Characteristic Function in Estimation of Probability Distribution Moments

Vladimir S. Timofeev

***Abstract*—**In this article the problem of distributional moments estimation is considered. The new approach of moments estimation based on usage of the characteristic function is proposed. By statistical simulation technique author shows that new approach has some robust properties. For calculation of the derivatives of characteristic function there is used numerical differentiation. Obtained results confirmed that author's idea has a certain working efficiency and it can be recommended for any statistical applications.

***Keywords*—**Characteristic function, distributional moments, robustness, outlier, statistical estimation problem, statistical simulation.

I. Introduction

As a rule, traditional methods of applied statistical analysis based on the use of the information contained in the moments of random variables which carried out. For example, the construction of regression equations by classical methods is associated primarily with the expectation value. Evaluating of the mathematical expectation is also required in the distributions identification task, conducting correlation, factorial, and many other types of statistical analysis. Therefore, one of the main objectives should be considered evaluation of the expectation or more generally – evaluation moments of random variables. Traditionally estimation by the arithmetic mean would not always lead to good results. So, if the sample contains at least a few outliers (observation that located on enormous distance from other values); then our estimate may have high deformation (bias). It is a well known fact that has is emphasized many times by many authors [1], [2].

One way for solving this problem should be go to robust methods, among which the most famous based on the median. Also there are the order statistics mean, mean by truncated sample, winsorized mean and others [3], [4]. Some of the research results of these methods can be found in [5]. However, in practice, the particular choice of estimation method need do carefully. For example, usage of median as a measure of mathematical expectation for a skew distribution can lead to unfortunate results with high errors.

Another solution involves the usage of so-called rejection methods. However, there could be obtained undesirable results. Examples of such situations can be found in [1]. Especially carefully rejection methods should be used when analyzing the relationships between factors, i.e. during correlation, regression and other multivariate statistical analysis. This is due to the fact that highlights the value of an attribute (factor) may be the result of simultaneous effect of a set of other factors and then it is very valuable observation that should not be discarded.

By this cause, the authors focused their attention on the alternative method of calculating the moments based on the usage of characteristic function. This method is known today only as a theoretical fact and is not widely applied in solving real problems. In this regard, the article presents the results illustrating the advantages of this approach when solving a problem of estimating the moments of random variables, and in regression analysis.

V. S. Timofeev is with the Applied Mathematics and Computer Science Department, Novosibirsk State Technical University, Novosibirsk, 630073 Russia (phone: +7(383)346-37-54; fax: +7(383)346-02-09; e-mail: v.timofeev@corp.nstu.ru).

II. Characteristic Function and Moments of Random Variable

It is well known [6] that the characteristic function of a random variable ξ with the density function $f_\xi(u)$ is defined as follows:

$$\varphi(t) = E\left[e^{itu}\right] = \int_{-\infty}^{\infty} e^{itu} f_\xi(u)du\,, \tag{1}$$

where $t \in R$, $i = \sqrt{-1}$ – the so-called imaginary unit. Seeing

$$\left|e^{itu}\right| = 1,\ \forall\, t \in R,\ -\infty < u < \infty\,, \tag{2}$$

then the characteristic function exists for any real random variable. Directly from the definition (1) for $\varphi(t)$ follows

$$\varphi(0) = 1,\ \left|\varphi(t)\right| \le 1,\ \varphi(-t) = \overline{\varphi(t)}\,. \tag{3}$$

From the inversion theorem [6], there is a one-one correspondence between the characteristic function and the density function. Therefore, knowing $\varphi(t)$ may be restored $f_\xi(u)$

$$f_\xi(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itu}\varphi(t)dt\,. \tag{4}$$

Note also that the characteristic function $\varphi_\eta(t)$ of a random variable η resulting as the linear transformation $\eta = a_0 + a_1\xi$

(a_0, a_1 – constant) is associated with the characteristic function $\varphi(t)$ of the random variable ξ.

$$\varphi_\eta(t) = e^{ita_0}\varphi(a_1 t). \quad (5)$$

There is a decomposition of the characteristic function in a series by moments. If there is the initial moments m_r to ν-th order inclusive for a random variable ξ, they are expressed in terms $\varphi(t)$

$$m_r = i^{-r}\varphi^{(r)}(0), \; r = 1,...,\nu, \quad (6)$$

where $\varphi^{(r)}(t)$ – derivative of r-order of the characteristic function. Then it is true the Maclaurin expansion [6]

$$\varphi(t) = 1 + \sum_{r=1}^{\nu} \frac{i^r m_r}{r!} t^r + R_\nu, \quad (7)$$

where R_ν – residual term. Optionally, can be written the expansion analogous to (7), but by the central moments [6].

On the basis of the realization of the random variable ξ (sample volume is N) is also possible to build the sample estimate of the characteristic function, which has the form [7]:

$$\hat{\varphi}(t) = \frac{1}{N}\sum_{j=1}^{N} e^{itx_j} = \frac{1}{N}\sum_{j=1}^{N}\left(\cos(tx_j) + i\sin(tx_j)\right), \quad (8)$$

where $x_1, x_2, ..., x_N$ – the sample values of random variable ξ. It should be noted that, in accordance with the law of large numbers [6] estimate (8) is consistent.

Application of (6) requires knowledge of the derivatives of the characteristic function. In solving real-world problems the calculation of the characteristic function by (1) is usually not possible because has no information about the form of distribution of these random variables. In this regard, the authors have put into practice numerical differentiation of the empirical characteristic function (8). To do this, it should be noted that (8) is a complex function of a real argument, which can be written as

$$\hat{\varphi}(t) = u(t) + iv(t). \quad (9)$$

In this cause the derivative is

$$\hat{\varphi}'(t) = u'(t) + iv'(t). \quad (10)$$

Next will used standard formulas for numerical differentiation [8] separately for the real and imaginary parts of the characteristic function. In simulation researches the step size is chosen equal $\frac{\pi}{32}$. Expressions for the first five derivatives of the real part at zero argument are the follows:

$$u'(0) = \frac{1}{2h}\left(u(h) - u(-h)\right), \quad (11)$$

$$u''(0) = \frac{1}{h^2}\left(u(h) - 2u(0) + u(-h)\right), \quad (12)$$

$$u^{(3)}(0) = \frac{1}{2h^3}\left(u(2h) - 2u(h) + 2u(-h) - u(-2h)\right), \quad (13)$$

$$u^{(4)}(0) = \frac{1}{h^4}\begin{pmatrix} u(2h) - 4u(h) + 6u(0) - \\ -4u(-h) + u(-2h) \end{pmatrix}, \quad (14)$$

$$u^{(5)}(0) = \frac{1}{2h^5}\begin{pmatrix} u(3h) - 4u(2h) + 5u(h) - \\ -5u(-h) + 4u(-2h) - u(-3h) \end{pmatrix}. \quad (15)$$

Derivatives of the imaginary part are calculated in the same manner.

III. Simulation Results

Now we consider the results of our research that have been carried out on the basis of statistical simulation techniques. Research has been related with the proposed approach for estimation of moments. In the computational experiments, it was assumed that a random variable have distribution function as follow

$$F(x) = (1-\mu)F_1(x, 0, \sigma_1) + \mu F_2(x, 0, \sigma_2), \quad (16)$$

where $F_i(x, 0, \sigma_i)$ – normal distribution functions with zero mean and variance σ_i^2; $\mu \in [0,1]$, $i = 1, 2$.

Note that the (16) allows simulate a random variable (sample) in the presence of different level of outliers. The parameter μ determines fractions of observations with the variances σ_1^2 and σ_2^2 respectively in the sample. In cases $\mu = 0$ and $\mu = 1$ the sample have normal distribution and contains no outliers. In executed computational experiments it is assumed that $\sigma_2^2 \geq \sigma_1^2$. However, the author's simulation process not requires fixing the variances σ_1^2 and σ_2^2 themselves there it was need to determinate their corresponding indicators, called noise levels. The term «noise level» is well known and is defined as the ratio of «noise» / «signal» percentage.

$$\rho = \frac{\sigma}{c}100\%, \quad (17)$$

where σ^2 – the variance of random variable; $c^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i^0 - \bar{x}^0)^2$ – signal rate.

First of all consider the accuracy of estimating the initial moments for the normal distribution by the traditional way, i.e. through the usual average operation

$$\hat{m}_r = \frac{1}{N}\sum_{i=1}^{N} x_i^r . \tag{18}$$

Next is carried out the accuracy of estimating on the basis of the empirical characteristic function, i.e. by (6). Table I shows the obtained computational experiments results for the first five initial moments, averaged over 100 samples containing 500 elements each. In rows «classic» are presented moments estimation results by (18) and in rows «Based on characteristic function» – by (6). In these computational experiments considered the standard normal distribution, contaminated the standard normal distribution (one percent of outliers, i.e. $\mu = 0.01$), and the normal distribution with zero mean and variance 4.

Columns number corresponds to the order of initial moments. For example, column with number «1» consists estimation results of mean, column with number «2» consists estimation results of variance (because random variable is centered) and etc.

Research for moments of more high orders lead us for the same patterns.

TABLE I
RESULTS OF INITIAL MOMENT ESTIMATION

Estimation method	1	2	3	4	5
Standard normal distribution					
Classic	-0.0043	0.997	-0.0259	2.981	-0.1591
Based on characteristic function	-0.0043	0.995	-0.0255	2.957	-0.1555
Normal distribution *N*(0,4)					
Classic	0.0043	3.962	-0.0057	46.403	-0.6875
Based on characteristic function	0.0043	3.925	-0.0041	44.997	-0.5361
Contaminated the standard normal distribution					
Classic	0.1094	2.5374	22.244	326.708	4756.2
Based on characteristic function	0.0772	2.2924	12.939	230.132	2274.7

Analysis of the results presented in Table I, led to form the following conclusions. For the standard normal distribution obtained estimates of moments are very close to each other. In this case, the even moments a little better estimated by the standard approach, and the odd - proposed, i.e. based on the characteristic function. The appearance of an outlier in the sample has a very strong influence on the accuracy of the estimation; it becomes especially noticeable on the higher order moments. The initial moment of the first order is also better estimated using the characteristic function. It allows consider the construction of the characteristic function as a kind of filter that provides a more robust estimation. The results obtained for the normal distribution with variance 4, also support this conclusion: higher-order moments are measured by (6) is more precise.

IV. Conclusion

The obtained results emphasizes that characteristic function is important statistical instruments that can found a useful applications in robust theory.

The considered idea of robust estimation moments of random variables can be widely used in the regression analysis. It is well known that traditional regression analysis approaches use the information contained in the first two moments (expectation and variance). Actively develop adaptive procedures are beginning to use information on the third and fourth moments. Furthermore the author develops the technology of adaptive parameter estimation of regression models based on universal distributions (generalized lambda distribution, stable distributions, etc.). In this regard, the proposed technique of estimation of probability distribution moments is of particular importance.

Acknowledgment

This research has been supported by the Ministry of Education and Science of the Russian Federation as part of the state task no 2014/138 (project No 1689).

References

[1] P. J. Rousseeuw, and A. M. Leroy, *Robust regression and outlier detection*. New York: John Wiley& Sons, 1987, 334p.
[2] P. J. Rousseeuw, and K. van Driessen, *Computing LTS regression for large data sets*. Dept.Mathematics, University of Antwerp, 1999.21p.
[3] R. Grübel, “The length of the shorth,” *Ann. Statist.*, vol. 16, no. 2, pp. 619-628, 1988.
[4] J. W. Tukey, and D. H. McLaughlin, “Less vulnerable confidence and significance procedures for location based on a single sample: Trimming/Winsorization,” *Sankhya*, Series A, vol. 25, pp. 331-352, 1963.
[5] V. S. Timofeev, and V. Iu. Shchekoldin, “On the estimation of the statistical characteristics in the analysis of multivariate objects,” [Ob otsenivanii statisticheskikh kharakteristik pri analize mnogofaktornykh ob"ektov]. Nauchnyi vestnik NGTU – *Scientific Herald of the Novosibirsk State Technical University*, vol. 24, no. 3, pp. 47-58,2006.
[6] V. S. Pugachev, *Probability theory and mathematical statistics* [Teoriia veroiatnostei i matematicheskaia statistika]. Moscow, Nauka Publ., 1979, 496 p.
[7] A. Feuerverger, and R.A. Mureika, “The empirical characteristic function and its applications,” *The annals of statistics*, vol. 5, no. 1, pp. 88-97, 1977.
[8] J. N. Lyness, and C. B. Moler, “Numerical differentiation of analytic functions,” *SIAM J. Numer. Anal.*, vol. 4, pp. 202-210, 1967.