MSC 90B06 DOI: 10.14529/mmp220411

SMOOTH APPROXIMATION OF THE QUANTILE FUNCTION DERIVATIVES

V.R. Sobol[1], *R.O. Torishnyy*[1]
[1]Moscow Aviation Institute, Moscow, Russian Federation
E-mail: vitsobol@mail.ru, arenas-26@yandex.ru

In this paper, a smooth approximation of the second-order derivatives of quantile function is provided. The convergence of approximations of the first and second order derivatives of quantile function is studied in cases when there exists a deterministic equivalent for the corresponding stochastic programming problem. The quantile function is one of common criteria in stochastic programming problems. The first-order derivative of quantile function can be represented as a ratio of partial derivatives of probability function. Using smooth approximation of probability function and its derivatives we obtain approximations of these derivatives in the form of volume integrals. Approximation of the second-order derivative is obtained directly as derivative of the first-order derivative. A numerical example is provided to evaluate the accuracy of the presented approximations.

Keywords: stochastic programming; probability function; quantile function and its derivatives.

Introduction

Wide range of design problems are represented as mathematical programming problems, where the optimization criterion represents the system performance, which depends on chosen optimization vector (strategy) and some undetermined parameters. Consideration of the undetermined effects as random parameters appears very effective in many engineering, financial, and social problems. Addition of random parameters in the model turns the objective and constraints to random functions. This fact leads us to consideration of optimization problems in which the objective or constraints are in the form of a probability function or a quantile function [1]. The probability function equals to probability that a specified level is not exceeded by the loss value. The quantile function is the minimum loss level, which is not exceeded with a given probability. The choice of the loss level or reliability level depends on the specifics of the system considered.

There are vast numerical methods and algorithms developed to solve non-stochastic optimization problems. These methods often assume that we can calculate or estimate first and/or second order derivatives of the criterion – estimate its gradient and Hessian. But this technique can not be applied directly to stochastic optimization problems. Calculation of a probability function gradient implies the integration over surface [2]. Only in specific cases direct calculation can be implemented through volume integration [3]. Other approaches to derive the gradient approximations are described for example in [4,5], but they are limited by the type of distribution or other stochastic mechanisms.

One effective approach to estimate the derivatives of probability function is to use the smooth approximation of probability function. The key idea is to replace the Heaviside function inside the probability function with its smooth approximation that is a sigmoid function [6]. Then the approximation of probability function along with its gradient takes a form of volume integrals. Smooth approximation of probability function and its derivatives along with the proof of convergence of the approximations were first presented in [6] for one-dimensional random vector with absolutely continuous distribution. In [7] the same results were obtained in the case of continuous random vector.

Smooth approximation of the second-order derivatives of probability function was first described in [8]. The application of second-order optimization algorithms for stochastic optimization problems was described. The approximations of the second-order derivatives were presented as expectations or volume integrals, which can be calculated using Monte-Carlo method. The proof of convergence of approximations of the second-order derivatives was not considered in [8], since exact formulas are not available. But numerical examples show that smooth approximations of the second-order derivatives converge to ones estimated using finite differences. The present paper continues the research considered in [8] and presents a smooth approximation of the second-order derivatives of quantile function. The convergence analysis in some basic setups is provided. Numerical example is given. The example shows that smooth approximations have lower variance than finite-difference estimates and provide more accurate estimates.

The paper is organized as follows. In Section 1 all necessary formulas and statements on the smooth approximations are provided. In Section 2 the convergence analysis is made for some basic cases of the loss function. In Section 3 numerical example is given, comparing finite-difference estimates with the smooth approximations of quantile function derivatives. In conclusion an overview of the paper is provided.

1. Theoretical Part

Consider a complete probability space $(\mathbf{\Omega}, \mathbf{F}, \mathbf{P})$, an absolutely continuous random vector X with the support $G \subseteq \mathbb{R}^n$, a probability density function $f(x)$, and a loss function $g(u,x)$ depending on a strategy $u \in U$, $U \subseteq \mathbb{R}^m$. The function $g(u,x)$ is considered to be smooth and strictly piecewise monotonic with respect to x. The probability function is equal to probability that a random value $g(u,X)$ do not exceed a specified level φ, while the quantile function is equal to minimal loss level, which is not exceeded with a given probability α:

$$P_\varphi(u) \triangleq \mathbf{P}\left\{g(u,X) \le \varphi\right\} = \int\limits_G \Theta(\varphi - g(u,x)) f(x) dx, \tag{1}$$

$$\varphi_\alpha(u) \triangleq \min\{\varphi : P_\varphi(u) \ge \alpha\}, \tag{2}$$

where $\Theta(\cdot)$ is the Heaviside function.

Since $g(u,X)$ is the loss function, we can consider two optimization problems [1]:

$$\varphi_\alpha(u) \to \min_{u \in U}, \tag{3}$$

$$P_\varphi(u) \to \max_{u \in U}. \tag{4}$$

According to [6] we replace Heaviside function with a sigmoid function to get a differentiable approximation of the probability function. The sigmoid function and its derivatives are defined as follows

$$S_\theta(t) = \frac{1}{1 + e^{-\theta t}}, \tag{5}$$

$$S'_\theta(t) = \theta S_\theta(t)(1 - S_\theta(t)), \qquad S''_\theta(t) = \theta^2 S_\theta(t)(1 - S_\theta(t))(1 - 2S_\theta(t)) \tag{6}$$

with the parameter $\theta > 0$ corresponding to the steepness of the sigmoid function.

The approximation of the probability function is defined as:

$$P^\theta_\varphi(u) = \int\limits_G S_\theta(\varphi - g(u,x)) f(x) dx = \mathrm{E}\left(S_\theta(\varphi - g(u,X))\right), \tag{7}$$

The approximation of the probability function derivatives are defined as:

$$\frac{\partial P_{\varphi}^{\theta}(u)}{\partial \varphi} = \int\limits_{G} S'_{\theta}(\varphi - g(u,x)) f(x) dx = \mathrm{E}\left(S'_{\theta}(\varphi - g(u,X))\right), \tag{8}$$

$$\frac{\partial P_{\varphi}^{\theta}(u)}{\partial u_i} = -\int\limits_{G} S'_{\theta}(\varphi - g(u,x)) \frac{\partial g(u,x)}{\partial u_i} f(x) dx = -\mathrm{E}\left(S'_{\theta}(\varphi - g(u,X)) \frac{\partial g(u,X)}{\partial u_i}\right). \tag{9}$$

The following statements were proven [6, 7]:

$$\lim_{\theta \to +\infty} P_{\varphi}^{\theta}(u) = \mathbf{P}\left\{g(u,X) \le \varphi\right\}, \tag{10}$$

$$\lim_{\theta \to +\infty} \frac{\partial}{\partial \varphi} P_{\varphi}^{\theta}(u) = \frac{\partial}{\partial \varphi} \mathbf{P}\left\{g(u,X) \le \varphi\right\}, \tag{11}$$

$$\lim_{\theta \to +\infty} \frac{\partial}{\partial u_i} P_{\varphi}^{\theta}(u) = \frac{\partial}{\partial u_i} \mathbf{P}\left\{g(u,X) \le \varphi\right\}. \tag{12}$$

Approximation of the second-order derivatives of the probability function [8] is

$$\frac{\partial^2 P_{\varphi}^{\theta}(u)}{\partial u_i \partial u_j} = \mathrm{E}\left(S''_{\theta}(\varphi - g(u,X)) g'_{u_i}(u,X) g'_{u_j}(u,X) - S'_{\theta}(\varphi - g(u,X)) g''_{u_i u_j}(u,X)\right). \tag{13}$$

Approximation of quantile function derivatives are given in [6]:

$$\frac{\partial \varphi_{\alpha}^{\theta}(u)}{\partial \alpha} \triangleq \frac{1}{\mathrm{E}\left(S'_{\theta}(\varphi_{\alpha}(u) - g(u,X))\right)}, \quad \frac{\partial \varphi_{\alpha}^{\theta}(u)}{\partial u_i} \triangleq \frac{\mathrm{E}\left(S'_{\theta}(\varphi_{\alpha}(u) - g(u,X)) g'_{u_i}(u,X)\right)}{\mathrm{E}\left(S'_{\theta}(\varphi_{\alpha}(u) - g(u,X))\right)}. \tag{14}$$

In the last ratio, we denote the numerator and denominator by $V(u)$ and $W(u)$, respectively:

$$V(u) \triangleq \mathrm{E}\left[S'_{\theta}(\varphi_{\alpha}(u) - g(u,X)) g'_{u_i}(u,X)\right], \tag{15}$$

$$W(u) \triangleq \mathrm{E}\left[S'_{\theta}(\varphi_{\alpha}(u) - g(u,X))\right]. \tag{16}$$

Approximation of the second partial derivative of the quantile function is defined as:

$$\frac{\partial^2 \varphi_{\alpha}^{\theta}(u)}{\partial u_i \partial u_j} = \frac{V'_{u_j}(u) W(u) - V(u) W'_{u_j}(u)}{W^2(u)}, \tag{17}$$

where the derivatives of the functions $V(u)$ and $W(u)$ are defined as:

$$V'_{u_j}(u) = \mathrm{E}\left[S'_{\theta}(\varphi_{\alpha}(u) - g(u,X)) g''_{u_i u_j}(u,X)\right] + \\ + \mathrm{E}\left[S''_{\theta}(\varphi_{\alpha}(u) - g(u,X)) g'_{u_i}(u,X) \left(\frac{\partial}{\partial u_j} \varphi_{\alpha}(u) - g'_{u_j}(u,X)\right)\right], \tag{18}$$

$$W'_{u_j}(u) = \mathrm{E}\left[S''_{\theta}(\varphi_{\alpha}(u) - g(u,X)) \left(\frac{\partial}{\partial u_j} \varphi_{\alpha}(u) - g'_{u_j}(u,X)\right)\right]. \tag{19}$$

2. Convergence Analysis

We assess the convergence of approximations for 3 cases, where the quantile function can be defined explicitly (the deterministic equivalent exists).

2.1. Separable Loss Function

Let us consider the separable loss function, where $u \in U \subset \mathbb{R}^m$ and $X \in \mathbb{R}^n$:

$$g(u, X) = g_1(u) + g_2(X).$$

In this case the quantile function and its derivatives are equal to

$$\varphi_\alpha(u) = g_1(u) + [g_2(X)]_\alpha, \qquad \frac{\partial \varphi_\alpha(u)}{\partial u_i} = \frac{\partial g_1(u)}{\partial u_i}, \qquad \frac{\partial^2 \varphi_\alpha(u)}{\partial u_i \partial u_j} = \frac{\partial^2 g_1(u)}{\partial u_i \partial u_j},$$

where $[g_2(X)]_\alpha$ denotes the quantile of the random variable $g_2(X)$. In this case

$$V(u) = \mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X)) g'_{1u_i}(u)\right] = g'_{1u_i}(u) \mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X))\right],$$
$$W(u) = \mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X))\right].$$

The ratio of $V(u)$ and $W(u)$ gives $g'_{1u_i}(u) = \frac{\partial}{\partial u_i} g_1(u)$:

$$\frac{\partial \varphi^\theta_\alpha(u)}{\partial u_i} = \frac{V(u)}{W(u)} = \frac{\partial \varphi_\alpha(u)}{\partial u_i}, \qquad \frac{\partial^2 \varphi^\theta_\alpha(u)}{\partial u_i \partial u_j} = \frac{\partial^2 \varphi_\alpha(u)}{\partial u_i \partial u_j}.$$

2.2. Product of Two Functions

Next, consider the loss function to be a product of two functions:

$$g(u, X) = g_1(u) g_2(X),$$

where $g_1(u) > 0$ for every $u \in U$. The quantile function and its derivatives are equal to

$$\varphi_\alpha(u) = g_1(u)[g_2(X)]_\alpha, \qquad \frac{\partial \varphi_\alpha(u)}{\partial u_i} = \frac{\partial g_1(u)}{\partial u_i}[g_2(X)]_\alpha, \qquad \frac{\partial^2 \varphi_\alpha(u)}{\partial u_i \partial u_j} = \frac{\partial^2 g_1(u)}{\partial u_i \partial u_j}[g_2(X)]_\alpha.$$

The auxiliary functions $V(u)$ and $W(u)$ are defined as:

$$V(u) = \mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X)) g'_{1u_i}(u) g_2(X)\right] = g'_{1u_i}(u) \mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X)) g_2(X)\right],$$
$$W(u) = \mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X))\right].$$

The smooth approximation of the quantile function derivative is equal to:

$$\frac{\partial \varphi^\theta_\alpha(u)}{\partial u_i} = \frac{\mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X)) g_2(X)\right]}{\mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X))\right]} \frac{\partial}{\partial u_i} g_1(u).$$

The smooth approximation of the quantile function gradient takes the following form:

$$\nabla \varphi^\theta_\alpha(u) = \nabla g_1(u) \cdot \frac{\mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X)) g_2(X)\right]}{\mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X))\right]}.$$

The vectors $\nabla \varphi^\theta_\alpha(u)$ and $\nabla g_1(u)$ are codirectional. Smooth approximation of the second-order derivative is a linear combination of the second-order derivative of the loss function and the product of the first-order derivatives with coefficient converging to zero.

2.3. One-Dimensional Case

Let $X \in \mathbb{R}^1$, i.e. $n = 1$, and suppose that $g(u, x)$ is a strictly increasing and left-continuous function with respect to x for every u. This case was studied in [1, Chapter 4]:

$$\varphi_\alpha(u) = g(u, [X]_\alpha).$$

Using the properties of the sigmoid function described in [6] we state that:

$$V(u) = \mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X))g'_{u_i}(u, X)\right] \to \frac{g'_{u_i}(u, x)}{|g'_x(u, x)|}\Big|_{x=[X]_\alpha} \text{ as } \theta \to \infty,$$

$$W(u) = \mathrm{E}\left[S'_\theta(\varphi_\alpha(u) - g(u, X))\right] \to \frac{1}{|g'_x(u, x)|}\Big|_{x=[X]_\alpha} \text{ as } \theta \to \infty,$$

$$V'_{u_j}(u) \to \frac{g''_{u_i u_j}(u, x)}{|g'_x(u, x)|}\Big|_{x=[X]_\alpha}, \qquad W'_{u_j}(u) \to 0 \text{ as } \theta \to \infty.$$

Therefore:

$$\frac{\partial \varphi^\theta_\alpha(u)}{\partial u_i} = \frac{V(u)}{W(u)} \to g'_{u_i}(u, [X]_\alpha) \text{ as } \theta \to \infty,$$

$$\frac{\partial^2 \varphi^\theta_\alpha(u)}{\partial u_i \partial u_j} = \frac{V'_{u_j}(u)W(u) - V(u)W'_{u_j}(u)}{W^2(u)} \to g''_{u_i u_j}(u, [X]_\alpha) \text{ as } \theta \to \infty.$$

3. Numerical Example

Let us compare smooth approximation of quantile function derivatives with the finite difference estimates based on sample quantiles. We use a sample of size 100000 to calculate sample quantiles and to calculate expectations in (14) – (19). The parameter α is equal to 0,8, θ is equal to 30. To find a sample quantile for a given vector u we use the sample $\{X_i\}_{i=1}^N$ of size N, then obtain a sample of loss function values $\{g_i(u)\}_{i=1}^N$, where $g_i(u) = g(u, X_i)$. Let $g_{(i)}(u)$ be the i-th element of the variational series:

$$g_{(1)}(u) \le g_{(2)}(u) \le \ldots \le g_{(N)}(u).$$

The sample quantile is a statistical estimate of the quantile function [9]:

$$\hat{\varphi}_\alpha(u) \triangleq g_{([\alpha N])}(u), \tag{20}$$

where $[\cdot]$ denotes the integer part.

Example 1. Let us consider a two-dimensional case with a bilinear loss function and multivariate normal distribution:

$$g(u, X) = u_1 X_1 + u_2 X_2, \qquad X \sim N(m, K),$$

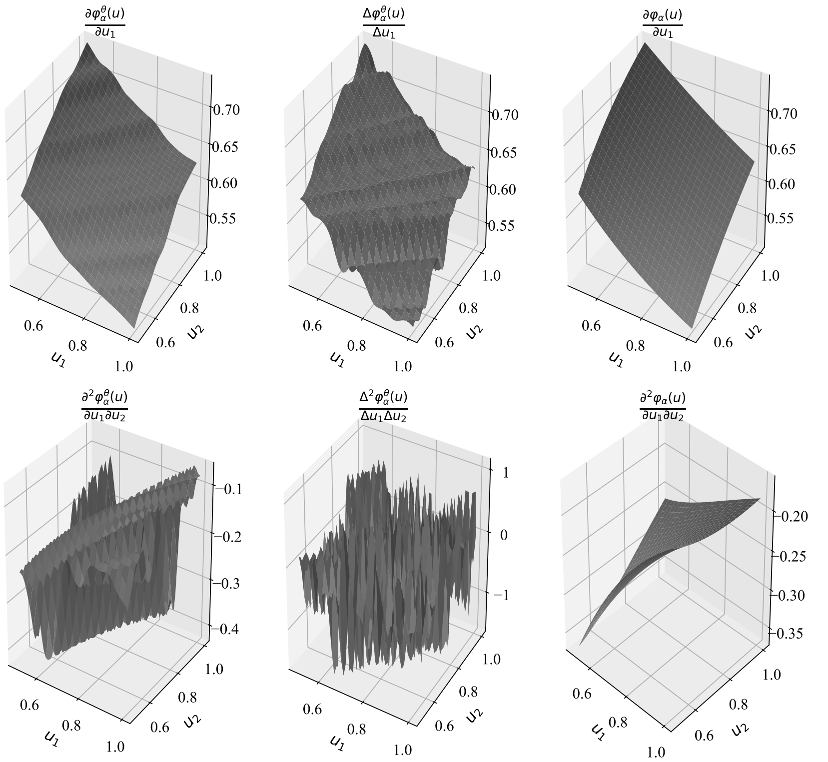
$$m = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad K = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 2 \end{pmatrix}.$$

The quantile function can be found explicitly:

$$\varphi_\alpha(u) = x_\alpha \sqrt{u_1^2 + u_1 u_2 + 2u_2^2},$$

where x_α is the α-quantile for standard Gaussian distribution.

The comparison of the smooth approximation of quantile function derivatives with the finite-difference estimates and with exact derivatives is presented in Figure. The smooth approximation of the first-order derivatives is close to the exact derivative and have lower variance. The smooth approximation of the second-order derivative has correct sign and slope, but is more affected by noise when compared to the first-order approximations.



First and second order derivatives of quantile function in Example 1

Conclusion

We provide a smooth approximation of the second-order derivatives of quantile function. Approximations of the first and second order derivatives of quantile function are expressed via partial derivatives of the probability function, which have a form of volume integrals and can be calculated using Monte-Carlo method. We provide convergence analysis for special cases of loss function and give a numerical example to analyse the accuracy of the smooth approximations of quantile function derivatives. Numerical example shows that the smooth approximation of quantile function derivatives is close to exact values.

Acknowledgements. *The reported research was funded by RSCF, project number 22-21-00213, https://rscf.ru/project/22-21-00213/.*

References

1. Kibzun A.I., Kan Yu.S. *Stochastic Programming Problems with Probability and Quantile Functions.* London, John Wiley & Sons, 1996.

2. Raik E. The Differentiability in the Parameter of the Probability Function and Optimization of the Probability Function via the Stochastic Pseudogradient Method. *Proceedings of Academy of Sciences of the Estonian SSR. Physics. Mathematics*, 1975, vol. 24, no. 1, pp. 3–9.

3. Uryas'ev S. Derivatives of Probability Functions and Some Applications. *Annals of Operations Research*, 1995, vol. 56, pp. 287–311. DOI:10.1007/BF02031712

4. Henrion R. Gradient Estimates for Gaussian Distribution Functions: Application to Probabilistically Constrained Optimization Problems. *Numerical Algebra, Control and Optimization*, 2012, vol. 2, no. 4, pp. 655–668. DOI:10.3934/naco.2012.2.655

5. Pflug G., Weisshaupt H. Probability Gradient Estimation by Set-Valued Calculus and Applications in Network Design. *SIAM Journal on Optimization*, 2005, vol. 15, no. 3, pp. 898–914. DOI:10.1137/S1052623403431639
6. Sobol V.R., Torishnyi R.O. On Smooth Approximation of Probabilistic Criteria in Stochastic Programming Problems. *SPIIRAS Proceedings*, 2020, vol. 19, no. 1, pp. 181–217. DOI:10.15622/sp.2020.19.1.7
7. Sobol V., Torishnyi R. Smooth Approximation of Probability and Quantile Functions: Vector Generalization and its Applications. *Journal of Physics: Conference Series*, 2021, vol. 1925, article ID 012034.
8. Torishyi, R. Application of the Second-Order Optimization Methods to the Stochastic Programming Problems with Probability Function. *Trudy MAI*, 2021, no. 121, 27 p.
9. Cox D.R., Hinkley D.V. *Theoretical Statistics.* London, Chapman and Hall, 1979.

Received May 13, 2022

УДК 519.856 **DOI: 10.14529/mmp220411**

О ГЛАДКОЙ АППРОКСИМАЦИИ ПРОИЗВОДНЫХ ФУНКЦИИ КВАНТИЛИ

В.Р. Соболь[1]***, Р.О. Торишный***[1]
[1]Московский авиационный институт (НИУ), г. Москва, Российская Федерация

В статье предложена гладкая аппроксимация вторых производных функции квантили. Сходимость аппроксимаций первых и вторых производных функции квантили исследуется в случаях, когда для соответствующей задачи стохастического программирования существует детерминированный эквивалент. Функция квантили является одним из основных критериев в задачах стохастического программирования. Производная первого порядка может быть представлена как отношение частных производных функции вероятности. Используя гладкую аппроксимацию функции вероятности и ее производных, эти производные аппроксимируются в форме объемных интегралов. Аппроксимация второй производной определяется непосредственно дифференцированием аппроксимации первой производной. Для оценки точности представленных аппроксимаций приведен численный пример.

Ключевые слова: стохастическое программирование; функция вероятности; функция квантили и ее производные.

Исследование выполнено за счет гранта Российского научного фонда № 22-21-00213, https://rscf.ru/project/22-21-00213/

Литература

1. Кибзун, А.И. Задачи стохастического программирования с вероятностными критериями / А.И. Кибзун, Ю.С. Кан. – М.: Физматлит, 2009.
2. Raik, E. The Differentiability in the Parameter of the Probability Function and Optimization of the Probability Function Via the Stochastic Pseudogradient Method / E. Raik // Proceedings of Academy of Sciences of the Estonian SSR. Physics. Mathematics. – 1975. – V. 24, № 1. – P. 3–9.
3. Uryas'ev, S. Derivatives of Probability Functions and Some Applications / S. Uryas'ev // Annals of Operations Research. – 1995. – V. 56. – P. 287–311.

4. Henrion, R. Gradient Estimates for Gaussian Distribution Functions: Application to Probabilistically Constrained Optimization Problems / R. Henrion // Numerical Algebra, Control and Optimization. – 2012. – V. 2, № 4. – P. 655–668.
5. Pflug, G. Probability Gradient Estimation by Set-Valued Calculus and Applications in Network Design / G. Pflug, H. Weisshaupt // SIAM Journal on Optimization. – 2005. – V. 15, № 3. – P. 898–914.
6. Соболь, В.Р. О гладкой аппроксимации вероятностных критериев в задачах стохастического программирования / В.Р. Соболь, Р.О. Торишный // Труды СПИИРАН. – 2020. – Т. 19, № 1. – С. 180–217.
7. Sobol, V. Smooth Approximation of Probability and Quantile Functions: Vector Generalization and its Applications / V. Sobol, R. Torishnyi // Journal of Physics: Conference Series. – 2021. – V. 1925. – № 012034.
8. Торишный, Р.О. О применении численных методов второго порядка к задачам стохастического программирования с функцией вероятности / Р.О. Торишный // Труды МАИ. – 2021. – № 121. – 27 с.
9. Кокс, Д. Теоретическая статистика / Д. Кокс, Д. Хинкли. – М.: Мир, 1978.

Виталий Романович Соболь, кандидат физико-математических наук, доцент, кафедра теории вероятностей и компьютерного моделирования, Московский авиационный институт (национальный исследовательский университет) (г. Москва, Российская Федерация), vitsobol@mail.ru.

Роман Олегович Торишный, ассистент, кафедра теории вероятностей и компьютерного моделирования, Московский авиационный институт (национальный исследовательский университет) (г. Москва, Российская Федерация), arenas-26@yandex.ru.

Поступила в редакцию 13 мая 2022 г.