i=1}^{n}\ln(Y_i-c)-an\ln b-\frac{1}{b}\sum_{i=1}^{n}Y_i+\frac{nc}{b}. \quad (10)$$

For statement 1, the estimate of the parameter a is found from the equation

$$\frac{\partial L}{\partial a}=-n\frac{d\ln\Gamma(a)}{da}+\sum_{i=1}^{n}\ln(Y_i-c)-n\ln b=0. \quad \text{(eleven)}$$

Let us introduce the logarithmic derivative of the gamma function

$$\Psi(a)=\frac{d\ln\Gamma(a)}{da}=\frac{1}{\Gamma(a)}\frac{d\Gamma(a)}{da}$$

(this special function is called the psi-function [19]). Then from (11) it follows that the estimate of the shape parameter for known scale and shift parameters is found as a solution to the equation

$$\Psi(a)=\frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{Y_i-c}{b}\right),$$

those. the asymptotically normal form parameter estimate is

$$a^* = G\left(\frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{Y_i - c}{b}\right)\right),$$

where the function G is the inverse of the psi-function. In [5, Sec. 4] the rules for calculating the values of the function G, the variance of the estimate a* and confidence limits are given, based on the use of tables and approximate formulas. At present, it is advisable to use appropriate computer programs.

For staging 2 rating *b**parameter b - solution of the equation

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial b}\left(-an\ln b - \frac{1}{b}\sum_{i=1}^{n}(Y_i - c)\right) = -\frac{an}{b} + \frac{1}{b^2}\sum_{i=1}^{n}(Y_i - c) = 0, \quad (12),$$

those.

$$b^* = \frac{1}{na}\sum_{i=1}^{n}(Y_i - c). \quad (13)$$

Tables and approximate (asymptotic) formulas for calculating the confidence limits for the parameter b are given in [5, Sec. 3]. At the present stage of development of applied statistics, computer programs are used for the same purpose.

Consider the estimation of the shape and scale parameters with a known shift parameter (statement 6). Maximum likelihood estimates are found by solving the system of equations (11) - (12). The estimate of the scale parameter is given by formula (13). Substituting it into (11), we get the equation for shape parameter

$$-n\frac{d\ln\Gamma(a)}{da} + \sum_{i=1}^{n}\ln(Y_i - c) - n\ln\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i - c)\right) + n\ln a = 0.$$

Let's transform this equation:

$$\ln a - \Psi(a) = \ln\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i - c)\right) - \frac{1}{n}\sum_{i=1}^{n}\ln(Y_i - c).$$

Therefore, the estimates of the shape and scale parameters with a known shift parameter have the form

$$a^* = H\left(\ln\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i - c)\right) - \frac{1}{n}\sum_{i=1}^{n}\ln(Y_i - c)\right),\ b^* = \frac{1}{na^*}\sum_{i=1}^{n}(Y_i - c)\,, \qquad (14)$$

where H is the function inverse to the function

$$Q(a) = \ln a - \Psi(a)\,.$$

IN[5, Section 7] describes the algorithms for calculating estimates (14), their variances and confidence limits, based on the corresponding tables and approximate (asymptotic) formulas developed in the preparation of GOST 11.011-83. As in the previously analyzed statements, these results of applied mathematics can now be used in the preparation of the corresponding software.

Using the linearization method (see also [20, pp. 83, 98]), one can show that for large n, up to infinitesimals of a higher order

$$M(a^* - a)^2 = \frac{a}{n\big(aI(a) - 1\big)},\ M(b^* = b)^2 = \frac{b^2 I(a)}{n\big(aI(a) - 1\big)},\ I(a) = \frac{d\Psi(a)}{da}\,.$$

Based on the properties of the gamma function [19], it was found in [5] that for large a

$$M(a^* - a)^2 = \frac{a(2a-1)}{n},\ M(b^* - b)^2 = \frac{2b^2}{n}\,,$$

up to infinitesimals of a higher order. Using Table 2, we conclude that the mean squares of errors for the estimates of the method of moments of the shape parameters a** and scale b** (with a known shift parameter)

$$M(a^{**} - a)^2 = \frac{2a(a+1)}{n},\ M(b^{**} - b)^2 = \frac{b^2}{n}\left(2 + \frac{3}{a}\right)$$

greater than the corresponding mean squared errors for the maximum likelihood estimates. Thus, for Statement 6, the maximum likelihood estimates have an advantage over the estimates of the method of moments. As established in the theory of classical mathematical statistics, this is the case in most parameter estimation problems. The maximum likelihood estimators are always the best asymptotically normal estimators, and the estimators of the method of moments are only in some cases.

One-step estimates (with unknown shift parameter)

In settings 3, 4, 5, 7, the shift parameter *With* un known and to be estimated.

Maximum Likelihood Estimation c*parameter c in Statement 3 can be determined from the condition

$$\frac{\partial L}{\partial c} = \frac{\partial L}{\partial c}\left((a-1)\sum_{i=1}^{n}\ln(Y_i - c) + \frac{nc}{b}\right) = -(a-1)\sum_{i=1}^{n}\frac{1}{Y_i - c} + \frac{n}{b} = 0. \qquad (15)$$

In Statement 4, the estimates of the scale and shift parameters are solutions of the system of maximum likelihood equations (12) and (15).The expression for the scale parameter is given in (13)). Substituting it into (15), we obtain an estimate for the shift parameter c* from the equation

$$-(a-1)\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i - c} + \frac{a}{\frac{1}{n}\sum_{i=1}^{n}(Y_i - c)} = 0. \qquad (16)$$

In formulation 5, the shape and shear parameter estimate sare found when solving the system of maximum likelihood equations (11) and (15). We substitute the solution of equation (11) into (15), we obtain an equation for estimating the shift parameter

$$\left\{G\left(\frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{Y_i - c}{b}\right)\right) - 1\right\}) - \frac{n}{b} = 0 \qquad (17)$$

where the function G is the inverse of the psi-function. The calculation rules are given in [5].

For three unknown parameters (statement 7), the maximum likelihood estimates should be found by solving the system of three equations (11), (12), and (15).

Thus, in a number of formulations, in order to obtain maximum likelihood estimates, it is necessary to solve the corresponding equations or systems of equations. It would seem that it is enough to apply common numerical methods. However, this raises a number of problems.

In many parameter estimation problems, there are no algorithms based on explicit formulas for finding maximum likelihood estimates, and numerical methods have to be applied.

In applied statistics apart from the method of moments and maximum likelihood estimators, many other types of estimators have been developed.

One-step estimates are obtained in the form of explicit formulas, i.e. for them, the computational problems considered above are removed. When calculating them, one can take the estimate of the method of moments as an initial approximation, and then find corrections based on the first iteration of the Newton-Raphson method.

For example, for the gamma distribution in statement 3, instead of solving equation (15), it is recommended to use a one-step estimate. As an initial approximation, the estimate of the method of moments is taken $\overline{Y} - ab$ .Calculation formulas are given in [5, pp. 8.15 - 8.18]. For statements 4 and 7, the rules for determining one-step estimates are given in [5, Sec. 8].

Development of GOST 11.011-83[5] gave impetus to the further development of the theory of one-step estimates [26–29]. To date, a detailed presentation of this theory has been included in textbooks (see [8, Section 6.2], [30, Section 3.25]), including its application to estimating the parameters of the gamma distribution, so in this paper we do not consider these scientific results.

Note that although one-step estimators, as well as maximum likelihood estimators, are asymptotically best, but for specific sample sizes, these estimators may be inferior to other estimators, for example, unbiased ones.[31].

Interval Data Statistics

When developing GOST 11.011-83, the foundations were laid for a developed area of modern mathematical statistics - statistics of interval data. The impetus was the analysis of data on the operating time of cutters to the limit state (in hours). They are given in [5] and further publications on this topic (see, for example, [18, Section 6.1]). It is striking that all the initial data are given either by natural numbers or by the sums of natural numbers and 0.5, for example, 127.5 hours). This means that the operating time of the cutters to the limit state is measured with an accuracy of 0.5 hours. Obviously, the presence of errors must be taken into account in the statistical processing of data.

The presence of measurement errors affects the accuracy of parameter estimates when applying certain calculation algorithms [5, Section 5].

Further development of interval data statistics is reflected in the articles [32 - 34]. To date, this area of mathematical statistics is sufficiently developed and covers all the main areas of statistical methods. It is an important part of systemic fuzzy interval mathematics [35, ch.7] and is described in detail in textbooks [18, ch.12], [36, ch.4], [37, ch.2.3].

Conclusion

Gamma distributions are widely used in various fields of science and practice, in particular, in reliability (for example, in the "load-strength" model [12]) and testing theory, in various fields of engineering and technology (including modeling the accuracy of a technological process [13]), in meteorology, etc. [40 - 43].

This article includes information about the main results in the development of methods for estimating the parameters of the gamma distribution based on limit theorems. The next stage is the construction and study of estimation algorithms for finite sample sizes. Note that, in comparison with the maximum likelihood estimates and one-step estimates, the unbiased estimates may turn out to be more accurate [31].

It is interesting to look at the development of research in this area from the point of view of science of science.. When solving a practical problem (i.e., developing GOST 11.011-83), problems were identified, for the solution of which new methods from a scientific point of view were proposed. Further development of new scientific results obtained in the course of solving a practical problem led to the creation of new scientific directions. We are talking about the statistics of interval data, as well as one-step estimates. Naturally, the development of the theory was followed by its practical applications.

This development of research is not an isolated case. Another example is the development of ASPPAP (automated system for forecasting and preventing aviation accidents).As explained in [44], its creation turned out to be possible only as a result of a large number of specific scientific works.

Those who wish to get acquainted with this work in Russian can do this in the work [1].

Literature

1. Orlov AI Estimation of gamma distribution parameters // August 2023, DOI:10.13140/RG.2.2.34932.12165,https://www.researchgate.net/publication/373482588

References

1. Orlov AI Ocenivanie parametrov gamma-raspredeleniya // August 2023, DOI:10.13140/RG.2.2.34932.12165,https://www.researchgate.net/publication/373482588